*Combinatorics, Probability and Computing* (1997) 6, 349–352. Printed in the United Kingdom
© 1997 Cambridge University Press

# Bipartite Coverings of Graphs

VOJTECH RÖDL[1] and ANDRZEJ RUCIŃSKI[1,2]†

[1]Department of Mathematics and Computer Science, Emory University, Atlanta GA 30322, USA
(e-mail: rodl@mathcs.emory.edu)

[2]Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland
(e-mail: ruciński@math.amu.edu.pl)

*Received 11 September 1995*

In this note we give a probabilistic proof of the existence of an $n$-vertex graph $G_n$, $n = 1, 2, \ldots$, such that, for some constant $c > 0$, the edges of $G_n$ cannot be covered by $n - c \log n$ complete bipartite subgraphs of $G_n$. This result improves a previous bound due to F. R. K. Chung and is the best possible up to a constant.

Let $G$ be a graph. *A covering of* $G$ is a set of subgraphs $G_1, \ldots, G_r$ of $G$ such that every edge of $G$ belongs to at least one of the graphs $G_i$, $i = 1, \ldots, r$. For a family of graphs $\mathscr{H}$, a covering $G_1, \ldots, G_r$ of $G$ is called *an* $\mathscr{H}$*-covering* if each graph $G_i$, $i = 1, \ldots, r$, belongs to family $\mathscr{H}$.

For instance, every graph on $n$ vertices has an $\mathscr{S}$-covering consisting of at most $n - 1$ subgraphs, where $\mathscr{S}$ is the family of stars. In this paper we shall be interested in $\mathscr{B}$-coverings, where $\mathscr{B}$ is the family of complete bipartite graphs. It is easy to observe that the complete graph $K_n$ has a $\mathscr{B}$-covering with only $\log_2 n$ subgraphs. However, there exist $n$-vertex graphs which require many more subgraphs in any $\mathscr{B}$-covering. Let $\rho(n)$ be the smallest integer such that every graph on $n$ vertices has a $\mathscr{B}$-covering with at most $\rho(n)$ subgraphs.

F. R. K. Chung [1] proved that $\rho(n) > n - n^{11/14+\varepsilon}$ for any positive $\varepsilon$ and sufficiently large $n$. That result implied the conjecture of J.-C. Bermond that $\lim_{n\to\infty} \rho(n)/n = 1$. On the other hand, a simple upper bound of the form $\rho(n) < n - \Theta(\log n)$ was pointed out by P. Erdős as a consequence of a well-known upper bound on the Ramsey number $r(k, k)$. Indeed, as $r(k, k) < 4^k$, every graph on $n$ vertices contains either an independent set or a clique of order $\Theta(\log n)$. In the former case $n - \Theta(\log n)$ stars suffice to cover the

† Research supported by NSF grants DMS-9401559 and INT-9406971, and by KBN grant 2 P03A 023 09.

https://doi.org/10.1017/S0963548397003064 Published online by Cambridge University Press

graph. In the latter, in addition one needs a $\mathscr{B}$-covering of the clique of order $\Theta(\log n)$, but this requires only $\Theta(\log\log n)$ subgraphs. (Tuza [2] improved this upper bound to $n - \lfloor \log_2 n \rfloor + 1$.)

The result proved in this note provides a matching lower bound on $\rho(n)$, improving the previous one due to Chung.

**Theorem.** *There exists a constant $c > 0$ such that*

$$\rho(n) > n - c\log n.$$

The proof consists of one deterministic observation and one probabilistic lemma.

A $\mathscr{B}$-covering $G_1, \ldots, G_r$ is called *strong* if there is no star among the graphs $G_i \setminus \bigcup_{j\neq i} G_j$, where $G_i \setminus \bigcup_{j\neq i} G_j$ is the graph without isolated vertices and with the edge set $E(G_i) \setminus \bigcup_{j\neq i} E(G_j)$.

**Fact.** *Let $G$ be an $n$-vertex graph and let $t < n$. If no subgraph of $G$ on at least $v \geqslant t$ vertices admits a strong $\mathscr{B}$-covering consisting of less than $v$ subgraphs, then every $\mathscr{B}$-covering of $G$ requires more than $n - t$ subgraphs.*

**Proof.** Let $G_1, \ldots, G_r$ be a $\mathscr{B}$-covering of $G$ with $r < n$. Set $G_i^* = G_i \setminus \bigcup_{j\neq i} G_j$, $i = 1, \ldots, r$. Define a new $\mathscr{B}$-covering $\bar{G}_1, \ldots, \bar{G}_r$ recursively as follows. Initially set all $\bar{G}_i = G_i$. Since $n > t$, there is a star among the graphs $G_i^*$. Let $G_{j_1}^*$ be such a star. Then reset $\bar{G}_{j_1}$ as the complete star of $G$ containing $G_{j_1}^*$. (A complete star of $G$ is a subgraph of $G$ consisting of all edges incident to a vertex.) Next, consider the graphs $G_i^{**} = G_i - \bigcup_{j\neq i} \bar{G}_j$, $i = 1, \ldots, r$, $i \neq j_1$. If there is a star among them, say $G_{j_2}^{**}$, then reset $\bar{G}_{j_2}$ to the complete star of $G$ containing $G_{j_2}^{**}$. Continue this way as long as possible. Finally, arrive at the point where the graphs $\bar{G}_{j_1}, \ldots, \bar{G}_{j_{r_0}}$ are complete stars of $G$, and there are no stars among the graphs $G_i^{**}$, $i \in R = \{1, 2, \ldots, r\} \setminus \{j_1, j_2, \ldots, j_{r_0}\}$. Then reset $\bar{G}_i$ to the smallest complete bipartite graph containing $G_i^{**}$, $i \in R$. After renumbering of vertices we may assume that the vertices $1, 2, \ldots, r_0$ are the centres of the stars $\bar{G}_{j_1}, \ldots, \bar{G}_{j_{r_0}}$. Observe that the graphs $\bar{G}_i$, $i \in R$, form a strong $\mathscr{B}$-covering of the subgraph of $G$ induced by the set $\{r_0 + 1, \ldots, n\}$. Since $r - r_0 < n - r_0$, this forces $n - r_0 < t$ which, in turn, implies that $r \geqslant r_0 > n - t$. □

**Lemma.** *There exists a constant $c > 0$ and, for sufficiently large $n$, an $n$-vertex graph $G$ such that, setting $t = c\log n$, no subgraph of $G$ on $v \geqslant t$ vertices admits a strong $\mathscr{B}$-covering consisting of less than $v$ subgraphs.*

**Proof.** Let $\mathbf{G}(n, 1/5)$ be the random graph on vertices $1, 2, \ldots, n$, where each edge appears independently with probability $1/5$. Our goal is to show that the probability that there exists a subgraph of $\mathbf{G}(n, 1/5)$ on $v \geqslant t$ vertices which admits a strong $\mathscr{B}$-covering consisting of less than $v$ complete bipartite subgraphs of $\mathbf{G}(n, 1/5)$ is less than 1. For a fixed $t \leqslant v \leqslant n$, let $\mathscr{E}_0$ be the event that the random graph $\mathbf{G} = \mathbf{G}(v, 1/5)$ admits a strong $\mathscr{B}$-covering with less than $v$ subgraphs and let $P_v = P(\mathscr{E}_0)$. From now on we focus on the

https://doi.org/10.1017/S0963548397003064 Published online by Cambridge University Press

random graph **G** and shall show that

$$\sum_{v=t}^{n} \binom{n}{v} P_v < 1.$$

Call a complete bipartite graph *good* if it is not a star, has at least $v/12$ edges and at least one of the partition sets is of order at least $v/1200$. Note that a minimal good graph has at least two but at most 100 vertices in the smaller partition set and less than $v/12 + 100$ edges. Let $\tau(G)$ be the size of the largest family of minimal good subgraphs of $G$ with mutually disjoint smaller partition sets. Let $\mathscr{E}_1$ be the event that $\tau(\mathbf{G}) < v/10^5$. Assume that the event $\mathscr{E}_1$ holds true (or, in other words, take $G \in \mathscr{E}_1$). Then there exists in $G$ an induced subgraph $H$ on at least $v - 100v/10^5 = 0.999v$ vertices and such that every complete bipartite subgraph of $H$ with at least $v/12$ edges and which is not a star, has both partition sets of order less than $v/1200$. Suppose that $G$ also admits a strong $\mathscr{B}$-covering $F_1, \ldots, F_r$, $r < v$, i.e. that $G \in \mathscr{E}_0$ (and hence $G \in \mathscr{E}_0 \cap \mathscr{E}_1$).

Let $\mathscr{E}_2$ be the event that every induced subgraph of **G** on $s \geqslant 0.999v$ vertices has more than $(1/5 - 0.001)\binom{s}{2}$ edges. Assume finally that the event $\mathscr{E}_2$ holds true (i.e. $G \in \mathscr{E}_0 \cap \mathscr{E}_1 \cap \mathscr{E}_2$). Then at least $(1/5 - 0.001)\binom{0.999v}{2} - v^2/12 \geqslant 0.01v^2$ edges of $H$ are covered by those graphs $F_i$ which have each at least $v/12$ edges. By the definition of $H$, each of these graphs $F_i$ has (in $H$) both partition sets smaller than $v/1200$. We conclude that in $G$ there are less than $v$ complete bipartite subgraphs, each with both partition sets smaller than $v/1200$, which together have at least $0.01v^2$ edges. Let us denote this event by $\mathscr{E}_3$. Summarizing, we have that $\mathscr{E}_0 \cap \mathscr{E}_1 \cap \mathscr{E}_2 \subseteq \mathscr{E}_3$ and hence

$$P_v \leqslant P(\neg\mathscr{E}_1) + P(\neg\mathscr{E}_2) + P(\mathscr{E}_3).$$

Now we shall estimate the three above summands.

If $\tau(\mathbf{G}) \geqslant v/10^5$ then there are at least $v/10^5$ edge-disjoint complete subgraphs, each with one partition set of order $k$, for some $2 \leqslant k \leqslant 100$, and the other partition set of order at least $v/12k$ (the smaller partition sets mutually disjoint), to ensure that the number of edges is at least $v/12$. If at least $v/(2 \times 10^5)$ of them have $k = 2$, certainly there are as many disjoint pairs of vertices, each pair having at least $v/24$ common neighbours, in particular, having at least $v/24 - v/10^5 > (v - v/10^5)(1/5)^2$ common neighbours outside the set of all these pairs (this makes the events independent). Let $X$ be the binomially distributed random variable with parameters $v - v/10^5$ and $1/25$. Then, by Chernoff's inequality, the probability that this happens can be bounded from above by

$$\binom{v}{2, \ldots, 2, v - v/10^5} \left(P\left(X \geqslant v/24 - v/10^5\right)\right)^{v/(2\times 10^5)} < v^{v/10^5} \left(e^{-c_1 v}\right)^{v/(2\times 10^5)} < e^{-c_2 v^2}.$$

If less then $v/(2 \times 10^5)$ complete subgraphs in question have $k = 2$, then the probability of this event can be bounded from above by

$$\left(\sum_{3\leqslant k\leqslant 100} \binom{v}{k}\binom{v}{v/12k} 5^{-v/12}\right)^{v/(2\times 10^5)} < c_3^{v^2},$$

https://doi.org/10.1017/S0963548397003064 Published online by Cambridge University Press

for some $0 < c_3 < 1$, as $\binom{v}{v/12k}5^{-v/12} \sim c_4^v$ for some $0 < c_4 < 1$ and $k \geqslant 3$. Hence,

$$P(\neg \mathscr{E}_1) < d_1^{v^2},$$

for some $0 < d_1 < 1$.

By Chernoff's inequality, setting $Y$ for a random variable with the binomial distribution with parameters $\binom{s}{2}$ and $1/5$, we have

$$P(\neg \mathscr{E}_2) < \sum_{s=0.999v}^{v} \binom{n}{s} P\left(Y < (1/5 - 0.001)\binom{s}{2}\right) < e^{-d_2 v^2},$$

for some $d_2 > 0$. This follows for $t = c \log n$ large enough (say, for $c > 3 \times 10^7$).

Finally, we focus on $\mathscr{E}_3$.

$$P(\mathscr{E}_3) \leqslant \binom{v}{v/1200}^{2v} \left(\frac{1}{5}\right)^{v^2/100} < d_3^{v^2},$$

for some $0 < d_3 < 1$.

Altogether, $P_v < d^{v^2}$ for some $0 < d < 1$, and thus, recalling that $t = c \log n$,

$$\sum_{v=t}^{n} \binom{n}{v} P_v < \sum_{v=t}^{n} (nd^v)^v < n\left(nd^t\right)^t = o(1)$$

if $c > 1/\log d$. This completes the proof of the lemma. □

The present argument does not seem to give a realistic constant $c$. We strongly believe that Tuza's upper bound $\rho(n) \leqslant n - \lfloor \log_2 n \rfloor + 1$ is much closer to the truth. It is natural to ask (as Tuza did [2, Problem 3]) if the limit of $\frac{n-\rho(n)}{\log n}$ exists and if so what is the limit.

## References

[1] Chung, F. R. K. (1980) On the coverings of graphs. *Discrete Mathematics* **30** 89–93.
[2] Tuza, Zs. (1984) Coverings of graphs by complete bipartite subgraphs; complexity of 0-1 matrices. *Combinatorica* **4**(1) 111–116.

https://doi.org/10.1017/S0963548397003064 Published online by Cambridge University Press