On Reliability in Stochastic Graphs

Ove Frank
Statistics Department, University of Lund, S-22007, Lund, Sweden
Wolfgang Gaul
Applied Mathematics Department, University of Bonn, 5300 Bonn, West Germany

A complete graph has randomly failing nodes and edges. All failures are independent, and there is a common node reliability and a common edge reliability. Generalizing an approach developed by Gilbert for reliable nodes and unreliable edges, we find formulas for various kinds of connectedness probabilities. Bounds and approximations to the probabilities are given.

I. INTRODUCTION

A simple model for a system of unreliable components is given by a stochastic graph G having N nodes and R random edges which correspond to the components of the system. The edges corresponding to functioning components are realized. Components are independently functioning with probability α. This model was used by Moore and Shannon [7] in their pioneering work and has since then been used extensively in the literature; references may be found, for instance, in Barlow and Proschan [1, 2], Wilkov [11], and Tainiter [9].

Two measures of system reliability which are commonly used are the probability P_N that the stochastic graph G is connected and the probability Q_N that two specified nodes are connected in G. The probability P_N is equal to

$$P_N = \sum_{k=0}^{R} A_{N,k}\alpha^k(1-\alpha)^{R-k}, \tag{1}$$

where $A_{N,k}$ is the number of connected realizations of G having exactly k edges. Another formula for P_N is given by

$$P_N = 1 - \sum_{k=0}^{R} B_{N,k}\alpha^k(1-\alpha)^{R-k}, \tag{2}$$

NETWORKS, Vol. 12 (1982) 119–126

CCC 0028-3045/82/020119-08$01.80

where $B_{N,k}$ is the number of disconnected realizations of G having exactly k edges. Neither of formulas (1) and (2) is convenient for numerical calculations for large N, and many algorithms have been proposed for determining P_N and Q_N; see for instance Shogan [8] for some references to various algorithms based on enumeration of realizations, paths, or cutsets of G.

For complete graphs with random edges there is another approach available. Gilbert [6] proved the recursive formula

$$P_N = 1 - \sum_{k=1}^{N-1} \binom{N-1}{k-1} P_k (1-\alpha)^{k(N-k)} \tag{3}$$

by which P_N can be successively determined for $N = 2, 3, \ldots$, using the initial value $P_1 = 1$. In order to determine Q_N, Gilbert used the relationship

$$Q_N = 1 - \sum_{k=1}^{N-1} \binom{N-2}{k-1} P_k (1-\alpha)^{k(N-k)} \tag{4}$$

for $N = 2, 3, \ldots$. Using a generating series for $A_{N,k}$, Gilbert obtained a generating series for P_N and a combinatorial expression giving P_N as a sum extended over all partitions of N:

$$P_N = \sum \left((-1)^{K-1} (K-1)!\,N!\,(1-\alpha)^{\binom{N}{2}-R} \Big/ \prod_{n=1}^{N} (n!)^{K_n} K_n! \right), \tag{5}$$

where K_n is the number of n-parts and

$$K = \sum_{n=1}^{N} K_n, \quad N = \sum_{n=1}^{N} nK_n, \quad R = \sum_{n=1}^{N} \binom{n}{2} K_n. \tag{6}$$

By using (6) it follows that the exponent $\binom{N}{2} - R$ in (5) is equal to the expression $\frac{1}{2}(N^2 - \Sigma n^2 K_n)$ which was used by Gilbert. The number of terms in (5) increases rapidly with N, and P_N is more easily computed by means of the recurrence relation (3).

The appropriateness of a model in which both nodes and edges are random has been discussed by Van Slyke, Frank, and Kershenbaum [10]. We shall introduce such a model and generalize Gilbert's approach in order to find various connectedness probabilities. In Section II we give the necessary definitions and notation, and in Section III we prove formulas for four different connectedness probabilities. Section IV gives some bounds and approximations to the probabilities.

II. A STOCHASTIC GRAPH MODEL

Consider a complete graph of order N and size $\binom{N}{2}$ in which the nodes and edges may be subject to failure. We shall say that a node or edge is *up* if it does not fail. Assume that nodes and edges are up independently of each other with probabilities β and α, respectively. Let G be the subgraph consisting of the edges which are up and all the N nodes. Let H be the subgraph of G induced by the nodes which are up. Then H consists of the nodes which are up and the edges which are up and not incident to any failing node. If we introduce indicator variables

$$y_i = \begin{cases} 1 & \text{if node } i \text{ is up,} \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

$$x_{ij} = \begin{cases} 1 & \text{if the edge between nodes } i \text{ and } j \text{ is up,} \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

for $i, j = 1, \ldots, N$, then G and H can be given by the adjacency matrices with elements x_{ij} and $z_{ij} = x_{ij} y_i y_j$, respectively. Here $x_{ij} = x_{ji}$ and by definition $x_{ii} = 0$. The N indicators y_i and the $\binom{N}{2}$ indicators x_{ij} for $i < j$ are independent Bernoulli variables with expectations β and α, respectively.

Stochastic graph models of this kind have also been investigated in statistical analysis of social networks. In this case G is an unknown graph and H is an observed subgraph induced by a random sample of nodes. The above model applies if the unknown graph is assumed to have a prior distribution given by independent edge occurrences with a common probability α and the nodes are Bernoulli sampled with selection probability β. See Frank [3] for further details and references.

The order of G is N, and the size of G is given by

$$R = \sum_{i<j}\sum x_{ij} \tag{9}$$

and is binomially distributed with parameters $\binom{N}{2}$ and α. The order of H is given by

$$n = \sum_{i=1}^{N} y_i \tag{10}$$

and is binomially distributed with parameters N and β. The size of H is given by

$$r = \sum_{i<j}\sum x_{ij} y_i y_j. \tag{11}$$

The expected value of r is equal to

$$\mathrm{E}\, r = \binom{N}{2} \alpha\beta^2. \tag{12}$$

By applying a moment formula given in [4] or by using results from [5], the variance of r is found to be

$$\mathrm{Var}\, r = \binom{N}{2} \alpha\beta^2(1 - \alpha\beta^2) + 6\binom{N}{3} \alpha^2\beta^3(1 - \beta). \tag{13}$$

The sizes and orders of G and H provide some crude reliability information since, for instance, $r < n - 1$ implies that H cannot be connected, and $r > \binom{n-1}{2}$ implies that H is connected.

Other reliability measures are provided by the probability P_N that G is connected, the probability Q_N that two specified nodes are connected in G, the probability R_N that H is connected, and the probability S_N that two specified nodes are connected in H. These probabilities are investigated in the next section.

III. CONNECTEDNESS PROBABILITIES

Consider first the probability P_N that G is connected, given by Gilbert's formula. This recurrence relation (3) can be given as

$$\sum_{k=1}^{N} \binom{N-1}{k-1} P_k(1-\alpha)^{k(N-k)} = 1, \tag{14}$$

and this relation is proved by noting that the kth term is the probability that a fixed node is contained in a connected component of order k in G. Another recurrence relation for P_N can be obtained by considering the event that a fixed node is contained in a connected component of order k in H. This event has probability

$$\binom{N-1}{k-1} \beta^k P_k [1 - \beta + \beta(1-\alpha)^k]^{N-k} \tag{15}$$

since the fixed node and $k - 1$ other nodes should be up and connected, and each of the remaining $N - k$ nodes should be either failing or up and not incident to any node in the connected component. The union of these events for $k = 1, \ldots, N$ is the event that the fixed node is up, and it follows that

$$\sum_{k=1}^{N} \binom{N-1}{k-1} \beta^k P_k [1 - \beta + \beta(1-\alpha)^k]^{N-k} = \beta. \tag{16}$$

It is also possible to prove (16) algebraically by using a binomial expansion of each term, changing the order of summation, and applying (14) to the inner sum.

The probability Q_N that two fixed nodes are connected in G is equal to

$$Q_N = \sum_{k=2}^{N} \binom{N-2}{k-2} P_k(1-\alpha)^{k(N-k)}, \tag{17}$$

where the kth term is the probability that the two fixed nodes are contained in a connected component of order k in G. Gilbert's relation (4) can be obtained from (14) and (17) by using the identity

$$\binom{N-1}{k-1} = \binom{N-2}{k-1} + \binom{N-2}{k-2}. \tag{18}$$

The probability R_N that H is connected is equal to

$$R_N = \sum_{k=0}^{N} \binom{N}{k} \beta^k P_k (1 - \beta)^{N-k}, \tag{19}$$

where the kth term is the probability that H has order k and is connected. If P_0 is defined as 1, then an empty graph H is vacuously connected, and if P_0 is defined as 0, then an empty graph H is vacuously disconnected.

The probability S_N that two fixed nodes are connected in H is equal to

$$S_N = \sum_{k=2}^{N} \binom{N-2}{k-2} \beta^k P_k [1 - \beta + \beta(1-\alpha)^k]^{N-k}, \tag{20}$$

where the kth term is the probability that the two fixed nodes are contained in a connected component of order k in H. Another expression for S_N is

$$S_N = \sum_{k=2}^{N} \binom{N-2}{k-2} \beta^k Q_k (1 - \beta)^{N-k}, \tag{21}$$

where the kth term is the probability that H has order k and that the two fixed nodes are connected in the subgraph of G induced by the nodes in H. It is also possible to prove (21) algebraically from (17) and (20).

IV. BOUNDS AND APPROXIMATIONS

We shall start by deriving bounds on P_N using the same technique as Gilbert. A slight modification of his lower bound on P_N will be obtained.

An upper bound on P_N is provided by the probability that no node is isolated in G. If A_i denotes the event that node i is isolated in G, then

$$P_N \leqslant 1 - P\left(\bigcup_{i=1}^{N} A_i\right). \tag{22}$$

By using the general inequality

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geqslant \sum_{i=1}^{N} P(A_i) - \sum_{i<j}\sum P(A_i \cap A_j) \tag{23}$$

and the equations

$$P(A_i) = (1-\alpha)^{N-1}, \quad P(A_i \cap A_j) = (1-\alpha)^{2N-3} \tag{24}$$

for $i < j$, it follows that

$$P_N \leqslant 1 - N(1-\alpha)^{N-1} + \binom{N}{2}(1-\alpha)^{2N-3}. \tag{25}$$

A lower bound on P_N can be obtained from (3) by bounding P_k with 1 for $k<N$. This yields

$$P_N \geqslant 1 - \sum_{k=1}^{N-1} \binom{N-1}{k-1} (1-\alpha)^{k(N-k)}$$

$$= 1 - \frac{1}{2} \sum_{k=1}^{N-1} \binom{N}{k} (1-\alpha)^{k(N-k)}. \tag{26}$$

Since

$$k(N-k) \geqslant \begin{cases} \frac{1}{2}[N+k(N-2)], & 1 \leqslant k \leqslant \frac{1}{2}N, \\ \frac{1}{2}[N+(N-k)(N-2)], & \frac{1}{2}N \leqslant k \leqslant N-1, \end{cases} \tag{27}$$

it follows that

$$(1-\alpha)^{k(N-k)} \leqslant (1-\alpha)^{[N+k(N-2)]/2} + (1-\alpha)^{[N+(N-k)(N-2)]/2} \tag{28}$$

for $1 \leqslant k \leqslant N-1$, and if (28) is used in (26) we obtain, after simplification,

$$P_N \geqslant 1 - \sum_{k=1}^{N-1} \binom{N}{k} (1-\alpha)^{[N+k(N-2)]/2}$$

$$= 1 - (1-\alpha)^{N/2} \{[1+(1-\alpha)^{(N-2)/2}]^N - 1 - (1-\alpha)^{N(N-2)/2}\}. \tag{29}$$

From (25) and (29) it follows that

$$P_N = 1 - N(1-\alpha)^{N-1} + O[N^2(1-\alpha)^{3N/2}]. \tag{30}$$

An upper bound on Q_N is provided by the probability that neither of the two specified nodes is isolated in G. This yields

$$Q_N \leqslant 1 - 2(1-\alpha)^{N-1} + (1-\alpha)^{2N-3}. \tag{31}$$

A lower bound on Q_N can be obtained from (4) by bounding P_k with 1 for $k<N$ and using (28); this gives

$$Q_N \geqslant 1 - 2(1-\alpha)^{N-1}[1+(1-\alpha)^{(N-2)/2}]^{N-2}. \tag{32}$$

From (31) and (32) it follows that

$$Q_N = 1 - 2(1-\alpha)^{N-1} + O[N(1-\alpha)^{3N/2}]. \tag{33}$$

Since P_N is bounded, we obtain from (30) that

$$P_N = 1 - N(1-\alpha)^{N-1} + N^2(1-\alpha)^{3N/2} M_N, \tag{34}$$

where M_N is bounded for all N, say $M_N \leqq M$. By application of (34) to each term in (19) and by using the well-known identity

$$\sum_{k=0}^{N} k \binom{N}{k} a^k b^{N-k} = Na(a+b)^{N-1} \tag{35}$$

we find that

$$R_N = \sum_{k=0}^{N} \binom{N}{k} \beta^k (1-\beta)^{N-k} [1 - k(1-\alpha)^{k-1} + k^2(1-\alpha)^{3k/2} M_k]$$

$$= 1 - N\beta(1-\alpha\beta)^{N-1} + \sum_{k=0}^{N} c_{Nk} M_k, \tag{36}$$

where

$$c_{Nk} = k^2 \binom{N}{k} [\beta(1-\alpha)^{3/2}]^k (1-\beta)^{N-k}. \tag{37}$$

Now $c_{Nk} \geqq 0$ and

$$\left| \sum_{k=0}^{N} c_{Nk} M_k \right| \leqq M \sum_{k=0}^{N} c_{Nk}, \tag{38}$$

i.e.,

$$\sum_{k=0}^{N} c_{Nk} M_k = O\left(\sum_{k=0}^{N} c_{Nk} \right). \tag{39}$$

By using the identities (35) and

$$\sum_{k=0}^{N} k(k-1) \binom{N}{k} a^k b^{N-k} = N(N-1)\, a^2 (a+b)^{N-2} \tag{40}$$

we find that

$$\sum_{k=0}^{N} c_{Nk} = N(N-1)\,[\beta(1-\alpha)^{3/2}]^2\,[\beta(1-\alpha)^{3/2} + 1 - \beta]^{N-2}$$

$$+ N\beta(1-\alpha)^{3/2}\,[\beta(1-\alpha)^{3/2} + 1 - \beta]^{N-1}. \tag{41}$$

It follows from (36), (39), and (41) that

$$R_N = 1 - N\beta(1-\alpha\beta)^{N-1} + O(N^2\gamma^N), \tag{42}$$

where

$$\gamma = 1 - \beta[1 - (1 - \alpha)^{3/2}]. \tag{43}$$

Since $\gamma < 1 - \alpha\beta$ we see that the approximation

$$R_N \approx 1 - N\beta(1 - \alpha\beta)^{N-1} \tag{44}$$

holds in the sense

$$R_N = 1 - N\beta(1 - \alpha\beta)^{N-1}[1 + o(1)]. \tag{45}$$

A similar investigation of S_N starts from (33) and (21) and leads to

$$S_N = \beta^2 - 2\beta^2(1 - \alpha)(1 - \alpha\beta)^{N-2} + O(N\gamma^N), \tag{46}$$

where γ is given by (43). In particular, we see that the approximation

$$S_N \approx \beta^2[1 - 2(1 - \alpha)(1 - \alpha\beta)^{N-2}] \tag{47}$$

holds in the sense

$$S_N = \beta^2 - 2\beta^2(1 - \alpha)(1 - \alpha\beta)^{N-2}[1 + o(1)]. \tag{48}$$

This work was supported by Sonderforschungsbereich 72, University of Bonn.

References

[1] R. E. Barlow and F. Proschan, *Mathematical Theory of Reliability*. Wiley, New York (1965).
[2] R. E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, New York (1975).
[3] O. Frank, Sampling and estimation in large social networks. *Soc. Networks* 1 (1978) 91–101.
[4] O. Frank, Moment properties of subgraph counts in stochastic graphs. *Ann. N.Y. Acad. Sci.* 319 (1979) 207–218.
[5] O. Frank and J. Ringström, Bayesian graph-size estimation. Statistics Dept., Univ. of Lund (1979).
[6] E. N. Gilbert, Random graphs. *Ann. Math. Statist.* 30 (1959) 1141–1144.
[7] E. F. Moore and C. E. Shannon, Reliable circuits using less reliable relays. *J. Franklin Inst.* 262 (1956) 191–208, 281–297.
[8] A. W. Shogan, A decomposition algorithm for network reliability analysis. *Networks* 8 (1978) 231–251.
[9] M. Tainiter, A new deterministic network reliability measure. *Networks* 6 (1976) 191–204.
[10] R. Van Slyke, H. Frank, and A. Kershenbaum, *Network Reliability Analysis: Part II. Reliability and Fault Tree Analysis*, R. E. Barlow, J. Fussell, and N. Singpurwalla, Eds., SIAM, Philadelphia (1975), pp. 619–650.
[11] R. S. Wilkov, Analysis and design of reliable computer networks. *IEEE Trans. Comm.* 20 (1972) 660–678.

Received October 1, 1979
Accepted January 22, 1981
Corrected proofs received February 27, 1982