CSS Analyses in Security Policy

CSS
ETH Zurich

Vol. 2 • No. 16 • June 2007

CRITICAL INFRASTRUCTURES: VULNERABILITIES, THREATS, RESPONSES

The terrorist attacks of 11 September 2001 have led to an increased focus on the vulnerability of modern societies in general and the protection of so-called critical infrastructures in particular. However, drafting efficient protection plans has proven to be a challenge: requirements include sophisticated situation analyses, better understanding of vulnerabilities, and a political consensus on how protection measures should be prioritized. Domestic and inter-state political cooperation as well as functioning public-private partnerships are also indispensable.

Information infrastructures – nerve centers of modern society *www.istockphoto.com*

The importance of protecting infrastructures has greatly increased in the security political debate of late, due in particular to the traumatic terrorist attacks in New York and Washington (2001), Madrid (2004), and London (2005). In all of these cases, the perpetrators exploited elements of the civilian infrastructure for the purpose of indiscriminate murder. In the case of the 11 September 2001 attacks in the US, they used the transport infrastructure by turning airplanes into weapons. In Europe, trains, underground railways, and train stations as well as commuters were targeted. This approach not only demonstrated the brutal nature of the "new terrorism", but also reinforced the view that traditional concepts of domestic security were no longer commensurate to contemporary requirements and needed to be adapted.

Long before these attacks, the protection of strategically important installations in the domestic economic and social sphere had already been an important part of national defense concepts under the label of "physical protection". The term Critical Infrastructure Protection (CIP), however, refers to a broader concept with a distinctly new flavor. First of all, it is no longer restricted to concrete defense against immediate dangers or criminal prosecution after a crime has been committed, but increasingly refers to preventive security measures as well. Furthermore, contemporary modern societies have become significantly more vulnerable, and the spectrum of possible causes of disruptions and crises has become broader and more diffuse. This is why CIP has become a crystallization point for current security policy debates.

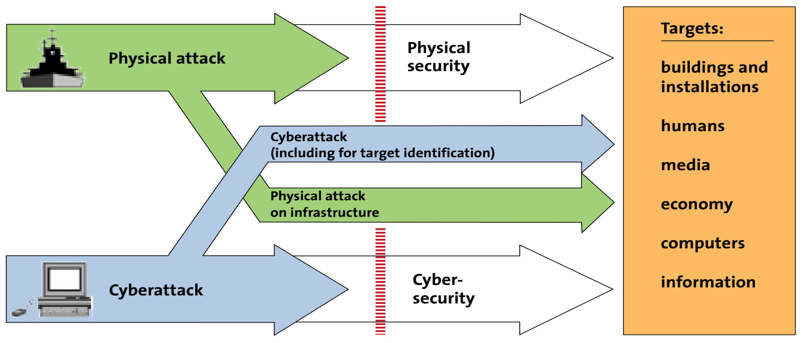
From threats to risks

The genesis and establishment of the concept of CIP is the result of two interlinked and at times mutually reinforcing factors: The expansion of the threat spectrum after the Cold War, especially in terms of malicious actors and their capabilities on the one hand, and a new kind of vulnerability due to modern society's dependency on inherently insecure information systems on the other.

During the Cold War, threats were mainly perceived as arising from the aggressive intentions of states to achieve domination over other states. Among other things, the end of the Cold War also heralded the end of unambiguous threat perceptions: following the disintegration of the Soviet Union, a variety of "new" threats were moved onto the security policy agendas of most countries. The main distinguishing quality of these "new" challenges is the element of uncertainty that surrounds them: uncertainty concerning the identity and goals of potential adversaries, the timeframe within which threats are likely to arise, the contingencies that might be imposed on the state by others, the capabilities against which one must prepare, and also about what type of challenge to prepare for. Clearly, the notion of "threat" as something imminent, direct, and certain no longer accurately describes these challenges. Rather, they can be characterized as "risks", which are by definition indirect, unintended, uncertain, and situated in the future, since they only materialize when they occur in reality.

As a result of these diffuse risks and due to difficulties in locating and identifying enemies, parts of the focus of security policies has shifted away from actors, capabilities, and motivations towards

Interdependencies of risks and vulnerabilities


Physical attack
Physical security
Cyberattack (including for target identification)
Physical attack on infrastructure
Cyberattack
Cyber-security
Targets:
buildings and installations
humans
media
economy
computers
information

Source: PCCIP report 1997, p. 6

general vulnerabilities of entire societies. The catchphrase in this debate is "asymmetry", and the US military has been a driving force behind the shaping of this threat perception in the early 1990s. The US as the only remaining superpower was seen as being predestined to become the target of asymmetric warfare. Specifically, those adversaries who were likely to fail against the American war machine might instead plan to bring the US to its knees by striking against vital points at home that are fundamental not to the military alone, but to the essential functioning of industrialized societies as a whole. These points are generally defined as critical infrastructures (CI). They are deemed critical because their incapacitation or destruction would have a debilitating impact on the national security and the economic and social welfare of a nation.

Fear of asymmetrical measures against such "soft targets" was aggravated by the second factor: the so-called information revolution. Most of the CI relies on a spectrum of software-based control systems for smooth, reliable, and continuous operation. In many cases, information and communication technologies (ICT) have become all-embracing, connecting other infrastructure systems and making them interrelated and interdependent. These technologies are in general regarded as inherently insecure: security has never been a system design driver, and pressure to reduce time-to-market is intense, so that a further explosion of computer and network vulnerabilities is to be expected, leading to the emergence of infrastructures with in-built instability, critical points of failure, and extensive interdependencies. At the same time, the spread of ICT was (and is) seen to make it much easier to attack the US asymmetrically, as big, specialized weapons systems or an army are no longer required. Borders, already porous in many ways in the real world, are nonexistent in cyberspace.

From hackers to terrorists

The US was the first nation to broadly address the new vulnerability of the vital infrastructures in a concerted effort. New risks in designated sectors like information and communications, banking and finance, energy, physical distribution, and vital human services were identified by the Presidential Commission on Critical Infrastructure Protection (PCCIP). The PCCIP concluded in 1997 that the US was so dependent on these infrastructures that the government had to view them through the lens of a "national security focus", since serious consequences for the entire nation were to be expected if these elements were unavailable for any significant amount of time.

According to this approach, critical infrastructures should be understood to include material and IT assets, networks, services, and installations that, if disrupted or destroyed, would have a serious impact on the health, security, or economic well-being of citizens and the efficient functioning of a country's government. Such infrastructures could be damaged by structural threats as well as by intentional, actor-based attacks. The first risk category would, for example, include natural catastrophes, human-induced catastrophes (e.g. dam failure, nuclear reactor accident), personnel shortages through strikes or epidemics, organizational shortcomings due to technical or personal failures, human error, technical outages, and dependencies and supply shortages. In the second category, the spectrum of possible attackers is extensive, ranging from bored teenagers, disaffected or dissatisfied employees, organized crime, fanatics and terrorist cells, to hostile states.

There is an equally broad range of attack options, including hacker attacks as well as the physical destruction of civilian or military installations. The main focus of early US CIP efforts was, however, directed towards the as-yet largely unknown risks emanating from cyberspace: The global information infrastructure appeared to facilitate anonymous attacks from anywhere in the world, while at the same time serving as a source for hacker tools for everyone. Based on this threat perception, a CIP policy crystallized under US President Bill Clinton that was largely directed towards information security.

However, since the terrorist attacks of 11 September 2001, there has been a noticeable return of the classical threat concept to the CIP debate. Especially from the US point of view, efforts have been made since then to tackle a series of structural threats within the framework of an increasingly actor-oriented counter-terrorism strategy. In the US, CIP became a key component of Homeland Security and is currently discussed predominantly with a view to developing strategies against Muslim terrorism. The physical aspects of CIP have been moved to the forefront, while the importance of information aspects has diminished. In the meantime, this CIP focus on counterterrorism has also become a hallmark of debates in the EU, which has recently begun to develop a CIP policy that consists mainly of coordinating the measures adopted by member states.

Challenges to an efficient CIP policy

The example of the US and experiences gathered by other countries allow us to identify four challenges to efficient CIP policy, some of which are closely interlinked. First of all, a sound assessment of the nature and scale of the relevant risks and threats is required. Instead of the current one-sided focus on terrorism, CIP should return to a broader approach and deal with the susceptibility of highly complex systems in general. The intelligence services have an important role to play in connection with providing nuanced assessments. This is all the more important considering that responsibilities may be located in various places depending on the danger involved, and that protective measures must be shaped accordingly on a case-by-case basis.

However, as long as there is no reliable data on the likelihood of threats, a focus on the likely effects of a failure of a specific infrastructure or asset and ways to mitigate them is a better approach. The reasoning for this is quite simple: from the perspective of maintaining reliable services, it is not so important whether the events that triggered the surprise originated from within or from outside the infrastructure. In practice, it is also often difficult to determine whether a particular detrimental event is the result of a malicious attack, of a component failure, or of an accident – a distinction that is often less important than the impact of the event. This demonstrates the value of an "all-hazards" approach, designed for comprehensive protection, irrespective of the nature of the threat, with a focus on the capability to respond to a whole spectrum of unanticipated events. The key is to create greater resilience, commonly defined as the ability of a system to recover from adversity, either returning back to its original state or in an adjusted state based on new requirements. Resilience is commonly embedded in processes, rather than individual physical assets or protection measures.

Secondly, a better understanding of vulnerabilities is required, including interdependencies between infrastructures. It is clear that comprehensive protection of all critical infrastructures – once they have been identified – against all threats and risks is impossible, not only for technical and practical reasons, but also because of the associated costs. Further prioritization, may, for example, involve a distinction between critical infrastructures that deserve a greater level of attention, or the identification of vital points within a critical infrastructure. Criteria used for prioritization can focus on the relative likelihood of the threat, on the criticality of an asset compared to another one, or on the relative cost of protection. It is clear that due to the high degree of system complexity, the existing methodology is not sufficient to grasp the entire range of the problem. From a strategic point of view, as opposed to the dominant "technical approach", the goal is often not so much to quantify and measure risks "objectively", but to understand them in their social, political-institutional, cultural, or economic context.

Third, there is a need to define what makes an infrastructure "critical". After 9/11, the list of critical infrastructures was vastly increased. "Critical" elements now also include such infrastructures the destruction of which would have an effect on the "national psyche" and the morale of the nation. This development poses near-insurmountable problems for the development of protective measures: How can one guarantee security when nearly everything is considered "critical" and therefore requires protection? Benchmarks for distinguishing between "normal" and "critical" should not be set too low. It is of key importance to establish sensible priorities. This, in turn, can only be done based on comprehensive risk analysis. The hypothetical vulnerability of a target is not a sufficient indicator for selecting targets to be protected. Instead, a sensible assessment of criticality must also be based on knowledge about concrete threats and about the scope and gravity of potential damage.

Fourth, CIP requires comprehensive cooperation. A functioning partnership between the state and the corporate sector is essential. Due to the liberalization of many public sectors since the 1980s, a large part of the critical infrastructure is privately administered today. Therefore, the private sector has a key role in defining and implementing protective policies, and nation states want operators to take on responsibility for the implementation of protection measures that are in accordance with the parameters or frameworks set by public authorities. In order to win the support of the private sector without having to fall back on heavy regulation, governments must strive to create a mutual win-win situation.

Fortunately states can provide a number of services that are in the interest of the private sector. For example, nation states can provide financial assistance, through funding of research on protection technologies and by contributing towards implementation costs. They can coordinate the intervention of law enforcement services regarding criminal matters and of emergency services for disaster relief, and can provide advice, guidance, or oversight concerning measures taken by other infrastructure operators to protect their facilities. Governments can provide non-technical analyses of the general risk situation provided by national and international intelligence services, such as information about the nature of criminal organizations. Further, private actors can profit from close contacts with the police (in particular, with high-tech crime units). Also, they can gain knowledge about incidents and lessons learned from an exchange with other private actors that is mediated by a "neutral" government entity.

However, the efficacy of national efforts remains limited: the vulnerability of modern societies has global origins and implications. Any adequate protection policy that extends to strategically important infrastructures will thus ultimately require transnational solutions. To create some kind of added value, international organizations can help develop and promulgate (information) security standards, or disseminate recommendations and guidelines on best practices. International law enforcement institutions and mechanisms, like Interpol, can be used for information exchange and investigations, with the aim of providing early warning of cyber-attacks by exchanging information between the public and private sectors. Enhanced cooperative policing mechanisms can be created. Multilateral conventions on computer crime, such as the Council of Europe convention, can be expanded and built on. However, it is of key importance not to duplicate efforts already undertaken at national level or below: the principles of subsidiarity and proportionality must be taken into account at all times. Key activities should concentrate on challenges that cannot be mastered by any single nation or region on its own; among these are global infrastructures, like the internet, or truly large-scale interdependencies.

Read more on this issue

Dunn, Myriam, "Critical information infrastructures: vulnerabilities, threats and responses", in: *Disarmament Forum*, 2007, no. 3 (forthcoming).

Author:
Myriam Dunn
dunn@sipo.gess.ethz.ch

Responsible editor:
Daniel Möckli
analysen@sipo.gess.ethz.ch

Translated from German:
Christopher Findlay

Other CSS Analyses / Mailinglist:
www.isn.ethz.ch

German and French versions:
www.ssn.ethz.ch