Refinements of a path-based efficient algorithm for network reliability estimation in the rare event case

Cancela, Héctor
Instituto de Computación, Facultad de Ingeniería, J. Herrera y Reissig 565, 11300 Montevideo, Uruguay. E-mail: cancela@fing.edu.uy

L'Ecuyer, Pierre
Département d'informatique et de recherche opérationnelle, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal (Québec), Canada, H3C 3J7. E-mail: lecuyer@iro.umontreal.ca

Lee, Matias
INRIA/IRISA Rennes, Campus universtaire de Beaulieu, 35042 Rennes Cedex, France. E-mail: mlee@irisa.fr

Rubino, Gerardo
INRIA/IRISA Rennes, Campus universtaire de Beaulieu, 35042 Rennes Cedex, France. E-mail: rubino@irisa.fr

Tuffin, Bruno
INRIA/IRISA Rennes, Campus universtaire de Beaulieu, 35042 Rennes Cedex, France. E-mail: btuffin@irisa.fr

This paper presents some refinements of a rare event simulation algorithm developped in [1] for estimating the probability of connection of two nodes s and t in an undirected graph G representing a communication network where nodes are perfect but links can fail. The method proposed in [1] makes use of disjoint paths (that is, with no common link) and samples a geometric variable representing the first time independent replications of a graph result in a configuration with at least one failed link in each path of a predetermined set. Generating such a random variable allows therefore to save a lot of computational time (but does not reduce variance) by avoiding the generation of all graphs. We propose here to investigate the variance reduction that can be obtained by replacing the geometric random variable by its expected value (conditional Monte Carlo). We also discuss how the set of disjoint paths can be determined and the robustness properties of the resulting estimator.

Model and previous work: using a geometric random variable to determine the first graph with a failed link on each path of a predefined set

Let $\forall h$, the probability that link h of graph G is failed be $u_h = a_h \varepsilon^{b_h}$ with $a_h, b_h > 0$ and ε a parameter representing rarity when it goes to zero. We want to quantify the probability u that the graph does not support the communication between two selected nodes s and t. The computation of u is known to be NP-hard, which naturally leads to make use of Monte Carlo simulation. The standard Monte Carlo method builds n independent copies of G, and simply estimates u by the proportion of graphs for which the selected nodes are not able to communicate. This estimator is unbiased, with variance is $\sigma_n^2 = u(1-u)/n$, but yields a relative variance $\sigma_n^2/u^2 \to \infty$ as $\varepsilon \to 0$ (because $u \to 0$), showing that it is an unrobust estimator when the system is highly reliable (i.e., when the failures are rare events).

In [1], a different estimator builds a set $\mathcal{P} = \{P_1, P_2, \cdots, P_H\}$ of elementary paths connecting nodes s and t, such that any pair of paths does not share any link. If $p_h = \prod_{i \in P_h} u_i$ is the probability that all links of path P_h work, the random variable F representing the first element in the sequence of independent graphs where every path in $\mathcal{P}$ has at least one link that does not work is geometrically distributed with parameter $q = \prod_{h=1}^{H}(1-p_h)$ ($\Pr(F > f) = (1-q)^f$, $f \geq 1$). Observe that $q \to 0$ as $\varepsilon \to 0$. On the average, we have to wait for $1/q$ samples to find one such graph and the F-th copy can easily be sampled, on a conditional basis, with no need to sample the $F-1$ first ones.

When sampling this way n graphs, the variance is still $\sigma_n^2 = u(1-u)/n$, but there is an important gain in computing time. To quantify this gain, we can look at the relative efficiency, as defined in [2]: $\mathrm{REff} = u^2/(\sigma_n^2 t_n)$ where t_n is the average time to build a n-sample. We say that we have bounded relative efficiency (BREff) if there exists a constant $d > 0$ such that REff is minored by d for all ε, meaning that we are able to obtain a bounded relative error for a given computational time budget, no matter how small ε is. A sufficient condition on $\mathcal{P}$ is provided in [2].

Variance reduction by considering the expected value of the geometric random variable

Consider now the case where F is replaced by its expected value. Then exactly one in $1/q$ independent graphs will have at least one failed link on each path of $\mathcal{P}$. Let $Z = qY$ be an (unbiased) estimator of u over such a block, where Y is a Bernoulli random variable that is 1 is the graph is failed and 0 otherwise, and such that at least one link on each path is failed. Getting a confidence interval for u is obtained by considering independent copies of Z and applying standard procedures.

Define p as the probability that $Y = 1$. Obviously, $u = qp$, and $\sigma^2[Z] = q^2\sigma^2[Y] = q^2 p(1-p)$.

Property: Comparing the ratios of efficiencies $1/(\sigma_1^2 t_1)$ of the estimator that considers the expected value over that of the one with the geometric distribution, we get (by assuming that the time to generate the geometric r. v. is negligible):

$$\frac{qu(1-u)}{q^2 p(1-p)} = \frac{1-qp}{1-p} > 1$$

because one geometric random variable F corresponds, on the average, to $1/q$ graphs. Considering the expected value therefore always yields an efficiency improvement that we are able to characterize.

On the choice of disjoint paths

A key issue is to determine for the two above methods a set of paths for which we have the BREff property.

Property: In the homogeneous case, that is assuming that $u_h = \varepsilon$ for each link h, a set of paths such that BREff is verified can be built as a by-product of a maximum flow computation, for instance using the Ford-Fulkerson algorithm. The proof consists of showing that the resulting set of paths verifies the sufficient condition in [2].

Property: In the heterogeneous case, when link unreliabilities u_h have different values of b_h, then it is not always possible to build a set of paths such that BREff is verified. A counter-example is provided in the next figure: all possible sets of link-disjoint paths lead to values of probabilities that are not sufficient to guarantee that the REff will remain bounded when epsilon goes to 0.

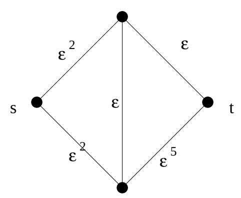

Figure 1: A “bridge” with the unreliabilities of its 5 edges

Keywords: Reliability analysis, rare events, Monte Carlo simulation.

REFERENCES

[1] M. El Khadiri and G. Rubino. An efficient formulation of the standard Monte Carlo simulation of binary systems reliability. In *2nd International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, 1996.

[2] H. Cancela, G. Rubino, and B. Tuffin. New measures of robustness in rare event simulation. In M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, editors, *Proceedings of the 2005 Winter Simulation Conference*, pages 519–527, 2005.

[3] L.R. Ford and D.R. Fulkerson. A simplex algorithm finding maximal networks flows and an application to the Hitchock problem, *Rand Report Rand Corporation*, Santa Monica, 1955.