INTERNATIONAL JOURNAL OF NETWORKS AND SECURITY
(ISSN: 2693-387X)
VOL.06 ISSUE 02 (2026)

Research Article

Adversarial Threat Modeling for SSO and MFA Infrastructure: A Systematic STRIDE-Based Security Evaluation

Amina Bello[1]

[1]Department of Artificial Intelligence and Data Analytics, Center for Computational and Intelligent Systems, Lagos, Nigeria

check for updates

Received: 12 June 2026
Revised: 2 July 2026
Accepted: 20 August 2026
Published: 01 September 2026

Doi:10.55640/ijns-06-02-02
Page No: 22-30

Abstract

Federated Single Sign-On (SSO) and Multi-Factor Authentication (MFA) infrastructures have become foundational components of contemporary identity and access management architectures. Their security, however, depends not only on the strength of individual authentication mechanisms but also on the interactions among identity providers, service providers, authentication protocols, tokens, sessions, recovery mechanisms, and trust relationships. This paper presents a systematic STRIDE-based security evaluation methodology for identifying and analyzing adversarial threats against SSO and MFA infrastructure. The methodology models authentication and federation workflows as security-sensitive information flows and maps potential threats to the STRIDE categories of spoofing, tampering, repudiation, information disclosure, denial of service, and elevation of privilege. Particular attention is given to trust-boundary crossings, authentication assertions, session state, credential recovery, token processing, and administrative interfaces. The supplied literature is primarily concerned with statistical modeling, forecasting, uncertainty estimation, and model evaluation rather than identity-security engineering; consequently, it provides methodological support for systematic analysis and uncertainty-aware evaluation rather than direct empirical evidence concerning SSO or MFA attacks. The study therefore positions STRIDE as a structured analytical framework and proposes risk-oriented evaluation criteria for comparing threat exposure across federated authentication components. The resulting framework emphasizes attack-surface decomposition, threat identification, severity assessment, control analysis, and residual-risk evaluation. The analysis demonstrates that authentication strength alone does not eliminate systemic risks created by federation trust, token handling, recovery paths, and privileged administrative functions.

Keywords: Federated Identity, Single Sign-On, Multi-Factor Authentication, STRIDE, Threat Modeling, Attack Surface, Identity Security, Security Evaluation, Authentication Infrastructure.

INTRODUCTION

Federated identity infrastructures enable users to authenticate through a trusted identity provider and subsequently access multiple relying or service applications without independently authenticating to each application. SSO therefore reduces authentication friction and can centralize identity administration. MFA further strengthens authentication by requiring multiple authentication factors rather than relying exclusively on a password. Nevertheless, centralization and federation introduce security dependencies: compromise of an identity provider, authentication transaction, assertion, token, session, recovery process, or administrative interface can potentially affect multiple downstream services.

The central security problem is consequently broader than determining whether an individual authentication mechanism is cryptographically or operationally strong. The security of an SSO/MFA ecosystem depends on how components interact across trust boundaries and how security assumptions propagate between them. A system may employ MFA while still exposing weaknesses through token substitution, session manipulation, account recovery, authorization inconsistencies, or compromised federation relationships. Threat modeling provides a means of examining these interactions before or alongside penetration testing and operational monitoring.

STRIDE is particularly suitable for this purpose because it classifies threats according to recognizable security properties: identity authenticity, integrity, accountability, confidentiality, availability, and authorization. A systematic STRIDE assessment can therefore transform a complex identity architecture into a collection of analyzable assets, processes, data flows, and trust boundaries. Ganapathy (2024) similarly frames federated SSO and MFA security through STRIDE-based analysis of attack vectors, providing the principal conceptual reference for applying the framework to the present problem.

The methodological challenge addressed by this paper is the development of a systematic evaluation process rather than a simple catalog of attacks. Threat modeling requires decisions concerning what constitutes an asset, where trust changes occur, which threats are plausible, how their consequences should be assessed, and which controls reduce residual exposure. The broader methodological literature supplied for this study emphasizes systematic model construction, evaluation, uncertainty, and forecasting performance. These principles are relevant to the organization of a repeatable security assessment even though the cited studies do not directly investigate identity infrastructure.

The objective of this research is therefore to develop a structured STRIDE-based evaluation model for adversarial threats affecting SSO and MFA infrastructure. The study focuses on five questions: (1) which architectural components constitute the principal attack surface; (2) how STRIDE categories manifest across federated authentication workflows; (3) which trust boundaries produce the most consequential security dependencies; (4) how threats can be systematically prioritized; and (5) what limitations must be recognized when using a model-based security evaluation.

The scope encompasses identity providers, service providers, federation interfaces, authentication exchanges, MFA mechanisms, tokens or assertions, sessions, recovery processes, and administrative interfaces. It does not attempt to claim empirical attack frequencies because the supplied references do not provide a corresponding cybersecurity incident dataset.

LITERATURE REVIEW

The supplied literature is dominated by research on time-series analysis, forecasting, prediction intervals, statistical inference, and model evaluation. This creates an important distinction between domain evidence and methodological evidence. The references cannot legitimately be interpreted as empirical studies of SSO, MFA, or STRIDE. Instead, they provide methodological concepts concerning systematic estimation, uncertainty, model comparison, and evaluation that can inform the design of a structured analytical methodology.

Brockwell and Davis (1991) and Cryer and Chan (2008) establish foundational treatments of time-series analysis, while Shumway and Stoffer (2010) provide a broader statistical framework incorporating computational analysis. These works demonstrate the importance of formally defining the analytical structure before interpreting observed patterns. Although their statistical subject differs from identity security, the same methodological principle applies to threat modeling: the analyst must define system components, relationships, assumptions, and evaluation dimensions before drawing

conclusions about security exposure.

Gooijer and Hyndman (2006) review developments in time-series forecasting and emphasize the evolution of forecasting methods and their comparative evaluation. Cerqueira, Torgo, and Mozetič (2020) explicitly investigate performance-estimation methods for forecasting models, highlighting the importance of evaluation methodology rather than relying solely on model construction. This principle is directly relevant to threat modeling because a STRIDE assessment should not end with threat enumeration; it requires a consistent basis for assessing the relative importance of identified threats.

Uncertainty is also central to the supplied literature. Chatfield (1993) examines interval forecasts, while Pan and Politis (2016) investigate bootstrap prediction intervals across different autoregressive settings. Kong, Gu, and Yang (2018) address prediction intervals through estimation of multi-step innovation distributions. Zhong (2024) examines statistical inference for innovation distributions in ARMA and multi-step-ahead prediction. Collectively, these works reinforce the methodological observation that analytical conclusions should account for uncertainty rather than presenting model outputs as inherently exact.

In the context of security evaluation, uncertainty arises because threat models depend on assumptions about attacker capability, architecture, configuration, user behavior, and control effectiveness. A conceptual risk score should therefore not be treated as an objective measurement of real-world attack probability. It is an analytical estimate conditioned on the model assumptions.

Inoue, Jin, and Rossi (2017) examine rolling-window selection for out-of-sample forecasting under time-varying parameters. Their work highlights the importance of recognizing changing conditions when evaluating predictive models. This principle has an analogous implication for identity security: an SSO/MFA threat model should be treated as a living assessment rather than a permanently valid representation. Federation partners, authentication policies, applications, recovery mechanisms, and administrative configurations evolve over time.

Fan and Gijbels (1996), Pierce (1971), Shao and Yang (2017), and Wang et al. (2014) provide additional statistical foundations involving estimation, regression errors, autoregressive structures, and efficient inference. Shao et al. (2024) further illustrates the use of longitudinal modeling for examining trends across changing conditions. While these studies do not establish cybersecurity findings, they collectively support the methodological importance of explicit assumptions, structured estimation, and longitudinal reassessment.

Petropoulos et al. (2022) and Abbasimehr, Paki, and Bahrini (2022) demonstrate the application of forecasting models to COVID-19 data, while R Core Team (2024) documents the computational environment used for statistical computing. Their relevance to this paper is primarily methodological: reproducibility and systematic computational analysis are important principles when developing repeatable evaluation procedures.

The most directly relevant supplied work is Ganapathy (2024), which specifically addresses threat modeling for federated SSO and MFA systems through STRIDE-based attack-vector analysis. It provides the conceptual basis for positioning STRIDE as a systematic framework for evaluating threats across federated authentication infrastructure.

The principal research gap is therefore clear. The supplied literature provides substantial methodological foundations for systematic analysis but very limited direct empirical evidence concerning SSO/MFA security. This paper addresses that gap at the framework level by integrating STRIDE-based identity threat analysis with disciplined model-

evaluation principles rather than treating statistical references as cybersecurity evidence.

METHODOLOGY

3.1 Research Design

This study adopts a systematic, architecture-oriented threat-modeling design. The analytical unit is not an individual vulnerability but a security-relevant interaction between an actor, component, process, data object, and trust boundary.

The methodology consists of five sequential stages: system decomposition, trust-boundary identification, STRIDE threat generation, risk-oriented evaluation, and control/residual-risk analysis. The approach is designed to be repeatable across different federation architectures.

The methodology follows the general principle that evaluation quality depends on a clearly specified analytical structure. This is consistent with the methodological emphasis on systematic model evaluation found in the supplied literature (Cerqueira et al., 2020).

3.2 System Decomposition

The first stage decomposes an SSO/MFA environment into security-relevant components. A representative architecture contains an end user, endpoint device, identity provider (IdP), MFA service, authentication interface, federation protocol interface, service provider (SP), authorization subsystem, session-management layer, token/assertion-processing mechanism, account-recovery service, and administrative console.

Each component is evaluated according to four properties: assets processed, trust assumptions, external dependencies, and consequences of compromise.

For example, the identity provider represents a high-value concentration point because it authenticates users and may issue authentication assertions consumed by multiple applications. A service provider may have a narrower scope but can still expose authorization weaknesses if it incorrectly interprets identity or privilege information received from the federation layer.

3.3 Trust-Boundary Analysis

Trust boundaries represent locations at which information crosses from one security domain into another. In federated authentication, significant boundaries may exist between the user's endpoint and IdP, IdP and MFA provider, IdP and SP, SP and application backend, and administrative users and management interfaces.

A boundary is security-critical when the receiving component relies on assumptions made by another component. For instance, an SP may trust an identity assertion generated by an IdP. If assertion integrity, origin, audience, or lifecycle assumptions are violated, the downstream application may accept an identity that should not have been authorized.

Ganapathy (2024) identifies the relevance of STRIDE-based attack-vector analysis to federated SSO and MFA environments; this study extends that conceptual approach by making trust boundaries an explicit evaluation layer.

3.4 STRIDE Threat Mapping

Threats are classified using six STRIDE categories.

Spoofing concerns unauthorized impersonation of users, administrators, identity providers, or trusted federation participants. In an SSO environment, spoofing may occur

if authentication credentials, session artifacts, recovery mechanisms, or federation identities are improperly trusted.

Tampering concerns unauthorized modification of authentication messages, assertions, tokens, configuration, policy, or session state. The principal analytical question is whether an attacker can modify security-sensitive information without detection or rejection.

Repudiation concerns insufficient evidence to establish who performed a security-relevant action. This is particularly important for privileged administrative actions, authentication-policy changes, account recovery, and federation configuration.

Information disclosure involves unauthorized exposure of credentials, authentication data, identity attributes, session information, recovery information, or other sensitive security material.

Denial of service addresses attacks that prevent legitimate authentication or authorization. Because SSO centralizes authentication, disruption of a central identity service may affect numerous dependent applications simultaneously.

Elevation of privilege concerns acquisition of permissions beyond those legitimately granted. In federation environments, this may result from incorrect mapping of roles, claims, groups, or administrative privileges.

The six categories should not be treated as isolated. A single attack chain may begin with spoofing, enable tampering, and ultimately result in elevation of privilege. Consequently, the methodology evaluates both individual threats and dependencies among threats.

3.5 Adversarial Scenario Construction

For each identified threat, the analyst defines an adversarial scenario containing an entry point, targeted asset, violated security property, potential consequence, and relevant control.

A hypothetical example is a compromised user endpoint where an attacker obtains an active session artifact. The threat may initially appear to be spoofing because the attacker attempts to impersonate the user. However, the analysis should also examine whether session integrity controls are sufficient, whether MFA remains meaningful after initial authentication, and whether downstream applications independently validate session state.

A second scenario involves an administrative interface. If an attacker obtains privileged administrative access, the consequence may extend beyond a single account. Changes to federation configuration, MFA policy, trusted applications, or recovery settings could alter the security posture of an entire identity ecosystem.

3.6 Risk Evaluation

The evaluation stage considers three dimensions: likelihood, impact, and control effectiveness. Because no empirical incident-frequency dataset is supplied, the resulting assessment should be treated as qualitative or semi-quantitative rather than as a statistically validated probability estimate.

Threat severity can conceptually be represented as a function of attacker feasibility and consequence:

Risk = Threat Likelihood × Security Impact × Exposure

The formulation is intended as an analytical prioritization mechanism rather than a claim of measured probability. Model-evaluation literature demonstrates the importance of

distinguishing analytical estimation from observed performance (Cerqueira et al., 2020).

Threats should be prioritized according to concentration of privilege, breadth of affected services, persistence potential, sensitivity of exposed information, and difficulty of detection. A compromise of a centralized IdP may therefore receive higher priority than a localized SP weakness because the former can have greater systemic consequences.

3.7 Control and Residual-Risk Assessment

The final methodological stage examines whether existing controls adequately mitigate each threat. Controls may include strong authentication policies, MFA enforcement, session expiration, token validation, administrative separation, logging, anomaly detection, recovery safeguards, and federation-policy restrictions.

The analysis distinguishes inherent risk from residual risk. Inherent risk represents exposure before controls, while residual risk represents the exposure that remains after controls are considered.

This distinction is important because MFA should not automatically be interpreted as eliminating identity threats. MFA can reduce certain credential-compromise scenarios while leaving other attack surfaces—particularly sessions, recovery processes, authorization mapping, and administrative controls—subject to attack.

The model should also be reassessed periodically. The time-varying nature of analytical environments highlighted by Inoue et al. (2017) provides a useful methodological analogy: security assessments should be updated when system conditions or assumptions change.

RESULTS

The systematic STRIDE evaluation produces five principal findings concerning the security characteristics of federated SSO and MFA infrastructure.

First, identity-provider concentration creates systemic attack impact. The IdP is not simply another application component; it frequently functions as a trust anchor for multiple downstream services. Consequently, threats affecting authentication decisions, federation configuration, or privileged IdP administration may propagate across the federation. This makes centralized identity infrastructure a particularly important component for threat prioritization.

Second, MFA does not eliminate the broader SSO attack surface. MFA primarily strengthens authentication by introducing additional verification requirements. It does not inherently guarantee the security of already-established sessions, authorization mappings, federation configuration, account recovery, or administrative interfaces. The evaluation therefore identifies a distinction between authentication assurance and overall identity-system assurance.

Third, trust boundaries are the primary locations for systematic STRIDE analysis. Federation interfaces, authentication exchanges, administrative interfaces, and recovery services represent points where security assumptions are transferred between components. The threat model indicates that these boundaries deserve greater analytical attention than treating each application independently. This finding is consistent with the STRIDE-oriented federated identity perspective of Ganapathy (2024).

Fourth, threat categories frequently form attack chains rather than independent events. Spoofing may facilitate privilege escalation; tampering may alter authorization information; information disclosure may enable subsequent impersonation; and denial-of-service attacks against centralized authentication can affect multiple applications

simultaneously. Consequently, a threat register organized only as independent STRIDE categories may underestimate compound attack effects.

Fifth, risk evaluation is inherently assumption-dependent. The supplied methodological literature demonstrates the importance of evaluating models systematically and recognizing uncertainty in analytical conclusions (Chatfield, 1993; Pan & Politis, 2016). In identity-security modeling, uncertainty arises from incomplete architectural information, changing configurations, unknown attacker capabilities, and uncertain control effectiveness. Therefore, the proposed framework should be interpreted as a structured decision-support mechanism rather than a precise numerical prediction of attack probability.

The findings also reveal a limitation in the evidence base. The supplied references do not contain empirical SSO/MFA incident datasets or experimental evaluations of STRIDE controls. Accordingly, the results represent analytical findings generated from systematic threat decomposition rather than statistically validated estimates of real-world attack frequency.

DISCUSSION

The findings demonstrate that a meaningful security evaluation of SSO and MFA infrastructure must move beyond the narrow question of whether MFA is enabled. Authentication security is a system property produced by the interaction of authentication mechanisms, federation trust, session management, authorization, recovery, administration, and monitoring.

Theoretically, STRIDE provides an effective taxonomy for transforming heterogeneous identity threats into recognizable security properties. Its principal strength is analytical completeness: spoofing, tampering, repudiation, information disclosure, denial of service, and elevation of privilege collectively encourage the analyst to consider confidentiality, integrity, availability, accountability, authentication, and authorization.

However, STRIDE is fundamentally a threat-classification framework rather than a complete quantitative risk model. It can identify what type of security failure may occur but does not independently establish how frequently that failure will occur or precisely how much organizational damage it will generate. This limitation makes risk prioritization an important secondary analytical activity.

The methodological literature supplied by the user reinforces this distinction. Research on forecasting evaluation emphasizes that model outputs require appropriate evaluation procedures rather than unquestioned acceptance (Cerqueira et al., 2020). Similarly, security-threat scores should be interpreted according to their assumptions and evaluation criteria. Prediction-interval research also demonstrates the broader importance of representing uncertainty in analytical conclusions (Chatfield, 1993; Pan & Politis, 2016). Applied to threat modeling, this implies that a high-risk designation should communicate analytical concern, not an empirically established probability of compromise.

Another important implication concerns temporal change. Identity environments evolve through new applications, federation partners, authentication methods, policy changes, and administrative configurations. A threat model constructed at deployment may therefore become incomplete as the architecture changes. The methodological emphasis on changing parameters and rolling evaluation found in Inoue et al. (2017) supports the principle of periodic reassessment.

The most significant practical implication is that security controls should be evaluated against complete attack paths. For example, strengthening initial authentication while leaving account recovery weak may create an asymmetry in which the attacker avoids the

stronger authentication mechanism altogether. Similarly, protecting user credentials while insufficiently protecting administrative federation configuration can leave the identity ecosystem vulnerable to higher-impact compromise.

The analysis also highlights the importance of residual risk. No control is universally effective against every STRIDE category. MFA may significantly strengthen authentication, but it does not automatically prevent configuration tampering, information disclosure, denial of service, repudiation, or authorization errors. Therefore, security programs should combine authentication controls with integrity protection, least privilege, administrative separation, logging, recovery safeguards, and continuous assessment.

A further limitation is the mismatch between the supplied literature and the cybersecurity research domain. Most supplied references concern statistical modeling and forecasting rather than SSO/MFA security. They should consequently be used to support methodological reasoning rather than as evidence for specific cyberattack claims. Ganapathy (2024) is the principal directly relevant source for the STRIDE-based SSO/MFA topic, while the remaining literature contributes methodological concepts concerning evaluation, uncertainty, model selection, and reproducibility.

Future empirical research should validate the proposed framework using real federation architectures, controlled attack scenarios, vulnerability assessments, and longitudinal security data. Such work could establish whether the qualitative prioritization proposed here correlates with observed security incidents and operational consequences.

CONCLUSION

This paper presented a systematic STRIDE-based methodology for evaluating adversarial threats in SSO and MFA infrastructure. The analysis demonstrates that federated identity security cannot be reduced to the strength of the authentication factor alone. Security depends on the integrity of federation relationships, authentication transactions, session state, authorization mappings, recovery processes, and privileged administrative functions.

The proposed methodology decomposes identity infrastructure into components and trust boundaries, maps threats to STRIDE categories, constructs adversarial scenarios, evaluates impact and exposure, and examines residual risk after controls. Its principal contribution is the integration of architecture-oriented threat decomposition with disciplined evaluation principles.

The analysis also establishes an important methodological limitation: the supplied references are predominantly statistical and forecasting studies and therefore cannot serve as empirical evidence for specific SSO/MFA attack frequencies. Their contribution is instead methodological, particularly regarding structured evaluation, uncertainty, changing conditions, and reproducibility. Ganapathy (2024) provides the principal direct conceptual foundation for applying STRIDE to federated SSO and MFA systems.

Future research should empirically validate the framework through representative federation deployments, controlled adversarial experiments, quantitative risk measurements, and longitudinal reassessment. Such validation would strengthen the transition from conceptual threat modeling toward evidence-based security evaluation.

REFERENCES

1. Abbasimehr H, Paki R, Bahrini A (2022). A novel approach based on combining deep learning models with statistical methods for COVID-19 time series forecasting. Neural Computing & Applications, 34: 3135–3149.

2. Brockwell P, Davis R (1991). Time Series Theory Methods, 2nd ed. Springer, New York.

3. Cerqueira V, Torgo L, Mozetič I (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. Machine Learning, 109: 1997–2028.

4. Chatfield C (1993). Calculating interval forecasts. Journal of Business & Economic Statistics, 11(2): 121–135.

5. Cryer DJ, Chan K (2008). Time Series Analysis: With Applications in R, 2nd ed. Springer, New York.

6. Fan J, Gijbels I (1996). Local Polynomial Modelling and Its Applications. Taylor & Francis.

7. Ganapathy, S. K. (2024). Threat Modeling for Federated SSO and MFA Systems: STRIDE-Based Analysis of Attack Vectors. International Journal of Data Science and Machine Learning, 4(02), 55-73. https://www.academicpublishers.org/journals/index.php/ijdsml/article/view/stride-analysis-sso-mfa

8. Gooijer J, Hyndman R (2006). 25 years of time series forecasting. International Journal of Forecasting, 22: 443–473.

9. Inoue A, Jin L, Rossi B (2017). Rolling window selection for out-of-sample forecasting with time-varying parameters. Journal of Econometrics, 196: 55–67.

10. Kong J, Gu L, Yang L (2018). Prediction interval for autoregressive time series via oracally efficient estimation of multi-step-ahead innovation distribution function. Journal of Time Series Analysis, 395: 690–708.

11. Pan L, Politis DN (2016). Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. Journal of Statistical Planning and Inference, 177: 1–27.

12. Petropoulos F, Makridakis S, Stylianou N (2022). COVID-19: Forecasting confirmed cases and deaths with a simple time series model. International Journal of Forecasting, 38: 439–452.

13. Pierce DA (1971). Least squares estimation in the regression model with autoregressive-moving average errors. Biometrika, 58: 299–312.

14. R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

15. Shao Q, Polavarapu M, Small L, Singh S, Nguyen Q, Shao K (2024). A longitudinal mixed effects model for assessing mortality trends during vaccine rollout. Healthcare Analytics, 6: 100347.

16. Shao Q, Yang L (2017). Oracally efficient estimation and consistent model selection for auto-regressive moving average time series with trend. Journal of the Royal Statistical Society, Series B, Statistical Methodology, 79: 507–524.

17. Shumway RH, Stoffer DS (2010). Time Series Analysis and Its Applications: With R Examples, 3rd ed. Springer Texts in Statistics, 2010.

18. Wang J, Liu R, Cheng F, Yang L (2014). Oracally efficient estimation of autoregressive error distribution with simultaneous confidence band. The Annals of Statistics, 42: 654–668.

19. Zhong C (2024). Statistical inference for innovation distribution in ARMA and multi-step-ahead prediction via empirical process. Journal of Nonparametric Statistics, 1–27.