ON THE PATH SEPARATION NUMBER OF GRAPHS

JÓZSEF BALOGH, BÉLA CSABA, RYAN R. MARTIN, AND ANDRÁS PLUHÁR

Abstract. A *path separator* of a graph G is a set of paths $\mathcal{P} = \{P_1, \dots, P_t\}$ such that for every pair of edges $e, f \in E(G)$, there exist paths $P_e, P_f \in \mathcal{P}$ such that $e \in E(P_e)$, $f \notin E(P_e)$, $e \notin E(P_f)$ and $f \in E(P_f)$. The *path separation number* of G, denoted $\mathrm{psn}(G)$, is the smallest number of paths in a path separator. We shall estimate the path separation number of several graph families, including complete graphs, random graph, the hypercube, and discuss general graphs as well.

Keywords: network reliability, test sets, path covering, path separation, path decomposition, trees
MSC numbers: 05C35, 05C70

1. Introduction

Given a transparent optical network G, localizing a failed edge is extensively studied problem in engineering (since it is unlikely that there are several failures at the same time, research in the area mostly focuses on having at most one failed connection). It is common to work with the following scenario: there is the *central controller* that can see whether a set of *optical tests* fail or not. An optical test can e.g. be a path between two vertices of the network, and a single test fails if the light cannot propagate from one of its endpoints to the other [1, 4]. Alternatively, one may use so called *monitoring-cycles*, or *monitoring-trails* (m-cycles or m-trails) for this purpose, or trees or even arbitrary connected subgraphs of G [9, 8]. The goal is to use as few tests as possible so that the central controller can unambiguously tell, which edge (optical connection) failed, if any. Clearly, one can test each and every edge in the network, however, this is very expensive in general.

More formally, our model is as follows. The set of graphs $G_1, \dots, G_t \subset G$ is a *test set* if for every $e \in G$ there is an $I_e \subset \{1, \dots, t\}$ such that $\cap_{i \in I_e} E_i = \{e\}$, where E_i is the edge set of G_i for $i = 1, \dots, t$. Note that this can be interpreted as follows: One unknown edge (but not more) may fault in G, and the central controller gets the list of G_i's containing e (this setup is also called combinatorial group testing). Either the controller can identify the faulty edge, or claim there is no such edge. In practice, most of the time one wants a small test set that contains either m-trails or connected subgraphs.

If there are no restrictions for the test subgraphs, t can be as small as $\lceil \log_2(m) \rceil$, where $m = |E(G)|$. This information theoretical lower bound can sometimes be achieved or at least approximated. Clearly, the stronger conditions we impose on the tests, the larger our test set will be. For example, under the relatively mild condition that G has two edge-disjoint spanning trees (if G is 4-edge-connected, it has this property), one can construct a (very small) test set containing at most $2\lceil \log_2(m) \rceil$ connected tests [4]. On the other hand, when one uses paths as tests, it is easy to see that $\Omega(m/n)$ tests are needed.

The first author's research is partially supported by NSF CAREER Grant DMS-0745185, TÁMOP-4.2.1/B-09/1/KONV-2010-0005, and Marie Curie FP7-PEOPLE-2012-IIF 327763.

The second and fourth authors were partially supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013). The second author was also supported in part by the ERC-AdG. 321104.

The third author's research partially was supported by NSF grant DMS-0901008 and by NSA grant H98230-13-1-0226.

Here we consider a closely related parameter, and derive results for the most natural restrictions. Let G be a graph with at least two edges. A *separator* of G is a collection of subgraphs $\mathcal{S} = \{S_1, \dots, S_t\}$ such that for every pair of edges $e, f \in E(G)$, there exist sets $S_e, S_f \in \mathcal{S}$ such that $e \in E(S_e)$, $f \notin E(S_e)$, $e \notin E(S_f)$ and $f \in E(S_f)$.

Observe that a separator is a test set. Note that for the non-restricted case basically the same $\lceil \log_2(m) \rceil$ bounds hold.

A *path separator* of G is a set of paths $\mathcal{P} = \{P_1, \dots, P_t\}$ such that $\mathcal{P}$ is a separator. The *path separation number* of G, denoted $\mathrm{psn}(G)$, is the smallest number of paths in a path separator. If G has exactly 1 edge, we say that $\mathrm{psn}(G) = 1$ and if G is empty, we say that $\mathrm{psn}(G) = 0$.

Denote $H_2(x)$ to be the *binary entropy function*, i.e. $H_2(x) = -x \log_2 x - (1-x)\log_2(1-x)$, where $x \in (0,1)$. Denote K_n to be the complete graph, and P_n to be the path on n vertices. The parameters $\delta(G)$ and $\Delta(G)$ denote the minimum and maximum degree of G, respectively.

Because of Fact 1 we will always assume that the graph G that we are working with is connected.

Fact 1. *If G is a graph that is the vertex-disjoint union of graphs G_1 and G_2, then* $\mathrm{psn}(G) = \mathrm{psn}(G_1) + \mathrm{psn}(G_2)$.

When G is a tree, we determine $\mathrm{psn}(G)$ in Theorem 4, otherwise Theorem 2 estimates it. Note that the proof of the lower bound in Theorem 2 does not use the structure of paths, only that a path has at most $n-1$ edges.

Theorem 2. *Let G be a graph on $n \geq 2$ vertices and $m \geq n$ edges, then*

$$\frac{m \ln m}{n \ln(en/2)} < \frac{\log_2 m}{H_2\left((n-1)/m\right)} \leq \mathrm{psn}(G) \leq 2n\lceil \log_2 \lceil m/n \rceil \rceil + n < 3n \log_2 n.$$

Theorem 3 establishes that the path separation number of the complete graph is at most $2n+4$, and Theorem 2 implies that it is at least $(1-o(1))n$.

Theorem 3. *For $n \geq 5$ we have* $\mathrm{psn}(K_n) \leq 4\lceil n/2 \rceil + 2 \leq 2n + 4$.

Theorem 4 gives an explicit formula for the path separation number of a forest, F, depending only on the degree sequence and the number of connected components of F that are, themselves, paths. A *path-component* of a graph is a connected component that is a path.

Theorem 4. *Let F be a forest with v_1 vertices of degree 1, v_2 vertices of degree 2 and p path-components. Then,* $\mathrm{psn}(F) = v_1 + v_2 - p$.

Corollary 5. *The smallest path separation number for a tree T on n vertices is $\lceil n/2 \rceil + 1$. This is achieved with equality if and only if (a) n is even and all the degrees of T are either* 1 *or* 3 *or (b) n is odd, T has one vertex of degree either* 2 *or* 4 *and all other vertices have degree either* 1 *or* 3.

Theorem 6 considers the graph of the d-dimensional hypercube Q_d, whose path separation number shows different behavior from our previous results.

Theorem 6. *For $d \geq 2$, let Q_d denote the d-dimensional hypercube. Then* $\frac{d^2}{2\log d} \leq \mathrm{psn}(Q_d) \leq 2d(d+1) - 8$.

In theorem 7, we address the Erdős-Rényi random graph in which each pair of vertices is, independently, chosen to be an edge with probability p. We say that a sequence of random events occurs *with high probability (whp)* if the probability of the events approaches 1 as $n \to \infty$.

Theorem 7. *Let $p = p(n) > 10 \log n/n$ and $s = 4 \log n/\log(pn/\log n)$. Then,* $\mathrm{psn}(G(n,p)) = O(psn)$*, whp.*

In particular, for $\alpha > 0$ and $p = p(n) > n^{\alpha - 1}$ this gives $\mathrm{psn}(G(n,p)) = \Theta(pn)$ *whp, and for $p = p(n) > 10 \log n/n$ it yields that* $\mathrm{psn}(G(n,p)) = O(pn \log n)$*, whp.*

In the rest of the paper we shall prove the above theorems. We close Section 1 with the following conjecture.

Conjecture 8. *For every graph G on n vertices* $\operatorname{psn}(G) = O(n)$.

In fact, for any $\varepsilon > 0$, there is an n-vertex graph whose path separation number must be as large as $(2-\varepsilon)n$: Using the entropy method, in a manner similar to the proof of Theorem 2, one can show the following:

For $a \le (n-1)/2$, the complete bipartite graph $K_{a,n-a}$ has a path separation number of at least $\left(\frac{a(n-a)}{2a+1}\right)\frac{\ln a(n-a)}{\ln(en/2)}$. In particular, if $0 < \alpha < 1$ is fixed, then, for $a = n^\alpha$, the path separation number of $K_{a,n-a}$ is at least $(1-o(1))\frac{1+\alpha}{2}n$.

2. Proofs

2.1. **Proof of Theorem 2: Lower bound.** We use the entropy method. For facts about the entropy method, see Section 22 of Jukna [5]. Let $\pi_1, \dots, \pi_t$ be the paths of a path separator of graph G. Let X_i be the event that a randomly-chosen edge is in path π_i. Since the joint distribution $(X_1, \dots, X_t)$ takes on m values, $H_2(X_1, \dots, X_t) = \log_2 m$. Using the subadditivity property of entropy,

$$\log_2 m \le H_2(X_1, \dots, X_t) \le \sum_{i=1}^{t} H_2(X_i) = \sum_{i=1}^{t} H_2\left(\frac{\ell(\pi_i)}{m}\right) \le tH_2\left(\frac{n-1}{m}\right),$$

because every path has length at most $n-1$. Writing $x = (n-1)/m$ we have

$$\begin{aligned} t &\ge \frac{\log_2 m}{H_2\left(\frac{n-1}{m}\right)} \ge \frac{\ln m}{-x\ln x - (1-x)\ln(1-x)} > \frac{\ln m}{-x\ln x + x} \\ &= \frac{m\ln m}{n-1} \cdot \frac{1}{\ln\left(\frac{m}{n-1}\right)+1} \ge \frac{m\ln m}{n-1} \cdot \frac{1}{\ln\left(\frac{n}{2}\right)+1} = \frac{m\ln m}{(n-1)\ln(en/2)}. \end{aligned}$$

Proof of Theorem 2: Upper bound. We use a classical theorem of Lovász [6]:

Theorem 9 (Lovász [6])**.** *The edges of a graph on n vertices can be covered by at most $\left\lfloor\frac{n}{2}\right\rfloor$ edge-disjoint paths and cycles.*

Since any cycle can be partitioned into two paths, we obtain the following corollary.

Corollary 10. *The edges of a graph on n vertices can be covered by at most n edge-disjoint paths.*

Apply Corollary 10 to G, and partition $E(G)$ into at most n paths, $P_1', \dots, P_n'$. We shall cut each path that is longer than m/n into paths of length $\lceil m/n \rceil$, and possibly one shorter path. The number of such paths is

$$\sum_{i=1}^{n} \left\lceil \frac{|P_i'|}{\lceil m/n \rceil} \right\rceil \le \sum_{i=1}^{n} \left\lceil \frac{n|P_i'|}{m} \right\rceil \le \sum_{i=1}^{n} \left(\frac{n|P_i'|}{m} + 1\right) = 2n$$

where $|P_i|$ is length (number of edges) of P_i. So we have that the new family, $\mathcal{P}$ consists of at most $2n$ paths. For each path $P \in \mathcal{P}$ label the the edges of P by binary vectors. Specifically, take an injective function $b_P : E(P) \to \{0,1\}^t \setminus \{0\}$, where $t = \lceil \log_2 \lceil m/n \rceil \rceil$. For each $i \in [t]$ define the graph G_i, whose vertex set is $V(G)$ and whose edge set consists of edges of path $P \in \mathcal{P}$ whose i-th digit in its vector is 1. We apply Corollary 10 to each G_i, giving a path partition $\mathcal{P}_i$. Observe that $\cup_i \mathcal{P}_i \cup \mathcal{P}$ is a path-separating system of G of size at most $2n\lceil \log_2 \lceil m/n \rceil \rceil + n$. The final inequality comes from using $m \le \binom{n}{2}$. □

2.2. **Proof of Theorem 3.** We start with a corollary of Theorem 9: The edges of K_n can be covered by at most $\lceil \frac{n}{2} \rceil$ edge-disjoint paths.

Note that Lovász [6] proved this when n is even. For n odd, Theorem 9 implies the existence of a partition of K_n into $\lfloor n/2 \rfloor$ Hamiltonian cycles, but one can check that it is always possible to choose one edge from each Hamilton cycle the way that these edges could be covered with one additional path.

Denote such a collection of $\lceil n/2 \rceil$ edge-disjoint paths by $\mathcal{P}_1 = \{P_1, \ldots, P_{\lceil n/2 \rceil}\}$. Select three random permutations α, β and γ uniformly and independently of each other from S_n, the set of n-element permutations. Let $\mathcal{P}_\alpha$, $\mathcal{P}_\beta$ and $\mathcal{P}_\gamma$ be the images of $\mathcal{P}_1$ under the permutations α, β and γ, respectively. That is, if $P_i = \{i_1, \ldots, i_n\} \in \mathcal{P}_1$, then, say, $\alpha(P_1) = \{\alpha(i_1), \alpha(i_2), \ldots, \alpha(i_n)\}$, assuming $V(K_n) = [n]$.

If a pair of edges, e, f are in different paths in $\mathcal{P}_1$, then they are separated by $\mathcal{P}_1$. Otherwise the probability that they are *not* separated by $\mathcal{P}_\alpha$ is at most $2/(n-2)$. The separations by $\mathcal{P}_\alpha, \mathcal{P}_\beta$ and $\mathcal{P}_\gamma$ are independent of each other, i.e. the probability that e and f are not separated by the system of paths

$$\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_\alpha \cup \mathcal{P}_\beta \cup \mathcal{P}_\gamma$$

is at most $(2/(n-2))^3$. The number of pairs that are not separated by $\mathcal{P}_1$ is at most $\binom{n-1}{2} \lceil \frac{n}{2} \rceil$, and so the expected number of pairs not separated by $\mathcal{P}$ is less than $(2/(n-2))^3 \binom{n-1}{2} \lceil \frac{n}{2} \rceil < 3$ for $n \geq 5$. We can finish the path separator with two additional paths that separate the remaining two pairs of edges. $\square$

2.3. **Proof of Theorem 4.** First, we prove that $\mathrm{psn}(P_k) = k-1$, where P_k is a path with k vertices $v_1, \ldots, v_k$, and edges $\{v_1v_2, v_2v_3, \ldots, v_{k-1}v_k\}$. For the upper bound, observe that the edge set itself is a path separator of size $k-1$. For the lower bound, one observes that $\{v_{k-1}, v_k\}$ must be a member of any path separator, otherwise the edge $v_{k-1}v_k$ cannot be separated from $v_{k-2}v_{k-1}$. Any path separator of $\{v_1, \ldots, v_{k-1}\}$ has size at least $k-2$ and the lower bound follows. By Fact 1, it is sufficient to prove only that $\mathrm{psn}(T) = v_1 + v_2$ for any tree different from a path with v_1 leaves and v_2 vertices of degree 2.

Let T be a non-path tree on $n \geq 4$ vertices with v_1 leaves and v_2 vertices of degree 2. Add an extra vertex (we may call it ∞) that is incident to all of the leaves of T. Call the new graph T^∞. The resulting graph is planar and we can form a path for every face of the embedding of T^∞ into the plane that begins and ends at a leaf and contains every vertex (other than "∞") of the face. Each leaf will be the endpoint of exactly two such paths. For each vertex w of degree 2, take one of these paths that passes through it and partition it into two subpaths (that is, subgraphs that are themselves paths), both of which have w as an endpoint. We continue this process until we have separated all edges incident to the same degree-2 vertex. In other words, the degree-2 vertices are 'cut points' of these paths.

The size of this family of paths is $v_1 + v_2$. This family clearly separates any pair of edges that are not on the same two faces (each edge is on exactly two faces). If two edges, e and f, share the same pair of faces, then there must be a path of degree 2 vertices (possibly with 0 additional edges, but containing at least one vertex) connecting them. But by the partition of the paths at vertices of degree 2, one of the original paths containing e and f was partitioned into at least two paths, one that contains e but not f and another that contains f but not e.

For the lower bound, we proceed by induction on the order of T. For the base case, the smallest tree which is not a path is a star on 4 vertices; easily a separating system of size 3 exists. If there is a leaf that is covered by a path of length 1, then both the leaf from the tree, and the path from the system can be removed. Assume that there is a tree T with a minimum-sized path separator $\{P_1, \ldots, P_t\}$ such that no leaf is in exactly one path. In this case we use a simple discharging argument, give each P_i a charge of 1 and discharge 1/2 to each of its endpoints in T. Every degree-2 vertex, w, must receive a charge of at least 1, otherwise the incident edges of w are not separated. Every leaf, x, must receive a charge of at least 1 because there are at least 2 paths that contain the edge incident to x. Thus, the number of paths is at least $v_1 + v_2$. □

2.4. **Proof of Corollary 5.** Lemma 11 with Theorem 4 implies that the smallest path separation number for a tree on n vertices is $\lceil n/2\rceil + 1$.

Lemma 11. *Let T be a tree on $n \geq 2$ vertices, with v_1 vertices of degree 1 and v_2 vertices of degree 2. Then $v_1 + v_2 \geq \lceil n/2\rceil + 1$. This is achieved with equality if and only if (a) n is even and all the degrees of T are either* 1 *or* 3 *or (b) n is odd, T has one vertex of degree either* 2 *or* 4 *and all other vertices have degree either* 1 *or* 3.

2.5. **Proof of Lemma 11.** Let us use the notation that v_i denotes the number of vertices of degree i in T. Then $\sum_i v_i = n$ and $\sum_i iv_i = 2n-2$ implies that $2v_1+v_2 \geq n+2$. For n even, v_1+v_2 is minimal if $v_1 = n/2+1, v_3 = n/2-1$. For n odd, $v_1 + v_2$ is minimal if $v_1 = (n+1)/2, v_2 = 1, v_3 = (n-3)/2$ or $v_1 = (n+3)/2, v_3 = (n-5)/2, v_4 = 1$. In all cases, $v_1 + v_2 \geq \lceil n/2\rceil + 1$. For all n, trees exist with each of the above degree sequences, therefore the inequalities are sharp. □

2.6. **Proof of Theorem 6.** For the lower bound we use Theorem 2, noting that Q_d has 2^d vertices and $d2^{d-1}$ edges:

$$\operatorname{psn}(Q_d) \geq \frac{\log_2(d2^{d-1})}{H_2((2^d-1)/d2^{d-1})}. \tag{1}$$

If $d = 2$, then it is easy to see that $\operatorname{psn}(Q_2) = 4 \geq \frac{2^2}{2\log_2 2}$.

If $d = 3$, then (1) gives $\operatorname{psn}(Q_3) \geq \frac{\log_2(12)}{H_2(7/12)} > 3.6 \geq \frac{3^2}{2\log_2 3}$.

If $d \geq 4$, then $H_2((2^d-1)/d2^{d-1}) \leq H_2(2/d) \leq -(2/d)\log_2(2/d) + 2/d$. Therefore,

$$\operatorname{psn}(Q_d) > \frac{\log_2(d2^{d-1})}{H_2(2/d)} > \frac{d-1+\log_2 d}{-(2/d)\log_2(2/d)+2/d} = \frac{d^2-1+d\log_2 d}{2\log_2 d} \geq \frac{d^2}{2\log_2 d}.$$

Consider the upper bound. We will show, by induction on d that $f(d) = 2d(d+1) - 8$ paths suffice. As we have established above, $\operatorname{psn}(Q_2) = 4$. It is helpful to view the vertices of Q_d as $\{0,1\}$-vectors of dimension d. Let Q^i denote the $(d-1)$-dimensional subcube whose vertices have 0 in the d^{th} coordinate, for $i = 1, 2$. Edges between vertices in Q^i are called *i-interior edges*, for $i = 0, 1$. Edges with one endvertex in Q^0 and the other in Q^1 are called *crossing edges*. Consider an edge e^0 in Q^0 and an edge e^1 in Q^1. We call such edges *mirror images* if the endvertices of e^0 can be made into the endvertices of e^1 simply by changing the d^{th} coordinate from 0 to 1. For a path in Q^0, the *mirror image path* is defined in an analogous way.

We construct three different types of paths. First, by the inductive hypothesis, Q^0 has a path separation set of size $f(d-1)$. Construct it and its mirror image in Q_1. Then, for each pair of mirrored paths, connect their final endpoints via a crossing edge. This is a set of paths of size $f(d-1)$.

With this set of paths, the following pairs of edges (e', e'') are separated: (1) if both are 0-interior edges or both are 1-interior edges, or (2) if e' is 0-interior and e'' is 1-interior but they are not mirror images.

So, there are only three types of pairs of edges (e', e'') that are not separated: (3) if e' and e'' are mirror images or (4) if e' is a crossing edge and e'' is an interior edge or (5) if both e' and e'' are crossing edges.

Second, we construct an arbitrary system of paths that covers the edges of Q^0. It is easy to prove by induction that there is such a system consisting of d paths. Construct the set of mirror image paths in Q^1. This second group of paths is of size $2d$ and separates any pair of mirror image edges. Observe that doing so for every crossing edge e' and interior edge e'' there is a path containing e'' but not e'.

We find a *separating family* of the 2^{d-1} crossing edges. That is, sets $S_1, \dots, S_{d-1}$ of crossing edges so that for each pair of crossing edges, there is an S_i that contains the first but not the second and an S_j that contains the second but not the first.[1]

For each S_j, we will construct two paths utilizing a Hamilton path $v_1^0, \dots, v_{2^{d-1}}^0$ in the $(d-1)$-dimensional hypercube, Q^0, (Hypercubes of dimension at least 2 are well-known to be Hamiltonian.) and its mirror image in Q^1. Let $S_j = \{v_{i_1}^0 v_{i_1}^1, v_{i_2}^0 v_{i_2}^1, \dots, v_{i_N}^0 v_{i_N}^1\}$ with $i_1 < i_2 < \cdots < i_N$. The first of the two paths related to S_j will begin by traversing the Hamilton path in Q^0, crossing at the first opportunity and then traversing the mirror image Hamilton path in Q^1, crossing at the next opportunity and continuing in Q^0. We continue this until we finish in Q^0:

$$v_1^0, \dots, v_{i_1}^0, v_{i_1}^1, v_{i_1+1}^1, \dots, v_{i_2-1}^1, v_{i_2}^1, v_{i_2}^0, v_{i_2+1}^0, \dots, v_{i_N-1}^1, v_{i_N}^1, v_{i_N}^0, v_{i_N+1}^0, \dots, v_{2^{d-1}}^0.$$

The second path is a mirror image:

$$v_1^1, \dots, v_{i_1}^1, v_{i_1}^0, v_{i_1+1}^0, \dots, v_{i_2-1}^0, v_{i_2}^0, v_{i_2}^1, v_{i_2+1}^1, \dots, v_{i_N-1}^0, v_{i_N}^0, v_{i_N}^1, v_{i_N+1}^1, \dots, v_{2^{d-1}}^1.$$

This third group of paths is of size $2(d-1)$ and completes the path separation. Let us do the same procedure with S_0, the set of *all* crossing edges. We add two new (Hamiltonian) paths to our system, first starting the alternating path from an $x^0 \in Q^0$, then from $x^1 \in Q^1$. These paths contain every crossing edge e', but every interior edge e'' is left out from at least one of those.

The total number of paths in the three groups is $f(d-1) + 4d$. Setting $f(d) = 2d(d+1) - 8$ satisfies $f(2) = 4$ and $f(d) = f(d-1) + 4d$. □

Remark. We believe that the lower bound is correct in the sense that $\mathrm{psn}(Q_d) = O(d^2/\log d)$. We think that a proof is likely in the same vein as the proof of the upper bound of $\mathrm{psn}(K_n)$. Note that $E(Q_d)$ can be covered by d paths, as we have described in the above proof. Fix such a path system $\mathcal{P} = P_1, \dots, P_d$. Now choose, randomly and independently, $100d/\log d$ automorphisms of Q_d, and apply it to $\mathcal{P}$. This will give a path system $\mathcal{P}^*$ of size $100d^2/\log d$. The system $\mathcal{P}$ does not separate at most $d2^{2d}$ pairs of edges, the ones which are in the same path P_i for some i. Unfortunately, we do not have a good estimate on the probability that a pair of edges (e', e'') are not separated in $\mathcal{P}^*$, unless we know that no P_i contains more than $O(d/\log d)$ edges that are crossing with respect to a given partition.

2.7. **Proof of Theorem 7.** The idea of the proof is to partition $G(n,p)$ into several random graphs, such that every pair of edges should be separated by them, and every pair of edges should be in several of the random graphs. Then by Vizing's theorem, we partition the edges of each of the random graphs into matchings, and using some other random graphs we connect them into paths.

Let G be a graph chosen according to the distribution $G(n,p)$. First we partition $E(G)$ into four random graphs. Let $f: E(G) \to \{1,2,3,4\}$ be a function so that each $f(e)$ is chosen uniformly from $\{1,2,3,4\}$ independently. Let $E_i : \{e : f(e) = i\}$ for $1 \le i \le 4$.

Next we form six random subgraphs with the same vertex set $V(G)$ and with edge sets as follows: $E(G_1^1) = E_1 \cup E_2$, $E(G_1^2) = E_3 \cup E_4$, $E(G_2^1) = E_1 \cup E_3$, $E(G_2^2) = E_2 \cup E_4$, $E(G_3^1) = E_1 \cup E_4$, $E(G_3^2) = E_2 \cup E_3$, These six random subgraphs have the property that for any pair of edges $e, f \in E(G)$ there is an i, j that $e, f \in E(G_i^j)$.

[1]Any separating family of a set, Σ, of size n can be found, e.g., by assigning binary codes of length $\lceil \log_2 n \rceil$ to each member of Σ and then placing an element into the j^{th} member of the family if and only if the j^{th} bit of its code is 1. Such a separating family has size $\lceil \log_2 n \rceil$ and the sizes of the sets are at most $n/2$.

Now fix a pair of indices (i,j). Without loss of generality, assume $i=j=1$ and consider G_1^1. Note that G_1^1 is itself a random graph, distributed according to $G(n,p/2)$. We will further partition the edgeset of G_1^1 into random subgraphs.

Fix $r=\lfloor 3pn/\log n\rfloor$ and let $g: E(G)\to\{1,\dots,r\}$ be a function so that each $g(e)$ is chosen uniformly from $\{1,\dots,r\}$ independently. Repeat this process $s=\lfloor 4\log n/\log(pn/\log n)\rfloor$ times. Because $p=\Omega(\log n/n)$, we have $sr\le \frac{4\log n}{\log(pn/\log n)}\cdot\frac{3pn}{\log n}=O(n)$ subgraphs $H_1,\dots,H_{sr}$, each of which is a copy of $G(n,p/(2r))$.

The set of graphs $\{H_\alpha:\alpha=1,\dots,sr\}$ will separate every pair of edges with high probability. To see this, the union bound gives

$$\mathbb{P}\left(\exists e,f\in E(G_1^1):\text{no } H_\alpha \text{ separates } e,f\right)\le\binom{n}{2}\left(\frac{1}{r}\right)^s=O\left(n^{-2}\right).$$

Furthermore, all of the graphs H_α have maximum degree less than $2n\frac{p}{2r}\le\log n/3$, noting that the average degree is $n\frac{p}{2r}$. By a Chernoff bound (Theorem 2.1 of Bollobás [2]) and the union bound,

$$\mathbb{P}(\exists\alpha:\Delta(H_\alpha)\ge pn/r)\le sr\exp\left\{-\frac{3pn}{8r}\right\}=o(1),$$

because $r\le 3pn/\log n$ and $sr=O(n)$.

The idea for the rest of the proof is that by Vizing's theorem, H_α can be partitioned into at most $\log n/3\ge\Delta(H_\alpha)+1$ matchings, and by using edges of another subgraph of G_1^2, we connect them into paths. The total number of paths used will be at most

$$(\log n/3)\cdot s\cdot pn/\log n=spn/3.$$

We will need an additional notion. Let $D(n,p)$ be the oriented directed graph on n vertices; i.e., for every *ordered pair* (x,y), we draw an edge with probability p, independently of each other. McDiarmid showed in [7] that the probability of the existence of a Hamiltonian cycle in $D(n,p)$ is not less than the same in $G(n,p)$, that is

$$\mathbb{P}(G(n,p)\text{ is Hamiltonian})\le\mathbb{P}(D(n,p)\text{ is Hamiltonian}).$$

The usual proofs for the Hamiltonicity of $G(n,p)$ (see, e. g. Chapter 8 of Bollobás [2]) give that

$$\mathbb{P}(G(n,p)\text{ is Hamiltonian})\ge 1-n^{-\omega(1)}. \tag{2}$$

So apply Vizing's theorem and decompose, say H_1, into $\Delta_0+1=\log n/3$ matchings, $M_1,\dots,M_{\Delta_0+1}$. Then for each $i\in\{1,\dots,\Delta_0+1\}$ we form a separating matching system $M_i^1,\dots,M_i^t$ for each M_i, where $t=\lceil\log_2(\Delta_0+1)\rceil$. I.e., for every pair of edges in M_i, there is a pair $\{M_i^{j_1},M_i^{j_2}\}$ such that one edge is in $M_i^{j_1}-M_i^{j_2}$ and the other is in $M_i^{j_2}-M_i^{j_1}$.

Now for each $i\in[\Delta_0+1]$ and $j\in[t]$ we find a path containing the edges of $M_i^j\subset E(G_1^1)$, connected by the edges of G_1^2. For this, we define an auxiliary oriented directed graph D. The vertex set of D will be the union of M_i^j and $V(G)-V(M_i^j)$. The edge set will be defined as follows: First assign a random orientation to the edges of M_i^j. For (oriented) edges $(xy),(uv)\in M_i^j$ we have the edge $(xy)(uv)\in E(D)$ if $yu\in E(G_1^2)$, and the edge $(uv)(xy)\in E(D)$ if $vx\in E(G_1^2)$. For an (oriented) edge $xy\in M_i^j$ and $u\in V(G)-V(M_i^j)$ we have in D the oriented edge $(xy)u$ if $yu\in E(G_1^2)$, and we have the oriented edge $u(xy)$ if $xu\in E(G_1^2)$. If $w,z\in V(G)-V(M_i^j)$, then both oriented edges (wz) and (zw) are in $E(D)$.

The key property of D is that if it contains a Hamilton path, then $M_i^j\cup E(G_1^2)$ will contain the desired path. The edge probability in D is the same as in $G(n,p/2)$,[2] therefore it contains a Hamilton path with probability at least $1-n^{-\omega(1)}$ by (2).

[2]We have an edge in D with probability $p/2$, independent of all the other edges.

Since the total number of such paths is at most $spn/3 = O(n \log n)$, the union bound gives that all of these paths can be constructed, whp. □

Added in Proof. Around the time that we were finishing writing up our results, a similar paper appeared in the arXiv by Falgas-Ravry, Kittipassorn, Korándi, Letzter, and Narayanan [3]. Their work is independent from ours, and they consider a different separation: for each pair of edges they are interested to find a separating set which contains exactly one of them. In many of the cases (like trees), this leads to a different behavior, for some other cases similar proof techniques as in our paper might be applied.

References

[1] S. Ahuja, S. Ramasubramanian and M. Krunz, Single-link failure detection in all-optical networks using monitoring cycles and paths, *IEEE/ACM Transactions on Networking*, **17**(4), 2009, pp. 1080–1093.

[2] B. Bollobás, *Random graphs*. Second edition. Cambridge Studies in Advanced Mathematics, 73, *Cambridge University Press, Cambridge*, 2001, xviii+48pp.

[3] V. Falgas-Ravry, T. Kittipassorn, D. Korándi, S. Letzter, B. P. Narayanan, Separating path systems, arXiv:1311.5051.

[4] N. Harvey, M. Patrascu, Y. Wen, S. Yekhanin and V. Chan, Non-Adaptive Fault Diagnosis for All-Optical Networks via Combinatorial Group Testing on Graphs, *INFOCOM 2007, 26th IEEE International Conference on Computer Communications*, 2007, pp. 697–705.

[5] S. Jukna, *Extremal Combinatorics. With Applications in Computer Science*. Second edition. Texts in Theoretical Computer Science. An EATCS Series. *Springer, Heidelberg*, 2011. xxiii+411pp.

[6] L. Lovász, On covering of graphs. 1968 *Theory of Graphs (Proc. Colloq. Tihany, 1966)* pp. 231–236 *Academic Press, New York*.

[7] C. McDiarmid, Clutter percolation and random graphs. Combinatorial Optimization II (Proc. Conf., Univ. East Anglia, Norwich, 1979). *Math Programming Stud.* No. 13 (1980), 17–25.

[8] J. Tapolcai, B. Wu, P.-H. Ho, On monitoring and failure localization mesh all-optical networks, *INFOCOM 2009*, 2009, pp. 1102–1113.

[9] B. Wu, P.-H. Ho, and K. Yeung, Monitoring trail: a new paradigm for fast link failure localization in WDM mesh networks, *IEEE GLOBECOM 2008*, 2008, pp. 1–5.

[10] H. Zeng and C. Huang, *Fault detection and path performance monitoring in meshed all-optical networks*, *IEEE GLOBECOM 2004*, Vol. 3, 2004, pp. 2014–2018.

Department of Mathematical Sciences, University of Illinois and Bolyai Institute, University of Szeged

E-mail address: jobal@math.uiuc.edu

Bolyai Institute, University of Szeged

E-mail address: bcsaba@math.u-szeged.hu

Department of Mathematics, Iowa State University, Ames, Iowa 50011

E-mail address: rymartin@iastate.edu

Department of Computer Science, University of Szeged

E-mail address: pluhar@inf.u-szeged.hu