Иван Иванов Благоев

КИБЕРСИГУРНОСТТА КАТО СИСТЕМА: ОТ ТЕСТОВЕ И ОДИТИ ДО РАЗУЗНАВАНЕ И ИЗКУСТВЕН ИНТЕЛЕКТ

Асоциация КИЕЕС,

Русе, 2026

Автор: Иван Иванов Благоев

Институт по информационни и комуникационни технологии

Българска Академия на Науките

Издателство: Асоциация КИЕЕС

ISBN 978-619-92403-7-3

eISBN 978-619-92403-8-0

Печат: УИЦ при Русенски университет “Ангел Кънчев”

Ivan Ivanov Blagoev

CYBERSECURITY AS A SYSTEM: FROM TESTING AND AUDITING TO CYBER INTELLIGENCE AND ARTIFICIAL INTELLIGENCE

2026

Ivan Ivanov Blagoev
Institute of Information and Communication Technologies
Bulgarian Academy of Sciences

ISBN 978-619-92403-7-3

eISBN 978-619-92403-8-0

Publisher: CIEES Association
Printed by: University Publishing Center at “Angel Kanchev” University of Ruse

Предговор

Настоящото изследване представя цялостен, систематичен и критичен анализ на съвременните подходи, инструменти и стратегии в областта на киберсигурността. Фокусът е поставен върху четири основни направления: средствата и технологиите за защита, методологиите за тестване на сигурността, добрите практики при провеждане на одити, както и върху използването на кибер полигони като симулационна и обучителна среда. В уводната част е обоснована актуалността на темата, разглеждайки нарастващата честота и сложност на кибератаките, както и необходимостта от интегриран и адаптивен подход към защитата на информационната инфраструктура.

Аналитичната част на изследването разглежда еволюцията на защитните технологии, от класическите мрежови средства до внедряването на изкуствен интелект в системите за откриване на заплахи. Внимание е отделено на различните методи за тестване, включително подходите „бяла кутия“, „черна кутия“ и „сива кутия“, както и на провеждането на проникващи тестове и автоматизирани сканирания за уязвимости. В частта за одитиране са разгледани вътрешни и външни одити, етапи на планиране, оценка и отчитане, както и приложими стандарти и рамки. Специално внимание е отделено на концепцията за кибер полигони, където са представени архитектурни модели, типови сценарии и възможности за използване при обучение, симулации и реалистично тестване на отбранителни механизми.

Разделът, посветен на кибер разузнаването, разширява приложния обхват на изследването, като разглежда структурата на източниците на разузнавателна информация, като това преминава от OSINT и SOCMINT до разузнаване в тъмната мрежа и връзката им с изграждането на проактивна защита. Отделно внимание е отделено на интеграцията на AI и ML модели в

киберсигурността, включително техните възможности, ограничения и етични предизвикателства.

Проведеното изследване представлява практически ориентирано и методологично съдържание за специалисти, изследователи и всички, ангажирани с повишаване на устойчивостта на цифровата среда. Представени са конкретни насоки за изграждане на съвременни модели на киберзащита в условията на динамично променящ се рисков пейзаж.

Foreword

This study presents a comprehensive, systematic, and critical analysis of contemporary approaches, tools, and strategies in the field of cybersecurity. The focus is placed on four principal areas: security technologies and protective mechanisms, security testing methodologies, best practices in conducting cybersecurity audits, and the use of cyber ranges as simulation and training environments. The introductory section substantiates the relevance of the topic by examining the increasing frequency and sophistication of cyberattacks, as well as the necessity for an integrated and adaptive approach to the protection of information infrastructure.

The analytical part of the study explores the evolution of security technologies, from classical network-based defenses to the integration of artificial intelligence in threat detection systems. Particular attention is given to various security testing methods, including white-box, black-box, and gray-box approaches, as well as the execution of penetration testing and automated vulnerability scanning. The auditing section addresses both internal and external audits, covering the stages of planning, assessment, and reporting, along with applicable standards and frameworks. Special emphasis is placed on the concept of cyber ranges, where architectural models, typical simulation scenarios, and opportunities for application in training, simulations, and realistic testing of defensive mechanisms are presented.

The section dedicated to cyber threat intelligence broadens the applied scope of the research by examining the structure of intelligence sources, ranging from OSINT and SOCMINT to dark web intelligence, and their role in building proactive defense capabilities. Additional focus is placed on the integration of artificial intelligence and machine learning models in cybersecurity, including their capabilities, limitations, and ethical challenges.

The conducted research provides practically oriented and methodologically grounded content for professionals, researchers, and all stakeholders engaged in enhancing the resilience of the digital environment. Concrete guidelines are presented for the development of modern cybersecurity models in the context of a dynamically evolving risk landscape.

СЪДЪРЖАНИЕ

Списък на използвани съкращения и означения

AI Artificial Intelligence (Изкуствен интелект)

AV Antivirus (Антивирусен софтуер)

BGP Border Gateway Protocol

CTI Cyber Threat Intelligence (Кибер разузнаване)

CVE Common Vulnerabilities and Exposures

DDoS Distributed Denial of Service (Разпределена атака за отказ от услуга)

DNS Domain Name System

DMZ Demilitarized Zone (Демилитаризирана зона в мрежата)

EDR Endpoint Detection and Response (Откриване и реагиране на крайни точки)

GDPR General Data Protection Regulation (Общ регламент за защита на данните)

IDS Intrusion Detection System (Система за откриване на проникванията)

IPS Intrusion Prevention System (Система за предотвратяване на проникванията)

I2P Invisible Internet Project (Анонимна мрежова инфраструктура)

IoC Indicators of Compromise (Индикатори за компрометиране)

IT Information Technology (Информационни технологии)

ML Machine Learning (Машинно обучение)

MFA Multi-Factor Authentication (Многофакторна автентикация)

MITM Man-In-The-Middle (Атака „човек по средата“)

NIST National Institute of Standards and Technology (Национален институт по стандарти и технологии – САЩ)

Nmap Network Mapper (Инструмент за сканиране на мрежи)

OSINT Open Source Intelligence (Разузнаване от отворени източници)

RDP Remote Desktop Protocol

SIEM Security Information and Event Management (Управление на събития и информация по сигурността)

SOC Security Operations Center (Център за управление на сигурността)

SOAR Security Orchestration, Automation and Response

SSH Secure Shell (Сигурен отдалечен достъп)

TOR The Onion Router (Анонимизираща мрежа)

URL Uniform Resource Locator

VPN Virtual Private Network (Виртуална частна мрежа)

YARA Yet Another Ridiculous Acronym (инструмент за описание на малуер шаблони)

Firewall Защитна стена

Email Електронна поща

УВОД

В условията на нарастваща цифровизация и постоянна свързаност, значението на киберсигурността непрекъснато расте. Организациите са изправени пред все по-сложни кибератаки, от злонамерен софтуер и рансъмуер до целенасочени мрежови пробиви и атаки от типа отказ от услуга. Според глобално проучване на SoSafe, 87 % от организациите съобщават, че през последната година са били засегнати от кибератаки, включително такива, използващи технологии с изкуствен интелект. Тези факти подчертават неотложната необходимост от надеждни мерки за киберсигурност и проактивен подход за защита на информационните ресурси.

Какво представлява киберсигурността? Киберсигурност наричаме съвкупността от технологии, политики, процеси и практически мерки, предназначени да защитят конфиденциалността, целостта и наличността на информацията и системите от кибератаки. Тя включва както технически средства (хардуерни и софтуерни решения), така и организационни дейности (обучения, процедури, одити). Ефективната киберсигурност изисква многостепенна защита, т.е. комбинация от различни средства, всяко адресиращо определен аспект от заплахите. Например, една организация може едновременно да използва мрежова защитна стена, системи за откриване на проникване, решения за защита на крайни точки, контрол на достъпа, криптиране на комуникациите и други, за да изгради т.нар. „дълбока защита“ (defense in depth) и да минимизира шансовете за пробив.

Но технологиите сами по себе си не вършат цялата работа. Затова са нужни и редовни проверки. Тестването на киберсигурността, като сканиране за уязвимости, етичен хакерски тестове и други. Също така помага да открием слабостите, преди някой злонамерен да се възползва от тях. Одитите пък дават независима оценка къде се намираме и дали сме в крак с най-добрите практики и стандарти, като ни предлагат план за подобрение. И не бива да забравяме човешкия фактор, той си остава критичен. Инвестициите в обучение и практически тренировки (например чрез кибер полигони) повишават подготовката на екипите, когато се случи реален инцидент.

Настоящата монография прави задълбочен, но практично насочен анализ на ключовите аспекти на киберсигурността: средствата за защита на инфраструктурата, методите за тестване и оценка, подходите за одитиране, възможностите за обучение чрез кибер полигони, стратегическите измерения на кибер разузнаването. Накрая са формулирани практически препоръки как да изградим по-устойчива цифрова среда.

1. СРЕДСТВА И ТЕХНОЛОГИИ ЗА КИБЕРСИГУРНОСТ

Съвременната киберсигурност разчита на богат набор от технически средства и решения. Те действат на различни нива, от периметъра на мрежата, през вътрешните мрежови сегменти и крайни устройства, чак до облачните ресурси и приложения. Често тези технологии се класифицират според ролята им в процеса на защита: превенция, откриване, реакция и възстановяване. Например, защитните стени и антивирусите работят основно като средства за превенция. Системи за мониторинг и анализ (като SIEM) се фокусират върху своевременното откриване на заплахи. Инструментите от тип SOAR автоматизират реакцията при инциденти, а решенията за архивиране и резервни копия подпомагат бързото възстановяване след атака. Разбира се, много от средствата покриват повече от една функция и границите помежду им се размиват.

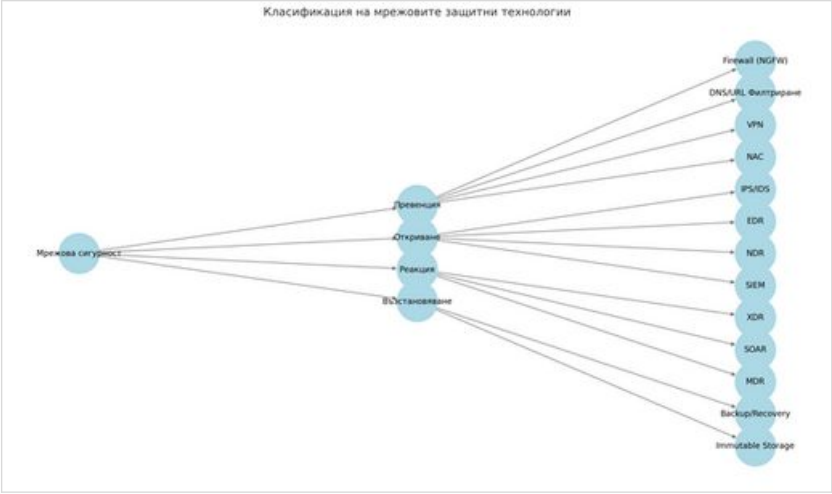
В тази глава се разглеждат основните категории защитни технологии, техните функции и примери за прилагане в различни организационни контексти. Ще анализираме както утвърдени комерсиални решения, така и достъпни алтернативи с отворен код, с акцент върху предимствата и недостатъците им, практическите аспекти на внедряване и влиянието на бюджетните ограничения.

1.1. Основни средства за мрежова сигурност

“Основни средства” обикновено се отнася до решенията, които директно защитават мрежовата инфраструктура и контролират потока от данни. Тук се включват мрежови защитни стени, системи за предотвратяване на проникванията, антивирусен софтуер, уеб и DNS филтри, инструменти за управление на атакуемата повърхност, VPN за отдалечен достъп, и контрол на

мрежовия достъп (NAC). Тези средства са предимно насочени към превенция, но някои имат и функции за откриване и реакция.

На Фиг. 1 е показана диаграма за класификация на мрежовите защитни технологии:



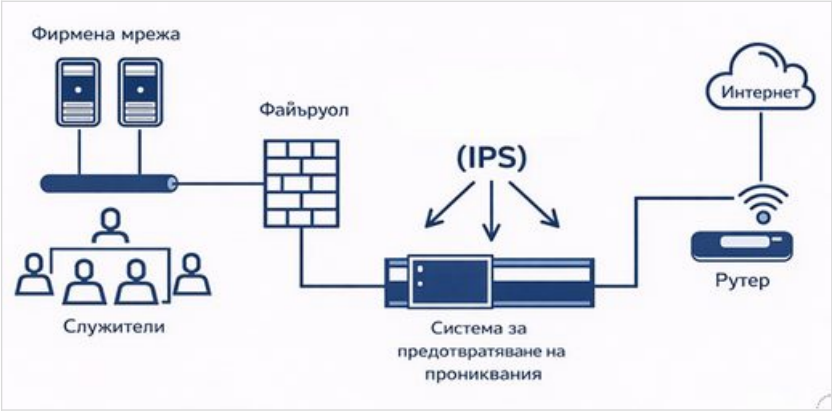
Фиг. 1. Класификация на мрежовите защитни технологии

Защитни стени (Firewalls). Мрежовата защитна стена е устройство или софтуер, който следи и контролира входящия и изходящия трафик между различни мрежови сегменти въз основа на предварително зададени правила за сигурност. Firewall действа като бариера между доверена вътрешна мрежа и недоверена външна (напр. интернет), като инспектира пакетите данни и решава дали да ги пропусне или блокира. Съвременните наследяващи защитни стени (Next-Generation Firewalls, NGFW) надграждат традиционните, като включват допълнителни функции с по-дълбок оглед на трафика, интегриран IPS (Intrusion Prevention System), антивирус, филтриране на уеб съдържание и др. Това обединяване на функционалности и позволява централизиран контрол с по-ефективни политики за сигурност.

Например, една NGFW може едновременно да блокира неоторизиран IP трафик, да разпознава и спира известни атаки и злонамерен софтуер, и да филтрира опасни уебсайтове, осигурявайки многостранна защита в една платформа. На пазара съществуват множество комерсиални firewall решения от производители като Cisco, Palo Alto, Check Point, Fortinet и др., които често интегрират горепосочените възможности. От друга страна, налични са и безплатни или отворени алтернативи като например pfSense и OPNsense (базирани на отворен код), които позволяват на по-малки организации да внедрят защитна стена върху стандартен хардуер (напр. Linux/BSD-базиран сървър) с по-ниски разходи.

Системи за предотвратяване на проникванията (Intrusion Prevention Systems, IPS).

IPS е технология, предназначена да разпознава и блокира зловредни активности в мрежата в реално време. Тя разширява концепцията на системите за откриване на проникванията (IDS - Intrusion Detection System), като не само алармира при подозрителен трафик, но и автоматично прилага мерки за блокиране. IPS наблюдава мрежовия трафик (включително и криптиран при по-новите решения) и сравнява модела на пакетите с база от известни сигнатури на атаки или аномалии в поведението. При съвпадение на например разпознаване на известна атака или необичайна последователност от заявки, системата може незабавно да отхвърли зловредния трафик и да уведоми администраторите. По този начин се предотвратява достигането на атаката до критичните ресурси. Съвременните IPS често предлагат и възможност за виртуални “пачове”, като временно блокиране на уязвимости на ниво мрежа, преди доставчиците да пуснат официални обновления, което повишава гъвкавостта на защитата.



Фиг. 2. Системи за предотвратяване на проникванията

Много от модерните защитни стени включват IPS функционалност, но съществуват и самостоятелни IDS/IPS системи. Пример за комерсиално IDS/IPS решение е Snort (първоначално разработен от Sourcefire, сега част от Cisco), също и Suricata, двете могат да се ползват и като open-source инструменти. Тези отворени платформи се внедряват често върху Linux-базирани сървъри и предоставят на организациите с ограничен бюджет способността да следят трафика за атаки, макар да изискват повече експертни усилия за настройка и поддръжка.

Антивирусен софтуер и „пясъчници“ (Sandboxing)

Антивирусните програми и технологиите за изолиране на изпълнението (sandbox) са ключови за идентифициране на зловредни файлове и програми. Антивирусът сканира файловете за известни сигнатури и модели на поведение на малуер, като блокира или карантинира тези, които съответстват на познати заплахи. “Sandbox” (букв. „пясъчник“) е защитена среда, в която подозрителни файлове могат да бъдат изпълнени, за да се наблюдава поведението им без риск за реалната система. Комбинацията от антивирус и sandbox позволява многослоен

подход: ако файл не бъде засечен от антивирус (например нов, непознат малуер), той може да бъде изолиран в sandbox, където евентуалните му зловредни действия (напр. опит за промяна на системни файлове или необичайна мрежова активност) ще индикират заплаха. Съвременните решения все по-често интегрират AI/ML техники, които ускоряват анализа и подобряват разпознаването дори на непознати атаки в рамките на части от секундата. На пазара има множество антивирусни решения на Symantec, McAfee, Kaspersky, Trend Micro и др., които често включват и облачни "sandbox" услуги за анализ на файлове. Като свободна алтернатива може да се посочи ClamAV (open-source антивирус), комбиниран с инструменти като Cuckoo Sandbox за динамичен анализ на злонамерен софтуер. Също така, експертите по сигурност разработват собствени YARA правила (patern) за откриване на специфични малуер заплахи в рамките на файлове и памет, което предоставя гъвкавост при реагиране на новоизникващи атаки.

Уеб и DNS филтриране

Тези технологии целят да предпазят потребителите и системите от достъп до злонамерени или нежелани интернет ресурси. DNS филтрирането блокира разрешаването на имена на домейни, които са свързвани с атаки като например фалшиви сайтове за phishing, командни сървъри на ботнети или домейни, използвани за разпространение на малуер. URL (уеб) филтрите контролират достъпа до конкретни уеб адреси или категории сайтове (напр. сайтове, съдържащи зловреден код, или неприемливо съдържание по корпоративните политики). На практика, ако потребител в организацията опита да отвори неблагонадежден уеб сайт, уеб филтърът първо проверява адреса спрямо база данни. Ако сайтът фигурира като опасен (например хостващ злонамерен софтуер), достъпът се предотвратява и потребителят получава предупреждение. Тези инструменти освен че пазят от инфекции, спомагат и за налагане на политики за допустимо

използване на интернет в рамките на организацията (например блокиране на определени категории сайтове с оглед сигурността или продуктивността). Комерсиалните уеб/DNS филтри обикновено са част от по-големи сигурностни платформи като напр. Cisco Umbrella (за DNS защита) или Blue Coat Web Security (сега част от Broadcom/Symantec). Налични са и прости решения с отворен код (като Pi-hole за DNS филтър на мрежово ниво, или конфигурация на прокси сървър Squid с филтриращи списъци), които могат да предоставят базова защита при ограничен бюджет.

Управление на атакуемата повърхност (Attack Surface Management, ASM)

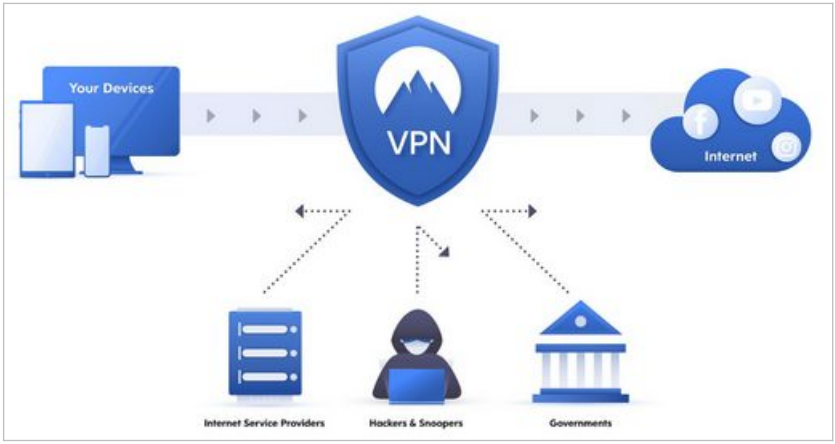
“Атакуема повърхност” наричаме всички точки (активи, устройства, услуги), през които една система може потенциално да бъде атакувана. Нов клас инструменти се появяват за автоматизирано идентифициране и управление на тези активи и уязвимости в мрежата. Те сканират средата, откриват всички налични устройства, приложения, IoT/OT устройства и др., и оценяват тяхното състояние от гледна точка на сигурността като например търсят известни уязвимости, неправилни конфигурации или липсващи актуализации. Резултатът е цялостен инвентар на потенциалните "точки на атака" и оценка къде съществуват рискове. Това позволява на организациите проактивно да намалят атакуемата повърхност, като поправят слабите места (например чрез коригиране на конфигурации, изключване на неизползвани услуги, затягане на контрола върху достъпа и др.). Често тези инструменти се интегрират с други системи за сигурност (например firewall/IPS платформи), за да предоставят препоръки за укрепване на защитата на база откритите пропуски. Сред популярните комерсиални ASM решения са платформи като Rapid7 InsightVM, Qualys VMDR или Microsoft Defender External Attack Surface Management, докато по-опростен подход с отворен код би включвал използване на

скенери като Nmap и OpenVAS, комбинирани с ръчен одит за обхващане на активите.

VPN за отдалечен достъп. Virtual Private Network (VPN)

VPN е технология, която позволява на потребители извън офиса (напр. работещи от дома или в движение) сигурно да се свързват към корпоративната мрежа през интернет.

На Фиг. 3 е показано схематично ролята на VPN услугата:



Фиг. 3. Криптиран тунел за свързаност през VPN услуга

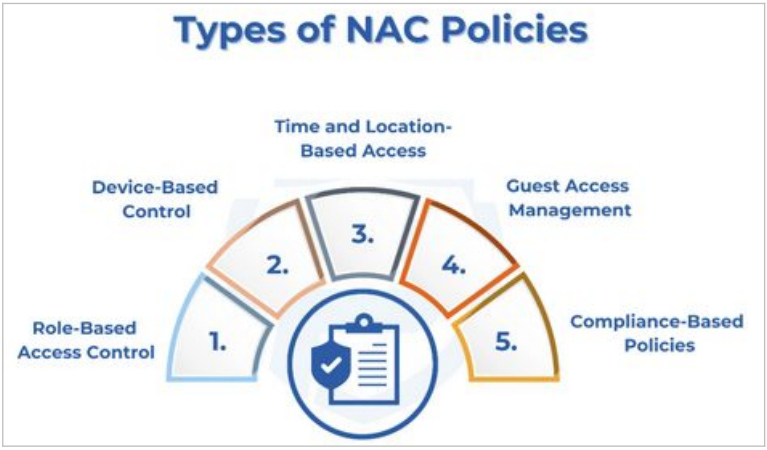
VPN създава криптиран тунел между устройството на потребителя и вътрешната мрежа, което гарантира, че пренасяните данни не могат да бъдат подслушани или променени от външни лица. По този начин служителите могат безопасно да достъпват вътрешни системи и ресурси, сякаш са на място в офиса. Отдалечените VPN решения станаха особено критични при преминаването към модели на дистанционна и хибридна работа, като те осигуряват контрол на достъпа и защита на информацията в ситуации, когато връзките идват от домашни или публични мрежи. С въвеждането на VPN, компаниите намаляват риска от прехващане на трафика и кражба на данни

при отдалечен достъп, тъй като всички комуникации са криптирани с надеждни алгоритми. Примери за корпоративни VPN платформи са Cisco AnyConnect, Palo Alto GlobalProtect, OpenVPN Access Server и др., докато общодостъпни решения като OpenVPN (отворен код) или WireGuard позволяват изграждане на VPN без лицензионни такси.

Контрол на достъпа до мрежата (Network Access Control, NAC)

NAC решенията налагат правила на кой и какво може да се свързва към мрежата. Те идентифицират устройствата при опит за достъп и проверяват дали отговарят на зададени политики за сигурност (напр. дали устройството има актуален антивирус, правилни настройки, нужните удостоверения и т.н.).

На Фиг. 4 е показана имплементация на NAC политики:



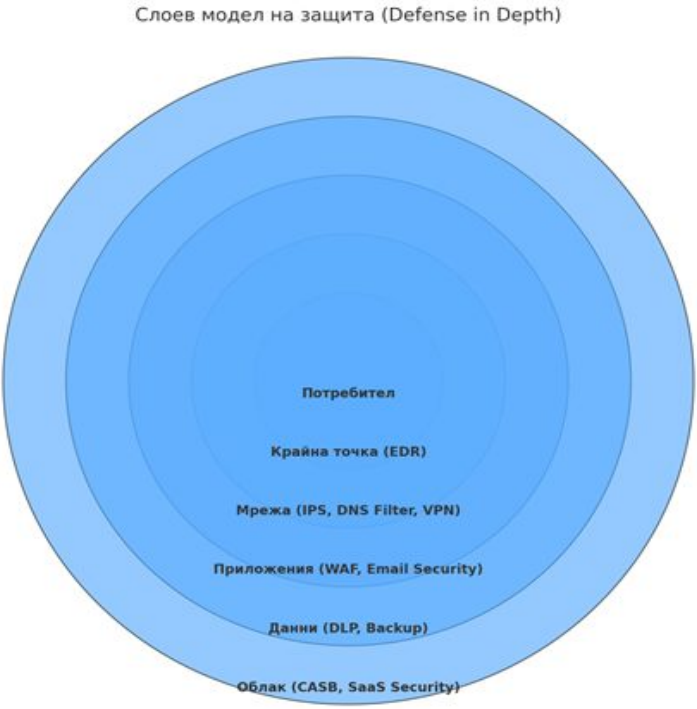
Фиг. 4. NAC политики, https://pristineinfo.com/blog-details/network-access-control

Само устройства и потребители, които са оторизирани и отговарят на изискванията, получават достъп до мрежовите

ресурси. Ако дадено устройство не покрива политиките, като например личен лаптоп без необходимата защита, то NAC може да го изолира в карантинна VLAN, да ограничи достъпа му или изцяло да го блокира. NAC играе ключова роля в корпоративната сигурност, особено с навлизането на разнообразни устройства (включително Internet of Things) и помага да се предотврати свързването на неподходящи или компрометирани устройства, които биха отворили вратичка за атаки. Така се поддържа целостта на мрежата, като всеки входящ елемент се проверява и контролира. Популярни NAC решения на пазара са например Cisco ISE (Identity Services Engine) и HPE Aruba ClearPass, докато като по-достъпна алтернатива с отворен код може да се използва PacketFence или комбинация от FreeRADIUS със скриптове за базов контрол на достъпа.

1.2. Свързани ТЕХНОЛОГИИ ЗА СИГУРНОСТ

Освен преките мрежови защити, съществуват и редица свързани технологии, които допълват киберсигурността на една организация. Те може да не действат пряко на мрежовия трафик, но защитават крайните устройства, данните и приложенията, които са част от общия риск за инфраструктурата. Тук спадат системи като защита на крайни точки (EDR), защита на електронната поща, предотвратяване на изтичане на данни (DLP), защита от DDoS атаки, сигурност на приложенията (напр. WAF), както и брокери за сигурност при достъп до облака (CASB). Макар различни по предназначение, повечето от тези инструменти изпълняват функции по откриване на заплахи и подпомагане на реакцията, а някои работят проактивно за превенция на инциденти.



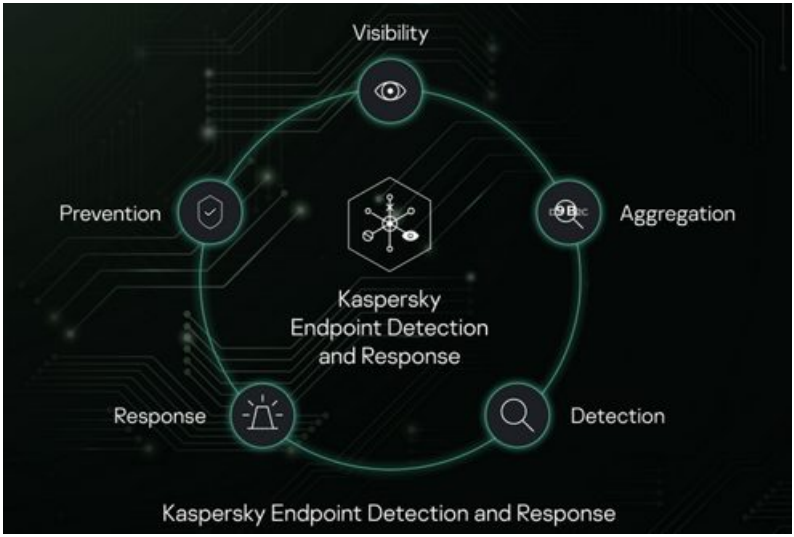
Фиг. 5. Свързани технологии за сигурност

EDR (Endpoint Detection and Response) Откриване и реагиране в крайни точки

EDR технологията непрекъснато наблюдава активността на крайни устройства (компютри, сървъри, мобилни устройства) с цел откриване на признаци на злонамерена дейност на самите тези устройства. EDR решенията събират телеметрични данни, процеси, мрежови връзки, промени по файлове и регистри и др. и използват анализ (включително поведенчески модели и машинно обучение), за да идентифицират подозрителни модели на поведение. Когато бъде засечена потенциална заплаха на крайна точка (например необичайно създаване на процес или опит за неразрешено повишаване на привилегии), EDR системата

може автоматично да реагира: да изолира устройството от мрежата, да прекрати злонамерения процес или да сигнализира екипа по сигурността. Освен откриване, тези системи улесняват и разследването на инциденти, предоставят детайлни журнали и данни за случилото се на засегнатата машина, което помага на анализаторите да разберат и отстранят първопричината. На практика, EDR разширява видимостта на защитата чак до нивото на крайното устройство и допълва мрежовите мерки, като гарантира, че дори ако една заплаха пробие периметъра, тя ще бъде уловена и спряна на хост ниво. Към 2025 г. EDR е стандартна компонента от корпоративната защита, като примери за водещи EDR платформи са CrowdStrike Falcon, Microsoft Defender for Endpoint, Palo Alto Cortex XDR и др.

На Фиг. 6 е показана диаграма на функцията на endpoint Detection и Response:



Фиг. 6. Endpoint Detection and Response, https://www.kaspersky.com/enterprise-security/wiki-section/products/kaspersky-endpoint-detection-and-response-edr

Съществуват и безплатни инструменти за мониторинг на крайни точки с по-ограничени способности като например OSSEC/ Wazuh (с отворен код) може да служи за базово откриване на подозрителна активност на хостове, макар да не предлага пълния набор от реакции на една EDR система.

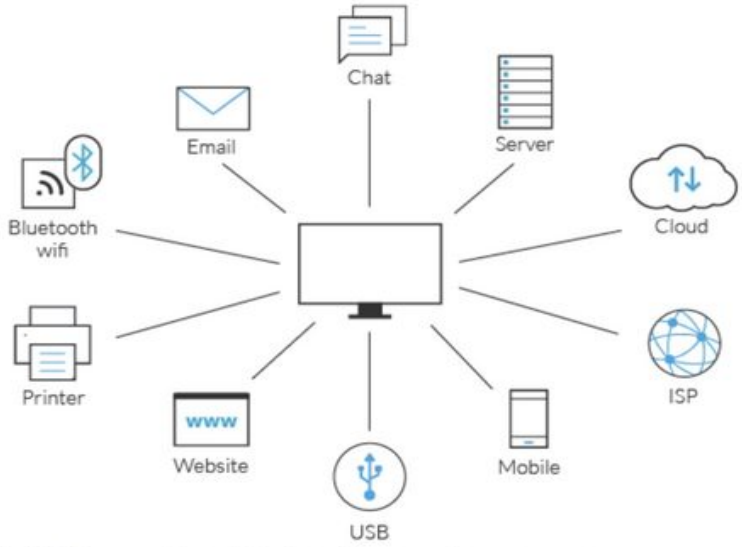
Защита на електронната поща (email security)

Решенията за email security се фокусират върху електронната поща, която е един от най-често експлоатираните канали за атака. Голям процент от пробивите започват с phishing имейл или злонамерено прикачен файл. Затова тези решения филтрират входящите имейли и ги проверяват за подозрително съдържание, известни зловредни прикачени файлове, опасни линкове, измамни податели и др. Системите за защита на пощата могат да поставят под карантина или изтриват опасни съобщения, преди да достигнат до потребителските пощенски кутии. Съвременните платформи използват комбинация от методи: черни списъци на известни вредоносни домейни, анализ на линкове чрез „sandbox“ за URL, валидиране на подателя (DMARC/DKIM/SPF), и дори изкуствен интелект за откриване на фишинг (силно таргетирани измамни писма). Целта е да се намали максимално рискът служител да кликне на компрометиран линк или файл. Електронната поща (Email) защитата е критична, защото човешката грешка (например подлъгване от добре изработено фишинг съобщение) може да заобиколи много други технически мерки, за това е най-добре рисковите съобщения изобщо да не стигат до потребителите. Примери за решения в тази категория са Proofpoint Email Security, Cisco Secure Email и облачните филтри в Microsoft 365 Defender. Малки организации понякога разчитат и на вградени или свободно достъпни средства, като например open-source филтри като SpamAssassin за спам и базов анти-малуер, но напредналите атаки изискват по-усъвършенствани платформи за защита на пощата.

DLP (Data Loss Prevention)

Предотвратяване на изтичане на данни. DLP обхваща инструменти и процеси, целящи да предпазят от неоторизирано изнасяне или разкриване на чувствителна информация.

На Фиг. 7 са показани улсуги, предмет на анализ от DLP:



Фиг. 7, DLP услуги за анализ, https://www.imperva.com/learn/data-security/data-loss-prevention-dlp/

DLP решенията обикновено сканират данните в системите (файлови хранилища, бази данни, имейли и др.), за да идентифицират поверителна информация като напр. лични данни, финансова информация, интелектуална собственост и следят за опити тази информация да напусне рамките на организацията. Например, ако служител се опита да изпрати по имейл файл, съдържащ списък с лични данни на клиенти, DLP системата може да засече това (чрез шаблони, ключови думи или цифрови отпечатъци на данните) и да блокира изпращането или

да алармира. DLP може да се прилага на няколко нива: на крайна точка (мониторинг на действията на потребителите на техните устройства), на мрежово ниво (следене на изходящия трафик за конфиденциални данни), и на ниво облачни услуги и хранилища (управление на правата за достъп и споделяне на файлове). В контекста на нарастващите регулации за защита на данните (GDPR, HIPAA и пр.), DLP е съществен компонент за поддържане на съответствие. Той помага проактивно да се предотвратят нарушения, вместо пасивно да се реагира след като щетата е настъпила. Типични представители на DLP решения са цялостните платформи като Forcepoint DLP, Symantec (Broadcom) DLP, McAfee Total Protection for DLP и др., които позволяват централизирано прилагане на политики. Open-source еквиваленти в пълен мащаб практически липсват, тъй като DLP изисква сложна интеграция. По-малки организации често ползват комбинация от строги политики, обучение на персонала и базови инструменти (например ограничения на USB устройствата, ръчни проверки) вместо скъпоструващ софтуер.

Защита от DDoS атаки

Разпределеният отказ от услуга (Distributed Denial of Service, DDoS) е атака, при която злонамерени лица използват множество компрометирани устройства (ботнет) за едновременно изпращане на огромен брой заявки към дадена система (сървър, уебсайт, услуга) с цел да претоварят ресурсите ѝ и да я направят недостъпна. DDoS атаките могат да причинят сериозни прекъсвания на услугите. Специализираните DDoS защити са системи, които разпознават и блокират подобен зловреден трафик, преди да е затормозил основната инфраструктура. Те работят като „буфер“ пред защитавания обект: наблюдават трафика в реално време и при откриване на аномалии (неестествено високо натоварване, характерно за DDoS) пренасочват или филтрират излишния трафик. Пример за такова решение е платформата Fortinet FortiDDoS, която анализира

входящите пакети и автоматично спира невалидните заявки, за да не достигнат те до сървърите. Доста често DDoS защитата включва и услуги в облака, тъй като ефективното отблъскване на масирани атаки може да изисква разпределена мрежа от филтриращи сървъри по света. Като услуга подобна защита предлагат компании като Cloudflare и Akamai/Prolexic, които имат глобална инфраструктура за поглъщане на атаките. Включването на DDoS защита гарантира, че критичните системи на организацията ще останат достъпни за легитимните потребители дори при опит за саботаж чрез трафик.

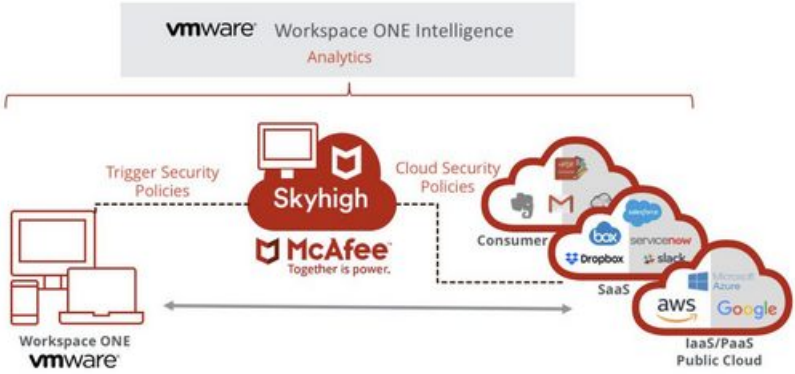
Сигурност на приложенията

Този термин обхваща разнообразни инструменти, които защитават уеб и други приложения от атаки, насочени към уязвимости в самия софтуер. Примери са уеб приложни защитни стени (Web Application Firewall, WAF), системи за защита на API, скенери за уязвимости на приложения, както и механизми, интегрирани в мрежовите устройства. Целта е администраторите да могат да разпознават и контролират специфичния трафик към приложенията, като кои заявки са легитимни и кои потенциално зловредни. Например, WAF може да анализира HTTP заявките към уебсайт и да блокира такива, които съдържат SQL инжекции или XSS скриптове. Някои мрежови решения използват протоколни декодери и сигнатури, за да идентифицират приложенията по техния трафик (напр. да различат трафика от Facebook, YouTube, корпоративна ERP система и т.н.) и позволяват налагане на политики, като определени нежелани приложения да се блокират или ограничават. Подобно приложение с ориентирано управление не само повишава сигурността, но и оптимизира използването на мрежовите ресурси (като ограничава развлекателния трафик и гарантира наличност за бизнес-критичните приложения). Сигурността на приложенията често е тясно свързана с методологиите на разработка (напр. тестване на кода за уязвимости, DevSecOps), но от гледна точка на

мрежовата защита тя се реализира чрез постоянен мониторинг и контрол на приложния трафик. На пазара се предлагат самостоятелни WAF решения (напр. Imperva SecureSphere, F5 BIG-IP WAF), услуги като AWS WAF в облачни платформи, а също и безплатни възможности като ModSecurity (open-source WAF), които могат да се внедрят с по-малко средства, макар и с повече ръчна настройка.

CASB (Cloud Access Security Broker)

С нарастването на използването на облачни услуги (Software-as-a-Service приложения, облачни платформи за данни и др.), възникна нужда от посредник, който да наложи корпоративните политики за сигурност върху тези външни услуги. CASB изпълнява именно тази роля се поставя между потребителя и облачната услуга, за да предостави видимост, контрол и защита при достъпа до облака. CASB решенията могат да работят в режим „прокси“, през който минава целият облачен трафик, или чрез директни интеграции с API-та на облачните доставчици. Те изпълняват няколко функции: удостоверяване на потребителите и гарантиране, че само оторизирани лица имат достъп; прилагане на класификации и ограничения върху данните (напр. да се предотврати качване на поверителни файлове към неуправлявани облачни приложения); откриване на така наречения “shadow IT”, използването на неоторизирани облачни услуги; и защита от заплахи, специфични за облачните приложения. По същество, CASB дава на организацията контрол върху SaaS средата, подобно на този, който има върху собствената си инфраструктура, осигурявайки съответствие и защита дори когато данните са извън прякото ѝ притежание.



Фиг. 8, CASB диаграма, https://vietnetco.vn/en/solutions-en/cloud-virt-sec-en/casb-en

Типични CASB продукти са McAfee MVISION Cloud (Skyhigh), Netskope CASB, Microsoft Defender for Cloud Apps и др. които често са включени като част от по-големи облачни сигурностни платформи. Поради сложността на тези решения, организации с ограничен бюджет може да се насочат към ограничени мерки като строги политики на достъп до определени облачни услуги и обучение на потребителите, ако не могат да си позволят пълноценен CASB.

Резервни копия и възстановяване

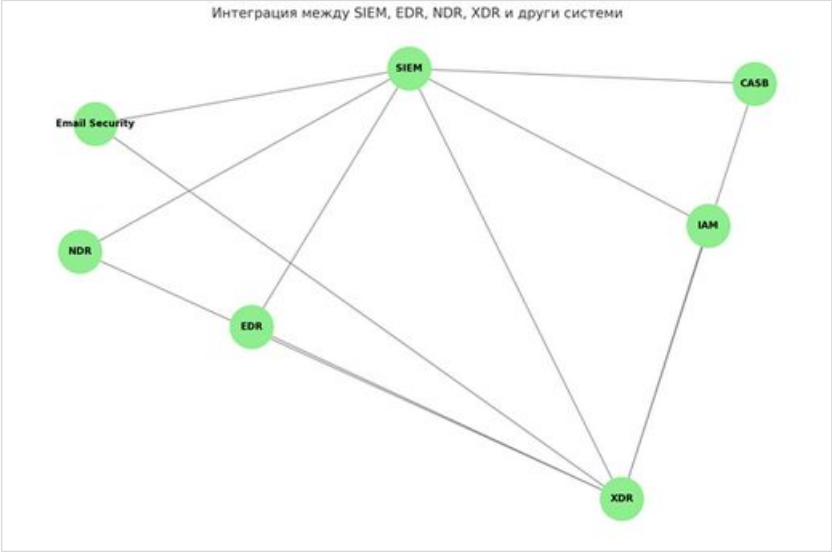
Макар често да се разглежда отделно от „киберсигурността“, технологията за създаване на резервни копия (backup) и планиране на възстановяването при инциденти е ключова част от способността на една организация да се възстанови след кибератака. Например, при рансъмуер (ransomware) атака, където данните биват криптирани от нападателите, наличието на актуални и защитени резервни копия може да бъде единственият начин засегнатата организация да възстанови информацията си без да плати откуп. Системите за бекъп периодично копират критичните данни и системни изображения на сигурно място, било то на физически носител, в отдалечен дата център или в

облачно хранилище. Важни характеристики на съвременните backup решения са автоматизацията (редовни планирани архиви), съхраняване на версиите (за да има копия назад във времето) и имунитетът към атаки (например, offline или неизменяеми копия, които да не могат да бъдат изтрити или криптирани от атакуващия). Освен самото съхранение, планът за възстановяване включва и редовно тестване на възстановителния процес, за да се гарантира, че резервните данни действително могат да бъдат върнати бързо в оперативна среда. Инструментите за backup варират от комерсиални продукти като Veeam Backup & Replication, Acronis Backup, Commvault и др., до open-source решения като Bacula и UrBackup. Изборът често зависи от мащаба и изискванията за надеждност: по-малките организации може да използват по-прости и евтини варианти, докато големите предприятия инвестират в комплексни системи с дублиране на данни в географски отдалечени центрове, което гарантира устойчивост и бързо възстановяване (Disaster Recovery) дори при мащабен инцидент. В контекста на киберсигурността, наличието на здрава стратегия за резервни копия е крайният пласт на защита, който не предотвратява атака, но смекчава най-тежките последици и позволява на бизнеса да продължи да функционира.

1.3. Корпоративни РЕШЕНИЯ ЗА СИГУРНОСТ

Големите организации с обширни и сложни ИТ инфраструктури използват корпоративни платформи за сигурност, които обединяват и корелират информация от много източници. Целта е да се постигне централизирано наблюдение и управление на сигурността в мащаб. Чрез такъв интегриран подход, инцидентите могат да се откриват по-бързо и в по-широк контекст, отколкото при изолирано следене на отделни системи. Тук основни примери са SIEM системите, както и решенията за мрежово и разширено откриване и реагиране (NDR, XDR).

Примерна диаграма за интеграция между SIEM, EDR, NDR и XDR може да се види на Фиг.9:

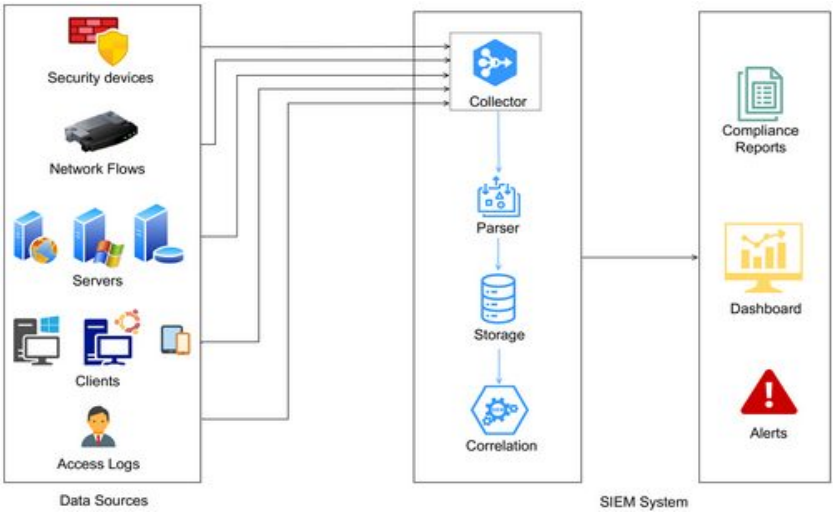


Фиг. 9 Интеграция между SIEM, EDR, NDR, XDR

SIEM (Security Information and Event Management)

SIEM е платформа, която събира и анализира логове и събития от различни системи – мрежови устройства, крайни точки, сървъри, приложения, бази данни и др., за да предостави единна картина на сигурността в организацията. Чрез агрегиране на информацията на едно място, екипите по сигурност (в рамките на Security Operations Center, SOC) могат по-лесно да откриват атаки, които биха останали скрити, ако се наблюдават само отделни системи.

На Фиг. 10 е диаграмата изобразява примерно SIEM решение:



Фиг. 10, Manzoor, J.; Waleed, A.; Jamali, A.F.; Masood, A. Cybersecurity on a budget: Evaluating security and performance of open-source SIEM solutions for SMEs. PLoS ONE 2024, 19, e0301183, https://doi.org/10.1371/journal.pone.0301183

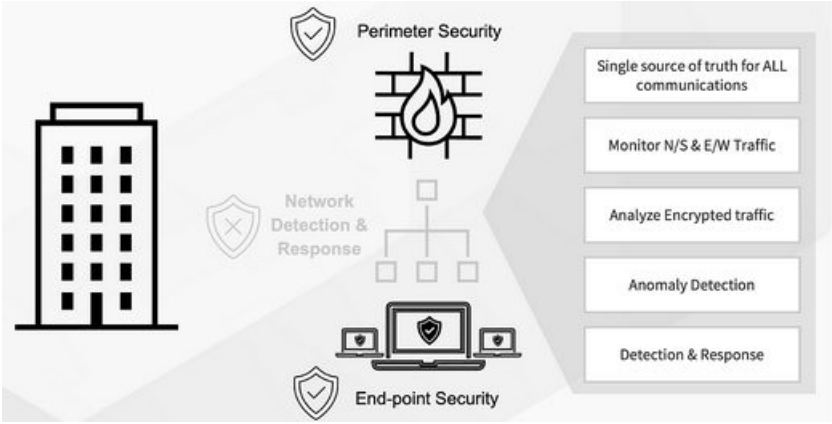
Съвременните SIEM използват правила и модели на заплахи, както и машинно обучение, за да открояят значимите инциденти сред огромния обем данни (т.нар. откриване на инциденти на база поведение). Например, SIEM може да корелира събития като многократни неуспешни опити за логин в различни системи, последвани от успешен вход и промяна на конфигурация и да ги маркира като потенциална атака. Освен откриване, SIEM подпомага и разследването (чрез възможности за търсене и анализ в натрупаните логове) и съответствието с регулации (чрез генериране на отчети за одит и спазване на стандарти). С една дума, SIEM намалява сложността на управлението на сигурността, като обединява многобройните аларми в единна платформа и така позволява по-бърза и информирана реакция. На пазара има множество SIEM решения като например Splunk Enterprise Security, IBM QRadar, ArcSight, както и продукти с отворен код като Elastic Stack (ELK), които в

комбинация могат да изпълняват SIEM функции. Предизвикателство при SIEM е, че внедряването и настройката му изискват значителни ресурси и експертиза, особено при голям мащаб на данните и нуждата от оптимизация на правила, за да се намали „шума“ от незначителни събития.

NDR (Network Detection and Response)

Докато EDR се фокусира върху крайните точки, мрежовото откриване и реагиране (NDR) наблюдава мрежовия трафик вътре в инфраструктурата.

На Фиг. 11 е представена диаграма на NDR решение:



Фиг. 11, https://www.progress.com/blogs/what-is-network-detection-and-response

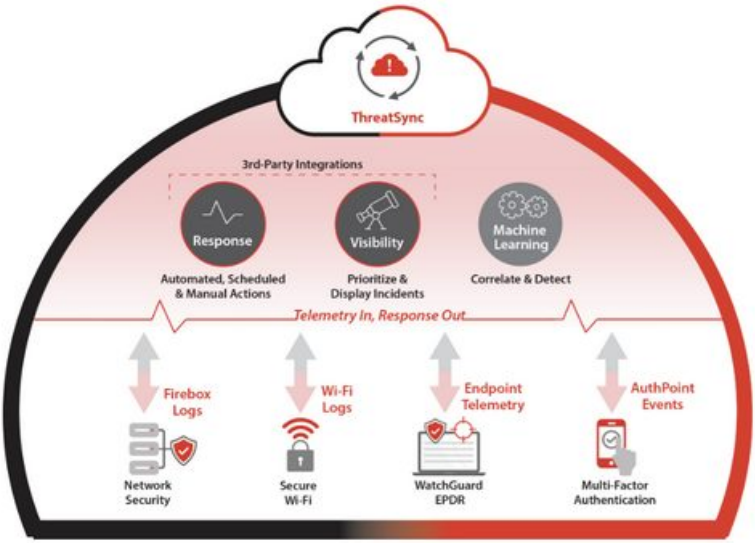
NDR системите събират данни за потока в мрежата, установяват базова линия за „нормално“ поведение и с помощта на аналитични методи (включително ML) търсят аномални модели, които може да са индикатор за проникване. Например, ако изведнъж вътрешен сървър започне да изпраща големи количества данни към външен адрес през необичаен порт, NDR би засекла това като отклонение. Особено ценни са тези решения

за откриване на нападатели, които вече са вътре в мрежата (например чрез компрометиран акаунт или устройство) и се опитват да се придвижват странично или да изнасят данни. При засичане на нещо подозрително, NDR системата може да генерира тревога и дори автоматично да предприеме действие като например да блокира подозрителния трафик или да изолира замесеното устройство. NDR често се интегрира със SIEM и други инструменти, за да обогати контекста на инцидентите. Така се осигурява непрекъснат мониторинг на „източно-западния“ трафик (вътре в организацията), който традиционните периметрови решения не обхващат достатъчно. Примери за NDR решения са Cisco Stealthwatch (Secure Network Analytics), Darktrace, RSA NetWitness NDR и др., те предоставят на големите организации повишена видимост за случващото се в мрежата им, но също изискват опитни анализатори, които да разбират и реагират на получените сигнали. Някои организации с ограничени средства прибягват и до open-source инструменти като Zeek (Bro) за мрежов мониторинг, макар че липсата на удобен интерфейс и автоматизация прави тази опция предизвикателна за по-широко приемане.

XDR (Extended Detection and Response)

Това е сравнително нова категория, която цели да разшири обхвата на откриване и реагиране отвъд единичния слой (само крайни точки, само мрежа и т.н.) към цялостна, многослойна видимост.

На Фиг. 12 е показана диаграма на XDR решение:



Фиг. 12. XDR,https://dolos.africa/xdr/

XDR платформите агрегатират данни и сигнали от различни източници и крайни точки (EDR), мрежов трафик (NDR), облачни услуги, идентичности на потребителите и др., като автоматично корелират отделни събития в единен инцидент. С други думи, XDR автоматизирано „сглобява“ парчетата на пъзела, които биха могли да изглеждат незабележими самостоятелно, но заедно сигнализират за сложна атака. Например, едновременно засечени странен процес на крайна точка и необичаен мрежов трафик, които XDR би обединил това в сигнал за проникване. Предимството е по-бързо разкриване на сложни атаки (като т.нар. APT - усъвършенствани постоянни заплахи), тъй като инструментът филтрира „шума“ и предоставя по-ясно приоритизирани инциденти. XDR често се разглежда като еволюция на концепцията SIEM + SOAR, ориентирана към по-голяма автоматизация и облекчаване работата на анализаторите. Чрез XDR организациите постигат по-висока ефективност в откриването и отстраняването на заплахи, понеже платформата

комбинира множество telemetry данни и може да реагира по-синхронизирано. Пример за XDR решение е Palo Alto Networks Cortex XDR (което комбинира EDR, мрежови данни и облачен мониторинг), както и Microsoft 365 Defender, Trend Micro Vision One и др. Много доставчици на сигурност активно развиват XDR функционалност, като я интегрират в портфолиото си, но пълните ползи се усещат, когато XDR се използва като част от цялостна стратегия с обучен персонал и ясно определени процеси за реакция.

1.4. Управлявани УСЛУГИ ЗА КИБЕРСИГУРНОСТ

Много организации, особено по-малки или такива без достатъчен експертен ресурс, се обръщат към външни доставчици за управлявани услуги в областта на сигурността. Тези услуги предоставят експертиза и инфраструктура „като услуга“, фирмите наемат специализирани екипи и платформи, които да наблюдават и защитават тяхната среда, вместо да изграждат и поддържат всичко сами. Типични примери тук са MDR (Managed Detection and Response), SOC-as-a-Service (SOCaaS) и управлявани услуги за защитни стени (Firewall as a Service, FWaaS).

MDR (Managed Detection and Response)

Управляваното откриване и реагиране представлява външно възложена услуга за денонощен мониторинг на сигурността, откриване на инциденти и реагиране при атаки. В рамките на MDR, доставчик (обикновено MSSP - Managed Security Service Provider) осигурява екип от специалисти и технологии, които наблюдават логовете, крайните точки и мрежата на клиента за признаци на заплахи. При засичане на инцидент, MDR екипът поема първоначалната реакция, като може да изолира засегнатите системи, да ограничи разпространението на атаката и да уведоми организацията и след това съдейства за пълното отстраняване и възстановяване. MDR услугите често включват и

проактивни дейности като лов на заплахи (threat hunting) и препоръки за подобряване на защитата. Основното им предимство е, че осигуряват на организации без собствен SOC достъп до високо ниво на експертиза и усъвършенствани инструменти, при това с по-ниски разходи, отколкото биха били нужни за изграждане на такава способност вътрешно. MDR е естествено продължение на EDR, като вместо компанията сама да следи и тълкува алармите от EDR/сензори, тя поверява тази задача на екип, който обслужва множество клиенти и е видял много инциденти, т.е. може по-бързо да реагира и да приложи най-добрите практики от опита си. На пазара много компании предлагат MDR услуги от специализирани фирми до големи играчи като Microsoft (с услугата Defender Experts) или CrowdStrike Falcon Complete. Изборът често зависи от доверието в доставчика и специфичните нужди (напр. някои MDR са насочени към определени платформи или индустрии).

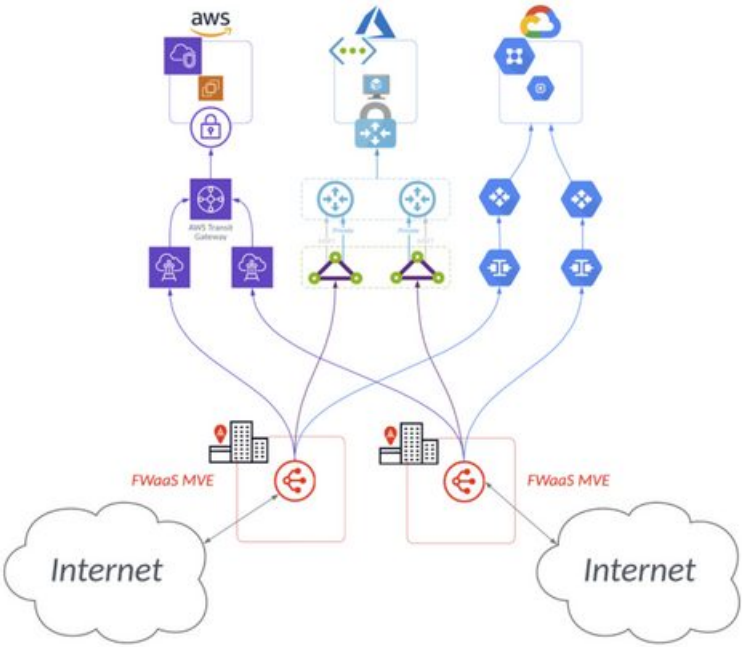
SOC-as-a-Service (SOCaaS)

Това понятие обобщава концепцията, при която целият Security Operations Center (SOC) хора, процеси, технологии се предоставя като услуга от външен партньор. Причините организациите да изберат SOCaaS включват: недостиг на квалифициран персонал по сигурността, високата цена за поддържане на денонощен екип, нуждата от скъпи платформи (SIEM, SOAR и т.н.) и тяхната поддръжка. При SOCaaS, доставчикът осигурява 24x7 наблюдение, обработка на аларми, разследване на инциденти, управление на уязвимости, отчитане и всичко, което един вътрешен SOC би правил, но в модела на аутсорсинг. Обичайно в обхвата влиза и отдалечено управление на сигурността, т.е. доставчикът може да предприема директни действия в средата на клиента (със съответните предварителни уговорки), например да изключи компрометиран потребителски акаунт, да приложи блокиращо правило на защитната стена и др. SOCaaS позволява на фирми, които нямат ресурси сами да поддържат пълен SOC, да

получат еквивалентно ниво на защита и непрекъснат отговор на инциденти, като разходите се споделят между много клиенти на доставчика. Разбира се, при този модел е критично детайлното договаряне на нивата на услуга (SLAs) и изграждането на доверие, тъй като външен екип ще борави с чувствителни данни и системи на организацията. SOCaaS е привлекателен за средни предприятия, които искат качествена защита без да изграждат скъп капацитет сами, и често се предлага от големи MSSP и консултантски компании (напр. IBM Managed Security Services, Dell Secureworks и др.).

Управлявани услуги за защитни стени (Firewall as a Service, FwaaS)

Управляваната защитна стена е пример за специфична услуга, при която конфигурирането, мониторингът и поддръжката на една или повече защитни стени на организацията се поемат от външен експертен екип.



Фиг. 13. Firewall as a Service, https://www.megaport.com/blog/what-is-firewall-as-a-service-fwaas/

В съвременния контекст това често е обвързано с облачни защитни стени, където доставчикът хоства и управлява виртуални firewall инстанции за клиента. Предимствата са няколко: организациите спестяват усилия по обновяване на правила, следене на логове и реагиране на аларми и всичко това се прави от доставчика, който има нужните умения и мащаб; освен това се гарантира, че настройките винаги са оптимизирани и в крак с най-новите заплахи. Управляваната услуга често включва и подробна видимост за клиента, достъп до портали с отчети за трафика, блокирани атаки и препоръки за подобрения, като по този начин организацията получава най-доброто от двата свята: силна защита и експертен надзор без да се налага да търси и назначава допълнителен персонал. При преминаването към облачна инфраструктура, FWaaS улеснява и уеднаквяването на

политиките, като организацията може да има едни и същи защитни правила, прилагани последователно както в локалните ѝ центрове за данни, така и в публичния облак, благодарение на управляемата услуга. В крайна сметка, услугите като MDR, SOCaaS и FWaaS позволяват на организациите да ускорят въвеждането на мерки за сигурност и да си осигурят високо ниво на защита, без да носят изцяло тежестта по управление на сложни технологии и екипи.

1.5. Сравнителен АНАЛИЗ НА ПОДХОДИТЕ ПРИ ИЗБОР НА ТЕХНОЛОГИИ

Съществуват различни подходи при изграждането на системата за киберсигурност, от използване на комерсиални готови решения, през залагане на инструменти с отворен код и собствени разработки, до комбинация от двете. Всеки подход има своите предимства и недостатъци. В Таблица 1 е направено сравнение на някои ключови характеристики на комерсиалните спрямо open-source (отворени) решения.

Таблица 1. Сравнение между комерсиални и open-source решения за сигурност

Аспект	Комерсиални решения	Open-source / Собствени решения
Цена (лицензи и Total Cost of Ownership (TCO))	Значителни първоначални разходи и текущи лицензионни такси. Включват се често договори за поддръжка.	Без (или ниски) разходи за лиценз. Изискват се обаче инвестиции във време и труд (конфигурация, поддръжка). По-нисък пряк разход, но скритите разходи (персонал) може да се натрупат.
Поддръжка и обновления	Професионална поддръжка от доставчика, редовни	Разчитане на общността или собствен екип за поддръжка. Обновленията

	обновления и кръпки. Service Level Agreement (SLA) гаранции за време за реакция при проблеми.	зависят от доброволци или вътрешен капацитет. Може да липсва официална гаранция или отговорност при инциденти.
Функционалност и интеграция	Завършени продукти с богати функционалности, често "всичко в едно". По-лесна интеграция с други продукти на същия доставчик. Ползват се стандартни интерфейси за връзка с външни системи.	Гъвкав избор на компоненти според нуждите. Възможност за интегриране на "best-of-breed" инструменти от различни източници. Необходимо е повече усилие за интеграция и тестване на съвместимостта между компонентите.
Гъвкавост и пригодимост	По-затворени системи – възможностите се определят от доставчика. Ограничена възможност за модифициране на кода или добавяне на нестандартни функционалности.	Пълен достъп до сорс кода позволява адаптиране към специфични нужди. Вътрешните екипи могат да разработват нови модули. Позволява се експериментиране, но изисква експертиза.
Изисквания за персонал	Не е необходим дълбок вътрешен експертен опит за базово използване. Доставчикът осигурява обучение, документация, понякога и управлявани услуги. Все пак е нужен екип	Изисква опитни ИТ/сигурност специалисти, които да внедрят, настроят и управляват системите. Необходими са повече време за обучение и поддръжка. Зависимост от наличието на ентусиасти/експерти в екипа.

	за мониторинг и реакция.	
Сигурност и надеждност	Изпитани във време решения, преминали през щателно тестване. Доставчикът поема отговорност при уязвимости (с издаване на пачове). Възможно е обаче vendor lock-in, заради зависимост от един доставчик за сигурността.	Прозрачност на кода, като уязвимостите могат да бъдат открити от общността. Липсва единна отговорност и сигурността зависи от усилията на организацията. Ако липсват навременни обновления, може да има по-висок риск.
Приложимост	Подходящи за организации с достатъчен бюджет, които искат бързо внедряване и гарантирана поддръжка. Често предпочитани при изисквания за сертифицирани технологии или когато липсва вътрешен екип за разработка.	Подходящи за организации с ограничени средства или специфични нужди, които не се покриват от готовите продукти. Дава контрол върху системата, но е оправдано, ако организацията може да си позволи експертизата за поддържането му.

От таблицата се вижда, че комерсиалните решения осигуряват по-бърз старт, професионална поддръжка и цялостен набор от функции „от кутията“, докато open-source подходът дава повече гъвкавост и спестява преки разходи, но изисква сериозен собствен ресурс (човешки и времеви). В практиката често се прилага комбиниран подход: например, базова инфраструктура като firewall и endpoint защита да са покрити от утвърдени

комерсиални продукти, докато за допълнителни мониторинг функции или нишови задачи (например специфични анализи с YARA правила, custom SIEM корелации с ELK и др.) се използват решения с отворен код. Вземането на решение зависи от приоритетите на организацията, някои ценят повече независимостта и ниската цена, други удобството и гаранциите от доставчика.

На Фиг. 14 изобразява графично сравнение между комерсиални и open-source кибер решения:

Фиг. 14. Сравнителна матрица на Комерсиални и Open-source решения

Друг аспект на сравнение е изборът между интегрирани платформи и самостоятелни решения от различни доставчици. Някои производители (напр. Cisco, Palo Alto, Microsoft) предлагат цялостни екосистеми, където различните компоненти на

сигурността (мрежова защита, крайни точки, облачен контрол и т.н.) работят заедно и обменят информация. Предимство тук е по-лесната интеграция и централизирано управление, например алармите от различни слоеве се събират на една конзола. Недостатък може да бъде зависимостта от една компания и възможни компромиси в някои модули (не винаги един доставчик е най-добрият във всяка категория). Алтернативно, организацията може да подбере най-добрия продукт за всяка отделна нужда (например отделен специализиран firewall, отделен топ SIEM, отделна EDR система). Това потенциално максимизира ефективността на всеки слой, но усложнява интеграцията и управлението, необходими са усилия различните системи да „си говорят“ (например чрез API или стандарти като STIX/TAXII за обмен на информация за заплахи). Често по-големите предприятия избират многофирмен подход, тъй като имат ресурси да интегрират и поддържат сложна среда, докато по-малките предпочитат цялостни решения за простота.

Няма универсално „най-добро“ решение, сравнението на подходите показва, че изборът трябва да балансира разходите, рисковете и капацитета на организационния екип. Понякога оптималното е комбинация: например, комерсиална платформа за критичните функции и отворени инструменти за допълнително подсилване и мониторинг.

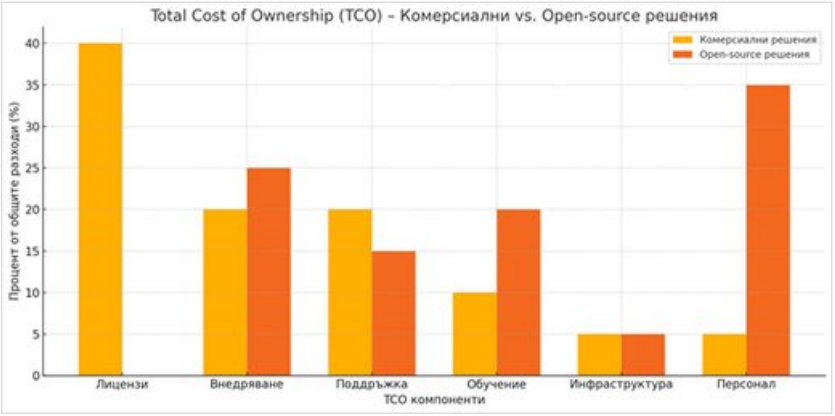
1.6. Финансови СЪОБРАЖЕНИЯ И ВЛИЯНИЕ ВЪРХУ ИЗБОРА НА ТЕХНОЛОГИИ

Финансовите ограничения играят ключова роля при планирането на киберсигурността. Бюджетът определя колко и какви технологии могат да се внедрят, а организациите често трябва да правят компромиси между желаното ниво на защита и достъпните средства. Някои важни аспекти на влиянието на финансите върху избора на решения са:

- **Приоритизация на инвестициите:** Обичайно първо се финансират мерките за превенция на най-критичните рискове (например защитни стени, основна защита на крайни точки), след което са средства за откриване (мониторинг, алармиране) и реакция. Ако бюджетът е ограничен, ръководството трябва да реши кои активи и данни са най-ценни и срещу кои заплахи е най-вероятно да се сблъска организацията, и да насочи инвестициите натам. Инструменти, които адресират множество рискове (например SIEM, който помага и за откриване, и за спазване на регулации) може да получат предимство пред тесни нишови решения.
- **Total Cost of Ownership (TCO):** При оценка на технологията трябва да се отчита не само цената на закупуване/лиценз, но и съпътстващите разходи за инфраструктура, внедряване, обучение на персонала, текуща поддръжка, и евентуални бъдещи ъпгрейди. Понякога на пръв поглед скъпо решение с добра автоматизация може да се окаже по-изгодно дългосрочно от евтина система, която изисква много ръчни усилия (което означава разходи за труд).
- **Отворен код и безплатни инструменти:** Организации с по-малки бюджети активно разглеждат безплатни или open-source инструменти като алтернатива на скъпите продукти. Както обсъдихме в предходната секция, този подход спестява преки разходи, но трябва да се подсигури наличието на квалифицирани хора, които да извлекат максимума от тези инструменти. Често собственият труд се използва вместо пари, като напр. настройка на Snort/Suricata вместо покупка на скъп IDS, или използване на ELK Stack вместо лиценз за Splunk. В много случаи този модел работи добре в малък мащаб, но с нарастване на организацията разходите за време и сложността може да надхвърлят спестените средства.

- **Облачни услуги и гъвкави модели:** Преминаването към облачни решения за сигурност (Security-as-a-Service) позволява избягване на значими капиталови разходи (CAPEX) и превръщането им в оперативни (OPEX), плаща се месечна такса според ползването. Това може да е по-достъпно за малки организации, тъй като не се налага да купуват скъпо оборудване. Например, вместо собствен център за мониторинг, абонамент за SOCaaS или MDR служба; или вместо закупуване на множество устройства, използване на облачна платформа за защита (като облачен уеб шлюз, облачен CASB и т.н.). Тези модели често са скалируеми, като една фирма може да започне с малка услуга и да увеличава обхвата си с времето, което спомага за оптимално използване на бюджета.
- **Цена на инцидентите:** При обосновката на разходи за киберсигурност, често се сравнява цената на превенцията със средната цена на един пробив. Проучвания показват, че пробив с изтичане на данни или продължителен срив може да струва на компанията от стотици хиляди до милиони (във валута, репутационни щети, глоби по GDPR и т.н.). В този контекст, инвестиция дори от десетки хиляди годишно в сигурност може да се окаже икономически оправдана, ако предотврати дори един сериозен инцидент. Такива аргументи се използват, за да се убеди ръководството да отдели достатъчен бюджет. Киберсигурността е добре да се разглежда не като разход без възвръщаемост, а като застраховка срещу много по-големи загуби.

На Фиг. 15 диаграмата представя взаимодействието между технологии, хора и процеси:



Фиг. 15. Взаимодействие между технологии, хора и процеси

Влиянието на финансовите съображения често се проявява и в неидеални компромиси: например, организация да избере “по-малкото зло”, като да приеме по-висок риск в една област, защото няма средства да го покрие напълно, и да се надява компенсиращите мерки да са достатъчни. Добра практика е ограничените ресурси да се насочват първо към основните хигиенни мерки (като обновяване на софтуера, обучение на персонала, базови защити), тъй като те дават най-голям ефект за вложените средства. След това, според възможностите, да се надгражда с по-усъвършенствани технологии. Накратко, бюджетът очертава рамките на защитата, за това стратегическото му управление и ясно подреждане на приоритетите е част от успешната киберсигурност.

1.7. Практически АСПЕКТИ И ПРЕДИЗВИКАТЕЛСТВА ПРИ ВНЕДРЯВАНЕ И ЕКСПЛОАТАЦИЯ

Изборът на правилните технологии е само първата стъпка към реалното внедряване и дългосрочната поддръжка на решенията за киберсигурност крият множество практически предизвикателства. Някои от основните аспекти, които

организациите трябва да управляват, са:

- **Внедряване и интеграция:** Настройването на нова система за сигурност в съществуващата ИТ среда може да бъде сложно. Необходимо е съвместимост с наличните мрежи, приложения и процеси. Понякога внедряването изисква временни прекъсвания или промени в конфигурациите, което трябва да се планира внимателно, за да не се наруши бизнесът. Интеграцията между различни инструменти (например свързване на EDR агенти със SIEM, или на DLP с пощенски сървър) често изисква допълнителни усилия по разработка или настройка. Липсата на интеграция пък води до "силозен" ефект, т.е. отделните системи работят изолирано, което затруднява екипа по сигурност. Затова проектите за внедряване трябва да включват фаза на тестване във възможно най-близка до реалната среда, както и план за поетапно пускане (например първо пилотно внедряване в част от мрежата, после постепенно разширяване).
- **Поддръжка и актуализация:** Киберсигурността не е еднократно усилие, внедрените средства изискват постоянна поддръжка. Това означава редовно обновяване на дефиниции (подписи за IPS/антималуер, списъци с опасни домейни), пачване на самите продукти, следене на нови версии и уязвимости в тях. Освен техническите обновления, нужно е и непрекъснато пренастройване на системите спрямо променящата се среда: например добавяне на нови правила при появата на нови видове атаки, коригиране на прагове за аларми ако се получават твърде много фалшиви позитиви, изключване на правила които създават шум без да носят стойност. Поддръжката включва също наблюдение на самото здраве на инструментите, дали работят правилно, събират ли се логовете, дали сензорите са онлайн и др. При липса на

добра поддръжка, дори скъпи решения могат да станат неефективни (напр. “забравен” IPS, който не е обновяван и не разпознава новите атаки, или SIEM, който е препълнен с невалидни аларми и се игнорира от персонала).

- **Обучение и осведоменост на персонала:** Човешкият фактор е критичен за успеха на всяка технология. ИТ персоналът, който оперира средствата за сигурност, трябва да бъде добре обучен да ги използва ефективно. Това включва както първоначално обучение при внедряване (от доставчика или вътрешни експерти), така и продължаващо професионално развитие, курсове, сертификати, участие в конференции, тъй като заплахите и инструментите се развиват постоянно. Недостатъчно квалифицираният екип може да конфигурира грешно системата или да не разпознае индикациите за инцидент навреме. Освен техническия екип, обучението на всички служители е част от стратегията: без значима осведоменост по въпросите на киберсигурността, потребителите могат несъзнателно да саботират дори най-добрите технологични защити (например да кликнат на фишинг линк или да пренебрегнат политики). Затова множество организации въвеждат регулярни тренинги и симулации на социално инженерство, за да повишат общата култура на сигурност. Това обаче е предизвикателство което изисква време, подходящи програми и понякога среща съпротива или незаинтересованост.
- **Недостиг на кадри и организационна подкрепа:** В глобален мащаб има дефицит на квалифицирани специалисти по киберсигурност. За много организации, особено извън ИТ сектора, е трудно да наемат или задържат опитни кадри, които да управляват сложните

технологии. Това води до претоварване на малкия съществуващ екип по сигурност, “прегаряне” (burnout) и риск от грешки. Една технология е толкова ефективна, колкото хората и процесите около нея, за това организациите трябва да инвестират не само в продукти, но и в изграждане на екип и добре дефинирани процедури (например планирани отговори при инцидент). Също така, висшето ръководство трябва да подкрепя инициативите за сигурност, без спонсорство и разбиране на важността, проекти за подобрения могат да затънат или да не получат нужните ресурси. Нерядко най-голямото предизвикателство не е техническо, а организационно, за да се внедри култура на сигурност, в която технологиите, хората и управленските процеси са синхронизирани.

Фиг. 16 Визуализираща триадата Хора - Процеси - Технологии

В обобщение, практическата страна на киберсигурността изисква постоянни грижи и адаптация. Внедряването на даден инструмент е проект, който продължава и след техническото му инсталиране – чрез поддръжка, актуализация и подобрение, както и чрез работа с хората, които го използват. Организациите трябва да бъдат подготвени за тази дългосрочна ангажираност: киберсигурността е процес, не еднократно решение.

1.8. Изводи

В този раздел е разгледан широк спектър от средства и технологии за киберсигурност от превантивните контроли по мрежовия периметър до усъвършенстваните системи за откриване и реагиране, както и организационните модели за тяхното използване. Анализирани са различни категории инструменти (превенция, откриване, реакция, възстановяване) и са посочени примери както от комерсиалния пазар, така и от решения с отворен код. Сравнението на подходите показа, че изборът на технологии винаги е въпрос на баланс: между сигурност и цена, между удобство и гъвкавост, между иновация и доказана стабилност.

Критично погледнато, технологиите са необходими, но не и достатъчни сами по себе си. Много организации са се проваляли в защитата си не защото не притежават даден инструмент, а защото не са го конфигурирали правилно, или не са имали подготвени хора да реагират. Затова внедряването на средство за киберсигурност трябва да върви ръка за ръка с ясни процеси и обучение. Също, приоритизирането е ключово, по-добре е да се приложат адекватно няколко основни защити, отколкото да се внедрят множество системи проформа, без да се използват пълноценно.

Следователно, съвременната киберсигурност изисква многопластов подход. Няма един-единствен инструмент или решение, което да елиминира всички заплахи. Ефективната защита се постига чрез комбинация от технологии на различни нива, подкрепени от добре обучени специалисти и управленска подкрепа. Финансовите ограничения неизбежно влияят, но дори малки организации могат да повишат значително сигурността си, ако внимателно подберат и правилно внедрят подходящите за тях средства. Киберзаплахите еволюират непрестанно, съответно и използваните технологии трябва да се преразглеждат и актуализират регулярно. Организация, която успее да изгради гъвкава, балансирана и устойчива стратегия за киберсигурност, съчетавайки хора, процеси и технологии, ще бъде по-добре подготвена да посрещне предизвикателствата на цифровата ера.

2. ТЕСТВАНЕ НА КИБЕРСИГУРНОСТТА

Наличието на множество защитни решения не гарантира автоматично неприкосновеност, от решаващо значение е те да бъдат редовно тествани и проверявани. Тестването на киберсигурността представлява съвкупност от подходи, техники и инструменти, чрез които се оценява ефективността на защитите и се откриват потенциални уязвимости, преди злоумишленици да се възползват от тях. Използват се различни сценарии, от тестване с пълно познаване на системата до напълно „сляпо“ атакиране отвън, за да се постигне цялостна представа за сигурността.

2.1. Подходи НА ТЕСТВАНЕ: БЯЛА, ЧЕРНА И СИВА КУТИЯ

„Бяла кутия“ (White Box) тестване. При този подход екипът за тестване разполага с пълна информация за системата, мрежова архитектура, сорсов код, конфигурации, акаунти и др. Целта е проверка „отвътре“: тестерите внимателно анализират логиката на приложенията, потока на данните и обработката на грешки, за да установят дали има отклонения от замисления дизайн или пропуски в кода. White box тестването често се асоциира с кодови ревюта и анализи на архитектурата, комбинирани с изпълнение на тестови случаи. Предимство е, че може да открие дълбоки логически грешки или тайни уязвимости, които биха останали скрити при едно външно black box тестване. Например, при white box тест на уеб приложение, екипът може да прегледа кода за автентикация и да намери уязвимост в начина на валидиране на сесиите. Този подход симулира ситуация, в която нападателят има вътрешни познания или достъп (например недоволен служител

или напреднал хакер, успял да придобие някаква вътрешна информация).

„Черна кутия“ (Black Box) тестване. Това е противоположният метод, като тестерите действат без никаква предварителна информация за системата, както би било при реален външен нападател. Те виждат тествания обект единствено откъм външния му интерфейс (например публично достъпен уебсайт, IP адрес на сървър и т.н.) и трябва сами да открият входните точки, уязвимостите и евентуално да ги експлоатират. Black box тестът емулира реална атака: започва се с разузнаване (reconnaissance), сканиране на мрежата за отворени портове, установяване на версии на софтуер, събиране на информация от публични източници и после се пристъпва към опити за проникване, без да се знае какво има зад фасадата. Това позволява да се види доколко системата е устойчива срещу външен нападател, който няма вътрешна информация. Black box подходът е добър за разкриване на пропуски в конфигурацията, неизвестни на екипа уязвимости на използвания софтуер и др. Например, чрез black box тест може да се открие отворен порт с услуга, която администраторите са забравили да обезопасят, или слаба парола на публичен акаунт.

„Сива кутия“ (Grey Box) тестване. Както подсказва името, това е комбинация от бял и черен кутия подходи. Тестерът получава ограничена информация, не пълно знание като при бял кутия, но и не нулево, както при черен. Например, може да му бъдат дадени потребителски учетни данни с базови права, или общ преглед на мрежовата архитектура, но не и достъп до кода. Целта е да се имитира ситуация, при която нападателят вече е постигнал начален достъп или има частична вътрешна информация (напр. чрез стандартен потребителски профил). Grey box тестовете са много полезни, защото от една страна спестяват време (някои базови познания са налице, което ускорява откриването на проблеми), а от друга проверяват сигурността от

гледна точка на един привилегирован, но не всесилен нападател. Например, тестер в режим grey box може да използва предоставените му ограничени права до системите, за да провери докъде може да стигне и дали един компрометиран потребител може да ескалира привилегиите си, колко щети би могъл да нанесе, до какви данни би имал достъп. Този вид тест комбинира предимствата на другите два, като е по-насочен е от black box (благодарение на начален достъп) и същевременно запазва елемента на неизвестност, защото тестерът все още трябва сам да открива допълнителни уязвимости.

Бяла Кутия	**Черна Кутия**	**Сива Кутия**
		
• Достъп до сорс код • Информация за частната инфраструктура • Достъп до мрежи и ресурси	• Гледна точка от атакуващ организацията • Липса на вътрешна информация • Разузнаване от отворени източници и кибер разузнаване	• Комбиниран подход • Ограничена информация за системата

Фиг. 17. Подходи за тестване: Бяла, Черна и Сива кутия

2.2. Методи И ТЕХНИКИ ЗА ТЕСТВАНЕ

Освен по степен на познаване на системата, тестването се класифицира и по използваните конкретни методологии и инструменти. В следващите параграфи описваме някои от най-разпространените видове проверки и тестове, които се прилагат в практическия пентестинг и одитиране на сигурността.

Мрежово сканиране. Това е основна стъпка и техника, прилагана както от администратори, така и от нападатели. Целта на мрежовото сканиране е да се идентифицират активните устройства и услуги в една мрежа. С други думи, чрез сканиране се получава „снимка“ на мрежата: кои IP адреси отговарят, какви портове са отворени на тях, какви операционни системи и приложения работят. Администраторите използват тази информация, за да открият евентуални неоторизирани устройства или забравени услуги, както и да сравнят реалната картина с очакваната. Атакуващите ползват същите техники (например с инструменти като Nmap) за картографиране на целта, като първо намират всички входни точки, после се съсредоточават върху потенциално уязвимите. Един типичен процес: нападател узнава IP диапазона на фирмата (чрез публична информация), сканира го, намира че на един адрес има отворени портове 80 и 443 (уеб), на друг порт 22 (SSH) и 3389 (RDP) и т.н. Тази информация после насочва кои експлойти да се пробват. От гледна точка на защитата, редовното контролирано сканиране (например вътрешно от екипа по сигурност) позволява предварително да се види какво би забелязал един атакуващ и да се запушат ненужните точки, преди да са станали уязвимости. Мрежовото сканиране е по-скоро предпоставка за останалите тестове, но е толкова фундаментално, че често се разглежда самостоятелно като метод.

Сканиране за уязвимости. Докато мрежовото сканиране показва „какво е там“, vulnerability scanning отива една стъпка по-напред и проверява за известни уязвимости в намерените

устройства и приложения. За целта се използват автоматизирани скенери (като Nessus, OpenVAS, QualysGuard и др.), които имат в базите си данни хиляди CVE (Common Vulnerabilities and Exposures) и техните характерни признаци. Скенерът идентифицира операционната система или софтуера (напр. уеб сървър Apache, версия еди-коя си) и го сравнява с известните уязвимости за тази версия. Ако версията е остаряла и фигурира в базата с публични слабости, инструментът ще докладва кои уязвимости са приложими. Освен по версии, много скенери действително изпробват леки тестове за уязвимост, като например ще пратят специално формирана заявка, за да видят дали сървърът реагира по начин, указващ SQL инжекция или препълване на буфер. Сканирането за уязвимости е ценен подход, защото бързо дава широка картина на слабите места, особено в големи инфраструктури. То обаче има ограничения, като обикновено открива само известни проблеми (нулев ден и логически грешки могат да убягнат) и може да произведе false positives, които да се нуждаят от допълнителна проверка. Независимо от това, като част от програмата за тестване, регулярното сканиране (често на месечна или тримесечна база) помага на организациите проактивно да адресират уязвимости (чрез пачове, конфигурационни промени) преди те да бъдат експлоатирани от атакуващи.

Разбиване на пароли (Password Cracking). Атаките, целящи отгатване или превземане на пароли, са сред най-старите, но все още ефективни методи за проникване. Тестерите по сигурността често включват в обхвата си опити за разбиване на пароли, за да оценят доколко сигурни или слаби са паролите, използвани в организацията. Съществуват два основни вида атаки за разбиване на пароли: речникова атака, при която се пробват пароли от предварително съставен списък (речник) на често използвани или изтекли пароли; и атака с пълно изчерпване (brute force), при която автоматизирано се генерират всички възможни комбинации символи до определена дължина,

докато се открие правилната парола. Например, ако се тества сигурността на паролите на служителите, може да се вземат хешовете на паролите (при спазване на правила за етика и законност) и да се пуснат през инструмент като Hashcat с речник, съдържащ милиони известни слаби пароли. Ако за кратко време се открият съвпадения (разбити пароли), това индикира, че потребителите използват недостатъчно надеждни пароли. Разбиването на пароли е не само нападателна техника, но и защитна проверка, като администраторите също могат да го ползват, за да идентифицират потребители с лесни за налучкване пароли и да наложат смяната им. В едно penetration test за клиента често се демонстрира как само с един компрометиран акаунт (с отгатната парола) може да се направи значителна атака, за да се подчертае важността на силните пароли и допълнителни механизми като многофакторна автентикация.

Етично хакерство. Терминът етично хакерство обобщава авторизираните опити за проникване в система, с цел разкриване на уязвимости и тяхното отстраняване, а не злоупотреба. Етичните хакери (наричани още white hats) използват същите техники и мислене като злонамерените хакери, но работят по заявка на организацията и при ясно определени граници (scope) и правила. Обикновено те подписват споразумения за неразкриване и детайлно се договаря кои системи могат да тестват, в какъв времеви прозорец, какви атаки са позволени (например може да е забранено генерирането на DDoS срещу производствени системи по време на теста, за да не се нарушава бизнесът). Етичното хакерство е концепция, която включва в себе си и другите посочени методи, един етичен хакер по същество провежда penetration test, използвайки комбинация от сканиране, социално инженерство, експлоатиране на уязвимости и т.н., но винаги спазвайки принципа да не причинява реална вреда. Целта му е да докладва всички намерени слабости на възложителя, заедно със стъпките как са били открити, за да могат да бъдат адресирани. Важен аспект са правилата на

ангажимента (rules of engagement) те гарантират, че етичният хакер няма да надхвърли позволеното. Например, може да му се разреши да се опитва да прониква в определена уеб система, но не и в мрежата на партньор през VPN, дори ако намери пароли за него. Етичните хакери често са сертифицирани професионалисти и спазват етичен кодекс. Тяхната работа е ценна, защото разкрива реалния облик на сигурността на една организация и показва какво би могъл да направи един истински нападател със същите средства, без обаче да се чака този нападател да се появи с лоши намерения.

Проникващо тестване (Penetration Testing). Това понятие обикновено се използва като обобщение на практиката на етичното хакерство, т.е. комплексен тест, при който експерти симулират разнообразни атаки срещу системите на организацията, опитвайки се реално да „проникнат“ през защитите. Penetration test-ът (пентест) има за цел не само да открие уязвимости, но и да демонстрира докъде може да се стигне чрез тяхното експлоатиране. Например, ако по време на теста се открие SQL инжекция в уеб приложение, екипът ще се опита да я използва, за да извлече от базата данни чувствителна информация или да получи администраторски достъп, и след това документира целия вектор на атака. Пентестът е един от най-практичните и реалистични начини за оценка на сигурността, тъй като обхваща целия процес от откриване на слабост, през компрометиране, до оценка на въздействието. В края на такива тестове обикновено се изготвя подробен доклад (Pentest Report), в който са описани намерените уязвимости, нивото на риск, което носят, и препоръки за отстраняването им. Добра практика е след пентеста организацията да предприеме незабавно стъпки за затваряне на установените пробойни, а впоследствие е добре да се планира редовно повторение (например ежегодно или при значими промени в инфраструктурата). По този начин се осигурява непрекъснат подобряване на защитата. В наши дни има и модели като PTaaS (Penetration Testing as a Service)

продължително, повтарящо се тестване, комбинирано с платформа за управление на откритите уязвимости, което прави процеса още по-интегриран с жизнения цикъл на сигурността.

3. ОДИТИ ПО КИБЕРСИГУРНОСТ

Докато тестването на сигурността често е техничен процес, ръководен от етични хакери или инженери, одитът по киберсигурност е по-формализирана проверка, която има за цел да оцени съответствието на организацията с определени критерии и стандарти, както и да даде независима експертна гледна точка за цялостното ѝ състояние на сигурност. Одитите обхващат технически контроли, процеси, политики, документация и хора. Те приключват с изготвяне на официален доклад с констатации и препоръки. В тази глава ще разгледаме защо одитите са важни, как се планират и провеждат, какво съдържат резултатите и как трябва да се използват на практика.

3.1. Цели и обхват на одита.

Целите на един одит по киберсигурност могат да бъдат разнообразни, но основната е да се извлече обективна представа за силните и слабите страни на защитата на организацията и да се провери спазването на определени изисквания. Много често одитите се правят, защото са задължителни по закон или стандарт като например, регулации като HIPAA (за здравни данни) или PCI DSS (за разплащателни карти) изискват регулярни одити като условие за съответствие. В други случаи, одит може да се проведе по инициатива на самата организация или нейни партньори, за да се провери нивото на сигурност е например, компания може да одитира доставчиците си, за да се увери, че те защитават адекватно споделените данни.

Обхватът на одита трябва ясно да се дефинира още в началото. Това включва кои системи, отдели, процеси и период от време ще бъдат предмет на проверка. Одитът може да бъде широк (на цялата организация) или целеви (напр. само на облачната инфраструктура, или само на политиките и процедурите). Важен елемент е изборът на критерии/референтна рамка за оценка. Обикновено това са популярни стандарти за киберсигурност или комбинация от тях като например NIST Cybersecurity Framework, ISO/IEC 27001, COBIT, CIS Controls и др. Често одиторите ползват checklist-и на база тези стандарти, за да проверят дали организацията има внедрени необходимите контроли и практики. Например, ако се използва рамката на NIST, одиторът ще иска да види доказателства за мерки в петте области Identify, Protect, Detect, Respond, Recover. Ако е ISO 27001, ще провери съответствието с изискванията на Система за управление на информационната сигурност (ISMS). Изборът на правилен стандарт зависи от сектора и рисковете, финансовите институции може да подлежат на специфични регулатори, доставчиците на облачни услуги на FedRAMP (за работа с правителството) и т.н.

Подготовка за одит. Организацията, подлежаща на одит, трябва да се подготви, като събере релевантната документация (политики, процедури, предишни одитни доклади) и информира персонала си, че ще се провеждат интервюта и проверки. Одиторите обикновено искат предварително списък на системите, мрежова диаграма, списък на критичните активи и др., за да се запознаят със средата. В някои случаи, преди същинския одит има фаза на самооценка, където организацията сама попълва въпросници или прави вътрешен преглед спрямо критериите, което помага да се идентифицират очевидни пропуски и да се коригират (ако е възможно) преди одиторите да дойдат. Добре планираният одит ще има точно определени дати за посещения, както и ангажирани съответните ключови лица (например ИТ мениджъра, мениджъра по сигурността, отговорника по съответствията и др.).

Трета страна или вътрешен одит? Одитите по киберсигурност могат да се извършат от независима трета страна (външни одитори или консултанти) или от вътрешен одит отдел на самата организация (ако е достатъчно голяма). Външният одит е ценен с независимостта и опита си от различни организации, а също понякога е изискване напр. PCI DSS задължава годишен одит от сертифициран външен оценител. Вътрешните одити, от друга страна, се провеждат от служители (обикновено звено, различно от ИТ, например Одит отдел подчинен директно на борда), което осигурява независимост вътре в структурата. Много организации комбинират и двата подхода, като правят вътрешни одити за собствена увереност и подобрение, а на интервали канят и външни експерти за допълнителна проверка.

3.2. Провеждане на одит – планиране, проверки и събиране на доказателства

Процесът на одита по киберсигурност включва няколко основни етапа: планиране, изпълнение (получаване на доказателства чрез различни методи) и анализ/докладване. Вече споменахме планирането (определяне на обхват и критерии, събиране на предварителна информация). Същинската изпълнителна фаза обикновено протича чрез комбинация от интервюта, преглед на документи и записи, и тестване на контроли.

Одиторите започват с **интервюта с ключов персонал** например да се разговарят със CISO (мениджъра по сигурността) относно рисковете и стратегиите; с ИТ администратори относно внедрените мерки; с ръководители на звена, за да разберат бизнес процесите и как те се обезпечават откъм сигурност. Целта е да се получи ясно разбиране как организацията възприема и управлява риска, както и да се събере контекст за оценка на адекватността на контрола. Например, одиторът може да попита: „Как предоставяте достъп на нов служител до системите и как го отнемате при напускане?“ това покрива контролите за управление на достъпа, и очакваният отговор би трябвало да

включва някаква процедура или система. Одиторите често търсят несъответствия между казаното от интервюираните и наличната документация или реалността, за да открият пропуски.

Следва **прегледът на документация и политики**. Одиторът изисква и чете различни документи: политики за пароли, политика за информационна сигурност, план за реакция при инциденти, мрежови схеми, протоколи от предишни учения, обучения по сигурността, резултати от вътрешни сканирания и тестове и т.н. Това служи за проверка, че организационната рамка е налице и формално добре дефинирана. Но важен принцип е „кажи-ми покажи-ми“, не стига документът да съществува, трябва да се види и дали се спазва на практика. Например, ако има политика, че всички критични сървъри получават ъпдейти в срок до 30 дни от излизането им, одиторът би искал да види регистри/отчети, че това действително се прави, или да интервюира администратора за процеса на обновяване.

Затова ключова част е **събирането на доказателства** чрез тестване на контроли. Одиторите ще поискат достъп (наблюдателен) до определени системи или извлечения от тях, за да потвърдят, че контролите работят. Например: могат да прегледат списъци с потребители и техните права, за да видят дали има бивши служители, чийто достъп не е спрян (което би било пропуск). Могат да проверят конфигурацията на firewall-а, за да се уверят, че правилата съответстват на политиките. Често те изпробват на случаен принцип определени контроли като например, одиторът може да поиска да бъде демонстрирано как се създава архивно копие на данни и как се възстановява, за да види дали този процес реално работи. Или да провери дали системните журнали от критичните сървъри се събират в SIEM и колко време се пазят (ако политиката изисква 1 година запазване).

Друг аспект е **техническото сканиране** и тестване като част от одита. Някои одити включват елементи на уязвимост

сканиране или дори контролен пентест, макар и не в дълбочина. Например, одиторът може сам или чрез технически екип да пусне сканиране за уязвимости върху няколко важни сървъра, за да валидира твърденията на организацията, че „всички системи са напълно патчнати“. Ако сканирането открие критични остарели уязвимости, това ще бъде отчетено като несъответствие, значи контролът по управление на обновления не е ефективен. Друга стандартна проверка е тест за социално инженерство и някои одити по сигурността включват опити от страна на одитора да изпрати фишинг имейл до малък брой служители (с предварително съгласуване) или да позвъни по телефона, представяйки се за ИТ поддръжка, и да поиска парола. Целта е да се провери доколко обучението по сигурност е дало резултат и дали служителите спазват политиката да не разгласяват пароли. Ако някой се поддаде, това отново влиза в констатациите.

През цялото време на одита е важно добро сътрудничество между одитори и одитирани. Одиторите трябва да събират и съхраняват всички доказателства (документи, скрийншоти, конфигурации, интервю записки) систематично, защото върху тях ще базират доклада си. Важно правило е „един въпрос - един и същ отговор“. Ако различни интервюирани дадат различна информация за един и същ процес, това е индикация за потенциален проблем. Също, ако политика казва едно, а практиката друго, това е пропуск.

3.3. Доклад от одита - КОНСТАТАЦИИ И ПРЕПОРЪКИ. ПОСЛЕДВАЩИ ДЕЙСТВИЯ И ПОДОБРЕНИЯ

Резултатът от одита се оформя в подробен одитен доклад. Одиторският доклад започва с обобщение (executive summary) за ръководството, което очертава основните открития и общото ниво на киберсигурност като например, дали е задоволително, има ли сериозни рискове, съответства ли организацията на изискванията или не. След това за всяка област от обхвата се

описват констатациите: това включва какви контроли са проверени, какво е открито (какви слабости или несъответствия) и оценка на риска или важността на всяка констатация. Обичайно констатациите се класифицират по ниво напр. нисък, среден, висок риск, или minor, major non-compliance спрямо стандарта.

Наред с констатациите, докладът съдържа и препоръки за всяка слабост и конкретни стъпки, които одиторите препоръчват на организацията да предприеме за подобрение. Например, ако констатацията е „Не се провежда обучение по информационна сигурност за новите служители“, препоръката ще е „Да се внедри задължително обучение по сигурност при постъпване на работа и ежегодни опреснителни обучения за целия персонал“. Ако констатацията е техническа напр. „Неприлагани актуализации на уеб сървъра в DMZ“, препоръката ще бъде „Да се приложат последните сигурностни актуализации на уеб сървъра и да се внедри процес за редовно пачване“. По този начин докладът дава на ръководството „план за действие“ което е списък от задачи за подобряване на защитата.

Важно е да се разбере, че одитният доклад не е самоцел и неговата стойност е в следващите стъпки. Една честа грешка е организацията да разгледа доклада, да го използва за непосредствени нужди (например, да го представи на регулатор или като отговор на изискване), но да не изпълни препоръките. Това обезсмисля целия процес. Вместо това, добро управление на сигурността изисква одитните находки да бъдат превърнати в конкретни задачи и проекти за адресиране. Може да се състави план за подобрение, където за всяка препоръка се определят отговорник, срок и стъпки за изпълнение. Например, ако има 10 констатации, организацията може да планира в следващите 6 месеца кои от тях ще отстрани, с приоритет на най-високорисковите.

Следенето на изпълнението на препоръките е отговорност на ръководството, обикновено CISO или мениджъра по

съответствие трябва да отчита прогреса пред висшето ръководство. Наличието на система за тикети или тракинг е полезно, като така всяка препоръка става тикет, който се затваря, когато работата е свършена, и се вижда статусът. Ако сроковете се изпускат, въпросът трябва да се ескалира на по-високо ниво, защото оставените без внимание констатации представляват известни рискове. Много организации интегрират тези задачи в общия си процес за управление на риска, т.е. третират всяка одитна слабост като риск, който трябва да бъде намален.

Последващи одити и непрекъснато усъвършенстване. Одитът не е еднократен, той е част от един цикъл на непрекъснато подобрение. След като препоръките се изпълнят, често се планира follow-up одит или поне преглед, за да се потвърди отстраняването на проблемите. Например, ако в доклада пише, че 5 от 10-те препоръки са критични, може след 6 месеца одиторите да проверят само тях, за да валидират напредъка. В следващия редовен одит (например догодина) ще се види дали миналогодишните констатации са били решени; ако не, те може да ескалират като по-сериозен проблем.

В по-широк план, ефективният одит допринася към цялостното управление на риска и съответствието. Той не бива да се възприема като контролен механизъм само за регулаторни цели, а като ценен инструмент, който помага на организацията да изгради по-зряла система за сигурност. Одитите осветляват слабите места в киберпространството, за да може организацията да постигне повишена кибер устойчивост. Но за да има полза, трябва организационна култура, която взема присърце обратната връзка. Това включва подкрепа от висшето ръководство, бордът и изпълнителното ниво трябва да са ангажирани да осигурят ресурси и приоритизация на инициативите по сигурността, произтичащи от одита. Само тогава одитът няма да бъде „упражнение на хартия“, а ще доведе до реално повишаване на

нивото на защита и способността за справяне с днешните сложни заплахи.

Иначе казано, одитът не е финал, а начало на процес по усъвършенстване. Той трябва да бъде последван от конкретни действия, наблюдение на напредъка и повторни проверки, вписани в един по-широк цикъл Plan-Do-Check-Act в управлението на информационната сигурност. Само така една организация може да гарантира, че не просто „покрива изисквания“, а действително изгражда едно гъвкаво и устойчиво на инциденти функциониране. В противен случай, одитът се свежда до формалност и отметка, без реален принос към сигурността, което особено в днешната динамична заплашителна среда, вече не е приемлив подход.

4. ЦИФРОВА КРИМИНАЛИСТИКА И РЕАГИРАНЕ ПРИ КИБЕРИНЦИДЕНТИ

Цифровата криминалистика се определя като процес на съхраняване, идентифициране, извличане и документиране на компютърни доказателства, които могат да бъдат използвани от съда. Казано с други думи, тя представлява наука за откриване и анализ на дигитални следи в електронни носители напр. компютри, мобилни телефони, сървъри, мрежови устройства и др. Цифровата криминалистика е критично направление както в правоохранителната дейност, така и в корпоративната сигурност, защото предоставя методи и инструменти за разкриване на факти и обстоятелства около инциденти в дигиталната среда. В съвременния свят практически няма сфера, която да не е засегната от цифровите данни, от наказателното право през трудови спорове до търговски дела, навсякъде могат да присъстват електронни доказателства. С повсеместното използване на интернет и ИТ технологии цифровите доказателства придобиват все по-голяма значимост и престъпления като кражби, изнудване или измами все по-често включват дигитална компонента. Поради това уменията за събиране и анализ на такива доказателства са ключови за разследващите екипи и специалистите по киберсигурност.

Правен контекст: За да бъдат годни за използване при разследване или в съда, цифровите доказателства трябва да бъдат събирани и съхранявани при спазване на законовите изисквания и с гаранция за тяхната цялост. Българският Наказателно-процесуален кодекс (НПК) дефинира доказателствата като факти от обективната действителност, индивидуализирани по време и място, независимо дали са материални обекти, документи, действия или бездействия. В този смисъл логовете, файловете, електронната кореспонденция и

други дигитални следи могат да се приемат за доказателства, ако са събрани по предвидения в закона ред и са годни да възпроизведат фактите в процеса. На практика обаче законодателството невинаги е детайлизирало процедурите за работа с електронни доказателства, тъй като много правни норми са приети преди ерата на съвременните технологии. Това поражда предизвикателства пред съдилищата при преценка на допустимостта и достоверността на цифровите доказателства. Ето защо в цифровата криминалистика се прилагат стриктни протоколи и стандарти, които да гарантират, че събраните данни не са изменени или компрометирани по време на обработка. Цифровите доказателства са крехки и краткотрайни и неправилното боравене с тях може лесно да доведе до промяната или унищожаването им. Поради тази нестабилност се изисква спазване на т.нар. верига на попечителство (chain of custody) за непрекъсната документирана проследимост на всяко действие с доказателството от изземването до представянето му. Това осигурява доказване на автентичността и целостта на електронните доказателства в съдебна или вътрешна проверка. Спазването на утвърдени стандарти и добри практики в областта (напр. ISO 27035, насоките на NIST и др., разгледани по-долу) осигурява структуpiran и законосъобразен подход към разследването на инциденти и повишава шансовете цифровите доказателства да бъдат признати за надеждни.

Цифровата криминалистика служи за установяване на обективната истина при инциденти в киберпространството, като позволява на експертите да извлекат максимално информация от наличните дигитални артефакти. Тя винаги върви ръка за ръка с процедурите по реагиране при киберинциденти, тъй като ефективното овладяване на един инцидент изисква не само спиране на вредоносните действия, но и разследване на причините, извършителите и засегнатите ресурси. По-нататък в тази глава разглеждаме типичните киберинциденти, етапите на разследването им, организационните аспекти на реагирането,

утвърдените методологии, както и практически техники и инструменти за цифрова криминалистика. Накрая са обсъдени примери от практиката, съвети и препоръки за успешна работа в реална среда.

4.1. Типови КИБЕРИНЦИДЕНТИ

Киберинцидент може да се определи като събитие или серия от събития, които застрашават конфиденциалността, целостта или наличността на информационните ресурси. Инцидентите могат да бъдат резултат от злонамерени атаки, вътрешни злоупотреби или неволни грешки. По-долу са описани някои от най-често срещаните типови киберинциденти:

- Пробиви в сигурността (security breaches): Това са инциденти, при които неоторизирани лица успяват да преодолеят защитите на система или мрежа. Пробивите най-често стават чрез експлоатация на уязвимости, кражба на удостоверения за достъп (пароли, ключове) или социално инженерство. В резултат може да се получи кражба на чувствителни данни, несанкционирано изменение на информация или поемане на контрол над системи. Например, пробив в база данни на компания може да доведе до изтичане на лични данни на клиенти. Разследването на такива инциденти цели да установи входната точка на атаката, използваните методи и обхвата на компрометиране.
- Вътрешни злоупотреби (insider threats): В тези случаи заплахата произтича отвътре като настоящ или бивш служител, контрактор или друг вътрешен потребител с достъп до системите извършва неправомерни действия. Примери са: кражба на фирмени данни от служител с цел предоставянето им на конкурент; злоупотреба с права от системен администратор (напр. неоторизирано разглеждане на чужди данни); умишлено повреждане

или изтриване на данни. Вътрешните злоупотреби са особено опасни, тъй като вътрешните лица често имат легитимен достъп и познават системите, което затруднява откриването им. Разследването тук се фокусира върху логовете на дейност на потребителите, версиите на файловете, комуникациите по имейл и др., за да се съберат доказателства за злонамерените действия.

- Атаки тип ransomware (изнудващ софтуер): Това са инциденти, при които злонамерен софтуер (вирус-рансъмуер) прониква в системата, шифрова файловете и изисква откуп (ransom) за тяхното разшифроване. Ransomware атаките през последните години са едни от най-разпространените киберзаплахи, засягащи както големи организации, така и обикновени потребители. Типично развитие на такъв инцидент: потребител отваря заразено прикачено писмо или уязвим сървър бива компрометиран, след което вредоносният код криптира масово данни и оставя съобщение с искане на откуп. Все по-често нападателите комбинират шифроването с кражба на данни (double extortion) и заплашват жертвата, че ще публикуват откраднатата информация, ако не плати. Реагирането при ransomware включва бързо изолиране на засегнатите машини, форензик анализ на зловредния код, за да се намери начин за възстановяване (напр. чрез бекъпи) и подобрение на защитите. Цифровата криминалистика в такива случаи се стреми да установи как е проникнал рансъмуерът, кои системи е засегнал, дали има ексфилтрация на данни и евентуално да извлече криптографски ключове или индикатори за компрометиране.
- Изтичане на данни (data leakage): Изтичането или нерегламентираното разкриване на данни е инцидент, при който конфиденциална информация става достъпна

за неоторизирани страни. Това може да стане по множество начини, чрез външен пробив (хакер публикува откраднатa база данни), чрез вътрешна злоупотреба (служител копира данни на външен носител или облак), или по невнимание (погрешно конфигуриран сървър оставя данните отворени в интернет). Последствията от изтичане на данни включват накърняване на личната неприкосновеност (при изтичане на лични данни), финансови загуби, репутационни щети и глоби по регулации като GDPR. Разследването на такива инциденти цели да проследи откъде и как са изтекли данните – например чрез анализ на логове за големи обеми изтеглена информация, чрез проверка на устройства за съхранение, иззети от заподозрян служител, или чрез търсене в мрежата (dark web) за копия на изтеклите данни. Важен аспект е и форензик анализа на файловете, проверка на метаданните може да покаже кой и кога е достъпвал или експортвал информацията.

Разбира се, съществуват и други видове киберинциденти, например DDoS атаки (разпределени атаки за отказ на услуга), компрометиране на електронна поща (Business Email Compromise), измами в онлайн плащания, атаки към веригата на доставките и др. Горните примери обаче обхващат едни от най-честите сценарии, при които се налага прилагането на цифрова криминалистика и организирано реагиране. В следващите раздели ще разгледаме как протича един процес на разследване на инцидент и кои са добрите практики за ефективна реакция.

4.2. Процес ПО РАЗСЛЕДВАНЕ НА ИНЦИДЕНТ (ФАЗИ)

Разследването на киберинцидент обикновено преминава през няколко основни фази. Моделите в литературата леко се различават в имената и броя на етапите, но целта е една и съща,

това е структуриран подход за да се гарантира пълнота и надеждност на анализа. Тук ще опишем пет ключови фази на процеса: идентификация, събиране, анализ, докладване и възстановяване. В реални условия тези фази може да се припокриват или повтарят итеративно, но разделянето им помага за методичното им изпълнение.

4.2.1. Идентификация НА ИНЦИДЕНТА

Първата стъпка е разпознаването и потвърждението, че има инцидент, който изисква реагиране. Идентификацията включва мониторинг и откриване на аномалии, сигнали или аларми, подсказващи нарушение на сигурността. Това може да стане чрез различни източници: системни логове, IDS/IPS системи за детекция на прониквания, антивирусен софтуер, аларми от SIEM платформи, доклади от потребители за необичайно поведение на системата и др. В тази фаза е от решаващо значение да се разграничат истинските инциденти от фалшивите аларми. Анализаторите провеждат първоначална оценка на събитията, напр. преглеждат логове от последните часове за индикатори на компрометиране (Indicators of Compromise), проверяват хеш суми на подозрителни файлове, извършват бърз одит на засегнатите системи. Целта е да се отговори на въпросите: Какво се е случило? Кои системи или данни са засегнати? Това инцидент ли е, изискващ по-нататъшно разследване? При потвърждение, че има киберинцидент, се преминава към организиране на официално разследване и активиране на план за реагиране.

4.2.2. Събиране НА ЦИФРОВИ ДОКАЗАТЕЛСТВА

След като инцидентът е идентифициран и потвърден, следва фазата на събиране на всички относими данни и артефакти, които могат да послужат като доказателства. Тази фаза често се нарича

още придобиване (acquisition) на доказателства. В нея се включва изземването или копирането на информация от засегнатите системи по начин, запазващ нейната цялост. Добра практика е първо да се събират летливи данни (volatile data) информация, която би се изгубила при изключване на машината. Това са например: съдържанието на оперативната памет (RAM), списъкът на активните мрежови връзки и отворени портове, активните процеси, потребители в системата в момента, съдържание на кеш памети, временни файлове и др. Тези данни се събират чрез инструменти за memory dump (например dd или специализирани програми за Windows) и команди като netstat, tasklist и др., като всичко се пренасочва към външен носител.

След летливите данни се преминава към постоянните данни: пълни образи (images) на твърди дискове или други носители, копия на лог файлове, конфигурации, бази данни, документи и всичко, което може да има отношение към инцидента. Изготвянето на криминалистичен образ на дискове се прави бит-по-бит (напр. с инструменти като dcfldd или Guymager) и обикновено се изчисляват хеш суми (SHA-256, MD5) преди и след копирането, за да се удостовери, че копието е идентично с оригинала. Оригиналните носители след това се запечатват и съхраняват безопасно, а всички анализи се извършват върху копията, за да не се допуска промяна на оригиналните данни. При събирането на данни се документира подробно кой, кога, къде и как иззема всеки артефакт, това е част от веригата на попечителство. В някои случаи събирането трябва да стане бързо и на място (on-site), например при живи системи, които не могат да се изключат. Тогава се използват форензик инструменти на място (като специализирани Live CD/USB с цифрово-следствени дистрибуции като CAINE, Kali Linux в режим Live и др.), които позволяват копиране на данни без да се нарушава (доколкото е възможно) състоянието на системата. При мрежови инциденти може да се включи и прехващане на мрежов трафик (packet capture) за определен период, за да се съберат пакети за анализ

във фазата на анализ. Всички тези дейности по събиране трябва да се провеждат внимателно, като всяко действие е добре да бъде повторяемо и одитируемо, така че по-късно да може да се удостовери в съда или пред мениджмънта, че данните са автентични.

4.2.3. Анализ НА СЪБРАНИТЕ ДАННИ

След като необходимите данни са събрани и обезпечени, основната работа преминава към анализ на тези данни. Анализът в цифровата криминалистика е детайлният процес на разглеждане на файлове, системни артефакти, логове, памет и други източници, с цел да се възстанови последователността на събитията и да се отговори на ключовите въпроси за инцидента: какво се е случило, как е било осъществено, кога, от кого и с какви последствия? Тази фаза често е най-времемката, тъй като включва множество аспекти:

- **Файлов анализ:** Проверяват се подозрителни файлове или програми, намерени в системите. Извършва се анализ на техните хеш стойности (сравнение с бази данни за известни доброкачествени или зловредни файлове), преглежда се структурата им с хекс редактор при нужда, търсят се скрити данни. Ако има изтрити файлове, използват се техники за възстановяване или (carving) сканиране на дисковия образ за характерни сигнатури на файлове (например начални байтове на PNG, PDF и др.), за да се извлекат остатъци от изтрити данни. Специално внимание се отделя на изпълними файлове или скриптове, които може да са малуер, при тях може да се наложи reverse engineering или динамичен анализ в изолирана среда (пясъчник) с цел да се разбере какво правят.

- **Анализ на системни артефакти:** Това включва преглед на системния регистър (Registry) при Windows, на конфигурационни файлове при Linux, списъци с инсталирани програми, планувани задачи (Scheduled Tasks/cron), записани пароли или ключове, и др. Целта е да се намерят следи от дейността на атакуващия или злонамерения потребител: например новосъздадени потребителски акаунти, модифицирани настройки на сигурността, инсталирани „backdoor“ програми за отдалечен достъп, необясними промени в конфигурации.
- **Лог анализ:** Лог файловете (журнали) са един от най-богатите източници на информация. Анализът им обхваща системни логове (напр. syslog при Linux, Windows Event Log при Windows), мрежови логове (от firewall, IDS, VPN, прокси сървъри), приложения (уеб сървъри, бази данни) и др. Чрез корелация на събитията във времето се изгражда хронология на инцидента. Например, анализът може да покаже, че в 02:13 ч. през VPN е влязъл потребител с администраторски права, в 02:15 ч. е изтеглил голям файл от сървъра, в 02:20 ч. е изтрил журналите. За подобна корелация често се използват специализирани инструменти, например log2timeline/Plaso, които от множество източници генерират обща времева линия. В процеса на лог анализа експертите търсят аномалии: неуспешни опити за логин, достъп в нетипично време, промени в системното време, съобщения за грешки или аларми от защитни системи и т.н. Това помага да се идентифицират ключови действия на нападателя или вътрешния нарушител.
- **Анализ на мрежов трафик:** Ако са налични PCAP записи (capture) на мрежовия трафик по време на инцидента или ако се е направило записване непосредствено след детектирането му, те подлежат на дълбок анализ.

Използват се инструменти като Wireshark (графичен) или TShark и tcpdump (конзолни) за филтриране и разглеждане на пакетите. Мрежовият анализ може да разкрие комуникация с външни сървъри (например извеждане на данни към IP на нападателя), управление на зловреден софтуер от дистанция (C2 трафик), сканиране на портове, необичаен протокол или голям обем криптиран трафик в необичайни часове. Чрез реконструиране на сесии от пакетите могат дори да се възстановят прехванати файлове, чат съобщения или имейли, ако не са били шифровани. Анализът на трафика допълва картината, особено при външни атаки, като показва кой с кого е комуникирал и какви данни са текли през мрежата.

- **Анализ на паметта:** В случаите на сложни инциденти (напр. подозрение за rootkit) се налага и тънък анализ на образите на памет (RAM), направени при фазата на събиране. Чрез рамкове като Volatility или Rekall се извлича информация за процесите, отворените файлове, мрежови връзки, заредени драйвери и модули в паметта. Това може да открие например зловреден процес, който не оставя файлове на диска (fileless malware), или скрито DLL инжектиране в легитимен процес. Данните от паметта често съдържат и откъси от текст или данни, които не са запазени никъде другаде, пароли в чист текст, криптографски ключове, части от документи и др.

При анализа експертите документират подробно всички значими находки. Всяка стъпка като например „намерен файл X, хеш Y, идентифициран като злоунищожаващ скрипт; от логовете се вижда, че файлът е качен от IP ...“ се записва. Това не само улеснява изготвянето на финалния доклад, но и гарантира повторяемост: друг специалист би могъл да проследи действията и да валидира заключенията. Важно е да се отбележи, че

анализът може да доведе и до разширяване на обхвата на разследването, като например откриване на нов компрометиран хост, което налага връщане към фаза събиране за този хост. Този итеративен процес продължава, докато екипът е уверен, че е събрал достатъчно данни за да разбере инцидента изчерпателно.

4.2.4. Докладване НА РЕЗУЛТАТИТЕ

След приключване на активния анализ, констатациите трябва да бъдат обобщени и представени пред съответната аудитория, било то ръководство на организацията, правоприлагащи органи или съд. Фазата на докладване (reporting) е съществена, защото добрият доклад превръща техническите детайли в разбираема и ползвана информация за вземане на решения. Един качествен форензик доклад обикновено включва:

- Резюме за ръководството: кратко описание на инцидента, какво се е случило и какви са последствията, написано на достъпен език (не-технически, доколкото е възможно).
- Обхват на разследването: кои системи, данни, период от време са обхванати; какво не е било достъпно (ако има ограничения).
- Методи на работа: описание на използваните техники и инструменти за събиране и анализ (напр. „Изготвени са образи на дисковете с tool X, хеши съвпадат; анализ на памет с Volatility; преглед на логове със Splunk“ и т.н.), за да се демонстрира надеждността на процеса.
- Хронология на събитията: таблица или описание по дати и часове на ключовите събития, които са установени (напр. „01:23 UTC - Вход в система A от потребител admin от IP ...; 01:24 - Изпълнение на команда Y; 01:30 - Настъпва срив на услугата...“).

- Детайлни находки: техническата част с доказателствата какви зловредни файлове са намерени, какви действия са извършени, показват се извадки от логове, скрийншоти или хексдати при нужда, резултати от хешове, YARA правила или антивирусни открития, и др. Тази част е ядро на доклада за специалистите.
- Заключение и оценка на въздействието: обобщава се причината за инцидента (root cause), идентифицират се евентуално извършителите или източниците (ако е възможно), и се описва какъв е ефектът какви данни са компрометирани, колко време системите са били неработоспособни, има ли риск от повторение.
- Препоръки: конкретни стъпки за подобрение, напр. пачване на уязвимости, усилване на паролната политика, обучение на персонала, внедряване на допълнителни средства за защита или мониторинг, план за възстановяване.

Докладът трябва да бъде написан безпристрастно и основан на факти, като в него не се правят необосновани заключения. Ако има хипотези, те се отбелязват като такива. Често се изготвят две версии: вътрешен подробен доклад за техническите екипи и съкратена версия (executive summary) за висшето ръководство или външни публики. При съдебно дело, експертът по цифрова криминалистика може да бъде призован да свидетелства като вещо лице, където докладът служи като основа на свидетелските му показания. Затова е критично съдържанието му да е точно и добре обосновано.

4.2.5. Възстановяване И ПОУКИ

Последната фаза, след приключване на непосредственото разследване и докладване, е възстановяване на нормалната работа и извличане на поуки (lessons learned). В контекста на

реагирането при инциденти, възстановяването включва всички действия по отстраняване на нанесените щети и превенция на бъдещи подобни инциденти. Това често означава:

- **Възстановяване на системите:** преинсталиране на компрометирани сървъри или устройства, възстановяване на данни от бекъп, декриптиране на файлове (в случай на ransomware, ако е възможно чрез ключ или бекъп), връщане към нормална работа. Преди пускане в експлоатация се уверяваме, че системите са чисти от зловреден код и задни вратички.
- **Прилагане на корекции и подобрения:** ако инцидентът е станал възможен заради определена уязвимост или слабост, тук се отстранява, например инсталиране на съответни пачове за сигурност, смяна на всички компрометирани пароли, повишаване на нивото на криптиране, спиране на незакърпени услуги и др. Това е част от ериадикация (eradication) на заплахата, гарантирайки, че нападателят вече няма достъп.
- **Мониторинг след инцидента:** известно време след възстановяването екипите наблюдават по-внимателно засегнатите системи за признаци на повторно компрометиране, тъй като нападателите понякога опитват отново или са оставили скрит малуер. Въвеждат се и допълнителни аларми или проверки.
- **Поуки и актуализиране на плана:** Много важно е екипът да проведе разбор на инцидента (post-incident review). На тази среща се обсъжда как е протекла реакцията, кои неща са се оказали ефективни и кои не, имаше ли пропуски в комуникацията или техническата координация, какво може да се направи по-добре следващия път. Отбелязват се и уроците: например „не разполагахме с достатъчно детайлен лог на VPN, трябва да увеличим нивото на логване“ или „служителите не

разпознаха фишинг имейла, следователно е нужно обучение“.

- **Обновяване на документацията:** На базата на наученото се актуализира планът за реагиране при инциденти, процедурите за екипа, ръководствата за ИТ персонала, и се въвеждат препоръките от доклада. Така всяко инцидентно разследване допринася към подобряване на цялостната киберустойчивост на организацията.

В някои класификации тази фаза се разглежда отделно като пост-инцидентна активност или уроци от инцидента. Важното е, че цикълът се затваря с възстановени системи и по-силна защита за в бъдеще.

4.3. Организация НА РЕАГИРАНЕТО ПРИ ИНЦИДЕНТИ

Ефективното справяне с киберинциденти изисква не само технически умения, но и правилна организация на екипите и процесите. В рамките на една организация трябва да съществува предварително определена структура за реагиране, често наричана Computer Security Incident Response Team (CSIRT) или просто Incident Response Team (IRT). Тук разглеждаме ролите, отговорностите, обученията и плановете, необходими за успешен отговор. Екип за реагиране (CSIRT): Обикновено това е мултидисциплинарен екип, който се активира при сериозен инцидент. В състава му могат да влизат: специалист по мрежова сигурност, системен администратор, експерт по цифрова криминалистика, представител на мениджмънта, специалист по комуникации/PR (ако инцидентът изисква публично оповестяване), юрист, отговорник по защита на данните (особено при утечки, свързани с лични данни) и др. Ясно трябва да се дефинират ролите на всеки член. Примери за роли:

- **Ръководител на инцидентния отговор (Incident Manager):** координира целия процес, разпределя задачи, следи за

спазване на процедурите и е основна точка за контакт с висшето ръководство.

- **Технически лидери/аналитици:** тези, които извършват техническото разследване, като анализират логове, имиджи, идентифицират причината. Може да има подразделение, например един фокус върху мрежата, друг върху системите, трети върху зловредния код.
- **Специалист по комуникации:** отговаря за вътрешното информиране (напр. да уведоми служителите, че има инцидент и какво да правят или да не правят) и външното, ако се наложи (прессъобщения, уведомяване на клиенти). Това е особено важно при инциденти с данни на клиенти или при които законът (напр. GDPR, NIS) изисква уведомяване на регулатор или потърпевши.
- **Правен консултант:** преценява правните аспекти, дали инцидентът изисква намесата на правоохранителни органи, как да се съблюдават изискванията за запазване на данни, какво може да се споделя публично без да се нарушат закони или да се поеме отговорност.
- **Отговорник по възстановяването (Recovery lead):** планира и ръководи действията по възстановяване на ИТ услугите, работи в тясна връзка с ИТ отделите за прилагане на корекции, рестартиране на системи, възстановяване от резервни копия.
- **Връзка с външни организации:** лице, което комуникира с външен CERT/CSIIRT (например националния CERT България), с доставчици на услуги (ако инцидентът касае техни услуги като напр. cloud провайдър), и с правоохранителните органи (ГДБОП, полиция) при необходимост. В много случаи организациите установяват контакт с външен експертен център, например ако нямат достатъчно ресурс, могат да ползват договор с Managed

Security Service Provider (MSSP) или да поискат съдействие от националния CERT.

План за реагиране при инциденти: Всеки такъв екип трябва да работи по предварително утвърден план или политика за инцидентен отговор. Този план (понякога наричан (IRP) Incident Response Plan) описва подробно стъпките, които да се следват при инцидент: от идентификация до възстановяване, кой уведомява кого, какви решения на какъв етап трябва да се вземат (напр. критерии кога да се изключи засегнат сървър от мрежата, или кога да се ескалира към по-високо ниво на управление). Планът включва и списък с контакти за спешни случаи (екип, мениджъри, външни партньори), готови шаблони за комуникация (например имейл до всички служители, че има фишинг кампания), и предварително определени сценарии и playbook-ове. Playbook представлява конкретна процедура за определен тип инцидент, например „ако се случи ransomware на работна станция, направи стъпки 1,2,3...; ако е DDoS атака срещу публичен сайт, следвай стъпки ...“.

Обучения и тестове: Организацията на отговора не се изчерпва с писането на план, необходимо е периодично да се провеждат обучения и тренировки. Екипът за реагиране трябва да е подготвен и сдобит с практически умения. Това се постига чрез:

- **Тренировъчни упражнения (drills):** Например tabletop exercise е сценарий на хартия, при който екипът дискутира как би действал при даден инцидент. Още по-добре са симулациите в реална или контролирана среда, т.нар. киберполигон (cyber range), където се разиграва симулиран инцидент върху тестова инфраструктура и екипът трябва да реагира сякаш е реално събитие. Такива упражнения разкриват слаби места и дават ценен опит.
- **Обучения и сертификации:** Членовете на екипа следва да се обучават по актуални програми, например

сертификации като Certified Incident Handler (GCIH), Certified Ethical Hacker, или курсове по цифрова криминалистика (CHFI, GCFA и др.). Това гарантира, че са запознати с последните техники и заплахи. Също така, персоналът в по-широк план (всички служители) трябва да получава обучение за киберхигиена, как да разпознава социално инженерство, към кого да се обърне при подозрителен инцидент, което помага за навременна идентификация.

- **Оценка на готовността (readiness assessments):** Редовно се правят проверки дали системите за детекция работят, дали резервните копия се пазят и могат да се възстановят, дали всеки в екипа знае ролята си. Част от това може да са одити на сигурността и проникващи тестове (pentest), които да предотвратят реални инциденти.

Накратко, организацията на отговора изисква и техническа подготовка, и административна уредба. При добре подготвена структура, когато възникне инцидент, екипът знае какво да прави, комуникира ефективно и избягва хаоса. Това значително намалява времето за реакция и щетите. Международните стандарти като ISO/IEC 27035 (Управление на инциденти) дават рамка за тези процеси, от подготовката, през откриването и реагирането, до пост-инцидентните дейности. Също така, NIST SP 800-61 предлага ръководство за изграждане на способност за реакция и подчертава важността на плановете и обученията. В следващия раздел ще разгледаме някои от тези методики и най-добри практики по-подробно.

4.4. Методики ЗА РАЗСЛЕДВАНЕ И РЕАГИРАНЕ (ДОБРИ ПРАКТИКИ, СТАНДАРТИ)

През последните две десетилетия са разработени множество методологии и стандарти, целящи да систематизират процеса на

управление на инциденти и провеждане на цифрови разследвания. Използването на утвърдени рамки помага на организациите да не „откриват топлата вода“ всеки път, а да следват доказани добри практики. Следва преглед на най-известните такива:

ISO/IEC 27035: Управление на инциденти в сигурността на информацията: Това е международен стандарт (състои се от няколко части), който описва концепциите и процесите за ефективно управление на инциденти. В него се поставя акцент върху подготовката (установяване на политика, отговорности, ресурси), идентифициране и докладване на инциденти, оценка и взимане на решения (дали едно събитие е инцидент, колко е сериозно, дали да се активира отговор), реагиране (съдържа стратегии за ограничаване, събиране на доказателства, отстраняване, възстановяване) и уроци.

ISO 27035 предоставя структурирана рамка, която може да се интегрира с цялостната система за управление на сигурността (ISO 27001). Предимството му е, че е агностичен към технологията като дава насоки на високо ниво, приложими за различни организации. Например, стандартът съветва да има определени критерии кога едно събитие да се класифицира като инцидент, да се води регистър на инцидентите, да се извършва първоначална оценка на риска от всеки инцидент и др. Следването на ISO 27035 улеснява и аудитирането показва, че организацията има зрял процес, което е важно и за съответствие с регулации.

NIST SP 800-61 (Computer Security Incident Handling Guide): Ръководство, издадено от Националния институт по стандарти и технологии (САЩ), широко използвано като де-факто стандарт особено в ИТ индустрията. NIST дефинира жизнен цикъл на реагиране с четири основни фази: Preparation (Подготовка), Detection and Analysis (Откриване и анализ), Containment, Eradication and Recovery (Ограничаване, премахване и

възстановяване), Post-Incident Activity (Дейности след инцидента). Тази циклична схема е изобразена на Фиг. 18.

Фиг. 18. Жизнен цикъл на реагиране при инциденти според NIST SP 800-61

NIST набляга на това, че подготовката (наличие на политика, екип, инструменти) е фундаментална за успеха. В ръководството се дават практически съвети, например: как да се изградят jump kit (полеви комплект) с необходимите софтуери и хардуери за разследване, как да се категоризират инциденти по ниво на сериозност, как да се комуникира по време на инцидент, как да се съхраняват доказателства. Специално внимание се отделя и на документацията по време на инцидента като поддържане на incident log на действията на екипа, което помага при анализа след това. NIST SP 800-61 също така съветва

организациите редовно да преразглеждат представянето си след инцидент и да обновяват плановете си.

Моделът включва четири основни фази (Подготовка; Откриване и анализ; Ограничаване, изчистване и възстановяване; Дейности след инцидента), които се изпълняват като непрекъснат процес на подобрение. Всяка фаза информира следващата, а научените уроци след инцидента подобряват подготовката за бъдещи инциденти.

SANS Institute предлага модел от шест стъпки: Друг популярен модел, преподаван от SANS, включва шест фази: Preparation; Identification; Containment; Eradication; Recovery; Lessons Learned. По същество той е сходен с NIST, но разделя ясно идентификацията като отделна стъпка (която при NIST е заедно с анализа) и добавя Lessons Learned като формализирана заключителна стъпка. SANS публикува и множество best practice ръководства, чеклистове и инструменти, които са полезни за екипите. Пример за най-добри практики от SANS е: винаги да има форензик копие на всички важни доказателства; да се ограничи достъпът до засегнатите системи (containment) възможно най-скоро, но без да се унищожават следи; да се комуникира „need-to-know“, т.е. информацията за инцидента да се споделя само с хората, които трябва да знаят, за да се избегнат паника или изтичане на информация към нападателя.

Други стандарти и насоки:

ISO/IEC 27037, 27041, 27042, 27043: серия стандарти, допълващи 27035, които се фокусират върху различни аспекти на цифровата криминалистика, от идентифициране, събиране и запазване на цифрови доказателства (27037) до анализ и оценка (27042) и общи принципи за разследване (27043). Те са по-специализирани и често се използват от експерти при съдебни експертизи.

ENISA насоки: Агенцията на ЕС за киберсигурност (ENISA) публикува регулярно практически ръководства, напр. за форензик в облачни среди, за цифрови доказателства, и за изграждане на CERT екипи. Тези документи често разглеждат специфични предизвикателства, като правните измерения в рамките на ЕС, трансграничното сътрудничество при инциденти, работа с доставчици на облачни услуги при разследване и др.

NIST Cybersecurity Framework (CSF): макар и фокусиран върху цялостния риск, CSF включва функция „Respond“ и „Recover“, които препращат към инцидентния отговор. Организациите често обвързват политиките си за инциденти с по-широката рамка на CSF.

FIRST Best Practices: FIRST (Forum of Incident Response and Security Teams) е международно сдружение на CERT/CSIRT екипи. Те публикуват препоръки и обменят опит. Например има ръководство за разследване на инциденти в индустриални мрежи (SCADA/ICS), което е ценно за екипи в енергетиката, производството и др.

Общото между всички тези методики е, че подчертават значението на подготвеността, структурирания подход и следването на формални процедури. Въвеждането на стандарт не гарантира, че инциденти няма да се случват, но гарантира, че когато се случат, реакцията ще бъде по-ефективна и контролирана. Добрите практики действат като контроли срещу хаоса: например, наличието на комуникационен план предотвратява риска някой от екипа да сподели чувствителна информация публично; следването на протоколи за събиране на доказателства предотвратява случайно изтриване на данни; създаването на доклад при всеки инцидент пък изгражда база знания за бъдещи случаи.

Препоръчително е всяка организация да се придържа към поне една от утвърдените рамки (ISO, NIST или друга) и да я адаптира към своя контекст. Това осигурява повторяемост,

съвместимост (с други екипи и регулатори) и качество на процеса на реагиране и криминалистично изследване.

4.5. Инструменти И ТЕХНИКИ В ЦИФРОВАТА КРИМИНАЛИСТИКА

За успешно провеждане на цифрово разследване, специалистите разчитат на разнообразен набор от инструменти и технически похвати. В този раздел разглеждаме ключови техники и примери за инструменти, използвани при анализ на файлове и дискове, метаданни, логове, мрежов трафик, стеганография, както и начини за автоматизация на тези задачи. Обръщаме внимание основно на безплатни и отворени инструменти, които са достъпни и често използвани в практиката.

Файлов И ДИСКОВ АНАЛИЗ

Анализът на файлови системи и дискови носители е централна задача в цифровата криминалистика. При нея се работи с копия (образи) на дискове, отделни файлове или други носители (USB памети, карти памет). Няколко важни аспекта и техники:

- Физически vs. логически образи: Форензик анализаторът може да създаде физически образ на носител (bitstream copy на целия диск) или логически образ (копие само на определени дялове или файлове). Физическият образ позволява по-задълбочен анализ по възстановяване на изтрити данни, анализ на неразпределено пространство и др., докато логическият е по-малък и удобен, но пропуска тази ниско ниво информация. Инструменти като dd или подобрения dcfldd (които работят на принципа на побитово копиране) се използват за образи, а GUI програми като FTK Imager или Guymager улесняват процеса, добавяйки хеширане и репорти.

- Анализ на дялове и файлови системи: След получаване на образ, се идентифицират дяловете (партициите), например с инструменти като mmls (от Sleuth Kit), които показват таблицата на дяловете. За всеки дял (NTFS, FAT32, exFAT, Ext4, HFS+ и т.н.) се използват специализирани инструменти или модули за файловата система. The Sleuth Kit (TSK) предоставя множество командни инструменти: fls (листинг на файловата система, включително изтрити записи), icat (извличане на файл по inode), istat (показва метаданните на файл), и др. С тях могат да се изследват съдържанието и атрибутите на файловете, включително тези, маркирани като изтрити (но чиито данни все още са на диска). Често се ползва и графичната среда Autopsy, която е интерфейс към Sleuth Kit – тя визуализира йерархията от файлове, позволява търсене по ключова дума, показва графики на типовете файлове, времеви линии на дейност и т.н.
- Възстановяване на изтрити данни (Carving): Както споменахме, дори когато файловете са изтрити (и записите им са премахнати от директориите), техните данни може да стоят на диска до презаписване. Инструменти като Foremost, Scalpel или PhotoRec сканират побитовия образ за подписи (signature) на известни типове файлове, като например JPEG изображения винаги започват с байтове FF D8 FF E0... (или FF D8 FF E1), PDF файловете започват с %PDF-, ZIP/Office с PK и т.н. Когато открият такъв подпис и последващо валидно съдържание, инструментите извличат тази последователност като възстановен файл. Carving-техниките могат да възстановят частично повредени или фрагментирани файлове, макар че ако файлът не е на съседни сектори, простото извличане може да не го сглоби коректно. Затова по-нови инструменти използват и файлова система информация (ако има) за по-добро възстановяване.

Анализ на съдържание на файлове: За всеки интересен файл се прилага съдържателен анализ. Това включва:

- Отваряне с хексадесимален редактор за оглед на магически числа и структура. Например, ако разширението е .jpg, но магическото число не съвпада, значи файлът е преименуван (маскиран) и реално може да е друг тип. Хекс редактори като xxd, hexedit или HxD (Windows) позволяват преглед и търсене на текст/десетични модели в шестнадесетичния изглед.
- Hash на файл MD5/SHA1/SHA256. Получените хеши се сравняват с известни бази данни. Например, NSRL RDS е списък от хеши на легитимни софтуерни файлове и ако хешът съвпадне, файлът може да се игнорира (known good). Обратно, в хранилища като VirusTotal, ако хешът е познат, може да индикира зловреден файл и дори да даде име на малуера.
- Търсене на ключови думи/регулярни изрази: Инструменти като bulk_extractor сканират дискови образи за e-mail адреси, URLи, номера на кредитни карти, телефонни номера и др. по зададени шаблони. Това автоматично открива полезни следи например, че на диска има запаметени пароли или че потребител е разглеждал определени сайтове (от URL в кеша).
- Проверка за известни уязвимости: Ако инцидентът включва изпълним файл (exe, dll, elf), може да се анализира дали той експлоатира известна уязвимост. Инструменти за статичен анализ и дезасемблиране (IDA Pro, Ghidra която е безплатна) помагат при нужда от дълбок анализ на зловреден код.
- Особености при SSD и модерни носители: Тук идват и предизвикателствата. SSD дисковете използват технология на изтриване (команда TRIM), която при изтриване на файл веднага почиства клетките памет, правейки

невъзможно възстановяването по-късно (за разлика от HDD, където данните стоят до презапис). Това означава, че при съмнение за инцидент на SSD е критично бързо да се направи образ, иначе автоматичните процеси може да изтрият нужното. Също, разпределението при SSD (wear leveling) прави точните битови копия по-сложни и понякога специализирани инструменти или фърмуерни функции са нужни, за да се обходят всички клетки. Друга новост са криптираните дискове е ако дискът е криптиран (BitLocker, LUKS и т.н.) и ключът не е наличен, анализаторът трябва да търси ключа в паметта (ако системата е била работеща) или чрез т.нар. атаки с речник/бройни комбинации, ако има подозрения за парола.

Файловият и дисков анализ предоставя базата данни от доказателства за почти всички разследвания. Затова и инструментите в тази област са богати. Освен споменатите, заслужава да се отбележат YARA (за сканиране на файлове по определени модели на байтове, често ползван за откриване на зловредни индикатори в множество файлове) и двоично обработване на файлови системи, като например TSK's timeline функционалност, която екстрахира MACB времена (Modified, Accessed, Changed, Birth timestamps) на всички файлове и позволява да се види активността на файловата система по време.

Анализ НА МЕТАДАННИ И EXIF ДАННИ

Всеки файл носи не само видимото съдържание, но и скрита информация за него чрез метаданни. Анализът на метаданни може да разкрие неочаквани подробности, които да са ценни за разследването. Няколко примера:

- EXIF в изображения: Снимки, направени с цифрови фотоапарати или телефони, обикновено съдържат вградени EXIF данни (Exchangeable Image File Format). Те включват дата и час на заснемане, модел на устройството (напр. “Canon EOS 5D Mark IV”), понякога GPS координати на мястото, настройките на камерата и дори име на софтуера за обработка, ако е редактирано. В криминалистиката това е полезно, за пример ако имаме снимка, публикувана от източник, EXIF може да покаже къде е направена (геолокация) или кога (разбира се, ако данните не са премахнати). Имало е случаи, в които престъпници се издават по EXIF, например снимка в социалните медии разкрива GPS локацията им в момента на снимане.
- Документни метаданни: Формати като Microsoft Office (Word, Excel, PowerPoint) и PDF също носят метаданни. Те могат да включват: име на автора (или името на компютъра/потребителя, регистриран в софтуера), организация, дата на създаване, дата на последна промяна, редактори на документа, версия на софтуера, пътища до шаблони. Например, един DOCX документ, изтекъл анонимно, би могъл чрез метаданните да издаде че автор е „Ivan Petrov“ и организация „CompanyX“, което насочва разследването към конкретен източник. PDF файловете често имат полета Producer, Creator, които посочват с коя програма са създадени. Това понякога помага да се разбере дали файлът е например сканиран физически документ (тогава Producer може да е „Canon Scanner XYZ“) или е конвертиран от Word (Producer: „Microsoft Word ...“).
- Файлова система метаданни: Вече споменатите MAC времена на файловете, те също са метаданни, макар и на ниво файловата система, не вътре във файла. Чрез тях се

изграждат таймлайни като например, ако един зловреден скрипт има Created time 10:15, а логът показва изпълнение на команда в 10:16, това се свързва. При Windows NTFS има и $MFT (Master File Table) с записи за файловете, при EXT4 са налични inodes и journaling информация, които дават исторически данни дори за изтрити вече файлове.

- Email headers: В контекста на метаданни можем да споменем и анализ на имейл хедъри. Те не са точно файлови метаданни, а част от съдържанието, но се анализират със сходна цел, за да се открие откъде е дошло писмото (полета Received: показват веригата от mail сървъри и IP адреси), дали автентично (справка по DKIM, SPF записи), кога е изпратено (дати в UTC и локално време, понякога издаващо часовата зона на подателя).

Инструменти за метаданни: Един от най-използваните е ExifTool, който поддържа извличане на метаданни от огромен брой формати: изображения, документи, видео/аудио файлове, дори изпълними. С exiftool <file> можем да получим списък с всички налични мета-полета. Например за снимка image.jpg ExifTool би върнал: Camera Model: iPhone X, Date/Time: 2025:04:01 14:33:21, GPS Lat/Long, etc. За Word документ: Author, Last Saved By, Company, Total Editing Time, и т.н. Друг инструмент е Pdfinfo/Pdfid за PDF, които дават основна информация и дали PDF съдържа скриптове, формуляри (потенциално зловредни).

Важно е да се внимава, че метаданните могат да бъдат манипулирани. Умел нападател може да изтрие EXIF или да сложи фалшиви стойности. Все пак, често при инциденти нападателите не се сещат да почистят метаданните, оставяйки следи. Затова практиката е първо да се правят копия на оригиналните файлове, след това да се пусне ExifTool и подобни, понеже понякога само отварянето на файл с обичаен софтуер

може да промени метаданни (например, отварянето на Word файл може да обнови „Last opened“ време).

Анализът на метаданни има и други специални случаи, като например във форензиката на оперативни системи: преглед на $LogFile и $UsnJrnl при NTFS (журнали на файловата система), или .plist файлове при macOS, които пазят предпочитания и история на действията, или база данни Registry в Windows (тя пази много „мета“ информация за системата и потребителските навици). Всички те допринасят за сглобяването на пъзела от улики.

Лог АНАЛИЗ И МОНИТОРИНГ

Както вече бе изтъкнато, логовете (журналите) са изключително ценни за всяко разследване. Те представляват автоматични записи на събития, водени от операционни системи, приложения и устройства. Ефективното им анализиране често решава случаите, затова ще разгледаме систематично подхода:

Централизирано събиране на логове: В една професионална среда обикновено се използват SIEM (Security Information and Event Management) или лог мениджмънт системи, които събират логове от разнородни източници на едно място (напр. Elastic Stack/ELK, Splunk, Graylog). Това улеснява претърсването и корелацията. Ако организацията има такъв инструмент, форензик анализаторът задължително го използва, като пуска заявки за определени IP адреси, потребители, грешки, във времеви интервали, за да види „следата“ на инцидента през системите. Ако няма централизиран сбор, тогава логовете трябва да се събират индивидуално от всеки хост или устройство (SSH в сървъри за /var/log, изнасяне на Event Viewer логове от Windows и т.н.).

Разбиране на различните видове логове:

- Системни логове: На Linux това включва auth.log (входи/изходи, sudo), syslog (общи съобщения), kern.log (ядрото), утилити логове (Apache има access.log и error.log, FTP сървърът свой лог, и т.н.). На Windows: Security log (логване, права), System log (събития на ОС), Application logs. Има също специфични: Event Tracing for Windows (ETW) за детайли, или PowerShell логове при включено аудити.
- Мрежови устройства: firewall-ът логва блокирани/разрешени връзки, IDS/IPS логват подозрителни пакети (напр. Snort правила, задействани), VPN сървърът пази сесии на потребители, уеб прокси пък заявките към външни сайтове, DLP системата пази евентуални опити за пращане на чувствителни данни, пощенски сървър пази следа за всички входящи/изходящи писма.
- Приложения: Бази данни (Oracle, MySQL) имат audit или general лог, който показва кой какви заявки е пускал; ERP/CRM системите имат собствени логове за транзакции; облачни услуги (O365, AWS), логове за действията на акаунтите.

Таймлайн и синхронизация: Един от първите проблеми е, че различните системи може да не са синхронизирани като време. Ако един сървър има часовник 5 минути напред, логовете му ще са 5 минути „в бъдещето“ спрямо останалите. Затова е много важно използването на NTP и следене за синхронизирани времена, а при анализ задължително да се отбележи часовата зона и отместванията. В противен случай, подреждайки събитията, може да се направят грешни заключения за последователността.

Филтриране и търсене на индикации: Форензик анализаторът често знае какво да търси: необичайно IP, име на файл, конкретен потребител. Например, ако от дисковия анализ

се установи наличие на файл malware.exe и от мрежовия връзка към badhost.com, то по логовете ще се търси кога този файл е изпълняван и кой го е създал, както и кога е имало DNS заявки или HTTP трафик към badhost.com. Използват се както прости команди като grep/findstr, така и мощни заявки в SIEM (напр. SQL-подобни или Lucene заявки). Примери:

- Търсене във firewall лог: src_ip: <IP_of_attacker> - за да се види дали този IP е правил и други опити.
- Търсене в Windows Security: Event ID 4624 (успешен логон) с LogonType=10 (remote/РДП) като показва дали някой се е логнал по RDP.
- Търсене в web server log: заявки с код 500 (грешки) или странни параметри (SQL инжекция признаци ' or '1'='1).
- Корелация: Използване на SIEM rule, което казва „давай ми всички случаи, където 1) привилегирован потребител се логва, 2) within 5 minutes модифицира конфигурация“. Така може да се улови сценарий, в който нападател влиза с админ акаунт и бързо прави промени.

Автоматизирани инструменти за лог анализ: Освен SIEM, има и форензик-ориентирани инструменти. Log2timeline (част от Plaso), който споменахме, обединява различни логове + артефакти (браузър истории, файл времена) в една хронология. OSQuery е интересно решение, което позволява на живо да се изпълняват "заявки" към система (Windows/Linux/macOS) като към база данни, извличайки рязко информация, която иначе изисква ръчни команди. Например, с OSQuery може да се изиска „дай ми всички стартирани процеси и техните пътища“ или „всички автостартиращи се програми“, като това е вид структурирано логване на състоянието.

Съхранение на логове: Практически съвет: при сериозен инцидент, копията на логове (особено ако са ротационни и могат да се презапишат) трябва веднага да се съхранят отделно.

Например, ако syslog ротира ежедневно, а инцидентът е на 5-и число, да се подсигури копие на логовете от 5-и, преди да дойде 6-и и да ги ротира. В контекста на веригата на попечителство, тези логове копия също се хешират и описват.

Лог анализът често разкрива завършващия акорд от разследването и свързва намерените на диска доказателства със събитията и действията. Например: на диска откриваме малуер с име X; от логовете виждаме, че този малуер е бил свален през уеб от конкретен адрес по време Y от потребител Z; от своя страна, това ни води до заключение, че потребител Z е жертвата на phishing имейл в този момент, който го е подмамил да свали файла. Всичко това идва именно от логовете и тяхната интерпретация.

Анализ НА МРЕЖОВ ТРАФИК

Докато логовете дават обобщена информация за мрежови събития, директният анализ на мрежовия трафик предоставя най-долния слой детайли като самите пакети, които са били разменени в мрежата. В ситуациите, когато имаме запазен трафик (PCAP файлове) или можем да го прехванем в реално време (например, ако инцидентът е текущ и решим да слушаме мрежата), този анализ е безценен.

Инструменти и подход:

- Wireshark: Най-популярният графичен анализатор на мрежов трафик. Позволява визуализиране на пакети, филтриране по разнообразни критерии (протокол, IP, порт, съдържание на байтове), реконструиране на TCP сесии (Follow TCP Stream), извличане на файлове от трафика (Export Objects за HTTP, SMB и др.). Wireshark разпознава стотици протоколи от Ethernet, IP, TCP/UDP до DNS, HTTP, TLS, FTP, SMTP, ... и др., и ги декодира разбираемо. Например, ако анализираме трафика на

компрометиран хост, можем да видим DNS заявки към подозрителни домейни или HTTP заявки, носещи изходяща информация към атакувача.

- TShark: Това е командната версия на Wireshark, която е удобна за анализ и филтриране в конзола. Например, tshark -r traffic.pcap -Y "http.request.method == POST" -T fields -e http.host -e http.content_length ще извлече всички HTTP POST заявки с хостовете и размера на съдържанието. Полезно е за бърз преглед на къде са пращани данни и колко. TShark може да използва същите филтри като Wireshark (Display filters) и е мощен за автоматизация (ще говорим повече за това в следващата секция).
- tcpdump: Ниско ниво инструмент, който най-често се използва за заснемане на трафик (capture). С tcpdump -w output.pcap на даден интерфейс може да се запише трафик, а с подходящи филтри (BPF филтри) може да се ограничи какво се записва (например само трафика от/до конкретен IP или порт, за да не става файлът огромен). Tcpdump също може да извежда в четим вид, но без детайлизиран decode като Wireshark, повече се използва за capture или за бързи проверки.
- NetworkMiner, Zeek (Bro): Това са също инструменти в арсенала. NetworkMiner (Windows) пасивно анализира pcap и опитва да извлече креденшъли, файлове и др. Zeek е мрежова платформа, която в реално време анализира трафик и генерира логове от по-високо ниво (например файл dns.log, http.log с обобщение на сесиите). Ако Zeek е бил разположен в мрежата, форензик аналитик може да ползва тези логове вместо голи pcaps, което често е по-лесно.

Какво се търси в мрежовия анализ:

- Командно-контролна комуникация (C2): Зловреден софтуер често се свързва към външен сървър за

инструкции. Това може да личи като периодични заявки към необичаен домейн (например hjuweir123.biz) или IP. Анализаторът идентифицира такива модели, например чрез статистика кой домейн колко пъти се появява, или използвайки списъци на известни зловредни домейни.

- Експлоатиране на уязвимости: Ако инцидентът е пробив чрез мрежова услуга, в pcap-а може да се видят самите зловредни пакети. Например, буферно препълване като в HTTP заявка се вижда подозрително дълъг стринг; SQL инжекция чрез URL с ' or 1=1; LFI/RFI пътеки като ../../etc/passwd. Разпознаването на такива шаблони помага да се разбере кой експлойт е използван.
- Теч на данни: Ако има изтичане, често данните трябва да „напуснат“ мрежата. Това може да стане през HTTP (качване към drop server), FTP, чрез e-mail, cloud услуга или дори DNS тунелиране. В трафика, анализаторът търси големи обеми към външни адреси или нехарактерни протоколи. Един пример: чрез филтър http.response.code == 200 && http.content_length > 1000000 се търсят отговори с над 1MB което може да са данни в отговор; или анализ на SMTP за прикачени файлове (MIME parts).
- Мрежова карта на инцидента: От pcap-а може да се изгради "кой говори с кого". Например, да видим, че заразената машина освен към командния сървър се опитва и да зарази други машини в локалната мрежа (ще има трафик от нея към множество вътрешни IP по порт 445, като опит за разпространение на ransomware по SMB например). Това води екипа да провери и тези други машини.
- Необичаен трафик: Това е широко понятие, но опитният анализатор има представа кое е нормално за дадена среда. Ако в pcap на уебсървър видим изходяща връзка към външен IP по порт 4444, това със сигурност е

подозрително (4444 често се ползва от Meterpreter reverse shell). Също, UDP трафик на големи порции към нестандартен порт може да е exfiltration.

- Дешифриране на трафик: Голямо предизвикателство е, че много от комуникациите сега са шифровани (HTTPS, SSH). Анализаторът може да не вижда съдържанието, а само метаданни (TLS SNI, сертификати, размери).

Стеганография, ЕНКОДИНГ И ЧИСЛОВИ СИСТЕМИ

В някои случаи атакуващите или вътрешните злонамерени лица използват техники за скриване или обфускация на информацията, за да затруднят откриването ѝ. Стеганографията е практика на укриване на данни в друг носител като например скрит текст в изображение или аудио файл, така че на пръв поглед файлът изглежда обикновен. Разследващите трябва да познават и тези методи. Типичен подход е анализ на изображения за LSB (Least Significant Bit) стеганография, при нея скритото съобщение се кодира в най-маловажните битове на пикселите и е невидимо за човека. Съществуват инструменти (StegExpose, StegoVeritas и др.), които автоматично откриват аномалии в статистическото разпределение на цветовете и могат да насочат към скрито съдържание. Например, ако при нормална JPEG картинка най-маловажните битове са случайни, но при стеганографирана те носят структура, това се открива чрез специализиран анализ.

Освен стеганография, често срещана е обфускацията чрез енкодинг като напр. вредоносен код или откраднати данни се преобразуват с Base64, хексадециално представяне, или други схеми, за да не бъдат разпознаваеми. Анализаторът трябва да умее да разпознава такива шаблони и да ги декодира. Например, един лог може да съдържа дълга низ от привидно безсмислени знаци, които всъщност са Base64-кодиран текст (често се разпознава по това, че съдържа букви, цифри, „+“ и „/“ и

дължината му е четима). След конвертиране от Base64 може да се окаже, че това е скрипт или URL. Подобно, ако в паметта се намери поредица като \x48\x45\x4C\x4F, трябва да се разчете, че това в ASCII кодировка е „HELO“. Понякога нападателите кодират съобщения като поредица от десетични или осмични кодове – умението бързо да се превеждат между двоична, осмична, десетична и шестнадесетична бройна система е полезно. Разбира се, това се подпомага от инструменти: почти всички езици имат команди за конверсия (например xxd за хекс -> ASCII, или онлайн/скриптови бази за Base64).

Binwalk е специфичен инструмент, който комбинира аспектите на стеганография и файлов анализ, използва се основно за сканиране на бинарни файлове (напр. фърмуер) за вградени подписи. С Binwalk можем да открием, че един уж JPEG файл всъщност съдържа appended ZIP архив след края на изображението (популярен трик за скриване). Това е пример за combining steganography and encoding: файлът работи нормално като картинка, но има допълнителни данни скрити след неговия край, които Binwalk ще идентифицира по ZIP сигнатурата. При подозрение за укрити данни, специалистите прилагат подобни инструменти и техники.

Както и при другите области, ключово е креативното мислене, нападателите могат да използват нестандартни формати или собствени алгоритми за стеганография. Тогава форензик анализаторът трябва да следи подозрителните признаци (например файл, чийто размер е по-голям от очакваното за даденото съдържание, или мрежов трафик, кодиран да изглежда като обикновен) и при нужда да пише собствени скриптове за декодиране.

Автоматизация ЧРЕЗ СКРИПТОВЕ И TSHARK

Предвид огромните обеми данни, които могат да се съберат при един инцидент (дискови образи стотици гигабайти, милиони лог записи, часове мрежови пакети), автоматизацията е най-добрият приятел на форензик експерта. Чрез скриптове и инструменти команден ред, повтарящи се задачи се изпълняват бързо и консистентно.

Как може да се ползва автоматизация на практика? Един пример е TShark, споменат по-горе – това е командния интерфейс на Wireshark, който позволява филтриране и извличане на информация от мрежови capture-и без човешка намеса. Да кажем, че разследването изисква да се прегледат 50 PCAP файла за присъствие на определен IP адрес или домейн. Ръчното отваряне на всеки във Wireshark би било непосилно. С един скрипт обаче (bash или python), можем да обходим всички файлове и за всеки да стартираме команда на TShark като: tshark -r fileX.pcap -Y "ip.addr == 192.0.2.123" -w suspicious_packets.pcap. Така машинно ще филтрираме и запишем всички пакети към/от този IP в нов файл. Подобно може да се направи за търсене на ключова дума в HTTP съдържание или конкретен порт.

Скриптовете (на Bash, Python, PowerShell и др.) се използват масово и за парсване на логове: например, да се минат всички системни журнали и да се изведат само записите с грешки или само тези между две дати. Python има библиотеки за обработка на JSON, CSV, XML формати, в които често се експортират данни от инструменти. Това позволява анализаторът да напише Python скрипт, който прочита даден JSON (да речем, износ от SIEM система) и автоматично извлича от него най-важното, като списък на уникални зловредни IP-та, часови диапазон на активност и т.н.

Автоматизацията пести време и намалява грешките. Където е възможно, ръчната работа следва да се допълва със скриптове. Инструменти като Log2timeline/Plaso на практика

представляват автоматизация, те вземат куп източници и сами генерират времева линия. Друга сфера е репортингът: след като анализаторът събере няколко ключови статистики или списъци, може чрез скрипт да ги форматира в таблица или графика за доклада.

В рамките на инцидентния отговор автоматизацията също помага за реакцията в реално време. Някои екипи подготвят „скриптови лечебни рецепти“, например ако на 100 машини трябва спешно да се изключи една услуга или да се блокира IP, скрипт или конфигурационен мениджър (Ansible, PowerShell script) го прави наведнъж. Това вече преминава към SOAR (Security Orchestration, Automation and Response) системи, които автоматизират и самия отговор.

Но дори и без специализиран софтуер, уменията за скриптиране са задължителни за криминалистите. Една просто написана програма може да прерови милиони данни за минути, което е задача, отнемаща дни на човек. Затова при обученията по цифрова криминалистика често се включват и модули по програмиране на скриптове и използване на CLI инструменти.

Безплатни и отворени инструменти под Linux:

Инструмент	Предназначение и приложение
Autopsy	Графичен интерфейс за Sleuth Kit е цялостна криминалистична платформа за анализ на дискови образи, извличане на файлове, генериране на доклади и времеви линии. Лесен за използване от начинаещи и разширяем с плъгини.
The Sleuth Kit	Набор командни инструменти и библиотеки за анализ на файлови системи и дисков. Включва утилити за листинг на директории, възстановяване на изтрити файлове, анализ на метаданни. Основен „двигател“ зад много други инструменти.
Foremost	Конзолен инструмент за *file carving*, който претърсва дискови образи за вградени файлове по техните сигнатури. Може да възстановява изтрити

	снимки, документи, видео и др., дори след като файловата система е забравила за тях.
Binwalk	Инструмент за анализ на бинарни файлове (особено фърмуери). Открива вградени файлове и данни в тях (напр. скрити ZIP архиви, изображения, скриптове). Полезен за стеганография и при reverse engineering на устройства.
YARA	Език и инструмент за дефиниране на шаблони на байтове или стрингове, по които да се разпознава злонамерен код. Използва се за сканиране на файлове или памет за наличие на известни малуер семейства чрез *YARA правила*.
ExifTool	Мощен инструмент за извличане на метаданни от файлове (изображения, документи, видео и др.). Показва EXIF, IPTC, XMP данни на снимки, автори и редакции на документи, и много други скрити полета. Незаменим при анализ на документи и снимки за скрита информация.
Volatility	Рамка за *memory forensics* за анализ на копия на оперативна памет. Има плъгини за извличане на списък с процеси, мрежови връзки, заредени DLL, кеширани пароли, и др. Помага за откриване на *fileless* малуер и атаки, оставящи следи само в паметта.
Log2timeline/ Plaso	Инструмент за автоматизирано изграждане на времева линия на събитията. Взема данни от файлове (напр. MFT от NTFS, регистратурата на Windows, логове) и генерира подредена последователност от събития, която може да се анализира с инструмента psort.
Wireshark	Популярен мрежов анализатор с графичен интерфейс. Позволява преглед в детайл на мрежови пакети, филтриране по протоколи, сглобяване на сесии и извличане на съдържание. Използва се за анализ на мрежови атаки, ексфилтрация на данни и преглед на подозрителен трафик.
Tcpdump	Класически инструмент за команден ред за прихващане и първичен анализ на мрежов трафик. Често се използва за бърз capture на пакети в полеви

	условия или на сървъри (например: `tcpdump -w traffic.pcap port 80` записва целия HTTP трафик).

Таблица 2: Примери за безплатни/open-source инструменти за цифрова криминалистика и инцидентен отговор под Linux

Много от тези инструменти имат версии и за Windows, но Linux е предпочитана среда заради гъвкавостта си и богатството на командно-редови утилити.

Списъкът не е изчерпателен, съществуват десетки други инструменти (Bulk Extractor, Caine Linux дистрибуцията, LiME за live memory capture, ddrescue за възстановяване на повредени дискове и др.). В практиката, експертите изграждат свой набор според нуждите и случая. Важното е, че наличието на отворен код позволява да се инспектира как работи инструментът (важно за свидетелстване в съда) и да се модифицира при нужда за специфични задачи.

4.6. Практически СЦЕНАРИИ И СЪВЕТИ ЗА РЕАЛНА СРЕДА

Теорията и инструментите са основата, но при прилагането им на практика възникват множество предизвикателства. Тук обобщаваме няколко практически сценария с акцент върху събиране на доказателства, съхранение и поддържане на верига на попечителство, както и други съвети, произтичащи от реални случаи:

Изземване на компютър на местопрестъпление: Представете си случай на офис служител, заподозрян в кражба на данни. Лаптопът му трябва да бъде иззет за анализ. Практическият съвет е да се изключи захранването възможно най-бързо, за да се запази състоянието на диска. При работещ компютър може да се обмисли форензично спиране: ако системата е заключена и няма криптиране на диск, дръпване на

захранването е ок; ако обаче дискът е криптиран и се опасяваме, че изключване ще заключи данните (без достъп до ключа), то тогава по-добре да се приспи (Sleep/Hibernate) или да се остави работещ и да се извлече памет и да се навигира внимателно. Документация на място: запишете точно време на изземване, от кого, серийни номера на устройството, състояние (включен/изключен), и обезопасете устройството (например с етикет върху портовете, че е иззето, за да не бъде включвано). Това е началото на веригата на попечителство, всеки който докосва доказателството от тук насетне трябва да остави "следа" в документация.

Събиране на данни в действаща среда: В корпоративна среда често не може просто да се изключат всички сървъри при инцидент, тъй като това би спряло всички процеси. Нужно е балансиране между запазване на доказателства и минимизиране на прекъсването. Добра практика е да се клонират трафика (порт mirroring) или дисковете, вместо да се спират, напр. ако уеб сървър е компрометиран, да се пусне tcpdump на изследваната машина (или на мрежовия комутатор) да събира трафик, докато сървърът продължава работа, или мигриране услугата на друг сървър, а оригиналния се изключва за анализ.

Remote forensics: Много инструменти позволяват събиране отдалечено, например с PSExec и disk2vhd на Windows може да вземем образ на машина през мрежата. Винаги обаче трябва да има копие, на което да се правят експериментите, а оригиналът да остане непокътнат. Ако трябва да се извърши анализ на жива система, задължително се записва кои команди са изпълнени, за да се знае какво евентуално би могло да промени по системата.

Съхранение на дигитални доказателства: След събирането, данните (дискови образи, памет, логове, трафик) трябва да се съхраняват сигурно. Това означава контрол на достъпа само от оторизирани лица от екипа да имат достъп,

съхранение на хеш сумите и регулярна проверка на тях (да се уверим, че файловете не са се променили). Често за по-големи случаи се изграждат Forensic Storage сървъри или използват WORM устройства (Write Once Read Many) за гаранция, че никой няма да промени записаното. Физическите носители (напр. иззет твърд диск) се слагат в антистатични пликчета, пломбират се и се държат в сейф. **Веригата на попечителство:** това понятие (chain of custody) означава, че за всяко доказателство има запис кой го е притежавал или предавал във всеки момент от събирането до представянето в съда. Целта е да не остане съмнение, че някой нерегламентирано е имал достъп и евентуално е подправил доказателството. На практика това става с формуляри или дигитални регистри, вписва се „Диск SN123, иззет от X на дата/час, предаден на аналитик Y на дата/час, съхранен в шкафа Z“. Ако по-късно се вади, пак се записва кой и защо. Всяко нарушение (липсващ запис, неоторизиран достъп) може да компрометира допустимостта в съда.

Поддържане на целостта на данните: Освен ограничаване на достъпа, се използват и технически средства, като хеширане и write-blockers. Write-blocker е устройство (или софтуерна настройка), което позволява да четем от диск, без да могат да се правят записи на него, задължително при свързване на иззети дискове към анализаторски компютър, за да не се промени нищо случайно. Хеширането (SHA-256, SHA-1 и др.) се прилага първо на място (напр. още преди копиране на диск, изчисляваме хеша на оригинала), после на копието хешовете се съпоставя и трябва стойностите да съвпадат. Същите хешове фигурират в доклада и уверяват читателя (ръководство или съдия), че данните, които са анализирани, са точно тези, които са иззети, без модификация. Ако по време на анализа се създадат нови артефакти (напр. извлечени файлове), те също могат да се хешират и приложат като доказателство.

Етика и поверителност: При работа с реални данни, особено в корпоративна среда, анализаторите могат да попаднат на чувствителна информация, несвързана с инцидента (лични файлове на служители, привилегирована информация като адвокатска кореспонденция и т.н.). Професионалният екип трябва да спазва принципа на минимално необходимото разглеждане, да се фокусира върху релевантните за разследването данни и да не разгласява странична информация. В много фирми членовете на реакцията подписват допълнителни споразумения за конфиденциалност. При инциденти, включващи лични данни, трябва да се вземе предвид и законодателството (напр. GDPR) и може да се наложи уведомяване на засегнатите лица или регулатора в определен срок, което е част от инцидентния отговор.

Координация с правоохранителни органи: Ако инцидентът е свързан с престъпление (кражба на данни, измама, саботаж), организацията трябва да прецени дали да уведоми полиция или специализиран отдел (в България това е ГДБОП, сектор „Киберпрестъпност“). В някои случаи това е задължително (напр. при пробив, засягащ голям брой лица). Добра практика е предварително да има изградени контакти, ако не формални, то поне да е ясно към кого да се обърне организацията. При привличане на външни разследващи, веригата на попечителство трябва гладко да се предаде, т.е. доказателствата да бъдат предадени с протокол на органите. Също така, трябва да сте подготвени, че при съдебни действия, вътрешните експерти могат да бъдат извикани да свидетелстват като вещи лица, където ще трябва да обяснят методите си и да защитят валидността на събраните данни.

Бърза реакция срещу задълбочен анализ: В реални условия често има напрежение между това да се възстанови работата (да „потушим пожара“) и това да се съберат всички улики за анализ. Затова е полезно да се разграничат ролите, един

под-екип може да работи по обезопасяване и възстановяване (например изолиране на мрежови сегмент, пускане на бекъпи), докато друг паралелно прави събиране на данни. Координацията помежду им е ключова: не бива, например, системата да се рестартира за възстановяване преди екипът по доказателствата да е свалил паметта и критичните логове. Затова планът трябва да предвижда тези стъпки, а ръководителят на инцидента да балансира приоритетите според ситуацията (напр. ако ransomware криптира всичко бързо, може да се жертват малко форензични детайли в замяна на бързо спиране на разпространението).

Учене от всеки случай: Практическият съвет, който често се дава, след всеки инцидент, независимо голям или малък, да се прави разбор с екипа. Това затвърждава наученото, и дори да не е имало сериозни грешки, винаги има нещо за подобряване, било то технически (например „да включим нов индикатор в системата за мониторинг“) или организационно („следващия път да уведомим по-рано отдел X“). Също така, да се поддържа база знания, вътрешна документация на срещаните атаки, използвани TTPs (тактики, техники и процедури) на нападателите, ефективни защити. Това постепенно ще ускори бъдещи разследвания, защото ще има референции как са решавани подобни проблеми в миналото.

Накратко, работата в реална среда изисква както стриктност и следване на процедура (за да са валидни доказателствата), така и гъвкавост и здрав разум (за да се минимизират щети). Всеки инцидент е различен, но основните принципи е да се пазят данните, да се действай бързо, документиране на всичко, мислене за безопасността.

4.7. Обобщения И ПРЕПОРЪКИ

Цифровата криминалистика и реагирането при киберинциденти формират заедно гръбнака на съвременната киберзащита. Тази глава разгледа дефинициите, значението и правния контекст на цифровата криминалистика, как тя служи за намиране на истината в цифровите данни и какви са предизвикателствата пред използването на електронни доказателства. Бяха описани често срещаните видове киберинциденти (пробиви, вътрешни злоупотреби, ransomware, утечки) и бе подчертана необходимостта към всеки случай да се подхожда както с мерки за спиране на инцидента, така и с мерки за разследване.

Разгледани са фазите на разследване, от идентификация през събиране, анализ, докладване до възстановяване и стана ясно, че макар теоретично разграничени, на практика те често се преплитат. Ключово е всяка фаза да се изпълнява внимателно: правилно да се идентифицира инцидента (за да не се пропусне или пък фалшива тревога да ангажира ресурс), изчерпателно да се съберат данните („по-добре повече, отколкото по-малко“, но и без излишен шум), дълбоко да се анализират (но с приоритети при спешност), ясно да се докладват и бързо да се възстанови функционирането. Организационните аспекти, наличие на екип, план, роли и комуникация, често се оказват разликата между успешен и хаотичен отговор.

Утвърдените стандарти (ISO 27035, NIST 800-61, SANS и др.) дават рамка и най-добри практики, които силно се препоръчва да бъдат възприети. Те не са само „за показ“, при реален инцидент следването на проверен план спестява ценно време и намалява грешките. Препоръката към организациите е да интегрират тези стандарти в своите политики, да тренират екипите си редовно и да осигурят нужните инструменти и достъп до експертиза (било то вътрешна или външна).

От технологична гледна точка, се разгледа широка гама от инструменти: анализ на файлове, метаданни, логове, мрежови

пакети, памет; откриване на скрити данни и обфускация; използване на скриптове и специализирани програми. Инструментариумът на един криминалист включва както универсални инструменти (напр. Wireshark, ExifTool), така и специфични за дадени задачи (Volatility за памет, YARA за малуер, Foremost за carving). Силно се препоръчва да се използват отворени инструменти поне в началната фаза, те са достъпни, с общностна поддръжка и позволяват проверка. В някои случаи разбира се, комерсиални решения (EnCase, FTK, X-Ways, Magnet AXIOM) може да предложат удобство и допълнителни функции, добрият специалист умее да работи с всякакви средства и да избере най-подходящото според ситуацията и бюджета.

Поддържането на веригата на попечителство и изрядната документация са абсолютно задължителни, ако резултатите от разследването ще се ползват във външни разследвания или съд. Дори за вътрешни цели, те гарантират доверие в находките, мениджмънтът е по-склонен да предприеме действия (дисциплинарни или инвестиции в сигурност), ако пред него има добре обоснован доклад с факти, отколкото неформални подозрения. В този дух, препоръката е: да се отнасяме към всяко разследване като потенциално съдебно, това дисциплинира процеса и повишава качеството.

Накрая, важно е да се осъзнае, че заплахите се развиват непрекъснато. Нови видове атаки (напр. към облачни услуги, IoT устройства, AI системи) ще изискват и нови подходи в криминалистиката. Затова специалистите трябва да продължават да учат и експериментират. Участието в общността (конференции, уъркшопи, онлайн форуми), споделянето на анонимизиран опит и следенето на актуални инциденти ще помогне на екипите да са подготвени за следващото предизвикателство.

В обобщение, изграждането на силна функция по цифрова криминалистика и инцидентен отговор е инвестиция, която се отплаща многократно чрез предотвратяване или бързо

овладяване на киберинциденти, намаляване на щетите, защита на данните и репутацията. Въоръжени със знанията от тази глава, от принципите и процесите до инструментите и практическите съвети, читателите могат уверено да навлязат по-дълбоко в тази критична област и да допринесат за по-сигурната цифрова среда на своята организация.

5. КИБЕР ПОЛИГОНИ

В областта на киберсигурността все повече се налага разбирането, че обучението на кадри и изпитването на реакции при инциденти трябва да се случва в реалистична, но контролирана среда, подобно на военните учения или авиационните симулатори. Такъв тип среда в ИТ сферата се нарича кибер полигон (или кибер тренажор, кибер симулатор, кибер учебен център). Кибер полигон е инфраструктура (хардуерна и софтуерна платформа), която позволява симулиране на мрежи, системи и атаки, върху които специалистите по сигурност могат да тренират уменията си, да изпробват инструменти и да проучват поведението на заплахите, без риск за реалните производствени системи. В тази глава разглеждаме какво представляват кибер полигоните, как се използват за обучение, и даваме примери за конкретни платформи и инициативи.

5.1. Същност И РОЛЯ НА КИБЕР ПОЛИГОНИТЕ ЗА ОБУЧЕНИЕ

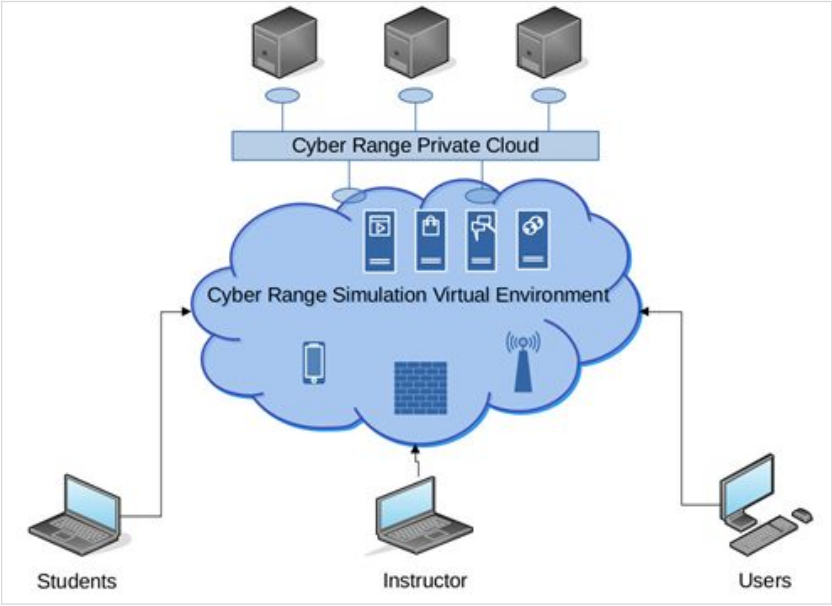
Дефиниция и цел. Кибер полигонът по същество е една виртуална „пясъчник“ среда в голям мащаб, където могат да се изграждат модели на корпоративни мрежи със сървъри, работни станции, мрежови устройства, софтуер и върху тях да се разиграват разнообразни сценарии за кибератаки и защита. Целта е участниците (които могат да бъдат екипи от различни организации или от една организация) да получат практически опит в справянето с инциденти: да разпознават атаки, да приложат противодействие, да комуникират и да възстановяват системи. Това се прави в контролирана среда, така че грешките и пораженията да са учебен опит, а не реална щета. Полигонът предоставя безопасно място, където да се експериментира

например, да се види как би се разпространил определен вид малуер, какви следи оставя, кое е най-ефективното средство да се спре, колко време отнема възстановяването на системите след него.

Компоненти. Кибер полигона обикновено включва: инфраструктурен слой, като хардуерни или облачни ресурси, на които се изпълняват виртуалните машини и симулациите; Симулатор е предварително подготвени сценарии на атаки/инциденти, които да се случват (напр. заразяване с рансъмуер, DDoS атака, проникване чрез уязвимост); набор от инструменти за участниците, като те имат на разположение обичайните средства, които биха имали и в реална ситуация (SIEM, антивирус, мониторинг, документация, playbooks); Модул за оценка често полигоните следят действията на екипите и оценяват ефективността им по определени критерии (време за реакция, успешност на предотвратяване, правилна комуникация и последователност на стъпките и т.н.).

Роля за повишаване на подготовката. Подобно на пожарникарските учения, редовното участие в кибер упражнения повишава способностите и увереността на екипите при реални инциденти. Те се научават да действат под стрес, да координират отговор между различни роли (напр. екипи по мрежи, по системи, по сигурност, ръководство), да мислят креативно срещу адаптивен противник. Добрите полигони предоставят сценарии, базирани на актуални реални заплахи. Например симулират атаки като тези, които са станали известни в последните месеци, за да може обучението да отразява съвременния нападателен метод. Това гарантира, че съдържанието е актуално спрямо най-новите хакерски техники. Такива сценарии могат да включват рансъмуер епидемия, проникване през supply chain (компрометиран софтуер, внесен отвътре), целенасочен APT (Advanced Persistent Threat) с дългосрочно присъствие в мрежата и др. Също така, много полигони насърчават разнообразие на

ролите: някои участници влизат в ролята на „червен отбор“ (attacker), други - „син отбор“ (defender), а понякога има и „бял отбор“ (съдии/наблюдатели). Това позволява интерактивно обучение, като защитниците реагират на ходовете на нападателите, което е по-достоверно от статична симулация. Умението да се мисли като хакер е ценен опит и за защитниците, познавайки манталитета и техниките на противника, те могат да подобрят и превантивната си работа.



Фиг. 19. Архитектура на кибер полигон

Хибридност и реализъм. С напредъка на технологиите, съвременните полигони често комбинират облачни ресурси с локални среди за максимална гъвкавост. Може да има локални сегменти за симулация на заводски мрежи или критична инфраструктура (OT - Operational Technology), съчетани с облачни компоненти за ИТ системите. Това отговаря на реалността, където

много системи са вече хибридни. Също полигони могат да интегрират реални потокови данни като например, някои платформи (като ThinkCyber, обсъдена по-долу) се свързват към глобални сензори за заплахи, които в реално време подават индикатори за актуални атаки, така че сценариите да могат да се напасват и обновяват на момента. Всичко това създава динамична и актуална обучителна среда, която надхвърля учебниците и статичните лабораторни упражнения.

5.2. Пример: ПЛАТФОРМАТА CYBERIUM (THINKCYBER) - РЕАЛИСТИЧНИ СЦЕНАРИИ, РОЛИ И ОЦЕНКА

За да се опише по-ясно концепцията за кибер полигон, ще разгледаме като пример една конкретна платформа Cyberium, разработена от академията ThinkCyber идваща от отбранителния сектор на Израел. Cyberium е цялостна арена за кибер обучения, която комбинира няколко ключови елемента: реалистични сценарии, интерактивни роли за участниците, интеграция с актуална информация за заплахите и система за оценка на представянето. Опитът, споделен от проведени обучения на тази платформа, дава ценни насоки за ефективността на подобни подходи.

Сценарии, базирани на реални заплахи. В основата на ThinkCyber модела стои идеята да се използват автентични казуси от действителни кибератаки. Вместо хипотетични или твърде опростени упражнения, сценариите в Cyberium симулират инциденти като рансъмуер атаки, пробиви на мрежата, компрометиране на верига на доставки и др., т.е. ситуации, които наистина са се случили някъде. Данните за тези сценарии се черпят от анализа на реални случаи чрез платформа наречена Specto AI, която агрегира глобална информация за заплахи. Това гарантира, че съдържанието на обучението е постоянно обновявано и релевантно към най-новите хакерски техники. Например, ако наскоро е имало вълна от атаки тип „ransomware

2.0“, включващи кражба на данни преди криптиране, Cyberium може да изгради сценарий, в който екипите трябва да реагират на подобен инцидент: първо необяснимо изтичане на информация, после криптиране и искане на откуп и да се види как биха се справили.

Интерактивни роли и обучение чрез правене. Участниците в Cyberium не са пасивни слушатели, а поемат конкретни роли в хода на симулацията. Обичайните роли включват: офанзивни (attacker) играещи ролята на червен отбор, дефанзивни (defender) синият отбор, отговорен за защитата, и форензик анализатори, които се фокусират върху събирането на доказателства и разследване след инцидента. Тази структура принуждава участниците да вземат решения под напрежение и да прилагат знанията си практически. Например, защитниците трябва да преценят кои системи да изолират, какви мерки да приоритизират (спиране на атака или запазване на логове), а нападателите как да останат незабелязани по-дълго или да заобиколят определена защита. Чрез това учене и чрез действие (learning by doing) се развиват не само техническите умения, но и стратегическото мислене и екипната комуникация. Участниците се насърчават да мислят нестандартно и да адаптират плановете си, защото в динамичното поле на киберсигурността шаблоните не винаги работят. По този начин, един кибер полигон служи и за тест на процесите на организацията: ако в реална ситуация плановете за реакция са тромави или неясни, това ще се прояви и по време на симулацията, давайки възможност за корекции и учене от грешките в безопасна среда.

Реално време и интегриране на заплахи. ThinkCyber Cyberium използва споменатата Specto AI платформа, която събира данни от глобални сензори („примамки“) за заплахи в реално време. Това означава, че докато тече обучението, може да бъдат вкарвани актуални елементи. Например, ако в реалния свят се появи нов вариант на малуер, обучаемите много скоро

могат да се срещнат с негова симулация. Сценариите се адаптират към текущия пейзаж на заплахите, осигурявайки, че трениращите се сблъскват с най-модерните атаки, а не с остарели примери. Освен това, платформата комбинира теоретично знание с практиката, като участниците имат достъп до богат набор от ресурси (книги, видео, проекти) по време на обучението. Така, ако срещнат нещо непознато, могат бързо да потърсят справка или насока. Това създава холистична учебна среда, където се насърчава не само „гасенето на пожара“, но и разбиране в дълбочина. Например, ако екипът установи атака чрез PowerShell скрипт, те могат да се консултират с налични материали за типовете зловредни PowerShell техники, подобрявайки знанията си докато действат.

Модулност и достъпност. Концептуалният модел на Cyberium е организиран в модули, което позволява постепенно надграждане на уменията, от основни нива на обучение до много сложни специализирани задачи. Така начинаещи могат да се запознаят с базови неща (напр. откриване на фишинг имейл), докато напреднали професионалисти могат да преминат към сложни разследвания. Платформата е достъпна през облака, което дава гъвкавост. Обучаемите могат да практикуват от всяка точка, без да е нужна физическа лаборатория на място. Това е особено ценно за глобални екипи или при ограничителни обстоятелства (като пандемии), където събирането на всички на едно място е трудно. Облачният подход също така улеснява мащабирането, като могат паралелно да се тренират множество екипи, всеки в отделна изолирана среда.

Оценка на ефективността (CES). Един иновативен аспект в Cyberium е Cyberium Effectiveness Score (CES), което е система за количествено измерване за представянето на участниците. По време на упражнението се следят ключови показатели: време за откриване на атаката, време за реагиране, колко правилни стъпки са предприети, колко грешки са допуснати, дали комуникацията

(напр. ескалация към мениджмънт) е станала своевременно, и т.н. На база на тези фактори, накрая всеки екип или участник получава оценка, която отразява тяхната ефективност и адаптивност. Това позволява идентифициране на силни и слаби страни като например, един екип може да е много бърз в техническата реакция (висок резултат там), но да има слабости в процеса на докладване и координация (нисък резултат по тези метрики). Чрез CES обучаемите получават конкретна обратна връзка и насоки къде да се подобрят. Така платформата не само обучава, но и измерва прогреса, което е важно за организациите, инвестиращи в тези обучения, те могат да видят дали след няколко сесии резултатите на екипа се вдигат, т.е. дали обучението дава резултат.

Фиг. 20. ThinkCyber Симулатор, конзола на кибер полигон

Архитектура и сигурност на самата платформа. Накратко ще спомена, че зад кулисите Cyberium е изграден върху сложна хибридна архитектура, комбинираща облачни и частни инфраструктури, за да гарантира сигурност и мащабируемост. Важно за доверие към една такава платформа е тя самата да е

добре защитена, тъй като симулира атаки, би било катастрофално, ако нападател успее да проникне в полигона и оттам в мрежите на участниците. ThinkCyber са заложили на изолация между средите, силни контроли за достъп и мониторинг на платформата. Те подчертават, че комбинацията от облак (гъвкавост, достъп отвсякъде) и частна инфраструктура (за чувствителни компоненти) им е позволила да създадат надеждна и сигурна учебна среда.

5.3. Други ПРИМЕРИ И ПЛАТФОРМИ ЗА КИБЕР ОБУЧЕНИЯ

Cyberium/ThinkCyber не е единственият кибер полигон, в глобален мащаб има множество инициативи, които също заслужават внимание, тъй като допринасят за колективната киберустойчивост.

Cyber Polygon (глобално учение). Забележка: Не бива да се бърка общият термин „кибер полигон“ с проекта Cyber Polygon, така е наречено международно онлайн киберучение, провеждано в партньорство със Световния икономически форум през 2020-2021 г. Cyber Polygon (с главни букви) представлява най-голямото по рода си техническо учение за корпоративни екипи, съчетано с онлайн конференция на високо ниво. В изданието му през 2021 г. участват стотици организации от различни страни, които заедно са тренирали реакция срещу симулирана глобална кибератака. Форматът комбинира три направления: технически тренировъчни сесии, стратегически учения и дискусии. Целта е повишаване на глобалната киберустойчивост чрез споделено учене и координация. Този вид упражнения показват, че значимостта на кибер полигоните е призната и на международно ниво, те са средство не само за обучение на отделни екипи, но и за изграждане на доверие и връзки между организации, които биха си партнирали при реални инциденти.

Правителствени и военни кибер полигони. Много държави инвестират в национални кибер полигони или т.нар. кибер диапазони (cyber ranges) за обучение на военни и държавна администрация. Например, НАТО разполага с Cyber Range платформа в Естония, където ежегодно се провежда известното учение „Locked Shields“, което е сложна симулация на кибервойна, включваща отбрана на критична инфраструктура. Израел, САЩ, Русия, Китай всички те имат собствени правителствени полигони, често разработени от водещи технологични компании или университети. Тези среди са обикновено класифицирани и изолирани, и се използват за трениране на кибер армии и специалисти, които биха действали при киберконфликти. От тях излизат много иновации, които по-късно се появяват и в комерсиалните продукти (например сценарии за защита на електроцентрали или водоснабдяване, днес са достъпни и за цивилни учения, но първо са моделирани във военните среди).

Академични и корпоративни полигони. В академичните среди, университетите изграждат собствени лаборатории/полигони за обучение на студенти по киберсигурност. Например, в България няколко университета работят в партьорство с ThinkCyber по такива проекти, създавайки виртуални лаборатории за студентски състезания (CTF - Capture The Flag) и практически занятия. Корпорациите също инвестират: големи компании като IBM, Cisco, Microsoft организират киберучения за своите клиенти или партньори, демонстрирайки продукти и методологии. Някои от тях предлагат комерсиални решения „кибер полигон под наем“, т.е. фирма може да наеме платформа и експерти, които да проведат учение специално за нейния екип, персонализирано към нейната инфраструктура и заплахи.

Отворени платформи. Има и open-source проекти за изграждане на кибер полигони. Например, Facebook (сега Meta)

пуснаха преди време Capture the Flag platform с отворен код, която позволява организиране на състезателни киберигри. С известна експертиза, организации могат сами да съчетаят различни инструменти (Docker контейнери за симулиране на услуги, скриптове за автоматизация на атаки) и да създадат мини-полигон за вътрешни нужди, без скъпи решения. Разбира се, това изисква много време и технически умения, затова много предпочитат готови решения.

Ползи и бъдеще. Накратко, кибер полигоните предоставят безценна възможност за обучение чрез преживяване, която допълва теоретичната подготовка. Те позволяват на екипите да направят грешки, да научат уроци и да изградят рефлекси в една контролирана обстановка, така че в реални условия да действат по-уверено и правилно. С нарастващата сложност на заплахите, можем да се очаква, че такива симулационни среди ще станат стандартен компонент от програмите по сигурност на всяка зряла организация, подобно на пожароизвестяването и аварийните упражнения. Инвестицията в кибер полигони е инвестиция в хората, най-важният ресурс в киберсигурността. А чрез споделени международни инициативи, тези упражнения допринасят и за глобална сигурност, тъй като киберзаплахите не признават граници.

6. КИБЕР РАЗУЗНАВАНЕ И ТЪМНАТА МРЕЖА: АНАЛИЗ, ТАКТИКИ, ИНСТРУМЕНТИ И РЕЗУЛТАТИ

Стратегическото значение на кибер разузнаването

В съвременната цифрова среда, в която глобалната свързаност и информационната зависимост се превръщат в основа на икономическия и държавния суверенитет, кибер разузнаването (Cyber Threat Intelligence - CTI) изпъква като ключов елемент от националната и корпоративна сигурност. Все по-често атаките не се ограничават до техническо нарушаване на инфраструктура, а включват целенасочена дезинформация, изтичане на данни с геополитическа стойност, както и кибершпионаж, финансиран от държавни актьори. В този контекст, CTI не е просто инструмент за предупреждение, то се превръща в анализатор на бъдещи събития, позволяващ трансформация на реактивните политики в проактивни стратегии за цифрова устойчивост.

С нарастващата сложност на заплахите, от APT групи (Advanced Persistent Threats), до престъпни синдикати, използващи модела „Ransomware-as-a-Service“, което прави традиционните мерки за сигурност недостатъчни. CTI допълва и надгражда защитните архитектури чрез контекстуализирана информация, която осигурява навременна осведоменост, ретроспективен анализ и стратегическо планиране.

6.1. Класификационни НИВА НА РАЗУЗНАВАНЕ

Съвременните модели на CTI приемат различни нива на класифициране, което обхваща не само времевия и технически хоризонт на информацията, но и нейната функция в процеса на вземане на решения. Най-често използваната типология е заложена в ENISA, NIST и ISO/IEC 27010 разделят разузнаването на четири основни нива:

- **Стратегическо разузнаване** – високото ниво на обобщение, насочено към управленския и политически елит в организацията. То анализира дългосрочни тенденции, геополитически рискове, регулаторни промени и развитието на заплахи, които могат да засегнат цели сектори. Например, анализ на последиците от въвеждането на NIS2 директивата в ЕС върху операторите на критична инфраструктура.
- **Тактическо разузнаване** – концентрира се върху TTPs (тактики, техники и процедури), използвани от нападатели. Осигурява мост между стратегията и оперативното прилагане, като служи на екипите по сигурност при изграждането на защитни сценарии. Пример тук е анализа на техниката „Living off the Land“, използвана от APT групи като APT29 (Cozy Bear).
- **Оперативно разузнаване** – в реално време проследява активни кампании, събития и компрометирани системи. Служи на Security Operations Center (SOC) за приоритетизация на инциденти и автоматизирани действия. Обикновено се интегрира със SIEM или SOAR платформи за бърза реакция.
- **Техническо разузнаване (Technical Threat Intel)** – свързано с конкретни артефакти като IP адреси, домейни, хешове, SSL сертификати, Indicators of Compromise (IoCs), които могат да бъдат директно използвани за блокиране или търсене на зловредна активност.

Разграничението между нивата не е строго, те често се преплитат и взаимно захранват, като всеки слой обогатява другия. Успешното CTI се изгражда именно върху способността за синхронизация между тези нива, включвайки анализ, автоматизация и организационно сътрудничество.

6.2. Източници НА ИНФОРМАЦИЯ: ОТ ОТКРИТИ ДАННИ ДО ТЪМНАТА МРЕЖА

Ефективността на всяка система за кибер разузнаване до голяма степен зависи от нейната способност да интегрира разнообразни, надеждни и навременни източници на информация. Тези източници трябва да отразяват както явните, така и скритите аспекти на заплахите, като осигуряват широка основа за анализ, предвиждане и вземане на решения. Класификацията на източниците на разузнавателна информация традиционно се извършва според степента на достъпност, правния им статус и характера на съдържанието, което предоставят. В съвременния контекст това включва както открити източници (OSINT), така и социални медии (SOCMINT), технически индикатори (TECHINT), човешки източници (HUMINT), както и високорискови среди като тъмната и дълбоката мрежа.

Откритото разузнаване, известно като OSINT, играе съществена роля в изграждането на начален профил на целеви обекти и в идентифицирането на потенциални рискове. То се основава на легално достъпни данни, които могат да бъдат получени от регистри на домейни, публични скенери на мрежи и портове като Shodan и Censys, хранилища за изходен код като GitHub, както и публикации на експерти по сигурност в блогове и форуми. Когато този тип информация бъде комбинирана с автоматизирани инструменти за събиране и анализ, тя предоставя широка картина на атакуемата повърхност на дадена организация, като често служи като входна точка за по-задълбочени разузнавателни дейности.

SOCMINT, или разузнаването от социални мрежи, се фокусира върху събирането на информация от платформи като Twitter/X, Telegram, Discord, Reddit и други. Въпреки че социалните мрежи не са източник на строго технически индикатори, те често съдържат признаци за подготвяни атаки, публикувани уязвимости, дискусии между участници в хакерски

кампании и дори координати на конкретни действия. Telegram, например, се утвърди като предпочитана среда за координиране на деструктивни операции от различни актьори, включително хактивистки групи, поради възможностите си за криптирана комуникация и анонимност на потребителите.

Техническото разузнаване, обозначавано като TECHINT, обхваща конкретни артефакти, които могат да бъдат използвани както за идентификация, така и за блокиране на заплахи. Такива индикатори включват IP адреси, URL адреси, криптографски хешове, SSL сертификати, експлойт кодове, както и конфигурационни елементи, свързани с ботнети или Command and Control сървъри. Информацията, извлечена от TECHINT, намира директно приложение в платформите за откриване и реагиране на инциденти, като се импортира във вид на сигнатури или правила за корелация.

Въпреки значителния напредък в автоматизацията, човешкият фактор остава неотменима част от разузнавателния процес. HUMINT, или разузнаването чрез човешки източници, включва достъп до закрити общности, конференции, неформални мрежи от експерти, както и сътрудничество с бивши участници в престъпни структури или представители на специализирани служби. Този тип разузнаване позволява разкриване на мотивационни и организационни аспекти на атакуващите групи, които често остават недостъпни за автоматизирани системи.

Особено важно, но и високорисково е разузнаването в рамките на тъмната и дълбоката мрежа. Тези среди, достъпни чрез анонимизиращи технологии като Tor и I2P, съчетават както легитимна дейност, така и значително количество нелегални операции. В тъмната мрежа се намират платформи за изтичане на чувствителна информация, форуми за търговия със зловреден софтуер, пазари за трафик на данни и услуги като „достъп до вътрешни мрежи срещу заплащане“. Информацията, достъпна там, е изключително ценна за стратегически анализ, но

събирането ѝ изисква прецизни процедури, правна обоснованост и висок професионализъм, за да се избегнат както етични, така и юридически нарушения. Анализът на тъмната мрежа позволява не само да се идентифицират заплахи в зародиш, но и да се разкрият мрежи от свързани дейности, включително икономическите връзки, които поддържат киберпрестъпността като услуга.

6.3. Инструменти И ПЛАТФОРМИ ЗА КИБЕР РАЗУЗНАВАНЕ

С нарастващия обем и сложност на разузнавателната информация, използването на специализирани платформи и инструменти се превръща в критичен елемент за успеха на всяка CTI стратегия. Възможността за интегриране на множество хетерогенни източници, тяхното нормализиране, обогатяване и визуализация, създава основата за аналитичен капацитет, който надхвърля механичното събиране на данни. Платформите за кибер разузнаване се развиват в три основни направления, събиране и агрегиране на информация, анализ и визуализация, както и автоматизирана интеграция със съществуващи системи за сигурност.

Една от най-разпространените и утвърдени платформи за визуален анализ на зависимости между обекти е Maltego. Тя предоставя възможност за графично представяне на връзките между домейни, IP адреси, имейли, организации и други цифрови обекти чрез трансформации, които извличат информация от публични и частни източници. Този подход позволява бързо идентифициране на скрити връзки и структури, което е особено полезно при разследване на целенасочени атаки и проследяване на кампании с продължителен характер.

Друг важен инструмент е платформата MISP (Malware Information Sharing Platform), която предоставя механизъм за

споделяне на индикатори за заплахи между организации в стандартизиран формат. MISP улеснява създаването на общностно знание и позволява по-бързо реагиране на нововъзникващи рискове. Нейното основно предимство се състои в поддръжката на формати като STIX и OpenIOC, които позволяват интеграция със SIEM и SOAR системи, без необходимост от ръчна обработка на данните.

Сред по-новите, но все по-широко възприемани платформи, е OpenCTI, разработена от френската национална агенция за киберсигурност. Нейната философия се основава на обединяването на технически, тактически и стратегически данни в единен контекстуален модел. OpenCTI предоставя възможност за създаване на виртуални профили на злонамерени актьори, описание на техните кампании, използвана инфраструктура и TTPs (тактики, техники и процедури), като използва моделиране по MITRE ATT&CK. Така се постига не само оперативна осведоменост, но и дългосрочна аналитична стойност, която подпомага стратегическото планиране на мерки за защита.

Съвременните CTI платформи не са самостоятелни инструменти, а елементи от по-широка екосистема, която включва автоматизация, машинно обучение, интеграция със системи за мониторинг и инцидентно реагиране. Те създават възможност за навременно превръщане на суровите данни в приложимо знание, което да служи като основа за информирано, ефективно и отговорно управление на киберсигурността.

6.4. Тъмната МРЕЖА КАТО ИКОНОМИКА НА КИБЕРПРЕСТЪПНОСТТА: СТРУКТУРИ, МОТИВИ, ДИНАМИКА

Тъмната мрежа (Dark Web) е дигитална територия, в която липсата на централизирана регулация, пълната анонимност и използването на криптовалутни механизми създават условия за

процъфтяване на нелегална икономика, изградена върху обмен на киберпрестъпни услуги, данни и инструменти. Тя се различава от „дълбоката мрежа“ (Deep Web), която обхваща ресурси извън обсега на стандартните търсачки, по това, че достъпът до тъмната мрежа изисква специфични технологии, като Tor, I2P или Freenet, и е насочен предимно към скриване на самоличността и дейността на участниците.

Една от най-устойчивите черти на тъмната мрежа е изграждането на паралелна икономическа система, която наподобява легалните пазари, но е ориентирана изцяло към нелегални или високорискови дейности. В тази среда се търгуват разнообразни ресурси, от данни за кредитни карти, идентификационни документи и потребителски акаунти, до експлойт комплекти, zero-day уязвимости и услуги като DDoS атаки по поръчка. Особено разпространен модел е т.нар. „Ransomware-as-a-Service“ (RaaS), при който автори на зловреден код предоставят наемна платформа, с която афилиирани лица могат да атакуват цели срещу процент от откупа. Този модел улеснява навлизането в киберпрестъпната екосистема на лица без сериозни технически умения, като по този начин разширява спектъра и мащаба на заплахите.

Тъмната мрежа функционира чрез утвърдени доверителни механизми, включително рейтингови системи за продавачи, услуги за ескроу (гарантирано посредничество при плащания), криптографски подписи за автентикация и системи за анонимна комуникация. Някои от пазарите имат професионална структура, с интерфейси, подобни на легитимните онлайн магазини, потребителска поддръжка, периодични отстъпки и дори програми за лоялност. Това създава устойчив модел, който трудно се разрушава чрез изолирани интервенции. Затварянето на големи пазари като AlphaBay и Hansa в предходното десетилетие не доведе до срив на системата, а по-скоро до нейното децентрализиране и адаптация.

Кибер разузнаването в контекста на тъмната мрежа изисква не просто технически достъп до тези платформи, а дълбоко разбиране на икономическите, културни и поведенчески особености на участниците. Анализът на форумни дискусии, обяви за продажба, цифрови отпечатъци от инструменти или повторно използвани псевдоними позволява да се картографират мрежи от свързани дейности, да се разкрият афилиации между различни хакерски групи и да се идентифицират нови тенденции. Така например, анализ на пазарната активност в Tor може да разкрие повишена търговия с определен вид експлойт, което да индикира предстояща вълна от атаки, използващи съответната уязвимост.

Важен аспект на анализа е проследяването на криптовалутните транзакции, които, макар и замаскирани чрез миксери и мултистепенни вериги, често оставят следи в публични блокчейни. Комбинирането на криптоанализ с разузнавателни данни от форумите може да осигури ценни доказателства за произхода, обхвата и бенефициентите на дадена атака. По този начин CTI се превръща не само в средство за превенция, но и в инструмент за разследване и събиране на дигитални доказателства, съвместно с правоприлагащи органи.

6.5. Автоматизация И ИЗКУСТВЕН ИНТЕЛЕКТ В ПРОЦЕСИТЕ НА КИБЕР РАЗУЗНАВАНЕ

Ускореното развитие на заплахите и тяхната все по-голяма динамика създават необходимост от автоматизация в обработката и анализа на разузнавателна информация. Тук на преден план излиза ролята на изкуствения интелект (AI) и машинното обучение (ML), които вече не са просто иновация, а необходимост в мащабируемите системи за кибер разузнаване.

Автоматизацията намира приложение във всички етапи на CTI, от събирането на информация чрез уеб скрейпъри и

ботове, през нормализирането на данни и извличането на индикатори, до тяхната класификация, оценка и приоритизиране. Системи, изградени с помощта на ML, могат да класифицират зловредни хостове по поведенчески модели, да разпознават аномалии в логове или да предсказват нови TTPs въз основа на исторически данни. Например, чрез наблюдение на IP адреси с кратък жизнен цикъл, системата може да идентифицира инфраструктура, използвана от fast-flux мрежи, типични за botnet операции.

Допълнително, NLP (Natural Language Processing) технологии се прилагат успешно при анализа на текстове от форуми, чатове и социални медии. Чрез извличане на ключови понятия, семантична класификация и анализ на контекста, системите могат автоматично да идентифицират нови заплахи, да открояят активни групи или да засекат подготовка на атаки. Такива възможности вече са интегрирани в множество комерсиални платформи за CTI, включително Recorded Future, ThreatConnect и IBM X-Force Exchange.

Използването на AI не отменя необходимостта от човешка експертиза. Напротив, то създава нови предизвикателства, свързани с обяснимостта на алгоритмите (explainability), управлението на false positives и необходимостта от непрекъснато обучение на моделите с актуални данни. Освен това, AI не е запазена територия само на защитниците, множество групи също използват ML за автоматизирано избягване на защити, генерация на фишинг съдържание или адаптация на малуер спрямо средата на жертвата. Това създава своеобразна технологична надпревара, в която разузнаването се превръща в игра на надмощие между човешка интуиция и изкуствена прецизност.

6.6. Правни И ЕТИЧНИ АСПЕКТИ НА КИБЕР РАЗУЗНАВАНЕТО В ТЪМНАТА МРЕЖА

Дейността по събиране на информация в контекста на кибер разузнаване, особено когато обхваща ресурси в тъмната мрежа, поставя на преден план сложна съвкупност от правни, етични и регулаторни въпроси, които не могат да бъдат игнорирани в една легитимна и отговорна практика. За разлика от чисто техническите аспекти на анализа, правният контекст е силно зависим от юрисдикцията, характера на организацията (държавна, частна, транснационална) и съществуващите рамки за защита на личните данни, поверителност и цифрови права.

В Европейския съюз, например, Общият регламент за защита на личните данни (GDPR) въвежда строги ограничения относно обработката на лична информация, дори когато тя е открита чрез OSINT или извлечена от компрометирани бази данни в тъмната мрежа. Това означава, че самият факт, че дадени данни са били публикувани от зловредни актьори, не отменя задължението за законосъобразна обработка, прозрачност и минимизация на събираната информация. Нарушения в този контекст могат да доведат до санкции не само от регулаторните органи, но и до тежки репутационни щети.

Допълнително, някои от техниките, използвани в практиката на CTI, като достъп до закрити форуми, използване на фалшиви самоличности, симулирани атаки или събиране на данни чрез скриптове в Tor мрежата, попадат в сивата зона между допустимото и недопустимото поведение. В редица страни, дори и пасивното наблюдение в нелегитимни среди може да се тълкува като съучастие, особено ако не е придружено от ясни вътрешни политики, юридическа обосновка и сътрудничество със съответните органи. Оттук произтича необходимостта всяка организация, извършваща дейности в сферата на CTI, да приеме вътрешен етичен кодекс, в който да се дефинират границите на допустимото поведение, процедурите за

одобрение на по-рискови действия и механизмите за контрол и отчетност.

Етичният въпрос не е само формален. В основата му стои напрежението между правото на защита и правото на неприкосновеност. Един добре функциониращ разузнавателен капацитет може да предотврати сериозни атаки, но ако е изграден върху съмнителни практики, той може да компрометира легитимността на самата организация. Например, използването на инфилтрирани агенти в тъмната мрежа, които участват в дискусии с цел събиране на информация, създава морална и юридическа отговорност, ако те неволно допринесат за реализацията на престъпни действия. Затова все по-често се препоръчва прилагането на принципа на пропорционалност, т.е. оценка дали дадено действие е необходимо, минимално инвазивно и балансирано спрямо очакваната полза.

На международно равнище липсва хармонизиран подход относно дейността в тъмната мрежа. Докато някои държави толерират агресивни методи от страна на частните структури за сигурност, други въвеждат ограничения, които изискват CTI да се извършва само от акредитирани субекти или в сътрудничество с разследващи органи. В тази разнородна правна среда е критично важно организациите да изградят собствен капацитет за правна експертиза, както и да поддържат активен диалог между техническите и юридическите екипи.

Следователно, правните и етични аспекти не са периферен въпрос, а фундаментален елемент от устойчивата стратегия за кибер разузнаване. Те изискват интегриран подход, в който техническата ефективност е балансирана от правна съобразност и етична отчетност. Само така CTI може да изпълни своята роля като инструмент за легитимна защита и доверие, а не като причина за нови рискове.

6.7. Кибер РАЗУЗНАВАНЕТО КАТО ОРГАНИЗАЦИОННА ФУНКЦИЯ: ИНТЕГРАЦИЯ, ЗРЕЛОСТ И ВЪЗДЕЙСТВИЕ

Извън чисто техническата си същност, кибер разузнаването трябва да се разглежда като организационна способност, съвкупност от процеси, роли, технологии и култура, които превръщат информацията в стратегическо предимство. Това предполага неговата интеграция не само в отдел „Сигурност“, но и в процесите на управление на риска, вземане на решения, разработка на продукти и комуникация с външни заинтересовани страни.

В зрели организации кибер разузнаването е вградено в структурата на Security Operations Center (SOC) или функционира като самостоятелна аналитична единица с пряка отчетност към ръководството. Основната функция на тази единица е да доставя навременна, контекстуализирана и приложима информация за заплахите, не като изолирани инциденти, а като част от цялостния риск профил на организацията. Така например, при идентифициране на индикатор за компрометиране в тъмната мрежа, CTI екипът може не само да уведоми екипа по сигурност, но и да подготви ръководството за комуникационна стратегия, да анализира потенциални регулаторни последствия и да оцени вероятното въздействие върху веригата на доставки.

Зрелостта на една CTI функция може да се измерва по няколко критерия: ниво на автоматизация; степен на интеграция с платформи като SIEM, SOAR и XDR; способност за корелация между външни и вътрешни източници; участие в информационни общности и инициативи за споделяне на заплахи (ISACs); както и наличието на утвърден аналитичен процес, основан на методологии като MITRE ATT&CK, Diamond Model или Kill Chain Framework. Важно е също така дали резултатите от разузнаването водят до реални действия – като обновяване на правила за защита, приоритизиране на уязвимости, или вземане на стратегически решения относно инвестиции в нови технологии.

Често срещан проблем е т.нар. „информационен шум“, прекомерно количество слабо релевантни сигнали, които затрудняват фокусираното реагиране. Ефективният CTI екип не е просто „агрегатор“ на данни, а интерпретатор, който умее да различава съществените от второстепенното, да разбира контекста на заплахите и да предлага конкретни препоръки. Това изисква висококвалифицирани кадри, способни да комбинират техническа експертиза с аналитично мислене, езикова и културна компетентност (особено при работа с международни източници), както и капацитет за писмено и визуално представяне на сложна информация пред различни аудитории.

В крайна сметка, CTI трябва да бъде разпознат не само като компонент от ИТ сигурността, а като стратегически актив, чието въздействие се простира от предотвратяването на атаки до изграждането на доверие у партньори, клиенти и регулатори. В една все по-нелинейна и рискова цифрова среда, разузнаването е не просто отговор, а предпоставка за предвиждане и адаптация.

6.8. Изводи

Кибер разузнаването в съвременната му форма представлява не просто реактивен инструмент за справяне със заплахи, а многопластов аналитичен процес, който свързва данните, технологиите и стратегическото управление в единен механизъм за дигитална устойчивост. Настоящият анализ на ролята на CTI и неговото взаимодействие с тъмната мрежа показа, че ефективната защита в киберпространството не може да бъде постигната единствено чрез технически средства. Необходим е холистичен подход, който интегрира проактивна осведоменост, юридическа съобразност, организационна готовност и аналитична култура.

В процеса на развитие на CTI като дисциплина, се очертават няколко основни тенденции, които трансформират както обхвата, така и функциите на разузнаването. Първо,

наблюдава се прогресивно разширяване на източниковата база, от публични ресурси до високорискови, но изключително информативни среди в тъмната мрежа. Това обогатяване на информацията води до по-добро разбиране на атакуващите групи, техните мотивации, структури и икономическа логика. Второ, изкуственият интелект и автоматизацията изместват границите на възможното в анализа, позволяват обработка на големи обеми от необработени данни, откриване на скрити зависимости и предсказване на бъдещи заплахи. В този контекст, човешката роля не изчезва, а се трансформира, от оператор към стратегически анализатор, способен да разчита поведението на сложни и често нелинейни системи.

В същото време обаче, нарастват изискванията за правна отговорност и етично поведение в рамките на разузнавателните дейности. Кибер сигурността не е изолирана сфера, а част от социалната и правна реалност, в която всяко действие може да има последствия отвъд непосредствения технически ефект. Следователно, зрелите разузнавателни структури не само познават ограниченията на своята дейност, но и изграждат вътрешни механизми за контрол, отчетност и легитимност. Включването на юридически експерти в процеса, разработването на вътрешни политики и участието в междуорганизационни формати за обмен на добри практики е не просто препоръчително, а наложително.

На организационно ниво, кибер разузнаването все по-ясно се позиционира като стратегически ресурс. То не е самостоятелна функция, а структурен компонент на модерното управление на сигурността. Неговата стойност се проявява не само в предотвратени атаки, но и в способността да подпомага вземането на информирани решения, да открива слабости в екосистемата и да изгражда дигитална хигиена, базирана на осведоменост, адаптация и непрекъснато учене. В този смисъл,

CTI е не просто реакция на риска, а основа за неговото преосмисляне.

Изграждането на устойчив разузнавателен капацитет изисква време, ресурси и междуфункционално сътрудничество. Това е инвестиция, чиито резултати не винаги са незабавни, но която гарантира стратегическа преднина в един все по-конкурентен и рисков киберсвят. Само чрез синергия между технологии, хора и процеси, организациите могат да отговорят на динамиката на заплахите с ефективност, скорост и етика.

В заключение, кибер разузнаването и анализът на тъмната мрежа не са просто елементи от техническия арсенал на съвременната сигурност, те са фундаментален слой на цифровата отбрана, в който се срещат знанието, технологията и отговорността. Прилагано в контекста на ясно дефинирани принципи и организационна зрелост, CTI може не само да предотвратява атаки, но и да преобрази културата на сигурността отвътре.

7. ИЗКУСТВЕН ИНТЕЛЕКТ В КИБЕРСИГУРНОСТТА: ВЪЗМОЖНОСТИ, ОГРАНИЧЕНИЯ И АТАКИ СРЕЩУ AI СИСТЕМИ

7.1. Ролята на НА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ В СЪВРЕМЕННАТА КИБЕРСИГУРНОСТ

Динамичното развитие на цифровите технологии и нарастващата сложност на киберзаплахите поставят под съмнение ефективността на традиционните, изцяло сигнатурно базирани подходи за защита. Съвременните атаки все по-често се характеризират с висока степен на адаптивност, автоматизация и целенасоченост, което значително затруднява тяхното своевременно откриване чрез класически средства. В този контекст изкуственият интелект и машинното обучение се утвърждават като ключови технологични компоненти за изграждане на следващо поколение системи за киберсигурност.

Основното предимство на AI-базираните решения е способността им да анализират големи обеми от хетерогенни данни в реално време и да откриват сложни зависимости и аномалии, които биха останали незабелязани при ръчен или статичен анализ. Логове от мрежов трафик, събития от крайни точки, потребителско поведение и телеметрични данни могат да бъдат корелирани автоматизирано, което позволява по-ранно идентифициране на потенциални инциденти и по-точна оценка на риска.

В системите за управление на сигурността (SIEM), за откриване и реагиране при инциденти (EDR/XDR), както и в центровете за операции по сигурността (SOC), AI постепенно се превръща от допълващ инструмент в основен аналитичен механизъм. Чрез поведенчески анализ и моделиране на нормално състояние на системите, алгоритмите могат да

разпознават отклонения, които са индикатор за компрометиране, дори при липса на известни сигнатури или предварително дефинирани правила.

Същевременно, нарастващото внедряване на AI в киберсигурността поражда и нови предизвикателства. Алгоритмите са силно зависими от качеството и представителността на данните, с които се обучават, а сложността на моделите често води до ограничена прозрачност при вземането на решения. Това поставя въпроси, свързани с доверие, обяснимост и контрол, особено в критични инфраструктури и среди с високи регулаторни изисквания.

Използването на изкуствен интелект в киберсигурността следователно не може да се разглежда единствено като технологично усъвършенстване, а като фундаментална промяна в подхода към защитата – от реактивни механизми към предиктивни и адаптивни модели. Именно тази трансформация налага задълбочен анализ както на възможностите, така и на рисковете, произтичащи от интеграцията на AI в сигурността на информационните системи.

7.2. Основни модели на ИЗКУСТВЕН ИНТЕЛЕКТ И МАШИННО ОБУЧЕНИЕ, ИЗПОЛЗВАНИ В КИБЕРСИГУРНОСТТА

Използването на изкуствен интелект в киберсигурността се основава предимно на техники от областта на машинното обучение, които позволяват автоматизирано извличане на закономерности от големи масиви от данни. В зависимост от начина на обучение и целите на анализа, тези модели се прилагат за различни задачи, като откриване на аномалии, класификация на събития, прогнозиране на поведение и подпомагане на процесите по вземане на решения в системите за сигурност.

Най-широко разпространени в практиката са моделите за надзиравано обучение, при които алгоритмите се обучават върху предварително етикетирани данни. В контекста на киберсигурността това включва примери за зловреден и легитимен трафик, известни атаки и нормално поведение на системите. Тези модели се използват успешно за класификация на зловреден софтуер, разпознаване на фишинг съобщения и детекция на познати модели на атаки. Основното им ограничение произтича от зависимостта от качеството и актуалността на обучаващите данни, което ги прави по-малко ефективни срещу нови и неизвестни заплахи. От особено значение за съвременната киберзащита са моделите за ненадзиравано обучение, които не изискват предварително етикетирани данни и се използват за откриване на аномалии. Чрез изграждане на модел на „нормално“ поведение на мрежата, системите или потребителите, тези алгоритми могат да идентифицират отклонения, които често са индикатор за компрометиране. Този подход е особено подходящ за среди с динамично поведение и голям обем от данни, каквито са корпоративните мрежи и центровете за операции по сигурността. Именно тук се наблюдава пряка връзка с разгледаните в предходните глави SIEM и UEBA системи, които използват подобни модели за корелация и анализ на събития.

В допълнение, дълбокото обучение (deep learning) намира все по-широко приложение в анализ на сложни и неструктурирани данни, включително мрежов трафик, бинарни файлове и текстово съдържание. Невронните мрежи с множество слоеве позволяват автоматично извличане на характеристики, без необходимост от ръчно дефинирани правила, което повишава ефективността при идентифициране на сложни атаки и прикрити зловредни активности. Въпреки това, тези модели често страдат от ограничена обяснимост, което поставя предизвикателства при тяхното използване в среди с високи изисквания за отчетност и контрол.

В контекста на киберразузнаването, разгледано в предходната глава, AI моделите се използват и за анализ на разузнавателна информация, включително автоматизирана обработка на текстови източници, форуми и данни от тъмната мрежа. Чрез техники за обработка на естествен език и клъстеризация, системите могат да идентифицират нови кампании, свързани актьори и тенденции в поведението на заплашителните групи. По този начин AI се превръща в ключов инструмент за трансформиране на сурови данни в приложима разузнавателна информация.

7.3. Приложения на ИЗКУСТВЕНИЯ ИНТЕЛЕКТ В ЗАЩИТАТА НА ИНФОРМАЦИОННИТЕ СИСТЕМИ

Практическото приложение на изкуствения интелект в киберсигурността обхваща широк спектър от защитни функции, които целят да повишат ефективността и скоростта на реакция при инциденти. В съвременните архитектури за сигурност AI е интегриран като аналитичен слой, който подпомага автоматизацията и намалява натоварването на човешките оператори.

В системите за управление на събития и информация по сигурността изкуственият интелект се използва за интелигентна корелация на събития от различни източници, включително мрежови устройства, крайни точки и приложения. Чрез анализ на времеви зависимости и поведенчески модели, AI-базираните SIEM платформи могат да намалят броя на фалшивите аларми и да приоритизират инцидентите с най-висок риск. Това е от ключово значение за ефективната работа на SOC екипите, особено в среди с ограничени ресурси и голям обем от сигнали.

При защитата на крайни точки и мрежи, AI намира приложение в EDR, NDR и XDR решения, където поведенческият анализ позволява откриване на атаки, които не оставят ясни

сигнатури. Наблюдението на процеси, системни повиквания и мрежови връзки в реално време дава възможност за ранно засичане на компрометиране и автоматизирано предприемане на мерки за ограничаване на щетите. Тези механизми се допълват от автоматизирани реакции, които могат да изолират засегнати системи или да блокират подозрителен трафик без човешка намеса.

Изкуственият интелект играе съществена роля и в автоматизацията на процесите по реагиране при инциденти чрез SOAR платформи. Чрез предварително дефинирани сценарии и адаптивни алгоритми, системите могат да координират действия между различни защитни инструменти, като по този начин се постига по-бързо и последователно реагиране. Това е в пряка връзка с концепциите за симулация и обучение, разгледани в главата за кибер полигоните, където AI може да се използва за генериране на динамични сценарии и оценка на ефективността на защитните реакции.

Не на последно място, AI намира приложение и в управлението на идентичности и достъп, където чрез анализ на потребителско поведение се откриват аномалии, свързани с компрометирани акаунти или вътрешни заплахи. Този подход допринася за по-фино управление на риска и допълва традиционните механизми за автентикация и контрол на достъпа.

Въпреки значителните предимства, практическото внедряване на AI в защитата на информационните системи изисква внимателно планиране, интеграция с вече съществуващите процеси и постоянен контрол върху качеството на входните данни. В противен случай съществува риск автоматизацията да доведе до погрешни решения или до фалшиво усещане за сигурност. Именно тези ограничения отварят пътя към следващия аспект на изследването – анализът на уязвимостите и атаките, насочени директно срещу AI системите, използвани в киберсигурността.

7.4. Ограничения и рискове при използване на изкуствен интелект в киберсигурността

Въпреки значителния потенциал на изкуствения интелект за повишаване на ефективността на защитните механизми, неговото приложение в киберсигурността е съпроводено с редица ограничения и рискове, които не могат да бъдат пренебрегвани. Едно от основните предизвикателства произтича от силната зависимост на AI моделите от качеството, обема и представителността на данните, използвани за обучение. Непълни, остарели или небалансирани набори от данни могат да доведат до неточни модели, които генерират фалшиви положителни или фалшиви отрицателни резултати, компрометирайки процесите по откриване и реагиране при инциденти.

Друг съществен риск е свързан с динамичния характер на киберзаплахите. Атакуващите постоянно променят своите тактики, техники и процедури, което води до т.нар. „деградация на модела“ с течение на времето. Ако AI системите не бъдат регулярно преобучавани и адаптирани към новите условия, тяхната ефективност може значително да намалее. Това поставя високи изисквания към процесите по поддръжка и управление на AI решенията и ги прави неразделна част от цялостната стратегия за сигурност, а не автономни инструменти.

Особено внимание заслужава и проблемът с обяснимостта на AI моделите. Много от използваните алгоритми, особено в областта на дълбокото обучение, функционират като т.нар. „черни кутии“, при които е трудно или дори невъзможно да се обясни по какъв начин е взето дадено решение. В контекста на киберсигурността това поражда сериозни въпроси, свързани с доверието в системите, възможността за одит и съответствието с регулаторни изисквания. Липсата на прозрачност може да

затрудни анализа на инциденти и да усложни процесите по отчетност, особено в среди с критична инфраструктура.

Допълнителен риск произтича от автоматизацията на реакциите. Макар автоматизираните действия да съкращават времето за реакция, те могат да доведат и до нежелани последствия, ако бъдат задействани на базата на неточен анализ. Автоматично изолиране на системи, блокиране на потребители или прекъсване на услуги може да окаже сериозно влияние върху бизнес процесите, ако решението е взето погрешно. Това подчертава необходимостта от внимателно балансиране между автоматизация и човешки контрол.

В контекста на разгледаните в предходните глави кибер полигони и симулационни среди, тези ограничения могат да бъдат адресирани чрез контролирано тестване и валидиране на AI-базирани решения. Симулирането на различни сценарии позволява оценка на поведението на моделите при необичайни условия и подпомага идентифицирането на слабости, преди те да бъдат експлоатирани в реална среда.

7.5. Атаки срещу AI системи в киберсигурността

С нарастващото внедряване на изкуствен интелект в защитните системи, самите AI модели се превръщат в цел на атаки, което разширява повърхността за компрометиране на киберсигурността. Този клас атаки, често обобщавани под термина „adversarial machine learning“, представлява съзнателен опит за манипулиране на поведението на моделите с цел заобикаляне, подкопаване или компрометиране на защитните механизми.

Един от основните подходи при атаки срещу AI системи е манипулирането на входните данни по начин, който води до погрешни изводи. Чрез внимателно подбрани изменения в мрежовия трафик, файловете или поведението на процесите,

атакуващите могат да предизвикат погрешна класификация, без това да бъде лесно забелязано от системата. Този тип атаки са особено опасни, тъй като могат да позволят на зловредна активност да остане незабелязана, дори при наличие на модерни AI-базирани защитни решения.

Друг съществен риск е свързан с компрометиране на процеса на обучение на моделите. Ако атакуващият успее да повлияе върху данните, използвани за обучение или преобучаване, резултатът може да бъде модел, който систематично допуска определен тип атаки или генерира некоректни аларми. Подобни сценарии са особено релевантни в среди, където AI системите се обучават автоматично върху данни от продукционна среда, без достатъчно механизми за валидация.

Съществуват и атаки, насочени към извличане на информация за самия модел, неговата структура или параметри. Чрез анализ на изходите на системата и наблюдение на нейните реакции, атакуващите могат да направят изводи за начина, по който тя функционира, и да използват тази информация за по-прецизно заобикаляне на защитата. Това е особено проблематично в облачни среди и при използване на AI като услуга, където моделите са достъпни чрез програмни интерфейси.

В контекста на киберразузнаването, разгледано в предходните глави, подобни атаки подчертават необходимостта от интегриране на информация за уязвимости в AI моделите като част от цялостната разузнавателна картина. Познаването на техниките за атака срещу машинното обучение позволява по-точно моделиране на заплахите и разработване на защитни стратегии, които отчитат не само традиционните вектори на атака, но и специфичните рискове, свързани с използването на AI.

Атаките срещу AI системите ясно показват, че изкуственият интелект не представлява универсално решение за всички проблеми на киберсигурността. Напротив, той въвежда

нов слой на сложност, който изисква задълбочено разбиране, постоянен мониторинг и интегриран подход към защитата. Именно това налага разглеждането на етичните, правните и стратегическите аспекти на използването на AI.

7.6. Злонамерено използване на изкуствения интелект от атакуващите

Паралелно с внедряването на изкуствен интелект в защитните системи, същите технологии все по-активно се използват и от атакуващите, което води до качествена промяна в характера на киберзаплахите. Изкуственият интелект предоставя на злонамерените актьори възможности за автоматизация, мащабируемост и адаптивност, които значително надхвърлят традиционните ръчни или скриптово базирани атаки. Това превръща AI не само в инструмент за защита, но и в мощно оръжие за извършване на по-прецизни, убедителни и трудни за откриване атаки.

Едно от най-очевидните направления на злонамерено използване на AI е автоматизацията на социалното инженерство. Чрез модели за обработка и генериране на естествен език, атакуващите могат да създават фишинг съобщения с висока степен на персонализация, съобразени с езика, контекста и поведението на конкретни жертви. Това значително намалява ефективността на традиционните защитни механизми и обучения, които често разчитат на разпознаване на шаблонни или граматически неточни съобщения. В комбинация със събрана разузнавателна информация от социални мрежи и публични източници, AI-базираният фишинг се превръща в целенасочен инструмент за компрометиране на конкретни организации или ключови служители.

Изкуственият интелект се използва и за усъвършенстване на зловредния софтуер, като позволява динамично адаптиране

на поведението му спрямо защитната среда. Чрез анализ на реакциите на антивирусни и EDR системи, злонамерените програми могат да променят своите характеристики, за да избегнат откриване. Този подход е особено ефективен срещу сигнатурно базирани и дори някои поведенчески механизми, тъй като атаката не следва статичен модел, а се развива в зависимост от конкретната среда.

Допълнително, AI се използва за автоматизирано разузнаване и откриване на уязвимости. Чрез анализ на публично достъпни данни, конфигурации и поведение на системи, атакуващите могат да идентифицират слаби места в инфраструктурата много по-бързо и в по-голям мащаб, отколкото при традиционните методи. Това има пряка връзка с разгледаните в предходните глави теми за киберразузнаване и тъмната мрежа, където AI подпомага обработката на големи обеми от разузнавателна информация и откриването на закономерности, които биха останали скрити за човешкия анализатор.

Особено притеснително направление е използването на AI за създаване на дълбоки фалшификации (deepfakes), които могат да бъдат използвани за компрометиране на доверие, манипулиране на процеси на вземане на решения или за извършване на измами. Генерирането на реалистични аудио и видео материали позволява симулиране на самоличността на ръководители или служители с високи привилегии, което увеличава риска от успешни атаки чрез социално инженерство и вътрешни злоупотреби. Този тип заплахи поставят под въпрос традиционните механизми за удостоверяване и изискват нови подходи за верификация на идентичността и автентичността на комуникацията.

Злонамереното използване на AI също така допринася за мащабирането на атаките. Автоматизираното генериране и изпълнение на атаки позволява едновременно насочване към

голям брой цели, като същевременно се запазва висока степен на персонализация. Това води до нарастване на натоварването върху защитните системи и SOC екипите, които са изправени пред значително по-голям обем от сложни инциденти, изискващи анализ и реакция.

В този контекст, използването на изкуствен интелект от атакуващите подчертава необходимостта от симетричен отговор от страна на защитата. Организациите следва да осъзнаят, че AI не е еднопосочен инструмент за подобряване на сигурността, а част от ескалираща технологична надпревара. Това налага интегриране на знания за злонамереното използване на AI в процесите по управление на риска, обучение на персонала и разработване на защитни стратегии. Кибер полигоните и симулационните среди, разгледани в предходните глави, могат да играят ключова роля в подготовката за подобни заплахи, като позволяват тестване на реакции срещу AI-базирани атаки в контролирана среда.

Злонамереното използване на изкуствения интелект ясно показва, че бъдещето на киберсигурността ще бъде белязано от все по-висока степен на автоматизация и интелигентност и от двете страни. Това не е просто технологично съревнование, а промяна в самия характер на заплахите, като те стават по-адаптивни, по-мащабни и по-трудни за предвиждане. Затова организациите не могат да разчитат единствено на традиционни защитни механизми или на статични правила. Необходимо е непрекъснато обучение, обмен на информация за новите методи на атака и изграждане на вътрешен капацитет за бърза реакция. В крайна сметка, въпросът не е дали AI ще се използва срещу нас, а колко добре сме подготвени да го посрещнем.

7.7. Етични, правни и регулаторни аспекти при използването на изкуствения интелект в киберсигурността

Интеграцията на изкуствения интелект в системите за киберсигурност поражда не само технологични, но и съществени етични и правни въпроси, които следва да бъдат разглеждани като неразделна част от процеса на внедряване и управление на тези технологии. В контекста на автоматизираното вземане на решения, особено когато те засягат достъпа до системи, обработката на лични данни или прекъсването на услуги, възниква необходимост от ясно дефинирани принципи за отговорност, прозрачност и контрол.

Един от основните етични проблеми е свързан с ограничената обяснимост на решенията, вземани от сложни AI модели. Когато дадена система автоматично класифицира поведение като злонамерено или предприема защитни действия, често е трудно да се предостави ясен и разбираем аргумент за това решение. Това може да доведе до недоверие в системите, затруднения при одитиране и проблеми при разследване на инциденти. В среди с високи регулаторни изисквания, като публичния сектор или критичната инфраструктура, липсата на обяснимост може да бъде сериозно ограничение за използването на AI.

От правна гледна точка, използването на изкуствен интелект в киберсигурността трябва да бъде съобразено с действащите нормативни рамки за защита на личните данни, информационната сигурност и правата на субектите на данни. Автоматизираният анализ на потребителско поведение, мрежов трафик и комуникации поставя въпроси относно пропорционалността на мерките и границите между сигурността и неприкосновеността на личния живот. Това налага въвеждане на ясни политики за обработка на данни, минимизиране на

събираната информация и осигуряване на адекватни механизми за контрол и отчетност.

В европейски контекст, развитието на регулаторни инициативи, насочени към изкуствения интелект, допълнително усложнява средата, в която се внедряват AI-базирани защитни решения. Класифицирането на определени приложения на AI като високорискови изисква организациите да прилагат допълнителни мерки за управление на риска, документация и човешки надзор. В този смисъл, използването на AI в киберсигурността не може да бъде изолирано от общата рамка на корпоративното управление и съответствието с регулаторните изисквания.

Етичните аспекти са тясно свързани и с въпроса за злоупотреба с технологиите. Както беше разгледано в предходното подразделение, същите AI инструменти, които подпомагат защитата, могат да бъдат използвани и за извършване на по-ефективни атаки. Това налага отговорно отношение при разработването и разпространението на AI решения, както и необходимост от международно сътрудничество и обмен на добри практики. В този контекст, киберразузнаването и споделянето на информация между организации и институции се утвърждават като ключови елементи за намаляване на рисковете, свързани с неконтролираното използване на AI.

7.8. Мястото на изкуствения интелект в бъдещите модели за киберзащита

Разгледаните възможности, ограничения и рискове ясно показват, че изкуственият интелект ще заема все по-централно място в бъдещите модели за киберзащита, но не като самостоятелно решение, а като интегриран компонент на цялостна, многослойна архитектура за сигурност. Ефективната

защита на информационните системи ще изисква синергия между AI-базирани аналитични механизми, традиционни защитни технологии, човешка експертиза и стратегическо управление на риска.

В бъдещите архитектури за сигурност се очаква изкуственият интелект да играе ключова роля в предиктивния анализ и ранното откриване на заплахи. Чрез комбиниране на данни от оперативни системи, киберразузнавателни източници и симулационни среди, AI може да подпомага не само откриването на текущи инциденти, но и прогнозиране на потенциални сценарии за атака. Това представлява естествено продължение на концепциите за проактивна сигурност, разгледани в предходните глави на монографията.

Това означава, че AI постепенно се превръща от инструмент за реакция в инструмент за превенция. Вместо само да отговаря на случилото се, системата може да предвижда вероятни насоки на атака и да подготвя защитата предварително. Разбира се, прогнозирането в киберпространството никога няма да бъде напълно точно, понеже заплахите се променят бързо, а нападателите също се учат. Но дори и приблизителната прогноза дава огромно предимство: спечелено време, по-добра приоритизация на ресурсите и възможност да се спрат атаки в ранен стадий, понякога още преди да са започнали.

Особено важно е мястото на AI в контекста на интегрираните платформи за сигурност, които обединяват функции за мониторинг, анализ и реакция. В такива среди изкуственият интелект може да служи като „свързващ елемент“, който осигурява контекстуализация на събитията и подпомага вземането на решения на стратегическо ниво. Това включва и подпомагане на процесите по управление на кибер риска, където AI може да допринася за по-точна оценка на въздействието и приоритизация на защитните мерки.

Важен аспект на бъдещото развитие е и ролята на кибер полигоните като среда за обучение, тестване и валидиране на AI-базирани защитни решения. Чрез симулиране на реалистични сценарии и атаки, организациите могат да оценят поведението на AI моделите в контролирана среда и да идентифицират потенциални слабости, преди те да бъдат експлоатирани в реални условия. Това допринася за изграждането на доверие в системите и за по-ефективно интегриране на AI в оперативната дейност.

В заключение, бъдещето на киберсигурността ще бъде белязано от все по-тясно взаимодействие между човека и машината. Изкуственият интелект ще предоставя мощни инструменти за анализ и автоматизация, но неговата ефективност ще зависи от качеството на данните, яснотата на процесите и отговорното управление на технологиите. Само чрез интегриран, етично обоснован и стратегически ориентиран подход AI може да се превърне в устойчив фактор за повишаване на сигурността, а не в източник на нови уязвимости.

Стратегически препоръки

На основата на извършеното изследване и анализа на съвременните практики в областта на киберсигурността могат да бъдат формулирани редица стратегически препоръки, насочени към организациите, които се стремят към изграждане на устойчива, адаптивна и интелигентна система за защита. В условията на динамично развиваща се заплашителна среда и нарастваща автоматизация на атаките, сигурността следва да се разглежда не като статично състояние, а като непрекъснат процес на управление на риска, основан на данни, разузнаване, симулации и информирано вземане на решения.

На първо място, организациите следва целенасочено да изградят и институционализират функцията по кибер разузнаване като неразделна част от цялостната архитектура за сигурност. Независимо дали чрез вътрешен екип или чрез стратегически партньорства, CTI трябва да бъде интегрирано в оперативните и управленските процеси, а не да се използва единствено реактивно при възникване на инциденти. Ефективното разузнаване изисква систематично използване на разнообразни източници от отворени данни и социални платформи до технически индикатори и контролирано наблюдение на тъмната мрежа, както и ясно дефинирани методологии за анализ, валидиране и разпространение на разузнавателните продукти вътре в организацията.

На второ място, стратегическата рамка за сигурност следва да бъде подкрепена от регулярни тестове, симулации и практически упражнения, които отразяват реалистични сценарии на атаки. Провеждането на тестове за проникване, Red Team и Blue Team дейности не бива да се възприема като формално изискване за съответствие, а като ключов механизъм за проверка на готовността на технологиите, процесите и хората. Особено важно е тези упражнения да включват сценарии, които отчитат

съвременните тенденции, като автоматизирани атаки, кампании с използване на изкуствен интелект, атаки по веригата на доставки и комбинирани хибридни заплахи. Само чрез редовно тестване и анализ на резултатите организациите могат да изградят реалистична представа за собствените си способности и ограничения.

На трето място, инвестициите в кибер полигони или достъпът до подобни симулационни среди представляват стратегически актив за дългосрочно повишаване на зрелостта на сигурността. Кибер полигоните предоставят уникална възможност за обучение на екипи, валидиране на политики и процедури, както и експериментиране с нови защитни технологии, включително AI-базирани решения. В контекста на нарастващата сложност на изкуствения интелект, тези среди позволяват контролирано тестване на поведението на алгоритмите при различни сценарии и подпомагат изграждането на доверие и разбиране за техните възможности и ограничения.

На четвърто място, стратегическото внедряване на изкуствен интелект в киберсигурността следва да бъде съпътствано от ясно дефинирани принципи за управление, етика и човешки контрол. Организациите не бива да възприемат AI като универсално решение, а като инструмент, който изисква постоянен надзор, качествени данни и прозрачност при вземането на решения. Особено внимание следва да се обърне на рисковете, свързани с атаки срещу AI системите, деградация на моделите и автоматизирани реакции с потенциално високо въздействие върху бизнес процесите. В този смисъл, комбинирането на автоматизация с експертна човешка намеса остава ключов принцип за устойчивата киберзащита.

На пето място, в условията на засилена регулация и обществена чувствителност към защитата на данните, всички дейности в областта на киберсигурността и киберразузнаването следва да бъдат юридически обосновани и етично допустими.

Това изисква тясно взаимодействие между техническите, правните и управленските екипи, разработване на вътрешни политики и процедури, както и въвеждане на механизми за отчетност и контрол. Особено важно е при използването на разузнавателни източници и AI-базирани анализи да се спазват принципите за защита на личните данни и да се избягват действия, които биха могли да доведат до правни или репутационни рискове.

На следващо място, активното участие в общности за споделяне на информация и добри практики следва да бъде насърчавано като стратегически приоритет. Чрез сътрудничество с национални и международни CERT структури, индустриални съюзи и специализирани общности за обмен на информация за заплахи, организациите могат значително да подобрят своята осведоменост и способност за реакция. В условията на координирани и трансгранични атаки, изолираните подходи към сигурността са неефективни и увеличават уязвимостта на отделните участници.

Не на последно място, сигурността следва да бъде възприемана като стратегическа функция на най-високо управленско ниво. Това предполага активна ангажираност на ръководството, интеграция на киберразузнавателните изводи в процесите по вземане на решения и включване на киберриска в общата рамка на корпоративното управление. Само чрез подобен холистичен подход е възможно изграждането на устойчиви организационни модели, способни да се адаптират към бързо променящата се цифрова среда.

Внедряването на тези стратегически препоръки не следва да се разглежда като еднократна инициатива, а като непрекъснат процес на изграждане и поддържане на зрялост в киберсигурността. В свят, в който заплахите се развиват с експоненциална скорост и все по-често се подпомагат от интелигентни технологии, именно адаптивността, проактивността

и знанието се утвърждават като ключови фактори за устойчиво и сигурно цифрово развитие.

Библиография

1. Anderson R., Security Engineering: A Guide to Building Dependable Distributed Systems, Wiley, ISBN 978-1119642787 (2020)
2. Bishop M., Computer Security: Art and Science, Addison-Wesley, ISBN 978-0321712332 (2018)
3. Stallings W., Cryptography and Network Security: Principles and Practice, Pearson, ISBN 978-0134444284 (2017)
4. Whitman M., Mattord H., Principles of Information Security, Cengage Learning, ISBN 978-0357508330 (2021)
5. Behl A., Cyberwarfare: Threats, Security, and Resilience, Springer, ISBN 978-3030191936 (2019)
6. Scarfone K., Mell P., Guide to Intrusion Detection and Prevention Systems (IDS/IPS), NIST SP 800-94, NIST (2007)
7. Cichonski P. et al., Computer Security Incident Handling Guide, NIST SP 800-61 Rev.2, NIST (2012)
8. Souppaya M., Scarfone K., Guide to Malware Incident Prevention and Handling, NIST SP 800-83, NIST (2013)
9. Kent K., Souppaya M., Guide to Computer Security Log Management, NIST SP 800-92, NIST (2006)
10. ISO/IEC, ISO/IEC 27001: Information Security Management Systems, International Organization for Standardization (2022)
11. ISO/IEC, ISO/IEC 27002: Information Security Controls, International Organization for Standardization (2022)
12. ENISA, Threat Landscape Report, European Union Agency for Cybersecurity (2023)
13. ENISA, Cyber Ranges: From Education to National Defense, ENISA Report (2020)
14. MITRE Corporation, MITRE ATT&CK Framework, MITRE (ongoing)
15. Verizon, Data Breach Investigations Report (DBIR), Verizon Enterprise (2024)
16. Mandiant, M-Trends Report, Google Cloud Security (2024)
17. CrowdStrike, Global Threat Report, CrowdStrike (2024)
18. IBM Security, X-Force Threat Intelligence Index, IBM (2023)
19. Kaspersky, Threat Landscape Report, Kaspersky Labs (2023)
20. Shostack A., Threat Modeling: Designing for Security, Wiley, ISBN 978-1118809990 (2014)

21. McGraw G., Software Security: Building Security In, Addison-Wesley, ISBN 978-0321356703 (2006)
22. Skoudis E., Liston T., Counter Hack Reloaded, Prentice Hall, ISBN 978-0137155880 (2011)
23. Casey E., Digital Evidence and Computer Crime, Academic Press, ISBN 978-0123742681 (2011)
24. Nelson B., Phillips A., Steuart C., Guide to Computer Forensics and Investigations, Cengage Learning, ISBN 978-1337568944 (2019)
25. Rogers M., The Role of Digital Forensics in Incident Response, Journal of Digital Forensics (2018)
26. Rid T., Cyber War Will Not Take Place, Oxford University Press, ISBN 978-0199314736 (2013)
27. Holt T., Bossler A., Cybercrime in Progress, Routledge, ISBN 978-1138572639 (2016)
28. Décary-Hétu D., Darknet Markets and Cybercrime, Palgrave Macmillan, ISBN 978-1137478574 (2018)
29. Moore T., Rid T., Cryptopolitik and the Dark Web, Survival Journal, Vol. 58 (2016)
30. Hutchins E., Cloppert M., Amin R., Intelligence-Driven Computer Network Defense, Lockheed Martin (2011)
31. SANS Institute, Cyber Threat Intelligence Best Practices, SANS Whitepaper (2022)
32. OpenCTI Consortium, OpenCTI Platform Documentation, ANSSI (2024)
33. MISP Project, Malware Information Sharing Platform Documentation, CIRCL (2024)
34. Shodan, The Internet Intelligence Platform, Shodan LLC (2024)
35. Censys, Internet Asset Discovery and Threat Intelligence, Censys Inc. (2024)
36. Russell S., Norvig P., Artificial Intelligence: A Modern Approach, Pearson, ISBN 978-0134610993 (2021)
37. Goodfellow I., Bengio Y., Courville A., Deep Learning, MIT Press, ISBN 978-0262035613 (2016)
38. Brundage M. et al., The Malicious Use of Artificial Intelligence, Oxford University Press (2018)
39. Papernot N. et al., Practical Black-Box Attacks against Machine Learning, ACM CCS (2017)

40. Biggio B., Roli F., Wild Patterns: Ten Years After the Rise of Adversarial Machine Learning, Pattern Recognition (2018)
41. MITRE, ATLAS: Adversarial Threat Landscape for AI Systems, MITRE (2023)
42. NIST, Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST (2023)
43. European Commission, AI Act Proposal, European Union (2024)
44. Floridi L. et al., AI Ethics Guidelines, Springer (2019)
45. Kott A., Artificial Intelligence for Cyber Defense, Springer, ISBN 978-3319728318 (2019)
46. Sommer R., Paxson V., Outside the Closed World: On Using ML for Network Security, IEEE S&P (2010)
47. Darktrace, AI in Cyber Defense, Darktrace Whitepaper (2023)
48. Palo Alto Networks, Cortex XDR Technical Overview, Palo Alto Networks (2024)
49. Microsoft, Digital Defense Report, Microsoft Security (2024)
50. Cloudflare, DDoS Threat Report, Cloudflare Inc. (2024)
51. Akamai, State of the Internet Security Report, Akamai (2023)
52. CERT-EU, Incident Response Best Practices, CERT-EU (2023)
53. FIRST.org, CSIRT Services Framework, FIRST (2022)
54. NATO CCDCOE, Cyber Defence and Artificial Intelligence, NATO Cooperative Cyber Defence Centre of Excellence (2022)
55. Barzev I., Borissova D., An Improved Static Analysis Approach for Malware Detection by Optimizing Feature Extraction Combining Different ML Algorithms, in: Intelligent Systems and Pattern Recognition (ISPR 2024), Communications in Computer and Information Science (CCIS), Springer, Cham (2025), DOI: 10.1007/978-3-031-82153-0_8
56. Barzev I., Borissova D., Buhtiyarov N., *Comparison of Different Binary Classification Algorithms for Malware Detection*, in: *Information Technology and Systems (ICITS 2024)*, Lecture Notes in Networks and Systems, Springer, Cham (2024), DOI: 10.1007/978-3-031-54235-0_33
57. Borissova D., Barzev I., Yoshinov R., Kotseva M., *Group Decision-Making Models for Selection of Virtual Machine Software for Malware Detection Purposes*, in: *12th Mediterranean Conference on Embedded Computing (MECO)*, IEEE (2023), DOI: 10.1109/MECO58584.2023.10155084

58. Axelsson S., *Intrusion Detection Systems: A Survey and Taxonomy*, Technical Report, Chalmers University of Technology (2000)
59. Lippmann R. et al., *The 1999 DARPA Off-Line Intrusion Detection Evaluation*, *Computer Networks*, Elsevier (2000)
60. Roesch M., *Snort: Lightweight Intrusion Detection for Networks*, in: *USENIX LISA* (1999)
61. Kruegel C., Vigna G., *Anomaly Detection of Web-based Attacks*, in: *ACM CCS* (2003)
62. Garfinkel S. L., *Digital Forensics Research: The Next 10 Years*, *Digital Investigation* (2010)
63. Carrier B., Spafford E. H., *An Event-Based Digital Forensic Investigation Framework*, in: *Digital Forensic Research Workshop (DFRWS)* (2004)
64. Casey E., Ferraro M., Nguyen L., *Investigation Delayed Is Justice Denied: Proposals for Expediting Forensic Examinations*, *Digital Investigation* (2009)
65. Reith M., Carr C., Gunsch G., *An Examination of Digital Forensic Models*, *International Journal of Digital Evidence* (2002)
66. Grispos G., Glisson W., Storer T., *Using Smartphones as a Proxy for Digital Evidence*, in: *IEEE ARES* (2011)
67. Quick D., Choo K.-K. R., *Impacts of Increasing Volume of Digital Forensic Data: A Survey and Future Research Challenges*, *Digital Investigation* (2014)
68. Arkin B., Stender S., McGraw G., *Software Penetration Testing*, *IEEE Security & Privacy* (2005)
69. Shin Y., Williams L., *Can Traditional Fault Prediction Models Be Used for Vulnerability Prediction?*, *Empirical Software Engineering* (2011)
70. Zimmermann T. et al., *Predicting Vulnerabilities Using Code Metrics*, in: *MSR / ICSE workshops* (2007–2010 линия публикации)
71. Chowdhury I., Zulkernine M., *Can Complexity, Coupling, and Cohesion Metrics Be Used as Early Indicators of Vulnerabilities?*, in: *SAC* (2010)
72. Alshaikh M., Maynard S., Ahmad A., Chang S., *An Investigation of Current Security Audit Practices and Challenges*, in: *IFIP SEC* (2018)
73. Siponen M., Willison R., *Information Security Management Standards: Problems and Solutions*, *Information & Management* (2009)

74. Da Veiga A., Eloff J., *An Information Security Governance Framework*, *Information Systems Management* (2007)
75. Von Solms R., Von Solms B., *From Information Security to Cyber Security, Computers & Security* (2018)
76. Herath T., Rao H. R., *Protection Motivation and Deterrence: A Framework for Security Policy Compliance*, *Decision Support Systems* (2009)
77. Puhakainen P., Siponen M., *Improving Employees' Compliance Through Information Systems Security Training*, *MIS Quarterly* (2010)
78. Sasse M. A., Brostoff S., Weirich D., *Transforming the "Weakest Link": A Human/Computer Interaction Approach to Usable and Effective Security*, *BT Technology Journal* (2001)
79. Jajodia S., Liu P., Swarup V., Wang C. (eds.), *Cyber Situational Awareness: Issues and Research*, Springer (2010)
80. Conti M., Dehghantanha A., Franke K., Watson S., *Internet of Things Security and Forensics: Challenges and Opportunities*, *Future Generation Computer Systems* (2018)
81. Barnum S., *Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX)*, in: *MITRE/industry-academic proceedings*
82. *Husák M. et al., Survey of Attack Projection, Prediction, and Forecasting in Cyber Security, IEEE Communications Surveys & Tutorials (2018)*
83. *Sabottke C., Suciu O., Dumitras T., Vulnerability Disclosure in the Age of Social Media: Exploiting Twitter for Predicting Exploits, in: USENIX Security (2015)*
84. *Bridges R. A., Glass-Vanderlan T., Iannacone M., Ferragut E., Automated Labeling for Cybersecurity using Weak Supervision, in: Workshop/Conference proceedings (2019–2021; подходящо към SIEM/ML)*
85. *Ring M. et al., A Survey of Network-based Intrusion Detection Data Sets, Computers & Security (2019)*
86. *Sharafaldin I., Lashkari A. H., Ghorbani A., Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization, in: ICISSP (2018)*
87. *Tavallaee M. et al., A Detailed Analysis of the KDD CUP 99 Data Set, in: IEEE CISDA (2009)*

88. *Garcia-Teodoro P., Diaz-Verdejo J., Maciá-Fernández G., Vázquez E., Anomaly-based Network Intrusion Detection: Techniques, Systems and Challenges, Computers & Security (2009)*
89. *Chandola V., Banerjee A., Kumar V., Anomaly Detection: A Survey, ACM Computing Surveys (2009)*
90. *Sommer R., Kreibich C., Paxson V., The Effect of Feature Choice on Intrusion Detection, in: RAID*
91. *Carlini N., Wagner D., Towards Evaluating the Robustness of Neural Networks, in: IEEE Symposium on Security and Privacy (2017)*
92. *Carlini N. et al., Extracting Training Data from Large Language Models, in: USENIX Security (2021)*
93. *Tramèr F. et al., Stealing Machine Learning Models via Prediction APIs, in: USENIX Security (2016)*
94. *Shokri R. et al., Membership Inference Attacks Against Machine Learning Models, in: IEEE S&P (2017)*
95. *Fredrikson M. et al., Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures, in: ACM CCS (2015)*
96. *Jagielski M. et al., Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning, in: IEEE S&P (2018)*
97. *Steinhardt J., Koh P. W., Liang P., Certified Defenses for Data Poisoning Attacks, in: NeurIPS (2017)*
98. *Eykholt K. et al., Robust Physical-World Attacks on Deep Learning Models, in: CVPR (2018)*
99. *Goodfellow I. J., Shlens J., Szegedy C., Explaining and Harnessing Adversarial Examples, in: ICLR (2015)*
100. *Brown T. B. et al., Language Models are Few-Shot Learners, in: NeurIPS (2020)*
101. *Weidinger L. et al., Ethical and Social Risks of Harm from Language Models, in: ACM FAccT (2022)*
102. *Pearce H. et al., Asleep at the Keyboard? Assessing the Security of GitHub Copilot's Code Contributions, in: IEEE Symposium on Security and Privacy (S&P) (2022)*

КИБЕРСИГУРНОСТТА КАТО СИСТЕМА: ОТ ТЕСТОВЕ И ОДИТИ ДО РАЗУЗНАВАНЕ И ИЗКУСТВЕН ИНТЕЛЕКТ

Автор: Иван Иванов Благоев

Институт по информационни и комуникационни технологии - Българска Академия на Науките

Научни рецензенти:

полк. доц. д-р Иван Николов Чакъров,

началник на катедра „Комуникационни и информационни системи“, факултет „Командно-щабен“,

Военна академия „Г. С. Раковски“.

доц. д-р Нончо Иванов Димитров,

ръководител на катедра „Национална и регионална сигурност“, факултет „Икономика на инфраструктурата“,

Университет за национално и световно стопанство (УНСС).

Народност българска

Първо издание

Технически редактор: проф. Даниела Борисова

Формат: А5

Тираж: 50 бр.

ISBN 978-619-92403-7-3

eISBN 978-619-92403-8-0

Издателство: Асоциация КИЕЕС

Печат: УИЦ при Русенски университет “Ангел Кънчев”