convex skeleton: generalization of network spanning tree

Lovro Šubelj
University of Ljubljana
Faculty of Computer and Information Science
Ljubljana, Slovenia

LaRichNet ’18
convex subgraphs of networks

**convex/non-convex** real functions, sets in $\mathbb{R}^2$ & subgraphs

Disconnected $\supseteq$ connected $\supseteq$ **induced** $\supseteq$ isometric $\supseteq$ **convex** subgraphs

**convex hull** $\mathcal{H}(S)$ is smallest convex subgraph including $S$.

**subset** $S$ is convex if it induces convex **subgraph**
expansion of convex subgraphs

grow subgraph $S$ by one node & expand $S$ to convex hull $\mathcal{H}(S)$

- $S = \{\text{random node } i\}$
- until $S$ contains $n$ nodes:
  1. select $i \notin S$ by random edge
  2. expand $S = \mathcal{H}(S \cup \{i\})$

$S$ quantifies (locally) **tree-like/clique-like** structure of networks
examples of convex expansion

\[ s(t) = \text{average fraction of nodes in } S \text{ after } t \text{ expansion steps} \]

\[ s(t) \approx \frac{t + 1}{n} \text{ in convex & } s(t) \gg \frac{t + 1}{n} \text{ in non-convex networks} \]

\( s(t) \) quantifies (locally) tree-like/clique-like structure of networks
measure of network convexity

\[ X_s = s - \sum_{t=1}^{sn-1} \max(s \Delta s(t) - 1/n, 0) \]

\[ s = \text{fraction of nodes in LCC} \]

\( X_s \) highlights tree-like/clique-like networks & synthetic graphs

\begin{tabular}{|l|l|l|l|l|l|l|}
\hline
          & \( n \) & \( \langle k \rangle \) & \( X_s \) &          & \( n \) & \( \langle k \rangle \) & \( X_s \) \\
\hline
Jazz musicians & 198 & 27.70 & 0.12 &          & 2500 & 10.00 & 0.00 \\
Network scientists & 379 & 4.82 & 0.85 & Random graphs & 1000 & 10.00 & 0.01 \\
Computer scientists & 239 & 4.75 & 0.64 &          & 225 & 10.00 & 0.03 \\
\textit{Plasmodium falciparum} & 1158 & 4.15 & 0.43 & Triangular lattice & 225 & 5.48 & 0.23 \\
\textit{Saccharomyces cerevisiae} & 1458 & 2.67 & 0.68 & Rectangular lattice & 225 & 3.73 & 0.13 \\
\textit{Caenorhabditis elegans} & 3747 & 4.14 & 0.56 & Core-periphery graph & 3747 & 4.48 & 0.39 \\
\hline
AS (January 1, 1998) & 3213 & 3.50 & 0.66 & Trees of cliques & 2500 & 5.97 & 1.00 \\
AS (January 1, 1999) & 531 & 4.58 & 0.49 &          & 1000 & 5.97 & 1.00 \\
AS (January 1, 2000) & 3570 & 3.94 & 0.59 &          & 225 & 6.01 & 1.00 \\
\hline
Little Rock Lake & 183 & 26.60 & 0.02 &          &          &          &        \\
Florida Bay (wet) & 128 & 32.42 & 0.03 &          &          &          &        \\
Florida Bay (dry) & 128 & 32.91 & 0.03 &          &          &          &        \\
\hline
\end{tabular}

\( X_s \) measures global & regional convexity in (disconnected) networks
convex skeletons of networks

convex skeleton = largest high-$X$ subnetwork (every $S$ convex)

**spanning tree** & **convex skeleton** of network scientists coauthorships

convex skeleton is **tree** of **cliques** extracted by edge removal
statistics of convex skeletons

\[ \langle C \rangle = \frac{1}{n} \sum_{i} \frac{2t_i}{k_i(k_i - 1)} \quad \langle \sigma \rangle = \frac{2}{n(n-1)} \sum_{i<j} \sigma_{ij} \quad X_s = \ldots \]

statistics of **convex skeletons** & **spanning trees** of networks

<table>
<thead>
<tr>
<th></th>
<th>clustering ( \langle C \rangle )</th>
<th>geodesics ( \langle \sigma \rangle )</th>
<th>convexity ( X_s )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>CS</td>
<td>ST</td>
</tr>
<tr>
<td>Jazz musicians</td>
<td>0.62</td>
<td>0.81</td>
<td>0.00</td>
</tr>
<tr>
<td>Network scientists</td>
<td>0.74</td>
<td>0.75</td>
<td>0.00</td>
</tr>
<tr>
<td>Computer scientists</td>
<td>0.48</td>
<td>0.54</td>
<td>0.00</td>
</tr>
<tr>
<td><em>Plasmodium falciparum</em></td>
<td>0.02</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em></td>
<td>0.07</td>
<td>0.10</td>
<td>0.00</td>
</tr>
<tr>
<td><em>Caenorhabditis elegans</em></td>
<td>0.06</td>
<td>0.12</td>
<td>0.00</td>
</tr>
<tr>
<td>AS (January 1, 1998)</td>
<td>0.18</td>
<td>0.21</td>
<td>0.00</td>
</tr>
<tr>
<td>AS (January 1, 1999)</td>
<td>0.18</td>
<td>0.27</td>
<td>0.00</td>
</tr>
<tr>
<td>AS (January 1, 2000)</td>
<td>0.20</td>
<td>0.25</td>
<td>0.00</td>
</tr>
<tr>
<td>Little Rock Lake</td>
<td>0.32</td>
<td>0.69</td>
<td>0.00</td>
</tr>
<tr>
<td>Florida Bay (wet)</td>
<td>0.33</td>
<td>0.79</td>
<td>0.00</td>
</tr>
<tr>
<td>Florida Bay (dry)</td>
<td>0.33</td>
<td>0.82</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**convex skeleton** is generalization of **spanning tree** retaining **clustering**

6/12
distributions of convex skeletons & spanning trees of networks

convex skeletons retain node distributions in contrast to spanning trees
position in convex skeletons

node position in **convex skeletons** & **spanning trees** of networks

**convex skeletons** retain node position in contrast to **spanning trees**
robustness of convex skeletons

MDS maps of convex skeletons, spanning trees & random graphs

networks allow robust extraction of convex skeletons & spanning trees
convex skeleton of coauthorships

convex skeleton $\sim$ network abstraction technique

convex skeleton of Slovenian computer scientists coauthorships

computer theory (◆), information systems (■), intelligent systems (▲), programming technologies (○) & other (●)
network backbones of coauthorships

convex skeleton $\gg$ high-betweenness & high-salience skeletons

properties of backbones of Slovenian computer scientists coauthorships

convex skeletons strengthen properties in contrast to other backbones
convex skeletons of networks

spanning tree

tree w/o cliques

convex skeleton

tree w/ cliques

**convex skeleton** $\Rightarrow$ network backbones

analysis, modeling, sampling, abstraction, visualization etc.
network convexity:

arXiv:1608.03402v3

convex skeletons:

arXiv:1709.00255v3

Lovro Šubelj
University of Ljubljana
lovro.subelj@fri.uni-lj.si
http://lovro.lpt.fri.uni-lj.si