Flows through random networks

David Aldous – April 2006

Title brings to mind many somewhat-related topics; is there a core theory?

- General setup

- 4 specific models/problems under study

Graph: has vertices and edges
Network: a graph with some context-dependent extra structure. We consider toy models of networks (**transportation/communication**) whose purpose is to move stuff/information from one place to another.

Assume edges have **lengths** or **costs**. Could take the default "edge-length = 1" but taking generic real lengths is more convenient because it gives **unique shortest paths**.

Study deterministic flows ("fluid", as in the max-flow min-cut theorem) with simultaneous flows between different source-destination pairs (**multicommodity flow**). Take simplest case: constant flow between each source-destination pair.

Given some notion of the **cost** of a flow (e.g. route-length) and some constraints (e.g. edge capacities) we seek the minimum-cost routing.

Deterministic algorithmic problems like this are studied as part of **network algorithms**; as multicommodity flow problems they are NP-hard in general. We take **statistical physics viewpoint** of modeling the network (topology, costs, constraints) as random and studying properties of optimal solution. We take transportation measure uniform on all (source,destination) pairs, so there's one parameter

$$\rho = \text{normalized traffic demand}$$

normalized with n so that flow volume across typical edge is order 1.

Seek to study (in different models on n-vertex networks) the $n \to \infty$ limit curves giving some quantitative measure of network performance vs ρ.

1. Optimal flows through the disordered lattice. (Preprint).

Order-of-magnitude calculation on $N \times N$ grid. Send volume ρ_N between each (source,destination) pair. Average flow volume $\bar{f}$ across edges is

$$(N^2 \times N^2) \times \rho_N \times N \approx \bar{f} \times N^2$$

To make $\bar{f}$ be order 1 we take

$$\rho_N = \rho N^{-3}$$

Open Problem. Take i.i.d. capacities (cap(e)) with $0 < c_- \leq \text{cap}(e) \leq c_+ < \infty$. Obvious: a feasible flow with normalized demand ρ exists for $\rho < \rho_-$ and doesn't exist for $\rho > \rho_+$. Prove there is a constant ρ_* depending on distribution of cap(e) such that as $N \to \infty$

$$\begin{aligned} P(\exists \text{ feasible flow, norm. demand } \rho) &\to 1 \quad , \ \rho < \rho_* \\ &\to 0 \quad , \ \rho > \rho_*. \end{aligned}$$

Instead of focussing on capacities, let's focus on congestion. In a network without congestion, the cost (to system; all users combined) of a flow of volume $f(e)$ scales linearly with $f(e)$. With congestion, extra users impose extra costs on other users as well as on themselves. So cost scales super-linearly with $f(e)$.

Model: The cost of a flow $\mathbf{f} = (f(e))$ in an environment $\mathbf{c} = (c(e))$ is

$$\text{cost}_{(N)}(\mathbf{f}, \mathbf{c}) = \sum_e c(e) f^2(e).$$

Theorem 1. $N \times N$ torus (for simplicity)
Large constant bound B on edge-capacity (for simplicity)
i.i.d. cost-factors $c(e)$ with

$$0 < c_- \leq c(e) \leq c^+ < \infty.$$

Let Γ_N be minimum cost of flow with normalized intensity $\rho = 1$. Then

$$N^{-2} E\Gamma_N \to \text{constant}(B, \text{dist}(c(e))).$$

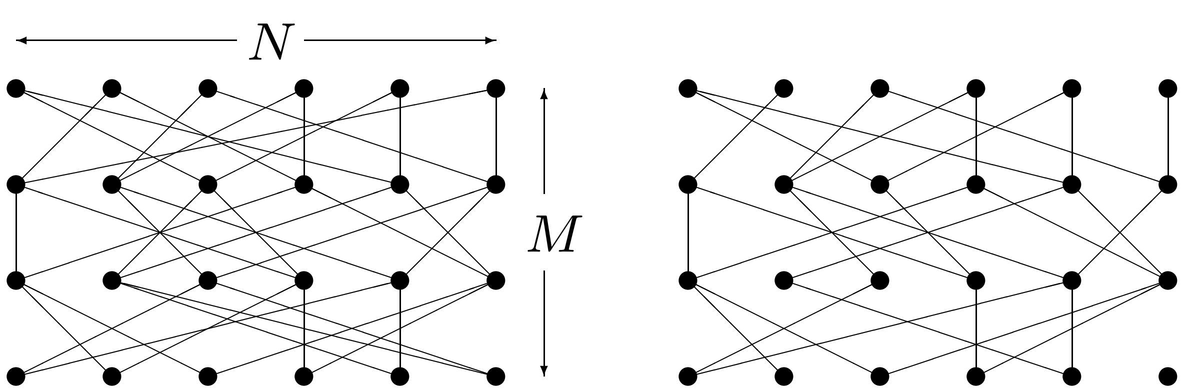
Comments. Methodology is to compare with flows across (boundary-to-boundary) $M \times M$ squares. Should work to prove existence of limits in other "optimal flows on $N \times N$ grid" models. But details are surprisingly hard to prove.

2. Cost-volume relationships for flows through a disordered network. (Preprint).

Consider a network with

- M layers
- N vertices per layer
- directed edges upwards from one layer to next
- edges between successive layers are placed randomly subject to each vertex having in-degree = out-degree = 2.

Within this model we'll consider a "special" and a "general" problem.



$p = 1$ $p = 0.8$

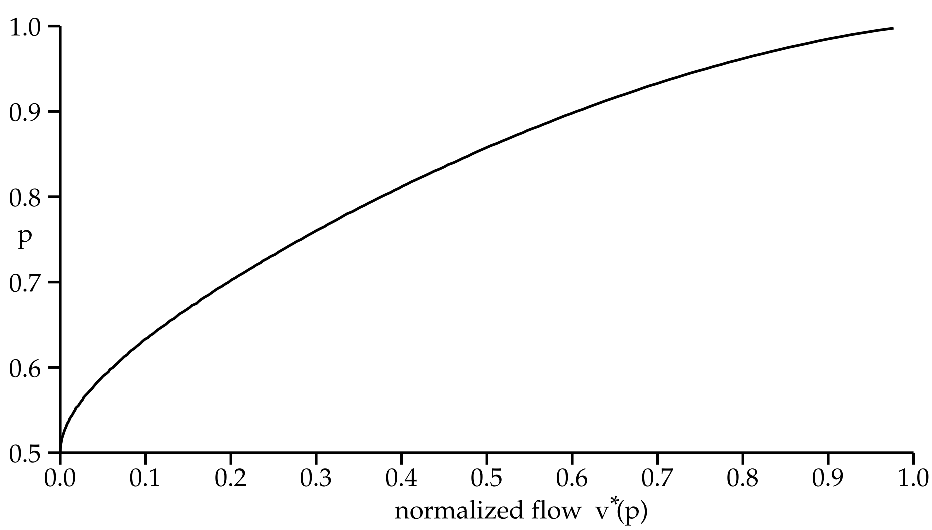
Special problem. Suppose
• edges have capacity = 1.
• retain each edge with probability p, delete with probability $1-p$.
Study maximum flow from bottom to top layers; same as maximum number of edge-disjoint paths from bottom to top layers. Clearly for $p = 1$ the maximum flow $= 2N$, so for general p we consider the relative flow

$$F_{N,M}(p) = \frac{1}{2N} \times \text{(max flow through network)}.$$

We anticipate a limit function

$$EF_{N,N}(p) \to v^*(p) \text{ as } n \to \infty.$$

Cavity method tells you how to write down an equation whose solution determines $v^*(p)$.



Cavity method from statistical physics provides a heuristic for obtaining solutions of various combinatorial optimization problems over random networks which are **locally tree-like**. This work is first explicit application to flow problems.

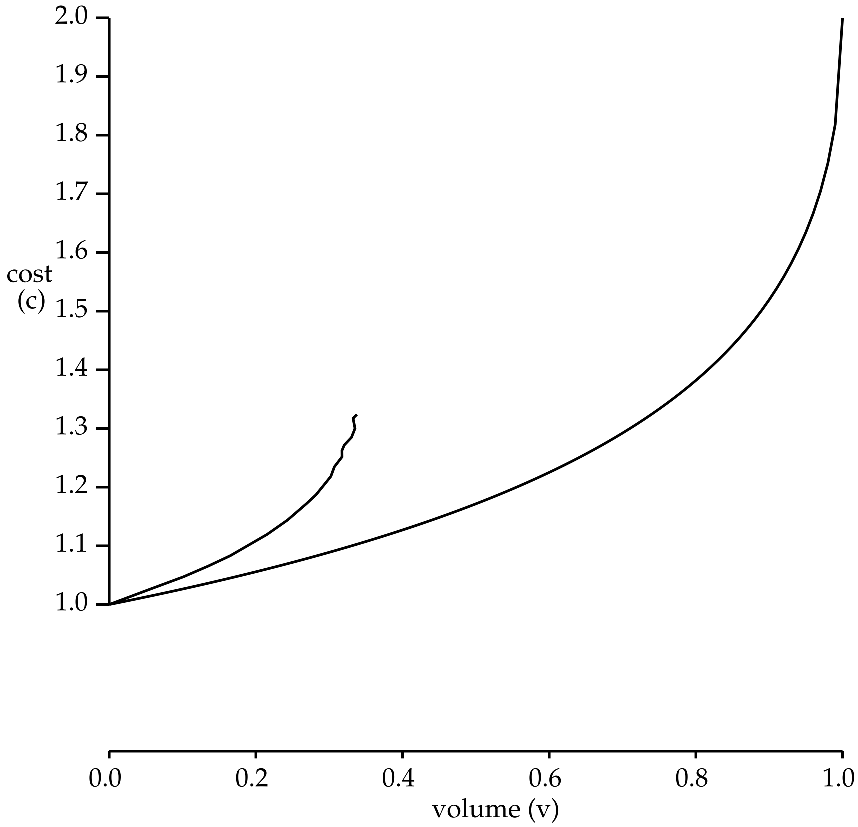
General problem. Same underlying random graph model: in-degree = out-degree = 2.

- On each edge there is a cost-volume function:

$\phi(v) =$ cost-per-unit flow when flow volume $= v$.

- The functions ϕ are i.i.d. over edges.

The cavity method lets us calculate (via numerical solution of an equation) the network cost-volume function $\psi(\cdot) =$ normalized total cost of flow when normalized total volume $= v$.



Here we take a particular form (long curve) for cost-volume function on an edge. This arises from a road-traffic model in which speed is decreasing linear function of density, cost = 1/(speed).

Make maximum volume be i.i.d. Exponential (1) over edges. Short curve shows the network cost-volume function, with maximum volume (congestion) around 0.34.

Models **3** and **4** are based on

The mean-field model of distance

Take complete graph on n vertices. Let each of the $\binom{n}{2}$ edges (i,j) have random length, independently, with Exponential (mean n) distribution. This model has several names:

- *Complete graph with random edge weights*
- *random link model*
- *stochastic mean-field model of distance.*

Within this model one can study classical combinatorial optimization problems such as TSP and MST. The length L_n of optimal solutions will scale as n.

Here is a systematic way to study many problems within the mean-field model. From a typical vertex, the distances

$$0 < \xi_{n,1} < \xi_{n,2} < \ldots < \xi_{n,n-1}$$

to other vertices, in increasing order, have a $n \to \infty$ limit in distribution

$$0 < \xi_1 < \xi_2 < \xi_3 < \ldots$$

which is the Poisson process of rate 1 on $(0, \infty)$.

In a certain sense (local weak convergence), the model has a $n \to \infty$ limit which we call the **PWIT** (Poisson weighted infinite tree).

3. Edge-flow distribution uder shortest-path routing. (Aldous - Bhamidi in progress).

In mean-field model of distance, easy to see that distance $D(i,j)$ between specified vertices i,j satisfies

$$D(i,j) = \log n \pm O(1) \text{ in prob.}$$

Send flow of volume $1/n$ between each pair (i,j) along shortest path. Each edge e gets some total flow $F_n(e)$. What is the distribution of edge-flows $(F_n(e) : e \text{ an edge})$?

Call edges of length $O(1)$ "short". Easy to see intuitively that short edges should get flow of order $\log n$.

Theorem 1 *As $n \to \infty$ for fixed $z > 0$,*

$$\frac{1}{n}\#\{e : F_n(e) > z \log n\} \to_{L^1}$$

$$G(z) := \int_0^\infty P(W_1 W_2 e^{-u} > z)\ du$$

where W_1 and W_2 are independent Exponential(1). In particular

$$\frac{1}{n}E\#\{e : F_n(e) > z \log n\} \to G(z).$$

Proof is intricate "bare-hands" calculations, exploiting i.i.d. Exponential edge-lengths.

Here is a heuristic argument for why the limit is this particular function $G(z)$.

Background fact: the process

$N(t)$ = number of vertices within distance t of a specified vertex

is (exactly) the Yule process in the PWIT, and (approximately) the Yule process in the finite-n model.

Consider a short edge e, and suppose there are $W'(\tau)$ vertices within a fixed large distance τ of one end of the edge, and $W''(\tau)$ vertices within distance τ of the other end. A shortest-length path between distant vertices which passes through e must enter and exit the region above via some pair of vertices in the sets above, and there are $W'(\tau)W''(\tau)$ such pairs. The dependence on the length L is more subtle. By the Yule process approximation, the number of vertices within distance r of an initial vertex grows as e^r, and it turns out that the flow through e depends on L as $\exp(-L)$ because of the availability of alternate possible shortest paths. So flow through e should be proportional to $W'(\tau)W''(\tau)\exp(-L)$. But (again by the Yule process approximation) for large τ the r.v. $e^{-\tau}W'(\tau)$ has approximately the Exponential(1) distribution W_1. And as $n \to \infty$ the normalized distribution $n^{-1}\#\{e : L_e \in \cdot\}$ of all edge-lengths converges to the σ-finite distribution of U_∞. This is heuristically how the limit distribution $W_1 W_2 \exp(-U_\infty)$ arises.

4. "Price of anarchy" in mean-field model of distance. (back-of-envelope, last week).

In previous model, suppose each edge e has an owner who sets a price-per-unit-volume $\pi(e)$ for using edge e. So from a customer's viewpoint the cost of using edge e is

$$\text{length}(e) + \pi(e)$$

and customers choose minimum-cost routes. The owners adjust prices to maximize their income

$$\pi(e) \times (\text{volume of flow across } e).$$

Expect equilibrium prices.

Recall in previous setting (no prices) the mean cost of routing (uniform source - destination) is $(1 + o(1)) \log n$. In the current setting we heuristically have a striking result in $n \to \infty$ limit

- for each edge e' we have $\pi(e') \to e = 2.718\ldots$

- mean cost $= (e + o(1)) \log n$.

Key idea: Difference between cost of minimum-cost route and second-minimum-cost route has limit distribution which is robust (up to scaling constants) to imposing random prices.