A LINE GRAPH as a model of a social network

Małgorzata Krawczyk, Lev Muchnik, Anna Mańska-Krasoń, Krzysztof Kułakowski

AGH Kraków
Stern School of Business of NY University
outline

- ideas, definitions, milestones
- line graphs
- LiveJournal data

Floor-level networking is often staged to look like friendly socializing. Nevertheless, it has an instrumental goal: useful contacts. Every networker knows this, so the instrumentality is not latent. It is not completely manifest, either. I presume that there is a mutual implicit contract of not breeching the situation by stating the instrumental goal aloud. The getting of contacts is a conscious but tacit function of socializing.

Juha Klemelä, Turku University

Managing Mixed Emotions in the Layered Ritual Reality of Networking Events, XVII ISA World Congress of Sociology, Gothenburg, 11-17 July 2010
b. 1971
B.Sc. in physics
Ph.D. in theoretical and applied mechanics

b. 1959
B. Sc. in mathematics
PhD in applied mathematics

D J Watts, S H Strogatz, Collective dynamics of 'small-world' networks,
Connectivity matrix („sociomatrix“)

\[
\begin{pmatrix}
0 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0
\end{pmatrix}
\]

Clustering coefficient

\[
C = \frac{1}{N} \sum_{i=1}^{N} \frac{2L_i}{k_i(k_i - 1)}
\]
Mean free path („diameter“)

\[ L(N) = \frac{1}{N(N-1)} \sum_{i,j=1}^{N} \text{shortest\_path}(i, j) \]

Small-world effect \( L(N) \propto \ln(N) \)

Erdős-Rényi networks

- Take $N$ nodes
- Connect each pair with probability $p$

Mean degree

$$<k> = (N-1)p$$

Degree distribution

$$P(k) = e^{-<k>} \frac{<k>^k}{k!}$$
b. 1967
MSc in physics and engineering
PhD in physics

PhD in complex networks 2001

Growing networks – construction for $M = 2$

Nodes to attach a new node are selected with probability:
- constant (exponential graphs)
- proportional to their degree (scale-free graphs)
<table>
<thead>
<tr>
<th>networks</th>
<th>N</th>
<th>#</th>
<th>d</th>
<th>C</th>
<th>γ</th>
</tr>
</thead>
<tbody>
<tr>
<td>actors</td>
<td>45x10^4</td>
<td>25x10^6</td>
<td>3.48</td>
<td>0.78</td>
<td>2.3</td>
</tr>
<tr>
<td>www Altavista</td>
<td>2x10^9</td>
<td>2x10^10</td>
<td>16.18</td>
<td>-</td>
<td>2.1/2.7</td>
</tr>
<tr>
<td>coauthorship - math</td>
<td>25x10^4</td>
<td>50x10^4</td>
<td>7.57</td>
<td>0.34</td>
<td>-</td>
</tr>
<tr>
<td>coauthorship - phys</td>
<td>5^10^4</td>
<td>25x10^4</td>
<td>6.19</td>
<td>0.56</td>
<td>-</td>
</tr>
<tr>
<td>Phone calls</td>
<td>47x10^6</td>
<td>8x10^7</td>
<td>-</td>
<td>-</td>
<td>2.1</td>
</tr>
<tr>
<td>Web chains in water</td>
<td>92</td>
<td>997</td>
<td>1.9</td>
<td>0.087</td>
<td>-</td>
</tr>
<tr>
<td>Protein interactions</td>
<td>2115</td>
<td>2240</td>
<td>6.8</td>
<td>0.071</td>
<td>2.4</td>
</tr>
<tr>
<td>Sexual contacts</td>
<td>2810</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.2</td>
</tr>
<tr>
<td>Words in sentences</td>
<td>46x10^4</td>
<td>17x10^6</td>
<td>-</td>
<td>0.44</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Assortativity

\[ r = \frac{\sum_{jk} (e_{jk} - q_j q_k)}{\sigma_q^2} \]

where

\[ q_k = \frac{(k + 1) p_{k+1}}{<k>} \]

\( e_{jk} \) – fraction of links between nodes \( j,k \)

<table>
<thead>
<tr>
<th>network</th>
<th>type</th>
<th>size</th>
<th>assortativity ( r )</th>
<th>error ( \sigma_r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>physics coauthorship</td>
<td>undirected</td>
<td>52909</td>
<td>0.363</td>
<td>0.002</td>
</tr>
<tr>
<td>biology coauthorship</td>
<td>undirected</td>
<td>1520251</td>
<td>0.127</td>
<td>0.0004</td>
</tr>
<tr>
<td>mathematics coauthorship</td>
<td>undirected</td>
<td>253339</td>
<td>0.120</td>
<td>0.002</td>
</tr>
<tr>
<td>film actor collaborations</td>
<td>undirected</td>
<td>449913</td>
<td>0.208</td>
<td>0.0002</td>
</tr>
<tr>
<td>company directors</td>
<td>undirected</td>
<td>7673</td>
<td>0.276</td>
<td>0.004</td>
</tr>
<tr>
<td>student relationships</td>
<td>undirected</td>
<td>573</td>
<td>-0.029</td>
<td>0.037</td>
</tr>
<tr>
<td>email address books</td>
<td>directed</td>
<td>16881</td>
<td>0.092</td>
<td>0.004</td>
</tr>
<tr>
<td>power grid</td>
<td>undirected</td>
<td>4941</td>
<td>-0.003</td>
<td>0.013</td>
</tr>
<tr>
<td>Internet</td>
<td>undirected</td>
<td>10697</td>
<td>-0.189</td>
<td>0.002</td>
</tr>
<tr>
<td>World-Wide Web</td>
<td>directed</td>
<td>269504</td>
<td>-0.067</td>
<td>0.0002</td>
</tr>
<tr>
<td>software dependencies</td>
<td>directed</td>
<td>3162</td>
<td>-0.016</td>
<td>0.020</td>
</tr>
<tr>
<td>protein interactions</td>
<td>undirected</td>
<td>2115</td>
<td>-0.156</td>
<td>0.010</td>
</tr>
<tr>
<td>metabolic network</td>
<td>undirected</td>
<td>765</td>
<td>-0.240</td>
<td>0.007</td>
</tr>
<tr>
<td>neural network</td>
<td>directed</td>
<td>307</td>
<td>-0.226</td>
<td>0.016</td>
</tr>
<tr>
<td>marine food web</td>
<td>directed</td>
<td>134</td>
<td>-0.263</td>
<td>0.037</td>
</tr>
<tr>
<td>freshwater food web</td>
<td>directed</td>
<td>92</td>
<td>-0.326</td>
<td>0.031</td>
</tr>
</tbody>
</table>

MEJ Newman, PRE 67 (2003) 026126
Why not ALL social networks are scale-free
Why social networks are different from other types of networks

In this paper we have argued that social and non-social networks differ in two important ways. First, they show distinctly different patterns of correlation between the degrees of adjacent vertices, with degrees being positively correlated (assortative mixing) in most social networks and negatively correlated (disassortative mixing) in most non-social networks. Second, social networks show high levels of clustering or network transitivity, whereas clustering in many non-social networks is no higher than one would expect on the basis of pure chance, given the observed degree distribution.

We have shown that both of these differences can be explained by the same hypothesis, that social networks are divided into communities, and non-social networks are not.
It is used to assume but it could be.
In the model proposed here, a social network is the line graph of an initial network of families, communities, interest groups, school classes and small companies. These groups play the role of nodes, and individuals are represented by links between these nodes.

M J Krawczyk et al, arXiv:1010.2460
line graph – the construction

(a) [Diagram of a graph]

(b) [Diagram of a graph]

(c) [Diagram of a complex graph]

(d) [Diagram of a simple graph]
Algorithm

1. assign numbers to the elements of the connectivity matrix above the diagonal. Make the matrix symmetric.

\[
\begin{array}{cccc}
0 & 1 & 0 & 0 \\
1 & 0 & 2 & 3 \\
0 & 2 & 0 & 4 \\
0 & 3 & 4 & 0 \\
\end{array}
\]

2. If \(i,j\) are in the same row or column, then the element \(C(i,j)\) of the transformed matrix is 1

\[
\begin{array}{cccc}
0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 \\
\end{array}
\]
Degree distribution of a line graph

\[ P_t(k) = \frac{\sum_{n=1}^{\infty} nP(n) \sum_{m=1}^{\infty} mP(m) \delta_{k,m+n-2}}{\sum_{n=1}^{\infty} nP(n) \sum_{m=1}^{\infty} mP(m)} = \frac{1}{\lambda^2} \sum_{n=1}^{k+1} nP(n)(k-n+2)P(k-n+2) \]

Degree distribution $P(k)$

Erdős-Rényi networks

\[ P_t(k) = \lambda^k e^{-2\lambda} \sum_{n=1}^{k+1} \frac{1}{(n-1)!(k-n+1)!} = \]
\[ = e^{-2\lambda} \frac{(2\lambda)^k}{k!} \]

exponential networks

\[ P_t(k) = \frac{(1-c)^4}{6} (k+1)(k+2)(k+3)c^k \]
The degree distribution of a line graph on a scale-free network

\[ P_t(k) \]

\[ \langle k \rangle = 6,16 \]
Clustering coefficient in line graphs

\[ C = \frac{1}{\lambda^2} \sum_{n=1}^{\infty} n P(n) \sum_{m=1}^{\infty} m P(m) \frac{(n-1)(n-2) + (m-1)(m-2)}{(n+m-2)(n+m-3)} \]
Line graphs - assortativity

\[
<k'(k) >= \frac{\sum_n nP(n) \sum_m mP(m) \sum_r rP(r)(n + m - 2)\delta_{k,m+r-2}}{\sum_n nP(n) \sum_m mP(m) \sum_r rP(r)\delta_{k,m+r-2}}
\]

Assortativity $<k'(k)>$

Erdős-Rényi networks

$$< k'(k) > = \lambda + \frac{2^{k-1}k}{2^k - 1} \approx \lambda + \frac{k}{2}$$

exponential networks
Assortativity of a line graph on a scale-free network

\[ \langle k'(k) \rangle \]
LiveJournal

LiveJournal is a remarkably popular platform for personal blog management, populated with over 8 million blogs and over 1 million of communities. LiveJournal was among the first of such platforms available online and it still remains one of the most active and popular. Its users manage personal blogs where they share their daily experiences, political views or discuss news events. Users can also comment on posts of other users.

We defined the network nodes to correspond to personal blogs. Directional links connecting these nodes represent the record that a particular user (owning one blog) monitors another blog (owned by another user).
LiveJournal: data
\[ N = 8.1 \times 10^6; \ # = 125 \times 10^6 \]

\[ P(k) \]

Simulations: \[ N = 9 \times 10^3 \]
LiveJournal:

assortativity

$\langle k'(k) \rangle$

$\langle k_{nn} \rangle$ - average neighbour's degree

$\langle k \rangle$, degree

$\langle k_{nn} \rangle$, average neighbour's degree

K in

K out

simulations
LiveJournal

$C(k)$

$k$

simulations
conclusions

The degree distribution of a line graph is close to the degree distribution of its initial network. Line graphs are clustered and assortative. The degree-dependent clustering coefficient $C(k)$ indicates the presence of cliques.

We have shown that LiveJournal, where $P(k)$ is scale-free, displays qualitatively the same features.

Thank you