TEORIA DEI SISTEMI A RETE

Giovanni Carrosio

gcarrosio@units.it

Anno accademico 2017-2018

Lezione 5 - 15 marzo 2108

The Small-World Problem

By Stanley Milgram

Fred Jones of Peoria, sitting in a sidewalk cafe in Tunis, and needing a light for his cigarette, asks the man at the next table for a match. They fall into conversation; the stranger is an Englishman who, it turns out, spent several months in Detroit studying the operation of an interchangeable-bottlecap-factory. "I know it's a foolish question," says Jones, "but did you ever by any chance run into a fellow named Ben Arkadian? He's an old friend of mine, manages a chain of supermarkets in Detroit . . ."

"Arkadian, Arkadian," the Englishman mutters. "Why, upon my soul, I believe I do! Small chap, very energetic, raised merry hell with the factory over a shipment of defective bottlecaps."

"No kidding!" Jones exclaims in amazement.

"Good lord, it's a small world, isn't it?'

TEORIA DEI SEI GRADI DI SEPARAZIONE

1967 - The Small-World Problem, Psychology Today, (May), pp. 61-67

1969 - An Experimental Study of the Small World Problem, Sociometry, Vol. 32, No. 4 (Dec), pp. 425-443

THE SMALL WORLD PROBLEM IN SINTESI

- A partire da due qualunque persone nel mondo, qual è la probabilità che si conoscano?
- Dal momento che le persone X e Z possono non conoscersi direttamente, possono condividere un conoscente comune, ovvero una persona che li conosce entrambi.
- Si può allora pensare ad una catena di conoscenze, con X che conosce Y e Y che conosce Z.
- Si può anche immaginare che X è legato a Z non attraverso un singolo legame, ma da una serie di legami, X-a-b-c-d….y-Z.
- Quindi, date due persone nel mondo X e Z, quanti legami di conoscenza sono necessari perché X si connetta con Z?

The underlying structure

Metodo sperimentale per tracciare una linea di conoscenza che lega due persone scelte a caso

A la persona da cui è iniziata la catena

Z la persona target da raggiungere

Scegliere una persona A reale tra i 200 milioni di abitanti negli USA e una persona Z reale

A deve inviare un messaggio raggiungendo Z utilizzando una catena di amici e conoscenti. Ogni nodo deve scegliere un conoscente che potrebbe essere più idoneo per arrivare a Z.

A da due città: Wichita (Kansas) e Omaha (Nebraska)

B: Cambridge (wife of a divinity school student) e Sharon, Massachussets, (stockbroker who worked in Boston)

Primo studio: Kansas Study

Secondo studio: Nebraska Study

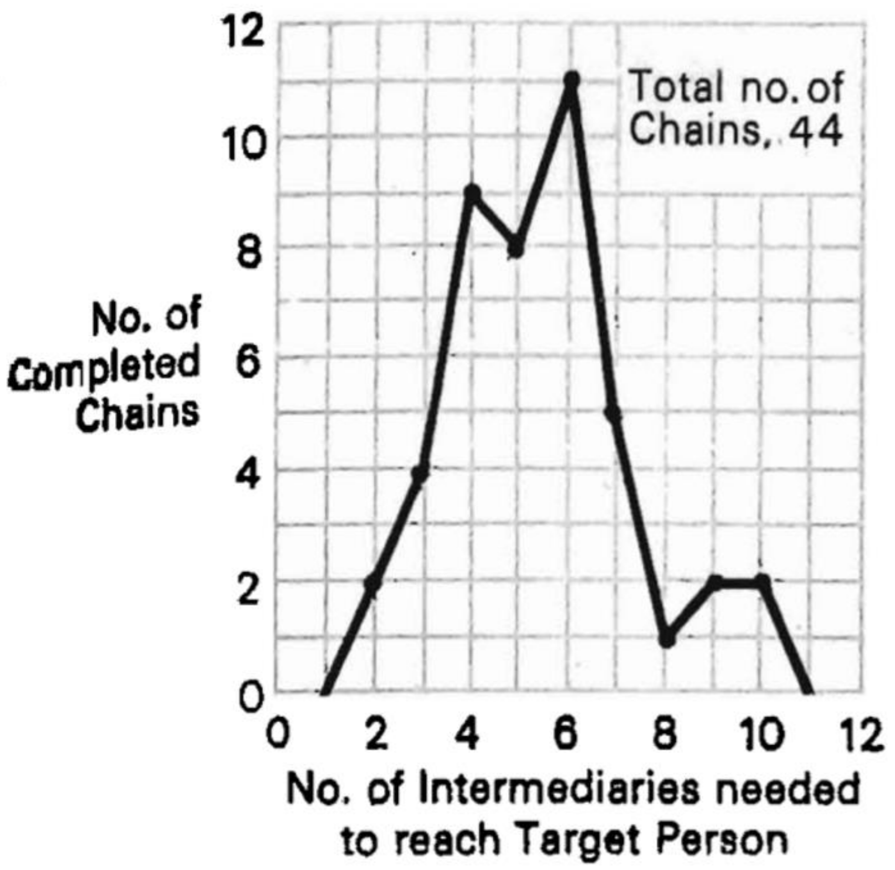
Il documento contiene: 1. il nome e qualche informazione di B 2. le regole del gioco 3. la lista sulla quale ogni persona della catena scrive il suo nome 4. 15 cartoline che ogni nodo della catena deve inviare a Milgram

Di 160 catene iniziate in Nebraska, 44 sono state completate, mentre 126 non sono arrivate a Z.

Distorsione di genere nella scelta del successivo nodo della catena.

In Kansas, 123 hanno inviato a conoscenti e soltanto 22 a parenti.

6 è la moda!!!



I GRADI SEPERAZIONE DI MESSENGER

Un nuovo esperimento simile a quello di Milgram è stato condotto nel 2008 da Leskovec and Horvitz

Utilizzati i dati relativi alle connessioni su Messenger di 240 milioni di utenti durante un periodo di 30 giorni

Hanno individuato una componente gigante con distanza media di 6.6

Planetary-Scale Views on an Instant-Messaging Network*

Jure Leskovec†
Machine Learning Department
Carnegie Mellon University
Pittsburgh, PA, USA

Eric Horvitz
Microsoft Research
Redmond, WA, USA

Microsoft Research Technical Report
MSR-TR-2006-186

June 2007

Abstract

We present a study of anonymized data capturing a month of high-level communication activities within the whole of the Microsoft Messenger instant-messaging system. We examine characteristics and patterns that emerge from the collective dynamics of large numbers of people, rather than the actions and characteristics of individuals. The dataset contains summary properties of 30 billion conversations among 240 million people. From the data, we construct a communication graph with 180 million nodes and 1.3 billion undirected edges, creating the largest social network constructed and analyzed to date. We report on multiple aspects of the dataset and synthesized graph. We find that the graph is well-connected and robust to node removal. We investigate on a planetary-scale the oft-cited report that people are separated by "six degrees of separation" and find that the average path length among Messenger users is 6.6. We also find that people tend to communicate more with each other when they have similar age, language, and location, and that cross-gender conversations are both more frequent and of longer duration than conversations with the same gender.

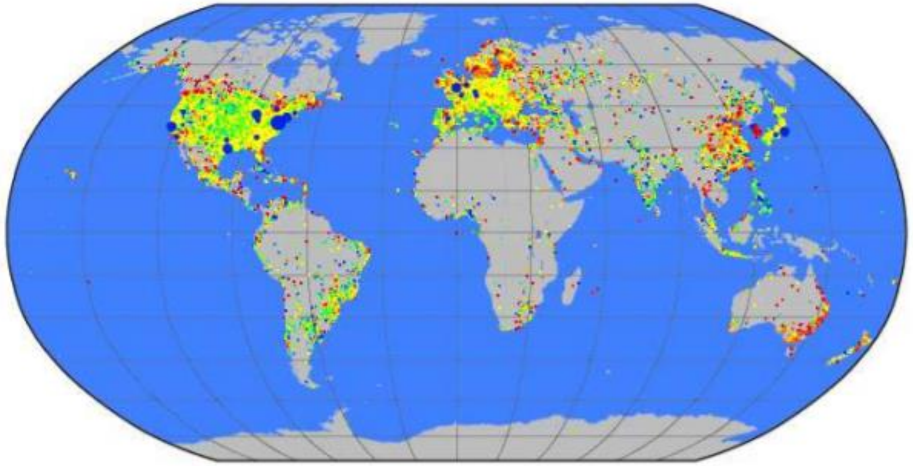


Figure 11: Number of users at particular geographic location superimposed on the map of the world. Color represents the number of users.

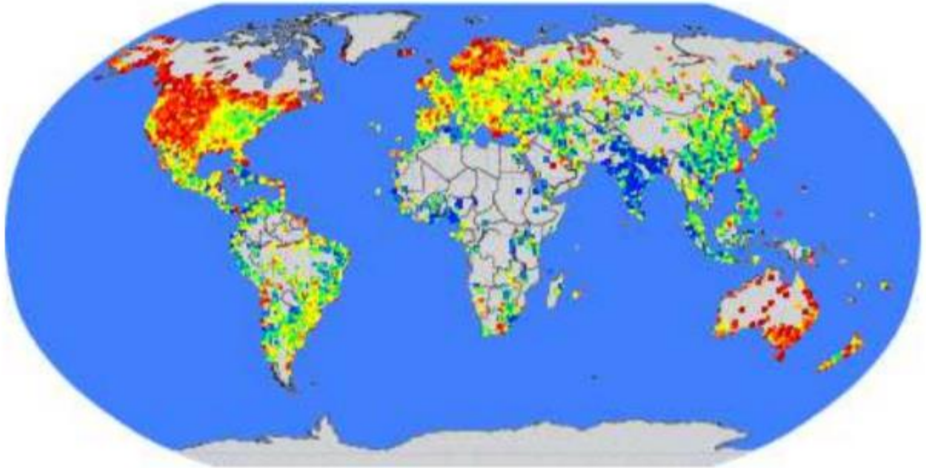


Figure 12: Number of Messenger users per capita. Color intensity corresponds to the number of users per capita in the cell of the grid.

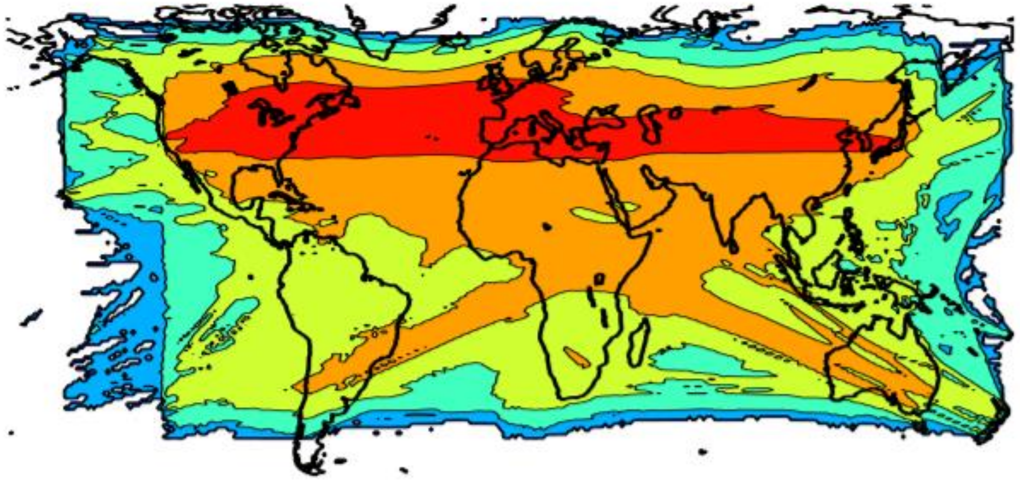


Figure 13: A communication heat map.

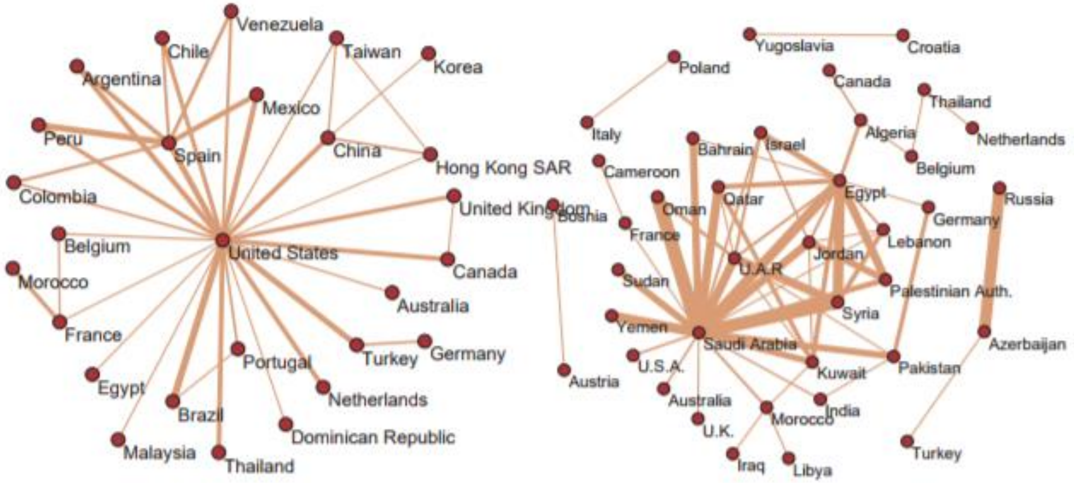


Figure 14: (a) Communication among countries with at least 10 million conversations in June 2006. (b) Countries by average length of the conversation. Edge widths correspond to logarithms of intensity of links.

7.1 How small is the small world?

Messenger data gives us a unique opportunity to study distances in the social network. To our knowledge, this is the first time a planetary-scale social network has been available to validate the well-known "6 degrees of separation" finding by Travers and Milgram (17). The earlier work employed a sample of 64 people and found that the average number of hops for a letter to travel from Nebraska to Boston was 6.2 (mode 5, median 5), which is popularly known as the "6 degrees of separation" among people. We used a population sample that is more than two million times larger than the group studied earlier and confirmed the classic finding.

Figure 20(a) displays the distribution over the shortest path lengths. To approximate the distribution of the distances, we randomly sampled 1000 nodes and calculated for each node the shortest paths to all other nodes. We found that the distribution of path lengths reaches the mode at 6 hops and has a median at 7. The average path length is 6.6. This result means that a random pair of nodes in the Messenger network is 6.6 hops apart on the average, which is half a link longer than the length measured by Travers and Milgram. The 90th percentile (effective diameter (16)) of the distribution is 7.8. 48% of nodes can be reached within 6 hops and 78% within 7 hops. So, we might say that, via the lens provided on the world by Messenger, we find that there are about "7 degrees of separation" among people. We note that long paths, *i.e.*, nodes that are far apart, exist in the network; we found paths up to a length of 29.

4 GRADI DI SEPARAZIONE

The first world-scale social-network graph-distance computations, using the entire Facebook network of active users (721 million users and 69 billion friendship links)

3,74 intermediaries, showing that the world is even smaller than we expected

Four Degrees of Separation

Lars Backstrom* Paolo Boldi† Marco Rosa† Johan Ugander* Sebastiano Vigna†

January 6, 2012

Abstract

Frigyes Karinthy, in his 1929 short story "Láncszemek" ("Chains") suggested that any two persons are distanced by at most six friendship links.[1] Stanley Milgram in his famous experiment [20, 23] challenged people to route postcards to a fixed recipient by passing them only through direct acquaintances. The average number of intermediaries on the path of the postcards lay between 4.4 and 5.7, depending on the sample of people chosen.

We report the results of the first world-scale social-network graph-distance computations, using the entire Facebook network of active users (≈ 721 million users, ≈ 69 billion friendship links). The average distance we observe is 4.74, corresponding to 3.74 intermediaries or "degrees of separation", showing that the world is even smaller than we expected, and prompting the title of this paper. More generally, we study the distance distribution of Facebook and of some interesting geographic subgraphs, looking also at their evolution over time.

The networks we are able to explore are almost two orders of magnitude larger than those analysed in the previous literature. We report detailed statistical metadata showing that our measurements (which rely on probabilistic algorithms) are very accurate.

1 Introduction

At the 20th World–Wide Web Conference, in Hyderabad, India, one of the authors (Sebastiano) presented a new tool for studying the distance distribution of very large graphs: HyperANF [3]. Building on previous graph compression [4] work and on the idea of diffusive computation pioneered in [21], the new tool made it possible to accurately study the distance distribution of graphs orders of magnitude larger than it was previously possible.

One of the goals in studying the distance distribution is the identification of interesting statistical parameters that can be used to tell proper social networks from other complex networks, such as web graphs. More generally, the distance distribution is one interesting *global* feature that makes it possible to reject probabilistic models even when they match local features such as the in-degree distribution.

In particular, earlier work had shown that the *spid*[2], which measures the *dispersion* of the distance distribution, appeared to be smaller than 1 (underdispersion) for social networks, but larger than one (overdispersion) for web graphs [3]. Hence, during the talk, one of the main open questions was "What is the spid of Facebook?".

Lars Backstrom happened to listen to the talk, and suggested a collaboration studying the Facebook graph. This was of course an extremely intriguing possibility: beside testing the "spid hypothesis", computing the distance distribution of the Facebook graph would have been the largest Milgram-like [20] experiment ever performed, orders of magnitudes larger than previous attempts (during our experiments Facebook has ≈ 721 million active users and ≈ 69 billion friendship links).

This paper reports our findings in studying the distance distribution of the largest electronic social network ever created. That world is smaller than we thought: the average distance of the current Facebook graph is 4.74. Moreover, the

*Facebook.

LETTURE

L'internet dell'energia, Le Scienze, dicembre 2017

If Osama's only six degrees away, why can't we find him?, Discover Magazine, febbraio 2008

VIDEO

Connected. The Power of Six Degrees

LA PROSSIMA LEZIONE

Trovare altri esperimenti sui 6 gradi di separazione

Come è nata la società in rete