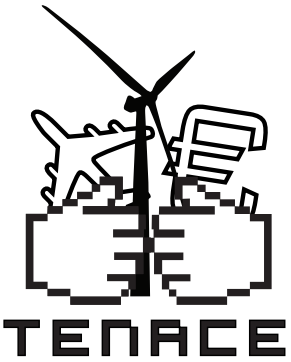
Critical Infrastructure Protection: Threats, Attacks and Countermeasures

Critical Infrastructure Protection
CERT
Threat Analysis
External Audit
Cyber Intelligence
ECIP
Internal Audit
National Strategy
Cyber Security
Intrusion Detection Systems
Cyber Threats
Penetration Testing

March 2014



Critical Infrastructure Protection: Threats, Attacks and Countermeasures

March 2014

Editors

Luca Montanari
Leonardo Querzoni

Authors

Giuseppe Ateniese
Roberto Baldoni
Domenico Daniele Bloisi
Andrea Bondavalli
Francesco Buccafurri
Andrea Ceccarelli
Alessandro Cilardo
Marcello Cinque
Luigi Coppolino
Domenico Cotroneo
Salvatore D'Antonio
Felicita Di Giandomenico
Cesario Di Sarno
Gianluca Dini
Valerio Formicola
Alessia Garofalo
Antonella Guzzo
Luca Iocchi
Gianluca Lax
Antonio Lioy
Paolo Lollini
Federico Maggi
Luigi Vincenzo Mancini
Ilaria Matteucci
Luca Montanari
Leonardo Montecchi
Luca Moretto
Daniele Nardi
Roberto Natella
Antonino Nocera
Nicola Nostro
Ida Claudia Panetta
Antonio Pecchia
Roberto Pietrantuono
Andrea Pugliese
Leonardo Querzoni
Luigi Romano
Domenico Rosaci
Stefano Russo
Marco Vallini
Nino Vincenzo Verde
Stefano Zanero

With the support of

Claudio Ciccotelli
Elizabeth Lee
Fabio Petroni

Contents

Foreword

Critical infrastructures (CI) are at the heart of any advanced civilized country. These infrastructures include among others: finance and insurance, transportation(e.g. mass transit, railways and aircrafts), public services (e.g., law enforcement, fire and emergency services), energy, health care. The recent virus attacks on the SCADA systems of the Iranian nuclear facilities as well as those targeting the telecommunication and power grid infrastructures of Estonia and Georgia show how cyber attacks against CIs are becoming increasingly prevalent and disruptive. In many respects, this results from growing exposure of the Information Technology (IT) employed within CIs to the Internet, which, in turn, is motivated by the desire to cut operational costs by switching to open networking technologies, and off-the-shelf computing equipment. All surveys from leading organizations of the security sector indicate that attacks are expected to increase in scale, to become more accurate and precise, and therefore to become real cyber weapons.

Organizations subject to attacks incur serious tangible and intangible costs, which for example in the context of financial institution could exceed 6 million US dollars per day according to some estimates. This is in addition to numerous intangible costs associated such as damage to reputation and degraded user experience. In the context of the energy and transportation sectors, cyber attacks to such infrastructures could also bring about to loss of human life. Improving cybersecurity knowledge, skills and capability of a nation will be essential for supporting an open society and for protecting its vital infrastructures such as telecommunication networks, power grid networks, industries, financial infrastructures etc.

This manuscript has been conceived and written in the context of the Italian TENACE project, funded under the PRIN 2010 program by the Italian Ministry of University and Research (MIUR). The TENACE project investigates the protection of national critical infrastructures from cyber threats, following a collaborative approach. TENACE addresses three scenarios: financial infrastructure, the power grid and transportation systems. These represent three widely different settings with distinct inter-depencies, threats, vulnerabilities and possible countermeasures.

TENACE has the objective of defining the collaborative technical and organizational methodologies necessary for increasing the protection of such CIs. Furthermore, there is the specific objective of looking at the common steps needed to be taken in steps necessary for developiong a unifying methodology and understanding of the underground economics fueling an attack. This study of specific CI vulnerabilities and related attacks results in the development of algorithms, models, architectures and tools as the means to enable the effective protection of critical infrastructure, enhancing their degree of security and dependability by considering a continuously evolving adversary.

The manuscript provides the reader, in an accessible manner, with a state-

of-the-art analysis of the protection of financial, power grid and transportation infrastructures from cyber attacks. It points out standard solutions for securing the specific critical infrastructure, the type of attacks and exploited vulnerabilities, strategies of protection and fault mitigation approaches. The manuscript has been edited by Luca Montanari and Leonardo Querzoni with content written by several contributing scientists belonging to TENACE consortium.

TENACE is composed of a multidisciplinary group of academic scientists from nine among of the most prestigious Italian universities (University of Rome La Sapienza, University of Naples Federico II, Polytechnic Institute of Milan, University of Trento, University of Florence, Polytechnic University of Turin, University of Naples Parthenope, University of Pisa, University of Reggio Calabria, University of Calabria) and the National Research Council (CNR).

Rome, 3 February 2014

Roberto Baldoni
TENACE Project Coordinator
Center of Cyber Intelligence and Information Security
Università degli Studi di Roma “La Sapienza”

Abstract

Today, most modern countries base their economic wealth and societal prosperity on several infrastructures. These infrastructures constitute the cornerstone of a country's growth, and thanks to this role they are considered critical assets that must be protected against possible attacks and malfunctioning. Being a topic treated independently by different countries around the world, there is no homogeneous concept of what protecting critical infrastructures means. This document tries to provide an analysis of the current worldwide situation regarding this topic, whilst maintaining a specific focus on the European scenario.

The document structure is divided in three parts. In the first part (Chapter 1), the document provides an overview of the various definitions of what a critical infrastructure is and what its protection means. This first part reports both on common knowledge about the topic and on how different countries adapted their legislative frameworks to encompass critical infrastructure protection issues. In the second part (Chapters 2, 3, 4 and 5), the document deals with more technical aspects. First it provides a background on threats, vulnerabilities and accidental faults that threaten these infrastructure in general and then it analyzes the peculiarities of three specific scenarios: financial systems, power grids and the transportation sector. The third part of the document (Chapter 6) concludes by analyzing the maturity of critical infrastructure protection in Italy, providing several figures and statistical data.

TENACE - Protecting National Critical Infrastructures from Cyber Threats is a research project funded by the Italian *Ministero dell'Istruzione, dell'Università e della Ricerca* under the program *Progetti di Ricerca di Interesse Nazionale* (project number 20103P34XC). Further information on TENACE is available on the official website at `http://www.dis.uniroma1.it/~tenace/`.

CHAPTER

1 Critical Infrastructure: Definitions and Concepts

The protection of national critical infrastructures is nowadays considered a paramount objective by all modern countries around the world. Assessing the complexity of the problems involved in reaching this goal must necessarily start from a coherent vision of what a critical infrastructure is and what critical infrastructure protection actually means. As of today, there is no universal recognized definition of critical infrastructure. This chapter provides an overview of the most common definitions, with a strong focus on the European case, which embraces all the European Union member states (Section 1.1). Some other definitions regarding other realities are reported. Furthermore, this chapter analyzes the governance and legislative aspects of the Italian critical infrastructure protection and provides an overview of some other developed countries (Section 1.3). Finally, this chapter discusses open problems (Section 1.4) and introduces a prominent problem in the field of critical infrastructure protection: managing the inter-dependencies among infrastructures.

1.1 Critical infrastructure definitions

With reference to the critical infrastructure (CI), despite the numerous attempts made so far, there is still no universally recognized definition, or at least a definition that provides a classification suiting the characteristics of each nation. A critical infrastructure is often identified as that infrastructure whose incorrect functioning, even for a limited time period, may negatively affect the economy of individual subjects or groups, involving economic losses and/or even expose people and things to a safety and security risk [57].

Within the European Union a Critical Infrastructure is defined as “an asset, system or part thereof located in member states which is essential for

the maintenance of vital societal functions, health, safety, security, economic or social well-being of people, and the disruption or destruction of which would have a significant impact in a member state as a result of the failure to maintain those functions"[17]. While a European Critical Infrastructure (ECI) is defined as a "critical infrastructure located in Member States the disruption or destruction of which would have a significant impact on at least two Member States. The significance of the impact shall be assessed in terms of cross-cutting criteria. This includes effects resulting from cross-sector dependencies on other types of infrastructure"[17].

The designation of a critical infrastructure as an ECI is the result of a complex technical-political process, which arises from the potential impact that can be caused by a failure/destruction of an infrastructure in terms of sectoral and inter-sectoral relevance. The inter-sectoral evaluation criteria relate to:

- potential victims, in terms of number of fatalities or injuries;
- potential economic effects, in terms of financial losses, deterioration of products or services, and environmental effects/damages;
- potential effects on population, in terms of impact on public confidence, physical suffering and disruption of daily life, including the loss of essential services.

What mainly emerges from the European directive[17] quoted above, is that the obligations of owners/operators for what concerns the security of their infrastructures should be made to prevent, or at least limit, the consequences on other nations. In other words, given the pan-European role played by such large infrastructure, security levels must conform to a high qualitative standard and, thus, the rules to be adopted are not defined only by the member state in which such infrastructures are located, but, to some extent, they are imposed at an European level. An essential component of the European Programme for Critical Infrastructure Protection (EPCIP) is the Critical Infrastructure Warning Information Network (CIWIN), a protected public internet-based information and communication system that allows subjects involved in Critical Infrastructure Protection (CIP) to share CIP-related information and good practices.

An interesting alternative definition, independent from the one provided in the EU directive, can be found in the "International CIIP Handbook 2008/2009" [57], where the CI are identified as "infrastructure whose incorrect functioning, even for a limited time period, may negatively affect the economy of individual subjects or groups, involving economic losses and/or even expose them to safety and security risk".

By looking at the United States scenario, the Public law 107–56 (October 26, 2001) of United States defines critical infrastructure as "systems and assets,

Energy	Nuclear Industry
ICT	Water
Food	Health
Financial	Transport
Chemical Industry	Space
Research Facilities	

Table 1.1: EU draft list of critical infrastructure activity areas [14].

whether physical or virtual, that are so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters".Despite the presence of a few differences, in essence both the above definitions look at identifying potential threats like human error, occasional accidents and attacks that can lead to a malfunction or onset of the crisis of the CI under observation.

In 2006 the European Commission defined network and information security as "the ability of a network or an information system to resist (...) accidental events or malicious actions that compromise the availability, authenticity, integrity and confidentiality of stored or transmitted data and the related services offered by or accessible via these networks and systems"[15]. The Critical Information Infrastructure Protection (CIIP) is thus crucial, both for autonomous infrastructure and for those that are functional to the operativeness of other critical infrastructures. CIIP includes "the programs and activities of infrastructure owners, operators, manufacturers, users, and regulatory authorities which aim at keeping the performance of critical information infrastructure in case of failures, attacks or accidents above a defined minimum level of services and aim at minimizing the recovery time and damage"[13].

Over the years the European government has drafted lists [14] identifying 11 areas in which critical infrastructure operate (see Table 1.1).

The motivations that brought the EU to list those CI are often straightforward: the banking and financial services play a vital role in the economy of each country, so that a violation would be a huge risk for the entire system. Also the energy sector is critical. The electrical energy has various features, including the ease of conversion into other forms of energy (mechanical, light, thermal, etc.), the ease and flexibility of transport, the possibility of a widespread distribution and, at the same time, it is storable, only in limited quantities. This means that, at any time, the demand must be balanced by the production of energy. The need to use ICT technologies exposes the mentioned areas to the risk of computer breaches. It should be noted that with the promulgation of the council directive 2008/114/EC [16], the European government accepted only two of the areas listed above, namely energy and transportation, as those where ECI operate. Furthermore, several countries

within the European Union independently provide lists of critical sectors at a national level.

1.2 EU cybersecurity strategy

The maintenance of a good level of cybersecurity in the EU context involves disparate sectors with different jurisdictions and responsibilities, both at national and EU level. Managing cybersecurity through centralized supervision at European level is not feasible. National governments have the main responsibility for the maintenance of a good level of security and must cooperate at EU level in case of risks and security breaches that extend beyond national boundaries.

The structures involved in the maintenance of cybersecurity are organized in three fundamental areas: Network and Information Security (NIS), law enforcement and defense. At national level member states should have already, or as a result of the European cybersecurity strategy, national structures in each of the aforementioned areas (see Figure 1.1). Member states are responsible for carefully defining the roles and responsibilities of such national structures.

The European strategy invites member states to encourage information sharing between national structures involved in cybersecurity and the private sector, so that they can have both a comprehensive vision of risks and security threats, and a better comprehension of cyber crime techniques so as to respond more rapidly and effectively.

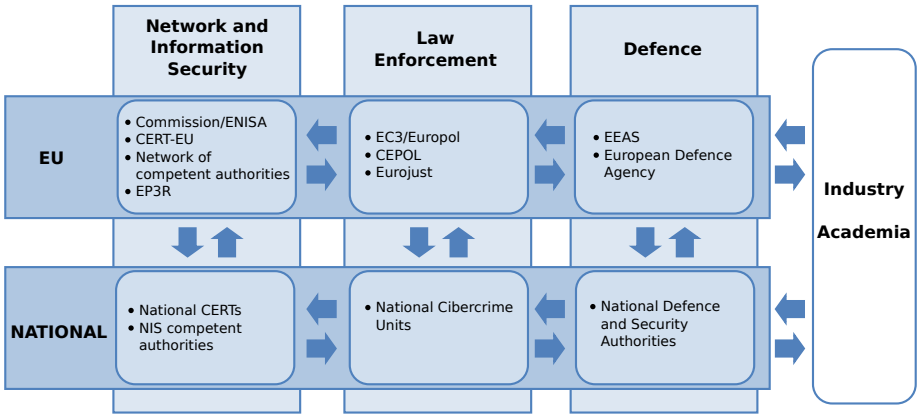


Figure 1.1: EU cybersecurity strategy: Interacting organizations at national and EU level [30].

Several organizations are involved at EU level. In the NIS area, the European Network and Information Security Agency (ENISA), established in 2004, is responsible for improving network and information security. Cur-

rently a new regulation [20] to strengthen ENISA and modernize its mandate is under examination by the Council of Europe and the European Parliament. ENISA will also be responsible for building expertise in security of industrial control systems, transport and energy infrastructure. A Computer Emergency Response Team at EU level (CERT-EU), responsible for the security of the IT systems of EU agencies and institutions, was established in 2012.

Furthermore, in March 2009, the European Commission established the European Public-Private Partnership for Resilience (EP3R) with the objective of encouraging sharing of NIS related information between interested parties in the public and private sector at European level. In the area of law enforcement, in 2013, the European Cyber Crime Centre (EC3) was formed within Europol to represent the European focal point of the fight against cyber crime. In particular, EC3 will provide analysis and intelligence, support investigations, provide high level forensics and facilitate cooperation and information sharing between the competent authorities of member states, the private sector and other stakeholders. Europol/EC3 and Eurojust will cooperate closely to improve their capability in fighting cyber crime. In the area of defense, the main responsibility for cyber defense at EU level is the European Defence Agency (EDA). The European strategy for cybersecurity supports cooperation and information sharing between these organizations, in particular ENISA, Europol/EC3 and EDA, and between these and their counterparts at national level.

Finally, at an international level the European Commission and the member states engage in dialogue with international partners and organizations such as the Council of Europe, OECD, OSCE, NATO and UN. ENISA provides a list of national cybersecurity strategies through its website [3].

1.3 National protection strategies

The regulations imposed at the European level, introduced in the previous sections have been accepted by the member states in different ways. This section describes several case studies of national protection strategies at the European level. Furthermore, the US case is described as it provides a meaningful comparative example of regulations that do not need to adhere to EU directives.

1.3.1 Italian governance and legislative overview

The intrinsic characteristics of cybersecurity require a national strategic plan for critical infrastructure protection and the identification of practices to realize it as well as response actions to threats with tools, also organizational, able to face the new socio-technological context and the interdependencies produced by cyberspace; in other words, cybersecurity governance. To achieve

this goal primary and secondary regulation which individuates specific competency areas, jurisdictional areas, involved subjects, types and modalities of a-priori and a-posteriori intervention is needed, thus applying extra-national regulations. The growing number of threats and security breaches has already caused considerable economic damages, thus reducing users' confidence in the use of new services and technologies and hindering the development of electronic commerce and the implementation of the so called "digital agenda" in Italy. In this area, Italy presents a slight delay in definition of cybersecurity governance; even though the cybersecurity issue has been debated in Italy since the early 2000's, significant improvements in the identification of a road-map for the implementation of a national strategy have been observed only recently. The main milestones that brought about the definition of roles and responsibilities for safeguarding Italian cybersecurity are reported in the following section.

The inter-ministerial decree of September 21, 1999 established a working group made up representatives from the Ministry of Communications (*Ministero delle Comunicazioni*), of Justice (*Ministero della Giustizia*), and of Internal Affairs (*Ministero dell'Interno*), with the task of operating in the sector of network security and communications protection as a support to administrative and regulatory interventions. To achieve preset goals the working group, after the analysis of the requirements, in terms of technical and regulatory support, resources, for a "safe" evolution of the telecommunications services, the nature of the relationships between public administration and telecommunication operators, mainly dealt with internationally harmonized regulations in the telecommunication sector, internationally harmonized.

In 2003, the working group was converted into an observatory for network and communications protection and security (*Osservatorio permanente per la sicurezza e la tutela delle reti e delle comunicazioni*), within the Ministry of Economic Development (*Ministero dello Sviluppo Economico*), with the aim of taking into account the technological and regulatory evolution of the different aspects of the telecommunications sector, with particular attention to security. It is permanently integrated with representatives of the Ministry of Defense (*Ministero della difesa*), the department of Public Service (*dipartimento per la funzione pubblica*), Department of Innovation and Technologies(*dipartimento per l'innovazione e le tecnologie*) and Ministry of Productive Activities(*Ministero delle attività produttive*). The observatory, among others, has played a supporting role aimed at transposing Directive 2002/58/EC into reality. This directive concerns the processing of personal data and the protection of privacy in the electronic communications sector, and the legislative decree concerning the Electronic Communications Code (Codice delle comunicazioni elettroniche) which was issued September 16, 2003.

In October 2001, the Technical Interdepartmental Committee of Civil Defense (*Commissione Inter Ministeriale Tecnica della Difesa Civile* - CITDC) was established as a political and military unit supporting organ for the tech-

nical coordination of civil defense activities in case of crises. It operates within the Department of Fire and Public Rescue and Civil Defense (*Dipartimento dei Vigili del Fuoco e del Soccorso Pubblico e della Difesa Civile*).It has the role of evaluating emergencies and planning the measures to be taken in the event of crisis. The committee also considers other hypotheses of risk, not directly related to malicious acts, which can lead to situations of crisis for the continuity of government as well as damage to the population and, in general, the security of the country.

In March 2003, the Ministry for Innovation and Technology established the Working Group on CIIP , in which representatives from ministries involved in critical infrastructure management (Interior, Infrastructure, Communication, etc.), major private providers (ABI, ASI, CESI, GRTN, RFI, Snam Rete Gas, Telecom Italia, Wind and others) and the research and academic world took part. In March 2004, this working group issued the document "Critical Information Infrastructure Protection: The Italian Situation" [12], in which the results of work carried out during the previous year are reported.

With regards to aspects strictly related to critical information infrastructure protection, the legislative decree (D.L.) n. 155 of 31/7/05 (the so called *Legge Pisanu*) conferred the jurisdiction to the Ministry of Interior identifying the Postal and Communications Police as the unit responsible for law enforcement initiatives against cyberattacks on critical information infrastructures. In 2008 the Ministry of Interior established a national center for cyber crime prevention for critical infrastructure called *Centro Nazionale Anticrimine Informatico per la Protezione delle Infrastrutture Critiche* (CNAIPIC) as a special unit within the Postal and Communication Police Service [18]. CNAIPIC acts as a police authority for all activities of prevention, repression and contrast of criminal actions committed against the different critical infrastructure through the cyberspace. For this purpose CNAIPIC and critical infrastructure maintain dedicated and protected exclusive telematic links, for a mutual and constant sharing of data and information relevant to the practice of the assessment, prevention and repression of threats and cyber crime.

Furthermore, the "Unit for Cyber Crime Analysis" (*Unità d'analisi del crimine informatico* -UACI) was established to study and analyze the phenomenon of cyber crime in partnership with major Italian universities. Territorial compartments have an organization similar to that service, with a more operative profile and more bounded to their jurisdictions. These compartments manage the legal cases and the emergencies arising from citizen reports to police hotlines.

In order to enhance the effectiveness of strategies against cyber crime, the police participates, with some of its representatives, in permanent working groups, established by government or international organizations, including the Inter-ministerial Group for Network Security (*Gruppo Interministeriale per la sicurezza delle reti*), G8, the European Community, the Council of Europe, OCSE, Interpol, Europol. Moreover, it cooperates with the institutions

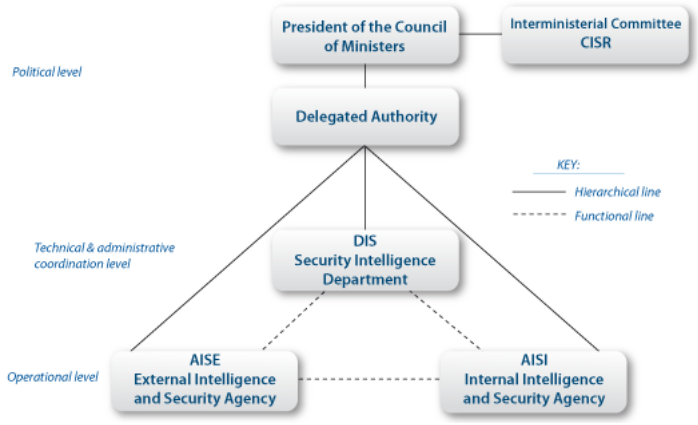


Figure 1.2: Organization of the Intelligence System for the Security of the Republic (www.sicurezzanazionale.gov.it).

(including the Ministry of Communications and Authority for the Communications) and private operators dealing with communications in general.

The Parliamentary Committee for the Security of the Republic (*Comitato parlamentare per la sicurezza della Repubblica*) (COPASIR), established by law n.124 of August 3, 2007, is aimed at ensuring, in a systematic and continuous way, that the activity of the system for information security, is carried out in compliance of the Constitution and laws, in the exclusive interest and for the protection of the Republic and its institutions. COPASIR has important advisory jurisdictions; in particular, the parliamentary body is required to express its non-binding opinion on every regulatory scheme concerning the organization and management of entities involved with sensitive information and security affairs. COPASIR and its president are recipients of an informative stream from the government and the intelligence agencies and, in such area, there is a formal obligation for them to inform in advance the president of COPASIR about the designation of the directors and vice-directors of DIS, AISE and AISI. COPASIR reports unlawful or irregular behavior, detected on the basis of performed controls, to the prime minister and to the presidents of the Upper and Lower House. In addition to an annual report, COPASIR may also submit to parliament urgent reports.

The same law (124/2007) deeply changed the Intelligence System for the Security of the Republic (*Sistema di informazione per la sicurezza della Repubblica*) (Figure 1.2), which currently consists of the complex of organs and authorities that have the task of ensuring the information activities for security in order to protect the Republic against each type of risk and threat both within the country and outside.

Line of action on the Italian cybersecurity. According to the Italian Digital Agenda (*Agenda Digitale Italiana*), the national cybersecurity strategy plans to act on the following areas:

- Educate citizens and enterprises: raise awareness of citizens, business and industry about the serious risks related to the use of the Web (e.g. UK initiative "Get Safe Online", public-private online campaign to raise awareness);
- Enhance threats detection and contrast tools: develop tools (intended as organizations, processes, legislation and applications) able to detect and contrast potential threats (e.g. the National Cyber security Centre of the Netherlands will adopt tools to enhance awareness and classification of threats and vulnerabilities through public-private information sharing);
- Promote education: create education paths able to provide the necessary competences from the early school levels (e.g. the United States has issued a draft for "National Initiative for cybersecurity Education Strategic Plan" which outlines the educational steps, beginning at primary school, for a career in cybersecurity);
- Strengthen public-private cooperation: create mechanisms of debate, sharing and coordination between the public and private sectors, especially with regard to critical infrastructure protection (e.g. Germany, in its strategy, envisaged a National Cybersecurity Council where representatives of the private sector are asked to participate as associate members);
- Strengthen mechanisms of international cooperation: involvement in international forums for the discussion of standards, policy and international principles on cybersecurity (e.g. Czech Republic's strategy envisages an active participation at EU and NATO forums);
- Create and enhance mechanisms for incident response: It is necessary to enhance, through the establishment of national CERTs (Computer Emergency Response Team) and, in some cases, to create, specialized structures able to respond to cyberattacks and incidents within national boundaries and able to coordinate with the corresponding centers at international level;
- Define a standard for the management of digital identities as well as guiding principles for the creation of a federal system at national and international level, able to satisfy the daily needs of digital citizens, including improved security for Internet payment systems;
- Stimulate the growth of an Italian cybersecurity industry, concerning both technologies/services, and skills and talents. This will allow not

only for the growth and the maintenance of specialized competences, but will also attract talents and experts from other countries.

1.3.2 The situation in other countries

Germany. The German Federal Government provides a substantial contribution to cybersecurity, maintaining and promoting economic and social prosperity in Germany. The latest German strategy, 2011, mainly focuses on civilian approaches and measures. These are complemented by the measures undertaken by the armed forces (Bundeswehr) aimed at protecting its capabilities and measures based on mandates to include cybersecurity as part of the preventive security strategy. The global nature of information and communication technologies raises the necessity for an international vision and coordination on security policy aspects with the aim of enhancing cybersecurity capabilities of the international community. For this purpose, Germany cooperates with the United Nations, the European Union, the Council of Europe, Nato, G8, OCSE and other international organizations.

The German strategic plan is organized in 10 specific strategic areas:

1. *Protection of CII (Critical Information Infrastructure)* – CIIs constitutes the central component of almost all critical infrastructure. Thus, protecting such infrastructure is the primary objective of cybersecurity. In order to support CIIs protection the introduction of new technologies is taken into consideration by the plan. Cooperation and information sharing between public and private sectors is also a priority.

2. *Security of IT systems* – Germany aims to support security of IT systems with informative intervention, to provide citizens and small and medium-sized businesses with consistent information concerning risks related to the use of IT systems, and by promoting the use of fundamental security functions, such as state certified electronic proof of identity and De-mail[1]. Furthermore, providers will have to make available to clients a basic collection of security products and services and might be subject to greater responsibilities.

3. *Strengthening IT security in the public administration* – The German plan for strengthening IT security in the public administration includes the creation of a common, uniform and secure network infrastructure in the federal administration to serve as the basis for electronic audio and data communications.

4. *Creation of a National Cyber Response Centre* – The National Cyber Response Centre aims to optimize cooperation between state authorities,

[1] De-mail is a German government communication service similar to the Italian certified e-mail service (*Posta Elettronica Certificata* - PEC).

thus improving response to IT incidents. Information sharing on vulnerabilities, form of attacks and profiles of attackers allow the National Cyber Response Centre to analyze IT incidents and provide recommendations for action to be taken in response to incidents. To favor readiness for IT incidents, the National Cyber Response Centre will submit recommendations to the National Cybersecurity Council both regularly and when specific incidents occur. In case of cybersecurity incidents that reach the level of a crisis the National Cyber Response Centre will directly inform the crisis management staff headed by the State Secretary at the Federal Ministry of the Interior.

5. *Creation of a National Cybersecurity Council* – The National Cybersecurity Council will coordinate preventive tools and the interdisciplinary cybersecurity approaches of the public and private sector. Several ministries of the state and representatives of the federal states (Lãnder) will participate in the council. Representatives from business and academia will be invited on specific occasions.

6. *Effective crime control in cyberspace* – The German strategic plan envisages the strengthening of the capabilities in fighting cyber crime of law enforcement agencies, the Federal Office for Information Security and the private sector. To deal with global cyber crime Germany will make an effort to achieve global harmonization in criminal law based on the Council of Europe Cyber Crime Convention, and will also examine whether new conventions on cyber crime should be adopted at UN level.

7. *Effective coordinated action to ensure cybersecurity in Europe and worldwide* – The German federal government recognizes the importance to conform to European and international standards related to cybersecurity. At European Union level Germany adopts measures based on an extension and moderate enlargement of the mandate of ENISA. Germany intends to shape its external cybersecurity policy so that German interests and ideas concerning cybersecurity will be pursued by international organizations, such as the United Nations, OSCE, the Council of Europe, OECD and NATO.

8. *Use of reliable and trustworthy information technology* – Given the importance of availability and reliability of IT systems, Germany intends to increase research into IT security and critical information infrastructure protection, in particular, by further developing its technologies in these areas. Moreover, Germany approves diversity in technology, combining, when necessary, its own resources alongside those of its partners and allies, favoring the use of technologies certified by international standards.

9. *Personnel development in federal authorities* – One of the priorities of the federal government is to examine whether authorities require additional

staff to enhance cybersecurity. In order to improve inter ministerial cooperation it will favor personnel exchange between federal authorities, providing appropriate staff training measures.

10. *Tools to respond to cyberattacks* – In order to achieve an adequate preparedness against cyberattacks, the German government recognizes the importance of the creation, in collaboration with the specific state authorities, of a collection of tools to effectively respond to cyberattacks.

The objective of the German government is the sustainable implementation of these strategic objectives to ensure freedom and prosperity in Germany. Technologies used in the area of IT security have short innovation cycles. Thus, the German Federal Government will periodically verify whether the objectives of the strategic plan have been achieved, under the control of the National Cybersecurity Council, and will conform them, if necessary, to national and international requirements.

United Kingdom. The UK strategy builds on more than ten years of development. The first step was carried out in 2001 by the Communications-Electronics Security Group (CESG). This group recognized that the increasing use of online services required the development of security measures to protect data and recommended the appointment of a central sponsor for information assurance of government data. Therefore, the government published its first national strategy in 2004, in which a network of Senior Information Risk Owners was established.

In 2009, the government recognized the risk of cyberthreats and published its first cybersecurity strategy. In 2010 the government ranked cyberattacks as a key risk for national security and announced a fund of 650 million pounds for a four-year National Cybersecurity Programme. Since 2011 the Cabinet Office has been responsible for cybersecurity. The most recent strategy was published in 2011 and set out how the government planned to deliver the National Cybersecurity Programme until 2015. Four objectives characterize the strategy:

- Tackling cyber crime and making the UK one of the most secure places in the world to do business;
- Making the UK more resilient to cyberattack and better able to protect its interests in cyberspace;
- Helping shape an open, stable and vibrant cyberspace which the UK public can use safely and which supports an open society;
- Building the UK's cross-cutting knowledge, skills and capability to underpin all cybersecurity objectives.

Six central departments and nine government organizations are responsible for delivery: Home Office, Serious Organized Crime Agency, Child Exploitation and Online Protection, Police Central e-crime Unit, Police force, National Fraud Authority, Department for Business, Innovation and Skills, Technology Strategy Board, UK Trade and Investment, the Cabinet Office, the Intelligence and Security Agencies, Ministry of Defense, Department for Culture, Media and Sport, Foreign and Commonwealth Office.

Concerning critical infrastructure protection in the United Kingdom, everything is delegated to the Centre for the Protection of National Infrastructure (CPNI). CPNI protects national security by providing protective security advice, in terms of personnel security, physical security and cybersecurity. CPNI takes into special consideration the policy context. Policy considerations are one of the building blocks of the mechanism of protective security advice provided by CPNI. In particular, several government policies influence CPNI's work:

- National security strategy: Establishes the strategies aimed at reacting effectively and rapidly to security threats, such as: acts of terrorism, attacks on UK cyberspace, natural accidents and disasters and international military crises that involve the United Kingdom and its allies.
- Strategic defense and security review: Establishes how the objectives of the national security strategy have to be pursued.
- Counter terrorism strategy: The UK's counter terrorism strategy is developed in four main directions: prevent, pursue, protect and prepare. CPNI's work falls within the "protect" category which aims at reducing the vulnerability of the UK to terrorist attacks.
- Cybersecurity strategy (as described above).
- National Risk Register: The National Risk Register is the public version of the confidential National Risk Assessment that registers the events that may cause damage to people or property, or disruption of essential services. Events are categorized in three broad areas: natural events, major accidents, malicious attacks.
- Resilience of infrastructure to natural hazards: In order to enhance critical infrastructure and essential services resilience to disruption due to natural hazard, the Civil Contingencies Secretariat within the Cabinet Office developed the Critical Infrastructure Resilience Programme (CIRP).

CPNI actively cooperates with partners in the public and private sector. In the public sector CPNI works closely with the National Technical Authority for Information Assurance (CESG) and, within the police, with the National

Counter Terrorism Security Office (NaCTSO) and with the Counter Terrorism Security Advisor (CTSA) network. Government departments are responsible for taking appropriate actions to improve security in their respective sectors. These departments are also responsible for the identification of critical infrastructure in their sectors in cooperation with CPNI and sector organizations. The departments involved are:

- Department for Business, Innovation and Skills;
- Department of Health;
- Department for Communities and Local Government;
- Department for Transport;
- Home Office;
- Department for Energy and Climate Change;
- HM Treasury;
- Department for the Environment, Food & Rural Affairs and Food Standards Agency;
- Cabinet Office.

Concerning cybersecurity, the U.K.'s government established in 2010 the Cybersecurity Operations Centre (CSOC) and the Office of Cybersecurity and Information Assurance (OCSIA). CPNI cooperates with CSOC, OCSIA and CESG in order to conduct the cybersecurity program for the UK government. In the private sector, CPNI interacts with the organizations that operate in the national infrastructure. The relationships, established over the years, between CPNI's security advisers and security managers in several sectors enable information sharing between trusted entities and, when appropriate, sharing of vulnerabilities and effective response measures in order to improve the protection of the national critical infrastructure and private organizations. Moreover, CPNI has established a partnership program, Risk Management Delivery Group, which aims to promote strong links between the principal UK consultancy partners.

France. The French president first presented the French strategy on defense and national security in June 2008 with the French White Paper on Defense and National Security. Given the unexpected emergence of cyberspace in the field of national security, in 2009 the government set up the French Network and Information Security Agency (*Agence nationale de la sécurité des systèmes*

d'information - ANSSI)[2]. In 2010 the president decided to give the agency, in addition to its security role, the responsibility for the defense of information systems. Four strategic objective characterize the French strategy:

- Becoming a cyber defense world power;
- Safeguarding France's ability to make decisions by means of the protection of information related to its sovereignty;
- Strengthening the cybersecurity of critical national infrastructure;
- Ensuring security in cyberspace.

In order to reach these objectives, seven areas of action have been identified by the French strategy:

- Effectively anticipate and analyze the environment in order to make appropriate decisions. Monitor the latest technology developments in order to understand and even anticipate the actions of public or private actors.
- Detect and block attacks, and alert and support potential victims. France is developing detection capability for attacks on information systems deployed within the ministry networks. These will enable the personnel to be alerted, assess the nature of attacks and create countermeasures. ANSSI has been equipped with an operations room to meet the challenges.
- Enhance and perpetuate French scientific, technical, industrial and human capabilities in order to maintain independence. Driving forward research into cryptology, formal methods and other security-related areas and creating cyber defense research centers in collaboration with industrial partners. Strategic investment funds will be provided by the state in order to promote the strengthening of industry.
- Protect the information systems of the nation and of the critical infrastructure to ensure better national resilience. The French strategy on security products and components has been redefined in order to take account of France re-joining NATO integrated command. Robust authentication systems will be integrated in the ministerial networks having a significant impact on the level of security. A public-private partnership will be set up in order to enhance the security of information systems of operators of critical infrastructure. The operators will benefit from the information gathered by the state on threat analysis

[2]Decree No. 2009-834 of 7 July 2009 creating the French Network and Information Security Agency* (ANSSI).

and the state will be able to ensure the appropriate level of protection of the infrastructure that is crucial to keep the country running properly.

- Adapt French legislation to incorporate technological developments and new practices: enact new rules to protect information systems and alert government authorities in case of incidents regarding operators of electronic communications. Enforcement of the General Security Framework in order to raise the protection level of the information systems of the public authorities.

- Develop international collaboration initiatives in the areas of information systems security, cyber defense and the fight against cyber crime in order to better protect national information systems and promote the sharing of essential data (information on vulnerabilities, services, threats) by establishing a wide network of foreign partners.

- Communicate, inform and raise understanding by the French population of the extent of the challenges related to information systems security, and ensure the awareness and motivation of individuals and organizations. ANSSI will conduct appropriate communication campaigns targeting the general public and companies.

USA. In May 2009, President Obama declared his intention to make cybersecurity a priority for his administration. This brought about the publication of a document entitled "Cybersecurity Policy Review" (CPR). In particular this document identifies 10 short-term actions:

1. Appointment of a cybersecurity policy official responsible for coordinating the nation's cybersecurity policies and activities.

2. Preparation for the president's approval of an updated national strategy to secure the information and communications infrastructure.

3. Designation of cybersecurity as one of the president's key management priorities and establishment of performance metrics.

4. Designation of a privacy and civil liberties official to the NSC cybersecurity directorate.

5. Conducting interagency-cleared legal analyses of priority cybersecurity related issues.

6. Initiating a national awareness and education campaign to promote cybersecurity.

7. Development of an international cybersecurity policy framework and the strengthening of international partnerships.

8. Preparation of a cybersecurity incident response plan and initiation of a dialogue to enhance public-private partnerships.

9. Development of a framework for research and development strategies which focuses on game-changing technologies that have the potential to enhance the security, reliability, resilience, and trustworthiness of digital infrastructure.

10. Building a cybersecurity based identity management vision and strategy, and leveraging privacy-enhancing technologies for the nation.

The achievement of such objectives must respect the Comprehensive National Cybersecurity Initiative (CNCI) [9], launched by President George W. Bush in January 2008, which consists of a set of initiatives aimed at strengthening US cybersecurity. President Obama established that CNCI had to be included and extended in the updated strategy of national cybersecurity, and that it would play a key role in the realization of the 10 objectives. During the 14 months following the issue of CPR many of the objectives were achieved:

- President Obama appointed a cybersecurity coordinator at the head of the Cyber Security Directorate created within the National Security Staff (NSS). This coordinator works closely with the Office of Management and Budget and the Office of Science and Technology Policy.

- The Cybersecurity Directorate started the development of an updated cybersecurity strategy that expands and implements the strategy envisaged by CPR and CNCI.

- A continuous and real-time monitoring of federal networks has been introduced, thus enabling faster detection of vulnerabilities and more effective infrastructure protection.

- According to CPR, a privacy and civil liberties official has been designated within the NSS.

- The National Initiative for Cybersecurity Education (NICE) has been released to improve the recruitment, training, and retention of cybersecurity professionals, to raise public awareness in cybersecurity, and to enhance cybersecurity education by expanding the education programme of CNCI.

- The United States is working to strengthen cooperation and dialogue with international partners. In cooperation with allied countries, the United States has taken on a leading role in international organizations, such as the United Nations, to make cybersecurity an international priority.

- The National Cyber Incident Response Plan (NCIRP) has been developed to enable a coordinated national response to cyber incidents.

- The administration has developed a research and development strategy based on three main themes: moving targets (systems that change continuously to increase their complexity, thus limiting attackers and exposition to vulnerabilities), tailored trustworthy spaces (trusted environments that allows the definition of tailored requirements) and cyber economic incentives (incentives to adopt appropriate cybersecurity solutions for individuals and organizations).

- A draft "National Strategy for Trusted Identities in Cyberspace" (NSTIC), aimed at reducing cybersecurity vulnerabilities through the use of trusted digital identities, has been released.

With regard to the US roadmap, in February 2013, President Obama issued an executive order to further improve the management of critical infrastructure cybersecurity. The aim of this executive order is to establish a new partnership with the critical infrastructure owners and operators in order to increase cybersecurity information sharing and collaboratively develop risk-based standards.

Information sharing on cybersecurity issues, such as suffered and foiled attacks, threats and vulnerabilities, between the public and private sector is the key factor in the improvement process envisaged by the executive order. The US government is responsible for improving such exchange of information in terms of volume, timeliness and quality of information shared with the private sector, thus enabling entities of the private sector to better protect themselves against cyberthreats. As a result of the executive order the Secretary of Homeland Security, the Attorney General[3] and the Director of National Intelligence will be responsible for ensuring the timely production of specific unclassified reports of cyberthreats to the US homeland. Moreover, classified reports will be delivered to authorized critical infrastructure entities. The Secretary of Homeland Security and the Attorney General, in coordination with the Director of National Intelligence, will be also responsible for setting up a system to track the production, dissemination and disposition of the reports. The aim is to maximize the utility of information sharing related to cyberthreats and attacks.

The executive order also addresses the protection of privacy and civil liberties. Important roles in this context are covered by the Chief Privacy Officer and by the Officer for Civil Rights and Civil Liberties (of the Department of Homeland Security). They are responsible for assessing the privacy and civil

[3]In the federal government of the United States, the Attorney General is a member of the Cabinet and as head of the Department of Justice is the top law enforcement officer and lawyer for the government (Wikipedia).

liberties risks of the functions performed by the Department of Homeland Security and for identifying and report ways to minimize such risks in a publicly available report to be released within one year from the issue of the executive order. In the production of the report the Chief Privacy Officer and the Officer for Civil Rights and Civil Liberties will consult the Privacy and Civil Liberties Oversight Board and the OMB.

The executive order issued by President Obama also envisages the creation of a Cybersecurity Framework aimed at reducing cyber risks to critical infrastructure. The Secretary of Commerce will direct the Director of the National Institute of Standard and Technology in the development of the framework. The Cybersecurity Framework will include a collection of standards and procedures to align policy, business and technological approaches to better address cyber risks. The framework will also include as much as possible industry best practices and will be available in final version by February 2014. The Secretary of Homeland Security will support the adoption of the framework by the owners and operators of the critical infrastructure and other interested entities.

1.4 Basic Security Issues of Critical Infrastructures

The critical infrastructure of every country, ranging from oil pipelines to the electricity grids, from gas to water networks, from transportation to financial/banking systems, to the public administration, are increasingly electronically managed. The progressive introduction of network management, monitoring and control systems, as well as the interdependence that has arisen, has certainly improved the performance level of such infrastructure, but it has also allowed access to cyber criminals, with consequent cyber attacks and the increased risk of a domino effect. Therefore, the scenario has become more and more complex in recent years, as the introduction of advanced technologies added new sources of potential risk alongside the traditional threats. An effective infrastructure protection includes threats identification, vulnerability reduction and attack source or damage origin identification. This activity aims at service downtime minimization and damage limitation.

Usually a cyberattack is launched to paralyze the critical infrastructure activities or to purloin its information assets. It is important to evaluate the possible attack targets to assess the consequences, also in terms of time required to restore normal behavior (resilience). The cyberthreat is an important challenge for the national economic system, both because it involves the digital domain and because of its transnational nature and, therefore, for the potential effects that it can produce. It is clear that when the attacks targets are critical infrastructure and warning systems, the consequences for the entire society could be disastrous. In light of this consideration and of

the awareness that it is a continuously changing environment, it is urgent to intervene, at the national level and beyond, against all cyber crime forms, which represent a growing threat to critical infrastructure, society, business and citizens.

In this context, a further prominent aspect is represented by the interaction and interdependencies among critical infrastructure. Understanding and analyzing interdependencies is of utmost importance, since they might be a source of threat to systems and contribute to risk uncertainty with an ensuing amplification in speed and size of loss following the occurrence of failures.

An interdependency is a bidirectional relationship between two infrastructure through which the state of each infrastructure influences or is correlated to the state of the other [149]. Infrastructure interdependencies can be characterized according to various dimensions in order to facilitate their identification, understanding and analysis. Six dimensions have been identified in [149], which include: a) the couplings among the infrastructure and their effects on their response behavior (loose or tight, inflexible or adaptive), b) the state of operation (normal, stressed, emergency, repair), c) the type of failure affecting the infrastructure (common-cause, cascading, escalating), and d) the types of interdependencies. Focussing on the type of interdependencies, four classes have been distinguished in [149]: physical, cyber, geographic, and logical.

- Physical interdependencies, which arise from physical links or connections among elements of the infrastructure. In this context disruptions and perturbations in one infrastructure can propagate to other infrastructure.

- Cyber interdependencies, which occur when the state of an infrastructure depends on information transmitted through the information infrastructure. Such interdependencies result from the increased use of computer-based information systems, such as SCADA systems, to support control, monitoring and management activities.

- Geographic interdependencies, which exist between two infrastructure when a local environmental event can create state changes in both of them. This generally occurs when the elements of the infrastructure are in close spatial proximity.

- Logical interdependencies, which gather all interdependencies that are not physical, cyber or geographic, caused for example by regulatory, legal or policy constraints.

The four types of interdependencies are not mutually exclusive, although each of them has its own characteristics. Other classifications have also been proposed in the literature [119, 110, 145]. For example, the classification proposed in [119] looks at both the involved systems and their potential interconnections that are characterized by two key factors: i) the character type of the

link, that identifies which elements of the systems are affected: physical, logical, human/organizational; ii) the layer of interaction: structural, functional, behavioral.

As discussed in [54], the focus of much of the debate and research about critical infrastructure is around the more classical power infrastructure. However critical information infrastructure (CII) can be an important source of interdependencies, for their role at the heart of the other infrastructure and inherent internal complexity due to the many involved sub-sectors (including information systems and network protection, instrumentation and control systems (SCADA), fixed communications and mobile communications). In fact, according to its definition by the OECD, a CII consists of those information and communication technology facilities, networks, services and assets which, if disrupted or destroyed, either (1) have a serious impact on the health, safety, security or economic wellbeing of citizens or the effective functioning of governments, or (2) causes the functioning of a critical infrastructure which it supports to be seriously disrupted. Additional complexity in terms of interdependencies arises when categorizing information infrastructure in the two dimensions: 1) service oriented view and 2) information and data. The former concerns the delivery of services to the end users and the latter the provision of information and data to ensure the correct and regular functioning of the services.

Although integrating critical infrastructure and synergically using them undoubtedly provide valuable benefits in terms of efficiency, quality of service and cost reduction, interdependencies increase the vulnerability of the corresponding infrastructure as they give rise to multiple error propagation channels from one infrastructure to another that increase their exposure to accidental as well as to malicious threats. Consequently, the impact of infrastructure components failures and their severity can be exacerbated and are generally much higher and more difficult to foresee, compared to failures confined to single infrastructure. As reported in [146], typically blackouts can be caused by the outage of a single transmission (or generation) element, which is not properly managed by automatic control actions or operator intervention, so gradually leading to cascading outages and eventually to the collapse of the entire system. Examples of cascading effects from infrastructure interdependencies leading to catastrophic events, across multiple infrastructure possibly spanning wide geographical areas, are reported in [146, 42].

CHAPTER

2 Background on Threats, Vulnerabilities and Accidental Faults

The evolution of ICT made it feasible and convenient to control CIs remotely (e.g., over the Internet). Therefore, industries and governments have been progressively adopting IT systems to consolidate the operation of CIs. As a result, CIs and IT systems have converged. This raises security concerns (and threats), because two previously isolated worlds, the Internet and CI systems, are now interconnected. Interestingly, the Internet is itself an underlying, critical asset of modern CIs, because their controlling systems are often distributed over remote, Internet-connected locations. This strong correlation between CIs and their IT comes at the cost of increased complexity and, as a consequence, increased risks of accidental faults. It is important to note here that the correct management of non-malicious faults is as important as the management of security risks. These two aspects, in fact, are strongly correlated, and should always be considered together when planning for the protection of CIs. In this chapter we report several instances of accidental faults, and their consequences, that have been experienced in real critical infrastructure over the last decades, focusing especially on cyber aspects. The remainder of this chapter is organized as follows: first the main threats (Section 2.1) and the motivations behind them, the vulnerabilities (Section 2.2) that make some attacks feasible (Section 2.3), and the current remediation approaches (Section 2.4) are explored. Then, the chapter describes current accidental threats CIs can be exposed to.

2.1 Threats

Well-known threats such as malware or denial of service attacks, which have been impacting on the security of Internet-connected devices, have become threats for CIs as well. In addition, unlike other Internet-connected devices

e.g., personal computers, mobile devices, servers, CIs can take real actions that can ultimately impact on the physical environment. This clearly poses serious safety risks, with the possibility lost production, equipment damage, information theft, and even loss of human life. However, it seems that, contrarily to dire predictions, the actors behind the events portrayed in the news as "cyberattacks" are probing without causing deliberate damage, as described in Section 2.3.1. Something, however, may be changing on this point, as described in Section 2.3.2.

The remainder of this section explores the actors and their motivations for attacking CI.

2.1.1 Actors

One of the reasons that make CI security a complex and pressing matter is that there exist several actors who pose a threat to CIs, possibly more than in traditional IT systems.

Below is a non-exhaustive, broad enumeration of classes of actors, in order of importance:

Nation states are an important new set of actors in the landscape of cyberattacks against CIs. Their importance derives from the fact that CIs are relevant targets in modern cyberwarfare. As described in Section 2.3.1 and 2.3.2, attacks against CIs, or high-value targets, can be politically or economically motivated. In this, nation states play an important role. An extension of this category of actors includes those attackers that are sponsored by nation states i.e., an external subject paid or supported by nationstate offices to compromise another nation's CIs.

Non-state organized threat groups usually labeled as "cyber-terrorists", are also a worrying threat. The potential for asymmetric warfare derives from the ease of attacking CIs through cyberwarfare means. The attack described in Section 2.3.2, while not attributable to a terrorist group with any certainty, is an example of what an organized, terror attack against an infrastructure could look like.

Hacktivists have been gaining lot of attention recently. The term hacktivist refers to an attacker, in many cases with limited technical skills, who relies on ready-to-use attack kits and services, or even third-party botnets, to cause damage to a system e.g., denial of service, defacement as a means of protest. Protests are often politically motivated. Although with different motivations than nation states, hacktivists also see CIs as an appealing target in their campaigns.

Business-oriented attackers refer to a more traditional category of attackers (i.e., those who would launch a denial of service attack against

a competitor's website). In the landscape of cyber-physical systems, business-motivated attackers are interested in performing abusive activities against competitor-controlled CIs in order to cause concrete damage and gain business advantages.

Casual attackers such as script kiddies, who in the past would launch a publicly available exploit against a random website for no real motivation, gain much more importance if considered in the context of CIs. Although casual attackers normally have little or no technical skills, launching attacks against an Internet-facing CI (e.g., maybe found through services such as SHODAN[7]) can cause serious damage, much higher than in the case of simple IT system (e.g., a website).

It is important to note that hacktivists, business-oriented attackers and casual attackers could also be tolerated by nationstates as allies in a low-intensity warfare against an opponent nation.

2.1.2 Motivation and goals

The aforementioned actors are driven by two broad categories of motivations.

Political, strategical warfare. From the (scarce) amount of reliable information that circulates regarding attacks against CIs, it can be concluded that most of the attacks have warfare or strategical motivations behind them. The most known and recent cases are Stuxnet (described in Section 2.3.1), Aramco (described in Section 2.3.2) and Duqu. Another type of attack had the goal of exfiltrating intelligence or secret information. For now, it cannot be stated with certainty what the final use of such information is; however it can be argued that the main motivations are of a political nature. Actors such as nation states and hacktivists fall in this category.

Financial. Business-oriented and nation states actors are also driven by economic reasons. This category of motivation also existed before CIs became an appealing and sensitive target. However, hitting valuable CI may result in a substantially higher financial impact than hitting a traditional IT systems.

2.2 Vulnerabilities

CIs are composed of critical components. Each component has to be analyzed from the point of view of possible risks and security aspects. Components meant to operate in safety-critical environments are usually designed to be fail-safe, but security vulnerabilities could be exploited by an attacker to thwart the fail-safe mechanisms.

Today's CIs have different classes of vulnerabilities. As explained in Section 2.2.1, besides the logical and design vulnerabilities due to the increased connectivity and open design of these network infrastructure, the use of commercial off-the-shelf (COTS) components, which were not built with security in mind, increases the attack surface. In addition, as described in Section 2.2.2, the application layer also exposes vulnerabilities and, more importantly, lacks security features.

2.2.1 Network and infrastructure layer

As CIs are controlled by several installations of interconnected networks systems, the infrastructure layer is a prominent issue in ensuring their cybersecurity. Due to several factors explained in the following, such layers are usually particularly vulnerable.

Increased connectivity. Until recently, control systems were electronically isolated (i.e., "air gapped") from all other networks. Therefore, industrial security was ensured mostly by enforcing physical security, so that attackers would not be able to access them [124, 154]. Nowadays, the growing demands of industry for increased connectivity between factory floors and corporate networks have altered the simple, isolated control network into a member of a complex inter-network such as the Internet.

Open design and use of COTS components. Control systems used to be based on proprietary solutions, which provided a weak form of security by obscurity. Over the years, CI operators, as well as the automation industry in general, have moved away from proprietary standards for SCADA communication protocols towards open international standards such as Ethernet or TCP/IP, and COTS hardware and software components.

The first effect of this is that the previously held belief that it would be difficult for attackers to gain access to information about control systems networks-the common defense, "the hackers don't know our systems"-is no longer true [162]. It should be noted, however, that relying on proprietary protocols and systems to ensure a form of protection was a rather misconceived notion since the beginning, as such obscure protocols and devices provided usually very little built-in security.

The migration of systems such as SCADA to TCP/IP facilitates interconnections between SCADA networks and corporate ICT infrastructure [134]. Conversion to standard protocols often happens by encapsulating established serial-line based protocols onto a TCP packet. Many of these protocols abandon any strict master/slave relationships traditionally seen in SCADA networks, and devices designed for these networks often provide additional application layer interfaces beyond the SCADA messaging protocol. These can include web-interface capability which, when coupled with the integration to

the corporate network, allows for convenient gathering of production information for higher-level management. Of course, inclusion of these services makes any devices on the SCADA network supporting them vulnerable to popular application layer and TCP/IP-based attacks. Even if it can be convenient and cost-effective from an operational point of view, this trend raises serious security issues. In fact, previously unprotected SCADA protocols can be severely exposed by attacks on the TCP/IP carrier. Also, attacks on a corporate network could then tunnel into a SCADA system and seriously threaten the controlled process.

COTS components also allow for cost saving and reduced design time, but they are not designed with security, or safety, in mind, thus offering a tempting target for attack. Being broadly installed, the knowledge base of readily available attacks for such system is surely wider.

Wireless sensor networks. CI protection requires monitoring mechanisms to detect failures and attacks as early as possible. Since many CIs have a large geographical span, CI protection needs monitoring mechanisms that scale well. In this context wireless sensor networks (WSNs) arise naturally as a potential solution. For example, the application of sensors to monitor the structural health of transmission lines is an important way to reduce power system vulnerability [112]. WSNs can be relatively easily deployed on a large scale, and as they are normally built from low-cost devices, they can provide the monitoring service in a cost-efficient manner since they do not require additional infrastructure. In addition, the distributed nature of a WSN increases the survivability of the network in critical situations, because a large-scale WSN is much less likely to be affected in its entirety by failures or attacks. In very critical situations WSNs may still provide sufficient information about the CI to help the operator prevent further damage and begin the recovery process.

It must be clear, however, that the usefulness of WSNs for CI protection is primarily determined by the dependability of the WSN itself [59]. A WSN that fails to report a faulty condition prevents the CI operator from carrying out the appropriate maintenance that may fix the problem before its consequences affect the CI. On the other hand, a WSN reporting too many false positives will lead to time and resources being wasted and endanger the benefit of using a WSN.

Security in WSNs is a more difficult long-term problem than in traditional distributed systems [61], for various reasons. First of all, WSNs are generally installed in unattended, possibly hostile, environments, which may be difficult to protect physically, especially in geographically large deployments or where conditions are unfavorable for humans. In addition, it may be economically unfeasible to make all of the nodes tamper-resistant. Thus, it cannot be excluded that an adversary may capture and compromise nodes, thus altering their behavior and potentially injecting fake messages into the network.

2.2.2 SCADA/ICS and embedded devices

One of the main vulnerabilities of SCADA and industrial control systems is the lack of security features in the protocols they use. As already mentioned, the SCADA system evolved from being segregated physically and logically from other networks, to being interconnected and migrated to standard and open protocols [70]. This of course changed the threat landscape and exposed vulnerabilities such as the ones described below.

Lack of authentication and authorization facilities. The absence of proper authentication and authorization schemes can let an unauthorized intruder create false control messages, thus causing major concerns for the correct operation of the system and possibly leading to dramatic consequences for public safety and health [101]. This situation demonstrates that SCADA systems need to support key security properties such as authentication, authorization, confidentiality, integrity, availability and non-repudiation.

In 2010, the United States Industrial Control Systems Cyber Emergency Response Team (ICS-CERT) published an alert stating that several security researchers successfully employed the SHODAN search engine[7] to identify Internet-facing SCADA systems that used insecure authentication and authorization mechanisms [19]. The alert did not only demonstrate that potentially unprotected control systems are readily accessible from the Internet but also that, by using tools such as SHODAN, the effort and resources required to identify them have been greatly reduced.

Lack of protocol protection mechanisms. Another source of vulnerability for control systems is represented by software bugs in SCADA devices. Even an input validation bug could leave a CI vulnerable to attack. Different fuzz-testing experiments demonstrated how deliberately malformed input data could be used to successfully crash SCADA equipment [79, 156]. For this reason, securing SCADA systems requires extensive testing for software vulnerabilities. This kind of testing, however, can be troublesome: vulnerabilities are usually not well understood by SCADA developers, whereas external security experts lack the necessary SCADA knowledge and resources to run thorough tests.

Security of the underlying embedded devices. The security of the different embedded devices that compose a control system must also be considered. Securing single devices can further enhance the overall system security. In fact, it can help to support particular security requirements that real-time applications alone could never be capable of addressing [101].
Furthermore, embedded devices may expose specific vulnerabilities that, if unresolved, could be exploited to compromise the whole system. This is a particularly relevant aspect, because embedded devices security is generally

overlooked since such devices are not managed like regular computers. For instance, as demonstrated in [144], an unprotected firmware upgrade utility in SCADA field devices could be used by an attacker to remotely install a malicious firmware. In this way, the attacker would have full control over the device functionalities and its interactions with the rest of the system, thus dramatically threatening the whole CI.

2.2.3 Applications

At the application layer, several vulnerabilities can afflict CIs, in particular due to their distributed nature.

Consider, for instance, where the application is deployed and where it is executed. In a distributed system an application can be deployed as a unique piece of software run by a unique trusted platform, or it can be deployed on different platforms as separate pieces of code that can communicate with one another. This leads to possible security violations. The exchange of messages among pieces of code deployed on different components of the system has to be regulated in such a way that no information is lost or leaked. Hence the channels must be protected against all possible attacks (e.g., man-in-the-middle, compromised sensor, etc.) in order to avoid relying on compromised information. Furthermore, each component cannot control each piece of code run by another component of the system. These two aspects lead to the necessity of having a trust model for managing the trust relations among the components of the CI. Distributed environments do not guarantee that the provided information is genuine. A trust management infrastructure is needed in order to provide some security guarantees.

Other security threats are related to the interdependencies among the components of a CI. Indeed, interdependencies can be exploited to attack the system in a coordinated way by several attackers located in different strategic points of the infrastructure. In fact, it is possible to distinguish and classify attacks into two main classes: attacks led in isolation by one component and attacks led by several components that cooperate in order to violate the system.

2.2.4 Business layer

Critical infrastructure and business-core applications can be attacked by means of many different vectors. Expanding on the previous analysis, it should be kept in mind that CI is, at an operative level, ordinary business with all the typical weaknesses that this implies. After all, Stuxnet, the malware that compromised hundreds of computers in Iranian nuclear factories in 2009-2010, was spread through a USB key an employee plugged in his workstation. Employees are indeed a well known point-of-failure for the security of an infrastructure, whatever its nature is.

Social engineering and similar attacks. A common technique used by attackers to infiltrate networks and operative environments is social engineering. Social engineering refers to any technique used to trick the user into giving away information or performing some apparently innocuous and legitimate actions that, instead, compromise the system or the network. This category of attacks is very broad, and operatively can be implemented in many different ways: from an e-mail asking for credentials or attaching a suspicious file, to a phone call during which the attacker pretends he or she is, for example, calling to provide some form of technical support. These attacks move the focus from the system and infrastructure vulnerabilities to the human who operates them. Somewhat surprisingly, while these attacks are often not very technical, a technical attack infrastructure exists to operate them. In the cybercrime black market, cyber crooks develop and trade platforms and frameworks to develop and deploy social engineering attacks, e.g., via e-mail. Citadel, for instance, is a very popular social platform that allows attackers to build their own attacks, discuss them and give feedback to their peers and the platform developers. In this way, a diversity of socially engineered attacks can be forged and deployed without any particular effort, distributing malware and stealing credentials (that may as well belong to a user account within a CI).

Spear phishing and targeted attacks. These types of attack can be directed both against the general population, when the attacker is interested in accumulating, for example, credit card PINs, and a particular user; in this case the attack is targeted against a particular person or organization for which the victim acts as a proxy for the attack. These attacks are usually more elaborate and well-thought out than non-targeted ones. They often require the attacker to have some pre-existent knowledge about the victim or the infrastructure she operates in. In particular, the term spear phishing refers to attacks where the attacker impersonates or spoofs the e-mail address or the contact information of a CI employee, often somewhere high up in the hierarchy, and uses it to trick email recipients to perform the compromising action. Imagine the situation when the head of the department or the company CEO sends an e-mail to employees, and tells them to visualize the attachment: who wouldn't open it?

This type of social engineering is clearly more targeted than more common phishing techniques. Other types of targeted attacks, however, can be much more technical. For example, exploitation of 0-day vulnerabilities known to affect some CI systems often indicates that the attack was explicitly directed against that infrastructure. These attacks are very difficult to detect and mitigate, because of their very nature. Their impact is also hard to assess. What data did the attacker steal? What else did he compromise after the first successful attack? Did the attacker install a silent malware the organization should detect and neutralize? The last case in particular is extremely tricky.

Once the attacker has gained access to the compromised system, he can install a silent, hard to detect software that monitors, audits, or simply waits for the attacker's commands before performing potentially disastrous actions. These threats are called Advanced Persistent Threats (APTs). APTs are particularly harmful because, even if discovered, it is very hard to assess the initial moment in which the threat was injected and what is its actual impact at the organizational level.

Cascading failures in interconnected systems. A cascading failure is a sequence of dependent failures that successively weaken a system. Due to their structure and interdependencies among components, CIs (i.e. power systems) are particularly subject to such type of failure. A common thing to see during a cascade failure is a walking failure, where sections go down, causing the next section to fail, after which the first section comes back up. This ripple can make several passes through the same sections or connecting nodes before stability is restored. The threat of cascading failures across critical infrastructure has been identified as a key challenge for governments. Cascading failure is seen as potentially catastrophic, extremely difficult to predict and increasingly likely to happen. Privatization of some CIs and the consequent profit-driven management can only increase the risk of such failures to happen.

2.3 Attacks

Starting from the vulnerabilities described in the previous sections, a list of possible violations is obtained. In particular, this document focuses on security properties related to maintaining confidentiality, integrity and availability. Examples of such kind of property are:

- Authorization properties stating which actions are allowed.
- Access-control properties that regulate the access to some resources. The decision can be taken according to the role of the user that requires the access, or to the usage of the required resource. Access control policies can also list the set of proscribed executions by stating the unacceptable operations.
- Bounded availability properties may be characterized as safety ones. An example is "one principal cannot be denied the use of a resource for more then D steps as a results of the use of that resource by other principals". Here, the defining set of partial executions contains intervals that exceed D steps and during which a principal is denied use of a resource.
- Chinese Wall policies regulate the access to resources that are classified in two different domains. In particular a Chinese wall policy guarantees

> that if a user has access to information of one set, that user cannot have access to the information belonging to the other set. The Chinese Wall policy combines commercial discretion with legally enforceable mandatory controls. It is required in the operation of many financial services organizations and is, therefore, perhaps as significant to the financial world as Bell-LaPadula's policies [50] are to the military.

Referring to Bell-LaPadula's policies, the set of information flow properties is introduced. In published works there are many definitions of these kind of properties. The basic idea is that the flow of information from high level users to low level users can be forbidden in such a way that the activity of high level users is transparent with respect to low level users. In terms of critical infrastructure, in which several components cooperate one another, the information flow policy could consist of a regulation of the flow of information among different components in such a way that sensitive information is not disclosed or leaked by a possible attacker.

2.3.1 Case study: Stuxnet

W32.Stuxnet [88], also simply known as Stuxnet, is a malware used in 2009-2010 to implement a targeted attack. This attack gained a lot of attention, in both the media and research community. After three years there are still many obscure points but from what is known Stuxnet was crafted specifically to propagate into and compromise a Siemens-branded ICS network. Thanks to a 0-day vulnerability, a Windows rootkit, a PLC rootkit, and many other advanced evasion and replication techniques, Stuxnet managed to infect many ICS-managed facilities. The main explanation for the reaction of the media, industry, governments and researchers, is that Iran's nuclear plants were the most infected target.

The goal of Stuxnet was to modify the functioning of PLCs (thanks to the first PLC rootkit ever found) in order to alter the operation of the equipment, possibly sabotaging the entire facility thus causing serious damage in the physical world (e.g., explosions, radiation).

A recent report by Symantec [123] describes that earlier versions of this sophisticated cyber weapon contained other known versions of the malicious code that were reportedly unleashed by the US and Israel several years ago, in an attempt to sabotage Iran's nuclear program. This indicates that Stuxnet was active about two years before the main incident. It also implies that neither of the two campaigns of Stuxnet (in 2007 and 2009–2010) had a serious impact on Iran's nuclear facilities, the avowed main target of the attack. Even though Stuxnet basically failed, an important fact remains true: Stuxnet was created (by nation states offices, as some experts argue) with careful planning and several resources.

2.3.2 Case study: Aramco

On 16 August 2012, Symantec and Kaspersky Lab [6], followed by several other vendors and researchers, described a novel, modular computer worm, which was dubbed Shamoon. The malware was part of a string of cyber espionage and sabotage attacks in the Middle East area (along with the previously described Stuxnet, see Section 2.3.1). It is not notable for its spreading mechanisms which exploit shared drives and folders, but rather for its quite unique payload.

Once a system is infected, Shamoon gathers files from specific locations on the system, sends the collected information back to the attacker, and replaces the files and the master boot record of the system with an image cropped from a picture of an American flag in flames.

The self-styled Cutting Sword of Justice group claimed responsibility for using Shamoon against 30,000 Saudi Aramco workstations, causing the company to spend a week restoring their services. Surprisingly, the attack did not hit any of the production control computers and networks, and was limited to the office and administration systems.

2.4 Remediation and protection approaches

The complexity, heterogeneity, adaptability, and mobility of critical infrastructure impose novel challenges on the design of risk mitigation systems and security mechanisms. Indeed, structures evolve to improve the quality of the provided services as well as to manage possible threats caused by new methods of attack.

A fundamental task needed for protecting a system is the execution of vulnerability assessments. This process can help to identify, quantify and rank the vulnerabilities of a system and to implement the security controls required to mitigate such vulnerabilities. While this operation is well-suited to traditional information systems, it can result unsatisfactory and limited in scope for CIs. In fact, while the downtime caused by vulnerability assessment may be acceptable for traditional systems, it becomes unacceptable for CIs because it risks disrupting controlled processes and damaging expensive equipment [168]. Furthermore, when vulnerabilities are identified and resolved, patching CI components is problematic for both the availability requirements and the large-scale nature of the systems [134]. The presented issues highlight the extreme need to design and develop CI systems with particular attention to security properties. Researchers have been developing testbeds, composed of both physical and virtual devices, that can help to identify common vulnerabilities and to verify the effectiveness of different protection approaches, without impacting on the operation of real CIs [168, 102].

A possible solution to avoid unauthorized access is represented by network segregation. In the CI scenario, this technique consists in separating the

control systems networks from the corporate networks, which are usually connected to the Internet. In this way unauthorized access from employees and remote intruders can be prevented. While it can be effective for enhancing system security, complete physical segregation is not a viable and future-proof solution for modern CI systems. The large-scale and distributed nature of these systems makes it necessary to remotely access them for management, monitoring and control purposes, even from mobile devices [134]. Nonetheless, logical network segregation mechanisms have to be implemented in order to protect CIs from unauthorized access. Control networks must be isolated from corporate networks by using filtering security controls such as firewalls. Internal monitoring and administration traffic can be further separated from normal LAN traffic by using VLANs. This method ensures virtual isolation of users that access critical data from the rest of traffic [102]. Finally, only authorized and protected remote access must be allowed. This can be achieved by implementing Virtual Private Networks (VPNs) using, for instance, IPsec tunnels [19]. Obviously network segregation alone does not provide complete protection of CIs. For example, physical access to the control systems networks, which might be achieved through social engineering attacks, overcomes any network segregation protection and can seriously threaten the whole system. For this reason other security mechanisms have to be implemented to further protect CIs.

A great improvement can result from a security-focused redesign of the communication protocols. Designing secure protocols is a delicate, time-consuming and costly task, but it must be seriously taken into account because unprotected protocols represent a major threat to CIs. However, designing new protocols from scratch may not be a satisfying solution in the short-term, because the adoption of these protocols could lead to unacceptable downtime and incompatibility with legacy systems. For this reason, researchers have been focusing their efforts in designing security solutions that respect existing protocol specifications and standards. In particular, Chandia et al. [70] propose the adoption of unused function fields in standard SCADA protocols (Modbus and DNP3) to provide confidentiality and integrity. This approach enhances CI security without losing compatibility with legacy systems. Another solution is represented by transparent tunneling techniques. By using these techniques, existing protocols can be wrapped in secure communication tunnels that provide fundamental security properties such as authentication, integrity and confidentiality. Tunnels can be implemented as an independent software layer in existing field devices or within special-purpose embedded components acting as gateways.

To further protect CI systems, traffic monitoring and anomaly detection mechanisms should be implemented. As in traditional information systems, these techniques can help to identify the data transported on the network, to monitor the transactions between the different components and to prevent or detect attack attempts. These techniques can also enhance CIs from a

functional point of view, for example to optimize the performance of a plant by monitoring the process behavior [70]. The traffic monitoring and anomaly detection mechanisms can be implemented by intrusion prevention systems (IPS) that make it possible to reduce malicious activity. A fundamental task of IPS deployment is the learning phase, which is used to analyze and collect data about normal network activity. After this phase, an IPS is able to recognize malicious activity and to react accordingly.

In traditional IT networks, where traffic is user-generated by complex communication patterns, this task can be extremely complex. However, the uniformity and low volumes of traffic in typical CI systems simplify the learning task so that it becomes feasible [70]. Yet, for an IPS to properly work, deep knowledge of the systems and protocols vulnerabilities is required. As previously stated, vulnerability analysis for CI systems is a difficult and ongoing operation that requires security experts to acquire field-specific knowledge and resources. Available results have already been used to develop attack signatures for standard SCADA protocols (Modbus, DNP3 and ICCP). These attack signatures are integrated into most commercial intrusion prevention systems [144].

2.5 Accidental faults

Although there is a growing attention devoted to malicious acts targeting critical infrastructure, accidental faults remain an important source of failures that may affect both the physical and the cyber aspects of CIs.

Three types of failures are of particular interest when analyzing interdependent critical infrastructure:

- Cascading failures which occur when a failure in one infrastructure causes the failure of one or more component(s) in a second infrastructure;

- Escalating failures which occur when an existing failure in one infrastructure exacerbates an independent failure in another infrastructure, increasing its severity or the time for recovery and restoration from this failure;

- Common cause failures which occur when two or more infrastructure are affected simultaneously because of some common cause.

Of course, besides analyzing the types of failures, it is important to understand the different causes that might lead to the occurrence of such failures. Once the cause of the failure is known, proper measures can be taken at the system control level of the infrastructure so as to prevent future occurrence of the same fault or at least mitigate its effects on the system.

From the reports on incidents in a number of critical sectors, such as global financing, energy distribution, transportation, there is empirical evidence that the risks from interdependencies seem to have not been sufficiently addressed or estimated. Although it is rather obvious that there are infrastructure interdependencies e.g., telecommunication needs power, water needs electricity for pumping, and a power station needs water to start-up, it is necessary to determine to what level dependencies are a significant contributor to the risk of exposing the infrastructure to catastrophic consequences. Assessing the importance of interdependencies and, more in general, the uncertainties in infrastructure interactions is a challenge, mainly due to the complexity, heterogeneity and scale of the involved systems. Many initiatives have been undertaken to tackle this challenge and initial approaches have been developed, including qualitative and quantitative analysis of incident data, as well as modeling and simulation solutions.

It is important to note that traditional research in either security or dependability has been developed almost always under the assumption of nominal behavior on one of two dimensions: no attacks are assumed while tolerating accidental faults and no faults are assumed to occur while facing attacks.

There are some partial exceptions to this attitude. The most remarkable being represented by byzantine fault tolerance where no limits are assumed on the behavior of faulty units (thus including deliberate malicious ones) whereby some proposals include non-homogeneous faults mixing deliberate malicious faults and plain accidental ones. As described earlier, when dealing with interdependencies, critical infrastructure show escalating and cascading failures due to combinations of attacks and non malicious faults, thus making the relationships and interplay between such attacks and non malicious faults very important to understand and absolutely critical to control. Therefore, besides keeping pace with the continuously emerging new threats and attacks, a new fundamental challenge appears when aiming at protecting critical infrastructure which demands the development of a body of knowledge with an integrated vision of all the threats, deliberate and accidental, that may hit critical infrastructure.

2.5.1 Overview of accidental faults and countermeasures

According to the Technical Committee on Fault-Tolerant Computing of the IEEE Computer Society and the IFIP Working Group 10.4, Dependable Computing and Fault Tolerance, which recently systematized basic concepts on dependable computing in [45], a fault is the cause (adjudged or hypothesized) of an incorrect system state. An incorrect state, called error, turns into a failure when the service delivered by the computer system deviates from correct service and negatively affects its users and other external systems.

Faults can be classified according to several dimensions. For instance, we can distinguish hardware and software faults, or permanent and transient

faults. Moreover, faults can be either internal or external of a system. Looking at the phenomenological cause, faults can be classified as natural (caused by natural phenomena, such as deterioration, unexpected radiation or noise, bad environmental conditions, etc.) or man-made (as a result from a human action). Man-made faults can be further categorized into malicious and non-malicious faults. Malicious faults are deliberately introduced during development or use with the explicit intention of causing harm to the system. Non-malicious faults can be instead the result of a human mistake or a bad decision, and can be categorized as accidental faults if introduced inadvertently, or incompetence faults if caused by the lack of professional competence. Clearly, all natural faults are accidental, being non-deliberate and not caused by humans.

In spite of engineering efforts for avoiding the occurrence of accidental faults, it is unfortunately impossible to prevent their occurrence. As a matter of fact, no amount of rigorous development activities can assure that large and complex computer systems that include several networked hardware and software components, which is the case of CI, will not fail due to accidental faults such as aging equipment, corollary of a protective device, environmental or man-made faults. For this reason, achieving dependable CI requires a combination both of rigorous engineering, to prevent accidental faults and keep their occurrence within reasonable limits, and of additional means for mitigating the impact of accidental faults, with the aim of removing accidental faults from the system and avoiding that they lead to more serious cascading failures of the critical infrastructure as a whole. The strategies that can be adopted for mitigating accidental faults are grouped in:

- **Fault tolerance** which avoids service failures by automatically identifying faults and recovering from them, for instance by isolating a faulty component or by replacing it using redundant components.
- **Fault removal** which reduces the number and severity of faults both during development, through rigorous testing, inspections or formal verification, and during the use of the system, through preventive and corrective maintenance activities.
- **Fault forecasting** which analyzes the incidence and the consequences of faults, in order to provide a qualitative/quantitative evaluation of the system behavior in the presence of faults. This evaluation provides useful feedback for improving the design of the system, for instance, by suggesting where to use fault tolerance for improving dependability.

2.5.2 Accidental Faults in Critical Infrastructure

In the following, several instances of CI accidental faults and related failures, where appropriate, are reported with reference to the application domains

relevant to the TENACE project, namely the power grid domain, the transportation domain (air traffic control and railways), and the financial domain. The resulting survey is the sum of the direct experiences of TENACE partners in past or present research projects and industrial collaborations involving CIs.

The financial domain. A financial system comprises a complex landscape of actors, including stakeholders, regulatory agencies, financial service providers and the communication networks linking them. These systems are quintessential to the functioning of modern national economies, and they can be definitely considered a CI of society. The financial ecosystem has relied upon IT resources and digital communications since the birth of commercial computing solutions (dating back to the sixties). This decades-long experience in the usage and management of complex IT resources, and the huge amount of know-how built in this timeframe by financial institutions' IT departments, make these players more prepared to face the challenges offered by today's interconnected world. Nevertheless, the widespread usage of applications that ubiquitously interconnect users with their bank accounts, and a rush towards high speed financial transactions, are putting financial infrastructure at severe risk. This infrastructure, being heavily based on IT systems, is prone to several different risks like hardware, network, and power failures; data loss caused by inadequate backup facilities or policies; poorly trained/skilled IT staff that lack sufficient knowledge; over-dependence on IT outsourcing; poor IT management practices; inadequate facilities or investments in IT.

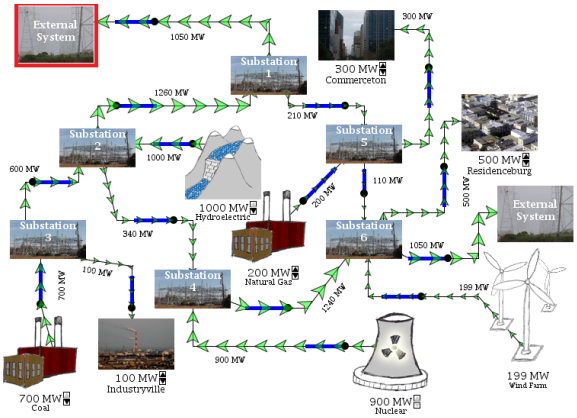
The 2012 IT incident at the Royal Bank of Scotland (RBS) [120] is a significant example of the unexpected problems that can arise when such risks are not adequately taken into account. In June 2012, the 16.7 million customers of three banks (RBS, NatWest and Ulster Bank) were left unable to access money in their accounts for four days. The incident was caused by a simple human error in managing a batch update job on a critical IBM mainframe that was used to manage more than 20 million transactions per day. The jobs were handled by a CA-7 scheduler that failed to update for three days in a row. During this period transactions were buffered without the possibility of being confirmed. A backlog of more than 100 million transactions were not paid in or out of bank accounts. This caused severe consequences for customers who were unable to access their accounts or use debit cards. The failure even condemned a man to spend weekend in jail as he was not able to pay his bail [100]. Banking experts said that the cost to RBS of dealing with IT problems, including extra staff costs as well as the money to reimburse customers, was likely to be between £50 million and £100 million. As a consequence of this accident, the UK's Financial Services Authority started to put pressure on UK banks to update their legacy systems to more modern and manageable technologies. RBS's crash, in fact, represented a stark wake-up call for global banks, many of which rely on decades-old IT systems

that become ever more complex as banks expanded through acquisition, often without fully integrating the systems they inherited [96].

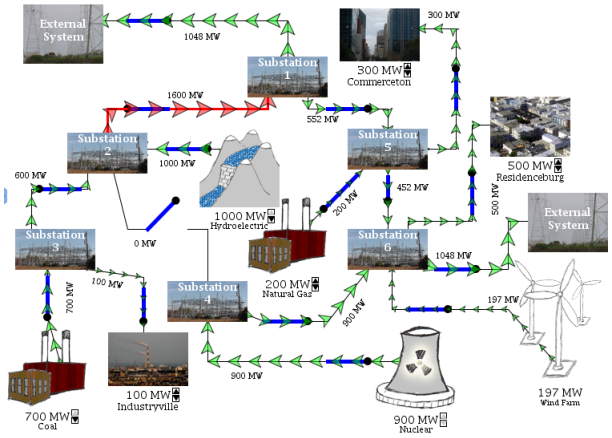
Even when IT infrastructure are maintained properly, severe incidents can still arise owing to the accidental interplay of independent systems, as shown by a recent study on the behavior of trading algorithms [105]. In algorithmic trading, high-performance machines run automated algorithms whose purpose is to track micro-fluctuations in financial markets and exploit them to perform quick market transactions (a single transactions can be prepared in less than 740ns [74]). Algorithmic trading represented, in 2012, more than 50% of US-based trade and more than 30% of EU-based trade. When performed correctly and at high speed, these transactions can easily lead to large incomes for the investor. However, such systems today work at such a high speed that human control and intervention over them is highly impractical. This problem, coupled with lack of mathematical models able to predict the collective behavior of these algorithms, gave rise to a group of competitive machines featuring crowds of predatory algorithms that cannot be fully controlled. The price is already being paid for this lack of control: in 2010, Wall Street suffered the so-called "Flash Crash", when the Dow Jones, S&P500 and Nasdaq indices suffered a close to 9% loss during the day. After five months of investigations, the SEC and the Commodity Futures Trading Commission presented a report that clearly described how high-frequency trading negotiations pushed a complex system, like the financial market is, in an unexpected direction [170]. Despite new regulations aimed at controlling the market, new cases arise frequently [115].

The power grid domain. A power grid is a system of producers and consumers of electricity [164]. It includes power generators, electricity users, switches that control the electricity, and Substations, power lines, and transformers that deliver the electricity. A community might have a generator to provide its power. The generator can vary its production as the usage of the customers changes. When the demand for energy is too high for the generator, the community buys electricity from another source. When the generator is making more electricity than the community is using, it can be sold to other communities.

A power grid is a system consisting of interconnected power generators, transmission systems and users that produce, transmit and consume electricity. Figure 2.1 shows a scenario which adopts coal, hydropower, natural gas, wind, and nuclear generators. Green arrows (Figure 2.1a) show the direction the power is moving, flowing out of the generators, through the Substations and into the communities. Bigger arrows indicate more power. Power from different generators is distributed to users in Commerceton, Industryville, and Residenceburg. Any power that is not used by the communities is sent to users in other systems (external systems, red rectangle in Figure 2.1a). An external



(a) Normal operation mode.



(b) A line transmission is overloaded.

Figure 2.1: Example of Power Grid Critical Infrastructure.

system can be a nation such as Italy that buys electricity from Switzerland. The sum of the power entering a Substation must be equal to the sum of the power exiting that Substation. For example, the coal generator sends 700 MW of power to Substation 3: Industryville receives 100 MW of that power, and 600 MW goes to Substation 2.

Major blackouts experienced in the last decade over several power grids worldwide have been caused by accidental faults. For example, the US-Canadian blackout of 14 August 2003, which affected approximately 50 million people in eight US states and two Canadian provinces, started with reactive power supply problems in the states of Indiana and Ohio, which were not promptly treated because of the lack of early warning caused by software problems. Another blackout occurred in continental Europe on 28 September 2003, resulting in a complete loss of power throughout Italy. It started with

the tripping of a major power line between Italy and Switzerland caused by a tree flashover (a natural cause), but the connection was not reestablished because the automatic breaker controls did not close the line (a man-made development mistake). The resulting cascading effect of lines tripping caused the collapse of the entire Italian system.

Figure 2.1b shows an example of what happened at that time. The transmission line between Substation 4 and Substation 2 got damaged and so the link between Substation 2 and Substation 1 was overloaded. In particular, this link was carrying about 1600 MW of power (red arrows). After the overload lasted some minutes the line transmission went down and the fault was propagated within the power grid until the blackout occurred.

The EU-FP6-027513 CRUTIAL project (2006-2009), involving the CNR research unit, addressed new networked ICT systems for the management of electric power grids. This project identified both accidental faults and malicious attacks threatening these CIs. Transient and permanent outages of electrical components have been considered those such as disconnection, overloads, reduction of production, increase or reduction of demand, voltage collapse. From the control infrastructure point of view, the failure model included omission, time, and byzantine failures, due to either accidental sources or attacks. In addition, the failure model of the interactions between the electrical grid and its control system has been accounted for. The impact of control system failures on the state of the grid has been analyzed, namely in terms of the topology and of the electrical parameters values such as voltage, active and reactive power, depending on the logical components affected by the failures, and on the type of the failures. Disruptions of the grid have been assumed to affect the control system by lessening its functionalities till complete failure in the extreme case the disruption is a total blackout.

The air traffic control domain. A civil Air Traffic Control (ATC) system is a typical software-intensive mission-critical system, that plays a key role in Air Traffic Management (ATM) [86]. It provides facilities and services to ground controllers and pilots for safely managing ground and en-route flight operations.

These systems need to meet stringent Quality of Service requirements in terms of availability in order to ensure, in their turn, the high availability of the whole infrastructure. To achieve this objective, software applications are require to distribute and replicate data e.g., flight routes on a number of nodes connected through a wide-area or local-area networks. Due to the nature of such systems, the replicas of a software application need to be strictly consistent in order to keep the same state in time, thus providing the same outputs to service requests. In such complex distributed systems, failures of individual components are frequent and have to be safely handled to ensure system survivability. Extensive testing during the design phase of the software

application cannot avoid the occurrence at operational time of faults that can lead to catastrophic consequences for the entire system.

A core ATC software component is the Flight Data Processing System (FDPS) which provides information such as flight routes, their current trajectory, airplane-related information, and meteorological data. The FDPS has been one of the main targets of experimentation conducted during recent research projects, such as the Italian PRIN DOTS-LCCI project[1] that involved a subset of TENACE partners, and the COSMIC Public-Private Regional Laboratory[2] in the Campania Region that involved the UNINA research unit and the SELEX-ES Finmeccanica company which develops the FDPS with other European partners.

The analysis of potential faults affecting an FDPS, reported in [143], points out the risk of failure of the overall system due to faults in individual software components, and encourages the adoption of fault mitigation strategies. The failure modes of FDPS software entities include process crashes, i.e., the entity stops providing service due to unexpected failure, passive hangs, i.e., the entity waits indefinitely for a resource which will never be released e.g., deadlocks, and active hangs, i.e., the entity indefinitely halts, but it keeps the system resources busy e.g., it is stuck in an infinite loop. In turn, the unavailability of air traffic control software can cause noticeable delays and service disruptions, and expose aircrafts to serious accidents. This was the case of the failure of a voice switching and control system, at the Los Angeles Air Route Traffic Control Center, which caused the loss of voice contact with airplanes, making it impossible to warn them of impeding dangers [95, 179]. This software failure affected 800 flights across United States, and in at least 5 cases airplanes came within the US Federal Aviation Administration's mandatory minimum separation distances, thus significantly increasing the risk of collisions.

The railways domain. Railway transportation is an important example of a domain where fault-tolerance and safety features play a vital role. Countless catastrophic failures have been experienced over time in railway systems due to infrastructure faults, speed, or erroneous signaling. Design approaches oriented to fault-tolerance and safety in the railway domain span of course many different fields of engineering. For the purposes of the project, computer-based railway control systems are focused on. There are essentially three classes of safety-critical railway control systems:

- Interlocking systems to manage train routes and signals in stations;
- Traffic management systems to manage train headways at the trackside level;

[1] `http://dots-lcci.prin.dis.unina.it/`
[2] `http://www.cosmiclab.it`

- Train control systems to manage train movement on board.

The evolution of computer-based systems brought about more complex failure modes, since each of the above mentioned systems is implemented as an increasingly complex computer platform, often in the form of heterogeneous, real-time embedded systems distributed over a large-scale infrastructure. This makes it very challenging to assure reliability, availability, maintainability, safety, and security (RAMSS) requirements.

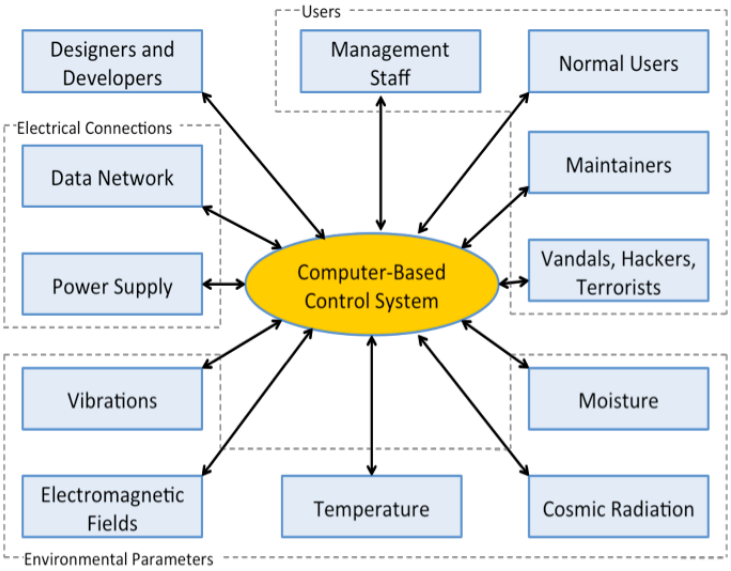


Figure 2.2: Dependability threats of a railway control system.

The threats to dependability of railway control systems can be summarized as shown in Figure 2.2. Faults can be introduced at the design and development stage, requiring a whole range of design approaches and best practices to be adopted before the system is put into the field. During normal system operation, faults can arise from a variety of sources. Human users, e.g., management, operators, maintainers, can cause system failures both unintentionally or intentionally. Infrastructure faults, including for instance power supply and data networks, also represent a threat for a railway control system. Finally, natural faults due to environmental conditions, e.g., temperature, moisture, cosmic radiation, etc. represent a variety of threats to take into account to ensure RAMSS requirements.

CHAPTER

3 Financial Systems

A financial system is defined by the set of institutions (markets and intermediaries) through which households invest their savings and corporations and governments obtain funding for their activities. Financial systems also exists to fuel the flow of funds from savers (lenders) to borrowers (investors or spenders) as part of a credit system, even to facilitate payments as part of a payment system. By providing a wide range of services that are the lifeline of the world economy, the financial system can be considered quintessential to the functioning of a modern nation's economy. Therefore this system can be definitely considered as a critical infrastructure of our society and, due to the continuously increasing penetration of the Internet in this infrastructure, it has to be protected from cyber attacks.

This chapter introduces the financial system (FS) with its main stakeholders and players and its requirements as infrastructure for the economy. Later on, it describes some protection strategies that aim at preventing attacks from having any effect, then, it concludes with an overview of the open problems in this field.

3.1 Description of the Critical Infrastructure

A financial system is largely an intangible asset that promotes economic growth by facilitating the transfer of funds from savers to borrowers and by facilitating payments. Even if the shape assumed by the financial system in each country can differ consistently, the financial system, with more or less efficiency, benefits the economy and in particular:

- Individuals. It provides the possibility of risk diversification of their investments, the liquidity of financial assets (ability to exchange a fi-

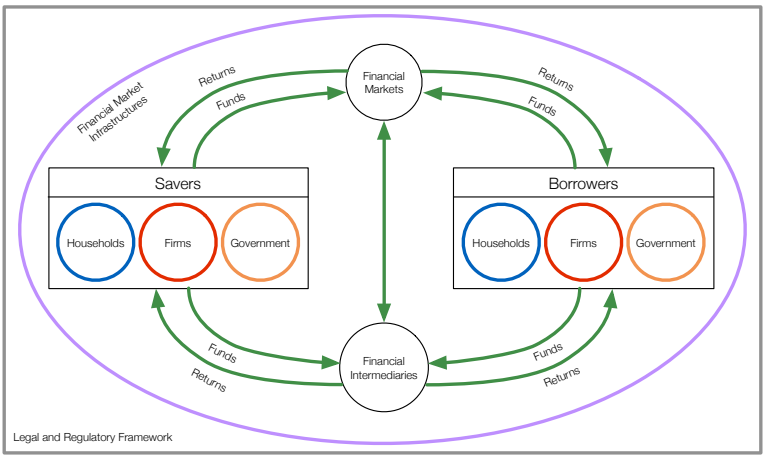


Figure 3.1: Schematic representation of a financial system.

nancial asset for cash at short notice and at low cost), and information in order to evaluate relevant data concerning the risk and return of various financial transactions (allowing for collection and communication of information savers and borrowers).

- Society. It facilitates the efficient allocation of scarce economic resources by providing an efficient credit system to transfer funds between savers and borrowers, which promotes economic growth, and it provides an efficient payment system to foster transactions.

The term system in financial system indicates a group of complex and closely linked institutions, agents, procedures, markets, transactions, claims and liabilities within a economy. In order to simplify we can consider the five main components that identify the financial system: i) financial instruments, ii) financial markets iii) financial intermediaries, iv) the legal and regulatory framework, v) financial market infrastructures (more detail are provided in next section).

Traditionally, schematic representations of financial systems, such as the one depicted in Figure 3.1, are designed to describe a given country's financial system. This is a simplified and limited view, since thanks to technological developments there is no longer any physical or technological impediment to the timely and relatively low cost flow of funds anywhere in the world. Furthermore, if we consider an area like the EU, we can talk about a financial system that transcends the national borders of each European country. By definition, and taking into account the role of the growth of the economy, the financial system can itself be considered a critical infrastructure; furthermore, as detailed in the following sections, the correct functioning of the FS relies on the correct functioning of all the elements involved and on the financial

infrastructures the system uses for its functioning (for instance the payment system and so on), nowadays more and more based on technology and IT systems. This renders it of utmost importance that the financial system is protected from cyber attacks.

3.1.1 Main Stakeholders and Players

As mentioned above the financial system can by analyzed through its principal component:

Financial instruments are all the products which are traded in a financial market/system; it refers to all financial assets, securities or other types of financial instruments according to the needs of investors and credit seekers. Modern financial markets are characterized by the presence of a variety of financial instruments, including securities (such as debt instruments and equities) and derivatives (such as futures, options and swaps).They indicate a claim on the settlement of principal down the road or payment of a regular amount by means of interest or dividend. Equity shares, debentures, bonds, derivatives etc are some examples.

Financial markets are the place in which financial assets/instruments are created and/or transferred. The purpose of a securities market is to bring together two groups of participants: those who have capital to invest (i.e. investors) and those who want to borrow that capital (e.g., firms and public bodies). Thus, as an alternative to borrowing money from an intermediary (e.g., a bank), firms and public bodies can raise funds directly from investors by issuing securities. Without financial markets, or underdeveloped ones, it is difficult for borrowers to find lenders. In this case, intermediaries assist or substitute the markets in this process (for instance, banks take deposits from investors and lend money from this pool of deposited money to people who need loans). For the securities market to work, it needs to be underpinned by arrangements and infrastructures for the handling of securities. This involves intermediaries, rules, procedures and processes, as well as organizations that provide trading, clearing and settlement services. It relies on institutions that provide securities accounts and related services. The securities trading landscape is changing, with the emergence of new markets and infrastructure. In addition to traditional exchanges, new recognized marketplaces (such as multilateral trading facilities) and other new trading venues such as electronic communication networks (ECNs), have been introduced. ECNs are order-driven, screen-based electronic markets for securities trading which bypass traditional market-makers. In addition, some investment firms are offering their customers sub-trading platforms for securities traded on several exchanges. Financial markets are generally categorized into money markets which handle short-term financial assets of less than a year and capital markets in which financial assets have a maturity period of more than a year.

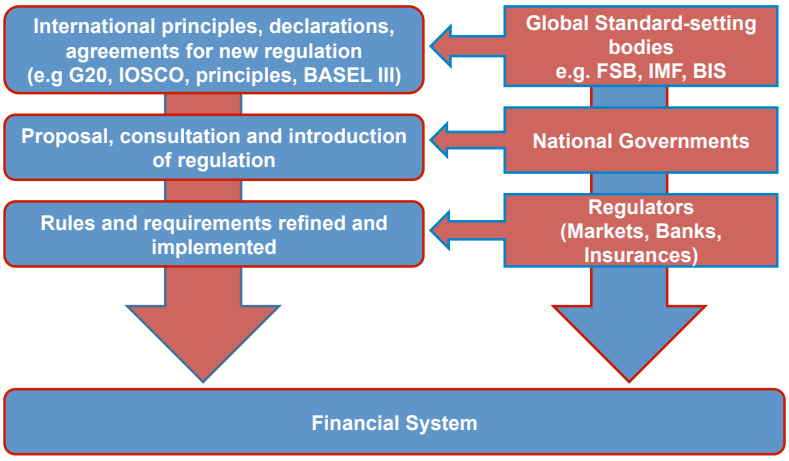


Figure 3.2: The legal and regulatory framework. Source: IOSCO 2013.

Financial intermediaries are those institutions which mobilize the savings of investors either directly or indirectly via financial markets (Figure 3.1), by making use of different financial instruments and using the services of numerous financial services providers; they are generally deeply regulated and supervised. A distinction is generally made between banks, non-banking financial companies (NBFCs), mutual funds and insurance organizations. Financial intermediaries or institutions provide a wide range of financial services, directly or through companies operating in the FS. The financial services sector offers a number of professional services like credit rating, venture capital financing, mutual funds, merchant banking, depository services, book building, etc.

The legal and regulatory framework plays a pivotal role in the operation of financial markets. The legal and regulatory framework ensure, among other things, a suitable structure for regulatory authority, appropriate powers to regulate and supervise markets and products, a credible regime for consumer protection, effective rules for transparent processes and sound governance and an efficient enforcement system. Furthermore, the legal system defines the shape assumed by each type of financial intermediary in each country. Since the global financial crisis began in 2007, mitigation of its causes has become a prominent task for global standard setters and national and regional governmental regulators. At the global level, governments established the Group of Twenty (G20) and called for regulatory reform of the entire financial sector to prevent the crisis from worsening and possibly reoccurring. Coordination of this large reform agenda was given to the newly created Financial Stability Board (FSB). Other standard setters, such as IOSCO and BIS, have been providing new global standards and principles. At the regional and national levels, governments are working on directives, laws and regulations to implement specific reforms (see Figure 3.2).

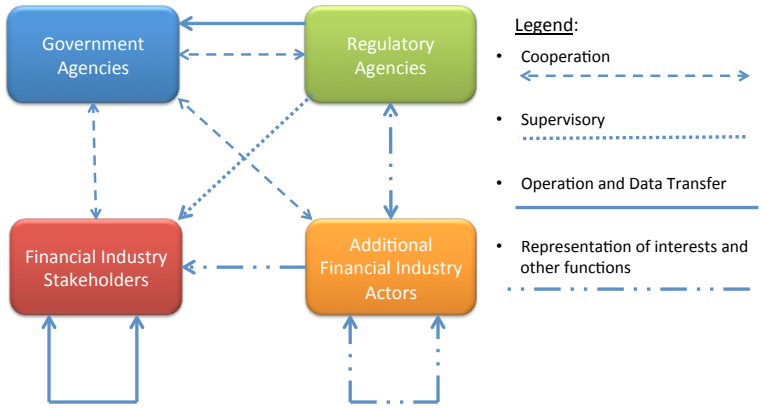


Figure 3.3: Simplified logical connections among players of the financial system.

Financial market infrastructures (FMIs) are defined as "a multilateral system among participating financial institutions, including the operator of the system, used for the purposes of recording, clearing, or settling payments, securities, derivatives, or other financial transactions"[1]. The principal objective of payment, clearing and settlement arrangements is to facilitate transactions between economic agents and to support the efficient allocation of resources in the economy. Market infrastructure for payments and financial instruments represents one of the transversal core components of the financial system, which connects all entities in the financial system. The complexity and importance of market infrastructure for the handling of payments and financial instruments has increased greatly in recent decades, owing not only to the tremendous increases observed in the volume and value of financial transactions, but also to the wealth of financial innovation and the advances seen in information and communication technologies. A payment system comprises three main elements or processes:

- payment instruments, which are a means of authorizing and submitting a payment (i.e., the means by which the payer gives its bank authorization for funds to be transferred or the means by which the payee gives its bank instructions for funds to be collected from the payer);
- processing (including clearing), which involves the payment instruction being exchanged between the banks (and accounts) concerned;
- a means of settlement for the relevant banks (i.e., the payer's bank has to compensate the payee's bank, either bilaterally or through accounts that the two banks hold with a third-party settlement agent).

[1]IOSCO, 2011

Main groups of stakeholders	Members	
Regulatory agencies	Financial supervisory authorities Tax and financial control office	
Government agencies	National Central Banks State treasuries	
FI stakeholders	**Money markets**	Banks
		Specialized Credit Institutions
		Co-operative Credit Institutions
		Savings Co-operatives
		Credit Co-operatives
		Financial Enterprises
	Capital markets	Investment Firms
		Investment Fund Managers
		Other Institutions
	Funds	Private Pension Funds
		Voluntary Pension Funds
		Health and Income-Replacement Funds
	Insurance companies	Proprietary Insurance Companies
		Mutual Insurance Companies
		Insurance brokers
		Insurance consultants

Table 3.1: Main Stakeholders of the Financial Industry.

A payment system also relies on institutions that provide payment accounts, instruments and services to customers (including consumers, businesses and public institutions) and on organizations that operate payment, clearing and settlement services (such as interbank funds transfer systems). Banks and other financial institutions are core actors in the market infrastructure. Banks are the principal providers of payment accounts, instruments and financial services to end users. In a relatively recent development, non-bank entities are now also entering the market, providing services at various stages in the initiation and processing of transactions. FMIs that facilitate the clearing, settlement, and recording of monetary and other financial transactions can strengthen the markets they serve and play a critical role in fostering financial stability. However, if not properly managed, they can pose significant risks to the financial system and be a potential source of contagion, particularly in periods of market stress. Interaction between the main components of financial system are hard to represent since each player may assume a different role in the system and operate with multiple counter-parties. A synthetic representation of this interaction is provided in Figure 3.3, taking into account the main type of logical interaction.

Finally, Table 3.1 summarizes the main stakeholders of the financial system stemming from this general model.

3.1.2 Requirements

Financial IT infrastructure is used to process, store and exchange critical and sensitive information, hence it is characterized by strict security requirements. Systems, networks, data and exchanged information should be protected against any type of malicious activity (such as interception, insertion of fake information, update, delete). The Financial IT infrastructure is a key critical infrastructure for financial operators and consequently needs to be dependable or trustworthy. The attributes of dependability/trustworthiness [159] refer to the degree of (quantifiable) user certainty that the system will operate as expected and that the system will not fail in normal use. Basically, the IT financial infrastructure has to satisfy the following dependability and security requirements and properties:

- Availability - the capacity to access systems, networks and critical data for the infrastructure survival anytime even if the infrastructure is operating under extreme conditions.
- Reliability - the capacity to ensure that a system or network will perform its intended functions without failures when operated under specific conditions for a specified time interval.
- Authentication - the capacity to identify a user that is appropriate to the specific information and service type.
- Access control - the capacity to ensure that only authorized users can access system and network resources.
- Data and message confidentiality - the capacity to ensure that only authorized users can access protected data and messages.
- Data and message integrity - the capacity to ensure that data managed by systems and messages transmitted over the network are not altered by unauthorized users or non guaranteed software or hardware.
- Reliable message delivery - the capacity to avoid message loss and replication, and guarantee ordered delivery, along with the ability to provide verifiable proof of delivery to both the endpoints of a communication.
- Non repudiation - the capacity to provide verifiable proof of message delivery to both the endpoints of a communication, in order to ensure that the sender of a message can not deny having sent the message and that the recipient can not deny having received the message.

In addition to dependability and security requirements, the financial infrastructure has to meet performance and Quality of Service (QoS) requirements, characterized by specific low level technical metrics for interconnection

networks, (such as packet drop, network latency round trip time, jitter, out-of-order delivery and transmission errors) as well as higher level business-level metrics (such as number of transmitted transactions, percentage of rejected transactions, number of incorrect transactions).

3.2 Standard Solutions for Securing the CI

The most challenging aspect in financial CIs is the new model that is being established for financial transactions. Unitil twenty years ago a financial transaction originated with a financial stakeholder (such as a bank) and was received through a complex communication network and few intermediate nodes by another financial stakeholder (such as another bank elsewhere). Communication networks at that time were quite controlled and secure. Nowadays the new model foresees online and real-time transactions that are generated by a non-financial stakeholder (usually a business customer), flow through financial stakeholders and intermediate nodes and sometimes arrive to another non-financial stakeholder (for example an enterprise or a SME). In this new model the communication network includes many different network types and quite often includes Internet as well. In such a case the communication network cannot be considered as intrinsically controlled and secure.

Communications among financial players are carried out through quite different technological solutions providing different performance, reliability and security levels: communications among financial institutions usually leverage dedicated leased lines, central bank offices are connected to local agencies through other dedicated lines or through secure Virtual Private Networks (VPNs) over Internet links.

Financial organizations are nowadays interconnected through extensive proprietary networks to provide their financial customers with advanced services and to exchange financial messages securely for business purposes (e.g., for cash management, funds transfer, credit advices, alerts). These networks are for financial transactions only and complex requirements related to security and privacy leads to proprietary and closed networks. Usually, financial networks are hierarchically interconnected according to a tree structure. In this interconnection model, each network can be considered as a tree node at a well-defined level. Therefore, two networks existing at the same level can communicate with each other by sending their messages to the network at the upper level, which guarantees a secure and reliable exchange of information.

Leased lines interconnecting financial institutions are specifically designed for high availability. Fault tolerance is provided by means of multiple redundancy. High dependability is also achieved through isolation of these dedicated communication lines with respect to the Internet traffic. This choice protects financial communications from availability issues. Dedicated communication lines used for information exchange among financial players can

provide a tightly monitored and controlled environment, in which it is possible to enforce performance-oriented policies. In this context, the possibility of performance guarantees is a direct consequence of the isolation of the dedicated communication lines with respect to the shared Internet. In isolated networks, it is quite simple to design and provide a communication infrastructure where the performance cannot be jeopardized by uncontrollable Internet phenomena and/or attacks that could result in the degradation of the communication channel performance. Moreover, the complete isolation of financial networks from other networks ensures a high level of security against intrusions or malfunctions from outside. However, it is often difficult to separate financial networks from the external ones because they have the need for the convenience of interconnecting to other networks to exchange essential data for financial purposes. Hence it is important to ensure maximum network interconnection security under these conditions, using suitable protection policies and technical solutions that guarantee full access and data exchange security. At the edge of financial CIs there are the connections among financial institutions and their customers. While high security guarantees can be achieved through dedicated channels, communications between a financial player and their customers are carried out through the Internet. Nevertheless, it is possible to guarantee authentication, non-repudiation, privacy and integrity by leveraging state of the art encryption and key distribution algorithms. VPNs can be established to enable secure communication between a known and authenticated user and (virtually) any host belonging to the internal network of a financial organization. This solution can be effectively used to enable secure (but not dependable) communication channels for customers or employees of a financial institution that is connected through the Internet. Transaction security will be implemented at platform and application level and performances may often not be guaranteed. Processes for establishing and securing the communication link and for managing transactions will be defined by financial institutions and then carefully implemented by customer (such as the use of OTP passwords to confirm transaction).

Communications that use the Internet as their backbone cannot be characterized by performance guarantees. It is possible to stipulate service level agreements (SLAs) when there is one provider among the financial institutions or when the traffic is confined in one autonomous system. In more general cases, however, it is impossible (or very tough) to guarantee SLA contracts when multiple autonomous system are involved between the communication endpoints. In fact, Internet traffic can be arbitrary delayed or dropped by the intermediate autonomous system that are based on a best effort routing service.

The Society for Worldwide Interbank Financial Telecommunication (SWIFT) is the most important worldwide financial communication infrastructure that enables the exchange of messages between banks and other financial institutions. It was founded in 1973 as a co-operative society owned by member

banks. SWIFT does not generate transactions, but it is responsible for providing a fast, secure, available and accurate means of transferring a variety of financial instructions on behalf of its international members. It is a private networks which provides the platform, products and services to connect and exchange financial information among financial organizations spread all over the world. Therefore, SWIFT can be considered as one of the nodes at the top of the tree model representing the interconnection structure of the European financial networks.

3.3 Types of attacks and exploited vulnerabilities

As modern society becomes more and more reliant on networked information systems, cyber-attacks to IT infrastructure gain the ability to target crucial services which are used by every citizen in their daily activities. Cyber-attacks against the IT infrastructure of financial institutions and their customers are representative of this trend. A large majority of financial activities are carried out by networked computers, and interaction between financial institutions and their customers are usually mediated via the open Internet. This landscape offers new opportunities for attackers. In particular, the ability to compromise the security of online financial transactions is especially alluring, since they give an attacker the opportunity to easily monetize a successful attack.

In this context, several different attack strategies have already been used in the recent past to prepare or to execute fraud and extortion against banks and their customers. These cyber-attacks are extremely heterogeneous, ranging from insider threats to network intrusion by an external attacker, from attacks targeted to a specific financial institution to widespread campaigns of SPAM and phishing, from the exploitation of vulnerabilities in software used by financial institutions to intrusions in customer's personal computers.

The most common attack strategies are: Man-in-the-Middle, portscan activities, distributed denial of service, session hijacking and malware-based attacks against financial institutions' customers. All these attacks share a common trait that makes them especially relevant: they involve multiple entities. Man-in-the-Middle attacks target multiple customers, and possibly multiple financial institutions. Portscan activities are routinely detected by virtually all financial institutions, and often performed by multiple, coordinated attackers. Distributed denial of services are a well known threat, which already targeted several financial institutions in the recent past, whose sources are geographically distributed. Session hijacking techniques can be used to compromise the integrity of financial transactions carried out by multiple customers. Finally, banking malware is usually represented by self-replicating software that attacks hundreds of thousands of vulnerable personal computers, thus targeting a high number of financial institutions' customers.

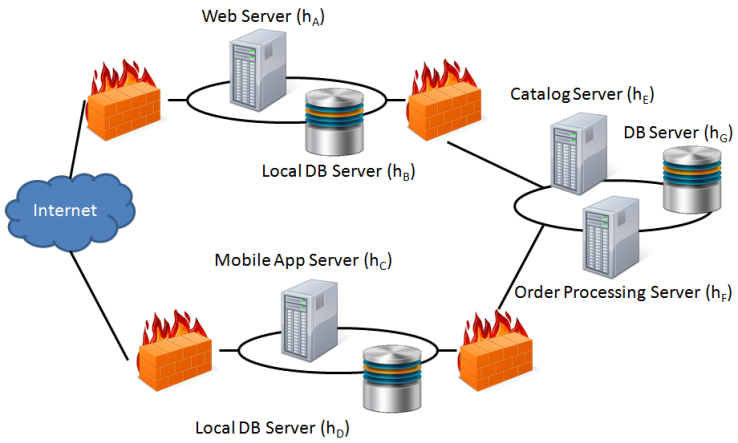


Figure 3.4: Example e-commerce network architecture

Moreover, there are many cases where attack strategies have more complex structures, defined by different patterns. Given a sequence of logged activities, it is then necessary to look for subsequences of the log that correspond to an attack. The models of the attacks must be capable of representing patterns with many alternatives and appropriate constraints. Consider for instance the example e-commerce network architecture shown in Figure 3.4. This network consists of three subnetworks delimited by firewalls. Two subnetworks include a host accessible from the Internet. The third subnetwork implements the business logic, and includes a central database server. An attacker who wants to steal sensitive data from the central database server will need to breach the firewalls and gain privileges on several hosts before reaching the target.

Now assume that a very simple attack pattern exists, comprising the following actions:

1. exploiting a vulnerability V_C on the mobile application server,
2. exploiting either a vulnerability V_D on the DB Server h_D or a vulnerability V_F on the order processing server,
3. exploiting a vulnerability V_G on the central DB server.

Also assume that moving through the order processing server triggers a basic security alert that prevents attackers from gaining access to the central DB server if the overall transaction time exceeds 20 time units. This pattern can be represented by the simple graph in Figure 3.5, where a time constraint is added between edges (V_C, V_F) and (V_F, V_G).

This very basic example shows that *(hyper-)graphs* can prove very effective as a basis for attack models. Our knowledge of the attacker's complex behavior can be formally encoded by means a hypergraph where vertices represent

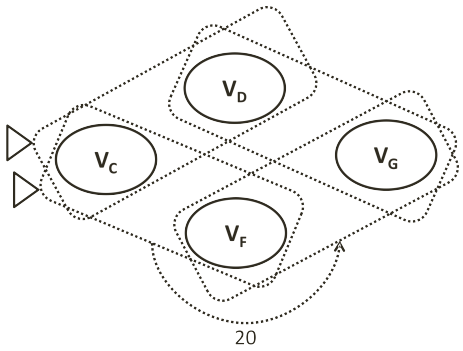


Figure 3.5: Example attack graph

possible group of events (i.e., actions of the attacker or exploitable vulnerabilities) while hyperedges among vertices specify the temporal sequence of the attack (as associations between vertices), with the intended meaning that the events belonging to the same hyperedge have to be completed within a specified temporal window, in any possible order. Moreover, attack models must allow various kinds of constraints to be specified during the possible attacks, in particular in terms of causal relationships that hold over hyperedges, and thus over the events they group together. It is then relevant to define and study forms of constraints that provide a high degree of flexibility in representing different security scenarios. Moreover, it is necessary to efficiently address the problems of checking the consistency and/or the redundancy of the attack models, and of efficiently detecting attack instances in sequences of logged activities.

3.4 Protection Strategies

This section covers protection strategies that aim at preventing attacks from having any effect. The strategies typically are twofold. The first strategy is to accept that a particular service is indeed vulnerable, and counter this vulnerability by making the potential rewards from an attack as small as possible. Another strategy is to reduce the vulnerability to near zero. Both strategies would make it unattractive to attack the service, and one would expect that attacks would not happen.

Standing order. The use of standing orders is increasing. A standing order is an agreement between the bank's customer and the bank which automatically effects payments, i.e. debit the customer's account and send a corresponding credit instruction to the creditor bank. There is one agreement for each debtor/creditor pair. The agreement often sets a limit as to the value of each payment. The agreement is ongoing, i.e. one and the same agreement applies to repeat payments, until the agreement is cancelled by the customer.

The creditor has to enter into an agreement with the bank also. A perpetrator would have to be accepted as a legitimate creditor for standing order payments as a first step in defrauding customers. The payments are effected automatically based on information stored centrally with the banks. The customer does not have to be logged on for the payment to take place. Hence there is limited scope for a Trojan, Man-in-the-Middle, Man-in-the-Browser when it comes to modifying transaction information. Increasing use of standing orders would limit the playing ground for fraudsters.

E-invoicing. The use of e-invoicing is slowly increasing in Europe. With e-invoicing, all payment instruction information is made available to the customer through Internet banking. In other words, the customer does not enter payment information, e.g., amounts, credit account number. The customer either accepts the e-invoice or declines it. The creditor and debtor have to enter into agreements with their respective banks in order to set up e-invoicing.

With both standing orders and e-invoicing the payer does not enter payment information and there is very little scope for a perpetrator when it comes to modifying payment information.

One time password. Several banks now use a one-time-password in connection with customer log-in to the Internet bank. The primary idea is that once the customer is logged in, no impostor having phished or otherwise got access to the code, may log in using the same code. By using more or less advanced phishing schemes however, impostors have been able to get access to the code before it has reached the bank. One instance of a scheme took this close to full automation, in that attackers presented banking customers with an automated front-end to the banking application, through which the log-on happened, giving the attackers full access to the customer account. In order to combat this kind of attack, some banks have introduced transaction authentication, this means that the customer has to enter a separate one time password (OTP) on submitting each transaction, i.e. after and in addition to the OTP submitted in connection with login to Internet banking. However, many OTP tokens (the devices that produce OTPs) are time synchronized with the machines that host Internet banking. Because clocks tend to differ slightly, and to allow for time to transmit and process the code, there is a time-window within which an OTP is valid. This time window allows attackers to phish two codes that subsequently prove to be valid; one for login and one for authenticating a (fraudulent) transaction.

Antivirus. Several banks now offer free antivirus software for the customers to download and run on their computers. It is given a prominent position on the home page of the banks, and customers are strongly encouraged to download it. Antivirus vendors are engaged in an arms race against malware

producers, and they are up against a well-organized and resourceful adversary. Lately we have seen the merger of two malware producers (Spyeye and Zeus). Advanced malware, e.g., polymorphous variants, change their signature dynamically and can escape antivirus software. Hence the effectiveness and efficiency of antivirus software is being questioned. In the aftermath of an attack against Norwegian online banking customers, security analysts reported that more than half of the investigated systems infected were running fully updated antivirus and operative system versions.

Transaction analysis. In order to become aware of and stop unauthorized transactions, banks perform back-end transaction analyses, both immediately and retrospectively. In a recent wave of attacks against banks in Norway, back-end analyses are known to have prevented losses. Transactions were compared against black lists of accounts. The footprint of the Trojan was identified, and transactions that matched the footprint were stopped. There are companies that are proficient in collecting traces and indications that authentication credentials have been compromised. The companies post these traces to so called drop sites which contain profiles of compromised users.

Computer Emergency Response Team. Several banks subscribe to a CERT. In several countries CERTs are government bodies. CERTs are manned with highly qualified technicians who analyze traffic and traffic patterns looking for possible attacks. In a recent attack in Norway, the Norwegian national CERT played a prominent role in analyzing the Trojan and also using its power to convince the ISPs to close down IP addresses of the command and control center of the trojan.

Protection from complex attacks. Most companies nowadays employ a variety of technologies to identify and thwart suspicious activity. Common technologies include the use of firewalls to detect them at the perimeter of a network, antivirus packages to identify malicious code entering companies systems from various sources, and intrusion detection software (IDS) to scan packets on networks and to monitor a variety of questionable activities on application servers and at the operating system level. These techniques are usually very effective when used to assess vulnerabilities as isolated actions. However, they do not offer the necessary level of protection against malicious activities that are performed in combination, thus giving rise to a complex (and generally more powerful) kind of attacks. In order to face complex attacks, companies might set up a specialized security team (or a CERT), but this is often rather expensive and its activities might be time and resource consuming. For these reasons, an emerging security paradigm is to adopt novel kinds of technologies based on a level-wise protection strategy. The basic idea is to encode the knowledge about specific complex attacks into a suitable

formal model describing an attacker's behavior as a combination of smaller attacks. The resulting model is used for analysis purposes and/or as a predictive model for detecting future attacks. In fact, best security practices demonstrate that attack models (e.g., attack graphs, trees, hypergraphs) can be applied to support offensive (e.g., penetration testing) and/or defensive (e.g., network hardening) strategies, in a wide spectrum of security contexts, including vulnerability analysis [129, 104], intrusion alarm correlation [172, 173], and attack response [71]. For example, models such as attack graphs are successfully used to identify all potential paths of vulnerability, thereby evidencing how attackers can penetrate through a network [37, 38, 41]. The approach consists in modeling the network configuration (topology, connectivity limiting devices such as firewalls, vulnerabilities, etc), and then in simulating attacks (in particular, penetrations) over the networks. Sets of attack paths (for the successful attacks) can be collected and organized in a graph based structure, which can be reused as a predictive roadmap for real attacks. In the context of intrusion detection, attack models have been also experimented for modeling computer networks that are protected by an IDS and for correlating the large numbers of alerts that are produced by that IDS with actions in the attack model [137, 175, 176]. Alerts that arrive in a sequence that is predicted when proceeding along a single attack path may indicate that an attacker is successfully performing the steps in that path. Early approaches have also experimented with the use of attack models in the context of anomaly detection, where a model encoding normal behavior as a combination of actions is used to automatically learn and detect normal/abnormal behaviors in an observed actions [174, 130].

On demand protection measure. In certain countries Internet banks use solutions for authenticating customers that are also used to authenticate the customers in other websites. In other words there is one authentication server serving all sites. The solution often employs OTP codes as part of the authentication. This implies that any OTP code would be valid for any one of the websites, including Internet banking. As the authentication solution is gaining ground, more and more websites ask the customers for their credentials. Previously the authentication credentials were submitted in the context of Internet banking only. Now, people are being prompted for login credentials in different contexts, e.g., in the context of various online shops, in the context of logging to public services etc. Under these new circumstances, it is harder for the public to exercise vigilance and to know who is behind this website asking for their credentials. As a result of not being able to exercise control, people will tend to become less critical when it comes to whom to submit login credentials to. This provides an example of phishing. An attacker would purport to be an online shop and phish login credentials and then turn around and use the credentials to log into Internet banking. Alternatively a website which is

authorized to accept credentials for authentication could use the credentials for unauthorized purposes, defrauding the customer. In order to combat this threat, the organization behind the authentication solution has come up with the idea of a context sensitive OTP. This means that the OTP will be issued per website, i.e. a phished OTP would be valid only in the context of one particular website. An OTP obtained in connection with online shopping, could not be used to log into an online bank. This of course would remove the prime motivation for the phisher.

Authentication roaming. Many online banking authentication solutions are roaming, i.e. the customer may gain access to the Internet bank from PCs anywhere using the same authentication mechanism. This provides a case for phishing, inasmuch as the phished authentication credentials may be instantly used by the attacker from the attacker's PC. To counter this portability aspect of authentication, banks are known to have built up a table with MAC addresses and matching logon-ids. They have built this table by recording the MACs that the user regularly uses. If there is an attack, by looking up any one customer in this table, the bank would check the MAC address and allow access only from this address or a limited number of other machines that the customer has been using.

3.5 Fault Mitigation Approaches

Due to the role played in the economic development, the financial system is can be considered one of the most regulated sector. The regulation framework involved uses financial IT requirement and risk management processes in order to establish a sound financial system. Infact, Financial IT Infrastructure is largely used to process, store and exchange critical and sensitive information, hence it is characterized by strict security requirements. Systems, networks, data and exchanged information should be protected against any type of malicious activity (such as interception, insertion of fake information, update, delete). The relevance of security requirements in the financial context is highlighted at both firm (financial intermediaries/institutions) and financial markets infrastructures level, focusing on operational risk management and/or business continuity management. Operational risk management focuses on every conceivable risk that could potentially affect the smooth operations of a system or service. In the financial sector, operational risk has wide ranging systemic implications given the increasingly large size, interconnectedness, and complexity of financial institutions which increase the possibility of errors and fraud. Disruptions to the flow of financial services because of impairment of all or part of the financial system may give rise to systemic risk and possible spill over effects to the real economy. Business continuity management looks at one aspect, operational failures, that could disrupt the delivery of

a key services. Therefore the two disciplines have generally similarities and overlap: business continuity management can be regarded as a specialist discipline which is at ones complementary to and part of the overall operational risk management process. Recognized best practices and standards suggest that an effective business continuity management programme should typically comprise the following four key elements: i) a business impact analysis with a view to identifying critical activities and determining recovery objectives; ii) a well-defined business continuity strategy; iii) appropriate plans and procedures to ensure the continuity of critical services; iv) the testing, maintaining and reviewing of existing plans in order to validate their effectiveness and ensure that they are kept up to date.

Regarding financial intermediaries, banks in particular, the main regulatory prescriptions in the field of operational risk are provided in the new capital accords, the so called Basel II and III [4] (BIS). Damages caused by security breaches within the financial IT infrastructure generally fall within this risk category, as defined in the Basel III first pillar[3] and Annex 9[2]. In particular, system security issues (such as hacking activities and data theft), are considered as examples of the external fraud event type, defined (among others) within the operational risks.

As a follow up to the financial crisis which erupted in 2008, the Committee of European Banking Supervisors (CEBS) issued guidelines covering either wholly or partially internal governance aspects for credit institutions and investment firms; in particular new guidelines on information and communication systems and business continuity management were added in the fifth chapter, "Systems and Continuity". Instead of formulating extensive requirements with regard to IT systems, the guidelines refer to generally accepted standards in this matter. The principles on business continuity are consistent with the BCBS "High Level Principles for Business Continuity". In Italy the Bank of Italy (responsible for issuing secondary legislation on technical matters regarding financial intermediaries and for interventions of a prudential nature) enforces both national law and European directive and regulation, all requirements needed in order to achieve a comprehensive management of IT systems both in term of security and of dependability. In this respect the 15th update (2 July 2013) of the "Circular No. 263 - New regulations for the prudential supervision of banks" is one of the most recent reference documents in this field; it disciplines the organization of internal controls, the functioning, roles and responsibilities related to the development and the management of Italian financial informative systems. This intervention transposes the CEBS's guidelines onto the Italian regulatory framework. Among others, the most important innovations are related to:

- The discipline of information systems. The discipline of information systems, taking into account the main developments which emerged on the international scene and setting the main constituents of governance

and organization of the information system, IT risk management, all the requirements to ensure the security and the system management of data. The provisions also provide that the definition of principals security for access to critical systems and services through the Internet channel are applicable Recommendations of the ECB in the field of security of online payments.

- The Business continuity discipline. The Business continuity discipline, by reorganizing provisions presently contained in different regulation sources. Among others a process of rapid escalation by accident in emergency was defined so as to ensure that the declaration of a state of crisis happened in the shortest possible time from the moment the accident is detected. The total time for recovery will not exceed four hours, including times for the stages of analysis, decision-making,technical assistance and verification.

- The formalization of the role of the CODISE Working Group. The formalization of the role of the CODISE Working Group (the business continuity working group set up in 2002) as the structure responsible for the coordination of crisis management operating in the Italian financial system. The group is coordinated by the Banca d'Italia in agreement with the CONSOB (the Italian stock exchange commission) and consists of representatives of the leading banking groups and the companies that manage infrastructures essential to the orderly working of the financial system.

The fault mitigation concepts are embodied in the risk analysis process and in the information availability sections of the document. Below the main principles are reported:

Risk Analysis. Risk analysis of ICT resources constitutes an instrument to grant efficiency and efficacy for their protection strategies. It allows for the regulation of mitigation measures based on the field in which the system operates. The risk to which ICT resources are exposed should be evaluated. This involves both the development of new structures and updating the pre-existent ones. The risk analysis will provide levels of classification, potential risks and residual, lists of considered threats, lists of individuated assets, and will be repeated periodically, according to the criticality of the ICT resources.

Information Availability. Information availability and ICT resources availability is guaranteed to users according to the service level agreements[2]. To this aim, all the processes of:

[2]The usage profile of the clients (known or estimated) during the operative hours and potential utilization spikes should be taken into account.

- architectural models design;
- software and infrastructure development;
- fault management;
- transmission capability planning and Monicatoring;
- processing capability planning and monitoring;
- providers management;

must take into account the following directives:

SLA monitoring
Service levels that should be observed are formally defined, specially regarding applications that have a higher level of criticality. The performance of components and assets that are required in order to obtain the target level of service must be regularly monitored.

Off-site backup
According to availability requirements of each assets, software or service, backup procedures for software and configurations, for data and for hardware systems must be defined. Off-site backup should be ready and preventively individuated.

No single point of failure
Strictly related to the previous point, each architecture must be designed considering security profiles of the hosted applications. All the ICT resources and supporting resources (electrical power, cooling systems, etc.) should be correctly redundant and robust, no single point of failure should be present. The higher availability should be granted for applications with higher levels of criticality, according also to the disaster recovery plans.

Interconnection redundancy
According to the risk profiling of the communication systems, of the application and of the services accessed, each bank or financial institution should redound the ICT links as well as specific solutions for the detection and blocking of malicious traffic, and the bank should evaluate procedures and instruments of dynamic allocation of the transmissive and computational power.

Standardization
The management of the ICT systems is properly automatized and uses, at most of the effort, standard procedures. The ordinary and extraordinary maintenance operation must be planned and timely spread to all interested users.

Area	Main contents
General organization	Principle 2 on governance requires that an FMI have robust governance arrangements that focus on the safety and efficiency of the FMI and that support the stability of the broader financial system, other public interest considerations and objectives of relevant stakeholders.
	Principle 3 on the framework for the comprehensive management of risks requires an FMI to take an integrated and comprehensive view of its risks, including those it bears from and poses to its participants, their customers and other entities
Operational risk	Principle 17 strengthens the requirements on operational reliability and resilience[3].
Access	Principle 18 provides guidance to an FMI for establishing appropriate access policies that provide fair and open access, while ensuring the FMI's own safety and efficiency.
Efficiency	Principle 22 on the use of communication procedures and standards. An FMI should use, or at a minimum accommodate, internationally accepted communication procedures and standards to enhance efficiency[4].

Table 3.2: PFMI principles related to security. Source: CPSS-IOSCO, 2012.

Capacity Planning

Information obtained through ICT resource monitoring regularly feed into the capacity planning and should be used in the designing and in the updating of the informative system.

Nonetheless an important role is played by a comprehensive business continuity plan (BCI), which should always be respected in order to ensure the availability of critical financial services.

As reported above, FMIs are important contributors to the removal of financial risks, but must ensure that they do not themselves become sources of unacceptable risk in the financial system, particularly in severe stress conditions. In this field the Committee on Payment and Settlement Systems (CPSS) and Technical Committee of the International Organization of Securities Commissions (IOSCO) have contributed to the set of standards, codes and best practices that are deemed essential for strengthening the financial architecture worldwide.

[3]For example, business continuity management should aim for timely recovery of operations and fulfillment of the FMI's obligations, including in the event of a wide-scale or major disruption. Business continuity plans should be designed to enable an FMI to complete settlement by the end of the day of the disruption even in extreme circumstances, and critical systems should be designed so that operations can be restored within two hours of a disruption.

[4]For an FMI that maintains cross-border operations or provides cross-border services, the use of internationally accepted communication procedures and standards is particularly important.

In April 2012, the important document "Principles for financial market infrastructures" was issued.[5].The twenty-four principles outlined in this report are categorized into nine broad categories: (a) general organization, (b) credit and liquidity risk management, (c) settlement, (d) CSDs and exchange-of-value settlement systems, (e) default management, (f) general business and operational risk management, (g) access, (h) efficiency, and (i) transparency. These broad categories encompass the major elements critical to the safe and efficient design and operation of FMIs. Table 3.2 highlights the most important principles related to our field of investigation.

CPSS and IOSCO have also recently (August 2013) published for public comment a consultative report on the recovery of financial market infrastructures. The report is intended:

- to provide supplemental guidance on, and a menu of tools for, observance of the PFMI, taking into account different type of FMI;
- to be consistent with the FSB's Key attributes of effective resolution regimes for financial institutions;
- to provide guidance on the recovery planning process and content of recovery plans.

At European level considering the FMIs, we can consider three area of major requirement can be considered:

- Retail payment systems: RPSs are used for the bulk of payments to and from individuals, and between individuals and firms; these systems are currently subject to major changes as a result of the implementation of the Single Euro Payments Area (SEPA). Even if many of them are not of systemic importance, they play a major role with respect to both the safety and efficiency of the financial system as a whole and citizens' confidence in the euro. In recognition of the relevance of retail payment systems, the Eurosystem has introduced "Oversight standards for euro retail payment systems", which distinguish between systemically important payment systems, prominently important payment systems and others, and specify which of the Core Principles are also of relevance for prominently important retail payment systems. In order to ensure a consistent application of these oversight standards by the different NCBs and the ECB, the Eurosystem has released a common methodology for the assessment of systems against the respective standards.

[5]The new standards replace the three existing sets of international standards set out in the Core principles for systemically important payment systems (CPSS, 2001); the Recommendations for securities settlement systems(CPSS-IOSCO, 2001); and the Recommendations for central counterparts (CPSS-IOSCO, 2004)

- Large-value payment systems: LVPSs form the backbone of the euro area market infrastructure. The Eurosystem applies the Core Principles for Systemically Important Systems of the CPSS and has refined them further by issuing "Business continuity oversight expectations for systemically important payment systems" that elaborates further on the business continuity aspects of the Core Principle on security, operational reliability and business continuity.

- Securities and derivatives clearing and settlement systems: System and process failures are particularly dangerous if they occur in the clearing and settling of financial transactions as well as in the trading and pricing of financial instruments. The infrastructures and arrangements for the handling of securities are, to some extent, more complex than those for the handling of payments. Since securities are, as a rule, delivered in exchange for payment, there are two delivery legs to consider, the cash leg and the securities leg. The handling of securities also involves a wider range of functions and participants.

Level of intervention	**Documents**	**Main content**
Payment systems	Core Principles for Systemically Important Payment Systems, Bank for International Settlements (adopted by the Governing Council of the ECB in January 2001).	In particular Core Principles VII: The System should ensure a high degree of security and operational reliability and should have contingency arrangements for timely completion of daily processing
	Business continuity oversight expectations for systemically important payment systems (SIPS), (ECB, June 2006).	The document constituted the revised implementation guidelines, which are described in this paper in the form of oversight expectations, identify key elements of business continuity management. They will contribute to ensuring a level of resilience on the part of SIPS across the euro area which is consistent with the objective set by CP VII.
	Regulation No 260/2012	Common technical standards established for processing SEPA payments, necessary to allow interaction and interoperability between IT systems and to ensure an automated processing of euro-denominated transactions between payment service providers (PSPs), referred to as "straight-through processing". The regulation requires the use of certain common standards and technical requirements, such as the financial services messaging standard ISO 20022 XML for all credit transfers and direct debits in euro in the EU.

Continued on next page

Table 3.3 – continued from previous page

Level of intervention	Documents	Main content
Payment instruments	Harmonised oversight approach and oversight standards for payment instruments (ECB, February 2009)	Standard 3: The scheme should ensure an adequate degree of security, operational reliability and business continuity. Adequate security controls should be in place to mitigate operational risks. In this regard, the governance authority should ensure that all relevant actors in the scheme focus on risk and security management, business continuity and outsourcing by ensuring that adequate technical standards and procedures are in place.
	Oversight frameworks for direct debit schemes, October 2010	The aim of the oversight framework for direct debit schemes and for credit transfer schemes is to ensure the soundness and efficiency of payments made with such instruments. Five standards have been identified that deal with legal issues, transparency, operational reliability, good governance and sound clearing and settlement processes; in particular standard n. 3 talks about the necessity to ensure an adequate degree of security, operational reliability and business continuity.
	Oversight framework for credit transfer schemes, October 2010	
	SEPA Cards Standardisation Volume - Book of Requirements Version 6.0 (EPC, January 2012)	The volume defines standards requirements for cards and terminals. It also defines the functional and security requirements, including requirements for the evaluation and certification methodology and architecture, that are recommended by the EPC for adoption throughout the card payment value chain to ensure interoperability within SEPA. Security requirements (including "card not present" and innovative payments) as well as certifications are included in the Volume.
	Recommendations for the security of mobile payments. draft document for public consultation (ECB, November 2013)	The report outlines 14 recommendations, constituting minimum expectations for promoting the security of mobile payments, organized into three categories.
Clearing and settlement systems	The Eurosystem's policy line with regard to consolidation in central counterparty clearing (ECB, September 2001)	Recommendations aim to promote efficient, safe and sound pan-European post-trading arrangements in order to increase confidence in securities markets, ensure better investor protection, contain systemic risk and foster financial stability
	Eurosystem statement on central counterparties and interoperability, terms of reference (ECB, March 2008)	

Continued on next page

Table 3.3 – continued from previous page

Level of intervention	Documents	Main content
	Recommendations for securities settlement systems and central counterparties in the EU (ESCB, -CESR, June 2009)	

Table 3.3: List of standards and best practices at EU Level on Payment Systems.

Table 3.3 summarizes the most relevant sources of standards and regulation framework regarding FMIs in European Union.

Finally greater attention has recently been paid to third-party service providers to which payment and settlement systems contract all or part of their operations (e.g., their IT infrastructure);those providers may be of critical importance for the functioning of those systems. For the Eurosystem, a key principle is that the individual systems retain full responsibility for any activity that is material to their operations, including responsibility for ensuring that the service provider complies with applicable oversight policies. Only when a service provider supplies important services to more than one key system will direct oversight activities be undertaken. For instance, this is the case for SWIFT, a global provider of interbank financial telecommunication services.

3.6 Open Problems

Financial services and organizations are subject, by their very nature, to a number of security concerns. Many standards and regulations are currently actively influencing the way organizations manage their internal and external security. Exploitation of software vulnerabilities is a common threat vector, responsible for a number of breaches that, unfortunately, often remain undetected by the victim for a long time [46]. In these cases, a piece of malware, a backdoor, a key-logger or some other malicious software can stay operative and undetected on the victim system for months if not years. The damage this may cause to a financial institution and its costumers may be enormous. For this reason, standards like Payment Card Industry Data Security Standard (PCI-DSS) regulate what vulnerabilities *must* be fixed (by means of a process called *patching*) by the organization. Amongst the most important reasons for being compliant, liability in case of accidents is surely one of the main factors that pushes an organization toward being certified.

Keeping PCI-DSS as an example, organizations must patch vulnerabilities according to their CVSS score, the industry standard for *vulnerability risk assessment* [126]. The CVSS score is used, in this context, to provide a "threshold of risk" above which all vulnerabilities must be fixed. This threshold is subject, clearly, to a trade-off. On one side, one wants to fix as many vulnerabilities as possible to cover most "risky" vulnerabilities. On the other, setting the threshold too low would generate an impossible-to-manage amount of vulnerabilities to fix. In fact, system administrators often have to deal with hundreds of vulnerabilities, which is not a trivial task to carry out.

Unfortunately, the CVSS score has been shown to be a poor indicator for actual exploitation [39], which generates a very high amount of false positives and false negatives: compliance for security makes you do much more work than you should actually do to stay secure, and also misleads you in skipping vulnerabilities that you really should fix [40]. In other words, despite the huge amount of work being compliant asks for, to what degree it addresses actual security is not clear. As a result, in order to stay compliant many organizations are forced to hire entire teams which sole purpose is to *justify why certain vulnerabilities have not been fixed.*

This adds additional costs and organizational overhead, and hardly contributes to enhance the security of the organization: security measures to drive compliance and security management have still much room for improvement both in the sense of *threat identification* and *organizational efficacy.*

CHAPTER

4 The Power Grid

The energy sector represents a crucial asset in most modern countries that base their industrial and societal growth on the continuous availability of energy in its various forms. As a consequence, most of the industries working in this sector are considered as critical infrastructure. The energy sector is listed as one of the two sectors representing ECI in Europe (see Section 1.1). It is also one of 16 critical infrastructure sectors established by Presidential Policy Directive 21 (PPD-21) [4] in the US.

This chapter analyzes the energy sector. It starts with an introduction of the power grid, its main stakeholders and players, and its requirements. It continues by describing some protection strategies that aim at preventing attacks from having any effect, then, it concludes with an overview of the open problems in this field.

4.1 Description of the Critical Infrastructure

The energy sector includes assets related to three key energy resources: electricity, petroleum, and natural gas. A power grid is an interconnected network for delivering electricity from suppliers to consumers. It is composed of three main components: power plant, transmission substation and distribution grid. A power plant produces simultaneously three different phases of AC power with 120 degrees offset from each other. The three-phase power feeds a transmission substation. This substation uses large transformers to increase the generator's voltage up to extremely high voltages to reduce transmission line losses over long distances. The distribution grid is the final stage of energy conversion before the electricity is supplied to end users. Power grids were designed in order to meet requirements that were defined in the 20th century when the goal was to "keep lights turned on". Today, the requirements

expected to be fulfilled by power grids have changed. The increasing load and consumption demands increase electricity issues, such as blackouts, and overloads. In July 2012, for two days, India experienced blackouts that involved a large portion of the country's power grid. Specifically, a 9% gap was estimated between the effective energy requirements and the available energy amount [152] [158]. In the afternoon of 8 September 2011, an 11 minutes-long system disturbance occurred in the Pacific Southwest, leading to cascading outages and leaving approximately 2.7 million customers without power. The failure of the power grid was due to bad redistribution of the power flow caused by the failure of a transmission line. Other examples of power grid blackouts due to different types of failure are reported in [25] [169] [108] [141]. In addition, it is estimated that power outages and power quality disturbances cost to the economy from 75 to 180 billion US dollars annually. The growth of electricity demand is only one of the motivations that makes the power grid an obsolete technology. In fact, in 2009, the Department of Energy defined other requirements for modern power grid design, known as smart grid [139] [114]. The requirements are:

- Enabling informed participation by customers. Traditional power grids provide a one-way communication model between the power plant and end users, so customers assume a passive role within the power grid infrastructure. Instead, a two-way communication model with the participation of users is encouraged. In fact, with bi-directional flows of energy and coordination through communication mechanisms, a smart grid helps balance supply and demand and enhance reliability by modifying the manner in which customers use and purchase electricity. The smart grid becomes an active electricity market that allows customers to shift load and to generate and store energy based on near real-time prices and other economic incentives.

- Accommodating all generation and storage options. The future power grid cannot be based only on a centralized power generation, but must also adopt diverse and widespread distributed energy resources such as solar, hydro-electric and wind. Of course the network architecture of smart grids must be designed in a flexible way to support different types of energy resources. Energy resources heterogeneity allows to alleviate peak load and to support back-up energy during emergencies.

- Enabling new products, services, and markets. The bidirectional communication between end users and operative central of a smart grid allows for the creation of new products and services customized to the customers. By using consumer-oriented smart appliances or Intelligent Electronic Devices (IEDs), for instance, customers or service providers can remotely control IEDs' power usage. Markets act as coordinators

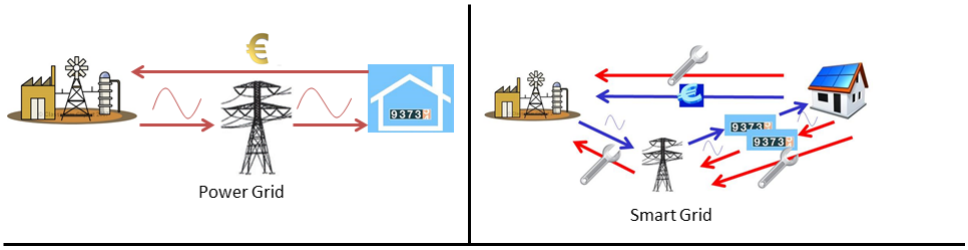


Figure 4.1: Evolution from Power Grid to Smart Grid

.

managing a series of independent grid parameters, such as time, capacity, the rate of change, service quality, etc.

- Providing the power quality for the range of needs. Power quality is a very important aspect of the modern power grid. In particular, mechanisms to avoid voltage flicker and momentary interruptions must be implemented. Is is also necessary to distinguish between power quality required by industries and residential users. So the architecture of smart grid must be designed to meet a wide range of power quality.

- Optimizing asset utilization and operating efficiently. The smart grid is a complex system composed of different subsystems that cooperate to ensure the requirements described. Each subsystems manages a variety of appliances, facilities, and distributed energy resources. So optimizing the utilization of those assets will reduce the whole life-cycle, investment costs and power consumption.

Figure 4.1 shows the simplified scheme of a classic power grid (left side) and a modern smart grid (right side). The communication flow of a classic power grid is unidirectional while the smart grid uses a communication bidirectional model i.e., the smart grid introduces feedback to regulate power distribution, generation and diagnose problems in the network. Nevertheless, increased connectivity is becoming more critical to the cybersecurity of the power system. In fact, many organizations are currently involved with the development of smart grid security requirements, including the North American Electrical Reliability Corporation – Critical Infrastructure Protection (NERC CIP), the International Society of Automation (ISA), the National Infrastructure Protection Plan (NIPP), and the NIST. Analysis [98] conducted in collaboration between Iowa State University and University of Illinois at Urbana-Champaign is focused on the definition of security requirements for smart grid infrastructure. In particular for each component of a smart grid infrastructure the following security requirements are identified:

- Advanced Metering Infrastructure (AMI). Consumer homes will be enhanced with the addition of smart meters which provide two way communication between the customer and utilities. AMI presents its own unique security requirements [24]. Confidentiality is of greater concern than in other grid domains due to the large quantities of end user billing and privacy data. Integrity is necessary for both the meter's operation and control, along with the communication of both pricing and status information. Authentication and nonrepudiation of both utility and consumer activities are critical.

- Distribution management systems (DMS). Management and automation systems are becoming increasingly important to meet the demands of the energy distribution infrastructure. DMS systems are geographically distributed; they communicate using the network and primarily perform control applications. So DMS demands both high integrity and availability of all supporting control and communication resources. In addition to integrity and availability demands, all critical system functions and messages must be authenticated to ensure malicious individuals cannot send fraudulent data or commands.

- Energy management systems (EMS). Unlike DMS systems, EMS focus on the bulk power system generation and transmission domain. EMS have historically utilized real-time communications for control and monitoring, with applications such as automatic generation control (AGC), state estimation, and flexible AC transmission systems (FACTS). EMS and networks maintain obvious requirements for strong integrity and availability. These attributes are especially important due to the criticality of the applications controlling the bulk power system. Additionally, strong authentication should be supported for all grid-related communications, especially remote field devices, such as IEDs and PLCs.

- Wide area measurement, protection, and control (WAMPAC). Phasor measurement units (PMUs) are devices used to monitor and to protect the modern power grid. The ability to perform real-time grid state measurements will enable the development of increasingly effective protection schemes and control functions. However, WAMPAC systems will be extremely dependent on high speed networks, additionally, phasor data concentrators (PDC) and gateways that can both authenticate and authorize the sharing of PMU readings with various utilities and independent system operators. The cybersecurity concerns and requirements for WAMPAC are well documented [23]. Authentication plays a critical role in WAMPAC environments. Proposed architectures such as NASPInet have identified the need for sophisticated access control mechanisms to limit the transmission of PMU measurements only to authorized parties. Availability and integrity are again critical for high

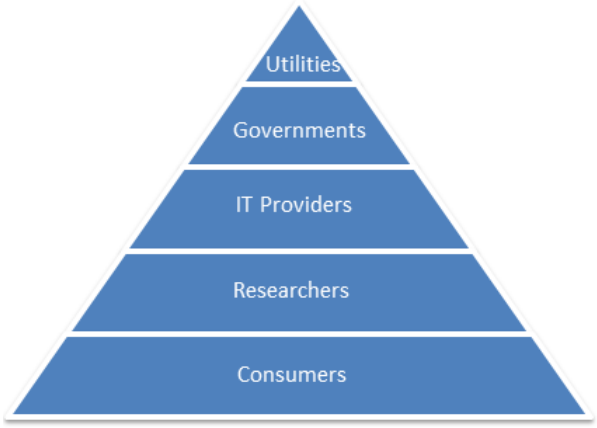


Figure 4.2: Main Players in the Smart Grid context

speed communication. Finally, PMU measurements depend on GPS technology for timestamp data. This dependency inherits additional security concerns from potential jamming or spoofing attacks.

4.1.1 Main stakeholders and players

Particular attention must be paid to the identification of the main stakeholders involved in the smart grid development. Stakeholders range from utility and energy producers to consumers, policy - makers, technology providers, and researchers [128] [82]. The primary benefit of a smart grid development to these stakeholders concerns the mitigation of energy prices, reduced dependence on foreign oil, increased efficiency, and reliability of power supply. Figure 4.2 shows the categories of stakeholders.

- Utilities. These are focused on the implementation and installation of technologies. They can provide more reliable energy, particularly during challenging emergency conditions.

- Governments. They establish new standards for operation, monitoring and interoperability and are also responsible of the creation of new regulations to improve smart grid infrastructure. Finally they mediate the needs of all parties involved in the smart grid development.

- IT providers. They develop new technologies for power grid enhancement. IBM and CISCO are the major players in the provision of IT equipment for the smart grids at a global level. In 2008, IBM was chosen to spearhead IT support and services for smart grid energy efficiency programs by American Electric Power, Michigan Gas and Electric, and Consumers Energy. CISCO has contributed with a new IP architecture.

CISCO describes the smart grid as a data communication network integrated into the electrical grid that collects and analyses data captured in near real time about power transmission, distribution, and consumption.

- Consumers. In the smart grid context they become both consumers and producers. In fact they are called prosumers. This new role of the consumer creates new business opportunities. In fact, the classic consumer can generate energy (e.g., through solar panels) and provide the stored energy to the network.

4.1.2 A smart metering scenario

This scenario is focused on the prototypical model of a private household with an Advanced Metering Infrastructure. In this context we identified the following list of key components: the smart meter (SM), the energy management system (EMS), the smart appliance (SA), the home area network (HAN) and the home gateway (HG), which are inside the home domain; the data communication network (DCN), the network gateway (NG) and the energy supplier server (ESS) which are in the domain of the energy supplier. The other two, the remote device for house management (REMS) and the energy generator (EG) are in separate domains.

- SM. SMs are devices that record the energy consumption of appliances within a home environment and communicate this information to energy suppliers via a DCN. For more details, see below.

- (Home) EMS. It is assumed that the EMS is one dedicated computer, more precisely, it is a web server, that allows the user to observe how much single appliances or rooms are consuming, and also their production and storage. Here, the user is also able to define policies, describing in some detail when to buy, sell, store, or consume energy. It also hosts data management applications and directly or indirectly (via the SM) controls the SA, see [28]. For simplicity, it is assumed that it communicates with all other elements in the house via a wireless HAN and it is connected to the internet, via the HG. The user is able to log on into a personal device (PC, Tablet, etc) and access the functionality of the EMS.

- REMS. Users can remotely access, via some mobile applications, their EMS, access their data, or modify their policies.

- SA. SAs are appliances that can be remotely monitored and controlled; as such, they inherently include appropriate monitoring modules. For the purpose of this report, thermostats, energy generation devices (like solar panels) or charging stations are treated as SAs: they receive control messages (say, commands) and send status information.

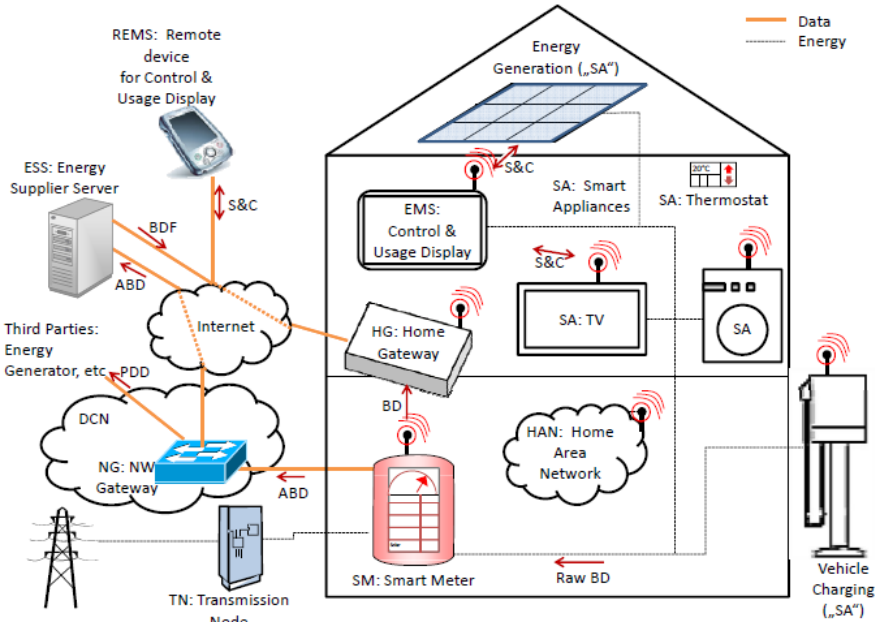


Figure 4.3: Case Overview: Entities and Steps.

- HAN.The wireless HAN, used by the SM and by smart devices to communicate with the EMS and by the EMS to communicate with the HG.
- HG. HGs are devices that can access the Internet, and also via the HAN, the SAs, electric switches, and the SM.
- ESS. This collects Aggregated Billing Data (ABD) from the smart meters, plus other data from the home gateway for added value services. The ESS (or another server of the same domain) also stores ABD information.
- The meter point operator (MPO). The MPO has a particular role, outsourced and controlled by the energy supplier, with the purpose of installing and maintaining the main devices of the advanced metering infrastructure, namely the SM and the EMS.
- DCN. This enables two-way communication between the SMs, on the prosumer side, and the NGs of the energy suppliers. A DCN is typically implemented using a public IP network.
- NG. The NG interconnects HGs within a specific area to other smart grid components, such as energy suppliers or transmission system operators.
- EG. These typically operate conventional or regenerative power plants (fossil, nuclear, solar power, etc.). Their importance is that they receive

aggregated data from the households, that is: how much energy was consumed or produced in a city area (say, at least 30 households) in a regular amount of time (say, every 5 minutes).

Description of the Environment Scenario

In this scenario the household is able to produce, store, and consume energy, and to regulate the corresponding amount of energy that is taken from the grid or fed into the grid depending on market prices. It is also able to communicate data for the energy network and other services, to use SAs which accommodate the household's behavior by reacting to personal preferences, and to use the electrical vehicle charging infrastructure, which supports the stabilization of the smart grid by providing an energy storage for the grid. Figure 4.3 depicts the an overview of the environment. All communication may be assumed to be sent encrypted, but for some connections the costs for providing the appropriate security infrastructure (software, hardware, key distribution, etc) may be too large. The keys in the smart meters and in the HGs may be assumed to be placed by the MPO, and the keys of the HAN to be managed by the customer (end user), but other choices could be better suited. In general, an interesting, but difficult question is the key management issue, see section 4.1.2.

- Raw billing data (BD). The raw data related to energy consumption, storage and production is gathered by the SM. How exactly is not relevant to this report and it is assumed there are no problems here.

- BD. Then, the SM processes and stores data. It communicates the data to the EMS (via the HAN), so that the energy consumption, storage, and production can be analyzed and adjusted by the prosumer inside the household. The data is also stored by the (local) EMS.

- ABD. The SM sends, on the other hand, ABD to the NG over the public DCN, which forwards it to the energy supplier. In some scenarios, like on-demand reading, BD could be collected at low frequencies (daily/weekly/monthly), but since power metrics are required to provide real time incentives for energy savings during consumption peaks, BD are sent at high frequencies, depending on the price calculation. ABD is to be considered as personally identifiable information: it can easily be linked to the users in the household.

- Power generation and distribution (PDD). Energy suppliers and energy generators use data for PDD purposes to obtain usage forecasts for certain sectors. PDD is aggregated BD from several households. As opposed to ABD, PDD should contain personal information only up to the point of aggregation, even if the collection frequency is high. Obviously,

a sufficiently large aggregation set of different households and a trustworthy aggregation party is required. The chosen granularity for the aggregation (town area) appears to be be appropriate. Thereby, future energy generation can be calculated and regularized.

- Billing data feedback information (BDF). It can be assumed that every five minutes the users are informed about their energy usage or generation volume, costs, revenues, and current rates. In this way, the EMS calculates the price of the energy that it needs for its workload, based on public energy rates.

- Status and control (S&C) messages. The user locally logs on into his EMS, views the status of his devices and sends commands to the SAs or modifies the energy management policies. For instance, the customer triggers the activation of the washing machine. The history of actions is also stored by the (local) EMS in a log. Furthermore, the SAs send status messages to the EMS. S &C messages are also sent and received when the user is not logged in, due to rules of policies, timeouts, etc. For instance the EMS reacts to requests of appliances, or activates an SA only if enough power is available from the solar panel or public energy supply rates are low.

- Remote S&C (RS&C). The user remotely logs on into his EMS via the Internet, using a REMS, say, a cellular phone or a remote PC which can be in an internet cafe. This offers the user access to (a large part) of the functionality of the native EMS. The history of actions is stored by the EMS (at home, not remotely) in a log.

Inside a household a variety of people live, whose routines differ from the ones of people in other households. Energy consumption patterns are closely tied to a persons' habits and preferences, as the SAs adapt the energy levels accordingly. Free time activities such as watching TV, but also charging the battery of an electrical vehicle can change the consumption volume significantly. Metering and energy generation devices (like solar panels) are installed by the MPO and calibrated yearly. Once installed, devices are left unattended at the disposal of the customer, unless exceptional behavior is detected. In an interesting variant of the scenario two households are in the same building, and the SMs are placed in a common room as the basement.

Suggestions for worst cases, attackers and threats

The following questions should help to analyze the environment and more precisely, to identify the main security requirements, the possible threats (not only from external attackers but also from insiders or normal participants of the system), unwanted situations (here referred as worst cases, although they can be merely unwanted), and possible requirements for in depth security .

- Question 1: Imagine a household inhabited by a family with children inside the premises described in Section 4.1.2. For an attacker, how easy can it be to know which and how many people are in the house? What smart-metering related data would he need? How can an attacker trace individuals inside by exploiting the SM infrastructure? What type of attacker can do this? Could the attacker use this information to plan or schedule a burglary? How could such an attack be possible? Is the household, in particular the children, by the misuse of, say, the billing data?

- Question 2: Nowadays, it should be assumed that different households have appliances of different types, providing different functionalities. In one household, the SMs, SAs and/or the home EMS monitor or control the entrance door locks or the windows, the oven, the microwave, electric appliances, etc. Furthermore, the same household may have a smart Android TV, which is a full-edged computer, with all its possible common vulnerabilities. Appliances are connected to the Internet (via the gateway) and may receive software and key updates, system control messages and send status reports. Do SAs introduce new threats to the SM infrastructure? If so, which attacker models and threats are identifiable in this scenario? Which security and privacy requirements have to be set?

- Question 3: The scenario initially assumes that all the communication is encrypted. But even if this is true, is the encryption of all communication enough to guarantee privacy requirements against insider and outsider threats?

- Question 4: Are impersonation attacks viable in the SM infrastructure? Could, in some cases, a customer impersonate another customer? Could an attacker impersonate a server? What could happen? In which cases might this become critical?

- Question 5: The scenario assumes that every communication is encrypted. How should the communication encryption be managed? Who chooses the keys? When and how are the keys communicated to the relevant parties? At least one implementation choice may be relevant: should the system use shared symmetric keys or asymmetric ones? Other choices could be: should the encryption be at the lower layers (network or transport layer) or at the application layer? How should the keys be secured inside the devices? (And secured against whom?) Considering all the parties involved and their possibilities to malfunction (un-)intentionally, it may be important once again to take into account the flexibility and cost effectiveness of the solution, such that it can be realized in the large scale of the SM infrastructure.

- Question 6: Electric mobility and the vehicle charging infrastructure will be an integral part of the future smart grid. Imagine that vehicle charging stations and vehicles exchange unique vehicle IDs (uvID). Which security and privacy breaches could result from this approach? Which requirements should be set to avoid breaches?

4.2 Standard solutions for securing the CI

In the context of smart grids there are not universally available and recognized solutions to solve cybersecurity problems. This is because smart grid technologies introduce many new components to the electric grid. So, controlling the global infrastructure is a very hard task. Moreover, many of these components are developed by different producers with different quality standards. This makes it very hard to ensure the interoperability and reliability of global systems. Note also that bidirectional communication occurs on normal networks. Thus security requirements such as confidentiality, integrity, and availability (CIA) of information must be ensured when this new communication model is used. Also, the resilience of critical infrastructure can be strongly undermined by the dependencies between infrastructure components, processes and procedures. The understanding of complex infrastructure interactions, their dependencies and the implications of these dependencies is therefore important for achieving resilient systems, both when designing them and when dealing with a crisis. Power grids, being among prominent representatives of critical infrastructure, have been, and are still currently, the subject of numerous studies and initiative to deal with the interdependencies problem, especially following the major incidents that occurred in the last decade (already described at the beginning of this chapter). In fact, most incidents in this sector have had serious consequences due to the presence of dependencies, which have amplified the phenomenon of cascading failures. For example, most major power grid blackouts which have occurred in the past were initiated by a single event (or multiple related events such as a power grid equipment failure that is not properly handled by the SCADA) that gradually leads to cascading failures and eventual collapse of the entire system. In the NIAC 2009 Report and Recommendation, the concept of infrastructure resilience is introduced as the ability to reduce the magnitude, impact, or duration of a disruption. Among the recommendations to favor resilience, there is emphasis on the need of understanding real-time interdependencies and the expectations and limitations of interconnected sectors, to minimize unforeseen circumstances. Approaches to the analysis of critical infrastructure interdependencies include primarily a range of modeling, simulation and analysis techniques. A review on research in infrastructure interdependency modeling and analysis can be found in [55], while some specific studies addressing the power grid sector are in [73, 43, 78, 150, 153, 62, 49]. The NIST [21] provides

key concepts and assumptions that are the foundation for the logical security architecture.

- Defence-in-depth strategy: Security should be applied in layers, with one or more security measures implemented at each layer. The objective is to mitigate the risk of one component of the defense being compromised or circumvented. This is often referred to as defense-in-depth. A defense-in-depth approach focuses on defending the information, assets, power systems, and communications and IT infrastructure through layered defenses (e.g., firewalls, intrusion detection systems, antivirus software, and cryptography). Due to the large variety of communication methods and performance characteristics, as well as because no single security measure can counter all types of threats, it is expected that multiple levels of security measures will be implemented.

- Power system availability: Power system resiliency to events potentially leading to outages has been the primary focus of power system engineering and operations for decades. Existing power system design and capabilities have been successful in providing this availability for protection against inadvertent actions and natural disasters. These existing power system capabilities may be used to address cybersecurity requirements.

A solution, widely accepted by the scientific community, is to use the information provided by wide area monitoring systems (WAMSs) to monitor the transmission grid and to prevent the spread of disturbances. WAMSs make use of devices distributed throughout the power grid that measure the key parameters for detecting anomalous conditions. Today PMUs are the most commonly used devices in WAMS. In particular, PMUs are devices that perform measurements of real-time phasors of voltages and currents to provide information about power grid status. Time synchronization between different PMUs is required to understand the global status of the power grid at the same time. This is because events occurring in one part of the grid affect operations elsewhere, and they also extend to other systems beyond the grid that rely on stable power. Time synchronized measurements produced by PMUs are called synchrophasors. In order to obtain simultaneous measurements of phasors detected from different PMUs installed across a wide area of the power system, it is necessary to synchronize these times, so that all phasor measurements belonging to the same time are truly simultaneous. Each PMU uses a GPS receiver [81] to take a unique timestamp within the global system. By providing real-time information on stability and operating safety margins, WAMS give early warnings of system disturbances for the prevention and mitigation of major blackouts. The continuing presence of existing measurement devices and the overlapping visibility of individual PMU lead some utilities to argue that WAMS are not cyber critical systems as defined by NERC CIP

[138] and therefore PMU and PDC need not be treated as cyber critical assets (CCA). WAMS data will also be used for new visualization systems and stored in a secure way for post incident analysis [116]. WAMS may also include power system event classification applications such as semantic driven knowledge discovery algorithms.

System integrity protection schemes (SIPS) [131] are installed to protect the integrity of the power system or strategic portions, as opposed to conventional protection systems that are dedicated to protecting a specific power system element. SIPS require multiple devices (actuators and detectors) installed over a wide area that communicate through network infrastructure. This control scheme is useful for detecting changes in load, generation, or system configuration and for attempting to take control actions to maintain system stability.

Intrusion detection systems (IDSs) [98] within the electric grid have gained significant attention in recent years. Two types of IDS are used and they perform anomaly or misuse detection. IDS effectiveness is demonstrated in [72] where the authors have identified malicious events within control systems by focusing on known, static network communication patterns. Additionally, [52] has demonstrated how specification based intrusion detection methods can be leveraged within AMI deployments to detect malicious communication patterns.

Security event and management (SIEM) systems are already widely adopted to protect the critical infrastructure, so they can also be used to protect the smart grid. The power of a SIEM system is that it analyses and correlates different events provided by many information sources in order to detect cyber attacks. SIEM architecture is composed of three main components: sensors, a server and a storage system. The sensors are deployed into infrastructure for monitoring purposes. The main goal of sensors is to gather events and send them to the server. Smart sensors proposed in literature, e.g., [151] allow: the gathering of syntactically heterogeneous event formats in order to process data generated from multiple layers of the infrastructure; the correlation of multiple layer data based on different semantics, e.g., not only IP addresses, port numbers, protocol types, payload signatures; the processing of data at the edge of SIEM architecture, for filtering out micro events generated by the infrastructure. The server performs complex correlation of events provided by different sensors in order to discover new attacks. The storage system stores alarms and events generated. The storage system is a very important component of the SIEM system. In fact, the analysis performed in on-line and off-line mode by a server or any analysis tool is correct if the integrity and unforgeability of data is ensured. So the storage architecture must be designed to ensure integrity and unforgeability of data even when some components of storage architecture are compromised [34, 35].

4.3 Types of attacks and exploited vulnerabilities

In [99], WAMS systems and their main vulnerabilities are described. The GPS system represents one of the most important vulnerabilities of WAMS systems that affects PMU devices. In fact each PMU uses a GPS receiver [81] to take a unique timestamp within the global system. However, GPSs are subject to three primary sources of interference: blocking, jamming, and spoofing. Jamming and blocking are the processes of generating noise signals that concatenate with GPS signals generating new signals that the receiver cannot understand [84, 148]. These types of attack are recognizable because the goal is to deny a specific offered service, timing in this case. GPS spoofing [94] is the process of feeding the GPS receiver with false information so that it computes an erroneous time or location. This type of attack is complex to detect because the GPS signal is forged in order to mislead the GPS receiver that uses it. GPS spoofing was discovered and highlighted in 2001 by the US Department of Transportation during a study performed on vulnerabilities of the transportation infrastructure that uses the GPS signal [11].

The first step needed to perform a GPS spoofing attack is to acquire and to track the GPS signals to obtain a reference signal. Then a forged signal is generated and summed to the original GPS signal. The new signal is used to synchronize the spoofed signal with the authentic signal received. So the attacker produces a signal perfectly aligned with the authentic signals but with lower power. The generated spoofed signal is comparable to the noise of the target receiver in terms of power. Then the attacker increases the power of the forged signal until it overcomes the authentic signal. In this way, the forged signal shows higher Signal-to-Noise Ratio (SNR). So, the GPS receiver tracks the fake GPS signal (instead of the authentic signal) due to its higher SNR. After that, the attacker has successfully taken control of the GPS receiver. Then he slowly moves the spoofed signal from the authentic signal. The GPS signal received is considered to be completely captured when the spoofed signal is delayed by 2 microseconds from the authentic signal as described in [157]. Thus the attacker could increase the time delay until the PMU is desynchronized. If an attacker forges the timestamps provided by GPS to a PMU, it could cause variations in measured phase angles. The difference in the phase angle between two PMUs indicates that the power between the regions measured by each PMU has changed. These variations could compromise the stability of the system in such a way that grid operators or automatic response systems would make incorrect decisions like powering up or shutting down generators. Incorrect decisions can cause blackouts or damage. In [98] possible attacks that can be performed in order to compromise a smart grid infrastructure are described. These attacks include:

- Protocol attacks: The network protocols used in the power system, such as ICCP, IEC 61850, and DNP3, could be potentially exploited to launch

cyberattacks if they are not secured properly. Since these protocols are used to control remote devices and substations, once an attacker is able to gain network access they could manipulate the communications to inject malicious system state and controls. Therefore, the grid requires secure versions of these protocols that not only provide security guarantees, but also meet the required latency and reliability guarantees needed by the grid applications.

- Routing attacks: This refers to cyber attacks on the routing infrastructure of the Internet and other wide area networks. By manipulating the routing of packets, attackers could perform man-in-the-middle (MITM) attacks, spoofing, or delaying the delivery of the authentic traffic. A massive routing attack could have consequences on real-time operations of the grid and on real-time markets that rely on wide area communication.

- Intrusions: This refers to exploiting vulnerabilities in the software and communication infrastructure of the grid which then provides access to critical system elements. Network intrusions are of specific concern due to recent reports identifying numerous weaknesses in software and networks used in the utility industry. An example intrusion scenario is gaining access to the control station bypassing security controls (firewalls, system passwords).

- Malware: This refers to malicious software that exploits vulnerabilities in system software, programmable logic controllers, or protocols. The malware generally scans the network for potential victim machines, exploits specific vulnerabilities in those machines, replicates the malware payload to the victims, and then self-propagates. In recent years, malware attacks are growing in numbers and sophistication (e.g., Stuxnet), and this has been a source of major concern for critical infrastructure systems including the power grid.

- Denial of service (DOS) attacks: A DOS is any attack that denies normal services to legitimate users. This could also mean denial of control or observability in the power grid's context. These attacks are typically created through massive resource exhaustion attacks that flood the communication network or the server with huge volumes of traffic or spurious workloads, thus denying service to legitimate users.

- Insider threats: The electric grid also faces risk from insider threats, such as those identified by the NERC HILF report. A malicious insider with access to a control system network could easily abuse their trusted status to install malware or directly inject malicious commands into the network. Malicious insiders are especially dangerous because they also

possess detailed knowledge about the system topologies and operations, and therefore could easily design an attack scenario that causes the system to operate outside safe operating points.

Other possible attacks concern AMI, indeed several new security challenges are introduced in this context [122, 125, 51]. The core elements of AMI are smart meters, which strongly differ from traditional meters in that they allow remote control of power consumption and demand. However, since smart meters have to be installed in each customer's building, AMI will consist of billions of such devices which, due to the huge size of the area to be covered, must be cheap (often at the expense of quality). Moreover, they will be placed in physically insecure locations and under the control of often disinterested, unsophisticated, or sometimes malicious users. For this reason, even though each meter is subjected to rigorous examination before the installation, once they are placed in the customer's premises they can no longer be consider reliable. Basically, the risk associated with using smart meters is related to both the fact that they were not built according to some security policy and that they are accessed remotely through disparate communication technologies. Often, the network infrastructure underlying an AMI has a mesh topology relying on various wireless networking media and protocols, such as WiFi, cellular, WiMAX, satellite communications, etc. This introduces further security problems in that an attacked meter may propagate malicious code to other meters in its neighborhood. Researchers have already shown that smart meters are vulnerable to attacks that may result in power disconnections, energy usage frauds, etc. In particular [125] describes different typologies of attack that can be performed to defraud the electrical grid by manipulating AMI systems. The authors of the report demonstrate that energy theft is still possible in AMI systems and that current AMI devices introduce new possibilities for achieving this goal. Smart meters are equipped with new anti-tamper solutions, however while these solutions are enough for ordinary honest people, they do not prevent malicious users from circumventing them. The approach described in [125] relies on the manipulation of the demand data. Basically, there are three ways to manipulate such data, each tailored to a specific state of data during the measuring: (i) the data is being stored in the device, (ii) the data is archived in the device, (iii) the data is traveling through the network to reach the utilities that manage power usage. On the basis of the particular state the attacks rely on, the latter can be classified in three corresponding categories. The first and second categories of attacks require access to the device in such a way as to overwrite the meter's firmware (first category) or to modify stored data (second category), this task is very hard and requires intense reverse engineering. The attacks belonging to the third category work by injecting modified values into communication between meters and utilities. Moreover, since the information sent by several meters is often collected into collector nodes located between meters and utilities, attacks on this side of

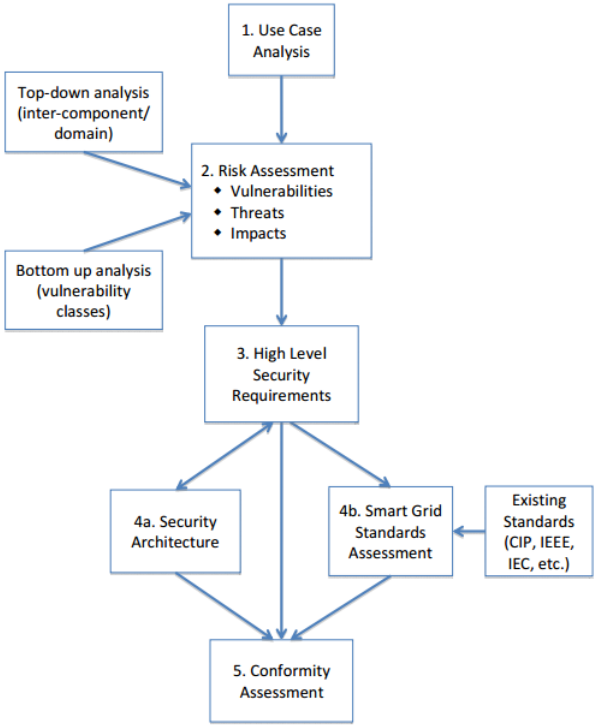


Figure 4.4: Tasks in the Smart Grid cybersecurity Strategy.

the network allow the modification of a large number of demand data thus increasing the damage. In [51] different, real-world smart meters are considered and analyzed from a security point of view. Each meter is seen as a sensor so that the overall AMI can be seen as a sensor network. Communication occurs through a low rate wireless personal network with a multi-hop routing scheme. This choice is low cost but increases the surface area to attack. In this setting, several attacks could be carried out as, for instance, black hole, gray hole and sybil attacks. The aforementioned authors focus on black hole attacks and show some precaution meters vendors should adopt to avoid this kind of attack.

4.4 Protection Strategies

NIST study [21] has defined some tasks to be performed in order to make smart grid secure. Implementation of a cyber-security strategy requires the definition and implementation of an overall cybersecurity risk assessment process. Risk is the potential for an unwanted outcome resulting from an incident, event, or occurrence, as determined by its likelihood and the associated impact. The smart grid risk assessment process is based on existing risk as-

sessment approaches developed by both the private and public sectors and includes identifying assets, vulnerabilities and threats and specifying impacts to produce an assessment of risk of the smart grid and to its domains and subdomains, such as homes and businesses. As the smart grid includes systems from the IT, telecommunications, and electric sectors, the risk assessment process is applied to all three sectors as they interact in the smart grid. The tasks that have been identified and should be performed in the implementation of the cybersecurity strategy are shown in Figure 4.4 and are detailed below.

- Use case analysis: The set of use cases provides a common framework for performing the risk assessment, developing the logical reference model, and selecting and tailoring the security requirements.
- Risk assessment: Risk assessment includes identifying vulnerabilities, assets and threats. Two approaches are possible: top-down and bottom-up analysis. The bottom-up approach focuses on well-understood problems that must be solved e.g., intrusion detection system for power equipment, users authorization and authentication in order to access to substation control. The top-down approach focuses on a logic model that must be ensured at architectural level. The output of the risk assessment phase is useful to select security requirements that must be ensured.
- High-level security requirements: For the assessment of specific security requirements and the selection of appropriate security technologies and methodologies, both cybersecurity experts and power system experts are required. The cybersecurity experts have a broad awareness of IT and control system security technologies, while the power system experts have a deep understanding of traditional power system methodologies for maintaining power system reliability.
- Security architecture: Secure smart grid architecture is designed and developed in agreement with requirements described in the previous steps.
- Smart grid standards assessment: In this phase the standards that have been identified as potentially relevant by the Priority Action Plan (PAP) teams are evaluated. This process highlight the gaps between security requirements and the standard identified. Also recommendations will be made for addressing these gaps.
- Conformity assessment: The last task is to define a conformity assessment program for security.

4.5 Fault mitigation approaches

Smart grids are usually equipped with a Wide Area Monitoring System (WAMS), often used to avoid catastrophic failures, such as, blackouts due to overloads.

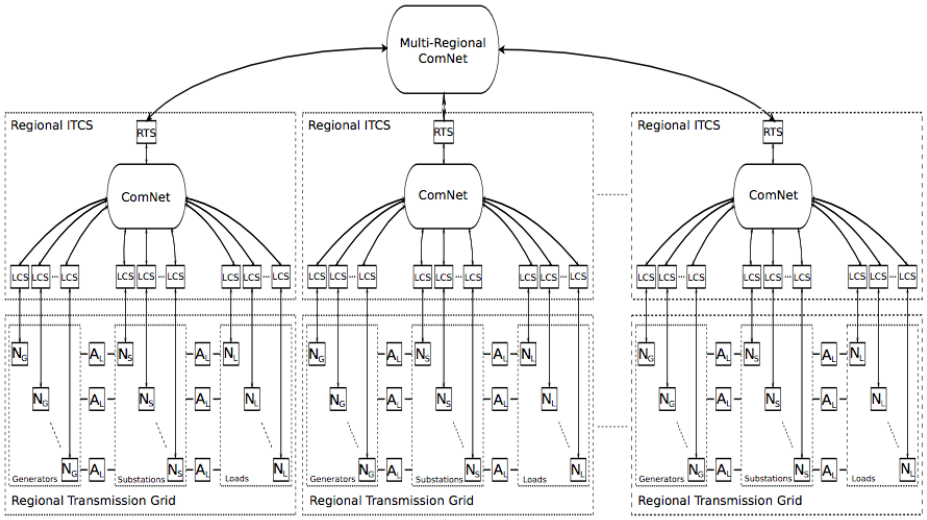


Figure 4.5: High-level model of a electric power grid.

WAMS uses different PMUs installed in different locations in order to perform distributed measurements within a power grid. PMUs are devices that perform measurements of real-time phasors of voltages and currents to provide information about the global status of a power grid. Time-synchronized measurements produced by PMUs are called synchrophasors. Each PMU uses a GPS receiver to take a unique timestamp within the global system. So it is possible to use different phasors provided by different power grid islands to check the correct behavior of global power grid and to avoid overloads that generate failures. In particular, the strategy proposed [160, 161] avoid failures by analyzing the variation of phasors detected by different PMU in the same time interval. In fact, if the phasors diverge, one or more transmission lines of a power grid are overloaded and the risk related to a possible blackout is increased. When the monitoring systems detect that a difference between phasors exceeds a threshold, then they notify a control station about the anomaly. The control station can then automatically reconfigure the power grid in order to avoid failures.

Other efforts to improve smart grids' dependability focused on the application of fault forecasting, especially to assess the exposure of electrical power grids to escalating and cascading failures due to the inherent interdependencies between electrical and information infrastructure. Fault forecasting allows for the analysis of critical scenarios in which internal or external faults in a segment of the information infrastructure provoke a serious impact on the controlled electric power infrastructure [49, 73]. To this aim, interdependencies between entities of the infrastructure can be represented using a model, that allows for the description of an infrastructure in terms of entities and

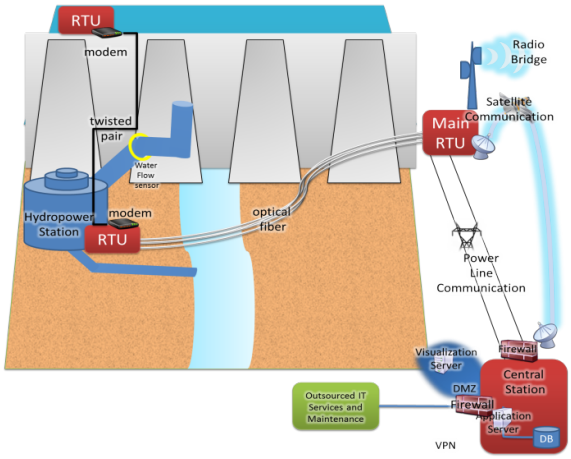


Figure 4.6: Hydropower Station Monitoring.

interdependencies from a high-level point of view (see Figure 4.5). From this high-level model, a stochastic model can obtained for analyzing in a quantitative way the impact of various types of failures that can occur in the presence of accidental and malicious faults in these entities. For instance, the likelihood (in terms of probability) of potential failure propagation phenomena among entities can be evaluated using such a modeling approach.

The case of a DAM control station. Infrastructure, such as dams or power generators is monitored by a control station (Figure 4.6). Generally, a SCADA system is used to monitor this type of infrastructure.

Several sensors are installed at different locations in order to perform measurements. The sensors are connected to Remote Terminal Units (RTUs), reading the measurements of analog and digital signals, and sending them to the main RTU. RTUs communicate via RS-232, Ethernet, optical fiber, GPS or GPRS. The main RTU sends the information gathered by RTUs to the control station in order to evaluate the security/safety of the DAM. The main RTU uses the power line in order to send information to the control station. In this scenario, the power line is a very important asset of the IT infrastructure. In fact, the loss of the power line for any reason could cause the interruption of data transmission.

A technique to avoid communication failures is based on the concept of redundancy. In particular, the transmission data system is designed in order to use more than one channel of communication. In this context, the main communication channel is the power line. If the channel impedance is infinite, the communication cannot be performed. The main RTU detects this issue and establishes a secondary channel, based on satellite communication, to

send the same information. In this case, the backup communication system is considered to be a cold replication mechanism.

4.6 Open Problems

Network approaches used in modern smart grids are limited to analyzing the network traffic in order to detect, mitigate and eliminate DoS attacks. However, valid countermeasures against integrity and confidentiality attacks are often only provided as an add-on to existing systems and this can result in weaknesses resulting from the integration of the heterogeneous components (typically provided by multiple vendors). The impact of classic cryptography algorithms on network performances should be analyzed, so as to assess the trade-off between security and timing constraints in order to adopt the most appropriate authentication schemes in the smart grid.

As already discussed, GPS spoofing attack represents a real threat to smart grid security. Today, several techniques have been proposed in order to detect such attack [177] attack. Such techniques are mainly based on:

- monitoring the absolute GPS signal strength. This technique is based on comparisons between the observed signal strength and the expected signal strength. If their difference is greater than a fixed threshold, an alert is generated.

- monitoring the signal strength received from each satellite. The idea is to compare the observed signal strength with the expected signal strength for each satellite. The attacker will generate a forged signal of equal strength for each artificial satellite through the GPS satellite simulator. Instead, the signals provided by real satellites will change over time for each satellite. So, an alert is generated if the signal characteristics are constant over time for each satellite.

- monitoring the relative GPS signal strength. This technique implies that the average signal strength is recorded and compared periodically. An alert is generated if a large change in relative signal strength is detected.

However, remediation techniques are not available to avoid PMU failures.

Since smart metering systems are new, no comprehensive database of attacks exists for them. Therefore, it is very important to design and implement a detection and diagnosis system capable of detecting unknown attacks accurately. In particular, approaches based on both signature and anomaly techniques could be considered to protect smart metering. Furthermore, the AMI communication network is not homogeneous and different technologies and protocols are used to allow the exchange of messages among AMI components. In order to address such heterogeneity it could be useful to realize a system capable of gathering security-relevant information from any crucial

component of the infrastructure and to analyze multi-protocol information in order to discover new attacks.

As described above, AMI generally consists of billions of smart meters placed in physically insecure locations, potentially vulnerable to cyberattacks. The significant size of this type of network, as well as the necessity for meters and control components to continuously exchange messages for data acquisition, control and supervision of the devices, leads to this scenario being considered very similar to that of a traditional SCADA system. By using distributed electronic controls and sensors to perform batch or repetitive tasks, SCADA alerts the operator if some system component needs attention or has exceeded pre-set parameters. While the first SCADA systems held all operations in one computer (generally a mainframe) and SCADA functions were limited to only monitoring sensors, the later SCADA systems use a distributed architecture since they often share control functions across multiple smaller computers. Moreover, if in the first distributed SCADA systems the nodes were connected by local area networks, current SCADA systems are usually networked, communicating through wide area networks, and often the clients can access the system using Internet and the Web, also via a wireless connection. Currently, SCADA systems are built upon a wireless sensor network . In such a context, security concerns are key issues when developing SCADA applications, as they are vulnerable to cyberattacks. Observe that these two levels of components in SCADA systems, namely, (i) the underlying sensor network and (ii) the high-level of supervision and control, are not equivalent from the viewpoints of security problems and possible countermeasures. For example, a SCADA system may collect a number of measures from different sensors, and may elaborate them, so that a simple detection of outliers in the temporal sequences of sensors is not immediately extensible to the other high SCADA levels, since the misbehavior of a higher level SCADA node could be much more complex with respect to that of a sensor node. Therefore, the aforementioned approaches working at sensor network level are not immediately applicable to the whole SCADA system. These issues arise also in the case of a smart metering system, in which the supervision and control components can be implemented as a SCADA system. On the basis of the above considerations, it is possible to analyze the possibility of applying a technique for detecting compromised nodes in a smart metering system, exploiting a trust-based strategy derived from the research into social agents. This possibility has been already explored in SCADA systems. For example, some architectures to stimulate a correct routing behavior have been proposed [58]. In this approach, each node receives a per-hop payment in every packet it forwards. Nodes store this payment information in an internal counter and this information is shared by nodes that directly interact, introducing a cooperation element in the security mechanism. This type of approach [118] has proposed the mitigation of routing misbehavior by detecting non-forwarding nodes and rating every path so those nodes are avoided when the routes are

recalculated. In this way, the non-routing nodes are not included in routing paths, since they are not going to cooperate, but they still can ask others to forward their messages. This scheme detects the misbehavior but it does not isolate it. Other versions of this approach [58] introduce a secure routing protocol, sitting over the chosen routing protocol which makes misbehavior less attractive for the nodes than proper routing. In particular, the nodes watch their neighbors for bad behavior and take this behavior into account using a local reputation system. Moreover, the nodes can also inform their trusted neighbors of misbehaving nodes. It is also possible to use a global (instead of local) reputation system [133]: some reputation information about the sensor nodes is transmitted all over the network, in order to warn all nodes about misbehaving nodes. In this approach, the compromised nodes are detected and isolated, since their reputation is rated as low. In any case, all the above approaches work at the level of sensor networks, without involving the high-level nodes. Designing a detection strategy that takes all the nodes of a AMI system (not only the meters) into account, can introduce two main advantages: (i) it is possible to immediately detect attacks that are directly addressed to high-level AMI nodes; (ii) it is possible to use different software components for high-level nodes (that are generally fully-equipped PC) and for lower-level nodes (that have limited resources).

Communications security involves the design of a key management scheme. However, the definition of a suitable key management scheme for smart grids is still an open issue. One of the main stumbling blocks is that devices in a smart grid system usually come with limited storage, low power and bandwidth, which require that the key management scheme should be efficient and flexible. From this standpoint, solutions based on PKIs are still far from being mature and effective for several reasons. First of all, the need for a long validity time of a certificate clashes with the need for a CRL of manageable size [182]. Furthermore, in resource-scarce devices, a PKI solution may create a conflict between the interoperability and scalability requirements. Indeed, in order to be largely interoperable, a device should, at least in principle, store the certificates of all possible certification subjects (e.g., manufacturers, distributors, or even users). The number of these certificates may turn out to be too large for the device's limited storage resources. A relevant example of this conflict is intrinsic to the ZigBee Smart Energy Profile. In order to overcome this problem, Dini et al. propose a new form of local cross-certification [80].

In the field of PKIs, ID-Based Encryption (IBE) may be particularly attractive for smart grids as it can be deployed without prior configuration, as the identity (ID) of a device is used to generate unique keys. This allows for the easy deployment of low powered devices such as sensors because they may start sending secure messages without the need to contact a key server. Alternatively, symmetric encryption could be used. For the symmetric key method, it is risky if all of the devices use the one same preloaded key since if one of these devices is compromised the attacker could know every device's shared

key. Instead of a preselected key, it is better to set up a trusted third party to distribute the shared key between two parties. The Kerberos system can be used in this environment to distribute the key for the components in the smart grid. However, smart grids present unique peculiarities with respect to the distributed computing environments for which Kerberos was originally conceived. The key distribution centre (KDC) in the Kerberos system cannot support the keys distribution when network or power outages occur. More importantly, it is too expensive and insecure to have a back-up server for the KDC considering the size of the smart grid. In general, a mixture of hierarchical, decentralized, delegated or hybrid security schemes may be feasible. Recent and relevant examples of this kind of key management schemes are [182, 109, 181]. Preferably, a candidate scheme should include secure bootstrapping protocols, i.e., it should provide effective means to initialize new devices. Furthermore, critical security operations, such as key updates, should preferably employ group key management techniques. Once again, the ZigBee Smart Energy Profile does not adequately address these requirements and, in particular, fails to fulfill the forward security requirement [182]. In fact, upon leaving the system (a node may be dismissed, sent to maintenance, lost, compromised, or supposed so), a node still remains able to access communication because the on-board keying material is not properly revoked. If not properly revoked, an adversary may exploit the keys on the device to mount severe attacks against system integrity and user privacy.

Smart grids introduce interconnections of previously independent subsystems. As an example, the smart meter sub-system offers current energy consumption monitoring information used by energy balancing controllers operating at medium-voltage (MV) and low-voltage (LV) levels. These interfaces provide new entrance points for malicious attacks. Furthermore, attacks can be based on affecting the interplay between these subsystems and hence may be very difficult to detect within the individual domain. Thus, detection and protection approaches have to consider the interaction of the previously independent sub-systems [21].

Smart grids contain a large number of sensor and actuator components as well as processing and interconnection systems [87]. In addition to the problem of scale, not all of the components are deployed under the control of a single stakeholder. Hence, security solutions for smart grids cannot rely on mandatory deployment of functionalities on all grid elements, but rather solutions are required to infer trust levels of components and information even if these cannot be modified, but only observed.

As smart grids introduce additional intelligence and not all malicious attacks can be prevented, reactive countermeasures need to be introduced that may need to rely on disabling the use of certain information and components if these are suspected to not be trustworthy. The detection of such scenarios and the support of fall-back operation modes therefore need to be supported. Though these will depend on the specific smart grid control application, there

can be supporting functionality for detection and reconfiguration; this motivates an approach which provides a security middleware with adequate interfaces that can be used by smart grid applications.

The electricity industry is lacking metrics for quantitatively evaluating smart grid security. Without suitable security metrics it is difficult to assess the effectiveness of deployed security mechanisms. Such metrics will also help to evaluate the cost efficiency of different solutions and provide criteria for investment.

Concerning interdependencies understanding and management, a variety of models and simulation studies have been developed. However, when comparing requirements for resilience assessment in modern and future grid infrastructure with existing approaches, it becomes clear that further research is still required. Most of the modeling research in CIs (including power grids) uses handcrafted reliability block diagrams, fault-trees, or stochastic Petri nets. The application of stochastic methods captures the continuous dynamics of the physical world and the discrete characteristics of the control infrastructure. However, further research is necessary to ensure the scalability of hybrid approaches. Accounting for heterogeneity, flexibility and dynamicity of modern smart grids with heavy penetration of distributed energy resources and renewable energy resources calls for advanced modeling efforts, possibly requiring a combination of different formalisms/techniques to describe the various components of a system and their dependencies. Heterogeneity also needs to be addressed at the level of vulnerability exposed by the different subsystems composing a critical infrastructure and to be properly represented in the model, e.g., the use of subsystems, such as wireless SCADA, which are known to be typically vulnerable to error and misuse. In fact, advances in technology and SCADA systems have enhanced critical sector operations but created additional vulnerabilities, which must be analyzed and addressed to adequately protect the critical infrastructure.

CHAPTER

5 Transportation

Transportation systems move people and goods within a country and between countries. The security and safety of these systems have always been important but became critical after the terrorist attacks of 2001. Again in 2004 and in 2005, Madrid and London were targets of attacks that involved the public transportation system. The transportation sectors includes severals subsectors, for example aviation, highway, shipping. This chapter focuses on three of them: air traffic control (ATC), the maritime transportation system (MTS) and the railway system.

5.1 Air traffic control

This section is organized as follows: 5.1.1 provides a brief introduction to ATC infrastructure; 5.1.2 discusses standard solutions for securing the ATC infrastructure and related open problems; finally 5.1.3 explains the open issues (also considering the related attacks) and discusses some possible solution.

5.1.1 Description of the critical infrastructure

An ATC system is a typical software-intensive mission-critical system, which plays a key role in air traffic management (ATM) [2]. It provides facilities and services to ground controllers and pilots for managing safely ground and en-route flight operations, with the aim of preventing collisions, organizing the flow of traffic, and providing support information to operators and pilots. From an architectural perspective, the design of an ATC system is divided into two major subsystems: en-route and terminal area. The en-route subsystem is devised for aircraft moving along the airway network, generally cruising at higher altitudes. In Europe, for instance, en-route ATC is segmented in

several Area Control Centers (ACCs), each responsible for a defined portion of the air space, with ATC systems in ACCs cooperating to guarantee the safety of flights. The terminal area subsystem handles aircraft flying at lower speeds and altitudes as they arrive at and depart from airports. It must also control, through Instrument Flight Rules (IFR), the traffic that is passing through a terminal area without landing. IFRs are a set of rules that permit the airplanes to fly in adverse conditions, such as in presence of obstacles and other airplanes.

Main stakeholders and players

Production and usage contexts of air traffic control systems involve several stakeholders and players. According to EUROCONTROL[1], the involved stakeholders are:

- Airspace users: Airlines, pilots, aircraft operators and passengers;
- Air navigation service providers: They are in charge of organizing and managing the flow of traffic in the air and on the ground in a dedicated airspace;
- Airports
- National and international aviation regulators: National supervisory authorities, and international regulators, such as the European Aviation Safety Agency;
- The aeronautics industry: Including manufacturers of aircrafts, avionics (aviation electronics) and air traffic management infrastructure (radio antennas and satellites for instance);
- International aviation organizations: Organizations such as the United Nations International Civil Aviation Organization (ICAO) or the European Civil Aviation Conference (ECAC).

Such a number of involved entities imposes an extremely high attention to possible security threats and risks that would have an immediate impact primarily in terms of safety, but also in economic terms.

Requirements

The ATC system is designed with a component-based approach; it has tens of thousands of requirements and it consists of many interacting Computer Software Configuration Items (CSCIs). The major equipment components that

[1] `http://www.eurocontrol.int/articles/stakeholders`

support these ATC facilities are surveillance radar, airborne transponders, navigation aids, computers, and communication links.

The main requirements of a modern ATC system includes: *dependability, robustness, and security* in order to prevent (malicious or non-malicious) threats from causing failures with a potentially disastrous impact, and to prevent failures in case of unexpected conditions; *performance*, which is increasingly required due to the more and more intensive traffic and to the consequent need of higher scalability (e.g., the ability of elaborating more and more flight data plans in a short time); *interoperability*, implied both by the complexity and size of the overall system including several CSCIs, and by the necessity to interact with other ATC systems across Europe, which is also a need felt by stakeholders; *maintainability*, to ease changes due to future integration/interoperability needs.

5.1.2 Standard solutions for securing the CI and open issues

The current air transportation system performs well, but it is susceptible to disturbances (e.g., due to weather) that can cause long delays. Moreover, the air transportation system is approaching its capacity limits. Without a transformation, the expected growth in air traffic will likely create costly flight delays and increased flight safety hazards.

The current ATM system in Europe is fragmented, which reduces efficiency and adds to the cost of flying. With over 40,000 daily flights a day predicted for 2020, the current ATM system will not be able to cope with this volume of traffic in an efficient manner.

The evolution of ATC systems starts from the concept of Single European Sky (SES), a set of legislative packages with the goal of creating a legislative framework for a unified European Aviation [8]. SES was born to organize the airspace shared by European countries into functional blocks. A primary goal of the SES project is the interoperability between European ATMs. In fact, according to Regulation (EC) No 549/2004 (the framework regulation), interoperability means a set of functional, technical and operational properties required of systems and constituents of the European air traffic management network (EATMN) and of the procedures for its operation, in order to enable its safe, seamless and efficient operation. Interoperability is achieved by making the systems and constituents compliant with the essential requirements. Regulation (EC) No 552/2004 (the interoperability Regulation) is focused on the interoperability of systems, constituents and associated procedures of the EATMN. It ensures that new validated concepts and technologies can be introduced timely and efficiently.

The following seven domains are addressed: (i) aeronautical information services, (ii) airspace management, (iii) air traffic flow management, (iv) air traffic services, (v) communications, (vi) navigation, (vi) surveillance, and (vii) meteorological information. The European Air Traffic Management Sys-

tem (eATMS) long-term program is being run to design a new generation of ATM/ATC systems that are compliant with the SES framework. eATMS goals include: i) optimizing system deployment and maintenance, ii) achieving the performance required to manage the traffic increase, and iii) converging towards interoperability with other European ATM systems as required by the Single European Sky ATM Research (SESAR) project [5]. The main eATMS non-functional requirements concern: (i) dependability and security, to provide continuous availability and integrity; (ii) robustness, to prevent failures in case of anomalous operating conditions; (iii) changeability, to support long-term evolution and integration/interoperability with other systems, as well as quick response to changes in operating environments; (iv) performance, to support the air traffic increases in European skies; (v) security, to prevent and counteract malicious attacks. In response to these growing concerns, the US Federal Aviation Administration Next Generation (NextGen) upgrade proposes a fundamental transformation that is intended to increase the capacity and safety of the air transportation system. The upgrade requires a fundamental transformation of the entire airspace system, including the incorporation of satellite-based technologies for surveillance operations to replace the legacy ground-based systems that are currently in use, as well as the upgrade of processing capacities of key en-route components responsible for the processing of flight data. Key components of the upgrade are the Automatic Dependent Surveillance Broadcast (ADS-B) system, and the Flight Data Plan Processor (FDP) system.

5.1.3 Types of attacks, exploited vulnerabilities (anatomy of an attack) and economic consequences

FDP is the (sub)system responsible for processing flight data plans, containing information such as the flight route, the current trajectory, airplane-related information, and meteorological data, and it is the system that predicts the direction of every aircraft based. At a high level, security threats may regard basically the communication between on-board systems and the ground stations; if exchanged information is altered by an intruder, consequences can be disastrous (e.g., the predefined trajectory is changed).

Security issues in ADS-B

The current air transportation system performs well, but, as noted above, is susceptible to disturbances (e.g., due to weather) that can cause long delays. Moreover, the air transportation system is approaching its capacity limits. Without a transformation, the expected growth in air traffic will likely create costly flight delays and increased flight safety hazards. In response to growing concerns, the US Congress established the Joint Planning and Development Office to manage the *NextGen upgrade*. The primary goal of the NextGen up-

grade is to significantly enhance the capacity and safety of air transportation operations. The upgrade requires a fundamental transformation of the entire national airspace system, including the incorporation of satellite-based technologies for surveillance operations to replace the legacy ground-based systems that are currently in use. A similar program has been launched in Europe within SESAR, the technological and operational dimension of the SES initiative to meet future airspace capacity and safety needs (see Section 5.1.2).

A key component of the upgrade is the ADS-B system. ADS-B provides continuous broadcast of aircraft position, identity, velocity and other information over unencrypted data links to generate a precise air picture for ATM. ADS-B incorporates surveillance techniques for precise aircraft tracking that replace antiquated capabilities. Indeed, the operational plans claim significant advancements in safety, efficiency and flexibility over the current airspace system infrastructure. ADS-B is designed to enhance air traffic control situational awareness, collision avoidance, surface runway incursion avoidance and air traffic control in non-radar environments (e.g., oceanic surveillance). Increased accuracy will allow tighter aircraft separation standards, higher probability of clearance requests and enhanced visual approaches, all of which will contribute to greater aircraft throughput. Additionally, ADS-B will result in more direct routings and optimized departures and approaches, which will increase capacity and save time and fuel. Finally, the ADS-B infrastructure relies on simple radio stations that are significantly cheaper to install and maintain than the mechanical infrastructure associated with traditional ground-based radar stations

ADS-B is designed to overhaul current air traffic surveillance techniques while providing new capabilities that would enhance ATM. ADS-B is automatic because it requires no pilot or controller intervention. It is dependent surveillance because an aircraft derives its own position from the global navigation satellite system. Moreover, it continually broadcasts aircraft position and other data to nearby ground stations, aircraft and surface vehicles (e.g., taxiing aircraft). ADS-B also affords improved accuracy over conventional radars (200 meters of precision compared with 300 meters at 60 nautical miles and with an accuracy that does not deteriorate as the range from the receiver increases.

In Europe, the implementation of ADS-B is part of the Implementation Package 1 (IP1), 2008–2013, of the SESAR ATM Master Plan. In Italy, ENAV has started a program, co-funded by the European Union, called *Programma di Implementazione Nazionale dell'ADS-B'* (National ADS-B Implementation Plan), within the Italy integration of communications and surveillance IP1 implementation plan, whose objective is the strengthening of the surveillance services by means of technologies based on ADS-B. The plan envisages the installation of fourteen ground stations all over Italy.

It is claimed that operational requirements necessitate the use of unencrypted data links and maintains that there is a low likelihood of malicious

exploitation. FAA conducted several analyses of the security aspects of ADS-B. The system was subject to certification and accreditation under National Institute of Standards and Technology (NIST) guidelines related to confidentiality, integrity and availability, and as well as other security goals. The FAA concluded that "using ADS-B data does not subject an aircraft to any increased risk compared to the risk that is experienced today" [89]. Moreover, the FAA believes that encryption would unnecessarily limit the international use of ADS-B.

However, the FAA report raises some major concerns from a security standpoint [121]. In particular, historical incidents have demonstrated that unencrypted data links can be exploited by a motivated adversary. As early as 2006, concerns were raised about the ability of hackers to introduce as many as 50 false targets on the radar screens of air traffic controllers [180]. In 2010, an iPhone and Android application (app) called Plane Finder AR was released that allows precise tracking of aircraft using ADS-B transmissions [135]. In 2012, Costing and Francillon demonstrated that attacks against ADS-B security are both easy and practically feasible, for a moderately sophisticated attacker [76]. Attacks range from passive attacks (eavesdropping) to active attacks (message jamming, replaying of injection). Finally, H. Teso has recently given a practical demonstration on how to remotely attack and take full control of an aircraft. ADS-B protocol was used during the discovery and information gathering phases [163].

For these reasons, deeper investigations about ADS-B security have been recently started. McCallie *et al.* have recently analyzed the security vulnerabilities associated with the ADS-B implementation and have provided a taxonomy of attacks including aircraft reconnaissance, ground station flood denial, ground station target ghost inject, aircraft flood denial, aircraft target ghost inject, and ground station multiple ghost inject [121]. The authors also examine the potential impact that the attacks may have on air transportation operations and provide recommendations that could enhance security if integrated into the ADS-B implementation plan. Furthermore, Finke *et al.* have explored the feasibility of employing format-preserving encryption, specifically the FFX algorithm, in the ADS-B environment [90]. The ability of the algorithm to confuse and diffuse predictable message input is examined using message entropy as a metric. Based on the analysis, recommendations are provided that highlight areas which should be examined for inclusion in the ADS-B upgrade plan.

However, crucial issues such as key management are still open. A single key leak compromises the entire system. Indeed, this is a major hurdle that must be overcome before considering the use of a symmetric cipher in a highly distributed system. However, a symmetric system is effectively employed by the military to encrypt Identification Friend or Foe (IFF) Mode-4 transmissions [107]. An ADS-B solution, designed to provide surveillance confidentiality, may be modeled following this example.

Security issues in flight fata processor (FDP)

The requirements of next generation ATC systems include improvements related to the communication between ground personnel and pilots, that consists in: (i) amending situational awareness of the pilots through the provision of the same kind of real-time air-traffic information as ATC controllers; (ii) preventing air-traffic conflict through detection and resolution; (iii) providing extremely accurate air traffic data for the pilots and the ground personnel. ATC systems rely on surveillance systems to share and broadcast flight data information among control personnel and pilots. Modern surveillance systems aim at improving safety of the air traffic management and control. Unfortunately, vulnerabilities in the surveillance and broadcasting system permit malicious attackers to exploit software components of the ATC systems, such as FDP. FDP is a system providing information such as the flight route, the current trajectory, airplane-related information, and meteorological data. It processes detailed aircraft information to predict flight plan profiles (e.g., the direction) of every aircraft. This information is combined with signals (e.g., radar track signal) for safety requirements such as medium and short term conflict alerts.

On-board systems interact with the Aircraft Communications Addressing and Reporting System (ACARS) to transmit messages with the ground stations. Meanwhile, the ground stations guide the aircraft during the mission and it is used for exchanging text messages between aircrafts and ground stations via worldwide transmission over radio (VHF) or satellite. Initially ACARS was used to automatically detect and report changes to the major flight phases referred as OOOI (Out of the gate, Off the ground, On the ground, and Into the gate). At the start of each flight phase, a digital message was transmitted to the ground containing the flight phase, the time at which it occurred, and other related information such as amount of fuel onboard or flight origin and destination. These messages are used to track the status of aircraft and crews. The industry started to upgrade the onboard maintenance computers in the 1990s to support the transmission of maintenance-related information in real-time through ACARS. This enabled airline maintenance personnel to receive real-time data associated with maintenance faults on the aircraft. The ACARS Management Unit was introduced to automatically perform all of the processing described above without flight crew intervention. Attacker can penetrate ACARS systems to send messages to airplanes on-board systems or changing the flight data plan managed by FDP (e.g., causing the deviation of the predefined path). Fallacies in the ACARS protocol are: (i) use of simple ciphers, (ii) exchange of very detailed aircraft information, such as a public database, local data and virtual aircrafts.

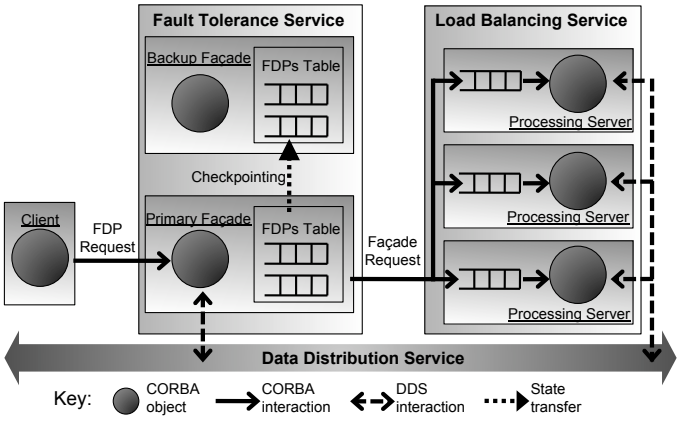


Figure 5.1: A simplified view of the architecture of FDPS.

5.1.4 Fault mitigation approaches

A relevant example of fault-tolerant architecture adopted in the ATC domain is represented by the Flight Data Processing System (FDPS) described below. The FDPS is a distributed software developed in C++ which uses CARDAMOM, a fault-tolerant CORBA-compliant middleware [75]. It is a part of an ATC system, in charge of managing Flight Data Plans (FDP). The goal of FDPS is to keep FDP up-to-date. For example, the FDPS has to analyze the actual position of aircrafts, retrieved from radar tracks, and update flight route, in order to efficiently allocate the flight space and to avoid flight collisions.

The architecture of the FDPS (Figure 5.1) is composed of a façade component, which acts as the frontend of the system and is replicated by the CARDAMOM Fault-Tolerance (FT) Service, and by a set of three Processing Servers (PSs), managed by the Load-Balancing (LB) Service. Service requests for inserting, deleting, and updating FDP are delivered to the façade through the middleware. The façade forwards requests to a specific PS according to a round robin scheduler. The selected server retrieves the specified FPD instance from a Data Distribution Service (DDS), compliant to the OMG Data Distribution Service standard [140], executes request-specific computations, and returns the updated FDP instance to the façade. Finally, the façade disseminates the updated FDP through the DDS, and replies to the clients.

The state of requests for each FDP is stored in an FDP table of the façade. The FT Service performs a *warm replication* of the façade process: the FDP table is checkpointed at each update and transmitted to backup replicas, which are activated in the case of failure of the primary replica. This warm replication mechanism, implemented in the FDPS, uses the CARDAMOM FT API. In the case of a failure of the main façade, such as a process crash, CAR-

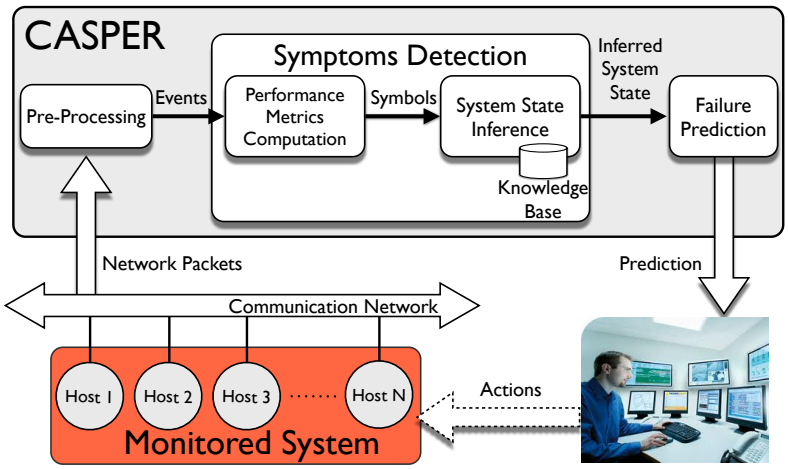


Figure 5.2: The CASPER failure prediction architecture.

DAMOM detects the failure and activates the replicated façade, which takes the place of the primary façade and assures service continuity.

The fault-tolerant architecture presented above reacts to faults once they have occurred and have been detected. In some cases, faults in the system can occur unnoticed, thus eluding fault-tolerance mechanisms and leading to system failures. Therefore, on-line failure prediction is a noteworthy approach to further improve the availability of mission-critical systems, such as ATC systems. It anticipates the occurrence of near-term failures during the system execution. When a failure is predicted to occur in the near future, the predictor produces an alert that allows for the timely triggering of some (human or automatic) recovery mechanism. Failure prediction is based on the run-time monitoring of symptoms, that is, out-of-norm behavior of system parameters that result as a side effect of faults in the system. CASPER (Figure 5.2) is an on-line failure prediction system which was adopted for failure prediction in ATC systems [48]. CASPER is a non-intrusive monitoring system, since it neither requires software to be installed nor logging in on the monitored system's hosts, thus avoiding privacy and security issues. Moreover, it is a black-box monitoring approach. It considers the monitored system components as black boxes, without requiring knowledge of the internals and logic of the system being monitored, and does not try to recognize causality paths among the boxes.

5.2 Maritime transportation system

This section is organized as follows: 5.2.1 provides a brief introduction to maritime transportation system (MTS) infrastructure; 5.2.2 discusses stan-

dard solutions for securing the MTS infrastructure and related open problems; finally 5.2.3 explains, using an example, a set of attacks that could be performed exploiting known vulnerabilities.

5.2.1 Description of critical infrastructure

Given definitions of CI reported in Chapter 1 and the fact that the maritime sector sustains society and the economy through the movement of people and vital goods, such as energy (transportation of oil and gas) and food, the maritime sector has to be considered a critical infrastructure [85].

On one hand, maritime infrastructure is critical to the employment of national maritime power, on the other hand it is a possible target for acts of terrorism. Indeed, a successful attack against a port could incur serious economic and military damage, generate an elevated number of casualties, and have serious long-term detrimental effects on the national economy.

Maritime critical infrastructure protection (MCIP) presents many challenges in today's asymmetric environment [178]. Previous models of maritime defense have focused on protecting ships from traditional naval attack that comes from the sea: even when ports and supporting infrastructure have been considered targets, emphasis was on defense against a military threat.

The post 9/11 scenario has created a new outlook on maritime defense. Many targets without any military importance in a conventional war, such as symbolic buildings and places, must now be considered in strategic defensive planning. Possible threats from the sea are wide-ranging and diverse, relying on a combination of asymmetric offensive tactics while exploiting the variety of the littoral.

The main threat categories for port facility security refer to [142]:

- Theft and sabotage;
- Terrorism;
- Illegal traffic and migration;
- Environmental threats and large-scale accidents.

In addition, new and emerging asymmetric threats have to be considered [142]:

- Political trans-national and international terrorism;
- Actions that may harm the safety of national and international transport systems;
- Individual or group actions of illegal access to informative data systems;
- Deliberate actions that can affect the credibility and seriousness of a nation;

- Deliberate acts of ecological sabotage.

Due to the complexity of the activities and large areas to survey in a port, it is highly necessary to implement a Joint Harbor Operations Centers (JHOCs) as a component of maritime anti-terrorist force protection [178]. The development of multi-agent maritime homeland security systems is a logical next step in the evolving problem of port security and defense.

Main stakeholders and players

Securing the critical infrastructure of the maritime sector is increasingly becoming a priority for the key European stakeholders, including the European Commission, European member state governments and the main actors from the private sector [85].

At the global level, the relevant stakeholders include, while not being limited to, various intergovernmental organizations such as the International Maritime Organization (IMO), the World Customs Organization (WCO) and the ICC International Maritime Bureau (IMB), which is a specialized division of the International Chamber of Commerce (ICC). Additionally, it is also important to mention the relevance of the International Maritime Security Corporation (IMSC) which focuses on actions to specifically protect ships, their crews, and their cargo against a variety of threats.

The lack of coordination between stakeholders at different levels e.g., European and national, brings about major discrepancies in the way maritime security is addressed.
The list of players in this sector includes both private and public agents:

- Industrial facility(ies) owners and operators;
- Industrial facility(ies) employees;
- Providers, customers, users;
- Emergency response agencies and personnel;
- Vessel crews;
- Shipping companies;
- Fisheries in the area;
- Tourism companies;
- Ferry companies;
- Local communities;
- Littoral state governments;

- Communities, users and companies down the supply chain;

From the above, it is clear that maritime critical infrastructure protection represents a challenging and hard to solve problem and that only a synergistic, joint, and coordinated effort can be effective.

Requirements

An integrated security system (ISS) for a maritime port area has to accomplish the following objectives [142]:

1. Advanced detection of any attempts to intrude into the port security areas;
2. Transmitting alarm and sabotage signals to the software, giving it the possibility to remotely control the activation and deactivation of security areas and to acknowledge alarm signals;
3. Surveillance for threats;
4. Security data dissemination at the port's local and central authority levels, as well as at the other institutions involved in disseminating security events.

To implement an ISS for a critical maritime infrastructure, it is necessary to specify that any port area can be physically characterized by different parts, namely perimeter boundary, access points, and infrastructure (such as transport, communication system, utilities, etc.).

In order to ensure the security and the safety of all these parts, the main function of the security system is to control the access flow inside the port area, which implies the necessity to prevent unauthorized accesses through access points.

The surveillance capability can be fused into one multi-agent system including:

- A coastal radar;
- Vessel traffic services (VTS) system;
- An automated identification system (AIS) processor;
- A port control camera system;
- A thermal imagery system for night-time operations.

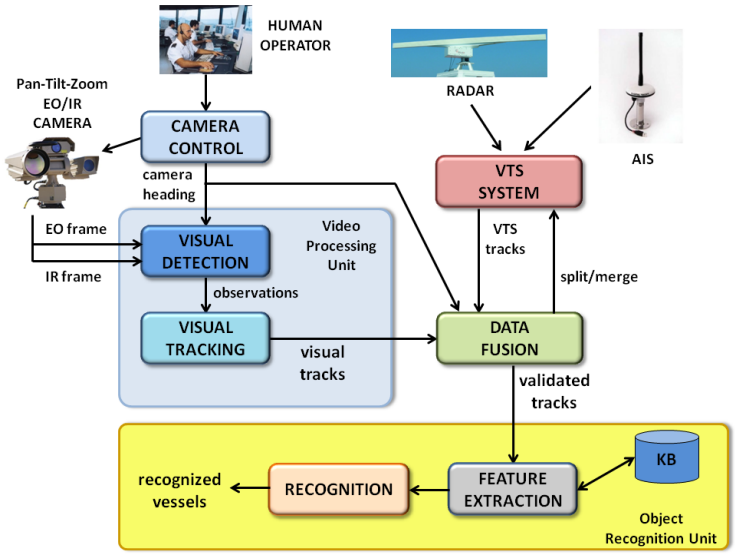


Figure 5.3: Functional architecture of the proposed framework. The data fusion module manages information coming from the visual processing unit and the VTS system and sends validated tracks to the object recognition unit.

A general framework for a multi-agent automatic surveillance system, allowing for data fusion of information coming from different sources, namely radar, electro optical (EO) day-light cameras, and infra-red (IR) night-time cameras, is shown in Figure 5.3.

An EO-IR visual device is the main sensor and it can be moved through a control module by a human operator. The control module provides orientation and field-of-view (FOV) from the device to the VPU, which is responsible for detecting and tracking vessels using visual information. The data fusion module receives data from both the VPU and the VTS system. It aims at associating the visual tracks coming from the video analysis with the tracks generated from radar and AIS data. In this way, it is possible to provide the user with a new visual dimension in addition to the traditional geo-referenced, radar-like VTS view. Moreover, the data fusion module sends feedback information to the VTS system, in order to improve the detection accuracy of the radar.

The tracks generated by the data fusion module are considered valid and the object recognition unit can classify them according to their visual features.

5.2.2 Standard solutions for securing the CI and open issues

The Cold War defense model assigns the responsibility of ports and outload operations to the Navy-Coast Guard. Defensive operations are managed by

co-locating coast guard and navy personnel in operation centers that would oversee all military operations (including load out operations and critical infrastructure protection) within the port during time of national emergency.

In the standard Cold War defense model risk was very much a matter of proportionality and the threat to critical maritime infrastructure was distinctly military. In considering the worst case scenario, planners envisioned enemy actions in the littoral focusing on submarine attack, offensive mining, and special operations attacks against critical military infrastructure. It was assumed that terrorist actions would be sponsored by the enemy state and, as part of the enemy strategy, would not be directed against targets with limited or no military significance.

Prior to 9/11 the coast guard port and offshore tactical constructs were divided into two separate areas of responsibility based on the type of law enforcement being conducted. Regulatory functions, such as vessel inspection, environmental response, licensing, etc. were performed by vessel inspectors who performed their duties unarmed. Operations of a more traditional law enforcement variety, such as counter-narcotics or fisheries enforcement, search and rescue, and other offshore operations were the responsibility of armed personnel.

The 9/11 terroristic attack has demonstrated the fragility of this standard defensive approach. The difference between pre and post 9/11 consists in how to interpret the concept of potential targets. As an example, in the pre 9/11 scenario, a strictly civilian target such as the World Trade Center would not have been considered a valid target in New York City. Indeed, the major weapons out load point at Earle, NJ, was Priority One for infrastructure protection. Obviously this has changed: maritime infrastructure that would not be considered critical in a Cold War scenario now has the potential to be targeted as a means of obtaining an economic or psychological victory.

Since a potentially infinite list of non-military targets can be chosen by the terrorist, it is impossible to have enough defensive forces to protect all the potential targets. This is not to say that the Cold War model is completely invalid, or that we cannot learn from the lessons of history. In the pre 9/11 model, military intelligence had to deal with a specific military threat against known target areas, with a response that was distinctly military. The new threat requires that the defensive model is expanded to consider all the players within the port vital for total protection.

Open issues

Two main open issues can be individuated:

- The existing maritime security standards, methodologies, and tools are monolithic and concentrate solely on physical security;

- Commercial ports are not considered as critical infrastructure and the security of their information and telecommunication systems is not organized.

Lessons from the past indicate that the key to effective defense is tactical coordination through dedicated multi-agent command and control [178]. Before 9/11 the model for command and control was to deal with a military threat from the sea, but this has changed with the emerging new asymmetric threats. The JHOC concept has proven to be effective in multi-agency intelligence fusion and coordinated tactical port operations essential for maritime critical infrastructure protection and should be considered a model for coordinated port defense.

5.2.3 Types of attacks, exploited vulnerabilities (anatomy of an attack) and economic consequences

The asymmetric threats to ports, in particular the terrorist ones, are a relatively new element in the spectrum of naval warfare. Until quite recently ports were composed of infrastructure that was relatively easy to replace or replicate, making them relatively low priority targets for an enemy dedicated to striking at maritime strength [178].

Today ports have become centers of highly technical, well-integrated infrastructure designed for the rapid loading and unloading of cargo, an evolution that has become highly complex in the era of containerization. Port cargo operations are also highly dependent on networked operations, making the disruption of the process far simpler for a potential attacker. This has made major ports more important economically and strategically while simultaneously making them more attractive targets for terrorist action.

The following aspects have to be considered when evaluating an attack against a port facility.

Economic impact. Imports and exports rely on shipment by sea. A successful attack on maritime infrastructure would affect this trade in far greater proportion than the actual physical damage. An attack on one port would have a cascade effect on other ports. Delay of shipping in loading and offloading cargo is one of the most costly elements of the shipping process.

High visibility. Ports are not isolated areas, but rather major centers of commerce, usually surrounded by large cities and economic activities. An attack on a port could be highly visible and potentially the scene of mass conflagration. As a result of urban development, most major ports are no longer confined to strictly industrial areas, but rather have become well-developed centers of commerce and entertainment, surrounded by built up waterside areas dedicated to tourism and recreation. Many of these facilities are located

next to volatile maritime infrastructure (fuel tanks, docks, etc.) that could create mass conflagration if attacked through large explosive force. Sympathetic detonation, fires, and other catastrophic effects would certainly create mass casualties.

Ease of attack. Commercial ports are not fortresses. The ocean itself presents a number of distinct advantages to a dedicated attacker, especially when employing maritime suicide terrorism or means to rapidly deliver large explosive force. Water is not only a tremendously efficient transport medium (allowing for rapid transit), but the large amount of legitimate commercial and recreational traffic in ports allows for an enemy to mask movements prior to an attack, making effective defense difficult.

5.2.4 Protection strategies

Given the importance of ports to the economy and military power, it is likely that ports will become a target for future terrorist attacks.

MCIP is a critical vulnerability that must be addressed by realizing a coordinated effort among different agents. This mission goes beyond the traditional port security operations or anti-terrorist force protection, and requires a command and control construct that can truly fuse the myriad of responsibilities and operations in ports.

The diversity of the threat against ports and the number of regulatory agencies that oversee critical infrastructure requires an expanded comprehensive command and control system that fuses multi-agency intelligence, has understanding of multi-agency capabilities, and can provide direction to these forces in the field.

Multi-agent JHOCs offer several advantages for merging effective port operations and critical infrastructure protection. This is evident in the areas of intelligence fusion, coordinated planning, and tactical command and control.

European member states have to exchange data related to the maritime context. Appropriate identification and categorization of relevant data is required, in order to facilitate an agreement on data exchange between the various stakeholders.

However, a series of difficulties have to be overcome in order to develop an effective JHOC. For example, regarding cybersecurity incidents and other cyber-related threats (e.g., fraud, e-crime, etc.) facing the maritime sector, the lack of information exchange between the involved actors represents a crucial problem [85]. Cybersecurity should not target only major ports. Even less developed ports should be offered the opportunity of implementing cybersecurity initiatives.

Moreover, the level of ICT implementation maturity varies greatly from one port to another, while security is not always a priority. Therefore, a first

step towards achieving cybersecurity at port level would be the implementation of ICT systems that are secure by design.

5.3 Railway system

This section is organized as follows: 5.3.1 provides a brief introduction on railway critical infrastructure; 5.3.2 discusses standard solution for securing the railway infrastructure and related open problems; 5.3.3 explains, using examples, a set of attacks that could be performed exploiting known vulnerabilities; finally 5.3.4 summarizes the open issues, suggesting some updates to the system, and introduces risk analysis to mitigate future attacks and threats.

5.3.1 Description of critical infrastructure

Railway systems move people and goods within a country and between countries. Before 1989, however, within the European Union different, often incompatible, control mechanisms were employed. The European Rail Traffic Management System (ERTMS) is an initiative that aims to overcome this situation, by defining a shared standard to enhance the interoperability among the railway systems of different countries. The ERTMS is composed of the European Train Control System (ETCS) and the GSM-R. The ETCS is a standard for in-cab train control that includes signaling and train protection systems. The GSM-R, an extension of the GSM, is the radio system for providing voice and data communication between the track and the train.

Since "disruptions of the railway infrastructure can have a significant negative impact on the economy and security of an individual country" [47] and the current railway system depends on ICT (typically to increase performance), the security and safety aspects (especially for wireless communications) become critical.

Main stakeholders and players

The railway system is a supranational critical infrastructure and it has a wide range of stakeholders and players. Considering the European railway system, the major stakeholders are: the European Commission, which defines guidelines for railway system integration; European member states, which supervise the system; private/public companies,which implement and manage the infrastructure (e.g., Rete Ferroviaria Italiana, RFI) and local communities, which benefit from the service to transport goods and people. Also the list of players involves several actors, from international/national companies to passengers. A simple, non-exhaustive, classification includes: public and private rail transportation companies for passengers and goods (e.g., Trenitalia); supply companies: e.g., railway signaling system, trains, IT services, etc.; shipping companies; passengers; local transportation companies (e.g., Gruppo

Torinese Trasporti, GTT); employees of different companies (e.g., Trenitalia, GTT).

Since the railway system includes several stakeholders and players from local communities to national ones, the economic interests are very high and subversive issues (e.g., terrorism) may expose the system, hence people and goods, to physical and electronic attacks.

Requirements

As widely known, the dependability of railway system is based on reliability, availability, maintainability, safety and security (RAMSS). These attributes drive the definition of requirements that must be fulfilled to avoid, or to limit, accidents and attacks. Although the railway system is composed of different components and a holistic approach is required to satisfy its dependability, this section mainly focuses on the communication aspects among electronic systems that have been widely discussed in literature and have been defined by European railway standards.

Two standards in particular are applicable to communication in safety-related electronic systems. The first is the CEI EN 50159-1 [68] for closed transmission systems (according to [68], a closed transmission system is constituted by a fixed number or fixed maximum number of participants linked by a transmission system with well known and fixed properties, where the risk of unauthorized access is considered negligible) and the CEI EN 50159-1 [69] for open transmission systems (according to [69] this is a transmission system with an unknown number of participants, having unknown, variable and non-trusted properties, used for unknown telecommunication services, and for which the risk of unauthorized access will be assessed).

The standard EN 50159-1 provides a set of requirements for closed transmission systems. Briefly, there are six main requirements that should be provided:

- safety protection will be applied to the generation of the data to be transmitted;
- safety reaction will be applied in case of misoperation. This shall be consistent with the safety requirements of the receiver;
- error detection mechanism will be applied at the receiver and will be consistent with the safety requirements of the receiver;
- the implementation of the safety reaction will be functionally independent of the non-trusted transmission system;
- the residual error rate of the safety-related transmission system for each information interchange between transmitter and receiver will be less

than a pre-defined value. This rate must be compatible with the safety integrity level of each receiver;

- the safety integrity level of the safety-related transmission system will be consistent with the highest safety integrity level of the safety processes;

Note that these requirements are safety-related requirements and do not explicitly mention security threats that may arise due to tampering, external attackers or malicious authorized users. This follows the definition of closed transmission system reported above.

The standard EN50159-2 provides a set of instructions and requirements for open transmission systems. In particular, seven possible security threats are identified: repetition, deletion, insertion, re-sequence, corruption, delay and masquerade. The standard presents guidelines for protecting the safety of the transmission system; these are sequence number, timestamp, time-out, source and destination identifiers, feedback message, identification procedure, safety code, and cryptographic techniques.

Note that this standard only considers unauthorized users while it does not address the possibility of malicious actions performed by authorized ones.

5.3.2 Standard solutions for securing the CI and open issues

The rail system is critical to the economic and social wellbeing of several, if not all, EU nations; the Italian railway system is one of the most important parts of the infrastructure of Italy, with a total length of above 24000 km. This section discusses in general terms to what extent security is considered an issue for the railway infrastructure and what is the current level of protection.

In closed transmission systems the risk of tampering is usually considered negligible, and the potential actions of malicious authorized users are mitigated by the internal logic of the system: the components which are in charge of taking decisions, e.g., the European Vital Computer (EVC), do not permit determined actions if these are not confirmed by sensors and other (human-independent) indicators (e.g., balises or the Radio Broadcast Center). An example of a closed system and its related security issues are reported in the section below.

In an open transmission system, instead, security is actually considered an issue, although we must remember that the standard EN50159-2 requires security in order to guarantee safety. That is, availability is not explicitly addressed in the standard: DOS attacks may result in blocking communications, and consequently forcing trains to stop, although having no impact on safety. This approach has been largely applied in the past, but it should be re-considered, given the relevance that nowadays critical infrastructure is acquiring. The unexpected stopping of a train results in delays (also with a cascading effect for all trains that share the same line), and ultimately loss of money: cybersecurity attacks that target the transmission system leading to

unavailability may be not longer acceptable. In fact, it is possible to exploit the fail-safe behavior of ERTMS and create a situation that causes a train to halt [53]. Although a DoS attack on ERTMS might not impact on safety, it could cause disruption or passenger discomfort. Therefore DoS attacks are relevant at least for the availability of the service.

In more detail, [53] considers the different components (and their interfaces) that interact with on-board ETCS system. The driver and train interfaces are only specified at a functional level and no other requirements are provided for their implementation. Both driver and train are considered trusted components because the driver could override the entire ERTMS/ETCS system and the train because it could have been sabotaged in other ways (e.g., compromising the braking system). However, an important issue is that the specifications do not define authentication on the communication channels that are used for these interfaces. Although, whether this approach is acceptable in a closed network (i.e., the interfaces connected only to ETCS), in a setting where the on-board systems are connected to a network that transports other services, for example, Internet access for passengers, these components may be compromised (e.g., by a malware). The balises are a part of the signaling system and are placed on the track. Although the balises are protected from accidental transmission errors and interferences and the ERTMS/ETCS provide different levels of data consistency checks, no authentication mechanism is provided. Again, this makes malicious attacks possible, for example, an attacker could tamper with a balise and send counterfeit data, subvert an existing balise or place a new balise on the track. The ERTMS make a distinction between linked and unlinked balises. The locations of linked balises are transmitted by radio over a secure channel to trains. If a train does not encounter a linked balise at the expected position, stops. On the contrary, unlinked balises could be encountered everywhere. Even if the trains accept a limited number of commands from an unlinked balises, some DoS attacks are still possible and some commands can be used to create a hazardous situation. The purpose of the Euroradio protocol [22] is to transmit the linked balises messages to the train over GSM-R network. In particular, the communication is established using a shared secret key ensuring authenticity and integrity of messages. However, the protocol does not guarantee confidentiality of transmitted information, and so if the GSM-R network were to be compromised, it would be possible for an attacker to eavesdrop on ERTMS messages and perhaps learn sensitive information using a man in the middle attack [53]. In addition, Euroradio suggests using Triple DES as the underlying cryptographic algorithm rather than a more effective algorithm such as AES. Another problem is managing key distribution since the interoperability specifications of ERTMS/ETCS only deal with secure key management between different key management domains, leaving key distribution within a key management domain to national implementation [53]. Although a new specification to mitigate this issue has been proposed, the current standards

adopt an off-line key management solution, which is not feasible for refreshing and revoking keys. The GSM-R is an extension of GSM that provides additional services required for railway operations. Since standardized numbers are used to address on-board functions, once an attacker has gained access to a GSM-R network disruption could be caused. In addition, [147] analyses the security of the GSM system concluding that the design of GSM security is weak since it adopts some cryptographic algorithms which come from the security through obscurity approach. Again, this weakness could be exploited by using a man in the middle attack. Also, an attack on the GSM-R network could bring down the ERTMS/ETCS system over a large area, creating a wide area DOS attack [53]. Finally, like other wireless communications systems, the GSM-R is also susceptible to radio interference from external sources. In [47] Baldini et al. discuss the fact that interference can potentially affect the entire railway infrastructure, because the movement of every train is correlated with the positions of other trains in the network, causing a potential service disruption.

Additionally, the railway infrastructure is currently rather exposed to the risk of terrorism, including cyber-terrorism. In fact, a key part of the rail transportation system presents an inviting target for terrorist attacks [60]. Terrorist attacks and criminal attacks are discussed in [60], [165]: sample attacks points are railroad tracks and switches (vulnerable to attacks by unbolting of joint bars or misalignment of switches), bridges (vulnerable to attack by explosives), tunnels (vulnerable to attacks by explosives and chem-bio agents), control and dispatching systems (vulnerable to explosive and to cyberattacks). An example of security analysis in an open system is reported in the section below.

5.3.3 Types of attacks and exploited vulnerabilities (anatomy of an attack) and economic consequences

This section discusses the types of attacks and exploitable vulnerabilities that affect railway infrastructure using two different examples: security in closed/open systems and railway/subway attacks. In particular, as discussed below, the security of a closed system could impact only service availability (e.g., train stops), otherwise in an open system security could become an issue. For example, this situation may occur when an IT infrastructure is shared between a critical service (e.g., communications between a train and a radio block center, RBC) and a non critical service (e.g., Internet access for passengers).

Security in closed systems: an example

Two main components of railway train-borne equipment are the European vital computer (EVC, in the ETCS) and the driver machine interface (DMI).

The EVC is the main core of the on-board automatic train control system: it supervises the movement of the train and sends information to the DMI. The train is supervised using information received both from the *eurobalises* (transmitters located along tracks) and from the RBC through a GSM-Railway network. The DMI acts like a bridge between the train driver and the EVC. It communicates with the EVC as a slave; it shows, using both audio and video devices, EVC messages and information to the driver, and communicates inputs from the driver to the EVC. The train driver interacts with the DMI using the DMI's LCD screen, audio devices and keyboard (or touchscreen); the train driver performs the two key roles of (i) information sink for EVC originated information, which is displayed by the DMI, and (ii) command source for commands to be delivered to the EVC by means of the DMI [65].

The EVC is a safety-critical component, its safety being assessed according to Safety Integrity Level 4 (SIL 4) as prescribed by the related CENELEC standard. This means that the tolerable hazard rate per hour (THR) is required to be between 10^{-9} and 10^{-8}. The DMI is instead a SIL 0 component.

The EVC and DMI are configured in master/slave, where EVC is the master and the DMI executes only on the basis of the orders sent by the EVC (audio and video information). The DMI sends data to the EVC only when explicitly requested e.g., because of data entry performed by the train driver, to select options from menu, to acknowledge messages. The protocol applied for the EVC-DMI communication is described in the standards [167], [166].

Security threats in this case are not considered an issue, because the DMI-EVC communication follows the definition of closed system reported above. The EVC merges information from the train driver, the balises, and the RBC: the EVC is able to guarantee safety of the train mission independently of the behavior of the train driver, and of possible incorrect inputs received by a tampered DMI. Note that any kind of misbehavior of the train driver, inconsistencies in the information received, or delay of the information, result in the train entering a safe state (e.g., a train stop), thus affecting availability.

Security in open systems: an example

The ALARP (a railway automatic track warning system based on distributed personal mobile terminals, [36]) project proposed to design, develop and validate an automatic track warning system (ATWS) able to: i) detect trains and rolling stocks approaching a worksite, and ii) notify their arrival to the workers, thus improving their safety. In fact, the safety of workers in railway scenarios is a serious concern, since vehicles are constrained to tracks, drivers have tight margins to react in if there are emergencies trackside workers can be exposed to injuries and fatalities. For example, between 1993-2002 , on US railways there were 460 fatal railroad-related work injuries among railroading workers and 761 fatal railroad-related work injuries involving workers not from the railroad field [83].

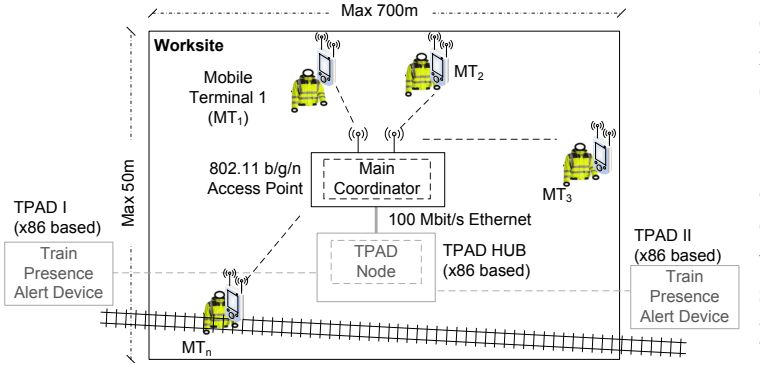


Figure 5.4: Overview of the ALARP communication system.

ALARP architecture is based on the following components (see Figure 5.4). One or more trackside train presence alert devices (TPADs) are placed external to the worksite; the TPADs sense approaching trains (rolling stock) on the monitored track. A set of distributed, real-time and wireless mobile terminals (MTs) [64], are worn by the workers, providing accurate, safe and real-time information about approaching trains and events that can put at risk the workers' safety (e.g., health problems). TPADs and MTs are connected through a base station.

When a train is approaching, it is detected by a TPAD, and a notification is sent to the MT. The MT dispatches an alert to the worker if he is in a dangerous area (called *red* zone) which is located close to the track in which the train is approaching, and a warning if the worker is located in a non-dangerous area (called a *green* zone).

The overall communication architecture in ALARP follows a centralized communication setup based on the IEEE 802.11 standard. This enables better predictability of the communication timing and simplified realization of synchronous communication channels at the worksite. This setup is primarily based on a fixed coordinator located at the worksite, with all MTs communicating via the coordinator. The deployed timed reliable wireless communication protocol uses the coordinator to implement its centralized communication algorithm and maintain allocation of necessary communication resources. Communication links between TPAD and the worksite can be enhanced by helper infrastructure in the form of additional relay nodes (or repeaters) at the transmission path. At the worksite, TPAD information is disseminated to MTs via the coordinator [117].

High reliability, timeliness and safety, despite the possible harsh conditions are mandatory requirements of ALARP communication, as alarms raised by the TPAD are safety-critical messages that need to be timely delivered to all the workers. For safety reasons, violations of timing bounds need to be detected, and workers are signaled to move to a green zone; this operational

procedure maintains safety but impacts on the working time, leading to loss of productivity [117].

A security analysis of the ALARP system and its transmission protocol has been performed in [63], first using the ADVISE method introduced in [111] for modeling the behavior of the attacker, and then combining it with a stochastic activity network (SAN) [155] model of the system behavior. To quantify the impact of external attacks, the main vulnerabilities of the system were analyzed, with particular attention to vulnerabilities of the communication architecture: vulnerabilities of the IEEE 802.11 standard in particular were observed. Results, reported in [63], show that external attacks do not impact on the probability of occurrence of a catastrophic failure (potential harm to workers) i.e., they do not reduce the safety of ALARP. But in order to protect from cybersecurity attacks, the MT enters a safe state, and alerts the worker to move to a green zone (safe position). This mechanism has a strong impact on the working time as already mentioned.

It is noteworthy that this quantitative investigation of security is not requested by railway certification standards, but still it is considered relevant in the context of an open, critical infrastructure using COTS components, where a multitude of threats can lead to severe safety hazards or availability drops.

Railway or subway station attack: an example

As an example application, consider a railway or subway station. The threats against the infrastructure that should be considered include damage to property and vandalism, theft and aggressions to personnel and passengers, microcriminality, manumission and forced service interruption (sabotage), bombing or spread of nuclear, bacteriologic, chemical or Radiologic (NBCR) contaminators (terrorism). An analytical taxonomy of these threats would look like the following:

- Vandalism
- Theft of PCs
- Bombing
- Hacking
- Gas attacks
- Infrastructure damage

Each of these threats can be associated with quantitative parameters:

- the frequency P of occurrence of the threat [events/year];

- vulnerability V of the system with respect to the threat, i.e., the probability that the threat will cause the expected consequences (damage), given that the threat has occurred;
- the expected damage D occurring after a successful attack in Euro. For instance, the expected damage, relating to a single attack, may be computed by predicting the expense needed to restore the assets and the possible consequences of service interruption.

It is assumed that the values are obtained by analyzing historical data of successful and unsuccessful attacks before and after adopting specific countermeasures, such data is usually available for comparable installations. On the other hand, it is necessary classify the available protection mechanisms, e.g., fence alarms, volumetric detectors, video-surveillance (internal), chemical detectors, intrusion detection, system explosive detectors. Each of these protection mechanisms must be associated with a few quantitative parameters:

- list of threat categories for which the mechanism is effective;
- expected protective, deterrent, and rationalizing effectiveness, i.e., percentage of risk reduction the mechanism enables;
- site, i.e., geographical reference, to which the mechanism applies;
- estimated coverage, e.g., percentage of the physical area or perimeter of the site;
- annual cost (acquisition, management, maintenance, etc).

These parameters account for the benefits brought about by the protection mechanisms as well as the cost they incur. One possible objective of a quantitative risk analysis and the risk management process may be the derivation of optimal security-related design choices, bringing the risk level below a required threshold value under given cost constraints.

5.3.4 Protection strategies

Although much effort has been made to increase safety and interoperability of European railway system (e.g., the standard ERTMS/ETCS), the currently available systems adopt a wide set of technologies which, at least for some of them, should be upgraded. In addition, the available standard defines the interfaces between components without additional requirements for their implementation, which typically is managed at a national level. This approach could work for interoperability but it is not enough to guarantee the safety and security of a critical infrastructure. While the substitution or upgrading of technologies is not a very complex task, the identification and deployment of shared strategies (e.g., among the European nations) to protect the railway

system is quite hard. Therefore, considering the European railway system, a common and holistic approach for safety and security becomes necessary. The adoption of shared risk analysis approaches helps to identify and manage critical aspects and threats of a modern and interoperable railway system.

The following sections introduce the risk analysis techniques and a set of technical solutions and upgrades useful for the railway system.

Risk analysis

Risk analysis is a central activity in the security assurance for critical railway transportation infrastructure and mass transit systems. Risk analysis can be performed using qualitative approaches, based on expert judgment and limited ranges for risk attributes [171]. However, model-based quantitative approaches [91] may be required in order to precisely determine the risk indices taking into account the frequency of occurrence of threats (e.g., considering historical data) and the consequences (damage of assets, service interruption, people injured, etc.). Quantitative approaches pose several issues, such as the availability of source data and the methodology for the analysis, which is not straightforward. Several approaches to the risk analysis of critical infrastructure are available in the literature [44, 56, 93, 113, 127, 132], although in most cases they are either qualitative, excessively abstract and general, or tailored to applications other than railway transportation. Risk assessment is the process of measuring the expected risk as a combination of threat occurrence probability, system vulnerability, and expected damage. Each of these aspects can be evaluated quantitatively by adopting suitable approaches, possibly based on well-established techniques borrowed from the reliability domain. To do so, an appropriate threat frequency model (e.g., based on BNs) is necessary, which quantifies the frequency of occurrence of the threat, measured in events/year; a threat vulnerability model (e.g., based on SPNs), enabling the quantification of the vulnerability of the system with respect to each threat, i.e., the probability that the threat will cause the expected consequences, or damage, given that the threat has occurred; a threat consequences model (e.g., based on event trees) which estimates in Euros the impact of the expected damage occurring after a successful attack.

Unfortunately, the above parameters involved in risk assessment are not easy to obtain. The analysis requires both procedural and modeling aspects. Procedural aspects include brainstorming sessions, site surveys, design review, statistic data analysis, expert judgment, etc. Formal modeling languages that can be used to analytically compute threat frequency, vulnerability and consequences include Attack Trees, Bayesian Networks, Stochastic Petri Nets and possibly other formalisms able to take into account the uncertainty inherently associated with the risk, as well as the possibility of strategic attacks [136]. In fact, these three parameters feature an inter-dependence which should be modeled, too. Risk management (or mitigation) is instead used to indicate

the process of choosing the countermeasures and predicting their impact on risk reduction. Protection mechanisms can reduce the risk as they have three main effects:

- protective, aimed at the reduction of the vulnerability levels;
- deterrent, aimed at reducing the frequency of occurrence of a given threat;
- rationalizing, aimed at the reduction of the expected damage.

Each mechanism has some impact on one or more threats. Again, it is necessary to quantify this impact, evaluating both the fraction of the asset or resource in the system actually protected by the mechanism and the quantitative reduction of the risk it enables. Furthermore, an installation/operation cost, associated to each mechanism, should be taken into account to evaluate its actual profitability. Consequently, in a real-world scenario suitable approaches for carrying out analytical evaluations of cost-benefit trade-offs and drive precise security-related design choices may be needed.

Technical solutions

Related works [53, 47] discuss a set of technical protection strategies that could mitigate threats that affect wireless communications of ERTMS.

First of all, the Euroradio protocol is built upon the GSM-R (a GSM extension) that has some security issues, as discussed in [147, 53]. At least two solutions are available: upgrade GSM and GSM-R or switch to another wireless technology; probably, the most simple solution is the former. Since Euroradio does not guarantee confidentiality of the messages [53], if the GSM-R network were to be compromised, an attacker could eavesdrop on ERTMS messages. Thus, even if the GSM and GSM-R were upgraded, this problem persists.

As discussed before, driver and train interfaces were in the past considered trusted, however in the current specifications no authentication is required on the communication channels that are used for these interfaces [53]. Therefore, considering a scenario where ETCS system were connected to a network that carries non-critical messages (e.g., Internet connection for passengers), authentication is recommended. Balises, like driver and train interfaces, also do not support authentication. This situation opens up the possibility of malicious attacks via the balise interface, since the data received from a balise is effectively trusted by the system [53]. Therefore, also for balises, authentication is recommended.

Baldini et al. [47] discusses the interference issue that affects GSM and GSM-R causing DoS attacks on the system. Their work describes a wireless monitoring system to detect interference sources and apply the appropriate

countermeasures. Their provided results are supported by a set of experiments conducted on some Italian railway stations.

5.3.5 Fault mitigation approaches

In the railway domain, components that are deployed across the infrastructure need to be certified according to strict safety standards in order to be allowed to operate. Such standards also provide guidelines on the kind of faults that should be taken into account for the purpose of system certification, and provide general guidelines in order to ensure that the system exhibits safe behavior despite the occurrence of faults.

In particular, the fail-safe concept has been commonly used from the early days of railway systems [67]. This concept is based on the use of components having well established failure modes, and on the achievement of a safe condition in case of failure of one of its parts. For high integrity systems in the railway infrastructure (i.e., SIL3 and SIL4 systems), the CENELEC standard EN 50129 [66] prescribes the application of this principle for the mitigation of any *single random hardware fault* which is recognized as possible. According to CENELEC EN 50129, the fail-safe principle can be achieved in different ways:

- *Composite fail-safety* assures that each safety-related function is performed by at least two items. Each of these items is independent from the others and the necessary number of items shall agree in order to progress.
- *Reactive fail-safety* assures safe operation by proper detection and negation of hazardous faults that occur in a single item that implements a function. The detection is regarded as a second item that shall be independent in order to avoid common-cause failures.
- *Inherent fail-safety* is considered if all non-negligible failure modes of a single item are non-hazardous.

Below can be found some fault-tolerance approaches that are typically applied in processing units within the railway infrastructure when implementing the above principles.

- *Single channel with self-test by software.* In a single channel architecture there is only one flow of computation on a single piece of hardware. Error detection is performed only by additional software functions that perform self-tests.If an error is found then the system is forced to a safe state. Note that it is very difficult to prove that a faulty hardware unit (microprocessor) can auto detect a failure in itself, reveal it and then assure a safe state (e.g., shutdown). Single channel systems are often

used together with a standby unit which takes over the computation when a failure is detected in the primary channel.

- *Coded processing.* This method uses a single channel architecture with special error detecting codes. All variables of the program consist of a value part and a control part. The program instructions handle both parts of the variables. Both inputs and outputs are in this encoded form. For example, to cover CPU related faults, the control part may consist of the following codes: (1) an arithmetic code to detect computing errors, (2) a static signature to detect addressing errors (operand error, operator error, variable confusion), (3) a timing signature to detect timing errors (incorrect number of loops etc.), (4) sequence signature to detect certain sequencing and branching errors. In a fail-safe controller the final signature is compared with a pre-calculated one and a safe state is enforced if they differ.

- *Multi-channel architectures.* These architectures incorporate independent processing channels (flows of computations) with facilities for cross-checking between channels to detect divergence and latent faults. Usually, physical independence between channels is utilized to form physical-fault containment regions, while design diversity is used to achieve design-fault containment regions in the channels. Inputs are copied to each channel or accessed independently (and synchronized before processing). Two channels can be used to provide safety (using inter-channel comparison and switching to a safe state) or improved reliability (based on intra-channel self-checking and failover between the channels). The increase in the number of channels can be motivated by the need to tolerate more intricate faults (e.g., in case of Byzantine faults at least four channels are required) or the need to survive a given mission time with high probability. In a common configuration, one channel is responsible for the output, which is monitored by the other channels, and each channel is allowed to decide if the output is faulty (a typical set-up is two-out-of-two). The usual fail-safety mechanism is to set the outputs of the channels to a predefined safe state in case of a failure. If there is no safe state, or the reliability has to be increased then majority voting can be used by defining the minimum number of channels that must agree (the most common set-up is the two-out-of-three).

- *Dual channel architecture with a diverse safety bag.* All actions are processed on a two channel basis with diverse software. The two channels are the logic and the safety channel (safety bag). Inputs in the logic channel are checked against operating and safety conditions and the computation begins only if the result is positive. Before output, it is checked again by the safety channel if the result would lead to hazardous operational conditions. If there are no problems, both channels

will provide the output. A separate comparison is performed before the outputs are used. Diversity between the channels is enforced by different specifications, different languages (e.g., procedural and rule-based). If any of the comparisons detects a disagreement, then the system goes into a safe state.

- *Reciprocal comparison by software.* Two redundant software processing units are used (potentially on the same hardware) which exchange data reciprocally (intermediate results, test data etc.) and the comparison of data is carried out in software in order to detect discrepancy and lead the system to a safe state. This is executed in both channels independently (dual voting),to assure a safe comparison.

It is also essential to enable fault forecasting at early development stages of the computer-based control system.The model used to represent the system should ensure many, often conflicting requirements, i.e., it should enable a realistic and detailed description of the system, it should be maintainable, relatively easy to apply, and efficient to process by means of suitable analysis tools. A few example of models used to represent the system with respect to RAMSS requirements include Bayesian Networks (BNs), Fault Trees (FTs), Repairable Fault Trees (RFTs), Petri Nets (PNs) and Generalized Stochastic Petri Nets (GSPNs). Classical examples of situations where the RAMSS requirements of a railway control system can be modeled by one or some of these formal approaches include the use of reliability models for on-board systems, (e.g., FTs), performability models for describing networks and software, (e.g., by means of GSPNs), maintainability models for trackside systems, (e.g., RFTs), or combined models (e.g., GSPNs, FTs, BNs) for evaluating the safety of redundant architectures in the presence of imperfect maintenance.

CHAPTER

6 The Maturity of Italian Critical Infrastructure

In the light of what has been presented in the previous chapters, it is evident the critical infrastructure is quite important in general. To better understand how relevant they are to Italy, this report tries to answer two simple questions that almost arise spontaneously: i) What is the degree of maturity of the Italian critical infrastructure? ii) What is the relevance of critical infrastructure for the life of the Italian nation? Answering these questions is not so easy. On one hand a complete and exhaustive answer is difficult due to the large number of different critical infrastructure that must be considered and, in addition it is not always possible to have a complete and updated overview, either because the available data are not always up to date, or because data are the proprietary of private companies which do not always make them available. This final chapter aims to provide answers using statistical data restricted to some specified area of interest, without entering in details that can be found in the provided references.

6.1 The relevance of CI in society

First of all, we may distinguish between *physical* and *cyber* CI. Physical infrastructure consist of a wide range of systems and facilities, in other words something concrete and easy to assess (e.g., energy, transportation, telecommunication, water supply, etc.), which is today complemented by a cyber part. Cyber CI is more abstract, intangible, sometimes virtual and usually tied to IT (e.g., financial services, e-health, e-government, etc.).

The various critical infrastructure subsets are too numerous to all be discussed in this report. Thus, energy production and distribution, transport and the financial infrastructure are presented as representative examples of CI subsets.

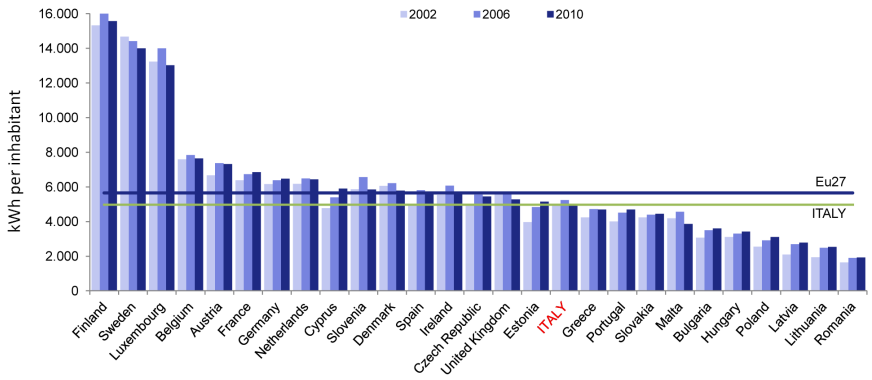


Figure 6.1: Energy utilization in EU countries. Source: Istat [31]

6.1.1 The energy sector

The energy sector has a key role in the sustainable economic development of a country, both in terms of the availability of sources, and for its impact on the environment. In 2011, electricity consumption in Italy amounted to 5,094.1 kWh per capita, an increase compared to 2010 of 0.8 %. Electricity consumption represents the energy supplied to the end users for all energy uses. Figure 6.1 shows the energy consumption, in kWh per capita, in the years 2002, 2006 and 2010, in the European member states, where the average consumption in 2010 is 5,652.4 kWh per capita. Italy has a value lower than other large European countries such as the UK, Spain, Germany and France.

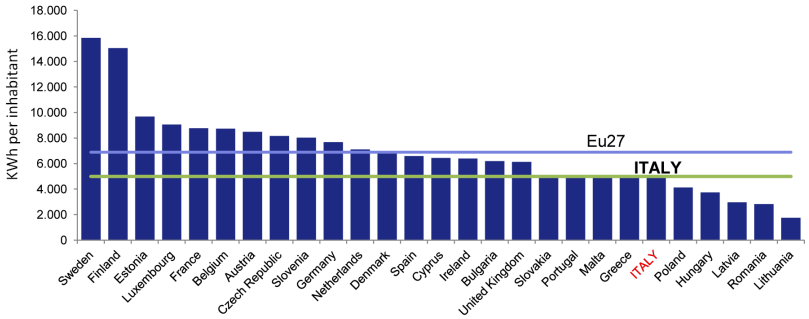


Figure 6.2: Energy production in EU countries (2010). Source: Istat [31]

Italy is characterized by its strong dependence on foreign energy markets, thus having a reduced energy production infrastructure. The internal production of electrical energy is a measure of energy autonomy. During 2011, 86.3% of the total Italian power demand was satisfied by domestic production, while the remainder by the balance between imports and exports. In the European context, with a production of 49,9 GWh per ten thousand people in 2010, Italy is below the EU27 average of 66,7 GWh (see Figure 6.2).

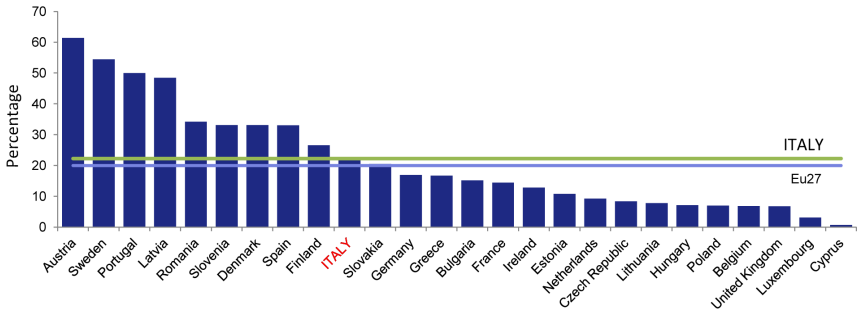


Figure 6.3: Energy consumption from renewable sources in EU countries (2010). Source: Istat [31]

In the context of the European strategy for the promotion of economic growth the development of renewable energy sources is a priority goal. In Italy, in 2010, the percentage of final energy consumption from renewable sources was 22.2%. Higher than EU average of 19.9% (see Figure 6.3). In 2011, the percentage rose to 23.8%, with an increase of 1.6 points. The target to achieve in 2020 is 26%, reflecting the fact that the trend towards the intensification and the exploitation of such energy infrastructure is growing.

Data from 2012 [33] show a strong impact on energy consumption caused by the enduring economic crisis. This impact is strongly focussed on energy consumption by industries while it only marginally affected consumption by domestic users. However, most of the trends described above are still valid; more specifically, the increasing importance of renewable energy sources together with the growth of their corresponding markets is confirmed by the latest data.

6.1.2 The transportation sector

Transportation and the related infrastructure play a key role in Italy. Just to give some numbers, the Italian motorway network, in 2010, covered 6,668 kilometers, representing about 10% of the European network. This means that Italy has a value higher than the European average.

The railway network counts an average of 5.5 km of rail for every 100 km^2 of surface area (2010 data). Compared to the average value of EU (4.9 km), Italy is ranked fifth on the EU-scale for kilometers of electrified double-track network. Figure 6.4 shows the data related to the overall railway and the electrified double-track network in the EU member states in 2010. In June 2012, the high-speed Italian line accounted for 1,434 km of track, of which 92 were under construction [10].

With regard to maritime transport, port infrastructure is becoming increasingly important in the context of new European policies for the transportation of goods and passengers. Figure 6.5 shows the data, related to

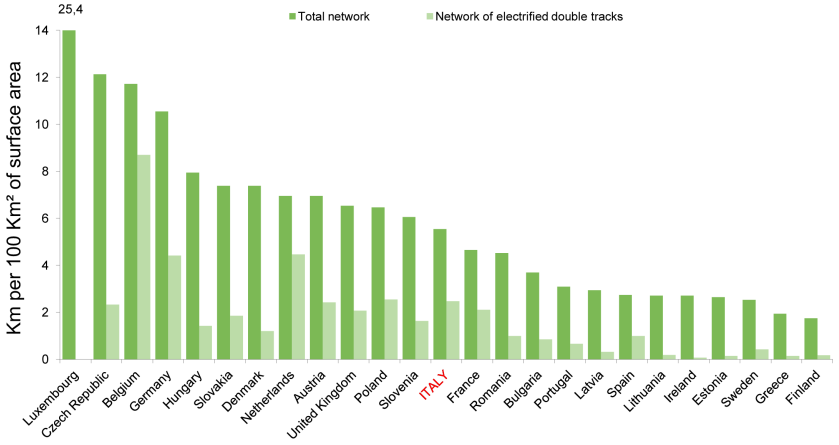


Figure 6.4: Railway network in EU countries (2010). Source: Istat [31]

2010, concerning the volume of containers transported and passengers arriving in and departing from EU ports. The figure shows that Italy is ranked fifth for the volume of container traffic and first for passenger transport (with more than 87.6 million passengers).

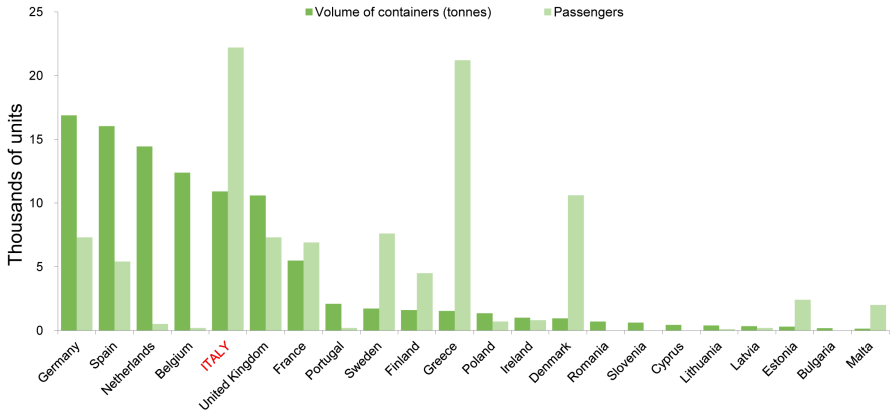


Figure 6.5: Volume of containers handled and passenger movements in EU ports (2010). Source: Istat [31]

Air transportation [31] [26] is used by larger and larger segments of the population for mobility over medium to long distances, thanks to the availability of low-cost fares. Compared to other means of transport, air transport experiences an higher level of dynamism, but it is limited by the fact that its infrastructure is close to saturation level. Figure 6.6 shows the growth, in percent, of total passenger air transport of EU member states for the years 2011 and 2012, with Italy slightly below the average. Figure 6.7 shows the

passenger traffic per resident in 2011, with Italy being significantly below EU average. Italian traffic is concentrated in Rome with 37.4 million passengers and 25.3% of Italian traffic and Milan's two airports, Malpensa Airiport with 19.1 million passengers and 12.9% of Italian traffic and Linate Airport with 9.1 million passengers and 6.1% of Italian traffic.

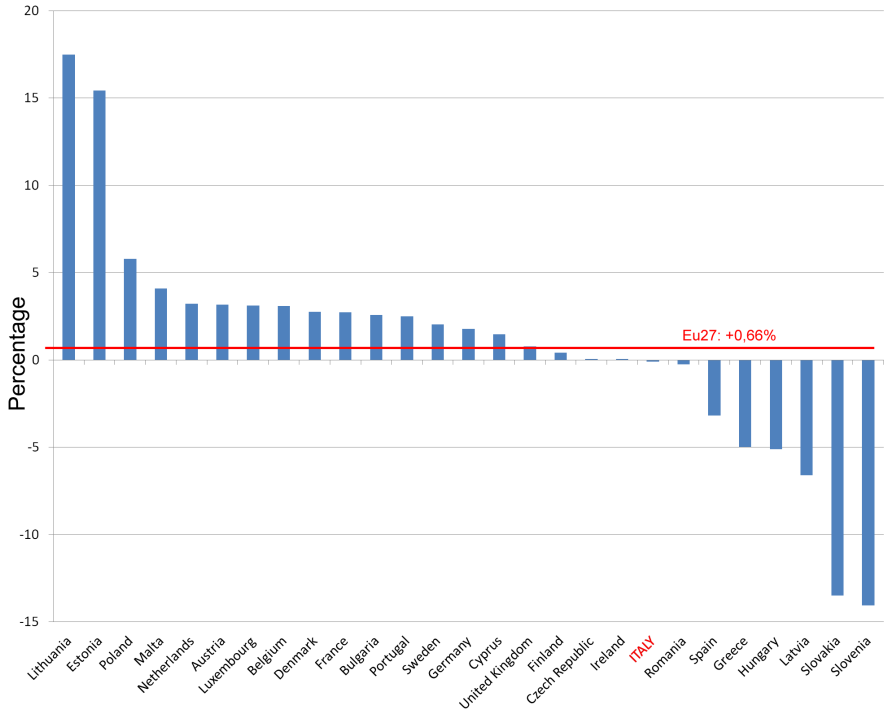


Figure 6.6: 2011/2012 growth in total passenger air transport of EU member states (in %). Source: Eurostat [1]

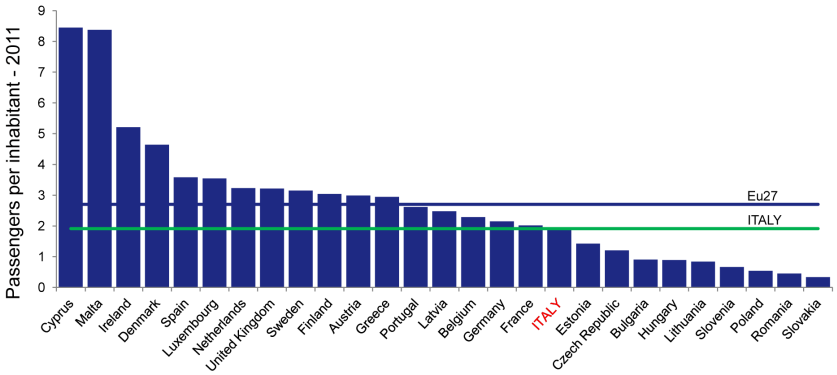


Figure 6.7: Total passenger per resident air transport of European member states (2011) [31]

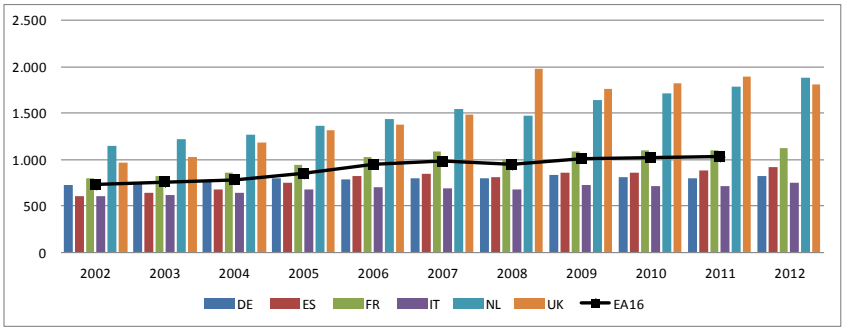


Figure 6.8: Financial interrelation ratio: asset of total economy as percentage of GDP. Source: Eurostat (non consolidated data).

6.1.3 The financial sector

In order to evaluate the importance of the financial system in Italy it is necessary to first consider some metrics generally used for this purpose: the financial interrelation ratio, the financial intermediation ratio, the credit intermediation ratio, the net financial interrelation ratio. These figures are compared with those provided by other EU countries, in order to better understand the level of financial intermediation reached in Italy. The level of financial intermediation of an economic system has been deeply analyzed by Goldsmith since the 1950s [97]. In particular, with the financial interrelation ratio (FIR), given by the ratio of the weight of the financial assets on the wealth of all sectors, Goldsmith devised a measure of the degree of financial intensity of an economic system. The higher it is, the wider is the financial deepening of the economic system. As shown in Figure 6.8 there is some evidence of a constant development of financial intensity in the Italian economy, even if, compared with other similar bank oriented system such as Germany, France and Spain, Italy has a lower percentage; the UK is historically considered a country with a market oriented financial system, and the FIR ratio seems to confirm this point of view.

Figure 6.9 shows the relation between the total assets held by financial institutions and GDP, signaling the relevance assumed by those institutions in the different countries considered. Italy, once again, falls behind compared with other countries.

Another Goldsmith's ratio, the financial intermediation ratio (FIR), analyzes the effective weight of the liabilities of the financial corporations. Comparing (see Figure 6.10) the values assumed by FIR in other EU countries, the increasing importance of intermediation in the Italian economic system is apparent. Indeed it reaches the level of other comparable countries.

By considering the credit intermediation ratio, Figure 6.11, the role assumed by the financial intermediaries in the economic system is focused on the ratio representing the weight of the loans granted by financial corporations

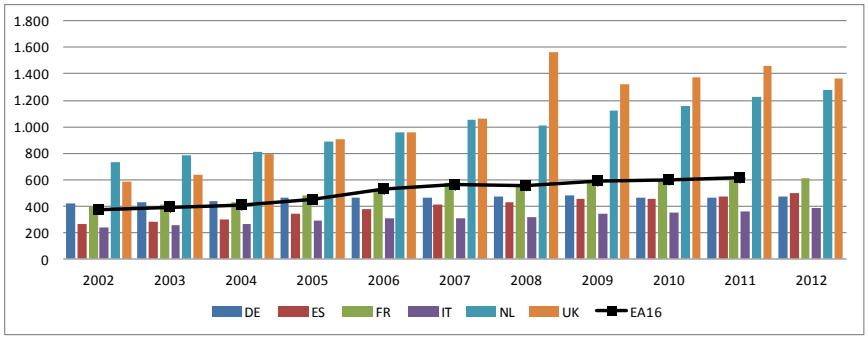


Figure 6.9: Asset of financial institutions as percentage of GDP. Source: Eurostat (non consolidated data).

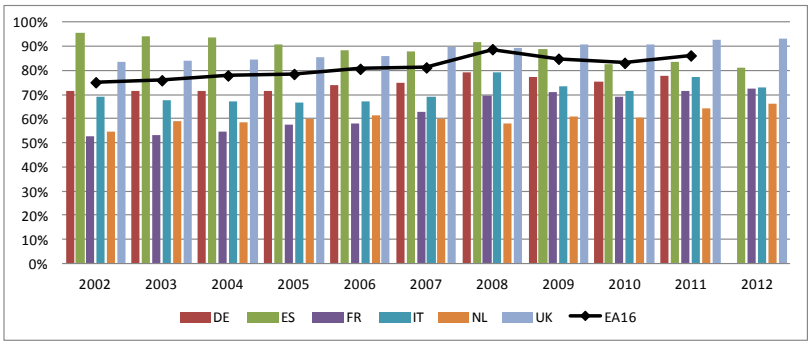


Figure 6.10: Financial intermediation ratio in selected EU countries. Source: Eurostat (non consolidated data).

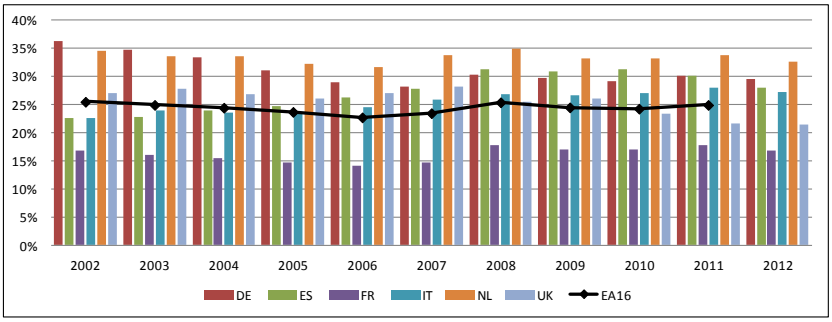


Figure 6.11: Credit intermediation ratio of selected EU countries. Source: Eurostat (non consolidated data).

against the liabilities issued by all the other sectors [92]. We note that the role played by financial intermediaries in Italy is in line with other countries considered in the analysis.

Furthermore, in Figure 6.12, the net financial interrelation ratio is considered in order to evaluate the financial deepening in Italy compared with other

EU countries. The ratio, calculated as the net financial wealth of the private sector (non financial corporations[1] and households [2]) by the GDP, confirms a good Italian performance in the period considered, even if there is a consistent reduction of the level of financial wealth compared to GDP due to the financial crisis and the economic downturn.

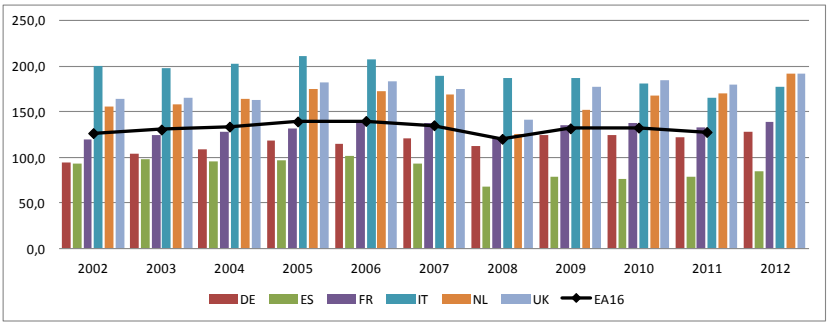


Figure 6.12: Net financial interrelation ratio of selected EU countries (%). Source: Eurostat (non consolidated data).

In Figure 6.13 the number and the typology of financial intermediaries operating in Italy in the last two years are represented.

It is evident that the banking system plays a prominent role in the Italian financial system: at the end of 2012, there were 706 banks with total assets of about 220% of GDP, of which 169 were part of 75 banking groups and account for almost 85% of total financial sector assets. The sector has become slightly more concentrated over the past decade (the five largest group own 49.4% of the assets of the banks and financial companies operating in Italy), following a major banking restructuring in the early 90s involving the divestment of state holdings. Nonetheless, there are still many small cooperative and local banks operating under different regional economic environments. Partly result of this, the system has a higher branch density (1,806 inhabitants per branch) than European peers (average of 2,168 inhabitants per branch). Another important constituent of the financial system is the insurance industry.

[1]The non-financial corporations sector comprises all private and public corporate businesses that produce goods or provide non-financial services for the market. Accordingly, the government sector excludes such public businesses and comprises central, state (regional) and local government and social security funds. The financial corporations sector comprises all private and public entities engaged in financial intermediation such as monetary financial institutions (broadly equivalent to banks), investment funds, insurance corporations and pension funds.

[2]The households sector comprises all households and includes household firms. These cover sole proprietorships and most partnerships that do not have an independent legal status. Therefore the households sector, in addition to consumption, also generates output and entrepreneurial income. In the European accounts, non-profit institutions serving households (NPISHs), such as charities and trade unions, are grouped with households. Their economic weight is relatively limited.

	Number of intermediaries	
	2011	2012
Banking Groups	77	75
Investment firm groups	20	19
Banks	740	706
limited company banks	214	197
cooperative banks (banche popolari)	37	37
mutual banks (banche di credito cooperativo)	411	394
branches of foreign banks	78	78
Investment firms	102	101
Asset management companies and SICAVs	190	172
Financial companies entered in the special register under Article 107 of the Consolidated Law on Banking	188	186
Financial companies entered in the general register under Article 106 of the Consolidated Law on Banking	782	658
Electronic money institutions	3	3
Payment institutions	34	44
Other supervised intermediaries (Bancoposta and Cassa Depositi e Prestiti)	2	2
Insurance Companies		135
operating in life insurance sector		52
In non-life insurance sector		69
In both sectors		14

Figure 6.13: The structure of the Italian financial system (2011-2012). Source: Banca d'Italia.

At the end of 2012, there were 135 Italian insurance companies (52 operating exclusively in life insurance, 69 in non-life insurance and 14 in both sectors). The degree of concentration of the insurance industry is high by European standards, in particular for the non-life sector. Banks play an important role in the ownership structures of Italian insurance companies, albeit not as great as in the other asset management sectors (investment funds and individually managed portfolios)[3]. The banking sector also plays a large role in the distribution of standardized insurance products.. In the last fifteen years, the insurance sector has assumed a prominent role in the asset management industry in Italy: the technical provisions have quadrupled since 1998, reaching €486 billion, and their asset under management has increased from 17 to 35%. Despite this rapid growth, Italy's insurance industry is still smaller compared with other European countries: insurance products amount to 12% of household wealth in Italy (5% in 1998), compared with 33% in France and 18% in Germany. These differences are basically due to the supply structure: while in Italy, France and Germany traditional policies, which offer subscribers a minimum guaranteed return, are life insurance companies' leading product, in the United Kingdom the main sector is index- and unit-linked policies, in which financial risk is typically borne by the insured party. As underlined above, banks tend to monopolize the financing of the whole economy in Italy; this is also testified to by the scarce recourse to the capital market for non financial corporations in order to finance their investments both with debt and equity capital. Italian companies made net issues worth €91 billion in 2012 (see Figure 6.14), and most of those were made by banks (€83 billion, as against €66 billion in 2011), while other financial institutions continued

[3]In 2012 the assets of insurance companies controlled by domestic banking and financial groups made up 19% of the total.

	Net Issues			Stocks			as a % of GDP
	2010	2011	2012	2010	2011	2012	
Banks	-11,8	66,33	83,153	807,045	873,618	956,739	61
Other financial corporations	-36,458	-4,328	-6,132	243,398	239,125	233,022	15
Non-financial corporations	12,373	-100	13,576	89,874	90,018	103,615	7
Total	**-35,885**	**61,902**	**90,597**	**1,140,317**	**1,202,761**	**1,293,376**	**83**

Figure 6.14: Medium- and long-term bonds of Italian banks and firms (2010-2012). Source: Banca d'Italia.

Change in prices	-48.7	20.7	-8.7	-24.0	10.2
Listed companies (number at end of year)	336	332	332	328	323
of which: Italian	*294*	*291*	*291*	*287*	*282*
Market capitalization of Italian companies	374,702	457,126	425,099	332,374	365,466
per cent of GDP	*23.8*	*30.1*	*27.4*	*21.1*	*23.3*
Percentage composition:					
industrials	*33*	*37*	*41*	*45*	*47*
insurance	*11*	*9*	*7*	*7*	*8*
banking	*25*	*26*	*20*	*17*	*18*
financials	*3*	*2*	*3*	*3*	*2*
services	*28*	*26*	*28*	*29*	*25*
Total	100	100	100	100	100
Dividends	39,072	21,309	16,036	17,009	13,207
Earnings/price ratio	15.6	5.3	7.6	9.0	7.2
Dividend yield	8.0	5.0	3.8	5.1	4.2

Figure 6.15: Main indicators of the Italian stock exchange (2012). Source: Banca d'Italia.

to make net redemptions. According to Dealogic data, gross placements on the international market by issuers belonging to Italian non-financial groups increased from €19 billion to €29 billion; nearly 80% of the new issues were accounted for by six large groups (Enel, Eni, Fiat, Snam, Telecom Italia and Terna).

The funds raised through capital increases by listed companies were slightly down on 2011, €10.1 billion, as against 11.9 billion (Figure 6.15). Once again financial institutions raised most part of the capital: one bank raised about three quarters of the total amount, insurance companies about one fifth and non-financial corporations the remainder. In 2012 the ratio of Italian companies' market capitalization to GDP rose from 21 to 23%, while in the other major advanced countries the ratio at the end of the year was much higher: 45% in Germany, 63% in France, 107% in the United States and 156% in the United Kingdom. The average daily turnover of shares on the Italian Stock Exchange was significantly lower than in the previous year.

6.2 Maturity of protection against cyberattacks

When it come to the protection of CI, we are living in a real global emergency in which nothing and no one can no longer be considered secure. Every field

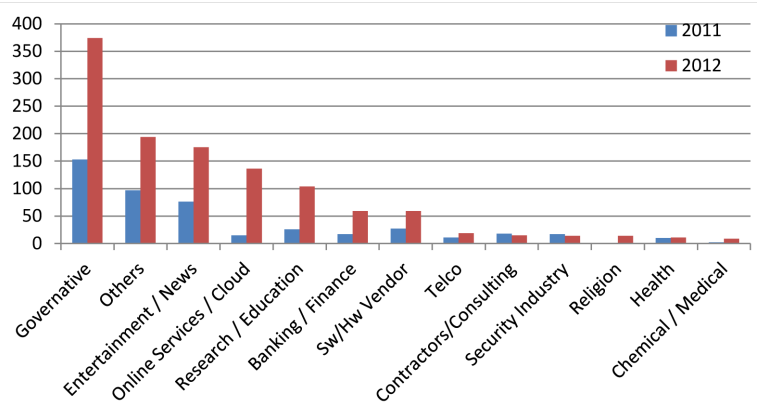


Figure 6.16: Known victims of cyberattacks in Italy. Source: Clusit [32].

has become a potential target: citizens, companies, governments. The report presented by McAfee and CSIS [7] in 2011, highlights an incredible increase in cyberattacks on critical infrastructure, which continue to be ill-prepared to address such threats. Conventional protection is no longer adequate to block threats, which are becoming more sophisticated and are beyond the majority of control systems.

Figure 6.16 shows the known victims of cyberattacks in Italy, classified by sector of competence. For each class the numerical data of 2011 and 2012 are shown, in order to highlight this recent trend.

Sector	**2011**	**2012**	**Total**	**Delta**
Mil, LEAs, Intelligence	153	374	527	144%
Others	97	194	291	100%
Entertainment/News	76	175	251	130%
Online services/Cloud	15	136	151	807%
Research/Education	26	104	130	300%
Banking/Finance	17	59	76	247%
Softw./Hardw. Vendor	27	59	86	119%
Telco	11	19	30	73%
Contractors/Consulting	18	15	33	-17%
Security industry	17	14	31	-18%
Religion	0	14	14	100%
Health	10	11	21	10%
Chemical/Medical	2	9	11	350%
Total	469	1183	1652	152%

Table 6.1: Evolution of number of attacks by sector. Source: Clusit [32].

The data show that just two sectors have witnessed a decrease in attacks, whilst the other sectors have all suffered from an increased number of attacks, sometimes even greater than 500%. Deltas for each category are detailed

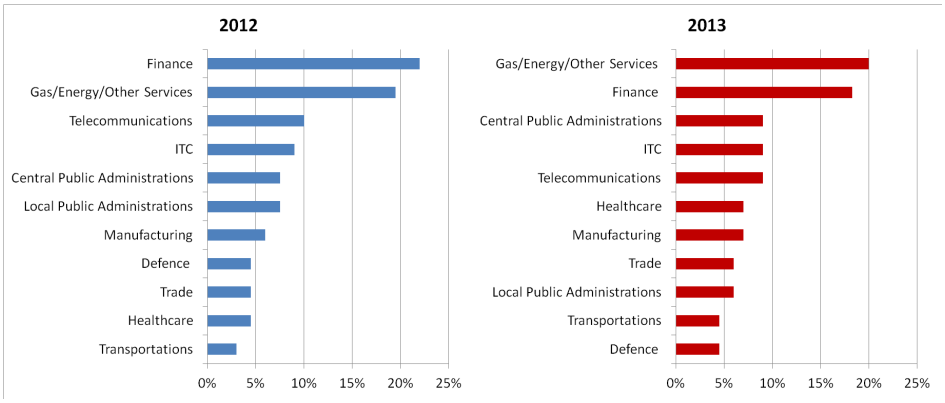


Figure 6.17: Critical infrastructure protection investments (2012-2013). Source: Clusit [32].

in Table 6.1. These data are useful to understand how the phenomenon of information security in every area, consequently also in CI, is very delicate and needs to be appropriately addressed.

The awareness of such a need is evident and despite the economic crisis, it can be observed that the ICT security market continues to have a stable and positive trend, as a sign that companies, users, and countries in general, are more and more aware of the need for security and thus invest in it. The CLUSIT data [32] provide an important comparison, using a sample of Italian companies, of investments made in 2012 and the forecast of investments in 2013. This is shown in Figure 6.17.

At the same time it appears evident that the role of governments should be crucial in encouraging security by collaborating with industry and by adopting proper regulations.

The status of computer security awareness in Italy is even worse when considering normal users and how much they protect against cyberattacks and cyberfrauds. Despite the widespread and increasing use of the Internet among Italians, there is still a low level of awareness of the risks associated with careless use of the Internet. Consequently people buy products and services which are inherently insecure, or implement and configure in an insecure manner, without any guarantee or protection. As reported in figure 6.18, the main consequence [27] in Italy is that about 44% of PCs are attacked by malware while browsing the Internet, compared with 20% in Denmark [106]. The main cause of the spread of attacks is the limited use of threat protection solutions. Only 33% of Italian users (the percentage rises to 44% on a global scale) actually use software able to ensure the necessary security of their data and only 45% of Italian users employ privacy settings to control the information they share with their contacts. In addition, 44% of users in Italy (about 40% in the world) do not use complex passwords or change their keywords frequently.

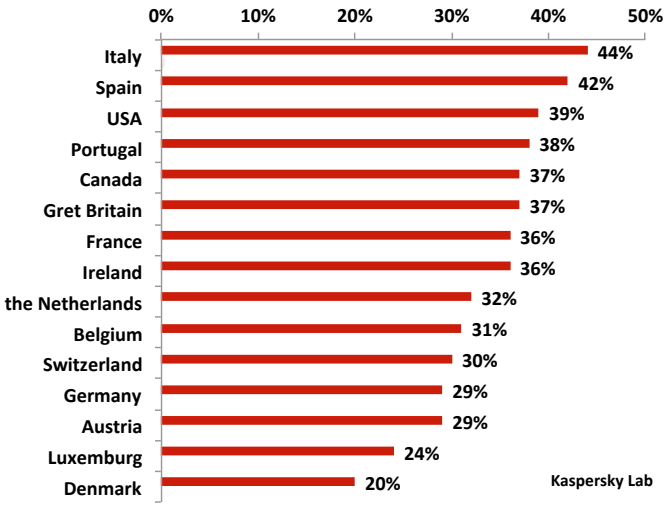


Figure 6.18: Percentage of personal computers attacked by malware while browsing the Internet. Source Kaspersky Lab, 2012.

The number of criminal activities and the level of sophistication of attacks do not correspond to a proportional growth of attention.

6.3 The cost of cybercrime in Italy

Costs associated with cybercriminal activities have been recently reported [77]. Currently, there are no official statistics on the cost of cybercrime in Italy. The only available statistics come from the private sector. According to the Norton Cybercrime Report [29] (September 2012), which analyzes the impact of cybercrime on consumer users, the total net cost of consumer cybercrime in Italy in the previous 12 months amounts to 2.45 billion euros, whereas the cost at global level amounts to 110 billion US dollars (about 85 billion euros). The report estimates the number of cybercrime victims to be 8.9 million people, about one third of Internet users active in Italy in 2012 [32]. This results in an average cost per person of 275 euros (more than the global average cost per person, estimated to be 197 US dollars). In particular, Norton registers an increasing number of victims among mobile and social network users, suggesting that cybercrime is evolving towards new technologies. Indeed, approximately 17% of adults in Italy have been victim of social or mobile cybercrime in 2012, and about 10% of social network users have had someone hack into their profile.

In the business context, an analysis led by the Ponemon Institute[103] estimates the cost of data breach in Italy, in terms of direct, indirect and opportunity costs incurred by an organization in response to data breach.

The analysis, conducted in 2011 and published in March 2012, reports the average cost of data breach per record (i.e., the total cost divided by the number of compromised records) and the average total organizational cost of data breach. As shown in figure 6.19, the average cost per record incurred by Italian organizations is 78 euros. This cost accounts for a range of business costs: detection (26 euros), notification (3 euros), ex-post response (22 euros) and lost business (27 euros). The majority of the total cost (41 euros) is due to indirect costs, while the remaining part (37 euros) is due to direct costs.

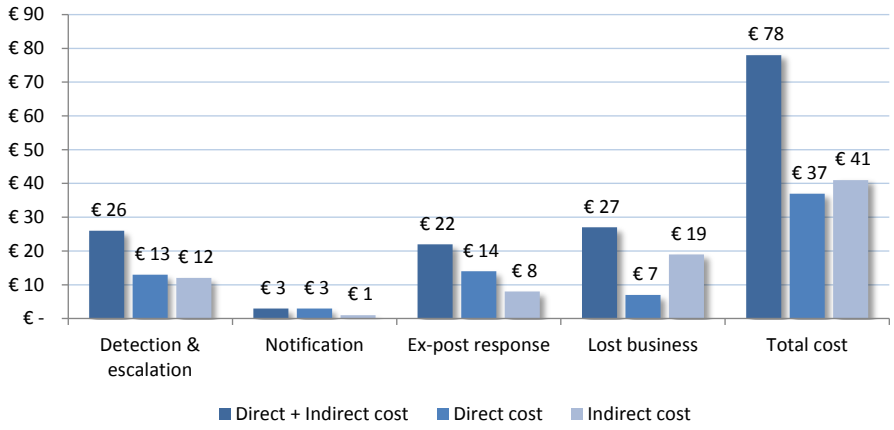


Figure 6.19: Average cost of data breach per record. Source: Ponemon Institute, 2011.

Figure 6.20 shows the average total organizational cost of a data breach (1,384,798 euros) and its constituent costs. Both figures show that the largest cost is represented by lost business. This cost is mainly due to abnormal turnover in customers (a higher than average loss of customers for the organization) and reputation loss. Indeed, customers often abandon the organization after a data breach. The analysis also revealed that the primary cause of data breach is negligence (39%), followed by system glitches (33%) and malicious or criminal attacks (28%). However, malicious attacks are on average the most costly.

6.4 Italian cybersecurity readiness

In 2013 the "Cyber Intelligence and Information Security Research Center" at the University of Rome La Sapienza published a report on the technical maturity of the main actors on the Italian cybersecurity landscape [77]. The report included a study conducted by interviewing several Italian CI actors in order to assess their level of preparation against cyberattacks. The study was performed by collecting answers to an extensive questionnaire in which several

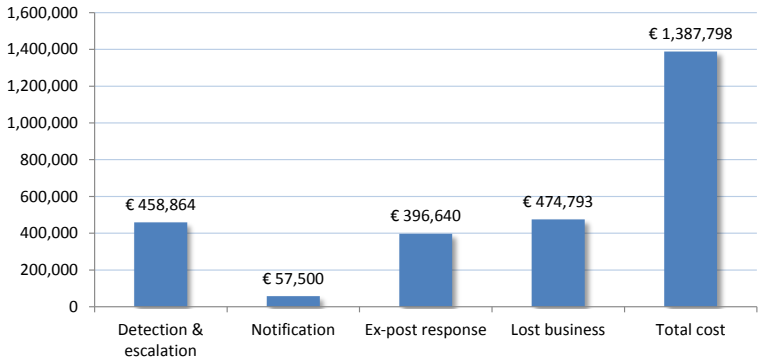


Figure 6.20: Average total organizational cost of data breach per record. Source: Ponemon Institute, 2011.

aspects of cybersecurity were investigated, ranging from organizational aspects to more technical issues. The questionnaire was submitted to organizations belonging to the set of critical infrastructure reported in section 1: an effective cyberattack on each of these companies could have serious consequences, from an economic point of view and/or from a safety point of view.

One of the more interesting results reported in that document is represented by the Cybersecurity Readiness Index, a complex metric based on the analysis of multiple questions included in the questionnaire, which the report authors introduced. It captures with a single score, multiple correlated aspects that impact the readiness of CI against cyberthreats.

The cybersecurity readiness index is a composite measure of the capacity and willingness of an organization to face cyberthreats. The index covers four distinct aspects: awareness, defense, policy and external independency. A positive cybersecurity habit for an organization is considered to be the one which is able to cover the largest area on a radar chart taking into account the four aspects. Thus the cybersecurity readiness index reflects the dimension of this area. The complete structure of the score system that computes the cybersecurity readiness index based on the four indexes is reported in [77].

Awareness index - This assesses the situational awareness related to cyberrisks of the organization. As an example an organization that monitors the security levels guaranteed by its subcontractors will possibly have a larger awareness index. On the contrary, a company that frequently experiences abnormal behaviors in its IT infrastructure that are not adequately analyzed will see its awareness score shrink.

Defense index - This assesses the capacity of an organization to protect itself from a cyberattack. This considered the evaluation of the defense mechanisms and tools employed by an organization. The questionnaire

contained a selection of well-know strategies employed by several organization to protect their assets. This index checks how well the organization is equipped with such tools and if it is adequately trained to use them. Notice that the defense index is somewhat correlated with the awareness index. Some responses that positively impact the defense index, also positively impact the awareness index. This correlation is well-grounded since the implementation of strong defense mechanisms implies a good level of awareness.

Policy index - This assesses the implementation of security related policies. A high score in this index shows compliance to several security policies and their constant update. As for the defense index there is a strong correlation of the policy index with the awareness index since the adoption of updated security policies show an increased awareness.

External independency index - This assesses the correlation between internal systems and external providers. A low score on this index indicates a negative correlation of the organization internal mechanisms to external providers since the fault of an external provider could impact on its possibility to deliver the core product of its business. A high score on this index shows an organization that relies minimally on external services that could impact on its security. Note that such high scores imply larger operational costs as the organization has to insource software services without the involvement of third parties.

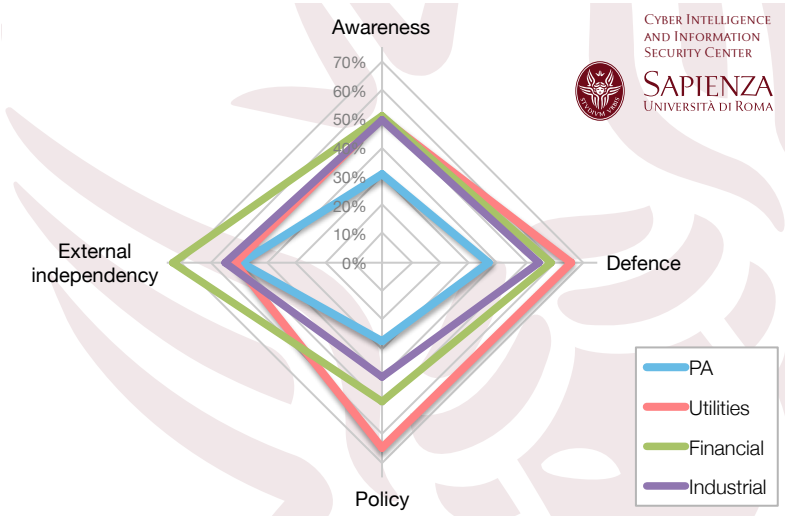


Figure 6.21: Cybersecurity readiness index: awareness, defense, policy and external dependencies indexes per group. Source: CIS, 2013 [77].

The analysis was performed by dividing the respondent organizations into four groups: public administration bodies (PA), utilities (energy and telecomunication), financial organizations (like banks), industrial companies. A radar chart is depicted in Figure 6.21 showing the results of the cybersecurity readiness index per group.

The utility group covers the largest area in the ranking. It scores better than other groups along two axes, namely defense and policy. It has also a high score on awareness. Nevertheless, it suffers from a low external independence; this problem is shared by all the other groups, with the exception of the financial group that still seems reluctant to heavily rely on external service providers.

The financial group also exhibits a large covered area in the radar chart by showing high values for external independency, defense and awareness indexes. Surprisingly, it does not score as expected on policy index. However, it should be noted that some of the questions that influenced the policy index were related to the specific policy imposed by the EU directive 2008/114/EC on European CI that financial organizations are not obliged to comply with.

The industrial group is the third one in the ranking, showing a high level of awareness and a good defense index while lagging behind in policy adoption. The PA group shows a low degree of cybersecurity readiness with respect to the other groups; indeed, the area covered by the radar plot is the smallest among all the groups. It has by far the lowest indexes on policy, defense and awareness.

Bibliography

[1] Air transport statistics - Statistics Explained (2013/12/10). Available at: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=avia_paoc&lang=en.

[2] Eurocontrol website. http://www.eurocontrol.int/articles/what-air-traffic-management.

[3] European union agency for network and information security (enisa), national cyber security strategy list, http://www.enisa.europa.eu/activities/Resilience-and-CIIP/national-cyber-security-strategies-ncsss/national-cyber-security-strategies-in-the-world.

[4] Presidential Policy Directive – Critical Infrastructure Security and Resilience | The White House.

[5] SESAR project website. http://ec.europa.eu/transport/modes/air/sesar/.

[6] Shamoon the Wiper - Copycats at Work - Blog post http://www.securelist.com/en/blog/208193786/Shamoon_the_Wiper_Copycats_at_Work.

[7] Shodan website. http://www.shodanhq.com.

[8] Single European Sky For a performant air traffic system in Europe. Eurocontrol dossier. Available at: http://www.eurocontrol.int/dossiers/single-european-sky.

[9] The Comprehensive National Cybersecurity Initiative. Whitehouse official website. Available at: http://www.whitehouse.gov/sites/default/files/cybersecurity.pdf.

[10] Wikipedia - Alta velocità ferroviaria.

[11] Vulnerability Assessment of the Transportation Infrastructure Relying on the Global Positioning System. Available at: http://www.navcen.uscg.gov/pdf/vulnerability_assess_2001.pdf. Volpe National Transportation Systems Center - U.S. Department of Transportation Research and Innovative Technology Administration, Cambridge, Massachusetts, 2001.

[12] Protezione delle infrastrutture critiche - la realtà italiana. Presidenza del Consiglio dei Ministri - Dipartimento per l'innovazione e le tecnologie - Gruppo di lavoro sulla protezione delle infrastrutture critiche informatizzate http://www.vigilfuoco.it/aspx/ReturnDocument.aspx?IdDocumento=2832, 2004.

[13] Directive of the council on a european programme for critical infrastructure protection. COM(2006) 786, Brussels., 2006.

[14] Directive of the council on the identification and designation of european critical infrastructure and the assessment of the need to improve their protection. COM(2006) 787, Brussels., 2006.

[15] A strategy for a secure information society – "dialogue, partnership and empowerment". COM(2006) 251, Brussels., 2006.

[16] Directive of the council on on the identification and designation of european critical infrastructures and the assessment of the need to improve their protection. COUNCIL DIRECTIVE 2008/114/EC, 2008.

[17] European union directive 2008/114/ec, 2008.

[18] Individuazione delle infrastrutture critiche informatiche di interesse nazionale. Italian Ministry for the Interior Decree G.U. 30 aprile 2008, n. 101, 2008.

[19] ICS-CERT - U.S. Department of Homeland Security - Alert (ICS-ALERT-10-301-01) Control System Internet Accessibility - http://ics-cert.us-cert.gov/alerts/ICS-ALERT-10-301-01, 2010.

[20] Proposal for a regulation of the european parliament and of the council concerning the european network and information security agency (enisa). COM(2010) 521, Brussels., 2010.

[21] Smart grid cyber security strategy, architecture, and high-level requirements. NISTIR 7628 Guidelines for Smart Grid Cyber Security: Vol. 1, 2010.

[22] UNISIG SUBSET-037, Euroradio FIS, Version 2.3.0. European Railway Agency. Available at: `http://www.era.europa.eu/Document-Register/Pages/UNISIGSUBSET-037.aspx`, 2010.

[23] Security profile for wide-area monitoring, protection, and control. version 0.8. The Advanced Security Acceleration Project (ASAP-SG), 2011.

[24] The Future of the Electric Grid. Available online: `http://web.mit.edu/mitei/research/studies/the-electric-grid-2011.shtml`. Massachusetts Institute of Technology (MIT), 2011.

[25] Arizona-Southern California Outages on September 8, 2011. FERC/NERC Staff Report on the September 8, 2011 Blackout, 2012.

[26] *Conto Nazionale delle Infrastrutture e dei Trasporti. Anni 2010 - 2011.* Istituto Poligrafico e Zecca dello Stato S.p.A. - Roma, 2012.

[27] Kaspersky securelist. `http://www.securelist.com/en/analysis/204792244/The_geography_of_cybercrime_Western_Europe_and_North_America`, 2012.

[28] Network of Excellence (NESoS) Deliverable: Selection and Documentation of the Two Major Application Case Studies. *Springer Lecture Notes in Computer Science*, 2012.

[29] Norton 2012 cybercrime report - italy. `http://now-static.norton.com/now/en/pu/images/Promotions/2012/cybercrimeReport/NCR-Country_Fact_Sheet-Italy.pdf`, 2012.

[30] Cybersecurity strategy of the european union: An open, safe and secure cyberspace. European Commission, Brussels, JOIN(2013) 1 final, 2013.

[31] *Noi Italia. 100 statistiche per capire il paese in cui viviamo.* ISTAT, 2013.

[32] *Rapporto Clusit 2013 sulla sicurezza ICT in Italia.* Cardi Editore galleria San Babila 4 20122 Milano, 2013.

[33] Relazione annuale sullo stato dei servizi. Technical report, Autorità per l'energia elettrica e il gas, 2013.

[34] M. Afzaal, C. Di Sarno, L. Coppolino, S. D'Antonio, and L. Romano. A resilient architecture for forensic storage of events in critical infrastructures. In *High-Assurance Systems Engineering (HASE), IEEE 14th International Symposium on*, pages 48–55, 2012.

[35] M. Afzaal, C. Di Sarno, S. Dantonio, and L. Romano. An Intrusion and Fault Tolerant Forensic Storage for a SIEM System. In *Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS)*, pages 579–586, 2012.

[36] ALARP - A railway automatic track warning system based on distributed personal mobile terminals - Project Contract FP7-SST-2010-234088, `http://www.alarp.eu`.

[37] M. Albanese, S. Jajodia, A. Pugliese, and V. S. Subrahmanian. Scalable analysis of attack scenarios. In *ESORICS*, pages 416–433, 2011.

[38] M. Albanese, A. Pugliese, and V. S. Subrahmanian. Fast activity detection: Indexing for temporal stochastic automaton-based activity models. *IEEE Transactions on Knowledge and Data Engineering*, 25(2):360–373, 2013.

[39] L. Allodi and F. Massacci. A preliminary analysis of vulnerability scores for attacks in wild. In *Proceedings of the 2012 ACM CCS Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*, 2012.

[40] L. Allodi and F. Massacci. How CVSS is DOSsing your patching policy (and wasting your money). *BlackHat USA*, 2013.

[41] P. Ammann, D. Wijesekera, and S. Kaushik. Scalable, graph-based network vulnerability analysis. In *ACM Conference on Computer and Communications Security*, pages 217–224, 2002.

[42] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal. Causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance. *IEEE Transactions on Power Systems*, 20(4):1922–1928, 2005.

[43] M. Anghel, K. Werley, and A. Motter. Stochastic model for power grid dynamics. In *40th Annual Hawaii International Conference on System Sciences (HICSS)*, pages 113–113, 2007.

[44] Asis International. General Security Risk Assessment Guideline. Available at: `http://www.asisonline.org/guidelines/guidelinesgsra.pdf`, 2008.

[45] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr. Basic concepts and taxonomy of dependable and secure computing. *Dependable and Secure Computing, IEEE Transactions on*, 1(1):11–33, 2004.

[46] W. Baker, M. Howard, A. Hutton, and C. D. Hylender. 2012 data breach investigation report. Technical report, Verizon, 2012.

[47] G. Baldini, I. N. Fovino, M. Masera, M. Luise, V. Pellegrini, E. Bagagli, G. Rubino, R. Malangone, M. Stefano, and F. Senesi. An early warning system for detecting GSM-R wireless interference in the high-speed railway infrastructure. *International Journal of Critical Infrastructure Protection*, 3(3–4):140 – 156, 2010.

[48] R. Baldoni, G. Lodi, L. Montanari, G. Mariotta, and M. Rizzuto. Online black-box failure prediction for mission critical distributed systems. In *SAFECOMP*, pages 185–197, 2012.

[49] M. Beccuti, S. Chiaradonna, F. Di Giandomenico, S. Donatelli, G. Dondossola, and G. Franceschinis. Quantification of dependencies between electrical and information infrastructures. *International Journal of Critical Infrastructure Protection*, 5(1):14–27, 2012.

[50] D. Bell and L. J. L. Padula. Secure computer systems: Unified exposition and multics interpretation. Technical report, MITRE Corp., Bedford, MA, Tech. Rep. ESD-TR-75-306„ 1975.

[51] C. Bennett and S. Wicker. Decreased time delay and security enhancement recommendations for AMI smart meter networks. In *Innovative Smart Grid Technologies (ISGT)*, 2010.

[52] R. Berthier and W. H. Sanders. Specification-based intrusion detection for advanced metering infrastructures. In *Proceedings of the 2011 IEEE 17th Pacific Rim International Symposium on Dependable Computing*, PRDC '11, pages 184–193, Washington, DC, USA, 2011. IEEE Computer Society.

[53] R. Bloomfield, R. Bloomfield, I. Gashi, and R. Stroud. How Secure Is ERTMS? In *Computer Safety, Reliability, and Security*, volume 7613 of *Lecture Notes in Computer Science*, pages 247–258. Springer Berlin Heidelberg, 2012.

[54] R. Bloomfield, N. Chozos, and P. Nobles. Infrastructure interdependency analysis: Requirements, capability and strategy. Technical report, Available online `http://www.csr.city.ac.uk/projects/cetifs/d418v13_public.pdf`, 2009.

[55] R. Bloomfield, K. Salako, D. Wright, N. Chozos, and P. Nobles. Infrastructure interdependency analysis: an introductory research review. In *Adelard document reference D/422/12101/4 available for download at* `http://www.csr.city.ac.uk/projects/cetifs.html`, 2009.

[56] J. Broder. Risk analysis and the security survey. Butterworth-Heinemann, 2006.

[57] M. Brunner and E. M. Suter. *International CIIP Handbook 2008/2009.* Center for Security Studies, ETH Zurich, 2008.

[58] S. Buchegger and J.-Y. Le Boudec. Performance analysis of the confidant protocol. In *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing*, MobiHoc '02, pages 226–236, New York, NY, USA, 2002. ACM.

[59] L. Buttyan, D. Gessner, A. Hessler, and P. Langendoerfer. Application of wireless sensor networks in critical infrastructure protection: challenges and design options. *Wireless Communications, IEEE*, 17(5):44–49, 2010.

[60] G. S. Capra and U. Center. *Protecting Critical Rail Infrastructure.* USAF Counterproliferation Center, Air University, 2006.

[61] A. A. Cardenas, T. Roosta, and S. Sastry. Rethinking security properties, threat models, and the design space in sensor networks: A case study in scada systems. *Ad Hoc Netw.*, 7(8):1434–1447, 2009.

[62] E. Casalicchio, E. Galli, and S. Tucci. Federated agent-based modeling and simulation approach to study interdependencies in it critical infrastructures. In *Distributed Simulation and Real-Time Applications, 2007. DS-RT 2007. 11th IEEE International Symposium*, pages 182–189, 2007.

[63] M. Casciaro. Modelling and analysis of security in an automatic work site protection system. Master's thesis, University of Florence, 2013.

[64] A. Ceccarelli, A. Bondavalli, J. Figueiras, B. Malinowsky, J. Wakula, F. Brancati, C. Dambra, and A. Seminatore. Design and implementation of real-time wearable devices for a safety-critical track warning system. In *HASE*, pages 147–154, 2012.

[65] A. Ceccarelli, I. Majzik, D. Iovino, F. Caneschi, G. Pintér, and A. Bondavalli. A Resilient SIL 2 Driver Machine Interface for Train Control Systems. In *DepCoS-RELCOMEX*, pages 365–374, 2008.

[66] CENELEC. Railway Applications–Communication, signalling and processing systems–Safety related electronic systems for signalling. *EN 50129*, 1999.

[67] CENELEC. Railway Applications–The Specification and demonstration of Reliability, Availability, Maintainability, and Safety (RAMS). *EN 50126*, 1999.

[68] CENELEC. *EN 50159-1 - Railway applications - Communication, signalling and processing systems - Part 1- Safety related communication in closed transmission systems*, 2001.

[69] CENELEC. *EN 50159-2 - Railway applications - Communication, signalling and processing systems - Part 2- Safety related communication in open transmission systems*, 2001.

[70] R. Chandia, J. Gonzalez, T. Kilpatrick, M. Papa, and S. Shenoi. Security strategies for SCADA networks. In *Critical Infrastructure Protection*, volume 253 of *IFIP International Federation for Information Processing*, pages 117–131. Springer US, 2007.

[71] Y. Chen, B. W. Boehm, and L. Sheppard. Value driven security threat modeling based on attack path analysis. In *HICSS*, page 280, 2007.

[72] S. Cheung, B. Dutertre, M. Fong, U. Lindqvist, K. Skinner, and A. Valdes. Using Model-based Intrusion Detection for SCADA Networks. In *Proceedings of the SCADA Security Scientific Symposium*, 2007.

[73] S. Chiaradonna, F. D. Giandomenico, and P. Lollini. Definition, Implementation and Application of a Model-based Framework for the Analysis of Interdependencies in Electric Power Systems Protection. *International Journal of Critical Infrastructure (IJCIP), Elsevier*, 4(1):24–40, April 2011.

[74] B. Conway. Wall Street's need for trading speed: The nanosecond age. Available at: `http://blogs.wsj.com/marketbeat/2011/06/14/wall-streets-need-for-trading-speed-the-nanosecond-age/`. *The Wall Street Journal*, 2011.

[75] A. Corsaro. CARDAMOM: a next generation mission and safety critical enterprise middleware. In *Third IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems (SEUS)*, pages 73–74, 2005.

[76] A. Costin and A. Francillon. Ghost in the Air(Traffic): On insecurity of ADS-B protocol and practical attacks on ADS-B devices. In *Black Hat*, Las Vegas, NV, USA, July 27–August 1 2012.

[77] Cyber Intelligence and Information Security Research Center of Sapienza - Università di Roma. *2013 Italian Cyber Security Report -*

Critical Infrastructure and Other Sensitive Sectors Readiness. Casa Editrice Università La Sapienza, 2013.

[78] S. Delamare, A. Diallo, and C. Chaudet. High-level modeling of critical infrastructures' interdependencies. *International Journal of Critical Infrastructures*, 5(1/2), 2009.

[79] G. Devarajan. Unraveling scada protocols: Using sulley fuzzer. In *DefCon 15 Hacking Conference*, 2007.

[80] G. Dini and M. Tiloca. Considerations on security in zigbee networks. In *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC)*, pages 58–65, 2010.

[81] D. Doberstein. *Fundamentals of GPS Receivers: A Hardware Approach*. Springer New York, 2011.

[82] D. V. Dollen. Report to NIST on the Smart Grid Interoperability Standards Roadmap. EPRI Contract No. SB1341-09-CN-0031-Deliverable 7, 2009.

[83] D. Druidi. Railroad-related work injury fatalities. Available at: `http://www.pcsforum.org/library/files/1159904563-TSWG_INL_CIP_Tool_Survey_final.pdf`. Technical Report Monthly Labor Review, 2007.

[84] H. M. El-Bakry and N. Mastorakis. Design of anti-GPS for reasons of security. In *Proceedings of the international conference on Computational and information Science*, CIS'09, pages 480–500, Stevens Point, Wisconsin, USA, 2009. World Scientific and Engineering Academy and Society (WSEAS).

[85] Enisa. Analysis of cyber security aspects in the maritime sector, 2011.

[86] Eurocontrol. What is air traffic management? Building skyways, managing traffic flows and capacity, handling flights. Available online: `http://www.eurocontrol.int/articles/what-air-traffic-management`, 2011.

[87] European Commission, Joint Research Centre. JRC, Smart Grid projects in Europe: lessons learned and current developments. Technical report, Joint Research Centre Reference Report `http://www.pcsforum.org/library/files/1159904563-TSWG_INL_CIP_Tool_Survey_final.pdf`, 2011.

[88] N. Falliere, L. O. Murchu, and E. Chien. W32. Stuxnet Dossier. Available online `http://www.symantec.com/content/en/us/enterprise/`

media/security_response/whitepapers/w32_stuxnet_dossier.pdf, 2011.

[89] Federal Aviation Administration. Automatic Dependent Surveillance Broadcast (ADS-B) Out Performance Requiremens to Support Air Traffic Control (ATC) Service. Final Rule, 14 CFR part 91, Federal Register 75 (103), 2010.

[90] C. Finke, J. Butts, R. Mills, and M. Grimaila. Enhancing the security of aircraft surveillance in the next generation air traffic control system. *International Journal of Critical Infrastructure Protection*, 6(1):pp. 3–11, 2013.

[91] F. Flammini, A. Gaglione, N. Mazzocca, and C. Pragliola. Quantitative security risk assessment and management for railway transportation infrastructures. In *Critical Information Infrastructure Security*, volume 5508 of *Lecture Notes in Computer Science*, pages 180–189. Springer Berlin Heidelberg, 2009.

[92] C.-B. G., C.-S. J., and S. L. *The measurement of financial intermediation in Japan.* Japan and the World Economy, 2008.

[93] M. Garcia. Vulnerability assessment of physical protection systems. Butterworth-Heinemann, 2005.

[94] A. Garofalo, C. Di Sarno, L. Coppolino, and S. D'Antonio. A GPS Spoofing Resilient WAMS for Smart Grid. In *Dependable Computing*, volume 7869 of *Lecture Notes in Computer Science*, pages 134–147. Springer Berlin Heidelberg, 2013.

[95] L. Geppert. Lost radio contact leaves pilots on their own. *Spectrum, IEEE*, 41(11):16–17, 2004.

[96] S. Goff. Banks told to fix systems after RBS crash. Financial Times, July 2012.

[97] R. W. Goldsmith. *Financial structure and development.* Yale University Press - New Haven, 1969.

[98] M. Govindarasu, A. Hahn, and P. Sauer. Cyber-Physical Systems Security for Smart Grid. PSERC Future Grid Initiatives Webinar Series, February 2012.

[99] M. Hadley, J. McBride, T. Edgar, L. O'Neil, and J. Johnson. Securing Wide Area Measurement Systems. Available at: http://energy.gov/oe/downloads/securing-wide-area-measurement-systems. U.S. Department of Energy Office of Electricity Delivery and Energy Reliability, 2007.

[100] J. Hall and G. Rayner. RBS computer failure condemns man to spend weekend in the cells. Telegraph, June 2012.

[101] M. Hentea. Improving security for scada control systems. *Interdisciplinary Journal of Information, Knowledge, and Management*, 3(1):73–86, 2008.

[102] J. Hong, S.-S. Wu, A. Stefanov, A. Fshosha, C.-C. Liu, P. Gladyshev, and M. Govindarasu. An intrusion and defense testbed in a cyber-power system environment. In *IEEE Power and Energy Society General Meeting*, pages 1–5, 2011.

[103] P. Institute. 2011 cost of data breach study. `http://www.ponemon.org`.

[104] C. R. Johnson, M. Montanari, and R. H. Campbell. Automatic management of logging infrastructure. In *National Centers of Academic Excellence - Workshop on Insider Threat*, St Louis, MO, USA, 2010.

[105] N. Johnson, G. Zhao, E. Hunsader, H. Qi, N. Johnson, J. Meng, and B. Tivnan. Abrupt rise of new machine ecology beyond human response time. *Sci. Rep.*, 3, 09 2013.

[106] Kaspersky Lab. The geography of cybercrime: Western europe and north america. Technical report, 2012.

[107] L. Kenney, J. Dietrich, and J. Woodall. Secure ATC surveillance for military applications. In *IEEE Military Communications Conference, 2008. MILCOM*, pages 1–6, 2008.

[108] S. Larsson and E. Ek. The black-out in southern Sweden and eastern Denmark, September 23, 2003. In *IEEE Power Engineering Society General Meeting*, pages 1668–1672 Vol.2, 2004.

[109] Y. W. Law, M. Palaniswami, G. Kounga, and A. Lo. WAKE: Key management scheme for wide-area measurement systems in smart grid. *IEEE Communications Magazine*, 51(1):34–41, 2013.

[110] E. E. Lee, J. E. Mitchell, and W. A. Wallace. Assessing vulnerability of proposed designs for interdependent infrastructure systems. In *7th Hawaii International Conference on System Sciences*, 2004.

[111] E. LeMay, M. Ford, K. Keefe, W. Sanders, and C. Muehrcke. Model-based Security Metrics Using ADversary VIew Security Evaluation (ADVISE). In *8th International Conference on Quantitative Evaluation of Systems (QEST)*, pages 191–200, September 2011.

[112] R. Leon, V. Vittal, and G. Manimaran. Application of sensor network for secure electric energy infrastructure. *IEEE Transactions on Power Delivery*, 22(2):1021–1028, 2007.

[113] T. Lewis. Critical infrastructure protection in homeland security: Defending a networked nation. John Wiley, 2006.

[114] J. Liu, Y. Xiao, S. Li, W. Liang, and C. L. P. Chen. Cyber security and privacy issues in smart grids. *Communications Surveys Tutorials, IEEE*, 14(4):981–997, 2012.

[115] M. Mackenzie and A. Massoudi. NYSE cancels trades after algo glitch. financial times. Available Online: `http://www.ft.com/intl/cms/s/0/bd5f2af8-dbe7-11e1-8d78-00144feab49a.html?siteedition=intl#axzz2f3Fy6h85`, August 2010.

[116] V. Madani and R. King. Strategies and roadmaps to meet grid challenges for safety and reliability. In G. Anders and A. Vaccaro, editors, *Innovations in Power Systems Reliability*, Springer Series in Reliability Engineering, pages 1–11. Springer London, 2011.

[117] B. Malinowsky, G. Gronbaek, H. Schwefel, A. Ceccarelli, A. Bondavalli, and E. Nett. Timed Broadcast via Off-The-Shelf WLAN Distributed Coordination Function for Safety-Critical Systems. In *European Dependable Computing Conference (EDCC)*, pages 144–155, 2012.

[118] S. Marti, T. J. Giuli, K. Lai, and M. Baker. Mitigating routing misbehavior in mobile ad hoc networks. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, MobiCom '00, pages 255–265, New York, NY, USA, 2000. ACM.

[119] M. Masera. An approach to the understanding of interdependencies. In *Power Systems and Communications Infrastructures for the Future (CRIS)*. 2002.

[120] B. Masters, E. Moore, and J. Pickard. The upgrade that downed Royal Bank of Scotland. Financial Times. Available online:, June 2012.

[121] D. McCallie, J. Butts, and R. Mills. Security analysis of the ads-b implementation in the next generation air transportation system. *International Journal of Critical Infrastructure Protection*, 4(2):pp. 78–87, 2011.

[122] P. McDaniel and S. McLaughlin. Security and Privacy Challenges in the Smart Grid. *IEEE Security and Privacy,*, 7(3):pp. 75–77, May 2009.

[123] G. McDonald, L. O. Murchu, S. Doherty, and E. Chien. Stuxnet 0.5:The Missing Link - Available online `http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/stuxnet_0_5_the_missing_link.pdf`.

[124] J. McDonald. *Power Substations Engineering*. CRC Press, Boca Raton, FL, 2007.

[125] S. McLaughlin, D. Podkuiko, and P. McDaniel. Energy theft in the advanced metering infrastructure. In *Critical Information Infrastructures Security*, pages pp 176–187, 2010.

[126] P. Mell, K. Scarfone, and S. Romanosky. A complete guide to the common vulnerability scoring system version 2.0. Technical report, FIRST, Available at `http://www.first.org/cvss`, 2007.

[127] J. W. Meritt. A method for quantitative risk analysis. Available online:`http://csrc.nist.gov/nissc/1999/proceeding/papers/p28.pdf`, 2008.

[128] J. Momoh. *Smart Grid: Fundamentals of Design and Analysis*. I E E Power Engineering Series. Wiley, 2012.

[129] B. Morin, L. Mé, H. Debar, and M. Ducassé. M2d2: A formal data model for ids alert correlation. In *RAID*, pages 115–127, 2002.

[130] B. T. Morris and M. M. Trivedi. Trajectory learning for activity understanding: Unsupervised, multilevel, and long-term adaptive approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(11):2287–2301, 2011.

[131] T. Morris, S. Pan, and U. Adhikari. Cyber security recommendations for wide area monitoring, protection, and control systems. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–6, 2012.

[132] J. Moteff. Risk management and critical infrastructure protection: Assessing, integrating, and managing threats, vulnerabilities and consequences. CRS Report for Congress, The Library of Congress, 2004.

[133] J. M. Moya, A. Araujo, Z. Bankovic, J.-M. D. Goyeneche, J. C. Vallejo, P. Malagon, D. Villanueva, D. Fraga, E. Romero, and J. Blesa. Improving Security for SCADA Sensor Networks with Reputation Systems and Self-Organizing Maps. In *Sensors. 9*, pages pp. 9380–9397, 2009.

[134] K. Munro. SCADA - A critical situation. *Network Security*, (1):4 – 6, 2008.

[135] NDTV. A phone application that threatens security. Press Trust of India, New Delhi, India, October 4 2010.

[136] D. M. Nicol, W. H. Saunders, and K. S. Trivedi. Model-based evaluation: From dependability to security. *IEEE Transactions on Dependable and Secure Computing*, 1(1):48–65, 2004.

[137] S. Noel, E. Robertson, and S. Jajodia. Correlating intrusion events and building attack scenarios through attack graph distances. In *ACSAC*, pages 350–359, 2004.

[138] North American Electric Reliability Corporation. Critical Infrastructure Protection Standards 002-3 - 009-3, 2009.

[139] U. D. of Energy. Smart grid system report. Technical report, U.S. Department of Energy, Feb 2012. `http://energy.gov/sites/prod/files/2010%20Smart%20Grid%20System%20Report.pdf`.

[140] OMG. Data Distribution Service specifications. `http://www.omg.org/spec/DDS/`, 2007.

[141] F. D. Oral. The Impacts of Natural Disasters on Power Systems: Anatomy of the Marmara Earthquake Blackout. *Acta Polytechnica Hungarica*, 7(2):107–118, 2010.

[142] M. A. Palfi. Efficient maritime port critical infrastructure protection: A project management and critical thinking perspective. *Synergy*, 2:237–253, 2008.

[143] A. Pecchia, R. Pietrantuono, and S. Russo. Criticality-driven component integration in complex software systems. In *Computer Safety, Reliability, and Security*, pages 452–466. Springer, 2011.

[144] D. Peck and D. Peterson. Leveraging ethernet card vulnerabilities in field devices. In *Proceedings of SCADA Security Scientific Symposium, Miami, USA*, 2009.

[145] P. Pederson, D. Dudenhoeffer, S. Hartley, and M. Permann. Critical infrastructure interdependency modeling: A survey of U.S. and international research. Technical Report INL/EXT-06-11464, `http://www.pcsforum.org/library/files/1159904563-TSWG_INL_CIP_Tool_Survey_final.pdf`, August 2006.

[146] P. Pourbeik, P. S. Kundur, and C. W. Taylor. The anatomy of a power grid blackout. *IEEE Power and Energy Magazine*, pages 22–29, September-October 2006.

[147] J. Quirke. Security in the GSM system. *AusMobile, May*, pages 1–26, 2004.

[148] G. D. Rash. GPS Jamming in A Laboratory Environment. In *Proceedings of the 53rd Annual Meeting of The Institute of Navigation*, pages pp. 389–398., 1997.

[149] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly. Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies. *IEEE Control Systems Magazine*, pages 11–25, December 2001.

[150] B. Robert, R. D. Calan, and L. Morabito. Modelling interdependencies among critical infrastructures. *IJCIS*, 4(4):392–408, 2008.

[151] L. Romano, S. D'Antonio, V. Formicola, and L. Coppolino. Protecting the wsn zones of a critical infrastructure via enhanced siem technology. In F. Ortmeier and P. Daniel, editors, *Computer Safety, Reliability, and Security*, volume 7613 of *Lecture Notes in Computer Science*, pages 222–234. Springer Berlin Heidelberg, 2012.

[152] J. J. Romero. Blackouts illuminate India's power problems. *IEEE Spectrum*, 49(10):11–12, 2012.

[153] V. Rosato, L. Issacharoff, F. Tiriticco, S. Meloni, S. D. Porcellinis, and R. Setola. Modelling interdependent infrastructures using interacting dynamical models. *International Journal of Critical Infrastructures*, 4(1/2):63, 2008.

[154] J. Salmeron, K. Wood, and R. Baldick. Analysis of electric grid security under terrorist threat. *IEEE Transactions on Power Systems*, 19(2):905–912, 2004.

[155] W. H. Sanders and J. F. Meyer. Stochastic activity networks: formal definitions and concepts. In *Lectures on formal methods and performance analysis*, pages 315–343. Springer-Verlag New York, Inc., New York, NY, USA, 2002.

[156] R. Shapiro, S. Bratus, E. Rogers, and S. Smith. Identifying vulnerabilities in SCADA systems via Fuzz-Testing. In *Critical Infrastructure Protection V*, volume 367 of *IFIP Advances in Information and Communication Technology*, pages 57–72. Springer Berlin Heidelberg, 2011.

[157] D. P. Shepard, T. E. Humphreys, and A. A. Fansler. Evaluation of the vulnerability of phasor measurement units to GPS spoofing attacks. *International Journal of Critical Infrastructure Protection*, 2012.

[158] A. Singh and S. Aasma. A Grid failure in Northern, Eastern and North-Eastern grid in 2012: Cause & its effect on economy of India And Review. *SAMRIDDHI-A Journal of Physical Sciences, Engineering and Technology (S-JPSET)*, 3(2), 2012.

[159] I. Sommerville. *Software engineering*. Addison-Wesley, 2001.

[160] J. Tate and T. Overbye. Line outage detection using phasor angle measurements. *IEEE Transactions on Power Systems*, 23(4):1644–1652, 2008.

[161] J. Tate and T. Overbye. Double line outage detection using phasor angle measurements. In *IEEE Power and Energy Society General Meeting 09*, pages 1–5, 2009.

[162] C.-W. Ten, G. Manimaran, and C.-C. Liu. Cybersecurity for critical infrastructures: Attack and defense modeling. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,*, 40(4):853–865, 2010.

[163] H. Teso. Aircraft hacking: Practical aero series. In *4th annual Hack in the Box Security Conference*, Amsterdam, The Netherlands, April 10–11 2013.

[164] The Power Grid. TCIPG: Trustworthy Cyber Infrastructure for the Power Grid website: `http://tcipg.org`.

[165] W. C. Thompson. Railroad infrastructure security, trb annual meeting, january 14 2002.

[166] UIC. *Subset 033 - rev. 2.0.0 - ERTMS-ETCS Class 1 - FIS for the Man-Machine Interface*, 2000.

[167] UIC. *Subset-026 - rev. 2.2.2 - ERTMS-ETCS Class 1 - System requirements specification*, 2002.

[168] V. Urias, B. Van Leeuwen, and B. Richardson. Supervisory Command and Data Acquisition (SCADA) system cyber security analysis using a live, virtual, and constructive (LVC) testbed. In *Military Communications Conference (MILCOM)*, pages 1–8, 2012.

[169] U.S.-Canada Power System Outage Task Force. *Final report on the August 14, 2003 blackout in the United States and Canada: causes and recommendations.* U.S. Dept. of Energy, Washington, D.C, 2004.

[170] U.S. Commodity Futures Trading Commission and U.S. Securities & Exchange Commission. Findings regarding the market events of may 6, 2010 - Available online: `http://www.sec.gov/news/studies/2010/marketevents-report.pdf`, September 2010.

[171] U.S. Department of Transportation. The Public Transportation Security & Emergency Preparedness Planning Guide. Federal Transit Administration, Final Report, 2003.

[172] G. Vigna. A Topological Characterization of TCP/IP Security. In *FME*, pages 914–939, 2003.

[173] G. Vigna and R. A. Kemmerer. NetSTAT: A Network-Based Intrusion Detection Approach. In *ACSAC*, pages 25–34, 1998.

[174] J. Wang, Z. Cheng, M. Zhang, Y. Zhou, and L. Jing. Design of a Situation-Aware System for Abnormal Activity Detection of Elderly People. In *AMT*, pages 561–571, 2012.

[175] L. Wang, A. Liu, and S. Jajodia. An efficient and unified approach to correlating, hypothesizing, and predicting intrusion alerts. In *ESORICS*, pages 247–266, 2005.

[176] L. Wang, S. Noel, and S. Jajodia. Minimum-cost network hardening using attack graphs. *Computer Communications*, 29(18):3812–3824, 2006.

[177] J. S. Warner and R. G. Johnston. GPS Spoofing Countermeasures. *Journal of Homeland Security*, 2003.

[178] R. Watts. Maritime Critical Infrastructure Protection: Multi-Agency Command and Control in an Asymmetric Environment. Homeland Security Affairs 1, issue 2, 2005.

[179] W. E. Wong, V. Debroy, A. Surampudi, H. Kim, and M. F. Siok. Recent Catastrophic Accidents: Investigating How Software Was Responsible. In *Fourth International Conference on Secure Software Integration and Reliability Improvement (SSIRI)*, pages 14–22, 2010.

[180] A. Wood. After ADS-B launch, security concerns raised. Aviation International News, July 2006.

[181] D. Wu and C. Zhou. Fault-tolerant and scalable key management for smart grid. *IEEE Transactions on Smart Grid*, 2(2):375–381, 2011.

[182] J. Xia and Y. Wang. Secure key distribution for the smart grid. *Smart Grid, IEEE Transactions on*, 3(3):1437–1443, 2012.