Industry White Book

AI-Driven Investigation Framework Based on Behaviour Chain Semantics

Let Ai Investigate, Let Human Decide

June 2025

Author: Qimin Zhao

Version: Public Review V1.1

Date: October 30, 2025

Patent Disclosure and Confidentiality Notice

This white paper is released for **public reference only.**

Any form of technical adoption, integration, or application requires prior written authorization from the patent holder.

The invention was formally filed as a patent application with the China National Intellectual Property Administration (CNIPA) and the United States Patent and Trademark Office (USPTO) on June 19, 2025, with a simultaneous international filing under the Patent Cooperation Treaty (PCT). The filing date constitutes the priority date, protected **under PCT international novelty provisions.**

Public citation is permitted only with proper attribution to the original source.

For industry standards organizations, government agencies, enterprises, or research institutions seeking to obtain the detailed technical version, **please contact the author directly for authorized access and further collaboration**.

Public Contents

I.Executive Summary

The proposed model, named the **Semantic Chain Security Model (SCSM)**, integrates:

1.Natural language understanding (LLM/NLP)

2.Structured behavior chain modeling

3.AI-based semantic reasoning

4.Micro-model anomaly scoring

5.Human-machine collaboration

6.Knowledge feedback and evolution mechanisms

This model transforms isolated logs into structured behavior chains, allowing AI to interpret and investigate attack paths in ways previously only possible for experienced human analysts.

At the core of the model is the belief that AI cannot investigate what it cannot understand—and without structured behavior sequences, AI lacks the semantic substrate required for reasoning.

The Key Innovations of this model Include:

1.Natural Language Interfaces

Analysts use English, Chinese, or German to initiate investigations—no more SPL or KQL.

2.Democratized Threat Investigation

Anyone who can describe a suspicious behavior can launch an AI-powered inquiry.

3.Behavior Chain Construction

Logs are chronologically and semantically transformed into coherent behavioral narratives for each actor (IP, host, user).

4.Semantic Reasoning Engine

AI doesn't just match patterns—it infers intent, detects anomalies in context, and identifies pivotal transitions.

5.Pivot Point Detection (BPP)

Novel concept of pivot strength identifies critical semantic transitions (e.g., from reconnaissance to exploitation), even when actions appear “normal.”

6.Expert Feedback + Knowledge Write-Back

Human-validated behavior chains are written back into a dynamic knowledge base, enabling memory, comparison, and cumulative learning.

7.Micro-Level Anomaly Models

Pluggable, role-specific models analyze single logs for character usage, API paths, time-of-day anomalies, and more.

8.Three-Field Mapping for Accurate Attribution

Introduces external_ip → srcip → dstip mapping to resolve actor ambiguity in edge device logs (e.g., WAF).

9.Investigative Layer Standardization Proposal

Recommends the formal inclusion of an "Investigative Layer" into security standards (SIEM/SOC/XDR), bridging the gap between alert generation and incident response.

For use authorization, contact QMZ.

II.Problem Statement and Problem Definition

A.Structural Gaps in the Current Industry Framework

1.Existing Security Operations Investigation Architecture:

Mainstream security platforms (such as SIEM, XDR, EDR) generally follow the operational workflow:

Log Collection → Anomaly **Detection** (via Rules/Models) → **Alert** Generation → **Manual Response**

2.Core Issues:

An alert is **not an investigation** result—it is merely a detection outcome.

A response is not an analytical judgment—it is often a default action.

The system lacks the ability to connect behavior fragments into a coherent "attack narrative"—analysts are forced to manually piece everything together.

B.The Missing Investigative Layer: A Semantic Blind Spot in the Industry

Table1：Global Standardization Landscape

Security Domain	Standardization Status	Representative Standards
Prevention	Mature / Established	NIST CSF, ISO/IEC 27001
Detection	Mature / Established	MITRE ATT&CK, NIST SP 800-94
Investigation	Absent — No globally unified architectural standard	(Gap — no established global framework)
Response	Partially Standardized	NIST SP 800-61 (Computer Security Incident Handling Guide)

1.No Behavior Chain Structure → Logs Are Fragmented

Each log entry is treated as an isolated "point"—there is no structural connection to form a "line" or "chain".

There is no structured perspective to answer:

Who performed the behavior? When did it happen? What actions were taken?

2.No Structural Language → No Standardized Modeling

For release authorization, contact QMZ.

Without an Entity Behavior Database (EBD), it is impossible to model behavior around a specific actor.

Behavior coordinates—such as which step, what stage, or is this a critical point—are entirely absent.

3.No Attack Path Mapping → No Stage-Based Intent Reasoning

Frameworks like MITRE ATT&CK define tactical stages of attacks, but the industry lacks mechanisms to map raw behavior sequences to these stages.

Can we determine if an attack has escalated? Or if it’s preparing for outbound communication?

Without structure, AI cannot make such judgments—it can only passively assign scores.

C.Semantic Disconnection: Alert ≠ Explanation, Detection ≠ Reasoning

Current Reality:

Detection systems are good at identifying anomalous behaviors,

but true investigation requires understanding the behavioral path, intent at each stage, and the actor behind the actions.

The absence of an investigative layer leads to semantic misinterpretation:

Table2：Traditional Detection Conclusion VS SCSM Semantic Interpretation

Example Behavior Chain	Traditional Detection Conclusion	SCSM Semantic Interpretation
Failed login ×50 → Successful login	Successful login is treated as “normal”	Successful login is a pivot point, indicating a shift in attack stage
3 failed logins → Brute-force attack detected	Normal fluctuation in login outcome is flagged as “suspicious”	Normal behavior by the same actor is misclassified as an attack

SCSM is not about detecting anomalies — it’s about reconstructing the attack story.

Table3：Seven Consequences of a Missing Investigative Layer

Missing Structural Element	Immediate Symptom	Systemic Impact
No Behavior Chain	Alerts are fragmented	No context; AI cannot reconstruct a coherent "story"
No Role-Based Path Mapping	Actor confusion across logs	Cannot determine if actions belong to the same entity
No Stage Coordinates	Logs lack positional reference	AI cannot judge whether behavior belongs to an attack phase
No Pivot Point Recognition	Phase transitions are vague	Attack escalations blend with normal activity
Missing Structural Element	Immediate Symptom	Systemic Impact
No Semantic Query Interface	No semantic search	Analysts must rely on handwritten

For use authorization, contact QMZ.

	capability	KQL/SPL queries
No Expert Feedback Mechanism	Human insight is lost	Judgments cannot be structured and written back for reuse
No Knowledge Base Growth Mechanism	AI has no "memory"	Cannot reuse or evolve historical investigative structures

Let’s take another look at three major unsolved challenges in today’s cyber security landscape &Industry Exploration Trends :

Table4：Three major unsolved challenges

No.	Threat Type / Problem	Current Limitation	Root Cause Analysis	Reasoning
1	Zero days False Positives / Missed Alerts / Alert Overload	Single-rule decisions lack context, easily bypassed or become ineffective	No behavior chain structure; rules operate without semantic context	SCSM does not rely on known signatures or predefined indicators. Instead, it identifies abnormal behavior chain structures and intent transition signals by reasoning over behavior chain paths and attack phases—even when the individual actions may appear normal, as is often the case in 0-day attacks.
2	APT (Advanced Persistent Threats)	Log fragmentation, no context, no chronological action reconstruction	No unified behavior chain per entity; semantic gaps remain	SCSM constructs role-centric temporal chains through the **Entity Behavior Database (EBD),** ensuring that the attack path remains continuous even across devices and systems.
3	DDoS (Distributed Denial of Service)	Knowledge can't accumulate; no reusable templates for identifying patterns across actors	No knowledge feedback or generalization mechanism; behaviors remain isolated	SCSM does not determine DDoS attacks based on access frequency alone; instead, it evaluates behavior chain similarity and overlapping multi-source activity paths to identify coordinated role-based attack behaviors.

While some vendors claim to offer "investigation" capabilities, Actually， their systems typically only aggregate or correlate data after alerts have been triggered. In the SCSM (Semantic Chain Security Model) framework, **true investigation requires**:

1.Reconstruction of behavioral sequences — not isolated or single-point queries

2.Role continuity across actions — not raw IP matching

3.Path-based semantics and event coordinates — not flat or surface-level correlations

4.Reasoning over tactical phase transitions — not rigid rule-based triggers

5.Feedback-integrated memory and structured knowledge base

Therefore, these so-called “investigation” features in commercial platforms do not constitute semantic chain investigation. They remain alert-centric, descriptive, and superficial, lacking the structural layer that should exist between detection and response.

For use authorization, contact QMZ.

For use authorization, contact QMZ.

Table5：Global Vendor Gap vs SCSM

Vendor	Investigation Capability Claim	Actual Layer Coverage	Semantic Behavior Chain Modeling	Pivot Strength / Phase Reasoning	Expert Feedback + Knowledge Memory	Comments
Google Chronicle	Natural language search + threat hunting	Alert Correlation Layer	No	No	No	Primarily enhances Alert Searchability
Microsoft Sentinel	KQL + Workbooks + Hunting Queries	Alert Correlation Layer	No	No	No	Focused on Alert Enrichment & Hunting
Splunk Security Suite	SPL-based correlation & dashboards	Alert/Detection Layer	No	No	No	Powerful detection but lacks Behavior Chain semantics
Crowdstrike Falcon	EDR + Threat Graph	EDR-level Process Tracking	Partial (Process Chains)	No	No	Strong EDR focus, lacks cross-system Behavior Chain and semantic investigation
Palo Alto Cortex XDR	Analytics + Playbooks	XDR Alert Handling	Partial (limited path correlation)	No	No	Emphasizes playbook-driven response, no semantic layer
SCSM	Semantic Chain Modeling + AI Reasoning	Dedicated Investigation Layer	Full Behavior Chain Modeling	Yes,Pivot Strength + Phase Reasoning	Yes,Expert Feedback + Knowledge Memory	Fills global architecture gap between Detection and Response

In Summary：

Whether it’s the gradual failure of traditional rule-based systems against novel attack scenarios, or the current industry's fragmented attempts to integrate AI, both trends point to a fundamental root cause: the absence of an investigative layer and structured behavior chains.

For use authorization, contact ...MZ.

III.Solution Theory&Solution Framework

Table6：Industry Paradigm Evolution for cyber security

Evolution Stage	Paradigm Innovation	Representative Technologies/Products	Industry Architectural Layer	SCSM Contribution	Impact
First Stage (1990s)	Perimeter Security (Firewall)	Cisco PIX, Checkpoint FW	Perimeter Layer	—	Defined boundary security
SecondStage (2000s)	Real-time Intrusion Detection (IDS/IPS)	Snort, Suricata, Bro/Zeek	Detection Layer	—	Introduced real-time detection
Third Stage (2005–2015)	Centralized Log Analysis (SIEM)	Splunk, ArcSight	Detection → Alert Layer	—	Enabled cross-device correlation
Fourth Stage (2015–2020)	Tactical Phase Modeling (MITRE ATT&CK)	MITRE ATT&CK Framework	Detection → Attack Understanding	—	—
Fifth Stage (2017–Present)	Automated Response (SOAR)	Cortex XSOAR, DFLabs	Response Layer	—	Enabled response automation
Sixth Stage (2025.6)	**Investigation Layer**	**SCSM**	New Investigation Layer	First structured investigation layer definition	Bridges Detection → response gap
Seventh Stage (2025.6)	**Knowledge Loop / AI-Evolving Layer**	**SCSM**	New Knowledge Memory Layer	First expert feedback + AI learning loop	Establishes AI-driven investigative paradigm

Over the past three decades, the cybersecurity field has undergone multiple paradigm shifts — each introducing a new architectural layer: perimeter defense, real-time detection, centralized analysis, tactical phase modeling, and automated response.

Yet **two critical architectural gaps remain**: the absence of a native Investigation Layer and a structured, AI-driven Knowledge Memory Layer.The Semantic Chain Security Model (SCSM) addresses **both gaps simultaneously** — defining the world’s **first structured Investigation Layer**, and **introducing an AI-evolving Knowledge Loop** that enables dynamic learning and memory within security operations.This marks not one, but two paradigm breakthroughs in cybersecurity architecture — moving beyond traditional detection and response toward **AI-driven investigation** and **knowledge-based adaptive security.**

A.Solution Theory

Original Semantic Constructs Introduced for AI-Driven Security Investigation Architecture.As the original author, I hereby introduce the following core semantic constructs into the AI-driven security incident investigation system. These foundational concepts are designed to support behavior chain modeling, path inference, human-AI consensus, and knowledge base evolution through write-back mechanisms.

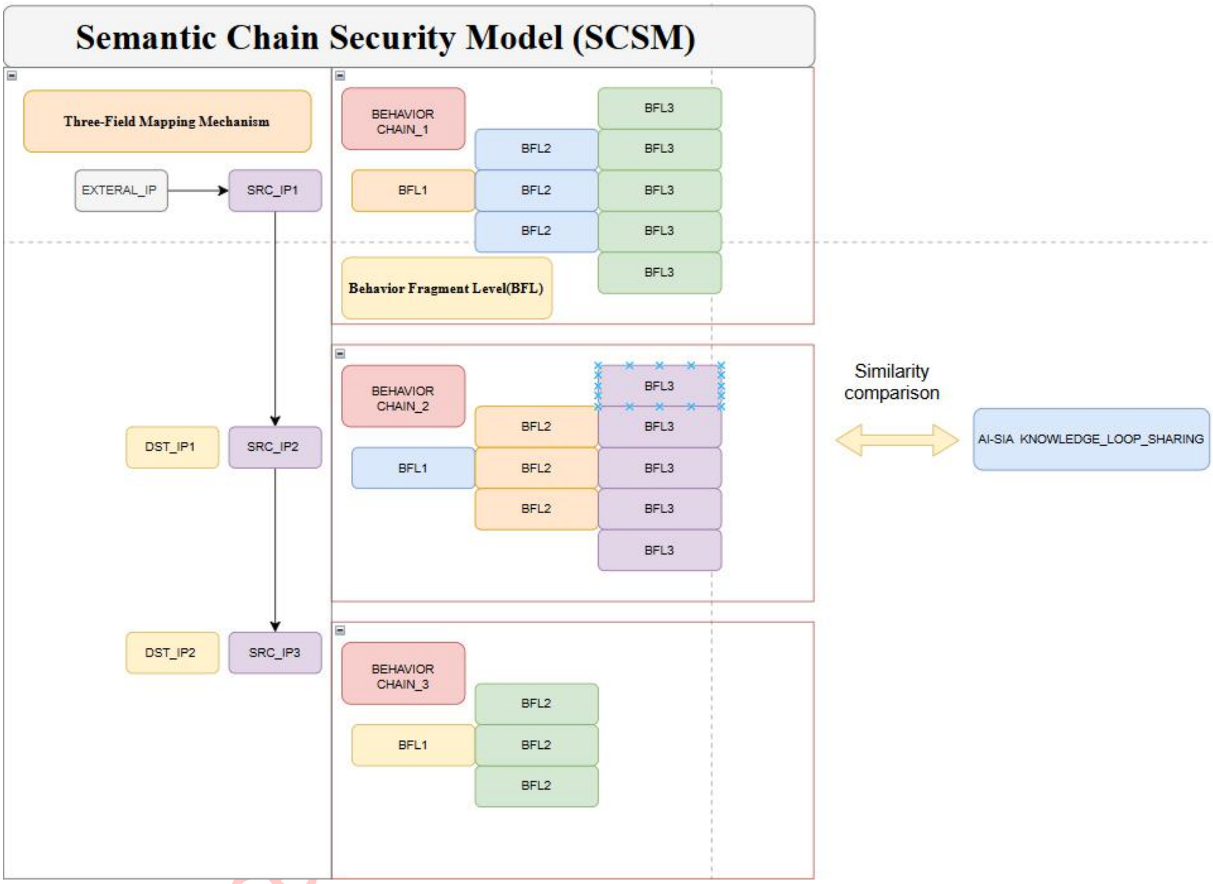

Figure1：Semantic Chain Security Model

1.Semantic Chain Security Model (SCSM)

The Semantic Chain Security Model (SCSM) is an AI-driven security operations framework centered on behavioral semantics. It establishes a full-cycle investigation process that spans from log ingestion, behavior reconstruction, and semantic inference to expert validation and knowledge write-back.

SCSM breaks away from the traditional SOC paradigm built on rule-based detection and manual log querying. By introducing a structural semantic layer, it empowers AI to understand behavior—enabling proactive risk assessment, attack path recognition, and continuous self-evolution.

2.Character/Entity Behavior Database (EBD)

The Entity Behavior Database (EBD) is a structured, role-centric data repository that chronologically records and organizes all observable actions performed by a specific entity—such as a user, IP address, host, or device.

It provides the foundational structure for behavior reconstruction and actor-centric semantic reasoning in security investigations.

3.Knowledge Write-Back Mechanism

The Knowledge Write-Back Mechanism refers to the process by which human security experts review, validate, and optionally revise the AI-generated behavior chains and risk evaluations.

This mechanism is the cornerstone of human-AI collaboration. It ensures that expert judgments are not lost but instead structurally recorded and fed back into the system's evolving knowledge base, thereby improving semantic inference over time.

4.Three-Field Mapping Mechanism

The Three-Field Mapping Mechanism addresses semantic ambiguities in traditional binary log field mappings (e.g., srcip, dstip), particularly in edge device logs (e.g., WAF, proxy).

It introduces a third semantic field, external_ip, to explicitly represent the origin of the access request. This results in a clarified behavioral structure:

External_ip → srcip → dstip

enabling accurate actor attribution and path-level modeling in a three-stage behavior chain format.

ation, contact QMZ.

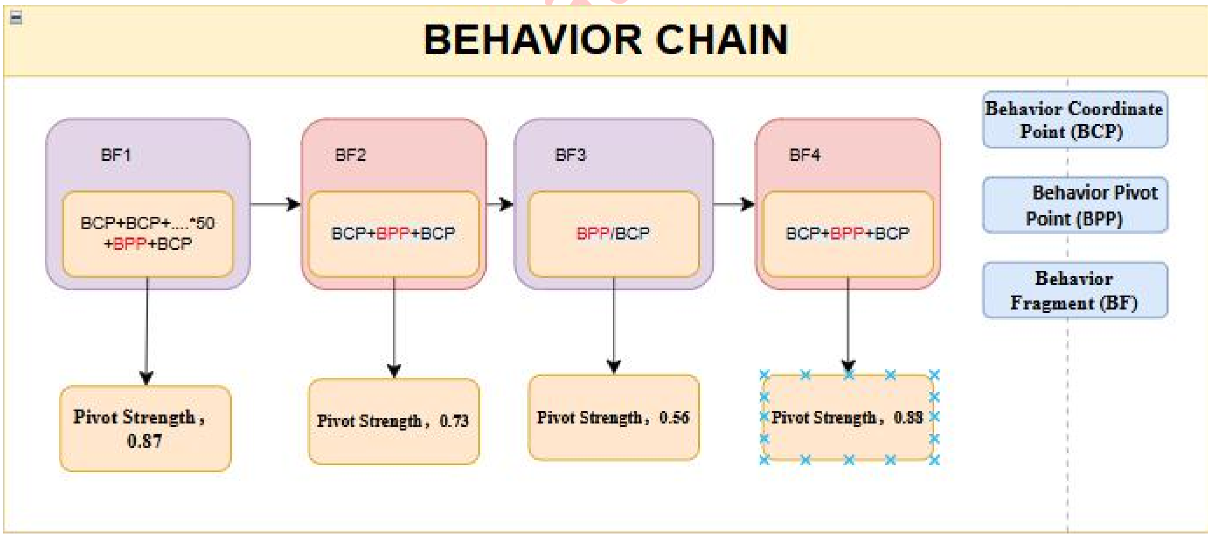

Figure2：Behavior Chain

5.Behavior Chain / User Behavior Chain (BC / UBC)

A Behavior Chain is a time-ordered, semantically consistent sequence of behavioral events linked to a single entity (e.g., user, IP, host, or device) within a defined time window. It integrates both normal and abnormal behaviors and provides the structural foundation for AI to perform reasoning, identify anomalous stages, and infer malicious intent. It is the precondition for building an attack chain.

6.Behavior Chain Path (BP)

A Behavior Chain Path (BP) represents the semantic linkage of behavior **chains across different hosts** or entities, describing how an intrusion or action propagates between them.When tracking activities across systems — for example, **from Host A to Host B** — the sequential transition A → B forms a Behavior Chain Path.This cross-host reconstruction enables end-to-end tracing of attacker behavior.Each BP is composed of multiple Behavior Chains (BCs) connected through source – destination relationships identified by the Three-Field Mapping Mechanism (e.g., SRC_IP, DST_IP).

Behavior Fragment (BF) Structural Components:

Behavior States: Different outcomes or phases of the same behavior (e.g., success, failure).

Behavior Coordinate Point (BCP): An individual state event in the process (e.g., one failed login).

Behavior Pivot Point (BPP): A special type of BCP representing a decisive transition to another phase (e.g., multiple failures followed by a final success).

Conditions: Contextual rules for evaluating the fragment (e.g., time window, frequency thresholds, file type).

Table1：Structural Variables Usable as Trigger Conditions

Category	Description	Example
Time-based	Specific time windows or event intervals	1:00–5:00 AM; high-frequency behavior within 10 minutes
Count / Threshold	Total number of BCPs; existence of a BPP	More than 100 failed logins AND at least one successful login
Category	Description	Example
Keyword / Regex	Presence of sensitive commands, paths, or indicators in logs	curl, /tmp/myscript.sh, wget,port
Log Type Ratio	Proportion of specific log types among all BCPs	60% from endpoint logs, 40% from network logs
Logical Formula	Boolean expression to define abnormal behavior patterns	(BPP = True) AND (BCP > 10)
AI Micro-Model	AI-based scoring and/or tag evaluation	Score > 0.75 AND tag = "Suspicious behavior"

For use authorization contact OMZ.

Behavior Fragment (BF) Key Features:

Captures the **entire lifecycle of a behavior**, not just isolated events.

Can be benign or malicious, depending on the combination of states.

Serves as the fundamental analytic unit for mapping behaviors to the attack lifecycle.

Examples of Behavior Fragments

Example 1: Login Behavior (BF) — Abnormal Brute-Force Behavior

States included:

Multiple failed login attempts (BCPs)

One successful login (BPP)

Condition: ⩾ 50 failed attempts within 24 hours

Interpretation: A complete login behavior fragment → successful brute-force attempt.

Example 2: Login Behavior (BF) — Normal Login Behavior

States included:

A few failed login attempts (BCPs)

One successful login (BPP)

Condition: < 10 failures within 1 hour

Interpretation: A complete login behavior fragment → normal login activity.

Behavior Fragment Level (BFL)

Behavior Fragment Level (BFL) is a semantic classification system that assigns each **Malicious Behavior Fragment** (**MBF**) to a distinct phase in the typical attack lifecycle.

Each level reflects a different stage of adversarial activity, semantic severity, and corresponding MITRE ATT&CK tactics.

It serves as a bridge between micro-level behavior detection and macro-level kill-chain reasoning, enabling:

1.Progressive threat scoring

2.AI reasoning sequence validation

3.Visualization of attack stage distribution

Level 1 (Early Stage / Initial Access & Discovery) Representative BF: **Login (Brute-Force)**

MITRE ATT&CK Mapping: Initial Access, Credential Access

Level 2 (Mid Stage / Execution, Persistence, Lateral Movement) Representative BF: **File Upload**

MITRE ATT&CK Mapping: Execution, Persistence, Lateral Movement

Level 3 (Late Stage / Command & Control, Exfiltration) Representative BF: **Command Execution**

MITRE ATT&CK Mapping: Command & Control, Exfiltration, Impact

For use authorization, contact OMZ.

Table2：Semantic Phase Mapping Between Behavior Fragment Levels (BFL) and MITRE ATT&CK Framework

BFL Semantic Phase	Definition	Corresponding MITRE ATT&CK Tactic(s)
Level 1	Also known as Recon Phase. This level describes the process from external probing to successful access. For example, brute-force attempts followed by successful login.	*Initial Access*, *Credential Access*, *Discovery*, *Reconnaissance*
Level 2	Represents the stage where the attacker leverages obtained access to execute tools, scripts, or malicious payloads within the environment.	*Execution*, *Privilege Escalation*, *Defense Evasion*, *Persistence (early)*
Level 3	After tool execution, multiple paths may emerge including lateral movement, credential theft, communication with C2 servers, or data exfiltration.	*Lateral Movement*, *Collection*, *Command & Control*, *Exfiltration*, *Impact*, *Persistence (sustained)*

The following three structural elements form the foundational prerequisites for enabling AI-driven knowledge base feedback:

1.A behavior chain must exist

Providing the structural backbone that organizes actions chronologically and semantically around a specific actor (user, IP, host, etc.).

2.Each behavior must have a semantic coordinate

Marking its position and meaning within the overall context (e.g., login attempt, file upload, lateral movement), enabling role-aware interpretation.

3.AI must identify pivot points

Detecting critical transitions that signify a shift in the attack phase, such as a successful login following multiple failures.

Only when these conditions are met can a human analyst validate the event severity based on the reconstructed path and perform accurate knowledge write-back.This is not just a technical process—it is the semantic essence of the system. These structural anchors are what allow the AI to truly "understand the attack story" and evolve from detection to reasoning.

(Note:Pivot Point Identification: From Binary Judgment to Weighted Confidence)

4.Pivot Strength

Pivot Strength refers to the semantic confidence score assigned to a specific behavior node when it is identified as a pivot point within an attack behavior chain. It reflects the logical support for judging whether the action represents a tactical phase transition—such as a shift from reconnaissance to execution—and serves as a key signal indicator in AI inference pathways.

In semantic investigation, whether a behavior constitutes a pivot point should not be treated as

For use authorization, contact MZ.

a binary decision (yes/no), but rather as a probabilistic weight known as Pivot Strength.

Three Key Decay Factors Impact **Pivot Strength**:

1.Time interval since the most recent failure — the longer the interval, the greater the decay (no ×1 multiplier applied).

2.Number of previous failed actions (BCPs) compared with the predefined malicious behavior fragment threshold — for example, the preset brute-force limit is 200 attempts, but the current attack contains 500 attempts, indicating a larger deviation.

3.Average time interval between consecutive failures, reflecting the attack's persistence and intensity.

B.Solution Framework

Table8：Five-Layer Foundation

Level	Name	Function Description
1	**Field Normalization Layer**	Standardizes heterogeneous log formats from multiple sources; constructs key fields (e.g., srcip, dstip, eventid, source_from, external_ip)
2	**Micro-Scoring Layer**	Applies multiple AI models to each log entry to generate threat score fields (model1 to model4) for downstream semantic computation
3	**Behavior Chain Modeling Layer**	Reconstructs behavior chains for the same entity based on temporal, semantic, and role consistency, forming inference material
4	**AI Semantic Reasoning Layer**	Performs path inference, attack chain recognition, and pivot point detection on behavior chains; supports natural language interface via LLM
5	**Expert Feedback & Knowledge Loop**	Integrates human feedback into a knowledge base with behavior chains, pivot points, labels, and attack names, enhancing future reasoning capabilities

For use authorization only QMZ.

For use authorization, contact QMZ.

IV.Use Cases

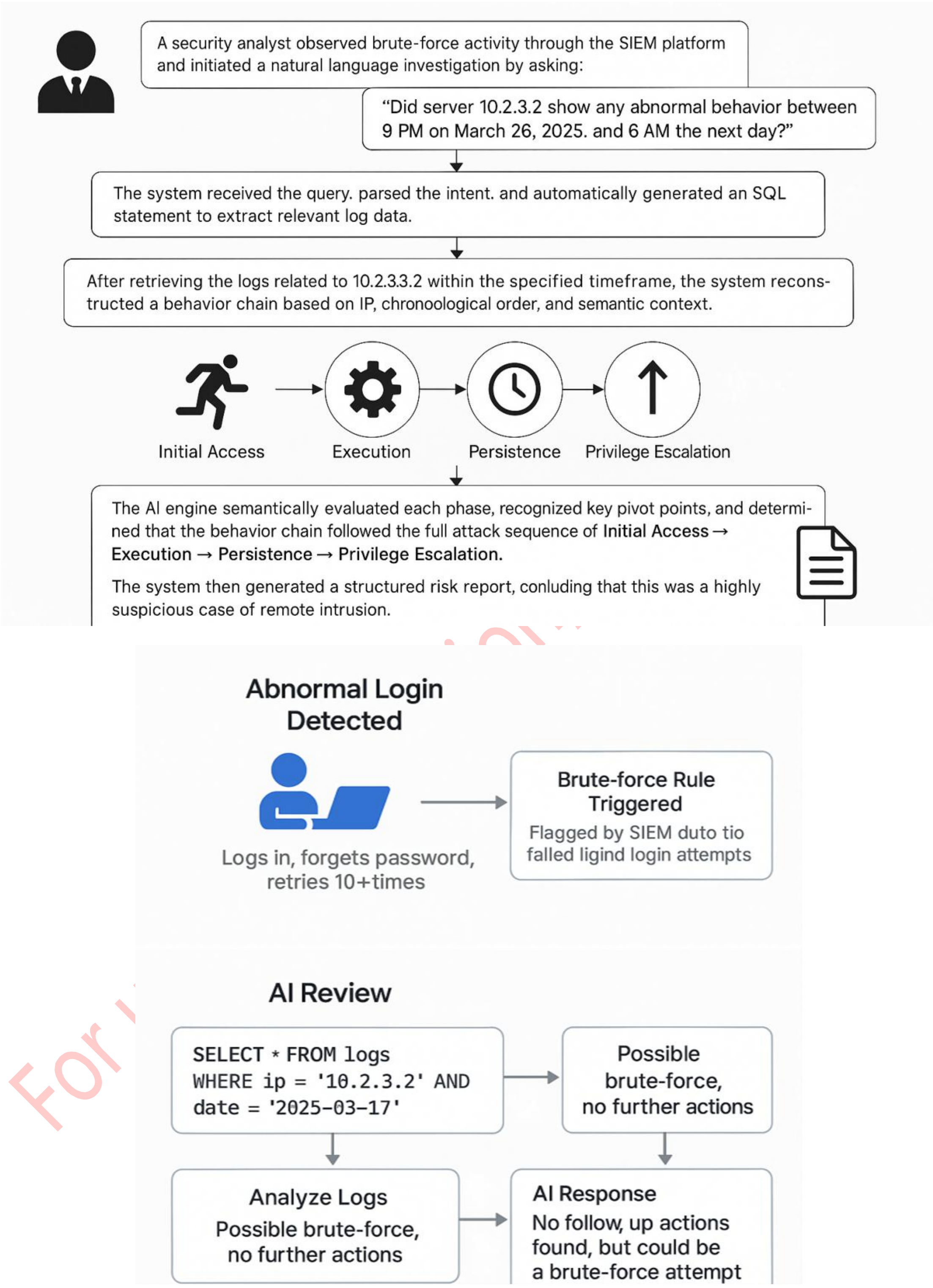

Figure8:AI-Driven Behavior Chain Analysis Flowchart(1.1)

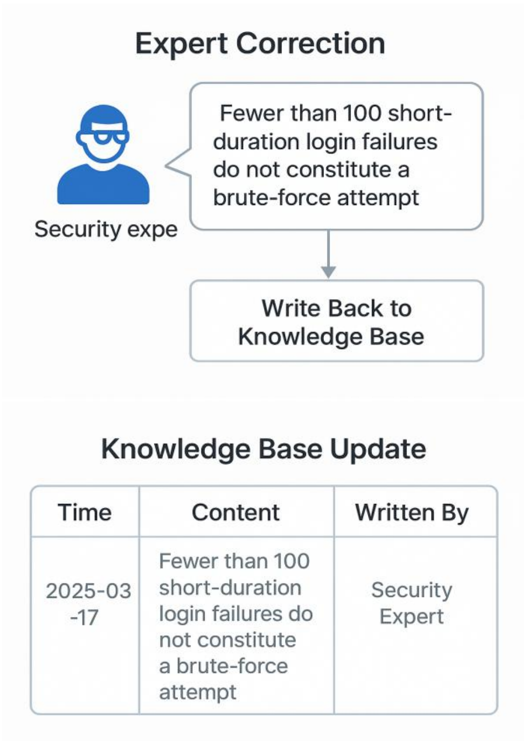

Figure9:Expert Feedback and Knowledge Base Update Diagram(1.2)

For use authorization, contact QMZ.

For use authorization, contact QMZ.

VI.Industry Impact

A.Impact

To address the lack of structured investigative semantics between "alert" and "response" in current SOC/SIEM architectures, it is recommended that the "Investigative Layer" be formally introduced as a distinct architectural component in industry standards. It mainly includes:

1.Inputs: Sequences of behavioral logs and alert events

2.Outputs: Structured behavior chains, attack path graphs, and AI-generated reasoning suggestions

3.Core Functions: Behavior modeling, pivot point identification, and expert knowledge feedback integration

4.Necessity: Enables the response layer to execute precise, explainable, and automated decisions

B.Systemic Value

1.Language replaces experience: Transitions from manual, experience-driven investigation to AI-driven semantic reasoning and interpretation.

2.Structure captures expertise: Uses behavior chains to structurally store human expertise and AI feedback

3.AI understands behavior: Enables AI to understand not just "logs" but actual "behaviors"

4.Adaptive to attack-defense evolution: Supports continuous evolution, knowledge accumulation, and reasoning over behavioral variants

5.Transforms traditional SIEM investigation: Fully replaces the manual "log piecing" paradigm—AI investigates, humans validate

For use authorization, contact QMZ.

Table10:Structured Solutions to Core Industry Challenges

Industry Challenge	Root Cause	Structured Solution
0day Attacks	0days fundamentally bypass fixed detection rules; in traditional defenses, rule bypass = full defense failure	Structural-semantic language does not rely on static rules. It identifies anomalies through behavior chain paths and AI semantic reasoning, making it resilient to rule bypasses
Pivot Point Detection	Traditional systems misclassify post-exploit "normal" behaviors as benign, failing to detect lateral movement stages	Based on the ATT&CK tactical framework, any "normal behavior" following a labeled attack stage is flagged as a pivot point. The system includes decay mechanisms and path-based scoring
Industry	**Root Cause**	**Structured Solution**

Challenge		
APT Attacks	Logs are fragmented and lack contextual continuity, making it impossible to reconstruct the full attack chain	Builds a Role-Based Behavior Database (EBD), storing logs by entity and sorting them chronologically. AI semantic analysis is applied to reconstruct the complete attack path
DDoS Attacks	Distributed, multi-source, concurrent access depletes resources. Traditional frequency-based thresholds are easily bypassed	Trains micro-models targeting specific DDoS features to detect patterns in path, frequency, and distribution. AI synthesizes these into behavior chains for organizational defense. Detection strategy includes: 1. Micro-models for fine-grained analysis, macro-logic for attack inference 2. Behavior chain knowledge base with expert feedback integration

Table11:SCSM vs Traditional SOC: Structural Innovation Comparison

Innovation Point	**Traditional SOC**	**SCSM Advantages**
Log Structure	Non-standard fields, inconsistent formats	Field semantic normalization with triple-field mapping (e.g., external_ip)
Detection Method	Rule-based matching, signature-dependent	AI-based micro-scoring, multi-dimensional models, self-evolving detection
Behavior Understanding	Evaluates each log independently	Constructs semantic chains via temporal and role-based behavior linking
Attack Recognition	Relies on static attack rules	Dynamically identifies attack chains and pivot points; supports variant inference
Expert Knowledge Retention	High loss of individual insights	Writes back into structured knowledge chains, enhancing system memory
LLM Integration	Absent	Supports full flow: natural language → SQL/semantic reasoning → structured output
Structural Layer Completion	Detection → Alert → Response (Investigation is manual)	Completes the gap: Detection → Alert → Investigation → Response

Table12:SCSM vs. Three Historical Protocol Standards

Protocol	**Initial Phase**	**Structural Release**	**Adopted by Standards Bodies**	**Global Ecosystem Expansion**
HTTP	Defined at CERN as URL structure	Adopted by W3C → Unified browser language	Referenced by all web protocols	Became the universal language of web communication
TLS	Netscape’s proprietary SSL	Extended by IETF as TLS standard	Adopted by all secure browser platforms	Became the de facto encrypted communication standard for the Internet
MITRE ATT&C	MITRE research	Released as an open tactical	Referenced by NIST, widely adopted by	Became the factual standard for threat detection and

K	project	structure model	security companies	behavior modeling
SCSM	Independently invented + patent-defined	Released via industry white paper + five supporting documents		

For use authorization, contact QMZ.

VII.Problem Solved

1.1 Security Operations Architecture Issue

Typical Process:

Log Collection → Detection (via Rules/Models) → (Human analysis) → Response

Core Problem:

The core task defined by the investigation layer is to Explain and reconstruct "**how events occur in the corporate environment**".Without the investigation layer, we can never explain "**what is the basis for response**".Investigation layer is neither a single event, nor an alert(log), nor a response(action), nor an intelligence technique(ioc,tactic). It output **Structured Evidence of investigative logic.(Ref. Page 25.Miner Script Json)**

1.2 Excessive False Positives in SOCs

Root Causes:

1.Rule-driven and event-centric: Each log entry is treated as an isolated event and analyzed statically.

2.No mechanism exists for identifying Behavior Pattern Points (BPP).

SCSM Approach:

SCSM reduces false positives not because it “uses AI” or “adds more rules,” but because it **fundamentally changes the unit of meaning** through which the entire security system interprets alerts.

Semantics-driven and role-centric: Each log is interpreted **as an action** performed by a subject (Behavior Chain Point – BCP, Behavior Pattern Point – BPP), enabling dynamic contextual analysis.

BPP recognition enables semantic transition points—linking isolated events into meaningful behavioral chains for reasoning and investigation.

Example:

Traditional model: 500 failed login attempts → 500 alerts

SCSM model: 500 failed logins + 1 success → 1 incident, with Behavior Fragment Level (BFL) = 1& Attck T1110

In **SCSM**, investigation is **not rule-based** but **behavior-segment-based**. For example, the “login” behavior exists in the local behavior repository in both **normal** and **abnormal** states (e.g., brute-force attack, see Px). In the **role-behavior database**, these 500 failed logins and one successful login are

For use authorization, contact Q

recognized as actions belonging to the same role.

When this entire **behavior segment** matches a **behavior fragment** in the local repository, it is classified as a single alert.

In summary:

Traditional model: 500 failed login attempts → 500 alerts

SCSM model: 500 failed logins + 1 success → **1 alert**, with **Behavior Fragment Level (BFL) = 1**

1.3 Advanced Persistent Threat (APT) attacks

Traditional SOC or SIEM systems can hardly *understand* long-chain APT attacks, because an APT is a **cross-stage, cross-device, and cross-time behavioral semantic chain**.

Each phase of an APT campaign appears in a traditional SIEM as a set of unrelated log entries.

The purpose of **SCSM** is to **reconstruct these fragmented logs into a coherent semantic behavior chain**.

Example:

Day 1: External IP performs port scanning

Day 5: Abnormal VPN login

Day 8: Internal host executes PowerShell

Day 10: Archive uploaded to an external address

A traditional SOC would treat these as **four isolated events**.

SCSM interprets them as **four behavioral segments of a single incident**, recognizing them as one **behavior chain**:

Reconnaissance (port-closed → BCP, port-open → BPP) →

Initial Access (login failed → BCP, login successful → BPP) →

Execution (PowerShell failed → BCP*(0-xx), PowerShell executed → BPP) →

Exfiltration (connection failed → BCP*(0-xx), connection successful → BPP)

The system links all actions of the **same role** (user / IP / host) according to **temporal** and **semantic continuity**, constructing a complete **Behavior Chain**.

SCSM is designed to solve this problem:

SCSM defines the **smallest logical unit of a security event** not as a *log entry*, but as a **Behavior Chain**.

1.4 0DAY attacks

Zero-Day Attacks and Unknown Vulnerabilities

Root-Cause Analysis：

Zero-day attacks exploit vulnerabilities that are not yet known to vendors or security communities.

For use authorization, contact OMZ.

Because there is no existing signature, IOC, or patch, traditional detection and response pipelines cannot recognize them in time.

Problem:

Rule-based systems rely on prior knowledge; when facing a 0-day exploit, they generate no alert or produce ambiguous anomalies.

SCSM Detection Philosophy: Semantic-Driven, Not Signature-Driven

SCSM detects threats **based on semantics rather than signatures**.

Its detection logic does not ask whether *a single event matches a vulnerability rule*; instead, it determines whether a **sequence of events forms a causal and self-contained behavior chain**.

Even when an attacker gains system access through a **zero-day exploit**, as long as subsequent actions occur, the **Entity Behavior Database (EBD)** will reveal a clear and traceable **attack path represented as a behavior chain**.

1.5 In Traditional SOCs, BPP Identification Relies Entirely on Human Expertise

In conventional SOC operations, the identification of Behavior Pattern Points (BPPs) is not formally defined within the system. It exists only in the analyst ’ s mind, relying on individual experience and intuition rather than a structured or machine-recognizable coordinate.

For use authorization, contact OMZ.