Network Probabilistic Connectivity: Expectation of a Number of Disconnected Pairs of Nodes

Alexey S. Rodionov* and Olga K. Rodionova

Institute of Computational Mathematics and Mathematical Geophysics
Siberian Division of the Russian Academy of Science
Novosibirsk, Russia
Tel.: +383-3326949
alrod@rav.sscc.ru

Abstract. The task of calculating the expectation of a number of disconnected pairs of nodes (EDP) in unreliable network is discussed. The task is NP-hard that is it requires complete enumeration of subgraphs. The techniques for decreasing a number of enumerated subgraphs by using the branching (factoring) method and taking advantage from possible structural features are discussed. Usage of chains, bridges, cutnodes and dangling nodes is considered.

1 Introduction

Random graphs is an acknowledged model for networks of different kinds. For the analysis of network's reliability the probability of connectivity is used mostly ([1,2,3,4,5,6,7], for example) while the expectation of a number of disconnected pairs of nodes (EDP) is sometimes more valuable and informative index. For example all trees are equal from the point of the probability of connectivity while they are not from the point of EDP. Complementary index to EDP is a number of *connected* pairs of nodes (ECP), their sum is equal to the whole number of pairs of nodes in a network. Finding EDP requires complete enumeration of network destructions for its calculation. This is one of the reasons why few of researches deal with it. We can refer to the paper [8] where this index is mentioned among other valuable indexes of a network reliability. We investigate how to decrease the enumeration by using the branching (factoring) method and taking advantage from possible structural features.

As in [1] we mostly gain from considering simple chains and dangling nodes. Existence of bridges or cutnodes can help also.

The rest of the paper is organized as follows: in Section 2 we give main definitions and notations. In Section 3 we discuss reduction of the task dimension by considering bridges, cutnodes and dangling nodes. In Section 4 we give the exact equations for EDP for some kinds of graphs and in Section 5 we present equations for branching by chain. Section 6 is the brief conclusion.

* Corresponding author.

M. Gerndt and D. Kranzlmüller (Eds.): HPCC 2006, LNCS 4208, pp. 101–109, 2006.

2 Definitions and Notations

We consider random graphs with reliable nodes and unreliable edges. Let us denote:

$G(n, m) = (V, U, P, WT)$ – non-oriented graph with a set of nodes V, set of edges U, matrix of edges reliabilities P and vector of nodes weights WT.
$n = |V|$, $m = |U|$ – number of nodes and edges, respectively.
$w_i = w(v_i)$ – weight of a node v_i, $WT = w_1, \ldots, w_n$.
$W(G)$ – total weight of all nodes of G.
p_{ij} – probability of an edge e_{ij} being existent (being in a working state, edge's *reliability*), $P = ||p_{ij}||$; $q_{ij} = 1 - p_{ij}$.
$M(G, P)$ **and** $N(G, P)$ – ECP and EDP of a random graph G.

We use simply G for a graph if its n and m are clear from the context. If needed we use $G(P)$, $G(WT)$ or $G(P, WT)$ also. If we need refer to the weight of i-th node in some graph G, then we use $WT_i(G)$. If we consider some special edges then we usually assign them personal numbers that is use notation e_k (k-th edge) instead of e_{ij} (an edge that connects v_i and v_j). Notation p_k is used for corresponding edge's reliability.

For simplifying some equations we assume that $\prod_{s=i+1}^{i} p_s \equiv 1$.

A weight w_i equal to a number of nodes that were contracted to form a special node v_i (initial weight of each node is 1) is needed for keeping the number of disconnected pairs in case when this node is separated from some other nodes in a graph.

It is obvious that

$$N(G) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} a_{ij} w_i w_j, \tag{1}$$

where a_{ij} is a probability of v_i and v_j be disconnected in G.

The branching (factoring) method for calculating an expected value of any function of a random graph G is based on the equation of composite probability by two alternative hypothesis: existence or non-existence of some edge e_{ij}. Thus

$$N(G) = p_{ij} N(G^*_{ij}) + (1 - p_{ij}) N(G \backslash e_{ij}), \tag{2}$$

where G^*_{ij} is a graph obtained from G by contracting v_i and v_j by an edge e_{ij}, $G \backslash e_{ij}$ – graph obtained from G by deleting the edge e_{ij}. Recursions go on until deriving a graph for which a $N(G)$ is easily obtained.

We say that a random graph is connected (disconnected) when its structure is connected, not realization. In last case we say that "realization of a random graph is connected (disconnected)."

3 EDP for Graphs of a Small Dimension ($n = 2, 3, 4$)

Case of $n = 2$ is obvious:

$$N(G) = (1 - p_{12}) w_1 \cdot w_2. \tag{3}$$

For $n = 3$ we use (1):

$$N(G) = w_1 w_2 (1 - p_{12})(1 - p_{23} p_{13}) + \\ w_1 w_3 (1 - p_{13})(1 - p_{12} p_{23}) + w_2 w_3 (1 - p_{23})(1 - p_{12} p_{13}). \quad (4)$$

For $n = 4$ we use the equation of composite probability considering all possible ways of a graph destruction as hypotheses. After collecting terms we have:

$$N(G) = \quad (5)$$
$$w_1 w_2 q_{12}[(1-p_{13}p_{23})(1-p_{14}p_{24})-p_{13}q_{14}q_{23}p_{24}p_{34}-q_{13}p_{14}p_{23}q_{24}p_{34}] +$$
$$w_2 w_3 q_{23}[(1-p_{24}p_{34})(1-p_{12}p_{13})-q_{12}p_{13}p_{14}p_{24}q_{34}-p_{12}q_{13}p_{14}q_{24}p_{34}] +$$
$$w_3 w_4 q_{34}[(1-p_{13}p_{14})(1-p_{23}p_{24})-p_{12}p_{13}q_{14}q_{23}p_{24}-p_{12}q_{13}p_{14}p_{23}q_{24}] +$$
$$w_1 w_4 q_{14}[(1-p_{12}p_{24})(1-p_{13}p_{34})-q_{12}p_{13}p_{23}p_{24}q_{34}-p_{12}q_{13}p_{23}q_{24}p_{34}] +$$
$$w_1 w_3 q_{13}[(1-p_{12}p_{23})(1-p_{14}p_{34})-p_{12}q_{14}q_{23}p_{24}p_{34}-q_{12}p_{14}p_{23}p_{24}q_{34}] +$$
$$w_2 w_4 q_{24}[(1-p_{23}p_{34})(1-p_{12}p_{14})-q_{12}p_{13}p_{14}p_{23}q_{34}-p_{12}p_{13}q_{14}q_{23}p_{34}].$$

4 Using Structural Peculiarities

As in the case of calculation of a probabilistic connectivity we can take advantage from some peculiarities of a graph under consideration.

First let us make the following derivation. During factoring process by equation (2) we may obtain several graphs with the same structure and matrix P but with different weights of nodes. In this case we can gain from the following useful lemma.

Lemma 1. *If during a graph G factoring process some subgraphs $G_1, \ldots, G_k$ with same structure and matrix P are obtained in which only one special node v_s has different weight w_{si} in G_i, $i = 1, \ldots, k$, then the total contribution of these subgraphs into $N(G)$ is equal to*

$$\left(\sum_{i=1}^{k} p_i \right) \cdot N(G^o), \quad (6)$$

where p_i is a probability of G_i's realization, and graph G^o has the same structure and P as G_i and

$$WT_s(G^o) = \sum_{i=1}^{k} p_i w_{si} / \sum_{i=1}^{k} p_i. \quad (7)$$

Proof. From (1) we have that for any selected index s

$$N(G) = \Big(\sum_{i \in \{1 \ldots n\} \setminus s} w_i \Big) w_s + \sum_{i \in \{1 \ldots n\} \setminus s} \sum_{(j>i) \& (j \neq s)} w_i w_j = A w_s + B. \quad (8)$$

So, if we change a weight w_s of v_s, then coefficients A and B remain unchanged. Thus the total contribution D of all G_i into $N(G)$ is

$$\begin{aligned} D &= \sum_{i=1}^{k} p_i N(G_i) = \sum_{i=1}^{k} p_i (A w_{si} + B) \\ &= A\Big(\sum_{i=1}^{k} p_i w_{si}\Big) + \sum_{i=1}^{k} p_i B = \sum_{i=1}^{k} p_i \Big(A \sum_{i=1}^{k} p_i w_{si} / \sum_{i=1}^{k} p_i + B\Big), \end{aligned} \tag{9}$$

from which we have what was to be proved. □

4.1 *k*-Component Graphs

If a graph G consist of k mutually disconnected subgraphs G_1, …, G_k then, obviously

$$N(G) = \sum_{i=1}^{k} N(G_i) + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} W(G_i) W(G_j). \tag{10}$$

4.2 Deleting Dangling Nodes

Let a connected graph $G(n, m)$ have a dangling node v_i adjacent to some node v_j. Then deletion of the edge e_{ij} leads to occurrence of $w_i \cdot W(G \backslash \{e_{ij}\}) = w_i \cdot \sum\limits_{k \neq i} w_k$ pairs of disconnected nodes. Hence,

$$N(G) = p_{ij} N(G^*_{ij}) + (1 - p_{ij})(w_i W(G \backslash \{e_{ij}\}) + N(G \backslash \{e_{ij}\})). \tag{11}$$

Note that in this case a graph G^*_{ij}, obtained by contracting nodes v_i and v_j by e_{ij}, is by its structure *the same as* $G \backslash \{e_{ij}\}$, and $WT_j(G^*) = w_i + w_j$. Thus from lemma 1 we obtain:

$$N(G) = N(G^o) + (1 - p_{ij}) w_i W(G \backslash e_{ij}), \tag{12}$$

where G^o — a graph which has the same structure and a weight of v_j

$$WT_j(G^o) = p_{ij}(w_j + w_i) + (1 - p_{ij}) w_j = w_j + p_{ij} w_i. \tag{13}$$

4.3 Using Cutnodes

Let G consists from two subgraphs (blocks) G_1 and G_2 that are jointed through a node v_s (cutnode). Then obviously for any pair of nodes v_i and v_j a path between them or lies in one of the blocks if these nodes are in the same block, or goes through v_s otherwise. Thus from (1) we have

$$N(G) = N(G_1) + N(G_2) + \sum_{i \in X(G_1)} \sum_{j \in X(G_2)} (a_{is} + a_{sj} - a_{is} a_{sj}) w_i w_j, \tag{14}$$

where a_{is} and a_{sj} are probabilities of v_i and v_j being disconnected with v_s in G_1 and G_2 correspondingly.

4.4 Bridge Removing

Let a connected graph $G(n, m)$ have an edge e_{st} such that its deletion leads to dividing the graph into two separated components $G_1(k, f)$ and $G_2(n - k, m - f - 1)$ (they are obviously connected graphs). Such edges are known as bridges. Then

$$N(G) = p_{st}N(G^*_{st}) + (1 - p_{st})\big[W(G_1)W(G_2) + N(G_1) + N(G_2)\big]. \tag{15}$$

Using (14) we obtain

$$N(G) = (1 - p_{st})\big[W(G_1)W(G_2)\big] + N(G_1) + N(G_2) + p_{st}\sum_{i\in X(G_1)}\sum_{j\in X(G_2)}(a_{is} + a_{sj} - a_{is}a_{tj})w_iw_j, \tag{16}$$

where a_{is} and a_{tj} are probabilities of v_i and v_j being disconnected with v_s in G_1 and with v_t in G_2 correspondingly.

4.5 Case of a Chain

Let us discuss the case of a chain with k nodes and $k-1$ edges (Ch_k). Let nodes be enumerated in increasing order thus that nodes v_1 and v_k have a degree 1 while all other nodes have degree 2. For simplicity we denote $e_{i,i+1}$ as e_i here.

From (1) we have

$$N(Ch_k) = \sum_{i=1}^{k-1}\sum_{j=i+1}^{k} w_iw_j\Big(1 - \prod_{s=i}^{j-1} p_s\Big). \tag{17}$$

4.6 Case of a Cycle

Now we consider a cycle with k nodes and k edges (C_k). Let nodes be enumerated in order thus that node v_{i+1} follows v_i, $i = 1, \ldots, k - 1$ and v_1 follows v_k. For simplicity we denote $e_{i,i+1}$, $i = 1, \ldots, k - 1$ as e_i and $e_{k,1}$ as e_k and denote reliability of e_i as p_i. Now we can use equation (1) directly. Between each pair of nodes v_i and v_j there are two pathes, clockwise and counterclockwise. Thus

$$N(C_k) = \sum_{i=1}^{k-1}\sum_{j=i+1}^{k} w_iw_j\Big(1 - \prod_{s=i}^{j-1} p_s\Big)\Big(1 - \prod_{s=j}^{k} p_s\prod_{s=1}^{i-1} p_s\Big). \tag{18}$$

4.7 Case of a Tree

The EDP for a n-nodes tree T_n we obtain from (1):

$$N(T_n) = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_iw_j\Big(1 - \prod_{e_{st}\in Pt_{ij}} p_{st}\Big), \tag{19}$$

where Pt_{ij} is a path from v_i to v_j.

For partial cases we can obtain simpler expressions. The case of a chain that is the special case of a tree has been discussed earlier. Now let us have a star-like tree S_k in which k nodes are adjacent to a single node (root). Thus we have some node v_0 with weight w_0 that is adjacent to dangling nodes v_i, $i = 1, \dots, k$ with weights w_i. For simplicity let us denote edges (v_0, v_i) as e_i and their reliabilities as p_i, correspondingly. Using (19) we obtain:

$$N(Ch_k) = w_0 \sum_{i=1}^{k} p_i w_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} w_i w_j (1 - p_i p_j). \tag{20}$$

5 Branching by Chain

As in [1] where we consider the probabilistic connectivity, we can gain from considering simple chains. Because of limited area of this paper we consider only the case of 2-edges chains here.

For further consideration we need the following lemma.

Lemma 2. *If during a graph G factoring process some subgraphs $G_1, \dots, G_k$ are obtained with probabilities $p_1, \dots, p_k$, such that they have the same structure and matrix P, and two special nodes v_s and v_t have weights w_{si} and w_{ti} in G_i, $i = 1, \dots, k$, then the total contribution of these subgraphs into $N(G)$ is equal to*

$$D = \sum_{i=1}^{k} p_i N(G^o) + a_{st} \frac{\sum_{i=1}^{k} p_i \sum_{i=1}^{k} p_i w_{si} w_{ti} - \sum_{i=1}^{k} p_i w_{si} \cdot \sum_{i=1}^{k} p_i w_{ti}}{\left(\sum_{i=1}^{k} p_i \right)^2}. \tag{21}$$

where graph G^o has the same structure and P as G_i and

$$WT_s(G^o) = \sum_{i=1}^{k} p_i w_{si} / \sum_{i=1}^{k} p_i, \tag{22}$$

$$WT_t(G^o) = \sum_{i=1}^{k} p_i w_{ti} / \sum_{i=1}^{k} p_i.$$

Here a_{st} is a probability of v_s and v_t being disconnected in G^o.

Proof. Proof is similar to that of lemma 1 but more complex because of existence of production $w_s w_t$ in (1). □

Now we continue to the main theorem.

Theorem 1. *If a connected random graph has a simple chain $C = e_{sx}, e_{xt}$ connecting nodes s and t through a node v_x with degree 2 then the following equation is true.*

$$N(G) = \left[p_{st}(1 - p_{sx} p_{xt}) + p_{sx} p_{xt}\right] N(G^*) + \tag{23}$$
$$\left(1 - p_{st} - p_{sx} p_{xt}\right) \left\{ N(G^o) + a_{st} \frac{(1 - p_{st})(1 - p_{xt}) p_{sx} p_{xt}}{(1 - p_{st} - p_{sx} p_{xt})^2} w_x^2 \right\}.$$

where a_{st} is a probability of v_s and v_t being disconnected in G.

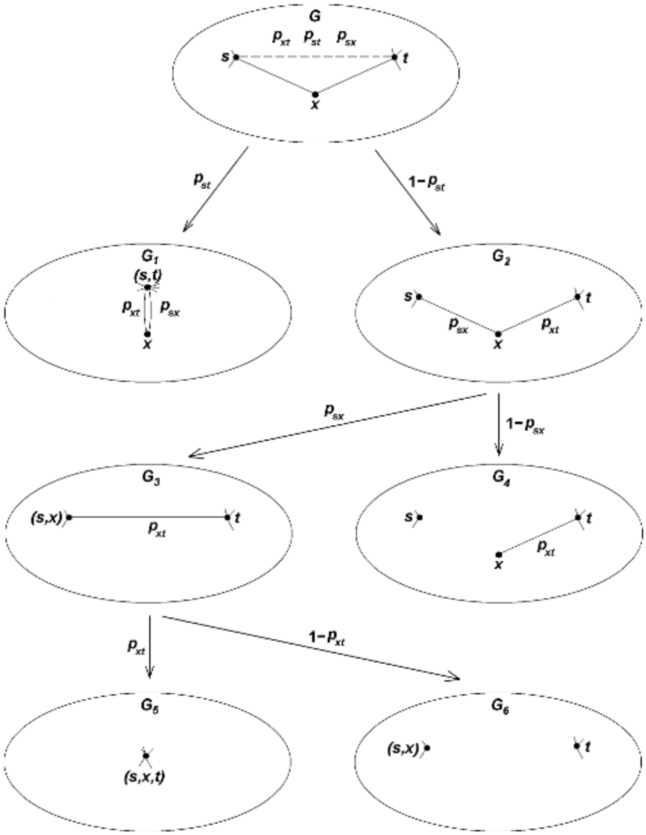

Fig. 1. Branching in a graph with a 2-edge chain

Proof. If e_{st} exists then we first choose it for factoring. Then we consequently make factoring by e_{sx} and e_{xt} (see Fig. 1). Two terminal graphs G_1 and G_4 we can easy replace by graphs with a structure of G_5 and G_6, correspondingly, using (12) and (13) (note that in G_1 we first replace multi-edge by a single one with equivalent reliability):

$$N(G_1) = N(G_5^0),\ WT(G_5^0, st) = w_s + w_t + (p_{sx} + p_{xt} - p_{sx}p_{xt})w_x, \quad (24)$$
$$N(G_4) = N(G_6^0),\ WT(G_6^0, t) = w_t + p_{xt}w_x.$$

For graphs G_5 and G_6 themselves we have:

$$WT(G_5, xst) = w_s + w_t + w_x, \quad (25)$$
$$WT(G_6, s) = w_s + w_x,$$
$$WT(G_6, t) = w_t.$$

It is clear that graphs G_1, G_4, G_5 and G_6 are obtained with probabilities p_{st}, $(1 - p_{st})(1 - p_{sx})$, $(1 - p_{st})p_{sx}p_{xt}$ and $(1 - p_{st})(1 - p_{xt})p_{sx}$, correspondingly.

According to lemma 1 we can change calculation of EDP for G_5 and G_5^0 to calculation of EDP for some graph G^*, that has a structure of G_5 with a weight of joint node

$$WT(G^*, sxt) = w_s + w_t + \frac{(p_{sx}p_{st} + p_{sx}p_{xt} + p_{st}p_{xt} - 2p_{sx}p_{st}p_{xt})}{p_{st} + p_{sx}p_{st} - p_{sx}p_{st}p_{xt}} w_x. \tag{26}$$

For G_6 and G_6^0 we use lemma 2 as two nodes v_s and v_t have different weights in them. According to this lemma a joint contribution of G_6 and G_6^0 into EDP of G is

$$\left(1 - p_{st} - p_{sx}p_{xt}\right)\left\{N(G^o) + a_{st}\frac{(1 - p_{st})(1 - p_{xt})p_{sx}p_{xt}}{(1 - p_{st} - p_{sx}p_{xt})^2} w_x^2\right\}. \tag{27}$$

where G^o has a structure of G without chain e_{sx}, e_{xt} and edge e_{st} and

$$WT_s(G^o) = w_s + w_x \frac{p_{sx}(1 - p_{xt})}{1 - p_{sx}p_{xt}}; \tag{28}$$
$$WT_t(G^o) = w_t + w_x \frac{p_{xt}}{1 - p_{sx}p_{xt}}.$$

From this and (26) we obtain what was to be proofed. □

If e_{st} is absent then the equation (23) can be simplified:

$$N(G) = p_{sx}p_{xt}N(G^*) + \left(1 - p_{sx}p_{xt}\right)\left\{N(G^o) + a_{st}\frac{(1 - p_{xt})p_{sx}p_{xt}}{(1 - p_{sx}p_{xt})^2} w_x^2\right\}, \tag{29}$$

where $WT(G^*, sxt) = w_s + w_t + w_x$.

Now let us discuss obtaining of a_{st} that is a probability of v_s and v_t being disconnected in G. Obviously

$$a_{st} = (1 - p_{st})(1 - p_{sx}p_{xt})P(v_s \text{ and } v_t \text{ are disconnected in } G_6). \tag{30}$$

There are well-known algorithms for finding 2-terminal probabilistic connectivity (see [6], for example). Note that a complexity of this task is obviously less then complexity of calculation of EDP for G_1 or G_4 which makes use of equation (23) effective.

6 Conclusion

Thus we have presented some useful equations that can help in calculating the EDP of a random graph. Most advantage for speeding up is gained from branching by chains. Note that chains are inevitable during the factoring process as a result of edge deletion and, sometimes, as a result of contracting by edge (refer to [1]). Experiments shows that calculation of EDP for 30 random $G(10, 15)$ is in average more than 20 times faster with use of our equations then by using equation (1).

References

1. Rodionova, O.K., Rodionov, A.S., and Choo, H.: Network Probabilistic Connectivity: Exact Calculation with Use of Chains. ICCSA-2004, Springer LNCS. **3046** (2004) 315–324
2. Krivoulets, V.G., Polesskii, V.P., "What is the Theory of Bounds for Network Reliability?," *Information Processes, ISSN: 1819-5822, Vol. 1, n. 2,* pp. 199-203, 2001.
3. Cancela, H., Petingi L., "Diameter constrained network reliability: exact evaluation by factorization and bounds," *Proc. of the Int. Conf. on Industrial Logistics, Japan, 2001*, pp. 359-366, 2001.
4. Lucet,C., Manouvrier, J.-F., "Exact Methods to compute Network Reliability, *Proc. of 1st International Conf. on Mathematical Methods in Reliability, Bucharest, Roumanie, Sep. 1997*, (paper available via http:// ramp.ucsd.edu/resources.html)
5. Satyanarayana, A. and Chang, M.K. "Network reliability and the factoring theorem," *Networks, Vol. 13,* pp.107120, 1983.
6. Shooman, A.M.: Algorithms for Network Reliability and Connection Availability Analysis. Electro/95 Int. Professional Program Proc. (1995) 309–333
7. Moore, E.F., Shannon, C.E., "Reliable Circuits Using Less Reliable Relays," *J. Franclin Inst., 262, n. 4b,* pp. 191-208, 1956.
8. Palmer C.R., Siganos G., Faloutsos M., Faloutsos C., and Gibbons P.: The connectivity and faulttolerance of the Internet topology. Workshop on Network-Related Data Management (NRDM-2001), http://www.research. att.com/ divesh/papers/cjfgs98-ir.ps.
9. Rodionov, A.S., Choo H.: On Generating Random Network Structures: Connected Graphs. International Conference on Information Networking ICOIN 2004, Revised selected papers. Springer LNCS. **3090** (2004) 483–491
10. Shooman, A.M., Kershenbaum, A.: Exact Graph-Reduction Algorithms for Network Reliability Analysis. Proc. GLOBECOM' 91. **2** (1991) 1412–1420