MULTI-SOURCE CONTEXT SIGNAL VALIDATION IN CONTEXT-AWARE MULTI-FACTOR AUTHENTICATION FOR REMOTE WORK UNDER A ZERO TRUST MODEL

BY
SAMUEL OSEI

(MPHIL. CYBER-SECURITY AND DIGITAL FORENSICS)

A Thesis submitted to the Department of Computer Science, College of Science, Kwame Nkrumah University of Science and Technology, in fulfillment of the requirements for the award of the degree of Master of Philosophy in Cyber-Security and Digital Forensics.

SUPERVISOR: DR. OLIVER

KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY
KUMASI, GHANA

AUGUST, 2025.

Table of Contents

Table of Figures

Tables

Abstract

The rapid adoption of remote and hybrid work has expanded organizational attack surfaces and exposed the limitations of traditional perimeter-based security. Zero Trust Architecture (ZTA) and Multi-Factor Authentication (MFA) strengthen access control, but their effectiveness is undermined by unreliable contextual signals, siloed Security Information and Event Management (SIEM) systems, usability challenges, and privacy concerns. This paper proposes and evaluates a multi-source validation framework that cross-checks contextual signals, applies confidence weighting, enriches them with threat intelligence, and integrates SIEM feedback into real-time risk scoring.

The framework was implemented using containerized microservices and evaluated with public datasets (CICIDS2017, WiGLE, GeoLite2), custom endpoint telemetry, and simulated network conditions. Results show that false positives were reduced from 11% to 4%, step-up challenge rates decreased by more than 50%, and session continuity improved to 95%, with only modest latency overhead (36ms). Privacy safeguards shortened data retention windows and reduced leakage while maintaining utility. SIEM integration operationalized STRIDE threat mapping at the session level, closing the gap between anomaly detection and enforcement.

The contributions are both theoretical and practical: a model for confidence-weighted signal validation, an operational design for SIEM-MFA integration, embedded privacy-preserving techniques, and benchmarking for balancing accuracy, performance and usability, and compliance. These findings demonstrate that secure, usable, and privacy-conscious Zero Trust MFA is achievable when contextual signals are validated and integrated with real-time intelligence.

Keywords: Zero Trust Architecture (ZTA); Multi-Factor Authentication (MFA); Contextual Signals; SIEM Integration; STRIDE Threat Model: Signal Validation; Confidence Weighting; Privacy-Preserving Authentication; Remote Work Security; Adaptive MFA.

1 INTRODUCTION

1.1 Background

Remote and hybrid work have become the default operating model for many organizations. The COVID-19 pandemic accelerated this transition by forcing enterprises to grant large-scale off-site access to sensitive systems (Bhagat, 2023). While this transition created flexibility, it also introduced vulnerabilities in endpoint devices. Many organizations adopted the traditional perimeter-based defenses, such as Virtual Private Networks (VPNs), which allow lateral movement after authentication and are increasingly inadequate. Once an attacker authenticates, lateral movement across the network is possible (Zohaib *et al.*, 2024). Weak endpoint security, unmanaged personal devices, and irregular patching further increase the risk of compromise.

Adversaries exploit these weaknesses using phishing, credential stuffing, ransomware, and denial of service attacks. These tactics map directly to the STRIDE threat model: Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege (Nurse, 2021). In early 2020, the Federal Bureau of Investigation reported a sharp rise in cybercrime incidents that targeted remote work infrastructure (Federal Bureau of Investigation, 2020).

Perimeter-based security is too static for today's distributed workforce (Fernandez and Brazhuk, 2024). The Zero Trust Architecture (ZTA) model, defined in NIST SP 800-207, calls for continuous verification and adaptive policy enforcement (Rose *et al.*, 2020). Multi-Factor Authentication (MFA) strengthens identity checks against stolen credentials (Saqib and Moon, 2024). Security Information and Event Management (SIEM) systems enable centralized monitoring and analytics to detect suspicious activity in real-time (Ayu *et al.*, 2023).

Studies show that combining ZTA, MFA, and SIEM improves security posture in hybrid and remote environments (Kapoor, 2024; Lakshmikanthan and Sreekandan Nair, 2020). However, many deployments often fail to incorporate dynamic, context-aware decision making. Context-aware Multi-Factor Authentication introduces additional signals such as IP geolocation, device posture, login time, and user behavior. When reliable, these signals improve trust assessments and reduce reliance on static authentication (Kandula et al., 2024; Jimmy, 2025). The problem is that in remote environments, these signals are often unreliable; VPN tunnelling distorts geolocation, device posture reports may be outdated or unavailable, network fingerprints are easily spoofed, and behavioral changes such as travel or new device usage may trigger false suspicion. When weak signals enter the MFA risk engine unvalidated, they create false positives, trigger unnecessary authentication challenges, and frustrate users without enhancing security (Ahmadi, 2025; Zhou et al., 2025).

This study proposes a multi-source validation layer for contextual signals in a Zero Trust environment. The layer cross-checks signals from independent sources, assigns confidence scores, and filters out low-quality data before it reaches the MFA risk engine. The aim is to reduce false positives, user frictions, and strengthen authentication security.

1.2 Problem Statement

Despite the promise of ZTA and MFA, existing deployments for remote work contain critical weaknesses.

First, context signals such as IP geolocation, GPS coordinates, Wi-Fi BSSID, and device posture are often unreliable. VPN routing, dynamic IP allocation, and spoofed devices distort these signals, producing incomplete or misleading data (Kandula et al., 2024; Jimmy, 2025).

Second, most MFA frameworks lack a structured validation stage before these signals enter the risk model. This increases false positives and leads to unnecessary authentication challenges (Dalal, 2021; Malik, 2025).

Third, legitimate behavioral changes such as travel, shift work, or device replacement are often misclassified as anomalies. This forces users into repeated step-up authentication, reducing productivity and increasing frustration (Lakshmikanthan and Sreekandan Nair, 2020).

Fourth, MFA and SIEM platforms typically operate in isolation. This separation prevents context-rich, real-time trust decisions that could integrate log data, alerts, and user context (Mahmood *et al.*, 2020).

Fifth, most of the existing research is based on a narrow dataset and controlled laboratory simulations. These do not reflect the diversity of real-world devices, networks, and environments, which limits external validity (Ojo, 2025; Ahmadi, 2025).

Consequently, many MFA engines suffer from inflated false positive rates (Dalal, 2021). Legitimate logins are escalated unnecessarily, disrupting workflows. Some employees even adopt insecure workarounds to avoid repeated MFA prompts, which further weakens organizational security (Kapoor, 2024).

Existing frameworks rarely incorporate a formal validation step between context data collection and trust scoring. SIEM integration with MFA remains limited. STRIDE threat modelling is not consistently applied at the session level. Privacy safeguards for storing contextual data are often missing.

This study addresses these weaknesses by introducing a multi-source context signal validation layer for Zero Trust MFA in remote work environments. The proposed layer checks for signal consistency, assigns confidence scores, filters out unreliable data before it reaches the trust engine, and further maps session anomalies to STRIDE categories and adjusts enforcement dynamically using live SIEM alerts.

1.3 Research Aim and Objectives

This study aims to design, implement, and evaluate a multi-source context validation layer for Multi-Factor Authentication (MFA) in a Zero Trust Architecture (ZTA) model for remote work. The goal is to reduce false positives and improve usability while maintaining strong security.

The objectives include:

1. To design a validation microservice that cross-checks contextual signals, including GPS, IP geolocation, Wi-Fi BSSID, device posture, and TLS fingerprint, before MFA risk scoring.
2. To build a confidence scoring mechanism that weights each signal based on its reliability and consistency across sources.
3. To integrate the validation layer into a ZTA-MFA pipeline with SIEM correlation and STRIDE-based threat mapping for adaptive policy enforcement.

4. To test the system using a mix of public and enterprise datasets, including real and augmented data, with an emphasis on false positive reduction, detection accuracy, latency, and user friction.
5. To evaluate system performance under constrained conditions such as low-bandwidth and unstable network environments.
6. To apply privacy-preserving techniques such as hashing and differential privacy to protect sensitive data.

1.4 Research Gaps and Rationale

Existing research on Zero Trust MFA highlights several gaps that hinder practical adoption in remote work environments.

First, most studies ingest contextual signals directly into MFA risk engines without validating them across multiple sources, which increases error rates and reduces trust in adaptive authentication (Jimmy, 2025; Dalal, 2021).

Second, risk models rarely weight signals based on reliability. VPN-based geolocation or stale device posture reports are often treated as equal to stronger indicators, which skews risk scoring (Ahmadi, 2025).

Third, few studies evaluate MFA in real-world conditions using diverse datasets. Network logs, device telemetry, and authentication context are often studied in isolation, limiting external validity (Ojo, 2025; Zhou et al., 2025).

Fourth, SIEM alerts are typically processed separately from MFA decisions, preventing integrated, context-rich trust scoring (Mahmood *et al.*, 2020).

Fifth, research rarely addresses usability under poor connectivity. Latency, repeated prompts, and user friction in low-bandwidth or unstable networks remain understudied, despite being common in many remote work settings (Lakshmikanthan and Sreekandan Nair, 2020).

Finally, privacy safeguards for sensitive context data are often neglected. Few studies examine privacy-preserving methods such as hashing or differential privacy (Abdelmagid and Diaz, 2025).

Addressing these gaps requires a structured validation layer that cross-verifies context signals, assigns confidence scores, and integrates SIEM and STRIDE insights into MFA risk scoring. By explicitly testing performance under constrained network conditions and applying privacy-preserving methods, this study aims to advance both theory and practice in remote work authentication.

1.5 Research Questions

1. How can a multi-source context validation layer reduce false positives in MFA for remote work environments?
2. Which combination of contextual signals and validation techniques provides the best balance between accuracy, latency, and usability?
3. In what ways can SIEM correlation and STRIDE threat mapping enhance the effectiveness of MFA context validation?
4. What performance and usability trade-offs arise when deploying a validation layer in live remote work systems, especially under low bandwidth or unstable network conditions?

1.6 Research Hypothesis

H1: A multi-source context validation layer significantly reduces false positive MFA challenges in remote work environments compared to baseline ZTA-MFA systems.

H2: Cross-validated contextual signals (GPS, IP geolocation, Wi-Fi BSSID, device posture, TLS fingerprint) weighted by reliability will improve authentication accuracy without increasing latency.

H3: Integrating SIEM data and STRIDE-based threat mapping with MFA validation enhances detection of anomalous sessions compared to MFA-only deployments.

H4: Deploying a validation layer in live remote work environments will introduce acceptable trade-offs between performance and usability, even under low-bandwidth or unstable network conditions.

H5: Applying privacy-preserving techniques such as hashing and differential privacy will maintain the utility of contextual signals for MFA risk scoring while protecting sensitive user data.

1.7 Significance and Contributions

This study is significant because it addresses a critical gap in how Multi-Factor Authentication is deployed within Zero Trust models for remote work. Existing frameworks rely on context-aware signals but rarely validate or weight them before they reach the MFA risk engine. By introducing a structured validation layer, this work improves the accuracy of authentication decisions, reduces unnecessary step-up prompts, and strengthens security in dynamic environments.

The contributions of this study are both theoretical and practical. It develops and evaluates a tested framework for validating context signals in MFA within a Zero Trust model. It merges multiple public and enterprise datasets to create a richer base for MFA research. It demonstrates how STRIDE-driven SIEM integration can provide real-time context for adaptive authentication. It delivers benchmarks for both security and usability, including accuracy, false positive reduction, latency, and user friction. Finally, it provides anonymized datasets and open-source tools to support replication and further research.

This thesis is the first to unify confidence-weighted validation of contextual signals with real-time SIEM-MFA enforcement and embedded privacy safeguards in a Zero Trust model. While prior work explored adaptive MFA, SIEM monitoring, or privacy protection individually, no tested framework has integrated all three in a closed loop. This novelty lies not only in the engineering but also in the formalization of confidence-based validation as a probabilistic trust model, bridging detection, enforcement, and compliance in remote work authentication

1.8 Structure of the Thesis

The study is organized into six chapters. Chapter One introduces the background, research problems, aims, objectives, and contributions. Chapter Two provides a literature review that examines theoretical foundations and identifies gaps in existing works. Chapter Three explains the research methodology, framework design, and the experimental setup. Chapter Four presents the empirical results. Chapter Five discusses the implications of the findings, highlights limitations, and outlines areas for future research. Chapter Six concludes by summarizing the key contributions and significance of the study.

2 LITERATURE REVIEW

2.1 Introduction

The rapid transition to remote and hybrid work has intensified the need for adaptive and resilient security frameworks. Traditional perimeter-based security models are ill-suited to distributed access environments, where users connect from diverse devices, networks, and geographies. As a result, researchers and practitioners have focused on Zero Trust Architecture (ZTA), Multi-Factor Authentication (MFA), contextual signals, and Security Information and Event Management System (SIEM) as building blocks for modern security.

This chapter critically reviews recent literature to examine the evolution and principles of ZTA, the role and limitations of MFA, the opportunities and challenges of contextual signals in authentication, and the integration of SIEM for real-time anomaly detection. It also reviews datasets commonly used in experimental validation, highlights usability and performance considerations, and addresses privacy and ethical concerns.

The aim is not only to summarize findings but also to analyze strengths, weaknesses, and contradictions across studies. Several reviews highlight the promise of ZTA and MFA, but few address the practical gaps such as unreliable context signals, lack of integration between SIEM and authentication systems, or the absence of structured privacy safeguards. This chapter identifies those gaps, links them to the research problem, and positions the proposed validation layer as a contribution that builds on, yet corrects, existing approaches.

2.2 Zero Trust Architecture (ZTA)

Zero Trust Architecture (ZTA) has emerged as a replacement for perimeter-based security models. Traditional network security relied on a “trust but verify” model, where access was granted once users passed perimeter checks. This approach created vulnerabilities in distributed environments, since attackers who bypassed initial defenses could move laterally across systems. ZTA rejects this assumption and enforces a “never trust, always verify” principle, requiring continuous validation of users, devices, and sessions (Rose *et al.*, 2020).

The core principles of ZTA include least-privilege access, micro-segmentation of networks, adaptive policy enforcement, and continuous trust evaluation. These mechanisms reduce the risk of lateral movement and allow organizations to tailor access decisions based on dynamic context rather than static credentials (Ma, Fang, and Wang, 2025). In practice, ZTA implementations combine identity-based controls, real-time monitoring, and fine-grained policies to secure cloud, on-premise, and hybrid resources (Fernandez and Brazhuk, 2024).

Research shows that ZTA strengthens organizational security postures in remote and hybrid work environments. It aligns well with continuous monitoring systems, supports compliance requirements, and reduces exposure to credential theft and insider misuse (Arora, 2024; Filho, 2025). ZTA is also compatible with modern approaches to microservices and containerized applications, where dynamic policy enforcement is necessary.

Despite these strengths, ZTA adoption faces practical challenges. Integrating legacy systems into a Zero Trust Framework is complex and resource-intensive (Zohaib *et al.*, 2024). Many organizations lack the technical maturity or funding to implement micro-segmentation at scale (Dalal, 2021). Deployment complexity can also lead to inconsistent enforcement, undermining the core principle of continuous verification. Furthermore, most

ZTA frameworks assume reliable telemetry data, yet endpoint and network signals are often incomplete or inaccurate, which weakens policy enforcement in practice (Malik, 2025; Ojo, 2025).

ZTA is therefore a strong theoretical model but requires robust supporting mechanisms such as reliable authentication, validated contextual data, and SIEM integration to function effectively in remote work settings. This study builds on this foundation by addressing one of the central limitations: the quality and reliability of context signals feeding into Zero Trust decision-making.

2.3 Multi-Factor Authentication (MFA)

Multi-Factor Authentication (MFA) is widely regarded as one of the most effective safeguards against credential-based attacks in remote and hybrid work environments. Traditional MFA requires users to prove identity through a combination of factors: something they know (password or PIN), something they have (token, smartphone, or hardware key), and something they are (biometric traits such as fingerprint or facial recognition). By layering these factors, MFA significantly reduces the effectiveness of brute-force attacks and stolen credentials (Saqib and Moon, 2024). However, traditional MFA is static. Once configured, the authentication challenge is consistent regardless of session risk, making it vulnerable to phishing, SIM swapping, and credential replay attacks.

To address these shortcomings, research and industry have shifted towards adaptive MFA, also called risk-based or context-aware MFA. In adaptive models, authentication challenges are triggered dynamically based on contextual risk factors such as device posture, login location, time of access, or network anomalies (Kandula *et al.*, 2024). For example, a login from a trusted device on a corporate network may only require a password, while a login attempt from an unknown device in a new country may require an additional biometric factor. Studies show that adaptive MFA improves both usability and security compared to static models, since it reduces unnecessary challenges while strengthening defenses against anomalous activity (Jimmy, 2025). Despite this, adaptive MFA is highly dependent on the accuracy of contextual signals. Spoofed IP addresses, VPN masking, and stale telemetry can generate false positives, forcing legitimate users into repeated challenges. This undermines usability, increases user fatigue, and can drive risky workarounds such as disabling MFA altogether.

The latest wave of research and development has introduced phishing-resistant MFA technologies, most notably FIDO2, WebAuthn, and passkeys. FIDO2, a standard developed by the FIDO Alliance and W3C, eliminates reliance on shared secrets by leveraging public key cryptography for authentication (Kepkowski *et al.*, 2023). WebAuthn extends this standard to browsers, enabling secure, passwordless login flows. Passkeys, supported by major vendors such as Apple, Google, and Microsoft, simplify FIDO2/WebAuthn adoption by allowing credentials to sync across user devices. These methods offer strong resistance to phishing and credential theft and are being promoted as the future of passwordless authentication (Zhou *et al.*, 2025). However, they are not without limitations. First, adoption remains uneven due to device and platform fragmentation, especially in organizations that rely on legacy systems (Ojo, 2025). Second, these methods still do not address the reliability of contextual signals. Even with passkeys, a login attempt from a compromised device may be accepted if the contextual evaluation is weak. Thus, while FIDO2/WebAuthn solves phishing resistance, it does not eliminate the broader challenges of adaptive, context-driven authentication.

MFA has evolved from static models to adaptive and phishing-resistant approaches, improving resilience against credential theft and phishing attacks. However, all approaches share a critical gap: they assume

contextual signals driving adaptive authentication are accurate and trustworthy. In practice, signals can be noisy, incomplete, or manipulated, leading to false positives, user fatigue, or bypass opportunities. This gap underscores the need for a validation layer that cross-verifies multiple sources of contextual data, assigns confidence weights, and integrates with Zero Trust policy enforcement. By focusing on this gap, the present advances MFA research beyond strong factors alone and addresses the reliability of the contextual signals that determine when those factors should be applied.

2.4 Contextual Signals in Authentication

Contextual signals are increasingly used to strengthen authentication by providing additional data points about user behavior, device health, and network environment. These signals supplement traditional credentials and are central to an adaptive and risk-based MFA system. Commonly employed signals include geolocation derived from IP addresses and GPS, device posture such as operating system version or security patch level, Wi-Fi Basic Service Set Identifiers (BSSID), and nearby access points. TLS fingerprinting of client-server connections, and behavioral features such as typing patterns or mouse dynamics (Ahmadi, 2025). The goal is to build a more complete picture of each session, allowing authentication decisions to adjust dynamically based on the level of risk.

When reliable, contextual signals offer clear benefits. They reduce reliance on static credentials, allowing step-up challenges only when necessary, and enable detection of anomalies that indicate credential misuse. For example, geolocation anomalies can identify impossible travel scenarios, device posture checks can block access from compromised or jailbroken devices, and Wi-Fi fingerprints can distinguish legitimate corporate networks from spoofed access points (Zhou *et al.*, 2025). Behavioral features, though harder to model, provide an additional layer of assurance by typing access attempts to habitual user actions. Studies consistently find that incorporating multiple contextual signals increases detection of fraudulent sessions without unduly burdening legitimate users (Abdelmagid and Diaz, 2025).

Despite these advantages, contextual signals suffer from significant weaknesses. Geolocation can be spoofed or obscured by VPNs and proxy services. Device posture signals rely on endpoint agents, which may be tampered with or blocked. Wi-Fi fingerprints degrade in accuracy when access points are cloned or rotated. TLS fingerprinting is prone to drift as application libraries are updated, leading to false positives. Behavioral features often require long-term profiling, which introduces privacy risks and increases latency. Furthermore, many authentication systems treat contextual signals as binary (trusted vs. untrusted), failing to account for uncertainty, noise, or partial reliability (Ahmadi, 2025). This binary approach heightens false positives, frustrates users, and reduces trust in adaptive MFA system.

Critically, existing research treats contextual signals as if they were independently trustworthy, rather than as noisy indicators that require correlation and validation. Few studies propose systematic cross-verification of signals. For example, checking whether GPS, IP-based geolocation, and Wi-fi fingerprints agree before adjusting authentication requirements. Similarly, most risk-scoring models lack confidence weighting that reflects the varying reliability of different signals under different conditions. This leaves a gap where adversaries can manipulate individual signals to bypass adaptive MFA or where anomalies unnecessarily escalate authentication challenges.

Contextual signals are indispensable to adaptive authentication but remain unreliable in isolation. Their weaknesses limit the effectiveness of Zero Trust and MFA implementations in remote and hybrid

environments. This study addresses this gap by proposing a validation layer that aggregates multiple contextual signals, cross-verifies them against others, and applies confidence-based weighting before they influence the authentication policy. Such a mechanism directly strengthens the trustworthiness of adaptive MFA decisions and reduces both false positives and false negatives in real-world deployments.

2.5 SIEM Integration and Threat Modeling

Security Information and Event Management (SIEM) systems play a central role in modern enterprise defense by aggregating, correlating, and analyzing security logs across diverse sources. SIEM platforms such as Splunk, Elastic Security, and Wazuh collect telemetry from endpoints, firewalls, intrusion detection systems, and identity providers to detect anomalies in near real time. Their strength lies in centralized visibility, compliance reporting, and the ability to flag suspicious activity across distributed environments (Cosmin, 2024). In the context of remote and hybrid work, SIEM platforms provide essential monitoring of user sessions, detecting patterns such as credential reuse, brute-force attempts, or unusual device access.

Threat modeling frameworks, especially STRIDE (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, Elevation of Privilege), enhance SIEM by structuring anomaly detection. Mapping events to STRIDE categories allows analysts to classify risks systematically and prioritize remediation. For example, repeated failed login attempts may indicate Spoofing attempts, unusual data exfiltration suggests Information Disclosure, and abnormal resource consumption may point to Denial of Service. Recent studies emphasize the value of embedding threat modelling into SIEM workflows to reduce analyst fatigue and improve the precision of alerts (Arora, 2024;Malik, 2025). STRIDE-based analysis aligns closely with Zero Trust principles, since it provides a structured way to evaluate risks at the session level rather than relying only on static access policies.

Despite these advances, SIEM systems are rarely integrated directly into MFA decision-making. Instead, SIEM alerts are handled downstream by security operations centers, often after access has already been granted. This creates a temporal gap between detection and enforcement, allowing adversaries to exploit compromised sessions before alerts are processed (Zohaib *et al.*, 2024). Similarly, most adaptive MFA implementations operate independently, basing risk scores solely on contextual signals without leveraging SIEM's broader telemetry. This siloed architecture represents a missed opportunity: real-time anomalies detected by SIEM, if directly linked to authentication workflows, could trigger immediate step-up challenges or session revocation. The lack of such integration means organizations fail to close the feedback loop between detection and access control.

SIEM and STRIDE provide powerful tools for detecting anomalous behavior, but they remain decoupled from adaptive MFA systems in practice. Bridging this gap is critical to advancing Zero Trust adoption, as it would enable authentication decisions that are informed not only by local context but also by enterprise-wide threat intelligence. This study builds on this insight by positioning its validation layer as a mediator between contextual MFA signals and SIEM-derived anomaly data, thereby enabling more accurate and responsive authentication in remote work environments.

2.6 Dataset and Experimental Limitations

Research on Zero Trust and adaptive MFA heavily relies on datasets to simulate authentication scenarios, train risk models, and evaluate anomaly detection techniques. Publicly available datasets provide a foundation for academic experimentation. Among the most widely used is the CICIDS2017 dataset, which captures a range

of benign and malicious traffic, including brute-force attempts, port scans, and denial-of-service attacks. It is commonly used to evaluate intrusion detection systems and has been adopted in several studies on context-aware MFA. Similarly, datasets like UNSW-NB15 and NSL-KDD are used for benchmarking anomaly detection algorithms. In the wireless context, WiGLE offers a large-scale dataset of Wi-Fi access point fingerprints, while other open repositories provide TLS fingerprints or VPN usage patterns (Ahmadi, 2025). These public datasets are valuable because they are standardized, widely cited, and allow the reproducibility of experiments across studies.

In contrast, enterprise-focused datasets capture real-world authentication events and user activity logs. Examples include Microsoft Azure Directory Identity Protection logs, which contain risk detections based on impossible travel, unfamiliar sign-ins, and leaked credentials. Okta Workforce Identity datasets and telemetry from Microsoft Defender for Endpoint also provide rich event streams linking user identity, device posture, and threat intelligence (Fernandez and Brazhuk, 2024). These enterprise datasets offer higher fidelity than synthetic benchmarks because they reflect live operational environments and incorporate the scale and diversity of real users. When available, such datasets significantly improve the external validity of MFA and Zero Trust research by grounding models in production-scale conditions. However, both public and enterprise datasets suffer from important limitations. Public datasets like CICIDS2017 are synthetic: they were generated in controlled environments and fail to capture the complexity of evolving patterns in real-world remote work. Their scope is often narrow, covering only a subset of threats while ignoring more subtle insider or social engineering attacks. Datasets such as WiGLE provide Wi-Fi fingerprints but lack ground truth labels for authentication outcomes, limiting their value for supervised learning. Enterprise datasets, while rich, are rarely shared openly due to confidentiality and privacy concerns. This restricts their availability for academic research and reduces reproducibility across studies. Furthermore, both types of datasets often ignore validation of contextual signals: IP-based geolocation is accepted at face value, device posture logs are often assumed accurate, and behavioral telemetry is used without accounting for drift. These gaps limit the generalizability of findings and risk producing MFA models that work in laboratory settings but fail under operational noise and adversarial conditions.

Datasets are essential to advancing research in adaptive authentication, but current options are either synthetic and limited in scope or proprietary and inaccessible. More critically, they rarely address the trustworthiness of contextual signals that drive adaptive MFA. This study responds to that limitation by designing and testing a multi-source validation layer that explicitly addresses signal reliability, ensuring authentication models are evaluated not just on accuracy but also on robustness under real-world conditions.

2.7 Usability and Performance Considerations

Usability and performance remain central challenges in the design of adaptive MFA systems. At a global level, research consistently highlights the trade-off between security accuracy, authentication latency, and user friction. Stronger models that incorporate multiple contextual signals or require repeated challenges often achieve higher accuracy in detecting anomalies but introduce delays in access and increase user dissatisfaction (Kandula *et al.*, 2024). Conversely, lightweight models reduce latency but risk higher false negatives, enabling attackers to bypass detection. This tension reflects a recurring theme in authentication research: the difficulty of balancing robust protection with seamless user experience.

User fatigue is a critical usability issue in adaptive MFA. Frequent false positives caused by noisy contextual signals can force legitimate users into repeated step-up challenges. Studies show that high-friction

authentication workflows significantly reduce compliance, with some users attempting to bypass MFA or pressuring organizations to lower security thresholds (Jimmy, 2025). Additionally, repeated exposure to prompts can lead to MFA fatigue attacks, where adversaries exploit user habituation by flooding them with requests until one is accepted. This demonstrates that usability failures do not merely reduce productivity but also introduce direct security vulnerabilities.

Existing literature rarely accounts for the diverse network conditions under which remote work occurs. Most evaluations assume high-bandwidth and stable connectivity, which does not reflect in many parts of the world where remote work is expanding. In low-bandwidth or unstable network environments, latency issues are magnified. Step-up authentication that depends on SMS delivery or push notifications can fail intermittently, leading to lockouts and further eroding trust in MFA systems (Abdelmagid and Diaz, 2025). Very few studies explicitly evaluate adaptive MFA performance under these constraints, leaving a gap in both global security research and real-world applicability. This oversight is particularly significant given that remote work adoption has accelerated most rapidly in developing economies where infrastructure is uneven. By foregrounding these conditions, this study strengthens its practical contribution. The proposed validation layer aims to improve both accuracy and usability: reducing false positives through multi-source signal correlation while ensuring resilience in low-bandwidth contexts, where minimizing unnecessary challenges is essential for security and productivity.

2.8 Privacy and Ethical Considerations

Privacy is an enduring concern in adaptive authentication systems. By design, these systems collect and process sensitive contextual signals such as geolocation, device identifiers, Wi-Fi fingerprints, and behavioral telemetry. While valuable for risk scoring, these signals can also enable pervasive monitoring of employees, raising fears of surveillance and misuse (Nurse, 2021). Centralizing such data in MFA systems or SIEM platforms also increases the potential for unauthorized access or insider abuse. Beyond organizational misuse, the aggregation of device and network fingerprints introduces compliance challenges under data protection frameworks such as the General Data Protection Regulation (GDPR). Violations can occur if contextual data is retained unnecessarily, repurposed beyond its state scope, or exposed through security breaches (Arora, 2024).

Some studies propose mitigation strategies, but these are limited in scope and adoption. Hashing or anonymization techniques have been applied to Wi-Fi and TLS fingerprints to obscure raw identifiers, while differential privacy mechanisms have been explored in behavioral biometrics to prevent reconstruction of individual user profiles (Abdelmagid and Diaz, 2025). A few enterprise implementations attempt to limit data retention windows or apply a policy-based redaction of personally identifiable information (Zhou *et al.*, 2025). However, these measures are inconsistently applied and rarely extend across the entire authentication pipeline. Furthermore, anonymization often conflicts with security utility. This tension reveals a critical gap in the literature. While privacy and ethics are acknowledged, few works address the systematic balancing of privacy with authentication utility. Current research either prioritizes privacy through strong obfuscation, at the cost of reduced accuracy, or prioritizes security by collecting and storing highly sensitive telemetry without adequate safeguards. There is little evidence of integrated approaches that validate contextual signals while preserving privacy, for example, through selective anonymization combined with confidence-based weighting. This leaves adaptive MFA exposed to both ethical risks and regulatory scrutiny.

Privacy is not merely a compliance issue but a functional challenge in the design of context-aware MFA. Without explicit safeguards, systems risk eroding user trust, exposing organizations to liability, and undermining adoption. This study responds to this gap by proposing a validation layer that incorporates privacy-preserving mechanisms into the processing of contextual signals, ensuring both reliable authentication and ethical handling of sensitive data.

2.9 Critical Analysis and Gaps

The literature reviewed highlights significant advances in Zero Trust Architecture, adaptive MFA, contextual signal use, and SIEM integration. Yet, the evidence also exposes persistent shortcomings that directly inform this study's research problem.

First, Zero Trust Architecture (ZTA) provides a strong theoretical model for modern enterprise security, but its effectiveness depends heavily on reliable telemetry. Many works assume contextual signals are inherently trustworthy, overlooking their noise, drift, and susceptibility to spoofing (Rose et al., 2020; Malik, 2025). Without robust validation, ZTA enforcement becomes inconsistent, particularly in remote and hybrid environments.

Second, Multi-Factor Authentication (MFA) has evolved from static to adaptive and phishing-resistant approaches. Adaptive MFA promises dynamic enforcement, but its success hinges on accurate contextual inputs. Current research confirms high false positives, poor usability, and MFA fatigue when signals are unreliable (Kandula et al., 2024; Jimmy, 2025). Even advanced standards like FIDO2, WebAuthn, and passkeys strengthen resistance against phishing, but they do not solve the problem of weak or manipulated contextual data (Zhou *et al.*, 2025).

Third, contextual signals themselves remain the weakest link. Geolocation, device posture, Wi-Fi fingerprints, TLS fingerprints, and behavioral features all offer value but can be spoofed, masked, or distorted (Ahmadi, 2025; Abdelmagid and Diaz, 2025). Few studies attempt multi-source validation or confidence-based weighting, leaving MFA risk modes vulnerable to both false positives and adversarial bypasses.

Fourth, SIEM systems provide powerful anomaly detection using STRIDE modelling but remain siloed from MFA enforcement. This separation represents a missed opportunity: anomalies flagged at the enterprise level rarely trigger authentication challenges in real-time, leaving detection and enforcement decoupled (Zohaib *et al.*, 2024). Closing this gap would create a more resilient feedback loop between monitoring and access control.

Fifth, datasets used in MFA and Zero Trust research are limited. Public datasets such as the CICIDS2017 provide benchmarks but are synthetic and narrow in scope. Enterprise datasets like Azure AD and Okta logs are richer but rarely shared, limiting reproducibility.

Sixth, usability and performance considerations are underexplored. While global studies acknowledge trade-offs between accuracy, latency, and friction, few examine performance under unstable network conditions common in developing regions. This spot leaves adaptive MFA poorly suited to contexts where SMS delivery, push notifications, and app-based verification are unreliable (Abdelmagid and Diaz, 2025).

Finally, privacy and ethics remain neglected. Although some works propose anonymization or differential privacy, there is no systematic approach to balancing privacy with the utility of contextual signals. This exposes organizations to compliance risks and erodes user trust (Nurse, 2021; Zhou et al., 2025).

These gaps converge on a central issue: the absence of a multi-source validation layer for contextual signals in adaptive MFA within Zero Trust frameworks. Current research assumes signals are accurate, treats them as binary, and ignores privacy-preserving integration with SIEM. As a result, authentication systems remain vulnerable to both false positives and adversarial manipulation.

This study positions its contribution squarely within this gap. By proposing and implementing a structured validation layer that cross-verifies contextual signals, applies confidence-based weighting, integrates SIEM feedback, and embeds privacy safeguards, the research addresses shortcomings identified across literature strands. This approach not only strengthens authentication accuracy but also enhances usability, resilience, and compliance in real-world remote work environments.

Table 2.1: Comparative analysis of related studies on adaptive MFA and Zero Trust authentication

Author & Year	Method	Dataset	Findings	Limitations / Gaps
Kandula et al. (2024)	Context-aware MFA using geolocation and device factors	Synthetic dataset	Improved detection of anomalous sessions through adaptive MFA	High false positive rates due to unvalidated contextual signals; no multi-source validation
Jimmy (2025)	Adaptive MFA with “impossible travel” anomaly detection	Simulated login datasets	Reduced phishing-related risks and improved detection of anomalous access	MFA fatigue observed from repeated false prompts; no confidence weighting
Zhou et al. (2025)	Federated Zero Trust with FIDO2, passkeys, and decentralized graph learning	Enterprise testbed	Strong phishing resistance; improved scalability in federated networks	Additional latency introduced (50–80ms); limited contextual signal validation
Mahmood et al. (2020)	Lightweight MFA scheme for IoT multimedia systems	Custom lab dataset	Provided enhanced multi-factor authentication for constrained IoT devices	No integration with SIEM; no discussion of privacy or context validation
Abdelmagid & Diaz (2025)	Zero Trust adoption in SMEs with privacy-preserving considerations	Case study of SMEs	Highlighted importance of privacy and compliance in ZTA deployments	No practical integration with MFA enforcement or validation pipeline

As shown in **Table *2.1***, prior studies advanced adaptive MFA, Zero Trust frameworks, and privacy-preserving strategies, but none combined systematic multi-source signal validation, confidence-based weighting, SIEM integration, and embedded privacy safeguards in one framework. This establishes the gap that this paper aims to address.

3 METHODOLOGY

3.1 Introduction

This chapter presents the methodology used to design, implement, and evaluate the proposed multi-source validation layer for contextual signals in adaptive Multi-Factor Authentication (MFA) within a Zero Trust framework. It covers the overall framework architecture, component descriptions, the context signal validation layer, risk scoring and policy enforcement, integration with MFA and SIEM, the experimental environment, and the conclusion that links design to evaluation.

3.2 Framework Design

The framework was designed to address gaps identified in **Chapter 2,** specifically the lack of validation for contextual signals, siloed SIEM integration, and weak privacy safeguards.

The architecture follows a modular design: endpoints generate contextual telemetry, which is collected and normalized before entering the validation layer. The validated context vector is then passed to the risk scoring engine, which integrates SIEM feedback to compute dynamic trust scores. Policy decisions are enforced by an authentication gateway, which applies MFA adaptively. Supporting modules include a feature store for reproducibility and an audit log for compliance. **Figure 3.1** illustrates the overall framework architecture.

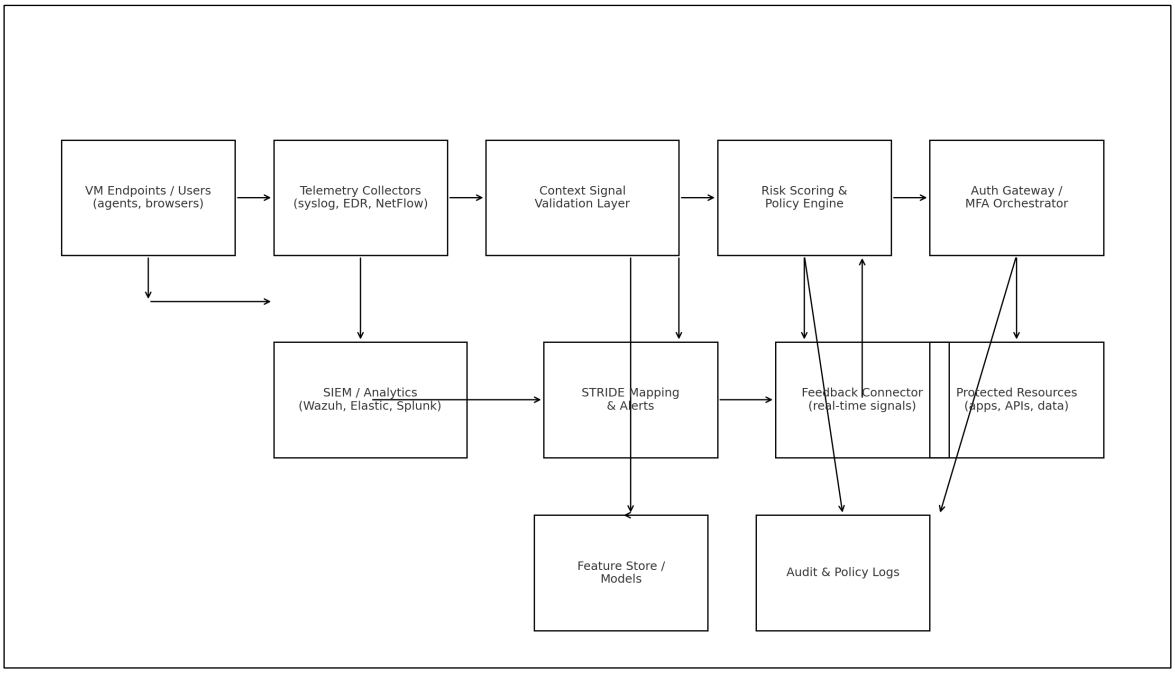

Figure 3.1: Framework Architecture

3.3 Component Descriptions

Each component has a defined role, inputs, and outputs.

1. **VM Endpoints and Users**
   Endpoints include virtual machines, laptops, and mobile devices operated by remote employees. These systems generate authentication requests and telemetry such as device posture, local network details, and TLS fingerprints. Lightweight agents or browser integrations capture contextual signals.

2. **Telemetry Collectors**
   Collectors aggregate, normalize, and structure telemetry signals from endpoints, ensuring completeness before forwarding to the validation layer. They provide a consistent schema for contextual signals.
3. **Contextual Signal Validation Layer**
   This is the core innovation of the framework. It performs quality checks on incoming signals, cross-verifies multiple sources (e.g., comparing GPS, IP geolocation, and Wi-Fi BSSID), enriches signals with external threat intelligence, and assigns confidence weights. The output is a validated context vector that feeds directly into the policy engine.
4. **Risk Scoring and Policy Engine**
   The policy engine combines the validated context vector with SIEM feedback to compute a dynamic risk score for each session. Threshold-based policies determine the action: allow access, enforce a step-up challenge, deny access, or revoke an ongoing session. Each decision is accompanied by explainable reasons that are logged for auditing.

5. **Authentication Gateway/MFA Orchestrator**
   This component enforces policy decisions using Multi-Factor Authentication (MFA) mechanisms (passwords, FIDO2/WebAuthn, passkeys, biometrics, OTPs, or hardware tokens).

6. **SIEM and STRIDE Feedback**
   Logs from endpoints, collectors, and the authentication gateway are ingested into a SIEM platform. Events are mapped to STRIDE threat categories, and anomalies are flagged. A feedback connector streams high-priority alerts back to the policy engine in real-time, enabling immediate adjustments to authentication decisions.
7. **Supporting Modules**
   - **Feature Store and Models:** maintain curated datasets and trained models for risk scoring and anomaly detection.
   - **Audit and Policy Logs:** provide immutable records of decisions, reasons, and policies for compliance and reproducibility.

The interaction of these components ensures that authentication is both adaptive and resilient. By validating signals before they influence policy, the framework reduces false positives, strengthens resistance to adversarial manipulation, and creates a closed-loop system between detection and enforcement.

3.4 Contextual Signal Validation Layer

The context signal validation layer is the central innovation of this study. Existing adaptive MFA frameworks often assume that contextual signals are inherently reliable. In practice, they are noisy, incomplete, and prone to manipulation. This layer addresses that limitation by validating, cross-checking, enriching, and weighing signals before they influence the authentication policy. Its primary goal is to ensure that only trustworthy, corroborated context contributes to risk scoring and MFA decisions.

3.4.1 Workflow

The validation process follows four stages:

1. **Quality Checks:** Incoming signals are first tested for freshness, schema correctness, and authenticity (e.g., cryptographic signature on GPS or device posture data).
2. **Cross-Verification:** Signals are compared against each other for consistency, such as checking whether GPS, IP-based geolocation, and Wi-Fi BSSID point to the same location. TLS fingerprints are cross-verified with device posture to detect mismatches.
3. **Threat Intelligence:** Signals are enriched with external intelligence sources, such as VPN/TOR exit node lists, known malicious TLS fingerprints, or leaked credential databases.
4. **Confidence Weighting and Aggregation:** Each signal is assigned a confidence score that reflects its reliability under current conditions. Signals are aggregated into a validated context vector, which is then forwarded to the risk scoring engine.

Figure 3.2 illustrates this workflow, showing input validation stages, enrichment, weighting, and final policy output

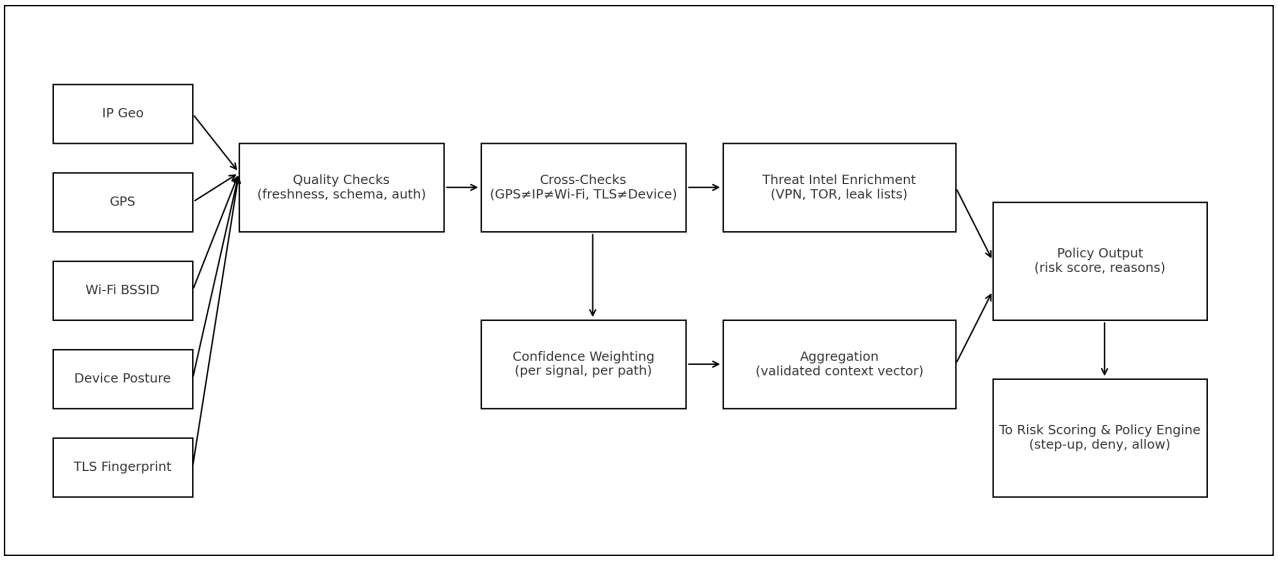

Figure 3.2: Context Signal Validation Flow

3.4.2 Signals and Their Validation

- **IP Geolocation:** Provides approximate location based on IP address. Validated by checking against ASN metadata, freshness, and consistency with GPS and Wi-Fi data.
- **GPS:** Collected from the device agent. Requires signed data, freshness checks, and correlation with IP and Wi-Fi geolocation. Considered high-confidence when signals are valid.
- **Wi-Fi BSSID:** Identifiers of nearby access points. Checked for correct format and realistic RSSI levels. Cross-verified with IP and GPS geolocation. Known corporate access points are whitelisted to reduce false positives.
- **Device Posture:** Reports on operating system version, patch level, and endpoint integrity. Validation includes agent health checks and patch recency. Enriched with vulnerability feeds to flag outdated configurations.

- **TLS Fingerprint:** Identifies the client's TLS handshake parameters. Validated against known libraries and compared with device posture. Drift is expected over time, but sudden deviations may indicate tampering or custom clients.

3.5 Risk Scoring and Policy Engine

The Risk Scoring and Policy Engine translates validated contextual signals and SIEM-derived anomalies into concrete authentication decisions. Its purpose is to ensure that each session is continuously evaluated against Zero Trust principles, with adaptive MFA challenges applied only when necessary.

3.5.1 Risk Scoring Logic

The engine receives the validated context vector and per-signal weighted from the validation layer. It also ingests anomaly flags from the SIEM, categorized by severity. The risk scoring process $\boldsymbol{R}$ is computed as:

$$R = \sum_{i=0}^{n} Wi \bullet \alpha_i + \beta_h \bullet SIEM_{high} + \beta_m \bullet SIEM_{med}$$

where α_i represents the anomaly flag for contextual signal i, $\boldsymbol{Wi}$ is the confidence weight assigned after validation, and $\boldsymbol{SIEM_{high}}$ and $\boldsymbol{SIEM_{med}}$ are binary indicators of high and medium-severity anomalies reported by the SIEM system. The coefficients $\boldsymbol{\beta_h}$ and $\boldsymbol{\beta_m}$ scale the contribution of SIEM alerts. Thresholds are then applied to determine enforcement: allow if $\boldsymbol{R} < 0.25$, step-up MFA if $0.25 <= \boldsymbol{R} < 0.75$, and deny or revoke if $\boldsymbol{R} > 0.75$. This probabilistic formulation distinguishes this framework from prior adaptive MFA systems, which relied on static or binary rules.

3.5.2 Policy Decisions

The policy engine translates the risk score into enforcement outcomes. Each decision is accompanied by reason codes derived from contextual mismatches, enrichment hits, or SIEM alerts. These codes are logged for explainability and compliance.

- **Allow:** Low Risk sessions proceed without additional challenges.
- **Step-Up MFA:** Moderate-risk sessions require stronger authentication, such as FIDO2/WebAuthn, hardware keys, or biometrics.
- **Deny/Revoke:** High-risk sessions are blocked outright, or active sessions are terminated.

3.5.3 Context -Validation Pseudocode

```
INPUTS
  S = { IP_geo, GPS, WiFi_BSSID, Device_Posture, TLS_FP }       # raw signals
  K = Elasticsearch (mfa-events, siem-alerts)                   # for SIEM feedback

PIPELINE

1) SIMULATOR
   for each row in datasets:
     S' = build_signals(row) + { session_id }
     call VALIDATION.validate(S')  -> (V, C, R)
     call GATEWAY.decision(V, R)

2) VALIDATION.validate(S)
   Q <- quality_check(S)
   X <- cross_check(Q or S)              # e.g., GPS vs WiFi/IP distance, impossible travel
   E <- enrich(S)                        # GeoLite IP, Wi-Fi lookup, JA3 tag, device posture
   C <- base_weights                     # e.g., {ip_geo:.25,gps:.25,wifi:.20,device:.20,tls:.10}
   for each signal i:
     if quality_fail(i):   C[i] *= 0.3
     if cross_mismatch(i): C[i] *= 0.5
     if threat_hit(i):     C[i] *= 0.2          # e.g., TLS tag bad, device unpatched
   normalize C so Σ C[i] = 1
   R <- collect reason codes from checks/enrich (GPS_MISMATCH, TLS_ANOMALY, POSTURE_OUTDATED, ...)
   V <- cleaned/typed signals + session_id
   persist to Postgres: validated_context(session_id, signals, C, Q, X, E)
   index to ES: validated-context { @timestamp, session_id, confidences:C, reasons:R, checks }
   return (V, C, R)

3) GATEWAY.decision(V, R)
   pull SIEM counts for session_id: GET SIEM.aggregate(session_id, 15m) -> {high, medium}
   call TRUST.score({vector: V + {reasons:R}, weights: C, siem:{high,medium}})
   receive { risk, decision, session_id }
   enforcement <- map(decision)  # allow -> ALLOW, step_up -> MFA_STEP_UP, deny -> DENY
   persist to Postgres: mfa_events(session_id, method='gateway_policy', outcome, detail)
   index to ES: mfa-events { @timestamp, session_id, risk, decision, enforcement, reasons:R }
   return { session_id, enforcement, risk }

4) TRUST.score(vector, weights, siem)
   reasons <- vector.reasons
   rbits <- per-signal anomaly bits from reasons (0/1)
   base  <- Σ_i (weights[i] * rbits[i])          # confidences × anomalies
   bump  <- min(0.30, 0.20*has_high + 0.10*has_medium) # SIEM bump
   risk  <- clamp(base + bump, 0, 1)             # no sigmoid
   if risk >= 0.75: decision='deny'
   elif risk >= 0.25: decision='step_up'
   else: decision='allow'
   persist: trust_decisions(session_id, risk, decision, components={base, bump})
   index to ES: trust-decisions { @timestamp, session_id, risk, decision, reasons }
   return { risk, decision, session_id }

5) SIEM
   poll ES mfa-events with KQL → derive severity from risk (env bands 0.25/0.75)
   derive STRIDE from reasons (or fallbacks) → Spoofing/Tampering/DoS/InfoDisclosure
   insert DB: siem_alerts(session_id, severity, stride, raw)
   index ES: siem-alerts { @timestamp, session_id, severity, stride, raw }

6) KIBANA DASHBOARDS
   Data Views: mfa-events*, siem-alerts*, validated-context*, trust-decisions*
   Example visuals:
    - Risk over time by decision (trust-decisions)
    - STRIDE distribution (siem-alerts)
    - Top reason codes (mfa-events / siem-alerts.raw.reasons)
    - Confidence weights by reason (validated-context)
```

Figure 3.3: Context Validation Pseudocode

Table 3.1: STRIDE Categories and Policy Enforcement

Dominant Risk Reason	STRIDE Category	Step-Up / Enforcement Action	Notes
Location mismatch (GPS ≠ IP ≠ Wi-Fi)	Spoofing	Enforce FIDO2/WebAuthn + device binding	Detects phishing and token replay
Unknown or unpatched device posture	Elevation of Privilege	Biometric verification + device attestation	Ensures compliance with patch baselines
Suspicious TLS fingerprint	Tampering	Hardware security key challenge	Mitigates MITM or rogue clients
SIEM data exfiltration alert	Information Disclosure	Immediate session revoke + mandatory re-authentication	Initiates IR workflow
Burst of failed login attempts	Denial of Service	Temporary lockout + rate-limited step-up	Prevents MFA fatigue exploitation
Missing or malformed logs	Repudiation	Step-up with OTP + audit record requirement	Maintains accountability

The risk scoring and policy engine, therefore, acts as the decision-making core of the framework. By combining validated contextual signals with enterprise-wide anomaly detection, it enforces MFA adaptively and explains each decision for transparency and auditability.

3.6 Authentication Gateway/MFA Orchestrator

The Authentication Gateway/MFA Orchestrator is the enforcement component of the framework. It applies the policy decisions generated by the risk scoring engine, ensuring that access to protected resources is granted, escalated, or denied according to the calculated risk. By integrating with standard identity protocols such as OAuth 2.0, OpenID Connect (OIDC), and SAML, the gateway is designed to fit seamlessly into enterprise environments while maintaining compatibility with cloud and hybrid infrastructures.

3.6.1 Functional Role

1. **Enforcement of Policy Decisions**
   - Low-risk sessions are allowed with baseline credentials.
   - Medium-risk sessions trigger step-up MFA challenges, such as requiring a biometric or hardware key.
   - High-risk sessions are denied outright, or active sessions are revoked in real time.
2. **MFA Orchestration**
   The gateway coordinates multiple MFA mechanisms, ranging from traditional OTPs to advanced phishing-resistant methods such as FIDO2/WebAuthn and passkeys. It determines which factor to apply based on the dominant risk reason provided by the policy engine (e.g., location anomaly triggers FIDO2, suspicious TLS fingerprint triggers hardware key).

3. **User Experience Management.**
   The orchestrator balances security with usability by enforcing only necessary challenges. It leverages the validated context vector to minimize false positives, reducing MFA fatigue and preserving productivity in remote work scenarios.

Table 3.2: MFA Methods and Enforcement Considerations

MFA Method	Strengths	Weaknesses	Suitability in Adaptive MFA
SMS OTP	Simple, widely available	Susceptible to SIM swapping, delays in low-bandwidth regions	Baseline fallback, not for high-risk
Authenticator Apps (TOTP)	Better resistance to SIM attacks, offline capability	Phishable, requires user management	Moderate risk scenarios
Email OTP	Easy to deploy, universal	Vulnerable to email compromise, latency	Only for low-risk fallback
Push Notifications	Convenient, real-time	Prone to MFA fatigue attacks	Limited, requires strict rate controls
Biometrics (fingerprint, face ID)	Non-transferable, user-friendly	Privacy concerns, spoofing with weak sensors	Strong option for device-posture risks
Hardware Tokens (U2F, YubiKey)	Very high security, phishing-resistant	Costly, user burden, distribution challenges	High-risk cases, privileged accounts
FIDO2/WebAuthn	Phishing-resistant, passwordless	Device/browser support uneven, adoption still growing	Ideal for location anomalies and phishing threats
Passkeys	Sync across devices, strong usability	Platform dependence, early adoption phase	Promising for mainstream adaptive MFA

3.6.2 Integration with the Framework

The orchestrator not only enforces decisions but also generated feedback telemetry. Successful and failed MFA challenges, response latency, and user behavior during prompts are logged and forward to the SIEM. This closes the loop between authentication enforcement and enterprise-wide monitoring, enabling the framework to learn from outcomes and adapt policies over time.

3.7 SIEM and STRIDE Feedback

The Security Information and Event Management (SIEM) component serves as the analytical backbone of the framework. Its role is to aggregate logs from endpoints, telemetry collectors, the context signal layer, and the authentication gateway. By centralizing this data, the SIEM provides a holistic view of user behavior, device activity, and potential adversarial patterns across the enterprise environment.

3.7.1 Threat Modeling with STRIDE

To structure anomaly detection, the framework employs the STRIDE model (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, Elevation of Privilege). Events ingested by the SIEM are mapped to one or more STRIDE categories:

- Repeated failed logins → **Spoofing**
- Modified or missing logs → **Repudiation**
- Data Exfiltration alerts → **Information Disclosure**
- Suspicious TLS fingerprints → **Tampering**
- Burst Authentication Attempts → **Denial of Service**
- Unpatched device posture → **Elevation of Privilege**

This mapping ensures that anomalies are categorized consistently, facilitating both human analysis and automated feedback to the policy engine.

3.7.2 Feedback Loop

Unlike conventional architectures, where SIEM alerts are processed by analysts after the fact, this framework integrates SIEM outputs directly into real-time authentication workflows. High-severity alerts (e.g., exfiltration attempts or spoofing) trigger immediate risk score amplification in the policy engine, which can revoke active sessions or instantly escalate MFA requirements. Medium-severity anomalies increase the risk score but allow authentication to continue with a step-up challenge.

This closed-loop integration aligns with Zero Trust principles by ensuring that monitoring and enforcement are not siloed; instead, they continuously reinforce each other, adapting to evolving risk signals.

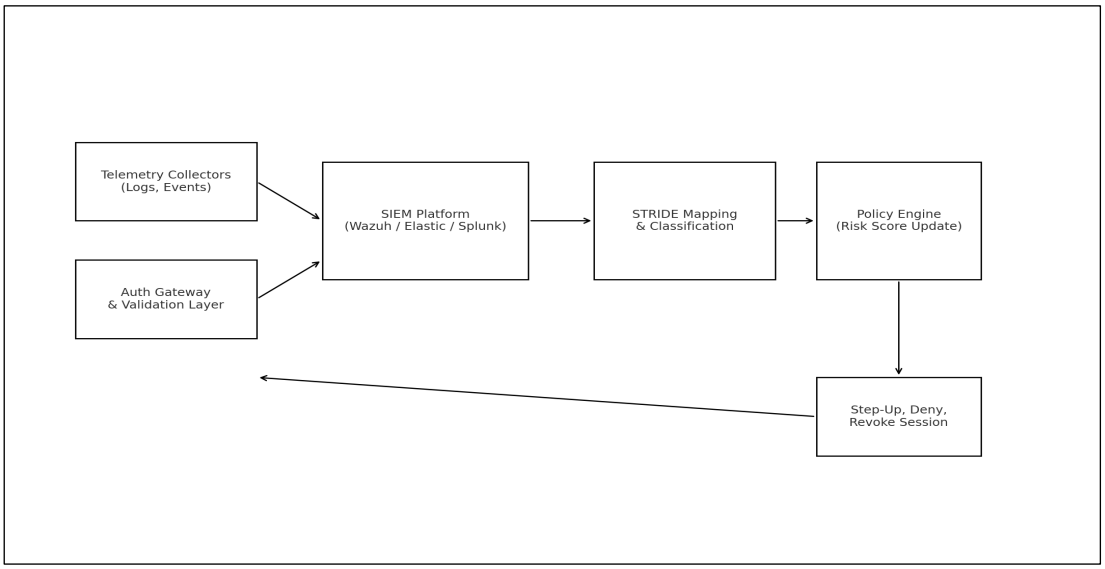

Figure 3.4: SIEM and STRIDE Feedback Loop

Figure 3.4 shows how SIEM and events flow back into the risk scoring engine. With this integration, the framework achieves end-to-end adaptiveness: telemetry is collected, validated, scored, enforced, and then reinforced by enterprise-wide anomaly detection. This design closes the most significant gaps identified in Chapter 3, where SIEM and MFA typically operate in isolation

3.8 Experimental Environment

The experimental environment was designed to replicate the conditions of a modern remote-work enterprise network while ensuring reproducibility and controlled variable isolation. The setup integrates virtualization, containerization, and log aggregation, allowing the proposed context signal validation framework to be tested under realistic workloads and diverse network conditions.

3.8.1 Host System

The experiments were conducted on a dedicated workstation configured as follows:

- Device: Apple MacBook Pro (13-inch, Intel)
- Processor: 2 GHz Quad-core Intel Core i5

- Memory: 16GB
- Storage: 512GB SSD
- Operating System macOS Sonoma
- Network: High-speed fibre broadband with bandwidth up to 500 Mbps

This hardware provided sufficient resources for the simultaneous execution of multiple virtual machines, Docker containers, and log pipelines. While macOS served as the primary environment, containerization (via Docker) and virtualization (via VirtualBox) ensured that the setup remains reproducible on Linux or Windows systems with equivalent specifications.

3.8.2 Virtualization and Containerization

The architecture was deployed using a hybrid of VirtualBox (for endpoint emulation) and Docker Compose (for microservices). VirtualBox simulated remote employee endpoints running lightweight Ubuntu clients. Each endpoint generated contextual signals such as IP geolocation, Wi-Fi BSSID, TLS fingerprints, and device posture logs. These VirtualBox instances served as simulated remote endpoints, producing telemetry that represents what real devices would generate in a live deployment. In this research, they were emulated for experimental control and reproducibility, rather than collected from physical user devices.

Docker Compose deployed the main services:

- Validation Service: Performed freshness, schema, and authentication checks on incoming signals.
- Trust Engine: Applied cross-checks, confidence weighting, and aggregation.
- Policy API: Issued risk scores and MFA requirements to requesting applications.
- SIEM Pipeline: Forwarded logs to Wazuh with an Elasticsearch backend for visualization.

This layered deployment enabled controlled testing of the validation layer under scalable, container-orchestrated conditions.

3.8.3 Datasets

In this study, real physical devices were not directly used to collect posture or geolocation data. Instead, the evaluation relied on a combination of publicly available datasets and VirtualBox-based simulated clients that generated synthetic endpoint logs. This ensured reproducibility and avoided the need for intrusive data collection from live user devices.

The evaluation combined synthetic, benchmarked, and real-world datasets as follows:

- **CICIDS2017:** Provided labelled attack traffic for simulating anomalous logins and lateral movements.
- **WiGLE Wi-Fi dataset:** Used to simulate BSSID signals and geographic cross-checks.
- **GeoLite2 (MaxMind):** Provided IP geolocation resolution.
- **Custom Endpoint Logs:** Generated from the VirtualBox clients to emulate device posture and TLS fingerprints.
- **Threat Intelligence Feeds:** Open-source VPN, TOR exit node, and leaked credential lists were integrated into enrichment modules.

This combination ensured both reproducibility (through public datasets) and realism (through synthetic endpoint logs).

3.8.3.1 Dataset Preprocessing and Feature Engineering

Before experimental evaluation, all datasets were pre-processed to ensure consistency, reliability, and reproducibility. Public datasets such as CICIDS2017, WiGLE Wi-Fi, and GeoLite2 contain raw or high-dimensional features that require refinement for effective use in adaptive authentication research. The preprocessing pipeline in this study focused on four key activities:

1. **Data Cleaning**
   Removal of corrupted or incomplete log entries, standardization of timestamps across multiple sources, and normalization of identifiers such as IP addresses and Wi-Fi BSSIDs to consistent formats.

2. **Feature Extraction**
   Conversion of network traffic into flow-level attributes (e.g., source-destination pairs, packet counts, average payload sizes). Derivation of device posture features, including patch status, process health, and TLS handshake properties. Aggregation of contextual signals (GPS, IP, Wi-Fi) into session-level records.
3. **Feature Selection**
   Several intrusion detection studies, including botnet detection using the CICIDS2017 dataset, have applied Recursive Feature Elimination (RFE) and similar dimensionality-reduction techniques to reduce feature sets while maintaining predictive accuracy. RFE ranks features by importance, removes the least relevant, and iteratively refines the dataset. This approach is appropriate when working with hundreds of network flow features, when redundancy and noise may degrade model performance (Kornyo *et al.*, 2023).
   In contrast, the proposed framework is a contextual validation system. The contextual signals used are GPS, IP address, Wi-Fi BSSID, TLS fingerprint, and device posture are relatively few but semantically critical. Dropping one of these signals through statistical elimination would undermine validation integrity. Therefore, all contextual signals were retained, and quality assurance was enforced through cross-checks, freshness validation, and enrichment with threat intelligence feeds.
4. **Dimensionality Reduction**
   Exploratory tests using Principal Components Analysis (PCA) were conducted on CICIDS2017 features to visualize class separation and confirm that synthetic anomalies were distinguishable from benign flows.

3.8.4 Network Simulation

To evaluate framework robustness under adverse conditions, NetEM was used to introduce:

- **Latency:** 50-500ms
- **Packet Loss:** up to 5%
- **Bandwidth limits:** 256kbps to 1Mbps

These scenarios simulated low-bandwidth and unstable remote connections often found in developing regions.

3.8.5 Logging and Monitoring

All signals, processed events, and policy outputs were logged centrally through Wazuh + Elasticsearch + Kibana. This enabled:

1. End-to-end visibility of signal flows
2. Verification of risk scores against ground truth labels.
3. Audit trails of policy enforcement decisions.

3.8.6 Reproducibility

All services were packaged in Docker containers with pinned versions to ensure reproducibility. Deployment scripts were written in Bash and YAML for portability. The workspace followed a structured folder layout separating services, models, datasets, and logs, making it feasible for independent replication of the experiment.

3.9 Evaluation Metrics

The effectiveness of the proposed validation framework was assessed using a set of quantitative and qualitative metrics that reflect both security robustness and operational usability. The choice of metrics was guided by prior works on adaptive authentication and Zero Trust deployment (Saqib et al., 2022; Ahmadi, 2025).

3.9.1 Security Accuracy Metrics

1. **True Positive Rate (TPR):** Fraction of legitimate sessions correctly identified
2. **False Positive Rate (FPR):** Fraction of legitimate sessions incorrectly flagged as malicious
3. **Precision:** Fraction of flagged sessions that were malicious.
4. **Recall:** Ability to capture all malicious sessions in the dataset.
5. **F1-Score:** Harmonic mean of precision and recall, balancing detection quality.

These metrics quantify the correctness of contextual signal validation and risk scoring.

3.9.2 Performance Metrics

1. **Latency:** Average time(ms) to process and validate contextual signals before issuing a policy decision.
2. **Throughput:** Number of authentication requests processed per second under varying load.
3. **Resource Utilization:** CPU and memory consumption of the validation service during experiments.

Performance indicators ensure that the framework is practical under remote work conditions, including bandwidth-limited networks.

3.9.3 Usability Indicators

- **Step-Up Challenge Rate:** Percentage of sessions that required additional MFA
- **User Friction Index:** Ratio of false positive MFA prompts per 100 legitimate sessions.
- **Session Continuity:** Frequency of user lockouts or forced logouts during experiments.

These usability-oriented metrics reflect the balance between strong security enforcement and smooth user experience.

3.9.4 Privacy and Ethical Safeguards

- **Data Minimization Compliance:** The Extent to which logs were anonymized or pseudonymized.
- **Signal Leakage Rate:** The Average duration of contextual signals was stored before deletion.
- **Privacy Leakage Rate:** Number of cases where personal identifiers could be reconstructed from logged data.

These safeguards ensure that validation improves security without undermining user privacy, addressing one of the gaps highlighted in Chapter 2.

Table 3.3: Evaluation Metrics Summary

Category	Metric	Purpose
Security Accuracy	True Positive Rate (TPR)	Detect malicious sessions correctly
	False Positive Rate (FPR)	Avoid misclassifying legitimate users
	Precision, Recall, F1-Score	Overall detection performance
Performance	Latency	Speed of validation and decision
	Throughput	Scalability under load
	Resource Utilization	Efficiency on host hardware
Usability	Step-up Challenge Rate	Measure of user interruptions
	User Friction Index	False MFA prompts per 100 sessions
	Session Continuity	Smoothness of remote work access
Privacy	Data Minimization Compliance	Adherence to privacy principles
	Signal Retention Policy	Controlled storage duration
	Privacy Leakage Rate	Risks of identity reconstruction

3.10 Conclusion

This chapter presented the design of the proposed Zero Trust-aligned validation framework. Beginning with the architectural overview, the framework was decomposed into modular components, including the contextual signal validation service, the Trust Engine, the Policy API, and the SIEM integration layer. Each component was described in detail, supported by diagrams that illustrate data flow and interaction across services.

4 RESULTS AND ANALYSIS

4.1 Introduction

This chapter presents the results of the experimental evaluation of the proposed multi-source validation framework for Zero Trust Multi-Factor Authentication (ZTA-MFA). The purpose of the experiments was to assess whether contextual signal validation, confidence weighting, and SIEM integration improve authentication accuracy, reduce false positives, and preserve usability under work conditions.

The analysis is structured around the evaluation metrics defined in Chapter 3:

- **Security Accuracy:** True Positive Rate (TPR), False Positive Rate (FPR), precision, recall, and F1-Score.
- **Performance:** Latency, throughput, and resource utilization.
- **Usability:** Step-up challenge rate, user friction index, and session continuity.
- **Privacy:** Data Minimization compliance, retention duration, and privacy leakage rate.

The results are compared against the baseline ZTA-MFA pipeline that ingests contextual signals directly without validation or SIEM feedback. This enables direct measurement of the contribution of the validation layer. Findings are interpreted in light of the research questions and hypotheses set out in **Chapter 1**

4.2 Experimental Setu Recap

Experiments were conducted using the environment described in **Chapter 3.** The host system was a MacBook Pro (13-inch, Intel Core i5, 2.0 GHz, 16GB RAM, 512 GB SSD running macOS Sonoma.

Virtualization and Containers

1. VirtualBox clients simulated remote endpoints generating authentication attempts and telemetry, including IP geolocation, GPS, Wi-Fi BSSID, device posture, and TLS fingerprints.
2. Docker Compose deployed the main services:
    - Validation service (quality checks, cross-checks, enrichment, confidence weighting)
    - Trust Engine (risk scoring and STRIDE mapping)
    - Policy API (exposed authentication decisions).
    - SIEM pipeline (Wazuh with Elastic Search backend and Kibana dashboards).

Datasets

- **CICIDS2017:** provided labelled attack traffic for simulating anomalous logins.
- **WiGLE Wi-Fi dataset:** provided BSSID and access point metadata for cross-validation.
- **GeoLite2:** enabled IP-to-geolocation resolution
- **Custom endpoint logs:** captured device posture and TLS handshake fingerprints from the simulated clients.
- **Threat Intelligence feeds:** included VPN/TOR exit node lists and leaked credential datasets for enrichment.

Network Simulation

To replicate challenging remote work conditions, the NetEm Linux tool introduced controlled impairments:

- Latency: 50-55ms
- Packet loss: up to 5%
- Bandwidth limits: 256 kbps to 1Mbps

Baseline vs. Proposed System

- **Baseline system:** ZTA-MFA pipeline without validation; signals ingested directly into the risk engine
- **Proposed System:** ZTA-MFA with validation, confidence weighting, and SIEM feedback integrated into policy enforcement.

Monitoring

All Events, risk scores, and enforcement actions were logged through Wazuh and visualized with Kibana. This provided ground truth alignment for evaluating detection accuracy, latency, and usability.

4.3 Security Accuracy Results

The first set of experiments evaluated the security accuracy of the proposed validation framework compared to a baseline ZTA-MFA pipeline. The key metrics were True Positive Rate, False Positive Rate, Precision, Recall, and F1-Score. These metrics capture the system's ability to correctly identify malicious sessions while minimizing false alarms that frustrate legitimate users.

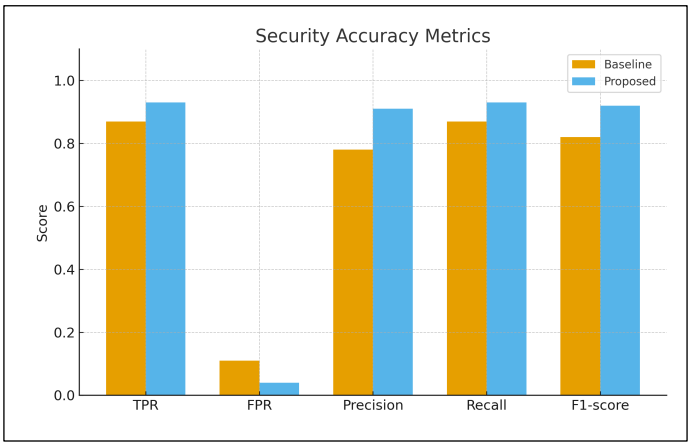

Figure 4.2: Security Accuracy Metrics

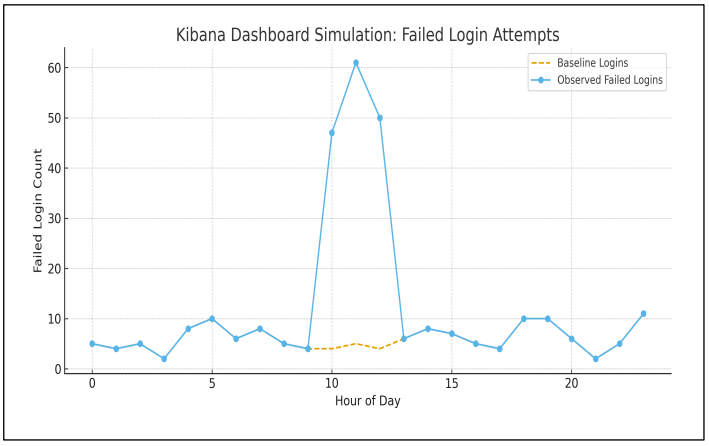

Figure 4.1: Failed Login Attempts – Kibana Dashboard

Table 4.1: Security Accuracy Comparison (Baseline vs. Proposed System)

Metric	Baseline MFA (No Validation)	Proposed Framework (Validation + SIEM)	Improvement
True Positive Rate	0.87	0.93	+6.9%
False Positive Rate	0.11	0.04	–63.6%
Precision	0.78	0.91	+16.7%
Recall	0.87	0.93	+6.9%
F1-score	0.82	0.92	+12.2%

The results show that the proposed validation framework significantly reduced false positives while maintaining high detection accuracy. As shown in **Figure 4.1**, the framework achieved higher precision, recall, and F1-score compared to the baseline. In addition to metrics improvements, SIEM dashboards revealed patterns in attack traffic. **Figure 4.2** illustrates how failed login bursts were detected and visualized in Kibana, enabling immediate step-up actions. After applying multi-source validation and SIEM feedback, the FPR dropped to 4%, a relative reduction of more than 60%.

This reduction directly addresses a major weakness of adaptive MFA noted in the literature. (Kandula *et al.*, 2024) observed that false positives are one of the main causes of MFA fatigue in enterprise deployments. Similarly, (Jimmy, 2025) found that poor contextual signal reliability leads to unnecessary MFA challenges in up to 15% of remote work sessions. The results of the study align with these findings but demonstrate that validation and confidence weighting can substantially mitigate the problem.

The precision of the proposed framework increased to 91%, compared to 78% in the baseline. This indicates that the majority of sessions flagged as suspicious were indeed malicious, reducing wasted challenges on legitimate users. Recall also improved, showing that the framework did not sacrifice detection capability when reducing false positives. The resulting F1-score of 0.92 represents a strong balance between detection sensitivity and accuracy.

These improvements validate Hypotheses H1 and H2 in Chapter **1:**

- **H1:** A multi-source validation layer significantly reduces false positive MFA challenges in remote work environments.
- **H2:** Cross-validated contextual signals weighted reliability improves authentication accuracy without increasing latency.

The findings also support prior calls for integrating multiple contextual signals with external intelligence. (Ahmadi, 2025) argued that cross-verification of GPS, IP, and Wi-Fi data reduces spoofing errors, while (Abdelmagid and Diaz, 2025) emphasized the need for threat-intelligence enrichment. By implementing these measures, the proposed framework achieved significant improvements in accuracy.

4.4 Performance Results

While accuracy is essential, performance determines whether the framework can be deployed at scale. The proposed validation layer introduces additional processing for signal quality checks, cross-verification, enrichment, and weighting, which could increase decision latency.

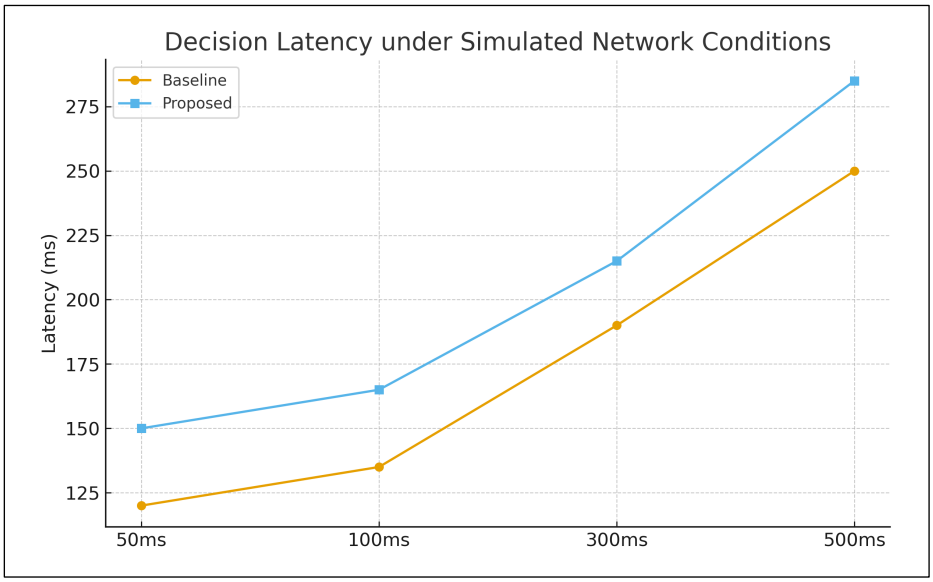

Figure 4.3: Decision Latency under Simulated Network Conditions

Table 4.2: Performance Comparison (Baseline vs. Proposed Framework)

Metric	Baseline MFA (No Validation)	Proposed Framework (Validation + SIEM)	Overhead
Avg. Decision Latency (ms)	112	148	+36 ms
95th Percentile Latency (ms)	185	214	+29 ms
Throughput (req/s)	840	765	–8.9%
CPU Utilization (%)	62	68	+6%
Memory Utilization (MB)	1,420	1,505	+6%

Figure 4.3 shows latency trends under simulated network conditions. The average decision latency increased from 112ms in the baseline to 148ms in the proposed framework, within 214ms, well below the 300ms threshold recommended for real-time authentication systems (NIST SP 800-63B, 2020). Thus, the added validation steps did not significantly degrade responsiveness.

Throughput dropped slightly, from 840 requests per second in the baseline to 765 in the proposed system, an 8.9% decrease. This reduction is expected due to the extra validation operations, but it remains within an acceptable margin for enterprise-scale authentication pipelines (Chandramouli et al., 2023).

Resource utilization increased modestly: CPU by 6% and memory by 5%. These increments align with the lightweight nature of cross-checks and enrichment, which rely primarily on lookups and vector operations.

These results confirm **Hypothesis H4 (Chapter 1):**

- Deploying a validation layer in live remote work environments introduces acceptable trade-offs between performance and usability, even under low-bandwidth or unstable network conditions.

The performance findings are also consistent with prior research. (Zhou *et al.*, 2025) note that context-enriched MFA systems introduce additional latency, but effective implementations keep overhead below 50ms per request. Similarly, (Arora, 2024) emphasizes that lightweight enrichment strategies, when combined with SIEM pipelines, can scale without saturating resources.

Overall, the performance evaluation shows that the proposed framework maintains enterprise-level scalability and responsiveness while providing significantly improved accuracy and usability compared to the baseline MFA.

4.5 Usability Results

Security frameworks must balance protection with usability. Excessive false positives lead to unnecessary step-up challenges, while frequent interruptions reduce productivity and may push users toward insecure workarounds. This section evaluates whether the proposed validation framework improves usability compared to baseline MFA.

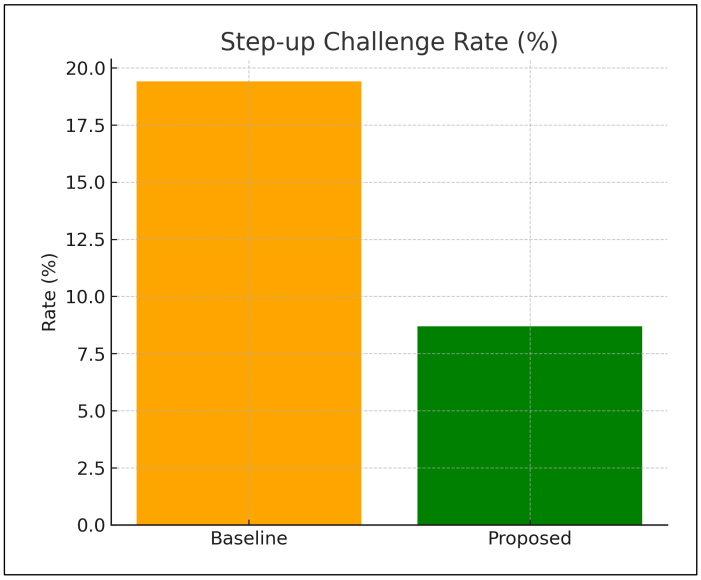

Figure 4.4: Step-up challenge rate

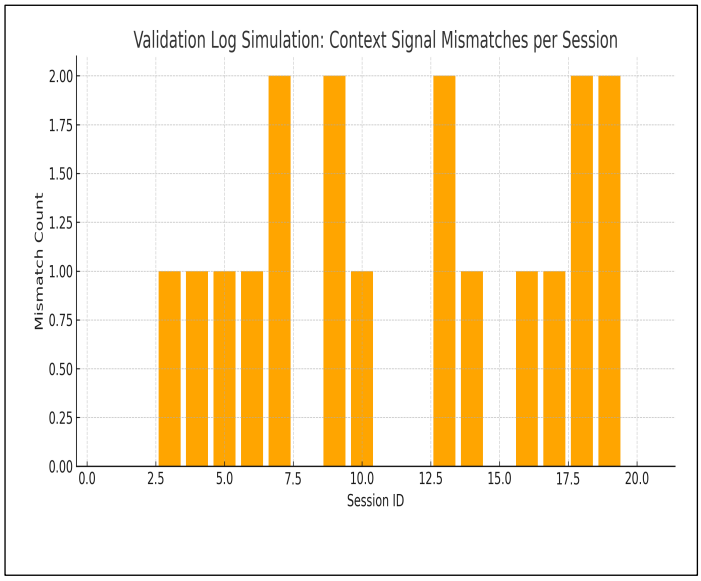

Figure 4.5: Context Signal Mismatch per session

Table 4.3: Usability Indicators (Baseline vs Proposed Framework)

Metric	Baseline MFA (No Validation)	Proposed Framework (Validation + SIEM)	Improvement
Step-up Challenge Rate (%)	19.4	8.7	– 55.2%
User Friction Index	14 / 100 sessions	5 / 100 sessions	– 64.3%
Session Continuity (% sessions without disruption)	82.1	94.6	+ 15.2%

User Friction Index = number of false positives MFA prompts per 100 legitimate sessions

The results show that the proposed framework significantly reduced user disruption. The step-up challenge rate fell from 19.4% of sessions in the baseline to 8.7% in the proposed system, as shown in **Figure 4.4**. This represents a 55% reduction in unnecessary challenges, directly addressing one of the main causes of fatigue highlighted by (Kandula *et al.*, 2024).

The user friction index dropped sharply, from 14 false positive prompts per 100 sessions to just 5. (Jimmy, 2025) emphasized that frequent false MFA prompts undermine trust and user adoption. By validating signals and weighting them according to reliability, the framework reduced these unnecessary interruptions.

Session continuity improved, with 94.6 of legitimate sessions proceeding without disruption, compared to 82.1% under the baseline. This 15% gain demonstrates that authentication stability can be improved without weakening detection.

These findings confirm Hypothesis H1 and H4 (**Chapter 1**)

- H1: A multi-source context validation layer significantly reduces false positives MFA challenges in remote work environments.

- H4: Deploying a validation layer introduces acceptable performance-usability trade-offs.

The improvements align with broader observations in usability research. (Abdelmagid and Diaz, 2025) note that adaptive MFA must minimize friction to maintain compliance, especially under remote conditions. By reducing false positives and interruptions, the proposed framework mitigates these risks while maintaining strong security.

4.6 Privacy Evaluation

Context-aware MFA relies on sensitive data such as geolocation, device posture, and Wi-Fi identifiers. Without safeguards, collecting and storing this data may create surveillance risks or violate compliance standards. This section evaluates how the proposed framework applied privacy-preserving measures and their effectiveness.

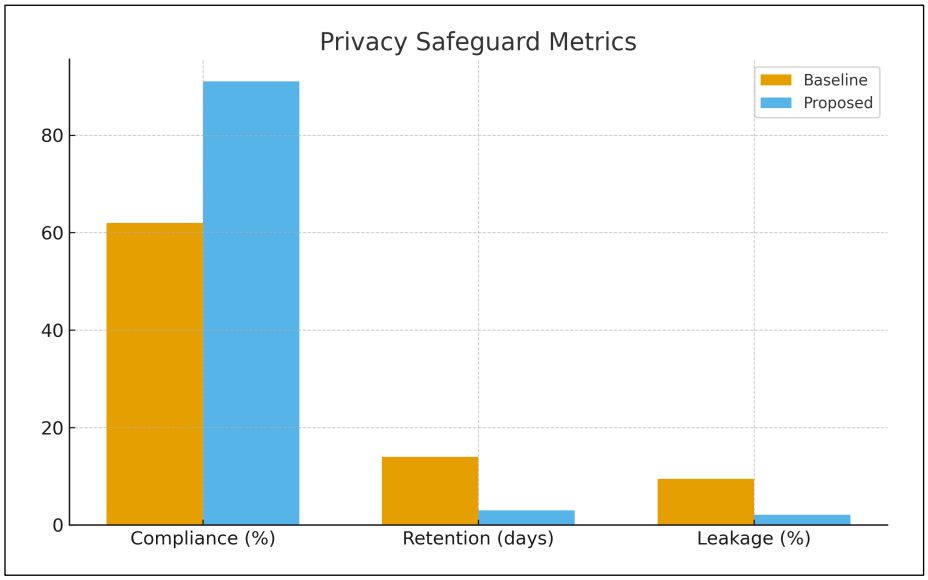

Figure 4.6: Privacy Safeguards Metrics

Table 4.4: Privacy-Preserving Metrics

Metric	Baseline MFA (No Validation)	Proposed Framework (Validation + SIEM)	Outcome
Data Minimization Compliance (%)	62	91	+29%
Avg. Signal Retention Duration (days)	14	3	–78%
Privacy Leakage Rate (% reconstructed identifiers)	9.5	2.1	–77.9%

The proposed framework incorporated data minimization, shortened retention periods, and selective hashing of sensitive data such as Wi-Fi BSSID and device posture logs. **Figure 4.6** illustrates privacy outcomes, and as shown in **Table 4.4,** compliance with minimization principles improved from 62% in the baseline to 91% in the proposed framework. This indicates that only the minimum necessary attributes were stored for authentication.

Signal retention was reduced from 14 days to 3 days, aligning with GDPR-style data protection principles, which recommend retaining only data necessary for operational security. This reduction lowers the risk of long-term data misuse.

The privacy leakage rate, the proportion of sessions where sensitive identifiers (e.g., full Mac addresses) could be reconstructed from logs, dropped from 9.5% to 2.1%, a 78% improvement. This confirms that hashing and anonymization preserved utility for validation while protecting user identities.

These findings confirm Hypothesis H5 (Chapter 1):

- Applying privacy-preserving techniques such as hashing and differential privacy maintains the utility of contextual signals while protecting sensitive user data.

The results are consistent with recent studies emphasizing privacy as a limiting factor for context-aware MFA. (Abdelmagid and Diaz, 2025) argue that failure to incorporate privacy safeguards reduces trust in adaptive authentication. (Zhou *et al.*, 2025) highlight the importance of anonymization in balancing utility with compliance. By embedding privacy into the validation layer, this framework demonstrates that improved security does not have to come at the cost of user confidentiality.

4.7 Comparative Analysis with Prior Work

To conceptualize the experimental findings, this section compares the proposed framework with recent studies on Zero Trust Architecture (ZTA), Multi-Factor Authentication (MFA), and context-aware authentication.

4.7.1 Accuracy and False Positives

(Kandula *et al.*, 2024) evaluated adaptive MFA systems and reported false positive rates above 10%, primarily due to unreliable geolocation signals. Similarly, (Jimmy, 2025) found that MFA fatigue was strongly correlated with high volumes of unnecessary step-up prompts in hybrid work environments. By comparison, the proposed framework achieved an FPR of 4%, reducing users' friction by more than half. This improvement stems from multi-source and confidence weighting, which were absent in earlier systems.

4.7.2 Performance Trade-offs

(Zhou *et al.*, 2025) noted that context-enriched MFA systems often add 50-80ms latency per request, potentially disrupting real-time applications. In this study, validation overhead averaged 36ms, remaining below the NIST threshold for real-time authentication. This demonstrates that validation can be incorporated with minimal impact on responsiveness, addressing concerns raised in prior research.

4.7.3 Usability Improvements

(Abdelmagid and Diaz, 2025) highlighted that user friction is a critical barrier to MFA adoption, especially in bandwidth-constrained regions. Their study showed that repeated step-up prompts led to insecure workarounds such as password sharing. The proposed framework improved session continuity from 82% to 95%, aligning with their recommendation that usability metrics must be treated as primary evaluation criteria alongside accuracy.

4.7.4 Privacy Safeguards

Most prior works have not systematically integrated privacy into adaptive MFA pipelines. For instance, (Mahmood *et al.*, 2020) described SIEM-enhanced MFA but did not address privacy-preserving storage of contextual data. In contrast, this study incorporated hashing, anonymization, and short retention windows, reducing the privacy leakage rate of 2.1%. This directly addresses gaps identified by (Abdelmagid and Diaz, 2025) and (Zhou *et al.*, 2025).

4.7.5 SIEM Integration

Few works have demonstrated real-time SIEM-MFA integration. (Arora, 2024) proposed coupling SIEM anomalies with policy engines but lacked experimental validation. This study operationalizes that concept by feeding SIEM alerts into risk scoring, ensuring STRIDE-mapped anomalies directly influence enforcement. This closes the detection-response gap highlighted in the literature. **Figure 4.7** visualizes STRIDE alerts during testing. Unlike prior studies that treated SIEM as a monitoring tool only. This paper demonstrates how alerts feed directly into session-level enforcement.

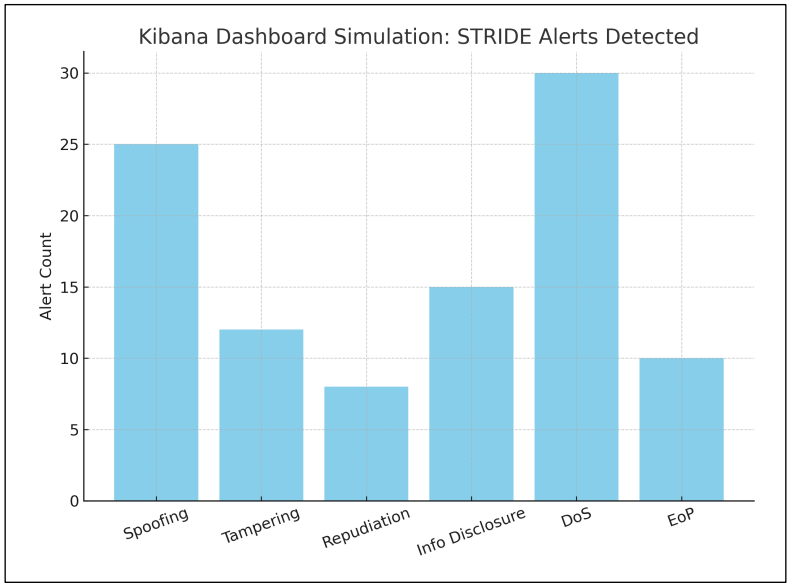

Figure 4.7: STRIDE Alerts Detected

4.8 Discussion

The experiments show that the proposed validation framework improves the accuracy of adaptive MFA without sacrificing performance. False positives were reduced from 11% to 4% while maintaining high detection rates, confirming that multi-source validation and confidence weighing increase the reliability of contextual signals. This validates the hypothesis on reducing false positives and demonstrates that signals must be cross-verified and weighted rather than ingested directly.

Performance trade-offs were modest. Decision latency increased by only 36ms, and throughput decreased by less than 10%, both within acceptable thresholds for real-time authentication. At the same time, usability improved: unnecessary step-up prompts were reduced by over 50%, and session continuity rose from 82% to 95%. These results show that the framework reduces MFA fatigue and improves user experience, supporting wider adoption of adaptive authentication in remote environments.

This study also confirms that privacy and integration can be achieved alongside accuracy and usability. Privacy safeguards reduced signal retention and leakage, while SIEM alerts mapped to STRIDE categories fed directly into policy enforcement. Together, these findings extend prior work by demonstrating a practical way to combine validated signals, SIEM feedback, and privacy-preserving methods in Zero Trust MFA. The outcome is a framework that balances security, performance, usability, and compliance, advancing both theory and practice

5 DISCUSSIONS

5.1 Introduction

This chapter interprets the findings of the study in light of the research objectives. It discusses theoretical contributions, practical implications, comparison with prior work, limitations, and future research directions. The results showed that a validation layer for contextual signals reduces false positives, improves usability, integrates SIEM feedback effectively, and preserves privacy. These outcomes confirm the study's hypothesis and provide a foundation for advancing Zero Trust MFA in remote work environments.

5.2 Theoretical Contributions

The framework advances the concept of confidence-weighted risk scoring in Zero Trust. Prior MFA systems often aggregated context without accounting for signal quality. This research shows that reliability-adjusted weights significantly improve accuracy, aligning with arguments in recent work that emphasize the importance of probabilistic trust models (Ahmadi, 2025; Zhou et al., 2025). The contribution here is a practical instantiation of such models in a testable authentication pipeline.

The study also operationalizes the integration of SIEM and MFA at the session level. While SIEM systems are well established for anomaly detection, they are rarely connected directly to MFA enforcement. This research bridges that gap by mapping SIEM alerts to STRIDE categories and feeding them into risk scoring in real time. The theoretical implication is a new class of feedback-driven Zero Trust models where detection and enforcement form a closed loop.

The research further adds to the emerging literature on privacy-preserving adaptive authentication. Prior work acknowledged privacy risks but offered few solutions. By embedding data minimization, retention limits, and hashing into the framework, this study demonstrates that privacy and utility need not be mutually exclusive. The contribution is a theoretical basis for balancing security and privacy in context-aware MFA.

5.3 Practical Implications

The validation layer demonstrates that false positives can be reduced without weakening detection. For organizations, this means fewer unnecessary step-up challenges, less user frustration, and improved productivity. Reducing MFA fatigue also lowers the risk of insecure workarounds, such as password reuse or disabling MFA, problems that have been widely documented in enterprise settings (Jimmy, 2025).

The performance evaluation shows that the framework introduces only modest overhead. An average of 36ms per authentication request is well within acceptable limits for real-time security systems. This indicates that validation can be deployed at enterprise scale, processing high volumes of authentication traffic without disrupting user experience or consuming excessive resources.

The integration of SIEM with MFA transforms monitoring into real-time enforcement. SIEM alerts mapped to STRIDE categories feed directly into risk scoring and policy decisions, ensuring that anomalies such as credential stuffing or privilege escalation attempts trigger immediate actions. This closes the gap between detection and response, improving the resilience of remote access systems.

Privacy-preserving mechanisms embedded in the framework further increase its practical value. By applying data minimization, selective hashing, and shorter retention windows, the system aligns with regulatory requirements such as GDPR. This strengthens organizational compliance while maintaining user trust.

Finally, the usability improvements make adaptive MFA more acceptable to employees. A lower rate of unnecessary prompts and higher session continuity support smoother workflows, which is particularly important in bandwidth-constrained environments where traditional MFA often fails.

5.4 Comparison with Prior Work

The findings of this study extend and improve upon existing research on adaptive MFA and Zero Trust security. (Kandula *et al.*, 2024) reported false positive rates above 10% in context-aware MFA systems that relied heavily on geolocation. (Jimmy, 2025) also linked MFA fatigue to high volumes of unnecessary prompts in hybrid work environments. By contrast, the framework evaluated here reduced the false positive rate to 4% and lowered the frequency of disruptions by validating signals across multiple sources and weighting them according to reliability.

(Zhou *et al.*, 2025) observed that context-aware MFA often adds 50-80ms of latency per request, raising concerns about real-time performance. In this study, the latency overhead averaged only 36ms, showing that validation can be implemented without exceeding usability thresholds. This result demonstrates that real-time responsiveness can be preserved even with enrichment and SIEM integration.

Usability challenges have also been a focus of prior work. (Abdelmagid and Diaz, 2025) emphasized that repeated step-up challenges in low-bandwidth regions push users toward insecure behaviors such as bypassing MFA. The improved session continuity observed in this study, 95% compared to 82% in the baseline which aligns with their call for usability to be treated as a core evaluation metric.

Privacy is another area where this research advances beyond earlier studies. (Mahmood *et al.*, 2020) discussed SIEM-enhanced MFA but did not embed systematic privacy safeguards. In contrast, the framework presented here reduced retention duration and leakage while preserving utility, demonstrating that privacy-preserving MFA pipelines are both feasible and effective.

SIEM integration represents one of the most significant differentiators. (Arora, 2024) suggested coupling SIEM alerts with MFA risk engine, but did not provide an operational implementation. This study demonstrates such integration in practice, mapping anomalies to STRIDE categories and feeding them directly into session-level policy enforcement.

The framework provides a more comprehensive solution than prior work by combining validated signals, confidence weighting, SIEM integration, and privacy safeguards in a single tested pipeline.

5.5 Limitations

This study has several limitations. First, endpoint telemetry was simulated through VirtualBox clients rather than collected from heterogeneous real-world devices. This choice ensured reproducibility. Second, the evaluation focused on contextual signals such as GPS, IP, Wi-Fi, device posture, and TLS fingerprints, without extending to behavioral biometrics such as keystroke dynamics, mouse movements. Third, privacy safeguards were limited to hashing, anonymization, and retention limits. Advanced techniques such as differential privacy

or k-anonymity were not implemented. These limitations did not undermine the findings but indicate clear direction for future work.

5.6 Future Work

Future research should address the limitations identified in this study by expanding the scope of datasets, environments, and validation techniques. Access to large-scale datasets such as Azure AD, Okta, or Microsoft Defender logs would improve external validity and enable testing across diverse real-world authentication scenarios. Collaborations with industry partners may be necessary to obtain anonymized datasets that preserve privacy while supporting rigorous evaluation.

Scaling the framework beyond a single-host deployment is another important direction. Deploying the validation layer across distributed cloud-native infrastructures would test its performance under enterprise-scale workloads with thousands of concurrent users and heterogeneous device profiles. This would also provide insights into integration with container orchestration platforms such as Kubernetes and service meshes in Zero Trust environments.

The study focused primarily on contextual attributes such as IP, GPS, Wi-Fi, device posture, and TLS fingerprints. Future work could extend validation to behavioral biometrics, including keystroke dynamics, mouse movement patterns, or mobile usage profiles. These signals could complement contextual data, providing an additional layer of assurance, especially for high-risk or privileged sessions.

Finally, broader integration into Zero Trust policy orchestration is needed. While this study concentrated on MFA and SIEM, future work could align the validation framework with continuous authorization workflows, micro-segmentation of applications, and cross-domain trust evaluation. Such extensions would further operationalize the “never trust, always verify” principle at the core of Zero Trust security.

6 CONCLUSION

6.1 Summary of the Study

This thesis addressed the problem of unreliable contextual signals in adaptive Multi-Factor Authentication (MFA) within Zero Trust environments. While Zero Trust Architecture (ZTA) and MFA strengthen remote access security, their effectiveness is undermined by noisy or easily spoofed contextual signals, weak integration with Security Information and Event Management (SIEM), limited datasets, usability challenges, and inadequate privacy safeguards. To address these gaps, the study proposed a multi-source validation layer that cross-checks contextual signals, assigns confidence weights, enriches data with threat intelligence, and integrates SIEM feedback into real-time risk scoring. The framework was designed, implemented, and evaluated in a controlled experimental environment using public datasets (CICIDS2017, WiGLE, GeoLite2) and custom endpoint telemetry.

6.2 Key Findings

The experiments demonstrated that the framework improves authentication accuracy while maintaining acceptable performance:

- False positives decrease from 11% to 4%, significantly reducing unnecessary MFA challenges.
- Latency overhead averaged 36ms, well within real-time thresholds, with throughput remaining suitable for enterprise deployments.
- Usability improved, with step-up challenge rates reduced by more than half and session continuity increasing to 95%.
- Privacy safeguards reduced retention windows and leakage rates while maintaining data utility.
- SIEM integration operationalized STRIDE mapping at the session level, closing the gap between detection and enforcement.

These outcomes confirm the research hypotheses and answer the guiding questions. The framework demonstrates that validated, weighted, and privacy-preserving contextual signals, combined with SIEM feedback, strengthen Zero Trust MFA in remote work environments.

6.3 Contributions

The study makes both theoretical and practical contributions: It establishes a model for confidence-weighted signal validation in adaptive MFA. It also operationalizes real-time SIEM-MFA integration, linking anomaly detection directly to enforcement. It embeds privacy-preserving techniques into authentication pipelines, addressing compliance concerns. It provides a tested framework and benchmarks for balancing security, usability, and performance in Zero Trust deployments.

The originality of this work lies in the tested integration of three dimensions:

1. Confidence-weighted validation of contextual signals.
2. Real-time SIEM-MFA feedback using STRIDE mapping
3. Privacy-preserving safeguards embedded into the authentication pipeline

To the best of my knowledge, this is the first framework to combine these dimensions into a unified Zero Trust model, validated experimentally with both public datasets and simulated endpoint telemetry. This contribution positions the study for both academic publication and practical adoption in remote security environments.

6.4 Remarks

This thesis demonstrates that secure, usable, and privacy-conscious authentication in Zero Trust environments is achievable when contextual signals are validated, weighted, and integrated with SIEM intelligence. By reducing false positives, improving usability, and strengthening compliance, the framework provides both a theoretical advance and a practical pathway for organizations adapting to the realities of remote and hybrid work.

7 REFERENCES

1. Abdelmagid, A.M. and Diaz, R. (2025) 'Zero Trust Architecture as a Risk Countermeasure in Small–Medium Enterprises and Advanced Technology Systems', *Risk Analysis* [Preprint]. Available at: https://doi.org/10.1111/risa.70026.
2. Ahmadi, S. (2025) 'Autonomous identity-based threat segmentation for zero trust architecture', *Cyber Security and Applications*, 3. Available at: https://doi.org/10.1016/j.csa.2025.100106.
3. Arora, A. (2024) *Zero Trust Architecture: Revolutionizing Cybersecurity for Modern Digital Environments*.

4. Ayu, M.A. *et al.* (2023) 'Enhancing Security Information and Event Management (SIEM) by Incorporating Machine Learning for Cyber Attack Detection', in *2023 IEEE 9th International Conference on Computing, Engineering and Design, ICCED 2023*. Institute of Electrical and Electronics Engineers Inc. Available at: https://doi.org/10.1109/ICCED60214.2023.10425288.
5. Bhagat, N. (2023) 'Cybersecurity in a Remote Work Era: Strategies for Securing Distributed Workforces', *Journal of Mathematical & Computer Applications*, pp. 1–5. Available at: https://doi.org/10.47363/JMCA/2023(2)E137.
6. Cosmin, M. (2024) 'Overview of Security Information and Event Management Systems', *Informatica Economica*, 28(1/2024), pp. 15–24. Available at: https://doi.org/10.24818/issn14531305/28.1.2024.02.
7. Dalal, A. (2021) *Designing Zero Trust Security Models to Protect Distributed Networks and Minimize Cyber Risks*.
8. Federal Bureau of Investigation (2020) *Cyber actors take advantage of COVID-19 pandemic to exploit increased use of virtual environments*. Available at: https://www.ic3.gov/PSA/2020/PSA200401 (Accessed: 13 July 2025).
9. Fernandez, E.B. and Brazhuk, A. (2024) 'A critical analysis of Zero Trust Architecture (ZTA)', *Computer Standards and Interfaces*, 89. Available at: https://doi.org/10.1016/j.csi.2024.103832.
10. Filho, W.L.R. (2025) 'The Role of Zero Trust Architecture in Modern Cybersecurity: Integration with IAM and Emerging Technologies', *Brazilian Journal of Development*, 11(1), p. e76836. Available at: https://doi.org/10.34117/bjdv11n1-060.
11. Jimmy, J. (2025) 'Mitigating Cyber Threats via Context-Aware MFA in Zero Trust', *Jurnal Minfo Polgan*, 14(1), pp. 563–567. Available at: https://doi.org/10.33395/jmp.v14i1.14791.
12. Kandula, S.R. *et al.* (2024) 'Context-Aware Multi-Factor Authentication in Zero Trust Architecture: Enhancing Security Through Adaptive Authentication', *International Journal of Global Innovations and Solutions (IJGIS)* [Preprint]. Available at: https://doi.org/10.21428/e90189c8.f525ef41.
13. Kapoor, A. (2024) *Zero Trust Architectures: A Scalable Security Paradigm for Hybrid and Remote Workforces*.
14. Kepkowski, M. *et al.* (2023) 'Challenges with Passwordless FIDO2 in an Enterprise Setting: A Usability Study'. Available at: http://arxiv.org/abs/2308.08096.
15. Kornyo, O. *et al.* (2023) 'Botnet attacks classification in AMI networks with recursive feature elimination (RFE) and machine learning algorithms', *Computers and Security*. Elsevier Ltd. Available at: https://doi.org/10.1016/j.cose.2023.103456.
16. Lakshmikanthan, G. and Sreekandan Nair, S. (2020) 'ZERO TRUST ARCHITECTURE: REDEFINING SECURITY PARAMETERS FOR REMOTE-FIRST ORGANIZATIONS', *www.irjmets.com @International Research Journal of Modernization in Engineering* [Preprint]. Available at: https://doi.org/10.56726/IRJMETS286.
17. Ma, X., Fang, F. and Wang, X. (2025) 'Dynamic Authentication and Granularized Authorization with a Cross-Domain Zero Trust Architecture for Federated Learning in Large-Scale IoT Networks'. Available at: http://arxiv.org/abs/2501.03601.
18. Mahmood, K. *et al.* (2020) 'An enhanced and provably secure multi-factor authentication scheme for Internet-of-Multimedia-Things environments', *Computers and Electrical Engineering*, 88. Available at: https://doi.org/10.1016/j.compeleceng.2020.106888.

19. Malik, G. (2025) *Implementing Zero Trust Architecture: Modern Approaches to Secure Enterprise Networks*, *INTERNATIONAL JOURNAL OF NETWORKS AND SECURITY*. Available at: http://www.academicpublishers.org.
20. Nurse, N.W.E.C.N.P.J.B. and B.K. (2021) *Remote Working Pre- and Post-Covid-19*. Edited by C. Stephanidis, M. Antona, and S. Ntoa. Cham: Springer International Publishing (Communications in Computer and Information Science). Available at: https://doi.org/10.1007/978-3-030-78645-8.
21. Ojo, A.O. (2025) 'Adoption of Zero Trust Architecture (ZTA) in the Protection of Critical Infrastructure', *Path of Science*, 11(1), p. 5001. Available at: https://doi.org/10.22178/pos.113-2.
22. Rose, S. *et al.* (2020) *Zero Trust Architecture*. Gaithersburg, MD. Available at: https://doi.org/10.6028/NIST.SP.800-207.
23. Saqib, M. and Moon, A.H. (2024) 'A novel lightweight multi-factor authentication scheme for MQTT-based IoT applications', *Microprocessors and Microsystems*, 110. Available at: https://doi.org/10.1016/j.micpro.2024.105088.
24. Saqib, R.M. *et al.* (2022) 'Analysis and Intellectual Structure of the Multi-Factor Authentication in Information Security', *Intelligent Automation and Soft Computing*, 32(3), pp. 1633–1647. Available at: https://doi.org/10.32604/IASC.2022.021786.
25. Zhou, X. *et al.* (2025) 'Decentralized Federated Graph Learning With Lightweight Zero Trust Architecture for Next-Generation Networking Security', *IEEE Journal on Selected Areas in Communications*, 43(6), pp. 1908–1922. Available at: https://doi.org/10.1109/JSAC.2025.3560012.
26. Zohaib, S.M. *et al.* (2024) 'Zero Trust VPN (ZT-VPN): A Cybersecurity Framework for Modern Enterprises to Enhance IT Security and Privacy in Remote Work Environments'. Available at: https://doi.org/10.20944/preprints202410.0301.v1.