Routing in the Internet and Navigability of Scale-Free Networks

Dmitri Krioukov
CAIDA
dima@caida.org
M. Boguñá, M. Á. Serrano, F. Papadopoulos, X. Dimitropoulos, kc claffy, A. Vahdat

Mallorca, June 9th, 2008

What the Internet does

- The Internet was designed for and exists to transfer information packets from *A* to *B*, where *A* and *B* are any two Internet-Protocol- (IP-)talking devices

IP packet format

+	Bits 0–3	4–7	8–15	16–18	19–31
0	Version	Header length	Type of Service (now DiffServ and ECN)	Total Length	
32	Identification			Flags	Fragment Offset
64	Time to Live		Protocol	Header Checksum	
96	Source Address				
128	Destination Address				
160	Options				
160 or 192+	Data				

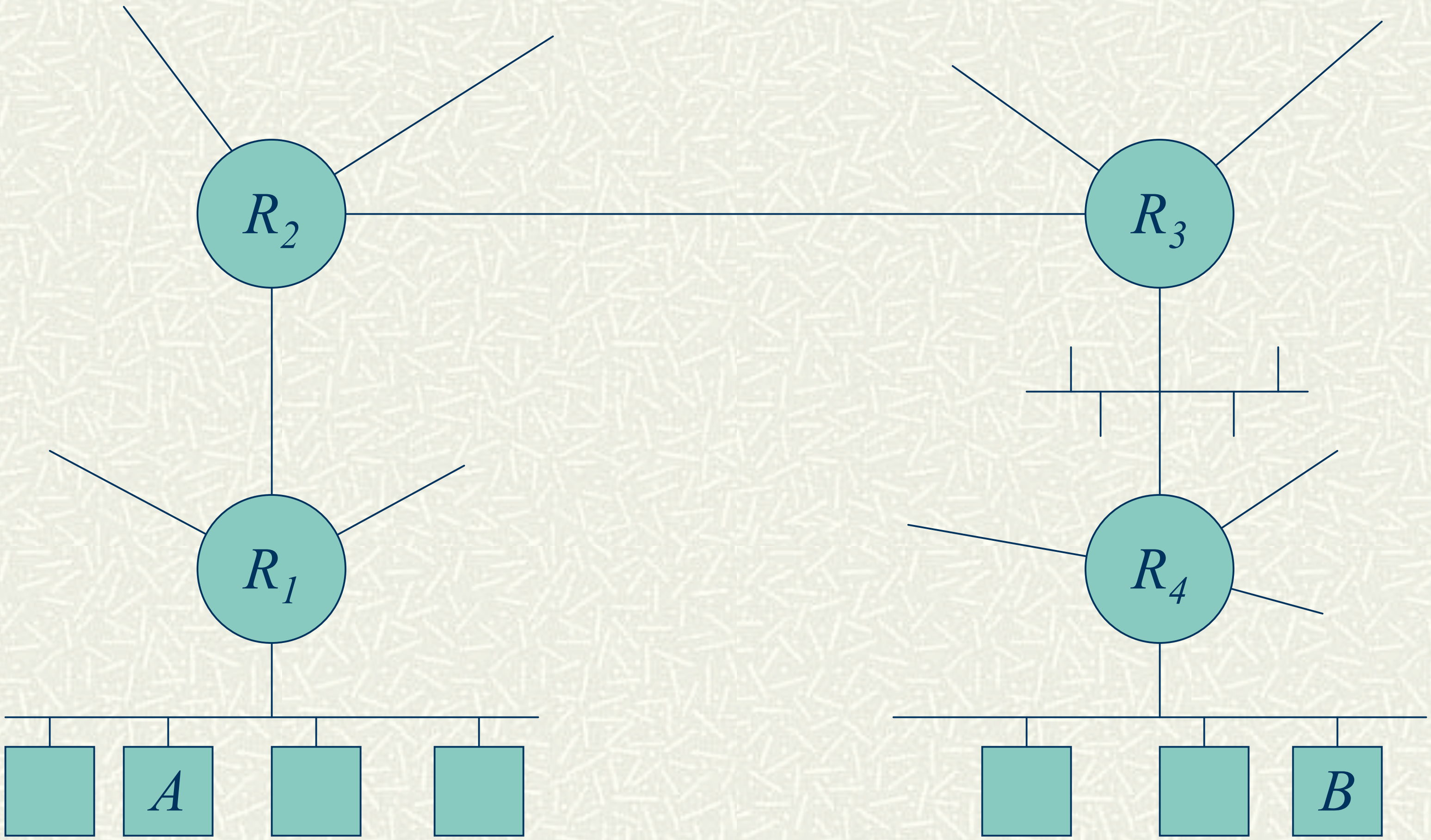
IP addresses

- $A = 161.116.80.85$
- $B = 192.172.226.78$

IP routes

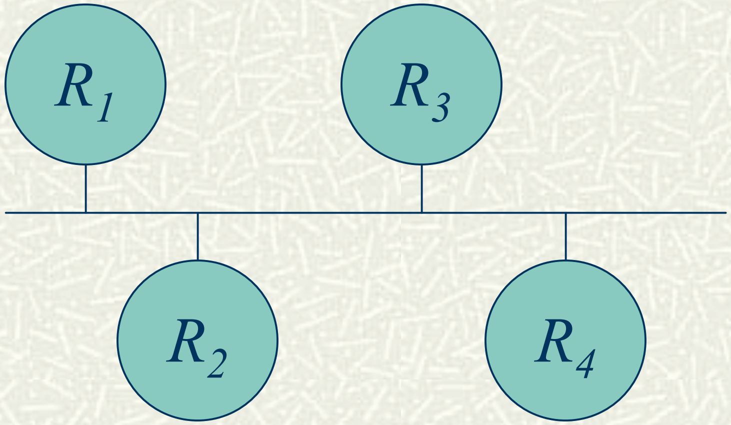
```
traceroute 192.172.226.78

  1    <1 ms    <1 ms    <1 ms  161.116.80.254
  2     *        *        *     Request timed out.
  3    <1 ms    <1 ms    <1 ms  161.116.221.14
  4     1 ms    <1 ms    <1 ms  192.168.3.250
  5     8 ms     1 ms     1 ms  84.88.18.5
  6     1 ms    <1 ms    <1 ms  130.206.202.29
  7    15 ms    15 ms    15 ms  130.206.250.25
  8    15 ms    15 ms    15 ms  130.206.250.2
  9    16 ms    15 ms    15 ms  62.40.124.53
 10    37 ms    37 ms    37 ms  62.40.112.25
 11    50 ms    45 ms    45 ms  62.40.112.22
 12   138 ms   138 ms   138 ms  62.40.125.18
 13   152 ms   152 ms   152 ms  64.57.28.6
 14   175 ms   175 ms   175 ms  64.57.28.43
 15   207 ms   217 ms   207 ms  64.57.28.44
 16   209 ms   208 ms   209 ms  137.164.26.132
 17   215 ms   215 ms   215 ms  137.164.25.5
 18   215 ms   215 ms   215 ms  137.164.27.50
 19   215 ms   215 ms   215 ms  198.17.46.56
 20   215 ms   215 ms   215 ms  192.172.226.78
```

IP routers

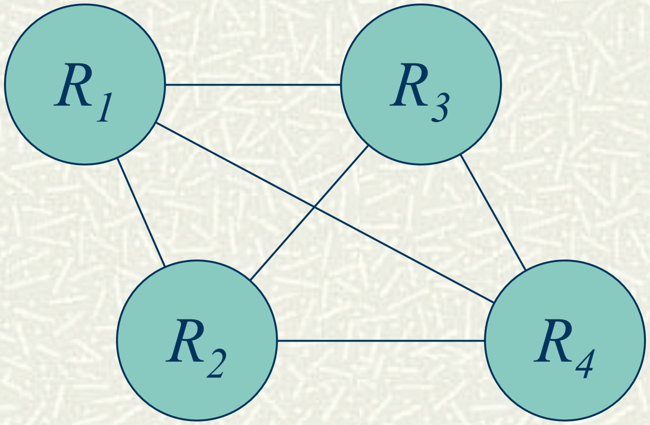


Broadcast media (e.g., ethernet)

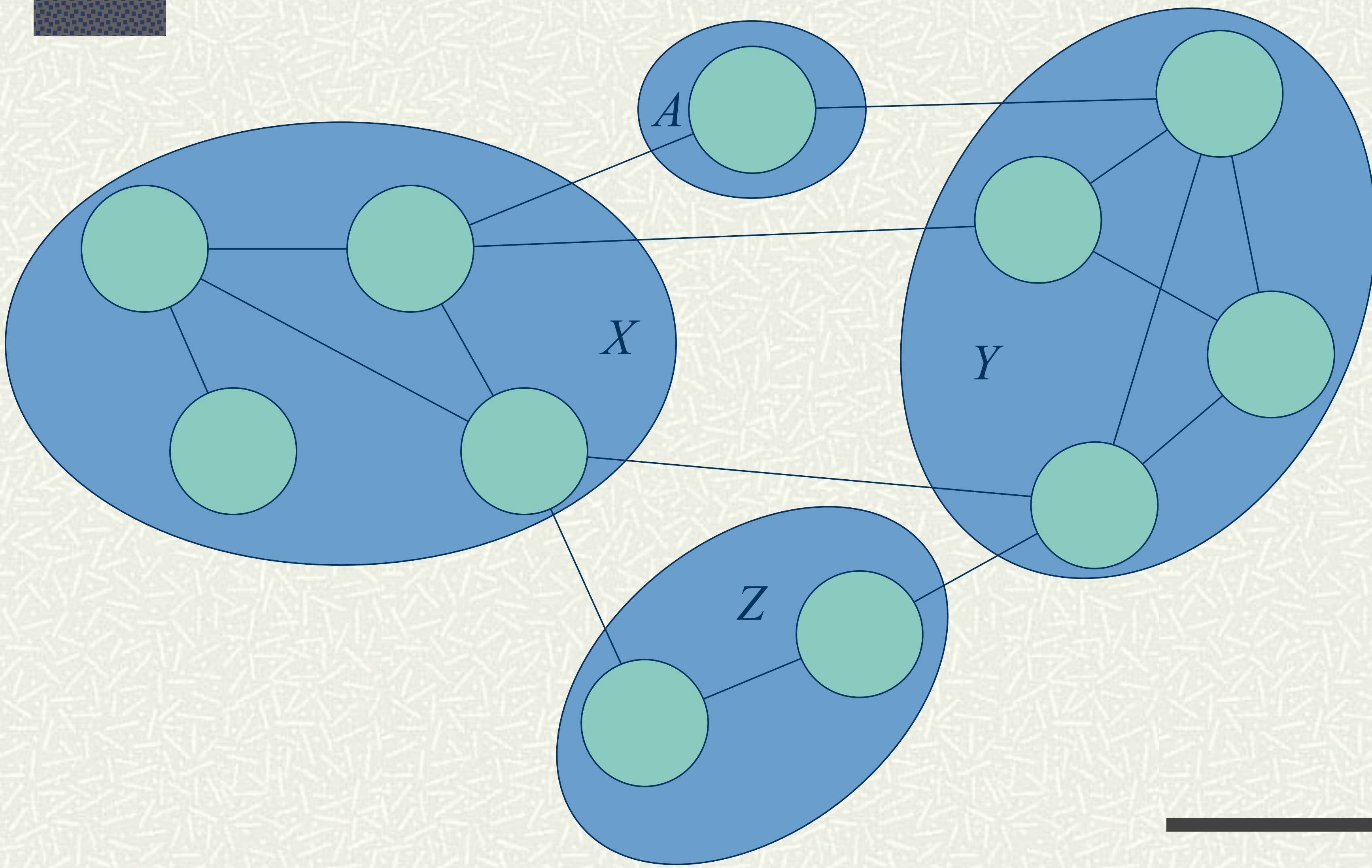
Reality



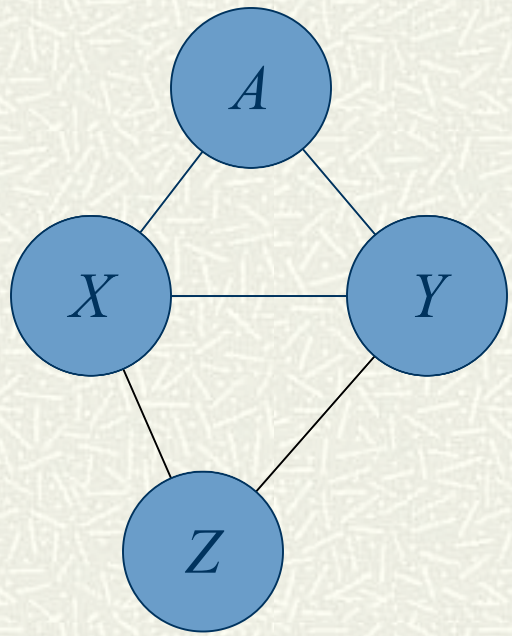
Perception



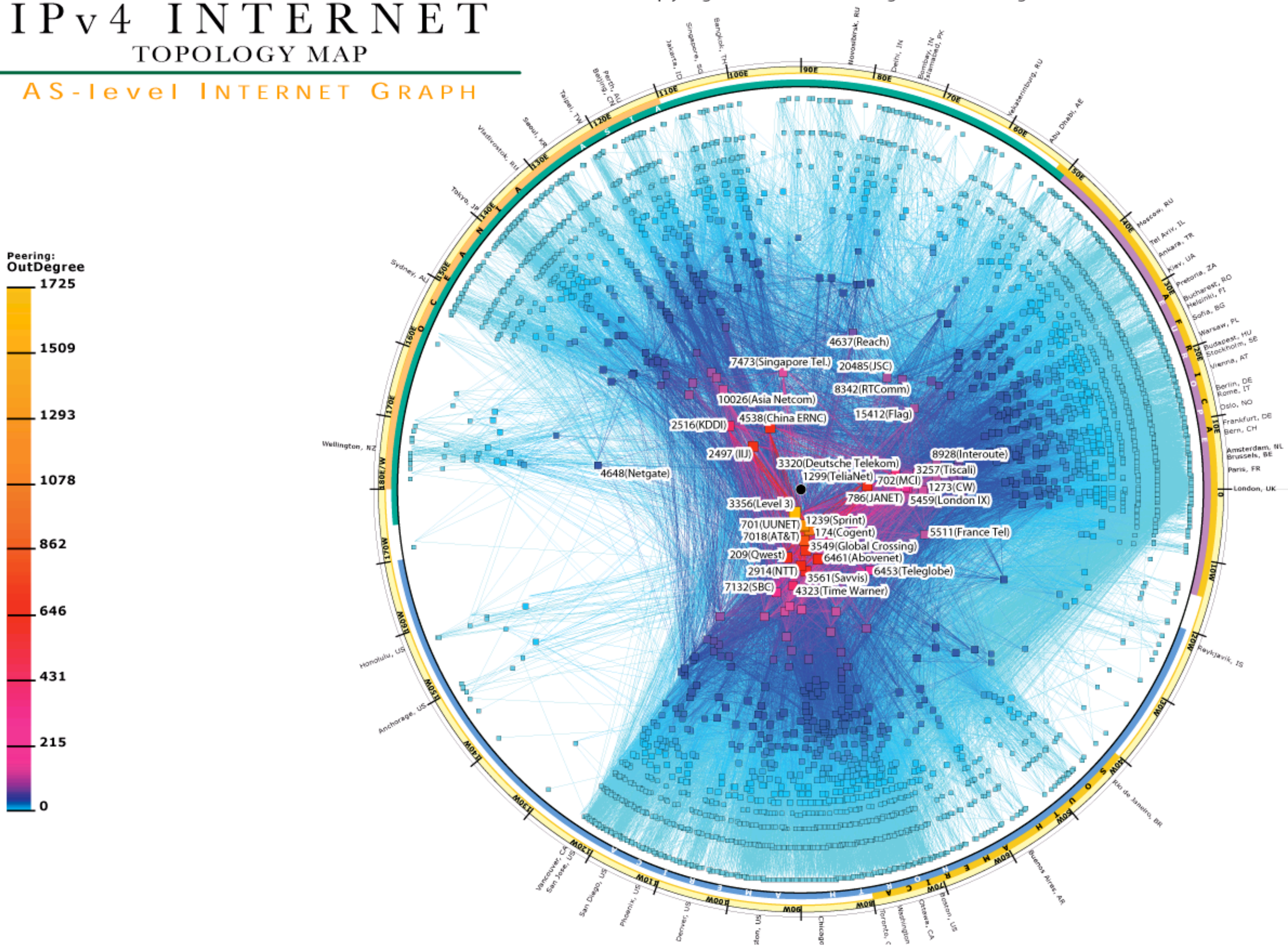
Autonomous Systems



AS topology



IPv4 INTERNET
TOPOLOGY MAP
AS-level INTERNET GRAPH



IP routing

- Intradomain (Interior Gateway Protocols (IGPs))
  - routing within an Autonomous System (AS)
  - protocols:
    - Open Shortest Path First (OSPF)
    - Intermediate System to Intermediate System (ISIS)
  - Links State (LS) routing protocols
- Interdomain (Exterior Gateway Protocols (EGPs))
  - routing between Autonomous Systems (ASs)
  - protocols:
    - Border Gateway Protocol (BGP)
  - Path Vector (PV) routing protocol

BGP

- Each AS advertises IP addresses that it has
  - AS 13041 (University of Barcelona) advertises: 161.116.0.0 – 161.116.255.255 (161.116.0.0/16)
- All neighboring ASs receiving this advertisement re-advertise them to their neighbors after pre-pending their AS numbers
- The result is that each AS *A* has a routing entry for 161.116.0.0/16 which looks like: 161.116.0.0/16: AS X_1, AS X_2, …, AS 766, where X_1 is a neighbor of *A*, X_2 is a neighbor of X_1, and so on.

AS relationships and BGP policies

- Each AS link is the relationship (i.e., business, contractual agreement) between the two ASs
- There are roughly three types of such relationships
  - customer-provider (c2p)
  - peer-peer (p2p)
  - sibling-sibling (s2s)
- Standard routing policies: to reach a destination, the route preference order is
  - routes via customers
  - routes via peers
  - routes via providers
- Standard route re-advertisement policies
  - re-advertising to provider or peer, an AS advertises only its own IP addresses and IP routes learnt from its customers
  - re-advertising to customer or sibling, an AS advertises everything
- BGP advertisement policy combinations vs. AS relationships
  - asymmetric combination: c2p
  - symmetric combinations: p2p and s2s

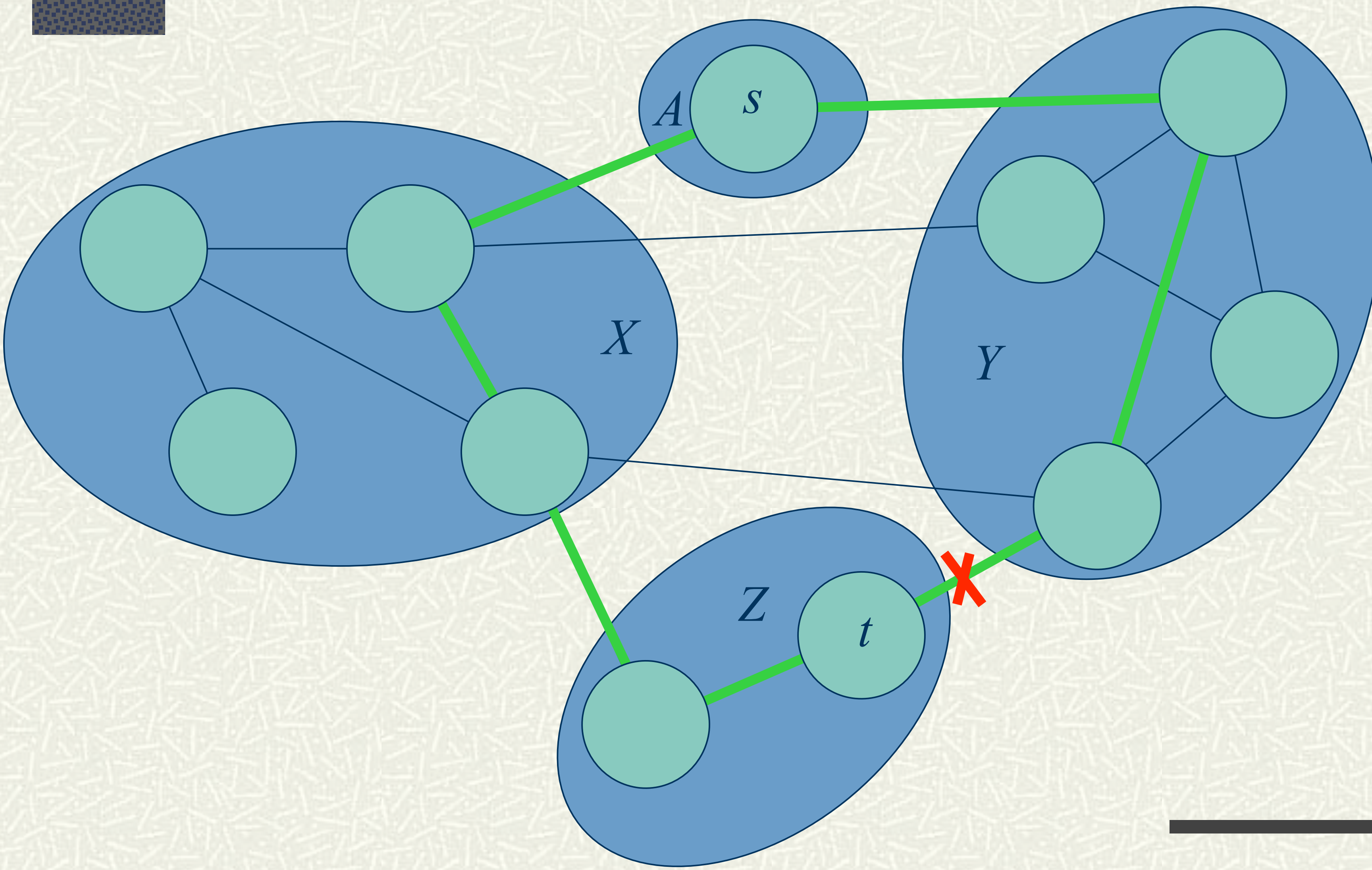
Hierarchy of valid paths

- Valid paths consists of the following potions
  - uphill: zero or more links from customer to provider
  - pass: zero or one link from peer to peer
  - downhill: zero or more links from provider to customer
  - any number of sibling links anywhere in the path
- Given a collection of paths observed in BGP routing tables, trying to assign relationships to AS links that minimize the number of invalid paths is a way to infer AS relationships

Internet topology

- Cumulative result of local, decentralized, and rather complex interactions between AS pairs
- Surprisingly, in 1999, it was found to look completely differently than engineers had thought: it shares all the main features of topologies of other complex networks (scale-free degree distributions and strong clustering)
- Routing protocols have to find ***and update*** valid paths to destinations through it

Simple routing event



BGP dynamics

- BGP updates
  - 2 per second on average
  - 7000 per second peak rate
- Convergence after a single event can take up to tens of minutes

Routing theory

- There can be no routing algorithm with the number of messages per topology change scaling better than linearly with the network size in the worst case
- Small-world networks are this worst case

- *Is there any workaround?*
- *If topology updates/convergence is so expensive, then may be we can route without them, i.e., without global knowledge of the network topology?*

Milgram's experiments

- Settings: random people were asked to forward a letter to a random individual by passing it to their friends who they thought would maximize the probability of letter moving "closer" to the destination
- Results: surprisingly many letters (30%) reached the destination by making only ~6 hops on average
- Conclusion:
  - People do not know the global topology of the human acquaintance network
  - But they can still find (short) paths through it

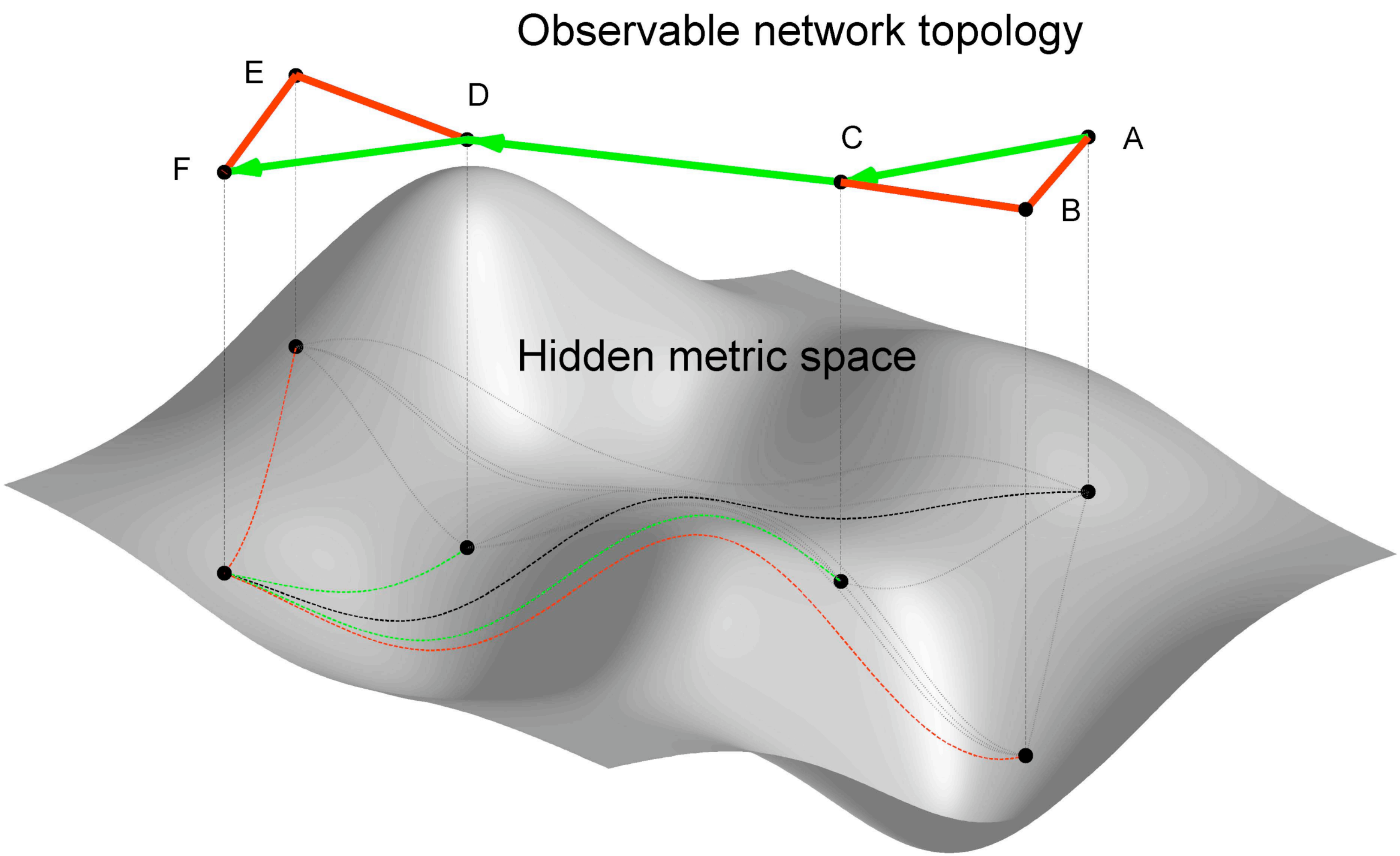
Hidden metric space explanation

- All nodes exist in a metric space
- Distance in this space abstract node similarities
- Network consists of links that exist with probability that decreases with the hidden distance
- More similar/close nodes are more likely to be connected
- The result is that all nodes exist in “two places at once”:
  - a network
  - a hidden metric space
- So that there are two distances between each pair of nodes
  - the length of shortest path between them in the network
  - hidden distance

Greedy routing (Kleinberg)

- To reach a destination, each node forwards information to the one of its neighbors that is closest to the destination in the hidden space

Hidden space visualized



Questions raised by the model

- What is the hidden space?
- What are the node positions in it?
- What is the connection probability?
- How efficient is the greedy routing process?
  - How often greedy-routing paths get stuck at nodes that do not have any neighbors closer to the destination than themselves
  - How close greedy-routing paths are to shortest paths in the network

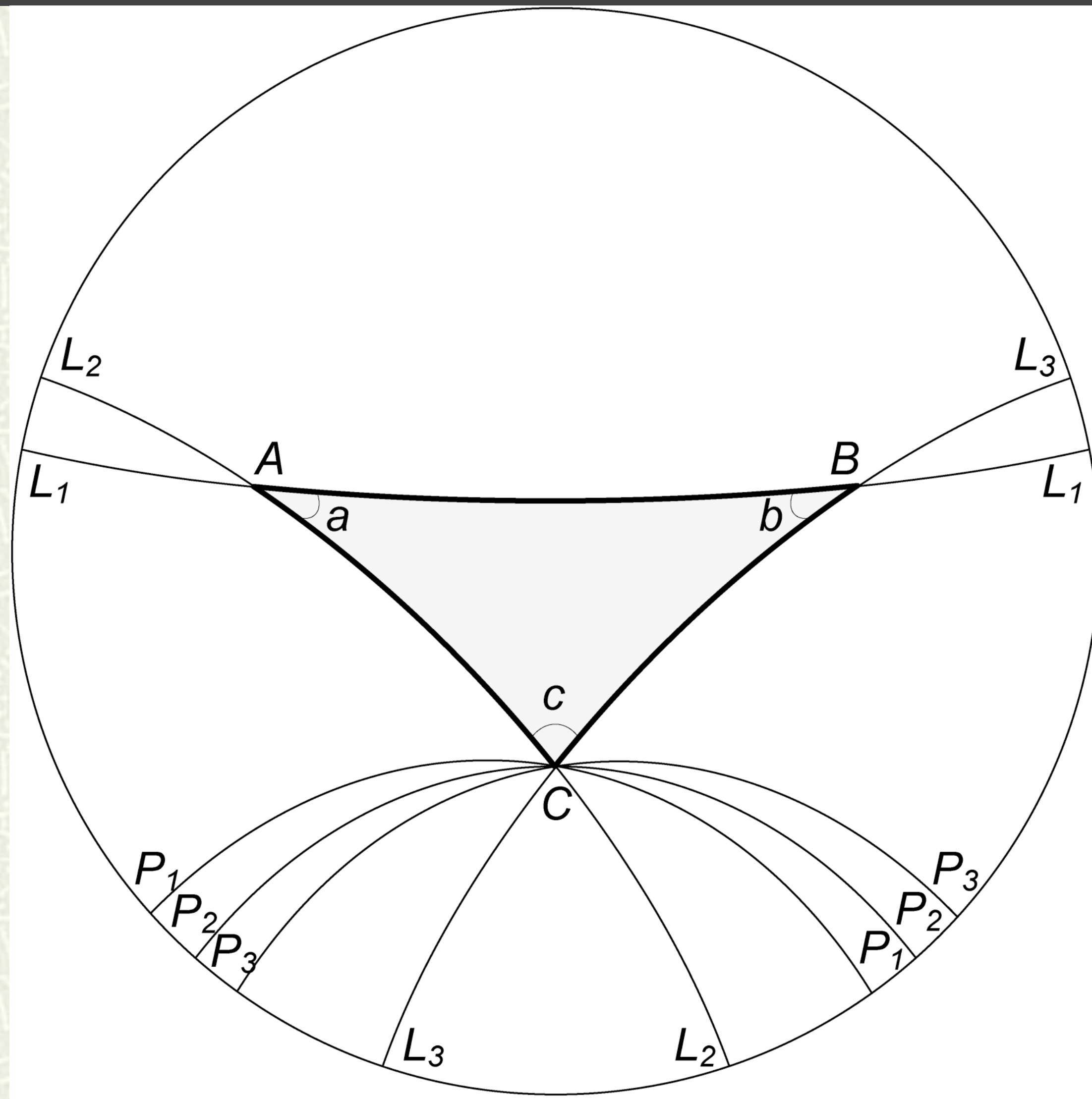
Hidden metric spaces are hyperbolic

- Network nodes can often be hierarchically classified
- Hierarchies are (approximately) trees
- Trees embed isometrically in hyperbolic spaces

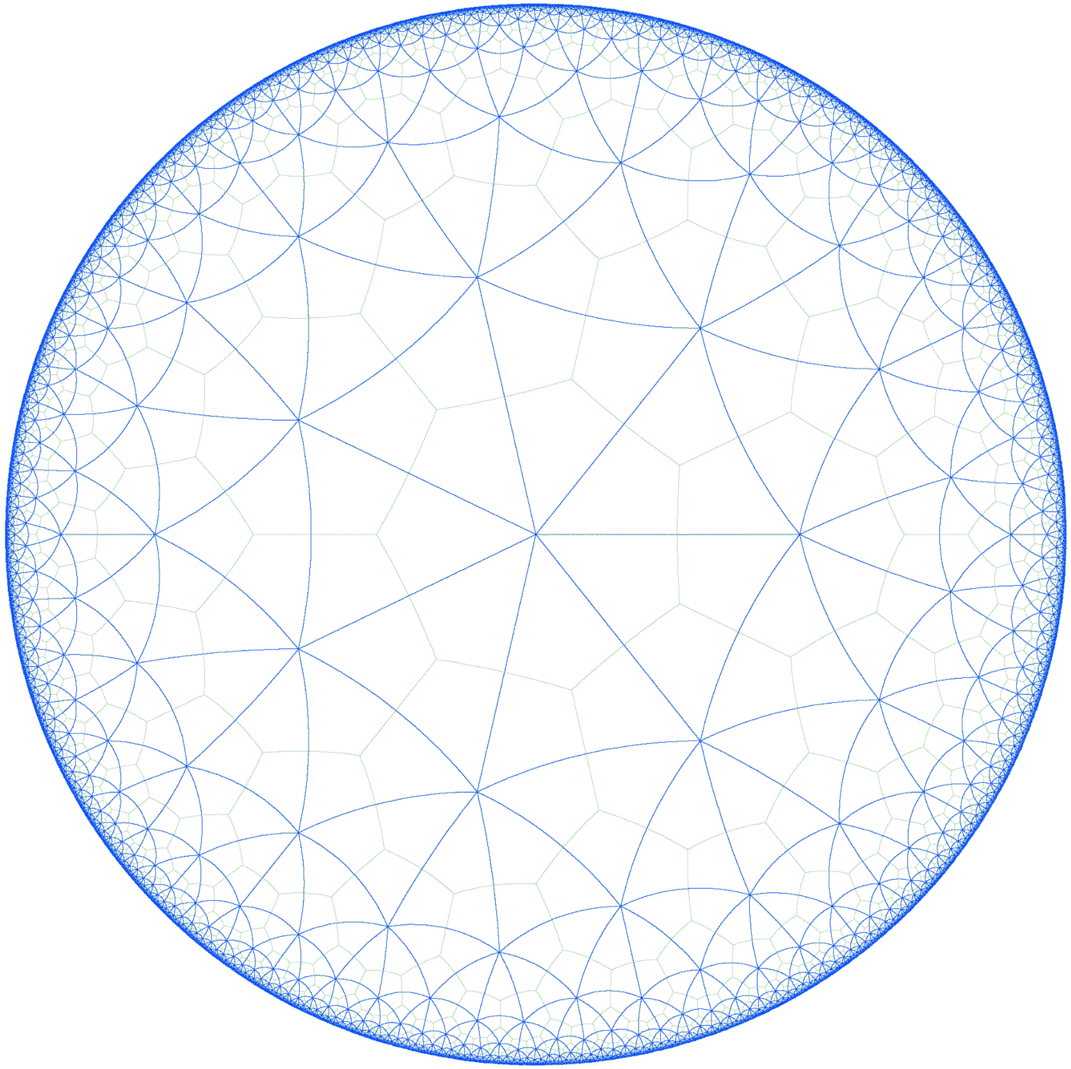
Hyperbolic geometry

- Geometry in which through a point not belonging to a line passes not one but infinitely many lines parallel to the given line

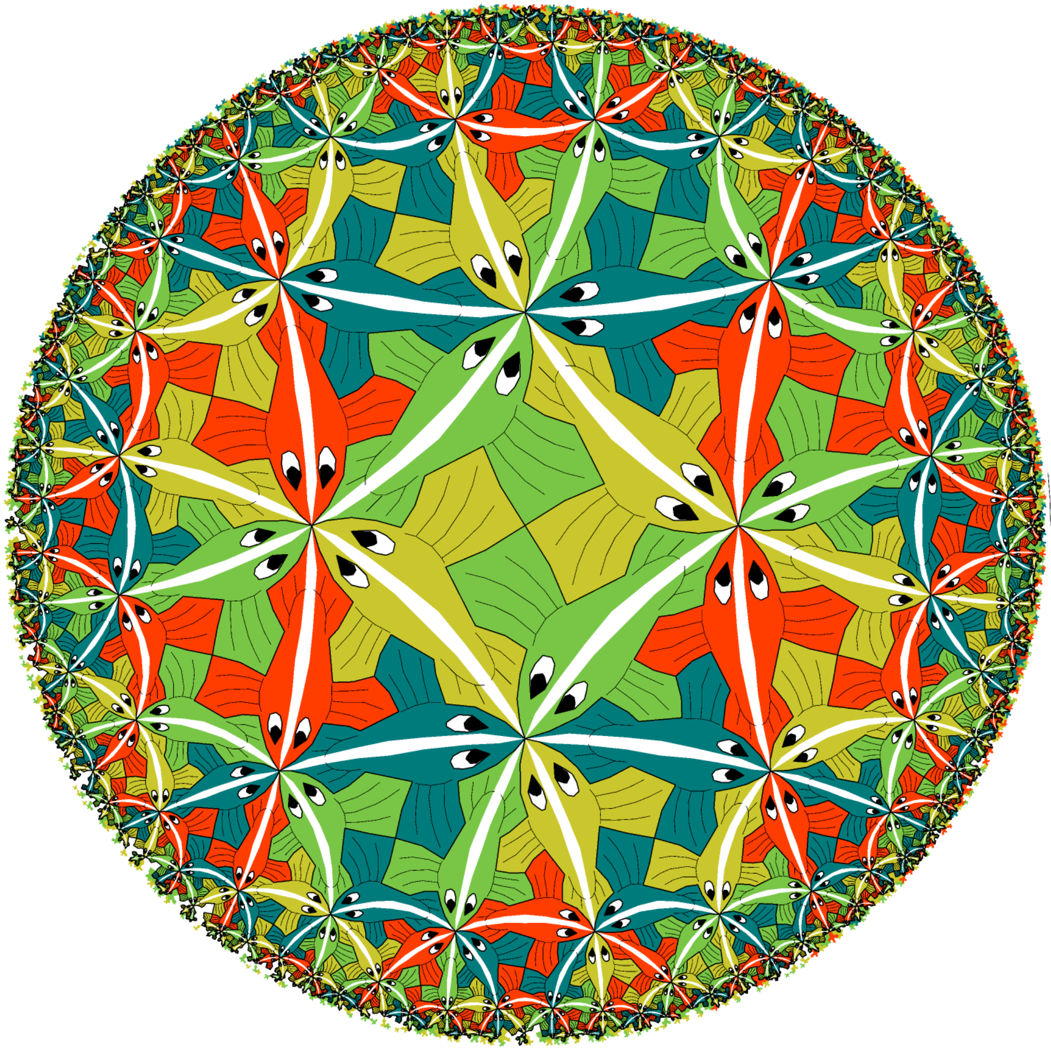
Poincaré disc model



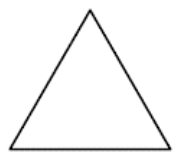
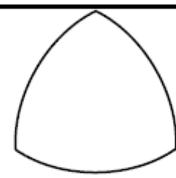
Tessellation and tree embedding



Tessellation art



Geometry properties

Property	Euclid.	Spherical	Hyperbolic
Curvature	0	1	-1
Parallel lines	1	0	∞
Triangles are	normal	thick	thin
Shape of triangles			
Sum of angles	π	$> \pi$	$< \pi$
Circle length	$2\pi R$	$2\pi \sin R$	$2\pi \sinh R$
Disc area	$2\pi R^2/2$	$2\pi(1 - \cos R)$	$2\pi(\cosh R - 1)$

Main hyperbolic property

- The volume of balls and surface of spheres grow with their radius *r* as

$$e^{\alpha r}$$

  where $\alpha = (-K)^{1/2}(d-1)$, K is the curvature and d is the dimension of the hyperbolic space
- The number of nodes in a tree within or at *r* hops from the root grow as

$$b^{r}$$

  where b is the tree branching factor

Hidden space in our model

- Hyperbolic disc of radius R, where $N = \kappa\, e^{R/2}$, N is the number of nodes in the network and κ controls its average degree

Node distribution

- Number of nodes *n(r)* located at distance *r* from the disc center is

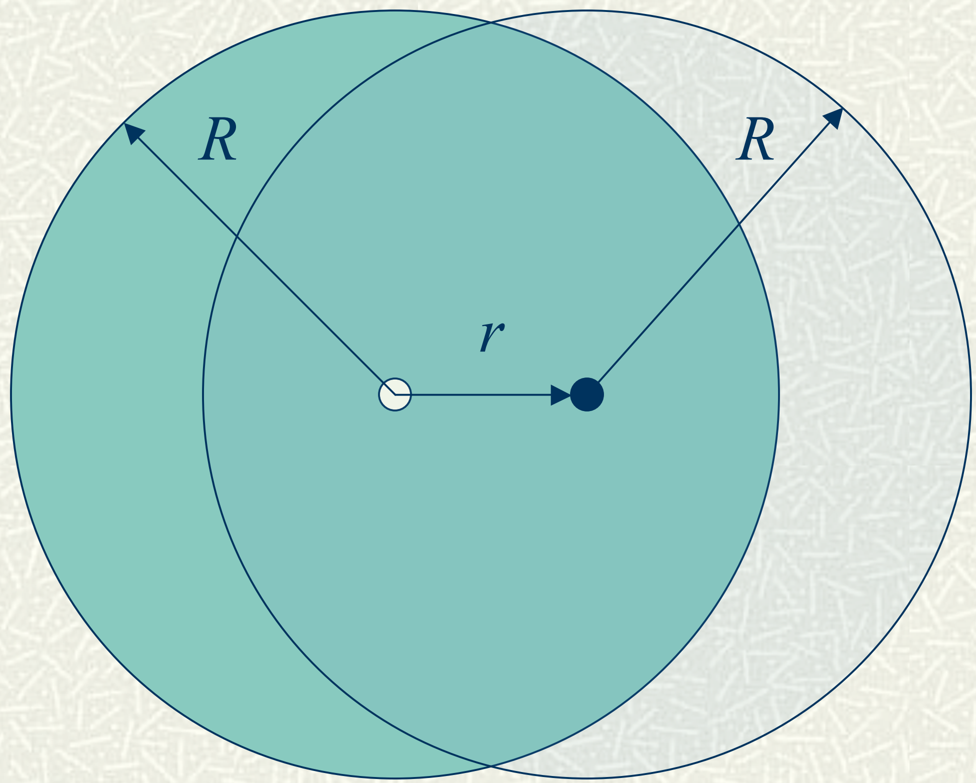
$$n(r) \sim e^{\alpha r}$$

where $\alpha = 1$ corresponds to the uniform node distribution in the hyperbolic plane of curvature -1

Connection probability

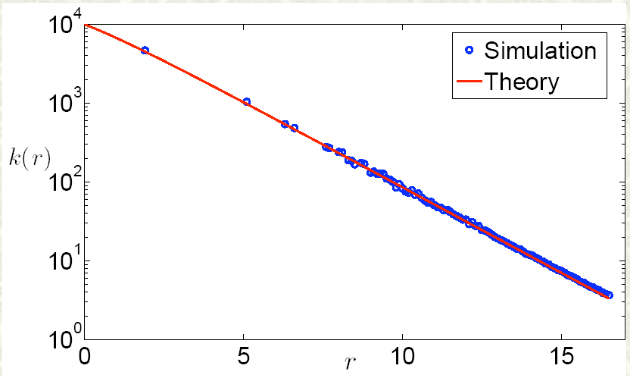
- Connected each two nodes if the distance between them is less than or equal to R

Average node degree at distance *r* from the disc center



Average node degree at distance *r* from the disc center

- For $\alpha = 1$, we obtain a terse but exact expression



- For other α:

$$k(r) \sim e^{-\beta r}$$

where

$\beta = \alpha$ if $\alpha \leq ½$

$\beta = ½$ otherwise

Node degree distribution

- Is given by the combination of exponentials to yield a power law

  $$P(k) \sim k^{-\gamma}$$

  where

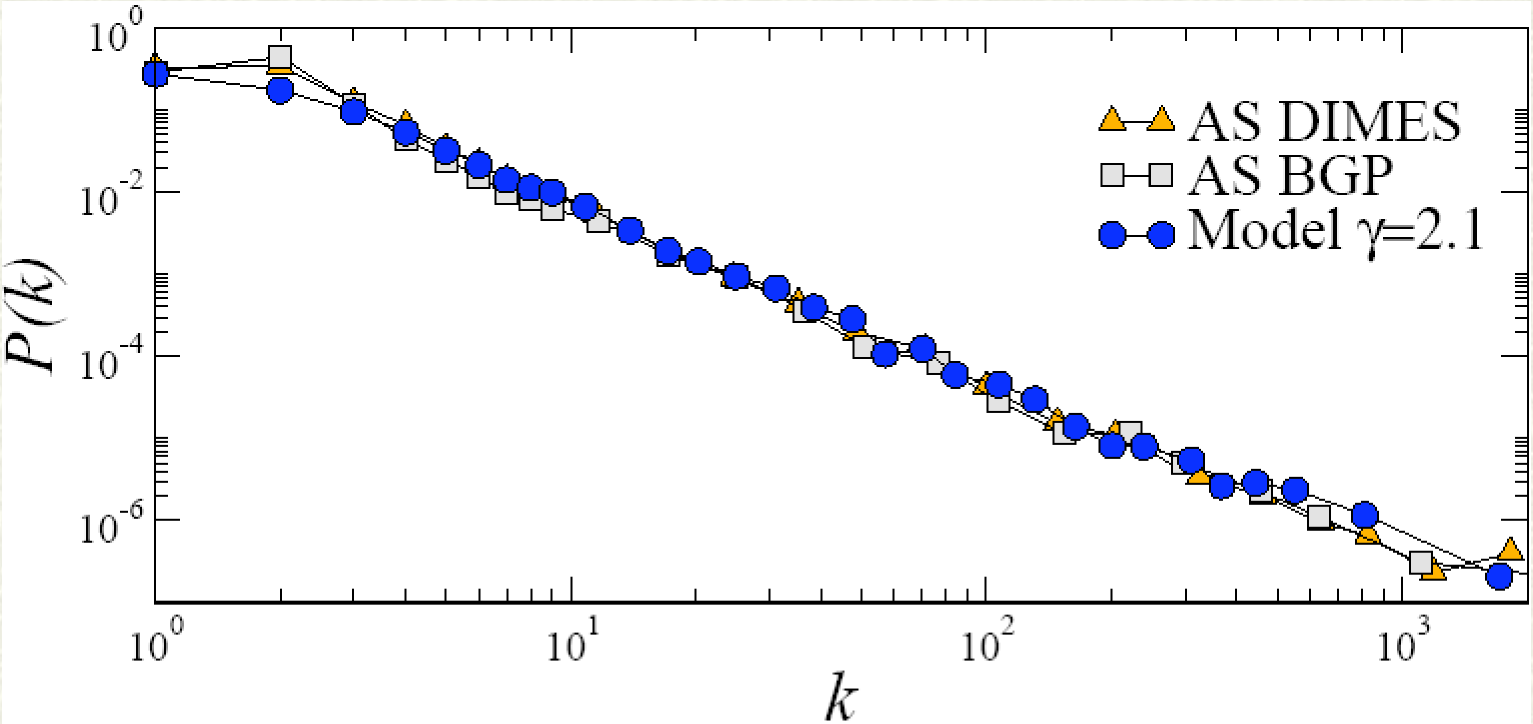
  $\gamma = 1 + \alpha/\beta =$

  2 if $\alpha \leq ½$; or

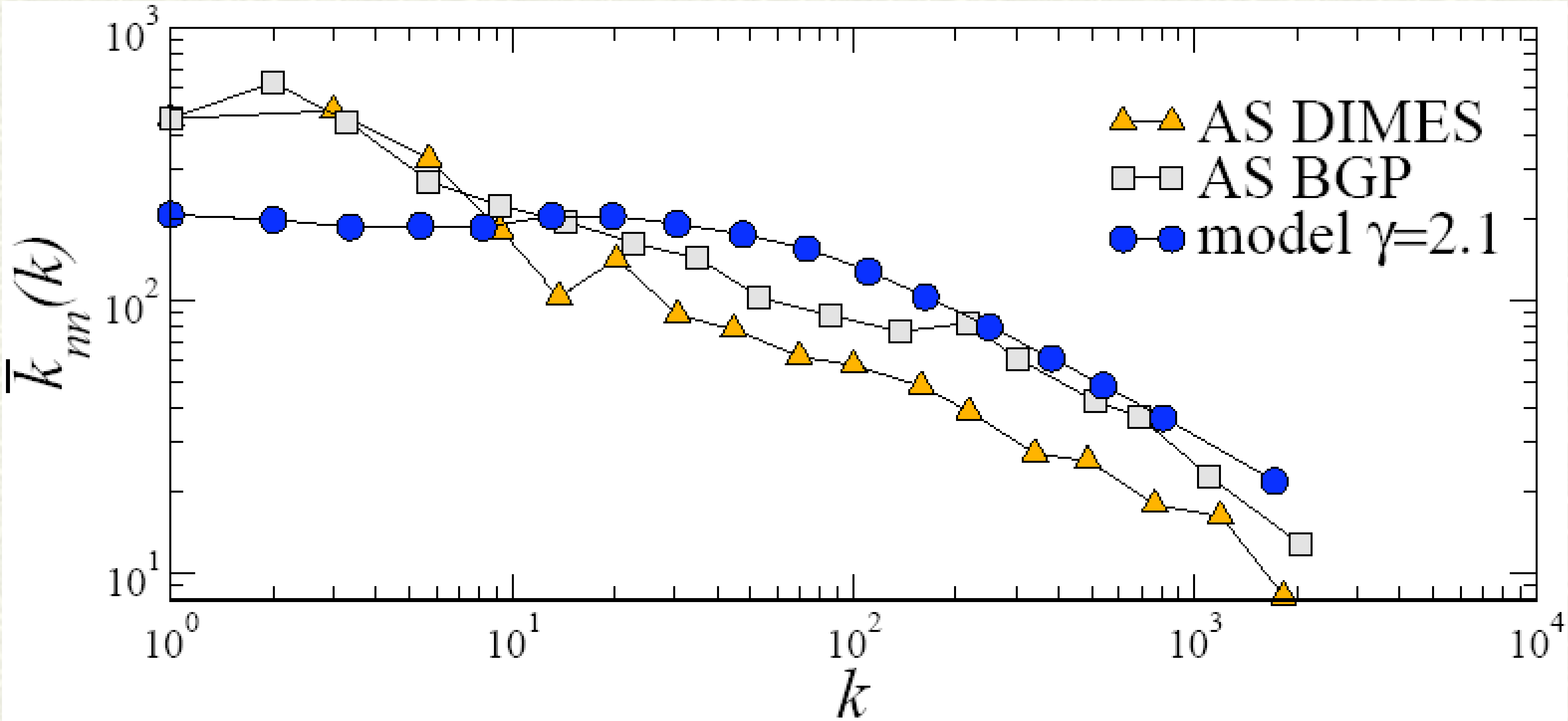
  $2\alpha + 1$ otherwise

- The uniform node distribution in the plane $(\alpha = 1)$ yields $\gamma = 3$

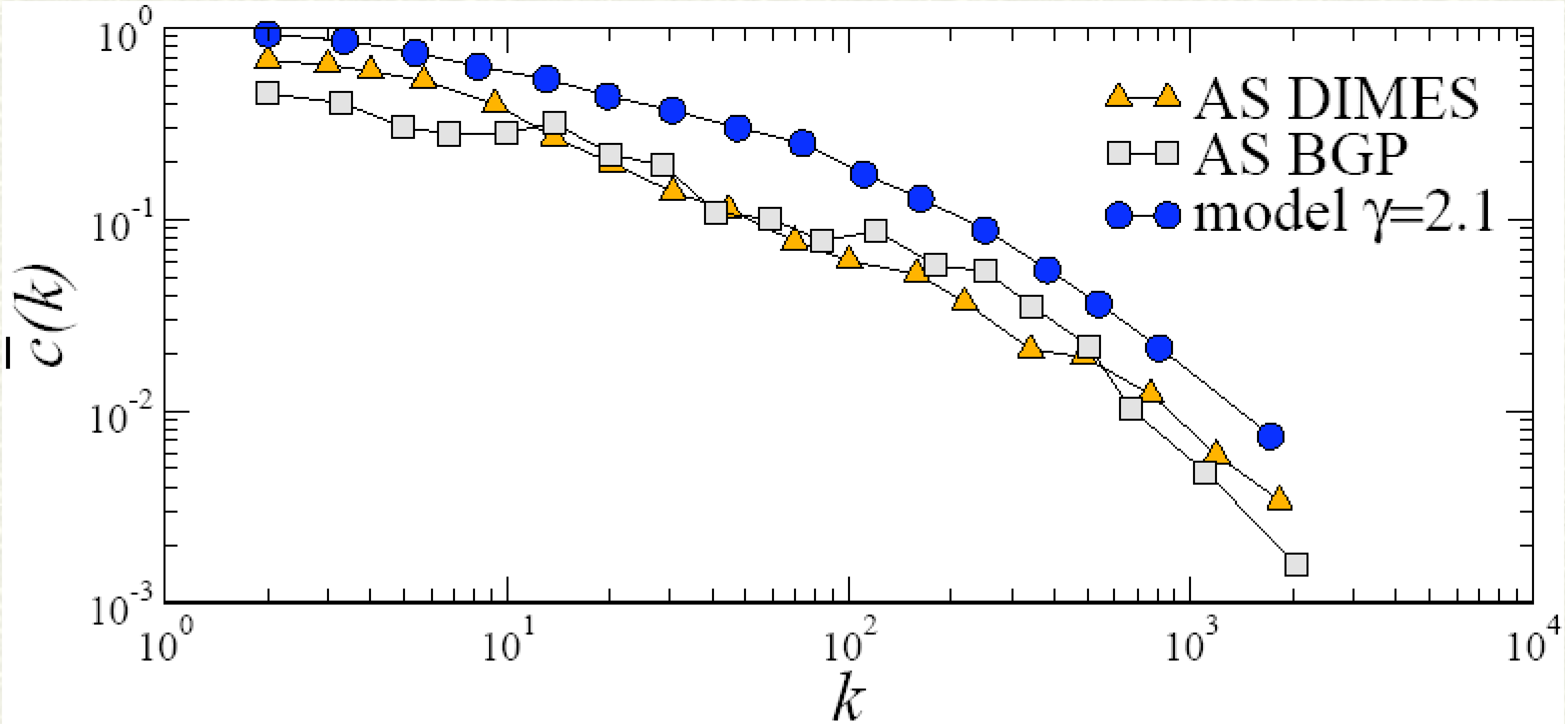
Node degree distribution in modeled and real networks



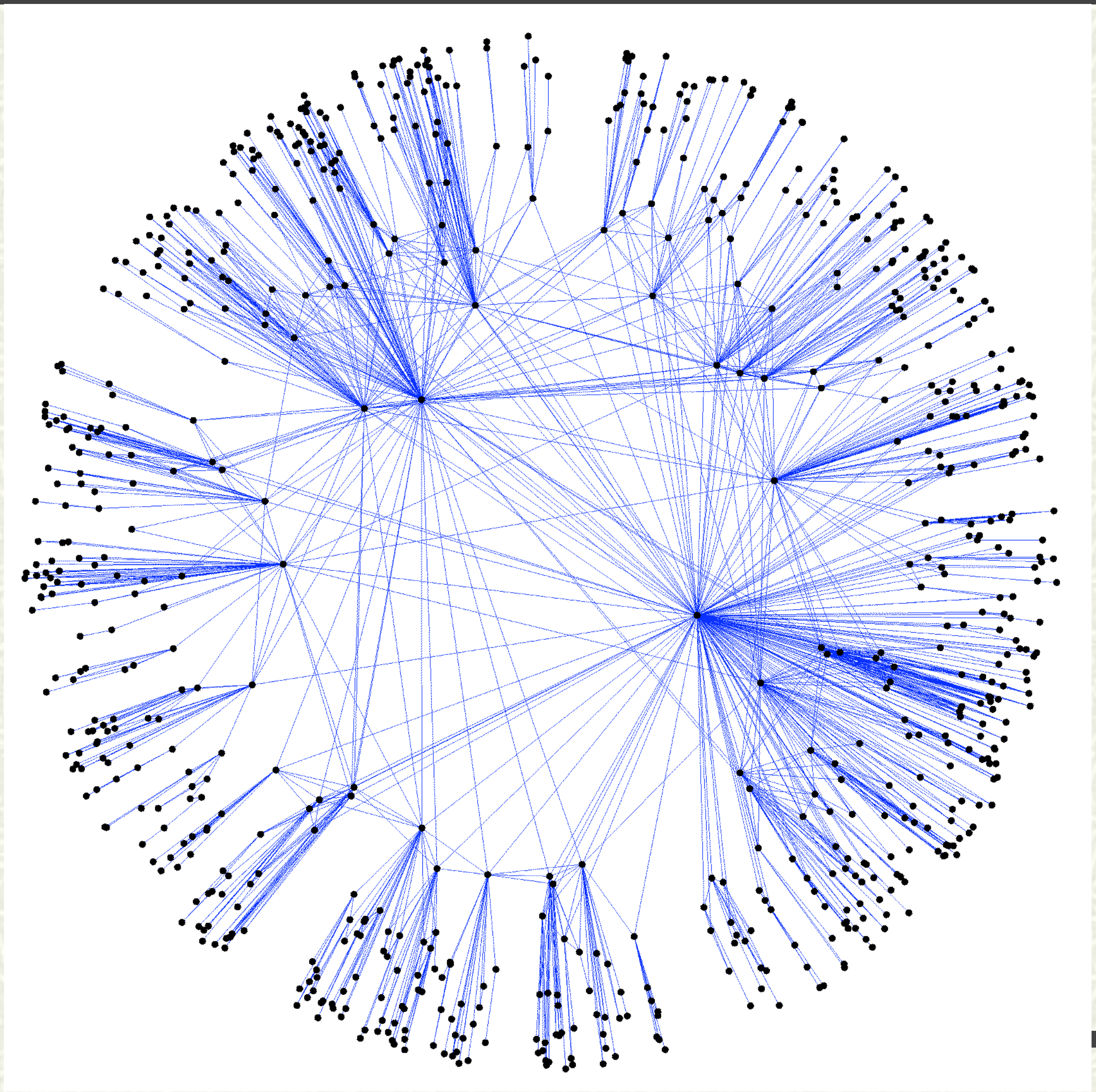
Degree correlations in modeled and real networks



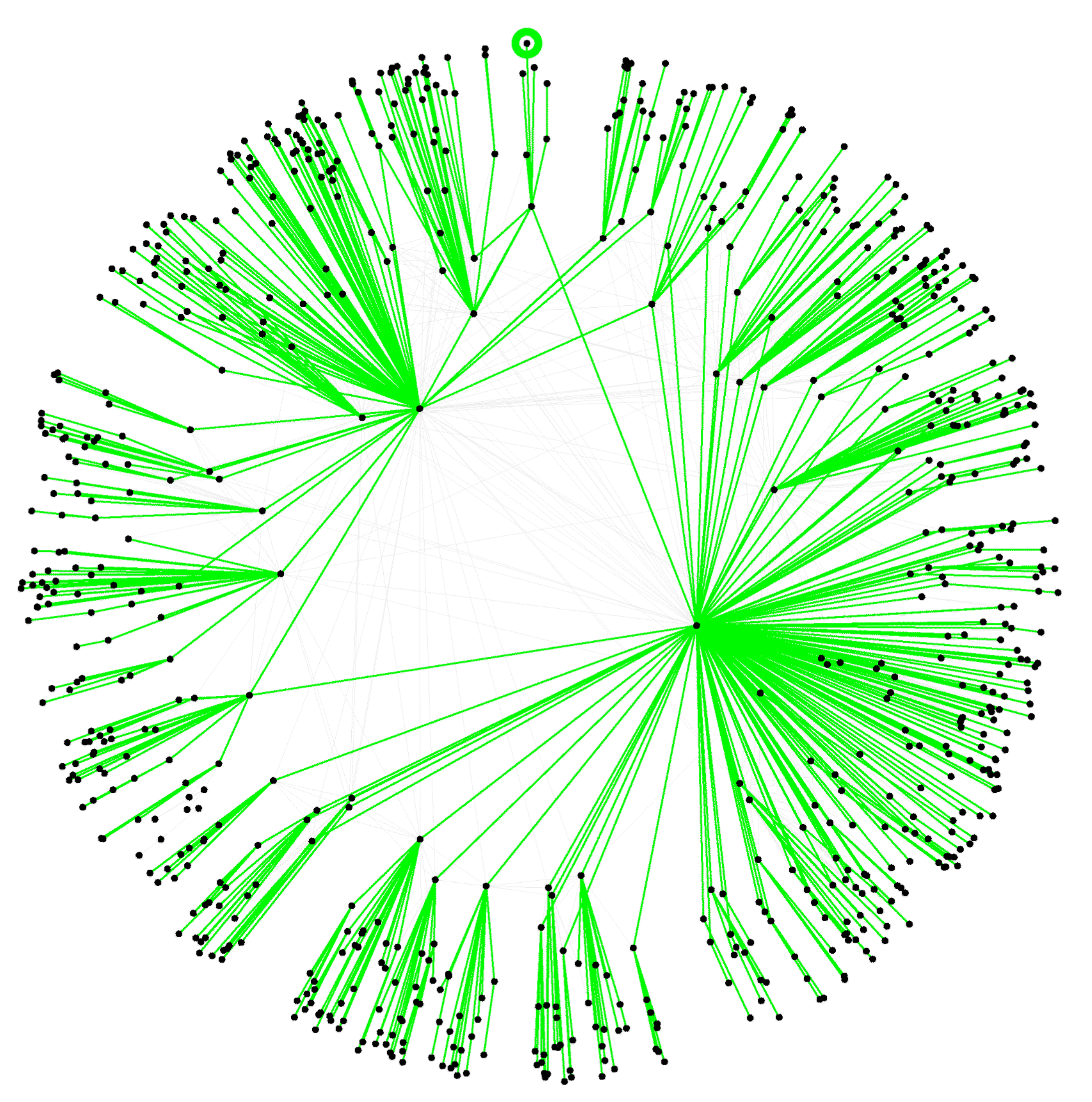
Clustering in modeled and real networks



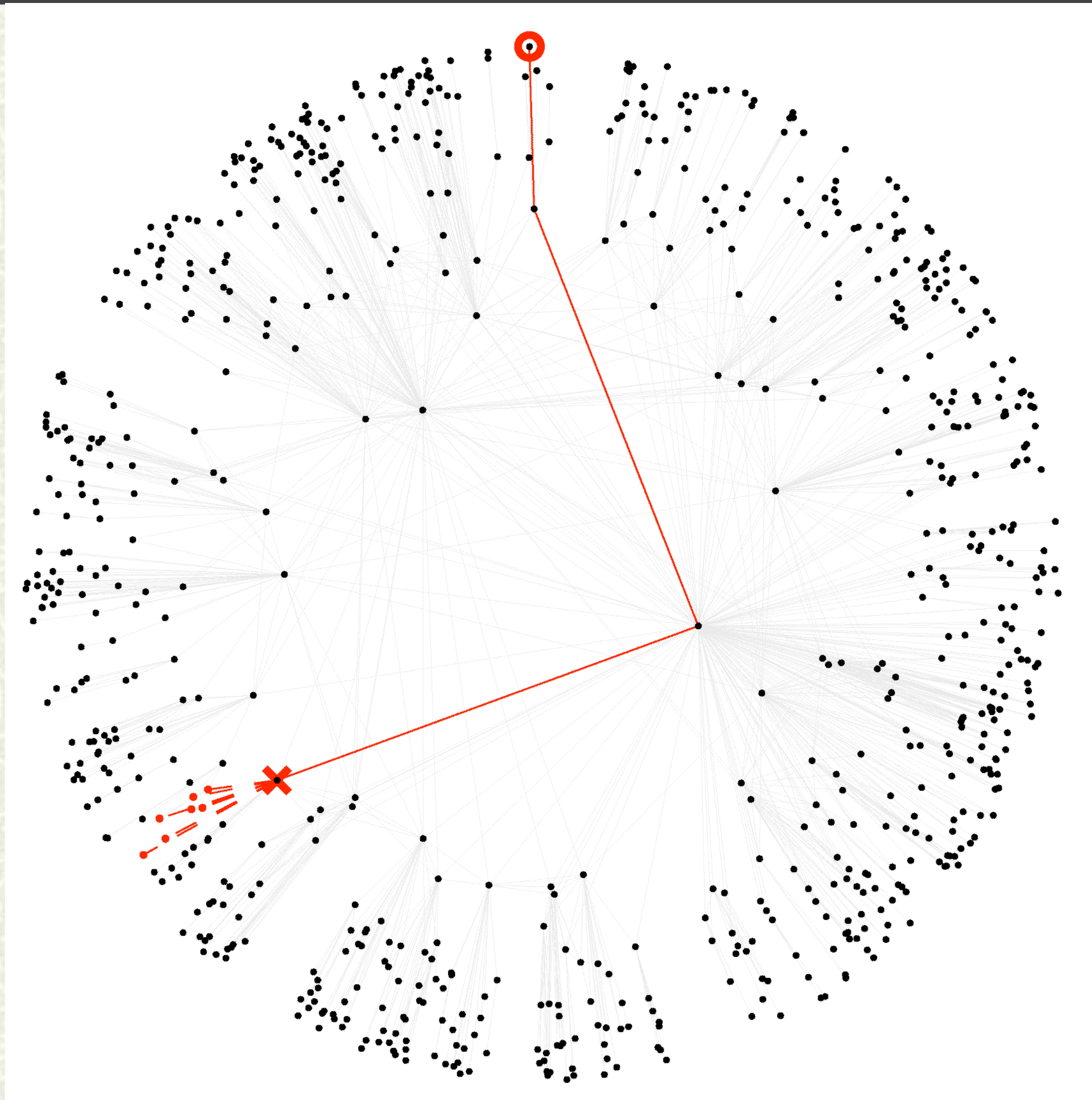
Visualization of a modeled network



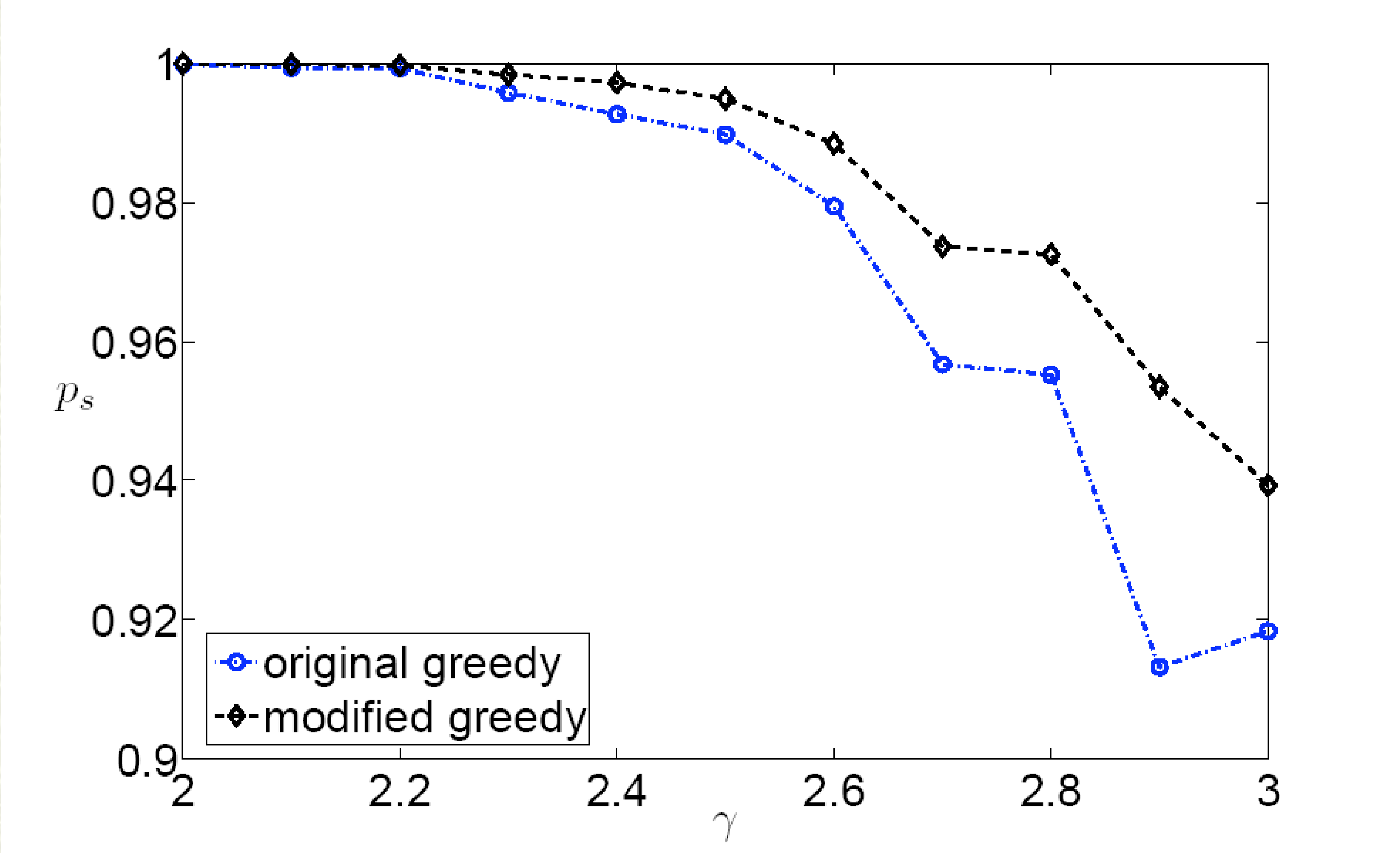
Successful greedy paths



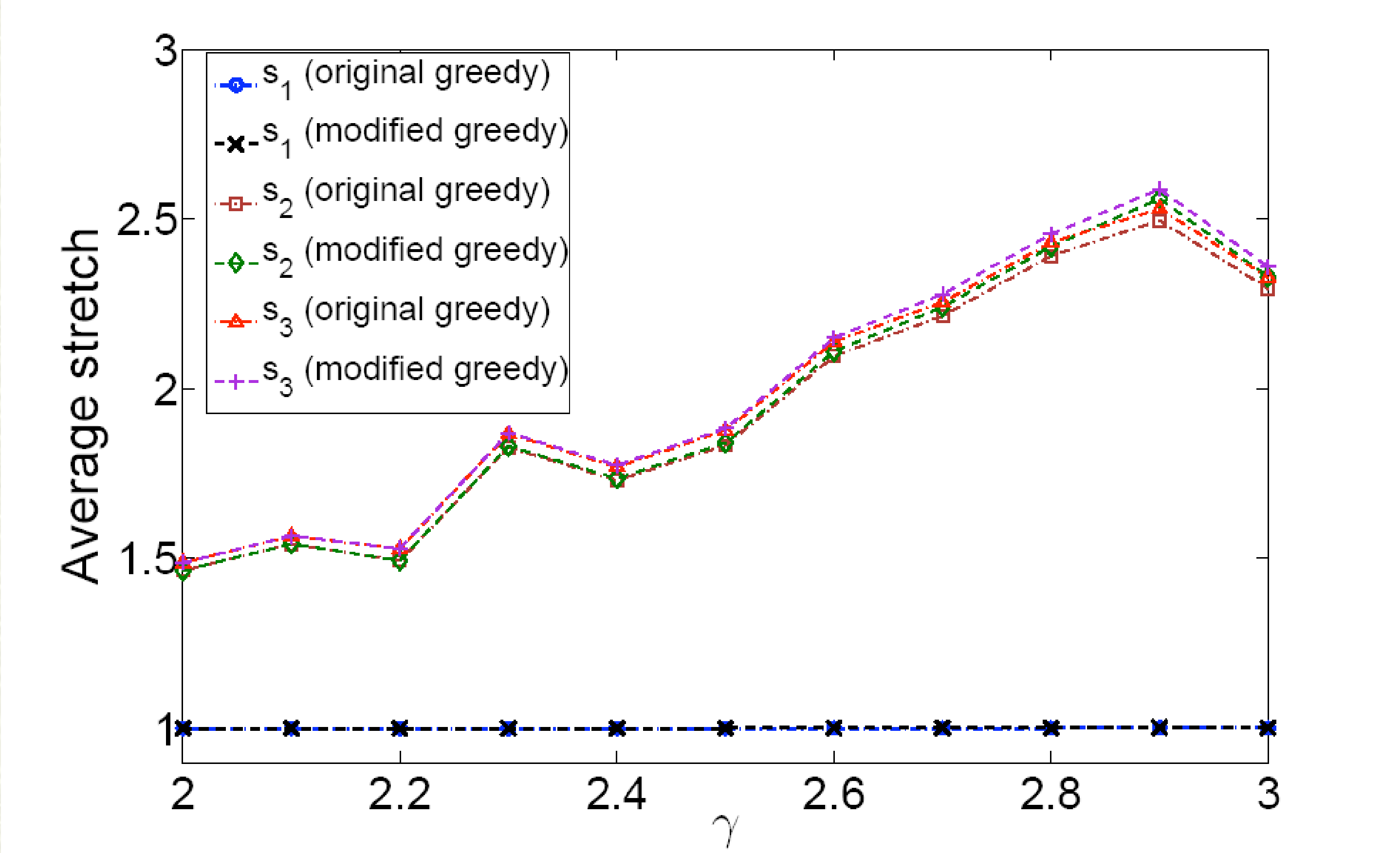
Unsuccessful greedy paths



Percentage of successful paths



Multiplicative stretch



Robustness of greedy routing w.r.t. network dynamics

- As network topology changes, the greedy routing efficiency characteristics deteriorate very slowly
- For example, for $\gamma \leq 2.5$, removal of up to *10%* of the links from the topology degrades the percentage of successful path by less than *1%*

In summary

- Scale-free networks are congruent w.r.t. hidden hyperbolic geometries
- This congruency is robust w.r.t. network dynamics/evolution

Conclusion

- Hidden hyperbolic metric spaces explain, simultaneously, the two main topological characteristics of complex networks
  - scale-free degree distributions
  - strong clustering
- Greedy routing mechanism in these settings may offer virtually infinitely scalable routing algorithms for future communication networks

Problems to solve

- Find the exact structure of hidden metric spaces underlying real networks
- Find the coordinates of nodes in them