

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2014 年 7 月 3 日 (03.07.2014)



(10) 国际公布号
WO 2014/101112 A1

- (51) 国际专利分类号:
H04L 29/06 (2006.01)
- (21) 国际申请号: PCT/CN2012/087848
- (22) 国际申请日: 2012 年 12 月 28 日 (28.12.2012)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 杨文宏 (YANG, Wenhong); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR,

CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

(54) Title: WEBSITE IDENTIFICATION METHOD, DEVICE, AND NETWORK SYSTEM

(54) 发明名称: 一种网站识别方法、装置及网络系统

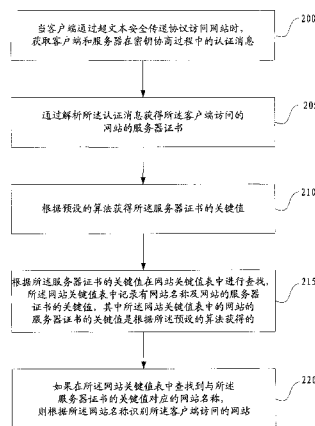


图 2 / FIG. 2

(57) Abstract: Embodiments of the present invention provide a website identification method, device, and network system. The website identification method comprises: when a client accesses a website through a Hypertext Transfer Protocol Secure (HTTPS) protocol, acquiring a server certificate of the accessed website through acquisition and analysis of an authentication message in a key negotiation process of the client and a server; acquiring a key value of the server certificate according to a preset algorithm; searching in a website key-value table according to the key value of the server certificate; and if a website name corresponding to the key value of the server certificate is found in the website key-value table, identifying the website accessed by the client according to the website name. The website identification method can also identify the name of the website accessed by a user when the client accesses the website through the HTTPS protocol.

(57) 摘要:

[见续页]



WO 2014/101112 A1

本发明实施例提供了一种网站识别方法、装置及网络系统，其中，所述网站识别方法在客户端通过 HTTPS 协议访问网站时，通过获取并解析客户端和服务器在密钥协商过程中的认证消息获得所述访问的网站的服务器证书，根据预设的算法获得所述服务器证书的关键值，并根据所述服务器证书的关键值在网站关键值表中进行查找，如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称，则根据所述网站名称识别所述客户端访问的网站。该网站识别方法在客户端通过 HTTPS 协议访问网站时也能识别出用户访问的网站名称。

一种网站识别方法、装置及网络系统

技术领域

本发明涉及通信技术领域，尤其涉及一种网站识别方法、装置及网络系统。

背景技术

在现有的网络运营过程中，为了对用户的网络访问行为进行管理，例如，对用户访问的网站进行过滤、根据用户访问的网站进行计费或实现运营商推出的各种网站套餐业务等等，需要能够识别用户访问的网站。当用户通过超文本传送协议(hypertext transport protocol, HTTP)访问网站时，由于报文是明文传输，因此可以通过分析用户的http报文的内容识别出用户所访问的网站，然而，当用户通过超文本安全传送协议(Hypertext Transfer Protocol Secure, HTTPS)访问网站时，由于应用层报文被封装在加密的传输隧道中，因此，无法从用户访问网站的报文中识别出用户所访问的网站。

发明内容

本发明实施例中提供了一种网站识别方法、装置及网络系统，能够在用户通过HTTPS协议访问网站时识别出用户访问的网站。

第一方面，本发明实施例提供一种网站识别方法，包括：

当客户端通过超文本安全传送协议HTTPS访问网站时，获取所述客户端和所述网站的服务器在密钥协商过程中的认证消息；

通过解析所述认证消息获得所述客户端访问的网站的服务器证书；

根据预设的算法获得所述服务器证书的关键值；

根据所述服务器证书的关键值在网站关键值表中进行查找，所述网站关键值表中记录有网站名称及网站的服务器证书的关键值，其中所述网站关键值表中的网站的服务器证书的关键值是根据所述预设的算法获得的；

如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站

名称, 则根据所述网站名称识别所述客户端访问的网站。

在第一方面的第一种可能的实现方式中, 所述网站识别方法还包括:

如果在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称, 则判断所述服务器证书的格式类型;

如果所述服务器证书的格式类型为X.509格式, 则在所述服务器证书的使用者属性中获得所述网站的域名, 并根据所述网站的域名识别所述客户端访问的网站;

如果所述服务器证书的格式类型为PGP格式, 则在所述服务器证书的用户标识属性中获得所述网站的域名, 并根据所述网站的域名识别所述客户端访问的网站。

结合第一方面, 在第一方面的第二种可能的实施方式中, 所述方法还包括:

如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称, 则不对所述客户端的流量进行计费统计;

如果在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称, 则对所述客户端的流量进行计费统计。

结合第一方面或第一方面的第一种可能实现的方式或第一方面的第二种可能实现的方式, 在第三种可能实现的方式中, 所述获取客户端和服务端在密钥协商过程中的认证消息包括:

监听所述客户端和所述服务器在密钥协商过程中的报文;

根据所述报文中的 HandshakeType 字段识别所述认证消息。

第二方面, 本发明实施例提供一种深度包检测设备, 包括:

获取模块, 用于当客户端通过超文本安全传送协议HTTPS访问网站时, 获取所述客户端和所述网站的服务器在密钥协商过程中的认证消息;

解析模块, 用于通过解析所述获取模块获取的所述认证消息获得所述客户端访问的网站的服务器证书;

计算模块, 用于根据预设的算法获得所述服务器证书的关键值;

查找模块, 用于根据所述计算模块获得的所述服务器证书的关键值在网站关键值表中进行查找, 所述网站关键值表中记录有网站名称及网站的服务器证书的关键值, 其中所述网站关键值表中的网站的服务器证书的关键值是根据所述预设

的算法获得的;

识别模块,如果所述查找模块在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称,则根据所述网站名称识别所述客户端访问的网站。

在第二方面的第一种可能的实现方式中,所述深度包检测设备还包括:

设置模块,用于接收并存储管理服务器发送的所述网站关键值表,其中所述网站关键值表中包含有至少一个网站的网站名称以及网站的服务器证书的关键值。

结合第二方面或第二方面的第一种可能的实现方式,在所述第二方面的第二种可能的实现方式中,所述深度包检测设备还包括:

判断模块,用于当所述查找模块在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称时,判断所述解析模块获得的所述服务器证书的格式类型,并触发所述识别模块;

所述识别模块,还用于当所述判断模块判断所述网站的服务器证书的格式类型为X.509格式时,在所述证书的使用者属性中获得所述网站的域名,并根据所述网站的域名识别所述客户端访问的网站;当所述判断模块判断所述网站的服务器证书的格式类型为PGP格式时,在所述证书的用户标识属性中获得所述网站的域名,并根据所述网站的域名识别所述客户端访问的网站。

结合第二方面或第二方面的第一种可能的实现方式,在第二方面的第三种可能的实现方式中,所述深度包检测设备还包括:

流量管理模块,用于当所述查找模块在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称时,不对所述客户端的流量进行计费统计;当所述查找模块在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称时,对所述客户端的流量进行计费统计。

结合第二方面、第二方面的第一种可能的实现方式、第二方面的第二种可能的实现方式或第二方面的第三种可能的实现方式,在第二方面的第四种可能的实现方式中,所述获取模块具体用于监听所述客户端和所述网站的服务器在密钥协商过程中的报文,并根据所述报文中的HandshakeType字段识别所述认证消息。

第三方面,本发明实施例提供一种网络系统,包括客户端、至少一个网站服务器以及上述任意一种实现方式提供的深度包检测设备,其中:

所述客户端，用于向所述至少一个网站服务器发送基于HTTPS协议的访问请求，与所述至少一个网站服务器进行密钥协商；

所述网站服务器，用于接收所述客户端发送的网络访问请求，并根据所述网络访问请求与所述客户端进行密钥协商。

本发明实施例中提供的网站识别方法，在客户端通过HTTPS协议访问网站时，通过获取客户端访问的网站的服务器证书，根据预设的算法获得所述服务器证书的关键值，并将获得的关键值在网站关键值表中进行查找，通过在该网站关键值表中查找到的与所述服务器证书的关键值对应的网站名称来识别客户端访问的网站，从而在客户端通过HTTPS协议访问网站时也能识别出客户端访问的网站。

附图说明

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。

图1是本发明实施例中提供的一种网站识别方法的应用场景图；

图2为本发明实施例提供的一种网站识别方法的流程图；

图3为本发明实施例提供的又一种网站识别方法的流程图；

图4为本发明实施例提供的客户端与网站的服务器进行密钥协商的信令图；

图5为本发明实施例提供的一种网站识别方法中的服务器证书分析方法的流程图；

图6为本发明实施例提供的一种深度包检测设备的物理结构示意图；

图7为本发明实施例提供的又一种深度包检测设备的结构示意图；

图8为本发明实施例提供的一种网站识别系统的结构示意图。

具体实施方式

为了使本技术领域的人员更好地理解本发明方案，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的实施例，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做出创造性劳动前提下所获得的所有其他实施例,都应当属于本发明保护的范围。

如图 1 所示,图 1 为本发明实施例的一种应用场景,在图 1 所述的一种应用场景中,其中,客户端 100 通过超文本安全传送协议(Hypertext Transfer Protocol Secure, HTTPS)访问网站。网关设备 105 能够对客户端 100 的网络访问行为进行网络接入控制。深度包检测(Deep Packet Inspection, DPI)设备 110 能够对客户端 100 的网络访问行为进行基于应用层的流量检测和控制,可用于检测、分析客户端 100 的网络访问流量,并能够实现策略及计费执行功能(Policy and Charging Enforcement Function, PCEF),及向在线计费系统(Online Charging System, OCS) 115 请求该用户的流量配额并上报客户端 100 的流量,以使 OCS 能够根据客户端 100 的访问流量对用户进行计费。其中, DPI 设备 110 部署于网关设备 105 的下行链路中,客户端 100 访问网站的数据包经过网关设备和 DPI 设备到达待访问的网站服务器 120, DPI 设备 110 对客户端 100 的流量进行分析以识别客户端 110 访问的网站,并根据识别出的网站对客户端 100 的网络访问行为进行控制,例如,若某用户参加了计费套餐业务,则 DPI 设备 110 根据策略与计费规则功能(Policy and Charging Rules Function, PCRF)设备 125 对该用户设定的套餐计费策略对该用户的客户端 100 的访问流量进行计费管理。如果该客户端 100 访问的网站是套餐内的网站,则 DPI 设备 110 不对该客户端 100 的网络流量进行流量统计,如果该客户端 100 访问的网站不是套餐内的网站,则 DPI 设备 110 会对该客户端 100 的网络访问流量进行流量统计,并将客户端 100 的流量统计结果上报给在线计费系统 115 以进行计费处理。可以理解的是,本应用场景下根据识别出的网站对客户端 100 的网络访问行为进行计费管理只是对客户端 100 的网络行为的一种管理,实际应用中,还可以根据识别出的网站对用户进行流量控制或过滤等等。

可以理解的是,客户端100可以是手机、计算机等能够实现网络访问的设备,网关设备105可以是网关GPRS支持节点(Gateway GPRS Support Node, GGSN)或分组数据网关(Packet Data Gateway, PDG)等,在此不做限定。

图2为本发明实施例提供的一种网站识别方法的流程图,该方法应用于客户

端通过HTTPS协议访问网站的场景中,可以由图1中的DPI设备110来执行,该方法包括:

步骤200,当客户端通过HTTPS协议访问网站时,获取所述客户端和所述网站的服务器在密钥协商过程中的认证消息,进入步骤205;

结合图1所示,当客户端100通过HTTPS协议访问网站时,客户端100与访问的网站服务器120会进行密钥协商,该密钥协商过程可以根据安全传输层(Transport Layer Security, TLS)协议进行,其中,TLS协议是在互联网上提供保密安全通道的加密协议,所述服务器证书是所述服务器从认证授权(Certificate Authority, CA)中心申请的一个用于证明服务器用途类型的证书,该证书只有用于对应的服务器的时候,客户端才信任此服务器。由于DPI设备110部署于网关设备105的下行链路中,DPI设备110可以监听获取客户端100和网站服务器120在密钥协商过程中的认证Certificate消息,

步骤205,通过解析所述认证消息获得所述客户端访问的网站的服务器证书,进入步骤210;

在有些情况下,如果获取的Certificate消息中包含有证书链CertificateList,则可以通过提取所述认证消息中的证书链CertificateList部分的第一个证书获得所述访问的网站的服务器证书。

步骤210,根据预设的算法获得所述服务器证书的关键值,进入步骤215;

在本发明实施例中,可以通过预先设置的算法来计算所述服务器证书的内容获得所述服务器证书的关键值,例如,可以采用数字摘要算法计算所述服务器证书的内容获得所述服务器证书的关键值,其中数字摘要算法可以采用信息-摘要算法5(Message-Digest Algorithm, MD5)算法、安全算列算法SHA-1以及RACE原始完整性校验消息摘要(RACE Integrity Primitives Evaluation Message Digest, RIPEMD)等算法,在此不做限定。

步骤215,根据所述服务器证书的关键值在网站关键值表中进行查找,所述网站关键值表中记录有网站名称及网站的服务器证书的关键值,其中所述网站关键值表中的网站的服务器证书的关键值是根据所述预设的算法获得的,进入步骤220;

在本发明实施例中,可以在DPI设备中预先设置网站关键值表,预设的网站

关键值表中记录有各网站名称及网站的服务器证书的关键值,其中,该网站关键值表中的服务器证书的关键值是根据与步骤210中采用的相同的算法对各网站的服务器证书的内容进行计算获得的,比如说采用相同的数字摘要算法。例如,如果步骤215中,预设的网站关键值表中的服务器证书的关键值是通过MD5算法获得的,则在步骤210中,在用户访问网站时,也需要通过MD5算法计算用户访问的服务器证书的关键值。当然,可以理解的是,该网站关键值表也可以预设在其他设备中,只要DPI设备能够访问即可。

步骤220,如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称,则根据所述网站名称识别所述客户端访问的网站。

由于MD5等数字摘要算法能够通过将原文信息经过哈希计算后获得一个数字摘要(即哈希值,也就是本实施例中的关键值),因此,本发明实施例利用数字摘要算法能够将不同的原文转换为不同的哈希值的特性,预先将需要识别的网站的服务器证书的内容通过预设的数字摘要算法计算获得服务器证书的关键值,并将获得的服务器证书的关键值以及该网站的名称对应存储于网站关键值表中,当DPI设备需要识别用户访问的网站时,可以通过将用户访问的数据流中获取的服务器证书的关键值与预设的服务器证书的关键值进行比较,如果在该网站关键值表中查找到与客户端访问的网站的服务器证书的关键值相同的网站的服务器证书的关键值,则可以根据网站关键值表获得与服务证书的关键值对应的网站的名称。

本发明实施例中提供的网站识别方法,在客户端通过HTTPS协议访问网站时,通过获取客户端访问的网站的服务器证书,根据预设的算法获得所述服务器证书的关键值,并将获得的关键值在网站关键值表中进行查找,通过在该网站关键值表中查找到的与所述服务器证书的关键值对应的网站名称来识别客户端访问的网站,从而在客户端通过HTTPS协议访问网站时也能识别出客户端访问的网站。

图3为本发明实施例提供的又一种网站识别方法的流程图,该方法应用于客户端通过HTTPS协议访问网站的场景中,可以由图1中的DPI设备110来执行,如图3所示,该方法包括:

步骤300,接收并存储管理服务器发送的网站关键值表,其中所述网站关键

值表中包含有至少一个网站的网站名称以及网站的服务器证书的关键值,进入步骤305;

在实际应用中,可以由管理服务器预先获取至少一个网站的服务器证书,根据预设的数字摘要算法计算所述获取的至少一个网站的服务器证书的关键值,将计算获得的服务器证书的关键值与网站名称记录在网站关键值表中,并将该网站关键值表发送给DPI设备。

具体的,管理服务器可以预先获取需要关注的网站的服务器证书,例如,在需要对用户流量进行计费管理时,可以预先获取网站套餐中规定的可访问的网站的服务器证书,例如Facebook、YouTube等网站的服务器证书,还可以根据需要预先获取需要控制访问的网站的服务器证书,具体可以根据需要来进行设定,在此不做限定。并且,管理服务器可以通过预先通过浏览器访问需要关注的网站,将该网站的服务器证书下载到本地,再通过预设的算法(例如,MD5算法)对下载的服务器证书进行计算获取的服务器证书的关键值,得到网站关键值表。其中,网站关键值表的形式可以如下表一所示:

序号	网站域名	证书校验值
1	Facebook	2C89277EE96E8B15E2F9A90FED69B0B2
2	YouTube	5F43086FACDB301B74334183A0472584
...	...	...

表一 网站关键值表

可以理解的是,当网站的服务器证书有更新、需要新增或删除需要关注的网站时,需要及时更新预设的网站关键值表,具体的,当关注的网站的服务器证书有更新时,可以根据更新后的网站服务器证书及时更新网站关键值表中的服务器证书的关键值;当需要新增关注的网站时,根据新增的网站的服务器证书增加网站关键值表中的服务器证书的关键值及网站名称;或者当需要减少需关注的网站时,可以删除网站关键值表中与该需要删除的网站名称对应的服务器证书的关键值。

步骤305,获取客户端和服务器的在密钥协商过程中的认证消息,所述认证消息中携带有客户端访问的网站的服务器证书,进入步骤310;

结合图1所示,当客户端100通过HTTPS协议访问网站时,由于DPI设备110部署于网关设备105的下行链路中,DPI设备可以监听获取客户端100和网站服务器

120在密钥协商过程中的认证消息，其中该认证消息中携带有客户端访问的网站的服务器证书。具体的，客户端与服务器基于TLS协议的密钥协商过程具体的可以如图4所示，该密钥协商过程可以包括：

步骤400，客户端向要访问的网站的服务器发送密钥协商请求消息ClientHello，进入步骤405；

在ClientHello消息中，客户端将其能够支持的加密方式通知服务器，进入步骤405。

步骤405，所述服务器向所述客户端返回密钥协商响应消息ServerHello，进入步骤410；

接收到ClientHello消息的服务器会向客户端返回一个ServerHello消息，在ServerHello消息中指定采用客户端能够支持的加密方式中的一种加密方式。

步骤410，所述服务器向所述客户端发送认证Certificate消息，在所述Certificate消息中包含有所述服务器的服务器证书，进入步骤415；

服务器在向客户端发送ServerHello消息后，会接着发送Certificate消息，在所述Certificate消息中会包含有所述服务器的服务器证书，以使所述客户端对所述服务器证书进行验证。其中，所述服务器证书是所述服务器从认证授权（Certificate Authority，CA）中心申请的一个用于证明服务器用途类型的证书，该证书只有用于对应的服务器的时候，客户端才信任此服务器。其中，Certificate消息的格式可以如下表二所述：

CertificateList			
Length Hi	HandshakeType	Length	
ContentType	Version		Length Lo
TCP			
IP			
Frame			

表二 Certificate消息格式

其中，HandshakeType字段的值为11表示该消息为Certificate消息。

需要说明的是，通常Certificate消息中会包含有证书链CertificateList，证书链中可以包含有多个证书，在这多个证书中，列在后面的证书用于验证列在前面的证书是否为通过权威机构认证的证书，例如：当通过HTTPS访问谷歌google网站时，会获取到多个证书，其中最一个证书的是google的服务器证书，后面的证

书可能是某个认证机构的证书, 认证机构的证书用来证明google的服务器证书为合法证书。因此, 当获得证书链CertificateList时, 则第一个证书为要访问的网站的服务器证书。

步骤415, 在所述客户端对所述服务器的服务器证书验证通过后, 所述客户端向所述服务器发送密钥交换ClientKeyExchange消息, 其中所述ClientKeyExchange消息中包含采用服务器的公钥进行加密的密钥, 进入步骤420;

步骤420, 所述客户端向所述服务器发送密钥协商完成Finished消息, 表示协商完成, 进入步骤425;

步骤425, 所述服务器向所述客户端返回密钥协商完成Finished消息, 表示协商完成, 进入步骤430;

步骤430, 所述客户端和所述服务器进行数据报文的交互, 其中所述数据报文为经过所述协商的密钥加密的数据报文。

可以理解的是, DPI设备还可以以旁路部署的方式部署于网络链路中, 当DPI设备以旁路方式部署时, 可以通过镜像或抄送等方式获得客户端100与网站服务器120在密钥协商过程中的协商报文, 从而获得客户端100和网站服务器120在密钥协商过程中的认证消息。

步骤310, 通过解析所述认证消息获得所述访问的网站的服务器证书, 进入步骤315;

步骤315, 根据预设的算法获得所述访问的网站的服务器证书的关键值, 进入步骤320;

具体地, 可以通过预先设置的算法来计算所述服务器证书的内容获得所述服务器证书的关键值, 例如, 可以采用数字摘要算法计算所述服务器证书的内容获得所述服务器证书的关键值, 其中数字摘要算法可以包括MD5或SHA-1等算法, 其中, 本步骤采用的算法需要与管理服务器预设网站关键值表时采用的算法相同。

步骤320, 根据所述服务器证书的关键值在所述网站关键值表中进行查找, 进入步骤325;

步骤325, 判断是否在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称, 如果查找到与所述服务器证书的关键值对应的网站名称, 则进入步骤330, 否则进入步骤335;

步骤330, 根据所述网站名称识别所述客户端访问的网站, 进入步骤340;

步骤335, 分析所述网站的服务器证书, 获得所述用户访问的网站的域名, 并根据所述网站的域名识别所述客户端访问的网站, 进入步骤340;

若根据获得的网站的服务器证书的关键值没有在预设的网站关键值表中查找到相同的服务器证书的关键值, 在某些情况下, 为了识别客户端访问的网站的名称, 可以通过在线分析步骤310中获得的所述访问的网站的服务器证书来识别客户端访问的网站的名称, 具体的分析方法可以如图5所示, 包括:

步骤505, 判断所述访问的网站的服务器证书的格式类型, 如果所述访问的网站的服务器证书为X. 509格式的服务器证书, 则进入步骤510, 如果所述访问的网站的服务器证书为PGP格式的服务器证书, 则进入步骤520;

本领域人员可以知道, 当前TLS协议支持的服务器证书包括两种格式的服务器证书, 一种是X. 509格式的服务器证书, 另一种是PGP格式的服务器证书。在具体分析服务器证书时, 需要先识别出步骤310中获得的服务器证书的格式类型, 再根据不同格式类型的证书的数据结构对该服务器证书进行具体分析。具体的, 在协商过程中, 可以通过ClientHello消息和ServerHello消息中的CertificateType字段来协商服务器证书的格式, 如果CertificateType字段为0, 表示服务器证书为X. 509格式的服务器证书, 如果CertificateType字段为1, 表示服务器证书为PGP格式的证书。其中, X. 509是被广泛使用的数字证书标准, 是由国际电联电信委员会 (ITU-T) 为单点登录 (Single Sign-on, SSO) 和授权管理基础设施 (Privilege Management Infrastructure, PMI) 制定的公钥基础设施 (Public Key Infrastructure, PKI) 标准; PGP (Pretty Good Privacy, 更好的保护隐私) 是一个基于RSA公匙加密体系的邮件加密软件, 它使用一对数学上相关的钥匙, 其中一个 (公钥) 用来加密信息, 另一个 (私钥) 用来解密信息。PGP中的每个公钥和私钥都伴随着一个密钥证书, 也就是PGP格式的密钥证书, 其中PGP格式包括开放的Open PGP格式的证书。

步骤510, 根据X. 509格式证书的数据结构解析所述访问的网站的服务器证书, 获得所述服务器证书中的使用者属性, 进入步骤515;

其中, X. 509格式的证书中一般包含有:

Certificate证书;

Version 版本;
Serial Number 序列号;
Algorithm ID 算法标识;
Issuer 颁发者;
Validity 有效期;
Subject 使用者;
Subject Public Key Info 使用者公钥信息;
Public Key Algorithm 公钥算法;
Subject Public Key 公钥;
Certificate Signature Algorithm 证书签名算法;
Certificate Signature 证书签名等信息。

其中,“Subject 使用者”属性中具有网站的域名信息。因此,若步骤310中获得的访问的网站的服务器证书为X.509格式的证书,则可以根据X.509格式的证书的数字结构获得该证书的Subject使用者属性。当然可以理解的是,在确定服务器证书的数字格式时,还需要确定该证书采用的版本,按照不同的版本格式解析证书。

步骤515,通过解析所述使用者属性,获得所述访问的网站的域名;

例如,可以通过解析所述访问的服务器证书的Subject使用者属性,获得该客户端访问的网站的域名,例如,在X.509格式的服务器证书中,Subject属性的格式可以为: www.facebook.com。

步骤520,通过解析PGP格式的证书获得用户标识属性,进入步骤525;

PGP的密钥证书一般包含以下内容:

密钥内容:通常为用长达百位的大数字表示的密钥;

密钥类型:表示该密钥为公钥还是私钥;

密钥长度:用于表示密钥的长度,通常以二进制位表示;

密钥编号:用以唯一标识该密钥;

创建时间:用于表示该证书创建的时间;

用户标识:密钥创建人的信息,如创建人的姓名、电子邮件等;

密钥指纹:为128位的数字,是密钥内容的提要表示密钥唯一的特征;

中介人签名：中介人的数字签名，声明该密钥及其所有者的真实性，包括中介人的密钥编号和标识信息。

其中，“用户标识”部分用于标识密钥创建人的信息，这些信息可以用创建人的姓名、电子邮件等来表示，因此，如果网站的服务器证书是PGP格式的密钥证书，则从服务器证书中的“用户标识”部分的信息中可以知道该网站的标识，从而能够识别出用户访问的网站。

步骤525，通过解析所述用户标识属性获得访问的网站的域名。

具体的，在PGP格式的服务器证书中，网站的域名包含在用户标识属性中，其中，用户标识可以是电子邮件地址、姓名等能够识别用户的标识。例如，在PGP格式的服务器证书中，用户标识可以表示为：administrator@facebook.com。

步骤340，根据识别出的所述客户端访问的网站对所述客户端访问的数据流进行管理。

当识别出客户端访问的网站后，可以根据识别出的该客户端访问的网站以及预设的控制策略对用户的网络访问行为进行管理，例如，可以根据访问的网站名称和预设的网站套餐策略确定是否需要对该客户端的流量进行计费管理，或对客户端的访问流量进行控制或过滤等，在此不做限定。

本发明实施例中的网站识别方法，当客户端通过HTTPS协议访问网站时，在通过预设的网站关键值表无法识别出客户端访问的网站时，可以通过在线分析客户端访问的网站的服务器证书的方式来识别客户端访问的网站。本发明实施例提供的网站识别方法更加全面，网站识别的准确性更高。并且，本发明实施例提供的网站识别方法，当识别出客户端访问的网站时，可以及时的根据识别出的网站名称对用户的网络访问行为进行管理，从而能够为用户提供更加及时、高效及精细化的网络管理。

在一种应用场景下，当使用本发明实施例所述的网站识别方法来对用户进行流量计费管理时，预设的网站关键值表中可以是用户预定的套餐中规定可以访问的网站的域名及服务器证书的关键值，当在步骤325中，判断在所述网站关键值表中查找到与所述访问的网站的服务器证书的关键值相同的网站的服务器证书的关键值时，说明该客户端访问的是套餐内的网站，可以不对客户端的访问进行计费流量统计，当在步骤325中，判断在所述网站关键值表中没有查找到与所

述访问的网站的服务器证书的关键值相同的网站的服务器证书的关键值时,说明客户端访问的不是套餐内的网站,从而可以直接对客户端的访问流量进行及时的流量统计,并将统计的流量上报给计费服务器进行计费。在这种应用场景下,当在步骤325中,没有在网站关键值表中查找到与该访问的网站的服务器证书的关键值相同的网站的服务器证书的关键值时,并不需要再进入步骤335来通过分析网站的服务器证书获得用户访问的网站的域名,而可以直接对客户端的流量进行控制管理。本发明实施例提供的这种网站识别方法当应用于计费管理时,这种通过比较服务器证书的关键值的方法获得客户端访问的网站名称,识别效率高,且操作简单,从而能对客户端的访问流量进行及时的控制管理,提高运营商的服务质量。

图6为本发明实施例提供的一种深度包检测(Deep Packet Inspection, DPI)设备的物理结构示意图,如图6所示,所述DPI设备60包括:

处理器(processor)610,通信接口(Communications Interface)620,存储器(memory)630,通信总线640。

处理器610、通信接口620以及存储器630通过通信总线640完成相互间的通信。

通信接口620,用于与网元通信,比如网关设备、在线计费系统(Online Charging System, OCS)或管理服务器等。

处理器610,用于执行程序632,具体可以执行上述图2至图3所示的方法实施例中的相关步骤。

具体地,程序632可以包括程序代码,所述程序代码包括计算机操作指令。

处理器610可能是一个中央处理器CPU,或者是特定集成电路ASIC(Application Specific Integrated Circuit),或者是被配置成实施本发明实施例的一个或多个集成电路。

存储器630,用于存放程序632。存储器630可能包含高速RAM存储器,也可能还包括非易失性存储器(non-volatile memory),例如至少一个磁盘存储器。

程序632中各功能模块的具体实现可以参见下述图7所述实施例中的相应模块,在此不再赘述。

图 7 为本发明实施例提供的一种 DPI 设备的结构示意图,如图 7 所示,所述 DPI 设备 70 包括:

获取模块700,用于当客户端通过超文本安全传送协议HTTPS访问网站时,获取所述客户端和所述网站的服务器在密钥协商过程中的认证消息;

结合图1所示,当客户端100通过HTTPS协议访问网站时,客户端100与访问的网站的服务器120会进行密钥协商,该密钥协商过程可以根据安全传输层(Transport Layer Security, TLS)协议进行,其中,TLS协议是在互联网上提供保密安全通道的加密协议,所述服务器证书是所述服务器从认证授权(Certificate Authority, CA)中心申请的一个用于证明服务器用途类型的证书,该证书只有用于对应的服务器的时候,客户端才信任此服务器。DPI设备中的获取模块700可以通过监听获取客户端和网站服务器在密钥协商过程中的认证消息,其中该认证消息中携带有客户端访问的网站的服务器证书。具体的,获取模块700可以通过监听所述客户端和所述服务器在密钥协商过程中的报文,并根据所述报文中的HandshakeType字段识别所述认证消息。

解析模块705,用于通过解析所述获取模块700获取的所述认证消息获得所述客户端访问的网站的服务器证书;

在有些情况下,如果获取的Certificate消息中包含有证书链CertificateList,则可以通过提取所述认证消息中的证书链CertificateList部分的第一个证书获得所述访问的网站的服务器证书。

计算模块710,用于根据预设的算法获得所述服务器证书的关键值;

本发明实施例中,可以通过预先设置的算法来计算解析模块705获得的所述服务器证书的内容,获得该服务器证书的关键值,例如,可以采用数字摘要算法计算所述服务器证书的内容得到该服务器证书的关键值,其中数字摘要算法可以采用信息-摘要算法5(Message-Digest Algorithm, MD5)算法、安全算列算法SHA-1以及RACE原始完整性校验消息摘要(RACE Integrity Primitives Evaluation Message Digest, RIPEMD)等算法,在此不做限定。

查找模块715,用于根据计算模块710获得的所述服务器证书的关键值在网站关键值表中进行查找,所述网站关键值表中记录有网站名称及网站的服务器证书

的关键值,其中所述网站关键值表中的网站的服务器证书的关键值是根据所述预设的算法获得的;

在本发明实施例中,可以在DPI设备中预先设置网站关键值表,其中,网站关键值表中的服务器证书的关键值是采用与计算模块710计算时采用的算法相同的算法,比如说采用相同的数字摘要算法。例如,如果预设的网站关键值表中的服务器证书的关键值是通过MD5算法获得的,则在用户访问网站时,计算模块710也需要通过MD5算法计算用户访问的服务器证书的关键值。当然,可以理解的是,该网站关键值表也可以预设在其他设备中,只要DPI设备能够访问即可。

识别模块720,如果查找模块715在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称,则根据所述网站名称识别所述客户端访问的网站。

由于MD5等数字摘要算法能够通过将原文信息经过哈希计算后获得一个数字摘要(即哈希值,也就是本实施例中的关键值),因此,本发明实施例利用数字摘要算法能够将不同的原文转换为不同的哈希值的特性,预先将需要识别的网站的服务器证书的内容通过预设的数字摘要算法计算获得服务器证书的关键值,并将获得的服务器证书的关键值以及该网站的名称对应存储于网站关键值表中,当DPI设备需要识别用户访问的网站时,可以通过查找模块715将用户访问的数据流中获取的服务器证书的关键值与预设的服务器证书的关键值进行比较,如果在该网站关键值表中查找到与客户端访问的网站的服务器证书的关键值相同的网站的服务器证书的关键值,则识别模块720可以根据网站关键值表获得与服务证书的关键值对应的网站的名称。

本发明实施例中的DPI设备,在客户端通过HTTPS协议访问网站时,通过获取客户端访问的网站的服务器证书,根据预设的算法获得所述服务器证书的关键值,并将获得的关键值在网站关键值表中进行查找,从而通过在该网站关键值表中查找到的与所述关键值对应的网站名称来识别用户访问的网站的名称。从而在客户端通过HTTPS协议访问网站时也能识别出用户访问的网站名称。

在另一种情形下,本发明实施例提供的DPI设备70还可以包括:

设置模块725,用于接收并存储管理服务器发送的网站关键值表,其中所述网站关键值表中包含有至少一个网站的网站名称以及网站的服务器证书的关键值,所述网站关键值表中网站的服务器证书的关键值根据所述预设的算法获得。

可以理解的是,当网站的服务器证书有更新、需要新增或删除需要关注的网站时,设置模块725需要及时更新预设的网站关键值表,具体的,当关注的网站的服务器证书有更新时,设置模块725可以根据更新后的网站服务器证书及时更新网站关键值表中的服务器证书的关键值;当需要新增关注的网站时,设置模块725可以根据新增的网站的服务器证书增加网站关键值表中的服务器证书的关键值及网站名称;或者当需要减少需关注的网站时,设置模块725可以删除网站关键值表中与该需要删除的网站名称对应的服务器证书的关键值。

在又一种情形下,本发明上述实施例提供的DPI设备70还可以包括:

判断模块730,用于当所述查找模块715在所述网站关键值表中没有查找到与所述访问的网站的服务器证书的关键值对应的网站名称时,判断所述解析模块705获得的所述服务器证书的格式类型,并触发所述识别模块720;

所述识别模块720,还用于根据判断模块730判断的不同格式的证书对所述访问的网站的服务器证书进行解析来识别客户端访问的网站。

具体的,如果判断模块730判断的所述网站的服务器证书的格式类型是X.509格式的证书,则识别模块720在所述证书的使用者属性中获得所述网站的域名,并根据所述网站的域名识别所述客户端访问的网站;如果判断模块730判断的所述网站的服务器证书的格式类型是PGP格式的证书,则识别模块720在所述证书的用户标识属性中获得所述网站的域名,并根据所述网站的域名识别所述客户端访问的网站。

在又一种情形下,本发明上述实施例提供的DPI设备70还可以包括:

流量管理模块735,用于根据识别模块720识别出的所述客户端访问的网站对所述客户端访问的数据流进行管理。

当识别出客户端访问的网站后,流量管理模块735可以根据识别模块720识别出的该客户端访问的网站以及预设的控制策略对用户的网络访问行为进行管理,例如,可以根据访问的网站的名称和预设的网站套餐策略确定是否需要对该客户端的流量进行计费管理或计费流量统计,或对客户端的访问流量进行控制或过滤等,在此不做限定。

本发明实施例中的DPI设备,当客户端通过HTTPS协议访问网站时,在通过预设的网站关键值表无法识别出客户端访问的网站时,可以通过在线分析客户端访

问的网站的服务证书的方式来识别客户端访问的网站。使得对网站的识别更加全面，网站识别的准确性更高。并且，本发明实施例提供的DPI设备，当识别出客户端访问的网站时，可以及时的根据识别出的网站名称对用户的网络访问行为进行管理，从而能够为用户提供更加及时、高效及精细化的网络管理。

在一种应用场景下，当使用本发明实施例所述的DPI设备来对用户进行流量计费管理时，预设的网站关键值表中可以是用户预定的套餐中规定可以访问的网站的域名及服务器证书的关键值，当查找模块715在所述网站关键值表中查找到与所述访问的网站的服务证书的关键值相同的网站的服务证书的关键值时，说明该客户端访问的是套餐内的网站，则流量管理模块735可以不对客户端的访问进行计费流量统计，而直接放行；当查找模块715在所述网站关键值表中没有查找到与所述访问的网站的服务证书的关键值相同的网站的服务证书的关键值时，说明客户端访问的不是套餐内的网站，从而流量管理模块735可以直接对客户端的访问流量进行及时的流量统计，并将统计的流量上报给计费服务器进行计费。在这种应用场景下，当查找模块715没有在网站关键值表中查找到与该访问的网站的服务证书的关键值相同的网站的服务证书的关键值时，并不需要再触发判断模块730和识别模块720来分析网站的服务证书获得用户访问的网站的域名，而可以直接对客户端的流量进行控制管理。本发明实施例提供的这种DPI设备当应用于计费管理时，该DPI设备可以通过比较服务器证书的关键值来获得客户端访问的网站名称，识别效率高，且操作简单，从而能对客户端的访问流量进行及时的控制管理，提高运营商的服务质量。

图8为本发明实施例提供的一种网络系统的结构示意图，如图8所示，所述网络系统包括：客户端100、深度包检测DPI设备110以及至少一个网站服务器120，所述DPI设备直路部署于客户端100与网络服务器120的通信链路中，其中，：

所述客户端100，用于向所述至少一个网站服务器120发送基于HTTPS协议的访问请求，与所述至少一个网站服务器120进行密钥协商；

所述深度包检测DPI设备110，用于当客户端通过超文本安全传送协议HTTPS访问网站时，获取所述客户端和所述网站的服务器在密钥协商过程中的认证消息，通过解析所述认证消息获得所述客户端访问的网站的服务器证书，根据预设的算

法获得所述服务器证书的关键值,并根据所述服务器证书的关键值在网站关键值表中进行查找,所述网站关键值表中记录有网站名称及网站的服务器证书的关键值,其中所述网站关键值表中的网站的服务器证书的关键值是根据所述预设的算法获得的,如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称,则根据所述网站名称识别所述客户端访问的网站;

所述网站服务器120,用于接收客户端100发送的网络访问请求,并根据所述网络访问请求与所述客户端进行密钥协商。

在另一种情况下,所述DPI设备110,还用于如果在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称,则判断所述服务器证书的格式类型,当所述服务器证书的格式类型为X.509格式时,在所述服务器证书的使用者属性中获得所述网站的域名,并根据所述网站的域名识别所述客户端访问的网站;当所述服务器证书的格式类型为PGP格式时,在所述服务器证书的用户标识属性中获得所述网站的域名,并根据所述网站的域名识别所述客户端访问的网站。

在又一种情况下,所述DPI设备110,还用于根据识别出的所述客户端访问的网站名称对所述客户端访问的数据流进行管理。

具体的,本发明实施例中的DPI设备110的详细描述可以参见前述相关实施例,在此不再赘述。

本发明实施例中的网络系统,在客户端通过HTTPS协议访问网站时,能够通过获取客户端访问的网站的服务器证书,根据预设的算法获得服务器证书的关键值,并将获得的关键值在网站关键值表中进行查找,从而通过在该网站关键值表中查找到的与所述服务器证书的关键值对应的网站名称来确定用户访问的网站名称。从而在客户端通过HTTPS协议访问网站时也能识别出用户访问的网站名称。

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。

所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的

设备和模块的具体工作过程，可以参考前述方法实施例中的对应过程描述，在此不再赘述。

在本申请所提供的几个实施例中，应该理解到，所揭露的设备和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个模块或组件可以结合或者可以集成到另一个设备中，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口，装置或模块的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块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，或者也可以分布到多个网络单元上。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，在本发明各个实施例中的各功能模块可以集成在一个处理模块中，也可以是各个模块单独物理存在，也可以两个或两个以上模块集成在一个模块中。最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各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：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部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范围。

权 利 要 求

1、一种网站识别方法，其特征在于，包括：

当客户端通过超文本安全传送协议HTTPS访问网站时，获取所述客户端和所述网站的服务器在密钥协商过程中的认证消息；

通过解析所述认证消息获得所述客户端访问的网站的服务器证书；

根据预设的算法获得所述服务器证书的关键值；

根据所述服务器证书的关键值在网站关键值表中进行查找，所述网站关键值表中记录有网站名称及网站的服务器证书的关键值，其中所述网站关键值表中的网站的服务器证书的关键值是根据所述预设的算法获得的；

如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称，则根据所述网站名称识别所述客户端访问的网站。

2、根据权利要求1所述的网站识别方法，其特征在于，还包括：

如果在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称，则判断所述服务器证书的格式类型；

如果所述服务器证书的格式类型为X.509格式，则在所述服务器证书的使用者属性中获得所述网站的域名，并根据所述网站的域名识别所述客户端访问的网站；

如果所述服务器证书的格式类型为PGP格式，则在所述服务器证书的用户标识属性中获得所述网站的域名，并根据所述网站的域名识别所述客户端访问的网站。

3、根据权利要求1所述的网站识别方法，其特征在于，还包括：

如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称，则不对所述客户端的流量进行计费统计；

如果在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称，则对所述客户端的流量进行计费统计。

4、根据权利要求1-3任意一项所述的网站识别方法，其特征在于，所述获取客户端和所述网站的服务器在密钥协商过程中的认证消息包括：

监听所述客户端和所述网站的服务器在密钥协商过程中的报文；

根据所述报文中的HandshakeType字段识别所述认证消息。

5、根据权利要求1-4任意一项所述的网站识别方法，其特征在于：所述预设的算法包括MD5算法或SHA-1算法。

6、一种深度包检测设备，其特征在于，包括：

获取模块，用于当客户端通过超文本安全传送协议HTTPS访问网站时，获取所述客户端和所述网站的服务器在密钥协商过程中的认证消息；

解析模块，用于通过解析所述获取模块获取的所述认证消息获得所述客户端访问的网站的服务器证书；

计算模块，用于根据预设的算法获得所述服务器证书的关键值；

查找模块，用于根据所述计算模块获得的所述服务器证书的关键值在网站关键值表中进行查找，所述网站关键值表中记录有网站名称及网站的服务器证书的关键值，其中所述网站关键值表中的网站的服务器证书的关键值是根据所述预设的算法获得的；

识别模块，如果所述查找模块在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称，则根据所述网站名称识别所述客户端访问的网站。

7、根据权利要求6所述的深度包检测设备，其特征在于，还包括：

设置模块，用于接收并存储管理服务器发送的所述网站关键值表，其中所述网站关键值表中包含有至少一个网站的网站名称以及网站的服务器证书的关键值。

8、根据权利要求6或7所述的深度包检测设备，其特征在于，还包括：

判断模块，用于当所述查找模块在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称时，判断所述解析模块获得的所述服务器证书的格式类型，并触发所述识别模块；

所述识别模块，还用于当所述判断模块判断所述网站的服务器证书的格式类型为X.509格式时，在所述证书的使用者属性中获得所述网站的域名，并根据所述网站的域名识别所述客户端访问的网站；当所述判断模块判断所述网站的服务器证书的格式类型为PGP格式时，在所述证书的用户标识属性中获得所述网站的域名，并根据所述网站的域名识别所述客户端访问的网站。

9、根据权利要求6或7所述的深度包检测设备，其特征在于，还包括：

流量管理模块，用于当所述查找模块在所述网站关键值表中查找到与所述服

务器证书的关键值对应的网站名称时,不对所述客户端的流量进行计费统计;当所述查找模块在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称时,对所述客户端的流量进行计费统计。

10、根据权利要求6-9任意一项所述的深度包检测设备,其特征在于:

所述获取模块具体用于监听所述客户端和所述网站的服务器在密钥协商过程中的报文,并根据所述报文中的HandshakeType字段识别所述认证消息。

11、根据权利要求6-10任意一项所述的深度包检测设备,其特征在于:

所述预设的算法包括MD5算法或SHA-1算法。

12、一种深度包检测设备,其特征在于:

所述深度包检测设备包括处理器、通信接口、存储器和通信总线;

其中,所述处理器和所述通信接口通过所述通信总线进行通信;

所述通信接口,用于与进行通信;

所述存储器用于存储程序;

所述处理器用于执行所述程序,以实现

当客户端通过超文本安全传送协议HTTPS访问网站时,获取所述客户端和所述网站的服务器在密钥协商过程中的认证消息;

通过解析所述认证消息获得所述客户端访问的网站的服务器证书;

根据预设的算法获得所述服务器证书的关键值;

根据所述服务器证书的关键值在网站关键值表中进行查找,所述网站关键值表中记录有网站名称及网站的服务器证书的关键值,其中所述网站关键值表中的网站的服务器证书的关键值是根据所述预设的算法获得的;

如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称,则根据所述网站名称识别所述客户端访问的网站。

13、根据权利要求12所述的深度包检测设备,其特征在于:

所述处理器,还用于当在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称时,判断所述网站的服务器证书的格式类型;

如果所述网站的服务器证书的格式类型为X.509格式,则在所述服务器证书的使用者属性中获得所述网站的域名,并根据所述网站的域名识别所述客户端访问的网站;

如果所述网站的服务器证书的格式类型为PGP格式，则在所述服务器证书的用户标识属性中获得所述网站的域名，并根据所述网站的域名识别所述客户端访问的网站。

14、根据权利要求12或13所述的深度包检测设备，其特征在于：

所述处理器，还用于如果在所述网站关键值表中查找到与所述服务器证书的关键值对应的网站名称，则不对所述客户端的流量进行计费统计；如果在所述网站关键值表中没有查找到与所述服务器证书的关键值对应的网站名称，则对所述客户端的流量进行计费统计。

15、根据权利要求12-14任意一项所述的深度包检测设备，其特征在于，所述处理器具体用于监听所述客户端和所述服务器在密钥协商过程中的报文，并根据所述报文中的HandshakeType字段识别所述认证消息。

16、根据权利要求12-15任意一项所述的深度包检测设备，其特征在于：所述预设的算法包括MD5算法或SHA-1算法。

17、一种网络系统，其特征在于，包括客户端、至少一个网站服务器以及如权利要求6-11任意一项所述的深度包检测设备，其中：

所述客户端，用于向所述至少一个网站服务器发送基于HTTPS协议的访问请求，与所述至少一个网站服务器进行密钥协商；

所述网站服务器，用于接收所述客户端发送的网络访问请求，并根据所述网络访问请求与所述客户端进行密钥协商。

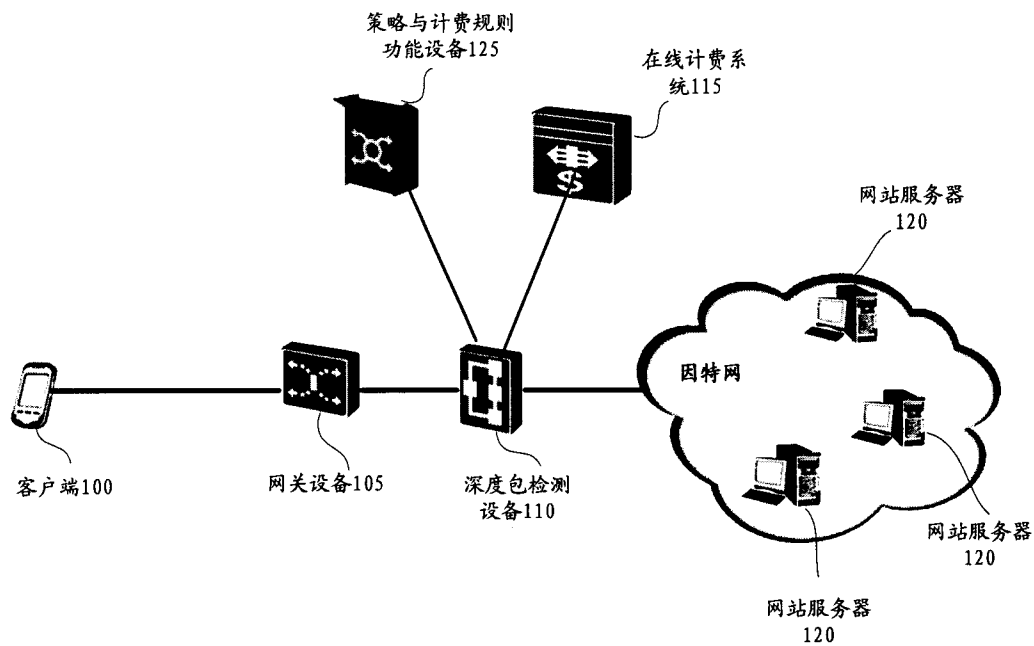


图 1

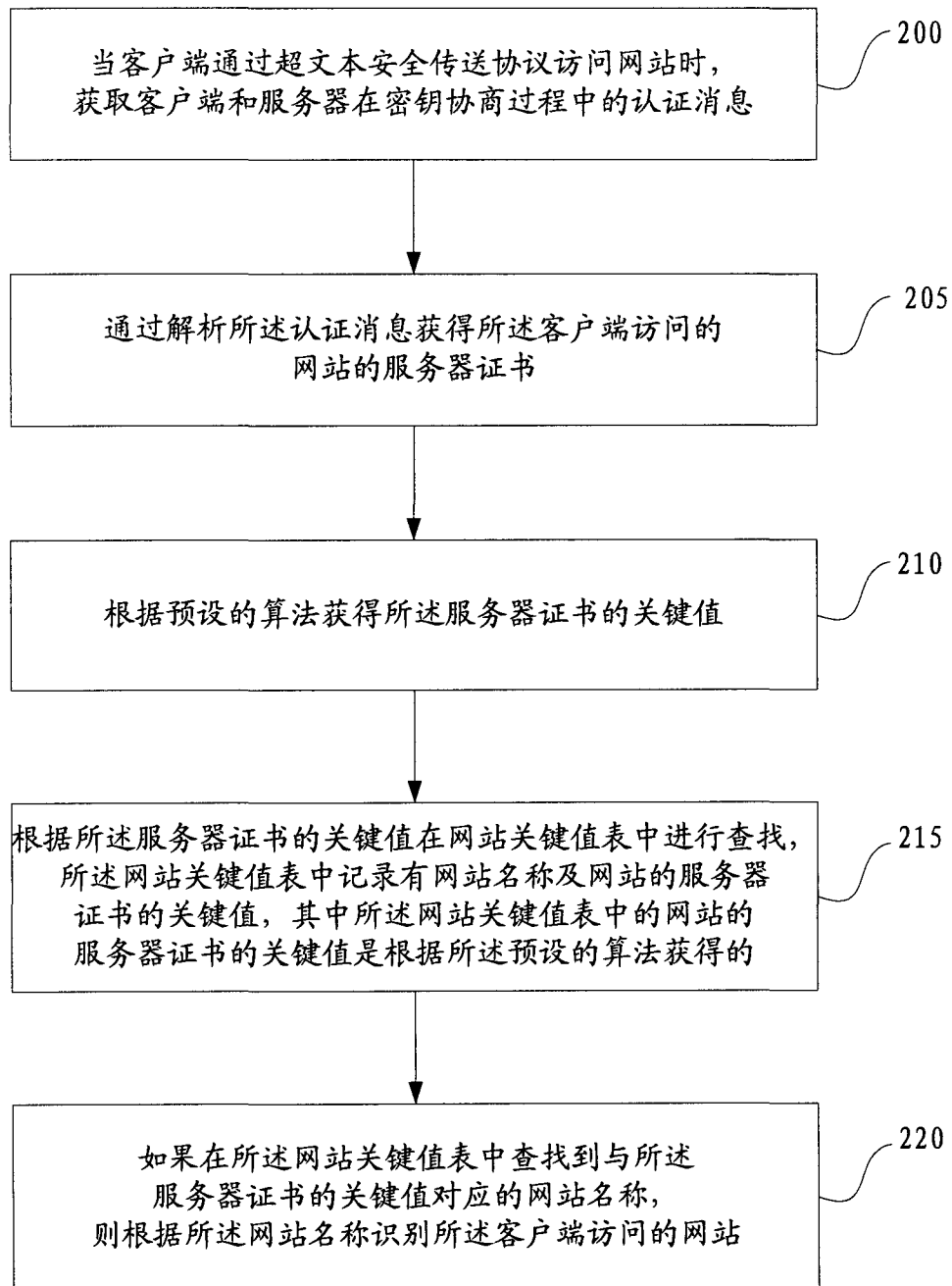


图 2

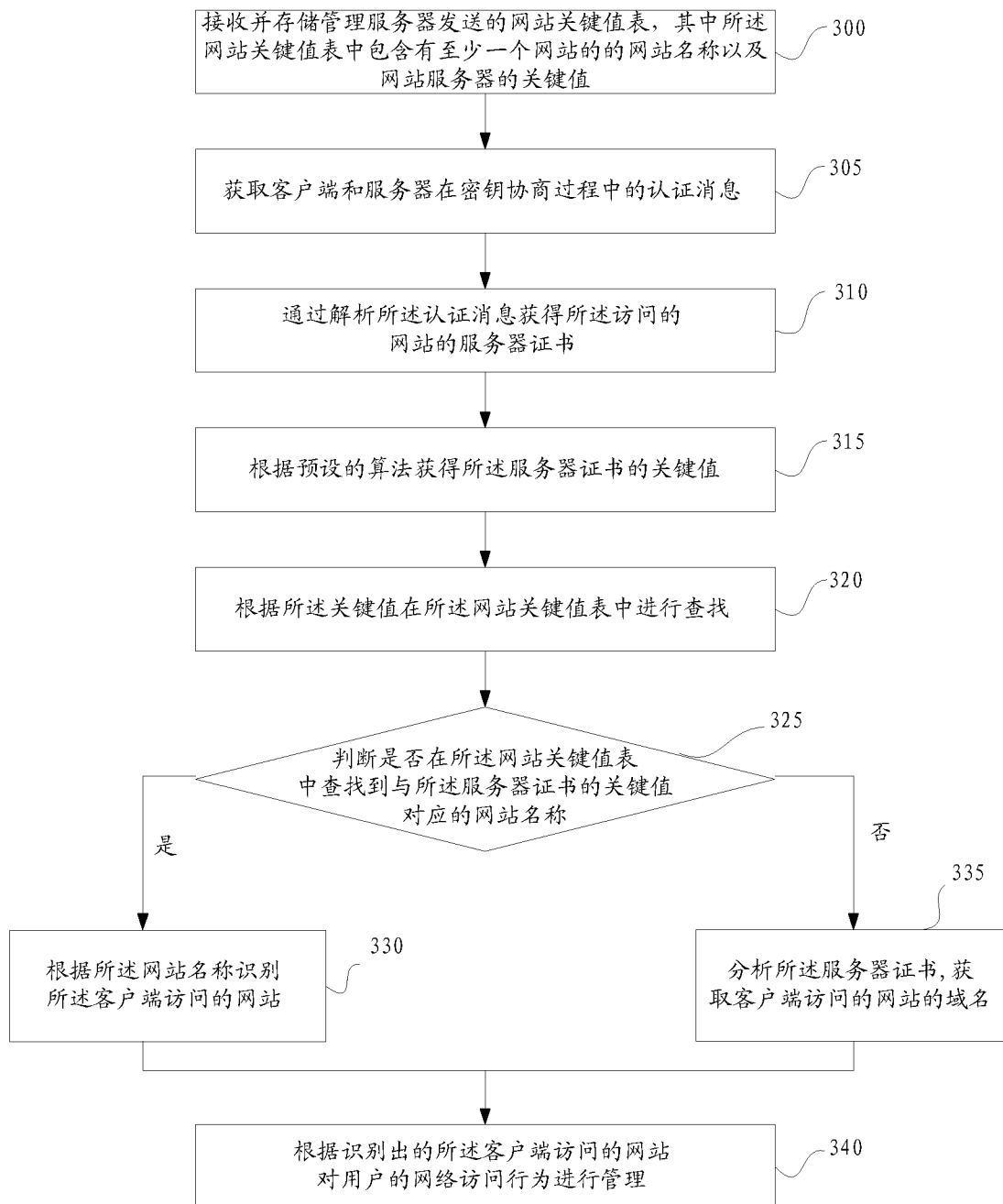


图 3

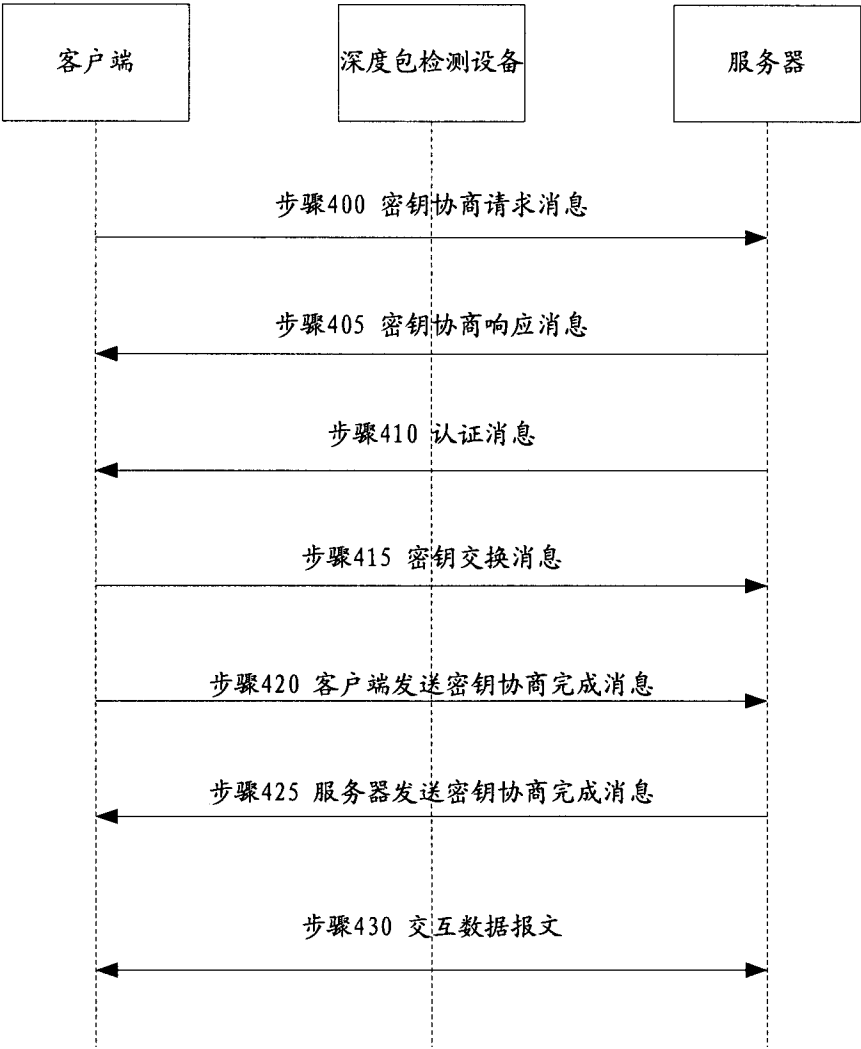


图 4

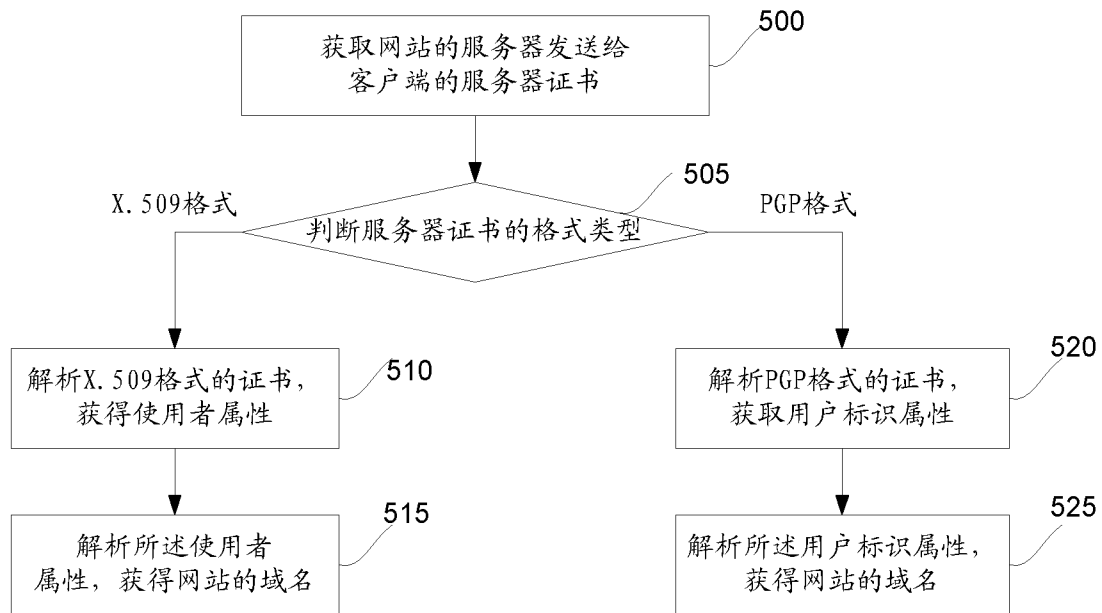


图 5

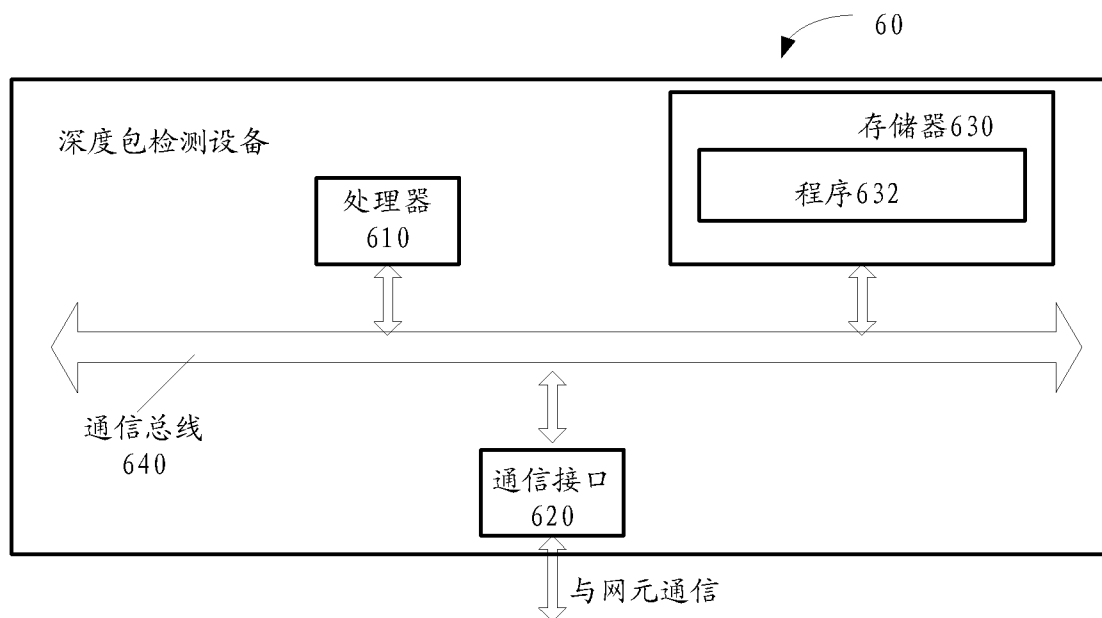


图 6

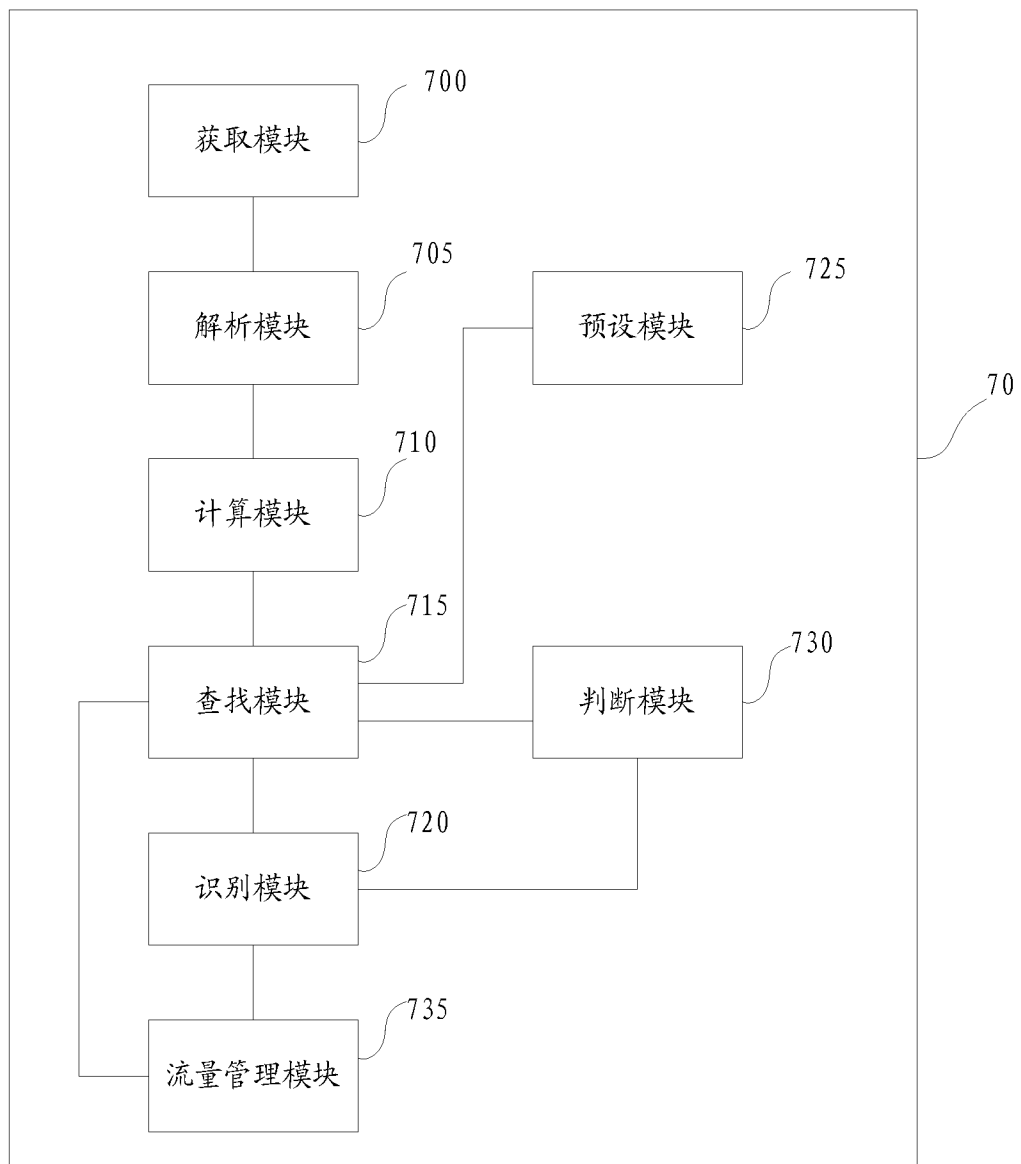


图 7

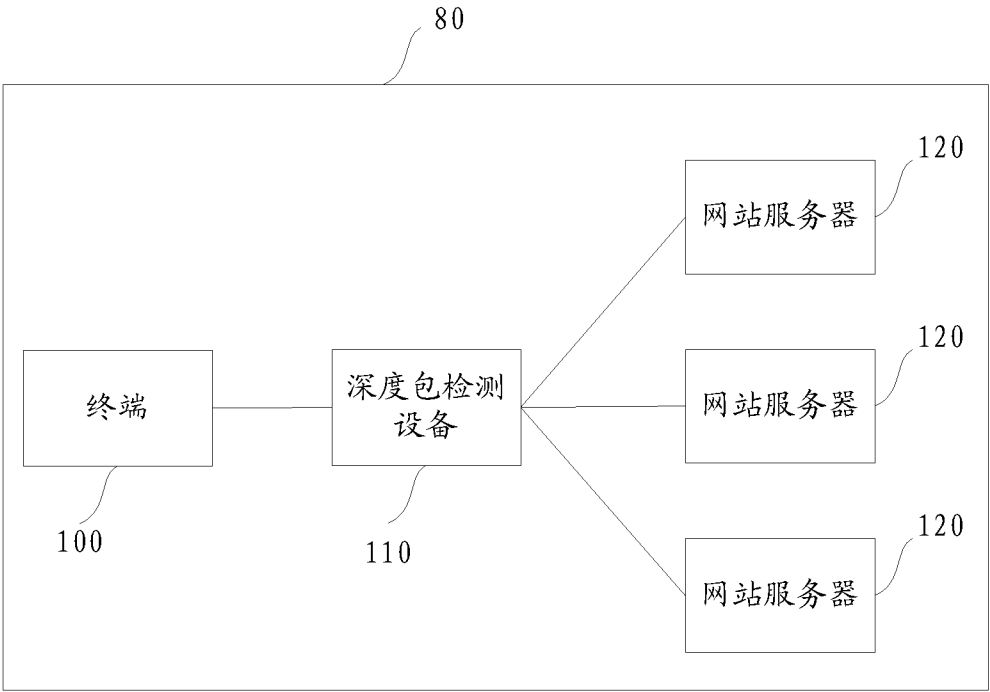


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2012/087848**A. CLASSIFICATION OF SUBJECT MATTER**

H04L 29/06 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CPRSABS, CNTXT, CNKI: HTTPS, SSL, TLS, web site, site, server, certificate, key-value, hash, digital abstract, domain name, checked value

VEN, GOOGLE: HTTPS, SSL, TLS, server, certificate, hash, abstract, domain name, web site

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101977235 A (BEIJING VRV SOFTWARE CO., LTD.), 16 February 2011 (16.02.2011), the whole document	1-17
A	CN 101026599 A (SANGFOR INC.), 29 August 2007 (29.08.2007), the whole document	1-17
A	US 6874084 B1 (INTERNATIONAL BUSINESS MACHINES CORPORATION), 29 March 2005 (29.03.2005), the whole document	1-17

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 27 September 2013 (27.09.2013)	Date of mailing of the international search report 17 October 2013 (17.10.2013)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer XIA, Caijie Telephone No.: (86-10) 62411284

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2012/087848

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101977235 A	16.02.2011	CN 101977235 B	27.03.2013
CN 101026599 A	29.08.2007	CN 100553242 C	21.10.2009
US 6874084 B1	29.03.2005	None	

国际检索报告

国际申请号

PCT/CN2012/087848

A. 主题的分类

H04L 29/06 (2006.01) i

按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04L G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CPRSABS,CNTEXT,CNKI: HTTPS, SSL, TLS, 网站, 站点, 服务器, 证书, 关键值, 哈希, 散列, 数字摘要, 域名, 校验值

VEN,GOOGLE: HTTPS, SSL, TLS, server, certificate, hash, abstract, domain name, web site

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 101977235 A (北京北信源软件股份有限公司) 16.2 月 2011 (16.02.2011), 全文	1-17
A	CN 101026599 A (深圳市深信服电子科技有限公司) 29.8 月 2007 (29.08.2007), 全文	1-17
A	US 6874084 B1 (INTERNATIONAL BUSINESS MACHINES CORPORATION), 29.3 月 2005 (29.03.2005), 全文	1-17

☐ 其余文件在 C 栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

27.9 月 2013 (27.09.2013)

国际检索报告邮寄日期

17.10 月 2013 (17.10.2013)

ISA/CN 的名称和邮寄地址:

中华人民共和国国家知识产权局

中国北京市海淀区蓟门桥西土城路 6 号 100088

传真号: (86-10)62019451

授权官员

夏彩杰

电话号码: (86-10) **62411284**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2012/087848

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN 101977235 A	16.02.2011	CN 101977235 B	27.03.2013
CN 101026599 A	29.08.2007	CN 100553242 C	21.10.2009
US 6874084 B1	29.03.2005	无	