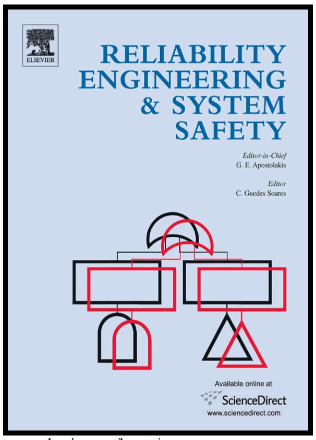
Author's Accepted Manuscript



www.elsevier.com/locate/ress

Accelerated Monte Carlo system reliability analysis through machine-learning-based surrogate models of network connectivity

R.E. Stern, J. Song, D.B. Work

PII: S0951-8320(17)30127-8
DOI: http://dx.doi.org/10.1016/j.ress.2017.01.021
Reference: RESS5743

To appear in: *Reliability Engineering and System Safety*

Received date: 5 May 2016
Revised date: 26 January 2017
Accepted date: 30 January 2017

Accelerated Monte Carlo system reliability analysis through machine-learning-based surrogate models of network connectivity

R. E. Stern[a], J. Song[b,*], D. B. Work[a]

[a] University of Illinois at Urbana-Champaign
[b] Seoul National University

Abstract

The two-terminal reliability problem in system reliability analysis is known to be computationally intractable for large infrastructure graphs. Monte Carlo techniques can estimate the probability of a disconnection between two points in a network by selecting a representative sample of network component failure realizations and determining the source-terminal connectivity of each realization. To reduce the runtime required for the Monte Carlo approximation, this article proposes an approximate framework in which the connectivity check of each sample is estimated using a machine-learning-based classifier. The framework is implemented using both a *support vector machine* (SVM) and a *logistic regression* based surrogate model. Numerical experiments are performed on the California gas distribution network using the epicenter and magnitude of the 1989 Loma Prieta earthquake as well as randomly-generated earthquakes. It is shown that the SVM and logistic regression surrogate models are able to predict network connectivity with accuracies of 99% for both methods, and are 1-2 orders of magnitude faster than using a Monte Carlo method with an exact connectivity check.

1. Introduction

In the aftermath of a natural disaster, a prompt response is critical to minimize economic damage and the loss of life [1, 2]. In order to improve post–disaster response times to hazards that impact lifeline networks such as gas, water, and electricity distribution systems, efficient methods are needed to quickly and accurately estimate the system-level network failure probability under a given set of individual component failure probabilities before crews have time to inspect the extent of the damage. This will enable first responders to make strategic and immediate risk-informed decisions about relief efforts before there is time to assess damage to the infrastructure based on comprehensive and visual inspections.

System reliability analysis (SRA) of infrastructure networks encompasses a number of methods to determine the probability that a network will be able to complete its designed function after a disaster event. The methods are used to analyze the resilience of the network with respect to failure under a catastrophic event [3, 4, 5, 6, 7].

To apply an SRA method, the infrastructure network is first modeled as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is the vertex (node) set and $\mathcal{E} \subseteq \{(i, j) : i, j \in \mathcal{V}\}$ is the set of edges with cardinality $n = |\mathcal{E}|$. This network is a *stochastic network* where the i^{th} edge fails randomly with probability $p_{f,i}$.

Given a network $\mathcal{G}$ and a vector of the correlated edge failure probabilities $p_f = [p_{f,1}, \cdots, p_{f,n}]^T$, the *two-terminal reliability* problem [8] is to find the probability that a connection (i.e., a path) remains between a source node $s \in \mathcal{V}$ and a terminal node $t \in \mathcal{V}$. This problem has been studied extensively for infrastructure, communications, and wireless networks [9, 10, 11, 12, 13, 14, 15].

For some special cases of graphs, the two-terminal reliability problem can be efficiently solved. For graphs consisting of series and parallel subgraphs, as well as graphs with a bounded path width (a measure of the decomposition of a graph), decomposition algorithms exist to solve reliability problems in linear time [16, 17]. However, in its most general form, the two-terminal reliability problem is $\sharp\mathcal{P}$-*complete* [18, 19], which is a complexity class of problems that involve counting the number of solutions to a problem in $\mathcal{NP}$. No exact polynomial-time algorithms are known to solve problems in $\sharp\mathcal{P}$-*complete*, which motivates the use of approximate or heuristic approaches.

One popular approximation approach is to use *Monte Carlo simulation* (MCS) to approximate the probability of disconnection. In this approach, a damaged network realization is simulated by removing edges in the network according to their failure probability. The connectivity of s and t are checked on each network realization, and the probability of disconnection of s and t is estimated as the fraction of disconnected samples compared to the total sample size. A bound on the number of samples required to achieve a converged estimate of the failure probability within a given error bound is provided by Karger [20].

To evaluate the connectivity of each realization, several exact methods exist [21, 22, 23, 24, 25]. However, these

*Corresponding author
Email address: junhosong@snu.ac.kr (J. Song)