Maximum number of 3-paths in a graph

Owen D. Byer
byer@emu.edu
Eastern Mennonite University
Harrisonburg, VA 22802

Abstract

In this paper we show that a graph $\mathbf{G}$ with $e \geq 6$ edges contains at most $h(h-1)(h-2)(h-3)/2$ paths of length three, where $h \geq 0$ satisfies $h(h-1)/2 = e$. It follows immediately that $\mathbf{G}$ contains at most $h(h-1)(h-2)(h-3)/8$ cycles of length four. For $e > 6$, the bounds will be attained if and only if h is an integer and $\mathbf{G}$ is the union of $\mathbf{K}_h$ and isolated vertices. The bounds improve those found recently by Bollobás and Sarkar.

1 Introduction

All graphs considered are finite, undirected, and simple. Given such a graph $\mathbf{G} = (V, E)$, $p_s(\mathbf{G})$ represents the number of subgraphs of $\mathbf{G}$ isomorphic to a path of length s, and $p_s(e) = max\{p_s(\mathbf{G}) : |E(\mathbf{G})| = e\}$. For vertices $v, x \in V$, we let $d(v, x)$ represent the distance between v and x, and if v and x are adjacent, we denote the edge by vx. Finally, $deg_{\mathbf{G}}(v) = deg(v)$ gives the degree of v, $Nbh(v)$ is the neighborhood of v, and for $S \subset V$, $\mathbf{G}[S]$ represents the subgraph of $\mathbf{G}$ induced by the vertices in S.

In this paper we show that if $e \geq 6$, then $p_3(e) \leq h(h-1)(h-2)(h-3)/2$, where $h = \frac{1+\sqrt{8e+1}}{2}$ satisfies $h(h-1)/2 = e$. Equality will be attained for $e > 6$ if and only if $\mathbf{G}$ contains the clique of size h as its only nontrivial component.

We first mention some related works. Clearly a graph of size e has at most $\binom{e}{2}$ paths of length two, as any two edges form a 2-path, and the star graph $\mathbf{K}_{1,e}$ is the unique extremal graph. Ahlswede and Katona [5] maximize $p_2(\mathbf{G})$ over graphs of fixed size and order, and in [4] the present author follows up by classifying the six types of extremal graphs.

ARS COMBINATORIA 61(2001), pp. 73-79

In [2], Bollobás and Sarkar give the following result:

Theorem 1 *Let* $\mathbf{G}$ *be a graph of size* e *containing no isolated vertices, with* $6 \le \binom{k}{2} \le e < \binom{k+1}{2}$. *Then* $p_3(\mathbf{G}) \le 2e(e-k)(k-2)/k$. *Equality holds if and only if either* $e = \binom{k}{2}$ *and* $\mathbf{G} \approx \mathbf{K}_k$ *or* $e = 6$ *and* $\mathbf{G}$ *is the graph obtained by joining a new vertex to two opposite vertices of a 4-cycle.*

In the same paper, Bollobás and Sarkar also give asymptotic (lower and upper) bounds for $p_s(e)$ when $e = \binom{k}{2}$ and $s \ge 4$, noting that the apparent extremal graphs are different depending on whether s is even or odd. In particular, when s is odd, $p_s(e)$ is roughly on the order of $k^{s+1}/2$, which indicates that the extremal graphs resemble the graph $\mathbf{K}_k$. When $s = 2t$ is even, $p_{2t}(e)$ is on the order of $C_t e^{t+1}$ (C_t is defined explicitly, with $\lim_{t\to\infty} C_t = \frac{1}{exp(1)t^2}$), and the extremal graphs appear to be bipartite. In [3], they show that $\mathbf{K}_{e/2,2}$ (or a slight variation if e is odd) is the unique extremal graph for $s = 4$ and e sufficiently large. Our bound on $p_3(e)$ can be used to remove the requirement that e is a binomial coefficient in their upper bound for $p_s(e)$ when s is odd, and we will discuss this more at the end of the paper.

Finally, in [1], Alon gives more general asymptotic results for the maximum number $N(e, \mathbf{H})$ of subgraphs isomorphic to a fixed graph $\mathbf{H}$ in a graph of size e. His bounds are asymptotically best possible when $\mathbf{H}$ has a spanning subgraph consisting of cycles and disjoint edges, and therefore match those found in [2] when $\mathbf{H}$ is a path of odd length.

2 Main Results

We now seek to improve the upper bound on $p_3(e)$ given in [2]. The main difference between our approach and theirs is that we first show that, except in certain instances, an extremal graph of size e can be found that does not contain vertices of degree one or two. This allows us to estimate $p_3(\mathbf{G})$ in a slightly different manner, and obtain a tighter bound. We begin with two lemmas showing how one can sometimes eliminate vertices of degree one or two without decreasing the number of paths of length three.

Lemma 2 *Let* $\mathbf{G} = (V, E)$ *be a connected graph of order* n *and size* $e \ge 4$, *and suppose* $\mathbf{G}$ *has vertices* v *and* x *such that* $Nbh(v) = \{x\}$. *If* $deg(x) \le n-2$, *then there exists a graph* $\mathbf{H} = (V, E')$ *of size* e *such that* $p_3(\mathbf{G}) \le p_3(\mathbf{H})$ *and* $deg_{\mathbf{H}}(v) = 0$.

Proof: Choose vertex $w \in V$ of maximum distance from v. Then $d(x, w) \ge 2$. Let $\mathbf{H} = (V, E')$, where $E' = E \cup \{wx\} \backslash \{vx\}$. If $d(x, w) \ge 3$, then each path of the form $vxbp$ in $\mathbf{G}$ is replaced by the path $wxbp$ in $\mathbf{H}$, so we see

$p_3(\mathbf{H}) \geq p_3(\mathbf{G})$. Now if $d(x,w) = 2$, then $\mathbf{G}$ has $k \geq 1$ additional 3-paths of the form vxa_iw, $i = 1,\ldots,k$, that are not 3-paths in $\mathbf{H}$. However, in this case $\mathbf{H}$ has $k(k-1)$ new 3-paths of the form a_iwxa_j, $i \neq j$, so we see that $p_3(\mathbf{H}) \geq p_3(\mathbf{G})$ if $k \geq 2$. If $k = 1$, then vxa_1w is the only 3-path in $\mathbf{G}$ for which we still need to find a replacement in $\mathbf{H}$. To this end, note that when $k = 1$, since $e \geq 4$ and w is a vertex of maximum distance from x, we necessarily have $deg(x) \geq 3$ or $deg(a_1) \geq 3$ ($\mathbf{G}$ cannot consist of the path vxa_1w). If $deg(x) \geq 3$, then $\{v, a_1, b\} \subseteq Nbh(x)$ for some $b \in V$, so $\mathbf{H}$ contains the new 3-path a_1wxb. Similarly, if $deg(a_1) \geq 3$, then $\mathbf{H}$ will contain the new 3-path xwa_1t for some $t \in Nbh(a_1)$. In each case, we have exhibited an injection from the 3-paths of $\mathbf{G}$ into the 3-paths of $\mathbf{H}$, so $p_3(\mathbf{G}) \leq p_3(\mathbf{H})$. ■

Lemma 3 *Let $\mathbf{G} = (V,E)$ be a connected graph of order n and size $e \geq 4$, and suppose there exist vertices v, x, and y in V such that $Nbh(v) = \{x,y\}$. Then*

a) if there exists $w \in V$ such that $d(x,w) \geq 2$ and $d(y,w) \geq 2$, then there exists a graph $\mathbf{H} = (V,E')$ of size e such that $p_3(\mathbf{G}) \leq p_3(\mathbf{H})$ and $deg_{\mathbf{H}}(v) = 0$, and

b) if $Nbh(x) \cap Nbh(y) \neq V \backslash \{x,y\}$, then either $\mathbf{G} = \mathbf{C}_5$ or there exists a graph $\mathbf{H} = (V,E')$ of size e such that $p_3(\mathbf{G}) \leq p_3(\mathbf{H})$, where $\mathbf{H}$ has a vertex v' of degree at most one.

Proof: To prove a, Let $\mathbf{H} = (V,E')$, where $E' = E \cup \{wx, wy\} \backslash \{vx, vy\}$. Note that each 3-path in $\mathbf{G}$ is either also a 3-path in $\mathbf{H}$, or, as shown in Table 1, uniquely determines a new 3-path in $\mathbf{H}$. This shows that $p_3(\mathbf{G}) \leq p_3(\mathbf{H})$, and since v is an isolated vertex in $\mathbf{H}$, proves part a.

3-paths in G	3-paths in H
$vxap$, $(p \neq w)$	$wxap$
$vyap$, $(p \neq w)$	$wyap$
$yvxa$	$ywxa$
$xvya$	$xwya$
$vxya; vyxa$, if $xy \in E$	$wxya; wyxa$
$vxaw$, if $d(x,w) = 2$	$ywax$
$vyaw$, if $d(y,w) = 2$	$xway$

Table 1: 3-path repacements in part a

To prove b, we note that by part a, we need only consider the case where there is a vertex $w \in Nbh(y)$, $w \notin Nbh(x)$ (the same argument will work if $w \in Nbh(x)$, $w \notin Nbh(y)$). Let $\mathbf{H} = (V,E')$, where $E' = E \cup \{wx\} \backslash \{vx\}$.

Table 2 shows most of the paths of length three in **G** that are not 3-paths in **H**.

3-paths in **G**	3-paths in **H**
$vxap$, $(p \neq w)$	$wxap$
$xvyt$	$xwyt$
$yvxa$	$ywxa$
$xvya$	$xwya$
$vxyw$, if $xy \in E$	$vyxw$
$xvyw$	$vywx$

Table 2: 3-path repacements in part b

The only remaining 3-paths in **G** that need "replacements" in **H** are those of the form vxa_iw, $i = 1, \ldots, k$. We have several cases to consider. If $deg_{\mathbf{G}}(w) \geq 3$, then for each i, **H** contains the new 3-paths a_ixwb_i, where $b_i \in Nbh_{\mathbf{G}}(w) \backslash \{y, a_i\}$. Therefore, we are done unless $Nbh_{\mathbf{G}}(w) = \{a_1, y\}$, which we now assume. If $deg(a_1) \geq 3$, the path vxa_1w in **G** can be replaced by the path xwa_1t, where $t \in Nbh(a_1)\backslash\{x, w\}$. If $deg(a_1) = 2$, we simply switch the roles of v and w and x and y in our current proof (i.e., w is the vertex of degree 2 with neighborhood $\{a, y\}$). Arguing as before, since x and a_1 have changed roles, we will be finished unless $deg(x) = 2$. Alas, in the final case, we are left with the graph **G** where vertices v, x, a_1, and w each have degree 2, and $\mathbf{G}[\{v, x, a, w, y\}]$ is a 5-cycle. If $deg(y) = 2$, then $\mathbf{G} \approx \mathbf{C}_5$. If $deg(y) \geq 3$, then by part a (letting x take the role of v, and realizing that any vertex in $Nbh(y)\backslash\{v, w\}$ is at distance at least two from each of v and a) there is a graph $\mathbf{H}'$ of size e with vertex set V, where $deg_{\mathbf{H}'}(x) = 0$ and $p_3(\mathbf{G}) \leq p_3(\mathbf{H}')$. We have checked all cases, and this finishes the proof of the lemma. ■

Theorem 4 *Let* **G** *be a graph of size* e, $e \geq 6$. *Then* $p_3(\mathbf{G}) \leq h(h-1)(h-2)(h-3)/2$, *where* $h \geq 0$ *satisfies* $h(h-1)/2 = e$. *For* $e > 6$, *equality will be attained if and only* **G** *contains* $\mathbf{K}_h$ *as its only nontrivial component.*

Proof: One can easily check that any graph of size 6 contains at most 12 paths of length three. Assume then that the result holds for all graphs of size e', $6 \leq e' < e$, and let **G** be a graph of size e. It clearly suffices to assume **G** is connected. Consider first the case where **G** has a vertex v of degree one, say $Nbh(v) = x$. By Lemma 2, we may assume $Nbh(x) \cup \{x\} = V$. In this case, each edge ab in $\mathbf{G}[Nbh(x)]$ lies on two 3-paths containing v (namely $vxab$ and $vxba$). Therefore, $p_3(\mathbf{G}) \leq 2|E(\mathbf{G}-x)| + p_3(\mathbf{G}-v)$, and applying the inductive hypothesis to $\mathbf{G} - v$, we get

$$\begin{aligned} p_3(\mathbf{G}) &\leq 2(e - deg_{\mathbf{G}}(x)) + h'(h'-1)(h'-2)(h'-3)/2 \\ &\leq 2(e - h') + h'(h'-1)(h'-2)(h'-3)/2, \end{aligned} \tag{1}$$

where $0 \leq h' \leq |V(\mathbf{G}-v)|$ satisfies $h'(h'-1)/2 = e-1$. One can then verify algebraically that the right-hand side of (1) is bounded (strictly) above by $h(h-1)(h-2)(h-3)/2$ by converting both expressions to functions of e.

Suppose next that $\mathbf{G}$ has a vertex v of degree two, and $Nbh(v) = \{x, y\}$. By Lemma 3, it suffices to consider the case where $Nbh(x) \cap Nbh(y) = V\backslash\{x, y\}$. Suppose that $\mathbf{G}$ has order n and $\mathbf{G}[V\backslash\{x, y\}]$ has size e'. Then $\mathbf{G}$ contains at most $4e' + 2(n-3)$ 3-paths beginning with v ($2e'$ each beginning with vx and vy but not containing xy, and potentially $2(n-3)$ more if $xy \in E$). In addition, $\mathbf{G}$ contains $n-3$ paths of length 3 of the form $xvya$ and $n-3$ of the form $yvxa$. Therefore, when $|E(\mathbf{G}-v)| \geq 6$ (a check of both graphs of size 7 with vertices v, x, and y as described shows that the result holds when $|E(G-v)| = 5$), by the inductive hypothesis we have,

$$\begin{aligned} p_3(\mathbf{G}) &\leq 4(e - 2(n-2)) + 2(n-3) + 2(n-3) + p_3(\mathbf{G}-v) \\ &\leq 4e - 4h' + h'(h'-1)(h'-2)(h'-3)/2, \end{aligned} \tag{2}$$

where $0 \leq h' \leq n-1$ satisfies $h'(h'-1) = e-2$. As before, one can verify that the right hand side of (2) is bounded (strictly) above by $h(h-1)(h-2)(h-3)/2$, as desired.

Finally, suppose each vertex of $\mathbf{G}$ has degree at least three. Using the idea found in [2], let $V = \{x_1, x_2, \ldots, x_n\}$, and for each i set

$$\begin{aligned} d_i &= deg(x_i) \\ c_i &= |E(\mathbf{G}[Nbh(x_i)])| \\ e_i &= |\{ab \ : a \in Nbh(x_i),\ d(b, x_i) = 2\}|. \end{aligned}$$

In [2], the authors then summed over all 3-paths with middle edge incident with x_i. This sum appears in the first line below, except theirs ran over only those vertices of degree at least two. They then needed to consider various parts of the sum separately, depending on the actual value of d_i. However, since we have $d_i \geq 3$ for each i, we can proceed as follows:

$$\begin{aligned} p_3(\mathbf{G}) &= \frac{1}{2}\sum_{i=1}^{n}((d_i-1)e_i + 2(d_i-2)c_i) \\ &\leq \frac{1}{2}\sum_{i=1}^{n}((d_i-1)e_i + 2(d_i-2)(e - e_i - d_i)) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n}((e+2)d_i - 2e)) - \sum_{i=1}^{n}(d_i^2 + (d_i - 3)e_i/2) \\
&\le (e+2)(2e) - 2en - \sum_{i=1}^{n}\overline{d}^2 \quad (\overline{d} \text{ is the average vertex degree}) \\
&= 2e^2 + 4e - 2en - n(2e/n)^2 \\
&\le 2e^2 + 4e - 2eh - 4e^2/h \\
&= h(h-1)(h-2)(h-3)/2.
\end{aligned}
$$

Note that equality will occur if and only if $d_i = \overline{d} = 2e/n$ for each i if and only if $\mathbf{G} \approx \mathbf{K}_h$ ($h = \overline{d} + 1 = n$ will be an integer). ■

3 Discussion of Results

The upper bound for $p_3(e)$ given in the theorem is tight when h is an integer, and lowers the bound given in [2]. Elementary calculus shows that the differences in the two bounds are unimodal as e varies between $\binom{h'}{2}$ and $\binom{h'+1}{2}$ for any integer $h' \ge 4$. In fact, the two bounds are equal at the endpoints, and have a maximum difference of $\frac{e}{\sqrt{8e+1}}$ when e is written as $h(h+\sqrt{h^2+1})/4$. Of course the number of edges and number of 3-paths must be integers, so we see that the actual maximum difference in bounds is about $\left\lfloor\sqrt{e/8}\right\rfloor$ when $e \approx h^2/2$. Another benefit of our bound is that it is more descriptive of the structure of the extremal graphs: the expression $h(h-1)(h-2)(h-3)/2$ is the natural one to use to count the number of paths of length three in the graph $\mathbf{K}_h$ with $h(h-1)/2$ edges, and we have shown that this expression also serves as an upper bound when h is not an integer. Apparently, the extremal graphs will be close in structure to a clique.

By "gluing together" an appropriate number of paths of length three and perhaps an additional edge, Bollobás and Sarkar showed in [2] that if r is a positive integer and $m = k(k-1)/2$ for some integer $k \ge 4r+4$, then

$$
\begin{aligned}
\frac{1}{2}(k)_{4r+2} &\le p_{4r+1}(m) \le 2^r m(p_3(m))_r = \frac{1}{2}k(k-1)((k)_4)_r \quad \text{and} \\
\frac{1}{2}(k)_{4r+4} &\le p_{4r+3}(m) \le 2^r(p_3(m))_{r+1} = \frac{1}{2}((k)_4)_{r+1}.
\end{aligned}
$$

We simply note here that our bound on $p_3(m)$ allows the removal of the requirement that k is an integer in their upper bounds for $p_{4r+1}(m)$ and $p_{4r+3}(m)$.

We conclude with an immediate corollary of the bound on $p_3(e)$ and give a related conjecture. Since a 4-cycle contains four paths of length three, we have:

Corollary 5 *A graph of size $e \geq 6$ contains at most $h(h-1)(h-2)(h-3)/8$ 4-cycles, where $h \geq 0$ satisfies $h(h-1)/2 = e$.*

Conjecture: *Let $\mathbf{G}$ be a graph of size $e > 8$, and write $e = \binom{k}{2} + t$, $1 \leq t \leq k$. Then $\mathbf{G}$ contains at most $3\binom{k}{4} + (k-2)\binom{t}{2}$ cycles of length four, with equality when $\mathbf{G}$ is the graph obtained by joining a new vertex to t vertices of $\mathbf{K}_k$.*

The graph $\mathbf{K}_{4,2}$ necessitates that $e > 8$ in the conjecture. Also, the (conjectured) extremal graphs are not unique. In particular, if $e = \binom{s}{2} + 3s$ for some integer s, then the graph obtained by deleting a triangle from $\mathbf{K}_{s+3}$ contains the same number of 4-cycles as the graph described in the conjecture.

References

[1] N. Alon, On the number of subgraphs of prescribed type of graphs with a given number of edges, *Israel J. Math*, **38** (1981), 116-130.

[2] B. Bollobás and A. Sarkar, Paths in Graphs, to appear.

[3] B. Bollobás and A. Sarkar, Paths of Length Four, to appear.

[4] O. Byer, Two path extremal graphs and an application to a Ramsey-type problem, *Discrete Mathematics*, **196** (1999), 51-64.

[5] R. Ahlswede and G. O. H. Katona, Graphs with Maximal Number of Adjacent Pairs of Edges, *Acta Mathematica Academiae Scientiarum Hungaricae* **32** (1978), 97-120.