



(12) 发明专利申请

(10) 申请公布号 CN 105553957 A

(43) 申请公布日 2016. 05. 04

(21) 申请号 201510907420. 2

(22) 申请日 2015. 12. 09

(71) 申请人 国家电网公司

地址 100031 北京市西城区西长安街 86 号

申请人 国网内蒙古东部电力有限公司信息
通信分公司

(72) 发明人 刘世民 齐四清 孙添资 朱继阳
高敏 任春雷 王磊 樊锐
郭立勇

(74) 专利代理机构 北京风雅颂专利代理有限公司
11403

代理人 李弘 杨红梅

(51) Int. Cl.

H04L 29/06(2006. 01)

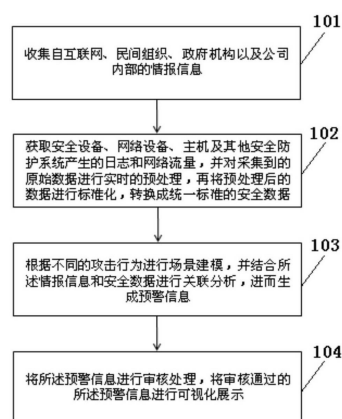
权利要求书2页 说明书7页 附图1页

(54) 发明名称

基于大数据的网络安全态势感知预警方法和系统

(57) 摘要

本发明公开了一种基于大数据的网络安全态势感知预警方法和系统,包括:收集自互联网、民间组织、政府机构以及公司内部的情报信息;获取安全设备、网络设备、主机及其他安全防护系统产生的日志和网络流量,并对采集到的原始数据进行实时的预处理,再将预处理后的数据进行标准化,转换成统一标准的安全数据;根据不同的攻击行为进行场景建模,并结合所述情报信息和安全数据进行关联分析,进而生成预警信息;将所述预警信息进行审核处理,将审核通过的所述预警信息进行可视化展示。本发明能够实现安全态势全面感知、安全威胁实施预警、安全事件及时处置和响应的能力,从而提升电力系统整体安全防护能力。



1. 一种基于大数据的网络安全态势感知预警方法,其特征在于,包括步骤:

收集自互联网、民间组织、政府机构以及公司内部的情报信息;

获取安全设备、网络设备、主机及其他安全防护系统产生的日志和网络流量,并对采集到的原始数据进行实时的预处理,再将预处理后的数据进行标准化,转换成统一标准的安全数据;

根据不同的攻击行为进行场景建模,并结合所述情报信息和安全数据进行关联分析,进而生成预警信息;

将所述预警信息进行审核处理,将审核通过的所述预警信息进行可视化展示。

2. 根据权利要求1所述的方法,其特征在于,所述情报信息包括:黑客攻击行为特征、漏洞库信息、信誉库信息。

3. 根据权利要求1所述的方法,其特征在于,所述预处理包括:数据去重、数据噪声去除、数据增强。

4. 根据权利要求1所述的方法,其特征在于,在进行所述标准化时,进一步从所述原始数据提取信息,提取的信息包括:设备基本信息、设备状态信息、网络安全事件、设备策略、通信命令。

5. 根据权利要求1所述的方法,其特征在于,所述结合所述情报信息和安全数据进行关联分析的步骤进一步包括:

通过预制多种不同的安全分析引擎,再借助大数据平台的超强计算能力,对海量内外网数据、事件、文件等进行关联分析和检索,实时在线检测、离线检测,发现高级的APT攻击和信息泄漏行为。

6. 一种基于大数据的网络安全态势感知预警系统,其特征在于,包括:

情报收集模块,用于收集自互联网、民间组织、政府机构以及公司内部的情报信息;

深度监测模块,用于获取安全设备、网络设备、主机及其他安全防护系统产生的日志和网络流量,并对采集到的原始数据进行实时的预处理,再将预处理后的数据进行标准化,转换成统一标准的安全数据;

数据分析模块,用于根据不同的攻击行为进行场景建模,并结合所述情报信息和安全数据进行关联分析,进而生成预警信息;

预警处置模块,用于将所述预警信息进行审核处理,将审核通过的所述预警信息进行可视化展示。

7. 根据权利要求6所述的系统,其特征在于,所述情报信息包括:黑客攻击行为特征、漏洞库信息、信誉库信息。

8. 根据权利要求6所述的系统,其特征在于,所述深度监测模块进一步用于:将采集到的原始数据进行数据去重、数据噪声去除、数据增强。

9. 根据权利要求6所述的系统,其特征在于,所述深度监测模块进一步用于:从所述原始数据提取信息,提取的信息包括:设备基本信息、设备状态信息、网络安全事件、设备策略、通信命令。

10. 根据权利要求6所述的系统,其特征在于,所述数据分析模块进一步用于:通过预制多种不同的安全分析引擎,再借助大数据平台的超强计算能力,对海量内外网数据、事件、文件等进行关联分析和检索,实时在线检测、离线检测,发现高级的APT攻击和信息泄漏行

为。

基于大数据的网络安全态势感知预警方法和系统

技术领域

[0001] 本发明涉及网络安全技术领域,特别是指一种基于大数据的网络安全态势感知预警方法和系统。

背景技术

[0002] 随着大数据、云计算、物联网、工业互联网等新兴互联网技术应用的不断深入,企业信息化程度也越来越高,对信息系统的依赖程度达到了前所未有的高度,与此同时,也导致了各种新型网络攻击、敏感信息泄露等恶意信息安全事件频繁发生。国家互联网应急中心调查显示,2015年涉及重要行业和政府部门的高危漏洞事件增多,基础应用或通用软件漏洞风险凸显,安全形势日趋严峻。特别是对于国家电网公司这样的特大型企业,企业信息系统规模属于全球企业前列,安全问题更加不容忽视,因为如果电力系统遭到网络安全攻击的威胁,则不单单是信息领域的安全问题,很有可能间接导致工业生产和社会生活的电力供应问题,从而影响国家安全。因此,为了不断应对新的安全挑战,国家电网公司先后部署了防火墙、UTM、IPS、IDS、漏洞扫描系统、防病毒系统、终端管理系统、WAF、DB-AUDIT以及安全监控平台等,构建起一道道安全防线。然而,形式并不乐观,现有电力系统的安全防御设施防御能力仍然不足,主要表现在以下三个方面:

[0003] 这些传统的安全产品都只能抵御来自某个方面的安全威胁,形成了一个的“安全防御孤岛”,缺乏对海量多维度的信息安全数据进行有效的融合关联分析,无法产生协同效应,不能使这些安全监测数据成为上层安全决策有效资源。

[0004] 这些传统的安全防御设施大多数都通过分析某些安全设备的日志对已经发生的攻击行为进行分析和监测,基本都是被动防御的思路,缺乏网络安全态势感知与联动预警的能力,当检测到网络攻击事件之后再采取相应的应急措施,往往为时已晚,因为此时网络攻击已经发生过去了,攻击已经造成了不可挽回的损失。

[0005] 这些复杂的IT资源及其安全防御设施在运行过程中不断产生大量的安全日志和事件,形成了大量的“信息孤岛”,有限的安全管理人员面对这些数量巨大、彼此割裂的安全大数据,操作着各种产品自身的控制台界面和告警窗口,显得束手无策,工作效率极低,难以发现真正的安全隐患。

发明内容

[0006] 有鉴于此,本发明的目的在于提出一种基于大数据的网络安全态势感知预警方法和系统,解决了传统安全防御手段存在的不足。

[0007] 基于上述目的本发明提供一种基于大数据的网络安全态势感知预警方法,包括步骤:

[0008] 收集自互联网、民间组织、政府机构以及公司内部的情报信息;

[0009] 获取安全设备、网络设备、主机及其他安全防护系统产生的日志和网络流量,并对采集到的原始数据进行实时的预处理,再将预处理后的数据进行标准化,转换成统一标准

的安全数据；

[0010] 根据不同的攻击行为进行场景建模，并结合所述情报信息和安全数据进行关联分析，进而生成预警信息；

[0011] 将所述预警信息进行审核处理，将审核通过的所述预警信息进行可视化展示。

[0012] 优选的，所述情报信息包括：黑客攻击行为特征、漏洞库信息、信誉库信息。

[0013] 优选的，所述预处理包括：数据去重、数据噪声去除、数据增强。

[0014] 优选的，在进行所述标准化时，进一步从所述原始数据提取信息，提取的信息包括：设备基本信息、设备状态信息、网络安全事件、设备策略、通信命令。

[0015] 优选的，所述结合所述情报信息和安全数据进行关联分析的步骤进一步包括：

[0016] 通过预制多种不同的安全分析引擎，再借助大数据平台的超强计算能力，对海量内外网数据、事件、文件等进行关联分析和检索，实时在线检测、离线检测，发现高级的APT攻击和信息泄漏行为。

[0017] 本发明还提供了一种基于大数据的网络安全态势感知预警系统，包括：

[0018] 情报收集模块，用于收集自互联网、民间组织、政府机构以及公司内部的情报信息；

[0019] 深度监测模块，用于获取安全设备、网络设备、主机及其他安全防护系统产生的日志和网络流量，并对采集到的原始数据进行实时的预处理，再将预处理后的数据进行标准化，转换成统一标准的安全数据；

[0020] 数据分析模块，用于根据不同的攻击行为进行场景建模，并结合所述情报信息和安全数据进行关联分析，进而生成预警信息；

[0021] 预警处置模块，用于将所述预警信息进行审核处理，将审核通过的所述预警信息进行可视化展示。

[0022] 优选的，所述情报信息包括：黑客攻击行为特征、漏洞库信息、信誉库信息。

[0023] 优选的，所述深度监测模块进一步用于：将采集到的原始数据进行数据去重、数据噪声去除、数据增强。

[0024] 优选的，所述深度监测模块进一步用于：从所述原始数据提取信息，提取的信息包括：设备基本信息、设备状态信息、网络安全事件、设备策略、通信命令。

[0025] 优选的，所述数据分析模块进一步用于：通过预制多种不同的安全分析引擎，再借助大数据平台的超强计算能力，对海量内外网数据、事件、文件等进行关联分析和检索，实时在线检测、离线检测，发现高级的APT攻击和信息泄漏行为。

[0026] 从上面所述可以看出，本发明提供的基于大数据的网络安全态势感知预警方法和系统，通过发现、预警、处置外部入侵事件、内部泄密事件、特权滥用事件保障电力系统业务安全运营。通过威胁情报共享、大数据、智能分析技术驱动网络安全态势感知与预警分析系统实现安全态势全面感知、安全威胁实施预警、安全事件及时处置和响应的能力，从而提升电力系统整体安全防护能力。

附图说明

[0027] 为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本

发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0028] 图1为本发明实施例的基于大数据的网络安全态势感知预警方法流程图;

[0029] 图2为本发明实施例的基于大数据的网络安全态势感知预警系统结构示意图。

具体实施方式

[0030] 为使本发明的目的、技术方案和优点更加清楚明白,以下结合具体实施例,并参照附图,对本发明进一步详细说明。

[0031] 本发明实施例提供了一种基于大数据的网络安全态势感知预警方法。参考图1,为本发明实施例的基于大数据的网络安全态势感知预警方法流程图。

[0032] 所述基于大数据的网络安全态势感知预警方法,包括以下步骤:

[0033] 步骤101、收集自互联网、民间组织、政府机构以及公司内部的情报信息。

[0034] 本实施例中。情报信息主要是黑客攻击行为特征、漏洞库、信誉库、其他安全威胁信息等。情报来源有两种,一种是来自互联网漏洞平台所发布的一些信息通过机器采集完成自动收集,另一种是通过人工录入;机器采集自动收集的情报信息需要通过验证后才能完成提交,生成一条真实的情报数据;所有生成的情报数据均需要和资产进行关联分析。

[0035] 具体的,通过数据采集器收集核心交换机镜像流量、NetFlow日志、安全设备Syslog日志以及文件等原始安全数据和漏洞信息、病毒库信息、攻击行为特征等情报数据,将收集的原始数据进行缓存并上传到数据预处理及存储层。数据采集需要根据不同的数据来源分别采用不同采集方式进行数据采集,比如一些网络安全设备,如防火墙、防病毒系统、IDS、IPS等,它们上报的原始信息需要被实时地采集和处理;而另一些网络安全设备,如漏洞扫描设备、流量分析设备、资产信息等,它们在一个时间段内的变化很小,不需要进行频繁地采集,更适合定点采集或者是当数据分析模块发出请求是才进行采集,这样可以节省获取数据所带来的系统处理时间开销和网络资源开销。面对不同安全产品对实时性的需求不一样,理想的方案是分开处理。对于需要被实时处理的原始信息,采用主动采集,实时上报;对于变化力度不大的安全信息进行定点采集或者是当数据分析模块发出请求时进行采集。根据对实时性的需求不一样,上报的原始事件可以分两类,被动采集事件和主动采集事件。

[0036] 步骤102、获取安全设备、网络设备、主机及其他安全防护系统产生的日志和网络流量,并对采集到的原始数据进行实时的预处理,再将预处理后的数据进行标准化,转换成统一标准的安全数据。

[0037] 首先获取安全设备、网络设备、主机及其他安全防护系统产生的日志和网络流量,并对采集到的原始数据进行实时的预处理和分析,对原始数据的预处理包括数据去重、数据噪声去除、数据增强,再将预处理后的数据进行标准化,转换成统一标准的安全数据。

[0038] 原始数据信息被送到数据预处理及存储层先要进行数据去重、数据标准化和数据加强,其中数据去重是确保所收集的数据均是唯一可靠的可信数据。

[0039] 其中数据去重的原理就是通过分析“脏数据”的产生原因和存在形式,利用现有的技术手段和方法去清洗“脏数据”,将“脏数据”转化为满足数据质量或应用要求的数据,从而提高数据集的数据质量。数据去重的手段有:去除源数据集中的噪声数据和无关数据,处

理遗漏数据和清洗“脏数据”，去除空白数据域和知识背景上的白噪声，考虑时间顺序和数据变化等，完成重复数据处理和默认数据处理，完成数据类型的转换。

[0040] 数据标准化是按照系统提前设定好的标准化规则将来自不同数据源的多源异构信息安全监测数据做数据归一化，实现数据映射。从共性角度看，需要从原始事件提取的信息包括以下几个方面：

[0041] (1)设备基本信息：基本信息包括设备的一些不经常变更的信息，仅在需要时上报，如设备名称、设备ID、版本信息等，必要时也可包括一些用户信息，如系统管理员、单位名称等。

[0042] (2)设备状态信息：状态信息包含设备运行状态信息，是一些经常变动的信息，需要定时上报以便于运维人员及时了解情况，如设备是否在线、CPU使用率、内存使用率等。

[0043] (3)网络安全事件：事件的上报是网络安全事件关联的重要组成部分，是进行下一步关联和分析的基础。

[0044] (4)设备策略：在各种网络安全设备中，安全策略是获取安全事件的重要依据，运维人员需要通过了解安全策略掌握安全状况。

[0045] (5)通信命令：许多设备通过代理、控制台的方式与安管平台进行通信，需要定义合理的命令保证各种信息的顺利传递。

[0046] 同时，标准化的安全事件数据格式还具有很好的扩展性和兼容性。系统建立了一套数据采集的流程规范，不但能够采集当前已有的设备或系统数据，而且对于将来加入的新的安全设备或安全监控系统都可以很容易实现扩展。

[0047] 数据加强是通过资产信息关联、外部漏洞库、情报库关联等对标准化后数据进行有条件加强，根据后续业务分析的需要有针对性的对部分数据进行加强。

[0048] 经过数据预处理后的数据需要进入分布式存储进行存储，分结构化存储和非结构化存储两类，其中结构化存储主要用于存储来自于外部信誉库、系统配置库、系统自身结构化数据等数据，非结构存储主要用于存储来自于Syslog、Flow、文件等内容数据。最后需要为所有存储的数据创建数据索引，一遍后续查询追溯使用。

[0049] 步骤103、根据不同的攻击行为进行场景建模，并结合所述情报信息和安全数据进行关联分析，进而生成预警信息。

[0050] 安全运维人员首先将现有的场景规则化，然后按照场景规则创建场景模型，并在场景模型中内置分析判断规则，通过相应的规则对标准化后的安全数据进行实时和离线关联分析，根据关联分析的结果产生相应的告警信息。

[0051] 本实施例的方法通过预制多种不同的安全分析引擎，如攻击分析引擎、脆弱性分析引擎、安全策略分析引擎、关联分析引擎等，再借助大数据平台的超强计算能力，对海量内外网数据、事件、文件等进行关联分析和检索，实时在线检测、离线检测，发现高级的APT攻击和信息泄漏行为。

[0052] 步骤104、将所述预警信息进行审核处理，将审核通过的所述预警信息进行可视化展示。

[0053] 所有的告警信息和预警信息均需要上传至总部管理中心进行研判，安全运维人员若发现预警或者告警是误报，则将其关闭；若发现是一般威胁，则将其提交为通告，如果是高危威胁，则将其提交为事件。事件提交后，审核员进行审核，审核通过的，则转为事件，状

态为待处置;待处置的事件需要下级单位进行处置,并将最终的处置结果提交到总部管理中心;最后总部管理中心的安全运维人员需要对提交的处置结果进行审核,如果审核不通过,需要重新处置,如果提交成果,则关闭事件处置流程。

[0054] 进一步的,采用可配置图形分析工具提供优秀的可视化展示功能,主要包括实时态势展示、安全态势预测、专项态势展示、历史状态展示、其他信息展示等。实时态势展示主要通过实时的数据信息,展示当前正在发生的实时安全行为,让安全运维人员对当前网络环境正在发生的安全行为有整体的实时的监测和感知能力,包括:安全攻击展示、安全漏洞展示、安全风险展示、区域转换;安全态势预测是通过历史存储的数据,结合当前的实时数据对未来各种指标的发展趋势进行预测,包括:攻击路径预测、攻击方式预测、攻击时间预测等,通过自动化的方式提供未来可能发生攻击行为;专项态势展示主要针对内部网络重大安全威胁与隐患进行可视化展示,包括:异常外联展示、事件回溯展示、隐秘通道展示、其他专项展示等;实时安全态势数据进行收集与存储形成历史安全状态数据,包括:安全攻击趋势、安全漏洞趋势、安全风险趋势。系统会根据设定发出通告信息、预警信息、告警信息、事件信息。这些信息按类别进行分类展示,还可通过查询来快速找到所需信息。

[0055] 另一方面,本发明实施例还提供了一种基于大数据的网络安全态势感知预警系统。参考图2,为本发明实施例的基于大数据的网络安全态势感知预警系统结构示意图。

[0056] 所述基于大数据的网络安全态势感知预警系统,包括:

[0057] 情报收集模块,用于收集自互联网、民间组织、政府机构以及公司内部的情报信息;

[0058] 深度监测模块,用于获取安全设备、网络设备、主机及其他安全防护系统产生的日志和网络流量,并对采集到的原始数据进行实时的预处理,再将预处理后的数据进行标准化,转换成统一标准的安全数据;

[0059] 数据分析模块,用于根据不同的攻击行为进行场景建模,并结合所述情报信息和安全数据进行关联分析,进而生成预警信息;

[0060] 预警处置模块,用于将所述预警信息进行审核处理,将审核通过的所述预警信息进行可视化展示。

[0061] 对于本实施例中的情报收集模块,最底层的数据源作为系统的数据输入,包括来自核心交换机的镜像流量、NetFlow、安全设备Syslog日志(防火墙、IPS、IDS等安全设备)以及其他安全监控系统的告警数据等,除此之外,数据源还应包括信息网络中所有资产的配置信息(如资产库)与外部情报(如漏洞库、病毒库、信誉库等)。输入的数据按照不同种类进入到数据收集层,采用分布式数据收集组件Flume对非结构化数据进行收集,收集到的数据暂时缓存到Kafka组件中以便后续数据处理使用。结构化数据通过Webservice进行收集,并存储到MySQL中。

[0062] 对于本实施例中的深度监测模块,通过流计算组件Storm对收集到的数据进行预处理,包括数据去重、数据噪声去除、数据增强。

[0063] 数据存储包括结构化存储和非结构化存储两部分。

[0064] 结构化存储主要存储业务功能模块数据及环境数据等产生的结构化数据。包括:情报数据存储、场景建模脚本存储、告警信息存储、事件及工单信息存储、指标信息存储、流程管理信息存储、系统管理信息存储以及环境数据存储等。

[0065] 非结构化存储主要存储平台处理过程中的原始数据、过程数据和结果数据,其中索引数据存储到Elastic search组件中。邮件的附件、沙箱运行报告和重度汇总数据存储到数据仓库Nana中;分布式数据库HBase用于存储原始数据(加强)以及轻度汇总后所产生的数据。

[0066] 对于本实施例中的数据分析模块,数据分析层主要包括大数据计算和大数据分析两部分。分析部分包括用图形化数据挖掘工具实现的可视化辅助分析、用ElasticSearch和Lucense实现的数据搜索分析、用数据融合组件Candy对分散保存在各处的结构化和非结构化数据进行融合汇总分析。大数据计算部分包括Nana组件对离线历史数据的挖掘以及Storm组件对数据流的实时计算。

[0067] 对于本实施例中的预警处置模块,数据展现层主要采用可配置图形分析工具Cherry提供优秀的可视化展示功能,主要包括实时态势展示、安全态势预测、专项态势展示、历史状态展示、其他信息展示等。包括安全态势、场景建模、事件处置、情报管理、告警分析、指标管理、流程管理和系统管理等。

[0068] 基于上述实施例可见,相比于现有技术,本发明至少包括以下优点:

[0069] (1)本发明采用开源的分布式计算架构,实现了海量安全信息数据的收集、存储、分析和展现。基于大数据的网络安全态势感知与安全预警分析系统能够处理PB级数据,并具有高可靠、高扩展、高效和高容错等特点。通过高速处理技术实现了对海量数据的处理,并且系统采用了分布式文件索引技术,保证了海量数据的处理。

[0070] (2)基于大数据的网络安全态势感知与安全预警分析系统将各类分析引擎的计算过程分散到不同的计算节点中,实现了分布式计算,从而为大数据手机集、存储、分析和展现提供了强有力的物质基础。通过分布式计算技术,系统可以将数据收集、存储、分析的功能均衡分配在分布式计算节点中,为了适应更高的速度,只须扩充计算节点数量即可。分布式架构由于其计算分散技术,单点失效不影响整体能力,因而具备了高可靠性和高容错性。

[0071] (3)基于大数据的网络安全态势感知与安全预警分析系统支持多种数据源类型,支持对结构化和非结构化数据的收集,具备针对多源异构数据的关联分析能力。系统实现了对所收集的数据的预处理和存储,将需要的数据转换为结构化数据,对非结构化数据进行索引和存储,将数据分别送至分布式文件系统中供分析层各分析引擎使用。丰富的基础安全数据,保障了数据源的多样性,为多种分析方法的结合和综合关联分析提供了数据基础。

[0072] (4)基于大数据的网络安全态势感知与安全预警分析系统支持对存储在分布式计算存储节点和数据库中的历史数据进行历史分析,结合多种数据挖掘算法及时发现过去未发现的安全问题,帮助安全运维人员进行调查分析,及时采取有效措施消除安全隐患。历史分析针对存储在分布式存储系统中的历史数据进行,可实现追溯分析、取证分析、查询统计,有效的弥补了传统数据库技术效率低下的问题。

[0073] (5)基于大数据的网络安全态势感知与安全预警分析系统通过丰富的可视化展示组件和功能,可对当前网络安全态势做出直观的展示,并结合相应的情报信息,对未来的一段时间的安全趋势以及安全威胁做出预警提示,便于安全运维人员及早采取措施,改善了传统安全运维平台早期预警能力偏弱的问题。

[0074] 所属领域的普通技术人员应当理解:以上任何实施例的讨论仅为示例性的,并非

旨在暗示本公开的范围(包括权利要求)被限于这些例子;在本发明的思路下,以上实施例或者不同实施例中的技术特征之间也可以进行组合,步骤可以以任意顺序实现,并存在如上所述的本发明的不同方面的许多其它变化,为了简明它们没有在细节中提供。因此,凡在本发明的精神和原则之内,所做的任何省略、修改、等同替换、改进等,均应包含在本发明的保护范围之内。

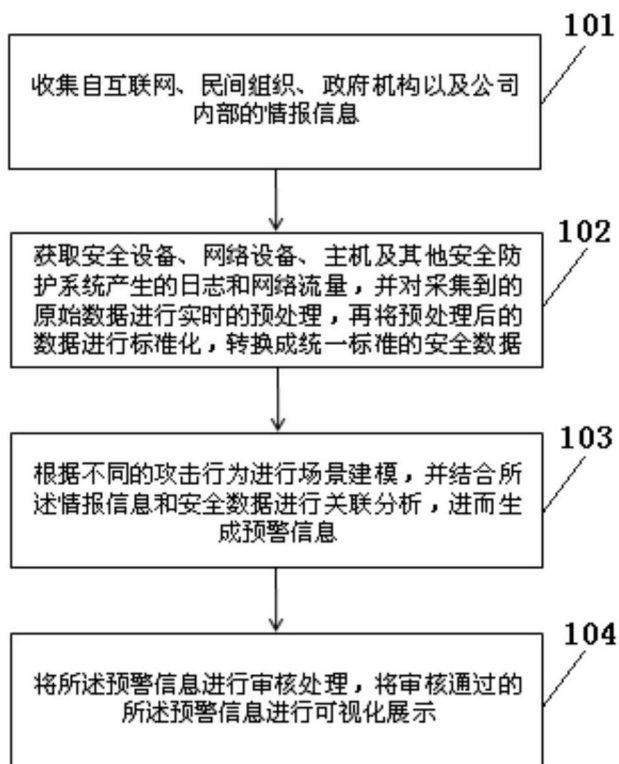


图1

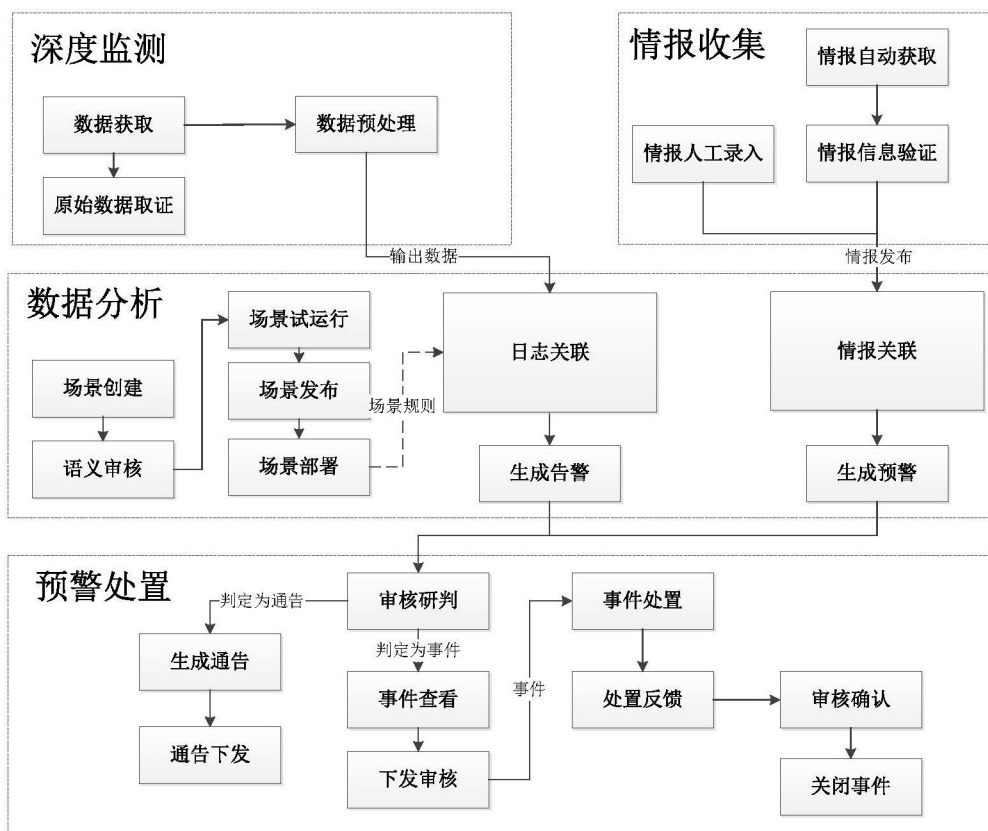


图2