DCAF Geneva Centre for Security Sector Governance

Introduction to Computer Security Incident Response Teams (CSIRTs):

Structures and Functions of Cybersecurity's First Responders

By Nebojša Jokić

This knowledge product was produced by DCAF – Geneva Centre for Security Sector Governance, in the context of the project, 'Good Governance in Cybersecurity in the Western Balkans', supported by the UK's Foreign, Commonwealth and Development Office (FCDO)

March 2023



About DCAF

DCAF – Geneva Centre for Security Sector Governance is dedicated to improving the security of states and their people within a framework of democratic governance, the rule of law, respect for human rights, and gender equality. Since its founding in 2000, DCAF has contributed to making peace and development more sustainable by assisting partner states, and international actors supporting these states, to improve the governance of their security sector through inclusive and participatory reforms. It creates innovative knowledge products, promotes norms and good practices, provides legal and policy advice and supports capacity-building of both state and non-state security sector stakeholders. DCAF's Foundation Council members represent over 50 countries and the Canton of Geneva. Active in over 70 countries, DCAF is internationally recognized as one of the world's leading centres of excellence for security sector governance (SSG) and security sector reform (SSR). DCAF is guided by the principles of neutrality, impartiality, local ownership, inclusive participation, and gender equality. For more information visit www.dcaf.ch and follow us on Twitter @DCAF_Geneva.

DCAF – Geneva Centre for Security Sector Governance
Maison de la Paix
Chemin Eugène-Rigot 2E
CH-1202 Geneva, Switzerland
Tel: +41 22 730 94 00
info@dcaf.ch
www.dcaf.ch
Twitter @DCAF_Geneva
Design & layout: DTP studio
ISBN: 978-92-9222-695-4

Contents

Foreword **7**

Preface **8**

What is a CSIRT? **9**

A bit of history 9

The first CERT 10

CERT, CSIRT, CIRT: Different names, one objective 10

Purpose 11

SOCs 11

Position within an organization 11

Constituents 12

Basic terminology 12

Tasks performed by CSIRTs **13**

Reactive tasks 14

Preventive tasks 14

Vulnerability assessment 14

Penetration testing 15

Threat hunting 16

Quality management tasks 16

Awareness raising 16

Cyber hygiene 17

The organization of a CSIRT **18**

Types of CSIRTs 18

National CSIRTs 18

RFC 2350: CSIRT description document 18

CSIRT operations framework **19**

Procedures 19

Operational capacities 19

SIM3 model 19

Legal issues 20

Cooperation with law enforcement agencies 21

The work environment **21**

Duties and expectations 21

Code of conduct 22

Non-disclosure agreements 22

Due care and due diligence 22

Ethical principles 22

Necessary knowledge and skills 23
Cybersecurity frameworks 24
Courses 26
Cyber exercises 27
Cyber range platforms 27
CTF 28
Information collection 28
Information handling 29
Storage 29
Categorization 29
Information sharing in the CSIRT community 30
What is exchanged and how? 30
International associations 32
Gathering information from reliable sources 32
Open-source intelligence (OSINT) 33
What are we protecting? 34
The supply chain 34
The three basic security properties 34
What are we defending against? 35
Threats, vulnerabilities, and risks 35
Cyber threat intelligence (CTI) 36
STRIDE model 36
The Common Vulnerability Scoring System (CVSS) 36
The Common Vulnerabilities and Exposures (CVE) Program 37
The Common Weakness Enumeration (CWE) 37
Coordinated vulnerability disclosure (CVD) 37
A brief overview of the most common techniques used by attackers 38
Tactics, Techniques, and Procedures 38
Malware 39
Denial of service (DoS) 40
Social engineering 40
DNS attacks 41
Password attacks 41
Web based attacks 42
Attacks in wireless networks 42
Terms related to more dangerous attacks 43
Sources of threat 43
How we defend ourselves 44

Security standards 44
Security controls 45
Security concepts 46
Cryptography 46
Data sanitization 49
Access control services 49
Security protocols 50
Protection of wireless networks 50
A brief overview of security devices 50
Security platforms 51
Security software 52
Data loss prevention (DLP) 52
Secure communications 53
VPNs 53
PGP 53
Incident handling 54
Preparation 54
Monitoring 54
Communications planning 56
Resources planning 56
User awareness 56
Detection 56
Receipt of incident reports 57
Triage 57
Prioritization 58
Incident numbers 58
Ticketing system 58
Indicators of compromise (IOCs) 59
Indicators of attack (IOAs) 60
Analysis 61
Analysing log files 61
Analysing malware 61
Digital forensics 63
An intruder's signature 64
Containment 64
Identifying an attacker 65
Eradication 65
Recovery 65

Closing a ticket 66
Post-event activities 66
Organizing incident resolution 66
Incident communication and coordination 66
Evidence gathering 67
Incident handling checklist 68
What next? **70**

Foreword

In society, we look at first responders as heroes – as women and men who rise to our rescue at times of crises, save lives and provide security. We see their good work: the fire fighters, the paramedics, law enforcement officers, the disaster responders.

But there is another group of first responders, dealing with emerging and difficult-to-predict crises, and therefore always on alert. They deal with risks and threats which stem from the digital space, and which influence the security and stability of our societies as well as our individual wellbeing. They are generally not very visible, at least not publicly. They are Computer Security Incident Response Teams (CSIRTs) or Computer Emergency Response Team (CERTs).

DCAF has had long experience with CSIRTs/CERTs. From this, we know what it takes to strengthen their governance, improve their planning, and build their capacities.

We are pleased to see an array of critical concepts and best practices come together coherently and clearly in this guidebook. When we embarked on this initiative, we had in mind a simple, easy-to-read primer for new CSIRT/CERT employees – and it has, we think, achieved this. Beyond this, and perhaps equally important is that this guidebook can help existing employees, because it presents an up-to-date look at all of the essential, rapidly evolving issues in this field.

These digital first responders are also connectors, across the public and private sectors, and indeed the general public. And DCAF has recognized, and created knowledge about, the importance of connections and collaboration among CSIRTs/CERTs nationally and internationally.

Good governance in cybersecurity means that all relevant actors have clearly prescribed mandates, have resources to conduct these mandates efficiently and effectively, and are overseen and held accountable for their work. This is certainly the case for CSIRTs/CERTs. They provide key services, with a broader scope than many people realize. Developing and improving these capacities at the national level, and their international cooperation, is a global priority, and we hope this guidebook will make a useful contribution towards this goal.

Darko Stančić

Head, Europe and Central Asia Division

DCAF – Geneva Centre for Security Sector Governance

Preface

This Guidebook is intended to help new employees on a Computer Security Incident Response Team (CSIRT) understand the structure in which they are working, and clarify what is expected of them and how to achieve it.

This Guidebook is not intended to train CSIRT employees, for example in the methods and techniques used to identify an attack or analyze malware, or how to use specific tools; rather, its purpose is to suggest what tools can be used and where to find them. Of course, such a Guidebook can also be used by long-term CSIRT employees as means of reviewing their professional knowledge. Employees with varying levels of knowledge and experience work on a CSIRT. Therefore, parts of this Guidebook may cover matters that some employees find basic or trivial, and they can simply pass over these parts.

Finally, it should be noted that the field of cybersecurity is evolving rapidly, and there tend to be many points of view. Even renowned institutions apply different conceptions, definitions, or classifications of the same terms. Thus, the parts of this Guidebook that represent the opinion of the author – shaped over many years of experience in information security and on a CSIRT – may diverge from the views of other individuals or institutions in this field.

What is a CSIRT?

A bit of history

The creation of the first Computer Emergency Response Team (CERT) is linked to the early creation, development, and expansion of computer networks, before the Internet as we know it today existed. This was a period of incredible and unprecedented growth in communication, and the birth of a new era of human civilization. Without the enthusiasm and dedication of those who created this network over decades, the rise of a global cyberspace would not have been possible, and the world would never had made such a 'quantum leap' in communications. But such rapid progress in any field leaves it open to certain weaknesses. In case of the Internet, functionality took priority over security, resulting in communication protocols that were relatively easy to abuse.

But let us return to the start of this story, in the 1960s. The world was in the midst of the Cold War, facing the existential danger that two opposing powers could resort to the use of nuclear weapons and enter into a conflict that would end in mass destruction. As in any war, the adversary's communication systems would be among the first targets, with the goal of cutting off their ability to exchange information and coordinate activities. One of those powers, the United States of America (US), recognized this problem and tasked its Defense Advanced Research Projects Agency (DARPA) with developing a system for information exchange even in the event of widespread nuclear destruction.

At the time, the most extensive and redundant communication network in the US was the telephone network. Many telephone companies had been founded in the US, each independently connecting certain cities and regions, so that even the furthest corners of the country were part of a far-reaching, nationwide tangle of copper telephone cables. DARPA concluded that it was this network that would form the best foundation for creating the information exchange system they sought and turned to the academic community for help.

The resulting collaboration was known as ARPA (Advanced Research Projects Agency), and the network ARPA developed was called ARPANET (ARPA Network). Its protocols allowed information of various kinds to be transmitted between any two nodes of the network, and importantly, the network itself was able to dynamically determine the best path between those two nodes and adapt itself in case of disruptions. For example, if some nodes on the path were destroyed during the transmission of information, the rest of the information was delivered by another, available path.

Having helped realize ARPANET, the academic community saw that such a network could also significantly improve the work and functioning of academia. This led to the creation of the CSNET (Computer Science Network) in 1981, which was built exclusively for civilian (academic) use and linked several prestigious US universities. The establishment of CSNET was met with enthusiasm from both professors and students, generating further expansion of the network and campus-wide discussions of the new possibilities it represented.

Several years later, in 1986, the next big step in Internet development occurred when supercomputers were connected to a computer network, establishing the NSFNET (National Science Foundation Network). This network connected key universities in the US with universities in Germany, Israel, and Japan. As interest in connecting to the network surged, transmission capacities had to be raised to 34 Mb/s by August 1988. The growth of NSFNET forced those who created and maintained it to concede that, inevitably, a commercial element had to be introduced into the network. Hence, private companies, primarily Merit Network, were involved in maintaining and further upgrading the network. Then, when final restrictions on commercial traffic were removed in 1995, the principles on which the Internet still operates today were established.

The first CERT

A crucial moment for the future of computer security incident management occurred in November 1988, just a few months after upgrades were made to the NSFNET communications capacity, when Cornell University student Robert Tappan Morris wrote a programme that used vulnerabilities in the BSD version of the UNIX operating system to examine active connections, forward a copy of itself over the network to another computer, and start its execution. In other words, Morris had created a worm, and though he claimed his intention had simply been to test the functionality of the network, he had executed his programme at the Massachusetts Institute of Technology (MIT) and not at Cornell. Upon launch, the worm did what it was intended to do, polling all the IP addresses available from the computer running the programme, transferring its copy wherever possible, and initiating the launch of those copies. But all those copies were doing the same thing, so after a short time, multiple copies of the same programme were running on any vulnerable computer, which slowed them down and eventually crashed the system.

It was later estimated that some 6,000 computers were infected with the worm, out of 60,000 that were on the network at the time. This was despite the fact that problems had quickly been detected after Morris launched his worm and teams at universities across the US had gotten swiftly to work to determine the cause, identifying it within the first 24 hours. The researchers at Berkeley and MIT who found the malware offered recommendations on how to prevent its further spread and clean affected systems, but doing so properly required that some segments of the network be disconnected and cleaned completely before they were reconnected. Though it took several weeks, the approach successfully cleared the network, but the Morris worm became something of a clarion call for universities. While the universities had managed to find a solution to the attack relatively quickly, they understood that similar incidents would occur in the future and may present much more serious problems to solve.[1]

It was therefore necessary to be prepared, and the first practical action in this regard was taken by Carnegie Mellon University, which formed a CERT to prepare for and respond to future computer security events. The US Government liked the idea and made an agreement with Carnegie Mellon's Software Engineering Institute (SEI) to implement a similar concept within government entities, with the goal of responding to incidents in the critical infrastructure of the US.[2]

CERT, CSIRT, CIRT: Different names, one objective

The establishment of the CERT at Carnegie Mellon and the subsequent support it received from the US government contributed to the normalization of these teams, and many more were formed across the US and around the world. The first CERT in Europe was formed in 1992, at SURFnet in The Netherlands[3] – which develops, implements, and maintains that country's national research and education network (NREN).

Though 'CERT' originally stood for Computer Emergency Response Team, the development of the CERT/CC led Carnegie Mellon to trademark the acronym. It is still possible to use the term CERT by signing an agreement with Carnegie Mellon, which does not seek any compensation in return but reserves the right to release the name only to teams that it deems meet certain standards.[4] Many teams choose instead to use the term Computer Security Incident Response Team (CSIRT), and this is the term that will be used in this text. Some teams also use Computer Incident Response Team (CIRT), Incident Response Team (IRT), and Security Incident Response Team (SIRT), among others.

1 Morris became the first person in the US to be sentenced under the Computer Fraud and Abuse Act of 1986, for damage estimated at between $100,000 and $10 million. He received a suspended prison sentence, a fine, and community service. However, Morris later became a professor at MIT.

2 See: Software Engineering Institute, 'The CERT Division: What We Do', https://www.sei.cmu.edu/about/divisions/cert/index.cfm#cert-division-what-we-do. NB: The URLs contained in this text are correct and accessible at the time of publication, but their enduring accuracy cannot be guaranteed.

3 See more at: https://www.surf.nl/en

4 Although there have been no instances of Carnegie Mellon University taking legal action against teams using the CERT name without signing an agreement, contacting the University is considered a matter of respect and is recommended for any team wishing to use the CERT name.

Purpose

No matter their name, the one thing all CSIRTs have in common is that they respond to computer incidents. The way they respond can vary widely, from 'soft' measures such as recording the incident and giving advice to victims, to 'hard' measures that include taking the lead in resolving the incident and walking the organization through the recovery process, giving instructions to victims of an attack, and analysing all elements of an incident to supply forensics and offer cooperation with the law enforcement authorities. Still, it is important to remember that members of a CSIRT are not the police, even if it is within their jurisdiction to collect evidence for use in legal proceedings and testify in court. A CSIRT does not conduct criminal-legal investigations, file charges, or arrest suspects. These are the responsibilities of cybercrime or high-tech crime units, and CSIRT employees merely cooperate with them as needed.

Commonly, a CSIRT has other responsibilities in addition to incident response. These may be of a preventive nature (such as security awareness raising) or may include monitoring and notification. Regardless of the specific responsibilities of any given CSIRT, its employees must be well-trained, alert, and capable of detecting and effectively responding to a cyber incident at any time.

It benefits a CSIRT if its purpose and competences are clearly stated and are supported by the management of the organization it serves. The responsibilities of a CSIRT can be expressed in a document that describes the remit of every unit within the larger organization, or in a separate document specific to the CSIRT. One of the first things any new CSIRT employee should do is familiarize themselves with these responsibilities.

SOCs

In addition to a CSIRT, an organization may form a Security Operations Centre (SOC), which is also meant to detect and respond to incidents. The distinction between the responsibilities of a CSIRT and a SOC can be confusing, but the primary task of a SOC is to monitor an organization's information and communications technology (ICT) infrastructure and detect security events and incidents. Accordingly, SOC teams typically work in shifts covering 24/7/365. In organizations with both a SOC and a CSIRT, the latter typically responds to security incidents detected by the former.

Position within an organization

Because the activity of a CSIRT is closely related to ICT, it is natural for CSIRTs to be organizationally situated within larger units dealing with information and communications. This streamlines cooperation with other entities directly connected to the work of a CSIRT, including data centres and units charged with website maintenance. However, as the work of a CSIRT is intertwined with organizational security as well, these teams are sometimes placed within departments responsible for corporate security, which better connects the team to security structures in the organization. This can be crucial when it is necessary to react quickly and compel other organizational units to apply necessary measures. In some cases, a CSIRT is also an independent entity, directly subordinate to the leadership of an organization.

Notwithstanding its place in an organization's structure, a CSIRT must have effective communication with the management of that organization and must have top-down support for both preventive and reactive pursuits. When serious incidents occur, it is sometimes necessary to implement measures that require additional effort from other personnel or inhibit them from performing their work effectively, and the resentment this may breed can only be overcome with the full support of management. This support from an organization's leadership is also crucial in implementing preventive measures like security awareness training, drills and tests for employees, or the prohibition of certain resources (such as USB ports or social networks). When there are no visible indicators of threat, employees tend to treat these trainings and exercises as a waste of time, and often perceive prohibitions as a personal attack or complain of a decrease in efficiency. The resistance of employees can represent a challenge for management if they are not themselves well informed about the purpose of such measures.

Constituents

Constituents represent one or more entities to which a CSIRT provides its services. In most cases, constituents are organizations that finance the work of a CSIRT. There are also examples of CSIRTs that do not have a limited number of constituents but provide services to anyone seeking help.

When establishing a CSIRT, the constituent(s) must be defined; but this definition can change and the set of constituents expanded or reduced.

Basic terminology

The field of IT is very dynamic, with new notions, concepts, and terms emerging every day. During your work on a CSIRT, you will encounter many new terms, and should understand their definitions. Yet, for many terms, there is no universal defining concept, and different definitions exist for the same term even in eminent sources. This may mean that you and your colleague from another CSIRT, or perhaps from the IT department of your own organization, don't fully understand each other when communicating about a problem, because you are using different terms for the same thing or the same term for different things.

To ensure that readers of this Guidebook understand the terms herein as intended, the short list of definitions that follows clarifies some key concepts. These definitions are mostly taken from the European Agency for Cybersecurity (ENISA)[5] and the US National Institute of Standards and Technology (NIST)[6]:

Information security – the protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide confidentiality, integrity, and availability. **Information security is a subset of cybersecurity.**

Network and information security – the ability of a network or an information system to resist, at a given level of confidence, accidental events or unlawful or malicious actions that compromise the availability, authenticity, integrity, and confidentiality of stored or transmitted data and the related services offered by or accessible via those networks and systems. **Network and information security is a subset of cybersecurity.**

Cybersecurity – all aspects of prevention, forecasting, tolerance, detection, mitigation, removal, analysis, and investigation of cyber incidents. Considering the different components of cyberspace, cybersecurity should cover the following attributes: availability, reliability, safety, confidentiality, integrity, maintainability (for tangible systems, information, and networks), robustness, survivability, resilience (to support the dynamicity of the cyberspace), accountability, authenticity, and non-repudiation (to support information security).

Cyber hygiene – practices that should be implemented to protect users and businesses from data loss due to attacks or theft.

Event – any observable occurrence in a system or network, including a user connecting to a file share, a server receiving a request for a web page, a user sending an email, and a firewall blocking a connection attempt.

Adverse event – any occurrence with a negative consequence, such as a system crash or packet flood, the unauthorized use of system privileges or unauthorized access to sensitive data, and the execution of malware that destroys data.

Cyber incident – any occurrence that has impact on any of the components of cyberspace or the functioning of the cyberspace, independent of whether it is natural or human

5 See: 'ENISA overview of cybersecurity and related terminology', Version 1, September 2017. Available at: https://www.enisa.europa.eu/publications/enisa-position-papers-and-opinions/enisa-overview-of-cybersecurity-and-related-terminology (When ENISA issued this publication in 2017, it was known by its original name, the European Agency for Network and Information Security/the European Network and Information Security Agency; in 2019, the name was changed but the ENISA acronym was retained).

6 See: Paul Cichonski et al., Computer Security Incident Handling Guide: Recommendations of the National Institute of Standards and Technology, Special Publication 800-61, Revision 2 (US Department of Commerce, 2012). Available as a PDF at: https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-61r2.pdf

made; malicious or non-malicious; deliberate, accidental, or due to incompetence; due to development; or due to operational interactions. Also, any adverse event generated by any cyberspace components even if the damage/disruption or dysfunctionality extends from outside the cyberspace; and any violation or imminent threat of violation of computer security policies, acceptable use policies, or standard security practices.

Cyber accident – any occurrence that causes significant damage to cyberspace or any other asset (i.e., has an impact on performance, or requires repairs or replacement), or causes personal injury.

Cyber attacks – cyber incidents triggered by malicious intent causing damage, disruption, or dysfunctionality.

Incident handling or incident response – the process of detecting and analysing incidents and limiting their effect.

Incident response team – the team responsible for providing incident response services to parts or entire organization, by receiving information on possible incidents, investigating, and taking action to ensure that any damage is minimized.

Cyber investigation – a process conducted for the purpose of cyber accident and incident prevention which includes the gathering and analysis of information, the drawing of conclusions (including the determination of causes) and, when appropriate, the making of safety and security recommendations.

Cybercrime – any crime/criminal activity facilitated by or using cyberspace.

Cyber sabotage – any sabotage activity facilitated by or using cyberspace.

Cyber espionage – any espionage facilitated by or using cyberspace, which falls into two categories: (a) state espionage (involving state actors), and (b) industrial espionage (involving commercial actors).

Cyber defence – a variety of mechanisms used to mitigate or respond to cyber attacks.

Cyberwarfare – any action by a state, group, or criminal organization facilitated by or using cyberspace to target a state.

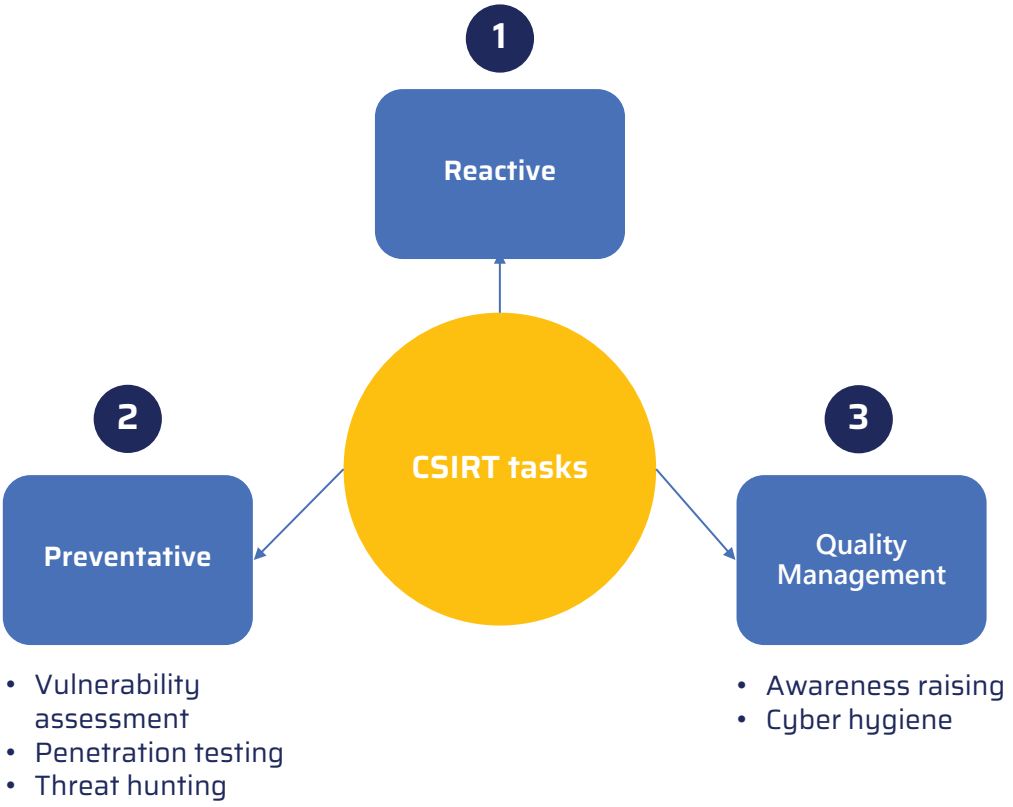
While these definitions will help you better understanding this Guidebook, what is most important is that you fully understand the language used by your own CSIRT colleagues. Indeed, it is not unusual for teams to develop their own definition of certain terms, and even endemic terms and definitions.

Tasks performed by CSIRTs

It has already been noted that incident response is the job of every CSIRT but that these teams can have many other responsibilities as well. Carnegie Mellon has published a manual on establishing computer response teams that lists the tasks for which a CSIRT should be competent, and a description of each.[7] These tasks are divided into three groups: reactive, preventive, and quality improvement.

Reactive tasks include all the work directed at remediating an ongoing incident, such as by coordinating incident response or helping an attacked system recover. Preventive tasks reduce the surface attackers can use for a potential attack, for example through vulnerability testing or security awareness. Quality improvement is achieved in tasks such as education or analysis.

7 Moira J. West-Brown et al., Handbook for Computer Security Incident Response Teams (CSIRTs), 2nd Edition, No. CMU/SEI-2003-HB-002 (Pittsburgh: Carnegie Mellon Software Engineering Institute, 2003). Available as a PDF at: https://resources.sei.cmu.edu/asset_files/Handbook/2003_002_001_14102.pdf



Reactive tasks

Reactive tasks, including any that represent a response to information about an incident or security threat, are the essence of why a CSIRT exists. This includes handling incidents in all phases (discussed in more detail later), addressing detected vulnerabilities, analysing artifacts, providing notifications and warnings about an incident or imminent threat, and etc.

Preventive tasks

The purpose of preventive tasks is to reduce the number of future cyber incidents and better prepare for them. This includes threat monitoring, risk analysis, the testing of security mechanisms and devices, the creation of specialized security tools, the configuration of security devices, and etc.

Vulnerability assessment

In a **vulnerability assessment,** ICT systems or products are examined in order to identify, prioritize, and manage vulnerabilities. A vulnerability assessment may be performed if there is any suspicion, when information about weaknesses is obtained, following significant changes in the system, after the implementation of security patches, or as a part of planned and systematic **vulnerability management.**

An element of vulnerability assessment is **vulnerability scanning,** in which a system is checked for vulnerabilities by using techniques that cannot compromise that system. Despite the non-invasive nature of vulnerability scanning, approval for this activity should always come from senior management and/or the owner of the system being tested.[8]

Nmap

Network Mapper (Nmap) is a free, open-source tool for network discovery, inventory, and

8 This approval may be implicit; for example, based on a document signed by someone with the proper authority.

mapping, but can also be used for security auditing.[9] Nmap can identify available hosts on a network, which ports are open on them, what services they offer, and what operating system they are running; and by combining different techniques, can detect network vulnerabilities. Nmap is therefore widely used as a tool for vulnerability scanning.

OpenVAS

Open Vulnerability Assessment System (OpenVAS) is a vulnerability scanner based on the Nessus scanning tool. While Nessus has transitioned from an open source to a commercial product, OpenVAS is still available under the GNU General Public License, as part of the Greenbone Vulnerability Manager software platform.[10]

OWASP

The Open Web Application Security Project (OWASP) is a non-profit foundation that works to improve the security of web applications.[11] Among other useful documents, OWASP publishes a Web Security Testing Guide[12] and a Mobile Application Security Testing Guide,[13] both of which can be used in planning vulnerability assessments. Furthermore, OWASP regularly updates a list of the ten most critical web application security risks.[14]

Penetration testing

A penetration test (or pen test) is a method of testing the security of a system by simulating a cyber attack on that system. While a vulnerability assessment is strictly limited and is carried out without any possibility of harming or compromising the tested system, a pen test is much more invasive and uses methods applied in real cyber attacks. Since they can have potentially destructive effects, and because social engineering methods may be applied to employees as part of a pen test, plans for these tests must be carefully developed and should be approved by an organization's management.

Penetration tests must be performed by proven experts (pen testers). This task is sometimes assigned to the CSIRT, but if CSIRT employees lack sufficient experience, external companies that specialize in pen testing can be hired. In that case, external actors who participate in pen testing at an organization must sign non-disclosure agreements before any activity commences.

There are three types of penetration testing:

- white box, for which pen testers know everything about the system;
- grey box, for which information about the system is only partially known; and
- black box, for which pen testers know nothing about the system prior to testing.

Metasploit

Metasploit (or the Metasploit Framework) is a tool used to test for vulnerabilities on computers or networks.[15] It is open source, available for most operating systems, easy to customize, and contains a large set of exploits and payloads. These can be used by pen testers, but also by cybercriminals.

Kali Linux

Kali Linux is a special Linux distribution with a suite of tools designed for penetration testing, vulnerability assessment, digital forensics, packet analysis, reverse engineering, and more.[16] It is

9 See: https://nmap.org/

10 See: https://github.com/greenbone/openvas-scanner

11 See: https://owasp.org/

12 Elie Saad and Rick Mitchell, Web Security Testing Guide, Version 4.2 (The OWASP Foundation, 2020). Available as a PDF at: https://github.com/OWASP/wstg/releases/download/v4.2/wstg-v4.2.pdf. Also see: https://owasp.org/www-project-web-security-testing-guide/

13 Sven Schleier et al., Mobile Application Security Testing Guide, Version 1.5.0 (The OWASP Foundation, 2022). Available as a PDF at: https://github.com/OWASP/owasp-mastg/releases/latest/download/OWASP_MASTG-v1.5.0.pdf

14 See: https://owasp.org/www-project-top-ten/

15 See: https://www.metasploit.com/

16 See: https://www.kali.org/

open-source and free-for-download software, maintained by Offensive Security.[17] Just some of the tools included in the Kali Linux package are Metasploit, Nmap, OWASP ZAP, Wireshark, Aircrack-ng, and John the Ripper.

Threat hunting

In fact, security systems are imperfect and sophisticated attackers can find a way through them. An attacker who has some initial success can slowly explore a system, raise their level of access, and extract important data. These activities could remain undetected for months, or even years. And, if an attacker deems it necessary to erase their tracks, they can demolish a system, which could lead to the collapse of an organization.

Threat hunting is aimed at identifying threats in a system. This activity requires skilled, human personnel who can recognize suspicious and malicious activities that go undetected by security incident and event management (SIEM) or endpoint detection and response (EDR) systems. Threat hunting may be conducted by searching a system for Indicators of Compromise (IOCs) or known malicious hash values, for example, or by examining network logs in search of communications with known malicious IP addresses or domain names. In the case of a suspected attack or attacker, a system should be checked for Indicators of Attack (IOAs) according to models provided in the Mitre ATT&CK framework (see the 'Tactics, Techniques, and Procedures' section).[18]

A good source for more information on threat hunting is a publication of the UK Home Office Digital, Data, and Technology Cyber Security Programme entitled, 'Detecting the Unknown: A Guide to Threat Hunting'.[19]

YARA

YARA, a tool created primarily to identify malware samples, can be used to classify files based on conditions defined in a structure called YARA rules. These rules consist of textual and binary patterns depicting malware and sets of strings and Boolean expressions that define the search method. The most comprehensive documentation on YARA and downloadable releases is provided by VirusTotal.[20] Some examples of YARA rules can also be found on Github.[21]

Quality management tasks

The kinds of tasks that fall under quality management are typically done by other units in an organization, but CSIRT employees may be engaged because their expertise is a benefit to the performance of these tasks. This includes the delivery of trainings, efforts to raise security awareness among employees, product certification, participation in business continuity planning, security-related consultations, etc.

Awareness raising

Awareness raising is vital among these tasks. After all, it is the human element that is the weakest link of all that makes up an organization's ICT system protection. This is why such a large percentage of cyber attacks are initiated through communication between an attacker and a potential (human) victim. Cybersecurity awareness raising in an organization refers to educating employees on how to recognize attempted cyber attacks, as well as how to respond and who to inform. The goal is to create a security-oriented mindset among employees and foster a cybersecurity culture across an organization. To that end, employees must be educated not only in how to properly protect the assets of the organization, but also their own personal data and devices.

17 See: https://www.offensive-security.com/

18 See: https://attack.mitre.org/

19 'Detecting the Unknown: A Guide to Threat Hunting', Version 2.0, March 2019. Available as a PDF at: https://hodigital.blog.gov.uk/wp-content/uploads/sites/161/2020/03/Detecting-the-Unknown-A-Guide-to-Threat-Hunting-v2.0.pdf

20 See: https://virustotal.github.io/yara/

21 See: https://github.com/Yara-Rules/rules

Truly mainstreaming this culture demands more than a single training session. Cybersecurity awareness education must be a process, and trainings must be reviewed periodically and constantly improved. Helpful tips on fostering cybersecurity awareness in an organization can be found in the Carnegie Mellon publication, 'Building a Cybersecurity Awareness Program'.[22]

A good practice in the EU, and in many other countries where October is a cybersecurity month, is the yearly focus brought to security awareness when ENISA and EU Member States promote cybersecurity by engaging in awareness raising activities and knowledge sharing throughout the month.

Gophish

Among the most common cyber attacks are phishing emails, which is why one of the most common topics in awareness raising is how to recognize such emails. An open-source tool, Gophish, offers a phishing framework with editable templates of phishing emails, controlled launch, and results tracking.[23] Beware that if such a simulated campaign is carried out within an organization, it must be handled with the utmost professionalism to avoid shaming any employees who fall victim, and management approval is mandatory.

Cyber hygiene

Cyber hygiene refers to practices that should be implemented by employees or by IT personnel to maintain the security and functionality of ICT systems. The goal of cyber hygiene programmes is to make it a regular habit to implement security measures, akin to hand washing. And while there is no commonly agreed set of cyber hygiene measures, some of those used most frequently include:

- the establishment of a regular backup procedure;
- the use of complex and separate passwords, which are changed frequently;
- regular patching and updating to operating systems and application software;
- the development of a plan for steps to take in the event of a cyber incident;
- awareness raising among employees and other users about cyber threats and cybersecurity;
- the use of secure configurations for all devices and software;
- the implementation of anti-malware software;
- the encryption of all sensitive data;
- the protection of Wi-Fi networks;
- the use of multi-factor authentication whenever possible;
- the implementation of firewalls;
- configuration standards for email and web browser security settings;
- training that encourages employees and users to take care when interacting with emails, especially when they contain attachments and web links; and
- ensuring the security of mobile devices.

Some useful tips related to cyber hygiene can be found in the DCAF publication, 'Guidebook on Staying Safe Online: Cyber Hygiene for Public Institutions and SMEs'.[24]

22 Angel Hueca, Brittany Manley, and Larry Rogers, 'Building a Cybersecurity Awareness Program', Carnegie Mellon CERT, 2020. Available as a PDF at: https://apps.dtic.mil/sti/pdfs/AD1112780.pdf

23 See: https://getgophish.com/

24 Vladan Babić and Aleksandar Bratić, 'Guidebook on Staying Safe Online: Cyber Hygiene for Public Institutions and SMEs', October 2022, DCAF. Available as a PDF at: https://www.dcaf.ch/sites/default/files/publications/documents/GuidebookStayingSafeOnline_CyberHygiene_sept2022_EN-2web.pdf

The organization of a CSIRT

A CSIRT may be made up of people specifically employed to be on the CSIRT. However, it can also include people who work in other parts of an organization and come together when needed, or may be supported by external experts when required. Volunteer members sometimes work on CSIRTs as well.

Types of CSIRTs

Depending on the constituents served by a CSIRT, different types are recognized. These are:

- National CSIRTs
- Public sector CSIRTs
- Private sector CSIRTs
- Military CSIRTs
- Academic CSIRTs
- CSIRTs providing services to third parties

National CSIRTs

A national CSIRT is established at the state level to provide services to an entire national community. In the event of an incident in critical infrastructure systems, it may play only a coordinating role, or it may be staffed and equipped to take over responsibility for resolving an incident through completion. The specific role of a national CSIRT depends primarily on its legal framework, but nevertheless, all national CSIRTs represent a technical point of contact for international cooperation.

The purpose and tasks of national CSIRTs are set out in a guidebook published by the Global Forum on Cyber Expertise (GFCE) and the Dutch government, entitled **Getting started with a National CSIRT.**[25]

RFC 2350: CSIRT description document

It is important that the constituents of a CSIRT, other CSIRTs, and anyone else who seeks support from or cooperation with a CSIRT can access details about its competences, constituents, and services, as well as contact information. A template presented by the Network Working Group in Request for Comments (RFC) 2350 of June 1998 can be used to systematically collect any information required for this purpose.[26] It should be completed by CSIRTs and published for open access and review.

25 Hanneke Duijnhoven et al., Getting started with a national CSIRT (Den Haag: TNO, 2021). Available as a PDF at: https://cybilportal.org/wp-content/uploads/2021/06/TNO-2021-Getting_started_with_a_national_CSIRT_FINAL.pdf

26 The document is available at: https://www.ietf.org/rfc/rfc2350.txt; see Appendix D for the template.

CSIRT operations framework



- Standard Operating Procedures (SOPs) for cyber attacks: equipment, backups, trainings, etc.
- Knowledge, clarity, management support, tools, etc.
- Security Incident Management Maturity (SIM3) supports self-assessment
- Need to know limits of the law and of privacy rules during stressful situations
- CSIRTs are not law enforcement
- However, they can pass on useful facts

Procedures

Procedures, or Standard Operating Procedures (SOPs), are stipulated in documents that explain every step to be taken in operational activities. In the context of a CSIRT for example, a procedure may describe the actions to take in the event of a cyber attack, how to install certain equipment, the process for performing backups, or how to prepare a CSIRT training. Procedures are essential to the operation of any CSIRT, and long-time members of a CSIRT usually know most of these procedures by heart (though, they should have their memory refreshed from time to time). Obviously, it is best for new employees to refer to procedures in a written form.

Every so often, a CSIRT's procedures should be practiced so that their application becomes more routine for members of the team. This can also help illuminate why a procedure may need to be changed. Further, it is necessary after incident response to analyse whether the relevant procedures were adequate in practice.

Operational capacities

From the moment a CSIRT is established, either formally or informally, the work begins to build the necessary capacities for successful operations. This is not only a matter of the technical knowledge of team members, as it also hinges on whether there is clarity regarding the services provided, how competences are defined, whether management has expressed support, the procedures established, the usable tools available to a CSIRT, and much more. Additionally, no CSIRT can push continuously forward without periods of stagnation or backsliding, which can be due to any number of causes, from changes in the legal framework or in the organization a CSIRT serves, to the departure of personnel or the deterioration of interpersonal relations, to a lack of management support or budget problems. Any such problems should be solved by CSIRT members with patience and persistence.

SIM3 model

One way to measure the maturity of an organization's security incident management is through a CSIRT self-assessment that utilizes the Security Incident Management Maturity Model, commonly

known as the SIM3 model.[27] Developed by the Open CSIRT Foundation, the model considers the organization, and its human resources, tools, and processes, and within each of these areas, scores several parameters from zero to four. A grade of zero is assigned to parameters for which there has been no activity or discussion within a CSIRT, while a grade of four is assigned if a parameter is well defined and verified by an authority above the CSIRT leadership. The SIM3 model includes a total of 44 parameters that must be assessed and provides each with a reference value. It is on this basis that TF-CSIRT, an international association of CSIRTs (see the 'International associations' section), certifies its members.[28]

The SIM3 model has also been embraced by ENISA. The Agency published a manual offering self-assessment recommendations as well as criteria for CSIRTs to assess their maturity level: as basic, intermediate, or advanced.[29]

Legal issues

CSIRT employees must perform their tasks in accordance with regulations, and they should take extreme care not to commit a criminal offence in the course of their work. Yet, some members of a CSIRT may be tempted to use their knowledge and skills to reach beyond the threshold of what is permissible, driven by a desire to defend their system from attack, especially in stressful and adrenaline-charged environments. This can lead to their own prosecution, the disbanding of their team, or even accusations against their country that it has violated international rules. Thus, members of a CSIRT must always be mindful to defend their own system without attacking others, even when it appears likely that a counterattack could prevent an attack on the system they are defending.

Privacy is another significant legal issue that is inescapably relevant to CSIRT employees. While performing CSIRT-specific tasks such as log analysis, a great deal of private user data may be available. Members of a CSIRT must understand that reading the content of communications is not part of their job and that such activities are not permitted. And if a CSIRT employee comes across the private data of other persons in the course of their work, they must know that they cannot share it with anyone or process it in any way.

A CSIRT must take responsibility for its actions. Otherwise, considerable problems may arise for the CSIRT itself, or its constituents. Irresponsible behavior by CSIRT employees can manifest, among other things, in a failure to respond to an incident report, only a cursory analysis of logs and other documents, a failure to fulfill legal obligations regarding the reporting of criminal activities, the publication or transmission of unverified and inaccurate information that may mislead recipients or damage reputations, and the communication of useless or outdated advice.

Due to the nature of their work, CSIRT employees may find it necessary to obtain legal advice at times. They should seek this advice from someone who is qualified from a legal standpoint, but who is also familiar with the specific issues that arise in the work of a CSIRT. Ideally, such a qualified legal adviser should be part of a CSIRT, but if this is not feasible, a legal adviser should be designated within the larger organization a CSIRT serves, so that CSIRT employees know where to turn if necessary. Additionally, legal advisers should have an influence on the procedures implemented in a CSIRT, at least by reviewing and approving SOPs prior to their adoption, and must be consulted when drafting any contracts, cooperation agreements, memorandums of understanding, service provision agreements, etc.

27 Don Stikvoort, 'SIM3: Security Incident Management Maturity Model', Open CSIRT Foundation, May 2019. Available as a PDF at: http://opencsirt.org/wp-content/uploads/2019/12/SIM3-mkXVIIIc.pdf

28 See: https://tf-csirt.org/tf-csirt/

29 ENISA CSIRT maturity assessment model, Version 2.0, 30 April 2019. See: https://www.enisa.europa.eu/publications/study-on-csirt-maturity

Cooperation with law enforcement agencies

Different countries treat cyberspace activities differently, but in general, it is safe to say that actions such as breaking into an ICT system, stealing personal and business data, making information inaccessible, or endangering critical infrastructure, are criminal offences almost everywhere. How evidence of a criminal offense is obtained and how judicial proceedings are conducted varies from state to state, as some countries have special units to combat cybercrime or high-tech crime, and special prosecutors' offices and courts to try cases in this area. It is important to note that CSIRT employees are not members of a cybercrime police unit and are not law enforcement agents.

In responding to an incident, CSIRT employees should take care to collect and save any facts about the incident so that the course of the attack can be successfully reconstructed and its cause determined. If an incident can be characterized as criminal according to national legislation, a report should be made to the relevant law enforcement authority, with appropriate documentation about the incident attached. The specific documentation that is submitted to law enforcement depends on organizational policy. This is always a sensitive issue and members of a CSIRT should be mindful not to reveal confidential information of the organization; yet, if national law prescribes that certain information must be provided, or there is a court order stipulating that all documentation be made available, a CSIRT is obliged to comply.

At the end of the day, a CSIRT is on the same side as law enforcement, and the two have complementary responsibilities. An atmosphere of mutual trust can enhance cooperation. Thus, it is common for a CSIRT to select a member to be the primary point of contact with law enforcement. This employee should be familiar with any relevant legal framework and obligations as well as any procedures for communicating with law enforcement authorities. If such a point person exists in a CSIRT, all initial communication with legal authorities should go through them; if not, the manager of the CSIRT serves as this point of contact.

In some states, police sometimes seek the assistance of CSIRTs to perform forensic analysis on evidence they have collected. CSIRT employees should be informed about the practice in their state, and whether the courts accept evidence obtained though these means.

The work environment

For any CSIRT, people are the most important resource. An organization may have premier security devices, and a CSIRT that is equipped with wonderful tools and the support of management, but if no one on the team can read logs, correlate events, or spot hidden malicious activity, attackers will be successful anyway. All the devices and tools in the world are no match for the human mind when it comes to understanding the desires and plans of an attacker and preventing further consequences.

Duties and expectations

Working on a CSIRT can be demanding and complicated, especially with numerous intense responsibilities. Organizations often emphasize this in documents describing the duties and requirements for CSIRT employees. New employees should receive this information within their first few days at a workplace, as well as any documents that must be signed. It is advised to read and understand everything in any document you are asked to sign or otherwise seek a senior colleague's assistance.

CSIRT employees must behave in accordance with codes of conduct established by any organization within which a CSIRT operates and show that they are capable of respecting policy and following procedure. Additionally, members of a CSIRT must possess the character and ability to work under stress, as well as an analytical spirit, an affinity for teamwork, the discipline to

protect personal and proprietary information, and avoid using their knowledge and experience in a reckless manner. CSIRT employees are also expected to have good oral and written skills in both English and their native language (if applicable), the capacity to communicate professionally with third parties, a sense of stewardship vis-à-vis the reputation of the CSIRT, a willingness to obtain additional and continuing education, and a readiness to ask for help and admit mistakes.

Code of conduct

Some organizations, and some CSIRTs, have codes of conduct that set out general expectations for all employees. A code of conduct may also describe how employees are expected to react in certain situations, and the kinds of interactions they can have with other employees or with people outside the organization. Long-time employees may behave in accordance with a code of conduct without having to think about it, but new employees may need guidance from more experienced colleagues.

One example is the Code of Professional Conduct for SEI Services issued by Carnegie Mellon University.[30] It sets out expectations and practices for anyone operating under license or other applicable agreement with the SEI. While this Code is comprehensive and detailed, such documents can be as brief as one page.

Non-disclosure agreements

A non-disclosure agreement (NDA) is a document underscoring that the signatory should not disclose information about a particular organization, process, work environment, etc. Because the work of a CSIRT involves access to considerable amounts of information and data, which may be very sensitive, many organizations ask new employees to sign NDAs. It is also not uncommon for a similar document to be presented to outgoing employees when they leave a CSIRT, to prevent them from later disclosing any information related to their work on the team.

Due care and due diligence

Due care is an important aspect of the behaviour of any employee, related to the expectation that they will behave like any other reasonable person in a given situation. For example, it assumes that employees will take the reasonable care to remedy issues that could have a negative impact if they are not rectified in an appropriate and timely manner. Bearing this in mind, the expectation of CSIRT employees, in the broadest sense, is to protect prudently the interests of the CSIRT and the organization it serves. Due diligence is the practical application of due care, and thus involves taking any steps required to discover all the facts about an issue and learn what is necessary to prevent similar negative outcomes in the future (such as the detection and investigation of a breach or incident).

Ethical principles

CSIRT employees are expected to have high moral standards. They must resist pressures to abuse the knowledge they have, the rights afforded them by their role, and the resources at their disposal. Thus, the Forum of Incident Response and Security Teams (FIRST) has published a list of ethical principles to which CSIRT employees should abide.[31] While this list is not binding, adherence to these principles are likely to improve any CSIRT and make them better partners within the wider CSIRT community:

- reliability in mutual relations;
- the coordinated detection of vulnerabilities through cooperation with vulnerable parties;
- compliance with information confidentiality requirements;
- the timely confirmation of receipt of requests or notifications, and responses to them;

30 Code of Professional Conduct for SEI Services, Version 1.0, No. CMU/SEI-2004-SR-009, September 2004. Available as a PDF at: https://resources.sei.cmu.edu/asset_files/SpecialReport/2004_003_001_14282.pdf

31 'EthicsfIRST: Ethics for Incident Response and Security Teams', 2019. Available as a PDF at: https://www.first.org/global/sigs/ethics/ethics-first-20191202.pdf

- a commitment to accessing and operating in only authorized systems;
- communication with constituents about ongoing security threats and risks, and opportunities to improve security;
- acting at all times with respect for human rights and intellectual property;
- the maintenance of a healthy, safe, and positive work environment to preserve the physical and mental health of CSIRT employees;
- the provision of resources to improve the scientific and technical knowledge of all members of a CSIRT;
- the collection, storage, and processing of only the data necessary for the work of a CSIRT, with respect for any legal provisions and the right to privacy;
- recognition and respect of the legal obligations, rights, and competences of all parties involved in incident response procedures; and
- acting on the basis of verifiable facts and not rumour.[32]

Necessary knowledge and skills

Attackers are very creative in finding new ways to bypass security measures, and a high volume of new malware appears every day, which means that CSIRT employees can never really complete their education and training as long as they are on the job. There is always something new to learn, ever another problem to solve. It is not realistic or possible for a single team member to be expert in everything, so the solution is specialization. In other words, since attackers often specialize in certain kinds of attacks, there should be CSIRT employees who also specialize, and are thus adequately trained to respond to distinct cyber challenges.

What CSIRT employees need to know to successfully perform their work depends on the competence and services provided by the team. Still, all CSIRT employees should possess basic relevant knowledge, and their specialties should be linked to their affinities.

At the least, CSIRT employees should have basic technical knowledge in the following areas:

- Internet architecture
- Network and routing protocols
- Domain Name System (DNS)
- E-mail and other communication systems
- Computer architecture
- Operating systems
- Programming principles
- Security principles, tools, and equipment
- Security risks, threats, and vulnerabilities
- Methods used for cyber attacks and ways to recognize and defend against them
- Cryptography techniques

As a team, the members of a CSIRT should combine to have in-depth knowledge and experience in each of these areas. Over time, members of a CSIRT inevitably gain additional knowledge and experience outside their specialty area(s).

New CSIRT employees may want to identify an area for which the team lacks specialization and aim to specialize in this area. While a CSIRT manager is ultimately responsible for making any final decisions about the training and role of each team member, the initiative and enthusiasm of an employee can influence these decisions. Remember, though, that courses and trainings alone are not enough to develop sufficient skill in an area if this learning is not paired with outside practice time and a genuine curiosity about the subject. It is also very important to apply new skills in practical ways in order to develop lasting knowledge.

32 Ibid.

Cybersecurity frameworks

Clearly, due to the nature of cybersecurity work, CSIRT members require continuing education and skills improvement throughout their careers. This is a burden for cybersecurity professionals, but also for CSIRT managers, who must identify the training needs of employees and implement appropriate educational modules and capacity building activities. This is complicated by a persistent shortage of labour in the cybersecurity sector, and many unfilled positions.

Forecasts that cyber threats will increase means the need for CSIRT professionals will only grow. Officials have recognized this problem and have embarked on efforts to define the positions, tasks, and requisite knowledge and skills of cybersecurity work. Two of the documents this has yielded are the European Cybersecurity Skills Framework (ECSF) and the National Initiative for Cybersecurity Education (NICE) Cybersecurity Framework, both of which are described below.

European Cybersecurity Skills Framework (ECSF)

In September 2022, ENISA published two documents which, together, form the ECSF.[33] The European Cybersecurity Skills Framework Role Profiles, identify 12 typical cybersecurity roles and their respective mission, tasks, skills, knowledge, and competences;[34] and the ECSF User Manual offers additional guidance and practical examples.[35]

The typical cybersecurity roles defined in the Framework are:

- Chief Information Security Officer (CISO)
- Cyber Incident Responder
- Cyber Legal, Policy and Compliance Officer
- Cyber Threat Intelligence Specialist
- Cybersecurity Architect
- Cybersecurity Auditor
- Cybersecurity Educator
- Cybersecurity Implementer
- Cybersecurity Researcher
- Cybersecurity Risk Manager
- Digital Forensics Investigator
- Penetration Tester

Depending on its mandate, a CSIRT may need to hire staff to cover several roles. For a new CSIRT employee who is not already a specialist in another role, or employed to cover a specific area of work, a role that may be of particular interest is that of Cyber Incident Responder. The primary purpose of this role is to monitor the state of cybersecurity in an organization, handle incidents during cyber attacks, and assure the continued operations of ICT systems. The mission of a Cyber Incident Responder is to:

- monitor and assess the cybersecurity state of systems;
- analyse, evaluate, and mitigate the impact of cybersecurity incidents;
- identify the root causes of cyber incidents and any relevant malicious actors; and
- restore the functionalities of systems and processes to an operational state, collecting evidence and documenting actions taken, according to the Incident Response Plan of an organization.



33 See: ENISA 'European Cybersecurity Skills Framework (ECSF)', https://www.enisa.europa.eu/topics/education/european-cybersecurity-skills-framework

34 ENISA, ECSF: European Cybersecurity Skills Framework, September 2022. See: https://www.enisa.europa.eu/publications/european-cybersecurity-skills-framework-role-profiles

35 ENISA, User Manual: European Cybersecurity Skills Framework (ECSF), September 2022. See: https://www.enisa.europa.eu/publications/european-cybersecurity-skills-framework-ecsf?v2=1

The main tasks of this role are:

- contributing to the development, maintenance, and assessment of the Incident Response Plan;
- developing, implementing, and assessing procedures related to incident handling;
- identifying, analysing, mitigating, and communicating cybersecurity incidents;
- assessing and managing technical vulnerabilities;
- measuring the effectiveness of cybersecurity incident detection and response;
- evaluating the resilience of cybersecurity controls and mitigation actions taken after a cybersecurity or data breach incident;
- adopting and developing incident handling testing techniques;
- establishing procedures for incident results analysis and incident handling reporting;
- documenting incident results analysis and incident handling actions;
- cooperating with SOCs and CSIRTs; and
- cooperating with key personnel for reporting of security incidents, according to the applicable legal framework.

A Cyber Incident Responder must have the skills to:

- practice all technical, functional, and operational aspects of cybersecurity incident handling and response;
- collect, analyse, and correlate cyber threat information originating from multiple sources;
- work on operating systems, servers, clouds, and relevant infrastructures;
- work under pressure;
- communicate, present, and report to relevant stakeholders; and
- manage and analyse log files.

Key knowledge for this role relates to:

- incident handling standards, methodologies, and frameworks;
- incident handling recommendations and best practices;
- incident handling tools;
- incident handling communication procedures;
- operating systems security;
- computer networks security;
- cyber threats;
- cybersecurity attack procedures;
- computer systems vulnerabilities;
- cybersecurity-related certifications;
- cybersecurity related laws, regulations, and legislation;
- the operation of SOCs; and
- the operation of CSIRTs.

Competences for this role are defined by European Standard 16234-1 (2019), the e-Competence Framework (eCF),[36] according to which a Cyber Incident Responder must be competent in:

- A.7. Technology Trend Monitoring (Level 3),
- B.2. Component Integration (Level 2),
- B.3. Testing (Level 3),
- B.5. Documentation Production (Level 3), and
- C.4. Problem Management (Level 4).[37]

36 European Committee for Standardization, 'e-Competence Framework - A Common European Framework for ICT Professionals in all sectors - Part 1: Framework', EN 16234-1, December 2019 (approved October 2019).

37 The e-CF lists competence levels from 1 to 5, indicating increasing knowledge and proficiency. For more on this, see: ENISA, User Manual: European Cybersecurity Skills Framework (ECSF), pp. 25-26.

NICE Framework

The National Initiative for Cybersecurity Education (NICE), a US effort to increase the size and capacity of the US cybersecurity workforce, is another useful resource. Initially focused on federal institutions, NICE later expanded to the private sector and is now a nationwide project involving government, academia, and the private sector, led by the National Institute of Standards and Technology (NIST). The NICE Workforce Framework for Cybersecurity (known simply as the NICE Framework) groups common cybersecurity functions into seven categories, as follows:

- Analyze (review and evaluate information to create intelligence data);
- Collect and Operate (gather cybersecurity information to develop intelligence, and perform specialized deception operations);
- Investigate (investigate cybersecurity events or crimes);
- Operate and Maintain (support, administer, and maintain IT systems);
- Oversee and Govern (lead, manage, direct, develop, and advocate for cybersecurity work);
- Protect and Defend (identify, analyse, and mitigate threats); and
- Securely Provision (conceptualize, design, procure, and build systems).[38]

The NICE Framework defines Specialty Areas (totalling 33) under each of these categories, and Work Roles (totalling 52) under each Specialty Area. These Work Roles are specific jobs in the cybersecurity field requiring knowledge, skills, and abilities relevant to a given Specialty Area.

One outcome of this project that is of particular value is an Education and Training Catalogue, which lists available cybersecurity courses aligned to the Specialty Areas of the NICE Framework.[39] Many of the online courses listed in the Catalogue are available in Europe.

Courses

Since CSIRT employees must receive quality continuing education on a regular basis, it is common for team members to take several courses and trainings per year, based on the needs and budget of the CSIRT. A broad selection of courses are available, ranging from general to very specific, and prices for similar courses can vary significantly. Notably, there are also many useful online courses accessible for free, and a list of these courses, as well as some lower cost options, are published on the NIST website.[40]

A CSIRT employee cannot expect that all their educational needs will be met through courses organized by the team or organization for which they work. Indeed, one of the characteristics of a good CSIRT member is a natural curiosity to seek more information and new knowledge in any areas of interest; and self-education of this sort is considered normal and expected in the field. Below is an overview of just some of the courses that have proven interesting and useful to CSIRT employees.

TRANSITS I and II

TRANSITS (Training of Network and Security Incident Teams Staff) courses provide 'state of the art, high quality training to CSIRT personnel... for commercial, governmental, military and national CSIRTs, as well as those in the research and education sector'.[41]

TRANSITS I is intended for new employees on CSIRTs, or people who aspire to become CSIRT employees.[42] The material was developed by TF-CSIRT, which regularly hosts this course; but it is sometimes hosted by other organizers. The TRANSITS I curriculum features four modules covering the most important aspects of CSIRT work:

38 See: National Initiative for Cybersecurity Careers and Studies, 'Workforce Framework for Cybersecurity (NICE Framwork)', 30 June 2022, https://niccs.cisa.gov/workforce-development/nice-framework; also see: National Institute of Standards and Technology, 'National Initiative for Cybersecurity Education (NICE)', https://www.nist.gov/itl/applied-cybersecurity/nice

39 National Initiative for Cybersecurity Careers and Studies, 'NICCS Education and Training Catalog', 30 June 2022, https://niccs.cisa.gov/education-training/catalog

40 National Institute of Standards and Technology, 'Free and Low Cost Online Cybersecurity Learning Content', 20 October 2022, https://www.nist.gov/itl/applied-cybersecurity/nice/resources/online-learning-content

41 TF-CSIRT, 'TRANSITS', https://tf-csirt.org/transits/

42 TF-CSIRT, 'TRANSITS I', https://tf-csirt.org/transits/transits-events/transits-i/

- Organizational (defining the structure and services of a CSIRT, its place within an organization, and communication strategies);
- Technical (discussing the methods and means of cyber attackers, various kinds of attacks, how protocols can be abused, system vulnerabilities, and investigative techniques);
- Operational (covering the incident handling process); and
- Legal (introducing European legislation relevant to the work of CSIRTs).

The material for the TRANSITS I course is publicly available on Github and may be used under the Creative Commons license.[43]

TRANSITS II is intended for slightly more informed CSIRT employees, with at least one year of work experience.[44] This course was also developed by the TF-CSIRT community and includes the following modules:

- NetFlow Analysis (i.e., how to analyse traffic flow log data captured in routers and switches);
- Forensics (i.e., how to collect evidence when networks and systems are compromised);
- Communication (i.e., how to liaise with constituents, funders, and management); and
- CSIRT exercises.

Unlike TRANSITS I, the materials for TRANSITS II are not publicly available and are the property of the authors.

ENISA training resources

Over the last 15 years, ENISA has compiled a selection of publicly available online training materials in four areas:

- Technical (e.g., analysis and forensics);
- Operational (e.g., incident handling);
- Setting up a CSIRT (e.g., staff recruitment and infrastructure development); and
- Legal and Cooperation (e.g., contacts and communication).[45]

These materials include handbooks for instructors, toolsets for students, and virtual images for hands-on exercises. To support the preparation and implementation of trainings, ENISA has also developed two publications: Good Practice Guide on Training Methodologies: How to become an effective and inspirational trainer,[46] and **Roadmap to provide more proactive and efficient Computer Emergency Response Team training.**[47]

Cyber exercises

Preferably, CSIRT employees will test their knowledge in exercises that help them prepare for a real incident. There are two general types of exercises:

- Table-top – the procedures, roles, and responsibilities involved in responding to an incident are reviewed through discussion of scenarios and answers to questions from participants
- Live – an incident is simulated on a pre-prepared platform and participants must complete certain tasks

Cyber range platforms

Cyber range platforms simulate a realistic local or online environment and are safe for practice because any activities take place in a closed environment without the possibility of spilling over into the real world. This is true even if participants can connect to that environment from a remote

43 See: https://github.com/GEANT/TRANSITS

44 TF-CSIRTS, 'TRANSITS II', https://tf-csirt.org/transits/transits-events/transits-ii/

45 ENISA, 'Training for Cybersecurity Specialists: Online training materials', https://www.enisa.europa.eu/topics/trainings-for-cybersecurity-specialists/online-training-material

46 Lauri Palkmets, Good Practice Guide on Training Methodologies: How to become an effective and inspirational trainer, No. TP-04-14-898-EN-N (ENISA, 2014). Available as a PDF at: https://www.enisa.europa.eu/publications/good-practice-guide-on-training-methodologies/at_download/fullReport (accessed 4 December 2022).

47 Miroslaw Maj and Don Stikvoort, Roadmap to provide more proactive and efficient Computer Emergency Response Team training (ENISA, 2012). Available as a PDF at: https://www.enisa.europa.eu/publications/roadmap-to-provide-more-proactive-and-efficient-cert-training/at_download/fullReport

location via the Internet. During exercises on cyber range platforms, attacks are simulated with realistic IOCs and participants are asked to identify all the factors and resolve the incident.

Tools for creating Internet simulations

Among the tasks of the Carnegie Mellon SEI is the development of cybersecurity trainings and exercises for US government institutions such as the Department of Defense, National Security Agency (NSA), and Federal Bureau of Investigation (FBI). As part of this activity, the SEI develops simulation tools, including:

- TopGen, an offline exercises simulator;
- GreyBox, an emulator of the Internet backbone;
- GHOSTS, a simulator of virtual, non-player users;
- vTunnel, a communication tool;
- WELLE-D, an emulator of wireless communications; and
- TopoMojo, a web application for the creation and deployment of virtual labs.

All the above tools can be downloaded for free from the Carnegie Mellon University repository.[48]

CTF

In a Capture the Flag (CTF) competition, contestants (individuals or teams) solve cybersecurity tasks (commonly called challenges) from different domains: cryptography, networking, operating systems, web, social engineering, mobile devices, IoT, privacy, and more. The goal of each task (or challenge) is to find and capture a 'flag' hidden in vulnerable software, a website, or even hardware. The flag may be hidden in the CTF organizer's infrastructure (jeopardy-style CTF), or competitors may steal flags from other competitors (attack/defence-style CTF).

These competitions simulate real-life scenarios, but in an isolated environment, and represent a practice platform for security analysts, forensics experts, pen testers, and other cybersecurity professionals, including CSIRT employees.

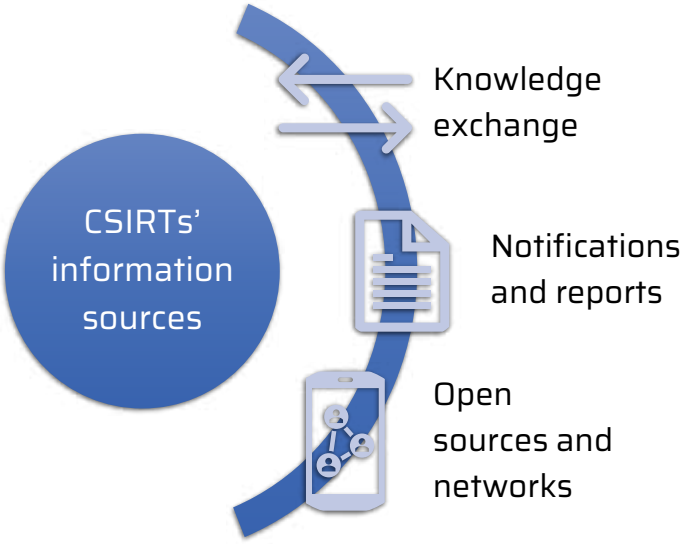
Information collection

A CSIRT must have reliable and accurate information whenever it is needed. For instance, the ability to quickly obtain information such as how to patch a newly discovered critical vulnerability can prevent serious security problems in an ICT system. In the event a cyber attack is underway, CSIRTs can respond more effectively if they know where to obtain information on whether a similar attack has occurred before and how it was countered. CSIRT staff may obtain some information independently, for example through log files or by analysing the behaviour of malware, but external sources are also indispensable to the functioning of a CSIRT.

There are three main ways CSIRTs obtain information from external sources:

- through knowledge exchange within the CSIRT community;
- via notifications and reports from known reliable sources outside the CSIRT community (including commercial and specialty sources); and
- from other open sources and personal sources (public web forums, social networks, private contacts, etc.).

48 See: https://github.com/orgs/cmu-sei/repositories



Any information should always be double-checked before it is used, no matter how it was obtained. The type or extent of verification process applied to information depends on its source and nature, and various other factors such as time constraints.

Information handling

Storage

The work of a CSIRT involves processing large amounts of information. This entails certain risks, including that important information may be overlooked or that information which needs to be protected is inadequately stored or transported. The accidental disclosure of even one piece of sensitive information can bring serious damage to the reputation and perceived trustworthiness of a CSIRT. In fact, when essential information is disclosed, it is sometimes not accidental but is the deliberate work of a malicious actor aiming to discredit a CSIRT.

As the storage and disposal of data is a key task of a CSIRT, with the aim to prevent any disclosure of potentially sensitive information, CSIRTs tend to keep any information related to every incident they have handled (and this may be a legal obligation if the incident is a criminal offense). It is a good practice to store all information on incidents in a separate repository. It is also a good practice to make backup copies of this information and store it separately. If ICT equipment must be handed over to another party (e.g., due to replacement, failure, etc.), permanent data storage should be physically destroyed or securely erased before it is handed over.

Categorization

Certain types of information are more sensitive than others and must be properly categorized in order to be appropriately protected. The process of placing information into categories is often called classification, and information placed into sensitive categories is usually referred to as classified information. If information is publicly available, it cannot be classified in any of the sensitive categories.

Different CSIRTs use different categorization schemes, and a single CSIRT may use multiple schemes in parallel – such as one for information sharing with other member entities in an organization to which the CSIRT belongs, a second for internal information categorization, and a third for information sharing with third parties. Any schemes used by a CSIRT must be well defined, and must determine aspects like categorization criteria and the relationships between categories in different schemes. Regardless of how many categorizations schemes a CSIRT uses, every CSIRT employee must be aware of them all and must understand them all in order to avoid making a mistake that could have serious consequences.

Categorization schemes delineate time periods in which information should be treated as sensitive, and after which it can be made available to the public. This does not mean information must be publicized after these periods expires, but additional measures to protect that information are no longer applied. If no such period is specified in a categorization scheme, information should be treated as sensitive until the owner of the information or another authorized person opts to remove the classification (i.e., de-classify the information). National classification schemes always define such a period for each classification level.

Some publicly known categorization schemes are:

- European Commission Decision on the security rules for protecting EU classified information[49]
- NIST Standards for Security Categorization of Federal Information and Information Systems[50]
- Carnegie Mellon University Guideline for Data Classification[51]

Information sharing in the CSIRT community

Cooperation between CSIRTs is vitally important, as it is only through networking and collaboration that important information and warnings can be spread quickly. Keep in mind that some information is not publicly available or published on the Internet, but is passed on informally, through personal contacts; and to be in a position to receive such information, a CSIRT must be seen as trustworthy, and its team members must establish relationships with members of other CSIRTs. These relationships tend to be established through real-world interactions, which is why CSIRT employees should attend events organized for CSIRTs, meet colleagues from other teams and maintain contact afterwards.

There is an expectation that a CSIRT member will share any relevant information they possess with their circle of acquaintances and friends from other CSIRTs. But there are examples of more extensive and deeper collaboration between CSIRTs as well, including on common projects (such as awareness raising) and in resolving incidents (to compensate for a lack of capacity). When CSIRTs frequently cooperate and share mutual trust, it is a good practice to sign a Memorandum of Understanding (MoU) that precisely defines the scope of cooperation and the obligations or expectations of each team.

Of course, this kind of cooperation is not always possible. Sometimes, constituents prefer that a CSIRT does not spend resources assisting another team or that information about attacks and vulnerabilities is not shared. For this reason, approval for collaborations with other CSIRTs should always be obtained from management in advance.

ENISA has issued a Good Practice Guide on Information Sharing.[52]

What is exchanged and how?

Hundreds of thousands of different variants of malware and other threats appear in the world every day. It is impossible for one CSIRT to collect or process all this information, so defending against attacks requires more than a single team. Thus, within the cybersecurity community, there is an intensive exchange of information about threats, but also about how to detect malware or intruders in a system and how to remove threats. Initially, a CSIRT with less knowledge and experience will only collect information for the purposes of better protecting their own systems and responding to incidents, but over time, they will gain enough skills to contribute to the community by releasing information and providing advice.

49 European Commission, Decision on the security rules for protecting EU classified information, Official Journal of the European Union No. 2015/444, 13 March 2015. Available as a PDF at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015D0444&from=GA

50 US Department of Commerce, National Institute of Standards and Technology, 'Standards for Security Categorization of Federal Information and Information Systems', FIPS PUB 199, February 2004. Available as a PDF at: https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.199.pdf

51 Carnegie Mellon University, 'Guidelines for Data Classification', 21 February 2021, https://www.cmu.edu/iso/governance/guidelines/data-classification.html

52 ENISA, 'Good Practice Guide: Network Security Information Exchanges', 13 June 2009. Available at: https://www.enisa.europa.eu/publications/good-practice-guide

Need to Know principle

Even when your organization agrees to share information with another CSIRT, and there are appropriate reasons and methods for doing so, you must be careful about the information you provide. When deciding on what to share, always be mindful of whether the information is useful to a recipient and whether there is a distinct purpose to sharing it. In short, do not provide any unnecessary information. Any information you do share should help solve a problem or clarify the bigger picture. This principle of sharing only minimal but necessary information is called the Need-to-Know principle.

Sometimes, beyond the Need-to-Know standard, the law restricts information sharing. If information possessed by a CSIRT is categorized as confidential and another CSIRT is not authorized to access information classified at that level, it cannot be shared, no matter how desperately it is needed.

TLP protocol

The Traffic Light Protocol (TLP) was designed to allow information providers to easily inform recipients as to how they should handle information.[53] The TLP stipulates four possible labels:

- TLP:RED, indicating that the information cannot be shared outside the group to whom the information provider shared it (this type of information is often shared verbally, in a closed meeting);
- TLP:AMBER, indicating that the information can be shared within the organizations to which any recipients of the information belong, and their clients (TLP:AMBER+STRICT is a variant of this label that restricts sharing to the organization only);
- TLP:GREEN, indicating that the information can be shared within the cybersecurity community, but must not be made public; or
- TLP:CLEAR (formerly TLP:WHITE), indicating that the information can be made publicly available.

This protocol is widely accepted in the international CSIRT community, and because trust is one of the core characteristics of the community, any label assigned by an information provider must be strictly respected. If a recipient is ever faced with a legal obligation or other reason to transfer information beyond the recipients defined by the TLP label received with that information, the information provider must be notified.

MISP

One of the most popular platforms for information exchange is MISP (this abbreviation has been retained even though its former name, Malware Information Sharing Platform, has been changed to Open-Source Threat Intelligence and Sharing Platform).[54] Information about threats, vulnerabilities, indicators of compromise, and malicious actors is shared on the platform, as well as intelligence data that can contribute to preventing an incident or resolving one more effectively.

MISP is maintained by the Computer Incident Response Center Luxembourg (CIRCL), which allows free access to the platform.[55] An installation version of the application is also available as a free download, and your own instance of MISP can be synchronized to exchange information with other MISP instances.[56]

Information Sharing and Analysis Centre (ISAC)

To improve the exchange of knowledge and information about cyber threats, the EU initiated the creation of ISACs. These non-profit organizations enable two-way communication between the

53 Forum of Incident Response and Security Teams, 'Traffic Light Protocol (TLP): FIRST Standards Definitions and Usage Guidance – Version 2.0', https://www.first.org/tlp/

54 See: https://www.misp-project.org/features/

55 See: https://www.circl.lu/

56 Learn where MISP can be downloaded at: https://www.misp-project.org/download/

private and public sectors about threats, incidents, methods, and experiences, mainly related to critical infrastructure. ISACs and CSIRTs have different but compatible tasks and their cooperation can be very mutually beneficial.

International associations

With the aim of facilitating communication and cooperation among CSIRTs, associations have been formed to bring teams together and provide them platforms for organizing, information exchange, and other activities that can improve their work. Among the most important of these international associations are FIRST and TF-CSIRT.

FIRST

The chaos caused in 1988 by the 'Morris worm', described earlier, exposed how unprepared the Internet community was for a cyber incident, as its response was uncoordinated, unsystematic, and inconsistent. Consequently, the CERT/CC was formed with the support of the US government, and in the following years, other similar teams were formed and began to cooperate. But bilateral cooperation, especially at an international level, was burdened by various problems including national legislation, communication in different languages, operations in different time zones, and more. The need to establish some form of joint coordination and communication was obvious, however, so the Forum of Incident Response and Security Teams (FIRST) was founded in 1990.[57] This community of CSIRTs has grown continuously since then, and now comprises more than 600 teams from around the world. Among other things, FIRST provides members with access to documents on best practices; organizes technical colloquiums, practical trainings, and conferences on incident response topics; provides various publications and web services; and hosts special interest groups.

Every CSIRT that is a member of FIRST is listed on the Forum's website.[58] Contact information for teams is available there as well, along with information such as the date a team was established, its constituency, and its PGP public key.

TF-CSIRT

In 2000, the Task Force on Computer Security Incident Response Teams (TF-CSIRT) was launched as an academic initiative intended to improve cooperation within the European CSIRT community.[59] Over time, TF-CSIRT has become a trusted forum for the exchange of experiences and knowledge and is not limited only to European teams.

Trusted Introducer, a service of TF-CSIRT, functions as a data exchange platform for all CSIRTs.[60] The service publishes a list of teams that are recognized as credible by other CSIRTs, and currently numbers more than 450.[61] Trusted Introducer also provides additional information about every team on the list, including contact information, its constituency, and its PGP key. On top of this, teams with significant capacities can go through a procedure and receive the status of an accredited or certified team, which verifies their level of maturity. Only accredited and certified teams have access to additional services that enable more efficient and effective interaction.

Gathering information from reliable sources

Beyond internal resources and CSIRT exchanges, trusted external sources can provide valuable information regarding detected incidents. External sources collect this information through threat intelligence activities and deliver it as real-time data streams called threat intelligence feeds. These feeds can be delivered in variety of formats (plain text, CSV, JSON, etc.) and provide information about malware URLs, phishing sites, C&C servers, infected machines, vulnerable

57 See: https://www.first.org/about
58 See: https://www.first.org/members/teams/
59 See: https://tf-csirt.org/
60 See: https://www.trusted-introducer.org/
61 See: https://www.trusted-introducer.org/directory/alpha_LICSA.html

services, and much more. They can be automatically implemented into security systems after download.

In addition to feeds, information about threats and vulnerabilities can be received in the form of reports, which cannot be automatically implemented into security systems.

A publication from ENISA entitled 'Proactive Detection – Measures and Information Sources' offers an analysis of methods, tools, activities, and information sources that can improve incident detection.[62]

Shadowserver

The Shadowserver Foundation collects a constant stream of information about attacks and suspicious activities from the significant number of honeypots it has established around the world,[63] which lure attackers (see the 'Indicators of Compromise' section). Shadowserver learns even more about malicious activity by scanning the entire IPv4 Internet over one hundred times a day and engaging at a high level with network providers, law enforcement agencies, and governments. Samples of any detected malware are analysed in sandboxes and any information obtained is stored in the Shadowserver database. According to the Shadowserver website, more than one million new unique malware samples are analysed every day and entered into a malware repository containing over 1.7 billion samples.

Shadowserver provides daily reports to its subscribers free of charge, and each of these reports is tailored to the subscriber (e.g., national CSIRTs receive reports filtered by country). National CSIRTs and law enforcement agencies can also benefit from additional Shadowserver services, such as data analysis, training, or investigation support.

Open-source intelligence (OSINT)

OSINT refers to information gathering from publicly available sources and is used by various actors for different purposes that may or may not be related to cybersecurity. In the context of cyber threat intelligence (CTI), the source most used for OSINT is the Internet. Some information can be gathered using regular browsers such as Chrome, Firefox, or Internet Explorer; though, these browsers can only collect information from the surface Internet, leaving the content of the deep web undiscovered.

Platforms that facilitate the collection and processing of open-source information include: Shodan,[64] theHarvester,[65] Maltego,[66] and recon-ng.[67]

Shodan

The Shodan search engine is designed to thoroughly search any web-connected device, including IoT devices like web cameras, baby alarms, home appliances, industry control systems, etc. Shodan works by trying to connect every possible port of every possible IP address on the Internet and asking for a response, in a constant loop. It collects responses that contain important metadata such as device name, IP address, port number, organization, location, and more; and after analyzing this data, it can detect various vulnerabilities. These features make Shodan a favourite tool of pen testers, security researchers, and law enforcement agencies, but also of cybercriminals.

With simple search options and for a limited number of searches, Shodan is free of charge.

62 Piotr Białczak, Paweł Pawliński, Kryzysztof Rydz, and Rossella Mattioli, 'Proactive Detection – Measures and Information Sources', ENISA, May 2020. See: https://www.enisa.europa.eu/publications/proactive-detection-measures-and-information-sources

63 See: https://www.shadowserver.org/

64 See: https://www.shodan.io/

65 See: https://github.com/laramies/theHarvester

66 See: https://www.maltego.com/

67 See: https://github.com/lanmaster53/recon-ng

Maltego

Maltego is an open-source intelligence tool that also features graphical link analysis, which helps connect information for investigative tasks.[68] Maltego offers a free Community plan with basic functionality and limited access to the database, but more advanced versions are paid.

What are we protecting?

In general, information security protects information, and preserves our ability to store, transmit, and process that information. Because various telecommunication and IT devices, as well as a range of software, are used to handle information, protecting these devices and software falls under the umbrella of information security. Hardware, software, and information in any form are collectively referred to as assets. These represent value to an organization; meaning, anything sufficiently valuable to an organization to warrant protection is an asset.

When information security measures are applied to assets, the measures applied to each asset, and to what extent, depend on the value of those assets to the organization. In theory, the cost of protecting an asset should not exceed the value of that asset. But it must be emphasized that, in this context, 'value' does not refer only to the financial worth of a device or service. For example, if attackers take down the server of a company that provides website hosting services, the company can compensate its customers for any time the service was unavailable; however, if similar attacks are repeated on several occasions, clients will begin to leave the provider. Then, damage to the company will exceed the sum it must pay back to clients for a service interruption. The fact that cyber attackers consider it a success to threaten the reputation of their victims should be considered when developing protection measures.

It is also important to keep in mind that there is no such thing as perfect protection. The protocols used are imperfect, devices are imperfect, and people are imperfect and often make mistakes. Indeed, even device and software manufacturers are unaware of many of the security vulnerabilities in our systems. Hence, CSIRT employees must always be vigilant about security issues, monitor information about vulnerabilities that have been discovered, and react quickly and efficiently in response to incidents.

In principle, information security measures in an organization are not established, maintained, or administered by the CSIRT, but the team can provide necessary professional and technical assistance to the people doing these jobs and can assess whether security problems exist despite the security measures and devices in place.

The supply chain

Because many large organizations implement strong security measures, attackers will sometimes focus their efforts on smaller companies that supply devices or software to these larger organizations but do not have the resources or security culture to implement the same strict security measures. In this way, through the vendor's product, a large organization enters a malicious code into its system that allows an attacker to access its resources.

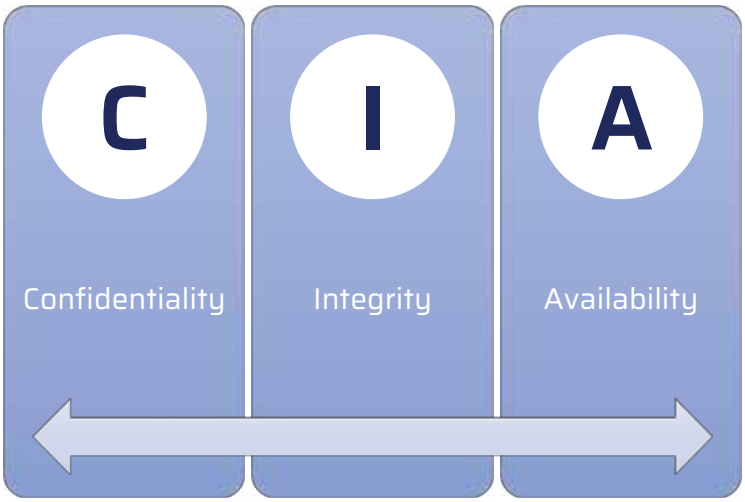
There have been several such attacks in recent years, and some have been extremely damaging to the victims. For this reason, significant efforts are being made to establish product verification and certification systems to reduce the risk of supply chain attacks.

The three basic security properties

Information security measures should assure three basic security properties: confidentiality, integrity, and availability – otherwise known as 'CIA', and often referred to as the 'CIA triad'.

68 See: https://www.maltego.com/

- Confidentiality implies that some information or other resources are inaccessible to everyone except authorized users.
- Integrity means that data has not been changed without authorization.
- Availability infers that data or other resources can be accessed by authorized users at will.



What are we defending against?

Attacks on ICT systems target assets. Some attackers apply one method of attack to a large number of users, hoping that potential victims will emerge from among less cautious users or those with less secure systems. Completely different and much more dangerous are attackers who study a specific victim to identify weaknesses in their information security measures, use various methods to enter their system unnoticed, then explore further opportunities within that system. Depending on their ultimate goals, these attackers typically extract information from a system for as long as possible while remaining undetected, or initiate visible activity only after they are fully prepared to put an organization in a position from which it cannot defend itself without consequences.

The tactic of remaining unnoticed for as long as possible is called **Advanced Persistent Threat (APT)** and represents one of the biggest dangers posed by cyber attackers to information security. In cases where attackers manage to gain complete control over a system, they can do damage that leads to the complete collapse of the targeted organization. This type of attack requires considerable resources and knowledge, so the prevailing opinion is that most attacks of this nature are sponsored by states.

Threats, vulnerabilities, and risks

If an asset has a security weakness, or if the equipment or method to protect that asset has such a weakness, then we refer to that asset as vulnerable, and to that weakness as a **vulnerability.** Another factor is a threat, which is any action that may endanger our assets, so that they are damaged, altered, disclosed, unavailable, or otherwise no longer under our absolute control. Moreover, inaction may represent a threat; for example, the inactivity of responsible persons to protect vulnerable assets.

Threat agents can be people, programmes, or hardware, deliberately aimed at exploiting vulnerabilities. When the threat agent is a person, they are commonly known as a 'malicious actor'. But not all threats are due to malicious intent. Natural disasters, fire, human error, and technical problems can represent threats to security as well.

A **risk** exists if there is a vulnerability in a system and a threat that can exploit this vulnerability. Thus, risk is assessed by judging the likelihood that a threat can exploit a vulnerability; and the higher this probability, the higher the risk.

If a malicious actor seeks to exploit a vulnerability, then an **attack** has occured. If that actor only manages to bypass security measures, this is called a **breach.** If, after this breach, they exploit the vulnerability, this is called a **penetration** (or **intrusion**).

Cyber threat intelligence (CTI)

Cyber threat intelligence is the process of gathering and analysing information with the goal of understanding the motives and methods used in attacks, and their targets. Some organizations take a more generalized approach to CTI that is intended for use by management and is based mostly on analysis of global trends and risks; but in most cases, the information collected in the context of CTI is much more nuanced, ranging from the tactics, techniques, and procedures used by malicious actors, to signs that an attack is in progress (such as indicators of reconnaissance or delivery).

CTI information often relates to the capabilities, motivations, and organizational and personal details of attackers, and corresponding IOCs. This information may be collected from a variety of sources, including publicly available websites, social media, paid services, internal logs, or past experiences. Due to the huge amount of information that is constantly collected for the purposes of CTI, machine learning is often used to support analysis.

STRIDE model

A tool developed by Microsoft to model the threats posed to operating systems and applications, known as the STRIDE model, defines six categories of threat:

- Spoofing – attackers present themselves as someone or something else (an individual, entity, asset), threatening authenticity.
- Tampering – attackers modify computer code or data, threatening integrity.
- Repudiation – attackers deny activities that cannot be proven otherwise, threatening non-repudiation.
- Information disclosure – attackers access data to which they are not entitled, threatening confidentiality.
- Denial of service – attackers disable or degrade services to users, threatening availability.
- Elevation of privilege – attackers gain unauthorized access, threatening authorization.[69]

Although this model is focused on software, it can also be applied to other types of threats, such as those to networks or data centres.

The Common Vulnerability Scoring System (CVSS)

The CVSS is an open framework developed by FIRST to evaluate the nature and severity of software vulnerabilities.[70] Software vulnerabilities are scored in three metric groups:

- Base – vulnerability characteristics that are invariant over time and across environments.
- Temporal – vulnerability characteristics that change over time but do not depend on user environment.
- Environmental – vulnerability characteristics that depend on a specific user environment.

A Base score is derived from assessment of Attack Vector (AV), Attack Complexity (AC), Privileges Required (PR), User Interaction (UI), Scope (S), Confidentiality Impact (C), Integrity Impact (I), and Availability Impact (A) metrics; a Temporal score from Exploit Code Maturity (E), Remediation Level (RL), and Report Confidence (RC) metrics; and an Environmental score from Confidentiality

69 See: 'Microsoft Threat Modeling Tool', 25 August 2022, https://docs.microsoft.com/en-us/azure/security/develop/threat-modeling-tool-threats

70 The first version of the CVSS was released in 2005 by the National Infrastructure Advisory Council (NIAC); soon afterward, FIRST was chosen to develop it going forward. See: https://www.first.org/cvss/user-guide

Requirement (CR), Integrity Requirement (IR), and Availability Requirement (AR) metrics. Individual metrics are scored, then scores for each metric group are calculated separately and are also combined into an overall score. The scores for each metric group, and the overall score, can range from 0 to 10 (from none to most severe).

Depending on the Base score, the severity of software vulnerability is rated in one of five categories:

- None (0.0)
- Low (0.1-3.9)
- Medium (4.0-6.9)
- High (7.0-8.9)
- Critical (9.0-10.0)

A CVSS score can be represented as a vector string showing the value of each metric.[71]

The Common Vulnerabilities and Exposures (CVE) Program

The CVE Program catalogues known software vulnerabilities using a system of unique identifiers. Each vulnerability is assigned a unique string, in the format:

CVE-<year>-<number>[72]

These strings are called CVE identifiers (CVE IDs), CVE names, CVE numbers, or simply CVEs. Every CVE ID, as well as other data, has been compiled into a CVE database that can be downloaded in various formats,[73] or searched online.[74] These records are maintained by the Mitre Corporation, which is the primary authority in the process of assigning CVEs, along with others such as CERT/CC, Microsoft, and Oracle.

The Common Weakness Enumeration (CWE)

Though it is community developed, the CWE is also operated by the Mitre Corporation. The CWE is a list of common weaknesses in software (e.g., buffer overflows, user interface errors, authentication errors, etc.) and hardware (e.g., issues associated with CPUs, graphics, access control, power, RTC, etc.).[75] The severity of these weaknesses is scored using two tools: the Common Weakness Scoring System (CWSS),[76] and the Common Weakness Risk Analysis Framework (CWRAF).[77]

The National Vulnerability Database (NVD)

The NVD, a US government repository of vulnerability information, is available on the NIST website.[78] It incorporates a CVSS calculator,[79] and offers examples of CVSS that break a score down into its component parts.[80] The NVD is maintained by NIST and sponsored by US-CERT.[81]

Coordinated vulnerability disclosure (CVD)

Vulnerability hunting is a daily task for many cybersecurity researchers and professionals, to identify weaknesses before malicious users can find and exploit them. Through CVD, information about vulnerabilities is shared among researchers, security professionals, vendors, and providers. The goal is to create solutions before exploitation occurs.

71 An example of a CVSS Base score vector is: AV:A/AC:H/PR:L/UI:R/S:C/C:L/I:L/A:N. This means: Attack Vector is Adjacent Network (AV:A), Attack Complexity is High (AC:H), Privileges Required is Low (PR:L), User Interaction is Required (UI:R), Scope is Changed (S:C), Confidentiality Impact is Low (C:L), Integrity Impact is Low (I:L), and Availability Impact is None (A:N). The CVSS Base score for these parameters is 4.0.

72 This number appears in sequence but can be an arbitrary length.

73 See: https://www.cve.org/Downloads

74 See: https://cve.mitre.org/cve/search_cve_list.html

75 See: https://cwe.mitre.org/data/index.html

76 See: https://cwe.mitre.org/cwss/cwss_v1.0.1.html

77 See: https://cwe.mitre.org/cwraf/

78 See: https://nvd.nist.gov/

79 See: https://nvd.nist.gov/vuln-metrics/cvss#

80 See: https://nvd.nist.gov/Vulnerability-Metrics/Calculator-Product-Integration

81 US-CERT is part of the Cybersecurity and Infrastructure Security Agency (CISA).

CVD analyses and guides have been published by many leading organizations in the field, including ENISA,[82] Carnegie Mellon University,[83] the US Cybersecurity and Infrastructure Security Agency (CISA),[84] and the National Cyber Security Centre (NCSC) of The Netherlands.[85]

A brief overview of the most common techniques used by attackers

Depending on the target of an attack, different techniques may be applied by the attacker(s). Perhaps the most well-known technique is the execution of malware on a victim's system, but there are many more, including password guessing, denial of service (DoS), social engineering, etc.

Tactics, Techniques, and Procedures

Tactics, Techniques, and Procedures (TTP) – a concept taken from the US Army – describes the behaviour of attackers in a hierarchical way; meaning that tactics refer generally to their behaviour, techniques to a more detailed layer of those tactics, and procedures to a more specific description of those techniques.

Mitre ATT&CK

Mitre ATT&CK is a publicly available knowledge base of tactics and techniques used by attackers.[86] The ATT&CK Matrix for Enterprise incorporates the following tactics:

- Reconnaissance
- Resource Development
- Initial Access
- Execution
- Persistence
- Privilege Escalation
- Defense Evasion
- Credential Access
- Discovery
- Lateral Movement
- Collection
- Command and Control
- Exfiltration
- Impact

For each of these tactics, several techniques (and sub-techniques) used by attackers are identified and described in detail. These descriptions include examples from practice, as well as recommendations for mitigation and detection.

Common Attack Pattern Enumeration and Classification (CAPEC)

CAPEC is a catalogue of common attack patterns used by attackers to exploit weaknesses.[87] Each attack pattern is logged with a description, an execution flow, the prerequisites for an attack of that type, and mitigation tips. Like ATT&CK, CAPEC is operated by the Mitre Corporation. The two tools are similar, but ATT&CK focuses on network defence while CAPEC focuses on application security.[88]

82 See: 'Coordinated Vulnerability Disclosure policies in the EU', 13 April 2022, https://www.enisa.europa.eu/news/enisa-news/coordinated-vulnerability-disclosure-policies-in-the-eu

83 Allen D. Householder et al., The CERT Guide to Coordinated Vulnerability Disclosure, Special Report CMU/SEI-2017-SR-022 (Carnegie Mellon University, 2017). Available as a PDF at: https://resources.sei.cmu.edu/asset_files/SpecialReport/2017_003_001_503340.pdf

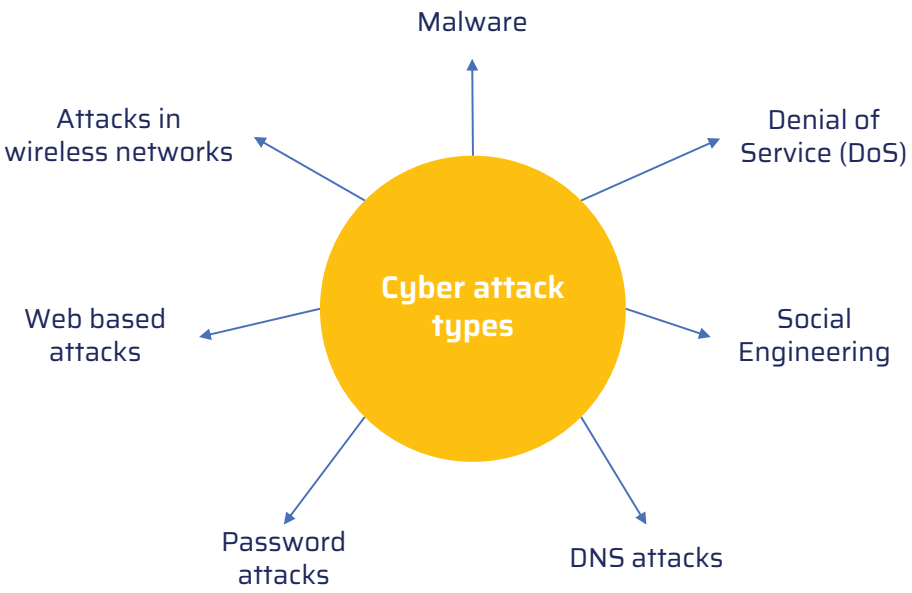
84 'CISA Coordinated Vulnerability Disclosure (CVD) Process', https://www.cisa.gov/coordinated-vulnerability-disclosure-process

85 NCSC, 'Coordinated Vulnerability Disclosure: the Guideline', October 2018. See: https://english.ncsc.nl/publications/publications/2019/juni/01/coordinated-vulnerability-disclosure-the-guideline

86 See: https://attack.mitre.org/

87 'About CAPEC', 4 April 2019, https://capec.mitre.org/about/index.html

88 See: CAPEC, 'ATT&CK Comparison', 16 October 2019, https://capec.mitre.org/about/attack_comparison.html



Malware

Malware is the common name for any piece of code written with malicious intent and harmful to the system on which it runs. Malware can be classified into several categories based on its method of transmission and impact on a victim's system, but in every case, an attacker must find a way to insert malware into a targeted system and initiate its execution. This allows an attacker to control a victim's computer, access their documents, exploit resources on their computer, or conduct other malicious activities, depending on the purpose and capabilities of the malware.

Viruses

A virus is a type of malicious code that needs a host file and typically spreads by seeking new hosts and copying itself to them. The virus code is merged with the host file code in such a way that lines of virus code are executed when the host file is executed. The virus contains a malicious payload that is activated when certain conditions are met.

Worms

Unlike viruses, worms do not need a host file to spread. Instead, worms spread by using vulnerabilities in network and information systems to copy their code to another computer and initiate execution. However, worms are very similar to viruses in other ways and are always looking for a means to spread further and activate their payload.

Trojans

Trojans, or Trojan Horses, are a type of malicious code that is embedded inside a legitimate programme. Unlike viruses and worms, Trojans are not designed to spread, only to activate their payloads when the programme in which they are embedded is launched.

Rootkits

Malware located in the part of a computer's memory where programmes are automatically executed at start-up is known as a rootkit. Inserting malicious code into this part of a computer's memory is difficult, but also makes it harder to detect.

Ransomware

Upon activation, this type of malware disables access to some of a user's resources, before requesting they pay a ransom to regain access. Ransomware can block access to some of the hardware resources of an information system, but in most cases, it encrypts a user's documents, images, videos, and other files, and demands a ransom in exchange for a decryption key.

Adware

The purpose of adware is simply to display advertisements on a computer screen. Most adware does this through a web browser, though other methods are used as well. While adware may not harm a victim's system in any way, and may advertise benign products and services, the fact that illegitimate means are used for it to enter a system and the user does not consent to the content displayed makes it a kind of malware.

Keyloggers

A keylogger is used to record and temporarily save each of a user's keystrokes and pass them on to an attacker. An attacker may use this method to try and extract usernames and passwords, pin codes, bank cards details, etc.

Crypto miners

This type of malware allows attackers to access the resources of a victim's computer in order to mine cryptocurrencies. An affected user may not only experience slower performance, but also receive higher energy bills. For greater efficiency, attackers often try to infect as many computers as they can, form a botnet, and distribute crypto mining tasks equally to each 'zombie' computer.

Fileless malware

Malware that never exists on a victim's hard drive or any other permanent storage, but only in RAM, is referred to as fileless malware and is very difficult for anti-malware software to detect. This type of malware uses legitimate system files, APIs, or registry files to install and execute malicious code.

Denial of service (DoS)

In a DoS attack, a large volume of nonsensical, erroneous, incomplete, or otherwise fallacious requests to a service are intended to prevent that service from responding to other, legitimate requests.

DDoS

To be effective, the malicious requests sent in a DoS attack must significantly outnumber the capacity of that service to process requests. As this cannot always be achieved from one computer, attackers frequently use a technique known as distributed denial of service (DDoS), in which they simultaneously execute such an attack from several computer systems. Previously infected 'zombie' computers are often used for this purpose, in a botnet network controlled by the attacker.

Social engineering

In social engineering attacks, the goal of attackers is to influence people to do or say something that will provide the information needed for the attackers to access secure systems. To do this, attackers impersonate others, entice people to do something unwise by offering them a reward or coercing them with consequences, convince people they are doing a good deed by providing certain information, etc. Even CSIRT employees are vulnerable to these attacks, despite their knowledge about the techniques used by attackers.

Indeed, instructing others on how to protect themselves does not mean that you are immune to cyber attacks as a CSIRT employee, and this is especially true when it comes to social engineering attacks. After all, CSIRT employees are only human, too. Still, when a CSIRT employee is a victim of any kind of cyber attack, it takes on a different dimension, as it may hurt their reputation among constituents. For this reason, it is very important that CSIRT members are fastidious about implementing security protocols for themselves, and fulfil at least the recommendations they make to others.

Phishing

One of the most common forms of social engineering is phishing. In a phishing attack, a potential victim receives an email in which an attacker convinces them to run a program, open a file attachment, or click a link to a web page. If the victim complies, malicious code will execute. Phishing usually targets many potential victims at once and is relatively easy to recognize. However, more serious attackers come prepared, having selected and researched potential victims in advance.

There are several subtypes of phishing, including: spear phishing (when a phishing email, often impersonating a friend, co-worker, or family member, targets a specific person and addresses them directly in the text); whaling (when attackers target people in high-level positions, i.e., the 'big fish'); vishing (social engineering carried out via a telephone conversation); or smishing (social engineering performed using SMS messaging).

DNS attacks

DNS poisoning

DNS poisoning or DNS spoofing is a type of attack in which attackers redirect users to an illegitimate website. This website may be very similar to the original, and may trick users into entering their usernames and passwords; or, it may show indecent or shocking content and have nothing in common with the original website.

DNS tunnelling

Firewalls and other security applications are sometimes set by default to forward DNS queries and responses without additional verification. Attackers can use this type of protocol as a tunnel, by injecting malicious code into it. If they manage to get into a victim's system in this way, attackers can remain undetected for long periods and steal valuable information.

Password attacks

Dictionary attack

Taking advantage of the fact that a considerable percentage of users never change default passwords, use simple or identity-based passwords, or use the same password for different accounts, some cybercriminals build lists of commonly used passwords (some with millions of entries) and then create corresponding programmes that attempt to penetrate systems using passwords from the list, one-by-one, until they are successful or exhaust the list.

Brute force

If a dictionary attack fails, an attacker seeking to find a user's password is left to try every possible combination of letters, numbers, and special characters. This type of password attack is called a brute force attack and may require enormous amounts of time, depending on the length and strength of the password in question.

Rainbow tables

A rainbow table is a spreadsheet containing passwords, along with corresponding hash values for different hash algorithms. To create such a table, a list of all possible passwords must first be compiled, and then hash values must be calculated for each of them using different hash algorithms. That way, when a particular hash value appears in a password protocol, it can be checked against this extensive collection of hashes to find the corresponding password (as calculating the original password from the hash value is infeasible). This approach is often used by password cracking software.

Web based attacks

SQL injections

In SQL injection attacks, databases are targeted by attackers who use unexpected inputs to compromise a web application and gain unauthorized access to the underlying database. If successful, attackers can bypass authentication, allowing them to access, change, add, or delete data in the database.

SQL injection attacks are relatively easy to prevent by limiting the data a user can enter into a web form, and by limiting the privileges of the database server account. Some examples of SQL injection attacks can be found on the OWASP website.[89]

Cross-site scripting (XSS)

In this type of attack, an attacker injects malicious code into the content accessed by website users. When users visit a website compromised by an XSS attack, they download this malicious code along with other website content. By running code on a victim's computer, an attacker may be able to obtain the victim's credentials and access other accounts, or control their computer and use it to attack other systems.

Cross-site request forgery (XSRF or CSRF)

While an attack using XSS is focused on injecting malicious code into a website, XSRF/CSRF is deployed to trick web browsers into performing actions that are not authorized by the user, such as online shopping, creating a session, downloading cookies, etc. In some variants of XSRF attacks, malicious code remains inactive until a specific website is accessed.

Attacks in wireless networks

We are surrounded by wireless networks. They are literally everywhere – our homes, hotels, airports, and even public places like parks and streets. Some of these networks are better protected than others, and some are not protected at all; but our communications, whether open or encrypted, are available to everyone in our vicinity. Yet, it is important to remember that we often have no knowledge of who controls publicly available access points, and who else can access them.

Man-in-the-middle (MitM)

MitM attacks are especially common in the case of unprotected (open) wireless networks, but they can be implemented in other types of communication networks as well. Attackers may simply eavesdrop on communications and gather information, or they may stand between two communicating parties by pretending to be the other party. Attackers can obtain personal information this way, including usernames, passwords, PINs, or other important information that may serve malicious intentions.

In a subvariant of the MitM attack called session hijacking, an attacker takes over the session between a user and a server on the Internet.

Evil twin

In this type of wireless attack, an attacker mounts a fake base station with the same access point SSID as an existing access point but with a higher power rate. When a user's device tries to connect to the wireless network, the evil twin spoofs the identity of the existing access point and offers an open connection. Since the authentication and encryption is dictated by the base station, the attacker is in a man-in-the-middle position and is able to eavesdrop on the user's communications.

89 See: https://owasp.org/www-community/attacks/SQL_Injection

Rogue access point

A rogue access point attack is similar to an evil twin attack, but does not use the same SSID as an existing access point (though it is often similar to an existing SSID). Instead, an attacker usually mounts a rogue access point in a public place, such as a cafeteria, hotel, park, airport, or other place where people gather. There are also examples, however, of rogue access points being set up on company premises in an attempt to trick careless employees or company visitors.

Terms related to more dangerous attacks

Zero-day

A **zero-day vulnerability** describes an unknown or unpatched security vulnerability that can be exploited by malicious actors. Upon the release of a corresponding patch, this vulnerability is no longer labelled zero-day.

A **zero-day exploit** is a method a malicious actor can use to exploit a zero-day vulnerability.

A **zero-day attack** is a cyber attack in which a malicious user attempts to exploit a zero-day vulnerability before a patch for that vulnerability is released.

Zero-day malware refers to malware for which a specific signature has not yet been identified in anti-malware databases, which means it is undetectable by signature-based anti-malware software.

APT

Advanced Persistent Threat (APT) is the term used for an attack in which the attacker, or (most often) group of attackers, aims to infiltrate an ICT system and remain undetected for as long as they can or want. This type of attack is complex, requiring various specialties on the part of the attackers and, usually, considerable financial resources. Hence, the target must be sufficiently attractive to the attackers, who usually seek to obtain sensitive and confidential information, intellectual property, or personal data, or sometimes intend to erase databases and destroy the ability of the victim to continue operations.

APT attacks have several phases, from the gathering of detailed information about the target, through initial infiltration and the careful elevation of access rights, to the exfiltration of data and complete control over the system. Many of these activities are performed manually by attackers, which separates an APT attack from other types of cyber attacks. Once attackers have acquired enough privileges, they typically establish one or more backdoors to the system, as an access point in case they are discovered or if they decide to temporarily disengage with the victim and return later.

Due to the knowledge, skills, and resources required to carry out an APT attack, they are typically undertaken by groups of attackers oriented and organized around these types of attacks. These APT groups each have their own methods, preferred targets, and tools. A list of APT groups can be found on the Mitre ATT&CK website.[90]

Sources of threat

Cyber threats come from a variety of sources:

- States – motivated by political or economic espionage, a desire to damage the critical infrastructure of another state, or the acquisition of money, the perpetrators of state-sponsored attacks are usually state-sponsored actors who receive funding, intelligence, and other support from that state.
- Companies – induced primarily by economic espionage and to target competing organizations.
- Criminals – driven by financial gain.

90 See: https://attack.mitre.org/groups/

- Terrorists – seeking to promote their ideology, do the kind of damage that can support their goals, or obtain funding for further activities.
- Hacktivists – motivated to emphasize a problem, or to make a political statement against organizations with which they disagree ideologically.
- Curious users – compelled by a sense of succeeding at something, and often associated with the need for recognition from others.
- Insider threats – inspired by diverse motives, whether those of a disgruntled employee, a third-party contractor, or a former employee with access rights, or due to sheer negligence, these threats are particularly dangerous because they come from someone with access rights within the protected area.

How we defend ourselves

Although CSIRT employees do not establish, maintain, or administer security devices and protection systems, they must be well informed regarding the way protection is organized in their organization, and must know which devices and tools are implemented. Only then can CSIRT employees understand the data they receive and effectively organize themselves to react in the event of an incident. It is also important to be knowledgeable about the security standards, principles, and measures that are applied, to be able to assess security and make suggestions on how to improve it.

Security standards

A technical standard is a set of requirements for achieving a certain level of quality and/or quantity in a certain area, and is described in a document that also sets out methods for measuring the fulfillment of those requirements. Security standards are a subset of technical standards related to security.

ISO/IEC 27000 family

The ISO/IEC 27000 family of standards relates to information security.[91] Published jointly by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC), these standards help organizations secure their assets and manage information flows. Standards from this family include:

- ISO/IEC 27000: Overview and vocabulary
- ISO/IEC 27001: Information security management systems — Requirements
- ISO/IEC 27002: Code of practice for information security controls
- ISO/IEC 27003: Information security management systems — Guidance
- ISO/IEC 27004: Monitoring, measurement, analysis and evaluation
- ISO/IEC 27005: Information security risk management
- ISO/IEC 27032: Guidelines for cybersecurity
- ISO/IEC 27033-1: Network security — Part 1: Overview and concepts
- ISO/IEC 27033-2: Network security — Part 2: Guidelines for the design and implementation of network security
- ISO/IEC 27033-3: Network security — Part 3: Reference networking scenarios — Threats, design techniques and control issues
- ISO/IEC 27033-4: Network security — Part 4: Securing communications between networks using security gateways
- ISO/IEC 27033-5: Network security — Part 5: Securing communications across networks using Virtual Private Networks (VPNs)
- ISO/IEC 27033-6: Network security — Part 6: Securing wireless IP network access
- ISO/IEC DIS 27033-7: Network security — Part 7: Guidelines for network virtualization

91 See: https://www.iso.org/search.html?q=27000

security

- ISO/IEC 27035-1: Information security incident management — Part 1: Principles of incident management
- ISO/IEC 27035-2: Information security incident management — Part 2: Guidelines to plan and prepare for incident response
- ISO/IEC 27035-3: Information security incident management — Part 3: Guidelines for ICT incident response operations
- ISO/IEC CD 27035-4: Information security incident management — Part 4: Coordination
- ISO/IEC 27037: Guidelines for identification, collection, acquisition and preservation of digital evidence
- ISO/IEC 27039: Selection, deployment and operations of intrusion detection and prevention systems (IDPS)
- ISO/IEC 27042: Guidelines for the analysis and interpretation of digital evidence
- ISO/IEC 27043: Incident investigation principles and processes

These standards are not freely available; they can only be accessed through paid download.

Security controls

Security controls (or countermeasures) are implemented to reduce risk by eliminating threats or vulnerabilities. For example, threats may be eliminated by installing fences, hiring security guards, and improving access control, and vulnerabilities may be eliminated by applying patches, implementing backups, or improving security policies.

When deciding on which security controls to implement, there are two principles to keep in mind:

1. **The value of a protected asset must be greater than the cost of implementing a security control.**
2. **The security control must make the cost of an attack greater than the value of the protected asset.**

A comprehensive list of security controls is available in NIST Special Publication 800-53, **Security and Privacy Controls for Information Systems and Organizations.**[92]

Security controls can be divided into three major groups – physical, logical (technical), and administrative:

- Physical controls constitute any measures that prevent or deter an attacker from attempting to come into direct contact with assets. These include fences around buildings, bars on windows, physical security, video surveillance cameras, alarms, door locks and padlocks, physical access control systems, and the like.
- Logical or technical controls comprise hardware and software designed to prevent or limit access to information contained in an ICT system or ICT system resources. These include methods for identification, authentication, and authorization (i.e., what I know, what I have, and what I am), network and security devices (routers, firewalls, IDS, IPS, and WAF), software (anti-malware and anti-spyware), virtual private networks (VPNs), file and communication encryption, data loss prevention (DLP) systems, and various security protocols.
- Administrative controls are prescribed by the management of an organization or represent a regulatory obligation to which employees are bound to comply. These include security protocols and procedures applied across the organization, such as the classification and labeling of sensitive data, the control and testing of employees, security checks during employment, efforts to raise the security awareness of employees, and other training and education in this area.

92 Joint Task Force, Security and Privacy Controls for Information Systems and Organizations, SP 800-53, Rev. 5 (NIST, 2020). Available as a PDF at: https://csrc.nist.gov/CSRC/media/Projects/risk-management/800-53%20Downloads/800-53r5/SP_800-53_v5_1-derived-OSCAL.pdf

Security concepts

Cryptography

Cryptography is used to achieve four major goals: confidentiality, integrity, authentication, and non-repudiation.

Confidentiality implies that information is inaccessible to everyone except users authorized to access it. The methods of applied cryptography must preserve confidentiality when information is in use (in the operative memory of a computer system), when it is in transit (in a communication network, as it travels from one point to another), and when it is at rest (in storage, such as a hard drive, USB, etc.).

Integrity means that information cannot be changed without authorization. By applying appropriate cryptographic methods, information is protected from intentional and unintentional changes, partial deletion, or addition.

Authentication proves the identity of a specific person. This is used to prove a user's identity when establishing a communication channel on a network, or to prove the origin of information and files.

Non-repudiation ensures that the sender of a message cannot subsequently deny sending the message.

Cryptography is, in essence, the application of algorithms on a set of data. The process of transforming this data so that it is unreadable is called **encryption,** and the opposite process is called **decryption.** There are three dominant types of algorithms that fulfil the four major goals of cryptography: symmetric key algorithms, asymmetric key algorithms, and hash algorithms.

Symmetric key algorithms

Symmetric key algorithms were used long before computers, but machines capable of performing millions of mathematical operations per second have brought a new approach to the field. Still, the principle has remained the same: both parties in a communication must possess the same algorithm and the same encryption key. The sender first encrypts a file using an algorithm and a key (which is called a secret key), after which the encrypted file is secure for transmission using any available communication channel (including public communication, such as the Internet). When a recipient receives the encrypted file, they can only decrypt it with the same algorithm and identical secret key.

Today, most of the algorithms used for public communication are well-known, which means security lies solely in the secret key. This secret key is a binary file, usually 128 to 256 bits long (but can be shorter or longer). A new secret key can be generated for each occasion, or a single secret key can be used over time for a group of users. These keys can be generated in a key generator, but a problem arises in distributing secret keys to other parties in a communication; long ago, the distribution of secret keys would have been carried out by couriers, but cryptography techniques are now used to solve this problem.

Currently, the most popular symmetric key algorithm is AES (originally Rijndael). Many others are also available, including IDEA, Twofish, Blowfish, RC4, RC5, Safer, and Skipjack.

Asymmetric key algorithms

Unlike symmetric key algorithms, asymmetric key algorithms use two keys, private and public. What is encrypted with a private key can only be decrypted with the corresponding public key, and vice versa. These private and public keys are always generated together using a special algorithm, and corresponding keys are called a key pair.

As the names of these keys suggest, the private key is the arbiter of secrecy, and the owner of the private key must be the only one who accesses it. On the other hand, the public key is intended for public distribution, meaning that anyone can access it. The owner of the private key is also the owner of the corresponding public key, but can transmit the public key by email or upload it

to an Internet repository without security consequences. The system is simple: if I want to send something secret to someone else, I use their public key for encryption, since they are the only person who has the corresponding private key to decrypt the file. If I want to prove that I am who I say I am on the Internet, I encrypt something known to everyone with my private key, which can only be decrypted with my public key, available to everyone.

Key lengths used for asymmetric key algorithms are longer than those used for symmetric key algorithms. For example, the strength of a symmetric key algorithm with a 128-bit key corresponds to the strength of the most commonly used asymmetric key algorithms, RSA and DSA, with a 3072-bit public key. Furthermore, the strength of a symmetric key algorithm with a 256-bit key corresponds to the strength of an asymmetric key algorithm with a 15630-bit public key. Asymmetric key algorithms require much more operative memory and much more processing power than symmetric key algorithms, and they take much longer to encrypt and decrypt. This is why symmetric key algorithms are more suitable for encrypting large amounts of data and are used to achieve confidentiality, while asymmetric key algorithms are used to achieve integrity, authentication, and non-repudiation.

Elliptic Curve Cryptography (ECC) algorithms use different mathematical operations (solving the discrete logarithm problem) than standard asymmetric key algorithms like RSA and DSA (which solve the factorization of a large integer). ECC algorithms use much shorter keys than RSA and DSA (3072-bit RSA or DSA keys are comparable to 256-bit ECC keys, and 15630-bit RSA or DSA keys are comparable to 512-bit ECC keys). Due to their lower need for memory space and processing power, ECC algorithms are widely used in smartphones for applications such as key exchange (ECDH algorithm) or digital signature (ECDSA algorithm). However, ECC algorithms are still much more demanding than symmetric key algorithms.

Hash algorithms

Hash algorithms are closely related to a category of mathematical functions called one-way functions. The main characteristic of this class of functions is that calculating the result from a given initial value is quite easy, while calculating the initial value if the result is known is very difficult or even impossible.

Based on one-way functions, hash algorithms thus calculate representations of files called hash values or message digests, from which it is very difficult or impossible to reconstruct the original file. The length of the hash value for a particular type of hash function is always the same regardless of the length of the file on which the calculation is performed. For example, hash algorithm SHA-1 will produce 160-bit hash values from a single-bit file or from a file of several megabytes. Of course, these two hash values will be different, and in both cases, the original set of data from which the hash value is calculated cannot be reconstructed.

In a good hash algorithm, if a single bit in the original file is changed, the new hash value is drastically different (meaning that approximately 50% of bits in the hash value are changed relative to the hash value of the original file). Moreover, it is infeasible to determine any set of data that will produce a particular hash value. These features are widely used to achieve file integrity. For instance, if the hash value of a particular file (i.e., an installation file, data set, etc.) is published in a secure place on the Internet, accessible by anyone, then someone who wants to use that file can download it, calculate the hash value of the downloaded file before any use, and compare it with the hash value published online. If these two values match, the integrity of the file can be trusted; but if they are different, the file has been altered and should not be used.

Among the most commonly used hash algorithms are SHA-256 and SHA-512 (with hash value lengths of 256 and 512 bits, respectively). Previously, the MD5 hash algorithm with a 128-bit hash value was widely used, but it is no longer considered secure.

Digital signatures

Digital signatures are intended to achieve authentication, integrity, and non-repudiation.

A digital signature is generated by the sender of a message. The sender first calculates the hash value of the message using a hash algorithm, then encrypts that hash value using an asymmetric key algorithm and a private key. After the sender transmits the message and encrypted hash value to the recipient, the recipient decrypts the hash value using the sender's public key (already received from the sender or downloaded from a repository on the Internet). The recipient can then calculate the hash value of the message and compare it to the previously decrypted hash value. If these two values are identical:

- the message has not been altered, either intentionally or unintentionally.
- the message originated from the sender whose public key was used to decrypt the received encrypted hash value; and
- the sender cannot deny sending the message because they are the only one who can access and use their private key.

This requires that the sender and receiver use the same hash algorithm and asymmetric key algorithm.

The NIST publication, Digital Signature Standard (DSS), details all aspects of digital signatures.[93]

Digital envelopes

The digital envelope is intended to achieve all four goals of cryptography. Thus, it is logical that it uses all three types of cryptographic algorithms.

The process of creating a digital envelope is similar to that used to create a digital signature, except in the case of a digital envelope, the message is encrypted before it is sent by using a symmetric key algorithm and a secret key. The secret key is encrypted with an asymmetric key algorithm and the recipient's public key. The sender then transmits a digital envelope consisting of:

- the secret key, encrypted with the recipient's public key, which only the recipient can decrypt using their private key;
- the message, encrypted using a symmetric key algorithm, which can only be decrypted using the associated secret key; and
- the hash value of the message, encrypted with the sender's private key, which can be decrypted using the sender's public key.

Digital certificates

One of the main problems we encounter when exchanging information with another person on the Internet is how we can ensure that they are who they say there are. Even if someone sends their public key via email, or you find it in an online repository, can you really know that someone is who they claim to be? This is where digital certificates are useful. They are issued by authorities that have the resources to verify data provided by a person or entity, and offer digital certification as proof that this data is valid.

The structure of digital certificates is determined by the international standard set out in Recommendation ITU-T X.509 (10/2019).[94] Accordingly, they contain at least the following information:

- the version of the X.509 standard being used;
- the serial number of the certificate;
- the signature algorithm ID;
- the name of the authority that issued the certificate;
- its period of validity;

93 US Department of Commerce, National Institute of Standards and Technology, Digital Signature Standard (DSS), FIPS PUB 186-4, July 2013. Available as a PDF at: https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.186-4.pdf

94 International Telecommunications Union, 'ITU-T Recommendations: ITU-T X.509 (10/2019)', https://www.itu.int/itu-t/recommendations/rec.aspx?rec=X.509

- the name of the person, organization, or machine that owns the digital certificate;
- the public key of the owner of the digital certificate; and
- the digital signature of the certificate itself.

The authority that issues a digital certificate is called a Certification Authority (CA), and the digital signature of a certificate is the hashed body of the certificate (including the public key of the owner) encrypted with the private key of the CA. With that digital signature, the CA guarantees that the attached public key belongs to the owner named in the digital certificate.

Some well-known CAs include Amazon Web Services, GlobalSign, Comodo, and GoDaddy.

Data sanitization

In the lifecycle of data, the last stage is deletion; but depending on the method of deletion used, deleted data may be permanently unrecoverable, partially recoverable, or fully recoverable. Deletion methods such as a simple data wipe, reformatting, factory reset, or unverified file shredding are not secure and may result in subsequent data recovery.

Removing information from storage media in a way that prevents further recovery is called sanitization. Sanitization methods can be divided into three groups:

- Software-based data erasure, which writes ones and zeroes to each part of the storage space selected for sanitization.
- Cryptographic erasure, which encrypts an entire data storage device with a certified encryption algorithm and a unique encryption key, and irreversibly destroys the key immediately after encryption.
- Physical destruction, whereby storage media are physically destroyed, either by shredding the media into small pieces or by degaussing magnetic media through exposure to a very strong magnetic field.

Securely erasing data is a crucial task, as a failure to do so can lead not only to the exposure of data, but also to potential intrusions into a system. Resources that offer useful information and advice on data sanitization include: guidance published by the UK National Cyber Security Centre,[95] guidelines developed by NIST,[96] and a policy manual issued by the NSA.[97]

Access control services

These services provide and control access to certain resources (data or devices), and fall into five categories:

- Identification – a user provides a system with insight into the credentials they possess to access that system.
- Authentication – a system, based on credentials, decides whether a user has the right to access the system.
- Authorization – a system, based on credentials, decides which resources a user has the right to access.
- Auditing – information about user activities while logged into a system is collected in log files.
- Accounting – information collected in log files is reviewed to determine whether a user has acted in accordance with their rights and authorizations.

95 'Secure sanitization of storage media', 23 September 2016, https://www.ncsc.gov.uk/guidance/secure-sanitisation-storage-media

96 Richard Kissel et al., Guidelines for Media Sanitization, Special Publication 800-88, Rev. 1 (NIST, 2014). Available as a PDF at: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf

97 NSA/CSS, 'Storage Device Sanitization and Destruction Manual', Unclassified Policy Manual 9-12, 4 December 2020. Available as a PDF at: https://media.defense.gov/2021/Oct/05/2002867897/-1/-1/0/NSACSS%20PM%209-12%2020201204.PDF

Security protocols

In general, a protocol is a series of steps carried out to process data. A security protocol is a series of steps that relate to providing security services for network communications in untrusted environments.

TLS/SSL

Secure Socket Layer (SSL) is a protocol originally developed by Netscape to provide secure communication between web servers and browser clients. After Microsoft began incorporating SSL into Internet Explorer, other browser vendors followed, and SSL became the standard for encrypting web traffic.

Despite the introduction of the Transport Layer Security (TLS) protocol in 1999, which featured many security enhancements, SSL remained the dominant protocol for encrypting web traffic for years, until a major SSL vulnerability was discovered in 2014. This vulnerability was so serious that many vendors and host providers abandoned the use of SSL in a relatively short time and switched to TLS entirely. Then, in 2015, the Internet Engineering Task Force (IETF) issued a memo declaring SSL 'comprehensively broken' and stipulating its deprecation.[98]

TLS is used to add encryption to the traffic of application-layer protocols such as HTTP or FTP, but it can also be used for VPN applications and some multimedia traffic. The latest version of the TLS protocol is 1.3, described in the IETF's RFC 8446.[99]

IPsec and IKE

Internet Protocol Security (IPsec) ensures the security of communications at the IP layer. This suite of protocols is primarily intended for secure point-to-point communication between two gateways (network-to-network), a host and gateway (a remote user and an enterprise network), or two hosts (end-to-end). IPsec is the most common protocol for VPN applications, but is also used to protect sensitive data in other Internet transmissions.

Internet Key Exchange (IKE) provides a dynamic key exchange for IPsec.

More detail about both IPsec and IKE is available from the IETF in RFC 6071.[100]

Protection of wireless networks

Wireless computer networks allow open authentication and unencrypted communication. This type of connection presents dangers, because of how easy it can be for malicious actors to eavesdrop on users. Open authentication is usually offered by public wireless networks in situations where there is no convenient way to provide users with passwords to access the network. Connecting to wireless networks of this type should be avoided, except for fairly trivial use, such as to check the news while waiting at an airport; and even then, only when more secure options are unavailable. It is always better to connect to the Internet through a mobile operator than through an open wireless network.

Several protocols support authentication and encryption in wireless networks, such as WEP, WPA, and WPA2. The most dominant of these is WPA2, while WEP and WPA are considered particularly insecure.

A brief overview of security devices

Firewalls

A firewall is a network device located at the border between network segments, the purpose of which is to filter traffic according to defined rules. It uses predetermined criteria to block packets, but also built-in logic to recognize certain types of attacks and prevent them. Firewalls

98 IETF, 'Deprecating Secure Sockets Layer Version 3.0', Request for Comments 7568, June 2015, https://www.rfc-editor.org/rfc/rfc7568

99 IETF, 'The Transport Layer Security (TLS) Protocol Version 1.3', Request for Comments 8446, August 2018, https://www.rfc-editor.org/rfc/rfc8446.html

100 IETF, 'IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap', Request for Comments 6071, February 2011, https://www.rfc-editor.org/rfc/rfc6071.html

are common due to their simplicity and applicability, and exist in both hardware and software versions.

WAFs

A Web Application Firewall (WAF) is a device intended to protect a website from specialized attacks aimed at accessing restricted parts of the site, modifying web pages, placing malicious content on the site, etc. This device protects both the content of the website and users who access it.

IPS and IDS

An Intrusion Prevention System (IPS) and Intrusion Detection System (IDS) are devices installed to identify intruders based on predefined models and on information about the previous behaviour of system users. Upon installation in a system, these devices must go through a period of learning in order to recognize the usual behaviour of system users, during which they must be supported by an administrator. The difference between an IPS and IDS is that an IPS is placed in the traffic flow and has the ability to interrupt traffic if it identifies irregular behavior, whereas an IDS is placed parallel to the traffic flow and can only raise an alarm if it detects irregular behaviour.

Hardware security modules (HSMs)

HSMs secure the storage of cryptographic keys. Designed to be tamper-proof, all contents of the device are securely deleted if the device is opened. This means that stored cryptographic keys never leave the device. Some HSMs have built-in cryptographic algorithms and can encrypt and decrypt files.

The Trusted Platform Module (TPM) is an implementation of HSM. A TPM is built into a computer's mainboard and its primary function is to support hard drive encryption. A hard drive encrypted using a TPM on a particular computer cannot be decrypted on another computer or by using a different TPM.

Security platforms

Demilitarized zone (DMZ)

A DMZ is a subnet where servers accessible from the Internet, such as web or email servers, are located. Internet access to these servers is protected by firewalls and other security devices, and a DMZ is also separated from other internal parts of the network.

SIEM

Log data from various devices, not only security devices but also various servers, end-user computers, and network devices, flows into a security incident and event management (SIEM) platform in real time. A SIEM system processes this information and correlates it by time to draw conclusions about incidents or threats to the system.

SOAR

Security Orchestration, Automation, and Response (SOAR) is a collection of various security tools, integrated into one. SOAR provides options for vulnerability management and incident response when security operations are automated.

Virtualization

Virtualization is not essentially a security platform, but can be used for a variety of security purposes. Virtualization facilitates the hosting of several different operating systems on the same computer, all of which can work independently of each other. For this reason, virtualization is often used for testing purposes, as it enables the relatively easy creation of an adequate environment and also the separation of the virtual machine from the host system, in order to disable any possible data exchange between them. In this way, a secure environment can be created in which to test suspicious files.

Virtual machines are often used for training purposes as well, so it is useful for any CSIRT employee to have a hypervisor installed on their laptop. Hypervisors that are free to download include VMware Workstation Player[101] and Oracle VM VirtualBox.[102] Microsoft Hyper-V is also built into some newer versions of Windows as an optional feature that can be enabled.[103]

EDR

Endpoint detection and response (EDR) technology monitors endpoints (computers, laptops, mobile phones, IoT devices, etc.) to detect suspicious activities. Unlike anti-malware solutions, most EDR solutions do not include mechanisms to protect endpoints, only to provide alerts when malicious activity is detected. Data created by EDR can be stored independently, in a dedicated database, or can be sent to a SIEM platform.

Cybersecurity company Comodo (now known as Xcitium) has developed an open-source EDR solution, OpenEDR, which includes real-time monitoring of endpoint file systems, the detection of fileless threats, the ability to investigate threat vectors, and the option to add custom rules.[104]

Security software

Anti-malware

Anti-malware software is used to detect, and then remove or disable, a variety of malware. It relies on signature based operation, comparing data from a memory source (operational or mass storage) to a database; and if a match is found, a given action is performed (for example, malware is quarantined or deleted). Anti-malware software also checks for behaviour patterns that are typical of malware in a system – a process known as a heuristic or behavioural analysis – and if such patterns are identified, an alarm is triggered and assigned actions are performed.

Because so much new malware, and variants and sub-variants, appear every day, it is essential to regularly update anti-malware software databases.

Anti-spyware

Spyware is a specialized malicious programme that, once activated on a victim's computer, collects information about that user's behaviour and activities and sends that information to an attacker – who controls the spyware. Spyware can collect information about everything from the websites we visit to the passwords we enter. Anti-spyware software recognizes programmes of this type and disables them.

Operating system security

For the purpose of security, operating systems are organized into rings of protection. Outer rings must communicate with inner rings in order to access restricted resources. Most of today's operating systems use four such rings, the innermost of which is called the kernel. This represents the part of the operating system that resides in computer memory, has the highest level of privilege, and can access every memory location and every file. The rest of the operating system is stored in the second ring, known as ring 1; device drivers and some system utilities are stored in the third ring, ring 2, with the privilege to access peripheral devices; and applications and other programs are stored in the fourth ring, ring 3.

Data loss prevention (DLP)

Data loss prevention is a set of measures aimed at detecting and preventing the unauthorized use, manipulation, or transfer of data. These include:

- data identification and classification,
- user authentication and authorization,

101 See: https://www.vmware.com/products/workstation-player.html

102 See: https://www.oracle.com/virtualization/technologies/vm/downloads/virtualbox-downloads.html

103 For several ways to enable Hyper-V, see: https://docs.microsoft.com/en-us/virtualization/hyper-v-on-windows/quick-start/enable-hyper-v

104 See: https://openedr.com/

- data encryption,
- a registry of user access to sensitive data,
- traffic monitoring, and
- endpoint monitoring.

DLP measures defend against both data loss (i.e., data is no longer available) and data leakage (i.e., an unauthorized person has come into possession of data), both of which typically result from:

- a malicious insider,
- an external breach, or
- the negligence of authorized users.

Because data classification largely depends on human discernment, and is therefore prone to possible error, many DLP solutions incorporate machine learning algorithms to detect abnormal data access and manipulation.

Secure communications

VPNs

A virtual private network (VPN) is an environment in which remote users or remote networks can communicate as if they were part of the same network. VPNs can be established between two remote users, a remote user and an enterprise network, or two distant networks. VPN applications are based on tunnelling protocols, which encapsulate protocol packets within packets of another protocol. Most VPNs use encryption to protect encapsulated data.

PGP

Pretty Good Privacy (PGP) is a secure communication system that uses both symmetric and asymmetric cryptography. Despite having been introduced in 1991, it remains secure and is widely used. And even though the author of PGP sold the product, some versions of PGP-compliant software (OpenPGP) are available for free under the GNU General Public License.[105]

RFC 4880, issued by the Network Working Group in 2007, offers more details about OpenPGP.[106] But briefly, the secure transmission of a file to another person using OpenPGP entails the following sequence: 1) a key is generated on the sender's system, 2) the file is encrypted using symmetric cryptography and the key that was generated, 3) that key is encrypted using asymmetric cryptography and the recipient's public key, and 4) the encrypted file and the encrypted symmetric key are sent to the recipient. A sender can also elect to sign the transmitted file with their private key. In this way, OpenPGP achieves the principles of confidentiality, integrity, authentication, and non-repudiation.

Beyond the sender's key pair and the file to be sent, the only things a sender needs in order to use OpenPGP are the recipient's public key and email address. The CSIRT community shares public keys in several public key repositories on the Internet, but it is also wise to exchange public keys with other professionals in the field while attending events and trainings, which allows you to ensure that a particular public key belongs to a particular person.

105 See: https://gnupg.org/

106 Network Working Group, 'OpenPGP Message Format', November 2007, https://www.ietf.org/rfc/rfc4880.txt

Incident handling

Incident handling involves a number of phases:



Depending on the model, some of these are grouped together. For example, NIST recognizes four phases of incident handling, combining detection and analysis into a second phase, and grouping containment, eradication, and recovery into a third.[107] Meanwhile, the SANS Institute envisions six phases, as it combines detection and analysis into a second 'identification' phase.[108] Regardless of the number of phases stipulated by a model, they all call for CSIRTs to carry out the same activities in the course of incident handling.

Preparation

To effectively resolve an incident, especially a significant incident, CSIRT members must be well prepared and coordinated. This starts with the adoption of an incident handling plan, which must contain details on what to do, who should do it, who needs to be informed and when, and how to communicate. This plan should be co-signed by the head of the organization the CSIRT serves or another person of authority, and it should be explained to constituents so that they understand their role and obligations. Importantly, the plan must be tested, so that its functionality is verified before it is deployed to address a real incident.

Monitoring

Monitoring is a continuous activity that provides real-time information about security incidents like network attacks, attempts at unauthorized access, or technical failures. Monitoring involves the collection and analysis of large amounts of data, to detect known attack patterns or unusual behaviours and communications. Continuous monitoring of this kind helps prevent breaches, improve response time, and detect vulnerabilities. The objects that are monitored include web

107 See: Cichonski et al., Computer Security Incident Handling Guide, p. 21.

108 Patrick Kral, 'Indident Handler's Handbook', SANS white paper, 5 December 2011. Available at: https://sansorg.egnyte.com/dl/6Btqoa63at

servers and applications, email servers, network devices, database servers, endpoint devices, and cloud storage.

Since the volume of data to be analysed is typically quite significant, this analysis must be carried out using specialized software or systems, such as SIEM platforms.

Snort

Snort is an open-source application capable of capturing and analysing network traffic in real time.[109] It can be configured to operate as an IDS based on rules defined by the administrator, but also to actively influence network traffic, making it an IPS. In addition, Snort can be configured to run in both sniffer mode and packet logger mode. Sourcefire, the company that created Snort, was purchased by Cisco in 2013, but the tool is still free to use under the GNU General Public License. Cisco Talos develops, tests, and approves a subscriber ruleset for Snort (for which users must pay), but there is also a community ruleset available, free to all, created by the Snort community and checked by Cisco Talos.

Suricata

Suricata, a project of the Open Information Security Foundation (OISF), was launched in 2009.[110] Fundamentally, Suricata is an open-source IDS and IPS solution, but it offers other network security monitoring capabilities as well. Moreover, it can be easily integrated with Logstash, Splunk, Kibana, and other open-source tools.

The Elastic Stack

One of the most popular platforms for log management is known as the Elastic Stack,[111] and consists of four related tools:

- ElasticSearch, a search and analytics engine;
- Logstash, an engine for log collection from multiple sources;
- Kibana, a data visualization engine used for results presentation; and
- Beats, an agent installed on servers and used to send operational data to ElasticSearch.

These tools are all open and free to use.[112] Their source code is also available on Github.[113]

Wazuh

Wazuh is an open-source SIEM platform with the capability to collect, aggregate, index, and analyse security data.[114] One of its core components, Wazuh agents, are installed on endpoints and act as host-based IDS solutions. Another core component, the Wazuh server, receives data from these agents, subjects it to threat intelligence analysis, and generates alerts. Two more components constitute the Wazuh platform: the Wazuh indexer, which indexes and stores alerts generated by the Wazuh server; and the Wazuh dashboard, a user interface for data visualization, analysis, and management.

Open Source Security Information Management (OSSIM)

OSSIM is an open-source SIEM platform created by AlienVault (which was acquired by AT&T in 2019 and became AT&T Cybersecurity).[115] It integrates several well-known open-source tools, as well as a correlation engine and other tools developed by AlienVault, with unified access under a web-based user interface. Some of the third-party tools integrated into OSSIM are:

- PRADS, for passive monitoring;
- Fprobe, for generating NetFlow data;
- NFSen/NFDump, for collecting and analyzing NetFlow information;

109 See: https://www.snort.org/

110 See: https://suricata.io/

111 This collection was previously referred to as the ELK Stack – an acronym for the names of the first three tools, before Beats was added.

112 See: https://www.elastic.co/downloads/

113 See: https://github.com/elastic

114 See: https://wazuh.com/install/

115 See: https://cybersecurity.att.com/products/ossim

- Nagios, for system monitoring;
- Suricata, for intrusion detection; and
- OpenVAS, for vulnerability assessment.

OSSIM is freely distributed under the GNU General Public License.[116] There is also a commercial version of the platform, known as USM Anywhere.[117]

Communications planning

One of the most important things to prepare for in advance of an incident is communication. Every team member must know who to inform about an incident, and by what means of communication. A secondary contact point must also be identified, in case a primary contact is unavailable. Additional contact information must be available to CSIRT members, including phone numbers, email addresses, encryption keys, authentication methods, etc.

The contact information of more experienced colleagues is priceless in situations when you are facing an ongoing incident alone. Be sure you have all the information you need before you find yourself in such a situation.

Resources planning

Incident handling also requires some advance preparation of certain resources:

- Laptops should be loaded with incident analysis tools.
- Access should be granted to all relevant logs.
- There must be sufficient storage capacity to collect all relevant data.
- Clean, removable disks and media must be available.
- Network diagrams and other documentation about your system must be accessible.
- Spare computer and network equipment (or virtualized equivalents) should be on hand so that actions which must be performed separately from the system (such as the activation of malware) can be performed.

You must know where to find and how to access these resources, for when an incident occurs, time itself is a critical resource that cannot be wasted searching for something that should be prepared and readily available.

User awareness

In many cases, attackers gain access to a system by tricking a user (often an employee) through social engineering methods. The easiest and most common way they achieve this is by sending phishing emails, but if the reward is attractive enough, attackers will also deploy other approaches – such as a tailgate, honey trap, quid pro quo, or an even more complex social engineering attack.

Educating users in your organization about how to recognize and avoid social engineering attacks will prevent many malicious attempts from succeeding. However, over time, even employees who have received training in this area will relax their habits, increasing the chances that they will become victims of social engineering attacks. Thus, many CSIRTs engage in a practice of continuous awareness raising.

Detection

The resolution of an incident begins the moment information about the incident is received. This information may be obtained through an incident report, or a security device may detect malicious activity and trigger an alarm, or an external actor may report that they have detected unusual activity originating from your system. No matter the source of information about an incident, measures should be taken from the moment it is received to suppress malicious activity, return the

116 See: https://cybersecurity.att.com/products/ossim/download
117 See: https://cybersecurity.att.com/products/usm-anywhere

system to its state before the incident, and remove weaknesses in the system so that the same incident does not reoccur.

Some signs of a potential incident are visible **before** it occurs and are called precursors. For example, there may be traces of a reconnaissance or vulnerability scan in your log files, information may have been published about a vulnerability in your system, or a patch may have been issued for software used by your organization and has not yet been applied.

Receipt of incident reports

There is usually a single point of contact in a CSIRT for the reporting of incidents, by phone, SMS, email, web form, or any other means. In larger CSIRTs, there may be an entire department dedicated to receiving this information, but it is more common for an on-call employee to be assigned this duty for a set period (e.g., a week or month). Even where the receipt of incident reports is automated, there must be a CSIRT employee assigned to verify these reports.

If incidents are to be reported at all, constituents must know how to report them. Often, this information is found in a publicly released RFC 2350 template, along with other information about a CSIRT. To avoid the possibility that the single point of contact for incident reporting may be bypassed, it is common for CSIRTs to publish contact data only for incident reporting contacts and not individual team members.

When reporting an incident, it should be mandatory (and therefore must be made possible) for reporting person(s) to leave their own contact data so that they can receive a response or be contacted by CSIRT employees if needed.

Triage

The CSIRT employee in charge of initiating incident report processing must review a report in a short amount of time and begin the procedure defined for the case at hand. This may involve the employee performing administrative actions, such as filing the report, prioritizing or rejecting the report, advising the reporting person, etc., or may simply forward the report to another CSIRT employee. This decision-making is known as triage.

A set of all the information and documents related to an incident is called a ticket. The receipt and triage of each incident report must be recorded, including the precise date and time a report was received, any response, and details about further processing. In situations where a cyber attack causes considerable damage, all elements of the incident response are reviewed, and it is very important to have precise data on all the measures taken by all actors.

Mistakes can occur during triage, such as:

- opening several different tickets upon receipt of several reports for the same incident;
- joining a new report to an already opened ticket, when in fact it concerns a different incident;
- neglecting or overemphasizing certain aspects of the incident, leading to poor decision-making regarding prioritization;
- rejecting an incident report due to misinterpreted information;
- forwarding an incident report to the wrong CSIRT employee; and
- offering bad advice to the person who reported the incident.

Sometimes, these mistakes also result from errors on the part of those who report an incident, for example if they report a new incident but use an old incident number in their report. In that case, a CSIRT employee may add new information to an old ticket during triage, combining two completely different events.

While CSIRT employees should of course be conscientious, mistakes do happen and are in fact an integral part of the job. What is important is that, when we recognize our mistake, we do our best to correct it as soon as possible, or inform a superior if we are unable to do so.

Prioritization

Not every incident is of equal importance, and incidents of a higher priority must be resolved before lower priority incidents. Different CSIRTs use different incident prioritization models that account for their constituency, the potential for damage, the source of an attack, the type of incident, and more. In general, if your CSIRT faces multiple active cyber attacks at the same time, priority will always be given to incidents that may cause a loss of lives or other severe consequences to humans (if this is a possible outcome at all). If several cyber attacks of the same priority occur at once, the principle of 'first come, first served' is usually applied.

It is not uncommon for the priority of an incident to change in the midst of incident resolution. The priority of an incident may be raised if there is an escalation of malware activity, unexpected damage caused by an attack, new information about the source or target of the attack, newly discovered vulnerabilities in the system under attack, or merely a call from senior management requesting that a higher priority be assigned to the incident. While incident de-prioritization during incident resolution is rare, it may occur if new information is received that reflects positively on the security of the system or if a mistake was made during the initial prioritization of the incident.

Incident numbers

Each incident report must be allocated a unique number. Often, CSIRTs use automated incident reporting systems that assign this number and send a first response to the reporting party, and every subsequent communication with the reporting party and constituents regarding the incident refers to this number. Incident reporting systems are also known as ticketing systems (covered in more detail below).

Though we refer to 'the number' of an incident, it is not necessarily composed solely of numerals and can consist of any combination of characters (numbers, letters, and special characters). Also, though incident numbers are unique within a CSIRT, different teams may use the same incident number creation scheme, which means it is possible that two CSIRTs could create the same incident numbers. Of course, these would represent different incidents at different times, as it is practically impossible that two CSIRTs could create the same incident number at precisely the same time.

When defining a scheme for the creation of incident numbers, it is important that no elements of the incident itself (date and time, reporting party, sorting elements, etc.) are revealed in the number, which is often used in correspondence. Many CSIRTs combine a unique identifier (e.g., their abbreviated name followed by a # sign) with an integer that represents each incident; but many teams do not assign these integers in sequence, instead using a non-repeating number generator to assign a random value.

Ticketing system

The purpose of a ticketing system is to collect all the data about an incident in one place, implement procedures more efficiently, determine whether a new incident is related to any previous or ongoing incidents, merge two or more separate entries from the same incident, make it easier for a CSIRT employee to continue solving an incident started by another employee, prevent an incident from being 'forgotten', and so on.

Each incident gets its own 'ticket' and each ticket has its own unique ticket (incident) number, covered above. Constituents are asked to always use this ticket number as a reference in their communications with CSIRT, to facilitate the work of CSIRT employees and save time. If a portal, website, or email system is used for communication, a ticket number can be used to identify information related to an incident and the CSIRT employee handling that incident can be notified of new information from the ticketing system.

In addition to a unique ticket number, each ticket must contain general information as required by the procedures of a given CSIRT, as well as any information collected from the incident, such as:

- its category (the type of malicious activity that caused the incident);
- its classification (the confidentiality of information);
- its priority (whether the incident should be resolved before other incidents);
- the name of the team member responsible for managing the incident;
- contact information for all involved parties (users who reported the incident, contacts in compromised systems, contacts in other CSIRTs who have assisted or offered assistance, etc.);
- the names and locations of files associated with the incident (remnant files, log files, backups, etc.);
- links to locations related to the incident on the local network or on the Internet (a local repository with the appropriate patch, a website with information about the malware, etc.);
- a communications history (all communications related to the incident, with precise dates and times);
- all actions taken (with precise dates and times, and a description of observed impacts);
- its current status (the phase of incident handling in progress, or when the ticket was closed); and
- any other information related to the incident (relevant legal obligations, notes on cooperation with other CSIRTs, problems observed in existing procedures, etc.).

In the past, many CSIRTs used the Request Tracker for Incident Response (RTIR) ticketing system.[118] Today, more modern ticketing systems tend to be popular.

The Hive

Among the most popular of these modern ticketing systems is The Hive.[119] This security incident response platform has many capabilities beyond its ticketing system. It offers opportunities to collaborate with other CSIRTs on simultaneous investigations, integration with Cortex and MISP (including rapid triage of imported information and export of IOCs to MISP instances), and more.

Indicators of compromise (IOCs)

IOCs are artifacts present on a system or network that indicate a potential intrusion. They can be found in logs or files, and their presence implies a potential or ongoing attack. Examples include:

- unusual inbound or outbound network traffic;
- communications with systems in unusual geographic areas;
- unusual DNS requests;
- unusual ports used;
- unusual size of HTML responses;
- non-human behaviour in communications;
- increased number of unsuccessful logins;
- unusual activities on the part of privileged user accounts;
- suspicious changes to registries, system settings, or system files;
- unexpected software updates;
- unknown files and processes in the system;
- increased access to databases; and
- unusual requests and read volume for some files.

Sometimes, to hide IOCs, attackers apply Base64 or similar encoding (or even repeated Base64 encoding) to malicious scripts.[120]

118 See: https://open-source-security-software.net/project/rtir

119 https://thehive-project.org/

120 See: Network Working Group, 'The Base16, Base32, and Base64 Data Encodings', Request for Comments 4648, October 2006, https://www.rfc-editor.org/rfc/rfc4648

STIX and TAXII

Structured Threat Information Expression (STIX) is a language and format for exchanging cyber threat intelligence.[121] It is used by CSIRTs for collaborative threat analysis, automated threat exchange, and automated detection and response.

Trusted Automated eXchange of Indicator Information (TAXII) is an application layer protocol used to exchange cyber threat intelligence over HTTPS.[122] TAXII fully supports the exchange of CTI in STIX format, but may also support other formats.

STIX and TAXII are both open standards developed and maintained by OASIS Open, a non-profit standards body.[123] The development of these standards, which together provide a means of automated expression of CTI, was community-driven.

Indicators of attack (IOAs)

While IOCs relate to how an attack is carried out, IOAs relate to an attacker's intentions and objectives. IOCs can be unreliable and may produce many false alarms, but IOAs are more dependable and can help defenders predict an attacker's next move.

IOAs include:

- communication of internal hosts with known malicious destinations;
- suspicious communication of internal hosts with external public servers;
- communication of internal hosts with DMZ hosts;
- connections that use non-standard ports, especially for the HTTP and HTTPS protocols;
- excessive SMTP traffic;
- excessive DNS requests from internal hosts;
- suspicious use of RDP protocol;
- suspicious use of SSH protocol;
- inter-host communications within a short time period;
- malware reinfection a short time after removal;
- logins from different geographical locations to the same user account;
- IPS or IDS alerts outside normal working hours; and
- network scans by internal hosts.

Honeypots

The role of honeypots is to serve as bait for attackers. A honeypot can be implemented as information, a database, a server, a system, or a service, but without any real functionality for the organization and lacking access to any sensitive information. Isolated from the production part of a system, a honeypot only pretends to be a functional part of that system in order to lure adversaries into attempting malicious activities. All honeypot activity is monitored and analysed to ascertain and understand the methods and intentions of attackers.

A piece of information, such as a file or database, used as a honeypot is called a **honeytoken**. Since the information has no use to the organization, any access to it must be intentional and is likely malicious.

Sometimes, several honeypots are organized into a separate network, imitating the internal network of an organization. This is called a honeynet and can be particularly attractive to attackers.

Detailed information on honeypots can be found in the ENISA publication, **Proactive Detection of Security Incidents: Honeypots.**[124]

121 See: https://oasis-open.github.io/cti-documentation/stix/intro

122 See: https://oasis-open.github.io/cti-documentation/taxii/intro.html

123 See: https://www.oasis-open.org/

124 CERT Polska et al., Proactive Detection of Security Incidents: Honeypots (ENISA, 2012). See: https://www.enisa.europa.eu/publications/proactive-detection-of-security-incidents-II-honeypots

Analysis

In the analysis phase, traces left by an attacker (e.g., remnant files or traces in log files) are examined, the behaviour of found malware is analysed (e.g., using sandboxes), and operating systems and other software environments on computers and other devices in the system are scrutinized. The nature of the analysis undertaken in this phase hinges on many factors, not the least of which is the ability and knowledge of CSIRT members, but also the severity of the incident, the number of incidents occurring simultaneously, the time and resources available, whether the method of attack is already known and documented or represents something new, etc.

Analysing log files

Log files (or simply logs) store information about events in a system. Today, almost every system creates some kind of log file. While the information recorded varies based on the system, in every case, the date and time of events are logged. Date and time data can be crucial to understanding the flow and elements of an incident, so it is very important that the date and time on a system is always accurate.

Depending on the type and configuration, a system may record millions of new lines in a log file every day. Most of these lines represent regular and benign system activity, but some are associated with harmful activities, and these are what CSIRT analysts are searching for. Analysts, who use filters to detect attacker activities, typically become more efficient with training and experience.

The review of log files can be automatic, in real time, using the tools built into systems that generate log files or using special devices or software. An alarm is then activated if any data in the log file exceeds a certain threshold. Still, it always remains a task for CSIRT experts to review any log file in question to determine the facts related to an incident.

It is sometimes necessary to send a log file to another organizational unit or even another CSIRT, if an incident concerns them. The rule in such cases is that only relevant parts of the log file are sent, never the entire log file. Log files contain information that may be sensitive, so it is necessary to apply appropriate protection measures whenever they are stored or transferred. Furthermore, log files are text based, which means that somebody could simply create one or edit one that already exists; which is why log files included as evidence in court must be verified to ensure they haven't been changed.

Wireshark

In the CSIRT community, Wireshark is a very popular network protocol analysis tool that can be used to analyse real-time traffic directly from various platforms (Ethernet, WiFi, FDDI, and many others), as well as stored log files of different formats in offline mode.[125] Wireshark has built-in decryption support for the most important protocols, and an intuitive graphical user interface.

Analysing malware

Malware analysis refers to activities carried out with the aim to understand how malware is executed and what effect it has on a system. There are two basic types of malware analysis, static and dynamic. In static malware analysis, malware is studied without execution; meaning that malicious files are searched for packets or embedded objects, PE headers are checked, etc., to make assumptions about potential malware functionality. In dynamic malware analysis, malware is executed in a controlled environment and its behaviour is observed.

Static analysis can also involve **reverse engineering**, during which malicious code is disassembled and investigated in detail. This is a comprehensive analysis, involving an examination of the part of a malware algorithm that would not be performed during a dynamic analysis, but can be very time consuming and requires considerable skill. The fundamentals of reverse engineering are described in the 'Malware Reverse Engineering Handbook', a publication of the NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE).[126]

125 See: https://www.wireshark.org/

126 Ahmet Balci, Dan Ungureanu, and Jaromír Vondruška, 'Malware Reverse Engineering Handbook', NATO CCDCOE, 2020. Available as a PDF at: https://ccdcoe.org/uploads/2020/07/Malware_Reverse_Engineering_Handbook.pdf

Notably, reverse engineering is also used by malicious actors, who reverse engineer legitimate applications in an attempt to find vulnerabilities.

In advanced dynamic analysis, malware is run in a **debugger** that can record its execution path. Whereas basic dynamic analysis includes an examination of network logs in order to discover the addresses of malicious entities on the Internet with which malware communicates, one of the goals of advanced dynamic analysis is to discover the pattern and content of inbound and outbound communications of malware, including by cracking encryption schemes. Advanced dynamic analysis takes more time than other types of analysis and requires an analyst with considerable knowledge and experience.

A sandbox is an isolated environment designed for the automated analysis of potential malware. A sandbox mimics a real system, and when malware is executed in it, every activity of that malware is recorded (e.g., which IP addresses it tries to connect to, which processes are created, how it spreads, etc.). Because a sandbox is usually implemented in a virtual environment, some malware is programmed to check whether it is executing on a physical machine and only then proceed further. This kind of malware simply puts a stop to all activities if it recognizes its execution in a virtual environment, so mechanisms are incorporated into sandboxes to make them appear to be physical machines.

Many sandboxes and other analysis services are offered online for free, requiring only that a user upload a malware sample. This can be useful, especially for inexperienced CSIRT employees, but keep in mind that any uploaded sample is known and available to the community from the moment it is uploaded, which may notify an attacker that you are aware of their attack.

Artifacts are files created on a victim's system during an attack. When a system is examined in the incident handling process, any suspicious files should be checked and treated as malicious until proven otherwise. While not every suspicious file can be attributed to an attack (they may be legitimate files created by an application), nor are all the files left by an attacker necessarily malicious, it is vital that all the files created during an attack are reviewed in order to understand how the attack was implemented, what weaknesses were exploited, and how it can be countered. Some artifacts have a very short lifespan, as attackers want to hide as much of their activity as possible; these files can only be captured using specialized tools.

IDA Pro

Interactive DisAssembler Pro (IDA Pro) is widely used for reverse engineering purposes, as a tool for translating executable (machine) code into assembly code, and for debugging. The first versions of IDA Pro were distributed as shareware, but the rights were later sold and it is now a commercial application. A free version (IDA Free) remains available, however, with limited but useful features.[127]

REMnux

REMnux is a collection of free tools designed for different aspects of malware analysis.[128] It includes tools for static and dynamic analysis, memory forensics, document analysis, and more, and can be downloaded as a virtual machine or run as a container.

x64dbg

x64dbg is an open-source tool for Windows, distributed under the GPL version 3 license, that can be used to debug both x64 and x32 applications.[129] It uses open-source libraries such as Qt, TitanEngine, Zydis, Yara, and Scylla.

127 See: https://hex-rays.com/ida-free/#download
128 See: https://remnux.org/
129 See: https://x64dbg.com/

Ghidra

Ghidra is an open-source and free-to-use reverse engineering tool developed by the NSA.[130] It has the functionality of traditional disassemblers, debugging features, and also decompiles assembly code back into source code (which is usually similar to the original code), making the job of reverse engineering much easier.

Cuckoo sandbox

Cuckoo is an open-source automated malware analysis system, free for download and use under GNU General Public License v3.[131] It can analyse many types of files, including executables and documents in Microsoft Office or Acrobat pdf formats. Cuckoo also analyses network traffic and memory, traces Windows API calls, saves files dropped by malware, and more.

Digital forensics

The focus of digital forensics is on 'identifying, acquiring, processing, analysing, and reporting on' electronically stored data.[132] It is an indispensable part of criminal investigations, but can be used in almost any other type of investigation, too. The field was once called computer forensics, but today, computer forensics is just one among many digital disciplines (alongside network forensics, mobile device forensics, and database forensics, among others).

Best practices and approaches to digital forensics are laid out in the Interpol's Guidelines for Digital Forensics First Responders.[133]

TSK and Autopsy

The Sleuth Kit (TSK) is a collection of command-line tools and C libraries that facilitate disk image analysis.[134] TSK is used in digital forensics to recover deleted files, search for files that meet certain criteria or are hidden, get information from slack space, etc.

Autopsy is a digital forensic platform that offers options for timeline analysis, keyword search, extraction of web activities from common browsers, registry analysis, and file type detection.[135] The most exploited feature of Autopsy is probably its graphical user interface (GUI), which is used for TSK, but also for other digital forensics tools.

TSK and Autopsy are built into other bigger platforms, such as Kali Linux or CAINE. Both tools are open source and free to use.

Freetool

The Freetool Project, launched in 2012 to coordinate development and enable sharing of appropriate and reliable digital forensics tools, offers a set of free digital forensics tools developed by law enforcement officers.[136] These tools are distributed only to government or public service bodies that have the appropriate legal mandate, such as police, national CSIRTs, or military actors.

CAINE

The Computer Aided Investigative Environment (CAINE) is a forensic environment that integrates several digital forensic tools and automates interactions between them.[137] CAINE is free to use under the GNU Lesser General Public License.

130 See: https://ghidra-sre.org/

131 See: https://cuckoosandbox.org/

132 INTERPOL, 'Digital forensics', https://www.interpol.int/How-we-work/Innovation/Digital-forensics

133 INTERPOL, Guidelines for Digital Forensics First Responders: Best practices for search and seizure of electronic and digital devices (INTERPOL ICDFL & Norwegian Ministry of Foreign Affairs, 2021). Available as a PDF at: https://www.interpol.int/content/download/16243/file/Guidelines%20to%20Digital%20Forensics%20First%20Responders_V7.pdf

134 See: http://www.sleuthkit.org/sleuthkit/

135 See: http://www.sleuthkit.org/autopsy/

136 See: https://freetool.ucd.ie/

137 See: https://www.caine-live.net/

An intruder's signature

Typically, malicious actors use the same matrix, or a similar one, to break into a system and progress further after penetration. The set of traces left behind (including scripts, names of the files, exploits, etc.) is referred to as an intruder's signature. If you find different sets of traces when analysing the same incident, it could mean that the intruder is being creative, that another incident is happening at the same time, or that there are multiple intruders sharing information about your system.

After solving an incident and collecting information about the intruder's signature, you can check earlier incidents and compare the traces to understand if that intruder has made any such attempts in the past.

Containment

After the initial analysis of an attack, measures must be taken to limit or completely prevent the further spread of malware or the possibility that the attacker, using knowledge acquired about the system, threatens other (still uncompromised) parts of that system. When determining what measures to apply, care must be taken not to unnecessarily jeopardize the functioning of the entire organization, nor to further endanger an already compromised part of the system. For example, disconnecting a compromised host from the network may trigger malware to wipe the entire contents of the hard drive of that host. Deciding on the most appropriate measures is not easy, as this depends on the type of attack and a number of other circumstances, but CSIRTs usually have procedures prepared in advance to deal with various scenarios.

Decisions about which measures to apply must take into account:

- the potential destructive effect of the attack (i.e., the scope of actions of the malware in question);
- potential damage to the organization (financial, reputational, loss of services, theft of sensitive and private information, etc.);
- the resources needed to implement the measures;
- the time needed to implement the measures;
- implications to the organization;
- implications to other parties; and
- how long the measures should be in force.

Some measures that could be implemented include:

- shutting down or disconnecting compromised parts of a system;
- changing administrator and user passwords;
- disabling the accounts of some users;
- blocking certain external IP addresses;
- blocking the organization's connection to the Internet;
- blocking specific services and ports; and
- blocking incoming network traffic.

Cortex

Cortex is a free, open-source software created by The Hive Project for the purpose of analysing collected observables in a single tool, which supports the process of analysis and containment during an incident.[138] Cortex makes it possible to analyse files, hashes, IP addresses, email addresses, domain names, or URLs through a selection of analysers in a web interface. These operations can even be automated, and Cortex can process large sets of information from The Hive, MISP, or custom scripts.

138 See: https://github.com/TheHive-Project/CortexDocs

Identifying an attacker

The primary objective of incident handling is to minimize the effects of the incident and bring the system back to its pre-incident state in the shortest possible time. Identifying an attacker is not a high-priority task, especially because it takes valuable time and offers no guarantee of success. But accomplishing this task could lead to fewer problems in the future.

Some methods to identify an attacker include:

- searching log files to determine the IP address or URL from which the attack originated;
- locating the geographical area of the attacker's IP address or URL;
- verifying that the attacker's IP address or URL is not spoofed by checking its connectivity;
- searching blacklists for that IP address or URL;
- gathering information about this IP address or URL through a browser's search engine; and
- monitoring possible communication channels of the attacker.

It is important to note that the attacker and the owner of an involved IP address or URL may not be the same person if the attacker used someone else's resources to carry out the attack.

VirusTotal

VirusTotal is a free online service that checks files and URLs/domains for signs of malware.[139] Upon uploading a file or copying a URL to VirusTotal, the tool communicates with over 70 online antivirus scanners and blocklisting services to return a report. VirusTotal is used primarily through its public web interface, but also offers desktop uploaders, browser extensions, and programmatic API.

Eradication

The next phase of incident handling is the removal of all existing malware from a system. Ordinarily, this is done with the help of anti-malware software, and care should be taken to apply the latest versions of this software with up-to-date databases. All security software and devices should also be rechecked, their versions and databases updated, and routines run to test the entire system. Additionally, any malicious files that are not automatically removed by the anti-malware software must be removed manually.

It is during this phase that any exploited vulnerabilities should be identified, as well as any hosts that need remediation. The mitigation of vulnerabilities should be completed as soon as possible to prevent future incidents of the same type.

Recovery

After all traces of malware have been removed from a system, it should be restored to its pre-incident state. Depending on the damage to a system, this may entail restoring data from backups, reinstalling operating systems, reinstalling application software, restoring disabled accounts, reconnecting disconnected subnets, establishing a connection to the Internet, installing patches, changing passwords, tightening the firewall ruleset, and the like. In some cases, damage caused by an attack is so extensive that the recovery phase stretches over months. This requires that priorities are set to quickly identify what is most critical for recovery, to prevent similar incidents.

Recovery is usually performed by system administrators and technicians, not by CSIRT employees.

139 See: https://www.virustotal.com/gui/home/upload

Closing a ticket

Every incident comes to an end. Generally, this means the CSIRT is no longer receiving new information related to any phase of the incident that can be added to the ticket, nor is it generating its own new information about the incident (that is, CSIRT employees have taken all the measures available to them and have no plans to take additional measures). A CSIRT may also end its activities in relation to an incident even when information about the incident is still incoming, when there are no further possible measures to be applied by the CSIRT. The CSIRT should notify all relevant parties of the completion of an incident, or it should inform them that it will consider the incident closed if no new information is received.

The end of an incident may be marked by closing the ticket, which is done by the CSIRT employee who handled the incident, or sometimes by a supervisor. The criteria and manner in which the end of an incident is marked differs among CSIRTs.

If necessary, a closed ticket can be re-opened and the resolution of an incident can continue. This may be done, for example, when significant new information about the incident emerges requiring the application of new measures or when the system is not thoroughly cleaned and the malware is reactivated.

Post-event activities

The final phase in incident handling is a look back, at the cause of the incident, the entire course of the response, and at the failures or vulnerabilities that made it possible for the incident to occur. This is the time to solve those problems so that a similar incident does not occur in the future. It is also necessary to review the procedural steps that were taken, to assess whether there are needs and opportunities to make changes to procedure, and if so, to implement them.

This final, reflective phase is often neglected due to a sense that, once an incident has ended, the work is done. However, one method to facilitate such analysis is a **lessons learned** meeting, which is usually organized by CSIRT leadership but should be attended by any employees of the organization who participated in resolution of the incident, not just CSIRT employees. Indeed, perhaps even some employees who did not participate in resolving the incident but could be useful during future incidents should attend. A meeting focused on lessons learning should be held within a few days after the closure of an incident ticket and should cover:

- what happened,
- when it happened,
- which vulnerabilities made it possible,
- whether response procedures were adequate,
- whether the information received was sufficient,
- whether communication within the organization was adequate during the incident, and
- what should be changed.

Reports from meetings such as this can be very useful to more efficiently resolving similar incidents in the future, but can also be used as training materials for junior employees in CSIRTs. Ask your senior colleagues or CSIRT leadership for these reports.

Organizing incident resolution

Incident communication and coordination

Coordinating cyber incident activities is a natural function of a CSIRT. For this to be possible, any actors working to supress an incident must be familiar with the responsibilities of the CSIRT and have good lines of communication with the team. This communication needs to be established before any incident occurs, as does a good relationship between units working to solve an incident, as this can be crucial to their ultimate effectiveness.

Also, it is important to have one CSIRT employee assigned to communicate with management. In crisis situations, managers want to know the details of a problem, the extent to which a system is compromised, how much time is needed for recovery, the potential consequences, and the current situation. In some cases, managers want to receive status reports at short time intervals, which can lead team members to become more concerned with this communication than with solving the problem. In the midst of an ongoing incident, the team should be focused on their work and the contact person should be familiar with the procedure, should be in the area where the team is working, and should deliver information about the situation with minimal disruption to the team.

It is necessary to designate a dedicated point of contact in an organization's management as well, to which findings and reports are delivered. In a crisis, many people may be interested in learning the most current information, but few may have that right. A CSIRT must have a clear policy as to who provides information from the CSIRT, and to whom they provide it.

Some situations require communication outside an organization, too. There is a good chance that an attack has already occurred elsewhere of the same or similar form, and that someone in the CSIRT community knows how to best defend against such an attack. International organizations of CSIRTs, such as TF-CSIRT or FIRST, can help in these cases. If your CSIRT is a member, communication should be initiated by the team representative listed as your contact person for these organizations; otherwise, communication can be initiated through the national CSIRT. There are other external actors that can be helpful in resolving an incident, too. For example, an Internet Service Provider (ISP) can help block some types of network attacks or trace the origin of an attack, a hosting provider can disable a malicious server in its data center being used to attack your system, and software and hardware vendors can help clarify the logs created by their products.

In some situations, it is necessary to inform the public about an incident. Decisions about this kind of communication must be made by someone competent for this within the organization the CSIRT serves, based on regulations in force in the organization and an assessment of the current situation and any consequences a cyber incident will produce. Organizations may choose not to release any information about an incident due to the risk of reputational damage, but this poses other significant risks if information about the incident becomes public in other ways.

Communication with the public must not jeopardize or hinder the activities of the CSIRT and any others working to handle a cyber incident, and should be aimed at eliminating the speculation that often emerges in such scenarios. This communication can be handled by a predetermined member of the CSIRT who is trained and prepared for the job, or someone from the organization in charge of public communication to whom the CSIRT delivers information. In either case, all the information coming out of a CSIRT goes through a single contact person, and other team members are not permitted to share any information about the incident to the public.

When a serious incident is in progress, the public tends to be hungry for new and detailed information, which is not necessarily provided by an official spokesperson. Journalists will not miss an opportunity to interview someone working directly on resolving the incident, if such an opportunity arises. Remember, if you do find yourself approached by reporters during a crisis, sharing any sensitive information could devastate the efforts of the CSIRT to resolve the incident. A simple 'no comment' is a polite and rational response. You should not expect that journalists will be satisfied with this, however, or even general information about an incident; it is their job to get you to talk.

Evidence gathering

A considerable amount of documentation is collected during incident resolution, and it must be handled carefully to be admissible in court. If legal proceedings take place, these documents represent evidence that must be treated in accordance with relevant regulations. Naturally, the first consideration of incident handlers is how to resolve an incident, but it is vital to think longer-term as well. If law enforcement authorities are to identify and prosecute the perpetrator, you

must provide them with quality evidence that is valid and irrefutable. Evidence must therefore be collected and stored in a way that aligns with procedures that meet all applicable laws.

From the moment it is obtained to when it is handed over to authorities, evidence should be safeguarded. Thus, it is necessary to clearly track who has had custody of evidence at all times. If evidence is transferred from one person to another, this should be documented in detail and signed by both parties. This process of protecting evidence from its collection to delivery is called maintaining the **chain of custody,** and a corresponding log file should contain **at least** the following data:

- information that can indisputably identify the evidence (e.g., the model name or serial number of equipment, a file name with the hash value and date and time of creation, IP addresses, MAC addresses, etc.);
- the name, position, and contact details of each person who collected or handled evidence during incident resolution;
- the precise date and time evidence was collected, and any changes in its handling; and
- the method and location of evidence storage during the period of incident resolution.

More information on evidence collection and preservation can be found in comprehensive publications issued by the US Department of Justice[140] and NIST.[141]

Incident handling checklist

When incidents occur, CSIRT employees are inevitably under stress. Hence, it is always useful to have a written list of actions to follow. This prevents wasted time and allows CSIRT members to think less about what to do, and to just do it. The basic checklist below – from NIST[142] – is one example of an incident handling checklist, which should be customized to the needs of a CSIRT and to the nature of each incident:

140 US Department of Justice, Computer Crime and Intellectual Property Section, Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations, 3rd Edition, OLE Litigation Series (Washington, DC: Office of Legal Education, Executive Office for US Attorneys, 2009). Available at: https://www.justice.gov/file/442111/download

141 Karen Kent et al., Guide to Integrating Forensic Techniques into Incident Response: Recommendations of the National Institute of Standards and Technology, Special Publication 800-86 (US Department of Commerce, 2006). Available as a PDF at: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-86.pdf

142 See: https://www.hhs.gov/sites/default/files/incident-handling-checklist.pdf

Action		Completed
Detection and Analysis		
1.	Determine whether an incident has occurred	
1.1	Analyze the precursors and indicators	
1.2	Look for correlating information	
1.3	Perform research (e.g., search engines, knowledge base)	
1.4	As soon as the handler believes an incident has occurred, begin documenting the investigation and gathering evidence	
2.	Prioritize handling the incident based on the relevant factors (functional impact, recoverability effort, etc.)	
3.	Report the incident to the appropriate internal personnel and external organizations	
Containment, Eradication, and Recovery		
4.	Acquire, preserve, secure, and document evidence	
5.	Contain the incident	
6.	Eradicate the incident	
6.1	Identify and mitigate all vulnerabilities that were exploited	
6.2	Remove malware, inappropriate materials, and other components	
6.3	If more affected hosts are discovered (e.g., new malware infections), repeat the Detection and Analysis steps (1.1, 1.2) to identify all other affected hosts, then contain (5) and eradicate (6) the incident for them	
7.	Recover from the incident	
7.1	Return affected systems to an operationally ready state	
7.2	Confirm that the affected systems are functioning normally	
7.3	If necessary, implement additional monitoring to look for future related activity	
Post-incident Activity		
8.	Create a follow-up report	
9.	Hold a lessons learned meeting (mandatory for major incidents, optional otherwise)	

What next?

By now, you undoubtedly understand that your work as a CSIRT employee demands a continuing commitment to education and skills strengthening. With each new training you gain new knowledge, with each new incident you are more informed by experience, and with each new workshop you collaborate more adeptly with colleagues. Still, the field of cybersecurity is incredibly broad and so much knowledge emerges every day, making it impossible to absorb it all. This is why working as part of a CSIRT is truly a team sport. You rely on your teammates, and the larger organization relies on your team.

So, be kind to your colleagues when they make a mistake; it can happen to anyone. But at the same time, be disciplined with yourself when you make a mistake, analysing why it happened and what you did wrong so that it won't happen again. Most importantly, learn, learn, learn.

Enjoy this demanding but glorious job, and good luck!

DCAF Geneva Centre for Security Sector Governance

DCAF Geneva Headquarters

P.O.Box 1360
CH-1211 Geneva 1
Switzerland

info@dcaf.ch
+41 (0) 22 730 9400

www.dcaf.ch

@DCAF_Geneva