Issues in Information Systems
Volume 24, Issue 4, pp. 311-327, 2023

DOI: https://doi.org/10.48009/4_iis_2023_124

Application of artificial intelligence and machine learning in a security operations center

Mohammad Anwarul Islam, *Middle Georgia State University, mohammad.islam@mga.edu*

Abstract

The security operations center's (SOC) mission is to protect digital assets (data, applications, infrastructure) from malicious attacks and breaches. The SOC accomplishes its mission through people, processes, and technologies in detecting, responding, and recovering from cyber-attacks. SOC depends on several hardware appliances and software tools such as firewalls, intrusion detection and prevention systems, sensors-based events, system logs, endpoint detection and response, threat intelligence, vulnerabilities scanner, etc. These tools and appliances generate an enormous volume of data in real-time. Therefore, tools such as security events and information management (SIEM) must analyze large volumes of data to detect malicious activities and security incidents. Machine learning and artificial intelligence technologies have the potential to detect anomalies and cyberattacks. This research focuses on how AI/ML is embedded in SOC tools.

Keywords: security operations center, artificial intelligence, machine learning, cyber-attack

Introduction

The primary mission of the Security Operations Center (SOC) is to defend the enterprise against breaches and attacks, keep enterprise assets (e.g., data, applications, infrastructure) secure, and facilitate normal operations around the clock (Palo Alto Networks, 2020). According to SANS Institute, "a SOC is a combination of people, processes, and technology protecting an organization's information systems through proactive design and configuration, ongoing monitoring of system state, detection of unintended actions or undesirable state, and minimizing damage from unwanted effects."

SOC is the central point of the enterprise to collaborate and coordinate efforts in monitoring, assessing, and mitigating malicious activities (Check Point, n.d.). The operational focus of SOC is security incident detection, analysis, and response. Johnson (n.d.) described the SOC as the heart of any security organization, and its essential functions are monitoring, analysis, and response. According to Johnson (n.d.), the proposed SOC framework is based on five core principles: (1) monitoring, (2) analysis, (3) incidence response and containment, (4) auditing and logging, and (5) threat hunting.

With increased digital transformation, the attack surface for enterprises has been growing tremendously. With the rise of advanced persistent threat (APT), attackers have been deploying complex and cutting-edge technologies to discover and exploit vulnerabilities in systems and networks, resulting in the exponential growth of cyberattacks, including zero-day ones. Furthermore, enterprises adopted mobile and cloud technologies, making traditional perimeter-based cyber defense less effective (Knerler et al., 2022). Despite all defensive security postures, organizations are prone to cyberattacks and security breaches, negatively impacting normal operations supporting missions. Due to security breaches, companies lose valuable data, including personally identifiable information (PII) about customers and employees. Therefore, the security operations center must adopt cutting-edge technologies to disrupt and identify attacks and mitigate security incidents.

The SOC system continuously monitors digital activities, and the enterprise infrastructure security prevents and detects malicious activities (Lindstrom, 2018). In some cases, it takes the appropriate actions automatically. Additionally, security analysts must review the system-generated alerts and take proper steps to prevent and respond to attacks already underway. Suitable machine learning algorithms utilize large volumes of collected data and identify anomalies and attacks as part of SOC systems, like Security Incident and Event Management (SIEM) (Sathana & Memamalini, 2022).

With the adoption of technologies, various devices, software, and cloud services (e.g., IaaS, PaaS, SaaS, etc.) and the exponential growth of sophisticated adversaries and cyber threats (Kumar et al., 2021), the security operation center must monitor inbound and outbound traffic to prevent, detect, respond and recover from any security indents (Yeshwanth et al., 2022). To identify cyber threats and attacks, SOC must gather data across the enterprise and perform real-time analysis of massive volumes of data from logs (e.g., network firewalls, web firewalls, network sensors, endpoints, etc.), intrusion detection and prevention systems, identity and access management systems, threat intelligence feeds, and numerous other sources. With traditional technologies and methodologies, analyzing large volumes of data to identify threats and attacks is challenging (Kim et al., 2020). Furthermore, conventional methods generate many alerts that overwhelm the security operations center analysts in prioritizing them and taking the appropriate remediation actions (Farooq & Otabi, 2018). Therefore, the SOC needs the capability to analyze large volumes of data in real-time or near real-time, identify attacks and breaches, and take corrective measures.

Purpose of the Study

This study aims to examine how artificial intelligence (AI) and machine learning (ML) techniques and algorithms have been used in a SOC to prevent, detect, respond, and recover from security incidents. Furthermore, the study explores various frameworks and their effectiveness in securing enterprise resources. This research also focuses on identifying the limitations of existing SOC frameworks and technologies and how AI/ML-driven technologies can enhance the effectiveness of SOC against malicious activities.

This research addresses the following research questions (RQ):

1. How effective are the current security operation center technologies and frameworks in detecting, preventing, and responding to cyber threats and malicious activities in real-time or near real-time?
2. How do the current SOC technologies incorporate AI/ML algorithms to enhance capabilities, particularly in identifying zero-day attacks?
3. Which SOC software vendors have adopted machine-learning algorithms and open-source libraries?

This study helps develop an integrated framework for security operations centers of enterprises, incorporating proven, cutting-edge machine learning algorithms and artificial intelligence approaches. The developed framework can be implemented with available machine learning algorithms and libraries, particularly open-source ones. ML algorithms and libraries can be incorporated into a Python-based integrated development environment (IDE) and tested using actual data before production deployment. The new solution allows SOC software vendors and security professionals to embed the appropriate ML libraries. The AI/ML-driven SOC software significantly enhances the capabilities to protect enterprises from attacks and breaches in real-time.

The rest of the research paper is organized as follows. A literature review is presented in the next section, followed by the methodology that includes the critical analysis of existing SOC technology frameworks and solutions and the development of AI/ML-driven solution architecture. After the methodology, the research results are depicted in the results section. Following the research results, an AI/ML-driven solution architecture for the SOC and its implementation guidelines have been presented. A discussion and conclusion follow the implementation guidelines.

Review of the Literature

Research articles review

Capgemini Research Institute (2019) studied how to bolster cybersecurity by adopting artificial intelligence (AI) in security systems to attain the security posture for the enterprise. This report also outlined the roadmap for implementing AI in cybersecurity. Capgemini identified several use cases, including deploying AI as part of security orchestration, automation, and response (SOAR). The AI-driven SOAR is a force multiplier empowering SOC analysts further in detecting, preventing, and responding to identified threats and attacks (Nayyer, 2022).

SOC analysts are essential personnel and usually perform tasks such as (1) continuous monitoring, (2) threat hunting, (3) threat intelligence analysis, and (4) threat incident response support (Trellix, n.d.). The current challenges for SOC analysts in preventing and recovering from cyber incidents were discussed by Mirilla (2018). The author proposed a Dynamic SOC Management (DSM) framework to improve alertness, early detection, and investigation of suspicious or anomalous activities. The proposed framework would activate an early incident response to either disrupt an attack in progress or contain an inbound one. It is safe to say that machine learning-based framework for security operation centers minimizes false positive alerts (Feng et al., 2017). Lindstrom (2018) described how security analysts with insufficient data or poor judgment could label an alert as a false positive when it was indeed positive. It can have a significant negative impact on the organization. The paper argued that machine learning-based SOC software could identify anomalies and detect intrusion. The author suggested using signature-based and ML-based software to minimize false positives and negatives.

It is an inherent strength of machine learning-based algorithms that they can learn from historical and current data, predict future security incidents, and help analysts detect and analyze threats quickly (Feng et al., 2017). However, SOC analysts' confidence in AI-predicated attacks and threats decreased due to the high rate of false alerts (Kim et al., 2020). SOAR systems utilize AI/ML to detect, mitigate, and prevent cyber threats and attacks. Many SOAR, SIEM, and other security software vendors have adopted AI/ML and have become force multipliers, further empowering SOC analysts (Kinyua & Awuah, 2021). The authors also identified several research areas for embedding AI/ML algorithms, such as SOC incidents classification and prioritization. Further research is warranted in applying the deep reinforced Learning (DRL) algorithm in SOAR systems.

The explainable AI (XAI) new model provides explainability through an interpretation of AI prediction results and a reliability analysis of AI predictions based on explainable artificial intelligence (Aslam et al., 2022). In addition, the authors proposed a method for screening high-quality data that could efficiently detect false predictions based on reliability indicators. The XAI model would help reduce false alerts, enabling SOC analysts to review and take appropriate remediation actions. Some cybersecurity tools are model-centric, lacking the focus of end-users and security analysts. As a result, SOC analysts spend more time understanding alerts than responding to incidents (Eriksson, 2022). The author proposed that explainable AI be deployed as part of SOC tools, which could help analyze the alerts for analysts to respond appropriately. The author used XMI methods LIME and SHAP to generate valid alerts for SOC analysts. The contemporary challenges of security operations centers might be daunting to detect, mitigate and recover from cyberattacks and threats (Shutock & Deirtich, 2022). The authors highlighted how machine learning-based algorithms could detect and mitigate cyberattacks and help the SOC take remedial action automatically. The security operations center consists of several layers: data collection sources, data processing, technologies (analysis), and display. Different machine-learning models can be used in different layers of security operations centers (Yeshwanth et al., 2022).

In general, the SOC operational model that utilizes a machine learning-based analytics framework comprises four layers: data collection, data processing, threat intelligence leveraging machine learning, and security dashboards and management (Yeshwanth et al., 2022). The authors adopted supervised and unsupervised learning-based algorithms. The supervised algorithms, such as a Support Vector Machine (SVM) and decision tree in the framework, have been deployed in SOC tools. The unsupervised algorithms incorporated in the model were K-Means Clustering and K-nearest Neighbor. The author demonstrated the application of the model and its ML algorithms through case studies. Prasad (2021) proposed a SOC user-centric machine learning framework to reduce false positive alerts in the real world. Ban et al. (2021) presented artificial intelligence-assisted tools to distinguish actual threats from false alarms to reduce alert fatigue. Sathana & Hemamalini (2022) discussed high false alerts in SOCs and proposed a user-centric machine learning framework for the Internet Safety Functional Center. The heart of the framework was the support vector machine (SVM) based machine learning algorithms which could learn from past data (logs, threat intelligence, etc.) and determine the threat level to the enterprise.

Based on the reviews, it is evident that many researchers have identified that SOC's capabilities can be significantly increased by deploying different AI/ML algorithms like deep reinforcement learning (DRL) in SOC tools, such as SOAR and SIEM. AI-driven automation of routine SOC tasks and workflows would substantially enhance the organization's security posture.

Review of industry research reports

Industry research groups like Gartner, Forrester, IDC, MITRE, etc., have been involved in developing new concepts, frameworks, architectures, and reports to guide both public and private sector enterprises to deploy technologies and secure their digital assets. Forrester (2020) presented the state of security operations and found that only 46% of decision-makers agreed that they were satisfied with their organization's ability to detect threats. Among the enterprises with SOC, only 13% used automation and machine learning for triage, analysis, and response. Though Extended Detection and Response (XDR) can overcome challenges, it has not been widely adopted across industries.

Gartner's security operations center model highlighted typical SOC capabilities and presented an implementation model example (Collins et al., 2021). Gartner's model included four broad capabilities of SOC, which are (1) monitoring and detection, (2) detection and automation engineering, (3) incidence response and threat hunting, and (4) threat intelligence. The traditional SOC approach in cybersecurity is not sustainable due to expanding security surface areas and large amounts of data that generate many false positive alerts. Artificial intelligence and machine learning must supplement manual processes to generate true positive alerts and could take remediation action automatically (Kissel, 2021). The author found that Check Point SOC tool utilized artificial intelligence and machine learning. Furthermore, the Check Point tool could use the MITRE ATT&CK framework. MITRE also developed strategies and implementation guidelines for building world-class cybersecurity operations centers (Knerler et al.,2022). It is evident that industry research groups have been promoting the use of artificial intelligence and machine learning technologies for security operations center tools to detect and prevent cyberattacks and take remediation actions.

Review of cybersecurity frameworks

Several international bodies and US organizations have developed cybersecurity frameworks, and the SOC is one component of the overall cybersecurity framework. NIST (2018) developed a cybersecurity framework (CSF), a widely used and cited framework. CSF is composed of three parts: (1) the framework core, (2) framework implementation tiers, and (3) the framework profile. The framework core presents industry standards, guidelines, and practices. The framework's core elements are *functions*, *categories*, *and subcategories*. The framework core consists of five concurrent and continuous functions: *identify, protect, detect, respond, and recover*, as shown in Figure 1.

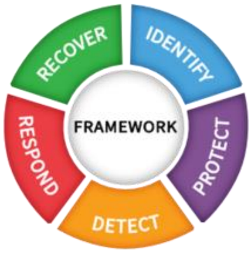
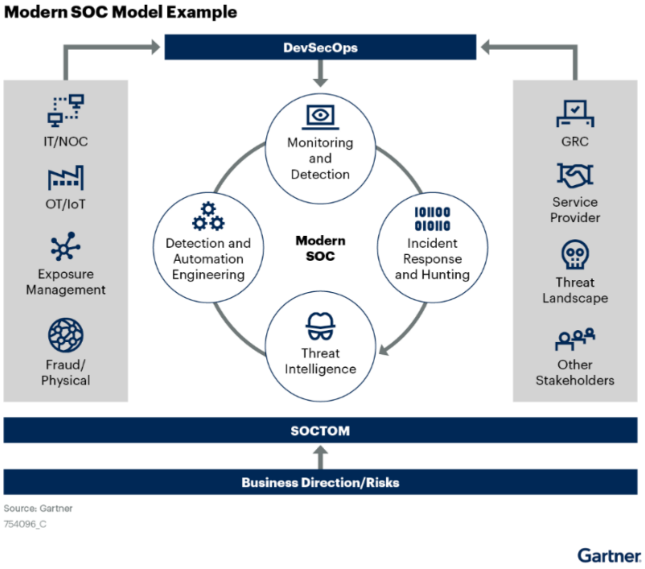


Figure 1: **NIST Cybersecurity Framework**

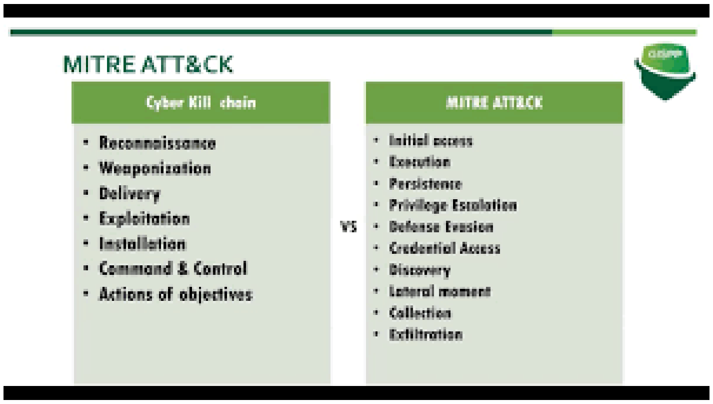
Gartner's cybersecurity research group outlined best practices for adopting a security operations center. Figure 2 shows Gartner's modern SOC model (Collins et al., 2021).



Source: Gartner
754096_C

Figure 2: **Modern SOC Model Example (by Gartner)**

Several other frameworks exist, notably MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) and Cyber Kill Chain. MITRE ATT&CK is classifying and describing attacks and associated remediations. The SOC technology vendors and enterprise SOC teams can utilize ATT&CK for threat intelligence, understanding methods used by malicious actors, and detection and mitigation techniques for preventing or identifying attacks (Alsheh, 2022). Cyber Kill Chain, developed by Lockheed Martin, traces the stages of cyberattack (as shown in Figure 3). The Cyber Kill Chain model comprises seven stages. SOC professionals can apply appropriate controls associated with the stage to prevent and detect cyberattacks before they penetrate the enterprise resources (Logsign, 2020). ATT&CK and Cyber Kill Chain frameworks can help SOC software vendors and SOC professionals incorporate appropriate controls to secure the Enterprise.



MITRE ATT&CK

Cyber Kill chain		MITRE ATT&CK
• Reconnaissance • Weaponization • Delivery • Exploitation • Installation • Command & Control • Actions of objectives	VS	• Initial access • Execution • Persistence • Privilege Escalation • Defense Evasion • Credential Access • Discovery • Lateral moment • Collection • Exfiltration

Figure 3: **Comparison between Cyber Kill Chain and MITRE ATT&CK (Adopted from YouTube)**

Review of SOC technologies

Security processes, functions, and technologies within SOC must be integrable or interoperable to keep enterprises safe from intrusion and attacks. The key functional elements of modern-day SOC are (1) log management, (2) security information and event management (SIEM), (3) Security orchestration, automation and response (SOAR), (4) vulnerability management, (5) endpoint detection and response (EDR) (ManageEngine, n.d.). Several technologies are essential for a fully functional SOC, such as user and entity behavior analytics (UEBA), cyber threat hunting, and cyber threat intelligence. Splunk (n.d.), a leading SOC vendor, provides end-to-end software and technologies for a modern security operations center. Splunk solutions include SIEM, SOAR, UEBA, and other third-party tools that can be incorporated under the Splunk Mission Critical module, as shown in Figure 4. Splunk's UEBA used an unsupervised machine learning algorithm to find unknown threats and anomalous behaviors across users, endpoints, and applications.

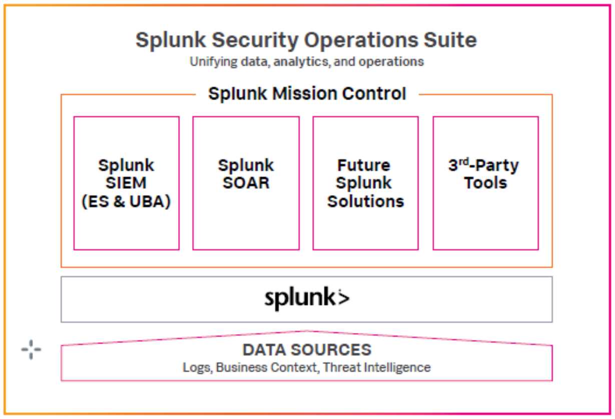

Splunk Security Operations Suite
Unifying data, analytics, and operations
Splunk Mission Control
Splunk SIEM (ES & UBA)
Splunk SOAR
Future Splunk Solutions
3rd-Party Tools
splunk>
DATA SOURCES
Logs, Business Context, Threat Intelligence

Figure 4: **Splunk Security Operations Center Architecture**

Exabeam (n.d.) discussed challenges in building a security operations center and highlighted how Exabeam security tools address these challenges. This vendor offers SIEM, GRC (governance, risk, and compliance), IDS/IPS (intrusion detection and prevention systems), firewalls, cyber threat intelligence, EDR (end-point detection and response), NTA (network traffic analysis), APM (application performance monitoring), UEBA, and SOAR. The heart of Exabeam SOC solution is next-generation SIEM, leveraging machine learning and behavioral analytics, data science, and data lakes.

Microsoft (n.d.) elaborated on essential functions of the security operations center to help prevent, respond, and recover from attacks. Microsoft outlined SOC functions such as continuous monitoring, threat intelligence, threat detection, log management, incident response, recovery and remediation, root cause analysis, security refinement, and compliance management. Microsoft offers cloud-based SOC tools and technologies that include SIEM, SOAR, XDR (extended detection and response), firewalls, logs management (as part of SIEM), vulnerability management, and UEBA. Microsoft utilizes machine learning and artificial intelligence both in SIEM and UEBA. The Microsoft SOC includes Microsoft Sentinel, Microsoft 365 Defender, Microsoft Defender Threat Intelligence, etc.

ArcSight (n.d.) offered solid security operation center software and tools, including SIEM with built-in SOAR, data platform, and layered analytics. ArcSight has a built-in AI engine to provide anomaly detection to find insider threats, zero-day attacks, and advanced persistent threats. Layered analytics perform real-time correlation, hypothesis, and analytics-based threat hunting, providing rich insights about malicious activities. IBM (n.d.) presented a suite of security software products that can be part of a fully-fledged security operations center. IBM security tools are SIEM, QRadar SOAR, QRadar NDR, and QRadar XDR. QRadar SIEM Security 4412-Q2A appliance accurately detects and prioritizes cybersecurity threats and internal user violations with embedded security AI, user behavior analytics, and machine learning technology. Some SOC vendors incorporated AI/ML into the technology suites to identify threats, anomalies, and attacks. This research investigates the effectiveness and efficiency of deployed AI/ML algorithms.

Research Methodology

The research was conducted by reviewing and analyzing academic articles, standards, and frameworks developed by national and inter-government bodies, industry research reports, and software and technologies available in the marketplace. The research methodology is visually depicted in Figure 5.

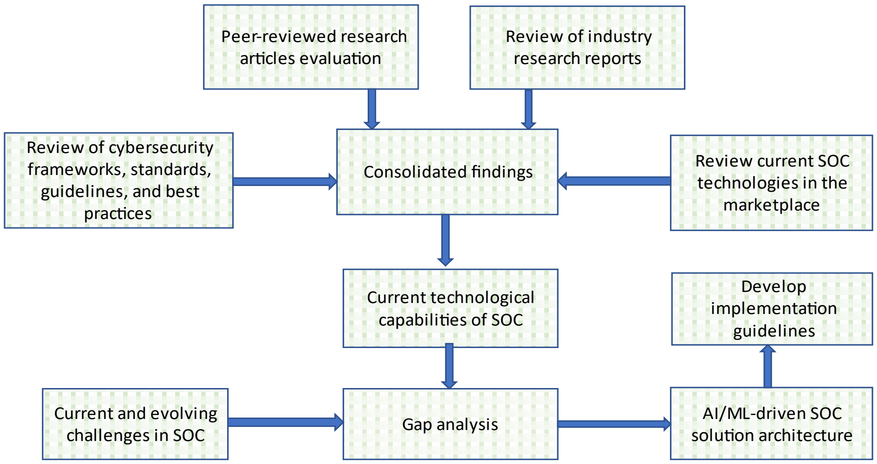

Peer-reviewed research articles evaluation
Review of industry research reports
Review of cybersecurity frameworks, standards, guidelines, and best practices
Consolidated findings
Review current SOC technologies in the marketplace
Current technological capabilities of SOC
Develop implementation guidelines
Current and evolving challenges in SOC
Gap analysis
AI/ML-driven SOC solution architecture

Figure 5: Graphical Representation of Research Methodology for AI/ML Embedded SOC Architecture

The research methodology helped to develop a solution architecture embedding AI/ML algorithms and their implementation in an actual security operations center. The research methodology steps are described below.

1. Review and analysis of peer-reviewed articles on the application of AI/ML in SOC from reputable research journals and conference proceedings.
2. Review and analysis of industry research reports on SOC published by industry research organizations (e.g., Gartner, Forrester, SAN Institute, etc.).
3. Review and analysis of cybersecurity frameworks, standards, guidelines, and best practices by national and international bodies (e.g., NIST, CISA, DoD, ISO, MITRE, ISC2, ISACA, etc.) applicable to SOC.
4. Review SOC software vendors' current offerings and capabilities, meeting SOC's requirements.
5. Development of the consolidated research findings.
6. Creation of a summary of the review and analysis of results in the context of the technological capabilities of SOC in the marketplace.
7. Enumeration and synthesis of the current and evolving challenges SOC faces in protecting enterprise digital resources.
8. Gap analysis for identification of existing tools' capabilities and evolving SOC challenges.
9. Development of Artificial Intelligence and Machine Learning centric solution architecture for SOC to close the gap between existing technologies and required capabilities.
10. Creation of implementation guidelines for the proposed solution, including AI/ML algorithms to be deployed.

Research Results

This study aimed to find the answers to three research questions to understand the capabilities, effectiveness, and security operations center technologies and tools in defending the enterprise. The findings have been summarized broadly under typical SOC functions, current capabilities, and evolving challenges in security operations centers. Regarding RQ1, this study concluded that security operations center tools are highly effective in defending data and digital resources against threats and attacks. Regarding RQ2, this research confirmed that machine learning and artificial intelligence algorithms enhanced the capabilities of SOC tools to identify zero-day attacks. In response to RQ3, this study deduced that leading security operations center technology vendors utilized machine learning algorithms to detect, remove, and recover from cyber incidents. However, the study could not reveal what type of machine learning or artificial intelligence (e.g., proprietary or open-source libraries) algorithms were embedded in SOC software tools. The detailed explanations of how the author derived the answers to research questions can be found in this paper's "current technological capabilities of SOC" section.

Typical SOC functions

Based on the literature review, these are significant functions performed by a typical security operations center (CheckPoint, n.d.; InfosecMatter, 2020): continuous monitoring, preventing threats and attacks, investigating alerts and events, root cause analysis, responding and mitigating, etc.

Current technological capabilities of SOC

Modern security operations centers can monitor and defend enterprise resources against malicious activities. Current capabilities have been summarized based on literature reviews and security operations centers tools and technology vendors' offerings (IBM, n.d.; Splunk, 2022; Microfocus, n.d.; Trellix, n.d.), etc. The technology stack for the security operations center has various components and modules, and different SOC may deploy a subset of these tools such as SIEM (Security Information and Event Management), SOAR (Security Orchestration, Automation, and Response), EDR (Endpoint Detection and Response), XDR (Extended Detection and Response), etc.

Most leading security operation centers' technology vendors incorporated advanced and near real-time detection of security incidences. Many vendors integrated machine learning-based anomaly detection algorithms in Security Information and Event Management (SIEM), Security Orchestration, Automation and Response (SOAR), and Extended Detection and Response (XDR) software. Furthermore, SOC technology vendors incorporated technologies like Threat Intelligence, Threat Hunting, etc., to enrich and augment the collected data to minimize false positive alerts. The literature review showed that most security technologies have core capabilities in detecting, preventing, responding, and recovering from security incidents. Many security incidents were detected and prevented in real-time and near real-time through firewalls, IDS and IPS, EDR/XDR, and SOAR technologies. Both corporate and public sector enterprises have been using different security operations center tools and technologies and have successfully defended against attacks on enterprise digital assets.

Security operations center tools provided real-time and near-time capabilities against threats and malicious activities. NIST Cybersecurity Framework (CSF) outlined five board functions, and at least three of these functions can be mapped to security operations center capabilities, as shown in Table 1. Most security vendors incorporated machine learning and artificial intelligence in their software tools. The analysis concluded that modern-day security operations center tools and software could detect and respond to security incidence in real-time or near real-time. Furthermore, these tools successfully detect, respond, and recover from security incidents. Therefore, the answer to research question one is that security operations center tools are highly effective in defending data and digital resources against threats and attacks.

Artificial intelligence, machine learning, and threat intelligence data helped the tools analyze large volumes of data and detect anomalies and potential security incidents. Tools using signature-based algorithms could not detect zero-day (previously unknown) attacks. Machine learning algorithms correlate and corroborate with data from different endpoints, network data, and log data from servers, firewalls, etc., to develop a pattern, determine the incidents and priorities, and immediately remediate the incidents. This answered research question two and confirmed that machine learning and artificial intelligence algorithms enhanced the capabilities of SOC tools to identify zero-day attacks.

Many SOC technology vendors in this study utilized machine learning algorithms in the SOC tool sets (e.g., SIEM, XDR, and SOAR). However, published articles or technology vendors' materials did not mention what type of machine learning or artificial algorithms were embedded in the tools. This answered research question three: leading security operations center technology vendors utilized machine learning algorithms to detect, remove, and recover from cyber incidents.

Current and evolving challenges of SOC

The current and evolving challenges were summarized based on literature reviews and published reports from several security technologies of reputable vendors such as Splunk, Microfocus, IBM, Veracode, Trend Micro, Rapid7, Exabeam, LogRhythm, CrowdStrike, Trellix, etc. SOC tools were primarily defensive mechanisms that passively monitor and respond (InfosecMatters, 2020). The manual investigation, alert prioritization, and threshold-based correlation rules were still challenges in many enterprises (Kaliyaperumal, 2021).

Security operation centers had ongoing challenges in four main areas: people (lack of skilled people, monotonous tasks, collaboration experts, integration of domain knowledge), processes (lack of standard procedures, adapting the generic IT process to SOC), technologies (increased complexity, wide variety of tools, visualization capabilities, insufficient level of automation) and governance and compliance (effective measure of SOC performance, lack of best practices and standards, privacy regulations) (Microfocus, n.d.). While organizations lacked skilled people, determined and highly skilled attackers could use the latest tools and technologies, including artificial intelligence and machine learning, to mount sophisticated attacks against organizations (Microfocus, n.d.). Even tier 3 or tier 4 SOC analysts found it challenging to investigate incidents when sophisticated attackers removed their digital footprint (IIoT World, 2022). The security operations center must handle enormous amounts of data from network traffic and logs from devices, applications, and networks. The parsing and ingesting of data into the data lake and then determining the malicious activities in real-time was challenging. Alert fatigue might happen due to many anomaly alerts without context or intelligence. The machine learning-based tool correlates with context data across the enterprise, minimizes false positives, and generates a priority list of alerts and incidents. Zero-day attack vectors could be challenging due to a lack of threat intelligence and undiscovered vulnerabilities. Machine learning-embedded behavior analytics can detect unusual behavior and unknown attacks (Microfocus, n.d.).

In some organizations, CISO/CIO opted for a best-of-breed set of security tools and software from several vendors. Too many security tools from multiple vendors without a unified framework, integration architecture, and sometimes disconnected work on silos could create duplicate, overlapping, and sometimes conflicting alerts and recommendations (Microfocus, n.d.). Many organizations lacked a complete inventory of digital infrastructure because different teams managed different components. Sometimes, SOC might not have full configuration details, firewall rules, and network diagrams. Even naming conventions for infrastructure components might not be standardized (InfosecMatter, 2020) in many organizations. Some organizations might not subscribe to any threat intelligence platform and lack indicators of compromises (IOC) data. IOC data was essential in defending against advanced persistent threats (APT) and determining malicious actors. (InfosecMatter, 2020). This research proposed a new solution architecture for the security operations center in the discussion section. The proposed solution architecture aims to overcome current and evolving challenges in securing enterprise digital assets.

Discussions and Conclusions

This research addressed the below deep-dive questions to understand the security operations center's existing capabilities, perceived gaps in security technologies, and future needs in the fast-evolving threats and technology landscape.

- What AI/ML tools made a significant impact on SOC activities?
- What security frameworks have been most adopted by the industry?
- What capabilities were lacking in SOC software tools?
- What were the emerging themes for future research on adopting AI/ML in SOC software?
- Industry research group's reports and recommendations – AI/ML algorithms, frameworks, software tools that had been adopted, and how effective were they?
- Why were SOC software vendors slow to adopt AI/ML technologies?

Technological gaps

The primary mission of an enterprise security regime is to protect digital resources from malicious activities. The core functions are: identify attack surfaces (digital assets) and vulnerabilities (known and zero days), gather threat intelligence and proactively hunt threats, protect the enterprise through safeguards (thus prevent attacks), detect breaches and attacks, respond to security incidents, recover from attacks and restore the typical operating environment. The security operations center should proactively detect and identify threats and take appropriate action(s) in real-time (Kaliyaperumal, 2021). Existing SOC tools were primarily passive and identified events after they occurred. Furthermore, there was inherent latency in current security tools in identifying threats, intrusions, and in-progress attacks. Though intrusion detection and prevention systems have capabilities at the network level, there are other attack vectors and avenues targeted by malicious actors to conduct their activities. SOC technologies were not mature enough to detect in-progress command and control activities of malicious actors in real-time, which occurred in the SolarWinds attack (Constantin, 2020). Security operations center tools should have robust and rapid prevention, detection, and response capabilities, including advanced controls, automation, and orchestration (Optiv, n.d.).

Many attackers are determined to steal data from the enterprise. Despite their best efforts, data loss prevention (DLP) and SOC tools may not detect data exfiltration effectively in all scenarios. Therefore, SOC needs advanced tools such as UEBA to prevent in-progress data exfiltration (Gonzalez, 2020). Furthermore, SOAR and XDR-type capabilities are required to remediate by terminating malicious sessions, deactivating user accounts, and disconnecting and isolating affected endpoints from the network to halt in-progress data exfiltration. The performance of a SIEM could be enhanced by adding various functionalities such as threat hunting, threat intelligence, and malware identification and prevention. These capabilities would reduce false positive alarms and increase the accuracy and efficiency of an organization's overall Security Operations process (Perera et al., 2021).

AI/ML-based SOC solution architecture

Though the security operations center is a mandatory function of an enterprise, its organizational structure varies widely. The architecture of a SOC can differ depending on the organization's size, budget, and security requirements, but several key components are common to most SOCs. This research proposed a holistic solution architecture for a SOC, as shown in Figure 6. Security operations centers usually deploy several software tools, many of which fall under one of these tool categories: SIEM, SOAR, and XDR.

SOC technology stack incorporates capabilities such as security information and event management, security orchestration, automation and response, and extended detection and response. Cobb (n.d.) discussed each technology type's focus, capabilities, and limitations. Gartner's SIEM, SOAR, and XDR definitions are similar (Gartner, n.d.). SIEM "supports threat detection, compliance, and security incident management through the collection and analysis of security events, as well as a wide variety of other event and contextual data sources." SOAR enables "organizations to collect inputs monitored by the security operations team." XDR is "a unified security incident detection and response platform that automatically collects and correlates data from multiple proprietary security components."

This research proposed a four-tier solution architecture for the security operations center, as shown in Figure 6. The proposed SOC architecture is vendor-neutral and leverages existing tools and technologies. Furthermore, evolving security tools and technologies can be added to this architecture.

*1. **Events and Log Collection***

In this layer, the SOC system gathers all security-related events and logs from security events from different sources, such as network devices, servers, and applications. The security-related data can also flow directly from the Extended Detection and Response system to the SOC central repository in addition to logs and events collections. The SOC polling tools gather the status of different systems and infrastructure elements and record them as events. XDR solutions consolidate data from various sources into a single platform, including endpoint detection and response, network detection and response, and cloud security tools. By doing so, XDR can provide a more comprehensive view of an organization's security posture and help identify threats that may be missed by individual security tools.

2. Central Data Repository

The events and logs collection tool inserts raw data into the staging area within the data repository. The next module (ingestion, normalization, aggregation, and assimilation) harmonizes a wide variety of data and adds the normalized data into the security data lake for further analysis. Additionally, the enterprise knowledge database gathers and normalizes data about security policies, vulnerabilities, assets inventory, and network/system. Furthermore, data from the threat intelligence platform and threat hunting tool flows directly to SOC central repository. Threat intelligence provides the SOC with information about known and emerging threats, such as malware, phishing attacks, and advanced persistent threats (APTs).

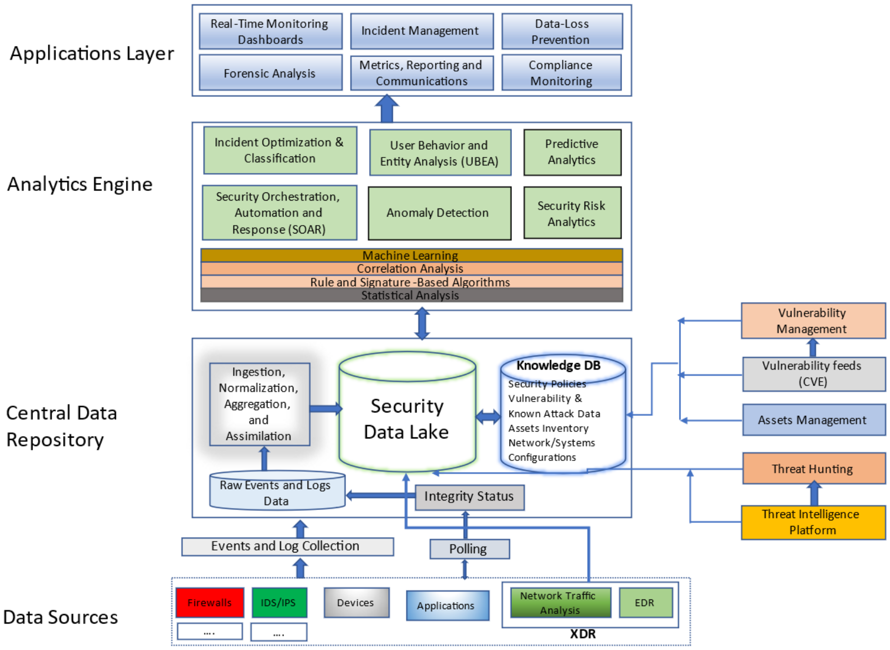


Figure 6: The proposed solution architecture for the security operations center

3. Analytics Engine

The analytics engine is the heart of the SOC technology stack to detect, respond and recover from cyber events and intrusions. The foundation of the analytics engine is based on machine learning, correlation analysis, rule and signature-based algorithms, and statistical analysis. The engine provides the capabilities such as correlation analysis among alerts and events across the enterprise, thereby detecting intrusions or malicious activities. Anomaly detection is a machine learning algorithm that can be trained to identify abnormal patterns in network traffic or user behavior, which may indicate a security threat. Anomaly detection models can learn from historical data to detect previously unknown threats and reduce false positives. Anomaly-based machine learning algorithms, such as unsupervised clustering algorithms, can establish baseline behavior. Any deviations can be corroborated and correlated with other data, such as threat intelligence and threat hunting, to provide proper context and prioritization of events.

User behavior and entity analysis tool uses advanced analytics and machine learning algorithms to identify anomalous user behavior and potential insider threats within an organization's network. SOAR automates and streamlines the incident response processes, especially repetitive tasks and workflows, such as incident triage and response, threat hunting, and vulnerability scanning. Incident optimization and classifications can remove false alerts and identify and prioritize them based on their impact and severity on the enterprise. Predictive Analytics is a machine learning model that can be trained to predict the likelihood of a security incident based on historical data.

4. Applications Layer

The applications layer is the front end of SOC, where analysts interact with security tools to detect and respond to security incidents. SOC analysts use several applications for their daily activities; a couple are shown in Figure 6. Real-time monitoring dashboards enable analysts to continuously monitor the organization's systems and network to detect suspicious or anomalous activity. It provides real-time visibility into the security posture of an organization. It enables SOC managers to monitor the performance of the SOC and make informed decisions to improve the organization's security posture.

The incident management system is used to track and manage security incidents. It allows SOC analysts to assign incident ownership, track the progress of the incident investigation, and manage communication with stakeholders. Reporting and metric components include tools and dashboards that help SOC managers monitor the organization's security posture, track key performance indicators (KPIs), and generate reports for executives and other stakeholders. These reports and metrics help the organization understand the SOC's effectiveness and identify improvement areas. Furthermore, these metrics help communicate security issues across the organization. The forensics analysis tool is used by security analysts to investigate security incidents and analyze security-related data to identify the incident's root cause. Compliance monitoring ensures that the SOC complies with relevant regulatory requirements and standards, such as the Payment Card Industry Data Security Standard (PCI DSS) and the General Data Protection Regulation (GDPR). All four tiers in the architecture are essential but modular, allowing organizations to select an appropriate portfolio of software tools from multiple vendors.

Implementation guidelines of the proposed architecture

Ideally, a security operations center should have a holistic software tool or an integrated technology stack with SIEM, SOAR, and XDR capabilities. XDR is to detect and effectively block malware on endpoints. SIEM correlates threat data across multiple systems and data sources (and holds historical log data for threat-hunting). SOAR is used to orchestrate response activity and to integrate these and other tools across the IT environment to provide an orchestrated response to threats (Eirevo, n.d.). Security operations centers may adopt the NIST cybersecurity framework's (CSF) core functionalities, identification, protection, detection, response, and recovery (RSI, 2021). For implementing this architecture, enterprises need a holistic design including technologies, functions, automation, and deployment model (in-house, managed security services providers), etc.

Limitations and future research directions

This research revealed the broad adoption of artificial intelligence, machine learning techniques, and algorithms into security operations center tool kits. AI/ML-based SOC tools successfully reduce false positive alerts and help security analysts to prioritize and take remediation steps to reduce the risk of malicious activities. As of 2018, almost 52% of SOCs reported using AI/ML in some capacity (Crowley & Pescatore, 2018). The literature review revealed that AI/ML-driven algorithms are mainly used in SIEM (removing false positives and bringing attention to actual positive alerts) for investigation and incident management. Security software vendors also incorporated machine learning algorithms in XDR, SOAR, and UEBA. The results suggested that security tool vendors and user organizations have adopted AI/ML-based SOC tools and continually adapt and modify cybersecurity defenses (Kaliyaperumal, 2021).

This research is a preliminary study to understand the applicability and effectiveness of artificial intelligence and machine learning in security operations center tools. This study is based on a literature review and published reports by academics, research organizations, and security software vendors. The effectiveness of security tools is based on published articles and reports. Furthermore, this study has not developed a comprehensive set of metrics to measure the effectiveness of security operations centers and their underlying tools and technologies. The research did not test the machine learning algorithms' performance detecting and preventing threats and attacks. Additionally, this study does not include the effectiveness of AI/ML embedded security tools in preventing zero-day attacks.

Furthermore, the study does not cover the effectiveness of AI/ML algorithms in taking automatic corrective actions in the security incident mitigation phase. This research has not focused on the SOC tools' capabilities in the recovery, digital forensics, and lesson-learned phases of incident management. This research has not covered an in-depth study of security tools from major technology vendors about the type of AI/ML algorithms, including open-source ML libraries, and how software makers implemented algorithms and libraries in their SOC tools. Furthermore, this research has not covered vendors' functional test results, such as detecting and protecting the enterprise's digital assets from malicious activities. The level of effort and complexities SOC operators face in configuring and setting up parameters for the tools were not in the scope of this study. The effectiveness of the security operations center must be measured through a standard set of metrics. Further research is needed to define the SOC performance metrics and associated formulae. Security professionals require guidelines and best practices for testing, such as Red-Teaming and penetration testing, to measure the effectiveness of SOC tools. This research reveals the software's strengths in detecting and protecting enterprise resources and latency in preventing attacks. The security tools should be capable of investigation, and digital forensics and research should consider how to incorporate AI/ML-based algorithms in this effort. Working with vendors, researchers should uncover the type of algorithms and document the vendors' functional test results to gain customers' confidence.

References

Alsheh, E. (2022, June 4). Creating a smarter SOC with the MITRE ATT&CK Framework. *CyberProof*. Retrieved January 29, 2023, from https://blog.cyberproof.com/blog/creating-a-smarter-soc-with-the-mitre-attck-frameworkArcSight (n.d.). *SIEM+SOAR for threats that matters*. Retrieved January 28, 2023, from https://www.microfocus.com/en-us/cyberres/secops

Aslam, N., Khan, I. U., Mirza, S., AlOwayed, A., Anis, F. M., Aljuaid, R. M., & Baageel, R. (2022). Interpretable machine learning models for malicious domains detection using explainable artificial intelligence (XAI). *Sustainability*, 14(12), 7375. MDPI AG. Retrieved from http://dx.doi.org/10.3390/su14127375

Ban, T., Ndichu, S., Takahashi, T., & Inoue, D. (2021). Combat security alert fatigue with AI-assisted techniques. *CSET '21: Cyber Security Experimentation and Test Workshop August 2021*, 9–16. https://doi.org/10.1145/3474718.3474723

Capgemini Research Institute (2019). *Reinventing cybersecurity with artificial intelligence, the new frontier in digital security*. Retrieved October 22, 2022, from https://www.capgemini.com/wp-content/uploads/2019/07/AI-in-Cybersecurity_Report_20190711_V06.pdf

Check Point (n.d.). *What is a security operations center (SOC)?* Retrieved October 8, 2022, from https://www.checkpoint.com/cyber-hub/threat-prevention/what-is-soc/

Cobb, M. (n.d.). SIEM vs. SOAR vs. XDR: Evaluate the differences. *TechTarget*. Retrieved February 4, 2023, from https://www.techtarget.com/searchsecurity/tip/SIEM-vs-SOAR-vs-XDR-Evaluate-the-differences

Collins, J., Schneider, M., & Shoard, P. (2021, October 19). *SOC model guide*. Gartner, ID G00754096. Retrieved January 28, 2023, from https://www.gartner.com/doc/reprints?id=1-2C6FPM26&ct=230103&st=sb

Constantin, L. (2020, December 15). SolarWinds attack explained: And why it was so hard to detect. *CSO Online*. Retrieved February 19, 2023, from https://www.csoonline.com/article/3601508/solarwinds-supply-chain-attack-explained-why-organizations-were-not-prepared.html

Crowley, C. & Pescatore, J. (2018). The definition of SOC-cess? *SANS 2018 Security Operations Center Survey*, *SANS Institute Reading Room*, *SANS Institute*. Retrieved January 28, 2023, from https://assets.extrahop.com/whitepapers/Survey_SOC-2018_ExtraHop.pdf

Eirevo (n.d.). *SIEM SOAR and XDR: Differences explained*. Retrieved February 4, 2023, from https://eirevo.ie/blog/siem-vs-soar-vs-xdr/

Eriksson, H.S. (2022), *A user-centric approach to explainable AI in a security operation center environment* [Masters thesis, University of Oslo]. https://www.duo.uio.no/handle/10852/96758

Exabeam (n.d.). *The SOC, SIEM, and other essential SOC tools*. Retrieved January 28, 2023, from https://www.exabeam.com/explainers/siem/the-soc-secops-and-siem/

Farooq, H. M., & Otaibi, N. M. (2018). Optimal machine learning algorithms for cyber threat detection. *2018 UKSim-AMSS 20th International Conference on Computer Modelling and Simulation (UKSim),* 32-37. https://doi.org/10.1109/UKSim.2018.00018

Feng, C., Wu, S., & Liu, N. (2017). A user-centric machine learning framework for cyber security operations center. *2017 IEEE International Conference on Intelligence and Security Informatics (ISI), Beijing, China*. 173-175. https://doi.org/10.1109/ISI.2017.8004902

Forrester (2020). *The 2020 state of security operations*. Retrieved February 25, 2023, from https://www.paloguard.com/datasheets/forrester-the-2020-state-of-security.pdf

Gartner (n.d.). *Information Technology Gartner Glossary*. Retrieved March 1, 2023, from https://www.gartner.com/doc/reprints?id=1-2C6FPM26&ct=230103&st=sb

Gonzalez, C. (2020, December 8). Data exfiltration threats and prevention techniques you should know. *Exabeam*. Retrieved February 18, 2023, from https://www.exabeam.com/dlp/data-exfiltration/

IBM (n.d.). *Cybersecurity products help protect your organization with intelligence analysis, fraud protection, and mobile security solutions*. Retrieved September 19, 2023, from https://www.ibm.com/security/products?

InfosecMatter (2020, September 20). *Security Operations Center: Challenges of SOC teams*. Retrieved February 3, 2023, from https://www.infosecmatter.com/security-operations-center-challenges-of-soc-teams/

IIoT World (2022, January 28). *What is a SOC? Top security operations center challenges*. Retrieved February 12, 2023, from https://www.iiot-world.com/ics-security/cybersecurity/top-challenges-soc-are-facing/

Johnson, J.T. (n.d.). Building an effective security operations center framework. *TechTarget*. Retrieved January 27, 2023, from https://www.techtarget.com/searchsecurity/tip/Building-an-effective-security-operations-center-framework

Kaliyaperumal, L.N. (2021, October 21). The evolution of security operations and strategies for building an effective SOC. *ISACA Journal*, 5. Retrieved February 3, 2023, from https://www.isaca.org/resources/isaca-journal/issues/2021/volume-5/the-evolution-of-security-operations-and-strategies-for-building-an-effective-soc

Kim, A., Park, M., & Lee, D. H. (2020). AI-IDS: Application of deep learning to real-time web intrusion detection. *IEEE Access*, *8*, 70245-70261. https://doi.org/10.1109/ACCESS.2020.2986882

Kinyua, J., & Awuah, L. (2021). AI/ML in security orchestration, automation and response: Future research directions. *Intelligent Automation & Soft Computing*, *28*(2), 527-545. https://doi.org/10.32604/iasc.2021.016240

Kissel, C. (2021, January). Uplevel the SOC with one tool and the insights behind it. *IDC Vendor Spotlight*. Retrieved January 28, 2023, from https://www.checkpoint.com/downloads/resources/idc-spotlight-uplevel-your-soc.pdf

Knerler, K., Parker, I., & Zimmerman, C. (2022). 11 strategies of a world-class cybersecurity operations center, *MITRE*. Retrieved October 8, 2022, from https://www.mitre.org/sites/default/files/2022-04/11-strategies-of-a-world-class-cybersecurity-operations-center.pdf

Kumar, S., Singh, B. P., & Kumar, V. (2021). A semantic machine learning algorithm for cyber threat detection and monitoring security. *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, 1963-1967, https://doi.org/10.1109/ICAC3N53548.2021.9725596

Lindstrom, O. (2018). *Next generation security operations center* [Masters thesis, Metropolia University of Applied Sciences].

Logpoint (2022, September 30). *What are the advantages of SIEM+SOAR vs. XDR (Extended Detection & Response)?* Retrieved February 4, 2023, from https://www.logpoint.com/en/blog/what-are-the-advantages-of-siemsoar-vs-xdr/

Logsign (2020, August 28). *How can Cyber Kill Chain be valid for a SOC team? (Part 1)*. Retrieved January 29, 2023, from https://www.logsign.com/blog/how-cyber-kill-chain-can-be-useful-for-a-soc-team/

ManageEngine Log360 (n.d.). *Tools and technologies used in SOCs*. Retrieved January 27, 2023, from https://www.manageengine.com/log-management/siem/soc-tools-technologies.html

Microfocus (n.d.). *What is a Security Operations Center (SOC)?* Retrieved February 3, 2023, from https://www.microfocus.com/en-us/what-is/security-operations-center

Microsoft (n.d.). *What is a security operation center (SOC)?* Retrieved January 28, 2023. from https://www.microsoft.com/en-us/security/business/security-101/what-is-a-security-operations-center-soc

Mirilla, D.F. (2018). *Slow incident response in cyber security: The impact of task disengagement in security operations centers* [Doctoral dissertation, Pace University]. ProQuest Dissertations and Theses Global.

Nayyar, S. (2022, April 4). How to optimize security operations with machine learning and artificial intelligence. *Forbes*. Retrieved October 8, 2022, from https://www.forbes.com/sites/forbestechcouncil/2022/04/04/how-to-optimize-security-operations-with-machine-learning-and-artificial-intelligence/

NIST (2018). *Framework for improving critical infrastructure cybersecurity*. Retrieved March 1, 2023, from https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf

Optiv (n.d.). *Is your cybersecurity tech stack creating more gaps*? Retrieved February 20, 2023, from https://www.optiv.com/your-cybersecurity-tech-stack-creating-more-gaps

Palo Alto Networks (2020, May 6). *Artificial intelligence and machine learning in the security operation center*. Retrieved October 8, 2022, from https://www.paloaltonetworks.com/resources/techbriefs/artificial-intelligence-and-machine-learning-in-the-security-operations-center

Prasad, S. (2021). Cyber security operations center ML framework for the needs of the users. *International Journal of Machine Learning for Sustainable Development, 3(*3), 11-20. Retrieved from https://ijsdcs.com/index.php/IJMLSD/article/view/46

Perera, A., Rathnayaka, S., Perera, N. D., Madushanka, W.W., & Senarathne, A.N. (2021). The next-gen security operation center. *2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021*, 1-9. https://doi.org/10.1109/I2CT51068.2021.9418136

RSI (2021, September 16). *NIST security operations center best practices.* Retrieved February 21, 2023, from https://blog.rsisecurity.com/nist-security-operations-center-best-practices/

Sathana, M., & Hemamalini, M. (2022). A thread based machine learning framework for cyber security operations center. *International Journal of Research Publication and Reviews, 3*(5), 3683-3687.

Shutock, M., & Dietrich, G. (2022). Security Operations Centers: A holistic view on problems and solutions. *Proceedings of the 55th Hawaii International Conference on System Sciences.* 7555-7563. https://doi.org/10.24251/HICSS.2022.907

Splunk (n.d.). *10 Essential capabilities of a modern SOC*. E-Book, Retrieved September 19, 2023, from https://www.splunk.com/en_us/form/10-essential-capabilities-of-a-modern-security-operations-center.html

Trellix (n.d.). *What is Endpoint Detection and Response (EDR*)? Retrieved February 6, 2023, from https://www.trellix.com/en-us/security-awareness/endpoint/what-is-endpoint-detection-and-response.html

Yeshwanth, M. V., Kalluri, R., Rao, M. S., Kumar, R. K. S., & Bindhumadhava, B. S. (2022). Adoption and assessment of machine learning algorithms in security operations center for critical infrastructure. *In: Pillai, R.K., Ghatikar, G., Sonavane, V.L., Singh, B.P. (eds) ISUW 2020. Lecture Notes in Electrical Engineering*, 847. Springer, Singapore. https://doi.org/10.1007/978-981-16-9008-2_38