NEW CHALLENGES IN NETWORK RELIABILITY ANALYSIS

Andrea Bobbio, Caterina Ferraris, Roberta Terruggia

Dipartimento di Informatica, Università del Piemonte Orientale[1]

Keywords: network reliability, Small scale, minpath, mincut, BDD, large scale, scale free networks

Abstract

Many complex physical, technological, social, biological and economical systems can be represented in the form of networks, where vertices are the entities of the system and the edges represent the relational links among the entities. Since complex networks display a high degree of tolerance to random failures, errors and attacks, network dependability has always been a challenging task in system reliability analysis. In the literature, network reliability is intended as the probability that two specific nodes (a source node and a destination node) are connected given the probability of the elements of the network (nodes, edges or both) of being up or down. A number of techniques have been developed to tackle this problem. However, in recent year, with the appearance of networks of giant dimensions (e.g. the internet and www) the exhaustive searching techniques are no more appropriate. A completely new field of research has emerged to study the statistical properties of these huge networks, together with the study of their robustness to random failures and attacks. This paper is aimed at illustrating the old a new challenges in network dependability evaluation.

1. Introduction

The present paper is aimed at investigating modelling and analysis techniques for the evaluation of the reliability of systems whose structure can be described in a form of a graph. For such systems we use the generic term of *networks*, where networks are characterized by a set of nodes (vertices) connected by directed or undirected arcs (edges). Systems in the form of networks are present in nature, in biological systems and in many technological fields. Their study has proved to be beneficial also in the analysis of social, economical, epidemic and political relations. With the growing dependence of our modern society on technological systems and information networks, there is an increasing demand for high dependability. The degree to which a system is able to provide the required operation needs to be quantitatively assessed by defining proper measurable quantities. The quantitative assessment of system dependability becomes essential in system design, planning, implementation, validation, manufacturing and field operation.

Traditional reliability analysis techniques in the area of networked systems used to look at networks of tens or hundreds of vertices [KAUF77, ABRA79, PAGE88, BALA03]. These techniques were based on exhaustive search algorithms intended to provide qualitative and quantitative information on the network connectivity, dependability, and vulnerability. A literature survey indicates that the approaches, which have been used to compute two-terminal reliability could broadly be classified into two paradigms: *i)* - the paradigm in which desired network reliability is directly calculated (series-parallel reduction [BOBB82a], pivotal decomposition using keystone components [PAGE88, HARD05]) and *ii)* - the paradigm in which all possibilities through which the two specified nodes can communicate (or not

[1] Via Bellini 25/G, 15100 Alessandria, Italy – bobbio@mfn.unipmn.it

communicate) with each other are first enumerated (path/cut set search [BALA03, LUOT98]) and then reliability (unreliability) expression is evaluated. In this last case, the knowledge of all the minpaths (mincuts) provides qualitative information about the connectivity properties of the network. If furthermore, the probability of failure of nodes and edges is known, the reliability expression can be calculated using different techniques, like inclusion-exclusion method, sum of disjoint products [ABRA79], and, more recently, Binary Decision Diagrams (BDD) [BRYA86]. BDDs provide an extraordinarily efficient method to represent complex binary structures by means of the Shannon's decomposition principle [BURC92]. Algorithms exploiting the direct use of BDDs to model the network connectivity, and the level of reachable complexity, needs to be investigated more deeply.

However, the complexity of real world today networks (the internet, the www, the public telecommunication networks), can reach millions or even billions of vertices. This change of scale forces a corresponding change in the analytic approach. Many of the approaches that have been applied in small or medium scale networks, and many of the questions that have been answered are simply not feasible in much larger networks. Recent years have witnessed a substantial new movement in network research [ALBE02a, BARA02a, DORO02a, NEWM03a], with the focus shifting away from the analysis of small graphs to consideration of large-scale statistical properties of graphs and with the aim of predicting what the behavior of complex networked systems will be on the basis of measured structural properties and the local rules governing individual vertices.

The shift, experienced in the past few years in the understanding of complex networks, was rapid and unexpected. Empirical studies, models and analytic approaches have enlighten that real networks display generic organizing principles shared by rather different systems. These advances have created a prolific branch of statistical mechanics, followed with equal interest by sociologists, biologists and computer scientists. Moreover, the structural organization of a complex network influences how the system reacts to occasional failures or to intentional attacks, and hence has a direct impact on the dependability and security of these structures. The paper will explore exhaustive analysis algorithms for small scale networks (from tens to hundreds of nodes) by enlightening the developed analytical techniques and the related software tools. For massive large-scale networks, the paper will mainly provide some pointers to the enormous amount of research material published in the very last few years.

2. Small Scale Networks: Exhaustive Analysis

We consider exhaustive analysis tools based on searching algorithms through the net. They can be distinguished in *qualitative* and *quantitative* algorithms. Qualitative algorithms provide information on the level of connectivity of two nodes and on the level of vulnerability. If failure probability value can be assigned to edges or nodes or both, we compute the reliability from the connectivity and we compute the unreliability from the vulnerability. The network is represented in the form of a graph $G=(V,E)$, where V is the set of vertices (or nodes) and E the set of edges (or arcs). We assume, in general, that edges can be undirected or directed, and that arcs or nodes or both can be failure prone. In order to study the reliability of a network we identify a single node as a source node O and a single node as a sink node Z. The analysis may be either qualitative or quantitative.

2.1 - Path sets and Cut sets

Definition 1 - For a given graph $G=(V,E)$, a *path* is a subset of components, arcs and/or nodes, that guarantees the source O and sink Z to be connected if all the components of this subset are functioning. A path H is a *minpath* if a subset of elements in H does not exist that is also a path.

Definition 2 - For a given graph $G=(V,E)$, a *cut* is a subset of components, arcs and/or nodes, whose failure disconnects the source O and sink Z. A cut K is a *mincut* if a subset of elements in K does not exist that is also a cut.

Property 1 - Any path contains at least one component from each cut, and any cut contains at least one component from each path [KAUF77].

Let h be the number of minpaths between a source O and a sink Z in a network $G=(V,E)$ and let $H_1, H_2, \ldots, H_h$ be the h minpaths. Since all the elements of a minpath must be working for the minpath to be connected, the elements of the minpath are logically connected in *AND*. However, since any one of the minpaths is sufficient to make the network working, the network connectivity S can be represented as the logical *OR* of its minpaths:

$$S = H_1 \cap H_2 \cap \ldots H_h$$

According to the above definition, the point-to-point reliability is:

$$R_S = Pr\{S\} = Pr\{ H_1 \cap H_2 \cap \ldots H_h \} \quad (1)$$

The probability of the union of non-disjoint events, as in Formula (1), can be computed by different techniques:

- The inclusion-exclusion expansion algorithm [TRIV01a];
- Prepocessing minpaths or mincuts for sum of disjoint products [ABRA79, LUOT98, HEDT02];
- Binary Decision Diagram (BDD) representation [BRYA86]: BDDs represent Boolean functions as directed acyclic graphs.

Example

A bridge network with directed arcs has the configuration presented in Figure 1. We consider that only arcs can fail. Minpaths are searched by means of the Dijkstra algorithm, while mincuts are searched by means of the YTL algorithm [YAN94].

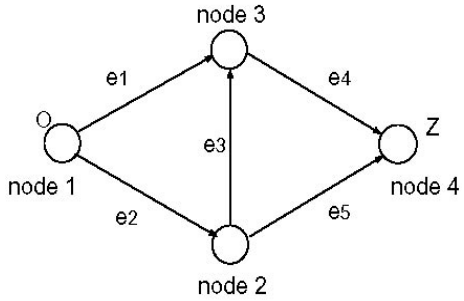

Figure 1 – A directed bridge network

The network possesses three minpaths and four mincuts. They are, respectively:

$$H_1 = \{ e_1, e_4 \} \quad ; \quad H_2 = \{ e_2, e_3, e_4 \} \quad ; \quad H_3 = \{ e_2, e_5 \}$$

$$K_1 = \{ e_1, e_2 \} \quad ; \quad K_2 = \{ e_2, e_4 \} \quad ; \quad K_3 = \{ e_4, e_5 \} \quad ; \quad K_4 = \{ e_1, e_3, e_5 \}$$

From the above formulas, the network connectivity can be expressed as:

$$S_{1\text{-}4} = e_1 e_4 \cap e_2 e_3 e_4 \cap e_2 e_5 \quad (2)$$

We assume that the elements of the graph have only two states *up* and down, and that the element failures are statistically independent. By denoting with pi the probability of the arc e_i of being working, by applying the classical inclusion-exclusion formula for calculating the probability we get from (2):

$$R_{1\text{-}4} = Pr\{S_{1\text{-}4}\} = p1\,p4 + p2\,p3\,p4 + p2\,p5 - p1\,p2\,p3\,p4 - p2\,p3\,p4\,p5 - p1\,p2\,p4\,p5 + p1\,p2\,p3\,p4\,p5 \quad (3)$$

2.2 - BDD Representation

A BDD is a way to represent a Boolean expression by means of the Shannon's decomposition principle: if F is a Boolean function on variables $x_1, x_2, \ldots, x_n$ the following decomposition holds:

$$F = x_1 \; F_{\{x1=1\}} \cap \neg x_1 \; F_{\{x1=0\}} \tag{4}$$

Where, ($\neg x_1$= *NOT* x_1); $F_{\{x1=1\}}$ is derived from F assuming x_1 is true, and $F_{\{x1=0\}}$ is derived from F assuming x_1 is false.

The functions $F_{\{x1=1\}}$ and $F_{\{x1=0\}}$ depend now on *(n-1)* variables; applying to them the Shannon decomposition (4), pivoting with respect to one of the remaining *(n-1)* variables, each one of the two functions $F_{\{x1=1\}}$ and $F_{\{x1=0\}}$ can be decomposed in two functions depending on *(n-2)* variables. Applying iteratively the Shannon's decomposition formula pivoting with respect to a complete sequence of all the variables, a complete decomposition can be obtained. The sequence of decompositions can be represented in graphical form using a binary tree. Each node of the tree represents the pivot variable with respect to which the decomposition is done. From each node spawn two branches: the left branch has the value *true (or 1)* and its descendent is the first term $F_{\{x1=1\}}$ of the decomposed function obtained by fixing the value of the pivot variable to 1, the right branch has the value *false (or 0)* and its descendent $F_{\{x1=0\}}$ is the second term of the decomposed function obtained by fixing the value of the pivot variable to 0.

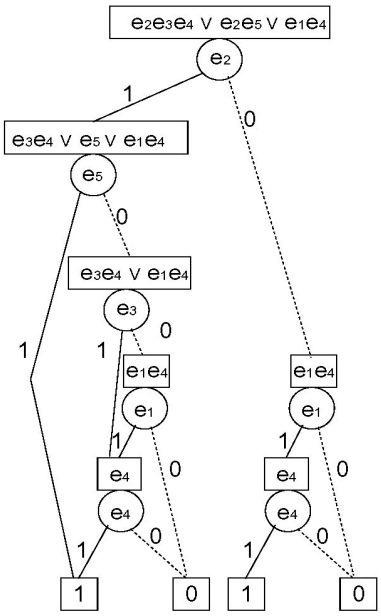

Figure 2- Shannon's decomposition of the connectivity of the bridge network (1)

We apply the Shannon's decomposition to the Boolean connectivity function of the directed bridge network expressed as the union of the minpaths in Formula (2). The ordering of the variables strongly affects the decomposition. With the ordering:

$$e_2 \; e_5 \; e_3 \; e_1 \; e_4$$

the Boolean connectivity expression (2) is built step by step in Figure 2. Occasionally, the binary tree contains identical subtrees. To make the representation more compact, identical subtrees may be folded. The compact representation in which duplicated subtrees are folded and the terminal nodes 1 and the terminal nodes 0 are merged is called the Binary Decision Diagram (BDD). In Figure (2), the subtrees generated at the node e_1 e_4 appear twice and can be folded to produce the BDD of Figure 3.

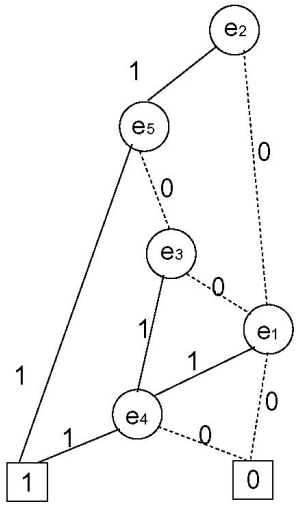

Figure 3- BDD representation of the bridge network (1)

The computation of the probability of the BDD of Figure 3, can again proceed recursively by resorting to the Shannon decomposition.

$$Pr\{F\} = p1\ \ Pr\{F_{\{x1=1\}}\} + (1 - p1)\ Pr\{F_{\{x1=0\}}\} \qquad (5)$$
$$= Pr\{F_{\{x1=0\}}\} + p1\ (Pr\{F_{\{x1=1\}}\} - Pr\{F_{\{x1=0\}}\})$$

where p_1 is the probability of the Boolean variable x_1 to be true and $(1-p_1)$ is the probability of the Bolean variable x_1 to be false. Recursive application of Equation (4) is pictorially shown in Figure 4. Computation of the probability values at the intermediate nodes (pr_1, pr_3, pr_5) provides the value of the network reliability $R_{1\text{-}4} = pr_2$ as in Formula (3).

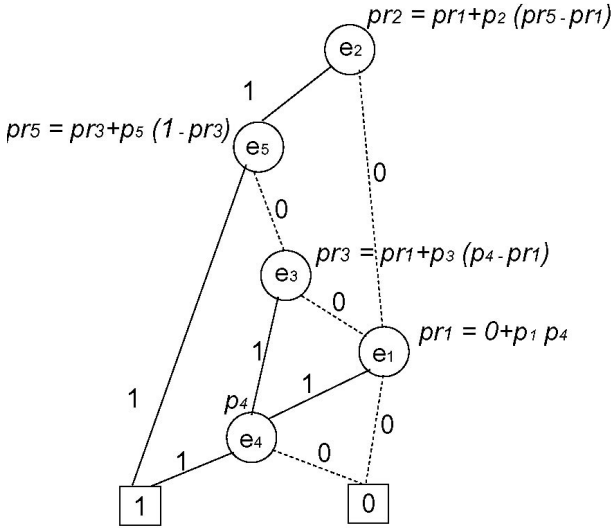

Figure 4 - Probability computation from the BDD of Figure 3

2.3 – Direct algorithms

Direct algorithms are intended to derive the connectivity logical function and the reliability of a network from the structure of the graph, without passing from a preliminary search for the minpaths or mincuts. Several techniques have been envisaged for this purpose. An algorithm based on the construction and saving of the pathsets in disjoint binary form is presented in [GOYN05]. The logical expression of the connectivity is directly derived from the graph as a sum of disjoint logical terms (SDP), so that the reliability expression can be calculated as sum of products. Furthermore, the paper in [GOYN05] proposes specific operators to modify the logical expression of the connectivity by adding (or removing) arcs in

the network. From this point of view, the paper is oriented to follow a step by step growth of a network. According to [GOYN05], the SPD expressions for the connectivity and for the reliability of nodes *1-4* in Figure 1 (denoting $q_i=1-p_i$) become:

$$S_{1\text{-}4} = e_1 e_4 \cap e_2 e_5 \neg e_1 \cap e_2 e_5 e_1 \neg e_4 \cap e_2 e_3 e_2 (\neg e_1 \neg e_5) \quad (6)$$

$$R_{1\text{-}4} = p1\ p4 + p2\ p5\ q1 + p2\ p5\ p1\ q4 + p2\ p3\ p4\ (q1\ q5) \quad (7)$$

Two algorithms have been recently proposed in the literature to derive directly the connectivity function in the form of a BDD, without deriving the corresponding logical expression. The algorithm in [SEKI95], revisited in [HARD05] is based on a suitable factorization algorithm, consisting in pivoting along a complete sequence of arcs and deriving the right subgraph when the arc is supposed working (the arc is contracted or short-circuited) and the left subgraph the arc is supposed not working (the arc is deleted or opened). Using as a scalable benchmark a lattice structure with the source *O* and the destination *Z* located on the external vertices of the main diagonal, both papers indicate a lattice of 12×12 (144 nodes and 264 edges) as an upper limit to computational power of the method.

The algorithm suggested in [ZANG00] generates the BDD directly, without explicitly forming the Boolean expression, via a recursive depth first search on the graph. An outline of the proposed algorithm is reported in Figure 5.

```
bdd_gen(start_node){
  T_bdd = 0
  set start_node in this_path
  for (edge_i in the set of edges starting from start_node){
    next_node = the other end of edge_i
    if(next_node == sink_node)
    subpath_bdd = edge_i_bdd
    else if (next_node is already in this_path)
      continue;
    else
      subpath_bdd = bdd_gen(next_node) AND edge_i_bdd
    T_bdd=T_bdd OR subpath_bdd
  }
  clear start_node in this_path
  return T_bdd
}
```

Figure 5 - Generation of the BDD via a recursive depth first search on the graph.

We have implemented the algorithm of Fig. 5 by resorting to the BDD library developed at the Carnagie Mellon University (http://www.cs.cmu.edu/~modelcheck/bdd.html). We have tested the algorithm using the scalable (directed) lattice graph proposed in [SEKI95, HARD05]. Table I displays the dimension of the lattice, the number of nodes and of edges of the graph, and the number of nodes of the BDD resulting from the application of the algorithm of Figure 5. Finally, from the BDD the list of minpaths is obtained by applying a transformation of the BDD according to a technique proposed in [RAUZ93].

TABLE I – Benchmark on Directed Lattice Graph

Lattice	# Lattice nodes	# Lattice arcs	# BDD nodes	# minpath
2 X 2	4	4	6	2
4 X 4	16	24	94	20
6 X 6	36	60	1034	252
8 X 8	64	112	8384	3432
10 X 10	100	180	56338	N.A.
12 X 12	144	264	342038	N.A.

The complexity of the resulting BDD, expressed as the number of the BDD nodes, versus the lattice dimension, is plotted in Figure 6. An exponential increasing complexity is evident from the plot, as theoretically argued in [SEKI95].

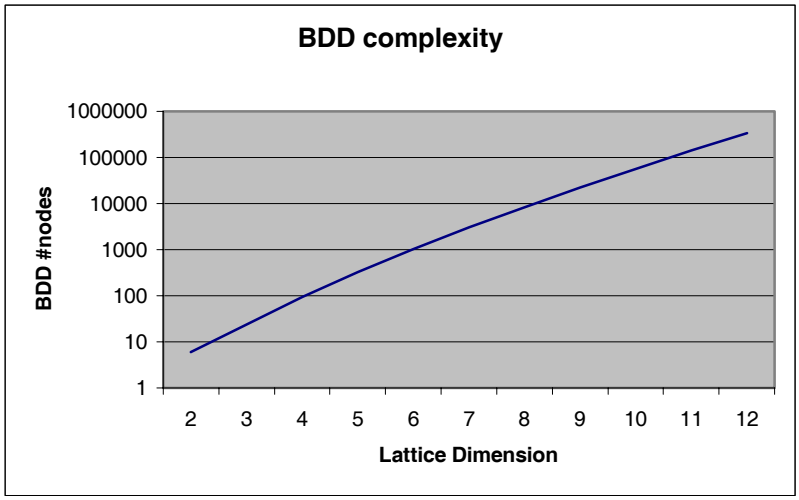

Figure 6 - BDD complexity of the connectivity function of a lattice graph.

Further comparison among different algorithms and experimentation of new data structures is needed. However, the search through a graph is know to be a NP-complete problem and the exhaustive methods grow in complexity very rapidly as a function of the network dimension (number of nodes and edges). Hence, it is hard to imagine how these methods can overcome the barrier of the order of hundreds of nodes. For large and very large networks new techniques are needed and new ideas are emerging.

3. Large Scale Networks

Massive large-scale networks have been the object of an enormous amount of research material published in the very last few years. The following lines of investigation are of particular interest in the dependability arena:

- Structural properties: A new mathematical foundation for complex interactive networks;
- Error and attack tolerance;
- Modeling and diagnosis methods for large-scale complex networks;
- Innovative techniques for defense against catastrophic random failures and preferential attacks.

Real word today networks can reach millions or even billions of vertices. The Internet and World Wide Web are perhaps the most known network systems that have influenced our every day life. But several examples belonging to a great class of large scale evolving networks can be mentioned: e.g. communications nets, collaboration networks, public relations nets, citations of scientific papers, some industrial networks, transportation

networks, many biological networks, etc.. Quite recently, it has become clear that their complex structure is a natural consequence of the principles of their evolution and of their growth. Some simple basic ideas have been proposed for studying the self-organization of growing networks and processes occurring within them. Some basic ideas, used later by physicists, were proposed long ago by the outstanding Hungarian mathematician Paul Erdős and his collaborator Rényi. Most of the results of their graph theory are related to random graphs. Random graphs exhibit a Poisson distribution of connections, so that the tail of the distribution decays exponentially. The study of large modern networks has assessed two main relevant factors:

- Large networks evolve dynamically.
- Interest to study how growing networks self organize into special (scale free) structures and the role of the mechanism of preferential linking.

The combination of the dynamical evolution with growing aggregation principles has led to the discovering that the network connectivity does not exhibit a Poisson distribution, but a power-law long-tailed distribution. The power-law decay in the connectivity distribution entails the presence of few nodes with an extremely large degree of connectivity dispersed into a net whose node have a relatively low degree of connectivity. The highly connected nodes, called hubs, play a special role in the properties of the network. After these findings, a number of researchers in different areas have started intensive study of evolving networks in various areas, and some extensive review articles have recently appeared [ALBE02a, BARA02a, DORO02a, NEWM03a].

3.1 – Dependability and Security

Many complex networks show a very high degree of tolerance to malfunction or errors, in the sense that a malfunction of an edge or of a node perturbs the behaviour of the system only locally, but has very little effect on the overall properties of the system. Hence, it became natural to extend the research interests to encompass studies on the behaviour of huge networks in the presence of failures or errors. Since exact point-to-point probability calculations are not possible due to the dimension of the system, the studies where directed to evaluating how overall statistical properties of the network are affected by local damage. A failure of an edge or a node is the loss of operation of the element and can be emulated by removing that element from the graph. Studies in this direction were addressed on topological aspects of robustness caused by edges or node removal [ALBE02a]. Edge removal affects only the link between the two connecting nodes and all the paths passing through that link, while node removal implies the disconnection of all the edges ingoing or outgoing from that node and inflicts more damage than edge removal.

Research was addressed to explore the network connectivity property as a function of the fraction of edges (or node) removal and the correlation between robustness and network topology. Calling f the fraction of edges removed from the structure, the studies were focussed on finding the existence of a critical fraction fc for which the resulting network is no more connected (were connectivity is defined and measured in a statistical sense). Two main cases of possible malfunction are considered:

- The element removal (either edges, nodes or both) occurs randomly; i.e. any element has the same probability of being removed. In this case we talk about random failure;
- The element removal (either edges, nodes or both) is preferential; i.e. there is an intentional attack targeted at nodes with the potential highest effect on the network connectivity.

The first case models physical failures, and an appropriate metric is the dependability, while the second one models malicious attacks and is more pertinent to the field of security. The structure of the network reacts differently to the two different types of failure. The main results are surveyed in [ALBE02a]. For random graphs a critical removal frequency fc can be identified such, that when $f>fc$, the network breaks into small isolated clusters. For scale free networks, in the limit of the network dimension going to infinity, a critical frequency cannot be identified since the networks remains connected through the hubs as the removal fraction tends to 1. This result can be summarized by saying that scale free networks display

an exceptional robustness against random failure and random node removal, and are more resilient than random graph structures. For intentional malicious attacks, the situation is reversed. Scale free networks are more fragile to maliciously preferential attacks since the damage of the highly connected nodes has a dramatic disruptive effect on the network. The presence of hubs makes the structure more vulnerable to intentional damage. A similar effect has been observed for the spreading of diseases in epidemic studies or the diffusion of viruses in information networks. For random networks a disease spreads in the whole network only if the infection rate is larger that a critical value, while in scale free networks the whole network is infected at any spreading rate [ALBE02a]. Thus again scale free networks are more vulnerable than random graphs.

In these aforementioned studies, element failures are considered as statistical independent and their presence does not modify the failure properties of the other, even neighbouring, elements. Often this is not the case, and a single point failure can produce a cascade effect on adjacent nodes and edges. The investigation of `cascade failures, their impact on the net properties and the relation between robustness and net structure are important fields of investigation.

4. Conclusion and Future Work.

We have divided our presentation in two main sections: small scale networks and large scale networks, and we have illustrated the most promising analysing techniques and results in the two cases. Of course the challenge is to make the algorithms implemented for small scale networks to cope with larger dimensions, by exploiting new and more compact data structures or introducing approximations (partition of the net in subnets and clusters, search for the paths of minimal length only), and find the way of evaluating point-to-point measures in huge systems. A new emerging challenging field is the study of the performance and dependability properties among different interacting networks, like the interaction of the power distribution network with the control information system, the telecommunication with the transportation systems.

Acknowledgements
This paper has been partially supported by the Italian Ministry of Education (MIUR) under the project FIRB-Perf

References

[ABRA79] - J.A. Abraham. An improved algorithm for network reliability. *IEEE Transaction on Reliability*, 28:58--61, 1979.

[ALBE02a] - R. Albert and A.L. Barabasi. Statistical mechanics of complex networks. *Review Modern Physics*, 74:47--97, 2002.

[BALA03] – A.O. Balan and L. Traldi. Preprocessing minpaths for sum of disjoint products. *IEEE Transaction on Reliability*, 52(3):289--295, September, 2003.

[BARA02a] - A.L. Barabasi. Linked: the new science of networks. 2002.

[BOBB82a] - A. Bobbio and A. Premoli, Fast algorithm for unavailability and sensitivity analysis of series-parallel systems, *IEEE Transactions on Reliability*, R-31, 359-361, 1982

[BRYA86] - R.E. Bryant.Graph-based algorithms for Boolean function manipulation. *IEEE Transactions on Computers*, C-35:677--691, 1986.

[BURC92] - Burch, J.R., Clarke, E.M., McMillan, K.L., Dill, D.L., & Hwang, J. Symbolic Model Checking: 10^20 States and Beyond , *Information and Computation* (Special issue for the best papers from LICS'90), Vol. 98, pp. 142-170, June 1992.

[DORO02a] - S.N. Dorogovtsev and J.F.F. Mendes. Evolution of networks. *Advances in Physics*, 51:1079--1187, 2002.

[GOYN05] - N.K. Goyal, Network reliability evaluation: a new modeling approach, *Int Conference on Reliability and Safety Engineering (INCRESE2005)*, 473-488, 2005

[HARD05] - G. Hardy and C. Lucet and N. Limnios, Computing all-terminal reliability of stochastic networks by Binary Decision Diagrams, *Proceedings Applied Stochastic Modeling and Data Analysis, ASMDA2005*, 2005

[LUOT98], T. Luo and K.S. Trivedi, An improved algorithm for coherent-system reliability, *IEEE Transactions on Reliability*, 47, 73-78, 1998

[KAUF77] - A. Kaufmann, D. Grouchko, and R. Cruon. Mathematical Models for the Study of the Reliability of Systems. Academic Press, 1977.

[NEWM03a] - M.E. Newman. The structure and function of complex networks. *SIAM Review*, 45:167--256, 2003.

[PAGE88] - L.B. Page and J.E. Perry, A practical implementation of the factoring theorem for network reliability", *IEEE Transactions on Reliability*, R-37, 259-267, 1988

[RAUZ93] - A. Rauzy, New algorithms for fault tree analysis, *Reliability Engineering and System Safety*, 40, 203-211, 1993

[SEKI95] - K. Sekine, H. Imai, A unified approach via BDD to the network reliability and path number, Department of Information Science, University of Tokyo, TR-95-09, 1995

[TRIV01a] - K. Trivedi. Probability & Statistics with Reliability, Queueing & Computer Science applications, Wiley, II Edition, 2001.

[YAN94] – Li Yan, H. A.Taha, T L. Landers. A recursive approach for enumerating minimal cutset in a network. *IEEE Transaction on Reliability*, 43(3):383--387, 1994.

[ZANG00] - X. Zang, H. Sun, K. Trivedi, A BDD-based algorithm for reliability graph analysis, Department of Electrical Engineering, Duke University, 2000