Large-scale curvature of networks

Onuttom Narayan[1] and Iraj Saniee[2]
[1]*Department of Physics, University of California, Santa Cruz, California 95064, USA*
[2]*Mathematics of Networks Department, Bell Laboratories, Alcatel-Lucent, 600 Mountain Avenue, Murray Hill, New Jersey 07974, USA*
(Received 26 August 2010; revised manuscript received 21 June 2011; published 13 December 2011)

Understanding key structural properties of large-scale networks is crucial for analyzing and optimizing their performance and improving their reliability and security. Here, through an analysis of a collection of data networks across the globe as measured and documented by previous researchers, we show that communications networks at the Internet protocol (IP) layer possess global negative curvature. We show that negative curvature is independent of previously studied network properties, and that it has a major impact on core congestion: the load at the core of a finite negatively curved network with N nodes scales as N^2, as compared to $N^{1.5}$ for a generic finite flat network.

DOI: 10.1103/PhysRevE.84.066108 PACS number(s): 89.75.Hc

I. INTRODUCTION

Large-scale data networks form the infrastructure for contemporary global communications. Increasingly, a single network may provide a variety of disparate services, flatter architectures (i.e., fewer controlling hubs) are used to achieve robustness against failure, and networks have to be dynamically and automatically reconfigurable to allow services to be set up quickly. With these trends in mind, it is impractical to perform detailed case-by-case simulations in order to predict and understand the behavior of such large-scale networks. Instead, one has to identify the key structural properties that affect network performance, reliability, and security. These structural properties can then be used to construct models that estimate network behavior in an efficient and scalable manner.

A key observation regarding large-scale communications and biological and social networks has been the "small-world" property [1–3]. More recent network models have focused on power-law degree distributions (PLDD) (for a few examples, see Refs. [4–6]) as an explanation of or correlated with the small-world property. Evidence for PLDD has been found in data networks at the Internet protocol (IP) layer [7], for the worldwide web [4], and for the virtual network of social connections [8]. Although these features are interesting and important, the impact of intrinsic geometrical and topological features of large-scale networks on performance, reliability, and security is of much greater importance. Intuitively, it is known that traffic between nodes tends to go through a relatively small core of the network, as if the shortest path between them is curved inward. It has been suggested that this property may be due to *global curvature* or hyperbolicity of the network [9].

In this paper, we consider how global (negative) curvature may be defined for large-scale networks. We use an extension of the definition of global negative curvature for infinite graphs [10]; our extension is slightly different from that proposed by earlier authors [11–13]. We show that the more traditional (Gaussian) curvature of a network, which has been examined earlier [13,14], is distinct from the definition of global curvature we use [10], and that the latter is of greater relevance for network analysis. Most importantly, we examine numerous publicly available networks at the IP layer [15] and demonstrate the existence of negative curvature in the sense defined here. With a few exceptions [14,16]—which use completely different notions of network curvature—the prior literature on network curvature has been based entirely on the study of models [11,17,18].

Second, we examine the impact of negative curvature on the performance of networks by calculating the load at each node, which is the traffic flowing through each node if one unit of traffic flows between each pair of nodes in the network and shortest path routing is used. (This is also called the *betweenness*, and is *not* the actual time-variable demand that is flowing through the nodes.) We show that negative curvature implies that the load at the core of the network scales with the number of nodes N as $\sim N^2$. In contrast, the load in a generic flat network scales as $\sim N^{1.5}$, which grows less rapidly as N is increased, although, as discussed in Sec. III A, it is possible to construct flat networks with $\sim N^2$ scaling of core load. [In this paper, we define the core as the set of m (greater than zero) nodes with the highest traffic. The precise value of m is not very important; although the results shown in Fig. 5 are for $m = 1$, we have verified that the conclusions are unaltered with $m = 10$.] Thus core congestion is worse in hyperbolic networks, and geodesic routing achieved with greedy algorithms on hyperbolic networks [17] is actually problematic. Furthermore, we argue that departure from shortest path routing in large negatively curved networks is liable to result in an unattractively large increase in the average network load. Finally, we provide a "taxonomy" of networks, establishing through examples how their various intrinsic features and properties are related and placing the different common models appropriately in this chart. In particular, we distinguish PLDD and negative curvature as distinct but possibly overlapping features of real networks, and that PLDD is neither required for nor implied by negative curvature. Based on the N^2 scaling of load observed in real networks and inherent in negative curvature, we assert that hyperbolicity is a more relevant companion feature to the small-world property of large-scale networks and that it leads to core congestion.

The rest of this paper is organized as follows. In Sec. II, we briefly review the definition of local curvature for networks and contrast it with the Gromov concept of curvature formally defined for infinite graphs [10]. We then introduce our extension of this notion of curvature to finite graphs and establish that measured networks at the IP layer [15] are hyperbolic. We also present results for three network models and for a few families of graphs. In Sec. III, we show that the load—or betweenness—at the core of a hyperbolic graph with shortest path routing scales as $\sim N^2$, with N being the number of nodes in the graph. In Sec. IV, we provide a taxonomy of networks and their models.

II. NEGATIVE CURVATURE OF NETWORKS

A. Gaussian curvature

The most natural way to define the curvature of a finite graph involves embedding it in a two-dimensional smooth orientable surface and using the Gaussian curvature. The genus g of the graph is the smallest number of handles that such a surface would need to have to allow an embedding of the graph without edges crossing. That such an embedding is always possible is assured by Ref. [19], which also shows that g need not be more than $\lceil (N-3)(N-4)/12 \rceil$ for a graph with N vertices. Further, this embedding can be strengthened so that each vertex is surrounded by polygonal faces (see Ref. [20], Chap. 3). Having completed such a strong embedding, it is now possible to extend the notion of vertex curvature defined for planar graphs (see Ref. [21]) to any finite graph:

$$k_G(v) = 1 - \frac{d(v)}{2} + \sum_{f:v\in f} \frac{1}{|f|}, \tag{1}$$

where v is any vertex, $d(v)$ is its degree, and $|f|$ counts the number of sides of the polygonal face f for all the faces that meet at the vertex v. The (combinatorial) Gauss-Bonnet theorem [21] then gives the sum of all vertex curvatures in terms of $\chi(G)$, the Euler characteristic of the graph:

$$\sum_{v\in G} k_G(v) = \chi(G) = 2 - 2g. \tag{2}$$

The combinatorial curvature above is best understood by assigning a unit length to each edge of the graph and treating the polygonal faces as if they were regular. Simple geometry then shows that the right-hand side of Eq. (1) is equal to the deficit angle at the node v (divided by 2π). The generalization to the case in which the edge lengths are different but non-negative is known as the *discrete curvature* of the graph (see, for example, [22]). It is well-defined when the graph is triangulated, i.e., when all faces on the minimal embedding are broken into triangles, and the edge lengths w_{ij} around each triangular face satisfy the triangle inequality.

The combinatorial and discrete curvatures are defined locally, and are natural extension of the Gaussian curvature to graphs. In Ref. [13], the curvature defined through embedding at the small scale is of this kind, calculated from the deficit angle. Reference [14] constructs a curvature at every node that is qualitatively like the local combinatorial curvature. Despite the appeal of an extended Gaussian curvature, we now provide a simple example showing that it is of limited usefulness for networks if one is interested in the flow of traffic on them.

Consider a square lattice graph, and then add edges along the diagonals of each primitive square. To embed this graph on a surface requires one to pull out one handle (thereby increasing the genus) for each edge crossing. Before the diagonals are added, if periodic boundary conditions are used, the average Gaussian curvature of the graph is zero. After the diagonals are added, the average Gaussian curvature is negative from the combinatorial Gauss-Bonnet theorem. (Similarly, one can add edges connecting each node to its six next-nearest neighbors on a triangular lattice graph to change the curvature defined in Ref. [14], although since their definition is normalized to lie between 0 and 1 there is no sign change.) But despite this change in curvature, the flow of traffic along geodesics (i.e., the shortest paths) is unchanged at a coarse-grained level, and therefore maximal congestion is essentially unaffected. In the next (sub)section, we shall consider an alternative definition of curvature that we show is more appropriate for distinguishing traffic flows on different networks.

B. Large-scale curvature

As an alternative to the combinatorial curvature, negative curvature of an (infinite) geodesic metric space is defined by Gromov [10] in terms of the "δ-thin triangle condition." The definition can be applied to graphs if a metric is provided. For any three nodes (ijk), the geodesics g_{ij}, g_{jk}, and g_{ki} of lengths d_{ij}, d_{jk}, and d_{ki} are constructed. A fourth node m is chosen, and the shortest distance between m and all the nodes on g_{ij} is defined as $d(m;ij)$. The distance $D(m;ijk)$ is defined as the maximum of $d(m;ij)$, $d(m;jk)$, and $d(m;ki)$. For any triangle $\Delta = (ijk)$, we define

$$\delta_\Delta = \min_m D(m;ijk). \tag{3}$$

Then if

$$\delta = \max_\Delta \delta_\Delta < \infty, \tag{4}$$

the graph is said to have negative curvature, i.e., to be δ-hyperbolic. Slightly different definitions of δ_Δ are possible, using the "slim triangle condition" (see, for example, Ref. [23]) or other constructions [13], but they are equivalent. One can show that various other properties that are linked to negatively curved spaces all follow from the thin triangle condition [24,25], which is why it is significant in geometry.

We note that in the example at the end of the previous subsection, Eq. (4) is not satisfied for the square lattice with or without extra diagonals, and therefore the curvature does not become negative when the diagonals are added. This suggests that we might have more success associating traffic on networks with negative curvature as defined by Eq. (4) than with the combinatorial curvature.

For a finite graph, δ in Eq. (4) is trivially finite. Recognizing this, an extension of the criterion for negative curvature in Eq. (4) was proposed in Refs. [11–13]. Using a slightly different definition of δ_Δ, they showed that the ratio of δ_Δ to the perimeter of the triangle Δ is bounded above by $3/2$ if the sides of the triangle curve inward. Therefore, the strict inequality was used as the criterion for the triangle to be in a region of negative curvature. While this provides a quantitative definition of negative curvature for finite graphs, the requirement that *every* triangle should be in a negatively curved region may be unduly restrictive for traffic flow on networks. Furthermore, in the limit of an infinitely large graph

TABLE I. Parameters of all the networks studied from the Rocketfuel database [15]. The column δ_{max} shows the maximum δ_Δ that was found through the random sampling described in Sec. II C.

Network ID	Name	No. nodes	No. edges	Diameter	Avg. geodesic	δ_{max}
1221	Telstra (Aust.)	2998	3806	12	5.53	2
1239	Sprintlink (US)	8341	14025	13	5.18	2
1755	EBONE (US)	605	1035	13	6.00	2
2914	Verio (US)	7102	12291	13	6.04	3
3257	Tiscali (EU)	855	1173	14	5.30	2
3356	Level 3 (US)	3447	9390	11	5.07	2
3967	Exodus (US)	895	2070	13	5.94	3
4755	VSNL (India)	121	228	6	3.20	1
6461	Abovenet (US)	2720	3824	12	5.72	2
7018	AT&T	10152	14319	12	6.95	3

with infinitely large triangles, it is possible for the inequality to be satisfied and yet for Eq. (4) to be violated.

Therefore, in this paper we approach the problem in a slightly different manner and introduce the concept of the "curvature plot" of a network: for every triangle $\Delta = (ijk)$ we calculate δ_Δ and l_Δ, where $l_\Delta = \min[d(ij), d(jk), d(ki)]$. If there are n geodesics between the vertices i and j, the probability of choosing any one is $1/n$. (This is consistent with the manner in which traffic is distributed when there are multiple geodesics between nodes, e.g., in Ref. [26].) This procedure yields $P_l(\delta)$, the probability distribution for δ at fixed l. If $\delta_a(l)$ is the average δ_Δ for all triangles with a given l, we say that the network is negatively curved in the large if and only if $\delta_a(l)$ appears to approach a limiting value as l is increased. Since we use the average of the distribution instead of the maximum as in Eq. (4), a statistical sampling of triangles with the vertices i, j, k chosen independently and randomly from all the nodes in the graph is robust. For a tree graph, it is easy to see that δ_Δ is zero for every triangle, and the graph is hyperbolic. Although one would expect that a flat space would have $\delta_a(l)$ increasing linearly with l, all that we require here is that it should diverge with l [27].

A few comments are in order here. First, because the average value of δ_Δ for all triangles with a given l_Δ is used instead of the maximum, it is possible in principle for an infinitely large graph to satisfy the condition given above and yet fail the test of Eq. (4). However, we believe that if this were to happen, occasional fat triangles would be unlikely to affect the flow of traffic and congestion significantly, and so the condition above is appropriate for our purpose. Second, the condition above is qualitative, and a sufficiently large range of l is required for it to yield conclusive results. Finally, the approach in this paper and in Refs. [11–13] analyzes the curvature of a graph without reference to any underlying manifold in which the graph could be embedded, and it can therefore be used even when an embedding is difficult to achieve. This is in contrast to Ref. [16], where the nodes of the graph are embedded in a d-dimensional hyperbolic space, similar to the strategy of Refs. [17,18], where negative curvature is *assumed* and a model constructed with a few extra simple assumptions is shown to share features such as PLDD with real networks. (Note that, as we will see in Sec. IV, PLDD's are neither necessary nor sufficient for hyperbolicity, so tying the two together is misleading.)

C. Numerical results

In the previous subsection, we have introduced the curvature plot and how it relates to graphs with negative curvature in the large. In this section, we present numerical results for the curvature plots for a variety of graphs: models that are known to be Euclidean or hyperbolic (triangular and square lattices and hyperbolic grids), two random network models (the Watts-Strogatz and Erdos-Renyi models), and real communications networks at the IP layer from the Rocketfuel database [15] that are cataloged in Table I. As mentioned in the Introduction, real networks have not been studied in much of the prior work on network curvature.

Figure 1 is a curvature plot for all these graphs. Triangles were randomly sampled to construct the curvature plot for each graph: for each triangle, all the vertices were chosen independently and randomly, with all nodes in the graph being equally likely. If two or three vertices of a triangle were coincident, the triangle was rejected and a new triangle was chosen. A total of 32×10^5 triangles were sampled for each graph to construct its curvature plot. (There was no significant

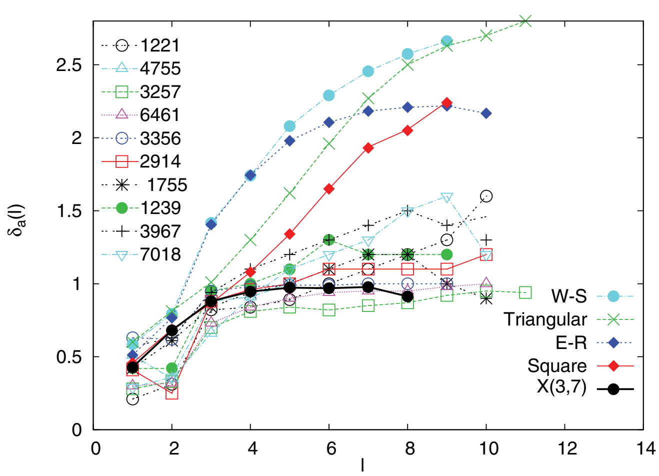

FIG. 1. (Color) The average δ as a function of l, $\delta_a(l)$, for the 10 IP-layer networks from the Rocketfuel database. Also shown are $\delta_a(l)$ for the Watts-Strogatz (WS) and Erdos-Renyi (ER) models, for a triangular and square lattice, and for the hyperbolic grid $X(3,7)$. The parameters used for the first ten and last six curves are given in Tables I and II, respectively.

TABLE II. Parameters for all the models in Fig. 1. For the first three, the results are the averages of 80 random graphs. For the last three, open boundary conditions are used. Only the largest component of each graph of 2000 nodes was used for the Erdos-Renyi model, so that the number of nodes was less than 2000 and was different for each graph. The column δ_{max} shows the maximum δ_Δ that was found through the random sampling described in Sec. II C.

Model	Parameters	No. nodes	No. edges	Diameter	Avg. geodesic	δ_{max}
Erdos-Renyi	$p = 4/N$	1960	7987	12.0	5.61	4
Watts-Strogatz	$p = 0.2$	1225	5880	9.89	5.69	4
Triangular lattice		127	612	12	5.95	4
Square lattice		64	224	14	5.33	3
Hyperbolic grid	$p = 3, q = 7$	232	798	8	5.82	2

change if this was increased tenfold.) For the three random network models, the average of 80 independent random graphs was taken. The probability of any edge being connected was $4/N$ for the Erdos-Renyi model, so that the average degree of the nodes was 4. For the Watts-Strogatz model, an extra 20% edges were added randomly to a square lattice with periodic boundary conditions.

As is seen from Fig. 1, even the plots for the Euclidean lattice graphs show some rounding near the top due to finite-size effects. Finite-size effects should be visible for triangles whose linear dimensions are comparable to those of the graph from which they are drawn. Accordingly, the parameters of the models in Fig. 1 are chosen so that all the graphs shown in the figure have a comparable average geodesic length (except the Rocketfuel network 4755 in Table I). Table II lists the parameters of all the models shown in Fig. 1.

Figure 1 shows a clear separation between the curvature plots for the Euclidean lattices and the hyperbolic grid $X(3,7)$. [The hyperbolic grid $X(p,q)$ is a tessellation of the Poincaré disk using identical p-gons, with q of the p-gons meeting at each node and $(p-2)(q-2) > 4$; projected onto the unit disk, the p-gons become progressively smaller as one moves further from the center.] The curves for the Euclidean lattices rise linearly with some rounding near the top due to finite-size effects, while the curve for the hyperbolic grid saturates rapidly. The figure also shows the curvature plots for the Rocketfuel networks We see that they are all clustered around the curve for the hyperbolic grid. Finally, the figure shows the curvature plot for the Erdos-Renyi and Watts-Strogatz models. Although these do not rise as far as the Euclidean lattices before leveling off, they are separated from the cluster of hyperbolic curves. A detailed numerical and analytical proof that Erdos-Renyi graphs (with fixed average degree) are not δ-hyperbolic has been recently obtained [28], and therefore we group the Erdos-Renyi curvature plot with those for the Euclidean lattices. Since the Watts-Strogatz model is an interpolation between a Euclidean lattice at short length scales and an Erdos-Renyi model at long length scales, we conclude that it too is not δ-hyperbolic.

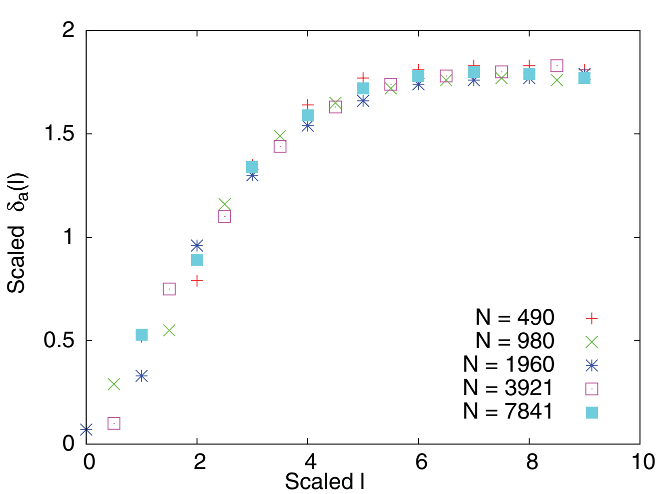

FIG. 2. (Color online) Scaling collapse of the curvature plot for Erdos-Renyi graphs with $p = 4/N$ and various values of N. Only the giant component was retained in the analysis; the N values shown in the figure are the average number of nodes in the giant component when the graph has $500, 1000, 2000, \ldots$ nodes. To achieve the scaling collapse, each successive curvature plot was shifted down by 0.22 and to the left by 0.5 relative to its predecessor.

As further evidence of this, Fig. 2 shows a scaled curvature plot for Erdos-Renyi random graphs [1] for various N with $p = 4/N$, the same value of p used in Table II and Fig. 1. Curvature plots for various system sizes collapse onto a single curve when they are shifted downward and to the left linearly with $\ln N$. This demonstrates that the leveling off of the curvature plot is a finite-size effect, and is pushed out to infinity as $\sim \ln N$. The $\sim \ln N$ dependence is expected, since the characteristic length scales in these graphs grow logarithmically with N. Figure 3 shows similar plots for the square lattice with open boundary conditions and the triangular lattice, also with the parameters given in Table II and Fig. 1. The curvature plot is shifted downward and to the left by an amount proportional to the diameter of each graph to achieve the same scaling collapse. Finally, Fig. 4 shows curvature plots for hyperbolic grids with $p = 3$, $q = 7$, and various N. In this case, no rescaling is performed, and the plateaus for all the system sizes are seen to coincide, confirming that this model is δ-hyperbolic. A scaling collapse for the Watts-Strogatz model is not provided, for the same reason as given in the previous paragraph.

III. TRAFFIC IN HYPERBOLIC NETWORKS

A. Core congestion

We now turn to the performance implications of δ-hyperbolicity. As discussed in the Introduction, we consider traffic flowing on finite graphs with unit demand between each pair of nodes. The traffic between any pair of nodes travels

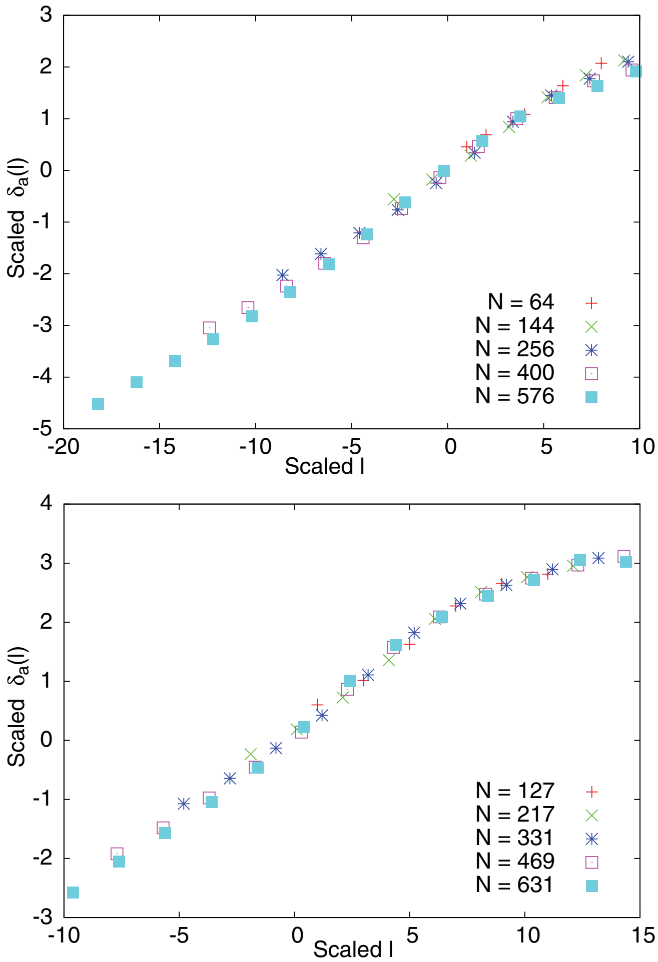

FIG. 3. (Color online) Scaling collapse of the curvature plot for a square lattice with open boundary conditions (first panel) and a triangular lattice (second panel) for various values of N. The square lattice graphs have side $8,12,16,\ldots$, and each successive curvature plot is shifted down by 1.25 and to the left by 4.8 relative to its predecessor. The triangular lattices are built outward to a distance $6,8,10,\ldots$ from a central node, and each successive curvature plot is shifted down by 0.85 and to the left by 2.9 relative to its predecessor.

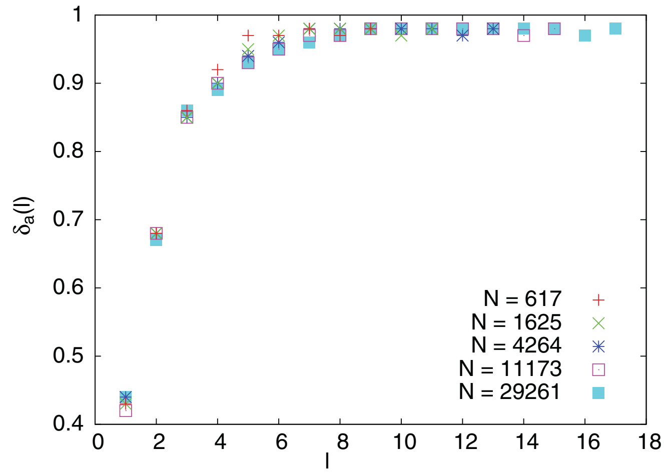

FIG. 4. (Color online) Curvature plots for hyperbolic grids with $p = 3$ and $q = 7$. The graphs are built outward to a distance $5,6,7,\ldots$ from a central node. The curvature plots saturate at the same height without any rescaling.

along the geodesic connecting them (evenly distributed over all geodesics in the case of ties). For the models, we construct graphs with the number of nodes N varying and study how the traffic through the core scales with N. For the Rocketfuel networks, the range of N provided by the ten different networks in the database is used to study the same question.

Although the core has been defined in terms of the traffic that flows through it, for all the graphs we have examined we have verified that the core is close to the geometric center, defined as the node whose average (geodesic) distance to all the other nodes in the graph is the smallest. For the Euclidean lattices and hyperbolic grid, symmetry requires that the core should be centered at the geometric center.

Physically, we expect that the core load $L_c(N)$ should scale as

$$L_c(N) \sim N^2 \tag{5}$$

for the hyperbolic graphs and $L_c(N) \sim N^{1.5}$ for the Euclidean graphs. This is because all traffic from the $\sim N$ nodes on the left of a Euclidean lattice to the $\sim N$ nodes on the right flows through the center across a line of length $\sim\sqrt{N}$, whereas for a hyperbolic graph the same traffic is pulled inward and flows within an $O(1)$ distance of the center. Mathematically, Eq. (5) can be obtained analytically for the continuum Poincaré disk truncated to a radius $r < 1$, converted to a graph by introducing a uniform distribution of nodes with each node connected to its neighbors. (This model is applicable to a hyperbolic network if the spacing between nodes is small compared to its radius of curvature.) In contrast, for a Euclidean disk with area N one can verify that $L_c(N) \sim N^{1.5}$.

It is well recognized that shortest path routing results in some highly congested nodes, and the connection between topology and congestion has been made in Refs. [11–13,29,30], but the precise form of the scaling for hyperbolic networks that we report in this paper has hitherto gone unnoticed. After the initial report of our work [31], a mathematical justification was proposed for the $\sim N^2$ and $\sim N^{1.5}$ scaling reported herein [32].

For tree graphs, which are trivially δ-hyperbolic, it is easy to see that the traffic flowing through a node that divides a graph with N nodes into two equal regions is $\sim N^2$. However, hyperbolic graphs can be quite far from being trees; indeed, all the nodes in a hyperbolic grid $X(p,q)$ have the same degree q. For the Rocketfuel networks, the ratio of the number of undirected edges to nodes ranges from 1.27 to 2.72; for comparison, the ratio for a tree and a square lattice is 1.0 and 2.0, respectively. Therefore, we investigate whether the results of the previous paragraph are true for the graphs we have studied in Sec. II.

Figure 5 shows $L_c(N)$ versus N for all the networks in the Rocketfuel database, demonstrating $\sim N^2$ scaling. The figure also shows results for the two random network models, which do not show $\sim N^2$ scaling. Similar results are obtained by taking the average load at the top ten nodes.

Thus hyperbolicity can significantly impact the performance of communications networks. Although the communications networks we have studied have a PLDD and

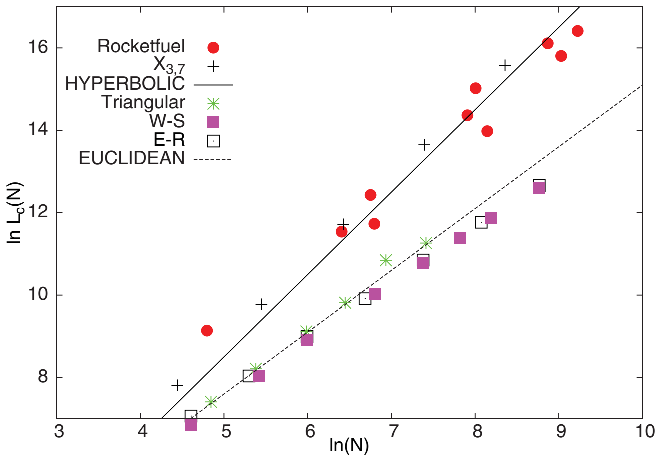

FIG. 5. (Color online) Plot of the maximum load $L_c(N)$ for each network in the Rocketfuel database as a function of the number of nodes N in the network. Also shown are the maximum load for the hyperbolic grid $X(3,7)$, the Watts-Strogatz model, and a triangular lattice for various N. The lines have slopes of 2.0 and 1.5, corresponding to the hyperbolic and Euclidean cases, respectively.

satisfy the small-world property, in Sec. IV we will see that δ-hyperbolicity is distinct from these.

Although hyperbolicity is sufficient to cause $\sim N^2$ scaling of the core load in (finite) networks, it is not necessary. It is easy to construct graphs in which narrow bridges connect large islands, with the size of each island increasing uniformly as N is increased. Clearly, a finite fraction of the total traffic passes through each bridge, even though each island—and therefore the graph—can have zero curvature. It remains an open question whether there is a broader category that includes δ-hyperbolic graphs, these "island-bridge" graphs, and perhaps other graphs as well, and it is both necessary and sufficient that a graph should belong to this category in order that its core load should scale as $\sim N^2$.

B. Other traffic patterns

The $L_c(N) \sim O(N^2)$ scaling obtained in the previous subsection is equivalent to the observation that a fixed percentage of *all* the network load is routed through the network core. Given the severity of this observation, one might ask if the scaling is specific to shortest path routing and would not occur if congestion avoidance algorithms are used. Indeed, various algorithms [33–38] have been proposed to alleviate the congestion that arises with shortest path routing. As a special case of such techniques, one can use shortest path routing but with different weights for the edges of the graph. For example, the inverse of the bandwidth or other engineered choices can be used for edge weights to achieve a balanced load, capacity optimization, or other traffic engineering purposes. In fact, scenarios in which edge weights are not all 1 are more the norm than the exception (see, for example, [39] and ensuing publications on Internet traffic engineering), and so this question is very relevant.

However, even in such cases, a trade-off has to be made between the level of load reduction from $O(N^2)$ and the length of new paths used, as measured by the number of nodes traversed. This is a consequence of a property of δ-hyperbolic spaces according to which a deviation d from a geodesic path results in an increased path length that is exponential in d. More precisely [40], a path between a pair of nodes, whose distance from the geodesic path between the same nodes is d, has a length at least $2^{(d-1)/\delta}$. In order to distribute the $O(N^2)$ paths through the core through a region of size N^α (to significantly reduce core congestion), the radius of the region has to be $\sim \ln N$, and therefore one would expect d to be $\sim \ln N$. This will increase the average path length from $\sim \ln N$ to $\sim N^\beta$ with $\beta > 0$, destroying the small-world property. With N^2 traffic streams and uniform demand, the average load at a node increases from $\sim N \ln N$ to $\sim N^{1+\beta}$. If β is small, congestion avoidance may be useful when N is not very large, but it becomes unattractive in the large-N limit.

In sum, link-metric engineering, even if successful in reducing the worst loads, is not a panacea, since—even apart from the caveat of [41]—one cannot significantly eliminate the $O(N^2)$ core congestion without losing the much prized small-world property of these networks.

In the discussion above, we defined the load by assuming one unit of traffic between all node pairs. Thus the load at any node is a geometric property of the graph, being equal to the number of geodesics passing through it, i.e., its betweenness. The actual traffic patterns in a communications network may be different; in particular, if loads are only nonzero between nodes and their close neighbors, the $L_c(N) \sim N^2$ scaling may be modified. However, we are not aware of any publicly available and detailed measurements of node-pair traffic for any realistic network, and therefore an analysis using actual traffic patterns is difficult. Furthermore, since network traffic is expected to evolve rapidly, the conclusions from actual traffic patterns may be even less robust than the results we have obtained here.

IV. TAXONOMY OF NETWORKS

In this section, we summarize the connections between the various network models and the features associated with each model and the IP layer networks we have studied. The results are shown in Fig. 6.

The placement of the two random network models follows from the previous sections in this paper: they are small-world networks, but not δ-hyperbolic. Both Euclidean and hyperbolic grids lie outside the PLDD region (since all nodes have the same degree), but only the hyperbolic grids are small-world and δ-hyperbolic. Bethe lattices and power-law trees are both small-world graphs and (being trees) are trivially hyperbolic, but only power-law trees have a PLDD. This justifies the location of all these points in Fig. 6.

Chains refer to minimal spanning trees of Euclidean lattices. For example, for a square lattice, all edges on the x axis or parallel to the y axis are connected. (An even simpler example is a line chain.) Being trees, they are δ-hyperbolic, but they clearly do not have a PLDD. The number of nodes $n(r)$ within a distance r of a node is proportional to r^2 for large r, and therefore these graphs do not have the small-world property either.

The "hairy" graphs require further explanation. They are constructed by taking a graph and adding leaf nodes (of degree 1) connected to each node in the original graph. The number of leaf nodes connected to each node in the original graph is an

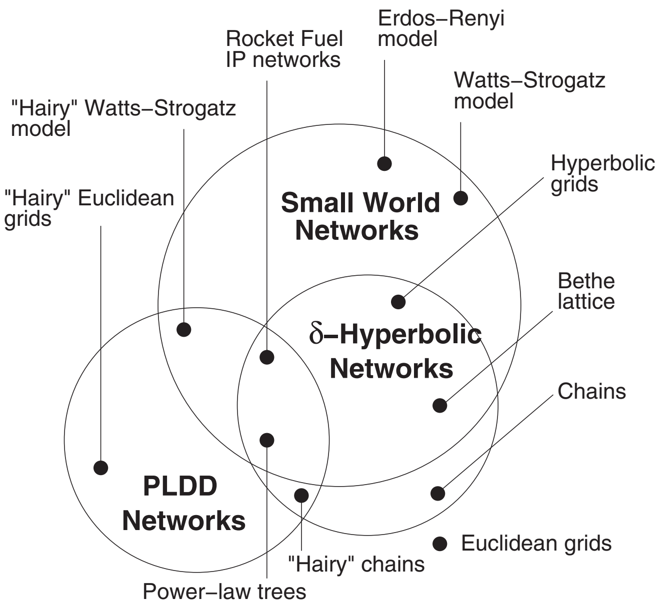

FIG. 6. Taxonomy of key characteristics of networks and their overlaps in a schematic diagram. "Hairy," as used in this figure, refers to the simple mechanism of making a grid power-law by adding to each node a set of singly connected nodes (hairs) whose number is drawn from any desired power-law distribution. PLDD refers to power-law degree distributions.

independent random variable drawn from a distribution with a power-law tail. By construction, the resultant hairy graph has a degree distribution with a power-law tail. On the other hand, geodesics are essentially unaffected by the addition of the hairs, except that an extra edge has to be added at the beginning (end) of the geodesic when it begins (ends) at a leaf node. Therefore, geodesic triangles are δ-thin if and only if they were that way in the original graph. One can also verify that the small-world property of a graph is unaffected by the addition of hairs: if there are $n(r)$ nodes within a distance r of a node in the original graph, the corresponding number for the hairy version is the sum of $n(r)$ random variables. In particular, the hairy Euclidean grid has a PLDD, but it is not small-world and not δ-hyperbolic because the Euclidean grid is neither. Similarly, the hairy Watts-Strogatz model has a PLDD and is small-world but not δ-hyperbolic like the original Watts-Strogatz model.

Since all the regions in Fig. 6 have been populated, they must all have nonzero size. The three circles are overlapping but distinct; networks that are δ-hyperbolic but not small-world are also known as "elementary" δ-hyperbolic networks. We conclude that δ-hyperbolicity is a nontrivial feature that is distinct from both small-world property and power-law degree distributions. Based on the $O(N^2)$ scaling of Sec. III, this property entails features that are important within the networking context. The networks we have studied from the Rocketfuel database appear to be δ-hyperbolic and small-world and also have a (limited) power-law degree distribution.

V. CONCLUSION

In this paper, we have examined the curvature of networks and extended the Gromov [10] criterion for negative curvature of infinite graphs to finite graphs. We have shown that the maximum traffic congestion in networks with negative curvature scales with the number of nodes N as $\sim N^2$, in contrast to the $\sim N^{1.5}$ scaling for planar graphs. We have verified that physical networks at the IP layer are negatively curved (δ-hyperbolic) and show $\sim N^2$ scaling of the core load. We have established a taxonomy of networks that shows that δ-hyperbolicity is distinct from other properties of networks that have been examined in the past. Even though carefully traffic-engineered link metrics may alleviate core congestion and remove points of maximum vulnerability, both in terms of security and reliability, this is achieved at the expense of very long paths (and consequently higher costs) and an effective loss of the prized small-world property that has simplified routing considerably.

ACKNOWLEDGMENTS

This work was funded by AFOSR Grants No. FA9550-08-1-0064 and No. FA9550-11-1-0278.

[1] P. Erdos and A. Renyi, Publ. Mathematicae **6**, 290 (1959).
[2] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 409 (1998).
[3] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[4] R. Albert, H. Jeong, and A-L. Barabasi, Nature (London) **401**, 130 (1999).
[5] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).
[6] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **64**, 066110 (2001).
[7] M. Faloutsos, P. Faloutsos, and C. Faloutsos, in *Proceedings of ACM SIGCOMM'99* (ACM, New York, NY, 1999), p. 251.
[8] A-L. Barabasi, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek, Physica A **311**, 590 (2002).
[9] Y. Baryshnikov (unpublished).
[10] M. Gromov, in *Essays in Group Theory*, Mathematical Sciences Research Institute Publication No. 8 (Springer, New York, 1987), p. 75.
[11] E. A. Jonckheere, P. Lohsoonthorn, and F. Ariaei, Appl. Math. Comput. **192**, 191 (2007).
[12] E. Jonckheere, P. Lohsoonthorn, and F. Bonahon, J. Graph Theor. **57**, 157 (2007).
[13] E. Jonckheere and P. Lohsoonthorn, *Proceedings of the 2004 American Control Conference* (IEEE, Piscataway, NJ, 2004).
[14] J.-P. Eckmann and E. Moses, Proc. Natl. Acad. Sci. USA **99**, 5825 (2002).
[15] N. Spring, R. Mahajan, and D. Wetherall, in *Proceedings of the ACM/SIGCOMM '02* (ACM, New York, NY, 2002), p. 133. Data archive at http://www.cs.washington.edu/research/networking/rocketfuel/.
[16] Y. Shavitt and T. Tankel, *Proceedings of the IEEE INFOCOM* (IEEE, Piscataway, NJ, 2004); IEEE/ACM Trans. Netw. **16**, 1 (2008).
[17] D. Krioukov, F. Papadopoulos, M. Boguna, and A. Vahdat, e-print arXiv:0805.1266 (2008).

[18] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguna, Phys. Rev. E **80**, 035101 (2009); D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguna, *ibid.* **82**, 036106 (2010).
[19] G. Ringel and J. Youngs, Proc. Natl. Acad. Sci. USA **60**, 438 (1968).
[20] B. Mohar and C. Thomassen, *Graphs on Surfaces* (Johns Hopkins University Press, Baltimore, 2001).
[21] Y. Higuchi, J. Graph Theor. **38**, 220 (1999).
[22] X. Yin *et al.*, in *Proceedings of Emerging Trends in Visual Computing 2008* (Springer, New York, 2009), p. 38.
[23] M. Bridson and A. Haefliger, *Metric Spaces of Non-positive Curvature* (Springer, Berlin, 1999).
[24] B. H. Bowditch, Lecture Notes 1, http://www.warwick.ac.uk/masgak/papers/bhb-hypnotes1.pdf (1990).
[25] B. H. Bowditch, Lecture Notes 2, http://www.warwick.ac.uk/masgak/papers/bhb-hypnotes2.pdf (1990).
[26] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).
[27] By choosing m in Eq. (3) to be on the shortest side of the triangle, it is easy to verify that $\delta_\Delta \leqslant l_\Delta/2$, so that positive curvature $[\delta_a(l)/l$ increasing faster than linearly with l for large $l]$ cannot be seen.
[28] O. Narayan, I. Saniee, and G. Tucci, e-print arXiv:1009.5700 (2010).
[29] S. Sreenivasan, R. Cohen, E. Lopez, Z. Toroczkai, and H. E. Stanley, Phys. Rev. E **75**, 036105 (2007).
[30] M. Lou, Ph.D. thesis, University of Southern California, 2008, http://gradworks.umi.com/3341712.pdf.
[31] O. Narayan and I. Saniee, e-print arXiv:0907.1478 (2009).
[32] E. Jonckheere, M. Lou, F. Bonahon, and Y. Baryshnikov, e-print arXiv:0911.2538 (2009).
[33] M. Schafer, J. Scholz, and M. Greiner, Phys. Rev. Lett. **96**, 108701 (2006).
[34] P. Echenique, J. Gomez-Gardenes, and Y. Moreno, Phys. Rev. E **70**, 056105 (2004).
[35] L. Danon, A. Arenas, and A. Diaz-Guilera, Phys. Rev. E **77**, 036103 (2008).
[36] Z.-X. Wu, W.-X. Wang, and K.-H. Yeung, New J. Phys. **10**, 023025 (2008).
[37] B. Danila, Y. Yu, J. A. Marsh, and K. E. Bassler, Phys. Rev. E **74**, 046106 (2006).
[38] B. Danila, Y. Yu, J. A. Marsh, and K. E. Bassler, Chaos **17**, 026102 (2007).
[39] B. Fortz and M. Thorup, in *Proceedings of the IEEE INFOCOM*, Vol. 2 (IEEE, Piscataway, NJ, 2000), p. 519.
[40] See Ref. [23], Proposition 1.6 on p. 400.
[41] Z. Wang and J. Crowcroft, ACM SIGCOMM Comput. Commun. Rev. **22**, 63 (1992).