Wireless Networks Science: An Optimization View

Alejandro Ribeiro
Dept. of Electrical & Systems Engineering, University of Pennsylvania

Communication networks are deployed to transport information from generating sources to intended destinations. There are different ways in which a given network can accomplish this task, but among these there are some that optimize properly chosen criteria. Hence, optimization is a natural basis for a Wireless Networks Science. In this paper we will show how optimization theory can be used to understand fundamental properties of wireless networks.

Currently, wireless communication networks are designed based on heuristics. While this has led to the emergence of revolutionizing technology, our scientific understanding of fundamental issues arising in wireless networks remains limited. Developing a formal theory of wireless networks will lead to rigorous approaches to wireless network's design. This shift from current heuristic based approaches to future rigorous formulations we foresee it as analogous to what has happened with digital communications in the last fifty years.

Optimization as a mathematical tool to understand networking protocols appears first in the form of the network utility maximization (NUM) framework [1]. The insight provided there is that congestion control algorithms can be understood as distributed implementations of algorithms that solve utility maximization problems. Over time this has been significantly generalized. The modern understanding is that to support information flows with some required level of service communication networks administer given resources. In conventional wired networks the resource given is a set of physical connections between nodes. Supporting information flows requires finding routes between source and destination, determining link sharing strategies, and controlling the amount of traffic injected into the network, among other subproblems. It was an early design specification to separate these problems in layers – routing, link and transport for the problems in the previous sentence – that operate more or less independently, and interact through standardized interfaces. While this was mostly a matter of ensuring inter-operability it is remarkable that this separation can be optimal. Specifically, it is possible to define separate per-layer optimization problems whose outcome coincides with the solution of a joint non-layered optimization. Mathematically, the separability comes from the fact that the wired networking problem is convex – a linear program in fact. The Lagrangian dual problem can thus be solved instead. As it frequently happens, the Lagrangian exhibits a separable structure, which, as it turns out coincides with the conventional layers.

Extending the NUM framework to wireless networking problems is not a simple pursuit. In a wireless network, the given resources are not connections but bandwidth and power. Therefore, on top of routes, link shares and rate control, a wireless networking problem entails determining which connections, among those possible for the given bandwidth and power, should be established to support the required level of service. Earlier approaches to wireless networking migrated the conventional layers and defined the power and frequency assignment as physical layer subproblems. This yields poor results though, and over time lead to the surge of cross-layer design as synonym of joint optimization across layers; see e.g., [2]. Ultimately, the poor performance of layered wireless networks stems from the non-convexity of the joint cross-layer networking problem. As in wired networks, the Lagrangian exhibits a separable structure that can be mapped to layers as has been observed in, e.g., [3] and [4]. But non-convex problems have positive duality gap explaining the markedly suboptimal performance of layered wireless networks.

In a recent paper, [5] we have proved that general wireless networking problems in the presence of fading, while non-convex, have zero Lagrangian duality gap. Exploiting the separability of the Lagrangian, this result yields the following fundamental separation principles:

First separation principle of wireless networking. This principle pertains to the separability of wireless networking problems into layers. It states that it is possible to define separate optimization problems to obtain optimal routes, optimal link capacity allocations, and optimal power/frequency assignments.

Second separation principle. Another difficulty in optimal wireless networking is the need to optimize jointly over all fading states. Given that fading coefficients take on a continuum of values, this is a variational problem that requires finding optimal functions of the fading coefficients. This principle asserts that the network optimization is further separable in per-fading-state subproblems. The practical importance of this result is that it is not necessary to find optimal functions but only the values of the functions for those channels actually observed.

It is worth remarking that these results depend on the existence of fading. In fact, in deriving the separation principles the randomness introduced by fading channels is determinant. In this sense these results are a manifestation of the surprising but nonetheless well known fact that randomness generates structure. This has been nowhere better stated than in Gnedenko and Kolmogorov's classical work on limit distributions: "the epistemological value of the theory of probability is based on this: that large-scale random phenomena in their collective action create strict, nonrandom regularity." It is often possible to infer properties of large-scale stochastic systems even if analogous deterministic counterparts are intractable. For wireless networks, randomness – inherently present because of fading – leads to the optimality of layered architectures and the separability in per-fading state subproblems.

In this paper we will survey the separation principles and describe how optimization theory and the structure introduced by randomness can be leveraged as the building blocks of a Wireless Network Science.

REFERENCES

[1] F. P. Kelly, A. Maulloo, and D. Tan., "Rate control for communication networks: shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, 1998.

[2] X. Lin, N. B. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1452–1463, August 2006.

[3] L. Xiao, M. Johansson, and S. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1136–1144, July 2004.

[4] M. Chiang, S. H. Low, R. A. Calderbank, and J. C. Doyle, "Layering as optimization decomposition," *Proc IEEE*, vol. 95, no. 1, pp. 255–312, January 2007.

[5] Alejandro Ribeiro and Georgios B. Giannakis, "Separation theorems of wireless networking," *IEEE Trans. Inf. Theory*, December 2007 (submitted), Available at http:// www.seas.upenn.edu/˜aribeiro/.