说 明 书

一种语义行为链重构与 AI 引导的安全调查方法与系统

技术领域

本发明涉及网络安全与人工智能交叉领域，尤其涉及一种语义行为链重构与 AI 引导的安全调查方法与系统。

背景技术

现有的安全运营平台主要依赖 SIEM 系统进行日志汇总与事件分析 其核心模式通常基于固定规则匹配单条日志以生成告警 在实际调查过程中，分析人员需手动编写查询语句 从多源日志中定位关键信息并人工拼接行为链以推断威胁路径 此方法存在语法门槛高 日志割裂 信息遗漏 响应滞后等问题，同时，各类日志格式异构、字段不统一 ，安全设备各自为战，缺乏上下文关联能力，导致无法对攻击过程进行完整还原，并且系统告警严重依赖静态规则，对未知攻击与高隐蔽性行为缺乏检测能力，对于 APT ZERO DAY 等高级威胁往往无法提前识别和阻断。

此外，当前 SOC 调查依赖经验型人员完成行为还原与风险判断，一旦人员变动，知识与流程缺乏沉淀，调查质量易波动且缺乏可复用的知识结构，同时传统平台不支持通过自然语言提问的交互方式，分析人员仍需掌握特定语法或规则语言查询系统，数据访问与推理能力存在高门槛。

由此可以得出，现有安全运营平台普遍存在以下结构性缺陷：

一、缺乏跨设备跨维度统一主体的行为建模能力；

二、无法自动构建行为链并对角色行为做上下文解释；

三、缺少智能化问答式交互机制与语义输出能力；

四、无法将历史攻击知识沉淀形成动态可复用知识库；

五、无法摆脱对规则的依赖 对高级威胁响应能力弱。

发明内容

本发明的目的在于克服上述问题，提出了一种基于语义行为链重构的 AI 引导安全调查方法，定义为当前安全运营架构中缺失的调查层，首次实现跨日志、跨设备、跨角色的结构化调查流程与知识演进闭环。为实现上述目的，本发明采用如下技术方案：

一种语义行为链重构与 AI 引导的安全调查方法，包括以下步骤：

步骤 S11：采集来自多个安全设备或系统的日志数据，以统一行为主体建模为目的字段标准化；

步骤 S12：构建与实体相关联的行为链结构，所述行为链结构为时间序列、图结构、标签链式结构或其等效路径结构；

步骤 S13：接收用户输入的自然语言安全调查问题；

步骤 S14：利用人工智能引擎将自然语言问题映射至与行为链结构相关的查询目标；

步骤 S15：基于行为链结构对查询目标执行语义推理、路径搜索或上下文分析操作；

步骤 S16：输出结构化或非结构化的查询结果，包括路径图、节点列表、行为片段、风险提示、专家反馈信息或提示式结果。

进一步地，在步骤 S11 中，所述日志数据在采集过程中进行字段标准化、实体统一、平台对齐、标识聚合，进行跨平台、跨系统日志行为链整合。

进一步地，在步骤 S12 中，所述行为链结构基于日志时间顺序、操作关联、标签联动、因果依赖或知识图谱路径建立，支持行为节点的图形化展现与实时或持久化存储；所述行为链结构嵌入增强信息包括节点风险评分、结构路径权重、行为意图标签、异常置信度；跃迁点标识，所述跃迁点标识用于标记行为链中攻击阶段变化的关键节点；行为坐标，所述行为坐标记录每一行为在链路径中的阶段位置、角色归属、时间位次；阶段跃迁强度评分，所述阶段跃迁强度评分用于度量节点间跃迁的可信度，综合考虑时间连续性、角色一致性与行为模式偏移度；所述增强信息用于优化人工智能模型的路径排序、异常识别与阶段跃迁推理能力。

进一步地，在步骤 S12 中，所述行为链结构为预先构建的结构或在查询过程中基于用户指定的实体标识、时间范围或行为特征条件所动态聚合形成的事件片段集合；所述事件片段集合在语义层面具有关联性或逻辑顺序；所述行为链结构作为构建基础，触发结构化事件调查流程，包括特定 IP 或用户的行为链重建、相似攻击路径比对、行为归因、协同分析任务；行为链节点构成包括通信动作、登录行为、文件操作、配置变更、进程调用、URL 访问、API 请求或其它可识别的安全行为事件。

进一步地，在步骤 S14 中，所述人工智能引擎包括大语言模型、NLP，图神经网络、规则-学习混合引擎或其等效语义推理系统；所述人工智能引擎将自然语言问题转化为结构化查询模板、图结构遍历请求或 SQL 语句，并结合 Prompt 工程进行自然语言与结构化查询的高效映射，支撑语义推理过程。

进一步地，用户以自然语言方式进行连续提问，人工智能基于行为链上下文状态维持语义会话，支持上下文缩放、链结构焦点切换、逆向提问生成与交互式行为路径演化，并支持跨轮次上下文记忆与专家反馈驱动的对话演化能力。

进一步地，在步骤 S16 中，输出结果包括攻击路径图、结构节点列表、时间窗口内行为轨迹、匹配片段、风险说明、推理失败提示或需人工确认的半结构结果。

进一步地，包括字段归一化模块、行为链建模模块、AI 语义推理模块、微模型评分模块以及专家反馈与知识回写模块；所述字段归一化模块、行为链建模模块、AI 语义推理模块、微模型评分模块以及专家反馈与知识回写模块构成逻辑反馈闭环，支撑行为链构建、攻击路径重构与人工智能语义分析过程；

所述字段归一化模块对采集的多源日志数据进行字段标准化、实体归一化、平台对齐与关键字段映射；

所述行为链建模模块基于时间顺序、语义一致性与角色一致性构建行为链结构，作为 AI 推理与路径重建的输入基础；

所述 AI 语义推理模块对行为链进行路径推理、阶段跃迁识别与风险评估，支持自然语言接口调用；

所述微模型评分模块对标准化日志字段执行多维度模型评分，生成威胁评分字段，支撑行为链建模与 AI 推理；

所述专家反馈与知识回写模块将经人工专家确认的行为链与阶段标签反馈至知识库，形成可持续演进的 AI 知识记忆层。

进一步地，在微模型评分模块中，日志数据处理与评分分析方法对结构化日志字段进行外部分析处理并统一写入数据库，包括以下步骤：

步骤 S21：日志预处理，接收原始日志数据后，利用本地脚本程序、应用程序或等效模块执行字段提取、数据清洗及标准化操作；

步骤 S22：接口调用，将预处理后的日志字段以结构化格式通过 API 接口发送至一个或多个外部分析服务；

步骤 S23：服务处理，外部分析服务包括微观模型服务、评分引擎、特征提取器、标签分类器或基于规则的评估器，用于执行行为评估、威胁打分、字段统计、上下文解析的处理操作；

步骤 S24：结果接收与封装，统一接收分析服务返回的结构化结果并进行封装处理；

步骤 S25： 数据库写入，将处理后的结果数据写入至数据库的预定义字段，供行为链建模、攻击阶段识别、可视化呈现或知识库更新使用。

进一步地，在字段归一化模块中，基于三元字段映射对安全日志进行处理，包括以下步骤：

步骤 S31：采集来自安全设备的日志数据，提取原始字段 srcip 与 dstip；

步骤 S32：构建 external_ip → srcip → dstip 的三元结构，当日志来源为可直接接收公网访问的边界设备时，将原始 srcip 映射为 external_ip，将原始 dstip 映射为逻辑 srcip，并保留原始 srcip 字段内容不用于行为链建模。

进一步地，专家反馈与知识回写模块包括结构化行为链知识库，所述结构化行为链知识库将已验证的攻击行为链及其关联信息写入知识库，并基于知识库的自动语义判断与行为链相似性检索；

所述专家反馈与知识回写模块中，专家通过人工智能辅助调查流程对安全事件进行判定并反馈知识，手动触发比对流程，将当前待排查的行为链与知识库中已存的历史行为链进行相似性比对，辅助判断为相似攻击路径或阶段跃迁风险，所述比对流程独立于 AI 自动推理流程运行；

所述知识库包括行为链复现库、字段语义解析库、风险标签与意图库、图谱关系数据库、路径语义坐标库及其他优化信息知识库；所述路径语义坐标库存储行为链中各节点的结构位次、阶段标签、行为坐标、跃迁点标识与跃迁强度评分；

所述结构化行为链知识库利用结构向量、语义标签或图结构重叠度进行相似性匹配，基于历史行为链进行高置信度攻击路径比对与阶段跃迁推理，匹配结果作为人工智能模型输入或安全分析师辅助参考；

所述结构化行为链知识库与人工智能模型形成闭环交互机制，在人工智能模型的训练阶段、更新阶段

及推理阶段，所述人工智能模型动态依赖知识库中的结构化行为链数据、路径语义坐标、跃迁点标识与阶段跃迁强度评分信息。

本发明的优点在于：

本发明通过采集多源日志数据并进行字段标准化、实体统一等处理，构建基于时间序列、图结构等的行为链结构，嵌入节点风险评分、行为坐标等增强信息，实现了跨平台、跨系统的日志行为链整合与完整攻击路径还原，解决了日志割裂、信息遗漏问题，为安全调查提供统一数据基础。

本发明利用大语言模型等 AI 引擎将自然语言问题转化为结构化查询，支持连续提问的语义会话与交互式行为路径演化，实现了智能化问答式交互，降低了安全调查的语法门槛，使分析人员无需掌握特定语法即可高效获取攻击路径、风险提示等结果。

本发明通过专家反馈与知识回写模块将验证后的行为链及标签写入结构化知识库，结合微模型评分模块对日志多维度评分，实现了历史攻击知识的动态沉淀与复用，提升了对未知攻击和高级威胁的检测能力，解决了传统平台依赖静态规则、响应滞后的问题。

附图说明

构成本申请的一部分的附图用来提供对本申请的进一步理解，使得本申请的其它特征、目的和优点变得更明显。本申请的示意性实施例附图及其说明用于解释本申请，并不构成对本申请的不当限定。

在附图中：

图 1 为实施例 1 中一种语义行为链重构与 AI 引导的安全调查方法的闭环示意图。

图 2 为实施例 1 中一种语义行为链重构与 AI 引导的安全调查系统的模块交互图。

图 3 为实施例 1 中一种语义行为链重构与 AI 引导的安全调查系统中行为链构建与 AI 上下文分析的原理图。

图 4 为实施例 1 中一种语义行为链重构与 AI 引导的安全调查系统中专家确认与知识库回写的原理图。

图 5 为实施例 1 中一种语义行为链重构与 AI 引导的安全调查系统中微模型评分模块的原理图。

图 6 为实施例 1 中一种语义行为链重构与 AI 引导的安全调查系统中自然语言输入与输出反馈的示意图。

图 7 为实施例 1 中一种语义行为链重构与 AI 引导的安全调查系统中映射方式的对比图。

具体实施方式

为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。通常在此处附图中描述和示出的本发明实施例的组件可以以各种不同的配置来布置和设计。

下面通过具体实施例对本发明进行详细和具体的介绍，以使更好的理解本发明，但是下述实施例并不限定本发明的保护范围。

实施例 1

如图 1-7 所示，一种语义行为链重构与 AI 引导的安全调查方法，包括以下步骤：

步骤 S11：采集来自多个安全设备或系统的日志数据，以统一行为主体建模为目的字段标准化；

步骤 S12：构建与实体相关联的行为链结构，所述行为链结构为时间序列、图结构、标签链式结构或其等效路径结构；

步骤 S13：接收用户输入的自然语言安全调查问题；

步骤 S14：利用人工智能引擎将自然语言问题映射至与行为链结构相关的查询目标；

步骤 S15：基于行为链结构对查询目标执行语义推理、路径搜索或上下文分析操作；

步骤 S16：输出结构化或非结构化的查询结果，包括路径图、节点列表、行为片段、风险提示、专家反馈信息或提示式结果。

进一步地，在步骤 S11 中，所述日志数据在采集过程中进行字段标准化、实体统一、平台对齐、标识聚合，进行跨平台、跨系统日志行为链整合。

进一步地，在步骤 S12 中，所述行为链结构基于日志时间顺序、操作关联、标签联动、因果依赖或知识图谱路径建立，支持行为节点的图形化展现与实时或持久化存储；所述行为链结构嵌入增强信息包括节点风险评分、结构路径权重、行为意图标签、异常置信度；跃迁点标识，所述跃迁点标识用于标记行为链中攻击阶段变化的关键节点；行为坐标，所述行为坐标记录每一行为在链路径中的阶段位置、角色归属、时间位次；阶段跃迁强度评分，所述阶段跃迁强度评分用于度量节点间跃迁的可信度，综合考虑时间连续性、角色一致性与行为模式偏移度；所述增强信息用于优化人工智能模型的路径排序、异常识别与阶段跃迁推理能力。

进一步地，在步骤 S12 中，所述行为链结构为预先构建的结构或在查询过程中基于用户指定的实体标识、时间范围或行为特征条件所动态聚合形成的事件片段集合；所述事件片段集合在语义层面具有关联性或逻辑顺序；所述行为链结构作为构建基础，触发结构化事件调查流程，包括特定 IP 或用户的行为链重建、相似攻击路径比对、行为归因、协同分析任务；行为链节点构成包括通信动作、登录行为、文件操作、配置变更、进程调用、URL 访问、API 请求或其它可识别的安全行为事件。

进一步地，在步骤 S14 中，所述人工智能引擎包括大语言模型、NLP，图神经网络、规则-学习混合引擎或其等效语义推理系统；所述人工智能引擎将自然语言问题转化为结构化查询模板、图结构遍历请求或 SQL 语句，并结合 Prompt 工程进行自然语言与结构化查询的高效映射，支撑语义推理过程。

进一步地，用户以自然语言方式进行连续提问，人工智能基于行为链上下文状态维持语义会话，支持上下文缩放、链结构焦点切换、逆向提问生成与交互式行为路径演化，并支持跨轮次上下文记忆与专家反馈驱动的对话演化能力。

进一步地，在步骤 S16 中，输出结果包括攻击路径图、结构节点列表、时间窗口内行为轨迹、匹配片

段、风险说明、推理失败提示或需人工确认的半结构结果。

进一步地，一种语义行为链重构与 AI 引导的安全调查系统，包括字段归一化模块、行为链建模模块、AI 语义推理模块、微模型评分模块以及专家反馈与知识回写模块；所述字段归一化模块、行为链建模模块、AI 语义推理模块、微模型评分模块以及专家反馈与知识回写模块构成逻辑反馈闭环，支撑行为链构建、攻击路径重构与人工智能语义分析过程；

所述字段归一化模块对采集的多源日志数据进行字段标准化、实体归一化、平台对齐与关键字段映射；

所述行为链建模模块基于时间顺序、语义一致性与角色一致性构建行为链结构，作为 AI 推理与路径重建的输入基础；

所述 AI 语义推理模块对行为链进行路径推理、阶段跃迁识别与风险评估，支持自然语言接口调用；

本模块基于本地部署的大语言模型，支持用户通过自然语言直接发起行为调查，无需编写 SQL 语句。系统通过 Prompt 工程生成结构化 SQL，具体流程为：封装标准 Prompt 模板，当用户输入自然语言问题时，系统调用 DeepSeek 等本地部署的大语言模型（LLM）对问题进行语义解析，并结合 LangChain 等技术框架将自然语言问题自动转译为对应的 SQL 查询语句，查询角色行为数据库中的相关行为数据。专家通过集成的聊天平台与 AI 进行交互式提问与结果复核，通过 post 请求发送到数据库执行查询，并将结果返回给用户。如 openchat 等 Ai 聊天平台，自建 Web 端、IM 工具嵌入式对话框，或安全运营平台内置 AI 对话模块，例如用户提问“在 2025 年 3 月 26 日 14 点到 18 点期间，IP 1.1.1.1 做了哪些动作，是否有异常行为？”，系统会将其转译为有效 SQL 查询该 IP 的角色行为数据库中的数据表并获取结构化日志。然后 AI 将基于行为链建模层中的行为序列、历史知识库中的攻击模式、微观评分结果、路径坐标信息等多维数据进行综合推理，并以结构化或图形化结果返回至聊天平台界面，供专家进一步分析确认。其中，微观评分结果可以是 Model_X 字段，路径坐标信息包括行为坐标点、行为跃迁点，行为坐标点是行为链中的最小语义单元，标记行为事件在链路中的位置与语义信息。行为跃迁点是行为链中发生阶段性跃迁的关键语义节点，通常位于攻击路径的“战术转折处”。

专家可在对话过程中通过追加自然语言问题实现上下文递进式调查，AI 能够保持行为链上下文状态，支持链结构焦点切换、阶段跃迁推理与相似性比对，最终辅助专家完成完整的调查闭环并将确认结果回写至结构化行为链知识库，持续优化 AI 的推理能力与知识积累。该模块实现了“人话即查询”的智能问答式调查体验，显著降低了安全分析的使用门槛。

基于“角色行为数据库”中每个角色的行为日志，通过 AI 自动重建时间顺序下的行为链。当用户通过自然语言发起查询时，AI 会在角色行为数据库中检索指定角色在指定时间内的所有日志，按时间先后顺序排列构建行为链。例如，以 300 条日志为返回示例，系统可识别如“登录失败→登录成功→上传→命令执行→外联”等潜在攻击路径。该建链方式不依赖于固定规则或设备特征，而是通过上下文理解与语义组合构建逻辑行为序列。系统自动融合主机、流量、安全等多类日志来源，具备跨时间段、多源异

构数据建链能力，能有效识别异步行为、延迟攻击等复杂战术。AI 对构建的行为链进行判断和数据分析后，返回最终的分析结果交由专家判断或进一步人机交互。建议在行为链每个节点附带评分值、日志来源与字段解释，以提升分析的可追溯性。

所述微模型评分模块对标准化日志字段执行多维度模型评分，生成威胁评分字段，支撑行为链建模与 AI 推理；

所述专家反馈与知识回写模块将经人工专家确认的行为链与阶段标签反馈至知识库，形成可持续演进的 AI 知识记忆层。

进一步地，在微模型评分模块中，日志数据处理与评分分析方法对结构化日志字段进行外部分析处理并统一写入数据库，包括以下步骤：

步骤 S21：日志预处理，接收原始日志数据后，利用本地脚本程序、应用程序或等效模块执行字段提取、数据清洗及标准化操作；

步骤 S22：接口调用，将预处理后的日志字段以结构化格式通过 API 接口发送至一个或多个外部分析服务；

步骤 S23：服务处理，外部分析服务包括微观模型服务、评分引擎、特征提取器、标签分类器或基于规则的评估器，用于执行行为评估、威胁打分、字段统计、上下文解析的处理操作；

步骤 S24：结果接收与封装，统一接收分析服务返回的结构化结果并进行封装处理；

步骤 S25： 数据库写入，将处理后的结果数据写入至数据库的预定义字段，供行为链建模、攻击阶段识别、可视化呈现或知识库更新使用。

系统通过 Python 调用 Flask/FastAPI 接口连接训练好的 TensorFlow 的模型，对每条日志进行微观威胁评分，该评分值会写回数据库指定字段，为后续分析提供量化依据。首先，系统通过 python+tensorflow 根据需求训练并自定义生成异常评分模型，这些模型涵盖多个维度，包括使用 autoken 和随机森林训练正常数据模型以评估日志正常或异常、使用恶意行为特征的数据集训练针对单条语句的恶意检测模型、基于工作时间和非工作时间的时间维度模型、特殊字符数量模型，以及通过统计所有 api 服务路径训练识别异常路径的访问异常路径模型。系统将这些机器学习后生成的 H5、Keras 模型通过 Flask/FastAPI 打包成为微服务，使用 python 脚本定时访问微服务对推送入库的每条数据进行打分，并将结果回写至 sql 数据库指定字段。此外，系统支持模型融合判断与可解释性输出，如未来接入 SHAP/LIME 等，此模块是数据驱动分析的核心支撑。

在字段归一化模块中，基于三元字段映射对安全日志进行处理，包括以下步骤：

步骤 S31：采集来自安全设备的日志数据，提取原始字段 srcip 与 dstip；

步骤 S32：构建 external_ip → srcip → dstip 的三元结构，当日志来源为可直接接收公网访问的边界设备时，将原始 srcip 映射为 external_ip，将原始 dstip 映射为逻辑 srcip，并保留原始 srcip 字段内容不用于行

为链建模。

本模块负责从多种安全设备中采集日志，并进行标准化处理，构建角色行为数据库。首先进行日志标准化，从各个不同来源的日志中提取相关字段，并映射至包括但不限于 Date、Time、external_ip、srcip、dstip、source_from、Event_ID、Detail、Model_X 等标准关键字段；随后为该日志在数据库中生成唯一标识（Log_id），并允许根据具体实现场景调整字段集，以保证字段之间的结构一致性和语义一致性，支撑后续行为链建模和 AI 语义分析过程。

其中，external_ip 字段用于根据原始边界设备日志中 srcip 字段的场景判断是否应用三元映射机制，若判定为外部地址，则记录对应的外部访问 IP 地址；srcip 字段用于表示内部行为主体；dstip 字段用于记录行为目标地址；source_from 字段用于标注日志来源设备信息；Detail 字段用于封装其余原始字段内容，采用 JSON 格式存储；Model_X 字段用于存储微观模型评分结果，所述评分结果通过 API 调用外部分析服务，并在获取评分结果后写回数据库。其中，外部分析服务包括但不限于微观模型服务、评分引擎、特征提取器、标签分类器或基于规则的评估器，用于执行行为评估、威胁打分、字段统计、上下文解析等处理操作。

随后进行关系型数据库初始化，以 SRCIP 为角色主体分别建立数据库表，构建角色行为数据库。

在本实施例中，使用 filebeat、logstash、python 根据日志主体不同分别写入各个角色数据库中所属的数据库表中，日志可按时间排序。通过映射至包括但不限于 Date、Time、external_ip、srcip、dstip、source_from、Event_ID、Detail、Model_X 等标准关键字段以及生成该条日志的唯一标识（Log_id），使不同来源日志在后续处理流程中具备统一结构语义，而增加的 source_from 字段则有助于后续建模溯源。角色数据库以 srcip 为主键，每个角色对应一张表，按时间序列排序存储，统一写入相关日志，建立以“行为主体”而非“事件”为核心的安全数据模型，这种结构天然支持攻击溯源、行为链重建与行为模式学习。

系统首先从多种安全设备采集原始日志，统一字段格式实现异构数据标准化，并通过引入三元字段映射明确行为主客体归属关系，为后续行为链建模与 AI 语义分析奠定统一数据基础。标准化日志随后进入微观模型威胁评分层，通过集成多维度威胁检测模型对每条日志进行细粒度评分，生成 Model_X 等评分字段，辅助后续行为链建模与风险判别。评分后的日志数据按角色实体进行聚合，进入 行为链建模层，系统基于时间顺序、行为语义一致性与角色连续性重建行为链，形成结构化行为序列，如：“A1 → A2 → A3 → A4”。

已建模的行为链进一步交由 AI 语义推理层处理，系统支持专家通过 自然语言输入向 AI 提问，如“某 IP 在某时间段是否存在异常登录行为？”，AI 将自然语言转化为结构化结果查询，如通过 langchain + LLM 输入，结果输出给 LLM 进行分析，专家通过自然语言与 AI 系统交互，由 AI 理解并分析指定行为主体在特定时间段内的动作序列是否存在可疑特征。分析过程中，AI 引擎会综合调用历史行为链复现库、微观模型评分库及关键字段解析库等多维数据资源进行语义推理判断。其中，微观模型评分库包括针

对单条日志或行为节点的细粒度威胁评分，例如 MySQL 中的 model4 字段，0% 表示正常，100% 表示恶意，供异常行为参考。

关键字段解析库则包含如 Host_eventid 等字段，如 Windows 安全日志中的 eventid 字段，用于辅助判断主机行为意图与阶段归属。结合行为链、历史知识与上下文信息进行路径推理、阶段跃迁（Pivot Point）识别与异常评估，生成分析结果与推理依据。

系统在完成 AI 引擎对行为链的初步分析后，将分析结果提交专家复核，形成“人机协同”的判断闭环。专家基于 AI 推理结果与自身分析，确认是否存在攻击行为或异常行为链，若确认，则将具备攻击特征的关键行为链路径及对应特征写入 AI 知识库，包括行为链复现库、风险标签与意图库、路径语义坐标库等模块，形成可供后续分析与模型优化的经验数据积累。例如，关键行为链路径可以为：“A1 爆破 → A2 登录成功 → A3 上传行为 → A4 命令执行 → A5 外联行为”。该知识库在持续演进过程中，可进一步升级为图数据库结构，支持高效的攻击相似性比对与知识推理，提升系统整体智能调查与辅助决策能力。

进一步地，所述专家反馈与知识回写模块包括结构化行为链知识库，所述结构化行为链知识库将已验证的攻击行为链及其关联信息写入知识库，并基于知识库的自动语义判断与行为链相似性检索。

所述专家反馈与知识回写模块中，专家通过人工智能辅助调查流程对安全事件进行判定并反馈知识，手动触发比对流程，将当前待排查的行为链与知识库中已存的历史行为链进行相似性比对，辅助判断为相似攻击路径或阶段跃迁风险，所述比对流程独立于 AI 自动推理流程运行。

所述知识库可包括行为链复现库、字段语义解析库、风险标签与意图库、图谱关系数据库、路径语义坐标库及其他优化信息知识库；所述路径语义坐标库存储行为链中各节点的结构位次、阶段标签、行为坐标、跃迁点标识与跃迁强度评分。

所述结构化行为链知识库利用结构向量、语义标签或图结构重叠度进行相似性匹配，基于历史行为链进行高置信度攻击路径比对与阶段跃迁推理，匹配结果作为人工智能模型输入或安全分析师辅助参考。

所述结构化行为链知识库与人工智能模型形成闭环交互机制，在人工智能模型的训练阶段、更新阶段及推理阶段，所述人工智能模型动态依赖知识库中的结构化行为链数据、路径语义坐标、跃迁点标识与阶段跃迁强度评分信息。

以上对本发明的具体实施例进行了详细描述，但其只是作为范例，本发明并不等同于以上描述的具体实施例。对于本领域技术人员而言，任何对本发明进行的等同修改和替代也都在本发明的范畴之中。因此，不脱离本发明的精神和范围下所做的均等变换和修改，都应涵盖在本发明的范围内。