Proc. of the 6th Intl. Symposium on Reliability Engineering and Risk Management (6ISRERM)
31 May – 1 June 2018, Singapore
Editor(s) Xudong Qian, Sze Dai Pang, Ghim Ping Raymond Ong, Kok-Kwang Phoon

Efficient approximation of the survival signature for large networks

JASPER BEHRENSDORF[1], SEBASTIAN BRANDT[1], MATTEO BROGGI[1] and MICHAEL BEER[1,2,3]

[1]*Institute for Risk and Reliability, Leibniz University Hannover, Callinstr. 34, Hanover, Germany.*
E-mail: behrensdorf@irz.uni-hannover.de, brandt@irz.uni-hannover.de, broggi@irz.uni-hannover.de, beer@irz.uni-hannover.de
[2]*Institute for Risk and Uncertainty, University of Liverpool, Peach Street, Liverpool, UK.*
[3]*International Joint Research Center for Engineering Reliability and Stochastic Mechanics (ERSM), Tongji University, Shanghai, China.*

The reliability analysis of complex networks, e.g. water supply networks, transportation networks or electrical distribution networks, is of key importance to the resilience of communities. The concept of survival signature provides a novel basis for analyzing complex networks efficiently. The survival signature outperforms traditional analyses techniques, in particular, when estimating the reliability of networks. Its most unique feature is the separation of the network structure from its probabilistic properties, opening pathways for the consideration of, for instance, general dependencies, common cause failures, or vaguely specified probabilities. However, the numerical effort to calculate the survival signature is still prohibitive for large systems. While the issue of numerical efficiency can be addressed well with analytical approaches such as the use of binary decision diagrams, these approaches are limited by the number of components and types. In this paper we propose an approximation of the survival signature using a combination of graph theory and Monte Carlo simulation. By application of graph theory, we are able to predetermine certain fractions of the survival signature without explicitly evaluating it. The remaining fraction is then analyzed with Monte Carlo simulation in a targeted manner, circumventing high-effort-low-contribution calculations. The developed approach excels, in particular, in cases with a large number of different component types. Using an example we highlight the significant reduction in computational effort required to accurately determine the survival signature.

Keywords: networks, reliability, Monte Carlo simulation, survival signature

1 Introduction

As the complexity of large infrastructure networks such as electrical, gas, and water distribution networks, traffic networks and communication networks increases, so does the need for efficient reliability analysis techniques for these networks. A plethora of algorithms for numerical reliability analysis of systems and networks is already available from literature (Zio, 2013). One common element in these algorithms is the requirement for some form of evaluation of the system structure. However, traditional techniques such as fault tree analysis or reliability block diagrams do not scale well with the size of the system, e.g., in the case of large network system and system of systems.

A recently developed method known as survival signature (Coolen & Coolen-Maturi, 2013), an extension of the system signature (Samaniego, 2007) allows for a more efficient analysis of the network structure, outperforming traditional techniques. The survival signature serves as a compressed representation of system availability and is especially useful due to a separation of the structural information and the probabilistic properties of a network. This in turn facilitates consideration of dependencies, common cause of failures or imprecision. Nonetheless, network size is still a substantial factor when calculating the survival signature.

Recently, several survival signature based network reliability algorithms have been published, see for example Feng et al., (2016), Patelli et al. (2017), or Behrensdorf et al. (2017). However, these publications are focused on performing the reliability analysis, without taking into account the amount of work required for the computation of the survival signature.

In Reed (2017) the author presents an efficient method to analytically calculate the survival signature, based on converting the fault tree representation of a system to a binary decision diagram. While this works excellently for small systems containing a limited number of component types, improving on the direct evaluation of the survival signature (see e.g. Aslett (2012) for an R implementation of the direct algorithm), the method becomes increasingly impractical with growing system complexity, as the computational cost of building the binary decision diagram grows exponentially with the number of system nodes. Nevertheless, in cases where the fault tree or even the binary decision diagram is already known, this method performs very well.

In this work we present a novel approach to the approximation of the survival signature based on graph theory and Monte Carlo simulation. The use of graph theory allows to pre-eliminate a significant number of system configurations for computing the signature while Monte Carlo simulation helps in avoiding high-effort-low-contribution calculations, thus increasing numerical efficiency.

The remainder of the paper is structured as follows. In section 2 we present the necessary theory on the survival signature. The proposed methodology is presented in section 3 followed by a numerical example in the subsequent section. The paper concludes with a summary and an outlook into possible future research.

2 Survival signature

This section presents the survival signature as introduced in Coolen & Coolen-Maturi (2013). Consider a system with m components where l out of m components are in a working state, then the survival signature is defined as

$$\Phi(l) = \binom{m}{l}^{-1} \sum_{\underline{x} \in S_l} \varphi(\underline{x}), \tag{1}$$

with $\underline{x} = (x_1, \ldots, x_m)$ denoting the state vector of the system where $x_i = 1$ and $x_i = 0$ indicate working and failed components respectively. Then, $\varphi(\underline{x})$ is called the structure function, which returns the system's state as a 1 or 0 for a given state vector.

Eq. (1) can easily be adapted to systems with Kcomponent types. Considering m_k components for each $k = 1, \ldots, K$ types and l_k components working, the resulting survival signature is defined by

$$\Phi(l_1, \ldots, l_k) = \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1} \right] \times \sum_{\underline{x} \in S_{l_1, \ldots, l_k}} \varphi(\underline{x}), \tag{2}$$

Next, the survival function, which returns the system reliability, is defined based on the signature as

$$P(T_s > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_k=0}^{m_k} \Phi(l_1, \dots, l_k) P\left(\bigcap_{k=1}^{K} \{C_t^k = l_k\}\right). \tag{3}$$

The aforementioned separation of structural and probabilistic information is clearly visible in the equation, highlighting the usefulness of the survival signature approach. Most research concerning network reliability is currently focused on devising numerically efficient algorithms to approximate the probabilistic (right) part of Eq. (3) (see for example, Patelli et al., 2017). However, these algorithms can only function if the survival signature can be computed in time, increasing the demand for fast and low-effort survival signature algorithms.

3 Approximation of the survival signature

This section introduces the new proposed method for the survival signature after further outlining the shortcomings of the classical analytical approaches. As evident from Eq. (1) and Eq. (2), the calculation of the survival signature is a combinatorial problem. Thus, in analytical methods, every combination of working and failed nodes of a system must be generated and the state of the corresponding system configuration must be evaluated. At present, the most robust method was developed by Reed (2017) and is based on binary decision diagrams (BDD). However, this technique is only suitable in cases where the BDD can be efficiently computed. As of now, even a system with *only* 61 nodes and 14 component types remains unsolved with the existing algorithm. However, large infrastructure networks often consist of multiple thousands of nodes and can potentially never be solved with analytical approaches. That is why, an algorithm to efficiently *approximate* the survival signature is required to solve problems of engineering interest. The proposed technique can be split into two distinct phases: (1) Apply graph theory to eliminate trivial portions of the survival signature. (2) Approximate the remaining entries by Monte Carlo simulation.

3.1 Phase 1: Graph theory

A system or network is defined as *working* if a path from *start* to *end* exists. Since systems can be made up of multiple start and end nodes, a super start node S^* is set before all incoming nodes and a super end node E^* after all outgoing nodes. These nodes serve as a means to assess the functionality of the analyzed network and are assumed to be always working. In a first step, the minimum number of working nodes required for the system to be in a working state are calculated using *Dijkstra's algorithm* (Dijkstra, 1959) to find the shortest path from S^* to E^*. Thus, all the entries of the survival signature where the sum of working components of any type is smaller than this minimum number of components are set to 0. Next, for all component types, it is checked if a path from S^* to E^* exists using only components of one type. If so, the combinations of the survival signature where all components of said type are functioning are set to 1. This concludes the elimination of trivial combinations through graph theory.

3.2 Phase 2: Monte Carlo simulation

After pre-eliminating significant chucks of the survival signature through the application of graph theory described in the previous phase, the survival signature entries of the remaining combinations are now approximated by Monte Carlo simulation. For each unknown survival signature entry $\Phi_{l_1,\dots,l_k}$ the following steps are repeated until either the maximum number of samples N_{MC} is reached or the coefficient of variation falls under a predefined threshold CoV_{limit}:

Algorithm 1. Approximation of the survival signature

```
Input: N (network), S (start nodes), E (end nodes), CoV_limit (limit of
the CoV), N_MC (maximum number of samples), α (factor to decide between
analytical solution and approximation)
Output: Φ (survival signature), CoV (coefficient of variation)

Phase 1:
1. Prepare empty arrays for Φ and CoV
2. Set S* before S and E* before E
3. Find minimum number of required components using Dijkstra's
   algorithm and set all entries of P where the sum of working
   components is smaller to 0.
4. For all types, check whether a path from S* to E* exists using only
   components of this type → set entries of P where all components of
   type work to 1

Phase 2:
/* loop over all remaining entries of Φ with unknown
Probability */
for each unknown signature entry Φ_{l1,...,lk} do
  CoV_{l1,...,lk} = 0
  n_{l1,...,lk} = 0
  w_{l1,...,lk} = 0
  N_pos = number of all possible combinations;
  if α · N_MC > N_pos then
    Φ_{l1,...,lk} = /* calculate analytically */
    CoV_{l1,...,lk} = 0
  else
    while CoV_{l1,...,lk} > CoV_limit && n_{l1,...,lk} ≤ N_MC do
      Generate a random system state for survival signature state
      vector l1,...,lk
      n_{l1,...,lk} = n_{l1,...,lk} + 1
      if (system is working) do
        w_{l1,...,lk} = w_{l1,...,lk} + 1
      end
      Φ_{l1,...,lk} = Eq. (4)
      CoV_{l1,...,lk} = Eq. (5)
    end
  end
end
```

Generate a random network state corresponding to the state vector $l_1, \dots, l_k$ and increase the number of used samples $n_{l_1,\dots,l_k}$ by 1. Then, if the system is in working condition, increase the number of working samples $w_{l_1,\dots,l_k}$ by 1 followed by updating the current approximation of the survival signature entry as

$$\Phi(l_1, \dots, l_k) \approx {w_{l_1,\dots,l_k}}/{n_{l_1,\dots,l_k}} \tag{4}$$

and the current coefficient of variation as

$$CoV_i = \frac{\sqrt{\left(\Phi_{l_1,\ldots,l_k} - \Phi^2_{l_1,\ldots,l_k}\right)/n_{l,\ldots,l_k}}}{\Phi_{l_1,\ldots,l_k}}. \tag{5}$$

A modification of the depth-first-search algorithm, where visited nodes are removed from the graph, is used to quickly evaluate if a path from S^* to E^* exists for any given network configuration. Note, that if the number of possible combinations of a state vector is smaller than or only slightly exceeds the maximum number of samples to be used, the survival signature entry is calculated analytically instead of approximating it. Algorithm 1 contains the full procedure used to approximate the survival signature.

Even though, this method is able to efficiently and accurately the survival signature, an outstanding issue is the means of storage. The space required to store the full survival signature of a system increases exponentially with the number of components and types, quickly leading to sizes of several gigabytes before even reaching the dimensions of large infrastructure networks such as electricity or communication networks.

4 Numerical example

Consider as an example the network given in Figure 1. This is a simplified representation of the Berlin metro system. 50 of the existing stations were chosen and arbitrarily divided into two subcategories, resulting in 26 nodes of component type one and 24 nodes of component type two. The start and end node used to assess the working condition of the system under different failure scenarios are chosen at will and circled in the image.

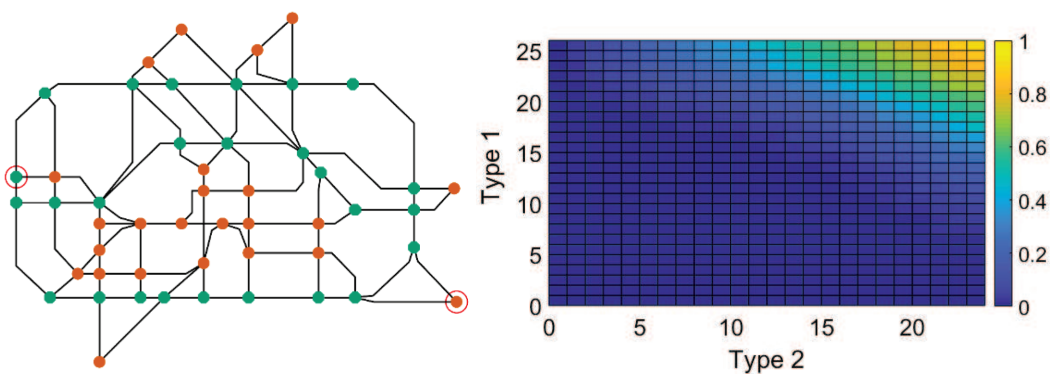

Figure 1. Simplified representation of the Berlin metro system with 50 nodes in two component types (left) and the plot of the corresponding survival signature (right).

In order to compute the exact solution to the survival signature, $2^{50} = 1.1259 \cdot 10^{15}$ combinations need to be evaluated. Using the algorithm introduced in the previous section with a soft limit of 2500 combinations and target coefficient of variation of 0.001, only a fraction of the combinations, around 1.000.000 to be exact, is actually evaluated. This results in a significant reduction of computational effort as less than one percent of all combinations are ever evaluated. However, this gain in efficiency does not come at the cost of accuracy. In fact, the difference of the approximated survival signature to the exact is only ~0.24 %.

5 Conclusion & Outlook

In this work we have presented a novel approach to the approximation of the survival signature in order to eliminate weaknesses of analytical approaches. Using a simple example, we have shown how little the approximation differs from the exact solution while at the same time requiring a significantly smaller number of combinations to be evaluated. This reduction of computational effort was achieved by pre-eliminating trivial parts of the survival signature through the application of graph theory methods and approximating the remaining entries by Monte Carlo simulation. Nonetheless, it was evident from this research, that an outstanding issue with the survival signature is the required storage space. With increasing network size, the storage space required scales exponentially to such a magnitude that the signatures of large infrastructure networks cannot be saved as a whole.

In the future, we must investigate additional advanced graph theory methods to hopefully reduce the amount of samples required for accurate approximation even further. In addition, intelligently splitting networks into sub-systems could potentially help in increasing efficiency.

Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – BE 2570/3-1 and BR 5446/1-1.

References

Aslett, L., ReliabilityTheory: Tools for structural reliability analysis, R package, http://www.louisaslett.com, 2012, Accessed: 24.01.2018

Behrensdorf, J., Broggi, M., Brandt, S. and Beer, M., Numerically efficient reliabilty analysis of interdependent networks, in *The 27th European Safety and Reliability Conference*, Jun, 2017

Coolen, F.P. and Coolen-Maturi, T., Generalizing the signature to systems with multiple types of components, in *Complex systems and dependability*, 115-130, Springer, Berlin, Heidelberg, 2013

Dijkstra, E.W., A note on two problems in connexion with graphs, *Numer Math*, 1(1), 269-271, Dec, 1959

Feng, G., Patelli, E., Beer, M. and Coolen, F.P., Imprecise system reliability and component importance based on survival signature, *Reliab Eng Syst Safe*, 150, 116–125, Jun, 2016

Patelli E, Feng G, Coolen FP and Coolen-Maturi T., Simulation methods for system reliability using the survival signature, *Reliab Eng Syst Safe,* 167, 327-337, Jun, 2017

Reed, S., An efficient algorithm for exact computation of system and survival signatures using binary decision diagrams, *Reliab Eng Syst Safe,* 165, 257-267, Sep, 2017

Samaniego, F.J., *System signatures and their applications in engineering reliability*, Springer Science & Business Media, New York, 2007

Zio E., *The Monte Carlo simulation method for system reliability and risk analysis*., Springer, London, Jan, 2013