95

Spanning Tree Edge Densities

Michael Ferrara[1] Ronald Gould [1] Charles Suffel[2]

Abstract

In the design of reliable and invulnerable networks, it is often a goal to maximize the number of spanning trees of a graph with a given number of vertices and edges. It is therefore logical to investigate the importance of individual edges to the number of spanning trees of a graph. Given a graph G, the spanning tree edge density or hereafter, simply density, of an edge e in G is the fraction of the spanning trees of G that contain e. Several bounds on the densities of the edges in a graph are given, including a lower bound on the maximum density based solely on edge-connectivity. It is also shown that the addition of edges to a graph decreases the densities of the preexisting edges, an important result from a vulnerability standpoint. Finally, some challenging open problems are presented as related to this new parameter.

1 Preliminaries

We consider undirected graphs G, possibly with multiple edges, and denote the vertex and edge sets of G by $V(G)$ and $E(G)$ respectively. For any v in $V(G)$ let $N(v)$ be the neighborhood of v. Also, we will denote the edge-connectivity of G by $\lambda(G)$. The shorthand P_n, C_n, K_n and $K_{m,n}$ will be used to denote the path, cycle, complete and complete bipartite graphs on the indicated number of vertices. A good reference for any other undefined terms is [3].

A spanning tree T of G is a connected acyclic spanning subgraph of G. Let $T(G)$ denote the set of spanning trees of G and define $t(G) = |T(G)|$. In general, it is not difficult to calculate $t(G)$ using elementary

[1]Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322

[2]Department of Mathematical Sciences, Stevens Insitutute of Technology, Hoboken, NJ 07030

CONGRESSUS NUMERANTIUM 154 (2002), pp. 155-163

linear algebraic methods, as outlined in [2] and elsewhere. We may consider any $e = uv$ in $E(G)$ and contract e by associating the vertices u and v in G and deleting the resulting loop to get the graph $G \cdot e$. It is well known that for any e,

$$t(G) = t(G - e) + t(G \cdot e)$$

which will prove to be useful in the sequel.

2 Spanning Tree Edge Densities

Recall that Cayley's Theorem (see [3] or numerous other sources) states that $t(K_4) = 4^{4-2} = 16$. However, if we were to remove any edge e from K_4 is is easily seen that only 8 spanning trees remain. Hence one can deduce that any edge of K_4 lies in $\frac{8}{16} = \frac{1}{2}$ of its spanning trees. The question that arises is therefore: what is the importance of an individual edge to the number of spanning trees of a graph? Is it possible to structure a graph to have a large number of spanning trees in such a way that the failure of any one of these edges does not significantly reduce the number of spanning trees? This leads to the general idea of the *density* of an edge in a graph.

Let $T(e; G)$ denote the set of spanning trees of G that contain e and $t(e; G) = |T(e; G)| = t(G \cdot e)$.

Definition 2.1 *For any edge e in a graph G, the* **spanning tree edge density** *or for our purposes, simply density of e, denoted $d_t(e; G)$, is defined as*

$$d_t(e; G) = \frac{t(e; G)}{t(G)}.$$

If G is disconnected, define $d_t(e; G)$ to be 0 for all e in $E(G)$.

We will simply write $d_t(e)$, $t(e)$, $orT(e)$ when the context is clear. One could also define $d_t(e)$ as

$$d_t(e) = 1 - \frac{t(G - e)}{t(G)}.$$

Given a graph G, we can construct the digraph associated with G by replacing each edge with two opposite arcs. It is shown in [5] that upon inverting the Kirchoff matrix of this digraph, one can calculate the density of any edge in G in constant time.

The following lemma will be useful in the sequel.

Lemma 2.2 *Let G be a connected graph on n vertices. Then*

$$\sum_{e \in E(G)} d_t(e) = n - 1.$$

PROOF: Consider the set of ordered pairs (t, e) where $t \in T(G)$ and $e \in E(t)$. Each spanning tree t has exactly $n - 1$ edges, so there must be $(n-1)t(G)$ such pairs. On the other hand, each edge e lies in exactly $t(e) = d_t(e)t(G)$ pairs. Hence

$$(n-1)t(G) = t(G) \sum_{e \in E(G)} d_t(e)$$

and the result follows. □

2.1 Weakly Edge Transitive Graphs

We will now calculate the densities of edges in a highly symmetric class of graphs. We say that a permutation π of $E(G)$ is an *edge automorphism* provided that for any edges e_1 and e_2 in G, e_1 and e_2 are adjacent if and only if πe_1 and πe_2 are adjacent.

Definition 2.3 *A graph G is* **weakly edge-transitive** *if for all e_1, e_2 in $E(G)$, there exists an edge automorphism π such that $\pi e_1 = e_2$.*

Alternatively, one may say that G is weakly edge-transitive if and only if the line graph of G is vertex-transitive. It is important to note that this is not equivalent to the standard definition of edge-transitivity, defined through vertex automorphisms. Indeed if G is edge-transitive it is certainly weakly edge-transitive, as every vertex automorphism induces an edge automorphism. The converse is false, as $K_3 \cup K_{1,3}$ is weakly edge-transitive, but not edge-transitive.

Take note of the fact that if π is an edge automorphism of G and $t \in T(G)$ then πt will induce a spanning tree in G as well. Moreover, if $\pi e_1 = e_2$ then $t(e_1) = t(e_2)$. This yields the following useful fact.

Fact 2.4 *If a graph G is weakly edge-transitive, the densities of all its edges are equal.*

Before proceeding, one should note that the converse of the above theorem is in general false. Indeed, simply examine two copies of K_3 joined at one vertex. Every edge has density $\frac{2}{3}$, but this graph is clearly not weakly edge transitive. However, combining Fact 2.4 with Lemma 2.2 we get another practical result.

Corollary 2.5 *If G is a weakly edge-transitive graph with n vertices and m edges, then for any e in $E(G)$,*

$$d_t(e) = \frac{n-1}{m}.$$

This corollary can be applied to a variety of common graphs. Although the following results are well known, the techniques employed thus far allow simple proofs requiring very little machinery. As both K_n and $K_{m,n}$ are edge-transitive, they are weakly edge-transitive so we may apply Corollary 2.5.

Corollary 2.6 *If $e \in E(K_n)$, then*

$$d_t(e) = \frac{2}{n}.$$

Corollary 2.7 *If $e \in E(K_{j,k})$, then*

$$d_t(e) = \frac{j+k-1}{jk}.$$

3 Spanning Tree Edge Dependence

Very often in the area of network vulnerability, we seek out the "worst-case" scenario. The following definition panders to that trend.

Definition 3.1 *The* **spanning tree edge-dependence** *, or simply* **dependence** *for our purposes, is given by*

$$dep(G) = max(d_t(e)), \qquad e \in E(G).$$

The dependence of a graph is the "most damage" that a hostile party could inflict (in terms of number of spanning trees) by destroying a single edge. It is interesting to note that $dep(C_n) \xrightarrow{n} 1$ and $dep(K_n) \xrightarrow{n} 0$. It may be of interest to investigate the asymptotic behavior of the dependences of other common families of graphs.

There is a natural link between the dependence of a graph and its edge connectivity. Indeed, let C be an edge cut-set in G and take note of the fact that every spanning tree of G must contain at least one edge from C. This allows us to give a useful bound on the dependence.

Theorem 3.2 *Let G be a graph. Then*

$$dep(G) \geq \frac{1}{\lambda(G)}$$

with equality holding if and only if $\lambda(G) = 1$.

PROOF: Let $\{e_1, e_2, \ldots, e_{\lambda(G)}\}$ be an edge cut-set in G. The above remark implies that

$$|\bigcup_{i=1}^{\lambda(G)} T(e_i)| = t(G).$$

However,

$$\sum_{i=1}^{\lambda(G)} t(e_i) = \sum_{i=1}^{\lambda(G)} |T(e_i)| \geq |\bigcup_{i=1}^{\lambda(G)} T(e_i)|,$$

so

$$\sum_{i=1}^{\lambda(G)} d_t(e_i) \geq 1.$$

Thus, as all edges are nonnegative rational numbers from $[0, 1]$, at least one of the e_i has density $\frac{1}{\lambda(G)}$ or greater.

If G contains a bridge b then we can see that $dep(G) = d_t(b) = 1$ so equality holds. It is easy to show that when $\lambda(G) > 1$ there exist spanning trees of G that contain two or more of the e_i. In this case,

$$\sum_{i=1}^{\lambda(G)} d_t(e_i) > 1$$

so at least one of the e_i has density strictly greater than $\frac{1}{\lambda(G)}$. □

4 Monotonicity

It is reasonable to investigate the behavior of reliability criterion when the graph is modified, for example when edges or vertices are added. In our case the addition of edges is of a great deal more interest, although the addition of vertices may be of interest to study in the future. We would like to know how the densities of edges behave when we either add new edges to a graph or delete existing edges. In order to do so, we first need to examine how subsets of edges behave with respect to spanning trees.

4.1 Densities of Subsets

Let S be a subset of $E(G)$ and define $T(S;G)$ to be the set of all spanning trees of G containing all of S. As above, let $t(S;G) = |T(S;G)|$.

Definition 4.1 *Let S be a subset of $E(G)$. The* **spanning tree edge density** *of S, denoted $d_t(S;G)$ is given by*

$$d_t(S;G) = \frac{t(S;G)}{t(G)}.$$

The next theorem is from Cayley, and can be found in [1].

Theorem 4.2 *Let T_k be any tree on k vertices. Then for any $n \geq k$*

$$d_t(T_k; K_n) = \frac{k}{n^{k-1}}.$$

It is also possible to relate the density of a pair of edges to the densities of the individual edges, as was shown in [6]. In the language of probability theory, which integrates nicely with the concept of densities, the next claim shows that densities are negatively correlated.

Theorem 4.3 (Pemantle, [6]) *Let G be a graph and let e, f be in $E(G)$. Then*

$$d_t(e,f) \leq d_t(e)d_t(f).$$

4.2 Densities are Monotone Decreasing

We would like to answer the following question:

Question 4.4 *Suppose G is a graph and H is a spanning subgraph of G such that $e \in E(H) \cap E(G)$. Is it then true that*

$$d_t(e;G) \leq d_t(e;H)?$$

Note that if e is a bridge in G or H then the answer is yes. First, we state a straightfoward number theoretic lemma that is provable using elementary algebra.

Lemma 4.5 *Let $a \leq b$ and $x \leq y$ be integers such that b and y are positive, $x < a$ and $y < b$. Then if $\frac{x}{y} \leq \frac{a}{b}$,*

$$\frac{a-x}{b-y} \geq \frac{a}{b}.$$

To answer Question 4.4, consider the following. Let e and f be distinct edges in G, where f is not a bridge and note that

$$d_t(e : G - f) = \frac{t(e;G) - t(e,f;G)}{t(G) - t(f;G)},$$

or equivalently,

$$d_t(e;G-f) = \frac{d_t(e;G) - d_t(e,f;G)}{1 - d_t(f;G)}.$$

Thus, Theorem 4.3 and Lemma 4.5 immediately imply the following.

Theorem 4.6 *Let H be a graph and let G be obtained from H by adding any number of edges. Then for any e in $E(H)$,*

$$d_t(e;G) \leq d_t(e;H).$$

Hence spanning tree edge densities are monotone decreasing with respect to adding edges.

It is not difficult to show, through a similar argument, that adding copies of existing edges to G will not increase the density of any edge. This does not imply, however, that adding edges to G will decrease the dependence of G, as we are unsure of the relationship between the densities of the previous edges and the density of the new edge. It would also be interesting to investigate the monotonicities of dependences.

One consequence of the above statement, when viewed together with Corollary 2.6, is that no simple graph can have an edge with density less than $\frac{2}{|V(G)|}$. The question is then, how quickly does the density of an edge approach this value? We can, at this time, provide only a partial answer.

Theorem 4.7 *Let G be a simple graph and let $|V(G)| = n$. If there exist $x, y \in V(G)$ such that $N(x) = V(G) - \{x\}$ and $N(y) = V(G) - \{y\}$ then $d_t(xy) = \frac{2}{n}$.*

Proof: Let $S = G - \{x,y\}, e = xy$ and note that every spanning tree in G induces a spanning forest in $H = \langle S \rangle$. We shall thus partition $T(G)$ and $T(G;e)$ over all spanning forests of H.

Consider such a forest $\mathcal{F}$ and let $c_1, \ldots, c_p$ denote the components of $\mathcal{F}$ having orders $n_1, \ldots, n_p$. Moreover, let $N = n_1 n_2 \ldots n_p$ and $T_{\mathcal{F}}(G;e)$ and $T_{\mathcal{F}}(G)$ denote the subsets of $T(G)$ and $T(G;E)$ respectively that induce $\mathcal{F}$ in H. Now observe that each tree in $T_{\mathcal{F}}(G;e)$ has the property that either x or y, but not both, is adjacent to exactly one vertex in each C_i. Thus

$$|T_{\mathcal{F}}(G;e)| = 2^p N.$$

On the other hand, each tree in $T_{\mathcal{F}}(G) - T_{\mathcal{F}}(G;e)$ has the property that both x and y are adjacent to exactly one C_t and exactly one of x or y is adjacent to each of the other C_i's. Thus

$$|T_{\mathcal{F}}(G) - T_{\mathcal{F}}(G;e)| = \sum_{t=1}^{p} 2^{p-1} n_t^2 \frac{N}{n_t} = 2^{p-1}(n-2)N.$$

Hence

$$|T(G;e)| = \sum_{\mathcal{F}} |T_{\mathcal{F}}(G,e)| = \sum_{\mathcal{F}} 2^p N,$$

and

$$\begin{aligned} |T(G)| &= \sum_{\mathcal{F}} |T_{\mathcal{F}}(G)| \\ &= \sum_{\mathcal{F}} \left(2^p N + 2^{p-1}(n-2)N\right) \\ &= \frac{n}{2} \sum_{\mathcal{F}} 2^p N. \end{aligned}$$

The result follows. □

5 Open Problems

In this final section we give two open problems that may be of interest.

5.1 Constructibility

One can show [4] that given any rational number r in $[0,1]$ there exists a multigraph that has an edge with density r. The question then arises, can one always find a simple graph that contains an edge with density r?

The above formulas show that we can construct a number of densities using familiar graphs. Unfortunately, these densities all have a tendency to fall near the boundary of $[0,1]$. With a few exceptions, we have viewed densities as a global property. We believe that progress could be made on this problem, and others cited above, by taking a more local approach. For instance, an examination of the densities of induced subgraphs may allow for a better handle on the densities of the edges therein.

5.2 Realizability

Another problem related to that of constructing densities is trying to determine when rationals occur together as densities in a graph. We say a sequence of rational numbers, $S : r_1, r_2, \ldots, r_m$ is *realizable* (with respect to spanning tree edge densities) if there exists a graph G with $E(G) = \{e_1, \ldots, e_m\}$ such that $d_t(e_i) = r_i$. We say that such a graph *realizes* S. The primary goal would be to find necessary and sufficient conditions for a sequence to be realizable.

From above results, the necessity of several conditions is clear. First, a realizable sequence should sum to an integer, specifically if G realizes S and has n vertices, then $\sum r_i$ should equal $n - 1$. It should also be clear that if some $r_i = 0$ then S is realizable if and only if it is the zero sequence. Additionally, note that contracting a bridge in G will not change any of the densities of the other edges. Hence, if any $r_i = 1$ and S is realizable, then so too is $S - \{r_i\}$.

In closing, it should be noted that there are a lot of potentially interesting questions connected to the idea of densities. It is our hope that this paper will pique the interest of some of its readers so that some progress will be made.

References

[1] Berge, Claude, **Graphs and Hypergraphs**, ©1973, North Holland Publishing Company, pp 44-46

[2] Biggs, Norman, **Algebraic Graph Theory** Second edition, *Cambridge Mathematical Library*, Cambridge University Press, 1993.

[3] Chartrand, Gary and Lesniak, Linda, **Graphs and Digraphs** 2^{nd} edition, CRC Press, 1996.

[4] Kahl, Nathan, Private Transmission, April 2000

[5] Myrvold, Wendy, et. al., *Finding the Most Vital Edges with Respect to the Number of Spanning Trees*, IEEE Transactions on Reliability, vol 43, no. 4, December 1994.

[6] Pemantle, Robin, *Uniform Random Spanning Trees*, **Topics in Contemporary Probability and its Applications**, Probability and Stochastics series, CRC Press, 1995, pp 1-54