

---

## Sommario

Capitolo 1 .....	1
Modelli di rete sociale:.....	1
dalle online social network alle reti Multilivello .....	1
1.1 Modelli di rete sociale .....	1
1.2 Online Social Network .....	5
1.3 Definizione di una rete sociale multilivello .....	7
Capitolo 2.....	8
Dataset e applicazione.....	8
del modello di rete sociale .....	8
multilivello.....	8
2.1 Descrizione dei dataset.....	8
2.1.1 UNICAL.....	11
2.1.2 UPB.....	11
2.1.3 LAPLAND.....	11
2.1.4 SASSY .....	12
2.1.5 Social Evolution.....	12
2.1.6 SIGCOMM.....	12
2.2 Applicazione del modello multi livello ai dataset.....	13
Capitolo 3.....	18
Strumenti di analisi: .....	18
approccio sociocentrico.....	18
ed egocentrico .....	18
3.1 Introduzione.....	18
3.2 Analisi Sociocentrica.....	18
3.2.1 Betweenness centrality.....	19
3.2.2 Closeness centrality.....	19
3.2.3 Eigenvector centrality .....	20
3.2.4 Louvain community detection .....	20

---

3.2.5	K-CLIQUE community detection .....	21
3.3	Analisi Egocentrica .....	21
3.3.1	Degree centrality .....	21
3.3.2	Ego betweenness centrality .....	22
3.3.3	Brokerage .....	22
3.3.4	Tie strenght .....	23
3.4	Analisi dei risultati .....	23
3.4.1	Centralità e correlazione .....	23
3.4.2	Similarità tra gruppi .....	30
3.4.3	Ruoli di intermediazione .....	31
3.4.4	Legami forti.....	35
Capitolo IV	.....	39
Il problema della	.....	39
Selfishness nelle reti opportunistiche:	.....	39
cause ed effetti	.....	39
4.1	Definizione .....	39
4.2	Simulations setup .....	41
4.3	Gli effetti della selfishness .....	43
4.3.1	Risultati su SASSY .....	44
4.3.2	Risultati su LAPLAND.....	46
Capitolo V	.....	48
SORSI: protocollo di instradamento Social – based Opportunistico con Rilevazione della Selfishness e meccanismi di Incentivazione	.....	48
5.1	Descrizione e funzionamento .....	48
5.2	Valutazione delle performance.....	55
5.2.1	Simulation setup.....	55
5.2.2	Risultati .....	58
Capitolo VI	.....	63
Conclusioni, Pubblicazioni	.....	63
e Sviluppi Futuri.....	.....	63
6.1	Conclusioni.....	63
6.2	Pubblicazioni.....	65

---

6.2.1	Journal.....	65
6.2.2	International Conference.....	65
6.3	Sviluppi futuri.....	66

## **Abstract**

In questa tesi si è proceduto ad approfondire le tematiche inerenti la trasmissione delle informazioni all'interno delle reti "ad hoc" e l'individuazione e lo studio di modelli e algoritmi per le reti sociali digitali. In particolare, a partire dall'individuazione di un'architettura di rete sociale multi – layer, basata su grafi, in cui i vari livelli possono essere rappresentati dalla rete dei contatti fisici, dalle online networks e dalla rete degli interessi, ci si è posto come obiettivo principale quello di analizzare la generazione di community per i vari livelli andando a proporre nuove metriche sociali. Caratterizzare il comportamento dei nodi di tali reti risulta infatti essere di importanza rilevante specialmente per l'individuazione dei nodi maggiormente sfruttabili ai fini dell'instradamento dell'informazione. Ulteriore tematica affrontata in questa tesi è quella relativa alla caratterizzazione e analisi degli effetti del comportamento egoistico dei nodi all'interno di una rete opportunistica rivolti alla formulazione di algoritmi di routing che ne gestiscano gli effetti negativi. Risulta chiaro come i nodi di tali reti, essendo prevalentemente dispositivi mobili dalle limitate capacità sia in termini di buffer che di batteria, mirino a preservare tali risorse. Per analizzare le prestazioni della rete in termini di delivery ratio, latenza media e overhead occorre pertanto tenere in considerazione i comportamenti egoistici dei nodi. A seguito di tale analisi è stato proposto SORSI (Social-based Opportunistic Routing with Selfishness detection and Incentive mechanisms), un algoritmo di gestione della "selfishness" nelle reti opportunistiche, e ne sono state valutate le performance in confronto con alcuni dei principali algoritmi sociali presenti in letteratura quali Bubble Rap e SPRINT-SELF.

## Introduzione

La visione di un futuro non troppo lontano in cui una moltitudine di dispositivi mobili trasportati da persone fisiche possa facilmente creare reti wireless locali esterne alla rete Internet classica sta attraendo diversi gruppi di ricercatori ad affrontare le tematiche nelle aree delle Delay Tolerant Networks (DTNs)<sup>1,2,3</sup>, reti opportunistiche<sup>4</sup> e le più recenti Do – it – yourself (DIY<sup>5</sup>) networks. Anche se Internet con la sua larga diffusione, ha rivoluzionato il modo in cui comunichiamo, esistono ancora diversi scenari in cui questa infrastruttura di rete non è sempre disponibile. Esempi di ciò, sono le situazioni di disaster/recovery, grandi eventi sportivi e festival musicali, o più in generale, aree in cui Internet è costoso, non disponibile o sovraffollato. In tali situazioni, un mezzo di comunicazione alternativo appare necessario. Vista l'ormai sempre più larga diffusione di smartphones, tablets e smartwatch, e l'impatto che il loro uso ha nella vita sociale di ogni individuo, lo studio di reti prive di infrastruttura che permettano comunicazioni wireless a corto raggio tra nodi, mediante tecnologie quali ad esempio il Bluetooth e il Wi – Fi, sta generando particolare interesse nella comunità scientifica.

Le DTN sono state progettate per permettere la comunicazione tra dispositivi distribuiti all'interno di uno scenario di rete privo di infrastrutture, con topologie di rete sparsa e contatti intermittenti. I nuovi meccanismi di instradamento applicati a queste reti sono chiamati store – carry – forward, ovvero, quando un nodo riceve un pacchetto da uno dei nodi tra i suoi contatti può memorizzare il messaggio, trasportarlo mentre si muove e instradarlo verso i nodi che incontra, che possono essere a loro volta intermediari o la destinazione. Una strategia simile è usata nelle reti opportunistiche in cui i dispositivi sono nodi tipicamente trasportati da persone ed instradano i messaggi durante le opportunità di incontro e contatto che si verificano nel momento in cui entrano nel reciproco raggio di comunicazione. Mentre nelle reti DTN ci sono casi in cui i punti di disconnessione sono conosciuti e l'instradamento può essere effettuato anche tramite l'infrastruttura Internet, nelle reti opportunistiche i percorsi verso la destinazione sono calcolati dinamicamente. Nel momento in cui un percorso end – to – end tra la sorgente e la destinazione di un

---

messaggio non è sempre disponibile, gli algoritmi di routing opportunistici si dimostrano spesso migliori rispetto a quelli utilizzati nella rete Internet e le reti ad hoc tradizionali. Ulteriori esempi di utilizzo delle reti opportunistiche sono la costruzioni di mappe di qualità dell'aria, del rumore, della temperatura, della concentrazione di CO<sub>2</sub><sup>6</sup>.

È inoltre pensiero comune che le reti opportunistiche siano caratterizzate da una natura sociale che può essere sfruttata per lo scambio di informazioni. Lavori recenti sulle reti opportunistiche hanno mostrato che la consegna dei messaggi può essere migliorata selezionando i migliori nodi intermediari considerando sia gli incontri wireless che avvengono nel mondo reale sia le interazioni sociali con gli altri nodi<sup>7,8,9,10</sup>. Similmente agli incontri wireless dai quali estrarre i comportamenti sociali offline dei nodi, le reti sociali online come Facebook, Twitter e LinkedIn, giusto per fornire alcuni esempi, stanno aumentando la disponibilità di dati addizionali utili per analizzare il comportamento sociale generale dei nodi nelle reti opportunistiche. Da questo punto di vista, l'analisi della socialità derivate sia dagli incontri wireless e dai dati online diventa un aspetto fondamentale all'interno di queste reti. Esistono diversi lavori che studiano gli incontri delle persone e le loro relazioni sociali, ma mentre molti di questi si focalizzano sull'analisi della socialità offline estratta dagli incontri rilevati dalla vicinanza wireless, o da relazioni sociali online come le amicizie Facebook o le interazioni Twitter, pochi lavori analizzano congiuntamente questi aspetti.

Nella tesi, differentemente rispetto agli approcci menzionati, mi sono focalizzato sull'analisi delle relazioni tra la socialità offline costruita sugli incontri wireless e la socialità online, in cui entrambi questi livelli sono costruiti sullo stesso gruppo di utenti. Negli ultimi anni, diversi dataset reali di mobilità per le reti opportunistiche sono stati collezionati per esplorare la mobilità di nodi trasportati dall'uomo e la socialità. Tali tracce dati, che includono in alcuni casi i profili sociali online dei nodi, sono tipicamente acquisite mediante esperimenti che monitorano un set di partecipanti che trasportano piccoli dispositivi mobili in campus, conferenze e svariati ambienti. La maggior parte di questi dati possono essere ottenuti attraverso l'archivio digitale CRAWDAD.

---

Nonostante il lavoro fatto per analizzare queste tracce dati mostri diversi importanti aspetti sui dati di mobilità, la relazione tra la socialità costruita sulla mobilità e sulle altre dimensioni sociali non è stata ancora completamente esplorata. In base agli studi fatti, la conoscenza quanto più ampia sul comportamento sociale degli utenti mobili è essenziale per creare efficienti algoritmi di instradamento dell'informazione nelle reti opportunistiche. In base a ciò, il cuore di uno dei filoni di ricerca proposti è rappresentato dall'uso di un approccio d'analisi basato su rete multilivello per comparare differenti livelli di rete sociale estratti da un set eterogeneo di sei tracce di mobilità che coprono diversi ambienti di rete: accademico, conferenza e scenari urbani.

Sociologi, antropologi e psicologi hanno largamente studiato il comportamento umano usando due diversi approcci d'analisi. L'approccio sociocentrico deriva dalla sociologia e implica, prendendo in considerazione un gruppo delineato di persone, la quantificazione delle relazioni all'interno di esso, mentre l'approccio d'analisi egocentrico deriva dall'antropologia; tale forma di analisi sociale della rete è quasi sempre incentrata sulle persone più che sui gruppi. Le reti egocentriche sono definite come delle reti formate di un singolo attore unitamente agli attori con i quali egli è direttamente connesso, definiti come vicini. Entrambi gli approcci permettono di cogliere aspetti importanti sui legami sociali tra persone. Uno degli obiettivi del mio lavoro è stato quello di studiare il comportamento sociale dei nodi usando entrambe le misure sociocentriche ed egocentriche<sup>11</sup>. Nello specifico sarà presentata un'analisi dettagliata di sei dataset di reti sociali mobili opportunistiche contenenti due livelli di rete sociale:

- Il grafo di rete sociale costruito sugli incontri wireless offline
- Il grafo di rete sociale costruito sulle amicizie Facebook

Uno dei dataset analizzati è stato generato mediante un esperimento che mi ha coinvolto in prima persona, portato avanti su un campione di studenti dell'università della Calabria che sarà discusso in seguito. In generale per tutti i dataset, applicando un approccio di tipo multi livello, si è mirato a contribuire all'allargamento della conoscenza sulla similarità tra i mondi online e offline nei differenti ambienti di reti opportunistiche. Questo ha rappresentato un miglioramento nell'analisi rispetto agli

---

studi presenti in letteratura<sup>12</sup>. Considerare una collezione di sei dataset differenti ha permesso di trarre conclusioni più significative rispetto ad analisi che ne prendono in considerazione solo uno.

A causa della natura intrinseca delle reti opportunistiche, i cui nodi sono prevalentemente dispositivi mobili con risorse limitate, si è passato ad un secondo filone di ricerca che verte sull'analisi del comportamento egoistico o selfish dei nodi che costituiscono tali reti. Appare probabile infatti che non tutti i nodi di tali reti garantiranno sempre un comportamento altruistico e partecipativo nella rete, ma tenderanno a preservare le proprie risorse. Gli effetti di tali comportamenti sono stati approfonditi durante il mio percorso di studi e hanno condotto alla proposta di un algoritmo per mitigare gli effetti negativi in tali scenari.

La tesi è organizzata come segue:

- Nel capitolo 1 sono illustrati i principali modelli di rete sociale presenti in letteratura e viene fornita la definizione utilizzata di rete sociale multilivello e il modello sfruttato nei successivi capitoli.
- Nel capitolo 2 vengono fornite le specifiche dei dataset presi in considerazione e viene applicato su di essi il modello di rete sociale multilivello esplicito nel primo capitolo
- Nel capitolo 3 vengono illustrati i principali strumenti di analisi utilizzati, sia in termini di centralità che di community detection e sono forniti i risultati ottenuti in termini di correlazione nelle misurazioni di centralità multilivello e communities similarity con relativa analisi e interpretazione dei risultati
- Nel capitolo 4 viene approcciato il problema della selfishness nelle reti opportunistiche e dimostrati i suoi effetti negativi; vengono, inoltre, presentati alcuni metodi per la mitigazione di tale problematica
- Nel capitolo 5 viene presentato SORSI, algoritmo social – based per l'instradamento nelle reti opportunistiche con meccanismo di rilevazione della selfishness



# Capitolo 1

## Modelli di rete sociale:

## dalle online social network alle reti

## Multilivello

### 1.1 Modelli di rete sociale

L'analisi delle social network è uno degli argomenti di maggior rilievo all'interno delle odierne ricerche scientifiche; la modellazione di tali reti complesse è stata portata avanti attraverso gli studi di Paul Erdős e Alfréd Rényi sui grafi. Nel 1959 i due scienziati descrissero attraverso il concetto di grafo casuale<sup>13</sup> la topologia di rete complessa. Secondo la loro teoria in tali grafi ogni nodo possiede pari probabilità di avere un arco verso un'altra coppia di nodi indipendentemente dagli altri archi presenti. Ad esempio dati  $N$  nodi, con  $N$  molto maggiore di 1, per ottenere una rete secondo il modello Erdős-Rényi (ER) occorre collegare con pari probabilità  $p$  tutte le coppie di nodi. Gli archi ottenuti saranno dunque circa:

$$\frac{pN(N-1)}{2}$$

Le formulazioni di tale modello sono due; nel primo caso si considera un grafo  $\Gamma_{n,N}$  che possiede  $n$  vertici etichettati  $P_1, P_2, \dots, P_n$  e  $N$  archi. Si suppone che questi  $N$

---

archi siano scelti casualmente fra  $\binom{n}{2}$  archi possibili, così che tutte le  $\binom{\binom{n}{2}}{N} = C_{n,N}$  scelte possibili siano equiprobabili. Così, se  $G_{n,N}$  denota uno qualsiasi dei  $C_{n,N}$  grafi formati dagli  $n$  punti etichettati ed aventi  $N$  archi, la probabilità che il grafo casuale  $\Gamma_{n,N}$  sia identico a  $G_{n,N}$  è  $\frac{1}{C_{n,N}}$ . Un'altra formulazione equivalente è la seguente: si suppone che siano dati  $n$  vertici etichettati  $P_1, P_2, \dots, P_n$ . Si scelga un arco a caso fra gli  $\binom{n}{2}$  archi possibili, cosicché tutti tali archi siano equiprobabili. Successivamente si scelga un altro arco fra i rimanenti  $\binom{n}{2} - 1$  e si continui questo processo cosicché, se già fissati  $k$  archi, uno qualsiasi degli  $\binom{n}{2} - k$  archi rimanenti abbia ugual probabilità di essere scelto come successivo.

Tale modellazione proposta dai due scienziati non tiene conto del fatto che la maggior parte delle reti complesse presenti nella vita reale risultino essere non regolari e, fattore ancora più determinante, non completamente casuali. Tuttavia, a causa della scarsità di potenza computazionale e di informazioni topologiche dettagliate riguardo le reti su larga scala, tale modello ha rappresentato un riferimento fino agli anni 90. Successivamente grazie alla grande mole di dati che è stato possibile acquisire riguardo le reti di grandi dimensioni, Albert-László Barabási e Réka Albert hanno dimostrato che, indipendentemente dalla natura del sistema e dall'identità dei suoi costituenti, la probabilità  $P(k)$  che un vertice nella rete sia connesso a  $k$  altri vertici si comporta come una legge di potenza<sup>14</sup>, seguendo  $P(k) \sim k^{-\lambda}$ . Ciò implica che le reti di grandi dimensioni si auto-organizzano in uno stato privo di scala o scale-free, caratteristica non presente nelle reti casuali precedentemente teorizzate<sup>15</sup>. Infatti, caratteristica comune a tutti questi modelli è che la distribuzione della connettività dei nodi  $P(k)$  abbia un andamento esponenziale, con un picco ad un valore medio. Barabási ed Albert hanno rilevato che esistono due caratteristiche che non si confanno alle reti reali, ovverosia che si parta con un numero fisso di vertici ( $N$ ) che vengono casualmente collegati senza mai variarne il numero. Nel mondo reale, la maggior parte delle reti sono dinamiche nel senso che durante la loro vita il numero di nodi che le compongono varia, dunque  $N$  cresce durante il periodo di vita della rete. Esempio tipico può essere rappresentato

---

dalla rete internet che cresce esponenzialmente nel corso del tempo grazie alla continua aggiunta di nuove pagine Web.

Ulteriore aspetto da tenere in considerazione è che nelle reti in esame la probabilità che due nodi siano connessi tra loro sia uniforme e casuale, ma questo non è sempre vero in quanto in molte reti vale il concetto di connettività preferenziale. In altre parole, è più probabile che all'interno di una pagina Web di recente creazione siano presenti link a file e documenti più conosciuti con una connettività alta o che ad esempio in un atto sia citata una fonte più conosciuta rispetto ad una meno quotata. Questi esempi mettono in risalto il fatto che non è uniforme la probabilità con cui un nuovo nodo si connette ai già esistenti, ma sarà più probabile che si colleghi ad un nodo maggiormente quotato, con una quantità rilevante di connessioni.

Tenendo presente le assunzioni e considerazioni proposte Barabási ed Albert svilupparono un modello per la generazione di reti che gode della proprietà scale-free. Si inizia con un numero ridotto di nodi ( $m_0$ ) e ad ogni intervallo temporale si procede ad aggiungere un nuovo nodo con  $m \leq m_0$  archi che lo connettono ad  $m$  nodi diversi già presenti nel sistema; il tutto per mappare la caratteristica di crescita della rete. Per quanto riguarda invece la caratteristica di annessione preferenziale, si assume che la probabilità  $P$  che un nuovo nodo sia connesso ad  $i$  dipenda dal grado, ovvero dal numero di archi in esso incidenti,  $k_i$  di quel nodo. In tal modo si otterrebbe:

$$P(k_i) = \frac{k_i}{\sum_j k_j}$$

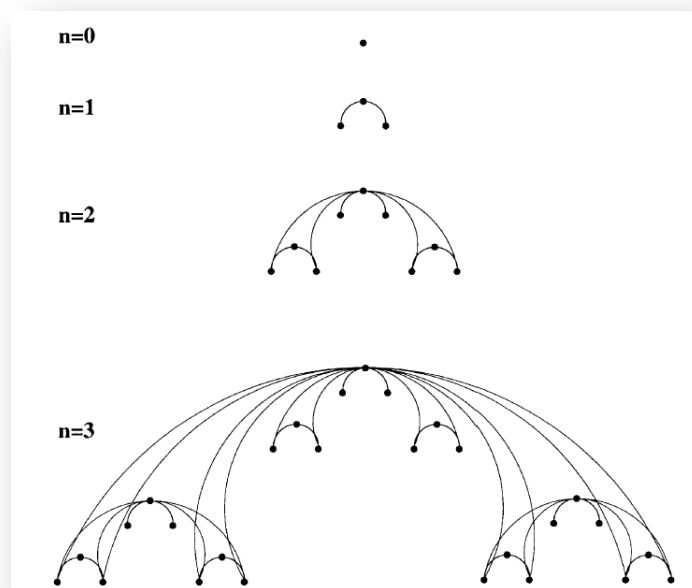
La probabilità che un nodo della rete sia connesso a  $k$  altri nodi si comporta come una legge di potenza ed è indipendente rispetto al tempo poiché, nonostante la crescita continua, il sistema stesso si organizza in uno stato stazionario scale-free.

In aggiunta all'algoritmo descritto, sempre Barabási ne descrive un altro per la generazione di grafi scale-free in maniera deterministica. La generazione della rete avviene attraverso iterazioni che utilizzano quanto generato nei passi precedenti secondo il seguente schema a passi:

- a Un nodo viene designato come radice del grafo.
- b Vengono aggiunti due ulteriori nodi, connessi alla radice.
- c Si aggiungono poi due gruppi di tre nodi ciascuno, ognuno dei quali esattamente uguale alla rete create nel passo 1, connettendo ognuno dei nodi inferiori di questi due gruppi alla radice come mostrato in fig.1. Dunque la radice possiederà quattro nuovi collegamenti.
- d Si aggiungono due gruppi composti di nove nodi ciascuno, identici ai gruppi generati nel passo precedente e si connettono tutti gli otto nodi inferiori dei due nuovi gruppi alla radice.

Generalizzando per il passo n-esimo:

- n Si aggiungono due gruppi di  $2^{n-1}$  nodi ciascuno, identici alla rete creata nel passo  $n - 1$ , e si connette ognuno dei  $2^n$  nodi inferiori di questi gruppi alla radice della rete.



**Figura 1 Esempio di costruzione per una rete scale-free deterministica per i primi 4 passi.**

Analizzando la distribuzione del grado dei nodi si determina il valore  $\gamma$  dell'algoritmo; nel dettaglio al passo  $i$  il grado dell'hub più connesso, la radice, è  $2^{i+1}-2$ . Nella successiva iterazione le due copie dell'hub vengono scritte nei due

---

gruppi aggiunti. Procedendo con l'iterazione all'n-esimo passo saranno presenti  $3^{n-i}$  copie dell'hub. In realtà le copie appena generate non cresceranno di grado con ulteriori passaggi, dunque dopo n iterazioni i nodi saranno pari a:

$$\frac{2}{3} 3^{n-1}$$

con grado  $s^{i+1} - 2$ . Dunque la distribuzione del grado segue:

$$P(k) \sim k^{-\frac{\ln 3}{\ln 2}}$$

Dunque l'esponente del grado è  $\gamma = \frac{\ln 3}{\ln 2}$ , che rimane fisso; per variare l'indice basta variare il numero di collegamenti connessi con la radice ad ogni passo<sup>16</sup>.

## 1.2 Online Social Network

Le OSN possono essere definite come servizi per lo più Web – based che permettono agli individui di costruire un profilo pubblico o semi pubblico all'interno di un sistema circostanziato, articolare una lista di utenti con i quali condividere una connessione e vedere e navigare le loro liste di connessioni. La natura e la nomenclatura di queste connessioni può variare da OSN ad OSN. Ciò che le rende uniche è che permettano agli utenti di articolare e rendere visibile le proprie reti sociali e di interesse. Questo può risolversi nella generazione di connessioni tra individui che altrimenti non si sarebbero mai incontrati e conosciuti, anche se, in molte OSN gli utenti vogliono prevalentemente comunicare con persone che fanno già parte della loro rete sociale reale. Ogni OSN possiede una serie di peculiarità tecniche, ma la struttura di base è formata dai profili visibili degli utenti, che mostrano una lista articolata di amici che sono altresì utenti del sistema. I profili sono pagine uniche in cui l'individuo può descrivere se stesso. Dopo essersi registrato, tipicamente tramite l'inserimento di username e password, all'utente verranno sottoposte una serie di domande attraverso le quali verrà generato un profilo; tale profilo conterrà elementi quali l'età, la sede, gli interessi e una sezione specifica alla quale genericamente si può fare riferimento col nome “about me”. La maggior parte delle OSN incoraggiano gli utenti ad effettuare l'upload di una foto per il profilo ed

---

alcune di esse permettono di migliorare il profilo aggiungendo contenuti multimediali o modificando caratterizzandolo l'aspetto del profilo. Altre ancora consentono di caricare dei moduli o applicazioni esterne per ottenere funzionalità o semplicemente dei giochi.

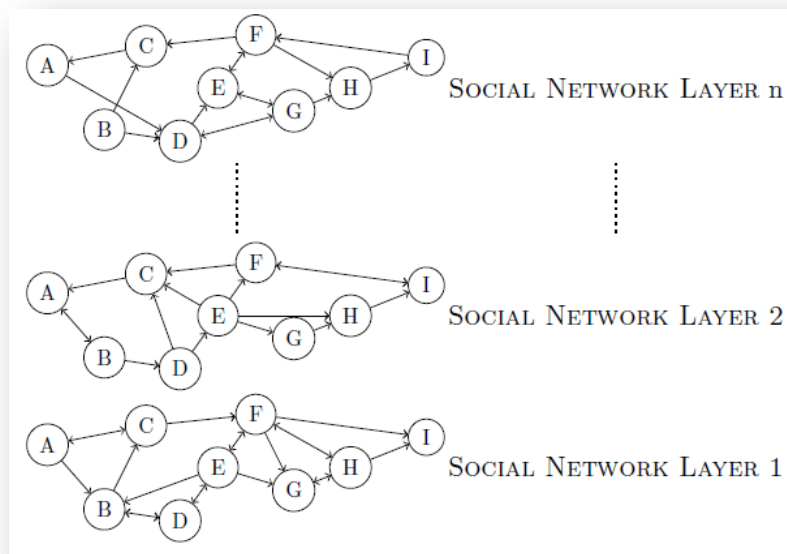
La visibilità di un profilo varia secondo le scelte dell'utente; di default, i profili sono ricercabili tramite motori di ricerca e ciò li rende assai più fruibili. Ogni utente, dopo essersi registrato e aver fatto l'accesso, viene invitato ad identificare e ricercare altri utenti nel sistema con i quali potrebbe avere un qualche legame o relazione. Le etichette per questo tipo di relazioni differiscono da OSN ad OSN. In generale il termine utilizzato è "amico" anche se tale definizione potrebbe fuorviare in quanto la connessione non significa necessariamente amicizia nel senso comune, ma la ragione per cui le persone possono aggiungersi sono molteplici. La lista delle amicizie contiene i link ai profili di ogni utente, permettendo ai visitatori di attraversare il grafo della rete cliccando sulla lista stessa. Sulla maggior parte dei siti la lista degli amici è visibile a chiunque abbia accesso al profilo stesso, anche se esistono delle eccezioni. La maggior parte delle OSN permettono anche un sistema di messaggistica istantanea e offline; più genericamente è possibile lasciare un messaggio sul profilo dell'utente contattato. Questa particolare caratteristica determina anche la possibilità che qualcun altro possa a sua volta commentare ciò che viene scritto sul profilo. Differente è il messaggio privato che non può essere visualizzato da nessun altro all'infuori dei due utenti interessati. Altra caratteristica più o meno comune è quella relativa alla condivisione di foto o video. Sebbene le OSN nascano con l'intento di essere accessibili a tutti, molti gruppi di utenti le utilizzano per creare delle comunità ristrette generando dei gruppi chiusi accessibili tramite invito e con motivazioni aggreganti basate su nazionalità, età, appartenenza ad un determinato gruppo d'istruzione o altri fattore che tipicamente caratterizzano la società, sebbene questo non fosse previsto in origine dai creatori del design delle OSN.

### 1.3 Definizione di una rete sociale multilivello

Se si estende il numero di contesti sociali e si considerano diverse reti sociali per un particolare gruppo di utenti, ciò che si ottiene è una struttura multilivello che rappresenta le connessioni di un singolo utente verso gli altri utenti su diversi livelli autonomi. Due utenti potrebbero essere contemporaneamente connessi su diversi livelli, ad esempio sulla rete Bluetooth e sulle reti Facebook, LinkedIn e Twitter, mentre gli altri utenti potrebbero essere connessi solo su un livello, come ad esempio dei collaboratori connessi solo tramite LinkedIn o degli amici connessi solo tramite Facebook. Il risultato è un sistema complesso in cui ci sono diversi livelli di rete sociale e in cui gli utenti si ritrovano coinvolti in diverse tipologie di relazione. La definizione<sup>17</sup> di un modello di rete sociale multilivello si basa su grafi in cui è possibile utilizzare i pesi sugli archi per definire la forza di una relazione.

**Definizione 1:** un livello di rete sociale  $L$  è un grafo  $G ( V, E )$ , i cui vertici  $V$  corrispondono agli utenti sulla rete sociale, mentre gli archi  $E$ , con  $E \subseteq V \times V$ , corrispondono ai link sociali tra utenti.

**Definizione 2:** una rete sociale multilivello è una tupla  $MLSN = ( L_1, L_2, \dots, L_n )$  in cui  $L_i = G_i ( V, E_i )$ , con  $i \in 1, \dots, n$ , sono i livelli di rete sociale.



# Capitolo 2

## Dataset e applicazione del modello di rete sociale multilivello

### **2.1 Descrizione dei dataset**

Sono stati presi in considerazione una collezione di cinque dataset reali che includono sia i dati relativi alla mobilità che i dati relativi alla socialità online. A questi è stato poi aggiunto il dataset relativo ad un esperimento portato avanti personalmente presso l'Università della Calabria tramite cui è stato possibile collezionare dati su un campione di studenti di un corso specialistico. In particolare sono stati tracciati i contatti wireless degli studenti stessi in termini di incontri Bluetooth intercorsi tra i nodi all'interno e all'esterno del gruppo di candidati che trasportavano smartphones. Inoltre è stato collezionato un trace dati, ovverosia un insieme di dati strutturati, contenente le amicizie Facebook tra i medesimi individui e un trace inerente i loro interessi. La maggior parte di questi dataset è accessibile sull'archivio CRAWDAD. In tab.1 vengono riassunte le principali caratteristiche dei dataset selezionati. Per ogni dataset sperimentale, il gruppo di ricercatori che ha portato avanti l'esperimento ha istruito i candidati affinché portassero con sé i



---

dispositivi wireless, rappresentati da telefoni o sensori, affinché venissero rilevati e creati i log dei nodi che si trovavano nel range di comunicazione durante tutta la durata dell'esperimento. Poiché i diversi dataset hanno durate differenti, si è scelto in una successiva fase di analisi di uniformarli considerando una durata massima di sette giorni, o, nel caso di esperimenti di durata inferiore, l'intera durata degli stessi.

In particolare, per i dataset più lunghi, l'analisi si è focalizzata sulla settimana di contatti wireless aventi durata di contatto maggiore, come si può vedere nella riga *Analyzed week* della tabella. Dunque, sono stati analizzati parte dei dati sui contatti wireless contenenti le interazioni più lunghe tra i dispositivi mobili. Come conseguenza, il numero totale di nodi, indicati nella riga *Overall # of nodes*, si è ridotta a causa dell'assenza di parte di essi durante la settimana considerata, come si può vedere nella riga *# of Analyzed nodes*. Scegliendo i collegamenti con la maggiore durata di contatto, si è assunto che tale scelta rappresenti uno scenario favorevole per lo scambio messaggi. Tale considerazione viene avvalorata quando, ad esempio, viene misurata la centralità su di un grafo i cui link rappresentano un elevato rate di contatto. In questo scenario un nodo con un'elevata degree centrality potrebbe essere considerato più centrale, ma tale misura potrebbe risultare falsata se la durata di contatto tra il nodo in questione e un altro nodo della rete fosse in realtà molto breve e non adeguata per lo scambio di messaggi. Un esempio pratico di ciò potrebbe essere un individuo che si sposta da lavoro verso casa tramite autobus. Il suo smartphone può effettivamente incontrare diversi altri dispositivi nel range di trasmissione, ma per pochi secondi passando da una fermata all'altra. Queste brevi interazioni non rendono questo nodo socialmente più idoneo alla trasmissione di informazioni. Infatti, scegliendo questo nodo come next hop per una trasmissione, potrebbe non esserci il tempo necessario ad instaurare una connessione adeguata per lo scambio messaggi se esso rileva un nodo attraverso ad esempio la sua interfaccia Bluetooth e successivamente la stessa connessione viene persa. Anche se il tempo di setup della connessione è sufficiente, si dovrà comunque procedere a frammentare il messaggio determinando un sovraccarico della rete e dei buffer dei nodi che avranno diverse copie del messaggio.

Tabella 1 Caratteristiche delle tracce dati wireless

<b>Experimental Dataset</b>	<b>UNICAL</b>	<b>UPB</b>	<b>LAPLAND</b>	<b>SASSY</b>	<b>Social Evolution</b>	<b>SIGCOMM</b>
<b>Environment</b>	Academic	Academic	Conference	Academic/ Urban	Academic/ Urban	Conference
<b>Device Type</b>	Phone	Phone	I-mote	T-mote	Phone	Phone
<b>Radio Range</b>	~ 10 m	~ 10 m	~ 10 m	~ 10 m	~ 10 m	[10-20] m
<b>Granularity</b>	180 s	[5–30] min	[120–600] s	6 .67 s	360 s	[120 ± 10 .24] s
<b>Overall Duration</b>	7 days	35 days	3 days	70 days	352 days	5 days
<b>Analyzed week</b>	from 28/01 to 22/02 2014	from 18/11 to 24/11 2011	from 09/08 to 11/08 2009	from 08/03 to 14/03 2008	from 02/03 to 08/03 2010	from 17/08 to 21/08 2009
<b>Overall # of nodes</b>	15	22	17	27	70	76
<b># of analyzed nodes</b>	15	15	17	24	55	67

A seguire una breve descrizione per ogni singolo dataset analizzato.

---

### 2.1.1 UNICAL

Il dataset UNICAL<sup>18</sup> contiene dati di prossimità Bluetooth collezionati attraverso una applicazione Android ad – hoc chiamata *SocialBlueConn*, i profili sociali in termini di amicizie Facebook e interessi di un gruppo di 15 studenti dell'Università della Calabria (Italy). L'esperimento, che ho personalmente portato avanti, ha avuto una durata di una settimana durante un corso specialistico, dal 28 Gennaio 2014 al 5 Febbraio 2014. Per ottenere le informazioni di prossimità, ogni partecipante è stato istruito a mantenere il dispositivo attivo dalle 12 AM alle 8 PM. Per collezionare i dati Facebook, ai partecipanti è stato chiesto di loggarsi con le loro credenziali Facebook ad un sito ad – hoc che utilizzando le API di Facebook. Una volta che gli studenti hanno effettuato il log in, le loro liste amici e i loro profili sociali sono stati collezionati e mandato ad un server centrale.

### 2.1.2 UPB

Il dataset UPB<sup>19</sup> si compone di una traccia dati relativa alle amicizie Facebook e una traccia relativa ai contatti Bluetooth. Questi dati sono stati collezionati durante un esperimento di 35 giorni, dal 18 Novembre 2011 al 22 Dicembre 2011. L'esperimento ha avuto luogo in ambiente accademico presso l'Università Politehnica of Bucharest (Romania). I 22 candidati reclutati sono stati istruiti affinché installassero sul proprio dispositivo Android un'applicazione che permettesse di loggare i contatti Bluetooth e mantenessero accessi i dispositivi dalle 8 AM alle 8 PM.

### 2.1.3 LAPLAND

Il dataset LAPLAND<sup>20</sup> è stato collezionato durante il workshop ExtremeCom09 tenutosi nel National Park a Padjelanta in Svezia. Esso contiene la traccia dati di posizionamento di 17 attendenti alla conferenza raccolta durante i quattro giorni consecutivi dell'esperimento, dal 9 al 12 Agosto 2009. In particolare sono stati analizzati solo i primi tre giorni in quanto durante il quarto giorno vi sono state poche connessioni. Ad ogni candidato è stato chiesto di portare il dispositivo

---

iMotes con sé per rilevare gli altri dispositivi nel range di comunicazione. Inoltre il dataset include la lista degli amici Facebook e gli interessi in termini di topic scientifici.

#### **2.1.4 SASSY**

Il dataset SASSY<sup>21</sup> si compone dei record di incontri relativi ad un gruppo di 27 partecipanti, di cui 22 studenti universitari, 3 studenti post laurea e 2 membri dello staff, che trasportavano sensori IEEE 802.15.4 (dispositivi T-mote), e le loro liste amici autodichiarate ad inizio esperimento. L'esperimento ha avuto luogo presso la "University of St. Andrews" (United Kingdom) per una durata complessiva di tre mesi tra il 15 Febbraio 2008 e il 29 Aprile 2008.

#### **2.1.5 Social Evolution**

Il dataset MIT Social Evolution<sup>22</sup> è relativo ad un esperimento effettuato per tracciare il più fedelmente possibile la vita di ogni giorno di un intero dormitorio di 70 studenti universitari. La durata dell'esperimento copre un intero anno accademico dal 6 Aprile 2009 al 23 Marzo 2010. Il dataset include dati Facebook, di prossimità, di posizione e log di chiamata, collezionati attraverso un'applicazione mobile per cellulare che scansionava i vicini access point Wi-Fi e i dispositivi Bluetooth.

#### **2.1.6 SIGCOMM**

Il dataset SIGCOMM<sup>23</sup> include i dati di posizionamento Bluetooth di 76 attendenti ad una conferenza collezionati da una applicazione social cellulare, chiamata *MobiClique*<sup>24</sup>. Ad ogni partecipante è stato chiesto di loggarsi in Facebook per ottenere la sua lista di amicizie e interessi. Questi dati sono stati collezionati durante la conferenza SIGCOMM tenutasi a Barcellona in Spagna, dal 17 al 21 Agosto 2009.

---

## 2.2 Applicazione del modello multi livello ai dataset

Partendo dai dati di mobilità e dai dati Facebook, si è proceduto a costruire un grafo di rete sociale multi livello per ogni dataset analizzato. Successivamente su questa struttura è stata portata avanti un'analisi basata sugli approcci sociocentrico ed egocentrico discussa in seguito. Il grafo costruito a partire dai dati online è non pesato per la natura dei link relativi alle amicizie Facebook. Riprendendo le definizioni fornite e caratterizzandole otteniamo:

- **Definizione 1: (Livello di rete Sociale)** Un livello di rete sociale  $L$  è un grafo non pesato  $G(V, E)$  che possiede un set di vertici  $V$  corrispondenti agli utenti sulla rete sociale e un set di archi  $E \subseteq V \times V$  corrispondenti ai collegamenti sociali o amicizie, tra gli utenti
- **Definizione 2: (Rete sociale multi livello)** Una rete sociale multilivello  $MLSN = (L_1, L_2, \dots, L_n)$  è una tupla in cui  $L_i = G_i(V, E_i)$ , con  $i \in 1, \dots, n$ , sono i livelli di rete sociale.

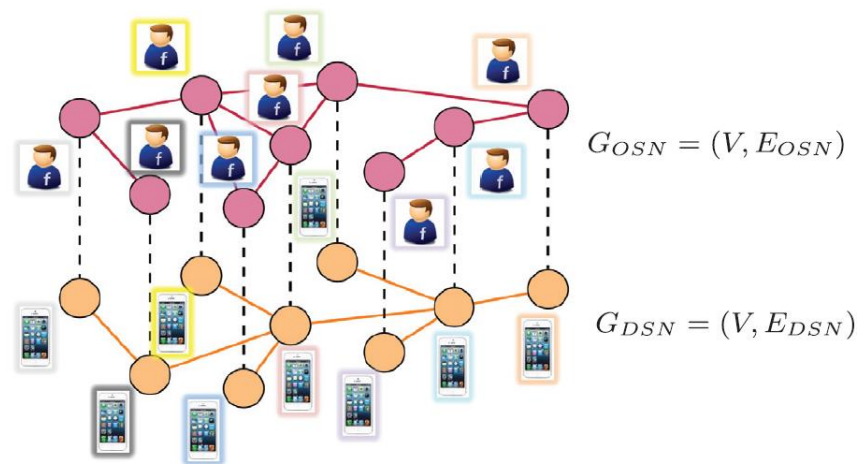
Per ogni dataset si è proceduto a formare una rete sociale a due livelli composta come segue:

- Un primo livello sociale composto dalla rete sociale offline rilevata attraverso gli incontri wireless, definita detected social network (DSN).
- Un secondo livello sociale composto dalla rete delle amicizie Facebook, definito come online social network (OSN).

Utilizzando i dati Facebook dei partecipanti nella forma di una lista con una riga rappresentata da  $\{\#NODE ID1, \#NODE ID2, \#FRIENDSHIP FLAG\}$ , dove  $ola$  "friendship flag" indica se due nodi sono amici Facebook o meno, è stato creato il grafo OSN, dove un arco esiste se due nodi sono amici.

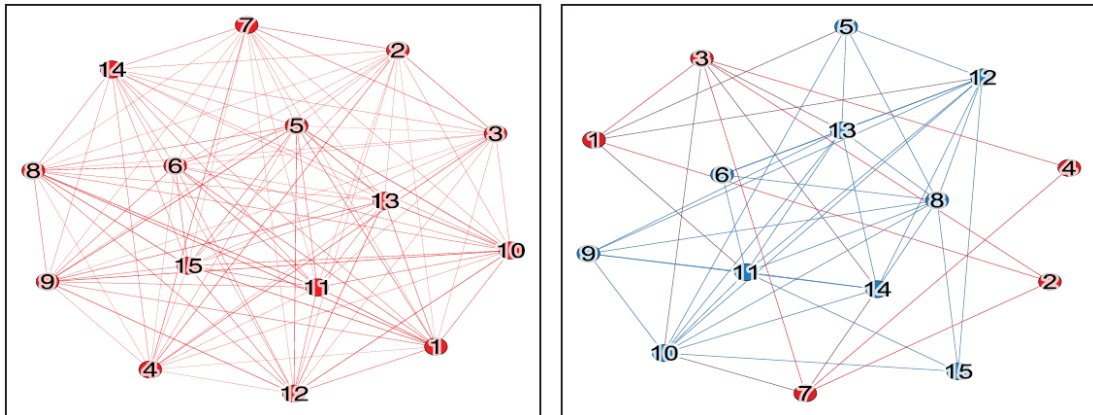
Allo stesso modo per come sono strutturati i dati relativi agli incontri wireless, la modellazione di un unico grafo di rete temporale è più complesso ed è un problema ancora aperto. Si è pertanto deciso di formare il grafo DSN stabilendo un arco tra due nodi se essi hanno avuto almeno un contatto durante la settimana analizzata, usando i

dati di contatto nella forma di una lista con una riga rappresentata da {#NODE ID1, #NODE ID2, #CONTACT TIMESTAMP}. Si sottolinea come il grafo DSN, anche se non pesato, è stato definito sulla finestra temporale della settimana in cui in ogni esperimento c'è stata la più elevata durata dei contatti. In altre parole, un collegamento tra due nodi nel grafo DSN rappresenta una durata di contatto elevata. In base a questa osservazione, se da un lato si perdono alcune informazioni sul comportamento sociale degli utenti, come ad esempio quanto sia durato il collegamento, d'altra parte si preserva l'aspetto relativo alla lunghezza del contatto ed è possibile comparare il grafo DSN con il grafo OSN.



**Figura 2** Esempio di rete sociale a due livelli in una rete sociale opportunistica mobile in cui ogni nodo mobile ha il suo corrispettivo sulla rete Facebook. I nodi possono avere collegamenti differenti per ogni livello

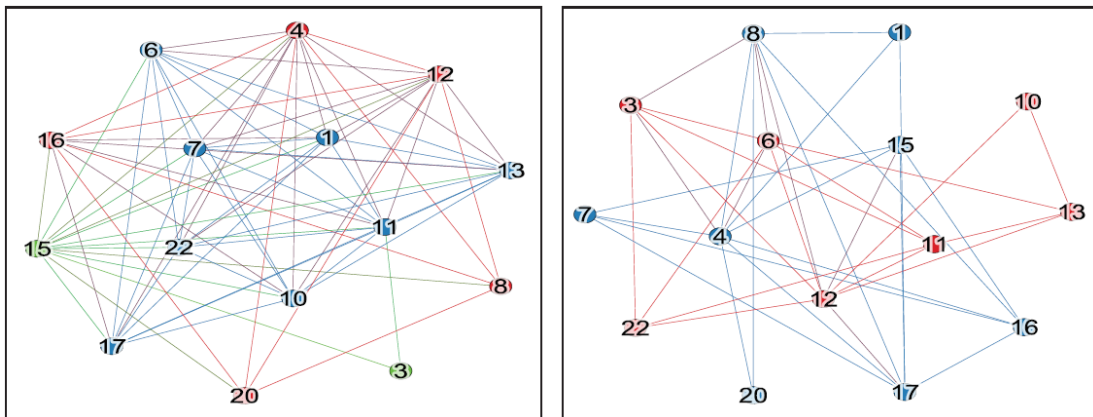
In fig.2 viene mostrato un esempio della struttura di rete multi livello utilizzato, mentre nelle fig.3,4,5,6,7,8, vengono rappresentati i grafi dei due livelli OSN e DSN utilizzando colori differenti per i nodi appartenenti a community differenti. Per individuare le community è stato utilizzato il metodo di Louvain discusso in seguito.



(a)

(b)

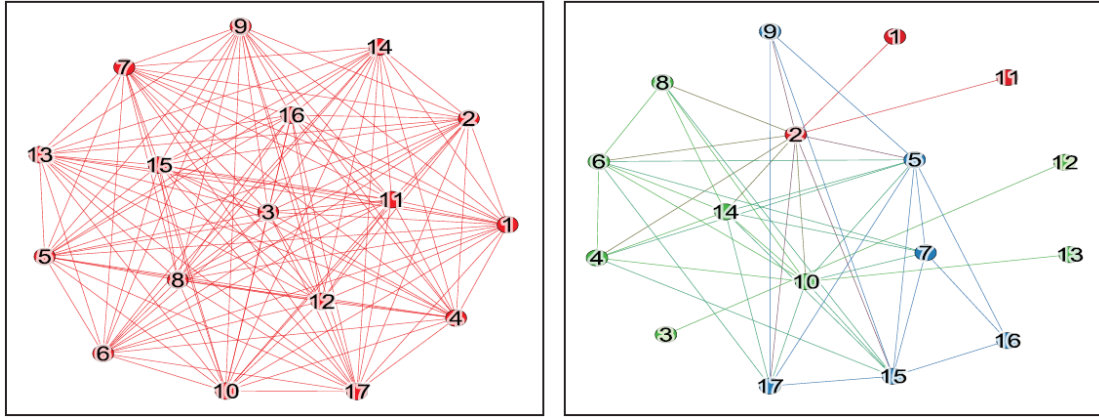
Figura 3 UNICAL (a) DSN and (b) OSN graph layers.



(a)

(b)

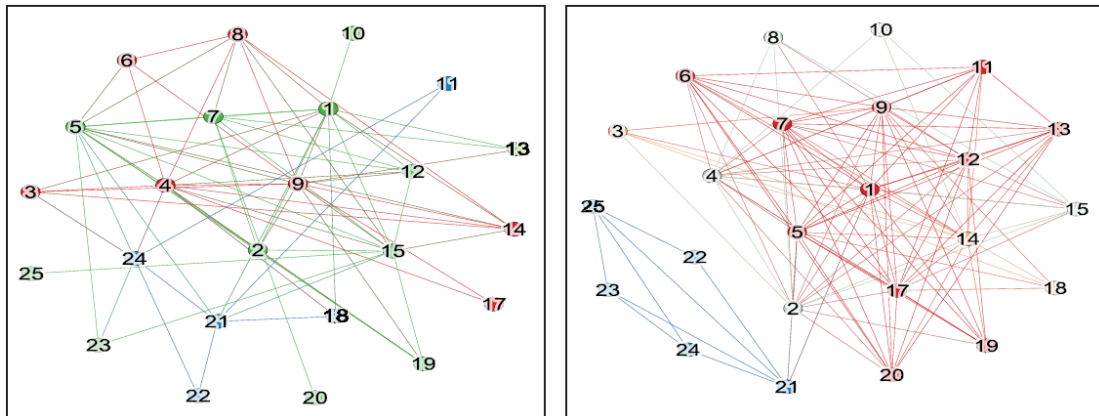
Figura 4 UPB (a) DSN and (b) OSN graph layers.



(a)

(b)

Figure 5 LAPLAND (a) DSN and (b) OSN graph layers.

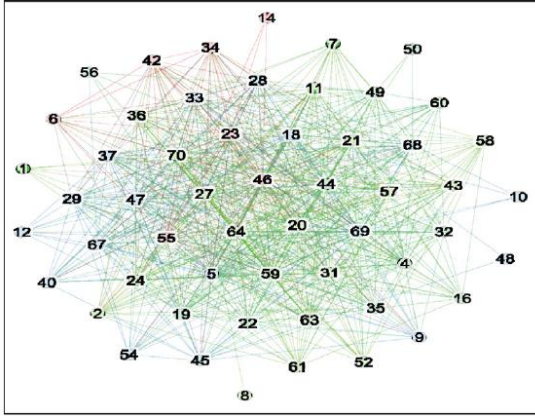


(a)

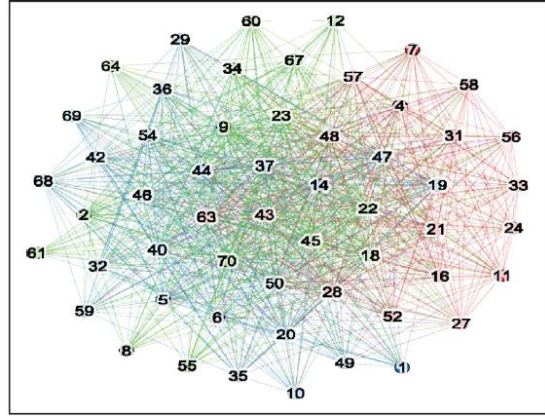
(b)

Figure 6 SASSY (a) DSN and (b) OSN graph layers.



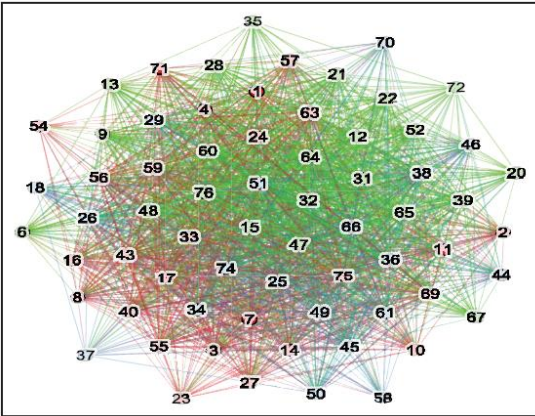


(a)

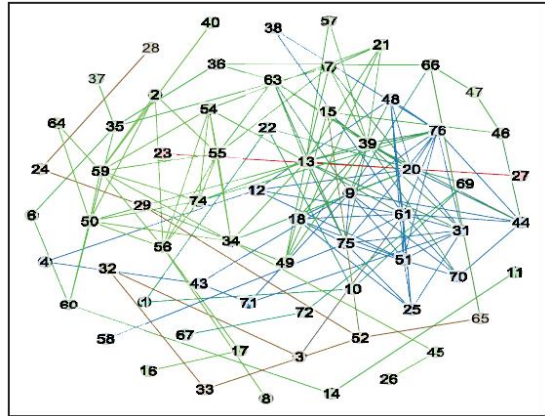


(b)

Figure 7 Social Evolution (a) DSN and (b) OSN graph layers.



(a)



(b)

Figure 8 SIGCOMM (a) DSN and (b) OSN graph layers.

# Capitolo 3

## Strumenti di analisi: approccio sociocentrico ed egocentrico

### **3.1 Introduzione**

Nella teoria delle reti sociali esistono due approcci distinti per l'analisi delle reti, che derivano da due tradizioni storiche distinte. L'approccio sociocentrico deriva dalla sociologia e implica, prendendo in considerazione un gruppo delineato di persone e la quantificazione delle relazioni all'interno di esso, mentre l'approccio d'analisi egocentrico deriva dall'antropologia; tale forma di analisi sociale della rete è quasi sempre incentrata sulle persone più che sui gruppi. Le reti egocentriche sono definite come delle reti formate di un singolo attore unitamente agli attori con i quali egli è direttamente connesso, definiti come vicini. Entrambi gli approcci permettono di cogliere aspetti importanti sui legami sociali tra persone.

### **3.2 Analisi Sociocentrica**

L'approccio d'analisi sociocentrico estende e complementa la scienza sociale tradizionale focalizzando l'attenzione sulla quantificazione delle interazioni tra un gruppo ben definito di persone e sull'identificazione di pattern strutturali globali. In

---

particolare, il metodo analizza gruppi di relazioni tra nodi che sono considerati come collettivi socialmente delimitati. Di seguito verranno brevemente descritte le misure di centralità sociocentrica e gli algoritmi di rilevazione delle community che sono stati scelti per valutare la similarità tra la rete online e la rete offline dei contatti. Le tre misure di centralità scelte sono considerate in letteratura le misure principali sui grafi per determinare l'importanza di un nodo all'interno della rete. Gli algoritmi di community detection scelti rappresentano due classi differenti, la prima è quella relativa al partizionamento delle reti senza sovrapposizioni di community e il secondo invece permette l'individuazione di community sovrapposte.

### 3.2.1 Betweenness centrality

La betweenness centrality<sup>25</sup>, anche definita sociocentric betweenness centrality, misura l'influenza che un nodo ha sul flusso di informazioni tra ogni coppia di nodi nel grafo rappresentante la rete, con l'assunzione che le informazioni fluiscano attraverso il percorso più breve che li colleghi. Formalmente, la betweenness centralità di un nodo  $i$  è definita come:

$$C_b(i) = \sum_{i \neq j \neq k} \frac{g_{jk}(i)}{g_{jk}}$$

Definita geodetica come curva più breve che congiunge due punti uno spazio,  $g_{jk}$  rappresenta il numero di percorsi geodetici da  $j$  a  $k$ ,  $g_{jk}(i)$  è il numero di percorsi più brevi da  $j$  a  $k$  che attraversano  $i$ .

### 3.2.2 Closeness centrality

La closeness centrality<sup>26</sup> misura l'inverso della somma dei percorsi più brevi di un nodo rispetto ad ogni altro.

$$C_c(i) = \frac{1}{\sum_{j=1}^N d(i,j)}$$

---

Dove  $d(i, j)$  è il percorso pesato più breve dal nodo di riferimento verso ogni nodo della rete. Una forte limitazione di questo parametro di misura è la mancanza di applicabilità verso le reti che risultano essere non completamente connesse: due nodi appartenenti a componenti diversi non hanno una distanza finita tra essi. Perciò, per superare tale problema, Dangalchev<sup>27</sup> ha ridefinito la formula in:

$$C_c(i) = \sum_{j \neq i} \frac{1}{2^{d(i,j)}}$$

### 3.2.3 Eigenvector centrality

La eigenvector centrality<sup>28</sup> è una misura di centralità definita in maniera circolare. La centralità di un nodo è proporzionale alla somma dei valori di centralità di tutti i suoi nodi vicini. In altre parole, una persona importante può essere caratterizzata da i propri link verso altre persone importanti. Questa misura è calcolata usando la matrice di adiacenza  $A$  del grafo non orientato per cercare i nodi centrali della rete. La eigenvector centrality per un nodo  $i$  è proporzionale alla somma dei valori di eigenvector centrality dei suoi nodi vicini. Essa è definita come:

$$C_e(i) = \frac{1}{\lambda} \sum_{j=1}^N A_{ij} C_e(j)$$

Dove  $\lambda$  rappresenta la più grande eigenvalue che assicuri che la centralità sia non negativa. Perciò,  $C_e(i)$  rappresenta l' $i$ esimo componente dell'eigenvector associato con il più grande eigenvector  $\lambda$  della rete.

### 3.2.4 Louvain community detection

Il metodo Louvain<sup>29</sup> suddivide il grafo di rete in community disgiunte; esso si basa su una tecnica di ottimizzazione greedy che cerca di ottimizzare la modularità di una partizione del grafo. Inizialmente, il metodo ricerca le community più piccole ottimizzando localmente la modularità. Successivamente, aggrega i nodi appartenenti alla stessa community e costruisce una nuova rete i cui nodi sono le community.

---

Questi passaggi sono ripetuti iterativamente finché viene raggiunto un valore massimo di modularità e viene prodotta la gerarchia delle community.

### 3.2.5 K-CLIQUE community detection

Tale metodo, anche conosciuto come “Clique Percolation Method” (CPM)<sup>30</sup>, ricerca le community sovrapposte in cui una community è definita come l’unione di tutte le k-cliques, ovverosia i sottografi completi con k nodi, che possono raggiungere gli uni con gli altri attraverso una serie di k-cliques adiacenti, in cui due k-cliques sono dette adiacenti se condividono k-1 nodi. Durante le analisi e a seguito di diversi esperimenti, si è scelto per lo studio portato avanti di settare k = 5 sia per la online network che per la offline network.

## 3.3 Analisi Egocentrica

Secondo tale approccio, se si vogliono capire le similarità e le differenze tra due reti focalizzandosi sui comportamenti degli individui, occorre prendere in considerazione le condizioni locali. Le reti egocentriche sono definite come delle reti formate da due elementi, il singolo nodo o attore di riferimento e i nodi o attori direttamente connessi ad esso. Il descrivere e quantificare la variazione tra gli individui in maniera tale da analizzare il modo in cui essi sono integrati in strutture sociali locali rappresenta l’obiettivo dell’analisi sociocentrica.

### 3.3.1 Degree centrality

La degree centrality<sup>31,32</sup> conta quante connessioni un nodo possenga e può essere considerata la base di ogni misura di centralità. Essa è definita per il nodo i come segue:

$$C_d(i) = \sum_{j=1}^N a_{ij}$$

Dove  $a_{ij} = 1$  se i nodi i e j sono connessi da un arco,  $a_{ij} = 0$  altrimenti.

---

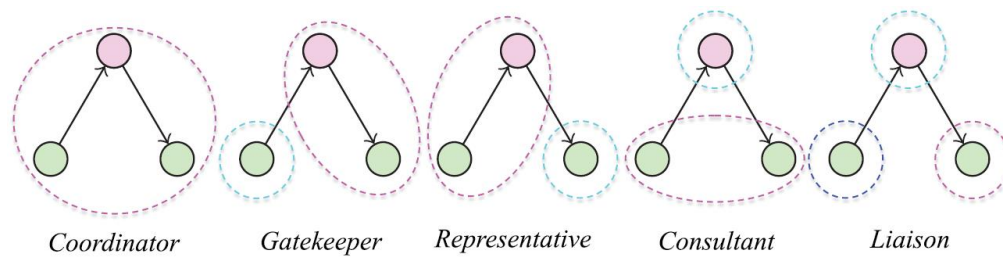
### 3.3.2 Ego betweenness centrality

La ego betweenness è calcolata usando solo la matrice di adiacenza ad un hop di un nodo, dunque il percorso più breve tra ogni coppia di nodi non adiacenti avrà dimensione 2. La metrica di ego betweenness centrality può essere calcolata efficientemente in maniera distribuita poiché per ogni nodo sono richieste solo informazioni locali. E' stato ampiamente dimostrato<sup>33,34</sup> che i valori di ego betweenness centrality hanno una forte correlazione ai valori di sociocentric betweenness per la maggior parte delle reti.

### 3.3.3 Brokerage

I ruoli che l'ego gioca nel connettere gli altri nodi in gruppi in qualità di intermediario è stato esplicito all'interno dei propri studi da parte di Gould e Fernandez<sup>35</sup>. Vengono dunque individuati cinque tipi di ruoli di intermediazione o di brokerage come mostrato in fig.9:

- Coordinator: il broker funge da mediatore tra due individui all'interno del proprio gruppo di appartenenza
- Gatekeeper: il broker funge da mediatore tra un contatto in arrivo da un gruppo esterno e un contatto presente in un gruppo interno
- Representative: il broker media tra un contatto in uscita da un gruppo interno verso un gruppo esterno
- Consultant: il broker funge da mediatore tra due individui appartenenti a gruppi differenti
- Liaison: il broker media tra due individui di differenti gruppi, nessuno dei quali appartenente al gruppo del broker stesso.



**Figura 9** Ruoli di mediazione, i nodi bianchi sono i brokers, le ellissi rappresentano i confini della comunità.

Il punteggio di brokerage per un dato nodo rispetto ad un dato ruolo rappresenta il numero di coppie ordinate aventi l'appartenenza all'appropriato gruppo mediato dal detto nodo.

### 3.3.4 Tie strenght

Il legame forte è una proprietà quantificabile che caratterizza i collegamenti tra due nodi. La nozione di legame forte, o tie strenght, fu introdotta da Granovetter<sup>36,37</sup> nel 1973 ed è definita come misura che prende in considerazione quattro differenti aspetti:

- Frequenza di contatto
- Durata di contatto
- Storia dei legami
- Numero di transazioni

Nello studio portato avanti la durata di contatto per ogni coppia di nodi è stata calcolata come la durata totale del contatto durante il periodo sperimentale.

## 3.4 Analisi dei risultati

### 3.4.1 Centralità e correlazione

Utilizzando le misure definite in precedenza, si è proceduto a calcolare i valori di centralità sulle reti multi livello costruite sui dataset definiti nel capitolo 2. Nello specifico, all'interno di ogni rete multi livello e per ogni misura di centralità considerata, è stato calcolato il coefficiente di correlazione di Pearson tra i valori di

centralità dei nodi sulla OSN e i valori di centralità sulle DSN. Il coefficiente di correlazione di Pearson è definito come:

$$\rho_{X,Y} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}$$

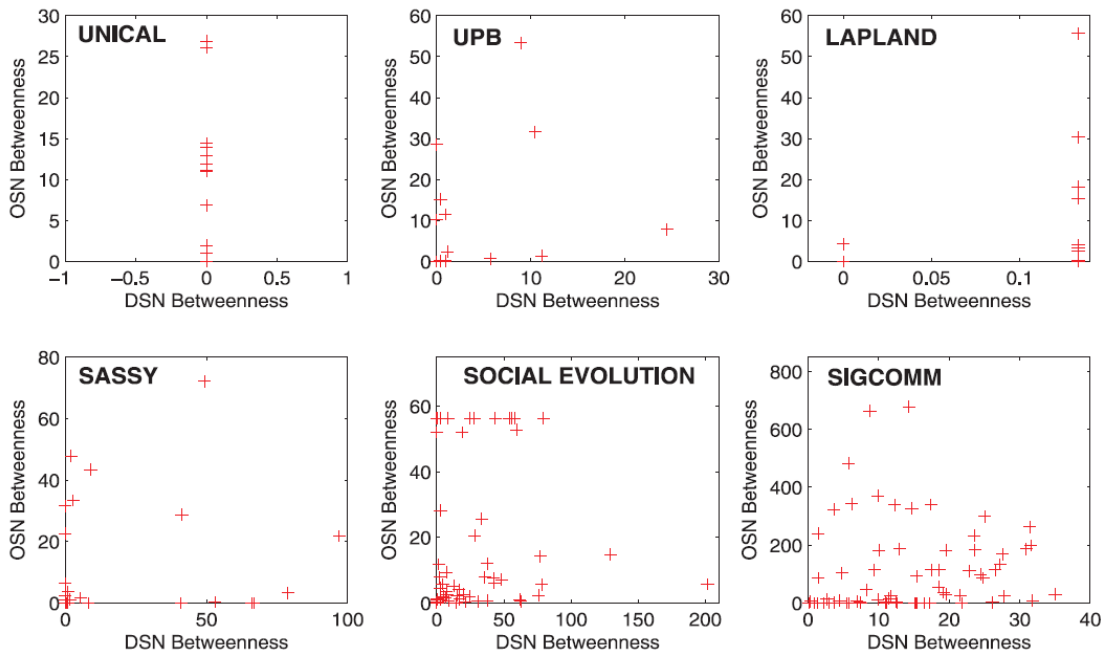
Dove  $COV(X, Y)$  rappresenta la covarianza tra le due variabili aleatorie  $X$  e  $Y$ , e  $\sigma_X$  e  $\sigma_Y$  sono le deviazioni standard. L'analisi di correlazione mira a ricercare relazioni lineari tra le stesse misure di centralità calcolate sui due livelli sociali.

**Tabella 2** Correlazione tra le misure di centralità calcolate sulle OSN e DSN

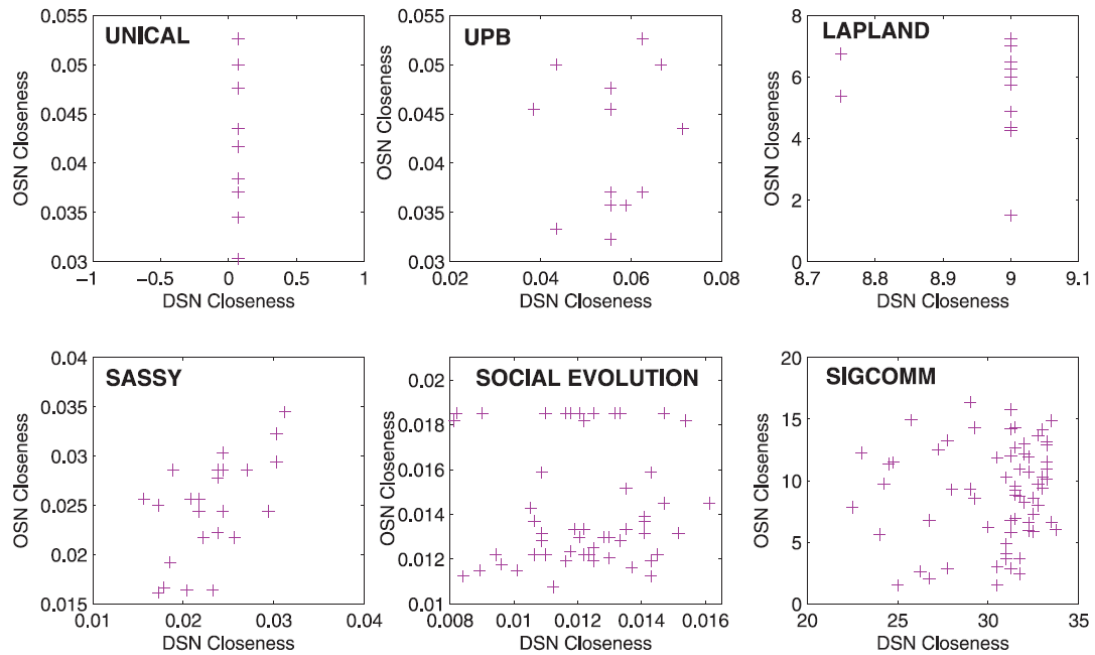
Experimental dataset	Correlation				
	Betweenness	Closeness	Eigenvector	Degree	Ego Betweenness
UPB	0.2151	0.0988	-0.015	0.1541	0.2587
LAPLAND	0.1446	-0.1454	-0.1498	-0.098	0.1455
SASSY	0.05	0.5791	0.5135	0.5251	0.6224
SOCIAL EVOLUTION	0.0492	0.0278	0.1058	0.089	0.0816
SIGCOMM	0.0533	0.1052	0.0268	0.0573	0.0012

In tab.2 vengono illustrati i valori di correlazione ottenuti per ogni dataset, mentre nelle fig.10,11,12,13,14, vengono rappresentate le relazioni tra le misure di centralità della DSN e della OSN. Si fa presente che i valori di correlazione per il dataset UNICAL poiché i valori di centralità sono pari 0 sulla DSN per le misure di betweenness ed ego betweenness fig.10 e 14, mentre le altre misure di centralità sono costanti fig.11,12,13; questa particolare situazione determina che la covarianza e la deviazione standard prodotte per le misure di centralità sulle OSN e DSN sono 0. Se si osserva la fig.3 (a) si può notare come il grafo DSN UNICAL sia un grafo completo, in cui ogni nodo può essere direttamente raggiunto da ogni altro nodo, dunque, non esistono percorsi minimi che coinvolgano nodi intermediari e questo determina che i valori di centralità siano 0.





**Figura 10** Correlazione tra OSN e DSN per la Betweenness centrality



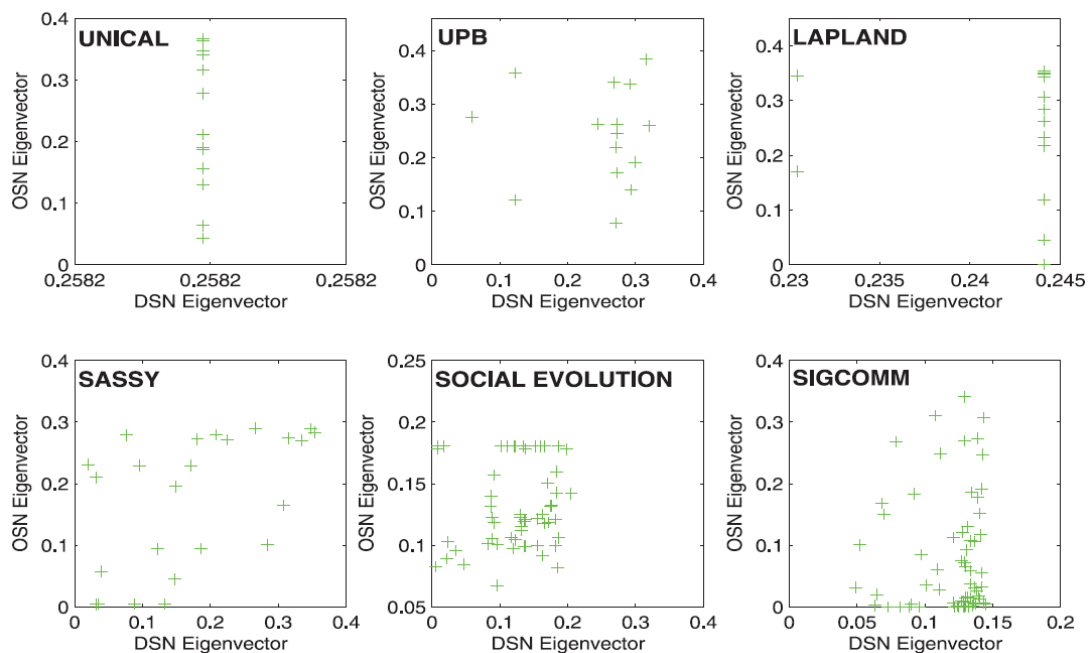
**Figura 11** Correlazione tra OSN e DSN per la Closeness centrality

Gli utenti del dataset UNICAL, infatti, sono stati spesso localizzati nella stessa classe durante le lezioni e questo ha determinato che i nodi mobili abbiano facilmente rilevato tutti gli altri nodi della rete dell'esperimento. Anche i valori costanti per la closeness, eigenvector e degree centrality sono chiaramente legati alla struttura

---

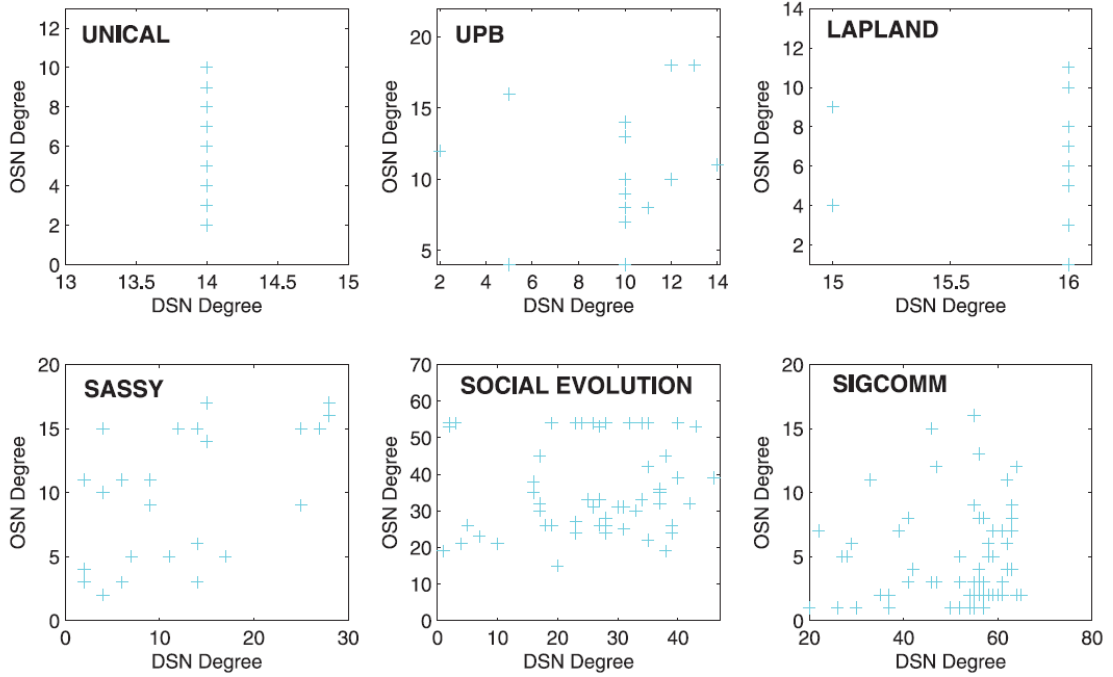
completa del dataset UNICAL. Al contrario, il grafo OSN UNICAL fig.3 (b) è più sparso poiché non tutti gli studenti coinvolti nell'esperimento erano amici Facebook; questo è giustificato dal fatto che provenissero da corsi differenti. Questa differenza ha determinato che tutti i valori di centralità per la OSN fossero non nulli. In definitiva, per il dataset UNICAL i comportamenti online e offline degli studenti sono differenti per tutte le misure considerate a causa della contemporanea presenza di tutti i candidati nella stessa aula e dal fatto che molti di loro non fossero amici Facebook.

Analizzando gli altri dataset, si nota che LAPLAND mostra comportamenti online e offline differenti in quanto i valori di correlazione per le diverse misure sono bassi. In questo caso la dimensione della rete, in cui sono presenti 17 nodi, e la struttura della DSN sono simili ad UNICAL, ma l'ambiente in cui è stata portata avanti la rilevazione è differente, in quanto UNICAL ha avuto luogo in un campus universitario mentre LAPLAND durante una conferenza. I comportamenti online ed offline risultano differenti in quanto i partecipanti sono membri di una conferenza che lavorano su aree di ricerca complementari, non sempre localizzati nello stesso posto e che non sono amici Facebook. Anche UPB, con una dimensione di rete ridotto di 15 nodi e rilevato in ambiente accademico come UNICAL, mostra deboli similarità tra le centralità online ed offline. Purtroppo, per questo dataset, non sono presenti dettagli relativi ai partecipanti, dunque, si ipotizza che i candidati possano essere studenti che seguivano corsi accademici differenti se si guarda al fatto che non tutti i nodi della DSN sono connessi e pochi di loro sono amici Facebook, apprezzabile in fig.4.

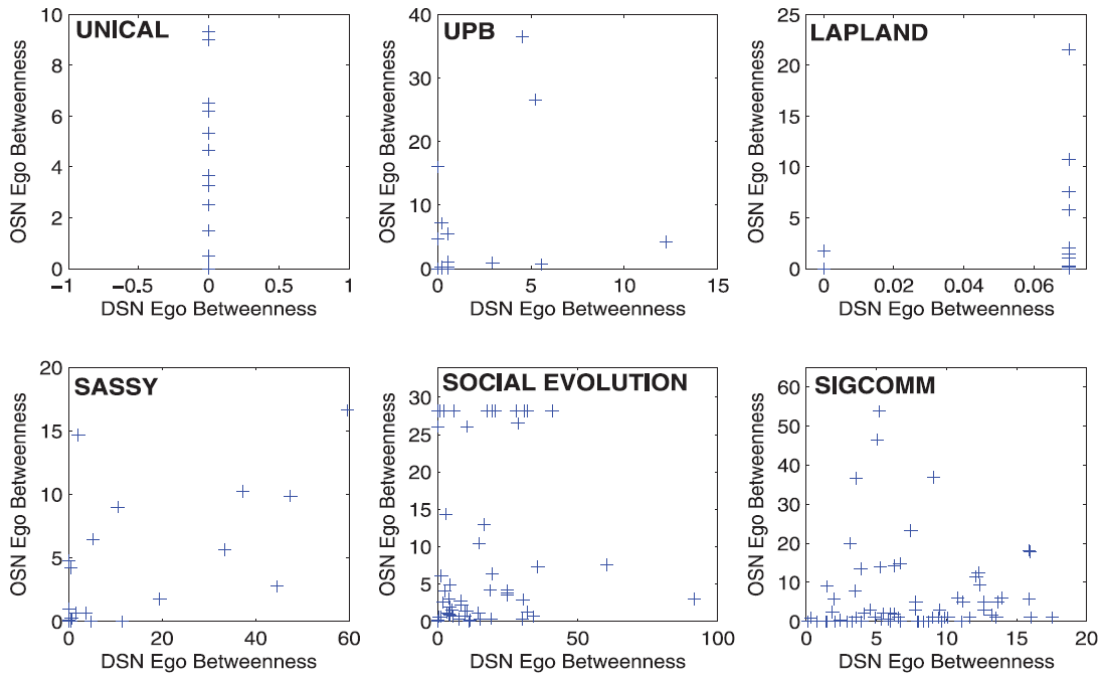


**Figura 12** Correlazione tra OSN e DSN per la Eigenvector centrality

Prendendo in considerazione SASSY, è possibile osservare che il dataset è caratterizzato da elevati valori di correlazione rispetto alle misure di closeness, eigenvector, degree centrality ed ego betweenness. In questo caso il gruppo di partecipanti mostra interessanti similarità online ed offline in grado di influenzare localmente il flusso dati. Nonostante il dataset denoti comportamenti online e offline simili per la maggior parte delle misure di centralità probabilmente in relazione al fatto che il gruppo di studenti fossero amici, se consideriamo tutti gli altri dataset, si può concludere che in generale esiste una correlazione debole tra le centralità OSN e DSN.



**Figura 13** Correlazione tra OSN e DSN per la Degree centrality



**Figura 14** Correlazione tra OSN e DSN per la Ego Betweenness centrality

I risultati di bassa correlazione ottenuti riflettono i comportamenti online e offline differenti, sia per le misure sociocentriche che egocentriche. Nel dettaglio, per i dataset SIGCOMM e Social Evolution, caratterizzati da un elevato numero di nodi

---

(rispettivamente 67 e 55), la correlazione tra ogni misura di centralità assume valori molto vicini a 0. Nel primo dataset, i partecipanti sono membri di una grande conferenza operanti per la maggior parte su topic di ricerca differenti che si trovavano in aree differenti durante l'esperimento, in quanto partecipavano a sessioni diverse ed inoltre pochi di loro erano amici Facebook, come si può apprezzare in fig.8 in cui il grafo OSN è molto sparso rispetto al grafo DSN. In Social Evolution i partecipanti sono studenti di un dormitorio e molti di loro sono amici Facebook come si può osservare sul grafo OSN in fig.7 (b). In generale i grafi OSN e DSN sono significativamente differenti in termini di centralità. Gli studenti coinvolti hanno più relazioni virtuali rispetto che incontri fisici come si può vedere in fig.7. Il risultato di questa analisi mostra chiaramente che le centralità della rete basata sulle rilevazioni Bluetooth differiscono da quelle della rete sociale Facebook. Questo accade in quanto il posizionamento all'interno del range di rilevazione di due nodi in un ambiente wireless si verifica sia per le connessioni tra nodi trasportati da individui che hanno una reale interazione, ad esempio mentre discutono o si conoscono, ma anche quando i nodi sono solo in prossimità, ad esempio due sconosciuti nella stessa stanza. Nel caso delle reti online come Facebook, un nodo, tipicamente, ha solo connessioni che ha stabilito intenzionalmente. Dunque, la struttura della DSN e della OSN sono differenti e determinano differenti valori di centralità.

Dai risultati di questa analisi, si può concludere che nel creare un algoritmo di rete opportunistico, la bassa correlazione tra i comportamenti online e offline dovrebbe essere presa in considerazione. Ad esempio, quando un algoritmo di instradamento social-based necessita di inizializzare il comportamento di un nodo nella fase di bootstrapping della rete, sono poche o assenti le informazioni sociali a disposizione, in quanto non esiste uno storico dei contatti wireless. In questi casi, l'algoritmo necessita di tempo per costruire il comportamento sociale di un nodo e sfruttarlo per migliorare la consegna dei messaggi. Dunque, in questo scenario, le informazioni sulla socialità online del nodo potrebbero essere una risorsa ed essere considerate. A seguito dei risultati ottenuti, la centralità online del nodo dovrebbe essere pesata in maniera adeguata con i dati offline sui contatti a disposizione, o in fase di

immagazzinamento, con l'obiettivo di trovare buoni percorsi di instradamento e migliorare la percentuale di messaggi consegnati.

### 3.4.2 Similarità tra gruppi

Per calcolare la similarità tra community appartenenti a due livelli di rete, è stata utilizzata la misura definita come *normalized mutual information*<sup>38</sup>.

$$NMI(A, B) = \frac{-2 \sum_{i=1}^{c_A} \sum_{j=1}^{c_B} N_{ij} \log\left(\frac{N_{ij}N}{N_i N_j}\right)}{\sum_{i=1}^{c_A} N_i \log\left(\frac{N_i}{N}\right) + \sum_{j=1}^{c_B} N_j \log\left(\frac{N_j}{N}\right)}$$

Dove  $C_A$  è il numero di community nella rete  $A$ ,  $C_B$  è il numero di community nella rete  $B$ ,  $N_{ij}$  è il numero di nodi nell'intersezione tra le community  $i$  della rete  $A$  e le community  $j$  della rete  $B$ ,  $N$  è il numero totale di nodi e  $N_i$  è il numero di nodi nella community  $i$  della rete  $A$  e la community  $j$  della rete  $B$ , rispettivamente.

$NMI(A, B)$  assume valori nel range tra 0 e 1, ed in particolare due community che abbiano tale valore uguale a 0 sono completamente differenti mentre se hanno valore pari a 1 sono identiche.

**Tabella 3 Similarità (Normalized Mutual Information) tra le community OSN e DSN.**

NMI (OSN,OSN)	Experimental datasets					
	UNICAL	UPB	LAPLAND	SASSY	SOCIAL EVOLUTION	SIGCOMM
Louvain	0.3975	0.5738	0.3192	0.2521	0.0864	0.3466
k- CLIQUE	0.5026	0.3849	0	0.1611	0	0.0103

L'indice NMI che quantifica la similarità tra livelli in termini di community per ogni metodo di community detection è riassunto in tab.3 I dataset UNICAL e UPB mostrano un significativo grado di similarità nel formare i gruppi online o offline, sia con il metodo Louvain, come è possibile apprezzare per le community DSN e OSN in cui, per entrambi i grafi, i nodi 6, 22, 11, 13 e 10 appartengono allo stesso gruppo e differiscono solo per due nodi, sia con il metodo k – CLIQUE. I dataset LAPLAND, SASSY e SIGCOMM mostrano invece un basso livello di similarità, mentre le community OSN e DSN di Social Evolution sono completamente

differenti. In generale, se ci si sofferma sull'utilizzo dei diversi metodi di rilevazione dei gruppi utilizzati, si può notare come i due metodi generino valori differenti di NMI. Questo porta a concludere che l'assunzione di gruppi distinti o sovrapposti influenza la similarità tra le community online e offline per singolo dataset. UNICAL, UPB e SASSY che sono stati rilevati tutti in ambiente accademico mostrano valori simili di NMI per i due metodi. Questo porta a concludere che i tre ambienti accademici determinano e condividono un comportamento simile anche se vengono utilizzati metodi di rilevazione dei gruppi differenti. In generale, se si considerano tutti i dataset, si può concludere che la struttura delle community online e offline è differente.

**Tabella 4 Correlazione tra OSN e DSN per i ruoli di intermediazione**

<b>Experimental dataset</b>	<b>Correlation</b>				
	<b>Coordinator</b>	<b>Gatekeeper</b>	<b>Consultant</b>	<b>Liaison</b>	<b>Total Brokerage</b>
<b>UPB</b>	0.1378	-0.1861	-0.1612	-	-0.0778
<b>LAPLAND</b>	0.0991	0.7619	-	-	0.0845
<b>SASSY</b>	0.1984	0.612	0.228	0.8036	0.6103
<b>SOCIAL EVOLUTION</b>	-0.038	0.0563	-0.0707	0.836	0.528
<b>SIGCOMM</b>	-0.0557	-0.0243	-0.0636	-0.0184	0.0199

### 3.4.3 Ruoli di intermediazione

Per valutare i ruoli di intermediazione, è stato calcolato il punteggio di brokerage per ogni ruolo e per ogni nodo nella DSN nella OSN raggruppando i nodi tramite l'algoritmo di rilevazione delle community di Louvain. In tab.4 vengono mostrati i valori di correlazione tra i ruoli di intermediazione della OSN e dalla DSN includendo anche il valore totale calcolato come frequenza totale di ogni tipologia di ruolo. Nelle fig.15,16,17,18,19 vengono inoltre raffigurate le relazioni tra i valori di intermediazione sulle DSN e OSN.

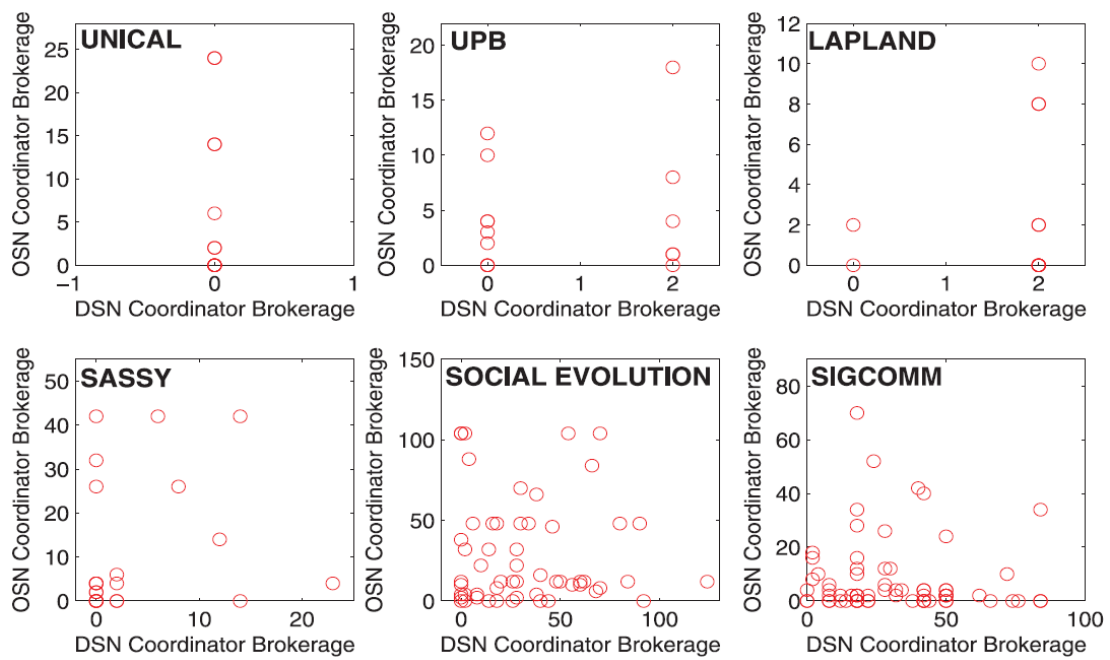


Figura 15 Correlazione del ruolo Coordinator tra OSN e DSN

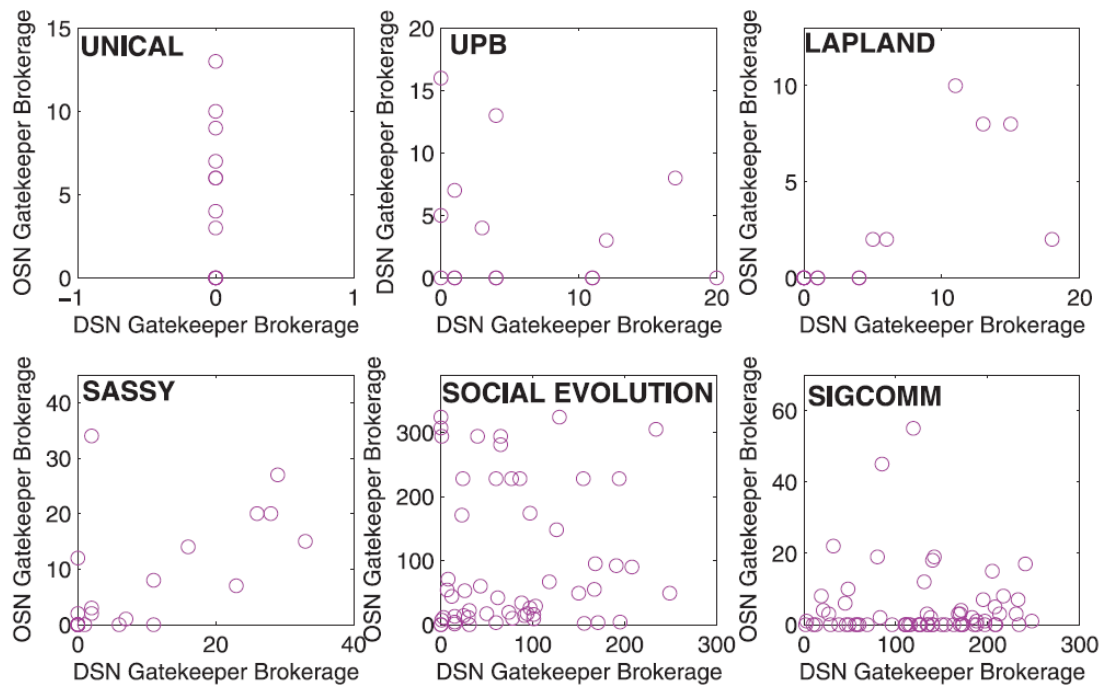


Figura 16 Correlazione del ruolo Gatekeeper tra OSN e DSN



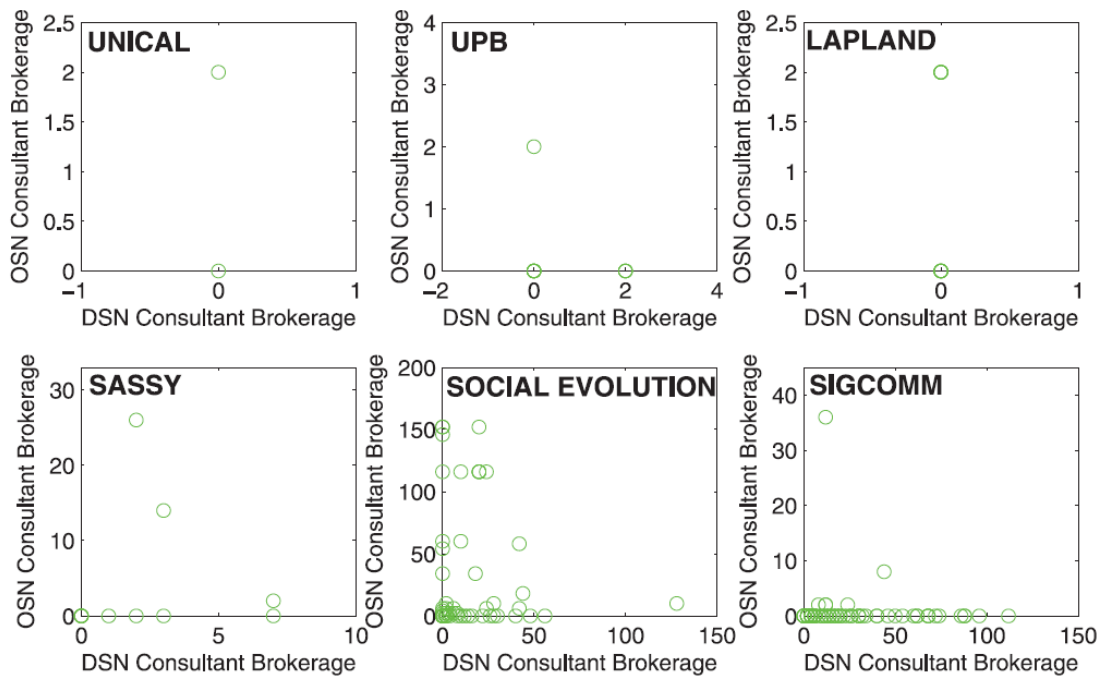


Figura 17 Correlazione del ruolo Coconsultant tra OSN e DSN

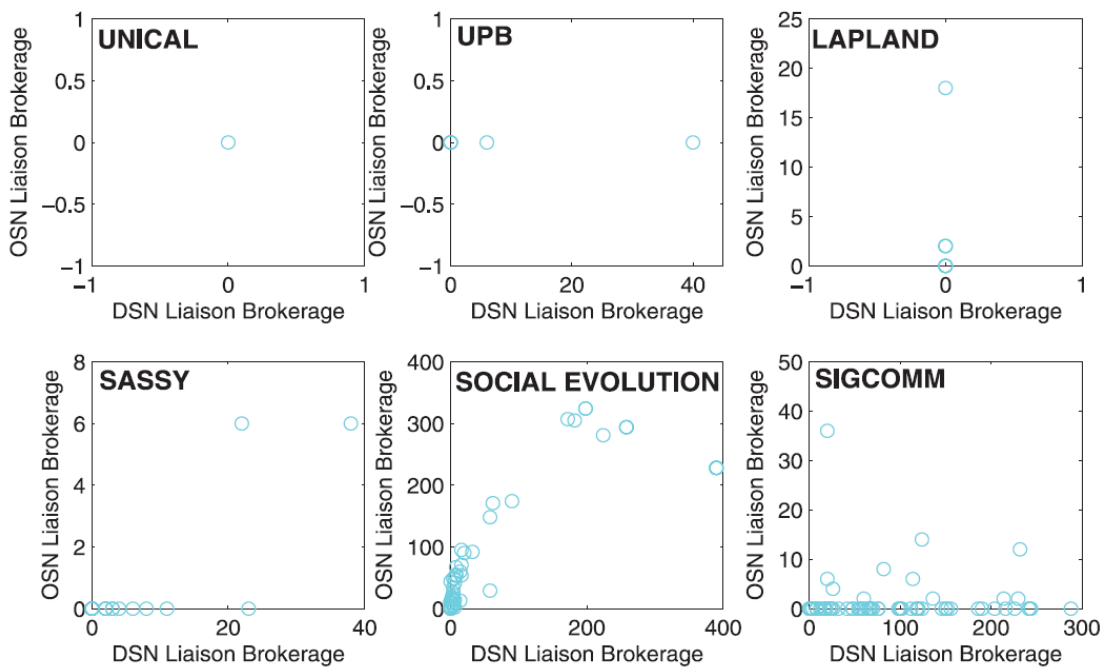
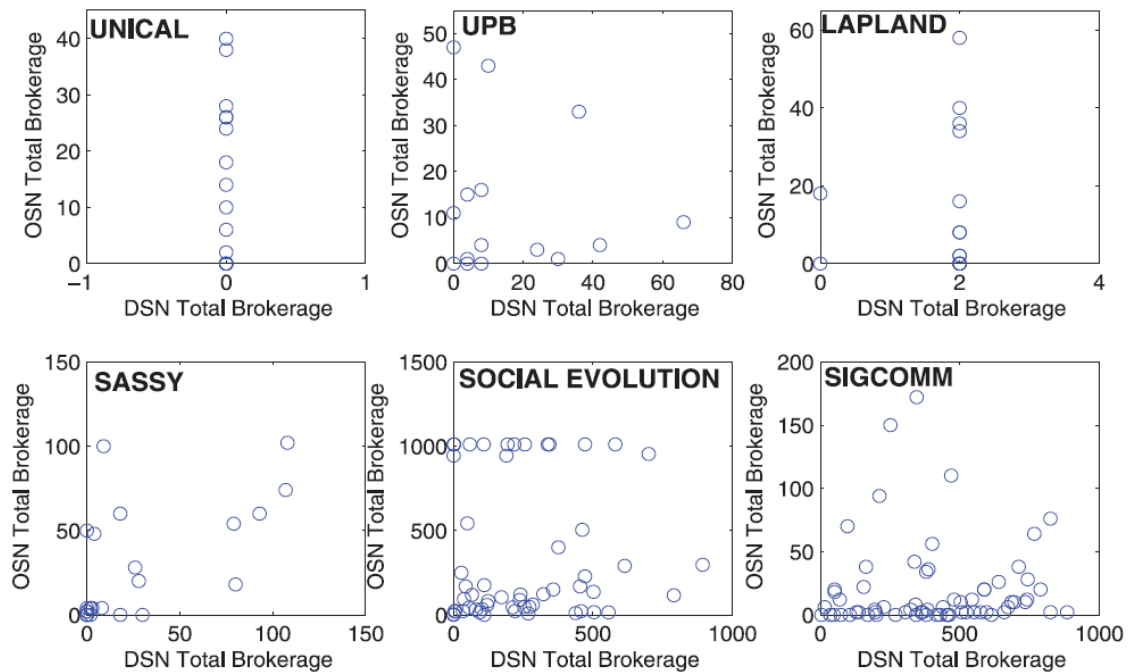


Figura 18 Correlazione del ruolo Liaison tra OSN e DSN



**Figura 19** Correlazione del total brookeage tra OSN e DSN

In generale, si denota un comportamento differente per ogni dataset. Così come per l'analisi sulla centralità, non è stato possibile riportare i valori di correlazione per il dataset UNICAL poiché i valori di intermediazione sulla DSN sono pari a 0. Questa assenza è dovuta al fatto che la DSN di UNICAL ha una struttura completa ed è stata dunque rilevata un'unica community. Anche per LAPLAND e UPB mancano alcuni ruoli, rispettivamente di liaison e consultant, e dunque non è stato possibile riportare i valori di correlazione. UPB e SIGCOMM mostrano in generale valori molto differenti per i ruoli sulla OSN e DSN, determinando dei valori di correlazione trascurabili. Al contrario SASSY mostra un valore di correlazione molto elevato per il ruolo di liaison, gatekeeper e total brookeage. Si fa presente che nei grafi DSN e OSN molti nodi che appartengono a determinate community sono connessi con coppie di nodi appartenenti a community distinte, ricadendo nel caso di liaison, o appartenenti alla stessa community, ricadendo nel caso di gatekeeper. Anche Social Evolution mostra per alcuni ruoli interessanti valori di similarità: il ruolo di liaison assume un valore di correlazione elevato e come conseguenza, il total brookeage è anch'esso caratterizzato da un valore elevato di correlazione. Infine, LAPLAND mostra valori di correlazione molto elevati per il ruolo di gatekeeper.

### 3.4.4 Legami forti

Dopo aver esaminato la centralità di rete online e offline, le community e i ruoli di intermediazione, si è passato ad analizzare i legami forti o strong ties. Nello specifico, è stata calcolata per ogni coppia di nodi all'interno dei due livelli considerati la durata di contatto, in termini di numero di intervalli di granularità, usando la granularità definita per dataset come unità di tempo. Successivamente sono stati ordinate le coppie di nodi in maniera decrescente per durata di contatto scegliendo i top 20 legami forti e analizzando la corrispondenza tra gli strong ties trovati sulla DSN e le coppie di nodi per cui esisteva un legame di amicizia Facebook sul livello OSN. Nelle fig.20,21,22,23,24,25 viene mostrato che, nella maggior parte dei dataset, la quasi totalità degli strong ties corrisponde ad amicizie Facebook. Nel dettaglio, UNICAL, UPB, SASSY, SOCIAL Evolution e SIGCOMM mostrano rispettivamente le seguenti percentuali di corrispondenza: 75%, 80%, 70%, 70% e 70%.

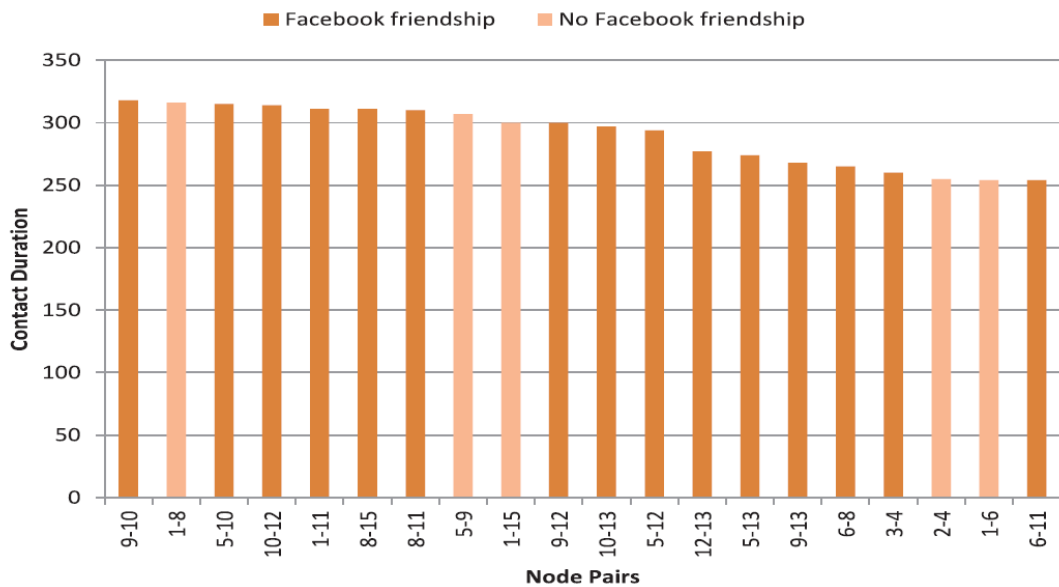


Figura 20 UNICAL – Relazioni tra strong ties con elevate durate di contatto e amicizie Facebook

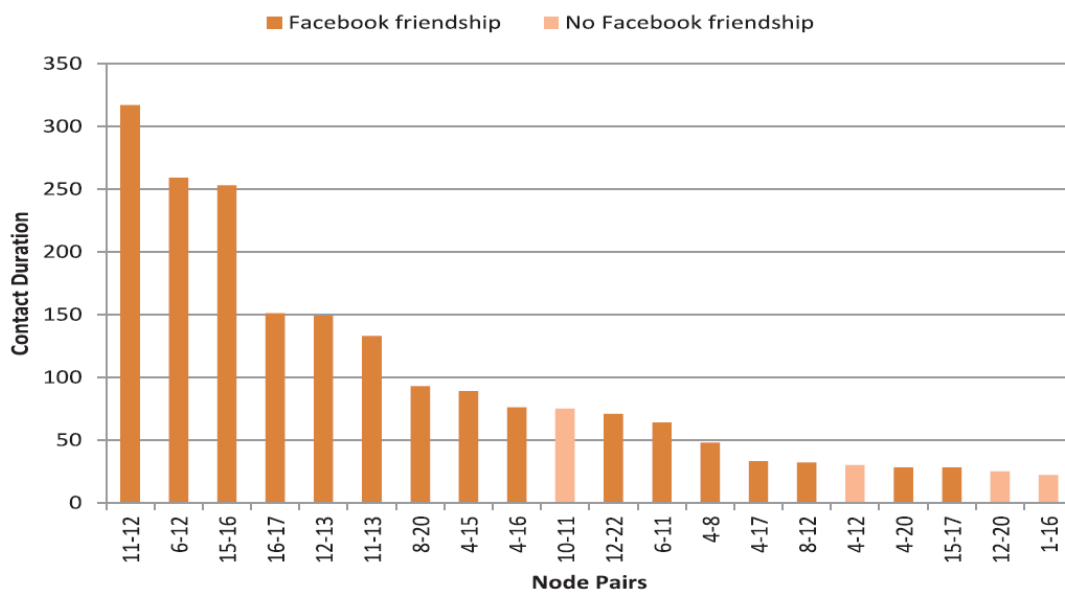


Figura 21 UPB – Relazioni tra strong ties con elevate durate di contatto e amicizie Facebook

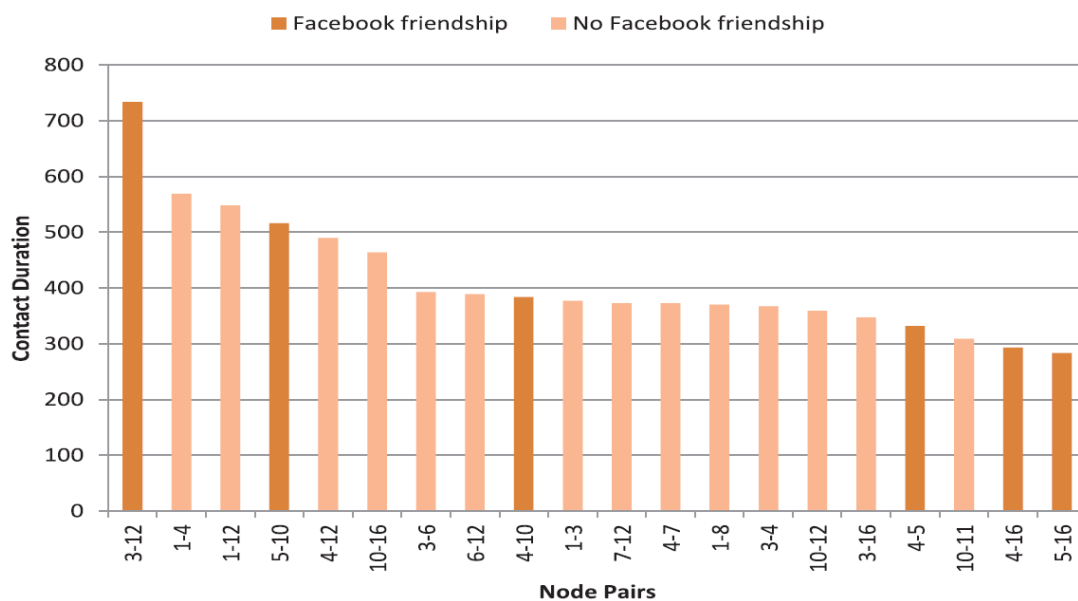


Figura 22 LAPLAND – Relazioni tra strong ties con elevate durate di contatto e amicizie Facebook

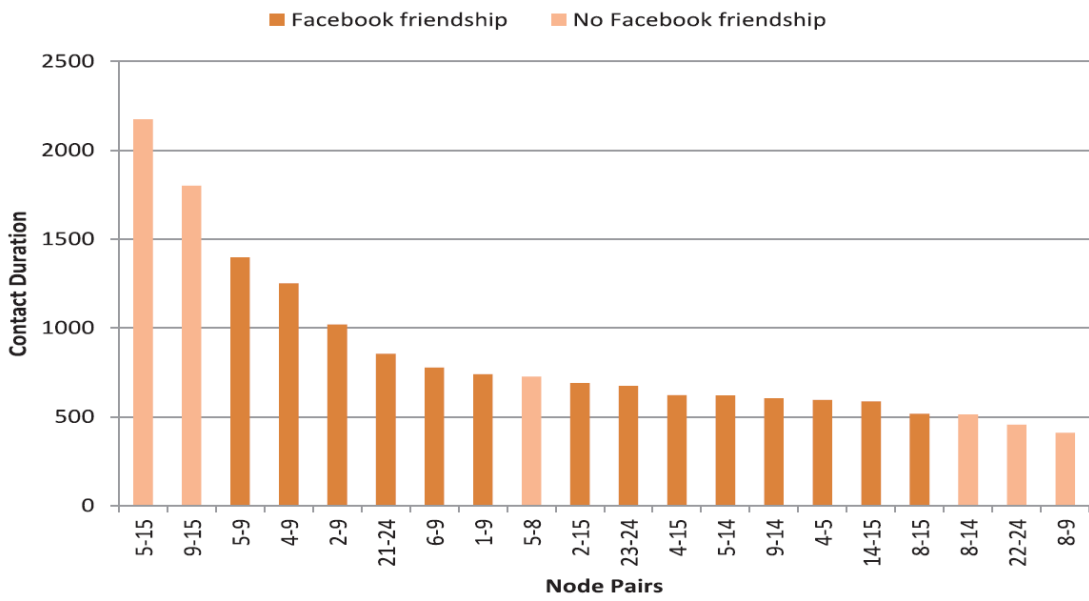


Figura 23 SASSY – Relazioni tra strong ties con elevate durate di contatto e amicizie Facebook

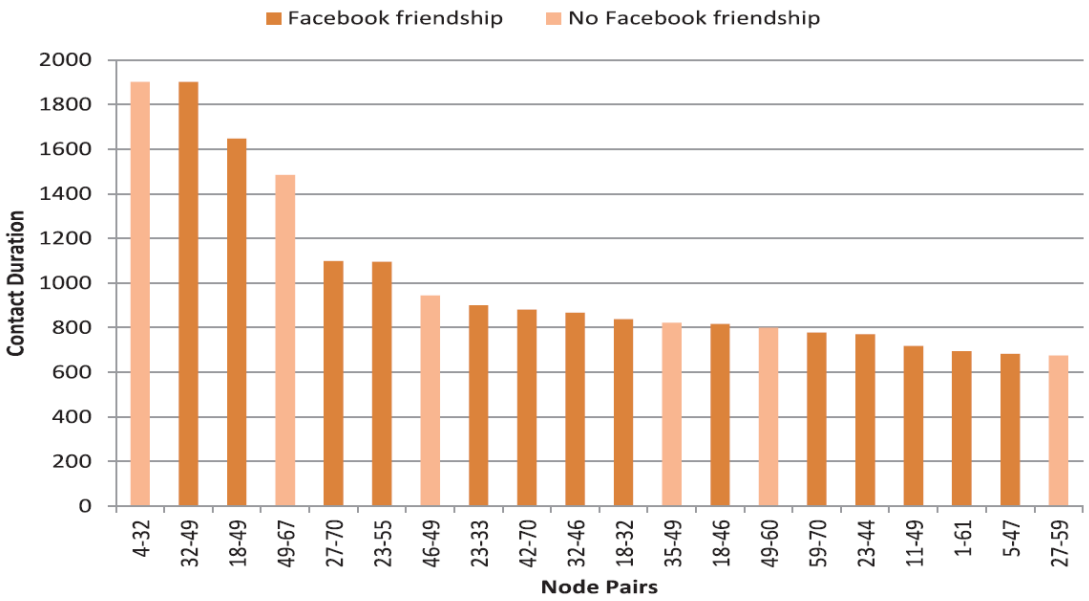
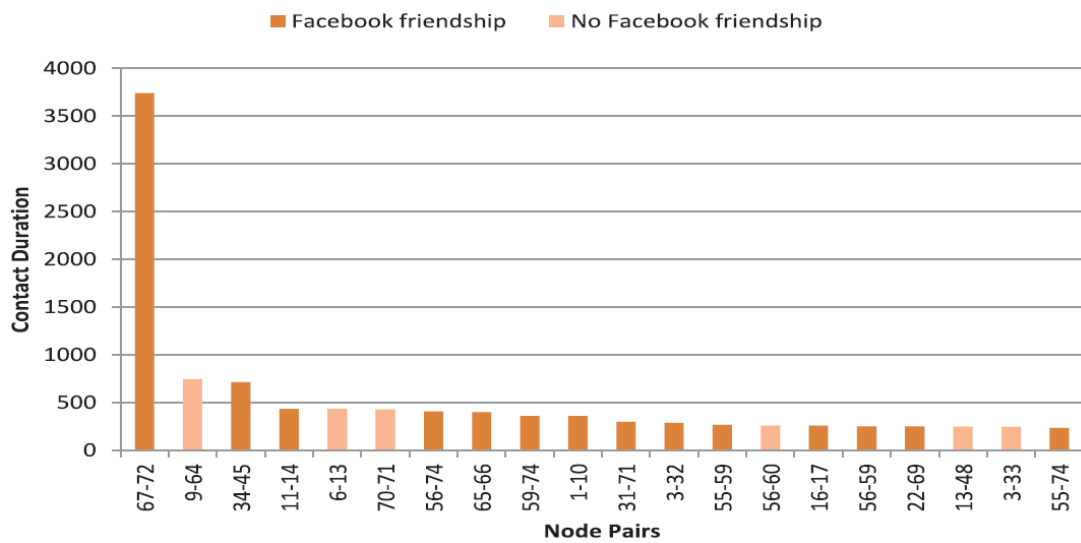


Figura 24 Social Evolution – Relazioni tra strong ties con elevate durate di contatto e amicizie Facebook



**Figura 25 SIGCOMM – Relazioni tra strong ties con elevate durate di contatto e amicizie Facebook**

A differenza delle altre metriche e algoritmi usati per valutare la similarità tra livelli di rete sociale, il comportamento online e offline dei nodi in termini di legami forti è comune per diversi dataset. Questo risultato è interessante in quanto trovando una buona corrispondenza tra la i legami forti della DSN e collegamenti sul livello della OSN, è possibile valutare se un collegamento sulla DSN è uno strong tie anche solo considerando la presenza/assenza del corrispondente link sul livello OSN e dunque evitando il calcolo delle durate di contatto.

# Capitolo IV

## Il problema della

## Selfishness nelle reti opportunistiche:

## cause ed effetti

### 4.1 Definizione

Dal punto di vista di un individuo, l'egoismo o selfishness è comunemente definita come una serie di attitudini e comportamenti rivolti unicamente a raggiungere i propri interessi. Un individuo selfish persegue i propri obiettivi al costo di danneggiare gli interessi degli altri. Ma come questo comportamento si traduce nel contesto delle reti opportunistiche? Molti ricercatori come Urpi et al.<sup>39</sup> considerano il compromesso tra il costo in termini di energia consumata e i benefici in termini di throughput di rete come un aspetto chiave per definire un eventuale comportamento selfish in una operazione di instradamento. Un nodo di una rete opportunistica può pertanto agire in maniera egoistica quando è soggetto a vincoli energetici. Un'altra valida ragione per un nodo per agire in maniera selfish è legato allo stato del buffer dei messaggi. Se un nodo ha il proprio buffer pieno, potrà agire in maniera egoistica decidendo di scartare prima i pacchetti di altri nodi rispetto che i propri. Inoltre, considerando un ambiente di rete opportunistico in cui i nodi hanno una elevata

mobilità e caratterizzato da frequenti disconnessioni, la durata di contatto tra i nodi stessi può essere assai variabile e, dunque, questo potrebbe essere un ulteriore motivo per agire in maniera egoistica. Un nodo, una volta che abbia valutato di avere molti contatti brevi, potrà decidere di sfruttare i contatti per scambiare prima i propri messaggi e solo successivamente i messaggi appartenenti ad altri nodi così da lasciare che solo i messaggi altrui vengano potenzialmente frammentati a causa dell'interruzione del contatto wireless. In fig.26 vengono illustrate le tre principali ragioni per far decidere un nodo di agire in maniera egoistica in ambiente di rete opportunistico.

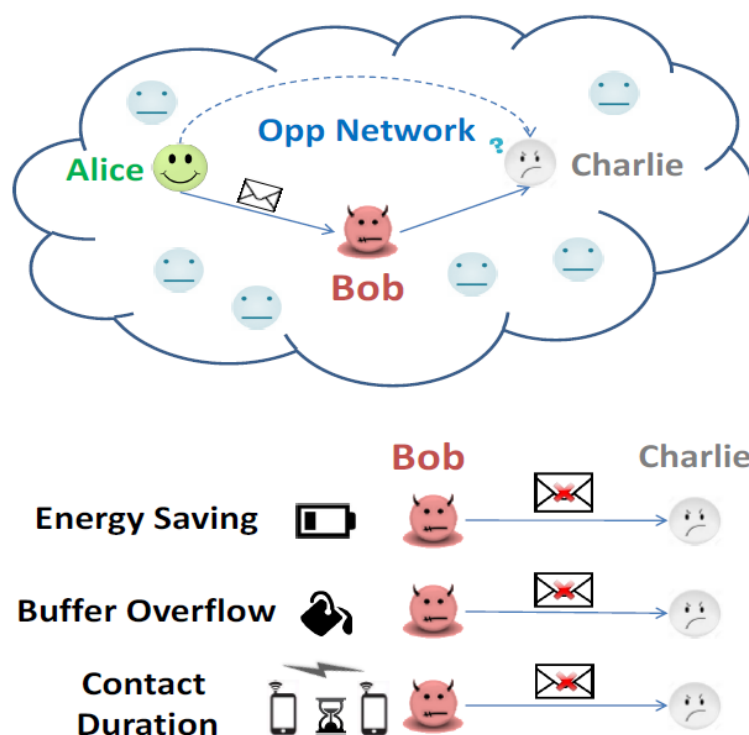


Figura 26 Alice vuole inviare un messaggio a Charlie usando la rete opportunistica (Opp Network). Alice invia il messaggio a Bob, ma Bob è selfish e non instrada il messaggio a Charlie. Le ragioni per le quali Bob può agire in maniera egoistica sono 1) preservare la sua batteria (Energy Constraints) 2) lasciare spazio nel buffer cancellando i messaggi appartenenti ad altri nodi (Buffer Overflow) 3) privilegiare i propri messaggi durante le fasi di instradamento a causa delle connessioni intermittenti (Contact Duration)

Ipotizzando che un nodo agisca in maniera selfish per una delle ragioni appena descritte, le performance di instradamento in rete opportunistica possono degradare in maniera vistosa generando seri problemi per le comunicazioni all'interno della rete stessa. Quanto affermato, sebbene possa essere intuitivamente accettato, verrà dimostrato nei paragrafi successivi in cui verranno misurate le performance di



---

differenti algoritmi di routing nel caso venissero affetti dalla selfishness dei nodi. Per modellare il comportamento egoistico, si è proceduto a scegliere in maniera randomica, tramite una distribuzione uniforme, un sottoinsieme dei nodi di rete identificandoli come selfish. Sono stati considerate le seguenti percentuali di nodi selfish: 25%, 50%, 75% e il caso limite in cui tutti i nodi lo sono. È stato ipotizzato che ognuno di questi nodi agisse in maniera egoistica per uno o più dei tre motivi enunciati. In questo studio iniziale è stato implementato un comportamento selfish dei nodi di base. Quando un nodo selfish incontra un altro nodo, essi scambiano i loro messaggi come avviene comunemente in uno scenario di rete opportunistico con un determinato schema di instradamento. Il nodo selfish scarnerà i messaggi ricevuti per i quali non è il destinatario.

## 4.2 Simulations setup

Per testare gli effetti della selfishness sull'instradamento opportunistico, sono state effettuate diverse simulazioni usando il simulatore Opportunistic Network Environment (ONE)<sup>40</sup>. Tale ambiente di simulazione permette di modellare un nodo con la propria interfaccia, memoria persistente, diversi possibilità di instradamento, un modello base di consumo di energia e differenti interazioni applicative, diversi modelli di movimento (incluse la simulazione di tracce realistiche di mobilità).

Proprio grazie a quest'ultima caratteristica, è stato possibile inserire le tracce di mobilità reale legate ad alcuni dataset già ampiamente discussi nel capitolo 2 di questa tesi: SASSY e LAPLAND. Nelle simulazioni, sono stati analizzati come benchmark tre protocolli di instradamento opportunistici: Epidemic routing<sup>41,42</sup>, Spray & Wait<sup>43</sup> e Bubble Rap<sup>44</sup>. Di seguito vengono riportate una descrizione e le ragioni per le quali sono stati scelti questi tre protocolli.

- L'Epidemic Routing è un protocollo basato sul flooding. Quando due nodi si incontrano, scambiano tutti i loro messaggi in maniera tale che questi messaggi si propagano come virus tra coppie di contatti. Questo protocollo è considerato un riferimento per l'instradamento opportunistico poiché determina un upper bound per la delivery dei messaggi.

- 
- Spray & Wait è un algoritmo di instradamento epidemico differente in quanto effettua la propagazione nella rete di un numero fissato di copie del messaggio. Il nodo sorgente invia  $L$  copie del messaggio ad  $L$  nodi distinti incontrati e aspetta sperando che uno di questi nodi trasporterà il messaggio al nodo destinazione. Se il nodo destinazione non viene trovato durante la fase di invio o spray, ognuno degli  $L$  nodi che ha una copia del messaggio, manterrà e instraderà il messaggio stesso solo se incontra il nodo destinazione. Tale algoritmo è stato dimostrato avere performance di instradamento vicine al più classico epidemic routing con un costo in termini di overhead inferiore.
  - Bubble Rap è un protocollo social – based considerato uno dei più efficienti protocolli in termini di messaggi consegnati e ritardo nel consegnarli. Esso usa due valori di centralità associati ad ogni nodo basati sulla sua popolarità globale nell'intera rete e la popolarità globale all'interno della propria community. Lo schema di instradamento usa questi valori di centralità in maniera tale che il messaggio si trasmette ai nodi con valori elevati di centralità globale finché il nodo che lo trasporta incontra un nodo della stessa community del nodo destinatario del messaggio. A questo punto, il messaggio viene instradato verso i nodi con il più alto valore di centralità locale finché non viene consegnato con successo al destinatario. Poiché le reti opportunistiche sono caratterizzate da una natura sociale a seguito del fatto che sono intrinsecamente legate alla mobilità umana, tale protocollo ha mostrato di ridurre il numero di repliche di messaggio iniettate nella rete, riuscendo a mantenere un buon valore di consegna messaggi. Per le simulazioni è stata scelta la  $C - Window\ degree$ <sup>44</sup> per implementare la centralità di un nodo in quanto può essere calcolata in maniera distribuita, fornendo una buona stima dell'idoneità di nodo ad avere il ruolo di relay. Questa è una misura di centralità cumulativa che media la degree centrality su un numero fissato di finestre temporali. Per rilevare le community, è stato scelto  $k - Clique$ . Anche se esistono in letteratura diversi algoritmi, la scelta è ricaduta su questo metodo in quanto, come già espresso nei capitoli precedenti, permette di trovare community sovrapposte, le più simili a quelle che si formano nella vita reale.

---

### 4.3 Gli effetti della selfishness

Ai fini di rilevare gli effetti della selfishness nelle reti sono stati utilizzate due metriche di performance. La prima metrica è la percentuale di consegna o delivery ratio, che è calcolata nella fattispecie come il numero di messaggi consegnati diviso il numero di pacchetti unici creati nel sistema. La seconda metrica è il ritardo di consegna o delivery delay, che considera solo i messaggi consegnati. Tale metrica misura il tempo impiegato da un messaggio per essere consegnato al nodo destinazione. Queste metriche sono state valutate in funzione del Tempo di vita o Time To Live, che rappresenta il tempo massimo che un messaggio può rimanere in un sistema dopo la sua creazione prima di essere scartato. Il TTL è fondamentale per studiare l'abilità di un protocollo di instradamento per cercare un numero adeguato di nodi relay in un tempo definito. Il valore dei principali parametri usati nelle simulazioni sono riassunti in tab.5 Ogni simulazione è stata ripetuta 20 volte ed è stata presa la media di ogni esecuzione come risultato.

Tabella 5 Valori dei parametri di simulazione

	Parametri	Valori
<b>Rete</b>	Dimensione Buffer	2000 MB
	Dimensione Messaggio	1 kB
	Intervallo di creazione di un messaggio	1800 s
<b>Spray &amp; Wait</b>	L	5
<b>Bubble Rap</b>	C-Window duration	6 hours
	C-Window # of windows	5
	k (k-Clique)	3

### 4.3.1 Risultati su SASSY

Prendendo in considerazione il dataset SASSY si è proceduto ad analizzare gli effetti del comportamento selfish dei nodi durante le operazioni di instradamento opportunistico. La delivery ratio e la latenza media per questo dataset sono evidenziate in fig.27. Analizzando la delivery ratio, si osserva che al crescere della porzione di nodi che agiscono egoisticamente nella rete, le performance della rete peggiorano per tutti i protocolli di routing analizzati. Questo significa che se il comportamento egoistico può essere in qualche modo rilevato e scoraggiato, può essere possibile raggiungere la stessa performance che si avrebbe se nessun nodo agisse egoisticamente.

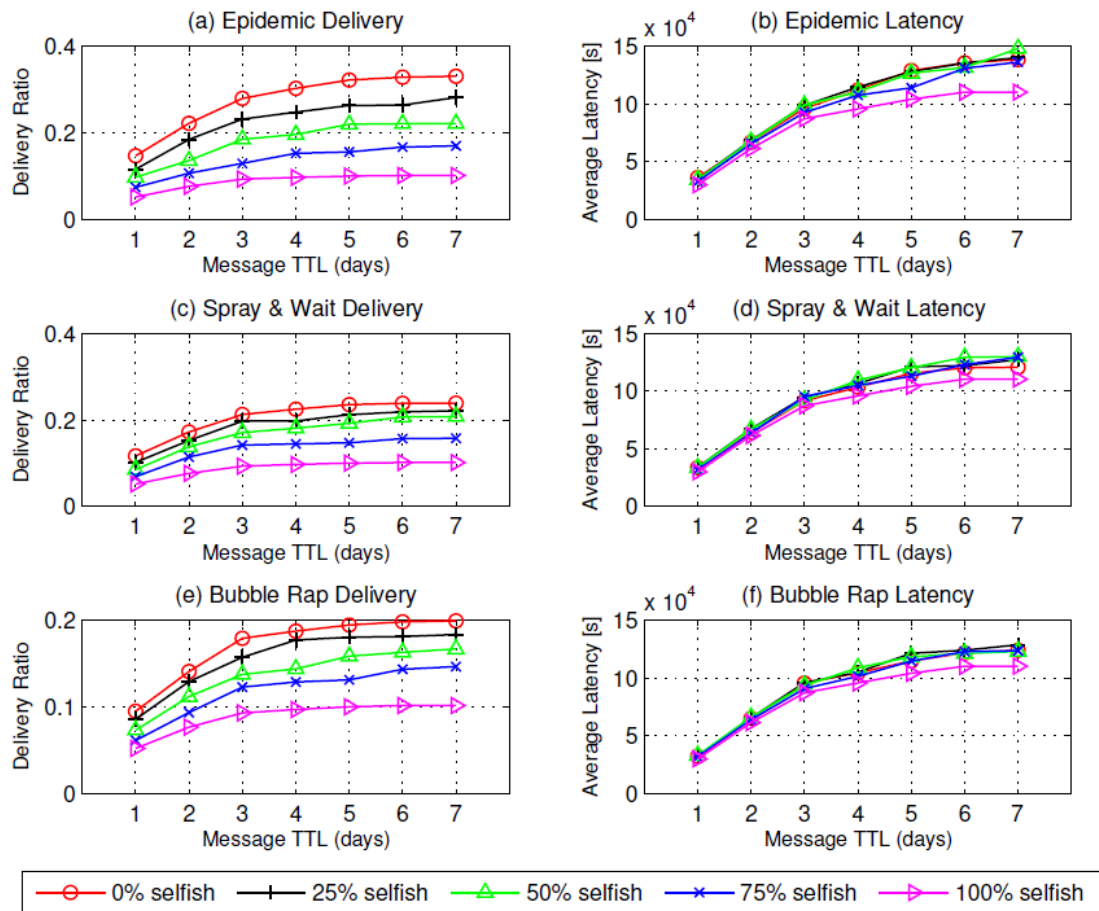


Figura 27 Performance di instradamento di SASSY per differenti % di nodi selfish

---

Il tasso di consegna dello Spray & Wait, comparato con il tasso di consegna dell'Epidemico che raggiunge il valore più alto in termini di consegna messaggio, è leggermente più basso come ci si poteva aspettare in quanto propaga un numero di copie limitato. Analizzando i risultati su Bubble Rap sono emersi interessanti risvolti: al crescere della percentuale di nodi selfish, la delivery ratio non varia significativamente. Per un TTL pari ad una settimana, ad esempio, il protocollo è caratterizzato da una differenza massima tra la delivery con totale assenza di nodi selfish e con tutti i nodi selfish di 0.1. Tale differenza è bassa se comparata con il protocollo di routing Epidemico e lo Spray & Wait. Questo protocollo social – based, se da un lato è caratterizzato da una percentuale di consegna dei messaggi più bassa a causa delle sue regole di instradamento più restrittive, d'altra parte è capace di gestire meglio la selfishness dei nodi. Ciò è stato rilevato poiché scegliendo i nodi relay che sono più sociali, l'effetto selfish può essere meglio bilanciato. Si pensa che scegliendo solo i nodi più centrali come relay, la probabilità di trovare un nodo relay che può essere anche selfish è ridotta. Nel caso di instradamento Epidemico, al contrario, dove ogni nodo incontrato è scelto come nodo relay, la probabilità di trovare un nodo selfish è più alta. Situazione analoga accade per lo Spray & Wait. In generale può essere osservato che tutti gli algoritmi consegnano più messaggi alle destinazioni al crescere del TTL. Tuttavia, quando il TTL diventa elevato, l'incremento nella percentuale di consegna dei messaggi è marginale, poiché la capacità della rete di instradare messaggi diventa il collo di bottiglia per le performance.

Per quanto riguarda il ritardo di consegna, la selfishness dei nodi causa un degrado delle performance al crescere del TTL. Per TTL maggiori di 3 giorni, la latenza media dipende molto dalla percentuale dei nodi egoisti che sono presenti all'interno della rete, mentre per TTL più bassi la selfishness non degrada particolarmente le performance di tutti i protocolli. Questo accade in quanto i TTL più lunghi determinano una più alta probabilità per i messaggi di essere consegnati ad un nodo egoista durante il percorso verso la destinazione. Tali pacchetti saranno scartati dai nodi selfish perciò incrementeranno la latenza media. Si può inoltre osservare che Bubble Rap è meno influenzato dalla percentuale di nodi selfish. La latenza media

---

varia leggermente al variare dallo 0% al 100% dei nodi selfish. Allo stesso modo che per la percentuale di consegna, le regole di instradamento social – based dimostrano di gestire meglio la selfishness. Infine, comparando i valori di ritardo raggiunti dai tre protocolli considerati, l'Epidemico ha la più alta latenza media seguito dallo Spray & Wait e da Bubble Rap, nell'ordine. Anche se il protocollo Epidemico ha tipicamente il più basso ritardo a causa dell'elevato numero copie del messaggio iniettate nella rete, in questo scenario di mobilità, il comportamento egoistico di alcuni nodi intermediari insieme alle particolari dinamiche di incontro dei nodi, che spesso assegna un pacchetto a questi intermediari malevoli, determina che i pacchetti consegnati abbiano un elevato ritardo.

### **4.3.2 Risultati su LAPLAND**

Il dataset LAPLAND copre un periodo d'esperimento più breve e ha un numero di nodi inferiore rispetto a SASSY. A seguito di tali osservazioni, il TTL è stato fatto variare da 1 a 7 ore. Come si può osservare dai risultati in fig.28, così come per SASSY, il crescere della percentuale di nodi selfish determina un abbassamento della delivery ratio per tutti i protocolli considerati. Si può dunque concludere che anche per i dataset più piccoli questa caratteristica è verificata. Bubble Rap mostra di gestire meglio la selfishness come può essere osservato in fig.28 (c). Si può aggiungere che tutti gli algoritmi consegnano più pacchetti alle destinazioni al crescere del TTL. La latenza media risulta confermare che il ritardo nella consegna incrementa al crescere del TTL e della percentuale dei nodi selfish. In particolare, per valori di TTL bassi, la latenza media è caratterizzata da un trend quasi lineare, comportamento rilevato anche su SASSY per TTL più bassi di 3 giorni. I risultati interessanti riguardano ancora una volta Bubble Rap: per questo protocollo social – based, i valori di latenza media sono simili sia se il sistema contiene nodi selfish, sia se il sistema ha tutti i nodi altruistici. Questo conferma che i protocolli social – based offrono una promettente linea di ricerca riguardo la selfishness nell'instradamento opportunistico. Comparando i risultati della latenza media ottenuti dai protocolli, si è riscontrato un trend opposto rispetto a SASSY. In questo caso, l'Epidemico ha la latenza media più bassa, seguito dallo Spray & Wait e da Bubble Rap. Questi risultati

sono più simili a quelli tipicamente trovati in letteratura poiché nella maggior parte dei casi, l'instradamento epidemico determina una latenza più bassa a costo di un maggior overhead.

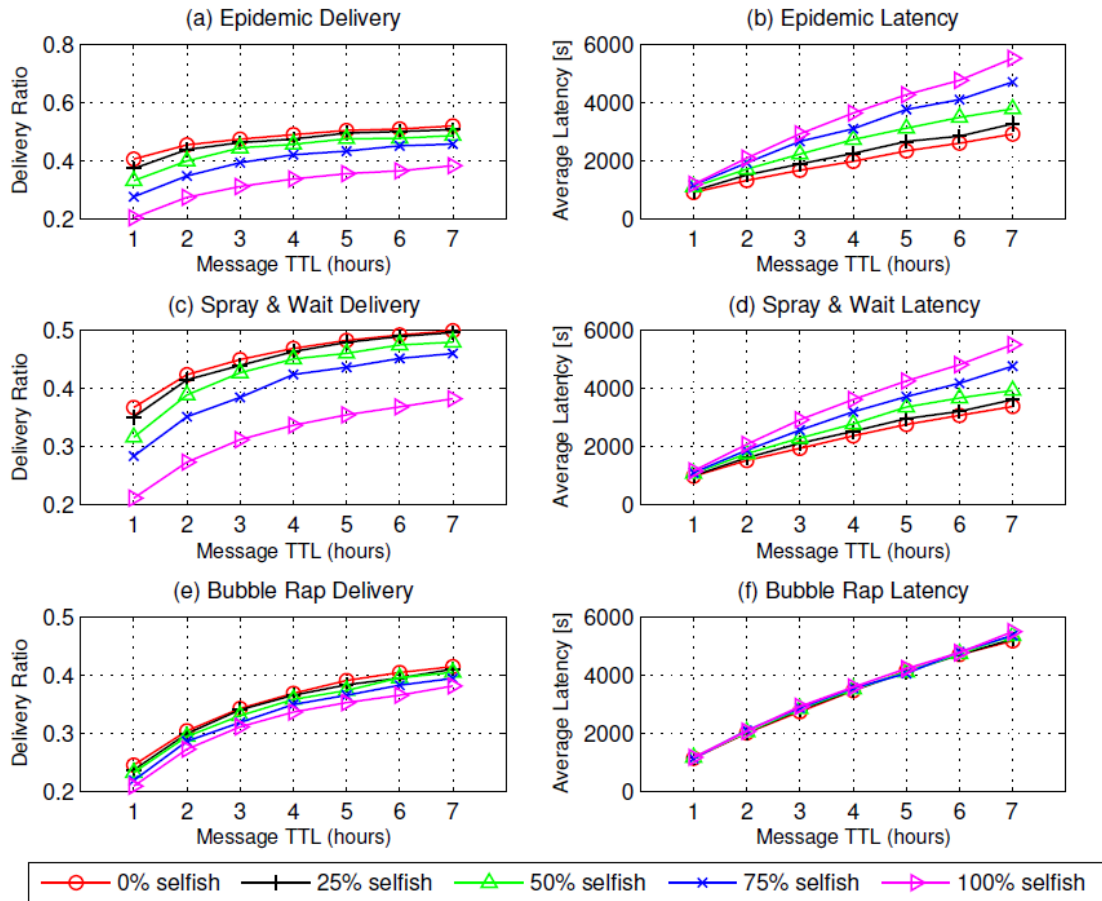


Figura 28 Performance di instradamento di LAPLAND per differenti % di nodi selfish

## Capitolo V

### SORSI: protocollo di instradamento

### Social – based Opportunistico con

### Rilevazione della Selfishness e

### meccanismi di Incentivazione

#### **5.1 Descrizione e funzionamento**

Dopo aver dimostrato tramite le simulazioni che la selfishness può fortemente degradare le performance dell'instradamento opportunistico, si passa a descrivere come questo problema possa essere mitigato attraverso un meccanismo di rilevazione dell'egoismo e l'incentivazione della collaborazione tra nodi. Lo schema proposto, definito col nome di SORSI (Social-based Opportunistic Routing with Selfishness detection and Incentive mechanisms) sfrutta la socialità dei nodi e lo storico degli incontri per rilevare la selfishness. Una volta che la selfishness è rilevata, i nodi non instradano i messaggi verso i nodi egoisti. In tal modo i nodi selfish sono incentivati a partecipare nel processo di instradamento se vogliono che i loro messaggi vengano a loro volta inviati. Le regole social – based di invio dei messaggi in ambito opportunistico, hanno mostrato di ottenere percorsi di instradamento a basso



---

costo<sup>45,46,47,48,49</sup>. Inoltre, per quanto dimostrato nel capitolo precedente, le analisi mostrano che Bubble Rap, che è un protocollo di instradamento opportunistico social – based di riferimento, è in grado di assorbire meglio gli effetti della selfishness. Come risultato di questo, si è scelto di creare lo schema di instradamento di SORSI su una logica social – based. Differentemente rispetto a Bubble Rap che considera solo la socialità offline costruita dagli incontri dei nodi, SORSI si basa su una metrica di instradamento, chiamata MLS, che sfrutta tre dimensioni sociali: la prossimità wireless, le amicizie online e gli interessi. Poiché ognuna di queste dimensioni è capace di rappresentare la socialità di un nodo, si è scelto di considerare l'insieme di queste caratteristiche sociali per ottenere una vista più approfondita del comportamento sociale di un nodo. Come dato di fatto, il modo in cui ci muoviamo, interagiamo online con i nostri amici e condividiamo i nostri interessi rappresenta la nostra socialità. La metrica MLS è perciò calcolata usando la combinazione di tre misure:

1. La centralità del nodo calcolata sulla DSN
2. I legami forti calcolati sulla OSN
3. Un link predictor calcolato sulla rete degli interessi

È stata modellata la socialità dei nodi della rete come una rete sociale multi livello in cui ogni livello è un grafo sociale che rappresenta una dimensione sociale su cui calcolare la metrica MLS. La tab.6 elenca i simboli usati per definire la metrica sociale. Si è considerata la centralità come uno dei fattori più importanti per scegliere un buon nodo intermediario per i messaggi. Nella teoria dei grafi e nella network analysis, la centralità quantifica l'importanza strutturale di un nodo all'interno di un grafo. Un nodo centrale ha tipicamente una capacità elevata di connettere altri nodi nella rete. È stata perciò calcolata la centralità di un nodo al livello DSN, in cui il corrisponde grafo sociale viene creato attraverso gli incontri tra dispositivi mobili. Il grafo sociale DSN della rete sociale multilivello è un grafo dinamico in cui un arco tra due nodi rappresenta un contatto wireless.

---

**Tabella 6 Lista di simboli usati per definire la metrica sociale di SORSI**

<b>Lista di simboli</b>	
$i$	$t^{esimo}$ nodo
$j$	$j^{esimo}$ nodo
$d$	nodo destinazione
$t$	slot temporale $t^{esimo}$
<b>T</b>	numero di slot temporali
<b>N</b>	numero di vicini del nodo $i$
<b>M</b>	numero di vicini del nodo $j$
$l$	livello di rette sociale online $l^{esimo}$
<b>L</b>	numero totale di livelli di rete sociale online
$e$	evento di incontro
<b>C<sub>Degree</sub></b>	temporal degree centrality
<b>C<sub>CDegree</sub></b>	degree centrality cumulativa
<b>TS</b>	tie strenght online
<b>TS<sub>TOT</sub></b>	tie strenght online totale
<b>LP</b>	link predictor
<b>CS</b>	Centrality utility score
<b>TSS</b>	Tie strenght utility score
<b>LPS</b>	Link predictor utility score
<b>MLS</b>	SORSI utility score
<b>ESS</b>	SORSI encounter-based selfishness score

---

Esistono diversi metodo per calcolare la centralità. La metrica sociale di SORSI calcola la centralità di un nodo  $i$ ,  $C_{CDegree}(i)$ , usando una stima cumulativa a lungo termine della degree centrality. Il vantaggio nell'usare questa misura è che può essere facilmente calcolata localmente considerando solo la rete ego di un nodo, mentre le altre misure, quali ad esempio la betweenness, la closeness o l'eigenvector, necessitano di una conoscenza globale della rete. Più nel dettaglio, SORSI calcola il numero di nodi unici visti in uno specifico slot temporale e media questa misura con un set di rilevazioni precedenti. La degree centrality per un nodo  $i$  durante uno slot temporale  $t$  è calcolata come segue:

---


$$C_{Degree}(i, t) = \sum_{j=1}^N e(i, j, t)$$

Dove

$$e(i, j, t) = \begin{cases} 1 & \text{se } i \text{ incontra } j \text{ durante il time slot } t \\ 0 & \text{altrimenti} \end{cases}$$

rappresenta un arco tra il nodo  $i$  e il nodo  $j$  sul grafo DSN corrispondente all'intervallo temporale considerato<sup>1</sup>, e  $N$  è il numero di nodi nel range di  $i$ . La degree cumulativa,  $C_{Degree}(i)$ , è calcolata mediando i valori di degree del nodo su un set di intervalli temporali  $T$  che includono il più recente intervallo temporale e tutti i precedenti:

$$C_{Degree}(i, t) = \frac{1}{T} \sum_{t=0}^T C_{Degree}(i, T - t)$$

In questo modo, SORSI fornisce un'approssimazione completamente decentralizzata per la degree centrality di un nodo, facile da calcolare. La centralità descritta in precedenza è misurata usando la storia dei contatti e non considerando la futura disponibilità dei link. Considerando che i link nella rete sono tempo variabili, un collegamento esistente verso un nodo centrale potrebbe non essere altamente disponibile. Dunque si è incluso un indicatore della tie strenght all'interno della metrica sociale di SORSI. Tale indicatore è capace di identificare i collegamenti che hanno un'elevata probabilità di essere attivati ed è misurato considerando i collegamenti online tra gli individui che trasportavano i dispositivi mobili. Questa scelta è mossa dalla considerazione che i legami sociali sui siti di reti sociali, come Facebook, Twitter o LinedIn, sono più stabili e dunque più forti rispetto ai legami sulla rete dei contatti fisici. Solitamente, un legame online tra due utenti non cambia nel tempo. In Facebook, ad esempio, le amicizie per così dire intermittenti sono molto improbabili. Inoltre, è stato dimostrato<sup>50,51</sup> che i nodi incontrano altri nodi socialmente connessi online con un'elevata probabilità. In conseguenza di ciò, i

---

<sup>1</sup> Considerando che il grafo DSN è un grafo temporale, si è deciso di formare un grafo statico per ogni slot temporale unendo tutti i contatti in quell'intervallo di tempo.

---

legami online possono essere considerati una buona misura di quando un link sulla DSN sarà attivato.

SORSI calcola la tie strenght tra un nodo  $i$  e un nodo  $j$  sul livello OSN  $l$  come:

$$TS(i, j, l) = \begin{cases} 1 & \text{se } i \text{ incontra } j \text{ al livello } l \\ 0 & \text{altrimenti} \end{cases}$$

La tie strenght totale tra due nodi è la somma degli indicatori misurati ad ogni livello OSN:

$$TS_{TOT}(i, j) = \sum_{l=1}^L TS(i, j, l)$$

Dove  $L$  è il numero totale di siti di reti sociali online considerati.

La metrica sociale di SORSI prende in considerazione una terza misura utile a predire future collaborazioni tra due nodi. Un link predictor è calcolato sul livello della rete degli interessi, dove un link tra due nodi esiste se essi hanno almeno un interesse in comune. Esaminando i vicini comuni di una coppia di nodi sul livello di rete degli interessi, è possibile predire futuri incontri. Diversi lavori sui co-autori o reti di collaborazione, dimostrano che la probabilità che due nodi diventino connessi tramite un link è più alta quando i nodi in questione hanno vicini in comune. Newman<sup>52</sup>, ad esempio, analizza una rete di collaborazioni scientifiche, mostrando che esaminare i co-autori di autori aiuta nel predire future collaborazioni. Poiché le reti di co-autori sono reti di collaborazioni scientifiche e dunque reti di interessi scientifici condivisi, si è scelto di utilizzare il livello di rete degli interessi come link prediction, assumendo che le reti di interesse scientifico e più in generale le reti degli interessi abbiano un comportamento simile. SORSI calcola il link predictor  $LP(i, j)$  come una possibile collaborazione tra i nodi  $i$  e  $j$  come una misura comune di vicinanza basata su coefficiente di Jaccard:

$$LP(i, j) = \frac{|N \cap M|}{|N \cup M|}$$

---

dove  $N$  rappresenta i nodi nel range di  $i$  e  $M$  rappresenta i nodi nel range di  $j$ .

Per ogni misura, SORSI determina l'utility score di un nodo  $i$  per consegnare un messaggio ad un nodo  $d$  rispetto ad un nodo  $j$  come segue:

$$CS(i, j) = \frac{C_{CDegree}(i)}{C_{CDegree}(i) + C_{CDegree}(j)}$$

$$TSS(i, j, d) = \frac{TS_{TOT}(i, d)}{TS_{TOT}(i, d) + TS_{TOT}(j, d)}$$

$$LPS(i, j, d) = \frac{LP(i, d)}{LP(i, d) + LP(j, d)}$$

La metrica sociale di SORSI è data dalla combinazione di questi contributi come segue:

$$MLS(i, j, d) = CS(i, j)[1 + TSS(i, j, d) + LPS(i, j, d)]$$

Come si può osservare,  $MLS$  cattura l'importanza di un nodo in qualità di intermediario comparata a quella di un nodo incontrato su tutti i livelli di rete sociale, in termini di centralità, tie strength e link predictor. Si noti anche che la centralità di un nodo è considerata come fattore predominante nell'instradamento di un messaggio. Sia i punteggi di tie strength che la tie predictor utility sono pesati con il punteggio di centrality utility e quindi aggiunti al punteggio di centrality utility. In questo modo, la tie strength e la link predictor utility rifletteranno il punteggio di centrality utility tra un nodo mittente e il nodo incontrato.

Il processo di instradamento in SORSI è dato dall'Algoritmo 1. I due nodi che si incontrano scambiano anche la lista dello storico di instradamento contenente l'attività di instradamento di ogni nodo incontrato. Ogni instradamento passato incluso nella lista di un nodo è loggato attraverso l'ID del nodo incontrato che instrada il messaggio a quel nodo, l'ID del messaggio, l'ID della sorgente del messaggio e l'informazione su quando il messaggio è stato ricevuto. Uno storico è perciò implementato come una lista contenente  $n$  delle seguenti righe [forwarderID,

---

msgID, msgSourceID, timestamp]. Dopo aver calcolato la metrica MLS per decidere quando instradare i messaggi ad un nodo incontrato N secondo la sua socialità e selfishness, il nodo verifica lo storico di instradamento del nodo incontrato per rilevare eventuali nodi selfish. Si fa presente che i messaggi nel buffer sono messaggi ricevuti che hanno passato con successo la verifica di selfishness. Se il nodo rileva un altro come selfish (ad esempio lo storico di N non contiene messaggi del nodo rilevato che sarebbero dovuti passare attraverso il nodo selfish), il nodo che effettua la rilevazione incrementa il punteggio di selfishness ESS per il nodo in questione.

---

**Algoritmo 1** SORSI instradamento di un messaggio con rilevazione della selfishness

---

```

1:   procedure encounterNode(N)
2:       exchangeCentralityValues()
3:       exchangeOnlineContactsLists()
4:       exchangeInterestNodeList()
5:       exchangeForwardingHistoryList()
6:       for every message m in message buffer do
7:           D ← m.destination()
8:           encounterMLS ← computePeerMLScore()
9:           myMLS ← computeMLScore()
10:          if encounterMLS ≥ myMLS || N==D then
11:              forwardMessage(m,N)
12:          end if
13:      end for
14:      for every message m in forwarding history list do
15:          if timestamp > last encounter with N && msgSourceID==myID then
16:              if last encounter with forwarderID > last encounter with N &&
forwarderID.receivedMessage(msgID)==false then
17:                  ESS ← forwarderID.selfishScore()
18:                  forwarderID.setSelfishScore(ESS+1)
19:              end if
20:          end if
21:      end for
22:  end procedure

```

---

---

Similmente, quando un nodo riceve e potrebbe potenzialmente instradare un messaggio per il quale esso stesso non è la sorgente (si veda l'Algoritmo 2), il suo altruismo è premiato decrementando il punteggio di selfishness. In questo modo si realizza un meccanismo di incentivazione per l'instradamento dei messaggi. I nodi non ricevono i messaggi di nodi che hanno un punteggio di selfish maggiore di una soglia. In conseguenza di ciò. In accordo con lo schema di SORSI, i nodi che sono stati etichettati come selfish saranno obbligati a instradare messaggi di altri nodi per poter instradare poi i propri.

---

**Algoritmo 2** SORSI ricezione di un messaggio con rilevazione della selfishness

---

```

1:   procedure RECEIVEMESSAGE( $m, N$ )
2:       if  $N \neq m.source()$  then
3:            $ESS \leftarrow N.selfishScore()$ 
4:            $N.setSelfishScore(ESS - 1)$ 
5:           receive( $m$ )
6:       else
7:           if  $N.selfishScore() < thr$  then
8:               receive( $m$ )
9:           else
10:              discard( $m$ )
11:          end if
12:      end if
13:  end procedure

```

---

## 5.2 Valutazione delle performance

### 5.2.1 Simulation setup

Si è proceduto a validare la proposta nuovamente attraverso il simulatore ONE usando le tracce di mobilità umana SASSY e LAPLAND. Nelle simulazioni SORSI è stato comparato alla sua versione priva della rilevazione della selfishness, denominata SORSI-NS, nell'ottica di quantificare i miglioramenti apportati dalla rilevazione della selfishness e dal meccanismo di incentivazione alla partecipazione,

---

oltreché all'instradamento Epidemico, a Bubble Rap e a SPRINT – SELF<sup>53</sup>. Per questa analisi, non è stato preso nuovamente in considerazione Spray & Wait poiché si è prestata maggiore attenzione sugli schemi social – based considerando che, attraverso SORSI, si vuole dimostrare i miglioramenti portati all'instradamento sociale mobile. Tuttavia è stato ancora utilizzato il routing Epidemico per avere un upper bound relativo alla consegna dei messaggi. Come specificato in precedenza, il protocollo SPRINT – SELF è di tipo social – based e utilizza un meccanismo di rilevazione della selfishness che scoraggia il comportamento egoistico dei nodi. Similmente a SORSI, esso utilizza le informazioni della social network, la predizione del comportamento dei nodi e un meccanismo di rilevazione della selfishness. Queste similitudini hanno determinato la scelta del protocollo come paragone per SORSI. Tuttavia, diversamente rispetto a SORSI, SPRINT – SELF non considera una struttura di rete sociale multilivello per modellare la rete opportunistica e rendere utili le conoscenze sulle community. Proprio le community, poiché è noto che la maggior parte dei nodi sono più portati ad interagire con i membri del proprio gruppo, sono parti fondamentali del protocollo. Entrambi i protocolli sono calcolati on-the-fly attraverso algoritmi di rilevazione delle community come k – Clique o attraverso i gruppi estratti direttamente dalle online social network come Facebook. Tuttavia, in uno studio recente<sup>39</sup>, si è dimostrato che non sempre le community online corrispondono a quelle offline. A partire da questo, le informazioni sui gruppi online non devono essere sempre usate e possono portare all'individuazione di percorsi di instradamento sub-ottimi. Inoltre, a causa del calcolo delle community attraverso gli algoritmi di rilevazione dei gruppi e a causa di un valore di utilità che prende in considerazione diverse informazioni sui contatti e sullo storico nello scambio dei messaggi, il costo computazionale di questo valore di utilità è sensibilmente più alto rispetto a quello in SORSI.

Per SPRINT – SELF, quando un nodo è etichettato come selfish si presuppone che il livello di carica della sua batteria sia maggiore di una soglia di tolleranza; ad esempio un nodo il cui livello di batteria è del 2% non è considerato selfish se non instrada i messaggi, mentre se il livello di batteria si attesta sul 30% lo è. Per le analisi, sono state utilizzate tre metriche di performance opportunistiche.



- La percentuale di pacchetti consegnati, anche definita come throughput di sistema.
- L'overhead, che misura il numero di pacchetti trasmessi diviso il numero di pacchetti unici creati. Tale indice è importante per lo studio in quanto i messaggi duplicati immessi nella rete consumano risorse, come batteria e memoria.
- Il ritardo di sistema o il ritardo di consegna

Come primo esperimento, attraverso una distribuzione uniforme, sono stati scelti casualmente un sottoinsieme di nodi di rete etichettandoli come selfish. Sono stati scelti di nuovo il 25%, 50%, 75% di nodi selfish e il caso limite quando tutti i nodi lo sono, ipotizzando che la selfishness sia guidata da una o più delle motivazioni descritte in precedenza. Per il routing Epidemico, Bubble Rap e SORSI – NS, un nodo selfish scarta sempre i messaggi per i quali non è il destinatario. Ogni simulazione è stata ripetuta 20 volte ed è stata anche in questo caso presa la media di ogni lancio come risultato. I valori dei principali parametri utilizzati nelle simulazioni sono mostrati in tab.7. Simulando differenti soglie di selfishness (fig.29), si è scelta una soglia uguale a 50 per la quale si sono ottenuti i migliori risultati per le performance di consegna in SORSI in entrambi i dataset. Come ci si aspettava, al crescere della soglia divenendo meno restrittiva con i nodi selfish, la percentuale di consegna inizia a diminuire. Come secondo esperimento, si è testato e comparato l'effetto del consumo di energia e l'occupazione del buffer. Per ogni nodo, è stato considerato un modello per il consumo di energia noto in letteratura e adottato in altri lavori<sup>54</sup>.

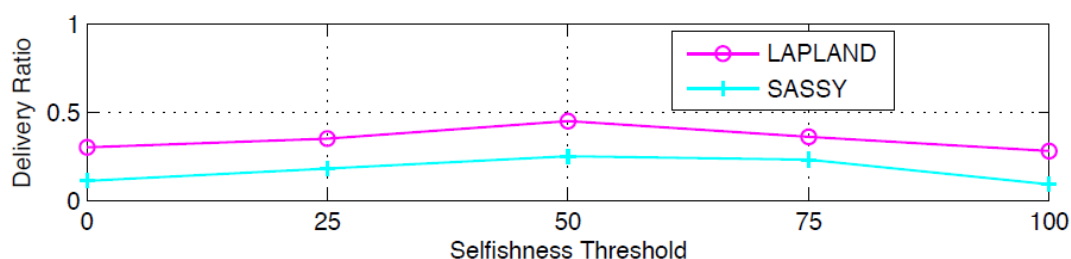


Figura 29 Effetto dell'applicazione della soglia di selfishness sulla consegna messaggi.

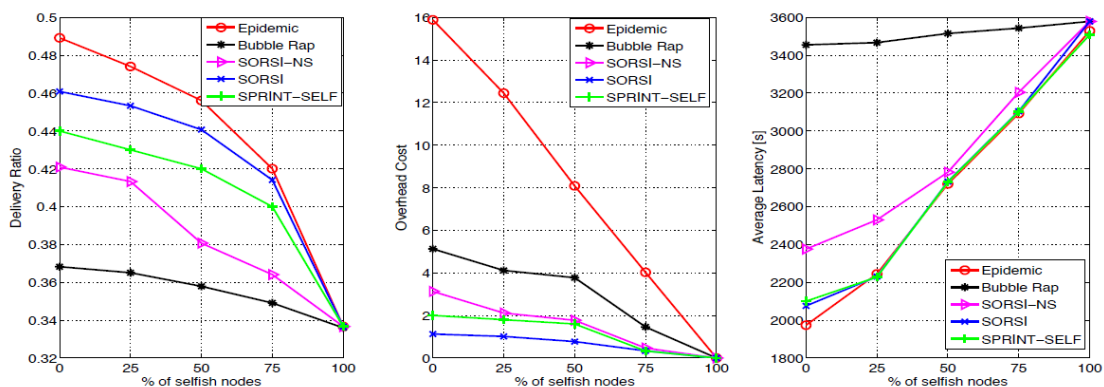
Tabella 7 Valorizzazione parametri di simulazione

	<b>Parametri</b>	<b>Valori</b>
<b>Rete</b>	Dimensione buffer	2000 MB
	Dimensione messaggio	1 kB
	Intervallo di creazione messaggio	1800 s
	TTL SASSY	4 giorni
	TTL LAPLAND	240 s
<b>SORSI-NS</b>	Time slot	6 ore
	T	5
<b>SORSI</b>	Soglia selfishness thr	50
	Punteggio selfish iniziale ESS (nodo altruista)	0
	Punteggio selfish iniziale ESS (nodo egoista)	100
<b>Bubble Rap</b>	T	5
	Durata C-Window	6 ore
	# di C-Window	5
	K (k-Clique)	3
<b>SPRINT-SELF</b>	$W_1$	0.7
	$W_2$	0.3
	K (k-Clique)	3
	Dimensione cache	100 MB

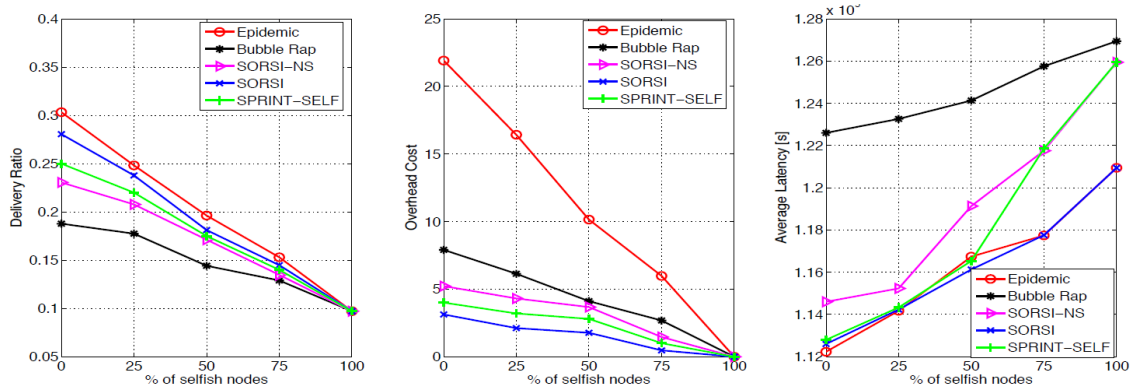
### 5.2.2 Risultati

Per valutare i risultati si è partiti analizzando la traccia di mobilità SASSY per il primo esperimento, in cui è stato scelto in maniera random se un nodo fosse selfish o meno. La percentuale di messaggi consegnati, l'overhead e la latenza media per questa traccia sono rappresentati in fig.30. Come ci si aspettava, quando aumenta la proporzione di nodi che agiscono egoisticamente, il tasso di consegna decresce per tutti i protocolli di instradamento considerati. Allo stesso modo, sempre al crescere della selfishness, degrada la latenza media, mentre l'overhead decresce poiché il comportamento egoistico dei nodi determina un minor numero di pacchetti instradati. Il protocollo di routing Epidemico è maggiormente influenzato dalla percentuale di nodi selfish, mostrando un più rapido decadimento delle sue performance al crescere

della percentuale dei nodi egoisti presenti nella rete. Si può notare inoltre che quando tutti i nodi sono selfish, i protocolli consegnano lo stesso numero di messaggi, ottenendo gli stessi costi e ritardi. Questo accade in quanto ogni nodo instrada solo i propri messaggi. Quando è presente una determinata percentuale di nodi selfish (25%, 50%, 75%), al contrario, il routing Epidemico ottiene i più alti valori in termini di percentuale di messaggi consegnati, seguito da SORSI che è capace di ottenere migliori risultati rispetto agli altri schemi social – based presi in considerazione. Inoltre, SORSI è caratterizzato da latenza media e tasso di consegna comparabili al routing Epidemico con un costo di overhead decisamente inferiore. Tale risultato mostra che SORSI è in grado di gestire la selfishness dei nodi e allo stesso tempo limitare sensibilmente il numero di repliche di messaggio e quindi anche il consumo di energia, evitando il propagarsi di comportamenti selfish. Inoltre, prendendo in considerazione SORSI con il 50% di nodi selfish, ad esempio, il protocollo è in grado di raggiungere le stesse performance in termini di percentuale di messaggi consegnati di Bubble Rap senza nodi selfish. Si sottolinea che la sua metrica sociale multi – livello insieme con la gestione della selfishness è capace di raggiungere performance di instradamento paragonabili ad un protocollo social – based di riferimento senza nodi selfish. Comparato con la sua versione senza rilevazione della selfishness, si può osservare come SORSI è capace di ridurre gli effetti della selfishness comportandosi similmente a SORSI-NS con un significativo numero di nodi egoisti inferiore.



**Figura 30 Performance di instradamento di SORSI comparate al routing Epidemico, a Bubble Rap, a SORSI-NS e a SPRINT-SELF per la traccia di mobilità SASSY**



**Figura 31 Performance di instradamento di SORSI comparate al routing Epidemico, a Bubble Rap, a SORSI-NS e a SPRINT-SELF per la traccia di mobilità LAPLAND**

In fig.31 sono evidenziati i risultati ottenuti per la traccia di mobilità LAPLAND. Tale traccia ha una durata d’esperimento più breve e ha un numero di nodi più piccolo rispetto a SASSY. Per questo motivo, si è scelto un TTL più breve pari a 4 minuti. Così come per SASSY, l’incrementare della porzione di nodi egoisti nella rete determina una più bassa percentuale di messaggi consegnati per tutti i protocolli considerati. Ancora una volta, SORSI supera Bubble Rap, SORSI-NS e SPRINT-SELF raggiungendo un tasso di consegna simile al routing Epidemico. A seguito di questi risultati si è concluso che anche nei dataset più piccoli tale caratteristica è verificata. Il costo in termini di overhead conferma che SORSI è capace di ridurre il numero di repliche di messaggio raggiungendo costi più bassi senza compromettere la sua efficienza in termini di ritardo end – to – end.

**Tabella 8 Performance dei protocolli su traccia di mobilità SASSY quando i nodi hanno il 50% di energia residua.**

	Epidemico	Bubble Rap	SORSI-NS	SORSI	SPRINT-SELF
<b>Delivery Ratio</b>	0.412	0.317	0.331	0.397	0.375
<b>Overhead</b>	7.121	3.136	1.951	1.875	1.056
<b>Latenza Media</b>	2406.54	3229.35	2451.779	2394.832	2398.569

**Tabella 9 Performance dei protocolli su traccia di mobilità LAPLAND quando i nodi hanno il 50% di energia residua.**

	Epidemico	Bubble Rap	SORSI-NS	SORSI	SPRINT-SELF
<b>Delivery Ratio</b>	0.151	0.091	0.129	0.147	0.137
<b>Overhead</b>	5.162	4.125	3.165	1.551	2.945
<b>Latenza Media</b>	$1.147 \cdot 10^5$	$1.211 \cdot 10^5$	$1.171 \cdot 10^5$	$1.145 \cdot 10^5$	$1.149 \cdot 10^5$

In tab.8 SORSI si dimostra superiore rispetto a SORSI-NS e Bubble Rap ottenendo performance più vicine al routing Epidemico in termini percentuale di messaggi consegnati. Anche la latenza media è più bassa in SORSI in paragone con gli altri protocolli. Tuttavia, l'overhead di SORSI è più alto rispetto a SPRINT-SELF. Ciò significa che il numero di nodi implicati nel trasferimento dell'informazione è maggiore in SORSI e SORSI-NS rispetto a SPRINT-SELF. Lo stesso trend può essere osservato in tab.9 in cui SORSI nello scenario LAPLAND supera tutte le tecniche di diffusione per tutte le metriche di performance.

**Tabella 10 Performance dei protocolli su traccia di mobilità SASSY quando i nodi hanno il 75% di energia residua.**

	Epidemico	Bubble Rap	SORSI-NS	SORSI	SPRINT-SELF
<b>Delivery Ratio</b>	0.461	0.387	0.404	0.443	0.427
<b>Overhead</b>	10.182	6.183	4.889	4.751	3.559
<b>Latenza Media</b>	2009.573	3275.596	2354.773	2010.551	2009.762

**Tabella 11 Performance dei protocolli su traccia di mobilità LAPLAND quando i nodi hanno il 75% di energia residua.**

	Epidemico	Bubble Rap	SORSI-NS	SORSI	SPRINT-SELF
<b>Delivery Ratio</b>	0.231	0.189	0.201	0.229	0.211
<b>Overhead</b>	13.769	5.125	4.751	3.161	4.125
<b>Latenza Media</b>	$1.122 \cdot 10^5$	$1.109 \cdot 10^5$	$1.131 \cdot 10^5$	$1.122 \cdot 10^5$	$1.123 \cdot 10^5$

Quando il numero di nodi egoisti aumenta come mostrato in tab.10 e tab.11, le metriche di performance relative alla latenza media e alla delivery ratio sono sempre migliori rispetto alle altre strategie. Tuttavia, l'overhead di SORSI è peggiore rispetto a SASSY, mentre è migliore nello scenario LAPLAND. Questo significa che quando sono permessi più TTL e lo scenario è piccolo, SORSI non ha sufficiente tempo per coinvolgere più nodi nell'instradamento.

**Tabella 12 Performance dei protocolli su traccia di mobilità SASSY quando i nodi hanno il 75% di buffer pieno.**

	Epidemico	Bubble Rap	SORSI-NS	SORSI	SPRINT-SELF
<b>Delivery Ratio</b>	0.452	0.385	0.404	0.449	0.421
<b>Overhead</b>	9.332	5.957	3.146	1.161	2.593
<b>Latenza Media</b>	2002.765	2754.925	2012.48	1984.275	1999.734

**Tabella 13 Performance dei protocolli su traccia di mobilità LAPLAND quando i nodi hanno il 75% di buffer pieno.**

	Epidemico	Bubble Rap	SORSI-NS	SORSI	SPRINT-SELF
<b>Delivery Ratio</b>	0.261	0.199	0.219	0.259	0.231
<b>Overhead</b>	5.125	4.165	3.165	1.447	2.169
<b>Latenza Media</b>	$1.139 \cdot 10^5$	$1.172 \cdot 10^5$	$1.154 \cdot 10^5$	$1.137 \cdot 10^5$	$1.138 \cdot 10^5$

Considerazioni simili possono essere fatte quando viene presa in considerazione la soglia di occupazione del buffer. I risultati apprezzabili in tab.12 e tab.13, mostrano come vengano influenzate le diverse metriche di performance. In questi casi, SORSI dimostra sempre una percentuale di messaggi recapitati e una latenza media migliori rispetto alle altre strategie di instradamento con l'eccezione del routing Epidemico che è comunque considerato il benchmark per la delivery ratio. Inoltre, l'overhead incrementa di poco nello scenario SASSY in cui SPRINT-SELF è più performante rispetto a SORSI e a SORSI-NS.

---

# Capitolo VI

## Conclusioni, Pubblicazioni e Sviluppi Futuri

### 6.1 Conclusioni

Attraverso il lavoro di tesi svolto, sono stati ottenuti importanti risultati nelle aree delle reti opportunistiche, delle reti sociali multi – layer e del routing opportunistico. In particolare è stata fornita una metodologia dettagliata per l’analisi di una collezione di sei tracce di mobilità reale per reti opportunistiche, utilizzando un approccio multi – layer. Attraverso questo studio si è mirato ad ottenere una maggiore comprensione del comportamento sociale degli utenti in termini di comportamenti egocentrici e sociocentrici che sono stati derivati, non solo dalla mobilità fisica, ma anche dalla disponibilità di informazioni addizionali ottenute dal livello di rete sociale costruito sulle amicizie Facebook. I risultati hanno mostrato che il comportamento online e offline per le misure di centralità di betweenness, ego betweenness, closeness, degree ed eigenvector non sono significativamente correlate sulla maggior parte dei dataset. In altre parole, la popolarità di un nodo cambia significativamente tra i mondi sociali online e offline. Anche la struttura delle community online e offline è diversa. Analizzando la rete sociale a due livelli costruita sui dati di mobilità e sulle amicizie Facebook, si è dimostrato che in alcuni casi, i ruoli di intermediazione assunti dai nodi tra i livelli online e offline sono

---

simili. Tuttavia, in generale, la correlazione tra tali ruoli è comunque assai variabile da un dataset all'altro e non è dunque possibile estrarre un trend comune. Analizzando invece i legami forti si è dimostrato come in cinque dataset su sei gli strong ties sul livello sociale costruito sugli incontri wireless corrispondono alle amicizie Facebook. I risultati ottenuti fanno riferimento ad una collezione di dataset che copre differenti ambienti, quali quello accademico, conferenziale e urbano, con diversi pattern di mobilità, durata, numero di nodi e grafi Facebook di amicizia. Con la scelta di questa collezione dati, si è mirato ad analizzare i comportamenti sociocentrici ed egocentrici degli utenti sulle reti sociali DSN e OSN, in ambienti in cui le applicazioni di rete opportunistica possono essere adottate come soluzioni per comunicare. Si crede dunque che i risultati ottenuti, rappresentativi di diversi ambienti d'esperimento, possano essere sfruttati dalla comunità di ricerca sulle reti opportunistiche quando si approciano gli aspetti sociali degli utenti. Avendo dimostrato che le centralità di rete e le community di un dato utente variano in maniera rilevante per i mondi sociali online e offline, nella creazione di futuri algoritmi sociali, queste caratteristiche dovrebbero essere prese in considerazione. Ad esempio, un algoritmo social – based per l'instradamento che sfrutti la centralità e le community come metriche sociali dovrebbe tenere in considerazione le informazioni sociali online nella fase di bootstrapping della rete in cui le informazioni sociali offline sono parziali. Mentre, la buona corrispondenza tra legami forti potrebbe essere sfruttata per predire gli strong ties della DSN semplicemente essendo a conoscenza se due utenti sono amici Facebook o meno. Questo permetterebbe ad un algoritmo di instradamento di evitare il calcolo di metriche pesanti a livello computazionale usando solo le informazioni sulle amicizie Facebook.

A partire dagli studi effettuati sui meccanismi che regolano l'instradamento nelle reti opportunistiche, è stata inoltre analizzata e approfondita l'analisi degli effetti della selfishness in tali reti. Una volta dimostrati i suoi effetti negativi sulle performance in tali reti, è stato proposto un algoritmo per mitigarne gli effetti. Una volta delineate le principali ragioni per le quali un nodo può agire in maniera egoistica e valutarne gli effetti sull'instradamento, è stato proposto SORSI, uno schema di routing



---

opportunistic selfishness-aware che usa le informazioni della rete sociale multi-layer e lo storico di instradamento per guidare le decisioni di routing. Utilizzando per le simulazioni due tracce di mobilità reali, si è dimostrato che SORSI è capace di superare Bubble Rap, SORSI-NS (versione senza la rilevazione della selfishness) e SPRINT-SELF, avvicinando il routing epidemico in termini di delivery ratio e con un più basso costo in termini di overhead. Inoltre, SORSI supera SPRINT-SELF e Bubble Rap anche nei casi in cui le soglie di buffer o di energia possono influenzare il comportamento selfish aumentando o diminuendo il numero di nodi che agiscono in maniera egoistica. In tali scenari, SORSI continua ad avere migliori performance in termini di delivery ratio riducendo la latenza media in comparazione con le altre tecniche di disseminazione dell'informazione in contesti sociali. In definitiva, SORSI è capace di aumentare il numero di nodi che cooperano all'interno della rete opportunistica quando alcuni di essi hanno un'attitudine iniziale ad essere selfish.

## **6.2 Pubblicazioni**

Il lavoro di tesi portato avanti ha permesso di ottenere diverse pubblicazioni di seguito suddivise per tipologia.

### **6.2.1 Journal**

*J1 A Socievole, F De Rango, and A Caputo. Opportunistic mobile social networks: From mobility and facebook friendships to structural analysis of user social behavior. Computer Communications, 87:1–18, 2016.*

### **6.2.2 International Conference**

*C1 A. Socievole, F. D. Rango, A. Caputo and S. Marano, "Simulating node selfishness in opportunistic networks", 2016 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS), Montreal, QC, Canada, 2016, pp. 1-6.*

- 
- C2 A. Socievole, A. Caputo and S. Marano “Multi-layer sociality in Opportunistic Networks: an Extensive Analysis of Online and Offline Node Centrality”, in *IEEE/SCS International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, Chicago, IL, USA, 8-11, July 26-29, 2015.
- C3 A. Socievole, A. Caputo, F. De Rango, S. Marano. “Do-it-yourself Networks: a Multi-layer Network approach to the Analysis of Mobile User Egocentric and Sociocentric Behaviors”, in *ACM MobiSys DIYNet Workshop*, Florence, Italy, May 18, 2015.
- C4 A. Socievole, F. De Rango and A. Caputo, “Wireless Contacts, Facebook Friendship and Interests: Analysis of a Multi – Layer Social Network in an Academic Environment”, *Wireless Days (WD)*, 2014 IFIP, Rio De Janeiro, Brazil, pp.1-7, Nov 12-14, 2014.

### 6.3 Sviluppi futuri

Nello sviluppo del lavoro di tesi sono emerse diverse problematiche che necessitano di ulteriori attenzioni in futuro. Di seguito vengono riportate le più rilevanti:

- L’utilizzo di un’architettura sociale multi – layer ha aperto le porte ad un’analisi su più dimensioni sociali. Nello studio portato avanti si sono considerate due dimensioni, ma risulta chiaro che, nell’ottica di acquisire maggiori informazioni sulla socialità dei nodi, possano essere considerate anche altre reti sociali.
- L’utilizzo degli approcci egocentrico e sociocentrico nell’analisi delle reti multi – layer ha delineato chiaramente i rapporti tra i mondi online e offline. Ciò che manca in letteratura è probabilmente la presenza di una misura di centralità del nodo applicabile sull’intera architettura a livelli e che fornisca un valore cumulativo o comunque rappresentativo dell’importanza del nodo stesso, sfruttabile per l’instradamento ma anche in diversi altri ambiti.

- 
- La proposta di SORSI è stata validata mediante l'utilizzo di due tracce di mobilità reali. Tale analisi può essere estesa ad altri dataset investigando anche l'impatto di altri parametri come il TTL e il rate di generazione dei messaggi. Vanno inoltre approfondite le ragioni teoriche che stanno dietro le performance di SORSI.

# Bibliography

- <sup>1</sup> V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, K. Durst R.and Scott, K. Fall, H. Weiss, Delay-tolerant networking architecture, RFC4838, April (2007).
- 2 K. Fall, A delay-tolerant network architecture for challenged internets, in: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications, in: SIGCOMM '03, ACM, New York, NY, USA, 2003, pp. 27–34, doi: 10.1145/863955.863960.
- 3 A. Socievole, F. De Rango, C. Coscarella, Routing approaches and performance evaluation in delay tolerant networks, in: Wireless Telecommunications Symposium (WTS), 2011, 2011, pp. 1–6, doi: 10.1109/WTS.2011.5960870.
- 4 L. Pelusi, A. Passarella, M. Conti, Opportunistic networking: data forwarding in disconnected mobile ad hoc networks, Commun. Mag. IEEE 44 (11) (2006) 134–141, doi: 10.1109/MCOM.2006.248176.
- 5 P. Antoniadis, J. Ott, A. Passarella, Do it yourself networking: an interdisciplinary approach, in: Dagstuhl reports, 2014, pp. 125–151.
- 6 D. Zhao, H. Ma, S. Tang, X.-Y. Li, Coupon: A cooperative framework for building sensing maps in mobile opportunistic networks, Parallel Distrib. Syst. IEEE Trans. 26 (2) (2015) 392–402.
- 7 A. Mtibaa, M. May, C. Diot, M. Ammar, Peoplerank: Social opportunistic forwarding, in: INFOCOM, 2010 Proceedings IEEE, 2010, pp. 1–5, doi: 10.1109/INFOCOM.2010.5462261.
- 8 R.I. Ciobanu, C. Dobre, V. Cristea, SPRINT: social prediction-based opportunistic routing, in: Proceedings of Seventh IEEE WoWMoM workshop on autonomous and opportunistic communications, AOC 2013, Madrid, Spain, 2013, pp. 1161–1166.
- 9 R.I. Ciobanu, C. Dobre, V. Cristea, Social aspects to support opportunistic networks in an academic environment, in: Ad-hoc, Mobile, and Wireless Networks, Springer, 2012, pp. 69–82.

- 10 G. Bigwood, T. Henderson, Bootstrapping opportunistic networks using social roles, in: *World of Wireless, Mobile and Multimedia Networks (WoW- MoM)*, 2011 IEEE International Symposium on a, 2011, pp. 1–6, doi: 10.1109/ WoWMoM.2011.5986139.
- <sup>11</sup> A. Socievole, S. Marano, Exploring user sociocentric and egocentric behaviors in online and detected social networks, in: *Future Internet Communications (BCFIC)*, 2012 2nd Baltic Congress on, 2012, pp. 140–147, doi: 10.1109/BCFIC. 2012.6217994.
- <sup>12</sup> A. Socievole, F. De Rango, A. Caputo, Wireless contacts, facebook friendships and interests: Analysis of a multi-layer social network in an academic environment, *Wireless Days (WD)*, IFIP, 2014.
- <sup>13</sup> Erdős R., Rényi A., “On the evolution of random graphs.”, (1960).
- <sup>14</sup> Albert-László Barabási, Réka Albert, Hawoong Jeong, “Mean-field theory for scale-free random networks.”, *Physica A* 272, pp 173-187 (1999).
- <sup>15</sup> Albert-László Barabási, Réka Albert, “Emergence of Scaling in Random Networks. *Science*.”, vol 286, October (1999).
- <sup>16</sup> Albert-László et Al., “Deterministic scale-free networks.”, *Physica A* 299, pp. 559-564 (2001).
- <sup>17</sup> M. Magnani, L. Rossi, The ML-model for multi-layer social networks, in: *Advances in Social Networks Analysis and Mining (ASONAM)*, 2011 International Conference on, IEEE, 2011, pp. 5-12.
- <sup>18</sup> Caputo, A. Socievole, F. De Rango, CRAWDAD data set unical/socialblueconn (v. 2015-02-08). <http://crawdad.org/unical/socialblueconn/>, 2015.
- <sup>19</sup> R.I. Ciobanu, C. Dobre, CRAWDAD data set upb/mobility2011 (v. 2012-06- 18) <http://crawdad.org/upb/mobility2011/>, 2012.
- <sup>20</sup> E. Yoneki, F.B. Abdesslem, Finding a data blackhole in Bluetooth scanning, 2009, (ExtremeCom).
- <sup>21</sup> G. Bigwood, D. Rehunathan, M. Bateman, T. Henderson, S. Bhatti, CRAWDAD trace set st\_andrews/sassy/mobile (v. 2011-06-03). [http://crawdad.cs.dartmouth.edu/st\\_andrews/sassy/mobile](http://crawdad.cs.dartmouth.edu/st_andrews/sassy/mobile), 2011.
- <sup>22</sup> <http://realitycommons.media.mit.edu/socialevolution.html>
- <sup>23</sup> A. Pietiläinen, C. Diot, CRAWDAD data set thlab/sigcomm2009 (v. 2012- 07-15), 2012, (Downloaded from <http://crawdad.cs.dartmouth.edu/thlab/sigcomm2009>).
- <sup>24</sup> A. Pietiläinen, E. Oliver, J. LeBrun, G. Varghese, C. Diot, Mobiclique: Middleware for mobile social networking, in: *Proceedings of the Second ACM workshop on Online social networks*, in: *WOSN '09*, ACM, New York, NY, USA, 2009, pp. 49–54, doi: 10.1145/1592665.1592678.
- <sup>25</sup> L. C. Freeman, “A set of measures of centrality based on betweenness”, *Sociometry*, 40:35-41, 1977
- <sup>26</sup> G. Sabidussi, The centrality index of a graph, *Psychometrika* 31 (4) (1966) 581–603, doi: 10.1007/BF02289527.
- <sup>27</sup> C. Dangalchev, Residual closeness in networks, *Physica A* 365 (2) (2006) 556–564.

- <sup>28</sup> P. Bonacich, Factoring and weighting approaches to status scores and clique identification, *J. Math. Sociol.* 2 (1) (1972) 113–120.
- <sup>29</sup> V. D. Blondel, J. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, *J. Stat. Mech. Theory Exp.* 2008 (10) (2008) P10 0 08.
- <sup>30</sup> G. Palla, I. Derényi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, *Nature* 435 (7043) (2005) 814–818.
- <sup>31</sup> J. Nieminen, On the centrality in a graph, *Scandinavian J. Psychol.* 15 (1) (1974) 332–336.
- <sup>32</sup> L. Freeman, Centrality in social networks conceptual clarification, in: *Social Networks*, Elsevier, 1978, pp. 215–239.
- <sup>33</sup> P. Marsden, “Egocentric and sociocentric measures of network centrality.”, *Social Networks*, vol. 24, no.4, pp. 407–422, 2002.
- <sup>34</sup> M. Everett and S. Borgatti, “Ego network betweenness.”, *Social Networks*, vol. 27, no. 1, pp. 31–38, 2005.
- <sup>35</sup> R.V. Gould, F. R.M., Structures of mediation: a formal approach to brokerage in transaction networks, *Sociol. Methodol.* (19) (1989) 91.
- <sup>36</sup> M.S. Granovetter, The strength of weak ties, *Am. J. Sociol.* (1973) 1360–1380.
- <sup>37</sup> E.M. Daly, M. Haahr, Social network analysis for information flow in disconnected delay-tolerant MANETs, *Mob. Comput. IEEE Trans.* 8 (5) (2009) 606–621, doi: 10.1109/TMC.2008.161.
- <sup>38</sup> L. Danon, A. Diaz-Guilera, J. Duch, A. Arenas, Comparing community structure identification, *J. Stat. Mech. Theory Exp.* 2005 (09) (2005) P09008.
- <sup>39</sup> A. Urpi, M. Bonuccelli, and S. Giordano. Modelling cooperation in mobile ad hoc networks: a formal description of selfishness. *WiOpt’03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, Mar 2003.
- <sup>40</sup> Ari Keränen, Jörg Ott, and Teemu Kärkkäinen. The ONE simulator for DTN protocol evaluation. In *Proceedings of the 2nd International Conference on Simulation Tools and Techniques, Simutools ’09*, pages 55:1–55:10, ICST, Brussels, Belgium, Belgium, 2009.
- <sup>41</sup> Amin Vahdat and David Becker. Epidemic routing for Partially-Connected ad hoc networks. Technical report, Duke University, April 2000.
- <sup>42</sup> F. De Rango, S. Amelio, and P. Fazio. Enhancements of epidemic routing in delay tolerant networks from an energy perspective. In *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 731–735. IEEE, 2013.
- <sup>43</sup> Thrasyvoulos Spyropoulos, Konstantinos Psounis, and Cauligi S. Raghavendra. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking, WDTN ’05*, pages 252–259, New York, NY, USA, 2005. ACM.

- <sup>44</sup> Pan Hui, Jon Crowcroft, and Eiko Yoneki. Bubble Rap: Social-based forwarding in delay-tolerant networks. *Mobile Computing, IEEE Transactions on*, 10(11):1576–1589, 2011.
- <sup>45</sup> A Socievole, E Yoneki, F De Rango, and J Crowcroft. ML-SOR: message routing using multi-layer social networks in opportunistic communications. *Computer Networks*, 81:201–219, April 2015.
- <sup>46</sup> A. Socievole, F. De Rango, and S. Marano. Face-to-face with Facebook friends: using online friendlists for routing in opportunistic networks. In *Proceedings of 24th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2013)*, September 2013.
- <sup>47</sup> F. De Rango, A. Socievole, and S. Marano. Exploiting online and offline activity-based metrics for opportunistic forwarding. *Wireless Networks*, pages 1–17, 2014
- <sup>48</sup> F. De Rango and F. Monteverdi. Social and dynamic graph-based scalable routing protocol in a DTN network. In *Performance Evaluation of Computer and Telecommunication Systems (SPECTS), 2012 International Symposium on*, pages 1–8. IEEE, 2012.
- <sup>49</sup> F. De Rango, A. Socievole, A. Scaglione, and S. Marano. Novel activity-based metrics for efficient forwarding over online and detected social networks. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International*, July 2013.
- <sup>50</sup> A Socievole, F De Rango, and A Caputo. Opportunistic mobile social networks: From mobility and facebook friendships to structural analysis of user social behavior. *Computer Communications*, 87:1–18, 2016.
- <sup>51</sup> Radu Ioan Ciobanu, Ciprian Dobre, Valentin Cristea, and Dhiya Al-Jumeily. Social aspects for opportunistic communication. In *2012 11<sup>th</sup> International Symposium on Parallel and Distributed Computing*, pages 251–258. IEEE, 2012.
- <sup>52</sup> MEJ Newman. Clustering and preferential attachment in growing networks. *Physical Review E-Statistical, Nonlinear, and Soft Matter Physics*, 64(2):251021–251024, 2001.
- <sup>53</sup> Radu Ioan Ciobanu, Ciprian Dobre, Valentin Cristea, Florin Pop, and Fatos Xhafa. Sprint-self: Social-based routing and selfish node detection in opportunistic networks. *Mobile Information Systems*, 2015, 2015.
- 54 A. Socievole and S. Marano. Evaluating the impact of energy consumption on routing performance in delay tolerant networks. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2012 8th International*, pages 481 – 486, Aug. 2012.