A HYPERBOLIC GEOMETRY APPROACH TO MULTIPATH ROUTING

E. A. Jonckheere*, P. Lohsoonthorn †

* Department of Electrical Engineering–Systems
University of Southern California
Los Angeles, CA 90089-2563
fax: (213) 821 1109
e-mail: jonckhee@eudoxus.usc.edu

† MED2002 Conference
e-mail: med2002@isr.ist.utl.pt
http://www.isr.ist.utl.pt/med2002/

Keywords: Computer network security, hyperbolic graphs.

Abstract

A hyperbolic geometry method is devised for routing packets of the same message along multiple paths. The primary motivation is to defeat the attack scenario in which a link is compromised by "eavesdropping." Rerouting is restricted to be along near optimal paths to mitigate out-of-order packet arrival, against which TCP is not quite robust. On a hyperbolic graph, near optimal paths are called quasi-geodesics and have the property that they remain within an identifiable neighborhood of the optimal path, so that the search for near optimum paths can be narrowed down. More specifically, yet another property of hyperbolic graphs allows the near optimum paths to be computed as locally optimum paths, with a minimum modification of distance vectoring.

1 Introduction

One of the many security concerns in modern data networks is eavesdropping, that is, unauthorized packet interception along a link with the potential of reconstructing the full message—if all packets are sent along the same optimum path from source to destination, as TCP does under normal conditions. One of the proposed patches to such a security breach is to send packets in a randomnized fashion along different, nonoptimal routes [6]. Since the many routes have different delays, out-of-order packet arrival at the destination could create drops if the arrival sequence is altered by more than 3 slots. Unless some robustified TCP protocol is implemented, there is a need to restrict the paths to have costs bounded away from the optimum cost. On a graph, or on a surface or manifold for that matter, these near optimum paths may or may not remain within an identifiable neighborhood of the optimum path. In fact, in classical Riemannian geometry, this behavior is encapsulated in the concept of curvature: In a negative curvature space, the near optimal paths—referred to as quasi-geodesics—remain in a neighborhood of the optimal path, while in positive curvature space the near optimal paths could potentially spread across the whole manifold. For these facts to be applicable to graphs, there is a need to define a curvature concept for nondifferentiable structures, a curvature concept upon which the features of Riemannian geometry can be extended to other structures. In what has been referred to as the most significant development in geometry over the past 20 years, the concept of negative curvature has been redefined in terms of such a more primitive concept as distance and has hence become applicable to graphs.

Here we restrict ourselves to hyperbolic (negative curvature) graphs, the chief reason being that

Monte Carlo simulation has indicated that the popular "growth, preferential attachment" model of Internet build up promotes negative curvature (see Section 3.6). Furthermore, from the point of view of network architecture, it appears desirable to design it hyperbolic, for the near optimal paths do not have to be sought across the whole network, but can be narrowed down to an identifiable neighborhood of the optimal path. The problem is that, while *existence* of bounds has been proved [5], they tend to be overly conservative, so that for engineering applications, tighter bounds must be sought (Section 3.3).

Even though the quasi-optimal paths can be sought, with reasonable efficiency, within an identifiable neighborhood of the optimum path, we prefer, however, to use yet another property—mainly that the quasi-geodesics can be computed as local geodesics, that is, locally optimum paths. As such the quasi-geodesics will be computable via a slight modification of distance vectoring.

2 Coarse geometry

2.1 Geodesic spaces

Let (X, d) be a metric space, that is, a set endowed with a reflexive, symmetric, positive definite, non-degenerate form satisfying the triangle inequality. A rectifiable arc is a continuous map $a : [0, l] \to X$ such that

$$\sup_{t_i < t_{i+1}} \sum_i d(a(t_i), a(t_{i+1})) < \infty,$$

the latter defining the length $\ell(a)$ of the arc. A geodesic γ is an isometric embedding $\gamma : [0, \ell] \to X$. It is easily seen that γ is the shortest length arc joining $\gamma(0)$ and $\gamma(\ell)$. (X, d) is said to be a (totally) geodesic space if any two points A, B can be joined by a geodesic of length $d(A, B)$. The Hopf-Rinow theorem [8, Th. 1.4.8], [2, Corollary 3.20] asserts that a *complete*, connected Riemannian manifold is a geodesic space. A finite graph in which every edge is assigned a "weight," as it is practiced in communication networks, is a geodesic space; the distance between two vertices is defined as the infimum of the weights of all paths joining A and B, where the weight of a path is the sum of the weights of the constituting edges of the path. An infinite graph need not be a geodesic space; for example, the non locally finite graph on two vertices A, B joined by countably infinitely many edges with weights $1 + 1/n, \; n = 1, 2, \ldots$ is not a geodesic space, because $d(A, B) = 1$, yet there does not exist a geodesic of length 1 joining A, B.

In general, a geometry has its "isometries." The isometries of Euclidean geometry are orthogonal matrices; the isometries of symplectic geometry are symplectic matrices; etc. In the geometry we are about to define, a crucial role is played by the so-called *quasi-isometries*. Two metric spaces $(X_1, d_1), (X_2, d_2)$ are said to be (λ, ϵ) quasi-isometric if there exists a (not necessarily continuous) function $f : X_1 \to X_2$ such that

$$\frac{1}{\lambda} d(x_1, y_1) - \epsilon \leq d(f(x_1), f(y_1)) \leq \lambda d(x_1, y_1) + \epsilon$$

and there exists a C such that

$$\inf\{d_2(x_2, f(X_1)) : x_2 \in X_2\} \leq C$$

It is easily seen that the lattice $\mathbb{Z}^2$ is quasi-isometric to the plane $\mathbb{R}^2$, the intuition behind it being that if one looks at the lattice $\mathbb{Z}^2$ from far away, it appears as $\mathbb{R}^2$. In other words, $\mathbb{R}^2$ is a "blurring" of $\mathbb{Z}^2$. Conversely, $\mathbb{Z}^2$ is a "coarsening" of $\mathbb{R}^2$. Beyond this trivial example, it is amazing that some geometry can be done at all up to quasi-isometries. Probably the most spectacular example comes from group theory. Let $\langle g_1, \ldots, g_r | R_1, \ldots, R_m \rangle$ be a presentation of a group Γ by generators and relators. The Cayley graph of this presentation is the graph rooted at the identity element 1, with branches corresponding to right multiplication by g_i, g_j^{-1}. The relators R_i destroy the tree structure of the graph and create loops. The distance between two words $d(w_1, w_2)$ is the minimum number of generators g_i, g_j^{-1} needed to construct $w_1^{-1} w_2$. It is easily seen that this distance is symmetric and that the Cayley graph is a geodesic space. Let $\langle g_1', \ldots, g_r' | R_1', \ldots, R_m' \rangle$ be *another* presentation of the *same* group obtained for example via a Nielsen transformation. It turns out that the Cayley graphs of the two presentations of the same groups are quasi-isometric. Of course, a communication network graph is far from a Cayley graph, the chief difference being that communication network graphs are more heterogeneous. Nevertheless, because Cayley graphs are so well understood and also because they are hyperbolic with probability 1, they provide an ideal testbed for new theories. For example, the rate of propagation of a worm on a graph appears to be slowed down by loops; this phenomenon is most easily analyzed on Cayley graphs of "small cancellation" groups (see [10, Chap. V]), because the number of loops and their sizes are easily controlled by the relators and their word length.

Among the concepts associated with a geometry up to quasi-isometry is the concept of *quasi-geodesics*. A (λ, ϵ) quasi-geodesic is simply a (λ, ϵ) quasi-isometric embedding $[0, \ell] \to X$. In Engineering, we would certainly trade a geodesic for a quasi-geodesic. The only problem is whether a quasi-geodesic is guaranteed to be close to the geodesic. This is the case if the geodesic space is hyperbolic. Another important concept in geodesic spaces is that of a k-local geodesic, defined to be a continuous map $a : [0, l] \to X$ such that the restriction $a|_{[t_1,t_2]}$ is an isometry for $|t_2 - t_1| < k$. Is a k-local geodesic a geodesic? The answer to this question is postponed to Section 3.3.

It is sometimes useful (see, e.g., Section 4) to have a concept of angle in a geodesic space. The approach of Buseman involves the important notion of *comparison triangle*. Given a geodesic triangle $[ABC]$ in some geodesic space, the *comparison triangle* [1, p. 19] is a triangle $\bar{A}\bar{B}\bar{C}$ in Euclidean space $(\mathbb{E}^n, \bar{d})$ or in constant sectional curvature standard space $(M_\kappa, \bar{d})$ such that $\bar{d}(\bar{A}, \bar{B}) = d(A, B)$, $\bar{d}(\bar{B}, \bar{C}) = d(B, C)$, $\bar{d}(\bar{C}, \bar{A}) = d(C, A)$. The comparison triangle also allows a concept of angle to be defined *solely in terms of the distance, independently of the concept of inner product*. The *Alexandrov angle* [2, Def. 1.12] at the vertex A of a geodesic triangle $[ABC]$ is the $\overline{\lim}_{\epsilon\to 0}$ of the usual Euclidean angle at the vertex $\bar{A}$ of the comparison triangle $\bar{A}\bar{B}_\epsilon\bar{C}_\epsilon$ of $[ABC]$, where $B_\epsilon \in [AB]$, $C_\epsilon \in [AC]$, and $d(A, B_\epsilon) = \epsilon d(A, B)$, $d(A, C_\epsilon) = \epsilon d(A, C)$.

Clearly, we have found a common structure that encompasses Riemannian manifolds and graphs. We have fallen, however, a bit short of encompassing those networks where the cost of communicating from A to B is not the same as the cost of communicating from B to A, also referred to as *digraphs*. This can be accomplished using some concepts of noncommutative geometry.

2.2 Embedding graphs on surfaces

To capture some of the many concepts of coarse geometry, consider a graph G, which we write on the sphere S^2, possibly with some edges crossing. (In fact, the only situation where no edge crossings occur is when the graph is planar.) For each edge crossing, "pull a handle" and draw one of the edges on the handle rather than on the sphere. After pulling a handle for every pair of crossing edges, the graph is written—without edge crossings— on a sphere with g handles, that is, the compact surface S_g of genus g. This process is of course nonunique, but among all such processes there is one that leads to a minimum number of handles, called the *genus of the graph* [12, Def. 6-9]. From here on, we will assume that the graph G has been embedded in S_g for the smallest genus g.

If $g > 2$, then it is well-known that S_g carries a hyperbolic metric g_{ij} of constant sectional curvature. If the graph is properly filling S_g, then it will be hyperbolic in the sense of Section 3.2.

3 Hyperbolic geometry

3.1 Riemannian curvature

In classical Riemannian geometry, the curvature is defined as the operator $R(X, Y) = \nabla_Y \nabla_X - \nabla_X \nabla_Y + \nabla_{[X,Y]}$ acting on sections of the tangent bundle. The important interpretation is that $-R(X, Y)Z$ is the difference between Z and the vector obtained by parallel displacement of Z around the parallelogram constructed on X, Y. Introducing the sectional curvature $\kappa(X, Y) = \frac{<R(X,Y)X,Y>}{||X||^2||Z||^2-<X,Y>^2}$, the preceding yields, for a surface, the Gauss-Bonnet theorem $\int\int_S \kappa dS = 2\pi - \sum_i \tau_i$, where the τ_i's are the turn angles on the boundary of S. If S is a geodesic triangle with internal angles α, β, γ, it is readily seen that $2\pi - \sum_i \tau_i = \alpha+\beta+\gamma-\pi$, so that "negative curvature" means that the sum of the internal angles of a geodesic triangle is less than π.

The important properties of negative sectional curvature spaces stem form the behavior of their geodesics. The Jacobi field J [7, Section 2.1] is by definition the infinitesimal variation of a geodesic relative to a nearby geodesic. Searching a solution of the form $J = jW$, where W is orthonormal to the nominal geodesic, the Riccati equation of the Jacobi field, for a constant sectional curvature space, yields $\frac{d^2j}{ds^2} + \kappa j = 0$, which, if $\kappa < 0$, yields $j(s) = j(0)\cosh\left(\sqrt{-\kappa}s\right)$. Turning the problem the other way round, under variation of the end point, the perturbed geodesic remains close to the nominal geodesic.

In contrast, a positively curved space, e.g., a sphere, does not have well behaved geodesics; think, for example, of the geodesic joining the South pole of a sphere to a point drifting around the North pole!

3.2 Hyperbolic graphs

Inspiring ourselves form Riemannian geometry, the question arises as to whether in, say, a communication network a quasi-geodesic would be guaranteed to be close to the geodesic. One would be tempted to say that this might be the case if the graph is "hyperbolic." The problem is that in classical Riemannian geometry, sectional curvature is a differential concept and a graph is certainly not a differential object! The approach developed by Cartan, Alexandrov, Rauch, and Toponogov [9, Sec. 3.2] is to single out a property of a hyperbolic manifold that can be reformulated solely in terms of distance function and geodesics. One such property is the bounded fatness of the geodesic triangles. To be specific, let M be a Riemannian manifold of sectional curvature $\kappa < 0$. The *fatness* of a geodesic triangle $[ABC]$ is defined as

$$\delta_F([ABC]) = \inf\left\{ d(X,Y) + d(Y,Z) + d(Z,X) : \begin{array}{r} x \in [BC] \\ y \in [AC] \\ z \in [AB] \end{array} \right\} \quad (1)$$

and the bounded fatness property [11, pp. 84-85] is that

$$\delta_F = \sup\{\delta_F([ABC]) : A, B, C, \in M\} < \frac{6}{\sqrt{-\kappa}} \quad (2)$$

Here we are at the crucial point. The classical Riemannian concept of negative sectional curvature has been reformulated in terms of distance and geodesics. As such it can be extended to geodesic spaces, in particular to graphs. Therefore, we will say that a graph is δ_F-negatively curved if the above holds. We could go one step further and say that a graph has curvature bounded above by $\kappa < 0$ if $\delta_F < \frac{6}{\sqrt{-\kappa}}$.

There are many alternative characterizations of negative curvature in terms of distance and geodesics, although not all of them are uniformly equivalent. One such characterization we will make ample use is that of slimness of a geodesic triangle, which is defined as the least δ such that every edge is contained within the union of the δ-neighborhoods of the other edges; formally,

$$\delta_S([ABC]) = \inf\left\{ \delta : \begin{array}{c} [AB] \subseteq N_{[AC]}(\delta) \cup N_{[CB]}(\delta) \\ [BC] \subseteq N_{[BA]}(\delta) \cup N_{[AC]}(\delta) \\ [CA] \subseteq N_{[CB]}(\delta) \cup N_{[BA]}(\delta) \end{array} \right\}$$

The slimness of the geodesic triangles is the property that

$$\delta_S = \sup\{\delta_S([ABC]) : A, B, C \in M\} < \frac{2}{\sqrt{-\kappa}} \quad (3)$$

It can be shown that $\delta_F < \infty \Leftrightarrow \delta_S < \infty$. However, bounds are harder to come by. The reason why some bounds are needed is that most of our results pertaining to good behavior of the geodesics on communication graphs involve δ_S, while computationally we found that δ_F is easier to come by (see Section 3.5). From (2),(3), we would be tempted to assert that $\delta_F = 3\delta_S$, which appears to be confirmed by numerical exploration. However, exact bounds are harder to come by. The way to go about the problem is to use the radius r of the circle inscribed to the triangle (Gromov [5, p. 162] rather speaks of the inscribed triangle), from which it follows that $r < \delta_S < 2r$

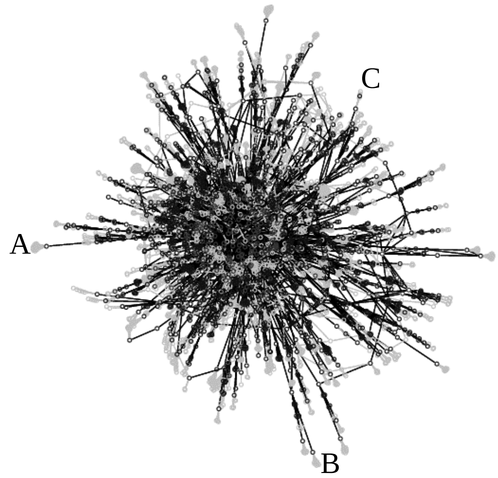

Figure 1: The ISP graph consisting of a highly connective core and long tendrils. Observe that the geodesic triangle $[ABC]$ is "slim."

and $2r < \delta_F < 6r$, and finally $1 < \delta_F \delta_S^{-1} < 6$. (See also [5, Sec. 6.5, 6.6] for similar bounds.) Yet another approach is the "thinness" δ_T of a geodesic triangle, in which a geodesic triangle is viewed as a δ_T-fattened version of a tripod (see [5, Sec. 6.3] for detail).

As an illustration, consider Figure 1, showing the ISP graph. Each node is a cluster of hosts with nearly matching IP addresses and the links are adjacent addresses as revealed by a traceroute. This figure is an illustration of the "core-concentric" property of many AS (Autonomous Systems) graphs as revealed by the CAIDA project. As with any graph consisting of a highly connective core and long tendrils, the geodesic lines joining three points A, B, C at the ends of the tendrils are forced to transit via the highly connective core, contributing to the slimness of the geodesic triangle ABC.

Besides fatness, slimness, and thinness of geodesic triangles, we mention the approach of Alexandrov [7, Sec. 2.3] and Buseman [7, Sec. 2.2].

3.3 Geodesics versus quasi-geodesics versus k-local geodesics

A crucial feature of δ negatively curved spaces, which is in a sense their motivation, is that k-local geodesics are quasi-geodesics and that quasi-geodesics are "close" to geodesics; hence k-local geodesics are "close" to geodesics. We first review the traditional results.

Specifically, a k-local, $k > 8\delta_S$, geodesic is a (λ, ϵ) quasi-geodesic for $\lambda = \frac{k+4\delta_S}{k-4\delta_S}$ and $\epsilon = 2\delta_S$ (see [2, Th. III.1.13]). Next, there exists a universal tolerance constant $r(\delta_S, \lambda, \epsilon)$ such that if γ is a geodesic and $\tilde{\gamma}$ a (λ, ϵ) quasi-geodesic with the same end points, we have $d_H(\tilde{\gamma}([0,\ell]), \gamma([0,\ell])) < r(\delta_S, \lambda, \epsilon)$ (see [2, Th. III.1.7]). A somewhat more specific bound that applies to $(\lambda, \epsilon = 0)$ quasi-geodesics is $r(\delta_S, \lambda, 0) = 100\delta_S(1 + \log_2 \lambda)$ (see [5]). Combining the above two results, it follows that, if γ, $\tilde{\gamma}_k$ are the geodesic and a k-local geodesic, respectively, we have

$$d_H(\tilde{\gamma}_k, \gamma) < r(\delta_S, \frac{k+4\delta_S}{k-4\delta_S}, 2\delta_S)$$

Here, we see the problem: even if we take $k \to \infty$, we get $d_H(\tilde{\gamma}_k, \gamma) < r(\delta_S, 1, 2\delta_S) \neq 0$, so that we cannot guarantee that the k-local geodesic has been forced to be a geodesic.

The good news is that we feel that, if we restrict ourselves to graphs, we can do better than that. Specifically, we conjecture that for a graph, taking $k = 2\delta_S$ would force the k-local geodesic to coincide with the geodesic. To get the feeling for this result, consider a graph consisting of a triangle $A'B'C'$ with extra edges $A'A, B'B, C'C$ attached to the vertices of the triangle. Clearly, $\delta_S = 1$. It is easily seen that to find the optimal path from A to B, it suffices to check optimality of the $(k = 2)$-local geodesics; this indeed rules out the path $A'C'B'$ that takes the detour along two edges of the triangle.

The formal argument–yet to be constructed–would proceed as follows: A δ_S graph is a δ_S fattened version of a tree; more precisely, the δ_S graph is $(\lambda = 1, \delta_S)$-quasi-isometric to a tree (see [5, Sec. 6.1] for similar concepts). The geodesic $[AB]$ is then quasi-isometric to any path joining A and B in the "fat tree." A nonoptimal path would go from one side to the other of a branch of the fat tree and this would be detected by a $2\delta_S$ local test of optimality.

To assess how conservative the traditional bounds are, we consider the Riemannian geometry case. Let

$\gamma, \tilde{\gamma} : [0, \ell] \to M$ be geodesic, λ quasi-geodesic, respectively, with the same end points, viz., $\gamma(0) = \tilde{\gamma}(0)$, $\gamma(\ell) = \tilde{\gamma}(\ell)$. The geodesic is parameterized by arc length s. The quasi-geodesic is parameterized as follows: Let $\tilde{X}_i \in \tilde{\gamma}$. Let $[\tilde{X}_i X_i]$ be the geodesic segment orthogonal to γ at the point $X_i \in \gamma$. Let $\gamma(s_1) = X_1$. Then $\tilde{\gamma}(s_1) = \tilde{X}_i$. This provides a unique parameterization of $\tilde{\gamma}$ provided we restrict ourselves to a neighborhood without focal points. Let $D(s) := d(\tilde{X}(s), \gamma)$. Let $X_1 = \gamma(s_1)$, $X_2 = \gamma(s_2)$ be a pair of arbitrarily close points on the geodesic and let $\tilde{X}_1 = \tilde{\gamma}(s_1)$, $\tilde{X}_2 = \tilde{\gamma}(s_2)$ be the corresponding points on the quasi-geodesic. Using the Jacobi field applied to the geodesics $[X_1\tilde{X}_1]$, $[X_2\tilde{X}_1]$ and the λ bound on the quasi-geodesic, it follows that

$$\begin{aligned} d^2(\tilde{X}_1, \tilde{X}_2) &= ds^2 \left(\cosh^2(\sqrt{-\kappa} D(s)) + (D'(s))^2\right) \\ &\leq \lambda^2 ds^2 \end{aligned}$$

Therefore, the λ quasi-geodesic that departs with a maximum speed from the geodesic is given by the differential equation

$$\cosh^2\left(\sqrt{-\kappa} D(s)\right) + (D'(s))^2 = \lambda^2, \qquad (4)$$

subject to the mixed boundary condition

$$D(0) = D(\ell) = 0$$

and

$$\cosh^2\left(\sqrt{-\kappa} D(s)\right) - \lambda^2 \geq 0$$

Therefore, from the latter, the bound is derived as

$$D_{max} = \frac{1}{\sqrt{-\kappa}} \cosh^{-1} \lambda \approx \frac{1}{\sqrt{-\kappa}} \log_e 2\lambda$$

Figures 2,3 show the results of two simulation runs of Equation (4). In the first case, the quasi-geodesic reaches its bound and remains at that constant distance away from the geodesic, while in the second case, the quasi-geodesic does not reach its bound.

The traditional bounds are derived as follows: Let γ be a geodesic and let $\tilde{\gamma}$ be the λ quasi-geodesic fastest departing from the geodesic with the same end points. Let D be the distance, which by the above argument is reached at the mid point $X_0 \in \gamma$ and let $\tilde{X}_0$ be the corresponding point on $\tilde{\gamma}$, that

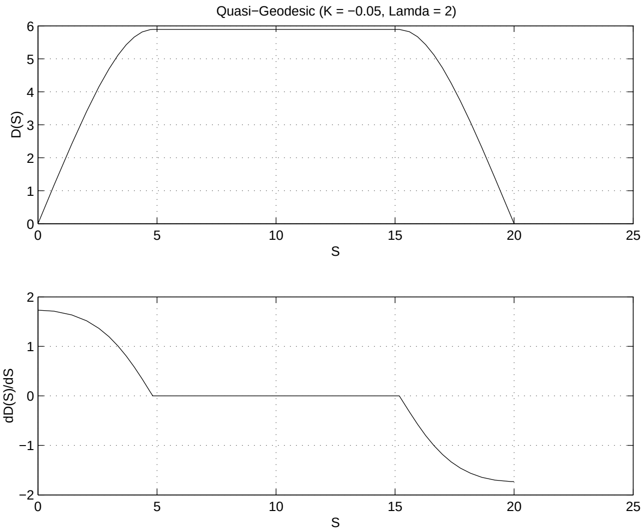

Figure 2: Envelope of all distance plots between λ quasi-geodesics and a nominal geodesic of length 20. The curvature is adjusted so that the solution reaches the bound, resulting in a continuously differentiable curve.

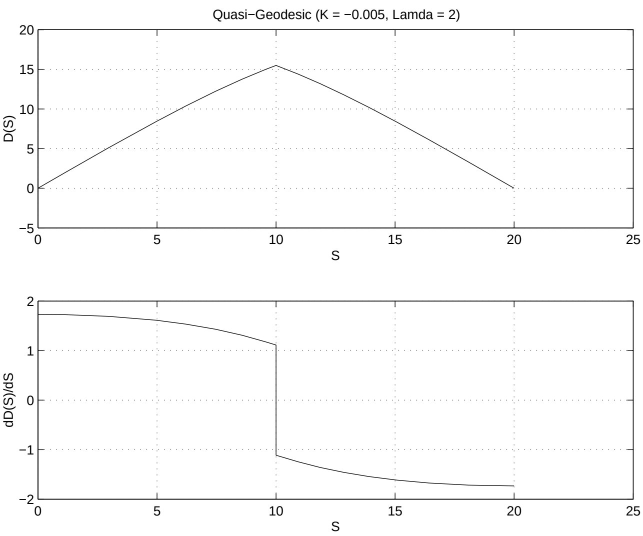

Figure 3: Envelope of all distance plots between λ quasi-geodesics and a nominal geodesic of length 20. The curvature is adjusted so that the solution does not reach the bound, resulting in a nondifferentiable curve.

is, $d(X_0, \tilde{X}_0) = D$. Choose points $X_1, X_2 \in \gamma$ along with their corresponding points $\tilde{X}_1, \tilde{X}_2 \in \tilde{\gamma}$ such that $d(X_0, X_1) = d(X_0, X_2) = d(X_1, \tilde{X}_1) = d(X_2, \tilde{X}_2) =: D_1$. Using the λ quasi-geodesic property, we get

$$\ell(\tilde{X}_1 \tilde{X}_2) \leq \lambda d(X_1, X_2) = 2\lambda D_1$$

On the other hand, using the Jacobi field, we get

$$2D_1 \cosh\left(\sqrt{-\kappa} D_1\right) \leq \ell(\tilde{X}_1 \tilde{X}_2)$$

Combining the above two yields

$$D_1 \leq \frac{1}{\sqrt{-\kappa}} \cosh^{-1} \lambda$$

Finally, using $D \leq \frac{1}{2}\ell(\tilde{X}_1 \tilde{X}_2) + D_1$, we get

$$D \leq \frac{\lambda + 1}{\sqrt{-\kappa}} \cosh^{-1} \lambda \approx \frac{\lambda + 1}{\sqrt{-\kappa}} \log_e 2\lambda$$

We now look at tight bounds in arbitrary geodesic spaces. Using the material of [2], it can be shown that a bound is given by

$$D_{max} = D_0(\lambda^2 + 1) + \frac{\lambda}{2}(2\lambda^2 + 3)$$

where D_0 is the maximum solution to

$$D_0 \leq \delta_S \log_2 \left(\frac{1}{\delta_S} \left(D_0 \left(6\lambda^2 + 2 \right) + \lambda \left(2\lambda^2 + 3 \right) \right) \right) + \frac{\delta_S}{2}$$

Using some equation solver, this bound is easily computed and compared with the traditonal Gromov bound in Figure 4, showing a substantial improvement by more careful analysis of the bounds.

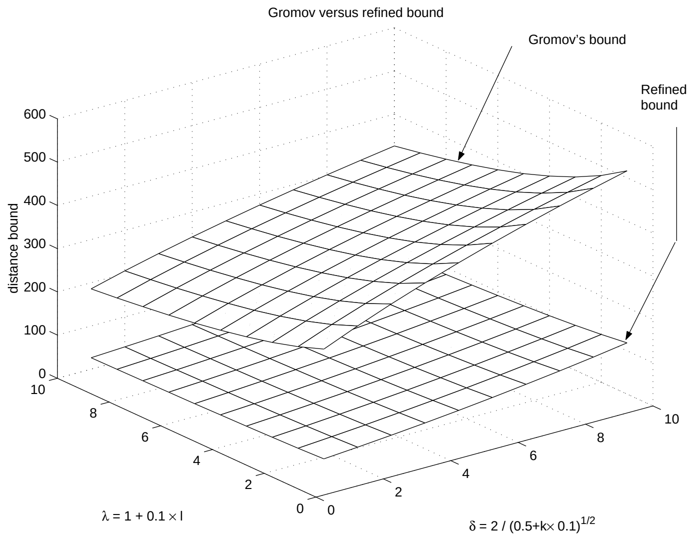

Figure 4: Gromov's bound versus refined bound.

3.4 Billiard dynamics curvature computation

The following provides an optics interpretation of the "fatness," where the angles are defined from the Riemannian inner product.

Theorem 1 *In a complete Riemannian manifold of nonpositive curvature and for a geodesic triangle $[ABC]$ without obtuse angles, the solution triangle XYZ of (1) is such that the incidence angles at X, Y, Z equal the corresponding reflection angles.*

Proof. We first check the first order variation. It is easily seen that, in general Riemannian geometry,

$$\begin{aligned} &\frac{\partial}{\partial d(A, Z)}(d(X, Y) + d(Y, Z) + d(Z, X)) \\ &= \cos\angle AZY - \cos\angle BZX \end{aligned}$$

Cancelling the first order variation yields

$$\begin{aligned} \text{incidence angle} &:= \frac{\pi}{2} - \angle AZY \\ &= \frac{\pi}{2} - \angle BZX \\ &=: \text{reflection angle} \end{aligned}$$

and the same holds at X, Y. Regarding the second order variation, we have, for a complete Riemannian manifold of nonpositive sectional ccurvature [8, Lemma 2.1.5],

$$\frac{\partial}{\partial d(A, Z)}(d^2(Y, Z) + d^2(Z, X)) \geq 4$$

which takes care of the second variation. ∎

We leave it to the reader to verify that, in case the triangle has an obtuse angle, say α, the solution to (1) is $X_\perp, Y = A, Z = A$ and the optimization problem (1) is no longer a differentiable one.

Theorem 1 is the well-known *Fermat Principle* of optics, saying that a light ray from X to Z reflecting at $Y \in [BC]$ minimizes $\int_X^Z n ds$ where n is the refraction index. Another interpretation is that the XYZ triangle is a *periodic orbit* of the billiard dynamics on a geodesic triangular table.

The problem is how to effectively construct the trajectory that reflects itself with an angle equal to the angle of incidence on every edge of the triangle.

We first consider the case of an ordinary Euclidean triangle $[ABC]$ without obtuse angle, that is, $\alpha, \beta, \gamma < \frac{\pi}{2}$. From A, draw the altitude $[AX_\perp]$, that is, the line segment such that $[AX_\perp] \perp [BC]$ and $X_\perp \in (BC)$. The altitudes $[BY_\perp]$ and $[CZ_\perp]$ are defined similarly. All three altitudes $[AX_\perp]$, $[BY_\perp]$ and $[CZ_\perp]$ are well known to intersect at a single point, say H, called orthocenter [3, Sec. 1.6, p. 17]. The following is easily proved:

Theorem 2 $\angle Z_\perp X_\perp H = \angle H X_\perp Y_\perp$, *with a similar result at* $Y_\perp, Z_\perp$, *so that* $X_\perp, Y_\perp, Z_\perp$ *is the solution to (1).*

The question arises as to whether the above can be extended to negative curvature geometries. It is not hard to show that the above remains true on a constant negative sectional curvature surface and for triangles with two vanishes angles and the remaining one acute; for example, take the Poncaré upper half plane model and consider a triangle with A on the real axis, B anywhere in the upper half plane and $C = \infty$. The edge $[AB]$ is an arc of the circle passing through A, B and orthogonal to the real axis; the edges $[CB]$ and $[AC]$ are straight lines perpendicular to the real axis; clearly, $\alpha = \gamma = 0$ and take B such that $\beta \leq \frac{\pi}{2}$. For such a triangle, the orthocenter exists and the above holds true. We do not know whether this extends to arbitrary triangles.

3.5 Linear programming curvature computation

The distance on a graph is initially defined over the vertex set, but it can easily be affinely extended to the edges. Then we have the following:

Theorem 3 *For a graph and with the distance* $d(\cdot, \cdot)$ *affinely extended to the edges, the solution* X, Y, Z *to (1) is reached at vertices of* $[AB], [BC], [CA]$*, respectively.*

Proof. Clearly, the domain in which (X, Y, Z) runs is $[AB] \times [BC] \times [CA]$ and the simplicial decomposition of the factors endows the product with the structure of a polyhedron. Next, because distances are measured along paths of edges on the graph, $d(X, Y) + d(Y, Z) + d(Z, A)$ is easily seen to be affine in X, Y, Z. Therefore, (1) is a linear programming problem, which has its solution on a vertex of the polyhedron $[AB] \times [BC] \times [CA]$, that is, X is on a vertex of $[BC]$, Y on a vertex of $[AC]$, and Z on a vertex of $[AB]$. ■

3.6 Negative curvature versus heavy tail

Besides the obvious core-concentricity, a more subtle process through which a graph could become hyperbolic is when its degree has a heavy tail distribution. To show that a heavy tail graph is more hyperbolic than a uniform graph, we consider the growth model of the Internet. We start from a core network G^{n_0} consisting of n_0 vertices. This start-up graph could have line topology, bidirectional ring topology, unidirectional ring topology, or star topology. We recursively introduce a new vertex v_*^n and successively connect this new vertex to m vertices, $v_1, ..., v_m$, of the already existing graph G^{n-1}, resulting in the graph G^n with n vertices and with adjacency matrix A^n. For the kth attachment of the new vertex v_*^n to an old vertex not previously attached to v_*^n, $v_i \in G^{n-1} \setminus \{v_1, ..., v_{k-1}\}$, we use either of the following rules:

1. the *uniform* attachment rule, that is,
$$p\left(a^n(v_i, v_*^n) = 1\right) = \frac{1}{n-k}$$
2. the *preferential* attachment rule, that is,
$$p\left(a^n(v_i, v_*^n) = 1\right) = \frac{d(v_i)}{\sum_{v_i} d(v_i)}$$

The preferential attachment rule is known to yield, as $n \to \infty$, a heavy tail graph. However, since here we generate finite graphs, the finiteness of δ_F is irrelevant and the real issue is whether δ_F is small relative to the diameter. We therefore plot $E(\delta_F/\text{diam})$, the “normalized” expectation of δ_F. Observe that for a

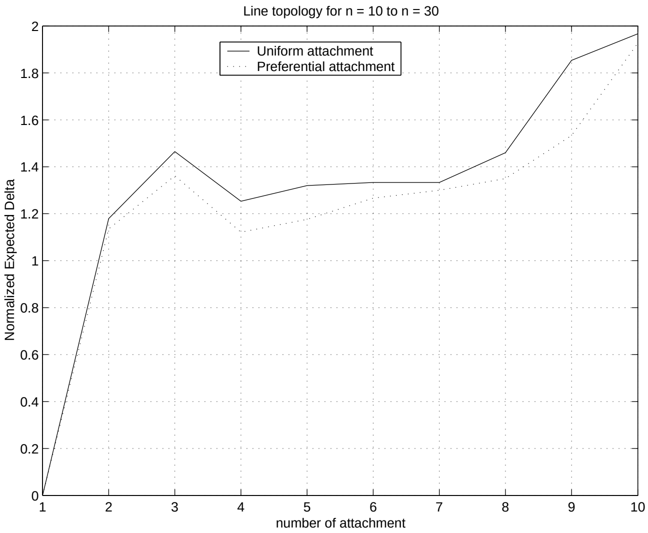

Figure 5: The normalized mathematical expectation of δ_F versus the number m of attachments at every step for the line start-up topology.

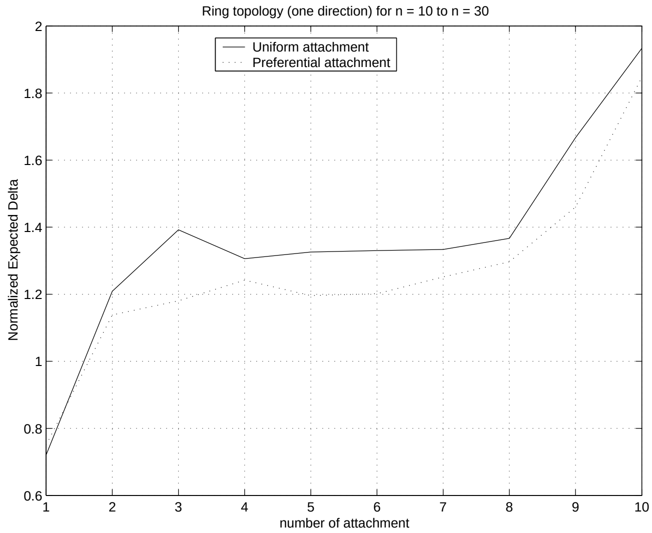

Figure 6: The normalized mathematical expectation of δ_F versus the number m of attachments at every step for the unidirectional ring start-up topology.

Euclidean equilateral triangle, the fattest possible triangle in nonpositive curvature, we have $\delta_F/\text{diam} = 3/2$ so that the issue is whether $E(\delta_F/\text{diam}) < 3/2$. In the simulation, we start with $n_0 = 10$ and run to $n = 30$. The results, for the line and unidirectional ring topologies, are shown in Fig. 5, 6, respectively, and reveal that the fatness is smaller for the heavy tail graph.

4 Negative versus positive curvature

For reliability purposes, it would be interesting to have a great many quasi-geodesics with costs within a small bound away from the geodesic cost. The number of quasi-geodesics is quantified by their distances to the geodesic, arguing that if they go far away there is plenty of space to pack them. This distance is normalized by the length of the λ quasi-geodesic, bounded as $\lambda\ell(\gamma)$ (see [5, Sec. 7.2, Prop 7.2.A]). For a differentiable hyperbolic structure, we could define the *"differential density of quasi-geodesics per unit geodesic length"* as

$$\frac{\partial d_H(\gamma, \tilde{\gamma})}{\partial \lambda} = \frac{100\delta_S \log_2 e}{\lambda}$$

while, for discrete hyperbolic structures, the *"average density of quasi-geodesics per unit geodesic length"* is defined as

$$\lim_{\lambda \to \infty} \frac{100\delta_S(1 + \log_2 \lambda)}{\lambda} = 0$$

(Another approach would be based on the isoperimetric inequality [2, III.2].) The latter observation points to the fact that the drawback of the good behavior of the geodesics in a hyperbolic space is that the density of quasi-geodesics is small. To have a good density of quasi-geodesics, we would have to go to the opposite spaces of positive curvature. Besides, positive curvature is more appropriate for small diameter graphs.

A positive curvature graph can be defined via the angles. Consider a situation where a vertex of the graph A is connected to vertices B_i, $i = 1, ..., n$, B_i is connected to B_{i+1}, and B_n is connected to B_1, with weights $d(A, B_i) = 1$, $d(B_i, B_{i+1}) = 1$, $d(B_n, B_1) = 1$. Let α_i be the angle at the vertex A of the (geodesic) triangle AB_iB_{i+1} and let α_n be the angle at the vertex A of the triangle AB_nB_1. The angles at the vertex A of the *graph* are most naturally defined following the procedure of Alexandrov in a comparison triangle in the model space of constant sectional curvature κ, M_κ [7], [2, Definition 2.15]. (In fact, a deeper result [2, Prop. 2.9] shows that, no matter what comparison space M_κ we choose, the angle is the same and hence equal to 60 degrees.)

Then the graph is locally positively curved at A if $\sum_i \alpha_i < 2\pi$. The graph is positively curved if it is locally positively curved at every vertex. Such graphs enjoy some of the properties of manifolds of positive sectional curvature.

Unfortunately, positive curvature is topologically much more constraining than negative curvature (because, for example, the Lichnerowicz-Weitzenböck formula [4, Th. 9.16], [11, p. 6] implies that the Dirac operator has vanishing index) and except for some scarce discrete geometry results (e.g., [5, Prop. 7.2.E]) and of course the Gromov-Lawson theory of manifolds with positive *scalar* curvature, the theory is on shaky mathematical ground.

5 Routing

The modification of the distance vector routing necessary for it to generate k-local geodesics (and hence quasi-geodesics) rather than geodesics is nearly trivial. Each router X, instead of keeping a vector L_X of distances from X to every other node, is only required to keep a vector of distances from X to all nodes within k hops. The Bellman-Ford algorithm is particularly well adpated to computing this reduced distance vector, since by definition it computes h hops optimal paths for arbitrary h. The vector W_X of weights of all links attached to X is unaffected. The "next hop" data is more involved, in the sense that for every node $Y \neq X$, there are several possible "next hops."

Acknowledgements

This work was supported by the Defense Advanced Research Project Agency (DARPA). The views presented here are the authors' and not those of DARPA.

References

[1] Marcel Berger. *Riemannian Geometry During the Second Half of the Twentieth Century*, volume 17 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2000.

[2] Martin R. Bridson and André Haefliger. *Metric Spaces of Non-Positive Curvature*, volume 319 of *A Series of Comprehensive Surveys in Mathematics*. Springer, New York, NY, 1999.

[3] H. S. M. Coxeter. *Introduction to Geometry*. Wiley Classics Library. John Wiley & Sons, New York, 1989.

[4] J. M. Gracia-Bondia, J. C. Varilly, and H. Figueroa. *Elements of Noncommutative Geometry*. Birkhauser, Boston, 2001.

[5] M. Gromov. *Essays in Group Theory: Hyperbolic Groups*, pages 75–263. Number 8 in Math. Sci. Res. Inst. Publ. Springer, New, York-Berlin, 1987.

[6] Joo Pedro Hespanha and Stephan Bohacek. Preliminary results in routing games. In *Proc. of the 2001 Amer. Contr. Conf.*, June 2001.

[7] J. Jost. *Nonpositive Curvature: Geometric and Analytic Aspects*. Lectures in Mathematics. Birkhauser, Basel-Boston-Berlin, 1997.

[8] J. Jost. *Riemannian Geometry and Geometric Analysis*. Universitext. Springer, Berlin, Heidelberg, New York, 1998. Second Edition.

[9] Michael Kapovich. *Hyperbolic Manifolds and Discrete groups*, volume 183 of *Progress in Mathematics*. Birkhauser, Boston, MA, 2001.

[10] R. C. Lyndon and P. E. Schupp. *Combinatorial Group Theory*. Classics in Mathematics. Springer, Berlin, Heildelberg, New York, 2001. Reprint of the 1977 edition.

[11] John Roe. *Index Theory, Coarse Geometry, and Topology of Manifolds*. Number 90 in Conference Board of the Mathematical Sciences (CBMS); Regional Conference Series in Mathematics. American Mathematical Society, Providence, RI, 1996.

[12] Arthur T. White. *Graphs, Groups, and Surfaces*, volume 8 of *Mathematical Studies*. North-Holland, Amsterdam, New York, Oxford, 1984.