What fraction of the vertex set of a graph can be covered by k neighborhoods?

János Pach

Submitted: May 27, 2026; Accepted: Nov 3, 2026; Published: Nov 20, 2026
© The author. Released under the CC BY-ND license (International 4.0).

Abstract

Let $k \geqslant 2$ be a fixed integer, $0 < c < 1$ a constant. Consider a graph G with n vertices and average degree cn. We answer a question of Simon Griffiths by showing that G has k vertices such that their neighborhoods together cover at least $\min\big(1-(1-c)^k, \sqrt{c}\big)n$ vertices. This result is essentially tight.

Mathematics Subject Classifications: 05C35, 05C69, 05D40

1 Introduction

Let G be a fixed graph with vertex set V and edge set E, where $n = |V|$. For any vertex $u \in V$, let $N(u)$ denote the *neighborhood* of u in G, that is $N(u) = \{x \in V : ux \in E\}$. The size of $N(u)$ is the *degree* of u, and is denoted by $d(u)$. Let $\Delta(G) = \Delta$ and $d(G) = d$ stand for the *maximum* degree and the *average* degree of the vertices of G. Clearly, we have $\Delta \geqslant d$.

Let $k \geqslant 2$ be a fixed integer, and assume that the average degree of the vertices is

$$d(G) = \sum_{u \in V} d(u)/n = cn.$$

We want to estimate the largest number $h(G)$ such that G has k vertices, $u_0, u_1, \dots u_{k-1}$ such that $|N(u_0) \cup \dots \cup N(u_{k-1})| \geqslant h(G)$.

We start with two simple examples. Consider a random graph $\mathbf{G}(n, c)$ on n vertices with edge probability c $(0 < c < 1)$. For a given set of k vertices, $u_0, \dots, u_{k-1}$, the probability that a fixed vertex different from them is not connected to any of them is $(1-c)^k$, Thus, the expected size

$$\mathbb{E}\left(\left|\bigcup_{i=0}^{k-1} N(u_i) \setminus \{u_0, \dots, u_{k-1}\}\right|\right) = (n-k)(1-(1-c)^k) = (1-(1-c)^k + o(1))n.$$

Alfréd Rényi Institute of Mathematics, Budapest, Hungary (pach@renyi.hu).

 https://doi.org/10.37236/15672

It easily follows by Chernoff's bound [1, 3] that, as $n \to \infty$, almost surely the size of the union of the neighborhoods of every k-tuple of vertices will be of the form $(1-(1-c)^k + o(1))n$.

Our second example is a graph G' that consists of a complete graph with $\lceil \sqrt{c}\, n \rceil + 1$ vertices and its remaining vertices are isolated. Clearly, the average degree of G' satisfies $d(G') \geqslant cn$, but the size of the union of any collection of neighborhoods in G' is at most $\lceil \sqrt{c}\, n \rceil + 1$.

At the problem session of the 60th birthday conference of Bruce Reed (Oxford, July 2022), Simon Griffiths conjectured that for every c, $0 < c < 1$, one of the above two graphs is asymptotically optimal, that is, the following theorem is true.

Theorem 1. *Let $k \geqslant 2$, and let $G = (V, E)$ be a graph on n vertices with average degree cn, where $0 \leqslant c \leqslant 1$. Then there exist vertices $u_0, \ldots, u_{k-1} \in V$ such that*

$$\big|N(u_0) \cup \cdots \cup N(u_{k-1})\big| \geqslant \min\big\{(1-(1-c)^k)n,\, \sqrt{c}\, n\big\}.$$

The problem was also featured as Questions 1 and 2 in the seminal paper of José D. Alvarado, Leonardo Gonçalves de Oliveira, and Simon Griffiths [2] (page 842).

In the next section, we prove the above result. Note that formally the theorem is true for $k = 1$, too, and also for $c = 0$ or $c = 1$, but these cases are trivial (in fact, the latter never happens).

Before turning to the proof, we remark that an easy averaging argument shows that the theorem is asymptotically true for $k = 2$, provided that G is *regular*, that is $d(u) = cn$ for every $u \in V$. Let H denote the complement of G, and let $N_H(w)$ stand for the neighborhood of w in H. Using the fact that $|N(u_1) \cup N(u_2)| \geqslant n - 2 - |N_H(u_1) \cap N_H(u_2)|$, it is sufficient to show that there exist $u_1, u_2 \in V$ such that

$$|N_H(u_1) \cap N_H(u_2)| \leqslant ((1-c)^2 + o(1))n. \tag{1}$$

To see this, estimate the sum of $|N_H(u_1) \cap N_H(u_2)|$ over all the $\binom{n}{2}$ unordered pairs of distinct vertices $\{u_1, u_2\}$. We have

$$\sum_{\{u_1,u_2\}} |N_H(u_1) \cap N_H(u_2)| = \sum_{v \in V} \binom{d_H(v)}{2} = n\binom{n-1-cn}{2}.$$

Dividing the right-hand side by $\binom{n}{2}$, inequality (1) follows.

After this initial success, we expected that, with a little extra effort, the argument could be extended to all degree sequences. We explored several plausible greedy strategies to find k neighborhoods that collectively maximize vertex coverage; however, their outcomes proved resistant to analysis. Ultimately, we returned to the averaging approach, albeit with a small tweak: we started the procedure with a vertex of maximum degree. Miraculously, this slight modification made the underlying calculations tractable. The devil resides in the details of the calculus!

2 Proof of Theorem 1

Choose a vertex $u_0 \in V$ of *maximum* degree Δ, and denote its neighborhood $N(u_0)$ by S. Write

$$d(u_0) = \Delta = sn, \qquad S := N(u_0), \qquad T := V \setminus S.$$

Since Δ is at least the average degree $d(G)$, we have $s \geqslant c$.

If $s \geqslant \sqrt{c}$, then already

$$|N(u_0)| = \Delta = sn \geqslant \sqrt{c}\, n,$$

and we are done.

Thus, we may assume throughout that

$$c \leqslant s < \sqrt{c}.$$

Since every vertex of S has degree at most Δ, we have

$$\sum_{u \in S} d(u) \leqslant |S|\, \Delta = \Delta^2 = s^2 n^2.$$

As the total degree sum is cn^2, it follows that

$$\sum_{u \in V(G)} |N(u) \cap T| = \sum_{u \in T} d(u) = cn^2 - \sum_{u \in S} d(u) \geqslant (c - s^2)n^2. \tag{2}$$

Choose $u_1, \ldots, u_{k-1} \in V(G)$ independently and uniformly at random. For any $u \in T$, the probability that u lies in at least one of the neighborhoods $N(u_1), \ldots, N(u_{k-1})$ is

$$1 - \left(1 - \frac{d(u)}{n}\right)^{k-1}.$$

Notice that the function

$$f(x) := 1 - (1 - x)^{k-1} \qquad (0 \leqslant x \leqslant 1) \tag{3}$$

is nonnegative, concave on the interval $[0, 1]$, and we have $f(0) = 0$, $f(1) = 1$.

Since $d(u)/n \leqslant \Delta/n = s \leqslant \sqrt{c} \leqslant 1$, by concavity we obtain

$$f\left(\frac{d(u)}{n}\right) \geqslant \frac{d(u)/n}{s} f(s).$$

Therefore, in view of (2), the expected number of vertices of T covered by $N(u_1) \cup \cdots \cup N(u_{k-1})$ is at least

$$\frac{f(s)}{s} \cdot \frac{1}{n} \sum_{u \in T} d(u) \geqslant \frac{1 - (1 - s)^{k-1}}{s} (c - s^2)n.$$

Hence, for a suitable choice of (not necessarily distinct) $u_1, \dots, u_{k-1} \in V$, we have

$$\big|(N(u_1) \cup \cdots \cup N(u_{k-1})) \cap T\big| \geqslant \frac{1-(1-s)^{k-1}}{s}\,(c-s^2)n.$$

Adding the neighborhood $S = N(u_0)$, we conclude that

$$\big|N(u_0) \cup N(u_1) \cup \cdots \cup N(u_{k-1})\big| \geqslant \left(s + \frac{1-(1-s)^{k-1}}{s}(c-s^2)\right) n.$$

Thus, it remains to prove that for $c \leqslant s \leqslant \sqrt{c}$, the last expression is at least $\min\{1-(1-c)^k, \sqrt{c}\}$. This is the content of the next lemma.

Let $k \geqslant 2$. To simplify the notation, set

$$F(s) := s + \frac{1-(1-s)^{k-1}}{s}(c-s^2), \qquad G(c) := 1-(1-c)^k.$$

Lemma 2. *Let $k \geqslant 2$ be fixed. For every $0 < c < 1$ and every $s \in [c, \sqrt{c}]$, we have*

$$F(s) \geqslant \min\{G(c), \sqrt{c}\}.$$

Proof. Set $x := 1-s$ and define the function $H(x) := F(1-x)$. Then we have

$$\begin{aligned} F(s) &= s(1-s)^{k-1} + c\frac{1-(1-s)^{k-1}}{s} \\ &= (1-x)x^{k-1} + c\sum_{i=0}^{k-2} x^i = H(x). \end{aligned}$$

Moreover, $s \in [c, \sqrt{c}]$ corresponds to

$$x \in [1-\sqrt{c}, 1-c] \subset (0,1).$$

Taking derivatives, we obtain

$$H'(x) = c\sum_{i=1}^{k-2} i\,x^{i-1} + (k-1)x^{k-2} - kx^{k-1},$$

where the sum is empty when $k = 2$. The leading coefficient of H' is negative, and all its other coefficients are positive. Thus, its coefficients have exactly one sign change. By Descartes' rule of signs, H' has at most one positive root.

On the other hand,

$$H'(0) = \begin{cases} 1 & \text{if } k = 2, \\ c & \text{if } k \geqslant 3. \end{cases}$$

Since

$$\lim_{x\to\infty} H'(x) = -\infty,$$

by continuity, H' has precisely one positive root x_0. Therefore, $H'(x) > 0$ for $0 < x < x_0$, and $H'(x) < 0$ for $x > x_0$. Consequently, H is increasing up to x_0 and decreasing thereafter. Its minimum on $[1-\sqrt{c}, 1-c]$ must occur at an endpoint.

Finally,

$$F(\sqrt{c}) = \sqrt{c}$$

and

$$F(c) = c + (1-c)\big(1-(1-c)^{k-1}\big) = 1-(1-c)^k = G(c).$$

Thus, for every $s \in [c, \sqrt{c}]$, we have

$$\begin{aligned} F(s) = H(1-s) &\geqslant \min\{H(1-\sqrt{c}), H(1-c)\} \\ &= \min\{F(\sqrt{c}), F(c)\} = \min\{\sqrt{c}, G(c)\}, \end{aligned}$$

as required. This completes the proof of Lemma 2 and, hence, the theorem. □

Remark 3. The validity and the asymptotic tightness of Theorem 1 do not require k and c to be fixed. In particular, if $c \in (0,1)$ is fixed and

$$\frac{\log(1-\sqrt{c})}{\log(1-c)} \leqslant k = k(n) \leqslant n,$$

then $\min\big\{(1-(1-c)^k)n, \sqrt{c}\,n\big\} = \sqrt{c}\,n$ and the theorem is asymptotically tight, as $n \to \infty$. On the other hand, if k is fixed and $c = c(n) \to 0$, then $\min\big\{(1-(1-c)^k)n, \sqrt{c}\,n\big\} = (k+o(1))cn$, and the theorem is again asymptotically tight, as long as the average degree $c(n)n$ tends to infinity.

Acknowledgements

János Pach was supported by ERC Advanced Grant 'GeoScape' No. 882971 and by the National Research, Development and Innovation Office, NKFIH, K-131529. He also acknowledges the hospitality of Academia Sinica, Taipei and the Matrix Research Institute, Creswick, where this note was written. He is grateful to Péter Frankl, Simon Griffiths, Bruce Reed, and Gábor Tardos for their valuable comments. He also wants to thank an anonymous referee for suggesting a shortcut in the proof of Lemma 2.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, 4th ed., Wiley, 2016.

[2] J. D. Alvarado, L. Gonçalves de Oliveira, and S. Griffiths, Moderate deviations of triangle counts in sparse Erdős–Rényi random graphs $G(n,m)$ and $G(n,p)$, *Probab. Theory Related Fields* 191 (2025), 779–851. doi:10.1007/s00440-024-01346-3.

[3] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley, 2000.