When Bad Things Happen to Good Trees

Manuel Aivaliotis, Gary Gordon, and William Graveman

LAFAYETTE COLLEGE
EASTON, PENNSYLVANIA 18042
E-mail: gordong@lafayette.edu

Received May 26, 1999; revised June 13, 2000

Abstract: When the edges in a tree or rooted tree fail with a certain fixed probability, the (greedoid) rank may drop. We compute the expected rank as a polynomial in p and as a real number under the assumption of uniform distribution. We obtain several different expressions for this expected rank polynomial for both trees and rooted trees, one of which is especially simple in each case. We also prove two extremal theorems that determine both the largest and smallest values for the expected rank of a (rooted or unrooted) tree, and precisely when these extreme bounds are achieved. We conclude with directions for further study. J Graph Theory 37: 79–99, 2001

Keywords: *antinatroid; expected rank polynomial*

1. INTRODUCTION

No one seems to notice systems when they are operating normally, but when something bad happens, everyone becomes interested. Suppose the edges in some network are working, but a sudden power surge, or earthquake, or tornado, or asteroid,. . . disables some of these edges. What is the expected size of the surviving network?

This problem has been well studied in many different contexts within combinatorics. Indeed, it is probably not an overstatement to say this problem has

Correspondence to: Gary Gordon
With apologies to Harold S. Kushner.

motivated much of reliability theory. For example, k-terminal reliability problems seek to determine the probability that k-specified terminals can communicate after something bad has happened to the network. A standard source for reliability theory is [7].

Traditional approaches to the problem concentrate on the number and size of the connected components of the surviving graph. However, if the graph is rooted at some distinguished vertex (a cable television network in which the root is the cable provider, for example), then the most important information may not be the size or number of components, but the size of the component containing the root. We will see that this information can be obtained from the greedoid rank function for the rooted graph. Thus, the setting for this paper is the crossroads of reliability theory and the interpretation of trees as greedoids. Although we won't use greedoids explicitly, they form the background for the results in this paper.

Although unrooted trees have probably been studied more extensively than rooted trees, it is more difficult to motivate a reliability interpretation in the unrooted case. We propose the following scenario as one possible application. Suppose several remote sensors are gathering information (for a scientific survey of Mars or a remote volcano or perhaps for a spying mission) and then relaying that information to each other. Some sensors may not be able to communicate with others directly (because of distances involved or other considerations), so a network is constructed with the sensors as the vertices of a tree. Further, the most remote sensors (i.e., the sensors corresponding to leaves in the tree) are the sensors we will access periodically, as these sensors are most accessible. When some of the communications (edges of the tree) break down, we are concerned with how far the information can be passed along from the leaf-sensors to the interior of the network, and vice versa. Our treatment of trees uses this interpretation for the rank of a subset of edges of a tree. This notion of rank is based on the *complements* of subtrees. (This is the rank function of the *pruning greedoid* associated with the tree.)

Generating statistics for trees and rooted trees has been done in a somewhat different context by Jamison. In a series of four papers [15, 16, 17, 18], Jamison computes several means of interest and proves some extremal results as well. While our approach is through probability, many of our results have a similar flavor to this interesting work.

Since rooted trees are easier to deal with than unrooted trees, we concentrate on them first. The polynomials we consider, $R(T)$ and $\bar{R}(T,p)$, give the expected rank of a tree or rooted tree. (In $R(T)$, each edge e is assumed to have a probability p_e of succeeding, while $\bar{R}(T,p)$ is a standard evaluation of $R(T)$ formed by assuming $p_e = p$ for all edges e.) When the tree is rooted, these polynomials have very simple combinatorial interpretations. In particular, Theorem 2.4 gives the following formula:

$$R(T) = \sum_{v \in V} \prod_{e \in P(v)} p_e,$$

where $P(v)$ is the unique path from the root to the vertex v. Thus, this expected-value polynomial is simply a generating function for paths from the root to the other vertices of T. Using this result, we show how to reconstruct the rooted tree T from this polynomial. Amin, Siegrist, and Slater [3] prove a similar result for the *pair-connected reliability* of a tree.

Section 3 is concerned with unrooted trees. Compared with the previous section, the results and their proofs are a bit more complicated. The main theorem, Theorem 3.3, again gives a rather simple form for the expected value polynomial $R(T)$. This form, which has (essentially) two terms for each edge of T, also has some interesting corollaries. It is again true (although somewhat more difficult to prove) that the polynomial $R(T)$ determines the tree T (Corollary 3.5). Our proof gives a recursive algorithm for reconstructing T from $R(T)$.

We use standard probabilistic interpretations in Sec. 4 to get numerical values for the expected rank of a tree or rooted tree. To do this, we assume p is a uniformly distributed random variable. The main results in this section are extremal theorems (Propositions 4.2 and 4.3). These results give upper and lower bounds for the expected value and determine precisely when these bounds can be achieved. Similar results hold for pair-connected reliability in rooted trees (Theorems 3 and 4 of [3]).

We conclude with several directions for further study in Sec. 5. Some of the suggested areas of research may have more immediate application than this work, which is more concerned with the combinatorial structure of the polynomial invariants associated with reliability.

2. ROOTED TREES

Let T be a rooted tree, rooted at *, with edge set E. Let $\mathcal{F}$ denote the subsets of E which are rooted subtrees of T. (These are the *feasible* sets of the associated rooted branching greedoid.) The *rank* of a subset $S \subseteq E$ is given by

$$r(s) = \max_{A \subseteq S}\{|A| : A \in \mathcal{F}\}.$$

Note that $r(E) = |E|$. We also remark that the maximum-size subtree A of a subset S is unique—this will be important in simplifying some of our formulas. (This is true because the union of feasible sets is always feasible—this property characterizes antimatroids.)

Our probabilistic interpretation is straightforward. Assume each edge e succeeds with probability p_e (and fails with probability $(1 - p_e)$). This expected rank of T is then a polynomial in the edge probabilities:

Definition 2.1. The *expected-rank polynomial* $R(T)$ of the rooted tree T is given by

$$R(T) = \sum_{S \subseteq E} r(S) \prod_{e \in S} p_e \prod_{e \notin S} (1 - p_e).$$

Let $M(F) = \{e \in T - F : F \cup \{e\}$ is a subtree of $T\}$, i.e., $M(F)$ are the edges which can be added to the subtree F. The next proposition gives a simpler expression for the expected rank polynomial $R(T)$.

Proposition 2.2.

$$R(T) = \sum_{F \in \mathcal{F}} |F| \prod_{e \in F} p_e \prod_{e \in M(F)} (1 - p_e).$$

Proof. To simplify notation, let $a(S)$ denote the contribution the subset of edges S makes to $R(T)$ in the definition:

$$a(S) = r(S) \prod_{e \in S} p_e \prod_{e \notin S} (1 - p_e).$$

In this same spirit, let $b(F)$ represent the contribution the subtree F makes to the sum on the right-hand side in the statement of the proposition:

$$b(F) = |F| \prod_{e \in F} p_e \prod_{e \in M(F)} (1 - p_e).$$

We also let $T(S)$ be the unique maximum-size subtree of S.

Now let F be a rooted subtree of T and note that the proposition would follow from showing $\sum_{S:T(S)=F} a(S) = b(F)$. But

$$\begin{aligned} \sum_{S:T(S)=F} a(S) &= |F| \prod_{e \in F} p_e \prod_{e \in M(F)} (1 - p_e) \prod_{e \notin F \cup M(F)} (p_e + (1 - p_e)) \\ &= |F| \prod_{e \in F} p_e \prod_{e \in M(F)} (1 - p_e) \\ &= b(F) \end{aligned}$$

since any edge $e \notin F \cup M(F)$ contributes both p_e and $1 - p_e$ to each subset S having $T(S) = F$. This completes the proof. ■

We assume $E = \{e_1, \ldots, e_n\}$ and write p_i for $p(e_i)$. When $p_i = p$ for all edges e_i, $R(T)$ becomes polynomial in p which we denote $\bar{R}(T; p)$:

$$\bar{R}(T; p) = \sum_{S \subseteq E} r(S) p^{r(S)} (1 - p)^{|E| - r(S)}.$$

Then Proposition 2.2 immediately yields a simpler expression for $\bar{R}(T; p)$, too.

Corollary 2.3.

$$\bar{R}(T; p) = \sum_{F \in \mathcal{F}} |F| p^{|F|} (1 - p)^{|M(F)|}.$$

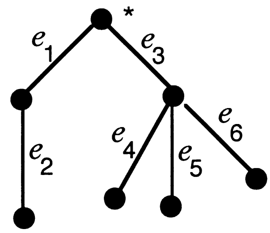

FIGURE 1. A rooted tree.

The polynomial $\bar{R}(T,p)$ has been studied before for ordinary graphs. In Sec. 2 of [4], a deletion–contraction recursion is established and the coefficient of the leading term of the polynomial is shown to be equal to $\pm\beta(G)$. We remark that this recursion remains valid for rooted graphs as well.

As an example, we compute $R(T)$ and $\bar{R}(T,p)$ for the rooted tree shown in Fig. 1. Using the definition or Proposition 2.2 and its corollary, we have $R(T) = p_1 + p_3 + p_1p_2 + p_3p_4 + p_3p_5 + p_3p_6$ and $\bar{R}(T,p) = 2p + 4p^2$.

The simple form for these polynomials suggests that a simpler expansion underlies the formulas given in 2.2 and 2.3. The next result shows that this is true. Recall that each vertex in a rooted tree is joined by a unique path to the root. Let $P(v)$ denote this path. We now prove the theorem.

Theorem 2.4. *Let T be a rooted tree. Then*

$$R(T) = \sum_{v\in V} \prod_{e\in P(v)} p_e.$$

Proof. For each nonroot vertex $v \in V$, we let $I(v)$ be an indicator function for whether v is reachable from the root using the surviving edges. This $I(v) = 0$ if there is no path connecting v to the root and $I(v) = 1$ if there is such a path. Let $Pr(v)$ denote the probability that v is reachable from the root. It is immediate that $E(I(v)) = Pr(v)$, where E is the expected-value operator. Furthermore, it is clear that $Pr(v) = \prod_{e\in P(v)} p_e$. Then

$$R(T) = E\left(\sum_{*\neq v\in V} I(v)\right) = \sum_{*\neq v\in V} E(I(v)) = \sum_{*\neq v\in V} Pr(v) = \sum_{v\in V} \prod_{e\in P(v)} p_e. \qquad \blacksquare$$

The proof we have given for Theorem 2.4 is essentially due to Amin, Siegrist, and Slater [3]. Their work assumes each edge has the same probability p of succeeding, as in Corollary 2.6. We remark that a combinatorial proof is also straightforward.

Recall that the definition of $R(T)$ involved an expansion via subsets of E—a calculation involving $2^{|E|}$ terms. Proposition 2.2 offers an improvement—the calculation uses the subtrees instead of the subsets. Unfortunately, this calculation will (in general) also be exponential in $|E|$. Theorem 2.4 gives a much more efficient way to compute $R(T)$ since we can determine all the paths in T in polynomial time.

Corollary 2.5. *Let T_1 and T_2 be rooted trees. Then $R(T_1) = R(T_2)$ if and only if T_1 and T_2 are isomorphic as labeled trees.*

Proof. We show that T can be reconstructed from $R(T)$. By Theorem 2.4, $R(T)$ gives a list of all the paths of T adjacent to *. Then all edges e_i adjacent to * are paths of length 1 (and hence appear as degree 1 monomials p_i in $R(T)$), and so these edges can be reconstructed, with labels. All edges e_j adjacent to these edges appear in $R(T)$ as degree 2 monomials $p_i p_j$ where e_i is adjacent to *. Thus, the labeled paths of length 2 can be reconstructed. This process can be continued until all edges are labeled and uniquely placed in T, and this completes the proof. ■

Let $d(*, v)$ be the distance from the root * to the vertex v.

Corollary 2.6. $\bar{R}(T; p) = \sum_{v \in V} p^{d(*,v)}$.

If T has n edges, then (from the corollary) $\bar{R}(1) = n$. This corresponds to the trivial case in which every edge survives, so the expected rank equals n. We can also use the last result to create nonisomorphic trees with the same rank polynomial. For example, let T_1 and T_2 be the trees in Fig. 2. Then $\bar{R}(T_1) = \bar{R}(T_2) = 2p + 2p^2$.

The *direct sum* $T_1 \oplus T_2$ of two rooted trees T_1 and T_2 is formed by identifying the two roots $*_1$ and $*_2$ of the respective trees. The next result follows immediately from the theorem.

Corollary 2.7. $R(T_1 \oplus T_2) = R(T_1) + R(T_2)$.

Another proof of Corollary 2.7 can be formulated as follows. Let

$$F(T) = \sum_{S \subseteq E} Pr(S) x^{r(S)},$$

where

$$Pr(S) = \prod_{e \in S} p_e \prod_{e \notin S} (1 - p_e).$$

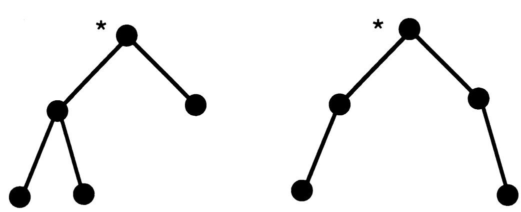

FIGURE 2. $\bar{R}(T_1) = \bar{R}(T_2) = 2p + 2p^2$.

Then it is clear from the definition of rank that $F(T_1 \oplus T_2) = F(T_1) + F(T_2)$. The proof follows by differentiating this equation with respect to x and then evaluating at $x = 1$.

3. UNROOTED TREES

We now turn our attention to unrooted trees. Trees are among the most studied classes of graphs, in part because they are among the simplest graphs that exhibit deep and interesting behavior. They are also extremely useful in modeling all sorts of systems, when there is no distinguished vertex. To apply the tools developed for rooted trees to unrooted trees, we need a good definition of rank for subsets of edges. As we did with rooted trees, we use a greedoid rank function.

Let T be the collection of subtrees of T and let $\mathcal{F}$ be the collection of all subtree *complements* of T. (The subtree complements are the feasible sets of the *pruning* greedoid associated with T.) Then the *rank* of a subset $S \subseteq E$ is given by

$$r(S) = \max_{A \subseteq S} \{|A| : A \in \mathcal{F}\}.$$

It is still true $r(E) = |E|$ and the maximum-size subtree *complement* A of the subset S is unique. (We use complements of subtrees instead of the subtrees themselves to preserve the antimatroid property. The union of subtree complements is a subtree complement.)

We can now define the polynomials $R(T)$ and $\bar{R}(T, p)$ exactly as before;

$$(1) \quad R(T) = \sum_{S \subseteq E} r(S) \prod_{e \in S} p_e \prod_{e \notin S} (1 - p_e)$$

$$(2) \quad \bar{R}(T; p) = \sum_{S \subseteq E} r(S) p^{r(S)} (1 - p)^{|E| - r(S)}.$$

Proposition 2.2 and Corollary 2.3 have analogs in the unrooted case. We state these results without proof; the proof of Proposition 3.1 is similar to the proof of Proposition 13(b) of [6].

Proposition 3.1. *Let T be a tree, $L(F)$ be the set of edges which are leaves of the subtree F, and let $\mathcal{T}$ be the collection of all subtrees of T. Then*

$$R(T) = \sum_{F \in \mathcal{T}} |E - F| \prod_{e \in E - F} p_e \prod_{e \in L(F)} (1 - p_e).$$

Corollary 3.2. $\bar{R}(T; p) = \sum_{F \in \mathcal{T}} |E - F| p^{|E - F|} (1 - p)^{|L(F)|}$.

Our main theorem for trees gives a much simpler expression for $R(T)$. As in Theorem 2.4, the new representation of $R(T)$ is linear in the number of edges of T

(instead of the exponential number of terms in the definition (1) or Proposition 3.1). The theorem will also allow us to prove that $R(T)$ is a complete invariant, i.e., nonisomorphic trees have different $R(T)$ polynomials (Corollary 3.5).

When an edge e that is incident to vertices v and w is deleted from a tree T, the tree is separated into two components. Call these components $C_e(v)$ and $C_e(w)$ and note that one of these components will have no edges when e is a leaf of T.

Theorem 3.3. *Let T be an unrooted tree with n edges and l leaves and leaf set $L(T)$. Write $p_S = \prod_{e\in S} p_e$. Then*

$$\begin{aligned} R(T) &= \left(\sum_{e\notin L(T)} p_e\big(p_{C_e(v)} + p_{C_e(w)}\big) + \sum_{e\in L(T)} p_e \right) - (n-l)p_T \\ &= \left(\sum_{e\in E(T)} p_e\big(p_{C_e(v)} + p_{C_e(w)}\big) \right) - np_T. \end{aligned}$$

Proof. The second equality follows from the first since, if e is a leaf and v is the vertex of degree 1 incident to e, $C_e(v) = \emptyset$, and $C_e(w) = T - \{e\}$. The proof of the first equality is similar to a combinatorial proof of Theorem 2.4 in that the key step is reversing the sum in the expansion for $R(T)$ given in Proposition 3.1.

$$\begin{aligned} R(T) &= \sum_{F\in\mathcal{T}} |E-F|_{p_{E-F}} \prod_{e\in L(F)} (1-p_e) \\ &= \sum_{F\in\mathcal{T}} |E-F| \sum_{S\subseteq L(F)} (-1)^{|S|} p_{E-(F-S)} \\ &= \sum_{\emptyset\neq F\in\mathcal{T}} p_{E-F} \sum_{S\subseteq M(F)} (-1)^{|S|} |E-(F\cup S)| + h(T) \\ &= \sum_{\emptyset\neq F\in\mathcal{T}} p_{E-F} \sum_{k=0}^{m} (-1)^k \binom{m}{k} (n-f-k) + h(T), \end{aligned}$$

where $f = |F|$, $m = |M(F)|$ and $h(T)$ is the contribution to the sum made when F is empty. The last equality follows from a result in lattice theory; the lattice of all subtrees of a tree is meet-distributive (so the interval $[F, F\cup M(F)]$ is boolean), and adding an edge in $M(F)$ and deleting an edge in $L(F)$ are inverse operations.

To complete the proof, we need to compute $Z = \sum_{k=0}^{m} (-1)^k \binom{m}{k} (n-f-k)$ as well as $h(T)$. But $Z = 0$ unless $m = 1$. It is a routine exercise to see that $|M(F)| = 1$ iff $F = C_e(v)$ for some edge e. When e (with vertices v and w) is not a leaf of T, there are two subtrees F with $M(F) = \{e\} : F = C_e(v)$ or $F = C_e(w)$. When e is a leaf, $M(F) = \{e\}$ can only occur when $F = T - \{e\}$.

It remains to be proved that $h(T) = -(n-l)p_T$. We first note that a subtree F will contribute to the coefficient of the term p_T in $R(T)$ iff F is empty or every edge of F is a leaf of F, i.e., iff $F \subseteq D(v)$ for some vertex v, where $D(v)$ is the subtree consisting of all edges incident to the vertex v. Summing over all subsets of $D(v)$ for all vertices v will include all these contributions, but it will count each edge (when $F = \{e\}$) twice (once for each vertex of e) and $\emptyset$ will be counted $n+1$ times (once for each vertex). The contribution of a single edge is (-1) $(n-1)$ and $\emptyset$ contributes n. Thus

$$h(T) = p_T\left((n-1)n - n^2 + \sum_{v\in V}\sum_{F\subseteq D(v)}(-1)^{|F|}(n-|F|)\right)$$
$$= p_T\left(-n + \sum_{v\in V}\sum_{k=0}^{d(v)}(-1)^k\binom{d(v)}{k}(n-k)\right),$$

where $d(v)$ is the degree of the vertex v. As before, $\sum_{k=0}^{d(v)}(-1)^k \binom{d(v)}{k}$ $(n-k) = 0$ unless $d(v) = 1$, in which case the sum equals 1. Thus, this sum contributes 1 if e is a leaf and 0 otherwise. This gives $h(T) = (l-n)p_T$ and completes the proof. ■

The invariant $l-n$ which appears as the coefficient of p_T in $R(T)$ is the β invariant of the tree. This invariant is associated with the number of "internal elements" of the combinatorial object under consideration. See [1, 8] for a relationship with finite subsets of $\Re^n$ and [11, 14] for the connection with trees.

We also remark that it is possible to formulate a probabilistic proof of Theorem 3.3, as was given for Theorem 2.4. To do so, define an indicator function on edges (instead of vertices) and note that (under the suitable assumptions) $Pr(e) = p_{C_e(v)} + p_{C_e(w)} - p_T$.

As in the rooted case, the formula of Theorem 3.3 has several corollaries. We first prove that $R(T)$ uniquely determines the labeled tree. Before providing this result, we need a simple lemma. A *star* is a tree in which every edge is a leaf, i.e., all the edges are incident to one vertex.

Lemma 3.4. *Let T be a tree which is not a star. Then there is a vertex v and a collection of edges $e_1, \ldots e_m, f$ which are all incident to v such that e_i is a leaf for $1 \leq i \leq m$ and one of the components of $T - \{f\}$ is $e_i, \ldots, e_m$.*

Proof. Remove all of the leaves from T and let v be any vertex of degree one in $T - L(T)$. Then there is an edge f that is not a leaf of T that is incident to v. Thus, in T, the vertices incident to v are $e_1, \ldots, e_m, f$ and clearly one of the components of $T - \{f\}$ is $e_1, \ldots, e_m$. ■

Corollary 3.5. *Let T_1 and T_2 be unrooted trees. Then $R(T_1) = R(T_2)$ if and only if T_1 and T_2 are isomorphic as labeled trees.*

Proof. As in Corollary 2.5, we show that T can be reconstructed from $R(T)$. By Theorem 3.3, the monomials of $R(T)$ give a list of all the labeled subtree complements having $|M(F)| = 1$. In particular, we can uniquely recover all the labeled leaves of T. If T is a star, then every edge is a leaf and we are done. Otherwise, by Lemma 3.4, there is a vertex v, a collection of leaves $e_1, \ldots, e_m$, and an edge f such that $C_f(v) = \{e_1, \ldots, e_m\}$. Then the product $p_{e_1} \cdots p_{e_m} p_f$ appears as a monomial term in $R(T)$ (where $e_1, \ldots, e_m$ have already been identified as leaves of T).

We now show how $R(T - \{e_1, \ldots, e_m\})$ can be obtained from $R(T)$. The result then follows by induction. First remove all the terms of the form p_{e_i} (for $1 \leq i \leq m$) from $R(T)$ – there will be one such term for each e_i since each leaf of T so appears in $R(T)$. Now Theorem 3.3 implies each remaining monomial of $R(T)$ either contains the entire product $p_{e_1} \cdots p_{e_m}$ as a factor or contains none of the p_{e_i} as factor. For each monomial of $R(T)$ containing $p_{e_1} \cdots p_{e_m}$ as a factor, we delete $p_{e_1} \cdots p_{e_m}$ from the monomial, and we leave unchanged the monomials that do not have any p_{e_i} as a factor (for $1 \leq i \leq m$). Call the new polynomial that results from this process $S(T)$.

We claim that $S(T) = R(T - \{e_1, \ldots, e_m\})$. The result then follows from the claim since we can inductively reconstruct the labeled tree $T - \{e_1, \ldots, e_m\}$ and then reattach $e_1, \ldots, e_m$ to f. To verify the claim, consider $C_e(u)$ computed in T and in $T - \{e_1, \ldots, e_m\}$. Either these sets are identical (if none of the e_i are in $C_e(u)$) or they differ precisely by the set $\{e_1, \ldots, e_m\}$. But $S(T)$ was created so that the corresponding terms match $C_e(u)$ exactly in $T - \{e_1, \ldots, e_m\}$. Further, the term corresponding to all of $T - \{e_1, \ldots e_m\}$ in $S(T)$ will have the correct coefficient; the number of internal edges of $T - \{e_1, \ldots, e_m\}$ is 1 less than the number of internal edges of T (since f is no longer internal), and the construction of $S(T)$ adjusts the coefficient of $P_{T-\{e_1,\ldots,e_m\}}$ accordingly. This completes the proof. ∎

We can use the inductive proof to obtain a recursive procedure for reconstructing T from $R(T)$. As in the proof, first identify all leaves of T from $R(T)$, then find a monomial having form $p_{e_1} \cdots p_{e_m} p_f$, where $e_1, \ldots, e_m$ are leaves (and f is not a leaf). (If all the edges of T are leaves, then T is a star and the reconstruction is trivial.) Now modify the polynomial as in the proof and iterate the process. We demonstrate the procedure with an example.

Suppose

$$R(T) = \sum_{i=1}^{6} p_i + p_4p_5p_7 + p_4p_5p_6p_7p_8 + p_1p_2p_3p_8 + p_1p_2p_3p_6p_7p_8 - 2\prod_{i=1}^{8} p_i.$$

Then there are 8 edges and e_i is a leaf for $1 \leq i \leq 6$. Now the term $p_1p_2p_3p_8$ gives a vertex in T which is adjacent to just e_1, e_2, e_3 and e_8. We form the derived polynomial $S(T)$ as in the proof:

$$S(T) = p_4 + p_5 + p_6 + p_8 + p_4p_5p_7 + p_6p_7p_8 - \prod_{i=4}^{8} p_i.$$

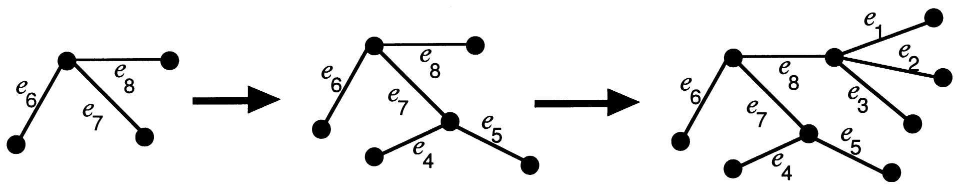

FIGURE 3. Reconstructing T from $R(T)$.

(Note that the coefficient of $\prod_{i=4}^{8} p_i$ changes from -2 to -1 in this process.) Now repeat the procedure using the term $p_4p_5p_7 : S(S(T)) = p_6 + p_7 + p_8$. At this point the tree is a star and the process terminates. See Fig. 3 for the reconstruction.

Corollary 3.6. *Let T be an unrooted tree with n edges, l leaves and let $I(T)$ denote the interior (nonleaf) edges. Then*

$$\begin{aligned}\bar{R}(T,p) &= lp + \sum_{e\in I(T)} \left(p^{|C_e(v)|+1} + p^{|C_e(w)|+1} - p^n\right)\\ &= \left(\sum_{e\in E(T)} \left(p^{|C_e(v)|+1} + p^{|C_e(w)|+1}\right)\right) - np^n.\end{aligned}$$

As was the case with rooted trees the corollary gives $\bar{R}(1) = n$. This again corresponds to the trivial case in which every edge survives. We can also construct examples that show it is impossible, in general, to reconstruct a tree T from $\bar{R}(T,p)$. In Fig. 4, the two trees T_1 and T_2 share the same $\bar{R}$ polynomial:

$$\bar{R}(T_1,p) = \bar{R}(T_2,p) = 4p + p^2 + p^3 + p^4 + p^5 - 2p^6.$$

Thus, as in the rooted case, $R(T,p)$ does not uniquely determine T. (Note that this invariant does not even determine the degree sequence of T.)

We can modify this example to produce more pairs of trees with the same $\bar{R}$: Let T_1 be a tree with three interior vertices of degrees a, b, and c (with a, b, and $c > 1$ and $a < c$), as in Fig. 5. (A tree in which all of the nonleaf vertices are arranged on a single path is called a *caterpillar*.) Now let T_2 be another tree with

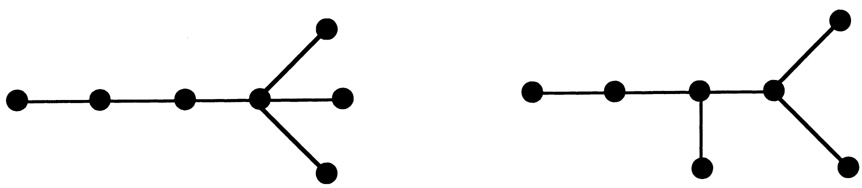

FIGURE 4. Two trees with $\bar{R}(T_1) = \bar{R}(T_2)$.

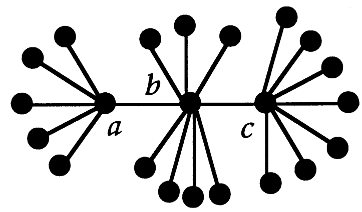

FIGURE 5. A caterpillar with three interior vertices.

three interior vertices of degrees a, $c-a+1$, and $a+b-1$, where the central vertex has degree $a+b-1$. (Each tree has $n=a+b+c-2$ edges.) Then

$$\bar{R}(T_1)=\bar{R}(T_2)=(n-2)p+p^a+p^c+p^{a+b-1}+p^{b+c-1}-2p^n.$$

4. EXPECTED VALUES FOR ROOTED AND UNROOTED TREES

Given a rooted or unrooted tree T, we can use the tools developed in Secs. 2 and 3 to associate an expected value for the rank of T. When $p_i=p$ and we assume p is uniformly distributed, the expected value $EV(T)$ of the rank polynomial $\bar{R}(T,p)$ is obtained from an integral:

$$EV(T)=\int_0^1 \bar{R}(T,p)dp.$$

This definition is valid for both rooted and unrooted trees and is consistent with the usual interpretation of expected value in probability.

Which rooted trees have the highest and lowest expected values under these assumptions? When do two nonisomorphic rooted trees T_1 and T_2 have the same expected value? What about unrooted trees? We explore these questions here, blending the discrete analysis of $\bar{R}(T,p)$ with continuous probability.

If T is a rooted tree, recall $M(F)$ is the set of edges that can be adjoined to a rooted subtree F. The formulas in the next proposition for $EV(T)$ when T is rooted follow from Corollaries 2.3 and 2.6, while the formulas in the unrooted case follow from Corollaries 3.2 and 3.6.

Proposition 4.1. *Let T be a rooted tree.*

(1) $EV(T)=\sum_{F\in\mathcal{F}}|F|\dfrac{|F|!|M(F)|!}{(|F|+|M(F)|+1)!}$;

(2) $EV(T)=\sum_{v\in V,v\neq *}\dfrac{1}{d(*,v)+1.}$

Let T be an unrooted tree with n edges.

$$(3) \quad EV(T) = \sum_{F \in T} (n - |F|) \frac{(n - |F|)!|L(F)|!}{(n - |F| + |L(F)| + 1)!};$$

$$(4) \quad EV(T) = \sum_{e \in E} \left(\frac{1}{|C_e(v)| + 2} + \frac{1}{|C_e(w)| + 2} \right) - \frac{n}{n+1}$$

As an application of formulas (2) and (4) of the proposition, we prove two extremal results. Proposition 4.2 is analogous to Theorems 3 and 4 of [3]. A *rooted path* is a rooted tree in which every nonleaf vertex has degree 2. A *rooted star* is a rooted tree in which every vertex is adjacent to the root.

Proposition 4.2. *Let T be any rooted tree with n edges. Then*

$$\sum_{k=2}^{n+1} \frac{1}{k} \leq EV(T) \leq \frac{n}{2}.$$

Furthermore, equality holds for the lower bound iff T is a rooted path and equality holds for the upper bound iff T is a rooted star.

Proof. Let A be the rooted path with n edges and let B be the rooted star. For the lower bound, we find a common labeling of the vertices of T and the vertices of A, showing that $d_T(*, v) \leq d_A(*, v)$ for all v. The result then follows from applying the formula (2) of Proposition 4.1 to T and A. To obtain a common labeling, assume that the vertices of T are already labeled $v_1, \ldots, v_n$ and note that there is a natural partial order on the vertices of $T : v_1 \leq v_2$ if the unique path from * to v_2 passes through v_1. Then label the vertices of A with the labels $v_1, \ldots, v_n$ so that this property is preserved: $v_1 \leq v_2$ in T implies v_1 is closer to * than v_2 in A. (This can always be done—we are just extending the partial order from T to a total order.) Then it is clear that $d_T(*, v_i) \leq d_A(*, v_i)$ for all i and the lower bound is established.

For the case of equality in the lower bound, note that we get equality in the preceding argument iff $d_T(*, v_i) = d_A(*, v_i)$ for all i. It is now easy to show inductively that T must be isomorphic to A in this case.

For the upper bound, we again find a common labeling of the vertices of T and the vertices of B, showing that $d_B(*, v) \leq d_T(*, v)$ for all v. But $d_B(*, v) = 1$ for all v, so any labeling of B will do. The result follows as above, including the case of equality. ∎

Proposition 4.3. *Let T be any unrooted tree with n edges. Then*

$$\sum_{k=2}^{n+1} \frac{2}{k} - \frac{n}{n+1} \leq EV(T) \leq \frac{n}{2}.$$

Furthermore, equality holds for the lower bound iff T is a path and equality holds for the upper bound iff T is a star.

Proof. Let A be the path with n edges and let B be the star. If e is any edge in a tree T with vertices v and w, let

$$h_T(e) = \frac{1}{|C_e(v)|+2} + \frac{1}{|C_e(w)|+2} - \frac{1}{n+1}$$

be the contribution e makes to $EV(T)$ in formula (4) of Proposition 4.1, so $EV(T) = \sum_{e \in E(T)} h_T(e)$.

We first establish the upper bound. Let $g : E(T) \to E(B)$ be any bijection between the edge sets of T and B. Note that $h_T(e) = \frac{1}{2}$ iff e is a leaf of T. Then

$$h_T(e) = \frac{1}{|C_e(v)|+2} + \frac{1}{|C_e(w)|+2} - \frac{1}{n+1} \leq \frac{1}{2} = h_B(g(e))$$

for all edges $e \in E(T)$. (The inequality is easily established by noting the maximum value of the function $F(x) = (x+2)^{-1} + (n+1-x)^{-1} - (n+1)^{-1}$ on the interval $[0, n-1]$ occurs at $x = 0$ or $x = n-1$.) The result now follows immediately from formula (4) of Proposition 4.1. Note that equality holds iff $h_T(e) = \frac{1}{2}$ for all edges e, i.e., iff T is a star.

For the lower bound, we first find a vertex u of degree $d \geq 2$ in T so that each subtree T_i (for $1 \leq i \leq d$) that is adjacent to u has at most $[\frac{n}{2}]$ edges. (It is easy to see that this is always possible.) Now let T' be the tree obtained from T by *straightening* each subtree T_i, i.e., T' is a tree in which every vertex, except possibly u, has degree at most 2 and the d subtrees adjacent to u are simply paths $P_1, \ldots, P_d$ of lengths $|E(T_1)|, \ldots, |E(T_d)|$. See Fig. 6 for an example.

Claim 1. $EV(T') \leq EV(T)$. To prove the claim, first consider the subtree T_1. Define a partial order on the edges of T_1 as follows: $x \leq y$ iff the unique path from u to the edge y passes through the edge x. Thus, edges "close to" the vertex u are less than edges farther away—this is equivalent to the partial order used in the proof of the lower bound in Proposition 4.2, using u as the root. This partial order

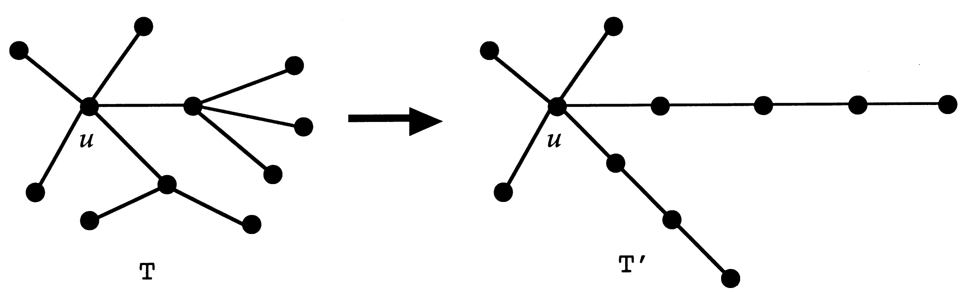

FIGURE 6. Constructing T' from T.

is well-defined for P_1 too—it yields a total order. Now let $g : E(T_1) \to E(P_1)$ be any order-preserving bijection. Thus, for example, x is adjacent to u in T_1 iff $g(x)$ is adjacent to u in P_1, and the leaf of P_1 corresponds to some leaf of T_1.

To complete the proof of the claim, we show that $h_{T'}(g(e)) \le h_T(e)$ for all $e \in T_1$. The full claim then follows by applying the same argument to T_i and P_i for all $i > 1$ and the fact that $EV(T) = \sum_{e \in E(T)} h_T(e)$. Let $e \in E(T_1)$ and let v and w be the two vertices adjacent to e. Assume w is closer to u than v is and also assume (to simplify notation) that v and w label the vertices adjacent to $g(x)$ in P_1, with w closer to u again. Then $|C_e(v)| \le |C_e(w)|$ in T_1 and $|C_{g(e)}(v)| \le |C_{g(e)}(w)|$ in P_1 because $|E(T_1)| = |E(P_1)| \le \lceil \frac{n}{2} \rceil$. However, as in the proof of Proposition 4.2, $|C_e(v)| \le |C_{g(e)}(v)|$ because g preserves order. Then the pair $\{|C_e(v)|, |C_e(w)|\}$ in T is more *imbalanced* than the pair $\{|C_{g(e)}(v)|, |C_{g(e)}(w)|\}$ in T'. Since $F(x) = (x+2)^{-1} + (n+1-x)^{-1} - (r+1)^{-1}$ is monotone decreasing on the interval $[0, \frac{n-1}{2}]$, we have $h_{T'}(g(e)) \le h_T(e)$ and the claim is established.

To finish the proof, we need one more claim.

Claim 2. $EV(A) \le EV(T')$. To prove this claim, we first let $m(e)$ be the smaller of the two numbers $|C_e(v)|$ and $|C_e(w)|$ and then label each edge e by $m(e)$. Note that $h_T(e)$ is completely determined by the label $m(e) : h_T(e) = \frac{1}{m(e)+2} + \frac{1}{n+1-m(e)} - \frac{1}{n+1}$. Further, $m(e_1) < m(e_2)$ iff $h_T(e_1) > h_T(e_2)$. It is now easy to construct a bijection $g : E(T') \to E(A)$ with the property that $m(e) \le m(g(e))$ for all $e \in E(T')$. (In particular, note that the labels $m(e)$ along the arm P_i are just intervals $[0, |T_i|]$ of consecutive integers.) See Fig. 7 for a labeling of T' and A. This finishes the proof of the claim.

Finally, if equality holds in the first claim, then the partial order in T_i is a total order for all i, i.e., T_i is a path for all i. If equality holds in the second claim, then $d = 2$ and T' is also a path. Putting all the pieces together gives the result. ■

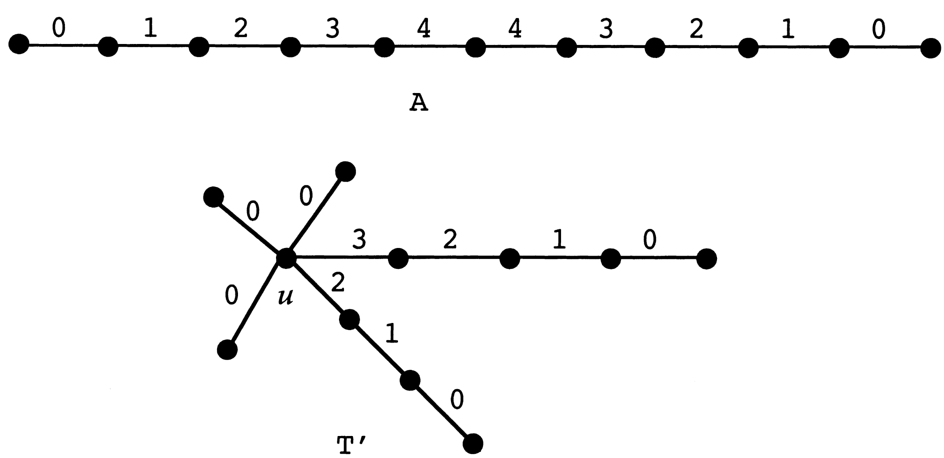

FIGURE 7. Using $m(e)$ to label A and T'.

It would be of interest to find a simpler proof for the lower bound above. In general, the proofs for unrooted trees tend to be more complicated than the corresponding proofs for rooted trees since the existence of the root allows easier inductive arguments.

Since integration is a linear operator, we also get the following:

Corollary 4.4. *Let T_1 and T_2 be rooted trees. Then $EV(T_1 \oplus T_2) = EV(T_1) + EV(T_2)$.*

In light of Corollary 2.6, it is easy to construct nonisomorphic rooted trees having the same expected value. Let T_1 and T_2 be the trees of Fig. 2. Then $EV(T_1) = EV(T_2) = \frac{5}{3}$. In fact, it is possible for $EV(T_1) = EV(T_2)$ even when T_1 and T_2 are not the same size. The two rooted trees of Fig. 8 both have $EV = \frac{4}{3}$.

In the unrooted case, it is generally true that trees having more leaves have higher expected values since leaves contribute more $(\frac{1}{2})$ to formula (4) of Proposition 4.1 than any other edges. This is true for all trees having 7 or fewer edges: If T_1 and T_2 each have $n \leq 7$ edges and T_1 has more leaves than T_2, then $EV(T_1) > EV(T_2)$. This fails for the two trees of Fig. 9, however. T_1 has 4 leaves and T_2 has 3 leaves, but $EV(T_1) = \frac{1937}{630} = 3.0746\ldots$ and $EV(T_2) = \frac{1565}{504} = 3.1051\ldots$

What is the probability that the rank of T equals a given target rank k after some edges fail? What is the probability that the rank never falls below some threshold value? These questions are frequently studied in many applications within reliability theory. We let $\bar{R}_k(T,p) = \sum_{r(S)=k} p^{|S|}(1-p)^{n-|S|}$ and compute the probability in the following way:

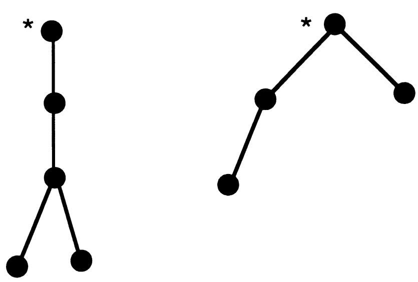

FIGURE 8. $EV(T_1) = EV(T_2) = \frac{4}{3}$.

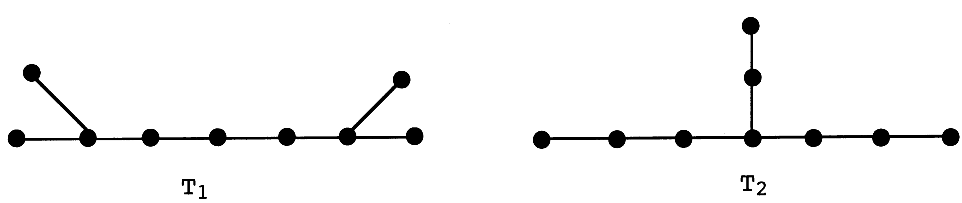

FIGURE 9. $EV(T_1) < EV(T_2)$ despite T_1 having more leaves.

TABLE I.

k	$\bar{R}_k(T_1, p)$	$Pr_1(k)$	$\bar{R}_k(T_2, p)$	$Pr_2(k)$
0	$p^2 - 2p + 1$	$\frac{1}{3}$	$p^2 - 2p + 1$	$\frac{1}{3}$
1	$-p^4 + 3p^3 - 4p^2 + 2p$	$\frac{13}{60}$	$2p^3 - 4p^2 + 2p$	$\frac{1}{6}$
2	$3p^4 - 6p^3 + 3p^2$	$\frac{1}{10}$	$p^4 - 4p^3 + 3p^2$	$\frac{1}{5}$
3	$-3p^4 + 3p^3$	$\frac{3}{20}$	$-2p^4 + 2p^3$	$\frac{1}{10}$
4	p^4	$\frac{1}{5}$	p^4	$\frac{1}{5}$

Definition 4.5. Let T be a (rooted or unrooted) tree with n edges with $k \leq n$. Then the probability that the rank of T equals k is given by

$$Pr_T(\text{rank} = k) = \int_0^1 \bar{R}_k(T, p)dp.$$

From this definition, we immediately get the following:

Proposition 4.6. *Let T be a (rooted or unrooted) tree with n edges. Then*

(1) $\bar{R}(T, p) = \sum_{k=0}^{n} k\bar{R}_k(T, p)$.
(2) $EV(T) = \sum_{k=0}^{n} kPr_T(rank = k)$.

We now give two examples, one rooted and one unrooted. For the rooted case, we again consider the trees T_1 and T_2 of Fig. 2. The computations of $\bar{R}_k(T, p)$ and $Pr_T(\text{rank} = k)$ are given in Table I. In the table, we write $Pr_1(k)$ and $Pr_2(k)$ for rooted trees T_1 and T_2, resp. Note that $\sum_{k=0}^{4} k\bar{R}_k(T_i, p) = 2p^2 + 2p$ and $\sum_{k=0}^{4} kPr_i(k) = \frac{5}{3}$ for $i = 1, 2$ as required by Proposition 4.6. As a check, also note that $\sum_{k=0}^{4} \bar{R}_k(T_i, p) = 1$ as polynomials, for $i = 1, 2$.

For the unrooted case, consider the trees T_1 and T_2 of Fig. 4. In Table II, we list all computations of $\bar{R}_k(T, p)$ and $Pr_T(\text{rank} = k)$. We note that $EV(T_i) = \frac{373}{140} = 2.6642857\ldots$, $\sum_{k=0}^{6} k\bar{R}_k(T_i, p) = 4p + p^2 + p^3 + p^4 + p^5 - 2p^6$, and $\sum_{k=0}^{4} \bar{R}_k(T_i, p) = 1$ for $i = 1, 2$.

Is there an easier formula for computing $\bar{R}_k(T, p)$ and $Pr_T(\text{rank} = k)$ in either the rooted or unrooted case? The formula given in Definition 4.5 is a sum over all subsets of rank k. We can collect terms to sum over all *subtrees* in a manner completely analogous to that given in Corollary 2.3 (in the rooted case) or Corollary 3.2 (in the unrooted case) by concentrating solely on subtrees of size k in the first formula of Proposition 4.1. We omit the straightforward proof of the next proposition.

TABLE II.

k	$\bar{R}_k(T_1,p)$	$Pr_1(k)$	$\bar{R}_k(T_2,p)$	$Pr_2(k)$
0	$(1-p)^4$	$\frac{1}{5}$	$(1-p)^4$	$\frac{1}{5}$
1	$4p-13p^2+15p^3-7p^4+p^5$	$\frac{11}{60}$	$4p-13p^2+15p^3-7p^4+p^5$	$\frac{11}{60}$
2	$7p^2-19p^3+18p^4-7p^5+p^6$	$\frac{67}{420}$	$7p^2-19p^3+17p^4-5p^5$	$\frac{3}{20}$
3	$8p^3-20p^4+16p^5-4p^6$	$\frac{2}{21}$	$8p^3-18p^4+12p^5-2p^6$	$\frac{4}{35}$
4	$8p^4(1-p)^2$	$\frac{8}{105}$	$7p^4(1-p)^2$	$\frac{1}{15}$
5	$6p^5(1-p)$	$\frac{1}{7}$	$6p^5(1-p)$	$\frac{1}{7}$
6	p^6	$\frac{1}{7}$	p^6	$\frac{1}{7}$

Proposition 4.7. *Let T be a rooted tree.*

(1) $\bar{R}_k(T,p) = \sum_{F\in\mathcal{F},|F|=k} p^k(1-p)^{|M(F)|}.$

(2) $Pr_T(rank = k) = \sum_{F\in\mathcal{F},|F|=k} \frac{k!|M(F)|!}{(k+|M(F)|+1)!}.$

Let T be an unrooted tree.

(3) $\bar{R}_k(T,p) = \sum_{F\in\mathcal{T},|F|=k} p^{n-k}(1-p)^{|L(F)|}.$

(4) $Pr_T(rank = k) = \sum_{F\in\mathcal{T},|F|=k} \frac{(n-k)!|L(F)|!}{(n-k+|L(F)|+1)!}.$

Unfortunately, we do not have a formula for Pr_T (rank $= k$) that is analogous to that given in Corollary 2.6 or 3.6. We mention some interesting questions concerning the sequence $\{Pr_T \text{ (rank} = k)\}_{k=0}^{n}$ in Sec. 5.

5. DIRECTIONS FOR FUTURE RESEARCH

5.1. Other Distributions

The expected value calculations in Sec. 4 assume the random variable p is uniformly distributed. There is no physical reason to believe this is the most appropriate distribution for p. In particular, if $g(p)$ is any density function for p defined on [0,1], then we could define the expected value with respect to this distribution as follows:

$$EV_g(T) = \int_0^1 \bar{R}(T,p)g(p)dp.$$

The methods developed in Sec. 4, applied to other distributions, could have more direct applications to real networking problems. In particular, the *beta* distribution $g(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$ can be used to model situations when p is more likely to be in a specified range. For example, if the characteristics of our network imply a probable range of say [0.85, 0.95] for p, then choose $\alpha = 9\beta$. (Different choices of α and β with the same ratio will give different distributions, all having the same expected value, but differing variances. This distribution is treated in most standard texts on statistics; see [5] for example)

5.2. Applications to Other Graphs

It is possible to use greedoid rank functions to apply the techniques here to rooted graphs and digraphs. Some work in this direction appears in [2]; in particular, rooted fans, rooted wheels, and rooted complete graphs generate natural questions. How fast does $EV(G)$ grow for the these graphs? Do the polynomials $R(G)$ and $\bar{R}(G)$ have simpler expressions?

The novelty of using the techniques developed here for rooted graphs and digraphs derives from using the greedoid rank function for these graphs. Extensive analysis of the case when G is not rooted appears in [7], where the rank function is matroidal. Rooted graphs and digraphs do not have matroidal rank functions; hence they are treated differently in reliability theory.

5.3. Random Trees

If T is a tree with n edges (rooted or not), then Propositions 4.2 and 4.3 show $EV(T)$ is bounded below by $\log n - 1$ and bounded above by $n/2$. What is the most likely value of $EV(T)$ for a random tree? Is there a bound on $EV(T)$ that depends on the length of the longest path in T? the maximum degree occurring in T? the number of leaves of T? Answers to these questions should give natural generalizations of Propositions 4.2 and 4.3.

5.4. Probability Sequences

The sequences of Sec. 4 deserve more complete study. We conjecture the following:

Conjecture 5.1.

(1) *Let T be a rooted tree. Then the sequence $\{Pr_T(rank = k)\}_{k=0}^n$ uniquely determines the rooted tree.*

(2) *Let T be an unrooted tree. Then the sequence* $\{Pr_T(rank = k)\}_{k=0}^{n}$ *uniquely determines the unrooted tree.*

This conjecture is true for all trees and rooted trees having eight or fewer edges. A weaker conjecture is the following:

Conjecture 5.2.

(1) *Let T be a rooted tree. Then the sequence* $\{Pr_T(rank = k)\}_{k=0}^{n}$, *together with the polynomial* $\bar{R}(T,p)$, *uniquely determines the rooted tree.*
(2) *Let T be an unrooted tree. Then the sequence* $\{Pr_T(rank = k)\}_{k=0}^{n}$, *together with the polynomial* $\bar{R}(T,p)$, *uniquely determines the unrooted tree.*

5.5. Applications to Other Antimatroids and Greedoids

There has been a sustained program organized to extend the Tutte polynomial to nonmatroidal structures; see [6, 11, 12, 13, 14] for a sample of this work. The probabilistic approach taken here should be applicable to many of the combinatorial structures considered here. This could include a reliability theory for:

- Partially ordered sets [9, 10]
- Rooted directed graphs [20]
- Convex point sets [1, 8]
- Simplicial shelling in a chordal graph [11]

See [19] for more examples of greedoids.

ACKNOWLEDGMENT

We thank Lorenzo Traldi for pointing out the simple proof of Corollary 2.7 using the rank generating function. We also thank Evan Fisher and Joseph Kung for useful discussions.

References

[1] C. Ahrens, G. Gordon, and E. McMahon, Convexity and the beta invariant, Discrete Comp Geo 22 (1999), 411–424.
[2] M. Aivaliotis, A probabilistic approach to network reliability in graph theory, Honors thesis, Lafayette College, 1998.
[3] A. Amin, K. Siegrist, and P. Slater, Pair-connected reliability of a tree and its distance degree sequences, Cong Numer 58 (1987), 29–42.
[4] J. Benashski, R. Martin, J. Moore, and L. Traldi, On the β-invariant for graphs, Cong Numer 109 (1995), 211–221.

[5] G. Casella and R. Berger, Statistical Inference, Wadsworth & Brooks/Cole, Belmont, California, 1990.

[6] S. Chaudhary and G. Gordon, Tutte polynomials for trees, J Graph Theory, 15 (1991), 317–331.

[7] C. Colbourn, The Combinatorics of Network Reliability, Oxford University Press, Oxford, 1987.

[8] P. Edelman and V. Reiner, Counting the interior of a point configuration, Discrete Comp Geo 23 (2000), 1–13.

[9] G. Gordon, A Tutte polynomial for partially ordered sets, J Combin Theory Ser B 59 (1993), 132–155.

[10] G. Gordon, Series-parallel posets and the Tutte polynomial, Discrete Math 158 (1996), 63–75.

[11] G. Gordon, A Beta invariant for greedoids and antimatroids, Electronic J Combin 4 (1997) R13.

[12] G. Gordon, E. McDonnell, D. Orloff, and N. Yung, On the Tutte polynomial of a tree, Cong Numer 108 (1995), 141–151.

[13] G. Gordon and E. McMahon, A greedoid polynomial which distinguishes rooted arborescences, Proc Am Math Soc 107 (1989), 287–298.

[14] G. Gordon and E. McMahon, A greedoid characteristic polynomial, Contemp. Math. 197 (1996), 343–351.

[15] R. Jamison, On the average number of nodes in a subtree of a tree, J Combin Theory Ser B 35 (1983), 207–223.

[16] R. Jamison, Monotonicity of the mean order of subtrees, J Combin Theory Ser B 37 (1984), 70–78.

[17] R. Jamison, Alternating Whitney sums and matchings in trees, Part 1, Discrete Math 67 (1987), 177–189.

[18] R. Jamison, Alternating Whitney sums and matchings in trees, Part 2, Discrete Math 79 (!989/90), 177–189.

[19] B. Korte, L. Lovász and R. Schrader, Greedoids, Springer-Verlag, Berin, 1991.

[20] E. McMahon, On the greedoid polynomial for rooted graphs and rooted digraphs, J Graph Theory 17 (1993), 433–442.