Network curvature, and its implications for network management

Iraj Saniee, Math of Networks & Comm. Dept., Bell Labs, Alcatel-Lucent, Murray Hill, NJ
Joint work with Onuttom Narayan, Dept. of Physics, UCSC, CA

DIMACS Workshop, November 12-13, 2009

*Fundied by AFOSR grant FA9550-08-1-0064
Understanding Large-Scale Networks

- Hard to visualize due to scale
- Unclear what is essential and what is not for overall performance, reliability and security
- Need better ways to “summarize” critical network information
- How?

A promising direction is to look at two main geometric characteristics of objects: their dimension & curvature

Rocketfuel dataset 7018 (AT&T)
10152 nodes, 28638 links, diameter 12
Local Dimension and Graph Genus

How “planar” is a given network? Turns out all graphs are *locally planar*.

**Ringel-Youngs.** (“All graphs with \(N \geq 3\) nodes are locally 2-dimensional.”) For \(N \geq 3\), any \(G=(N,L)\) can be embedded in \(T_g\), a torus with \(g\) holes, where

\[
g \leq \left\lfloor \frac{(N - 3)(N - 4)}{12} \right\rfloor
\]

**Issue.** Even if computation of \(g\) were easy, R-Y wouldn’t help much because the advantage of local 2-D structure is obscured by impact of holes on the large-scale properties of graph. *How so?*
Our Data Source -- Rocketfuel

Look at scaling of the average shortest path length $<h>$. In 2-dim grid, $<h> \sim \sqrt{N}$ (or $\sim N^{1/\Delta}$ in $\Delta$-dim grid) but many real networks have slower scaling of $<h>$ than $\sqrt{N}$ -- e.g., think of the “small world” E-R random graphs.

- Look at “Rocketfuel” data, [Washington University researchers’ detailed connectivity data from various ISPs 2002-2003]
- Clearly, $<h>$, average SPs with respect to the hop metric, do not scale like $\sqrt{N}$
- Need to consider something else!

<table>
<thead>
<tr>
<th>Network ID</th>
<th>Network Name</th>
<th>Size node - #links</th>
<th>Average Shortest Path Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1221</td>
<td>Telstra (Australia)</td>
<td>2998 - 7612</td>
<td>5.53</td>
</tr>
<tr>
<td>1239</td>
<td>Sprintlink (US)</td>
<td>8341 - 28050</td>
<td>5.18</td>
</tr>
<tr>
<td>1755</td>
<td>EBONE (US)</td>
<td>605 - 2070</td>
<td>6.0</td>
</tr>
<tr>
<td>2914</td>
<td>Verio (US)</td>
<td>3045 - 10726</td>
<td>6.0</td>
</tr>
<tr>
<td>3257</td>
<td>Tiscali (EU)</td>
<td>855 - 2346</td>
<td>5.3</td>
</tr>
<tr>
<td>3356</td>
<td>Level 3 (US)</td>
<td>3447 - 18780</td>
<td>5.0</td>
</tr>
<tr>
<td>3967</td>
<td>Exodus (US)</td>
<td>895 - 4140</td>
<td>5.9</td>
</tr>
<tr>
<td>4755</td>
<td>VSNL (India)</td>
<td>121 - 456</td>
<td>3.2</td>
</tr>
<tr>
<td>6461</td>
<td>Abovenet (US)</td>
<td>2720 - 7648</td>
<td>5.7</td>
</tr>
<tr>
<td>7018</td>
<td>AT&amp;T (US)</td>
<td>10152 - 28638</td>
<td>6.9</td>
</tr>
</tbody>
</table>
Other Locally 2-Dimensional Embeddings: The Poincaré Disk $H^2$

Consider the unit disk $\{x \in \mathbb{R}^2; |x| < 1\}$ with length metric given by

$$ds^2 = \frac{dx^2 + dy^2}{(1 - x^2 - y^2)^2}$$

the hyperbolic metric.

**Advantages.** In the small scale it is 2-dimensional, but has slower scaling of geodesics (shortest paths) than $\mathbb{R}^2$.

**Relationship to graphs?** The Poincaré disk comes with numerous natural “scaffoldings” or tilings.
Scaffoldings of $H^2$: Hyperbolic Regular Graphs

Consider $X_{p,q}$ tilings (isometries) of $H^2$, that at each vertex consist of $q$ regular $p$-gons for integers $p \& q$ with $(p-2)(q-2)>4$ (flat with equality)

Examples:

- $X_{3,7}$
- $X_{4,5}$
- $X_{6,4}$

Note. Since networks of interest to us are typically finite, we’ll consider truncations of $X_{p,q}$, the part within a (large enough) radius $r$ from the center. Call this $TX_{p,q}$. 

6 DIMACS Workshop, Nov. 12-13, 2009
Iraj Sanlee
All Rights Reserved © Alcatel-Lucent 2009, ####
Some Key Properties of $TX_{p,q}$

1. **Negative local curvature.** The local angular deviation from $2\pi$ (Gauss-Bonnet curvature) of $X_{p,q}$ at each node

   \[2\pi\{4 - (p-2)(q-2)\} < 0\]

   gives the (combinatorial curvature) at node $v$

   \[\kappa_v = \frac{1}{p} + \frac{1}{q} - \frac{1}{2} < 0\]

2. **Exponential growth.** Number of nodes within a ball of radius $r$ is proportional to $\lambda^r$

   for some $\lambda = \lambda(p,q) > 1$ (e.g., for $X_{3,7}$, $\lambda = \phi$, the golden ratio) or equivalently

2’. **Logarithmic scaling of geodesics.** For a (finite truncation of) $X_{p,q}$ with $N$ nodes, the average geodesic (shortest path length) scales like $O(\log(N))$
Local Curvature versus Curvature in the Large

Computation of total curvature of non-flat networks with varying nodal degrees via $\sum_{v \in G} \kappa_v$ does not appear to be possible/easy to provide information about the large-scale properties of the network.

A more direct definition of (negative) curvature in the large is the thin-triangle condition for a geodesic metric space (or a CAT(-$\kappa$) space):

(M. Gromov’s Thin Triangle Condition for a hyperbolic geodesic metric space) There is a (minimal) value $\delta \geq 0$ such that for any three nodes of the graph connected to each other by geodesics, each geodesic is within the $\delta$-neighborhood of the union of the other two.

**Example.** For $H^2$, $\delta = \ln(\sqrt{2} + 1)$. [Sketch. Largest inscribed circle must be in largest area triangle, $\text{Area}_H(ABC) = \pi - (\alpha + \beta + \gamma)$, maximized to $\pi$ when $\alpha, \beta, \gamma = 0$ or when A, B, & C are on the boundary.]
What About Communication Networks?

Communication networks are geodesic metric spaces via reasonable link metrics (e.g., the hop metric)

Is there evidence for negative curvature in real communication networks?

We consider 10 Rocketfuel networks and some prototypical flat and curved famous networks to test this hypothesis

<table>
<thead>
<tr>
<th>Network ID</th>
<th>Network Name</th>
<th>Size #node - #links</th>
<th>Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1221</td>
<td>Telstra (Australia)</td>
<td>2998 - 7612</td>
<td>12</td>
</tr>
<tr>
<td>1239</td>
<td>Sprintlink (US)</td>
<td>8341 - 28050</td>
<td>13</td>
</tr>
<tr>
<td>1755</td>
<td>EBONE (US)</td>
<td>605 - 2070</td>
<td>13</td>
</tr>
<tr>
<td>2914</td>
<td>Verio (US)</td>
<td>3045 - 10726</td>
<td>13</td>
</tr>
<tr>
<td>3257</td>
<td>Tiscali (EU)</td>
<td>855 - 2346</td>
<td>14</td>
</tr>
<tr>
<td>3356</td>
<td>Level 3 (US)</td>
<td>3447 - 18780</td>
<td>11</td>
</tr>
<tr>
<td>3967</td>
<td>Exodus (US)</td>
<td>895 - 4140</td>
<td>13</td>
</tr>
<tr>
<td>4755</td>
<td>VSNL (India)</td>
<td>121 - 456</td>
<td>6</td>
</tr>
<tr>
<td>6461</td>
<td>AboveNet (US)</td>
<td>2720 - 7648</td>
<td>12</td>
</tr>
<tr>
<td>7018</td>
<td>AT&amp;T (US)</td>
<td>10152 - 28638</td>
<td>12</td>
</tr>
<tr>
<td>TX(3,7)</td>
<td>Synthetic</td>
<td>4,264-15,022</td>
<td>14</td>
</tr>
<tr>
<td>Power-law</td>
<td>Synthetic</td>
<td>10,000 - 39,994</td>
<td>9</td>
</tr>
<tr>
<td>(Albert-Barabasi)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strogatz-Watts (2D)</td>
<td>Synthetic (p=0.2)</td>
<td>80x80 - 26,578</td>
<td>20</td>
</tr>
<tr>
<td>Truncated (3,6) Grid</td>
<td>Synthetic</td>
<td>469 - 2,520</td>
<td>24</td>
</tr>
<tr>
<td>Truncated (4,4) Grid</td>
<td>Synthetic</td>
<td>80x80 - 25,280</td>
<td>158</td>
</tr>
</tbody>
</table>
Experiments and Methodology

We ran experiments on all Rocketfuel networks plus a few prototypical flat/curved networks to test our key hypothesis:

1. Dimension. “Growth test” - Polynomial or exponential?
   - Consider the volume V(r) as a function of radius r for arbitrary centers
   - [In flat graphs volume growth is typically polynomial in radius r]

2. Curvature. “Triangle test” - Are triangles are universally δ-thin
   - Randomly selected 32M, 16M, 1.6M triangles for networks with more than 1K nodes and exhaustively for the remainder
   - For each triangle noted shortest side L and computed the δ
   - Counted number of such triangles, indexed by δ and L
   - [In flat graphs δ grows without bound as the size of the smallest side increases]

We conduct “growth” and “triangle” tests
1. Growth Charts

Recall that:

Euclidean growth $V(r) \approx r^D$
Exponential growth $V(r) \approx \theta r$

Volume (number of points with distance $r$) as a function of radius $r$ from a “center” of the graph. Flattening of curves for larger $r$ is due to approach to boundary / finite size of network.
2. Triangle Test - Rocketfuel 7018 (AT&T) & Triangular Grid

(a) Probability $P_L(\delta)$ for randomly chosen triangles whose shortest side is $L$ to have a given $\delta$ for the network 7018 (AT&T network) which has 10152 nodes and 14319 bi-directional links and diameter 12. The quantities $\delta$ and $L$ are restricted to integers, and the smooth plot is by interpolation.

(b) Similar to (a), for a (flat) triangular lattice with 469 nodes and 1260 links. (The smaller number of nodes is sufficient for comparing with (a) since the range for $L$ is large due to the absence of the small world effect.)
Summary of Triangle Tests for Rocketfuel Networks

The average $\delta$ as a function of $L$, $E[\delta](L)$, for the 10 IP-layer networks studied here, and for the Barabasi-Albert model with $k = 2$ and $N = 10000$ (11th curve) and the hyperbolic grid $X3.7$ (12th curve). On the other hand, a Watts- Strogatz type model on a square lattice with $N = 6400$, open boundary conditions and 5% extra random connections (13th curve) and two flat grids (the triangular lattice with diameter 29 and the square lattice with diameter 154) are also shown.
Where to go from here?

- OK, these ten RF datasets and some “well-known” large-scale networks exhibit
  
  - Exponential growth / logarithmic scaling of shortest paths
  
  - Negatively curvature in the large

So what?

- Turns out negatively-curved structures exhibit specific features that affect critical properties of networks
  
  - Existence of a “core”
The Downside of Hyperbolicity: Quadratic Scaling of Load ("Betweenness Centrality" and Existence of "Core")

Plot of the maximum load $L_c(N)$ -- maximal number of geodesics intersecting at a node -- for each network in the Rocketfuel database as a function of the number of nodes $N$ in the network. Also shown are the maximum load for the hyperbolic grid $X_{3,7}$, the Barabasi-Albert model with $k = 2$, the Watts-Strogatz model and a triangular lattice, for various $N$. The dashed lines have slopes of 2.0 and 1.5, corresponding to the hyperbolic and Euclidean cases respectively.
Key Claims: Network Curvature -> Congestion, Reliability and Security

Congestion is not necessarily a manifestation of the heavy-tailed property of the distribution of the node degrees, but rather of a more subtle feature, given the (negative) curvature of the network.

Numerical studies show that congestion does not necessarily occur at vertices of high degree, nor at the so-called highly connected nodes but rather at the points relative to which the network has minimum inertia (the “core”).

Geodesics (shortest paths)

- (Upside) Are very effective, as diameter is small compared to \(N\), e.g., TTL of ~20 good enough for all of the Internet!
- (Downside) Lead to
  - congestion
  - non-random failure can be very severe
  - certain nodes exhibit significant security threats

Need for non-geodesic routing to avoid the downside

\[X_{3,7}\]

Nodal loads need not be related to nodal degrees

Geodesic triangles on negatively curved spaces: thin regardless of the edge lengths
Taxonomy for Large-Scale Networks

Taxonomy of key characteristics of networks and their overlaps in a schematic diagram.
To Summarize

- Networks, in addition to the small world property -- that has been widely verified -- and the power law degree distribution -- that is often but not always observed -- have another signature: curvature

- Rocketfuel ISP networks exhibit negative curvature and small world behavior and (within numerical accuracy) power law degree distribution

- Synthetic network models can show negative curvature (B-A) or not (S-W)

- Analytical prediction on continuum and hyperbolic lattices shows $N^2$ scaling load at their “core”. This agrees with nodal power-law and Rocketfuel, but not with S-W.

- Thus hyperbolicity seems sufficient AND necessary for an important network performance measure. Real networks show this behavior, so it is relevant!

- Need for routing protocols that don’t fall into geodesic grooves
References
