(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 106295366 B
(45)授权公告日 2020.11.24

(21)申请号 201610671817.0
(22)申请日 2016.08.15
(65)同一申请的已公布的文献号
申请公布号 CN 106295366 A
(43)申请公布日 2017.01.04
(73)专利权人 北京奇虎科技有限公司
地址 100088 北京市西城区新街口外大街28号D座112室(德胜园区)
专利权人 北京奇安信科技有限公司
(72)发明人 刘敬良 黄凌志
(74)专利代理机构 北京律诚同业知识产权代理有限公司 11006
代理人 王玉双
(51)Int.Cl.
G06F 21/60(2013.01)
H04L 29/08(2006.01)
(56)对比文件
CN 105468990 A,2016.04.06
US 2014195798 A1,2014.07.10
US 8776249 B1,2014.07.08
审查员 李思彤

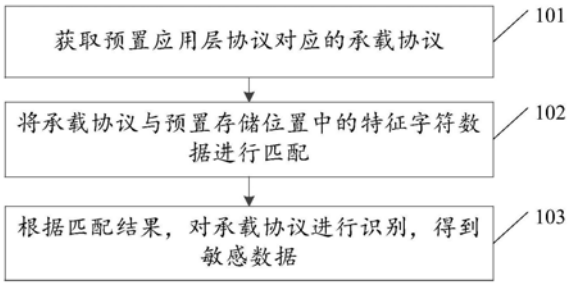
权利要求书2页 说明书10页 附图3页

(54)发明名称

敏感数据识别方法及装置

(57)摘要

本发明公开了一种敏感数据识别方法及装置,涉及信息技术领域,主要目的在于能够提升敏感数据的精度以及能够提升敏感数据的安全性。所述方法包括:获取预置应用层协议对应的承载协议;将所述承载协议与预置存储位置中的特征字符数据进行匹配,所述预置存储位置保存有不同的特征字符数据;根据匹配结果,对所述承载协议进行识别,得到敏感数据。本发明适用于敏感数据的识别。



CN 106295366 B

1.一种敏感数据识别方法,其特征在于,包括:

获取预置应用层协议对应的承载协议;

将所述承载协议与预置存储位置中的特征字符数据进行匹配,所述预置存储位置保存有不同的特征字符数据;

根据匹配结果,对所述承载协议进行识别,得到敏感数据;

根据所述敏感数据对应的数据类型配置与所述敏感数据对应的加密策略,其中,所述数据类型为动态数据和静态数据;

根据所述加密策略对所述敏感数据进行加密处理,得到加密文件,其中,当所述敏感数据对应的数据类型为静态数据,配置密钥管理加密策略对敏感数据进行加密,当所述敏感数据对应的数据类型为动态数据,配置通过替换数据的方式对敏感数据进行加密的加密策略;

所述预置存储位置中还保存有不同的交互操作类型,所述将所述承载协议与预置存储位置中的特征字符数据进行匹配包括:

将所述承载协议与所述预置存储位置中的交互操作类型进行匹配;

根据交互操作类型匹配结果,对所述承载协议进行识别,得到所述承载协议与所述预置应用层协议之间的交互操作数据,包括:

从承载协议中提取与预存储位置中的交互操作类型匹配成功的数据,得到所述承载协议与所述预置应用层协议之间的交互操作数据;

将所述交互操作数据与预置存储位置中的特征字符数据进行匹配。

2.根据权利要求1所述的方法,其特征在于,所述获取预置应用层协议对应的承载协议包括:

根据预置代理网关设备获取预置应用层协议对应的承载协议。

3.根据权利要求1所述的方法,其特征在于,所述将所述承载协议与预置存储位置中的特征字符数据进行匹配包括:

通过预置多模式匹配算法将所述承载协议与预置存储位置中的特征字符数据进行匹配。

4.根据权利要求1所述的方法,其特征在于,所述根据匹配结果,对所述承载协议进行识别,得到敏感数据包括:

从所述承载协议中提取与预置存储位置中的特征字符数据匹配成功的数据,得到敏感数据。

5.根据权利要求1-4任一项所述的方法,其特征在于,所述预置应用层协议为超文本传输协议http,所述预置应用层协议对应的承载协议为超文本传输安全协议https。

6.一种敏感数据识别装置,其特征在于,包括:

获取单元,用于获取预置应用层协议对应的承载协议;

匹配单元,用于将所述获取单元获取的承载协议与预置存储位置中的特征字符数据进行匹配,所述预置存储位置保存有不同的特征字符数据;

识别单元,用于根据所述匹配单元的匹配结果,对所述承载协议进行识别,得到敏感数据;

配置单元,用于根据所述识别单元得到的所述敏感数据对应的数据类型配置与所述敏

感数据对应的加密策略，其中，所述数据类型为动态数据和静态数据；

加密单元，用于根据所述配置单元配置的所述加密策略对所述敏感数据进行加密处理，得到加密文件，其中，当所述敏感数据对应的数据类型为静态数据，配置密钥管理加密策略对敏感数据进行加密，当所述敏感数据对应的数据类型为动态数据，配置通过替换数据的方式对敏感数据进行加密的加密策略；

所述匹配单元包括：

匹配模块，用于将所述承载协议与所述预置存储位置中的交互操作类型进行匹配；

识别模块，用于根据交互操作类型匹配结果，对所述承载协议进行识别，得到所述承载协议与所述预置应用层协议之间的交互操作数据，包括：从承载协议中提取与预存储位置中的交互操作类型匹配成功的数据，得到所述承载协议与所述预置应用层协议之间的交互操作数据；

所述匹配模块，还用于将所述交互操作数据与预置存储位置中的特征字符数据进行匹配。

7.根据权利要求6所述的装置，其特征在于，

所述获取单元，具体用于根据预置代理网关设备获取预置应用层协议对应的承载协议。

8.根据权利要求6所述的装置，其特征在于，

所述匹配单元，具体用于通过预置多模式匹配算法将所述获取单元获取的所述承载协议与预置存储位置中的特征字符数据进行匹配。

9.根据权利要求6所述的装置，其特征在于，

所述识别单元，具体用于从所述承载协议中提取与预置存储位置中的特征字符数据匹配成功的数据，得到敏感数据。

10.根据权利要求6-9任一项所述的装置，其特征在于，所述预置应用层协议为超文本传输协议http，所述预置应用层协议对应的承载协议为超文本传输安全协议https。

敏感数据识别方法及装置

技术领域

[0001] 本发明涉及信息技术领域，特别是涉及一种敏感数据识别方法及装置。

背景技术

[0002] 随着信息技术的不断发展，云存储服务随之出现，云存储服务是由互联网公司推出的在线存储服务，向用户提供互联网的应用数据的存储、访问、备份、共享等数据管理功能。基于互联网的应用也越来越多，一些基于互联网的应用会涉及到用户的敏感数据，该敏感数据可以为用户账户的密码、用户的身份证号码和姓名等数据，用户的敏感数据通常不希望被其他用户观看到或者窃取到。为了保证敏感数据的安全性，通常需要对敏感数据进行加密处理。

[0003] 目前，在进行敏感数据识别时，通常对基于网络层获取的应用数据进行识别得到敏感数据。然而，在后续应用中用户还会输入敏感数据，基于网络层获取的应用数据包含的敏感数据不全面，若对基于网络层获取的应用数据进行识别得到敏感数据，会造成敏感数据的精度较低，且会造成无法对部分敏感数据进行加密处理，从而导致敏感数据的安全性较低。

发明内容

[0004] 有鉴于此，本发明提供一种敏感数据识别方法及装置，主要目的在于能够提升敏感数据的精度以及能够提升敏感数据的安全性。

[0005] 依据本发明一个方面，提供了一种敏感数据识别方法，包括：

[0006] 获取预置应用层协议对应的承载协议；

[0007] 将所述承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；

[0008] 根据匹配结果，对所述承载协议进行识别，得到敏感数据。

[0009] 依据本发明另一个方面，提供了一种敏感数据识别装置，包括：

[0010] 获取单元，用于获取预置应用层协议对应的承载协议；

[0011] 匹配单元，用于将所述获取单元获取的承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；

[0012] 识别单元，用于根据所述匹配单元的匹配结果，对所述承载协议进行识别，得到敏感数据。

[0013] 借由上述技术方案，本发明实施例提供的技术方案至少具有下列优点：

[0014] 本发明实施例提供一种敏感数据识别方法及装置。首先获取预置应用层协议对应的承载协议；然后将所述承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；根据匹配结果，对所述承载协议进行识别，得到敏感数据。与现有技术中在进行敏感数据识别时，通常对基于网络层获取的应用数据进行识别得到敏感数据相比，本发明实施例通过获取预置应用层协议对应的承载协议；然后根据承

载协议与预置存储位置中的特征字符数据的匹配结果，对承载协议进行识别得到敏感数据，能够保证识别到全部敏感数据，从而能够提升敏感数据的精度，且能够保证对全部敏感数据进行加密处理，进而能够提升敏感数据的安全性。

[0015] 上述说明仅是本发明技术方案的概述，为了能够更清楚了解本发明的技术手段，而可依照说明书的内容予以实施，并且为了让本发明的上述和其它目的、特征和优点能够更明显易懂，以下特举本发明的具体实施方式。

附图说明

[0016] 通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本发明的限制。而且在整个附图中，用相同的参考符号表示相同的部件。在附图中：

[0017] 图1示出了本发明实施例提供的一种敏感数据识别方法流程图；

[0018] 图2示出了本发明实施例提供的另一种敏感数据识别方法流程图；

[0019] 图3示出了本发明实施例提供的一种对交互操作进行识别的示意图；

[0020] 图4示出了本发明实施例提供的另一种对交互操作进行识别的示意图；

[0021] 图5示出了本发明实施例提供的一种敏感数据识别装置结构示意图；

[0022] 图6示出了本发明实施例提供的另一种敏感数据识别装置结构示意图。

具体实施方式

[0023] 下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例，然而应当理解，可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反，提供这些实施例是为了能够更透彻地理解本公开，并且能够将本公开的范围完整的传达给本领域的技术人员。

[0024] 本发明实施例提供了一种敏感数据识别方法，如图1所示，所述方法包括：

[0025] 101、获取预置应用层协议对应的承载协议。

[0026] 其中，预置应用层协议可以为超文本传输协议(Hyper Text Transfer Protocol，HTTP)，文件传输协议(File Transfer Protocol，FTP)等；所述承载协议可以为超文本传输安全协议(Hyper Text Transfer Protocol over Secure Socket Layer，HTTPS)等。

[0027] 对于本发明实施例，由于承载协议通常承载着涉及敏感数据的交互操作，通过获取预置应用层协议对应的承载协议，能够保证识别到全部敏感数据，从而能够保证对全部敏感数据进行加密处理，进而能够提升敏感数据的安全性。

[0028] 102、将承载协议与预置存储位置中的特征字符数据进行匹配。

[0029] 其中，所述预置存储位置保存有不同的特征字符数据。所述特征字符数据可以为涉及用户隐私的特征字符数据。例如，特征字符数据可以为姓名字符数据、密码字符数据、身份证号码字符数据等。

[0030] 103、根据匹配结果，对承载协议进行识别，得到敏感数据。

[0031] 其中，在整个匹配过程中，只要预置存储位置中存在与承载协议匹配成功的特征字符数据，就返回匹配结果；然后从承载协议提取与承载协议匹配成功的特征字符数据作为敏感数据。所述敏感数据为对于用户或者企业具有重要意义的关键数据。例如，敏感数据

可以为用户账号的密码、用户的身份证号码和姓名、联系方式、银行卡号等，也可以为企业中涉及商业机密的重要数据等。具体地，敏感数据可以以文本的形式存在。

[0032] 在本发明实施例中，通过对承载协议进行识别，得到敏感数据，能够实现只对敏感数据进行加密，无需对应用数据中的非敏感数据进行加密，能够实现显示非敏感数据，从而能够实现非敏感数据被用户观看到。

[0033] 对于本发明实施例，所述方法还包括：对敏感数据进行加密处理。

[0034] 在本发明实施例中，可以通过密钥管理加密策略和标记化替代加密策略，对敏感数据进行加密。所述密钥管理是现有云服务提供商可以提供的基于加密密钥方案来保护用户的数据，具体可以包括保护密钥存储，使得数据在存储、传输和备份中都受到保护，还可以包括访问密钥存储，限制只有特定需要单独密钥的实体可以访问密钥存储，还可以包括密钥的备份和恢复，以便更好的保护数据。所述标记化替代是通过将敏感数据采用标记化字符进行替代从而实现加密处理，避免敏感数据被泄露的风险，本发明实施例对这里的标记化替代字符的形式不做限定，具体可以根据实际需求进行选取。

[0035] 需要说明的是，对于发明实施例，可以将加密密钥保存在本地客户端，以便对该敏感数据进行解密，可以实现只有通过本地客户端才能对该敏感数据进行解密，通过其他客户端无法解密该敏感数据，进而保证了该敏感数据的安全性。

[0036] 本发明实施例提供的一种敏感数据识别方法。首先获取预置应用层协议对应的承载协议；然后将所述承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；根据匹配结果，对所述承载协议进行识别，得到敏感数据。与现有技术中在进行敏感数据识别时，通常对基于网络层获取的应用数据进行识别得到敏感数据相比，本发明实施例通过获取预置应用层协议对应的承载协议；然后根据承载协议与预置存储位置中的特征字符数据的匹配结果，对承载协议进行识别得到敏感数据，能够保证识别到全部敏感数据，从而能够提升敏感数据的精度，且能够保证对全部敏感数据进行加密处理，进而能够提升敏感数据的安全性。

[0037] 本发明实施例提供了另一种敏感数据识别方法，如图2所示，所述方法包括：

[0038] 201、获取预置应用层协议对应的承载协议。

[0039] 对于本发明实施例，步骤201具体还可以为：根据预置代理网关设备获取预置应用层协议对应的承载协议。

[0040] 其中，预置代理网关设备为配置在客户端的一个提供登录安全服务器网关的设备，是代理服务器的一种，它能够根据用户携带的域名登陆自己的代理网关，进而建立数据连接，将待处理的数据的上传至代理网关，根据代理网关实现数据的转发，从而上传至云服务器，以便云服务器为用户构建更稳定、更安全的应用，然后根据承载协议中的特征字符数据识别出承载协议中的敏感数据，进而获取敏感数据。其中，安全服务器为一个公网服务器中的一个或多个服务器，每个安全服务器用于向云服务器转发固定区域内的客户端发送的数据处理请求，或者向客户端转发云服务器返回的数据处理响应，并且每个安全服务器都拥有自己的域名及IP地址，安全服务器之间互相为主备关系。通过预置代理网关设备获取预置应用层协议对应的承载协议，能够实现将经过加密的敏感数据通过安全服务器上传到云服务器中，从而保证敏感数据的安全性。

[0041] 202、将承载协议与预置存储位置中的特征字符数据进行匹配。

[0042] 其中，所述预置存储位置保存有不同的特征字符数据。

[0043] 对于本发明实施例，当所述预置存储位置中还保存有不同的交互操作类型时，步骤201具体可以为：将所述承载协议与所述预置存储位置中的交互操作类型进行匹配；根据交互操作类型匹配结果，对所述承载协议进行识别，得到所述承载协议与所述预置应用层协议之间的交互操作数据；将所述交互操作数据与预置存储位置中的特征字符数据进行匹配。具体地，所述根据交互操作类型匹配结果，对所述承载协议进行识别，得到所述承载协议与所述预置应用层协议之间的交互操作数据可以包括：从承载协议中提取与预置存储位置中的交互操作类型匹配成功的数据，得到所述承载协议与所述预置应用层协议之间的交互操作数据。

[0044] 其中，交互操作可以为预置应用层协议与承载协议之间的交互操作。所述交互操作类型可以为上传、删除、复制、移动、回收、共享等操作。需要说明的是，执行交互操作识别的主体可以为深度应用操作识别设备，该深度应用操作识别设备基于应用层对交互操作的识别。

[0045] 对于本发明实施例，通过识别预置应用层协议与承载协议之间的交互操作，可以实现获知应用在应用层内的具体内容，实现了应用访问的更精细化控制，进一步加强了应用的安全防护力度。

[0046] 进一步地，为了提升匹配速度以及获取到敏感数据的速度，步骤202具体可以为：通过预置多模式匹配算法将所述承载协议与预置存储位置中的特征字符数据进行匹配。其中，通过预置多模式匹配算法是通过构造字典树的方式进行匹配的，整个匹配过程直接按照字典树的顺序匹配，根本不需要回溯字符串，减小了匹配过程的复杂度，从而提升了匹配速度进而提升了获取到敏感数据的速度。

[0047] 203、根据匹配结果，对所述承载协议进行识别，得到敏感数据。

[0048] 对于本发明实施例，步骤203具体可以为：从所述承载协议中提取与预置存储位置中的特征字符数据匹配成功的数据，得到敏感数据。

[0049] 204、根据所述敏感数据对应的数据类型配置与所述敏感数据对应的加密策略。

[0050] 其中，所述数据类型可以分为动态数据和静态数据，静态数据可以为当用户在查看数据时已生成，并没有与服务器数据库进行交互的数据，主要指硬盘、存储空间中的数据等，动态数据可以为在系统应用中随时间变化而改变的数据，与服务器数据库有交互的数据，如用户访问的数据、流量数据等。所述加密策略可以为通过预先配置的加密算法对敏感数据进行加密，也可以通过替换数据的方式对敏感数据进行加密。

[0051] 例如，当所述敏感数据对应的数据类型为静态数据，可以配置密钥管理加密策略对敏感数据进行加密。由于不同应用场景下的操作产生的数据类型有所不同，本发明实施例敏感数据的数据类型的不同采取不同的加密方式，对于磁盘上静态数据或者生产数据库中的静态数据的加密尤为重要，因为这样可以用来防止恶意的云服务提供商、恶意的邻居“租户”及某些类型应用的滥用。这些用户控制并保存密钥，在自己需要的情况下解密数据。

[0052] 需要说明的是，由于静态数据的特点是由系统分配固定大小的存储空间，在传输过程中，存储空间和容量都不会发生改变，因此静态数据相对比较稳定，由于密钥管理的加密策略对当前加密的敏感数据配置有相应的解密密钥，因此对于稳定性较高的静态数据采用的密钥管理的加密策略，保证在数据处理过程中无需经常对静态数据进行解密。

[0053] 再例如，当所述敏感数据对应的数据类型为动态数据，可以配置通过替换数据的方式对敏感数据进行加密的加密策略。对于加密传输中的动态数据，如信用卡号、密码和私钥等，虽然云提供商网络可能比开放网络安全，但是他们使用其特有的、由许多不同的组成部分构成的架构，且由不同的组织共享云。因此，即便实在云提供商的网络中，保护这些传输中的敏感数据和受监管信息也是非常重要的。

[0054] 需要说明的是，由于动态数据的结构不确定总的数据存储量，而是现有的每一个数据元素定义一个确定的初始大小的空间，若干个数据元素分配若干个同样大小的空间，当数据发生变化时，数据的存储空间也会发生变化，因此静态数据相对不固定，由于标记化替代的加密策略对当前加密的敏感数据的随机性比较，并且不会配置有相应的解密密钥，需要在原始加密处进行标记化解密，才能获取解密文件，因此对于时常变化的动态数据采用的标记化替代的加密策略，更能够保证数据的安全性。

[0055] 205、根据所述加密策略对所述敏感数据进行加密处理，得到加密文件。

[0056] 对于本发明实施例，通过对不同数据类型的敏感数据采用不同的加密策略，提高了加密精度，使得不同数据类型的数据能够被正确的加密，更有效的防止用户的敏感数据泄露，进一步提高了数据在云存储应用的传输过程中的安全性。

[0057] 对于本发明实施例，具体可以应用到如下场景，但不限于此包括：提供了一种对交互操作进行识别的示意图和另一种对交互操作进行识别的示意图，如图3和图4所示，当客户端通过http协议向云存储服务请求上传数据或者请求删除数据时，可以通过深度应用操作识别设备从http协议的数据包中识别http协议承载的应用操作，识别结果为http post upload(上传)，http post delete(删除)，即http协议承载的应用操作为上传数据操作，或者请求删除数据操作。然后，可以判断识别结果是否命中了云盘交互操作，其中，云盘交互操作可以为上传、删除、复制、移动、回收、共享等交互操作。通过图4可以获知，识别结果命中了云盘的交互删除操作，因此，可以识别删除操作数据中敏感数据；然后通过标记化替换的方式对敏感数据进行加密处理，得到加密文件，最后对报文进行重组后通过发送给云服务器。具体地，可以通过安全服务器将加密后的敏感数据发送给云服务器。

[0058] 需要说明的是，在通过安全服务器将加密后的敏感数据发送给云服务器之前，客户端需要登录安全服务器的域名后，通过域名解析系统(Domain Name System，DNS)的域名服务器解析出安全服务器的互联网协议(Internet Protocol，IP)地址，然后根据所述IP地址建立客户端与安全服务器之间的数据连接。

[0059] 此外，当云存储服务通过http协议向客户端发送请求响应http response，时，可以通过深度应用操作识别设备从http协议的中识别http协议承载的请求响应，如上传数据请求响应为或者删除数据请求响应，然后将识别出来的请求响应发送给客户端。

[0060] 本发明实施例提供的另一种敏感数据识别方法。首先获取预置应用层协议对应的承载协议；然后将所述承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；根据匹配结果，对所述承载协议进行识别，得到敏感数据。与现有技术中在进行敏感数据识别时，通常对基于网络层获取的应用数据进行识别得到敏感数据相比，本发明实施例通过获取预置应用层协议对应的承载协议；然后根据承载协议与预置存储位置中的特征字符数据的匹配结果，对承载协议进行识别得到敏感数据，能够保证识别到全部敏感数据，从而能够提升敏感数据的精度，且能够保证对全部敏感数

据进行加密处理，进而能够提升敏感数据的安全性。

[0061] 本发明实施例提供了一种敏感数据识别装置，如图5所示，所述装置包括：获取单元31、匹配单元32和识别单元33。

[0062] 获取单元31，可以用于获取预置应用层协议对应的承载协议。

[0063] 匹配单元32，可以用于将所述获取单元31获取的承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据。

[0064] 识别单元33，可以用于根据所述匹配单元32的匹配结果，对所述承载协议进行识别，得到敏感数据。

[0065] 需要说明的是，本发明实施例提供的一种敏感数据识别装置所涉及各功能单元的其他相应描述，可以参考图1中的对应描述，在此不再赘述。

[0066] 本发明实施例提供的一种敏感数据识别装置。首先获取预置应用层协议对应的承载协议；然后将所述承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；根据匹配结果，对所述承载协议进行识别，得到敏感数据。与现有技术中在进行敏感数据识别时，通常对基于网络层获取的应用数据进行识别得到敏感数据相比，本发明实施例通过获取预置应用层协议对应的承载协议；然后根据承载协议与预置存储位置中的特征字符数据的匹配结果，对承载协议进行识别得到敏感数据，能够保证识别到全部敏感数据，从而能够提升敏感数据的精度，且能够保证对全部敏感数据进行加密处理，进而能够提升敏感数据的安全性。

[0067] 本发明实施例提供了另一种敏感数据识别装置，如图6所示，所述装置包括：获取单元41、匹配单元42和识别单元43。

[0068] 获取单元41，可以用于获取预置应用层协议对应的承载协议。其中，所述预置应用层协议为超文本传输协议http，所述预置应用层协议对应的承载协议为超文本传输安全协议https。

[0069] 匹配单元42，可以用于将所述获取单元41获取的承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据。

[0070] 识别单元43，可以用于根据所述匹配单元42的匹配结果，对所述承载协议进行识别，得到敏感数据。

[0071] 所述匹配单元42包括：匹配模块4201和识别模块4202。

[0072] 匹配模块4201，可以用于将所述承载协议与所述预置存储位置中的交互操作类型进行匹配。

[0073] 识别模块4202，可以用于根据交互操作类型匹配结果，对所述承载协议进行识别，得到所述承载协议与所述预置应用层协议之间的交互操作数据。

[0074] 所述匹配模块4201，还可以用于将所述交互操作数据与预置存储位置中的特征字符数据进行匹配。

[0075] 所述获取单元41，具体可以用于根据预置代理网关设备获取预置应用层协议对应的承载协议。

[0076] 所述匹配单元42，具体可以用于通过预置多模式匹配算法将所述获取单元41获取的所述承载协议与预置存储位置中的特征字符数据进行匹配。

[0077] 所述识别单元43，具体可以用于当所述承载协议与预置存储位置中的特征字符数

据匹配成功时，则将所述获取单元41获取的所述承载协议确定为敏感数据。

[0078] 进一步地，所述装置还包括：加密单元44。

[0079] 加密单元44，可以用于对所述识别单元43得到的所述敏感数据进行加密处理。

[0080] 进一步地，所述装置还包括：配置单元45。

[0081] 所述配置单元45，可以用于根据所述识别单元43得到的所述敏感数据对应的数据类型配置与所述敏感数据对应的加密策略。

[0082] 所述加密单元44，具体可以用于根据所述配置单元45配置的所述加密策略对所述敏感数据进行加密处理，得到加密文件。

[0083] 需要说明的是，本发明实施例提供的另一种敏感数据识别装置所涉及各功能单元的其他相应描述，可以参考图2中的对应描述，在此不再赘述。

[0084] 本发明实施例提供的另一种敏感数据识别装置。首先获取预置应用层协议对应的承载协议；然后将所述承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；根据匹配结果，对所述承载协议进行识别，得到敏感数据。与现有技术中在进行敏感数据识别时，通常对基于网络层获取的应用数据进行识别得到敏感数据相比，本发明实施例通过获取预置应用层协议对应的承载协议；然后根据承载协议与预置存储位置中的特征字符数据的匹配结果，对承载协议进行识别得到敏感数据，能够保证识别到全部敏感数据，从而能够提升敏感数据的精度，且能够保证对全部敏感数据进行加密处理，进而能够提升敏感数据的安全性。

[0085] 在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述的部分，可以参见其他实施例的相关描述。

[0086] 可以理解的是，上述方法及装置中的相关特征可以相互参考。另外，上述实施例中的“第一”、“第二”等是用于区分各实施例，而并不代表各实施例的优劣。

[0087] 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，上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

[0088] 在此提供的算法和显示不与任何特定计算机、虚拟系统或者其它设备固有相关。各种通用系统也可以与基于在此的示教一起使用。根据上面的描述，构造这类系统所要求的结构是显而易见的。此外，本发明也不针对任何特定编程语言。应当明白，可以利用各种编程语言实现在此描述的本发明的内容，并且上面对特定语言所做的描述是为了披露本发明的最佳实施方式。

[0089] 在此处所提供的说明书中，说明了大量具体细节。然而，能够理解，本发明的实施例可以在没有这些具体细节的情况下实践。在一些实例中，并未详细示出公知的方法、结构和技术，以便不模糊对本说明书的理解。

[0090] 类似地，应当理解，为了精简本公开并帮助理解各个发明方面中的一个或多个，在上面对本发明的示例性实施例的描述中，本发明的各个特征有时被一起分组到单个实施例、图、或者对其的描述中。然而，并不应将该公开的方法解释成反映如下意图：即所要求保护的本发明要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说，如下面的权利要求书所反映的那样，发明方面在于少于前面公开的单个实施例的所有特征。因此，遵循具体实施方式的权利要求书由此明确地并入该具体实施方式，其中每个权利要求本身都作为本发明的单独实施例。

[0091] 本领域那些技术人员可以理解，可以对实施例中的设备中的模块进行自适应性地改变并且把它们设置在与该实施例不同的一个或多个设备中。可以把实施例中的模块或单元或组件组合成一个模块或单元或组件，以及此外可以把它们分成多个子模块或子单元或子组件。除了这样的特征和/或过程或者单元中的至少一些是相互排斥之外，可以采用任何组合对本说明书(包括伴随的权利要求、摘要和附图)中公开的所有特征以及如此公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述，本说明书(包括伴随的权利要求、摘要和附图)中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

[0092] 此外，本领域的技术人员能够理解，尽管在此所述的一些实施例包括其它实施例中所包括的某些特征而不是其它特征，但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如，在下面的权利要求书中，所要求保护的实施例的任意之一都可以以任意的组合方式来使用。

[0093] 本发明的各个部件实施例可以以硬件实现，或者以在一个或者多个处理器上运行的软件模块实现，或者以它们的组合实现。本领域的技术人员应当理解，可以在实践中使用微处理器或者数字信号处理器(DSP)来实现根据本发明实施例的敏感数据识别装置中的一些或者全部部件的一些或者全部功能。本发明还可以实现为用于执行这里所描述的方法的一部分或者全部的设备或者装置程序(例如，计算机程序和计算机程序产品)。这样的实现本发明的程序可以存储在计算机可读介质上，或者可以具有一个或者多个信号的形式。这样的信号可以从因特网网站上下载得到，或者在载体信号上提供，或者以任何其他形式提供。

[0094] 应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制，并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中，不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了若干装置的单元权利要求中，这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名称。

[0095] 本发明公开了A1、一种敏感数据识别方法，包括：

[0096] 获取预置应用层协议对应的承载协议；

[0097] 将所述承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；

[0098] 根据匹配结果，对所述承载协议进行识别，得到敏感数据。

[0099] A2、如A1所述的方法，所述预置存储位置中还保存有不同的交互操作类型，所述将所述承载协议与预置存储位置中的特征字符数据进行匹配包括：

[0100] 将所述承载协议与所述预置存储位置中的交互操作类型进行匹配；

[0101] 根据交互操作类型匹配结果，对所述承载协议进行识别，得到所述承载协议与所述预置应用层协议之间的交互操作数据；

[0102] 将所述交互操作数据与预置存储位置中的特征字符数据进行匹配。

[0103] A3、如A1所述的方法，所述获取预置应用层协议对应的承载协议包括：

[0104] 根据预置代理网关设备获取预置应用层协议对应的承载协议。

[0105] A4、如A1所述的方法，所述将所述承载协议与预置存储位置中的特征字符数据进行匹配包括：

[0106] 通过预置多模式匹配算法将所述承载协议与预置存储位置中的特征字符数据进行匹配。

[0107] A5、如A1所述的方法，所述根据匹配结果，对所述承载协议进行识别，得到敏感数据包括：

[0108] 从所述承载协议中提取与预置存储位置中的特征字符数据匹配成功的数据，得到敏感数据。

[0109] A6、如A1所述的方法，所述方法还包括：

[0110] 对所述敏感数据进行加密处理。

[0111] A7、如A6所述的方法，所述方法还包括：

[0112] 根据所述敏感数据对应的数据类型配置与所述敏感数据对应的加密策略；

[0113] 所述对所述敏感数据进行加密处理包括：

[0114] 根据所述加密策略对所述敏感数据进行加密处理，得到加密文件。

[0115] A2、如A1-A7任一项所述的方法，所述预置应用层协议为超文本传输协议http，所述预置应用层协议对应的承载协议为超文本传输安全协议https。

[0116] B9、一种敏感数据识别装置，包括：

[0117] 获取单元，用于获取预置应用层协议对应的承载协议；

[0118] 匹配单元，用于将所述获取单元获取的承载协议与预置存储位置中的特征字符数据进行匹配，所述预置存储位置保存有不同的特征字符数据；

[0119] 识别单元，用于根据所述匹配单元的匹配结果，对所述承载协议进行识别，得到敏感数据。

[0120] B10、如B9所述的装置，所述匹配单元包括：

[0121] 匹配模块，用于将所述承载协议与所述预置存储位置中的交互操作类型进行匹配；

[0122] 识别模块，用于根据交互操作类型匹配结果，对所述承载协议进行识别，得到所述承载协议与所述预置应用层协议之间的交互操作数据；

[0123] 所述匹配模块，还用于将所述交互操作数据与预置存储位置中的特征字符数据进行匹配。

[0124] B11、如B9所述的装置，

[0125] 所述获取单元，具体用于根据预置代理网关设备获取预置应用层协议对应的承载协议。

[0126] B12、如B9所述的装置，

[0127] 所述匹配单元，具体用于通过预置多模式匹配算法将所述获取单元获取的所述承载协议与预置存储位置中的特征字符数据进行匹配。

[0128] B13、如B9所述的装置，

[0129] 所述识别单元，具体用于从所述承载协议中提取与预置存储位置中的特征字符数

据匹配成功的数据，得到敏感数据。

[0130] B14、如B9所述的装置，所述装置还包括：

[0131] 加密单元，用于对所述敏感数据进行加密处理。

[0132] B15、如B14所述的装置，所述装置还包括：配置单元，

[0133] 所述配置单元，用于根据所述识别单元得到的所述敏感数据对应的数据类型配置与所述敏感数据对应的加密策略；

[0134] 所述加密单元，具体用于根据所述配置单元配置的所述加密策略对所述敏感数据进行加密处理，得到加密文件。

[0135] B16、如B9-B15任一项所述的装置，所述预置应用层协议为超文本传输协议http，所述预置应用层协议对应的承载协议为超文本传输安全协议https。

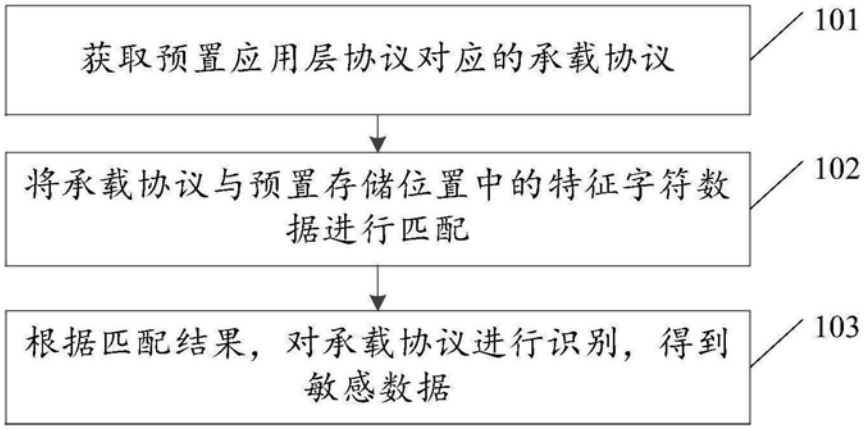

获取预置应用层协议对应的承载协议
101
将承载协议与预置存储位置中的特征字符数据进行匹配
102
根据匹配结果，对承载协议进行识别，得到敏感数据
103

图1

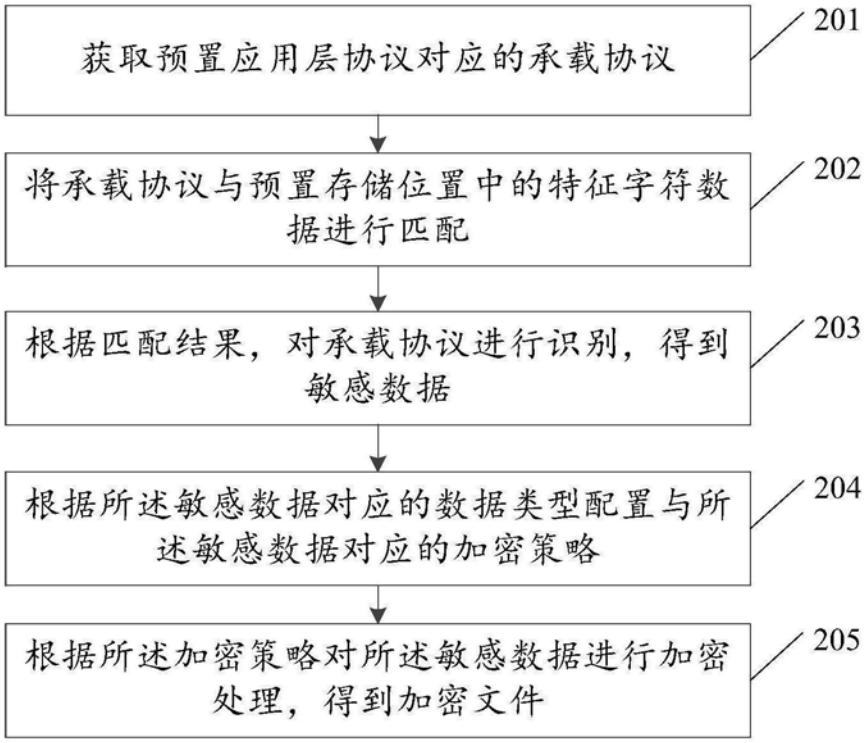

获取预置应用层协议对应的承载协议
201
将承载协议与预置存储位置中的特征字符数据进行匹配
202
根据匹配结果，对承载协议进行识别，得到敏感数据
203
根据所述敏感数据对应的数据类型配置与所述敏感数据对应的加密策略
204
根据所述加密策略对所述敏感数据进行加密处理，得到加密文件
205

图2

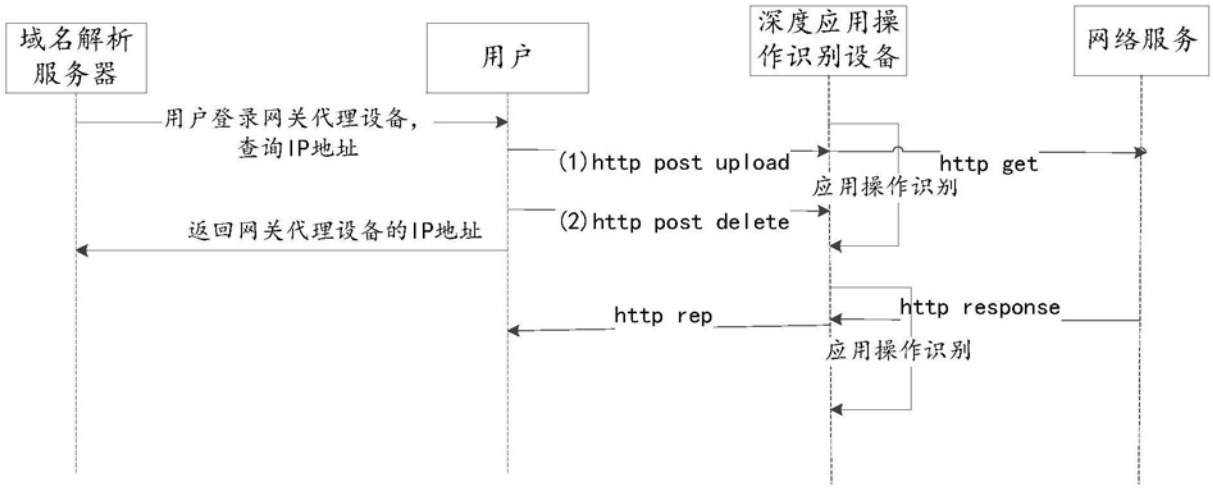

域名解析服务器
用户
深度应用操作识别设备
网络服务
用户登录网关代理设备，查询IP地址
(1)http post upload
http get
应用操作识别
返回网关代理设备的IP地址
(2)http post delete
http rep
http response
应用操作识别

图3

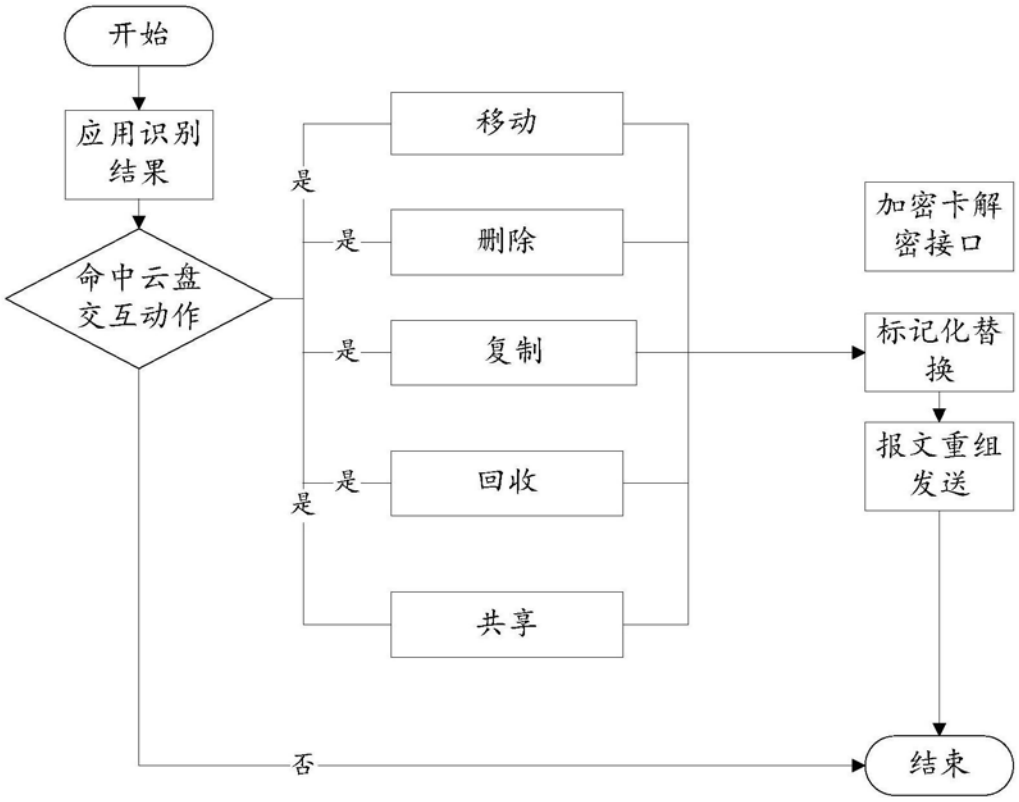

开始
应用识别结果
命中云盘交互动作
是
移动
是
删除
是
复制
是
回收
是
共享
加密卡解密接口
标记化替换
报文重组发送
否
结束

图4

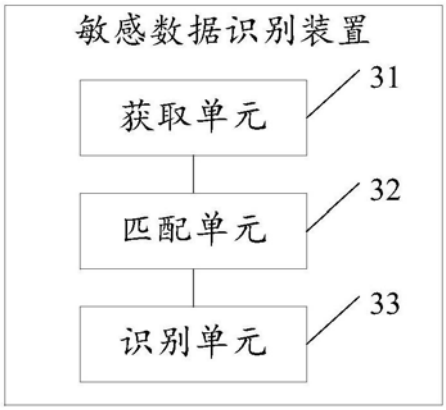

敏感数据识别装置
获取单元
31
匹配单元
32
识别单元
33

图5

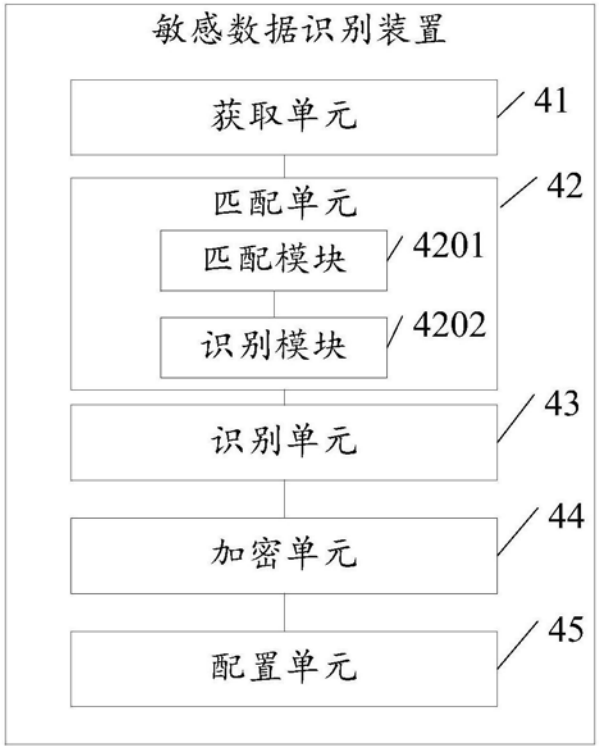

敏感数据识别装置
获取单元
41
匹配单元
42
匹配模块
4201
识别模块
4202
识别单元
43
加密单元
44
配置单元
45

图6