

ABSONES: Agent Based SOcial NEtwork Simulator

Simone Ciccolella

s.ciccolella@campus.unimib.it

Mat. 762234

Daniele Bellani

d.bellani1@campus.unimib.it

Mat. 780675

Sommario. In questo progetto proponiamo un modello multi-agente per la simulazione di social network, in particolare la piattaforma di microblogging Twitter. Il nostro approccio prevede la rappresentazione del sistema come un grafo diretto e l'identificazione degli utenti come agenti di tale sistema complesso, che si attivano e compiono azioni sulla base di probabilità ed estrazioni di tipo Monte Carlo. Gli obiettivi che ci siamo posti sono stati la prevenzione del *completamento* della rete e il mantenimento del principio di *invarianza di scala*, così da simulare, in modo quanto più fedele, il comportamento della vera rete sociale. Con il modello così costruito, è possibile studiare fenomeni sociologici usando scenari simulati nel modo più attinente possibile alla realtà. In particolare, abbiamo provato a simulare il “salto” di popolarità di un individuo poco conosciuto in seguito all'interazione con un utente più “in vista”, traendo ispirazione da quanto avvenuto nel 2012 tra l'On. Maurizio Gasparri e l'utente Daniele Termite.

Indice

1	Introduzione	3
2	Letteratura correlata	3
3	Descrizione modello	5
4	Visione multi-agente	5
5	Accenni teoria dei grafi	5
6	Esperimenti	5
7	Conclusione	5
8	Sviluppi futuri	5
8.1	Fabrication	5
8.2	Experimental set-up	5
9	Results and interpretation 2-3 pages	5
9.1	Classical regime	5
9.2	Quantum regime	6
10	Discussion 1/2-1 page	6
11	Some LaTeX tips	7
11.1	How to Include Figures	7
11.2	How to Make Tables	7
11.3	How to Write Mathematics	7
11.4	How to Make Sections and Subsections	7
11.5	How to Make Lists	7

1 Introduzione

I social network sono uno strumento che è diventato parte integrante della vita di tutti. Nel giugno 2017 la piattaforma Facebook ha passato la soglia dei due miliardi di utenti attivi su base mensile [1]; nel frattempo, Twitter si è imposto come mezzo di comunicazione principale tra i cosiddetti “influencer”, con un grosso impatto sull’opinione pubblica (per esempio, giocando un ruolo fondamentale nell’elezione di Donald Trump [2]). L’aspetto principale di questa forma di comunicazione è la produzione, da parte degli utenti, di un’enorme mole di dati: le acquisizioni di YouTube da parte di Google e di LinkedIn da parte di Microsoft testimoniano l’interesse delle grandi aziende per queste sorgenti di dati. L’analisi di questi ultimi aiuta a determinare strategie di mercato, a personalizzare raccomandazioni di prodotti, studiare e prevedere il sentimento su un evento o un prodotto, oppure condurre studi di stampo sociologico.

Sorgono però alcune difficoltà. In primo luogo, raramente questi dati vengono rilasciati in formato aperto (Open Data). Inoltre, le loro enormi proporzioni ne rendono difficile l’analisi e la gestione. Da qui la necessità di svolgere delle simulazioni *in-silico* con modelli sviluppati in modo da essere il più possibile veritieri. L’attendibilità di queste simulazioni dipende dal modello su cui vengono eseguite e dalle inevitabili assunzioni che sono state fatte durante la sua costruzione. L’obiettivo è quindi produrre sistemi artificiali che siano repliche fedeli di sistemi complessi reali. Una possibile strada è quella dei *sistemi multi-agente*. Questi si basano sulla definizione di *agente* (vedi sezione 4), un’entità virtuale [6] capace di compiere azioni ed interagire con altri agenti all’interno di un ambiente (*environment*). Si può ottenere una rappresentazione multi-agente di una rete sociale rappresentando l’ambiente come un grafo (connesso o non connesso) e gli agenti come nodi di questo grafo. Gli archi rappresentano le interazioni tra i vari agenti, siano esse relazioni di “amicizia” (es. Facebook), oppure di “subscription” (es. Twitter).

Negli ultimi anni sono stati proposti alcuni modelli per la simulazione multi-agente di social network, con risultati alterni e non definitivi, dovuti alla complessità del tema. Da questi abbiamo tratto ispirazione per il nostro modello. Nella prima sezione esamineremo alcune delle metodologie proposte, dopoiché passeremo all’illustrazione della nostra soluzione e della visione agent-based dello stesso. Nelle ultime due sezioni esporremo gli esperimenti condotti e ne commenteremo i risultati.

2 Letteratura correlata

La letteratura riguardante modellazione multi-agente di social network è relativamente scarsa. Nonostante la pubblicazione in formato open di dati provenienti da alcune piattaforme, risulta comunque molto difficile avere una visione generale di ciò che accade in sistemi così ampi e complessi.

Un primo lavoro degno di nota è quello pubblicato da Hamill e Gilbert [3]. In primo luogo, gli autori stabiliscono quali sono le caratteristiche che una rete sociale simulata dovrebbe avere, tra cui:

Bassa densità di rete La densità di una rete [3] (*network density*) è definita come il rapporto tra il numero di archi esistenti e il numero massimo di archi possibili. Un utente medio è collegato con un numero di utenti dell’ordine delle centinaia o poche migliaia, numero che, se confrontato con le centinaia di milioni di utenti (se non miliardi) di tutto il sistema, risulta essere piuttosto basso.

Assortatività positiva Con questo termine, gli autori indicano la tendenza dei nodi con più connessioni ad essere collegati con altri nodi molto connessi.

Presenza di comunità Ovvero, la tendenza a formare *clusters*, gruppi di nodi fortemente connessi tra di loro ma debolmente connessi con il resto del sistema. Viene introdotto, a questo proposito, il *coefficiente di clustering* (*clustering coefficient*, vedi sezione ??)

Lunghezza ridotta dei cammini Un cammino (*path*) [4] tra i nodi v_0 e v_n è una lista ordinata di archi $P = \{(v_0, v_1), (v_1, v_2), \dots, (v_{n-1}, v_n)\}$, e n corrisponde alla lunghezza di questo cammino. Secondo gli autori, in media si può raggiungere un utente di un social network partendo da un qualsiasi altro nodo compiendo solo pochi passi, ovvero percorrendo un cammino ridotto. La lunghezza dipende dalle proporzioni della rete.

Vengono esposti inoltre diversi tipi di rete, emersi nel corso degli anni in letteratura:

Regular lattice Ogni nodo è collegato ad un numero fisso di suoi vicini

Random network Ogni nodo è collegato in media ad un certo numero di altri nodi

Small world network Basato sul modello *regular lattice*, aggiunge o riarrangia collegamenti in modo casuale

Scale-free network Descritta per la prima volta da Barabási & Bonabeau [5], prevede che pochi nodi abbiano molti collegamenti.

Gli autori indicano la costruzione *scale-free* come la migliore tra le quattro, in quanto presenta tutte le caratteristiche elencate in precedenza, con l'eccezione dell'*assortatività*, non particolarmente riflessa nel modello. Passano quindi all'esposizione della loro proposta, un modello ad agenti basato sul concetto di *social circles* [3]: ogni agente può stabilire un "link" con un altro agente solo se quest'ultimo può fare altrettanto. Quest'idea di *reciprocità* si adatta bene alla modellazione di alcune piattaforme (es. Facebook), mentre si adatta meno su altre: un esempio è Twitter, dove per stabilire un collegamento, il fatto che due utenti si conoscano direttamente è poco rilevante.

Un altro lavoro che abbiamo recuperato è quello di Liben-Nowell & Kleinberg [7]. Il problema da loro affrontato è quello della previsione di collegamenti (*link prediction*) nei social network, ovvero prevedere, dato lo stato del sistema in un dato istante di tempo, la formazione di nuovi collegamenti negli istanti immediatamente successivi. Un aspetto fondamentale in questo caso è il criterio con cui viene valutata la *similarità* tra i nodi del grafo, ovvero come viene assegnato uno *score* a una coppia di nodi (arco) in modo che ne misuri la "distanza" rispetto ad una particolare proprietà o caratteristica. Nodi simili ma non ancora collegati avranno infatti un'alta probabilità di stabilire un collegamento negli istanti di tempo più prossimi.

Nell'articolo viene fatta una rassegna dei metodi principali; l'obiettivo di questi metodi è costruire la matrice di similarità :

Neighborhood-based methods Di questa categoria fanno parte tutti quei metodi che si basano sul *vicinato* (*neighborhood*) dei nodi di cui si vuole calcolare la similarità. La misura di distanza può essere il numero di nodi adiacenti in comune, oppure la probabilità che i due nodi in esame abbiano un vicino in comune, per esempio usando la *distanza di Jaccard*. In alternativa, questa probabilità può essere calcolata in modo proporzionale alla dimensione dei vicinati dei due nodi candidati, per esempio moltiplicandone la cardinalità.

Paths-based methods Anche la distanza intesa come lunghezza di un cammino tra due nodi può essere intesa come misura di similarità; per esempio, in [8] viene descritta una metrica che prende in considerazione la lunghezza tutti i cammini esistenti tra due nodi per quantificarne la similarità. Anche alcuni celebri algoritmi, come il *PageRank* di Google [9], fanno parte di questa categoria.

Altre tecniche possono essere usate in congiunzione con i metodi sopra elencati, per semplificare la computazione oppure per irrobustirne la previsione. Procedure di *clustering* o *matrix factorization* possono aiutare ad eliminare i collegamenti meno significativi prima dell'effettivo calcolo delle metriche.

Essendo il nostro obiettivo la simulazione di una *social network*, non si può ignorare la (forte) componente sociale del sistema considerato. Ogni nodo rappresenta un utente, e dunque ne eredita le caratteristiche personali come interessi e carattere. La letteratura specializzata in *social sciences* ha da tempo definito il concetto di *omofilia* (*homophily*) [10], inteso come la tendenza di ogni individuo a stringere legami con altri individui dalle caratteristiche simili. Nel computo della misura di similarità (o dissimilarità) bisognerebbe quindi tenere conto di tali caratteristiche e trovare quindi il modo di rappresentarle nel sistema.

3 Descrizione modello

4 Visione multi-agente

5 Accenni teoria dei grafi

6 Esperimenti

7 Conclusione

8 Sviluppi futuri

8.1 Fabrication

Explain a step-by-step recipe for fabrication here. How long did you etch and why? What is an Ohmic contact?

8.2 Experimental set-up

Explain the experimental set-up here. Use a schematic picture (make it yourself in photoshop, paint, ...) to show how the components are connected. Briefly explain how a lock-in amplifier works.

9 Results and interpretation 2-3 pages

Show a graph of the longitudinal resistivity (ρ_{xx}) and Hall resistivity (ρ_{xy}) versus magnetic field, extracted from the raw data shown in figure ???. You will have the link to the data in your absalon messages, if not e-mail Guen (guen@nbi.dk). Explain how you calculated these values, and refer to the theory.

9.1 Classical regime

Calculate the sheet electron density n_s and electron mobility μ from the data in the low-field regime, and refer to the theory in section ??. Explain how you retrieved the values from the

data (did you use a linear fit?). Round values off to 1 or 2 significant digits: $8.1643 = 8.2$. Also, $5\text{e-}6$ is easier to read than 0.000005 .

!OBS: This part is optional (only if you have time left). Calculate the uncertainty as follows:

$u(f(x, y, z)) = \sqrt{(\frac{\delta f}{\delta x}u(x))^2 + (\frac{\delta f}{\delta y}u(y))^2 + (\frac{\delta f}{\delta z}u(z))^2}$, where f is the calculated value (n_s or μ), x, y, z are the variables taken from the measurement and $u(x)$ is the uncertainty in x (and so on).

9.2 Quantum regime

Calculate n_s for the high-field regime. Show a graph of the longitudinal conductivity (ρ_{xx}) and Hall conductivity (ρ_{xy}) **in units of the resistance quantum** ($\frac{h}{e^2}$), depicting the integer filling factors for each plateau. Show a graph of the plateau number versus its corresponding value of $1/B$. From this you can determine the slope, which you use to calculate the electron density. Again, calculate the uncertainty for your obtained values.

10 Discussion 1/2-1 page

Discuss your results. Compare the two values of n_s that you've found in the previous section. Compare your results with literature and comment on the difference. If you didn't know the value of the resistance quantum, would you be able to deduce it from your measurements? If yes/no, why?

Item	Quantity
Widgets	42
Gadgets	13

Tabella 1: An example table.

11 Some LaTeX tips

11.1 How to Include Figures

First you have to upload the image file (JPEG, PNG or PDF) from your computer to writeLaTeX using the upload link the project menu. Then use the `includegraphics` command to include it in your document. Use the `figure` environment and the `caption` command to add a number and a caption to your figure. See the code for Figure ?? in this section for an example.

11.2 How to Make Tables

Use the `table` and `tabular` commands for basic tables — see Table 1, for example.

11.3 How to Write Mathematics

L^AT_EX is great at typesetting mathematics. Let X_1, X_2, \dots, X_n be a sequence of independent and identically distributed random variables with $E[X_i] = \mu$ and $\text{Var}[X_i] = \sigma^2 < \infty$, and let

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_i^n X_i \quad (1)$$

denote their mean. Then as n approaches infinity, the random variables $\sqrt{n}(S_n - \mu)$ converge in distribution to a normal $\mathcal{N}(0, \sigma^2)$.

The equation 1 is very nice.

11.4 How to Make Sections and Subsections

Use section and subsection commands to organize your document. L^AT_EX handles all the formatting and numbering automatically. Use `ref` and `label` commands for cross-references.

11.5 How to Make Lists

You can make lists with automatic numbering ...

1. Like this,
2. and like this.

...or bullet points ...

- Like this,
- and like this.

...or with words and descriptions ...

Word Definition

Concept Explanation

Idea Text

We hope you find write \LaTeX useful, and please let us know if you have any feedback using the help menu above.

Riferimenti bibliografici

- [1] Welch, Chris (2017). "Facebook crosses 2 billion monthly users". *The Verge. Vox Media*. Retrieved June 27, 2017.
- [2] Bulman, May (2016). "Donald Trump's 'celebrity-style' tweets helped him win US presidential election, says data scientist". *The Independent* Retrieved November 28, 2016.
- [3] Hamill, L., & Gilbert, N. (2010). Simulating large social networks in agent-based models: A social circle model. *Emergence: Complexity and Organization*, 12(4), 78.
- [4] Barabási, A. L. (2016). *Network science*. Cambridge university press.
- [5] Barabási, B. A. L., & Bonabeau, E. (2003). Scale-free. *Scientific American*, 288(5), 50-59.
- [6] Ferber, J. (1999). *Multi-agent systems: an introduction to distributed artificial intelligence* (Vol. 1). Reading: Addison-Wesley. ISO 690
- [7] Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. *journal of the Association for Information Science and Technology*, 58(7), 1019-1031.
- [8] Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 18(1), 39-43.
- [9] Brin, S., & Page, L. (2012). Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer networks*, 56(18), 3825-3833.
- [10] McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1), 415-444.