(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 109428857 B
(45)授权公告日 2021.01.05

(21)申请号 201710729894.1
(22)申请日 2017.08.23
(65)同一申请的已公布的文献号
申请公布号 CN 109428857 A
(43)申请公布日 2019.03.05
(73)专利权人 腾讯科技(深圳)有限公司
地址 518057 广东省深圳市南山区高新区科技中一路腾讯大厦35层
(72)发明人 马立伟 王月强 李志豪 张刚 王朝飞
(74)专利代理机构 深圳市深佳知识产权代理事务所(普通合伙) 44285
代理人 王仲凯
(51)Int.Cl.
H04L 29/06(2006.01)

(56)对比文件
CN 106131071 A,2016.11.16
CN 106330861 A,2017.01.11
CN 104601556 A,2015.05.06
US 8959643 B1,2015.02.17
CN 101841523 A,2010.09.22

审查员 白生斌

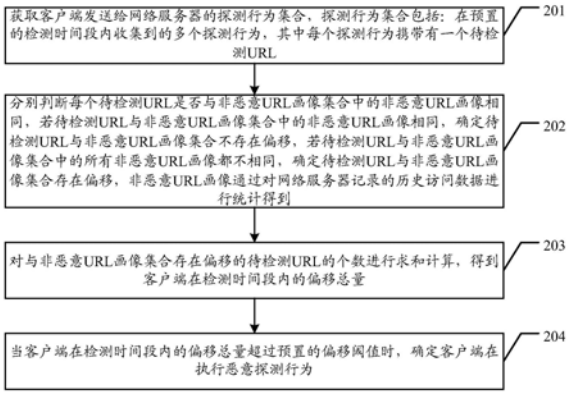
权利要求书5页 说明书20页 附图7页

(54)发明名称

一种恶意探测行为的检测方法和装置

(57)摘要

本发明实施例公开一种恶意探测行为的检测方法和装置，可提高恶意探测行为的检测效果。在该方法中，获取客户端发送给网络服务器的探测行为集合；分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若待检测URL与非恶意URL画像集合中的非恶意URL画像相同，确定待检测URL与非恶意URL画像集合不存在偏移，若待检测URL与非恶意URL画像集合中的所有非恶意URL画像都不相同，确定待检测URL与非恶意URL画像集合存在偏移；对与非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量；当偏移总量超过偏移阈值时，确定客户端在执行恶意探测行为。



CN 109428857 B

1.一种恶意探测行为的检测方法,其特征在于,包括:

获取客户端发送给网络服务器的探测行为集合,所述探测行为集合包括:在预置的检测时间段内收集到的至少一个探测行为,其中每个探测行为携带有一个待检测统一资源定位符URL;

分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同,若所述待检测URL与所述非恶意URL画像集合中的非恶意URL画像相同,确定所述待检测URL与所述非恶意URL画像集合不存在偏移,若所述待检测URL与所述非恶意URL画像集合中的所有非恶意URL画像都不相同,确定所述待检测URL与所述非恶意URL画像集合存在偏移,所述非恶意URL画像通过对所述网络服务器记录的历史访问数据进行统计得到;

对与所述非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算,得到所述客户端在所述检测时间段内的偏移总量;

当所述客户端在所述检测时间段内的偏移总量超过预置的偏移阈值时,确定所述客户端在执行恶意探测行为。

2.根据权利要求1所述的方法,其特征在于,所述非恶意URL画像集合,包括:常用访问URL画像子集合、临时访问URL画像子集合、特殊权限URL画像子集合和空URL画像子集合中的至少一种;

所述分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同,包括:

分别判断每个待检测URL是否与所述常用访问URL画像子集合中的常用访问URL画像相同,若所述待检测URL与所述常用访问URL画像子集合中的常用访问URL画像相同,确定所述待检测URL与所述常用访问URL画像子集合不存在偏移,若所述待检测URL与所述常用访问URL画像子集合中的所有常用访问URL画像都不相同,确定所述待检测URL与所述常用访问URL画像子集合存在偏移;

分别判断每个待检测URL是否与所述临时访问URL画像子集合中的临时访问URL画像相同,若所述待检测URL与所述临时访问URL画像子集合中的临时访问URL画像相同,确定所述待检测URL与所述临时访问URL画像子集合不存在偏移,若所述待检测URL与所述临时访问URL画像子集合中的所有临时访问URL画像都不相同,确定所述待检测URL与所述临时访问URL画像子集合存在偏移;

分别判断每个待检测URL是否与所述特殊权限URL画像子集合中的特殊权限URL画像相同,若所述待检测URL与所述特殊权限URL画像子集合中的特殊权限URL画像相同,确定所述待检测URL与所述特殊权限URL画像子集合不存在偏移,若所述待检测URL与所述特殊权限URL画像子集合中的所有特殊权限URL画像都不相同,确定所述待检测URL与所述特殊权限URL画像子集合存在偏移;

分别判断每个待检测URL是否与所述空URL画像子集合中的空URL画像相同,若所述待检测URL与所述空URL画像子集合中的空URL画像相同,确定所述待检测URL与所述空URL画像子集合不存在偏移,若所述待检测URL与所述空URL画像子集合中的所有空URL画像都不相同,确定所述待检测URL与所述空URL画像子集合存在偏移。

3.根据权利要求2所述的方法,其特征在于,所述对与所述非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算,得到所述客户端在所述检测时间段内的偏移总量,包

括：

对与所述常用访问URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第一偏移量；

对与所述临时访问URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第二偏移量；

对与所述特殊权限URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第三偏移量；

对与所述空URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第四偏移量；

将所述第一偏移量、所述第二偏移量、所述第三偏移量和所述第四偏移量进行相加计算，得到所述客户端在所述检测时间段内的偏移总量。

4.根据权利要求2所述的方法，其特征在于，所述分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同之前，所述方法还包括：

通过日志采集系统收集所述网络服务器的访问日志，所述访问日志包括：通过客户端发送的目的URL；

将所述目的URL划分到如下四种非恶意URL画像中的至少一种：常用访问URL画像、临时访问URL画像、特殊权限URL画像和空URL画像。

5.根据权利要求4所述的方法，其特征在于，所述将所述目的URL划分到如下四种非恶意URL画像中的至少一种：常用访问URL画像、临时访问URL画像、特殊权限URL画像和空URL画像，包括：

将满足第一条件的目的URL划分到常用访问URL画像，所述第一条件包括：在预置的第一单位时间内访问成功的次数大于第一次数阈值；

将满足第二条件的目的URL划分到临时访问URL画像，所述第二条件包括：在预置的第一单位时间内访问成功的次数大于0、且小于或等于所述第一次数阈值；

将满足第三条件的目的URL划分到空URL画像，所述第三条件包括：在预置的第二单位时间内访问失败的次数小于第二次数阈值；

将满足第四条件的目的URL划分到特殊权限URL画像，所述第四条件包括：在预置的第三单位时间内访问成功的用户所属的集合包括的用户个数小于用户个数阈值。

6.根据权利要求1至5中任一项所述的方法，其特征在于，所述对与所述非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到所述客户端在所述检测时间段内的偏移总量，包括：

当存在至少两种类型的所述非恶意URL画像时，分别统计与所述至少两种类型的所述非恶意URL画像存在偏移的待检测URL的总个数，得到对应于不同类型的非恶意URL画像的待检测URL的总个数；

获取为每种类型的所述非恶意URL画像分别配置的权重参数；

根据所述权重参数对所述对应于不同类型的非恶意URL画像的待检测URL的总个数进行加权计算，得到所述客户端在所述检测时间段内的偏移总量。

7.根据权利要求6所述的方法，其特征在于，所述确定所述客户端在执行恶意探测行为之后，所述方法还包括：

获取对所述客户端的恶意探测行为处理结果，并根据所述恶意探测行为处理结果对所述每种类型的非恶意URL画像配置的权重参数进行优化调整。

8.根据权利要求1至5中任一项所述的方法，其特征在于，所述确定所述客户端在执行恶意探测行为之后，所述方法还包括：

获取对所述客户端的恶意探测行为处理结果，并根据所述恶意探测行为处理结果对所述偏移阈值进行优化调整。

9.一种恶意探测行为的检测装置，其特征在于，包括：

探测行为获取模块，用于获取客户端发送给网络服务器的探测行为集合，所述探测行为集合包括：在预置的检测时间段内收集到的至少一个探测行为，其中每个探测行为携带有一个待检测统一资源定位符URL；

URL判断模块，用于分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若所述待检测URL与所述非恶意URL画像集合中的非恶意URL画像相同，确定所述待检测URL与所述非恶意URL画像集合不存在偏移，若所述待检测URL与所述非恶意URL画像集合中的所有非恶意URL画像都不相同，确定所述待检测URL与所述非恶意URL画像集合存在偏移，所述非恶意URL画像通过对所述网络服务器记录的历史访问数据进行统计得到；

偏移总量计算模块，用于对与所述非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到所述客户端在所述检测时间段内的偏移总量；

检测模块，用于当所述客户端在所述检测时间段内的偏移总量超过预置的偏移阈值时，确定所述客户端在执行恶意探测行为。

10.根据权利要求9所述的装置，其特征在于，所述非恶意URL画像集合，包括：常用访问URL画像子集合、临时访问URL画像子集合、特殊权限URL画像子集合和空URL画像子集合中的至少一种；

所述URL判断模块，包括：

第一判断子模块，用于分别判断每个待检测URL是否与所述常用访问URL画像子集合中的常用访问URL画像相同，若所述待检测URL与所述常用访问URL画像子集合中的常用访问URL画像相同，确定所述待检测URL与所述常用访问URL画像子集合不存在偏移，若所述待检测URL与所述常用访问URL画像子集合中的所有常用访问URL画像都不相同，确定所述待检测URL与所述常用访问URL画像子集合存在偏移；

第二判断子模块，用于分别判断每个待检测URL是否与所述临时访问URL画像子集合中的临时访问URL画像相同，若所述待检测URL与所述临时访问URL画像子集合中的临时访问URL画像相同，确定所述待检测URL与所述临时访问URL画像子集合不存在偏移，若所述待检测URL与所述临时访问URL画像子集合中的所有临时访问URL画像都不相同，确定所述待检测URL与所述临时访问URL画像子集合存在偏移；

第三判断子模块，用于分别判断每个待检测URL是否与所述特殊权限URL画像子集合中的特殊权限URL画像相同，若所述待检测URL与所述特殊权限URL画像子集合中的特殊权限URL画像相同，确定所述待检测URL与所述特殊权限URL画像子集合不存在偏移，若所述待检测URL与所述特殊权限URL画像子集合中的所有特殊权限URL画像都不相同，确定所述待检测URL与所述特殊权限URL画像子集合存在偏移；

第四判断子模块，用于分别判断每个待检测URL是否与所述空URL画像子集合中的空

URL画像相同，若所述待检测URL与所述空URL画像子集合中的空URL画像相同，确定所述待检测URL与所述空URL画像子集合不存在偏移，若所述待检测URL与所述空URL画像子集合中的所有空URL画像都不相同，确定所述待检测URL与所述空URL画像子集合存在偏移。

11.根据权利要求10所述的装置，其特征在于，所述偏移总量计算模块，包括：

第一计算子模块，用于对与所述常用访问URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第一偏移量；

第二计算子模块，用于对与所述临时访问URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第二偏移量；

第三计算子模块，用于对与所述特殊权限URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第三偏移量；

第四计算子模块，用于对与所述空URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第四偏移量；

第五计算子模块，用于将所述第一偏移量、所述第二偏移量、所述第三偏移量和所述第四偏移量进行相加计算，得到所述客户端在所述检测时间段内的偏移总量。

12.根据权利要求10所述的装置，其特征在于，所述恶意探测行为的检测装置还包括：日志收集模块和行为对比模块，其中，

所述日志收集模块，用于所述URL判断模块分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同之前，通过日志采集系统收集所述网络服务器的访问日志，所述访问日志包括：通过客户端发送的目的URL；

所述行为对比模块，用于将所述目的URL划分到如下四种非恶意URL画像中的至少一种：常用访问URL画像、临时访问URL画像、特殊权限URL画像和空URL画像。

13.根据权利要求12所述的装置，其特征在于，所述行为对比模块，包括：

第一划分子模块，用于将满足第一条件的目的URL划分到常用访问URL画像，所述第一条件包括：在预置的第一单位时间内访问成功的次数大于第一次数阈值；

第二划分子模块，用于将满足第二条件的目的URL划分到临时访问URL画像，所述第二条件包括：在预置的第一单位时间内访问成功的次数大于0、且小于或等于所述第一次数阈值；

第三划分子模块，用于将满足第三条件的目的URL划分到空URL画像，所述第三条件包括：在预置的第二单位时间内访问失败的次数小于第二次数阈值；

第四划分子模块，用于将满足第四条件的目的URL划分到特殊权限URL画像，所述第四条件包括：在预置的第三单位时间内访问成功的用户所属的集合包括的用户个数小于用户个数阈值。

14.根据权利要求9至13中任一项所述的装置，其特征在于，所述URL判断模块，包括：

分类统计子模块，用于当存在至少两种类型的所述非恶意URL画像时，分别统计与所述至少两种类型的所述非恶意URL画像存在偏移的待检测URL的总个数，得到对应于不同类型的非恶意URL画像的待检测URL的总个数；

权重获取子模块，用于获取为每种类型的所述非恶意URL画像分别配置的权重参数；

加权计算子模块，用于根据所述权重参数对所述对应于不同类型的非恶意URL画像的待检测URL的总个数进行加权计算，得到所述客户端在所述检测时间段内的偏移总量。

15.一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行如权利要求1-8任意一项所述的方法。

一种恶意探测行为的检测方法和装置

技术领域

[0001] 本发明涉及计算机技术领域，尤其涉及一种恶意探测行为的检测方法和装置。

背景技术

[0002] 随着互联网技术的快速发展，用户在互联网上进行的操作会越来越频繁，与此同时，一些恶意分子开发自动访问程序在互联网上进行恶意行为，例如破解用户标识的密码，利用破解的用户标识发送垃圾消息和盗取财产等，因此需要对恶意行为进行有效检测。

[0003] 目前，网络(web)扫描探测是黑客进入内网后进一步入侵渗透的常用手段，例如，通过爬虫获取统一资源定位符(Uniform Resource Locator，URL)资源树，或者探测网络服务器的管理后台都是黑客踩点的常用手段，当前业内检测恶意探测行为的方法具体可以为：某一用户通过URL访问网络中对应的网页内容时，统计目的URL在单位时间内被访问的总次数，当该总次数超过某个阈值时确定为恶意探测行为。其中恶意探测行为常用于各类欺诈、仿冒、钓鱼或挂马网页，当用户不慎访问此类网页时，就可能对用户造成如经济上的损失、个人隐私信息的泄露或是使当前电脑感染木马病毒等不利影响。

[0004] 现有技术中，对恶意探测行为进行检测主要依赖于对目的URL的访问总次数的统计以及与阈值的判断，但是这种方案无法发现访问总数小于该阈值的恶意探测行为，且因用户正常的访问请求也会恶意探测行为夹杂在一起，从而会造成大量误报。假如阈值设置的过小，则会产生大量误报，若该阈值设置的过大，则会无法有效检测到恶意探测行为，因此现有技术提供的恶意探测行为的检测方案存在检测效果差的问题。

发明内容

[0005] 本发明实施例提供了一种恶意探测行为的检测方法和装置，用于有效解决恶意探测行为的发现问题，提高恶意探测行为的检测效果。

[0006] 为解决上述技术问题，本发明实施例提供以下技术方案：

[0007] 第一方面，本发明实施例提供一种恶意探测行为的检测方法，包括：

[0008] 获取客户端发送给网络服务器的探测行为集合，所述探测行为集合包括：在预置的检测时间段内收集到的至少一个探测行为，其中每个探测行为携带有一个待检测统一资源定位符URL；

[0009] 分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若所述待检测URL与所述非恶意URL画像集合中的非恶意URL画像相同，确定所述待检测URL与所述非恶意URL画像集合不存在偏移，若所述待检测URL与所述非恶意URL画像集合中的所有非恶意URL画像都不相同，确定所述待检测URL与所述非恶意URL画像集合存在偏移，所述非恶意URL画像通过对所述网络服务器记录的历史访问数据进行统计得到；

[0010] 对与所述非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到所述客户端在所述检测时间段内的偏移总量，

[0011] 当所述客户端在所述检测时间段内的偏移总量超过预置的偏移阈值时，确定所述

客户端在执行恶意探测行为。

[0012] 第二方面，本发明实施例还提供一种恶意探测行为的检测装置，包括：

[0013] 探测行为获取模块，用于获取客户端发送给网络服务器的探测行为集合，所述探测行为集合包括：在预置的检测时间段内收集到的至少一个探测行为，其中每个探测行为携带有一个待检测统一资源定位符URL；

[0014] URL判断模块，用于分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若所述待检测URL与所述非恶意URL画像集合中的非恶意URL画像相同，确定所述待检测URL与所述非恶意URL画像集合不存在偏移，若所述待检测URL与所述非恶意URL画像集合中的所有非恶意URL画像都不相同，确定所述待检测URL与所述非恶意URL画像集合存在偏移，所述非恶意URL画像通过对所述网络服务器记录的历史访问数据进行统计得到；

[0015] 偏移总量计算模块，用于对与所述非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到所述客户端在所述检测时间段内的偏移总量，

[0016] 检测模块，用于当所述客户端在所述检测时间段内的偏移总量超过预置的偏移阈值时，确定所述客户端在执行恶意探测行为。

[0017] 本申请的第三方面，提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有指令，当其在计算机上运行时，使得计算机执行上述各方面所述的方法。

[0018] 从以上技术方案可以看出，本发明实施例具有以下优点：

[0019] 在本发明实施例中，首先获取客户端发送给网络服务器的探测行为集合，探测行为集合包括：在预置的检测时间段内收集到的至少一个探测行为，其中每个探测行为携带有一个待检测URL，然后分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若待检测URL与非恶意URL画像集合中的非恶意URL画像相同，确定待检测URL与非恶意URL画像集合不存在偏移，若待检测URL与非恶意URL画像集合中的所有非恶意URL画像都不相同，确定待检测URL与非恶意URL画像集合存在偏移，对与非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量，当客户端在检测时间段内的偏移总量超过预置的偏移阈值时，确定客户端在执行恶意探测行为。由于非恶意URL画像可以通过对网络服务器记录的历史访问数据进行统计得到，使用该非恶意URL画像作为参考量，从而通过衡量待检测URL与非恶意URL画像集合是否产生偏移可以计算出客户端在检测时间段内的偏移总量，最后通过对偏移总量的阈值判断可以确定出客户端是否在执行恶意探测行为。对于发送URL访问总次数较少的用户，也可以通过判断偏移总量确定该客户端是否在执行恶意探测行为，因此可以有效解决恶意探测行为的发现问题，提高恶意探测行为的检测效果。

附图说明

[0020] 为了更清楚地说明本发明实施例中的技术方案，下面将对实施例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，对于本领域的技术人员来讲，还可以根据这些附图获得其他的附图。

[0021] 图1为本发明实施例提供的恶意探测行为的检测方法所应用系统的组成结构示意图；

[0022] 图2为本发明实施例提供的一种恶意探测行为的检测方法的流程方框示意图；

[0023] 图3为本发明实施例提供的恶意探测行为的检测方法所应用系统的一种实现场景下的架构示意图；

[0024] 图4-a为本发明实施例提供的一种恶意探测行为的检测装置的组成结构示意图；

[0025] 图4-b为本发明实施例提供的一种URL判断模块的组成结构示意图；

[0026] 图4-c为本发明实施例提供的一种偏移总量计算模块的组成结构示意图；

[0027] 图4-d为本发明实施例提供的另一种恶意探测行为的检测装置的组成结构示意图；

[0028] 图4-e为本发明实施例提供的一种行为对比模块的组成结构示意图；

[0029] 图4-f为本发明实施例提供的另一种URL判断模块的组成结构示意图；

[0030] 图4-g为本发明实施例提供的另一种恶意探测行为的检测装置的组成结构示意图；

[0031] 图4-h为本发明实施例提供的另一种恶意探测行为的检测装置的组成结构示意图；

[0032] 图5为本发明实施例提供的恶意探测行为的检测方法应用于服务器的组成结构示意图。

具体实施方式

[0033] 本发明实施例提供了一种恶意探测行为的检测方法和装置，用于有效解决恶意探测行为的发现问题，提高恶意探测行为的检测效果。

[0034] 为使得本发明的发明目的、特征、优点能够更加的明显和易懂，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，下面所描述的实施例仅仅是本发明一部分实施例，而非全部实施例。基于本发明中的实施例，本领域的技术人员所获得的所有其他实施例，都属于本发明保护的范围。

[0035] 本发明的说明书和权利要求书及上述附图中的术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，以便包含一系列单元的过程、方法、系统、产品或设备不必限于那些单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它单元。

[0036] 以下分别进行详细说明。

[0037] 如图1所示，本发明实施例提供的恶意探测行为的检测方法可应用于图1所示的系统架构，该系统内可以包括：恶意探测行为的检测装置、网络(web)服务器和客户端，其中，客户端向网络服务器发送至少一个探测行为，每个探测行为可以携带一个待检测URL，在恶意探测行为的检测装置所执行的检测时间段内，网络服务器可以收集到至少一个探测行为，定义为探测行为集合。恶意探测行为的检测装置可以通过与网络服务器的交互获取到探测行为集合，该恶意探测行为的检测装置可以是独立于网络服务器的一个单独设备，也可以作为网络服务器内部集成实现的一个检测模块，图1中以恶意探测行为的检测装置为一个独立设备进行举例说明。恶意探测行为的检测装置用于通过对该探测行为集合中待检测URL的分析确定客户端的操作用户是否正在执行恶意探测行为。web扫描探测是黑客进入网络服务器的内网后进一步入侵渗透的常用手段，本发明实施例中对于URL访问总次数较

少的用户也可以进行有效检测，并且提高对恶意探测行为的检测效果。该恶意探测行为的检测装置可以如下实施例中所描述的恶意探测行为的检测方法，详见后续实施例中的举例说明。

[0038] 请参阅图1所示，本发明一个实施例提供的恶意探测行为的检测方法，可以通过对恶意探测行为的检测，从而可以保护网络服务器的安全。可以包括如下步骤：

[0039] 201、获取客户端发送给网络服务器的探测行为集合，探测行为集合包括：在预置的检测时间段内收集到的至少一个探测行为，其中每个探测行为携带有一个待检测URL。

[0040] 在本发明实施例中，客户端可以同时或者异步的方式向网络服务器发送至少一个探测行为，每个探测行为可以携带一个待检测URL，在恶意探测行为的检测装置所执行的检测时间段内，网络服务器可以收集到至少一个探测行为，定义为探测行为集合，从该网络服务器可以先获取到探测行为集合，该探测行为集合是本发明实施例中分析客户端的操作用户是否执行恶意探测行为的基础数据。在实际应用中，每个探测行为除了携带待检测URL之外，还可以携带如下信息中的至少一种：访问时间、源互联网协议(Internet Protocol，IP)地址、源设备名、源用户名、目的IP地址、目的域名。

[0041] 需要说明的是，在本发明实施例中，在检测时间段内可以从网络服务器提取到至少一条的探测行为，从而通过对多条的探测行为的URL的解析，从而确定是否产生恶意探测行为。其中需要收集的探测行为的条数可以根据实现场景来确定，例如可以收集到多条的探测行为，从而通过对多条的探测行为的分析确定是否产生恶意探测行为，通常情况下，同一个客户端发送的探测行为越多，越能够准确的探测到其是否正在执行恶意探测行为。

[0042] 202、分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若待检测URL与非恶意URL画像集合中的非恶意URL画像相同，确定待检测URL与非恶意URL画像集合不存在偏移，若待检测URL与非恶意URL画像集合中的所有非恶意URL画像都不相同，确定待检测URL与非恶意URL画像集合存在偏移，非恶意URL画像通过对网络服务器记录的历史访问数据进行统计得到。

[0043] 在本发明实施例中，通过前述步骤201获取到至少一个探测请求之后，可以计算判断每个待检测URL与非恶意URL画像集合之间是否产生偏移，例如是否产生偏移的判断方式可以包括：若待检测URL与非恶意URL画像集合中的非恶意URL画像相同，例如某一个待检测URL与非恶意URL画像集合中的某一个非恶意URL画像相同时说明该待检测URL与非恶意URL画像集合匹配成功，即待检测URL与非恶意URL画像集合不存在偏移。例如某一个待检测URL与非恶意URL画像集合中的所有非恶意URL画像都不相同时说明该待检测URL与非恶意URL画像集合匹配失败，即待检测URL与非恶意URL画像集合存在偏移。其中，为了准确判断出每个待检测URL是否存在恶意的可能性，将待检测URL与非恶意URL画像集合进行画像比对分析，从而可以确定出各个待检测URL是否与非恶意URL画像集合产生了偏移。该非恶意URL画像集合可以通过对网络服务器记录的历史访问数据进行统计得到，由于网络服务器记录的历史访问数据是网络服务器在一段时间内的真实用户在网络上产生的访问数据，通过对历史访问数据的分析可确定出非恶意URL集合，这些非恶意URL集合中的所有非恶意URL可以作为参考量来判断待检测URL是否产生偏移，本发明实施例中将通过历史访问数据统计出的非恶意URL的样本定义为非恶意URL画像，非恶意URL画像是指非恶意URL的描述信息，例如非恶意URL画像可以包括：非恶意URL的地址、对应的域名、访问用户的集合等，通过对网

络服务器的历史访问数据进行分析确定出的所有非恶意URL画像构成非恶意URL画像集合。在本发明实施例中，非恶意URL画像集合可以在步骤202执行之前就预先生成即可，从而可以将该非恶意URL画像集合作为参考量，通过衡量待检测URL与非恶意URL画像集合中的非恶意URL画像是否相同来确定同一个客户端的多个待检测URL与非恶意URL画像集合之间是否产生偏移。

[0044] 在本发明的一些实施例中，非恶意URL画像集合可以有多种实现情况，即本发明实施例中非恶意URL画像集合可以包括多种的非恶意URL图像。举例说明如下，非恶意URL画像集合，可以包括：常用访问URL画像子集合、临时访问URL画像子集合、特殊权限URL画像子集合、空URL画像子集合。通过对网络服务器记录的历史访问数据进行统计可以归纳出前述的四种URL画像子集合，这四种URL画像子集合可以构成非恶意URL画像集合作为对待检测URL是否存在恶意的判断标准。其中，常用访问URL画像子集合包括了多个常用访问URL画像，常用访问URL画像是指对网络服务器记录的历史访问数据进行统计后得到的访问量很大的一个或多个URL，常用访问URL画像也可以称为“目的网站常用URL画像”。临时访问URL画像子集合包括了多个临时访问URL画像，临时访问URL画像是指对网络服务器记录的历史访问数据进行统计得到的访问量很小的一个或多个URL，特殊权限URL画像子集合包括了多个特殊权限URL画像，特殊权限URL画像是指对网络服务器记录的历史访问数据进行统计得到的只由少数用户(即特定的用户标识)发起访问的一个或多个URL，空URL画像子集合包括了多个空URL画像，空URL画像是指对网络服务器记录的历史访问数据进行统计得到的空符号。

[0045] 当非恶意URL画像集合，包括：常用访问URL画像子集合、临时访问URL画像子集合、特殊权限URL画像子集合、空URL画像子集合等上述四种非恶意URL图像子集合时，步骤202分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，包括：

[0046] A1、分别判断每个待检测URL是否与常用访问URL画像子集合中的常用访问URL画像相同，若待检测URL与常用访问URL画像子集合中的常用访问URL画像相同，确定待检测URL与常用访问URL画像子集合不存在偏移，若待检测URL与常用访问URL画像子集合中的所有常用访问URL画像都不相同，确定待检测URL与常用访问URL画像子集合存在偏移；

[0047] A2、分别判断每个待检测URL是否与临时访问URL画像子集合中的临时访问URL画像相同，若待检测URL与临时访问URL画像子集合中的临时访问URL画像相同，确定待检测URL与临时访问URL画像子集合不存在偏移，若待检测URL与临时访问URL画像子集合中的所有临时访问URL画像都不相同，确定待检测URL与临时访问URL画像子集合存在偏移；

[0048] A3、分别判断每个待检测URL是否与特殊权限URL画像子集合中的特殊权限URL画像相同，若待检测URL与特殊权限URL画像子集合中的特殊权限URL画像相同，确定待检测URL与特殊权限URL画像子集合不存在偏移，若待检测URL与特殊权限URL画像子集合中的所有特殊权限URL画像都不相同，确定待检测URL与特殊权限URL画像子集合存在偏移；

[0049] A4、分别判断每个待检测URL是否与空URL画像子集合中的空URL画像相同，若待检测URL与空URL画像子集合中的空URL画像相同，确定待检测URL与空URL画像子集合不存在偏移，若待检测URL与空URL画像子集合中的所有空URL画像都不相同，确定待检测URL与空URL画像子集合存在偏移。

[0050] 其中，步骤A1至步骤A4中分别对常用访问URL画像子集合、临时访问URL画像子集合、特殊权限URL画像子集合、空URL画像子集合作为判断待检测URL是否产生偏移的考量标

准时，对同一个客户端在检测时间段内收集到的所有待检测URL分别执行上述步骤A1至步骤A4。以步骤A1中一个待检测URL的偏移判断为例，该待检测URL表示为URL1，常用访问URL画像子集合中包括有多个常用访问URL画像，若常用访问URL画像子集合中有一个常用访问URL画像和该URL1相同，则确定URL1与常用访问URL画像子集合不存在偏移，若常用访问URL画像子集合中所有常用访问URL画像和该URL1都不相同，则确定URL1与常用访问URL画像子集合存在偏移，记录与该常用访问URL画像子集合存在偏移的URL1。

[0051] 进一步的，在执行前述步骤A1至步骤A4的实现场景下，在本发明的另一些实施例中，步骤202分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同之前，本发明实施例提供的恶意探测行为的检测方法还可以包括如下步骤：

[0052] B1、通过日志采集系统收集网络服务器的访问日志，访问日志包括：通过客户端发送的目的URL；

[0053] B2、将目的URL划分到如下四种非恶意URL画像中的至少一种：常用访问URL画像、临时访问URL画像、特殊权限URL画像、空URL画像。

[0054] 其中，步骤B1至步骤B2对四种非恶意URL画像的生成方式进行了举例说明。日志收集系统可以从网络服务器获取到网络数据包，通过该网络数据包获取到网络服务器的访问日志，该访问日志可以包括：通过客户端发送的目的URL，访问日志除了携带目的URL之外，还可以携带如下信息中的至少一种：访问时间、源IP地址、源设备名、源用户名、目的IP地址、目的域名、目的端口、是否成功，返回码。其中，是否成功表示了本次对目的URL的访问是否成功，返回码是指超文本传输协议(HyperText Transfer Protocol，HTTP)状态码(Status Code)，该状态码表示HTTP服务器对于请求HTTP响应状态的3位数字代码。它由一系列以编号排定的文件(Request For Comments，RFC)2616规范定义的，并得到RFC 2518、RFC 2817、RFC 2295、RFC 2774、RFC4918等规范扩展。通过日志收集系统获取到访问日志之后，从该访问日志中可以获取到通过客户端发送的目的URL，还可以获取到该目的URL是否访问成功以及发起URL访问的用户名和请求访问的域名。接下来可以根据访问日志将目的URL划分到如下四种非恶意URL画像中的至少一种：常用访问URL画像、临时访问URL画像、特殊权限URL画像、空URL画像。将访问日志中携带的目的URL划分到至少一种的非恶意URL画像中，从而可以实现对非恶意URL画像的动态更新，使得非恶意URL画像作为待检测URL是否产生偏移的考量标准时能够更准确的判断出待检测URL是否产生了偏移。

[0055] 进一步的，在执行前述步骤B1至步骤B2的实现场景下，在本发明的另一些实施例中，步骤B2将目的URL划分到如下四种非恶意URL画像中的至少一种：常用访问URL画像、临时访问URL画像、特殊权限URL画像、空URL画像，包括如下步骤：

[0056] B21、将满足第一条件的目的URL划分到常用访问URL画像，第一条件包括：在预置的第一单位时间内访问成功的次数大于第一次数阈值；

[0057] B22、将满足第二条件的目的URL划分到临时访问URL画像，第二条件包括：在预置的第一单位时间内访问成功的次数大于0、且小于或等于第一次数阈值；

[0058] B23、将满足第三条件的目的URL划分到空URL画像，第三条件包括：在预置的第二单位时间内访问失败的次数小于第二次数阈值；

[0059] B24、将满足第四条件的目的URL划分到特殊权限URL画像，第四条件包括：在预置的第三单位时间内访问成功的用户所属的集合包括的用户个数小于用户个数阈值。

[0060] 其中，步骤B21至步骤B24对四种非恶意URL画像的生成过程进行详细说明，分别定义了第一条件、第二条件、第三条件、第四条件，每种条件描述了是否可以将目的URL划分到某一种具体的非恶意URL画像。举例说明，以步骤B21的实现过程为例，判断目的URL在预置的第一单位时间内访问成功的次数是否大于第一次数阈值，若大于该第一次数阈值，则可确定该目的URL满足第一条件，则该目的URL可以划分到常用访问URL画像，若小于或等于第一次数阈值，则该目的URL不能划分到常用访问URL画像。依次执行步骤B21至步骤B24的判断过程，可以将目的URL划分到四种非恶意URL画像中的至少一种。

[0061] 在上述步骤B24的实现场景下，第四条件可以包括：在预置的第三单位时间内访问成功的用户所属的集合包括的用户个数小于用户个数阈值。即若某个目的URL仅仅有少数的用户能够访问成功，那么说明该目的URL是需要特殊权限才能访问成功，并不是所有的用户都可以访问，其中，用户个数阈值的取值可以根据应用场景来确定，例如该用户个数阈值可以为3或者4。需要说明的是，第一单位时间、第二单位时间、第三单位时间均可以指的是某一种具体的时间长度，具体的时间长短可以根据应用场景灵活选择，详见后续实施例中的举例说明。第一次数阈值和第二次数阈值均可以指的是具体的次数门限值，具体的次数门限可以根据应用长颈鹿灵活选择，详见后续实施例中的举例说明。

[0062] 步骤203、对与非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量。

[0063] 在本申请实施例中，通过前述步骤202可以判断出同一个客户端的所有待检测URL是否与非恶意URL画像集合存在偏移，然后根据步骤202的判断结果可以对与非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量。在本发明实施例中，计算出每个待检测URL与非恶意URL画像集合是否产生偏移之后，针对非恶意URL图像集合作为参考量统计出存在偏移的待检测URL个数，再对存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量，例如对所有的与非恶意URL画像集合存在偏移的待检测URL的个数进行累计可以得到偏移总量，也可以对所有与非恶意URL画像集合存在偏移的待检测URL的个数进行累计之后再通过对累计结果进行调整从而得到偏移总量。其中，本发明实施例中描述的偏移总量可以认为是客户端在检测时间段内产生的所有探测行为是否为恶意探测行为的考量依据，判断客户端在检测时间段内的偏移总量是否超过预置的偏移阈值。其中，偏移阈值是一个门限值，具体取值可以根据不同场景下的恶意探测行为的爆发量来确定，也可以根据历史经验进行设置，此处不做限定。

[0064] 在本发明的一些实施例中，在执行前述步骤A1至步骤A4的实现场景下，步骤203对与非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量，包括：

[0065] C1、对与常用访问URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第一偏移量；

[0066] C2、对与临时访问URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第二偏移量；

[0067] C3、对与特殊权限URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第三偏移量；

[0068] C4、对与空URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第四

偏移量；

[0069] C5、将第一偏移量、第二偏移量、第三偏移量和第四偏移量进行相加计算，得到客户端在检测时间段内的偏移总量。

[0070] 其中，所有待检测URL可以分别按照步骤A1至步骤A4中的方式进行是否产生偏移的判断，分别对应于步骤A1至步骤A4，可以执行步骤C1至步骤C4，将各个步骤中计算出的偏移量分别定义为“第一偏移量”、“第二偏移量”、“第三偏移量”、“第四偏移量”。以步骤A1中第一偏移量的计算为例，假设共有n个待检测URL，分别为第1个待检测URL、第2个待检测URL、…、第i个待检测URL、…和第n个待检测URL，以第i个待检测URL为例，计算出第i个待检测URL与常用访问URL画像子集合之间是否产生偏移，按照此方式，分别计算n个待检测URL与常用访问URL画像子集合之间是否产生偏移，统计出n个待检测URL中与常用访问URL画像子集合之间产生偏移的待检测URL个数，得到第一偏移量，该第一偏移量为以常用访问URL画像作为考量标准时，所有待检测URL中产生偏移的待检测URL个数之和。步骤A2至步骤A4的计算过程与前述A1的举例过程相类似，不再赘述。最后可以将第一偏移量、第二偏移量、第三偏移量和第四偏移量进行相加计算，得到客户端在检测时间段内的偏移总量。

[0071] 在本发明的另一些实施例中，步骤203对与非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量，包括：

[0072] D1、当存在至少两种类型的非恶意URL画像时，分别统计与所述至少两种类型的所述非恶意URL画像存在偏移的待检测URL的总个数，得到对应于不同类型的非恶意URL画像的待检测URL的总个数；

[0073] D2、获取为每种类型的非恶意URL画像分别配置的权重参数；

[0074] D3、根据权重参数对对应于不同类型的非恶意URL画像的待检测URL的总个数进行加权计算，得到客户端在检测时间段内的偏移总量。

[0075] 其中，当存在至少两种类型的非恶意URL画像时，需要针对每种类型的非恶意URL画像计算出对应于不同类型的非恶意URL画像的待检测URL的总个数，例如前述步骤C1至C4中对于不同的非恶意URL画像所计算出的偏移量。步骤D2中可以通过权重系统为每种类型的非恶意URL画像分别配置权重参数，举例说明，四种非恶意URL画像包括：常用访问URL画像、临时访问URL画像、特殊权限URL画像、空URL画像，这四种非恶意URL画像分别有对应的权重参数。步骤D3中，根据每种类型的非恶意URL画像对应的权重参数对所有待检测URL的偏移量进行加权计算，可以得到客户端在检测时间段内的偏移总量。举例说明，共有m种类型的非恶意URL画像，第i种类型的非恶意URL画像对应的权重参数为p_i，对应于第i种类型的非恶意URL画像的待检测URL的总个数为n_i，则可以通过如下方式计算出偏移总量H：

[0076] $$H=\sum_{i=1}^{m} p_i * n_i$$ 。

[0077] 204、当客户端在检测时间段内的偏移总量超过预置的偏移阈值时，确定客户端在执行恶意探测行为。

[0078] 在本发明实施例中，通过步骤203可以计算出偏移总量，该偏移总量可以认为是客户端在检测时间段内产生的所有探测行为是否为恶意探测行为的考量依据，判断客户端在检测时间段内的偏移总量是否超过预置的偏移阈值。在客户端在检测时间段内的偏移总量超过预置的偏移阈值时，可以确定客户端在执行恶意探测行为。通过对偏移总量的阈值判

断可以确定出客户端是否在执行恶意探测行为。对于发送URL访问总次数较少的用户，也可以通过判断偏移总量确定该客户端是否在执行恶意探测行为，因此本发明实施例可以有效解决恶意探测行为的发现问题，提高恶意探测行为的检测效果。

[0079] 在本发明的另一些实施例中，步骤204确定客户端在执行恶意探测行为之后，本发明实施例提供的恶意探测行为的检测方法还可以包括如下步骤：

[0080] E1、触发报警系统发出异常报警。

[0081] 其中，在检测出用户在执行恶意探测行为之后，还可以触发报警系统发出异常报警，报警系统可以接收异常并发出报警，使得网络服务器接收到存在恶意探测行为的报警。

[0082] 进一步的，在执行前述步骤D1至步骤D3的实现场景下，在本发明的另一些实施例中，步骤204确定客户端在执行恶意探测行为之后，本发明实施例提供的恶意探测行为的检测方法还可以包括如下步骤：

[0083] F1、获取对客户端的恶意探测行为处理结果，并根据恶意探测行为处理结果对每种类型的非恶意URL画像配置的权重参数进行优化调整。

[0084] 其中，确定客户端在执行恶意探测行为之后，可以对该客户端是否存在恶意探测行为进行进一步的处理，根据该恶意探测行为处理结果可以对非恶意URL画像配置的权重参数进行优化调整，使得优化后的权重参数更能够准确的判定出用户是否存在恶意探测行为。

[0085] 在本发明的另一些实施例中，步骤204确定客户端在执行恶意探测行为之后，本发明实施例提供的恶意探测行为的检测方法还可以包括如下步骤：

[0086] G1、获取对客户端的恶意探测行为处理结果，并根据恶意探测行为处理结果对偏移阈值进行优化调整。

[0087] 其中，确定客户端在执行恶意探测行为之后，可以对该客户端是否存在恶意探测行为进行进一步的处理，根据该恶意探测行为处理结果可以对偏移阈值进行优化调整，使得优化后的偏移阈值更能够准确的判定出用户是否存在恶意探测行为。

[0088] 在本发明的前述步骤F1以及步骤G1的实施例中，优化调整方式在于对权重参数的优化以及对偏移阈值的优化，具体的优化方式可以是通过经验微调权重参数或者偏移阈值，然后重新验证恶意探测行为的检测结果是否符合预期表现，若不符合，继续进行优化调整。偏移阈值和权重参数的优化调整是一个动态过程，具体的调整方式可以结合场景进行相应的数值调整，详见后续实施例中的举例说明。

[0089] 通过以上实施例对本发明实施例的描述可知，首先获取客户端发送给网络服务器的探测行为集合，探测行为集合包括：在预置的检测时间段内收集到的至少一个探测行为，其中每个探测行为携带有一个待检测URL，然后分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若待检测URL与非恶意URL画像集合中的非恶意URL画像相同，确定待检测URL与非恶意URL画像集合不存在偏移，若待检测URL与非恶意URL画像集合中的所有非恶意URL画像都不相同，确定待检测URL与非恶意URL画像集合存在偏移，对与非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量，非恶意URL画像通过对网络服务器记录的历史访问数据进行统计得到，当客户端在检测时间段内的偏移总量超过预置的偏移阈值时，确定客户端在执行恶意探测行为。由于非恶意URL画像可以通过对网络服务器记录的历史访问数据进行统计得到，使用该

非恶意URL画像作为参考量，从而通过衡量待检测URL与非恶意URL画像集合是否产生偏移可以计算出客户端在检测时间段内的偏移总量，最后通过对偏移总量的阈值判断可以确定出客户端是否在执行恶意探测行为。对于发送URL访问总次数较少的用户，也可以通过判断偏移总量确定该客户端是否在执行恶意探测行为，因此可以有效解决恶意探测行为的发现问题，提高恶意探测行为的检测效果。

[0090] 为便于更好的理解和实施本发明实施例的上述方案，下面举例相应的应用场景来进行具体说明。

[0091] 本发明实施例可以应用于企业内部web扫描行为的检测。本发明实施例中提出一种基于web业务的被访问URL进行web恶意探测行为发现方案，例如可以发现爬虫、探测管理后台等恶意探测行为，通过网络服务器全部的内网访问数据生成四种非恶意URL画像：目的网站常用URL画像、特殊权限目录URL画像、空URL画像、临时访问URL画像，然后对比个人用户访问行为与以上4个非恶意URL画像的偏移，判定客户端是否在进行web扫描探测，从而有效解决web恶意探测行为的发现问题。

[0092] 在本发明的一种实现场景下，首先通过日志采集系统收集全部的访问日志，并格式化访问日志，包括：时间、源IP、源设备名、源用户名、目的IP、目的域名、目的URL、目的端口、是否成功，返回码等。内网访问数据进入URL画像生成系统后可以生成4类的非恶意URL画像，包括：常用访问URL画像、特殊权限URL画像、临时访问URL画像、空URL画像。然后对比单位时间内网用户访问的目的URL是否为常用访问URL画像、特殊权限URL画像、临时访问URL画像等，通过计算目的URL与上述每一种非恶意URL画像之间的偏离量，可以得到偏移总量，再根据偏移总量是否超过偏移阈值确定是否触发报警。

[0093] 请参阅图3所示，为本发明实施例提供的恶意探测行为的检测方法所应用系统的一种实现场景下的架构示意图。接下来对模块功能进行描述：

[0094] 日志采集系统，包括：日志格式化系统，和网络数据包采集系统。该日志采集系统可用于记录访问日志，并解析为规范格式。然后将全部内网用户的日志数据发送给URL画像生成系统，将单个内网用户的日志数据发送给行为对比画像系统。

[0095] URL画像生成系统可用于生成非恶意URL画像，包括：常用访问URL画像、特殊权限URL画像、空URL画像、临时访问URL画像。

[0096] 行为对比画像系统可用于计算常用访问URL画像偏移、特殊权限URL画像偏移、临时访问URL画像偏移、空URL画像命中。行为对比画像系统可用于计算实时用户访问的目的URL是否偏移非恶意URL画像，另外该行为对比画像系统中还可以设置权重系统，通过该权重系统可以为每种类型的非恶意URL画像配置权重参数，通过偏移量和权重参数计算出偏移总量，若该偏移总量达到偏移阈值则触发报警。

[0097] 报警系统可用于接收异常并通过报警单元报警，然后由应急响应人员进行梳理白名单或进行阈值调整。

[0098] 接下来对上述的日志采集系统、URL画像生成系统、行为对比画像系统和报警系统的具体实现流程进行举例说明。本发明实施例提供的具体流程如下：

[0099] 1、日志采集系统记录并输出规范格式的访问日志。

[0100] A、使用入侵检测系统(Intrusion Detection Systems，IDS)等类似设备记录网络传输的数据包。为保证数据采集完整性，实际应用中尽量做到双机备份。如有条件可在每台

web服务器部署采集器，采集web服务器的访问日志。

[0101] B、解析并格式化访问日志，解析IDS获取的网络包，根据TCP/IP协议+HTTP解析，如为HTTPS则只能从web服务器获取访问日志，获取到6元组并格式化为：时间、源IP、源设备名、源用户名、目的IP、目的域名、目的URL、目的端口、是否成功，返回码。例如20170101221245、192.168.1.2、lennonma-pc1，lennonma、10.14.14.14、www.oa.com、www.oa.com/Index.html，、80、Y、400。

[0102] 2、日志上传到URL画像生成系统，生成常用访问URL画像、临时访问URL画像、特殊权限URL画像、空URL画像，具体方法如下：

[0103] A、常用访问URL画像：

[0104] 统计访问日志，规定单位时间(如15min)内访问次数大于10、且访问成功的URL集合，例如访问成功时HTTP返回码为200。生成过程举例如下：

[0105] 常用访问URL画像：

[0106]

目的域名	目的URL
www.oa.com	Index.html,/login/login.jsp,/host/list/jpg
www.fuli.com	/Welfare/SBCMyWelfare/health.aspx, /Welfare/SBCMyWelfare/welfareitemdetail.aspx?itemId=50, /My
kk.oa.com	/articles/show/317500,/q?kmref=km_header, /discovery?kmref=km_header, /group/16280/articles/show/297736?kmref=discovery_page

[0107] B、临时访问URL画像：

[0108] 统计访问日志，规定单位时间(如15min)内访问次数>0且<＝10的访问成功的URL集合，其中，访问成功HTTP返回码为200，生成过程举例说明如下：

[0109] 临时访问URL画像：

[0110]

目的域名	目的URL
www.oa.com	/host/home/10.14.13.213
www.fuli.com	/forum/3835/thread/view/375596
kk.oa.com	/task/ctr_module/index
L.oa.com	Index.php,/list/cc.html

[0111] C、空URL画像：

[0112] 统计访问日志，规定单位时间(如24h)内访问次数<2的访问失败的URL集合，其中，访问HTTP返回码为404，生成用户偶尔输错等造成的空URL画像，生成过程举例说明如下：

[0113] 空URL画像：

[0114]

目的域名	目的URL
www.oa.com	/host/temp/temp.jsp
www.fuli.com	/test/test.php

kk.oa.com	/soc/ip.js

[0115] D、特殊权限URL画像:

[0116] 持续计算,统计n天(如5日)的网络数据,将同一URL只被同一群用户访问成功的URL集合,该群人数可以小于m,m可以为3人。生成过程举例说明如下

[0117] 特殊权限URL画像:

[0118]

目的域名	目的URL	访问人集合
www.oa.com	/login/login.jsp	Pony,tony,lw
www.fuli.com	/manager	Sy,ck,lennon
kk.oa.com	/special/ll.php	Liu,li,DD

[0119] 3、计算实时用户访问的目的URL是否与非恶意URL画像产生偏移,并统计偏移总量,若偏移总量超过偏移阈值,则触发报警系统进行报警。

[0120] A、计算用户访问(只看成功访问)和常用访问URL画像的偏移。默认权重为0.4,如下表1所示,Y表示是(Yes),N表示否(No):

[0121]

用户访问域名	用户访问URL	目的域名画像	目的URL画像	是否偏移
www.oa.com	Index.jsp,/login/login.jsp,/host/list/jpg,	www.oa.com	Index.html,/login/login.jsp,/host/list/jpg	Y(例如,Index.jsp,不存在于目的URL画像中)
www.fuli.com	/Welfare/SBCMyWelfare	www.fuli.com	/Welfare/SBCMyWelf	Y(例如,/forum/3835/t

[0122]

	/health.aspx, /forum/3835/thread/view/375596		are/health.aspx, /Welfare/SBCMyWelfare/welfareitemdetail.aspx?itemId=50, /My	hread/view/375596，不存在于目的 URL 画像中)
kk.oa.com	/articles/show/12345 , /group/16280/articles/show/297736?kmref=discovery_page	kk.oa.com	/articles/show/317500, /q?kmref=km_header, /discovery?kmref=km_header, /group/16280/articles/show/297736?kmref=discovery_page	Y(例 如 , /articles/show/12345 不存在于目的 URL 画像中)
L.oa.com	Index.php,/list/cc.html	/	/	Y(完全不存在于目的 URL 画像)
Y. oa.com	1.jsp	/	/	Y(完全不存在于目的 URL 画像)

[0123] 其中,表1中加下划线的,例如“Index.jsp,”表示用于举例说明的待检测URL。

[0124] 通过上述表1计算偏移量,上表1最后一列,有一个Y就算偏移量为1,则偏移量计算为(Y+Y+Y+Y+Y)*0.4＝2。

[0125] B、计算用户访问(只看成功访问)和临时访问URL画像的偏移,默认权重为0.8,如

下表2所示：

[0126]

用户访问域名	用户访问URL	目的域名画像	目的URL画像	是否偏移
www.oa.com	Index.jsp,/login/login.jsp,/host/list/jpg,	www.oa.com	/host/home/10.14.13.213	Y(均未出现在目的URL画像中)
www.fuli.com	/Welfare/SBCMyWelfare/health.aspx,/forum/3835/thread/view/375596	www.fuli.com	/forum/3835/thread/view/375596	N(已在常用访问URL画像中的不再对比，且/forum/3835/thread/view/375596在画像中)
kk.oa.com	/articles/show/12345 ,/group/16280/articles/show/297736?kmref=discovery_page	kk.oa.com	/task/ctr_module/index	Y(均未出现在目的URL画像中)
L.oa.com	Index.php,/list/cc.html	L.oa.com	Index.php,/list/cc.html	N (与画像匹配)
Y. oa.com	1.jsp	/	/	Y(完全不存在于目的URL画像)

[0127]

[0128] 通过上述表2计算偏移量，上表2最后一列，有一个Y就算偏移量为1，则偏移量计算为(Y+Y+Y)*0.8＝2.4。

[0129] C、计算用户访问(只看访问失败)和空URL画像的偏移，默认权重为1.5，如下表3所示：

[0130]

用户访问域名	用户访问URL	目的域名画像	目的URL画像	是否偏移
www.oa.com	/host/temp/temp.jsp	www.oa.com	/host/temp/temp.jsp	N (与画像匹配,定义为输错)
www.fuli.com	/login.html,/login/login.jsp	www.fuli.com	/test/test.php	Y
/	/	kk.oa.com	/soc/ip.js	N

[0131] 通过上述表3计算偏移量,上表3最后一列,有一个Y就算偏移量为1,则偏移量计算为(Y)*1.5=1.5。

[0132] D、计算用户访问(不区分访问成功或者失败)和特殊权限URL画像的偏移,默认权重为2,如下表4所示:

[0133]

用户名	用户访问域名	用户访问URL	目的域名画像	目的URL画像	画像访问人集合	是否偏移
ck	www.oa.com	Index.jsp,/login/login.jsp,/host/list/jpg,/host/temp/temp.jsp	www.oa.com	/login/login.jsp	Pony,tony,lw	Y (与画像访问人集合不匹配)
ck	www.fuli.	/Welfare/	www.fuli.	/manager	Sy,ck,lenno	N(完全

[0134]

	com	SBCMyWelfare/health.aspx,/forum/3835/thread/view/375596，/login.html	com		n	不在画像内)
ck	kk.oa.com	/articles/show/12345，/group/16280/articles/show/297736?kmref=discovery_page	kk.oa.com	/special/ll.php	Liu,li,DD	N(完全不在画像内)
ck	L.oa.com	Index.php,/list/cc.html	/	/	/	N(完全不在画像内)
ck	Y. oa.com	1.jsp	/	/	/	N(完全不在画像内)

[0135] 通过上述表4计算偏移量，上表4最后一列，有一个Y就算偏移量为1，则偏移量计算为(Y)*2＝2。

[0136] E、通过前述A、B、C、D的举例说明，接下来根据前述的权重系统计算是否产生报警。

[0137] 本发明实施例中采用如下的判断方式：常用访问URL画像偏移+临时访问URL画像偏移+空URL画像偏移+特殊权限URL画像偏移>偏移阈值。举例说明如下，假设偏移阈值设置为5，常用访问URL画像偏移+临时访问URL画像偏移+空URL画像偏移+特殊权限URL画像偏移＝2+2.4+1.5+2＝7.9>5，则偏移总量超过偏移阈值。

[0138] 最后，通过上述方式确定用户存在恶意探测行为时，可以将异常发送到报警系统，报警系统产生安全报警应急。响应人员处理异常，确定异常或误报，进一步优化偏移阈值和

权重参数。

[0139] 需要说明的是，对于前述的各方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本发明并不受所描述的动作顺序的限制，因为依据本发明，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定是本发明所必须的。

[0140] 为便于更好的实施本发明实施例的上述方案，下面还提供用于实施上述方案的相关装置。

[0141] 请参阅图4-a所示，本发明实施例提供的一种恶意探测行为的检测装置400，可以包括：探测行为获取模块401、URL判断模块402、偏移总量计算模块403和检测模块404，其中，

[0142] 探测行为获取模块401，用于获取客户端发送给网络服务器的探测行为集合，所述探测行为集合包括：在预置的检测时间段内收集到的至少一个探测行为，其中每个探测行为携带有一个待检测统一资源定位符URL；

[0143] URL判断模块402，用于分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若所述待检测URL与所述非恶意URL画像集合中的非恶意URL画像相同，确定所述待检测URL与所述非恶意URL画像集合不存在偏移，若所述待检测URL与所述非恶意URL画像集合中的所有非恶意URL画像都不相同，确定所述待检测URL与所述非恶意URL画像集合存在偏移，所述非恶意URL画像通过对所述网络服务器记录的历史访问数据进行统计得到；

[0144] 偏移总量计算模块403，用于对与所述非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到所述客户端在所述检测时间段内的偏移总量；

[0145] 检测模块404，用于当所述客户端在所述检测时间段内的偏移总量超过预置的偏移阈值时，确定所述客户端在执行恶意探测行为。

[0146] 在本发明的一些实施例中，所述非恶意URL画像集合，包括：常用访问URL画像子集合、临时访问URL画像子集合、特殊权限URL画像子集合、空URL画像子集合；在这种实现场景下，如图4-b所示，所述URL判断模块402，包括：

[0147] 第一判断子模块4021，用于分别判断每个待检测URL是否与所述常用访问URL画像子集合中的常用访问URL画像相同，若所述待检测URL与所述常用访问URL画像子集合中的常用访问URL画像相同，确定所述待检测URL与所述常用访问URL画像子集合不存在偏移，若所述待检测URL与所述常用访问URL画像子集合中的所有常用访问URL画像都不相同，确定所述待检测URL与所述常用访问URL画像子集合存在偏移；

[0148] 第二判断子模块4022，用于分别判断每个待检测URL是否与所述临时访问URL画像子集合中的临时访问URL画像相同，若所述待检测URL与所述临时访问URL画像子集合中的临时访问URL画像相同，确定所述待检测URL与所述临时访问URL画像子集合不存在偏移，若所述待检测URL与所述临时访问URL画像子集合中的所有临时访问URL画像都不相同，确定所述待检测URL与所述临时访问URL画像子集合存在偏移；

[0149] 第三判断子模块4023，用于分别判断每个待检测URL是否与所述特殊权限URL画像子集合中的特殊权限URL画像相同，若所述待检测URL与所述特殊权限URL画像子集合中的

特殊权限URL画像相同，确定所述待检测URL与所述特殊权限URL画像子集合不存在偏移，若所述待检测URL与所述特殊权限URL画像子集合中的所有特殊权限URL画像都不相同，确定所述待检测URL与所述特殊权限URL画像子集合存在偏移；

[0150] 第四判断子模块4024，用于分别判断每个待检测URL是否与所述空URL画像子集合中的空URL画像相同，若所述待检测URL与所述空URL画像子集合中的空URL画像相同，确定所述待检测URL与所述空URL画像子集合不存在偏移，若所述待检测URL与所述空URL画像子集合中的所有空URL画像都不相同，确定所述待检测URL与所述空URL画像子集合存在偏移。

[0151] 在本申请的一些实施例中，请参阅图4-c所示，偏移总量计算模块403，包括：

[0152] 第一计算子模块4031，用于对与所述常用访问URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第一偏移量；

[0153] 第二计算子模块4032，用于对与所述临时访问URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第二偏移量；

[0154] 第三计算子模块4033，用于对与所述特殊权限URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第三偏移量；

[0155] 第四计算子模块4034，用于对与所述空URL画像子集合存在偏移的待检测URL的个数进行求和计算，得到第四偏移量；

[0156] 第五计算子模块4035，用于将所述第一偏移量、所述第二偏移量、所述第三偏移量和所述第四偏移量进行相加计算，得到所述客户端在所述检测时间段内的偏移总量。

[0157] 在本申请的一些实施例中，请参阅图4-d所示，基于图4-b所示的URL判断模块，所述恶意探测行为的检测装置400还包括：日志收集模块404和行为对比模块405，其中，在图4-d中没有对URL判断模块402的组成部分进行示意说明，详见图4-b所示。

[0158] 所述日志收集模块404，用于所述URL判断模块402分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同之前，通过日志采集系统收集所述网络服务器的访问日志，所述访问日志包括：通过客户端发送的目的URL；

[0159] 所述行为对比模块405，用于将所述目的URL划分到如下四种非恶意URL画像中的至少一种：常用访问URL画像、临时访问URL画像、特殊权限URL画像、空URL画像。

[0160] 进一步的，请参阅图4-e所示，基于4-d所示的恶意探测行为的检测装置400，其中，在图4-e中没有对恶意探测行为的检测装置400的组成部分进行示意说明，所述行为对比模块405，包括：

[0161] 第一划分子模块4051，用于将满足第一条件的目的URL划分到常用访问URL画像，所述第一条件包括：在预置的第一单位时间内访问成功的次数大于第一次数阈值；

[0162] 第二划分子模块4052，用于将满足第二条件的目的URL划分到临时访问URL画像，所述第二条件包括：在预置的第一单位时间内访问成功的次数大于0、且小于或等于所述第一次数阈值；

[0163] 第三划分子模块4053，用于将满足第三条件的目的URL划分到空URL画像，所述第三条件包括：在预置的第二单位时间内访问失败的次数小于第二次数阈值；

[0164] 第四划分子模块4054，用于将满足第四条件的目的URL划分到特殊权限URL画像，所述第四条件包括：在预置的第三单位时间内访问成功的用户所属的集合包括的用户个数小于用户个数阈值。

[0165] 在本发明的一些实施例中，请参阅图4-f所示，所述URL判断模块402，包括：

[0166] 分类统计子模块4025，用于当存在至少两种类型的所述非恶意URL画像时，分别统计与所述至少两种类型的所述非恶意URL画像存在偏移的待检测URL的总个数，得到对应于不同类型的非恶意URL画像的待检测URL的总个数；

[0167] 权重获取子模块4026，用于获取为每种类型的所述非恶意URL画像分别配置的权重参数；

[0168] 加权计算子模块4027，用于根据所述权重参数对所述对应于不同类型的非恶意URL画像的待检测URL的总个数进行加权计算，得到所述客户端在所述检测时间段内的偏移总量。

[0169] 请参阅图4-g所示，相对于图4-a所示，所述恶意探测行为的检测装置400还包括：第一优化模块406，其中，

[0170] 所述第一优化模块406，用于所述检测模块403确定所述客户端在执行恶意探测行为之后，获取对所述客户端的恶意探测行为处理结果，并根据所述恶意探测行为处理结果对所述每种类型的非恶意URL画像配置的权重参数进行优化调整。

[0171] 请参阅图4-h所示，相对于图4-a所示，所述恶意探测行为的检测装置400还包括：第二优化模块407，其中，

[0172] 所述第二优化模块407，用于所述检测模块403确定所述客户端在执行恶意探测行为之后，获取对所述客户端的恶意探测行为处理结果，并根据所述恶意探测行为处理结果对所述偏移阈值进行优化调整。

[0173] 通过以上实施例对本发明实施例的描述可知，首先获取客户端发送给网络服务器的探测行为集合，探测行为集合包括：在预置的检测时间段内收集到的至少一个探测行为，其中每个探测行为携带有一个待检测URL，然后分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若所述待检测URL与所述非恶意URL画像集合中的非恶意URL画像相同，确定所述待检测URL与所述非恶意URL画像集合不存在偏移，若所述待检测URL与所述非恶意URL画像集合中的所有非恶意URL画像都不相同，确定所述待检测URL与所述非恶意URL画像集合存在偏移，对与所述非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量，非恶意URL画像通过对网络服务器记录的历史访问数据进行统计得到，当客户端在检测时间段内的偏移总量超过预置的偏移阈值时，确定客户端在执行恶意探测行为。由于非恶意URL画像可以通过对网络服务器记录的历史访问数据进行统计得到，使用该非恶意URL画像作为参考量，从而通过衡量待检测URL与非恶意URL画像集合是否产生偏移可以计算出客户端在检测时间段内的偏移总量，最后通过对偏移总量的阈值判断可以确定出客户端是否在执行恶意探测行为。对于发送URL访问总次数较少的用户，也可以通过判断偏移总量确定该客户端是否在执行恶意探测行为，因此可以有效解决恶意探测行为的发现问题，提高恶意探测行为的检测效果。

[0174] 图5是本发明实施例提供的一种服务器结构示意图，该服务器1100可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上中央处理器(central processing units，CPU)1122(例如，一个或一个以上处理器)和存储器1132，一个或一个以上存储应用程序1142或数据1144的存储介质1130(例如一个或一个以上海量存储设备)。其中，存储器1132和存储介质1130可以是短暂存储或持久存储。存储在存储介质1130的程序可以包括一

个或一个以上模块(图示没标出),每个模块可以包括对服务器中的一系列指令操作。更进一步地,中央处理器1122可以设置为与存储介质1130通信,在服务器1100上执行存储介质1130中的一系列指令操作。

[0175] 服务器1100还可以包括一个或一个以上电源1126,一个或一个以上有线或无线网络接口1150,一个或一个以上输入输出接口1158,和/或,一个或一个以上操作系统1141,例如Windows ServerTM,Mac OS XTM,UnixTM,LinuxTM,FreeBSDTM等等。

[0176] 上述实施例中由服务器所执行的恶意行为的检测方法步骤可以基于该图5所示的服务器结构。

[0177] 另外需说明的是,以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到至少一个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。另外,本发明提供的装置实施例附图中,模块之间的连接关系表示它们之间具有通信连接,具体可以实现为一条或多条通信总线或信号线。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0178] 通过以上的实施方式的描述,所属领域的技术人员可以清楚地了解到本发明可借助软件加必需的通用硬件的方式来实现,当然也可以通过专用硬件包括专用集成电路、专用CPU、专用存储器、专用元器件等来实现。一般情况下,凡由计算机程序完成的功能都可以很容易地用相应的硬件来实现,而且,用来实现同一功能的具体硬件结构也可以是多种多样的,例如模拟电路、数字电路或专用电路等。但是,对本发明而言更多情况下软件程序实现是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在可读取的存储介质中,如计算机的软盘、U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述的方法。

[0179] 综上所述,以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照上述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对上述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

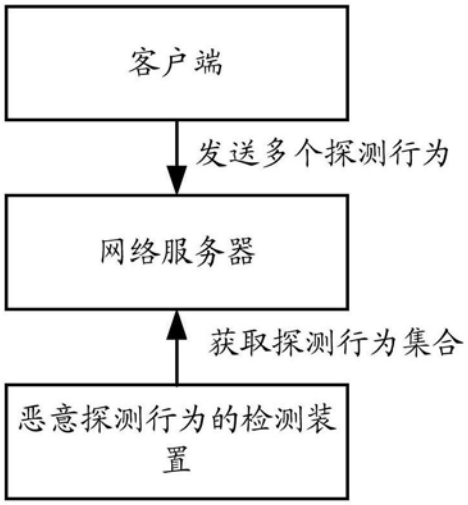

客户端
发送多个探测行为
网络服务器
获取探测行为集合
恶意探测行为的检测装置

图1

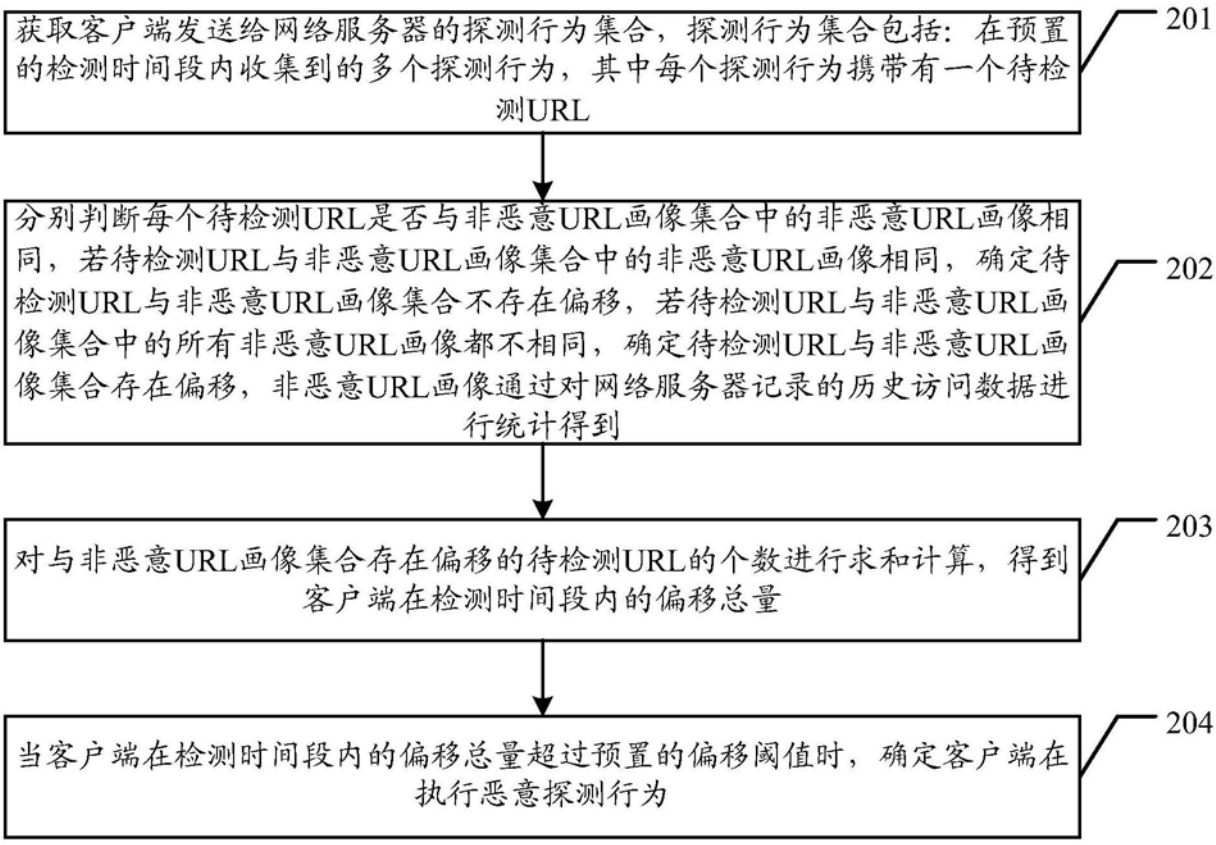

获取客户端发送给网络服务器的探测行为集合，探测行为集合包括：在预置的检测时间段内收集到的多个探测行为，其中每个探测行为携带有一个待检测URL
201
分别判断每个待检测URL是否与非恶意URL画像集合中的非恶意URL画像相同，若待检测URL与非恶意URL画像集合中的非恶意URL画像相同，确定待检测URL与非恶意URL画像集合不存在偏移，若待检测URL与非恶意URL画像集合中的所有非恶意URL画像都不相同，确定待检测URL与非恶意URL画像集合存在偏移，非恶意URL画像通过对网络服务器记录的历史访问数据进行统计得到
202
对与非恶意URL画像集合存在偏移的待检测URL的个数进行求和计算，得到客户端在检测时间段内的偏移总量
203
当客户端在检测时间段内的偏移总量超过预置的偏移阈值时，确定客户端在执行恶意探测行为
204

图2

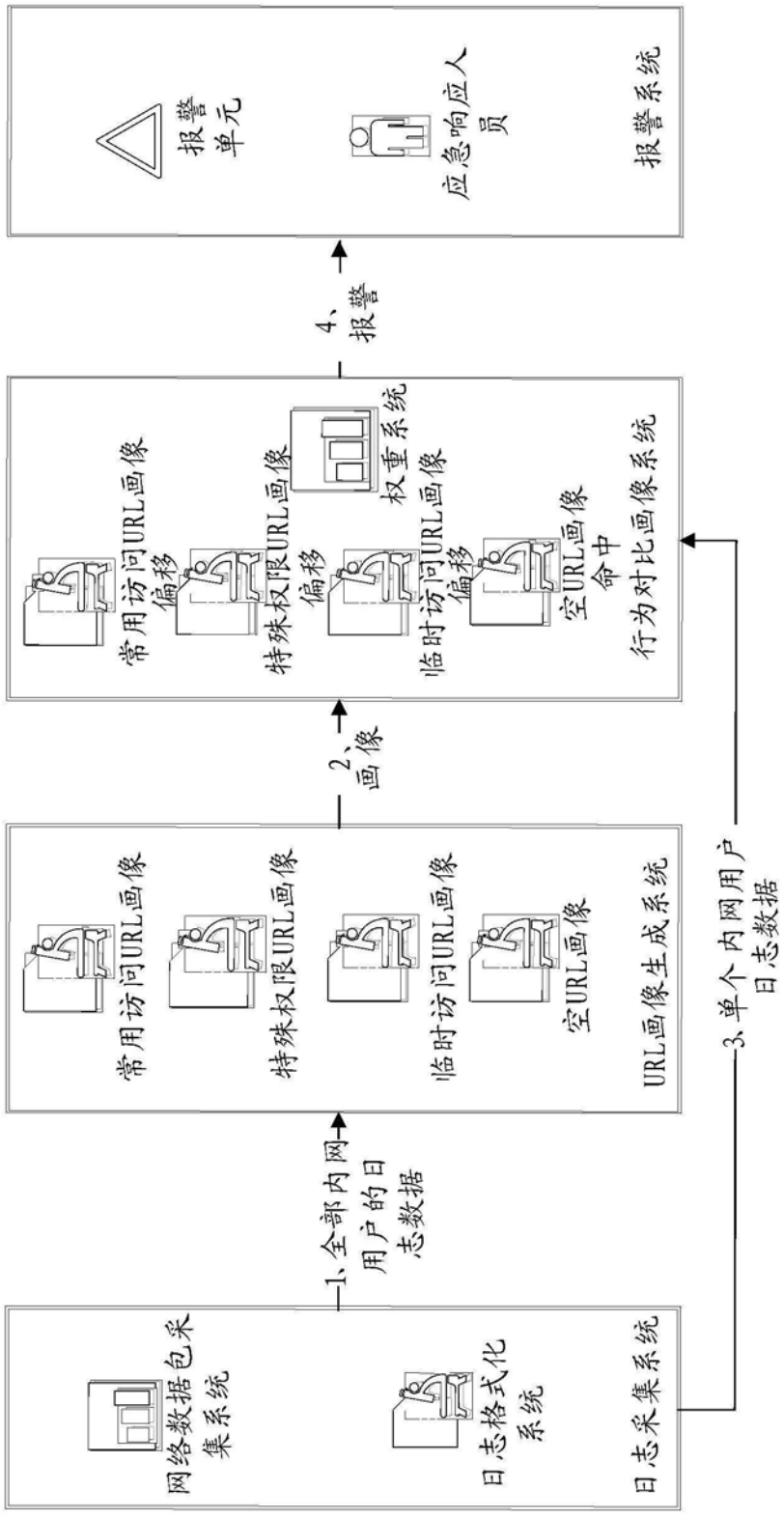

网络数据包采集系统
日志格式化系统
日志采集系统
1、全部内网用户的日志数据
常用访问URL画像
特殊权限URL画像
临时访问URL画像
空URL画像
URL画像生成系统
2、画像
常用访问URL画像
偏移
特殊权限URL画像
偏移
临时访问URL画像
偏移
空URL画像
命中
权重系统
行为对比画像系统
3、单个内网用户日志数据
4、报警
报警单元
应急响应人员
报警系统

图3

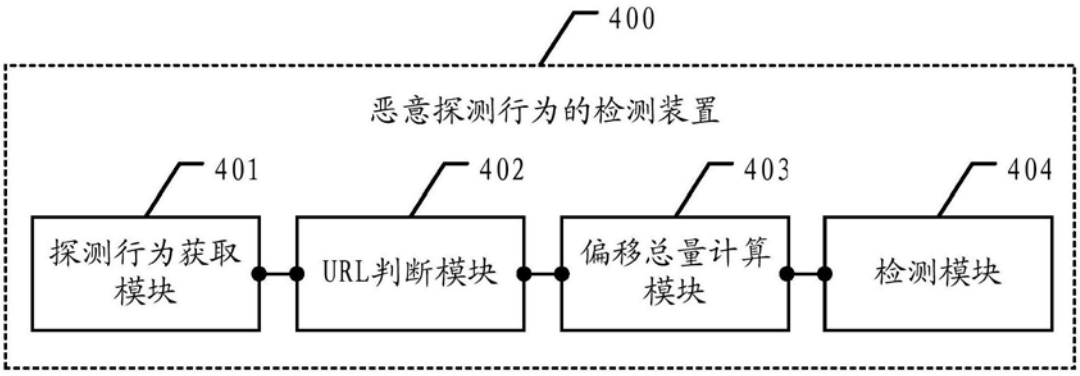

400
恶意探测行为的检测装置
401
探测行为获取模块
402
URL判断模块
403
偏移总量计算模块
404
检测模块

图4-a

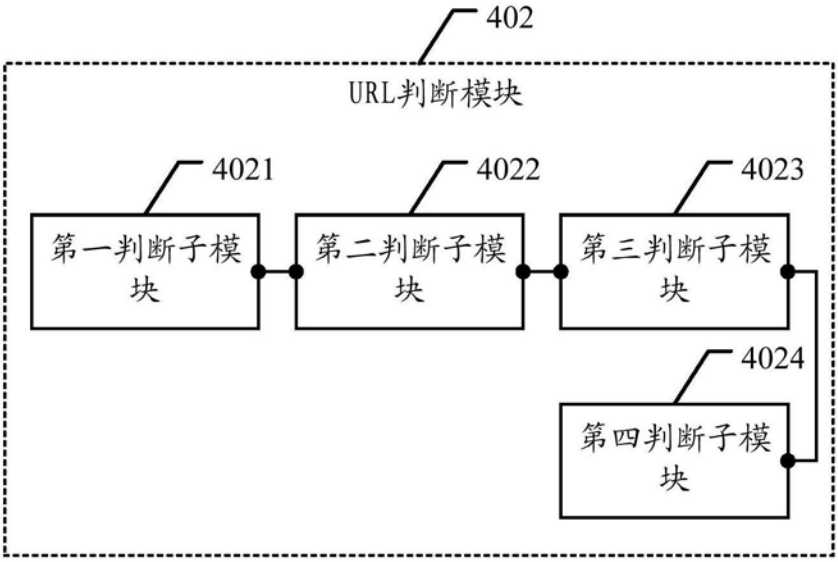

402
URL判断模块
4021
第一判断子模块
4022
第二判断子模块
4023
第三判断子模块
4024
第四判断子模块

图4-b

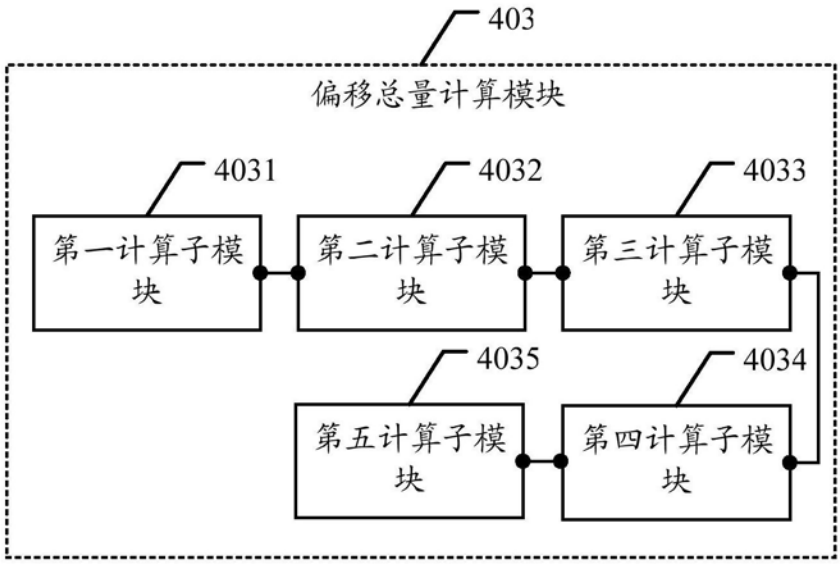

403
偏移总量计算模块
4031
第一计算子模块
4032
第二计算子模块
4033
第三计算子模块
4035
第五计算子模块
4034
第四计算子模块

图4-c

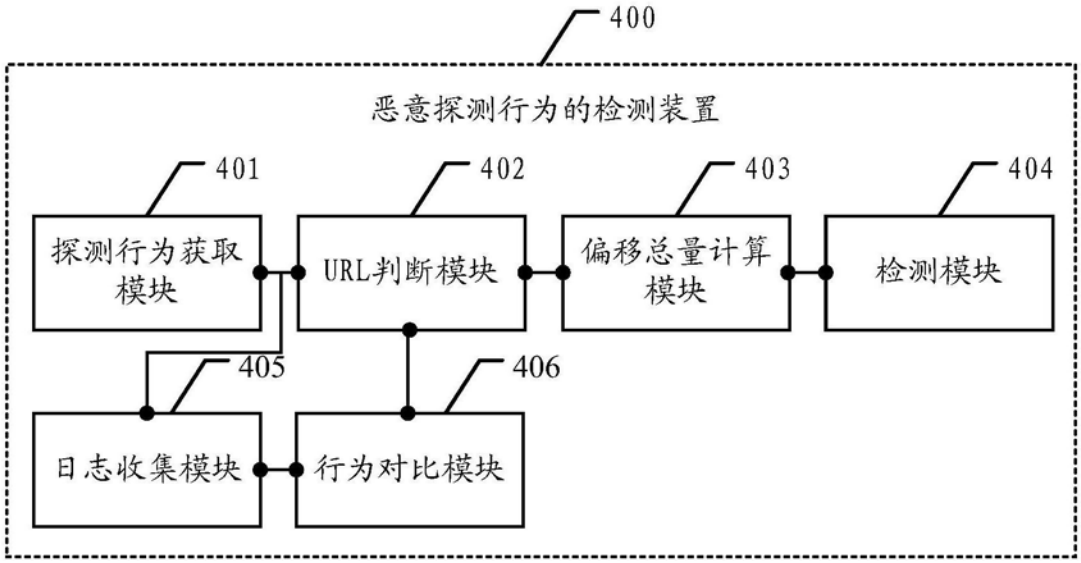

400
恶意探测行为的检测装置
401
探测行为获取模块
402
URL判断模块
403
偏移总量计算模块
404
检测模块
405
日志收集模块
406
行为对比模块

图4-d

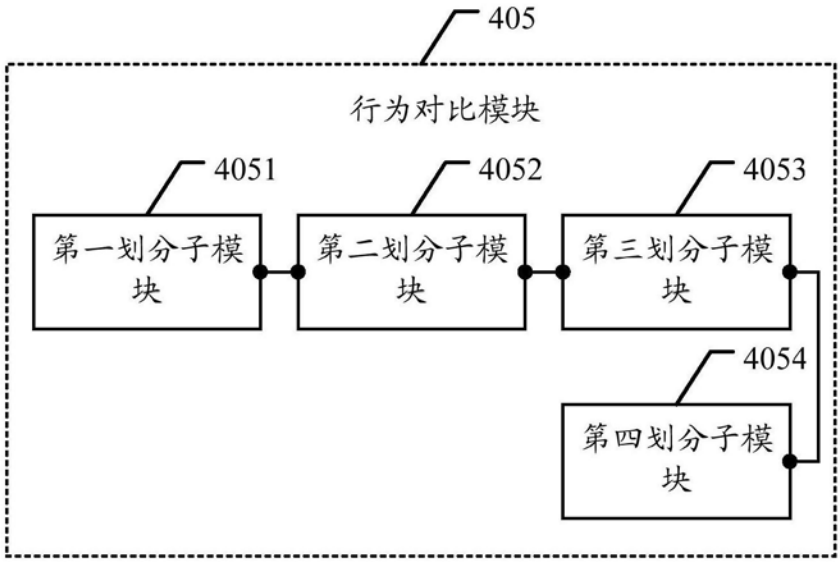

405
行为对比模块
4051
第一划分子模块
4052
第二划分子模块
4053
第三划分子模块
4054
第四划分子模块

图4-e

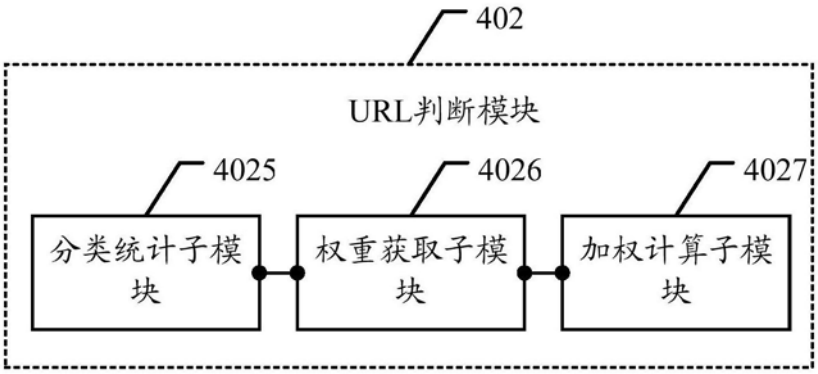

402
URL判断模块
4025
分类统计子模块
4026
权重获取子模块
4027
加权计算子模块

图4-f

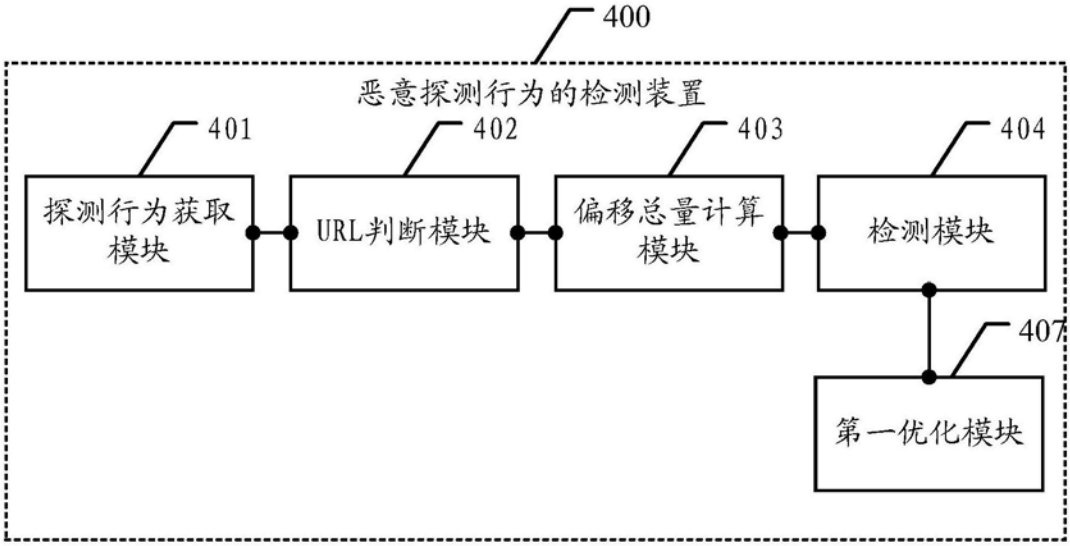

400
恶意探测行为的检测装置
401
探测行为获取模块
402
URL判断模块
403
偏移总量计算模块
404
检测模块
407
第一优化模块

图4-g

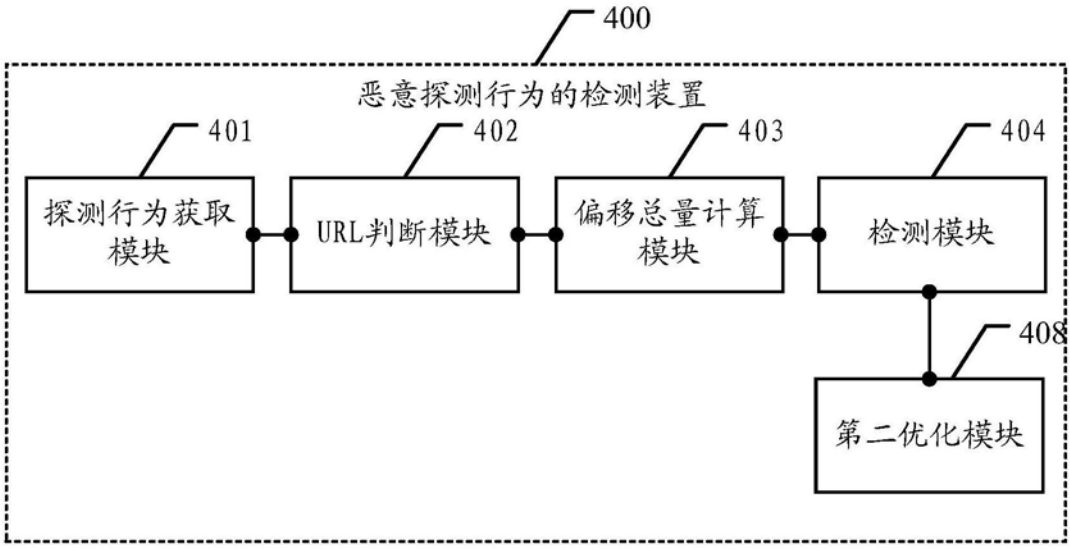

400
恶意探测行为的检测装置
401
探测行为获取模块
402
URL判断模块
403
偏移总量计算模块
404
检测模块
408
第二优化模块

图4-h

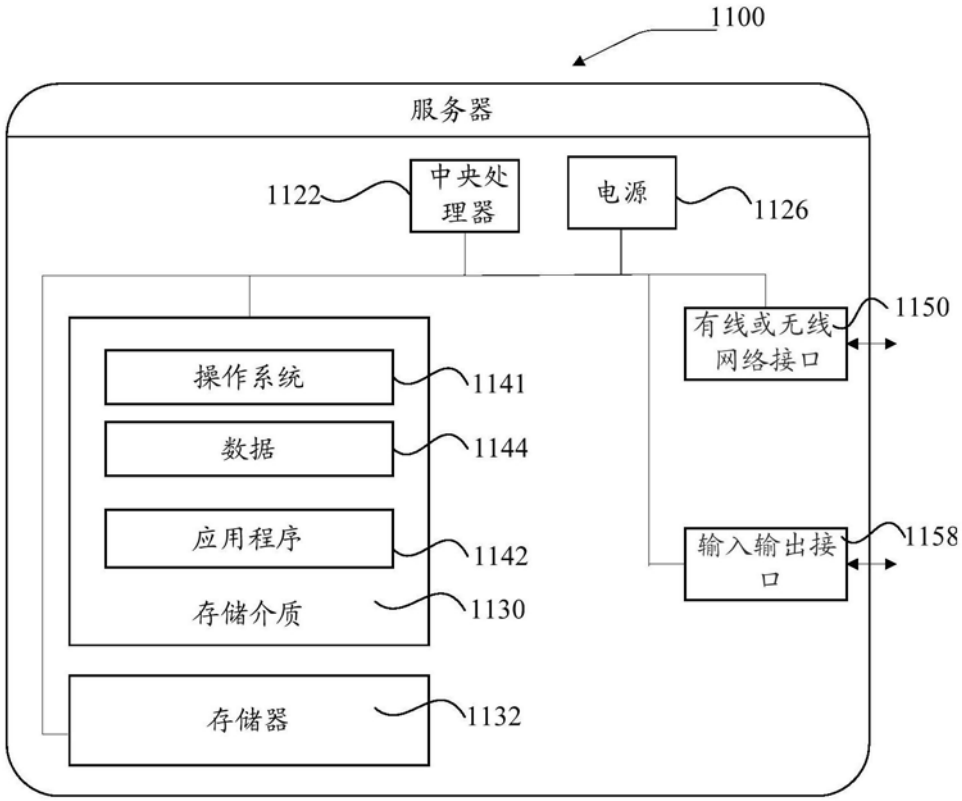

1100
服务器
1122
中央处理器
电源
1126
有线或无线网络接口
1150
操作系统
1141
数据
1144
应用程序
1142
存储介质
1130
输入输出接口
1158
存储器
1132

图5