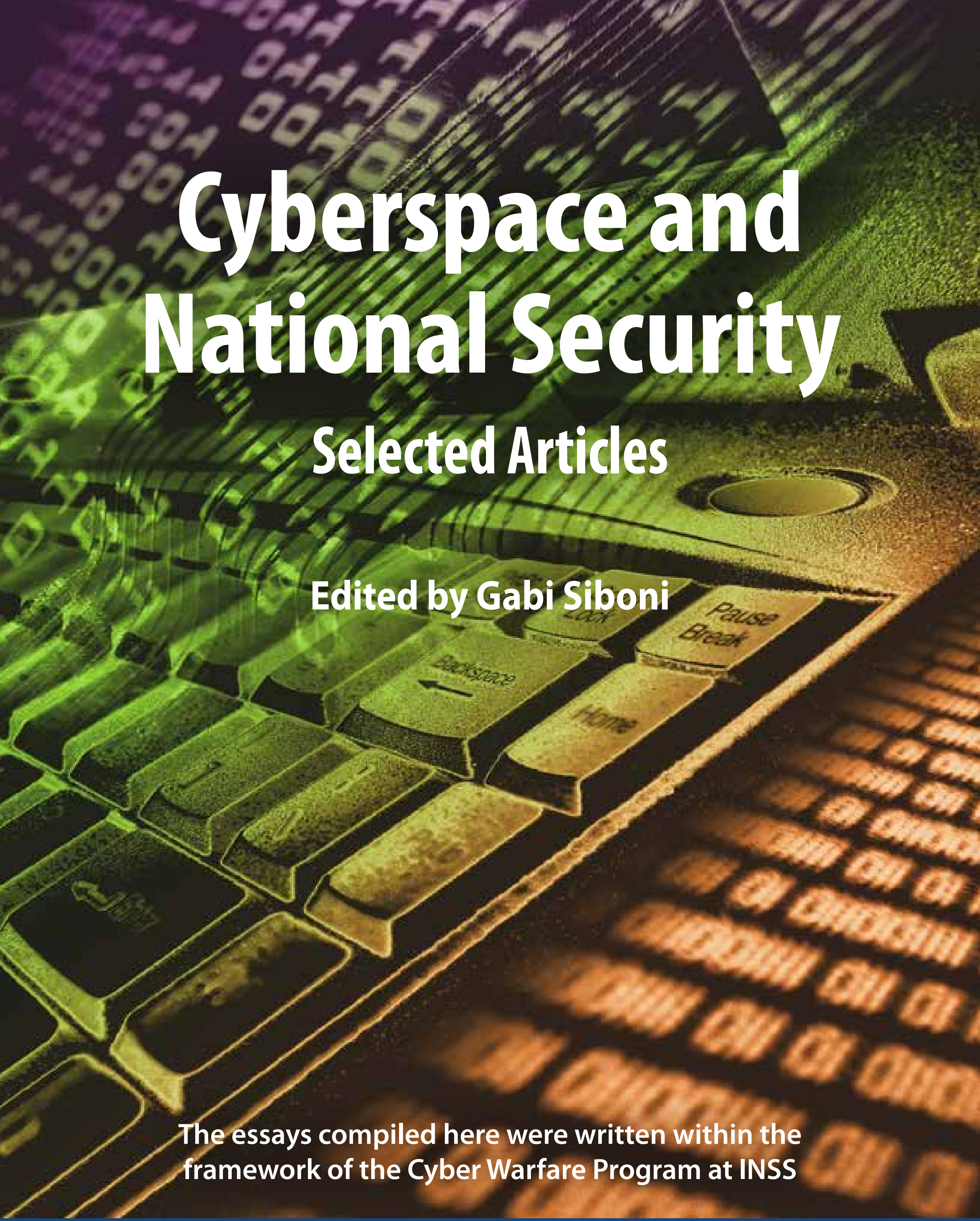


Cyberspace and National Security

Selected Articles

Edited by Gabi Siboni

The essays compiled here were written within the framework of the Cyber Warfare Program at INSS

INSS
המכון למחקרי ביטחון לאומי
THE INSTITUTE FOR NATIONAL SECURITY STUDIES
INCORPORATING THE JAFFEE
CENTER FOR STRATEGIC STUDIES
TEL AVIV UNIVERSITY
אוניברסיטת תל-אביב

Cyberspace and National Security

Selected Articles

Edited by Gabi Siboni

The essays compiled here were written within the framework of the Cyber Warfare Program at INSS

INSS
המכון למחקרי ביטחון לאומי
THE INSTITUTE FOR NATIONAL SECURITY STUDIES
INCORPORATING THE JAFFEE
CENTER FOR STRATEGIC STUDIES
TEL AVIV UNIVERSITY
אוניברסיטת תל-אביב

Graphic design: Michal Semo-Kovetz
Printing: Elinir

Institute for National Security Studies
(a public benefit company)
40 Haim Levanon Street
POB 39950
Ramat Aviv
Tel Aviv 6997556

Tel. +972-3-640-0400
Fax. +972-3-744-7590

E-mail: info@inss.org.il
http://www.inss.org.il

June 2013

ISBN: 978-965-7425-51-0

INSS Institute for National Security Studies

The Institute for National Security Studies (INSS), incorporating the Jaffee Center for Strategic Studies, was founded in 2006.

The purpose of the Institute for National Security Studies is first, to conduct basic research that meets the highest academic standards on matters related to Israel's national security as well as Middle East regional and international security affairs. Second, the Institute aims to contribute to the public debate and governmental deliberation of issues that are – or should be – at the top of Israel's national security agenda.

INSS seeks to address Israeli decision makers and policymakers, the defense establishment, public opinion makers, the academic community in Israel and abroad, and the general public.

INSS publishes research that it deems worthy of public attention, while it maintains a strict policy of non-partisanship. The opinions expressed in this publication are the author's alone, and do not necessarily reflect the views of the Institute, its trustees, boards, research staff, or the organization and individuals that support its research.

Contents

Foreword

Since most systems in developed societies depend on computer and information infrastructures, nations are increasingly vulnerable to criminal as well as hostile elements in the realms of computer and communications capabilities and the global proliferation of information systems. Indeed, the growing dependence on information technology and communications results in a situation in which damage to computers and the flow of information may lead to tangible, physical damage. It is possible to disrupt management and command and control systems through changes in computer software, rendering physical attacks unnecessary. Israel's well developed technological capabilities in computers and communications give it a tremendous edge in all fields, especially security, and allow it to act in cyberspace both to foil attacks and gain advantages on the modern battlefield. At the same time, the growing dependence on computers is also a potential Achilles' heel requiring solutions.

This volume, prepared in advance of the Institute for National Security Studies 2013 conference on cyberspace, compiles eight essays published previously in *Military and Strategic Affairs*. Written primarily by INSS researchers, the essays present some of the research produced in the framework of the INSS Cyber Warfare Program, which is supported by the Philadelphia-based Joseph and Jeanette Neubauer Foundation. This research program deals with a range of aspects, such as a framework for basic concepts in cyber warfare and an analysis of cyber warfare capabilities as part of specifically designated research on countries of special interest to Israel and the world at large, e.g., China and Iran. Other topics in the research program include the influence of cybercrime on national security, the proliferation of cyber weapons, the failure of existing cyber defense methods and investigation of new directions for comprehensive cyber defense, and a proposal to promote cyber defense regulation in the civilian sector.

This past year INSS launched a Cyber Policy and Strategy Forum to address the current gap in the discourse between two expanses – the active technological world, which has witnessed exponential growth in the knowledge reservoir both in Israel and abroad, and the world of strategy and policy formulation. This forum allows for direct discourse between technology companies and strategists and policymakers, thereby generating new insights to enhance cyber defense in Israel and to promote the relevant research both in Israel and around the world.

Gabi Siboni
Head of the Cyber Warfare Program at INSS, June 2013

Protecting Critical Assets and Infrastructures from Cyber Attacks

Gabi Siboni

The impact of computer and communications systems in recent decades has not bypassed the national security of states in general, and the State of Israel in particular. Most systems in developed societies rely on computer and information infrastructures, and this growing dependence on information and communication technologies means that a blow to computers and information flow processes is liable to disrupt, paralyze, and sometimes even cause substantive physical damage to essential systems. Computer-based capabilities and their near-global ubiquity expose states to harm in cyberspace by various elements, including hostile countries, terrorist organizations, criminal elements, and even individuals driven by personal challenges or anarchist motives. The threat is particularly acute as management, control, and monitoring systems can be disrupted through changes to a computer program, and no physical attack is needed. Thus, it stands to reason that the face of future conflicts will be transformed beyond recognition.

The strength of a sovereign state is a function of economic, societal, and scientific strength combined with military strength, and the purpose of the military strength is to protect the state's territory and its citizens so that they can cultivate and maintain economic strength. The vulnerability of computers and communications systems to cyber attacks entails a dramatic change in the concept of military strength. For the first time, it is possible to mortally wound national economic strength by paralyzing economic and civilian systems without using firepower and force maneuvers. Thus, the ability of states to operate in cyberspace for both defensive and offensive purposes coincides with classic military capabilities to play a significant role.

Dr. Col. (ret.) Gabi Siboni is head of the Military and Strategic Affairs Program and head of the Cyber Warfare Program at INSS.

This article was first published in *Military and Strategic Affairs* 3, no. 1 (2011): 93-101.

In the past two decades, states, along with their progress, profitability, and wellbeing – and their production and provision of national services in particular – have been exposed to new threats, yet insufficient attention has been paid to the appropriate means of confronting such threats. In the recent past, industry (private and public) was protected by the state. For example, excluding workplace accidents, power stations producing electricity, whether in private hands or publicly owned, were exposed to physical damage only if the state encountered a physical war, and it was the state's job to protect such infrastructures along with economic institutions, industrial facilities, and so forth. Public institutions were protected by the state by virtue of their existence in the territorial space under its authority and control. That has changed. In addition, the trend in recent decades to privatization has placed a large portion of the infrastructure plants that were traditionally in the hands of the government in private hands, including those relating to communications, transportation, electricity, energy, and heavy industry. Moreover, traditional industries have in recent decades been joined by new industries in the hi-tech realm that constitute a significant component of states' GDP.

Due to the universal understanding that "he who defends everything defends nothing,"[1] various countries have developed ways of protecting infrastructures and systems that are critical to their functioning. In 2002, the State of Israel established the Information Security Authority, "in charge of professional direction of the bodies for which it is responsible regarding securing essential computer infrastructures from the threats of terrorism and sabotage to the security of classified information, and from the threats of espionage and exposure."[2] In this context, a steering committee was established in the National Security Council whose role is to examine the risks in information security. It was also decided that the rules of the steering committee would apply to a number of bodies and institutions whose information systems are defined as critical, including the electric company, banks, government offices, and the like, and the committee is authorized to add to this list.[3]

The public service bodies that are required to protect themselves from a cyber attack have been under the direction of the Information Security Authority for quite a while. At the same time, changes in the structure of the Israeli economy and the emergence of elements, processes, assets, and projects – which if damaged could potentially cause significant harm on a national level – have exposed and increased the range of weak points and the targets for cyber attacks. Moreover, potential damage is not restricted to what can be quantified in financial terms or what impacts on the GDP: significant damage can also be caused to assets and values that have Israeli and Jewish national importance. Thus, for example, in the United States, defensive plans also apply to heritage and memorial sites.[4]

Consequently, it is highly important to be able to examine which additional entities require guidance by the Information Security Authority. This article proposes an approach that will make it possible to implement a systematic process using existing statutory tools, in order to identify other bodies (mainly from the private sector) whose damage might impact on national security, and therefore requires them to operate appropriate defensive mechanisms for their critical assets and infrastructures.

What Should be Protected?

In a US Department of Homeland Security document,[5] Patrick Beggs[6] reviews how authorized officials in the United States see the interface between defense-critical infrastructures and resources and their physical and cyber infrastructures.

In the United States, the mapping of defense-critical infrastructures covers water, energy, communications, transportation, the chemical industry, agriculture and the food industry, information systems, banking, commercial and financial services, health services, and finally, areas of importance to the American collective memory (national monuments, heritage sites, and so on). These sectors are grounded on two basic infrastructure components: the first regards physical infrastructure components, such as power stations, dams, airports and sea ports, roads, railroads tracks, various types of delivery infrastructures,[7] hospitals, factories, and the like. The second component concerns cyber infrastructures, including software and hardware systems, internet servers, command and control systems, and information systems.

In order to enable an appropriate basis for formulating defense plans, the US uses a methodology called Cyber Resiliency Review (CRR) of institutions and critical infrastructures that belong to the sectors described above. This approach makes it possible to assess a number of aspects, including the definition of defense-critical assets, management of communications, continuity of services, technological management, dependence on external components, management of unforeseen incidents and accidents, ability to assess the situation, and identification and management of weak points. From this review, decision makers can formulate a plan of action to improve the cyber resiliency of the organization.

The process is organized and well ordered once the organization or body is identified for review through this methodology. However, lacking is an effective way to identify these bodies and organizations. The situation in Israel is fairly similar. From time to time, the Information Security Authority brings additional bodies to the steering committee of the National Security Council that will need to examine and meet the agreed upon guidelines. At the same time, there is no binding systematic statutory process that allows these organizations to be identified.

Because an area or a sector that constitutes a critical national infrastructure comprises a large number (hundreds, and sometimes thousands) of organizations and systems, protecting a "sector" is meaningless. Rather, in practice, protection entails actions taken by specific organizations, companies, facilities, and processes. Therefore, the question is how is it possible to locate these bodies, since almost every company or government office interfaces with sectors that are defined as defense-critical infrastructures. For example, protection of water supply and water quality infrastructures in Israel does not only affect processes in Mekorot, Israel's national water company, but also dozens of other water suppliers, associations, water corporations, desalination and delivery facilities, sewage and wastewater treatment facilities, and so forth. A large number of these facilities are operated by private entrepreneurs who do not see activating protective mechanisms as a top priority. The situation is similar in other industries.

Furthermore, in many cases it is also necessary to protect interfacing systems that are connected to the supervised bodies. For example: an industrial factory that has been declared an essential component of a particular sector works under the direction of the Information Security Authority. Sometimes this factory is dependent for its operations on other manufacturers (smaller satellite manufacturers) that supply input (sometimes critical) for the production process of this protected factory. In many cases, some of these satellite manufacturers are not included in the group of critical infrastructures for protection and therefore they do not use satisfactory information defense processes. Thus, it is possible that cyber damage to one of these manufacturers will cause significant damage to a protected factory.

The use of information technologies in Israel is widespread, both in the public and the private sectors. As such, Israel offers a wide range of targets for a potential cyber attack. Therefore, identifying additional bodies for guidance by the Information Security Authority is an essential task for building an optimal defense system. Reviews taken from time to time and information from various government offices are essential to this process, but they are not sufficient. A built-in mechanism must be created that will allow a significant improvement in these processes, especially concerning certain projects in the private sector that if exposed to cyber damage could suffer extensive damage that might have an impact on national security.

The Proposed Process: Use of Existing Statutory Tools

The principal proposal aims to make cyber protection a built-in component of the existing statutory process, both in the establishment stages (i.e., the approval of the projects in the various planning commissions) and in the operational process (the business licensing law). It is proposed that in the framework of the national planning processes, every project submitted to the planning commissions for approval will be required to submit a Cyber Resiliency Assessment. This assessment will constitute the main statutory tool for examining the project's exposure to the possibility of cyber attacks and the measures protecting against these exposures. This assessment will also provide the Information Security Authority a tool for identifying and managing the critical infrastructures for defense. At the same time, in the framework of the business license, which is a license requiring periodic renewal, the relevant authority can check the ongoing compliance with cyber protection instructions of the body under review.

The establishment of every project in Israel, including national infrastructure projects, requires compliance with the customary processes of statutory planning. Thus, projects that are required to build facilities and structures must be approved by various planning commissions in accordance with the relevant regulations on the local, regional, and national levels. Review of the planning documents submitted for approval is the planning authorities' central tool of control over these projects. Among the documents submitted for review by the planning commissions today are reports concerning firefighting, public health issues, environmental aspects, handling of hazardous materials, home front defense, and so forth. These documents define the steps that the project initiator will take in order to comply with the necessary requirements in each of the areas described above. These steps are then relayed to the authorized regulatory authorities, which employ experts to ensure that at the end of the process, the project is implemented with public interests in mind and that public security is maintained throughout the various spheres. In Israel, dozens of projects that if damaged might harm national security are discussed every year, including infrastructure facilities, water and sewage treatment facilities, delivery systems, transportation projects, energy facilities, and communications. Expansion and establishment of industrial factories and a wide range of other projects are discussed as well. Cyber damage to some of the projects and ventures is liable to harm the country's economy, not only directly, such as through the inability to supply an essential service, but also in the form of commercial damage, e.g., the inability of Israeli companies that were attacked to supply their products for a given period.

An example that clarifies the proposed process is the requirement to submit an Environmental Impact Assessment. The goal of the assessment is to identify the environmental hazards that are likely to be caused by the project, along with ways to minimize this damage to a tolerable level. Submission of the review is anchored in the planning and building regulations (of 1982, and in its final version of 2003). The idea for this review originated in the enhanced public awareness in the United States of environmental issues, which in 1970 led to legislation requiring preparation of an Environmental Impact Assessment as part of the planning process.

Together with the planning component of new projects, it is also possible to make use of the business licensing process, which requires periodic renewal to ensure that over the years the project meets the necessary criteria in various spheres, including protection from cyber attacks. According to Justice Mishael Cheshin, "the goal of the [business licensing] law is to preserve and protect various values that our society considers important . . . such as the value of public safety, with the value of maintaining public health and safety, and the value of preserving the environment and quality of life . . . protecting the goals of society."[8] Use of the tools provided by the business licensing law for cyber protection and upholding its goals provides the Information Security Authority with an additional legal tool to ensure that existing activities are required to meet the necessary criteria. In certain cases, there has even been a demand of private business owners to submit a Cyber Resiliency Assessment and a requirement to meet security guidelines.

Projects in the pre-establishment process and in certain cases those that have already been set up will be required to submit a Cyber Resiliency Assessment to the Information Security Authority, which can ensure that essential protection instructions are followed. A number of guidelines can be proposed for the content of this assessment and for those authorized to submit and those authorized to check it. From a statutory point of view, the review process must be applied comprehensively and govern all requests, unless the authorized authority grants an exemption. However, from a practical point of view, the Information Security Authority will be required to draft criteria that define the projects and ventures for which an assessment must be submitted. These criteria could address a number of components, such as the size of the project, its sector (for example, the energy sector, natural gas, and the like), the project's interfaces with elements already under the purview of the Information Security Authority, and the expected damage in the event of a cyber attack.

When a decision is made that the body must submit a Cyber Resiliency Assessment, the process will adhere to a defined procedure, as follows:

a. *Assessment guidelines*. It is the responsibility of the Information Security Authority to prepare guidelines for carrying out the assessment. These guidelines must be suited to the project or the specific body and cover a number of components, including: mapping the potential damage from a cyber attack; mapping the weak points of the project/plan; and issuing instructions that will make it possible to minimize exposure and damage.
b. *Assessment preparation*. The assessment will be prepared under the auspices and with the funding of the project initiator. For this purpose, there will be consultants from a group of designated consultants trained and authorized by the Information Security Authority. These consultants will work according to the assessment preparation guidelines.
c. *Checking the assessment*. By virtue of its responsibility, the Information Security Authority can use external advisors trained and authorized to check the reviews, with the cost charged to the project initiator. In this process, it is possible that there will be a number of rounds of questions and answers between officials in the Information Security Authority and the party under review.

d. *Approval of the assessment*, meaning examination and review by the authority's officials and a decision on guidelines in this context for the project. This approval can also address aspects of the stipulations for the business license, as well as instructions that should be applied to the project initiator's plans.

Similarly, the business licensing law also constitutes an appropriate platform for implementing instructions and guidelines in the realm of protection from cyber attack. Due to the restrictions applying to the security and flow of information, it will be necessary to define this process as a departmentalized process that is not open to the wider public, but only to specific authorized officials.

Conclusion

Threats to civilian companies have grown not only because of increased competition in the marketplace but also because of their exposure to attacks by hostile elements. Hostile parties identify the potential damage to the country's economic infrastructure inherent in attacking these companies. States tend to protect mainly bodies that have a direct connection to national security, which traditionally included primarily government offices; intelligence and security bodies; organizations engaged in sensitive classified security manufacturing; and classical critical infrastructures, such as electricity, water, transportation, and so on. The logic that defined the criterion of this privileged class was derived from the classic strategic concept: a list of national infrastructures susceptible to disaster in the event of war, and which if damaged could cause direct harm to the country's fighting ability and resiliency. However, what will be the fate of civilian companies such as Teva Pharmaceutical Industries, or food manufacturing companies such as Tnuva, the Strauss Group, and the like? And what of cable companies and insurance companies, not to mention memorial and heritage sites? A quick examination shows that damage to these organizations is liable to cause significant damage to the country and harm the fabric of civilian life.

The establishment of the Information Security Authority and the steering committee of the National Security Council were first steps in the right direction. Now, with the increasing realization that cyberspace is becoming a combat zone before our eyes, the ability of the State of Israel and its economy to weather attacks of this type must be enhanced. Introducing cyber defense in the statutory processes can allow ongoing, systematic monitoring of the immunity of Israel's cyber security system.

Notes

[1] This saying is usually attributed to Frederick the Great.

[2] The website of the Information Security Authority, http://www.shabak.gov.il/about/units/reem/pages/default.aspx.

[3] Gal Mor, "Plan for Information Security Approved by Government," *Ynet*, December 11, 2002, http://www.ynet.co.il/articles/1,7340,L-2310234,00.html.

[4] Patrick Beggs, "Securing the Nation's Critical Cyber Infrastructure," US Department of Homeland Security, February 25, 2010.

[5] Ibid.

[6] Patrick Beggs is the director of Cyber Security Evaluations – National Cyber Security Division in the US Department of Homeland Security.

[7] The term "delivery systems" serves to describe infrastructures that conduct materials: water, sewage, waste water, gas, oil, electricity, communications fibers, and the like.

[8] Justice Mishael Cheshin, Criminal Appeals Authority (CAA) 4270/03, State of Israel vs. Tnuva.

Cyberspace and Terrorist Organizations

Yoram Schweitzer, Gabi Siboni, and Einav Yogev

In a scene in the 1990 movie *Die Hard 2*, terrorists take control of computer, traffic control, and aerial communications systems, impersonate flight inspectors, and feed in false data, thus leading the pilot and passengers to their death in the midst of a snowstorm with the plane crashing on the runway. Security personnel are helpless, incapable of providing a response; the movie's hero, John McClane (played by Bruce Willis), lacks the means to save the doomed flight and is left standing powerless in the fog on the landing strip, waving two improvised beacons at the approaching aircraft. At first it would seem that the movie is nothing but another Hollywood fantasy, dismissible as a wild exaggeration carried to yet further extremes in the sequel, *Die Hard 4*. However, the events of 9/11 and the changes in the nature of security threats over the last decade indicate that even the most far-fetched scenarios crafted in Hollywood studios are liable to find real-life expression in the public and security sphere in this day and age.

The use of cyberspace as a primary warfare arena between enemies or hostile nations has always been fertile ground for fantasy and lurid scenes on the silver screen. However, cyberspace is rapidly becoming a genuine central arena for future wars and hostile actions undertaken by various types of adversaries. These may include terrorist organizations, although until now they have relied primarily on physical violence to promote their own goals and those of their sponsors. In light of such threats, many nations in the West have in recent years established special authorities to use innovative technological means to prepare for war-like actions against strategic infrastructure targets.

Yoram Schweitzer is head of the Terrorism and Low Intensity Conflict Program at INSS. Dr. Col. (ret.) Gabi Siboni is head of the Military and Strategic Affairs Program and head of the Cyber Warfare Program at INSS. Einav Yogev is a research assistant in the Terrorism and Low Intensity Conflict Program at INSS.

This article was first published in *Military and Strategic Affairs* 3, no. 3 (2011): 39-47.

This essay focuses on an analysis of the factors that are likely to make terrorist organizations use cyber tools to perpetrate attacks on critical infrastructures of sovereign institutions and symbols, commercial and industrial infrastructures and systems, and public civilian targets. In addition, it examines the question of whether the threat is actual and imminent, or whether it is a far-fetched possibility that surfaces from time to time in the general discourse on the subject.[1]

The Cyber Threat from Terrorist Groups

Today there are five main groups that use or have the potential for future use of cyber attack tools: 1) states developing offensive and defensive capabilities as a growing part of their force capabilities; 2) criminal elements motivated primarily by illegal commercial interests; 3) commercial companies, primarily in the defensive mode (as the scope of cyber attacks in the commercial context is significantly growing), though some may resort to offensive moves against competitors; 4) terrorist organizations, out of cost-benefit considerations and other inherent advantages, are liable to try to carry out cyber attacks; and 5) anarchists opposed to the existing establishment who are interested in undermining it from within and without, and who endeavor to attack the entire system of computerization, which today is the basis for managing life as we know it, in order to disrupt or even destroy states' current social order and their fabric of life.

Cyber offense has the potential to change society's balance of power because it empowers those engaged in asymmetrical conflicts that operate from a position of inferiority, especially terrorist organizations. Capabilities in this sphere may enable them to attack installations, systemic processes, and sites while causing heavy physical damage and wielding a significant psychological impact on the society and public under attack. They thus acquire capabilities other than those familiar from conventional terrorist attacks, such as suicide bombings, booby traps, hostage situations, hijackings, and kidnappings.

Cyber offense affords several advantages. First, it removes the necessity of physical presence at the target. It is possible to damage communications networks and control systems of installations and processes from afar and thus avoid physical barriers and human systems. Second, it affords a wider scope of damage. Cyber attacks occur not only in the physical space but also carry the potential for severe and sustained damage to control and infrastructure systems. Thus, while most conventional terrorist attacks are limited in time and space,[2] a cyber attack magnifies terrorism's psychological impact through fear and intimidation. Third, it is easier to conceal the identity and source of the attack; in cyberspace, identities and boundaries between states are more easily blurred. Terrorists attacking in cyberspace can not only conceal their identity but can also feed false information as to the source of the attack, for example, by attacking a site inside the target state using addresses of a friendly nation. Fourth, cyberspace attacks are cost effective. Using the cyber platform for attacks maximizes the cost-benefit ratio from the perspective of a terrorist organization, endowed with fewer resources and capabilities than the states it targets. Assuming that terrorist organizations would prefer less defended targets rather than well-protected ones, they presumably would be able to gain access and insert malicious code into target sites, or use technologies that are becoming ever more accessible to wider audiences. Fifth, cyber terrorism can be non-lethal. It can cause significant damage without direct fatalities or physical injury, granting terrorists success by means of intimidation and disruption of the routine. This gives the perpetrators the ability to devise a defense and logical explanations for their deeds, which after all did not spill blood but were only an indirect cause of lost lives. The innovativeness represented by such action would also garner terrorist organizations widespread media coverage and enable them to engage in non-lethal threats in which a price would be extorted in exchange for removing the threat of a cyber attack.

It has been claimed that terrorist organizations are not interested in cyberspace because they prefer showcase attacks with much higher visibility rather than the anonymity that supposedly is conferred by attacks in this domain.[3] However this claim does not take into account the basic rationale of terrorism strategy, which holds that terrorist activity should focus on minimizing the power differential in the struggle against a stronger enemy with more powerful means, carry out destructive actions while identifying the weaknesses in the enemy's defense, and achieve a position of superiority at tolerable costs given the relatively poor means at the disposal of the perpetrators. Already today global jihad terrorist organizations are making use of cyberspace, though still in limited and relatively undeveloped fashion, to realize these advantages. A study examining the cyberspace warfare capabilities of jihadist organizations[4] identified a number of major features that serve to build and improve the organizational and operational infrastructures of terrorist organizations in the following fields:

a. Propaganda: using the web to disseminate ideas, decrees, directives, speeches, and opinion pieces by clergy and terrorist leaders.
b. Recruitment and training: using the web to identify and recruit potential members as well as to transmit instructional and training materials.
c. Fundraising and financing: using the web to fundraise under the guise of charities and aid organizations as well as to steal identities and credit cards.
d. Communications: using the web for operational communications while employing a range of tools, including accessible encryption tools.
e. Identifying targets and intelligence: using information available on the web to identify targets and gather intelligence.

It is thus clear that an essential upgrade of cyberspace tools available to terrorist organizations, from logistical and propaganda tools to actual operational tools, is liable to generate an innovative, dramatic, and relatively cheap type of attack with the power to effect severe damage, even if carried out with a low signature or in total anonymity. Therefore every terrorist organization, especially one seeking fame and wanting to affect the public psyche and morale in the targeted enemy, sees such an attack as an important and worthy challenge. Innovation would also guarantee the perpetrators international fame and transform them into role models. Thus, sub-state entities with more limited technological capabilities than the nations with which they are at war are liable to join the trend of using advanced technology needed for cyber warfare for their own benefit, either by receiving assistance from supportive nations or by acquiring such capabilities themselves in the future, by recruiting and operating individuals with the necessary skills in this field.

As for states supporting terrorism, cyberspace is very attractive for use of proxy organizations because of the anonymity afforded by the domain, the difficulty in proving the identity of the perpetrator, the high level of deniability by states about their involvement, and the satisfaction of causing severe damage to the enemy. Even if suspicions are aroused, it is still hard to prove guilt. Furthermore, the public under attack may perceive a cyber attack to be less outrageous than a terrorist attack that employs firearms and causes direct death and destruction – even if the damage caused is greater, more destructive of property, and takes more lives than a violent terrorist act.

Despite these advantages of cyber attacks, to date no such attack has been traced to a terrorist organization. Development of significant capabilities in this field requires surmounting a considerable intelligence and technological threshold. At this stage one may assume that terrorist organizations find it hard to identify, harness, and maintain such high technological capabilities and access that would allow them to cross that bar. It is true that this limitation can be partially overcome through the assistance of state supporters of terrorism, but at least for now this is not enough to give terrorist organizations the significant, stable technological platform required for maintaining effective cyber attack capabilities. In addition, terrorist organizations face limitations posed by cyber surveillance and state intelligence and technological capabilities that enable them to identify suspicious conduct on the web, identify attempts at organization, and mount a defense against them and against threats to specific targets.

Weaknesses and Responses

Although to date terrorist organizations have not been able to overcome the difficulties in achieving offensive cyber capabilities, civilian systems and routine civilian life presumably remain their preferred targets, because these are much more difficult to protect than security systems. Strengthening defenses of critical national infrastructures such as electric, water, and communications supply networks would likely encourage terrorists to seek out less protected targets in the civilian and commercial sectors. Even though systems in these sectors are usually not included in the rubric of critical and protected infrastructures, from the terrorist perspective an attack against them could be effective, by breaching ordinary citizens' basic sense of security and enhancing the terrorists' image by instilling fear.

A significant part of constructing a defense against cyber attacks is general and independent of the source of the threat, whether terrorist, state or criminal. This is reflected organizationally – consider Israel's Information Security Authority and ministries specializing in cyber defense in various nations – and also in certain components of defense from the fields of information systems and general security. In contrast, in fighting terrorist organizations it is also necessary to activate two designated components that require sustained development and improvement.

The first is intelligence. Effective gathering of accurate, high quality intelligence requires using a range of sources, including open sources and material from the terrorists' own computers and networks. To this end it is necessary to develop capabilities of infiltrating these systems covertly and inserting information effectively and continuously. The challenge that must be overcome is the widespread global deployment typical of terrorist organizations that use many chat rooms and transmit messages using unique code words. Intelligence agencies must be able to intercept these transmissions and decode them within the relevant timeframes and at the same time provide cyber defense systems with the tools needed to protect against and even disrupt the planned actions.

The second component is disruption. Unlike defense systems, which do not try to prevent an attack but rather obstruct its success once it has already been launched, the goal of disruption is to thwart the execution of the attack or to hamper its progress. Establishing an effective disruption structure against cyber attacks by terrorist organizations requires intelligence monitoring and control that can identify the organization of an attack before it takes place and operate effectively to foil it. This aspect relies primarily on tactical intelligence gathering capabilities, both from computers and from communications networks used by terrorist organizations.

Disruption attempts can also be directed towards damaging the organizational infrastructures of the organization. An example of this occurred in England when British intelligence hacked the online issue of the British al-Qaeda magazine *Inspire*. In addition, in recent years the various components of the electronic jihad have been targeted for occasional cyber attacks largely attributed to Western governments: the Taliban's website has been hacked time and again, as have exclusive jihadist forums and high profile fundamentalist websites. Meanwhile, American, Saudi Arabian, and Dutch authorities have extracted valuable information about potential Islamic terrorism from jihadist websites serving as honey traps for high quality intelligence.[5]

At the same time, it is necessary to deepen the defenses of civilian systems that represent the greatest weakness and therefore are also preferred terrorist targets. For example, the British government began taking legislative steps that include authorizing the use of invasive techniques such as telephone wiretaps, surveillance of emails in police files connected to crimes of terrorism, torpedoing internet radicalization processes, and specialized training of police units to confront cyber threats.[6] Nonetheless, in most states the defense of civilian systems is still in its infancy. Most states' cyber defense resources are allocated to security systems and to what are considered critical national infrastructures. Deepening the defense of civilian systems requires radical changes on a national scale that must be supported by appropriate regulation.[7]

Conclusion

In December 2001, at a meeting in New York shortly after the 9/11 attacks, the philosopher Jacques Derrida presented his understanding of the changes generated in the world as a result of those events. According to Derrida, the attacks were still part of the "archaic theater of violence," the real, visible world, in which events are still conducted in "clear and great order." However, according to him, cyberspace presents us with a more potent threat to our political and physical world; the dangers inherent in it change the relationship between terrorism, in the psychological and historical sense of a violent attack, and the concept of territory. Now, in the new techno-scientific world, the threat we knew in the past as real has become an invisible, quiet, and swift threat, devoid of bloodshed, which, according to Derrida, is worse than the 9/11 attacks, which at least were directed against a known location at a particular point in time. Now we are facing a challenge that threatens the social and economic fabric of life that connects all of us and upon which all of us depend in every place and at every moment.[8]

The rapid technological developments and innovations of recent years in the domain of cyberspace have indeed created a battlefield that simultaneously brings together many varied populations, local and international, representing a desirable target and fertile ground of activity by sub-state entities. Since thus far there has been no known cyber attack perpetrated by a terrorist organization, the threat does not seem acute. The challenge facing those who would try to use cyberspace for malicious purposes is three-pronged: attaining high level intelligence, the ability to crack computerized systems protected with advanced technology (or accessibility to such ability), and very high levels of calculation and computerization skills.

However, the advantages afforded by attaining cyberspace capabilities as described in this essay are liable to serve as an incentive for terrorists to develop, acquire, or harness such capabilities in the future. Gaining control of the advanced technological and intelligence capabilities required in cyberspace is likely to give these elements who seek to seriously damage their enemies by causing massive destruction and sowing terror and intimidation in the public at large the ability to disrupt the normal routine of civilian life, undermine civilian trust in their governments, and of course gain valuable prestige and media stature.

Therefore, Western nations must work diligently to meet this threat and improve the effective intelligence and defensive capabilities of civilian systems, while at the same time construct accurate intelligence gathering capabilities and the ability to disrupt cyberspace organization and attack by terrorists. Neglecting the civilian cyberspace domain, which is an attractive target for terrorists, is liable to prove disastrous in the future and place security personnel, when the time comes, in the same position as that fictional Hollywood hero of *Die Hard 2* trying to save airplanes from crashing using nothing other than improvised beacons.

Notes

[1] The use of the term cyber terrorism in this essay refers to the use of cyber tools liable to be used by terrorist organizations to attack economic infrastructures and civilian systems in targeted nations.

[2] There are of course important exceptions: the 9/11 attacks in the United States had a global effect on flight security systems.

[3] Shmuel Even and David Siman-Tov, *Cyber Warfare: Concepts, Trends, and Implications for Israel*, Memorandum No. 109 (Tel Aviv: Institute for National Security Studies, 2011, p. 42).

[4] *Examining the Cyber Capabilities of Islamic Terrorist Groups*, Institute for Security Technology Studies at Dartmouth College, Technical Analysis Group, March 2004.

[5] Adam Rawnsley, "Stop the Presses! Spooks Hacked al-Qaida Online Mag," *Wired*, June 3, 2011, http://www.wired.com/dangerroom/2011/06/stop-the-presses-spooks-hacked-al-qaida-online-mag/June 4, 2011.

[6] "Warning of Rise in Cyber-terrorism," *The Independent*, July 12, 2011, http://www.independent.co.uk/news/uk/crime/warning-of-rise-in-cyberterrorism-2312434.html.

[7] Gabi Siboni, "Protecting Critical Assets and Infrastructures from Cyber Attacks," *Military and Strategic Affairs* 3, no. 1 (2011): 93-101, http://www.inss.org.il/upload/(FILE)1308129638.pdf.

[8] Jacques Derrida, in Giovanna Borradori, *Philosophy in a Time of Terror: Dialogues with Jürgen Habermas and Derrida* (Hebrew translation, United Kibbutz Press, 2004), pp. 173-74; also available (in English) at http://www.

press.uchicago.edu/Misc/Chicago/066649.html: “One will be able to do even worse tomorrow, invisibly, in silence, more quickly and without any bloodshed, by attacking the computer and informational networks on which the entire life (social, economic, military, and so on) of a ‘great nation,’ of the greatest power on earth, depends. One day it might be said: ‘September 11’ – those were the (‘good’) old days of the last war. Things were still of the order of the gigantic: visible and enormous! What size, what height! There has been worse since. Nanotechnologies of all sorts are so much more powerful and invisible, uncontrollable, capable of creeping in everywhere. They are the micrological rivals of microbes and bacteria. Yet our unconscious is already aware of this; it already knows it, and that’s what’s scary.”

Critical Infrastructure Protection against Cyber Threats

Lior Tabansky

Introduction

A functioning modern society depends on a complex tapestry of infrastructures: energy, communications, transportation, food, and many others. This article discusses the developing cyber threat to critical infrastructure while focusing on several dimensions: aspects to the threat that require an interdisciplinary approach; defense against the threat; the existing Israeli response; and the developing challenges. An informed public debate is likely to lead to improved protection of national infrastructures in the civilian and public sectors.[1]

The article begins by defining the subject of critical infrastructures, and discusses the origins, uniqueness, and innovativeness of the threat to them. It then discusses levels of coping with the threat, using conceptual parallels to the world of military content. The existing Israeli response will be reviewed briefly, with an emphasis on the central challenges the cyber threat poses to public policy. Finally, directions for future research and action will be presented.

What are Critical Information Infrastructures

An infrastructure is a system that combines various facilities and enables certain activities, for example, a pipeline that conducts water from wells to homes and fields, paved roads, bridges and intersections that allow movement of people and goods, flight, communications, fuel, and health services. One of the properties of an infrastructure is the dependence of various spheres

Lior Tabansky, a former Neubauer research fellow at INSS, is a doctoral student in the Department of Political Science at Tel Aviv University.

This article was first published in *Military and Strategic Affairs* 3, no. 2 (2011): 61-78.

of activity on it. In the past, the dependence stemmed from physical or geographical relationships only. With the development of cyberspace, which includes data communication systems and computerized methods of automatic command and control, there are additional relationships, which in turn create further vulnerability. These are computerized relationships (for example, command and control by remote electronic means) and logical relationships (such as the international financial market as a factor influencing inputs and outputs of critical infrastructures), which are innovations that would not exist without information technologies. It is therefore worth distinguishing between infrastructures in the traditional sense and the modern use of this concept, which includes a cyber dimension.

In the information age, traditional infrastructures become information infrastructures because they incorporate computers. In addition, new critical infrastructures have been created that are purely information infrastructures: computerized databases that contain important data, such as records of capital in the banking system, scientific and technical intellectual property, and the programmed logic that manages production processes and various business processes. In the information age, the concept of "infrastructure" also includes computerized components, and thus "infrastructure" today necessarily refers to an information infrastructure.

Infrastructure is defined as critical when it is believed that disrupting its function would lead to a significant socio-economic crisis with the potential to undermine the stability of a society and thereby cause political, strategic, and security consequences. Different countries have offered a variety of definitions of critical infrastructures.[2] What all have in common is the existence of a computerized element upon which other physical systems are dependent and which, if harmed, would likely cause widespread damage in physical terms.[3]

Three factors can define a critical infrastructure. The first is the symbolic importance of the infrastructure. Thus, several democratic countries include heritage sites, museums, archives, and monuments among critical infrastructures that should be protected from cyber threats.[4] Another source of symbolic power is the perceived control of a government. For example, a hostile disruption of traditional media used by the state for communicating with its citizens will immediately harm the government's ability to function. Moreover, in the longer term, such disruption may diminish the citizens' confidence in the existing government, or even the general form of government or regime.

The second factor is the immediate dependence on infrastructure, such as the electricity grid or telecommunications network, which is obvious for most processes in society. The emergence and prevalence of cyberspace created a situation in which computerized networks constitute an infrastructure in and of themselves. Cyberspace is a representative example of an infrastructure that has become critical because of the interface of most of society's activity with computerized communications networks.

The third factor involves complex dependencies. The accelerated trend toward adding connectivity capabilities enables unanticipated effects beyond the local level (the "butterfly effect").[5] The relationships among various infrastructures are presumably not fully known, and the failure of one component is liable to cause a wide range of results and damage. The types of failure fall into three classes:

a. *Common cause failure*. For example, various facilities (fuel storage, airports, and power stations) that are located in geographic proximity are likely to be harmed from a single incident of flooding. It is hard to imagine a cyber attack that would directly cause a failure of this type.

b. *Cascading failure*. Disruption of a control system in one infrastructure (for example, water) leads to disruption of a second infrastructure (for example, in transportation, the flooding of a railway line), and then a third (for example, food supply chain) and so on, even if it is not directly dependent on it. A cyber attack could directly cause such a failure.

c. *Escalating failure*. Disruption of one infrastructure (for example, a communications network) harms the effort to fix other infrastructures that have been damaged by another entity (emergency services, commerce).[6] A cyber attack could directly cause this type of failure.

The commercial aviation sector, which has attracted the attention of enemies of the developed states and prompted noticeable acts of hostility – hijacking of commercial planes, the September 11 attacks, and other terrorist attacks using civilian airplanes – can illustrate the importance of critical infrastructures and the significance of an attack on them. Civil aviation is a basic infrastructure for developed societies: in 2009, commercial air transport carried more than 2 billion passengers on 28 million flights on 27,000 airplanes operating from 3,670 commercial airports around the world.[7] In addition to commercial flights, military aircraft (some unmanned) also populate the skies. Intra-state laws, regulations, and procedures, along with international cooperation, regulate the administrative aspect of the airline industry. Airports are connected to each other through scheduled air traffic, and the air traffic control system in each given location is part of the international aviation infrastructure. Air traffic control is based on computerized systems: methods of detection, monitoring, surveillance, automation, communications, command and control, and so on. Disrupting the proper functioning of air traffic control systems would harm all air traffic.

The Novelty of the Threat

Recent years have brought increased concern over the potential vulnerability of the infrastructures that are the basis of developed modern societies,[8] yet the fact that this discussion is taking place now is surprising. Critical infrastructures have always been critical and their importance is obvious. International and internal conflicts are not new to the world, and in war it is reasonable to anticipate attempts to harm the adversary's critical infrastructures with the goal of weakening and defeating it. In 1917, during the Bolshevik Revolution, Lenin and Trotsky ordered their activists to take over the post office, telegraph systems, bridges, and train stations. In prolonged wars, such as the Second World War, attempts have been made to harm critical infrastructures in order to interfere with the enemy's fighting ability and spirit.[9] A country's critical infrastructures, whatever they are, are elemental targets during a conflict, and therefore organizations and states have labored throughout history over defense systems for their infrastructures: camouflage, guarding, fortification, defensive forces, deterrence, and so on. Why, then, is there a growing fear of damage to critical infrastructures, particularly in the strongest countries?[10]

A critical infrastructure is a tempting target for an enemy, be it a terrorist organization or a hostile state. However, the developed countries currently enjoy total military superiority over their respective enemies. The US and Europe have not experienced wars on their territories in recent decades. Israel is the only developed country that is under ongoing military threat that is manifested in a variety of ways (missile attacks in 1991, rockets in the north and south of the country,[11] and suicide bombers in 2000-2005). Several developed countries have been harmed by hostile acts that directly attack the civilian population by circumventing the military that was supposed to protect it. The terrorist attacks could not threaten the countries attacked, but they did succeed in causing a change in their policy in one way or another.

In all forms of traditional warfare, the identity of the enemy is disclosed following the attack because in order for the attack to be carried out, the weapons must physically reach the target. In the event of a missile launch as well, there is no doubt as to the location of the launch site. The hijacking of commercial aircraft in the 1970s, the suicide bombings in Israeli population centers, the attacks in the United States in September 2001, and the attacks in Madrid in 2004 and London in 2005 all required the attackers to be physically present at of the attacks.

Identifying the enemy is critical for response and deterrence. Thus what prevented harm to critical infrastructures in the past was the defensive force placed in the path of the enemy, and even more so, the deterrence that promised to exact a heavy price. This familiar state of affairs came to an end with the development of cyberspace. For the first time in history, it is possible to attack strategic targets (such as critical infrastructures) without physically being in the place where they are located, without confronting defensive forces, and without exposure. In today's reality, the existing computerized infrastructure can be exploited through penetration of communications networks or the software or hardware of the command and control computers in order to disrupt, paralyze, or even physically destroy a critical system.[12] The threat stems from the vulnerability inherent in the properties of cyberspace,[13] and because of these special characteristics, the cyber threat challenge differs fundamentally from the challenges of traditional threats.

Levels in Confronting the Threat

This article focuses on the cyber threat to the computerized part of the infrastructures, based on the realization that such a threat has become possible, available, significant, and is liable to disrupt the functioning of developed society.

Confronting the threat to critical information infrastructures includes prevention, deterrence, identification and discovery of the attack, response, crisis management, damage control, and a return to full function. When examining ways to confront threats to national security, the accepted practice is to divide the discussion into the tactical, operational, and strategic levels. Proposed here is a division of methods for confronting the threat to critical communications infrastructures into a number of levels: technological, technical-tactical, operational, and national-strategic.

The technical level focuses on an organization's computerized system, which is the most common activity in this realm. Given the large volume of activity, the technical aspect of "information security" is often emphasized, though it is actually a concept that deals with both defense of critical infrastructures and cyber security in general. In addition, activity that examines the issue from a comprehensive national perspective, referred to below as the national level of cyber security, is underway.

All the levels are required to confront the threat, but given the different focus, it is worthwhile distinguishing between these levels of protection. The proposed division will help identify the essence of the challenges of protecting critical infrastructures particular to cyber security.

The Technical Levels: Tactical and Operational Levels

Since the threat is derived from the properties of computer technologies, the response to the threat is generally sought among computer experts. As expected, the proposed solutions are also based on computer technologies. The problem is perceived as a technical problem, and therefore, the proposed solution is an engineering solution. The technical and operational levels for confronting the cyber threat, which come from engineering, mathematics, and computers, focus on identifying vulnerabilities in an organization's computerized systems and seek an engineering solution that reduces this vulnerability.

Table 1 displays common issues confronted by the technical levels of protection.[14]

The primary means of attempting to build resilience[15] is to invest in backup, redundancy, air gap, and the like. Accordingly, important computer systems are built twice, in separate locations, in order to enable continued function in the event of physical damage to the system.

Today, most solutions to the engineering problems identified are implemented through the private market. Information security is a wide ranging field, and describing it is beyond the scope of this article. In the division proposed here, information security lies in the technical-operational levels. Information security is a developing discipline that brings together many resources for research and development, consulting services and outsourcing, a security product industry, and the like. The worldwide

Table 1. Types of Vulnerability and Responses

Vulnerability	Response
Access passwords for devices and systems are not changed from the default.	Password management
Passwords are saved and sent without encryption.	
Access passwords are not changed periodically.	
Physical security is lacking.	Physical access security
People who do not deal with critical equipment have access to it.	
Faulty management of user permissions gives a low level employee access to a critical process.	Computer access security
A firewall configured improperly allows unnecessary types of communication.	
The process network is not separated from the office network.	
The possibility of remote access to the computer system has been left open.	
The computer system can be accessed from a wireless network.	
The remote access process uses an open protocol and weak passwords.	
The manufacturer of the system supplied security updates but they were not installed in the system.	Configuration management
Administrator rights were given to regular users.	
Access to critical system components was not monitored; no log information was collected.	
Information log is not checked on an ongoing basis.	

information security market is expected to grow, and some market analysts claim (perhaps with some exaggeration) it will reach $125 billion in 2015. Most of these revenues will go to US and European companies that offer combined solutions of technical goods and services, together with technological-business consulting.[16]

The issue of cyber security, and especially of critical infrastructure protection, came about as a result of technological change. At first, it was expected that the solution to a problem of technical origin would be technical. However, there is a growing understanding that this problem cannot be dealt with on a technical-operational level only, since a precise engineering formula for dealing with the cyber threat is not possible: society's structure, values, and institutions are integral parts of the environment.

The Top Level: The National Strategic Level

The national strategic level examines the threat to critical infrastructures in the framework of national security, with a national focus that goes beyond the boundaries of an organization or a business process. This approach sees the protection of critical information infrastructures as part of the protection of society as a whole. Protection of information infrastructures actually becomes protection of an information-based society.[17] Information security, which is at the center of the technical level, is a necessary but by itself insufficient part of the strategic vision. The highest national level is based on technical and operational foundations, but in a broader approach it is not sufficient to fix local problems of organizational systems. As in the military, the strategic level needs an appropriate operational level, but this is not sufficient to achieve the strategic goal.

In a wider national perspective, a comprehensive national policy on protecting critical infrastructures is needed, which in addition to the engineering foundations will take into account the complex social, political, economic, and organizational aspects. An organizational entity capable of taking into account the complex of relationships between critical infrastructures and a functional society and the state is also required. The national level of protection requires cross-organizational activities, backed by effective authority. Without a doubt, this is a complex challenge for public policy, considering the structural limitations of public service on the one hand and a required level of strategic focus of those in the private sector, on the other. Just as the state defends its entire physical space, it also sees an increasing need to protect cyberspace fully, in spite of its special characteristics, which make the task more difficult.

Issues for Policymakers

The information revolution continues to change the strategic environment, and it affects a range of social, cultural, and economic issues in complex ways. Cyber security, and in particular, protection of critical infrastructures, is already on the agenda. The development of cyber threats to a national security issue makes governments into the main customers of protection services. Even limited experience shows that there are differences in the framework of the discussion and the types of solutions proposed in different countries, in spite of the great similarity in the source of the threat. Since the threat is similar, the explanation for the differences must be the role social institutions play in the discussion and in determining the response. What follows are the main issues concerning cyber threats that call for a public debate.

Which infrastructure is critical?[18] Any discussion on protection and defense measures must begin with prioritization. Assessing and measuring the level of the threat to components, computers, and systems is a necessary precondition for effectively confronting the threat. The exact sciences and engineering have mathematical methods for measuring the relationships and the dependence between components and the system. These tools are also used in the technical levels of protection of critical infrastructures. Nevertheless, more comprehensive methods are needed for assessing risks that stem from the intricate relationships among complex technological systems that critical infrastructures contain.

An assessment of how critical an infrastructure is on a national level must address the full matrix of social values, goals, and interests. Therefore, the relative importance of infrastructure and the amount of public investment needed to protect it are not derived from an engineering formula, and require a wide ranging and informed public discussion. Representative political institutions are the place for such a discussion in a democratic society. Given the constraints of the political system, such a discussion will presumably be lengthy and at times frustrating. Nevertheless, only through a joint political process will it be possible to design an optimal response to the threat for the long term.

Cyber vulnerability: technical issue, economic risk, or security threat? What is the potential significance of the growth of cyberspace in general, and the harm to critical cyber infrastructures in particular? The topic clearly goes beyond the scope of computers, engineering, and information security to the question of the role of the state in cyber protection of critical infrastructures. Is this task military, partially civilian, "homeland defense," or civilian-commercial? The answer directly affects the solution proposed, and it has wide political, budgetary, and organizational consequences. Until recently, the common assumption was that this is mainly a technical issue, and the response therefore was placed in the hands of computer experts. Commercial companies provided technical solutions for the military, commercial, and civilian sector, and governments did not play a significant role. Today it is clear that the optimal answer can be found only in a joint discussion between various sectors in society because it is derived from the values of the society, its political and social structure, and its national security concept.

A political process for finding the balance between the values of freedom, market ideology, and security requirements: Critical infrastructures and the information necessary for their proper functioning affect all areas of a citizen's life. They raise many issues that affect civil rights, such as privacy, confidentiality, and due process; the relative strength of the state, citizens, and corporations; and allocation of public funds. Therefore, the central challenge in designing a policy to protect critical infrastructures from cyber threats is not technical or operational, rather a challenge of a comprehensive national-strategic vision. Critical infrastructure protection is not the exclusive preserve of systems engineers and computer experts. The optimal response to the cyber threat in general and the threat to critical infrastructures in particular will be created only through a broad public discussion in the framework of a democratic political system.

The private market and cyber security: The cyber threat is affected by the decentralized nature of economic activity in an era of rapid technological change, globalization, and privatization. The global market economy has created the situation in which large parts of the critical infrastructures are privately owned.[19] The unprecedented mutual dependence in international trade is one of the prominent expressions of globalization and privatization. The industrialized nations import most of the raw food that their citizens consume and export finished products and services. Food retailers do not keep inventory beyond several days' worth of typical consumption, and they depend on the continued undisturbed function of the extensive logistical supply chain to satisfy demand within a short time.[20] Given that disruptions in food supply would be a grave problem of wide social implications, this supply chain could be perceived as a "critical information infrastructure" and become an urgent policy issue.

Open societies[21] with free economies shy away from state intervention in business processes. In the world of free markets, any attempt at government intervention in market processes is viewed with suspicion. Thus, for example, the arguments against government regulation of the internet originate with the ideology that goes along with a free market. The solution adopted thus far was focused on regulation: in the United States, since the mid-1990s detailed standards have been developed and adopted for securing information in various sectors and industries,[22] and organizations for supervision and control have been established. However, the world financial crisis of 2008 illustrated the dangers of private ownership of critical infrastructures, even if subject to regulation.

In the past year, the critical infrastructures protection policy in the United States has shifted from an emphasis on market mechanisms and voluntary "private-public cooperation" to a model that gives the government broad powers to guide business institutions and supervise implementation.[23] Israel too has regulation of critical infrastructures, and there was a proposal to expand it to small businesses.[24]

The computer products market and cyber security: The state of the market in this area is not encouraging. Security is secondary, as opposed to quick time to market. Furthermore, it is much more difficult to make the effort necessary for resilience and reliability testing in a private commercial environment, because achievements are measured by the length of time it takes to receive a return on the initial investment and the reduction of expenditures not connected to the core activity, and there is protection of limited liability only. Today, manufacturers of computer systems have no incentive to invest in increased reliability and protection. Security is seen as an external function, an addition to the core system, sometimes from another manufacturer that does not receive the cooperation of the original manufacturer.

The level of reliability and information security in most software, hardware, and computer system communication is thus lacking today, and this broad vulnerability has undoubtedly contributed to the rise of the cyber threat. Security systems must be easy for any user to operate, require minimal computer resources, and not harm the functionality of the core system or the user experience. Given the legal, economic, and competitive circumstances, it is difficult to expect productive voluntary cooperation between private companies in these fields. However, nationalization is not the answer, nor should it be expected as a condition for increasing cyber security. In light of the cyber threats, what is needed is developing government policies to direct the market towards a greater level of security overall.

The Israeli Response

Securing sensitive information and protecting computer infrastructures are not new issues for the State of Israel, and there are Cabinet decisions dating back to 1996 on defense against cyber threats.[25] The format for protecting computer infrastructures was laid out in decision B/84 of the Ministerial Committee on National Security, "Responsibility for protecting computerized systems in the State of Israel" on December 11, 2002. To this day, this decision serves as the basis of the Israeli response to the cyber threat to critical information infrastructures. The response mandated by the decision includes establishment of a steering committee which, from time to time, examines the identity of the institutions that it is critical to protect, and the establishment of a government unit to protect civilian computerized infrastructure, the Information Security Authority[26] (RE'EM). RE'EM was established within the Israel Security Agency (Shabak) in order to comply with legal restraints on government intervention in business, since by law only civilian authorities, such as the police or the GSS, can intervene in private businesses. RE'EM oversees IT security in institutions that have been defined as critical: provides guidance, oversees implementation, and is authorized to institute sanctions against those that violate its directives. The institutions bear the costs of the protection required. Other important institutions that are under the responsibility of a government ministry operate according to RE'EM professional guidelines but are not legally overseen by it. The IDF and intelligence community protect their specific infrastructures independently, with RE'EM formal guidance

In comparison with the situation abroad, it appears that at the time this decision was made and implemented, Israel was relatively advanced in designing and implementing protection of critical infrastructures on a national level. However, cyberspace has continued to develop rapidly since then, and new systems and relationships have developed that cannot necessarily be defined as critical national infrastructures. One example is small and mid-sized businesses dependency on commercial communications providers and open internet. The bloom of commercial and consumer "cloud computing" applications raises new issues and indicates yet again the increasing importance of cyberspace in all realms of life.

The Israeli policy for critical infrastructure protection was set up nearly a decade ago and served it well. Nowadays it may lack a comprehensive view of the interconnectivity developing in cyberspace that serves all civilian commercial activity. It is therefore worth reexamining the existing and anticipated challenges and the desired response. Last year, the government launched a National Cyber Initiative to advise the government on cyber security issues.[27] The National Cybernetic Task Force, an expert committee of academics and practitioners working under the auspices of the National Council for Research and Development in the Ministry of Science and Technology, formulated recommendations.[28] On August 7, 2011 the government of Israel decided:

> To work to promote the national capability in cyberspace and to better confront the current and future challenges in cyberspace: to improve protection of national infrastructures that are critical for normal life in the State of Israel and to protect them, to the extent possible, from cyber attack, while promoting Israel's status as a center for developing information technologies, encouraging cooperation between academia, industry, and the private sector, government ministries, and special institutions...Accordingly, pursuant to decision number B/84 of the Ministerial Committee on National Security, dated December 11, 2002, and without prejudice to the authority given to any other party under any other law or Cabinet decision [it is decided]:
>
> 1. To establish a national cyber headquarters in the Prime Minister's Office.
> 2. To arrange responsibility for handling the cyber field.
> 3. To promote the ability to protect cyberspace in Israel and to promote research and development in the cyber field and in supercomputing.[29]

The Cabinet decision is likely to lead to improved regulation for an Israeli response to the cyber threat in general, and the threat to critical infrastructures in particular.

Conclusion

The renewed discussion on critical national infrastructure protection focuses on the cyber dimension. Since all infrastructures have been affected by the information revolution and all now include computerized components that are mainly for command and control, this rapid technological change has created a new, additional security threat. The nature of cyberspace allows an attacker to disrupt the functioning of critical infrastructures without being physically near the target and without risking unequivocal discovery by the party attacked.

Although at first glance it appears that the subject of protecting critical information infrastructures belongs in the realm of computer engineering, upon further examination it becomes clear that it should be expanded beyond the technical aspect. Indeed, the major challenge in protecting critical infrastructures from cyber threats is not technical, but strategic and political. Today most states have legal and technical regulation for selected sectors. Since 2002, through the oversight and guidance of a particular organization, the State of Israel has been protecting infrastructures it deems critical. However, the development of cyberspace has left its civilian and non-critical sectors unprotected, and at the same time, raised both the level of vulnerability and the potential severity of effects. The recommendations of the new National Cyber Initiative are expected to set a policy process in motion.

The cyber threat to critical infrastructure is perhaps the most significant issue in the realm of cyber security. Only a thoughtful, informed process can design a policy of effective critical infrastructure protection from cyber threats and thus reduce the risk confronting the State of Israel and other developed countries from cyberspace. The major recommendation, therefore, is to broaden the public discussion of cyber security to include social and cultural aspects, which will make it possible to cope with the threat optimally on a national-strategic level with a comprehensive national perspective.

Notes

[1] This article was written before the launch of the National Cyber Initiative, which also dealt at length with the topic discussed here. However, the recommendations of the National Cyber Initiative have not yet been released publicly.

2 Critical information infrastructures are systems and facilities whose destruction or interference (by means of computers) would: "a. cause catastrophic health effects or mass casualties comparable to those from the use of a weapon of mass destruction; b. impair Federal departments and agencies' abilities to perform essential missions, or to ensure the public's health and safety; c. undermine State and local government capacities to maintain order and to deliver minimum essential public services; d. damage the private sector's capability to ensure the orderly functioning of the economy and delivery of essential services; e. have a negative effect on the economy through the cascading disruption of other critical infrastructure and key resources; or f. undermine the public's morale and confidence in our national economic and political institutions." See U.S. Government, White House, Homeland Security, Presidential Directive 7: *Critical Infrastructure Identification, Prioritization, and Protection*, December 17, 2003, http://www.dhs.gov/xabout/laws/gc_1214597989952.shtm#content.

3 Elgin Brunner and Manuel Suter, *International CIIP Handbook 2008/2009: An Inventory of 25 National and 7 International Critical Information Infrastructure Protection Policies* (Zurich: Center for Security Studies, ETH Zürich [Swiss Federal Institute of Technology], 2008); John Moteff, Claudia Copeland, and John Fischer, *Critical Infrastructures: What Makes an Infrastructure Critical?* (Washington, D.C.: Congressional Research Service, Library of Congress, 2002); Myriam Dunn, "The Socio-Political Dimensions of Critical Information Infrastructure Protection (CIIP)," *International Journal of Critical Infrastructures* 1, no. 2-3 (2005); U.S. Department of Homeland Security, *National Infrastructure Protection Plan (NIPP) 2009*, http://www.dhs.gov/xprevprot/programs/editorial_0827.shtm; Tyson Macaulay, *Critical Infrastructure: Understanding Its Component Parts, Vulnerabilities, Operating Risks and Interdependencies* (Boca Raton, FL: CRC Press, 2009); Robert Radvanovsky, *Critical Infrastructure: Homeland Security and Emergency Preparedness* (Boca Raton, FL: CRC/Taylor & Francis, 2006).

4 For example, Australia and the United States, which are countries that clearly attribute great importance to their political history as a central element in their collective national identity and social and political strength. *International CIIP Handbook 2008/2009*, Table 1; U.S. Department of Homeland Security, U.S. Department of the Interior: *National Monuments & Icons: Critical Infrastructure and Key Resources, Sector-Specific Plan,* 2010, http://www.dhs.gov/xlibrary/assets/nipp-ssp-national-monuments-icons.pdf.

5 This refers to a tenet of chaos theory describing how tiny variations affect complex systems. The chaos theory attempts to describe the phenomena through mathematical methods.

6 Harm to the government's level of functioning, which harms services to citizens, creates escalation: public confidence in the government drops, and this is liable to be expressed in political change (a change of government in a representative regime) or even regime change (a revolt against an

authoritarian regime or a change in the structure of the regime in a democracy).

7 IATA (International Air Transport Association), *Air Transport Facts (2009),* http://www.iata.org/pressroom/facts_figures/fact_sheets/Pages/economic-social-benefits.aspx. The IATA represents 93 percent of scheduled air traffic in the world.

8 The United States was a pioneer in this field, initiating a discussion on the presidential level in 1996: United States, President's Commission on Critical Infrastructure Protection, *Critical Foundations: Protecting America's Infrastructures: The Report of the President's Commission on Critical Infrastructure Protection* (Washington, D.C.: U.S. G.P.O., 1997).

9 In the ″strategic bombing campaign″ in World War II, the allies concentrated their aerial effort on attacking German factories producing ball bearings and lubricating oils, refining facilities, and railroad junctions. The operation was intended to harm the critical infrastructure for weapons manufacturing.

10 The United States has led the response to cyber vulnerability since the mid-1990s, having enormous technological and military strength and being the only superpower.

11 Since 2001, terrorist organizations have launched rockets and mortars from the Gaza Strip at towns in the Negev. The rockets have thus far caused nineteen deaths, and the mortars ten, and they have seriously disrupted life in the region. Following an escalation, Israel launched Operation Cast Lead in December 2008, which ended with a military victory. High trajectory fire from the Gaza Strip continues to this day, although there is less than before the operation.

12 The feasibility of using cyber means to cause physical damage has been shown in experiments. A CNN broadcast that discussed the Aurora experiment, ordered by the US Department of Homeland Security and conducted at Idaho National Labs, noted that broadcasting instructions to the command and control system of the electricity generating system caused a generator to stop working and then to explode.

13 Following is a summary of the challenges stemming from the characteristics of cyberspace as it exists today: the major vulnerability of computerized systems; the difficulty in distinguishing between a glitch and an attack, making the connection between an event and the result, tracing the source of the damage, and identifying the attacker, even if the source of the damage is known; and the widespread use of off-the-shelf commercial technologies. For a discussion of cyberspace in the context of national security, see Lior Tabansky, "Basic Concepts in Cyber Warfare," *Military and Strategic Affairs* 3, no. 1 (2011): 75-92.

14 Jason Stamp et al., *Common Vulnerabilities in Critical Infrastructure Control Systems* (Albuquerque, NM: Sandia National Laboratories, 2003), http://energy.sandia.gov/wp/wp-content/gallery/uploads/031172C.pdf.

15 Resilience is the system's ability to absorb an attack and return to proper function quickly. In computerized systems, the result is achieved by restoring the original situation (going back in time) or by quickly adjusting to new constraints (adaptation).

16 See http://www.strategyr.com/Information_Security_Products_and_Services_Market_Report.asp.

17 James Der Derian and Jesse Finkelstein, "Critical Infrastructures and Network Pathologies: The Semiotics and Biopolitics of Heteropolarity," in Myriam Dunn Cavelty and Kristian Søby Kristensen, eds., *Securing "The Homeland": Critical Infrastructure, Risk and (In)Security* (London and New York: Routledge, 2008).

18 There is a great difference between the definition of critical infrastructure and the means taken to protect it in the various countries. See Brunner and Suter, *International CIIP Handbook 2008/2009*. The civilian aspect of protection of critical infrastructures in Israel is grounded in the Laws to Regulate Security in Public Places, 1998. The law authorizes the General Security Services to instruct various public institutions in physical security, information security, and essential computer system security, according to details appearing in annexes to the law. This law set punishments for failure to follow its instructions, including a civil fine and incarceration. In 2003, the government Information Security Authority was established, which is "charged with professional guidance of the institutions under its responsibility in the area of protecting critical computer infrastructures from the threats of terrorism and sabotage, in the area of classified information security, and in threats of espionage and exposure." See http://www.shabak.gov.il/about/units/reem/pages/default.aspx.

19 Most public transportation in the United States and more than 85 percent of the country's energy sector are controlled by private commercial companies. Some 85 percent of the communications of the US Defense Department uses commercial networks. See http://training.fema.gov/EMIWeb/IS/IS860a/CIKR/energy1.htm.

20 The State of Israel, by virtue of its geopolitical situation, keeps an inventory of food and equipment in order to assure the needs of the economy in an emergency. The Supreme Emergency Economy Authority – Food and General Economy, which is part of the Ministry of Industry and Trade, is the body responsible for this issue today.

21 This concept comes from philosopher of science Karl Popper. See Karl Popper, *The Open Society and its Enemies* (Routledge: 2011).

22 See, for example, the publications of the US National Institute of Standards and Technology, http://csrc.nist.gov/publications/PubsFL.html, as well as the electrical standards of the North American Electric Reliability Corporation (NERC), CIP-002-3 through CIP-009-3, http://www.nerc.com/fileUploads/File/Standards/Revised_CIP-002-1_FAQs_20090217.

pdf and http://www.nerc.com/fileUploads/File/Standards/Revised_Implementation_Plan_CIP-002-009.pdf.

23 CSIS Commission on Cybersecurity for the 44th Presidency, *Cybersecurity Two Years Later: A Report of the CSIS Commission on Cybersecurity for the 44th Presidency* (Washington, D.C.: Center for Strategic and International Studies, 2011).

24 Gabi Siboni, "Protecting Critical Assets and Infrastructures from Cyber Attacks," *Military and Strategic Affairs* 3, no. 1 (2011): 93-101.

25 See, for example, Cabinet decision1886 BK/9 from March 20, 1997: Establishment of a steering committee on computerization in every government ministry; Cabinet decision 3582 BK/77 from March 16, 1998: Responsibility for the subject of information security in government ministries; Cabinet decision 4956 BK/179 from March 23, 1999: Establishment of a council to secure sensitive information in the Prime Minister's Office; Cabinet decision TM/80 from November 26, 2000, on responsibility for computer information security in the IDF and cooperation with civilian authorities; Cabinet decision TM/14 from July 18, 2001: A secure internal network for government ministries.

26 http://www.shabak.gov.il/about/units/reem/Pages/default.aspx.

27 The National Cyber Initiative deals in part with the subject of protecting civilian cyberspace.

28 The committee recommended establishing a national cyber headquarters to report directly to the Prime Minister, with a budget of NIS 100 million; establishing an office to deal with the country's infrastructure and the civil sector; policy and regulatory change to encourage the cyber industry; encouraging cyber R&D; developing centers of excellence and encouraging academic and industrial research. Shmulik Shelah, "'Israel Vulnerable to Cyber Attack on Civilian Targets,'" *Globes* July 5, 2011, http://www.globes.co.il/serveen/globes/docview.asp?did=1000660740.

29 Cabinet secretary announcement at the end of the Cabinet meeting of August 7, 2011, paragraph 4: Promoting national capability in cyberspace, http://www.pmo.gov.il/PMO/Secretarial/Govmes/2011/08/govmes070811.htm. At the time of writing the organization is not yet functioning.

What Lies behind Chinese Cyber Warfare

Gabi Siboni and Y. R.

兵之形，避實而擊虛
"Avoid strength, attack weakness."
Sun Tzu, *The Art of Warfare*

Introduction

Over the past several years China has been developing operational capabilities in the field of cyberspace warfare. A cyber attack may be defined as the unauthorized penetration of computer and communications systems belonging to individuals or organizations for the purpose of espionage and information theft, in order thereby to damage or disrupt the functioning of these systems or to damage other systems dependent on them, even to a point of causing actual physical damage. Despite denials by the Chinese government, researchers posit that China is behind a string of cyber attacks[1] against the United States,[2] Japan,[3] France,[4] Australia,[5] and other Western nations.[6]

Chinese activity in the field of cyberspace warfare is intensive and aggressive. It appears that China, focusing on extensive collection of intelligence and commercial information in various fields, is targeting a range of companies – from those with specific technological expertise to organizations with financial and economic knowledge, such as in the cyber attack on the International Monetary Fund in late 2011.[7] However, the fact that companies and organizations providing essential services and communications infrastructures have also been attacked suggests that there

Dr. Gabi Siboni is head of the Military and Strategic Affairs Program and head of the Cyber Warfare Program at INSS. Y. R. is a senior figure at the Prime Minister's Office.

This article was first published in *Military and Strategic Affairs* 4, no. 2 (2012): 49-64.

many be other motives in play. If so, what underlies these attacks, and is it possible to identify the strategic principle with which China operates in the West in general and the United States in particular? To this end, one must examine China's cyber warfare strategy, the Chinese organizations involved in recent years, and the resources invested to realize China's goals through this type of warfare.

It is commonly assumed that before 2009, most of the attacks attributed to China were directed against the American military and the administration, such as Operation Titan Rain against American government agencies[8] and Operation Ghost Net against diplomatic targets in the UN. By contrast, in recent years the attacks attributed to China have been directed against civilian targets, including national infrastructures of critical importance, companies forming a part of the chain of access to those targets, and companies that if attacked, generate an outcome that serves an economic or commercial need.

In recent years there has also been a quantitative leap in attacks against infrastructures. The first was the Shady RAT series of attacks from mid-2006 until February 2011.[9] The second series was Operation Aurora, an especially sophisticated series targeting Google, a critical infrastructure at the global level. These started in mid-2009 and lasted until the end of that year. The third, which received a great deal of media attention, was against RSA, a company specializing in information security and internet servers providing secure ID and one-time password services.

This essay argues that an analysis of the publicly available information about the more recent attacks makes it possible to establish that China does in fact stand behind these attacks and also makes it possible to identify the link between China's cyberspace warfare strategy and its choice of targets. The analysis includes an examination of the companies attacked to identify possible motives for the attacks. For example, attacking companies and organizations supplying technology allows access to general cutting-edge technology, military technology, and so on. The motives for these attacks are presumably to steal capabilities and conduct industrial espionage against nations and commercial competitors. Attacking companies and organizations in the financial and even political sectors allows access to valuable intelligence in these fields. By contrast, the intelligence value for immediate use in attacking companies providing critical infrastructures and communications services is usually relatively low. Rather, gaining access, if only to some providers of communications and internet services in the West and the United States, is liable to give attackers the ability to damage these services.

China's Cyberspace Warfare Strategy

China's strategy of cyberspace warfare was formulated in the previous decade as part of a profound modernization process undertaken by the Chinese military. Based on the awareness that when it comes to kinetic warfare the Chinese armed forces are structurally inferior to the armed forces of the West, such as the United States military, the strategy reflects the understanding that in order to confront an enemy with technological superiority in the area of information flow, it is necessary to disrupt the enemy's access to this information. The approach involves dealing an opening blow comprising a cyber attack, an electronic attack, and a kinetic attack on the enemy's information web and military technology centers. Such a blow will lead to the creation of blind spots on the enemy's part, allowing Chinese forces to operate with greater efficiency.[10] The Chinese assumption is that by disrupting the flow of information it is possible to cause significant damage to the capabilities of a sophisticated enemy and gain an advantage in the early stages of a confrontation.

The strategy developed by China in the last decade sees integrated network operations[11] as a key platform for the field. The strategy is based on a combination of four types of operations:[12] attacks on computer networks; electronic warfare, including anti-electronic and anti-radar measures; computer network protection; and computer network exploitation.[13] One of the key components in the Chinese strategy is controlling the enemy's flow of information, on the operating assumption that China's enemies (especially Western nations, with an emphasis on the United States) are highly dependent on information flow-based technology. The assumption is that during a confrontation, the ability to damage the flow of information would allow China to attain an advantage in the physical battlefield. This integrated approach gives China interdisciplinary operational capabilities, allowing it to use force effectively to attack an enemy.

Selected publications have undertaken detailed analyses of the most important institutions in the Chinese military in terms of network operations.[14] This essay describes two of these central military bodies: the Third Bureau (in the General Staff of the People's Liberation Army), responsible for SIGINT, and the Fourth Bureau, responsible for ELINT and electronic warfare. The Third Bureau employs experts in many fields: technicians, computer experts, language experts, intelligence experts, and more. Indeed, several Western researchers have surmised that the manpower operating in the Third Bureau numbers over 130,000 personnel.[15] The vast scope of the bureau's activity and the range of missions with which it is charged make it eminently fit to carry out cyber operations on the web. This bureau has many "collection stations" throughout China; it is responsible for gathering intelligence from voice and related data, and fully processing and assessing it. The department is also apparently responsible for internal intelligence gathering in the Chinese military for the purpose of internal information security and protection. The Fourth Bureau, responsible for ELINT, i.e., electronic intelligence operations and electronic warfare, seems to operate also in the field of integrated network operations.[16] It appears that the Third Bureau is the body coordinating overall activity in this field.

In addition to the military organization, China also has a very large hacker community,[17] including hackers who have claimed responsibility for a number of cyber attacks and are apparently involved in operations driven by national goals. Although the Chinese government presumably takes steps to enforce Chinese law, which prohibits this type of activity, it often turns a blind eye to the phenomenon and even provides material support for some of it, in a type of outsourcing of government cyber activity.[18] In addition, the Chinese army recruits civilians – from the hacker community and hi-tech industry – to its web militia units.[19] The web militia is integrated with the regular military, though its members are unpaid volunteers.

In contrast to the common perception of Chinese cyber activities, some researchers claim that these activities are designed first and foremost for internal needs, and that Western nations need not be overly concerned about the threat to their cyberspace. In this view, the Chinese have developed capabilities primarily to monitor opponents to the regime and control information available to Chinese citizens, essentially for political needs largely directed at preserving the regime.[20] However, while totalitarian regimes, including China, indeed use cyberspace capabilities for internal political ends,[21] this is only part of the picture, as evidenced by the series of cyberspace incidents emanating from China in recent years.

One of the main components of China's cyberspace strategy is the critical need for access to enemy communications infrastructures; without this access it is difficult to plant powerful blind spots. Attaining effective access to communications networks requires extensive and long term work on infrastructures. An attack on enemy communications networks is possible only if there is regular access to them over time, providing attackers with high quality intelligence that allows them secretly to install malware for use when the time comes. Such access requires long term maintenance and preservation because of the constant changes enemies make in their communications and information set-ups, and because they continually install new defensive systems designed to uncover malicious activity.

China's Cyber Attacks

The last six years have seen more than a few cyberspace attacks attributed to China, which apparently were intelligence gathering operations. An analysis of these attacks affords a means to identify China's basic attack techniques and infer its policy and methods. The attacks portray a world power intent not on focusing on a specific target, rather on gaining wide infrastructure access. In the case of Operation Aurora, the goal was to gain access to Google's password mechanism and the versions control software. In the RSA attack, the goal was to gain access to the internal network in which all information relating to secure ID was managed; such access could in the future be used to mount a more effective attack on other companies using the system, including security companies and companies engaged in sensitive activity.

The techniques identified in the well organized attacks were highly similar, using social engineering,[22] exploiting software weaknesses, and inserting delay mechanisms to expand intra-organizational access and extract information. The fact that China has taken these measures in a consistent, systematic manner over the past several years strengthens the assertion that the attacks were designed deliberately and that the same organizations were responsible, and weakens the claim that the attacks were the work of random hackers. Further substantiation may be found in the analysis made by the Northrop Grumman Corporation,[23] which noted several criteria:

a. *Similarity in keyboard behavior*. Similar behavioral characteristics or patterns in the attackers' methods in the various attacks were identified, e.g., attacking similar information parts and using similar tools.

b. *Scope of preliminary preparations*. The attacks comprised actions requiring preparation and prior knowledge, stemming apparently from preliminary action taken over several months before the actual attack. For example, familiarity with the architecture of the attacked networks was clearly evident.
c. *Attacker discipline*. The attackers were highly disciplined, e.g., they did not open files to scan the contents initially before copying them, indicative of the probability that they were operating on the basis of prior information.

Operation Nitro

Operation Nitro involved a series of attacks that occurred primarily from late July 2009 until mid-September 2009, when Symantec published information about it.[24] Its main purpose, likely technological espionage, was carried out in several consecutive waves, distinguishable by their targets. At first, human rights organizations in China were attacked, followed by motor industries; in the final stage, 29 chemical companies were targeted. The targeted companies were Fortune 100 companies working in chemical R&D and special materials for application in military vehicles and companies involved in the construction of infrastructures for chemical industries and the manufacturing of advanced materials. The attack method was similar to the method used in other attacks launched by the Chinese and included the following components:

a. Malicious code usually disguised as a security update. A great deal of non-personalized email was sent to organizations, unlike other operations in which great efforts were made to direct the email to individual email addresses.
b. Insertion of a back door (Trojan horse) into the targeted computers.
c. Increased access to the networks attacked while using remnants of passwords found on the attacked computers in order to gain control of central network computers.
d. Collection of material on interim servers and dispatch of this material outside the network.

In all, some 100 computers were attacked, 29 in the chemicals field and 19 belonging to the security sector. Most of the companies attacked were in the United States (about 30 percent), Bangladesh (about 20 percent), and the United Kingdom (15 percent), with the remaining located in some 20 different states around the world.

Operation Aurora

Operation Aurora included a series of attacks beginning in mid 2009 and continuing until December of that year. In January 2010, Google was the first to report it. The company announced that the attackers had hacked into Gmail accounts belonging to Chinese dissidents active in the United States, Europe, and China.[25] Adobe also reported attacks in the same operation, which targeted at least 34 organizations and companies.[26] McAfee, the information security company, analyzed the attacks. The findings indicated that the purpose of the attacks was to gain access to source codes of the attacked companies, especially the version management software Periscope used by hundreds of large software companies. McAfee discerned several stages in the attack:[27]

a. The operators of the attacked computer would receive a harmless-looking email or notification from what appeared to be a safe source.
b. The operator would take the bait and click on the link attached to the notification leading to a server containing malware.
c. The web browser in the attacked computer would download a binary code camouflaged inside a picture file and operate a back door that would connect to a control server located in Taiwan.
d. As a result, the attackers would gain full control of the computer and thus also to sensitive information communicated through the network.

This method was widely used in many of the attacks known as APTs (advanced persistent threats). At first, the term indicated sophisticated attacks on military and government networks, but currently the term is used to mean attacks of high intensity (i.e., state-level intensity) on a civilian target.

The Night Dragon and Shady RAT Attacks

These waves of attacks started in mid 2006 and continued until February 2011. McAfee, which gained access to one control server used by the attackers, identified the server after a log file analysis[28] and determined that some 70 targets had been attacked.[29] Given that McAfee gained access to only one control server, the attack presumably targeted many others as well. The analysis mapped the companies attacked and the time frames that the computers were controlled by a server through which the attackers extracted sensitive information. The targets included: 21 government organizations, 6 industrial and energy companies, 13 communication, computer, and electronics companies, 13 security companies, and 6 financial companies. In this context, the attacks on the Norwegian oil and gas companies are particularly noteworthy.[30] Attacks on companies considered national infrastructures, such as energy companies, could be evidence of the desire to create access for the purpose of damaging them at some point in the future.

RSA Attack

The RSA attack provides the basis for an in-depth analysis because one of the servers involved was a botnet[31] of some 2,000 computers. Penetrating the botnet's central server made it possible to analyze the list of infected computers; the analysis generated a list of 763 companies.[32] The attack was first reported by RSA in March 2011.[33] The stages of the attack, typical of other attacks as well, can be charted as follows:

Extensive infrastructure intelligence gathering →	Constructing the profile of the attacked computer's owner →	Sending email to attacked computer's owner →
Installing a back door in the computer →	Gathering initial information and expanding the attack →	Extensive information gathering

The first stage involves extensive gathering of infrastructure intelligence about the organization targeted. This intelligence is usually gathered from social networks and other open sources. The purpose of the information is to identify potential individual targets, as they will serve as the optimal channels to work within the attacked organization. For example, in the RSA attack, two small groups of employees were selected. They were not necessarily the final targets of the attack but were apparently selected because the attackers felt it would be convenient to start the attack with them.

The next stage involves constructing the profile of the attacked computers' owners: after identifying the penetration points, a profile of those to be attacked is constructed. This requires constructing a full enough picture that allows for the creation of an ostensibly harmless email that would not arouse any suspicion on the target's part. Such information gathering and the construction of a suitable profile require widespread, focused information gathering based on good organizational skills and resources (and especially English language skills).

This is followed by sending malicious email especially adapted to the attacked computer's owner (ZeroDate spear phishing email), which requires two steps. The first entails constructing a formula, structure, and look of a harmless message that would not immediately be erased by the user and would in fact prompt the user to open its links. Email is sent to specific groups of selected employees. At times the message is adapted to every individual user according to the profile constructed. The second action is including an attachment to the email with a security weakness and back door. Weaknesses are software security breaches through which attackers can insert their malicious code. At times the weakness is original, identified in the attacker's weakness identification process (apparently the case with Aurora); at other times, the weakness is well known (ZeroDate) and the attacker relies on the possibility that the targeted computer has not yet installed the patches to fix the weakness.[34] For example, in the RSA attack, the subject line of the email was "Recruitment Plan 2011" and had an Excel document attached, "Recruitment Plan 2011.xls." The ZeroDate weakness was CVE-0609-2011 in Adobe Flash. The moment one of the employees opened the file, the computer was infected via a back door. During the attack the weakness was considered unknown and there was no security update. The update was distributed about a week after the attack.

Installing a back door in the computer: Malicious code is inserted into the infected computer, which allows attackers to control it via a control server.[35] Usually back doors link the attacked computer to the attacker's server, and from there the computer is operated according to instructions from that server based on the commands of the human operators, usually working in shifts. This direction of communication – from within to outside the organization – makes it very difficult to identify the communication.

At this point the attackers gather initial information. Every attacked computer is matched with an attacker group analyzing the computer's contents and trying to assess how to gather information from the attacked computer and what information to gather. At this stage there is usually an assessment of the attacked computer's access to servers and other sources of information within the organization in order to identify the network map and learn how to expand the attack.

The central information gathering stage takes place after access to the company's servers has been gained and the desired information identified. The transfer of large amounts of information in a way that does not arouse suspicion and does not allow identification by monitoring software usually installed by large organizations is highly complex. It is generally done by means of another computer in the network whose access and permissions levels are high enough so that it upgrades the permissions of the servers to export information while using information-compressing encryption and algorithms. For example, in the case of RSA, the attackers finally arrived at a computer that stored sensitive information about the secure ID system, which later allowed the attackers access to information at other companies,[36] all of this bypassing the monitoring systems' warnings about illegal actions.[37]

The approach described herein requires the allocation of many professional resources. It seems that two groups working in tandem with different tools participated in this attack. The first identified the targeted information in the company's network, while the second worked separately to manufacture the channel for extracting the information. A third group, designated to preserve access for later use in the future, may also have been involved. Such an approach reflects the thinking of a world power working with a very high degree of professionalism while investing heavily in resources, such as highly skilled manpower and intelligence capabilities. Indeed, in this attack it is possible to discern some elements suggesting that a world power – presumably China – was behind it. These elements include:

a. *Infrastructure access*: Breaking into a company's one-time password mechanism (OTP) in order to gain access to other companies indicates a desire for extensive action requiring major resources.
b. *Scope of attack*: Open publications reported 763 infected computers found on one of the servers involved in the RSA attack. At least some of the targets required preliminary manual action, i.e., it was necessary to gather preliminary data about the target, construct emails in English that served as bait, and conduct a preliminary analysis of accessibility. An attack of such intensity would have required the organization of infrastructures at the level of a world power, indicating that this was not the work of individual hackers.
c. *The Sykipot back door program*:[38] This program, a variant of PoisonIvy, served Chinese attacks since 2006 (in similar versions) and through

early 2012.[39] The use of similar software (with relatively few changes) indicates organizational coordination among the various attackers over the last several years.

d. *Identifying marks*: The back door programs had strong links to China. According to an analysis of the software text, there were clear markers for the Chinese language, including remnants of information in Chinese in binary code (debug information). In addition, error messages in Chinese were identified. Finally, the only user's guide for the back door is in Chinese.

e. *The control servers*: An analysis of the sites where the control servers were placed and from where the attacked computers were controlled showed that most of them were located in China (299 of the 329 control servers).[40]

These findings strengthen the hypothesis that China is behind attacks requiring an extensive, systematic organizational and infrastructure system. Given this, one should not be surprised by the announcement made by General Keith Alexander, the Director of the NSA, which confirmed that China was behind the RSA attack.[41]

The list of 763 companied appearing on one of the servers involved in the RSA attack was analyzed. The analysis included identifying the companies through the internet and characterizing their activities according to three categories: technology companies apparently attacked for the purpose of technological espionage; financial and economic companies that would yield commercial information; and communications providers. These findings usually mean that the infected computer was linked to a public internet service provider (ISP). The analysis showed that close to 80 percent of the companies and organizations attacked were communications providers, while the other 20 percent were split between technological, financial, and other companies. The data indicates a typical botnet breakdown, which includes a very large number of infected computers belonging to private individuals who connected to the internet using an ISP. The rest of the attacked computers were distributed among some 90 countries, including five in Israel.

Concluding Insights

The series of attacks since 2006 indicate a transition to attacking critical infrastructures, both in the communications and energy fields. Regarding the RSA attack, it is possible that the list of companies on the server included a random botnet list compiled by the Chinese in a lengthy process before the attack was discovered in order to serve as an infrastructure for future attacks. It is possible to send attack email from every infected computer, transfer files, and hide the attacker's identity. However, it is also possible that some of the list is not random and includes companies that are explicitly targeted for attack.

The findings about the attacks in recent years strengthen the research hypothesis that the attacks described are part of a systematic, orderly campaign underway by China. China's cyberspace warfare strategy suits the choice of some of the attack targets, most of all those connected to critical infrastructures. The attack against Google in Aurora, the Shady RAT attacks, and especially the RSA attacks all signal a transition to a systemic approach that targets communications and critical infrastructures. China's strategy, designed to damage the enemy's weaker and lesser-protected realms in a move prior to using kinetic force, requires extensive activity to create long term access to critical infrastructures, including communications. Unlike normally noisy information gathering operations discovered from time to time, it is more difficult to discover operations aimed at infrastructures and gaining access to them for use at some time in the future. It is quite possible that they will never be discovered.

In addition to the attacks discussed above, in April 2011 China was accused of intercepting no less than 15 percent of all internet traffic.[42] Therefore, this activity is likely part of attacks designed to create intelligence access to internet traffic and intercept transmissions before they are encrypted. Moreover, the conclusions of this essay are based on knowledge accrued as the result of analysis of information about attacks that were discovered and publicized. Because some attacks are not discovered and others are discovered but not publicized, one may assume that China is running other cyberspace operations. It is hard to know what exactly is taking place at the companies under attack. One possibility is that they have been fitted with back doors different from the ones used to preserve access and that this back door will be put into action at the attackers' discretion in order to damage the relevant communications infrastructure. Moreover, a sleeper back door is virtually undetectable by existing defensive technologies such as various anti-virus programs.[43]

This is particularly serious with regard to the United States, where there tends not to be a physical separation of communications networks. In other words, the so-called civilian internet[44] is also frequently used in the computer systems of sensitive installations and organizations, and even critical national infrastructures such as electricity producing nuclear reactors and transportation infrastructure control systems. Furthermore, in some cases the United States security systems make extensive use of civilian internet infrastructures, and the separation of networks of sensitive operational systems is not sufficiently developed. This is an essential security weakness allowing attackers a great deal of access to these infrastructures by means of attacking less protected civilian systems. This means the creation of the ability to severely disrupt information transmission at some unspecified future date. Because of this weakness, preliminary damage to communications and telephony infrastructures during a confrontation is liable to disrupt operational and security systems based on these infrastructures.

The response to this weakness requires adopting a comprehensive systemic approach. Attempts to improve the defenses of communications infrastructure providers are insufficient to prevent future attacks. The use of the internet for communications of sensitive systems cannot be based solely on access permissions. No matter how protected, these permissions represent a severe security breach. One of the important components of a response to the weakness described herein lies in differentiated communications networks. It seems advisable to isolate operational networks of the whole gamut of critical systems, such as security systems, operational communications systems, and command and control systems of installations identified as critical national infrastructures. The ability to operate control systems of critical installations through the internet is liable to prove to be a serious problem the moment a sophisticated attacker decides to use back doors at some future point.

Notes

[1] Steve DeWeese, *Capability of the People's Republic of China to Conduct Cyber Warfare and Computer Network Exploitation*, Northrop Grumman, October 9, 2009, p. 67.

[2] Kenneth Lieberthal and Peter W. Singer, *Cybersecurity and U.S.-China Relations*, The Brookings Institution, February 2012.

3 On the attack on Mitsubishi Ltd. in Japan in August 2011, see Hiroko Tabuchi, "U.S. Expresses Concern about New Cyberattacks in Japan," *New York Times*, September 21, 2011, http://www.nytimes.com/2011/09/22/world/asia/us-expresses-concern-over-cyberattacks-in-japan.html?_r.

4 "Chinese Hacked French Ministry for G20 Data," *The Week*, March 8, 2011, http://www.theweek.co.uk/technology/7229/chinese-%E2%80%98hacked-french-ministry-g20-data%E2%80%99.

5 Erik Helin, "Fingers Point to China in Australian Prime Minister Hack," *Brick House Security,* March 30, 2011, http://blog.brickhousesecurity.com/2011/03/30/australia-pm-hack.

6 On the attack on Canadian government sites, see Greg Weston, "Hackers Attack Canadian Government**,"** *CBS News,* February 16, 2011, http://www.cbc.ca/news/politics/story/2011/02/16/pol-weston-hacking.html.

7 John Markoff and David Sanger, "IMF Reports Cyberattack Led to 'Very Major Breach,'" *New York Times,* June 11, 2011, http://www.nytimes.com/2011/06/12/world/12imf.html.

8 Nathan Thornburgh, "Inside the Chinese Hack Attack," *Time US*, August 25, 2005, http://www.time.com/time/nation/article/0,8599,1098371,00.html.

9 Dimitri Alperovitch, "Revealed: Operation Shady RAT," Version 1.1, McAfee, 2011, http://blogs.mcafee.com/mcafee-labs/revealed-operation-shady-rat.

10 DeWeese, *Copability of the People's Republic of China to Conduct Cyber Warfare*, p. 69.

11 Integrated network electronic warfare.

12 Tim Stevens, "Breaching Protocol: The Threat of Cyberespionage," *Jane's Intelligence Review*, March 2010, pp. 8-13.

13 Timothy L. Thomas, "Chinese and American Network Centric Warfare," *Joint Forces Quarterly* 38, p. 77, http://www.dtic.mil/doctrine/jel/jfq_pubs/1538.pdf.

14 DeWeese, *Copability of the People's Republic of China to Conduct Cyber Warfare*, p.31; Mark A. Stoke, Janny Lin, and L. C. Russell Hsiao, *The Chinese PLA Signal Intelligence and Cyber Reconnaissance Infrastructure,* Project 2049 Institute, 11, 2011, pp. 6-14.

15 It is difficult to verify this assessment.

16 James Mulvenon, "PLA Computer Network Operations: Scenarios, Doctrine, Organizations, and Capability," in Roy Kamphausen, David Lai, and Andrew Scobell, eds., *Beyond the Strait: PLA Missions Other than Taiwan* (Washington, DC: National Bureau of Research, 2009), p. 273.

17 In Mandarin: Hikè 黑客, literally "black guest."

18 Stevens, "Breaching Protocol," pp. 8-13.

19 Timothy L. Thomas, "Comparing US, Russian and Chinese Information Operations Concepts," *Foreign Military Studies Office*, Fort Leavenworth, KS 66048, February 2004, pp. 12-13.

20 Thomas Rid, "Think Again: Cyberwar," *Foreign Policy*, March/April 2012, http://www.foreignpolicy.com/articles/2012/02/27/cyberwar?page0,6.

21 See publications on China's cyberspace espionage against the Tibetan government in exile and the break-in of the Dalai Lama's computer infrastructure; Stevens, "Breaching Protocol," pp. 8-13.

22 In the context of this essay, this term denotes the ability to deceive the owner of the computer under attack by creating a posture that fits the user's profile so that the computer will take action that interests the attacker, e.g., respond to email addressed to the owner in a way that is contrary to the security policy of the organization in which s/he works.

23 DeWeese, *Copability of the People's Republic of China to Conduct Cyber Warfare,* p. 60.

24 Eric Chien and Gavin O'Gorman, *The Nitro Attacks, Stealing Secrets from the Chemical Industry,* Symantec Security Respond, 2011, www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/the_nitro_attacks.pdf.

25 It is possible that there was no connection between the hacking of the Gmail accounts of individuals and the attack designed to access the Google and Adobe source codes.

26 Ariana Eunjung Cha and Ellen Nakashima, "Google China Cyberattack Part of Vast Espionage Campaign, Experts Say," *Washington Post*, January 14, 2010, http://www.washingtonpost.com/wp-dyn/content/article/2010/01/13/AR2010011300359.html.

27 McAfee Labs and McAfee Foundstone Professional Services, *Protecting Your Critical Assets, Lessons Learned from "Operation Aurora,"* http://www.mcafee.com/us/resources/white-papers/wp-protecting-critical-assets.pdf.

28 Log files are files that continuously and automatically document defined computer activity.

29 Alperovitch, "Revealed," p. 3.

30 "Hackers Attack Norway's Oil, Gas and Defence Businesses," *BBC News*, November 18, 2011, http://www.bbc.co.uk/news/technology15790082-.

31 A botnet is a collection of software agents installed on host computers. In many cases these are infected computers that contracted the software agent without the computer owner's knowledge. The software agents can be operated under previously defined conditions or by commands coming from a control server.

32 Brian Kerbs, "Who Else Was Hit by the RSA Attackers," October 2011, http://krebsonsecurity.com/2011/10/who-else-was-hit-by-the-rsa-attackers.

33 Uri Rivner, "Anatomy of an Attack," April 1, 2011, http://blogs.rsa.com/rivner/anatomy-of-an-attack.

34 ZeroDate weaknesses are software security breaches publicly identified and noted. Usually, as soon as the breach becomes known, the software developer provides a response in the form of a security patch distributed to the public. There is generally a gap between the time the patch is distributed and the time it is actually installed on users' computers. The window of opportunity for attackers starts when the weakness is announced and lasts

until the patch is installed on the targeted computer. During this timeframe, attackers can insert malicious code through the breach.

35 Around November 2010, some of the computers of the companies under attack were already in communication with the attackers' control networks.

36 One of the companies attacked using information gathered in the RSA attack was Lockheed Martin. See Mathew J. Schwartz, "Lockheed Martin Suffers Massive Cyberattack," *Information Week*, May 31, 2011, http://www.informationweek.com/news/government/security229700151.

37 Large organizations usually have systems that monitor computer network traffic in order to identify behavior that is illegal according to predetermined rules. Such systems have different commonly used names, including SEIM (security event and information management) and NBA (network behavior analysis). These programs have a set of rules designed to alert administrators to non-permitted or unusual network behavior and also to prevent it from occurring.

38 Stephen Doherty et al., "The Sykipot Attacks," December 14, 2011, http://www.symantec.com/connect/blogs/sykipot-attacks.

39 Mathew J. Schwartz, "More Sykipot Malware Clues Point to China," *Information Week*, December 21, 2011, http://www.informationweek.com/news/security/attacks232300940/.

40 Kerbs, "Who Else Was Hit by the RSA Attackers."

41 Nicholas Hoover, "NSA Chief: China behind RSA Attacks," *Information Week*, March 27, 2012, http://www.informationweek.com/news/government/security232700341/.

42 Stew Magnuson, "Cyber Experts Have Proof that China hs Hijacked U.S.-Based Internet Traffic," *National Defense,* December 11, 2010, http://www.nationaldefensemagazine.org/blog/Lists/Posts/Post.aspx?ID249=.

43 Gunter Ollmann, *Serial Variant Evasion Tactics Techniques Used to Automatically Bypass Antivirus Technologies*, Damballa, 2009,http://www.damballa.com/downloads/r_pubs/WP_SerialVariantEvasionTactics.pdf.

44 The concept of "civilian internet" denotes internet communications networks used by the public at large and having no particular protection.

Cybercrime: A National Security Issue?

Lior Tabansky

Cyberspace, an offshoot of the development of computer and digital communications technologies, has in recent decades become part and parcel of our lives. Computerization is invaluable in improving and streamlining processes related to work, learning, and entertainment, and it affects virtually every field of human endeavor. Once the internet became commercial in 1988, it quickly turned into a mainstay of cyberspace, offering inexpensive and immediate access to many sources of information, information sharing, joint long distance work, and more.

The implications of cyberspace crime for national security derive from the way technology is used by hostile elements. This article proposes a policy directed examination of the meaning of cyberspace crime and its impact on national security, without focusing on the widespread monetary assessments of the damage caused by cybercrime. It includes a profile of cooperation among criminals, organized crime, and hostile organizations, and discusses the commercialization of cyber reconnaissance and cyber attack capabilities, made possible by ever-developing technologies and the growth of a black market in IT services. Currently, cybercrime is hardly significant beyond the realms of IT risk management and law enforcement. However, this article identifies two separate conditions where cybercrime could become a substantial threat to national security.

Public demand for cyber security rises in proportion to the growing recognition of the menace. Even in the absence of an objective increase in the scope of crime, this demand is not expected to decrease. The state's responsibility to provide security to its citizens cannot stop at the threshold

Lior Tabansky, a former Neubauer research fellow at INSS, is a doctoral student in the Department of Political Science at Tel Aviv University.

This article was first published in *Military and Strategic Affairs* 4, no. 3 (2012): 117-36.

of cyberspace, and in this realm too the practical expressions of such responsibility must be defined as part of a democratic political process on a firm factual basis.

The Cybercrime Phenomenon

Computerization allows tasks to be broken down into small units and decentralizes processing; networking allows global access to information and focus on knowledge as a valuable product. Computerized technologies are implemented to change and enhance the efficiency of creative and working processes in every aspect of life, and the world of crime is no exception. The proposed definition of cybercrime is: "The use of cyberspace for illegal ends, while exploiting unique cyberspace features, such as speed and immediacy; remote operation; encryption and obfuscation, making it difficult to identify the operation and the operator."

The debate on cybercrime continues. Over a decade ago, Grabovsky wondered what was new about cybercrime, whether it was not merely an old phenomenon making use of new tools.[1] But most researchers try to analyze cybercrime as a unique phenomenon. Majid Yar categorizes it according to the object targeted: property, people, or the state.[2] Shinder and Cross distinguish between types of crime according to the level of violence involved: violent and potentially violent crime, non-violent crime (drug trade, money laundering), and crime (still) perceived to fall within the white collar category (computer break-ins, theft, and fraud).[3] According to Wall, cybercrime is "the transformation of criminal or harmful behaviour by networked technology,"[4] i.e., it developed as a result of the evolution of computerization and cyberspace and consequent new opportunities to attain, disrupt, or manipulate information for gain. Wall further classifies cybercrime into three categories: crime involving the integrity and good working order of computer systems (hacking); crime making use of cyberspace (encrypted communications among criminals, the sale of counterfeit pharmaceuticals); and crime involving computerized information contents (theft of secrets, dissemination of harmful contents).

Table 1 categorizes crime on the basis of the role played by the computer in the commission of the crime,[5] a position similar to that adopted by the European Convention on Cybercrime.[6]

Table 1. The Computer in Cybercrime

The computer as a tool in the commission of crime

Access to and dissemination of contents	Malicious disruption or modification of data	Use of communications
• Secrets • Knowledge/data • Harmful contents	• Identity theft • Fraud • Sabotage	• Harassment • Trade in forbidden materials • Spam

The computer as a target of crime

Unauthorized access	Inserting malicious code	Disruption of operation	Theft of service
• Hacking	• Malware, spyware, viruses	• Distributed denial of service (DDoS)	• Unauthorized use

There is nothing unique or new in much of cybercrime – harassment, fraud, unlawful propaganda, pornography, theft, money laundering, espionage, and so on – except the use of cyberspace. But there is another level of crime that could not exist without cyberspace: spam, click fraud, various types of malware, networks of captive computers (botnets),[7] digital identity theft, camouflage and encryption[8] of data and communications, computerized breaches of highly valuable secure facilities, and automatic, long term espionage in secure organizations, depriving them of control of intellectual property. Cyber criminals are exploiting the increasing value of digital data in all its forms, and the legal and judicial ways in which different countries handle cyberspace.

Crime has always been a widespread social phenomenon. Criminological explanations combine motivation, opportunity, and the existence of a "guarding" factor. Two different sources of human motivation can be identified.[9] Many motives for criminal behavior are intrinsic and are not determined through a cost benefit analysis. There is no reason to believe that greater use of one technology or another would change human behavior. It is therefore not surprising that people also use cyberspace to realize their needs and pursue their goals in legitimate activities – study, entertainment, education, work – as well as in the age-old human pursuits of warfare and crime.

The classic doctrine of criminology is based on the concept of free choice and a rational assessment of anticipated gain versus the risk of punishment; accordingly, the motivation for committing a crime is a rational economic decision.[10] Economists and psychologists analyze human behavior, including criminal behavior, as a derivative of a rational cost-benefit analysis. The ever-changing array of external circumstances may encourage cybercrime; this happens when someone identifies a growth in potential gain and estimates the cost – the risk of punishment – as being lower than that gain. The combination of greater digital connectivity in its current insecure form, and the increased value of computerized data results in a situation in which extrinsic motivation for criminal behavior rises.

Although developed nations have instituted regulated law enforcement mechanisms, state responses have not kept up with the pace of technological changes in cyberspace. A good example is the "traditional" bank heist as compared to cyber theft. In a traditional bank robbery security arrangements must be subdued as the chance of a confrontation with armed guards is likely. Even if the robbery itself is successful, the authorities will pursue the robbers for years to come. As cyberspace has developed, the exploitation of its vulnerability has also come to encompass bank robbery. For example, the use of botnets comprising tens of thousands of personal computers[11] for extended theft of identification details to banking sites, which are then used to steal small amounts of money, is quite common. Given the attribution problem in cyberspace, the chances of identifying the criminal are slim.[12] Financial institutions are well aware of the risk to their business interests and, together with regulatory bodies, are taking steps to protect themselves, investing in IT security to minimize the scope of opportunity available to cybercriminals. But even so, the immediate physical risk is still substantially lower for the cyber thief than it is for the "traditional" thief. The risk of legal punishment is lower as well, since cyber fraud is generally perceived by the judicial system as a non-violent "white collar" offense and treated accordingly.

The Scope of Cybercrime and Subsequent Damage: Problematic Assessments

The cybercrime phenomenon is usually examined from a variety of perspectives: legal (legislation and penalties), criminological (motivation and organization), economic (incentives and value), or technical (data security). Jurists deal with setting the limits of what constitutes acceptable behavior and legal issues of prevention and enforcement. Criminologists apply their professional knowledge to understanding new phenomena. Economists describe the set of incentives affecting decision making by rational players. And data security experts deal with the technical aspects of technological infrastructures – software, hardware, and communications – while focusing on various vulnerabilities and ways to protect them. One thing that jurists, economists, and data security experts all agree on is that the scope and impact of cybercrime are constantly and rapidly on the rise. This assessment is based on the fact that the scope of digital data is increasing exponentially, as is connectivity between computerized facilities. Cyberspace contains more information with more potential access points for unauthorized breaches. The ordinary conclusion is that every breach exposes a growing scope of data.

Financial estimates of the scope of damage resulting from cybercrime have been issued since the 1990s, with security companies spearheading research into the subject and publishing numerous reports. There are dozens of different assessments emanating from the commercial and government sectors in the United States, England, and other developed nations.[13] An FBI report estimated damage to American business in 2005 at $65 billion.[14] In 2009, US Secretary of Commerce Gary Locke claimed that annual damage to American companies as a result of counterfeiting and piracy (i.e., illegal use of computer codes) was in the neighborhood of $200-250 billion.[15] A 2011 British report put damage at 27 billion pounds annually: the damage per annum to British citizens was estimated at 3.1 billion pounds, to the business sector at 21 billion pounds, and to the government at 2.2 billion pounds.[16] A recent report by Symantec, a leading global computer security software provider, estimated the direct damage caused by cybercrime at $114 billion annually in 24 nations.[17] Other estimates speak of hundreds of billions of dollars annually.[18]

These astronomical sums have raised question marks and doubts, but to date the impact of the criticism has been limited. Recently, two researchers at Microsoft published a position paper criticizing the shaky statistical infrastructure underlying assessments of cybercrime damage, which is typically estimated by surveys.[19] How have these estimates actually been carried out? An examination of research methods reveals how easy it is to produce inflated damage assessments. First of all, there is no information about the use made (or not made) of data that was accessed. Those incidents where firm knowledge exists are few, whereas the scope of potential damage is huge. Let us assume that a PC storing a database of one thousand entries is breached; let us also assume that the database is not encrypted and the entries are written in plain text. Every entry represents a valid credit card, including all the information needed to use it: the number, CVC code,[20] expiry date, full name, ID number, and address of the cardholder, as well as the card issuer's bank information. In this scenario the thief sees a complete and real picture of the information on file. Yet even under these optimal circumstances, are we able to fully estimate the financial value of the information accessed? Can the thief properly assess the true value of the stolen information? Can the victim do so?

When it comes to the theft of intellectual property – the product of long research and development efforts – the victim tends to identify as damage the maximum profit he would have liked to make on completion of the R&D, manufacturing, and marketing process. Surveys, which are an appropriate method for clarifying hard-to-observe phenomena, are the main method of learning about the scope of damage. Surveys allow researchers to reach a larger, more diverse group of respondents providing their own estimates of the number of incidents and the damage, but they are also a method containing some serious drawbacks that concern social scientists and statisticians.[21] Secondly, in the absence of sufficient data, researchers use statistical methods to derive assessments from partial data.

Measurement problems affect every aspect of the debate on cyberspace threats, particularly attempts to help the discussion by quantifying damage in monetary terms. There is an inherent difficulty in estimating damage and so far it seems that monetary assessments – created by a crude use of statistical methods to present suppositions on the basis of insufficient data – are inclined to be inflated. In addition to questions of reliability of the research methods, the credibility of sources of information and the suitability of the statistical method to this type of research, there is also another problem. Monetary estimates often include indirect components of damage: whether to the reputation of the victimized organization, negative impact on consumer behavior with macro-economic implications, issues of torts, insurance, attendant expenses, or others.

Some questions central to understanding the phenomenon remain unanswered. Does it make sense to assess damage on the basis of use actually made of the stolen information rather than maximum potential use? Perhaps it makes sense to relate to the monetary value of creating information instead of assessing its market value, present or future? And what about the cost of security and a return to normal functioning? The picture obtained from the usual sources is less than credible and the damage of inflated assessments is liable to result in a counter response of failing to take the power of cybercrime seriously enough. Basing the cybercrime debate on estimates of monetary damage detracts from a rational, intelligent, and informed debate on the problem and the ability to formulate appropriate public policy.

Cooperation between Criminals and Terrorist Organizations

The interface between professional criminals and organized crime on the one hand, and terrorist organizations on the other, is likewise not a new phenomenon. Even if we look only at the Israeli reality, we can see that such cooperation causes damage at the national level. Since 1996, the media campaign over pirated CDs has claimed that profits are used to fund Palestinian terrorism,[22] as part of a close connection between money laundering and its consumers such as terrorist organizations.[23] The widespread phenomenon of auto theft from Israel by West Bank thieves has been a feature of life in Israel for many years: the problem has hardly been confronted at national level because the threat was never considered to be a national security issue; the damage was covered by the insurance companies, which rolled it over onto the insured parties; the police took no action outside of sovereign Israeli territory; and the army – operating permanent security checkpoints on major roads – preferred to avoid dealing with a criminal population whose motivation was merely monetary, rather than nationalistic. During the "suicide bombers intifada" years the modus operandi of these criminals changed: terrorist organizations recruited the expertise of Palestinian car thieves in order to obtain cars with Israeli license plates to reach their destinations, and also to find routes to evade security checks and deliver explosives and suicide bombers into the heart of Israel's cities.

The possibilities of crossing over the fenced Gaza Strip border were more limited than between the West Bank and Israel. Tunnels were dug towards the Rafiah Egyptian border crossing to provide various kinds of smuggling channels. Smuggling generates large profits for the tunnels operators and this activity persists despite Israel's efforts to put a stop to it. The tunnels also became a national security problem when they were used to smuggle weapons from the Sinai Peninsula to the Gaza Strip and terrorists from the Gaza Strip to Sinai.[24] It was the criminal organizations' expertise in digging tunnels that made the June 25, 2006 attack on Kerem Shalom possible, in which two soldiers were killed and a third was taken hostage by Hamas. This was a clear case of criminal technical know-how used to damage Israel's national security.

Some Bedouins in Sinai make a living from their expertise as guides and scouts, and have for decades provided smuggling services into Israel. The "goods" smuggled included, in the not too distant past, hundreds of East European women for the sex industry, as well as drugs. In recent years, tens of thousands of African migrant workers and some refugees have been guided to the Israeli border. Some believed these cases posed significant challenges but were not a national security issue. However, as the smugglers' expertise is increasingly applied to enable terrorist attacks on Israel, that assessment is changing.[25] The smuggling of terrorists from the Gaza Strip through Sinai to Israel made the August 18, 2011 attack on Route 12 possible, resulting in the killing of eight Israelis and the wounding of four. Smuggling terrorists and weapons has placed Eilat within rocket range.[26] Hence smuggling grew to become a clear and present danger to Israel's national security.

A Reexamination of the Meaning of Cybercrime

Any current examination of cybercrime reveals comparable commercial cooperation. In recent years a black market of technical experts and botnet "herders" has emerged, developing and providing technical tools and services for a price.[27] The black market of cyberspace services (Crimeware as a Service, or CaaS) causes economic damage in developed nations, though the usual monetary damage estimates are greatly exaggerated.

Anyone who prefers to operate alone and lacks R&D resources finds cyberspace weapons (toolkits of malicious software)[28] available for downloading from the internet, usually for payment of anywhere from tens to several thousands of dollars. Knowledge is an inexhaustible product, a "non-rival good" for economists, so sharing the capabilities that were available with others to you does not diminish your own strength.[29] As a result, we see a situation in which powerful tools are available to anyone at marginal cost. The widespread impression that cyberspace makes it easier to rake in huge profits from criminal enterprises has not been lost on organized crime.[30]

Growth in computing power and the ubiquitous internet have created a new tool for extensive cybercrime: the botnet. This is a collection of internet-connected PCs whose defenses have been breached by malware and control ceded to a malicious third party, who is able to remotely control and exploit these computers on demand, usually without disrupting their normal functioning. Cybercriminals usually infect internet-connected computers with malware by exploiting known vulnerabilities that users and system administrators have failed to deal with. In 2007, McAfee estimated that some 5 percent of all internet-connected personal computers were botnet captives.[31] Large scale supply makes the cost of using a botnet affordable to virtually anyone.[32]

A newer phenomenon is the advanced persistent threat (APT), also known as adaptive persistent attack (APA)[33] – a complex, multi-stage use of cyberspace weapons for the purpose of ongoing clandestine attacks. The attacker does not operate statistically on a broad scale to exploit known vulnerabilities; instead the objective is well defined. The attacker uses a range of custom made tools, often using a valuable "zero-day" (never used before) attack mechanism. Such attacks comprise several stages and can last months or even years. The attacker begins to gather intelligence about the organizational structure of the target, and identifies people holding senior positions with access permissions for sensitive information. The gathering of personal information is usually accomplished by open source intelligence (OSInt): accessing public information and shared personal information on social networks and the news media. Once the key players are identified, a concerted effort is undertaken to steal their credentials and infect their computers.

One method is spear phishing, or inserting a remote access tool (RAT) by an email from a trusted sender with relevant content, which thus manages to bypass spam filtering mechanisms by using the personal information gathered. Opening the email allows the insertion of the Trojan horse into a trusted endpoint inside the organization's corporate network, thus gaining access to more internal resources. In a common crime, once access is accomplished, the average attacker moves quickly to retrieve valuable information and use it.

However, this is not the case with an APA attack: here the purpose is clandestine long term access, ignoring immediate monetary temptations. The attack lasts a long time, in part to overcome defense systems designed to prevent information leaks. In the course of the attack, attackers perform tests to identify the system's response thresholds and usually adapt the exfiltration methods of the stolen information. The data is divided into small packages, camouflaged inside legitimate communications, and thus leaks through the system without triggering defenses. An APA is much rarer than statistical attacks because it is much more expensive, requiring systematic intelligence gathering, planning, and adapting capabilities and the patience to carry out a long term task. Correspondingly, the damage of an APA is of a different scale.[34]

From the economic perspective, in terms of supply, hacker groups that have succeeded in developing and using software tools to control tens of thousands of computers have in fact created a service of economic value. In terms of demand, various customers – other hackers, private investigators, criminals, espionage organizations, and transnational criminal organizations – have found various uses for the product. This has created the "Crimeware as a Service" (CaaS) model, the black market counterpart to "Software as a Service" (SaaS) which has served the IT industry since 2001.[35] Over the years the model has undergone several transformations; the current buzzword for it is "cloud computing." The economic justification of the model is clear: from now on, the customer no longer needs to buy computer equipment in order to use computer services; he can simply buy the specific service he needs from large operators and use it over standard communications. The scope of the global market for this type of computer service was estimated at $14.5 billion in 2012.[36]

Let us examine the black market phenomenon from the national security perspective. The existence of a black market of cyber weapons, outsourcing research and development, quality assurance services, and technical support means that the requisite level of technical skills to become a cyber criminal has dropped. No longer is it necessary to have the competence to develop tools and methods for breaching computers oneself. The technological infrastructure needed to breach and make unauthorized use of computers is the same, regardless of whether the breach is aimed at profit, sabotage, terrorism, or destruction.[37] This reveals another risk: the use of existing tools for terrorist activity and damaging critical infrastructures – rather than the expected fraud targets for theft and quick profits – threatens to damage national security. The continuing development of cybercrime mechanisms is therefore becoming a natural security problem.

Critical infrastructures protection (CIP) is the most important issue in cyberspace security, and the black market in cyber weapons makes the need for it even more acute. This commercialization of technical and operational capabilities allows access for many factors – including small terrorist organizations and even isolated individuals – to powerful resources with potential cyber attack application. The reference group of threats is therefore expanding beyond states and known terrorist organizations to include any element capable of purchasing commercial services available on *DarkMarket*. Nonetheless, when there is ongoing state-sponsored investment in R&D, the technological capabilities openly available on the market naturally lag behind those being developed by the security forces and a nation's institutions of higher education. Therefore the capabilities available on the market will be inferior to those accessible to state-sponsored organizations with independent R&D means, enjoying state backing in terms of resources and organization.

Towards Realizing the State's Responsibility for Cyber Security

The meaning of the cybercrime phenomenon needs to be clarified for researchers and policymakers. For the reasons stated above, monetary damage assessments do not provide a firm factual basis for understanding the concept or formulating policy. Therefore, a reassessment of cybercrime is required to design appropriate national policy.

Even in the absence of agreement on the scope of direct and indirect damage caused by cybercrime, it certainly affects how citizens, organizations, and society as a whole function. Citizens and small businesses are variously damaged by cybercrime. Spam, internet fraud, digital identity theft, invasion of privacy, blackmail, economic espionage, and damage to intellectual property all are widespread and harm some citizens and organizations. Although monetary assessments seem to be exaggerated, the development of cyberspace increases numbers of potential victims and expands even further ways of committing crimes against citizens and groups. Given rising awareness of the problem and the actual increase in cybercrime, citizens of developed countries will reasonably demand the state take steps to provide personal, communal, and national cyber security. Growing media exposure of data breaches and cyber attacks is indicative of a proportionate growth of interest in the risks posed by cybercrime.

The state is fundamentally responsible for law and order and for the safety of its citizens, and is required to act to minimize damage to them. Policy should develop on the basis of understanding the broad implications of the phenomenon and a rational, informed public debate. Below are some pointers for developing such a debate.

The majority of the common phenomena classified as cybercrime have nothing to do with national security. What, then, is the significance of spreading hatred and incitement against Jews or the State of Israel while defacing Israeli websites, disseminating propaganda by means of social media and spam, hijacking social networks accounts, and creating internet videos and campaigns offensive to the public? Citizens will be vulnerable in cyberspace and the dignity of the nation and many of its citizens will be subjected to slander and defamation. However, experience shows that the public is not easily shaken by such acts. Beyond the professional realm of public relations, the damage at the national level is negligible.

What is the significance of common fraud – digital identity theft and unauthorized use of means of payment information aimed at stealing from citizens? When a citizen becomes a crime victim, the state authorities are expected and required to address the crime and deal with it. The state authorities have a range of methods to this end and the meaning of the events needs to be clarified so as to determine the appropriate policy. But from the perspective of national security, it is hard to see damage at national level as long as the rate of cybercrime is relatively low, even if it is higher than the more conventional crime rate. If, however, cybercrime grows to become a lasting and widespread phenomenon, citizens might lose their faith in state authorities that seem unequal to providing a safe and secure environment.

The current situation in developed nations is far from satisfactory. If "obedience in exchange for protection" is the condensed version of the social contract between citizens and the sovereign, then in the cybercrime area the state is defaulting on its side of the contract. Response to the new challenges requires, first and foremost, a clear understanding of the different phenomena and their implications and ramifications. Response processes and the formulation and enforcement of policy require updated regulation and legislation. Legislation, which by definition lags behind technological developments, lies within the sole purview of the state. The sovereign enforcement bodies operating on the basis of national legal infrastructures will have to allocate more resources to the prevention, investigation, and punishment of cybercrime. Despite the international nature of cyberspace, the state is the sole source of responsibility for the personal security of its citizens. International treaties such as the European Council's Budapest Convention on Cybercrime[38] and initiatives being developed in the UN,[39] the OECD,[40] the EU,[41] and the International Telecom Union[42] are all boosting cooperation among sovereign authorities. International cooperation may contribute to arming sovereign authorities in the fight against cybercrime, but international treaties cannot substitute for independent sovereign policy.

First, cooperation among nations in the anarchic international arena is possible only to a very limited extent and only on the basis of common interests. It may be that developed democracies will be able to formulate arrangements among themselves, but the gap between them and authoritarian regimes in terms of defining the threat seems too great. The American debate on the issue focuses on ongoing industrial espionage of intellectual property, the product of R&D in the commercial and government sectors in the United States. Over the years, senior personnel in the business and government community have become increasingly concerned about the loss of America's global economic and strategic advantage as the leading scientific-technological innovator and superpower. In fact, "loss" is not the right word, because the knowledge is not actually lost, but rather stolen through systematic, well-organized and widespread state-sponsored theft, and the culprit is China, a nation determined to catapult its economic and military might forward by copying the secrets of American research.[43] Hence discussion of the issue clearly shifts from focusing on the economy, data security, and the law, to an almost combative security dialogue.[44] For its part, China rejects these allegations outright and is worried about undermining the foundations of its regime by use of the West's internet in the name of freedom of expression.

Second, the authority and sovereignty of a state within its borders allows that state to promote independent policy: legislation and law enforcement are not dependent on international arrangements. In Israel, an incident known as the "Saudi hacker affair" demonstrates how the debate spills over from data security into national security. In early January 2012, someone calling himself OxOmar published a list containing the personal information and credit card numbers of thousands of Israeli citizens.[45] The information published was overwhelmingly outdated, and out of 380,000 entries only a few thousand were valid. The direct damage to cardholders was zero: the credit companies cancelled the cards and issued new ones, and in any case the law obliges them to cover unauthorized use. The scope of the information revealed was also not exceptional: every day, millions of such entries are stolen on the internet. The details are bundled according to different parameters and sold as dumps[46] to black market customers, as described above.

It soon became clear this was a simple attack: spyware had been inserted into a number of commercial Israeli websites, which transferred data stored by the site operators with gross disregard for data security. Although the attack lacked complexity and no real damage was incurred by the Israeli citizenry, the extensive media coverage of it lasted some three weeks and was initially tinged with panic and hysteria. The event was presented as anti-Israeli terrorism, because instead of realizing monetary profits from the information, the attacker chose to use it to propagate fear in the target country.

This event can be analyzed in any number of different ways. One may claim that citizens are unaware of data security; that the media are irresponsible and blow a marginal event out of all proportion, sowing panic; that website owners were careless or even criminally negligent in failing to secure the data in their possession; that the state neglected to create a safe environment for internet commerce and secure personal data. But in any analysis, the inevitable conclusion is that the personal and collective security of Israel's citizens in cyberspace needs to be upgraded. At the end of the day, that demand is directed at the state, which is responsible for its citizens' security and safety.

It is possible, even desirable, to discuss the definition of unwanted and criminal phenomena in cyberspace, the proper level of security, the division of responsibility, heightened user awareness, the limits of state involvement, and other dilemmas relevant to the matter. In a democracy, such issues are clarified through public discourse and political process. It cannot be assumed that the demand for cyberspace security will disappear, that the problem will go away, or that the state will be able to shrug off its responsibility towards citizens. In the aforementioned Israeli case, nothing exempts the state authorities from responding to various citizen demands and undertaking legal and regulatory changes to increase data security on commercial websites. Failure to regulate and enforce law and order in cyberspace will enable a range of cybercrime to flourish, to the point of real threats to national security: providing service to hostile elements aiming to carry out cyber attacks and increasing the scope of crime to the point of compromising both personal security and the nation's business environment.

A Dangerous Interface: Cybercrime as a National Security Threat

Cybercrime continues to grow and challenges developed nations in different ways. Existing information about cybercrime is acquired from periodic reports by consulting, IT and information security companies, and law enforcement agencies. Given the problems inherent in identifying the phenomenon, the crude use of statistical methods for a quantitative analysis, and the inclusion of indirect damage in monetary assessments, it is apparent that existing information is not reliable. It seems that monetary assessments are consistently inflated. Nonetheless, that there is great potential danger in cybercrime cannot be overlooked.

The analysis in this article shows that in effect a large range of cybercrime does not represent a threat to national security. Phenomena such as theft and industrial espionage, fraud, harmful contents, hate crime, destruction of websites, denial of service, and so on are liable to become a national security problem only if there is a marked increase in their incidence and their effects are lasting. Therefore, now is the time to take action to reduce the risk and make it more difficult for cybercriminals to operate in this realm.

Past experience shows that hostile elements recruit criminal expertise to achieve operational goals. Because of the pace of technological developments, what today are advanced IT capabilities will within very few years become inexpensive, off-the-shelf commodities. The black market of computer services makes advanced capabilities readily accessible. The evidence exacerbates the concern that in cyberspace too, cooperation among criminal elements and hostile entities exists and is on the increase.

On the basis of this analysis, focus on two major interfaces between cybercrime and national security is recommended. First, the nation state is the entity responsible for the personal and collective safety and security of its citizens. Cybercrime causes various kinds of damage to citizens and organizations. The scope of such damage is unclear and the various damage estimates proffered in the debate are largely unreliable and exaggerated. But even without agreement on the scope and damage incurred by citizens, organizations, and states, the state must still respond to the opportunities and challenges of the reality as it unfolds. With the ongoing entry of cyberspace into every walk of life, it is safe to assume that demands on the state to assure personal and national security in cyberspace will also grow. Despite the global nature of cyberspace, the state will be forced to expand its involvement considerably. The outline of state involvement in cyberspace has been emerging in recent years, one of the more loaded issues being the mutually contradictory values of privacy and national security. In a democracy, the process for formulating a government policy on cybercrime involves public debate, political battles, and long term legal treatment.

Second, the commercialization of technical and operational capabilities is lowering the threshold for entering the cyber warfare arena, expanding the reference threats beyond states and large terrorist organizations, and placing a very heavy burden on national security authorities. Cyber criminal organizations offer resources, infrastructures, and even customer service at reasonable cost. This is a market that can be exploited not only to commit crime for financial profit but also to carry out direct attacks on national security. Defending critical infrastructures against cyberspace threats is a key issue in cyber security and its importance is even greater given the prevalence of potential elements of risk capable of acquiring cyberspace weapons and recruiting "fighters" on the cyber criminal black market.

Given the analysis of the phenomenon's significance and the identification of dangerous interfaces between cybercrime and national security presented herein, the immediate state focus should be on dealing with the threat in order to prevent it becoming more acute. The state must upgrade its involvement in creating cyberspace security, but it cannot solve the problem alone. The successful realization of state responsibility for cyberspace security necessitates the cooperation of all interested parties in the business, academic, public, and security sectors, so as to provide national and personal cyberspace security to the state and its citizens.

Notes

[1] P. N. Grabosky, "Virtual Criminality: Old Wine in New Bottles?" *Social & Legal Studies*, 10, no. 2 (2001): 243-49.

[2] Majid Yar, *Cybercrime and Society: Crime and Punishment in the Information Age* (London: SAGE Publications, 2006).

3 D. L. Shinder and M. Cross, *Scene of the Cybercrime* (Burlington, MA: Syngress, 2008).

4 David S. Wall, *Cybercrimes: The Transformation of Crime in the Information Age* (Cambridge: Polity, 2007), p. 10.

5 A. Alkaabi, G. M. Mohay, A. J. McCullagh, and A. N. Chantler, "Dealing with the Problem of Cybercrime," Conference Proceedings of 2nd International ICST Conference on Digital Forensics & Cyber Crime, October 4–6, 2010, Abu Dhabi, http://eprints.qut.edu.au/38894/1/c38894.pdf.

6 CoE, "Convention on Cybercrime," Budapest, 2001, http://conventions.coe.int/Treaty/en/Treaties/html/185.htm.

7 A botnet is a collection of internet-connected computers whose defenses have been breached and control ceded to a malicious party gaining distance control and using these computers' capabilities. A botnet is commonly used for sending spam, attacking DDoS, and continuous data theft. See https://www.checkpoint.com/products/anti-bot-software-blade/anti-bot-software-blade-landing-page.html.

8 Asymmetric key cryptography is the basis of the RSA algorithm developed by Leonard Adelman, Adi Shamir, and Ron Rivest, and presented publicly in 1978. Its patent expired in 2000. PGP (Pretty Good Privacy) developed by Phil Zimmermann in 1991 was the first software to allow free use of strong encryption using this method . The common web security standards (HTTPS, TLS/SSL, SSH, Bitcoin) are employing the same public key cryptography principle.

9 Richard M. Ryan and Edward L. Deci, "Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions," *Contemporary Educational Psychology* 25, no. 1 (2000): 54-67.

10 A. R. Piquero and Stephen G. Tibbetts, eds., *Rational Choice and Criminal Behavior: Recent Research and Future Challenges* (New York: Routledge, 2002).

11 The number of infected computers is itself no indication of the network's power or potential damages. See Daniel Plohmann, Elmar Gerhards-Padilla, Felix Leder, *Botnets: 10 Tough Questions* (ENISA, 2011).

12 Wall, *Cybercrime,* p. 221.

13 See for example the GAO-07-705-Cybercrime Report, June 17, 2007, pp. 16-17, http://www.gao.gov/assets/270/262608.pdf.

14 "2005 FBI Computer Crime Survey," p.10, www.fbi.gov/publications/ccs2005.pdf.

15 Melissa E. Hathaway, "Falling Prey to Cybercrime: Implications for Business and the Economy," ch. 6, in *Securing Cyberspace: A New Domain for National Security* (Queenstown: Aspen Institute, February 2012).

16 Office of Cyber Security & Information Assurance in the UK Cabinet Office and BAE Detica,: "The Cost of Cyber Crime," 2011, http://www.cabinetoffice.gov.uk/sites/default/files/resources/the-cost-of-cyber-crime-full-report.pdf.

17 "Norton Study Calculates Cost of Global Cybercrime: $114 Billion Annually," http://www.symantec.com/about/news/release/article.jsp?prid=20110907_02.

18 M. Lesk, "Cybersecurity and Economics," *IEEE Security & Privacy,* 9, no. 6 (2011), p. 76; Carl Bialik, "A Cybercrime Stat's Nine Lives," *Wall Street Journal,* September 26, 2007, http://blogs.wsj.com/numbersguy/a-cybercrime-stats-nine-lives-194/tab/print/.

19 Dinei Florêncio and Cormac Herley, "Sex, Lies and Cybercrime Surveys," Microsoft Research, 2012. The study was condensed and appeared as an op-ed piece in Dinei Florêncio and Cormac Herley, "The Cybercrime Wave That Wasn't," *New York Times*, April 15, 2012, https://www.nytimes.com/2012/04/15/opinion/sunday/the-cybercrime-wave-that-wasnt.html?_r=3&hpw.

20 Card Verification Code – the secret three-digit code printed on the back of credit cards, used to verify the validity of the card details when the card is not being read magnetically.

21 This discussion exceeds the scope of the present article. For a good overview, see the chapter on surveys in Francis C. Dane, *Evaluating Research: Methodology for People Who Need to Read Research* (Los Angeles: Sage, 2011).

22 "Counterfeit CDs are Money for Islamic Terrorism," *Ynet*, January 16, 2003, http://www.ynet.co.il/articles/0,7340,L-2378873,00.html.

23 J. Hunt, "The New Frontier of Money Laundering: How Terrorist Organizations Use Cyberlaundering to Fund Their Activities, and How Governments Are Trying to Stop Them," *Information and Communications Technology Law* 20, no. 2 (2011): 133-52.

24 Israel Security Agency, "Report on Hamas' Use of Underground Passages in the Gaza Strip," November 2008, http://www.shabak.gov.il/publications/study/Pages/hamas-tunnel-report.aspx.

25 Israel Security Agency, "Smuggling Weapons to the Gaza Strip from Iran via Sudan and Sinai," http://www.shabak.gov.il/publications/study/Pages/Sudan120511.aspx?webid=a3db3c16-25d8-423d-98df-eb1b9253ab93 .

26 Meir Amit Intelligence and Terrorism Center, http://www.terrorism-info.org.il/malam_multimedia/Hebrew/heb_n/html/ipc_272.htm.

27 Nir Kshetri, "The Global Cybercrime Industry and Its Structure: Relevant Actors, Motivations, Threats, and Countermeasures," in Nir Kshetri, ed*, The Global Cybercrime Industry: Economic, Institutional and Strategic Perspectives* (Heidelberg; London: Springer, 2010); Misha Glenny, *Darkmarket: Cyberthieves, Cybercops, and You* (New York: Alfred A. Knopf, 2011).

28 Cyber weapons may be categorized by their intended usage: malware – malicious software meant to disrupt the normal workings of a computerized system clandestinely, thereby damaging the process controlled by that system; spyware – malicious software meant to gather data clandestinely and sometimes transfer it over the internet; scanners to identify known vulnerabilities; remote and local exploits – to exploit known vulnerabilities;

network sniffers – to eavesdrop on communications; backdoor tools, Trojan horses – for distance access and data retrieval.

[29] See Isaac Ben-Israel and Lior Tabansky, “An Interdisciplinary Look at Security Challenges in the Information Age,” *Military and Strategic Affairs* 3, no. 3 (2011), p. 24, http://www.inss.org.il/upload/(FILE)1333532835.pdf.

[30] Phil Williams, “Organized Crime and Cybercrime: Synergies, Trends and Responses,” *Global Issues* 6, no. 2 (2001): 5.

[31] McAfee, “Virtual Criminology Report: Organized Crime and the Internet,” December 2007, www.mcafee.com/us/research/criminology_report; C. Czosseck, G. Klein, and F. Leder, “On the Arms Race around Botnets: Setting up and Taking Down Botnets,” paper presented at the Cyber Conflict (ICCC), 2011 3rd International Conference, June 7-10, 2011.

[32] “Kaspersky Reveals Price List for Botnet Attacks,” July 23, 2009, http://www.computerweekly.com/news/1280090242/Kaspersky-reveals-price-list-for-botnet-attacks. It seems that the cost continues to drop. See Plohmann, Gerhards-Padilla, and Leder, *Botnets: 10 Tough Questions.*

[33] Jeffrey Carr, November 2, 2011, http://jeffreycarr.blogspot.com/2011/11/words-matter-dump-apt-for-apa.html.

[34] All high profile cases of cyber espionage, such as “Gh0st RAT,” RSA/Lockheed-Martin, and “Flame” are examples of an APA.

[35] *Software as a Service: Strategic Backgrounder* (Washington, D.C.: Software & Information Industry Association, February 28, 2001), http://www.siia.net/estore/pubs/SSB-01.pdf.

[36] https://www.gartner.com/it/page.jsp?id=1963815.

[37] Lior Tabansky, “Basic Concepts in Cyber Warfare,” *Military and Strategic Affairs* 3, no. 1 (2011): 75-92, http://www.inss.org.il/upload/(FILE)1308129610.pdf.

[38] CoE, “Convention on Cybercrime.” Since 2001, the convention has been ratified by 30 of the 46 signatory nations.

[39] T. Maurer, “Cyber Norm Emergence at the United Nations: An Analysis of the UN’s Activities regarding Cybersecurity,” Belfer Center for Science and International Affairs, Harvard Kennedy School, September 2011.

[40] OECD, “Communiqué on Principles for Internet Policy-Making,” June 29, 2011.

[41] EU, Europol, the European Cybercrime Centre (EC3) officially commenced its activities on January 1, 2013, https://www.europol.europa.eu/ec3.

[42] ITU, *National Cybersecurity Strategy Guide*, September 2011.

[43] Mike McConnell, Michael Chertoff, and William Lynn, “China’s Cyber Thievery is National Policy-and Must Be Challenged,” *Wall Street Journal*, January 27, 2012; Richard Clarke, “How China Steals our Secrets,” *New York Times,* April 2, 2012; Nathan Gardels, “Cyberwar: Former Intelligence Chief Says China Aims at America’s Soft Underbelly,” *New Perspectives Quarterly* 27, no. 2 (2010):15-17; Joel Brenner, *America the Vulnerable: Inside the New Threat Matrix of Digital Espionage, Crime, and Warfare* (New York: Penguin

Press, 2011); U.S.-China Economic and Security Review Commission (USCC), *2009 Report to Congress of the U.S.-China Economic and Security Review Commission.*

44 See Myriam Anna Dunn and Kristian Søby Kristensen, eds., *Securing "the Homeland": Critical Infrastructure, Risk and (In)Security* (London: Routledqe, 2007).

45 Ro'ee Goldenberg, "The Bank of Israel: Details of 15,000 Credit Cards Stolen," *Globes*, January 3, 2011, http://www.globes.co.il/serve/globes/printwindow.asp?did=1000712125; Yazan al-Saadi, "Saudi 0xOmar: Hackers of the World Unite Against Israel," *al-Akhbar English,* January 16, 2012, http://english.al-akhbar.com/node/3413.

46 Dump: a stolen credit card or bank account and the associated customer data. T. J. Holt, and E. Lampke, "Exploring Stolen Data Markets Online: Products and Market Forces," *Criminal Justice Studies* 23, no. 1 (2010).

Iran and Cyberspace Warfare

Gabi Siboni and Sami Kronenfeld

Introduction

Throughout the world decision makers and the general public have undoubtedly realized in recent years that cyberspace must be treated as a genuine realm of warfare. As such, it allows considerable room for maneuvering and has vulnerabilities that can be breached by hostile elements seeking to derail information systems or even inflict physical damage on critical infrastructures controlled by industrial control systems. In the wake of this new understanding, many countries are investing increasingly in safeguarding their cyber resources (particularly in the fields of defense, intelligence gathering, and offense capabilities). Since the Stuxnet attack – one of the most destructive cyber attacks to date – Iran has been working hard to improve its cyberspace defenses on the one hand, while building up cyberspace intelligence gathering and offensive capabilities on the other.

The Iranian cyberspace defense program has a dual objective: first, it hopes to prevent another attack like Stuxnet and intelligence-directed penetration of Iranian computers by viruses such as Duqu and Flame. In this sense, the goal of the Iranian program is similar to that of many other nations seeking to protect their critical infrastructures. The second objective is the regime's desire to ensure its survival by means of surveillance and blocking of information and services originating with the Iranian public. In many cases the two goals are achieved with the same tools, e.g., the Iranian effort to create a separate Iranian web or the disabling of Google services in that country.[1]

At the same time, Iran is also in the midst of a concerted effort to construct offensive capabilities, on the assumption that in any future confrontation

Dr. Gabi Siboni is the head of the Military and Strategic Affairs Program and head of the Cyber Warfare Program at INSS. Sami Kronenfeld is an intern in the Cyber Warfare Program at INSS.

This article was first published in *Military and Strategic Affairs* 4, no. 3 (2012): 77-99.

the use of cyberspace will have a critical impact on achieving success against the enemy. Gathering information openly about Iranian cyberspace capabilities, especially offensive ones, is by definition extremely difficult. But the country's cyberspace activities have recently been in the spotlight because of suspicions of Iranian involvement in some serious cyberspace incidents, including the theft of internet security permissions, an attack on the Saudi Arabian oil company's organizational network, and not least, the penetration of computers at some leading American banks.

This article examines the current situation regarding various elements of Iran's cyberspace development process. The first section analyzes the country's cyberspace strategy, while the second section describes the organizational and operational response to the formulated strategy. This comprises three components: infrastructures for training and developing technological manpower for work in cyberspace; technological developments that have already been introduced; and the overall processes of cyberspace force construction. Finally, the article focuses on a number of cyberspace incidents attributed to Iran, attempts to gain some insight into the way Iran conducts its cyberspace activities, and examines implications for Israel and other Western nations.

Iran's Cyberspace Strategy

The role of the communications and information networks in the outbreaks that followed the 2009 Iranian presidential election and those that erupted as part of the "Arab Spring," as well as the cyber attacks on Iran made the cyberspace arena tremendously important to the Iranian regime's overall security doctrine. Evidence of the subject's significance in the minds of Iran's decision makers was proffered by none other than the Supreme Leader himself, Khamenei, in a direct reference to the opportunities and dangers of cyberspace when, in March 2012, he announced the establishment of a Supreme Cyberspace Council composed of senior government representatives charged with planning and implementing a single integrated cyberspace strategy.[2] While the work of this Council began only quite recently, an analysis of Iranian cyberspace activity in recent years indicates the existence of an Iranian cyberspace strategy with clear goals and objectives.

Two fundamental assumptions underlie Iran's approach to its modus operandi in cyberspace. The first concerns the development of defensive capabilities to withstand attacks by hostile nations and entities, alongside the development of operational capabilities against opponents of the regime on the home front; the second concerns the development of offensive capabilities to enable Iran to combat what it sees as American superiority and control of global internet capabilities and infrastructures.

In the defense arena, Iran is working to accomplish two main goals in cyberspace.[3] First, it aims at an effective, comprehensive, advanced technological protective system to defend critical infrastructures and sensitive data against cyber attacks such as Stuxnet, which compromised the Iranian uranium enrichment program and shut down more than 1,000 centrifuges at the enrichment facility in Natanz.[4] Second, Iran is trying to curb and foil the cyberspace activities of domestic opposition parties and opponents of the regime, for whom cyberspace is an important communications platform for disseminating information and organizing anti-government activities. In addition, the regime hopes to prevent the cyberspace penetration of Western ideas and information that conflict with its interests, thereby blocking "soft revolution" processes that are liable to damage the regime's stability and hold on the state. In the context of defensive capabilities, the news about Iranian plans to develop a separate, independent communications network is noteworthy.[5] Although this has at times been denied by Iranian officials,[6] as time goes by it seems to take on more validity.[7]

On the offensive front, Iran's cyberspace strategy sees this arena first and foremost as central in the asymmetrical doctrine of warfare, a key principle in Iran's perception of the use of force. Iran sees cyberspace warfare, in a similar way to more obvious asymmetrical tactics such as terrorism and guerilla warfare, as an effective tool to inflict significant damage on the enemy's home front with military or geostrategic superiority. Experts estimate that in the event of an escalation in the confrontation between Iran and the West over the Iranian military nuclear program, Iran would attempt a cyber attack against major infrastructures – such as power plants, financial institutions, and transportation systems – on American soil.[8] An article published in July 2011 in the Iranian newspaper *Kayhan* (which is closely identified with Khamenei) hinted at such a possibility by warning that the United States must take care lest "an unknown player somewhere in the world" carry out an attack on its most vital infrastructures.[9]

Beyond the military-strategic aspect, the Iranian regime and its supporters also use offensive cyberspace warfare to impair the cyber activities of Western countries and opponents to the regime in Iran. Iranian hackers, who usually have no official affiliation with the establishment but are linked to it nonetheless, consistently engage in cyber attacks causing internet crashes, inserting pro-Iranian material, stealing information, committing credit card fraud, damaging service providers, and rerouting internet traffic.[10] Propaganda is another part of the cyberspace warfare strategy. The Iranian regime understands well the importance of cyberspace in shaping the points of view and attitudes of large groups of people inside Iran and abroad, and invests major efforts in creating a sizable and effective propaganda machine extolling the regime and maligning its enemies. To realize these strategic goals, Iran is investing considerable resources in creating a tight, skilled, multi-layered structure that includes impeding, monitoring, controlling, and offensive capabilities in cyberspace.

Iran's Organizational and Operative Response

With its cyberspace strategy goals in mind, Iran set about applying itself vigorously to strengthening its cyberspace capabilities. There are reports of investments amounting to some $1 billion in the development and acquisition of technologies and in recruitment and training of experts to advance and strengthen both defensive and offensive cyberspace capabilities.[11] There are various interconnected components in the processes of building an operative and organizational cyberspace response: first, building up a training and development manpower base at research institutes and institutions of higher education; second, efforts towards large scale technological development; and third, processes of force buildup, including development of a doctrine, establishment of organizations, and formulation of a hierarchy of authority to implement the doctrine.

Manpower Training and Development

The infrastructures for the technological training and development of Iranian cyberspace are found primarily in the country's universities and technological institutes. Iran has many institutions of higher education and academic research engaged in research and training in the fields of IT, computer engineering, and communications.[12] Leading universities in this area include: Sharif University of Technology in Tehran, offering advanced degrees in computer engineering, electronic engineering, and mathematics,[13] and which is also the site of two advanced research institutes in communications and information technologies (the Advanced Information and Communication Technology Center[14] and the Advanced Communication Research Institute[15]); and Amikabir University of Technology, also in Tehran, with large departments of mathematics, computer sciences, computer engineering, and information technology. It seems that Amikabir specializes in data security; the computer engineering department offers several advanced courses in security information,[16] and also operates a research lab specializing in data security[17] and a separate research lab specializing in secure systems analysis.[18]

In addition to academic research and training, the Iranian regime invests significant sums in the promotion and support of IT and computer communications companies. Such investments are made directly by government organizations such as the Science Ministry, and indirectly via the financing and establishment of greenhouses for hi-tech companies in which the government has an interest.[19] The Iran Telecommunications Research Center is a key government body in the IT field; it specializes in research in information and communications technology and is the research and professional arm of the Information and Communications Ministry. The center operates and trains advanced research teams in many fields, including data security.[20] Another government body promoting research in IT is the Technology Cooperation Office, which belongs to the Presidential Bureau. Its stated objective is to improve technological cooperation with other nations. It directs and initiates research projects in many areas, including information technologies.[21] The EU and other Western sources have singled it out as being involved in the nuclear program.[22]

Apart from direct investments by government bodies, the Iranian regime also operates hi-tech greenhouses engaged in data security research. Prominent among such hi-tech centers is the Pardis Technology Park, also known as the Iranian Silicon Valley. Established in 2001 by the Presidential Bureau and the Technology Cooperation Office, it houses more than 400 companies involved in communications and IT.[23] Another hi-tech greenhouse is Guilan Science and Technology Park, a support center for startups and home to a number of companies working on information security.[24]

Technological Empowerment

Beyond developing and training a strong cyberspace workforce, Iran has also been focusing on technology to promote its strategic goals in cyberspace. One target of major investment is intra-state cyberspace and information flow. In recent years, the Iranian regime has bought and developed advanced technological systems allowing it to conduct surveillance and monitor information traffic on computer and mobile networks in the country. The largest government controlled telecom corporation (the Telecommunications Company of Iran) bought a surveillance system from the Chinese ZTE Corp. The system, capable of monitoring information on telephone lines, computer networks, and cellular lines, was acquired as part of a comprehensive deal between the two companies estimated at $130 million. The deal covered products of the ZMXT system, which the Chinese company describes as an integrated monitoring system. The products purchased enable voice communications eavesdropping, text message surveillance, and monitoring of web surfing.[25]

In addition to surveillance and monitoring, the Iranian government is also developing website blocking and filtering technologies, since international sanctions prevent Iran from buying Western-manufactured data filters. Amnafzar Ltd., an IT company with links to the regime, developed a data filter called Separ, which is updated constantly and frequently changes its filtering strategy so as to evade efforts to circumvent it.[26] Using this technology, the regime has succeeded in significantly limiting the flow of information into and within the country. Research published in March 2009 by the OpenNet Initiative (a joint project by a number of institutions, including Harvard University and the University of Toronto) identified Iran as one of the leading nations in website filtering and blocking, alongside nations such as China, North Korea, Syria, and Myanmar.[27]

These technologies allow Iran relatively close control of the state's cyberspace, but the regime nonetheless strives for outright control of information, ideas, and access to Iranian cyberspace. To this end Iran embarked on a project of establishing an independent and separate national network, isolated from the World Wide Web. The idea is that the establishment of this national web, named Halal, will allow the regime full control of contents for public exposure and will also cause serious damage to opponents of the regime conducting widespread activities on the internet. It will also make virus attacks and other cyber attacks on Iranian infrastructures much more difficult. The national network project first came into being in 2009, when the Iranian authorities instructed domestic companies to move their network activities to servers and data centers on Iranian soil. During 2012 it was reported that Iran is developing an internal email service, an independent operating system, a search engine, and other tools for use on the new network.[28] In August 2012 Iranian Communications Minister Reza Taghipour announced that Iran would disconnect from the World Wide Web within 18 months.[29] However, Western experts believe it will be difficult for the regime to sever all connections with the global network.[30]

Iran is also seeking to isolate networks in the security establishment and construct a national intelligence communications network separate from the global web.[31] The first indication of this effort is Basir, the intra-organizational network of the Revolutionary Guards, whose existence became public knowledge in March 2012. Reports describe it as a closed cellular network, possibly operated by designated relay stations. The network supposedly affords the organization efficient, encrypted lines of communication, even in a scenario of a comprehensive cyber attack on the country's communications and information infrastructures. Thus far it is unclear if it is also an information network or a voice system only.[32]

Force Buildup

As for cyberspace force buildup processes, the many training and development facilities available to Iran have allowed the Islamic Republic to establish a large cyberspace configuration with multiple capabilities, both defensive and offensive. In the last decade, Iran embarked on a strategic expansion of its national cyber constellation, with cyberspace agencies and organizations established for almost every relevant government ministry. The goal is to create a hierarchical and diverse organizational alignment with a clear plan of action, well thought out resource allocations, distribution of responsibility and the ability to preserve and disseminate information, know-how, and data.

The crowning glory in the construction of Iran's cyberspace force is the establishment of the Supreme Cyberspace Council. The Council was set up in March 2012 at the behest of Supreme Leader Ayatollah Khamenei and serves as the ultimate authority on all of the nation's cyberspace issues.[33] The Iranian President heads the Council and its members comprise senior government representatives and others, including the senior commander of the Revolutionary Guards, the head of the Majlis, the Ministers of Science, Communications and Culture, the chief of police, and the president of the Islamic propaganda organization. The Council has the authority to determine national cyber policy and its directives are binding on all Iranian institutions operating in the field. The Council plans to establish a National Cyber Center under its auspices, to integrate all Iranian cyberspace activity, gather and disseminate information and instructions, and oversee the enforcement of the Council's directives by all relevant bodies.

Iran's cyberspace structure comprises many cyberspace organizations working in various fields and officially affiliated with establishment organizations. One central organization with a defensive orientation is the Cyberspace Defense Command, which operates in the context of the Passive Defense Organization belonging to the general staff of the armed forces.[34] Alongside military personnel, this cyberspace organization also comprises government ministry representatives (the Communications, Defense, Intelligence, and Industry ministries). Its main objective is to develop a comprehensive defensive doctrine for state institutions and infrastructures against cyber threats.[35] The organization is primarily defensive, and currently does not seem to be involved in offensive cyber activity.

Another defensive cyberspace entity is the Center for Information Security, known as MAHER, established and operated as part of the Communications and Information Technologies Ministry. This center is primarily responsible for activating computer security incident response teams in the event of emergencies and cyber attacks. In addition, the center trains skilled manpower, develops response mechanisms to cyber crises, and stores and disseminates data security know-how. It is responsible for defending all government websites, as well as those of private companies operating officially and listed with the Communications Ministry. The center's teams were called on to impede and foil the work of the Flame and Stuxnet viruses that attacked Iran.[36]

Other cyberspace organizations focus on enforcement and control of intra-Iranian cyber activities that run counter to the regime's interests. In July 2009, the Supreme Council of the Cultural Revolution, which is subject to the supreme leader, founded the Committee to Identify Unauthorized Websites. Among its members are the Attorney General, the chief of police, the supervisor of government media, and various government ministers (from the Intelligence, Communications, Culture, and Science ministries, among others). The committee's purpose is to identify websites whose contents and activities are incompatible with the regime's requirements and wishes, and it is authorized to block access to such sites.[37] In 2011, the police established its own cyberspace unit, FETA,[38] to combat cybercrime – fraud, data theft, threats, and so on – but it is also authorized to take action against political and security criminals in cyberspace, and it is actually this latter task that primarily occupies it.[39] In addition, FETA is further charged with monitoring and controlling internet users in Iran, especially those in internet cafes around the country, where web surfing can be relatively anonymous.[40]

As for the offensive capabilities of Iran's cyberspace resources, the picture is less clear. Naturally, the Revolutionary Guards are crucial in the establishment and operation of offensive cyberspace warfare. Western experts place Revolutionary Guards capabilities in the top tier of cyberspace warfare worldwide.[41] A 2008 analysis by the research institute Defense Tech[42] estimated that the Revolutionary Guards cyberspace warfare program employed some 2,400 professionals and at that time had a budget of $76 million. Among capabilities that Defense Tech attributed to the Revolutionary Guards were: developing infected software by inserting malicious codes into counterfeit computer software; developing capabilities to block communications and WiFi networks; developing malicious codes (viruses and worms) capable of reproducing in networks and attacking target computers; developing tools for penetrating computers and networks to gather intelligence and pass it on to remote servers; and developing delay mechanisms installed in target computers to be operated by a predetermined schedule or by command from control servers.

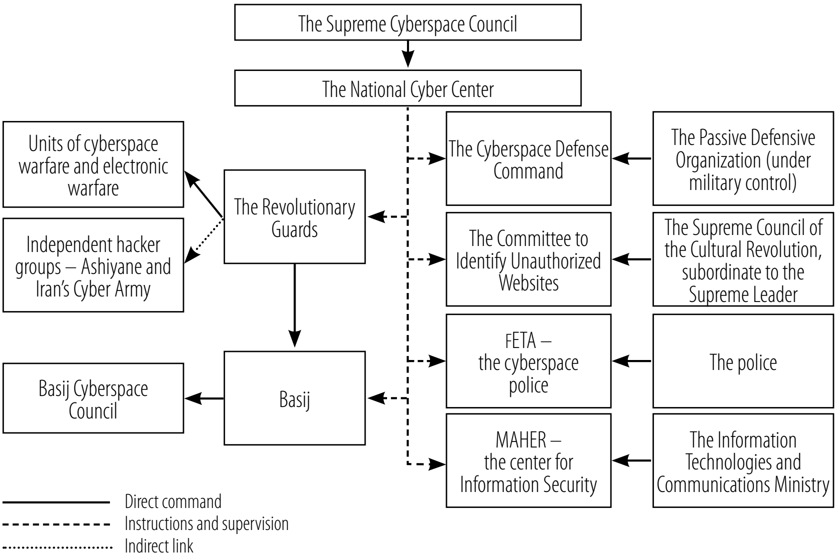
In addition to information warfare capabilities, the Revolutionary Guards are also creating an electronic warfare system capable of blocking radar and communications. The organization is investing large sums in the acquisition of electronic warfare systems[43] that, in conjunction with existing cyberspace warfare capabilities, will serve as an effective tool for compromising the electronic systems of the United States and its allies during a military confrontation.[44] According to declarations by the Revolutionary Guards, Iran has exhibited its prowess in the realm of cyberspace warfare with the capture of an unmanned aerial espionage vehicle in December 2011.[45]

Other than the Revolutionary Guards cyberspace warfare units, there is evidence linking the Revolutionary Guards and groups of Iranian hackers active against domestic and global enemies of the regime. The use of outsourcing allows the Revolutionary Guards and Iran to maintain distance and refute any allegations of Iranian involvement in cyberspace warfare and cybercrime. Experts have identified one group of Iranian hackers involved with the Revolutionary Guards as the Ashiyane Digital Security Team,[46] whose members are motivated by an ideology supporting the Iranian regime and the revolution, and who aim their attacks at the regime's enemies. The Ashiyane Team trains hackers and gives them significant capabilities,[47] which are then used for political activities (including the insertion of pro-Iranian propaganda into Western and Israeli websites and causing them to crash), as well as criminal enterprises (credit fraud, identity theft, and infiltration of databases and financial institutions). Furthermore, the group hosts a forum called War Games, which holds hacker competitions whose targets include American infrastructures companies.[48]

Another hacker group believed to be linked to the Revolutionary Guards is Iran's Cyber Army,[49] which consists of hackers and computer experts using fictitious identities and declaring themselves part of an organization. The group's main activities include breaking into Western websites with the aim of inserting pro-Iranian contents, seizing control of and redirecting information traffic, infiltrating Western data security companies, and damaging websites of the regime's opponents.

The Basij organization, which is subordinate to the Revolutionary Guards, has also become active in cyberspace and in 2010 established the Basij Cyberspace Council. Basij focuses primarily on creating pro-Iranian propaganda in cyberspace. It recruits and trains thousands of Iranians to write contents, afterwards deploying organized computer groups for tens of thousands of pro-regime bloggers. They also write talkbacks and other materials supporting the regime in the new media, on major forums, and on websites in Iran and abroad.[50] Nevertheless, Basij plans to further advance its cyberspace capabilities and is using experts from the Revolutionary Guards' cyberspace units to train hackers with high offensive capabilities.[51]

All of this clearly illustrates that in recent years Iran has established an extensive cyberspace structure encompassing many areas of activity, and has a wide range of capabilities at its disposal. The organizational flowchart below demonstrates the hierarchical configuration of the state's cyber establishment, as described above.


The Supreme Cyberspace Council
The National Cyber Center
Units of cyberspace warfare and electronic warfare
Independent hacker groups – Ashiyane and Iran's Cyber Army
The Revolutionary Guards
Basij Cyberspace Council
Basij
The Cyberspace Defense Command
The Committee to Identify Unauthorized Websites
FETA – the cyberspace police
MAHER – the center for Information Security
The Passive Defensive Organization (under military control)
The Supreme Council of the Cultural Revolution, subordinate to the Supreme Leader
The police
The Information Technologies and Communications Ministry
Direct command
Instructions and supervision
Indirect link

Clearly there have been significant advances in Iran's cyber activities. On the defense front all energies are focused on creating a defensive and isolation capability adequate for coping with any attempts at infiltrating the country's vital networks and infrastructures. Although it is hard to gain an entirely reliable picture of the development of offensive cyber capabilities, the following section of this article looks at several such activities.

Cyberspace Activities Attributed to Iran

In December 2011, an expose broadcast in an investigative program on the Univision television network led to an American inquiry into the involvement of official Iranian personnel in a cyber plot against the United States. The network's investigative reporters managed to infiltrate a group of Mexican hackers operating against US targets and secretly videotaped a meeting between their representatives and the Iranian Ambassador to Mexico. At the meeting, held at the Iranian Embassy, the hackers asked about the possibility of receiving support and financing from the Iranian government in order to carry out cyberspace attacks on American targets, such as the Pentagon, the CIA, the FBI, and various American nuclear installations. The video shows then-Iranian Ambassador to Mexico Muhammad Hassan

Ghadari asking questions and proposing additional courses of action. The Ambassador stressed that Iran wants information on the possibility of an American attack on Iran. At the end of the conversation, he expressed his desire to stay in touch with the hackers and promised to forward the proposal to his superiors.[52] It may be assumed that this attempt was not an isolated one and that Iran is actively recruiting hackers and others around the world to further its offensive cyberspace goals.

A decisive determination of the identity of cyberspace attackers is complex and requires resources and international cooperation. Therefore, it is hard to say with absolute certainty who is behind many cyberspace actions. Nonetheless, it is often possible, using circumstantial diagnostics, to identify those responsible with a high degree of certainty. This article highlights three incidents: an attack on two data security companies aimed at stealing security permissions; an attack on large financial institutions in the United States; and an attack on the Saudi Arabian oil company Aramco.

The Attack on DigiNotar and Comodo

In 2011 two attacks took place on companies providing SSL (secure sockets layer)[53] permissions. The first, in March 2011, targeted the American company Comodo Ltd. Several permissions were stolen, among them domain permissions of internet mail services such as Google, but these were withdrawn before being used by the attacker. In fact, someone with authority in the mail.google.com domain can steal Gmail passwords and hijack users' accounts. Someone with a stolen authorization for the Microsoft.com domain can install malicious software in victims' computers. According to the company, the following findings came to light about this incident:[54]

a. The attack lacked features typical of cybercrime.
b. The attackers were organized and knew precisely what they were seeking before the attack, indicating the involvement of a state organization in the attack.
c. The source of the attack was primarily Iran (based on identification of the IP address).
d. The website where the stolen permissions were checked is located in Iran and was immediately removed from the web after Comodo discovered the attack.

The attack on Comodo failed to achieve its goal: it was identified and neutralized before the stolen permissions could be used. However, this was not the case with DigiNotar, the major Dutch SSL permissions provider.

The company's databases came under attack from June through August 2011. During the attack, which came to be known by the name Black Tulip, certifications for website verification were stolen, including the certification serving to verify the google.com domain, thus allowing the attacker to assume this identity and reroute Gmail servers.[55]

An analysis ordered by DigiNotar (which went bankrupt and shut down operations after the attack) showed that 531 certificates were stolen and fabricated and that most stolen permissions were used to penetrate users' email accounts, especially in Iran. The analysis further revealed that the attack managed to penetrate more than 300,000 computers, which were overwhelmingly Iranian (more than 99 percent).[56] It is hard to determine the source of the attack with absolute certainty, but experts believe that it was Iran and that it was apparently intended for internal security purposes.[57] What led to this conclusion were the targets and extensive scope of users attacked and messages left on the company's website indicating Iranian involvement in the attack.

The Attack on American Financial Institutions

A report issued in the United States in September 2012 shows that at around the same time, several US financial institutions also came under attack, including sites belonging to the Bank of America, Morgan Chase, and Citigroup. Assessments by American sources concluded that the cyber attacks against the American financial institutions did not originate from random hackers, but were most likely financed by Iran and carried out by way of retaliation against sanctions imposed on Iran by the United States.[58]

As a result, the Financial Services Information Sharing and Analysis Center[59] issued an alert to banks in the United States about cyber attacks designed to steal identities via email, Trojan horses, and malicious tools for registering keystrokes and to retrieve user and employee names and passwords. Although large banks were also attacked, most of the victims were small and medium businesses, small banks, and credit companies. A group called the Izz ad-Din al-Qassam Cyberspace Fighters announced that it had attacked the Bank of America and the New York Stock Exchange in retaliation for a September 2012 movie expressing disrespect for the prophet Muhammad. These attacks, as described in the warning, indicate that the attackers succeeded in obtaining a great deal of information from the banks' networks, at least in some cases, and also accessed employees' entry permissions, thereby circumventing defensive mechanisms.[60]

The Attack against Aramco

In August 2012, apparently with insider help from someone with a high level of access to company computers, some 30,000 computers belonging to the Saudi Arabian oil company Aramco and the Qatari natural gas company ResGas were attacked by a computer virus called Shamoon. According to experts, this was one of the most devastating attacks carried out against any single company. The virus spread through the company's servers and attacked information stored in them. In-house computer experts say that the damage was limited to office computers and did not affect the company's operational and control systems.[61]

Symantec identified the virus for the first time in August 2012. An analysis by their experts and other security companies reveals the following findings:[62]

a. The Shamoon virus was designed to attack computers of an organizational computerized system (IT) rather than a control system. The virus is not in the same category of sophisticated cyberspace warfare tools such as Stuxnet, which attacked the Iranian nuclear program in 2010.
b. The purpose of the viral attack was not espionage or intelligence gathering but rather the complete and total destruction of data and target computers.
c. The writers of the malicious code do not seem to belong to the top tiers (such as the writers of Stuxnet and Flame), and there are indications that those behind it do not have a very high professional profile, since it was riddled with coding errors. They were, on the other hand, skilled enough to create a particularly destructive code.
d. The virus penetrated the company's computers with the help of a collaborator inside the company with direct access to the system and who seems to have used a USB device for the purpose.
e. The writers of the code used a section of a picture of a burning American flag to hide the contents of the files in the infected computers, indicating a political and/or religious (Islamic) affiliation.
f. The code of Shamoon's deletion mechanism contained the word Wiper. A similar name was used in the virus code of Flame, which attacked the Iranian oil company. This parallel raises a suspicion that the attack on Aramco was an Iranian retaliation to the Flame attack.

A group called The Cutting Sword of Justice claimed responsibility for the Aramco attack, declaring it was aimed at the main source of income in Saudi Arabia, a country accused of committing crimes against Syria and Bahrain. The group further claimed that the virus allowed it to access many secrets, but to date no relevant information on the issue has been reported. Reports on similar attacks on oil and gas companies in the Persian Gulf raised suspicions that the attacks were part of a concerted national effort. US Secretary of Defense Leon Panetta recently hinted at Iranian involvement in the attack. A former senior member of the American administration spoke out more directly when he claimed the administration believes Iran was behind the attacks in the Gulf.[63]

An analysis carried out by American cyberspace security expert Jeffrey Carr[64] raises a number of allegations linking Iran to the attack. It is the only country with access to the original Wiper code, which seems to have formed the basis for the Shamoon virus. According to a report issued by Kaspersky,[65] the Wiper code used in the attack on the Iranian Energy Ministry in April 2012 was also used by Shamoon's creators. Iran is highly motivated to attack the Saudi Arabian oil company because of harsh sanctions in place against Iran in the energy field. Furthermore, a suspicion of Hizbollah involvement in the attack was also investigated, and several Lebanese employees of Aramco were arrested and interrogated.

Conclusion

Iran's developed and developing cyberspace warfare capabilities should be a source of concern to Israel and, of course, the United States, as well as other Western nations. Because of the audacity demonstrated by the attempt on the life of the Saudi Arabian Ambassador to the United States, American experts feel that Iran's intentions and capabilities in daring to attack critical infrastructures in the United States should not be dismissed. Like the rest of the world, one may assume that Iran too – victim of one of the most destructive cyberspace attacks ever – has learned the lessons of Stuxnet and understands the destructive potential inherent in the development of an offensive tool that could damage industrial control systems, thereby causing physical destruction.

The development of the Iranian strategy and the subsequent force buildup processes indicates systematic preparations and organization with a view to becoming a major cyberspace warfare player. Experts report constant progress in Iran's cyberspace capabilities and operations. Following reports of the cyber attack on the American financial institutions attributed to Iran, one such expert stated, "[Iran's cyberspace program] is similar to the nuclear program: it isn't particularly sophisticated but it moves forward every year."[66] It would be a mistake not to take Iranian technological capabilities seriously. The country's science infrastructure is highly developed and there is a great deal of skilled manpower. One must therefore assume that before too long Iran will represent a significant threat in this area on the global level.

This assessment was further reinforced by the attack on Aramco, after which James A. Lewis, a specialist on cyberspace security, said that Iran was quicker in developing offensive capabilities and more daring in their use than anyone expected.[67] Usually, any activity that is exposed is no more than the tip of the iceberg of concealed activity. Furthermore, Iran's growing defensive sophistication requires interested parties to prepare to operate in an environment of isolated networks or an Iranian network isolated from the World Wide Web. Although the challenge of establishing such a network and achieving total isolation is enormous, such activity is also discernible. This defensive doctrine will represent a very tough challenge indeed for anyone interested in conducting activity in Iranian cyberspace.

The actions attributed to Iran as described above lead to several insights. Iran's attempts to secure SSL permissions indicate work against large groups of citizens rather than focused targets, such as nations or companies and organizations; they are apparently aimed at identifying and monitoring domestic targets. Nevertheless, the cumulative experience gained from such actions will also enable activity against more focused targets, such as nations and organizations. At the same time, although the detected activity indicates a certain degree of organization and systematic planning, it seems that Iran has yet to cross the threshold into the most sophisticated technological and organizational level. Nevertheless, the country's motivation, force buildup, and technological capabilities will enable it to make very rapid strides in that direction.

The attack on Aramco elicits further conclusions, the first being the fact that conventional defenses against internet threats are not enough. Most experts assume that the company had invested in protection against internet threats. The destructive virus was not discovered by virus protection systems and seems to have been inserted by a company insider possessing the appropriate permission. Current standard protective systems are not built to supply protection against focused threats (APT) and unknown malicious codes (Zero Date and others). Therefore, there is a growing need to develop tools capable of offering better protection against such threats. One such direction lies in developing tools based on the identification, blocking, and neutralization of anomalous and undesirable behavior in the computers under attack. Such tools can neutralize threats even after the malicious code has managed to enter the target computer. A second insight concerns the targets of the attack, which was aimed primarily at the mass and indiscriminate destruction of data in the tens of thousands of computers belonging to the Saudi Arabian oil company, rather than at intelligence gathering. If intelligence gathering in cyberspace may be considered legitimate in some cases, Iranian mass destruction of a civilian target is a sign that Iran is transitioning to retaliation. This should worry those in charge of defense in many nations. Leon Panetta's statement about the need to settle accounts with those behind the attack is one such illustration.[68] But of course actions will speak louder than words.

As the victim of one of the world's most destructive cyberspace attacks, one may assume that Iran fully understands the potential inherent in this realm, and accordingly will work to develop similar capabilities of its own. In that case, the systematic force construction described in this article will very quickly turn Iran into a significant player on the cyberspace battlefield; this will include attacking critical infrastructures in hostile nations, such as the United States and Israel, while creating maximum separation in the event of exposure of such activity. Iran uses so-called civilian hacker communities to try to create a distance between cyber activities and the regime and official Iranian organizations. A similar approach is adopted elsewhere in the world, e.g. China and Russia, allowing those nations to deny responsibility and lay the blame at civilian doors. Therefore the major challenge of connecting Iran to cyberspace offensives will continue.

Iran's focus of cyberspace activity on Israel and other Western countries requires designated defensive responses. All the countries in question need an updated doctrine on cyberspace defense and protection. The attackers' sophistication necessitates intelligence-based defense activity in addition to generic protections. Therefore, and in light of Iran's development processes, Israel must place Iranian cyberspace high on its list of intelligence priorities, preempting and foiling offenses before they can be carried out. In a way comparable to the Iranian nuclear program, the challenge is not Israel's alone but faces many nations in the West, as well as the Gulf states, as evidenced by the attack on Aramco. Hence, international cooperation of the widest scope possible should be initiated toward intelligence and preemption of Iranian cyberspace activity.

At the same time, Israel must continue to build an effective defensive response focused on three relevant national layers of cyberspace. The first is security organizations, which constantly need to test exposure to Iranian cyberspace capabilities and ensure they are not succeeding in damaging the critical capabilities of the defense establishment. The second concerns the network of critical infrastructures guided by the Information Security Authority by virtue of an Israeli government decision. Here too, the challenge requires constant activity, especially in terms of understanding the threat, adapting the response to it, and sharing information among the various institutions. Finally, one must not dismiss Iran's capabilities and possible attempts to damage non-governmental commerce and industry. Private sector commercial and industrial corporations usually take steps primarily to safeguard their data assets. It is hard to demand that they protect themselves against the possibility of a cyberspace attack from a foreign nation such as Iran. Hence the critical role of the recently established National Cyberspace Staff as an integrating entity capable of promoting processes of regulation, information sharing, and intelligence on the basis of the evolving map of threats.

Notes

[1] Art Keller, "The Great Persian Firewall, "*Foreign Policy,* September 2012, p.28, http://www.foreignpolicy.com/articles/2012/09/28/Iran_firewall_google?page=full.

[2] Khamenei's statement announcing the establishment of the council on his official website, http://farsi.khamenei.ir/message-content?id=19225.

[3] Ilan Berman, "The Iranian Cyber Threat to the U.S. Homeland," Statement before the US House of Representatives Committee on Homeland Security Subcommittee on Cybersecurity, Infrastructure Protection, and Security Technologies and Subcommittee on Counterterrorism and Intelligence, April 26, 2012, pp. 1-3, http://homeland.house.gov/sites/homeland.house.gov/files/Testimony%20-%20Berman.pdf.

[4] CBS News, "Iran Confirms Stuxnet Worm Halted Centrifuges," November 29, 2010, http://www.cbsnews.com/stories/2010/11/29/world/main7100197.shtml.

[5] Kevin McCaney, "Iran Building a Private, Isolated Internet, but Can it Shut out the World?" *GCN*, April 10, 2012, http://gcn.com/articles/2012/04/10/iran-building-separate-isolated-internet.aspx.

[6] Agence France Presse, "Iran Denies has Plan to Cut Internet Access," *AFP*, April 10, 2012, http://www.google.com/hostednews/afp/article/ALeqM5h4e57x6CYbsavza1PeDuQP7Bf9Vg.

[7] Amir Taheri, "Iran will Launch its National Internet Next Week but not for the Reasons you Might Think," September 20, 2012, http://www.opednews.com/articles/Iran-will-launch-its-natio-by-Amir-Taheri-120919-83.html.

[8] Brian Ross, "What Will Happen to the US If Israel Attacks Iran?" *ABC News*, March 5, 2012, http://abcnews.go.com/Blotter/israel-attacks-iran-gas-prices-cyberwar-terror-threat/story?id1584852.

[9] Berman, "The Iranian Cyber Threat to the U.S. Homeland," p. 4.

[10] Reza Marashi, "The Islamic Republic's Emerging Cyber War," National Iranian American Council, April 30, 2011, http://www.niacouncil.org/site/News2?page=NewsArticle&id=7318.

[11] Yaakov Katz, "Iran Embarks on $1b. Cyber-Warfare Program," *Jerusalem Post*, December 18, 2011, http://www.jpost.com/Defense/Article.aspx?id=249864.

[12] J. P. Patterson and M. N. Smith, *Developing a Reliable Methodology for Assessing the Computer Network Operations Threat of Iran*, Master's Thesis (Monterey, CA: Naval Postgraduate School, 2005), pp. 17-22, www.fas.org/irp/eprint/cno-iran.pdf.

[13] Sharif University website: http://www.sharif.ir/web/en.

[14] Institute website: http://www.aictc.com/web/content/main.

[15] Institute website: http://acri.sharif.ir/en/Default.asp.

[16] Advanced course descriptions: http://ceit.aut.ac.ir/autcms/courses/courseOfferingView.htm?level=M.Sc&depurl=computer-engineering&lang=en&cid=70317.

[17] The data security lab website: http://ceit.aut.ac.ir/autcms/labs/verticalPagesAjax/labHome.htm?id=3350532&depurl=computer-engineering&lang=en&cid=147776.

[18] The secure systems analysis lab website: http://ceit.aut.ac.ir/autcms/labs/verticalPagesAjax/labHome.htm?id=3369580&depurl=computer-engineering&lang=en&cid=147732.

[19] Patterson and Smith, *Developing a Reliable Methodology for Assessing the Computer Network Operations Threat of Iran*, pp. 29-35.

[20] For more on the center's activity in information security, see http://www.itrc.ac.ir/itrc-secure-en.php.

[21] Reference to investments in information technologies at the Technological Cooperation Office website, http://citc.ir/newpages/page27.aspx?lang=Fa.

[22] Iran Watch, "The Wisconsin Project on Nuclear Arms Control," January 3, 2011, http://www.iranwatch.org/suspect/records/technology-cooperation-office.htm.

23 The list of companies at Pardis Technology Park is available at http://www.techpark.ir/?/content/142.

24 Guilan Science and Technology Park website: http://www.gstp.ir/modules.php?name=Content&pa=showpage&pid=16.

25 *S*teve Stecklow, "Chinese Firm Helps Iran Spy on Citizens," *Reuters*, March 22, 2012, http://graphics.thomsonreuters.com/12/03/IranChina.pdf.

26 Marashi, "The Islamic Republic's Emerging Cyber War." Informational literature presenting the Separ technology and indicating the link between the regime and the technology's development may be found at http://www.iranascience.com/1-home/newsletters/21-Web%20Filters.pdf.

27 OpenNet Initiative, *Internet Filtering in Iran*, June 16, 2009, http://opennet.net/research/profiles/iran.

28 McCaney, "Iran Building a Private, Isolated Internet."

29 Robert Tait, "Iranian State Goes Offline to Dodge Cyber-Attacks," *The Telegraph*, August 5, 2012, http://www.telegraph.co.uk/news/worldnews/middleeast/iran/9453905/Iranian-state-goes-offline-to-dodge-cyber-attacks.html.

30 Cyrus Farivar, "Security Researcher Unearths Plans for Iran's Halal Internet," *Ars Technica*, April 17, 2012, http://arstechnica.com/tech-policy/2012/04/iran-publishes-request-for-information-for-halal-internet-project/.

31 Tait, "Iranian State Goes Offline to Dodge Cyber-Attacks."

32 Ali Akbar Dareini and Brian Murphy, "Iran Internet Control: Tehran Tightens Grip on Web," *Huffington Post*, April 16, 2012, http://www.huffingtonpost.com/2012/04/16/iran-internet-control_n_1429092.html?ref=world.

33 Emily Alpert and Ramin Mostaghim, "Iran's Supreme Leader Calls for New Internet Oversight Council," *Los Angeles Times,* March 7, 2012, http://latimesblogs.latimes.com/world_now/2012/03/iran-internet-council-khamenei.html.

34 "Structure of Iran's Cyber Warfare," *BBC Persian*, p. 1, http://nligf.nl/upload/pdf/Structure_of_Irans_Cyber_Operations.pdf.

35 "Iran is Formulating Strategic Cyber Defense Plan: Official," *Tehran Times*, June 15, 2012, http://tehrantimes.com/politics/98761-iran-is-formulating-strategic-cyber-defense-plan-official.

36 The center's structure and functions are described on its official website: http://www.certcc.ir/index.php?newlang=eng.

37 "Structure of Iran's Cyber Warfare", pp. 4-5.

38 "Iran to Crack Down on Web Censor-Beating Software," *Hürriyet Daily News*, September 22, 2012. http://www.hurriyetdailynews.com/iran-to-crack-down-on-web-censor-beating-software.aspx?pageID=238&nID=22789&NewsCatID=374.

39 "Structure of Iran's Cyber Warfare," p. 4.

[40] In January 2012 the regime passed a set of laws for monitoring and surveillance of web surfers at internet cafes throughout the country. These laws allow FETA to create a user log of all temporary surfers in the country and monitor anti-regime activities in cyberspace. Farnaz Fassihi, "Iran Mounts New Web Crackdown," *Wall Street Journal*, January 6, 2012, http://online.wsj.com/article/SB10001424052970203513604577142713916386248.html.

[41] Berman, "The Iranian Cyber Threat to the U.S. Homeland," p. 4.

[42] Kevin Coleman, "Iranian Cyber Warfare Threat Assessment," *Defense Tech*, September 23, 2008, http://defensetech.org/2008/09/23/iranian-cyber-warfare-threat-assessment.

[43] Stephen Trimble, "Avtobaza: Iran's Weapon in Alleged RQ-170 Affair?" *The DEW Line*, December 5, 2011, http://www.flightglobal.com/blogs/the-dewline/2011/12/avtobaza-irans-weapon-in-rq-17.html.

[44] Frank J. Cilluffo, "The Iranian Cyber Threat to the United States," A Statement before the U.S. House of Representatives Committee on Homeland Security, Subcommittee on Counterterrorism and Intelligence and Subcommittee on Cyber Security, Infrastructure Protection and Security Technologies. April 26, 2012, p. 5, http://homeland.house.gov/sites/homeland.house.gov/files/Testimony%20-%20Cilluffo.pdf.

[45] Scott Peterson, "Iran's Cyber Prowess: Could it Really have Cracked Drone Codes?" *Christian Science Monitor*, April 24, 2012, http://www.csmonitor.com/World/Middle-East/2012/0424/Iran-s-cyber-prowess-Could-it-really-have-cracked-drone-codes.

[46] Cilluffo, "The Iranian Cyber Threat to the United States," p. 5.

[47] Patterson and Smith, *Developing a Reliable Methodology for Assessing the Computer Network Operations Threat of Iran*, pp. 44-49.

[48] Iftach Ian Amit, "Cyber [Crime|War]," paper presented at DEFCON 18 Conference, July 31, 2010, http://www.defcon.org/images/defcon-18/dc-18-presentations/Amit/DEFCON-18-Amit-Cyber-Crime-WP.pdf.

[49] Khashayar Nouri, "Cyber Wars in Iran," *Institute for War & Peace Reporting*, July 23, 2010, http://iwpr.net/report-news/cyber-wars-iran.

[50] Golnaz Esfandiari, "Basij Members Trained to Conquer Virtual World," *Payvand Iran News*, August 21, 2010, http://www.payvand.com/news/10/aug/1206.html.

[51] Jeffrey Carr, "Iran's Paramilitary Militia is Recruiting Hackers," *Forbes*, January 12, 2011, http://www.forbes.com/sites/jeffreycarr/2011/01/12/irans-paramilitary-militia-is-recruiting-hackers/.

[52] Bob Beauprez, "Iranian Cyber-Attack Plot against U.S. Exposed in Mexico," *Townhall*, December 13, 2011, http://finance.townhall.com/columnists/bobbeauprez/2011/12/13/iranian_cyber attack_plot_against_us_exposed_in_mexico/page/full/.

[53] SSL is a protocol for security communications on the internet, making sure that the server a client is contacting is in fact the right server, while

encrypting the information between the browser and the server. SSL keys can be purchased from authorized providers. The theft of keys would allow the thief (with control of the network's infrastructure) to divert surfers to counterfeit websites masquerading as legal sites and thereby access confidential information about the user.

54 Report issued by Comodo Ltd., March 13, 2011, http://www.comodo.com/Comodo-Fraud-Incident-2011-03-23.html.

55 Eva Galperin, Seth Schoen, and Peter Eckersley, "A Post Mortem on the Iranian DigiNotar Attack," *Electronic Frontier Foundation,* September 13, 2011, https://www.eff.org/deeplinks/2011/09/post-mortem-iranian-diginotar-attack.

56 Fox-It, Interim Report, "DigiNotar Certificate Authority Breach 'Operation Black Tulip,'" September 5, 2011.

57 Toby Sterling, "Iran Involvement Suspected in DigiNotar Security Firm Hacking," *HuffPost Tech*, September 5, 2011, http://www.huffingtonpost.com/2011/09/05/iran-diginotar-hack_n_949517.html.

58 Gerry Smith, "Cyber Attacks Against US Banks Sponsored by Iran, Lieberman Says," *Huffington Post,* September 9, 2012.

59 The FS-ISAC is an organization whose role is to analyze and share information among financial institutions about threats to critical financial services in the United States.

60 Jaikumar Vijayan, "U.S. Banks on High Alert against Cyber Attacks," *Computerworld*, September 20, 2012, http://www.computerworld.com/s/article/print/9231515/U.S._banks_on_high_alert_against_cyber ttacks.

61 Jim Finkle, "Exclusive: Insiders Suspected in Saudi Cyber-Attack," *Reuters*, September 7, 2012, http://in.reuters.com/article/2012/09/07/net-us-saudi-aramco-hack-idINBRE8860CR20120907.

62 Kelly Jackson Higgins, "Shamoon Code 'Amateur' but Effective," *Dark Reading*, September 11, 2012, http://www.darkreading.com/advanced-threats/167901091/security/attacks-breaches/240007179/shamoon-code-amateur-but-effective.html; Nicole Perlroth, "Cyber Attack on Saudi Firm, U.S. Sees Iran Firing Back," *New York Times*, October 23, 2012, http://www.nytimes.com/2012/10/24/business/global/cyber attack-on-saudi-oil-firm-disquiets-us.html?_r=1&adxnnl=1&pagewanted=all&adxnnlx=1351084069-1i53F0BCczNEGcP8ut3n4A&.

63 Associated Press, "Panetta Hints Iran behind Gulf Cyber attacks," *CBS News*, October 12, 2012, http://www.cbsnews.com/8301-202_162-57531088/panetta-hints-iran-behind-gulf-cyber attacks.

64 Jeffrey Carr, "Who's Responsible for the Saudi Aramco Network Attack?" Blogspot, August 27, 2012, http://jeffreycarr.blogspot.co.uk/2012/08/whos-responsible-for-saudi-aramco.html.

65 Global Research & Analysis Team, "Shamoon the Wiper – Copycats at Work," *Kaspersky Lab Expert*, August 16, 2012, https://www.securelist.com/en/blog?print_mode=1&weblogid=208193786.

66 Reuters, "Iranian Hackers Attacked Three Largest U.S. Banks as Part of Cyber Campaign: Sources," September 21, 2012, http://news.nationalpost.com/2012/09/21/iranian-hackers-attacked-three-largest-u-s-banks-as-part-of-cyber-campaign-sources.

67 Perlroth, "Cyber Attack on Saudi Firm, U.S. Sees Iran Firing Back."

68 Associated Press, "Panetta Hints Iran behind Gulf Cyber Attacks."

The Proliferation of Weapons in Cyberspace

Daniel Cohen and Aviv Rotbart

Introduction

Cyberspace is a phenomenon whose fundamental nature is to utilize an electromagnetic field for human purposes by means of technology. This article argues that such technology is a type of weapon. A common dictionary definition of "weapon" is "any instrument used in combat" or "any means employed to get the better of another."[1] A "cyber weapon," therefore, is one that strikes with the purpose of vanquishing another by attacking systems connected to cyberspace. Cyber weapons can be used as non-lethal weapons and have the ability to cause tremendous destruction and serious damage without destroying physical infrastructures or human life. The cyber-strategic environment includes the use of cyber weapons in order to penetrate the enemy's systems for purposes of espionage, psychological warfare, deterrence, and damage to information technology systems or physical targets.

We distinguish between the broad and prolonged capability to attack strategic targets that have a high degree of defensive capability and an attack that is liable to cause local or temporary damage. Currently, offensive capability of the former kind is restricted to a limited number of states, and requires major resources. In contrast, the latter type of capability costs little, and consequently, there are already signs that weapons are being mass produced, are available on the open market, and are used by terrorist and criminal organizations.

Cyber warfare is rapidly becoming one of the popular offensive methods used by states seeking to protect their interests from hostile states or

Daniel Cohen is the coordinator of the Military and Strategic Affairs Program at INSS.
Aviv Rotbart is a Neubauer research fellow at INSS.

This article was first published in *Military and Strategic Affairs* 5, no. 1 (2013): 59-80.

organizations. This is apparent in the recent cyber attacks covered by the media, such as the attack, attributed to Iran, on oil companies in the Persian Gulf and on American banks; or the attacks on Iran's nuclear facilities, attributed to the United States and Israel.[2] There are a number of reasons for this, including the ability to carry out a targeted attack, the attacker's ability to camouflage itself, and the victim's ability to conceal the incident, thus avoiding the need to strike back. Cyberspace allows states with resources and high level technological capabilities to employ an arsenal of weapons for a cyber attack. Similarly, states lacking resources can also equip themselves with offensive weapons and operate in cyberspace, although on a more limited scale and with less potential for damage.

A unique aspect of cyberspace not found in other arenas of combat is the ability to defend against viruses or other malicious codes[3] that have already been used in the past and discovered by security bodies.[4] Ostensibly, cyber weapons can be used only once, as they become useless the moment they are identified and signed.[5]

That said, do all the man-years invested in developing sophisticated malicious codes go down the drain as soon as an attack is discovered and signed? This article shows that they do not. As cyber attacks increase, cyber tools and capabilities proliferate around the world. One of the main reasons for this is that cyber weapons, for example, malicious code used in one attack, can be used for other attacks as well after they are converted. In a term borrowed from the world of biology, this is called "mutated code." Such code has functional characteristics similar to the original code from which it was created (and can even be totally identical). The difference between the original code and the mutated code is syntactical (structural) only and not semantic, where it is intended to evade the radar of software that identifies attackers.

From this we can conclude that if malicious code falls into the hands of an adversary with motivation and capability, it provides the attacked party with a weapon that, if it arms itself appropriately while executing complex actions such as reverse engineering, can be exploited for repeated use.[6] In addition, an attacker who understands the weapon can use it effectively and change it according to his needs to carry out further attacks.

We are in the throes of a silent cyber war, and while very few details have been leaked to the media, the mystery cannot be maintained forever. Consider, for example, the development of the field of unmanned aerial vehicles, or drones. In its early days, the field was cloaked in secrecy. Few states had the ability to operate drones for espionage and subsequently for attack, and they made calculated and careful use of the technology in order not to reveal it to their adversaries. With the increasing use of unmanned tools, the wall of mystery has been breached, and today, thanks to the media, detailed descriptions of the countries that use drones, the targets of this type of attack, and drones capabilities and limitations are available. Terrorist organizations too have closely studied the new-old weapons that states use against them, and have developed means of defending themselves.

Another result of the extensive use of drones and the resulting media exposure is that an arms race has commenced, with many countries attempting to join the exclusive club of those in possession of these weapons for espionage and offensive purposes.[7] State supporters of terrorism have also entered the race, and terrorist organizations operating under the sponsorship of these states also enjoy the fruits of the investment. For example, Iran has acquired the ability to operate drones, and it did not take long for this capability to make its way to the Hamas and Hizbollah terrorist organizations.[8]

According to estimates, only a limited number of states currently possess the ability to carry out an attack in cyberspace in order to disrupt industrial control systems and cause physical damage, as with the Stuxnet virus, which damaged the centrifuges in the Iranian nuclear reactors, and many other states have joined the race to achieve this capability. Thus, a new type of combat weapon is being acquired for the purpose of causing damage and destruction from a great distance.

Carrying out an attack that will damage an industrial process is not overly complex, and it can be perpetrated by junior engineers. In contrast, understanding the industrial process that occurs at the target under attack and performing an in-depth analysis of it requires the full intelligence and penetration capabilities of a state.

Non-state actors in cyberspace, particularly criminal and terrorist organizations, can make use of, or already have made use of, variations of existing malicious codes and convert them so as to serve the organization's purposes. This is what happened in 2012 when criminal organizations made their own changes to two existing viruses, Zeus and SpyEye, and managed to withdraw some 78 million dollars from banks around the world.[9]

The greater the accessibility of existing codes and the greater the ability of individuals or small organizations to perform conversions and modifications, the greater the proliferation of malicious codes for attacks on the financial world and for economic gain for criminal organizations. Furthermore, these codes will also spread to terrorist organizations that wish to accomplish social, ideological, and political goals through intimidation and the disruption of normal civilian life.

Capabilities of Actors in Cyberspace

The transition from the industrial age to the information age has produced a new product in the shape of cyberspace. The development of the information age is connected to the growth of communications, control, and computer technologies, which have deep social and economic significance. The year 2008 has symbolic significance in that it was the year in which, for the first time, the number of home computers (most of them connected to the internet) passed the billion mark. That same year, it was reported that the number of people in the world possessing cell phones exceeded the number of people without cell phones. Every such computer or phone can serve as a gateway to cyberspace and a weapon for a potential attacker (or itself become a target for attack).[10]

The rapid technological developments of the information age create unique characteristics and features in cyberspace that make it possible to work quickly against adversaries located far from the attacker. These developments may also change the face of the modern battlefield, creating theaters of combat in which the non-state actor is the main actor and exerts its influence on the policy of governments and international institutions to a greater extent than in the past. For example, the fighting in Kosovo between 1996 and 1999 was dubbed "the first internet war." State and non-state actors used the internet to disseminate information and propaganda and to demonize their adversaries. Hackers used the internet during the fighting as a tool against both other former Yugoslavia states and NATO, interfering with government computer systems and taking over government websites. Individuals and activists used the web to disseminate messages from the combat zone.[11]

Another example can be found in the attacks in Estonia. Commencing in April 2007, Estonia was attacked for three weeks with a DDoS, or distributed denial of service. The wave of attacks targeted the websites of government institutions, banks, and newspapers. Since it began after a clash with Russia over demonstrations by the Russian minority in Estonia, Estonian and NATO officials hinted that there had been Russian state intervention in carrying out the attacks.[12]

Cyberspace has broad significance with regard to the use of military force, terrorist activity, organized crime, espionage, and intelligence. Concerning the use of force, an attack on computers does not require a state base; it can be carried out by organizations and even individuals. In addition, a cyber attack can also be perpetrated between friendly states competing for diplomatic and economic intelligence.

A unique trait of cyber warfare is the ability of both attacker and victim to conceal almost perfectly the fact that an attack did indeed take place. Because of the nature of cyberspace, the attacker can carry out the offensive action at a great distance from the target and use concealment techniques to prevent exposure almost entirely. The victim, for its part, can always claim that the damage to its systems was the result of a hardware or software problem, thereby avoiding tarnishing its image and responding or threatening to respond.

A direct result of the ability to hide in cyberspace is very limited media exposure of attacks. From the little that is published in the press, however, we can see an increase in the number and sophistication of cyber attacks. All the major powers are already involved in cyber warfare in one way or another, and many other countries are investing in developing attacks and defense capabilities in cyberspace.[13] Cyber warfare is being perfectly integrated into the new "Cold War" that is underway between East and West because it allows the adversary to be threatened or harmed without compelling it to respond. A cyber attack that is not reported and for which no one claims responsibility is an attack to which the victim does not feel obligated to respond; nonetheless, it is totally cognizant of the hint sent by the attacker. This is the essence of a cold war.

On the defensive side, with the expanded use of cyber weapons, there is greater awareness of the dangers of these weapons and the potential damage they can wreak in terms of security, economics, and image. As a result of this awareness, more resources are being invested in developing software systems that are better protected and more secure, as well as in securing facilities and critical infrastructures in various countries. As in any battle between attackers and defenders, in cyberspace too the attackers had the upper hand when cyber warfare began to develop. Now, however, it appears that the gap is narrowing, as more and more organizations are working to secure their IT infrastructures.

One of the characteristics of cyberspace is the difficulty in identifying the attacker. This contrasts, for example, with the attack on Pearl Harbor by Japanese Imperial Air Force bombers in 1941, which led the United States to declare war on Japan. After the large cyber attack such as that on Aramco in August 2012, the identity of the attacker is still being debated by security experts, even though an accusatory finger is being pointed at a state actor (Iran).[14] The characteristics of cyberspace also make it difficult to distinguish between intentional harm and a glitch, and to attribute an operation to a particular actor, thereby making it problematic for victims to respond to an attack. Some people argue that the characteristics of cyberspace today are still more advantageous for the attacker than for the defender.[15]

Groups that Employ Cyber Attack Tools

There are five main groups that employ cyber attack tools today or have the potential to use them in the future.[16]

States develop offensive and defensive capabilities as part of their exercise of power. Reasonable estimates are that some 40 states are acquiring cyber warfare capabilities or have already acquired them, including the ability to carry out cyber attacks. Most of the national programs are covert, and there is no consensus on the extent to which existing international law, which is valid for an armed conflict, is supposed to apply to this new type of attack.[17]

In the information age, there is increasing state intervention in the economy, civilian infrastructures, national security, civilian security, inter-organizational communication, management of government institutions, education, and so forth. As a result, countries around the world are increasing their investment in the defense of computerized systems, which is reflected in the resources allocated to the issue and to the development of specialized technologies and security concepts.[18] At the same time, defense and intelligence agencies are adopting the tools of cyberspace in order to achieve their goals. Information technologies are also providing state intelligence services with a wide range of ways and means to perform the task. States have the ability to gain access to closed computer systems by infiltrating or activating an agent and by intervening in the supply system and introducing “infected” components into the enemy target.

The same characteristics of cyberspace that make it difficult to identify the attacker can also provide the attacking state with an advantage by utilizing a proxy to carry out an attack or take responsibility for attacking a state or a business enterprise in a rival country.

For example, in state cyberspace, three new programs that employ malicious code were exposed in 2012: Flame, Gauss, and miniFlame. Flame is an example of complex malware that existed undetected for some time, and collected data and information. At 20 MB, Flame is a large program for a virus, as most viruses rely on their small size to avoid detection. The program includes properties of a Trojan horse, allowing those who activate it to open a "back door" to computer systems in order to collect information and pass it to remote servers around the world. In addition, Flame is capable of recording audio by means of the computer's microphones, of taking screen shots, and of connecting to Bluetooth devices in the area of the attack.

This type of attack, which, because of its complexity is attributed to a state, affects not only government institutions, but also businesses and the infrastructures of business enterprises that have ties with government institutions.[19]

Criminal organizations are driven mainly by criminal and business interests. Organized crime uses hackers for profit: identity theft, fraud, spam, pornography, concealment of criminal activity, money laundering, and the like. Some 80 percent of internet crime is perpetrated by criminal organizations.[20]

Former Interpol president Khoo Boon Hui claimed that banks in the United States are losing 900 million dollars every year as a result of computer crime.[21] During the first quarter of 2012, it was reported that criminal organizations had created variations of SpyEye and Zeus for an attack on banks in Europe and the United States. The attack was first identified in Italy, where the code was tailored specifically to attack different banks. Later, a similar type of attack was identified against German and Dutch banks. The attacks then spread to Latin America and the United States. The attackers managed to steal at least 78 million dollars in transfers from the accounts of some 60 financial institutions.[22]

According to the assessment of senior analysts, hackers manage to steal about one billion dollars a year from financial institutions. There are those who estimate that three of the major crime gangs operating in this field have succeeded in stealing some 100 million dollars a year by means of computer systems, while according to the FBI, in 2010, only 43 million dollars were stolen from American banks by non-cyber methods.[23]

Business enterprises mainly operate defensively since the scope of cyber attacks in the business context is growing significantly. However, some of them could elect to attack competitors for the purpose of industrial espionage – or have already done so. In addition, business enterprises face technological challenges in cyber defense such as protecting online payments, video broadcasts in real time, smartphone apps, and many others.

Terrorist organizations exploit the advantages of using cyberspace in order to pass coded messages, recruit supporters, acquire targets, gather intelligence, conceal operations, and the like. Out of cost-benefit considerations, terrorist organizations also use cyberspace to carry out cyber attacks, which help them influence public opinion so as to convey political messages and create demoralization and intimidation in order to disrupt citizens' lives. Terrorist organizations focus their offensive cyber operations on symbols of power such as the websites of government and media institutions.

One of the first documented attacks by a terrorist organization against state computer systems was carried out in Sri Lanka by the Tamil Tigers guerrilla fighters in 1998. For two weeks, Sri Lankan embassies around the world were flooded with some 800 e-mails per day saying, "We are the internet Black Tigers and we're doing this to disrupt your communications."[24] Some argue that this message induced fear at the embassies.[25] In Israel in January 2012, a group of pro-Palestinian hackers calling themselves "Nightmare" brought down the websites of the Tel Aviv Stock Exchange and El Al Airlines for a short time, and disrupted activity on the website of the First International Bank of Israel. Referring to this hacking incident, a Hamas spokesman in the Gaza Strip announced that the organization had initiated a new field of resistance against the Israeli occupation.[26]

Finally, *anarchists*, who oppose the existing institutional system, are eager to sabotage it from within or without, and will seek to attack the computer systems that are the basis for running it in order to disrupt and even destroy the social order and the fabric of life in the country. For example, groups of activists or individuals could attack websites in order to plant a political message, or endeavor to breach censorship mechanisms and reveal secrets.

In November 2012, during Operation Pillar of Defense in Gaza, government officials in Israel announced that there had been 100 million attempted cyber attacks against Israeli government internet services.[27] Anonymous, an organization that represents a theoretical concept of a community of hackers and activists, took responsibility for bringing down Israeli websites and leaking the credit card numbers of Israeli citizens during the conflict. Anonymous also published a list of more than 650 Israeli websites that it claimed were taken down or defaced as a result of the attacks by "hacktivists."[28]

A US government official has stated that "a couple dozen talented programmers wearing flip-flops and drinking Red Bull can do a lot of damage."[29] However, the ability to attack strategic targets of an enemy with advanced defensive capabilities differs from the ability to cause local, tactical damage. The various actors are acquiring cyber weapons in accordance with their capabilities and their limitations with regard to setting up a cyber force with offensive capabilities, and this has also been influenced by the interests and needs of each actor.

Table 1 charts cyber weapon capabilities of the various actors. Currently, there is a limited number of states with the capabilities and high level technological resources with the ability to use cyber weapons to attack both physical and cyber strategic targets. However, there is a low threshold of entry, and there are cyber weapons with the ability to cause tactical damage. Such weapons can be mass produced quickly and at a relatively low cost, and some of them are even available on the open market. States exploit cyberspace in order to gain an advantage and to promote their interests by collecting information, achieving the capacity to strike at the capabilities of anyone considered an enemy, and so forth. Non-state actors such as terrorist and criminal organizations can also leverage cyberspace for their purposes, and they benefit because it affords small actors influence that is disproportionate to their size.

The table shows that the state actor is capable of achieving offensive capabilities in all categories. States have diverse needs such as espionage and damaging industries in an enemy state. States also have restraints such as avoiding harm to innocents and avoiding a great deal of environmental damage. This leads to the development of cyber weapons for cyber attacks rather than physical attacks, or weapons for a psychological attack such as a warning before a bombing that makes it possible to avoid harming civilians.

Table 1. Basket of Cyber Weapon Capabilities of the Various Actors in Cyberspace

	Use of cyber weapons to attack physical equipment	Use of cyber weapons to attack in cyberspace	Use of cyber weapons for espionage	Denial of service and psychological warfare
States	Because of the large resources required, only a limited number of states today have the ability to carry out a cyber attack that causes physical damage to the defender. (According to reasonable estimates, the United States, Israel, Russia, China, and Britain have this capability.) Many other states are attempting to reach the threshold of physical attack capability or are keen to do so.	Medium-sized resources are required for this, and the number of states with electronic attack capabilities is greater than the number with the ability to attack physical equipment. States can carry out an electronic attack and/or use proxies to carry out such an attack.	The leading states in the field of industrial espionage and espionage for intelligence gathering are Russia and China, and, according to some, the United States and Israel. Since large resources are needed to make use of this capability, only a limited number of states possess it. If we assume that espionage is the second oldest profession in the world and that most states engage in espionage in one way or another, spyware to be used on targets inside and outside the country will become more common as access to technologies that provide cyber espionage capabilities increases.	This capability is relatively simple, and any state is likely to use it during a conflict with another state or by means of proxies.

	Use of cyber weapons to attack physical equipment	**Use of cyber weapons to attack in cyberspace**	**Use of cyber weapons for espionage**	**Denial of service and psychological warfare**
Terrorist organizations	Terrorist organizations today lack the resources required to realize this capability. Nevertheless, there are states that use terrorist organizations to carry out terrorist attacks, and it is therefore not inconceivable that they have been used or will be used to carry out physical cyber attacks as well.	Terrorist organizations lack the resources required to realize this capability, other than those acting as a proxy for a state.	Large resources are required to realize this capability. However, since it is one of the needs of terrorist organizations, they might attempt to acquire it (even though ostensibly, this weapon requires resources that are relatively complicated to acquire).	Used by terrorist organizations in order to disrupt routine life and sow anxiety and panic among civilians.
Criminal organizations		Used by criminal organizations in order to perpetrate financial crimes and extort business organizations and the wealthy.	Carry out the espionage activities necessary to perpetrate other crimes: identity theft, credit cards.	
Business organizations			Today, spyware is used to provide a business with an edge over a competitor.	A capability that can be exploited to harm competitors – for example, by bringing down a competitor's website or service.
Anarchists				A capability used by activists to convey messages by disrupting governmental and civilian systems.

The other actors in cyberspace have more focused interests and needs: terrorist organizations have more limited capabilities and resources, and are driven by the desire to accomplish political and ideological goals by means of damage to physical systems (even though no such incident has yet taken place), espionage, or psychological warfare. Business organizations, in contrast, are interested mainly in industrial espionage, and sometimes also in disrupting the activities of their competitors. Criminal organizations are interested primarily in obtaining assets and money fraudulently, and therefore focus on attacking cyber systems and on espionage that supports such activity (collecting credit cards and identity-linked information for an attack).

The Threat of the Repeated Use of Cyber Weapons

Every new cyber attack that is revealed brings cyber weapons closer to belonging to the public domain. As the use of cyber warfare tools increases, it is not inconceivable that more sophisticated cyber weapons with the ability to cause strategic damage will become commonplace, with various versions finding their way into the hands of state sponsors of terrorism and terrorist organizations.[30] An example of this is the Stuxnet virus attack on Iranian nuclear facilities. The attack continued in secret for several years, but the moment it was discovered, it led to the in-depth study and analysis of the virus's code and an attempt to understand everything that enabled it to be successful. The results of the analysis could have been used immediately to develop new viruses based on similar principles. The secret was exposed and the weapon disseminated. Theoretically, an analysis of malicious code by security companies and security experts could divulge the virus to various actors, ranging from states to terrorist organizations. Cyber weapons will not always remain the province of the few.

There is a belief that cyber weapons can be used only once, and that this will restrain their use and retard the development of new cyber warfare tools because it is imperative to innovate constantly and to avoid using weapons that have already been discovered and signed by protection software. This belief has not proven itself; in fact, the opposite is usually the case. In other words, there is widespread repeated use of cyber warfare tools, which undergo changes to allow them to evade the radar of protection software. Cyber attacks depend on successful exploitation of a vulnerability in the system attacked.[31] The vulnerability can reside in a software component whose code was written without sufficient attention being paid to security, in a hardware component that can be penetrated and programmed to carry out destructive actions, or in a non-secure communications protocol.

In order for a system to be considered secure, all the aspects noted must be checked and secured separately. The only thing that is required in order to penetrate and take over the entire system is a small breach in one of them. Let us suppose, for example, that there is a website that contains sensitive information and is very highly secured, so that it is not vulnerable to attacks such as XSS, SQL Injection, and the like. Let us also suppose that there is another website, unimportant and totally unsecured, on the same server on which this secure site is located. In such a case, an attack can be launched on the other site, meaning that the computer where the sites are stored can be accessed through it. Once the computer has been taken over, none of the systems protecting the secure site are relevant any longer, and the secure site is compromised.

While cyber weapons that have been discovered and signed are blocked from being used in their original form, this is still a far cry from blocking them totally and rendering all the code that was developed irrelevant. First, every offensive weapon is composed of a number of modules (software components), including the module responsible for concealing the weapon in the attacked system, various information-gathering modules, an information-storage module, and a module for sending information to the command and control servers of the weapon. If a Trojan horse is discovered and signed, some of its modules can be reused by incorporating them in the code of another Trojan horse. Such a combination creates a new attack weapon that is likely to evade the radar of the anti-virus systems. Another way to reuse malicious code is by concealing it using methods known in the world of software as obfuscation[32] and packing.[33] These can sometimes change the malicious code so that it will not be discovered by protection software. Finally, even if the code that has been discovered cannot be reused, a mutated code, which is based on similar ideas and methods of operation and exploits the same vulnerabilities as the original code, can be developed.

This claim is supported by the use of different variations of the Flame virus, which has recently been publicized in the media. Even after the original virus was discovered, various derivatives of it continued to attack the target computers until they were discovered.[34] Stuxnet, which is considered the most sophisticated virus discovered up to this point, opened the door for many others that imitate its modes of operation.[35] In fact, we can say with a high degree of probability that Flame and Stuxnet combined demonstrate in the clearest manner the ability to reuse malicious code because they have a large amount of code in common.[36] Although they were designed for completely different purposes (espionage and causing damage to industrial control systems, respectively), there are a number of functions that both must fulfill. These are penetrating the organization's computer system, concealing the existence of the weapon, analyzing the organization's network, and propagating within the network in order to find valuable target computers. Both weapons can carry out these functionalities by using the same code, which was written and checked only once.

Since the process of producing cyber weapons is long and expensive, the advantages of being able to use the same code for two different tools are enormous. However, this is a process that does not guarantee a positive result, despite the amount of effort that has been expended on it. Furthermore, even when a vulnerability is discovered, in order to exploit it and use it to penetrate the computer system, a great deal more work is required to write the appropriate code and build the files that can take advantage of the vulnerability. [37] It is also possible that no way will be found to do so because of the complexity of the vulnerability, and then further research will be necessary so as to identify another vulnerability that is easier to exploit. Therefore, when a creator of cyber weapons develops the ability to penetrate a system, his intention is to exploit it in several different scenarios and with several different tools in order to maximize the profit from his investment. However, the greater and more varied the use of a particular secret capability, the greater the chances that it will be exposed and blocked. This is a restraining factor in the considerations of the cyber weapon creator with regard to propagating the tools and using the capability in other scenarios.

On the face of it, it might be expected that after malware is discovered and the existence and exploitation of the vulnerabilities become public, the programs in which the vulnerabilities were discovered (for example, Windows Operating System) would be updated immediately and the update sent to every computer on which the system is installed, thereby rendering all computers immune to the malicious code that exploits the vulnerabilities in question. This is not what happens, however. The process of protecting systems from malicious code that has been discovered comprises four main stages: discovering the vulnerability exploited by the code; closing the gap in the system; distributing a security patch to all users of the software; and only then installing it on all computers. Closing the gap through which the malicious code infiltrated the system is complex because after this is done, the programmers must also make sure that the performance of the system has not been affected by the change that has been made. The effects of the change must be carefully examined and various test scenarios run in order to make sure that the problem has been resolved. Depending on the complexity of the system, the process could take many weeks or even months.

Furthermore, even after a security update (patch) has been developed and distributed, many people do not update their computers automatically; this is especially true of companies that have an internal communication network that is not connected to the internet. In such cases, computers on the internal network will be updated only after the individual in charge of security acquires the software update or patch from the internet in order to perform the update. For these reasons, vulnerabilities can be exploited long after they have been discovered and publicized.

There is an interesting catch-22 phenomenon associated with security updates. When Microsoft, for example, encounters a security problem in its operating system, it develops a security update and seeks to provide it to all users who have been exposed to the problem. However, the moment the update is distributed, hackers and writers of malicious code become aware of its existence. They can analyze it in order to understand which security problem it solves, and then write malicious code that exploits the security gap that Microsoft itself has revealed. Of course, the malicious code can work only in systems on which the security update has not yet been installed, but surprisingly, there are quite a few like that, belonging not only to private users who do not bother to update their computers frequently but also, and particularly, to companies whose computer personnel are responsible for taking action in order to update the company's computer system. This creates a window of several days or more during which the hackers can exploit the security gaps before they are closed.

The scenario described above is an example of the reuse of malicious code that is facilitated by the abuse of the security update distribution process. In general, Microsoft distributes security updates for its programs on the second Tuesday of each month, and this is called "Patch Tuesday."[38] The following day is called "Exploit Wednesday," because hackers analyze the security updates and begin to exploit them in order to penetrate computers that have still not been updated.

The ability to create new cyber weapons based on existing weapons or on a vulnerability that has been publicized is not always that simple and immediate. Hackers who exploit Microsoft's security updates in order to discover vulnerabilities in Windows must invest time in analyzing the patch and comparing the files that it corrects with the original files in order to identify where exactly the corrections have been made, since that is where the vulnerability lies. Finally, they must also find a way to exploit that vulnerability. This process can take anywhere from days to weeks, depending on the complexity of the patch and the determination of the hacker.

In contrast, an in-depth analysis of a sophisticated tool such as Flame would require more time and more professional and experienced personnel. In general, such an analysis is performed by states or security companies rather than by private individuals. An example is the cyber weapon, MiniFlame, which was analyzed in depth by the internet security firm, Kaspersky Lab.[39] This analysis, which took several months and required a large amount of manpower, was performed in order to develop protection against the weapon and to distribute it to the company's customers. However, the products of the analysis could serve as a basis for mutated code that utilizes similar techniques and sometimes even part of the code from the original cyber weapon. If these products were to leak from Kaspersky Labs to cyber weapon developers, it would not be surprising to discover new tools that share code with MiniFlame but are used by other attackers against other targets, and possibly even against the original creator of the weapon, in a boomerang effect.

In recent years, there has been an increase in cyber attacks that require broad and prolonged offensive capability against strategic targets with a high level of defensive capability. Only a few states have this capability today, but it is not inconceivable that this trend will persist and that other states will achieve such capabilities for both defensive and offensive purposes. The trend is also evident in the global cyber crime market.[40] In Russia, for example, there are signs indicating that organized crime organizations have begun to join forces to increase their profits by sharing data and tools.[41] The Kaspersky Lab's 2012 Security Bulletin revealed that the number of malicious code attacks on the internet among the company's clients almost doubled between 2011 and 2012 (from 946,393,693 attacks in 2011 to 1,595,587,670 in 2012). These attacks took place in 202 countries. Criminal organizations used 6,537,320 unique domains as tools for perpetrating financial attacks, some 2.5 million more than in 2011.[42]

Conclusion

Many states and non-state actors are participants in a secret arms race in cyberspace. The map of interests of the various actors indicates that different kinds of attacks in cyberspace require state actors to be prepared for a range of possible attacks. At the same time, characteristics and properties of the cyber battlefield pose dilemmas for the attacker. Cyber weapons are reusable. When an attacker uses them, it reveals its capabilities to the victim, who can then reuse them, possibly even against the attacker itself (the boomerang effect). Weapons with strategic destruction capability, such as Stuxnet, are liable to fall, or have already fallen, into the hands of terror-supporting states and terrorist and criminal organizations, and will serve as a basis for cyber attacks. Independent development of cyber attack weapons or their purchase on the black market is liable to provide these elements with the ability to cause widespread damage, even if the tools obtained in this way do not reach the level of sophistication of the cyber weapons created by advanced states.

Both the possession of cyber weapons by private entities and the resulting uncontrolled proliferation are problematic. For example, a senior security researcher claimed that Stuxnet's code is found online – and even offered to share it with others.[43] On another occasion, an expert who had analyzed Stuxnet claimed that the code was equivalent to a powerful weapon, but when asked why he did not destroy the copy in his possession, he preferred not to answer.

Aside from a discussion of ethical and moral questions, we believe that it is appropriate to implement both an intra-state and an international arrangement with regard to this issue in order to activate the regulation and enforcement mechanisms against proliferation of malicious code. Consideration should be given to limiting, and in certain cases, even banning, the possession of malicious computer codes so that they do not fall into the wrong hands. On this subject, we can perhaps learn from the war that is being waged against the illegal distribution of copyrighted intellectual property such as films and music.

Today, the arsenal of cyber weapons with the ability to cause tactical damage is reducing the procurement gap between states and non-state actors. Conversely, the gap between states with an arsenal of offensive capabilities against strategic targets on the one hand and states and actors that do not have the ability to achieve the high threshold for entry on the other is growing. It is not inconceivable that states and other actors will pursue the acquisition of cyber weapons that can cause physical damage, and there must be means of dealing with the dramatic increase in threats in cyberspace. Thus, there is an urgent need to discuss the concept of reusable cyber weapons that can be exploited for other attacks.

Notes

[1] The American Heritage Dictionary of the English Language.

[2] Mark Ambinder, "Did America's Cyber Attack on Iran Make Us More Vulnerable?" *The Atlantic*, June 5, 2012, http://www.theatlantic.com/national/archive/2012/06/did-americas-cyber-attack-on-iran-make-us-more-vulnerable/258120/.

[3] Computer codes written for the purpose of carrying out an action on a computer system, usually data theft or the disruption of processes in the system, which is run without the knowledge or approval of the system's owner.

[4] For example, when a malicious program is discovered by an anti-virus company, an electronic signature of the virus is created and sent to all the company's clients. This way, when another client is attacked by the same virus, the anti-virus program will identify the attack by the signature and block it effectively.

[5] Shmuel Even and David Siman-Tov, *Cyber Warfare: Concepts and Strategic Trends*, Memorandum No. 117 (Tel Aviv: Institute for National Security Studies, May 2012).

[6] The process of discovering the technological and engineering principles of a product by analyzing its structure and modus operandi. Generally, this process includes dismantling the product and analyzing in detail how each component works.

[7] Drone Wars UK, "Mapping Drone Proliferation: UAVs in 76 Countries," *Global Research*, September 18, 2012, http://www.globalresearch.ca/mapping-drone-proliferation-uavs-in-76-countries/5305191.

[8] William Troop, "Got Drones? The Problem with UAV Proliferation," *The World*, March 26, 2012, http://www.theworld.org/2012/03/drones-proliferation/.

[9] Dave Marcus and Ryan Sherstobitoff, "Dissecting Operation High Roller," *McAfee & Guardian Analytics*, 2012, http://www.mcafee.com/us/resources/reports/rp-operation-high-roller.pdf.

[10] Martin C. Libicki, *Cyber Deterrence and Cyberwar*, Rand Corporation, Project Air Force No. 3 (2009), http://www.rand.org/content/dam/rand/pubs/monographs/2009/RAND_MG877.pdf.

[11] Dorothy E. Denning, "Activism, Hacktivism and Cyberterrorism: The Internet as a Tool for Influencing Foreign Policy," in *Networks and Netwars: The Future of Terror, Crime, and Militancy*, ed. John Arquilla (Santa Monica: Rand Corporation, 2001), p. 240, http://www.rand.org/content/dam/rand/pubs/monograph_reports/MR1382/MR1382.ch8.pdf.

[12] Ian Trainor, "Russia Accused of Unleashing Cyberwar to Disable Estonia," *Guardian*, May 17, 2007.

[13] Even and Siman-Tov, *Cyber Warfare*.

[14] In this incident, malicious code was inserted on August 15, 2012 into the computer system of Aramco, the government-owned Saudi oil company. According to reports, some 30,000 computers were damaged.

[15] Isaac Ben-Israel and Lior Tabensky, "An Interdisciplinary Look at Security Challenges in the Information Age," *Military and Strategic Affairs* 3, no. 3 (2011): 21-37, www.inss.org.il/upload/(FILE)1333532835.pdf.

[16] Yoram Schweitzer, Gabi Siboni, and Einav Yogev, "Cyberspace and Terrorist Organizations," *Military and Strategic Affairs* 3, no. 3 (2011): 39-47, http://cdn.www.inss.org.il.reblazecdn.net/upload/%28FILE%291333532806.pdf.

[17] James A. Lewis and Katrina Timlin, "Cybersecurity and Cyberwarfare," *UNIDIR Resources*, 2001, www.unidir.org/pdf/ouvrages/pdf-1-92-9045-011-J-en.pdf.

[18] Rami Efrati and Lior Yafe, "The Challenges and Opportunities of National Cyber Defense," *Israel Defense*, August 11, 2012, http://www.israeldefense.com/?CategoryID=512&ArticleID=1557.

[19] For instance, attacks against civilian targets, including critical national infrastructures, companies that are links in a chain of access to those targets, and companies on which an attack serves an economic need.

[20] Eli Senior, "Interpol: 1,000 Cyber Attacks per Minute in Israel," *Ynet*, May 8, 2012, http://www.ynet.co.il/articles/0,7340,L-4226242,00.html.

[21] Ibid.

[22] See Marcus and Sherstobitoff, "Dissecting Operation High Roller."

[23] Greg Farrell and Michael A. Riley, "Hackers Take $1 Billion a Year as Banks Blame Their Clients," *Bloomberg*, August 5, 2011, http://www.bloomberg.com/news/2011-08-04/hackers-take-1-billion-a-year-from-company-accounts-banks-won-t-indemnify.html.

[24] Dorothy E. Denning, "Cyberterrorism: Testimony before the Special Oversight Panel on Terrorism," Committee on Armed Service, US House of

Representatives, May 23, 2000, http://www.cs.georgetown.edu/~denning/infosec/cyberterror.html.

25 See Denning, "Activism, Hacktivism and Cyberterrorism," p. 269.

26 Guy Grimland et al., "Cyber Attack," *The Marker*, January 16, 2012.

27 Or Hirshauga and Nati Tucker, "Cyber Wars against Israel: 100 Million Attacks, No Significant Achievements," *The Marker*, November 22, 2012, http://technation.themarker.com/hitech/1.1871058.

28 John D. Sutter, "Anonymous Declares Cyberwar on Israel," *CNN*, November, 2012, http://edition.cnn.com/2012/11/19/tech/web/cyber-attack-israel-anonymous/index.html?hpt=hp_c1.

29 See Even and Siman-Tov, *Cyber Warfare*, p. 19.

30 For example, Hizbollah's cyber program. See Ward Carroll, "Hezbollah's Cyber Warfare Program," *DEFENSETECH*, June 2, 2008, http://defensetech.org/2008/06/02/hezbollahs-cyber-warfare-program/.

31 Vulnerability is a characteristic of a software/hardware/protocol component that makes it possible to use the component for a purpose other than the one for which it was intended, in a way that is advantageous to the person exploiting this feature. The advantage can be obtained in one or more of the following ways: taking control of the system, disrupting the system, or obtaining information from the system.

32 Code obfuscation is a technique from the software world that takes existing computer code that is intended to carry out a certain task and changes it in such a way that its functionality is not harmed, but the result is sufficiently different from the original so that an anti-virus program will not be able to identify it as a virus. Anti-virus programs that are based on identifying signatures in the code (a signature in this context is from a section of code intended to carry out a particular action, which can be attributed to a malicious program with a high degree of probability) will find it difficult to identify as a virus the code that was successfully obfuscated because none of the signatures known to it will appear in the result of the obfuscation process.

33 Code packing is a sophisticated type of code obfuscation. In the packing process, malicious computer code undergoes a radical change in form so that it no longer looks anything like a running code, but more like an innocent text file. This method almost completely prevents the anti-virus programs from discovering the malicious code before it begins to carry out its operation (for example, during the virus's penetration of the computer, it will not be discovered). Packing code works through an innocent utility that, when it starts to run, calls the text file in which the malicious code is hiding, translates the text into run commands, and in fact, turns itself into a virus. This can be compared to a virus from the world of biology, which takes over a living cell and exploits all of the cell's mechanisms for its needs.

[34] Renana Ashuah, "Kaspersky Exposes miniFlame—Malicious Code Planned for Espionage Operations," *YedaTech*, October 15, 2012, http://www.yedatech.co.il/yt/news.jhtml?value=19827.

[35] For an article on Stuxnet's successors, see Steven Cherry, "Sons of Stuxnet," *IEEE Spectrum*, December 14, 2011, http://spectrum.ieee.org/podcast/telecom/security/sons-of-stuxnet.

[36] On Flame, see Aleks, "The Flame: Questions and Answers," *SECURELIST*, May 28, 2012, http://www.securelist.com/en/blog/208193522/The_Flame_Questions_and_Answers.

[37] Exploits are computer codes or files intended to exploit a vulnerability in a particular system in a manner that enables the writer of the exploit to penetrate or disrupt the system under attack. An example would be a program for viewing images on a computer screen that has a particular vulnerability that allows a code to be run on the computer under attack. Such a vulnerability is likely to be exploited in the form of an image file that includes code the attacker is eager to run on the computer under attack. Of course, such an image file must not only contain the code, but must also know how to exploit the vulnerability or the weak point of the image viewing software.

[38] A patch is a system update.

[39] Global Research and Analysis Team, Kaspersky Labs, "MiniFlame aka SPE: Elvis and His Friends," *SECURELIST*, October 15, 2012, http://www.securelist.com/en/analysis/204792247/miniFlame_aka_SPE_Elvis_and_his_friends.

[40] This market was estimated in 2011 at more than 12.5 billion dollars, with Russia's portion of the cake some 2.3 billion dollars (nearly double its absolute value the previous year). See Group-IB, "State and Trends of the Russian Digital Crime Market, 2011," http://group-ib.com/images/media/Group-IB_Report_2011_ENG.pdf.

[41] Frank J. Cilluffo, Sharon L. Cardash, and George C. Salmoiraghi, "A Blueprint for Cyber Deterrence: Building Stability through Strength," *Military and Strategic Affairs* 4, no. 3 (2012): 3-23, http://cdn.www.inss.org.il.reblazecdn.net/upload/%28FILE%291362315050.pdf.

[42] Denis Maselnnikov and Yuri Namestinkov, "Kaspersky Security Bulletin 2012: The Overall Statistics for 2012," *SECURELIST*, December 2012, http://www.securelist.com/en/analysis/204792255/Kaspersky_Security_Bulletin_2012_The_overall_statistics_for_2012.

[43] The authors were present at a meeting with a figure from the security company.

The Classic Cyber Defense Methods Have Failed – What Comes Next?

Amir Averbuch and Gabi Siboni

Introduction

The classic defense methods employed throughout the world in recent decades are proving unsuccessful in halting modern malware attacks that exploit unknown (and therefore still unsolved) security breaches called "zero-day vulnerabilities." Viruses, worms, backdoor, and Trojan horses (remote management/access tools – RATs) are some examples of these attacks on the computers and communications networks of large enterprises and providers of essential and critical infrastructure and services.

The classic defense methods, which include firewall-based software and hardware tools, signatures and rules, antivirus software, content filters, intruder detection systems (IDS), and the like, have completely failed to defend against unknown threats such as those based on zero-day vulnerabilities or new threats. These sophisticated and stealth threats impersonate reliable and legal information and data in the system, and as a result, the classic defense methods do not provide the necessary defense solution. The current defensive systems usually protect against known attacks, creating heuristic solutions based on known signatures and analysis that are already known attacks,[1] but they are useless against the increasing number of unfamiliar attacks that lack any signature. Solving this problem requires different thinking and solutions. This article proposes an up-to-date approach, based on an analysis of sensitive information that must be protected, for the purpose of identifying anomalous behavior.[2] The

Prof. Amir Averbuch is a faculty member at the Blavatnik School of Computer Science at Tel Aviv University and a researcher in the Cyber Warfare Program at INSS. Dr. Gabi Siboni is the head of the Military and Strategic Affairs Program and head of the Cyber Warfare Program at INSS.

This article was first published in *Military and Strategic Affairs* 5, no. 1 (2013): 45-58.

analyzed information includes an organization's data silos as a means of understanding unusual (anomalous) activity that in most cases indicates the presence of malware in the system. The article further proposes relying on the data to be protected as a source of knowledge for developing the defense system. An analytical analysis of massive data (big data analytics) will make it possible to identify such malware, while constructing a model that will provide a high degree of reliability in identifying and minimizing false positives, which pose a challenge to every defense system.

Development of Threats and the Limitations of the Traditional Defense Systems

The first cyber attacks on computer systems were based on viruses or worms that reproduced themselves and spread rapidly. Antivirus technology, however, completely failed to detect Trojan horses, whose behavior was entirely different than that of viruses. Traditionally, defense systems were developed to protect against known viruses, because it is quite difficult to identify such viruses by their behavior rather than their signatures. In this way, it became possible to create a database of virus signatures, and to compare files and communications reaching computers with these signatures. This approach required manufacturers of defensive software to continually monitor the development of viruses in order to create their signatures and distribute updates to their customers for the purpose of enabling them to update as quickly as possible the systems on which the protective software based on these signatures was installed. The burgeoning development of various forms of viruses and malware and the enormous growth in their number rendered this process virtually impossible, because major investments of resources in the continual updating of signature data for antivirus software were required.

The cyber attack hazards can be roughly divided into the following families: malware, spyware, worms, and Trojan horses (which open "backdoors"[3]). A classification that relates more to the object of an attack includes advanced persistent threats (APTs), which began with countries launching cyber attacks against other countries' military networks and the networks of government agencies, and in recent years developed into an attack by one country directed at another's organizational network of critical civilian infrastructure, and attacks against computer-operated industrial supervisory control and data acquisition (SCADA) systems – such as the

Stuxnet attack. Essential infrastructure systems controlled by industrial control systems in which control is exercised by the SCADA protocol are therefore exposed to attacks that are liable to paralyze the essential services, and could even suffer physical damage. Other possibilities include attacks against wireless systems and mobile broadcasting stations, the use of social networks for the purpose of spreading spyware and malware, and an attack against storage and cloud computing services.

The realm of attack in cyberspace can be divided into two types of attacks that exploit numerous weaknesses, including zero-day vulnerabilities:

a. *Broadcast attacks* are attacks that try to damage computers indiscriminately. They also feature extensive infection of software agents in order to create an entire network of computers (Botnet), with the aim of making these computers execute independent commands at a later stage or retrieve commands from a control server. As noted above, when information about new threats reaches the antivirus companies, they identify the signature or investigate them heuristically. By means of regular updates, the computers can be protected against these attacks. Given the extensive target community, the information about such threats will undoubtedly reach the relevant companies rapidly and be inserted into future versions of their products. In some cases, the goal of an attack of this kind is to reach a large number of computers – for example, employees (in the case of an attack against an organizational network) or customers (in the case of an attack against a financial institution, an attempt to steal credit cards via the internet, and so on). After the computer is infected, a Trojan horse is installed on it, making it possible to steal information or access the computer from a remote location. These attacks include various types of malicious code, even codes that vary from one infection to another in order to render identification through a signature more difficult (polymorphic viruses). There is still no complete defense since Trojan horse developers regularly check whether the antivirus software programs have already identified the hostile code and created the signature or group of heuristic rules to intercept it. In most cases, if the detection systems manage to identify the hostile code, the developers change the way it spreads or the way it operates in order to prevent its detection. In this way, many Trojan horses consistently succeed in evading detection by the leading defensive software.

b. *Targeted attacks* are planned especially for a specific need, and exploit unknown weaknesses in the operating systems or widely known software packages while independently spotting new weaknesses. The vast majority of antivirus software, which is by nature based on signature defense, is incapable of identifying and preventing this type of attack, and the limited target community enables such attacks to evade the "radar" of antivirus manufacturers. It should be noted that threats are rapidly developing in the direction of focused attacks on high caliber targets.

The volume of data transmitted on a modern communications network is very large, owing to the need to provide many services to various kinds of end stations, including PCs, work stations, servers, switches and communications equipment, and many other diverse units. Such networks have many users, most of whom have no security awareness at all. As a result, APT attacks focus on people as well as on machines – via social networks, for example. The attack on the RSA company, which targeted the people in the organization, succeeded in penetrating the most secure systems.[4]

In recent years, we have seen a dramatic rise in the volume of new, undocumented, sophisticated attacks of a stealth nature. This is reflected both in the group of general attacks and in focused attacks. These attacks are overcoming all the classic standard defenses of the companies currently leading the protection sector. Major investments by countries and organized crime are responsible for the development of these attack methods, and the resulting damage is extensive.[5] The quantity of malware successfully penetrating all the existing defense systems and overcoming all the signature and rule-based classic defenses is increasing by leaps and bounds. The rate of increase has been in the three-digit percentages from 2011 until the present time.[6]

The existing systems are based mainly on preventing and thwarting known threats through the use of signatures and rules that are known in advance. Having no known signature at any given moment, these systems cannot detect zero-day attacks. They also find it difficult to identify Trojan horses and backdoors, and many sophisticated stealth attacks have no known signatures. Because they appear to be legal data and code, and do not look like malware, they can penetrate almost any computer system. The attacks succeed in penetrating organizational networks and end-user computers despite all the defense systems; this is attributable to the fact that the initial appearance and behavior of the malware appears to be legal and proper. Furthermore, most of today's operating systems are built to handle a certain kind of attack, and are unable to deal with a broad range of attacks with mutations and secondary attacks.

In conventional software, one way of detecting unfamiliar and unsigned attacks is by identifying abnormal behavior of codes residing in the organizational systems, which differs from the way most normal data behave. This different behavior is what betrays hostile codes. The notion of the irregular behavior of a software element attempting to conduct unauthorized activity could serve as a possible basis for identifying and preventing attacks. Software producers worldwide understand the challenge and are taking steps to furnish such identification capabilities. This, however, is where the most significant challenge lies, namely, the difficulty in providing a reliable tool that will not produce false alarms or affect the user experience in an extremely negative manner. False alarms, which constitute one of the most significant challenges in defense systems, are created when the system issues a warning for a legal code with normal behavior and defines it as a hostile or suspicious code. If the load of such false alarms is too heavy, it will significantly harm the working capability of the computer systems, and is liable to cause the user to lose confidence in the defense system.

The second challenge is finding a solution for malicious code that evades the defense system. This phenomenon is called a false negative – when a result is obtained that appears negative, but is actually positive (comparable to a bearer of a serious virus who receives a negative test result from a laboratory when the virus is actually present in his body). These two challenges lie at the heart of defense systems in general, particularly in the use of analysis of the anomalous behavior of hostile code in an information system.

Identifying Anomalies as an Approach to an Operative Solution

This article focuses on the protection-based detection of anomalies in communications networks at various levels. The problem is broader, however, and includes the need to identify anomalies of hostile codes that have penetrated weak points in software programs and applications. This approach is not discussed in the present article, unless the hostile code is exposed in the organizational communications. Regardless of the above, one can assume that some of the ideas mentioned are also suitable for detecting anomalies in software and applications.

Anomalies first proposed in 1987[7] are deviations from the expected behavior, which is the normal behavior. The basic assumption for any system seeking anomalies posits that malicious data have characteristics that are not found in the normal behavior specified during the learning phase. Since 1987, additional theories and methodologies have been developed, based on machine learning approaches and on the theory of information,[8] such as nervous systems,[9] a support vector machine,[10] genetic algorithms,[11] and many others. There are also numerous approaches that utilize data mining in order to find hostile code.[12] A general review of finding anomalies appears in an article by Chandola and Banerjee,[13] and there is a study of methods for spotting hostile code.[14]

One approach to detecting attacks on data from communications networks entails monitoring anomalies in network activity by finding the deviation from a normal profile learned from benign (proper non-malware) data. This methodology is based on tools retrieved from studies in machine learning,[15] mathematical and stochastic analysis,[16] statistics, data mining, graph theory, information theory, geometry, probability theory and random processes, and so on. Machine learning and data mining tools, combined with the above methodologies, are used successfully in many other fields, such as systems for recommending Amazon products,[17] Netflix,[18] optical character recognition,[19] translation of a natural language,[20] and identifying junk e-mail (spam).[21] Machine learning deals with the development of algorithms that enable a computer to learn, based on examples. Supervised learning of data known in advance, in which the correct significance of the parameters is known ahead of time, namely, labeled data, already exists. In unsupervised learning, the goal of the algorithms is to find a simple representation of the data without labels. Supervised learning is more limited with respect to the data content being learned. On the other hand, the results are more reliable, and it is therefore preferable.

Learning first takes place with a "healthy" group of data, which presumably contains no malware at all. This is called the "training set." It is usually best for the learning method to be able to detect whether part of the training set contains malware up to a given percentage of all the data. Obviously, if most of the training set contains malware, it will be identified as normal data. As part of the filtering process, a process called "outlier removal" is used, which removes data that appear to be noise or infected from the training set.

The training set is analyzed by a variety of existing mathematical methods combined with innovative methods. The normal characteristics of the examined data can be identified through this process. This type of learning is called "one class." Another method, in which the characteristics are learned through comparison with a training set containing both clean and unclean data (e-mail with and without spam, for example) is called "binary class." The training set is derived from a mass of data accumulated and protected in an organization, together with continually guarded new data. For this purpose, methods of learning the data characteristic of normal behavior have been developed. While understanding the geometry of the learned data is one of the analysis methods, other methods also exist. For example, the following process describes a possible general structure of algorithms used as well as the processors of the training set in order to find the characteristics of normal (proper) behavior:

a. Breaking down each basic unit of communications or event data into characteristics (features, parameters).

b. Quantifying the relationships among the characteristics. There are a number of methods of characterizing such relationships. The kernel method[22] is one of the most common methodologies for defining them. Mathematical distance functions are usually used to define these relationships, which are near/far relationships with a range of characteristics existing between them. After this stage, the relationships between the communications data or events are guarded.

c. Lowering the dimension of the data. The dimension of the data is usually high, and is determined according to the number of characteristics making up a basic communications unit or basic event unit. The dimension of the data[23] is therefore lowered (from ten dimensions to two, for example), while preserving the relationships and coherence among the characteristics that were identified at the preceding stage. This is similar to sampling, in which only a small, reliably representative part of the original data is logically selected. Mathematical, algorithmic, and conceptual innovation is required in order to process data from a high dimension that will suit a computer and reliably represent the original data. The sampling, which is aimed at reducing the volume of data, can be random, and it can be proved that the coherence of the data is maintained. There are many mathematical methods for achieving this objective. One of the methods for streamlining the computations in order to construct a compact representative of multi-dimensional data is the construction of dictionaries in order to speed up calculations while maintaining the relationships and features identified before the dimension was lowered. Other methods for speeding up computations facilitate sparsification of the data. The goal of these approaches is to specify a normal profile for the data from the training set while overcoming heavy computational problems in processing the training set. The learning action is usually computationally heavy. This action is conducted offline, and need not take place in real time. Common methods include PCE,[24] LLE,[25] ISOMAP,[26] and so forth.

The methods described above make it possible to effectively process the training set, which is "heavy" and liable to make calculations impossible. The goal of processing the training set is to specify the training data's ordinary (normal) behavior, based on an examination of the training set and the relationships defined between the characteristics of the data and the events of the training set. This assumes that the learning and the conclusions derived from it will reflect the normal behavior of all the future new data that are not part of the training set. As the volume of data in the training set increases and its characteristics become more numerous and diverse, the normal behavioral characteristics derived from the training set become more reliable. The calculation is more complicated, however, and it is therefore necessary to invest a great deal of effort in producing algorithms that are computationally effective and can handle large volumes of data.

The process described above specifies a possible learning model that generates a specification of the normative behavior of future data with the help of the training set's normal profile. From there on, the characteristics of all new information arriving, or of a new event, are examined. These characteristics are processed in order to see whether they deviate from the normative profile learned and determined during the learning (an anomaly). Deviations from the normal profile make it necessary to identify the attacks characterized as zero-day attacks. The method described thus far does not use signatures; it finds behavioral deviations from the normal profile generated by processing the training set.

Figure 1 is a procedural description of the learning process described above. The chart also presents the range of sources from which the information has been retrieved for the purposes of the initial learning.

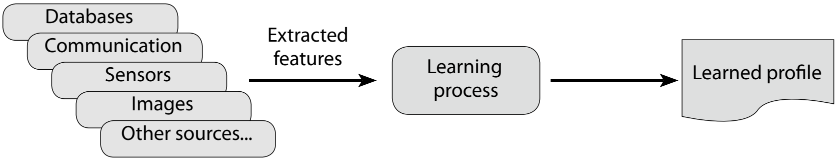


Figure 1. The Learning Process Chart

These methods and their derivatives for finding malware by monitoring the behavior of the data can be used in two different and complementary ways. The common denominator in these two ways consists of offline learning of the communications data from the protocol through which the data reach the organization (for example, port 443 [HTTPS], UDP port 53 [DNS], TCP, and TCP port 80 [HTTP], which are also web protocols) and constructing a profile that describes the normative behavior of the data of a given protocol that must be checked, according to the training set.[27]

a. *Operation in real time*. The algorithm for finding anomalies in communications data (accomplished in software or hardware) is located at the entrance to the organization. After data pass through the ordinary IPS Firewalls and IDS defense tools (signatures and rules allow them to enter), the algorithm checks each communications unit – whether its behavior matches the normal profile learned from the training set. If it proves to be an anomaly, its path into the organization is blocked. Since signatures are not used, the analysis of the substance of the anomaly can be performed either automatically or manually.

b. *Offline operation* – finding malware offline. Communications data that entered the organization through all the defense systems appear to be legal data, and subsequently begin to operate. An example of this is a spyware network absorbed into the environment with the aim of operating in the future. For this purpose, logs and events that occurred previously and are occurring now should be processed. In order to process information from both preserved and newly arrived logs, security information and event management (SIEM) technology is used. SIEM, an information security monitoring system commonly used in organizational networks, serves as a central location for preserving and decoding logs and events of communications data. SIEM, an archive of all the communications data and events, helps conduct forensic analysis in order to find anomalies.

The above-mentioned methods of finding anomalies can be applied to the data collected by SIEM. Other data mining tools can also be applied to the SIEM data. SIEM contains two functions for security management: security information management (SIM) and security event management (SEM). The method that employs SIEM data should constantly apply the methodology for finding anomalies in order to identify the operation of malware when it is activated at some future date.

Figure 2 describes processes for checking information, given the results of the learning analysis:

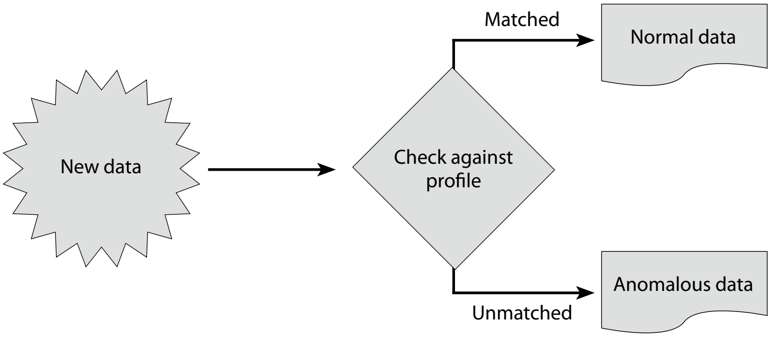


Figure 2. The Identification Process Chart

The Use of Big Data to Find Anomalies: The Data and Events Dictate the Identification Method

As described above, the main idea on which finding anomalies is based is specifying the behavior of the data in the training set and drawing conclusions from it with regard to the behavior of the data that did not participate in the training set, that is, characterizing the newly arrived data. In other words, the data dictate the processing, as reflected in the algorithms whose task was to learn the data as they are, and to adapt to them. This is in contrast to all the existing defenses against malware, which seek patterns of already familiar malware and are unrelated to the behavior of the data. In the case of communications data, the data from each information unit of the protocol being monitored are analyzed. The relationships between the data are found by using the kernel method, and they are stationed in non-linear fashion in spaces with a lower dimension. The dimension of the data, which is usually high, is lowered in this way, thereby creating an effective way of finding anomalies.

Today, the data in which we look for anomalies are referred to as "big data," that is, a huge volume of data collected from all the information sources available on the organizational network. In many organizations, they are guarded by SIEM methodology. According to former Google CEO Eric Schmidt, the quantity of data created between the dawn of civilization and 2003 was five exabytes.[28] Schmidt asserts that this quantity is now created every two days. The following are a number of examples of the creation of big data every single day: the New York Stock Exchange (NYSE) creates one terabyte of data, Facebook creates 20 terabytes of compressed data, and the CERN particle accelerator in Switzerland creates 40 terabytes of data. According to a published report,[29] the volume of data doubles every year, and at least half of all businesses keep their data for at least three years for analytic purposes. Some of them are legally required to keep these data for a number of years. New sources of enormous quantities of data are constantly emerging in various businesses such as utilities. The bulk (80 percent) of these data is unstructured, which means that the organization is therefore unable to use them effectively. Big data have become a source of data mining that facilitates the identification of malware. Many well known companies such as Facebook, Google, Amazon, LiveJournal, and Wikipedia possess quotidian big data, and this list is far from complete. Today, big data are kept in the cloud. The quantity of data stored in each organization is huge, and is constantly growing. In order to handle large data silos, tools have been developed for processing big data that are unrelated to data mining or finding anomalies, such as Hadoop,[30] MapReduce,[31] and Memcached[32] – enormous parallel databases[33] that facilitate rapid data queries. In addition, many communications "pipelines" are being developed (by the Mellanox company for instance) for high speed transmission of these quantities of data. A great deal of effort is being expended on developing advanced tools for effective processing of big data. Big data can therefore serve as a source for finding a broad range of sophisticated behavioral anomalies of different varieties of malware.

Conclusion

In order to process big data and effectively identify "high quality" malware, it is necessary to combine all the methods listed above. Tools – most of which are non-linear – were mentioned for reducing the volume of multi-dimensional big data without affecting the coherence of the data, at the same time maintaining the efficiency of the algorithms, for the purpose of handling huge volumes of data. The methods mentioned in this article that should be added are: learning from a small group of data; and using the kernel method on data, thereby determining the relationships (distances) between the sample points and reducing the dimension of the data by means of discrete or random sampling. This thins out the data, thereby obtaining an effective "housing project" of multidimensional big data in a significantly lower dimensional space in which anomalies are identified. Constructing dictionaries and using sophisticated and effective algorithms, together with big data processing tools, create many possibilities for finding malware in any organization by specifying the normative behavior and identifying deviations from it.

The proposed approach is a combination of computationally effective big data analysis and advanced tools for finding anomalies that are malware of zero-day attacks that do not yet have known signatures and behavior patterns. The methodology discussed here requires finding a needle in a haystack of data.[34] The point of departure states that the proposed algorithms adapt themselves and become accustomed to the data themselves. The data dictate how the algorithm operates. The methodology proposed in the article combines an understanding of the data structure by learning from a small group and drawing conclusions about the future behavior of the data that were not included in the learning set. This methodology is capable of detecting both malware whose activity is immediate, and malware, such as Trojan horses, that has entered the organization and will become operational at a later date.

Notes

[1] "Heuristically" means through rules that help detect the harmful code.

[2] Anomalous behavior of software code or information is unusual (uncharacteristic) behavior that arouses suspicion of malware in a system.

[3] A security breach facilitates access to a computer without the need to verify an identity. This can result from a software error, a deliberate breach in the original code, or the installation of special software (such as a Trojan horse).

[4] Gabi Siboni and Y. R., "What Lies Behind Chinese Cyber Warfare," *Military and Strategic Affairs* 4, no. 2 (2012): 43-56.

[5] Symantec, "Internal Security Threat Report," *2011 Trends* 17, April 2012.

[6] "FireEye Advanced Threat Report – 1H," *Source* 2012, http://www2.fireeye.com/advanced-threat-report-1h2012.html.

[7] D. E. Denning, "An Intrusion-Detection Model, IEEE Trans," *Software Engl* SE-13, no. 2 (1987): 222-32.

[8] W. Lee and D. Xiang, "Information-Theoretic Measures for Anomaly Detection," in *Proc. IEEE Symposium on Security and Privacy* (2001).

[9] Z. Zhang, J. Li, C. Manikopoulos, J. Jorgenson, and J. Ucles, "HIDE: A Hierarchical Network Intrusion Detection System Using Statistical Preprocessing and Neural Network Classification," in *Proc. IEEE Workshop on Information Assurance and Security* (2001).

[10] W. Hu, Y. Liao, and V. R. Vemuri, "Robust Anomaly Detection Using Support Vector Machines," in *Proc. International Conference on Machine Learning* (2003).

[11] C. Sinclair, L. Pierce, and S. Matzner, "An Application of Machine Learning to Network Intrusion Detection," in *Proc. Computer Security Applications Conference* (1999).

[12] M. A. Siddiqui, *Data Mining Methods for Malware Detection*, PhD dissertation, University of Central Florida (2008).

[13] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Computing Surveys (CSUR)* 41 no. 3, Article 15 (2009).

[14] N. Idika and A. P. Mathur, "A Survey of Malware Detection Techniques," Department of Computer Science, Purdue University (2009).

[15] R. Sommer and V. Paxon, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection," in *Proc. IEEE Symposium on Security and Privacy* (May 2010).

[16] Stochastic processes are processes whose development over time includes a certain element of randomness at any given moment.

[17] G. Linden, B. Smith, and J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet Computing* 7, no. 1 (2003): 76-80.

[18] J. Bennet, S. Lanning, and N. Netflix, "The Netflix Prize," in *Proc. KDD Cup and Workshop* (2007).

[19] L. Vincent, "Google Book Search: Document Understanding on a Massive Scale," Proc. International Conference on Document Analysis and Recognition, 2007; R. Smith, "An Overview of the Tesseract OCR Engine," in *Proc. International Conference on Document Analysis and Recognition* (2007).

[20] F. J. Och and H. Ney, "The Alignment Template Approach to Statistical Machine Translation," *Comput. Linguist* 30, no. 4 (2004): 417-49.

[21] P. Graham, "A Plan for Spam," in *Hackers &* Painters: Big Ideas for the Computer Age (O'Reilly, 2004).

[22] B. Scholkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond* (Cambridge: MIT Press, 2002).

23 M. Elad, *Sparse Redundant Representations: From Theory to Applications in Signal and Image Processing* (New York: Springer, 2010).

24 I. T. Jolliffe, *Principal Component Analysis* (New York: Springer, 1986).

25 S. T. Rowels and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science* 290, no. 5500 (2000): 2323-26.

26 J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A Global Geometric Framework for Non-Linear Dimensionality Reduction," *Science* 290, no. 5500 (2000): 2319-23.

27 This approach also facilitates performance monitoring, an analysis of users' behavior, an analysis of man-machine relationships, and control of processes.

28 1 exabyte = 1 billion billion bytes.

29 M. G. Siegler, "Eric Schmidt: Every 2 Days We Create as Much Information as We Did up to 2003," *TechCrunch*, August 4, 2010, http://techcrunch.com/2010/08/04/schmidt-data/.

30 Web page: hadoop.apache.org.

31 J. Dean and S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," *OSDI* (2004).

32 L. Gavish, *New Caching Policies for MEMCACHED*, MSc Thesis, Tel Aviv University (2012); B. Fitzpatrick, "Distributed Caching with MEMCACHED," *Linux Journal*. 2004, no. 124 (2004): 5.

33 Hadapt, http://hadapt.com/.

34 M. Baker, D. Turnbull, and G. Kaszuba, "Finding Needles in Haystacks (the Size of Countries Blackhat)," Amsterdam, The Netherlands, March 14-16, 2012.

INSS Memoranda, April 2012 – Present

No. 127, May 2013, Zvi Magen, *Russia and the Middle East: Policy Challenges*.

No. 126, April 2013, Yehuda Ben Meir and Olena Bagno-Moldavsky, *The Voice of the People: Israeli Public Opinion on National Security 2012*.

No. 125, March 2013, Amos Yadlin and Avner Golov, *Regime Stability in the Middle East: An Analytical Model to Assess the Possibility of Governmental Change* [Hebrew].

No. 124, December 2012, Shlomo Brom, ed. *In the Aftermath of Operation Pillar of Defense: The Gaza Strip, November 2012*.

No. 123, December 2012, Shlomo Brom, ed. *In the Aftermath of Operation Pillar of Defense: The Gaza Strip, November 2012* [Hebrew].

No. 122, September 2012, Emily B. Landau and Anat Kurz, eds., *Arms Control Dilemmas: Focus on the Middle East*.

No. 121, July 2012, Emily B. Landau and Anat Kurz, eds., *Arms Control Dilemmas: Selected Issues* [Hebrew].

No. 120, July 2012, Meir Elran and Alex Altshuler, eds., *The Complex Mosaic of the Civilian Front in Israel* [Hebrew].

No. 119, June 2012, Meir Elran and Yehuda Ben Meir, eds., *Drafting the Ultra-Orthodox into the IDF: Renewal of the Tal Law* [Hebrew].

No. 118, June 2012, Zvi Magen, *Russia in the Middle East: Policy Challenges* [Hebrew].

No. 117, May 2012, Shmuel Even and David Siman-Tov, *Cyber Warfare: Concepts and Strategic Trends*.

No. 116, April 2012, Yoel Guzansky, *The Gulf States in a Changing Strategic Environment* [Hebrew].