AN APPROACH TO MODEL COMPLEX INTERDEPENDENT INFRASTRUCTURES

Stefano Panzieri [*,1] **Roberto Setola** [**,1]
Giovanni Ulivi [*,1]

[*] *Dip.Infomatica e Automazione, Università "Roma Tre", Via della Vasca Navale, 79, 00144, Roma, Italy*
[**] *Università CAMPUS Biomedico di Roma, Via E. Longoni, 83, 00155, Roma, Italy*

Abstract: Developed countries rely on many infrastructures as energy transportation, water supply, telecommunication, etc., which are more and more mutually dependent. This phenomenon represents a new and very dangerous vulnerability: an accidental or malicious (e.g., terroristic attack) fault could spread across, amplifying its negative consequences. This imposes to develop methodologies and tools to support decision makers and infrastructures' stakeholders in the analysis of these new scenarios, and in defining suitable protection strategies. To this end, in this paper, we propose an approach to model interdependent infrastructures which, on the bases of mostly qualitative information, is able to set up a (rather sophisticated) simulator.

Keywords: Complex Systems, Large-scale Systems, Modelling, Networks, Simulation, Critical Infrastructures

1. INTRODUCTION

The welfare of large segments of population depends in developed countries on many technological infrastructures as energy distribution, telecommunications, water supply networks, transportation, etc. (Dunn, 2004; U.S., 2003*a*)

In the very last years, for a lot of economical, social, political and technological reasons, we observed a rapid change in their organizational, operational and technical structures. Among other reasons, this transformation is due to the wide spread of ICT technologies and brought to an increased level of interdependency.

Unfortunately this phenomenon represents a new and very dangerous vulnerability. Indeed, due to the presence of coupling among the different infrastructures, an accidental or malicious failure in one of them may easily spread across, amplifying its negative consequences, and affecting remote (from geographical and/or logical point of view) users. As an example, in 1998 the failure of the telecommunication satellite Galaxy IV produced, beyond several problems in telecommunication and air transportation (due to absence of high-altitude weather reports), also difficulties on the highway: drivers could not perform refuel because gas-stations lost the capability to process credit cards (Rosenbush, 1998).

Other examples about negative effects of interdependencies can be discovered analysing consequences of blackouts occurred in 2003. Specifically, in Italy on Sept. 27th there was a considerable delay in power recovery caused by the cascade failure of the telecommunication systems: SCADA

[1] Member of the Italian Government's Working Group on Critical Information Infrastructure Protection.

operators weren't able to tele-control the generator plants and they had to use manual procedures for the re-starting (waiting also the time needed by the operators to reach the plants!).

A last episode happed on Jan. 2nd 2004 in Rome: a failure into the air conditioned system of an important telecom node led to a large blackout into mobile and fixed communication systems, causing the quitting of the financial transaction into 5.000 banks and in 3.000 postal offices, and also difficulties at the international airport where about 70% of check-in desks were closed.

Any infrastructure is a complex, highly non-linear, geographically dispersed cluster of systems. Interaction is within the cluster and also with their human owners, operators and users. As expected, the presence of interdependencies (many of them hidden or poorly understood) augments of many orders of magnitude the complexity to a level that, as stressed in (Amin, 2002), the conventional mathematical methodologies, that underpin today's modelling, simulation and control paradigms, are unable to handle. However, due to the relevance of the topic many authors are proposing modelling and simulation techniques devoted to the study of this class of systems (Dunn, 2004).

In the literature we find, substantially, two main classes of modelling approaches: *Interdependencies Analysis* and *System Analysis*.

The first one encompasses some qualitative approaches used to help analysts to identify critical infrastructures, and is devoted to better emphasize their interdependencies. On the other side, System Analysis techniques are simulation-intensive approaches able to discover hidden interdependencies and to generate (more or less precise) crisis scenarios. These latter approaches suffer, beside the problem of defining appropriate models, the difficulties of acquiring detailed quantitative information about each infrastructure. Indeed, the more detailed is a model, the greater the number of parameters it encompasses. Some of them, moreover, may be considered sensitive information and infrastructure stakeholders appear generally very reluctant to their disclosure.

To overcome these difficulties, we propose a sort of hybrid approach which, on the bases of mostly qualitative information elicited from infrastructures stakeholders, is able to set up a (rather sophisticated) interdependent infrastructures simulator.

2. INTERDEPENDENT INFRASTRUCTURES MODELLING

Modeling procedures and simulation techniques of individual infrastructures represent a rather well developed field. Numerous products are commercially available to analyse each single infrastructure at different abstraction level, on multiple time scale and with a selectable level of details.

However, modelling and simulation of multiple, interdependent infrastructures are immature by comparison, even thought a number of approaches are under development to directly address interdependencies and to offer insight views into the operational and behavioural characteristics. As stressed in (U.S., 2003*a*), such techniques must be employed to develop creative approaches and enable complex decision support, risk management, and resource investment.

These studies are primarily devoted to determining the downstream consequences of the loss of elements in an infrastructure, such as which other infrastructures are affected (cascading and higher order effects), the geographical extend of the infrastructure outages, economic loses, etc. In particular, they are useful to display how infrastructures react to extreme and rare events, such as major natural disasters, or a catastrophic terroristic attacks. Given the rarity of these events, and the great and rapid innovation that characterizes the today techno-social scenario, the very limited record of historical data available is insufficient to base adequate strategies. Multiple simulations with stochastic variations could provide useful information on structural characteristics and their impact on the welfare of the population (U.S., 2003*a*).

Obviously, no simulation will be predictive, i.e., able to accurately portray the exact consequences associated with each single event. But simulations will provide useful inputs to recovery plans, reconstruction strategies and mitigation plans.

As noted in the Introduction, the modelling proposed in the literature can be divided into two main classes. The first one, **Interdependencies Analysis**, includes qualitative techniques which help to analyse infrastructures' interdependencies. In particular in (Rinaldi, 2001), the authors emphasize how interdependencies should be analysed with respect to six dimensions: Type of Failure, Infrastructure Characteristics, State of Operation, Type of Interdependencies, Environment, Coupling and Response Behaviour. In (Ezell, 2000) the Hierarchical Holographic Modelling is adopted: the whole model is obtained considering a multitude of mathematical and conceptual models each of them devoted to represent

a particular aspect of the system: hierarchy, functions, components, operations, etc.

These models are generally obtained via experts interview, round-table or workshop, and/or with the help of suitable questionnaires. Models are relatively easy to obtain but they are not able to discover hidden critical elements (i.e., elements not explicitly considered by the experts).

The other approach is the so called **System Analysis**. These techniques are quantitative approaches, and need sophisticated computer simulations.

An example of this approach is the project under development by the Los Alamos, Sandia and Argonne Laboratories with the creation of NISAC (National Infrastructures Simulation and Analysis Centre) to model and simulate the system composed by all the infrastructures, and their interdependencies, critical for U.S..

However, due to the huge complexity of the problem, one of the most challengeable task is the development of suitable models able to generate useful predictive information.

To overcome these difficulties, many authors suggest the use of bottom-up approach: the whole system is described starting from its individual parts (Rinaldi, 2001). This kind of approach is generally referred as Complex Adaptive Systems (CAS): i.e., independent networked systems (generally named agents) that autonomously elaborate information and resources in order to define their outputs.

The environment in which each agent acts is defined by the interaction with all other agents, and its reactions are conditioned to the exchanged signals. Usually an agent is conscious of its environment, producing a feedback to those *stimuli* that are in its field of cognition.

Interaction among agents produces the emergence of behaviours that are not predictable by the knowledge of any single agent.

CAS approach is largely used in bio-complexity researches and is particularly useful for situations, including the case of infrastructures interdependencies, with sparse or non-existent macro-scale information.

One disadvantage of these simulation models is that the complexity of the computer programs tends to obscure the underlying assumptions and the inevitable subjective inputs (Dunn, 2004).

An other disadvantage is, as mentioned in the introduction, the difficulty to acquire detailed information about each single infrastructure. This task appears, by its own, a difficult challenge (Moteff, 2003), because this kind of information is considered very sensible by infrastructure stakeholders due to the relevance for their business.

The approach we propose in this paper may be collocated on the borderline between the two classes: it is based on simulation approach but to facilitate information gathering, technical and specific (perhaps sensitive) data are kept to a minimum; the goal is to use coarse grain information obtained by interviewing the managers, in order to have the maximum level of abstraction in the description of internal mechanism and processes of each element.

Actually, the scope is limited to the study of faults propagation and performance degradation in a system composed by heterogeneous interdependent infrastructures and its main aims are:

- Evaluate of the short-term effects of one or more faults;
- Help analysts in what-if analysis;
- Discover the critical elements (i.e., those whose faults produce maximum impact).

Therefore we do not take into account neither fixing or recovery activities nor plant wearing, and assume that human habits are stable.

3. THE PROPOSED APPROACH

The considerations exposed in the previous section suggested us to adopt CAS approach: the model is obtained considering a population of nonlinear mutually dependent systems (agents), each of them representing a macro-component of a given infrastructure.

At the same time, in order to handle many heterogeneous infrastructures into an single framework, we described the behaviour of agents with a sufficiently high level of abstraction to allow the use of a small set of common quantities:

- Operative Level (OL): the ability of the agent to perform its required job. It is only an internal measure of the potential production/service, e.g., for an energy production plant, OL=100% does not means that it is providing the maximum nominal power, but that it could, if required.
- Requirements (R): what the node needs to reach OL=100%.
- Faults (F): the level of failure that affects, for each type of fault, the agent.

These quantities represent the state (memory) of each agent.

Interaction among agents is performed using three inputs:

- Induced faults ($IN.F$): faults propagated to it from its neighbourhoods (described in terms of type and magnitude);
- Input Requirements ($IN.R$): amount of resources requested by other objects;
- Input Operative Level ($IN.OL$): the operative level of those objects whose resources are used in it,

and three outputs:

- Propagated faults ($OUT.F$): faults propagated from the object to its neighbourhoods;
- Output Requirements ($OUT.R$): amount of resources requested to other objects;
- Output Operative Level ($OUT.OL$): the OL of the object itself.

The internal behaviour is related to the interconnected dynamics shown in fig. 1. One is associated to the service that the agent provides (Element Dynamic): input requirements ($IN.R$) coming from subsequent agents, merged with the resources available from foregoing ones ($IN.OL$) and the current operative level (OL), define both the output operative level ($OUT.OL$) and the level of resources it needs ($OUT.R$). Moreover, OL depends on the level of failure of the object (OL is set to zero when F is 100%). The second dynamic (Failure dynamic) is a mix of propagation (from $IN.F$ to $OUT.F$) and an internally generate condition related to agent's memory.

Notice that these dynamics capture the functionality of the node w.r.t. its IN/OUT rather than its mathematical (e.g., differential equations) model. This description of agent's behaviour is highly abstracted, but, at the same time, the formulation is sufficiently rich to leave the infrastructure's expert free to model the element dynamics in the most appropriate way.

Relations among agents are based on their *location* that we characterize in terms of agent's dependencies. In particular, we consider five different kind of dependencies, each of them described via an $n \times n$ binary incidence matrix (where n is the number of agents). In particular, we define:

- An **Operative Level Incidence Matrix** (m_{OL}); where the i-th row represents the set of nodes that need the output of the i-th node to perform their activities;
- A **Requirement Incidence Matrix** (m_R); where the i-th row represents the set of nodes providing the needed resources to i-th agent. Note that even thought generally $m_{OL} = m_R^T$ we did not exploit this feature in order to guarantee a more general formulation;
- Three **Fault Incidence Matrices** (FIMs); where the presence of a 1 in the ij-th position means that a fault may be propagated from the i-th node to the j-th one.

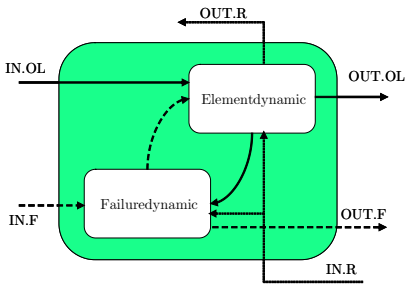

Fig. 1. Agent dynamics.

Note that the existence of a propagation path does not imply in a straight way that a fault in the i-th node induces a failure in the j-th node. Indeed, as better explained later, the target node shall consider also the type of the fault.

In accordance with (Rinaldi, 2001) we consider FIMs which correspond to three different types of interdependencies, namely:

- **Physical FIM** (m_{FP}) that describes faults propagation via the physical linkages (i.e., those related to exchange of physical quantities) between the input and the output of two agents. This kind of fault may be generated or may afflict any kind of agent. Note that

$$m_{FP}(i,j) = 1 \Rightarrow m_{OL}(i,j) = 1$$

  but the converse is not true.
- **Geographical FIM** (m_{FG}) emphasizes that faults may propagate among nodes that are in close spatial proximity. Events such as an explosion or fire could create correlated disturbances to all the systems localised in the spatial neighbour. The matrix m_{FG} exhibit a pattern of 1s characterized by isolated clusters. Inside each cluster, generally, we have a fully connected structure.
- **Cyber FIM** (m_{FC}), this matrix describes the propagation of faults associated with the cyberspace (e.g., virus, worm, etc.). Only a subset of the agents may be affected by this class of fault, i.e., computers and apparatus directly connected to the cyberspace. Obviously, any physical failure is propagated, instead, via m_{FP} or m_{FG}. Cyber-dependency defines, at first approximation, a unicum giant cluster fully connected. This characteristic emphasizes that the cyber-dependency is a global properties (Rinaldi, 2001), i.e., a system that uses the cyberspace is directly connected with any other system that uses the virtual space.

The use of three different FIM matrices, beyond the emphasis on different characteristics of each type of dependency, simplifies the interdependencies' discovery. Indeed, physical interdependencies, and then the possibility of failure propaga-

tion across the underlined channels, are, generally, well known to infrastructure's experts and could be read from the functional schemas. On the other side, geographical interdependencies are less understood by experts, but they can be discovered superimposing infrastructures' maps.

Cyber interdependencies are the less understood and the less considered into risk management plans, but, for some aspects, the most important from a security point of view (U.S., 2003*b*) and the most difficult to model too. Indeed, the hypothesis that cyber-dependency is a global dependency (i.e., nodes are fully connected) is only a rough approximation (even thought this is the better model we have at hand). To have a more precise modelling, we should consider carefully also the topological structure of the cyber-space and its scale-free or small world characteristics (Newman, 2000; Rèka, 2001).

Finally, this class of systems, and specifically their interdependencies, are characterised by a high degree of uncertainties. While it is relatively easy to obtain, at least via experts interviews, qualitative information on them, it is an hard challenge to discover quantitative and precise information. These considerations suggested us the use fuzzy numbers (Kaufmann, 1991) to describe F, R and OL quantities.

4. CRITICAL INFRASTRUCTURE SIMULATION BY INTERDEPENDENT AGENTS

In order to validate the proposed approach, we are developing CISIA: **C**ritical **I**nfrastructure **S**imulation by **I**nterdependent **A**gents. This simulator has been designed for analysing the short-term effects of a failure both in terms of faults propagation and with respect to performance degradations (Panzieri, 2004).

It has been implemented using REpast, a software framework that provides a library of classes for creating, running, displaying and collecting data from an agent based simulations. It is distributed under the GNU General Public Licence, and it has been used to model complex infrastructures by Argonne Laboratories (Argonne).

CISIA extends the Java classes of REpast defining a new class for each type of macro component present into any infrastructure: e.g., electric power plant, transmission line, telecommunication channel, waste-water system, etc.. Each class defines the behavioural's roles of the element and its input/output quantities in term of which resources the agent needed and supply. Moreover, the class defines which type of failure can be propagated to (generated from) the agent. Notice that, as shown in fig. 2, an agent may propagate different types of failure to different set of neighbourhoods.

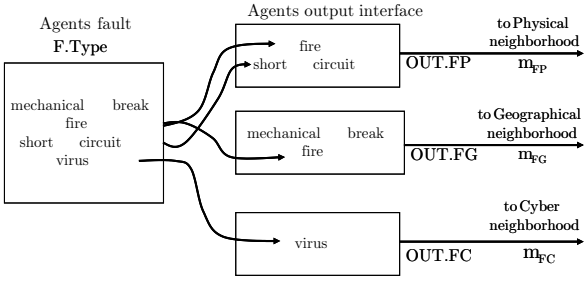

Fig. 2. Faults propagated from an agent. For each class, each kind of failure uses one or more FIM matrices.

Any agent in CISIA is an instance of one of these classes with a given set of parameters (i.e., nominal values). Then a CISIA model is composed by a set of agents that interact with their neighbourhoods, where the agent's neighbourhoods are specified via the incidence's matrices described before.

During simulation, each agent communicates via messages. At every time instant an agent sends messages to its neighbourhoods in order to specify its needs (requirements), communicate its level of service (operative level) and/or propagate faults (physical-faults, geographical-faults and cyber-faults).

CISIA implements an easy-linkage/black box philosophy: any model is obtained connecting together agents without any modification of their internal structure. In particular, no information on the nature, size and type of the target (source) agent has to be explicitly included into the source (target) agent. Model consistence and model coherence are automatically checked at run time.

In CISIA we adopt a triangular representation for the fuzzy number, and OL and R are normalised w.r.t. the corresponding nominal values. However in the presence of overload condition, these variables may assume also values greater than 1. This assumption facilitates the analysis of simulation results because the deviation of a variable from the unit represents an anomaly which calls for more careful investigations.

In order to validate our approach we scaled and particularized it to simulate the system composed by the interdependent infrastructures existing in the University Campus of one of the authors. For sake of simplicity we have considered only the power supply, information and air conditioned infrastructures and focused the attention on their most relevant components. In this way we obtain the model composed by eight agents shown in fig. 3.

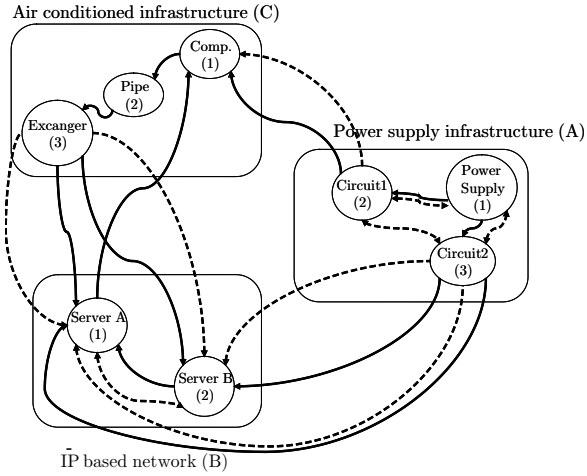

Fig. 3. CISIA model of the test-bed; inks related to operative level incidence matrix (continuos line) and the Physical FIM incidence matrix (dotted line).

Even thought this case study is very simple, and actually the results offered by the simulator did not increase our knowledge of the system, it helped us to validate our modelling approach and to test the correctness of the simulator.

5. CONCLUSIONS

Modelling a system composed by the different and heterogeneous infrastructures at the base of our society is a great challenge for the next years. The intrinsic complexity of each infrastructure, their multi-scale and geographic dispersed nature, the absence of global control mechanisms, and the presence of many physical and logical interdependencies among them, make extremely difficult to predict their behaviour.

However, new and dangerous threats impose to improve the robustness of this network with respect to accidental and malicious (specifically terrorist) actions.

To this end it is mandatory to develop analytical tools able to emphasize the more critical elements and skilled enough to help us to discover hidden interdependencies. Indeed, the presence of these links certainly constitutes the less perceived element of the whole risk, and then one of the major vulnerabilities for the complex system.

In this paper we propose an approach to model heterogeneous interdependent infrastructures using CAS approach. In particular, our approach is devoted to analyse performance degradation and fault propagation immediately after one or more failures (no recovery or repair procedures are take into account).

Even thought the approach uses computer simulation to analyse the different scenario, the modelling of each component is highly abstracted in order to simplify the phase of information gathering. This approach has been used to develop CISIA: a critical interdependent infrastructure simulator.

Work in progress is devoted to validate our approach and CISIA, but also to analyse how intelligent reaction, and autonomy capabilities (e.g., decentralised control strategies),might be used to improve the robustness of the system of system's composed by different heterogeneous and interdependent infrastructures.

REFERENCES

Amin, M. (2002). Modelling and control of complex interactive networks. *IEEE Control System Magazine* pp. 22–27.

Argonne Laboratory, U.S. http://www.dis.anl.gov/msv/msv_cas.html

Dunn, M., I. Wigert (2004). *International CIIP Handbook 2004*. A. Wenger, J. Metzeger, eds. ETH, the Swiss Federal Institute of Technology. Zurich.

Ezell, B., J. Farr, I. Wiese (2000). Infrastructure Risk Analysis Model, *Int. Journal of Infrastructure Systems*, pp. 114–117.

Kaufmann, A., M.M. Gupta (1991). *Introduction to Fuzzy Arithmetic Theory and Application*, Van Nostrand Reinhold, New York.

Moteff, J., G. Stevens (2003). Critical infrastructure information: Disclosure and homeland security. *Report for Congress RL31547*. The Library of Congress. Washington, USA.

Newman, N. (2000). Models of the small world. *Cond-mat/0001118v2*.

Panzieri, S., R. Setola, G. Ulivi (2004). An agent based simulator for critical interdependent infrastructures. *Proc. 2nd Int. Conf. on Critical Infrastructures*.

Rèka, A., A. Baràbasi (2001). Statistical mechanics of complex network. *Cond-mat / 0106096v1*.

REpast *REcursive Porous Agent Simulation Toolkit*. http://repast.sourceforge.net.

Rinaldi, S., J. Peerenboom, T. Kelly (2001). Identifying, understanding and analysing critical infrastructure interdependencies. *IEEE Control Systems Magazine* pp. 11–25.

Rosenbush, S. (May 21st 1998). Satellite's death puts millions out of touch. *USA Today*.

U.S. (2003*a*). *The National Strategy for The Physical Protection of Critical Infrastructures and Key Assets*. The White House, Washington, USA, http://www.whitehouse.gov/pcipb/physical.html.

U.S. (2003*b*). *The National Strategy to Secure Cyberspace*. The White House, Washington, USA, www.whitehouse.gov/pcipb.