Operational Incompleteness in AI-Native Cybersecurity

Adversarial Uncertainty, Hidden Operational Structure, and the Limits of Observability

Hassan Nasreddine
Fortisec Research

Foundational Research White Paper

Abstract

Artificial intelligence is transforming cybersecurity from a reactive monitoring discipline into an adaptive operational environment composed of autonomous agents, probabilistic reasoning systems, distributed telemetry pipelines, and continuously evolving adversarial states. Existing cybersecurity architectures increasingly rely on AI-assisted analysis, automated orchestration, behavioral classification, and large-scale anomaly detection. Yet despite these advances, most defensive systems remain grounded in a fundamental operational assumption: observable operational behavior sufficiently characterizes the underlying security state of the system.

This assumption becomes increasingly fragile in adversarial environments. Organizations may appear operationally compliant while hidden attack surfaces, latent dependencies, adversarial persistence channels, or unobserved operational sectors remain structurally outside the reasoning model of the defensive system. Similarly, AI systems may exhibit apparently aligned or stable behavior while hidden operational dynamics remain inaccessible to observable telemetry.

This white paper investigates the possibility that modern AI-native cybersecurity systems are intrinsically vulnerable to forms of operational incompleteness under adversarial uncertainty.

Contents

1. Introduction

Artificial intelligence is rapidly transforming cybersecurity into a distributed reasoning environment composed of autonomous agents, adaptive orchestration systems, probabilistic inference mechanisms, and continuously evolving adversarial structures. AI-assisted threat detection, automated security operations, behavioral analysis, and intelligent remediation increasingly participate directly in operational defensive infrastructures.

Despite these advances, most modern cybersecurity systems remain deeply dependent upon observable operational behavior. Security telemetry, compliance indicators, network events, access logs, behavioral analytics, and system alerts collectively define the operational surface through which defensive reasoning occurs.

Yet adversarial systems are not necessarily exhausted by their observable operational states. A system may appear operationally compliant while hidden attack surfaces, latent adversarial persistence channels, unmanaged identities, implicit dependencies, or structurally unobserved operational sectors remain outside the reasoning model of the defensive architecture itself.

This problem becomes increasingly significant in AI-native cybersecurity environments where autonomous systems introduce additional layers of operational opacity through hidden reasoning chains, latent probabilistic states, emergent behavioral dynamics, autonomous adaptation, and adversarial manipulation of operational context.

The central question motivating this paper is therefore the following:

> Can AI-native cybersecurity systems ever be operationally complete under adversarial uncertainty?

The work presented here should be interpreted as a foundational systems-theory perspective rather than a product proposal, implementation roadmap, or finalized mathematical framework. The objective is to investigate whether concepts related to operational completeness, distinguishability, entropy-sensitive uncertainty, and hidden adversarial structure may provide a meaningful conceptual foundation for future AI-native cybersecurity systems.

The present work emerges from a broader research trajectory investigating operational completeness, hidden admissible structure, entropy-sensitive distinguishability, and constrained observability in adversarial operational systems. While these investigations originated within abstract operational and information-theoretic settings, the present paper explores whether analogous limitations of observability and operational sufficiency may emerge within AI-native cybersecurity environments characterized by autonomous reasoning systems, incomplete observability, and adaptive adversarial behavior.

2. Conceptual Origins and Operational Completeness

The ideas explored in this paper did not emerge from conventional cybersecurity research alone. They originate from a broader investigation into operational completeness, hidden admissible structure, distinguishability, and entropy-sensitive reasoning developed in adversarial operational settings.

Recent research conducted by the author examined whether observable operational behavior is sufficient to characterize the full admissible structure of an adversarial system. A central insight emerging from this work is that observable consistency alone may fail to guarantee operational completeness. Hidden operational sectors may remain structurally admissible despite apparently stable or self-consistent observable behavior.

While these investigations were originally developed within abstract operational and information-theoretic settings, they raise broader questions extending far beyond their original mathematical context.

> Can observable behavior alone ever fully characterize the operational integrity of an adversarial system?

This question becomes deeply relevant in modern AI-native cybersecurity environments.

3. Operational Consequences of Incomplete Observability

If the concerns explored in this paper hold even partially true, then the implications for AI-native cybersecurity may extend beyond conventional questions of detection accuracy or telemetry coverage.

Modern defensive systems increasingly operate under the assumption that sufficiently large observable operational surfaces eventually converge toward sufficiently reliable operational understanding. Security telemetry, behavioral analytics, identity monitoring, anomaly detection, and automated orchestration systems collectively attempt to construct a coherent representation of organizational cyber risk.

Yet it is not obvious that observable operational consistency necessarily implies operational completeness.

A system may appear operationally stable while important forms of cyber exposure remain structurally absent from the reasoning model itself. In practical environments, this may emerge through unmanaged identities, shadow SaaS integrations, latent trust relationships, hidden dependencies, or operational sectors that defensive telemetry never meaningfully captures in the first place.

This problem becomes more significant in AI-native defensive environments where autonomous systems increasingly participate directly in operational reasoning. AI systems necessarily inherit the limitations of the operational surfaces through which they observe the environment. If those surfaces are incomplete, manipulated, or structurally insufficient, then the resulting operational understanding may also remain incomplete even when observable behavior appears internally consistent.

The issue therefore may not simply be one of insufficient data volume.

The deeper difficulty may concern whether adversarial environments can ever be fully exhausted by their observable operational states alone.

This raises broader questions concerning the future of AI-driven cyber risk assessment.

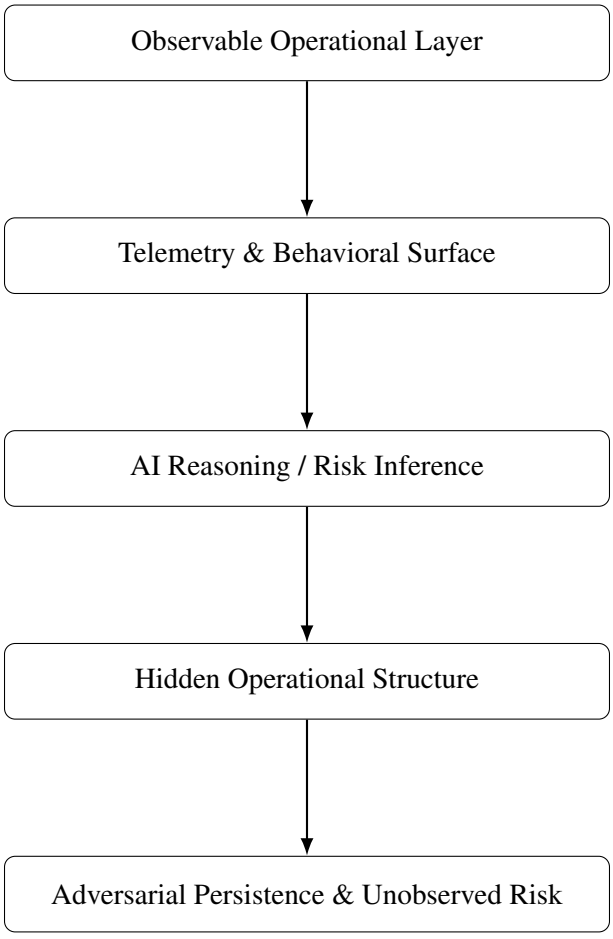
Future defensive systems may increasingly require the ability not only to classify observable threats, but to reason about hidden operational exposure, incomplete observability, and the possibility that structurally relevant operational states remain outside existing defensive inference mechanisms altogether.

In this sense, AI-native cybersecurity may gradually become less a problem of isolated threat detection and more a problem of uncertainty-aware operational inference under adversarial conditions.

Fortisec Research investigates whether future defensive architectures may require new forms of operational inference capable of identifying hidden exposure pathways and structurally unobserved forms of cyber risk beyond conventional observable telemetry.

4. A Minimal Operational Model

The relationship between observable telemetry, AI-native defensive reasoning, and hidden operational structure may be conceptualized as follows:



The framework explored in this paper suggests that observable operational behavior may not fully characterize the admissible operational structure of adversarial systems.

AI-native defensive systems necessarily reason through projected operational surfaces constructed from telemetry, behavioral inference, identity relationships, and observable operational signals. If those surfaces remain incomplete, adversarially manipulated, or structurally insufficient, then operationally relevant states may persist outside the reasoning model itself despite apparently stable observable behavior.

In this sense, the problem may extend beyond simple telemetry insufficiency alone. The deeper difficulty concerns whether adversarial operational environments can ever be fully exhausted by their observable operational representations.

5. Conclusion

The framework explored in this paper should be interpreted as a foundational investigation into the limits of operational observability in AI-native cybersecurity systems.

The central concern is not whether modern defensive architectures can continue improving detection accuracy, telemetry coverage, or automated analysis capabilities. Rather, the deeper question concerns whether observable operational behavior alone can ever fully characterize the admissible operational structure of adversarial environments.

As cybersecurity systems become increasingly dependent upon AI-assisted reasoning, behavioral classification, probabilistic inference, and autonomous operational decision-making, the limitations of observable operational surfaces themselves may become increasingly important.

AI systems necessarily inherit the constraints of the environments through which they observe and reason about operational state. If those operational surfaces remain incomplete, adversarially manipulated, or structurally insufficient, then the resulting operational understanding may also remain fundamentally incomplete despite apparently stable observable behavior.

In this sense, future cyber risk assessment may increasingly require defensive architectures capable not only of identifying observable threats, but of reasoning under incomplete observability and uncertainty itself. This may become increasingly important as organizations continue shifting toward highly distributed, identity-centric, AI-assisted operational environments whose risk surfaces are not always fully observable through conventional defensive telemetry alone. The present work should therefore be viewed not as a finalized cybersecurity framework, but as part of a broader research direction concerned with operational completeness, uncertainty-aware defensive reasoning, hidden operational structure, and the limits of observable risk in AI-native cybersecurity systems.

Fortisec Research investigates whether these questions may contribute toward the development of future AI-native operational cyber risk intelligence systems capable of reasoning about hidden exposure pathways and structurally unobserved forms of cyber risk beyond conventional telemetry-driven defensive models.

Selected Research Context

The cited works are included as part of the broader research context motivating the operational questions explored in this paper, particularly those related to observability, operational sufficiency, uncertainty, and hidden operational structure in adversarial systems.

[1] Hassan Nasreddine. *Essential Duality and Maximal Non-signalling Extensions in Algebraic Quantum Field Theory*. arXiv:2605.00075.

[2] Hassan Nasreddine. *Entropy Moduli and Support-Sensitive BKM Coercivity for Rank-Deficient Non-Commutative Markov Semigroups*. arXiv:2604.16616.

[3] Hassan Nasreddine. *Midpoint BKM Estimates and Boundary Coherence*. arXiv:2605.11024.

[4] Hassan Nasreddine. *Quadratic Stability of Entropy Minimizers under Block-Separable Convex Constraints*. arXiv:2512.16192.