Bootstrap percolation on $G(n,p)$

Svante Janson
(joint work with Tomasz Łuczak, Tatyana Turova, Thomas Vallier)

Bootstrap percolation on a graph G is defined as the spread of *activation* or *infection* according to the following rule, with a given threshold $r \geq 2$: We start with a set $\mathcal{A}(0) \subseteq V(G)$ of *active* vertices. Each inactive vertex that has at least r active neigbours becomes active. This is repeated until no more vertices become active, i.e., when no inactive vertex has r or more active neigbours.

We are mainly interested in the final size A^* of the active set, and in particular whether eventually all vertices will be active or not. If they are, we say that the initial set $\mathcal{A}(0)$ *percolates*. We will study a sequence of graphs of order $n \to \infty$; we then also say that (a sequence of) $\mathcal{A}(0)$ *almost percolates* if the number of vertices that remain inactive is $o(n)$, i.e., if $A^* = n - o(n)$.

Bootstrap percolation has been studied on various graphs, both deterministic and random; one can study either a random initial set or the deterministic problem of choosing an initial set that is optimal in some sense. For example, a classical folklore problem is to find the minmal percolating set in a two-dimensional grid; see Balogh and Pete [3] and Bollobas [5]. (These references also treat higher-dimensional grids.) Some further references for random initial sets on various graphs are Cerf and Manzo [6], Holroyd [7] (grids); Balogh and Bollobás [1] (hypercube); Balogh, Peres and Pete [2] (infinite trees); Balogh and Pittel [4] (random regular graphs).

We here study bootstrap percolation on the Erdös-Rényi random graph $G_{n,p}$ (which somewhat surprisingly seems to have been neglected so far in this context), with an initial set $\mathcal{A}(0)$ consisting of a given number a vertices chosen at random. This was first studied by Vallier [8]; we here present a simple method that allows us to both simplify the proofs and improve the results.

In order to analyze the bootstrap percolation process on $G_{n,p}$, we change the time scale and consider at each time step the spread of activation from one vertex only. Choose $u_1 \in \mathcal{A}(0)$ and give each of its neighbours a *mark*; we then say that u_1 is *used*, and let $\mathcal{Z}(1) := \{u_1\}$ be the set of used vertices at time 1. We continue recursively: At time $t+1$, choose a vertex $u_{t+1} \in \mathcal{A}(t) \setminus \mathcal{Z}(t)$ (provided this set is non-empty). We give each neighbour of u_{t+1} a new mark. Let $\Delta\mathcal{A}(t+1)$ be the set of inactive vertices with r marks; these now become active and we let $\mathcal{A}(t+1) = \mathcal{A}(t) \cup \Delta\mathcal{A}(t+1)$ be the set of active vertices at time t. We finally set $\mathcal{Z}(t+1) = \mathcal{Z}(t) \cup \{u_{t+1}\} = \{u_i : i \leq t+1\}$, the set of used vertices.

The process stops when $\mathcal{A}(t) \setminus \mathcal{Z}(t) = \emptyset$, i.e., when all active vertices are used. We denote this time by T;

$$T := \min\{t \geq 0 : \mathcal{A}(t) \setminus \mathcal{Z}(t) = \emptyset\}. \tag{1}$$

Thus the final infected set is $\mathcal{A}(T) = \mathcal{Z}(T)$, and its size is

$$A^* := |\mathcal{A}(T)| = |\mathcal{Z}(T)| = T. \tag{2}$$

Hence, the set $\mathcal{A}(0)$ percolates if and only if $T = n$, and $\mathcal{A}(0)$ almost percolates if and only if $T = n - o(n)$.

Since $|\mathcal{Z}(t)| = t$ and $\mathcal{Z}(t) \subseteq \mathcal{A}(t)$ for $t = 0, \dots, T$, we also have, with $A(t) := |\mathcal{A}(t)|$, the number of active vertices at time t,

$$T = \min\{t \geq 0 : A(t) = t\}. \tag{3}$$

We analyze this process by the standard method of revealing the edges of the graph $G_{n,p}$ only on a need-to-know basis. We thus begin by choosing u_1 as above and then reveal its neighbours; we then find u_2 and reveal its neighbours, and so on. Let, for $i \notin \mathcal{Z}(s)$, $I_i(s)$ be the indicator function that there is an edge between the vertices u_s and i. This is also the indicator that i gets a mark at time s, so if $M_i(t)$ is the number of marks i has at time t, then

$$M_i(t) = \sum_{s=1}^{t} I_i(s), \tag{4}$$

at least until i is activated (and what happens later does not matter). Note that if $i \notin \mathcal{A}(0)$, then, for every $t \leq T$, $i \in \mathcal{A}(t)$ if and only if $M_i(t) \geq r$.

The crucial feature of this description of the process, which makes the analysis simple, is that the random variables $I_i(s)$ are i.i.d. $\mathrm{Be}(p)$. We have defined $I_i(s)$ only for $s \leq T$ and $i \notin \mathcal{Z}(s)$, but it is convenient to add further (redundant) variables so that $I_i(s)$ are defined, and i.i.d., for all $i \in V_n$ and all $s \geq 1$.

Define, for $i \in V_n \setminus \mathcal{A}(0)$,

$$Y_i := \min\bigl\{t : M_i(t) \geq r\bigr\}. \tag{5}$$

If $Y_i \leq T$, then Y_i is the time vertex i becomes active, but if $Y_i > T$, then Y_i never becomes active. Thus, for $t \leq T$,

$$\mathcal{A}(t) = \mathcal{A}(0) \cup \{i \notin \mathcal{A}(0) : Y_i \leq t\}.$$

By (4) and (5), each Y_i has a negative binomial distribution $\mathrm{NegBin}(r, p)$;

$$\mathbb{P}(Y_i = k) = \mathbb{P}\bigl(M_i(k-1) = r-1,\, I_i(k) = 1\bigr) = \binom{k-1}{r-1} p^k (1-p)^{r-k};$$

moreover, these random variables Y_i are i.i.d.

We let, for $t = 0, 1, 2, \dots$,

$$S(t) := |\{i \notin \mathcal{A}(0) : Y_i \leq t\}|,$$

so

$$A(t) = S(t) + A(0) = S(t) + a. \tag{6}$$

By (3), (2) and (6), it suffices to study the stochastic process $S(t)$. Note that $S(t)$ is a sum of $n - a$ i.i.d. processes $\mathbf{1}[t \geq Y_i]$, each of which is 0/1-valued and jumps from 0 to 1 at time Y_i. The fact that $S(t)$, and thus $A(t)$, is a sum of i.i.d. processes makes the analysis easy; in particular, for any given t,

$$S(t) \sim \mathrm{Bin}\bigl(n - a, \mathbb{P}(Y_1 \leq t)\bigr).$$

We have, for any given t_0,

$$T \geq t_0 \iff \min_{t<t_0}(A(t)-t) > 0 \iff a+\min_{t<t_0}(S(t)-t) > 0 \iff a > -\min_{t<t_0}(S(t)-t).$$

(Note that this is exact; so far no approximation has been done.)

To find the threshold for (almost) percolation, we thus only have to find the minimum $\min_{t<t_0}(S(t)-t)$ for $t_0 = n$ or t_0 close to n. Standard concentration results show that $S(t) \approx \mathbb{E}\, S(t)$, where

$$\mathbb{E}\, S(t) = (n-a)\,\mathbb{P}(Y_1 \leq t) = (n-a)\,\mathbb{P}(M_1(t) \geq r),$$

and explicit results are easily found.

For notational simplicity we state the result for $r = 2$ only. In this case, $\mathbb{E}\, S(t) - t$ has a minimum $1/(2np^2)$ at $t = 1/(np^2)$ (asymptotically), and we obtain the following result.

Theorem 1. *Let $r = 2$, and assume $n^{-1} \ll p = p(n) \ll n^{-1/2}$. Then the threshold for (almost) percolation is*

$$a_* := \frac{1}{2np^2}.$$

More precisely, for any fixed $\delta > 0$,

(i) *If $|\mathcal{A}(0)| \leq (1-\epsilon)a_*$, then whp $A^* \leq 2|\mathcal{A}(0)|$.*

(ii) *If $|\mathcal{A}(0)| \geq (1+\epsilon)a_*$, then whp $A^* = n - o(n)$. If further $np \geq \log n + \log\log n + \omega(n)$ for some $\omega(n) \to \infty$, then whp $A^* = n$, so $\mathcal{A}(0)$ percolates completely.*

Moreover, $S(t) - \mathbb{E}\, S(t)$ converges after normalization to a Gaussian process, and it is easy to refine the results above and obtain very precise information on the width of the critical window (which is of the order $\sqrt{a_*}$); we also obtain a Gaussian limit law for the final size A^* in the subcritical case.

Details will appear.

References

[1] J. Balogh and B. Bollobás, Bootstrap percolation on the hypercube. *Probab. Theory Related Fields* **134** (2006), no. 4, 624–648.

[2] J. Balogh, Y. Peres and G. Pete, Bootstrap percolation on infinite trees and non-amenable groups. *Combin. Probab. Comput.* **15** (2006), no. 5, 715–730.

[3] J. Balogh and G. Pete, Random disease on the square grid. *Random Structures Algorithms* **13** (1998), no. 3-4, 409–422. 60C05

[4] J. Balogh and B. G. Pittel, Bootstrap percolation on the random regular graph. *Random Structures Algorithms* **30** (2007), no. 1-2, 257–286.

[5] B. Bollobás, *The Art of Mathematics. Coffee Time in Memphis.* Cambridge University Press, New York, 2006.

[6] R. Cerf and F. Manzo, The threshold regime of finite volume bootstrap percolation. *Stochastic Process. Appl.* **101** (2002), no. 1, 69–82.

[7] A. E. Holroyd, Sharp metastability threshold for two-dimensional bootstrap percolation. *Probab. Theory Related Fields* **125** (2003), no. 2, 195–224.

[8] T. Vallier, *Random graph models and their applications.* Ph. D. thesis, Lund University, 2007.