解决方案简介

绿盟科技天枢实验室引入英特尔® 流量分析开发工具套件，打造针对 Web 攻击的 AI 高性能检测方案

“在网络安全隐患日益凸显的大环境下，利用 AI 技术综合提升 Web 防护产品的防御能力，来弥补传统规则匹配的 Web 攻击缺陷已成为业内共识，但 AI 引擎的加入也增加了对性能的要求。为此，我们基于英特尔® TADK 打造了针对 Web 攻击的 AI 高性能检测方案。该方案具有较高检测准确率及泛化能力，为持续创新开发 Web 防护产品提供了宝贵经验。”

顾杜娟
主任研究员
绿盟科技天枢实验室

近年来广受关注的 Web 防护产品，如 Web 应用程序和 API 保护平台 (WAAP) 等，能在应用层——即开放式系统互联 (Open System Interconnection，OSI) 模型的第七层提供针对黑客攻击的有效方案。但传统 Web 防护解决方案通常依靠规则引擎等方法来防御 Web 攻击，而随着技术迭代速度的加快以及企业业务场景的变化，基于 Web 应用构建的业务场景具有越来越高的复杂度，使基于规则构建的安全策略的缺陷日益凸显，迫使企业需要不断调整规则来应对攻击者的挑战。然而，这不仅极为依赖专家资源，使运维成本居高不下，同时漏报率、误报率等问题也不断攀升。

得益于人工智能 (Artificial Intelligence，AI) 技术的发展，结合 AI 技术来提升 Web 防护产品的防御能力已逐渐成为业界共识。借助 AI 安全分析引擎对 Web 访问数据进行智能学习和建模，能有效增强识别未知威胁的能力，提升 Web 防护产品的检测准确率。因应这一趋势，绿盟科技集团股份有限公司 (以下简称“绿盟科技”) 天枢实验室基于数据智能安全的前沿研究，不断探索将各类 AI 算法应用于包括 Web 防护产品在内的安全解决方案中。

为进一步提升融合 AI 分析引擎的 Web 防护产品的应用效能，并加速商用化落地，绿盟科技天枢实验室与英特尔合作，基于英特尔® 流量分析开发工具套件 (Traffic Analytics Development Kit，TADK) 打造针对 SQL 注入 (SQL Injection，SQLI) 攻击、跨站脚本 (Cross Site Scripting，XSS) 攻击的 AI 高性能检测方案。下述一系列测试结果表明，新方案中基于英特尔® TADK 获得的 AI 模型有着良好的检测准确率[1]，并具有较高的泛化能力 (Generalization Ability)。

背景与挑战：Web 防护解决方案需要借助人工智能来提升效能

各类 Web 应用，包括各类互联网站点、企业信息化系统等在为人们的日常生活、企业的业务发展带来更多便利的同时，也正在成为黑客攻击的重要目标。攻击者利用形形色色的攻击手段，诸如 SQL 注入攻击、XSS 攻击等对用户的隐私、企业关键数据进行窃取或内容篡改，不仅严重侵害用户利益，也对企业信息安全造成巨大危害。

为应对新型攻击，越来越多的企业与组织正将面向 Web 安全构建的高级 Web API 防护、Web 应用安全防护等方案放在越来越重要的位置。而这些方案中，基于规则引擎、语义引擎等方法的产品已在市场中获得了广泛的运用。

规则引擎是目前 Web 防护产品中识别和阻止已知攻击的常见检测方法，具有解释性好、检出问题后有明确处置建议等优势。以业内 Web 应用防护系统为例，得益于所积累的大量静态规则，系统可对风险特征、行为进行精准防护，具备较强的匹配能力。而当静态规则与基于上下文的动态语义分析相结合时，还可基于规则对网络层、HTTP 请求 / 应答中的已知威胁进行检测，并对基于嵌套编码的攻击进行解码。通过词法、语法进行深层次分析，还可以识别隐藏极深的威胁，弥补基于规则引擎的 Web 防攻击技术的不足。

随着 WEB 应用所涉及的技术方案不断迭代、业务流程日益复杂，且应用场景更为多样化，其提供的安全保护也日益受到挑战。一方面，攻击者的攻击频次越来越高，攻击手段越来越多变且日趋智能化；另一方面，规则引擎和语义分析与客户环境缺乏实时关联，使漏报、误报的缺陷也逐渐凸显出来，甚至在 0day 漏洞防护时显得略为滞后。

AI 应用具备自主学习和自我更新的能力，能根据数据的迭代，通过不断学习来对模型参数进行调整，实现自我更新进化。而在 Web 应用防护中，同样也可借助 AI 技术，以现网流量数据来自我学习，理解应用的业务逻辑，从而为 Web 应用建立安全访问基线，成为阻断未知 Web 威胁的有效方法。通过与规则引擎、语义引擎融合，形成联合防御机制，全新的 AI 安全分析引擎能有效提升 Web 防护产品的工作效能。

拥抱智能化趋势，包括绿盟科技在内的安全厂商积极探索在规则引擎与语义分析之上，引入 AI 安全分析引擎来协同检测 Web 攻击，进而通过加持 AI 能力的 Web 防护产品为 Web 安全保驾护航。但随着各类 AI 安全分析引擎在 Web 防护产品中的运用越来越广泛，也带来了新的挑战。在 AI 引擎对流量明细、条件参数等进行学习，进而输出算法模型，产生防护“智慧”的过程中，会产生较大的性能开销。特别是在基于深度学习的 AI 引擎中，参数往往达到上万甚至上亿个，训练过程中有着大量张量运算。此时训练过程就会占用较高的处理器资源，使训练过程十分漫长，需要进一步通过优化来提升性能。

为有效应对这一挑战，英特尔与绿盟科技天枢实验室一起展开深入合作，借助英特尔® TADK 内置的多个优势功能特性与能力引擎，并与其它英特尔 AI 加速技术和 AI 框架优化进行协同，使基于深度学习方法构建的 AI 高性能检测方案提升了性能表现及检测能力，从而有力推动 AI 安全分析模型从实验室走进网络安全攻防实战。

解决方案：基于英特尔® TADK，打造应对 Web 攻击的 AI 高性能检测方案

为了在 Web 防护产品中构建高性能、高可用的 AI 安全分析引擎，英特尔与绿盟科技天枢实验室一起，基于英特尔® TADK 全新打造 AI 高性能检测组件，并与传统的规则引擎或语义分析组件进行集成。例如，针对 Web 攻击中最常见的 SQL 注入攻击、XSS 攻击等，英特尔与绿盟科技天枢实验室借助英特尔® TADK 提供的优势能力，并基于 URLNet[2] 打造了应对 Web 攻击的深度学习高性能检测方案。

方案基于内容对流量进行分析。首先，对于 Web 应用中常见的字符串信息，方案将生成字典 (vocabulary) 并进行高效的分词 (Tokenization)，然后输入到运行在英特尔® oneAPI Deep Neural Network (英特尔® oneDNN) 和 OpenVINO™ 工具套件上的 URLNet 模型中。

在这一过程中，英特尔® TADK 提供了一系列高可用、高性能的分析工具，例如词法分析库 (Lexical Parser)、分词器 (Tokenizer) 等，显著提升了方案性能；而 OpenVINO™ 工具套件对于模型推理的性能加速，也使 SQL 注入攻击、XSS 攻击的判别推理速度大幅提高[3]。与此同时，方案还引入了自动超参数调整能力，这一能力可在确保用户数据精确度的情况下，进一步提升模型推理速率，且不需要用户手动介入，提升了运行效率。

作为高性能的流量分析开发套件，英特尔® TADK 提供了可在 Web 应用中部署的端到端 AI 流水线关键组件，支持对 NGINX、ModSecurity 等开源软件的集成，并提供了基于 AI 方法的流量分类、Web 应用防火墙等应用能力。

英特尔® TADK 的架构如图 1 所示，其核心库由以下六个主要组件构成：

流特征提取库（Flow Feature Extraction Library, FFEL）

可获取数据中的元数据和统计信息，包括数据包特征、协议特征以及词袋（Bag of Words，BOW）等；

词法分析器（Lexical Parser，Tokenizer）

具有一个基于确定有限状态自动机（Deterministic Finite automaton，DFA）的分词器和分词编码器，可通过配置文件 / 字典来生成运行时的 DFA 引擎；

流分类器（Flow Classifier）

可执行基于 5 元组的双向流分类，并具有基于时间轮的流老化机制；

协议检测（Protocol Detection）

可用于检测和解析 HTTP、IPv4、UDP、TCP、DNS 等常见 Web 协议；

AI 引擎（AI Engine）

通过封装英特尔® oneDNN、OpenVINO™ 工具套件、英特尔® oneAPI Data Analytics Library（英特尔® oneDAL）等，可实现面向深度学习和机器学习的分类器，并提升大数据分析效率；

DPI 引擎（DPI Engine）

可从 TLS 流中获取 SNI 字段和证书字段，并利用 Hyperscan 对相关规则集进行匹配，对现网流量数据进行分类。

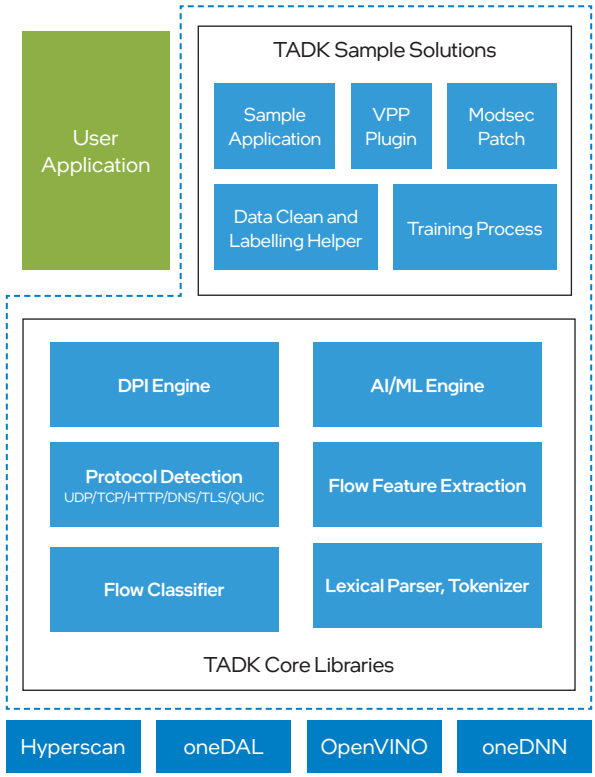

图 1 英特尔® TADK 架构

同时，英特尔® TADK 还可与英特尔提供的一系列 AI 加速技术和 AI 框架优化实现协同，来为 AI 安全分析模型的性能提供保障，这些能力既包括在英特尔® 至强® 可扩展处理器平台中集成的英特尔® 深度学习加速（英特尔® DL Boost）技术、英特尔® 高级矢量扩展 512（英特尔® AVX-512）等 AI 加速技术和指令，也包括面向英特尔® 架构优化的 AI 框架（例如英特尔® oneDAL、英特尔® oneDNN 等）、OpenVINO™ 工具套件等。

例如，在特征提取阶段，英特尔® TADK 能充分挖掘英特尔® 至强® 可扩展处理器中集成的英特尔® AVX-512 指令集所具备的优化能力，进行实时流量特征提取。英特尔® AVX-512 在数据寄存器宽度、数量以及融合乘加（Fused Multiply Add，FMA）单元的宽度上的优势，可以使方案实现通用计算能力和并行计算能力的双提升，并有效提升吞吐量性能表现[4]。

方案价值与收益

为验证基于英特尔® TADK 打造的针对 Web 攻击的 AI 高性能检测方案的安全能力表现，绿盟科技天枢实验室基于实际流量数据，进行了 SQL 注入攻击检测和 XSS 攻击检测的准确性及性能测试。数据来源于绿盟科技的 WAF、IPS 等防护设备中检测到的真实历史告警数据，告警数据产生时间越久，占比随之越低。

测试基于使用英特尔® TADK 的 URLNet 模型，利用训练数据，通过自动参数选择训练出一个成熟模型 (即训练获得拟合训练数据较好，训练准确率、验证准确率较高的模型)，并利用测试数据对模型进行测试，最终得到测试结果 (准确率)。测试方案中的攻击数据包含多种请求类型，包括多种类型的 SQL 命令注入方式及 XSS 攻击方式，覆盖多数的攻击类型变种：

训练数据

- SQL 注入攻击数据：10000 条；其他类型数据：18000 条；
- XSS 攻击数据 10000 条 其他类型数据 18000 条
- 其他类型数据包含：正常流量数据和命令注入、路径穿越等其他类型的 Web 攻击数据。

测试数据

- SQL 注入攻击数据：2022 条；其他类型数据：16000 条；
- XSS 攻击数据: 2022 条; 其他类型数据: 16000 条。

数据经由安全专家分析和打标签，构造出质量较高的训练数据集，保障了测试的可靠性。训练结果表明，训练出的 AI 模型对数据的拟合效果较好。由测试结果可知，SQL 注入攻击检测准确率达到 99.05%[5]，XSS 攻击检测准确率达到 99.6%[6]。

测试结果表明，基于 TADK 训练出的 AI 模型在 SQL 注入攻击、XSS 攻击检测方面具有较高检测准确率，对于正常流量的误报率较低，表明 TADK 方案具备较高的实际部署价值。

展望

一直以来，绿盟科技天枢实验室积极探索研究 AI 引擎在网络安全产品，尤其是 Web 防护产品中的应用，以求推进 AI 能力加持的 Web 防护解决方案不断走向成熟，提升产品应用效能。

得益于英特尔® TADK 提供的 AI 加速能力，绿盟科技天枢实验室对于探索以 AI 技术为核心能力的网安产品应用进入新的研究阶段。从实验室走入生产环境，绿盟科技天枢实验室也在积极推进 AI 分析引擎与规则引擎、语义引擎在 Web 防护产品中的融合应用，与英特尔联合推动网络安全防御体系迈向高度智能化、自动化，推动 Web 防护产品在生产环境中的检测准确率不断攀升，并降低耗时、提升 AI 性能和效率，加速融入 AI 能力的 Web 防护解决方案商业化落地。

[1, 3] 相关测试结果参见本文【方案价值与收益】部分内容。
[2] URLNet：是基于卷积神经网络 (Convolutional Neural Networks，CNN) 的一种 Web 防护模型，用于检测恶意 URL。
[4] 性能因用途、配置和其他因素而异。更多信息请访问 :https://www.intel.cn/content/www/cn/zh/architecture-and-technology/avx-512-overview.html
[5, 6] 测试配置：处理器：双路英特尔® 至强® Gold 6438N 处理器，32 核 64 线程；内存：256GB；操作系统：Ubuntu20.04；英特尔® TADK 版本：22.09。绿盟科技于 2022 年 11 月 20 日测试，且可能并未反映所有公开可用的安全更新。