Nikolaos Fountoulakis[1] Konstantinos Panagiotou[1]

3-connected cores of random 2-connected graphs

[1] Max-Planck-Institut für Informatik, Germany

Laboratoire d'Informatique de Paris-Nord, 16 November 2010

The structure of random graphs

Erdős and Rényi proved that around $p_{cr} = \frac{1}{n}$ a change in the structure of $G(n, p)$ occurs:

if $p < \frac{1-\varepsilon}{n}$, then all components of $G_{n,p}$ contain $O(\log n)$ vertices;
if $p > \frac{1+\varepsilon}{n}$, then there exists a unique component with $\Theta(n)$ vertices, whereas every other component has $O(\log n)$ vertices.

The structure of random graphs

Erdős and Rényi proved that around $p_{cr} = \frac{1}{n}$ a change in the structure of $G(n, p)$ occurs:

if $p < \frac{1-\varepsilon}{n}$, then all components of $G_{n,p}$ contain $O(\log n)$ vertices;

if $p > \frac{1+\varepsilon}{n}$, then there exists a unique component with $\Theta(n)$ vertices, whereas every other component has $O(\log n)$ vertices.

The structure of random graphs

Erdős and Rényi proved that around $p_{cr} = \frac{1}{n}$ a change in the structure of $G(n,p)$ occurs:

if $p < \frac{1-\varepsilon}{n}$, then all components of $G_{n,p}$ contain $O(\log n)$ vertices;
if $p > \frac{1+\varepsilon}{n}$, then there exists a unique component with $\Theta(n)$ vertices, whereas every other component has $O(\log n)$ vertices.

Random planar graphs

Let $\mathcal{P}_n$ be the set of labeled all planar graphs on n vertices and let P_n denote a graph taken at random from $\mathcal{P}_n$ with probability $1/|\mathcal{P}_n|$.

How many edges does P_n typically have?

Giménez and Noy (2009) showed that

$$\frac{e(P_n) - \kappa n}{\sqrt{2\pi\lambda n}} \xrightarrow{d} N(0,1),$$

where $\kappa \approx 2.21326$ and $\lambda \approx 0.43034$ are constants.

Random planar graphs

Let $\mathcal{P}_n$ be the set of labeled all planar graphs on n vertices and let P_n denote a graph taken at random from $\mathcal{P}_n$ with probability $1/|\mathcal{P}_n|$.

How many edges does P_n typically have?

Giménez and Noy (2009) showed that

$$\frac{e(P_n) - \kappa n}{\sqrt{2\pi\lambda n}} \xrightarrow{d} N(0,1),$$

where $\kappa \approx 2.21326$ and $\lambda \approx 0.43034$ are constants.

Random planar graphs

Let $\mathcal{P}_n$ be the set of labeled all planar graphs on n vertices and let P_n denote a graph taken at random from $\mathcal{P}_n$ with probability $1/|\mathcal{P}_n|$.

How many edges does P_n typically have?

Giménez and Noy (2009) showed that

$$\frac{e(P_n) - \kappa n}{\sqrt{2\pi\lambda n}} \xrightarrow{d} N(0,1),$$

where $\kappa \approx 2.21326$ and $\lambda \approx 0.43034$ are constants.

Random planar graphs

What is the typical component structure of P_n?

Giménez and Noy (2009) proved that the number of connected components of P_n is asymptotically distributed as

$$1 + X,$$

where $X \overset{\mathcal{L}}{=} \mathrm{Po}(\nu)$, and $\nu \approx 0.037439$.

Typical structure of a random planar graph

There is a unique giant component containing $n - O_C(1)$ vertices, whereas the remaining components are “small".

Random planar graphs

What is the typical component structure of P_n?

Giménez and Noy (2009) proved that the number of connected components of P_n is asymptotically distributed as

$$1 + X,$$

where $X \stackrel{\mathcal{L}}{=} \texttt{Po}(\nu)$, and $\nu \approx 0.037439$.

Typical structure of a random planar graph

There is a unique giant component containing $n - O_C(1)$ vertices, whereas the remaining components are "small".

Random planar graphs

What is the typical component structure of P_n?

Giménez and Noy (2009) proved that the number of connected components of P_n is asymptotically distributed as

$$1 + X,$$

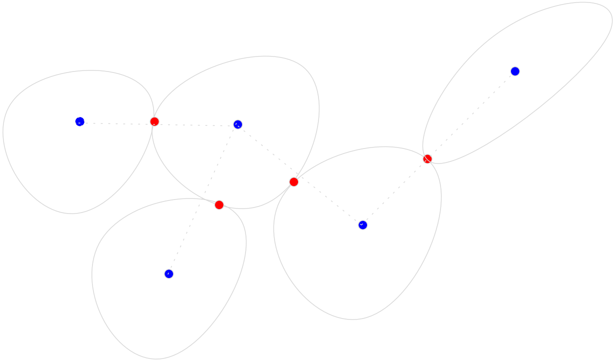
where $X \stackrel{\mathcal{L}}{=} \mathrm{Po}(\nu)$, and $\nu \approx 0.037439$.

Typical structure of a random planar graph

There is a unique giant component containing $n - O_C(1)$ vertices, whereas the remaining components are "small".

The structure of random connected graphs

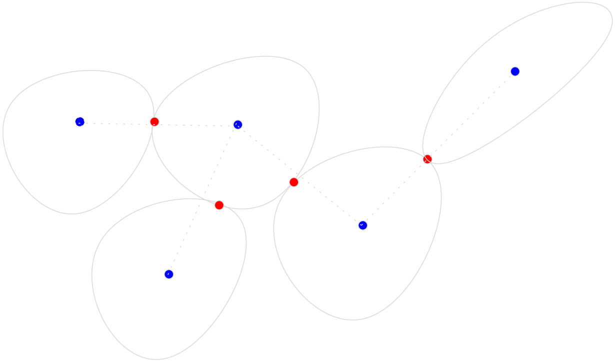
A connected graph consists of its *biconnected components*:
the maximal subgraphs of *connectivity at least 2*.



A connected graph

The structure of random connected graphs

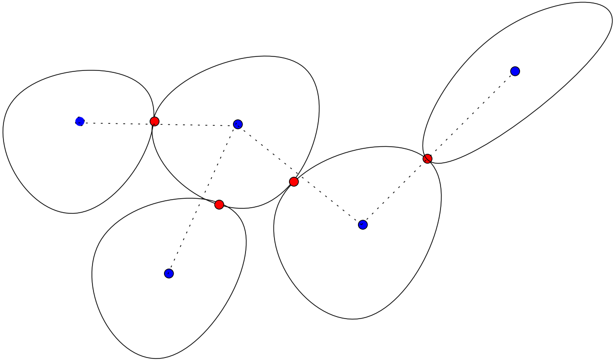
A connected graph consists of its *biconnected components*: the maximal subgraphs of *connectivity at least 2*.



A connected graph

The structure of random connected graphs

A connected graph consists of its *biconnected components*: the maximal subgraphs of *connectivity at least 2*.



A connected graph

The structure of random connected graphs

Assume that we sample a connected graph uniformly from the family of graphs whose biconnected components belong to the family $\mathcal{B}$.

Let $B(x)$ be the enumerating generating function of $\mathcal{B}$:

$$B(x) = \sum_{n=1}^{\infty} \frac{|\mathcal{B}_n|}{n!} x^n,$$

and let $\rho_{\mathcal{B}}$ be its radius of convergence.

The structure of random connected graphs

Assume that we sample a connected graph uniformly from the family of graphs whose biconnected components belong to the family $\mathcal{B}$.

Let $B(x)$ be the enumerating generating function of $\mathcal{B}$:

$$B(x) = \sum_{n=1}^{\infty} \frac{|\mathcal{B}_n|}{n!} x^n,$$

and let $\rho_{\mathcal{B}}$ be its radius of convergence.

The structure of random connected graphs

Assume that we sample a connected graph uniformly from the family of graphs whose biconnected components belong to the family $\mathcal{B}$.

Let $B(x)$ be the enumerating generating function of $\mathcal{B}$:

$$B(x) = \sum_{n=1}^{\infty} \frac{|\mathcal{B}_n|}{n!} x^n,$$

and let $\rho_{\mathcal{B}}$ be its radius of convergence.

Random connected graphs

It turns out that what determines the structure of a random connected graph sampled from this family is

$$\rho_{\mathcal{B}} B''(\rho_{\mathcal{B}})$$

Random connected graphs

It turns out that what determines the structure of a random connected graph sampled from this family is

$$\rho_{\mathcal{B}} B''(\rho_{\mathcal{B}})$$

Random connected graphs

Theorem [Panagiotou and Steger (2009)]

Let C_n be a random graph sampled uniformly from the family of connected graphs on n vertices with biconnected components in $\mathcal{B}$. With probability $1 - o(1)$

- if $\rho_{\mathcal{B}} B''(\rho_{\mathcal{B}}) > 1$, then all biconnected components have size $O(\log n)$;
- if $\rho_{\mathcal{B}} B''(\rho_{\mathcal{B}}) < 1$, then there exists a unique biconnected component of order $\Theta(n)$, but every other component has $o(n)$ vertices.

Random biconnected graphs

Question

If C_n is a random graph on n vertices sampled from a certain class of biconnected graphs, what is the typical distribution of its 3-connected buildings blocks?

What is a building block of a 2-connected graph?

Networks (Trakhtenbrot–Tutte)

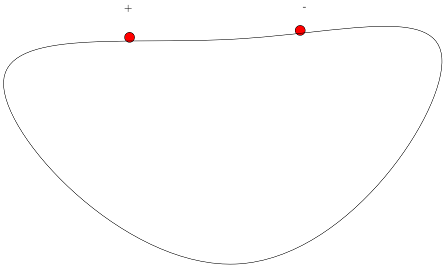
Definition - Networks

A *network* is a graph with two distinguished vertices which we call *poles*, so that if we add an edge between them, then the resulting (multi)graph belongs to the certain class of biconnected graphs.

Networks (Trakhtenbrot–Tutte)

Definition - Networks

A *network* is a graph with two distinguished vertices which we call *poles*, so that if we add an edge between them, then the resulting (multi)graph belongs to the certain class of biconnected graphs.

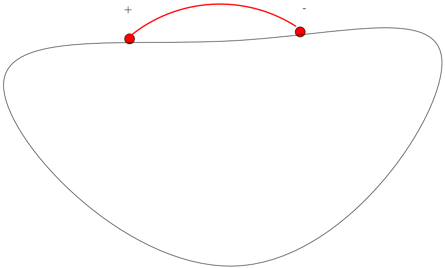


A network

Networks (Trakhtenbrot–Tutte)

Definition - Networks

A *network* is a graph with two distinguished vertices which we call *poles*, so that if we add an edge between them, then the resulting (multi)graph belongs to the certain class of biconnected graphs.



A network

Network decomposition

A network is:

- an edge;
- a *series* network (type S);
- a *parallel* network (type P);
- a *core* network (type H).

Network decomposition

A network is:

- an edge;
- a *series* network (type S);
- a *parallel* network (type P);
- a *core* network (type H).

Network decomposition

A network is:

- an edge;
- a *series* network (type S);
- a *parallel* network (type P);
- a *core* network (type H).

Network decomposition

A network is:

- an edge;
- a *series* network (type S);
- a *parallel* network (type P);
- a *core* network (type H).

Network decomposition

A network is:

- an edge;
- a *series* network (type S);
- a *parallel* network (type P);
- a *core* network (type H).

Series networks

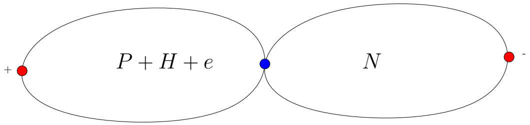
A *series* network is:

Series networks

A *series* network is:

Series networks

A *series* network is:



A *series* network

Parallel networks

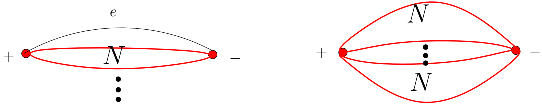
A *parallel* network is:

Parallel networks

A *parallel* network is:

Parallel networks

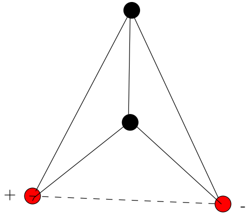
A *parallel* network is:



The two types of *parallel networks*

Core networks

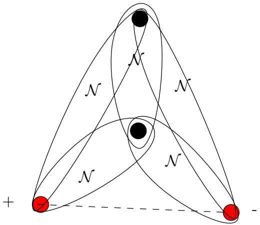
A *core* network is:



The underlying 3-connected graph is called a *core*.

Core networks

A *core* network is:



The underlying 3-connected graph is called a *core*.

Generating functions

Let

- $N(x, y)$ be the enumerating generating function of the class of networks;
- $T(x, y)$ be the e.g.f. of the class of 3-connected graphs from which we choose the cores.

Generating functions

Let

- $N(x, y)$ be the enumerating generating function of the class of networks;
- $T(x, y)$ be the e.g.f. of the class of 3-connected graphs from which we choose the cores.

Generating functions

These functions satisfy:

$$\Phi(x, y, N(x, y)) = 0,$$

where

$$\Phi(x, y, z) = T(x, z) - \log\left(\frac{1+z}{1+y}\right) + \frac{xz^2}{1+xz}.$$

Random networks

Aim

We study random networks on n vertices, where the cores are sampled from a given class of 3-connected graphs.

Random networks

A certain correspondence between networks and the resulting class of biconnected graphs yields the following:

Rough equivalence

If a property holds a.a.s. for such a class of networks, then it also holds a.a.s. for the corresponding class of biconnected graphs.

Random networks

What determines the typical structure of a random network on n vertices is

$$\text{the sign of } \Phi_z(\rho_N(1), 1, N(\rho_N(1), 1)),$$

where $\rho_N(1)$ is the radius of convergence of $N(x, 1)$.

Random networks

Theorem [F. and Panagiotou]

Let N_n be a random network on n vertices. If $\Phi_z(\rho_N(1), 1, N(\rho_N(1), 1)) > 0$, then

all cores of N_n have $O_C(\log n)$ vertices.

If $\Phi_z(\rho_N(1), 1, N(\rho_N(1), 1)) < 0$, then for some $\gamma_C > 0$

there is a unique core with $\gamma_C n + o_p(n)$ vertices, but every other core has $o_p(n)$ vertices.

Random networks

We have calculated

- asymptotic counts for the number of small cores;
- the order of the "giant" core;
- asymptotic distribution for the number of edges of N_n as $n \to \infty$.

Random networks

We have calculated

- asymptotic counts for the number of small cores;
- the order of the “giant" core;
- asymptotic distribution for the number of edges of N_n as $n \to \infty$.

Random networks

We have calculated

- asymptotic counts for the number of small cores;
- the order of the "giant" core;
- asymptotic distribution for the number of edges of N_n as $n \to \infty$.

Random networks

We have calculated

- asymptotic counts for the number of small cores;
- the order of the "giant" core;
- asymptotic distribution for the number of edges of N_n as $n \to \infty$.

Example - Biconnected Random Planar Graphs

If we sample the cores from the class of 3-connected planar graphs, the resulting network corresponds to random biconnected planar graphs;

- It turns out that this class of networks falls into the second "category":

Theorem

A random biconnected planar graph on n vertices has a unique core of order $c_p n + o(n)$, where $c_p = 0.765...$, whereas every other core has $O(n^{2/3})$ vertices, with probability $1 - o(1)$.

This was also shown recently by Giménez, Noy and Rué, with the use of analytic methods.

Example - Biconnected Random Planar Graphs

If we sample the cores from the class of 3-connected planar graphs, the resulting network corresponds to random biconnected planar graphs;

- It turns out that this class of networks falls into the second "category":

Theorem

A random biconnected planar graph on n vertices has a unique core of order $c_p n + o(n)$, where $c_p = 0.765...$, whereas every other core has $O(n^{2/3})$ vertices, with probability $1 - o(1)$.

This was also shown recently by Giménez, Noy and Rué, with the use of analytic methods.

Example - Biconnected Random Planar Graphs

For every $4 \le \ell = O\left(\left(\frac{n}{\log n}\right)^{2/5}\right)$ the number of cores with ℓ vertices is for any $\varepsilon > 0$

$$c_\ell n(1 \pm \varepsilon),$$

with probability $1 - o(1)$, where c_ℓ is determined by the generating function of the class of 3-connected planar graphs.

Proof techniques

- We analyse the output of *Boltzmann samplers* which are randomised algorithms that generate networks;
- In our case, these are a collection of randomized algorithms that call each other recursively, reflecting the recursive construction of a network.

Proof techniques

- We analyse the output of *Boltzmann samplers* which are randomised algorithms that generate networks;
- In our case, these are a collection of randomized algorithms that call each other recursively, reflecting the recursive construction of a network.

Proof techniques

Example: Boltzmann sampler for Networks

$$\Gamma N(x,y): \quad \begin{array}{ll} \Gamma N \leftarrow e & w.p.\ \frac{y}{N(x,y)}; \\ \Gamma N \leftarrow \Gamma S(x,y) & w.p.\ \frac{S(x,y)}{N(x,y)}; \\ \Gamma N \leftarrow \Gamma P(x,y) & w.p.\ \frac{P(x,y)}{N(x,y)}; \\ \Gamma N \leftarrow \Gamma H(x,y) & w.p.\ \frac{H(x,y)}{N(x,y)}; \end{array}$$

Proof techniques

Example: Boltzmann sampler for Core Networks

$$
\begin{aligned}
\Gamma H(x,y): \quad & T \leftarrow \Gamma T(x, N(x,y)) \\
& \text{for each edge } e \text{ of } T \\
& \quad \gamma_e \leftarrow \Gamma N(x,y) \\
& \text{replace every } e \text{ in } T \text{ by } \gamma_e \\
& \text{Return } T\text{, relabeling randomly its vertices.}
\end{aligned}
$$

Proof techniques

- We are able to show the concentration of the number of calls of each routine;
- Let
  - $A_{\mathtt{Net}}$ be the number of calls of the network routine;
  - $A_{\mathtt{Ser}}$ be the number of calls of the series networks routine;
  - $A_{\mathtt{Par}}$ be the number of calls of the parallel networks routine;
  - V_T and E_T denote the total number of vertices and edges in cores.

Proof techniques

- We are able to show the concentration of the number of calls of each routine;
- Let
  - $A_{\mathtt{Net}}$ be the number of calls of the network routine;
  - $A_{\mathtt{Ser}}$ be the number of calls of the series networks routine;
  - $A_{\mathtt{Par}}$ be the number of calls of the parallel networks routine;
  - $V_{\mathcal{T}}$ and $E_{\mathcal{T}}$ denote the total number of vertices and edges in cores.

Proof techniques

- We are able to show the concentration of the number of calls of each routine;
- Let
  - $A_{\texttt{Net}}$ be the number of calls of the network routine;
  - $A_{\texttt{Ser}}$ be the number of calls of the series networks routine;
  - $A_{\texttt{Par}}$ be the number of calls of the parallel networks routine;
  - $V_{\mathcal{T}}$ and $E_{\mathcal{T}}$ denote the total number of vertices and edges in cores.

Proof techniques

- We are able to show the concentration of the number of calls of each routine;
- Let
  - $A_{\texttt{Net}}$ be the number of calls of the network routine;
  - $A_{\texttt{Ser}}$ be the number of calls of the series networks routine;
  - $A_{\texttt{Par}}$ be the number of calls of the parallel networks routine;
  - $V_{\mathcal{T}}$ and $E_{\mathcal{T}}$ denote the total number of vertices and edges in cores.

Proof techniques

- We are able to show the concentration of the number of calls of each routine;
- Let
  - $A_{\texttt{Net}}$ be the number of calls of the network routine;
  - $A_{\texttt{Ser}}$ be the number of calls of the series networks routine;
  - $A_{\texttt{Par}}$ be the number of calls of the parallel networks routine;
  - $V_{\mathcal{T}}$ and $E_{\mathcal{T}}$ denote the total number of vertices and edges in cores.

Proof techniques

- We are able to show the concentration of the number of calls of each routine;
- Let
  - $A_{\texttt{Net}}$ be the number of calls of the network routine;
  - $A_{\texttt{Ser}}$ be the number of calls of the series networks routine;
  - $A_{\texttt{Par}}$ be the number of calls of the parallel networks routine;
  - $V_{\mathcal{T}}$ and $E_{\mathcal{T}}$ denote the total number of vertices and edges in cores.

Proof Techniques

We show that for each $Z \in \{A_{\mathtt{Net}}, A_{\mathtt{Ser}}, A_{\mathtt{Par}}, V_{\mathcal{T}}, E_{\mathcal{T}}\}$ we have

$$\mathbb{P}\left(|Z - zn| < \varepsilon n\right) > 1 - e^{-C\varepsilon^2 n},$$

where $z \in \{a_{\mathtt{Net}}, a_{\mathtt{Ser}}, a_{\mathtt{Par}}, v_{\mathcal{T}}, e_{\mathcal{T}}\}$ and the vector $a := [a_{\mathtt{Net}}, a_{\mathtt{Ser}}, a_{\mathtt{Par}}, v_{\mathcal{T}}, e_{\mathcal{T}}]^T$ is the solution of the system $Ma = r$, where

$$M = \begin{bmatrix} \frac{1}{N(x,y)} & \frac{\rho_N N(x,y)}{S(x,y)} & \frac{N(x,y)-1}{2P(x,y)} & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ \frac{S(x,y)}{N(x,y)} & -1 & \frac{S(x,y)N(x,y)}{P(x,y)} & 0 & 0 \\ \frac{P(x,y)}{N(x,y)} & \frac{\rho_N P(x,y)N(x,y)}{S(x,y)} & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 \end{bmatrix}, r = \begin{bmatrix} \mu \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{1}$$

and $\mu = -\frac{\rho_N'(1)}{\rho_N(1)}$.

- In particular, if A_H is the number of calls of the *Core Networks* routine, then

$$A_H = \alpha_H n + o_p(n),$$

where

$$\alpha_H = 2\mu H(x, y).$$

- In particular, if A_H is the number of calls of the *Core Networks* routine, then

$$A_H = \alpha_H n + o_p(n),$$

where

$$\alpha_H = 2\mu H(x, y).$$

Proof techniques

- We treat $\Gamma H(x, y)$ as a deterministic algorithm that reads its inputs from a list
$$(T_1, T_2, \ldots),$$
where the $\{T_i\}_{i \geq 1}$ are independent samples from the class of cores, distributed according to the Boltzmann distribution.
- If $C_k(n)$ is the number of cores of size k in a network with n vertices, we are able to bound it by looking inside
$$(T_1, \ldots, T_{\lceil \alpha_H n + \varepsilon n \rceil})$$
and use Chernoff bounds.

Proof techniques

- We treat $\Gamma H(x, y)$ as a deterministic algorithm that reads its inputs from a list
$$(T_1, T_2, \ldots),$$
where the $\{T_i\}_{i \geq 1}$ are independent samples from the class of cores, distributed according to the Boltzmann distribution.
- If $C_k(n)$ is the number of cores of size k in a network with n vertices, we are able to bound it by looking inside
$$(T_1, \ldots, T_{\lceil \alpha_H n + \varepsilon n \rceil})$$
and use Chernoff bounds.

Outlook

- Use these results for algorithmic applications;
- determine the asymptotic distribution of the order and the size of the giant core;
- higher connectivities???

Outlook

- Use these results for algorithmic applications;
- determine the asymptotic distribution of the order and the size of the giant core;
- higher connectivities???

Outlook

- Use these results for algorithmic applications;
- determine the asymptotic distribution of the order and the size of the giant core;
- higher connectivities???

Thank you!