LEVEL 2

AD A108826

DTIC
ELECTE
DEC 23 1981
S D
D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

K-CONNECTIVITY IN RANDOM UNDIRECTED GRAPHS

BY

John H. Reif and Paul G. Spirakis

TR-19-81

12 14

September 1981

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

DTIC
S ELECTE D
DEC 23 1981
D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO. AD-A108826

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
K-Connectivity in Random Undirected Graphs

5. TYPE OF REPORT & PERIOD COVERED
Technical Report

6. PERFORMING ORG. REPORT NUMBER
TR-19-81

7. AUTHOR(s)
John H. Reif
Paul G. Spirakis

8. CONTRACT OR GRANT NUMBER(s)
N00014-80-C-0674

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Harvard University
Cambridge, MA 02138

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

12. REPORT DATE
1981

13. NUMBER OF PAGES
14

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
same as above

15. SECURITY CLASS. (of this report)

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
unlimited

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
umlimited

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
connectivity, random graphs, k-blocks

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
See reverse.

81 12 22 117

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

20.

This paper concerns vertex connectivity in random graphs. We present results bounding the cardinality of the biggest k-block in random graphs of the $G_{n,p}$ model, for any constant value of k. These results generalize those of [Erdös, Renyi, 60] and [Karp, Tarjan, 80] for $k = 1$ and 2. We furthermore prove here that the cardinality of the biggest k-block is $\geq n-\log n$ with probability $\geq 1-n^{-2}$ for $p \geq c_1(k)/n$ and $c_1(k) > k+2$. We also show that if $p \geq c(k)\frac{\log n}{n}$ with $c(k) > 32k^2$ then the graph $G_{n,p}$ is k-connected with probability $\geq 1-2n^{-d'(k)}$, $d'(k) > 1$.

K-CONNECTIVITY IN RANDOM UNDIRECTED GRAPHS*

by

John H. Reif and Paul G. Spirakis
Harvard University
Aiken Computation Laboratory
Cambridge, MA 02138

*This work was supported in part by the National Science Foundation Grant NSF-MCS79-21024 and the Office of Naval Research Contract N00014-80-C-0674

1. Summary

This paper concerns vertex connectivity in random graphs. We present results bounding the cardinality of the biggest k-block in random graphs of the $G_{n,p}$ model, for any constant value of k. These results generalize those of [Erdös, Renyi, 60] and [Karp, Tarjan, 80] for $k=1$ and 2. We furthermore prove here that the cardinality of the biggest k-block is $\geq n-\log n$ with probability $\geq 1-n^{-2}$ for $p \geq c_1(k)/n$ and $c_1(k) > k+2$. We also show that if $p \geq c(k) \frac{\log n}{n}$ with $c(k) > 32k^2$ then the graph $G_{n,p}$ is k-connected with probability $\geq 1-2n^{-d'(k)}$, $d'(k) > 1$.

2. Introduction

A *graph* $G = (V,E)$ consists of a finite nonempty set V of *vertices* together with a prescribed set E of unordered pairs of *distinct* elements of V (set of edges). (We allow no loops neither multiple edges). The *vertex connectivity* $k(G)$ of an undirected graph G is the minimum number of vertices whose removal results in a disconnected graph or a *trivial* graph (consisting of just one vertex). Note that we follow here [Matula, 78] in defining k-connectivity, which we find to be most natural. [McLane, 37] gives a (somewhat different) definition of triconnectivity so that he can have the theorem that a graph is planar if its triconnected components are. [McLane, 37] shows that his triconnected components are homeomorphic to 3-blocks. Vertex k-connectivity seems to be a fundamental property of a graph and has numerous applications to other graph problems (such as planarity testing, routing problems etc). It is relevant to questions concerning vulnerability of a graph to separation. Cluster analysis methods considering the nature and inherent reliability of proximity data use the theory of k-connectivity to find groups of likes and dislikes in object pair association graphs ([Matula, 77], [Matula, 78] also [Jardine, Sibson, 71]).

A *k-block* of an undirected graph G is a maximal k-connected subgraph. A k-block is *trivial* if it has only one vertex [Matula, 78]. Clearly, each k-block consists of $\geq k$ vertices or it is trivial.

[Matula, 78] examined certain properties of k-blocks in graphs (number of them, separation lemma) and [Erdös, Renyi, 60] and [Karp, Tarjan, 80] examined the distribution of the size of the bigge[illegible] 1 and 2-blocks in random graphs $G_{n,p}$ with $p \geq \frac{c}{n}$ and $G_{n,N}$ with $N \geq cn$. They proved that there is a giant k-block for $k=1,2$, with exponentially decaying probability of error. For $p > \frac{1}{2}\frac{\log n}{n}$ [Erdös, Renyi 60] showed that $G_{n,p}$ becomes almost surely 2-connected.

In our paper we examine k-connectivity in the model $G_{n,p}$, defined precisely as follows: For $0 \leq p \leq 1$ and $n \geq 0$ let $G_{n,p}$ be a random variable whose values are graphs on the vertex set $\{1, 2, \ldots, n\}$. If $e = \{u,v\}$ and $u,v, \in \{1, 2, \ldots, n\}$ then $\text{Prob}\{e \text{ is an edge}\} = p$ and these probabilities are independent for different e.

We prove that for each constant $k \geq 0$ and for each ε $(0 \leq \varepsilon \leq 1)$ and $\alpha > 1$, there is a k-block of cardinality $\geq \varepsilon n$ in $G_{n,p}$ with $p \geq \frac{c(k, \varepsilon, \alpha)}{n}$ with probability $\geq 1 - e^{-\alpha \cdot n}$. We furthermore prove that for any $k > 0$ and $0 \leq m < \frac{n}{2k}$ there are constants $c(k), d(k) > 0$ such that the size of the biggest k-block of $G_{n,p}$ where $p \geqslant c(k)\frac{\log n}{n}$ is equal to n-m with probability $n^{-m \cdot d(k)}$. From that we get as corollaries, that there are $c(k), d(k) > 0$ and $d'(k) > 1$ such that the size of the biggest k-block of $G_{n,p}$ is $\geq n - \log n$ with prob $> 1-2n^{1-d(k)\log n}$ and that $G_{n,p}$ is k-connected with prob $> 1-2n^{-d'(k)}$.

Finally, we prove that for any $m = o(n)$ $\exists c_1(k) > k+2$ and a function $t(n) \geq \frac{c_1(k) \log n}{m}$ such that, if $p \geq \frac{t(n)}{n}$ then the biggest k-block of $G_{n,p}$ has size $> n-m$ with probability $> 1 - n^k/e^{t(n)m} \to 1$ as $n \to \infty$. A corollary is that if $p \geq \frac{c_1(k)}{n}$ then the biggest k-block of $G_{n,p}$ has cardinality $> n - \log n$ with probability $\geq 1-n^{-2}$. These results were known by [Erdös, Renyi, 60] only for $k=1$ and $c(1) \geq \frac{1}{2}$.

3. Properties of k-blocks

PROPOSITION 1 [Matula, 78] For each $k \geq 0$, any two k-blocks have no more than $k-1$ vertices in common.

DEFINITION [Matula, 78] A *separation set* S of G is a vertex subset $S \subseteq V(G)$ such that $G - S$ is disconnected. A *minimum separating set* $S \subset V(G)$ has $|S| = k(G)$.

DEFINITION Let G be a graph (V,E) and let $S \subseteq V$ be a set of vertices. Then by $\langle S \rangle$ we denote the subgraph induced by S on G.

LEMMA 1 [Matula, 78] (Block separation lemma) Let $S \subseteq V(G)$ be a minimum separating set of the noncomplete graph G with $\langle A_1 \rangle, \langle A_2 \rangle, \ldots, \langle A_m \rangle$, $m \geq 2$ the components of $G - \langle S \rangle$ and let $k \geq k(G) + 1$. Then each k-block of G is a k-block of $\langle A_i \cup S \rangle$ for precisely one value of i, and each k-block of $\langle A_i \cup S \rangle$ for every i is a k-block of G.

For a proof, see [Matula, 78].

REMARK [Matula, 78] shows that for each $k \geq 1$ the total number of nontrivial k'-blocks for $1 \leq k' \leq k$, is $\leq \left\lfloor \frac{2n-1}{3} \right\rfloor$ for any graph G with n vertices.

4. Giant k-blocks in Random Graphs

In the following we introduce special notation for very large subgraphs. For each ε, $0 \leq \varepsilon \leq 1$, a subgraph H of a graph G of n vertices is called an ε-*giant of* G if the cardinality of the vertex set of H is $\geq \varepsilon n$.

DEFINITION: Given a vertex set $S \subseteq V$ in the graph $G = (V,E)$, the *boundary vertices* of S is the set $B(S) = \{u \in S \mid \exists v \in V-S$ such that $\{u,v\} \in E\}$.

DEFINITION: Let X be a random variable whose values are the cardinality of the maximum k-block of instances of $G_{n,p}$. Let $F_{n,p,k}(a) = \text{Prob}\{X \leq a\}$ be the distribution function of X.

THEOREM 1: For every ε on $(0,1)$, $\alpha > 1$ and $k > 0$ there is a $c = c(k,\varepsilon,\alpha) > 0$ such that, for $p \geq \frac{c}{n}$, $F_{n,p,k}(\varepsilon n) \leq e^{-\alpha n}$. In other words, the random graph $G_{n,p}$ with $p \geq \frac{c}{n}$ has an ε-giant k-block with probability at least $1 - e^{-\alpha n}$. To prove this theorem, we shall need the following definition and lemma.

DEFINITION: If $G = (V,E)$ and A,B are subsets of V, then $E(A,B) = \{e = \{u,v\} \in E \mid u \in A$ and $v \in B\}$.

LEMMA 2: For any α_1, ε_1, $\varepsilon_2 > 0$ where $\varepsilon_1 + \varepsilon_2 \leq 1$ and $\alpha_1 \geq 1$ there are constants c, ε_3, $\varepsilon_4 > 0$ such that a random graph $G_{n,p}$ with $p \geq \frac{c}{n}$ has the property (*) with probability $\geq 1 - e^{-\alpha_1 n}$.

(*): If A,B are any two vertex subsets of V such that $|A| \geq \lfloor \varepsilon_1 n \rfloor$, $|B| \geq \lfloor \varepsilon_2 n \rfloor$ and $A \cap B = \emptyset$ then $|E(A,B)| > 0$.

PROOF OF LEMMA: The complement of (*) is: "There are two vertex subsets A,B such that $|A| \geq \lfloor \varepsilon_1 n \rfloor$, $|B| \geq \lfloor \varepsilon_2 n \rfloor$, $A \cap B = \emptyset$ and

$E(A,B) = \emptyset$". Clearly

$$\text{Prob}\{E(A,B) = \emptyset\} \le (1-p)^{\varepsilon_1 n \varepsilon_2 n} \le \left((1 - \frac{c}{n})^n\right)^{\varepsilon_1 \varepsilon_2 n} \le e^{-c\varepsilon_1 \varepsilon_2 n}$$

Since there are at most $\frac{1}{2} \cdot 4^n$ ways to select these A,B, and upper bound on the probability of the complement of (*) is

$$\sum_{\text{all } A,B} \text{Prob}\{E(A,B) = \emptyset\}$$

$$\le \frac{1}{2}\left(4e^{-c\varepsilon_1\varepsilon_2}\right)^n \le e^{-\alpha_1 n}$$

for

$$c \ge \frac{\alpha_1 + \log_e 4}{\varepsilon_1 \quad \varepsilon_2}$$

$\square$

Now we return to the proof of the Theorem 1. Let $G = (V,E)$ be an instance of the random graph $G_{n,p}$. Let $\mathcal{E}_1$ be the event "G has no ε-giant k-block". Assume event $\mathcal{E}_1$ be true in the instance G of $G_{n,p}$. Let initially the set $A = \emptyset$. Do the following construction just until A has cardinality $\ge \varepsilon' n/2$, where $\varepsilon' = \min(\varepsilon, 1-\varepsilon)$.

(a) Find a minimum separating set S of G. Let $\langle A_1\rangle, \ldots, \langle A_m\rangle$ $m \ge 2$ be the components of $G-S$. Let $\langle A_i\rangle$ be the smallest of them. Let $A \leftarrow (A_i \cup S) \cup A$. Let B be the union of the rest of the components and let $G \leftarrow$ the graph induced by $B \cup S$. If $|A| < \varepsilon' \cdot \frac{n}{2}$, then go to (a).

By the above method of constructing A, each addition of a component in A adds at most $k-1$ vertices to $B(A)$ (i.e. the vertices of the cut) and at least one vertex to $A - B(A)$ (by the block separation lemma and by the fact that k-blocks have $\geq k$ vertices if they are non-trivial) or causes the transformation of a boundary to a nonboundary vertex. Thus, at least $1/k$ of the vertices of A are not in $B(A)$.

By this construction, finally the k-blocks of G are going to be separated. Because all k-blocks have been assumed to have cardinality $\leq \varepsilon n$, we will finally have

$$\varepsilon' \quad \frac{n}{2} \leq |A| \leq \min\left[\varepsilon' \quad \frac{n}{2} + \varepsilon n,\ \varepsilon' \quad \frac{n}{2} \quad \frac{3}{2}\right]$$

So

$$|A - B(A)| \geq \frac{\min(\varepsilon, 1-\varepsilon)}{2k} \cdot n$$

and

$$|V - A| \geq n\left(1 - \min\left[(\varepsilon + \quad \varepsilon'/2\ ), \left(\ 3\varepsilon'/4\ \right)\right]\right.$$

(obviously $|V - A| > 0$ for any ε on $(0,1)$). Let $Y = A - B(A)$ and $Z = V - A$ then $|Y| \geq \varepsilon_1 n$ and $|Z| \geq \varepsilon_2 n$ where $\varepsilon_1 = \frac{\varepsilon'}{2k}$, $\varepsilon_2 = 1 - \min\left[\left(\varepsilon + \frac{1}{2}\varepsilon'\right), \left(3\varepsilon'/4\right)\right]$ and $E(Y,Z) = \emptyset$ by construction

Hence, there are disjoint sets $Y' \subseteq Y$ and $Z' \subseteq Z$ such that $|Y'| = \varepsilon_1 n$, $|Z'| = \varepsilon_2 n$ and $E(Y',Z') = \emptyset$. Call $\mathcal{E}_2$ the above event. We have just shown $\mathcal{E}_1$ implies $\mathcal{E}_2$. So,

$$\text{prob}\{\mathcal{E}_1\} \leq \text{Prob}\{\mathcal{E}_2\} \leq e^{-\alpha \cdot n}$$

by Lemma 2.

$\square$

NOTE: According to Lemma 2, any $\alpha \geq 1$ and $c \frac{\alpha + \log_e 4}{\varepsilon_1 \varepsilon_2}$ satisfy the theorem. Replacing $\varepsilon_1, \varepsilon_2$ with the expressions found, we get

$$c \geq 2k \left[\frac{\alpha + \log_e 4}{\varepsilon' \cdot \left(1 - \min(\varepsilon + \frac{1}{2}\varepsilon' \quad \frac{3}{4}\varepsilon')\right)} \right]$$

5. k-blocks of dense random graphs.

This section considers edge density $p \geq c \frac{\log n}{n}$.

THEOREM 2. For any constant integer $k > 0$ and any n and $m < \frac{n}{2k}$ there are constants $c(k), d(k) > 0$ such that the cardinality X of the biggest k-block of the graph $G_{n,p}$ with $p \geq c(k) \frac{\log n}{n}$ satisfies the property

$$\text{Prob}\{X = n-m\} \leq n^{-md(k)}$$

PROOF: Let G be an instance of $G_{n,p}$ and let the event $X = n-m$ be true in that instance. Let A be a k-block with $|A| = X$. For every $u \in V-A$, we have that

$$\left|\left\{\{u,v\} \in E(G) : v \in A\right\}\right| \leq k - 1$$

(since, otherwise u would belong to A). Let

$$A_1 = \left\{v \in A : \exists u \in V-A : \{u,v\} \in E(G)\right\}$$

then

$$|A_1| \leq (k-1)\ |V-A| = (k-1)m \quad .$$

Let $A_2 = A - A_1$. We get

$$|A_2| \geq n-m - (k-1)m = n-km$$

Furthermore, there is no edge from $V-A$ to A_2.

Let $\mathscr{E}$ be the above event. The probability of $\mathscr{E}$ is bounded above by

$$u(m,n) = \binom{n}{m}\binom{n-m}{n-km}(1-P)^{(n-km)m} \tag{***}$$

But
$$(1-p) \leq \left(1 - \frac{c \log n}{n}\right) \leqslant e^{-c \frac{\log n}{n}}$$

since
$$p \geqslant \frac{c \log n}{n} .$$

Also
$$\binom{n-m}{n-km} \leq \binom{n-m}{(k-1)m} \leq e^{(k-1)m \log(n-m)}$$

since
$$(k-1)m < \frac{n-m}{2}$$

and
$$\binom{n}{m} \leq e^{m \log n}$$

since
$$m < \frac{n}{2}$$

Thus $u(n,m) \leq n^{-d(n,m)}$ where $d(n,m) =$

$$cm\left(1 - \frac{km}{n}\right) - m - (k-1)m\frac{\log(n-m)}{\log n}$$

$$> cm\left(1 - \frac{km}{n}\right) - m - (k-1)m$$

$$> \frac{c}{2}m - km \qquad \text{(by our assumption).}$$

So, $d(n,m) > md(k)$ where $d(k) = \frac{c}{2} - k$. Note that $d(k) > 0$ iff $c(k) > 2k$.

So, $\text{Prob}(\mathscr{E}) \leq n^{-m\, d(k)}$.

$\square$

THEOREM 3: For any constant integer $k > 0$ and any $n >> k$ there is a constant $c(k) > 0$ and a $d(k) > 0$ such that the cardinality X of the biggest k-block of the graph $G_{n,p}$ with $p \geq c(k) \frac{\log n}{n}$ satisfies the property

$$\text{Prob}\{X \leq n - \log n\} < 2n^{(1-d(k)\log n)}$$

PROOF: By using theorem 2, we get

$$\text{Prob}\left\{\log n \leq n - X < \frac{n}{2k}\right\} = \sum_{m=\log n}^{n/2k} n^{-md(k)}$$

with $$d(k) = \frac{c(k)}{2} - k > 0 \quad \text{for} \quad c(k) > 2k.$$

So, $$\text{Prob}\left\{\log n \leq n - X < \frac{n}{2k}\right\} < n \cdot n^{-\log n \cdot d(k)} < n^{1-d(k)\log n}.$$

Also, by theorem 1 and using $\varepsilon = \frac{1}{2k}$ we get

$$\text{Prob}\left\{n - X > \frac{n}{2k}\right\} < e^{-\alpha \cdot n}$$

for any $\alpha > 1$ and $c(k) \geq \frac{\alpha + \log_e 4}{\varepsilon_1 \varepsilon_2}$ and $\varepsilon_1 \varepsilon_2 = \frac{1}{2k}\left(1 - \frac{3}{8k}\right)$.

So, for $$c(k) > \max\left(2k, \frac{\alpha + \log_e 4}{\varepsilon_1 \varepsilon_2}\right)$$

or $$c(k) > (\alpha + \log_e 4)16k^2$$

we get

$$\text{Prob}\{\log n \leq n - X\} < e^{-\alpha \cdot n} + n^{1-\log n \cdot d(k)}$$

or

$$\text{Prob}\{X \leq n - \log n\} < 2n^{1-d(k)\cdot \log n}$$

for sufficiently large n.

□

NOTE: Theorem 3 says that for $p \geq c(k)\frac{\log n}{n}$ the graph $G_{n,p}$ has a k-block of size $\geq n - \log n$ with probability limiting to 1 as $n \to \infty$.

THEOREM 4: For any constant integer $k > 0$ and $n >> k$ there are constants $c(k) > 0$, $d'(k) > 1$ such that the random graph $G_{n,p}$ with $p \geq c(k)\frac{\log n}{n}$ is k-connected with probability

$$\geq 1 - 2n^{-d'(k)} .$$

PROOF: Let $R = n - X$ where X = cardinality of the biggest k-block of $G_{n,p}$. By using theorems 2, 3 and $c(k) > 2 + \max\left(2k, \frac{\alpha+\log_e 4}{\varepsilon_1\varepsilon_2}\right)$ with $\varepsilon_1\varepsilon_2 = \frac{1}{2k}\left(1 - \frac{3}{8k}\right)$ we get that

$$\mathrm{Prob}\{1 \leq R\} < e^{-\alpha \cdot n} + n^{1-(\frac{c}{2} - k)} .$$

Let

$$d'(k) = \frac{c(k)}{2} - k - 1 .$$

Then $d'(k) > 1$ for $c(k) > 2 + \left(\max 2k, \frac{\alpha + \log_e 4}{\varepsilon_1\varepsilon_2}\right)$

and

$$\mathrm{Prob}\{1 \leq R\} \leq e^{-\alpha n} + n^{-d'(k)} < 2n^{-d'(k)}$$

for large n.

Hence

$$\mathrm{Prob}\{R = 0\} > 1 - 2n^{-d'(k)}$$

$\square$

6. k-blocks for intermediate edge densities.

Let $\frac{c}{n} \leq p \leq c' \frac{\log n}{n}$. We wish to study the k-connectivity of this class of random graphs.

THEOREM 5. For any constant $k \geq 0$ and any $m = o(n)$ there is a constant $c_1(k) > 0$ and a function $t(n) > \frac{c_1(k) \log n}{m}$ such that, if $p > \frac{t(n)}{n}$ then if X is the cardinality of the biggest k-block of $G_{n,p}$ then

$$\operatorname{Prob}\{X \leq n - m\} \leq \frac{n^k}{e^{t(n)\, m}} \to 0 \quad \text{as} \quad n \to \infty .$$

PROOF: Assume that in the instance G of $G_{n,p}$ the cardinality X of the biggest k-block satisfies the inequality $X \leq n - m$. Then, we can find two sets Y, Z (as in proof of theorem 3) such that $|Y| = m$, $|Z| = n - km$ and no edge between them. This event is above bounded by the probability $1 - q$ where

$q = \operatorname{Prob}\{$for every pair of disjoint sets Y,Z of vertices of the above sizes, there is at least one edge between $Y, Z.\}$

We shall show $q \to 1$ as $n \to \infty$. Let us enumerate all possible pairs of sets of vertices of the above sizes. Call them

$$(Y_1,Z_1),\ Y_2,Z_2),\ \ldots,\ (Y_g,Z_g)$$

where

$$g = \binom{n}{m}\binom{n-m}{n-km} = \binom{n}{m}\binom{n-m}{(k-1)\ m}$$

We have that $q =$

$$\mathrm{Prob}\{E(Y_1,Z_1) \neq \emptyset \wedge \ldots, \wedge E(Y_g,Z_g) \neq \emptyset\}$$

where $E(Y,Z)$ = set of edges between Y,Z.

So, by Baye's formula, $q =$

$$\mathrm{Prob}\,\{E(Y_1,Z_1) \neq \emptyset\} \quad \mathrm{Prob}\left\{\frac{E(Y_2,Z_2) \neq \emptyset}{E(Y_1,Z_1) \neq \emptyset}\right\} \cdots \mathrm{Prob}\left\{\frac{E(Y_g,Z_g) \neq \emptyset}{\bigwedge_{i=1,\ldots g-1} E(Y_i,Z_i) \neq \emptyset}\right\}$$

We need the following enumeration lemma:

LEMMA 3: For every two sets Y_i, Z_i having at least one edge e between them, there are at least

$$g_i = \binom{n-2}{m-1}\binom{n-2-(m-1)}{(k-1)m-1}$$

pairs of sets of sizes m, $n - km$ which also have this edge between them.

This lemma can be proved easily by taking out the two vertices of e and enumerating.

COROLLARY: There is a suitable enumeration of the sets in the q product such that for every term i not equal to 1 the next g_1 or more terms (conditioned on the existence of an edge from A_i to B_i) will be equal to 1.

Hence, the value of q is

$$q \geq \left[\mathrm{Prob}\{E(Y_1,Z_1) \neq \emptyset\}\right]^{g/g_1}$$

But

$$g/g_1 \leq \left(\frac{n}{m}\right)^k \quad \text{as} \quad n \to \infty .$$

Hence,

$$q \geq \left[1 - (1-p)^{m(n-km)}\right]^{(n/m)^k}$$
$$\geq \left[1 - \left((1-p)^{1/p}\right)^{pm(n-km)}\right]^{(n/m)^k}$$

or

$$q \geq \left(1 - e^{-pm(n-km)}\right)^{\left(\frac{n}{m}\right)^k}$$
$$\geq 1 - \left(\frac{n}{m}\right)^k e^{-t(n)m}$$

or

$$q \geq 1 - e^{-[t(n)m - k\log n]} > 1 - n^{-2}$$

if $c_1(k) > k + 2.$

(Since $t(n)m > c_1(k)\log n > (k+2)\log n$)

So,

$$\text{Prob}\{X < n - m\} < e^{-[t(n)m - k\log n]} \to 0 \quad \text{as} \quad n \to \infty$$

for the above values of $c(k)$

COROLLARY: For $m = \log n$ and $t(n) \geq c_1(k) > k + 2$ we get: For each $k > 0$, the graph $G_{n,p}$ with $p \geq \frac{c_1(k)}{n}$, has a k-block of cardinality $> n - \log n$ with probability $\geq 1 - n^{-2}$.

REFERENCES

Erdös, P. and A. Renyi, "On Random Graphs." Publicationes Mathematicae 6, (1959), pp. 290-297.

Erdös, P. and A. Renyi, "On the evolution of random graphs." Publ. Math. Inst. Hung. Acad. Sci. 5A (1960), pp. 17-61.

Jardine, N. and R. Sibson, "Mathematical Taxonomy." Wiley, London (1971).

Karp, R.M. and R.E. Tarjan, "Linear expected time algorithms for connectivity Problems." 12th annual ACM Symposium on Theory of Computing, Los Angeles, (1980).

Matula, D., "K-blocks and ultrablocks in graphs." Journal of Combinatorial Theory, b24 (1978), pp. 1-13.

Matula, D., "Graph theoretic techniques for cluster analysis algorithms," in Classification and Clustering, edited by F. Van Ryzon. Acad. Press, NY (1977), pp. 95-129.

MacLane, S., "A structural characterization of planar combinatorial graphs." Duke Math. J., 3 (1937), pp. 340-472.