Centrality, treeness and miscellaneous

Social and Technological Networks

Rik Sarkar

University of Edinburgh, 2016.

Centrality

- How ‘central’ is a node in a network?
  - A notion of importance of the node
- E.g. degree, pagerank, beweenness..

- Degree centrality
  - Degree of a vertex

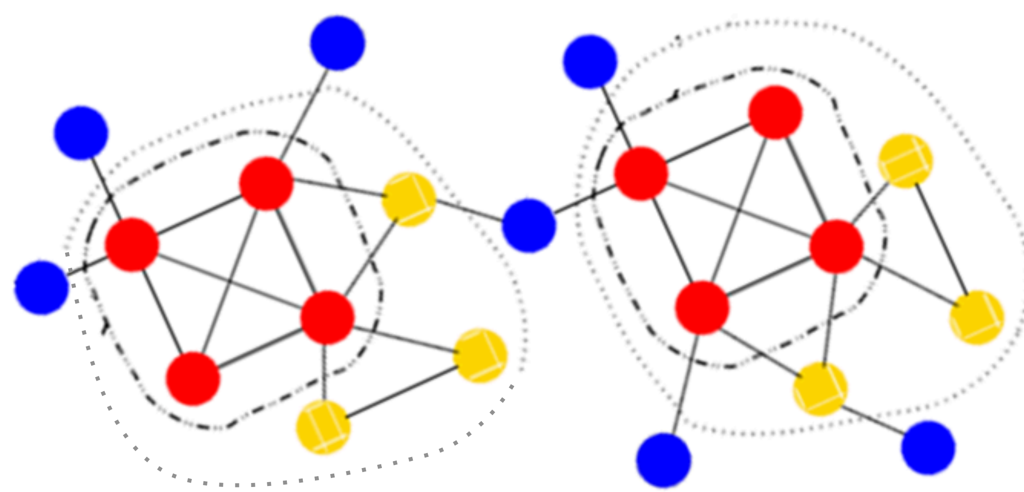
- Closeness centrality
  - Average distance to all other nodes $\ell_x = \frac{1}{n} \sum_y d(x, y)$
    - Decreases with centrality
    - Inverse is an increasing measure of centrality

$$C_x = \frac{1}{\ell_x} = \frac{n}{\sum d(x, y)}$$

- Betweenness centrality
  - The number of shortest paths passing through a node
    - (see slides from strong and weak ties)
- Pagerank
  - See slides on web graphs and ranking pages
  - Pagerank is a type of Eigenvector centrality
  - Another eigen centrality is Katz centrality, which we will not discuss

k-core of a graph G

- A maximal connected subgraph where each vertex has a degree at least k
  - *Inside that subgraph.*
- Obtained by repeatedly deleting vertices of degree less than k



Internet

- An interconnection network of “network of routers”
- Thousands of networks together form the Internet
- The “center” consists of big routers in highly connected networks, many connections between adjacent networks
- Outer layers have smaller routers and sparser connections

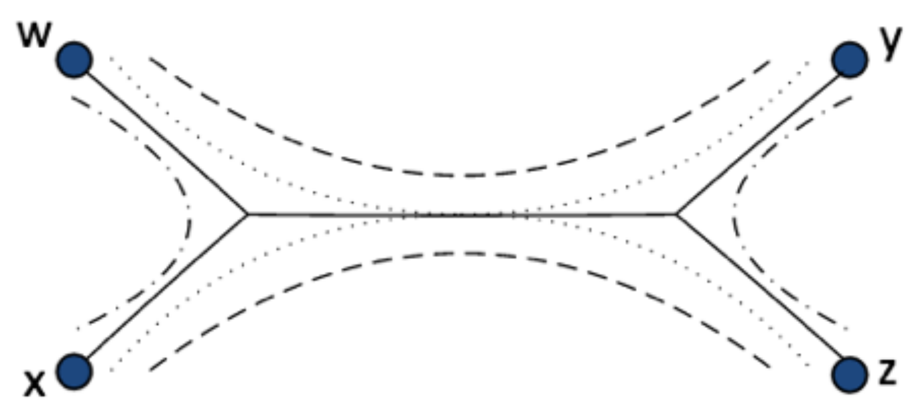
Internet

- Has a layered structure with higher connectivity at the core
  - A routed packet tends to use high connectivity regions to get shorter/faster routes
  - Effectively a tree-like structure

- Known to have power law distribution of degrees

A test for tree metrics

- A metric is a tree metric if and only if it satisfies this 4 Point Condition:
- Any 4 nodes (points in the metric space) can be ordered as w,x,y,z such that:
- $d(w,x) + d(y,z) \le d(w,y) + d(x,z) \le d(w,z) + d(x,y)$ and
- $d(w,y) + d(x,z) = d(w,z) + d(x,y)$

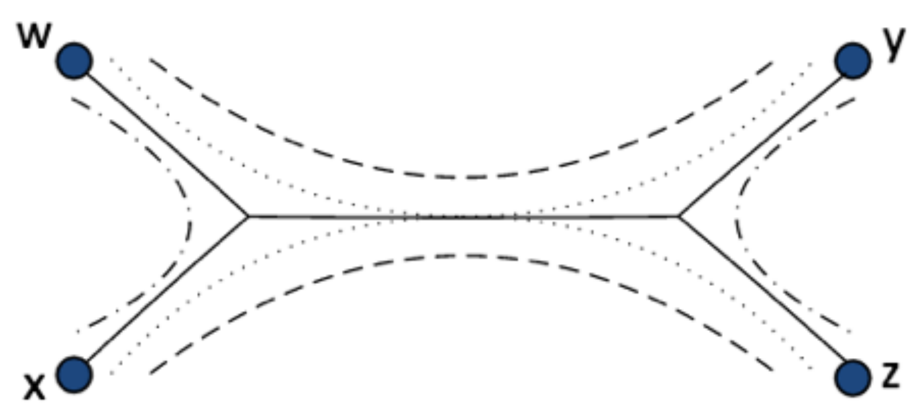


Trees tend to have high loads in "center"

- Since many routes will have to go through the center

Almost tree metrics

- Real networks are not exactly trees
- Let's measure how far a network is from a tree
- 4PC-ε for a set of 4 nodes is the smallest ε that satisfies:
- $d(w,x) + d(y,z) \leq d(w,y) + d(x,z) \leq d(w,z) + d(x,y)$ and
- $d(w,z) + d(x,y) \leq d(w,y) + d(x,z) + 2\varepsilon \cdot \min\{d(w,x),d(y,z)\}$

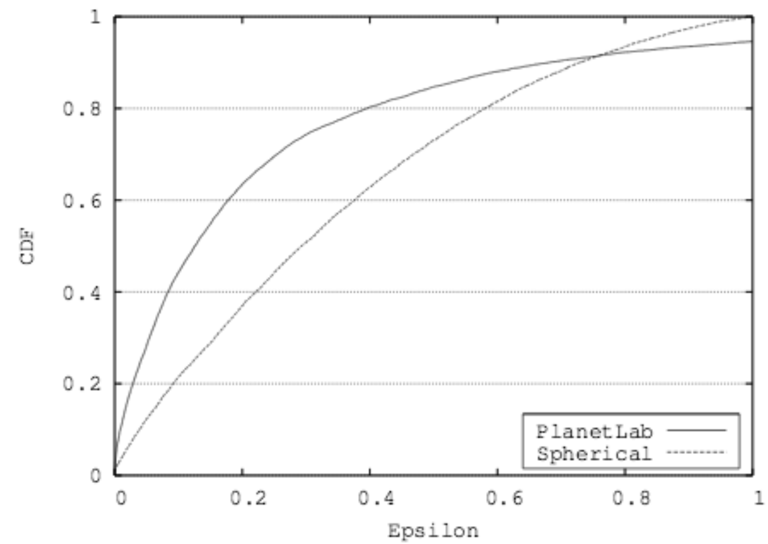


Almost tree metrics

- A tree has $\varepsilon = 0$
- A metric space with smaller ε implies that it is more similar to a tree
  - Theorem: A metric space with small ε can be embedded into a tree with correspondingly small distortion
  - Ref: I Abraham et al. Reconstructing approximate tree metrics, PODC 07.

Treeness of Internet

- PlanetLab: A distributed collection of servers around the world
- Experiment based on latency (communication delay) as an estimate of distance
- Shows the distance metric between servers is similar to a tree, and far from a sphere

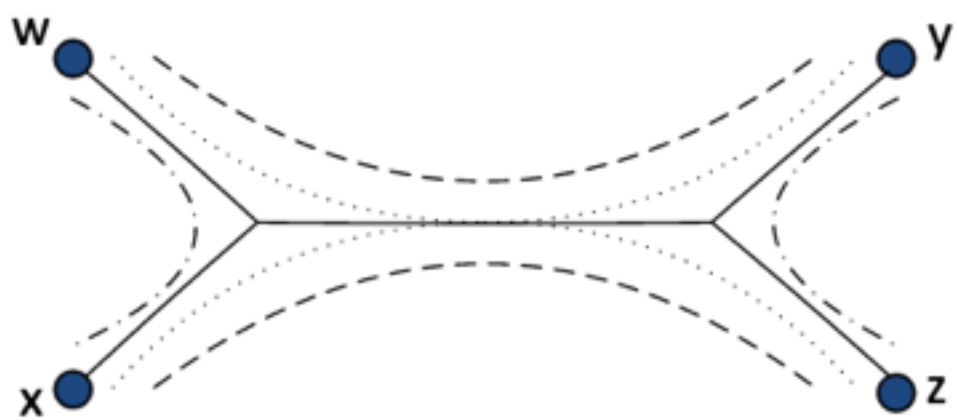


- Ref: I Abraham et al. Reconstructing approximate tree metrics, PODC 07.
- V. Ramasubramanian etal. On treeness of internet latency and bandwidth, Sigmetrics 09.

Treeness of metrics

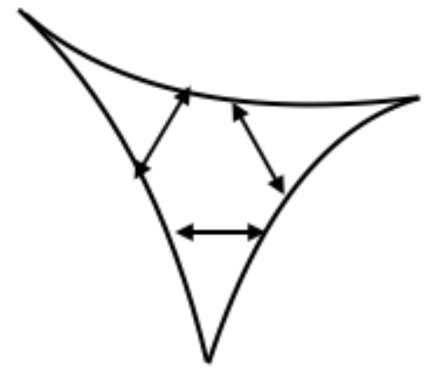
- δ-hyperbolic metrics
- $d(w,x) + d(y,z) \leq d(w,y) + d(x,z) \leq d(w,z) + d(x,y)$ and
- $d(w,z) + d(x,y) = d(w,y) + d(x,z) + \delta$

- Uses an absolute value δ
- Instead of a multiplicative factor



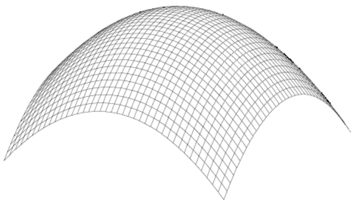
δ-hyperbolic metrics: Thin triangles

- Alternative definition
- Any point on a triangle must be within distance δ of one of the *other* sides
- The middle of the triangles are squeezed together
- trees have δ = 0: most hyperbolic

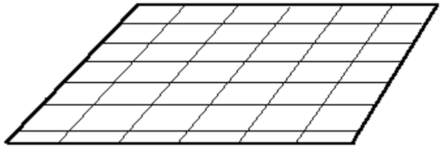


Curvatures of spaces

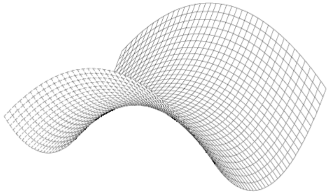
- Spherical : +ve curvature
- Triangle centers are “Fat”

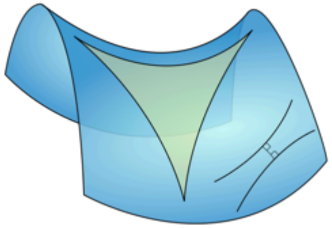


- Flat (Euclidean): 0 Curvature



- Hyperbolic: -ve curvature

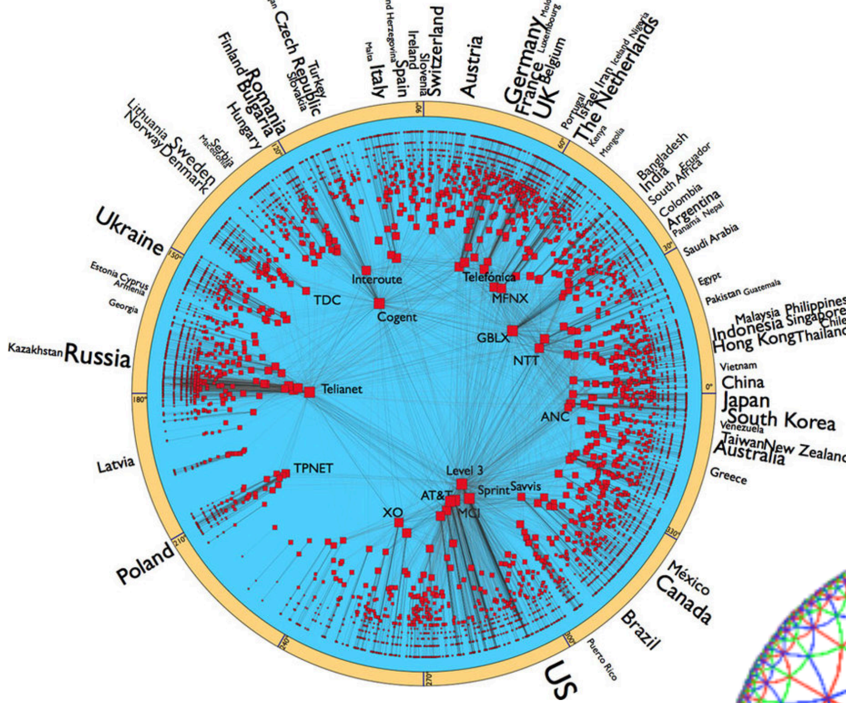




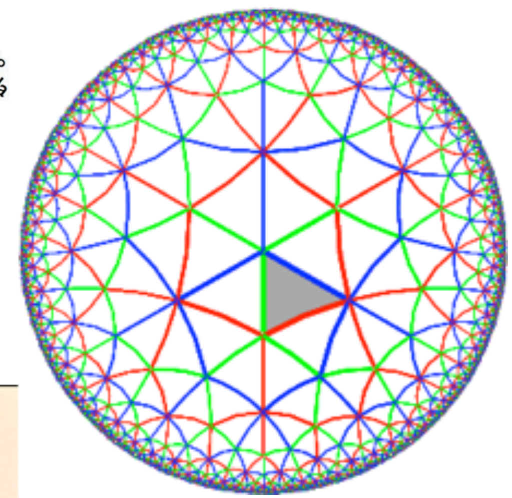
- Any hyperbolic space is δ-hyperbolic for some finite delta
- Not the case for Euclidean and spherical spaces
- For more on δ-hyperbolic spaces, See: Gromov hypoerbolic spaces

- Any tree can be embedded in a hyperbolic space with a low distortion
- R. Sarkar. “Low distortion delaunay embedding of trees in hyperbolic plane.” GD 2011.

- Internet has good embedding in hyperbolic spaces



- Ref. Shavitt and Tankel 2008, Narayan and Saniee 2011



- Model for power law social networks with clustering properties
- Place nodes in hyperbolic plane
  - Later nodes are farther away from center
  - At random angle from center
  - Nodes connect probabilistically to nodes closer in hyperbolic distance
- Ref: Popularity vs similarity in growing networks
  - Papdopoulos et al. Nature 2012

Course

- Project:
  - Submission tomorrow
  - **Individual submissions and report**
  - **Read submission instructions carefully**
  - Submit early. Do not keep for the last moment. You can always resubmit
- Lecture:
  - Last lecture Friday
  - Discussion of course, study material, exams
- Exam material:
  - Slides.
  - Items in “Reading”.
  - Not “Additional” reading, or references in slides.