DNS查询流量异常的检测方法、装置和系统

申请号：201010034253.2
申请日：2010-01-18

申请(专利权)人　中国科学院计算机网络信息中心
地址　100190 北京市海淀区中关村南四街四号北京349信箱6分箱
发明(设计)人　李晓东 毛伟 王正 王欣 金键

主分类号　H04L12/26(2006.01)I
分类号　H04L12/26(2006.01)I H04L29/12(2006.01)I
公开(公告)号　101841435A
公开(公告)日　2010-09-22
专利代理机构　北京同立钧成知识产权代理有限公司 11205
代理人　刘芳

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 101841435 A
(43)申请公布日 2010.09.22

(21)申请号 201010034253.2

(22)申请日 2010.01.18

(71)申请人 中国科学院计算机网络信息中心
地址 100190 北京市海淀区中关村南四街四号北京 349 信箱 6 分箱

(72)发明人 李晓东 毛伟 王正 王欣 金键

(74)专利代理机构 北京同立钧成知识产权代理有限公司 11205
代理人 刘芳

(51)Int.Cl.
H04L 12/26(2006.01)
H04L 29/12(2006.01)

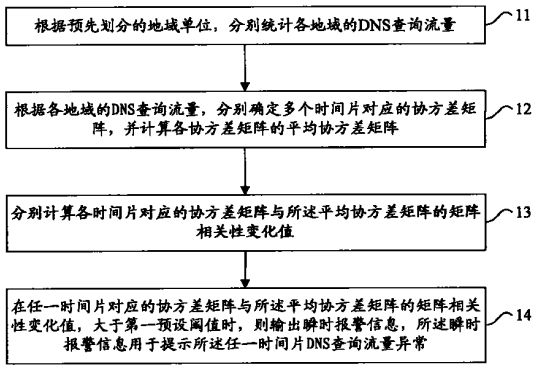
权利要求书 2 页 说明书 10 页 附图 3 页

(54)发明名称

DNS 查询流量异常的检测方法、装置和系统

(57)摘要

本发明涉及一种 DNS 查询流量异常的检测方法、装置和系统，属于互联网技术领域；该检测方法包括：根据预先划分的地域单位，分别统计各地域的 DNS 查询流量；根据各地域的 DNS 查询流量，分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵；分别计算各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值；在任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值时，则输出瞬时报警信息，用于提示所述任一时间片 DNS 查询流量异常。本发明有利于降低域名系统查询流量异常的漏检率。



CN 101841435 A

1. 一种 DNS 查询流量异常的检测方法，其特征在于，包括：

根据预先划分的地域单位，分别统计各地域的 DNS 查询流量；

根据各地域的 DNS 查询流量，分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵；

分别计算各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值；

在任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值时，则输出瞬时报警信息，所述瞬时报警信息用于提示所述任一时间片 DNS 查询流量异常。

2. 根据权利要求 1 所述的检测方法，其特征在于，在任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值时，还包括：

根据所述任一时间片对应的协方差矩阵以及所述平均协方差矩阵分别包括的各元素，分别计算所述任一时间片内各地域的属性相关性变化值，任一元素为任两个地域的 DNS 查询流量的协方差；

对所述任一时间片内各地域的属性相关性变化值进行聚类分析，得到异常地域属性，所述异常地域属性对应所述任一时间片 DNS 查询流量异常所在的地域。

3. 根据权利要求 1 所述的检测方法，其特征在于，还包括：

对所述任一时间片对应的协方差矩阵进行平滑处理，以使所述任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，小于或等于第一预设阈值。

4. 根据权利要求 1 所述的检测方法，其特征在于，所述多个时间片的长度相同或不同，且相邻的两个时间片在时间轴上相连或相互重叠。

5. 根据权利要求 1-4 任一所述的检测方法，其特征在于，还包括：

检测在一预设时间段包括的各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第二预设阈值的矩阵数量，或者大于第二预设阈值且小于或等于第一预设阈值的矩阵数量；

在所述矩阵数量大于第三预设阈值时，则输出时间段报警信息，所述时间段报警信息用于提示所述时间段 DNS 查询流量异常。

6. 根据权利要求 5 所述的检测方法，其特征在于，任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，等于所述任一时间片对应的协方差矩阵与所述平均协方差矩阵的距离。

7. 一种 DNS 查询流量异常的检测装置，其特征在于，包括：

地域流量统计模块，用于根据预先划分的地域单位，分别统计各地域的 DNS 查询流量；

矩阵确定模块，用于根据各地域的 DNS 查询流量，分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵；

矩阵相关性计算模块，用于分别计算各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值；

瞬时报警模块，用于在任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值时，则输出瞬时报警信息，所述瞬时报警信息用于提示所述任一时间片 DNS 查询流量异常。

8. 根据权利要求 7 所述的检测装置，其特征在于，还包括：

属性相关性计算模块，用于根据所述任一时间片对应的协方差矩阵以及所述平均协方差矩阵分别包括的各元素，分别计算所述任一时间片内各地域的属性相关性变化值，任一元素为任两个地域的 DNS 查询流量的协方差；

异常地域属性确定模块，用于对所述任一时间片内各地域的属性相关性变化值进行聚类分析，得到异常地域属性，所述异常地域属性对应所述任一时间片 DNS 查询流量异常所在的地域。

9. 根据权利要求 7 所述的检测装置，其特征在于，还包括：

平滑处理模块，用于对所述任一时间片对应的协方差矩阵进行平滑处理，以使所述任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，小于或等于第一预设阈值。

10. 根据权利要求 7 所述的检测装置，其特征在于，还包括：

时间片确定模块，用于确定所述多个时间片，所述多个时间片的长度相同或不同，且相邻的两个时间片在时间轴上相连或相互重叠。

11. 根据权利要求 7-10 任一所述的检测装置，其特征在于，还包括：

矩阵数量确定模块，用于检测在一预设时间段包括的各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第二预设阈值的矩阵数量，或者大于第二预设阈值且小于或等于第一预设阈值的矩阵数量；

时间段报警模块，用于在所述矩阵数量大于第三预设阈值时，则输出时间段报警信息，所述时间段报警信息用于提示所述时间段 DNS 查询流量异常。

12. 一种 DNS 查询流量异常的检测系统，其特征在于，包括如权利要求 7-11 任一所述的 DNS 查询流量异常的检测装置。

13. 根据权利要求 12 所述的检测系统，其特征在于，还包括：分别与所述检测装置以可通信方式连接的多个 DNS 分节点；

任一所述 DNS 分节点，用于统计各地域的 DNS 查询流量，并将统计到的 DNS 查询流量上报给所述检测装置。

DNS 查询流量异常的检测方法、装置和系统

技术领域

[0001] 本发明涉及互联网技术领域，特别是涉及一种 DNS 查询流量异常的检测方法、装置和系统。

背景技术

[0002] 域名系统 (Domain Name System，DNS) 主要用于完成从域名到 IP 地址的映射及其它互联网资源的解析，是当今互联网中重要的基础设施之一。由于网络流量逐年快速增长和网络拓扑的复杂性，在宏观上直接监测互联网流量变得越来越困难；而 DNS 查询流量作为整个互联网流量的重要反映，对其的监测更易实现、更具可行性。

[0003] 在进行 DNS 查询流量异常检测过程中，需要能够描述 DNS 服务器所在网络系统的正常情况下的行为，以建立正常行为的分布规律，并且要能够快速准确的发现此网络系统中出现的异常行为，如攻击行为，以便网络系统有足够的响应时间。

[0004] 现有技术主要选择一些检测特征，以区分 DNS 查询流量的正常状态和异常状态，主要包括以下三种：

[0005] 1) 域名解析失败率特征：

[0006] 虚假域名请求最终会解析失败，而正常情况下域名解析一般能够正确执行。随着搜索引擎的不断进步，正常情况下，需要用户手工输入域名的情况以及输入错误的情况都不断减少，且单台计算机单位时间内发起的域名请求数也不会太大，而当发生异常，如发生攻击时单个 IP 会有大量没有成功解析的域名。

[0007] 2) 网络流量特征：

[0008] 正常情况下，DNS 服务器的查询流量在同一时间段内比较稳定，浮动在特定范围之内。而发生异常，如发生攻击时查询流量会迅速增大，例如，在某一时间段内存在大量目标地址为 DNS 服务器、端口为 53 的域名请求，从而引起 DNS 服务器的查询流量反常和突变。

[0009] 3) 大量请求的域名请求存在规律特征：

[0010] 正常情况下，域名请求具有相对稳定性和随机性的规律特征。稳定性是指存在几个访问比较集中的域名，如大型知名网站的域名；随机性是指访问数量较少，如用户通过搜索引擎链接到的域名地址，这些地址多是用户用来查询信息偶然访问到的域名。

[0011] 发明人在实现本发明实施例过程中发现，现有技术主要针对各区域的总流量特征进行异常检测，而在实际发生异常或攻击中，总流量特征的变化往往不明显，因而漏检的几率较大。

发明内容

[0012] 本发明实施例提供一种 DNS 查询流量异常的检测方法、装置和系统，用以降低域名系统查询流量异常的漏检率。

[0013] 本发明实施例提供了一种 DNS 查询流量异常的检测方法，包括：

[0014] 根据预先划分的地域单位，分别统计各地域的 DNS 查询流量；

[0015] 根据各地域的 DNS 查询流量，分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵；

[0016] 分别计算各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值；

[0017] 在任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值时，则输出瞬时报警信息，所述瞬时报警信息用于提示所述任一时间片 DNS 查询流量异常。

[0018] 本发明实施例还提供了一种 DNS 查询流量异常的检测装置，包括：

[0019] 地域流量统计模块，用于根据预先划分的地域单位，分别统计各地域的 DNS 查询流量；

[0020] 矩阵确定模块，用于根据各地域的 DNS 查询流量，分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵；

[0021] 矩阵相关性计算模块，用于分别计算各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值；

[0022] 瞬时报警模块，用于在任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值时，则输出瞬时报警信息，所述瞬时报警信息用于提示所述任一时间片 DNS 查询流量异常。

[0023] 本发明实施例还提供了一种 DNS 查询流量异常的检测系统，包括上述 DNS 查询流量异常的检测装置。

[0024] 本发明实施例基于地域属性作为异常检测的特征，如果所有地域的总 DNS 查询流量整体变化不大，但个别地域的 DNS 查询流量发生异常，则可在本实施例矩阵相关性变化值上反映出来，因而有利于降低域名系统查询流量异常的漏检率。

附图说明

[0025] 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性的前提下，还可以根据这些附图获得其他的附图。

[0026] 图 1 为本发明第一实施例提供的 DNS 查询流量异常的检测方法流程图；

[0027] 图 2 为本发明第二实施例提供的 DNS 分布式系统的结构示意图；

[0028] 图 3 为本发明第三实施例提供的 DNS 查询流量异常的检测方法流程图；

[0029] 图 4 为本发明第四实施例提供的 DNS 查询流量异常的检测装置结构示意图。

具体实施方式

[0030] 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，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有付出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0031] 图 1 为本发明第一实施例提供的 DNS 查询流量异常的检测方法流程图。如图 1 所

示，本实施例提供的 DNS 查询流量异常的检测方法包括 ：

[0032] 步骤 11 ：根据预先划分的地域单位，分别统计各地域的 DNS 查询流量。

[0033] 各地域的 DNS 查询流量的统计方式不受限制，例如，可以一定时间粒度，如以分钟为单位，分别统计各地域的 DNS 查询流量 ；可有某一装置集中统计 ；或者，可采用分布式系统结构，由 DNS 分节点统计 DNS 分节点各自对应的地域的 DNS 查询流量，再由各 DNS 分节点将统计得到的 DNS 查询流量汇总到某一 DNS 中心节点，由该 DNS 中心节点进行后续的检测。

[0034] 步骤 12 ：根据各地域的 DNS 查询流量，分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵。

[0035] 本发明实施例不需要限定时间片的数量，不需要限定时间片的长度，也不需要限定时间片的起始点 ；在实际应用中，本领域技术人员依据实际需要进行设定即可。可选的，上述多个时间片的长度相同或不同，相邻的两个时间片在时间轴上相连或相互重叠。

[0036] 例如 ：从某一检测时刻开始，取 5 个时间片 ；其中，前三个时间片的长度为 1 分钟，后两个时间片长度为 3 分钟 ；第一个时间片从当前时间点至前 1 分钟、第二个时间片从前 1 分钟至前 2 分钟、第三个时间片从前 2 分钟至前 3 分钟、第四个时间片从前 2 分钟到前 5 分钟，第五个时间片从前 4 分钟到前 6 分钟。

[0037] 本步骤分别计算各时间片对应的协方差矩阵。每个时间片对应的协方差矩阵以基于地域的属性信息为元素，如以在该时间片内划分的各时间间隔内观察到的各地域的 DNS 查询流量变化方差为元素 ；各时间片对应的协方差矩阵的维数相同。计算各时间片对应的各协方差矩阵的平均协方差矩阵。该平均协方差矩阵可经长期学习预先得到，可近似作为各地域 DNS 查询流量处于正常状态时的协方差矩阵。

[0038] 步骤 13 ：分别计算各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值。

[0039] 将任一时间片对应的协方差矩阵与平均协方差矩阵进行比较，计算二者的相关性变化，本发明实施例将该相关性变化称为矩阵相关性变化值，其值可等于该协方差矩阵与平均协方差矩阵的距离。

[0040] 步骤 14 ：在任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值时，则输出瞬时报警信息，所述瞬时报警信息用于提示所述任一时间片 DNS 查询流量异常。

[0041] 任一时间片对应的协方差矩阵与平均协方差矩阵的矩阵相关性变化值，反映了该时间片内各地域的 DNS 查询流量变化。因此，以矩阵相关性变化值为特征检测 DNS 查询流量是否异常，相当于以各地域的 DNS 查询流量等地域属性为特征，检测 DNS 查询流量是否异常。

[0042] 第一预设阈值可根据长期学习得到的经验值预先确定。当矩阵相关性变化值大于第一预设阈值时，表示该时间片对应的协方差矩阵相对于平均协方差矩阵的相关性变化大，该时间片内各地域的 DNS 查询流量统计特性，相对于平均协方差矩阵反映出的各地域 DNS 查询流量处于正常状态下的统计特性，存在较大偏差。该情形下，可输出瞬时报警信息，用于提示该时间片内各地域的 DNS 查询流量异常。

[0043] 本实施例基于地域属性作为异常检测的特征，如果所有地域的总 DNS 查询流量整体变化不大，但个别地域的 DNS 查询流量发生异常，则可在本实施例矩阵相关性变化值上

反映出来，因而有利于降低域名系统查询流量异常的漏检率。

[0044] 图 2 为本发明第二实施例提供的 DNS 分布式系统的结构示意图。如图 2 所示的 DNS 分布式系统中包括：一个 DNS 中心节点和多个 DNS 分节点，每个 DNS 分节点分别与 DNS 中心节点以可通信的方式连接，如通过因特网 (Internet) 通信连接。各 DNS 分节点将各自统计到的各地域的 DNS 查询流量汇总到 DNS 中心节点上，由 DNS 中心节点检测 DNS 查询流量是否发生异常。

[0045] 图 3 为本发明第三实施例提供的 DNS 查询流量异常的检测方法流程图。本实施例以图 2 所示的分布式系统为例，说明本发明 DNS 查询流量异常的检测方法。如图 2 和图 3 所示，本实施例检测方法包括：

[0046] 步骤 31：各 DNS 分节点以一定的时间粒度分别统计各自地域的 DNS 查询流量，并将统计到的 DNS 查询流量上报给 DNS 中心节点。

[0047] DNS 分节点在接收到 DNS 查询数据包时，获取 DNS 查询数据包中的源 IP 地址，利用 DNS 分节点本地预先建立的地理信息数据库，将源 IP 地址映射到一定的地域单位，如将源 IP 地址映射到某省级行政单位；以一定的时间粒度，如以 1 分钟为周期，统计各地域的 DNS 查询流量。当一个时间粒度终结时，将该时间粒度统计的各地域的 DNS 查询流量上传到 DNS 中心节点。例如：以省为行政单位，将某国划分为 34 个省级行政单位，每个时间粒度需上传 34 个地域属性，每个地域属性为 1 个省级行政单位在该时间粒度内的 DNS 查询流量。

[0048] 本领域技术人员可以理解，虽然本实施例是以分布式系统为例进行说明，但在只应用单个 DNS 节点的系统中也可实现；区别在于，在应用单个 DNS 节点的系统中，由该单个 DNS 节点，如本步骤由 DNS 中心节点根据预先划分的地域单位，分别统计各地域的 DNS 查询流量。

[0049] 步骤 32：DNS 中心节点分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵。

[0050] 在实际应用过程中，可基于地域属性，即根据各地域的 DNS 查询流量构造协方差矩阵。以下对基于地域属性构造协方差矩阵的具体算法实现举例说明，但显然本领域技术人员还可基于地域属性采用其他方法构造协方差矩阵，本实施例以下示例不应理解为对本发明实现方式的具体限制。

[0051] 假设：预先将某国划分为 p 个地域单位，每个地域在某一时间片内的 DNS 查询流量构成一地域属性，分别表示为：$f_1, \cdots, f_p$。某一时间片内的各地域属性构成一个随机向量 $X = (f_1, \cdots, f_p)'$。在同一时间片内观察多次，如观察 n 次，令 $x_1, \cdots, x_n$ 为 n 个观察向量，$\boldsymbol{x_j} = \left(f_1^j, ..., f_p^j\right)'$ 为 1 个时间片 1 内的第 j 个观察向量。$f_i^{1,j}$ 表示在 1 个时间片 1 内的第 j 个观察的第 i 个地域的 DNS 查询流量，其中，1 表示时间片的长度；i 表示地域，且 $1 \leqslant i \leqslant p$；j 表示观察向量的序号，且 $1 \leqslant j \leqslant n$。定义一个新的变量 y_1：

[0052]

$$y_l = \begin{pmatrix} f_1^{l,1} & \cdots & f_p^{l,1} \\ f_1^{l,2} & \cdots & f_p^{l,2} \\ \vdots & \ddots & \vdots \\ f_1^{l,n} & \cdots & f_p^{l,n} \end{pmatrix}$$

[0053] 本实施例不妨定义协方差矩阵 M_{y1} 来描述变量 y_1：

[0054]

$$M_{y_l}=\begin{pmatrix}\sigma_{f_1^l f_1^l} & \sigma_{f_1^l f_2^l} & \cdots & \sigma_{f_1^l f_p^l}\\ \sigma_{f_2^l f_1^l} & \sigma_{f_2^l f_2^l} & \cdots & \sigma_{f_2^l f_p^l}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{f_p^l f_1^l} & \sigma_{f_p^l f_2^l} & & \sigma_{f_p^l f_p^l}\end{pmatrix}$$

[0055] M_{y1} 是一个 p 行 p 列的矩阵，其中，每一元素对应两两地域 DNS 查询流量的协方差，如元素 $\sigma_{f_i^l f_j^l}$ 表示第 l 个时间片内第 i 个地域的 DNS 查询流量与第 j 个地域的 DNS 查询流量的协方差，且

[0056] $$\sigma_{f_i^l f_j^l}=\mathrm{cov}\left(f_i^l,f_j^l\right)=E\left[\left(f_i^l-E\left(f_i^l\right)\right)\left(f_j^l-E\left(f_j^l\right)\right)\right]$$

[0057] 其中，$E(f_i^1)$ 为第 l 个时间片内第 i 个地域的 DNS 查询流量均值。

[0058] 如果在第 l 个时间片内对第 i 个地域的 DNS 查询流量共观察 n 次，则：

[0059] $$E\left(f_i^l\right)=\sum_{k=1}^{n}f_i^{l,k}/n$$ 。

[0060] 依据上述定义可得到多个时间片的协方差矩阵，计算多个协方差矩阵的平均协方差矩阵，平均协方差矩阵表示为 $E(M_{y1})$。

[0061] 如果 L 个时间片中，各个时间片的长度均为 1，在时间轴上相连且不重叠，则平均协方差矩阵 $E(M_{y1})$ 可采用下式确定：

[0062] $$E\left(M_{y_l}\right)=\left(\sum_{l=1}^{L}M_{y_l}\right)/L$$

[0063] 如果 L 个时间片中，部分时间片的长度相同，而部分时间片的长度不同，则可对各个时间片的协方差矩阵进行归一化和拓展处理，得到最小时间片长度对应的协方差矩阵的平均协方差矩阵。

[0064] 步骤 33：DNS 中心节点分别计算各时间片对应的协方差矩阵，与平均协方差矩阵的矩阵相关性变化值。

[0065] 任一时间片对应的协方差矩阵，与平均协方差矩阵之间的矩阵相关性变化值可采用二个矩阵之间的距离 z_1 表示，z_1 为：

[0066] $$z_l=\left\|M_{y_l}-E\left(M_{y_l}\right)\right\|=\sqrt{\sum_{1\le i,j\le p}\left(a_{i,j}-b_{i,j}\right)^2},\forall a_{i,j}\in M_{y_l},\forall b_{i,j}\in E\left(M_{y_l}\right),$$ $1\leqslant i,j\leqslant p$。

[0067] 步骤 34：DNS 中心节点判断是否存在矩阵相关性变化值大于第一预设阈值的协方差矩阵，如果存在，执行步骤 35；否则，执行步骤 37。

[0068] 步骤 35：DNS 中心节点根据所述任一时间片对应的协方差矩阵以及所述平均协方差矩阵分别包括的各元素，分别计算所述任一时间片内各地域的属性相关性变化值，任一元素为任两个地域的 DNS 查询流量的协方差；对所述任一时间片内各地域的属性相关性变化值进行聚类分析，得到异常地域属性，所述异常地域属性对应所述任一时间片 DNS 查询流量异常所在的地域。

[0069] 某一时间片对应的协方差矩阵与平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值，则说明该时间片内统计得到的各地域的 DNS 查询流量异常，因而可得到发生 DNS 查询流量异常的时间信息。为了提高异常检测的准确性，可选的，还可对相应时间片内各地域的属性相关性变化值进行聚类分析，以确定该时间片内 DNS 查询流量异常的地域信息，即确定该时间片内，哪个地域的 DNS 查询流量发生异常。

[0070] 发明人在实践本发明实施例过程中发现，当某一时间片内 DNS 查询流量异常时，通常不是所有地域属性共同导致的，而是部分地域属性导致的。

[0071] 不妨令任一时间片对应的协方差矩阵，与平均协方差矩阵之间的相对均值的协方差变化矩阵为 ΔM_{y1}，且：

[0072]

$$\Delta M_{y_l} = M_{y_l} - E\left(M_{y_l}\right) = \begin{pmatrix} \Delta\sigma_{f_1^l f_1^l} & \Delta\sigma_{f_1^l f_2^l} & \cdots & \Delta\sigma_{f_1^l f_p^l} \\ \Delta\sigma_{f_2^l f_1^l} & \Delta\sigma_{f_2^l f_2^l} & \cdots & \Delta\sigma_{f_2^l f_p^l} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta\sigma_{f_p^l f_1^l} & \Delta\sigma_{f_p^l f_2^l} & & \Delta\sigma_{f_p^l f_p^l} \end{pmatrix}$$

[0073] 其中，$\Delta\sigma_{f_j^l f_i^l}$ 为协方差变化矩阵 ΔM_{y1} 第 i 行 j 列的元素，表示第 l 个时间片内，第 i 个地域的 DNS 查询流量与第 j 个地域的 DNS 查询流量的协方差，相对于均值协方差的变化。

[0074] 为了有效定位发生异常的地域，可计算某一时间片内，任一地域的 DNS 查询流量相对于该地域的 DNS 正常查询流量的相关性变化，该相关性变化即为本发明实施例所述的属性相关性变化值，以下表示为 C_i^1，且：$C_i^l = \sqrt{\sum_{j=1}^{n}\left(\Delta\sigma_{f_j^l f_i^l}\right)^2}$ 其中，$1 \leqslant i, j \leqslant n$；

[0075] 其中，$\Delta\sigma_{f_j^l f_i^l}$ 为协方差变化矩阵 ΔM_{y1} 第 i 行 j 列的元素，表示第 l 个时间片内，第 i 个地域的 DNS 查询流量与第 j 个地域的 DNS 查询流量的协方差，相对于均值协方差的变化；C_i^1 表示第 l 个时间片对应的协方差矩阵中与第 i 个地域相关的协方差变化，即第 l 个时间片内，第 i 个地域的属性相关性变化值。

[0076] 可采用聚类分析的方法区分属性相关性中的正常属性和异常属性，如把各地域的属性相关度变化值划分为正常和异常两类，将属性相关度变化值大于某一预设阈值的地域属性作为异常地域属性，剩余地域属性作为正常地域属性，具体可采用 K-Means 算法进行聚类分析，得到异常地域属性，从而确定在该时间片内发生 DNS 查询异常的地域。

[0077] 由于采用本实施例所述的方法不仅可确定发生 DNS 查询流量异常的时间信息，还可确定发生 DNS 查询流量异常的地域，因此，提高了 DNS 查询流量检测的准确性。

[0078] 步骤 36：DNS 中心节点对矩阵相关性变化值大于第一预设阈值的协方差矩阵进行平滑处理，以使该协方差矩阵与平均协方差矩阵的矩阵相关性变化值，小于或等于第一预设阈值；执行步骤 39。

[0079] 平滑处理的具体实现方式不受限制，例如：可将该协方差矩阵各地域属性元素进行等比例缩小，以使该协方差矩阵与平均协方差矩阵的矩阵相关性变化值，小于或等于第

一预设阈值。

[0080] 步骤 37 :检测在一预设时间段包括的各时间片对应的协方差矩阵与平均协方差矩阵的矩阵相关性变化值,大于第二预设阈值且小于或等于第一预设阈值的矩阵数量。

[0081] 本步骤中的预设时间段可由多个时间片组成。在该时间段内,虽然各个时间片内的矩阵相关性变化值没有超过第一预设阈值,但有可能临近第一预设阈值,处于准异常状态。为了有效提供预警信息,可根据长期学习得到的经验值预先确定第二预设阈值,且第二预设阈值小于第一预设阈值。

[0082] 当某一时间片的矩阵相关性变化值大于第二预设阈值且小于或等于第一预设阈值时,表示该时间片内虽然未输出瞬时告警,但发生异常的几率较高。该情形下,可统计该时间段内矩阵相关性变化值大于第二预设阈值且小于或等于第一预设阈值的矩阵数量,并进行时间段 DNS 查询流量检测,确定是否需要进行时间段告警,以进一步提高异常检测的漏检率。

[0083] 步骤 38 :在所述矩阵数量大于第三预设阈值时,则输出时间段报警信息,所述时间段报警信息用于提示所述时间段 DNS 查询流量异常 ;执行步骤 39。

[0084] 时间段异常报警是另一种异常报警方式。本实施例可根据长期学习得到的经验值预先确定第三预设阈值。在采用步骤 37 所述的方法确定的矩阵数量大于第三预设阈值时,说明在该时间段内处于准异常状态的时间片密度较大,则可输出时间段报警信息,用于该时间段 DNS 查询流量异常。如果采用步骤 37 所述的方法确定的矩阵数量小于或等于第三预设阈值,则说明在该时间段内处于准异常状态的时间片密度较小,该情形下不需要进行时间段报警。

[0085] 可选的,在时间段报警过程中,可对步骤 37 确定的时间片分别对应的、矩阵相关性变化值大于第二预设阈值且小于或等于第一预设阈值的协方差矩阵的地域属性元素进行聚类分析,确定 DNS 查询流量准异常的地域,具体实现方式与步骤 35 相似,这样,可得到准异常的时间信息和地域信息。

[0086] 步骤 39 :待下一检测周期到来时,执行步骤 32。

[0087] 在实际应用中攻击者很难人为篡改地域特征,本实施例在时间片瞬时异常报警的基础上,还可以基于聚类分析的异常属性,有效定位该时间片内异常的发生地域,提高了异常检测的准确性。进一步的,本实施例可将时间片瞬时异常报警和时间段异常报警相结合,一方面及时提示 DNS 查询流量发生异常的时间信息和地域信息,另一方面可对某一时间段内处于准异常状态的时间片密度较大时,及时进行时间段报警以提示该时间段 DNS 查询流量异常,因此本实施例明显降低了异常检测的漏检率,有效提高了异常检测的准确性。

[0088] 可选的,上述对时间段异常检测的过程中,还可在另一实施例中统计矩阵相关性变化值大于第二预设阈值的矩阵数量,将该矩阵数量与第三预设阈值进行比较,从而确定该时间段是否发生 DNS 查询流量异常 ;其实现方式与本实施例相似,在此不再赘述。

[0089] 对矩阵相关性变化值大于第一预设阈值的矩阵进行平滑处理的好处在于 :如果没有对矩阵相关性变化值大于第一预设阈值的协方差矩阵进行平滑处理,则可能造成虚警,影响后续检测的准确性。例如 :在进行异常检测时,如果存在远远大于第一预设阈值的矩阵相关性变化值,则在进行时间段异常检测时,可能仅由这些变化值造成其临近的时间段异常,而不是由在这些时间段内密集出现的大于第二预设阈值的变化值造成时间段异常,从

而在实际应用中造成虚警，影响检测的准确度。因此本实施例通过本步骤对矩阵相关性变化值大于第一预设阈值的协方差矩阵进行平滑处理，使得这些协方差矩阵的矩阵相关性变化值不大于第一预设阈值，既在一定程度上保留了这些变化值信息，又使其不至于对时间段异常检测造成过大的影响。

[0090] DNS 查询流量检测的目标主要是及时发现异常，以利于告警和采取相应的处置措施。这种 DNS 流量异常的来源通常可以分为两个方面 ：一是 DNS 系统或网络本身的异常，如配置错误，攻击等等 ；另一方面涉及到网络舆情，即网络系统本身是正常的，而网络用户群体性的行为特征发生异常。

[0091] 本发明实施例由 DNS 中心节点收集由各个 DNS 分节点统计的各地域属性信息，或者，由 DNS 中心节点统计各地域属性信息，然后基于地域属性作为异常检测的特征，这种特征不仅对于 DNS 系统或网络本身的异常很敏感，而且可以有效反映网络舆情的异常，因此可以大大减少上述两种异常的漏检率，尤其后者是现有方法所没有解决的。

[0092] 其中，网络舆情定义为通过互联网传播的公众对现实生活中某些热点、焦点问题所持的有较强影响力、倾向性的言论和观点，主要通过 BBS 论坛、博客、新闻跟帖、转帖等实现并加以强化。由于互联网具有虚拟性、隐蔽性、发散性、渗透性和随意性等特点，越来越多的网民愿意通过这种渠道来表达观点、传播思想。由于网络媒体本身的开放、快捷、虚拟、易复制等特征，网络舆情也有别于传统舆情。同时，互联网涌现性特性使得大量网络舆情的爆发，某种程度上成为舆情安全的重要因素。

[0093] 当前，网络舆情分析的研究对象集中在网络媒体内容、具体网络应用的用户互动关系以及网络突发流量方面。域名访问是信息内容传播过程中必不可少的环节，域名系统能够准确记录用户访问信息通信服务的地址特性、时间特性，而且能够反映出用户的行为特性，具备构建相应舆情态势分析和预警机制的技术基础。利用域名访问行为中蕴藏的丰富信息，探索和发现信息网络中的行为模式和异常情况，有利于准确的掌握信息网络的宏观舆情态势，可视为现有基于网络媒体内容相关方法的一个重要的补充。利用域名访问日志提供的用户信息，开展针对特定地域和人群的舆情态势分析技术研究，弥补了现有方法对网络用户主体信息缺乏了解的缺点，加强了网络舆情分析的主体针对性。

[0094] 本发明上述实施例利用基于地域信息相关性作为异常检测的特征，不仅对于 DNS 系统或网络本身的异常很敏感，而且可以有效反映网络舆情的异常，因此可以大大减少上述两种异常的漏检率。

[0095] 图 4 为本发明第四实施例提供的 DNS 查询流量异常的检测装置结构示意图。如图 4 所示，本实施例 DNS 查询流量异常的检测装置包括 ：地域流量统计模块 41、矩阵确定模块 42、矩阵相关性计算模块 43 和瞬时报警模块 44。

[0096] 地域流量统计模块 41 用于根据预先划分的地域单位，分别统计各地域的 DNS 查询流量。

[0097] 矩阵确定模块 42 用于根据各地域的 DNS 查询流量，分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵。

[0098] 矩阵相关性计算模块 43 用于分别计算各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值。

[0099] 瞬时报警模块 44 用于在任一时间片对应的协方差矩阵与所述平均协方差矩阵的

矩阵相关性变化值，大于第一预设阈值时，则输出瞬时报警信息，所述瞬时报警信息用于提示所述任一时间片 DNS 查询流量异常。

[0100] 在上述技术方案的基础上，可选的，检测装置还可包括：属性相关性计算模块 45 和异常地域属性确定模块 46。

[0101] 属性相关性计算模块 45 用于根据所述任一时间片对应的协方差矩阵以及所述平均协方差矩阵分别包括的各元素，分别计算所述任一时间片内各地域的属性相关性变化值，任一元素为任两个地域的 DNS 查询流量的协方差。

[0102] 异常地域属性确定模块 46 用于对所述任一时间片内各地域的属性相关性变化值进行聚类分析，得到异常地域属性，所述异常地域属性对应所述任一时间片 DNS 查询流量异常所在的地域。

[0103] 可选的，本实施例检测装置还可包括：平滑处理模块 47。

[0104] 平滑处理模块 47 用于对所述任一时间片对应的协方差矩阵进行平滑处理，以使所述任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，小于或等于第一预设阈值。

[0105] 可选的，本实施例检测装置还可包括：时间片确定模块 48。

[0106] 时间片确定模块 48 用于确定所述多个时间片，所述多个时间片的长度相同或不同，且相邻的两个时间片在时间轴上相连或相互重叠。

[0107] 在上述技术方案的基础上，可选的，本实施例检测装置还可包括：矩阵数量确定模块 49 和时间段报警模块 410。

[0108] 矩阵数量确定模块 49 用于检测在一预设时间段包括的各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第二预设阈值的矩阵数量，或者大于第二预设阈值且小于或等于第一预设阈值的矩阵数量。

[0109] 时间段报警模块 410 用于在所述矩阵数量大于第三预设阈值时，则输出时间段报警信息，所述时间段报警信息用于提示所述时间段 DNS 查询流量异常。

[0110] 本实施例提供的 DNS 查询流量异常的检测装置，基于地域属性作为异常检测的特征，如果所有地域的总 DNS 查询流量整体变化不大，但个别地域的 DNS 查询流量发生异常，则可在本实施例矩阵相关性变化量上反映出来，因而有利于降低域名系统查询流量异常的漏检率。本实施例可确定 DNS 查询流量异常的时间信息和地域信息，有利于提高异常检测的精准性。进一步的，本实施例可将时间片瞬时异常报警和时间段异常报警相结合，可明显降低了异常检测的漏检率，有效提高了异常检测的准确性。本实施例检测装置的具体表现实体不受限制，如可为 DNS 服务器、分布式系统中的 DNS 中心节点等，其实现 DNS 查询流量异常的检测机理，可参见图 1 和图 3 对应实施例的记载，在此不再赘述。

[0111] 本发明还提供了一种包括如图 4 所示的 DNS 查询流量异常的检测系统。可选的，该检测系统可具有分布式结构，如可包括一检测装置和多个 DNS 分节点，该检测装置相当于 DNS 中心节点，系统结构如图 2 所示。每个 DNS 分节点与检测装置以可通信方式连接，用于 DNS 分节点，用于统计各地域的 DNS 查询流量，并将统计到的 DNS 查询流量上报给所述检测装置，由检测装置进行 DNS 查询流量异常的检测。本发明提供的检测系统实现 DNS 查询流量异常的检测机理，可参见图 1 和图 3 对应实施例的记载，在此不再赘述。

[0112] 本领域普通技术人员可以理解：附图只是一个实施例的示意图，附图中的模块或

流程并不一定是实施本发明所必须的。

[0113] 本领域普通技术人员可以理解：实施例中的装置中的模块可以按照实施例描述分布于实施例的装置中，也可以进行相应变化位于不同于本实施例的一个或多个装置中。上述实施例的模块可以合并为一个模块，也可以进一步拆分成多个子模块。

[0114] 上述本发明实施例序号仅仅为了描述，不代表实施例的优劣。

[0115] 本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于一计算机可读取存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0116] 最后应说明的是：以上实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前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：其依然可以对前述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实施例技术方案的精神和范围。

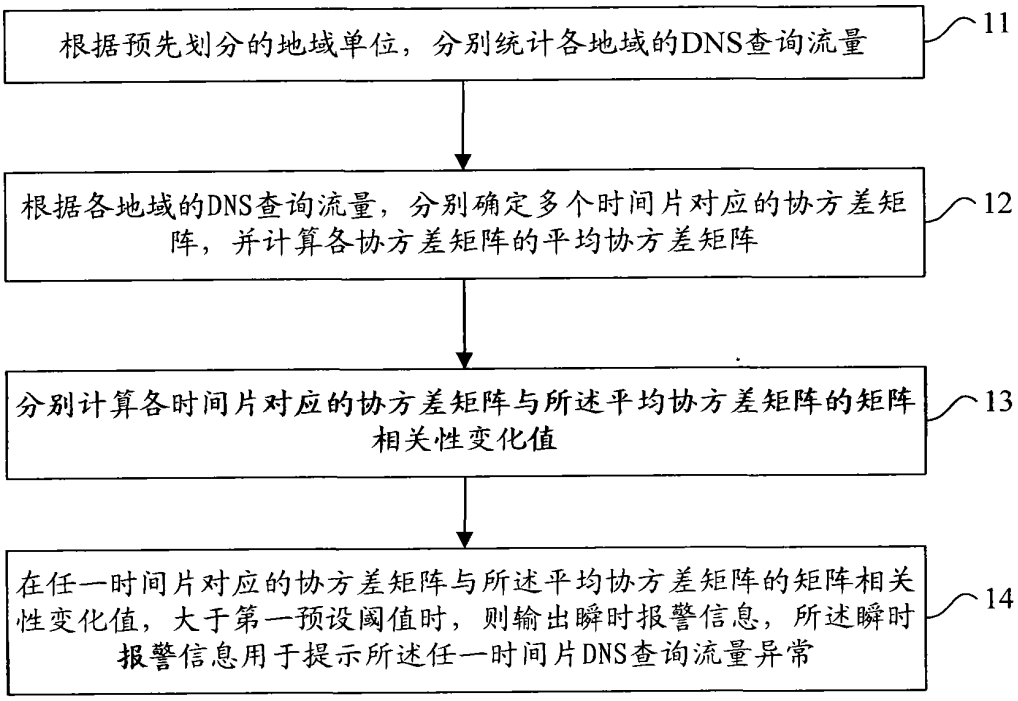

根据预先划分的地域单位，分别统计各地域的DNS查询流量
11
根据各地域的DNS查询流量，分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵
12
分别计算各时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值
13
在任一时间片对应的协方差矩阵与所述平均协方差矩阵的矩阵相关性变化值，大于第一预设阈值时，则输出瞬时报警信息，所述瞬时报警信息用于提示所述任一时间片DNS查询流量异常
14

图 1

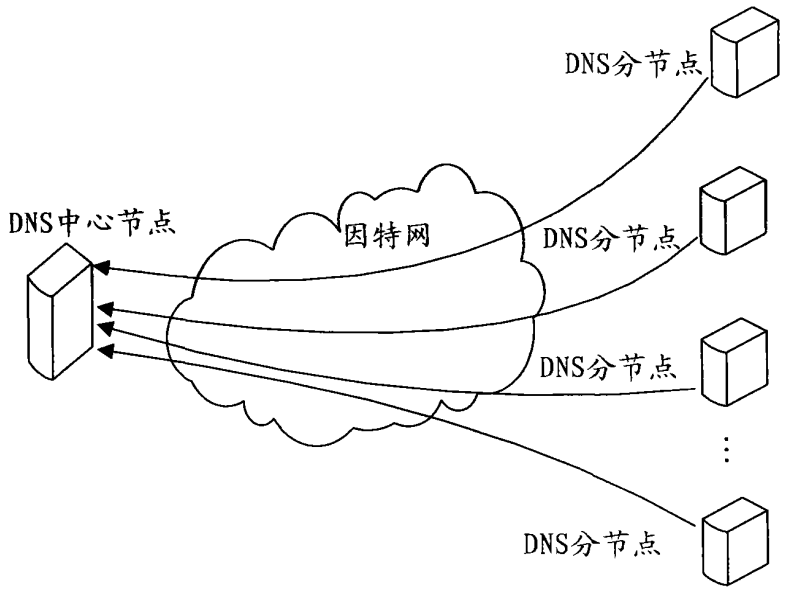

DNS分节点
DNS中心节点
因特网
DNS分节点
DNS分节点
DNS分节点

图 2

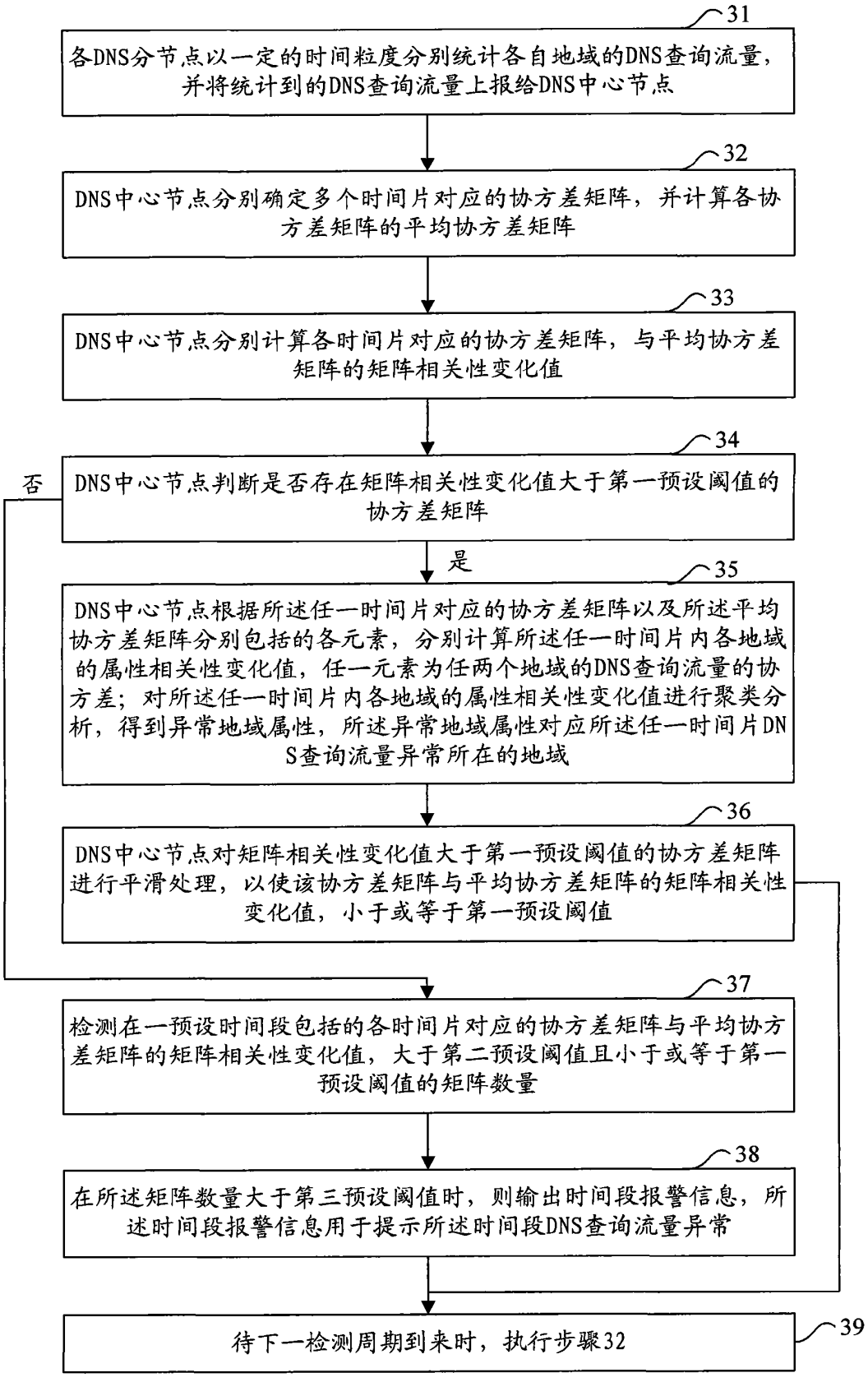

31
各DNS分节点以一定的时间粒度分别统计各自地域的DNS查询流量，并将统计到的DNS查询流量上报给DNS中心节点
32
DNS中心节点分别确定多个时间片对应的协方差矩阵，并计算各协方差矩阵的平均协方差矩阵
33
DNS中心节点分别计算各时间片对应的协方差矩阵，与平均协方差矩阵的矩阵相关性变化值
34
否
DNS中心节点判断是否存在矩阵相关性变化值大于第一预设阈值的协方差矩阵
是
35
DNS中心节点根据所述任一时间片对应的协方差矩阵以及所述平均协方差矩阵分别包括的各元素，分别计算所述任一时间片内各地域的属性相关性变化值，任一元素为任两个地域的DNS查询流量的协方差；对所述任一时间片内各地域的属性相关性变化值进行聚类分析，得到异常地域属性，所述异常地域属性对应所述任一时间片DNS查询流量异常所在的地域
36
DNS中心节点对矩阵相关性变化值大于第一预设阈值的协方差矩阵进行平滑处理，以使该协方差矩阵与平均协方差矩阵的矩阵相关性变化值，小于或等于第一预设阈值
37
检测在一预设时间段包括的各时间片对应的协方差矩阵与平均协方差矩阵的矩阵相关性变化值，大于第二预设阈值且小于或等于第一预设阈值的矩阵数量
38
在所述矩阵数量大于第三预设阈值时，则输出时间段报警信息，所述时间段报警信息用于提示所述时间段DNS查询流量异常
39
待下一检测周期到来时，执行步骤32

图 3

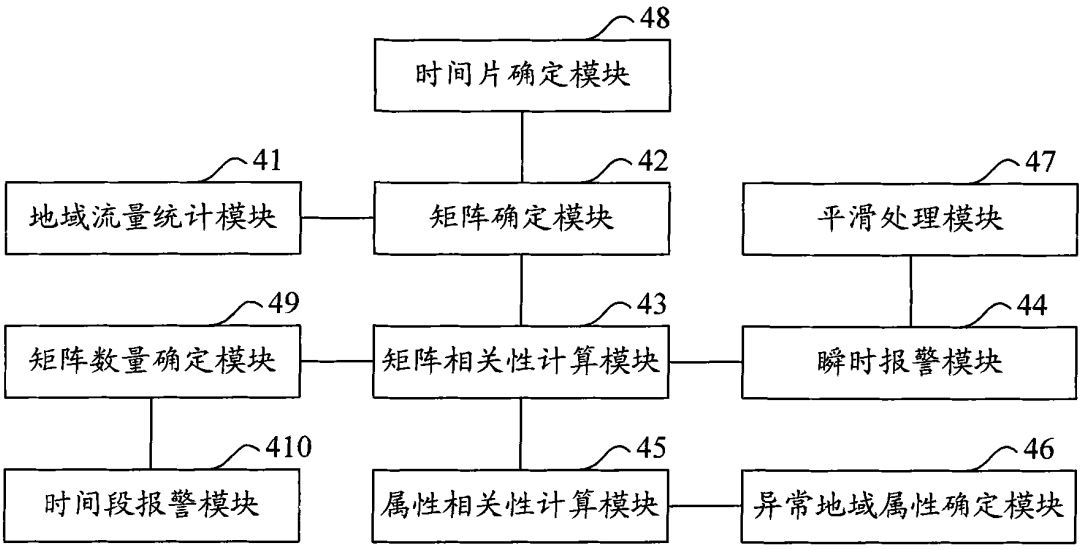

48
时间片确定模块
41
地域流量统计模块
42
矩阵确定模块
47
平滑处理模块
49
矩阵数量确定模块
43
矩阵相关性计算模块
44
瞬时报警模块
410
时间段报警模块
45
属性相关性计算模块
46
异常地域属性确定模块

图 4