



(12) 发明专利

(10) 授权公告号 CN 102130800 B

(45) 授权公告日 2013. 08. 28

(21) 申请号 201110083016. X

US 2010071061 A1, 2010. 03. 18,

(22) 申请日 2011. 04. 01

CN 101150581 A, 2008. 03. 26,

(73) 专利权人 苏州赛特斯网络科技有限公司

CN 101826996 A, 2010. 09. 08,

地址 215300 江苏省苏州市昆山开发区前进  
东路科技广场大楼 15 楼

审查员 高旭

(72) 发明人 逯利军 钱培专

(74) 专利代理机构 上海智信专利代理有限公司

31002

代理人 王洁 郑暄

(51) Int. Cl.

H04L 12/26 (2006. 01)

H04L 12/70 (2013. 01)

(56) 对比文件

TW 200522627 A, 2005. 07. 01,

US 2006047807 A1, 2006. 03. 02,

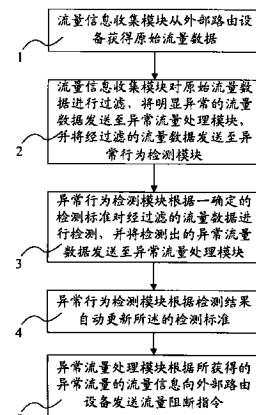
权利要求书3页 说明书10页 附图4页

(54) 发明名称

基于数据流行为分析的网络访问异常检测装  
置及方法

(57) 摘要

本发明涉及一种基于数据流行为分析的网络访问异常检测装置，其包括流量信息收集模块、异常行为检测模块和异常流量处理模块，流量信息收集模块的分别连接异常行为检测模块和异常流量处理模块，异常行为检测模块连接异常流量处理模块。本发明还涉及一种利用该装置的方法，该方法先过滤掉明显异常的流量数据，再利用一网络行为模型对经过滤的流量数据进行检测，并自动更新网络行为模型；最后根据检测结果阻断流量。利用本发明的装置和方法能建立正常网络行为模型，将该模型跟实时数据进行比较，以检测实时流量是否异常；并动态修正网络行为模型，分析异常流量来源，对异常流量进行阻断，从而快速有效地识别异常流量，提高检测的准确性。



1. 一种利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，所述的装置包括流量信息收集模块、异常行为检测模块和异常流量处理模块，所述的流量信息收集模块的输入端连接该检测装置外部的路由设备，所述的流量信息收集模块的输出端分别连接所述的异常行为检测模块和异常流量处理模块的输入端，所述的异常行为检测模块的输出端连接所述的异常流量处理模块的输入端，所述的异常流量处理模块的输出端连接该检测装置外部的路由设备，其特征在于，所述的方法包括以下步骤：

- (1) 所述的流量信息收集模块从外部路由设备获得原始流量数据；
- (2) 所述的流量信息收集模块对原始流量数据进行过滤，将明显异常的流量数据发送至所述的异常流量处理模块，并将经过滤的流量数据发送至所述的异常行为检测模块，具体包括以下步骤：
  - (21) 所述的流量信息收集模块对原始流量数据进行解析，获得流量数据信息；
  - (22) 所述的流量信息收集模块将明显异常的原始流量数据的流量数据信息存入一异常流量数据库；
  - (23) 所述的流量信息收集模块将其余的流量数据存入一待检测流量数据库；
- (3) 所述的异常行为检测模块根据一确定的检测标准对所述的经过滤的流量数据进行检测，并将检测出的异常流量数据发送至所述的异常流量处理模块，具体包括以下步骤：
  - (31) 所述的异常行为检测模块读取所述的待检测流量数据库中的流量数据；
  - (32) 所述的异常行为检测模块基于前一周期内一确定时间段的数据流量值生成本周期内对应时间段的流量数据的预测值；
  - (33) 所述的异常行为检测模块将本周期该时间段的预测值与本周期内该时间段的流量数据实际值比较，判断两者差距是否大于预设的阈值，若大于，则确定该流量数据为异常流量数据，并进入步骤(34)，若不大于，则进入步骤(4)；
  - (34) 所述的异常行为检测模块将该异常流量数据的信息存入所述的异常流量数据库，并进入步骤(5)；
  - (4) 所述的异常行为检测模块根据检测结果自动更新所述的检测标准；
  - (5) 所述的异常流量处理模块根据所获得的异常流量的流量信息向外部路由设备发送流量阻断指令。

2. 根据权利要求 1 所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，所述的原始流量数据为 netflow v5 格式数据或 sFlow 格式数据。

3. 根据权利要求 1 所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，所述的流量数据信息包括源 IP 地址、源端口、目的 IP 地址、目的端口、协议类型、端口号、字节数、数据包数及数据流产生时间。

4. 根据权利要求 1 所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，所述的步骤(32)具体是指：

设第  $m$  个周期的第  $i$  个时间段内的实际数据流量值为  $T_{(i, m)}$ ，则根据以下公式得到在第  $m+1$  个周期内对应的第  $i$  个时间段内的预测数据流量值  $P_{(i, m+1)}$ ：

$$P_{(i, m+1)} = a_{(i, m)} + b_{(i, m)},$$

其中：

$$a_{(i, m)} = 2S'_{(i, m)} - S''_{(i, m)},$$

$$b_{(i, m)} = \frac{\alpha}{1-\alpha} (S'_{(i, m)} - S''_{(i, m)}),$$

$S'_{(i, m)}$  与  $S''_{(i, m)}$  分别为在第  $m$  个周期中第  $i$  个时间段的预测参数：

$$S'_{(i, m)} = \alpha T_{(i, m)} + (1-\alpha) S'_{(i, m-1)},$$

$$S''_{(i, m)} = \alpha S'_{(i, m)} + (1-\alpha) S''_{(i, m-1)},$$

$\alpha$  为预设的预测敏感系数。

5. 根据权利要求 4 所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，在第 1 个周期中第  $i$  个时间段的预测参数  $S'_{(i, 0)}$  与  $S''_{(i, 0)}$  分别为： $S'_{(i, 0)} = S''_{(i, 0)} = T_{(i, 1)}$ 。

6. 根据权利要求 4 所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，步骤(33)具体包括以下步骤：

(33-1) 所述的异常行为检测模块判断  $|T_{(i, m)} - P_{(i, m)}|$  是否大于预设的阈值；

(33-2) 若大于，则确定第  $m$  个周期的第  $i$  个时间段的流量数据为异常流量数据，并进入步骤(34)；

(33-3) 若不大于，则进入步骤(4)。

7. 根据权利要求 6 所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，步骤(34)具体包括以下步骤：

(34-1) 所述的异常行为检测模块依确定的选取规则采集所述的第  $m$  个周期的第  $i$  个时间段的流量数据的流量数据信息；

(34-2) 所述的异常行为检测模块并将所采集的流量数据信息存入所述的异常流量数据库，并进入步骤(5)。

8. 根据权利要求 7 所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，所述的选取规则具体为以下之一：

(1) 采集该时间段内流量数据字节或数据包较大的数据流的流量数据信息；

(2) 采集非关键端口产生的数据流的流量数据信息；

(3) 综合采集非关键端口产生的数据字节或数据包较大的数据流的流量数据信息。

9. 根据权利要求 4 所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，所述的步骤(4)具体是指：

所述的异常行为检测模块将第  $m+1$  个周期的第  $i$  个时间段的流量数据预测值  $P_{(i, m+1)}$  设置为：

$$P_{(i, m+1)} = T_{(i, m)}.$$

10. 根据权利要求 1 至 9 中任一项所述的利用基于数据流行为分析的网络访问异常检测装置实现基于数据流行为分析的网络访问异常检测方法，其特征在于，所述的步骤(5)具体包括以下步骤：

(51) 所述的异常流量处理模块读取所述的异常流量数据库中的流量数据信息；

(52) 所述的异常流量处理模块统计所述的流量数据信息中的外部 IP 地址、协议类型及端口号；

(53) 所述的异常流量处理模块向外部路由设备发送针对所述的外部 IP 地址的流量阻断指令。

## 基于数据流行为分析的网络访问异常检测装置及方法

### 技术领域

[0001] 本发明涉及网络技术领域,特别涉及网络访问异常检测装置及方法领域,具体是指一种基于数据流行为分析的网络访问异常检测装置及方法。

### 背景技术

[0002] 目前,现有的网络流流量分析技术分为以下几类:

[0003] 1、通过对网络流量的总计,并设置阀值来判断网络流量是否有异常。例如,通过网络设备上的 SNMP 接口,定期获取相关网口的数据流量,例如,单播包总包数,单播包总字节数等等,并通过预设的阀值进行比较,判断是否异常。

[0004] 2、通过对单个或数个连续的数据包进行分析,判断数据包是否属于异常流量。例如,通过对数据包的协议、端口以及大小来进行判断。例如,SQL Slammer 蠕虫是通过 UDP 1434 端口发送大小为 376 字节的数据包进行扫描来加以判断。

[0005] 3、基于对历史流量的分析,统计网络流量的行为并进行预测,将一种数据流量分析的方法应用到各类网络设备中,该方法分析一段时期内历史数据流量,通过计算机算法挖掘出流量的变化规律,同时对下一周期内流量的情况进行预测,一旦实际测量值与预测值产生较大的差异,则认为产生了异常的访问流量。

[0006] 上述各种现有方法的缺点在于:方法 1 只能对总流量的大小进行简单识别,无法区分这些流量中那部分是正常流量,哪部分是异常流量。方法 2 虽然可以识别出具体的异常流量数据包,但是由于是通过预设的数据包特征模式,无法识别出新的变异后的异常流量。方法 3 考虑到了流量的历史访问规律,能检测出严重违背历史经验的流量访问突变,因此并不需要知道所有可能导致流量异常的数据包的特征模式。但是,目前大部分属于方法 3 的数据流行为检测技术,仅仅考虑了对数据流行为的统计,并未考虑到一旦检测到有异常数据流之后,如何进一步的阻止相关的异常流量。同时,由于属于方法 3 的这些技术是针对于历史数据流的变化规律进行统计,但未考虑到攻击方可能会采用一个缓慢的不断增加异常流量的过程,即在这样的情况下,缓慢增加的攻击流量,会导致算法不断的修正对历史流量规律的统计,从而不断增加系统预测流量的大小,从而当异常流量达到显著数量的情况下,系统仍然无法进行检测。所以,现有的方法都存在不同的缺陷,难以应用在大规模数据流量的网络访问异常检测之中。

### 发明内容

[0007] 本发明的目的是克服了上述现有技术中的缺点,提供一种能在大规模数据流量分析的应用环境中快速高效地总结流量行为,识别异常流量,并有效避免无法检测出缓慢异常流量增加的情况,从而提高检测准确性,且应用方式较为简单,应用成本低廉,且适用范围广泛的基于数据流行为分析的网络访问异常检测装置及方法。

[0008] 为了实现上述的目的,本发明的基于数据流行为分析的网络访问异常检测装置具有如下构成:

[0009] 该装置包括流量信息收集模块、异常行为检测模块和异常流量处理模块，所述的流量信息收集模块的输入端连接该检测装置外部的路由设备，所述的流量信息收集模块的输出端分别连接所述的异常行为检测模块和异常流量处理模块的输入端，所述的异常行为检测模块的输出端连接所述的异常流量处理模块的输入端，所述的异常流量处理模块的输出端连接该检测装置外部的路由设备。

[0010] 本发明所提供的利用所述的装置实现基于数据流行为分析的网络访问异常检测方法，其包括以下步骤：

[0011] (1) 所述的流量信息收集模块从外部路由设备获得原始流量数据；

[0012] (2) 所述的流量信息收集模块对原始流量数据进行过滤，将明显异常的流量数据发送至所述的异常流量处理模块，并将经过滤的流量数据发送至所述的异常行为检测模块；

[0013] (3) 所述的异常行为检测模块根据一确定的检测标准对所述的经过滤的流量数据进行检测，并将检测出的异常流量数据发送至所述的异常流量处理模块；

[0014] (4) 所述的异常行为检测模块根据检测结果自动更新所述的检测标准；

[0015] (5) 所述的异常流量处理模块根据所获得的异常流量的流量信息向外部路由设备发送流量阻断指令。

[0016] 该基于数据流行为分析的网络访问异常检测方法中，所述的原始流量数据为 netflow v5 格式数据或 sFlow 格式数据。

[0017] 该基于数据流行为分析的网络访问异常检测方法中，所述的步骤(2)具体包括以下步骤：

[0018] (21) 所述的流量信息收集模块对原始流量数据进行解析，获得流量数据信息；

[0019] (22) 所述的流量信息收集模块将明显异常的原始流量数据的流量数据信息存入一异常流量数据库；

[0020] (23) 所述的流量信息收集模块将其余的流量数据存入一待检测流量数据库。

[0021] 该基于数据流行为分析的网络访问异常检测方法中，所述的流量数据信息包括源 IP 地址、源端口、目的 IP 地址、目的端口、协议类型、端口号、字节数、数据包数及数据流产生时间。

[0022] 该基于数据流行为分析的网络访问异常检测方法中，所述的步骤(3)具体包括以下步骤：

[0023] (31) 所述的异常行为检测模块读取所述的待检测流量数据库中的流量数据；

[0024] (32) 所述的异常行为检测模块基于前一周期内一确定时间段的数据流量值生成本周期内对应时间段的流量数据的预测值；

[0025] (33) 所述的异常行为检测模块将本周期该时间段的预测值与本周期内该时间段的流量数据实际值比较，判断两者差距是否大于预设的阈值，若大于，则确定该流量数据为异常流量数据，并进入步骤(34)，若不大于，则进入步骤(4)；

[0026] (34) 所述的异常行为检测模块将该异常流量数据的信息存入所述的异常流量数据库，并进入步骤(5)。

[0027] 该基于数据流行为分析的网络访问异常检测方法中，所述的步骤(32)具体是指：

[0028] 设第 m 个周期的第 i 个时间段内的实际数据流量值为  $T_{(i, m)}$ ，则根据以下公式得到

在第  $m+1$  个周期内对应的第  $i$  个时间段内的预测数据流量值  $P_{(i, m+1)}$  :

[0029]  $P_{(i, m+1)} = a_{(i, m)} + b_{(i, m)},$

[0030] 其中 :

[0031]  $a_{(i, m)} = 2S'_{(i, m)} - S''_{(i, m)},$

[0032]  $b_{(i, m)} = \frac{\alpha}{1-\alpha} (S'_{(i, m)} - S''_{(i, m)}),$

[0033]  $S'_{(i, m)}$  与  $S''_{(i, m)}$  分别为在第  $m$  个周期中第  $i$  个时间段的预测参数 :

[0034]  $S'_{(i, m)} = \alpha T_{(i, m)} + (1-\alpha) S'_{(i, m-1)},$

[0035]  $S''_{(i, m)} = \alpha S'_{(i, m)} + (1-\alpha) S''_{(i, m-1)},$

[0036]  $\alpha$  为预设的预测敏感系数。

[0037] 该基于数据流行为分析的网络访问异常检测方法中, 在第 1 个周期中第  $i$  个时间段的预测参数  $S'_{(i, 0)}$  与  $S''_{(i, 0)}$  分别为 :  $S'_{(i, 0)} = S''_{(i, 0)} = T_{(i, 1)}$ 。

[0038] 该基于数据流行为分析的网络访问异常检测方法中, 步骤 (33) 具体包括以下步骤 :

[0039] (33-1) 所述的异常行为检测模块判断  $|T_{(i, m)} - P_{(i, m)}|$  是否大于预设的阈值 ;

[0040] (33-2) 若大于, 则确定第  $m$  个周期的第  $i$  个时间段的流量数据为异常流量数据, 并进入步骤 (34) ;

[0041] (33-3) 若不大于, 则进入步骤 (4)。

[0042] 该基于数据流行为分析的网络访问异常检测方法中, 步骤 (34) 具体包括以下步骤 :

[0043] (34-1) 所述的异常行为检测模块依确定的选取规则采集所述的第  $m$  个周期的第  $i$  个时间段的流量数据的流量数据信息 ;

[0044] (34-2) 所述的异常行为检测模块并将所采集的流量数据信息存入所述的异常流量数据库, 并进入步骤 (5)。

[0045] 该基于数据流行为分析的网络访问异常检测方法中, 所述的选取规则具体为以下之一 :

[0046] (1) 采集该时间段内流量数据字节或数据包较大的数据流的流量数据信息 ;

[0047] (2) 采集非关键端口产生的数据流的流量数据信息 ;

[0048] (3) 综合采集非关键端口产生的数据字节或数据包较大的数据流的流量数据信息。

[0049] 该基于数据流行为分析的网络访问异常检测方法中, 所述的步骤 (4) 具体是指 :

[0050] 所述的异常行为检测模块将第  $m+1$  个周期的第  $i$  个时间段的流量数据预测值  $P_{(i, m+1)}$  设置为 :

[0051]  $P_{(i, m+1)} = T_{(i, m)}.$

[0052] 该基于数据流行为分析的网络访问异常检测方法中, 所述的步骤 (5) 具体包括以下步骤 :

[0053] (51) 所述的异常流量处理模块读取所述的异常流量数据库中的流量数据信息 ;

[0054] (52) 所述的异常流量处理模块统计所述的流量数据信息中的外部 IP 地址、协议类型及端口号 ;

[0055] (53) 所述的异常流量处理模块向外部路由设备发送针对所述的外部 IP 地址的流量阻断指令。

[0056] 采用了该发明的基于数据流行为分析的网络访问异常检测装置及方法，其装置包括流量信息收集模块、异常行为检测模块和异常流量处理模块，所述的流量信息收集模块的输入端连接该检测装置外部的路由设备，所述的流量信息收集模块的输出端分别连接所述的异常行为检测模块和异常流量处理模块的输入端，所述的异常行为检测模块的输出端连接所述的异常流量处理模块的输入端，所述的异常流量处理模块的输出端连接该检测装置外部的路由设备，该方法在流量信息收集模块从外部路由设备获得原始流量数据后；先进行过滤，将明显异常的流量数据发送至所述的异常流量处理模块，并将经过滤的流量数据发送至所述的异常行为检测模块；异常行为检测模块根据一确定的检测标准对所述的经过滤的流量数据进行检测，并将检测出的异常流量数据发送至所述的异常流量处理模块，而后，异常行为检测模块根据检测结果自动更新所述的检测标准；最后异常流量处理模块根据所获得的异常流量的流量信息向外部路由设备发送流量阻断指令。从而能利用本发明的装置和方法过滤掉可能的攻击所产生的异常数据，建立运营商或是服务器的正常网络行为模型；进而可以把这个模型跟实时数据进行比较，以检测实时网络流量的行为是否为异常；如行为正常，则通过收集这部分的实时数据以动态修正网络行为模型，实现自学习的功能，如行为异常，则从数据流中分析出异常流量的来源，首先保障用户的应用服务，然后再根据用户设置的策略，对异常流量进行阻断。从而真正实现了在大规模数据流量分析的实际应用环境中，快速有效的总结流量行为，识别异常流量，并且避免无法检测出缓慢异常流量增加的情况，提高检测的准确性。本发明的基于数据流行为分析的网络访问异常检测装置及方法应用方式较为简单，应用成本低廉，且适用范围广泛。

## 附图说明

[0057] 图 1 为本发明的基于数据流行为分析的网络访问异常检测装置的结构示意图。

[0058] 图 2 为本发明的基于数据流行为分析的网络访问异常检测方法的步骤流程图。

[0059] 图 3 为本发明的基于数据流行为分析的网络访问异常检测装置中的异常行为检测模块进行数据流异常探测的流程图。

[0060] 图 4 为本发明的基于数据流行为分析的网络访问异常检测装置中异常行为检测模块进行实际数据流量值和预测数据流量值比较的流程图。

## 具体实施方式

[0061] 为了能够更清楚地理解本发明的技术内容，特举以下实施例详细说明。

[0062] 请参阅图 1 所示，为本发明的基于数据流行为分析的网络访问异常检测装置的结构示意图。

[0063] 在具体实施方式中，该装置包括流量信息收集模块、异常行为检测模块和异常流量处理模块，所述的流量信息收集模块的输入端连接该检测装置外部的路由设备，所述的流量信息收集模块的输出端分别连接所述的异常行为检测模块和异常流量处理模块的输入端，所述的异常行为检测模块的输出端连接所述的异常流量处理模块的输入端，所述的异常流量处理模块的输出端连接该检测装置外部的路由设备。

[0064] 本发明还提供了一种利用所述的装置实现基于数据流行为分析的网络访问异常检测方法。该方法的一种实施方式，如图 2 所示，其包括以下步骤：

[0065] (1) 所述的流量信息收集模块从外部路由设备获得原始流量数据；

[0066] (2) 所述的流量信息收集模块对原始流量数据进行过滤，将明显异常的流量数据发送至所述的异常流量处理模块，并将经过滤的流量数据发送至所述的异常行为检测模块；

[0067] (3) 所述的异常行为检测模块根据一确定的检测标准对所述的经过滤的流量数据进行检测，并将检测出的异常流量数据发送至所述的异常流量处理模块；

[0068] (4) 所述的异常行为检测模块根据检测结果自动更新所述的检测标准；

[0069] (5) 所述的异常流量处理模块根据所获得的异常流量的流量信息向外部路由设备发送流量阻断指令。

[0070] 其中，所述的原始流量数据为 netflow v5 格式数据或 sFlow 格式数据。

[0071] 在一种优选的实施方式中，该方法的步骤 (2) 具体包括以下步骤：

[0072] (21) 所述的流量信息收集模块对原始流量数据进行解析，获得流量数据信息；

[0073] (22) 所述的流量信息收集模块将明显异常的原始流量数据的流量数据信息存入一异常流量数据库；

[0074] (23) 所述的流量信息收集模块将其余的流量数据存入一待检测流量数据库。

[0075] 其中，所述的流量数据信息包括源 IP 地址、源端口、目的 IP 地址、目的端口、协议类型、端口号、字节数、数据包数及数据流产生时间。

[0076] 该方法的步骤 (3) 具体包括以下步骤：

[0077] (31) 所述的异常行为检测模块读取所述的待检测流量数据库中的流量数据；

[0078] (32) 所述的异常行为检测模块基于前一周期内一确定时间段的数据流量值生成本周期内对应时间段的流量数据的预测值；

[0079] (33) 所述的异常行为检测模块将本周期该时间段的预测值与本周期内该时间段的流量数据实际值比较，判断两者差距是否大于预设的阈值，若大于，则确定该流量数据为异常流量数据，并进入步骤 (34)，若不大于，则进入步骤 (4)；

[0080] (34) 所述的异常行为检测模块将该异常流量数据的信息存入所述的异常流量数据库，并进入步骤 (5)。

[0081] 其中，步骤 (32) 具体是指：设第  $m$  个周期的第  $i$  个时间段内的实际数据流量值为  $T_{(i,m)}$ ，则根据以下公式得到在第  $m+1$  个周期内对应的第  $i$  个时间段内的预测数据流量值  $P_{(i,m+1)}$ ：

$$P_{(i,m+1)} = a_{(i,m)} + b_{(i,m)},$$

[0083] 其中：

$$a_{(i,m)} = 2S'_{(i,m)} - S''_{(i,m)},$$

$$b_{(i,m)} = \frac{\alpha}{1-\alpha} (S'_{(i,m)} - S''_{(i,m)}),$$

[0086]  $S'_{(i,m)}$  与  $S''_{(i,m)}$  分别为在第  $m$  个周期中第  $i$  个时间段的预测参数；

$$S'_{(i,m)} = \alpha T_{(i,m)} + (1-\alpha) S'_{(i,m-1)},$$

$$S''_{(i,m)} = \alpha S'_{(i,m)} + (1-\alpha) S''_{(i,m-1)},$$

[0089]  $\alpha$  为预设的预测敏感系数。

[0090] 该方法的步骤 (4) 具体是指 :所述的异常行为检测模块将第  $m+1$  个周期的第  $i$  个时间段的流量数据预测值  $P_{(i, m+1)}$  设置为 :

[0091]  $P_{(i, m+1)} = T_{(i, m)}$ 。

[0092] 该方法的步骤 (5) 具体包括以下步骤 :

[0093] (51) 所述的异常流量处理模块读取所述的异常流量数据库中的流量数据信息 ;

[0094] (52) 所述的异常流量处理模块统计所述的流量数据信息中的外部 IP 地址、协议类型及端口号 ;

[0095] (53) 所述的异常流量处理模块向外部路由设备发送针对所述的外部 IP 地址的流量阻断指令。

[0096] 在进一步优选的实施方式中,所述的步骤 (33) 具体包括以下步骤 :

[0097] (33-1) 所述的异常行为检测模块判断  $|T_{(i, m)} - P_{(i, m)}|$  是否大于预设的阈值 ;

[0098] (33-2) 若大于,则确定第  $m$  个周期的第  $i$  个时间段的流量数据为异常流量数据,并进入步骤 (34) ;

[0099] (33-3) 若不大于,则进入步骤 (4)。

[0100] 所述的步骤 (34) 具体包括以下步骤 :

[0101] (34-1) 所述的异常行为检测模块依确定的选取规则采集所述的第  $m$  个周期的第  $i$  个时间段的流量数据的流量数据信息 ;

[0102] (34-2) 所述的异常行为检测模块并将所采集的流量数据信息存入所述的异常流量数据库,并进入步骤 (5)。

[0103] 在一种更为优选的实施方式中,在第 1 个周期中第  $i$  个时间段的预测参数  $S'_{(i, 0)}$  与  $S''_{(i, 0)}$  分别为 : $S'_{(i, 0)} = S''_{(i, 0)} = T_{(i, 1)}$ 。

[0104] 在另一种进一步优选的实施方式中,所述的选取规则具体为以下之一 :

[0105] (1) 采集该时间段内流量数据字节或数据包较大的数据流的流量数据信息 ;

[0106] (2) 采集非关键端口产生的数据流的流量数据信息 ;

[0107] (3) 综合采集非关键端口产生的数据字节或数据包较大的数据流的流量数据信息。

[0108] 在本发明的应用中,本发明的基于数据流行为分析的网络访问异常检测装置主要包括三个主要的部分 :流量信息收集模块 (DCC), 异常行为检测模块 (DEC), 异常流量处理模块 (APC), 其结构如图 1 所示。

[0109] 在该装置中, DCC 提供从路由器上收集流量数据, 流量数据可以使用多种表示形式, 包括但不限于 CISCO 的 netflow v5 格式, 其具体格式如下 (参见 CISCO 公布的相关格式文档 :[http://www.cisco.com/en/US/docs/net\\_mgmt/netflow\\_collection\\_engine/3.6/user/guide/format.html](http://www.cisco.com/en/US/docs/net_mgmt/netflow_collection_engine/3.6/user/guide/format.html)) :

[0110] 1) NetFlow V5 的数据头格式 :

[0111]

字节序	内容	描述
0-1	version	网络数据流输出格式版本号
2-3	count	流输出端口号 (1-30)
4-7	SysUptime	输出设备启动后，数毫秒内的流经时间
8-11	unix_secs	从 0000 UTC 1970 开始数秒内的流量计数
12-15	unix_nsecs	从 0000 UTC 1970 数纳秒内的残余
16-19	flow_sequence	总可见流的顺序计数
20	engine_type	流开关驱动类型
21	engine_id	流开关驱动号
22-23	sampling_interval	采样模式中取开头 2 字节；保留 14 个字节放置采样区间值

[0112] 2) NetFlow V5 的流量数据格式：

[0113]

字节序	内容	描述
0-3	srcaddr	源 IP 地址
4-7	dstaddr	目标 IP 地址
8-11	nexthop	下一跳跃路由器的 IP 地址
12-13	input	输入端的 SNMP 索引
14-15	output	输出端的 SNMP 索引
16-19	dPkts	流内数据包
20-23	dOctets	总层数, 流内数据包内的 3 个字节
24-27	First	流开始的 SysUptime
28-31	Last	收到流内最后数据包的 SysUptime
32-33	srcport	TCP/UDP 源地址端口号或等效值
34-35	dstport	TCP/UDP 目标地址端口号或等效值
36	pad1	未使用 (zero) 字节

37	tcp_flags	TCP 标志累计
38	prot	IP 协议类型 (例如, TCP = 6 ; UDP = 17)
39	tos	IP 服务类型 (ToS)
40-41	src_as	源地址的自动系统编号, 非原级即同级
42-43	dst_as	目标地址的自动系统编号, 非原级即同级
44	src_mask	源地址前置掩码字节
45	dst_mask	目标地址前置掩码字节
46-47	pad2	未使用 (zero) 字节

[0114] DCC 接收的流量数据也可以为 sFlow 格式的数据或其它格式的数据, sFlow 的具体格式如下 (参见 <http://www.ietf.org/rfc/rfc3176.txt>) :

[0115]

int_32 sFlow 版本号 (2 4 5)
int_32 IP 版本 (1 为 IPV4, 2 为 IPV6)
sFlow 代理 IP 地址 (IPV4 占用 4 字节, IPV6 占用 16 字节)
int_32 代理子 id
int_32 数据包序列号
int_32 交换机运行时间
int_32 数据采样包的个数
数据采样包的信息序列 (可能会占用多个字节)

[0116] DCC 的功能是将以这些格式表示的数据流信息进行解析, 从而得到实际的数据流信息。数据流的信息至少包括数据流的源地址, 目的地址, 源端口, 目的端口, 协议, 字节数, 包数, 数据流产生的时间, 以及一些必要的标志位。在还原为数据流信息后, DCC 会根据事先定义的异常流量的特征模式进行检测, 去掉明显属于异常流量的数据流信息, 而将通过检测的流量存入到数据库中, 作为下一步的流量行为分析与异常检测的输入数据。DCC 在发现具有明显特征的异常流量情况下, 会将这些异常流量的相关信息保存到数据库中, 为后续阻止异常流量提供依据, 这些信息通常包括源 IP 地址, 源端口, 目的 IP 地址, 目的端口, 协议, 端口号, 字节数, 包数, 数据流产生的时间等。

[0117] DEC 的功能是对原始流信息进行异常检测。在流量异常的检测中, DEC 需要根据预先设定的周期, 以及间隔的时长进行学习。通常, 用户的流量行为具有一定的周期性。例

如,以一个礼拜为周期,通常在工作日(星期一到星期五)的流量相对较多,而周末(周六到周日)的流量相对较少,到下个礼拜流量行为开始进行有规律的循环。在一个周期内,我们需要设定一个学习的时间间隔,时间间隔作为流量行为学习的最细颗粒度,DEC会将时间间隔中的流量进行统计,以便与下个周期中的相同时间间隔内的流量进行比较。例如,我们设置7天为一个周期而每天作为一个时间间隔,那么在DEC系统中,就会对每个礼拜一的数据流量进行比较与学习。

[0118] 设一个周期内共有多个时间间隔,且当前处于第m个周期的第i个时间段内,则在这个时间段结束时,DEC会根据前一个(第m-1个)周期内的第i时间段内的所有流量的学习值做出预测,产生预测值 $P_{(i,m)}$ ,同时DEC会从实时的流量信息中统计出真实的流量值 $T_{(i,m)}$ ,DEC系统将 $P_{(i,m)}$ 与 $T_{(i,m)}$ 进行比较,一旦两者的差值达超过了预设的阀值,则判断出有异常流量的产生。DEC所实现的功能的流程图如图3与图4所示,其具体算法如下:

$$[0119] P_{(i,m+1)} = a_{(i,m)} + b_{(i,m)}$$

$$[0120] a_{(i,m)} = 2S'_{(i,m)} - S''_{(i,m)}$$

$$[0121] b_{(i,m)} = \frac{\alpha}{1-\alpha} (S'_{(i,m)} - S''_{(i,m)})$$

$$[0122] S''_{(i,m)} = \alpha S'_{(i,m)} + (1-\alpha) S''_{(i,m-1)}$$

$$[0123] S'_{(i,m)} = \alpha T_{(i,m)} + (1-\alpha) S'_{(i,m-1)}$$

[0124] 其中: $P_{(i,m+1)}$ 为根据以往第m个周期中第i个时间段数据流量的学习结果预测出的第m+1个周期中第i个时间段的数据流量;

[0125]  $T_{(i,m)}$ 为第m个周期时,第i个时间段实际测量的数据流量值;

[0126]  $S'_{(i,m)}$ 与 $S''_{(i,m)}$ 为在第m个周期中第i个时间段的预测参数,根据上个周期计算出的预测参数以及上个周期的实际测量值计算得到。并且,在系统初始运行时, $S'_{(i,0)} = S''_{(i,0)} = T_{(i,1)}$ ;当系统检测出某个周期存在异常流量时, $S'_{(i,m)} = S'_{(i,m-1)}$ , $S''_{(i,m)} = S''_{(i,m-1)}$ ;

[0127]  $\alpha$ 为预测系统的敏感系数,当 $\alpha$ 越大时,系统更加依赖于上一周期的数据,当 $\alpha$ 越小时,系统更加依赖于之前第m个周期的历史数据。

[0128] 一旦判断出有异常流量产生,DEC会将这个时间段内的真实流量按照一定的策略收集其中的部分数据流信息,存入到异常流量数据。策略可以采用:1、将这个时间段内的流量按照数据字节或者数据包的大小进行排序,并选取排名靠前的数据流;2、将这个时间段内的流量按照端口进行分类,选取对服务器而言非关键端口产生的流量数据;3、将策略1与策略2进行组合使用。

[0129] APC根据从异常流量数据表中获得异常流量信息,统计出导致异常流量的外部IP地址,协议,端口等信息,通过访问控制列表(ACL)策略、防火墙(Firewall)策略或者专业的防火墙设备配置防御性策略,禁止这些外部IP地址的访问,从而达到保护本地服务器正常运行的目的。在进行异常流量控制时,APC会根据当前首先要保护的服务,进行优先放行:APC会根据网络中运行的网络服务的优先级进行排序,然后根据优先级由低到高依次进行策略配置,直到在某个时间段内DEC认为当前流量恢复异常。同时,在运行一段时间后,APC也会尝试减少一些防御性的策略配置,一旦减少的配置导致了DEC在连续的多个时间段内仍然检测出异常流量,则恢复此防御性的策略配置,否则不进行恢复。这样可以保证网络中

实用的防御性策略最少,从而提高相关网络设备的可维护性。

[0130] 采用了该发明的基于数据流行为分析的网络访问异常检测装置及方法,其装置包括流量信息收集模块、异常行为检测模块和异常流量处理模块,所述的流量信息收集模块的输入端连接该检测装置外部的路由设备,所述的流量信息收集模块的输出端分别连接所述的异常行为检测模块和异常流量处理模块的输入端,所述的异常行为检测模块的输出端连接所述的异常流量处理模块的输入端,所述的异常流量处理模块的输出端连接该检测装置外部的路由设备,该方法在流量信息收集模块从外部路由设备获得原始流量数据后;先进行过滤,将明显异常的流量数据发送至所述的异常流量处理模块,并将经过滤的流量数据发送至所述的异常行为检测模块;异常行为检测模块根据一确定的检测标准对所述的经过滤的流量数据进行检测,并将检测出的异常流量数据发送至所述的异常流量处理模块,而后,异常行为检测模块根据检测结果自动更新所述的检测标准;最后异常流量处理模块根据所获得的异常流量的流量信息向外部路由设备发送流量阻断指令。从而能利用本发明的装置和方法过滤掉可能的攻击所产生的异常数据,建立运营商或是服务器的正常网络行为模型;进而可以把这个模型跟实时数据进行比较,以检测实时网络流量的行为是否为异常;如行为正常,则通过收集这部分的实时数据以动态修正网络行为模型,实现自动学习的功能,如行为异常,则从数据流中分析出异常流量的来源,首先保障用户的应用服务,然后再根据用户设置的策略,对异常流量进行阻断。从而真正实现了在大规模数据流量分析的实际应用环境中,快速有效的总结流量行为,识别异常流量,并且避免无法检测出缓慢异常流量增加的情况,提高检测的准确性。本发明的基于数据流行为分析的网络访问异常检测装置及方法应用方式较为简单,应用成本低廉,且适用范围广泛。

[0131] 在此说明书中,本发明已参照其特定的实施例作了描述。但是,很显然仍可以作出各种修改和变换而不背离本发明的精神和范围。因此,说明书和附图应被认为是说明性的而非限制性的。

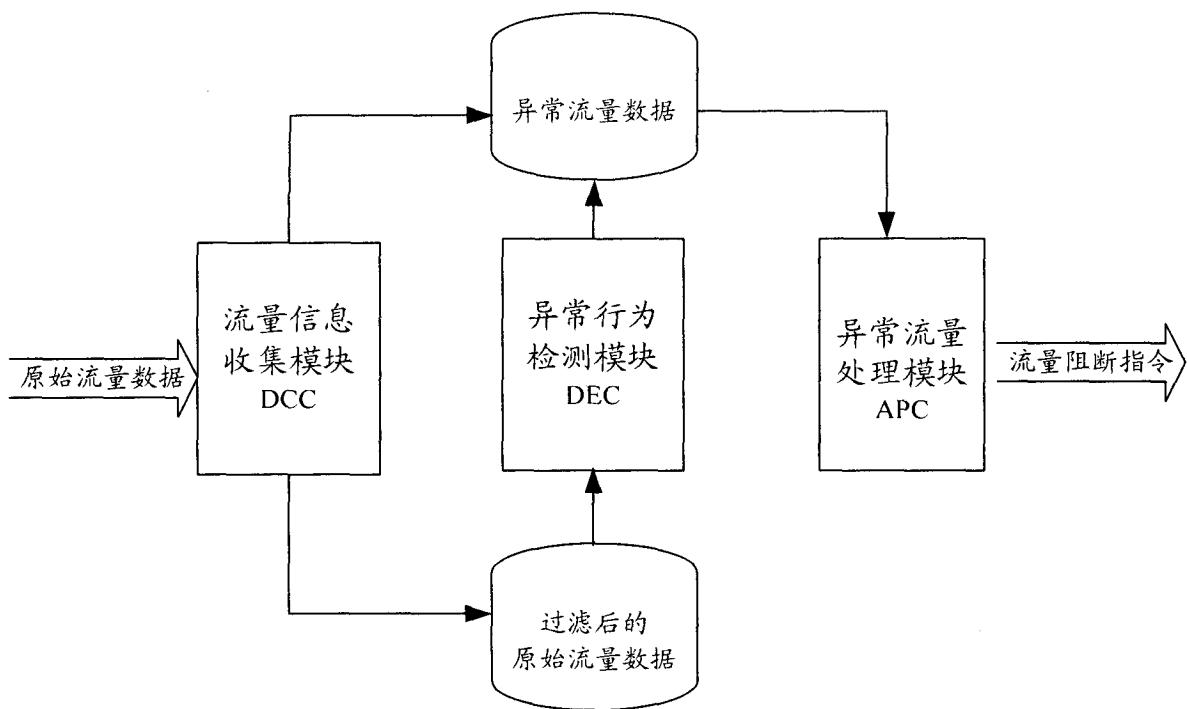


图 1

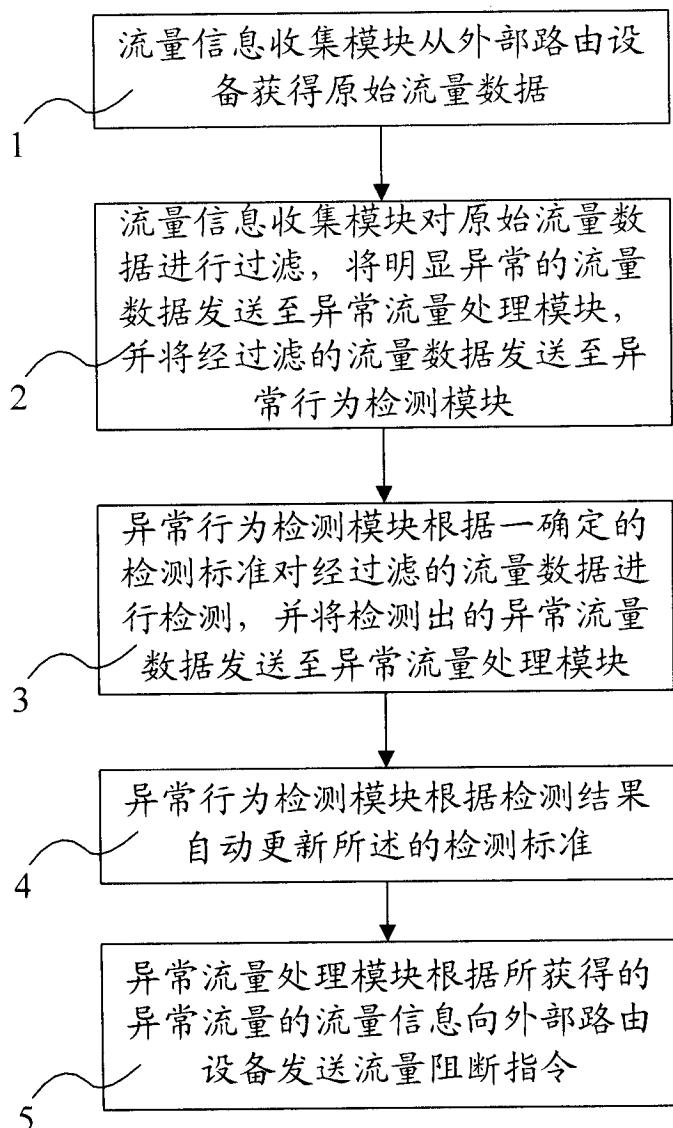


图 2

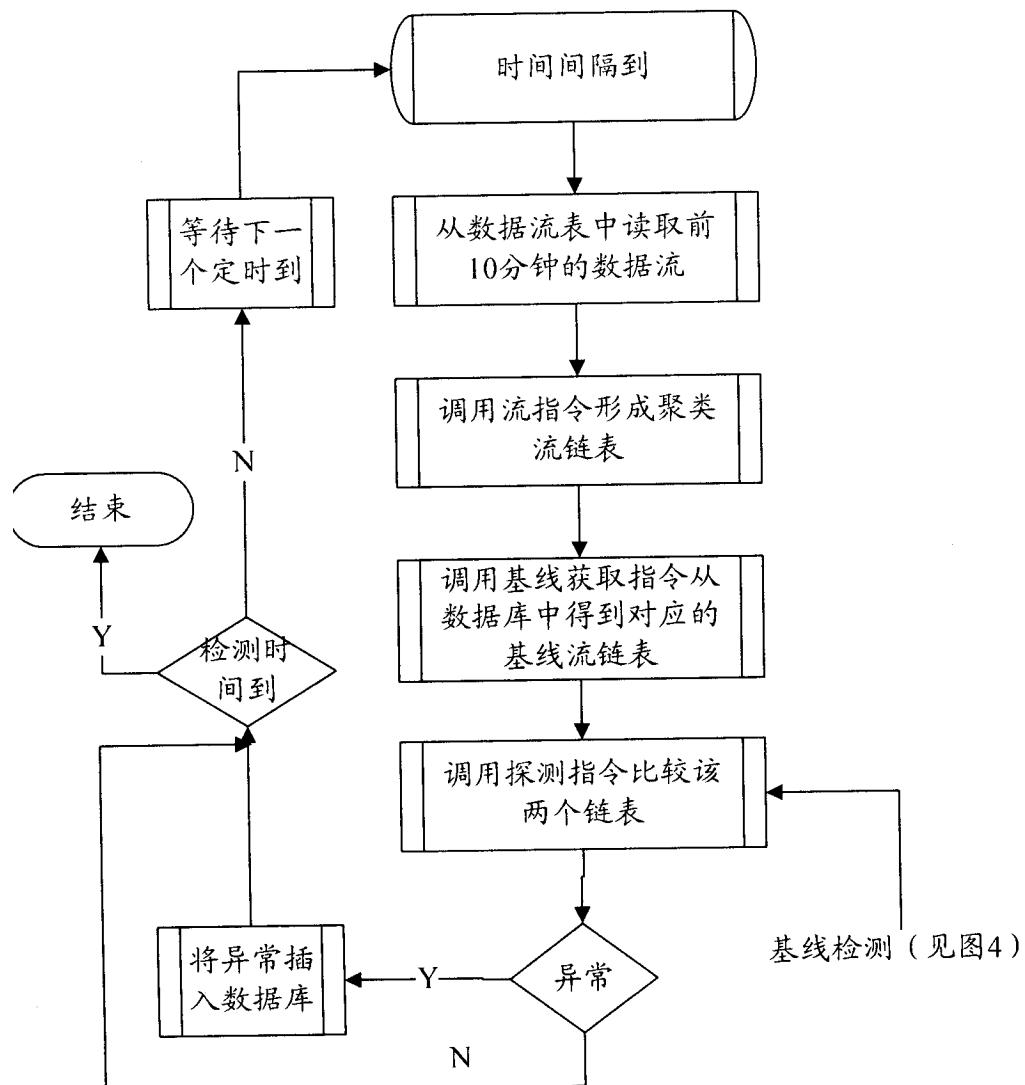


图 3

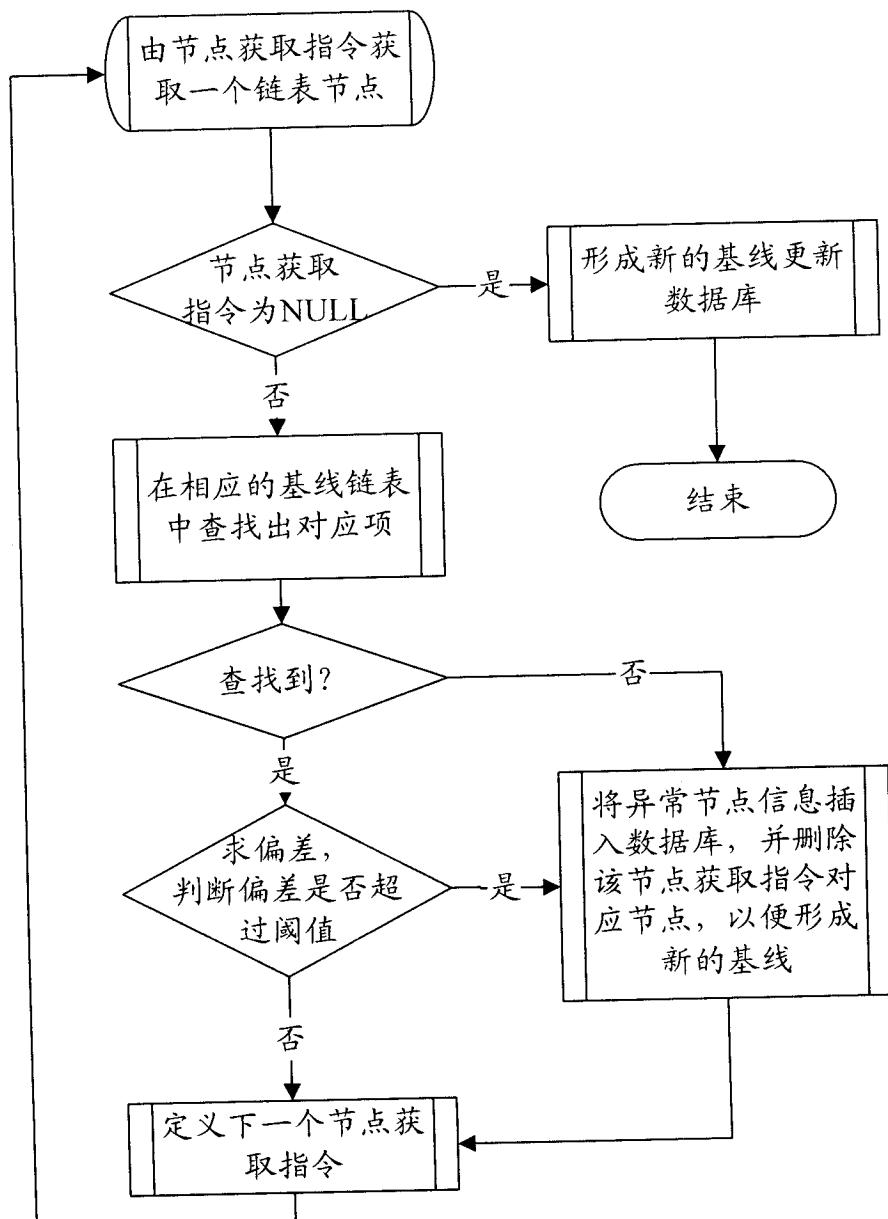


图 4