[1]Prasanth Alluri

Behavior-Based Cyber Defense Architectures for Enhancing the Resilience of Defense and National Critical Infrastructure

JES Journal of Electrical Systems

Abstract: Defense and national critical infrastructure systems are increasingly targeted by sophisticated cyber adversaries whose tactics exploit behavioral weaknesses rather than known technical vulnerabilities. Traditional signature-based security controls, while effective against previously observed threats, struggle to detect novel attacks, low-and-slow intrusions, insider misuse, and adaptive adversary campaigns that intentionally evade static indicators. This limitation poses a significant risk to mission-critical defense operations and essential services such as energy, transportation, telecommunications, and government systems, where cyber incidents can result in cascading operational, safety, and national security consequences. Despite growing interest in anomaly detection and machine learning for cybersecurity, there remains a lack of integrated, defense-oriented architectural frameworks that systematically operationalize behavior-based cyber defense while accounting for governance, resilience, and mission constraints. This article addresses this gap by proposing and analyzing a behavior-based cyber defense architecture tailored to defense and national critical infrastructure environments. The approach emphasizes continuous behavioral monitoring across users, hosts, networks, and operational technology assets, combined with adaptive analytics that establish dynamic baselines of normal activity. Rather than relying solely on static rules or signatures, the architecture integrates statistical profiling, machine learning-based anomaly detection, and adversary behavior modeling aligned with recognized threat frameworks. These components are embedded within a layered reference architecture that spans data collection, behavioral modeling, risk scoring, and controlled response, all governed by explicit assurance, audit, and privacy controls. Key findings from the synthesis of existing empirical studies and operational frameworks indicate that behavior-based approaches significantly improve the detection of advanced persistent threats, lateral movement, credential misuse, and anomalous control actions in cyber-physical systems when compared to signature-only defenses. When integrated with contextual risk scoring that incorporates asset criticality and mission impact, behavior-based analytics enable more effective prioritization of alerts and reduce analyst overload. The analysis further highlights that resilience outcomes, such as reduced dwell time, improved containment speed, and enhanced continuity of operations, are achievable when behavioral detection is coupled with governance workflows that manage model validation, deployment, and retraining in response to concept drift and adversarial adaptation. The article also identifies critical implementation challenges, including data quality limitations, false positive management, adversarial manipulation of learning models, and the need to balance behavioral monitoring with privacy and civil liberties protections. Addressing these challenges requires robust governance mechanisms, human-in-the-loop decision processes, and alignment with established cybersecurity and risk management standards. Overall, this work contributes a structured architectural and analytical foundation for behavior-based cyber defense, offering practical guidance for defense organizations and critical infrastructure operators seeking to enhance cyber resilience against evolving and adaptive threats.

***Keywords*:** Behavior analytics; anomaly detection; cyber resilience; critical infrastructure; defense cybersecurity; Zero Trust; OT security; SOC analytics; MITRE ATT&CK.

1. Introduction

1.1 Background: Why Defense and National Critical Infrastructure Are High-Value Targets

Defense systems and national critical infrastructure represent some of the most strategically valuable and operationally sensitive assets of modern states. These systems underpin military command and control, energy generation and distribution, telecommunications, transportation networks, public safety, and essential government services. Their disruption can produce consequences that extend beyond financial loss to include threats

[1]Principle IT Security Architect, California, USA

to national security, public safety, and geopolitical stability. As digital transformation accelerates across these sectors, the attack surface continues to expand, driven by increased connectivity, automation, and convergence between information technology (IT) and operational technology (OT) environments (Bhamare et al., 2020; Humayed et al., 2017).

State-sponsored adversaries, organized cybercriminal groups, and insider threats increasingly view critical infrastructure as a means of exerting strategic pressure. Advanced persistent threats target long-term espionage, prepositioning, and covert manipulation of systems, while ransomware operators exploit operational dependencies to maximize coercive leverage (Al-Sada et al.; Hutchins et al., 2011). In contrast to traditional enterprise environments, failures within national infrastructure systems can trigger cascading effects across interdependent sectors, amplifying the impact of even localized cyber incidents (Linkov C Kott, 2019; Woods, 2015).

The high value of these targets is further reinforced by their mission-critical nature and limited tolerance for downtime. Defense and infrastructure operators must maintain continuous availability while complying with strict safety, regulatory, and governance constraints. These conditions complicate the deployment of conventional cybersecurity controls and necessitate detection and response mechanisms that can operate effectively without disrupting essential services (Stouffer et al.; Ross et al., 2016).

1.2 Why Signature-Based Security Is Insufficient for Modern Threats

Signature-based security mechanisms, including traditional intrusion detection systems and antivirus tools, rely on predefined patterns derived from known threats. While effective against previously observed attacks, these approaches are fundamentally reactive and poorly suited to the dynamic and adaptive nature of contemporary cyber adversaries (Garcia-Teodoro et al., 2009; Patcha C Park, 2007). In defense and critical infrastructure environments, attackers deliberately design campaigns to evade static signatures through polymorphism, living-off-the-land techniques, and low-and-slow behaviors that blend into legitimate system activity (Sommer C Paxson, 2010).

Advanced persistent threats frequently operate over extended timeframes, emphasizing stealth and persistence rather than immediate exploitation. Such campaigns may involve legitimate administrative tools, credential abuse, and subtle configuration changes that produce no recognizable malicious signature (Al-Sada et al.; Caltagirone et al., 2013). Similarly, insider threats and supply chain compromises often manifest as deviations in behavior rather than explicit indicators of compromise, rendering signature-based detection ineffective (Anderson, 1980; Denning, 1987).

The limitations of signature-based security are particularly pronounced in OT and cyber- physical systems, where proprietary protocols, legacy devices, and safety constraints restrict the deployment of endpoint agents and frequent updates. In these environments, reliance on known signatures leaves operators blind to novel attack paths and emerging threat techniques, increasing dwell time and operational risk (Ding et al., 2018; Kriaa et al., 2015).

1.3 What "Behavior-Based Cyber Defense" Means in This Paper

In this paper, **behavior-based cyber defense** refers to security architectures and analytic approaches that focus on identifying deviations from established patterns of legitimate system, network, user, and process behavior rather than matching activity against predefined threat signatures. This paradigm builds upon foundational intrusion detection concepts that emphasize behavioral profiling and anomaly detection as core mechanisms for identifying malicious activity (Denning, 1987; Chandola et al., 2009).

Behavior-based defense encompasses multiple analytic layers, including user behavior analytics, host and process behavior monitoring, network traffic profiling, and OT process state analysis. These layers are integrated with machine learning and statistical models capable of adapting to evolving baselines and detecting subtle indicators of compromise, such as unusual authentication sequences, anomalous command execution, or protocol misuse (Ahmed et al., 2016; Buczak C Guven, 2016).

Importantly, behavior-based cyber defense is not positioned as a replacement for signature- based controls, but as a complementary and resilience-oriented capability. By emphasizing intent inference, contextual risk scoring, and continuous learning, behavior-based architectures enable earlier detection of novel threats and support informed response decisions aligned with mission and safety constraints (Mitchell C Chen, 2014; Rose et al., 2020).

1.4 Research Aims, Scope, and Contributions

The primary aim of this research is to develop a structured understanding of how behavior- based cyber defense architectures can enhance the resilience of defense and national critical infrastructure systems. The study focuses on environments characterized by high criticality, hybrid IT-OT integration, and constrained response options. It does not attempt to propose a single algorithmic solution, but rather to synthesize architectural principles, analytic methods, and governance considerations relevant to real-world deployment.

The key contributions of this paper are fourfold. First, it provides a threat-informed analysis of adversary behaviors targeting defense and critical infrastructure sectors, emphasizing operational impacts rather than purely technical indicators. Second, it articulates a reference architecture for behavior-based cyber defense that integrates data collection, analytics, decision support, and governance. Third, it examines sector-specific constraints and design considerations that shape the effectiveness of behavioral detection in OT and cyber-physical contexts. Finally, it identifies implementation challenges and future research directions necessary to operationalize behavior-based cyber defense at national scale (Linkov C Kott, 2019; Stouffer et al.).

1.5 Paper Organization

The remainder of this paper is organized as follows. Section 2 analyzes the contemporary threat landscape affecting defense and national critical infrastructure, including adversary categories, objectives, and operational impacts. Section 3 presents the theoretical foundations of behavior-based cyber defense and anomaly detection. Section 4 introduces a reference architecture for behavior-based cyber defense systems. Section 5 maps observed behaviors to adversary tactics and operational risk. Section 6 discusses sector- specific design considerations. Section 7 outlines implementation and governance workflows. Section 8 proposes evaluation metrics and assessment approaches. Section 9 addresses policy, privacy, and ethical considerations. Section 10 discusses limitations and future research directions, and Section 11 concludes the paper.

2. Threat Landscape for Defense and National Critical Infrastructure

2.1 Adversary Categories

Cyber threats targeting defense and national critical infrastructure originate from a diverse set of adversaries with distinct capabilities and objectives. **Advanced persistent threats** are typically state-aligned actors that conduct long-term campaigns involving reconnaissance, lateral movement, and stealthy persistence. These actors prioritize intelligence collection, strategic positioning, and the ability to disrupt operations during periods of geopolitical tension (Hutchins et al., 2011; Al-Sada et al.).

Insider threats, whether malicious or negligent, pose a significant risk due to their access to sensitive systems and knowledge of operational procedures. Behavioral deviations associated with insider misuse are often subtle and context-dependent, making them difficult to detect using traditional security controls (Anderson, 1980; Denning, 1987).

Supply chain compromises exploit trusted relationships between organizations and vendors. By inserting malicious components into software updates or hardware, adversaries can bypass perimeter defenses and operate within trusted environments for extended periods (Bhamare et al., 2020). **Ransomware operators**, while often financially motivated, increasingly target critical infrastructure to exploit high availability requirements and induce rapid compliance with extortion demands (Pinto et al.).

2.2 Attack Objectives and Operational Impacts

The objectives of attacks against defense and critical infrastructure extend beyond data theft. Adversaries seek to disrupt missions, degrade service availability, manipulate physical processes, and erode public trust. In defense contexts, successful intrusions may compromise command integrity or degrade situational awareness. In energy, transportation, and telecommunications sectors, cyber incidents can propagate across interconnected systems, leading to cascading failures and safety hazards (Woods, 2015; Kriaa et al., 2015).

Operational impacts are amplified by the tight coupling between cyber and physical processes. Malicious manipulation of control signals, timing parameters, or safety interlocks can result in equipment damage,

environmental harm, or loss of life. These risks necessitate early detection mechanisms that prioritize behavioral anomalies indicative of intent rather than relying solely on confirmed malware signatures (Ding et al., 2018; Stouffer et al.).

2.3 OT and Cyber-Physical Context: Why Detection and Response Are Harder

Operational technology and cyber-physical systems introduce unique challenges for cybersecurity. Many OT environments rely on legacy devices with limited computational capacity and long operational lifecycles, constraining the deployment of modern security agents and frequent patching (Humayed et al., 2017). Proprietary protocols and real-time performance requirements further complicate visibility and response.

Response actions that are acceptable in IT environments, such as isolating hosts or restarting services, may be unsafe or infeasible in OT contexts. As a result, detection accuracy and contextual understanding become critical. Behavior-based approaches offer advantages by leveraging passive monitoring and process-aware analytics to identify deviations without interfering with system operation (Mitchell C Chen, 2014; Sadeghi et al., 2015).

2.4 Typical Telemetry Sources in National Infrastructure Environments

Effective behavior-based cyber defense relies on diverse telemetry sources. Network-level data includes flow records, protocol metadata, and communication patterns. Endpoint telemetry encompasses process execution, file access, and memory usage. Identity and access data provide insight into authentication behavior, privilege escalation, and account misuse. In OT environments, sensor readings, control commands, and process state variables serve as critical indicators of normal and abnormal behavior (Ahmed et al., 2016; Stouffer et al.).

The fusion of these data sources enables multi-layered behavioral analysis, supporting correlation across domains and improving confidence in detection outcomes. However, this integration must be governed by strict access controls, data minimization principles, and operational constraints to ensure safety and compliance (Ross et al., 2016; Rose et al., 2020).

Table 1. Threat classes and behavior indicators across defense and national critical infrastructure sectors

Threat type	Key behavior signals	Impacted sectors
APT lateral movement	Unusual authentication sequences, abnormal administrative access, anomalous inter-host communication	Defense, energy, telecom
Insider misuse	Deviations in access timing, privilege escalation anomalies, atypical data access patterns	Defense, government, energy
Ransomware staging	Suspicious process injection, command execution chains, abnormal file system activity	Healthcare, transport, government
OT manipulation	Protocol misuse, abnormal control commands, deviation from physical process baselines	Energy, water, transportation

3. Foundations of Behavior-Based Cyber Defense

Behavior-based cyber defense is grounded in the premise that malicious activity, regardless of the tools or exploits used, ultimately manifests as **abnormal behavior** relative to an expected operational baseline. In defense and national critical infrastructure environments, where adversaries are often sophisticated, persistent, and adaptive, reliance on static signatures or known indicators is insufficient. Instead, resilient cyber defense requires continuous observation of system, network, and user behaviors, combined with analytical methods capable of detecting subtle deviations that may indicate emerging or stealthy threats.

This section establishes the conceptual foundations of behavior-based cyber defense by examining behavioral baselining, multi-level behavior modeling, detection paradigms, adversarial adaptation, and the operational metrics that govern real-world deployment.

3.1 Behavioral baselines and anomaly detection concepts

At the core of behavior-based cyber defense lies the concept of a **behavioral baseline**, which represents a statistical or analytical characterization of normal system activity over time. Unlike static rule sets, baselines are learned from observed data and capture patterns such as typical login frequencies, process execution sequences, network flows, and operational rhythms of cyber-physical systems.

Anomaly detection operates by comparing current observations against this baseline to identify deviations that exceed an acceptable threshold. In defense and critical infrastructure environments, anomalies may include unexpected privilege escalations, unusual lateral movement patterns, abnormal command sequences in operational technology systems, or deviations in traffic timing and volume that suggest covert data exfiltration or reconnaissance.

Importantly, anomaly detection does not assume prior knowledge of specific attack signatures. This makes it particularly well suited for identifying zero-day exploits, insider threats, and long-dwell advanced persistent threats that deliberately avoid known indicators. However, not all anomalies are malicious. Maintenance activities, mission surges, system upgrades, and emergency responses can also produce deviations. As a result, effective behavioral baselining must incorporate contextual awareness, asset criticality, and mission timelines to distinguish benign anomalies from genuine threats.

3.2 Behavior modeling levels: user, host, network, application, OT process

Behavior-based cyber defense operates across multiple analytical layers, each capturing different aspects of system activity. Modeling at a single level is rarely sufficient in isolation, particularly in complex defense and national infrastructure systems.

- **User-level behavior modeling** focuses on identity-driven actions, including authentication patterns, access timing, command usage, and role deviations. This level is essential for detecting credential misuse, insider threats, and account compromise, especially in environments with privileged operators or shared access constraints.
- **Host-level behavior modeling** examines endpoint activities such as process creation, memory access, file system interactions, and system calls. Abnormal host behavior can indicate malware execution, persistence mechanisms, or exploitation attempts that may not yet generate observable network signatures.
- **Network-level behavior modeling** analyzes communication patterns, including connection frequency, protocol usage, session duration, and data flow directionality. In critical infrastructure networks, where traffic patterns are often stable and predictable, even small deviations may signal reconnaissance, command-and- control activity, or lateral movement.
- **Application-level behavior modeling** focuses on logical workflows, API usage, transaction sequences, and error patterns. This layer is particularly relevant for defense logistics systems, command platforms, and government service applications, where misuse may occur without obvious infrastructure-level anomalies.
- **Operational technology process modeling** captures the behavior of industrial control systems, sensors, actuators, and physical processes. Deviations at this level may indicate unsafe commands, process manipulation, or cyber-physical attacks with real-world safety implications. Because OT environments prioritize availability and safety, behavioral models must be conservative and tightly aligned with engineering constraints.

The integration of these modeling levels enables cross-correlation, improving detection confidence while reducing reliance on any single data source.

3.3 Detection paradigms: rules plus analytics, statistical profiling, ML- based anomaly detection

Behavior-based cyber defense encompasses several complementary detection paradigms, each with distinct strengths and limitations.

- **Rule-based detection augmented by analytics** combines predefined logic with contextual enrichment. Rules encode known unsafe behaviors or policy violations, while analytics provide dynamic thresholds and contextual scoring. This approach offers transparency and control, which are critical in regulated or mission-sensitive environments, but it remains limited by rule coverage.

- **Statistical profiling** constructs mathematical representations of normal behavior using measures such as averages, variances, distributions, and correlations. Deviations are flagged when observations exceed statistically derived bounds. Statistical methods are computationally efficient and interpretable, making them attractive for large-scale monitoring, but they may struggle with complex, high- dimensional behaviors.

- **Machine learning-based anomaly detection** leverages supervised, unsupervised, or hybrid learning techniques to identify complex patterns that are difficult to encode manually. Unsupervised approaches are particularly valuable in defense settings where labeled attack data are scarce or classified. However, machine learning models introduce challenges related to explainability, training data quality, and susceptibility to adversarial manipulation.

In practice, defense-grade systems increasingly adopt **hybrid detection architectures** that combine rules, statistical methods, and machine learning. This layered approach balances transparency, adaptability, and detection coverage, while mitigating the weaknesses of any single paradigm.

3.4 Drift, deception, and adversarial behavior: why models must adapt

Behavioral models operate in environments that are inherently non-stationary. Legitimate system behavior evolves over time due to mission changes, software updates, infrastructure modernization, and shifts in operational tempo. This phenomenon, commonly referred to as **concept drift**, can degrade model accuracy if not properly managed.

Adversaries further complicate this landscape by deliberately attempting to evade detection. Techniques such as low-and-slow attacks, living-off-the-land behaviors, and mimicry of legitimate user actions are specifically designed to blend into established baselines. In some cases, attackers may actively probe detection thresholds to shape model behavior over time.

As a result, behavior-based cyber defense systems must incorporate **adaptive mechanisms**, including continuous model retraining, feedback from analyst decisions, and periodic baseline reassessment. Adaptation must be governed carefully, particularly in defense and critical infrastructure contexts, to prevent attackers from poisoning training data or inducing unsafe responses.

Controlled adaptation, supported by governance checkpoints and human oversight, is therefore a foundational requirement rather than an optional enhancement.

3.5 Evaluation metrics and operational constraints

The effectiveness of behavior-based cyber defense cannot be assessed solely through technical accuracy metrics. Operational realities impose additional constraints that directly affect mission success.

- **False positives** represent a primary concern, as excessive alerts can overwhelm analysts and erode trust in detection systems. In high-stakes defense environments, alert fatigue may lead to delayed responses or missed genuine threats.

- **Detection latency**, often measured as time-to-detect, is equally critical. Advanced threats frequently seek to establish persistence before acting, and delayed detection increases the risk of mission disruption or physical damage in cyber-physical systems.

- **Operational scalability** must also be considered. Behavior-based analytics must process large volumes of heterogeneous data without introducing unacceptable delays or resource overhead.

- Finally, **actionability** determines whether detected anomalies can realistically be investigated and mitigated. Alerts that lack context, confidence scoring, or recommended response actions impose additional cognitive burden on analysts.

Balancing detection sensitivity against operational burden is therefore a central design challenge in behavior-based cyber defense.

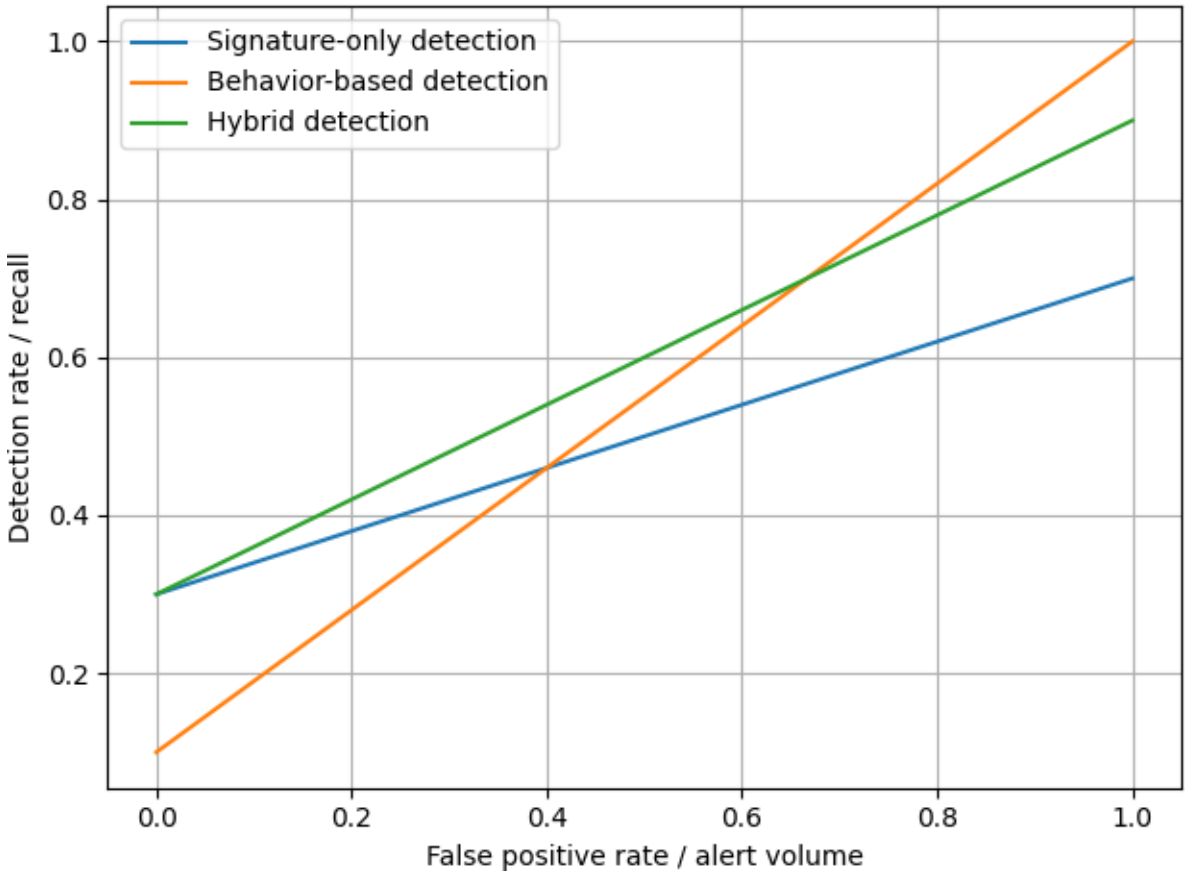

Figure 1. Detection tradeoff between sensitivity and operational burden

The chart illustrates the tradeoff between detection sensitivity and alert volume across three monitoring approaches. Signature-only systems show relatively low alert volume but limited detection capability against novel or stealthy threats. Behavior-based systems achieve higher detection rates but at the cost of increased alert volume. Hybrid approaches balance sensitivity and operational feasibility by combining behavioral analytics with contextual filtering and rule-based controls.

4. Reference Architecture for Behavior-Based Cyber Defense

This section presents a **defense-grade reference architecture** for behavior-based cyber defense, designed to enhance resilience across military systems and national critical infrastructure. The architecture emphasizes continuous visibility, adaptive detection, controlled response, and governance by design. It reflects operational realities in high-risk environments where availability, safety, and mission assurance are as critical as confidentiality.

4.1 Architecture goals: resilience, visibility, and controlled response

The primary goal of a behavior-based cyber defense architecture is **resilience rather than perimeter prevention**. Instead of assuming breaches can be fully avoided, the architecture assumes adversaries will penetrate systems and focuses on early detection, impact containment, and rapid recovery.

Three core objectives guide the architecture:

- **Resilience:** Maintain mission-critical functions despite intrusions by enabling rapid detection, isolation, and recovery. The system must degrade gracefully rather than fail catastrophically.
- **Visibility:** Establish continuous, cross-layer visibility into user, system, network, and operational technology behavior. This visibility is essential for identifying subtle deviations that signature-based tools miss.

- **Controlled response:** Ensure that defensive actions are proportionate, explainable, and safe, especially in environments where automated responses can disrupt physical processes or military operations.

These objectives directly address challenges in defense and critical infrastructure systems, where false positives, unsafe automation, or opaque decision-making can be as damaging as cyberattacks themselves.

4.2 Data layer: sensors, logs, packet metadata, EDR, identity, OT telemetry

The **data layer** forms the foundation of behavior-based defense by aggregating heterogeneous telemetry across IT, OT, and cyber-physical environments.

Key data sources include:

- **Network telemetry:** Flow records, packet metadata, DNS activity, and protocol-level summaries that reveal lateral movement, command-and-control patterns, and anomalous traffic behavior.
- **Endpoint and workload data:** EDR signals, process creation events, memory access patterns, file system activity, and kernel-level behaviors that indicate exploitation or persistence.
- **Identity and access data:** Authentication logs, privilege escalation events, access timing anomalies, and identity graph relationships that expose compromised or misused accounts.
- **Operational technology telemetry:** Sensor readings, control commands, and protocol states from ICS and SCADA systems, captured in a passive and safety-aware manner.
- **System and application logs:** Operating system logs, middleware events, and application-specific audit trails that provide context for behavioral deviations.

To support defense environments, the data layer must be **tamper-resistant, time- synchronized, and resilient to partial outages**, ensuring continuity of monitoring even under attack.

4.3 Analytics layer: feature engineering, models, correlation, threat intelligence mapping

The **analytics layer** transforms raw telemetry into interpretable behavioral insights. It combines statistical methods, machine learning, and contextual correlation to distinguish benign variability from malicious activity.

Core functions include:

- **Feature engineering:** Extraction of temporal, relational, and sequence-based features such as session duration, command frequency, peer group deviation, and protocol state transitions.
- **Behavioral modeling:** Construction of baselines at multiple levels including user, host, network segment, and OT process. Models may be unsupervised, semi- supervised, or hybrid depending on data availability.
- **Cross-domain correlation:** Linking events across identity, endpoint, network, and OT domains to detect coordinated attack behavior rather than isolated anomalies.
- **Threat intelligence alignment:** Mapping observed behaviors to adversary tactics and techniques using frameworks such as **MITRE ATTsCK**, enabling intent inference and campaign-level understanding.

This layer prioritizes **explainability and adaptability**, recognizing that defense operators must trust and understand analytic outputs under high-pressure conditions.

4.4 Decision layer: risk scoring, confidence, policy constraints, human-in- the-loop

The **decision layer** converts analytic findings into actionable security judgments while balancing automation with human oversight.

Key elements include:

- **Risk scoring:** Aggregation of detection confidence, asset criticality, mission impact, and potential blast radius into a composite risk score.
- **Confidence assessment:** Explicit representation of uncertainty, allowing analysts to distinguish high-confidence

intrusions from low-confidence anomalies.

- **Policy constraints:** Enforcement of mission, safety, and regulatory rules that limit which actions can be taken automatically, especially in OT and defense contexts.
- **Human-in-the-loop controls:** Analyst review points for high-impact decisions, ensuring accountability and preventing unsafe or irreversible actions.

This layer acts as a **control buffer**, preventing raw analytics from directly triggering disruptive responses without appropriate validation.

4.5 Response layer: containment, segmentation, account controls, OT- safe actions

The **response layer** executes defensive actions in a graduated and context-aware manner. Responses are selected based on risk level, system criticality, and operational constraints.

Typical response actions include:

- **Network containment:** Dynamic segmentation, traffic throttling, and micro- isolation to limit lateral movement.
- **Identity controls:** Session termination, credential revocation, step-up authentication, and privilege reduction.
- **Endpoint actions:** Process suspension, file quarantine, or controlled system reboot where operationally acceptable.
- **OT-safe actions:** Passive monitoring escalation, alerting, and operator-assisted intervention rather than aggressive automation.

The emphasis is on **containment and stabilization first**, followed by eradication and recovery once mission safety is assured.

4.6 Assurance layer: governance, audit, privacy, and compliance controls

Surrounding all functional layers is the **assurance layer**, which ensures that the architecture remains trustworthy, lawful, and aligned with defense governance requirements.

This layer provides:

- **Governance mechanisms:** Model approval workflows, change management, and version control for analytics and response logic.
- **Auditability:** Comprehensive logging of detections, decisions, and actions to support forensic analysis, accountability, and external review.
- **Privacy protections:** Data minimization, role-based access, and privacy-preserving analytics to limit unnecessary exposure of sensitive personnel or operational data.
- **Compliance alignment:** Mapping of controls and processes to national cybersecurity frameworks, defense standards, and critical infrastructure regulations.

By embedding assurance directly into the architecture, behavior-based cyber defense systems can scale without eroding trust or oversight.

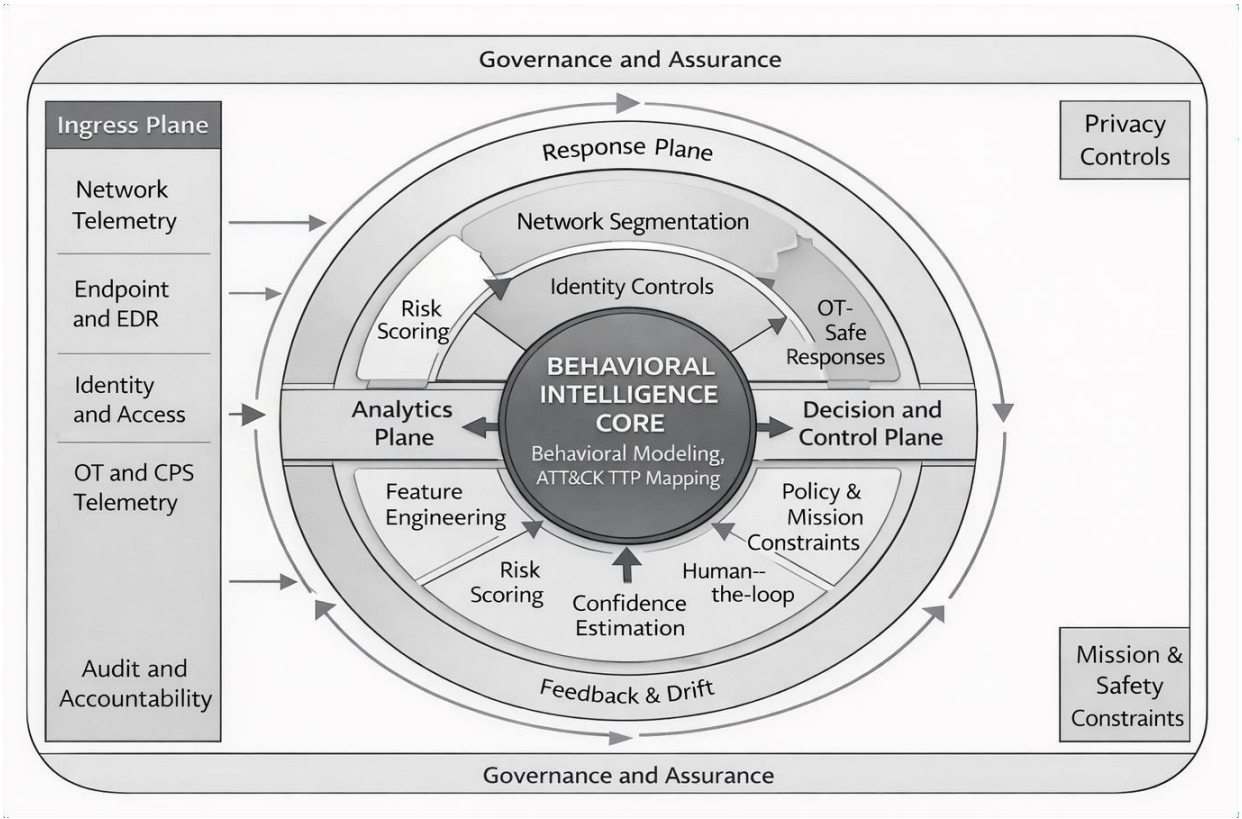

Figure 2. Behavior-based cyber defense reference architecture

This figure presents a conceptual reference architecture for behavior-based cyber defense tailored to defense and national critical infrastructure environments. The architecture is organized around a central behavioral intelligence core that performs behavioral modeling and adversary tactic mapping. Concentric layers represent analytics, decision and control, and response functions, emphasizing continuous adaptation and resilience rather than linear detection workflows. A dedicated ingress plane supplies multi-source telemetry, while feedback mechanisms enable model refinement and drift management. Governance and assurance mechanisms surround the architecture to enforce auditability, privacy protection, and mission and safety constraints.

5. Mapping Behaviors to Adversary Tactics and Operational Risk

5.1 Translating observed behaviors into TTPs (MITRE ATTsCK alignment)

Behavior-based cyber defense starts by converting low-level signals into semantically meaningful adversary actions. Raw telemetry such as authentication logs, process creation events, network flow metadata, and OT protocol traces is first normalized and contextualized against asset roles and expected baselines. These observed behaviors are then aligned to tactics, techniques, and procedures (TTPs) using the MITRE ATTCCK knowledge base.

Alignment does not rely on single indicators. Instead, behavior clusters are mapped to ATTCCK techniques when multiple reinforcing signals appear within a bounded time window and asset context. For example, a combination of abnormal Kerberos ticket requests, privilege escalation attempts, and unusual east-west traffic patterns supports mapping to credential access and lateral movement tactics. This structured translation improves analytic explainability and allows defenders to reason in terms of adversary intent rather than isolated alerts, which is essential for defense and national infrastructure environments where false positives are costly.

5.2 From alert to campaign understanding: correlation and intrusion chain logic

Individual alerts rarely represent the full scope of an intrusion. Behavior-based architectures therefore employ correlation logic to connect alerts across hosts, identities, and network segments into intrusion chains. Temporal sequencing, shared infrastructure artifacts, and repeated behavioral motifs are used to infer campaign-level activity.

Intrusion chain logic supports the transition from reactive alert handling to proactive campaign disruption. By reconstructing sequences such as initial access followed by persistence, discovery, and lateral movement, security operations centers can identify the current phase of adversary activity and anticipate likely next steps. This approach reduces analyst workload by collapsing dozens of low-confidence alerts into a small number of high- confidence campaign narratives, enabling earlier containment and more targeted response.

5.3 Risk scoring model: mission impact, asset criticality, confidence, and blast radius

To support operational decision-making, mapped behaviors and correlated campaigns are evaluated through a multi-dimensional risk scoring model. The model integrates four primary factors:

- **Mission impact:** The potential effect on operational continuity, safety, or national security objectives if the behavior progresses unchecked.
- **Asset criticality:** The role of the affected system within defense or infrastructure operations, including dependencies and failover capacity.
- **Analytic confidence:** The strength of evidence supporting the behavioral interpretation, based on signal consistency and historical accuracy.
- **Blast radius:** The estimated scope of propagation across interconnected systems, identities, or physical processes.

These dimensions are combined into a normalized risk score that supports consistent prioritization across heterogeneous environments. In OT and mission systems, risk scoring is explicitly constrained by safety and availability considerations, ensuring that high-risk detections do not automatically trigger disruptive responses.

5.4 Prioritization in high-noise environments: SOC workflow integration

Defense and national infrastructure SOCs operate under sustained alert pressure. Behavior- based prioritization integrates directly into SOC workflows by presenting analysts with ranked campaigns, contextual evidence, and recommended actions rather than raw alerts.

This integration supports tiered analysis. Automated analytics handle routine correlation and scoring, while human analysts focus on high-risk cases requiring judgment, coordination, or command authorization. Playbooks linked to ATTCCK techniques guide response selection, while audit logging preserves evidence for post-incident review and compliance. The result is a measurable reduction in alert fatigue and faster time-to-decision without sacrificing analytic rigor.

Table 2. ATTsCK-informed behavior-to-tactic mapping

Observed behavior	ATTsCK tactic / technique	Likely objective	Recommended response	Risk notes for OT or mission systems
Abnormal privileged login outside baseline hours	Credential Access / Valid Accounts	Account takeover for lateral movement	Enforce step-up authentication, isolate account	Avoid immediate lockout on mission-critical identities
Repeated service creation on multiple hosts	Persistence / Create or Modify System Process	Maintain long-term foothold	Disable service, verify binaries	Validate operational dependencies before removal
Unusual east-west traffic between enclaves	Lateral Movement / Remote Services	Expand access across segments	Segment isolation, traffic filtering	Ensure segmentation does not disrupt control traffic
Unauthorized protocol commands in OT network	Impact / Manipulation of Control	Alter physical process behavior	Switch to safe mode, block command source	Prioritize safety and manual override capability
Coordinated data staging prior to exfiltration	Exfiltration / Data Staged	Prepare sensitive data theft	Quarantine host, monitor egress	Assess national security or safety sensitivity

6. Sector-Specific Design Considerations

6.1 Defense enterprise networks

Defense enterprise networks are characterized by segmented enclaves, classified domains, and strict access controls. Behavior-based detection must emphasize identity abuse, credential misuse, and cross-domain movement. Analytics are designed to operate under partial visibility and classification boundaries, with federation across enclaves enabling pattern recognition without centralizing sensitive data.

6.2 Energy systems and ICS or SCADA

Energy infrastructure prioritizes availability and safety over aggressive response. Behavior models focus on deviations in control commands, timing irregularities, and unexpected interactions between IT and OT layers. Responses are constrained by safety engineering principles, favoring observation, manual confirmation, and controlled isolation over automated shutdown.

6.3 Telecommunications and backbone networks

Telecommunications systems generate high-volume telemetry and require scalable analytics. Behavior-based defense emphasizes traffic pattern anomalies, signaling abuse, and routing manipulation. Models must distinguish malicious deviation from legitimate load variation to avoid service disruption at national scale.

6.4 Transportation systems

Transportation environments exhibit tight coupling between digital systems and physical movement. Behavior analytics target anomalies in scheduling systems, signaling interfaces, and vehicle telemetry. Design must account for cascading risk, where localized cyber events propagate into broader operational delays or safety hazards.

6.5 Government services and public safety

Government and public safety systems demand high availability and strong data protection. Behavior-based defense focuses on insider misuse, unauthorized data access, and coordinated service degradation. Analytics are integrated with privacy controls and legal oversight to maintain public trust while supporting rapid incident response.

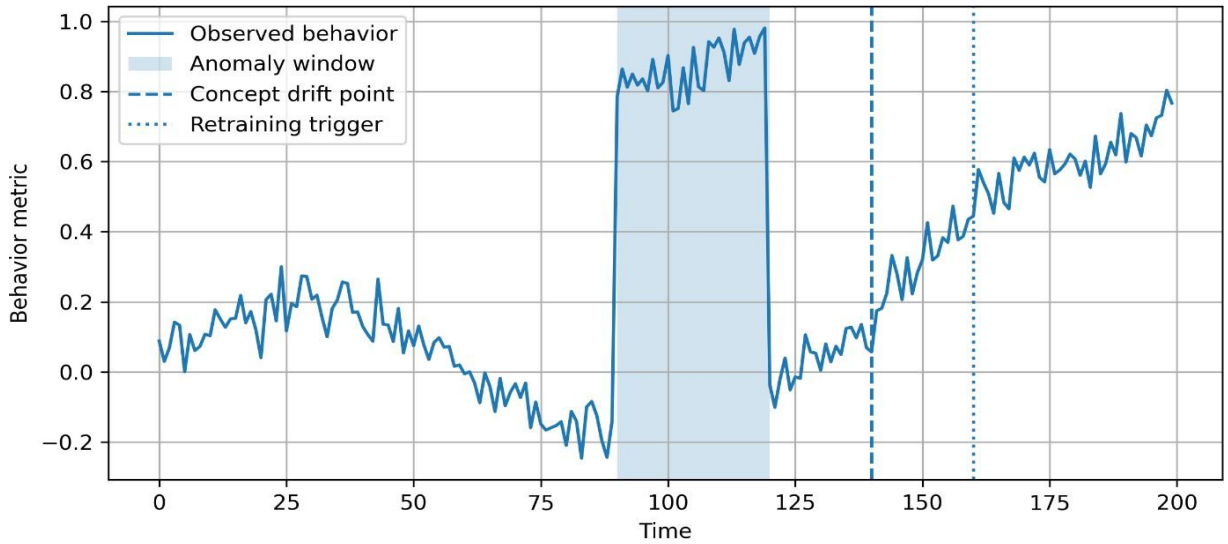

Figure 3. Example baseline deviation over time

The graph presents a time-series representation of a behavioral metric under normal

operating conditions, where system activity remains within an established baseline range.

The shaded anomaly window highlights a period of significant deviation from expected behavior, reflecting abnormal activity that exceeds predefined thresholds and warrants analyst attention. Such deviations typically correspond to emerging threats,

misconfigurations, or abnormal user or system actions within defense or critical infrastructure environments.

Beyond the anomaly phase, the graph identifies a concept

drift point, indicating a sustained shift in behavioral patterns rather than a transient event. This drift necessitates model review and retraining to ensure continued detection accuracy as operational conditions evolve. The retraining trigger marker illustrates how behavior- based cyber defense systems adapt over time, maintaining resilience by balancing sensitivity to new threats with stability in high-availability and mission-critical systems.

7. Implementation Blueprint

Implementing behavior-based cyber defense architectures within defense and national critical infrastructure environments requires more than algorithmic capability. Effective deployment depends on architectural choices, disciplined data engineering, controlled model governance, human operational workflows, and explicit resilience engineering. This section translates the conceptual architecture into an operationally viable implementation blueprint.

7.1 Deployment Models: Centralized SOC, Distributed Monitoring, and Hybrid Architectures

Deployment models for behavior-based cyber defense must reflect organizational scale, mission sensitivity, and network topology.

A **centralized Security Operations Center (SOC)** model consolidates behavioral analytics, correlation engines, and response coordination within a single operational hub. This approach enables consistent policy enforcement, unified visibility, and efficient expertise utilization, making it suitable for defense ministries or national-level operators with stable connectivity and mature SOC capabilities. However, centralized models can introduce latency and create single points of failure if connectivity to monitored environments is disrupted.

In contrast, **distributed monitoring architectures** place behavioral detection components closer to assets, such as at military bases, substations, transport hubs, or tactical networks. Localized analytics improve responsiveness, reduce data exfiltration risks, and support operations in disconnected or degraded environments. The primary challenge lies in maintaining consistent detection logic and governance across heterogeneous sites.

A **hybrid deployment model** combines centralized oversight with distributed analytics. Behavioral baselines and anomaly detection operate locally, while higher-level correlation, campaign analysis, and governance functions are centralized. This model is increasingly favored for national critical infrastructure, as it balances resilience, scalability, and strategic coordination while limiting excessive data aggregation.

7.2 Data Engineering: Collection, Normalization, Storage, Retention, and Access Control

Behavior-based cyber defense depends fundamentally on reliable and representative data pipelines. Data engineering must be treated as a security-critical function rather than a supporting technical detail.

- **Data collection** spans multiple layers, including network telemetry, endpoint activity, identity and access logs, application behavior, and OT or cyber-physical process signals. Sensor placement must reflect mission criticality and safety constraints, especially in industrial and defense systems where intrusive monitoring may be unacceptable.
- **Normalization and enrichment** are required to align heterogeneous data sources into a unified behavioral feature space. Time synchronization, schema harmonization, and contextual enrichment, such as asset criticality and mission role, are essential to avoid misleading anomaly detection outcomes.
- **Storage and retention policies** must balance forensic value with legal, privacy, and operational constraints. Defense environments often require extended retention for attribution and post-incident analysis, while critical infrastructure operators must comply with sector-specific regulations. Tiered storage models are commonly used, retaining high-fidelity recent data and aggregating older records.
- **Access control and segregation** are necessary to prevent misuse of sensitive behavioral data. Role-based access, audit logging, and strict separation between detection, investigation, and administrative functions reduce insider risk and support compliance requirements.

7.3 Model Lifecycle: Training, Validation, Tuning, and Controlled Rollout

Unlike static security rules, behavior-based models evolve continuously and therefore require formal lifecycle governance.

- **Model training** begins with baseline establishment using historical and live data that reflect normal operational behavior. For defense and infrastructure systems, training datasets must explicitly exclude known incident periods to prevent normalization of malicious behavior.
- **Validation and risk review** assess detection performance, false positive rates, and potential operational impact. This phase is critical in OT and mission-critical environments, where incorrect detections can disrupt safety or continuity.
- **Controlled rollout** introduces models in monitoring or advisory mode before enabling automated responses. Phased deployment allows analysts to assess trustworthiness and adjust thresholds while maintaining operational stability.
- **Ongoing tuning and retraining** are triggered by detected drift, environmental changes, or evolving adversary tactics. All changes must be documented, auditable, and approved through governance checkpoints to prevent uncontrolled model behavior.

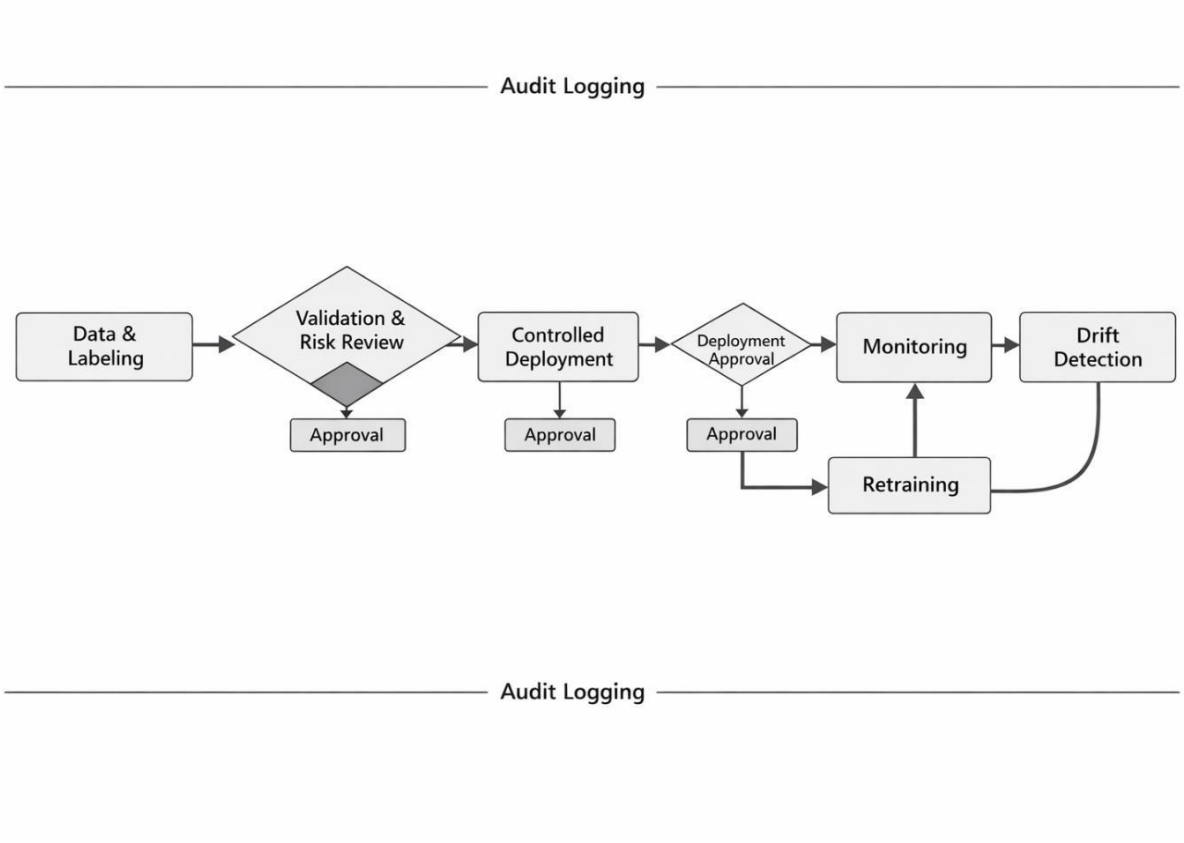

Figure 4. Governance Workflow for Model Lifecycle Management in Behavior-Based Cyber Defense Systems

This diagram illustrates a governance-controlled lifecycle for behavior-based cyber defense models. Models progress from data collection and labeling through validation, risk review, and controlled deployment under formal approval gates. Continuous monitoring and drift detection ensure that changes in system behavior or threat patterns trigger retraining, while audit logging spans all stages to support accountability, compliance, and trustworthy operation in defense and national critical infrastructure environments.

7.4 Human Workflow: Triage Playbooks, Escalation, and Evidence Packaging

Human operators remain central to behavior-based cyber defense, particularly in defense and national infrastructure contexts where automated actions may have strategic or safety implications.

- **Triage playbooks** guide analysts in interpreting behavioral alerts, distinguishing benign deviations from malicious intent, and correlating signals across domains. Well-designed playbooks reduce cognitive load and ensure consistent responses across shifts and teams.
- **Escalation paths** define when alerts transition from monitoring to investigation, containment, or command-level decision making. In defense environments, escalation may involve operational commanders, legal advisors, or safety officers, emphasizing the need for clear thresholds and responsibilities.
- **Evidence packaging** transforms behavioral detections into defensible incident records. Structured timelines, behavioral summaries, and mapped adversary actions support forensic analysis, compliance audits, and potential legal or policy responses. Proper evidence handling also strengthens collaboration with external partners and national authorities.

7.5 Resilience Engineering: Graceful Degradation and Continuity Under Attack

Behavior-based cyber defense architectures must themselves be resilient to attack, overload, or partial failure.

- **Graceful degradation** ensures that core detection and monitoring capabilities persist even when advanced analytics or centralized coordination are unavailable. For example, local anomaly detection may continue operating with reduced fidelity during connectivity loss.
- **Continuity planning** integrates cyber defense into broader resilience strategies, including failover mechanisms, backup SOC capabilities, and predefined operational modes for degraded conditions. Defense and infrastructure operators must assume that cyber incidents will coincide with physical, geopolitical, or environmental stressors.

By embedding resilience engineering principles into system design, behavior-based defenses contribute not only to threat detection but also to sustained mission assurance.

8. Evaluation Framework and Metrics

Evaluating behavior-based cyber defense requires a multidimensional framework that captures technical performance, operational effectiveness, and resilience outcomes.

8.1 Technical Metrics

Technical metrics assess detection accuracy and responsiveness:

- **Precision and recall** quantify the balance between false positives and missed detections.
- **Time to detect** measures how quickly behavioral deviations are identified after onset.
- **Time to respond** captures the interval between detection and effective mitigation.

These metrics must be interpreted in context, as aggressive detection thresholds may improve recall at the cost of operational feasibility.

8.2 Operational Metrics

Operational metrics evaluate the impact on security teams and processes:

- **Analyst workload** and alert volume per shift
- **False positive cost**, including investigation time and disruption
- **Coverage**, reflecting the proportion of assets and behaviors monitored
- **Dwell time reduction**, indicating how long adversaries remain undetected

Improvements in these metrics demonstrate whether behavior-based analytics translate into practical SOC gains.

8.3 Resilience Metrics

Resilience metrics focus on mission and service outcomes:

- **Service continuity** during cyber incidents
- **Recovery time objectives** and actual recovery performance
- **Mission impact reduction**, particularly for defense and safety-critical systems

These measures align cyber defense evaluation with organizational resilience goals rather than purely technical success.

8.4 Testing Approaches

Robust evaluation combines multiple testing methods:

- **Red teaming** to simulate realistic adversary behavior
- **Purple teaming** to iteratively improve detection and response
- **Attack emulation** using known tactics and techniques
- **Tabletop exercises** to assess decision making, escalation, and coordination Together, these approaches validate both technical systems and human workflows.

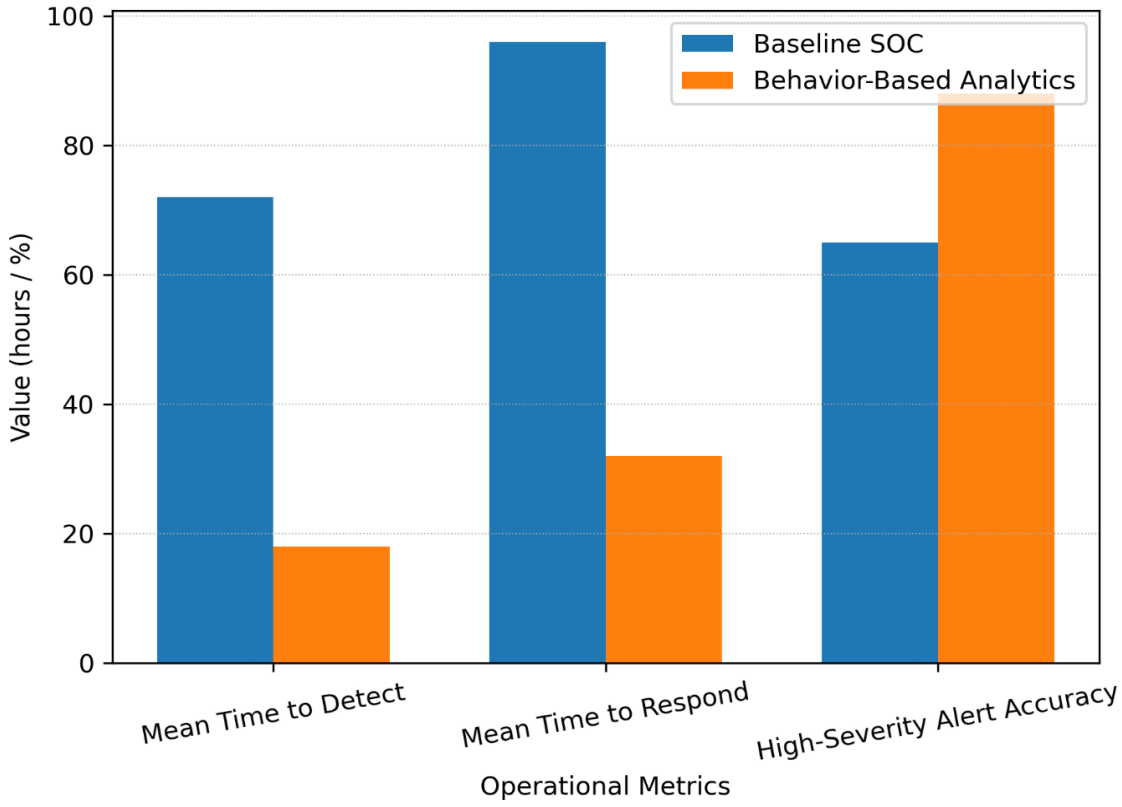

Figure 5. Operational Performance Improvements After Integrating Behavior-Based Analytics into the Defense SOC Workflow

Figure 5 illustrates the operational impact of integrating behavior-based analytics into a defense Security Operations Center. Compared with a baseline SOC configuration, the behavior-based architecture significantly reduces mean time to detect and mean time to respond while improving the accuracy of high-severity alerts. These improvements indicate enhanced situational awareness, more efficient analyst workflows, and stronger cyber

resilience in defense and national critical infrastructure environments.

G. Governance, Privacy, and Policy Alignment

Behavior-based cyber defense architectures for defense and national critical infrastructure must operate within strict governance, privacy, and policy constraints. Unlike purely technical security controls, behavior analytics directly observe patterns of human and system activity, which raises heightened concerns around access control, civil liberties, accountability, and long-term trust. This section explains how such architectures can be aligned with Zero Trust principles, privacy-preserving practices, regulatory compliance, ethical safeguards, and public-sector procurement realities, while still delivering meaningful resilience benefits.

G.1 Zero Trust Alignment and Access Governance

Zero Trust security models assume that no user, device, or system should be inherently trusted, regardless of network location. Behavior-based cyber defense naturally complements this paradigm by continuously evaluating trust based on observed activity rather than static credentials or perimeter assumptions.

Within a Zero Trust architecture, behavior analytics strengthen access governance in several ways. First, they enable continuous authentication and authorization by monitoring deviations from established behavioral baselines, such as unusual login times, abnormal command sequences, or atypical access paths across segmented environments (Rose et al., 2020). Second, behavioral risk scores can be integrated into policy decision points to dynamically adjust access privileges, such as triggering step-up authentication, session isolation, or temporary privilege reduction when anomalous behavior is detected. Third, these mechanisms improve visibility across hybrid IT and OT environments, where traditional identity-centric controls alone may be insufficient.

For defense and national infrastructure operators, governance is critical to ensure that behavior-driven access decisions remain transparent, auditable, and aligned with mission priorities. Clear policies must define which behaviors influence access control, how confidence thresholds are set, and when human approval is required before enforcement actions occur. This governance layer ensures that Zero Trust enforcement enhances resilience without introducing operational fragility or unintended denial of service.

G.2 Privacy-Preserving Monitoring

Behavior-based monitoring introduces legitimate privacy concerns, particularly when applied to government employees, contractors, or operators of critical infrastructure. Effective governance therefore requires privacy-by-design principles to be embedded into system architectures from the outset.

Data minimization is the primary control. Only telemetry strictly necessary for security objectives should be collected, retained, and analyzed. For example, metadata about access patterns or command frequency may be sufficient for anomaly detection, without capturing content or sensitive payloads (Ahmed et al., 2016). Retention periods should be explicitly defined and aligned with operational and legal requirements.

Privacy-enhancing techniques can further reduce risk. k-anonymity may be applied to aggregated behavioral datasets used for model training, ensuring that individual users cannot be re-identified within large populations (Sweeney, 2002). In environments where large-scale analytics or cross-organizational sharing is required, differential privacy mechanisms can be introduced to add controlled noise while preserving statistical utility (Dwork, 2006). While such techniques may slightly reduce detection precision, they significantly improve societal and legal acceptability.

Importantly, privacy controls must be enforceable and verifiable. Defense organizations should document which data fields are collected, how they are protected, and how privacy risks are assessed during system updates or model retraining. This transparency is essential for maintaining trust among personnel and oversight bodies.

G.3 Compliance, Auditability, and Evidence Handling

National critical infrastructure operators are subject to extensive regulatory, contractual, and policy obligations. Behavior-based cyber defense systems must therefore support strong compliance and auditability features to be viable in real-world deployments.

Alignment with established standards such as NIST SP 800-53, NIST SP 800-207, and ISO/IEC 27001 ensures that behavior analytics are treated as part of an integrated control environment rather than an ad hoc technical add-on (Joint Task Force, 2020; Rose et al., 2020; ISO/IEC, 2022). Logging, access decisions, model outputs, and response actions should all be traceable and time-stamped to support internal audits and external reviews.

Evidence handling is particularly important in defense and national security contexts, where cyber incidents may have legal, disciplinary, or strategic implications. Behavioral alerts and analytics outputs must be preserved in a manner that supports chain-of-custody requirements, forensic reconstruction, and post-incident review (Cichonski et al., 2012).

This includes documenting how models generated specific alerts, what data sources were involved, and which policies governed the response.

By embedding auditability into system design, organizations can demonstrate due diligence, justify enforcement decisions, and reduce institutional risk associated with automated or semi-automated security actions.

G.4 Ethical Boundaries in Behavior Monitoring and Insider Threat Programs

Behavior-based cyber defense inevitably intersects with insider threat detection, making ethical boundaries a central concern. While monitoring is necessary to protect critical systems, excessive or opaque surveillance can erode morale, trust, and organizational legitimacy.

Ethical governance requires clear separation between security monitoring and performance evaluation. Behavioral analytics should focus on security-relevant indicators rather than productivity, personal habits, or non-security-related activities. Access to detailed behavioral data should be restricted to authorized security personnel, and escalation pathways should include human review, particularly for high-impact actions such as account suspension or disciplinary referral.

Proportionality is another key principle. Monitoring intensity should reflect asset criticality and threat level, rather than applying uniform surveillance across all users and systems. Transparent communication with staff about monitoring objectives, safeguards, and oversight mechanisms helps reduce misunderstanding and resistance.

In defense environments, where insider threats can have severe consequences, ethical restraint is not a weakness but a resilience factor. Programs that are perceived as fair, limited, and accountable are more likely to gain cooperation from personnel and produce reliable security outcomes (Hollnagel et al., 2006).

G.5 Procurement and Interoperability for National Infrastructure Operators

Public-sector and critical infrastructure operators face unique procurement and interoperability challenges when adopting behavior-based cyber defense solutions. Systems must integrate with heterogeneous legacy environments, comply with procurement regulations, and avoid vendor lock-in.

Interoperability is essential. Behavior analytics platforms should support open standards, common data formats, and well-documented interfaces to integrate with existing SIEM, SOAR, identity management, and OT monitoring systems (Strom et al., 2018). This reduces deployment risk and allows organizations to evolve their security posture incrementally rather than through disruptive replacements.

Procurement processes should emphasize transparency, explainability, and lifecycle support. Vendors must be able to demonstrate how their models operate, how updates are governed, and how systems can be audited over time. For national infrastructure operators, long-term maintainability and alignment with national standards often outweigh short-term performance gains.

Ultimately, governance-aware procurement ensures that behavior-based cyber defense architectures enhance resilience at scale, across multiple operators and jurisdictions, without fragmenting security ecosystems or undermining policy coherence.

10. Limitations and Future Research Directions

Despite the promise of behavior-based cyber defense architectures for strengthening the resilience of defense and national critical infrastructure, several limitations remain. These constraints arise from data, modeling, operational safety requirements, and the adaptive nature of adversaries. Addressing these limitations is essential to ensure that such systems remain trustworthy, effective, and aligned with mission and safety objectives.

10.1 Data quality, labeling scarcity, and coverage gaps

Behavior-based cyber defense systems depend fundamentally on the availability of high- quality telemetry across network, endpoint, identity, and operational technology domains. In practice, data collected from defense and national infrastructure environments is often incomplete, noisy, or inconsistent due to legacy systems, segmented networks, and strict access controls (Ahmed et al., 2016; Garcia-Teodoro et al., 2009). These conditions complicate the construction of reliable behavioral baselines and increase the risk of both false positives and missed detections.

A persistent challenge is the scarcity of labeled data representing real-world attacks, particularly for advanced persistent threats and insider misuse. Many high-impact incidents are rare, classified, or insufficiently documented, limiting the feasibility of supervised learning approaches (Sommer C Paxson, 2010). As a result, behavior-based systems frequently rely on unsupervised or semi-supervised models, which may struggle to distinguish malicious deviations from legitimate operational changes.

Coverage gaps further limit detection effectiveness. Telemetry blind spots commonly occur in legacy OT systems, proprietary industrial protocols, and air-gapped or intermittently connected defense networks (Bhamare et al., 2020; Stouffer et al.). These gaps reduce situational awareness and may allow adversaries to operate undetected within less monitored segments of critical infrastructure.

Future research must focus on robust data engineering strategies, improved cross-domain data fusion, and principled approaches for learning under limited or weak supervision. Advances in transfer learning and self-supervised representation learning may help mitigate labeling constraints while improving generalization across heterogeneous environments.

10.2 Model drift, adversarial adaptation, and evasion

Behavior-based detection models are not static. Over time, normal system behavior evolves due to software updates, infrastructure modernization, policy changes, and shifting mission requirements. This phenomenon, commonly referred to as concept drift, can degrade detection performance if models are not continuously monitored and updated (Chandola et al., 2009). In defense and national infrastructure settings, where change management is often slow and highly regulated, timely model retraining presents both technical and governance challenges.

In parallel, sophisticated adversaries actively adapt their tactics to evade behavioral detection. Attackers may intentionally mimic legitimate user behavior, throttle malicious actions to remain below detection thresholds, or exploit blind spots in telemetry collection (Al-Sada et al.; Hutchins et al., 2011). Adversarial machine learning techniques further increase the risk that detection models can be manipulated or poisoned, particularly in environments where data integrity cannot be fully assured (Biggio C Roli, 2018).

These dynamics highlight the need for resilient detection strategies that combine multiple behavioral signals, incorporate adversary-aware modeling, and maintain human oversight. Continuous validation, red teaming, and attack emulation exercises are critical for assessing robustness against adaptive threats and preventing overreliance on automated analytics (Scarfone C Mell, 2007).

10.3 OT safety constraints and limited response options

In operational technology and cyber-physical system environments, the consequences of automated response actions can extend beyond data loss to physical damage, environmental harm, or risks to human safety. As a result, behavior-based cyber defense architectures face strict constraints on permissible response mechanisms in sectors such as energy, transportation, and defense manufacturing (Humayed et al., 2017; Kriaa et al., 2015).

Unlike IT systems, where isolation or shutdown may be acceptable, OT systems often require continuous operation and deterministic control. Automated containment actions such as device quarantine, network segmentation, or process termination may disrupt critical services or violate safety certifications (Stouffer et al.). These limitations

necessitate a cautious, policy-driven approach to response orchestration, with a strong emphasis on human-in-the-loop decision-making.

Future work must explore response strategies that balance cyber risk reduction with operational continuity. This includes the development of safety-aware response policies, simulation-based testing, and closer integration between cybersecurity teams and engineering operators.

10.4 Future directions: federated analytics, digital twins, explainable behavior models, autonomous response guardrails

Several emerging research directions offer promising pathways to address current limitations. Federated analytics and distributed learning approaches can enable collaborative behavior modeling across organizations or infrastructure operators without centralizing sensitive data, supporting privacy and sovereignty requirements in defense contexts (Linkov C Kott, 2019).

Digital twins of critical infrastructure systems represent another important avenue. By creating virtual replicas of physical and cyber assets, defenders can simulate behavioral baselines, test detection models, and evaluate response strategies under realistic conditions without risking live operations (Ding et al., 2018). Such environments also support continuous training and validation against evolving threat scenarios.

Explainable behavior models are essential for trust, governance, and operational adoption. Analysts and decision-makers must understand why a behavior is flagged as anomalous, particularly when actions carry mission or safety implications (Bishop, 2006). Research into interpretable machine learning and transparent risk scoring will remain critical.

Finally, autonomous response capabilities must be bounded by explicit guardrails. Rather than fully autonomous cyber defense, future systems should emphasize constrained automation, where predefined policies, confidence thresholds, and safety constraints govern response actions (Rose et al., 2020). This approach aligns automation with accountability and resilience objectives.

11. Conclusion

11.1 Summary of contributions

This article examined behavior-based cyber defense architectures as a foundational approach for enhancing the resilience of defense and national critical infrastructure systems. By shifting from static, signature-based detection toward continuous behavioral analysis, these architectures provide improved visibility into stealthy, adaptive, and previously unknown threats. The study synthesized insights from anomaly detection, cyber resilience engineering, and operational security to present an integrated architectural and governance perspective.

11.2 Key takeaways for defense and national critical infrastructure resilience

Several key conclusions emerge. First, behavior-based detection significantly strengthens early threat identification, particularly against advanced persistent threats and insider activity that evade traditional controls (Denning, 1987; Sommer C Paxson, 2010). Second, resilience is achieved not solely through detection accuracy, but through the integration of analytics with governance, risk assessment, and controlled response mechanisms. Third, sector-specific constraints, especially in OT and cyber-physical systems, require tailored architectures that prioritize safety and mission continuity alongside security.

Importantly, behavior-based cyber defense should be viewed as a complement to, rather than a replacement for, established controls such as Zero Trust architectures, standards- based governance, and incident response processes (Rose et al., 2020; Joint Task Force, 2020).

11.3 Practical recommendations for phased adoption

For defense organizations and critical infrastructure operators, phased adoption offers a pragmatic path forward. Initial deployments should focus on visibility and behavioral baselining, with analytics used primarily for decision support rather than automated enforcement. Subsequent phases can introduce risk scoring, ATTCCK-aligned correlation, and limited response automation under strict policy controls.

Investment in data quality, analyst training, and cross-disciplinary collaboration between cybersecurity, engineering, and operations teams is essential. Continuous evaluation through exercises and controlled testing should accompany each phase to ensure alignment with mission objectives and evolving threat conditions.

In conclusion, behavior-based cyber defense architectures represent a critical capability for modern defense and national infrastructure protection. When thoughtfully designed and governed, they provide a scalable, adaptive foundation for resilience in an increasingly contested and complex cyber environment.

References

1. Ahmed, M., Mahmood, A. N., C Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19-31.

2. Anderson, J. P. (1980). Computer security threat monitoring and surveillance. Technical Report, James P. Anderson Company.

3. Bhamare, D., Zolanvari, M., Erbad, A., Jain, R., Khan, K., C Meskin, N. (2020). Cybersecurity for industrial control systems: A survey. computers C security, 89, 101677.

4. Biggio, B., C Roli, F. (2018, October). Wild patterns: Ten years after the rise of adversarial machine learning. In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (pp. 2154-2156).

5. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

6. Buczak, A. L., C Guven, E. (2015). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications surveys C tutorials, 18(2), 1153-1176.

7. Caltagirone, S., Pendergast, A., C Betz, C. (2013). The diamond model of intrusion analysis.

8. Chandola, V., Banerjee, A., C Kumar, V. (2009). Anomaly detection: A survey. ACM computing surveys (CSUR), 41(3), 1-58.

9. Chio, C., C Freeman, D. (2018). Machine learning and security: Protecting systems with data and algorithms. " O'Reilly Media, Inc.".

10. Cichonski, P., Millar, T., Grance, T., C Scarfone, K. (2012). Computer security incident handling guide. NIST Special Publication, 800(61), 1-147.

11. Conti, M., Dehghantanha, A., Franke, K., C Watson, S. (2018). Internet of Things security and forensics: Challenges and opportunities. Future Generation Computer Systems, 78, 544-546.

12. Denning, D. E. (1987). An intrusion-detection model. IEEE Transactions on software engineering, (2), 222-232.

13. Ding, D., Han, Q. L., Xiang, Y., Ge, X., C Zhang, X. M. (2018). A survey on security control and attack detection for industrial cyber-physical systems. Neurocomputing, 275, 1674-1683.

14. Dwork, C. (2006). Differential privacy. In Proceedings of the International Colloquium on Automata, Languages and Programming (pp. 1–12). Springer.

15. Force, J. T. (2020). Security and privacy controls for information systems and organizations (No. NIST Special Publication (SP) 800-53 Rev. 5 (Withdrawn)). National Institute of Standards and Technology.

16. Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G., C Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. computers C security, 28(1-2), 18-28.

17. Goodfellow, I. J., Shlens, J., C Szegedy, C. (2014). Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572.

18. Hochreiter, S., C Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

19. Hollnagel, E., Woods, D. D., C Leveson, N. (Eds.). (2006). Resilience engineering: Concepts and precepts. Ashgate Publishing, Ltd..

20. Humayed, A., Lin, J., Li, F., C Luo, B. (2017). Cyber-physical systems security—A survey. IEEE Internet of Things Journal, 4(6), 1802-1831.

21. Hutchins, E. M., Cloppert, M. J., C Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Leading Issues in Information Warfare C Security Research, 1(1), 80.

22. ISO/IEC. (2022). ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection: Information security management systems: Requirements. International Organization for Standardization. https://www.iso.org/obp/ui/en/#iso:std:iso-

iec:27001:ed-3:v1:en

23. ISO/IEC. (2022). ISO/IEC 27002:2022 Information security, cybersecurity and privacy protection: Information security controls. International Organization for Standardization. https://www.iso.org/obp/ui/en/#iso:std:iso-iec:27002:ed-3:v2:en

24. Jajodia, S., Liu, P., Swarup, V., C Wang, C. (2009). Cyber situational awareness. Springer US.

25. Kriaa, S., Pietre-Cambacedes, L., Bouissou, M., C Halgand, Y. (2015). A survey of approaches combining safety and security for industrial control systems. Reliability engineering C system safety, 139, 156-178.

26. Liao, H. J., Lin, C. H. R., Lin, Y. C., C Tung, K. Y. (2013). Intrusion detection system: A comprehensive review. Journal of network and computer applications, 36(1), 16-24.

27. Linkov, I., C Kott, A. (2019). Fundamental concepts of cyber resilience: Introduction and overview. In Cyber resilience of systems and networks (pp. 1-25). Springer, Cham.

28. M Bishop, C. (2006). Pattern recognition and machine learning.

29. Mitchell, R., C Chen, I. R. (2014). A survey of intrusion detection techniques for cyber-physical systems. ACM Computing Surveys (CSUR), 46(4), 1-29.

30. Papernot, N., McDaniel, P., Goodfellow, I., Jha, S., Celik, Z. B., C Swami, A. (2017, April). Practical black-box attacks against machine learning. In Proceedings of the 2017 ACM on Asia conference on computer and communications security (pp. 506- 519).

31. Patcha, A., C Park, J. M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer networks, 51(12), 3448-3470.

32. Peltier, T. R. (2005). Information security risk analysis. Auerbach publications.

33. Rose, S., Borchert, O., Mitchell, S., C Connelly, S. (2020). Zero trust architecture. NIST special publication, 800(207), 1-52.

34. Ross, R., McEvilley, M., C Oren, J. (2016). Systems security engineering:

Considerations for a multidisciplinary approach in the engineering of trustworthy secure systems (No. NIST Special Publication (SP) 800-160 (Withdrawn)). National Institute of Standards and Technology.

35. Sadeghi, A. R., Wachsmann, C., C Waidner, M. (2015, June). Security and privacy challenges in industrial internet of things. In Proceedings of the 52nd annual design automation conference (pp. 1-6).

36. Scarfone, K., C Mell, P. (2007). Guide to intrusion detection and prevention systems (idps). NIST special publication, 800(2007), 94.

37. Shackleford, D. (2015). Who's using cyberthreat intelligence and how. SANS Institute.

38. Sommer, R., C Paxson, V. (2010, May). Outside the closed world: On using machine learning for network

intrusion detection. In 2010 IEEE symposium on security and privacy (pp. 305-316). IEEE.

39. Stoneburner, G., Goguen, A., C Feringa, A. (2002). Risk management guide for information technology systems. Nist special publication, 800(30), 800-30.

40. Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., C Thomas, C. B. (2018). Mitre attCck: Design and philosophy. In Technical report. The MITRE Corporation.

41. Sultana, N., Chilamkurti, N., Peng, W., C Alhadad, R. (2019). Survey on SDN based network intrusion detection system using machine learning approaches. Peer-to- Peer Networking and Applications, 12(2), 493-501.

42. Sweeney, L. (2002). k-anonymity: A model for protecting privacy. International

journal of uncertainty, fuzziness and knowledge-based systems, 10(05), 557-570.

43. Taorui Guan, "Evidence-Based Patent Damages," 28 Journal of Intellectual Property Law (2020), 1-61.

44. Tibshirani, R., C Friedman, J. H. (2001). The elements of statistical learning

[electronic resource]: data mining, inference, and prediction: with 200 full-color illustrations (Vol. 9). Springer.

45. Umer, M. A., Junejo, K. N., Jilani, M. T., C Mathur, A. P. (2022). Machine learning for intrusion detection in industrial control systems: Applications, challenges, and recommendations. International Journal of Critical Infrastructure Protection, 38, 100516.

46. Uppuluri, V. (2019). The Role of Natural Language Processing (NLP) in Business Intelligence (BI) for Clinical Decision Support. ISCSITR-INTERNATIONAL JOURNAL OF BUSINESS INTELLIGENCE (ISCSITR-IJBI), 1(2), 1-21.

47. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... C Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

48. Woods, D. D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. Reliability engineering C system safety, 141, 5-9.