NDMI011: Combinatorics and Graph Theory 1

Lecture #12
Bounding the number of edges in graphs without certain subgraphs

Irena Penev

1 Graphs without K_3 as a subgraph

A graph is said to be *triangle-free* if it does not contain K_3 as a subgraph. Equivalently, a graph G is triangle-free if $\omega(G) \leq 2$.

The following theorem is a special case of "Turán's theorem."

Theorem 1.1. *Let n be a positive integer. Then*

(a) any triangle-free graph on n vertices has at most $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil = \lfloor \frac{n^2}{4} \rfloor$ edges;

(b) there exists a triangle-free graph on n vertices that has precisely $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil = \lfloor \frac{n^2}{4} \rfloor$ edges.

Proof. First, let us check that $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil = \lfloor \frac{n^2}{4} \rfloor$. If n is even, then this is obvious. If n is odd, then there exists a non-negative integer k such that $n = 2k+1$, and we compute

$$\lfloor \tfrac{n}{2} \rfloor \lceil \tfrac{n}{2} \rceil \quad = \quad \lfloor \tfrac{2k+1}{2} \rfloor \lceil \tfrac{2k+1}{2} \rceil \quad = \quad k(k+1) \quad = \quad k^2 + k$$

and

$$\lfloor \tfrac{n^2}{4} \rfloor \quad = \quad \lfloor \tfrac{(2k+1)^2}{4} \rfloor \quad = \quad \lfloor \tfrac{4k^2+4k+1}{4} \rfloor \quad = \quad k^2 + k,$$

and we deduce that $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil = \lfloor \frac{n^2}{4} \rfloor$.

For (b), we observe that the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is triangle-free[1] and has precisely n vertices and $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil$ edges.

It remains to prove (a). We assume inductively that the claim holds for graphs on fewer than n vertices, i.e. that for all positive integers $\widetilde{n} < n$, any triangle-free graph on $\widetilde{n}$ vertices has at most $\lfloor \frac{\widetilde{n}^2}{4} \rfloor$ edges. It is clear that the

[1] Indeed, all bipartite graphs are triangle free.

theorem holds for $n = 1$ and $n = 2$. So, we assume that $n \geq 3$, we fix a triangle-free graph G on n vertices, and we show that G has at most $\lfloor \frac{n^2}{4} \rfloor$ edges. If G has no edges, then this is obvious. So assume that G has at least one edge, say uv. Then $G \setminus \{u, v\}$ is triangle-free and has $n - 2$ vertices, and so by the induction hypothesis, it has at most $\lfloor \frac{(n-2)^2}{4} \rfloor$ edges. Further, since G is triangle-free, a vertex in $V(G) \setminus \{u, v\}$ can be adjacent to at most one of u, v, and so the number of edges between $\{u, v\}$ and $V(G) \setminus \{u, v\}$ is at most $|V(G) \setminus \{u, v\}| = n - 2$. Since the edges of G are precisely the edges of $G \setminus \{u, v\}$, plus the edges between $\{u, v\}$ and $V(G) \setminus \{u, v\}$, plus the edge uv, we see that

$$\begin{aligned} |E(G)| &\leq \lfloor \tfrac{(n-2)^2}{4} \rfloor + (n-2) + 1 \\ &= \lfloor \tfrac{n^2-4n+4}{4} \rfloor + n - 1 \\ &= \lfloor \tfrac{n^2}{4} \rfloor, \end{aligned}$$

which is what we needed to show. □

2 Graphs without C_4 as a subgraph

In what follows, we will use the Cauchy-Schwarz inequality (below).

The Cauchy-Schwarz inequality. *All real numbers $a_1, \dots, a_n, b_1, \dots, b_n$ satisfy*

$$\Big(\sum_{i=1}^{n} a_i b_i \Big)^2 \leq \Big(\sum_{i=1}^{n} a_i^2 \Big) \Big(\sum_{i=1}^{n} b_i^2 \Big).$$

An *isolated vertex* is a vertex that has no neighbors.

Theorem 2.1. *Let $n \in \mathbb{N}$. Any graph on n vertices that does not contain C_4 as a subgraph has at most $\frac{1}{2}(n + n^{3/2})$ edges.*

Proof. Let G be a graph on n vertices, and assume that G does not contain C_4 as a subgraph. Clearly, we may assume that G has no isolated vertices.[2] Let $d_1, \dots, d_n$ be the degrees of the vertices of G;[3] since G is has no isolated veetices, we see that $d_1, \dots, d_n \geq 1$.

Let $M = \{(v, A) \mid v \in V(G), A \in \binom{N_G(v)}{2}\}$.[4] Now, we will count the number of elements of M in two ways.

[2] Why?

[3] The d_i's are not necessarily distinct; d_i is the degree of the i-th vertex of G.

[4] In other words, M is the set of all ordered pairs $(v, \{u_1, u_2\})$ such that $v \in V(G)$, and $u_1, u_2 \in V(G)$ are two distinct neighbors of v. Note also that $(v, \{u_1, u_2\}) \in M$ if and only if u_1, v, u_2 is a (not necessarily) induced two-edge path of G. So, $|M|$ is in fact the number of (not necessarily induced) two-edge paths in G.

First, for each $v \in V(G)$, there are precisely $\binom{d_G(v)}{2}$ choices of A such that $(v, A) \in M$. So, $|M| = \sum\limits_{v \in V(G)} \binom{d_G(v)}{2} = \sum\limits_{i=1}^{n} \binom{d_i}{2}$.

We now bound $|M|$ above, as follows. Note that the second coordinate of any element of M is simply an element of $\binom{V(G)}{2}$; since $|V(G)| = n$, there are at most $\binom{n}{2}$ choices for the second coordinate of an element of M. On the other hand, since G contains no C_4 as a subgraph, we see that no two distinct elements of M have the same second coordinate. Indeed, suppose that (v_1, A) and (v_2, A) are distinct elements of M; we then set $A = \{u_1, u_2\}$, we and observe that v_1, u_1, v_2, u_2, v_1 is a (not necessarily induced) C_4 in G, a contradiction. So, $|M| \leq \binom{n}{2}$.

We now have that

$$\sum_{i=1}^{n} \binom{d_i}{2} \leq \binom{n}{2}.$$

Obviously, $\binom{n}{2} \leq \frac{n^2}{2}$, and since $d_1, \dots, d_n \geq 1$, we see that $\binom{d_i}{2} \geq \frac{(d_i-1)^2}{2}$ for all $i \in \{1, \dots, n\}$; consequently,

$$\sum_{i=1}^{n} \frac{(d_i-1)^2}{2} \leq \sum_{i=1}^{n} \binom{d_i}{2} \leq \binom{n}{2} \leq \frac{n^2}{2},$$

and it follows that

$$\sum_{i=1}^{n} (d_i - 1)^2 \leq n^2.$$

We now compute:

$$\begin{aligned}
\sum_{i=1}^{n} (d_i - 1) &= \sum_{i=1}^{n} (d_i - 1) \cdot 1 \\
&\leq \sqrt{\sum_{i=1}^{n} (d_i - 1)^2} \sqrt{\sum_{i=1}^{n} 1^2} && \text{by the Cauchy-Schwarz inequality} \\
&= \sqrt{\sum_{i=1}^{n} (d_i - 1)^2} \sqrt{n} \\
&\leq \sqrt{n^2} \sqrt{n} && \text{because } \sum_{i=1}^{n} (d_i - 1)^2 \leq n^2 \\
&= n^{3/2}.
\end{aligned}$$

It now follows that

$$|E(G)| = \frac{1}{2} \sum_{i=1}^{n} d_i = \frac{1}{2}\Big(n + \sum_{i=1}^{n} (d_i - 1)\Big) \leq \frac{1}{2}(n + n^{3/2}),$$

which is what we needed to show. □