(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 112527888 B
(45)授权公告日 2024.04.05

(21)申请号 202011547280.X

(22)申请日 2020.12.24

(65)同一申请的已公布的文献号
申请公布号 CN 112527888 A

(43)申请公布日 2021.03.19

(73)专利权人 恒安嘉新(北京)科技股份公司
地址 100098 北京市海淀区北三环西路25号27号楼五层5002室

(72)发明人 叶辉 蔡琳 杨满智 王杰 孟宝权 王伟 范磊波 梁彧 田野 傅强 金红 陈晓光

(74)专利代理机构 北京品源专利代理有限公司 11332
专利代理师 孟金喆

(51)Int.Cl.
G06F 16/26(2019.01)
G06F 16/9035(2019.01)
G06F 9/54(2006.01)
G06F 11/30(2006.01)

(56)对比文件
CN 103793859 A,2014.05.14
CN 111726358 A,2020.09.29
CN 110458743 A,2019.11.15
CN 108040493 A,2018.05.15
CN 110443037 A,2019.11.12
CN 110598180 A,2019.12.20
CN 111160738 A,2020.05.15
US 2019222604 A1,2019.07.18
US 9571510 B1,2017.02.14
介贺彤.基于事故树分析法的电网企业安全生产风险管控系统.《电力信息与通信技术》.2019,第17卷(第6期),25-30.
P.A.S. Ralston et al..Cyber security risk assessment for SCADA and DCS networks.《ISA Transactions》.2007,583-594.

审查员 何承恩

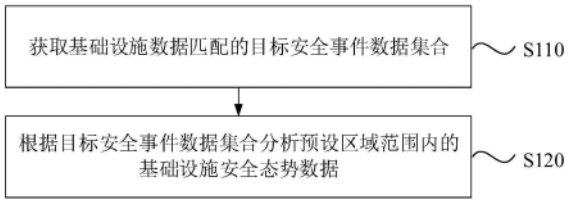
权利要求书2页 说明书11页 附图3页

(54)发明名称

一种数据分析方法、装置、电子设备及存储介质

(57)摘要

本发明实施例公开了一种数据分析方法、装置、电子设备及存储介质。所述数据分析方法,包括:获取基础设施数据匹配的目标安全事件数据集合;基础设施数据为基础设施单位的设施数据;根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据。本发明实施例的技术方案提高了基础设施安全事件数据的综合分析能力,丰富了数据分析维度。



CN 112527888 B

1.一种数据分析方法,其特征在于,包括:

获取基础设施数据匹配的目标安全事件数据集合;所述基础设施数据为基础设施单位的设施数据;

根据所述目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据;

所述获取基础设施数据匹配的目标安全事件数据集合,包括:

获取所述基础设施单位的基础设施网络数据;

根据安全事件定位数据获取目标区域范围内的安全事件数据集合;

根据所述基础设施网络数据和所述安全事件数据集合确定所述目标安全事件数据集合;

所述根据所述目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据,包括:

确定基础设施安全指数模型;

将所述目标安全事件数据集合输入至所述基础设施安全指数模型中,以根据所述基础设施安全指数模型的输出结果确定所述基础设施安全态势数据;

所述确定基础设施安全指数模型,包括:

根据安全事件类型和基础设施单位类型确定二维权重矩阵;

根据所述二维权重矩阵确定各类安全事件类型对应所述基础设施单位类型的权重系数;

根据所述权重系数确定所述基础设施安全指数模型。

2.根据权利要求1所述的方法,其特征在于,所述安全事件定位数据包括IP和/或域名数据;所述基础设施网络数据包括基础设施的IP和/或域名数据。

3.根据权利要求1所述的方法,其特征在于,所述根据所述权重系数确定所述基础设施安全指数模型,包括:

确定所述各类安全事件类型或所述基础设施单位类型对应的目标安全事件数量以及目标基础设施单位数量;

根据所述权重系数、所述目标安全事件数量和所述目标基础设施单位数量确定所述基础设施安全指数模型。

4.根据权利要求3所述的方法,其特征在于,所述根据所述权重系数、所述目标安全事件数量和所述目标基础设施单位数量确定所述基础设施安全指数模型,包括:

基于如下公式确定所述基础设施安全指数模型:

$$Score=\sum_{i=1}^{A} W_i*\left(B-\ln\left(\frac{X_i}{N}+1\right)*C\right)$$

其中,Score表示所述基础设施安全态势数据W_i,表示所述权重系数、X_i表示所述目标安全事件数量,N表示所述目标基础设施单位数量,A、B和C为常数。

5.根据权利要求1所述的方法,其特征在于,所述获取基础设施数据匹配的目标安全事件数据集合,包括:

确定数据分析时间窗口;

根据所述数据分析时间窗口获取所述基础设施数据匹配的目标安全事件数据集合。

6.一种数据分析装置,其特征在于,包括:

数据集合获取模块,用于获取基础设施数据匹配的目标安全事件数据集合;所述基础设施数据为基础设施单位的设施数据;

安全态势数据分析模块,用于根据所述目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据;

数据集合获取模块,具体用于:

获取所述基础设施单位的基础设施网络数据;

根据安全事件定位数据获取目标区域范围内的安全事件数据集合;

根据所述基础设施网络数据和所述安全事件数据集合确定所述目标安全事件数据集合;

安全态势数据分析模块,具体用于:

确定基础设施安全指数模型;

将所述目标安全事件数据集合输入至所述基础设施安全指数模型中,以根据所述基础设施安全指数模型的输出结果确定所述基础设施安全态势数据;

安全态势数据分析模块,具体用于:

根据安全事件类型和基础设施单位类型确定二维权重矩阵;

根据所述二维权重矩阵确定各类安全事件类型对应所述基础设施单位类型的权重系数;

根据所述权重系数确定所述基础设施安全指数模型。

7.一种电子设备,其特征在于,所述电子设备包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-5中任一所述的数据分析方法。

8.一种计算机存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1-5中任一所述的数据分析方法。

一种数据分析方法、装置、电子设备及存储介质

技术领域

[0001] 本发明实施例涉及互联网安全技术领域，尤其涉及一种数据分析方法、装置、电子设备及存储介质。

背景技术

[0002] 随着互联网技术的发展，基础设施的互联网通信安全越来越受到重视。而基于基础设施的数据分析是基础设施互联网通信安全的重要技术要点。

[0003] 现有技术中，针对全国基础设施、各省基础设施、各城市基础设施、以及各行业基础设施的安全事件分析的数据挖掘深度以及数据综合分析能力不足。不能实现对基础设施的安全事件的量化分析。

发明内容

[0004] 本发明实施例提供一种数据分析方法、装置、设备及存储介质，提高了基础设施安全事件数据的综合分析能力，丰富了数据分析维度。

[0005] 第一方面，本发明实施例提供了一种数据分析方法，包括：

[0006] 获取基础设施数据匹配的目标安全事件数据集合；基础设施数据为基础设施单位的设施数据；

[0007] 根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据。

[0008] 第二方面，本发明实施例还提供了一种数据分析装置，包括：

[0009] 数据集合获取模块，用于获取基础设施数据匹配的目标安全事件数据集合；基础设施数据为基础设施单位的设施数据；

[0010] 安全态势数据分析模块，用于根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据。

[0011] 第三方面，本发明实施例还提供了一种电子设备，电子设备包括：

[0012] 一个或多个处理器；

[0013] 存储装置，用于存储一个或多个程序；

[0014] 当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现本发明任意实施例所提供的数据分析方法。

[0015] 第四方面，本发明实施例还提供了一种计算机存储介质，其上存储有计算机程序，该程序被处理器执行时实现本发明任意实施例所提供的数据分析方法。

[0016] 本发明实施例，在确定与获取的基础设施数据匹配的目标安全事件数据集合之后，进一步根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据，解决了现有技术对基础设施安全事件的数据挖掘深度不足以及数据综合分析能力较差的问题，达到了根据业务需求在预设区域范围内对基础设施的安全事件进行分析的效果，提高了基础设施安全事件数据的综合分析能力，丰富了数据分析维度。

附图说明

[0017] 图1是本发明实施例一提供的一种数据分析方法的流程图；

[0018] 图2是本发明实施例二提供的一种数据分析方法的流程图；

[0019] 图3是本发明实施例二提供的一种多维数据分析示意图；

[0020] 图4是本发明实施例三提供的一种数据分析装置的示意图；

[0021] 图5为本发明实施例四提供的一种电子设备的结构示意图。

具体实施方式

[0022] 下面结合附图和实施例对本发明作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释本发明，而非对本发明的限定。

[0023] 另外还需要说明的是，为了便于描述，附图中仅示出了与本发明相关的部分而非全部内容。在更加详细地讨论示例性实施例之前应当提到的是，一些示例性实施例被描述成作为流程图描绘的处理或方法。虽然流程图将各项操作(或步骤)描述成顺序的处理，但是其中的许多操作可以被并行地、并发地或者同时实施。此外，各项操作的顺序可以被重新安排。当其操作完成时所述处理可以被终止，但是还可以具有未包括在附图中的附加步骤。所述处理可以对应于方法、函数、规程、子例程、子程序等等。

[0024] 实施例一

[0025] 图1是本发明实施例一提供的一种数据分析方法的流程图，本实施例可适用于根据对一定区域范围内的基础设施安全态势数据进行分析的情况，该方法可以由数据分析装置来执行，该装置可以由软件和/或硬件的方式来实现，并一般可集成在电子设备中。相应的，如图1所示，该方法包括如下操作：

[0026] S110、获取基础设施数据匹配的目标安全事件数据集合。

[0027] 其中，基础设施数据为基础设施单位的设施数据。

[0028] 其中，基础设施单位可以是基础设施下属的单位。例如，基础设施可以包括电信、广播、能源、金融、运输、铁路、民航、邮政、水利、应急、卫生、社会、国防、政府、教育、工业、互联、新闻、环境、公共、食品及化工等市政公用工程设施和公共生活服务设施。基础设施单位的设施数据可以是基础设施单位所有设施的相关数据。示例性的，基础设施数据可以包括但不限于基础设施的名称、基础设施的网络系统数据以及基础设施的身份标识信息等。基础设施的网络系统数据可以是基础设施在互联网建立的相关服务器的系统数据。基础设施的身份标识信息可以是基础设施身份的唯一标识，与基础设施一一对应。基础设施的身份标识信息可以用于对基础设施进行定位。目标安全事件数据集合可以是与基础设施数据匹配的安全事件数据集合。安全事件数据可以是互联网中威胁网络安全的事件数据。

[0029] 在本发明实施例中，在获取基础设施数据匹配的目标安全事件数据集合之前，首先获取基础设施单位的设施数据，对基础设施单位的设施数据进行数据解析以及数据处理，进一步将经过数据处理的基础设施数据与安全事件数据集合进行匹配，与安全事件数据集合匹配成功的基础设施数据对应的安全事件数据集合，作为目标安全事件数据集合。

[0030] 示例性的，基础设施单位的设施数据获取的具体过程为：首先获取基础设施单位的相关数据，例如，获取基础设施单位的单位名称、单位备案、单位地址、联系人、归属的省、城市以及县等基础信息。进一步根据基础设施单位的相关数据获取对应单位的设施数据，

例如该单位的网站、服务器的IP(Internet Protocol,国际互联协议)以及网络设备的域名和网络设备参数等数据。本发明实施例对基础设施单位的设施数据的具体数据内容不做限定。

[0031] 在本发明的一个可选实施例中,获取基础设施数据匹配的目标安全事件数据集合,可以包括:获取基础设施单位的基础设施网络数据;根据安全事件定位数据获取目标区域范围内的安全事件数据集合;根据基础设施网络数据和安全事件数据集合确定目标安全事件数据集合;其中,安全事件定位数据包括IP和/或域名数据;基础设施网络数据包括基础设施的IP和/或域名数据。

[0032] 其中,基础设施网络数据可以是基础设施进行互联网通信的网络数据。安全事件定位数据可以是对目标安全事件数据集合所属区域范围进行定位的数据。目标区域范围可以是需要进行目标安全事件数据集合分析的区域范围。目标区域范围可以包括但不限于目标安全事件数据集合所属地理区域范围,以及目标安全事件数据集合所属类型范围。安全事件数据集合可以是涉及安全事件相关数据的数据集合。例如,安全事件数据集合可以包括但不限于安全事件的类型、发生地点、发生时间以及定位数据等。

[0033] 在本发明实施例中,确定目标安全事件数据集合的具体过程为:首先根据基础设施单位的设施数据解析出基础设施单位的基础设施网络数据,进一步根据安全事件定位数据获取目标区域范围内的安全事件数据集合,以根据基础设施网络数据和安全事件数据集合进行数据匹配,与安全事件数据集合匹配成功的基础设施网络数据对应的安全事件数据集合,作为目标安全事件数据集合。其中,可以用于确定目标安全事件数据集合的基础设施网络数据可以包括基础设施的IP和/或域名数据。可以用于确定目标区域范围内的安全事件数据集合的安全事件定位数据,可以包括IP和/或域名数据。

[0034] 示例性的,根据向全国、各省、各市以及各行业下发的数据上报指令获取相应的基础设施单位的设施数据,通过解析基础设施单位的设施数据获取基础设施的IP和/或域名数据,进一步通过全国、各省、各市以及各行业的安全事件数据集合获取相应安全事件的IP和/或域名数据,以根据基础设施的IP和/或域名数据与安全事件数据集合的匹配结果,确定基础设施的安全事件数据集合。具体的,将匹配成功的基础设施的IP和/或域名数据对应的基础设施的安全事件数据集合作为目标安全事件数据集合。其中,安全事件数据集合所对应的安全事件的类型可以包括但不限于僵尸木马、DDOS(Distributed Denial Of Service,分布式拒绝服务)攻击、钓鱼、放马、恶意代码传播、恶意邮件、鬼影、蠕虫、网页篡改以及Web(Webite,网页)后门等。

[0035] 在本发明的一个可选实施例中,获取基础设施数据匹配的目标安全事件数据集合,可以包括:确定数据分析时间窗口;根据数据分析时间窗口获取基础设施数据匹配的目标安全事件数据集合。

[0036] 其中,数据分析时间窗口可以是某一时间范围。数据分析时间窗口可以用于确定某一时间范围内的目标安全事件数据集合。

[0037] 在本发明实施例中,在获取基础设施数据之前,可以首先确定数据分析时间窗口,即确定获取基础设施数据的时间范围,在获取数据分析时间窗口内的基础设施数据之后,将与基础设施数据匹配的安全事件数据集合确定为目标安全事件数据集合。

[0038] S120、根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数

据。

[0039] 其中,预设区域范围可以是根据业务需要预先设定的区域范围。区域范围可以包括地理区域范围以及类型区域范围等,本发明实施例对区域范围的划分类型不做具体限定。基础设施安全态势数据可以是表征基础设施安全情况的数据。

[0040] 在本发明实施例中,在获取基础设施数据匹配的目标安全事件数据集合之后,确定需要进行安全态势分析的区域范围即预设区域范围,以进一步获取预设区域范围内的目标安全事件数据,根据目标安全事件数据对应的基础设施的安全态势数据对该预设区域范围的安全态势进行分析。预设区域范围可以根据实际需求确定,如全国范围、全省范围或全市范围等,本发明实施例并不对预设区域范围的具体范围数据进行限定。

[0041] 在本发明实施例中,可以分析在全国、各省以及各市内的目标安全事件数据集合对应的基础设施安全态势数据。根据基础设施安全态势数据可以对相应的区域范围(全国、各省以及各市)的基础设施单位给出具有针对性的指导,并加强关键安全事件监控。还可以分析各行业的目标安全事件数据集合对应的基础设施安全态势数据。因此,本发明实施例可以实现对基础设施安全态势数据从所属地理范围以及所属行业等方面进行数据分析,将出现相同安全事件的基础设施单位进行关联分析,提升了基础设施安全事件的数据综合分析能力,并扩展了数据维度。

[0042] 本实施例的技术方案,在确定与获取的基础设施数据匹配的目标安全事件数据集合之后,进一步根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据,解决了现有技术对基础设施安全事件的数据挖掘深度不足以及数据综合分析能力较差的问题,达到了根据业务需求在预设区域范围内对基础设施的安全事件进行分析的效果,提高了基础设施安全事件数据的综合分析能力,丰富了数据分析维度。

[0043] 实施例二

[0044] 图2是本发明实施例二提供的一种数据分析方法的流程图,本实施例以上述实施例为基础进行具体化,在本实施例中,给出了根据所述目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据的具体可选的实施方案,相应的,如图2所示,该方法包括如下操作:

[0045] S210、获取基础设施数据匹配的目标安全事件数据集合;基础设施数据为基础设施单位的设施数据。

[0046] S220、根据安全事件类型和基础设施单位类型确定二维权重矩阵。

[0047] 其中,安全事件类型可以是安全事件所属范围的类型,安全事件类型可以用于对安全事件进行分类。基础设施单位类型可以是基础设施单位所属基础设施的类型。二维权重矩阵可以是基于安全事件类型和基础设施单位类型建立的安全事件权重矩阵。

[0048] 具体的,对安全事件根据安全事件所属范围进行分类得到安全事件类型,进一步建立维度分别为安全事件类型以及基础设施单位类型的二维矩阵,矩阵中的元素表征基于安全事件类型和基础设施单位类型的安全事件权重。

[0049] 示例性的,二维权重矩阵中的各矩阵元素可以根据大数据分析或者打分的形式给出。二维权重矩阵可以包括基于全国、各省、各市以及各行业的安全事件二维权重矩阵。

[0050] S230、根据二维权重矩阵确定各类安全事件类型对应基础设施单位类型的权重系数。

[0051] 其中,权重系数可以通过根据二维权重矩阵中的矩阵元素进行数学运算得到。各安全事件类型对应基础设施单位类型。

[0052] 具体的,首先确定所要计算的权重系数对应的安全事件类型以及基础设施单位类型,以根据安全事件类型以及基础设施单位类型查找二维权重矩阵对应的矩阵元素,进一步对该矩阵元素进行数学运算得到权重系数。

[0053] 示例性的,可以基于如下公式确定权重系数:

[0054] $$W_i=\frac{(a_i)^{\frac{1}{2}}}{\sum_{i=1}^{A}(a_i)^{\frac{1}{2}}}$$

[0055] 其中,W_i表示权重系数,a_i表示二维权重矩阵中的矩阵元素。A表示基础设施单位类型下所包括的安全事件类型的总数。

[0056] 示例性的,计算全国电信行业的僵尸木马类安全事件的权重系数,即安全事件类型为僵尸木马且基础设施单位类型为电信行业。首先找到全国安全事件二维权重矩阵,在该矩阵中查询出与电信行业僵尸木马安全事件类型对应的元素即a_i,该元素的算术平方根与电信行业各类安全事件对应元素的算术平方根的和,进行除法运算得到电信行业僵尸木马类安全事件的权重系数。

[0057] S240、根据权重系数确定基础设施安全指数模型。

[0058] 其中,安全指数模型可以是生成目标安全事件数据集合对应安全态势数据的数学模型。

[0059] 具体的,将权重系数以及预设参数输入至对应的数学模型生成基础设施安全指数模型。其中,预设参数与基础设施安全指数模型的类型有关。基础设施安全指数模型不同,则预设参数也不同。基础设施安全指数模型的类型可以根据目标安全事件数据集合所属区域范围来确定。

[0060] 示例性的,基础设施安全指数模型可以包括全国基础设施安全指数模型、各省基础设施安全指数模型、各市基础设施安全指数模型以及各行业基础设施安全指数模型。上述安全指数模型可以分别用于输出全国基础设施安全态势数据、各省基础设施安全态势数据、各市基础设施安全态势数据以及各行业基础设施安全态势数据。

[0061] 在本发明的一个可选实施例中,根据权重系数确定基础设施安全指数模型,可以包括:确定各类安全事件类型或基础设施单位类型对应的目标安全事件数量以及目标基础设施单位数量;根据权重系数、目标安全事件数量和目标基础设施单位数量确定基础设施安全指数模型。

[0062] 其中,目标安全事件可以是基础设施单位的设施发生过的安全事件。目标安全事件数量可以是与基础设施单位类型对应的目标安全事件的数量。目标基础设施单位数量可以是需要进行基础设施安全态势分析的单位的数量。

[0063] 具体的,在确定基础设施安全指数模型之前,根据目标安全事件数据集合确定各类安全事件类型对应的目标安全事件数量以及目标基础设施单位数量,或者根据基础设施单位类型确定与基础设施单位类型对应的目标安全事件数量以及目标基础设施单位数量,以进一步根据权重系数、目标安全事件数量和目标基础设施单位数量确定与安全事件类型

对应的基础设施安全指数模型,或者与基础设施单位类型对应的基础设施安全指数模型。

[0064] 在本发明的一个可选实施例中,根据权重系数、目标安全事件数量和目标基础设施单位数量确定基础设施安全指数模型,可以包括:基于如下公式确定基础设施安全指数模型:

[0065] $$Score=\sum_{i=1}^{A} W_i*\left(B-\ln\left(\frac{X_i}{N}+1\right)*C\right)$$

[0066] 其中,Score表示基础设施安全态势数据,W_i表示权重系数、X_i表示目标安全事件数量,N表示目标基础设施单位数量,A、B和C为常数。

[0067] 其中,A可以是安全事件数据集合中安全事件的类别数。B可以为使基础设施安全指数模型输出数据符合行业评估的一个常数。C可以为使基础设施安全指数模型输出的数据成符合一定数据分布的系数。数据分布形式可以包括但不限于正态分布、泊松分布以及卡方分布。$\left(\frac{X_i}{N}+1\right)$的运算是为了防止当$X_i$为0时,$\ln\left(\frac{X_i}{N}\right)$无意义。

[0068] 在本发明实施例中,对目标安全事件数量除以目标基础设施单位数量进行数据运算是为了保证公平性,上报的目标基础设施单位数量多则一般会导致目标基础设施单位对应的目标安全事件数量多,而目标安全事件数量多则将导致基础设施安全态势数据的数值变低。如果目标安全事件数量不与目标基础设施单位数量做除法,则对于上报目标基础设施单位的省、市或者行业来说是不公平的。为了保证一个更高的得分,上报数据的单位的省、市或者行业会降低上报基础设施单位的数量。因此为了保证公平性,在基础设施安全指数模型中,将目标安全事件数量与目标基础设施单位数量做除法运算的考虑是必要的。

[0069] 示例性的,当B为100,C为3.5,N为1,X_i为全国目标安全事件数量,A为全国目标安全事件的类别总数,例如,当A为27时,可以得到全国基础设施安全指数模型。全国基础设施安全指数模型如下面公式所示:

[0070] $$Score=\sum_{i=1}^{27} W_i*(100-\ln(X_i+1)*3.5)$$

[0071] 示例性的,当B为100,C为3.5,N为各省对应的目标基础设施单位数量,X_i为各省目标安全事件数量,A为各省目标安全事件的类别总数,例如,当A为27时,可以得到各省基础设施安全指数模型。各省基础设施安全指数模型如下面公式所示:

[0072] $$Score=\sum_{i=1}^{27} W_i*\left(100-\ln\left(\frac{X_i}{N}+1\right)*3.5\right)$$

[0073] 示例性的,当B为100,C为3.5,N为各市对应的目标基础设施单位数量,X_i为各市目标安全事件数量,A为各市目标安全事件的类别总数,例如,当A为27时,可以得到各市基础设施安全指数模型。各市基础设施安全指数模型如下面公式所示:

[0074] $$Score=\sum_{i=1}^{27} W_i*\left(100-\ln\left(\frac{X_i}{N}+1\right)*3.5\right)$$

[0075] 示例性的,当B为100,C为15,N为各基础设施单位类型对应的目标基础设施单位数量,X_i为各基础设施单位类型对应的目标安全事件数量,A为各各基础设施单位类型对应目标安全事件的类别总数,例如,当A为27以及$0\leq\sum_{i=1}^{27}X_i\leq108$时,可以得到各基础设施单位类型对应的基础设施安全指数模型。各基础设施单位类型对应的基础设施安全指数模型可以简称为行业基础设施安全指数模型,各行业基础设施安全指数模型如下面公式所示:

[0076] $$Score=\sum_{i=1}^{27}W_i*\left(100-\ln\left(\frac{X_i}{N}+1\right)*15\right)$$

[0077] 示例性的,当B为75,C为3,N为各基础设施单位类型对应的目标基础设施单位数量,X_i为各基础设施单位类型对应的目标安全事件数量,A为各基础设施单位类型对应目标安全事件的类别总数,例如,当A为27以及$\sum_{i=1}^{A}X_i>108$时,可以得到各行业基础设施安全指数模型,各行业基础设施安全指数模型如下面公式所示:

[0078] $$Score=\sum_{i=1}^{27}W_i*\left(75-\ln\left(\frac{X_i}{N}+1\right)*3\right)$$

[0079] 在本发明实施例中,各行业基础设施安全指数模型的约束条件是为了使基础设施安全指数模型输出数据符合一定的数学分布。

[0080] S250、将目标安全事件数据集合输入至基础设施安全指数模型中,以根据基础设施安全指数模型的输出结果确定基础设施安全态势数据。

[0081] 其中,输出结果可以是安全指数模型输出的用于分析基础设施对应的安全事件的数据。

[0082] 具体的,首先确定要进行基础设施安全态势分析的安全事件的类型、基础设施单位类型以及数据分析时间窗口,进一步获取数据分析时间窗口内的目标安全事件数据集合,根据安全事件类型或者基础设施单位类型确定对应的基础设施安全指数模型,从而将目标安全事件数据集合中的权重系数、目标安全事件数量和目标基础设施单位数量,输入至基础设施安全指数模型,基础设施安全指数模型的输出结果可以作为基础设施安全态势数据。

[0083] 示例性的,在数据分析时间窗口内,统计各省份上报的目标安全事件数据集合,根据目标安全事件数据集合计算各省份目标安全事件的数量,以及各省份目标基础设施单位数量,进一步将各省份目标安全事件的数量、各省份目标基础设施单位数量以及各类安全事件对应的权重系数输入至各省基础设施安全指数模型,输出的数据作为各省的基础设施安全态势分析数据。例如,可以针对某一行业的各类安全事件进行分析,也可以针对某一类安全事件对各个受攻击行业进行分析,进而完成该行业横向的安全状况对比以及纵向的安全状况对比。以此类推,本发明实施例也可以完成全国、各省以及各市的横向的安全状况对比以及纵向的安全状况对比。

[0084] 图3是本发明实施例二提供的一种多维数据分析示意图,在一个具体的例子中,如图3所示,获取全国、各省、各市以及各行业的基础设施单位数据,进一步根据上述基础设施单位数据解析出全国、各省、各市以及各行业的基础设施单位的设施数据,获取全国各安全

事件的IP和/或域名数据,以根据基础设施单位的设施数据中的IP和/或域名数据结合大数据分析,关联出基础设施对应的目标安全事件数据集合,最终将目标安全事件数据集合输入至基础设施安全指数模型,计算出全国、各省、各市以及各行业的基础设施安全态势数据。

[0085] 在本发明技术方案中,可以将目标安全事件数据集合输入至多种基础设施安全指数模型中,根据不同基础设施安全指数模型的输出结果可以形成基础设施安全态势的多维数据,即基于全国、各省、各市以及各行业的安全态势数据,通过多维数据分析可以提升基础设施安全事件数据的综合分析能力。

[0086] 需要说明的是,以上各实施例中各技术特征之间的任意排列组合也属于本发明的保护范围。

[0087] 实施例三

[0088] 图4是本发明实施例三提供的一种数据分析装置的示意图,如图4所示,所述装置包括:数据集合获取模块310以及安全态势数据分析模块320,其中:

[0089] 数据集合获取模块310,用于获取基础设施数据匹配的目标安全事件数据集合;所述基础设施数据为基础设施单位的设施数据;

[0090] 安全态势数据分析模块320,用于根据所述目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据。

[0091] 可选的,数据集合获取模块310,具体用于获取所述基础设施单位的基础设施网络数据;根据安全事件定位数据获取目标区域范围内的安全事件数据集合;根据所述基础设施网络数据和所述安全事件数据集合确定所述目标安全事件数据集合;其中,所述安全事件定位数据包括IP和/或域名数据;所述基础设施网络数据包括基础设施的IP和/或域名数据。

[0092] 可选的,安全态势数据分析模块320,具体用于确定基础设施安全指数模型;将所述目标安全事件数据集合输入至所述基础设施安全指数模型中,以根据所述基础设施安全指数模型的输出结果确定所述基础设施安全态势数据。

[0093] 可选的,安全态势数据分析模块320,具体用于根据安全事件类型和基础设施单位类型确定二维权重矩阵;根据所述二维权重矩阵确定各类安全事件类型对应所述基础设施单位类型的权重系数;根据所述权重系数确定所述基础设施安全指数模型。

[0094] 可选的,安全态势数据分析模块320,具体用于确定所述各类安全事件类型或所述基础设施单位类型对应的目标安全事件数量以及目标基础设施单位数量;根据所述权重系数、所述目标安全事件数量和所述目标基础设施单位数量确定所述基础设施安全指数模型。

[0095] 可选的,安全态势数据分析模块320,具体用于基于如下公式确定所述基础设施安全指数模型:

[0096] $$Score=\sum_{i=1}^{A} Wi*\left(B-\ln\left(\frac{Xi}{N}+1\right)*C\right)$$

[0097] 其中,Score表示所述基础设施安全态势数据,Wi表示所述权重系数、Xi表示所述目标安全事件数量,N表示所述目标基础设施单位数量,A、B和C为常数。

[0098] 可选的,数据集合获取模块310,具体用于确定数据分析时间窗口;根据所述数据分析时间窗口获取所述基础设施数据匹配的目标安全事件数据集合。

[0099] 本实施例的技术方案,在确定与获取的基础设施数据匹配的目标安全事件数据集合之后,进一步根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据,解决了现有技术对基础设施安全事件的数据挖掘深度不足以及数据综合分析能力较差的问题,达到了根据业务需求在预设区域范围内对基础设施的安全事件进行分析的效果,提高了基础设施安全事件数据的综合分析能力,丰富了数据分析维度。

[0100] 上述数据分析装置可执行本发明任意实施例所提供的数据分析方法,具备执行方法相应的功能模块和有益效果。未在本实施例中详尽描述的技术细节,可参见本发明任意实施例提供的数据分析方法。

[0101] 由于上述所介绍的数据分析装置为可以执行本发明实施例中的数据分析方法的装置,故而基于本发明实施例中所介绍的数据分析方法,本领域所属技术人员能够了解本实施例的数据分析装置的具体实施方式以及其各种变化形式,所以在此对于该数据分析装置如何实现本发明实施例中的数据分析方法不再详细介绍。只要本领域所属技术人员实施本发明实施例中数据分析方法所采用的装置,都属于本申请所欲保护的范围。

[0102] 实施例四

[0103] 图5为本发明实施例四提供的一种电子设备的结构示意图。图5示出了适于用来实现本发明实施方式的电子设备412的框图。图5显示的电子设备412仅仅是一个示例,不应对本发明实施例的功能和使用范围带来任何限制。

[0104] 如图5所示,电子设备412以通用计算设备的形式表现。电子设备412的组件可以包括但不限于:一个或者多个处理器416,存储装置428,连接不同系统组件(包括存储装置428和处理器416)的总线418。

[0105] 总线418表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线结构的局域总线。举例来说,这些体系结构包括但不限于工业标准体系结构(Industry Standard Architecture,ISA)总线,微通道体系结构(Micro Channel Architecture,MCA)总线,增强型ISA总线、视频电子标准协会(Video Electronics Standards Association,VESA)局域总线以及外围组件互连(Peripheral Component Interconnect,PCI)总线。

[0106] 电子设备412典型地包括多种计算机系统可读介质。这些介质可以是任何能够被电子设备412访问的可用介质,包括易失性和非易失性介质,可移动的和不可移动的介质。

[0107] 存储装置428可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器(Random Access Memory,RAM)430和/或高速缓存存储器432。电子设备412可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例,存储系统434可以用于读写不可移动的、非易失性磁介质(图5未显示,通常称为“硬盘驱动器”)。尽管图5中未示出,可以提供用于对可移动非易失性磁盘(例如“软盘”)读写的磁盘驱动器,以及对可移动非易失性光盘(例如只读光盘(Compact Disc-Read Only Memory,CD-ROM)、数字视盘(Digital Video Disc-Read Only Memory,DVD-ROM)或者其它光介质)读写的光盘驱动器。在这些情况下,每个驱动器可以通过一个或者多个数据介质接口与总线418相连。存储装置428可以包括至少一个程序产品,该程序产品具有一组(例如至少一个)程序模

块,这些程序模块被配置以执行本发明各实施例的功能。

[0108] 具有一组(至少一个)程序模块426的程序436,可以存储在例如存储装置428中,这样的程序模块426包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块426通常执行本发明所描述的实施例中的功能和/或方法。

[0109] 电子设备412也可以与一个或多个外部设备414(例如键盘、指向设备、摄像头、显示器424等)通信,还可与一个或者多个使得用户能与该电子设备412交互的设备通信,和/或与使得该电子设备412能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入/输出(Input/Output,I/O)接口422进行。并且,电子设备412还可以通过网络适配器420与一个或者多个网络(例如局域网(Local Area Network,LAN),广域网Wide Area Network,WAN)和/或公共网络,例如因特网)通信。如图所示,网络适配器420通过总线418与电子设备412的其它模块通信。应当明白,尽管图中未示出,可以结合电子设备412使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、磁盘阵列(Redundant Arrays of Independent Disks,RAID)系统、磁带驱动器以及数据备份存储系统等。

[0110] 处理器416通过运行存储在存储装置428中的程序,从而执行各种功能应用以及数据处理,例如实现本发明上述实施例所提供的数据分析方法:获取基础设施数据匹配的目标安全事件数据集合;所述基础设施数据为基础设施单位的设施数据;根据所述目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据。

[0111] 本实施例的技术方案,在确定与获取的基础设施数据匹配的目标安全事件数据集合之后,进一步根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据,解决了现有技术对基础设施安全事件的数据挖掘深度不足以及数据综合分析能力较差的问题,达到了根据业务需求在预设区域范围内对基础设施的安全事件进行分析的效果,提高了基础设施安全事件数据的综合分析能力,丰富了数据分析维度。

[0112] 实施例五

[0113] 本发明实施例五还提供一种存储计算机程序的计算机存储介质,所述计算机程序在由计算机处理器执行时用于执行本发明上述实施例任一所述的数据分析方法:获取基础设施数据匹配的目标安全事件数据集合;所述基础设施数据为基础设施单位的设施数据;根据所述目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据。

[0114] 本发明实施例的计算机存储介质,可以采用一个或多个计算机可读的介质的任意组合。计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。计算机可读存储介质例如可以是但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(Read Only Memory,ROM)、可擦式可编程只读存储器((Erasable Programmable Read Only Memory,EPROM)或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本文件中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0115] 计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,

其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0116] 计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于无线、电线、光缆、射频(Radio Frequency,RF)等等,或者上述的任意合适的组合。

[0117] 可以以一种或多种程序设计语言或其组合来编写用于执行本发明操作的计算机程序代码,所述程序设计语言包括面向对象的程序设计语言—诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言,诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络,包括局域网(LAN)或广域网(WAN)连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0118] 注意,上述仅为本发明的较佳实施例及所运用技术原理。本领域技术人员会理解,本发明不限于这里所述的特定实施例,对本领域技术人员来说能够进行各种明显的变化、重新调整和替代而不会脱离本发明的保护范围。因此,虽然通过以上实施例对本发明进行了较为详细的说明,但是本发明不仅仅限于以上实施例,在不脱离本发明构思的情况下,还可以包括更多其他等效实施例,而本发明的范围由所附的权利要求范围决定。

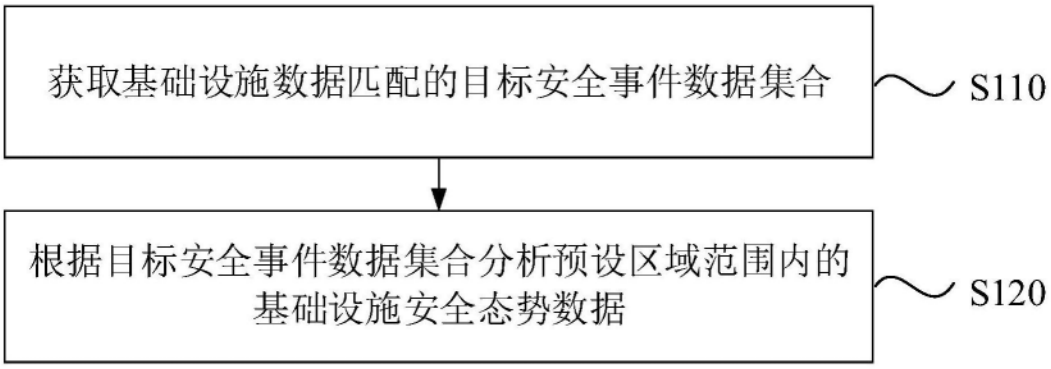

获取基础设施数据匹配的目标安全事件数据集合
S110
根据目标安全事件数据集合分析预设区域范围内的基础设施安全态势数据
S120

图1

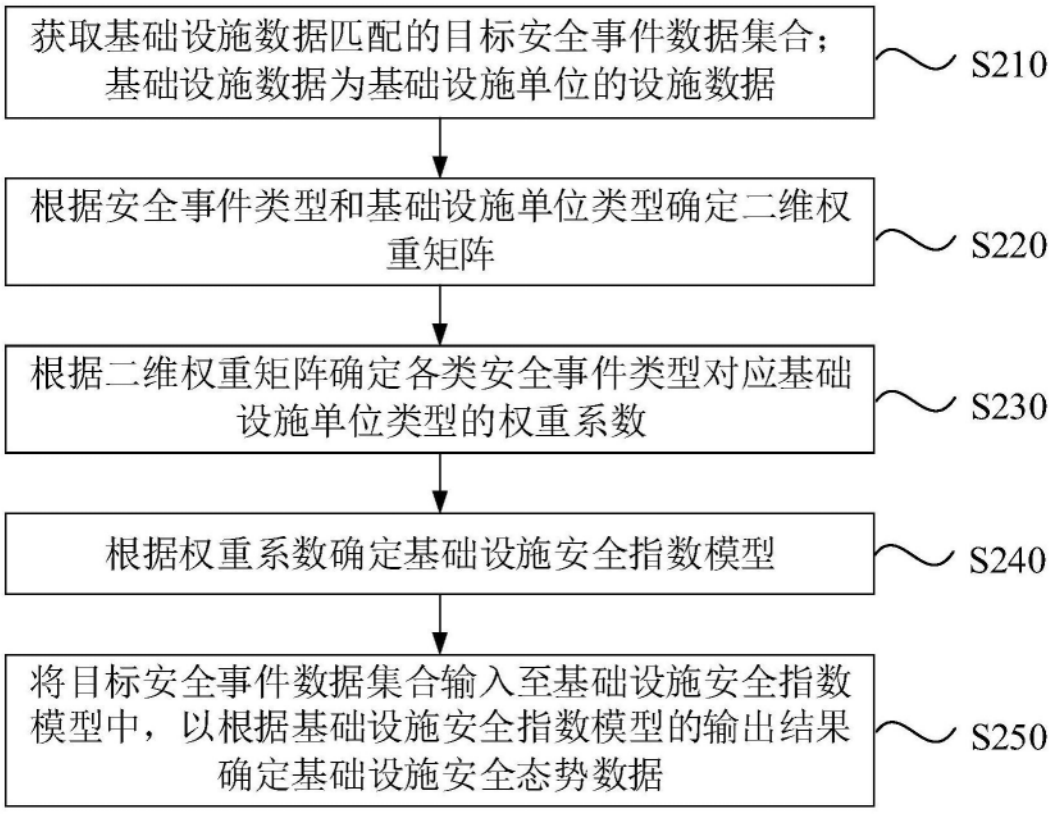

获取基础设施数据匹配的目标安全事件数据集合；基础设施数据为基础设施单位的设施数据
S210
根据安全事件类型和基础设施单位类型确定二维权重矩阵
S220
根据二维权重矩阵确定各类安全事件类型对应基础设施单位类型的权重系数
S230
根据权重系数确定基础设施安全指数模型
S240
将目标安全事件数据集合输入至基础设施安全指数模型中，以根据基础设施安全指数模型的输出结果确定基础设施安全态势数据
S250

图2

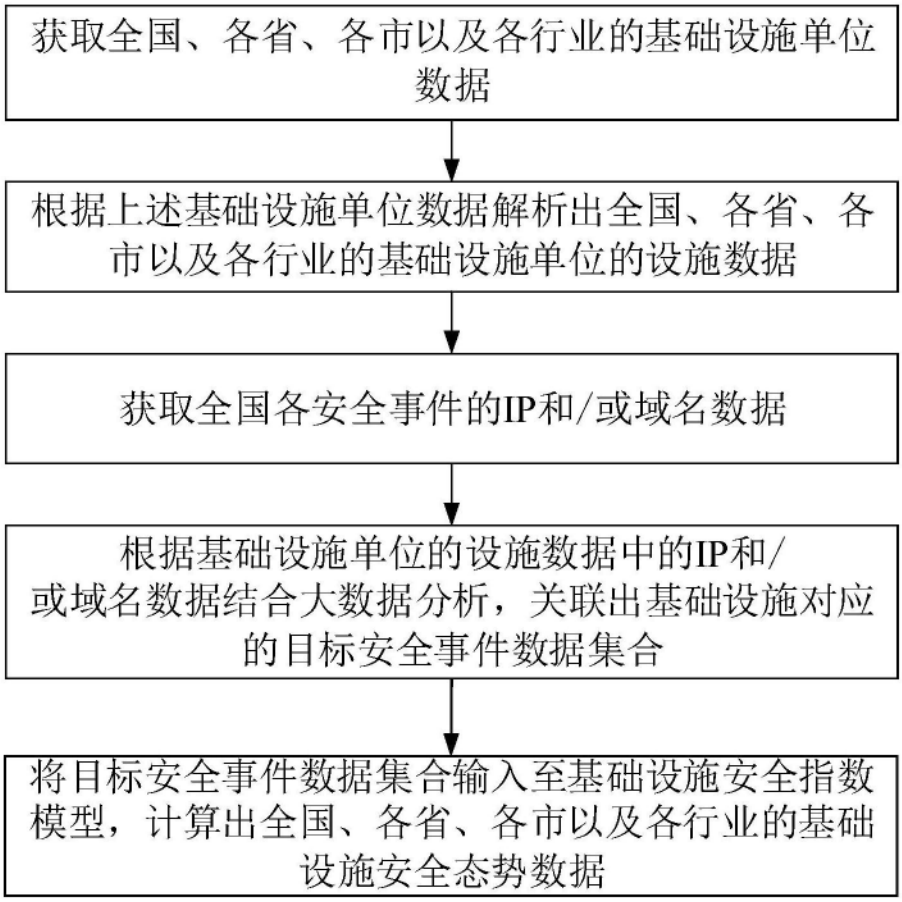

获取全国、各省、各市以及各行业的基础设施单位数据
根据上述基础设施单位数据解析出全国、各省、各市以及各行业的基础设施单位的设施数据
获取全国各安全事件的IP和/或域名数据
根据基础设施单位的设施数据中的IP和/或域名数据结合大数据分析，关联出基础设施对应的目标安全事件数据集合
将目标安全事件数据集合输入至基础设施安全指数模型，计算出全国、各省、各市以及各行业的基础设施安全态势数据

图3

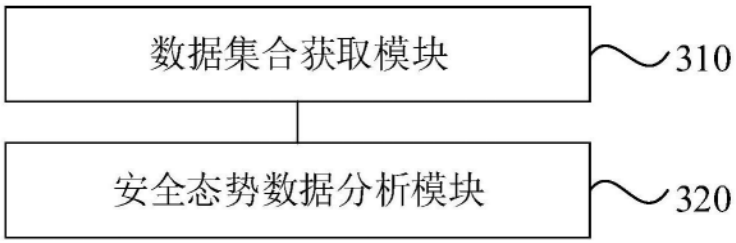

数据集合获取模块
310
安全态势数据分析模块
320

图4

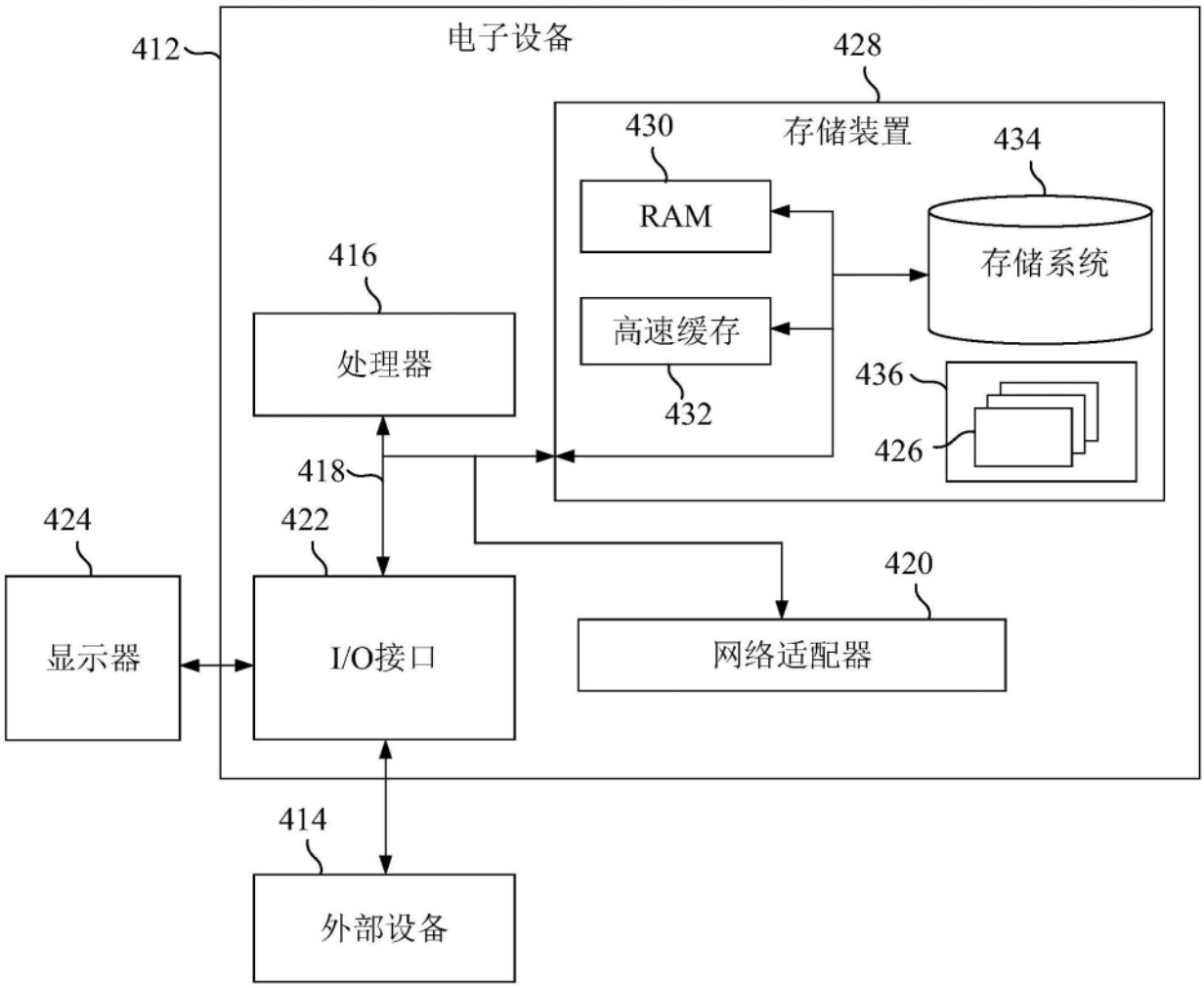

412
电子设备
428
430
存储装置
434
RAM
存储系统
416
处理器
高速缓存
436
432
426
418
424
422
420
显示器
I/O接口
网络适配器
414
外部设备

图5