Laplacian Eigenfunctions for Spatial Network Traffic Analysis

Raúl Leonardo Landa Gamiochipi, Lionel Sacks

University College London

Abstract: This paper addresses the general theory behind the use of the Laplacian eigenfunctions as a means for spatial network traffic analysis. The basics of the theory for node field analysis is then applied to measurements from the Géant network.

1 Introduction

Most traffic analysis techniques are fundamentally temporal, placing great emphasis in the behavior of traffic over single links and its change in time. However, the analysis of network traffic dynamics implies the correlation of **temporal** and **spatial** patterns across the network, and this calls for algorithms that are sensitive to its topology and geometry. Recent work in shape matching [8] and geometry processing [7] have underscored the role of the Laplace operator in exploring the structure of topological and metric spaces. Furthermore, in the context of spectral graph theory [2], the spectrum of the Laplacian has proven extremely valuable in the calculation of graph diameters, cuts and colorings. The extension of the methods of harmonic anlaysis to discrete spaces with arbitrary topologies and its use in network measurement are promising directions for further research, and this paper addresses its basic theory.

2 Discrete Differential Operators

The usual operators of vector calculus can be adapted to be used in the context of arbitrary network topologies [12][9]. In essence, the network can be thought of as a discretised manifold, and thus imbued with its own calculus. In particular, the network analogue of a vector field over $\Re^N$ is a function that assigns each link a real value (a **Link Field**), while the analogue of an scalar field is a function that assigns each node a real value (a **Node Field**). Thus, nodes take the role of points in a manifold, and links define the topological structure of this manifold, including its genus and orientability.

In practice, this approach allows us to define matrices that can be used to calculate the gradient, divergence, curl and both the scalar and vector Laplacians of fields defined on networks. In this paper we address only node fields over unweighted, symmetric directed graphs, and their gradient, divergence and Laplacian.

3 $\nabla \cdot \vec{F}$ and ∇f

The divergence is a linear operator (in our case, simply a matrix) that operates on link fields and produces a node field measuring the "traffic density" exiting the network at each node. The gradient, being the adjoint operator of the divergence, is in this case simply its matrix transpose, and maps node fields to a link field measuring the rate of change of the scalar field over each of the links.

In its most elementary form, the divergence D is directly related to the **Incidence Matrix** of the network [12]: It is a matrix that is indexed over the links in its columns, over the nodes in its rows and that assigns the value $+1$ for a node-link pair if the link is incoming on the node, -1 if it is outgoing and 0 if there is no incidence. This formulation can be extended to incorporate weights over the links or the nodes that take the role of the metric tensor over a manifold. When the gradient and the divergence are modified in this way, they can be use to define a Laplacian compatible with the ideas of spectral and differential geometry and spectral graph theory [2]. By defining W_L, the link weight matrix, as

$$W_L(l_i, l_j) = \begin{cases} w_i & \text{if } l_i = l_j \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where w_i is the weight assigned to link i, we can define the weighted divergence, D_W as

$$D_W = D\, W_L^{\frac{1}{2}} \tag{2}$$

Both the weighted and the unweighted divergences can be normalised by treating each row as a vector and making it of unitary norm.

For the gradient as G, we have that

$$G_W = D_W^T \tag{3}$$

where T denotes matrix transposition.

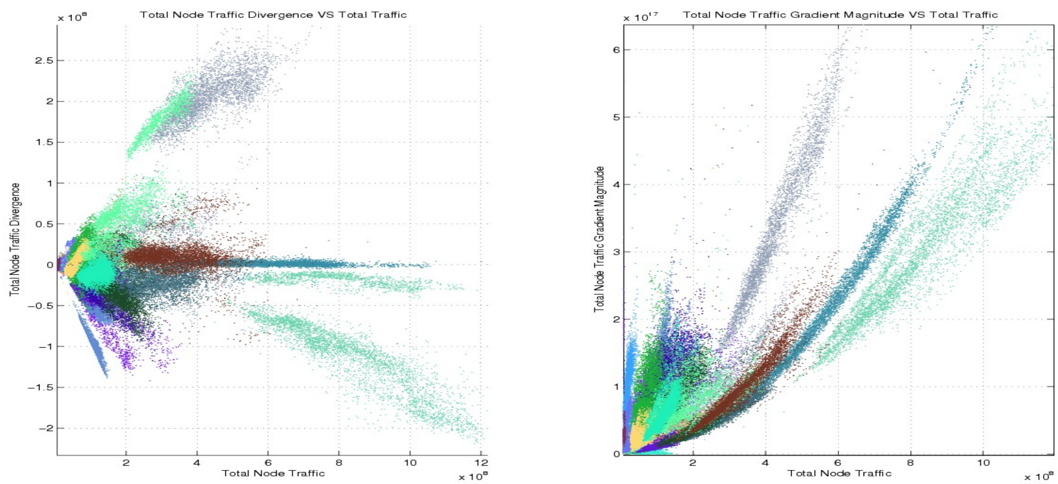

Figure 1: Divergence and Gradient Squared Magnitude as a magnitude of Total Traffic on the Géant Network. Each point corresponds to a 15 minute measurement interval on a given node over a time interval between May 4, 2005 and May 18, 2005.

In Figure 1 it can be readily seen that even though load levels vary over wide ranges even during normal network operation, the relationship between load and the gradient magnitude or divergence tends to remain stable. The source of this relationship needs to be further studied, as it suggests that simple, load dependent models can yield insight on the spatial traffic dynamics of networks during normal operation.

4 The Scalar Laplacian ($\nabla^2 f$)

As usual, the divergence of the gradient ($\nabla \cdot \nabla f$) is the **Laplacian** for scalar fields, a self adjoint operator that measures the difference of the value of a function in a point (in a node) with the average value of the function in an sphere in $\Re^N$ around it (the neighbourhood of the node). In terms of our previously defined matrices:

$$L = G_W D_W = W_L^{\frac{1}{2}} D^T D W_L^{\frac{1}{2}} \tag{4}$$

The operator that generalizes the Laplacian to general manifolds is the Laplace-Beltrami [11] operator, and it has a direct connection with the fundamental modes of vibration of a hypotetical drum in the shape of the manifold: its eigenvalues take the role of resonant frequencies, while its eigenfunctions take the role of fundamental modes of vibration that span its natural space and can be used to decompose functions on it.

The complex eigenfunctions of the Laplace operator are the main tools of **Fourier Analysis**. In particular, Fourier series are defined as a set of projections of a given function in the time/space domain into frequency domain complex eigenfunctions. Usually, when projection onto real eigenfunctions is needed, complementary functions at a same frequency are selected by the boundary conditions imposed on the Laplacian - thus arise the sine and cosine Fourier series. This same analysis is possible in the context of arbitrary network topologies by using the eigenvectors of the Laplace operator, that are always real (the Laplacian is a symmetric, positive semidefinite operator).

In the case of a network, the nullspace of the Laplacian is of dimension 1 [5], and it represents the network equivalent of a constant function (the typical 'Direct Current' component in circuit analysis). The rest of the node vector space dimensions ($|N| - 1$) is spanned by the eigenvectors corresponding to nonzero eigenvalues (thus, the node Laplacian has rank $|N| - 1$), and these are equivalent to the sine and cosine functions that span the real line.

Moreover, it has been shown that it is possible to use the Laplacian eigenfunctions as a basis for graph embedding on a manifold [1], and this can be exploited to display the network in a way that underscores the topological and geometric properties of the Laplacian eigenfunctions.

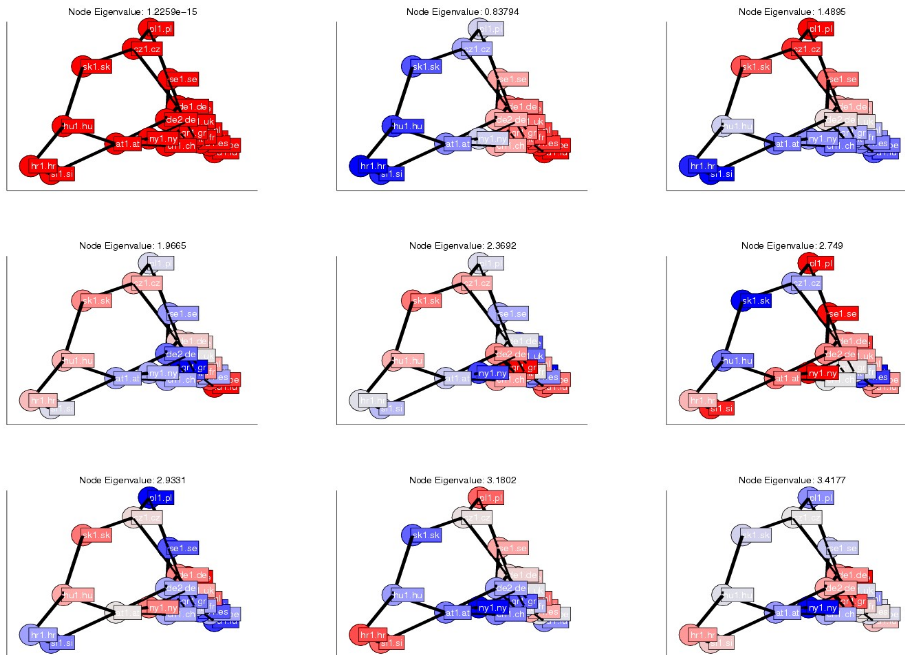

Figure 2: Node Laplacian Eigenfunctions for the first 9 eigenvalues of the Géant Network. It is clear that the Node Laplacian eigenvectors define natural modes of vibration of the underlying space, and that as the eigenvalue increases, the spatial variability of the eigenfunctions increases as well, supporting the interpretation of the eigenvalues as natural frequencies of vibration. Shades of red imply positive values, shades of blue negative values. The spatial layout of the network is defined by projecting each of the Laplacian eigenvectors onto $\Re^3$ and using the projections as coordinates.

5 Harmonic Decomposition of Network Traffic

By analysing the total traffic carried by each of the routers of the Géant network in terms of its projection over the Laplacian eigenfunction basis it is possible to measure the spatial variablity of total node traffic from a geometric standpoint. The data set we use is based on traffic matrices provided by the TOTEM project [10]. After shortrest path routing and eigenfunction projection, we analyse the resultant time series by using the Fast Fourier Transform [3].

In Figure 3, the distinct line around .042 cycles per hour corresponds to diurnal network activity periods of 24 hours. It is interesting to note that this temporal periodicity is not present in all spatial normal modes. Additionally, it is evident that some spatial modes (modes 1, 11 and 19) tend to exhibit higher temporal frequency components, or at least slower temporal frequency decay. This correlation between temporal and spatial dynamics is particulary important for network behaviour baselining, and will be the subject of further study.

6 Conclusions

The spatial analysis of network traffic can yield important insights regarding the normal and anomalous operation regimes of real networks. Thus, the development of techniques to analyse the topological and geometrical aspects of networks and the ways that they influence the structure of traffic flows within them is of great importance. The link Laplacian and its eigenfunctions, the adaptation of the Convolution Theorem to general network link and node topologies, and the study of eigenspace spatial filtering, geodesics, and the geometry of nonhomogeneous spaces (weighted graphs) remain topics for further work.

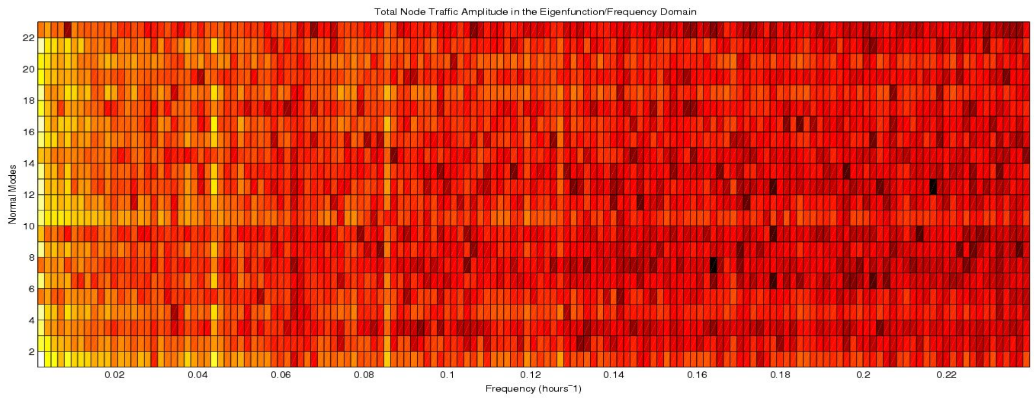

Figure 3: Partial Eigenmode-Frequency representation for link level traffic an the Géant Network. The logarithm of the temporal PSD (Power Spectral Density) of 23 fundamental spatial modes of the Laplacian basis and the first 150 discrete frequency samples of the total traffic per node on the Géant network over the 14 day period from May 4, 2005 to May 18, 2005.

Acknowledgments

Part of this work has been possible through support of the Foreign Commonwealth Office through the Chevening programme and the AlBan programme by the European Commission.

References

[1] M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering, 2002.

[2] F. R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, 1994.

[3] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19:297–301, 1965.

[4] C. Gordon, D. L. Webb, and S. Wolpert. One cannot hear the shape of a drum. *Bulletin of the American Mathematical Society*, 27:134, 1992.

[5] S. Guattery and G. L. Miller. Graph embeddings and Laplacian eigenvalues. *SIAM Journal on Matrix Analysis and Applications*, 21(3):703–723, 2000.

[6] M. Kac. Can one hear the shape of a drum? *Amer. Math. Monthly*, 73(4, part II):1–23, 1966.

[7] B. Levy. Laplace-beltrami eigenfunctions: Towards an algorithm that "understands" geometry. *IEEE International Conference on Shape Modeling and Applications 2006 (SMI'06)*, 0:13, 2006.

[8] M. Reuter, F.-E. Wolter, and N. Peinecke. Laplace-spectra as fingerprints for shape matching. In *SPM '05: Proceedings of the 2005 ACM symposium on Solid and physical modeling*, pages 101–106, New York, NY, USA, 2005. ACM Press.

[9] A. Smola and R. Kondor. Kernels and regularization on graphs. In B. Schoelkopf, editor, *Learning Theory and Kernel Machines*, pages 243–257, Berlin - Heidelberg, Germany, 2003. Springer Verlag.

[10] S. Uhlig, B. Quoitin, J. Lepropre, and S. Balon. Providing public intradomain traffic matrices to the research community. *SIGCOMM Comput. Commun. Rev.*, 36(1):83–86, 2006.

[11] R. H. wasserman. *Tensors and Manifolds with Applications to Physics*. Oxford University Press, 2nd edition, 2004.

[12] D. Zhou and B. Schölkopf. *Discrete Regularization*. MIT Press, Cambridge, MA, 2006.