THRESHOLD FUNCTIONS FOR SUBGRAPHS OF GIVEN TYPE OF THE BICHROMATIC RANDOM GRAPH

by

ILONA PALÁSTI

Dedicated to Professor T. Gallai
at his 50th birthday.

P. ERDŐS and A. RÉNYI in their paper [1] consider the evolution of the random graph $\Gamma_{n,N}$ having n labelled vertices (or point) and N edges which are chosen at random in such a way that each possible choice has the same probability. If A is some property which the random graph may or may not posses let $\mathbf{P}_{n,N}(A)$ denote the probability of the random graph possessing the property A. They treated the "typical" structures arising at a given stage of the evolution (i.e. if N is a given function of n). (A typical structure means a structure the probability $\mathbf{P}_{n,N}(A)$ of which tends to 1 for $n \to \infty$.) If A is a property such that

$$\lim_{n\to\infty} \mathbf{P}_{n,N}(A) = 1\,, \tag{1}$$

then we say that "almost all" graphs have this property which can be formed from n given labelled points and N edges. If there exists a function $A(n)$ tending monotonically to $+\infty$ for $n \to +\infty$, such that

$$\lim_{n\to+\infty} \mathbf{P}_{n,N(n)}(A) = \begin{cases} 0, & \text{if} \quad \lim\limits_{n\to+\infty} \dfrac{N(n)}{A(n)} = 0 \\ 1, & \text{if} \quad \lim\limits_{n\to+\infty} \dfrac{N(n)}{A(n)} = +\infty, \end{cases} \tag{2}$$

then this function $A(n)$ is called a threshold function for the property A.

A graph is called a balanced graph, if it has no subgraph of a degree larger than its own. (If the graph G has n given points and N edges, then the number $\frac{2N}{n}$ is called the degree of the graph.) Two graphs are called isomorphic, if there exists a one-to-one mapping of the vertices carrying over these graphs into another.

A graph is bichromatic if it has m given labelled points $P_1, P_2, \ldots, P_m$ of the first colour and n given labelled points $Q_1, Q_2, \ldots, Q_n$ of the second colour and N edges connecting points having different colours.

A bichromatic graph is called a bichromatic complete graph of order (k, l) if it has k vertices of the first colour, l vertices of the second one and kl edges. A connected bichromatic graph which has k points of the first colour and l points of the second one, and $k+l-1$ edges, is called a bichromatic (k, l)-tree. It can be seen easily that the complete graphs and trees are balanced graphs.

The aim of our paper is to determine the threshold function for a certain type of graphs, which are subgraphs of the bichromatic random graph.

Let the bichromatic graph $\Gamma_{m,n,N}$ consist of m given labelled vertices $P_1, P_2, \ldots, P_m$ of one colour, n given labelled points $Q_1, Q_2, \ldots, Q_n$ of the other colour and N edges connecting vertices of different colours only, chosen at random in such a way that all admissible choices have the same probability.

Increasing the number of the edges N so that it remains very small compared with m and n, for example if $N = o\left(\sqrt{\frac{mn}{m+n}}\right)$ it is very probable that $\Gamma_{m,n,N}$ consists of isolated points and isolated edges. Especially if $m \sim cn$, then in the case of $N = o(\sqrt{n})$ it will be very probable, that the random graph consisting of isolated points and isolated edges only. Namely the probability that at least two edges of $\Gamma_{m,n,N}$ have a common point equals to 1 minus the probability that there are no edges having a common point. As the probability of the choice of any admissible set of N edges is the same, namely $1\Big/\binom{mn}{N}$, thus the probability that the edges have no common point is

$$\frac{N!\binom{m}{N}\binom{n}{N}}{\binom{mn}{N}}. \tag{3}$$

Accordingly the probability that at least two edges of $\Gamma_{m,n,N}$ will have a point in common is

$$1-\frac{N!\binom{m}{N}\binom{n}{N}}{\binom{mn}{N}}. \tag{4}$$

Using the relation

$$\binom{n}{k} \sim \frac{n^k}{k!} e^{-\frac{k^2}{2n}} \tag{5}$$

valid for $k = o(n^{2/3})$, we obtain that for $N = o\left(\sqrt{\frac{mn}{m+n}}\right)$

$$1-\frac{N!\binom{m}{N}\binom{n}{N}}{\binom{mn}{N}} = O\left[N^2\left(\frac{1}{m}+\frac{1}{n}-\frac{1}{mn}\right)\right] = o(1). \tag{6}$$

If however in the case $m \sim cn$, $N \sim c_1\sqrt{n}$, where $c_1 > 0$, is a constant not depending on n then the appearance of trees consisting of three points, that is (1,2)- or (2,1)-trees has a probability with a positive limit for $n \to \infty$, but the appearance of a connected subgraph consisting of more than 3 vertices remains still very improbable. If N increases while n is fixed, the situation will change only if N reaches the order of magnitude of $n^{2/3}$. Then the (1,3)-, (2,2)- and (3,1)-trees having four points appear. Generally (in the case $m \sim cn$) the threshold function for the appearance of the trees of order (k, l) is $n^{\frac{k+l-2}{k+l-1}}$. This result is contained in the following

Theorem. *Let $k \geq 1, l \geq 1$, and v be positive integers ($k + l - 1 \leq v \leq kl$). Let $\mathcal{B}_{k,l,v}$ denote any non empty class of connected balanced bichromatic random graphs which contain k points of the first colour, l points of the second one and v edges (connecting points having different colours). The threshold function concerning the property of the bichromatic random graph, that it contains a subgraph isomorphic with some element of $\mathcal{B}_{k,l,v}$ is equal to $n^{2-\frac{k+l}{v}}$, supposing that $m \sim cn$, where $c > 0$ and does not depend on n.*

Proof. The proof of the theorem is similar to that of the analogous theorem of Erdős and Rényi concerning the threshold for subgraphs of given type of one coloured random graphs. Let $B_{k,l,v}$ denote the number of all graphs which can be formed from k labelled points of the first colour and l labelled points of the second colour and belonging to the class $\mathcal{B}_{k,l,v}$. If $\mathbf{P}_{m,n,N}(\mathcal{B}_{k,l,v})$ denotes the probability that $\Gamma_{m,n,N}$ has at least one such subgraph which is isomorphic to some element of $\mathcal{B}_{k,l,v}$ then obviously

$$\mathbf{P}_{m,n,N}(\mathcal{B}_{k,l,v}) \leq \binom{m}{k}\binom{n}{l} B_{k,l,v} \frac{\binom{mn-v}{N-v}}{\binom{mn}{N}}. \tag{7}$$

As a matter of fact at first we choose k points of one colour (this can be done in $\binom{m}{k}$ ways) and l points of the other colour (which can be done in $\binom{n}{l}$ different ways); from these we form all graphs which are isomorphic to some element of the class $\mathcal{B}_{k,l,v}$ (this may be done in $B_{k,l,v}$ different ways); then the number of graphs $G_{m,n,N}$ which contain the selected graph as a subgraph is equal to the number of ways in which the remaining $N - v$ edges can be chosen from the $mn - v$ possible other edges. (In this way are repeatedly taking into consideration those graphs which have more than one subgraph isomorphic with some element of the class $\mathcal{B}_{k,l,v}$.)

From (7) we obtain

$$\mathbf{P}_{m,n,N}(\mathcal{B}_{k,l,v}) = O\left(\frac{N^v}{m^{v-k}\, n^{v-l}}\right). \tag{8}$$

From (8) and the supposition $m \sim cn$, if $N = o\left(n^{2-\frac{k+l}{v}}\right)$ it is clear that

$$\mathbf{P}_{m,n,N}(B_{k,l,v}) = o(1) \tag{9}$$

and thus we proved the first assertion of the theorem.

In proving the second part of the theorem let us denote by $\mathscr{B}_{k,l,v}^{(m,n)}$ the set of all subgraphs of the bichromatic complete graph which contains m points of the first colour and n points of the second colour, which are isomorphic to some element of $\mathscr{B}_{k,l,v}$. If $S \in \mathscr{B}_{k,l,v}^{(m,n)}$, let the random variable $\varepsilon(S)$ be equal to 1 if S is a subgraph of $\Gamma_{m,n,N}$ and $\varepsilon(S) = 0$ otherwise. Then the mean value of the number of subgraphs of $\Gamma_{m,n,N}$ which are isomorphic to an element of $\mathscr{B}_{k,l,v}$, is

$$\mathbf{M}\Big(\sum_{S \in \mathscr{B}_{k,l,v}^{(m,n)}} \varepsilon(S)\Big) = \sum_{S \in \mathscr{B}_{k,l,v}^{(m,n)}} \mathbf{M}(\varepsilon(S)) = \tag{10}$$

$$= \binom{m}{k}\binom{n}{l} B_{k,l,v} \frac{\binom{mn-v}{N-v}}{\binom{mn}{N}} \sim \frac{B_{k,l,v}}{k!\,l!}\;\frac{N^v}{m^{v-k}\,n^{v-l}}\,.$$

If S_1 and S_2 are two elements of $\mathscr{B}_{k,l,v}^{(m,n)}$ which have no common edge then the mean value of the product $\varepsilon(S_1)\,\varepsilon(S_2)$ is

$$\mathbf{M}\big(\varepsilon(S_1)\,\varepsilon(S_2)\big) = \frac{\binom{mn-2v}{N-2v}}{\binom{mn}{N}}\,. \tag{11}$$

If however S_1 and S_2 have r common edges ($1 \leqq r \leqq v-1$), then

$$\mathbf{M}\big(\varepsilon(S_1)\,\varepsilon(S_2)\big) = \frac{\binom{mn-2v+r}{N-2v+r}}{\binom{mn}{N}} = O\left(\frac{N^{2v-r}}{(mn)^{2v-r}}\right). \tag{12}$$

On the other hand if i and j denote the number of common points of S_1 and S_2 of the first and second colour respectively, and $i+j=s$, then — as the intersection of S_1 and S_2 is a subgraph of S_1 (and S_2) and as by our supposition each S is balanced — we obtain $\frac{r}{s} \leqq \frac{v}{k+l}$ that is $s \geqq \frac{r(k+l)}{v}$; thus the number of such pairs of subgraphs S_1 and S_2 does not exceed the sum

$$B_{k,l,v}^2 \sum_{i+j \geqq \frac{r(k+l)}{v}} \binom{m}{k}\binom{k}{i}\binom{m-k}{k-i}\binom{n}{l}\binom{l}{j}\binom{n-l}{l-j}. \tag{13}$$

By making use of the supposition $m \sim cn$ we obtain

$$B_{k,l,v}^2 \sum_{i+j \geqq \frac{r(k+l)}{v}} \binom{m}{k}\binom{k}{i}\binom{m-k}{k-i}\binom{n}{l}\binom{l}{j}\binom{n-l}{l-j} = O\left(n^{2(k+l)-\frac{r(k+l)}{v}}\right). \tag{14}$$

Consequently we obtain

$$\mathbf{M}\Big(\big(\sum_{S\in\mathcal{B}_{k,l,v}^{(m,n)}}\varepsilon(S)\big)^2\Big)=\sum_{S\in\mathcal{B}_{k,l,v}^{(m,n)}}\mathbf{M}(\varepsilon(S))+$$

(15)
$$+\frac{B_{k,l,v}^2\, m!\, n!}{k!^2\, l!^2(m-2k)!\,(n-2l)!}\,\frac{\binom{mn-2v}{N-2v}}{\binom{mn}{N}}+O\left[\left(n^{\frac{N^v}{2v-(k+l)}}\right)^2\frac{1}{n}\right]+$$

$$+O\left[\left(n^{\frac{N^v}{2v-(k+l)}}\right)^2\sum_{r=1}^{v}\left(\frac{n^{2-\frac{k+l}{v}}}{N}\right)^r\right].$$

The meaning of the sum of the second and the third term of (15) is the following: the mean value of the number of such pairs of subgraphs which have no common edges but have $i+j$ common points does not exceed the following

(16)
$$B_{k,l,v}^2\,\frac{\binom{mn-2v}{N-2v}}{\binom{mn}{N}}\sum_{i+j=0}^{k+l-1}\binom{m}{k}\binom{k}{i}\binom{m-k}{k-i}\binom{n}{l}\binom{l}{j}\binom{n-l}{l-j}.$$

Let us divide the sum (16) into two parts

(17)
$$B_{k,l,v}^2\binom{m}{k}\binom{m-k}{k}\binom{n}{l}\binom{n-l}{l}\frac{\binom{mn-2v}{N-2v}}{\binom{m\,n}{N}}+$$

$$+B_{k,l,v}^2\,\frac{\binom{mn-2v}{N-2v}}{\binom{mn}{N}}\sum_{i+j=1}^{k+l-1}\binom{m}{k}\binom{k}{i}\binom{m-k}{k-i}\binom{n}{l}\binom{l}{j}\binom{n-l}{l-j}.$$

Wherefrom we obtain

(18)
$$\frac{B_{k,l,v}^2\, m!\, n!}{k!^2\, l!^2(m-2k)!\,(n-2l)!}\,\frac{\binom{mn-2v}{N-2v}}{\binom{mn}{N}}+O\left[\left(\frac{N^v}{n^{2v-(k+l)}}\right)^2\sum_{i+j=1}^{k+l-1}\frac{1}{n^{i+j}}\right].$$

For the second term of (15) it is easy to see that

(19)
$$\frac{m!\, n!}{k!^2\, l!^2(m-2k)!\,(n-2l)!}\,\frac{\binom{mn-2v}{N-2v}}{\binom{mn}{N}}\le\binom{m}{k}^2\binom{n}{l}^2\frac{\binom{mn-v}{N-v}^2}{\binom{mn}{N}^2}.$$

And if we suppose that

$$\frac{N}{n^{2-\frac{k+l}{v}}} = \omega \to +\infty, \quad n \to \infty. \tag{20}$$

it follows using the relation $\mathbf{D}^2(\xi) = \mathbf{M}(\xi^2) - \mathbf{M}^2(\xi)$ where $\mathbf{D}^2(\xi)$ denotes the variance of ξ that

$$\mathbf{D}^2\Big(\sum_{S\in\mathscr{B}_{k,l,v}^{(m,n)}} \varepsilon(S)\Big) = O\left(\frac{\Big(\sum_{S\in\mathscr{B}_{k,l,v}^{(m,n)}} \mathbf{M}(\varepsilon(S))\Big)^2}{\min(\omega, n)}\right). \tag{21}$$

On the other hand from the inequality of Chebyshev we obtain

$$\mathbf{P}_{m,n,N}\left(\Big|\sum_{S\in\mathscr{B}_{k,l,v}^{(m,n)}} \varepsilon(S) - \sum_{S\in\mathscr{B}_{k,l,v}^{(m,n)}} \mathbf{M}(\varepsilon(S))\Big| > \frac{1}{2}\sum_{S\in\mathscr{B}_{k,l,v}^{(m,n)}} \mathbf{M}(\varepsilon(S))\right) = O\left(\frac{1}{\min(\omega, n)}\right) \tag{22}$$

and thus

$$\mathbf{P}_{m,n,N}\left(\sum_{S\in\mathscr{B}_{k,l,v}^{(m,n)}} \varepsilon(S) \leqq \frac{1}{2}\sum_{S\in\mathscr{B}_{k,l,v}^{(m,n)}} \mathbf{M}(\varepsilon(S))\right) = O\left(\frac{1}{\min(\omega, n)}\right). \tag{23}$$

It is clear from (10) that if $\omega \to \infty$, then

$$\sum_{S\in\mathscr{B}_{k,l,v}^{(m,n)}} \mathbf{M}(\varepsilon(S)) \to +\infty. \tag{24}$$

Thus it follows not only that the probability of the graph $\Gamma_{m,n,N}$ containing at least one subgraph isomorphic to some element of $\mathscr{B}_{k,l,v}$ tends to 1, but also that the number of subgraphs of $\Gamma_{m,n,N}$ which are isomorphic with some element of $\mathscr{B}_{k,l,v}$ tends to $+\infty$ in probability.

Thus the theorem is proved.

If we put $v = k + l - 1$ into the threshold function given by the theorem we obtain

Corollary 1. *The threshold function for the property that the bichromatic random graph contains a (k, l)-tree is $n^{\frac{k+l-2}{k+l-1}}$ supposing that $m \sim cn$.*

This agrees with our result in [2].

And if we write kl instead of v we obtain

Corollary 2. *The threshold function for the property that the bichromatic random graph contains a bichromatic complete subgraph of order (k, l) is $n^{2-\frac{k+l}{kl}}$.*

It is evident that the latter threshold function is equal to the threshold for the property that a random graph (which is not coloured) contains a saturated even subgraph (i.e. a subgraph consisting of $k + l$ points $P_1, P_2, \ldots, P_k$ and $Q_1, Q_2, \ldots, Q_l$ and containing all edges $P_i Q_j$) given by P. Erdős and A. Rényi in [1].

I am indebted to Professor Rényi for his valuable remarks.

(Received April 16, 1962.)

REFERENCES

[1] ERDŐS, P.—RÉNYI, A.: "On the evolution of random graphs". *A Magyar Tud. Akadémia Matematikai Kutató Intézetének Közleményei* **5.** (1960) A. 17—61.

[2] PALÁSTI, I.: "On the distribution of the number of trees which are isolated subgraphs of a chromatic random graph". *A Magyar Tud. Akadémia Matematikai Kutató Intézetének Közleményei* **6** (1961) A. 405—409.

ГРАНИЧНЫЕ ФУНКЦИИ ДЛЯ ПОДГРАФОВ ДВУХЦВЕТНЫХ СЛУЧАЙНЫХ ГРАФОВ ДАННОГО ТИПА

I. PALÁSTI

Резюме

Пусть $\Gamma_{m,n,N}$ есть случайный граф четного обхода, который состоит из m данных пронумерованных вершин $P_1, P_2, \ldots, P_m$, окрашенных в один цвет, n данных пронумерованных точек $Q_1, Q_2, \ldots, Q_n$ окрашенных в другой цвет, и N наугадвыбранных ребер. Предположим, что точки одного цвета нельзя соединить ребром и что все возможные выборы графов одинакого вероятны. В связи с увеличением числа ребер $N(m, n)$ имеет место следующее: Если $k \geq s$, $l \geq 1$, v положительные целые числа $(k + l - 1 \leq v \leq kl)$ и $\mathscr{B}_{k,l,v}$ обозначает любой не пустой класс таких связных, уравновешенных, случайных графов с четным обходом, которые содержат k точек одного цвета и l точек другого цвета, соединенных v ребрами, то граничная функция относительно того свойства случайного двухцветного графа, что он содержит хотя бы один подграф, изоморфный некоторому элементу $\mathscr{B}_{k,l,v}$, равна $n^{2-\frac{k+l}{v}}$, при условии, что $m \sim cn$ (где $c > 0$ и не зависит от n).