The Geometric Theory of Network Routing: Final Grant Report

Gábor Rétvári (retvari@tmit.bme.hu)

Introduction

With the proliferation of new devices (mobile, Internet-of-Things), new system architectures (Software Defined Networks), and new applications (social, video, etc.), the fundamental design principles that underlie the today's telecommunications network architectures are challenged on a daily basis. These new challenges led many to propose a major rethinking of the Internet. Of course, such a brave undertaking needs a well-established theoretical foundation, one that in many respects seems to go missing today. A key mission of the research work under the support of the grant was to *augment the toolset of network engineers with new scientific principles and engineering methodology*, in order to help operating today's increasingly complex networks and designing a more versatile future Internet.

The main goal of any communications networks is the high-quality, reliable and economical transport of traffic. In a large-scale network, even the very discovery of existing transport paths between users is already a difficult challenge. And even if such paths are found, load must be distributed on these paths in a way as to avoid that the traffic of independent user sessions, competing for the scarce resources, interfere at the congested bottlenecks of the network infrastructure. A key observation and organizing scheme behind the research covered by the report was the revelation that a geometric view can be beneficial for answering these (and many other) questions related to today's network practice.

The model

In the below description we briefly recall the geometric theory of network routing as we introduced it in [1]. The basic idea is to associate geometric objects with the most important entities participating in the process of traffic forwarding, namely, the network topology, the users, the traffic demands, and the routing algorithms, which allows us to study their relationship in purely geometric terms.

Suppose we are given a network G and a set of users. As the first ingredient of the model, consider two Euclidean spaces: the flow space $\mathcal{M}$, describing the set of all possible assignments of flows to the paths in G, and the demand space $\mathcal{T}$ representing the set of traffic demand combinations that can be posed by the users. These two spaces serve as the "playing grounds" for competing users. Then, we introduce two geometric objects: the *flow polytope* $M(G)$ is the subset of M that corresponds to feasible routings subject to link capacity constraints and the *demand polytope* $T(G) \subset \mathcal{T}$ is the set of demand combinations that can be realized in the network without link overload. The intuition is that any point contained in $M(G)$ represents a routing one can establish in the network without overloading any of the links, and any point in $T(G)$ represents a particular combination of user demands for which there is a routing in $M(G)$. We observe that both $M(G)$ and $T(G)$ are (under very mild regularity conditions) compact, convex, full-dimensional, and down-monotone polyhedra. Furthermore, they are intricately related to each other, in that $T(G)$ is an affine projection of $M(G)$.

The second key ingredient in network geometry is the notion of routing functions. A routing algorithm, in certain sense, is a way to map user demands to actual physical paths, and this can be illustratively represented by a routing function $\mathcal{S} : \mathcal{T} \mapsto \mathcal{M}$. Intuitively, $\mathcal{S}$ embodies the very knowledge of a routing algorithm of how to choose the right amount of user traffic to be sent to individual paths in the network. Notationally, given some traffic matrix θ, the corresponding routing is obtained by applying $\mathcal{S}$ to θ: $u = \mathcal{S}(\theta)$.

Armed with these notions, we can now pose the fundamental task of network routing as *to route as many different combinations of user demands as possible without congestion.* This is spectacularly easy to express in terms of the geometric construct: the task is to map as large part of the interior of the demand polytope *inside* the flow polytope as possible, i.e., to solve the geometric optimization problem $\min_{\mathcal{S}} \alpha : \mathcal{S}(T(G)) \subseteq \alpha M(G)$.

The geometry of routing: polyhedral view

In this section, we describe the research work that was centered at the model for "routing geometry" as posed in the previous section. In the subsequent section, we shall discuss certain digressions from this main line of thought which we followed throughout the 3 years of grant work, in order to widen the range of questions that we can effectively tackle using the geometric model.

Centralized vs distributed routing architectures

First, we describe how the model was used to investigate one of the most notorious domains of systems science, namely, the organization of the routing logic [2]. There are many ways one can engineer a routing architecture and the decisions the architect makes at the initial phase of the design process greatly affects the eventual efficiency, dependability, and sustainability of the resultant protocols.

One such design axis is whether to apply a *static or an adaptive* scheme. In the context of network routing, a static scheme is just what it is, agnostic to the actual user demands; routes are fixed for the life-time of the network and change at only a very coarse time scale (e.g., upon a topology change). In contrast, in an adaptive scheme routes are constantly optimized with respect to the actual user demands, availability of free capacity, etc. Easily, a static scheme is simple but may not be robust, while an adaptive scheme may be more efficient but requires a complex control mechanism. A remarkable related finding is due to Räcke, who showed that, rather unexpectedly, there exists a special static routing scheme, the so called *(demand-)oblivious routing*, simultaneously optimized for *all* possible traffic demands, so that we pay only a logarithmic factor in congestion compared to the best possible adaptive scheme that is fully aware of demands.

Another question is the *centralized vs distributed* debate; there is permanent struggle inside the networking community whether to go with a centralized architecture where all control decisions are made at one central place that has all necessary network state readily available, or a distributed one that considers the network as a massively parallel processor and distributes control logic to the network nodes proper.

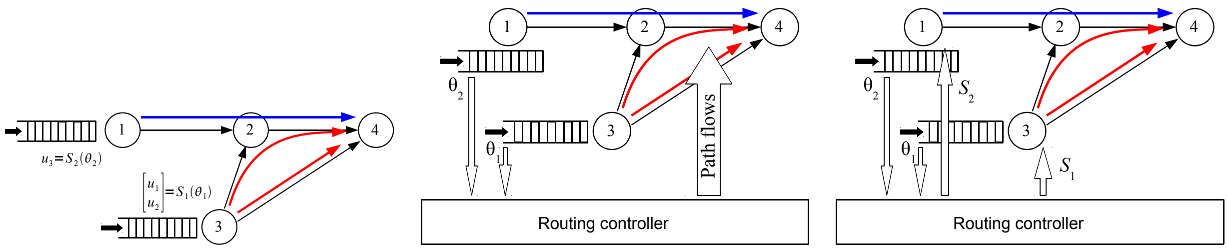
The main result of our work under the grant was the *description of adaptive, static, centralized, and distributed routing architectures in a single mathematical framework as provided by our geometric model and investigate the related fundamental trade-offs.* The main contributions were as follows.

- We introduced the notion of generalized demand-oblivious routing, in which a routing function is pre-computed to minimize congestion with respect to *any* user demands and we cast conventional oblivious routing in this model.
- We proved that for any network a centralized adaptive routing scheme exists that can route any combination of feasible user demands with zero congestion.
- We presented the first theoretical upper bound on the complexity of any optimal centralized algorithm.
- We gave a hybrid distributed-centralized algorithm that unifies the advantages of the two worlds and we proved that our hybrid scheme is asymptotically optimal.

The corresponding work was published in *Wiley Networks* [2]. The paper's material effectively covers the first two years of the work plan.

The first part effectively gives a tutorial into the model what we call "the geometry of networks" in this report, introduces the essential geometric objects with concise notation, collects the necessary theoretical and methodological background from the literature, and states important geometric properties. The second part introduces the notion of *generalized oblivious routing* as the routing architecture under the below general design principles:

- Principle-1: The routing function does not change with the demands.
- Principle-2: The amount of flow placed to a path by the routing function depends on the traffic matrix exclusively.
- Principle-3: The routing function minimizes the maximum link utilization over any traffic matrix that can appear in the network.



(a) Distributed routing architecture (b) Centralized routing architecture (c) Hybrid routing architecture

Distributed, centralized, and hybrid generalized oblivious routing architectures.

The key observation is that, depending on the way we choose the routing function, we can obtain different routing architectures and we can reason about the pros and cons *within a single mathematical framework*.

The second part contains these main results. We show that generalized oblivious routing simplifies into conventional oblivious routing when the routing function is taken as *singular and block-diagonal*, which corresponds to the "static/distributed" extreme in the design spectrum (see Fig. 1). Then, we show that *compound, affine* routing functions give rise to an "adaptive/centralized" routing architecture and we show that for any network there is a globally optimal and continuous routing that routes any demand with zero-congestion, despite that *the routing function is still precomputed*. We also give two methods to compute this routing function, one based on multiparametric linear programming and another purely geometric one using boundary triangulations. Finally, we observe that *compound, block-diagonal* routing functions give rise to a completely new model for routing, the so called *hybrid distributed-centralized demand-oblivious routing architecture*, which occupies an interesting sweet spot in the design space. The nice touch is that this scheme requires minimal and very coarse-grained central control and the lions' share of work is still done in a very scalable and robust distributed mechanism, yet, as we could prove, this scheme can still be asymptotically optimal. We give a cutting plane method to compute the optimal control regions for obtaining efficient hybrid routing functions.

Finally, we evaluate the algorithms in extensive numerical studies and we find that distributed schemes cause prohibitive congestion and centralized schemes can become overly complex in large networks, while our hybrid scheme performs well in many realistic scenarios with only reasonable and tolerable complexity.

A summary and evaluation of generalized oblivious routing is given in the below table.

Routing architecture	Routing function	Pros	Cons
Distributed	singular block-diagonal	scalable, continuous, stable	inefficient
Centralized	compound general affine	optimal, optimizable, continuous	complex, unstable
Hybrid	compound block-diagonal	simple, asymptotically optimal, stable	not optimizable, piecewise continuous

Geometric routing: software tool

The Matlab Routing toolboX (`mrx`) is a collection of data structures, utilities, algorithms, and visualization tools we have written throughout our work on the project. The toolset contains comprehensive collection of classes to compute oblivious routings in networks, competitive ratios, and different forms of performance indexes for distributed, centralized, and hybrid routing schemes. In the course of the project work we made `mrx` available to the research community as open source software (under the GLPv2), `https://github.com/ng201/mrx`.

A disclaimer

In preparation and under different phases of submission are two additional journal papers. The first one is a survey paper on the main geometric concepts that underlie the theory, intended to serve as a reference for future work. This paper gathers the information existing scattered throughout the vast literature on operations research, combinatorics, and computational geometry, under the umbrella of our geometric model. The second paper presents a new application of the geometric model, a statistical characterization of the efficiency of general rate-adaptive routing algorithms. As a main contribution, the approach allowed us to show that in certain networks the probability that a randomly chosen admissible traffic matrix causes congestion with oblivious routing rapidly approaches 1 with the increase of the number of users, meaning that the competitiveness of oblivious routing, an important cornerstone of today's network theory, is an illusion. The first paper was close to submission and the second one was accepted (with revisions) to *Elsevier Performance Evaluation* (a short version is available as [3]).

Unfortunately, due to circumstances beyond my control I had to withdraw both papers. The reason is a severe act of plagiarism and intellectual property theft committed against me by the PhD student who was working with me full-time on this research. This not just precluded the publication of these papers all together but also ceased our collaboration abruptly. This has been a major setback in my scientific career and a huge shock I will hardly ever leave behind entirely.

Approximation of the throughout polytope

Only a single TODO item has remained for the final year of the project: efficient approximation of the geometric bodies underlying the theory. Using the travel support available to me under the grant, answering to the invitation of Prof Michael Schapira I visited the Hebrew University of Jerusalem for roughly 3 months during 2014 and 2015 to seek collaboration for solving this problem. I had the chance to present my work to one of the most renowned experts of the field, Prof Yair Bartal, the father of approximate metric embedding theory. We agreed that a key to the efficient approximation of the demand polytope is oblivious routing, based on the observation that the feasible region of the oblivious routing already gives a logarithmic approximation to the demand polytope. Collaborating with Marco Chiesa, from the Roma Tre University, we came up with a fundamentally new direct proof for this claim. I plan to finalize the work and publish the results in the coming months.

The geometry of routing: alternative approaches

During the 3 years of the project, I have made substantial research efforts to using geometric arguments in various problem domains in networking. This line of research does not fall strictly under the above geometric routing model, but since there has been significant cross-fertilization between my activities I feel important to give a short overview. Below I report on two papers that we wrote during the project's lifetime and what I believe constitute the two most important scientific successes of my entire professional career.

The insights gained in the course of the research on routing geometry led us to study the hidden metric spaces behind large-scale networks in nature. Such massive networks include metabolic networks, road networks, the Internet, or the human brain, and a common organizing scheme behind these networks that there is an inherent geometric space, a hidden metric space, that governs their growth and the process of information transport and navigation. Using game theory, we were able to show that minimalistic networks designed to maximize the navigation efficiency at minimal cost share basic structural properties with these real networks. These idealistic networks are Nash equilibria of a network construction game whose purpose is to find an optimal trade-off between network cost and navigability, subject to the geometric space into which nodes are embedded.

The results were published in *Nature Communications* [4] and got substantial media coverage, see e.g.:

- "Hatalmas sikert értek el a BME kutatói", `http://index.hu/tudomany/2015/07/06/bme_vik_nature_halozatok_nash_equilibrium`.
- "Újabb lépés az olcsóbb és megbízhatóbb internet felé", `http://mta.hu/tudomany_hirei/ujabb-lepes-az-olcsobb-es-megbizhatobb-internet-fele-136593`.

- "Researchers find the organization of the human brain to be nearly ideal", `http://phys.org/news/2015-07-human-brain-ideal.html`.

We presented the work at the meeting of the *Section of Engineering Sciences* of the *Hungarian Academy of Sciences* (MTA) and demonstrated the results at the *ITU Telecom World 2015* held in Budapest.

The geometric approach sparked an interesting side-project. Since most networks show remarkable geometric structure, even very simple navigation schemes usually prove very efficient, even ones that make zero use of global knowledge on network topology, number of users, etc. This led us to argue that routing tables, the data structures needed to be stored at network nodes for facilitating communication, may be highly structured and, correspondingly, compressible. The corresponding paper, studying the compressibility of Internet forwarding tables, was the first one to point out the intricate and deep relations between geometry, information-theory, and the fundamental scalability of the Internet routing architecture. We published this paper at what is undeniably the highest-ranked conference in the field of computer communications, *ACM SIGCOMM*, and I had the chance to present the work to an audience of more than 700 people in Hong Kong [5]. To demonstrate the scale of this result, note that there has been only one occasion that a Hungarian paper made it to SIGCOMM and this happened more than 10 years ago. The paper sparked quite a number of interesting conversations at the conference, as everyone seemed to be intrigued by the possibility of linking networking to the well-established and deep field of information theory. Correspondingly, we published a substantial body of further work on the information-theoretical aspects of network routing in a series of 5 conference papers and a journal paper, see [6, 5, 7, 8, 9, 10]. See the ensuing discussion at `layer9.org` and in the *ACM Computer Communication Review*:

- `http://www.layer9.org/2013/08/sigcomm2013-compressing-ip-forwarding.html`
- `http://www.sigcomm.org/ccr/papers/2013/January/2427036.2427043`

Further results

Together with János Tapolcai, we have successfully won and finished a project under the illustrious Google Faculty Research Award programme. Google Research Awards are one-year awards structured as unrestricted gifts to support the work of world-class researchers around the globe, working on projects that might have considerable impact on Google's topics of interest. Our project was aimed at building efficient succinct bitstring encoders, both software-wise and hardware-wise, that could help making compressed data structures go mainstream. See the media coverage at:

- "Lendületes kutatócsoport nemzetközi elismerése",
  `http://mta.hu/mta_hirei/lenduletes-kutatocsoport-nemzetkozi-elismerese-132586`.
- "Google elismerés a BME VIK TMIT kutatócsoportjának",
  `https://www.vik.bme.hu/hir/562-google-elismeres-a-bme-vik-tmit-kutatocsoportjanak`.

I also achieved some important personal career milestones during the three years of the project. The Budapest University of Technology and Economics (BME), Faculty of Electrical Engineering and Informatics (VIK), recognized my research work in 2013 by awarding me the Senior Research Fellow (tudományos főmunkatárs) status. I became regular member of the *IEEE Infocom* Technical Program Committee, I visited the Montreal University, and I spent 3 months visiting the Hebrew University of Jerusalem, one of the highest-ranked universities around the world. Since last year, I also teach a full-time MSc course on the workings of the Internet Ecosystem.

Publication activities

Apart from these activities, my "regular" research projects have also been ongoing. This research, ranging from traffic-engineering to fast IP resilience, has also born fruit, in the form of publications at top-notch conferences (*ACM Hotnets 2012 and 2014*, *IEEE Infocom'13*, *IFIP Networking'13*, *IEEE ICNP 2014 and 2015*, etc.) and journals (*Springer Distributed Computing*, *Elsevier Computer Communications*, *Springer Telecommunication Systems Journal*, *IEEE-ACM Transactions on Networking*, etc.)

References

[1] G. Rétvári and G. Németh, "Demand-oblivious routing: Distributed vs. centralized approaches," in *IEEE INFOCOM 2010*, March 2010.

[2] G. Németh and G. Rétvári, "Rate-adaptive multipath routing: Distributed, centralized, and hybrid architectures," *Networks*, pp. 1–12, 2015.

[3] G. Németh and G. Rétvári, "Towards a statistical characterization of the competitiveness of oblivious routing (short paper)," *SIGMETRICS Perform. Eval. Rev.*, vol. 40, pp. 387–388, June 2012.

[4] A. Gulyás, J. J. Bíró, A. Kőrösi, G. Rétvári, and D. Krioukov, "Navigable networks as Nash equilibria of navigation games," *Nature Communications*, vol. 6, July 2015.

[5] G. Rétvári, J. Tapolcai, A. Kőrösi, A. Majdán, and Z. Heszberger, "Compressing IP forwarding tables: towards entropy bounds and beyond," in *ACM SIGCOMM 2013*, pp. 111–122, 2013.

[6] G. Rétvári, Z. Csernátony, A. Kőrösi, J. Tapolcai, A. Császár, G. Enyedi, and G. Pongrácz, "Compressing IP forwarding tables for fun and profit," in *ACM HotNets-XI*, ACM, October 2012.

[7] A. Kőrösi, J. Tapolcai, B. Mihálka, G. Mészáros, and G. Rétvári, "Compressing IP forwarding tables: Realizing information-theoretical space bounds and fast lookups simultaneously," in *Proc. IEEE ICNP*, 2014.

[8] B. Mihálka, A. Kőrösi, and G. Rétvári, "Compressing virtual forwarding information bases using the trie-folding algorithm," in *Advances in Communication Networking* (Y. Kermarrec, ed.), vol. 8846 of *Lecture Notes in Computer Science*, pp. 121–133, Springer International Publishing, 2014.

[9] G. Rétvári, J. Tapolcai, A. Kõrösi, A. Majdán, and Z. Heszberger, "Compressing IP forwarding tables: Towards entropy bounds and beyond," *Networking, IEEE/ACM Transactions on*, vol. PP, no. 99, pp. 1–1, 2014.

[10] G. Rétvári, D. Szabó, A. Gulyás, A. Kőrösi, and J. Tapolcai, "An information-theoretic approach to routing scalability," in *ACM HotNets-XIII*, pp. 2:1–2:7, 2014.