Weak Concentration for First Passage Percolation Times on Graphs and General Increasing Set-valued Processes

David Aldous

2 September 2015

Recall the **method of bounded differences;** for a RV Z of the form

$$Z = f(\xi_1, \dots \xi_n); \quad \text{for independent } (\xi_i)$$

where f has the property

$$|f(\mathbf{x}) - f(\mathbf{x}')| \leq 1 \text{ whenever } \mathbf{x},\ \mathbf{x}' \text{ differ in only one coordinate}$$

we have

$$\mathbb{P}(|Z - \mathbb{E}Z| \geq \lambda n^{1/2}) \leq 2\exp(-\lambda^2/2).$$

Basic example of a general **concentration inequality –** a key point is that one can bound on a difference $|Z - \mathbb{E}Z|$ even when you don't know $\mathbb{E}Z$.

I will present a simple but very specialized concentration inequality, with $2 + 2\varepsilon$ applications.

Consider a finite-state continuous-time Markov chain and a hitting time $T = T_A$ on some subset A of states, and suppose

$$h(i) := \mathbb{E}_i T < \infty \ \forall i.$$

Neither "finite" nor "continuous" is actually important here. What is important is the next assumption:

$$h(j) \leq h(i) \text{ for each possible transition } i \to j. \tag{1}$$

This is a very restrictive assumption – not obvious that **any** "interesting" chain satisfies this assumption.

Lemma

Under condition (1), for any initial state,

$$\frac{\mathrm{var}\ T}{\mathbb{E} T} \leq \max\{h(i) - h(j) : \ i \to j \textit{ a possible transition}\}.$$

We can rewrite this as

$$\kappa := \max\{h(i) - h(j) : \ i \to j \text{ a possible transition}\}$$

$$\frac{\mathrm{s.d.}(T)}{\mathbb{E}T} \le \sqrt{\frac{\kappa}{\mathbb{E}T}}$$

where the right side is always ≤ 1. So we get a weak concentration inequality if $\kappa/\mathbb{E}T$ is small.

The proof is easy if you think martingales. We have a chain $(X_t, 0 \le t < \infty)$ with transition rates $q(i,j)$. There is a martingale

$$M_t := \mathbb{E}(T|X_s, s \le t) = h(X_{t \wedge T}) + t \wedge T \quad (2)$$

and M_t^2 has a Doob-Meyer decomposition of into a martingale Q_t and a predictable process. We can write expressions for dM_t and for the predictable process, then applying the optional sampling theorem leads to the general formulas

$$\mathbb{E} T = \mathbb{E} \int_0^T b(X_t)\ dt, \quad \mathrm{var}\ T = \mathbb{E} \int_0^T a(X_t)\ dt$$

$$a(i) := \sum_j q(i,j)\ (h(i) - h(j))^2$$

$$b(i) := \sum_j q(i,j)(h(i) - h(j)).$$

Setting

$$\kappa = \max\{h(i) - h(j) : \ i \to j \text{ a possible transition}\}$$

we clearly have $a(i) \le \kappa b(i)$; note this is where we use the monotonicity hypothesis (1).

Our applications are all in the context of chains (Z_t) whose values are subsets S of a given discrete space and whose transitions are of the form $S \to S \cup \{v\}$. In words "increasing set-valued processes".

And our applications use hitting times T of the form

$$T := \inf\{t : Z_t \supset B \text{ for some } B \in \mathcal{B}\}$$

for a specified collection $\mathcal{B}$ of subsets B.

The rest of the talk is example which fit this context, ordered as

uninteresting/interesting/uninteresting/interesting.

Example: a general growth process (Z_t) on the lattice $\mathbb{Z}^2$.
The states are finite vertex-sets S, the possible transitions are $S \to S \cup \{v\}$ where v is a vertex adjacent to S. For each such transition, we assume the transition rates are bounded above and below:

$$0 < c_* \leq q(S, S \cup \{v\}) \leq c^* < \infty. \qquad (3)$$

Initially $Z_0 = \{\mathbf{0}\}$, where $\mathbf{0}$ denotes the origin. The "monotonicity" condition we impose is that these rates are increasing in S:

$$\text{if } v, v' \text{ are adjacent to } S \text{ then } q(S, S \cup \{v\}) \leq q(S \cup \{v'\}, S \cup \{v, v'\}) . \qquad (4)$$

Note that we do not assume any kind of spatial homogeneity.

Proposition

Let B be an arbitrary subset of vertices $\mathbb{Z}^2 \setminus \{\mathbf{0}\}$, and consider $T := \inf\{t : Z_t \cap B \text{ is non-empty.}\}$. Under assumptions (3, 4),

$$\mathrm{var}\ T \leq \mathbb{E} T / c_*.$$

Proof. Condition (4) allows us to couple versions (Z'_t, Z''_t) of the process starting from states $S' \subset S''$, such that in the coupled process we have $Z'_t \subseteq Z''_t$ for all $t \geq 0$. In particular, $h(S) := \mathbb{E}_S T$ satisfies the monotonicity condition (1). To deduce the result from Lemma 1 we need to show that, for any given possible transition $S_0 \to S_0 \cup \{v_0\}$, we have

$$h(S_0) \leq h(S_0 \cup \{v_0\}) + 1/c_*. \qquad (5)$$

Now by running the process started at S_0 until the first time T^* this process contain v_0, and then coupling the future of that process to the process started at $S_0 \cup \{v_0\}$, we have $h(S_0) \leq \mathbb{E}_{S_0} T^* + h(S_0 \cup \{v_0\})$. And $\mathbb{E}_{S_0} T^* \leq 1/c_*$ by (3), establishing (5).

Two recent recent arXiv posts deal with specific models of such non-homogeneous growth processes.

Asymptotic behavior of the Eden model with positively homogeneous edge weights by Bubeck and Gwynne

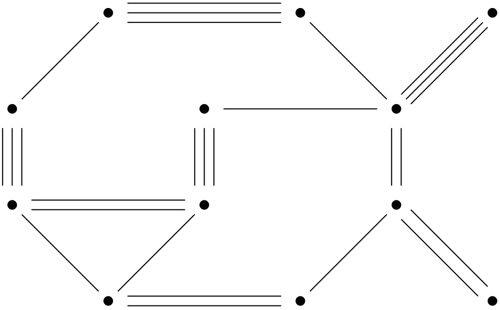
Nucleation and growth in two dimensions by Bollobas et al.

Example: A multigraph process

Take a finite connected graph $(\mathbf{V}, \mathbf{E})$ with edge-weights $\mathbf{w} = (w_e)$, where $w_e > 0 \ \forall e \in \mathbf{E}$.

Define a multigraph-valued process as follows. Initially we have the vertex-set $\mathbf{V}$ and no edges. For each vertex-pair $e = (vy) \in \mathbf{E}$, edges vy appear at the times of a Poisson (rate w_e) process, independent over $e \in \mathbf{E}$.

So at time t the state of the process, Z_t say, is a multigraph with $N_e(t) \geq 0$ copies of edge e, where $(N_e(t), \ e \in \mathbf{E})$ are independent Poisson(tw_e) random variables.



Background motivation: model of imperfectly observed network. How well can we estimate some aspect of the true network from observed multigraph?

Leads to a huge range of questions; one basic issue is to say something about connectivity properties of true network. This must relate to connectivity properties of the observed multigraph.

We study how long until Z_t has various connectivity properties. Specifically, consider

- $T'_k = \inf\{t : Z_t \text{ is } k\text{-edge-connected}\}$
- $T_k = \inf\{t : Z_t \text{ contains } k \text{ edge-disjoint spanning trees.}\}$

Here we regard the $N_e(t)$ copies of e as disjoint edges. Remarkably, Lemma 1 enables us to give a simple proof of a "weak concentration" bound which does not depend on the underlying weighted graph.

Proposition

$$\frac{\mathrm{s.d.}(T_k)}{\mathbb{E}\, T_k} \leq \frac{1}{\sqrt{k}}, \ \ k \geq 1.$$

Via a continuization device, the same bound holds in the discrete-time model where edges e arrive IID with probabilities proportional to w_e.

We conjecture that some similar result holds for T'_k. But proving this by our methods would require some structure theory (beyond Menger's theorem) for k-edge-connected graphs, and it is not clear whether relevant theory is known.

Proof. Here the states S are multigraphs over $\mathbf{V}$, and $h(S)$ is the expectation, starting at S, of the time until the process contains k edge-disjoint spanning trees. Monotonicity property is clear. What are the possible values of $h(S) - h(S \cup \{e\})$, where $S \cup \{e\}$ denotes the result of adding an extra copy of e to the multigraph S?

Consider the "min-cut" over proper subsets $S \subset \mathbf{V}$

$$\gamma := \min_S w(S, S^c)$$

where $w(S, S^c) = \sum_{v \in S, y \in S^c} w_{vy}$. Because a spanning tree must have at least one edge across the min-cut,

$$\mathbb{E} T_k \geq k/\gamma. \tag{6}$$

On the other hand we claim

$$h(S) - h(S \cup \{e\}) \leq 1/\gamma.$$

Given this, Lemma 1 establishes the proposition.

$$\text{Claim :} \qquad h(S) - h(S \cup \{e\}) \leq 1/\gamma.$$

To prove this, take the natural coupling (Z_t, Z_t^+) of the processes started from S and from $S \cup \{e\}$, and run the coupled process until Z_t^+ contains k edge-disjoint spanning trees. At this time, the process Z_t either contains k edge-disjoint spanning trees, or else contains $k-1$ spanning trees plus two trees (regard as edge-sets $\mathbf{t}_1$ and $\mathbf{t}_2$) such that $\mathbf{t}_1 \cup \mathbf{t}_2 \cup \{e\}$ is a spanning tree. So the extra time we need to run (Z_t) is at most the time until some arriving edge links $\mathbf{t}_1$ and $\mathbf{t}_2$, which has mean at most $1/\gamma$. This establishes the Claim.

Example: Coverage processes. The topic of *coverage processes* is centered upon spatial or combinatorial variants of the coupon collector's problem. Classical theory concerns low-parameter models for which the cover time T_n of a "size n" model can be shown to have a limit distribution after scaling: $(T_n - a_n)/b_n \to_d \xi$ for explicit a_n, b_n. In many settings, Lemma 1 can be used to give a weak concentration result for models with much less regular structure. For instance, a one-line proof gives

Proposition

Let G be an arbitrary n-vertex graph, and let $(V_i, i \geq 1)$ be IID uniform random vertices. Let T be the smallest t such that every vertex is contained in, or adjacent to, the set $\{V_i, 1 \leq i \leq t\}$. Then $\mathrm{var}\ T \leq n\, \mathbb{E} T$.

For a sequence (G_n) of sparse graphs, $\mathbb{E} T_n$ will be of order $n \log n$, so the bound says that $\mathrm{s.d.}(T_n)/\mathbb{E} T = O(1/\sqrt{\log n})$.

Main example: the general FPP model. Start with a finite connected graph $(\mathbf{V}, \mathbf{E})$ with edge-weights $\mathbf{w} = (w_e)$, where $w_e > 0 \ \forall e \in \mathbf{E}$. To the edges $e \in \mathbf{E}$ attach independent Exponential(rate w_e) random variables ξ_e. For each pair of vertices (v', v'') there is a random variable $X(v', v'')$ which can be viewed in two equivalent ways:

- viewing ξ_e as the length of edge e, then $X(v', v'')$ is the length of the shortest route from v to v';
- viewing ξ_e as the time to traverse edge e, then $X(v', v'')$ is the "first passage percolation" time from v to v'.

Taking the latter view, we call this the *FPP model*, and call $X(v', v'')$ the *FPP time* and ξ_e the *traversal time*. This type of model and many generalizations have been studied extensively in several settings, in particular

- FPP with general IID weights on $\mathbb{Z}^d$
- FPP on classical random graph (Erdős-Rényi or configuration) models
- and a much broader "epidemics and rumors on complex networks" literature.

However, this literature invariably starts by assuming some specific graph model; we do not know any "general results" which relate properties of the FPP model to properties of a general underlying graph. As an analogy, another structure that can be associated with a weighted finite graph is a finite reversible Markov chain; the established theory surrounding mixing times of Markov chains does contain "general results" relating properties of the chain to properties of the underlying graph.

Here we study the "weak concentration" property: when it is true that $X(v', v'')$ is close to its expectation? We can reformulate the FPP model as a set-valued process and then Lemma 1 immediately implies the following result, classical on $\mathbb{Z}^d$.

Proposition

$\mathrm{var}\ X(v', v'') \le \mathbb{E}X(v', v'')/w_*$ *for* $w_* := \min\{w_e : \ e \in \mathbf{E}\}$.

The Proposition implies that on any unweighted graph ($w_e = 1$ for all edges e), the spread of $X = X(v', v'')$ is at most order $\sqrt{\mathbb{E}X}$. For many specific graphs, stronger concentration results are known.

- For $\mathbb{Z}^2$ there is extensive literature on the longstanding conjecture that the spread is order $(\mathbb{E}X)^{1/3}$.
- For the complete graph and for sparse random graphs on n vertices the spread of $X/\mathbb{E}X$ is typically of order $1/\log n$.

In contrast, we study the completely general case where the edge-weights w_e may vary widely over the different edges $e \in \mathbf{E}$. Here is the main result (conjectured several years ago). Given a pair (v', v''), there is a random path $\boldsymbol{\pi}(v', v'')$ that attains the FPP time $X(v', v'')$. Define $\Xi(v', v'') := \max\{\xi_e : \ e \in \boldsymbol{\pi}(v', v'')\}$ as the maximum edge-traversal time in this minimal path. Recall the "L^0 norm"

$$||V||_0 := \inf\{\delta : \mathbb{P}(|V| > \delta) \leq \delta\}.$$

Theorem

There exist functions ψ^+ and $\psi_- : (0,1] \to (0,\infty)$ such that $\psi^+(\delta) \downarrow 0$ as $\delta \downarrow 0$, and $\psi_-(\delta) > 0$ for all $\delta > 0$, and such that, for all finite connected edge-weighted graphs and all vertex pairs (v', v''),

$$\psi_- \left(\left|\left| \frac{\Xi(v', v'')}{\mathbb{E}X(v', v'')} \right|\right|_0 \right) \leq \frac{\mathrm{s.d.}(X(v', v''))}{\mathbb{E}X(v', v'')} \leq \psi^+ \left(\left|\left| \frac{\Xi(v', v'')}{\mathbb{E}X(v', v'')} \right|\right|_0 \right).$$

In words, $X/\mathbb{E}X$ has small spread if and only if $\Xi/\mathbb{E}X$ is small. Intuition for this result comes from the "almost disconnected" case where the path $\boldsymbol{\pi}(v', v'')$ must contain a specific "bridge" edge e with small w_e; if $1/w_e$ is not $o(\mathbb{E}X)$ then the contribution to X from the traversal time ξ_e is enough to show that X cannot have weakly concentrated distribution.

This FPP model fits our "increasing set-valued process" framework by considering Z_t = the set of vertices reached by time t. This requires our Exponential assumption.

The issue is the upper bound. Details of proof are rather intricate (8 pages).

Outline of proof. Step 1 is to make an "approximate" version of our Lemma 1; this shows it is enough to prove that in **most** transitions $S \to S \cup \{y\}$ the decrements $h(S) - h(S \cup \{y\})$ are $o(\mathbb{E}T)$.

So suppose not, that is suppose that for some transitions $S \to S \cup \{y\}$ the decrements $h(S) - h(S \cup \{y\})$ are $\Omega(\mathbb{E}T)$.

(rest of proof is special to this example).

So suppose not, that is suppose that for some transitions $S \to S \cup \{y\}$ the decrements $h(S) - h(S \cup \{y\})$ are $\Omega(\mathbb{E}T)$.

Step 2 shows that in some such transitions, the used edge vy will (with non-vanishing probability) have traversal time ξ_{vy} also of order $\Omega(\mathbb{E}T)$.

(This is not obvious because, although the weight w_{vy} on the used edge must have $1/w_{vy} = \Omega(\mathbb{E}T)$, the actual realization of the r.v. ξ_{vy} might perhaps be much smaller.)

Step 3 shows that some such edges will (with non-vanishing probability) be in the minimal path.

(This is not obvious because the set-valued process (Z_t) may have many more transitions than are in the minimal path from v' to v'').