International Journal of Performability Engineering Vol. 5, No. 1, January 2009, pp. 71-84.

Printed in India

Application of Advanced Computational Techniques to the Vulnerability Assessment of Network Systems exposed to Uncertain Harmful Events

CLAUDIO M. ROCCO S.[1*], DANIEL E. SALAZAR A.[2] and ENRICO ZIO[3]

[1]*Universidad Central de Venezuela, Caracas, Venezuela*
[2] *Universidad de Las Palmas de Gran Canaria, Spain*
[3] *Politecnico di Milano, Italy*

(Received on October 8, 2007)

Abstract: This paper presents the application of advanced computational techniques developed by the authors for evaluating the vulnerability characteristics of network systems exposed to harmful events. The physical system is modeled as a network (graph) of nodes interconnected by links. Uncertainties on the propagation and effects of an attack are modeled by probability distributions on the times of propagation through the network links and the numbers of people affected at the network nodes reached by the hazard. The impact of an attack is quantified by simulating the propagation of the hazard through the network nodes and links, by means of a combination of cellular automata and Monte Carlo simulation. The vulnerability assessment is embedded within a systematic multiple-objective optimization analysis aimed at identifying the optimal protective scheme which minimizes the average impact in terms of entities affected and hazard propagation time. The vulnerabilities and relative protection schemes of two networks of realistic size are systematically analyzed by the proposed approach for testing the procedure and identifying its strengths and weaknesses.

Keywords: *Distributed networks, vulnerability assessment, Monte Carlo simulation,*

1. Introduction

The reliance of current society on critical infrastructures makes them particularly sensitive to partial or complete incapacitation of such infrastructures, due to internal or external sources of failures or attacks. While reliability engineering and risk analysis provide tools and procedures for estimating, preventing and handling undesired failure events that occur at random in complex systems, risks from intentional attacks constitute a new challenge due to the involvement of "a malevolent intelligence directed towards maximum social disruption" (Apostolakis and Lemon 2005, p. 361).

Many efforts have been devoted in recent years towards the development of a new paradigm for analyzing the safety and security of critical infrastructures so as to be able to set up the adequate protections against natural disasters and/or intentional attacks (Bier *et al* 2005, Hausken 2007, Korczak *et al* 2005, Levitin 2007, Levitin and H. Ben-Haim 2007, Levitin and K. Hausken 2007, Haimes and Longstaff, 2002).

Since most critical infrastructures present a distributed network configuration (see Birchmeier 2007 and references therein for examples), much research has focused on the

* Corresponding author's email: croccoucv@gmail.com

protection of complex networks against terrorist or intentional attacks to, *e.g.*, water supply networks (Skolicki *et al* 2006, Wadda *et al* 2004) or electricity networks (Holgrem *et al* 2007, Holgrem 2006, Johnson 2007).

With regards to intentional attacks, network infrastructures are susceptible to at least two modes of attack. On the one hand, an attack may be directed to damaging the infrastructure itself by impacting its components. On the other hand, an antagonist could take advantage of the infrastructure as a vector of propagation of a hazard to the people and the environment (*e.g.*, a contaminant or a virus injected into a water supply network).

In this latter case, hazard propagation modeling becomes a quite relevant task for providing the necessary information to devise effective countermeasures to the attacks. With respect to the example of an attack to a water distribution system by contaminating the water supply, real-time decisions must be undertaken by the responsible security officials to take actions for minimizing the impacts of such attack, *e.g.,* by shutting off selected distribution lines at specific times (Wadda *et al*. 2004).

For hazard propagation modeling, the use of simulation techniques allows the identification of the critical vulnerabilities of the network system, *i.e.,* those nodes where an "attack" can cause the worst damage. This in turn provides information for the optimal allocation of protective countermeasures. Given the uncertainty on the effectiveness of such countermeasures and on the antagonist's actions, hazard propagation models must handle uncertainty.

To protect from attacks effectively, several "immunization" schemata with different characteristics could be proposed, aimed at minimizing the impact of an attack (see for example Levitin and Ben-Haim, 2007); however, realistically one should consider several objectives in the search for the optimal protection scheme, *e.g.,* including cost, vulnerability, reliability.

In this paper, the multiple objective (MO) formulation of the problem of system security protection from hazard propagation is embraced (Rocco *et al* 2007, Zio *et al* 2007). Its solution is sought through an optimization approach which leads to finding a set of alternative protection schemes which are optimal in the sense of Pareto optimality with respect to the set of predefined objectives driving the solution search. Based on the results obtained, the Decision Maker (DM) can rationally decide on a robust protective scheme to defend those nodes identified as most critical, *i.e.,* whose protection most reduces the damage.

The physical system of interest is abstractly modeled as a network (graph) of nodes interconnected by links. Uncertainties on the propagation and effects of an attack are modeled by probability distributions on the times of propagation through the network links and the numbers of people affected at the network nodes reached by the hazard.

The MO problem of protecting the system from the hazard propagation is solved by a Multiple Objective Evolutionary Algorithm (MOEA) (Zio *et al* 2007) which conjugates the concepts of Pareto dominance with the typical heuristic search mechanisms of evolutionary algorithms. To this aim, the vulnerability assessment is embedded within a systematic multiple-objective optimization analysis aimed at identifying the optimal protective scheme which minimizes the average impact in terms of entities affected and hazard propagation time. For each alternative solution of network system protection proposed by the search algorithm, the impact of the attack and timing of the hazard propagation are quantified by simulating the propagation of the hazard through the network nodes and links, by means of a combination of cellular automata (CA) and Monte Carlo simulation which properly account for the uncertainties involved.

The remainder of the paper is organized as follows. In Section 2, the decision-making problem regarding the network system security protection from hazard propagation is introduced and the computational model of hazard propagation, based on cellular automata and Monte Carlo simulation, is illustrated in details. Section 3 sets the problem in its multiple-objective formulation and provides a short description of the MOEA heuristic optimization tool employed for its solution. Section 4 goes through the systematic application of the procedure to two networks of realistic size. This allows drawing some insights on the capabilities offered by the proposed scheme of analysis and the relative limitations, which are discussed in the closing Section 5.

2 The Hazard Propagation Problem

2.1 The Modelling Framework

A generic network $G(N;E)$ is composed of a set $N = \{n_i\}$ of n nodes linked by a set of edges $E = \{e_{ij}\}$, each of which connects two generic nodes n_i with n_j in a directed or undirected manner (Shier 1991). This abstraction can be applied to model the topology of numerous types of interconnected systems. In particular the model proposed in (Zio and Rocco 2008) associates nodes to sets of entities (beings or assets) that can be simultaneously damaged by an attack and can propagate it, and edges to propagation channels.

From the viewpoint of the modeling of the process of hazard propagation following an attack, several features may be considered, like the edges' transmission capacity, the intensity of the hazard propagated or the existence of different modes of attack, among others. Other realistic aspects to be considered would be, for example, constraints on the capacity of links and nodes (which arise for example in electric (Bier *et al*. 2006) or water distribution systems (Wadda *et al*. 2004)) or considering that when a disturbance occurs, the network starts to shed loads.

In the present work, the modeling is limited to capturing those aspects of the hazard propagation process which need to be accounted for, at a minimum, when analyzing network protection schemes. These relate to the number of network nodes which can be attacked by the antagonist, the number of entities at each node which can be potentially damaged by the hazard propagation and the time of propagation of the attack's hazardous effects through the network links. The modeling of only these generic aspects of the hazard propagation dynamics through the network topology allows concentrating the work on the optimization of the network protective measures for security.

According to the modeling viewpoint adopted, as soon as an attack takes place its harmful effects begin to propagate through the network, from node to node with the consequent impact on the entities associated to such node. When the generic node n_i is hit by the hazard, it propagates the attack to an adjacent node n_j through link e_{ij}, with a time delay TD_{ij}. As we shall see, in this work time delays are assumed, for simplicity but without loss of generality, to be integer random variables of known distributions; the time evolution of hazard propagation can then be evaluated by a combination of cellular automata and Monte Carlo simulation (Zio and Rocco, 2008).

From a defender point of view, the possible strategies against such hazard propagation are to prevent the antagonist from performing the attack or to implement a set of countermeasures to neutralize or mitigate the impact of the attack once it is performed. The decision-making problem considered in this work corresponds to this latter situation: given an attack, the defender aims at minimizing the impact on the network, subject to his

or her amount of available resources R_D; on the other hand, the antagonist is assumed to be rational so that his or her selection of targets is aimed at maximizing the impact subject to the amount of available resources R_A. Hence, the defender problem can be formulated as the identification of protections to be allocated on the network for minimizing the impact of an attack, subject to the amount of available resources R_D.

Naturally, the pattern of attack remains uncertain, even when some information about the preferences of the attacker and their resources R_A can be estimated through intelligence gathering.

The key issue for the optimization of the security protective measures is the definition of the impact of an attack to the network. For instance, in (Korkzak *et al* 2005, Levitin 2007, Levitin and Ben-Haim 2007, Levitin and Hausken 2007) the impact is defined in terms of a utility function that models the expected damage and the objective of the optimization of the protective measures is to minimize such quantity. Various quantities can be defined to describe the effects of the hazard propagation. When embedded in the evaluation of the protective measures, the underlying idea shared by the different definitions of impact is that the antagonist would want to maximize the impact, whereas the defender aims at minimizing it.

In this work, two quantities are used to define the impact of an attack to the network system depending on the hazard intensity and speed of propagation as well as on the distribution of persons and/or entities on the network nodes (Zio and Rocco 2008):

Time To Reach All network Destination nodes (TTRAD): it is the time that it takes for the hazard to propagate to all nodes of the network. From the point of view of a security protection scheme, short times of propagation are to be avoided. This measure bears some similarities with the 'all-terminal network reliability', often used in network reliability analysis.

Average Number of Affected Persons (ANAP) or Average Number of Affected Entities (ANAE): it is the average number of people or entities which are affected by the propagation of the hazard through the network. From the point of view of the network protection, the goal of the decision maker responsible for the safety investment is to minimize the number of affected persons and/or entities. In the face of the uncertainty in the actual consequences of an attack, due to uncertainties in the hazard propagation timing and mechanisms, the average impact, in terms of numbers of persons and/or entities affected by the attack, is taken as the representative value. If most of the persons and/or entities which are potential targets of the attack are gathered near the node where the attack begins, the number of persons and/or entities affected is likely to be large, in which case the ANAP and/or ANAE values would be large. Conversely, if most persons and/or entities are distributed far from the point of attack, the ANAP or ANAE would take small values.

The ANAP (ANAE) measure fully captures the negative consequences from the impact of an attack, albeit on average because of the uncertainties associated to the hazard propagation and effects. Yet, when analyzing alternative network protection strategies, the available time for emergency action may become relevant. Hence, the need to consider also a measure of the propagation timing, *e.g.,* the TTRAD. For example, from the point of view of the protection or mitigation strategy effectiveness, a scenario of attack that affects on average 1000 persons (as captured by the ANAP measure) is considered worse if the time that it takes for them to be affected by the hazard is short (as captured by the

TTRAD measure) since the time to curtail the attack or mitigate its effects would be short as well.

2.2 Hazard Propagation Modeling by CA

To model the hazard propagation through the network, under the limiting assumptions defined in the previous Section, a Cellular Automata approach has been undertaken. CA are mathematical models of dynamic systems. The dynamics of CA unfolds at discrete time steps on a discrete lattice of cells L, typically assumed homogeneous (all cells bear the same properties) (Wolfram, 1985). For example, in a three-dimensional cellular state space the state at the discrete time t of the generic cell ijl, of co-ordinates x_i, y_j, z_l with $i,j,l \in Z$, is described by the state variable $s_{ijl}(t)$. Each cell of L is a *finite automaton* which can assume one of a finite number of discrete values in a *local value space* S≡{0,1,2,…,k-1}.

The generic cell ijl interacts only with a fixed number n of cells that belong to its predefined local neighborhood N_{ijl}. At the next discrete time $t+1$, the cell ijl updates its state $s_{ijl}(t+1)$ according to a transition rule $\phi : S^n \rightarrow S$, which is a function of the state variables at time t of the n cells in N_{ijl}, *viz.*, $s_{ijl}(t+1) = \phi[s_{rsp}(t), \quad rsp \in N_{ijl}]$. Notice that the homogeneity assumption implies that the functional form of the rule is assumed to be the same everywhere in the cellular state space, *i.e.*, there is no space index attached to ϕ. Differences between what is happening at different locations are due only to differences in the values of the state variables of the local neighborhood, not to the update rule. The rule is also homogeneous in time. One "iteration step" of the dynamical evolution of the CA is achieved after the simultaneous application of the rule ϕ to each cell in the lattice L. The temporal evolution of this CA is obtained by: a) specifying the finite size of the lattice L; b) specifying the boundary conditions; c) specifying the initial condition $\vec{s}(0) = [s_1(0), s_2(0), ..., s_M(0)]$ and d) simultaneously applying the rule ϕ to each of the L lattice cells, in an iterative manner.

For example, consider a network of m binary nodes whose function is to deliver a given throughput from a source S to a destination node D (Rocco and Zio 2005).

Within a CA computational scheme, each node i is mapped into a spatial cell whose neighborhood N_i is the set of network elements which provide their input to it. The state variable s_i of cell i is binary, assuming the value of 1 when node i is operating (active) and of 0 when not operating (passive). Initially all the cells state values are passive.

The hazard propagation problem analyzed in this paper must take into account that the activation of a node is delayed by the time required to propagate the attack from node to node. Hence, the CA becomes dynamic. Indeed, a cell is activated if it is connected to and receives input from at least one active cell or node in its neighborhood. When accounting for the hazard propagation process, the cell activation concerning the hazard also depends on the time required to propagate the attack: the arrival time of the propagated attack is determined as the sum of the current time plus the time delay. If several nodes can propagate the attack to a given node, the arrival time of the attack at such node is determined by the minimum of the times of propagation from all connected nodes in its neighborhood.

Assume now that the generic connecting element (arc) ji from node j to i can be in two states, active ($w_{ji}(t) = 1$) or passive ($w_{ji}(t) = 0$). The ji arc state variable $w_{ji}(t)$

defines the "operational" state of the arc. Initially all $w_{ji}(t) = 0$. As soon as node j is reached by the attack, the state of $w_{ji}(t)$ changes from 0 to 1, for $t = t + TD_{ji}$.

The transition rule governing the evolution of the generic cell i consist of the application of the following rule:

$$s_i(t) = [w_{pi}(t) \vee w_{qi}(t) \vee ... \vee w_{ri}(t)] \quad , p,q,..r \in N_i \tag{1}$$

To account for the time to reach every node in the network, an additional node is introduced that is activated only when all nodes are activated. Finally, to account for the uncertainties in the time delays of hazard propagation, these are assumed to be distributed according to predefined probability distributions which are repeatedly sampled by Monte Carlo simulation (Marseguerra and Zio, 2002) and for each set of sampled values the CA propagation model is run (Zio and Rocco, 2008).

3. Multiple Objective Optimization

3.1 Formulation

In all generality, a Multiple-objective Optimization problem (MO) considers a vector $F(\mathbf{x})$ of objective functions $f_i(\mathbf{x})$, i=1,2,…,k, possibly under specified equality ($h(\mathbf{x})$) and inequality ($g(\mathbf{x})$) constraints:

$$\text{Opt } [F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_n(\mathbf{x}))^t]$$
$$\text{s.t.: } g_j(\mathbf{x}) \le 0,\ j=1,2,..,q;\ h_j(\mathbf{x}) = 0,\ j=1,2,..,r \quad (q + r = m)$$

where

$\mathbf{x}=(x_1, x_2, \ldots, x_n)^t \in \mathbf{X}$ is the vector of decision variables, and $\mathbf{X}$ is the feasible domain.

The solution to a MO can follow two different approaches. The first treats the MO problem as it is, looking for solutions which are simultaneously optimal with respect to all specified criteria (*e.g.*, minimize ANAP and maximize TTRAD), in terms of dominance and Pareto-optimality; the second approach transforms the original MO into a set of single-objective optimization problems to be solved sequentially and then jointly represented (Martorell *et al.* 2004, Ramírez-Rosado and Bernal-Agustín 2001). The former approach is here undertaken within a heuristic scheme of optimality search based on an evolutionary algorithm.

3.2 Multiple Objective Evolutionary Algorithms

Multiple-Objective Evolutionary Algorithms (MOEA) are evolutionary algorithms especially tailored to deal with multiple-criteria problems. They merge the potentiality of metaheuristics with the principles of multi-criteria decision-making, thus yielding algorithms of outstanding capabilities. MOEA are able to deal with non-continuous, non-convex and/or non-linear spaces, as well as problems whose objective functions are not explicitly known (*e.g.,* the output of Monte Carlo simulation runs). State-of-the-art MOEA comprises very efficient optimizers like SPEA2 (Zitzler *et al*. 2001), PAES (Knowles and Corne 2000), PESA-II (Corne *et al* 2000) and NSGA-II (Deb *et al* 2001), among others.

Basically, modern MOEA make multiple iterations of search for the optima following the evolutionary principles of genetic algorithms, simultaneously handling sets of optimal

(non-dominated) solutions instead of single optima. During the process, non-dominated solutions are probabilistically favored over dominated ones through different implementation strategies characteristic of the optimizer. Furthermore, these algorithms allow for the incorporation in the search of elitism by means of a secondary or external population of non-dominated solutions: when performing the recombination of solutions currently available at a given iteration step (*e.g.*, by 'genetic', 'evolutionary' operations such as selection, crossover, mutation), to generate the new candidate solutions for the successive step in the search, a fraction of the solutions to be recombined are taken from the elitist population so that the generated new solutions drive more effectively the search towards the Pareto frontier, *i.e.*, the set of overall non-dominated solutions.

4. MO Optimization of the Protection of a Network System from Hazard Propagation: Two case Studies

As previously stated, the issue of security risk assessment regards "a malevolent intelligence directed towards maximum social disruption" (Apostolakis and Lemon 2005). In this view, it must be expected that an attacker would choose attack those targets that maximize the amount of harm delivered. Adopting a multiple-objective optimization perspective, the attacker would want to arrive at the identification of a set of optimal points of attack in the network which are non-dominated with respect to the multiple-objectives defining the impact of the attack. The defender viewpoint to contrast such situation entails first the assessment of the vulnerabilities of the network system by 'imagining' 'How can someone make something go wrong?', so as to identify all the possible scenarios (Garrick 2002, Kaplan 1997, Kaplan and Garrick 1981) and then 'questioning' 'How can he or she maximize the havoc?', so as to identify the protective actions to take in order to minimize the maximal possible damage.

In practice, the nature of the protections which can be implemented on a network system is strongly dependent on the type of system under consideration. For example, in water distribution networks the attacks may be avoided or rejected by single surveillance of some critical nodes of the network and the propagation may be curtailed by interrupting selected lines of propagation (Skolicki *et al*. 2006).

Nevertheless, the number and nature of the protective actions that can be effectively implemented in practice is constrained by many factors, namely economic, technical or political among others. Assuming for simplicity that these constraints are such to allow protecting only one single attack point, the question is which one should be protected. If we are confident that a 'rational' attacker would target the non-dominated points maximizing the damage, any rational protection scheme should aim at the minimization of the maximal impact, within a so-called min-max policy. For the specific measures of impact here adopted, ANAP and TTRAD, the min-max decision-making problem amounts to finding the node n_p to be protected (see Salazar 2008):

$$n_p = \arg \left\{ \min_{n_p} \max_{n_a} ANPA \wedge \max_{n_p} \min_{n_a} TTRAD \right\}$$

where indexes p and a indicate the nodes to protect and to attack, respectively.

Notice that the min-max criterion is not the only criterion applicable to decide where to place the protection, although it seems to fit best to the management of high-consequence events and is considered "particularly appropriate in the design of robust military system" networks (Shier 1991). Moreover, the min-max is the typical criterion

adopted for optimizing the robustness of discrete domain systems (Salazar 2008, Salazar *et al.* 2006).

In the MO problem at stake, the min-max criterion for guiding the decision on where to place optimally the protection does not lead to a single node selection as one might think, but to a non-dominated set of nodes forming the Pareto frontier. Then, upon the protection of a given node it is possible to analyze the reduction in the impact of an attack to the network by looking at the displacement of the Pareto frontier towards the defender's ideal point (∞, 0) in the (TTRAD, ANAP) plane of attack impact, *i.e.,* the point with maximum TTRAD (infinite time) and minimum number of entities affected (no entity affected). The amount of displacement between the non-dominated Pareto frontiers of the unprotected system and the protected one gives an insight of the actual reduction in the impact of network attacks that can be achieved with the protection of each individual node.

To gain insights on the feasibility of the approach delineated in this paper to tackle the problem of network protection within a MO formulation, two case studies are considered. The first is related to a 52-node network (Manzi *et al*. 2001) and the second to a 332-node network (The topology corresponds to the US airports network (332 nodes and 2126 bi-directional links [http://vlado.fmf.uni-lj.si/pub/networks/pajek/data/gphs.htm]). Through their realistic size, these networks computationally challenge the systematic vulnerability analysis procedure proposed. On the other hand, all other realistic aspects of network operation are neglected in the modeling of the hazard propagation, as underlined at the beginning of Section 2.1.

As mentioned above, the protection strategy considered assumes that budget constraints are such to allow the protection of only one single node in the network and the protection results in its immunisation, *i.e.*, the protected node does not propagate further the hazard.

NETWORK 1

Let us assume that all the network nodes are equally vulnerable, *i.e.*, the cost and the difficulty in delivering the attack is the same for all of them, so that the probability of a node being attacked only depends on the impact resulting from such attack on the whole network.

As for the uncertainties characterizing the propagation of the hazard and its consequences, delays in the time of propagation through the network links are assumed, without loss of generality, to be discrete integer random variables which follow a uniform distribution U(0,10). Furthermore, the numbers of persons affected at the different nodes are also assumed to be uniformly distributed.

For the sake of simplicity only scenarios made by attacks on individual nodes in the network are considered. A total of 500 Monte Carlo evaluations of the CA-based hazard propagation model are performed to assess the consequences of an attack, while accounting for the uncertainties in the hazard propagation and consequences.

Single Objective Analysis

Let us first consider the protection of the network from the consequences of the impact of an attack as measured by the ANAP, with no consideration given to the time available for stopping the hazard propagation, i.e., the TTRAD. Fig. 1a (light shaded trend) shows the ANAP values resulting from the attacks of each of the constituent nodes of Network 1 ‘unprotected’. The maximum value is attained by attacking node 24. The

profile of the maximum impact is also shown in Fig. 1a (white trend), in terms of ANAP, on Network 1 in which one single node (indicated on the x-axis) at a time is protected, i.e., hazard propagation stops at that node. The numerical results of the maximal ANAP given the protected node have been obtained generating profiles similar to that of the unprotected network, but with one network node at a time protected. The actual reduction in the maximal ANAP attainable when introducing a protection is dotted area in Fig. 1a.

The information gained allows rationally devising the protection of a single node so as to minimize the maximal impact. In the case under consideration, this would lead to allocating the protection to node 18.

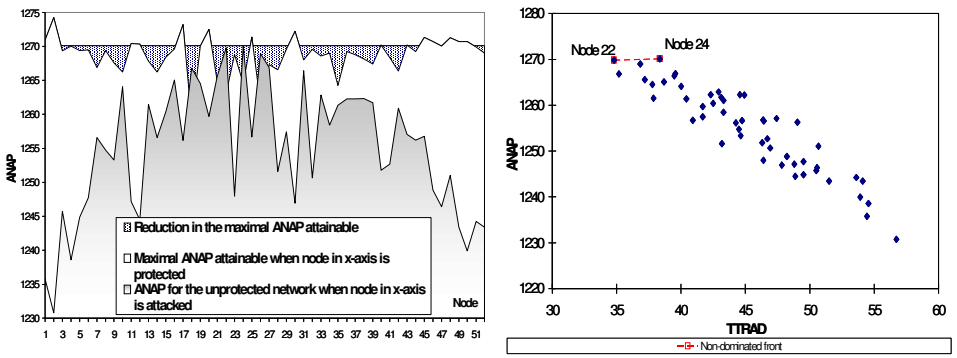


Figure 1: Network 1: a) ANAP profile for the unprotected; b) Attacker's efficient frontier (Rocco *et al.* 2007)

Multiple Objective Analysis

Let us turn now to considering the effectiveness of the protective action against hazard propagation also with respect to the time available for action implementation, as measured by TTRAD. With respect to the previous single-objective analysis focused only on the ANAP measure, the MO analysis extended to the TTRAD measure enriches the information given to the DM in that the inclusion of the TTRAD measure provides him or her with additional elements which may be useful for the identification of the best protective actions. Indeed, the TTRAD provides an indication of the available reaction time against the attack.

Figure 2a represents the two-objective consequence of each hazard propagation scenario following an attack to the various nodes of the network as points on the two-dimensional plane of the objective functions TTRAD and ANAP. Each point depicted represents the TTRAD and ANAP values resulting from an attack to one of the nodes in the unprotected network. The antagonist's aim of finding the set of target nodes of maximum impact amounts to finding among the points in Figure 2a those on the Pareto frontier which are non-dominated with respect to maximizing the ANAP and minimizing the TTRAD. In the case under consideration, nodes 22 and 24 represent optimal targets as their attack would lead to the largest damage in terms of TTRAD and ANAP.

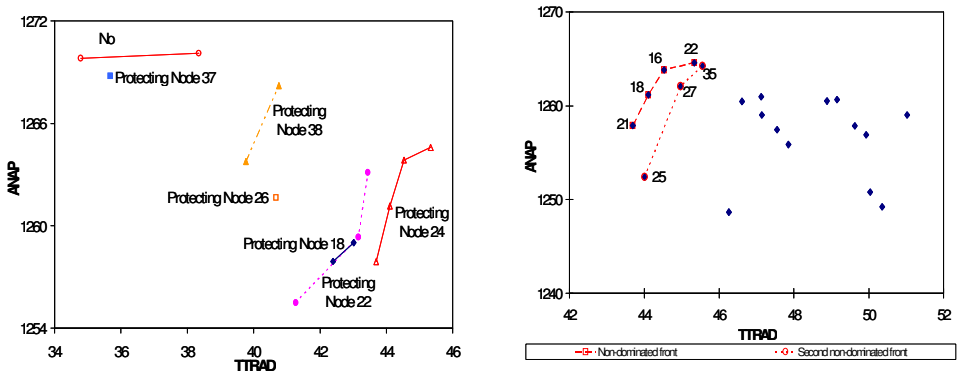


Figure 2: Network 1:a) Selected antagonist's Pareto frontiers for different individual node protections; b) Antagonist's efficient frontiers when node 24 is protected (labels represent nodes)

Figure 2b shows some selected Pareto frontiers for protected network configurations found by the min-max approach illustrated previously. The results can be analyzed either by visual inspection or introducing appropriate metrics. The first way is not applicable in many practical situations. In the case here of interest, the application of the ε-indicator (Laummans *et al* 2002a, b) leads to conclude that the protection of nodes 18, 22 and 24 would minimize the maximal impact, although it does not allow to differentiate which one among these nodes is best to protect. Hence, the rational contribution of the DM is crucial for choosing the final alternative. For instance, if node 24 is selected as the one to deserve protection, an attack on nodes 21, 18, 26 or 22 would produce the maximal impact on the protected network (Figure 2b). These nodes are topological neighbors in Network 1. This additional insight points at a whole section in the topology of the network that might deserve surveillance in order to reduce the risk of incurring in maximal damage from an attack in this area. If the resources were available to protect all the 5 nodes in such area, the maximal impact produced (by attacking nodes 25, 27 or 35) would be distinctly lower than that resulting from the sole protection of node 24.

NETWORK 2

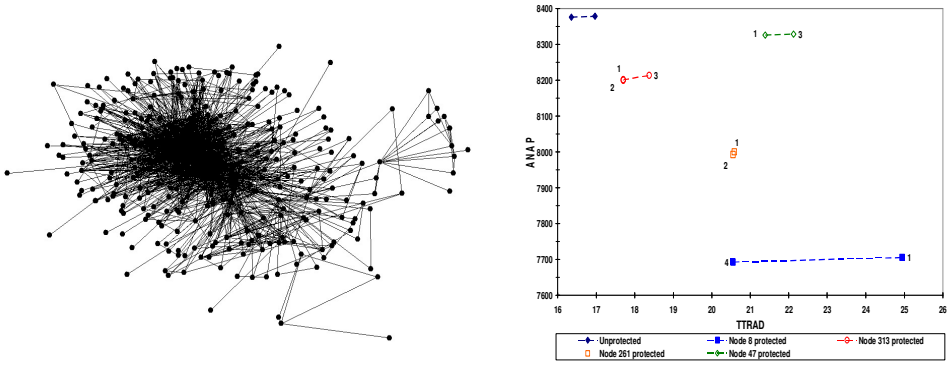
In this case study, the proposed MO approach to network vulnerability analysis and protection optimization is challenged by a network of large size (Fig. 3a). However, except for the topology, no other physical aspect of the work relates to the behavior of the real network.

Again, time delays in the hazard propagation through the network are assumed, without loss of generality, to be distributed according to a discrete uniform distribution U(0,10). The numbers of persons or entities affected at the different nodes are distributed as a discrete uniform distribution U(10,40). The number of MC samples to give due account to such uncertainties is 500.

As in the previous case study, the analysis is restricted to considering scenarios of one attack on a single node at a time and the number of nodes which can be protected is one. A complete enumeration of 332^2=110224 scenarios was considered by combining the attack and protected nodes. For each protected node, the hazard propagation on the

protected network due to the attack of each node is analyzed by CA-MC simulation and the non-dominated nodes, with respect to maximizing ANAP and minimizing TTRAD, are determined. The Pareto frontier obtained constitutes the worst-case, given the protection of that particular node. Finally, among the 332 Pareto frontiers generated in correspondence of the protection of the 332 nodes, one at a time, the one that is farthest from the reference Pareto frontier of the unprotected network provides the most robust protection choice.

Fig. 3b shows some selected Pareto frontiers, including that of the unprotected network. The protection of node 8 leads to the most robust choice in the sense that the reduction in the consequences is maximal, under the assumption of attacking only one single node.



a) b)

Figure 3: Network 2: a) Graph; b) Pareto frontiers for single attacks at different nodes (labels beside the points indicate the nodes of attack).

An additional insight that comes out of this analysis is the identification of special nodes (*cut-nodes*) whose removal disconnects the network, generating islands. As a consequence, if one of these nodes is protected, no matter what node the antagonist attacks, the network can never be affected in its entirety. Conversely, an attack on an unprotected *cut-node* will cause the immediate disconnection of the network, with the concomitant effects; something that could be appealing under certain circumstances. Sound interpretations from both the defensive and the antagonistic viewpoint of this issue and its incorporation into the presented methodology is a subject open to further research.

5. Conclusions and Further Research

The work in this paper has concerned the application of a multiple-objective approach to the vulnerability analysis of a network system exposed to terrorist attacks of uncertain consequences, for the optimization of its protection.

The approach is founded on the power of cellular automata for modeling the dynamics of hazard propagation, Monte Carlo sampling for handling the associated uncertainties and Multiple-Objective Evolutionary Algorithms for searching optimal solutions of network protection.

The viewpoint adopted for addressing the problem of system protection is consistent with the general probabilistic safety assessment and management framework which passes through the identification of accident scenarios that may keep a system from

accomplishing its mission and the quantification of the likelihood and consequences of such scenarios (Garrick *et al*. 2004, Kaplan 1997, Kaplan and Garrick 1981).

The potentials of the systematic procedure of analysis of the protection schemes are demonstrated on two networks of realistic size. Several insights can be drawn. On one side, the methodological approach is of general application, provided that the proper propagation model is introduced and the adequate objectives defining the attack's impact against which to protect are defined. On the other side, the computational efforts required by the approach may limit its applicability. In the case studies analyzed, for simplicity the scenarios considered have been restricted to attacks only on one node at a time and the hazard propagation model has been highly simplified with respect to the physical characteristics of real networks. Extensions to multiple attacks and protection actions on networks of realistic size, like the ones here considered, could lead to a combinatorial explosion in the number of alternative protection strategies to be evaluated as potential optimal solutions and if the evaluation of the hazard propagation model is time-consuming even the 'intelligent' MOEA here adopted for the search of solutions could run into problems.

Hence, as future line of methodological research aimed at reducing the computational effort involved in determining the optimal, robust protective scheme, it could be interesting to obtain a priori information on the most vulnerable nodes of the network through the use of fast 'screening analysis', *e.g.* those which rely on the so called centrality measures to quantify the topological relevance of the network nodes. Similarly, the role of *cut-nodes* in the determination of the robust protective scheme must be investigated further.

Finally, the procedure of vulnerability analysis proposed should always be corroborated by the attentive interpretation by the decision maker of the insights thereby gained, as this could lead to identifying neighborhoods of nodes to be optimally protected for minimum attack impact.

References

[1] Apostolakis, G. E., Lemon, D. M. (2005), "A Screening Methodology for the Identification and Ranking of Infrastructures Vulnerability Due to Terrorism." *Risk Analysis*, vol. 25(1), pp. 361-376.

[2] Bier, V. M., Nagaraj, A., Abhichandani, V. (2005) "Protection of simple series and parallel systems with components of different values". *Reliab Eng Sys Safety*, 87(3):315-323.

[3] Birchmeier, J. "Systematic assessment of the degree of criticality of infrastructures." Specialisation topics, (Eds T. Aven and J. E. Vinnem), in Risk, Reliability and Societal Safety series, vol. 1, 2007, Pages 859-864 (Taylor & Francis, London).

[4] Corne, D.W., Jerram, N.R., Knowles, J.D., Oates, M.J. (2001) "PESA-II: region based selection in evolutionary multiobjective optimization". In: Proceedings of the genetic and evolutionary computation conference (GECCO-2001). Los Altos, CA: Morgan Kaufmann Publishers. p. 283–90.

[5] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T. (2002) "A fast and elitist multiobjective genetic algorithm: NSGA-II". IEEE T Evol. Comput. 6(2):182-197.

[6] Garrick, B.J., (2002), "Perspectives on the Use of Risk Assessment to Adress Terrorism" *Risk Analysis*, vol. 22(3), pp. 421-423.

[7] Garrick, B. J., Hall, J. E., Kilger, M., McDonald, J. C., O'Toole, T., Probst, P. S., Rinskopf, P. E., Rosenthal, R., Trivelpiece, A.W., Van Arsdale, L.A., Zebroski, E.L.

(2004), "Confronting the risks of terrorism: making the right decisions." *Reliab Engnng Sys Safety*, vol. 86, pp. 129-176.

[8] Haimes, Y.Y., Longstaff, T. (2002), "The Role of Risk Analysis in the Protection of Critical Infrastructures Against Terrorism." *Risk Analysis*, vol. 22(3), pp. 439-444.

[9] Hausken, K. (2007) "Protecting infrastructures from strategic attackers". In T. Aven and J. E. Vinnem, editors, Risk, Reliability and Societal Safety, volume 1: Specialisation Topics, pages 881-887. Taylor & Francis. ISBN: 978-0-415-44783-6.

[10] Holmgren, Å.J. (2006) "Using Graph Models to Analyze the Vulnerability of Electric Power Networks". *Risk Analysis*, 26(4):955-968.

[11] Holmgren, Å.J., Jenelius, E., Westin, J. (2006) "Evaluating Strategies for Defending Electric Power Networks Against Antagonistic Attacks". *IEEE Transactions on Power Systems*, 22(1):76-84, February.

[12] Johnson, C.W. (2007) "Understanding the interaction between public policy, managerial decision-making and the engineering of critical infrastructures". *Reliab Engng Sys Safety,* 92(9):1141-1154, September.

[13] Kaplan, S. (1997) "The Words of Risk Analysis." *Risk Analysis*, vol. 17(4), 407-417.

[14] Kaplan, S., Garrick B.J. (1981) "On The Quantitative Definition Of Risk." *Risk Analysis*,1(1), pp. 11-27.

[15] Knowles, J.D., Corne D.W. (2000) "Approximating the nondominated front using the Pareto archived evolution strategy". *Evol Comput* 8(2):149–72.

[16] Korczak, E., Levitin, G., Haim, H.B. "Survivability of series-parallel systems with multilevel protection". *Reliab Engnng Sys Safety*, 90(1), pp. 45-54.

[17] Laumanns, M., Thiele, L., Zitzler, E., Deb., K. (2002) "Archiving with Guaranteed Convergence and Diversity in Multi-Objective Optimization". In E. C.-P. *et al*, editor, Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2002), pages 439-447, San Francisco, California, USA, July. Morgan Kaufmann Publishers.

[18] Laumanns, M., Thiele, L., Deb, K., Zitzler, E. (2002) "Combining Convergence and Diversity in Evolutionary Multi-objective Optimization". *Evol Comput*, 10(3):263-282, Fall.

[19] Levitin, G. (2007) "Optimal Defense Strategy Against Intentional Attacks". *IEEE Trans Reliab*, 56(1):148-157, March.

[20] Levitin, G. and H. Ben-Haim (2007) "Importance of protections against intentional attacks". *Reliab Engnng Sys Safety*, Vol.93(4), pp 639-646.

[21] Levitin, G. Hausken, K. (2007) "Protection vs. redundancy in homogeneous parallel systems". *Reliab Engnng Sys Safety*, Vol 93(10), pp. 1444-1451.

[22] Manzi, E., Labbé, M., Latouche, G., Maffioli, F. (2001), "Fishman's Sampling Plan for Computing Network Reliability," *IEEE Trans Reliab*, vol. R-50, pp. 41-46.

[23] Marseguerra, M., Zio, E. (2002), Basics of the Monte Carlo Method with Application to System Reliability, *LiLoLe- Verlag GmbH* (Publ. Co. Ltd.), Hagen, Germany 2002

[24] Martorell, S., Sánchez, A., Carlos, S., Serradell, V. (2004), "Alternatives and challenges in optimizing industrial safety using genetic algorithms." *Reliab Engnng Sys Safety*, Vol. 86(1), pp. 25-38.

[25] Ramírez-Rosado, I.J., Bernal-Agustín, J.L. (2001), "Reliability and Costs Optimization for Distribution Networks Expansion Using an Evolutionary Algorithm." *IEEE Trans on Power Systems*, vol. 16, pp. 111-118.

[26] Rocco S., C.M., Zio, E. (2005), "Solving advanced network reliability problems by means of cellular automata and Monte Carlo sampling" *Reliab Engnng Sys Safety*, vol 89(2), pp. 219-226.

[27] Rocco S., C.M., Zio, E., Salazar A., D.E. (2007), "Networks Vulnerability: A Multiple-Objective Optimization Approach." *Safety and Reliability Conference ESREL 2007*, 25-27 june 2007. Stavanger, Norway.
[28] Salazar A., D.E., Rocco S., C.M., Galván G., B.J., (2006), "Robustness Analysis: An Information-Based Perspective." *In: Newsletter of the European Working Group 'Multiple Criteria Decision Aiding' (Robustness Analysis Forum).* Series 3, nº 14, Fall 2006 pp. 17-22.
[29] Salazar A., D.E. (2008) "On Uncertainty and Robustness in Evolutionary Optimization-based Multi-Criterion Decision Making". PhD Thesis. University of Las Palmas de Gran Canaria, Spain.
[30] Shier, D.R. (1991), "Network Reliability and Algebraic Structures" Oxford University Press. Great Britain.
[31] Skolicki, Z., Wadda, M. M., Houck, M. H., Arciszewski, T. (2006), "Reduction of Physical Threats to Water Distribution Systems", Journal of Water Resources Planning and Management, Special Issue: Drinking Water Systems Security, Vol. 132, Issue 4, pp. 211-217, July/August
[32] Wadda, M., Skolicki, Z., Arciszewski, T. (2004) "Generation of Terrorist Scenarios for Water Distribution Systems: An Evolutionary Computation Approach". In A. Woodcock and C. Pommerening, editors, Proceedings of the workshop "Critical Infrastructure Protection Project", pp162-170. G.Mason Univ.
[33] Wolfram, S. (1985), "Origins of randomness in physical systems," *Phy. Rev. Lett.*, vol. 55, pp. 449-452.
[34] Zio, E., Rocco S., C. M. (2008) "Security assessment in complex networks exposed to terrorist hazard:" Int. Journal of Crit. Infrastructures, 4(1/2):80-95.
[35] Zio, E., Rocco S., C.M., Salazar A., D.E., Müller, G. "Complex networks vulnerability: a multiple-objective optimization approach." In 2007 Proc. Ann. Reliability & Maintainability Symp., Orlando, USA. Jan 22-25, 2007.. pp 196-201.
[36] Zitzler, E, Laummans, M, Thiele, L. (2001) "SPEA2: improving the strength Pareto evolutionary algorithm". TIK Report No. 103, Swiss Federal Institute of Technology (ETH), Computer Engineering and Networks Laboratory (TIK).

Claudio M. Rocco received his B.Sc. in Electrical Engineering and M.Sc. in Electrical Engineering (Power Systems) from Universidad Central de Venezuela and his Ph.D. from The Robert Gordon University, Aberdeen, Scotland, U.K. He is a full professor at Universidad Central de Venezuela in Operational Research post-graduate courses.

Daniel Salazar has recently been awarded with a Ph.D. in Computational Engineering by the University of Las Palmas de Gran Canaria, Spain. He also holds a BSc in Chemical Engineering and an MSc in Operational Research from Universidad Central de Venezuela.

Enrico Zio (B.S. in nuclear engineering, Politecnico di Milano, 1991; M.Sc. in mechanical engineering, UCLA, 1995; Ph.D. in nuclear engineering, Politecnico di Milano, 1995; Ph.D. in nuclear engineering, MIT, 1998) is a full professor of nuclear engineering and director of the Ph.D. school at the Politecnico di Milano. He is currently active in the IEEE Reliability Society as European Technical Operations Chair and Italian Chapter Chair. He is co-author of two international books and of more than 100 papers in international journals and serves as referee of more than 10 international journals.