Gossip Algorithms

Gossip Algorithms

Devavrat Shah

Massachusetts Institute of Technology
Cambridge, MA, USA
devavrat@mit.edu

now
the essence of knowledge

Boston – Delft

Full text available at: http://dx.doi.org/10.1561/1300000014

Foundations and Trends® in Networking

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is D. Shah, Gossip Algorithms, Foundation and Trends® in Networking, vol 3, no 1, pp 1–125, 2008

ISBN: 978-1-60198-236-0
© 2009 D. Shah

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, mechanical, photocopying, recording or otherwise, without prior written permission of the publishers.

Photocopying. In the USA: This journal is registered at the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923. Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by now Publishers Inc. for users registered with the Copyright Clearance Center (CCC). The 'services' for users can be found on the internet at: www.copyright.com

For those organizations that have been granted a photocopy license, a separate system of payment has been arranged. Authorization does not extend to other kinds of copying, such as that for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. In the rest of the world: Permission to photocopy must be obtained from the copyright owner. Please apply to now Publishers Inc., PO Box 1024, Hanover, MA 02339, USA; Tel. +1-781-871-0245; www.nowpublishers.com; sales@nowpublishers.com

now Publishers Inc. has an exclusive license to publish this material worldwide. Permission to use this content must be obtained from the copyright license holder. Please apply to now Publishers, PO Box 179, 2600 AD Delft, The Netherlands, www.nowpublishers.com; e-mail: sales@nowpublishers.com

Foundations and Trends® in Networking

Volume 3 Issue 1, 2008

Editorial Board

Editorial Scope

Foundations and Trends® in Networking will publish survey and tutorial articles in the following topics:

- Ad Hoc Wireless Networks
- Sensor Networks
- Optical Networks
- Local Area Networks
- Satellite and Hybrid Networks
- Cellular Networks
- Internet and Web Services
- Protocols and Cross-Layer Design
- Network Coding
- Energy-Efficiency Incentives/Pricing/Utility-based
- Games (co-operative or not)
- Security
- Scalability
- Topology
- Control/Graph-theoretic models
- Dynamics and Asymptotic Behavior of Networks

Information for Librarians

Foundations and Trends® in Networking, 2008, Volume 3, 4 issues. ISSN paper version 1554-057X. ISSN online version 1554-0588. Also available as a combined paper and online subscription.

Foundations and Trends® in
Networking
Vol. 3, No. 1 (2008) 1–125

DOI: 10.1561/1300000014

Gossip Algorithms

Devavrat Shah

Massachusetts Institute of Technology, Cambridge, MA, USA,
devavrat@mit.edu

Abstract

Unlike the Telephone network or the Internet, many of the next generation networks are not *engineered* for the purpose of providing efficient communication between various networked entities. Examples abound: sensor networks, peer-to-peer networks, mobile networks of vehicles and social networks. Indeed, these emerging networks do require algorithms for communication, computation, or merely spreading information. For example, estimation algorithms in sensor networks, broadcasting news through a peer-to-peer network, or viral advertising in a social network. These networks lack infrastructure; they exhibit unpredictable dynamics and they face stringent resource constraints. Therefore, algorithms operating within them need to be extremely simple, distributed, robust against networks dynamics, and efficient in resource utilization.

Gossip algorithms, as the name suggests, are built upon a *gossip* or *rumor* style unreliable, asynchronous information exchange protocol. Due to their immense simplicity and wide applicability, this class of algorithms has emerged as a canonical architectural solution for the next generation networks. This has led to exciting recent progress to understand the applicability as well as limitations of the Gossip algorithms. In this review, we provide a systematic survey of many of these recent results on Gossip network algorithms. The algorithmic results described here utilize interdisciplinary tools from Markov chain theory, Optimization, Percolation, Random graphs, Spectral graph theory, and Coding.

Contents

1

Introduction

The twentieth century has seen a revolution in terms of our ability to communicate at very long distances at very high speeds. This has fundamentally changed the way we live in the present world. The development of reliable and high-performance massive communication networks has been at the heart of this revolution. The telephone networks and the Internet are prime examples of such large networks. These networks were carefully engineered (and are still being engineered) for the single purpose of providing efficient communication given the available resources. In contrast to these networks, there has been a sudden emergence of different types of large networks in the past few years where the primary purpose is not that of providing communication. Examples of such networks include sensor networks, peer-to-peer (P2P) networks, mobile ad-hoc networks, and social networks.

A sensor network, made of a large number of unreliable cheap sensors, is usually deployed for the purpose of ‘sensing’, ‘detecting’ or ‘monitoring’ certain events. For example, smoke sensors capable of wireless transmission deployed for smoke detection in a large building, or a collection of interconnected camera sensors deployed for surveillance in a secure facility. The ability to deploy such networks anywhere with minimal cost of infrastucture has made them particularly attractive for these applications. Clearly, the primary purpose of such networks is to collect and process the sensed information by sensors rather than provide efficient communication.

The peer-to-peer networks are formed by connecting various users (e.g., computers or handheld devices) over an already existing network such as the Internet. Usually such networks are formed with minimal infrastructural support. The peers (or neighbors) are connected over an existing network and hence the advantage of using such networks is not in terms of efficiency of utilizing resources. However, a significant benefit arises in terms of reduced infrastructural support in situations like wide information dissemination. For example, in the absence of a P2P network an Internet content provider (e.g., BBC) needs to maintain a high bandwidth 'server farm' that 'streams' a popular movie or a TV show to a large number of users simultaneously. In contrast, in the presence of a P2P network a user is likely to obtain the desired popular content from a 'nearby' peer and thus distributing a large cost of 'streaming' from the 'server farm' to many 'peers'. Therefore, such an architecture can reduce the cost of content dissemination for a content provider drastically. Of course, it is likely to come at an increased cost of the network utilization. Now, whether or not the benefits obtained in terms of reduced infrastructure by utilizing P2P network for a content provider offset the increased network cost incurred by the network provider is indeed intriguing both in an engineering and an economic sense. While the recent trend suggests that it is indeed the case (e.g., advent of the BBCiPlayer [70] and adaptation of Korean ISPs [31]), the equilibrium solution is yet to be reached.

The mobile ad-hoc network formed between vehicles arises in various scenarios, including future smart cars traveling on road, or fleets of unmanned aerial vehicles deployed for surveillance. These networks, by design, are formed for a purpose other than communication. They need algorithms for the purpose of co-ordination, consensus or flocking (e.g., see classical work by Tsitsiklis [69], more recently [6, 32, 63]).

Finally, we have noticed a very recent emergence of massive social networks between individuals connected over a heterogenous collection of networks. Until recently, an individual's social network usually involved only a small number of other acquintances, relatives or close friends. However, the arrival of 'social network applications' (e.g., Orkut, Facebook, etc.) has totally changed the structure of existing social networks. Specifically, the social network of an individual now includes many more acquintances than before thanks to these online applications. Furthermore, the use of handheld devices like smart phones are likely to create new ways to 'socialize' through P2P networks formed between them in the near future. Naturally, this 'globalization' and 'ubiquitous presence' of social networks bring many exciting opportunities along with extreme challenges. To realize these opportunities and to deal with the challenges, we will need new algorithms with efficient effective social communication under uncertain environmental conditions.

1.1 NextGen Networks: Through an Algorithmic Lens

Algorithms are key building blocks of any network architecture. For example, the Internet provides efficient communication between users through a collection of algorithms operating at the end-users and inside the network. Popular instances of such algorithms are the Transmission Control Protocol (TCP) for congestion control or Border Gateway Protocol (BGP) for routing. The above discussed emerging or next generation networks are not designed to provide efficient communication between the entities or the users networked by them. But, they do require algorithms to enable their primary applications. For example, a sensor network may require an estimation algorithm for event detection given the sensor observations; a P2P network may require a dissemination algorithm using peer information; a network of aerial vehicles may need an algorithm to reach consensus to co-ordinate their surveillance efforts, and an advertiser may need a social network algorithm for efficient 'viral' advertisement.

In most of these next generation networks, algorithms usually need to operate under an 'adverse' environment. First of all, since these networks are not built for providing communication, there is usually a lack of a reliable network infrastructure. Second, these networks are highly dynamic in the sense that nodes may join the network, leave the network, or even become intermittently unavailable in an unpredictable manner. Third, the network is usually highly resource constrained in terms of communication, computation and sometimes energy resources.

The highly constrained environment in which algorithms are operating suggest that the algorithm must posses certain properties so as to be implementable in such networks. Specifically, an algorithm operating at a node of the network should utilize information 'local' to the node and should not expect any static infrastructure. It should attempt to achieve its task iteratively and by means of asynchronous message exchanges. The algorithm should be robust against the network dynamics and should not prescribe to any 'hard-wired' implementation. And finally, the algorithm should utilize minimal computational and communication resources by performing few logical operations per iteration as well as require light-weight data structures. These constraints naturally lead to 'Gossip' algorithms, formally described next, as a canonical algorithmic architectural solution for these next generation networks.

1.2 The Formal Agenda

We shall formally describe the quest for algorithm design for the next generation networks in this section. This will give rise to the formal definition of 'Gossip' algorithms, which will serve as the canonical algorithmic solution.

To this end, let us consider a network of n nodes denoted by $V = \{1,\dots,n\}$. Let $E \subset V \times V$ denote the set of (bidirectional) links along which node pairs can communicate. That is, $(i,j) \in E$ if and only if nodes $i,j \in V$ can communicate with each other. Let this network graph be denoted by $G = (V,E)$. This network graph G should be thought of as changing over time in terms of V and E. As the reader will notice, the algorithms considered here will not utilize any static property of G and hence will be applicable in the presence of explicit network dynamics. For simplicity of the exposition, we shall not model the network dynamics explicitly. Let d_i denote the degree of node i in G, i.e., $d_i = |\{j \in V : (i,j) \in E\}|$. We will assume that the network G is connected without loss of generality; or else we can focus on different connected components separately.

We consider a class of algorithms, called 'Gossip' algorithms, that are operating at each of the n nodes of the network. Now, we present the formal definition of these algorithms.

Definition 1.1 (Gossip algorithms). Under a Gossip algorithm, the operation at any node $i \in V$, must satisfy the following properties:

(1) The algorithm should only utilize information obtained from its neighbors $\mathcal{N}(i) \overset{\triangle}{=} \{j \in V : (i,j) \in E\}$.
(2) The algorithm performs at most $O(d_i \log n)$ amount of computation per unit time.
(3) Let $|F_i|$ be the amount of storage required at node i to generate its output. Then the algorithm maintains $O(\mathrm{poly}(\log n) + |F_i|)$ amount of storage at node i during its running.
(4) The algorithm does not require synchronization between node i and its neighbors, $\mathcal{N}(i)$.
(5) The eventual outcome of the algorithm is not affected by 'reasonable'[1] changes in $\mathcal{N}(i)$ during the course of running of the algorithm.

We wish to design Gossip algorithms for computing a generic network function. Specifically, let each node have some information, and let x_i denote the information of node $i \in V$. The node $i \in V$ wishes to compute a function $f_i(x_1,\ldots,x_n)$ using a Gossip algorithm. Also, it would like to obtain a good estimate of $f_i(x_1,\ldots,x_n)$ as quickly as possible. The question that is central to this survey is that of identifying the dependence of the computation time of the Gossip algorithm over the graph structure G and the functions of interest $f_1,\ldots,f_n$.

Before we embark on the description and organization of this survey, some remarks are in order. First, property (3) rules out 'trivial' algorithms like *first collect values* $x_1,\ldots,x_n$ *at each node and*

---

[1] By a reasonable change, here we mean dynamics that allow for a possibility of eventual computation of the desired function in a distributed manner. For example, if a node i becomes disconnected from the rest of the graph forever, then it will consist of unreasonable change as per our terminology.

then compute $f_i(x_1,\ldots,x_n)$ *locally* for functions like summation, i.e., $f_i(x_1,\ldots,x_n) = \sum_{k=1}^{n} x_k$. This is because for such a function the length of the output is $O(1)$ (we treat storage of each distinct number by unit space) and hence collection of all n items at node i would require storage $\Omega(n)$ which is a violation of property (3). Second, the computation of complex function (e.g., requiring beyond poly$(\log n)$ space) are beyond this class of algorithms. This is to reflect that the interest here is in functions that are easily computable, which is usually the case in the context of network applications. Third, the definition of a Gossip algorithm here should be interpreted as a rough guideline on the class of simple algorithms that are revelant rather than a very precise definition.

1.3 Organization

In the remainder of this survey, we provide a systematic description of the class of network functions that can be computed by means of a Gossip algorithm. A salient feature of the analysis of the algorithms described in this survey is the ability to describe the precise dependance of computation time on the network graph structure G and the function of interest. These dependancies are described in terms of 'spectral-like' graph properties. Therefore, we start with *Preliminaries* on graph properties and some known results that will be useful in the algorithm design and analysis. These are explained through examples of a collection of graph models throughout the survey.

The network functions for which we describe Gossip algorithms in this survey are naturally designed in a 'layered' fashion. At the bottom of the layer lies the design of a robust information layer using a Gossip algorithm. This is described in detail in *Information dissemination*. Here we will describe information dissemination Gossip algorithm for both unicast and multicast types of traffic scenarios. We will describe a natural relation between Percolation on graphs, information dissemination and certain spectral-like graph properties.

The simplest class of iterative algorithms, built upon an unreliable information layer, are based on linear dynamics. These algorithms have been used for solving consensus or multi-agent co-ordination problems classically. We provide a detailed account on the optimal design and analysis of such algorithms in *Linear computation*. Here, we shall describe the interplay between Markov chain theory, mixing times and Gossip algorithms. We also report some advances in the context of Markov chain theory due to considerations from the viewpoint of Gossip algorithms.

Linear function computation is an instance of, and essentially equivalent to, separable function computation. The quest for designing the fastest possible Gossip algorithm, in terms of its dependence on the graph structure, for separable function computation, which will be left partly unresolved by the linear dynamics based algorithms, will be brought to a conclusion in *Separable function computation*. Here, we shall describe an algorithm based on an 'extremal' property of the Exponential distribution. This algorithm will utilize the unreliable information layer designed in *Information dissemination* for the purpose of information exchange. The appropriately quantized version of this algorithm as well as information theoretic arguments suggesting its fundamental optimality will be discussed (see 'Summary') as well.

Next, we consider Gossip algorithm design for the task of scheduling in constrained queueing networks. This is a key operational question for networks such as those operating over a common wireless medium. For such a network a scheduling algorithm is required for the media access control (MAC). We describe Gossip scheduling algorithm in *Network scheduling*. This algorithm builds upon the separable function computation algorithm using clever randomization.

Network resource allocation is another fundamental problem that is faced while operating a communication network. Under flow-level modeling of a network, this involves solving certain network-wide or global constrained convex optimization problems. Therefore, we consider the question of designing a Gossip algorithm for a class of convex optimization problems in *Network convex optimization*. This algorithm, like network scheduling, builds upon the separable function computation algorithm. Specifically, it utilizes the separable function computation algorithm to design a 'distributed computation' layer.

In summary, the algorithms presented in this survey provide 'layers' of computation in a network. The key reason for the existence of such a ‘layered’ algorithmic architecture lies in the ability to ‘functionally decompose’ many interesting problems with separable function computation central to the decomposition. For this reason, Gossip algorithm for separable function computation becomes a key ‘subroutine’ in designing Gossip algorithms for many seemingly complex network computation problems. For these reasons, in addition to applications described in this survey, the separable function computation algorithm can be used to design Gossip algorithms for other important applications including spectral decomposition (using the algorithm of Kempe and McSherry [38] and the separable function computation algorithm) and Kalman filtering.

References

[1] D. J. Aldous, “Some inequalities for reversible Markov chains,” *Journal of the London Mathematical Society*, vol. 25, pp. 564–576, 1982.

[2] P. Assouad, “Plongements lipschitziens dans $\mathbf{R}^n$,” *Bulletin de la Société Mathématique de France*, vol. 111, no. 4, pp. 429–448, 1983.

[3] O. Ayaso, “Information theoretic approaches to distributed function computation,” PhD thesis, Massachusetts Institute of Technology, 2008.

[4] Y. Bartal, J. W. Byers, and D. Raz, “Fast, distributed approximation algorithms for positive linear programming with applications to flow control,” *SIAM Journal on Computing*, vol. 33, no. 6, pp. 1261–1279, 2004.

[5] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.

[6] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, “Convergence in multiagent coordination, consensus, and flocking,” in *Joint 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC’05)*, December 2005.

[7] S. Boyd, P. Diaconis, and L. Xiao, “Fastest mixing Markov chain on a graph,” *SIAM Review*, 2004.

[8] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” *IEEE/ACM Transaction on Networking*, vol. 14, no. SI, pp. 2508–2530, 2006.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[10] P. Chaporkar, K. Kar, and S. Sarkar, “Throughput guarantees through maximal scheduling in wireless networks,” in *43rd Allerton Conference on Communication Control and Computing*, 2005.

[11] L. Chen, S. H. Low, M. Chang, and J. C. Doyle, "Optimal cross-layer congestion control, routing and scheduling design in ad-hoc wireless networks," in *IEEE INFOCOM*, 2006.

[12] J. Considine, F. Li, G. Kollios, and J. W. Byers, "Approximate aggregation techniques for sensor databases," in *20th IEEE International Conference on Data Engineering*, April 2004.

[13] J. Dai and B. Prabhakar, "The throughput of switches with and without speed-up," in *Proceedings of IEEE Infocom*, pp. 556–564, 2000.

[14] S. Deb, M. Médard, and C. Choute, "Algebraic gossip: A network coding approach to optimal multiple rumor mongering," *IEEE/ACM Transactions on Networking*, vol. 14, 2006.

[15] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Springer, Second Edition, 1998.

[16] P. Diaconis, S. Holmes, and R. Neal, "Analysis of a non-reversible Markov chain sampler," *Annals of Applied Probability*, vol. 10, pp. 726–752, 2000.

[17] P. Diaconis and L. Saloff-Coste, "Moderate growth and random walk on finite groups," *Geometric and Functional Analysis*, vol. 4, no. 1, pp. 1–36, 1994.

[18] A. G. Dimakis, A. D. Sarwate, and M. J. Wainwright, "Geographic gossip: Efficient aggregation for sensor networks," in *5th International ACM/IEEE Symposium on Information Processing in Sensor Networks (IPSN '06)*, April 2006.

[19] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal throughput-delay scaling in wireless networks-part I: The fluid model," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2568–2592, 2006.

[20] L. Elsner, I. Koltracht, and M. Neumann, "On the convergence of asynchronous paracontractions with applications to tomographic reconstruction from incomplete data," *Linear Algebra and Its Applications*, no. 130, pp. 65–82, 1990.

[21] M. Enachescu, A. Goel, R. Govindan, and R. Motwani, "Scale free aggregation in sensor networks," in *International Workshop on Algorithmic Aspects of Wireless Sensor Networks*, 2004.

[22] A. Eryilmaz, A. Ozdaglar, D. Shah, and E. Modiano, "Distributed cross-layer algorithms for the optimal control of multi-hop wireless networks," *IEEE/ACM Transactions on Networking* (accepted to appear), 2009.

[23] P. Flajolet and G. N. Martin, "Probabilistic counting algorithms for data base applications," *Journal of Computer and System Science*, vol. 31, no. 2, pp. 182–209, 1985.

[24] R. G. Gallager, "A minimum delay routing algorithm using distributed computation," *IEEE Transactions on Communications*, vol. COM-25, no. l, pp. 73–85, 1977.

[25] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*.

[26] N. Garg and N. E. Young, "On-line end-to-end congestion control," in *IEEE FOCS*, pp. 303–312, 2002.

[27] P. Giaccone, B. Prabhakar, and D. Shah, "Randomized scheduling algorithms for high-aggregate bandwidth switches," *IEEE Journal of Select Areas*

Communication High-performance Electronic Switches/Routers for High-speed Internet, vol. 21, no. 4, pp. 546–559, 2003.

[28] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transaction on Information Theory*, vol. 46, no. 2, pp. 388–404, March 2000.

[29] B. Hajek and G. Sasaki, "Link scheduling in polynomial time," *IEEE Transactions on Information Theory*, vol. 34, 1988.

[30] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.

[31] http://www.reuters.com/article/pressrelease/idUS55597+16-Jan-2008+BW 20080116.

[32] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[33] K. Jung and D. Shah, "Low delay scheduling in wireless network," in *IEEE ISIT*, 2007.

[34] K. Jung, D. Shah, and J. Shin, "Minimizing rate of convergence for iterative algorithms," *IEEE Transactions on Information Theory* (accepted to appear), 2009.

[35] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus," (in press).

[36] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, March 1998.

[37] D. Kempe, A. Dobra, and J. Gehrke, "Gossip-based computation of aggregate information," in *FOCS '03: Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, p. 482, Washington, DC, USA: IEEE Computer Society, 2003.

[38] D. Kempe and F. McSherry, "A decentralized algorithm for spectral analaysis," in *Symposium on Theory of Computing*, ACM, 2004.

[39] M. Koubarakis, C. Tryfonopoulos, S. Idreos, and Y. Drougas, "Selective information dissemination in P2P networks: Problems and solutions," *SIGMOD Record*, vol. 32, no. 3, pp. 71–76, 2003.

[40] X. Lin, N. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *Submitted,* Available Through csl.uiuc.edu/rsrikant, 2006.

[41] X. Lin and N. B. Shroff, "Impact of imperfect scheduling in wireless networks," in *IEEE INFOCOM*, 2005.

[42] N. Linial and A. Wigderson, "Lecture notes on Expander Graphs," http://www.math.ias.edu/∼avi/BOOKS/expanderbookr1.pdf.

[43] L. Lovasz and P. Winkler, "Mixing times," in *Microsurveys in Discrete Probability, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, (D. Aldous and J. Propp, eds.), pp. 85–133, AMS, 1998.

[44] M. Luby and N. Nisan, "A parallel approximation algorithm for positive linear programming," in *Proceedings of the 25th Annual ACM Symposium on Theory of Computing*, pp. 448–457, 1993.

[45] R. Madan, D. Shah, and O. Leveque, "Product multi-commodity flow in wireless networks," *IEEE Transactions on Information Theory*, vol. 54, no. 4, pp. 1460–1476, 2008.

[46] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "Tag: A tiny aggregation service for ad-hoc sensor networks," *SIGOPS Operating Systems Review*, vol. 36, no. SI, pp. 131–146, 2002.
[47] L. Massoulié and M. Vojnovic, "Coupon replication systems," in *ACM SIGMETRICS/Performance*, 2005.
[48] N. McKeown, "iSLIP: A scheduling algorithm for input-queued switches," *IEEE Transactions on Networking*, vol. 7, no. 2, pp. 188–201, 1999.
[49] N. McKeown, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch," in *Proceedings of IEEE Infocom*, pp. 296–302, 1996.
[50] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*. London: Springer-Verlag, 1993.
[51] E. Modiano, D. Shah, and G. Zussman, "Maximizing throughput in wireless network via gossiping," in *ACM SIGMETRICS/Performance*, 2006.
[52] R. Montenegro and P. Tetali, "Mathematical aspects of mixing times in Markov chains," *Foundations and Trends in Theoretical Computer Science*, vol. 1, no. 3, pp. 237–354, 2006.
[53] D. Mosk-Aoyama, T. Roughgarden, and D. Shah, "Fully distributed algorithms for convex optimization problems," in *International Symposium on Distributed Computation (DISC)*, 2007.
[54] D. Mosk-Aoyama and D. Shah, "Information dissemination via network coding," in *IEEE ISIT*, 2006.
[55] D. Mosk-Aoyama and D. Shah, "Fast distributed algorithms for computing separable functions," *IEEE Transactions on Information Theory*, vol. 54, no. 7, pp. 2997–3007, 2008.
[56] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *LIDS Report 2755, to appear in IEEE Transactions on Automatic Control*, 2008.
[57] C. Papadimitriou and M. Yannakakis, "Linear programming without the matrix," in *Twenty-Fifth Annual ACM Symposium on the Theory of Computing*, 1993.
[58] M. Penrose, *Random Geometric Graphs. Oxford Studies in Probability*, Oxford: Oxford University Press, 2003.
[59] D. Qiu and R. Srikant, "Modeling and performance analysis of bittorrent-like peer-to-peer networks," in *ACM SIGCOMM*, pp. 367–378, 2004.
[60] S. Rajagopalan, D. Shah, and J. Shin, "Network adiabatic theorem: An efficient randomized protocol for contention resolution," in *ACM SIGMETRICS/Performance*, 2009.
[61] O. Reingold, A. Wigderson, and S. Vadhan, "Entropy waves, The zig-zag graph product, and new constant-degree expanders and extractors," *Annals of Mathematics*, 2002.
[62] T. Rockafellar, *Network Flows and Monotropic Optimization.* Wiley-Interscience, (republished by Athena Scientific, 1998), 1984.
[63] K. Savla, F. Bullo, and E. Frazzoli, "On traveling salesperson problems for Dubins' vehicle: stochastic and dynamic environments," in *IEEE CDC-ECC*, pp. 4530–4535, Seville, Spain, December 2005.

[64] D. Shah, “Stable algorithms for input queued switches,” in *Proceedings of Allerton Conference on Communication, Control and Computing*, 2001.

[65] D. Shah and D. J. Wischik, “Optimal scheduling algorithm for input queued switch,” in *IEEE INFOCOM*, 2006.

[66] R. Srikant, *The Mathematics of Internet Congestion Control*. Birkhäuser, 2004.

[67] L. Tassiulas and A. Ephremides, “Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks,” *IEEE Transactions on Automatic Control*, vol. 37, pp. 1936–1948, 1992.

[68] L. Trevisan, “Non-approximability results for optimization problems on bounded degree instances,” in *ACM STOC*, 2001.

[69] J. Tsitsiklis, “Problems in decentralized decision making and computation,” PhD dissertation, Lab. Information and Decision Systems, MIT, Cambridge, MA, 1984.

[70] www.bbc.co.uk/iplayer.