



UNIVERSITÀ
DEGLI STUDI
FIRENZE

Scuola di Scienze Matematiche, Fisiche e Naturali

Corso di Laurea in Fisica e Astrofisica

Tesi di Laurea

L'INFLUENZA DELLA PERCEZIONE DEL
RISCHIO NELLE DINAMICHE EPIDEMICHE:
UN APPROCCIO CON LA
MODELLIZZAZIONE AD AGENTI

INFLUENCE OF RISK PERCEPTION IN
EPIDEMIC SPREADING: AN AGENT-BASED
APPROACH

GIULIA LEONE

Relatore: Prof. *Franco Bagnoli*

Anno Accademico 2019-2020

Giulia Leone: *L'influenza della percezione del rischio nelle dinamiche epidemiche: un approccio con la modellizzazione ad agenti*, Versione 0.0.1, Corso di Laurea in Fisica e Astrofisica, © , Università degli Studi di Firenze, Anno Accademico 2019-2020

A <Nome>,
frase di dedica.

"Citazione coltissima" Autore, "Titolo", 19xx, rif bibliografia.

INDICE

| | |
|---|----|
| Introduzione | 1 |
| 1 TITOLO DEL PRIMO CAPITOLO | 2 |
| 1.1 Il modello SIR | 4 |
| 1.2 Modellizzazione ad agenti | 6 |
| 2 TITOLO DEL SECONDO CAPITOLO | 8 |
| 3 PERCEZIONE DEL RISCHIO | 15 |
| 3.1 Il modello e l'approssimazione di campo medio | 15 |
| 3.1.1 Generazione di una rete poissoniana | 16 |
| 3.1.2 Generazione di una rete scale-free | 17 |
| 3.2 Risultati | 20 |
| 4 CONCLUSIONI | 21 |
| A CODICE SVILUPPATO: COMPLETO E COMMENTATO | 22 |
| Acronimi | 23 |
| Bibliografia | 27 |

*"Citazione colta,
ma colta colta,
in italiano."*

*"Citazione colta,
ma colta colta,
in inglese."
— Opera, Autore*

INTRODUZIONE

PROVA DI CITAZIONE

[1]

TITOLO DEL PRIMO CAPITOLO

L'obiettivo dietro alla creazione di un modello matematico di una malattia infettiva è quello di arrivare a comprenderne e a descriverne il processo di trasmissione. In linea del tutto generale, possiamo andare a semplificarlo come segue:

- a. in primo luogo, uno o più soggetti infetti vengono introdotti in una popolazione di individui suscettibili (a rischio, cioè, di contrarre la malattia);
- b. un individuo che viene infettato può inizialmente rimanere asintomatico, per poi mostrare i sintomi; può guarire, sia grazie all'assunzione di medicinali che all'azione del sistema immunitario, ed acquisire così una protezione nei confronti di una possibile reinfezione;
- c. quando il bacino dei potenziali suscettibili viene sufficientemente svuotato, la diffusione inizia a rallentare fino a fermarsi; se vengono aggiunti nuovi soggetti alla popolazione, che sia a seguito di flussi migratori o di nascite, l'epidemia può persistere per un lungo periodo di tempo e diventare così endemica.

La modellazione matematica si è rivelata di centrale importanza nel saper rispondere alle domande che possono sorgere all'alba di quella che, a tutti gli effetti, potrebbe rivelarsi una nuova epidemia (se non una pandemia), come, ad esempio, quale possa essere il numero di persone bisognose di cure ospedaliere o quali effetti possa sortire l'imposizione di una quarantena. La sua forza sta anche nel fatto che gli approcci tradizionali, quello statistico e quello sperimentale, in questo frangente non si rivelano altrettanto utili: se da una parte diventa complicato riprodurre in laboratorio il comportamento su grande scala di una malattia infettiva,

che può coinvolgere un gran numero di persone distribuite in aree geografiche spazialmente estese, dall'altra è difficile fare affidamento su di un'analisi statistica se i dati raccolti non sono completi o accurati (basti pensare alla difficoltà di reperire informazioni su soggetti asintomatici). Lo scopo della modellazione si fa, dunque, triplice:

1. come già detto, consentire una migliore comprensione dei meccanismi di trasmissione dell'infezione;
2. riuscire, di conseguenza, a predirne l'andamento futuro;
3. infine, individuare delle modalità di contenimento per tenere sotto controllo la diffusione [2].

Il processo che si mette in atto consiste di una serie di passi, che vanno dalla formulazione di assunzioni sulla trasmissione della malattia, a partire dalle quali si può costruire un primo modello, alla validazione dello stesso mediante i dati raccolti. C'è, tuttavia, da tenere conto del fatto che un modello matematico non è che un'approssimazione, basata sulle ipotesi che facciamo a seguito di quanto siamo riusciti ad osservare; ciò si traduce nella necessità, imposta anche da una non adeguata conoscenza della malattia in questione, di fare ricorso a delle semplificazioni: risulta, pertanto, chiara l'esigenza di andare a confermare coi dati, qualora possibile, il modello che si è messo in piedi.

Ci sono generalmente tre approcci che si possono seguire:

- quello dei modelli statistici, ampiamente usati in epidemiologia, ma con lo svantaggio, come abbiamo già sottolineato, di necessitare di grandi campioni di dati;
- quello dei modelli deterministici, che prendono in considerazione soltanto l'evoluzione temporale dei valori medi delle dimensioni delle varie popolazioni; risultano meno affidabili se queste ultime constano di pochi individui, tuttavia sono matematicamente maturi e meno dipendenti dai dati;
- infine, quello dei modelli stocastici, che ben si adattano ad essere impiegati nel caso in cui si abbia a che fare con gruppi ristretti, ma che, al contempo, hanno bisogno di un gran numero di simulazioni numeriche [3].

In primo luogo, si va a suddividere la popolazione sotto esame in gruppi mutualmente esclusivi - come se fossero, per l'appunto, *compartimenti*

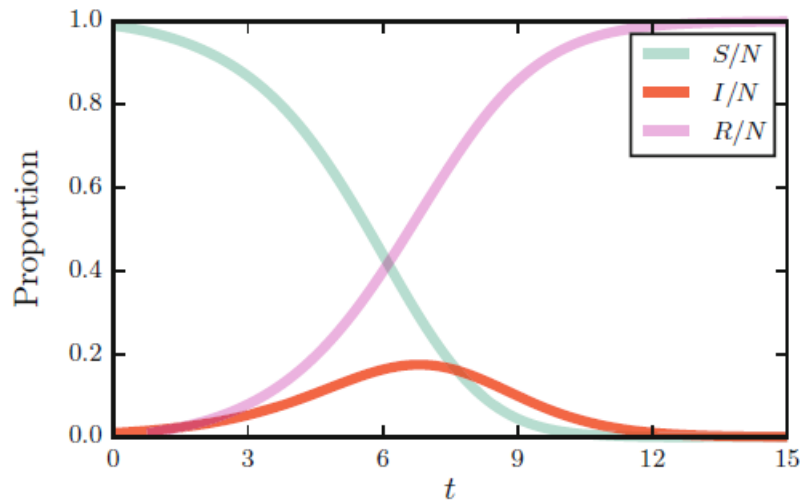


Figura 1: Dipendenza temporale delle dimensioni relative dei compartimenti rispetto alla totalità della popolazione; $\gamma = 1$ e $\beta = 1.6$ [4].

stagni - così che questi possano riflettere caratteristiche osservabili del processo di infezione. Come evidenziato in [4], quello che ci si propone di fare è descrivere nel modo più formale possibile le transizioni da una classe all'altra, così che si possa tener traccia del numero di individui che vi appartengono; detto risultato può essere raggiunto in modo diverso a seconda di quale tra i già citati approcci si sta seguendo: nella fattispecie, quello deterministico, che richiede la scrittura di un sistema di equazioni differenziali ordinarie ove le variabili in gioco sono le dimensioni dei vari compartimenti, consente di conoscere in modo esaustivo il comportamento di una popolazione a partire dalle sue condizioni iniziali, mentre quello stocastico si basa sull'esistenza implicita di una rete di connettività fra gli individui e, pertanto, tratta le informazioni relative all'infezione in funzione del rate di spostamento fra gli elementi di un reticolo.

1.1 IL MODELLO SIR

Sulla falsariga del modello più noto e forse storicamente più rilevante, quello delineato da Kermack e McKendrick nel 1927 [5], possiamo ora introdurre il modello SIR: esso comporta la ripartizione del campione in esame in tre gruppi, ovvero

- S , i suscettibili

- I, gli *infetti*
- R, interpretabile sia come *guariti* (recovered) che come *rimossi* (removed)

la cui evoluzione temporale - messa in evidenza nella figura 1 - è ben descritta dal seguente sistema:

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I(t) \frac{S(t)}{N}; \\ \frac{dI(t)}{dt} = \beta I(t) \frac{S(t)}{N} - \gamma I(t); \\ \frac{dR(t)}{dt} = \gamma I(t). \end{cases} \quad (1)$$

laddove β e γ , entrambe costanti > 0 , sono rispettivamente il rate d'infezione e di guarigione/rimozione, mentre $S(t) + I(t) + R(t) = N$ fornisce la dimensione della popolazione, anch'essa una quantità costante. Al sistema soprastante vengono aggiunte anche le opportune condizioni iniziali: $S(0) = S_0 > 0$, $I(0) = I_0 > 0$, $R(0) = 0$.

Osserviamo che il modello non perde di senso fintanto che $S(t)$ e $I(t)$ rimangono non negativi; qualora anche solo uno dei due diventi pari a zero, affermiamo che il sistema si arresta: in effetti, una qualsiasi soluzione ($I = 0$, $S + N = R$) è un punto fisso per il sistema. Notiamo, inoltre, che $\frac{dS(t)}{dt} < 0$ per ogni possibile valore di t , mentre $\frac{dI(t)}{dt} > 0$ se e solo se $\frac{\beta S(t)}{\gamma} > 1$: ciò significa che, fin quando quest'ultima disuguaglianza rimane vera, $I(t)$ cresce, ma in seguito, in virtù del fatto che $S(t)$ diminuisce a prescindere da t , inverte il suo comportamento e finisce per tendere a zero.

Quanto appena detto ci dà il via per andare a definire una quantità fondamentale nello studio della diffusione di una malattia infettiva, il *numero di riproduzione di base* \mathcal{R}_0

$$\mathcal{R}_0 = \frac{\beta N}{\gamma} \quad (2)$$

che conta quanti individui, in media, vengono contagiati da un singolo infetto in una popolazione che altrimenti sarebbe totalmente sana. In aggiunta, è un buon indicatore dell'eventuale esplodere o meno dell'epidemia; più precisamente

$$\text{se } \mathcal{R}_0 \begin{cases} < 1, & \text{allora } I < 0 \text{ e il processo d'infezione si arresta} \\ > 1, & \text{allora } I \text{ cresce e l'epidemia prende piede} \end{cases}$$

La transizione fra uno scenario e l'altro si ha per $\mathcal{R}_0 = 1$, punto che prende il nome di *soglia epidemica*.

Combinando le prime due equazioni [6] che compaiono in (1), si ottiene

$$\log \frac{S_0}{S_\infty} = \frac{\beta}{\gamma} [N - S_\infty] = \mathcal{R}_0 \left[1 - \frac{S_\infty}{N} \right] \quad (3)$$

che mette in luce la relazione fra il numero di riproduzione di base e la dimensione finale dell'epidemia, $[N - S_\infty]$.¹

1.2 MODELLIZZAZIONE AD AGENTI

Quando parliamo di *modellizzazione ad agenti* intendiamo un tipo di modellizzazione computazionale nel quale un fenomeno viene rappresentato in termine di quelli che vengono chiamati agenti e delle loro interazioni; con *agente* andiamo ad indicare un individuo o un oggetto autonomo in possesso di determinate proprietà e in grado di compiere certe azioni.

I rapporti fra le varie entità che popolano il modello avvengono localmente, cioè solo fra quelli che vengono considerati *vicini*: è immediato, pertanto, pensare che i soggetti sotto esame si muovano su di una rete e che possano necessariamente avere a che fare solo con chi si trova in loro prossimità.

Il primo grande vantaggio derivante dall'uso di questo approccio risiede nel fatto che risulta di più semplice comprensione rispetto ad una rappresentazione di tipo matematico: dal momento che consente di proiettare negli agenti quella che è la nostra esperienza personale, modulata in termini delle interazioni con gli individui coi quali entriamo in contatto, è evidente quanto possa essere più vicino al nostro linguaggio e al nostro modo di pensare. D'altro conto, non si può negare che un'equazione, se risolubile, fornisce in modo diretto un risultato senza il bisogno di far girare un modello, che, specialmente se costituito di un gran numero di agenti, può richiedere un tempo di esecuzione talmente lungo da renderlo poco funzionale [7]. Un ulteriore merito del quale è necessario prendere atto è la sua capacità di mettere in luce il fenomeno dell'*emergenza*, cioè di tutto quell'insieme di comportamenti e proprietà che vengono fuori dall'interazione di singoli individui e che non risultano predicibili a partire dalle caratteristiche di questi ultimi; ciò è possibile perché si tratta di un approccio "across-levels", secondo il quale gli agenti non

¹ Qui abbiamo sottinteso di star cercando delle soluzioni costanti della coppia di equazioni differenziali, (S_∞, I_∞) ; partendo dal presupposto che, a $t = 0$, $S_0 + I_0 = N$ e ricordando che $\lim_{t \rightarrow \infty} I(t) = I_\infty = 0$, risulta che $\lim_{t \rightarrow \infty} S(t) + I(t) = S_\infty$ [6].

vengono considerati avulsi dall'ambiente in cui sono immersi, ma viene dato particolare rilievo alla reciproca influenza fra i due.

Non sarà sorprendente, a questo punto, affermare che il processo di modellizzazione sottostà ad una serie di passi che possono venire iterati più volte - dando luogo a quello che chiamiamo "modeling cycle" - in modo da rendere il modello il più efficace ed efficiente possibile:

1. *formulazione delle domande*, per meglio mettere a fuoco il problema che si vuole analizzare;
2. *costruzione delle ipotesi*, che devono inizialmente essere il più semplici possibili, per venir poi rinforzate ed arricchite più avanti nel processo;
3. *scelta delle variabili di stato e dei parametri*, per iniziare a mettere per iscritto il modo in cui ci aspettiamo che il nostro modello si comporti;
4. *implementazione del modello*, che consente di esplorare e valutare se le assunzioni fatte sono valide ed hanno prodotto un che di utile;
5. *analisi, test e revisione*, così da apportare modifiche o migliorie;
6. *comunicazione del modello* [8].

In ultimo, è interessante sottolineare, seppur brevemente, quale sia il ruolo epistemico della simulazione computazionale: è, infatti, possibile attribuirle un potere sia *esplicativo*, poiché permette di comprendere come si siano verificati alcuni eventi, sia *predittivo*, dal momento che rende possibile immaginare il comportamento futuro di un sistema sotto determinate circostanze, che *esplorativo*, in quanto fa sì che le informazioni che intendiamo rappresentare possano essere condivise e rese note ad altri [9].

All'interno di questo lavoro di tesi faremo uso di Netlogo, che è sia un IDE, ovvero un ambiente di sviluppo integrato, che un linguaggio di modellazione; è stato progettato nel 1999 col fine di rendere il più semplice possibile la costruzione di modelli basati su agenti. Nella fattispecie, utilizzeremo la versione 6.2.0; per ulteriori dettagli tecnici, rimandiamo alla documentazione ufficiale [10].

TITOLO DEL SECONDO CAPITOLO

Le assunzioni secondo cui ogni individuo possa entrare in contatto con chiunque (*mixing omogeneo*) e il numero di interazioni di ciascun soggetto sia confrontabile con quello degli altri non sono realistiche: la probabilità che si verifichi un incontro fra due individui presi a caso è praticamente infinitesima, ma, al contrario, quella che due compartimenti distinti entrino in relazione l'uno con l'altro è senza dubbio finita. Di norma, i contatti fra individui sono limitati ad un numero ristretto di persone (familiari, colleghi, etc), mentre tutto il resto della popolazione viene ignorato; questo li rende particolarmente adatti ad essere rappresentati tramite una rete [11].

Andiamo ad introdurre una serie di definizioni che in seguito ci risulteranno utili.

Definizione 2.1 (*Grafo*). Un *grafo* (o una *rete*) è un insieme di elementi detti vertici o *nodi* che possono essere collegati fra loro da segmenti detti archi o *link* [12].

Consideriamo una rete non orientata - cioè una rete in cui i link possono essere percorsi indistintamente in un verso e nell'altro - con n vertici, che andiamo ad etichettare da 1 a n . Se indichiamo con (i, j) l'arco fra i nodi i e j , allora l'intera rete può essere descritta in funzione della

Definizione 2.2 (*Matrice di adiacenza* I). La matrice di adiacenza \mathbf{A} relativa ad un grafo semplice è una matrice i cui elementi sono così definiti [13]:

$$A_{ij} = \begin{cases} 1, & \text{se esiste un arco fra } i \text{ e } j; \\ 0, & \text{altrimenti.} \end{cases}$$

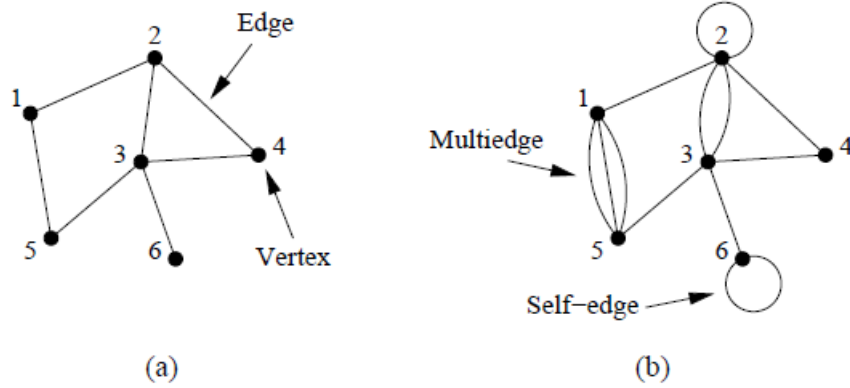


Figura 2: (a) Un grafo semplice, cioè privo di loop o link multipli. (b) Un grafo che presenta entrambi [13].

Se, ad esempio, prendiamo la rete (a) in figura 2, assumerà la seguente forma:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}. \quad (4)$$

Possiamo osservare che risulta essere una matrice simmetrica con tutti 0 sulla diagonale.

Qualora invece ne avessimo sotto esame una più simile alla rete (b) della medesima figura, dovremmo tener conto del fatto che sono presenti link multipli e loop; si stabilisce di assegnare ai primi un numero pari alla loro molteplicità e ai secondi il valore 2. Ne risulta una matrice, ancora una volta, simmetrica:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 3 & 0 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{pmatrix}. \quad (5)$$

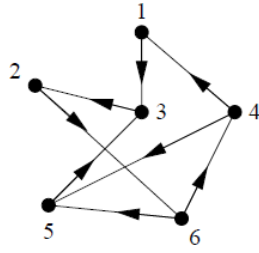


Figura 3: Un digrafo [13].

La questione cambia leggermente se si va a considerare una rete diretta o *digrafo*.

Definizione 2.3 (Digrafo). Un *digrafo* è un tipo di grafo in cui ogni arco ha una direzione, punta cioè da un vertice ad un altro [13].

Ciò ci porta a rivedere quanto detto per la matrice di adiacenza: affermiamo, quindi, che

Definizione 2.4 (Matrice di adiacenza II). La matrice di adiacenza **A** relativa ad un grafo orientato è una matrice i cui elementi sono così definiti [13]:

$$A_{ij} = \begin{cases} 1, & \text{se esiste un arco da } j \text{ a } i; \\ 0, & \text{altrimenti.} \end{cases}$$

Relativamente alla figura 3, la matrice assumerà, pertanto, la seguente forma:

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (6)$$

Ci soffermiamo poi su un particolare tipo di grafo, che ci risulterà più utile in seguito.

Definizione 2.5 (Grafo bipartito). Un grafo si dice *bipartito* se i suoi vertici possono essere suddivisi in due sottoinsiemi X e Y tali che ogni link ha un'estremità in X ed una in Y [14].

In modo del tutto equivalente a quanto abbiamo già fatto, possiamo andare a definire per un grafo del genere una matrice che lo va a descrivere.

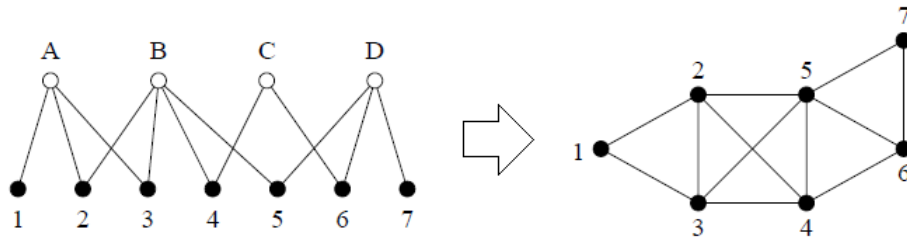


Figura 4: Passaggio da una rappresentazione tipica di una rete bipartita ad una in cui compaiono solo i vertici [13].

Definizione 2.6 (*Matrice di incidenza*). La matrice di incidenza \mathbf{B} è una matrice $g \times n$, dove g è il numero di sottoinsiemi e n quello dei vertici che ne fanno parte, i cui elementi sono definiti come segue [13]:

$$B_{ij} = \begin{cases} 1, & \text{se il vertice } j \text{ appartiene al gruppo } i; \\ 0, & \text{altrimenti.} \end{cases}$$

Uno dei pregi di una rete bipartita è quello di consentire in modo fluido il passaggio da una rappresentazione in cui compaiono sia vertici che gruppi ad una in cui, come viene messo in evidenza in figura 4, sono coinvolti soltanto i vertici; questa possibilità permette di visualizzare il formarsi di quelle che prendono il nome di cricche (*cliques*).

Definizione 2.7 (*Cricca*). Una *cricca* è un sottografo completo, cioè un sottografo in cui ogni vertice è connesso a tutti gli altri [12].

L'emergere di questi raggruppamenti, di conseguenza, mette in rilievo l'esistenza di gruppi fortemente coesi all'interno della rete; questo aspetto assume ulteriore valore, in particolare, quando cerchiamo di rispondere ad una domanda che ci siamo posti nel precedente capitolo, ovvero come poter rappresentare una popolazione nella quale stia prendendo piede un certo tipo di malattia infettiva, perché ci consente non solo di dare forma agli individui che la costituiscono, ma anche all'intensità dei legami che intercorrono fra di essi: ci aspettiamo, infatti, che membri di una famiglia o colleghi di lavoro vadano a costituire una cricca.

Definizione 2.8 (*Grado di un vertice*). Diciamo *grado* di un vertice il numero di archi ad esso connessi [15].

Se abbiamo un grafo non orientato con n vertici, possiamo esprimere il grado dell' i -esimo vertice come

$$k_i = \sum_{j=1}^n A_{ij}. \quad (7)$$

Osserviamo adesso che, se una rete consta di m link, il numero totale di estremità che possiamo contare - pari a $2m$ - altro non è se non la somma dei gradi di tutti i vertici:

$$m = \frac{1}{2} \sum_{i=1}^n k_i, \quad (8)$$

il che ci porta ad esprimere il grado *medio* di un vertice come

$$c = \frac{1}{2} \sum_{i=1}^n k_i = \frac{2m}{n}. \quad (9)$$

Indichiamo poi la *densità* di un grafo semplice come la frazione di archi effettivamente presenti:

$$\rho = \frac{m}{\binom{n}{2}} = \frac{2m}{n(n-1)} = \frac{c}{n-1} \quad (10)$$

con $0 \leq \rho \leq 1$ e $\binom{n}{2}$ numero massimo di link che è possibile tracciare. Ci soffermiamo brevemente sulle stesse grandezze nel caso in cui si stia esaminando una rete orientata:

- si può distinguere fra il grado esterno (relativo agli archi uscenti da un vertice)

$$k_i^{\text{out}} = \sum_{j=1}^n A_{ij} \quad (11)$$

e quello interno (relativo agli archi entranti)

$$k_j^{\text{in}} = \sum_{i=1}^n A_{ij} \quad (12)$$

- il numero di link presenti è pari a

$$m = \sum_{i=1}^n k_i^{\text{in}} = \sum_{j=1}^n k_j^{\text{out}} = \sum_{ij} A_{ij} \quad (13)$$

- il grado medio esterno è equivalente a quello interno

$$c^{\text{in}} = \frac{1}{n} \sum_{i=1}^n k_i^{\text{in}} = \frac{1}{n} \sum_{j=1}^n k_j^{\text{out}} = c^{\text{out}}. \quad (14)$$

Il concetto di grado ci è immediatamente utile per definirne un altro: quello di *assortatività* o di *mixing assortativo*. Con questo termine indichiamo una tendenza, da parte dei vertici, a legarsi con altri nodi che, in qualche modo, somigliano loro; è piuttosto immediato pensare che un aspetto che possono avere in comune sia proprio il grado e, in effetti, in questo contesto si parla di *mixing assortativo per grado* [13]: è possibile andarlo a quantificare mediante un coefficiente di correlazione,

$$r = \frac{\sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) k_i k_j}{\sum_{ij} \left(k_i \delta_{ij} - \frac{k_i k_j}{2m} \right) k_i k_j}. \quad (15)$$

Una rete reale, tuttavia, non possiede alcun tipo di regolarità, almeno ad un primo sguardo; la difficoltà sta proprio nel riuscire a riprodurre nel miglior modo possibile la disposizione casuale degli archi. A questo proposito andiamo ad introdurre il concetto di *grafo aleatorio*, che viene descritto da una serie di parametri, alcuni dei quali fissati ed altri liberi.

Definizione 2.9 (*Modello G (n, m)*). Un grafo aleatorio di questo tipo consiste di n nodi connessi da m link posizionati casualmente.

Definizione 2.10 (*Modello G (n, p)*). All'interno di questo modello, ciascuna coppia di nodi fra gli n presenti risulta legata con una probabilità p .

Poiché il calcolo di tutta una serie di quantità caratteristiche risulta più semplice nel caso del secondo modello, ci concentreremo essenzialmente su quello [11].

La probabilità che uno qualunque di tutti i possibili grafi semplici G appaia è

$$P(G) = p^m (1 - p)^{\binom{n}{2} - m} \quad (16)$$

ove m è il numero di archi; pertanto, la probabilità totale di disegnare un grafo con m link a partire dal nostro insieme segue una distribuzione binomiale standard

$$P(m) = \binom{\binom{n}{2}}{m} p^m (1 - p)^{\binom{n}{2} - m}. \quad (17)$$

Il numero atteso di archi, di conseguenza, sarà semplicemente:

$$\langle m \rangle = \sum_{m=0}^{\binom{n}{2}} m P(m) = \binom{n}{2} p, \quad (18)$$

mentre il grado medio verrà così calcolato:

$$\langle k \rangle = \sum_{m=0}^{\binom{n}{2}} \frac{2m}{n} P(m) = (n-1)p \doteq c. \quad (19)$$

Sulla base di quanto appena detto, quindi, p è la probabilità che un vertice si leghi con uno qualunque degli altri $n-1$ vertici; ciò ci porta ad esprimere la probabilità totale di essere connesso ad esattamente k nodi è data da:

$$p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}; \quad (20)$$

per $n \rightarrow \infty$, possiamo andare ad approssimarla come segue

$$p_k \simeq \frac{(n-1)^k}{k!} p^k e^{-c} = \frac{(n-1)^k}{k!} \left(\frac{c}{(n-1)} \right)^k e^{-c} = \frac{c^k e^{-c}}{k!} \quad (21)$$

cioè $G(n, p)$ ha una *distribuzione di grado* poissoniana [13]. Distinguiamo due casi limite [11]:

- se $p = 0$, $\langle k \rangle = 0$ e nessun link è presente;
- se $p = 1$, $\langle k \rangle = 1$ e ogni vertice risulta connesso a tutti gli altri.

Va da sé che, se indichiamo con n_g la dimensione del più grosso cluster connesso all'interno della rete, nel primo scenario $n_g = 1$, mentre nel secondo $n_g = n$; al crescere di $\langle k \rangle$ e di conseguenza del rapporto $\frac{n_g}{n}$, infatti, viene ad emergere quella che indichiamo col nome di *componente gigante* [16].

PERCEZIONE DEL RISCHIO

Un elemento sul quale vorremo concentrare la nostra attenzione è la *percezione del rischio*, poiché ci aspettiamo che possa rappresentare un fattore chiave nell'andamento dell'epidemia. Si osserva, infatti, che, qualora ci si trovi di fronte ad una malattia i cui sintomi sono evidentemente manifesti, le persone modificano il proprio comportamento e le proprie abitudini sulla base della propria soggettiva sensazione di pericolo; ad esempio, potrebbero prediligere l'uso di un mezzo di trasporto privato, rispetto ad uno pubblico, per recarsi a lavoro, in modo da limitare le occasioni di contatto con estranei. È altresì vero che la percezione del rischio varia anche col grado di intimità che si ha con un determinato individuo: in linea di massima, la tendenza è quella di temere meno l'infezione se ci si trova in compagnia di un familiare o di un amico, mentre in presenza di estranei il livello di allerta si fa ben più alto.

Per amore di semplicità, in questo lavoro assumeremo che tutti i soggetti percepiscano il rischio e reagiscano ad esso allo stesso modo: facciamo, cioè, un'ipotesi di *popolazione omogenea*. Aggiungiamo due ulteriori supposizioni:

1. la malattia si rende visibile nel medesimo momento in cui infetta qualcuno;
2. la percezione del rischio è proporzionale alla frazione di contatti con individui infetti rispetto alla totalità dei contatti [17].

3.1 IL MODELLO E L'APPROSSIMAZIONE DI CAMPO MEDIO

Poniamo la nostra attenzione su reti costituite da N nodi e $2mN$ link. Come già fatto in precedenza, andiamo ad indicare con a_{ij} l' ij -esimo elemento della matrice di adiacenza - pari a 1 se esiste un arco da j a i e 0 altrimenti - e con $k_i = \sum_j a_{ij}$ la connettività del sito i - così che risulta

$\langle k \rangle = 2m$; denotiamo, invece, quella dei suoi vicini come $j_1^i \dots j_{k_i}^i$. Andiamo a sostituire alla probabilità d'infezione netta, indicata con la lettera τ , la quantità $u(s, k)$: nello specifico, questa può esplicitamente essere scritta come

$$u(s, k_i) = \tau \exp(-J \frac{s}{k_i}) \quad (22)$$

e va a rappresentare la probabilità che il sito i , avente connettività k_i , contragga l'infezione da uno degli s vicini malati [18]. Quanto appena scritto riflette in pieno quanto asserivamo poco sopra, ovvero che l'eventualità di venire contagiati diminuisca in funzione della percezione del rischio, data dalla percentuale di vicini infetti e modulata da un fattore J , che può dipendere dal tipo di misure di precauzione adottate.

La più semplice approssimazione di campo medio che è possibile formulare richiede di trascurare l'eventuale correlazione fra le variabili; in altre parole, si assume che non ci siano loop.

3.1.1 Generazione di una rete poissoniana

Poniamoci nel caso in cui la rete che rappresenta i contatti sia aleatoria: ciò significa che ciascun nodo forma un numero m di link con altrettanti vertici scelti in modo casuale e che la distribuzione di grado segue una poissoniana. Inoltre, stabiliamo che k sia fissato: allora, se indichiamo con c la frazione di infetti al tempo t e con c' quella al tempo immediatamente successivo $t + 1$, possiamo scrivere che

$$c' = \sum_{s=0}^k \binom{k}{s} c^s (1-c)^{k-s} p(s, k), \quad (23)$$

laddove $p(s, k) = 1 - [1 - u(s, k)]^s$ è la probabilità di ammalarsi se sono presenti s vicini infetti su k [18].

Apriamo, a questo punto, una doverosa parentesi ed introduciamo un ulteriore concetto caro alla teoria dei grafi: la *percolazione*, ovvero la rimozione di un vertice di una rete (*p. di sito*) oppure quella di un arco (*p. di legame*). In questo contesto, possiamo affermare che un link viene occupato con una probabilità ϕ ; link di questo tipo sono quelli lungo cui può essere trasmessa l'infezione. Al crescere di ϕ , nodi inizialmente isolati si fanno via via più grandi, fino ad arrivare a creare un cluster unico che altro non è se non quella che abbiamo chiamato componente gigante e che raggiunge la sua dimensione massima per $\phi = 1$; questo

fenomeno prende il nome di *transizione di percolazione*. Quello che si può arrivare a dire è che la soglia di percolazione su di una rete coincide con la soglia epidemica per la diffusione di una malattia sulla medesima rete: infatti, per valori piccoli di ϕ il primo infetto appartiene necessariamente ad un cluster di dimensioni modeste e, di conseguenza, la maggior parte della popolazione non viene infettata; al contrario, quando ϕ è grande, aumenta il numero di vertici facenti parte della componente gigante e, di pari passo, la probabilità che il paziente zero vi faccia parte e dia l'avvio ad una vera e propria ondata epidemica [13].

Torniamo al nostro problema. Ci aspettiamo che in prossimità della soglia di percolazione/epidemica la (22) sia piccola, il che ci consente di approssimare

$$c' \simeq \sum_{s=0}^k \binom{k}{s} c^s (1-c)^{k-s} s \tau \exp(-J \frac{s}{k}); \quad (24)$$

ponendo $a \doteq \exp(-\frac{J}{k})$, si trova che

$$c' = \tau a k (ca + 1 - c)^{k-1}. \quad (25)$$

La soglia critica J_c corrisponde allo stato stazionario $c = c'$, nel limite $c \rightarrow 0$:

$$\tau = \frac{1}{k} \exp(\frac{J_c}{k}), \quad J_c = k \ln(k\tau). \quad (26)$$

La nostra predizione è piuttosto precisa: ce lo confermano le simulazioni eseguite, come possiamo osservare in figura 5.

3.1.2 Generazione di una rete scale-free

Vediamo adesso cosa cambia se prendiamo in esame una rete a invarianza di scala o *scale-free*: reti siffatte sono descritte da una distribuzione di grado data da $P(k) = Ak^{-\alpha}$, con A costante e, in genere, $2 \leq \alpha \leq 3$ ¹. Imponiamo che il grado di ciascun nodo sia limitato, $m \leq k \leq K$ ²; consci

-
- ¹ In realtà, la distribuzione di grado di una rete scale-free non segue pedissequamente l'andamento di una legge di potenza: questa ultima, infatti, risulta essere sempre una funzione monotona decrescente su tutto il proprio dominio, mentre capita che $P(k)$ si comporti come atteso solo in prossimità delle code della distribuzione.
 - ² Può succedere che $P(k)$ devii anche in corrispondenza di k grande; si introduce perciò un taglio K di qualche tipo in modo da andare a limitare il grado massimo dei vertici [13].

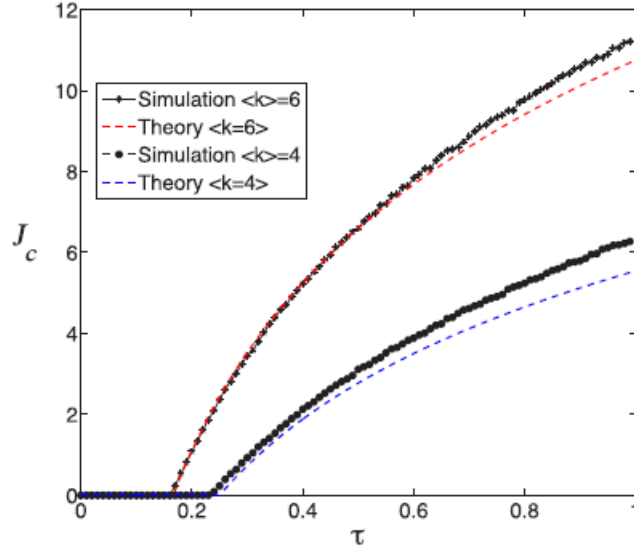


Figura 5: Confronto fra l'approssimazione di campo medio e le simulazioni condotte per una rete aleatoria al variare di $\langle k \rangle$ [18].

del fatto che $\sum_{k=0}^K p_k = 1$, possiamo normalizzare e ricavare il valore di A , cioè $A = \frac{\alpha-1}{m^{1-\alpha} - K^{1-\alpha}}$.

Utilizzando lo stesso formalismo impiegato in precedenza, possiamo prima di tutto esprimere lo stato di un nodo con connettività k al tempo $t+1$ nel modo seguente:

$$c'_k = \sum_{s_1 \dots s_k=0}^1 \sum_{j_1 \dots j_k=0}^{\infty} \prod_{i=1}^k C(j_i, k) I(s_i, j_i) T(k|s_i), \quad (27)$$

dove l'indice $i = 1, \dots, k$ corre sul numero dei vicini e $s_i = 0, 1$ esprime il loro stato (rispettivamente, sani e infetti). $C(j, k)$ è la probabilità che un nodo con connettività j sia legato ad un altro con connettività k , $I(s_i, j_i)$ quella che l' i -esimo vicino sia malato e $T(k|s_i)$, infine, quella che possa trasmettere l'infezione al nodo sotto esame.

Osservando che è possibile riscrivere

- $C(j, k) = \frac{jP(j)}{\langle k \rangle}$ (in quanto per reti simmetriche vale $jC(j, k)P(j) = kC(k, j)P(k)$),
- $I(s_i, j_i) = c_{j_i}^{s_i} (1 - c_{j_i})^{1-s_i}$,
- $T(k|s_i) = 1 - \left(1 - \tau \exp(-\frac{I s_i}{k})\right)^s \simeq s \tau \exp(-\frac{I s_i}{k})$

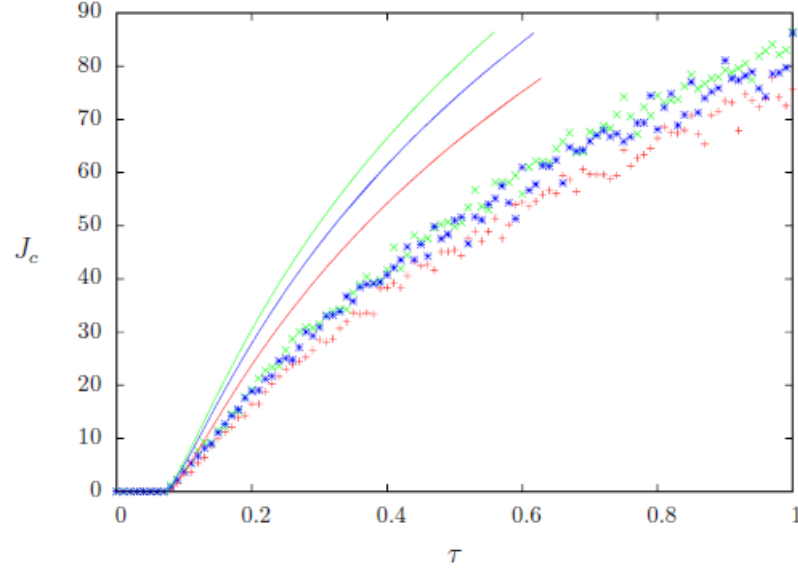


Figura 6: Confronto fra l'approssimazione di campo medio (le curve) e le simulazioni condotte per 3 diversi casi di reti scale-free con $\alpha = 2.4$, $N = 10000$, $m = 2$, $K = 300$ [18].

e definendo $\tilde{c} \doteq \sum_j j \frac{P(j)c_j}{\langle k \rangle}$, possiamo concludere che [18]

$$c'_k = \tilde{c} k \tau \exp\left(-\frac{I_s}{k}\right) \left(\tilde{c} \exp\left(-\frac{I_c}{k}\right) + 1 - \tilde{c} \right). \quad (28)$$

In prossimità della soglia epidemica, $\tilde{c} \rightarrow 0$; questo ci permette di ricavare nuovamente la relazione fra τ_c e J_c :

$$\tau_c(J_c) = \frac{\langle k \rangle}{\sum_k k^2 P(j) \exp\left(-\frac{J_c}{k}\right)}. \quad (29)$$

L'espressione soprastante può essere esplicitata se ci si pone nel limite continuo e si mettono in atto le opportune sostituzioni; infatti, per $K \gg m$, si ottiene che

$$\tau_c(J_c) = J_c^{3-\alpha} \left[\Gamma\left(\alpha - 3, \frac{J_c}{K}\right) - \Gamma\left(\alpha - 3, \frac{J_c}{m}\right) \right] \quad (30)$$

che risulta evidentemente divergente per $K \rightarrow \infty$ - il che ci fa concludere che $J_c = 0 \forall \tau$ -. Una rete reale, d'altra parte, possiede sempre un numero finito di nodi, per cui anche il taglio deve essere tale; questo ci consente di valutare τ per $J_c = 0$:

$$\tau_c(0) = \frac{\alpha - 3}{\alpha - 2} \frac{m^{2-\alpha} - K^{2-\alpha}}{m^{3-\alpha} - K^{3-\alpha}} \simeq \frac{3 - \alpha}{\alpha - 2} \frac{m^{2-\alpha}}{K^{3-\alpha}} \quad (31)$$

Notiamo in figura 6 che, stavolta, c'è un buon accordo fra il modello teorico e i risultati ottenuti dalle simulazioni solo per $J_c = 0$; oltretutto, il valore di $\tau_c(0)$ ottenuto a partire dall'approssimazione continua (31), $\tau(0) \simeq 0.037$, è piuttosto diverso da quello calcolato mediante le simulazioni, $\tau(0) \simeq 0.08$.

3.2 RISULTATI

Quello che possiamo affermare è che, a prescindere dal tipo di rete che andiamo a generare, l'adozione di un livello di precauzione J sufficientemente alto da parte di tutti gli agenti consente di fermare la diffusione epidemica anche se si è oltrepassata la soglia τ_c [18].

CONCLUSIONI

A

CODICE SVILUPPATO: COMPLETO E COMMENTATO

ACRONIMI

ELENCO DELLE FIGURE

| | | |
|----------|--|----|
| Figura 1 | Dipendenza temporale delle dimensioni relative dei compartimenti rispetto alla totalità della popolazione; $\gamma = 1$ e $\beta = 1.6$ [4]. | 4 |
| Figura 2 | (a) Un grafo semplice, cioè privo di loop o link multipli. (b) Un grafo che presenta entrambi [13]. | 9 |
| Figura 3 | Un digrafo [13]. | 10 |
| Figura 4 | Passaggio da una rappresentazione tipica di una rete bipartita ad una in cui compaiono solo i vertici [13]. | 11 |
| Figura 5 | Confronto fra l'approssimazione di campo medio e le simulazioni condotte per una rete aleatoria al variare di $\langle k \rangle$ [18]. | 18 |
| Figura 6 | Confronto fra l'approssimazione di campo medio (le curve) e le simulazioni condotte per 3 diversi casi di reti scale-free con $\alpha = 2.4$, $N = 10000$, $m = 2$, $K = 300$ [18]. | 19 |

ELENCO DEI CODICI

ELENCO DELLE TABELLE

BIBLIOGRAFIA

- [1] Franco Bagnoli, Daniele Lorini e Pietro Lió: *Modeling Social Groups, Policies and Cognitive Behavior in COVID-19 Epidemic Phases. Basic Scenarios*. Substantia, vol. 4(no. 1): 914, Jun. 2020. <https://doi.org/10.13128/Substantia-914>. (Cited on page 1.)
- [2] Daryl J. Daley e Joseph Gani: *Epidemic Modelling: An Introduction*. Cambridge University Press, 2001. (Cited on page 3.)
- [3] Michael Y. Li: *An Introduction to Mathematical Modeling of Infectious Diseases*. Springer, 2018. (Cited on page 3.)
- [4] István Z. Kiss, Joel C. Miller e Péter L. Simon: *Mathematics of Epidemics on Network: From Exact to Approximate Models*. Springer, 2017. (Cited on pages 4 and 24.)
- [5] William O. Kermack e Anderson G. McKendrick: *A contribution to the mathematical theory of epidemics*. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 115(772):700–721, Aug. 1927. <https://doi.org/10.1098/rspa.1927.0118>. (Cited on page 4.)
- [6] Fred Brauer, Carlos Castillo-Chavez e Zhilan Feng: *Mathematical Models in Epidemiology*. Springer, 2019. (Cited on page 6.)
- [7] William Rand e Uri Wilensky: *An Introduction to Agent-Based Modeling*. MIT Press, 2015. (Cited on page 6.)
- [8] Volker Grimm e Steven F. Railsback: *Agent-Based and Individual-Based Modeling*. Princeton University Press, 2019. (Cited on page 7.)
- [9] Giuseppe Primiero: *On the Foundations of Computing*. Oxford University Press, 2019. (Cited on page 7.)
- [10] Uri Wilensky: *Netlogo itself*, 1999. <https://ccl.northwestern.edu/netlogo>. (Cited on page 7.)
- [11] Albert-László Barabási: *Network Science*, 2016. <http://networksciencebook.com/>. (Cited on pages 8, 13, and 14.)

- [12] Allan Bickle: *Fundamentals of Graph Theory*. American Mathematical Society, 2020. (Cited on pages 8 and 11.)
- [13] Mark E. J. Newman: *Networks: An Introduction*. Oxford University Press, 2010. (Cited on pages 8, 9, 10, 11, 13, 14, 17, and 24.)
- [14] John A. Bondy e Uppaluri S. R. Murty: *Graph Theory*. Springer, 2008. (Cited on page 10.)
- [15] Gary Chartrand, Linda Lesniak e Ping Zhang: *Graphs and Digraphs*. CRC Press, 2016. (Cited on page 11.)
- [16] Paul Erdős e Alfréd Rényi: *On Random Graphs I*. *Publicationes Mathematicae*, 6:290–297, 1959. (Cited on page 14.)
- [17] Franco Bagnoli, Pietro Liò e Luca Sguanci: *Risk perception in epidemic modeling*. *Phys. Rev. E*, 76(6):061904, Dec. 2007. (Cited on page 15.)
- [18] Franco Bagnoli e Emanuele Massaro: *Epidemic spreading and risk perception in multiplex networks: A self-organized percolation method*. *Phys. Rev. E*, 90(5):052817, Nov. 2014. <https://doi.org/10.1103/PhysRevE.90.052817>. (Cited on pages 16, 18, 19, 20, and 24.)

RINGRAZIAMENTI

Ringraziamento strappalacrime 1.

Ringraziamento strappalacrime 2.