Sensitivity of the RVR Monte-Carlo method to cutset selection strategy

Héctor Cancela

Universidad de la República
Montevideo, Uruguay

Mohamed El Khadiri

IUT of Saint Nazaire, University of Nantes
Saint-Nazaire, France

RESIM
7th International Workshop on Rare Event Simulation
September 24-26 2008, Rennes, France

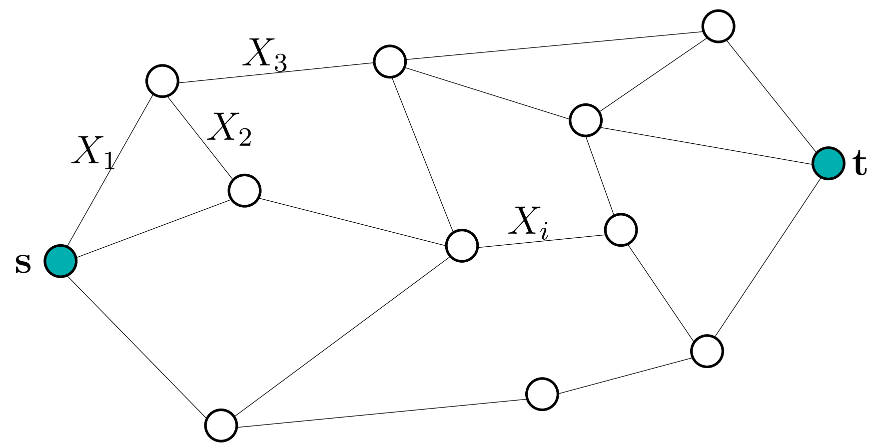
Talk Outline

- Network reliability model.
- Standard Monte Carlo simulation.
- Recursive Variance Reduction method.
- Cutset selection strategies.
- Experimental results.
- Conclusions and future work.

Classical network reliability model

- Network topology: simple graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{K})$.
- Network sites: set of nodes $\mathcal{V}$, $|\mathcal{V}| = n$.
- Network links: set of edges $\mathcal{E}$,$|\mathcal{E}| = m$.
- Set of terminals $\mathcal{K}$ (a fixed subset of $\mathcal{V}$).
- Hypothesis:
  - Nodes do not fail.
  - Links fail independently, with probability q_i.

Reliability model



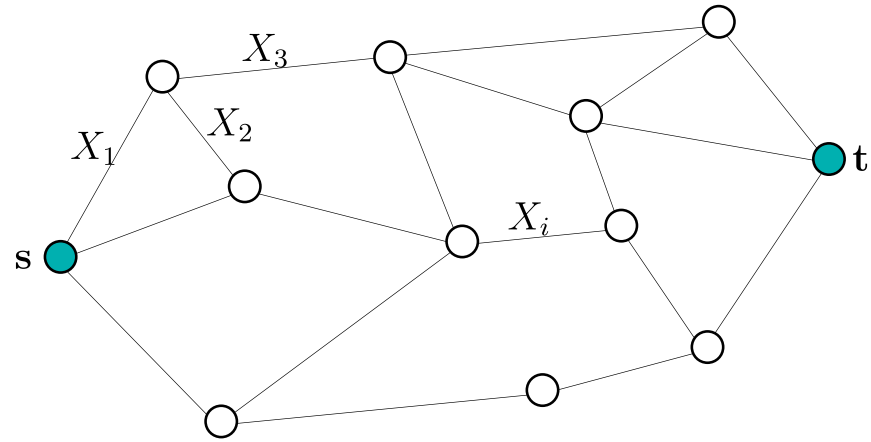
Links' state vector: $\mathbf{X} = (X_1, X_2, \ldots, X_m)$

$$X_i = \begin{cases} 1 \rightarrow & \text{link } i \textbf{ operational} \quad \mathbb{P}\{X_i = 1\} = r_i \\ 0 \rightarrow & \text{link } i \textbf{ failed} \quad \mathbb{P}\{X_i = 0\} = q_i = 1 - r_i \end{cases}$$

Network structure function, $\Phi(\mathbf{X})$

$$\Phi(\mathbf{X}) = \begin{cases} 1 \rightarrow & \textbf{operational} \text{ network (i.e., } \mathcal{K}\text{-connected).} \\ 0 \rightarrow & \textbf{failed} \text{ network (i.e., } \mathcal{K}\text{-unconnected)} \end{cases}$$

Reliability model



$\mathcal{K}$-terminal reliability / unreliability parameters:

$$\begin{cases} R(\mathcal{G}) = \mathbb{P}\{\text{network is } \mathcal{K}\text{-connected}\} = \mathbb{E}\{\Phi(\mathbf{X})\} \\ Q(\mathcal{G}) = \mathbb{P}\{\text{network is not } \mathcal{K}\text{-connected}\} = \mathbb{E}\{1 - \Phi(\mathbf{X})\} \end{cases}$$

Evaluation of $R(\mathcal{G})$ and $Q(\mathcal{G})$ is NP-hard.
Exact methods can only be applied to small networks.

Standard Monte Carlo simulation

- $\mathbf{X}^{(j)}$, $1 \leq j \leq N$: iid replications of r.v. $\mathbf{X} = (X_1, \ldots, X_m)$

$$\widehat{R} = \frac{1}{N}\sum_{j=1}^{N} \Phi(\mathbf{X}^{(j)}) \qquad \widehat{Q} = \frac{1}{N}\sum_{j=1}^{N}(1 - \Phi(\mathbf{X}^{(j)}))$$

$$\mathbb{V}\left\{\widehat{R}\right\} = \mathbb{V}\left\{\widehat{Q}\right\} = \frac{R(\mathcal{G})Q(\mathcal{G})}{N}$$

- Highly reliable network:
  - "most" $X_i^{(j)} = 1$, "almost always" $\Phi(\mathbf{X}^{(j)}) = 1$.
  - $\Phi(\mathbf{X}^{(j)}) = 0$ **rare event**, $Q \to 0$:

Relative error $RE = \dfrac{\mathbb{V}\{\widehat{Q}\}^{1/2}}{\mathbb{E}\left\{\widehat{Q}\right\}} = \left(\dfrac{1-Q}{NQ}\right)^{1/2} \approx \dfrac{1}{(NQ)^{1/2}} \longrightarrow \infty$

Variance reduction simulation approaches

- Based on reliability bounds (leads to sampling in subset of links' vector state space). Van Slyke and Frank; Kumamoto, Tanaka and Inoue; Fishman.
- Based on antithetic sampling or generalizations (improve efficiency in generation of uniform variates and lowers the varianza). Kumamoto, Tanaka and Inoue; Rubino and El Khadiri; Wei-Chang Yeh.
- Based on graph evolution models (stochastic processes), with importance sampling to reduce variance. Wong and Easton; Elperin, Gertsbakh and Lomonosov. Cross-Entropy based variants to optimize IS parameters. Hui, Bean, Kraetzl, and Kroese.

Variance reduction simulation approaches (2)

- Based on partitioning links' vector state space, or on reformulating problem in terms of other random variables with smaller variance. Karp and Luby; Jun and Ross; Cancela and El Khadiri; Tuffin, Rubino et al.
- Reformulations of the standard method to improve computational efficiency. Rubino and El Khadiri.

Recursive Variance Reduction method

This estimator is a recursive method which uses at each call:

- series–parallel reductions which preserve $Q(\mathcal{G}) = Q(\tilde{\mathcal{G}})$ where $\tilde{\mathcal{G}}$ network resulting from applying series and parallel reductions to $\mathcal{G}$;
- a conditioning process exploiting a $\mathcal{K}$-cutset $C_{\tilde{\mathcal{G}}}$ in the network $\tilde{\mathcal{G}}$ for transforming the initial problem into smaller sub-problems where new series–parallel reductions may appear;
- adequate selection of the sub-problems for subsequent recursive calls.

The recursion terminates on trivial cases.

Preliminary definitions

- $C \subseteq \mathcal{E}$ is $\mathcal{K}$-cutset of $\mathcal{G}$ if subnetwork $\mathcal{G}' = (\mathcal{V}, \mathcal{E} - C, \mathcal{K})$ is not $\mathcal{K}$-connected .
- $E_0(C)$: event "all components in C are in failed state";
- $q_0(C) = \mathbb{P}\{E_0(C)\} = \prod_{i=1}^{|C|} q_{l_i}$;
- $E_v(C)$: event "all components in $\{l_1, \ldots, l_{v-1}\}$ are failed and l_v is up", $1 \leq v \leq |C|$;
- $q_v(C) = \mathbb{P}\{E_v(C)\} = (1 - q_{l_v}) \prod_{i=1}^{v-1} q_{l_i}$, $1 \leq v \leq |C|$;
- For a given link l in $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{K})$, $\mathcal{G} - l$ denotes the network with node-set $\mathcal{V}$ and link-set derived from $\mathcal{E}$ by removing link $l \in \mathcal{E}$.
- For a given link l in $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{K})$, $\mathcal{G} * l$ denotes the network derived from $\mathcal{G}$ by contracting link $l = \{u, v\} \in \mathcal{E}$ (eliminate l and merge its extremities u and v into a new node w).

RVR definition

$$F(\mathcal{G}) = \begin{cases} 1 & \text{if } \mathcal{G} \text{ not } \mathcal{K}\text{–connected;} \\ 0 & \text{if } \mathcal{G}\ \mathcal{K}\text{–connected;} \\ 1 - \prod_{l \in \tilde{\mathcal{G}}} (1 - q_l) & \text{if } \mathcal{G} \text{ is } \textit{sp-reducible}; \\ q_0(C_{\tilde{\mathcal{G}}}) + (1 - q_0(C_{\tilde{\mathcal{G}}}))F(\mathcal{G}_{V(C_{\tilde{\mathcal{G}}})}) & \text{otherwise} \end{cases}$$

where $C_{\tilde{\mathcal{G}}}, l_{|C_{\tilde{\mathcal{G}}}|}\}$ a fixed $\mathcal{K}$-cutset in $\tilde{\mathcal{G}}$,
$V(C_{\tilde{\mathcal{G}}})$ independent r.v. with distribution
$\mathbb{P}\left\{V(C_{\tilde{\mathcal{G}}}) = v\right\} = (1 - q_{l_v}) \prod_{i=1}^{v-1} q_{l_i} / (1 - q_0(C_{\tilde{\mathcal{G}}}))$, $1 \leq v \leq |C_{\tilde{\mathcal{G}}}|$;
and $\mathcal{G}_v$: $\left(\tilde{\mathcal{G}} - l_1 - \ldots - l_{v-1}\right) * l_v$, $1 \leq v \leq |C_{\tilde{\mathcal{G}}}|$.

RVR main properties

For any network $\mathcal{G}$, let $F(\mathcal{G})$ be the r.v. defined in previous slide. Then $F(\mathcal{G})$ verifies

$$\mathbb{E}\{F(\mathcal{G})\} = Q(\mathcal{G})$$

and

$$\mathbb{V}\{F(\mathcal{G})\} \leq (Q(\mathcal{G}) - q_0(C_{\tilde{\mathcal{G}}}))R(\mathcal{G}) \leq Q(\mathcal{G})R(\mathcal{G}) = \mathbb{V}\{Y(\mathcal{G})\}.$$

If $F(\mathcal{G})^{(1)}, \dots, F(\mathcal{G})^{(N)}$ are s-independents trials of $F(\mathcal{G})$, we define a sample mean

$$\widehat{F}(\mathcal{G}) = \frac{1}{N}\sum_{k=1}^{N} F(\mathcal{G})^{(k)}$$

unbiased estimator of $Q(\mathcal{G})$, more accurate than standard Monte Carlo estimator $\widehat{Y}(\mathcal{G})$.

Cutset strategies

From $\mathbb{V}\{F(\mathcal{G})\} \leq (Q(\mathcal{G}) - q_0(C_{\tilde{\mathcal{G}}}))R(\mathcal{G}) \leq Q(\mathcal{G})R(\mathcal{G}) = \mathbb{V}\{Y(\mathcal{G})\}$, variance-reduction ratio of RVR with respect to SMC is at least

$$Q(\mathcal{G}) / \left(Q(\mathcal{G}) - q_0(C_{\tilde{\mathcal{G}}})\right).$$

Bound depends on $\mathcal{K}$-cutset $C_{\tilde{\mathcal{G}}}$ chosen, is maximal when $C_{\tilde{\mathcal{G}}}$ has largest $q_0(C_{\tilde{\mathcal{G}}})$.
We compare two strategies :

- RVR-SOURCE, set of adjacent links to one of the nodes in terminal-set;
- RVR-MAX, $\mathcal{K}$-mincutset $C_{\tilde{\mathcal{G}}}$ with largest $q_0(C_{\tilde{\mathcal{G}}}) = \prod_{i=1}^{|C_{\tilde{\mathcal{G}}}|} q_{l_i}$ or equivalently a minimum cost $\mathcal{K}$-mincutset $C_{\tilde{\mathcal{G}}}$ where each link l is valued by $-log(q_l)$.

Experiments

- Complete graph topology with 10 nodes and equi-reliable links.
- Three different cases for the $\mathcal{K}$-terminal set: $\mathcal{K} = \{1, 10\}$, $\mathcal{K} = \{1, 3, 5, 7, 10\}$ and $\mathcal{K} = \mathcal{V}$.
- Three links' elementary reliabilities levels: 0.50, 0.90, 0.95 (respectively $q_l = 0.50, 0.10, 0.05$).
- Sample size $N = 10^6$.

RVR-SOURCE results

q_l	$\mathcal{K}$	$Q(\mathcal{G})$	$\widehat{F}(\mathcal{G})$	$RE_{\widehat{F}(\mathcal{G})}(\%)$
0.50	$\{1, 10\}$	$4.03755e-03$	$4.03323e-03$	$1.07e-02$
0.50	$\{1, 3, 5, 7, 10\}$	$9.98206e-03$	$9.54892e-03$	$4.34e-02$
0.50	$\mathcal{V}$	$1.95508e-02$	$1.86542e-02$	$4.59e+00$
0.10	$\{1, 10\}$	$2.00000e-09$	$2.00000e-09$	$4.30e-05$
0.10	$\{1, 3, 5, 7, 10\}$	$5.00000e-09$	$2.00000e-09$	$6.00e+01$
0.10	$\mathcal{V}$	$1.00000e-08$	$2.00000e-09$	$8.00e+01$
0.05	$\{1, 10\}$	$3.90625e-12$	$3.90643e-12$	$4.63e-03$
0.05	$\{1, 3, 5, 7, 10\}$	$9.76563e-12$	$3.90643e-12$	$6.00e+01$
0.05	$\mathcal{V}$	$1.95313e-11$	$3.90643e-12$	$8.00e+01$

‘

RVR-MAX results

q_l	$\mathcal{K}$	$Q(\mathcal{G})$	$\widehat{F}(\mathcal{G})$	$RE_{\widehat{F}(\mathcal{G})}(\%)$
0.50	$\{1,10\}$	$4.03755e-03$	$4.03920e-03$	$4.09e-02$
0.50	$\{1,3,5,7,10\}$	$9.98206e-03$	$9.98653e-03$	$4.48e-02$
0.50	$\mathcal{V}$	$1.95508e-02$	$1.95598e-02$	$4.57e-02$
0.10	$\{1,10\}$	$2.00000e-09$	$2.00000e-09$	$2.26e-05$
0.10	$\{1,3,5,7,10\}$	$5.00000e-09$	$4.99971e-09$	$5.77e-03$
0.10	$\mathcal{V}$	$1.00000e-08$	$1.00030e-08$	$2.97e-02$
0.05	$\{1,10\}$	$3.90625e-12$	$3.90625e-12$	$3.07e-07$
0.05	$\{1,3,5,7,10\}$	$9.76563e-12$	$9.76474e-12$	$9.10e-03$
0.05	$\mathcal{V}$	$1.95313e-11$	$1.95453e-11$	$7.19e-02$

Comparison

- Source-terminal networks have smaller relative errors for both RVR-SOURCE and RVR-MAX variants.
- RVR-SOURCE has comparable (sometimes smaller) relative errors than RVR-MAX for source-terminal networks.
- RVR-SOURCE has much larger errors for networks with larger number of terminals. Also, estimations are identical to those for two terminal networks, showing this cutest selection strategy does not adapt well.
- RVR-MAX has good behaviour - small errors even with larger number of terminals.
- RVR-MAX has quite stable errors for the same topology and number of terminals when link reliabilities vary.

Conclusions

- RVR results based on RVR-SOURCE selection strategy already amongst best results in literature for source-terminal reliability.
- Cutset selection strategy has large impact on RVR estimation quality.
- For multiple terminals, RVR-SOURCE is not suitable, but RVR-MAX has good results.
- RVR-MAX not necessarily the best strategy (in particular, in source-terminal cases, RVR-SOURCE is sometimes better).
- RVR-MAX maximizes the "one-step" variance reduction bound; but the overall variance reduction results from the interaction of all recursive steps.

Future work

- More extensive experimentation, to better understand the tradeoff between improved bounds at each recursive step vs. the interaction among successive recursive calls.
- Search for other characteristics that good cutset selection strategies should have.
- Test computational performance accelerations to be attained using pre-computed cutsets.

Questions?