AI 大模型技术背景下网络数据安全协同防护体系构建

文丨 宋超 王静芬

随着数字经济深入发展，网络数据安全威胁持续升级，数据泄露、对抗攻击等复合风险对传统防护体系构成严峻挑战。AI 大模型凭借强大的语义理解与智能推理能力，为网络数据安全协同防护提供了全新技术路径。对典型案例、关键技术应用及防护效能评估开展系统研究，有助于推动网络数据安全协同防护体系的科学构建。

典型案例

某大型互联网平台网络数据安全协同防护项目，针对海量用户数据在存储、传输及跨域调用过程中面临的多维复合威胁，以 AI 大模型作为核心驱动引擎，围绕元宇宙场景下数字身份认证、虚拟数据资产保护及跨平台数据交互安全等新兴风险议题，系统地部署融合大数据智能分析、语义推理感知及多智能体协同联动的新型防护架构。该项目着重应对对抗样本注入、深度伪造内容渗透及非授权访问隐私数据等高隐蔽性威胁，依靠大模型跨模态理解能力，对异常流量和可疑行为开展动态识别；同时，引入隐私计算技术对敏感数据全生命周期进行智能管控，构建起覆盖感知层、传输层及决策层的立体化主动防御体系，为网络数据安全协同防护关键技术的深化应用提供实践参照。

网络数据安全协同防护的关键技术应用

大模型驱动的网络威胁智能感知

在上述典型案例所展示的协同防护实践基础上，大模型驱动的网络威胁智能感知，成为整个防护体系的技术起点，和传统基于规则匹配的静态检测方式有所不同，大模型依靠强大的语义理解与上下文推理能力，对深度伪造指令、变形恶意载荷及隐蔽渗透流量等高对抗性威胁，进行语义层面的精准捕捉，从而达成从特征匹配到语义理解的检测范式跨越。在威胁感知的量化建模当中，异常流量检测置信度评估采用如下公式：

$$S=\frac{1}{1+\mathrm{e}(\beta_0+\beta_1 x_1+\beta_2 x_2++\beta_n x_n)}$$

式中，S 为威胁置信度评分（无量纲，取值范围［0～1］）；x_i 为第 i 个网络行为特征向量（单位：bps，即比特每秒）；β_i 为对应特征的权重系数（无量纲）；β_0 为模型偏置项（无量纲）。大模型对各特征维度的语义权重进行动态调整，使感知模型面对元宇宙场景下多模态交互流量时，依然能够维持稳定的检测精度，在某智慧城市网络安全项目中，依托上述感知架构，对跨平台数字身份伪造攻击进行实时识别，显著提升了复杂威胁环境下的感知覆盖深度，威胁感知能力得到强化，为后续多源异构数据的安全协同防护奠定技术前提。

多源异构数据的安全协同防护

随着大数据技术的深化应用，网络环境中结构化日志、半结构化接口数据及非结构化多媒体内容同时存在，多源异构数据的安全协同防护，成为体系建设关键环节，针对异构数据在采集、聚合及分析链路中，存在语义鸿沟和安全边界模糊问题，以大模型为中枢的数据协同防护机制，借助跨模态语义对齐技术，实现对不同数据源安全状态的统一感知与联动研判。在元宇宙场景中，虚拟空间产生的三维交互数据、行为轨迹数据及物理空间的传感器数据深度交织，传统单模态防护手段难以满足此类高维异构数据流的安全管控需求。为此，大模型依靠多头注意力机制对异构特征开展跨维度关联分析，识别隐藏在数据融合节点中的潜在攻击路径，在某工业互联网安全防护案例中，该机制对来自操作日志、网络流量及设备状态三类异构数据源的协同威胁进行联合研判，有效解决了跨域“数据孤岛”带来的防护盲区问题，为数据安全流转的隐私保护提供了坚实技术支撑。

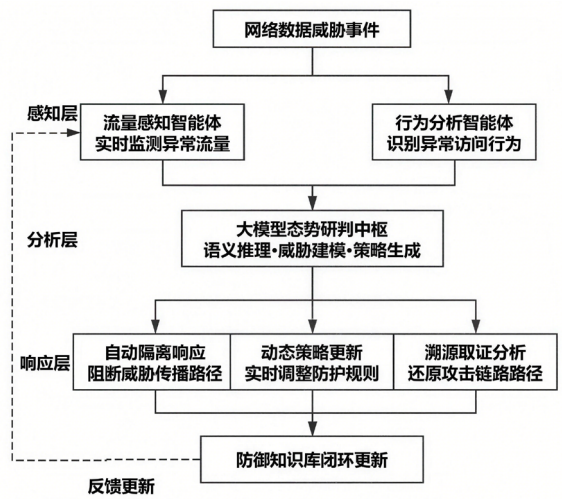
隐私计算与数据安全流转融合

数据安全流转是网络数据安全协同防护体系中连接感知、分析及响应各环节的核心纽带。在数据跨组织、跨域共享场景中，隐私泄露风险与合规约束的双重压力，让安全流转机制的设计格外复杂。隐私计算技术将差分隐私、联邦学习及同态加密这三类核心机制有机融合，构建出“数据可用不可见”的安全流转范式，差分隐私在数据发布阶段向统计结果注入可控噪声，从数学层面保障个体隐私免受推断攻击。联邦学习使各参与节点在本地完成模型训练，仅交换梯度参数而非原始数据，实现隐私保护前提下的协同建模，同态加密支持在密文状态下，对敏感数据直接执行安全计算，保障数据在传输与处理全链路中的机密性。在元宇宙数字资产流

转场景中，上述三类技术的协同应用，能够有效规避跨平台数据交换时的隐私暴露风险，数据安全流转能力的提升，为智能体联动与主动防御机制的协同部署创造了条件。

智能体联动的主动防御机制

在 AI 大模型技术赋能下，智能体联动的主动防御机制，以多智能体协同调度为核心，构建起覆盖威胁感知、态势研判、策略生成及自动响应的全链路闭环防御体系。如图 1 所示，该机制借助感知层、分析层及响应层智能体的分层协作，达成对网络数据安全威胁的动态处置。感知层智能体承担对网络流量和用户行为数据进行实时采集与异常标注的任务。分析层大模型态势研判中枢依靠语义推理，对威胁意图进行深度解析并生成应对策略。响应层智能体按照策略分类自动执行隔离处置、规则更新及溯源取证三类差异化操作，处置结果会同步反馈至防御知识库，以驱动模型持续迭代。在元宇宙跨域交互场景下，该机制能够针对虚实融合环境中的数字身份劫持等新型攻击实现快速联动响应，把被动处置方式转化为预测性的主动防御，有效提升了协同防护体系的整体智能化水平。



来源：根据网络公开数据整理

‖ 图1 混合集成算法优化架构图

协同防护体系的效能评估

协同防护体系的威胁拦截效果评估

在大模型智能感知、多源异构数据协同防护及智能体联动防御等关键技术全面部署的基础上，以典型互联网平台协同防护项目当作评估对象，从深度伪造、攻击拦截等多个维度对防护体系拦截能力进行系统评估，结果如表 1 所示。由表 1 可知，协同防护体系在五项核心指标上均较传统模式实现显著跃升，元宇宙场景异常识别率和跨域威胁联动响应率提升幅度最突出，这表明大模型语义推理与多智能体协同机制的融合，有效突破传统防护模式在高维异构场景下覆盖盲区，并为协同防护效能的量化分析提供了数据支撑。

‖ 表1 协同防护体系威胁拦截效果评估

评估维度	传统防护模式	协同防护体系	提升幅度
深度伪造攻击拦截率（%）	71.3	96.8	+25.5
恶意流量识别准确率（%）	78.6	97.2	+18.6
数据泄露阻断成功率（%）	74.2	95.6	+21.4
跨域威胁联动响应率（%）	62.4	93.7	+31.3
元宇宙场景异常识别率（%）	58.9	94.3	+35.4

来源：根据网络公开数据整理

协同防护效能的量化评估分析

在威胁拦截效果评估的基础上，进一步聚焦响应时效、资源调度效率及防护稳定性等综合运行指标，对协同防护体系的整体效能开展量化分析，结果如表 2 所示。由表 2 可知，协同防护体系在平均威胁响应时延方面的压缩幅度最为显著，跨平台数据协同处理效率和防护稳定性也得到大幅提升，这表明，以大模型为中枢、融合隐私计算及多智能体联动的协同架构，在响应时效和运行稳定性等效能维度都取得了实质性突破，为网络数据安全协同防护体系的工程化落地提供了有力的量化依据。

‖ 表2 协同防护体系效能量化评估对比

评估指标	传统防护模式	协同防护体系	优化幅度
平均威胁响应时延（ms）	843	127	−84.9%
智能体协同调度成功率（%）	69.5	96.1	+26.6
隐私数据安全流转合规率（%）	75.8	98.3	+22.5
高并发场景防护稳定性（%）	67.3	94.7	+27.4
跨平台数据协同处理效率（%）	61.2	92.8	+31.6

来源：根据网络公开数据整理

结束语

网络数据安全协同防护体系的构建，是应对当前复杂网络威胁环境的必然选择。通过融合 AI 大模型语义推理、隐私计算、大数据智能分析及多智能体协同联动等前沿技术，有效解决了传统防护模式在威胁感知、数据安全流转及跨域协同响应等方面的结构性不足，实现了从被动响应向主动防御的体系转型。在国家数字安全战略深入推进的背景下，网络数据安全治理已上升为战略优先议题。面向未来，随着联邦学习、可信执行环境及元宇宙安全治理机制的持续成熟，协同防护体系将向自主感知、动态演化及智能决策方向深化演进，为网络空间数据安全治理提供更坚实的技术支撑。

作者简介：宋超 王静芬 崂山实验室

责任编辑：杨佳宇 投稿邮箱：zhouhl@staff.ccidnet.com