Efficient routing with partial information

Xiaoran Yan
Network Science Institute (IUNI)
Indiana University
Bloomington, Indiana
xiaoran.a.yan@gmail.com

Andrea Avena-Koenigsberger
Department of Psychological and Brain Sciences
Indiana University
Bloomington, Indiana

Olaf Sporns
Department of Psychological and Brain Sciences
Network Science Institute (IUNI)
Indiana University
Bloomington, Indiana

ABSTRACT

We propose a series of mathematical frameworks for modeling routing and communication patterns on large weighted networks. By controlling the information available to each vertex, these models produce a continuous spectrum of dynamics that converge onto shortest-path and random diffusion processes at the extremes. We demonstrate that efficient global routing is possible with limited local or partial information. In the future, we plan to use Markov decision process optimization to provide efficient solutions to find such routing strategies.

KEYWORDS

Network routing, Information diffusion, shortest path, Markov decision process

ACM Reference Format:
Xiaoran Yan, Andrea Avena-Koenigsberger, and Olaf Sporns. 2018. Efficient routing with partial information. In *Proceedings of Workshop on Graph Techniques for Adversarial Activity Analytics (GTA3 2018).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The function of many real world complex networks relies on information exchange between their constituent elements. Efficient routing of information across large distances is essential to technological, social and biological systems. Previous work on optimal routing in networks highlighted the importance of small-world topologies for promoting short communication pathways at low wiring cost [6]. Yet, such a communication model assumes that system's elements having information about the global topology of the network [5]. Therefore, an appropriate analysis of routing strategies should not only concern about communication efficiency, but also takes into account the cost of storing and using topological information. In this paper, we call the cost of storing and using topological information as the informational cost.

A drastically different picture emerges if we consider communication models with minimal local topological information, i.e. the simple random walk or diffusion process [3, 7]. While the diffusion has no associated cost of storing global topological information, communication is inefficient if measured in terms of the time needed for a signal to arrive at a specific destination. We call the time cost of information routing as the transmission cost.

Between the two extremes of diffusion and shortest path routing, a spectrum of communication processes deserve greater attention. For example, a preferential choice policy where nodes preferentially route messages towards high degree nodes can decrease search times significantly [1], yet the informational cost is small since nodes only need to know the degree of their neighbors. In general, biased random walks can generate relatively efficient communication processes and are able to account for navigation rules that are observed in real world systems [4, 7].

Here, we propose a series of mathematical frameworks based on random walks. These frameworks allow us to explore a continuous spectrum of dynamics that converge onto shortest-path routing at one extreme, and diffusion processes at the other. We will empirically investigate the trade-off between informational and transmission costs, and demonstrate that efficient global routing is possible with limited local information. In the future, we plan to optimize the routing strategies under these models using a Markov decision process with finite horizons, which would also allow us to scale the experiments to larger networks.

2 A COMMUNICATION MODEL WITH LOCAL INFORMATION

We consider a weighted and undirected graph G composed of N vertices and M edges, with positive weights. The graph topology is described by the symmetric weight matrix A, where the generic element $A_{uv} = A_{vu} > 0$ if there is a weighted edge between nodes u and v, while $A_{uv} = A_{vu} = 0$, otherwise. The simple discrete time *unbiased random walk* (URW) on graph $G = (V, E, A)$, can be represented using a $N \times N$ stochastic matrix $P = AD^{-1}$, where D is the diagonal out-degree matrix, with entries $d_u = \sum_v A_{uv}$. We define a distance matrix W by inverting the non-zero entries of A, $W_{uv} = 1/A_{uv}$. Under the shortest path routing, we can also capture the pair-wise distances between vertices using a matrix S.

To capture the informational cost, we adopt a simplified explicit routing model from the computer networking literature [9]. Each vertex u has a routing table of size $O(|\delta(u)| * |\delta^k(u)|)$, where $\delta^k(u)$ with $k = [1, 2, \ldots]$ being the $k - hop$ neighborhood of vertex u. When $k = 1$, each vertex only has routing information about their direct neighbors, leading to the null model of a simple URW. For a unweighted graph, when $k = \max_{uv}[S_{uv}]$, each vertex has a complete routing table containing information of the shortest-paths towards any target in the graph. For a given k, we assume that the

GTA3 2018, Feb 2018, Los Angeles, California

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

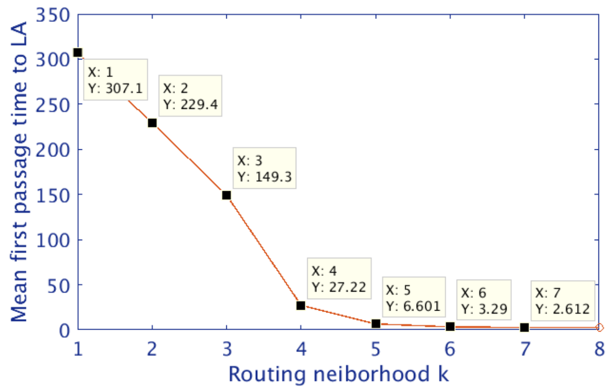

Figure 1: The average MFPT to the Los Angeles International Airport with different local routing neighborhood size

routing strategies follows a hybrid rule, i.e. conducting simple URW until the target is reachable within the $k - hop$ neighborhood of an inter-mediate vertex, then switch to the shortest-path using the routing tables.

To quantify the transmission cost, we calculate the mean first passage time (MFPT) of the random walk following the aforementioned hybrid strategy. To illustrate the effect of an increasing k, we try it on a global airline network. The airline network is constructed from public data on flights between major airports around the world, with the edge weight representing the number of flights as well as the capacity of the plane operating each flight [8].

As figure 1 demonstrates, when averaged over all possible source airport, the MFPT of the routing strategy quickly drops as k increase. The furthest airport can be reached from LA in 8 hops, leading to a averaged MFPT=2.527 at $k = 8$.

3 A CONTINUOUS MODEL WITH PARTIAL INFORMATION

While the simple hybrid routing strategy can model communication models with local information, a continuous interpolation between URW and shortest-path routing strategies can be achieved using biased random walks (BRW). Consider the random walk,

$$P_\lambda(u, v, t) = \frac{1}{Z_i^t} \exp[-(\lambda(W_{uv} + S_{vt}) + W_{uv})] , \qquad (1)$$

where λ is the interpolation parameter. As $\lambda \to \infty$, it approaches the shortest-paths. We recover the URW with $\lambda = 0$.

To capture the informational cost, we calculate the Kullback-Leibler divergence between the routing strategy $P_\lambda(u, v, t)$ and the reference P of an URW. The KL-divergence intuitively measures the additional bits of information required to manipulate the outgoing probabilities whenever a random walker deviates from the reference URW. This informational cost is incurred at each jump, until the target vertex is reached. We normalize the total cumulative informational cost by the walk length. Transmission cost is captured by the same MFPT of the interpolated BRW.

To illustrate the trade-off between transmission and informational costs of the BRW with different λ, we tested it on structural

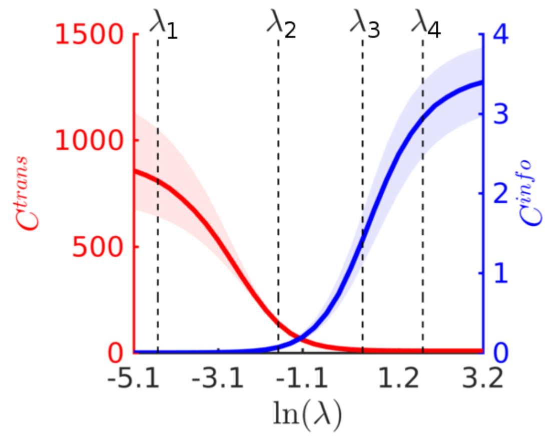

Figure 2: The trade-off between transmission and informational costs of the BRW model with different λ.

brain networks built from diffusion MRI scans[2]. As figure 2 demonstrates, when averaged over all source-target pairs, both costs can be kept at low levels by choosing a well balanced λ value, avoiding both extremes.

4 FUTURE WORK

While the two communication models illustrate the feasibility of efficient global routing with limited information, algorithmic solutions for learning such routing strategies are still difficult. In their seminal work[10], the authors developed an efficient solver for Markov decision process. Our models can be elegantly casted as a special case, with KL-divergence being the control cost and finite horizon cumulative costs for local routing neighborhoods. We plan to develop a unifying Markov decision process formulation, which would allow for efficient solutions and analysis of global navigation patterns using local/partial information on large systems.

REFERENCES

[1] Lada A Adamic, Rajan M Lukose, Amit R Puniyani, and Bernardo A Huberman. 2001. Search in power-law networks. *Physical review E* 64, 4 (2001), 046135.

[2] Richard F Betzel, Alessandra Griffa, Andrea Avena-Koenigsberger, Joaquín Goñi, Jean-Philippe Thiran, Patric Hagmann, and Olaf Sporns. 2013. Multi-scale community organization of the human structural connectome and its relationship with resting-state functional connectivity. *Network Science* 1, 03 (2013), 353–373.

[3] Fan Chung. 2007. The heat kernel as the pagerank of a graph. *Proceedings of the National Academy of Sciences* 104, 50 (2007), 19735–19740.

[4] Özgür Şimşek and David Jensen. 2008. Navigating networks by using homophily and degree. *Proceedings of the National Academy of Sciences* 105, 35 (2008), 12758–12762.

[5] Joaquín Goñi, Andrea Avena-Koenigsberger, Nieves Velez de Mendizabal, Martijn P van den Heuvel, Richard F Betzel, and Olaf Sporns. 2013. Exploring the morphospace of communication efficiency in complex networks. *PLoS One* 8, 3 (2013), e58070.

[6] Jon M. Kleinberg. 2000. Navigation in a small world. *Nature* 406, 6798 (Aug. 2000), 845–845. https://doi.org/10.1038/35022643

[7] Renaud Lambiotte, J.-C. Delvenne, and Mauricio Barahona. 2008. Laplacian dynamics and multiscale modular structure in networks. *arXiv preprint arXiv:0812.1770* (2008). http://arxiv.org/abs/0812.1770

[8] J. Patokallio. 2009. Openflights data. *http://openflights.org/* (2009).

[9] David Peleg and Eli Upfal. 1989. A trade-off between space and efficiency for routing tables. *Journal of the ACM (JACM)* 36, 3 (1989), 510–530.

[10] Emanuel Todorov. 2009. Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences* 106, 28 (July 2009), 11478–11483.