Available online at www.sciencedirect.com
SCIENCE DIRECT®

Discrete Applied Mathematics 128 (2003) 555–571

DISCRETE APPLIED MATHEMATICS

www.elsevier.com/locate/dam

Expected value expansions in rooted graphs☆

Alison Bailey, Gary Gordon, Matthew Patton, Jennifer Scancella

Mathematics Department, Lafayette College, Easton, PA 18042, USA

Received 15 December 2000; received in revised form 14 February 2002; accepted 29 July 2002

Abstract

When G is a rooted graph where each edge may independently succeed with probability p, we consider the expected number of vertices in the operational component of G containing the root. This expected value $EV(G;p)$ is a polynomial in p. We present several distinct equivalent formulations of $EV(G;p)$, unifying prior treatments of this topic. We use results on network resilience (introduced by Colbourn) to obtain complexity results for computing $EV(G;p)$. We use some of these formulations to derive closed form expressions for $EV(G;p)$ for some specific classes of graphs. We conclude by considering optimality questions for rooted graphs, root placement and some counterexamples.

Keywords: Reliability; Probabilistic graph

1. Introduction

When the edges in a rooted graph or digraph fail, the number of vertices adjacent to the root may be diminished, and this may have important consequences for repair of the network. As a dramatic example, on the evening of October 17, 1989, millions of baseball fans tuned in to watch the third game of the World Series between the San Francisco Giants and the Oakland Athletics. At 5:04 p.m. Pacific Daylight Time, just before the game was scheduled to begin, however, a strong earthquake rocked the San Francisco Bay area and the television signal from Candlestick Park was interrupted. The earthquake affected a variety of networks: water supply, sewage treatment and drainage, electrical supply, natural gas supply, computer networks, and so on. In short,

☆ This work was supported by NSF Grant DMS-9912325.
E-mail address: gordong@lafayette.edu (G. Gordon).

PII: S0166-218X(02)00568-1

almost every standard application in network reliability was affected in this natural catastrophe.

We are interested in the expected number of vertices which remain connected to the root vertex in a probabilistic network. Our assumptions are the standard assumptions of reliability theory: Each edge of the network independently succeeds with the same probability p (and so fails with probability $1 - p$), and the vertices of the network cannot fail. While these assumptions are certainly unrealistic in most real-world applications, it is not difficult to modify our models to allow vertex failures, and it is straightforward to allow the edge e to succeed with a probability p_e that depends on e. The assumption of independence, which is more difficult to relax, is reasonable for many networks. When dependence is reasonably well-understood, it is possible to use the model given here by modifying the edge probabilities. See [18] for an approach to dependent edge failures that concentrates on approximating reliabilities. A short discussion can also be found in [26].

While the calculation of the reliability of a network has a large and varied literature (see [12,24,26]), there has been relatively little attention directed towards the expected value polynomial of a rooted network. This polynomial has many interesting properties of its own, however, and we believe further study is warranted. A closely related invariant is Colbourn's *network resilience* for an (unrooted) graph [13], which is the expected number of node pairs that remain connected. Amin, Siegrist and Slater have also explored this topic (using the term *pair-connected reliability*) in a series of papers; see [3,4,27,28] for a sample of their results. That work is motivated by two-terminal reliability problems, which have been well-studied. Most of that literature is concerned with estimating (as a real number) the two-terminal reliability of a graph with two distinguished nodes (a source and a sink). For example, see [9,19] for lower and upper bounds on the probability that two given nodes are connected.

Expected value is, of course, a very simple statistical measure. It would be of interest to examine different probability distributions with different spreads, that give the same expected value. Introduction of measures of spread could give rise to new invariants that are refinements of those considered here. As an example of the issues that can arise, consider a rooted star T on n edges in which the distinguished vertex is a leaf. If each edge has a 70% chance of survival, then the expected number of vertices that remain reachable from the root (which is equivalent to our definition of expected rank for rooted graphs) is approximately $n/2$ (evaluating the expected rank polynomial at $p=0.7$), but there is obviously a 30% chance that no vertices will be adjacent to the root.

As a sample, the *beta* distribution $g(p)=(\Gamma(\alpha+\beta)/\Gamma(\alpha)\Gamma(\beta))p^{\alpha-1}(1-p)^{\beta-1}$ can be used to model situations when some range of values for p is known. For example, if the characteristics of our network imply a probable range of say $[0.85, 0.95]$ for p, then choosing $\alpha = 9\beta$ gives an expected value of 0.9. More importantly, 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 [10] for example.

In this paper, we study a new measure $u(G)$ for rooted graphs—the *mean value of the expected rank*: $u(G) = \int_0^1 EV(G; p)\,\mathrm{d}p$. This invariant, which was introduced in

[1], can be added to the long list of performance measures for a network. In addition to reliability and resilience, various authors have introduced and studied *cohesion* [23], *edge* and *node toughness* [11], *persistence* [8], and *reachability* [5]. Another related topic, *broadcasting,* refers to sending a signal from a source node to all the other nodes in the network, and is thus related to our investigations since our networks have a single, distinguished node [17,20]. Much of the work in this field is concerned with scheduling and timing questions, which do not concern us here.

In Section 2 we develop several equivalent ways to compute the expected value for the number of vertices that remain connected to the root vertex. When each edge has the same independent probability p of succeeding, the expected value is a polynomial in p. A deletion–contraction recursion (Proposition 2.3) (closely related to the *factoring* result of [25]), a related subtree expansion (Proposition 2.7) and a probabilistic expansion (Proposition 2.7) are all given. For many rooted graphs, the probabilistic expansion is the most efficient, but complexity results (Proposition 2.9) adapted from Colbourn's work on resilience show that calculating the expected value (for fixed p) is #P-complete, even for planar networks. Efficient reduction algorithms do exist for special classses—see [14] for consideration of delta-wye reductions, for example. We conclude the section with a brief discussion of polynomial dominance and optimality. Proposition 2.11 finds the *strongly optimal* graphs having n vertices and $n-1$ edges (rooted stars) and those having n vertices and n edges (a rooted star joined to a rooted triangle).

Section 3 gives explicit, closed form expressions for $EV(G;p)$ for rooted trees, circuits, fans and wheels. These results are based on calculations of Amin et al. [4], but can be derived in other ways. They are included to give an indication of the kind of formulas that hold for relatively simple rooted graphs.

In Section 4, we define the mean value of the expected rank $u(G)$. This measure corresponds to the situation when no information about the distribution of p exists, so we assume p is uniformly distributed. Then, among all graphs on n vertices and m edges, there is a unique value $u_{n,m}$ corresponding to the maximum possible $u(G)$. A graph achieving this maximum is called *mean-optimal*. Corollary 4.4 characterizes the mean-optimal trees and the mean-optimal unicyclic graphs. We also present an example to show how vertex placement can affect both $EV(G;p)$ and $u(G)$.

We conclude with some counterexamples in Section 5, showing that $u(G)$ is not necessarily maximized when the root is adjacent to all other vertices. We briefly indicate the results of a computer search that finds all uniformly optimal graphs when n and m are manageable.

There are several possibilities for extending this work in the future:

- Application of the techniques developed here to rooted directed graphs could have implications for directed network design. Many of our results apply to the directed case.
- The assumption that p is uniformly distributed is not realistic for most applications. It would be of interest to use real-world data to apply our optimality definition to situations where the distribution is not uniform (although results in this situation would probably be heuristic).

- Characterization of other optimal rooted graphs should be possible. In particular, it is natural to extend the characterization of Proposition 2.11 (or Corollary 4.4) to other classes, for example, graphs on n vertices and m edges for $n+1 < m < \binom{n}{2}$. Presumably, a partial catalog of optimality results could be developed.

While much of our motivation is rooted in the traditional reliability literature, we emphasize the underlying mathematical interest of this topic. Attempts to apply this work to real-world situations will require many adaptations and approximations, especially in view of the computational complexity results of Proposition 2.9(2). While we will not discuss these issues in depth, we believe such potential applications hold promise.

2. Definitions and examples

Let G be a rooted graph, let E denote the edges of G and suppose each edge of the graph has the same independent probability p of succeeding after some catastrophic event. Although it is not difficult to generalize so that each edge e has an independent probability p_e of success, we do not treat this situation here. Nearly all of the formulas we derive have analogous expressions in the more general situation. For a subset of edges $S \subseteq E$, we let $r(S)$ denote the number of vertices (besides the root) in the component of the subgraph S that contains the root. (We simultaneously refer to S as a subset of edges and as a *subgraph*.)

The rank $r(S)$ is the *greedoid* rank when the rooted graph is considered as an edge branching greedoid. Greedoids are generalizations of matroids, which, in turn, simultaneously generalize graphs and matrices. For more information on greedoids, the interested reader is referred to [7,21]. We will not explicitly refer to greedoids in this paper.

Definition 2.1. Let G be a rooted graph. The expected value $EV(G;p)$ is

$$EV(G;p) = \sum_{S \subseteq E} r(S) p^{|S|} (1-p)^{|E-S|}.$$

For the most part, we will consider simple rooted graphs (i.e., graphs without loops and multiple edges) throughout this paper. We remark, however, that all of the work here easily generalizes to graphs with loops and multiple edges, and we will relax this restriction when loops and multiple edges arise in certain computations, as in the deletion–contraction algorithm of Proposition 2.3. From the viewpoint of network reliability, building networks with multiple edges can be viewed as ‘toughening’ a single edge (i.e., making it less likely to fail).

Definition 2.1 is consistent with the usual interpretation for expected value. For example, if G is a rooted triangle, then the reader can check the rank of each of the 8 subsets of edges. Expanding the resulting polynomial gives $EV(G;p) = 2p + 2p^2 - 2p^3$. Note that $EV(G;1) = 2$, corresponding to the situation in which each edge is certain to survive.

The next result collects some easy consequences of the definition. We omit the proofs.

Proposition 2.2. *Let G be a connected, simple, rooted graph with n vertices, and suppose the root vertex has degree d. Then*

(1) $EV(G;p)$ *has degree at most* n,
(2) $EV(G;p) = dp + p^2g(p)$ *for some polynomial* $g(p)$,
(3) $EV(G;0) = 0$, *and*
(4) $EV(G;1) = n - 1$.

There are several equivalent ways to calculate $EV(G;p)$. We present a recursive procedure based on the familiar operations of deletion and contraction to compute $EV(G;p)$. Recall that if e is an edge in G, then the deletion $G - e$ is the graph obtained from G by simply removing the edge e; the contraction G/e is obtained by removing e and then identifying the two endpoints of e. (We will not contract loops in graphs that arise in this process.)

Proposition 2.3 (Deletion–contraction). *Let G be a rooted graph and let e ($\neq$ loop) be an edge adjacent to the root. Then*

$$EV(G;p) = (1-p)EV(G-e;p) + pEV(G/e;p) + p.$$

Proof. We use a conditional probability argument. Let X_G be the random variable giving the number of vertices in the component containing the root (i.e., X_G gives the rank of the surviving rooted subgraph). Then

$$E(X_G) = pE(X_G|e \text{ succeeds}) + (1-p)E(X_G|e \text{ fails}).$$

But $X_G = X_{G/e} + 1$ whenever e succeeds and $X_G = X_{G-e}$ whenever e fails. □

A similar deletion–contraction formula holds for unrooted graphs where the rank of a subset of edges S is the size of the largest acyclic subset of S—the matroid rank ([6, Proposition 2.1]). Further, this procedure seems to have developed in the reliability literature quite apart from matroid theory (although a common motivation for both treatments is the chromatic polynomial of a graph). In particular, deletion–contraction is often referred to as *factoring* [25].

We can use Proposition 2.3 to collect terms from Definition 2.1 to reduce the number of terms in the expansion. The term grouping we describe corresponds to an interval partition of the power set 2^E so that each interval gives rise to one term in the expansion. Further, there is a one-to-one correspondence between intervals in 2^E and rooted subtrees of G (see Proposition 2.4).

We give an algorithmic description of the recursive deletion–contraction procedure which leads to a rooted subtree expansion of $EV(G;p)$. By recursively applying Proposition 2.3, we eventually obtain a collection of rooted graphs, all of which will have rank zero. Resolving the rooted graph G into these rank zero minors gives rise to a *computation tree* (as in [15]). The terminal nodes in this computation tree correspond to the rank zero minors, since no non-loop edge will be adjacent to the root in any such minor, and hence no edge will be available for deleting or contracting (see Example 2.6).

Let S denote the set of rank zero minors obtained in this way, and, for each $s \in S$, partition the edges of G into 3 classes: C_s will be those edges that were contracted in arriving at the particular rank zero minor s, D_s will be those that were deleted, and the remaining edges (that were neither deleted nor contracted) will be denoted N_s.

Proposition 2.4 (Gordon and McMahon [15, Theorem 2.5]). *Let S denote the set of rank zero minors obtained by some deletion–contraction resolution of G, with $\{C_s\}_{s\in S}$ and $\{N_s\}_{s\in S}$ as above. Then $\{C_s\}_{s\in S}$ is the collection of all rooted subtrees of G. Further, the intervals $[C_s, C_s \cup N_s]$ partition the Boolean lattice 2^E.*

In light of Proposition 2.4, we can use the collection of rooted subtrees as our index set: replace the labels C_s by the rooted subtrees T, and the edges that were neither deleted nor contracted N_s by the more suggestive label N_T.

We summarize the procedure in the following algorithm:

Deletion–contraction resolution algorithm

(1) Use deletion and contraction repeatedly to resolve your rooted graph G into a collection of rank zero minors (in which the only surviving edges are loops or are disconnected from the root).
(2) For each rank zero minor, keep track of the elements contracted along the way; these will be the rooted subtrees of G.
(3) For each rooted subtree T, let N_T be the edges which were neither deleted nor contracted in arriving at T.
(4) Then the subtree T contributes the term $|T|p^{|T|}(1-p)^{|E|-|T|-|N_T|}$ to the polynomial $EV(G;p)$. The exponent $|E|-|T|-|N_T|$ represents the edges that were deleted in arriving at T.

Proposition 2.5 (Subtree expansion). *Let G be a rooted graph, and let $\mathscr{T}$ denote the collection of all rooted subtrees of G, where each subtree has the same root that G has. For each $T \in \mathscr{T}$, we let N_T denote the set of edges of T that were neither deleted nor contracted in the algorithm given above. Then*

$$EV(G;p) = \sum_{T\in\mathscr{T}} |T|p^{|T|}(1-p)^{n-|T|-|N_T|}.$$

Proof. By Proposition 2.4, we have a partition of all edge subsets into intervals. Let $S \subseteq E$ be a subset of edges. Then, $r(S) = |T_s|$ for all S such that $T_s \subseteq S \subseteq T_s \cup L_s$. Thus

$$EV(G;p) = \sum_{S\subseteq E} r(S)p^{|S|}(1-p)^{|E-S|}$$

Table 1

T	N_T	$\lvert T\rvert p^{\lvert T\rvert}(1-p)^{\lvert E\rvert-\lvert T\rvert-\lvert N_T\rvert}$
$\emptyset$	c	0
a	$\emptyset$	$p(1-p)^2$
b	$\emptyset$	$p(1-p)^2$
a,b	c	$2p^2$
a,c	$\emptyset$	$2p^2(1-p)$
b,c	$\emptyset$	$2p^2(1-p)$

$$
\begin{aligned}
&= \sum_{T_s\in\mathscr{T}} \sum_{S:T_s\subseteq S\subseteq T_s\cup L_s} r(S)p^{|S|}(1-p)^{|E-S|}\\
&= \sum_{T_s\in\mathscr{T}} |T_s|p^{|T_s|}(1-p)^{|E|-|T|-|N_T|}\sum_{k=0}^{|L_s|}\binom{|L_s|}{k}p^k(1-p)^{|L_s|-k}\\
&= \sum_{T_s\in\mathscr{T}} |T_s|p^{|T_s|}(1-p)^{|E|-|T|-|N_T|}. \qquad \square
\end{aligned}
$$

In Example 2.6, there are 6 subtrees that correspond to 6 intervals in the Boolean lattice. Four of these intervals are trivial, and the remaining two ($[\emptyset, c]$ and $[ab, abc]$) have length one.

Note that the set N_T depends not only on the subtree T, but also on the specific deletion–contraction resolution used. The fact that this apparent dependence on the order in which we operate on the edges leads to a polynomial which is independent of this order is directly analogous to Tutte's famous basis activities approach to the Tutte polynomial. More connections between order and Tutte polynomials can be found in [16] and [29]. An explicit connection between activities and reliability can be found in [31]. We give a small example to illustrate the definition and these two propositions.

Example 2.6. Let G be the rooted triangle at the top of Fig. 1. In the picture, we use the convention that the left-hand child of a graph minor is obtained by contraction and the right-hand child is obtained by deletion.

Using Definition 2.1 to calculate $EV(G)$ requires 8 terms. The subtree expansion of Proposition 2.5 requires 6 terms. In Table 1, we give the term corresponding to each subtree. The resulting polynomial is $EV(G) = 2p + 2p^2 - 2p^3$.

While Proposition 2.5 is an improvement over the definition in the sense that it involves fewer terms, this improvement is not significant since an arbitrary graph may have an exponential number of rooted subtrees. A much more efficient expansion can be obtained by using methods of Colbourn [13] or of Amin et al. [3]. We omit the straightforward proof of the next proposition.

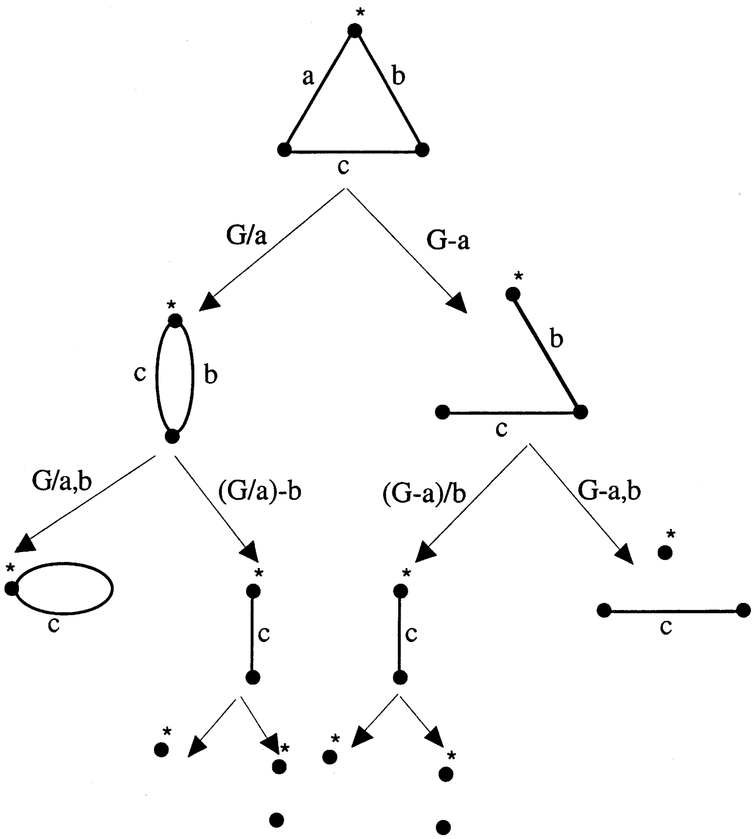

Fig. 1. A deletion–contraction resolution.

Proposition 2.7 (Probabilistic vertex expansion). *Let G be a rooted graph with root vertex $*$, and let V denote the vertices of G. For $v \in V$, let $Pr(v)$ be the probability that v is in the same component of G as the root. Then*

$$EV(G;p) = \sum_{* \neq v \in V} Pr(v).$$

We mention one corollary of this result. The *direct sum* $G_1 \oplus G_2$ of the rooted graphs G_1 and G_2 is the rooted graph formed by identifying the two roots.

Corollary 2.8 (Direct sum). $EV(G_1 \oplus G_2; p) = EV(G_1; p) + EV(G_2; p)$.

While Proposition 2.7 gives this result immediately, we remark that an inductive proof which uses Proposition 2.3 is also routine. Proposition 2.7 also provides an inductive proof of the deletion–contraction formula given in Proposition 2.3.

Proposition 2.7 gives a much simpler way to calculate $EV(G; p)$, provided $Pr(v)$ is easy to calculate. For example, if G is a rooted tree, then $Pr(v) = p^{d(*,v)}$, where $d(*, v)$ is the distance from the root to v. Since determining the distance from a given vertex to any other vertex in a tree can be done in linear time, this computation is very efficient. Consequences of this are explored in [2,3].

While calculation of $EV(G; p)$ is immediate for rooted trees, it is much harder for arbitrary planar graphs. The next proposition adapts two complexity results of Colbourn [13] to the computation of $EV(G; p)$. A graph is *series-parallel* if it contains no subgraph homeomorphic to K_4.

Proposition 2.9. (1) *Suppose G is a rooted series-parallel graph with n vertices. Then $EV(G; p)$ can be computed in* $O(n)$ *time.*

(2) *For rooted planar graphs G, computation of $EV(G; p)$ is #P-complete.*

Proof. (1) This result is [13, Theorem 3.4].

(2) Let $Res(G)$ be the network resilience of G as defined in [13], and let V be the vertex set. Then $Res(G) = \frac{1}{2}\sum_{x\in V} EV(G_x; p)$, where G_x is the graph G rooted at x. Hence, if there were a polynomial algorithm for computing $EV(G_x; p)$, there would be one for computing $Res(G)$. But resilience of planar networks is #P-complete by Theorem 4.3 of [13]. □

We turn our attention to optimal rooted graphs. We will need the following definitions.

Definition 2.10. (1) For polynomials $f(p)$ and $g(p)$ defined on $[0, 1]$, we say $f(p)$ *dominates* $g(p)$ if $g(p) \leqslant f(p)$ for all $p \in [0, 1]$, and the inequality is strict for some $p \in [0, 1]$.

(2) Let $\mathscr{G}_{n,m}$ be the class of all simple graphs on n vertices and m edges. Call a graph $G \in \mathscr{G}_{n,m}$ *strongly optimal* if, for any $H \in \mathscr{G}_{n,m}$ with $H \neq G$, we have $EV(G; p)$ dominates $EV(H; p)$.

Recall that a *rooted star* is a rooted tree in which every vertex is adjacent to the root.

Proposition 2.11. (1) *If $G \in \mathscr{G}_{n,n-1}$, then G is strongly optimal iff G is a rooted star.*

(2) *If $G \in \mathscr{G}_{n,n}$ with no isolated vertices, then G is strongly optimal iff G is a rooted star with one additional edge.*

(3) *Suppose $G_1, G_2 \in \mathscr{G}_{n,m}$, and let d_i be the degree of the root in G_i $(i = 1, 2)$. If $d_1 < d_2$, then $EV(G_1; p) < EV(G_2; p)$ if p is sufficiently small.*

Proof. (1) Let S_n be the rooted star on n vertices (including the root) and suppose $G \in \mathscr{G}_{n,n-1}$ is not a rooted star. Let E denote both the edges of G and S_n. If $A \subseteq E$, then $r_G(A) \leqslant r_{S_n}(A)$. Further, this inequality will be strict whenever A does not form a rooted subtree in G. Thus, $EV(G; p) \leqslant EV(S_n; p)$ for all $0 \leqslant p \leqslant 1$, and the inequality is strict for some p. Thus, S_n is strongly optimal.

(2) We let $G, H_n \in \mathscr{G}_{n,n}$, where G has no isolated vertices and H_n is the rooted star with one additional edge. (H_n is formed as the direct sum of $n-3$ edges and a single rooted triangle.)

Again, we let E denote both the edges of G and H_n. If G is disconnected, then G must contain an edge e which cannot be reached from the root. We may assume e is the unique edge of H_n that is not adjacent to the root. Then, as in the proof of part (1) above, if $A \subseteq E$, we again have $r_G(A) \leqslant r_{H_n}(A)$, and the inequality is strict for $A = \{e, f\}$ for at least one edge $f \in E$. Thus, when G is disconnected, $EV(G; p) \leqslant EV(H; p)$ for all $0 \leqslant p \leqslant 1$, with strict inequality for some p.

Now suppose G is connected. Then we may assume $G = G_1 \oplus \cdots \oplus G_k$, where G_1 has a (unique) cycle and $G_2, \ldots, G_k$ are rooted trees. By part (1) above (and Corollary 2.8), replacing $G_2 \oplus \cdots \oplus G_k$ by the rooted star S_j (where j is the number of vertices of $G_2 \oplus \cdots \oplus G_k$) produces a more dominant expected value polynomial. Thus, we may assume each G_i is a single edge for all $i \geqslant 2$.

We now show that we may further assume G_1 is a rooted cycle. If this is not the case, let $e_1, \ldots, e_l$ are the edges of G_1 which are not in the unique cycle. Then form a graph K_1 as the direct sum of a cycle (the same size as the cycle in G_1) and l single edges. Then the same rank comparison argument as before ($r_{G_1}(A) \leqslant r_{K_1}(A)$) shows $EV(K_1)$ dominates $EV(G_1)$.

Finally, it remains to show that $EV(H_k)$ dominates $EV(C_k)$ when $k > 3$. This follows from another edge correspondence argument, as given above. In particular, let edges a, b and c be the edges of the rooted triangle in H_k, with c the edge that is not adjacent to the root, and let a and b be the edges adjacent to the root in C_k, with c adjacent to a. Complete the edge correspondence arbirarily. Then, as before, we have $r_{C_k}(A) \leqslant r_{H_k}(A)$ for all subsets of edges A, and the inequality is strict for some subsets (e.g., $A = \{c, d\}$ for any $d \neq a, b$).

If $EV(G; p) = EV(H_n; p)$ for all p, then equality is forced at each step, and $G \cong H_n$.

(3) Write $EV(G_i) = d_i p + p^2 g_i(p)$ for $i = 1, 2$, where $g_i(p)$ are polynomials in p, as in Proposition 2.2(2). The result is then immediate. □

We will apply Proposition 2.11 in Section 4.

3. Some classes of rooted graphs

We turn our attention to the calculation of $EV(G; p)$ for some specific classes of graphs. Although the classes we consider in this section are rather simple, the associated polynomials exhibit interesting behavior. Obviously, more complicated graphs will have more complicated associated polynomials. It is possible to adapt our results to more complicated graphs via the deletion–contraction algorithm (for example), especially when the graph under consideration is ‘almost’ in one of the classes computed below.

A *rooted fan* F_n is the graph obtained from joining one vertex (the root) to every vertex on the path P_n with n vertices. Thus, F_n has $n+1$ vertices and $2n-1$ edges. The *rooted wheel* W_n is obtained by joining one vertex (again, the root) to every vertex of the cycle C_n. Thus, W_n has $n+1$ vertices and $2n$ edges. (Equivalently, W_n is obtained by adding one edge to F_n.)

Rooted fans and wheels are useful models for network configurations since they have a relatively small number of edges, but allow every vertex direct access to the root (which might be a server in the application). For example, suppose a satellite needs to communicate to a series of groundstations configured in a path. Then the rooted fan models this situation. Rooted wheels model situations in which stations are linked in a ring, with a central server having direct access to each station.

In the next proposition, we list $EV(G;p)$ for rooted trees, cycles, fans and wheels. For simplicity, we set $q=1-p$ in the formulas for the fan and wheel.

Proposition 3.1. *Let T be a tree, C_n be a rooted cycle on n edges, F_n a rooted fan with $n+1$ vertices and W_n a rooted wheel with $n+1$ vertices. Then*

(1) **Rooted trees**: $EV(T;p)=\sum_{v\in V} p^{d(*,v)}$, *where $d(*,v)$ is the distance from v to the root*,

(2) **Rooted cycles**:

$$EV(C_n;p)=\frac{2p-(n+1)p^n+(n-1)p^{n+1}}{1-p},$$

(3) **Rooted fans**:

$$EV(F_n;p)=\frac{np(1-pq)(1+p^3-p^2(1+(pq)^n))-2p^2q^2(1-(pq)^n)}{(1-pq)^3},$$

(4) **Rooted wheels**:

$$EV(W_n;p)=n\left(\frac{p(1-(pq)^n)}{1-pq}+\frac{p^2q^2(1-n(pq)^{n-1}+(n-1)(pq)^n)}{(1-pq)^2}\right).$$

Proof. (1) This appears as [3, Theorem 1]. Also see [2, Corollary 2.6].

(2) First note that $Pr(v)=p^a+p^b-p^n$, where $a+b=n$ and a and b are the distances from the root to v in the two different directions along the cycle. This immediately gives $EV(C_n;p)=2p+\cdots+2p^{n-1}-(n-1)p^n$. The formula given is a closed form version of this.

(3) Let $v_1,\ldots,v_n$ be the vertices of the path, each of which is joined to the root. By [4, Eq. (10)],

$$Pr(v_i)=1-\frac{q(q+p^2(pq)^{i-1})(q+p^2(pq)^{n-i})}{(1-pq)^2}.$$

To complete the proof, we need only compute $\sum_{i=1}^{n} Pr(v_i)$. The formula given is the result.

(4) If v is any vertex besides the root, then $Pr(v)$ is given in [4, Eq. (7)]. The formula for $EV(W_n;p)$ is a closed form for $nPr(v)$. $\square$

We remark that it is also quite easy to verify the formulas for $EV(C_n;p)$ and $EV(T;p)$ (for a tree T) using induction and Proposition 2.3. The derivations for rooted fans and wheels can also be done this way, but the details are a bit messy.

4. Mean-optimal rooted graphs, crossings and vertex placement

Within reliability theory, polynomials are evaluated frequently at various values of p. For example, if it is known that network connections are very reliable, then a high value for p can be assumed in computing the reliability or the expected value.

Our approach differs from the standard one in that we do not assume any specific value for p, but specify a *distribution* of values. This is a Bayesian approach in that it depends on the prior distribution of p. When viewing p as a parameter in this way, any new information that could be incorporated into such a model would, of course, be of interest and might facilitate more accurate prediction of edge failures. For example, if no prior information about the reliability of edges is available, then it is reasonable to assume that p is a random variable with uniform distribution. Then calculating the expected number of vertices that remain joined to the root amounts to computing an integral. This motivates the next definition.

Definition 4.1. Let G be a rooted graph with expected value polynomial $EV(G; p)$. Then the *uniform expected rank* $u(G)$ is defined by

$$u(G) = \int_0^1 EV(G; p)\, dp.$$

Of course, if the distribution of p is known as some density function $\delta(p)$, then we could compute $\int_0^1 EV(G; p)\delta(p)\, dp$ to give a more accurate measure of the expected number of vertices that remain joined to the root. It would be an interesting exercise with potentially wide application to apply this definition to real-world problems in which good data exist to estimate $\delta(p)$. A good model for this would be the two-parameter beta family, which allows p to be centered arbitrarily in $[0, 1]$ with specified spread. We concentrate exclusively on the uniform case here, however.

One fundamental practical problem in network design concerns the location of the root. More specifically, given a graph G, find the vertex (or vertices) v such that $u(G)$ is maximized when the root is placed at v. The next example shows how our reliability measures $EV(G; p)$ and $u(G)$ can vary depending on the root placement.

Example 4.2. Let G be the graph of Fig. 2. Write G_v for the graph rooted at the vertex v. Then $EV(G_5, p) \geqslant EV(G_v, p)$ for p in the range $0 \leqslant p \leqslant 0.355$ (approx.) and all vertices v. We say vertex 5 is *locally optimal* for the range $0 \leqslant p \leqslant 0.355$. For this graph, we find vertex 6 is locally optimal for $0.355 \leqslant p \leqslant 0.906$, and vertex 7 is optimal for $0.906 \leqslant p \leqslant 1$.

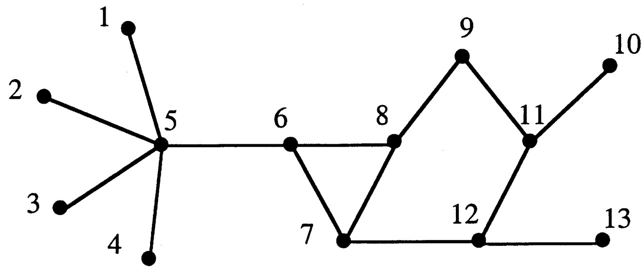

Fig. 2. Dependence on root location.

We also compute $u(G_v)$ for vertices 5, 6, 7 and 8. We find

$$u(G_5) = 4.68254\ldots,$$

$$u(G_6) = 4.86706\ldots,$$

$$u(G_7) = 4.81825\ldots,$$

$$u(G_8) = 4.78373\ldots.$$

Thus, vertex 6 is the optimal location for the root when using $u(G_v)$ as our criterion.

The dependence of vertex location on the value of p is analogous to finding 'crossings' of the graphs of two reliability polynomials. See [22] for examples that show there is no most reliable network for specified parameters.

It follows from Proposition 2.2(2) that for small values of p, the locally optimal vertex is the vertex with the largest degree. For larger values of p, other factors having to do with the general idea of 'centrality' of a vertex play a role in optimality. For example, the vertex (or vertices) with the minimum average distance to other vertices is frequently the globally optimal vertex, i.e., the vertex that maximizes $u(G_r)$.

Among all rooted graphs on n vertices and m edges, there is a unique maximum value for $u(G)$. We call a graph that achieves this maximum *mean-optimal*. More formally, we make the following definition.

Definition 4.3. $G \in \mathscr{G}_{n,m}$ is *mean-optimal* if $u(G) \geqslant u(H)$ for all $H \in \mathscr{G}_{n,m}$. If G is mean-optimal, we write $u_{n,m}$ for $u(G)$.

The next result follows easily from Proposition 2.11.

Corollary 4.4. (1) *If $m < n(n-1)/2$, then $u_{n,m} < u_{n,m+1}$.*

(2) *If $G \in \mathscr{G}_{n,n-1}$, then $u(G) \leqslant (n-1)/2 = u_{n-1,n}$, with equality iff G is a rooted star.*

(3) *If $G \in \mathscr{G}_{n,n}$ with no isolated vertices, then $u(G) \leqslant (3n-2)/6 = u_{n,n}$, with equality iff G is a rooted star with one additional edge.*

We can say (slightly) more: If G is strongly optimal, where $G, H \in \mathscr{G}_{n,m}$ with $H \neq G$, then $\int_0^1 EV(H; p)\delta(p)\,dp < \int_0^1 EV(G; p)\delta(p)\,dp$ holds for any continuous density function $\delta(p)$.

Corollary 4.4(2) characterizes mean-optimal trees. Similar extremal results hold for pair-connected reliability ([3, Theorem 3]) and unrooted trees ([2, Proposition 4.3]). Corollary 4.4(3) characterizes mean-optimal unicyclic graphs. For pair-connected reliability, Siegrist, Amin and Slater ([28, Theorem 4.1]) show that the optimal unicyclic graphs are *star cycles* in which the root is in a cycle of length j, where j varies depending on p.

Allowing multiple edges in our rooted graphs can allow more flexibility in designing mean-optimal networks. In particular, let R_n be the direct sum of one pair of parallel edges with $n-2$ single edges (so R_n has n vertices and edges). Then $EV(R_n; p) = np - p^2$, and so $u(R_n) = (3n-2)/6$, as in part (3) of Corollary 4.4. We give the following interpretation for such a graph as a network model. Rather than physically building parallel edges in a network, increase the reliability of a single edge from p to $2p - p^2$ (the value for the reliability of a pair of edges). Then expected value calculations in the network will match the calculations in the multigraph. A similar adjustment to other values can be used to extend this interpretation to an arbitrary multigraph.

This idea can be formalized in the following sense: In recursively computing $EV(G)$ using the deletion–contraction procedure of Proposition 2.3, multiple edges will arise. Then using series-parallel reductions can reduce the number of steps needed to compute $EV(G)$. This involves replacing a set of edges in parallel with a single edge having a new edge probability (as above), and performing a similar operation on edges in series (which will arise when the parallel reduction is performed). Although this will not make the entire calculation tractable (by Proposition 2.9(2)), it can be a useful tool for specific examples. See [30] for details.

The restriction of having no isolated vertices in Corollary 4.4(3) is rather mild from a physical viewpoint; it makes no sense at all to build networks with vertices isolated from the root. Nevertheless, it is easy to construct examples of rooted graphs $G, H \in \mathscr{G}_{n,m}$ in which G is connected, H is not connected, but $u(G) < u(H)$.

5. Edge flipping and counterexamples

Given two connected, rooted graphs $G, H \in \mathscr{G}_{n,m}$, how can we quickly decide which graph is 'better'? The answer depends, of course, on the application under consideration, and what we mean by 'better'. If p is small, then the rooted graph whose root has higher degree is 'better' (Proposition 2.11(4)). If we compare $u(G)$ and $u(H)$, we may not wish to select the rooted graph whose vertex has the higher degree.

As a simple example, let T_n be the rooted star with the root placed at a vertex of degree one, and let C_n denote the rooted cycle, as before. Then C_n is better than T_n when p is small, but $u(C_n) \sim 2\log n$, while $u(T_n) \sim n/3$.

A more interesting set of examples can be constructed as follows. Suppose e is an edge in the rooted graph G with endpoints v and w, and assume e is not adjacent to the root. Then replace the edge e by an edge e' joining the root to the vertex

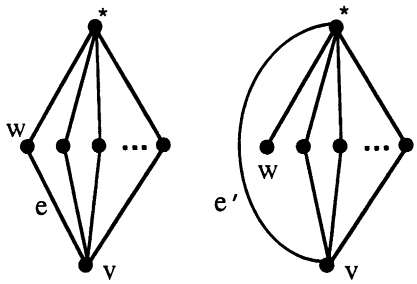

Fig. 3. Edge flipping.

Table 2

n	$u(D_n)$	$u(D_n(e_v))$
4	1.57	1.67
5	2.27	2.33
6	2.96	3
7	3.65	3.67
8	4.32	4.33
9	5.01	5
10	5.68	5.67

v. We call this operation *flipping e at the vertex v* and denote the resulting graph $G(e_v)$. (Similarly, we also define $G(e_w)$.) Edge flipping increases the degree of the root, and so it obviously increases $EV(G;p)$ when p is small. Does it also increase $u(G)$? More generally, must any (connected) mean-optimal graph be such that every vertex is adjacent to the root?

We answer both of these questions negatively in the next example.

Example 5.1. Let D_n be the rooted diamond graph of Fig. 3. D_n has n vertices and $2n-4$ edges. (D_n is isomorphic to the complete bipartite graph $K_{2,n-2}$, with the root at one of the two vertices comprising one of the color classes.) Let v denote the unique vertex of distance 2 from the root (i.e., v is the other vertex in the same color class as the root) and consider the graph $D_n(e_v)$ obtained from D_n by flipping any edge e adjacent to v at v (see Fig. 3).

Table 2 gives the (rounded) values of $u(D_n)$ and $u(D_n(e_v))$ for several small values of n. Note that edge flipping increases the integral for $n<9$, but then edge flipping no longer increases $u(G)$. This pattern continues: $u(D_n) > u(D_n(e_v))$ for all $n \geqslant 9$.

The preceding example shows that edge flipping does not always increase the value of $u(G)$. In fact, more is true: D_9 is the unique mean-optimal graph on 9 vertices and 14 edges, so $u_{9,14} = 64,447/12,870 = 5.00754\ldots$. Thus, mean-optimal graphs need not have the root joined to every other vertex.

A C++ program was written to catalog the mean-optimal graphs for small values of m and n and the values $u_{m,n}$. See the web site
http://www.cs.lafayette.edu/∼pattonm/reu/reu.html
for the optimal graphs and details on the implementation of the program. (We generate all of the optimal graphs on 4, 5, 6 and 7 vertices, the optimal graphs with from 9 to 15 edges and 23 to 27 edges on 8 vertices, and the optimal graphs with between 9 and 13 edges on 9 vertices.)

Acknowledgements

We thank Lorenzo Traldi for several very helpful discussions, Evan Fisher for valuable comments concerning probability and the anonymous referees for several very useful suggestions, including the conditional probability argument used to prove Proposition 2.3.

References

[1] M. Aivaliotis, A probabilistic approach to network reliability in graph theory, Honors Thesis, Lafayette College, 1998.
[2] M. Aivaliotis, G. Gordon, W. Graveman, When bad things happen to good trees, J. Graph Theory 37 (2001) 79–99.
[3] A. Amin, K. Siegrist, P. Slater, Pair-connected reliability of a tree and its distance degree sequences, Congr. Numer. 58 (1987) 29–42.
[4] A. Amin, K. Siegrist, P. Slater, Exact formulas for reliability measures for various classes of graphs, Congr. Numer. 58 (1987) 43–52.
[5] M. Ball, S. Provan, Calculating bounds on reachability and connectedness in stochastic networks, Networks 13 (1983) 253–278.
[6] J. Benashski, R. Martin, J. Moore, L. Traldi, On the β-invariant for graphs, Congr. Numer. 109 (1995) 211–221.
[7] A. Björner, G. Ziegler, in: N. White (Ed.), Introduction to Greedoids Matroid Applications, Encyclopedia of Mathematics and its Applications, Vol. 40, Cambridge University Press, London, 1992, pp. 284–357.
[8] F. Boesch, F. Harary, J. Kabell, Graphs as models of communication network vulnerability: connectivity and persistence, Networks 11 (1981) 57–63.
[9] T. Brecht, C. Colbourn, Lower bounds on two-terminal network reliability, Discrete Appl. Math. 21 (1988) 185–198.
[10] G. Casella, R. Berger, Statistical Inference, Wadsworth & Brooks/Cole, Belmont, CA, 1990.
[11] V. Chvatel, Tough graphs and hamiltonian circuits, Discrete Math. 5 (1973) 215–228.
[12] C. Colbourn, The Combinatorics of Network Reliability, Oxford University Press, Oxford, 1987.
[13] C. Colbourn, Network resilience, SIAM J. Algebraic Discrete Methods 8 (1987) 404–409.
[14] T. Feo, J.S. Provan, Delta-wye transformations and the efficient reduction of two-terminal planar graphs, Oper. Res. 41 (1993) 572–582.
[15] G. Gordon, E. McMahon, Interval partitions and the greedoid Tutte polynomial, Adv. Appl. Math. 18 (1997) 33–49.
[16] G. Gordon, L. Traldi, Generalized activities and the Tutte polynomial, Discrete Math. 85 (1989) 167–176.
[17] N. Hall, W-P. Liu, J. Sidney, Scheduling in broadcast networks, Networks 32 (1998) 233–253.
[18] P. Hansen, B. Jaumard, G.-B. Nguetse, Best second order bounds for two-terminal network reliability with dependent edge failures: the satisfiability problem. Boolean functions, (Certosa di Pontignano, 1996), Discrete Appl. Math. 96/97 (1999) 375–393.

[19] D. Harms, C. Colbourn, Renormalization of two-terminal reliability, Networks 23 (1993) 289–297.
[20] S.M. Hedetnieme, S.T. Hedetnieme, A. Liestman, A survey of gossiping and broadcasting in communication networks, Networks 18 (1988) 319–349.
[21] B. Korte, L. Lovász, R. Schrader, Greedoids, Springer, Berlin, 1991.
[22] W. Myrvold, K. Cheung, L. Page, J. Perry, Uniformly-most reliable networks do not always exist, Networks 21 (1991) 417–419.
[23] R. Ringeisen, M. Lipman, Cohesion and stability in graphs, Discrete Math. 46 (1983) 191–198.
[24] F. Roberts, F. Hwang, C. Monma (Eds.), Reliability of Computer and Communication Networks, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 5, 1991.
[25] A. Satyanarayana, M. Chang, Network reliability and the factoring theorem, Networks 13 (1983) 107–120.
[26] D. Shier, Network Reliability and Algebraic Structures, Oxford University Press, Oxford, 1991.
[27] K. Siegrist, Expected value expansions in random subgraphs with applications to network reliability, Combe. Probab. Comput. 7 (1993) 465–483.
[28] K. Siegrist, A. Amin, P. Slater, The optimal unicyclic graphs for pair-connected reliability, Discrete Appl. Math. 41 (1993) 235–243.
[29] L. Traldi, A note on reliability and expected value, Congr. Numer. 133 (1998) 95–99.
[30] L. Traldi, Series and parallel reductions for the Tutte polynomial, Discrete Math. 220 (2000) 291–297.
[31] L. Traldi, Activities for reliability. I, preprint.