On the Number of Monochromatic C_4 Subgraphs in a Randomly 2-colored K_n

Kelin Zhu

November 2023

Abstract

In this paper, we study the occurrences of monochromatic 4-cycles in 2-colored complete graphs. We will find explicit formulas for the average and standard deviation of the number of monochromatic 4-cycle subgraphs in randomly and uniformly 2-colored complete graphs. Then, we also introduce a theorem that establishes the minimum number of monochromatic 4-cycle subgraphs in any 2-coloring of a complete graph, which can be readily generalized. Lastly, we introduce a fast algorithm for counting the number of monochromatic 4-cycles subgraphs of a 2-colored graph. Our study has applications to future investigation of monochromatic cycle subgraphs in colored complete graphs.

Introduction

Ramsey Theory is the study of order arising from large structures in mathematics. It has many applications in the study of "graph" structures involving vertices and edges joining pairs of vertices. Originating from the 20th century, Ramsey Theory has allowed us to make statements of the form "any 2-coloring of the complete graph on x vertices (referred to as a K_x) has either a monochromatic Y subgraph of color 1 or monochromatic Z subgraph of color 2" for any pair of graphs Y, Z. The most studied case of this theorem is when Y and Z are complete graphs. We will focus our attention on the case where $Y = Z = C_4$, the 4-cycle. It is known that any 2-coloring of a K_6 contains a monochromatic 4-cycle subgraph [1]; we will strengthen this result to prove that any 2-coloring of a K_n contains at least $\frac{n(n+1)(n+2)(n+3)}{1512}$ monochromatic 4-cycle subgraphs for $n \geq 6$. Our method of proof can be generalized to compute the minimum number of monochromatic X subgraphs in any c-coloring of a K_n for any graph X and any positive integer c. We will also prove exact formulas for the mean and standard deviation of the number of 4-cycle subgraphs in random 2-colorings of K_n, and introduce an algorithm to check the number of monochromatic 4-cycles in 2-colored graphs in $O(n^3)$. Our study has applications to future studies in monochromatic cycles in graph colorings, and lead to a greater understanding of graph theory in general.

Mathematical Results

We present the following two theorems alongside their proofs.

Theorem 1: when each edges of a K_n is randomly colored between red and blue with equal probability (in short, the K_n is "randomly colored"), the expected number of monochromatic 4-cycles is $\frac{3}{8}\binom{n}{4}$.

Theorem 2: the SD of the number of monochromatic 4-cycles in a randomly colored K_n is $\sqrt{\frac{n^5}{128} - \frac{31n^4}{512} + \frac{43n^3}{256} - \frac{101n^2}{512} + \frac{21n}{256}}$.

Proof of Theorem 1: For each 4-cycle, there are $\binom{n}{4}$ ways to choose its 4 vertices; let x be one vertex, then, there are 3 ways to choose which of the other 3 vertices is not adjacent to x. Hence, there are $3\binom{n}{4}$ distinct 4-cycles in a K_n. Each 4-cycle has 16 equally probable colorings, two of which are monochromatic, hence the probability that any single 4-cycle in the coloring is monochromatic is $\frac{1}{8}$. By the Linearity of Expectation, the expected number of monochromatic 4-cycles is $3\binom{n}{4} \cdot \frac{1}{8}$. ■

Proof of Theorem 2: Let X be the random variable determined by the number of monochromatic 4-cycles of a randomly colored K_n. Then, $\text{SD}(X) = \sqrt{\text{Var}(X)} = \sqrt{E[X^2] - E[X]^2}$.

Note that $E[X^2] = \sum_{C_1 \in S} \sum_{C_2 \in S} E[C_1, C_2 \text{ are both monochromatic}]$, where S is the set of all 4-cycle subgraphs. We will compute the sum with casework on the number of shared edges between the pair (C_1, C_2).

4 shared edges: This means that $C_1 = C_2$, then $E[C_1, C_2 \text{ are both monochromatic}] = \frac{1}{8}$, and there are $3\binom{n}{4}$ such pairs.

3 shared edges: This case does not exist. Any two 4-cycles sharing 3 edges will also share the fourth edge since the 4 vertices in the cycle and their order are uniquely determined by the 3 edges.

2 shared edges: The probability that all $4 + 4 - 2 = 6$ edges in $C_1 \cup C_2$ are red is $\frac{1}{64}$. Similarly, the probability that all edges in $C_1 \cup C_2$ are blue is $\frac{1}{64}$, hence $E[C_1, C_2 \text{ are both monochromatic}] = \frac{1}{32}$. If the shared edges are not adjacent in C_1, then the two cycles share all 4 vertices. There are 2 ways to choose C_2 from the set of vertices of C_1 so that it is distinct from C_1, and $3\binom{n}{4}$ choices of C_1, hence $2 \cdot 3\binom{n}{4} = 6\binom{n}{4}$ pairs of this form exist. If the shared edges are adjacent in C_1, then the two cycles share exactly the 3 vertices covered by the shared edges. From the set of all vertices, we can choose a set of 5 vertices in $C_1 \cup C_2$ in $\binom{n}{5}$ ways, and from those 5 vertices, we can choose the vertices of C_1 in $\binom{5}{4}$ ways, determine C_1 from the set of its vertices in 3 ways, determine the pair of adjacent edges contained in $C_1 \cap C_2$ in 4 ways. Then, C_2 is determined by connecting the vertex in $(C_1 \cup C_2) \setminus C_1$ to the degree 1 endpoints of the edges in $C_1 \cap C_2$. Hence, $\binom{n}{5} \cdot \binom{5}{4} \cdot 3 \cdot 4 = 60\binom{n}{5}$ pairs of this form exist. In total, there are $6\binom{n}{4} + 60\binom{n}{5}$ pairs in this case.

1 or 0 shared edges: In the case of 1 shared edge, the probability that all $4{+}4{-}1 = 7$ edges in $C_1 \cup C_2$ are red is $\frac{1}{128}$, and the probability that all edges in $C_1 \cup C_2$ are blue is $\frac{1}{128}$, hence $E[C_1, C_2 \text{ are both monochromatic}] = \frac{1}{64}$. In the case of 0 shared edges, the events "C_1 is monochromatic" and "C_2 is monochromatic" are independent and each occur with probability $\frac{1}{8}$, and $E[C_1, C_2 \text{ are both monochromatic}] = \frac{1}{64}$ also holds. Hence, $E[C_1, C_2 \text{ are both monochromatic}] = \frac{1}{64}$ whenever C_1 and C_2 do not share at least 2 edges.

Therefore, we have that

$$E[X^2] - E[X]^2 = \frac{3\binom{n}{4}}{8} + \frac{6\binom{n}{4} + 60\binom{n}{5}}{32} + \frac{\left(3\binom{n}{4}\right)^2 - \left(3\binom{n}{4} + 6\binom{n}{4} + 60\binom{n}{5}\right)}{64} - \left(\frac{3}{8}\binom{n}{4}\right)^2$$

which expands to $\frac{n^5}{128} - \frac{31n^4}{512} + \frac{43n^3}{256} - \frac{101n^2}{512} + \frac{21n}{256}$. ■

Computer-assisted Results

We introduce the following algorithm for counting the number of monochromatic 4-cycle subgraphs of 2-colored graphs.

Theorem 3: Label the vertices of a 2-colored n-vertex graph G as vertices 1 through n. For any $1 \le x < y \le n, 1 \le i \le 2$, let $f(x, y, i)$ be the number of vertices z such that $z \ne x, y$ and the edges xz, yz both have color i. Then, the number of monochromatic 4-cycle subgraphs of G is equal to $\frac{1}{2} \sum_{1 \le x < y \le n, 1 \le i \le 2} \binom{f(x,y,i)}{2}$

Proof of Theorem 3: For fixed values of x, y, i, $\binom{f(x,y,i)}{2}$ counts the number of pairs (w, z) such that x, y, w, z are distinct, and xz, yz, xw, yw have color i. Hence, $\binom{f(x,y,1)}{2} + \binom{f(x,y,2)}{2}$ counts the number of monochromatic 4-cycles subgraphs of G having xy as a diagonal. Then, the expression $w = \sum_{1 \le x < y \le n, 1 \le i \le 2} \binom{f(x,y,i)}{2}$ counts the number of pairs (C, d), where C is a monochromatic 4-cycle subgraph of G and d is a diagonal of C. Each monochromatic 4-cycle has 2 diagonals, so $\frac{w}{2}$ counts the number of monochromatic 4-cycle subgraphs of G. ■

We can find $f(x, y, i)$ for all $1 \le x < y \le n, 1 \le i \le 2$ in $O(n^3)$ by brute-forcing all values of $z \ne x, y$, checking the colors of edges xz, yz using an adjacency list, and setting $f(x, z, i) += 1$ if both edges have color i. We perform $O(n)$ operations and use $O(1)$ memory for each pair $1 \le x < y \le n$, and there are $O(n^2)$ such pairs. We also use $O(n^2)$ memory and time to input and store the adjacency list of G. We can compute the expression $\frac{1}{2} \sum_{1 \le x < y \le n, 1 \le i \le 2} \binom{f(x,y,i)}{2}$ in $O(n^2)$. Hence, the total time complexity of an algorithm based on theorem 3 would be $O(n^3)$, and the total memory complexity would be $O(n^2)$.

Using our algorithm, we are able to establish the following result:

Lemma 4: All 2-colorings of a K_6 contain at least 2 monochromatic 4-cycle subgraphs.

Proof of Lemma 4: Use exhaustive search through all $2^{\binom{6}{2}}$ 2-colorings of a K_6, as per [2]. ■

Finally, we will present our second main theorem of this paper.

Theorem 5: For all $n \geq 6$, every 2-coloring of K_n has at least $f(n) = \frac{n(n+1)(n+2)(n+3)}{1512}$ monochromatic 4-cycles subgraphs.

Proof of Theorem 5: We proceed with induction, where the base case $n = 6$ is proven by our algorithm.

For the inductive step, let H be a 2-colored K_{n+1}. consider any n vertex induced subgraph G of H. We know G has at least $f(n)$ monochromatic 4-cycle subgraphs by the inductive hypothesis. Hence, there are at least $4f(n)$ pairs of the form (v, C), where v is any vertex in G and C is a monochromatic 4-cycle containing v that's a subgraph of G. Therefore, by the pigeonhole principle, there is a vertex u in G such that there are at least $\frac{4}{n}f(n)$ pairs of the form (u, D), where D is a monochromatic 4-cycle subgraph of G containing u. By the inductive hypothesis, there are at least $f(n)$ monochromatic 4-cycles subgraphs of $H \setminus u$. Hence, we have that $f(n+1) \geq (1 + \frac{4}{n})f(n)$, and the inductive step is complete. ■

Discussion of Results

An immediate corollary of Theorems 1 and 2 by Chebyshev's inequality is that: for any positive real number k, if a random 2-coloring of K_n has x monochromatic 4-cycle subgraphs, with probability at least $1 - \frac{1}{k^2}$ we have

$$\frac{3}{8}\binom{n}{4} - k\sqrt{\frac{n^5}{128} - \frac{31n^4}{512} + \frac{43n^3}{256} - \frac{101n^2}{512} + \frac{21n}{256}} < x < \frac{3}{8}\binom{n}{4} + k\sqrt{\frac{n^5}{128} - \frac{31n^4}{512} + \frac{43n^3}{256} - \frac{101n^2}{512} + \frac{21n}{256}}$$

Also, theorem 5 can be generalized to the following form using our proof method.

Theorem 5 (Generalized) Let L be any graph. Let l be the number of vertices in L. Let $c, n_0, m_0 \in \mathbb{N}$. Assume that for all c-colorings of K_{n_0} there are at least m_0 monochromatic L subgraphs. Then, for all $n \geq n_0$, all c-colorings of K_n contain at least $f(n) = \frac{m_0 n(n+1)\cdots(n+h-1)}{n_0\cdots(n_0+h-1)}$ monochromatic L subgraphs.

Proof of Theorem 5 (Generalized) We proceed with induction, where the base case $n = n_0$ is given.

For the inductive step, let H be a 2-colored K_{n+1}. consider any n vertex induced subgraph G of H. We know G has at least $f(n)$ monochromatic L subgraphs by the inductive hypothesis. Hence, there are at least $l \cdot f(n)$ pairs of the form (v, C), where v is any vertex in G and C is a monochromatic L containing v that's a subgraph of G. Therefore, by the pigeonhole principle, there is a vertex u in G such that there are at least $\frac{l}{n}f(n)$ pairs of the form (u, D), where D is a monochromatic L subgraph of G containing u. By the inductive hypothesis, there are at least $f(n)$ monochromatic L subgraphs of $H \setminus u$. Hence, we have that $f(n+1) \geq (1 + \frac{l}{n})f(n)$, and the inductive step is complete. ■

Acknowledgements

I thank Dr. William Gasarch for his mentorship of my project. I thank the University of Maryland, Department of Computer Science for hosting the preparatory reading of my project. I thank Ms. Angelique Bosse for her continued guidance of my project, and I thank my peers for their suggestions to my project.

References

[1] Stanisław Radziszowski. "Small Ramsey Numbers". In: *The Electronic Journal of Combinatorics* (2021). DOI: https://doi.org/10.37236/21.

[2] Kelin Zhu. *2-Colored Complete Graph Simulation.* URL: https://ide.usaco.guide/Nct5ipmFWfc6xNVx0Ct.