The maximum order of a strong matching in a random graph

El Maftouhi. A
Laboratoire de Recherche en Informatique
Université de Paris Sud, Centre d'Orsay
Bât, 490, 91405 Orsay Cedex, FRANCE

L. Marquez Gordones
School of Computing Sciences
Faculty of Sciences, Universidad Central de Venezuela,
47002 Los Chaguaramos, Caracas 1041-A, VENEZUELA

Abstract

A strong matching S in a given graph G is a set of disjoint edges $\{e_1, e_2, ..., e_m\}$ such that no other edge of the graph G connects an end-vertex of e_i with an end-vertex of $e_j, (e_i \neq e_j)$.
Let $G_{n,p}$ be the random graph on n vertices with fixed edge probability p, $0 < p < 1$. It is shown that, with probability tending to 1 as $n \to \infty$, the maximum size β of a strong matching in $G_{n,p}$ satisfies

$$\log_d n - \frac{1}{2}\log_d \log_d n - c_1 \leq \beta \leq \log_d n - \frac{1}{2}\log_d \log_d n + c_2$$

where c_1 and c_2 are constants depending only on p, and $d = \frac{1}{1-p}$.

Résumé

Un couplage fort S dans un graphe G est un ensemble d'arêtes disjointes $\{e_1, e_2, ..., e_m\}$ tel qu' aucune autre arête du graphe G ne relie une extremité de e_i avec une extremité de $e_j, (e_i \neq e_j)$.
Soit $G_{n,p}$ le graphe aléatoire à n sommets et de probabilité d'arête fixée p, $0 < p < 1$. On montre qu'avec une probabilité qui tend vers 1 quand $n \to \infty$, la taille maximum β d'un couplage fort dans $G_{n,p}$ verifie

$$\log_d n - \frac{1}{2}\log_d \log_d n - c_1 \leq \beta \leq \log_d n - \frac{1}{2}\log_d \log_d n + c_2$$

où c_1 and c_2 sont deux constantes dépendant seulement de p, et $d = \frac{1}{1-p}$.

Australasian Journal of Combinatorics 10(1994), pp.97-104

1 Introduction

Let $G = (V, E)$ denote a graph with vertex set V and edge set E. Let S be a subset of $E(G)$, $S = \{e_1, e_2, ..., e_m\}$. We say that S is a *strong matching* if $e_1, e_2, ..., e_m$ are disjoint, and no other edge of the graph connects an end-vertex of e_i with an end-vertex of e_j $(e_i \neq e_j)$. We shall call a strong matching of size m a m-strong matching.

In what follows $G_{n,p}$ denotes the random graph on n vertices with fixed edge probability p, $0 < p < 1$. We put $d = \frac{1}{1-p}$. Here, *almost always* means with probability tending to 1 as $n \to \infty$.

One of the surprising results in random graph theory was discovered by Matula [Mat 76], see also [Boll 85] (pages 251-257). He proved that, almost always, the independence number α of the random graph $G_{n,p}$ achieves only two possible values. More precisely, for every $\epsilon > 0$, almost always

$$\lfloor 2\log_d n - 2\log_d \log_d n + 1 + 2\log_d(\frac{e}{2}) - \epsilon \rfloor \le \alpha$$
$$\le \lfloor 2\log_d n - 2\log_d \log_d n + 1 + 2\log_d(\frac{e}{2}) + \epsilon \rfloor .$$

A similar problem devoted to maximal induced trees in $G_{n,p}$ was considered by P. Erdös and Z. Palka [E. P. 83]. They proved that for every $\epsilon > 0$, almost always, $G_{n,p}$ contains a maximal induced tree of order r if

$$(1 + \epsilon)\log_d n \le r \le (2 - \epsilon)\log_d n$$

but, almost always, $G_{n,p}$ does not contain a maximal induced tree of order smaller than $(1 - \epsilon)\log_d n$ or greater than $(2 + \epsilon)\log_d n$.

P. Erdös and B. Bollobás [B. E. 76] proved a similar result for maximal complete subgraphs in $G_{n,p}$.

Ruciński [Ru. 87] considered the following more general case. Let $\mathcal{F} = \{F_k\}$ be a family of graphs where F_k has v_k vertices and e_k edges, $k = 1, \ldots$. He showed that the order T_n of the largest induced copy of a graph from $\mathcal{F}$ in $G_{n,p}$ satisfies

$$\frac{T_n}{\log n} \to \frac{2}{A} \quad \text{as} \quad n \to \infty \text{ in } \text{ probability}$$

and, if F_k is an induced subgraph of F_{k+1}, $k = 1, \ldots,$ then

$$\frac{t_n}{\log n} \to \frac{1}{A} \quad \text{as} \quad n \to \infty \text{ in } \text{ probability}$$

where $A = a\log\left(\frac{1}{p}\right) + (1-a)\log\left(\frac{1}{1-p}\right)$ and $a = \lim_{k\to\infty} e_k/\binom{v_k}{2}$.

One can easily deduce from Ruciński's result that, almost always, the maximum size β of a strong matching in the random graph $G_{n,p}$ satisfies

$$(1 - \epsilon)\log_d n < \beta < (1 + \epsilon)\log_d n$$

where ϵ is an arbitrary positive constant.

The purpose of this paper is to estimate more precisely using the second moment method the parameter β in $G_{n,p}$. We shall prove that β achieves only a finite number of values as the following theorem shows.

Theorem *Let $G_{n,p}$ be the random graph with edge probability p fixed, $0 < p < 1$. Let $d = \frac{1}{1-p}$. There exist positive constants c_1 and c_2 depending only on p and not on n such that*
1) if $m \leq \log_d n - \frac{1}{2}\log_d \log_d n - c_1$ then, almost always, $G_{n,p}$ contains a strong matching of size m.
2) if $m \geq \log_d n - \frac{1}{2}\log_d \log_d n + c_2$ then, almost always, $G_{n,p}$ does not contain a strong matching of size m.

In sections 2 and 3 we compute the expectation and the variance of the number M of strong matchings of G, and in section 3 we conclude the proof of the theorem.

2 Expectation of the number of m-strong matchings

Proposition 1 *Let $G_{n,p}$ denote the random graph on n vertices with edge probability p, $0 < p < 1$. Let $d = \frac{1}{1-p}$. Then, the expectation $E(M)$ of the number of m-strong matchings in $G_{n,p}$ satisfies.*
i) $E(M) \to \infty$ as $n \to \infty$ if $m < \log_d n - \frac{1}{2}\log_d \log_d n + \frac{1}{2}\log_d\left(\frac{ep}{2}\right)$.
ii) $E(M) \to 0$ as $n \to \infty$ if $m \geq \log_d n - \frac{1}{2}\log_d \log_d n + \frac{1}{2}\log_d\left(\frac{ep}{2}\right)$.

Proof. Let $M = M_m$ be the number of m-strong matchings in $G_{n,p}$. Clearly, we have

$$E(M) = \binom{n}{2m}\binom{2m}{2,\ldots,2}\frac{1}{m!}p^m(1-p)^{2(m^2-m)}$$

where $\binom{n}{2m}\binom{2m}{2,\ldots,2}\frac{1}{m!}$ is the total number of m-strong matchings in the complete graph on n vertices and $p^m(1-p)^{2(m^2-m)}$ is the probability that $G_{n,p}$ contains any fixed m-strong matching.

$$E(M) = \frac{n!}{m!(n-2m)!}\left[\frac{p(1-p)^{2(m-1)}}{2}\right]^{2m^2}.$$

Stirling's formula gives

$$E(M) \simeq \frac{1}{\sqrt{2\pi m}}\left[\frac{en^2p(1-p)^{2m(-1)}}{2m}\right]^m. \tag{1}$$

Then $E(M) \to 0$ if $m \to \infty$ and for large n

$$\frac{en^2 p(1-p)^{2m(-1)}}{2m} \leq 1. \qquad (2)$$

Taking the log of both sides of (2) we get

$$m \geq -\frac{\log n}{\log(1-p)} + \frac{1}{2}\frac{\log m}{\log(1-p)} - \frac{\log\left(\frac{ep}{2}\right)}{\log(1-p)}. \qquad (3)$$

By setting $d = \frac{1}{1-p}$, we get

$$m \geq \log_d n - \frac{1}{2}\log_d m + \frac{1}{2}\log_d\left(\frac{ep}{2}\right)$$

The lower bound is asymptotic to $\log_d n$. We substitute this value in the r.h.s. of (3) and find

$$m \geq \log_d n - \frac{1}{2}\log_d \log_d n + \frac{1}{2}\log_d\left(\frac{ep}{2}\right).$$

Similarly, if

$$m < \log_d n - \frac{1}{2}\log_d \log_d n + \frac{1}{2}\log_d\left(\frac{ep}{2}\right)$$

then $E(M) \to \infty$ □

3 Variance of the number of m-strong matchings

Let S_1 and S_2 be two fixed strong matchings of size m. We denote by c the cardinality of the common part of the strong matchings S_1 and S_2, by a the number of edges belonging to S_1 and not to S_2 (a is also the number of edges belonging to S_2 and not to S_1) and by b the number of vertices which are incident to two distinct edges, one in S_1 and the other in S_2 (see Figure 1, next page). The parameters a, b, c satisfy $a+b+c=m$. Let us denote by $E_{a,b,c}$ the expectation of the number of pairs (S_1, S_2) corresponding to the above notation. Then we have the following proposition.

Proposition 2. *Let M denote the number of strong m-matchings in $G_{n,p}$. We have*

$$E(M^2) = \sum_{a+b+c=m} E_{a,b,c}$$

where

$$E_{a,b,c} = \frac{n!}{(n-4a-3b-2c)!(a!)^2 b! c!}\left(\frac{p}{2}\right)^{2a+c} p^{2b}(1-p)^{\frac{1}{2}[8m^2-8m+b+4c-(b+2c)^2]} \qquad (4)$$

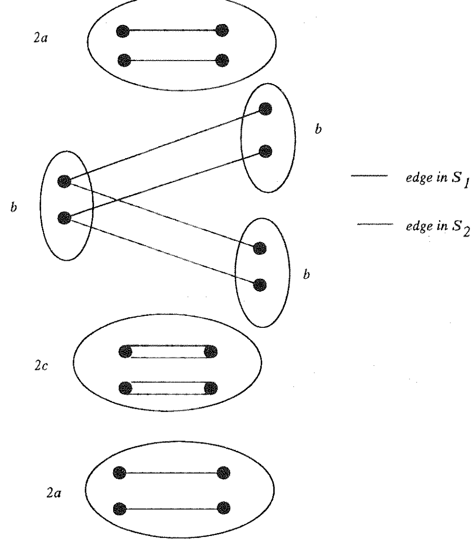

Figure 1:

Proof. Let a,b and c be fixed. Then the number of possible pairs (S_1, S_2) of m-strong matchings corresponding to the parameters a, b, c is

$$\binom{n}{2a, 2a, 2c, b, b, b, n-4a-3b-2c}\left[\binom{2a}{2,\ldots,2}/a!\right]^2\binom{2c}{2,\ldots,2}(c!)^{-1}(b!)^2.$$

The probability π that $G_{n,p}$ contains any fixed pair (S_1, S_2) of m-strong matchings is

$$\begin{aligned}\pi &= p^{2a+2b+c}(1-p)^{2\binom{2m}{2}-\binom{2c+b}{2}-(2a+2b+c)} \\ &= p^{2a+2b+c}(1-p)^{\frac{1}{2}[8m^2-8a-7b-4c-(b+2c)^2]}.\end{aligned}$$

After some calculation we get (4), and for $4a+3b+2c = o(\sqrt{n})$, we have

$$E_{a,b,c} \simeq \frac{n^{4m-(b+2c)}}{(a!)^2 b! c!} 2^{-2a-c} p^{2a+2b+c}(1-p)^{8m^2-8m+b+4c-(b+2c)^2}.$$

4 End of the proof

Let us prove that if $\log_2 n - \frac{1}{2}\log_2\log_2 n < \alpha$ where α is a positive constant which will be specified later, then

$$\frac{\sigma^2(M)}{E^2(M)} \to 0 \quad \text{as } n \to \infty. \tag{5}$$

This implies the Theorem, using Chebyshev's inequality.
By relations (1) and (2) we have, for all (a,b,c) ,

$$\begin{aligned}\frac{E_{a,b,c}}{E^2(M)} &\simeq \frac{(m!)^2 n^{-2c-b}}{(a!)^2(b!)c!} 2^{2m-2a-c} p^{2a+2b+c-2m} (1-p)^{\frac{1}{2}[b+4c-(b+2c)^2]} \\ &\simeq \frac{(m!)^2 n^{-2c-b}}{(a!)^2(b!)c!} 2^{2b+c} p^{2m-c} (1-p)^{\frac{1}{2}[b+4c-(b+2c)^2]}.\end{aligned}$$

By writing $n = (1-p)^{-\log_d n}$, the above relation gives

$$\begin{aligned}\frac{E_{a,b,c}}{E^2(M)} &\simeq \frac{(m!)^2}{(a!)^2 b! c!} 2^{2b+c} p^{-c} (1-p)^{\frac{1}{2}(b+2c)[2\log_d n-(b+2c)]+b+4c} \\ &\simeq \frac{(m!)^2}{(a!)^2 b! c!} 2^{2b+c} \left(\frac{1}{p}-1\right)^c (1-p)^{\frac{1}{2}(b+2c)[2\log_d n-(b+2c)+1]}.\end{aligned}$$

Thus

$$\frac{E_{a,b,c}}{E^2(M)} \simeq \frac{(m!)^2}{(a!)^2 b! c!} \left(\frac{1}{p}-1\right)^c (1-p)^{\frac{1}{2}(b+2c)[2\log_d n-(b+2c)+1-\log_d 2]}. \tag{6}$$

Clearly, for $a = m$ and $b = c = 0$, we have

$$\frac{E_{m,0,0}}{E^2(M)} \to 1 \quad \text{as} \quad n \to \infty.$$

So, it remains to prove that

$$\sum_{a+b+c=m, a\neq m} \frac{E_{a,b,c}}{E^2(M)} \to 0 \text{ as } n \to \infty.$$

Since the number of terms of the above sum is smaller than m^2, we need to prove, for all (a,b,c) with $a \neq m$

$$\frac{E_{a,b,c}}{E^2(M)} = o(\frac{1}{m^2}).$$

Set $x = 2b + c$, $1 \leq x \leq 2m$. Let $f(x)$ be the function defined by

$$f(x) = x(2\log_d n - x + 1 - \log_d 2).$$

So, for $1 \leq b + c \leq 2\log_d m$, we have

$$f(2b+c) \geq f(1) = 2\log_d n - \log_d 2.$$

Therefore

$$\begin{aligned}\frac{E_{a,b,c}}{E^2(M)} &\leq m^{2\log_d m} \left(\frac{1}{p}-1\right)^c (1-p)^{2\log_d n-\log_d 2} \\ &\leq m^{2\log_d m} \left(\frac{1}{p}\right)^{2\log_d m} (1-p)^{2\log_d n-\log_d 2} \\ &= o\left(\frac{1}{m^m}\right).\end{aligned}$$

While, for $2\log_d m \le b+c \le 2m$ and for m sufficiently large, we have

$$\begin{aligned} f(2b+c) &\ge f(2m) = 2m[2\log_d n - 2m + 1 - \log_d 2] \\ &\ge 2m[\log_d \log_d n + 2\alpha + 1 - \log_d 2]. \end{aligned}$$

This bound can be applied to equation (6) to obtain

$$\begin{aligned} \frac{E_{a,b,c}}{E^2(M)} &\le \frac{m!m^m}{a!b!c!}\left(\frac{1}{p}-1\right)^c (1-p)^{m[\log_d \log_d n + 2\alpha+1-\log_d 2]} \\ &\le \frac{m!}{a!b!c!}\left(\frac{1}{p}-1\right)^c (1-p)^{m[\log_d \log_d n - \log_d m + 2\alpha+1-\log_d 2]} \\ &\le \frac{m!}{a!b!c!}\left(\frac{1}{p}-1\right)^c (1-p)^{m[2\alpha+1-\log_d 2+o(1)]}. \end{aligned}$$

It follows that

$$\begin{aligned} \sum_{b+c>2\log_d m} \frac{E_{a,b,c}}{E^2(M)} &\le (1-p)^{m[2\alpha+1-\log_d 2+o(1)]} \sum \frac{m!}{a!b!c!}\left(\frac{1}{p}-1\right)^c \\ &\le \left[\left(1+\frac{1}{p}\right)(1-p)^{2\alpha+1-\log_d 2+o(1)}\right]^m. \end{aligned}$$

So,

$$\sum_{b+c>2\log_d m} \frac{E_{a,b,c}}{E^2(M)} = o(1)$$

if

$$\left(1+\frac{1}{p}\right)(1-p)^{2\alpha+1-\log_d 2+o(1)} < 1. \tag{7}$$

Any constant $\alpha > \frac{1}{2}[\log_d(1+\frac{1}{p})+\log_d 2-1]$ satisfies the inequality (7). This concludes the proof of the theorem. $\square$

5 Open problems

Problem 1

Find the minimum size of a maximal strong matching in the random graph $G_{n,p}$, where p is fixed.

Problem 2

Find estimates for the maximum size of a strong matching in the random graph $G_{n,p}$ with edge probability $p = \frac{c}{n}$ where $c > 0$ is a positive constant.

Problem 3

The strong chromatic index $sci(G)$ of a graph G is the smallest integer k such that the edge set of G can be partitioned into k induced matchings. If $e(G)$ denotes the number of edges of G then

$$sci(G) \geq \frac{e(G)}{\beta(G)}.$$

From our result one can deduce immediately that, almost always, the strong chromatic index $sci = sci(G_{n,p})$ of the random graph $G_{n,p}$, p fixed, satisfies

$$sci \geq (1 - o(1))\frac{pn^2}{\log_d n}.$$

Is true that, almost always, in $G_{n,p}$ we have

$$sci \leq (1 + o(1))\frac{pn^2}{\log_d n}?$$

Acknowledgments

We would like to express our thanks to Zs. Tuza for suggesting this problem, and W. Fernandez de la Vega for helpful discussions concerning this paper.

We would like also to thank an anonymous referee whose suggestions lead to a substantial improvement of the results of the paper.

References

[Mat 76] D. W. Matula, *The Largest Clique Size in Random Graph,* Technical Report, Southern Methodist University, Dallas (1976)

[Boll 85] B. Bollobás, *Random graphs,* Academic press (1985).

[B. E. 76] B. Bollobás, P. Erdös, *Cliques in Random Graphs*, Math. Proc. Camb. Phil. Soc. 80 (1976), 419-427.

[E. P. 83] P. Erdös, Z. Palka, *Trees in Random Graphs*, Discrete Math. 46 (1983). *Addendum to Trees in Random Graphs*, Discrete Math. 48 (1984), 331.

[Ru. 87] A. Ruciński, *Induced Subgraphs in Random Graphs*, Annals of Discrete Mathematics. 33 (1987), 275-296.

(Received 27/10/93; revised 17/3/94)