doi: 10.7690/bgzdh.2021.01.005

网络安全态势感知平台架构设计

糜　旗
（中国航天科技集团第八研究院上海航天动力技术研究所，上海 201109）

摘要：为提高网络安全防范能力，设计网络安全态势感知平台架构。详细阐述其架构与功能模块设计，利用大数据技术将异构日志源数据进行存储、处理，采用数据挖掘、机器学习算法等进行分析、整合，并用可视化技术将结果呈现给用户。通过该平台，可建立针对网络未知威胁的动态安全监控与防御体系，避免因网络攻击导致的数据泄露、信息系统被破坏等安全问题。

关键词：安全态势感知；架构；机器学习
中图分类号：TP393.081　　**文献标志码：**A

Network Security Situation Awareness Platform Architecture Design

Mi Qi
(*Shanghai Space Propulsion Technology Research Institute*,
No. 8 *Academy*, *CASC*, *Shanghai* 201109, *China*)

Abstract: In order to improve network security prevention capabilities, the network security situation awareness platform architecture is designed. It elaborates its architecture and functional module design, uses big data technology to store and process heterogeneous log source data, uses data mining and machine learning algorithms to analyze and integrate, and uses visualization technology to present the results to users. Through this platform, a dynamic security monitoring and defense system against unknown network threats can be established to avoid security issues such as data leakage and information system destruction caused by network attacks.

Keywords: security situational awareness; architecture; machine learning

0　引言

随着互联网技术在我国的快速发展和普及，有组织、有政治目的的网络攻击也明显增多。我国面临大量来自境外的钓鱼、后门、植入木马以及僵尸网络等攻击行为，给国家的基础设施和关键信息系统带来严重的威胁和挑战。尤其 2013 年斯诺登爆出的“棱镜门”事件[1-2]，揭露了美国对多个国家和人民实施的长期监听，引发了国际社会的强烈反对。这只是美国安全监控的冰山一角，可见网络安全对于国家安全的重要性。面对如此复杂、严峻的网络安全环境，网络管理者在网络中部署了大量的网络安全设备（如防火墙、入侵检测、漏洞扫描、防病毒设备等）。虽然各类安全设备或系统为保障网络安全发挥了重要作用，但是各类设备相对独立，无法反映总体的网络安全状况[3-5]及面临的威胁。针对上述网络安全问题，近年来网络安全态势分析已经成为国内外研究的焦点[6]；因此，笔者对其平台架构设计进行论述。

1　平台架构设计

平台架构设计思路是采用大数据和分布式的架构，实现海量日志的实时处理、分析和存储。面向数据库、系统、应用、中间件、网络安全设备等采集日志，支持 syslog 协议、snmp、odbc、命令行、客户端代理等多种方式采集。将日志采集到服务端，支持 worm 固化防篡改，可设定保留周期，满足至少保留 6 个月全量日志的要求。

采用 ELK（elasticsearch logstash kiabana）架构及 NoSql 数据库，可实现日志清洗与查询，具备高性能搜索引擎，实现海量日志秒级查询。具备日志大数据分析能力[7]，结合机器学习、SPL 语言以及关系图谱等进行高级分析。既可展现原始日志，也可用各类图表、仪表盘等展现分析结果，支持 api 对接，支持分析配置定制以及开发定制[8]。

1.1　平台架构

网络安全态势感知平台总体架构如图 1。

网络安全态势感知平台分为信息资源层、应用

收稿日期：2020-08-27；**修回日期：**2020-10-17
作者简介：糜　旗（1979—），男，上海人，硕士，高级工程师，从事信息安全、系统架构、逆向工程研究。E-mail: zeror@163.com。

支撑层和应用层。

1) 信息资源层：支持整合不同来源数据，包括文本日志、应用接口(应用数据)、网络设备、存储设备、数据库记录和安全设备等[9-11]。提供应用系统的日志生成规范，包括日志格式、存储、字段内容等，通过日志采集完成各种信息资源的集中存储。

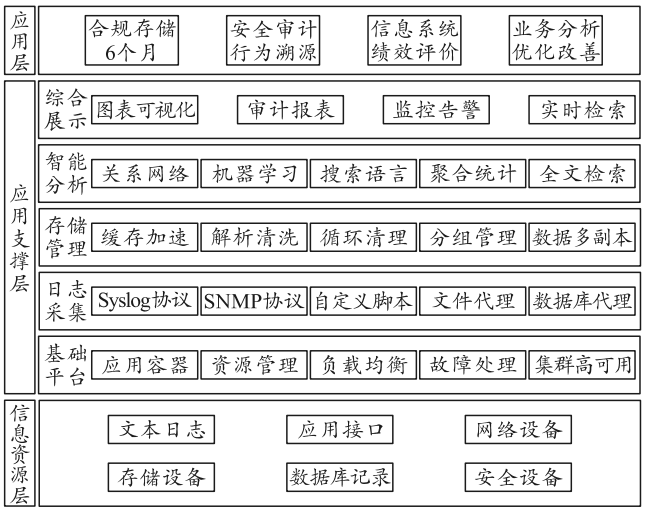

图 1 网络安全态势感知平台总体架构

2) 应用支撑层：包括基础平台、日志采集、存储管理、智能分析、综合展示，负责数据采集、处理、分析与呈现[12]。

3) 应用层：面向业务场景交付，包括日志合规的分析结果、安全审计、绩效评估和业务分析等方面，是在开发过程中不断迭代优化的部分[13]。

根据总体架构图的指导，将功能模块分解为基础平台、日志采集、存储管理、智能分析、综合展示。

1) 基础平台：主要解决高可用和资源交付，主要包含应用容器、资源管理、负载均衡、故障处理和集群高可用[14]。通过 docker 微服务化的模式实现高可用、故障处理以及资源交付，基于微服务容器方式交付，使整个网络安全态势感知平台的维护及未来的扩展都具有可控性。

2) 日志采集：采集对象主要包括主机系统、数据库、网络设备、各类安全设备、审计业务相关应用系统等[15-17]。支持通过 syslog 协议、SNMP 协议、自定义脚本、文件代理、ODBC 和数据库代理等多种方式采集。

3) 存储管理：包含数据的缓存加速(解决大量数据的峰值处理)、解析清洗(将不规则或者非结构化数据整理为结构化数据)、循环清理、分组管理和数据多副本[18]。

4) 智能分析：包含关系网络分析、机器学习、搜索语言、聚合统计和全文检索，可以面向不同的应用需求进行分析。

5) 综合展示：包含图表可视化、审计报表、监控告警和实时检索[19]。

在应用层，根据不同的业务需求以仪表盘或者 APP 的方式展现，在功能设计层面增加了仪表盘和 APP 的样式。

1.2 平台主要功能

1.2.1 全量的日志数据采集与解析

全面采集基础架构层、业务数据层和网络设备层的全量日志数据，通过内置的多种解析方式，实现不同格式日志的深度结构化。

1.2.2 实时的动态视图与监控告警

自定义的动态视图、灵活的监控告警策略以及实时的全局索引搜索，帮助用户全方位洞察日志数据的趋势变化，并可快速应用日志数据。

1.2.3 智能的日志关联分析与预测

通过搜索处理语言来实现逻辑钻取查询，辅以外部日志数据关联分析，可以综合处理复杂的联动事件，机器学习也可用于预测事件的趋势。

1.2.4 高效的日志采集部署与响应

线性可扩展的日志采集集群部署方式，可支持传输多达 50 TB 的日志数据。即便是 TB 级的超大数据量，也将于数秒内反馈精准的请求结果。

2 平台功能模块设计

2.1 日志采集方式设计

日志采集如图 2 所示。

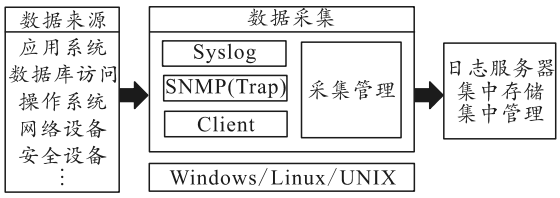

图 2 日志采集

通过提供丰富的日志采集方式，能够收集来自各种应用系统和网络设备产生的日志内容，以此来掌握整个网络环境的运营活动数据[20]。

2.2 日志解析清洗设计

日志本身是非结构化的，将日志文件从非结构化转换为结构化数据。通过解析，一条完整的日志

会被拆分为多个字段，每个字段都可作为单独的搜索过滤条件。不同格式的日志文件解析方法有所不同，解析的完整性和准确性也大有不同。日志解析清洗设计如图 3 所示。

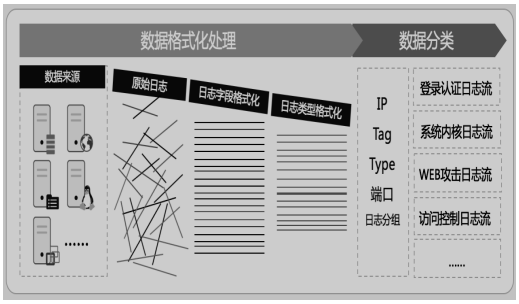

图 3　日志解析清洗

2.3　数据存储

使用分布式、可伸缩、实时搜索和分析引擎作为底层存储系统，所有存储在上面的数据都是高度可用，通过简单的增加节点进行扩展，通过数据分区和复制透明的技术处理来防止节点损失。

2.4　可视化统计分析

提供丰富的图表可视化功能，对日志数据进行高级数据分析及展示，让海量数据更易被理解，能够方便地创建、保存、分享图表数据，如柱形图、折线图、散点图、直方图、饼图和地图等。

通过将各类系统日志数据进行统计分析并生成各类实时报表，以日常合规审计视角进行多维度展示日志数据。

2.5　监控告警设计

执行定期或实时的监控告警任务，并依据指定的分析结果，通过邮件发出各项警示通知。告警分析依赖已保存的搜索，结合自定义的字段数值检测，能够实现多种维度的异常分析[21]。

2.6　关系图谱智能分析

关系图谱技术的数据流转过程如图 4。在基于统计分析的基础上，要实现智能化和先进性的分析，需要具备关系图谱、机器学习等高级分析的能力。关系图谱支持以实体建模的方式实现不同对象的关联关系，可以通过交互式点击下钻的方式进行分析，继而发现规律或者异常。

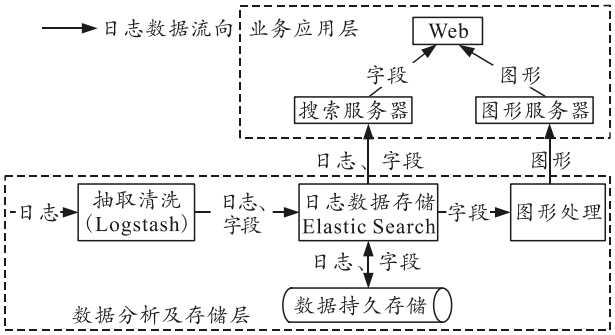

图 4　关系图谱技术的数据流转过程

2.7　机器学习算法库

传统的日志审计分析技术手段使用的是专家经验库规则分析模式，依据事先设计好的条件和规则去发现线索并提前预警，面对多变的用户行为动作和环境差异，经验库的局限性非常大、限制条件多，很难真正有效发挥作用。

机器学习技术能够极好地解决海量数据、复杂条件下的异常特征发现，非常适合运用在日常行为的异常识别，如 APT 攻击、欺诈行为识别、攻击链追溯，或者是用户、设备、应用的异常行为识别。也能结合网络环境中以用户为视角收集的所有行为数据，进行分类、聚类、离群值检测等分析，获得用户画像标签，找出异常用户、异常行为，或者是找出不同用户之间较为明显的差异特征[22]。数据处理流程中的模型训练和算法库如图 5 所示。

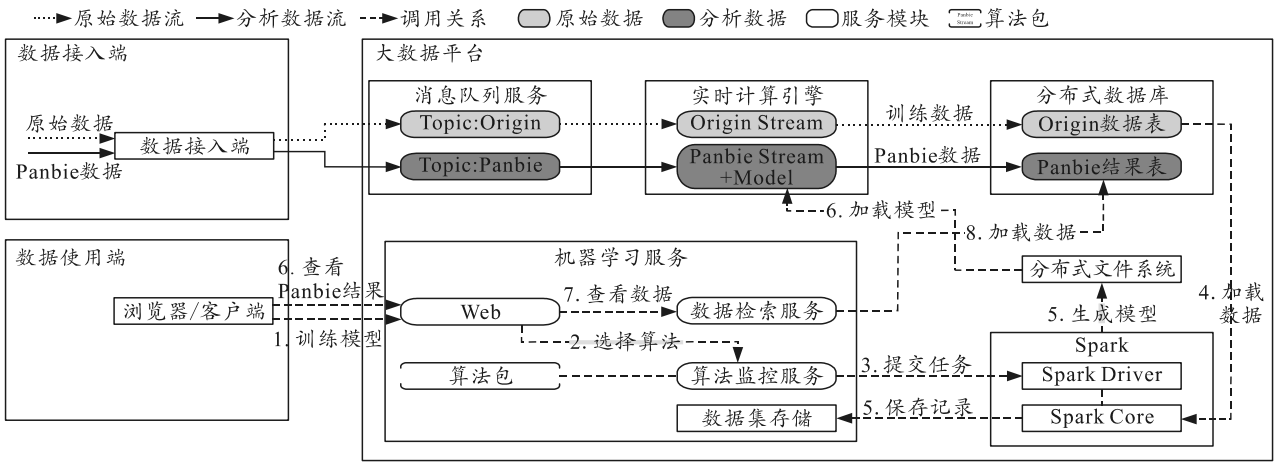

图 5　数据处理流程中的模型训练和算法库

2.8 模型训练

模型训练是平台的重要组成部分，由机器学习服务负责进行界面展示和任务提交，训练任务执行由批处理引擎和机器学习算法库 2 个模块负责，架构如图 6、图 7 所示。

图 6 机器学习任务创建界面

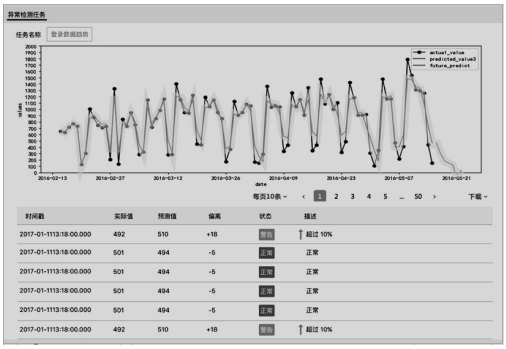

图 7 机器学习任务分析结果界面

2.9 系统管理功能

平台还具备其他常规管理功能，如存储空间管理、日志分组管理、用户权限管理和日志数据下载等[23]，此处不再赘述。

3 实现效果

大屏展示的实时动态数据如图 8 所示，最终可实现效果如下：

1) 实现自动化日志采集，并且支持定时/实时采集，全量、增量采集；

2) 实现日志合规存储，可按照用户预期设定保留周期，不仅满足保存 6 个月的要求，而且可通过 worm 固化实现日志数据防篡改；

3) 实现日志查询与分析，可实现秒级查询能力，快速获取所需的信息；

4) 结合现有系统与模板，输出定制化报表，实现日志合规审计；

5) 可扩展支持系统安全分析、业务分析等，为后续业务状态评估与优化打好基础。

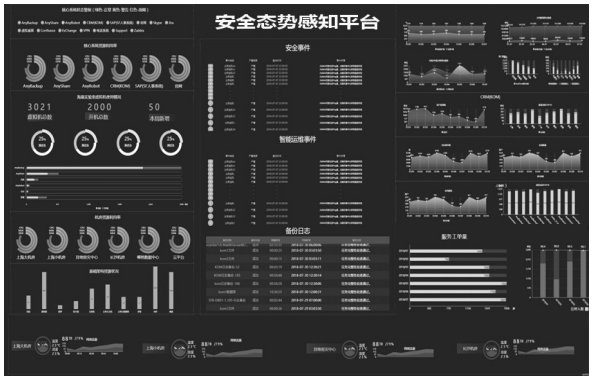

图 8 大屏展示实时动态数据

4 结束语

网络安全态势感知是近年来在信息安全领域新兴的一个技术方向，核心理念是构建安全威胁大数据分析平台。笔者论述了一种网络安全态势感知平台架构，通过数学建模与机器学习发现网络中潜在的入侵和攻击等安全威胁，建立针对未知网络威胁的动态安全监控与防御体系，从而避免因网络攻击导致的数据泄露、信息系统被破坏等安全问题。

参考文献：

[1] 刘念, 刘勇, 李涛, 等. 基于免疫的网络安全态势感知关键技术研究[J]. 四川大学学报(工程科学版), 2009, 41(6): 141-146.

[2] 韦勇, 连一峰, 冯登国. 基于信息融合的网络安全态势评估模型[J]. 计算机研究与发展, 2009, 46(3): 353-362.

[3] 韦勇, 连一峰. 基于日志审计与性能修正算法的网络安全态势评估模型[J]. 计算机学报, 2009, 32(4): 763-771.

[4] 胡威, 李建华, 陈秀真, 等. 可扩展的网络安全态势评价模型优化设计[J]. 电子科技大学学报, 2008, 38(1): 113-116.

[5] 王慧强, 赖积保, 朱亮, 等. 网络态势感知系统研究综述[J]. 计算机科学, 2006, 33(10): 5-10.

[6] 杨练. 网络安全态势评估与趋势感知的分析研究[J]. 信息化建设, 2016(8): 85.

[7] 赵国生, 王慧强, 王健. 基于灰色 Verhulst 的网络安全态势感知模型[J]. 哈尔滨工业大学学报, 2008, 40(5): 797-801.

[8] 刘效武, 王慧强, 梁颖, 等. 基于异质多传感器融合的网络安全态势感知模型[J]. 计算机科学, 2008, 135(18): 69-73.

[9] 姚军. 基于病例分型的医疗质量管理信息系统设计研究[D]. 上海: 复旦大学, 2009: 23-45.

[10] 赖积保, 王慧强, 朱亮. 网络安全态势感知模型研究[J]. 计算机研究与发展, 2006, 43(z2): 456-460.

[11] 梁颖, 王慧强, 赖积保. 一种基于粗糙集理论的网络安全态势感知方法[J]. 计算机科学, 2007, 34(8): 95-97.
[12] 张彬. 基于Spark大数据平台日志审计系统的设计与实现[D]. 济南: 山东大学, 2015: 14-17.
[13] 谷雨, 徐宗本, 孙剑, 等. 基于PCA与ICA特征提取的入侵检测集成分类系统[J]. 计算机研究与发展, 2006, 43(3): 633-638.
[14] 杨晨. 轻量级虚拟化高可用集群的研究与设计[D]. 上海: 上海交通大学, 2016: 27-29.
[15] 孙惟皓, 凌宗南, 陈炜忻. 日志智能分析在银行业 IT 安全运维管理中的应用[J]. 信息技术与网络安全, 2018, 4(7): 13-17.
[16] 岳剑. 基于态势感知技术的电子政务网络健康度评测平台搭建[J]. 中国信息化, 2018, 22(6): 44-48.
[17] 武宗涛, 赵进, 叶忠. 基于基线的 APT 检测分析平台研究与设计[J]. 网络安全技术与应用, 2017(9): 35-37.
[18] 吕威. 数据质量和隐私保护中聚类分类算法的应用研究[D]. 广州: 中山大学, 2008: 37-38.
[19] 郭山清, 谢立, 曾英佩. 入侵检测在线规则生成模型[J]. 计算机学报, 2006, 29(9): 1523-1532.
[20] 田大新, 刘衍珩, 李宾, 等. 基于动态分类算法的人侵检测系统[J]. 吉林大学学报(信息科学版), 2006, 24(2): 197-203.
[21] 陶新民, 陈万海, 郭黎利. 一种新的基于模糊聚类和免疫原理的入侵监测模型[J]. 电子学报, 2006, 34(7): 1329-1332.
[22] 苏璞睿, 冯登国. 基于进程行为的异常检测模型[J]. 电子学报, 2006, 34(10): 1809-1811.
[23] 屈立军. 集装箱码头商务信息管理系统的设计与实现[D]. 哈尔滨: 哈尔滨工业大学, 2011: 19-20.

（上接第 2 页）

3　实验验证

现行主要弹头涂色标机设备的生产节拍为80～100 发/min。在设计中经测算，从弹头涂完色标进入粘尖丝杠传送开始到最后进入下料滑槽时间为27～32 s，综合生产节拍能够满足要求。热风管道出风口可调温度经测量为 70～100 ℃，在此条件下进行弹头涂色标的实验。结果表明：色标不脱落，烘干迅速，从而得出的合理烘干温度和时间趋势如图 7 所示。

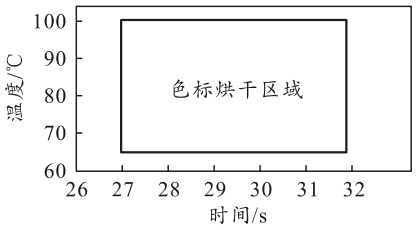

图 7　色标的烘干温度与时间趋势

按上述工艺条件和配置进行弹头涂色标的实际生产，然后将烘干的弹头进行检测，检测结果如表 1 所示，可以看出结果较理想，连续批量生产时弹头涂色标的合格率可达 98%以上，满足工艺要求。

表 1　弹头涂色标的检测结果

序号	检测项目	技术要求	检测结果	序号	检测项目	技术要求	检测结果
1	色标颜色	黄色	符合要求	4	色标附着力	不脱落	符合要求
2	色标高度	按标本	符合要求	5	色标完整性	完整均匀	符合要求
3	色标亮度	按标本	符合要求	6	烘干时间	≤30 s	符合要求

4　结束语

笔者设计的弹头涂色标工艺已应用在某兵器厂，实现了弹头尖部涂漆液均匀，涂色后不流挂、不堆色漆、无气泡、无缩孔。与直接运用加热管烤干相比，弹头色标受热更均匀，产生的刺激性气体经抽风处理排出，热风浴烘干安全快速，且烘干温度可调，范围为 70～100 ℃。烘干后的产品碰撞、挤压(弹头压入弹壳装配成成弹时)不会掉色漆，具有参考和借鉴价值。

参考文献：

[1] 段国发, 赵强, 宋建华. 某特种弹头涂双色工艺研究及应用[J]. 表面技术, 2004, 33(5): 63-66.
[2] 张怀智, 曹宏安, 黄鹏波. 炮弹标志自动印刷系统研究与开发[J]. 包装工程, 2011, 32(5): 26-28.
[3] 姚重阳, 郭鹏伟, 李爱华. 枕形包装机色标定位方法与实现[J]. 武汉工业学院学报, 2013, 32(4): 18-22.
[4] 张建蓉, 钱雄伟. 基于 PLC 及触摸屏的涂装前处理线控制[J]. 工业控制计算机, 2009, 22(2): 91-92.
[5] 杨青艳. 光电色标传感器的研究与应用[J]. 中国包装工业, 2013(10): 84.
[6] 范庆辉, 杨志清, 李作武, 等. 弹头自动装配机控制系统的设计与应用[J]. 兵工自动化, 2019, 39(8): 78-81.
[7] 陈月华, 张利萍, 邹志云. 智能电加热温控仪的设计制作及应用[J]. 控制工程, 2004, 11(S1): 22-24.
[8] 李敏, 孟臣. 基于 K 型热电偶的数字式温控仪设计[J]. 自动化仪表, 2004, 25(10): 21-23.
[9] 刘浩. 电加热温度控制系统设计与实验研究[D]. 杭州: 中国计量学院, 2013.