Estimating network resilience, a performability metric

G. Rubino, Inria, Rennes, France

Paris, June 28

MCM'23

Talk's content

1. The resilience metric
2. Monte Carlo standard and rare events
3. Back to the resilience
4. Conclusions

Talk's content

Introduction

- In the quantitative analysis of systems, two broad territories appear: performance analysis (“how much” the system does, *assuming it is perfect*) and dependability studies (how it behaves face to failures and in some cases repairs, *ignoring its work*).
- In some cases, we take simultaneously into account both aspects of systems, and we then speak about performability.
- In this work, we consider a dependability analysis on a static setting, and we work with an extension of classic reliability metrics to the performability family.

- The abstract model consists of a graph (assumed undirected here, to simplify), where nodes are perfect but edges fail independently of each other (the most used assumption; many other ones are possible).
- For each edge i, r_i is the probability that it works and $1 - r_i$ is the probability that it doesn't, behaving as absent from the graph. This means a binary world.
- Classic reliability consider connectivity-based metrics such as
  - $R_{s,t} = \mathbb{P}(\text{there is at least a working path between nodes } s \text{ and } t)$,
  - $R_{\text{all}} = \mathbb{P}(\text{there is at least a working path between any pair of nodes})$.
- They are all-nothing measures. How to do better tan that? We consider R_{all} and instead of the zero-one associated property, we look at the number of pairs of nodes that can communicate (the # of pairs of nodes for which there is a working path connecting them).

Context

- Origin: project with Nokia on dependability propzrties of new types of 5G/6G radio networks.
- Observation: in this context, it is some times hard to discriminate between different architectures using reliability, because of rarity of system failures.
- 2 ideas here: move to performability, and move to conditional metrics (evaluate the metrics conditioning on the fact that something failed), and working if possible with static models.
- (In the project, we finally had to move to dynamical models – transient analysis of stochastic models.)
- Here, we keep the static context and discuss some work following the two ideas above.

Resilience

- Let NCP = number of pairs of nodes that can communicate in G (obviously assumed to be connected).
- The resilience Res of the model is

$$Res = \mathbb{E}(NCP).$$

- This is a typical performability metric. Instead of a binary situation, we can distinguish between a large number of possibilities (of performance levels), from 1 to $1 + \binom{n}{2}$ if there are n nodes in the model.

Example

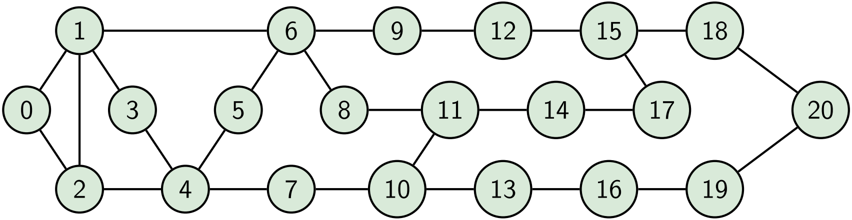

Figure: A widely used Arpanet topology in network reliability, from the history of this famous communication network.

Some properties of resilience

- Range: $0 \leq NCP \leq \binom{n}{2}$.
- $\mathbb{P}(NCP = 0) = \prod_{\text{all edge } i}(1 - r_i)$.
- $\mathbb{P}\left(NCP = \binom{n}{2}\right) = R_{\text{all}}$.
- $Res \,|_{\forall \text{ edge } i \text{ s.t. } r_i=0} = 0, \qquad Res \,|_{\forall \text{ edge } i \text{ s.t. } r_i=1} = \binom{n}{2}$.
- Immediate bounds: $\binom{n}{2} R_{\text{all}} \leq Res \leq \binom{n}{2}$.

- Define the r.v. $Y_{s,t} = 1\big(\text{there is a path connecting nodes } s \text{ and } t\big)$. We have $R_{s,t} = \mathbb{P}(Y_{s,t} = 1) = \mathbb{E}(Y_{s,t})$ and then,

$$NCP = \sum_{\text{all nodes } s,t,\ s<t} Y_{s,t},$$

leading to

$$Res = \sum_{\text{all nodes } s,t,\ s<t} R_{s,t}.$$

- We can normalize the metric, dividing $\mathbb{E}\big(NCP\big)$ by $\binom{n}{2}$, thus leading to an index in $[0,1]$.
- The *scaled resilience* of the network, $ResScaled$, is then $ResScaled = 2Res/\big(n(n-1)\big)$, and we have

$$R_{\text{all}} \le ResScaled \quad (\le 1).$$

Example: the bridge topology

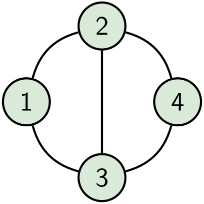

Figure: The bridge topology

- Homogeneous case.
- Using brute force (listing all 2^5 possible links' states), we obtain $Res = r\big(5 + 8r - 14r^3 + 7r^4\big)$.
- $ResScaled = \dfrac{Res}{6} = \dfrac{5 + 8r - 14r^3 + 7r^4}{6}$.

Example: a path with $n \geq 2$ nodes

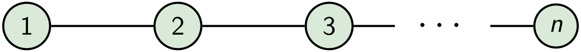

Figure: A path with $n \geq 2$ nodes

- Homogeneous case. Conditioning (factorization), we obtain

$$\mathit{Res}_n = \frac{r\left[n(1-r) - (1-r^n)\right]}{(1-r)^2} = \frac{r\left[n-1-r(n-r^{n-1})\right]}{(1-r)^2}.$$

- Scaling,

$$\mathit{ResScaled}_n = \frac{2r\left[n(1-r) - (1-r^n)\right]}{n(n-1)(1-r)^2}.$$

- After some more combinatorics, using $r = 1 - \varepsilon$, we have

$$1 - \mathit{ResScaled}_n = 1 - \frac{n+1}{3}\varepsilon + O(\varepsilon^2).$$

Example: a ring with $n \geq 3$ nodes

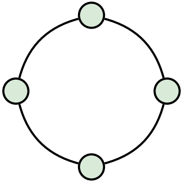

Figure: Here, a ring with 4 nodes

- Homogeneous case. Using series-parallel formulas, we obtain (also holds if $n = 2$)

$$\mathit{Res}_n = nr\left(\frac{1 - r^{n-1}}{1 - r} - \frac{n-1}{2} r^{n-1}\right).$$

- Scaling, $\mathit{ResScaled}_n = \dfrac{2r\left(1 - r^{n-1}\right)}{(n-1)(1-r)} - r^n$.
- With $r = 1 - \varepsilon$, $1 - \mathit{ResScaled}_n = 1 - \dfrac{n+1}{6}\varepsilon^2 + O(\varepsilon^3)$.

Talk's content

Standard Monte Carlo

- Let us consider a classic metric such that $R_{s,t}$ or R_{all}, here R.
- Let $X_i = 1\big(\text{line } i \text{ works}\big)$ and $Y = 1\big(\text{the network } \textbf{doesn't} \text{ work}\big)$. We have $\mathbb{P}(X_i = 1) = r_i$ and $\mathbb{P}(Y = 1) = 1 - R = \gamma$ (the target).
- Let $E = \#$ of edges (the components of the systems) and denote by X the vector $X = (X_1, X_2, \cdots, X_E)$.
- Make N independent copies $X^{(1)}, X^{(2)}, \cdots, X^{(N)}$ of X; for each copy $X^{(n)}$, compute the network state $Y^{(n)}$.
- We then estimate γ by

$$\widetilde{\gamma}_N = \frac{1}{N} \sum_{n=1}^{N} Y^{(n)}.$$

- This works fine if γ is not (too) close to 0, and becomes even useless when $\gamma \approx 0$ (the rare event case).
- The relative error of this estimation is $\Theta\big((N\widetilde{\gamma}_N)^{-1/2}\big)$, that $\to \infty$ as $\gamma \to 0$, for fixed N.
- This is typically resumed in sentences such as "the standard estimator behaves poorly in the rare event case". We claim here that this is not precise enough, it depends on how we implement the estimator.

A specific implementation

- Suppose we implement the standard Monte Carlo procedure the following (terrible) way:
  - we will first build a (huge) $N \times (E+1)$ table with
    - N rows (think $N \gg 1$), one per realisation,
    - and $E+1$ columns, one for each component's state and the last one for the system's state;
  - we fill the N rows ignoring column $E+1$;
  - element (n, i), $1 \leq i \leq E$, contains a realisation of the binary random variable $X_i^{(n)}$;
  - once the first E columns filled, column $E+1$ is filled with, at row n, the corresponding value of $Y^{(n)}$;
  - once the table filled, we compute the estimation $\widetilde{\gamma}_N$.

	X_1	X_2		$\cdots$	X_E	Y
1	1	1	1	$\cdots$		0
2	1	1	1	$\cdots$		0
3	1	1	0	$\cdots$		1
$\vdots$				$\vdots$		$\vdots$
$N\ (\gg 1)$	1	1	1	$\cdots$		1

Remarks

- Consider the case of $r_i \approx 1$ for all i, so $R \approx 1$ as well, and $\gamma \approx 0$.
- Let's now look at the table but in a column-by-column way. Assume, to simplify the presentation, $N = \infty$.
- Let F_i be the first element in column $i \in \{1, \cdots, E\}$ with a '0'; we have $\mathbb{P}(F_i = f) = r_i^{f-1}(1 - r_i), \quad f \geq 1.$
- Let $M = \min\{F_1, \cdots, F_E\}$; M is also geometric:
$$\mathbb{P}(M = m) = r^{m-1}(1 - r), \quad m \geq 1, \quad r = r_1 r_2 \cdots r_E.$$

A different viewpoint

- This suggests a different way of implementing the same standard Monte Carlo estimator, using previous observations and the fact that the geometric r.v.s on $\mathbb{N}_{\geq 1}$ are memoryless.
- Sample M. Call m the obtained value. Interpretation: the first $m-1$ values of last column are all 0s.
- For row m, we need to sample vector X, knowing that in F there is at least one component down (at least one X_i is equal to 0).

- Let J be the first component down in X and denote by Z the number of components down. We have ($j \geq 1$):

$$\mathbb{P}(J = j \mid Z \geq 1) = \frac{r_1 r_2 \cdots r_{j-1}(1 - r_j)}{1 - r}.$$

- For row m we sample from this conditional distribution obtaining some $j \in \{1, 2, \ldots, E\}$, we set to 1 the first $j - 1$ values, to 0 the jth, and we sample the rest according to their Binary a priori law, and independently. Then, we can evaluate $Y^{(m)}$.
- After this first round, we sample M again, obtaining m', and repeat the process, "virtually filling" rows $m + 1, m + 2, \ldots, m + m'$" of the table, etc.
- We can then prove that the estimator obtained this way (truncating the "virtual infinite table" to its first N rows), is the standard one (in distribution).
- So the variance remains identical, but the cost in time is reduced. We performed a time reduction, not a variance one.

Cost

- The standard implementation of the standard Monte Carlo estimator needs $O(NE)$ operations.
- After some algebra, using $\mathbb{E}(M) = (1-r)^{-1}$, the new implementation needs $O(N(1-r)E)$ operations.
- Dividing, the gain is

$$\approx \frac{NE}{N(1-r)E} = \frac{1}{1-r}.$$

Talk's content

1. The resilience metric
2. Monte Carlo standard and rare events
3. Back to the resilience
4. Conclusions

Standard estimator of Res

- Call NCC the number of connected components of the underlying graph. Evaluating NCC on a given graph costs $O(E)$ (a single DFS or BFS does the job), and measuring their sizes (in # of nodes) has no complexity overhead (with respect to E).
- Suppose a given graph has ncc connected components, with sizes $c_1, c_2, \ldots, c_{ncc}$. Then, the # of communicating pairs in the graph is

$$ncp = \sum_{h=1}^{ncc} \binom{c_h}{2}.$$

- Estimating Res using the standard estimator $\widehat{Res}$ provides the confidence interval at 95% $(\widehat{Res} \pm 1.96\,S)$, where

$$S^2 = \frac{1}{N(N-1)} \sum_{n=1}^{N} NCP^{(n)} - \frac{1}{N-1} \sum_{n=1}^{N} \widehat{Res}^2.$$

In a nutshell: improvement

- We can improve the proposed implementation of the standard estimator $\widehat{Res}$ following the same approach.
- For instance, assume that the breadth b of the graph[1] (whose evaluation is polynomial) is known. Instead of defining F as the first row in the virtual table where $Z \geq 1$, we can use $Z \geq b$, increasing the gain.
- Another possibility: find a covering tree of the underlying graph with minimal weight, where the weights are multiplicative (the elementary reliabilities). This is again, a polynomial task. Then, redefine F accordingly. We also tested this idea and the gains can be considerable.
- Previous idea can be extended to finding several covering trees.

[1]= the size of any mincut of minimal size.

In a nutshell: sensitivities

- Computing sensitivities can be more interesting for the engineers than the metrics itself.
- In a nutshell, the "virtual table" view allows to do this task with very little overhead, by means of the result that follows.
- Define
$$\sigma_i = \frac{\partial Res}{\partial r_i} \qquad \text{and} \qquad \sigma_{s,t;i} = \frac{\partial R_{s,t}}{\partial r_i}.$$
- Denote, similarly as for usual reliability metrics,
$$Res_i^c = Res(\mathcal{G} \mid X_i = 1)$$
(this is $\neq Res(\mathcal{G}_i^c)$ where $\mathcal{G}_i^c$ is the result of contracting i in $\mathcal{G}$), and
$$Res_i^d = Res(\mathcal{G} \mid X_i = 0) = Res(\mathcal{G}_i^d).$$

- Theorem 1: $\sigma_i = \dfrac{Res_i^c - Res}{1 - r_i} = \dfrac{Res - Res_i^d}{r_i}$.
- Theorem 2: The expression

$$\widehat{\sigma}_i = \frac{X_i - r_i}{r_i(1 - r_i)} NCP$$

  defines an unbiased estimator of σ_i.
- This means that the computation of the gradient of Res comes almost for free following the "virtual table" approach.

More metrics

- In the rare event case, we obviously have Res very close to $\binom{n}{2}$.
- A different way of having a deeper view of what happens in this case is to explore the situation where there is something broken, given the fact that we are using performability metrics here.
- We propose to look, for instance, at conditional metrics such as $\mathbb{E}(NCP \mid NCC \geq 2)$.
- Another possible direction to look at is to count the number of pairs of nodes between which there are at least 2 edge-disjoint working paths, which we denote by $NCP2$ here, and, for instance, to explore the conditional metric $\mathbb{E}(NCP2 \mid NCC \geq 2)$.
- Common point with previous metrics? Their evaluation (and that of their gradients) are immediate using the virtual table implementation of the standard estimator (and the costs are always polynomial).

An example

To provide a few illustrations of previous results, let us consider the following model:

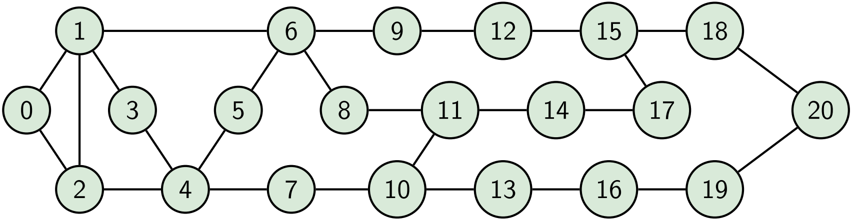

Figure: A widely used Arpanet topology in network reliability, from the history of this famous communication network.

Consider the i.i.c. case (independent and identical components), with common elementary reliabilities $= p$.

- We take $p = 0.999$.
- Using the breadth, here $b = 2$, the gain with respect to the standard estimation of Res was of several hundreds, varying with the implementation, using a library for some computations (`numpy`) or not, `python` or `C`, etc.

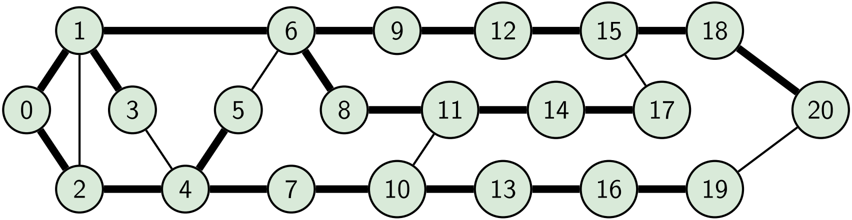

Figure: A covering tree on the Arpanet model.

- Using this tree, we moved the gain up to some more hundreds (typically to 500 or 600).

Conditional metrics

- To get an idea of the discriminatory power of the conditional metrics, look at this table (for the Arpanet graph, the number of nodes is $n = 21$ and $\binom{n}{2} = 210$):

Table: Three metrics on the Arpanet, i.i.c. case (homogeneous links).

p	0.91	0.95	0.99	0.995	0.999
$\mathbb{E}(NCC)$	1.251	1.073	1.0027	1.00066	1.000026
$\mathbb{E}(NCP) = Res$	199.8	207.2	209.9	209.977	209.9991
$\mathbb{E}(NCP \mid NCC \geq 2)$	160.3	167.6	174.5	175.2	175.5

Sensitivities

Consider this graph:

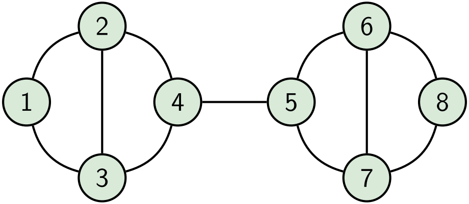

Figure: Two bridges connected by a “bridge” link. We assume i.i.c., with elementary reliability $p = 0.999$.

For instance,

edge	$\{1, 2\}$	$\{2, 3\}$	$\{4, 5\}$
sensitivity	$7.01\ 10^{-3}$	$3.79\ 10^{-5}$	16.0

Talk's content

Some conclusions:

- Interest in a performability viewpoint.
- Power of the "virtual table" approach, and the idea that the problem of the standard estimator in the rare event case concerns the combination time-variance, not the variance alone.
- The method presented here offers a good improvement on the classic one, but of course, recent techniques (e.g. the so-called Zero Variance IS) go much further in efficiency. The method here has the interest of allowing a direct analysis of many other quantities, mentioned below.
- Results on sensitivity analysis.
- Idea of conditional performability metrics.

Talk based on the chapter "*Network Reliability, Performability Metrics, Rare Events and Standard Monte Carlo*", in "Advances in Modeling and Simulation – Festschrift for Pierre L'Ecuyer", edited by Zdravko Botev, Alexander Keller, Christiane Lemieux and Bruno Tuffin, published by Springer, December 1st, 2022, DOI https://doi-org.passerelle.univ-rennes1.fr/10.1007/978-3-031-10193-9.