

Gianluigi Greco\*

\* Direttore dipartimento di Matematica e Informatica (DeMaCS)  
Università della Calabria



## Dinamiche evolutive e rischi nelle reti sociali visti dall'osservatorio di un informatico

# Illusione della maggioranza, diffusione delle opinioni

Lo sviluppo di algoritmi e metodologie capaci di monitorare utilizzi maliziosi dei sistemi di networking

Comprendere come un comportamento collettivo possa formarsi ed emergere attraverso le interazioni che occorrono tra singoli individui è un argomento tradizionalmente studiato nell'ambito dell'economia, della psicologia sociale e delle scienze politiche. Più recentemente, a causa della crescente e oramai capillare diffusione delle piattaforme di social networking, quali Facebook e Twitter, tali problematiche hanno catturato l'attenzione anche degli informatici. Molte conferenze e riviste internazionali del settore, infatti, promuovono spazi di confronto e di analisi dedicate sia alle dinamiche evolutive delle reti sociali, sia ai rischi che possono annidarsi negli ecosistemi tecno-sociali. Parallelamente, sul fronte più strettamente tecnologico, sta diventando sempre più attuale la sfida per sviluppare infrastrutture di elaborazione in grado di analizzare, in tempo reale, l'enorme mole di dati che su tali reti ed ecosistemi viene generata.

Accomunato in questa sorte a tanti miei colleghi, anch'io negli ultimi anni ho subito il fascino di queste interessanti tematiche, che ho iniziato ad affrontare – in una prospettiva interdisciplinare – mosso dal desiderio di comprendere se tecniche e metodologie proprie del mio campo di ricerca, cioè l'Intelligenza Artificiale, possano essere proficuamente utilizzate per ragionare sui meccanismi di formazione e diffusione delle opinioni nelle reti sociali.

Nell'avvicinarmi a questo mondo, e con tale prospettiva, sono risultati certamente decisivi numerosi spunti di riflessione, strettamente legati all'attualità, che palesavano quanto fossero necessari studi sistematici e approfonditi sulle dinamiche di diffusione virale delle cosiddette fake news e quanto fosse urgente l'implementazione e l'utilizzo su larga scala di algoritmi per la loro identificazione automatica. Ancor più decisive sono state, tuttavia, le letture di alcuni lavori scientifici che evidenziano interessanti fenomeni, in qualche modo paradossali, che prendono forma quando diversi individui scambiano le proprie opinioni all'interno di un contesto sociale.

Tra tutti questi lavori, mi sento di suggerire anche a un pubblico di non addetti uno studio<sup>1</sup> di Lerman, Yan e Wu, dall'evocativo e suggestivo titolo "The Majority Illusion in Social Networks". Nel loro interessante lavoro, gli autori presentano all'attenzione del lettore un paradosso di natura matematica, chiamato appunto "l'illusione della maggioranza", che illustra gli effetti sorprendentemente distorsivi che possono ingenerarsi nelle reti sociali. Spiegherò il paradosso con un esempio, attraverso il quale proverò a trarre alcuni utili spunti per analizzare il complesso tema delle reti sociali viste dalla prospettiva di un informatico e per discutere anche di alcuni studi condotti in questo ambito presso

il dipartimento di Matematica e Informatica dell'Università della Calabria.

### Dove andiamo a cena stasera?

Supponiamo che alcuni membri di un dipartimento debbano organizzare una cena per una ricorrenza e che stiano valutando se prenotare in un ristorante oppure in una pizzeria. Voglio essere più analitico. I membri del dipartimento sono sette e si chiamano A, B, C, D, E, F e G. Ognuno di loro ha degli amici (ad esempio, A è amico di C e D, ma non di B, di E, di F e di G) e ci è dato inoltre sapere che per ognuno di essi, tranne che per G, la maggioranza dei suoi amici ha una preferenza per la pizzeria.

In altri termini: la maggioranza degli amici di A amici vuole andare in pizzeria; la maggioranza degli amici di B vuole andare in pizzeria; la maggioranza degli amici di C vuole andare in pizzeria; ...e via dicendo. Solo G fa eccezione: la maggioranza dei suoi amici, infatti, preferirebbe andare in un ristorante.

Questo è, dunque, lo scenario che propongo come esempio e attraverso il quale invito a ragionare intorno a una semplice questione: mettendo ai voti tra A, B, C, D, E, F e G se si debba andare in pizzeria o al ristorante, quale sarà l'esito della votazione?

Devo ammettere che, prima di immergermi nello studio delle dinamiche che regolano le reti sociali, sarei stato incline a ritenere che il gruppo di membri del dipartimento alla fine avrebbe organizzato la cena in pizzeria. Infatti, la constatazione che su 7 membri ben 6 siano concordi nel dire che la maggioranza dei loro amici vuole andare in pizzeria sembrerebbe implicare logicamente che sia questa l'opinione più diffusa all'interno della rete sociale, e pertanto essa risulterà vincitrice nella votazione. Di più, si potrebbe addirittura pensare che la stessa deduzione logica resti ancora valida pur se solo 4 membri su 7 avessero indicato che la maggioranza dei loro amici sono inclini a preferire la pizzeria. E, tuttavia, l'apparenza inganna... Questi ragionamenti, infatti, sono del tutto errati!

Per comprendere perché la realtà possa essere completamente differente dalla percezione "sociale" che la rete delle relazioni ci manifesta, consideriamo la Figura 1 nella quale vengono rappresentati i sette membri A, B, C, D, E, F e G con le loro opinioni. Nella figura, un collegamento tra due membri rappresenta una relazione (simmetrica!) di "amicizia". Pertanto, A è amico di C e D; B è amico di G; C è amico di A, D e G; D è amico di A, C e G; E è amico di G; F è amico di G; e, infine, G ha come amici B, C, D, E e F. Dall'analisi della figura si nota che per ogni individuo, tranne che per G, la maggioranza degli amici ha una preferenza per la pizzeria. D'altra parte, emerge chiaramente che quattro membri su sette preferiscono il ristorante e, pertanto, la scelta del ristorante sarà l'esito finale della votazione! E tuttavia, fintantoché non saranno conteggiati i voti, ogni membro del dipartimento,

1 <https://doi.org/10.1371/journal.pone.0147617>

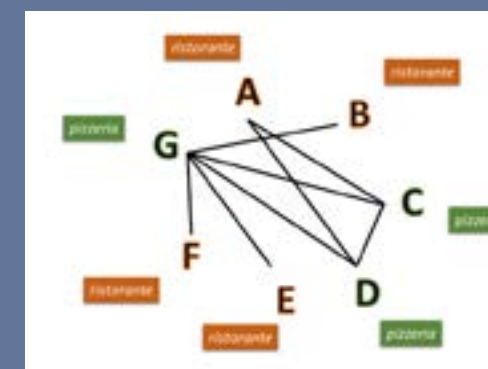


Figura 1: Un esempio dell'illusione della maggioranza.



facebook

Login

Sign Up for Facebook



fatto salvo G, sarà incline a pensare che l'esito della votazione porterà ad andare in pizzeria. E anche noi, che osserviamo dall'esterno quanto accade, fintantoché non analizzeremo per bene il quadro complessivo e smetteremo di dare ascolto ai "rumors" della rete sociale, saremo inclini a pensare – erroneamente – che la preferenza per la pizzeria sia l'opinione dominante. La riflessione su paradossi come quello che ho appena esemplificato – dicevo nell'introduzione – ha acceso il mio interesse per questo affascinante ambito di analisi e di ricerca. In una rete sociale, ogni individuo ha una visione molto parziale dell'ambiente nel quale è immerso e l'aggregazione di queste visioni parziali può portare a dipingere un quadro incompleto e non affidabile. Comprendere le dinamiche delle interazioni sociali pone dunque ai ricercatori molte sfide, sia per quanto attiene alla modellazione di tali dinamiche sia (e forse ancor di più) per quanto riguarda la capacità di predire la formazione e l'evoluzione stessa delle opinioni. Molte sono le strade che la ricerca deve ancora esplorare, ma negli anni sono stati raggiunti importanti risultati, alcuni dei quali, senza pretesa di completezza, proverò ora a passare in rassegna.

#### Dalle illusioni alla diffusione delle informazioni

A tutti è capitato di sperimentare nei rapporti sociali che l'interazione con altri individui, ad esempio con i nostri amici, possa determinare forme coscienti o inconsapevoli di pressione, capaci nel tempo di indurci a modificare le nostre opinioni. In effetti, molte azioni quotidiane risentono di una innata tendenza ad adattare (talvolta a conformare) comportamenti e opinioni a quelli dei nostri vicini e amici più prossimi. Ritornando alla rete sociale che abbiamo descritto nella figura 1, ci si potrebbe chiedere ad esempio se il fatto che gli amici di A (cioè C e D) desiderino andare in pizzeria abbia un qualche effetto sull'opinione che A stesso andrà ad esprimere quando prima o poi sarà chiamato a decidere sul da farsi. Magari A è un tipo risoluto per il quale l'opinione dei propri amici non ha alcuna influenza, ma molto più probabilmente A potrebbe risentire di una forma di "pressione sociale" che lo porterà ad adattare la sua opinione a quella di C e D. Dunque, l'opinione "pizzeria" potrebbe diffondersi da C e D influenzando anche A. Tale processo potrebbe poi continuare ed interessare altri individui, con un effetto di propagazione virale, detto anche a cascata.

Un esempio di una diffusione a cascata è riportato nella figura 2. Lo scenario presenta ancora gli individui A, B, C, D, E, F e G. Rispetto alla figura 1, tuttavia, le relazioni sociali e le opinioni di tali individui sono differenti. All'inizio, come evidenziato nel riquadro nella parte alta a sinistra, solo A e B desiderano andare al ristorante (colore arancio). L'individuo C, che vorrebbe andare in pizzeria (colore verde), inizia a risentire della pressione di A e B e, nel tempo, cambia opinione. Lo stesso accade poi, in successione, a D, E, F e infine G. Al termine di questo processo di diffusione, tutti gli individui avranno cambiato la propria opinione e si sarà raggiunta l'unanimità sull'opinione di andare al ristorante, opinione invece chiaramente minoritaria nella configurazione iniziale. In astratto, è accaduto che le interazioni sociali sono riuscite a ribaltare completamente l'esito di una eventuale votazione.

Fenomeni di "contagio sociale", come quello appena esemplificato, sono stati presi in considerazione sin dagli anni cinquanta, quando il termine rete sociale fece per la prima volta la sua comparsa in un lavoro di Barnes apparso sulla rivista *Human Relations Journal*. In una rete sociale, infatti, le opinioni non sono statiche, ma proprio a causa delle relazioni di "amicizia", esse possono fluire nella rete e propagarsi in modo pressoché analogo con i meccanismi che molto hanno in comune coi meccanismi che sottendono alla propagazione di epidemie e virus (da cui anche il riferimento al termine "virale" oggi comunemente

utilizzato nel contesto dei social media).

Nel corso degli anni, numerosi modelli sono stati proposti e analizzati per comprendere appieno come si materializzi il processo di diffusione delle opinioni su una rete sociale. Il modello che ha avuto il maggior impatto nell'ambito della comunità degli informatici, seppur inizialmente studiato in sociologia ed economia, è stato quello proposto da Mark Granovetter nel 1978, chiamato threshold<sup>2</sup> o anche tipping model. In questo modello, ogni individuo può essere in uno di due stati: attivo (se l'individuo ha adottato una certa opinione) o non attivo. Ogni individuo ha degli amici e diventerà attivo quando una frazione di essi (ad esempio, la maggioranza) si sarà precedentemente attivata, concretizzando così la pressione sociale evidenziata in precedenza. L'idea è che, inizialmente, solo un certo insieme di individui seed sia attivo, così che l'opinione sulla rete si propagerà in base alla capacità di contagio propria di questi seed. Ad esempio, nella figura 2, se l'opinione che siamo interessati a diffondere è quella relativa alla preferenza per il ristorante, allora gli individui A e B possono essere considerati come i seed del processo di propagazione (attivi nella configurazione iniziale). Abbiamo constatato che questi due seed sono in grado di portare tutta la rete a convergere all'unanimità, attraverso 5 passi di propagazione in cascata. Con seed diversi, invece, potrebbe non essere possibile raggiungere il consenso. Ad esempio, B e C non sono in grado di influenzare gli altri individui e infatti, dalla figura, emerge chiaramente il ruolo cruciale di A nel processo di propagazione. Con riferimento al modello di Granovetter appena discusso, non sarà sorprendente apprendere che molti sforzi sono stati intrapresi in letteratura per studiare algoritmi che, dato un certo insieme di individui seed, siano in grado di prevedere la configurazione finale che verrà a determinarsi e di calcolare in quanto tempo, cioè dopo quanti passi di propagazione, tale configurazione potrà essere raggiunta. Altre domande studiate analiticamente, o mediante strumenti di simulazione in ambito informatico, sono:

possono i seed far convergere la rete all'unanimità su una qualche opinione?

anche se incapaci di portare la rete ad un consenso, possono i seed propagare la loro opinione alla maggioranza degli individui (in altri termini, sono essi in grado di far vincere una elezione in un contesto sociale)?

avendo la possibilità di scegliere dei seed, abbiamo la possibilità di capire quali siano quelli da coinvolgere affinché la capacità del contagio sia la massima possibile (in altri termini, quali sono gli individui cui chiedere aiuto per vincere una elezione in un contesto sociale)?

quali sono gli individui più influenti di una rete sociale e quali le sotto-comunità nelle quali è più complesso propagare una certa opinione?

Al di là dell'interesse speculativo per queste domande, è del tutto evidente che esse rivestono rilevanza pratica quando si affrontano problematiche di manipolazione delle opinioni nelle reti sociali. I risultati delle ricerche in questo ambito possono, ad esempio, essere utilizzati per definire campagne di marketing che, sfruttando le dinamiche virali relative alla propagazione delle opinioni, con un piccolo sforzo iniziale di identificazione di poche decine di seed, consentano di influenzare milioni di individui in tutto il mondo. Ma è ancor più significativo il fatto che le attuali conoscenze

possono essere utilizzate anche per condizionare tangibilmente la percezione sociale di determinati eventi e per alterare gli esiti di votazioni (o mettere in campo le opportune contromisure per prevenire che tali condizionamenti e alterazioni possano avvenire...).

Nonostante il vasto ambito di applicazioni solo accennato, il modello di Granovetter presenta tuttavia alcune limitazioni tali da renderlo inadatto o comunque non completamente affidabile per predizioni su larga scala. In primo luogo, si tratta di un modello, completamente deterministico, che non tiene conto di alcuna forma di incertezza. Nella pratica, invece, è piuttosto irrealistico pensare di poter definire per ogni individuo quale sia l'esatta frazione di amici in grado di determinare un cambio di opinione. Nel nostro esempio, abbiamo infatti sempre ragionato in termini di "maggioranze". Invero, alcuni individui nella rete sociale potrebbero essere più risoluti e essere propensi ad adottare una certa opinione solo se la quasi totalità dei propri



2 <https://doi.org/10.1086/226707>

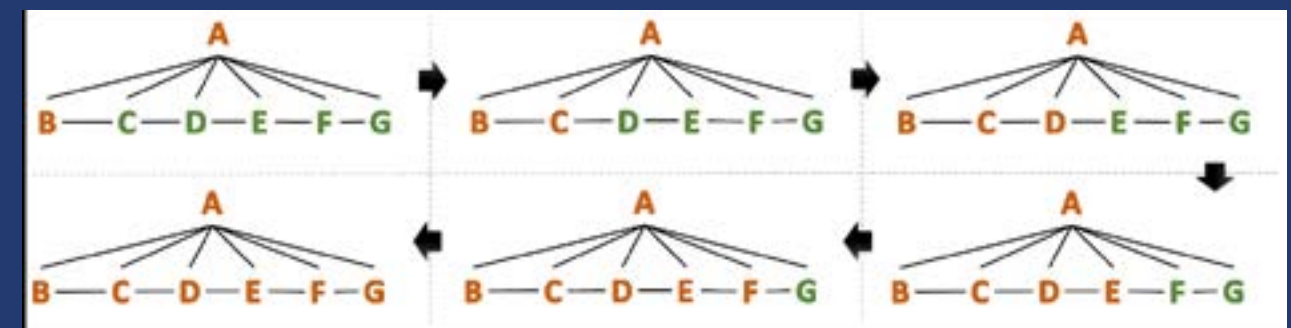


Figura 2: Diffusione a cascata delle opinioni.



amici ha già manifestato la propria preferenza per tale opinione. Altri potrebbero, invece, essere più inclini al cambiamento. Altri ancora potrebbero essere influenzati selettivamente solo da alcuni amici. In altri termini, il grado di “influenzabilità” di ogni individuo ben difficilmente può essere considerato come un parametro chiaramente definibile nel modello; più verosimilmente esso deve essere stimato empiricamente (sulla base di osservazioni della rete sociale) o addirittura – in assenza di ulteriori informazioni – assunto come elemento casuale. Modelli basati su queste assunzioni sono stati proposti nell’ambito dell’informatica già a partire da un lavoro<sup>3</sup> di Kempe, Kleinberg e Tardos del 2003 e costituiscono oggi lo standard de-facto per lo studio dei modelli di diffusione delle opinioni su reti sociali in ambito informatico.

La seconda limitazione del modello di Granovetter è che esso risulta intrinsecamente monotono, nel senso che si basa sull’assunto che, una volta acquisita una certa opinione, l’individuo non possa più ritornare sui suoi passi. Il contagio viene dunque modellato e analizzato come un processo irreversibile, assunzione questa piuttosto irrealistica quando calata nell’esperienza quotidiana dove, invece, le opinioni (che non devono essere necessariamente manifestate) sono strettamente connesse alla funzione del tempo e variano rispetto ad esso proprio in relazione alle interazioni sociali. Modelli che tengono conto di questo dinamismo intrinseco nella formazione delle opinioni sono tipicamente basati su elementi e concetti propri della “Teoria dei Giochi”, poiché muovono dall’idea che ogni individuo deciderà in modo razionale di cambiare la propria opinione ogni qual volta questo risulterà più conveniente. In casi limite, gli individui potrebbero continuare a cambiare continuamente opinione senza convergere mai su una posizione stabile. Più pragmaticamente, tuttavia, anche il continuo cambio di opinione ha un costo (di esposizione sociale) che rende le dinamiche più inerziali al trascorrere del tempo, o comunque esiste un limite temporale superato il quale l’opinione dovrà essere manifestata (come nel caso di una votazione). Dunque anche questi processi sono nella realtà convergenti.

#### Ricerche DeMaCS sulla diffusione delle opinioni

La ricchezza semantica dei modelli di diffusione delle opinioni basati sulla “Teoria dei Giochi”, prima accennata, rende il loro studio molto complesso e articolato. Numerose problematiche sono infatti ancora oggi completamente aperte, in particolare nell’ottica di ri-considerare le domande formulate sul modello di Granovetter (Quali seed scegliere? Come evolverà la rete?).

Negli ultimi anni ho cercato di fornire alcune risposte a queste domande, attraverso una serie di ricerche condotte nel dipartimento di Matematica e Informatica (DeMaCS) dell’Università della Calabria in stretta collaborazione con l’Università di Salerno. Un primo filone ha riguardato lo studio del numero minimo di seed necessario per far convergere la rete all’unanimità verso una opinione desiderata.

Nel lavoro “Reasoning about Consensus when Opinions Diffuse through Majority Dynamics”, scritto assieme ai colleghi Vincenzo Auletta e Diodato Ferraioli, appunto dell’Università di Salerno, è stato dimostrato che – qualunque siano le relazioni sociali – esiste sempre una configurazione iniziale formata da (al più) metà degli individui che è in grado di forzare il consenso unanime sulla propria opinione attraverso la propria pressione sociale (e, sotto opportune condizioni, addirittura con un unico passo di propagazione). Il risultato è concettualmente interessante e in qualche modo sorprendente. Infatti, l’intuizione suggerirebbe che la possibilità di raggiungere il consenso sia intimamente correlata con la struttura delle interazioni. Al contrario, la ricerca non solo dimostra che esiste sempre una maggioranza in grado di prevalere sull’opinione opposta, ma fornisce anche un algoritmo efficiente, che può

scalare su reti reali di grandi dimensioni, in grado di individuare precisamente questa maggioranza. Inoltre, viene dimostrato che per individuare un insieme iniziale di seed capace di assicurare il consenso e che contenga anche un solo individuo in meno rispetto alla metà del numero totale di individui della rete, è necessario in generale utilizzare algoritmi estremamente complessi che potrebbero richiedere enormi tempi di esecuzione. Al di là dell’interesse matematico per i risultati, questa ricerca ha un evidente impatto pratico rispetto allo sviluppo di algoritmi di manipolazione delle opinioni sulla rete e di identificazione degli individui più influenti. A testimonianza di ciò, va sottolineato che la ricerca è stata insignita nel 2018 del prestigioso “IJCAI Distinguished paper award”, riconoscimento che viene attribuito ai migliori lavori presentati nella conferenza IJCAI, il principale forum internazionale di discussione sull’Intelligenza Artificiale che vede ogni anno partecipare migliaia di ricercatori da tutto il mondo per discutere degli avanzamenti e delle nuove frontiere della disciplina.

Un secondo filone di ricerca ha riguardato lo studio delle dinamiche di evoluzioni sociali in contesti in cui sono tre le opinioni a contendersi la maggioranza. I risultati di questa ricerca, condotta oltre che con i colleghi Auletta e Ferraioli, anche con la collega Valeria Fionda dell’Università della Calabria, sono confluiti nel lavoro “Maximizing the Spread of an Opinion when Tertium Datur Est”. Esso ha evidenziato che la presenza di una terza opinione può modificare radicalmente la dinamica di propagazione rispetto al caso in cui siano presenti solo due possibili alternative. In particolare, è stato osservato che per massimizzare la diffusione di una data opinione, non è sempre vantaggioso insistere nella sua propagazione perché, in alcuni casi, risulta opportuno consentire che le altre due opinioni si diffondano in modo da limitarsi reciprocamente. L’effetto individuato, che abbiamo definito “guadagno entropico”, in sostanza tende a premiare la capacità di propagare una certa opinione all’aumentare del bilanciamento nella rete delle altre due opinioni che cercano di contrastarla. Questo effetto è emerso sia dall’analisi teorica sia da numerosi esperimenti condotti su reti reali, in particolare su Facebook.

Un terzo ambito di ricerca, infine, ha riguardato lo studio della propagazione di opinioni in contesti dove oltre agli agenti “conformisti” sono presenti anche agenti che reagiscono negativamente alla conformità sociale tendendo perciò a distinguere la propria opinione da quella dei propri “vicini”. I risultati sono stati pubblicati nel lavoro “Group Reasoning in Social Environments”, scritto con il collega Marco Manna del dipartimento di Matematica e Informatica (DeMaCS) e con il collega Erman Acar, per molti mesi ospite della struttura. Differentemente dai due contesti descritti in precedenza, la maggiore ricchezza semantica del dominio applicativo ha palesato la possibilità che emergano scenari “instabili”, ove cioè qualunque sia l’evoluzione specifica della rete esisterà sempre un qualche agente insoddisfatto e desideroso di modificare la propria opinione.

#### Conclusioni

Molto è stato fatto dalla comunità scientifica, ma tante dinamiche delle reti sociali sfuggono ancora alla nostra comprensione e alla nostra capacità di modellazione e predizione.

La ricerca informatica in questo ambito pone ancora molte affascinanti sfide intellettuali. Essa è preziosa per le sue immediate ricadute pratiche nello sviluppo di algoritmi e metodologie che consentano di monitorare e circoscrivere utilizzi maliziosi dei sistemi di social networking. Ma ancor di più è rilevante che le dinamiche su cui tale ricerca ha già consentito di far luce siano discusse e divulgate tra i non addetti ai lavori, in particolare le giovani generazioni, perché la migliore arma di prevenzione dalle manipolazioni resta comunque un approccio consapevole all’utilizzo degli ecosistemi digitali nei quali sempre più siamo calati.

