

Analisi della topologia e dei contenuti delle comunità virtuali per lo studio del fenomeno delle camere d'eco

Relatore: Dr. Marco Viviani

Correlatrice: Prof.ssa Gabriella Pasi

Tesi di Laurea Magistrale di:
Giacomo Villa
Matricola 807462

Anno Accademico 2019-2020

Al Dr. Marco Viviani e alla Prof.ssa Gabriella Pasi per avermi consigliato, seguito e dedicato il loro tempo durante questi mesi di lavoro e di incertezze.

A mia Mamma e a mio Papà per avermi sostenuto in questi anni di studio, non mi avete mai fatto mancare nulla e ve ne sarò per sempre grato.

A mio fratello Tommaso e a mia sorella Anita, possa io essere per voi un semplice punto di partenza e non un limite irraggiungibile.

Ai compagni conosciuti durante il percorso, con i quali ho affrontato tutti questi anni di studio e di sacrifici. Mi avete reso una persona migliore.

Ai miei amici di sempre, per avermi donato momenti di felicità e spensieratezza anche nei periodi più duri. Non è tanto dell'aiuto degli amici che abbiamo bisogno, ma del poter far affidamento su quell'aiuto.

Alla mia amata Ludovica per avermi sempre supportato e sopportato con amore e delicatezza. Ti seguirei su Marte con le lacrime alle guance.

A mia nonna Agnese, mi hai insegnato che non esiste maggior ricchezza della cultura, educandomi alla gentilezza e al ripudio della superbia. Spero che ovunque tu sia possa essere orgogliosa di me.

Abstract

La nascita dei social media ha portato a cambiamenti senza precedenti nella portata e nella velocità con cui le persone condividono e hanno accesso alle informazioni. I *feed* dei social media sono strumenti chiave per accedere a volumi elevati di notizie, opinioni e informazioni pubbliche. Tuttavia, favorendo una tale proliferazione di informazioni a cui le persone sono esposte, i social media possono interagire negativamente con i *bias* di selezione cognitiva, amplificando fenomeni indesiderati come l'estremismo delle opinioni e la diffusione di disinformazione. Ciò può avvenire quando, data una modellazione delle interazioni degli utenti come rete sociale, si giunge ad una cosiddetta *segregazione* della rete (quindi a una divisione della rete sociale in due o più gruppi scollegati o scarsamente connessi) e a una *polarizzazione* delle opinioni (l'elevata omogeneità delle opinioni all'interno di tali gruppi). Tutto questo si trova alla base della formazione delle cosiddette camere d'eco, realtà dove le informazioni, idee e opinioni vengono continuamente ripetute in un sistema chiuso, in cui i partecipanti vengono esposti alla stessa tipologia di informazioni e tendono quindi, in maniera automatica, a respingere tutto ciò che si trova in disaccordo con questa.

I processi cognitivi umani alla base delle camere d'eco, intesi come esposizione preferenziale, omofilia e influenza sociale, sono chiari da tempo in letteratura. Tuttavia anche alcuni meccanismi dei *social media* tendono a spingere verso una polarizzazione e segregazione della rete sociale: gli algoritmi dei *feed* di notizie favoriscono le news con le quali gli utenti hanno maggiori probabilità di interagire, in base al coinvolgimento passato. I sistemi di raccomandazione (ad esempio di nuove amicizie) suggeriscono nuovi legami basati su interessi, credenze e amici comuni. Infine, i social media consentono agli utenti di sciogliere i legami che tendono a collegarli con coloro con i quali non sono d'accordo, andando di fatto a eliminare una possibile fonte di informazione opposta rispetto a quella a cui si è principalmente esposti.

Questa tesi si pone come obiettivo quello di analizzare il fenomeno delle camere d'eco analizzando contenuti e relazioni sociali che caratterizzano gli utenti su una piattaforma sociale, dato un certo *topic* di discussione; lo scopo è dunque quello di verificare la presenza di camere d'eco valutando informazioni sia topologiche che semantiche, analizzare il contenuto e la struttura delle partizioni identificate al fine di validare e confermare la effettiva presenza di camere d'eco per poi cercare di mitigare tale fenomeno.

Indice

1	Introduzione al lavoro di tesi	1
1.1	Contesto applicativo	1
1.2	Obiettivo	2
1.3	Struttura	3
2	I <i>Social Media</i> come strumento di diffusione dell'informazione	5
2.1	Il <i>Web</i> e i <i>Social Media</i>	5
2.1.1	Introduzione al <i>Web</i> e alla sua evoluzione	5
2.1.2	Il <i>Web</i> 2.0 e lo sviluppo del <i>Web</i> sociale	7
2.1.3	La diffusione di informazione sui <i>Social Media</i>	9
2.1.4	La <i>Social Media Analytics</i> (SMA)	10
2.2	Rappresentare e analizzare la topologia delle reti sociali	11
2.2.1	Le reti sociali	11
2.2.2	Cenni di teoria dei grafi	13
2.2.3	<i>Social Network Analysis</i> (SNA)	15
2.2.4	<i>Community detection</i>	18
2.3	Rappresentare e analizzare i contenuti scambiati	19
2.3.1	<i>Sentiment analysis</i>	19
2.3.2	<i>Topic Modeling</i>	21
2.3.3	<i>Social Content Analysis</i> (SCA)	23
3	Il problema delle camere d'eco: Definizione e Stato dell'arte	25
3.1	Il fenomeno delle camere d'eco	25
3.1.1	Introduzione	25
3.1.2	Le camere d'eco	26
3.1.3	Possibili approcci al problema	27
3.2	Stato dell'arte	29
3.2.1	Identificazione delle camere d'eco	29
3.2.2	Superamento delle camere d'eco	39
3.2.3	Discussione	42
4	La soluzione proposta	45
4.1	Modellazione delle reti sociali	46
4.1.1	<i>Twitter</i> come fonte di dati	46
4.1.2	Dataset allo stato dell'arte	47
4.1.3	Dati Covid-19	48

4.1.4	Dati Vaccinazioni	51
4.2	<i>Community Detection</i>	52
4.2.1	Introduzione all'attività di <i>Community detection</i> nell'ambito delle camere d'eco	52
4.2.2	<i>METIS</i>	53
4.2.3	Grafi utilizzati nell'attività di <i>Community detection</i>	55
4.2.4	Considerare l'informazione semantica nel processo di <i>Community detection</i>	59
4.2.5	Metriche di valutazione topologiche	65
4.3	Quantificazione controversia	66
4.3.1	Metriche	67
4.3.2	Discussione	69
4.4	Riduzione controversia	70
4.4.1	Soluzioni basate sulla prossimità	70
4.4.2	Soluzioni basate sulla centralità	71
4.4.3	Modellazione <i>sentiment</i>	73
4.4.4	Aggiunta archi	73
5	Risultati	75
5.1	Community detection	75
5.1.1	Risultati partizionamento delle reti sociali	76
5.1.2	<i>Topic Modelling</i>	82
5.2	Analisi metadati	86
5.2.1	Covid-19	86
5.2.2	Vaccinazioni	89
5.3	Valutazione controversia	91
5.3.1	#BeefBan, #Ukraine e #NationalKissingDay	91
5.3.2	Covid-19	93
5.3.3	Vaccinazioni	93
5.3.4	Discussione	94
5.4	Link prediction	94
5.4.1	#BeefBan e #Ukraine	95
5.4.2	Covid-19	98
5.4.3	Vaccinazioni	98
5.4.4	Valutazione del <i>Sentiment Boost</i>	100
5.4.5	Analisi post <i>Link prediction</i>	104
5.4.6	Discussione	112
6	Conclusioni	115
6.1	Modellazione e partizionamento della rete	115
6.2	Valutazione quantitativa della controversia	117
6.3	Superamento polarizzazioni	118
7	Possibili sviluppi futuri	121
7.1	Superamento del partizionamento bipartito, valutazione semantica e valutazioni di SNA	121

7.2	Valutazione tratti della personalità e modifica <i>Sentiment Boost</i>	122
7.3	Hybrid <i>Link prediction</i>	123
7.4	Approfondimento dei risultati ottenuti	123
A	Immagini Analisi	125
A.1	Grafi Garimella	125
A.2	Covid-19	127
B	Immagini <i>Sentiment Boost</i>	129
B.1	Covid-19	129
B.2	Vaccination	133
Elenco delle figure		137
Elenco delle tabelle		139
Bibliografia		141

Capitolo 1

Introduzione al lavoro di tesi

In questo capitolo iniziale viene introdotto il lavoro di tesi, descrivendo il contesto, l'obiettivo e la strutturazione in capitoli. In particolare la suddivisione è la seguente:

- Nella Sezione 1.1 viene introdotto il contesto applicativo nel quale il lavoro di tesi si posiziona.
- Nella Sezione 1.2 viene descritto l'obiettivo del lavoro.
- Nella Sezione 1.3 si fornisce una guida al lettore circa la strutturazione del lavoro di tesi.

1.1 Contesto applicativo

Lo sviluppo delle tecnologie del Web 2.0 e la diffusione dei social media che su di esse si basano, come ad esempio Facebook¹, Twitter² ed Instagram³, hanno permesso a qualsiasi persona di condividere pensieri ed idee online, dando vita a una delle più grandi rivoluzioni dei nostri tempi. I *feed* dei social network sono la chiave di accesso a un grandissimo volume di news e opinioni di vario genere; le persone sono quindi esposte a un continuo proliferare di informazioni. La struttura stessa dei *social* permette agli utenti di collegarsi tra loro e decidere con chi effettivamente condividere i propri contenuti. Si vengono quindi a creare delle comunità virtuali tra utenti che utilizzano una certa piattaforma; le comunità, in una realtà virtuale, rappresentano i gruppi di nodi (dunque gli utenti) che mostrano caratteristiche simili (*topic* sui quali disquisiscono) e interagiscono tra di loro (amicizia, like, retweet, citazioni, ecc.) Fersini (2019).

All'interno delle comunità virtuali, può succedere che si vengano a creare gruppi fortemente polarizzati rispetto a certe posizioni ed idee. Quando ciò accade, si assiste ad una progressiva perdita di pensiero critico del singolo individuo a favore di comportamenti gregari: si può quindi assistere a fenomeni di estremismo e diffusione di disinformazione; quest'ultima sta diventando pervasiva nell'ambito dei social media, tanto che è stata definita dal World Economic Forum (2016) come una delle maggiori minacce per la nostra società. Inoltre, al fine di sottolineare ulteriormente la pericolosità di questo fenomeno, in

¹<https://www.facebook.com/>

²<https://twitter.com/>

³<https://www.instagram.com/>

accordo con la teoria dell'*esposizione selettiva* di Dieter (1986) gli individui tendono a consumare informazioni con contenuti che più somigliano alle proprie credenze, anziché accettare punti di vista contrastanti alle stesse; questo sicuramente aiuta ulteriormente la creazione di gruppi polarizzati consentendo a *fake news* di essere ritenute *real news* Viviani et al. (2017). Può così emergere il cosiddetto fenomeno delle *camere d'eco*, in base al quale certe informazioni, idee o credenze vengono continuamente amplificate o rafforzate dalla loro continua ripetizione all'interno di un sistema chiuso. Dunque, quando si parla di camere d'eco, ci si riferisce a quel fenomeno sociale nel quale il contenuto consumato dall'utente è fortemente inclinato verso l'opinione dello stesso andando dunque a rafforzarla, sia che questa sia quella corretta oppure no.

Al fine di identificare una camera d'eco come tale, all'interno di una comunità virtuale definita su di una rete sociale, risulta necessario valutare il livello di *controversy* che caratterizza la stessa. Possiamo vedere la *controversy* come una misura volta a quantificare il corretto flusso di informazione tra i diversi gruppi identificati nella realtà di interesse; nel caso in cui questo valore sia sufficientemente grande, si potrà concludere che si è identificata una realtà dotata di camere d'eco. Diversi lavori allo stato dell'arte valutano il livello di *controversy* tra diversi gruppi di utenti in una comunità virtuale, verificando come la possibile circolazione dell'informazione venga reso possibile all'interno delle reti sociali utilizzando diverse tecniche; altri ancora valutano caratteristiche strettamente strutturali al fine di costruire un modello di apprendimento supervisionato.

Al fine di mitigare il fenomeno delle camere d'eco, alcuni lavori di letteratura hanno investigato come sia quindi necessario lavorare sul livello di *controversy* andando a minimizzarla; in particolare, alcuni di questi lavori lo affrontano come un problema di ottimizzazione (spesso affrontato dal punto di vista teorico) andando di fatto a effettuare operazioni di *Link prediction* che possono essere più o meno sofisticate.

1.2 Obiettivo

L'obiettivo di questo lavoro di tesi è quello di progettare una *pipeline* volta ad identificare, analizzare e limitare la formazione di camere d'eco presenti su reti sociali, in particolare in piattaforme di *microblogging* come *Twitter*; a tal proposito, verranno impiegate tecniche di *Community detection*, integrando l'informazione semantica, nel tentativo di migliorarne le performance e identificare più puntualmente eventuali partizioni polarizzate.

L'informazione semantica sarà oggetto anche di un'analisi post *Community detection*, al fine di verificare il tipo di contenuto che circola nelle partizioni identificate. Questo può aiutare e quindi indirizzare, verso il motivo per il quale la potenziale polarizzazione ha avuto luogo.

Ai fini della quantificazione della *controversy*, vengono introdotti diversi indici attualmente presenti allo stato dell'arte e ne vengono proposti di nuovi, volti a cogliere e sfruttare i possibili diversi aspetti topologici della rete. Nel tentativo di minimizzare il livello di *controversy*, vengono utilizzate tecniche di *Link prediction* in funzione di caratteristiche strutturali dei nodi, tenendo in considerazione anche la probabilità di accettazione da parte dei membri della rete sociale di un nuovo possibile arco di collegamento tra due utenti.

1.3 Struttura

In questa sezione viene illustrata la struttura della tesi in capitoli, si fornisce dunque una rapida panoramica sul contenuto degli stessi in modo da facilitare la comprensione da parte del lettore.

Capitolo 2: in questo capitolo viene introdotto il contesto tecnologico nel quale le camere d'eco vedono la loro comparsa, vale a dire il cosiddetto *Social Web* ed in particolare i *Social Media*.

Capitolo 3: in questo capitolo si descrivono i principali lavori presenti allo stato dell'arte dai quali si è tratta ispirazione per la soluzione proposta al problema considerato. Vengono dunque analizzati i lavori che trattano la definizione e l'identificazione delle camere d'eco, con la consecutiva attività di superamento delle stesse, l'attività di *Community detection* con *focus* sul tentativo di inserimento di informazioni semantiche nella stessa mediante tecniche di *Sentiment analysis* e *topic modelling*.

Capitolo 4: in questo capitolo si mostrano e discutono le soluzioni implementate relative al *focus* del lavoro di tesi, in particolare vengono discusse l'attività di modellazione delle informazioni in reti sociali, l'attività di *Community detection*, vengono presentate le metriche di valutazione della stessa, le attività di quantificazione della *controversy* e le soluzioni proposte circa la riduzione della stessa.

Capitolo 5: vengono qui forniti i risultati ottenuti per tutte le attività precedentemente menzionate, viene inoltre fornita un'analisi circa i contenuti e le informazioni disponibili date le partizioni definite nell'attività di *Community detection*; si ritengono importanti tali analisi in quanto forniscono ulteriori conferme circa la qualità delle comunità polarizzate identificate.

Capitolo 6: in questo capitolo viene analizzato ogni *step* della soluzione proposta, si procederà quindi alla valutazione, con occhio critico, delle scelte prese e/o migliorabili.

Capitolo 7: in quest'ultimo capitolo, discussi alcuni limiti delle assunzioni effettuate nel lavoro di tesi, si propongono dei possibili sviluppi futuri e miglioramenti circa la soluzione proposta, consapevoli di essere autori dell'aggiunta di un singolo tassello al mosaico più ampio che costituisce il problema analizzato.

Capitolo 2

I *Social Media* come strumento di diffusione dell'informazione

In questo capitolo si fornisce un'introduzione al contesto tecnologico in cui questo lavoro di tesi si inserisce. In particolare:

- Nella Sezione 2.1 si fornisce un'introduzione al *Web* e ai *Social Media* in cui il problema affrontato in tesi si verifica e viene studiato; passando poi alla definizione dell'attività di analisi effettuabile nel contesto considerato.
- Nella Sezione 2.2 vengono introdotte le definizioni teoriche e le metodologie che permettono di modellare i dati, intesi come relazioni tra i protagonisti della realtà di interesse; si fornisce poi un approfondimento delle analisi strutturali, appartenenti al più ampio ventaglio di analisi dei *social media* precedentemente definito.
- Nella Sezione 2.3 si introducono le definizioni e le modellazioni in funzione del contenuto semantico scambiato; fornendo quindi un approfondimento circa le possibili analisi sul contenuto stesso, secondo pilastro delle analisi effettuabile nel contesto considerato.

2.1 Il *Web* e i *Social Media*

In questa sezione vengono fornite la definizione e la presentazione della storiografia del *Web*, concentrandosi poi sulla possibilità di interazione online tra gli utenti (Sezione 2.1.1).

Viene quindi presentato il concetto di *Web Sociale*, dallo stesso si passerà poi alla definizione di *Social Media* (Sezione 2.1.2). Viene poi affrontato il problema della diffusione dell'informazione sui *Social Media* (Sezione 2.1.3), terminando con la presentazione dell'attività di analisi in tale contesto (Sezione 2.1.4).

2.1.1 Introduzione al *Web* e alla sua evoluzione

Il *World Wide Web* (abbreviato in WWW o semplicemente *Web*) costituisce uno dei preminenti servizi associati ad *Internet*. Il concetto fondante del World Wide Web è da attribuirsi a *Tim Berners-Lee*; lo stesso lavorò a un sistema volto alla condivisione di documentazione scientifica, che potesse prescindere dalla piattaforma di visualizzazione. Il

proposito era definire un sistema di comunicazione efficace per *European Organization for Nuclear Research* (CERN).

La nascita del Web è datata 1991; da quel momento l'evoluzione non si è mai fermata e continua inesorabilmente, utilizzando e facendosi affiancare da tutto quell'universo di nuove tecnologie, che si sviluppano grazie alla crescita del contesto informatico. La funzionalità base del Web è la possibilità di navigare e usufruire di un insieme molto vasto di contenuti (multimediali e non), i quali risultano essere tra loro interconnessi, permette anche di avere accesso a tutta una serie di servizi (2019), i quali cambiano e sono definiti dall'epoca che si analizza. Risulta infatti possibile suddividere la tecnologia in esame, in diverse versioni, le quali sanciscono delle vere e proprie “*epoches*” informatiche.

La Figura 2.1 rappresenta l’evoluzione passata e la possibile futura evoluzione del Web, mettendo sugli assi due elementi fondamentali: il livello di *socializzazione* inteso come possibilità di socializzare con altre entità facenti parte della rete modellata dal Web e l’utilizzo della *semantica* nei servizi offerti.

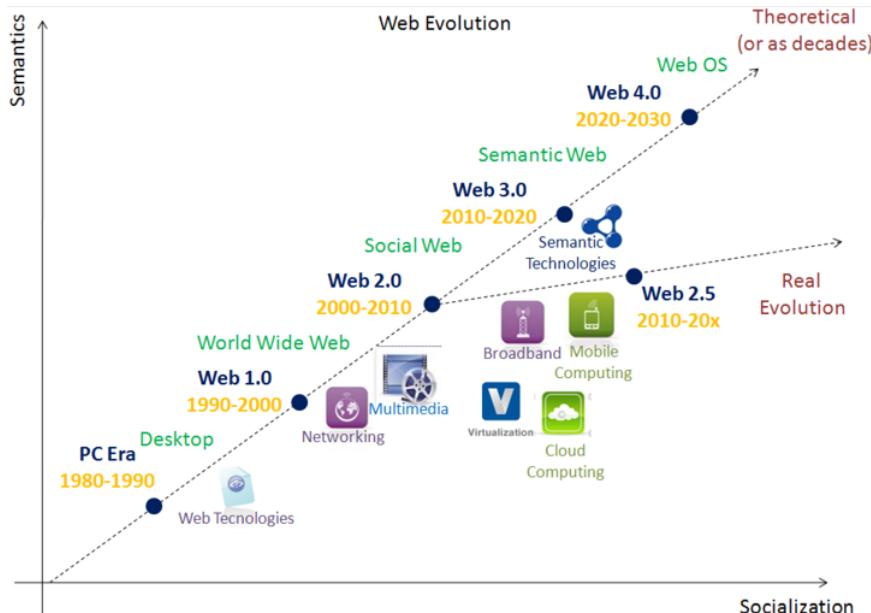


Figura 2.1: Evoluzione del Web^[37]

Web 1.0 Partendo dalla prima generazione, il Web 1.0 (denominazione attribuita successivamente alla nascita dello stesso), forniva la possibilità agli utenti di fruire di certi contenuti intesi come risorse testuali, l'utilizzo di questi si limitava però solo alla visualizzazione; di fatti, solo una piccola porzione di utenti costituiva i cosiddetti “*creators*”, i quali avevano il potere di creare nuovi contenuti.

Il Web 1.0 è conosciuto anche come “*internet dei contenuti*”, contraddistinto da pagine web *statiche*, dunque pagine che vengono visualizzate dall'utente così come sono state memorizzate sul server, senza elementi di personalizzazione in funzione del visualizzatore o di altri fattori e con frequenza di aggiornamento molto ridotta.

Web 2.0 La seconda generazione, il Web 2.0, ha permesso agli utenti di interagire e collaborare tra loro, generando e scambiando contenuti; si entra quindi nella fase

denominata “*read-write web*”. Con l’introduzione dei linguaggi di programmazione dinamici, gli sviluppatori hanno permesso all’utenza di interagire con i contenuti dei siti. Si inizia a introdurre il concetto di *usabilità* di un sito, questa Era è quindi caratterizzata da una partecipazione attiva degli utenti, nell’ottica di costruire dei contenuti, di classificarli e quindi distribuirli. Nascono i *Blog*, le *Wiki*, i *Social Network* e i Forum di discussione.

Web 3.0 Dal 2010 in poi viene introdotto il concetto di Web 3.0, il quale porta tutta una serie di nuovi aspetti evolutivi come il concetto di *web semantico*: forte associazione dei documenti (intesi come pagine HTML, file, immagini e così via) a tutta una serie di metadati che ne specificano e ne definiscono il contesto semantico, aprendo quindi la possibilità di elaborazione automatica e facilitando le attività di interrogazione e interpretazione. In questa realtà vedono l’utilizzo tutta una serie di linguaggi che fanno leva sulla logica dei predicati, questo al fine di poter modellare la conoscenza (e.g., le triple RDF).

Vengono inoltre introdotte funzionalità di *intelligenza artificiale*, consentendo ai motori di ricerca di effettuare interrogazioni efficienti partendo dal linguaggio naturale, dunque orientando tali tecnologie nei confronti del reperimento di documenti, tenendo in considerazione le peculiarità, le preferenze e la storia degli utenti; inoltre vi è una nuova e maggiore capacità di calcolo e nuovi algoritmi volti a creare mondi “*paralleli*” a quello reale (e.g., SecondLife).¹

Web 4.0 Definizione ancora in fase di chiara definizione e per questo piuttosto "instabile" è quella di Web 4.0, essendo una previsione per il futuro delle tecnologie che lo costituiranno. In questo nuova generazione si potrà assistere a una realtà aumentata predominante, affiancata dall’introduzione delle nuove tecnologie 5G. Nuove interfacce e nuovi dispositivi collegati alla rete, andranno a definire nuove realtà domotiche e reti *IoT*. Si potranno definire dei veri e propri alter ego digitali, con documenti che si aggiornano e si collegano fra loro, permettendo una personalizzazione e un accesso più puntuale a ciò di cui siamo in cerca o di cui abbiamo bisogno, il tutto causerà un controllo sempre maggiore delle informazioni, che andrà ad influenzare in maniera non indifferente, non solo il mondo digitale, ma anche la vita che si conduce nel mondo reale.

2.1.2 Il Web 2.0 e lo sviluppo del Web sociale

Dopo aver fornito una panoramica circa la storia e il futuro del *Web*, risulta necessario focalizzarsi sui paradigmi tecnologici che più di tutti hanno permesso agli utenti di interagire attivamente, comportando poi la nascita del fenomeno che questa tesi si pone di studiare.

Il Web 2.0, permette agli utenti di effettuare tutta una serie di attività di interazione e collaborazione, permettendo a questi di creare comunità virtuali basate su comuni interessi e passioni, nelle quali vengono generati e scambiati contenuti; la comunicazione diventa quindi bidirezionale, del tutto in contrasto con la generazione Web 1.0, in cui gli utenti erano limitati, nella maggioranza dei casi, ad una semplice visualizzazione delle risorse.

¹<https://secondlife.com/>

Nel contesto del Web 2.0, vede la luce il cosiddetto “*Web Sociale*”: insieme di relazioni sociali che permettono l’interconnessione delle persone. L’aspetto sociale della comunicazione viene favorito dalle tecnologie poste alla base del Web 2.0 (2019), di fatto permettono di facilitare l’interazione tra persone che:

- Condividono gli stessi gusti e/o interessi.
- Si trovano ad operare assieme in contesti geografici specifici o per scopi specifici.

Il Web 2.0 si dimostra essere terreno fertile per lo sviluppo dei *Social Media*. Una possibile definizione di *social media* è fornita nel lavoro di Kietzmann et al. (2011): “*I social media sono tecnologie interattive mediate da reti di computer che facilitano la creazione e la condivisione di informazioni, idee, interessi e altre forme di espressione tramite community e virtual network*”.

Con l’avanzata continua e inesorabile della digitalizzazione sociale, lo studio scientifico dei *social media* e delle dinamiche interne agli stessi è necessario, risultando poi un mezzo utile al fine di comprendere gli aspetti delle società contemporanee. Dagli studi proposti da Kaplan et. al (2010), volti alla analisi dei diversi social media, sono state definite delle caratteristiche comuni agli stessi: i *social media* sono applicazioni Web 2.0 interattive, che facilitano lo sviluppo di social network online collegando i profili degli utenti (creati dagli stessi) tra di loro, i contenuti sono tutti *user generated*.

Particolare attenzione deve essere posta su quest’ultimo punto, con *User-Generated Content* si intende una qualsiasi forma di contenuto creato dagli utenti di un sistema o servizio online, spesso reso disponibile tramite l’utilizzo dei *social media*. Le motivazioni alla base, circa la creazione di *UGC*, si possono suddividere in due classi:

- **Incentivi impliciti:** gli incentivi impliciti consentono all’utente di sentirsi “bene” in quanto membro attivo della comunità, includono e si basano su relazioni tra gli utenti (e.g., amicizia, *followee/follower*), sulla condivisione di contenuto riguardante l’utilizzo di un qualche prodotto (e.g., recensioni su Amazon), inoltre possono permettere di raggiungere uno “status” all’interno del *social* in questione (e.g., spunta blu, *local guide*). Diversi studi circa la valutazione di queste motivazioni sono presenti in letteratura (Lee et. al. 2012), gli stessi cercano di definire sempre meglio le ragioni per cui si è portati a condividere e creare contenuti, senza magari ricevere nulla di tangibile come ricompensa.
- **Incentivi esplicativi:** a differenza di prima, si parla di premi tangibili, gli stessi hanno un valore immediato (e.g., iscrizione che assicura l’ottenimento di un qualche bonus). Un possibile svantaggio di questa realtà è costituito dal grande *effort* che il *social media* deve infondere, nel mantenere alto il livello di partecipazione all’interno dello stesso; questo perché, mediante l’incentivo esplicito, l’utente potrebbe non essere motivato socialmente alla partecipazione, in quanto interessato solo al bene materiale.

La valutazione degli *UGC* può essere sfruttata per diversi fini come ad esempio valutazioni di marketing o per effettuare analisi dei contenuti intesa come *Sentiment analysis* e *topic modeling*; tutte queste attività sono racchiuse nel *task* di *Social Content Analytics*, che verrà introdotta nella Sezione 2.3.

La presenza predominante degli *UGC* permette quindi di attuare attività volte a “*misurare*” il contenuto scambiato (e.g., commenti, risposte, frequenza uso parole). Altri elementi che accomunano i *social media* riguardano l’accessibilità agli stessi tendenzialmente facile e gratuita, l’immenso e internazionale bacino d’utenza, la possibilità di fruire, velocemente, di una grandissima quantità di informazioni (2019).

Analizzando ora le differenze che possono presentare i *social*, emerge in *primis* la *teoria della presenza sociale* (Williams et al. 1976), questa afferma che i *social media* differiscono nel grado di *presenza sociale*, intesa come il livello di contatto acustico, visivo e fisico che può essere raggiunto tra gli utenti del *social* stesso (e.g., il livello di presenza sociale è inferiore per comunicazioni e-mail rispetto alle comunicazioni live chat, così come è maggiore per comunicazioni *face-to-face* che per comunicazioni telefoniche). Un’altra differenza risiede nella *ricchezza informativa* dei *social* (Daft et al. 2010), intesa come: quantità e qualità delle informazioni trasmesse in un certo intervallo di tempo; risulta quindi che alcuni *social media* siano più efficaci di altri nel risolvere ambiguità e incertezza.

Ulteriori differenze risiedono nelle tipologie di contenuti che possono essere condivisi e creati dagli utenti: articoli, *post*, *tweet*, *thread*, *review*, immagini e video.

2.1.3 La diffusione di informazione sui *Social Media*

L’analisi proposta circa il Web 2.0, fornita nella precedente sezione, ha introdotto il concetto di *User Generated Content* sottolineando poi la rivoluzionaria possibilità di interagire con questi contenuti, relazionandosi quindi con altri utenti e iniziando veri e propri *thread* di discussione.

Dall’avvento delle tecnologie Web 2.0 in poi, molta informazione ha iniziato a circolare sulle nuove piattaforme sociali (e.g., forum, *blog*, ecc.); l’aumento poderoso di questa è dovuta al già citato orientamento al *UGC*. Questa nuova realtà permette a chiunque di fruire velocemente di moltissima informazione, andando a semplificare e potenzialmente aumentare il processo di reperimento di conoscenza; con l’avvento di logiche volte alla raccomandazione in funzione del comportamento degli utenti, questa grande possibilità può tramutarsi in “un’arma a doppio taglio”.

Le macchine stanno andando sempre più a sostituire l’uomo nell’ambito giornalistico, soprattutto in quell’attività volta alla distribuzione di notizie (Nechushtai et. al. 2019). Nei media tradizionali, ciò a cui le persone erano esposte era determinato principalmente dagli editori che detenevano il controllo sul flusso quotidiano delle notizie, le persone ricevono passivamente i contenuti delle notizie distribuiti nei tradizionali canali dei media (come una sorta di Web 1.0). Ora, applicazioni come *Google News* stanno sempre più assumendo il ruolo di “*gatekeeping*” delle notizie, andando a effettuare attività di selezione della *news* e posizionamento della stessa nel *feed* dell’utente, il tutto basandosi sull’utente stesso, eseguendo quindi operazioni di raccomandazioni, novità assoluta rispetto ai media tradizionali.

Questa logica di raccomandazione viene trasportata anche sulle realtà definite dai *social network*; gli algoritmi di Facebook, per esempio, danno la priorità a determinati contenuti in funzione del profilo, delle preferenze e della storia (intesa come *history*) dell’utente. Questo può portare a situazioni nelle quali all’utente venga sempre proposto materiale simile, andando dunque a minimizzare la metrica di “*serendipity*”, la quale

misura quanto il sistema di raccomandazione sia in grado di “sorprendere” l’utente con materiale nuovo, che lo stesso “non si aspettava”. Questo fenomeno può quindi, dati certi argomenti, essere una della cause che porta alla polarizzazione degli utenti di una rete sociale.

Risulta quindi necessario domandarsi, per poi trovare, che tipologia di “*gatekeeping*” delle notizie risulti essere la migliore al fine di permettere un più ampio ventaglio di informazioni ottenibili, cercando di evitare l’esposizione continua alla stessa tipologia di informazione.

2.1.4 La *Social Media Analytics* (SMA)

Con il termine “*Social Media Analytics*” si intendono le attività di monitoraggio, analisi, misurazione e interpretazione delle interazioni digitali e le relazioni delle persone, temi, idee e contenuti che avvengono sulle piattaforme digitali. Le interazioni si verificano all’interno di comunità prestabilite e/o costituite spontaneamente².

In particolare, la *Social Media Analytics*, racchiude al suo interno due gradi famiglie di analisi:

- **Social Network Analysis:** questa tipologia di analisi si concentra più sull’aspetto strutturale, date le interazione tra gli utenti del *social media* analizzato. Questa tipologia di analisi verrà introdotta e definita nella Sezione 2.2.
- **Social Content Analysis:** questa tipologia di analisi si concentra sul contenuto semantico scambiato e generato tra gli utenti del *social media* analizzato, dunque valuta esclusivamente gli *User-Generated Content*. Questa tipologia di analisi verrà introdotta e definita nella Sezione 2.3.

Come sottolineato precedentemente, l’attività di *SMA* è una possibile occasione di business per le aziende, le stesse possono comprendere i modelli di comportamento e trend di mercato, in modo da ottenere un vantaggio competitivo nei confronti della concorrenza.

In particolare, in questo contesto, vengono effettuate delle operazioni di *Social Media Listening*, ovvero attività volte a identificare e tracciare le conversazioni online intorno a specifiche *keywords* (e.g., nome *brand*, nome settore di riferimento). Bisogna sottolineare che, il *Social Media Listening*, non è sinonimo di *Social Media Monitoring*; il secondo è di fatto solo uno *step* del primo, dove vengono semplicemente raccolti dati relativi alle *mention*, *like*, *sentiment* e di tutti quei dati volti a identificare e quantificare il *social media engagement* (2019).

In generale l’attività di *Social Media Analytics* può avere diverse inclinazioni, in funzione poi dell’entità, intesa come persona, azienda o associazione, che ne richiede l’utilizzo. Un aspetto molto importante da tenere in considerazione, che spesso passa in sordina oscurato dalle motivazioni economiche, riguarda l’occasione di studio delle relazioni umane che le attività di *SMA* rendono possibile.

Tramite le attività che compongono la *Social Media Analytics*, risulta possibile andare a studiare come si creino e come si sviluppino le relazioni umane, si può studiare in che modo il grado di istruzione possa andare a influire sui pensieri, il sentimento racchiuso nei

²<https://www.gartner.com/en/information-technology/glossary/social-analytics>

contenuti scambiati, ma anche sulla probabilità che una notizia falsa possa essere ritenuta vera o viceversa, apprendo quindi tutta una pagina di studi circa la struttura della *news*: che cosa colpisce prima? Perché?

Risultano quindi importanti, oltre al contenuto condiviso dall'utente e le interazioni dello stesso, tutta la serie di metadati che descrivono ogni singolo protagonista, informazioni come: la data di nascita, l'origine, il percorso di studi, il percorso lavorativo possono essere fondamentali per delineare i tratti personali di determinati gruppi. In questo senso, la *SMA* risulta fondamentale nella studio delle *camere d'eco* (le quali verranno introdotte nel Capitolo 3) in quanto l'analisi dei contenuti scambiati insieme alle valutazioni circa la strutturazione delle relazioni permette, e ha permesso, di delineare e comprendere meglio le contrapposizioni identificate.

2.2 Rappresentare e analizzare la topologia delle reti sociali

In questa sezione vengono introdotti il concetto di rete sociale (Sezione 2.2.1) e i concetti teorici circa la rappresentazione delle stesse (sezione 2.2.2), con l'obiettivo di fornire le basi per la comprensione e definizione di queste particolari *reti complesse* che vedono il loro sviluppo sui *Social Media*. Si passa poi alla presentazione dell'attività di *Social Network Analysis*, volta allo studio delle caratteristiche strutturali presentate da queste reti (Sezione 2.2.3).

2.2.1 Le reti sociali

Una rete sociale è una struttura sociale costituita da un insieme di “attori sociali” (e.g., utenti di una determinata piattaforma sociale) collegati tra di loro da interazioni sociali (e.g. *like*, commenti, amicizia).

Le reti sociali appartengono a quella famiglia di reti chiamate *reti complesse*, le quali modellano sistemi reali che spaziano da sistemi biologici, ai *climate network* alle reti sociali appunto. Le *reti complesse* sono reti che si differenziano nettamente dalle reti casuali e dalle reti regolari, le quali sono molto studiate in letteratura.

Le reti casuali presentano dei nodi collegati da una certa probabilità p , i nodi generalmente presentano tutti lo stesso numero di vicini; queste reti catturano qualche aspetto delle reti complesse, quindi delle reti reali.

Nelle reti regolari ogni nodo risulta essere interconnesso ad un certo numero fisso e costante di altri nodi, hanno quindi schemi regolari all'interno della struttura (il grado di casualità dei collegamenti è nullo), queste reti non somigliano affatto alle reti reali, dunque alle reti complesse.

Le reti complesse hanno caratteristiche diverse da entrambe le tipologie citate, in generale una rete complessa è una rete con caratteristiche topologiche non “banali”, ovvero caratteristiche che non si verificano in reti regolari o reti casuali, ma che spesso si verificano in reti che rappresentano sistemi reali. Inoltre, a differenza delle reti regolari, in cui ogni nodo è connesso a un numero fisso di nodi, o a reti casuali, in cui le connessioni tra i nodi

sono basate su una certa probabilità p , le reti complesse sono strutture più “sofisticate”. Tra le caratteristiche delle reti complesse troviamo:

- ***Small-world***
- ***Clustering***
- ***Forza dei legami deboli***
- ***Invarianza di scala***

Le declinazioni assunte nelle definizioni delle proprietà proposte di seguito, si rifanno a reti complesse intese come reti sociali, tutti gli esempi e spiegazioni dunque ipotizzeranno tale tipologia di rete complessa.

Small-world La proprietà denominata “*Small-world*” deriva dagli studi condotti dallo psicologo Stanley Milgram (1967), questa sancisce che tutte le reti complesse presenti in natura sono tali che, presa una coppia qualsiasi di nodi, si possa sempre trovare un percorso relativamente “corto” che colleghi le due estremità. Con l'espressione “*sei gradi di separazione*” si fa riferimento al concetto di “*Small-world*”: si ipotizza che tutte le persone siano a sei o meno legami sociali l'una dall'altra, con una catena di relazioni “*amico di amico*” di dimensioni sei, risulterebbe possibile collegare chiunque in una rete complessa.

Clustering La proprietà del “*Clustering*”, afferma che nelle reti complesse (quindi nelle reti sociali), vi è la naturale tendenza a creare raggruppamenti (cluster appunto). Tale proprietà è quantificata dal “*Coefficiente di Clustering*” (Watts et al. 1998), tale indice misura quanti degli amici di un certo nodo, risultino essere amici tra di loro; in una rete complessa tale valore è tendenzialmente alto. Sebbene prima si fosse affermato che le reti casuali, in parte, riuscissero a catturare aspetti delle reti reali, nel caso specifico di “*Coefficiente di Clustering*”, i valori che tendenzialmente le reti casuali presentano sono molto bassi, se paragonati a una rete complessa.

Forza dei legami deboli Anche questa proprietà deriva dagli studi “esterni” al mondo informatico/matematico, la proprietà della “*Forza dei legami deboli*” è stata proposta dagli studi di Granovetter (1973). Questa proprietà afferma che la forza di un determinato legame, che nelle reti sociali può essere modellato come *like*, amicizia ecc., viene definito in funzione di una combinazione di tempo, intensità emotiva, intimità (intesa come fiducia e confidenza reciproca) e dallo scambio di servizi che caratterizza il legame stesso.

Vengono quindi definiti “*legami deboli*”, quei legami presenti in reti informali di persone, gli stessi risultano (in termini di integrazione sociale) più importanti dei “*legami forti*”, quest'ultimi sono invece instaurati con persone facenti parti delle cosiddette reti primarie (famiglia, istituzioni, organizzazioni).

In particolare, la “*forza dei legami deboli*”, permetterebbe un più semplice e facile flusso dell'informazione, consentendo inoltre la nascita di collegamenti con nodi diversi rispetto al contesto sociale costituito dai “*legami forti*”, di fatto rappresentano dei “ponti” verso parti della rete inesplorate.

Invarianza di scala Si osserva la presenza di “*invarianza di scala*” quando, data una rete, la relazione tra il numero di nodi ed il numero delle loro connessioni è di tipo “esponenziale negativo” (Barabási et al. 1999). Dunque vi sono pochi nodi con un alto valore di grado e molti con un basso valore di grado. Questa situazione può essere causata dal fatto che, quando un nodo in una rete sociale vuole stabilire un nuova connessione (e.g. un *like* o un commento), sarà portato a farlo nei confronti di un nodo che ha già tanti collegamenti.

Dunque nelle reti complesse vi è la presenza di nodi *Hub*, gli stessi tendono a collegare zone della stessa che sarebbero in caso contrario separate, andando a supportare la proprietà dello *Small-world*.

Un'altra peculiarità, questa volta propria delle reti sociali, è la cosiddetta “*omofilia*”: gli individui tendono a collegarsi, dunque intrecciare relazioni con persone che condividono caratteristiche comuni alle proprie (e.g., credenze, valori, educazione, ecc.). Come precedentemente accennato, questa proprietà è una delle cause che porta alla creazione del fenomeno oggetto di studio di questa tesi.

La veloce panoramica fornita, suggerisce un’orizzonte molto ampio e con molti fattori e proprietà da tenere in considerazione nel momento in cui si intende intraprendere un’analisi su di una rete sociale.

Date le notevoli dimensioni di un sistema complesso, nella fattispecie una rete sociale, lo studio del comportamento di questo non può basarsi su un’analisi dei singoli elementi che insieme compongono il sistema, risulta necessario una valutazione dei comportamenti macroscopici chiamati *comportamenti emergenti*; si tratta di fenomeni collettivi, che si presentano spontaneamente e non in funzione di un’organizzazione centralizzata.

Le reti, siano esse regolari, casuali o complesse, vengono solitamente rappresentate mediante un *grafo*. Per questo motivo, i principali concetti formali legati alla teoria dei grafi verranno introdotti e descritti nella prossima sezione.

2.2.2 Cenni di teoria dei grafi

Data la natura del lavoro che si intende svolgere in questa tesi, i possibili dati che potrebbero essere utilizzati per svolgere le analisi e gli studi proposti devono essere necessariamente modellabili come un grafo. Un grafo G è un insieme di elementi detti nodi o vertici che possono essere collegati fra loro da linee chiamate archi, lati o spigoli; in maniera più formale diremo che un grafo G è costituito da una coppia di insiemi (V, E) dove V è detto insieme dei nodi $V = \{v_1, v_2, v_3, \dots, v_n\}$ e E è detto insieme di archi ed è un sottoinsieme del prodotto cartesiano $V \times V$, ossia un insieme di coppie (v_i, v_j) con $v_i, v_j \in V$.

Dunque i nodi possono essere visti, in generale, come elementi del dominio di interesse che interagiscono tra di loro, l’interazione tra questi è modellata mediante degli archi. Il tipo di interazione e le eventuali informazioni aggiuntive, che possono andare ad arricchire il grafo, dipendono unicamente dalla realtà di interesse modellata; possono quindi variare e non seguire regole fisse e prestabilite.

Data la definizione generale di grafo, possono essere derivate due diverse tipologie di questo; la definizione di relazione simmetrica è posta alla base della differenza. In matematica, una relazione binaria R in un insieme X è *simmetrica* se e solo se, presi due elementi qualsiasi a e b , vale che se a è in relazione con b allora anche b è in relazione con a . In forma concisa:

$$\forall a, b \in X, \quad aRb \Rightarrow bRa \quad (2.1)$$

La presenza, o assenza, di questa proprietà causa la formazione di due tipologie diverse di grafi:

- Grafi non orientati (o indiretti): un grafo $G = (V, E)$ si dice non orientato (o indiretto) se e solo se, E è una relazione simmetrica.
- Grafi orientati (o diretti): un grafo $G = (V, E)$ si dice orientato (o diretto) se e solo se, E non è una relazione simmetrica.

Dunque un grafo diretto è un grafo dove il senso dell'arco ha un significato forte, impostando quindi un ordine di percorrenza; un grafo non diretto, invece, risulta essere più “morbido”, permettendo di muoversi in qualsiasi senso, a patto che vi sia effettivamente un arco.

Un ulteriore e importante differenza, che prescinde dal tipo di orientamento del grafo, riguarda la possibilità che E sia un *multiinsieme*. Un *multiinsieme*, in matematica, e più in particolare nella combinatoria, nella logica matematica e nella teoria degli insiemi, è una generalizzazione del concetto basilare di insieme. Si potrebbe definirlo come un elenco che ammette elementi ripetuti; formalmente, un multiinsieme è definito dalla seguente coppia:

$$M = (A, m) \mid m : A \rightarrow \mathbb{N} \quad (2.2)$$

Dunque, A è un insieme e m è una funzione a valori naturali positivi; A viene detto insieme sostegno del multiinsieme, i suoi elementi si dicono elementi del multiinsieme, ed m è definita molteplicità del multiinsieme. Si può dire che la funzione molteplicità associa ad ogni elemento del multiinsieme un numero di ripetizioni che costituiscono il multiinsieme stesso.

Dunque, tornando alla possibile differenza causata dal fatto che E possa essere un multiinsieme, avremo che il grafo G sarà definito *multigrafo* se e solo se E risulti essere un *multiinsieme*. Avremo quindi, nel caso di multigrafo, la possibilità che dati due nodi questi possano essere collegati da uno o più di archi.

Un'altra possibile differenziazione dei grafi può essere dovuta alla presenza o meno dei *cappi*: un grafo G si definisce un *grafo semplice* se e solo se sprovvisto di cappi; possiamo definire cappio un arco che parte da un nodo e termina sul medesimo. Più formalmente:

$$(v_i, v_j) \in E \mid v_i = v_j \wedge v_i \in V \quad (2.3)$$

Riteniamo doveroso e necessario introdurre anche il concetto di *grado* di un nodo: si definisce *grado* di un nodo, su di un grafo non orientato, il numero di archi incidenti in quel nodo; nel caso in cui il grafo fosse orientato, al concetto di *grado* si affiancano i concetti di *grado in entrata* e di *grado in uscita*. Il grado in entrata è il numero di archi

incidenti diretti verso il nodo preso in esame, viceversa il grado di uscita è il numero di archi incidenti che partono dal nodo preso in esame.

Ulteriori e, sicuramente, più approfondite definizioni riguardanti i grafi potrebbero essere fatte; tuttavia, data la natura di questa tesi, si ritiene che quando detto possa essere sufficiente al fine di introdurre correttamente le definizioni dei grafi modellanti diverse reti; gli stessi risultano essere fondamentali in quanto sono posti alla base del lavoro e ne determinano la correttezza.

A questo punto è facile dedurre che le interazioni tra gli utenti di un determinato social media, qualsiasi siano queste interazioni, così come gli utenti stessi, possano essere modellati mediante un grafo, diventando dunque delle reti sociali.

2.2.3 *Social Network Analysis (SNA)*

Come introdotto nella sezione 2.1.4, sotto al cappello di *Social Media Analytics* vi è un'importante famiglia di analisi, che si concentra più sulle caratteristiche strutturali della rete sociale, quindi sull'investigazione mediante l'uso di reti e teoria dei grafi; questa analisi prende il nome di *Social Network Analysis (SNA)*.

Molto importante in questa fase sono quindi le caratteristiche strutturali, le quali possono arricchirsi di significato in funzione del contesto.

In generale tutte le analisi che fanno riferimento al concetto di “struttura della rete”, trovano qui il loro impiego. Esistono tre grandi famiglie di possibili metriche (2019), le stesse risultano essere:

- **Connessione:** valutazione circa le modalità con le quali i protagonisti della realtà analizzata (dunque i nodi) si collegano tra loro.
- **Distribuzione:** valutazioni circa la fluidità di movimento dell'informazione sulla rete.
- **Segmentazione:** valutazione circa le modalità di “clusterizzazione” dei componenti (dunque i nodi) della rete sociali.

Metriche di connessione

Le metriche di connessione racchiudono al loro interno diverse valutazioni circa diversi aspetti. Una proprietà molto studiata è quella dell'*omofilia*, ovvero lo studio di quella tendenza dei nodi a formare legami con nodi a loro simili, aspetto fondamentale nello studio delle camere d'eco. Si può infatti ipotizzare che gli individui nelle relazioni omofile abbiano peculiarità comuni, mentre individui non appartenenti a una certa relazione omofila no; questa proprietà è alla base di possibili situazioni di polarizzazione.

In questo senso, si trova legata allo studio della proprietà dell'*omofilia* anche lo studio della *chiusura della rete* o *chiusura triadica*, volta a definire una *trust network*: dati i nodi *A*, *B* e *C*, se *A* si “fida” del nodo *B*, con questo che si fida del nodo *C*, allora molto probabilmente *A* ha le condizioni per fidarsi di *C* e quindi si connetterà a questo (proprietà transitiva).

Nelle reti sociali la *chiusura triadica* avviene perché c'è una maggior possibilità, dati i tre nodi precedentemente introdotti, che A si connetta a C creando quindi un legame (molto probabilmente “debole”).

Viene inoltre studiata la *propinquity*, ovvero in che modo la vicinanza fisica tra i protagonisti di una rete sociale influisca sulla probabilità che questi vedano la presenza di un legame; presupposto fondamentale è la presenza di metadati circa la posizione geografica.

Un'altra proprietà, studiata da questa tipologia di metriche, prende il nome di *multiplexity*, ovvero valutazioni circa la forza del legame presente tra due nodi di una rete.

A tale proprietà si affiancano quelle di *mutualità* e *reciprocità* delle relazioni (quindi degli archi) presenti tra i nodi.

Metriche di distribuzione

Le metriche di distribuzione tengono in considerazioni analisi circa la presenza di buchi strutturali e quindi l'identificazione di quei nodi i cui legami definiscono “ponti” volti a unire cluster che in caso contrario non risulterebbero essere connessi; all'interno di queste metriche troviamo tutte le definizioni di centralità di un nodo, insieme poi a valutazioni circa la densità in funzione dei legami presenti.

A questa famiglia appartengono anche stime circa la distanza, intesa come numero minimo di legami che collegano due nodi (“sei gradi di separazione”) e analisi volte a definire la forza di un legame.

Di particolare importanza sono le valutazioni circa il concetto di “centralità”, questa proprietà può assumere diverse sfaccettature e diverse possibili declinazioni. La centralità è un concetto relativo, varia in base al contesto o allo scopo, esistono diverse misure di centralità ognuna di queste descrive un fenomeno ben preciso, coglie una caratteristica dei nodi che vengono considerati.

Ogni misura di centralità da un'informazione diversa rispetto alle altre. La definizione di “centrale” non è quindi univoca. Verranno quindi velocemente introdotte le più importanti misure di centralità presenti nelle attività di *Social Network Analysis*.

Centralità di grado All'interno di un grafo non diretto definiamo *centralità di grado* di un certo nodo, il numero di archi incidenti dato quel nodo; possiamo inoltre dire che il nodo con il grado più alto è considerabile come il più “importante”.

Questa tipologia di centralità quantifica la capacità di un nodo di comunicare con altri (quindi contagiare altri nodi); è una misura locale che guarda solamente, dato un nodo, il vicinato di primo ordine (quindi i nodi direttamente a lui collegati). Non viene quindi fornita nessuna informazione circa la capacità del nodo di raggiungere la periferia della rete (o l'estremità opposta). La formula 2.4 mostra come viene calcolata, dato un generico nodo x , appartenente a un grafo con N nodi, che osserva k archi incidenti, la centralità di grado.

$$DC(x) = \frac{k}{N - 1} \quad (2.4)$$

Nel caso di grafo orientato, si parlerà anche di *centralità di grado in uscita* (numero di archi che partono dal nodo sotto studio) e *centralità di grado in entrata* (numero di archi

diretti nel nodo sotto studio).

La *centralità di grado*, in una rete sociale, può indicare se un nodo risulti essere un “*influencer*”, l’identificazione di questi utenti può permettere tutta una serie di analisi che tengono in considerazione anche aspetti semantici: che tipo di contenuto è prodotto dal nodo “*influencer*”? Che tipo di influenza ha sul contenuto generato dai nodi collegati?

Betweenness Una delle metriche di centralità utilizzate per misurare la relazione che esiste tra un nodo e tutti gli altri nella rete, è la *Betweenness*. La struttura topologica della rete ha un impatto molto forte su questa misura. Tale metrica conta il numero di volte in cui, un determinato nodo, agisce come “ponte” lungo un path che collega due nodi. La *Betweenness* di un nodo è così calcolata:

$$B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (2.5)$$

Dove con $\sigma_{st}(v)$ si definisce il numero di cammini tra i nodi s e t , che vedono al loro interno il nodo v , invece σ_{st} è il numero totale di cammini tra s e t .

Al fine di comparare i nodi in funzione della *Betweennes*, viene effettuata un’operazione di normalizzazione volta a mappare il valore di tale centralità nel range $[0, 1]$ secondo la seguente formula:

$$B(v)' = \frac{B(v)}{[(n-1)(n-2)/2]} \quad (2.6)$$

Possiamo vedere un nodo centrale, secondo la misura di *Betweennes*, come un nodo “ponte” ovvero un nodo che risulta essere più volte parte del cammino che congiunge due nodi.

La *Betweenness* indica quei nodi che compaiono più volte nei cammini che congiungono diverse coppie di nodi, nelle reti sociali questi nodi possono rappresentare “ponti” che facilitano lo scambio di informazioni nella rete. Identificare tali attori, ipotizzando già un’attività di partizionamento, permette di comprendere le dinamiche circa le connessioni extra-comunitaria e intra-comunitarie; permettendo poi tutta una serie di analisi circa le conseguenze dovute all’eliminazioni di certi nodi: come si modificano le dinamiche interne in seguito alla cancellazione di nodi centrali per *betweenness*? E quelle esterne?

Closeness Nel caso in cui fosse necessario identificare quali nodi risultino essere in mezzo al flusso informativo, potremmo utilizzare la *Closeness*. La *Closeness Centrality* si basa sulla lunghezza della media degli *shortest-path* tra un vertice e tutti gli altri vertici del grafo:

$$C_c(v) = \left[\sum_{j=1}^N d(v, j) \right]^{-1} \quad (2.7)$$

La versione normalizzata, con l’obiettivo di rendere il valore di questa misura confrontabile, è dato dalla seguente formula:

$$C'_c(v) = C_c(v)/(N - 1) \quad (2.8)$$

Possiamo intendere la *Closeness Centrality* come un livello di efficienza, dato un certo nodo, nel diffondere informazioni a tutti gli altri elementi del grafo: più un nodo è al centro della rete, più questo avrà capacità di diffondere informazione.

Metriche di segmentazione

Le metriche di segmentazione valutano il grado di *coesione strutturale* degli attori della rete sociale, dunque valutazioni circa il numero minimo di membri che se “eliminati” dalla rete, causerebbero uno “scioglimento” della stessa, inteso come una disconnessione dei membri appartenenti alla rete.

L'utilizzo del *coefficiente di clustering* e conteggio delle *cricche* e *circoli sociali*, sono altri aspetti tenuti in considerazione da queste metriche. Dove con il termine “cricche” si intendono quelle realtà di forte interconnessione tra i membri della rete, con “circoli sociali” invece si intende quei gruppi di individui legati meno strettamente rispetto a quelli appartenenti a una cricca.

2.2.4 *Community detection*

Sebbene faccia parte della *Social Network Analysis*, merita una sezione dedicata l'attività di *Community detection*. Infatti, scoprire e studiare la struttura delle comunità che possono esistere nelle reti sociali, è un passo cruciale verso una migliore comprensione dei sistemi complessi, delle quali le reti sociali fanno parte. Attività non semplice, specie se ci si trova in contesto in cui la numerosità degli individui, facenti parte della realtà di studio, insieme ai loro collegamenti, risulti essere importante, portando poi a una difficile visualizzazione della rete stessa.

L'attività di *Community detection* è fondamentale rispetto al problema considerato; questo perché il risultato di tale *task* permetterà di definire i gruppi da cui far partire le analisi volte a quantificare le eventuali situazioni di polarizzazione.

La *Community detection*, nella sua concezione più generale, si occupa di definire dei raggruppamenti di nodi basandosi solo su informazione topologiche, dunque solo sui collegamenti che i protagonisti della rete osservano; si ipotizza che queste suddivisioni siano costituite da una grande densità di archi interni, contro una bassa densità di archi esterni, che collegano i nodi ad altre comunità.

La suddivisione in gruppi viene sempre eseguita, anche se la rete non dovesse presentare comunità e dunque il risultato risultasse essere poco “sensato” (le metriche volte a quantificare la qualità delle partizioni verranno presentate nei prossimi capitoli). In questo senso si introduce il problema noto come: *partizionamento del grafo*.

Partizionamento del grafo

L'attività di *Community detection* poggia le sue fondamenta sull'attività di partizionamento di un grafo, una definizione circa tale attività è proposta da Schulz et al. (2016):

Problema 1. *Dato un certo numero $k \in N$ con $k > 1$ e un grafo indiretto con pesi non negativi sugli archi, il problema della partizione di un grafo richiede di definire dei gruppi di nodi, tali $V_1, V_2, V_3, \dots, V_k$ dove:*

$$V_1 \cup \dots \cup V_k = V \quad (2.9)$$

$$V_i \cap V_j = \emptyset, \forall i \neq j \quad (2.10)$$

Può essere anche inserito, in funzione poi del contesto di applicazione, un vincolo circa la numerosità delle partizioni imponendo la stessa numerosità o meno delle suddivisioni identificate. Inoltre, un nodo $v \in V_i$ collegato a un nodo $w \in V_j$ con $i \neq j$, è chiamato “nodo di confine”, un arco che collega due nodi non appartenenti alla stessa comunità è chiamato “arco del taglio”. Le partizioni che si ricercano, tendenzialmente, sono tali per cui si vada a ottimizzare una determinata funzione obiettivo volta a quantificare la bontà dell’attività di partizionamento^[9].

Vi sono differenti possibilità circa la funzione da ottimizzare, nel Capitolo 4 si fa riferimento a quelle utilizzate in questo lavoro di tesi.

2.3 Rappresentare e analizzare i contenuti scambiati

In questa sezione vengono presentate e definite le principali attività che hanno a che vedere con l’analisi dell’informazione semantica, tali *task* rispondono al nome di *Sentiment analysis* e *Topic Modeling* (Sezione 2.3.1 e 2.3.2).

Queste due attività fanno parte la più generale analisi nota come *Social Content Analysis*, la quale verrà quindi introdotta alla fine di questa sezione (Sezione 2.3.3).

2.3.1 Sentiment analysis

Il task di *Sentiment analysis* ha come obiettivo quello di interpretare e classificare le emozioni presenti in un certo dato non strutturato utilizzando tecniche di analisi del testo; nel caso più semplice, lo stesso poi utilizzato in questa tesi, si classificherà il dato come positivo, negativo o neutro.

Vi sono una serie di componenti, che definiscono la realtà di studio, da tenere in considerazione al fine di poter implementare correttamente una pipeline di *Sentiment analysis*:

- **Opinion holder:** persone, organizzazioni o entità che hanno una determinata opinione su di un particolare oggetto, dato il contesto di studio.
- **Object:** modo in cui viene espressa un’opinione.
- **Opinion:** legato al *object*, è necessario definire se l’opinione è positiva o negativa.
- **Social:** modi in cui gli utenti interagiscono tra di loro.

Il problema principale riguarda la rappresentazione del dato e l’ambiguità intrinseca all’interno del linguaggio naturale.

Al fine di condurre una classificazione corretta, riducendo anche i tempi computazionali e se possibile l’ambiguità, viene sempre predisposta una pipeline di *text preprocessing*; in generale gli *steps* sono i seguenti:

- **Uniformazione del testo:** solitamente, al fine di superare la logica *case sensitive*, si setta il testo in *lowercase*.

- **Tokenizzazione:** si “rompe” il testo in *token*, questi possono essere intesi come singole parole (*uni-gram*), due parole (*bi-gram*), tra parole (*tri-gram*) e via discorrendo. Senza di questi è difficile immaginare di estrarre informazioni di alto livello dal testo. Dividere i testi in parole può essere difficile per un computer, il motivo è che certi caratteri sono dotati di delimitatori e altri no; dipende dall'applicazione. La divisione in *token* è *language-dependent*.

- **Filtering:** eliminazione di tutte quelle parole che non risultano essere utili per la classificazioni, solitamente si è in possesso di un elenco in funzione della lingua.

Spesso vengono eliminati i numeri: questo perché il numero delle differenti tipologie di *token* potrebbe crescere troppo. Può anche essere utile tenere in considerazione la frequenza di apparizione delle parole: parole troppo o troppo poco frequenti, potrebbero essere inutili.

- **Stemming e Lemmatization:** una volta ottenuti i *token* definitivi, un'ulteriore *step* è convertire ogni *token* nella forma *standard*. La *Lemmatization* è una tecnica morfologica volta a osservare la morfologia e ridurre le parole in una forma più basica, lo *Stemming* invece prevede il taglio delle parole al fine di ottenere una stessa radice.

La *Lemmatization* è complessa da implementare e non sempre fornisce aumenti delle prestazioni. Lo *Stemming* è molto più “rude”, ma utile perché con la stessa radice è possibile andare a recuperare le varie rappresentazioni.

Gli *step* elencati non sono da ritenere obbligatori, dipende molto dal contesto di applicazione ma soprattutto dal metodo di classificazione utilizzato; a tal proposito, esistono diverse metodologie di classificazione.

Approccio basato sul lessico Un approccio basato sul lessico e sui dizionari si basa fortemente sulla presenza delle parole e sul valore associato a queste: ad alcune parole sarà associato un valore positivo, ad altre un valore negativo; la somma di questi valori determinerà la positività o la negatività della frase (stesso numero di parole negative e positive denota una frase neutra).

Solitamente si è in possesso di un insieme di parole annotate, questo al fine di poter associare un valore corretto ad ogni singola parola; la qualità e l'attenzione riposta nella definizione del lessico ha, ovviamente, ripercussioni sulla qualità della classificazione.

Questo approccio è molto semplice da implementare, il *text preprocessing* può essere presente ma non è di vitale importanza.

Approcci automatici I metodi automatici, al contrario dei metodi basati sui lessici, non fanno affidamento a delle regole scritte ma a tecniche di *Machine Learning*; solitamente il task è modellato come un problema di classificazione date diverse categorie (e.g., positività, negatività, neutralità).

Nel processo di *training*, il modello viene allenato al fine di associare un particolare input a un certo output in funzione a un dataset, il quale risulta essere etichettato; dunque prerogativa fondamentale è la presenza di un dataset.

Metodologia più raffinata rispetto agli approcci basati sul lessico, in questo caso il *text preprocessing* è fondamentale al fine di eliminare ciò che può appesantire la fase di computazione e per fornire al modello un dato strutturato, quindi offrire una chiave di lettura per far “comprendere” il testo al modello, il quale potrà poi processarlo.

In particolare la fase di *tokenizzazione*, dunque la fase di estrazione delle *feature* che verranno utilizzate, è di vitale importanza e intacca direttamente la qualità del risultato.

2.3.2 *Topic Modeling*

Il *topic modeling* è un tipo di modello statistico per scoprire gli “argomenti” (*topic*) astratti di cui si parla in una raccolta di documenti (corpus di documenti) o in un singolo documento. Di fatto, mediante il *topic modeling* si cerca di determinare eventuali connessioni semantiche nascoste all’interno dei testi; il tutto con l’obiettivo di organizzare, comprendere e riassumere il dato non strutturato sul quale si applica questa metodologia.

Il *topic modeling* è stato adattato a molte tipologie di dati, in generale permette di:

- Scoprire concetti nascosti che caratterizzano il dato non strutturato.
- Annotare i documenti in funzione dei temi nascosti.
- Usare annotazioni per organizzare, riassumere, cercare e fare predizioni.

Un singolo *topic* è modellato solitamente come una lista di parole dove ogni parola è associata un valore non negativo che rappresenta la probabilità della parola di occorrere dato il relativo *topic*. Possiamo vedere un *topic* come un tema, argomento, concetto, insieme di parole che tende ad occorrere insieme.

Il numero di *topic* deve essere definito a priori, la definizione di tale parametro può non essere immediato e può necessitare di diverse prove al fine di identificare il corretto valore; in questa tesi è stata implementata una misurazione volta a validare il numero di *topic*, risulta tuttavia necessario implementare anche una valutazione più “umana” al fine di identificare il corretto numero degli stessi.

L’approccio utilizzato in questa tesi verrà discusso nelle sezioni successive.

Il task di *topic modeling* può essere inserito nella classe di attività di *Natural Language Processing (NLP)*; tutta la fase di *preprocessing* è simile a quella trattata nella sezione precedente, per tanto non verrà ripetuta.

Valutazione Topic Modeling Al fine di valutare la correttezza delle attività di *topic modeling*, esistono diverse tecniche; in ogni caso risulta fondamentale anche una valutazione “umana” del contenuto di ogni *topic*. Una delle metriche più utilizzata prende il nome di *Topic Coherence*.

La valutazione implementata prevede un approccio iterativo, basato sulla definizione di un certo range di possibili *topic*; per ogni possibile valore, dunque per ogni possibile numero di *topic*, verrà valutata la *Topic Coherence*.

Più alto sarà il valore della *Topic Coherence*, più i *topic* saranno da considerarsi effettivamente come strutture latenti nei documenti in oggetto.

L'approccio adottato in questa tesi fa riferimento al lavoro proposto nel paper Röder et all. (2015); la pipeline è composta da diversi *step*, come mostrato nella Figura 2.2.

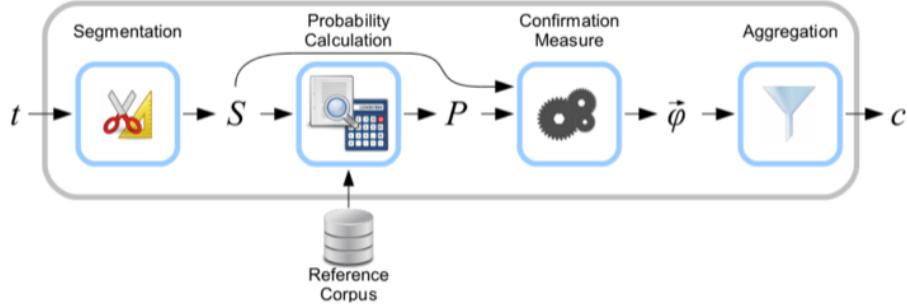


Figura 2.2: Pipeline valutazione topic coherence proposta da Röder et all. (2015)

Il *workflow* è quindi composto da quattro differenti fasi:

- Per ogni *topic* t identificato, le parole dello stesso vengono segmentate in un insieme di coppie di sottoinsiemi di parole S . Le classi di sottoinsiemi sono le seguenti:

$$S_{one}^{one} = \{(W', W^*) | W' = w_i; W^* = w_j; w_i, w_j \in W; i \neq j\} \quad (2.11)$$

$$S_{one}^{all} = \{(W', W^*) | W' = w_i; w_i \in W; W^* = W \setminus w_i\} \quad (2.12)$$

$$S_{one}^{any} = \{(W', W^*) | W' = w_i; w_i \in W; W^* = W \subseteq w_i\} \quad (2.13)$$

$$S_{any}^{any} = \{(W', W^*) | W', W^* \subset W; W' \cap W^* = \emptyset\} \quad (2.14)$$

Un possibile esempio di suddivisione, parziale, è mostrato dalla Figura 2.3.

- Vengono quindi, dati gli insiemi definiti (W' e W^*), computate le *word probability* P in funzione del corpus di riferimento.

La stima delle probabilità viene effettuata creando prima dei documenti “virtuali” mediante una *Boolean Sliding Window* di una certa dimensione, dati questi documenti viene applicata una *Boolean Document* (P_{bd}) che computa le probabilità di una parola dato il numero totale di documenti (rapporto tra il numero di documenti dove compare rispetto al numero totale di documenti).

Nel caso di più parole la probabilità congiunta è ottenuta valutando il numero di documenti che contengono entrambe le parole, dato il numero totale di documenti.

- Sia la serie di sottoinsiemi di parole S , che le probabilità calcolate P , sono utilizzate al fine di computare la validità del *topic*. Dato una singola coppia $S_i = (W', W^*)$, con le relative probabilità associate ai due insiemi, si vuole calcolare quanto, l'insieme di parole W^* , supporta W' ; questo è ottenibile mediante misurazioni dirette e non dirette, nel paper hanno utilizzato una misurazione non diretta.
- In ultimo vengono aggregate tutte le misure di validità, di tutti i *topic*, generando la *Topic Coherence* dell'insieme di *topic* identificati. Di fatto viene applicata una media aritmetica delle misure di validità ottenute.

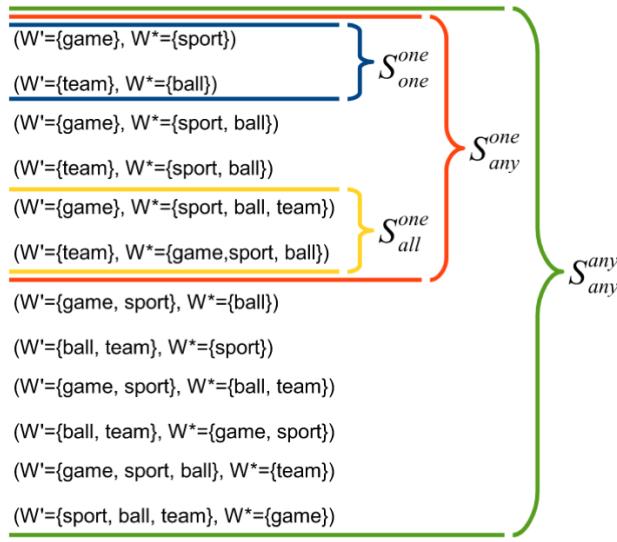


Figura 2.3: Esempio di segmentazione parziale dato il *topic* $\{game, ball, sport, team\}$
Röder et al. (2015)

2.3.3 *Social Content Analysis* (SCA)

Le due attività presentate, tali *Sentiment analysis* e *Topic Modeling*, sono racchiuse nel *task* noto come *Social Content Analysis*, lo stesso fa parte della famiglia di analisi effettuabili su reti sociali (*Social Media Analysis*).

Focus di queste analisi sono quindi il contenuto semantico scambiato e quindi generato (*UGC*), dati i protagonisti della rete sociale modellata.

I social media, grazie all'utilizzo di tecniche *SCA*, offrono un'opportunità senza precedenti per entrate in contatto diretto con i pensieri espressi dagli utenti di una determinata rete sociale, in funzione di un determinato *social media*; lo studio dell'informazione prodotta e scambiata nelle reti sociali, può essere importante per diverse ragioni, le stesse dipendono dall'entità interessata allo studio e dal tipo di metadati che arricchiscono le interazioni online tra gli utenti.

Si possono identificare diverse ragioni per le quali l'analisi dei contenuti scambiati risulta essere una possibile fonte di ricchezza:

- Permette di comprendere meglio la platea alla quale si vuole offrire (o si è già offerto) un servizio. Ad una migliore comprensione corrisponde una diminuzione della probabilità di insuccesso.
- Studiare il *feedback* su una determinata questione, permette di calibrare le attività future con l'obiettivo di comprendere ciò ha la più alta probabilità di "funzionare", in termini di servizi o pensieri.
- Lo studio delle differenze di reazione degli utenti, in funzione di un *competitors*, permette di migliorare la propria strategia di interazione.

È stato quindi mostrato come le analisi possano spaziare da valutazioni circa il sentimento dietro a determinati contenuti scambiati (e.g., rabbia, felicità, tristezza, positivo,

negativo, neutro), valutazione dei *topic* sui quali, dato l'argomento, si discute e in funzione di quelli valutare il sentimento, passando poi anche per uno studio dell'influenza sui comportamenti dei consumatori, causata dal contenuto scambiato sui *social*.

Nella fattispecie, dato lo studio proposto in questa tesi, lo studio dei contenuti scambiati dato un determinato argomento, può essere utilizzato come ulteriore metrica volta a quantificare la qualità delle attività di identificazioni delle partizioni identificate. Inoltre, l'analisi dell'informazione semantica scambiata, permette anche di capire quali aspetti, dato l'argomento che si ipotizza essere controverso, sono quelli che più vengono dibattuti nelle comunità definite.

Si ricorda che, la preziosa informazione semantica, in aggiunta a quella topologica, permette di studiare al meglio le modalità circa la diffusione dell'informazione in ambito intra-comunitario ed extra-comunitario.

Capitolo 3

Il problema delle camere d'eco: Definizione e Stato dell'arte

In questo capitolo si presenta la problematica centrale, la quale risulta essere oggetto di studio in questo lavoro di tesi: le camere d'eco. Il capitolo risulta così suddiviso:

- Nella Sezione 3.1 si procede a una presentazione del fenomeno, con una valutazione ad alto livello circa i possibili approcci volti al superamento dello stesso.
- Nella Sezione 3.2, si presenta lo stato dell'arte, descrivendo tutti quei lavori consultati che hanno fornito la possibilità di delineare la *pipeline* risolutiva che in questa tesi viene proposta; fornendo infine una panoramica sulla soluzione proposta, la quale verrà successivamente presentata e approfondita nel Capitolo 4.

3.1 Il fenomeno delle camere d'eco

In questa sezione è presentato il fenomeno noto con il termine camere d'eco. Viene fornita un'introduzione circa tale problematica (Sezione 3.1.1), per poi passare a una presentazione più dettagliata della stessa (Sezione 3.1.2). In ultimo vengono valutati, ad alto livello, i possibili approcci volti alla gestione di tale fenomeno (Sezione 3.1.3).

3.1.1 Introduzione

La mancanza di un costo e di regole chiare e condivise circa la diffusione online delle informazioni, unita all'effettiva mancanza di un'autorità centrale e autoritaria volta a valutare la qualità delle stesse, crea un "effetto livellamento" che pone molte informazioni su un livello di qualità apparentemente uguale, quando invece possono registrare un livello di accuratezza molto diverso.

Con gli ultimi sviluppi tecnologici, la digitalizzazione e le interazioni online, l'accesso e la condivisione di informazioni è reso sempre più semplice ed immediato, permettendo alle persone di essere soggette, dato un certo argomento, a opinioni diverse e anche contrastanti tra loro. Questa opportunità dovrebbe poter permettere di migliorare il processo decisionale quotidiano di ognuno, con l'obiettivo (forse utopistico) di una "democrazia deliberativa".

Tuttavia, data la teoria dell'esposizione selettiva, già accennata precedentemente e formalizzata dai lavori di Dieter (1986), si verifica la naturale tendenza delle persone a favorire le informazioni di conferma e ad evitare informazioni dissonanti. Si evidenzia come, data la facilitazione del controllo circa la propria “dieta informativa”, i *social media* possano facilitare la creazione delle cosiddette “camere d'eco”, nella quali le persone sono perennemente esposte solo ad opinioni a loro gradevoli e consonanti, comportando così una chiusura mentale a tutto ciò che è diverso o che si oppone alla ideologia centrale. Tutto questo porta a una crescente frammentazione sociale e polarizzazione ideologica nella nostra società.

La preoccupazione circa la questione, ha motivato la ricerca volta a trovare il modo di incoraggiare le persone a cercare informazioni, che mettano in discussione le loro opinioni nei più disparati campi. Ora più che mai è quindi necessario studiare tale fenomeno, al fine di poter migliorare l'esperienza e la sicurezza online degli utenti, cercando poi di minimizzare la diffusione di informazioni e notizie false o non accurate.

3.1.2 Le camere d'eco

Una possibile definizione accurata di camera d'eco è fornita dal lavoro di ricerca effettuato da Bruns (2017), lo stesso definisce una camera d'eco come: “*una realtà dove un gruppo di individui sceglie di connettersi e interagire preferenzialmente tra loro, escludendo gli estranei. Più questa realtà risulta coesa (cioè, più connessioni vengono create all'interno del gruppo e più connessioni con gli estranei vengono interrotte), più il gruppo è isolato dall'introduzione di punti di vista esterni, mentre i punti di vista dei suoi membri sono in grado di circolare ampiamente al suo interno.*

In tutti gli studi analizzati, volti all'analisi del fenomeno in oggetto, la suddivisione della rete sociale rappresentate un determinato *topic* controverso di discussione, produce solitamente due differenti *side* contrapposti. Tale modellazione è giustificata in quanto, tendenzialmente, intorno a un qualsiasi argomento controverso si definiscono sempre due pensieri estremi e contrapposti. Inoltre, data una grande discussione su di un determinato *topic*, ogni eventuale nuovo utente che si inserisce nella stessa, tende ad affiancarsi a una determinata posizione piuttosto che a crearne una nuova. Questo discorso è da intendere come generalizzazione, risulta triviale comprendere come, all'interno di una delle due partizioni iniziali, esistano diverse correnti e “sotto-pensieri” che però condividono un base comune.

Lo studio delle camere d'eco può essere valutato da diversi punti di vista e da diverse profondità; il tipo di esposizione a una certa informazione, il modo in cui ci si pone nei confronti dell'informazione, le credenze personali e il grado presunto (o conosciuto) di accuratezza dell'informazione consumata, giocano un ruolo decisivo nella possibile formazione delle stesse. Lo studio proposto da Liao et al. (2014) si pone come obiettivo la valutazione dell'esposizione e della scelte dell'informazione, da parte di un certo insieme di utenti, fornendo insieme all'informazione stessa indici di competenza (intesa come autorevolezza della fonte) e di posizione effettiva, dato il *topic*, della fonte.

Vengono quindi identificati diversi argomenti controversi, in funzione degli stessi vengono

selezionati diversi commenti opposti per ognuno, dotandoli poi di un livello di competenza e una posizione personale dell'autore dello stesso (risultato ottenuto mediante una collezione e categorizzazione manuale, ben documentata nel lavoro stesso); viene quindi definito un campione di utenti ai quali viene effettuato uno screening, dati i *topic* che saranno poi oggetto dell'analisi, sulla posizione circa gli stessi.

Dai risultati di questo lavoro emerge come l'indicatore di competenza, tenda a far sì che l'utente sia più indirizzato verso la scelta dell'informazione con un maggior livello del suddetto indice, rispetto a livelli più bassi dello stesso; l'utilizzo combinato degli indici di posizione personale e di competenza dell'autore del commento, rendono ancora più evidente tale tendenza.

Si è anche verificato come l'indicatore di “autorevolezza” abbia anche avuto un effetto evidente sul miglioramento della competenza percepita, da parte degli utenti, delle fonti indicate come esperte e sulla diminuzione di quella delle non esperte, anche se il contenuto informativo dei due non differiva in termini di qualità; emerge però che, nonostante il livello identico di “autorevolezza” di commenti opposti come contenuto, le fonti consonanti a quelle dell'utente osservato, vengano costantemente percepite come più esperte e ricevano un maggiore consenso rispetto a quelle dissonanti.

Viene quindi evidenziato come la creazione delle camere d'eco sia causata, in misura importante, da una tendenza umana all'esposizione selettiva delle informazioni; inoltre, in uno studio proposto da Sasahara et al. (2019), si sottolinea anche come altri fattori umani identificati come la omofilia (la tendenza di legare con persone simili) e la influenza sociale (la tendenza a diventare simili ad altre persone in seguito a interazione sociali) tendono a guidare verso una polarizzazione delle opinioni e a una segregazione sociale, elementi alla base di una camera d'eco.

Agire quindi a un livello così profondo non è semplice e non è compito degli algoritmi di raccomandazione di contenuti dei *social media*, ma è compito della società la quale deve educare all'apertura mentale del singolo, facendo passare per positiva la tendenza, giustificata, al cambio di posizione ed idee; in parallelo dovrebbe anche educare a una ricerca responsabile e curata, delle fonti dalle quale si attinge. L'inserimento però di un indice di “autorevolezza” dati i contenuti, come suggerito dallo studio di Liao et al. (2014), può essere un valido suggerimento al fine di migliorare la qualità di informazione circolante sui *social network*; l'utilizzo quindi di diversi indici potrebbe essere un modo per mitigare la problematica circa “l'effetto livellamento” presentato all'inizio di questa sezione.

3.1.3 Possibili approcci al problema

L'identificazione e il superamento, o l'eventuale prevenzione, di situazioni polarizzate può essere un'attività tutt'altro che banale. Le attività di definizione del modello che struttura e formalizza le relazioni e interazioni tra gli elementi del dominio di interesse è di fondamentale importanza, al fine di non introdurre *bias* che può andare a pregiudicare i risultati e le conclusioni, fornendo quindi una visione distorta, se non del tutto scorretta, della realtà analizzata. La precedente introduzioni delle reti sociali, particolari rete complesse, fornisce già una panoramica della difficoltà circa l'approccio a questa tipologia di modellazione.

Ipotizzando dunque una certa modellazione attuata mediante un grafo, rappresentante quindi una rete sociale, diverse possibilità possono essere tenute in considerazione al fine di mitigare la polarizzazione, in funzione di determinati contenuti, degli utenti; difatti risulta possibile intervenire lavorando direttamente sulle *bolle di filtraggio*, ponendosi a monte del problema o lavorando sulle *camere d'eco*, ponendosi quindi a valle del problema.

Bolle di filtraggio Come accennato nel capitolo precedente, i motori di raccomandazione sono ormai presenti in quasi tutti i servizi online. Mediante l'utilizzo di questi, risulta possibile adattare la fornitura di contenuti all'utente, nel tentativo di fornire le risorse che più probabilmente andrà a “*matchare*” meglio con lo stesso, in base a tutta una serie di informazioni raccolte.

Questo può portare a situazione di polarizzazione in quanto, se l'utente stesso non si interessa ad accingere conoscenza da altre fonti, riceverà sempre lo stesso tipo di informazioni. Già con queste poche righe si comprende che tale situazione porta inevitabilmente alla creazione di camere d'eco.

Tuttavia risulta possibile prevenire situazioni di polarizzazione andando a lavorare su tali *bolle di filtraggio* (generate cioè dagli algoritmi di *filtering*), dunque sul come gli utenti risultino essere esposti all'informazione, dato un determinato *topic* controverso. L'obiettivo è quello di fornire un'esposizione *bipartisan* che, ipotizzando la presenza di due punti di vista opposti, permetta ad ognuno di essere soggetto a contenuto proveniente da entrambi.

In questo senso si può ragionare sulla modifica dei sistemi di raccomandazione, nell'ottica di definirne alcuni volti a fornire materiale opposto a quello che, dato l'utente in studio, risulta essere normalmente esposto.

Questo approccio è sicuramente più complicato del precedente, difatti risulta necessario osservare continuamente l'evoluzione della rete sociale, in modo tale da verificare che le diversificazione di raccomandazioni effettuate, portino effettivamente a situazioni non polarizzate. Ad ogni modo, l'attuazione di soluzioni basate su questa logica, sono forse le più promettenti in termini di riduzione di probabilità circa il pericolo di polarizzazioni.

Camere d'eco In questo caso si presuppone di porsi in una possibile situazione di “fatto compiuto”, dove la polarizzazione sia già avvenuta o comunque sia in atto. Bisognerà quindi identificare puntualmente e accuratamente le camere d'eco, tale attività è paragonabile all'attività di *Community detection*, alla quale è tra le altre cose legata in diversi punti e idee di base; la principale problematica è la validazione puntale dei risultati che, per ovvie ragioni, non può passare da una “*true table*”, l'utilizzo di metriche è dunque fondamentale al fine di assicurare che le attività abbiano prodotto risultati validi, dunque partizioni opposte e polarizzate.

Le metriche utilizzabili hanno come obiettivo la quantificazione della “controversia”, la stessa può essere intesa in diversi modi, in generale può essere vista come una misura di “segregazione” che le partizioni identificate attuano vicendevolmente.

Inoltre l'utilizzo di tecniche di *Social Media Analytics*, può costituire ulteriori metriche volte a verificare l'effettiva contrapposizione delle partizioni identificate.

Una volta identificate le partizioni contrapposte e appurato che vi sia una situazione polarizzata che oppone due visioni, sarà necessario studiare strategie volte a superare tali polarizzazioni. L'attività principe in questa fase può essere la *Link prediction*, questo task ha come obiettivo il proporre dei collegamenti tra determinati elementi, i quali non risultavano prima direttamente collegati. Solitamente viene utilizzata una qualche metrica, volta a quantificare un determinato score del potenziale arco, in funzione di questo score possono essere adottate logiche volte all'effettivo inserimento dell'arco.

Quindi si procede a porre diversi collegamenti artificiali con l'obiettivo di superare le polarizzazioni e abbassare il livello di controversia identificato; risulta ovviamente necessario cercare di aggiungere il minor numero possibile di archi, in quanto è triviale che un grafo totalmente connesso sarà privo di contrapposizioni, risulta egualmente triviale che tale situazione, in una rete sociale, è praticamente impossibile.

L'approccio basato sulla gestione delle *camere d'eco* è forse il più semplice tra i due, tuttavia è forse quello più strettamente teorico, che quindi può scontrarsi più duramente con la realtà, minimizzando poi la capacità di risoluzione proposta.

Risulta, fin da subito, importante sottolineare che l'approccio scelto per questo lavoro di tesi è il primo. Dunque si lavorerà direttamente sulle camere d'eco, con l'obiettivo di minimizzare la controversia e quindi superare le polarizzazioni.

3.2 Stato dell'arte

In questa sezione vengono descritti i principali lavori di letteratura, che si occupano del problema affrontato in questa tesi da diversi punti di vista; alcuni lavori presentati sono stati utilizzati come guida e punto di riferimento per questa tesi, altri hanno permesso di ampliare la conoscenza e definire meglio il contesto, fornendo una consapevolezza maggiore circa il problema posto.

Si presentano prima tutti quei lavori volti a identificare le camere d'eco, investigando la possibile presenza di polarizzazioni (Sezione 3.2.1), in seguito si valutano le tecniche volte al superamento di tali situazioni (Sezione 3.2.2) per terminare con una veloce introduzione circa la soluzione proposta (Sezione 3.2.3).

3.2.1 Identificazione delle camere d'eco

L'identificazione delle camere d'eco è un campo di ricerca molto recente, questo è dovuto al fatto che solo negli ultimi anni i *social media* sono diventati così capillari e utilizzati, conseguenzialmente il mercato creato da queste tecnologie si è ampliato sempre di più comportando un interesse, da parte dei provider dei vari servizi, a raggiungere più persone possibili.

A causa di questa situazione il panorama di lavori circa l'identificazione delle camere d'eco risulta essere ancora poco esplorato, tale tendenza sembra però che si stia ribaltando a causa della digitalizzazione crescente della società.



Figura 3.1: Diversi stage proposti da *Garimella et al. (2017)*

Ad ogni modo, al fine di verificare l'effettiva presenza di camere d'eco, risulta possibile definire una serie di “*step*” che insieme vanno a formare una pipeline di identificazione. I possibili diversi passi vengono mostrati in Figura 3.1. In particolare, essi riguardano:

- **Costruzione del grafo:** questa fase è forse la più importante in quanto pregiudica, in quanto primo stage, tutte le attività future. La modellazione del dato, al fine di definire un grafo, può seguire diversi paradigmi e basarsi su diversi aspetti e fattori (e.g. *retweet*, *tag*, risposte, commenti, *followee/follower*). In questa fase si va quindi a definire la rete sociale rappresentante le interazione tra gli utenti, qualsiasi queste siano.
- **Suddivisione del grafo:** la suddivisione del grafo che non è nient'altro che l'attività di *Community Detection*, con il vincolo relativo alla numerosità delle comunità. Una volta definite le due partizioni è necessario comprendere se vi siano effettivamente polarizzazioni, compito questo del terzo *step*.
- **Misurazione controversia:** Ultima fase della *pipeline*, la misurazione della controversia ha come obiettivo verificare se le due comunità identificate possano essere definite camere d'eco o meno. Si tratta dunque di valutazioni circa il livello di polarizzazione.

Costruzione del Grafo

La modellazione della rete sociale, che porta quindi alla costruzione di un grafo che definisce la stessa, può essere un'attività tutt'altro che banale. Difatti, il tipo di dato di cui si è in possesso, potrebbe non permettere di modellare relazioni che possano notificare l'effettiva presenza di controversia.

A tal proposito, il lavoro proposto da Garimella et al. (2017), propone una modellazione di una rete sociale mediante un grafo $G = (V, E)$ non orientato; partendo da dati provenienti da *Twitter* riguardanti diversi *topic*, vengono generati diverse reti sociali in funzione del *topic* disquisito. La raccolta di questi dati è stata resa possibile dagli *hashtag* utilizzati autonomamente dagli utenti del suddetto *social*.

In particolare, nel lavoro proposto, gli autori si sono concentrati sui dati facenti riferimento ai *retweet* e alla relazione *followee/follower*; i primi tipicamente, stando a quanto gli autori dichiarano, indicano un “*endorsement*” di un utente nei confronti dell'opinione dell'altro, il secondo invece modella le relazioni di interesse a seguire i contenuti tra due utenti; la definizione di quest'ultimo grafo è “*topic-based*” (vengono tenute in considerazione solo le relazioni tra gli utenti i quali hanno espresso un *tweet* con l'*hashtag* relativo al *topic* di interesse).

L'assunzione di base è che utenti collegati da un arco, indipendentemente dal tipo di relazione, saranno anche più vicini per ciò che concerne il loro punto di vista su un

determinato *topic*

Di fatto vengono proposte, per ogni singolo *topic*, due modellazioni di reti sociali basate su due diverse tipologie di relazione. A detta degli autori, entrambi i grafi possono notificare la presenza di polarizzazioni, con lo stesso livello di efficacia. Dati i *topic* modellati, gli autori erano già a conoscenza della possibile presenza di polarizzazione o meno.

Dunque, ricapitolando, i grafi definiti da Garimella et al. (2017) vedono come nodi gli utenti e come archi o un *retweet* o una relazione di *followee/follower*. Più formalmente: viene definito un grafo $G = (V, E)$ dove V è l'insieme di utenti della rete sociale, mentre $e = (v_i, v_j) \in E$ può modellare una relazione di *retweet* pervenuta tra i due nodi, o una relazione di *followee/follower* vigente tra i due nodi, il tutto in funzione del *topic* modellato.

Altri lavori, volti sempre alla strutturazione di reti sociali nel contesto delle camere d'eco, sfruttano in maniera più profonda la strutturazione e la tipologia di interazioni tra gli utenti.

Nel lavoro proposto da Coletto et al. (2017), si propone la definizione, per ogni *topic*, di tre diversi grafi. I dati raccolti provengono da diverse pagine facenti riferimento a diversi domini (news, politica, gossip, intrattenimento), ancora una volta la provenienza degli stessi è da ricondurre a *Twitter*, questa è una scelta naturale data la facilità di reperimento degli stessi e la possibilità di strutturare in maniera semplice il dato grezzo.

Una volta identificate le pagine, per ogni *tweet* raccolto in ognuna (definito *root post*), viene ricostruito il *thread* di discussione andando a effettuare un'operazione di *crawling* circa le risposte al *tweet* iniziale. Il valore k identificante la “profondità” del numero di commenti, data la discussione, può assumere valori pari a 2, 3 e 10 (incluso l'autore del post). I dati vengono manualmente valutati come possibili controversi o meno; la logica di etichettamento manuale, prevedeva che un *topic* venisse ritenuto controverso nel caso in cui, dato il contesto, l'opinione espressa risultasse essere personale (un'idea o un'opinione).

In seguito alla raccolta ed etichettamento dei dati, gli autori provvedono a definire diverse strutture logiche utilizzate successivamente:

- ***user graph***: $G = (U, E)$ dove U è l'insieme di utenti e un arco $e = (u_i, u_j) \in E$ indica che l'utente u_i segue l'utente u_j .
- ***reply tree***: $T = (C, R)$ dove C è l'insieme dei *tweet* dato un *thread* e un arco $r = (c_i, c_j) \in R$ indica che il *tweet* c_i è in risposta al *tweet* c_j .
- ***reply graph***: proiettando l'albero T sul grafo G si ottiene un grafo $R = (U, I)$ dove un arco $e = (u_i, u_j) \in I$ indica che l'utente u_i ha risposto a un *tweet* dell'utente u_j .

Anche un altro lavoro, proposto da Kumar et al. (2018), vede la modellazione della rete sociale mediante l'utilizzo di un *reply tree* che coinvolge, dato un *post*, tutte le interazioni presenti.

Partizionamento del Grafo

Secondo passo della pipeline mostrata in Figura 3.1, la suddivisione del grafo nient'altro è che l'attività di *Community detection*. Un'introduzione al *task* è fornito nel lavoro di

Lancichinetti et al. (2010) nel quale si sottolinea l'importanza, negli ultimi 10 anni, della cosiddetta “scienza delle reti” nel campo dei sistemi complessi; infatti, molti sistemi complessi (e.g. le reti sociali), possono essere rappresentati come reti (i.e. quindi grafi), dove le parti elementari di un sistema e le loro mutue interazioni sono rispettivamente nodi e archi.

I sistemi complessi sono generalmente organizzati in compartimenti, che hanno un proprio ruolo e/o funzione. Nella rappresentazione della rete, tali compartimenti appaiono come insiemi di nodi con un'alta densità di collegamenti interni, mentre presentano una densità relativamente inferiore con gli altri compartimenti.

Questi sottografi sono chiamati comunità. La ricerca di compartimenti può far luce sull'organizzazione di sistemi complessi e sulla loro funzione.

Sono stati sviluppati molti algoritmi volti all'identificazione delle comunità, si è fatto utilizzo di strumenti e tecniche provenienti da discipline come la fisica, la biologia, la matematica applicata, l'informatica e scienze sociali, con l'obiettivo ultimo di fornire una metodologia unica volta all'identificazione puntuale delle comunità all'interno di una rete complessa; nella stragrande maggioranza dei casi, gli algoritmi utilizzati vedono l'utilizzo esclusivo dell'informazione topologica.

Ad ogni modo, non è ancora chiaro quali algoritmi siano completamente affidabili e debbano essere utilizzati nelle diverse applicazioni. La problematica sottolineata nel lavoro di Lancichinetti et al. (2010) è che la maggior parte dei test sporadici tendenzialmente eseguiti, al fine di valutare un algoritmo di *Community detection*, coinvolgono piccole reti con struttura di comunità nota o grafi artificiali strutturalmente semplici, questo è una situazione molto rara nei sistemi complessi reali, dunque poco utile al fine di comprendere quanto un algoritmo possa essere puntuale nell'individuazione delle *community*.

Un altro importante aspetto da tenere in considerazione è lo sforzo computazionale richiesto in questa attività, il quale può essere importante.

Come si diceva poc'anzi, gli algoritmi più tradizionali di *Community detection* tengono in considerazione quasi esclusivamente caratteristiche topologiche, data la rete sulla quale vengono applicati. Un grafo modellante una rete sociale, rappresenta mediante i vertici gli utenti del contesto e tramite gli archi i collegamenti e interazioni che sussistono tra gli stessi.

Come affermato nello studio proposto da Natarajan et al. (2013), un collegamento tra due utenti aumenta le possibilità che i due condividano interessi comuni ma non necessariamente lo implica. Inoltre, non è detto che due utenti che non condividono un collegamento non abbiano nessun interesse comune.

Pertanto, per determinare le comunità in maniera ancora più puntuale nell'ambito delle reti sociali, sta diventando sempre più fondamentale sfruttare sia la struttura dei collegamenti della rete (cioè che i tradizionali algoritmi di *Community detection* fanno) che la tendenza di consumo e produzione di contenuti dati i vari utenti.

Si andranno quindi a proporre anche lavori che tentano di sfruttare l'informazione semantica nell'attività di *Community detection*, in modo da migliorare le performance degli algoritmi di partizionamento del grafo.

Approccio tradizionale di *Community detection* Come si diceva precedentemente, gli approcci più tradizionali di *Community detection* vedono l'utilizzo esclusivo dell'informazione topologica presente nella rete, tenendo in considerazione solo di come i protagonisti della rete interagiscono tra di loro.

In generale esistono diversi approcci, basati esclusivamente su aspetti topologici, che si possono trovare alla base dei tradizionali algoritmi di *Community detection*. Verranno qui esposti quelli più importanti.

Node Centric Ogni nodo, per appartenere a una determinato gruppo, deve essere dotato di determinate proprietà. Queste proprietà possono riguardare concetti come, ad esempio, la *completa mutualità* (un nodo apparterrà a una *community* se è collegato con tutti gli altri nodi della stessa) o la *raggiungibilità dei membri* (un nodo per appartenere a una *community* ha vincoli circa la distanza da ogni altro nodo della stessa).

Un approccio *Node Centric* è preferibile per reti di piccole dimensioni, dunque non andrebbe bene per le reti sociali. Un esempio di algoritmo basato su logiche *Node Centric* è chiamato *k-clique* (Palla et al. 2005), questo algoritmo valuta proprietà circa la *raggiungibilità dei membri* con il fine di definire le comunità.

Group Centric Una comunità, per definirsi tale, deve nel suo complesso soddisfare un qualche tipo di proprietà, come ad esempio un certo livello di *densità*. Gli algoritmi basati su questo paradigma, modellano un problema di ottimizzazioni in funzione della proprietà che il gruppo, nella sua interezza, deve rispettare.

Anche in questo caso, l'approccio *Group Centric* è più indicato per grafi di modeste dimensioni.

Network Centric In questo caso vengono tenute in considerazione le connessioni dei nodi a livello globale, dunque tenendo in considerazione la rete nella sua interezza.

In questo contesto si vede l'utilizzo della matrice di adiacenza, ovvero una matrice $m \times m$ dove m è il numero di nodi della rete; ogni riga i rappresenta un nodo, una generica cella (i, j) conterrà il valore 1 (o più in generale il peso del collegamento) se il nodo i è collegato con il nodo j , 0 in caso contrario.

Si fornisce quindi una valutazione vettoriale di ogni nodo.

In questo senso si può valutare la *node similarity* tra i nodi al fine di identificare una *community*, si valuta quindi la similarità vettoriale mediante l'utilizzo di metriche come la *cosine similarity*.

O ancora approcci basati sullo *Spectral Clustering* o sulla massimizzazione della modularità (la *modularity* verrà introdotta e approfondita nel corso della tesi).

Un esempio di algoritmo basato sulle logiche *Group Centric* può essere il *Modularity Maximization* (Clauset et al. 2004).

Hierarchy Centric Gli algoritmi basati su un approccio *hierarchy centric* cercano di definire un'organizzazione gerarchica all'interno della rete.

In funzione dell'algoritmo si può procedere, con un approccio divisivo, dall'insieme totale dei nodi (quindi da una *community* unica), suddividendo via via in partizioni sempre più piccole. In caso opposto, si procede partendo dai singoli nodi (tante comunità quanti sono i nodi della rete) con l'obiettivo di unirli mediante un approccio agglomerativo.

Viene di fatto creato un dendogramma che subirà un taglio orizzontale a una certa altezza, in funzione di questa verranno definite le comunità; un possibile esempio di questa procedura è rappresentata nella Figura 3.2.

Un esempio di algoritmo basato su un approccio *Hierarchy Centric* è costituito dall'algoritmo *Girvan-Newman* (Girvan et al. 2002), in particolare si tratta di un approccio *Hierarchy Centric* con logica divisiva.

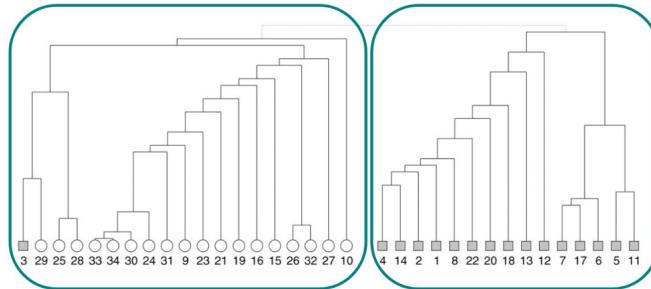


Figura 3.2: Esempio di dendrogramma con taglio volto a identificare le due partizioni ottenute Fersini (2019).

Considerare il sentimento nelle attività di *Community detection* Si ipotizza che la valutazione del *sentiment* degli utenti, in funzione di quanto condividono direttamente nella rete, possa indirizzi in maniera più puntale l'identificazione di comunità contrapposte, quindi a una realtà dotata di controversia con conseguente caratterizzazione di camere d'eco.

In questo frangente il lavoro proposto da Sawhney et al. (2017) ha fornito un valido punto di partenza al fine di strutturare questo tipo di informazione. In questo lavoro si cerca di modellare come peso la differenza tra il *sentiment* medio che viene attribuito a un utente; partendo dalle *reviews* contenute nel Yelp Dataset¹, si valutava la media (espressa con un punteggio da 1 a 5) del sentimento di ogni utente date le reviews a lui attribuite; dato poi il grafo $G = (V, E)$, con $e = (v_i, v_j) \in E$ a rappresentare la relazione di amicizia tra i due utenti v_i e v_j , si andava ad arricchire e con un nuovo peso dato dalla seguente formula:

$$weight_s(i, j) = scaling(5 - ||s_i - s_j||) \quad (3.1)$$

Dove s_j e s_i sono i sentimenti attribuiti ai due utenti i e j , il termine “*scaling*” indica la successiva attività di *rescaling* effettuata dagli autori volta a non disturbare le proprietà

¹<https://www.yelp.com/dataset>

strutturali.

Data questa modellazione appare chiaro che utenti *sentiment* simili verranno premiati con un peso più alto, utenti con *sentiment* diverso invece non vedranno particolari aumenti di peso dell'arco. Dopo aver definito questi nuovi pesi, vengono utilizzati i tradizionali approcci di *Community detection*, indicando il valore di *weight_s* come peso dell'arco, imponendo quindi all'algoritmo utilizzato, di calibrare il suo operato su questo nuovo valore.

I risultati ottenuti con questa modellazione sono poco più bassi di un approccio tradizionale, dunque senza l'aggiunta di informazioni semantiche.

Considerare i *topic* nelle attività di *Community detection* Anche in questo caso, il lavoro proposto da Sawhney et al. (2017), si pone come obiettivo l'inserimento di informazioni circa i *topic* discussi nell'attività di *Community detection*. La modellazione è molto simile a quanto mostrato precedentemente, sempre per questo lavoro, per ciò che concerne il *sentiment*; i dataset di partenza è infatti il medesimo.

Gli autori in primo luogo, mediante il modello LDA, identificavano i *topic* presenti date le recensioni espresse dagli utenti della rete. In seguito all'identificazione degli stessi, nella fattispecie 21 differenti *topic*, procedevano con la definizione di un nuovo peso volto a decorare i collegamenti presenti tra gli utenti; il valore di tale peso è il seguente:

$$weight_{LDA}(i, j) = scaling(21 - ||topic_i - topic_j||) \quad (3.2)$$

Lo *scaling*, in maniera del tutto analoga a quanto mostrato per la modellazione del *sentiment*, è reso necessario per non disturbare la struttura della rete. Tale modellazione ha l'intento di premiare, donando un valore più alto a *weight_{LDA}*, gli utenti che disquisiscono degli stessi argomenti nelle loro *reviews*.

Ancora una volta si predilige una modellazione volta a fornire un nuovo peso su cui calibrare le attività di *Community Detection*, le quali possono quindi essere effettuate con i tradizionali algoritmi che accettino un peso volto a quantificare l'importanza di un collegamento.

Un altro studio, volto a fornire un *topic* model con l'obiettivo di identificare, facendo leva sia sul contenuto, quindi sui contenuti generati dagli utenti, che sulla struttura di una rete sociale, è proposto da Natarajan et al. (2013), in particolare si pone di lavorare in reti sociali ottenute dalla modellazioni delle interazioni tra gli utenti su *twitter*. La rappresentazione più immediata del modello avviene, per ovvie ragioni, mediante la rappresentazione a piattini dello stesso proposta nella Figura 3.3.

Dato un *social network* $\langle U, S, D \rangle$, dove con U si denotano l'insieme degli utenti numerati, con $S \subseteq U \times U$ le relazioni di *followee* tra gli utenti (gli archi sono orientati) e con D un insieme di documenti (con D_u si indicano i documenti condivisi dall'utente u , i.e. i post a lui attribuiti).

Per ogni documento $d \in D_u$, il modello scegliere una certa *community* k data la distribuzione sulle comunità δ_u . Si passa poi alla generazione di un *topic* z data la distribuzione θ_k e infine la parola in d da ϕ_z . In maniera del tutto analoga, per ogni utente v nell'insieme L_u , il quale contiene tutti gli utenti che u segue, il modello sceglie una comunità da δ_u per poi, data la distribuzione ψ_k sugli utenti della *community*, scegliere v . Gli iperparametri ν, μ, α, β sono utilizzati per parametrizzare le distribuzioni per le variabili latenti $\delta_u, \psi_k, \theta_k, \phi_z$ rispettivamente. Tutti i nodi “segnati” rappresentano variabili osservate.

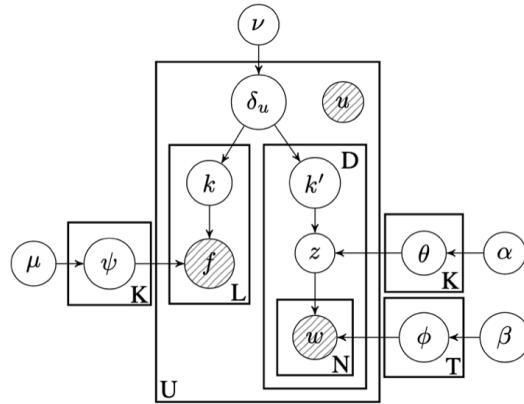


Figura 3.3: Rappresentazione a piattini del modello proposto da Natarajan et al. (2013)

L'attività di inferenza viene effettuata mediante il *Gibbs sampling* (2012), la sua accuratezza, efficienza e semplicità hanno guidato la scelta degli autori; lo pseudocodice fornito dagli autori è mostrato in Figura 3.4.

```

1 /*Initialize*/
2 for u ∈ U do
3   for v ∈ Lu do
4     k ~ uniform[1 . . K]
5     | Assign community k to link u → v
6   for d ∈ Du do
7     k ~ uniform[1 . . K]
8     z ~ uniform[1 . . T]
9     | Assign community-topic pair ⟨k, z⟩ to d
10 /* Burn-in */
11 I ← number of burn-in iterations
12 i ← 0
13 while i < I do
14   for u ∈ U do
15     for v ∈ Lu do
16       k ~  $\frac{M_{uk}^- + F_{uk}^- + \nu_k}{M_{u^-} + F_{u^-} + \sum_{k'} \nu_{k'}} \frac{F_{kv}^- + \mu_v}{F_{k^-} + \sum_{u'} \mu_{u'}}$ 
17       | Assign community k to link u → v
18     for d ∈ D do
19       ⟨k, z⟩ ~  $\frac{M_{uk}^- + F_{uk}^- + \nu_k}{M_{u^-} + F_{u^-} + \sum_{k'} \nu_{k'}} \frac{M_{kz}^- + \alpha_z}{M_{k^-} + \sum_{z'} \alpha_{z'}}$ 
20        $\prod_{w \in d} \prod_{i=1}^{n_{dw}} \frac{N_{zw}^- + i - 1 + \beta_w}{N_{zw}^- + \sum_{w' < w} n_{dw'} + i - 1 + \sum_{w'} \beta_{w'}}$ 
21       | Assign community-topic pair ⟨k, z⟩ to d
22   i = i + 1
23 /* Collect samples */
24 A ← number of samples to be collected
25 B ← number of iterations to be skipped
26 for i ∈ 1 . . A do
27   for j ∈ 1 . . B do
28     | Sample ⟨k, z⟩, ∀d ∈ D and k, ∀v ∈ Lu, u ∈ U
29   Collect sample

```

Figura 3.4: Gibbs sampling dato il modello Link-Content proposto da Natarajan et al. (2013)

I risultati ottenuti dall'approccio probabilistico, in termine di *community* identificate dati i *topic* scoperti, sono significativamente importanti, tenendo in considerazione che la definizione delle *community* e dei *topic* può essere, dato il modello, effettuata simultaneamente.

Un altro aspetto molto importante da tenere in considerazione, trasversale alle differenti tipologie di *Community detection* tradizionali, riguarda la possibilità di *overlapping*: il partizionamento del grafo crea suddivisioni disgiunte o no?

Nell'ambito dello studio delle camere d'eco, per definizione delle stesse, ogni utente deve appartenere esclusivamente a una delle partizioni identificate.

Passando ora un lavoro posto all'interno della problematica analizzata in questa tesi, Garimella et al. (2017), al fine di identificare le partizioni che potrebbero essere camere d'eco, utilizza un particolare algoritmo, tale METIS, proposto da Karypis et al. (1995). La presentazione dettagliata, dato l'utilizzo dello stesso in questo lavoro di tesi, è rimandata al Capitolo 4.

Misurazione Controversia

In seguito alla costruzione della rete sociale, con la conseguenziale suddivisione in partizioni, l'ultima attività proposta nella pipeline è la quantificazione della controversia.

Tale attività si può intendere come verifica, circa la presenza di gruppi polarizzati, data la rete sociale definita.

Un primo approccio proposto da Coletto et al. (2017), date le tre strutture dati precedentemente definite, ha come obiettivo la definizione di un modello che faccia leva su diverse caratteristiche delle strutture create, in particolare si è tenuto in considerazione informazioni strutturali (numero di nodi, numero di archi e distribuzione di grado), informazioni di propagazione (profondità del *reply tree* e grado dell'autore del *root post*), informazioni temporali (tempo passato tra una risposta e l'altra e percentuale di nodi che rispondono entro un'ora, dato il *reply tree*) e informazioni locali (quanti utenti che si rispondono sono connessi da una relazione di *followee/follower*).

Questa metodologia, a differenza dell'approccio proposto da Garimella et al. (2017) il quale fa invece leva, come a breve verrà descritto, sulla struttura della rete valutando la possibilità di "movimento" all'interno della stessa in seguito a un'operazione di *clustering*, propone un'analisi delle caratteristiche locali della rete senza effettuare assunzioni circa la presenza di comunità, di fatto non eseguendo nessuna attività di *Community detection*. L'ipotesi portante del lavoro è che le diverse caratteristiche e peculiarità di G , T e R (introdotti nelle metodologie circa la costruzione del grafo a inizio sessione) possano essere utilizzata la fine di distinguere realtà dotate di controversia o meno in maniera automatica, senza passare dall'attività di partizionamento del grafo.

Una volta formalizzate le *feature* si è passati alla valutazioni di diversi modelli volti all'attività di classificazione, come realtà dotate di controversia o meno, dei *topic*. I risultati ottenuti sono positivi e, come ulteriore conclusione, sono stati evidenziati comportamenti caratterizzanti dei *topic* controversi come la tendenza a interagire con utenti non facenti

parte dell'insieme di *followee/follower*, la densità maggior di contenuto scambiato nel breve periodo con la conseguenziale profondità maggiore per il *reply tree*.

Un altro lavoro, proposto da Du et al. (2016), indaga la presenza di camere d'eco mediante un approccio di confronto temporale. Vengono difatti raccolti e modellati tramite rete sociale *followee/follower* dati provenienti da *Twitter*, in funzione di un determinato utente *seed*, dal quale viene effettuata un'attività di *crawling* fin tanto che non si è raggiunta una quantità sufficiente di dati.

Dopo un determinato periodo di tempo viene rieseguita la stessa attività di *crawling*, dallo stesso nodo *seed*. A questo punto le appartenenze comunitarie della prima rete vengono copiate sulla seconda, si passa quindi a un conteggio circa i nuovi archi extra-comunitari e intra-comunitari, creatosi ed eliminatosi. Emerge come la maggior parte di nuovi collegamenti sia intra-comunitario, così come quelli eliminati.

Questo fatto viene giustificato in quanto il numero di archi intra-comunitario è più alto, dunque risulta più probabile la cancellazione di questi archi che di archi extra-comunitari. A ogni modo i risultati sono giustificati dall'utilizzo di un modello casuale volto all'eliminazione (o all'aggiunta) di archi.

I risultati mostrano come l'aggiunta casuale di archi sulla rete, ipotizzando un moltiplicatore di crescita pari a 100, comporti una crescita di archi intra-comunitari molto più bassa rispetto a quanto registrato. Anche per ciò che concerne l'eliminazione degli archi, viene mostrato come una cancellazione totalmente casuale registra un numero di cancellazioni extra-comunitario, simile a quanto registrato dato il confronto temporale.

In generale risulta che le *community* precedentemente identificate abbiano subito un “potenziamento”, notificando la creazione di camere d'eco.

Questo lavoro richiede sicuramente uno sforzo computazionale notevole in termini di recupero dati, inoltre non è proprio un meccanismo di soluzione orientato a metodologie “*online*”. Tuttavia ha fornito un'ulteriore prova del fatto che, su *social media* come *Twitter*, l'utenza sia orientata alla definizione di partizioni che tendono a rinforzarsi sempre di più nel tempo. Portando quindi a sempre meno interconnessione con realtà esterne, questo può causare possibili polarizzazioni.

L'approccio che viene invece proposto da Garamella et al. (2017), ha l'obiettivo di definire una serie di metriche volte a quantificare la controversia il cui valore possa quindi rispondere alla domanda “*è presente polarizzazione tra le partizioni identificate dato il topic analizzato?*”.

Diverse sono le metriche utilizzate nello studio, al fine di poter quantificare la presenza di controversia; quella che sembra essere più promettente si basa sulla logica dei *Random Walks*: cammini casuali che partono da un determinato nodo; in questo modo gli autori cercano di quantificare quanto si sia esposti a informazione che nasce nella partizione di appartenenza o meno.

Tuttavia questo non è l'unica metrica proposta, la *Boundary Connectivity* (Guerra et al. 2013) asserisce che una rete priva di polarizzazione vede la presenza di nodi “frontiera” (quindi nodi che hanno collegamenti con vertici appartenenti a entrambe le partizioni) con un egual numero di connessioni intra ed extra comunitarie.

In maniera analoga vede la definizione anche una metrica basata sulla misura di

centralità *Betweenness*, la stessa ipotizza che reti dotate di controversia vedano i nodi facenti parti del taglio (inteso come gli archi che connettono elementi appartenenti a due partizioni diverse) essere dotati di un grado di *Betweenness* maggiore rispetto al resto del grafo.

Vengono poi proposte logiche “*embedded*” (*Embedding Controversy*) volte a valutare la distanza “vettoriale” tra nodi appartenenti alla stessa partizione e nodi appartenenti a partizioni opposte.

In ultimo si fornisce una misura basata sulla nozione fisica del *dipolo elettrico* (Morales et al. 2015), a grandi linee: una misura della polarità complessiva, in funzione degli elementi che compongono le due partizioni

3.2.2 Superamento delle camere d'eco

In questa sezione vengono descritti alcuni approcci proposti in letteratura per il superamento delle camere d'eco, con l'obiettivo di limitare le polarizzazioni alla base del fenomeno stesso.

Diversi lavori si sono concentrati sull'attività volta a riduzione della controversia; la presenza importante della controversia è sintomo di camere d'eco, pertanto l'attività di minimizzazione del valore di tale metrica coincide con l'apertura delle camere d'eco presenti, tentando quindi di rendere più fluida la diffusione dell'informazione in tutta la rete sociale definita.

Un primo approccio analizzato, proposto da Munson et al. (2013), prevede l'utilizzo di un *widget* da *browser* il quale notifica all'utente quanto, date le pagine visitate, se ci si trova sbilanciati circa i contenuti di queste; nella fattispecie con focus su contenuti ritenuti liberali e contenuti ritenuti conservatori.

Tale metodologie non prevede la modellazione tramite una struttura logica della rete sociale (i.e. la definizione di un grafo), si pone di mitigare il possibile fenomeno delle camere d'eco lavorando direttamente sul singolo individuo, a grana molto fine.

Al fine di valutare i risultati, il campione identificato, è stato monitorato prima dell'utilizzo del *widget* al fine di verificare quanti utenti fossero effettivamente sbilanciati (più contenuto liberale che conservatore o viceversa).

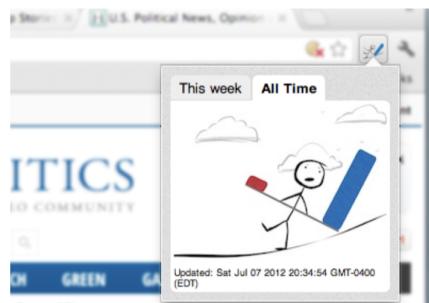


Figura 3.5: Animazione del widget proposto da Munson et al. (2013)

Soluzione sicuramente interessante ma che presenta delle criticità di fondo:

- Troppo “potere” all’utente, il quale si ipotizza essere “educato” all’autocorrezione circa i contenuti da lui consumati.
- Difficoltà nel classificare in maniera automatica e online le pagine come liberali, conservatrici o neutre.

In seguito al periodo di utilizzo, dato un campione di utenti, di questo *widget* è emerso come degli effettivi miglioramenti, seppur limitati, fossero avvenuti tra gli utenti ritenuti più polarizzati.

Un approccio volto a prevenire discussioni, sui *social media*, che possono portare a scontri e dunque causare polarizzazioni, viene proposto da Kumar et al. (2018).

In questo caso, la rete sociale viene modellata direttamente dai dati di *Reddit*², mediante un *reply graph*, in maniera simile ad altri lavori proposti. In questo *social media*, gli utenti condividono e creano contenuti direttamente in gruppi etichettati chiamati *subreddit*, queste sono di fatto pagine identificanti comunità che possono disquisire di un certo *topic*.

In particolare gli studi si sono concentrati su *subreddit* facenti parte dello stesso insieme (e.g. due *subreddit* facenti riferimento alla produzione di *meme*), in quanto è stato verificato come in questo contesto (stessa comunità virtuale) avvengano i maggiori confronti.

Il lavoro proposto cerca di definire uno strumento utile ai moderatori, al fine di identificare contenuti che possano dare il via a una discussione, che si può tramutare in scontro e dunque creare polarizzazioni (quindi portare a camere d’eco).

Viene definito un modello *LSTM* (*Long short-term memory*), il quale prende in input, in rappresentazione vettoriale, il testo prodotto, l’utente che lo ha prodotto e le *subreddit* di origine e di target del testo prodotto.

Difatti, è stato studiato che su questo *social media* gli scontri e discussioni tra gli utenti avvengano in merito alla creazione di un contenuto in una *subreddit* che faccia riferimento esplicito (quindi *linkando*) un’altra *subreddit*; questo tipo di contenuto viene definito di *cross-linking* o di *mobilizzazione*.

In particolare l’utente è caratterizzato dalle *subreddit* sulle quali ha postato, invece una *subreddit* è caratterizzata dagli utenti che hanno su questa creato contenuti. L’attività di *embedding* è ottenuta mediante approccio GloVe (Pennington et al. 2014).

Il dataset sul quale il modello opera è stato etichettato manualmente dagli autori dell’articolo proposto, partendo dalle informazioni circa le interazione tra gli utenti su *Reddit*; le modalità è l’approccio di etichettamento è ben documentato nel lavoro proposto.

Elemento fondamentale dello studio, soprattutto nella logica dell’etichettamento, è la definizione delle *negative mobilization* e *neutral mobilization*, ovvero tutti i messaggi di *mobilizzazione*. Al fine di identificare tali post di *cross-linking*, viene tenuto in considerazione il *sentiment* associato al contenuto generato.

Per tale valutazione sono state utilizzate, tra le altre cose, anche lessici quali *VADER* (Hutto et. al. 2014) e *LIWC* (Tausczik et al. 2010).

Vengono quindi condotte analisi circa le *negative mobilization*, sono emersi interessanti aspetti circa le peculiarità degli utenti autori del *negative mobilization*.

²<https://www.reddit.com/>

Difatti, pare che gli autori dei contenuti di *cross-linking* negativi, siano utenti molto attivi nella *subreddit* di origine (attivi in termine di creazione di contenuti). Tuttavia, il ruolo predominante nelle discussioni che nascono successivamente, è ricoperto da quell'insieme di membri che non risulta essere così attivo. Questo viene giustificato ipotizzando che gli utenti meno attivi, colgano questa occasione di scontro per aumentare il loro *status* all'interno del gruppo, mostrando quanto “valgono”.

In generale il lavoro svolto da Kumar et al. (2018), fornisce una panoramica circa gli scontri che avvengono sui *social media*, cerca poi di definire uno strumento che cerca di fornire una probabilità circa la possibilità di verifica di tali scontri, in modo tale da evitare contrapposizioni e polarizzazioni, che possono poi portare alla creazione di camere d'eco; le quali vengono tra le altre cose effettivamente identificate, mediante una logica volta alla quantificazione delle polarizzazione con metriche, nel *reply graph* generato dai *cross-link negative mobilization*.

In generale l'approccio risolutivo non cerca di semplificare l'interfacciamento tra utenti facenti parte di *community* diverse, viene semplicemente tentato di bloccare le interazioni negative (*negative mobilization*).

Nel lavoro proposto da Garimella et al. (2017) si valuta come ridurre la polarizzazione sui *social media*, dati *topic* controversi, mediante la creazione di “ponti” (quindi di archi) che possano connettere le due opposte partizioni, di fatto andando ad attuare una strategia di *Link prediction*.

L'ipotesi di base è che andando a collegare diversi vertici, facenti parti delle due partizioni opposte, si possa ridurre la controversia identificata all'interno del grafo; il grafo $G = (V, E)$, sul quale si opera, vede come vertici gli utenti che disquisiscono su un determinato *topic* e come archi i *retweet* intercorsi tra gli utenti (in modo da catturare la *user-to-user endorsement*).

L'obiettivo del lavoro è quindi identificare l'insieme più piccolo possibile di archi che possa minimizzare la controversia del grafo, la quale è quantizzata dall'utilizzo della *Random Walk Controversy* (RWC). L'algoritmo proposto dallo studio è mostrato nella Figura 3.6.

Gli autori sono consci del fatto che non tutti gli archi abbiano la stessa possibilità di comparire naturalmente nella rete; viene quindi proposta una modellazione di un valore di probabilità di accettazione che viene valutata al fine di pesare ulteriormente, dunque ordinare, gli archi da aggiungere. Il modello proposto si basa sulla polarità associata all'utente valutando, ogni qualvolta si voglia aggiungere un arco, quanti archi effettivamente esistono che collegano due utenti con quella polarità. Questa modellazione si pone dunque di proporre archi che minimizzano la *controversy* e che inoltre possano essere più possibilmente verosimili.

Ad ogni modo urge ricordare, come precedentemente accennato, che le persone sono più propense a selezionare informazioni che rinforzano le loro attitudini, piuttosto che informazioni contrarie alle stesse. La modellazione di tale valore può sicuramente ritenersi corretta, ma è pur sempre un'approssimazione forzata di ciò che poi guida l'accettazione, dunque la nascita, di un determinato collegamento.

Algorithm 1: Algorithm for k -EDGEADDITION

Input: Graph G, number of edges to add, k ; k_1, k_2 high degree vertices in X, Y respectively
Output: List of k edges that minimize the objective function, RWC

```

1 Initialize: Out ← empty list ;
2 for  $i = 1:k_1$  do
3   vertex  $u = X[i]$ ;
4   for  $j = 1:k_2$  do
5     vertex  $v = Y[j]$ ;
6     Compute  $\delta\text{RWC}_{u \rightarrow v}$ , the decrease in RWC if the edge  $(u, v)$  is added;
7     Append  $\delta\text{RWC}_{u \rightarrow v}$  to Out;
8     Compute  $\delta\text{RWC}_{v \rightarrow u}$ , the decrease in RWC if the edge  $(v, u)$  is added;
9     Append  $\delta\text{RWC}_{v \rightarrow u}$  to Out;
10 sorted ← sort(Out) by  $\delta\text{RWC}$  by decreasing order ;
11 return top  $k$  from sorted;
```

Figura 3.6: Pseudocodice algoritmo di riduzione controversia proposto da Garimella et al. (2017)

In generale, avendo come obiettivo ultimo la riduzione della polarizzazione, una delle attività attuabili prevede di rendere più coesivo il grafo analizzato, andando ad aggiungere archi; di fatto si modifica la struttura topologica dello stesso. Molti lavori allo stato dell'arte hanno avuto come obiettivo l'aumento della coesività di un certo grafo, dove tale metrica viene quantificata come *shortest paths* Papagelis et al. (2011), Parotsidis et al. (2015), *closeness centrality* Parotsidis et al. (2016), *diametro* Demaine et al. (2010), *eccentricità* Perumal et al. (2013), *comunicabilità* Arrigo et al. (2015), *sincronizzabilità* Zeng et al. (2012) e *natural connectivity* Chan et al. (2014).

Sebbene questi lavori non avessero come obiettivo ultimo la riduzione della controversia, non ponendosi proprio nel contesto applicativo delle camere d'eco, sono utili spunti al fine di valutare in che modo muoversi al fine di rendere più coeso un grafo nel tentativo ultimo di minimizzare, con il minor sforzo possibile, i valori di polarizzazione.

3.2.3 Discussione

In questa sezione sono stati descritti diversi lavori, il cui intento era quello di identificare in maniera puntuale le polarizzazioni o, date situazione di polarizzazione, superare queste.

Gli studi letti, e qui proposti, hanno inevitabilmente influenzato il lavoro portato avanti in questa tesi.

Rimandando poi per maggiori dettagli al Capitolo 4, il quale descrive approfonditamente la soluzione proposta, risulta possibile anticipare che è stato deciso di definire una *pipeline*, volta alla quantificazione della controversia con la consecutiva riduzione della stessa.

In particolare la modellazione della rete sociale avverrà mediante un grafo che rappresenterà l'interazione *online* tra gli utenti, definendo così la rete sociale; in seguito alla rappresentazione della rete si passerà al partizionamento della stessa.

L'attività di *Community detection* vedrà l'utilizzo dell'informazione semantica come proposto nel lavoro di Sawhney et al. (2017), dunque mediante la definizione di una nuova serie di pesi.

Data l'identificazione delle partizioni, la quale verrà in qualità quantificata, si procederà alla valutazione circa la presenza di controversia dato il partizionamento migliore, come proposto nel lavoro di Garimella et. al (2017); la quantificazione della controversia verrà effettuata utilizzando sia metriche allo stato dell'arte, sia metriche definite durante questo lavoro.

Per ciò che concerne invece il superamento delle eventuali polarizzazioni, si procederà con un'attività di *link prediction* come proposto nel lavoro di Garimella et al. (2017). Verrà inoltre tenuta in considerazione una “probabilità di accettazione”, che possa modificare lo *score* di un potenziale arco, in funzione della verosimiglianza di questo rispetto allo stato della rete.

L'obiettivo di questa attività è quello di identificare l'approccio che, a parità di archi aggiunti, possa permettere un superamento dell'eventuale polarizzazione, andando a minimizzare la *controversy* identificata.

Il superamento del fenomeno verrà in qualità quantificato, valutando la strutturazione in *community* in seguito all'attività di *linking*.

Capitolo 4

La soluzione proposta

Come illustrato nel Capitolo 3, alcuni approcci sono stati proposti in letteratura per cercare di esplorare e mitigare il fenomeno delle camere d'eco, si sono inoltre discusse metodologie volte a inserire informazioni di carattere semantico all'interno delle attività di *Community detection* infatti, in questo lavoro di tesi, ci si è focalizzati sul cercare di prendere in considerazione sia gli aspetti topologici sia legati ai contenuti delle reti sociali al fine di caratterizzare meglio la composizione delle *community* che si andranno ad identificare, nel tentativo poi di verificare se queste siano effettivamente camere d'eco.

In questo capitolo viene presentata la soluzione proposta per la modellazione e l'arricchimento del grafo rappresentante la realtà di interesse, per poi passare alle attività che hanno caratterizzato la quantificazione e la riduzione della controversia identificata. Tutti i risultati ottenuti sono descritti nel Capitolo 5 di questa tesi.

In particolare la suddivisione in sezioni di questo capitolo è la seguente:

- Nella Sezione 4.1 vengono introdotti i dataset utilizzati in questo lavoro di tesi, e illustrate le modellazioni che hanno permesso la definizione delle reti sociali rappresentanti un determinato *topic*.
- Nella Sezione 4.2 viene introdotta la scelta adottata per il partizionamento del grafo, in particolare si descrive l'algoritmo utilizzato e le metriche volte a quantificare la qualità del risultato ottenuto. In questa sezione si presenterà anche la modellazione dell'informazione semantica, con l'obiettivo di inserire la stessa nel processo di *Community detection*.
- Nella Sezione 4.3 vengono introdotte formalmente le metriche utilizzate nelle attività di quantificazione della controversia, date le partizioni identificate.
- Nella Sezione 4.4 si definiscono le possibili modalità di *Link prediction* utilizzabili e si pone l'attenzione anche sulla definizione di un moltiplicatore volto a modellare una “probabilità di accettazione” dell’arco. Infine viene descritto come si procederà con l’effettiva aggiunta dei nuovi archi.

4.1 Modellazione delle reti sociali

Questa sezione illustra come sono state rappresentate le reti sociali considerate nel lavoro di tesi. In particolare viene descritto il *social media* utilizzato come fonte di dati (Sezione 4.1.1), fornendo una panoramica circa le interazioni attuabili dagli utenti dello stesso.

In seguito vengono presentati i dati utilizzati in funzione dello stato dell'arte (Sezione 4.1.2), per poi passare a presentare due ulteriori dataset facenti riferimento al Covid-19 e alle Vaccinazioni (Sezioni 4.1.3 e 4.1.4)

4.1.1 Twitter come fonte di dati

I dataset scelti per condurre le sperimentazioni contengono tutti informazioni ottenute dal *social network Twitter*; *Twitter* è un servizio di *microblogging* su cui gli utenti postano e interagiscono tra di loro con messaggi chiamati *tweet*. Su questo *social* è possibile seguire ed essere seguiti da altri utenti, in particolare le diverse azioni/interazioni possono essere di tipo:

- *General Tweet*: un messaggio postato contenente testo, foto, *GIF* e/o video.
- *Followee/Follower*: gli utenti possono seguire o essere seguiti da altri utenti, non è detto che se si segue un utente questo ricambierà.
- *Mention*: un *tweet* generale con un riferimento ad un altro account *twitter*; il riferimento è ottenuto mediante l'accostamento del simbolo “@” all’*username* dell’utente che si intende menzionare.
- *Replies*: messaggio lasciato come commento a un certo *tweet*.
- *Retweet*: condivisione di un *tweet* di un altro utente; sullo stesso sarà possibile commentare. Restano visibili le informazioni circa l’utente che per primo ha scritto il messaggio.

Le interazioni che risultano essere interessanti per questo studio sono le ultime tre, questo perché si prestano ad essere modellate per mezzo di un grafo che rappresenta la vera e propria rete sociale; su questo grafo sarà possibile poi effettuare tutte le operazioni volte a identificare camere d’eco, quantificare e minimizzare la *controversy*.

Tuttavia, avere informazioni circa la prima tipologia di messaggio (*General Tweet*) risulta essere estremamente utile in quanto, un’analisi dei contenuti, può fornire indizi circa il pensiero del singolo e quindi permettere di effettuare indagini riguardanti il sentimento e i contenuti condivisi stand-alone dai componenti di un certa comunità.

Un’ulteriore considerazione deve essere fatta circa la relazione *followee/follower*: in diversi studi nell’ambito delle camere d’eco questa tipologia di relazione viene impiegata nella modellazione delle reti sociali; in particolare, nei dataset allo stato dell’arte, era possibile usufruire anche di questa tipologia di relazione. A livello pratico, in questo lavoro, non verrà utilizzata in quanto:

- Si ritiene più utile, al fine di identificare le camere d’eco, l’utilizzo di informazioni che modellino le interazioni semantiche tra utenti (e.g. *retweet*, *mention*).

- Non si era in possesso, per i dataset Covid-19 e *Vaccination*, di sufficienti informazioni circa le relazioni *followee/follower*.

In questo lavoro di tesi sono state utilizzate, al fine di modellare le interazioni tra gli utenti, le informazioni circa i *retweet* per i dataset allo stato dell'arte, mentre per i dataset Covid-19 e *Vaccination* sono state utilizzate quelle legate alla relazione di *mention* (i *general tweet* sono stati utilizzati principalmente per attività di *SCA*).

Si ritengono confrontabili le modellazioni per *retweet* e *mention*, circa i risultati, in quanto entrambi modellano un tipo di interazione che permette a un utente di attirare l'attenzione di uno o più utenti su un determinato contenuto semantico; permettono quindi di esprimere e condividere un'opinione personale.

L'utilizzo delle relazioni di tipo *followee/follower*, quando disponibili, potrebbero essere utilizzate come “rafforzativo” rispetto alle interazioni di *retweet/mention*, al fine di rendere più importante l'attività di interfacciamento con utenti che si seguono.

4.1.2 Dataset allo stato dell'arte

Ai fini dello studio sono stati utilizzati diversi dati, in primis citiamo i dataset impiegati da Garimella et al. (2017) in diversi lavori circa la quantificazione della *controversy*. I dati sono stati raccolti in funzione dei diversi *hashtag* utilizzati; gli *hashtag* sono molto utilizzati dagli utenti di *Twitter* a mo di “etichetta” e vengono posti direttamente sul *tweet* condiviso. Gli *hashtag* permettono di identificare in maniera semplice un determinato *topic*.

Sui dati raccolti, dati gli studi originali, sono state poi modellate diverse tipologie di relazioni: *retweet*, *followee/follower*, *content* e *hybrid content*. I dataset utilizzati, dallo studio di Garimella et al. (2017), risultano essere solo quelli basati sulle relazioni di *retweet* e *followee/follower* in quanto ritenuti più utili per la quantificazione della *controversy*; in particolare si è deciso di utilizzare, in questa tesi, solo i grafi ottenuti dall'analisi dei *retweet*.

I dati raccolti fanno già riferimento alla *giant component*, vale a dire la componente connessa di dimensione maggiore; la strutturazione del dato è la seguente:

$$user1, user2, weight$$

Si è quindi in possesso degli *username* (univoci) degli utenti e del numero totale di *retweet* intercorsi tra i due. Si ipotizza che non vi sia un senso di lettura riguardante il collegamento tra gli utenti in quanto, presa a caso una qualsiasi coppia di utenti indipendentemente dal peso associato, si avrà sempre e solo *user1 user2 weight* e mai *user2 user1 weight*; si ipotizza dunque che il peso sia dato dal numero di *retweet* intercorsi, in un certo periodo di osservazione, tra i due utenti.

Nel grafo che rappresenta la rete sociale ogni nodo rappresenta un singolo utente e sarà etichettato con lo *username* univoco di questo; ogni arco rappresenta la presenza di *retweet*, la stessa viene quantificata da un'etichetta posta sullo stesso (*weight*). La Figura 4.1 rappresenta la modellazione di una singola coppia di nodi.

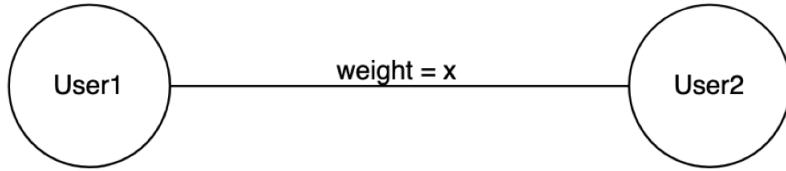


Figura 4.1: Esempio modellazione su singola riga dei dati utilizzati da *Garimella et al.*

I grafi ottenuti da questi dati verranno utilizzati come una sorta di *ground truth* in quanto risultano essere divisi in funzione della presenza o assenza della *controversy*, quindi divisi in funzione dell'assenza o presenza di *community* polarizzate.

Ipotizziamo dunque che pipeline sia grado di notificare la presenza o assenza della *controversy* su questi grafi, si potrà passare a una valutazione puntuale e corretta anche sugli altri grafi utilizzati in questo lavoro di tesi.

La costruzione della rete non ha visto l'eliminazione di nodo alcuno; questo perché, come precedentemente detto, i dati sono già stati ripuliti e comprendono la componente connessa di dimensione maggiore.

La Tabella 4.1 mostra informazioni di carattere quantitativo legate alla modellazione della rete sociale per tutti i dataset scelti.

I valori presenti nei campi $|E|$ e $|E|$ reali fanno riferimento, rispettivamente, al numero di archi indipendentemente dal peso e al numero di archi effettivo in funzione del peso. I dati sono pubblici e accessibili¹.

Hashtag	Con <i>controversy</i>				Periodo (2015)
	$ V $	$ E $	$ E $ reali	Topic	
#ukraine	5495	9452	29248	Guerra del Donbass.	27 Feb - 2 Mar
#beefban	1610	1978	4949	Divieto, da parte del governo Indiano, di consumare carne di manzo.	2 - 5 Mar
Senza <i>controversy</i>					
#kissingday	4638	4816	10241	Giornata internazionale del bacio.	19 Giu

Tabella 4.1: Informazioni sulle Reti Sociali dati i lavori condotti da *Garimella et al. (2017)*

4.1.3 Dati Covid-19

Un altro insieme di dati che si è utilizzato riguarda un ambito più specifico e sicuramente attuale: quello del Covid-19; in particolare è stato utilizzato un dataset collezionato per uno studio collaterale contenente *tweet* scaricati dal 12 Gennaio 2020 fino al 15 Marzo 2020 che contenevano la stringa “coronavirus” nel testo². Il dataset contiene 10.343.121 *tweets* scritti in diverse lingue, generati da 3.837.794 utenti.

¹<https://github.com/gvrkiran/controversy-detection>

²Il quel periodo il termine *coronavirus* era ancora molto utilizzato per discutere del Covid-19

Data la conformazione dei dati, è stato possibile modellare solo il grafo che, in generale, riguarda le *mention* degli utenti, quest'ultimo è comunque in grado di rappresentare eventuali camere d'eco ed eventuali polarizzazioni, questo perché tendenzialmente il menzionare un altro utente in un *tweet* permette di comunicare con questo, portando la sua attenzione sul *tweet* stesso. Non è stato possibile modellare una rete di *retweet* in quanto mancava l'informazione circa tale interazione.

I dati disponibili sono quindi i seguenti:

username_sender, tweet, mentions, #retweet, #likes

Vi sono informazioni circa l'utente autore del *tweet* (*username_sender*, univoco), il contenuto testuale dello stesso, la lista di utenti citati (potrebbe essere anche vuota nel caso di nessuna citazione) e il numero di *like* e *retweet* che il *tweet* ha ricevuto. Quindi, sostanzialmente, una riga rappresenta un singolo *tweet*.

Si è inoltre in possesso di una serie di metadati riguardanti gli utenti, questi sono i seguenti:

username, creazione, default_profile default_image, descrizione, #likes, #followee, #follower, #liste, luogo, #stati, verificato

I metadati riguardano la data di creazione del profilo, se il profilo ha i *setting* e l'immagine di default, la descrizione del profilo, il numero totale di *like* messi e di *retweet* fatti, il numero di *followee* e di *follower*, informazioni geografiche, il numero di gruppi a cui si è iscritti, la *location*, numero di *tweet* (inclusi *retweet*) fatti dal profilo e se il profilo è verificato; questi sono stati utili per condurre analisi più di carattere quantitativo. Si passa ora alla definizione dei grafi ottenuti in fase di modellazione della rete sociale.

Per questo dataset, è possibile tenere in considerazioni molti più aspetti rispetto a quanto visto precedentemente nella modellazione della rete sociale. Questo è dovuto al fatto che si è in possesso di un maggior numero di informazioni per ogni *tweet* e per ogni *user* della rete.

Il grafo risultante vede come nodi gli utenti con tutti i metadati riguardanti gli stessi, mentre gli archi sono di fatto i *tweet* (rappresentanti o *general tweet* o *mention*) con tutte le informazioni citate sopra.

La Figura 4.2, mostra il tipo di modellazione in funzione alla citazione presente nel *tweet*.

Si procede ora ad un'analisi dettagliata di tale modellazione, date le possibili casistiche:

- Nel caso in cui un utente non citi nessun altro, si verrà a creare un arco *self-loop* dato l'utente come mostrato nella Figura 4.2a. Un utente, quindi un nodo, potrà essere dotato sia di archi in uscita verso altri utenti, archi di entrata provenienti da altri utenti che archi *self-loop*.
- Nel caso di una citazione singola si verrà a creare un arco orientato in funzione della citazione, come mostrato in Figura 4.2b; l'arco stesso verrà arricchito di tutte le informazioni circa il *tweet* stesso.
- Nel caso di più di una citazione si andranno a creare n archi dove n è il numero di utenti citati, come mostrato in Figura 4.2c. Ogni arco conterrà le informazioni circa il *tweet* sotto esame.

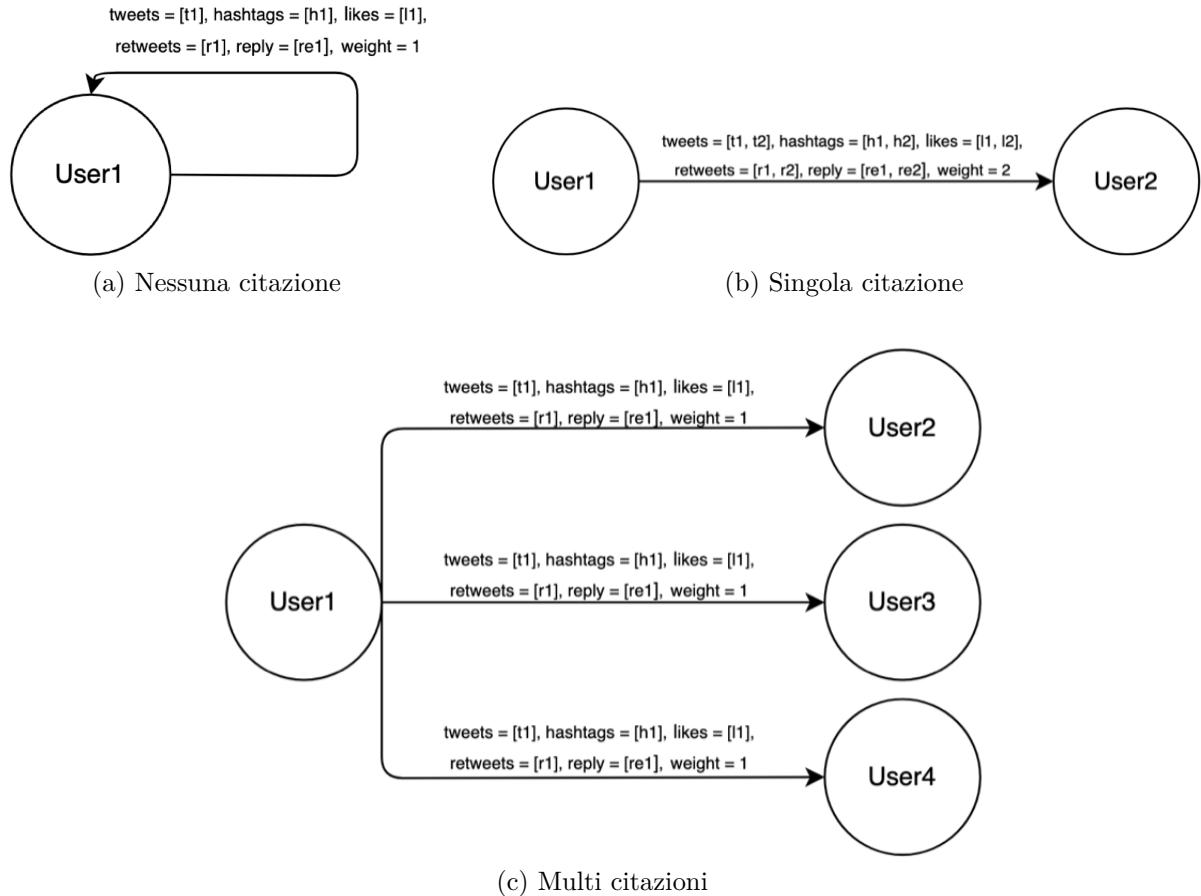


Figura 4.2: Modellazione in funzione delle citazioni trovate

Ovviamente ad ogni aggiunta di un nuovo arco si valuterà la presenza del/i nodo/i interessati, nel caso in cui non fossero presenti si procederà inserendoli. In ogni caso possiamo osservare la creazione di diverse liste, le stesse andranno a caratterizzare l'arco; tutte le informazioni presenti all' i -esima posizione delle liste faranno tutte riferimento al i -esimo *tweet*.

Ad ogni nuovo *tweet*, se gli estremi dello stesso sono già presenti, si procederà arricchendo l'arco di nuove informazioni aumentando la lunghezza delle liste. L'attributo *weight* è semplicemente il numero di *tweet*, è stato necessario definire questo attributo in quanto ha reso più comode le fasi di costruzione del grafo; inoltre, gli algoritmi di *Community detection*, al fine di poter notificare il peso hanno bisogno di rifarsi direttamente al nome dell'attributo, su cui calibrare l'algoritmo stesso.

Data la grande mole di dati iniziale, si parla di circa 3 milioni di *tweet*, è stato deciso di tenere in considerazione solo gli archi con un peso associato maggiore di 3; si vuole inoltre sottolineato che, in questo lavoro di tesi, sono stati tenuti in considerazione solamente i *tweet* scritti in lingua inglese (le dimensioni riportate a inizio sezione riguardavano tutte le lingue).

Le motivazioni circa i vincoli definiti in termini di partecipazione, è da attribuirsi alla politica di “eliminazione” del possibile rumore. In seguito alla procedura di eliminazione è

stata mantenuta la componente connessa di dimensione maggiore.

Vengono quindi forniti maggiori informazioni circa la rete sociale effettivamente ottenuta:

Covid-19			
Grafo	V	E	E reali
Grafo orientato	39336	85176	1186709

Tabella 4.2: Informazioni sulla Rete Sociale - Covid 19

Solo 2705 nodi non risultano dotati di nessuno dei metadati precedentemente introdotti.

4.1.4 Dati Vaccinazioni

Il terzo ed ultimo insieme di dati utilizzati riguarda un tema che, negli ultimi anni, ha diviso molto l’opinione pubblica ed è quindi terreno di dibattito spesso molto polarizzato su certe posizioni e costellato di *fake news*: le vaccinazioni³.

Le informazioni in possesso sono del tutto analoghe a quando visto per il secondo dataset:

username_sender, tweet, mention, #replies, #retweet, #like, hashtags

Vi è dunque l’*username* (univoco) dell’utente che ha scritto il *tweet*, il testo del *tweet* stesso, una lista contenente gli *username* degli utenti menzionati nel *tweet* (nel caso in cui non vi fossero citazioni il campo risulterà vuoto), il numero di commenti, *retweet* e *like* che il post ha ricevuto ed gli eventuali *hashtag* del *tweet*. Dunque, anche in questo caso, il tipo di rete sociale modellabile è una rete delle citazioni intercorse tra gli utenti.

I dati sono stati raccolti dal 20 Giugno 2019 al 3 Ottobre 2019, sono stati estratti tutti i *tweet* i quali avevano la stringa “*vaccination*” nel testo.

Partendo da questi dati è stata generata la rete sociale, la tipologia di rete generata è del tutto analoga a quella presentata per Covid-19.

Il grafo rappresentante la rete sociale è quindi dotato di tutte le informazioni che possono essere ricavate dal dataset sotto esame; i nodi saranno costituiti dagli utenti presenti, gli archi e il senso di questi sarà dettato dal senso della citazione.

Gli archi stessi saranno arricchiti con tutta l’informazione di carattere quantitativo presente: numero *like*, *retweet*, *replies* e *hashtags*.

Nel caso in cui tra due utenti , data la realtà “compatta” del grafo, intercorressero più *tweet* si verrà a creare una lista per ogni elemento rappresentato sull’arco. Tale liste avranno che, dato *i-esimo* l’elemento di ognuna, questo farà riferimento all’*i-esimo* tweet.

Nel caso in cui un utente citi, all’interno di un determinato *tweet*, più di un singolo altro utente si avrà la generazione di tanti archi contenenti la stessa informazione, tanti quanti sono gli utenti citati.

Nel caso in cui un utente, nel *tweet*, non citi nessuno si genererà un *self-loop* con tutte le informazioni circa il *tweet* stesso.

³<https://www.kaggle.com/keplaxo/twitter-vaccination-dataset>

Le immagini esplicative circa la modellazione sono del tutto analoghe a quelle mostrate in Figura 4.2 per il dataset riguardante il Covid-19; rimandiamo dunque il lettore alla stessa.

Una volta definita la rete sociale, come sopra mostrato, si è poi proceduto all'estrazione della *giant component*.

Vengono fornite le informazioni di carattere quantitativo data la rete sociale ottenuta:

Vaccination			
Grafo	V	E	E reali
Grafo orientato	39604	95126	221264

Tabella 4.3: Informazioni sulla Rete Sociale - *Vaccination*

4.2 *Community Detection*

In questa sezione viene descritta nel dettaglio la soluzione adottata per la fase di partizionamento del grafo. In particolare viene introdotta tale attività nel contesto delle camere d'eco (Sezione 4.2.1), passando poi alla presentazione dell'algoritmo utilizzato in questo lavoro di tesi (Sezione 4.2.2).

Si introduciranno quindi le strutture dati parallele, date le reti sociali definite, sulle quali sono state portate avanti le attività di *Community detection* (Sezione 4.2.3); per poi passare a proporre la modellazione dell'informazione semantica adottata (Sezione 4.2.4). In ultimo verranno formalmente definite le metriche utilizzate, in ambito di quantificazione della qualità delle partizioni ottenute (Sezione 4.2.5).

4.2.1 Introduzione all'attività di *Community detection* nell'ambito delle camere d'eco

L'identificazione di un algoritmo ottimale per la *Community detection* è stata resa difficile dal fatto che i grafi sul quale bisognava applicare tali algoritmi risultavano essere dotati di notevoli dimensioni, di conseguenza i tempi computazionali potevano subire incrementi importanti; inoltre la maggior parte degli algoritmi non permettevano di definire il numero di *community* richiesto, rendendo di fatto molti algoritmi non utili rispetto al problema considerato, in quanto risultava necessario impostare un valore pari a due circa le *community* ricercate.

Il vincolo di due *community* può introdurre *bias*, date le operazioni di identificazione delle comunità, in quanto forza ogni nodo ad appartenere a una delle due *community*; di fatto si può pensare di star generalizzando e perdendo informazioni imponendo un determinato valore.

Anche gli studi di Garimella et al. (2017) sollevano e discutono tale problematica, sempre negli stessi vengono fornite delle giustificazioni andando a sottolineare come, nei dibattiti politici che stanno alla base degli studi ed in generale rispetto a certi argomenti controversi, sia sempre possibile identificare due grandi visioni opposte, che ad esempio in ambito politico vanno a identificare i due grandi universi politici di destra e sinistra.

In ogni caso, come verrà mostrato nei risultati dell'attività, ciò che si ottiene tenendo conto anche delle analisi effettuate sui metadati, denotano un'effettiva presenza di *community* sensate e contrastanti per ciò che concerne i contenuti scambiati o gli utenti presenti all'interno delle stesse.

Uno scenario con più piccole camere dell'eco ha motivo di esistere ma è meno probabile che appaia, questo perché gli utenti, sulle piattaforme dei *social media*, tendono a selezionare affermazioni che aderiscono al loro sistema di credenze e ad ignorare le informazioni dissidenzienti.

La conseguenza di questo processo potrebbe essere una sorta di massimizzazione dell'omogeneità che è meglio rappresentata da due comunità quasi disgiunte piuttosto che da vari gruppi frammentati.

In ogni caso, in questa fase, si è andati ad effettuare attività di *Community detection* in funzione del peso strutturale associato all'arco; nelle prossime sezioni verrà anche mostrato il come si è modellata l'informazione semantica mediante la definizione di nuovi pesi. Al fine di identificare l'approccio migliore si è pensato di valutare le differenze in funzione di metriche circa la qualità delle attività: modularità, copertura, tempi computazionali, numerosità delle *community* identificate, distribuzione del *sentiment* e dei *topic* discussi nelle due partizioni identificate.

Inizialmente, oltre all'algoritmo che verrà presentato nella prossima sezione, si sono valutati altri algoritmi di *Community detection*; in particolare quelli noti come *Kernighan/Lin Bisection* (Kernighan et al. 1970) e *FluidC* (Parés et al. 2017).

Il primo è stato scartato in quanto i tempi computazionali richiesti, già solo sulle reti sociali allo stato dell'arte, risultavano essere molto elevati; l'utilizzo del *Kernighan/Lin Bisection* sui grafi facenti riferimento a *Covid-19* e *Vaccination*, i quali risultano essere di molto più densi delle reti allo stato dell'arte, avrebbe richiesto un quantitativo enorme di tempo.

Il secondo è stato scartato in quanto molto dipendente dalla scelta iniziale: impostando diversi *seed*, volti a identificare il nodo dal quale far partire l'attività di partizionamento, il risultato finale poteva cambiare di molto. Questo poteva produrre delle *community* che, data la rete priva di effettiva opposizioni (#*NationKissingDay*), osservavano valori alti circa le metriche volte a quantificare la qualità del partizionamento; tali risultati hanno spinto a preferire l'algoritmo che viene descritto nella prossima sezione.

4.2.2 METIS

L'algoritmo utilizzato in questo lavoro è noto col nome di *METIS* ed è stato proposto da Karypis et al. (1995); la scelta è ricaduta su questo algoritmo in quanto è stato utilizzato nelle studi circa la quantificazione delle controversie sui *social media* di Garimella et al. (2017), si è quindi ritenuto opportuno utilizzare lo stesso strumento, al fine di avere un'unità di misura validata dalla presenza di un *paper* circa l'impiego dello stesso, nelle attività di identificazione delle camere d'eco.

METIS è un algoritmo volto al partizionamento dei grafi non orientati, in funzione

delle caratteristiche topologiche della rete; il partizionamento è basato su una logica *k-way* multilivello, che vede quindi una progressiva riduzione del grafo, con una successiva “ricrescita” alle dimensioni originali.

È stato stimato che le partizioni ottenute da *METIS* risultano essere, tra il 10% e il 50%, migliori rispetto a suddivisioni effettuate da algoritmi di partizionamento spettrale; inoltre, il tempo *performing* richiesto da *METIS* è molto inferiore rispetto ad algoritmi facenti parte della stessa famiglia.

L’idea di fondo di questo algoritmo è semplice e si compone di diversi passi, partendo da un generico grafo non orientato G si predispongono le seguenti fasi:

- **Fase di riduzione:** il grafo G è ridotto a poche centinaia di vertici, questo viene ottenuto con una serie di passaggi sequenziali, si diminuirà progressivamente le dimensioni del grafo. Partendo quindi da un grafo iniziale $G_l = (V_l, E_l)$ si passerà a un grafo $G_{l+1} = (V_{l+1}, E_{l+1})$ tale per cui $|V_l| > |V_{l+1}|$. Si effettua quindi un *mapping* a un grafo con una dimensionalità ridotta con l’obiettivo ultimo di arrivare a poche centinaia di vertici.

In questo fase non più di due vertici vengono collassati sullo stesso, i vertici che non subiscono una aggregazione vengono semplicemente copiati nel grafo successivo; quando due vertici $u, v \in V_l$ vengono compattati a formare un vertice $w \in V_{l+1}$, il peso di quest’ultimo sarà pari alla somma dei pesi dei due iniziali, gli archi incidenti in w sono l’unione degli archi incidenti in u e v .

Nel caso in cui lo stesso arco fosse incidente in entrambi i vertici iniziali, il peso di questo verrà impostato pari alla somma dei pesi dei due archi iniziali. La riduzione viene ottenuta con una metodologia chiamata *heavy-edge matching* (HEM), questa definisce il *mapping* a una dimensione ridotta cercando la configurazione con peso massimo degli archi, dato un nodo scelto casualmente, questo verrà aggregato a un altro nodo con il quale condivide l’arco con peso maggiore.

- **Fase di partizionamento:** la seconda fase dell’algoritmo multilivello prevede la suddivisione del grafo in due parti il più possibile uguali (a livello di numerosità).

Il taglio definito includerà quegli archi i quali, dati loro pesi, rappresentano di fatto il taglio con peso minimo. La partizione del grafo può essere ottenuta utilizzando differente tecniche, in particolare *METIS* utilizza logiche di *graph growing heuristics*; vi è la possibilità di utilizzare approcci basati sulla bisezione spettrale, i quali però ottengono performance leggermente peggiori.

- **Fase di crescita:** durante questa fase, il partizionamento ottenuto sul grafo, data lo *step* precedente, inizia ad essere proiettato verso il grafo iniziale. Fintanto ogni vertice $u \in V_l$ non rappresenti un singolo vertice, si continuerà a proiettare.

Bisogna tenere in considerazione che un certo partizionamento con taglio locale minimo su G_l , potrebbe non essere un taglio locale minimo su G_{l-1} ; è presente quindi anche una fase di eventuale miglioramento del partizionamento ottenuto mediane l’utilizzo di algoritmi di raffinamento della partizione, questi identificano due insiemi di vertici (date le due partizioni) e valutano quale delle due ha taglio con peso minore.

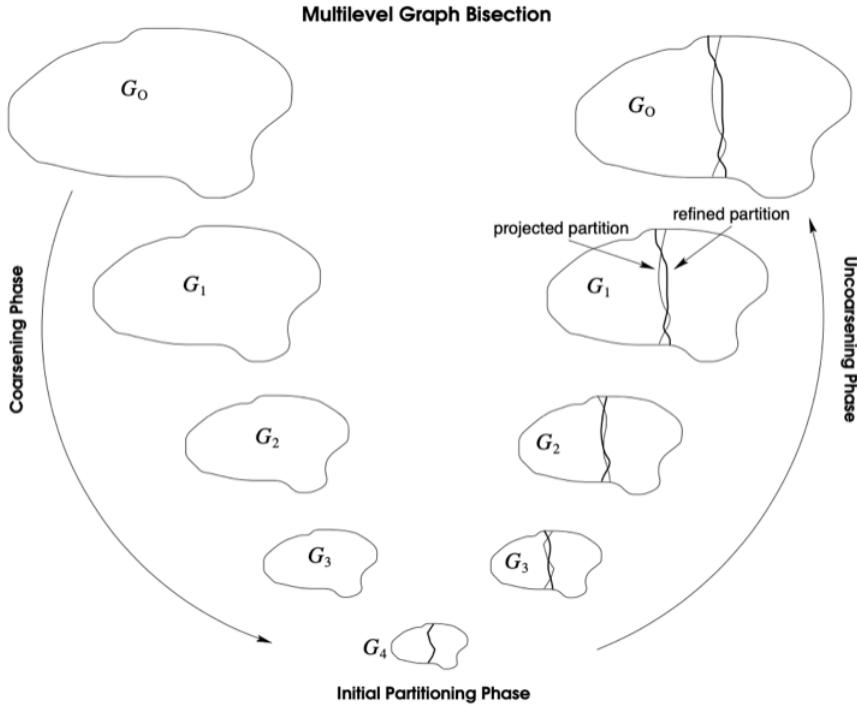


Figura 4.3: Fasi dell’algoritmo METIS dato il *paper* di presentazione Karypis et al. (1995)

Dunque *METIS*, al fine di calibrare il proprio operato, tiene in considerazione esclusivamente il peso degli archi.

Data la necessità di operare su grafi non orientati, è stato reso necessario definire delle rappresentazioni alternative a quelle presentate nella Sezione 4.1. Queste verranno introdotte nella sezione successiva.

4.2.3 Grafi utilizzati nell’attività di *Community detection*

Dato il funzionamento dell’algoritmo di *Community detection*, insieme all’implementazione delle metriche di valutazione (che verranno presentate nella sezione 4.2.5), è stato reso necessario definire delle rappresentazioni alternative a quelle delle reti sociali presentate nella Sezione 4.1.

Grafi per i dati allo stato dell’arte

Per questi grafi, dati gli studi su di questi condotti, si era già in possesso di una rappresentazione non orientata.

La ragione di tale situazione è da ricercarsi nel fatto che, dati i già citati studi condotti su queste reti sociali, l’algoritmo di *Community detection* impiegato nei lavori era appunto *METIS*.

Si anticipa dunque che le modellazioni delle reti sociali di Covid-19 e *Vaccination*, data l’intenzione di utilizzo di *METIS*, saranno poste sulla falsa riga di quanto mostrato nella Sezione 4.1.2.

Dunque, per queste reti sociali è stato necessario solo effettuare una modellazione di un multigrafo rappresentante la rete mostrata in Figura 4.1; non vi sarà più una compattazione degli archi ma, in funzione al valore di “*weight*”, si avrà un determinato numero di collegamenti non orientati che congiungeranno due nodi.

La definizione di questo grafo è stata resa necessaria dall’implementazione *software* della metrica nota come *coverage*, utilizzata in questo lavoro di tesi al fine di quantificare la qualità delle partizioni effettuate.

Questa sarà introdotta formalmente nelle prossime pagine (Sezione 4.2.5).

L’utilizzo della Figura 4.4 è sufficiente al fine di mostrare la modellazione a multigrafo, rappresentante la rete sociale.

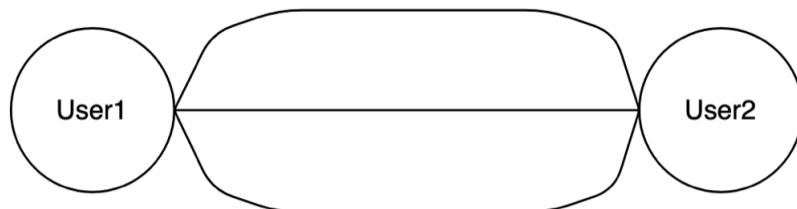


Figura 4.4: Esempio modellazione multigrafo data una singola riga dei dati utilizzati da *Garimella et al.* ipotizzando *weight* pari a 3.

Grafi per i dati Covid-19

La definizione dei nuovi grafi, per questa rete sociale, ricalca la modellazione delle reti sociali in funzione dei dataset allo stato dell’arte.

La modellazione del multigrafo è invece analoga a quanto illustrato precedentemente.

A ogni modo, partendo dal grafo utilizzato nelle attività di *Community detection*, si avrà che i nodi andranno a rappresentare sempre gli utenti e gli archi sempre i *tweet*, tuttavia vi sono delle modifiche nella strutturazione di quest’ultimi.

Gli archi non sono orientati, dunque le citazioni tra due utenti verranno compattate in un singolo arco, il cui peso sarà la somma dei pesi dei due archi; inoltre l’informazione aggiuntiva riguardante i *tweet* e i metadati degli utenti non è presente in questo grafo, rendendolo di fatto meno complesso e adatto alle operazioni di *Community detection*. La Figura 4.5 fornisce un chiarimento.

Si può osservare che la compattazione dei due archi crei un singolo arco non orientato; lo stesso, come detto in precedenza, non sarà dotato di nessun tipo di metadato associato.

Il peso dell’arco non orientato, dati due utenti, è caratterizzato dalla somma degli eventuali due archi direzionati preesistenti; questa scelta è dettata dal fatto che si ritiene corretto procedere come detto in quanto, indipendentemente dalla provenienza del maggior numero di citazioni, la somma dei pesi degli archi direzionati fornisce una stima del numero di collegamenti intercorsi tra gli utenti considerati.

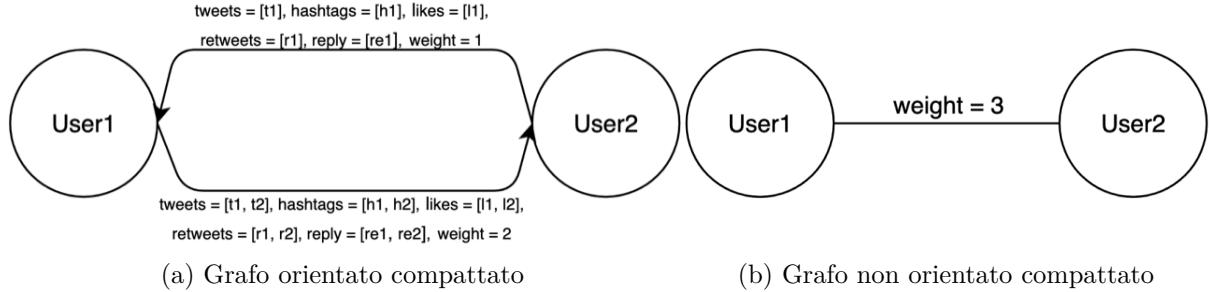


Figura 4.5: Passaggio da grafo orientato a non orientato

Si pensa che, anche nel caso in cui le citazioni seguissero un solo verso, la semplice addizione modelli correttamente la forza del nuovo collegamento prodotto, in quanto quantifica un certo livello di interazione esistente tra due utenti.

Un’ulteriore giustificazione proviene dagli studi condotti da Garimella et al. (2017), nei quali la modellazione delle relazioni di *retweet* è la medesima qui mostrata.

Risulta ora necessario, se non doveroso, sottolineare che la creazione dei differenti grafi (i.e. la rete sociale mostrata nella Sezione 4.1 e questa rappresentazione) ha seguito un percorso totalmente parallelo e non sequenziale.

Il grafo non orientato con archi dotati dell’attributo *weight*, oggetto di questa sezione, non discende dalla rete sociale precedentemente definita, ma direttamente dai dati originali.

In fase di costruzione, in funzione del grafo in oggetto, si è quindi proceduto in maniera differente una volta letti i dati dal dataset; in questo modo si è in possesso di un controllo circa la correttezza dei grafi ottenuti che, necessariamente, devono osservare lo stesso numero di nodi.

Un’ultima, ma necessaria, precisazione: questo grafo, essendo utilizzato principalmente per le attività di *Community detection*, è privo di *self-loop*. I *tweet* privi di citazione non sono presenti, in quanto inutili alle attività nelle quali questo grafo sarà protagonista.

L’ultimo grafo definito, utile principalmente alle attività di *Community detection*, è il multigrafo non orientato.

Ancora una volta si modella mediante i nodi gli utenti e con archi le *mention*; anche in questo caso le informazioni sui *tweet* e i metadati degli utenti, fatta eccezione per lo *username*, sono state eliminate.

Il valore di *weight* di un determinato arco che collegava due nodi, viene qui modelizzato con un numero di archi pari al valore *weight* dell’arco che collegava i due utenti. La modellazione di questo grafo ha seguito il paradigma introdotto precedentemente, rappresentato in Figura 4.4.

Ancora una volta, rispetto alla rete sociale, vi è una notevole diminuzione di informazione circa i nodi e gli archi.

Questo grafo, per motivi di comodità, discende dal grafo precedentemente illustrato; una volta definito il grafo non orientato con pesi sugli archi, si è proceduto valutando il

peso di ogni arco e creando un numero di archi pari allo stesso. Si è così ottenuto il grafo oggetto di questa sezione.

Si fornisce quindi un confronto quantitativo completo delle tre reti definite:

Covid-19			
Grafo	V	E	E reali
Rete Sociale	39336	85176	1186709
Grafo non orientato	39336	449678	449678
Multigrafo non orientato	39336	60163	449678

Tabella 4.4: Informazioni riguardanti la *giant component* del dataset riguardante Covid 19

Si ricorda che le notevoli differenze riguardanti il numero di archi tra il primo e il secondo grafo, è da attribuirsi sia alla traduzione da grafo orientato a non, ma soprattutto all'eliminazione, nel secondo grafo, degli archi *self-loop*.

Grafi per i dati Vaccination

In maniera del tutto analoga a quanto mostrato precedentemente, si è proceduto alla modellazione di grafi in funzione dell'utilizzo degli stessi nel *task* di *Community detection*.

Nel caso del grafo non orientato con archi pesati, l'immagine chiarificatrice mostrata in precedenza, Figura 4.5, anche in questo caso, è valida.

Allo stesso modo, dato il multigrafo non orientato, la Figura 4.4 già proposta precedentemente può chiarire ogni dubbio.

La costruzione di questi due nuovi grafi ha seguito le modalità già precedentemente citate.

In Tabella 4.5 viene fornita una visione finale quantitativa circa la struttura delle reti definite.

Vaccination			
Grafo	V	E	E reali
Rete sociale	39604	95126	221264
Grafo non orientato	39604	86848	210211
Multigrafo non orientato	39604	210211	210211

Tabella 4.5: Informazioni sui grafi ottenuti dal dataset sulle vaccinazioni

Dato il funzionamento di *METIS* e le modellazioni alternative qui definite, risulta essere adeguato l'approccio scelto, che verrà presentato nella Sezione 4.2.4, di rappresentare dell'informazione semantica a mo di nuovo peso volto ad arricchire il collegamento tra due nodi. Tale peso verrà applicato al grafo non orientato (Figura 4.5).

4.2.4 Considerare l'informazione semantica nel processo di *Community detection*

In questa sezione si introducono gli approcci utilizzati al fine di valorizzare l'informazione semantica in nostro possesso nel processo di identificazione delle comunità, passando poi alla modellazione di tale informazione al fine di modificare il comportamento dell'algoritmo di *Community detection* utilizzato.

Sentiment analysis

VADER La natura dei *social media* ha definito una serie di nuove sfide per i *task* di *Sentiment analysis* a causa della natura dei contenuti scambiati, in questo contesto si inserisce *VADER*, un modello *rule-based* proposto da C.J. Hutto et al. (2014) durante una *AAAI Conference* riguardante i *Weblogs* e i *Social Media*.

VADER (*Valence Aware Dictionary and sEntiment Reasoner*) è un modello facente parte della famiglia di soluzioni di *Sentiment analysis* con approccio basato sul lessico, particolarmente adatto a realtà *social* e di *microblogging* (per l'appunto, *Twitter*).

VADER utilizza una lista di *feature* lessicali (e.g. parole); le stesse risultano essere etichettate con un valore numerico, in funzione dell'orientamento semantico inteso come positività o negatività.

È stato riscontrato un buon livello di efficacia alla prova dei fatti, soprattutto per ciò che concerne i corpus di testi provenienti dai *social media*; sono stati infatti condotti, nel *paper* presentato, diversi confronti con altri approcci sia automatici che basati sul lessico; da tali confronti emerge come *VADER* riesca ad ottenere, mediamente, risultati migliori. In particolare, presenta svariati vantaggi, tra cui:

- Funziona molto bene sui corpus di origine *social*, riesce inoltre a generalizzare e quindi essere *domain-free*.
- Non richiede nessun tipo di *training*, poggia le sue fondamenta su un lessico curato dall'uomo, generalizzabile e *gold-standard*.
- È molto veloce e non necessita di un *text preprocessing* troppo curato in quanto riesce a trarre vantaggio da molti elementi del testo. *Stopwords* come “*but*”, punteggiatura come i punti esclamativi, parole in *upper case*, modificatori di grado (e.g. “*such as very*”, “*rather*”, “*fairly*”, “*quite*”) aiutano la classificazione.
- Il testo non deve essere suddiviso a priori in *tokens*, ma durante la fase di valutazione viene esaminato il testo valutando i *tri-gram*, al fine di identificare meglio ogni possibile inversione di polarità nel testo.
- Ha regole definite per la valutazione di *emojis*, *slangs* ed *emoticons*, in linea con la sua natura di lessico per corpus provenienti da realtà di *micro-bloggin* e *social*.

Una volta che viene richiesto di valutare una frase, *VADER* fornisce differenti score alla stessa:

- **Positivo, negativo e neutrale:** per ogni classe vi sarà un valore compreso tra 0 ed 1 che rappresenta la proporzione di testo che cade nella categoria, data la stringa fornita in input.

- **Compound:** metrica che calcola la somma di tutto il lessico, data la frase, normalizzando poi la somma in un range che va da -1 (estremamente negativo) a +1 (estremamente positivo).

Ai fini di questa tesi, sotto consiglio degli stessi sviluppatori dato il loro manifesto su *VADER*⁴, è stato deciso di utilizzare la metrica *compound* la quale fornisce una valutazione unidimensionale circa la misura di sentimento all'interno di un testo. Infine, un testo si definisce neutro nel caso in cui il suo *compound* score sia compreso nell'intervallo [-0,05, 0,05].

Modellazione sentiment Al fine di introdurre delle informazioni circa la semantica presente nei grafi Covid-19 e *Vaccination*, sono stati estratti per ogni utente i *tweet* da lui scritti valutando poi la polarità degli stessi mediante *VADER*; il processo è tenuto direttamente sul grafo rappresentante la rete sociale.

È possibile riassumere gli *step* di questa modellazione come segue:

- Per ogni nodo (quindi per ogni utente) vengono selezionati tutti gli archi uscenti e prelevati i testi dei *tweet*, in questa fase è stata posta particolare attenzione nel selezionare una sola ripetizione per *tweet*.
- In seguito, per ogni *tweet*, si è proceduto a una fase di *text preprocessing* al fine di eliminare i link contenuti nello stesso; si è poi *tokenizzato* al fine di eliminare anche determinati caratteri non presenti nel lessico di *VADER* (caratteri cinesi, caratteri in cirillico).

Terminata l'eliminazione di quei caratteri non utili all'attività di *scoring*, si è proceduto ricompattando il *tweet*.

- Il *tweet* verrà quindi valutato con *VADER*; si terrà in considerazione il *compound score* riscalando il valore in un range -30 (estremamente negativo), +30 (estremamente positivo).

Questo è stato reso necessario al fine di notificare, mediante cifra intera, le differenti polarità di ogni utente. Così facendo si ottiene anche uno spettro più ampio di *score* positivo e *score* negativo.

Per effettuare lo *scaling* è stata utilizzata la seguente formula:

$$scaling(x) = \frac{(x - oldMin) \cdot newRange}{oldRange} + newMin \quad (4.1)$$

Si è quindi utilizzata una conversione lineare.

- Una volta processato ogni *tweet*, dato il nodo preso in considerazione, il valore di *sentiment* associato all'utente è calcolato come segue:

$$S(user) = \frac{1}{n} \cdot \sum_{i=1}^n scaling(x_i) \quad (4.2)$$

Con *n* si intende il numero di *tweet* univoci riconducibili all'*user*.

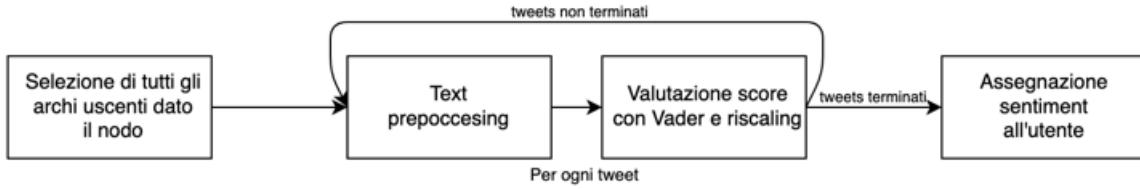


Figura 4.6: Pipeline di assegnazione score *sentiment*

La Figura 4.6 riassume il processo appena definito.

Una volta assegnato ad ogni utente uno score di sentimento, il quale potrà variare tra i valori limite +30 e -30, si è provveduto ad arricchire gli archi del grafo non diretto, mostrato in Figura 4.5, con un ulteriore peso.

Prima di tutto, dati due utenti congiunti da un arco, si calcolerà la *Sentiment Similarity* come segue:

$$SS(user1, user2) = 60 - |S(user1) - S(user2)| \quad (4.3)$$

Si va dunque ad attribuire un valore da 0 fino al 60, in funzione della similarità tra i sentimenti; utenti con sentimento simile vedranno un valore di *SS* maggiore rispetto a utenti con sentimento discordante. Questo è volto a premiare e rendere più forti utenti con sentimento simile.

A questo punto, con l'obiettivo ultimo di dare maggior importanza al sentimento, si è provveduto, al fine di calcolo del nuovo *weight* in funzione del sentimento, a valutare la presenza (in senso binario) dell'arco tra due utenti, indipendentemente dal peso di questo.

Si è quindi provveduto ad attribuire, data la presenza di un collegamento tra due utenti, un nuovo peso (*Sentiment Weight*) che tenesse in considerazione la similarità di *sentiment* associata ai due utenti:

$$SW(user1, user2) = weight_{binary}(user1, user2) + SS(user1, user2) = 1 + SS(user1, user2) \quad (4.4)$$

Avremo quindi che utenti che si sono scambiati pochi *tweet* ma che hanno un *sentiment* simile, vedranno un incremento del peso dell'arco che li congiunge; invece, utenti con *sentiment* diverso, non vedranno particolari incrementi sul peso dei loro archi.

Si ipotizza che le eventuali differenze di sentimento tra gli utenti, possano migliorare il *task* di *Community detection* permettendo di identificare in maniera più puntuale le *community* rispetto a un approccio strettamente strutturale.

Si pensa che le *community* dovrebbero includere al loro interno utenti con sentimento simile, si vuole quindi cercare la contrapposizione mediante la valutazione del *sentiment* medio per ogni utente.

Topic Modeling

Definizione del modello LDA Una delle tecnologie di *topic modeling* più utilizzata e applicata è *LDA* (*Latent Dirichlet Allocation* 2003); l'assunzione di base del modello è: ogni documento ha più *topic*, ogni *topic* è una distribuzione su di un insieme di parole, ogni documento è un insieme di termini appartenenti a diversi possibili *topic*, una parola

⁴<https://github.com/cjhutto/vaderSentiment>

può appartenere a più *topic*.

Dunque un *topic* è visto come una distribuzione di parole $W_{1,N}$ sopra un vocabolario (il vocabolario consiste in un insieme N di parole utilizzate per rappresentare il corpus di documenti). Sono differenti le distribuzioni che caratterizzano statisticamente *LDA* e, più in generale, il *topic modeling*; viene fornita una veloce panoramica.

Un *topic* è generalmente identificato dalla sua distribuzione di probabilità condizionata:

$$P(W_{i,N}|TOPICX) \quad w \in W_{i,N} \quad (4.5)$$

Ogni documento d dato un corpus di documenti D è associato con una *document-topics distribution*:

$$\forall d \in D \quad P(T_{1:K}|d) \quad (4.6)$$

Abbiamo inoltre, per un corpus D , una *corpus-topics distribution*:

$$P(T_{1:K}|D) \quad (4.7)$$

Le distribuzioni sopra mostrate vengono apprese grazie a *LDA*.

Una possibile rappresentazione del modello *LDA*, mediante un grafico a “piattini”, può essere quella fornita dalla Figura 4.7.

Possiamo dunque notare la presenza di differenti elementi, si fornisce una rapida panoramica degli stessi:

- α : iper-parametro che rappresenta, per ogni documento la probabilità a priori di appartenere a un determinato *topic*. Solitamente questo valore è impostato identico per ogni *topic*.
- β : iper-parametro che rappresenta la distribuzione di probabilità a priori per l'appartenenza di una certa parola a un certo *topic*. Solitamente questo valore è impostato identico per ogni parola.
- Z : probabilità del documento di appartenere a un determinato *topic*.
- W : parola che ci interessa “classificare”, come appartenente a un certo *topic*.
- N : numero di parole dato il documento.
- θ : *topic distribution* per un certo documento m .
- M : numero di documenti.
- ϕ : word distribution per un certo *topic* k .
- K : numero di *topic*.

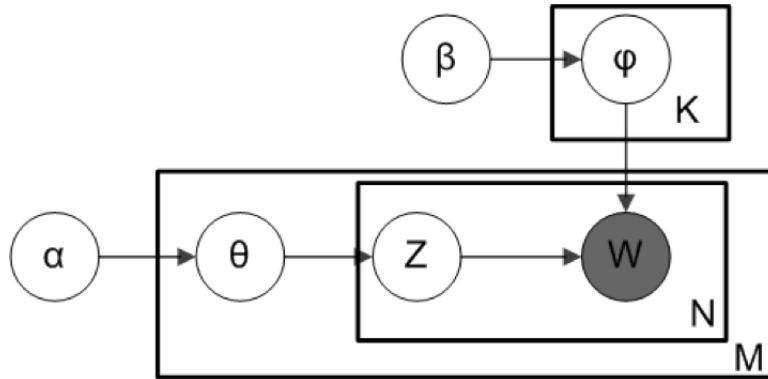


Figura 4.7: Rappresentazione modello *LDA*^[12]

Modellazione *topic* Al fine di introdurre, all'interno della modellazione del grafo, informazioni circa i *topic* disquisiti dai diversi utenti, è stato utilizzato il modello *LDA* sviluppato e mantenuto dal progetto *MALLET* di Andrew Kachites McCallum (2002).

L'obiettivo di tale inserimento è presto spiegato: si ipotizza che utenti che parlano degli stessi *sotto-topic*, dato il *topic* della rete sociale e.g. *Covid-19* o *Vaccination*, siano posizionati all'interno della stessa *community*.

Le operazioni effettuate, per ciò che concerne il *preprocessing*, sono state effettuate sulla rete sociale dotata di tutte le informazioni semantiche; le attività stesse sono le seguenti:

- Dati tutti gli archi, si è provveduto ad estrarre ogni singolo *tweet* presente nella rete, in questa fase si è posta attenzione all'univocità di ogni *tweet*.
- Si è effettuata una *tokenization*, questo è stato reso possibile utilizzando il *Tweet-Tokenizer* offerto da *NLTK* (2009); questo ha permesso una migliore definizione dei *token*, tenendo in considerazione la natura *social* propria dei *tweet*.
- Si è provveduto all'eliminazione delle *stopword* inglesi, si è inoltre definito un insieme proprietario di *stopword* al fine di inserire parole in carattere cirillico presenti; si è inoltre provveduto ad eliminare tutti gli *URL* presenti.

Questa fase è analoga a quanto fatto nella valutazione del *sentiment*, precedentemente mostrata.

- In ultimo è stato predisposta un'attività di *stemming*, mediante il *PorterStemmer* definito da Michael Porter (1980).

La decisione di definire uno *stemmer* è volta a ridurre le possibili rappresentazioni della stessa parola, riducendo tutto alla forma radice, detta “tema”.

A questo punto si dispone di una serie di *tweet*, quindi di documenti, “puliti” sui quali attuare le operazioni volte a identificare i *topic*; al tal proposito, è stata utilizzata la libreria *Gensim* (2010) che ha permesso di implementare un modello *LDA Mallet*.

Dunque, una volta ottenuti i documenti, è stato definito un dizionario con un'associazione circa la frequenza di apparizione, di una ogni *token* identificato. Si è anche

provveduto, al fine di eliminare i *token* non utili data la frequenza, eliminando tutti quei termini che comparivano in meno del 2% dei documenti e in più del 50% dei documenti, infine si è deciso di tenere in considerazione solo i primi 100.000 *token* data la frequenza.

Queste operazioni sono state rese necessarie in quanto, in caso di non applicazione delle stesse, i tempi computazionali subivano un incremento significativo; si è ritenuto non necessario tenere in considerazione parole che, date le loro frequenze, potessero risultare poco utili ai fini del *topic modelling*.

Una volta definito il vocabolario è stato dunque allenato il modello, il numero di *topic* ha visto la definizione di un range partendo da 2 arrivando fino a 30.

Per ogni possibile numero di *topic*, si è valutata la *Topic Coherence* e salvato il modello; il salvataggio del modello è reso necessario in quanto, al fine di identificare il numero ottimale di *topic*, si sono valutati sia una metrica circa la qualità dei *topic* (*Topic Coherence*), sia il senso e i termini che comparivano nei diversi *topic*, mediante quindi un approccio “*human-based*”.

Ottenuto il modello ottimale, si è provveduto a valutare i *topic* dei singoli utenti definendo, in funzione dei *tweet* scritti dall’utente, un insieme di *topic* associati allo stesso.

Ogni utente, in seguito alle operazioni di *topic modelling*, sarà dotato di un insieme contenente l’identificativo di ogni *topic* associato allo stesso dati i suoi *tweet* da lui pubblicati; l’associazione di un *topic*, dato un *tweet*, si basa sulla valutazione della probabilità restituita dal modello: un certo *tweet* parlerà di un certo *topic*, solo se quel *topic* presenterà probabilità più alta dato il *tweet*.

Dopo aver assegnato ad ogni utente, dati i *tweet* scritti, un certo insieme di *topic* si è proceduto a codificare la similarità tra gli utenti, dati i *topic* discussi. Dati due utenti collegati da un arco, si è calcolata una *Topic Similarity* in questo modo:

$$TS(user1, user2) = \#TotalTopic - card((TOPIC(user1) \Delta TOPIC(user2))) \quad (4.8)$$

Al fine di assegnare il nuovo peso all’arco (*Topic Weight*), si è proceduto, in maniera del tutto analoga a quanto mostrato precedentemente (Sezione 4.2.4), valutando un la presenza dell’arco strutturale con l’aggiunta della *Topic Similarity* precedentemente calcolata:

$$TW(user1, user2) = weight_{binary}(user1, user2) + TS(user1, user2) = 1 + TS(user1, user2) \quad (4.9)$$

Modellazione Ibrida

Data la definizione, tra loro indipendenti, di una modellazione che facesse leva sul *sentiment* e una che facesse leva sui *topic*, è stato deciso di utilizzare contemporaneamente queste due informazioni, creando di fatto un nuovo peso.

L’obiettivo di tale modellazione era quello di “allontanare” due utenti i quali, parlando dello stesso *topic*, utilizzassero sentimento diverso; viceversa si voleva “avvicinare” ulteriormente utenti con stessi *topic* e con sentimento simile.

La modellazione ibrida ha visto la definizione di un ulteriore peso (*Hybrid Weight*),

lo stesso verrà successivamente utilizzato nelle attività di *Community detection*. Il peso è, di fatto, una combinazione delle due misure di similarità precedentemente definite:

$$HW(user1, user2) = TW(user1, user2) + SW(user1, user2) \quad (4.10)$$

4.2.5 Metriche di valutazione topologiche

Al fine di valutare la qualità delle *community* ottenute, sono state introdotte diverse valutazioni riguardanti sia aspetti circa l'esecuzione e la dimensione delle partizioni ottenute, sia metriche più specifiche come la modularità e la copertura.

La valutazione combinata di modularità e copertura assicura un'visione completa circa la qualità della suddivisione effettuata dall'algoritmo.

La valutazione della modularità, data una determinata partizione, viene effettuata sul grafo non diretto con pesi sugli archi (Figura 4.5); la valutazione della copertura, data una determinata partizione, viene effettuata sul multigrafo non diretto (Figura 4.4).

Modularità

La modularità è una misura volta a quantificare la qualità della suddivisione di una rete; modella quindi la correttezza della suddivisione in *community*.

Grafi con un alto valore di tale misura saranno considerati come divisi in sotto-strutture (quindi in *community*) dense cioè dotate di un buon livello di connessione interno, ma tra loro sparse e dunque con meno collegamenti extra-community.

Oltre che a validare l'attività di partizione di un grafo, la modularità è utilizzata come misura da ottimizzare alla base di diversi algoritmi di *Community detection*.

Possiamo quindi vedere la modularità di una partizione come una quantità scalare che misura la densità di archi all'interno delle comunità individuate in paragone al valore atteso di tale densità.

L'idea alla base è: quantificare la misura in cui una determinata partizione, in comunità, presenta una tendenza sistematica ad avere più collegamenti intracomunitari rispetto a quelli che avrebbe, la stessa suddivisione in comunità, se gli archi fossero ottenuti utilizzando un modello casuale. Valori positivi della modularità indicano che una frazione statisticamente valida di archi in una rete cade all'interno delle comunità individuate dalla partizione^[58].

La misura di modularità utilizzata in questa tesi prende il nome di *Louvain Modularity* e fa riferimento al paper di Newman et al. (2004).

Per un grafo pesato, come nel caso delle strutture utilizzate in questa tesi, la modularità viene calcolata come segue:

$$M = \frac{1}{2m} \sum_{i,j} \left[A_{i,j} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (4.11)$$

Dove:

- $A_{i,j}$ rappresenta il peso tra i nodi i e j .

- k_i e k_j rappresentano il grado dei nodi i e j (e.g. il numero di archi incidenti nel nodo).
- m è la somma di tutti i pesi degli archi nel grafo.
- c_i e c_j sono le community dei due nodi i e j .
- δ è la funzione Kronecker delta tale per cui:

$$\delta_{ij} = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{altrimenti} \end{cases} \quad (4.12)$$

Tale misura può assumere valori compresi tra $[-\frac{1}{2}, 1]$, tanto sarà maggiore il valore, tanto le strutture identificate e che si sta validando sono lontane da un modello casuale.

In fase di calcolo di tale metrica, viene indicato il nome del peso degli archi, facente riferimento al forza delle connessioni intercorse tra due utenti, come parametro da utilizzare al fine di pesare correttamente il valore di un determinato arco.

Copertura

La copertura, date delle *community* definite su di un grafo, è un valore scalare volto a quantificare la percentuale di archi che cadono all'interno di una delle *community*; metrica proposta e basata al lavoro di Santo Fortunato (2010).

Ipotizzando due *community*, tale valore è così quantificato:

$$C = \frac{|E_1| + |E_2|}{|E|} \quad (4.13)$$

Dove $|E_1|$ e $|E_2|$ sono il numero di archi che cadono all'interno, rispettivamente, della *community* uno e due; $|E|$ è invece il numero totale di archi presenti nella rete.

Tale indice varia in un range $[0, 1]$, dove valori tendenti al massimo notificano che la maggior parte degli archi risulta interno alle *community*.

Questa metrica, da sola, non è sufficiente ad affermare che un certo partizionamento è totalmente corretto; tuttavia, affiancando tale valore a quello di *modularity*, si avrà una valutazione più puntuale circa la qualità della attività di *Community detection*.

4.3 Quantificazione controversia

In questa sezione si introducono le metriche utilizzate per l'attività di quantificazione della controversia, su tutti i grafi precedentemente introdotti (Sezione 4.3.1).

Tale attività, la quale è profondamente condizionata dai risultati ottenuti dal *task* di *Community detection*, è fondamentale al fine di comprendere se vi sia effettivamente presenza di controversia date le partizioni identificate.

In seguito verrà proposta una riflessione circa le assunzioni che le metriche presentate, prendono al fine di computare il valore volto a quantificare la controversia (Sezione 4.3.2)

4.3.1 Metriche

Random Walk Controversy

Questa misura usa la nozione di *random walk* sul grafo al fine di quantificare la controversia; l'idea di base è che in una discussione, dunque dato un *topic*, che porta alla creazione di visioni opposte e polarizzate, ci siano dei nodi “autorevoli” da entrambi gli schieramenti; questa misura va dunque a catturare quanto è probabile che un utente casuale, di una certa comunità, possa essere esposto al contenuto espresso da un nodo “autorevole” della comunità opposta.

Dato quindi il grafo, definito nel primo stage data la Figura 3.1, $G=(V, E)$ e le due partizioni X e Y ottenute nel secondo stage, si va a definire per entrambe le comunità i nodi con un alto valore di grado; si assume dunque che la centralità di grado possa essere utilizzata come indice di autorevolezza del nodo.

Si seleziona dunque un vertice in maniera casuale e si considera il *random walk* che parte da questo; un cammino termina quando si raggiunge uno dei nodi “autorevoli” precedentemente identificati.

Viene quindi definita la *Random Walk Controversy* (RWC) la differenza di probabilità di due eventi: (i) i *random walk* partiti da una partizione finiscono nella stessa e (ii) i *random walk* partiti da una partizione terminano nell'altra. La misura è così quantificata:

$$RWC = P_{XX}P_{YY} - P_{XY}P_{YX} \quad (4.14)$$

Dove P_{AB} con $A, B \in X, Y$ è una probabilità condizionata così definita:

$$P_{AB} = Prob[inizia nella partizione A | termina nella partizione B] \quad (4.15)$$

Il valore di questa metrica varia tra -1 e 1, più è vicino a -1 più sarà probabile passare all'altra partizione (assenza di controversia), più è vicina 1 più sarà probabile rimanere nella partizione di partenza (presenza di controversia). Tale metrica è stata introdotta nello studio condotto da Garimella et al. (2017), sulla quantificazione della controversia.

Di fatto questa misura vuole andare a quantificare quanto sia probabile che, dato un movimento totalmente casuale, ci si ritrovi esposti a un esponente “autorevole” della partizione opposta; tale movimento va di fatto a rappresentare una navigazione, mediante le *mention* modellate dalle varie reti sociali definite, dell'utente sui diversi tweet presenti nella rete.

In questo lavoro di tesi si è provveduto a inserire un meccanismo di *restart*: se dopo un *random walk* di lunghezza pari all'*average shortest path* del grafo moltiplicato per due, non si fosse raggiunto ancora nessun nodo “autorevole” di nessuna delle due partizioni, il risultato del nodo analizzato non viene conteggiato e si passa alla selezione casuale di un altro nodo.

Il numero di nodi selezionati per entrambe le partizioni è equivalente, lo stesso ammonta al 60% dei nodi facenti parte di ogni partizione; infine: un nodo viene definito “autorevole” se il grado dello stesso, ordinato insieme agli altri in senso decrescente, si posiziona nella top 15% data la comunità di appartenenza.

Boundary Connectivity

Questa metrica, utilizzata nel paper Garimella et al. (2017), è da attribuire a Guerra et al. (2013); la misura si basa sul concetto di “vertici interni” e di “frontiera”.

Definiamo $u \in X$ un vertice nella partizione X ; u appartiene alla “frontiera” se e solo se è connesso con almeno un vertice della partizione Y e con almeno un vertice della partizione X che non è però connesso con nessun vertice della partizione Y . Viene quindi definito l’insieme dei nodi “frontiera” come $B = B_x \cup B_Y$, al contrario i vertici $I_X = X - B_X$ sono chiamati “vertici interni” della partizione X (in maniera del tutto analoga si definisco i “vertici interni” per la partizione Y). Si definisce quindi insieme dei “vertici interni” l’insieme $I = I_x \cup I_Y$.

Se le due partizioni dovessero costituire una realtà di camere d’eco, l’insieme B dovrebbe essere costituito da vertici che sono più fortemente connessi con gli elementi dell’insieme I piuttosto che con nodi dell’insieme B . La seguente formula formalizza il concetto appena espresso:

$$GMCK = \frac{1}{|B|} \sum_{u \in B} \frac{d_i(u)}{d_b(u) + d_i(u)} - 0.5 \quad (4.16)$$

Con $d_i(u)$ si vuole definire il numero di archi tra il vertice u e gli elementi dell’insieme I , con $d_b(u)$ si definisce il numero di archi tra il vertice u e gli elementi dell’insieme B . Questa metrica fa quindi pressione sul fatto che le partizioni identificate siano poco interconnesse tra loro, la valutazione tiene in considerazione tutti i nodi presenti nel grafo al fine di identificare correttamente l’insieme I e l’insieme B .

Random Walk Top Degree

Questa misura di controversia, introdotta in questo lavoro di tesi, deriva direttamente da quella proposta nel lavoro di Garimella et al. (2017) e presentata all’inizio di questa sezione.

Difatti il tipo di misurazione e le modalità di attuazione sono del tutto identiche, l’unica differenza riguarda i punti di partenza dei *Random Walk*; se prima la selezione era del tutto casuale tra gli utenti delle due diverse partizioni, ora si parte solo dai nodi definiti “autorevoli”, dove la definizione di autorevole è la medesima fornita precedentemente.

In maniera del tutto analoga a quanto già visto, il *random walk* terminerà una volta che verrà raggiunto un vertice facente parte dell’insieme dei nodi “autorevoli” di una o dell’altra partizione.

In questo modo si cerca di quantificare quanto i nodi autorevoli di una partizione siano esposti a individui simili, ma appartenenti alla partizione opposta.

Change Side Controversy

Ultima metrica proposta e utilizzata in questa tesi, la *Change Side Controversy* si basa ancora una volta sul concetto di *Random Walk* ma agisce in maniera diversa nel conteggio e nell’attuazione del cammino stesso.

Per ogni partizione vengono selezionati casualmente un numero pari al 60% dei vertici facenti parte di ognuna; viene quindi quantificato l'*average shortest path* e moltiplicato per due, questo valore andrà a definire la lunghezza del *Random Walk* che verrà effettuato da ogni nodo.

Quindi si sarà dotati di un insieme N di nodi dai quali far partire la valutazione. La misura di controversia sarà quindi così quantificata:

$$CSC = \frac{\sum_{v \in N} 1 - (\frac{\kappa(v)}{l_G * 2})}{|N|} \quad (4.17)$$

Dove l_G è l'*average shortest path* per il grafo G sul quale si sta misurando la controversia, $\kappa(v)$ è invece il numero di passi del *Random Walk*, partiti da ogni v , che hanno visto la *community* del nodo di arrivo diversa dalla *community* del nodo di partenza.

Con questa metrica si cerca di osservare il comportamento del *Random Walk*, considerando quanti cambi di partizioni sono effettuati durante lo stesso.

Tale metrica varia in una range 0/1, nel caso in cui questa assuma valore massimo, significa che non vi è stato mai un cambio di *community* durante tutto il *Random Walk*; questo permette di affermare se vi è una segregazione nella rete o meno, in modo tale da poter verificare la presenza di controversia.

4.3.2 Discussione

Verificare la presenza di controversia non è un compito semplice per una serie di ragioni, in primo luogo dipende tutto dall'attività di *Community detection*: nel caso in cui questa non sia accurata (a causa dell'algoritmo utilizzato, ad esempio), la valutazione della controversia ne è subito influenzata, producendo risultati attinenti data la partizione non ottimale identificata; per risolvere questa problematica risulta necessario tornare sull'algoritmo volto a partizionare il grafo.

In secondo battuta, le diverse assunzioni circa la controversia possono produrre diverse metriche (come poi mostrato in questa tesi e in altri lavori di letteratura), le stesse possono andare a catturare una porzione della definizione della misura di *controversy*, lasciando però inesplorate altre.

Ad esempio: la RWC esegue un conteggio in funzione dell'ultimo nodo raggiunto, supponendo dunque che questo, facente si parte di un certo insieme autorevole identificato, possa risultare “influenzante” al fine di presa di posizione, dato un utente, su una certa questione o argomento.

Questo non è però sempre vero: potrebbe essere più importante il tipo di informazioni a cui si è stati esposti per tutto il *Random Walk*, di conseguenza l'ultimo passo dello stesso potrebbe essere insufficiente a definire correttamente una misura volta a quantificare la controversia.

Tuttavia questa è sempre una supposizione, come una supposizione è misurare quanto ben collegati possano essere i nodi “influenti”, ipotizzando che meglio questi risultino essere collegati, più *bipartisan* sia l'informazione condivisa, informazione alla quale gli utenti “normali” saranno esposti.

Si ritiene perciò più utile un approccio volto a tenere in considerazione, nello stesso momento, più misure diverse; in questo modo, ogni possibile frazione circa la definizione della controversia, e di come questa possa essere percepita, risulterà essere tenuta in considerazione; si cercherà perciò di valutare tutte le misure qui citate, al fine di avere una visione completa della situazione.

4.4 Riduzione controversia

L’attività successiva, in seguito alla verifica circa la presenza di controversia, è stata quella di riduzione della stessa; come precedentemente affermato, si è deciso di definire una strategia basata sull’attività di *Link prediction*.

In questa sezione verranno formalizzate le metriche volte a quantificare uno score per i potenziali archi, tenendo in considerazione fattori di prossimità e di centralità (Sezioni 4.4.1 e 4.4.2).

Si passerà poi a definire il modo in cui si è voluto modellare un moltiplicatore in funzione della definizione di una “probabilità di accettazione” (Sezione 4.4.3), infine si mostrerà come si è progettata la strategia di aggiunta archi (Sezione 4.4.4).

4.4.1 Soluzioni basate sulla prossimità

Vengono ora presentati gli indici basati sulla prossimità utilizzati in questo lavoro di tesi; tutti gli indici sono presi dalla letteratura.

Coefficiente di Jaccard

Il coefficiente di Jaccard è un indice utilizzato per quantificare la similarità in tantissimi e diversi campi, nel contesto nel quale ci troviamo servirà per attribuire lo *score* a un potenziale arco, tenendo in considerazione il rapporto tra l’intersezione dei vicini comuni di due nodi e l’unione di questi.

Proposto da Liben-Nowell et al. (2004) nel contesto della *Link prediction*, l’utilizzo dello stesso fornisce un punto di riferimento circa l’efficacia di indici che valutano la prossimità.

Il valore attribuito ai potenziali archi è dato dalla seguente formula:

$$JC_{(x,y)} = \frac{\Gamma(x) \cap \Gamma(y)}{\Gamma(x) \cup \Gamma(y)} \quad (4.18)$$

Dove con $\Gamma(u)$ denota l’insieme dei vicini di u .

Si è dunque proceduto, dati gli archi non presenti nella rete e che congiungessero nodi facenti parte delle due partizioni opposte, a definire un insieme $\Psi = \{(x, y, JC_{(x,y)})\}$ contenente, come si deduce, i possibili archi da aggiungere con il relativo *score*.

Adamic Adar Index

L’Adamic Adar Index proposto da Adamic et al. (2001), raffina il semplice conteggio dei vicini comuni tra due nodi, andando considerare come più importanti i nodi con un numero limitato di vicini.

Il concetto è quindi il seguente: elementi comuni dotati di un vicinato di notevoli dimensioni (e.g. *influencer*) risultano essere meno significativi ogni volta si preveda una connessione tra due nodi, rispetto agli elementi condivisi con un vicinato più ridotto.

L'indice è quindi definito come segue:

$$AAI_{(x,y)} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{\log|\Gamma(z)|} \quad (4.19)$$

Dove $\Gamma(u)$ denota l'insieme dei vicini di u , questo indice porta a una divisione di zero per i nodi collegati solo tramite *self-loop*. È dunque concepito per essere utilizzato quando non sono presenti questi tipo di collegamenti.

Sempre valutando gli archi non presenti tra nodi delle partizioni opposte, si è ottenuto un insieme A contenente tutti i possibili archi con i relativi score dell'*Adamic Adar Index*.

4.4.2 Soluzioni basate sulla centralità

Si passa ora alla definizione delle soluzioni volte a valutare le centralità e le peculiarità dei nodi, al fine di modellare uno *score* che verrà tenuto in considerazione al fine della *Link prediction*.

Top Degree Index

Questa prima misura, proposta in questo lavoro di tesi, ha come obiettivo quello di collegare, nel caso in cui non lo fossero già, i nodi delle due partizioni che osservano un valore alto del grado, data la partizione stessa. Questo ha come obiettivo cercare di collegare quei nodi che sono di fatto degli *hub* nella *community* di riferimento.

A tal fine, date le partizioni P_1 e P_2 , vengono definiti due insiemi TD_{P_1} e TD_{P_2} , dove $v \in TD_i$ se e solo se $v \in P_i$ e osserva un grado maggiore del 75% dei nodi della sua partizione.

Una volta identificata tali nodi si procede a verificare per quanti, prelevati a coppie $(v_x, v_y) | v_x \in TD_i \wedge v_y \in TD_j \wedge i \neq j$, sia già presente un arco che li unisce.

Nel caso in cui tale arco non dovesse sussistere, viene calcolato lo score dell'eventuale arco secondo la seguente formula:

$$TDI_{(x,y)} = \text{Degree}(x) + \text{Degree}(y) \quad (4.20)$$

Alla fine della procedura si sarà identificato un insieme E composto da una serie di tuple così formate $(x, y, TDI_{(x,y)})$; tale insieme è costituito dai possibili archi con il relativo score definito dal *Top Degree Index*.

Come inizialmente accennato, tale indice ha come scopo l'identificazione dei nodi che partecipano in maniera più attiva, dato il *topic* modellato, alla discussione.

Collegare nodi, per ragioni di grado “autorevoli”, si ipotizza possa influenzare in senso indiretto i nodi non facenti parte dell'insieme TD ; l'obiettivo ultimo è quello di omogenizzare il contenuto condiviso dai nodi, che più probabilmente risultano essere influenti nella rete.

Top Betweenness Index

In maniera analoga a quanto visto precedentemente, questo indice proposto nel lavoro di tesi vuole collegare quei nodi i quali osservano un valore di centralità, secondo la *betweenness*, alto.

Nella fattispecie si è proceduto, per ogni nodo del grafo, al calcolo della misura di centralità tenendo in considerazione, per motivi strettamente computazionali, il 60 % dei vertici.

Una volta identificati i valori di centralità per ognuno, si è definito un insieme $TB = \{v \in B \Leftrightarrow B(v) \geq \overline{B(V)}\}$ dove con $\overline{B(V)}$ si intende la media della misura di centralità in oggetto.

Una volta definito l'insieme TB si valutano coppie di nodi non facenti parte della stessa partizione e non collegati da nessun arco, per ognuno viene calcolato il *Top Betweenness Index* come segue:

$$TBI_{(x,y)} = \text{Betweenness}(x) + \text{Betweenness}(y) \quad (4.21)$$

Si otterrà dunque un insieme $B = \{(x, y, TBI_{(x,y)})\}$ composto, in maniera analoga a prima, da una serie di tuple.

In questo modo si cerca di collegare tra di loro quei nodi che di fatto risultano essere dei “passaggi obbligati”, nel tentativo di rendere più probabile un passaggio da una partizione all’altra.

Top to Normal Index

Ultimo indice proposto in questa tesi, il *Top to Normal Index* è l'indice più promettente sulla carta; il compito di questo è quello di collegare i nodi ritenuti “normali” per grado, dunque non facenti parte dell'insieme $TD = TD_{P_1} \cup TD_{P_2}$, con quei nodi che invece fanno parte dell'insieme stesso.

L'obiettivo è quindi quello di collegare direttamente gli utenti di una partizione ai nodi “autorevoli” per grado dell'altra, con l'obiettivo ultimo di esporre, in maniera più *bipartisan*, gli utenti a entrambe le partizioni identificate.

Si è proceduto quindi, in maniera analoga a quanto già fatto per il *TDI*, a definire per ogni partizione il relativo insieme TD_{P_i} ; a questo punto si è proceduto valutando le coppie:

$$(v_x, v_y) | v_x \in TD_{P_i} \wedge v_y \in P_j \wedge i \neq j \wedge v_y \notin TD_{P_j} \quad (4.22)$$

Che non presentavano archi, per ognuna di queste è stata quindi computata il *Top to Normal Index* come segue:

$$TNI_{x,y} = \text{Degree}(x) \quad (4.23)$$

Definendo infine un insieme $\Upsilon = \{(x, y, TNI_{(x,y)})\}$; il valore $\text{Degree}(x)$ che caratterizza il *TNI*, va ad indicare il valore del grado facente riferimento al nodo appartenente all'insieme TD .

4.4.3 Modellazione *sentiment*

Come introdotto all'inizio di questa sezione, riprendendo poi quanto espresso nel lavoro di Garamella et al. (2017); la semplice aggiunta di archi valutata in funzione di un qualche indice, risulta essere un approccio strettamente teorico che va quasi a semplificare la complessità di tale attività, soprattutto nel contesto delle reti sociali.

L'aggiunta "forzata" di un arco non è veritiera data la realtà di interesse, risulta perciò necessario provare a inserire una metrica volta a quantificare, in una qualche maniera, una sorta di "probabilità di accettazione" del collegamento proposto.

Con tale obiettivo si è scelto di implementare un cosiddetto "*Sentiment Boost*", volto a modificare gli score precedentemente definiti; si è quindi ipotizzato che la *sentiment similarity* potesse essere un valido indice volto a quantificare la probabilità che un arco potesse realmente esistere.

Dato un generico insieme di possibili archi $O = \{(x, y, Score_{(x,y)})\}$, si è analizzato ogni singolo arco, concentrandosi sulle due estremità dello stesso, dunque sugli utenti che l'arco si apprestava a collegare.

Si è dunque valutata la parte intera del *sentiment similarity* $SS_{(x,y)}$ che il potenziale arco avrebbe registrato, andando poi a valutare quanti archi presenti nel grafo presentassero, sempre concentrandosi sulla parte intera, tale valore. Il tutto si sintetizza con la seguente formula:

$$SB_{(x,y)} = \frac{|E_{SS_{(x,y)}}|}{|E|} \quad (4.24)$$

Dove con $|E_{SS_{(x,y)}}|$ si intende il numero di archi presenti che osservano una *sentiment similarity* identica a quella del potenziale nuovo arco, con $|E|$ il numero totale di archi tenuti in considerazione.

A questo punto, dato l'insieme di possibili archi $O = \{(x, y, Score_{(x,y)})\}$, si è ottenuto un nuovo insieme O_{SB} costituito da una serie di tuple così formate $(x, y, Score_{(x,y)} * SB_{(x,y)})$.

In questo modo si cerca di tenere in considerazione la natura sociale della rete in esame, inoltre è interessante notare come l'aggiunta di tale moltiplicatore possa poi influire nella riduzione della controversia identificata.

4.4.4 Aggiunta archi

In quest'ultima sezione verrà velocemente spiegato il come si è proceduto all'inserimento degli archi dato un generico insieme $O \in \{\Psi, A, E, B, \Upsilon\} \cup \{\Psi_{SB}, A_{SB}, E_{SB}, B_{SB}, \Upsilon_{SB}\}$; si ricorda che con la dicitura X_{SB} si intende un insieme di possibili archi, data una certa metodologia, che tiene in considerazione il *Sentiment Boost*.

Date le tuple caratterizzanti l'insieme stesso, si è proceduto ad ordinare in senso decrescente per il terzo elemento di ognuna, dunque per lo score che l'arco presenta; in questo modo si inizierà ad aggiungere dall'arco che ha registrato uno score più alto, dato l'indice caratterizzante l'insieme preso in esame, in poi.

Prima di procedere all'effettivo inserimento si è valutato in che misura mediamente, dato il grafo preso in esame, due nodi risultassero essere tra loro collegati; in questo modo

è stato possibile quantificare in maniera verosimile il parametro *weight* volto a definire il numero di *mention* intercorse tra i due utenti che verranno collegati.

In ultimo, sempre tenendo in considerazione l'obiettivo finale di un aumento minimo del numero di collegamenti, si andranno ad aggiungere un numero di archi, quindi una somma dei *weight*, tali per cui non si veda un aumento superiore al 30% degli archi già esistenti; ogni qual volta che si sarà incrementato del 2% la quantità di archi, verrà quantificata la controversia al fine di verificare la riduzione registrata durante tutto il processo.

Capitolo 5

Risultati

In questo capitolo vengono mostrati e discussi i risultati ottenuti relativamente a: (*i*) il partizionamento del grafo, (*ii*) la quantificazione dei livelli di controversia che portano alla definizione delle camere d'eco e (*iii*) il superamento delle stesse.

In particolare il capitolo è composto dalle seguenti sezioni:

- Nella Sezione 5.1 vengono mostrati i risultati delle attività di *Community detection*, gli stessi comprendono anche tutte quelle attività “accessorie” all’attuazione del *task* (e.g. *topic modelling*)
- Nella Sezione 5.2 sono riportate tutta una serie di analisi effettuate, data la metodologia di partizionamento adottata, sulle *community* identificate. Queste si sono concentrate unicamente su quelle reti delle quali eravamo in possesso di metadati e informazioni associati ai contenuti scambiati.
- Nella Sezione 5.3 si forniscono, per ogni rete sociale definita, i risultati circa la quantificazione della controversia. In questa fase sono stati utilizzati tutti gli indici presentati precedentemente.
- Nella Sezione 5.4 vengono mostrati i risultati ottenuti rispetto all’attività di *Link prediction*; viene anche fornita una panoramica sulla presenza di strutture interne sulle nuove reti “arricchite” di nuovi archi, in funzione della metodologia di *linking* migliore.

5.1 Community detection

In questa sezione si riportano tutti i risultati circa l’attività di *Community detection* con tutte le sotto attività a essa collegata.

Si procede mostrando i risultati, per ogni rete sociale precedentemente definita, circa il partizionamento della rete; fornendo i valori di tutte le metriche volte a quantificare la qualità delle *community* identificate, insieme a informazione di carattere qualitativo e quantitativo delle stesse 5.1.1.

In seguito si forniscono i risultati ottenuti in funzione del *task* di *topic modelling* (Sezione 5.1.2), quest’attività ha permesso di definire un peso circa la “similarità di *topic*” tra gli utenti, peso utilizzato poi per il partizionamento del grafo.

5.1.1 Risultati partizionamento delle reti sociali

Ukraine

Questo grafo, stando agli studi riassunti nella Tabella 4.1, dovrebbe presentare *controversy*; quello che ci si aspetta è che le attività di *Community detection* producano dei risultati buoni per ciò che concerne le metriche di quantificazione della qualità delle partizioni.

Partendo dalla valutazione circa la numerosità delle due controparti, mostrata in Figura 5.1, non si rimane sorpresi dalla dimensione pressoché identica delle due *community*; questo sarà poi comune a tutti i grafi sui quali è stato applicato *METIS*.

Questo perché, come spiegato precedentemente, *METIS* opera nell'ottica di creare *community* di dimensione simile cercando anche di minimizzare il più possibile il taglio che identifica le stesse.

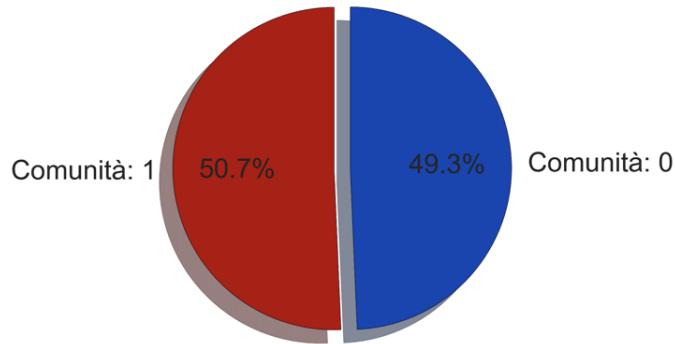


Figura 5.1: Suddivisione dei nodi date le *community* - Guerra del Donbass

Un'altra valutazione, riguardante la distribuzione di grado, è mostrata in Figura A.1.

Possiamo notare come la *community* 1 sembri presentare, al crescere del numero di *tweet*, valore di frequenza maggiori rispetto alla *community* 0; data l'assenza di informazioni semantiche e di direzionalità dei *retweet*, non è stato possibile indagare ulteriormente nel tentativo di giustificare meglio tale situazione.

I risultati ottenuti, data la fase di *Community detection*, sono i seguenti:

#Ukraine				
Community 0	Community 1	Modularità	Copertura	Tempo (s)
2710	2785	0.4685	0.9697	0.0607

Tabella 5.1: Risultati metriche in funzione dell'attività di *Community detection* - #ukraine

Come ci si aspettava, le metriche di modularità e copertura sono alte e suggeriscono quindi una suddivisione sensata del grafo in due parti contrapposte; il tempo di esecuzione, complice sicuramente la dimensionalità ridotta del grafo e le buone performance in generale di *METIS*, è molto basso.

Beefban

Anche per questo secondo grafo, sempre dati gli studi proposti da Garimella et al. (2017), dovrebbe presentare la presenza di *community* contrapposte.

I risultati mostrati nella Tabella 5.2, promuovono i risultati ottenuti da *METIS* per ciò che concerne l'identificazione delle *community*.

#BeefBan				
Community 0	Community 1	Modularità	Copertura	Tempo (s)
824	786	0.4758	0.9798	0.0192

Tabella 5.2: Risultati metriche in funzione dell'attività di *Community detection* - #beefban

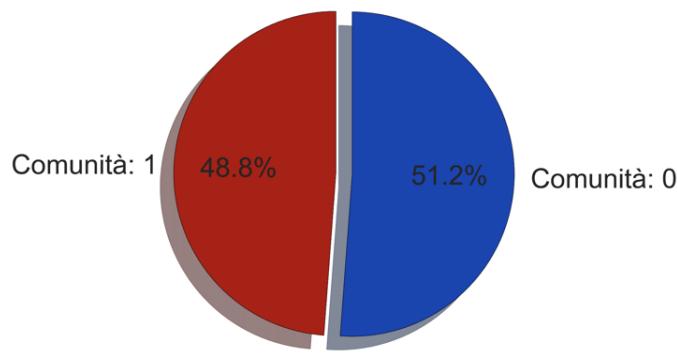


Figura 5.2: Suddivisione dei nodi date le *community* - Divieto di vendita carne di mucca in India

La Figura 5.2 mostra la numerosità delle due partizioni identificate, invece la Figura A.2 mostra, anche per questo grafo, la distribuzione di grado all'interno delle due *community*.

In questo caso le frequenze circa il numero di *retweet* sembrano identiche; ancora una volta, la mancanza di informazioni semantiche rappresenta un ostacolo alla interpretazione di tale situazione.

NationalKissingday

Ultimo grafo tratto dai dati forniti da Garimella, è l'unico grafo non dotato di *controversy*. Ci si aspetta quindi di osservare valori per le metriche definite, inferiori rispetto a quanto precedentemente osservato dati grafi con situazioni di *controversy*.

L'assenza di *controversy* suggerisce l'assenza di gruppi polarizzati su determinate posizioni contrapposte, di conseguenza non dovrebbero esistere strutture di aggregazione all'interno della rete.

La Figura 5.3, mostra la suddivisione per ciò che concerne la numerosità delle *community*; ancora una volta la partizione definita da *METIS* permette di creare due controparti di dimensioni simili.

Risulta invece interessante osservare i valori di modularità e copertura registrati sulle partizioni; la Tabella 5.3 mostra tali valori.

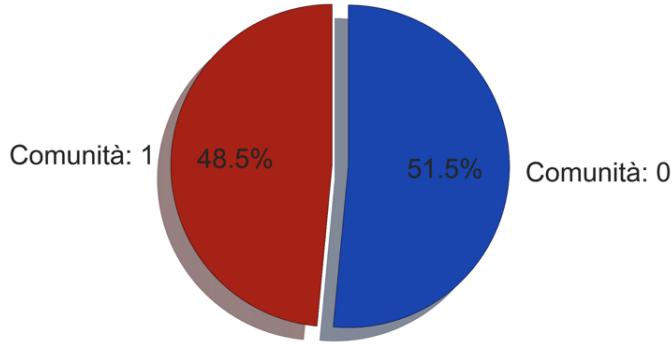


Figura 5.3: Suddivisione dei nodi date le *community* - Giornata internazionale del bacio

#NationalKissingDay				
Community 0	Community 1	Modularità	Copertura	Tempo (s)
2388	2250	0.2608	0.7747	0.0392

Tabella 5.3: Suddivisione nodi date le *community* - #nationalkissingday

Il valore di modularità, data l'assenza di *controversy*, è molto più basso rispetto ai due grafi precedenti; questo prova come i collegamenti di questo grafo siano più vicini a un modello casuale di collegamento.

Anche il livello di copertura è molto più basso rispetto a quanto registrato precedentemente, questo significa che vi sono molti archi che congiungono nodi delle due *community* definite, dunque archi che vivono al di fuori delle partizioni definite.

La valutazione delle tecniche di *Community detection* su questi grafi è importante in quanto può, dato il loro utilizzo e la conoscenza circa la presenza o meno di *controversy*, rappresentare un modello al fine di poter affermare, dati i grafi successivamente presi in considerazione, la possibilità che esista effettivamente una situazione di polarizzazione.

Viene infine fornita, data in Figura A.3, una visualizzazione grafica delle reti definite in funzione delle partizioni trovate e commentate in questa sezione. La visualizzazione grafica è stata resa possibile dall'utilizzo dal software per l'esplorazione e la manipolazione di reti presentato da Bastian et al. (2009), in particolare è stato utilizzato ForceAtlas 2 di Jacomy et al. (2014) a fine di rendere più chiara la visualizzazione delle *community* identificate da *METIS*.

Possiamo notare come le *community* definite su A.3a e A.3b abbiano una strutturazione più classica se si pensa alla contrapposizione tra due comunità; vi sono di fatto due gruppi ben distinti di nodi che vedono una forte connessione interna a fronte di una più scarsa connessione *extra-community*.

Questa situazione va a confermare ulteriormente i risultati ottenuti per ciò che riguarda la modularità e la copertura.

Differenti discorsi invece per la partizione su A.3c, dove la strutturazione della rete vede la presenza di differenti *ego-network* interconnesse tra loro; non è presente una vera e propria contrapposizione tra due partizioni ma più un agglomerato di nodi.

Covid-19

Dopo aver valutato il valore delle metriche volte a quantificare la qualità della *Community detection* su grafi con e senza *controversy*, si è passati ai grafi dei quali non si sapeva nulla circa la possibile presenza di polarizzazioni.

In questa fase è stata posta particolare attenzione alla valutazione delle performance tenendo in considerazione i diversi nuovi pesi definiti in funzione del *sentiment*, dei *topic* e dell'approccio ibrido.

Si sono quindi applicate le tecniche di *Community detection* in funzione dei diversi pesi definiti, per ogni partizione ottenuta si sono valutati i valori delle metriche già proposte precedentemente, al fine di verificare eventuali miglioramenti; si è anche valutata la differenza tra i risultati ottenuti, in termini di distribuzione di *sentiment intra-communitario* e di distribuzione circa i *topic* associati ad ogni utente, in funzione del suo contenuto generato.

Vengono quindi forniti, in Tabella 5.4, i risultati circa le partizioni ottenute in funzione dei diversi pesi definiti.

La colonna di nome *Cambi* quantifica il numero di nodi che, rispetto all'approccio strutturale, hanno subito un cambio di *community*.

	Covid-19					
Tipo	Community 0	Community 1	Modularità	Copertura	Tempo (s)	Cambi
Strutturale	19096	20240	0.4348	0.9351	0.4573	0
Sentiment	19163	20173	0.4403	0.9534	0.471	6297
Topic	19094	20242	0.4396	0.9403	0.4329	6920
Ibrido	20240	19096	0.4322	0.9224	0.489	32226

Tabella 5.4: Risultati metriche in funzione dell'attività di *Community detection* - Covid-19

Sembra che, l'approccio strettamente strutturale, riesca già a identificare delle partizioni che permettono di ottenere delle buone performance circa la modularità e la copertura; in ogni caso, ogni altro approccio proposto, permette di migliorare quelli in prima battuta ottenuti.

L'approccio migliore si rivela essere quello che certa di iniettare nella rete informazioni circa il sentimento, tale metodologia, inoltre, provoca una minore migrazione di utenti da una *community* all'altra.

La Figura 5.4 mostra invece la distribuzione di *sentiment* dati i vari approcci.

Si nota subito che l'approccio strettamente strutturale, in Figura 5.4a, va ad identificare due *community* dotate di distribuzione di sentimento non troppo diverse tra loro; la *community* 1 è infatti dotata di un sentimento tendente al negativo e al neutrale, a differenza invece della *community* 0 la quale è tendenzialmente dotata di un sentimento neutrale.

Si ricorda che la tendenza ad essere neutrale è da attribuirsi alla grande presenza di nodi volti all'attività di diffusione di *news*.

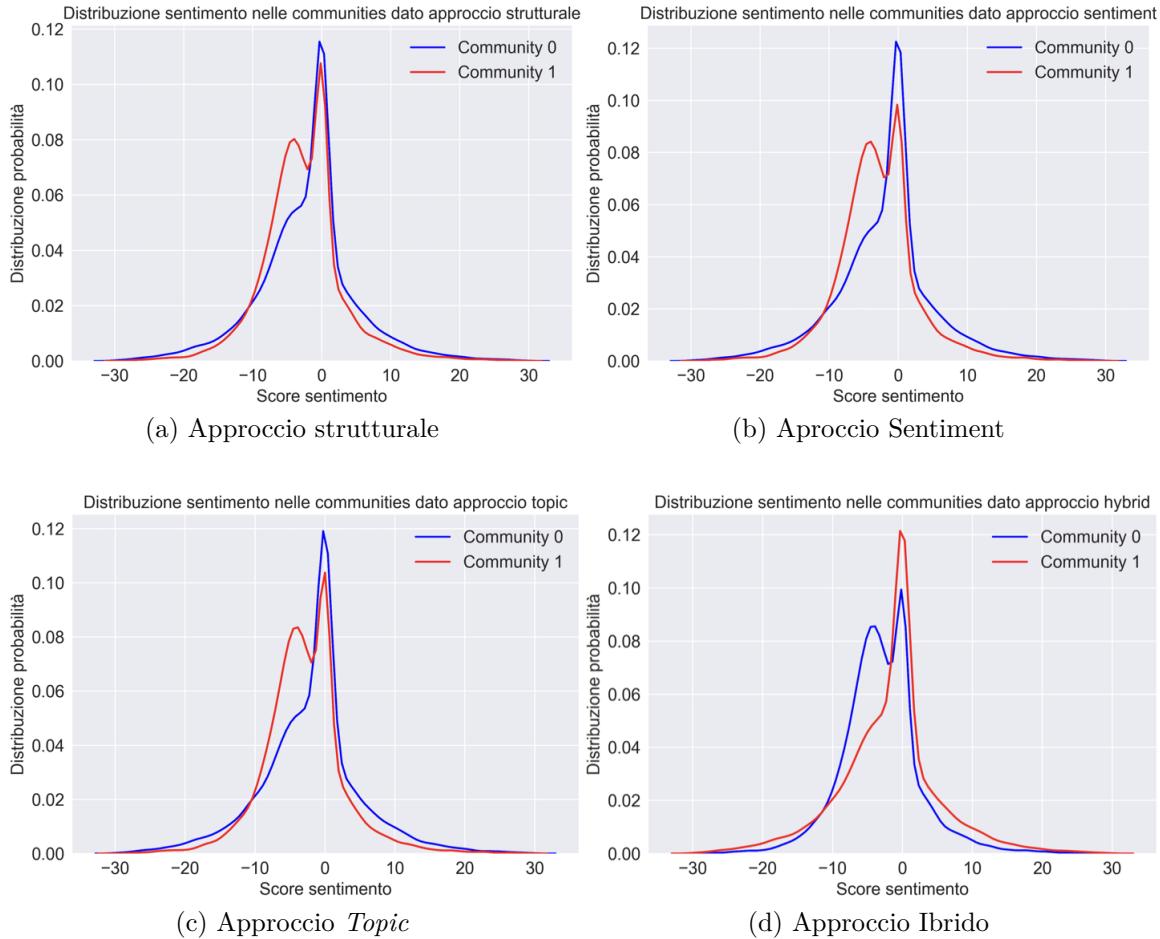


Figura 5.4: Distribuzione sentimento intra-community dati i diversi pesi utilizzati

Si può notare come l'approccio basato sul *sentiment*, mostrato in Figura 5.4b, “levigghi” il picco di utenti della *community* 1 con sentimento neutrale in maniera importante; si può quindi pensare che, dato il numero di cambi eseguiti, questi utenti vengano spostati nella *community* 0, la stessa vede infatti il proprio picco sui neutrali alzarsi.

Questo appiattimento del picco di utenti con score di *sentiment* 0 della *community* 1, viene leggermente perso dall'approccio basato sui *topic*, illustrato in Figura 5.4c.

In ultimo, l'approccio ibrido mostrato in Figura 5.4d, vede un risultato che sembra quasi un ribaltamento della situazione ottenuta con l'approccio basato sul *sentiment*; ricordiamo che il numero di utenti che, data questa metodologia, ha visto un cambio di *community* è pari a 32.226, dunque circa l’82% dei nodi.

È stato dunque deciso di utilizzare l'approccio basato sull'utilizzo del *sentiment*, dunque le *community* sulle quali si andrà a valutare la presenza della *controversy* saranno quelle definite tenendo in considerazione l'informazione circa il sentimento oltre a quella topologica.

Questo è stato deciso in quanto i risultati ottenuti sono i migliori per ciò che concerne le metriche impostate, inoltre è verificabile come vi sia una modifica della distribuzione di

sentiment che va a identificare due tendenze opposte.

Viene infine fornita, in Figura 5.5, una rappresentazione grafica, ottenuta mediante *Gephi*, del partizionamento della rete Covid-19 utilizzando questo approccio.

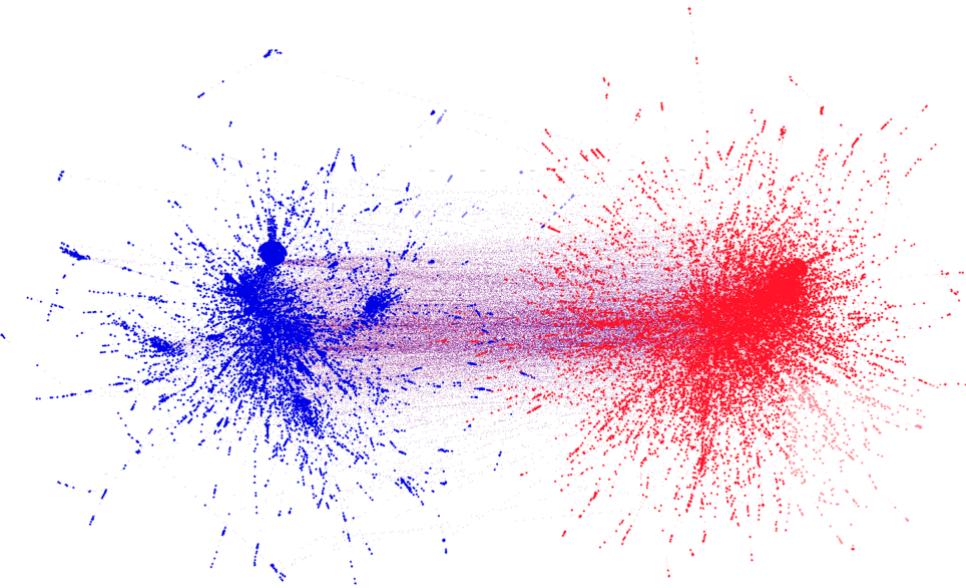


Figura 5.5: Contrapposizione tra le due *community* identificate con approccio basato sul *sentiment* - Covid-19

Vaccination

Ultimo grafo utilizzato in questo lavoro, l'attività riguardante la *Community detection* è del tutto analoga a quella mostrata con il grafo Covid-19.

Sono stati valutate, per tutti i pesi definiti precedentemente, le partizioni ottenibili e calcolate le metriche.

	Vaccinazioni						
Tipo	Community 0	Community 1	Modularità	Copertura	Tempo (s)	Cambi	
Strutturale	19808	19796	0.2953	0.9837	0.4927	0	
Sentiment	19525	20079	0.2935	0.9853	0.4675	35945	
Topic	20061	19543	0.2907	0.9853	0.7641	4169	
Ibrido	19970	19634	0.2912	0.985	2.1296	3776	

Tabella 5.5: Risultati *Community detection* con *METIS* - *Vaccination*

I risultati ottenuti, illustrati in Tabella 5.5, sono a metà strada tra la situazione con e senza *controversy* registrata utilizzando i grafi costruiti sui dati degli studi di Garimella et al. (2017).

Difatti, benché il valore di copertura sia quello più alto registrato, la modularità presenta un livello più tendente a una realtà come quella mostrata in *#NationalKissingDay*. Ad ogni modo, i valori così alti di copertura notificano che quasi tutti gli archi siano contenuti all'interno delle *community*, cosa che si allinea ai dati di Garimella dove viene registrata la presenza di *controversy*.

Si ricorda che i dati riguardanti le vaccinazioni non sono stati raccolti in seguito a un certo evento specifico di particolare interesse; risulta dunque una situazione sicuramente meno chiaramente polarizzata rispetto a ciò che si è ottenuto per Covid-19.

Anche in questo caso è stato deciso di utilizzare le *community* definite dato l'approccio basato sul *sentiment*, questo perché tale approccio vede un incremento della copertura a fronte di un leggerissimo calo della modularità; ci si allinea dunque a quanto fatto per Covid-19. Si fornisce in ultimo la rappresentazione grafica delle comunità identificate:

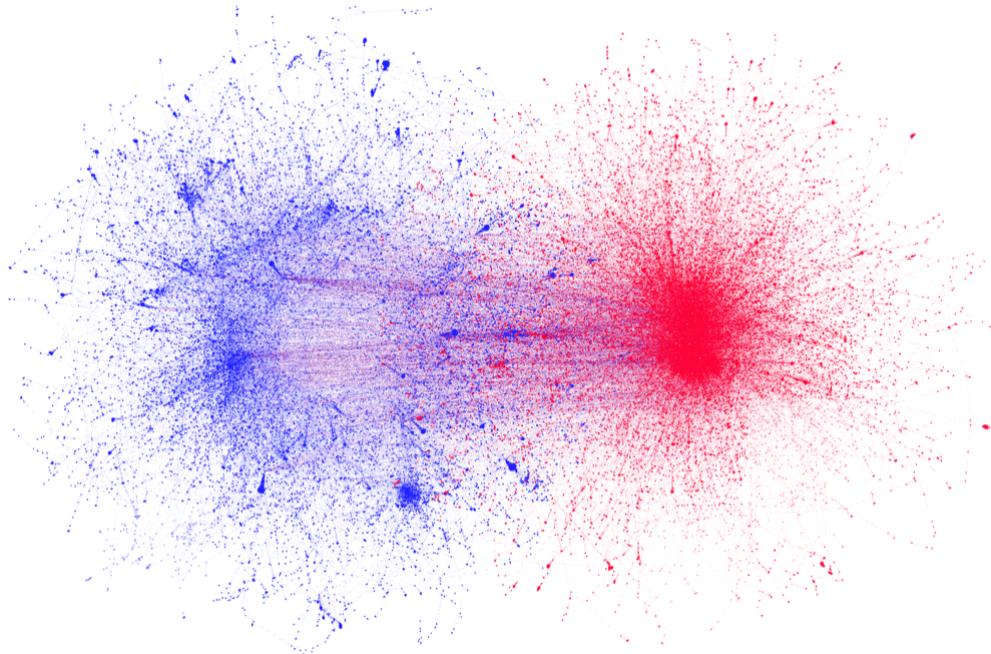


Figura 5.6: Contrapposizione tra le due *community* identificate con approccio basato sul *sentiment* - *Vaccination*

Dall'immagine è possibile notare una minore distanza, tra le due *community* identificate, rispetto ad altre realtà come quelle mostrate precedentemente. Notare inoltre la densità minore degli archi.

5.1.2 *Topic Modelling*

Al fine di definire il peso denominato *Topic Weight (TW)* è stata, come precedentemente sottolineato, effettuata un'operazione di *topic modelling*. I risultati di questa attività sono riportati di seguito.

Topic per Covid-19

Per quanto riguarda la modellazione dei *topic* circa il dataset facente riferimento ai dati del Covid-19, si è identificato un numero ottimale di *topic* pari a 25; con questo numero di *topic* si ottiene lo score di *Topic Coherence* più alto come mostrato in Figura 5.7, inoltre le *key word* facenti riferimento ai vari *topic* sembrano modellare differenti argomenti abbastanza significativi.

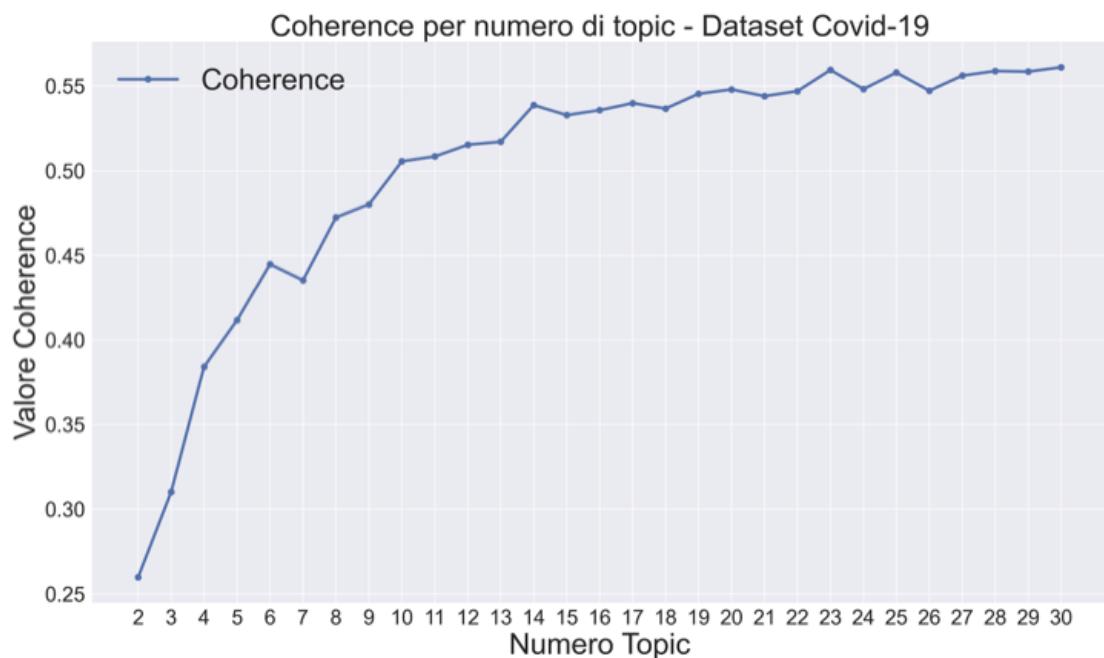


Figura 5.7: Score di *Topic Coherence* registrata per Covid-19

I valori, tendenzialmente crescenti e alti della *Topic Coherence* suggeriscono la presenza di *topic* significativi e ben divisi tra loro; andiamo a riportare, per ogni *topic* definito, le *key word* più probabili che compaiono nello stesso (Tabella 5.6).

Topic per Covid-19	
ID topic	Top 10 key word
1	viru, spread, corona, stop, world, break, news, deadli, continu, friday
2	peopl, infect, chines, die, doctor, warn, kill, million, dead, year
3	time, inform, read, post, chang, share, import, data, good, prepar
4	market, fear, stock, global, point, price, drop, fall, week, worri
5	china, outbreak, govern, countri, epidem, control, fight, india, prevent, measur
6	vaccin, expert, question, develop, start, month, work, cure, drug, treatment
7	health, emerg, public, world, declar, threat, minist, global, nation, intern
8	diseas, human, caus, sar, transmiss, scientist, studi, origin, sourc, expert
9	iran, home, medic, work, risk, care, stay, worker, sick, famili

10	test, patient, posit, hospit, symptom, isol, day, result, contact, neg
11	scare, amid, concern, cancel, plan, fear, event, year, grow, major
12	death, china, report, toll, number, rise, increas, infect, hubei, provinc
13	itali, close, school, region, student, lockdown, countri, italian, univers, shut
14	thing, good, happen, money, hope, wait, life, save, love, feel
15	impact, economi, global, industri, econom, bank, demand, expect, compani, product
16	travel, south, flight, korea, countri, singapor, airlin, australia, airport, restrict
17	case, confirm, report, total, recov, bring, number, germani, suspect, today
18	state, offici, washington, announc, counti, hong, kong, person, york, health
19	quarantin, ship, cruis, japan, passeng, american, princess, evacu, diamond, california
20	china, wuhan, epicent, victim, video, citi, citizen, insid, show, beij
21	mask, protect, face, hand, prevent, cough, disinfect, food, wash, wear
22	trump, respons, presid, blame, penc, democrat, administrat, lie, hoax, american
23	covid, ncov, pandem, asia, sarscov, canada, wuhanviru, coronavirusa, wuhanpneumonia,
	wuhan
24	updat, news, press, latest, watch, today, follow, outbreak, confer, panic
25	polit, polit, problem, claim, real, make, fact, believ, wors, danger

Tabella 5.6: Top 10 *key words* modello *LDA* con *Topic Coherence* più alta per *Covid-19*

I *topic* definiti, in linea di massima, tendono a essere caratterizzati da *key words* richiamanti differenti concetti e situazioni: il *topic* 1 sembra parlare della storia del medico che per primo cercò di avvisare le autorità circa la presenza di un nuovo virus, il *topic* 21 sembra parlare delle nuove abitudini di vita e dello *smartworking*, il 22 ha connotati politici circa gli Stati Uniti d'America, il *topic* 19 parla del caso della nave da crociera *Diamond Princess*, i *topic* 4 e 15 sembrano racchiudere le previsioni e paure a livello economico e il 13 parla del nostro paese e delle misure preventive adottate (e.g. chiusura delle università).

In generale, quasi tutti i *topic*, sembrano toccare un determinato argomento o aspetto riguardante il Covid-19, che si vada dai nuovi stili di vita imposti dal virus, alle problematiche economiche, fino a argomenti circa il blocco degli spostamenti e lo studio di un vaccino.

Il problema principale, registrato in questa fase, riguarda la presenza di molti utenti facenti riferimento a testate giornalistiche o a profili di divulgazione, i quali quindi condividevano *tweet* che di fatto riguardavano *news* sul virus; le operazioni preliminari hanno permesso di eliminare una serie di *tokens* che si ripetevano un numero non utile di volte, tuttavia molti *topic* sembrano richiamare a realtà di *breaking news*.

Topic per Vaccination

Data la modellazione dei *topic*, circa i *tweet* facenti riferimento al dataset sulle vaccinazioni, sono stati identificati un totale di 14 differenti *topic*.

I picchi maggiori di *Topic Coherence* sono stati registrati presso i valori 14 e 17, come mostrato nella Figura 5.8; una valutazione delle *key word* ha guidato la scelta del

modello.

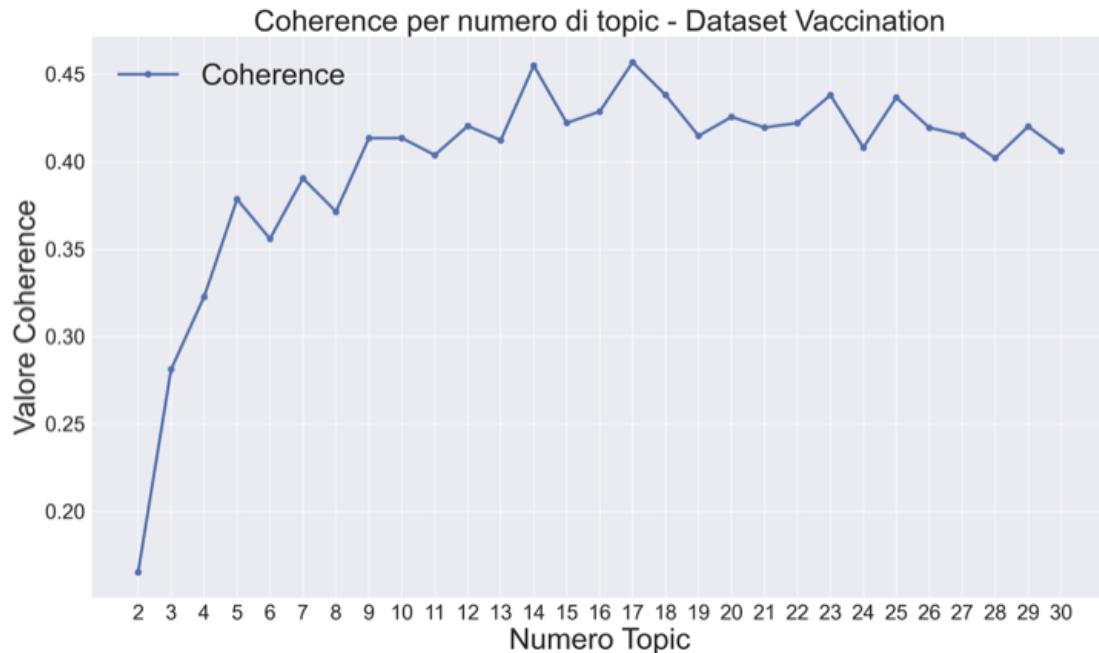


Figura 5.8: Score di Topic Coherence registrata per Vaccination

A differenza dei *tweet* su Covid-19, i valori della metrica circa la qualità delle operazioni di *topic modeling* è tendenzialmente più bassa, questo è dovuto principalmente al contesto di raccolta dei dati: se i dati circa il Covid-19 sono stati raccolti nel periodo di maggiore discussione circa la pandemia, i dati riguardanti le vaccinazioni sono stati raccolti senza che fossero accaduti fatti di particolare interesse pubblico, inoltre in questa rete non sono presenti nodi facenti riferimento a testate giornalistiche, di conseguenza anche i numero di *tweet*, in generale, è molto più basso.

Ovviamente, come verrà a breve mostrato, emergono argomenti propri riguardanti le vaccinazioni e lo scontro che può essere presente in questo contesto; tuttavia, il periodo di raccolta e la tipologia di utenti facenti parte della rete possono giustificare, anche se parzialmente, i valori più bassi registrati.

In maniera analoga a quanto fatto prima, vengono fornite le informazioni circa le *top key word* che caratterizzano i *topic* definiti in questa fase:

Topic per Vaccination	
ID topic	Top 10 key word
1	<i>children, protect, time, live, child, life, safe, adult, save, month</i>
2	<i>anti, peopl, scienc, believ, reason, chang, danger, thing, vaxxer, call</i>
3	<i>measl, diseas, immun, outbreak, infect, herd, viru, prevent, respons, dose</i>
4	<i>evid, person, point, post, problem, good, ignor, gener, base,a gre, mass</i>
5	<i>forc, mandatori, stop, peopl, choic, make, pharma, mandat, wrong, american</i>

6	year, polio, prevent, cancer, campaign, vaccineswork, start, million, free, girl
7	rate, risk, countri, world, high, increas, coverag, schedul, smallpox, global
8	caus, kid, child, autism, babi, die, kill, happen, harm, ill
9	import, today, great, awar, good, check, share, cours, stori, offe
10	death, studi, case, report, effect, follow, influenza, number, data, reduc
11	work, program, support, human, badger, programm, rabi, control, area, cull
12	health, care, commun, provid, patient, famili, recommend, includ, improv, servic
13	school, medic, parent, state, exempt, public, inform, requir, doctor, educ
14	research, issu, fact, read, question, differ, safeti, news, test, polici

Tabella 5.7: Top 10 *key word* modello *LDA* con *Topic Coherence* più alta per *Vaccination*

In generale, sembra possibile dividere i *topic* come pro o contro le vaccinazioni e la caratteristica obbligatorietà delle stesse; *topic* identificati dai numeri 6, 9, 11, 12 sembrano appoggiare i vaccini, sottolineandone l'utilità, la necessarietà e rigettando le tendenze *no-vax* che sembrano espresse nei *topic* 2, 4, 5, 7, 8, 10.

Solo i *topic* 3, 13 e 14 risultano essere i meno chiari o più generali; è stato in ogni caso deciso di utilizzare tale modello in quanto altri, non solo quello che modellava 17 *topic*, presentavano *topic* non immediati o comunque non facilmente interpretabili.

5.2 Analisi metadati

Prima di passare alla quantificazione della *controversy* tra le comunità identificate col metodo introdotto precedentemente e alla riduzione della stessa, è stato deciso di effettuare un'attività di analisi circa i metadati date le reti Covid-19 e *Vaccination* (Sezioni 5.2.1 e 5.2.2); questo tipo di attività è servito per applicare una metrica più “*human-based*” volta a validare ulteriormente il risultato della *Community detection*.

Per i grafî di Garimella non si è effettuata questo tipo di attività in quanto tutta l'informazione semantica risulta mancante, inoltre quei grafî risultano già “validati” dagli studi eseguiti precedentemente sugli stessi.

5.2.1 Covid-19

Questa rete, data la maggior presenza di metadati, ha sicuramente permesso di ottenere informazioni molto interessanti in grado di modellare e delineare meglio le *communit* definite.

Partendo dalle descrizioni dei profili utente, sono state effettuate delle analisi volte a creare delle *wordcloud* al fine di verificare eventuali differenze tra i membri delle *community*; la descrizione personale che un utente associa al suo profilo, può contenere informazioni e *keyword* importanti per definire le sue credenze e i suoi pensieri. La Figura 5.9 mostra le *wordcloud* in oggetto.



Figura 5.9: *Wordcloud* date le descrizione dei profili degli utenti in funzione della *community*

Salta subito all’occhio come, in Figura 5.9b facente riferimento alla *community* 1, vi siamo parole che richiamano chiaramente a un orientamento politico volto al sostegno dell’attuale presidente degli Stati Uniti di America: *Donald John Trump*; parole come *MAGA* (*Make America Great Again*) o *KAG* (*Keep America Great*), richiamano agli spot elettorali dell’attuale presidente.

Inoltre, parole come *Conservative*, *Patriot*, *Christian* e *God* richiamano certi valori fondanti sui quali si basa la destra negli Stati Uniti.

Dall’altra parte invece, la Figura 5.9a, risulta essere dotata di parole richiamanti più una comunità scientifica; parole come *research*, *Health*, *science*, *healthcare*, *Public Health*, *Journalist* richiamano questa visione.

Tenendo in considerazione il grafico mostrato precedentemente, circa la distribuzione del sentimento medio degli utenti date le *community* (5.4b), si può concludere come la *community* che sembra *pro-Trump* presenti un sentimento molto più negativo rispetto invece alla *community* che sembra più scientifica o di divulgazione.

Questo fatto è relativamente in linea a quanto ci si potrebbe aspettare, questo perché è ormai da anni che esponenti di spicco della destra Americana (e non solo) abituano l’elettorato a un linguaggio molto violento e diretto, questo *modus operandi* viene poi assimilato dall’elettorato in quanto viene di fatto giustificato dall’alto.

Un’altra interessante valutazione riguardante la possibile contrapposizione circa l’affidabilità dell’informazioni scambiata nella rete, emerge dalla valutazione del numero di utenti, per ogni *community*, dotati di un profilo verificato.

Il *badge* blu associato ad un certo *account* su *Twitter*, permette alle persone di sapere che un account di interesse pubblico è autentico.

Come mostrato in Figura 5.10, il numero di account verificati per la *community* 0, quella opposta alla *pro-Trump*, oscilla tra il 22% e il 23% sul totale degli account. La partizione opposta invece, presenta solo un 12% di account verificati.

Questo potrebbe celare una presenza di account *fake* promulgatori di *fake news*, questo viene da pensare in seguito ai precedenti durante la campagna elettorale del 2016 che ha visto l’elezione di *Trump* come 45° presidente.

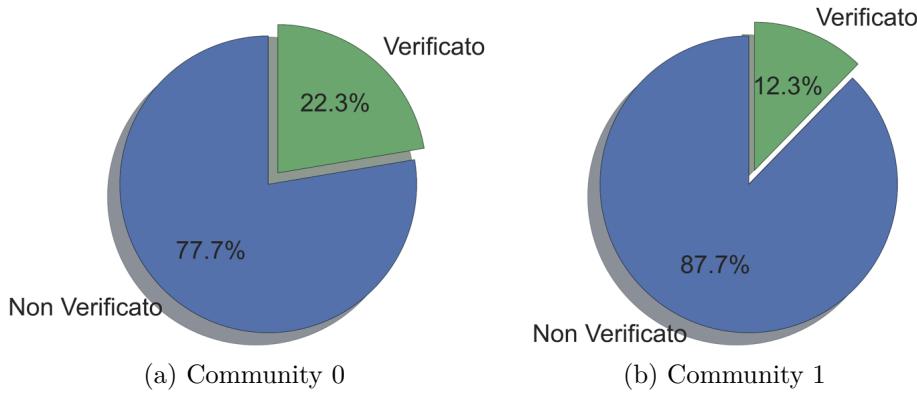


Figura 5.10: Percentuali di account verificati per ogni community

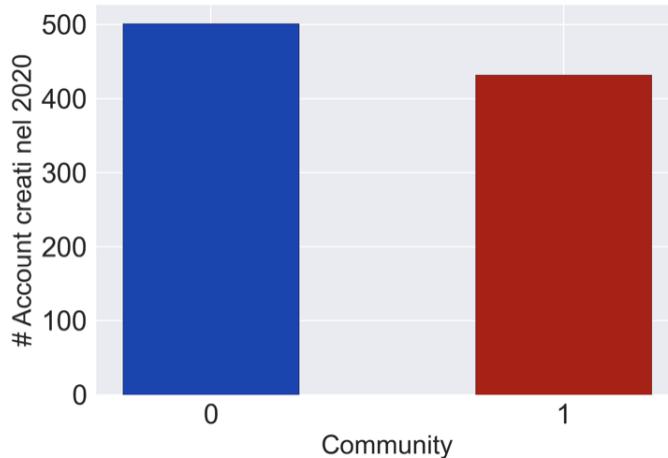


Figura 5.11: Numero di utenti con anno di creazione profilo nel 2020

Al fine di supportare questa ipotesi, le Figure 5.11 e A.4 rappresentano rispettivamente il numero di *account* per ogni *community*, creati nel 2020 durante il periodo di inizio pandemia e il numero di profili con immagine utente di default (dunque non personale).

Possiamo notare come, a fronte di un numero maggiore di nuovi utenti nella *community* 0, la *community* 1 presenti il numero maggiore di utenti con immagine del profilo di default.

Peculiarità dei profili *fake*, possibili promulgatori di *fake news*, è l’anonimato del profilo, il quale può rimanere per del tempo spoglio e non personalizzato prima di entrare totalmente in attività.

Un’altra interessante analisi è stata condotta in funzione del *sentiment* medio caratterizzante, per ogni partizione, gli utenti dotati del *badge* azzurro che di fatto “verifica” l’autenticità di un determinato profilo.

Come si può osservare dalla Figura 5.12, gli utenti verificati della *community* identificata come “*pro-Trump*” presentano un *sentiment* medio più negativo rispetto agli utenti verificati dell’altra *community*, i quali invece sono dotati di un sentimento più neutrale.

Capitolo 5. Risultati

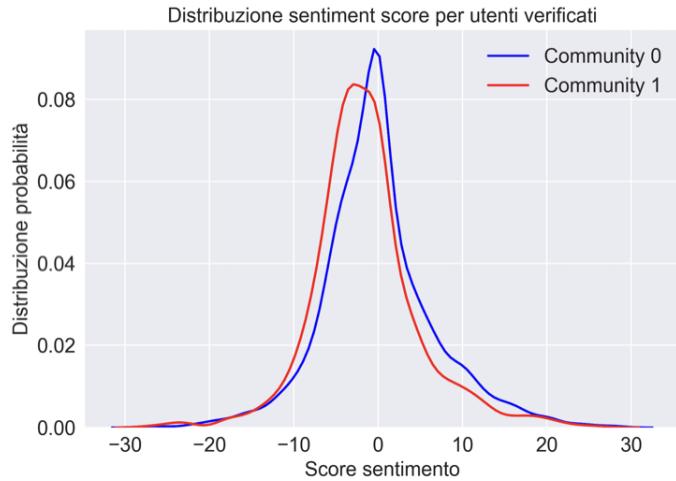


Figura 5.12: Distribuzione *sentiment score* per utenti verificati

La suddivisione in *community*, eseguita da *METIS*, sembra mettere effettivamente in contrapposizione due fazioni che non per forza risultano essere complementari, ma che possono osservare effettiva differenza di pensiero.

5.2.2 Vaccinazioni

Per questo grafo, oltre alle informazioni circa il contenuto scambiato, non si era purtroppo in possesso di ulteriori significativi dati riguardanti i diversi utenti nella rete.

Partendo, sulla falsa riga di quanto fatto prima, dai *wordcloud* facenti questa volta riferimento al contenuto dei *tweet*, si ottengono informazioni interessante rispetto alle *keyword* maggiormente utilizzate:

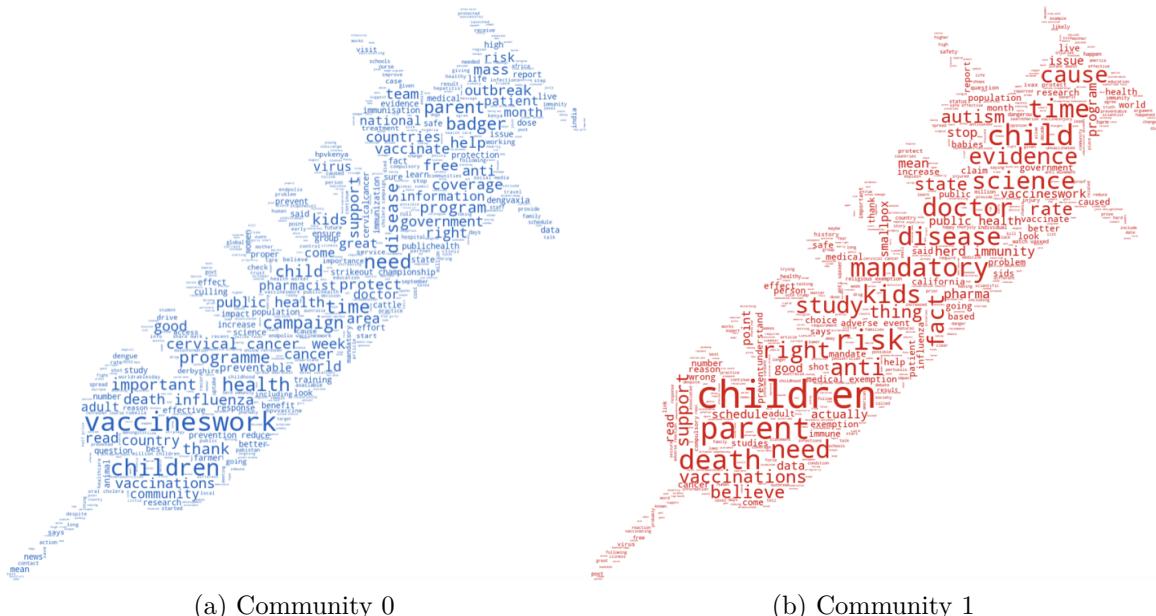


Figura 5.13: *Wordcloud* dati i *tweet* scambiati all'interno delle communities definite

Sembrerebbe delinearsi uno scontro tra un gruppo più tendente ad essere *pro-Vax* verso uno *no-Vax*; la Figura 5.13a presenta parole che sembrano portare verso il primo gruppo

identificato, parole come *vaccineswork*, *important*, *health*, *children* e *need* suggeriscono tale allineamento.

Viceversa, in Figura 5.13b, vi sono parole che richiamano la visione opposta: *disease*, *death*, *cause*, *anti*, *risk*, *autism* sembrano essere la conferma di tale ipotesi di contrapposizione.

Ulteriore conferma, volta a modellare le *community* come pro- e no-Vax, arriva dal numero di utenti, per ogni partizione, associato a ogni *topic*.

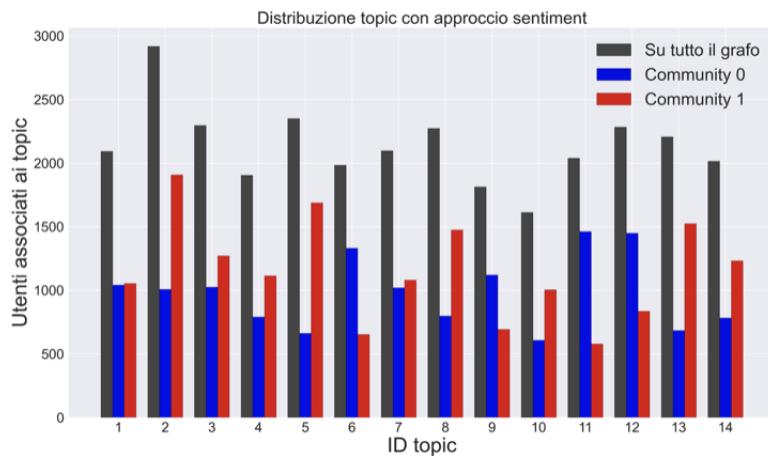


Figura 5.14: Numero di utenti per *community* che hanno prodotto *tweet* riconducibili ad un *topic*

Stando ai *topic* proposti nella Tabella 5.7 risulta possibile osservare come, *topic* che sembrano più tendenti ad essere *no-Vax*, presentino un numero maggiore di utenti disquisitori nella *community 1* (e.g. *topic* 2, 4, 5, 7, 8, 10); di contro invece *topic* con *keyword* più positive, sono discussi in misura maggiore nella *community 0* (e.g. *topic* 6, 9, 11, 12).

Tale situazione fa dunque pensare che vi sia effettivamente una contrapposizione sul tema della sicurezza dei vaccini, dunque la partizione ottenuta da *METIS* sembra, anche a livello semantico, sensata.

Un’ulteriore analisi effettuata ha riguardato le centralità per ciò che concerne il Grado e la *Betweenness*. I concetti di centralità, nella fattispecie quelli *Grado*, *Betweenness* e *Closeness*, sono stati introdotti nel Capitolo 2. Si rimanda quindi il lettore a consultare le definizioni fornite per tali misure.

Tornando alle analisi circa la centralità, la *community 0* vede la presenza principale di nodi i quali si professano dottori, ricercatori, cliniche; molto più interessante sono le informazioni che emergono dallo studio dei nodi centrali per la *community 1*, anche qui vi sono nodi che si professano appartenenti alla comunità scientifica, vi sono però anche nodi apertamente *no-vax*.

Vengono fornite ora degli *screenshot* facenti riferimento alle descrizioni dei loro profili utente:

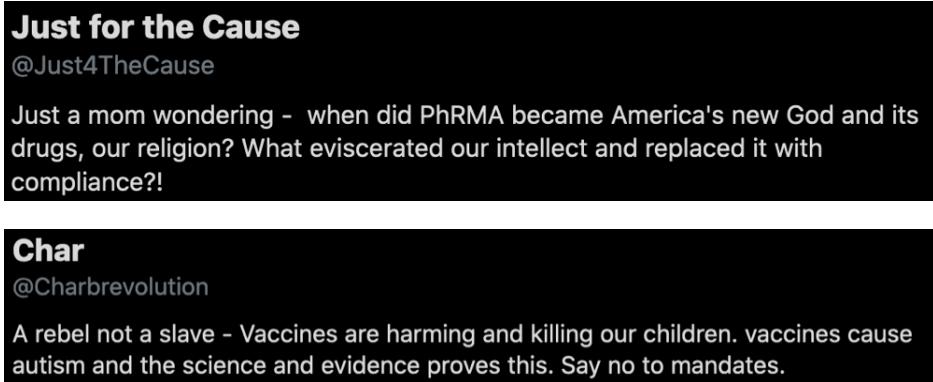


Figura 5.15: Descrizione profili utente no-vax presenti nella *community* 1 come nodi centrali, rispettivamente, per grado e per *Betweenness*

È stato anche osservato come questi profili siano molto attivi nella condivisione e risposta ad altri utenti.

Sembrerebbe dunque che, la partizione eseguite da *METIS* sul grafo delle vaccinazioni, abbia identificato una *community* strettamente scientifica (*community* 0) e una che invece presenta al suo interno un certo numero di profili volti a propagandare la pericolosità delle vaccinazioni (*community* 1).

Bisogna comunque sottolineare come, data la *community* 1, questa presenti comunque esponenti della comunità scientifica volti a scomporre le ipotesi complottiste dei *no-vax*.

5.3 Valutazione controversia

La valutazione della controversia ha visto l'utilizzo di tutte le metriche presentate nel Capitolo 4 nella sezione facente riferimento alla quantificazione della *controversia*, la quale può notificare la presenza di polarizzazioni.

Vengono ora riportati i risultati ottenuti per le reti modellate in funzione degli studi di Garimella et al. (Sezione 5.3.1), per poi passare alle reti sociali facenti riferimento a Covid-19 e *Vaccination* (Sezioni 5.3.2 e 5.3.3).

In ultimo verrà proposta una discussione circa i risultati ottenuti (Sezione 5.3.4)

5.3.1 #BeefBan, #Ukraine e #NationalKissingDay

La quantificazione della controversia su questi grafi ha assunto un notevole livello di importanza, questo a causa degli studi già condotti sugli stessi, i quali hanno permesso di trattare i risultati, in funzione dell'attività in oggetto, come una sorta di *benchmark* per tutti gli altri grafi.

Si ricorda che i grafi *#Ukraine* e *#BeefBan* sono dotati di controversia e invece *#NationalKissingDay* no, per i grafi successivi ci si aspetta di registrare risultati più simili a quelli dei due grafi iniziali.

Vengono ora forniti, nella Tabella 5.8, i risultati ottenuti in funzione delle seguenti

metriche: *Random Walk Controversy (RWC)*, *Boundary Connectivity (GMCK)*, *Random Walk Top Degree (RWTD)* e *Change Side Controversy (CSC)*. Tutte le metriche specificate hanno valutato la controversia, partendo e tenendo in considerazione, le partizioni ottenute con l'approccio strutturale (l'unico possibile dati questi grafi).

Grafo	Garimella			
	RWC	GMCK	RWTD	CSC
#BeefBan	0.986	0.2132	0.8583	0.9895
#Ukraine	0.9624	0.1729	0.8889	0.9837
#Nationalkissingday	0.4904	0.0213	0.6875	0.7427

Tabella 5.8: Risultati metriche quantificazione controversia - Garimella

È possibile fin da subito notare come, per tutte le metriche utilizzate, i risultati facenti riferimento al grafo privo di controversia registrino risultati molto più bassi rispetto agli altri due.

I divari più importanti e significativi sono registrati dalle metriche proposte dagli studi di Garimella et al. (2017), tali *RWC* e *GMCK*; tuttavia, anche le altre due, che sono state proposte in questa tesi, riescono a modellare due realtà distinte, ottenendo valore diversi in funzione della controversia identificata.

Possiamo notare come nei grafi dotati di controversia il valore della *RWC* sia molto più alto rispetto a quello registrato su di un grafo privo di controversia; questo ci suggerisce che gli utenti di una determinata partizione risultano essere esposti quasi unicamente ai nodi “autorevoli” della propria comunità.

Tale situazione viene ulteriormente confermata dai risultati registrati data la *CSC*, gli stessi permettono di affermare che, in una situazione dotata di controversia, durante una navigazione casuale sono praticamente assenti i cambi di partizione; risultano quindi quasi assenti esposizioni a contenuto promulgato dalla partizione opposta.

Nel caso in cui ci si concentrassesse nella valutazione di *Random Walk* eseguiti solo dai nodi “autorevoli”, andando a valutare quindi l’interconnessione tra i nodi più centrali per grado, si nota come la tendenza ad essere esposti a nodi centrali della propria comunità risulta essere ulteriormente confermata; in ultimo il valore registrato da *GMCK* notifica una bassa interconnessione tra i nodi di frontiera delle due partizioni identificate.

Questi risultati sono importanti al fine di pesare correttamente le affermazioni sulle prossime reti sociali, proposte in questo lavoro di tesi.

Dati i risultati simili ottenuti nella fase di *Community detection*, per i grafi *Covid-19* e *Vaccination* nei confronti di #BeefBan e #Ukraine, ci si aspetta di osservare valori simili a quelli qui presentati, date le metriche di quantificazione.

5.3.2 Covid-19

Verranno presentati ora, nella Tabella 5.9, i risultati ottenuti per il grafo Covid-19 data la fase di quantificazione della controversia; si ricorda che in questo caso, le partizioni sulle quali si è effettuato tali attività, sono state ottenute mediante l'approccio basato sul *sentiment* di *Community detection*.

La decisione è stata precedentemente illustrata in Sezione 5.1.1.

Covid-19			
RWC	GMCK	RWTD	CSC
0.9533	0.1813	0.854	0.9807

Tabella 5.9: Risultati metriche quantificazione controversia - Covid-19

Come ci si poteva aspettare i risultati ottenuti, per tutte le metriche utilizzate, sono in linea con quelli registrati per *#BeefBan* e *#Ukraine*; ancora una volta parrebbe che gli utenti siano esposti solo ed esclusivamente all'informazione interna alla propria partizione, subendo poche interazioni, dato l'argomento modellato, con utenti facente parte della community opposta.

Tali osservazioni sono giustificate dalle metriche *RWC*, *RWTD* e *CSC*. La leggera differenza presentata dal valore del *RWC*, non viene ritenuta statisticamente rilevante in quanto inferiore allo 0,1.

Il valore rilevato da *GMCK*, risalta ulteriormente la situazione di segregazione presente nella rete; infatti sembrerebbe che i nodi “frontiera” siano collegate di più con i nodi “interni” che tra loro, indice appunto della poca interconnessione registrata tra le due partizioni ottenute.

5.3.3 Vaccinazioni

Ultimo grafo tenuto in considerazione nel lavoro di tesi, per *Vaccination* si è proceduto in maniera del tutto analoga a quanto mostrato per i precedenti.

La Tabella 5.10 mostra i risultati ottenuti, date le metriche utilizzate.

Vaccination			
RWC	GMCK	RWTD	CSC
0.953	0.235	0.8616	0.9871

Tabella 5.10: Risultati metriche quantificazione controversia - Vaccination

Ancora una volta, i risultati sono in linea con quelli presentati sui grafici costruiti sui dati degli studi di Garimella che presentavano situazioni di controversia; viene quindi ancora notificata una scarsa tendenza all'interazione tra gli utenti delle due partizioni identificate.

Oltre ai valori in linea delle metriche basate sui *Random Walk*, si vuole sottolineare che *GMCK* segnala per questo grafo, il più alto valore registrato.

Risulta che questa rete veda la più bassa interconnessione tra i nodi definiti di “frontiera”, ulteriore indice volto a notificare la segregazione e polarizzazione della rete.

5.3.4 Discussione

L'attività di quantificazione della controversia compiuta in questa fase, ha permesso di verificare l'effettiva presenza di comunità polarizzate; in questo senso l'utilizzo dei grafi di Garimella è stato fondamentale al fine di avere valori di paragone per le due reti sociali, Covid-19 e *Vaccination*, delle quali non si avevano informazioni in merito.

Si è potuto quindi definire una sorta di intervallo di confidenza per la *controversy*, da poter utilizzare al fine di indicare un partizionamento su un grafo come caratterizzato da camere d'eco, dunque polarizzato.

Interessante poi notare la presenza, dati i valori di *modularity* e *coverage* registrati, di una relazione diretta tra l'attività di *Community detection* e di quantificazione della controversia; difatti si evidenzia come, ad alti valori associati alle metriche utilizzate per validare l'attività di partizionamento del grafo, corrispondano alti valori di metriche utilizzate per quantificare la controversia.

Questo poteva sembrare ovvio in quanto: partizioni che ottengono buoni valori di *modularity* e *coverage* sono molto probabilmente dotate di più collegamenti interni che di collegamenti esterni. Tuttavia, l'ulteriore conferma va di fatto a caratterizzare una nuova possibile forma di validazione, date le *community* ottenute.

In ultimo si vuole evidenziare come, data una modellazione di una discussione controversa o meno, tutte le metriche utilizzate abbiano registrato valori in linea alla realtà analizzata; l'utilizzo combinato delle stesse è sempre preferibile, in quanto va a modellare tutte le possibili sfaccettature della definizione di controversia, fornendo poi una misura complessiva più robusta.

Ad ogni modo, il comportamento individuale di ognuna può essere molto accurato, come poi dimostrato dai risultati ottenuti.

5.4 Link prediction

In quest'ultima sezione verranno riportati i risultati circa l'attività finale volta a minimizzare, con l'aggiunta del minor numero possibile di archi, la controversia precedentemente identificata.

Verrà quindi analizzata, per ogni metrica di controversia descritta ed utilizzata e per ogni algoritmo di *Link prediction* proposto, la riduzione di polarizzazione; l'obiettivo è quello di identificare la tecnica di collegamento che più minimizza la segregazione tra le due partizioni, andando dunque a permettere, a livello teorico, una maggior libertà di movimento nella rete stessa.

Si procederà mostrando i risultati sui grafi “allo stato dell’arte” (Sezione 5.4.1), per poi passare alle reti sociali Covid-19 e *Vaccination* (Sezioni 5.4.2 e 5.4.3).

Si passerà poi a una valutazione circa le possibili differenze di riduzioni in seguito all'utilizzo del *Sentiment Boost* (Sezione 5.4.4); per poi concludere con un'analisi, dato l'algoritmo di *Link prediction* migliore, delle reti risultanti la *Link prediction* (Sezione 5.4.5).

In ultimo si commenteranno i risultati ottenuti alla luce delle differenze strutturali delle reti sociali (Sezione 5.4.6).

5.4.1 #BeefBan e #Ukraine

Ancora una volta questi due grafi hanno rappresentato il terreno di “prova” nell’attività di *Link prediction*; seppure la dimensione degli stessi non risulti essere paragonabile a quella degli altri due grafi definiti, si ipotizza che i risultati qui osservati si possano applicare anche a realtà diverse.

Link prediction #BeefBan

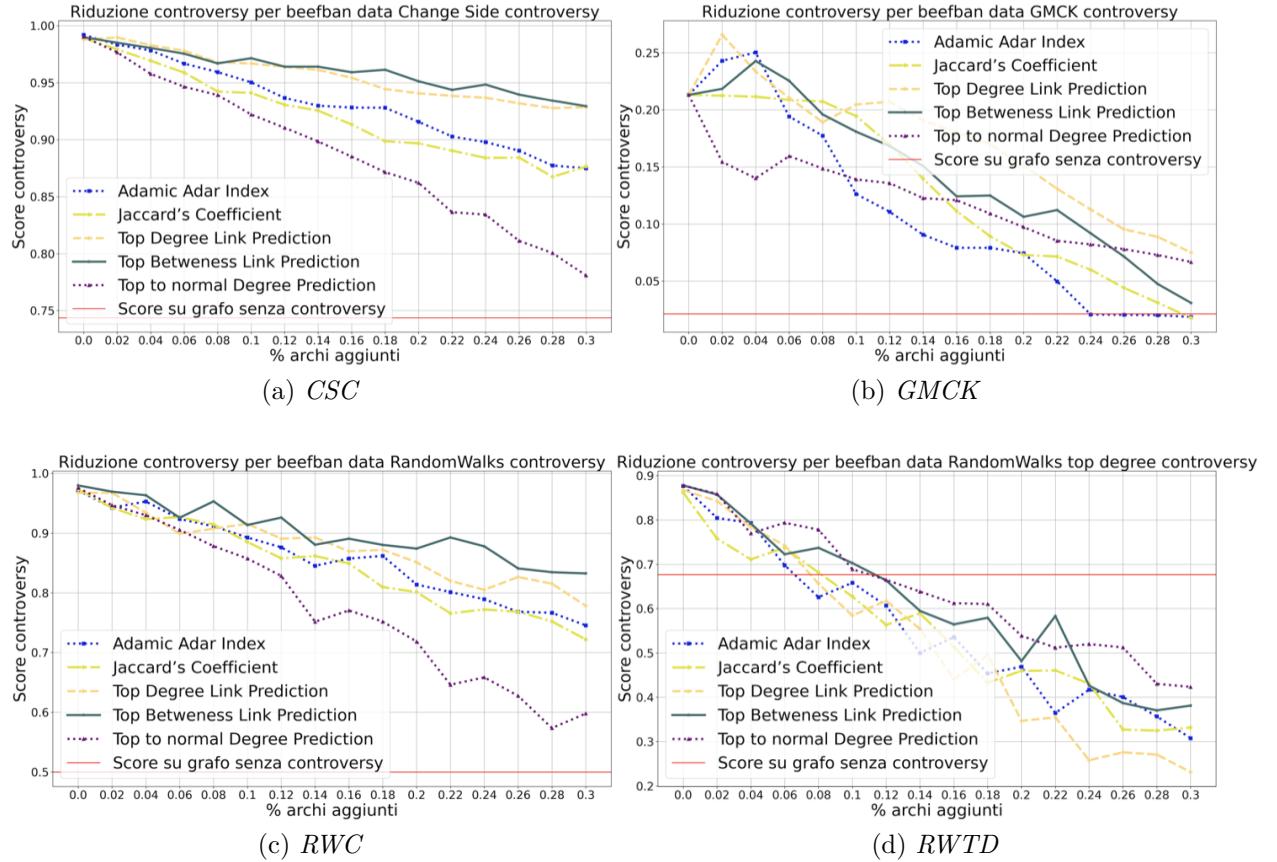


Figura 5.16: Riduzione controversia - #BeefBan

La Figura 5.16 mostra, per ogni metrica di quantificazione utilizzata e presentata precedentemente, i risultati ottenuti dall’attività di *Link prediction*.

Per ogni grafico, la linea rossa orizzontale, fa riferimento ai valori di controversia registrati per una realtà che non vede la presenza di camere d’eco; l’intendo è dunque mostrare un valore di controversia accettabile al fine di definire il grafo come privo di polarizzazioni.

Passando ora all'analisi dei risultati, sembrerebbe che l'attività di *Link prediction* effettuata mediante il *Top to Normal Linking* sia quella che ha registrato, con il numero minore di collegamenti, la riduzione più importante di controversia.

Date le metriche *CSC* e *RWC* (Figure 5.16a e 5.16c rispettivamente), vi è un vero e proprio "stacco", con l'aggiunta di un 17/18 % di archi, nei confronti delle altre soluzioni. Le uniche incertezze circa l'efficacia sono notificate dai grafici illustrati nelle Figure 5.16b e 5.16d che vedono, mediante l'utilizzo della *Top to Normal Linking*, una riduzione più contenuta rispetto alle altre modalità di *Link prediction*, la stessa però è comunque importante e si avvicina verso valori accettabili di controversia.

Riteniamo importante notificare che, date le dimensioni molto contenute di questo grafo, il numero di nuovi archi aggiunti è di molto inferiore a ciò che i prossimi grafi presenteranno; il raggiungimento parziale, date le metriche, di valori accettabili di controversia può ed è sicuramente dovuto da questa circostanza.

Link prediction #Ukraine

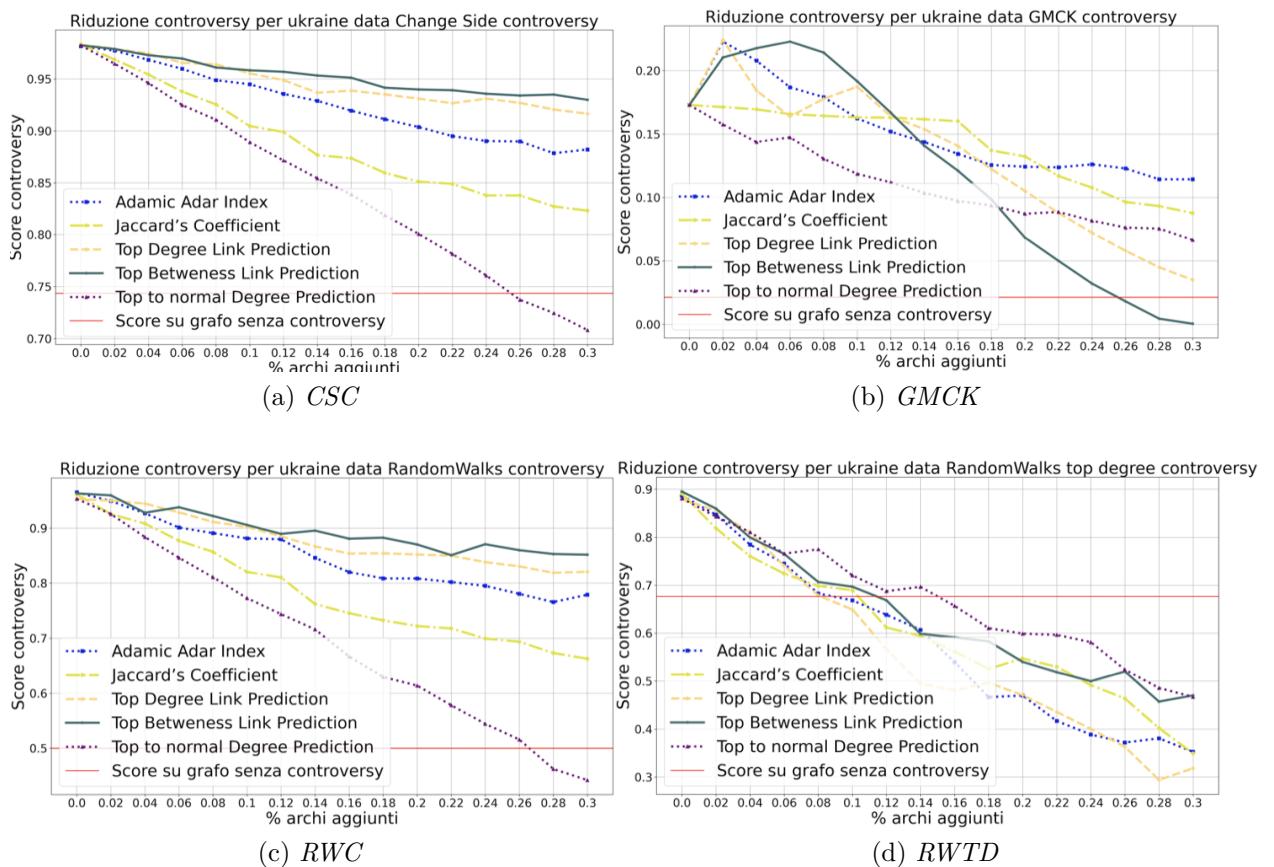


Figura 5.17: Riduzione controversia - *#Ukraine*

Per l'ultimo grafo allo stato dell'arte, la Figura 5.17 mostra ancora una volta come l'attività di *linking* proposta da *Normal to Top* paia registrare i risultati migliori sulle metriche *CSC* e *RWC* (Figure 5.17a e 5.17c); andando poi a causare un'importante

riduzione dei valori di *GMCK*, permettendo dunque una più diretta connessione tra i nodi frontiera delle due partizioni identificate.

L'unica metrica dove si posiziona ultima, in termini di riduzione di controversia, è la *RWTD* (Figura 5.17d) dove il *linking* effettuato dalla tecnica *Top Degree*, volta al collegamento tra i nodi “autorevoli” per grado, registra per ovvie ragioni i valori migliori.

Bisogna tuttavia notificare che tale metrica di quantificazione di controversia è quella che presenta, tra quelle basate sul concetto di *Random Walks*, il divario minore tra realtà con controversia e non; inoltre, stando ai risultati ottenuti anche su *#BeefBan*, qualsiasi logica di *Link prediction*, porta a una riduzione significativamente importante di tale metrica.

Le attività di *Link prediction* hanno permesso di ottenere, per ogni metrica utilizzata, valori di controversia sotto la soglia presentata dal grafo privo di polarizzazioni; è da notare che le dimensione del grafo *#Ukraine* risultano essere moderatamente superiori a quelle presentate da *#BeefBan*, di conseguenza il numero di archi inserito, sempre pari al 30% degli archi pre-esistenti, sembra causare una decrescita più significativa della controversia.

Discussione

La valutazione dei risultati ottenuti dalle tecniche di *Link prediction* sui grafi allo stato dell'arte, costituiscono degli utili indici per comprendere la validità delle logiche che si sono adottate; inoltre forniscono una panoramica circa le metriche più propense alla quantificazione della controversia.

Dati i valori registrati sulle reti presentate, sembrerebbe che la riduzione più costante e importante della controversia, venga effettuata mediante un collegamento tra utenti “normali” di una comunità e utenti “autorevoli” per grado dell'altra.

Le modalità *Top Betwennnes Linking* e *Top Degree Linking*, registrano risultati degni di nota solo nelle metriche alle quali sono più orientati: *GMCK* e *RWTD*; questo di fatto non è utile in quanto sarebbe necessario identificare una tecnica di *Link prediction* volta alla massima minimizzazione, o comunque sia ad ottime performance, indipendentemente dalla metrica di quantificazione della controversia utilizzata.

Per ciò che concerne le soluzioni basate sulla prossimità, i risultati ottenuti sono mediamente, per le metriche *CSC* e *RWC*, migliori rispetto alle tecniche di *linking* sopra discusse; tuttavia, non riescono mai a ottenere risultati degni di nota se non, in parte, per la valutazione della controversia quantificata mediante *GMCK* in *#BeefBan* (Figura 5.16b) e per *RWTD*. Ad ogni modo in quelle metriche il risultato migliore è detenuto, come si diceva prima, dalle tecniche di *Top Betwennnes Linking* e *Top Degree Linking* rispettivamente.

Si passa ora, tenendo sempre a mente i risultati qui ottenuti, allo studio dei risultati registrati per i grafi Covid-19 e *Vaccination*.

5.4.2 Covid-19

Vengono ora forniti, in maniera analoga a quanto mostrato precedentemente per i grafi *#BeefBan* e *#Ukraine*, i risultati circa la riduzione della controversia per il grafo *Covid-19* (Figura 5.18).

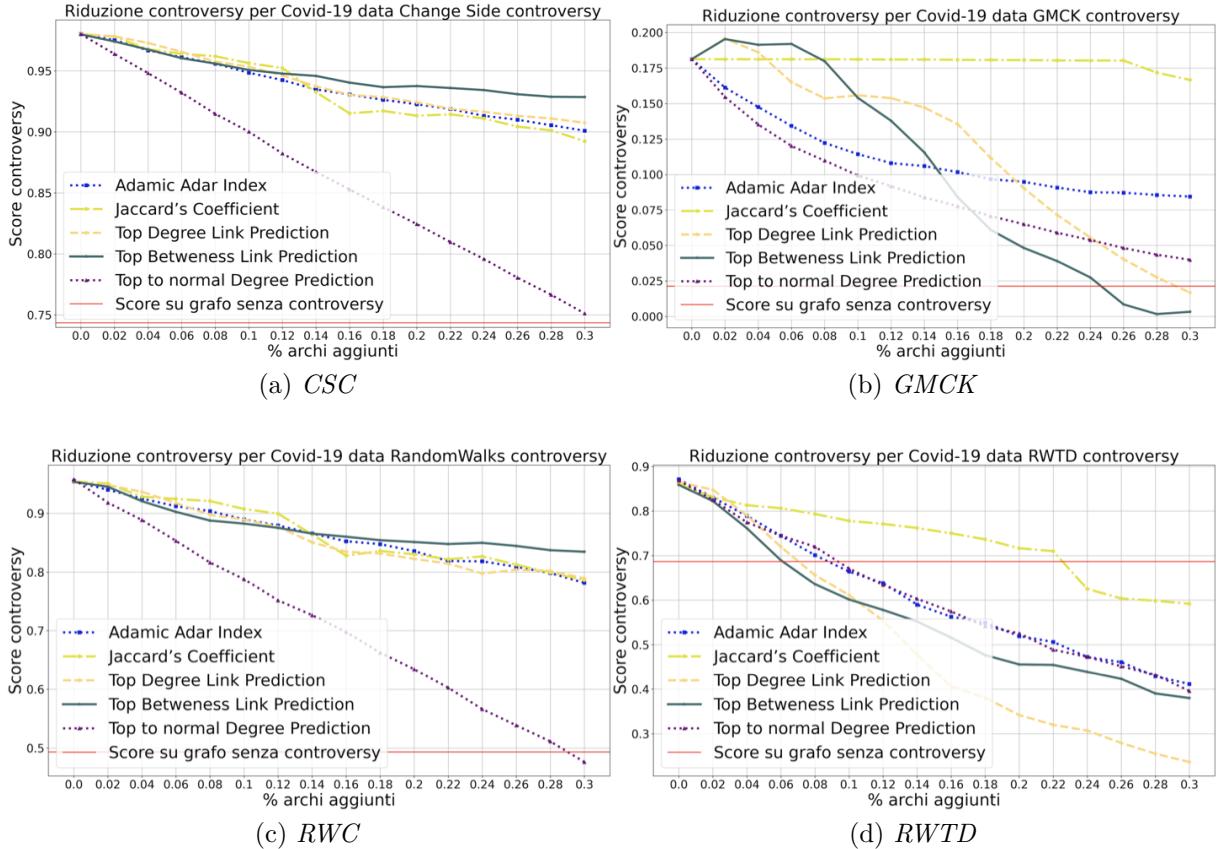


Figura 5.18: Riduzione controversia - *Covid-19*

La tendenza evidenziata precedentemente viene qui ulteriormente confermata, la pratica di *Link prediction* con logica *Top to Normal Linking* sembrerebbe superare significativamente tutte le altre in termini di riduzione di controversia. Solamente la valutazione fornita dalle metrica *GMCK* vede primeggiare l'approccio *Top Betweeness Linking*, lo stesso poi è più orientato ad abbassare la controversia quantificata dalla metrica utilizzata.

Interessante osservare, dati i grafici in Figure 5.18b e 5.18d, le attività di *Link prediction* che osservano le riduzioni minori sono quelle basate sulla valutazione della prossimità; questo può dunque suggerire che tali metodologie, su grafi particolarmente grossi rappresentanti reti sociali, possano non essere indicate al fine di rendere più fluido il movimento all'interno dei grafi, superando quindi situazioni di polarizzazione.

5.4.3 Vaccinazioni

Ultimo grafo utilizzato in questa tesi, *Vaccination* è stata forse il più “particolare” dei grafi utilizzati e sul quale, solo a fronte di un’analisi sui metadati, è stato possibile verificare come le *community* identificate risultassero essere valide.

Capitolo 5. Risultati

Si ricorda infatti che i dati raccolti circa le vaccinazioni, non sono stati raccolti durante un periodo storico di rilevanza a fronte di un determinato fatto di cronaca; il numero di archi in funzione al numero di nodi è estremamente basso se viene poi confrontato con il grafo definito per *Covid-19* tale tendenza viene ulteriormente evidenziata (Tabelle 4.2 e 4.3).

I nodi tra di loro risultavano poco interconnessi, situazione opposta a quanto mostrato per *Covid-19*.

Si ricorda anche che i valori di *modularity* registrati per le partizioni identificate su questo grafo, mostrati nella Tabella 5.5, sono più tendenti a realtà non dotata di strutture interne ben definite; tuttavia, i valori di *coverage* hanno suggerito di proseguire con l'analisi.

Infatti, l'analisi sul contenuto scambiato nelle due comunità ha registrato la presenza di due diverse visioni.

Ad ogni modo vengono mostrati nella Figura 5.19, anche per questo grafo, i risultati di riduzione di controversia a fronte dell'attività di *Link prediction*.

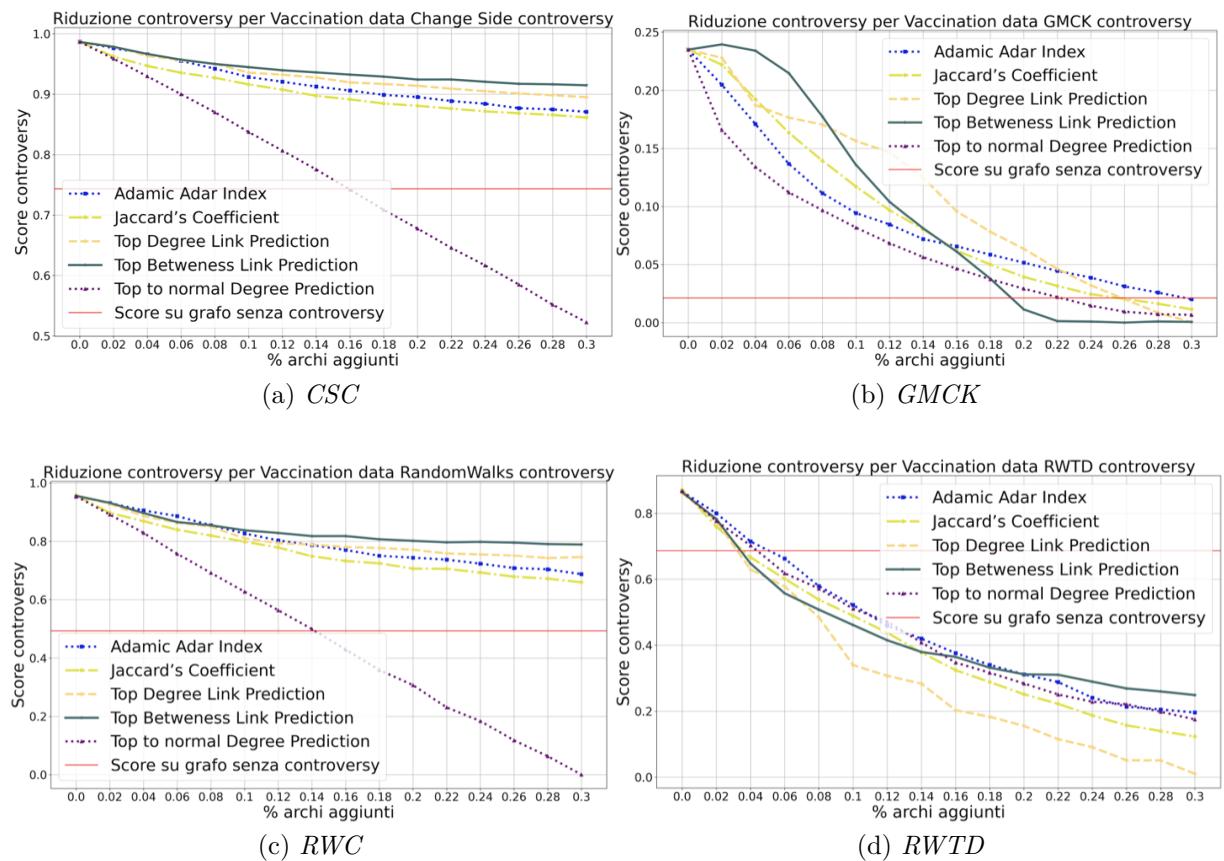


Figura 5.19: Riduzione controversia - Vaccination

Se l'approccio di *Link prediction* che sembra ottenere le migliori performance di riduzione della controversia si riconferma essere il *Top to Normal Linking*; quello che per la prima volta viene osservato è una riduzione importante della controversia date le metriche *CSC*, *RWC* e *RWTD* (rispettivamente le Figure 5.19a, 5.19c e 5.19d).

A fronte di quanto ottenuto in termini di riduzione, sulle precedenti reti, i risultati che vengono registrati potrebbero essere stati direttamente causati dalla struttura del grafo, la quale modella un argomento che è passato alla cronaca per essere molto controverso, soggetto a polarizzazioni tra i partecipanti alle possibili discussioni che si creano intorno allo stesso, ma che di fatto vede poca partecipazione in termini di contenuti scambiati in funzione dei dati in nostro possesso.

5.4.4 Valutazione del *Sentiment Boost*

Nell'ottica di fornire una “probabilità di accettazione” è stato modellato, come approfondito nella Sezione 4.4.3, un moltiplicatore da applicare allo score di *linking* per rendere più verosimile l'aggiunta di un arco.

Verranno ora forniti, per le reti che modellano i dati su Covid-19 e *Vaccination*, i confronti più significativi circa la riduzione di controversia. Il resto dei grafici sarà presentato nell'appendice B.

Covid-19

In Figura 5.20 vengono forniti i risultati circa le differenze tra le riduzioni apportate dall'attività di *Link prediction* con e senza *Sentiment Boost*, per le logiche basate sulla prossimità.

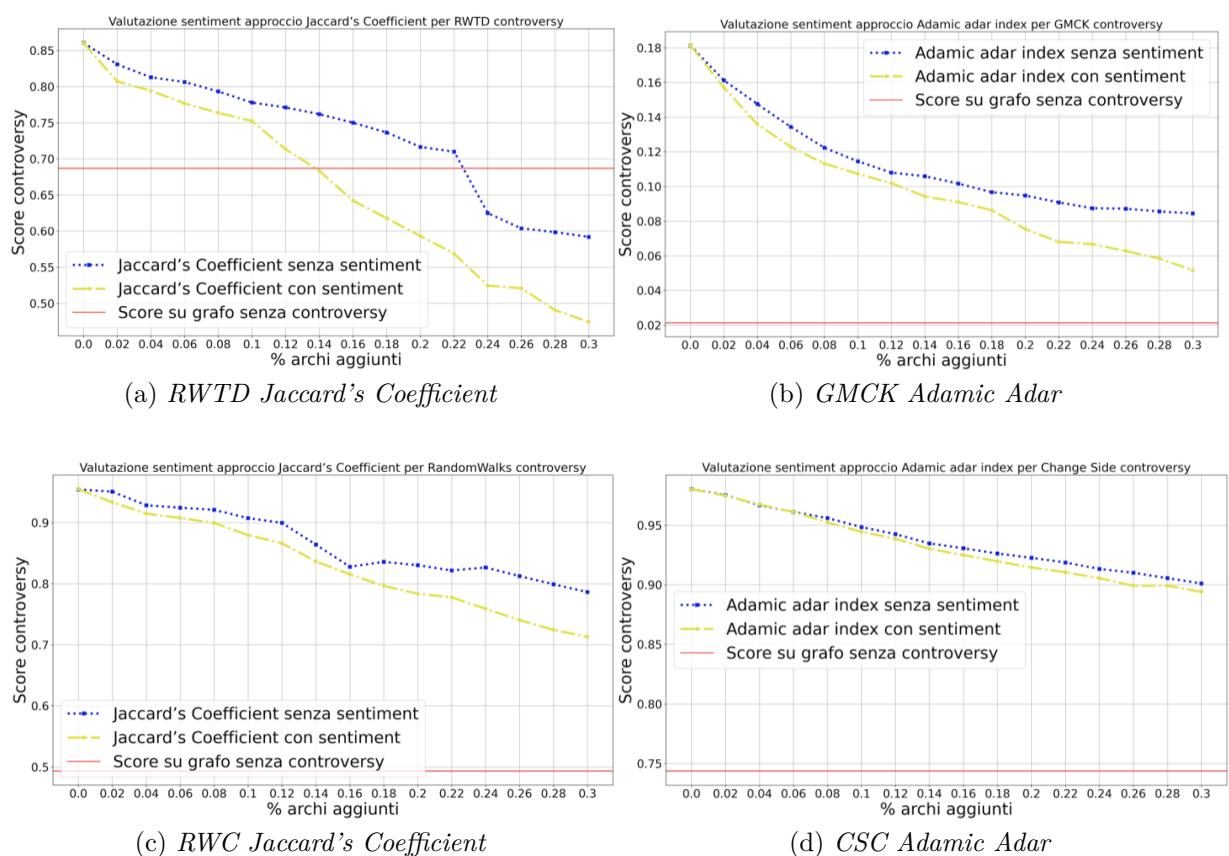


Figura 5.20: Valutazione *Sentiment Boost* per approcci basati sulla prossimità - Covid-19

Come si può notare da tutti e quattro i grafici proposti, pare che le logiche di *Linking* basate su approcci che valutano la prossimità, vedano un aumento significativo di *performance* (intesa come riduzione della controversia) ottenuto grazie al *Sentiment Boost*.

Non si prevedeva tale comportamento in quanto, essendo l'attività di *Link prediction* volta all'aggiunta dei migliori archi secondo un qualche metrica, si pensava che l'utilizzo di un moltiplicatore con lo scopo di rendere più verosimile l'aggiunta di un arco comportasse un calo delle *performances*.

L'aumento delle *perfomance* è significativamente presente da subito: in Figura 5.20a già dall'aggiunta pari al 2% di archi si nota l'inizio dello stacco che porterà a un divario di quasi 0.2 sulla controversia finale; lo stesso si può notare per le Figure 5.20c e 5.20b.

Sembra che in particolare il *linking* basato sul *Jaccard's Coefficient* sia più soggetto a questo miglioramento rispetto all'approccio basato sull'*Adamic Adar Index*; la causa di tale comportamento è probabilmente causata dalla formula volta a quantificare lo score associato a un nuovo potenziale arco.

Invece, risultati in linea con le aspettative per le logiche di *linking* basate sulla centralità e sulle peculiarità dei nodi, l'utilizzo del *Sentiment Boost* ha registrato un decremento delle *perfomance*.

La Figura 5.21 fornisce una rapida panoramica sugli andamenti più significativi date le quattro metriche utilizzate.

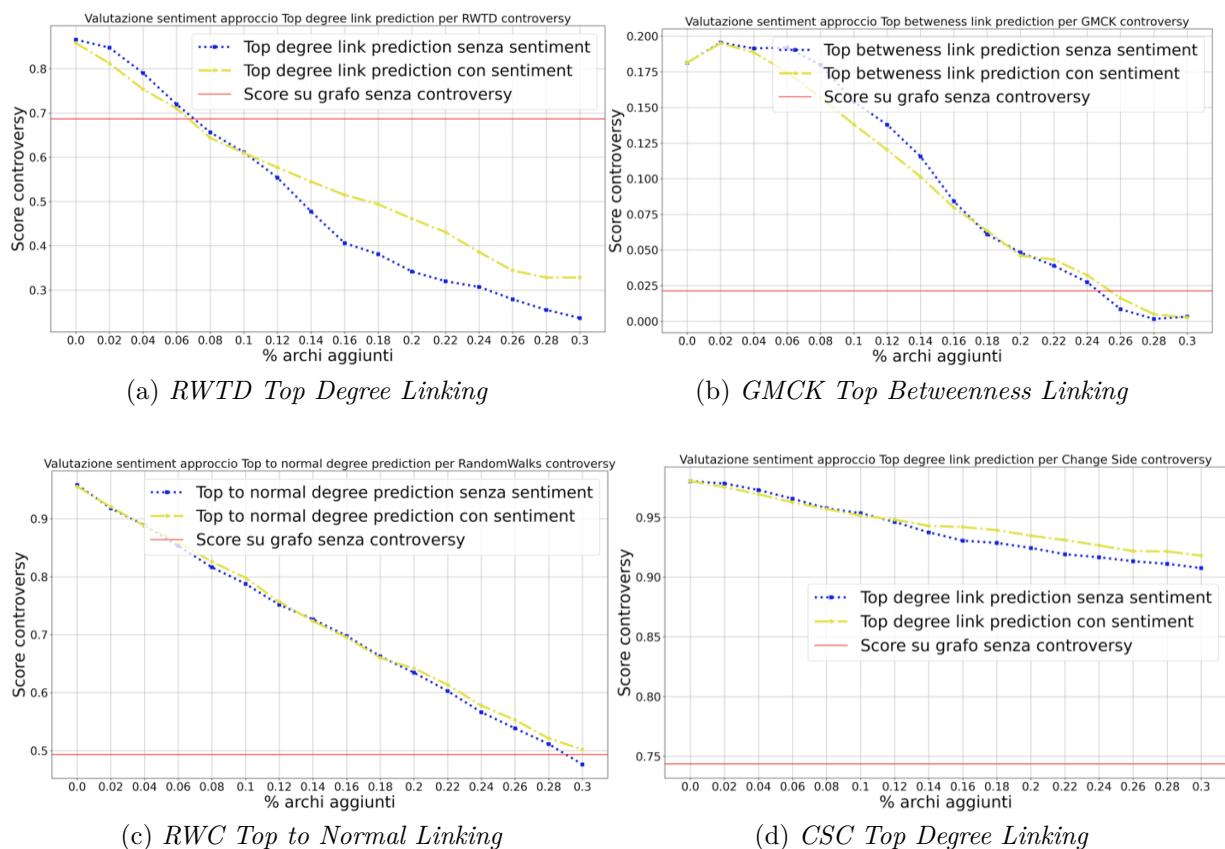


Figura 5.21: Valutazione *Sentiment Boost* per approcci basati sulla centralità - Covid-19

Si osserva come l'approccio che valuta il moltiplicatore *Sentiment Boost* sembra sempre notificare un calo più o meno marcato delle *performance*, questo è riconducibile al fatto che l'aggiunta mediante i diversi approcci di *Linking* proposti dipende da un certo *score*, questo *score* è computato in maniera del tutto teorica, non tenendo quindi in conto di nessuna peculiarità o similitudine degli oggetti collegati: l'obiettivo è semplicemente quello di massimizzare la funzione obiettivo.

Tenendo in conto una qualche proprietà di similitudine, al prezzo di un certo calo delle performance, si inserisce una modellazione circa la possibile veridicità del collegamento posto. Risulta interessante notare come l'approccio *Top to Normal Linking* non vede particolari peggioramenti delle performance (Figura 5.21c); questo probabilmente è dovuto alla natura del collegamento che permette, indipendentemente poi dall'arco posto, di esporre un certo utente di una partizione ai contenuti diffusi dai nodi centrali per grado dell'altra; questo comportamento sembrerebbe quello più premiato, in generale, per riduzione di controversia.

Vaccinazioni

In maniera analoga a quanto fatto per *Covid-19*, vengono fornite in Figura 5.22 le differenze di performance registrate dalle metodologie di *linking* basate sulla prossimità, dato l'utilizzo del *Sentiment Boost*.

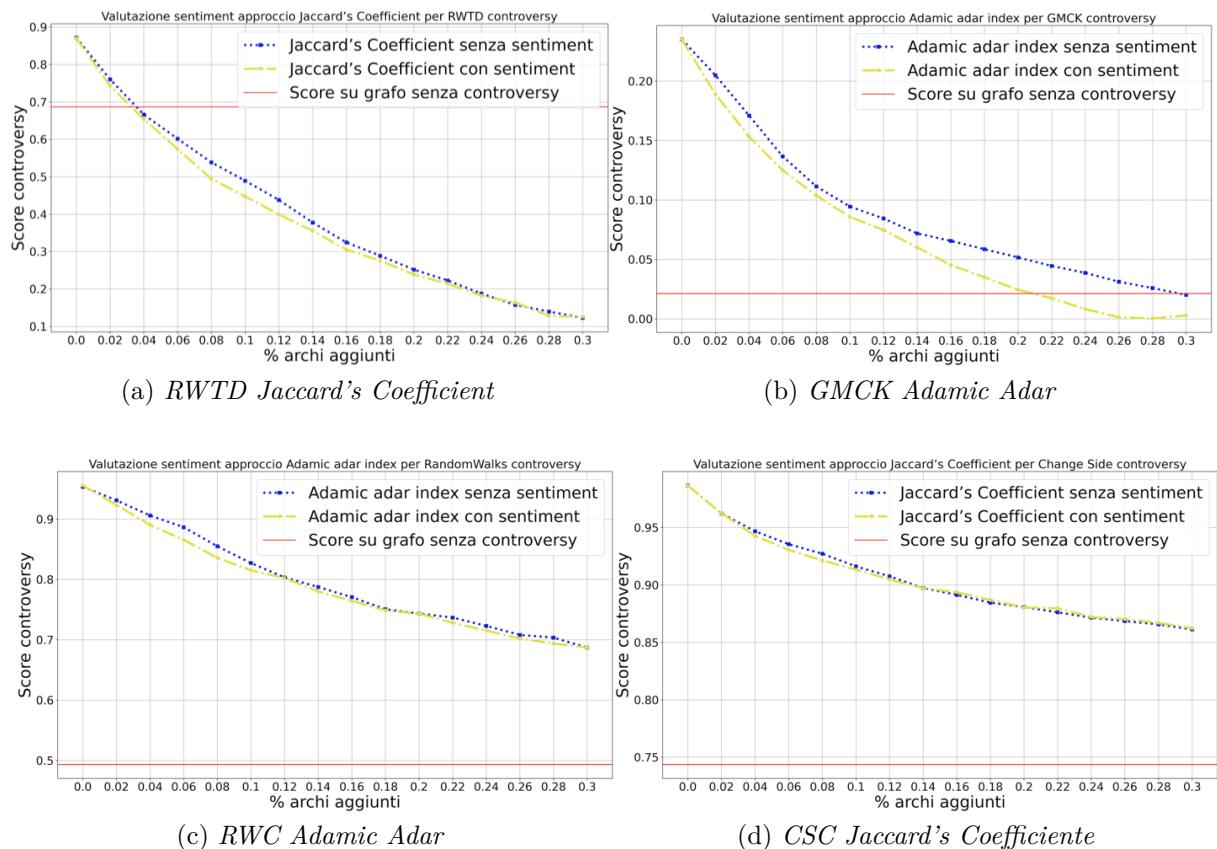


Figura 5.22: Valutazione *Sentiment Boost* per approcci basati sulla prossimità - Vaccination

A differenza di quanto osservato per *Covid-19*, l'utilizzo del *Sentiment Boost* non sembra causare una grande differenza in termini di riduzioni di controversia; sicuramente vi sono delle differenze sull'andamento registrato, ma questo non è confrontabile con quanto osservato precedentemente.

La struttura poco densa di questo grafo fa sì che l'eventuale aggiunta, di un qualsiasi arco con una logica basata sulla prossimità, che congiunge due nodi appartenenti alle due opposte partizioni, porti a una inevitabile riduzione di controversia. Probabilmente, data la poca densità di questo grafo, la valutazione della prossimità potrebbe non essere la metodologia migliore per superare la controversia identificata.

L'unica metrica che, in linea a quanto visto precedentemente, sembra giovare del *Sentiment Boost* è *GMCK* con l'utilizzo dell'indice *Adamic Adar*.

Viene infine mostrata, in Figura 5.23, le più significativa differenze di performance registrate in funzione di ogni metrica, dato l'utilizzo del *Sentiment Boost*, per gli approcci basati sulla centralità.

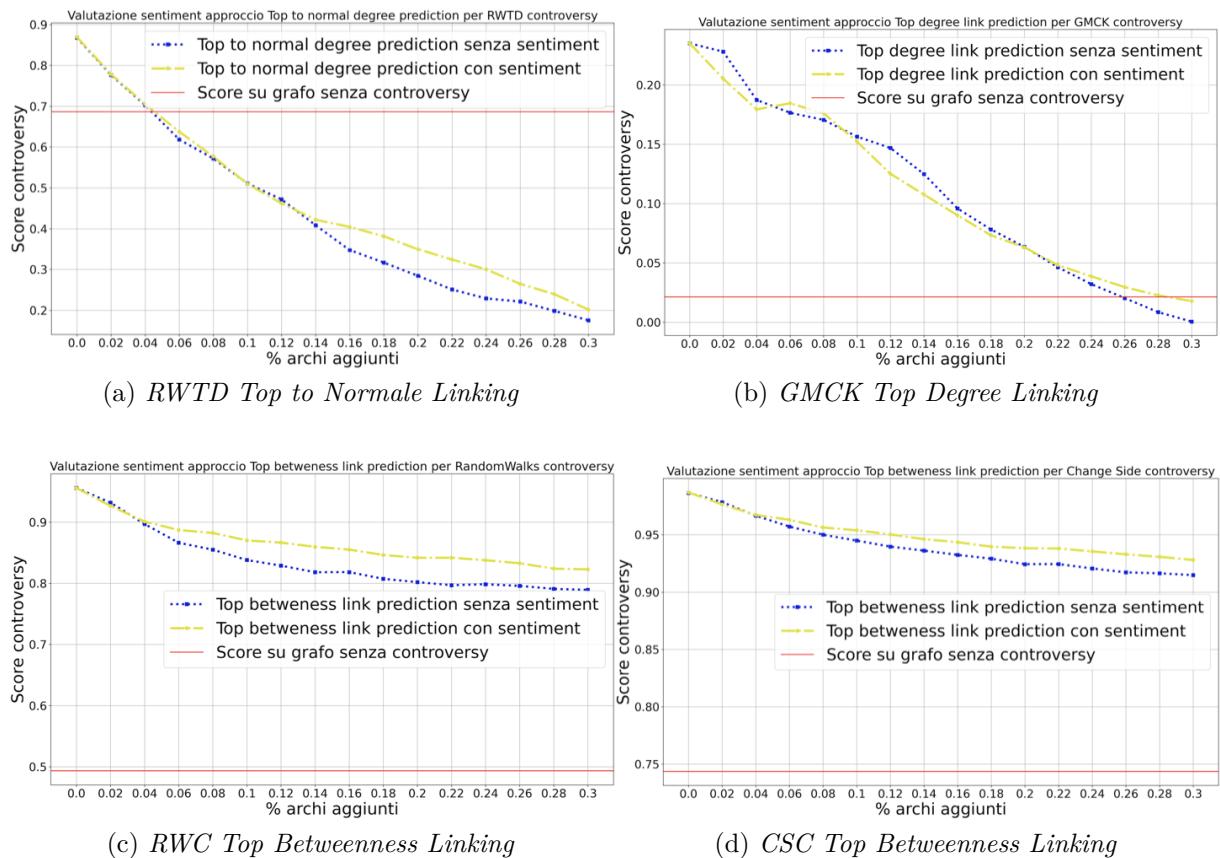


Figura 5.23: Valutazione *Sentiment Boost* per approcci basati sulla centralità - Vaccination

I risultati ottenuti e mostrati sopra, rivelano come anche in questo caso l'utilizzo di un indice volto a quantificare una sorta di “probabilità di accettazione”, porti a una riduzione delle *performance*; si ricorda che la valutazione di tale moltiplicatore è fondamentale, nel tentativo di aumentare le probabilità circa l’effettiva creazione dell’arco.

5.4.5 Analisi post *Link prediction*

Al fine di valutare ulteriormente l'attività di *Link prediction*, con l'obiettivo di verificare se la logica di *linking* che permetteva la riduzione più veloce e significativa di controversia, fosse poi quella che comportava anche “l'indebolimento” delle partizioni identificabili nella rete, si è proceduto ad attuare delle analisi volte a quantificare questo aspetto.

Si è quindi ripetuta l'attività di *Community detection*, andando a verificare se vi fosse effettivamente una riduzione di *modularity* e *coverage*, valutando anche i contenuti scambiati in modo da verificare l'eventuale omogeneità degli stessi e confrontarli con quelli ottenuti prima dell'attuazione di questo *task*.

#BeefBan e #Ukraine

Data l'assenza di informazione semantica scambiata, le analisi effettuate dopo l'attività di *Link prediction* si sono limitate alla verifica circa l'eventuale indebolimento, date tutte le metodologie utilizzate, delle eventuali strutture interne delle reti; si va di fatto a ripetere l'attività di *Community detection*, al fine di valutare gli score di *Modularity* e *Coverage*.

#BeefBan La Tabella 5.11 mostra i risultati raccolti in seguito alle analisi fatte:

Modalità di Linking	#BeefBan			
	Modularity	Coverage	Differenza Modularity	Differenza Coverage
<i>Adamic Index</i>	0.4422	0.9671	-0.0336	-0.0127
<i>Jaccard's Coefficient</i>	0.4611	0.9737	-0.0147	-0.0061
<i>Top Degree Linking</i>	0.4032	0.9079	-0.0726	-0.0719
<i>Top Betweenness Linking</i>	0.4065	0.9092	-0.0693	-0.0706
<i>Normal to Top Degree Linking</i>	0.3523	0.8551	-0.1235	-0.1247

Tabella 5.11: Risultati *Community detection* post *Link prediction* -
#BeefBan

Risulta possibile osservare un calo delle metriche di *Modularity* e *Coverage* qualsiasi sia l'approccio utilizzato; sembra però che le metodologie basate sulla prossimità, nonostante i buoni risultati circa la minimizzazione della controversia presentati precedentemente, registrino le riduzioni minori per gli indici volti a quantificare la qualità delle partizioni identificate.

Questo può essere dovuto al fatto che tali metodologie, basandosi appunto sulla prossimità, vadano a collegare solo nodi relativamente “vicini” tra loro (in termini di *path length*); utilizzando tali metriche non si osserverà mai collegamenti tra nodi i quali non condividono nemmeno un vicino.

Le modalità basate sulla prossimità sembrerebbero essere in grado di portare a realtà dove la nuova rete, può vedere la definizioni di nuove, fittizie, *community*.

Approcci basati invece sulla centralità, tendenzialmente, collegano anche nodi relativamente “distanti”, andando di fatto a indebolire le strutture che si possono andare a creare

nella rete.

Interessante notare come la modalità di *linking* che più di tutte sembrava ridurre la controversia, tale *Top to Normal Linking*, ha fatto registrare i valori più bassi per *Modularity* e *Coverage*.

In generale, stando a quanto riportato in Tabella 5.3, le partizioni identificate presentano valori tendenti, molto più di prima, a una realtà priva di controversia e quindi di gruppi polarizzati.

#Ukraine Seconda rete facente parte dell'insieme di grafi caratterizzanti il “benchmark” per le attività attuate in questa tesi, anche per *#Ukraine* verranno mostrati i risultati successivi all’attività di *Link prediction*, volti a quantificare la qualità delle nuovi partizioni ottenute.

La Tabella 5.12 mostra i risultati ottenuti.

Modalità di Linking	#Ukraine			
	Modularity	Coverage	Differenza Modularity	Differenza Coverage
<i>Adamic Index</i>	0.4384	0.9418	-0.0301	-0.0279
<i>Jaccard's Coefficient</i>	0.4541	0.9742	-0.0144	-0.0045
<i>Top Degree Linking</i>	0.3704	0.9001	-0.0981	-0.0696
<i>Top Betweenness Linking</i>	0.3762	0.8782	-0.0923	-0.0915
<i>Normal to Top Degree Linking</i>	0.3604	0.8605	-0.1081	-0.1092

Tabella 5.12: Risultati *Community detection* post *Link prediction* - *#Ukraine*

I valori mostrati sembrano del tutto in linea a quanto visto precedentemente, le metriche basate sulle logiche di prossimità registrano la diminuzione più bassa in termini di *modularity* e *coverage*; tutte le altre tre metodologie di *linking* basate sulla centralità, vedono indebolimenti più significativi delle *community* definite.

L’approccio che più di tutti, in seguito al ripetersi delle attività di *Community Detection*, ha visto la diminuzione più significativa delle metriche volte a quantificare la qualità delle partizioni identificate, è il *Normal to Top Degree Linking*.

Dai risultati ottenuti su queste due reti sembrerebbe emergere un fatto importante: le metodologie di *linking* basate sulla prossimità, sebbene registrino una diminuzione di controversia generalmente maggiore rispetto a logiche orientate alla centralità, non sembrerebbero causare un indebolimento importante delle strutture interne alla rete.

Questo comportamento potrebbe essere causato dal funzionamento “social oriented” di queste metodologie allo stato dell’arte.

Covid-19

Vengono fornite le analisi portate avanti in seguito all’attività di *linking*, in questo caso le analisi non si sono limitate sole alla valutazione della “forza” delle nuove strutture interne, ma si è andato a valutare anche le conseguenze circa l’utilizzo del *Sentiment Boost*.

Si è inoltre potuto procedere alla valutazione dei metadati, quali *wordcloud* circa le descrizioni dei profili all'interno dei nuovi partizionamenti, *sentiment distribution* e informazioni circa gli account verificati.

Il grafo sul quale verranno portate avanti le analisi circa i metadati, sarà quello sul quale verrà registrato il maggior “indebolimento” delle strutture interne identificate.

Community detection La Tabella 5.13 mostra i risultati ottenuti per l'attività di *Community detection*, effettuata in seguito all'aggiunta dei nuovi collegamenti.

Vengono forniti, nella parte superiore della tabella, i risultati per l'approccio strettamente “teorico”, che non tiene in considerazione la “probabilità di accettazione” che si è tentato di modellare mediante l'utilizzo del *Sentiment Boost*; nella seconda parte della tabella vengono invece proposti i risultati che hanno visto l'utilizzo di tale moltiplicatore.

Modalità di Linking	Senza <i>Sentiment Boost</i>			
	Modularity	Coverage	Differenza Modularity	Differenza Coverage
<i>Adamic Index</i>	0.3662	0.9032	-0.0741	-0.0372
<i>Jaccard's Coefficient</i>	0.4291	0.9305	-0.0112	-0.0099
<i>Top Degree Linking</i>	0.3277	0.8388	-0.1125	-0.1016
<i>Top Betweenness Linking</i>	0.374	0.8744	-0.0663	-0.066
<i>Normal to Top Degree Linking</i>	0.3036	0.8037	-0.1367	-0.1367
Con <i>Sentiment Boost</i>				
<i>Adamic Index</i>	0.3705	0.9049	-0.0698	-0.0355
<i>Jaccard's Coefficient</i>	0.4327	0.932	-0.0076	-0.0084
<i>Top Degree Linking</i>	0.3117	0.8294	-0.1286	-0.111
<i>Top Betweenness Linking</i>	0.3656	0.8687	-0.0747	-0.0717
<i>Normal to Top Degree Linking</i>	0.3023	0.8033	-0.138	-0.1371

Tabella 5.13: Risultati *Community detection* post *Link prediction - Covid-19*

In maniera del tutto analoga a quanto mostrato precedentemente, tutte le modalità di *linking* vedono una diminuzione delle metriche volte a quantificare la qualità delle partizioni identificate, tale risultato sottolinea come le strutture interne alla rete vedano di fatto, data l'aggiunta di nuovi archi, un indebolimento della loro struttura.

Ancora una volta approcci basati sulla prossimità vedono una riduzione molto più contenuta, rispetto agli altri approcci.

Passando poi alla valutazione dei risultati ottenuti in seguito all'attività di *Link prediction* con inserimento del *Sentiment Boost*, si rimane sorpresi dal fatto che metodologie che sembravano, in termini di riduzione di controversia, beneficiare del moltiplicatore, osservano peggioramenti delle performance: le comunità che vengono identificate registrano valori di *modularity* e *coverage* maggiori rispetto a quanto visto precedentemente, indice che le partizioni sono più “forti” rispetto a quanto mostrato nella parte superiore della tabella.

Al contrario, le metodologie che vedevano un peggioramento delle performance in termini di riduzione della controversia, osservano ora un miglioramento in termini di riduzione di *modularity* e *coverage*: le partizioni identificate sono più “deboli”.

Rispetto alla relazione riduzione controversia/riduzione *modularity*, solo la metodologia *Normal to Top Degree* sembra allinearsi alle supposizioni iniziali; di fatti, date le riduzioni di controversia sostanzialmente identiche con e senza *Sentiment Boost*, le partizioni ottenute non sembrano differire di molto in termini di metriche.

Tale situazione necessita di essere ulteriormente investigata, al fine di comprendere correttamente la relazione tra riduzione controversia/riduzione *modularity* (e *coverage*).

Ad ogni modo, la metodologia che più di tutte sembrerebbe comportare una maggior diminuzione delle metriche volte a quantificare la qualità delle partizioni, risulta essere la *Normal to Top Degree* con *Sentiment Boost*; sono quindi state condotte analisi circa i metadati in nostro possesso, al fine di verificare che le partizioni risultino essere più omogenee rispetto a quanto inizialmente osservato.

Analisi metadati Le analisi effettuate in questa fase hanno visto l'utilizzo di parte dei metadati già utilizzati precedentemente, permettendo così un più facile confronto.

Partendo da una valutazione circa la distribuzione del *sentiment Medio* degli utenti verificati, mostrata in Figura 5.24, è possibile notare come la differenza precedentemente identificata, sembra essere stata del tutto superata in seguito all'attività di *Link Prediction*.

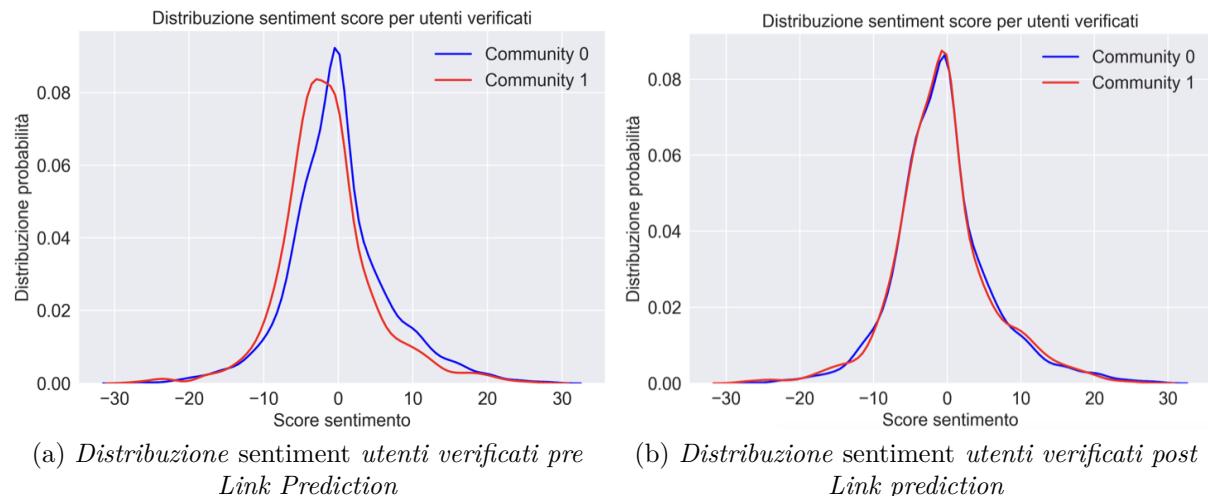


Figura 5.24: Valutazione *sentiment distribution* utenti verificati pre e post *Link prediction* - *Covid-19*

Altra utile analisi è stata condotta sulle descrizioni dei profili facenti parte delle due nuove partizioni identificate (Figura 5.25).

Pare subito chiaro che le due partizioni vedano sostanzialmente le stessa parole, se prima parole come *KAG*, *MAGA*, *Conservative*, *Trump Health*, *Science*, *Research* caratterizzavano o una o l'altra comunità, ora la situazione sembrerebbe più omogenea.

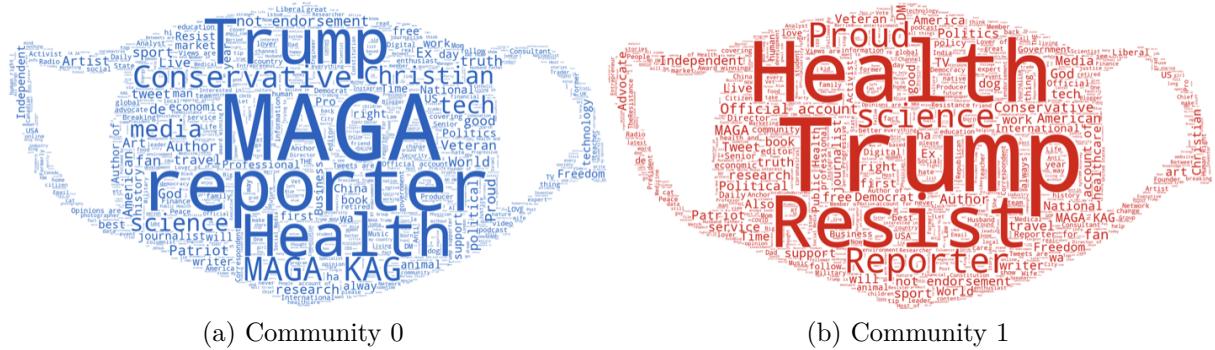


Figura 5.25: *Wordcloud* date le descrizioni dei profili degli utenti in funzione della *community* post *Link prediction - Covid-19*

Analizzando poi il numero di account verificati per ogni partizione, pare chiara una riduzione del *gap* precedentemente registrato, rispetto alla Figura 5.10 in precedenza mostrata, dove si registrava un divario pari a 10 punti.

Con il nuovo partizionamento ottenuto, mostrato in Figura 5.26, il *gap* si è ridotto a poco più di 5 punti.

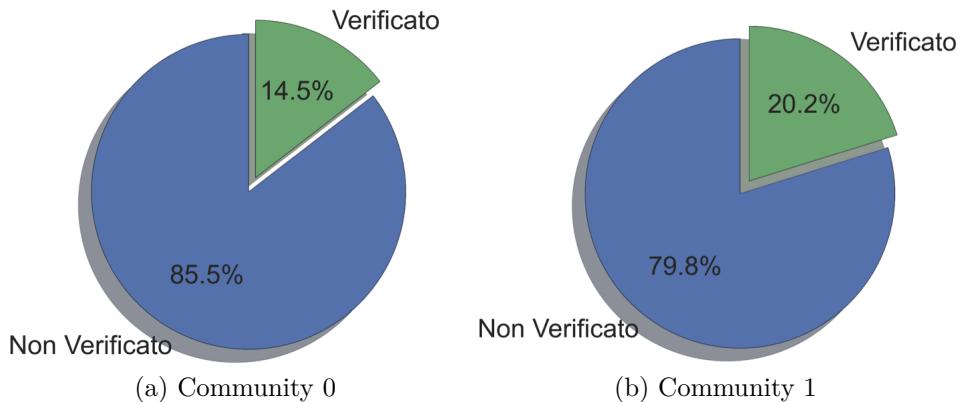


Figura 5.26: Percentuale account verificati in funzione della community post *Link prediction - Covid-19*

In ultimo si vuole mostrare la distribuzione del *sentiment* medio tra gli utenti delle due partizioni identificate, al fine di permettere il confronto con la situazione precedente.

Dalla Figura 5.27 si può notare come ci si stia muovendo verso una totale omogeneità nel sentimento racchiuso nei contenuti prodotti.

Vaccinazioni

Ultimo grafo utilizzato in questo lavoro, come più volte ripetuto in tutto in tutta la tesi, questa rete è forse la più particolare in termini di struttura; la partecipazione alla discussione circa il l'argomento modellato è molto bassa.

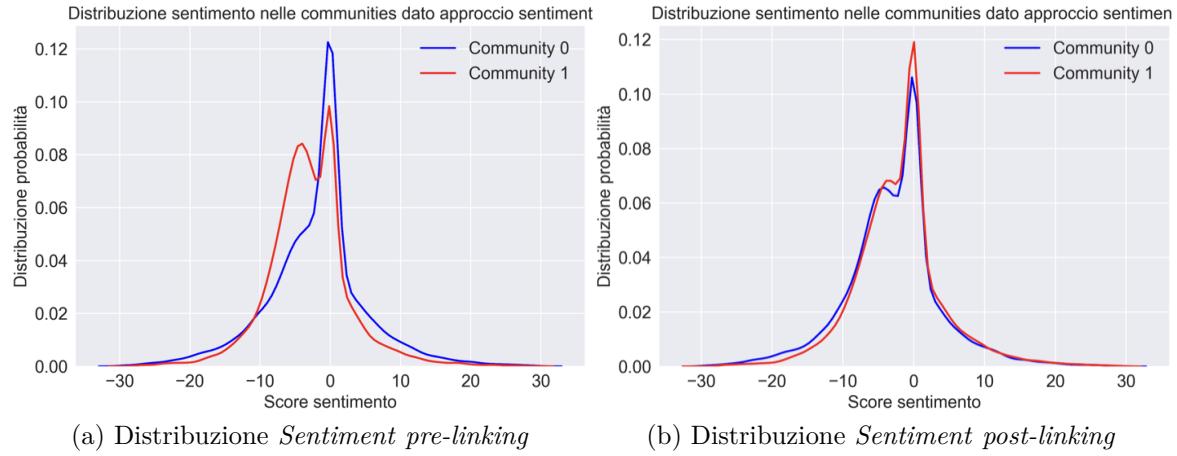


Figura 5.27: Distribuzione dello score medio di *Sentiment* delle communities identificate - *Covid-19*

L'argomento “vaccinazioni” è per sua natura terreno fertile per polarizzazioni, tuttavia i dati raccolti potrebbero non essere stati sufficientemente numerosi per modellare la contrapposizione che si tende a formare intorno al *topic* stesso.

Passiamo ora, in maniera analoga a quanto visto precedentemente, a presentare le analisi eseguite in seguito all’attività di *Link prediction*; anche in questo caso ci è consentito effettuare analisi sia circa la qualità dell’attività di *Community detection post linking*, sia circa l’informazione semantica scambiata nella rete.

Community detection La Tabella 5.14 mostra i risultati ottenuti dal partizionamento ottenute mediante *Metis*, per ogni modalità di *linking* vengono fornite i risultati.

Modalità di Linking	Senza <i>Sentiment Boost</i>		Differenza Modularity	Differenza Coverage
	Modularity	Coverage		
<i>Adamic Index</i>	0.2486	0.9548	-0.045	-0.0305
<i>Jaccard’s Coefficient</i>	0.355	0.9702	0.0615	-0.015
<i>Top Degree Linking</i>	0.2	0.9371	-0.0935	-0.0481
<i>Top Betweenness Linking</i>	0.3587	0.9611	0.0652	-0.0242
<i>Normal to Top Degree Linking</i>	0.2667	0.8938	-0.0268	-0.0914
Con <i>Sentiment Boost</i>				
<i>Adamic Index</i>	0.2913	0.95	-0.0022	-0.0353
<i>Jaccard’s Coefficient</i>	0.3466	0.9647	0.0531	-0.0206
<i>Top Degree Linking</i>	0.2411	0.9208	-0.0524	-0.0645
<i>Top Betweenness Linking</i>	0.3388	0.937	0.0452	-0.0482
<i>Normal to Top Degree Linking</i>	0.2723	0.8973	-0.0212	-0.088

Tabella 5.14: Risultati *Community detection post Link prediction - Vaccination*

I risultati ottenuti sono molto “controversi”, stando ai grafici mostrati nella Sezione 5.4.3, la metodologia che più di tutte sembrava minimizzare la controversia pareva essere la *Top to Normal linking*, portando anche a situazioni di quasi ribaltamento totale (più RandomWalks che terminavano nella partizione opposta).

Dai risultati emerge come la metodologia *Top Degree linking* permetta di ottenere le riduzioni migliori in termini di *modularità*, tale situazione potrebbe essere direttamente causata dalla struttura della rete che modella l’argomento in esame.

Si ricorda infatti che la rete vede poca partecipazione da parte degli utenti; tale problematica sarà discussa e motivata in maniera più ampia nella Sezione 5.4.6 di questa tesi.

Ancora più “controversi” sono poi i risultati ottenuti mediante *Jaccard’s Coefficient* e *Top Betweenness Linking* che, stando ai grafici proposti in Figura 5.19, mostrano di fatto una riduzione di controversia (soprattutto per la prima metodologia).

Le partizioni identificate in seguito all’attività di *linking* osservano valore di *modularity* maggiori rispetto a quanto ottenuto precedentemente; i valori che vengono ora registrati denotano delle strutture interne degne di nota, che potrebbero aver senso di esistere.

Passando alla valutazione del *Sentiment Boost*, i valori registrati per le metodologie che hanno rilevato, con la valutazione del moltiplicare per la definizione dello *score*, una riduzione minore della controversia, vedono una diminuzione più contenuta dei valori di *modularity* e *coverage*.

Unico valore non in linea risulta essere quello registrato dall’*Adamic Index*, il quale vede una forte differenza tra ciò che viene registrato con e senza valutazione del *Sentiment Score*.

Risulta infine interessante notare come, tolto *Top to Normal linking*, le riduzioni per la *coverage* risultano essere molto limitate.

I valori alti di questa metrica avevano spinto inizialmente a proseguire alla valutazione e allo studio di questa rete; trattandola di fatto, anche a fronte dei sensati risultati sui metadati in nostro possesso, come dotata di controversia e polarizzazioni.

A fronte dei risultati ottenuti, forse questa rete non è la più utile al fine di comprendere una metodologia corretta volta a quantificare e minimizzare la controversia; ad ogni modo, ritenendo più importanti per questa rete i valori di *coverage* rispetto a quelli della *modularity*, è stato deciso di effettuare le analisi successive sulla rete ottenuta in seguito all’attività di *linking*, effettuata con logica *Top to Normal Linking* senza valutazione *sentiment*.

Analisi metadati Si è quindi proceduto alle analisi circa i metadati date le due nuove partizioni. La Figura 5.28 mostra il numero di utenti associati ai *topic*, date le due differenti partizioni.

Risulta fin da subito possibile notare come, per quasi la totalità dei *topic*, vi sia una situazione più omogena circa le discussioni interne alle partizioni.

Facendo poi riferimento alla Tabella 5.7, possiamo notare che i *topic* che sembravano essere più orientati a posizioni di *no-vax* (2, 4, 5, 7, 8, e 10) vedono un *gap* minore tra le

due *community*; stesso discorso può essere fatto per i *topic* più *pro-vax* (6, 9, 11 e 12), anche per loro il divario è quasi assente.

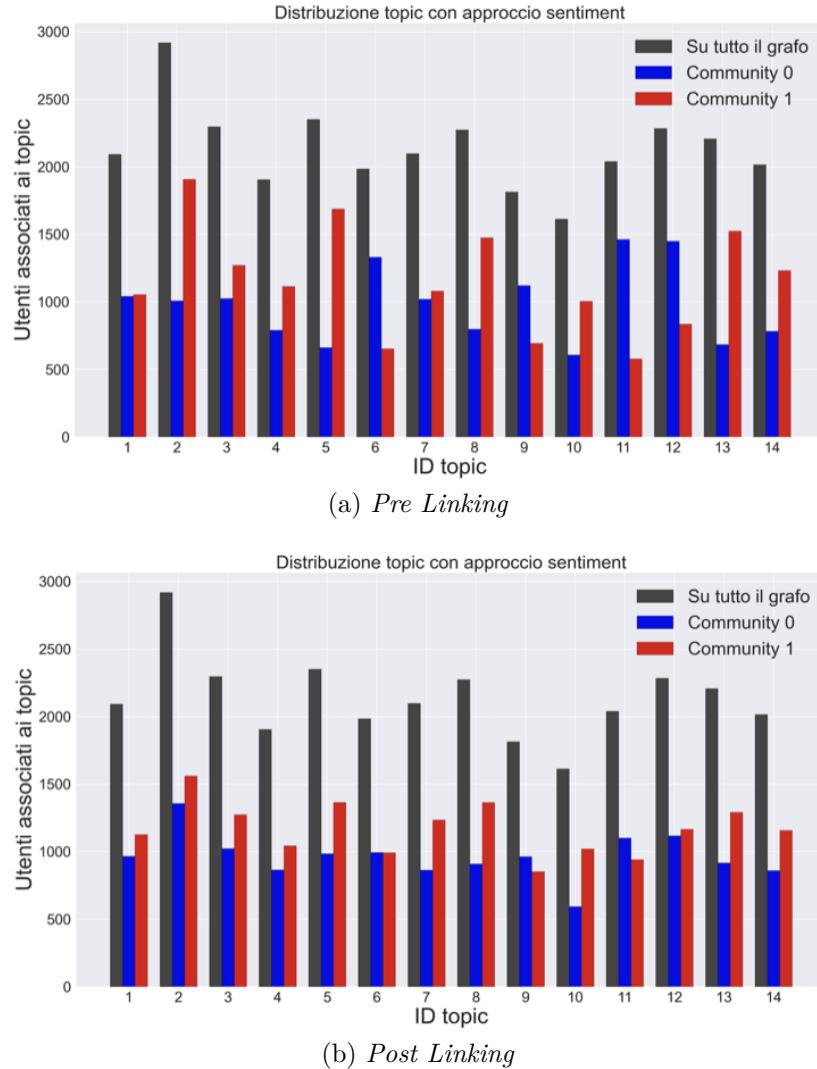


Figura 5.28: Numero di utenti per comunità associati ai *topic* - Vaccination

Questo risultato è molto importante, in quanto certifica come le due partizioni identificate vedano la stessa quantità di informazione condivisa per ogni *topic*, andando a superare quanto registrato precedentemente.

Un’ulteriore analisi svolta, volta a verificare l’effettiva situazione di omogeneità dei contenuti scambiati, ha visto la valutazione delle *keyword* più utilizzate all’interno dei *tweet* scambiati nei limiti della propria *community*. Tenendo a mente i risultati registrati precedentemente, raffigurati nella Figura 5.13, vengono fornite le due *wordcloud* facenti riferimento alle due nuove partizioni identificate.

Dalle *keyword* che emergono sembrerebbe essere superata la situazione precedentemente identificata; molte parole sono ora condivise tra le due *community*.

In particolare la parola *Vaccineswork* è ora presente in entrambe le partizioni, così come le parole *Health*, *Public Health*, *Support* e *Prevent*.

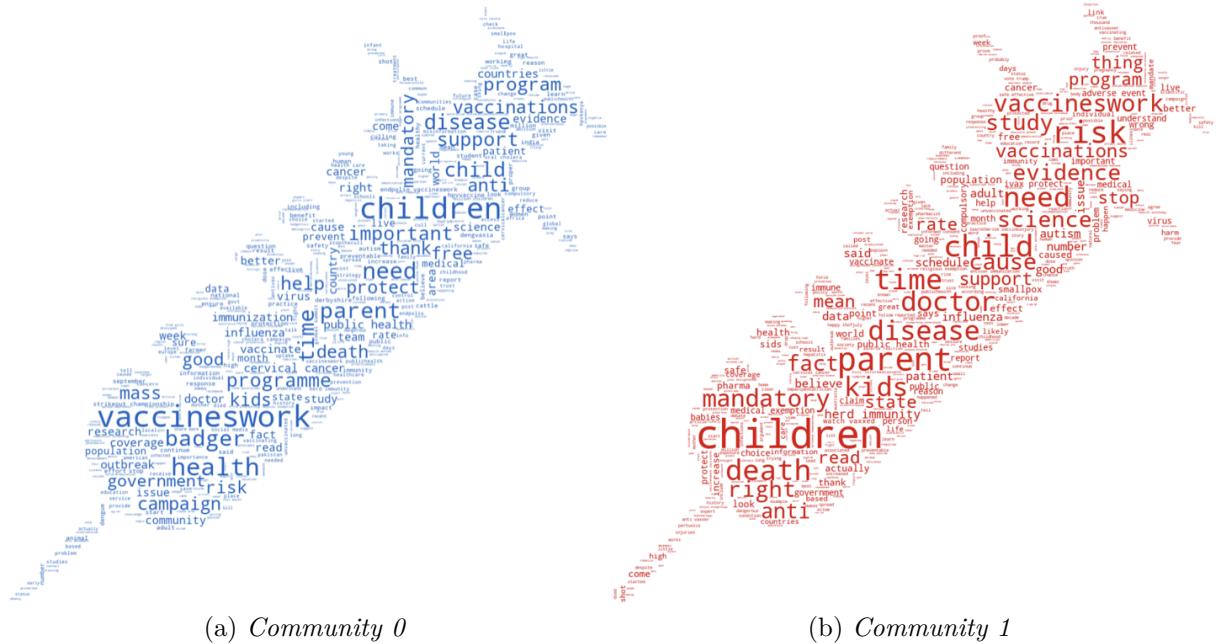


Figura 5.29: Wordcloud tweet intra-community - *Vaccination*

Questo risultato è in linea con quanto mostrato precedentemente, circa il numero di utenti per ogni *topic* date le partizioni identificate.

5.4.6 Discussione

L'attività di *Link prediction* ha permesso di valutare quali, tra i differenti approcci di *linking* proposti e utilizzati, riuscisse a indebolire le strutture delle possibili *community* che possono essere racchiuse in una rete.

I risultati ottenuti dalle metriche volte a quantificare la qualità delle partizioni, sembrano premiare approcci che cercano di collegare i nodi in funzione delle loro peculiarità strutturali, senza tenere conto di vicinato o similitudini in termini di collegamenti preesistenti. Si è infatti notato come gli approcci basati sulla prossimità, *Adamic Index* e *Jaccard's Coefficient*, registrano tendenzialmente delle riduzioni maggiori (anche se di poco) di controversia rispetto a *Top Degree Linking* e *Top Betweenness Linking*, tuttavia le riduzioni di *modularity* e *coverage* risultano essere molto più contenute.

Una possibile giustificazione di tali risultati è da ricercarsi nella natura stessa dello *score* attribuito dagli approcci basati sulla prossimità: un potenziale nuovo collegamento avrà uno *score* diverso da zero, solo nel caso in cui i nodi che si accinge a collegare osservano un vicino o più in comune. Questa politica di *linking* è molto *social-oriented*, partendo dalla supposizione che due utenti che condividono il vicinato possano essere più simili di nodi che non condividono collegamenti (che risultano dunque più lontani), i *social media* tengono in considerazione questo aspetto nell'attività di raccomandazione di nuove

amicizie, o più in generale di nuovi collegamenti (promozioni, sponsorizzazioni, ecc..); se questo può sembrare lecito e più verosimilmente potrà essere accettato, non sembra però essere adatto a ridurre la qualità delle strutture interne identificate.

Sicuramente uno studio più approfondito circa la relazione esistente tra riduzione controversia/indebolimento strutture interne alla reti, risulta necessario.

Un discorso totalmente a parte deve essere fatto per i risultati e le attività effettuate sulla rete *Vaccination*. Su questa rete si è assistito a un aumento delle metriche facenti riferimento alla qualità delle partizioni identificate, in seguito all'attività di *linking*; si è precedentemente giustificata tale situazione, andando definire la rete poco “densa”. Si fornisce quindi, in Figura 5.30, la distribuzione di grado su tutte le reti utilizzate.

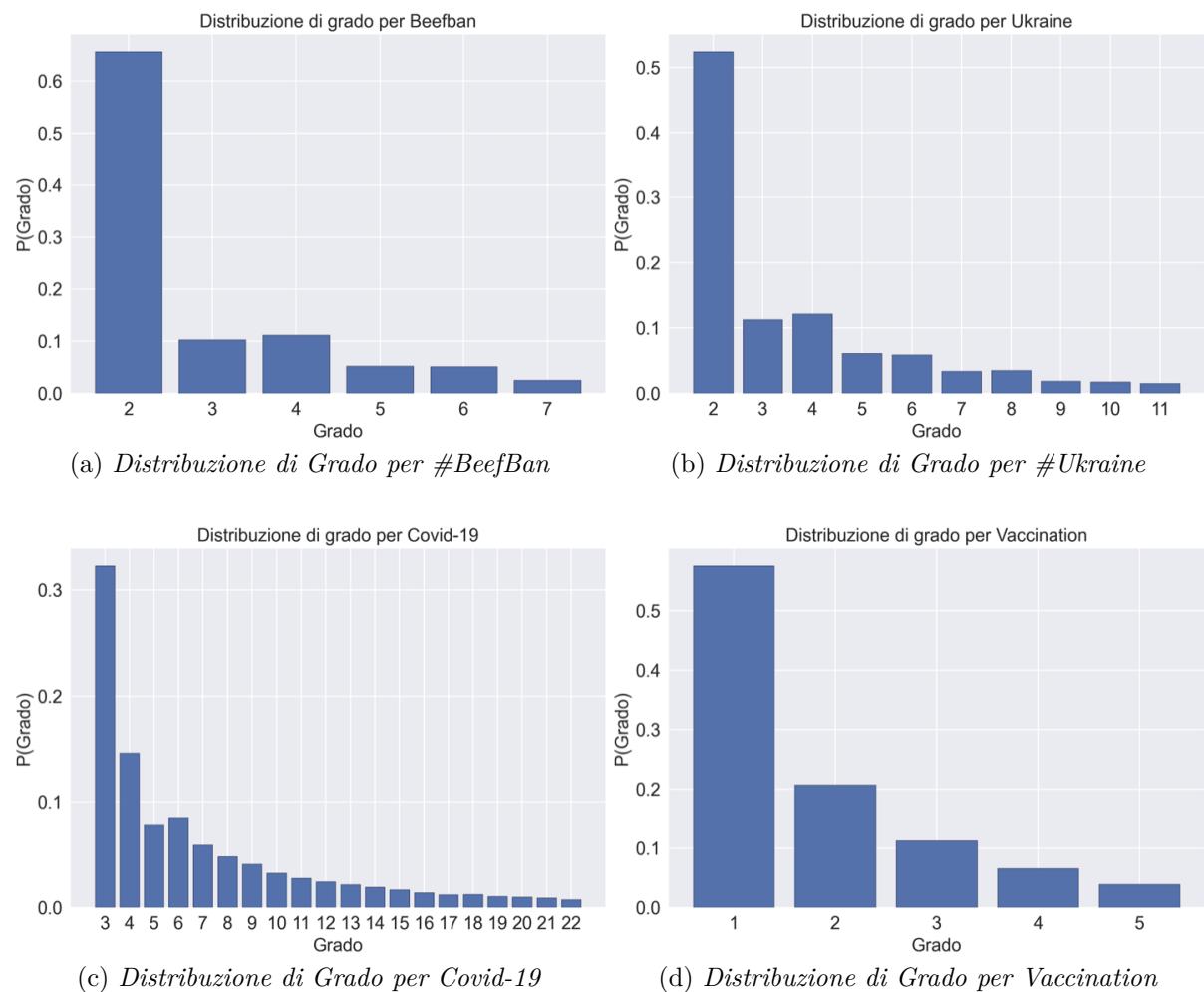


Figura 5.30: Distribuzione di grado per le reti dotate di Controversia

Si può subito notare come, a differenza di tutte le altre reti, *Vaccination* osservi quasi il 60% dei nodi dotati di un singolo collegamento. Questi valori fanno pensare a una struttura orientata a una serie di *ego-network* tra loro collegati. Tale differenza con gli altri grafi, i quali non vedono mai meno di due archi per nodo, permette di capire che la rete modellante l'argomento, potrebbe non essere sufficientemente grossa per il

lavoro effettuato. Tuttavia, i risultati ottenuti dalle analisi dei matadati *pre e post linking*, sembrano andare a confermare la presenza di controversia e la successiva riduzione della polarizzazione.

In ultimo è necessario sottolineare il fatto che, in tutto il lavoro di tesi, per l'attività di *Community detection* è sempre stato utilizzato un solo algoritmo. Tale scelta è giustificata dal fatto che, al fine di confrontarsi con lo stato dell'arte e fornire risultati comparabili a quanto già presente in letteratura, si è deciso di utilizzare lo stesso algoritmo proposto negli studi di Garimella et al. (2017).

Tuttavia questo potrebbe aver introdotto *bias* nella valutazione delle attività *pre e post linking*; altri algoritmi potrebbero identificare diverse (e magari anche migliori) partizioni; il tutto si riconduce però al quesito fondamentale alla base dell'attività *Community detection*: qual è il modo migliore e più puntale per identificare correttamente e completamente le *community* in una rete?

Capitolo 6

Conclusioni

In questo capitolo vengono discusse le fasi che hanno composto la *pipeline* di identificazione e riduzione della controversia, con l’obiettivo di superare le polarizzazioni; si analizzano, in particolare, le problematiche che ogni scelta può aver comportato.

In particolare, il capitolo risulta così suddiviso:

- Nella Sezione 6.1 si analizzano le scelte fatte in merito alla modellazione della rete sociale, sottolineando le assunzioni e le eventuali problematiche di queste. Viene poi analizzata la fase di partizionamento, con l’obiettivo di identificare le eventuali criticità.
- Nella Sezione 6.2 si passa alla valutazione delle attività di identificazione delle polarizzazioni, mediante l’utilizzo di metriche di *controversy*. Viene quindi ripreso il discorso circa la difficoltà nel modellare metriche volte a comprendere le distanze di visione e pensiero nelle comunità virtuali.
- Nella Sezione 6.3 viene discussa la soluzione adottata per la mitigazione del fenomeno; in particolare viene sottolineato come la natura teorica dell’attività di *Link prediction*, rappresenti un problema nella sua applicazione nell’ambito di reti complesse di interazioni. Vengono poi analizzati i limiti della “probabilità di accettazione” modellata mediante il *Sentiment Boost*.

6.1 Modellazione e partizionamento della rete

La modellazione della rete sociale è stato il primo passo nel lavoro proposto in questa tesi, se la definizione dei nodi e degli archi ha trovato ovvia e ragionevole modellazione, la scelta di lavorare poi su un grafo non orientato può essere una scelta contestabile.

La ragione del “compattamento” degli archi, che ha portato da una realtà orientata a una non orientata, è da ricercarsi in parte nel task di *Community detection* e in parte da un’assunzione effettuata: se un determinato utente cita un secondo e questo cita il primo, la modellazione orientata risulta essere superflua; un arco non orientato con la somma delle citazioni intercorse, può essere sufficiente a modellare la forza della relazione.

Situazione diversa per il caso in cui un utente citi molto un secondo e questi risponda in misura minore, o assente, al primo; in questo caso la modellazione mediante un “com-

pattamento” fa di fatto perdere la direzione dell’interazione e in parte la forza del legame.

Tuttavia, le relazioni *social* volte a identificare una realtà dove tanti utenti seguono un certo “*influencer*”, vedono tendenzialmente un interesse maggiore da parte dei tanti utenti (quindi dei *fan*) nei confronti del singolo nodo centrale (quindi *l'influencer*).

Modellare “compattando” e ponendo un singolo peso pari al numero di archi presenti, ottiene comunque il risultato di notificare un forte interesse tra i nodi, a patto di perdere, parzialmente, l’informazione di chi cita chi. Si è utilizzato il termine “parzialmente” in quanto, data una situazione dove un certo nodo *X* veda tanti collegamenti “forti” con altri nodi e questi altri nodi tendenzialmente presentino, tolto il nodo *X*, collegamenti non forti e poco numerosi, sarà possibile concludere che molto probabilmente è il nodo *X* ad essere molto citato.

In seguito alla modellazione della rete sociale si è effettuato il partizionamento della stessa, che costituisce forse l’attività più importante di tutta la pipeline definita, difatti il risultato di questo passaggio è andato inevitabilmente a condizionare tutto ciò che segue. Come anticipato precedentemente, questa attività ha condizionato la modellazione delle reti sociali, questo perché non è stato possibile identificare nessun algoritmo che tenesse in considerazione contemporaneamente, per calibrare il suo operato, il numero di partizionamenti richiesti, il senso delle relazioni esistenti e l’eventuale peso di queste.

Conseguenzialmente è stato scelto di utilizzare *METIS* in quanto, a patto di lavorare con grafi non orientati, permetteva di definire il numero di partizioni ricercate e di definire un attributo sul quale calibrare il proprio operato (peso dell’arco); inoltre è l’algoritmo utilizzato in diversi lavori allo stato dell’arte, sempre nel contesto di studio delle camere d’eco.

Come evidenziato più volte nel lavoro di Lancichinetti et al. (2010), l’attività di *Community detection* non è semplice, vi sono diversi algoritmi che operano in maniera diversa e effettuano supposizioni di base diverse.

L’utilizzo di metriche volte alla quantificazione della qualità delle partizioni identificate, può essere un buon espediente al fine di giustificare il partizionamento ottenuto da un determinato algoritmo; è stato tuttavia mostrato come, data la rete facente riferimento alle *vaccinazioni*, che non tutte le metriche prese singolarmente possono costituire una prova certa della qualità di un partizionamento: la *modularity* identificata in seguito alle partizioni sul già citato grafo, presentava valori molto bassi, gli stessi potevano far pensare che non vi fossero strutture significative all’interno della rete.

La successiva valutazione della *coverage*, congiunta alla conoscenza di dominio circa la possibilità di polarizzazioni dato il contesto, ha consentito di proseguire e portare alla luce un’effettiva presenza di idee contrastanti (vedi risultati *topic modelling* e *wordcloud* per *Vaccination* nel Capitolo 5).

Dunque l’attività di analisi dei metadati di cui si è in possesso, risulta essere fondamentale e di grande supporto per interpretare i risultati delle metriche presentate.

In questa fase poteva aver senso inserire un confronto circa i risultati di altri algoritmi, senza affidarsi unicamente ad un singolo; in questo modo si sarebbe potuto comprendere meglio quale suddivisione in *community* potesse essere quella ottimale. Questo costituisce effettivamente un aspetto della soluzione proposta che merita ulteriori indagini.

Ad ogni modo, i risultati ottenuti in questa fase sono buoni e denotano un'effettiva presenza di *community* polarizzate ed opposte; tali affermazioni vengono ulteriormente supportate, dove possibile, dallo studio dei metadati che è stato condotto.

6.2 Valutazione quantitativa della controversia

Dopo aver identificato e analizzato le due partizioni opposte, mediante anche una valutazione di diverse peculiarità delle partizioni identificate, il secondo *step* è stato quello di verificarne il livello di controversia.

La controversia si può intendere come una misura circa il grado di libertà di movimento all'interno di una rete sociale, dato un determinato argomento modellato mediante il grafo. Tuttavia questa è solo una delle possibili definizioni, difatti si può intendere anche come il livello di interconnessione tra i nodi di "frontiera" e i nodi "interni" o ancora come il grado medio di esposizione a utenti (quindi ai loro contenuti) di entrambe le partizioni.

Come è stato mostrato, sono state utilizzate diverse metriche al fine di coprire il più possibile ogni possibile modellazione del concetto di controversia della quale si è entrati in contatto; è stato poi ritenuto utile valutare il grado di polarizzazione, anche sull'unica rete priva di reali contrapposizioni.

Si è voluto definire un limite inferiore da porre, al fine di definire un argomento come "controverso" e al contempo si è definito un valore "obiettivo" per la fase successiva di superamento delle polarizzazioni.

Si è inoltre valutato come generalmente, a grandi valori di *modularity*, corrispondessero alti valori di *controversia* date tutte le metriche messe in campo. L'unica voce fuori dal coro è costituita dalla valutazione del grafo *Vaccination*, il quale presentava valori di *modularity* simili a quelli dell'unico grafo privo di camere d'eco (#Nationalkissingday), tuttavia al contempo registrava valori alti circa le metriche volte a quantificare la *controversia*.

Tale situazione suggerisce che può ed è necessario studiare meglio la relazione tra *controversia* e qualità delle partizioni (intesa come *modularity*).

Al fine di superare le problematiche circa l'eventuale insufficienza di una sola metrica per l'identificazione della *controversy*, si è deciso di lavorare in logica *compound*, andando a valutare tutti i valori delle metriche di quantificazione utilizzate.

Si ipotizza che quelle più orientate all'identificazione puntuale della polarizzazione, siano quelle basate sui *RandomWalk*, volti a simulare un movimento condotto dall'utente nel suo personale *feed* di notizie.

Tuttavia, un limite di tale modellazione è da ricercare nel fatto che ogni arco ha la stessa probabilità di essere scelto e quindi seguito; in accordo alla teoria di selezione selettiva formulata da Dieter (1986), tale scelta può essere troppo superficiale e può portare a risultati troppo ottimistici circa la controversia identificata.

Un ulteriore, e forse più marcato, limite riguarda la supposizione che un utente possa essere profondamente influenzato da un certo utente "autorevole" della partizione opposta (*RWC* e *RWTD*) o da una grande esposizione a contenuto e profili di utenti della partizione opposta (*CSC*).

La modellazione proposta per le metriche *RWC* e *RWTD* potrebbe essere poco accurata, questo perché non si tiene conto delle possibili credenze pre-esistenti dell’utente dal quale parte il *RandomWalk*; se a un utente profondamente convinto di una certa idea o teoria, viene fatto leggere un contenuto opposto alla stessa, prodotto da un utente “autorevole” (e.g. a un no-vax viene fatto leggere l’intervista di un pediatra) non è detto che il primo utente ne sarà influenzato.

Per quanto riguarda la metrica che modella il cambio di partizione (tale *CSC*), questa semplicemente conta quanti cambiamenti di partizione vengono effettuati durante un generico *RandomWalk*, supponendo che in un’azione controversa si tenda a rimanere in una partizione piuttosto che a “saltare” da una all’altra con grande facilità. Non viene tenuto però in considerazione, seppure si tratti di una situazione rara, il fatto che un *RandomWalk* partito da una partizione possa rimanere “incastrato” nella partizione opposta, restituendo così un valore quantificante controversia, quando di fatto non è vero.

Ad ogni modo i risultati ottenuti in questa fase sottolineano una differenza marcata tra reti dotate di polarizzazioni e non, andando a registrare divari importanti tra i valori delle varie metriche utilizzate.

6.3 Superamento polarizzazioni

L’attività finale effettuata nella pipeline proposta, prevedeva un’attività di riduzione della controversia (dunque della polarizzazione) mediante l’utilizzo di diverse metodologie di *Link prediction*; l’obiettivo fondamentale era quello di capire quale approccio, con l’aggiunta del minor numero di archi (valore stimato essere il 30% del numero di archi che la rete presentava) potesse permettere di superare la polarizzazione precedentemente quantificata.

Per questa attività, come mostrato, sono stati utilizzati approcci allo stato dell’arte (metodologie basate sulla prossimità) e approcci nuovi, proposti in questa tesi; quest’ultimi si basavano molto sulle peculiarità dei nodi in funzione della strutturazione della rete.

Stando ai risultati presentati nel Capitolo 5, la metodologia che più di tutte sembra ridurre e omogeneizzare il contenuto scambiato nella rete è quella basata sulla logica del *Top to Normal linking*; collegare utenti “comuni” di una partizione, con utenti “autorevoli” dell’altra sembra costituire il modo migliore per adempiere a questa attività.

Dato che la *Linking Prediction*, effettuata come descritto, fornisce un risultato strettamente teorico (non è detto che poi l’arco veda effettivamente luce), si è modellato una sorta di “probabilità di accettazione” con l’obiettivo di definire un moltiplicatore che potesse donare un livello di veridicità a un possibile arco. Siamo consci che questa modellazione si basi su un’assunzione che potrebbe rilevarsi insufficiente o limitata, ma ciò costituisce un primo tentativo di tenere in considerazione questa problematica.

È bene tuttavia sottolineare che basarsi solo sul *sentiment score* medio potrebbe essere una modellazione troppo superficiale: due persone appartenenti a due partiti politici opposti, potrebbero parlare bene e utilizzare vocaboli “positivi” per descrivere il proprio partito, provare a collegarli porterebbe facilmente a un rifiuto.

Un ulteriore punto, a cui forse non si è dedicata troppa discussione, riguarda la definizione del significato dietro al singolo collegamento posto; questo può essere concepito in due diverse maniere: consiglio di amicizia o sponsorizzazione. In particolare, la modellazione pensata e proposta in questo lavoro di tesi, dato il numero di archi aggiunto, è quello di sponsorizzazione: se viene aggiunto un arco da A a B con peso pari a x , con x media dei collegamenti tra utenti sulla totalità della rete, allora verranno fornite, ad A e B , x sponsorizzazioni circa i contenuti prodotti da, rispettivamente, A e B . In questo modo si cerca di esporre i due utenti a dell'informazione proveniente dall'altra partizione.

Per il valore di x si è deciso di utilizzare una media di tutti i collegamenti presenti nella rete (esclusi i *self-loop*). Si può pensare a logiche più “raffinate” come valutare la media dei collegamenti solo tra gli utenti delle due partizioni identificate o ancora valutare un $+/-\delta$ da applicare ad x .

L'attività di *linking* ha permesso di ottenere risultati apprezzabili in termini di superamento delle polarizzazioni precedentemente identificate; se non sempre le metriche di qualità delle *community* hanno potuto notificare la riduzione, analisi più dettagliate dei metadati in nostro possesso hanno confermato un effettivo superamento delle forti contrapposizioni.

Capitolo 7

Possibili sviluppi futuri

Consapevoli di essere autori dell'aggiunta di un singolo tassello al mosaico più ampio che costituisce il problema analizzato, viene fornita in questo capitolo una serie di possibili studi e approfondimenti che non si sono potuti affrontare in questo lavoro di tesi.

In particolare:

- Nella Sezione 7.1 vengono esposte le proposte circa valutazioni più raffinate in funzione del possibile partizionamento, proponendo poi un utilizzo più completo dell'informazione semantica e un'analisi più dettagliata delle informazioni topologiche.
- Nella Sezione 7.2 vengono introdotte possibili migliorie al *Sentiment Boost* proposto, viene inoltre suggerita una valutazione più attenta delle peculiarità personali, in funzione del materiale semantico prodotto.
- Nella Sezione 7.3 viene sottolineata l'importanza di modalità ibride volte ad assegnare lo *score* a un possibile arco, data l'attività di *Community detection*.
- Nella Sezione 7.4 si discute circa l'importanza di considerare i risultati ottenuti come base per ulteriori sperimentazioni e analisi sui risultati stessi.

7.1 Superamento del partizionamento bipartito, valutazione semantica e valutazioni di SNA

Gli attuali lavori di letteratura facenti riferimento al concetto delle camere d'eco, tendono a effettuare una partizione bipartita del grafo rappresentante la rete sociale. Tale decisione è giustificata dal fatto che, tendenzialmente, intorno a un determinato argomento controverso si sviluppano due pensieri opposti.

Questa supposizione può essere corretta nella maggior parte dei casi, tuttavia può costituire un limite troppo stringente in quanto obbliga a vedere tutto come "o bianco o nero". Tale criticità emerge soprattutto se si è intenzionati a effettuare uno studio più a grana fine, costituendo diversi livelli di pensiero intorno a una determinata questione. Potrebbe quindi essere sensato provare a superare la logica bipartita, andando a effettuare un partizionamento in un numero più alto di *community*, al fine di poter identificare diverse possibili posizioni su un determinato argomento.

In quest'ottica però risulterebbe fondamentale l'utilizzo puntuale e accurato della semantica presente nella rete, al fine di poter modellare in maniera più precisa le *community* intorno a una certa posizione. In questo senso, l'arricchimento degli archi proposto in questa tesi, potrebbe non essere sufficiente a tale scopo; sarebbe utile valutare modelli generativi che si basano su *LDA*, i quali dovrebbero tenere in considerazione anche la struttura topologica della rete, al fine di effettuare un'attività di *Community detection* che tenga in considerazione informazioni topologiche e semantiche, in funzione della rete sociale modellata.

Nel Capitolo 3 sono stati introdotti dei lavori, data la sezione facente riferimento alla *Community detection*, che vedevano l'utilizzo di modelli *LDA-oriented* al fine di ottenere il partizionamento del grafo.

In ultimo, nel Capitolo 2, sono state introdotte le analisi che appartengono alla *Social Network Analysis* che tengono in considerazione proprietà strutturali e si dividono in tre famiglie di sotto-analisi.

Seppur in questo lavoro di tesi si sono impiegate metriche facenti riferimento alla *SNA*, potrebbe essere interessante approfondire questa tipologia di analisi al fine di definire nuove misure volte a quantificare la qualità delle partizioni (valutazione *Coefficiente di Clustering* e *Coesione*).

Inoltre l'utilizzo più capillare della *SNA* apre a tutta una serie di nuove analisi volte a meglio identificare caratteristiche strutturali delle camere d'eco (presenza *buchi strutturali*, valutazioni circa *shortest path* e tipologie di *chiusura di rete*).

7.2 Valutazione tratti della personalità e modifica *Sentiment Boost*

Un ulteriore interessante analisi delle partizioni identificate, nell'ottica di arricchire ulteriormente la comprensione delle possibili differenze che possono esistere e che portano a delle polarizzazioni, potrebbe essere costituita da una valutazione dei tratti della personalità basandosi sul contenuto semantico condiviso.

Ponendo come base il lavoro proposto da Celli (2012), la valutazione della possibile personalità “dominante” in ogni partizione, può aiutare a comprendere meglio il contesto di studio.

Inoltre, mediante un'attività di *Sentiment analysis* e *Topic Modeling*, si rende possibile tutta una serie di analisi, a grana più fine, sulle possibili differenze di come gli utenti possano discutere e approcciarsi a un determinato argomento; il tutto con l'obiettivo ultimo di migliorare e ampliare la comprensione umana, dati i dinamismi e le attività sociali online effettuate dalle persone.

Per quanto riguarda invece il moltiplicatore che è stato proposto, tale *Sentiment Boost*, potrebbe essere interessante cercare di migliorarlo tenendo in considerazione più aspetti che potrebbero facilitare un'accettazione, da parte degli utenti, di un determinato collegamento.

In questo senso: valutazione circa i *topic* discussi, luogo geografico di appartenenza, estrazione sociale, studi condotti, età degli utenti potrebbero essere delle valide informazioni

sulle quali fare leva, al fine di migliorare la modellazione di una “probabilità di accettazione”. Si potrebbe anche pensare di utilizzare informazioni circa i tratti della personalità, ipotizzando quindi una modellazione in funzione della semantica di tale interessante aspetto.

Ovviamente risulta necessario essere in possesso di più metadati possibili, gli stessi non risultano però sempre accessibili con facilità.

7.3 Hybrid *Link prediction*

Le attività di *linking* hanno visto l'utilizzo di diverse metodologie, alcune orientate alla valutazione del vicinato dei potenziali nodi da collegare, altre più tendenti a valutazione in funzione delle peculiarità del nodo in termini di centralità.

Le diverse politiche presentate e utilizzate hanno viaggiato su binari paralleli, sarebbe interessante provare una combinazione delle stesse: valutazioni più ibride potrebbero ereditare i vantaggi delle singole metodologie e svilupparne di nuove.

In questo senso sarebbe utile modellare un indice che tenga in considerazione, anche proponendo pesi diversi, più peculiarità sia in riferimento a concetti basati sulla prossimità (vicini comuni, dimensione del vicinato dei vicini comuni) e sia su peculiarità intese come misure di centralità.

In particolare, come indici di centralità in questa tesi sono stati utilizzati solo la centralità di grado e la *Betweenness*. Il possibile utilizzo di altri indici potrebbe portare a nuovi e migliori risultati: utilizzare la *Closeness centrality* potrebbe essere interessante al fine di identificare i nodi più “vicini” a tutti, andando poi a collegare questi (un po' con l'approccio *TDL*).

A tale proposito anche combinare i valori di diversi indici di centralità, al fine di assegnare un *compound-score* a un potenziale arco, potrebbe costituire una possibilità percorribile al fine di ridurre, ancora più puntualmente e velocemente, le polarizzazioni identificate.

7.4 Approfondimento dei risultati ottenuti

Con l'obiettivo ultimo di comprendere sempre meglio la modalità migliore di *Link Prediction* da utilizzare per eliminare polarizzazioni, uno studio più approfondito circa i risultati ottenuti risulta essere necessario.

I risultati registrati dalla pipeline sono stati mostrati e commentati, i giudizi sugli stessi derivano però da un ragionamento basato sui valori ottenuti. Andare a studiare come, in funzione dei nodi congiunti, la riduzione della controversia possa cambiare, potrebbe essere fondamentale al fine di comprendere meglio il problema.

Si affianca poi la necessità di uno studio circa i rapporti esistenti tra la qualità, intesa come “forza”, delle partizioni identificate e il valore della *controversia*. Difatti è stato mostrato, sia nel contesto di *Sentiment Boost* che con la rete sociale delle *vaccinazioni*, che a una diminuzione di un certo livello della *controversia*, non sembra sempre equivalere

una diminuzione di *modularity* (metrica molto utilizzata nell’ambito di valutazione della qualità di una partizione).

Se per *Vaccination* si è fornita un possibile motivazione volta a giustificare situazioni, se vogliamo anche più “controverse”, dove si osservava addirittura un miglioramento delle strutture interne; per la gestione del *Sentiment Boost* invece non si è trovata una spiegazione sufficientemente supportata da evidenze.

Dunque, potrebbe essere utile comprendere e approfondire le differenze in termini di archi, valutando la presenza o assenza del moltiplicatore che modelli la “probabilità di accettazione”, andando quindi oltre alla sola valutazione della possibile differenza di riduzione di *controversia*.

Appendice A

Immagini Analisi

In questa sezione vengono fornite tutte quelle Figure, facenti riferimento al Capitolo 5, che per ragioni di sintesi non sono state inserite all'interno del capitolo stesso.

A.1 Grafi Garimella

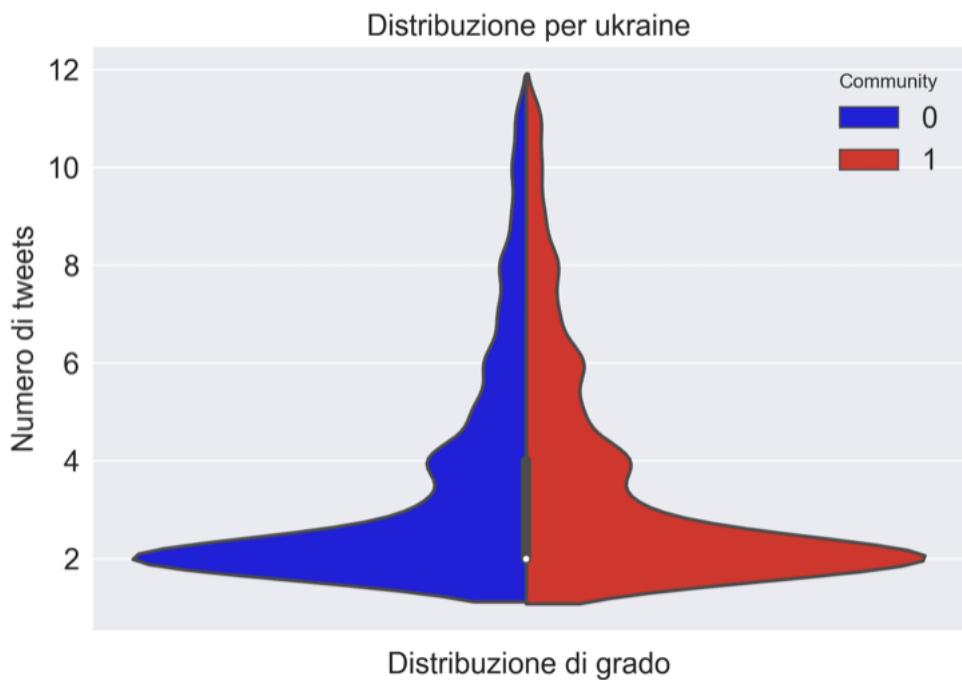


Figura A.1: Distribuzione di grado intra-community - #Ukraine

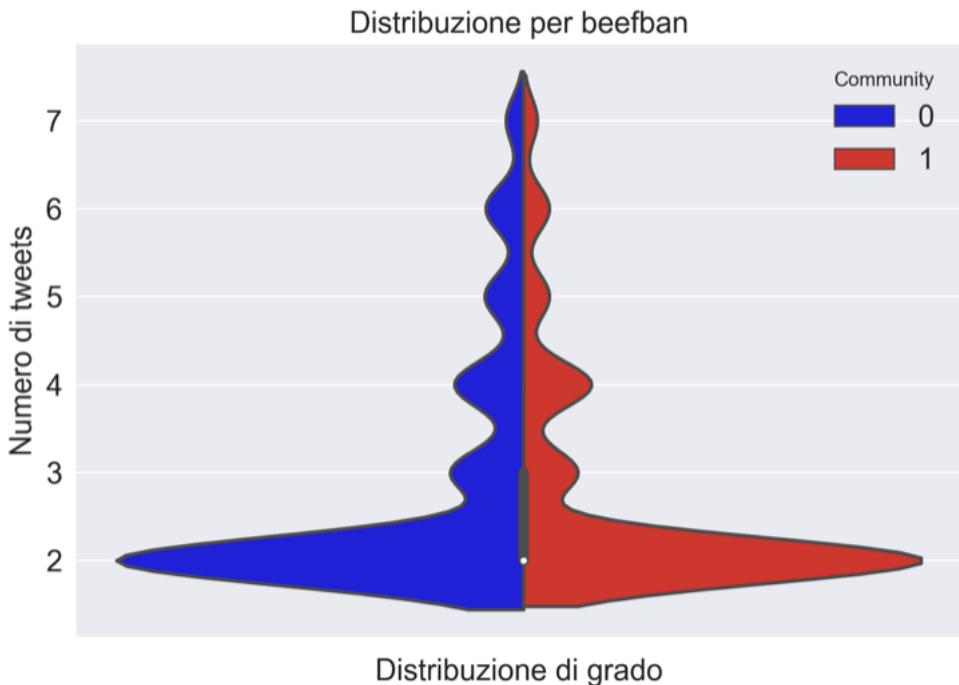


Figura A.2: Distribuzione di grado intra-community - #Beefban

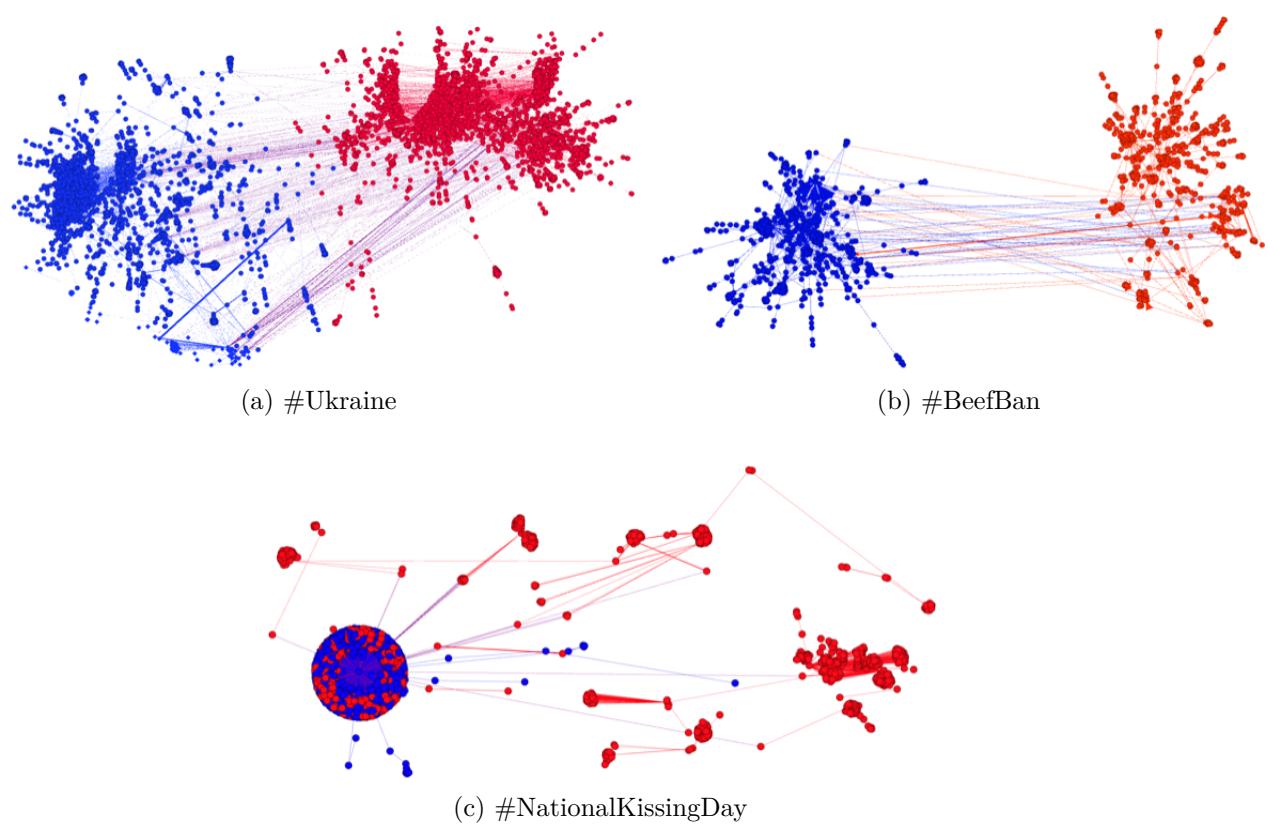


Figura A.3: Partizioni definite da METIS sui grafi di Garimella

A.2 Covid-19

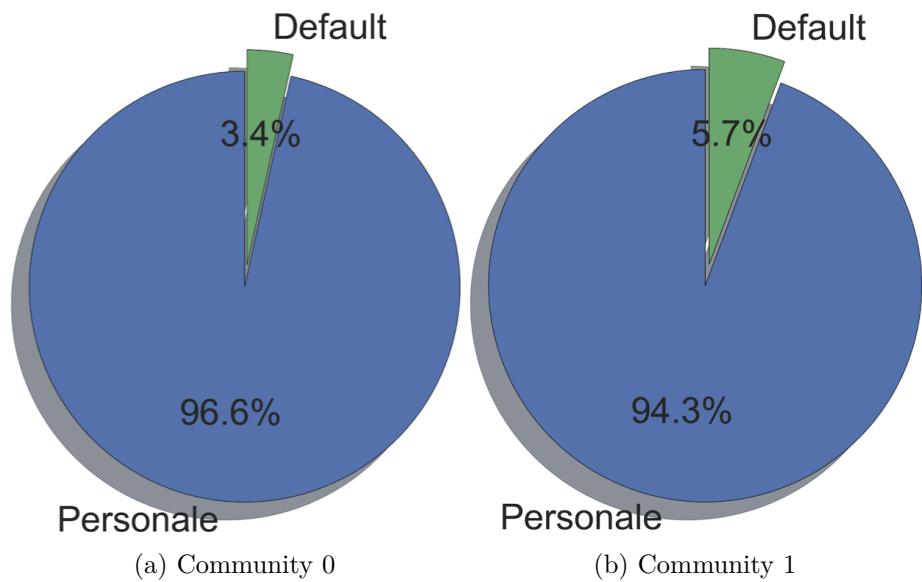


Figura A.4: Distribuzione utenti con immagine di default date le *community*

Appendice B

Immagini *Sentiment Boost*

Vengono qui mostrati tutti quei grafici che per questioni di importanza sono stati omessi nella Sezione 5.4.4 del Capitolo 5; gli stessi sono inseriti per completezza in questo appendice.

B.1 Covid-19

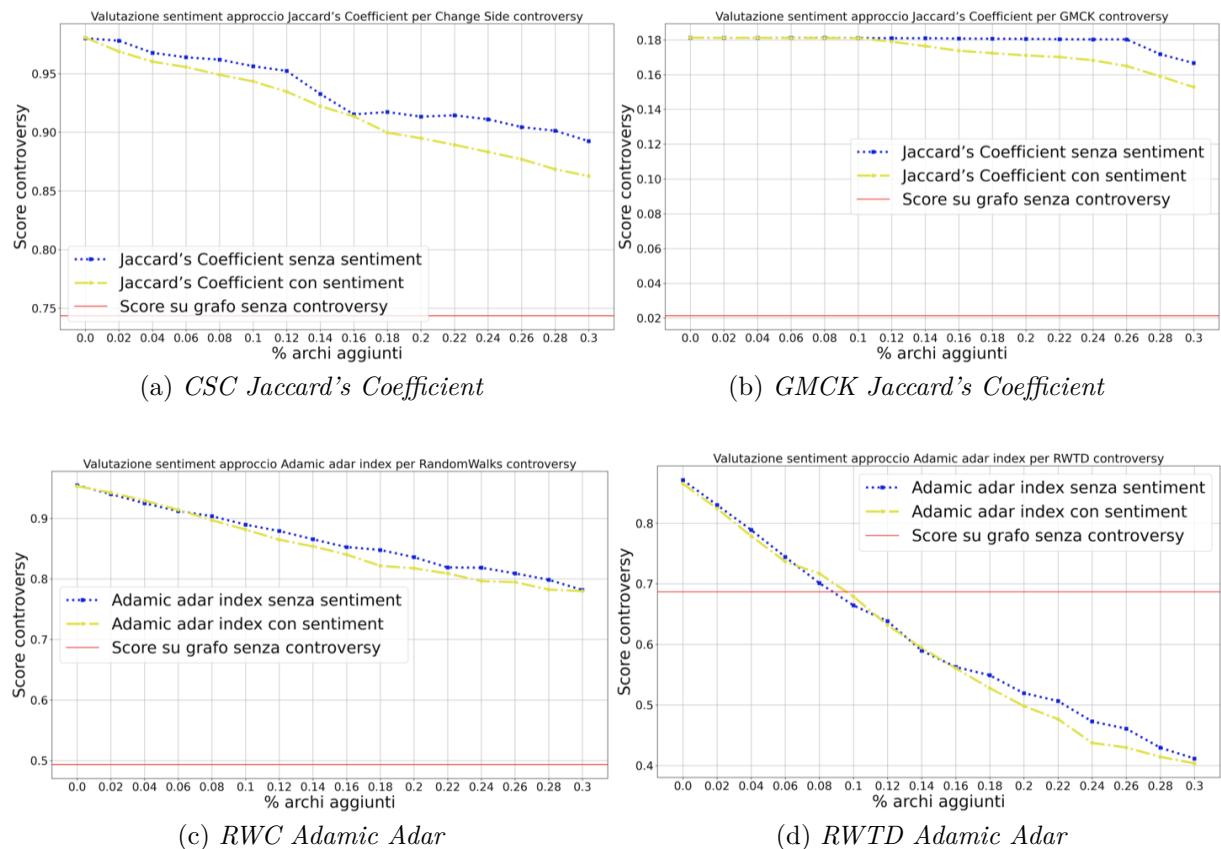
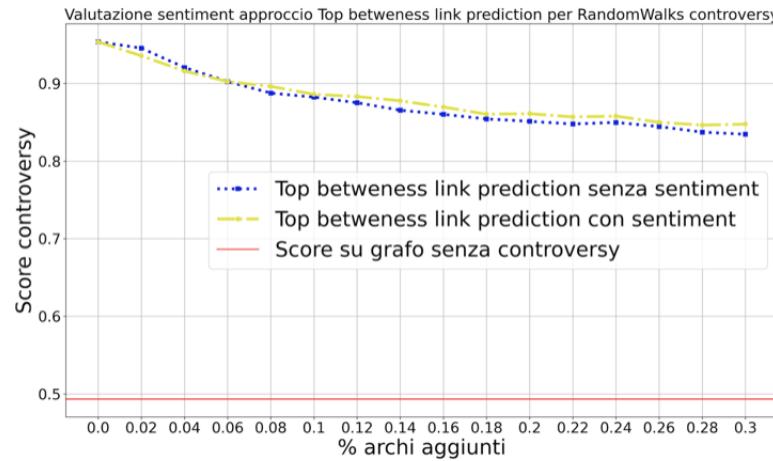
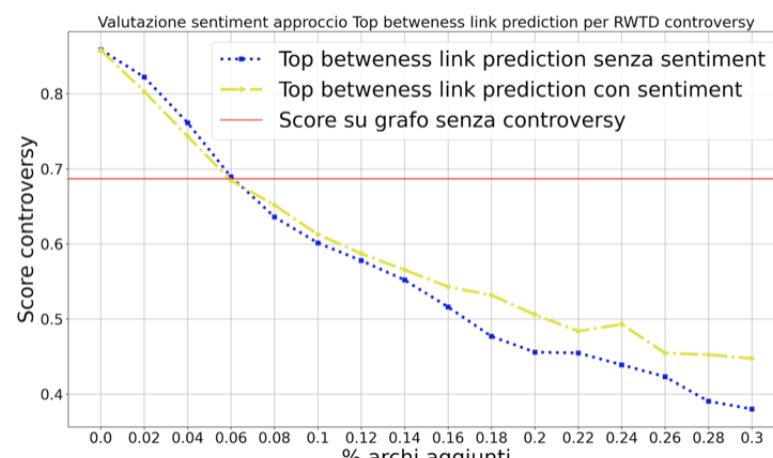


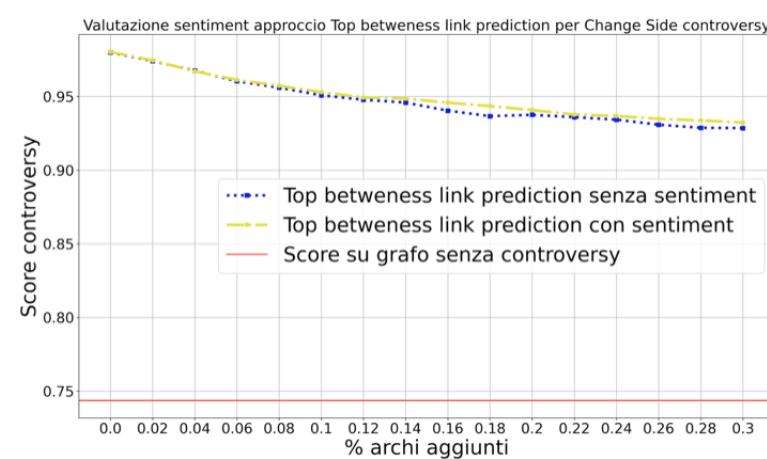
Figura B.1: Valutazione *Sentiment Boost* per *Jaccard's Coefficient* e *Adamic Adar* - Covid-19



(a) *RWC Top Betweenness Linking*



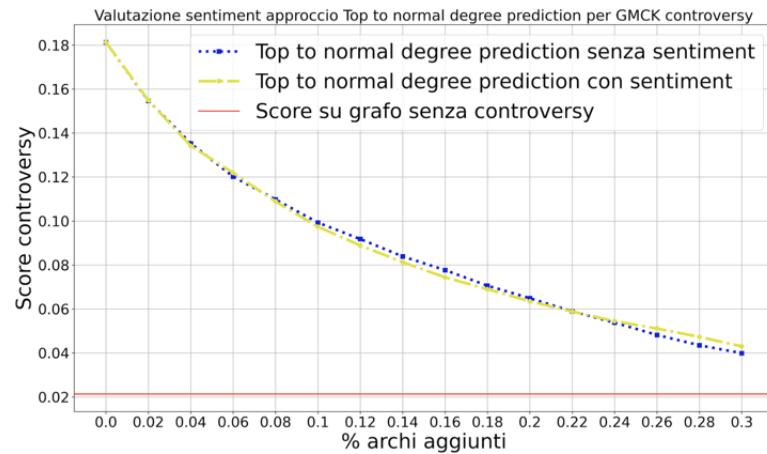
(b) *RWTD Top Betweenness Linking*



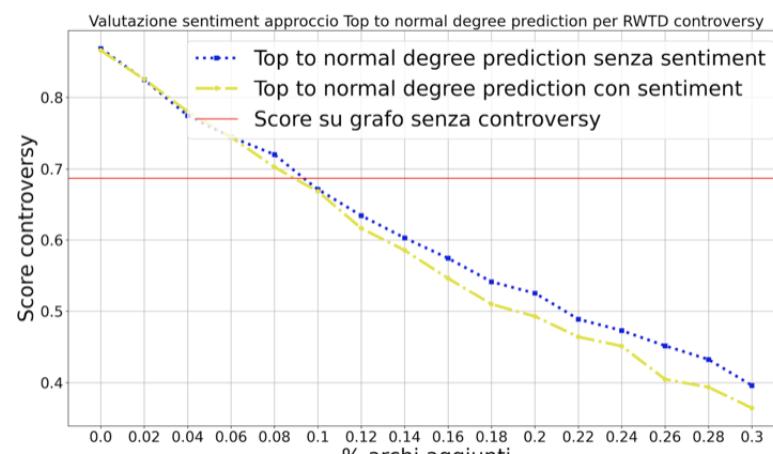
(c) *CSC Top Betweenness Linking*

Figura B.2: Valutazione *Sentiment Boost* per *Top Betweenness Linking* - Covid-19

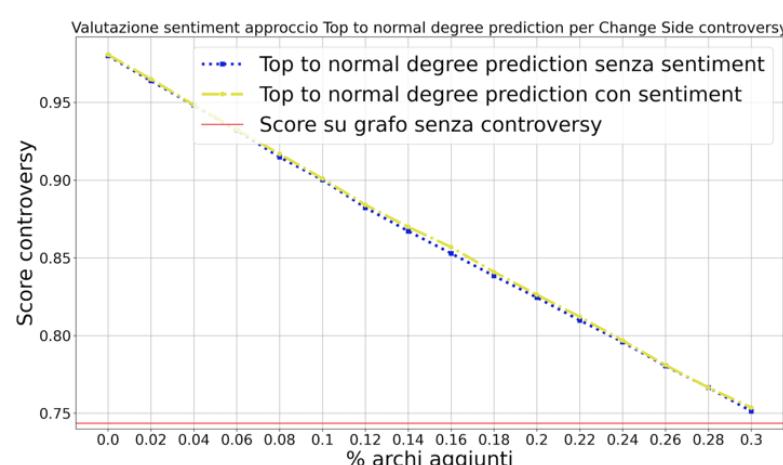
Appendice B. Immagini Sentiment Boost



(a) GMCK Top to Normal Linking



(b) RWTD Top to Normal Linking



(c) CSC Top to Normal Linking

Figura B.3: Valutazione *Sentiment Boost* per *Top to Normal Linking* - Covid-19

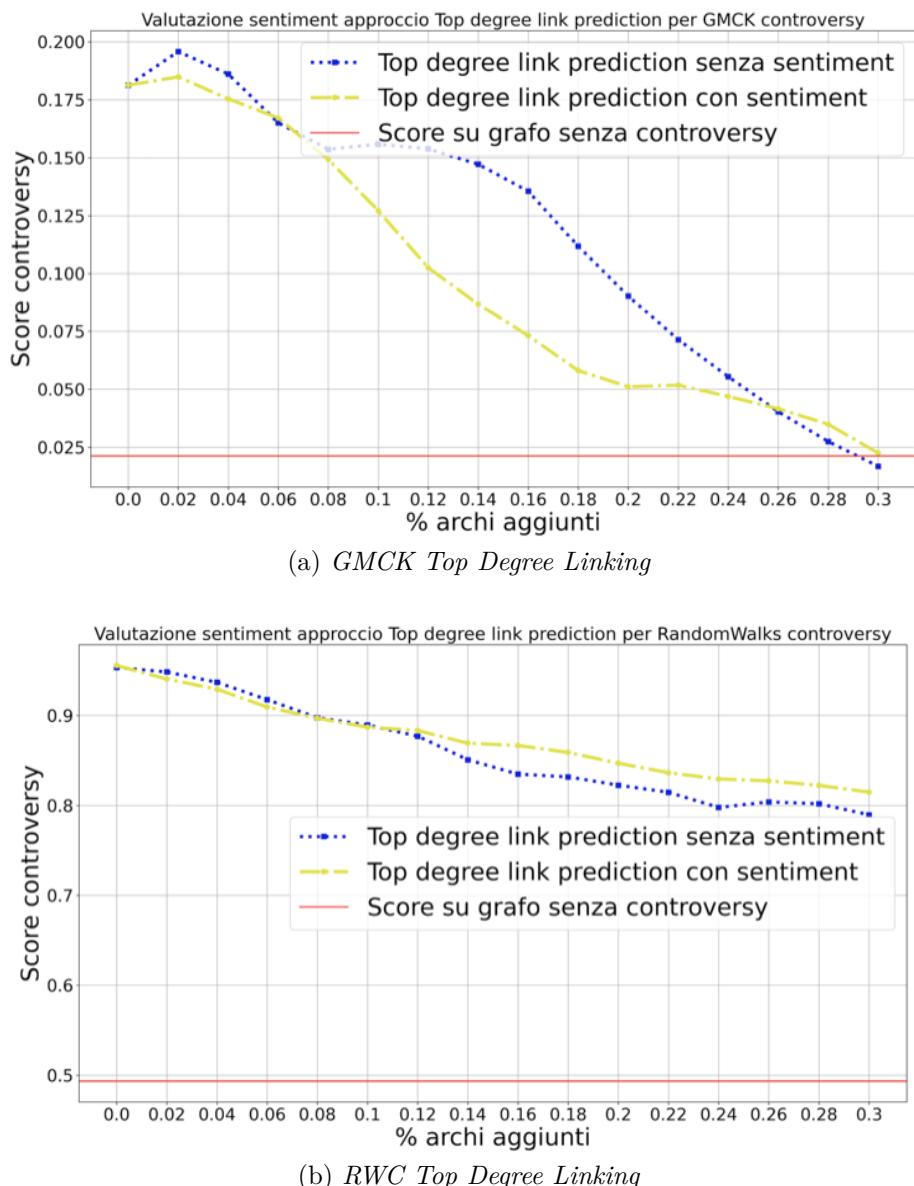


Figura B.4: Valutazione *Sentiment Boost* per *Top Degree Linking* - Covid-19

B.2 Vaccination

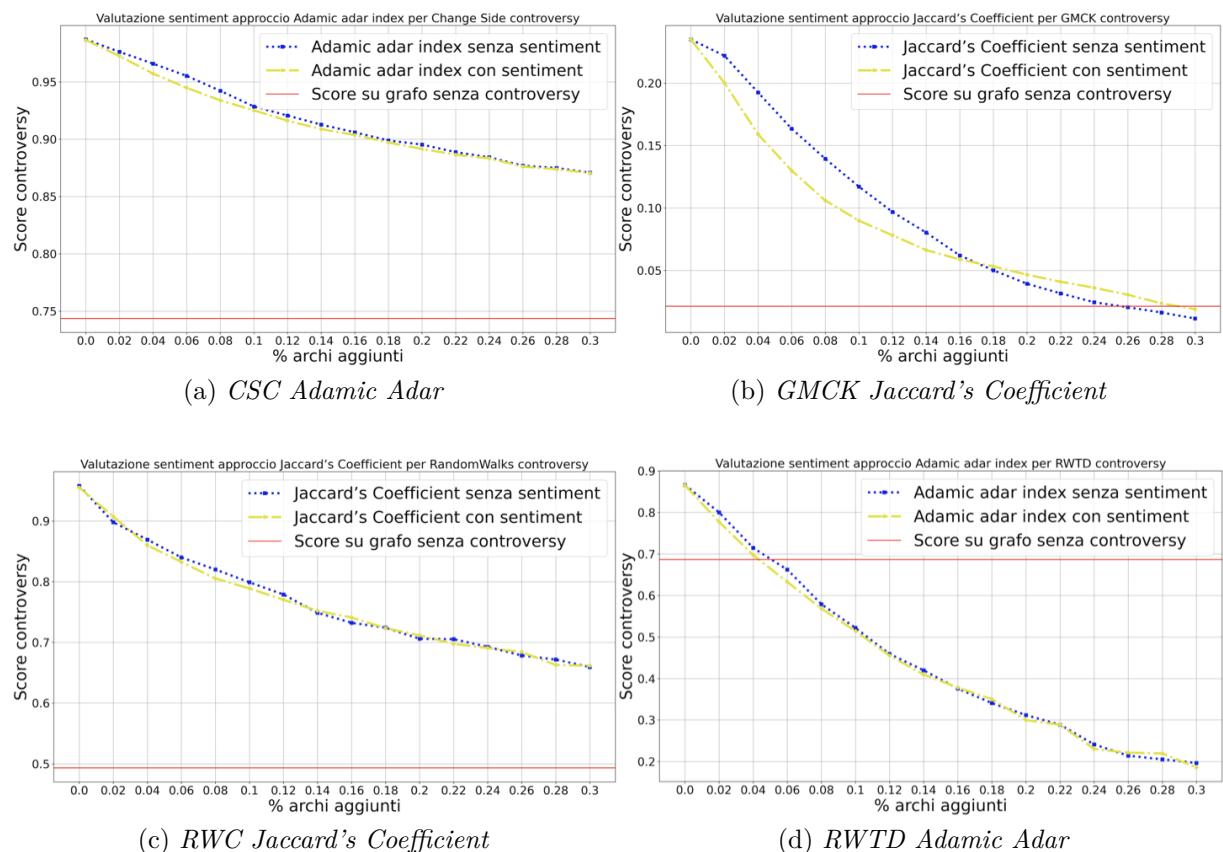
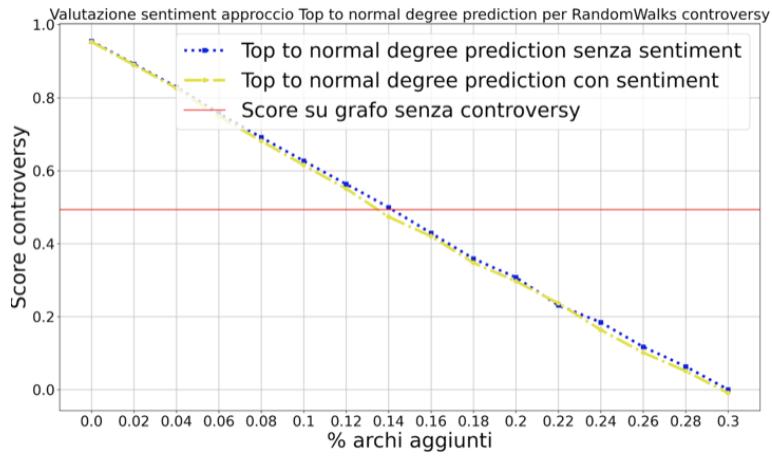
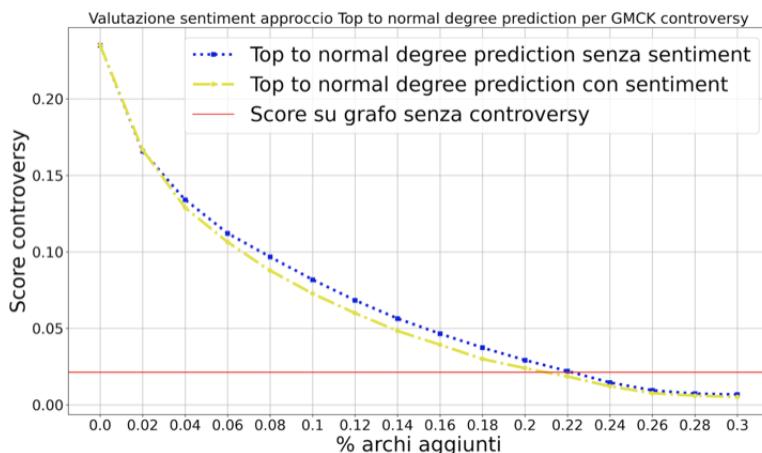


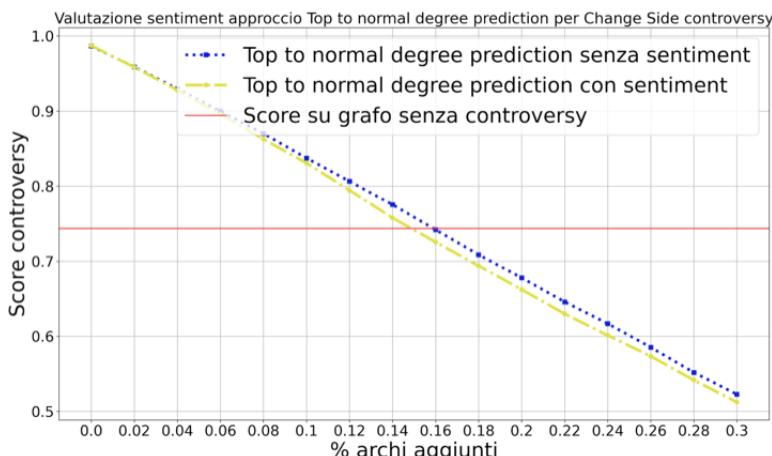
Figura B.5: Valutazione *Sentiment Boost* per *Jaccard's Coefficient* e *Adamic Adar* - Vaccination



(a) *RWC Top to Normal Linking*



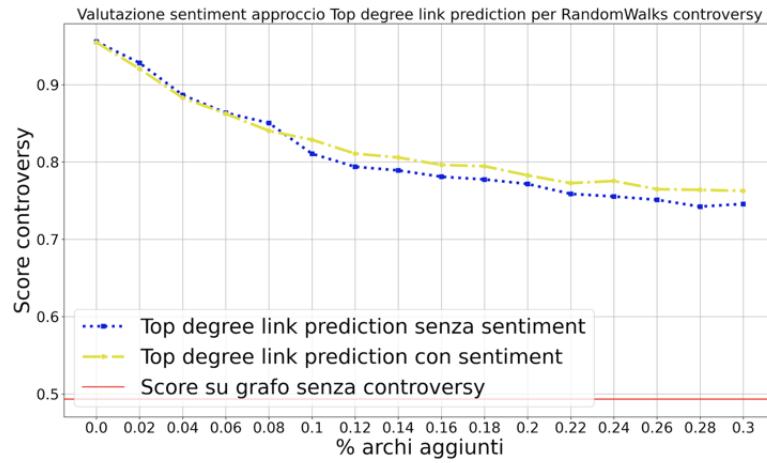
(b) *GMCK Top to Normal Linking*



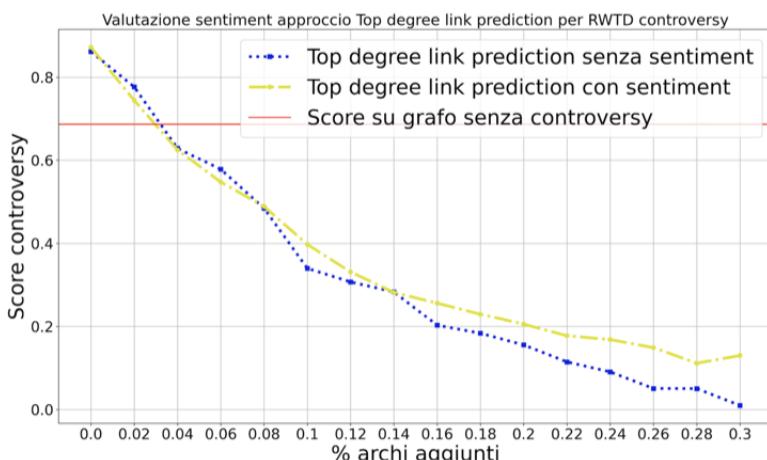
(c) *CSC Top to Normal Linking*

Figura B.6: Valutazione *Sentiment Boost* per *Top to Normal Linking* - Vaccination

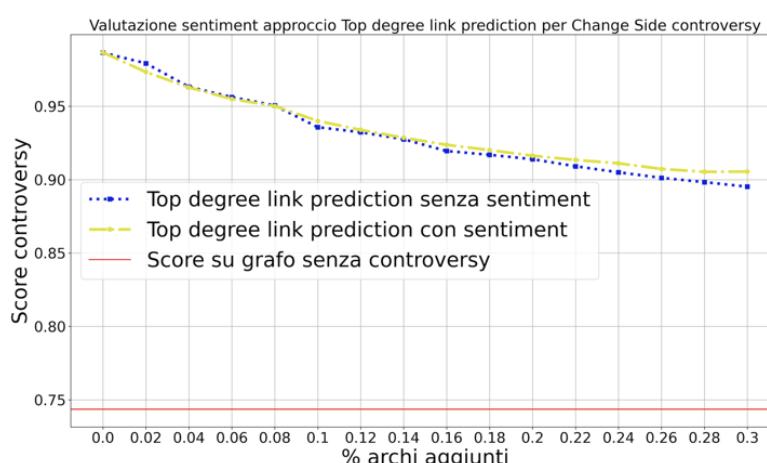
Appendice B. Immagini Sentiment Boost



(a) *RWC Top Degree Linking*

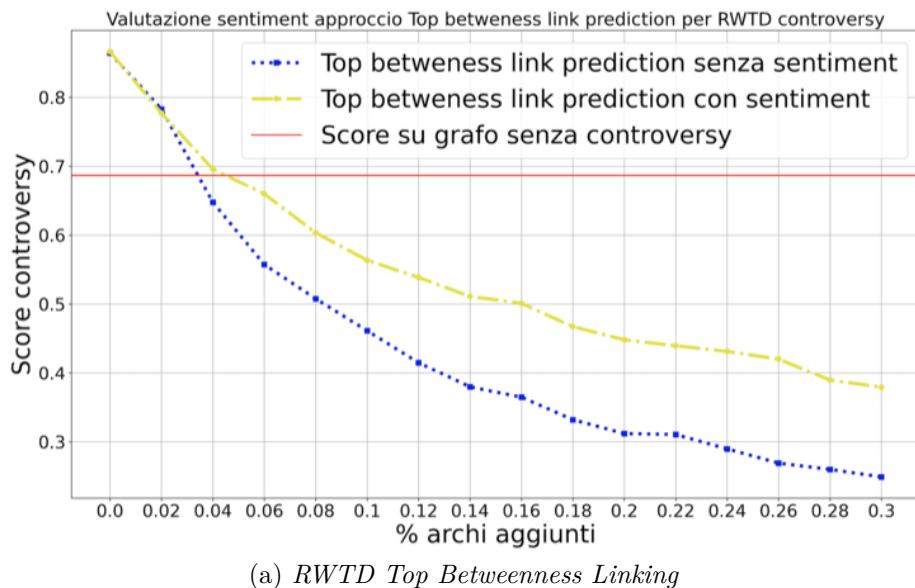


(b) *RWCTD Top Degree Linking*

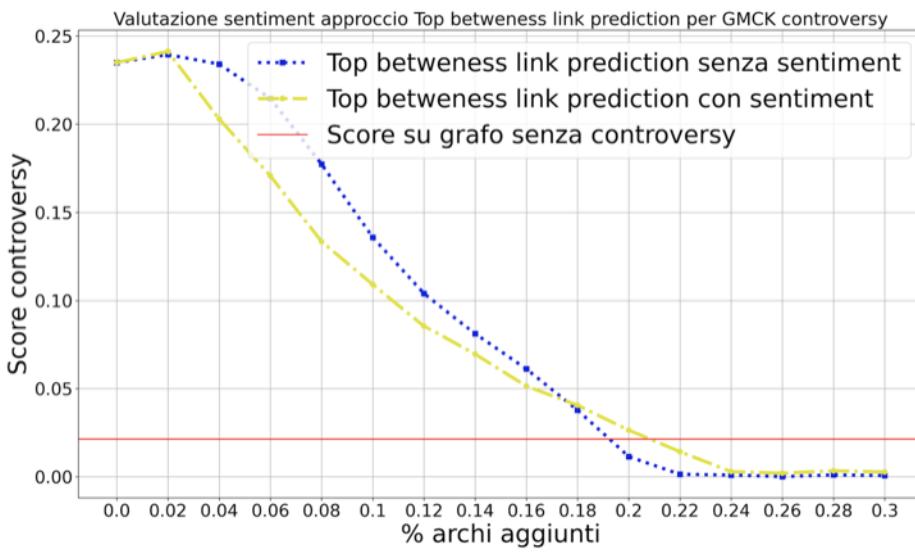


(c) *CSC Top Degree Linking*

Figura B.7: Valutazione *Sentiment Boost* per *Top Degree Linking* - Vaccination



(a) RWTD Top Betweenness Linking



(b) GMCK Top Betweenness Linking

Figura B.8: Valutazione *Sentiment Boost* per *Top Betweenness Linking* - Vaccination

Elenco delle figure

2.1	Evoluzione del Web ^[37]	6
2.2	Pipeline valutazione topic coherence proposta da Röder et all. (2015)	22
2.3	Esempio di segmentazione parziale dato il <i>topic</i> { <i>game</i> , <i>ball</i> , <i>sport</i> , <i>team</i> } Röder et all. (2015)	23
3.1	Diversi stage proposti da <i>Garimella et al.</i> (2017)	30
3.2	Esempio di dendogramma con taglio volto a identificare le due partizioni ottenute Fersini (2019).	34
3.3	Rappresentazione a piattini del modello proposto da Natarajan et al. (2013)	36
3.4	Gibbs sampling dato il modello Link-Content proposto da Natarajan et al. (2013)	36
3.5	Animazione del widget proposto da Munson et al. (2013)	39
3.6	Pseudocodice algoritmo di riduzione controversia proposto da Garimella et al. (2017)	42
4.1	Esempio modellazione su singola riga dei dati utilizzati da <i>Garimella et al.</i>	48
4.2	Modellazione in funzione delle citazioni trovate	50
4.3	Fasi dell'algoritmo METIS dato il <i>paper</i> di presentazione Karypis et al. (1995)	55
4.4	Esempio modellazione multigrafo data una singola riga dei dati utilizzati da <i>Garimella et al.</i> ipotizzando <i>weight</i> pari a 3.	56
4.5	Passaggio da grafo orientato a non orientato	57
4.6	Pipeline di assegnazione score <i>sentiment</i>	61
4.7	Rappresentazione modello <i>LDA</i> ^[12]	63
5.1	Suddivisione dei nodi date le <i>community</i> - Guerra del Donbass	76
5.2	Suddivisione dei nodi date le <i>community</i> - Divieto di vendita carne di mucca in India	77
5.3	Suddivisione dei nodi date le <i>community</i> - Giornata internazionale del bacio	78
5.4	Distribuzione sentimento intra-community dati i diversi pesi utilizzati . .	80
5.5	Contrapposizione tra le due <i>community</i> identificate con approccio basato sul <i>sentiment</i> - Covid-19	81
5.6	Contrapposizione tra le due <i>community</i> identificate con approccio basato sul <i>sentiment</i> - <i>Vaccination</i>	82
5.7	<i>Score</i> di <i>Topic Coherence</i> registrata per Covid-19	83
5.8	<i>Score</i> di <i>Topic Coherence</i> registrata per <i>Vaccination</i>	85

5.9	<i>Wordcloud</i> date le descrizione dei profili degli utenti in funzione della <i>community</i>	87
5.10	Percentuali di account verificati per ogni <i>community</i>	88
5.11	Numero di utenti con anno di creazione profilo nel 2020	88
5.12	Distribuzione <i>sentiment score</i> per utenti verificati	89
5.13	<i>Wordcloud</i> dati i <i>tweet</i> scambiati all'interno delle communities definite	89
5.14	Numero di utenti per <i>community</i> che hanno prodotto <i>tweet</i> riconducibili ad un <i>topic</i>	90
5.15	Descrizione profili utente no-vax presenti nella <i>community</i> 1 come nodi centrali, rispettivamente, per grado e per <i>Betweenness</i>	91
5.16	Riduzione controversia - <i>#BeefBan</i>	95
5.17	Riduzione controversia - <i>#Ukraine</i>	96
5.18	Riduzione controversia - <i>Covid-19</i>	98
5.19	Riduzione controversia - <i>Vaccination</i>	99
5.20	Valutazione <i>Sentiment Boost</i> per approcci basati sulla prossimità - Covid-19	100
5.21	Valutazione <i>Sentiment Boost</i> per approcci basati sulla centralità - Covid-19	101
5.22	Valutazione <i>Sentiment Boost</i> per approcci basati sulla prossimità - Vaccination	102
5.23	Valutazione <i>Sentiment Boost</i> per approcci basati sulla centralità - Vaccination	103
5.24	Valutazione <i>sentiment distribution</i> utenti verificati pre e post <i>Link prediction</i> - <i>Covid-19</i>	107
5.25	<i>Wordcloud</i> date le descrizione dei profili degli utenti in funzione della <i>community</i> post <i>Link prediction</i> - <i>Covid-19</i>	108
5.26	Percentuale account verificati in funzione della <i>community</i> post <i>Link prediction</i> - <i>Covid-19</i>	108
5.27	Distribuzione dello score medio di <i>Sentiment</i> delle communities identificate - <i>Covid-19</i>	109
5.28	Numero di utenti per comunità associati ai <i>topic</i> - <i>Vaccination</i>	111
5.29	<i>Wordcloud</i> tweet intra-community - <i>Vaccination</i>	112
5.30	Distribuzione di grado per le reti dotate di Controversia	113
A.1	Distribuzione di grado intra- <i>community</i> - <i>#Ukraine</i>	125
A.2	Distribuzione di grado intra- <i>community</i> - <i>#Beefban</i>	126
A.3	Partizioni definite da METIS sui grafi di Garimella	126
A.4	Distribuzione utenti con immagine di default date le <i>community</i>	127
B.1	Valutazione <i>Sentiment Boost</i> per <i>Jaccard's Coefficient</i> e <i>Adamic Adar</i> - Covid-19	129
B.2	Valutazione <i>Sentiment Boost</i> per <i>Top Betweenness Linking</i> - Covid-19	130
B.3	Valutazione <i>Sentiment Boost</i> per <i>Top to Normal Linking</i> - Covid-19	131
B.4	Valutazione <i>Sentiment Boost</i> per <i>Top Degree Linking</i> - Covid-19	132
B.5	Valutazione <i>Sentiment Boost</i> per <i>Jaccard's Coefficient</i> e <i>Adamic Adar</i> - Vaccination	133
B.6	Valutazione <i>Sentiment Boost</i> per <i>Top to Normal Linking</i> - Vaccination	134
B.7	Valutazione <i>Sentiment Boost</i> per <i>Top Degree Linking</i> - Vaccination	135
B.8	Valutazione <i>Sentiment Boost</i> per <i>Top Betweenness Linking</i> - Vaccination	136

Elenco delle tabelle

4.1	Informazioni sulle Reti Sociali dati i lavori condotti da <i>Garimella et al.</i> (2017)	48
4.2	Informazioni sulla Rete Sociale - Covid 19	51
4.3	Informazioni sulla Rete Sociale - <i>Vaccination</i>	52
4.4	Informazioni riguardanti la <i>giant component</i> del dataset riguardante Covid 19	58
4.5	Informazioni sui grafi ottenuti dal dataset sulle vaccinazioni	58
5.1	Risultati metriche in funzione dell'attività di <i>Community detection</i> - #ukraine	76
5.2	Risultati metriche in funzione dell'attività di <i>Community detection</i> - #beefban	77
5.3	Suddivisione nodi date le <i>community</i> - #nationalkissingday	78
5.4	Risultati metriche in funzione dell'attività di <i>Community detection</i> - Covid-19	79
5.5	Risultati <i>Community detection</i> con <i>METIS</i> - <i>Vaccination</i>	81
5.6	Top 10 <i>key words</i> modello <i>LDA</i> con <i>Topic Coherence</i> più alta per <i>Covid-19</i>	84
5.7	Top 10 <i>key word</i> modello <i>LDA</i> con <i>Topic Coherence</i> più alta per <i>Vaccination</i>	86
5.8	Risultati metriche quantificazione controversia - Garimella	92
5.9	Risultati metriche quantificazione controversia - Covid-19	93
5.10	Risultati metriche quantificazione controversia - Vaccination	93
5.11	Risultati <i>Community detection</i> post <i>Link prediction</i> - #BeefBan	104
5.12	Risultati <i>Community detection</i> post <i>Link prediction</i> - #Ukraine	105
5.13	Risultati <i>Community detection</i> post <i>Link prediction</i> - Covid-19	106
5.14	Risultati <i>Community detection</i> post <i>Link prediction</i> - Vaccination	109

Bibliografia

- [1] Santo Fortunato Andrea Lancichinetti. Community detection algorithms: a comparative analysis. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 80(5 Pt 2):056117, 2010.
- [2] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272(1-2):173–187, 1999.
- [3] Mathieu Bastian, Sébastien Heymann, and Mathieu Jacomy. Gephi: An open source software for exploring and manipulating networks, 2009. URL <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>.
- [4] Edward Loper Bird Steven and Ewan Klein. *Natural Language Processing with Python*. 2009.
- [5] Axel Bruns. Echo chamber? what echo chamber? reviewing the evidence. 2017.
- [6] F. Celli. Unsupervised personality recognition for social network sites. In *ICDS 2012*, 2012.
- [7] Aaron Clauset, Mark EJ Newman, and Christopher Moore. Finding community structure in very large networks. *Physical review E*, 70(6):066111, 2004.
- [8] Mauro Coletto, Venkata Rama Kiran Garimella, A. Gionis, and C. Lucchese. A motif-based approach for identifying controversy. In *ICWSM*, 2017.
- [9] Federico Comotto. *The Phenomenon of Echo Chambers in Social Media: Quantification and Reduction of Controversy*. Master’s thesis, Università degli studi di Milano-Bicocca, Dipartimento di Informatica, Sistemistica e Comunicazione Corso di Laurea Magistrale in Informatica, 2020.
- [10] Jon Kleinberg David Liben-Nowell. The link prediction problem for social networks. *Journal of the American society for information science and technology*, pages 1019–1031, 2004.
- [11] Michael I. Jordan David M. Blei, Andrew Y. Ng. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022, 2003.
- [12] Tyler Doll. Lda topic modeling: An explanation. *Towards data science*, 2018.

- [13] Siying Du and Steve Gregory. The echo chamber effect in twitter: does community polarization increase? In *International workshop on complex networks and their applications*, pages 373–378. Springer, 2016.
- [14] Morteza Zadimoghaddam Erik D. Demaine. *Minimizing the Diameter of a Network Using Shortcut Edges*. Springer, 2010. pages 420-431.
- [15] Elisabetta Fersini. Community detection. Lesson slides, 2019.
- [16] Santo Fortunato. Community detection in graphs. *ArXiv*, abs/0906.0612:75–174, 01 2010.
- [17] Michele Benzi Francesca Arrigo. Edge modification criteria for enhancing the communicability of digraphs. *SIAM J. Matrix Anal. Appl.*, 37:443–468, 2015.
- [18] Dieter Frey. *Recent research on selective exposure to information*. In *Advances in experimental social psychology*, volume 19. Elsevier, 2 edition, 1986. pages 41-80.
- [19] Kiran Garimella, Gianmarco De Francisci Morales, and Michael Mathioudakis Aristides Gionis. Quantifying controversy on social media. *ACM Transactions on Social Computing*, 1:1 – 27, 09 2017.
- [20] Venkata Rama Kiran Garimella, G. Morales, A. Gionis, and M. Mathioudakis. Reducing controversy by connecting opposing views. In *WSDM '17*, 2017.
- [21] Vipin Kumar George Karypis. Unstructured graph partitioning and sparse matrix ordering system. 08 1995.
- [22] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12):7821–7826, 2002.
- [23] Mark S Granovetter. The strength of weak ties. *American journal of sociology*, 78(6):1360–1380, 1973.
- [24] Pedro H. Calais Guerra, Wagner Meira Jr., Claire Cardie, and Robert Kleinberg. A measure of polarization on social media networksbased on community boundaries. *ICWSM*, 2013.
- [25] Hanghang Tong Hau Chan, Leman Akoglu. Manipulating directed networks for bettersynchronization. 2014.
- [26] C. Hutto and E. Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *ICWSM*, 2014.
- [27] Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PLoS ONE*, 9, 06 2014.
- [28] A. Kaplan and M. Haenlein. Users of the world, unite! the challenges and opportunities of social media. *Business Horizons*, 53:59–68, 2010.

- [29] Suvadip Paul Kartik Sawhney, Marcella Cindy Prasetio. Community detection using graph structure and semantic understanding of text. 2017.
- [30] Brian W Kernighan and Shen Lin. An efficient heuristic procedure for partitioning graphs. *The Bell system technical journal*, 49(2):291–307, 1970.
- [31] Jan H. Kietzmann, Kristopher Hermkens, I. McCarthy, and Bruno S. Silvestre. Social media? get serious! understanding the functional building blocks of social media. *Business Horizons*, 54:241–251, 2011.
- [32] Srijan Kumar, William L Hamilton, Jure Leskovec, and Dan Jurafsky. Community interaction and conflict on the web. In *Proceedings of the 2018 World Wide Web Conference*, pages 933–943, 2018.
- [33] EytanA dar Lada A. Adamic. Friends and neighbors on the web. *Social Networks*, pages 211–230, 2001.
- [34] Chei Sian Lee and Long Ma. News sharing in social media: The effect of gratifications and prior experience. *Computers in human behavior*, 28(2):331–339, 2012.
- [35] Qingzi Vera Liao and Wai-Tat Fu. Expert voices in echo chambers: effects of source expertise indicators on exposure to diverse opinions. In *CHI '14*, 2014.
- [36] M. Girvan M. E. J. Newman. Finding and evaluating community structure in networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 69 2 Pt 2:026113, 08 2004.
- [37] Vivek Madurai. Web evolution from 1.0 to 3.0. *Medium*, 2018.
- [38] Andrew Kachites McCallum. Mallet: A machine learning for language toolkit. <http://mallet.cs.umass.edu>, 2002.
- [39] Stanley Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.
- [40] Alfredo Jose Morales, Javier Borondo, Juan Carlos Losada, and Rosa M Benito. Measuring political polarization: Twitter shows the two sides of venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3):033114, 2015.
- [41] Vineet Chaoji Nagarajan Natarajan, Prithviraj Sen. Community detection in content-sharing social networks. *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)*, pages 82–89, 2013.
- [42] Efrat Nechushtai and Seth C Lewis. What kind of news gatekeepers do we want machines to be? filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, 90:298–307, 2019.
- [43] Gergely Palla, Imre Derényi, Illés Farkas, and Tamás Vicsek. Uncovering the overlapping community structure of complex networks in nature and society. *nature*, 435 (7043):814–818, 2005.

- [44] M. Papagelis, F. Bonchi, and A. Gionis. Suggesting ghost edges for a smaller world. In *CIKM '11*, 2011.
- [45] Ferran Parés, Dario Garcia Gasulla, Armand Vilalta, Jonatan Moreno, Eduard Ayguadé, Jesús Labarta, Ulises Cortés, and Toyotaro Suzumura. Fluid communities: A competitive, scalable and diverse community detection algorithm. In *International Conference on Complex Networks and their Applications*, pages 229–240. Springer, 2017.
- [46] Nikos Parotsidis, E. Pitoura, and Panayiotis Tsaparas. Selecting shortcuts for a smaller world. In *SDM*, 2015.
- [47] Nikos Parotsidis, E. Pitoura, and Panayiotis Tsaparas. Centrality-aware link recommendations. In *WSDM '16*, 2016.
- [48] Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014.
- [49] M. Porter. An algorithm for suffix stripping. *Program*, 14:130–137, 1980.
- [50] Radim Řehůřek and Petr Sojka. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta, May 2010. ELRA. <http://is.muni.cz/publication/884893/en>.
- [51] M. Röder, A. Both, and A. Hinneburg. Exploring the space of topic coherence measures. In *WSDM '15*, 2015.
- [52] Kazutoshi Sasaharaa, Wen Chen, Hao Peng, Giovanni Luca Ciampaglia, Alessandro Flammini, and Filippo Menczer. On the inevitability of online echo chambers. *ArXiv*, abs/1905.03919, 2019.
- [53] Christian Schulz. *Graph Partitioning and Graph Clusteringin Theory and Practice*. 2016.
- [54] Paul Resnick Sean A. Munson, Stephanie Y. Lee. Encouraging reading of diverse political viewpoints with a browser widget. *ICWSM*, 2013.
- [55] Ziyu Guan Senni Perumal, Prithwish Basu. Minimizing eccentricity in composite networks via constrained edge additions. *MILCOM 2013 - 2013 IEEE Military Communications Conference*, pages 1894–1899, 2013.
- [56] J. Short, Ederyn Williams, and B. Christie. The social psychology of telecommunications. 1976.
- [57] Yla R Tausczik and James W Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54, 2010.

- [58] UNISA. Communities. Lesson slides. <http://www.di-srv.unisa.it/professori/lg/RS/SN-communities.pdf>.
- [59] Marco Viviani. Materiale social media analytics (sma). Lesson slides, 2019.
- [60] Marco Viviani and Gabriella Pasi. Credibility in social media: opinions, news, and health information—a survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(5):e1209, 2017.
- [61] Duncan J Watts and Steven H Strogatz. Collective dynamics of ‘small-world’ networks. *nature*, 393(6684):440–442, 1998.
- [62] Helena Webb, Marina Jirotka, Bernd Stahl, William Housley, Adam Edwards, Matthew Williams, Rob Procter, Omer Rana, and Pete Burnap. Digital wildfires: hyper-connectivity, havoc and a global ethos to govern social media. *ACM SIGCAS Computers and Society*, pages 193–201, 01 2016.
- [63] Ilker Yildirim. Bayesian inference: Gibbs sampling. 2012. *Technical Note, University of Rochester*.
- [64] An Zeng, Linyuan Lü, and Tao Zhou. Manipulating directed networks for better synchronization. 2012.