(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
2019 年 5 月 23 日 (23.05.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/096308 A1

(51) 国际专利分类号:
H04L 29/06 (2006.01)

(21) 国际申请号: PCT/CN2018/116207

(22) 国际申请日: 2018 年 11 月 19 日 (19.11.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201711141610.3 2017年11月17日 (17.11.2017) CN

(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: 赵 咏 (**ZHAO, Yong**); 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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, 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,

(54) **Title:** METHOD AND DEVICE FOR IDENTIFYING ENCRYPTED DATA STREAM

(54) 发明名称: 一种识别加密数据流的方法及装置

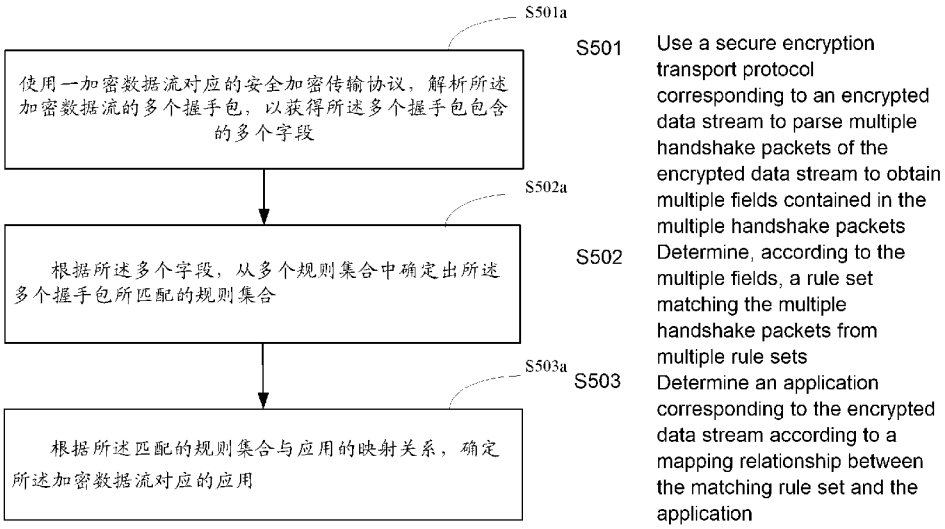

图 5a

(57) **Abstract:** The application describes a method for identifying an encrypted data stream. The method comprises: parsing, according to a secure encryption transport protocol corresponding to an encrypted data stream, multiple handshake packets of the encrypted data stream to obtain multiple fields contained in the multiple handshake packets; determining, according to the multiple fields, a rule set matching the multiple handshake packets from multiple rule sets; and determining an application corresponding to the encrypted data stream according to a mapping relationship between the matching rule set and the application. The method achieves more accurate identification of encrypted traffic in a network, namely, more accurate identification of an encrypted data stream at a greater volume.

WO 2019/096308 A1

[见续页]

RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, 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))。

(57) 摘要: 本申请描述了一种识别加密数据流的方法，该方法包括：根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的多个握手报文，以获得所述多个握手报文包含的多个字段；根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合，以及根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用，该方法能够更准确地识别出网络中的加密流量，即能够更多更准地识别出加密数据流。

说 明 书

一种识别加密数据流的方法及装置

技术领域

本发明涉及计算机领域，尤其涉及一种识别加密数据流的方法和装置。

背景技术

因特网（Internet）凭借其开放性、共享性等特点迅速普及并发展壮大，越来越多的新型网络应用应运而生，Internet 的开放性特点也意味着任何符合其技术标准的设备或软件都可以不受限制地接入互联网。为了对网络进行有效监督和管理，增强网络的可控性，例如有效利用带宽，并提供更好的服务质量（quality of service，QoS），或者提高网络的安全性，减少网络犯罪等行为的发生等，其中要解决的一个关键问题是对网络中的流量进行识别，例如从数据流的角度将流量分类，更具体的可以是确定网络中的数据流属于哪个应用，即一数据流携带的是哪个应用的数据。

常见的应用层之下的安全加密传输协议有安全套接字层（Secure Socket Layer，SSL）协议、传输层安全（Transport Layer Security，TLS）协议以及数据报传输层安全（Datagram Transport Layer Security，DTLS）协议等，其中，TLS 协议可以视为是 SSL 协议的升级版本，DTLS 协议基于 TLS 协议，用于保护 UDP 连接上数据的传输安全，网络中通过安全加密传输协议传输的数据流称为加密数据流（简称加密流），加密流的流量也被称为加密流量。加密流较难在除数据流的发端和收端被解析，因此识别网络中传输的加密流量所属的应用是目前业界主要的技术难题。

以 TLS 协议为例，识别加密流量可以通过解析握手报文中的 SNI（Server Name Indication）字段进行识别，该 SNI 字段是握手报文 ClientHello 中的字段，该字段用于指明握手报文 ClientHello 所在的加密流对应的域名（Host Name）。或者，还可以通过解析握手报文中 Common Name 字段进行识别，该 Common Name 字段是握手报文 Certificate(证书)的 subject 域中的字段,该字段中包括指示握手报文 Certificate 所在的加密流对应的域名的信息。

但是很多场景下，上述字段可以被设置为携带错误的信息或者模糊的信息（如通配符等），上述字段也可能在 TLS 报文中不存在，使用单一的 SNI 字段或者 Common Name 字段识别，使得对加密流量都不能正确识别出其对应的应用。

发明内容

本申请实施例提供一种识别加密数据流的方法及装置，能够更准确地识别出网络中的加密流量，即能够更多更准地识别出加密数据流，其中，识别是指识别出加密的数据流所对应的应用。

第一方面，本申请记载一种识别加密数据流的方法，该方法包括：根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的多个握手报文，以获得所述多个握手报文包含的多个字段；根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个握手报文中多个字段的顺序，所述多个握手报文中的字段满足所述

匹配的规则集合中的规则；根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用。

通过解析加密数据流中的多个握手报文中的字段，以及匹配这些字段与多个规则集合，找到该多个数据包对应的规则集合，再通过规则集合确定与之映射的应用，可以不再依赖握手报文中容易被篡改或者不准确的域名信息，能更多地识别出加密数据流对应于哪个应用，也能提高识别加密数据流的准确性。

应理解，该多个字段来自该多个握手报文。

应理解，多个规则集合就是多于一个的规则集合。则该“从多个规则集合中确定出所述多个握手报文所匹配的规则集合”，就是从该多于一个的规则集合中，确定出所述多个握手报文所匹配的规则集合。也就是说，该多个规则集合中也包括字段规则和顺序规则中的至少一种规则。

应理解，“根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的多个握手报文”，是本领域技术人员应理解的。例如一种实现方式下，就是对照一加密数据流对应的安全加密传输协议中对握手报文的字段的定义等信息，来解析出该加密数据流的多个握手报文中的多个字段。

一种实现方式下，多个规则集合中的每个规则集合都对应唯一的应用。

一种实现方式下，该多个规则集合中的每个规则集合与一个应用具有映射关系。

一种实现方式下，该多个握手报文满足所述匹配的规则集合中的所有规则。例如，该多个握手报文匹配的规则集合中只包括字段规则和顺序规则中的至少一种规则，该多个握手报文中的字段满足该匹配的规则集合中的全部规则。一种实现方式下，该多个握手报文满足所述匹配的规则集合中的部分规则。这种实现方式针对的是某些情况下，由于某些规则集合的特殊设置，无需匹配完一个规则集合中的全部规则（比如有些规则可选），就能确定该多个握手报文的对应的规则集合，以及就可以确定该加密数据流对应的应用了。

一种实现方式下，所述多个字段包括多个分组，所述多个分组中的每个分组对应一个握手报文，所述根据所述多个字段，从多个规则集合中确定所述多个握手报文所匹配的规则集合，包括：按照所述多个握手报文的接收顺序，将所述多个分组与所述多个规则集合中的规则匹配，以从所述多个规则集合中得到所述多个握手报文所匹配的规则集合。类似的，一种实现方式下，也可以按照所述多个握手报文的时序，将所述多个分组与所述多个规则集合中的规则匹配，以从所述多个规则集合中得到所述多个握手报文所匹配的规则集合。其中时序就是规定的握手阶段的握手报文交互顺序。这样，可以无需等待后续的握手报文都到达就可以开始识别加密流，且由于不同的规则集合中包括的规则对应的握手报文不同，也可以更快地识别出一些规则只分布在前几个握手报文的应用。其中，作为第三方面所述的装置的一种实现方式，第三方面所述的装置中的匹配模块可用于实现该步骤。

一种实现方式下，所述多个规则集合中包括字段规则和顺序规则，所述根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合包括：将所述多个字段与多个规则集合中的字段规则和顺序规则分别进行匹配，以从多个规则集合中确定出所述多个握手报文所匹配的规则集合。这样，分种类地对规则集合转给你

的规则进行匹配，可以批量处理，提高匹配的速度。其中，作为第三方面所述的装置的一种实现方式，第三方面所述的装置中的匹配模块可用于实现该步骤。

一种实现方式下，所述匹配的规则集合包括多个子集，所述多个子集中的每个子集对应至少一个应用，所述根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用，包括：根据所述多个子集中每个子集与应用的映射关系，得到所述多个子集对应的多个应用集合；求所述多个应用集合的交集，以得到所述规则集合对应的唯一的应用，所述应用为所述加密数据流对应的应用。其中，一种实现方式下，该多个规则集合中的每个规则集合中，可包括多个子集，每个中的规则对应的数据包的名称相同，同一规则集合中的不同子集对应的数据包的名称不同。也就是说，规则集合与应用的映射关系的表示形式是多种多样的，这样，就可以对这多个数据包一个一个并行或者串行地匹配规则，以及映射应用，可以加快对加密流的识别。其中，作为第三方面所述的装置的一种实现方式，第三方面所述的装置中的确定模块可用于实现该步骤。

一种实现方式下，所述多个规则集合中包括以链表形式保存的多条规则，其中，所述链表中的每一个节点保存对应一个握手报文的名称的规则。这样，能更好更有序地管理规则集合中的规则，可以方便一个握手报文一个握手报文的匹配，能更快速地识别加密流。

一种实现方式下，所述解析一加密数据流的多个握手报文，包括：对所述加密数据流的多个握手报文中的每一个握手报文，执行以下操作:解析所述每一个握手报文的名称；根据所述每一个握手报文的名称，从所述多个规则集合中，确定出对应于所述每一个握手报文的名称的至少一条规则；解析所述每一个握手报文中，所述至少一条规则所指示的字段。这样，可以只解析后续步骤需要的字段，减少解析步骤的处理量，使得识别过程更快，也节省解析步骤耗费的资源。其中，作为第三方面所述的装置的一种实现方式，第三方面所述的装置中的解析模块可用于实现该步骤。

一种实现方式下，所述方法还包括：通过机器学习算法训练多条加密数据流的多个样本，以得到目标应用对应的至少一个规则集合，其中，所述多个样本已知是否为所述目标应用对应的加密数据流的握手报文，所述多个样本中包括所述目标应用对应的加密数据流的握手报文。也就是说，规则集合，以及规则集合与应用的映射关系，可通过机器学习算法训练得到。这样，得到的上述信息能够更准确地描述数据包与应用间的对应关系。另外，也可以通过这种方式更新规则集合，以及规则集合与应用的映射关系。其中，作为第三方面所述的装置的一种实现方式，第三方面所述的装置可以还包括一训练模块，该训练模块可用于实现该步骤。

第二方面，本申请记载又一种识别加密数据流的方法，该方法包括：根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的一握手报文，以获得所述握手报文包含的多个字段；根据所述多个字段，从多个规则集合中确定出所述握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个数据包中多个字段的顺序，所述握手报文中的字段满足所述匹配的规则集合中的规则；根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的

应用。第二方面的方法，通过多个字段的特征和顺序中的至少一种可以不再依赖握手报文中容易被篡改或者不准确的域名信息，能更多地识别出加密数据流对应于哪个应用，也能提高识别加密数据流的准确性。

可见，第二方面与第一方面描述的方法很类似，区别在于第二方面使用一个握手报文就可以识别出该加密数据流对应的应用。则对第一方面中的解释以及技术效果，也可以用在第二方面中的相应部分，而第二方面中对一个数据包的解析，或者确定握手报文匹配的规则集合以及，确定加密数据流对应的应用这些步骤的实现方式的描述，也适用于第一方面中处理多个数据包。

一种实现方式下，所述多个规则集合是对应于所述握手报文的名称的规则集合。那么一种实现方式下，所述解析一加密数据流的一握手报文，包括：从所述多个规则集合中，确定出对应于所述每一个握手报文的名称的至少一条规则；解析所述每一个握手报文中，所述至少一条规则所指示的字段。这样，可以只解析后续步骤需要的字段，减少解析步骤的处理量，使得识别过程更快，也节省解析步骤耗费的资源。其中，作为第四方面所述的装置的一种实现方式，第四方面所述的装置中的解析模块可用于实现该步骤。

一种实现方式下，所述多个规则集合中包括字段规则和顺序规则，所述根据所述多个字段，从多个规则集合中确定出所述握手报文所匹配的规则集合，包括：将所述多个字段与多个规则集合中的字段规则和顺序规则分别进行匹配，以从多个规则集合中确定出所述握手报文所匹配的规则集合。这样，分种类地对规则集合转给你的规则进行匹配，可以批量处理，提高匹配的速度。其中，作为第四方面所述的装置的一种实现方式，第四方面所述的装置中的匹配模块可用于实现该步骤。

一种实现方式下，该方法还包括通过机器学习算法训练多个样本，以得到目标应用对应的至少一个规则集合，其中，所述多个样本已知是否为所述目标应用对应的加密数据流的握手报文，所述多个样本中包括所述目标应用对应的加密数据流的握手报文。其中，作为第四方面所述的装置的一种实现方式，第四方面所述的装置可以还包括一训练模块，该训练模块可用于实现该步骤。

下面列举几个第一方面和第二方面的方法均适用的实现方式。

上述两个方中，握手报文可以是执行该方法的装置主动抓取的，也可以是由其他网络节点发送的，该多个握手报文还可以是根据网络中传输的握手报文复制得来的。

上述两个方面中的解析步骤就是使用该加密数据流对应的安全加密传输协议，解析该加密数据流的握手报文，以获得所述握手报文包含的多个字段。其中，该加密数据流对应的安全加密传输协议为 SSL 协议、TLS 协议、DTLS 协议或者其他运行于应用层下的安全加密传输协议。则相应的，一种实现方式下，所述多个握手报文为安全套接字层 SSL 握手报文、传输层安全 TLS 握手报文或者数据报传输层安全 DTLS 握手报文。

一种实现方式下，多个规则集合，以及规则集合与应用的映射关系存储在识别信息库中。该识别信息库可以在执行识别加密数据流方法的设备中，也可以在其他设备中。也就是说，作为一种实现方式，第三方面和第四方面所述的装置中，还可以包括一个存储模块用于存储识别信息库中的信息。

一种实现方式下，对一条字段规则，所指示的字段的特征为字段的长度，或者字段的类型，或者字段的长度和类型，或者字段的类型和值，或者字段的长度、类型和值。一种实现方式下，字段的特征为字段的长度，类型和值中的至少一个。

一种实现方式下，所述多个规则集合包括多条规则，所述根据所述多个字段，从多个规则集合中确定与所述握手报文（或者多个握手报文）匹配的规则集合包括：将所述多个字段与所述多条规则进行匹配，以得到所述多个字段对应的一组规则；将所述一组规则与所述多个规则集合进行匹配，以得到与所述握手报文（或者多个握手报文）匹配的规则集合。其中，作为第三方面以及第四方面所述的装置的一种实现方式，第三方面以及第四方面装置中的匹配模块可用于实现该步骤。

一种实现方式下，根据所述多个字段，从多个规则集合中确定与所述握手报文（或者多个握手报文）匹配的规则集合包括：将所述多个规则集合与所述多个字段进行匹配，直到得到与所述握手报文（或者多个握手报文）匹配的规则集合。其中，作为第三方面以及第四方面所述的装置的一种实现方式，第三方面以及第四方面装置中的匹配模块可用于实现该步骤。

第三方面，本申请记载一种识别加密数据流的装置，所述装置包括：解析模块，所述解析模块用于根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的多个握手报文，以获得所述多个握手报文包含的多个字段；匹配模块，所述匹配模块用于根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个握手报文中多个字段的顺序，所述多个握手报文中的字段满足所述匹配的规则集合中的规则；确定模块，所述确定模块用于根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用。

通过解析加密数据流中的多个握手报文中的字段，以及匹配这些字段与多个规则集合，找到该多个数据包对应的规则集合，再通过规则集合确定与之映射的应用，可以不再依赖握手报文中容易被篡改或者不准确的域名信息，能更多地识别出加密数据流对应于哪个应用，也能提高识别加密数据流的准确性。

一种实现方式下，该识别加密数据流的装置中，保存有该多个规则集合以及规则集合与应用的映射关系。

由于第三方面是第一方面对应的装置，有关第三方面的各种名词解释、实现方式以及技术效果，请参见第一方面中的描述，本申请不再赘述。

第四方面，本申请记载一种识别加密数据流的装置，所述装置包括：解析模块，所述解析模块用于根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的一握手报文，以获得所述握手报文包含的多个字段；匹配模块，所述匹配模块用于根据所述多个字段，从多个规则集合中确定出所述握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个数据包中多个字段的顺序，所述握手报文中的字段满足所述匹配的规则集合中的规则；确定模块，所述确定模块用于根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应

的应用。

一种实现方式下，该识别加密数据流的装置中，保存有该多个规则集合以及规则集合与应用的映射关系。

由于第四方面是第二方面对应的装置，有关第四方面的各种名词解释、实现方式以及技术效果，请参见第二方面中的描述，本申请不再赘述。

第五方面，本申请记载一种识别加密数据流的设备，所述设备包括处理电路、接口电路和存储介质，所述接口电路用于通过所述存储介质中的指令与其他设备收发数据包，所述处理电路用于运行所述存储介质中的指令控制所述接口电路，以实现如第一方面及其实现方式描述的各种方法。

由于第五方面是第一方面对应的设备，有关第五方面的各种名词解释、实现方式以及技术效果，请参见第一方面中的描述，本申请不再赘述。一种实现方式下，第三方面描述的装置运行在第五方面的设备中。

第六方面，本申请记载又一种识别加密数据流的设备，所述设备包括处理电路、接口电路和存储介质，所述接口电路用于通过所述存储介质中的指令与其他设备收发数据包，所述处理电路用于运行所述存储介质中的指令控制所述接口电路，以实现如第二方面及其实现方式描述的各种方法。

由于第六方面是第二方面对应的设备，有关第六方面的各种名词解释、实现方式以及技术效果，请参见第二方面中的描述，本申请不再赘述。一种实现方式下，第四方面描述的装置运行在第六方面的设备中。

本申请的又一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。

本申请的又一方面提供了一种包含指令的计算机程序产品，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。

上述两方面的各种名词解释、实现方式以及技术效果的描述具体可以参见上述对相应方面的技术效果的相关描述，此处不再赘述。

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。

图 1 为本申请记载的一种数据流传输示意图；

图 2 为本申请记载的一种网络架构图；

图 3 为本申请记载的一种监控设备的架构图；

图 4a 为本申请记载的 TLS 流的一种握手流程的示意图；

图 4b 为本申请记载的 TLS 流的又一种握手流程的示意图；

图 5a 为本申请记载的一种识别加密流的方法的示意图；

图 5b 为本申请记载的另一种识别加密流的方法的示意图；

图 6 为本申请记载的一种 TLS 流的握手报文的匹配过程的示意图；

图 7 为本申请记载的一种识别加密流的装置的示意图；

图 8 为本申请记载的一种可用作识别加密流的设备的物理机的示意图。

具体实施方式

本文中字符“/”，一般表示前后关联对象是一种“或者”的关系。例如，A/B 可以理解为 A 或者 B。

本发明的说明书和权利要求书中的术语“第一”和“第二”等是用于区别不同的对象，而不是用于描述对象的特定顺序。

在本发明的描述中，除非另有说明，“多个”的含义是指两个或两个以上。

此外，本发明的描述中所提到的术语“包括”和“具有”以及它们的任何变形，意图在于覆盖不排他的包含。例如包含了一系列步骤或单元的过程、方法、系统、产品或设备没有限定于已列出的步骤或单元，而是可选地还包括其他没有列出的步骤或单元，或可选地还包括对于这些过程、方法、产品或设备固有的其它步骤或单元。

以下描述中，为了说明而不是为了限定，提出了诸如特定系统架构，装置架构，技术之类的具体细节，以便透切理解本发明。然而，本领域的技术人员应当清楚，在没有这些具体细节的其它实施例中也可以实现本发明。在其它情况中，省略对众所周知的装置、电路以及方法的详细说明，以免不必要的细节妨碍本发明的描述。

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详细地描述。

为了更加清楚地理解本发明实施例的各种实现方式，下面首先对本发明实施例中涉及的技术术语进行定义/说明。

安全加密传输协议：背景技术中提及的 SSL 协议、TLS 协议、DTLS 协议或者其他运行于应用层下的安全加密传输协议。由于技术的演进，每种安全加密传输协议都可以有多个版本。

握手报文（Handshake packet）：用户端和客户端之间用于通过安全加密传输协议建立连接的报文。握手报文出现在 TLS、SSL 或者 DTLS 等连接的握手阶段。在 TLS、SSL 以及 DTLS 的各个版本（如 TLS1.0~TLS1.3,SSL1.0~SSL3.0 以及 DTLS 1.0~DTLS1.3 等等）里都有关于握手阶段和握手报文的介绍，可以参照如 RFC 5246、RFC 6176、RFC6083、RFC6347 等等。例如下列报文都是握手报文：ClientHello、HelloVerifyRequest、ServerHello、Certificate、ServerKeyExchange、ServerHelloDone、ClientKeyExchange、HelloRequest、ChangeCipherSpec、EncryptedHandshakeMessage、NewSessionTicket、Alert、Finished、Certificatetatus、CertificateVerify、CertificateRequest。

应理解，一些情况下，由于网络环境或者通信协议的限制，上述一个握手报文的信息可以使用多个包或报文承载，也就是说，可以多个包或报文通过解析后共同组成一个完整的握手报文所包括的信息。应理解，本申请中的握手报文的名称或者报文的名称，指的是该握手报文通过对应的安全加密传输协议解析出的，在 Handshake Type 字段的值，或者握手报文解析后，记录在“Handshake Protocol: ”这一行的冒号后的内容，例如：“Handshake Protocol: Client Hello”，Client Hello 为该握手报文的名称；或者在其他版本的加密协议中，用于表示相同含义的字段的值。例如 Client

Hello、Certificate、ServerKeyExchange 等都是握手报文的名称。

报文（packet）：有时也翻译为包，例如数据包也是种报文。本申请中用户端和服务端之间的加密数据流中包括多个报文。本申请中的识别加密数据流对应的应用，其实就是通过识别加密数据流中握手报文对应的应用实现的。

数据流：一组有顺序的、有起点和终点的字节集合。

SSL 协议、TLS 协议以及 DTLS 协议都是当前使用非常广泛的应用层之下的安全加密传输协议。其中，TLS 协议是 SSL 协议的升级版本，二者都用于保护 TCP 连接的数据，故业界有时也 SSL 和 TLS 并用；而 DTLS 协议基于 TLS 协议，用于保护 UDP 连接的数据。例如，根据国外部分研究机构的数据显示，已有接近 60%的网络流量采用 SSL/TLS 进行加密保护。上述的安全加密传输协议用以保障数据传输的安全，利用数据加密(Encryption)的技术，可以确保数据在网络上的传输过程中不会被截取及窃听。

使用上述的安全加密传输协议，可以在网络中传输各种应用的数据。本申请描述的技术方案，就是用于识别使用上述的安全加密传输协议的数据流（下文简称加密流或者数据流）所对应的应用。

加密流的发送端和接收端可以采用户端/服务端（C/S，Client/Server）这种软件系统体系结构架构进行描述，一条加密流在一个应用（application）的用户端与服务端之间传输。其中，用户端和服务端在这种架构上指的都是应用，具体的，作为用户端（client）的应用请求服务，作为服务端（server）的应用为作为用户端（client）的应用提供服务。本申请中描述的加密流的用户端和服务端是同种应用。这样多个应用可以部署在相同或者不同的物理设备上，从而充分利用 client 和 server 所在的硬件环境的优势。也就是说本申请描述识别加密数据流的方法，并不限定该数据流的 client 和 server 部署在何种设备上，如可以是终端，也可以是服务器，也可以是云平台，也就是说数据流可以是传输于端（terminal）云(cloud)之间、也可以是传输于云到云或端到端等。SSL/TLS 协议位于 TCP/IP 协议与各种应用层协议之间，为数据通讯提供安全支持。应用数据需要先经过 SSL 层的处理，数据才能通过 TCP/IP 层发送出去。也就是说，client 和 server 间建立 TLS 连接之前，需要先建立 TCP 连接。DTLS 协议则位于 UDP 协议与各种应用层协议之间。以数据流在终端与服务器之间传输为例，图 1 为数据流的传输示意图，应用部署在应用层，其中，终端与服务器之间可以有监控设备。而另一方面，由于数据被加密，能用于识别加密数据流的信息很有限。本申请的一些实施例中，使用握手报文中的信息识别一加密流。本申请的一种场景中，用户端与服务端是相同应用的相同版本，另一种场景中，用户端与服务端是相同应用的不同版本。

另一方面，本申请的方法可以应用于网络中的监控设备，监控设备用于识别流量，而网络的具体组网方式本申请不做限制，例如，监控设备可以如图 2 所示的设置在需要检测的链路中，例如一个网关，也可以是旁路于需要检测的链路，例如具有识别流量功能的旁路分析设备，例如一个服务器中运行有具有识别流量功能的软件，则该服务器可视为监控设备。也就是说，从该需要检测的链路的某个节点复制出从该节点传输的需要识别的数据流输入到该监控设备。例如，图 2 是一种可以使用本申请方法的

网络的架构图。其中，互联网用户使用的终端通过本地网络连到接入网关。接入网关将用户的数据发送到骨干网（backbone network）的路由器上。骨干网（Backbone Network）是用来连接多个区域或地区的高速网络。每个骨干网中至少有一个和其他骨干网进行互联互通的节点点。例如不同的网络供应商都拥有自己的骨干网，用以连接其位于不同区域的网络。一般情况下，互联网上的各种服务器也通过接入网关连入骨干网的路由器。在互联网用户和服务器对应的接入网关之间存在多跳路由器，在这些接入网关和路由器之间均可以加入流量监控设备。这样流量监控设备就可以分析接收到的报文，以得出识别结果。

下面结合图 3 描述一种实现方式下，监控设备的架构，监控设备就是用于执行本申请描述的识别加密流的方法的设备，也就是用于识别加密流的设备，其软件部分可以看做是本申请记载的识别加密流的装置（例如图 7）的一种实现方式。监控设备包括一输入接口和一输出接口，网络中的数据流可从输入接口流入该监控设备，再从输入接口流出。图 3 中，为便于理解，输入接口与输出接口是分离的，而实际的设备中，输入接口和输出接口可以是同一接口，如 I/O（输入/输出）设备，输入接口与输出接口可以是天线，接口电路等等。该监控设备可以分析从输入接口流入的数据流（实际上是多个报文），也可以将输入接口流入的数据流复制或者镜像一份进行分析，而从输入接口流入的数据流则从输出接口流出。例如图 3 中两个虚线椭圆框中，示意性化为多项开关的装置，表示用于控制流量的单元，两开关在竖直位置的情况下，图 3 示意从输入接口流入的数据流通过下述的包处理器等模块进行识别；两开关在水平位置的情况下，图 3 示意输入接口流入的数据流未被分析，直接从输出接口流出，该情况下图 3 中未示意镜像的数据流通过下述的模块被分析。

下面结合图 3 简要描述一种实现方式下，分析数据流过程中的主要步骤：通过包处理器解析数据流中的报文，其中，包处理器使用 TCP/IP 协议栈或者 UDP 协议栈解析报文，以得到报文中的字段。可以解析数据流中的一部分报文，以及可以解析报文中的一部分字段。解析后的报文通过流处理器确定解析后的报文属于哪个数据流，即解析后的报文哪些属于同一个数据流，例如可以通过解析出的报文中的五元组（其中包括源 IP 地址、目的 IP 地址和端口号）来确定，例如给具有相同的五元组的报文分配同一个流的标识。监控设备可以通过哈希表记录解析后的报文哪些属于同一个流。流量识别引擎以流为单位分析解析后的报文，即分析一个流的一个或多个报文，其中包括使用用于识别加密数据流的规则集合对该一个或多个报文进行匹配，以确定出该一个或多个报文匹配的规则集合，根据该规则集合，以及规则集合与应用的映射关系，确定该流对应的应用。

其中，规则集合就是规则的集合，识别信息库用于保存上述的多个规则集合以及规则的集合与应用的映射关系。上文中描述的包处理器、流处理器、流量识别引擎以及识别信息库，都是软件模块，本文为了便于理解从功能角度如此划分，而实际实现上，只要可以实现本申请所描述的方法，并不限定监控设备中软件模块的划分方式。当然，一种实现方式下，监控设备可以和其他设备配合完成上述功能，例如上述的包处理器、流处理器、流量识别引擎以及识别信息库可以以云化的方式。或者分布式的方式分布在多个节点，这时候，也可以认为实现上述功能的一系列模块构成了一个监

控系统以实现上述的监控功能，或者也可以将整个监控系统当做云化的或者分布式的监控设备。即，本文中的监控设备不限定是单个的物理节点。例如，保存有多个规则集合以及规则集合与应用的映射关系的识别信息库可以保存在其他的服务器或存储节点上，监控设备中只加载一部分甚至不保存识别信息库中的信息。监控设备通过与其他的服务器或者存储节点通信使用该识别信息库。

使用 TLS 协议的连接称为 TLS 连接，使用 DTLS 协议的连接称为 DTLS 连接。

每个 TLS 连接或者 DTLS 连接在传输数据前，都要通过在 client 与 server 间传输握手报文以建立连接。握手完成后，client 与 server 间才进行数据传输。

下面举例示意性描述使用 SSL/TLS 协议传输数据，用户端（client）与服务端（server）之间的握手过程，对于 DTLS 协议的握手过程，请参见标准 RFC6347 的相关内容，本申请不做赘述。

应理解，TLS 协议有多个版本，也适用于多种场景，进一步的细节可以参看不同版本的 TLS 协议中的相关段落。例如图 4a 所示的是 TLS1.2 中 ECDHE-ECDSA 的场景下的一个例子, ECDHE-ECDSA 表示一 server 和一 client 首次连接时的 TLS 握手流程，图 4b 所示的是 TLS1.2 中 PSIC 的场景下的一个例子，PSIC 表示一 server 和一 client 非首次连接（如连接后又断开，或者前一个连接未断开又协商一个新的连接等）时的 TLS 握手流程。可见不同的场景下，TLS 握手过程不同。

图 4a 中，client（用户端）与 server（服务端）按照时序交互握手报文。其中包括:

client 向 server 发送 Client Hello 报文(图 4a 和图 4b 中均译为用户端问候报文), 该报文可以用于向 server 告知该 client 所支持的加密算法。

server 向 client 回复 Server Hello 报文（图 4a 和图 4b 中均译为服务端问候报文文）,Certificate 报文（图 4a 中译为证书报文）,Server Key Exchange 报文（图 4a 中均译为服务端密钥交换报文）和 Server Hello Done 报文（图 4a 和图 4b 中均译为服务端问候完成报文）。其中，Server Hello 报文可以用于向 client 告知此次握手使用的加密算法；Certificate 报文可以向 client 表明该 server 身份的合法性和真实性，例如不是被假冒的 server；Server Key Exchange 报文可以用于向 client 发送密钥以进行密钥协商；Server Hello Done 报文可以用于告知 client，server 这一组报文均已发送，该 client 可以发送接下来的报文。

client 接收到这些报文后，向 server 发送 Client Key Exchange 报文（图 4a 和图 4b 中均译为用户端密钥交换报文）和 Change Cipher Spec 报文（图 4a 和图 4b 中均译为更改密码报文）。server 接收到这两个报文后，向 client 发送 Change Cipher Spec 报文和 Encrypted Handshake Message 报文（图 4a 和图 4b 中均译为加密握手信息报文）。Client Key Exchange、Change Cipher Spec 以及 Change Cipher Spec 可以用于 client 和 server 协商接下来数据传输过程中使用的密钥。Server 发送的 Encrypted Handshake Message 报文可以用于告知 client 握手过程结束。

图 4b 描述的 PSIC 的场景的握手交互过程相对简单。client 向 server 发送 Client Hello 报文，该报文可以用于向 server 告知该 client 所支持的加密算法。

server 向 client 回复 Server Hello 和 Server Hello Done 报文。其中，Server Hello

报文可以用于向 client 告知此次握手使用的加密算法。

client 接收到这些报文后，向 server 发送 Client Key Exchange 和 Change Cipher Spec 报文。server 接收到这两个报文后，向 client 发送 Change Cipher Spec 和 Encrypted Handshake Message 报文。Client Key Exchange、Change Cipher Spec 以及 Change Cipher Spec 可以用于 client 和 server 协商接下来数据传输过程中使用的密钥。Server 发送的 Encrypted Handshake Message 报文可以用于告知 client 握手过程结束。

client 解析这两个报文的信息后，client 与 server 之间的握手过程结束。之后，client 和 server 可以进行数据交互，数据以 TLS 流的方式在网络中传输，而从指令交互角度，则图 4a 和图 4b 都示例了一种实现方式。即，client 发送 client Request（图 4a 和图 4b 中均译为用户端请求报文）以请求 server 发送数据，server 发送 Server Response（图 4a 和图 4b 中均译为服务端应答报文）以向 client 发送数据。可见，TLS 流的握手报文可以指示该 TLS 流。

在上述的 TLS 握手的报文交互过程之前，client 和 server 需要完成 TCP 层的三次握手的交互过程，图 4a 和 4b 中以 TCP START（图 4a 和图 4b 中均译为 TCP 连接开始）示意。在 client 和 server 间完成了本次连接中的 TLS 数据交互后，可以断开该 TLS 连接所依赖的 TCL 连接，图 4a 和 4b 中以 TCP FINISH（图 4a 和图 4b 中均译为 TCP 连接结束）示意。

本申请描述的识别加密数据流所使用的报文，就是来自握手阶段。可见，不同场景的握手阶段所包括的握手报文不同。例如，在协议 TLS1.2 中，上述握手阶段的握手报文大多是明文的（即监控设备获得握手报文后可以解析出大多字段的信息），而在协议 TLS1.3 中，握手阶段的明文只有 Client Hello 报文和 Server Hello 报文。

例如现有技术中，通过一 TLS 流的握手报文中的 SNI 字段或者 Common Name 字段，确定发送该 TLS 流的域名，从而识别该 TLS 流对应的应用。显然，这是由于这些字段中携带有域名(Host name)，由于域名携带发送该流的数据的服务器的信息，故域名中的一部分可以指示应用，如域名为 book.google.com，则可知是 google 的图书应用，又如域名为 mail.google.com，则可知是 google 的 gmail 应用。也就是说这两个字段就携带有与应用有关的信息，现有技术是通过解析字段的取值（域名）的含义指示应用的。

然而，由于以下原因，现有技术能识别出的数据流很有限，识别的准确率也较低。

1、上述字段中的域名信息在很多情况下是错误的，比如欺诈、VPN、病毒类的应用会刻意设置错误的 SNI，用以逃避应用识别。

2、越来越多的域名中不再携带加密流对应的应用的信息，或者不足以通过域名判断出应用。例如 Common Name 中的域名并不是精确地域名，大部分情况下是带有“*”通配符的域。这样能够保证一个证书同时支持一系列的域名。但是同时也给应用识别带来了困难。以上文提及的 google 的两个应用为例，而谷歌的服务器发送的 Certificate 报文中 Common Name 字段一般设置为*.google.com，此时是无法依靠 Common Name 字段区分该流属于谷歌下的哪个应用的。

3、很多被传输的数据（如图片、视频等）被存储在云平台（如 Amazon S3）上，

那么域名就只能指示出云平台的信息，例如域名为 a248e.akamai.net 或者 s3.amazon.com，不同的数据通过域名后面的具体路径来区分。例如，应用 A 和 B 均在 s3.amazon.com 上部署了自己的图片内容，但是用户使用 HTTPS 访问应用 A 和 B 时产生的流量使用的 SNI 都是 s3.amazon.com。这时该技术是没有办法区分应用 A 和 B 的。

4.另外，SNI 字段或者 Common Name 字段很可能不在报文中携带或者无法解析，则无法通过域名识别加密数据流对应的应用。SNI 字段是 ClientHello 报文的可选字段。例如，对于使用安全加密传输协议传输 FTP、邮件等通信场景，一般不设置 SNI 字段。即使是使用安全加密传输协议传输 HTTP 内容的 HTTPS 场景，很多时候也不存在 SNI。例如 IE6 浏览器作为 Windows XP 内置的浏览器，其 HTTPS 通信流量并不存在 SNI 字段。正如图 4a 和 4b 所示，Certificate 报文不是在每个 TLS 流中出现的。TLS 存在握手重用机制，握手重用的情况下不需要传输证书（Certificate），直接利用上次握手的信息。以及，TLS 的下一代标准 TLS1.3 将调整 TLS 协议的加密机制，其中 ClientHello 和 ServerHello 报文之后的所有报文都会加密传输。Certificate 握手报文变成了加密传输，作为网络设备的网关或者监控设备是无法解析出该握手报文中 Common Name 字段携带的域名的。

上述的缺陷都是由于现有技术仅依赖 SNI 字段或者 Common Name 字段进行加密流识别而带来的，但是 SNI、Common Name 这些字段可以不设置、可以设置错误、也可以设置为模糊的值，这时会造成无法识别或者识别错误。而本申请中，使用到其他字段来识别加密流，而并不依赖 SNI 字段或者 Common Name 字段，而其他字段并不携带能够指示应用的信息。因此本申请的方案不依赖字段的取值的含义，而是通过匹配规则集合，以及利用规则集合与应用的映射关系，能够识别出更多的数据流对应的应用，识别的准确率大大提高。

事实上，本申请中，用于识别加密数据流所使用的握手报文，可以是以下报文中的至少一种：ClientHello、HelloVerifyRequest、ServerHello、Certificate、ServerKeyExchange、ServerHelloDone、ClientKeyExchange、HelloRequest、ChangeCipherSpec、EncryptedHandshakeMessage、NewSessionTicket、Alert、Finished、Certificatetatus、CertificateVerify、CertificateRequest。也就是说，下文中描述的用于识别加密数据流的多种多样的规则，是通过检测上述一个或多个握手报文中的字段是否与其匹配，来识别该握手报文对应的加密流。

安全加密传输协议是一种 TLV (Type 类型, Length 长度, Value 值)格式的协议。故解析后，握手报文中的字段一般包括类型（Type）、长度(Length)和值(Value)三个部分，这三个部分中的每一个或者任意组合都可以称为该字段的特征。

其中，类型可以有不同的级别，例如对扩展字段 extension:ec_point_formats(len=2),extension 是类型，ec_point_formats 也是类型。通常用一个字段的类型来代指一个字段，例如称 extension: ec_point_formats(len=2)为扩展字段，或者更具体的称为扩展字段中的 ec_point_formats 字段。当然实际的报文中，有些字段的三个部分中可以有某些部分缺省。网络设备可以解析出接收到的报文中未加密的字段的类型、长度和值的具体内容，例如 Client Hello 报文中以 11 表示

一扩展字段的类型，通过查表可以得知 11 所代表的是 ec_point_formats。在匹配一个字段规则时，如果该字段规则包括对这三个部分的限定，而被匹配的报文中，该字段缺省了其中的一个或者多个部分，缺省部分可以默认为满足相应的规则。上述的缺省情况，可以是一个字段在协议中的定义就缺少一个部分，也可以是某些部分在协议中是定值，故不携带在报文中。如在 Certificate 报文中，id-at-countryName=US 字段仅有类型（id-at-countryName）和值（US）两部分内容，而 algorithm(rsaEncryption) 字段由于 TLS 协议已经规定了在固定位置出现该类型字段，且字段长度固定，因此在该字段中仅携带字段值（rsaEncryption）。

另一方面，握手报文的字段可以是嵌套的，体现为 TLV 格式的嵌套使用。例如 Client Hello 报文中的一个扩展（Extension）字段,其类型（Type）为 ec_point_formats(11),其中 11 为未解析的报文中用于表示该字段的类型的数字，ec_point_formats 为解析得到的具体类型，该字段的长度为 2，其值为 EC point formats Length:1 和 Elliptic curves point formats(1),其中，Elliptic curves point formats(1)下一级字段的 TLV 格式中的类型，而 EC point formats Length:1 是下一级字段的 TLV 格式中的长度。上述信息在解析出的字段中可以以下方式记载：

◢ Extension：ec_point_formats(len=2)
Type: ec_point_formats(11)
Length:2
EC point formats Length:1
▷ Elliptic curves point formats(1)

其中◢ 表示对 extension 字段内容的展开，▷ 表示 Elliptic curves point formats(1) 字段下还有其他内容未展开。

图 5a 和图 5b 分别描述了一识别加密数据流的方法。其中，图 5a 的方法，通过一加密数据流的多个握手报文来识别加密数据流，图 5b 的方法，通过一加密数据流的一个握手报文来识别加密数据流。图 5a 包括步骤 S501a，S502a 和 S503a；图 5b 包括步骤 S501b，S502b 和 S503b。应理解，图中的序号和步骤间的连接关系，不限定步骤的执行顺序。以及，本申请不限定图 5a 的方法和图 5b 的方法的具体实现方式。

例如图 5a 中的 S501a 和 S502a，某些实施方式中是交错往复执行的。这种交错往复执行也是具有多种实现方式的，本申请仅做示意性说明，不一一列举。一种实现方式下，可以解析一个报文，根据其中的字段从该多个规则集合中确定出多个匹配的规则集合，再解析一个报文，再从该多个规则集合中确定出多个匹配的规则集合，然后求交集，如果交集不是唯一的规则集合，则继续解析报文进行匹配，直到得到一与该多个报文匹配的规则集合。

一种实现方式下，可以解析一个报文，根据其中的字段从该多个规则集合中确定出多个匹配的规则集合，再解析一个报文，再从该多个匹配的规则集合中确定出与该两个报文匹配的规则集合，如果与该两个报文匹配的规则集合不是唯一的规则集合，则继续解析报文与该两个报文匹配的规则集合中进行匹配，直到得到一与该多个报文匹配的规则集合。当然，可选的，也可以在其中一次匹配中判断，如果已匹配出的规则集合中，有规则是该流的其他报文不满足的，那么就舍弃该已匹配出的规则集合，

重新在多个规则集合中匹配，或者重新在之前匹配出的规则集合中匹配。

再例如一种实现方式下，可以解析该多个报文中的一部分字段，根据该一部分字段与该多个规则集合进行匹配，得出与该一部分字段匹配的规则集合，再解析该多个报文中的另一部分字段，根据该另一部分字段与得出的匹配的规则集合进行匹配。

再例如一种实现方式下，可以先解析报文中字段的属性，根据多个字段的属性及其顺序，匹配规则集合中的顺序规则（顺序规则可存在或不存在），再解析报文的中字段的其他特征，以匹配其中的字段规则。当然，这种实现方式更适用于多个规则集合中包括顺序规则的情况。

再例如图 5a 中的 S502a 和 S503a，可能是先匹配一个报文，得到该报文对应的规则集合，以及该报文对应的规则集合中，该报文满足的规则子集，根据该满足的规则子集与应用的匹配关系，得到该报文对应的应用集；再匹配该加密数据流的其他报文，对应用集不断限缩或者得到若干应用集求交（具体不做限制），最终得到该加密数据流匹配的应用。

又例如图 5a 中的 S501a,S502a 和 S503a，对多个报文，步骤可以并行，即，可能对有的报文执行 S501a 时，对另一报文在执行 S502a，而对又一报文在执行 S503a。一种实现方式下，可以是解析该加密数据流中多个报文中的一个，根据该报文中的字段从多个规则集合中确定该报文匹配的至少一个规则集合，得到该匹配的至少一个规则集合中该报文满足的规则子集各自对应的应用集合，

类似的，例如图 5b 中的 S501b 和 S502b，某些实施方式中是交错往复执行的。这种交错往复执行也是具有多种实现方式的，本申请仅做示意性说明，不一一列举。

例如一种实现方式下，可以每解析一个字段，就将该字段与多个规则集合中的字段规则匹配，再根据与解析的字段所匹配的规则集合中其余的规则所涉及的字段，去解析相应的字段。

再例如，一种实现方式下，可以将一个握手报文分段解析，每解析出一段的若干字段，就将该若干字段与多个规则集合中的规则进行匹配，得到至少一个与该若干字段匹配的规则集合，再解析出其他段的字段进行匹配，以得到该报文匹配的规则集合。

关于图 5a 和图 5b 中的方法，本申请的发明内容部分也有对其的一些实现方式的描述，以及，下文中有以 TLS 流为例的实现方式的描述，本领域技术人员通过这些实现方式可得知其他安全加密传输协议的加密流的识别的相应的实现方式，故本申请不再赘述。

上述的各种实现方式，可以用于不同的场景，如识别网络中的加密流对应什么应用，再例如从网络中传输的数据流中，识别出对应某个应用的加密流，在不同的场景下，可能匹配速度较快的实现方式不同。

下面对图示以及上述实现方式中出现的名词以及一些具体实现方式进行描述。应理解，本申请描述多种对一个握手报文的处理的实现方式，可以用在使用多个握手报文识别，也可以用在使用一个握手报文识别，本文不再分别描述。

本领域技术人员应当理解，本申请中，使用安全加密传输协议解析出一个报文的字段，不仅能解析出这些字段的特征，也能解析出这些字段的顺序。

需说明的是，对前文描述的，根据多个字段，从多个规则集合中确定出多个或一

个报文匹配的规则集合，意思是说：该多个或一个报文所匹配的是一个规则集合，该匹配出的规则集合与该多个规则集合中的某个规则集合相同，即，这个匹配既约束该多个或一个报文匹配若干条规则，还约束这若干条规则的组合是该多个规则集合中的一个。以多个报文匹配举例，该多个报文匹配编号为 1，3，5，17，18，23 以及 37 的规则，但是多个规则集合中，只有{1,5,18,23,37}这样的规则集合，那么，该多个报文匹配的规则集合就是{1,5,18,23,37}，而不是{1，3，5，17，18，23，37}。

一条字段规则用于指示一个字段的特征。一个字段的特征可以是一个字段的类型、长度和取值中的至少一种。例如，可以是一个字段的长度，一个字段的类型和长度，或者一个字段的类型、长度和取值。

握手报文中存在许多的长度。本申请描述的规则中，将长度视为整数，则匹配一个字段的长度，就是匹配该长度的值。例如可以是解析出要匹配的字段长度后，与用于匹配该字段的长度的规则进行比较运算，比较结果为“真（true）”则命中该规则，反之则不满足该规则。

其中匹配字段的长度的规则，可以使用等于、不等于（如不等，大于等于，小于等）或者某个范围（如用区间或者集合表示）来表示。例如，规则可以是某个握手报文中的某个字段的长度等于某值；也可以是某个握手报文中的某个字段的长度大于或者小于某值；还可以是某个握手报文中的某个字段的长度的值在某个范围内。

本申请描述的规则中，将字段的值视为字符串，则匹配一个字段的值，可以是整串匹配，也可以是子串匹配。

整串匹配就是该字段的所有值都要与规则中定义的一致才算匹配。整串匹配可以有多种表现形式。其中，对于嵌套的 TLV 结构，一种实现方式下，整串匹配只匹配与该字段的长度和属性属于同一层的值。以上文提及的 Client Hello 报文中的 Extension 举例，整串匹配匹配的是由 EC point formats Length:1 和 Elliptic curves point formats(1)这两部分组成的字符串，该字符串不包括 Elliptic curves point formats(1) 展开后的内容，如 EC point formats Length:1 Elliptic curves point formats(1)，即不将 Elliptic curves point formats(1)展开后的其他行的内容作为需要匹配的字符串。另一种实现方式下，对于嵌套的 TLV 结构，整串匹配不仅要匹配与该字段的长度和属性属于同一层的值，还要匹配该值中嵌套的未展开的字段。以上文提及的 Client Hello 报文中的 Extension 举例，整串匹配匹配的是 EC point formats Length:1 和 Elliptic curves point formats(1)，以及 Elliptic curves point formats(1)展开后的内容组成的字符串。

一个字符串中，任意个连续的字符组成的子序列称为该串的子串。子串匹配则是字段的值中，只需包括与字段规则中指示的值匹配的子串即可。例如，字段的值有多行，可以是其中的一行或者几行与规则匹配，也可以是某行中的某个子串，或者跨行的字符组成的子串，本申请不做限制。

本申请还描述一种被称为顺序规则的规则。一条顺序规则用于指示一个握手报文中多个字段的顺序。顺序规则中，以一个字段的属性（Type）来指示该字段，而不考虑字段的长度（Length）和值（Value）。也就是说，匹配一条顺序规则，要求该顺序规则所指示的报文中，包括一组字段，这组字段的属性以及这些属性在报文中排列

的顺序与该顺序规则中描述的相同，否则，该报文不匹配（或者称为未命中）该条顺序规则。

一种实现方式下，这多个字段往往是一报文中有一定相似性的一组字段，比如某个报文中的多个扩展（extension）字段。例如，Client Hello 报文中的 6 个扩展（Extension）字段可视为一组字段；Client Hello 报文中的多个 cipher suits 字段可视为一组字段，该多个 cipher suits 字段保存有多种加密算法的名称；Server Hello 报文中的多个扩展字段可视为一组字段，例如将包括有 8 个扩展字段(extension: 8 items)的 Certificate 报文中的某 5 个扩展字段视为一组字段，也就是说，顺序规则可以规定一个报文中有一定相似性的多个字段中的一部分字段的顺序。再例如，Certificate 报文中 subject 类型下的多个 RDNSequence item 字段也可视为一组字段，Certificate 报文中 issuer 类型下的多个 RDNSequence item 字段也可视为一组字段等等。

需说明的是，顺序规则可以是多种形式，本申请并不限制，只要规定了一个报文中多个字段顺序的规则都是顺序规则。下面是几类顺序规则的举例：

1. 对有 6 个扩展（Extension）字段的 Client Hello 报文，规定该 6 个扩展（Extension）字段的顺序。也就是规定一个报文中，有确定数目的多个字段的全部多个字段的顺序。
2. 对有 8 个扩展字段(extension: 8 items)的 Certificate 报文，规定该 Certificate 报文中 8 个扩展字段(extension: 8 items)中的第 1、第 3、第 4、第 5 和第 7 字段的类型。也就是规定一报文中一类字段中的某几个有确定位置的字段的类型。
3. 对有 8 个扩展字段(extension: 8 items)的 Certificate 报文，规定该 Certificate 报文中 8 个扩展字段(extension: 8 items)中 5 个连续字段的类型，而不限定该 5 个连续字段是 8 个扩展字段中的哪 5 个扩展字段。也就是规定一报文中一类字段中的一部分相对位置固定的字段的类型。
4. 规定 Certificate 报文中 subject 类型下的多个 RDNSequence item 字段的顺序，以及该 Certificate 报文中 issuer 类型下多个 RDNSequence item 字段。
5. 规定 ClientHello 报文中多个 cipher suits 字段的类型，以及多个 extension 字段的类型。

需要说明的是，上述规则中可能存在通配规则，对该通配的规则所对应的某个名称的报文，该名称的合法的报文都满足。另一方面，在匹配一字段规则的过程中，被匹配的报文中，该字段规则对应的字段缺省了类型、长度和值中的一个或者多个部分，而该字段规则中恰恰包括对缺省部分的指示，那么缺省部分可以默认为满足相应的规则。上述的缺省情况，可以是一个字段在协议中的定义就缺少一个部分，也可以是某些部分在协议中是定值，故不携带在报文中。

一种实现方式下该多个规则集合中的每一个都对应一个应用，而一个应用可对应多个规则集合。本申请对规则集合与应用之间的映射关系如何表现不做限制，可以使用多种数据结构来维护，例如表。以及，规则集合与应用间的映射关系，可以表现为规则集合与一应用间的映射，也可以表现为规则集合中的若干子集与多个应用集的对

应关系，子集可依据报文的名称划分，或者依据规则的种类划分等等，本申请不做限制。

上文中提及的规则，以及规则集合和应用的映射关系可以以库（如上文提及的匹配信息库）的方式预置在监控设备中，是通过机器学习得到的。一种实现方式下，还可以从其他的接口定期或者不定期向监控设备输入新的规则集合和应用的映射关系，或者监控设备定期或者不定期的采样以进行机器学习，以更新库。

事实上，可以通过机器学习的方式得到某个应用对应的规则集合，为描述方便，该应用称为目标应用。这个机器学习的过程可以是离线的。下面简单描述通过机器学习算法获得目标应用对应的规则集合的过程，该过程使用已知的目标应用的握手报文（即 client 和 server 在握手阶段传输的报文），以及其他应用的报文作为样本来学习。本申请记载的方案使用 C/S 模型，也就是说，一个样本对应的应用应当是唯一的。应理解，样本可以用多种方式获得，本申请不做限制。例如在测试环境中，在该目标应用的用户端与服务器之间传输数据流，其中包括使用该目标应用建立握手等，再将该数据流中的包拷贝或者拦截下来以供分析，当然，还需要一些其他应用的报文和该目标应用的握手报文共同作为样本输入到机器学习平台中训练，例如下文中所说的正样本和负样本。一种实现方式下，为了使学习出的规则集合能更准确更全面地识别出目标应用的报文，可以使测试环境中传输的报文更贴近实际需要识别的报文产生的条件，例如，使用实际组网中该目标应用可能的多种版本产生报文；使该目标应用运行在实际组网中该应用可能的运行的操作系统中，以收集样本；使用实际组网中该目标应用可能所在的国家或者地区的局域网发送或者接收样本等等。

再例如在实际组网环境中，拦截或者拷贝多个报文，依靠经验人为地将该目标应用的握手报文识别出来，做上标记，将其他应用的报文做上另一种标记以供机器学习平台处理。一种较为方便的方式是在终端侧将发送和接收的报文拷贝下来。用于训练规则的样本需要达到一定的数量，通常是几百或者几千个流的包。例如，一种实现方式下，用于训练规则的样本为，1000 个目标应用的流的报文以及 1200 个其他应用的流的报文。一种实现方式下，为了提高机器学习的效率，可以先人为地按照经验对得到的样本进行筛选。

将上述的样本使用安全加密传输协议进行解析，以得到报文中字段的长度、属性以及值，当然，还有握手报文中成组的字段的顺序。将这些解析后得到的信息（即握手报文中字段的长度、属性以及值，以及握手报文中成组的字段的顺序）使用机器学习平台进行处理，以训练出对应于该目标应用的一个或多个规则集合。其中，每个规则集合中包括一条或多条规则，只有数据流中的握手报文满足该规则集合中所有规则，该数据流才对应于该目标应用。另一方面，该目标应用的一个或者多个规则集合之间是或的关系，也就是说，只要一数据流中的握手报文满足其中任意一个规则集合，该数据流就对应于该应用。本申请对使用的机器学习算法不做限制，例如可以使用决策树类机器学习算法，具体的可以是随机森林算法，C4.5 算法，以及这些算法的变种等等。使用决策树类的机器学习算法，输出的就是一个或多个树状模型。一种实现方式下，对一个树状模型，树状模型中的每个节点表示一个应用集合，根节点表示未经过规则匹配的样本可能对应的应用集合，该应用集合可以预置，也可以根据此次机器

学习输入的样本设置。树状模型中的两个节点之间的连接表示一个或多个规则，也就是说，从根节点到叶子节点的方向，有连接关系的两个节点可解释为：从一个节点对应的应用集合中，根据该连接对应的规则，筛选出的应用子集，该应用子集可保存在该连接的另一端的节点。到叶子节点，就只剩下一个应用，可以在叶子节点上设置标识来标明对应目标应用的叶子节点。这样，每个叶子节点都对应一个规则集合，其中包括从该叶子节点到根节点的各连接所代表的规则。从一个表示该目标应用的叶子节点到根节点所经过的所有节点所对应的规则集合，就是该目标应用对应的一个规则集合。则，在一树状模型包括多个表示该目标应用的叶子节点的情况下，该树状模型就可以得到多个该目标应用对应的规则集合。

当然，可以用多组样本进行机器学习，该多组样本中包括该目标应用的不同的流的握手报文，这样学习得到的规则集合可能是不同的。可以将多组样本学习得到的不同的规则集合都作为用于识别该应用的规则集合，这样就可以使得确定出的该应用对应的规则集合更加全面，在识别网络中的报文时，能够提高识别率。

一种实现方式下，可以将从多个报文中解析出的上述信息以矩阵形式表达，以便使用机器学习算法对矩阵进行训练。例如矩阵的行可以用于描述一个流的一个或多个握手报文，矩阵的列包括用于描述一个报文中一个字段的长度、类型、值或者字段的顺序的列。样本中包括目标应用的握手报文，例如该目标应用为应用 A，还包括其他应用的报文，例如应用 B，应用 E 等，这样，应用 A 的握手报文就标记为正样本，其他应用的报文就标记为负样本。相应的，矩阵中也专用一列来记录每一行的报文是正样本还是负样本。

另一方面，在将通过机器学习方法得到多个规则集合，以及这些规则集合与应用的映射关系加入库时，可以做一些筛选和管理，以使得在识别数据流时更加准确，减少后期的再识别、再判断，这样就能使上述的方法更具实用性。例如，可能学习出的某个规则集合以及映射关系与库中原有的规则集合重复，那么就可以不添加，再例如，可能相同的规则集合在不同次的学习中学习出对应不同的应用，那么可以不添加该规则集合以及相应的映射关系，再例如，可能学习出的某个规则集合是库中原有的规则集合的子集，那么可以不添加该规则集合以及相应的映射关系，再例如可能学习出的某个规则集合包括与库中原有的规则集合相同的子集，那么可用该学习出的规则集合及相应的映射关系替代原本库中相应的规则集合，本申请对如何筛选和管理不做限制。

一种实现方式中，维护的库中的每个规则集合都只对应一个应用，每个规则集合都不重复，且各个规则集合之间不存在相互包含的关系。这样的库显然能够通过规则集合更准确地指示数据流对应的应用。

上述的机器学习方法并不限定处理何种应用的报文，由于样本可以通过上述多种方式得到，故对网络中拦截或者拷贝的握手报文，就可以用上述确定出的规则集合中的规则去匹配，可以大大提高那些现有技术无法识别或者识别准确率低的应用的加密数据流的识别率和准确率。例如，目前使用名为 meek 的 TLS 隧道技术混淆流量，以逃避防火墙的检测的应用有几十种，这类应用在网络中传输的报文中携带任意值的 SNI，而 CDN(Content Delivery Network, 内容分发网络)服务器上会部署多种应用的

数据，该 CDN 服务器的域名（通常携带在 Common Name 字段）不能指示应用。另一方面由于 CDN 服务器上会部署多种应用的数据，该 CDN 服务器作为 server 接收到 client 端的握手报文，即使该报文中 SNI 是错误的或不准确的，CDN 服务器仍然可能进行接下来的建立连接以及发送数据的操作。这样，通过解析 SNI 或 Common Name 携带的信息来确定报文对应的应用的方式，就无法识别或者无法准确识别网络中传输的这类应用的报文。这类应用中，常见的如 Tor、Psiphon 等应用被当做躲避网络审查的 VPN 工具。

下面，以 TLS 协议为例，描述识别加密数据流的方法，以及相关的技术细节。其中，可能使用 TLS 流来指代加密数据流。结合下文描述的技术细节、方法以及装置等，对使用其他安全加密传输协议的加密流，例如 SSL 协议、或 DTLS 协议等，其识别方法和很多实现方式以及说明都是类似的，也是本领域技术人员可以想到如何实施的，本申请为篇幅考虑，不再详细描述识别使用其他安全加密传输协议的加密流。

以 Psiphon 为例，按照上述方法对 Psiphon 一个版本发送的握手报文进行机器学习以提取该应用对应的规则集合，该例子中，Psiphon 应用通过 TLS 流传输数据。提取出的一个规则集合包括以下 8 个规则，也就是说，一个 TLS 流的握手报文满足下述 8 个规则，则这个 TLS 流是 Psiphon 的数据流。应理解，下述描述的各种规则的表达形式（如格式，连接各不同层间的属性所使用的符号，连接不同字段的值所使用的符号等），仅仅是为便于理解的例子，本申请并不限定规则的表达形式。其中，对于 Client Hello 报文，有以下三个规则，分别是:

（1）、ClientHello_len=185

该规则表示 Client Hello 报文中类型为 len 的字段的值为 185。其中，类型为 len 的字段表示 Client Hello 报文的长度，则该规则表示 ClientHello 报文的长度是 185 字节。

（2）、

ClientHello_cipher_suits=0xc02b_0xc00a_0xc009_0xc023_0xc007_0xc02f_0xc014_0xc013_0xc011_0xc005_0xc004_0xc003_0x0039_0x0033_0x003d_0x0035_0x003c_0x002f_0x000a_0x0005_0x0004_0x00ff

该规则表示 Client Hello 报文中类型为 cipher_suits 的字段的值等于等号右边的字符串，即等于

0xc02b_0xc00a_0xc009_0xc023_0xc007_0xc02_0xc014_0xc013_0xc011_0xc005_0xc004_0xc003_0x0039_0x0033_0x003d_0x0035_0x003c_0x002f_0x000a_0x0005_0x0004_0x00ff

（3）、ClientHello_extension_0x0023 =0

该规则表示 Client Hello 报文中类型为 0x0023 的扩展字段（extension）的值为 0。其中，这里的报文类型使用的是未解析的报文中携带的信息来指示的，而未转化成经解析得到的，由 0x0023 所指示的类型。

对于 ServerHello 报文，有以下一个规则:

（4）、ServerHello_len=80

该规则表示 Server Hello 报文的长度为 80

对于 Certificate 报文，有以下三个规则：

（5）、Certificate_subject_CN=ssl334326.cloudflaressl.com

该规则与传统方案类似，其中 CN 为 Certificate 报文中的 Common Name 字段的缩写，表示 Certificate 报文中的 Common Name 字段为特定的值，即 ssl334326.cloudflaressl.com。

（6）、Certificate_subject_publickey_algorithm=RSA

该规则表示 Certificate 报文中的 subject 类型下 publickey algorithm 字段的值为 RSA。

事实上，这个条件中的 publickey algorithm 类似版本号字段，是固定字段，其类型与长度都是固定值，在握手报文中是缺省的。故，该规则是针对上文记述的缺省字段的。

（7）、

Certificate_extension_order=authKeyIdExt_subjKeyIdExt_basicConstraintsExt_keyUsageExt_crlDistPointsExt_certPoliciesExt

该规则表示的是一个顺序规则，即 Certificate 报文中包括 6 个按照规则中顺序（order）出现的 6 个扩展（extension）字段，该顺序为 authKeyIdExt，subjKeyIdExt，basicConstraintsExt，keyUsageExt，crlDistPointsExt，certPoliciesExt

对于 ServerHelloDone 报文，有以下一个规则：

（8）、ServerHelloDone_ver=0x0303

该规则表示 ServerHelloDone 报文的类型为版本（ver）的字段取值为 0x0303。这个字段属于协议中默认必选的字段，类型固定，长度固定为 2 字节。这也可以说明为什么该条件只匹配该字段的版本和值。

上述 8 个规则包括在一个规则集合内，也就是说，一个 TLS 流中的握手报文需要全满足上述全部的 8 个规则，该 TLS 流才被认为对应 Psiphon 应用，或者说该 TLS 流才被认为匹配 Psiphon 应用。而现有技术中的方案，则是通过解析出握手报文中的 SNI 或者 Common Name 字段以得到 TLS 流的域名。然而，由于 Psiphon 的 SNI 随机变化，故不能够通过解析 SNI 的内容是被出一 TLS 流是否匹配 Psiphon 应用。另一方面，本方案同样使用了 Common Name 作为条件之一，但是由于 Psiphon 使用通用的 CDN 作为服务器，证书（Certificate）并不是 Psiphon 专用的，故按照现有技术中的仅依靠 Common Name 的内容很可能将一对应于 Psiphon 的 TLS 流识别成对应其他应用，识别的准确率很低。而本申请的方案中，Common Name 的内容与其他 7 个规则包括在一个规则集合内，只有一个 TLS 流满足这 8 个规则才与 Psiphon 匹配，故能够够准确地识别出匹配 Psiphon 应用的流。

使用本申请描述的机器学习方法分析多条 Psiphon 应用的 TLS 流的握手报文，可以获得多个与 Psiphon 应用对应的规则集合。例如，使用上述的包括 8 条规则的规则集合，在多种不同的网络环境下测试使用该规则集合识别网络中对应于 Psiphon 应用的流的效果，即在多种网络环境下使用 Psiphon 应用传输 TLS 流，使用该规则集合匹配这些网络环境中 Psiphon 应用传输的 TLS 流。例如，网络环境包括：使用在 Windows 7、Windows 10、Android4.4.2、Android5.1.1、Android6.0.0 五种不同的

操作系统下运行的 Psiphon 应用传输 TLS 流，Psiphon 用户端使用 65、91、103、108、112、113、114、123、125、128、130、以及 133 这多个版本，使用 6 个不同网络接入点分别传输 TLS 流，网络接入点具体位于中国、日本、美国、英国、新加坡、和荷兰。在上述多种网络环境下，对 Psiphon 应用的流的识别准确率为 100%，召回率（Recall Rate，也称查全率）99.19%。其中，准确率表示识别出的 Psiphon 应用的流确实是 Psiphon 应用的流的比例，即是否识别正确；召回率表示识别出的网络中的 Psiphon 应用的流占该识别过程中截获或者拷贝的数据流中 Psiphon 应用的流的比例，即将存在的目标应用的流出别出来了多少。可见，确定出的规则集合反映了 TLS 握手报文中字段的内容的组合（值，长度，属性，多个字段的顺序等）与 Psiphon 应用之间的映射关系，因此相比于现有的方案，使用本申请描述的规则集合去识别 TLS 流，能够大大提高 TLS 流的识别率和准确率，达到了准确识别的目的。

当然，正如前文和图 5b 描述的，一些情况下，只匹配一个握手报文，就能够识别该握手报文的加密流是否对应某个应用。例如，只匹配 ClientHello 报文，或者只匹配 ServerHello 报文,或者只匹配 Certificate 报文等。这种情况下，某个应用的某个规则集合中的规则都来自同一报文。下面描述一个应用的一个规则集合，该规则集合中，所有规则都来自 ClientHello 报文，也就是说，对一个数据流，只需其 ClientHello 报文匹配上下面的规则集合，该数据流就可被识别为对应该应用。该规则集合包括以下规则：

ClientHello_ver=0x0303

ClientHello_CipherSuitesLength=32

ClientHello_CipherSuites=cca8cca9c02fc030c02bc02cc013c009c014c00a009c009d002f0035c012000a

ClientHello_CompressionMethodsLength=1

ClientHello_CompressionMethods=0

ClientHello_ExtensionsNumbers=7

以上六条规则为该规则集合中的字段规则，对每一条规则的含义此处不再详细叙述，可以参考 TLS 协议中对 ClientHello 报文各字段的定义以及前文有关 Psiphon 应用中对规则的解释进行理解。以下一条规则为顺序规则，即规定了该 ClientHello 报文的一组扩展字段 Extensions 的顺序，其中以每个扩展字段的种类（Type）来表示该扩展字段。

Extensions_order=0x0005_0x000a_0x000b_0x0023_0x000d_0xff01_0x0012

例如，一 ClientHello 报文中包括下述一组扩展字段 Extensions，则该 ClientHello 报文匹配该顺序规则。

Extensions Items (Type, Length, Value): 0x0005 5 0100000000

Extensions Items (Type, Length, Value): 0x000a 10 0008001d001700180019

Extensions Items (Type, Length, Value): 0x000b 2 0100

Extensions Items (Type, Length, Value): 0x0023

Extensions Items (Type, Length, Value): 0x000d 14

000c040104030501050302010203

Extensions Items (Type, Length, Value): 0xff01 1 00
Extensions Items (Type, Length, Value): 0x0012 0

另一方面，某个应用的某个规则集合中的规则也可以只有一种，也就是说，该规则集合中都是顺序规则或者都是字段规则。即，对某些流，可以仅通过顺序规则或者字段规则识别。规则集合中只包括字段规则的一种实现方式下，包括多条字段规则，且至少一条字段规则中包括对长度的约束。下面描述一个例子，该例子中的规则集合包含一条规则，该规则为顺序规则。该规则集合为：

Certificate_subject_order=CountryName_StateOrProvinceName_LocalityName_OrganizationName_OrganizationUnitName_CommonName

一个流匹配该规则，则表示该流中的 Certificate 报文中的一组字段 subject 的顺序为 CountryName_StateOrProvinceName_LocalityName_OrganizationName_OrganizationUnitName_CommonName

例如包括下述的一组 subject 字段的 Certificate 报文就满足上述规则集合，

subject:rdnSequence(0)

◢ rdnSequence: 6 items (id-at-commonName=www.update.microsoft.com...)
▷ RDNSequence item: 1 item (id-at-countryName=US)
▷ RDNSequence item: 1 item （id-at-stateOrProvinceName=Washington）
▷ RDNSequence item: 1 item (id-at-localityName=Redmond)
▷ RDNSequence item: 1 item (id-at-organizationName=Microsoft)
▷ RDNSequence item: 1 item (id-at-organizationUnitName=DSP)
▷ RDNSequence item: 1 item (id-at-commonName=www.update.microsoft.com)

在恶意软件或者流氓软件批量生成证书（Certificate）的情况下，Certificate 报文中 Common Name 的值会被篡改或者伪造，按照现有技术，通过识别 Common Name 的值中的域名（如上述例子中的 microsoft）来识别软件，就很容易识别错误。而使用类似上述例子的规则集合，其中的规则限定了 Certificate 报文中 subject 字段的顺序，无论如何篡改这些字段的值，其顺序都不会改变，那么也就能识别出该报文正确对应的应用。

下面以 TLS 握手报文为例，描述监控设备识别加密数据流的过程，其中，监控设备接收到的加密流都使用 TLS 协议。应理解，使用其他安全加密传输协议的加密流，也可以采用类似的实现方式，本申请不再赘述。

由于一条字段规则或者顺序规则都是针对一个握手报文，故一种实现方式下，监控设备确定截获的是一个握手报文（如通过解析该握手报文），就匹配该握手报文而无需等待该数据流的其他握手报文。按照前文对握手阶段的示例性描述可知，握手报文的传输是有顺序的，则可以将接收到的握手报文解析后，将其与库中的规则集合进行匹配，如果根据库中规则集合与应用的映射关系，得出该握手报文匹配出的规则集合唯一地映射一个应用，那么该握手报文所在的流对应该应用，匹配过程结束，不再匹配后续得到的该流的握手报文，而认为往后该监控设备接收到的该流的报文都对应该应用，这样，可以更快速地识别流，且可以节约监控设备的资源。一种实现方式下，可以解析了截获的多个握手报文后，再进行匹配，匹配过程可以参考下文中匹配一个

握手报文的说明，例如对多个握手报文按照规则的种类匹配等。这样可以简化匹配流程，批量处理，无需对每个握手报文分别进行一次匹配流程。

由于 ClientHello 报文从时序上看是 TLS 握手过程中第一个传输的握手报文且必定在 TLS 握手过程中存在，一种实现方式下，故可以通过解析先判定出是否接受到一个流的 ClientHello 报文，这样，可以确定是否收到的是一个 TLS 流的报文，若是 TLS 流（即获得了该流的 ClientHello 报文），则从 ClientHello 报文开始匹配规则，显然可以更快地识别一部分流对应的应用，这些应用对应的规则集合中存在规则全部用于匹配 ClientHello 报文的规则集合，而若没有获得一个流的 ClientHello 报文，则该流不是 TLS 流，可直接判定无法使用规则集合识别，即匹配失败。事实上，在本申请所描述的技术方案中，很多种应用对应的规则集合中，都包括对应 ClientHello 报文中的字段的规则子集。当然，也有些应用对应的规则集合中不包括对应 ClientHello 报文中的字段的规则子集，这就意味着对这些应用，可以跳过对 ClientHello 报文的匹配。

具体的，可以在确定接收到一握手报文（可称为 A 报文，如 ClientHello 报文）后对其进行解析，得到其中的字段长度、属性和值，以及多个字段的顺序等信息，然后，使用库中的规则对这些信息进行匹配。可以是解析出握手报文的所有字段，也可以是解析出握手报文中的部分字段。一种实现方式下，在解析出该握手报文的名称，即确定出该握手报文是什么握手报文后，可以只解析库中保存的该类握手报文的规则涉及的字段。也就是说，可以从保存的规则集合与应用的映射中，得到该类握手报文的多条规则对应的字段，解析该报文中有规则对应的字段，这样，可以有目的选择报文需要解析的部分。减轻监控设备的负担，在实际网络环境这种多条流并行传输的场景下，能够提高对报文的处理效率。那么类似的，对接收到多个握手报文一起解析的情况，也可以在确定出该多个握手报文的名称后，根据库中保存的规则集合，或者根据用户或者业务需求指示的要识别的应用所对应的规则集合，有选择地解析这些规则集合中规则对应的报文，以及报文中的字段。

一种实现方式下，可以一条规则一条规则地匹配，匹配上一条规则，查询该规则对应的应用子集，待对该报文的规则匹配完成后，将得到的多个应用子集求交（∩），则得到该 A 报文匹配的应用集合，如果该应用集合中只有一个元素，那么该元素就是该 A 报文所在的 TLS 流对应的应用，识别过程结束；如果该应用集合中包括多个应用，可以再按照上述的方式或者下述的其他方式匹配该 A 报文所在的数据流的另一个报文（可称为 B 报文，如 ServerHello 报文或者 Certificate 报文），例如，仅匹配与上述应用集合对应的规则集合中，与 A 报文相关的规则可匹配 A 报文的规则集合，从而对上述的应用集合进一步限缩，如果该限缩后的应用集合中只有一个元素，那么该元素就是该 A 报文和 B 报文所在的 TLS 流对应的应用，识别过程结束；如果该限缩后的应用集合仍然包括多个应用，则可对该 A 报文和 B 报文所在的 TLS 流中的其他握手报文进行上述类似的匹配过程，直到识别 A 报文所在的 TLS 流对应的应用。

当然，如果上述过程中出现了为空的应用子集，那么可以结束该匹配流程，认为与库中的应用匹配失败。

一种实现方式下，也可以一类规则一类规则地进行匹配，一种实现方式下，将规则分为两类，即如上文所述的字段规则和顺序规则；另一种实现方式下，将规则分为

三类，为上文所述的顺序规则，字段规则中用于匹配长度的规则，以及字段规则中除用于匹配长度的规则之外的规则，当然，本申请对如何划分规则的种类并不限制。仍旧以 A 报文进行说明，其中，可以多类规则并行进行匹配，各得出一个匹配的规则子集，其中，规则子集可为空集，表示该报文在这类规则中没有匹配成功的规则，再将这多个规则子集求并（U）以得到该报文对应的规则集合；或者可以按照一定的顺序来匹配各类规则，不断将匹配的规则子集并规则集合中，直到得到 A 报文匹配的规则集合。例如。先匹配一类规则（称 A 类规则。可以是前文中提及的任一类规则），筛选出 A 报文匹配 A 类规则的规则子集 A，再匹配另一类规则（称 B 类规则，可以是前文中提及的任一类与 A 类规则不同的规则），将与 A 报文匹配的 B 类规则加入规则子集 A 中；或者匹配该另一类规则时，基于先前得出的规则子集 A，以及库中的规则集合，得到与该规则子集有组合关系的 B 类规则，如规则集合中包括规则子集 A 的有{{A},{B}，{C}}，{{A},{D}}，{{B},{C},{F}}。其中，规则子集{B}和{D}都是针对 A 报文的 B 类规则，{C}为针对 B 报文的一类规则，那么使用规则子集{B}和{D}对 A 报文进行匹配，发现规则子集{B}与 A 报文匹配，这样，得到 A 报文可能对应的规则集合{{A},{B}，{C}}；再用规则子集{C}中的规则对接收到的 B 报文进行匹配；如果匹配成功，则 A 报文和 B 报文所在的流对应的应用就是规则集合{{A},{B}，{C}},对应的应用，如果匹配失败，则 A 报文和 B 报文所在的流识别失败。

需要说明的是，可以按上文描述的依照获得握手报文的顺序，或者依照握手过程握手报文的交互顺序，依次进行匹配，这样的好处是收到时序靠前的握手报文后，不用等待收到后续的握手报文，就可以进行匹配，如果匹配成功，就可以不用考虑后续的握手报文而得出匹配的应用，这显然节省监控设备的资源，以及能够更快更及时地识别一个流，例如，可以在收到 ClientHello 报文后开始匹配。当然，也可以在接收到多个握手报文后一起匹配，例如在收到 ClientHello 报文和 ServerHello 报文后开始匹配，或者在收到 ClientHello 报文和 Certificate 报文后开始匹配，或者在收到 ClientHello 报文、ServerHello 报文和 Certificate 报文后开始匹配，如并行匹配多个握手报文各自的规则，再根据多个握手报文匹配的规则和匹配信息库中的规则集合，得到该多个握手报文匹配的规则集合，再得到该加密流对应的应用。或者并行匹配报文各自的规则，得到该多个报文对应的多个规则子集所对应的多个应用子集，对这多个应用子集求交。这是由于很多规则集合中的规则都是针对这三个报文的，并行处理可以提高效率。当然也可以在收到所有的握手报文后进行匹配等等。

综上，监控设备可以并行处理同一个流的多个报文，也可以并行用多种类的规则处理一个报文或者多个报文，参考上文，本领域人员应理解，匹配过程有多种实现方法，本申请对监控设备具体如何匹配一个流的一个或者多个报文的流程和时序不做限制。另一方面，监控设备可能同时收到多条流的握手报文，监控设备可以并行处理不同流的握手报文，请参考上文对一个流的握手报文的匹配的实现方式的描述，本申请同样对此不做限制，也不再赘述。

为了便于理解，下面结合图 6，再举例说明对一个 TLS 流的握手报文的匹配过程。应理解，对使用其他安全加密传输协议的数据加密流，等同的也可以使用下述的过程和匹配信息的管理方式，也就是使用链表、数组以及表，当然，本申请对库中的匹配

信息如何保存，如何管理并不限定。

握手过程中时序靠前的 ClientHello 报文（简写为 CH）、ServerHello 报文（简写为 SH）和 Certificate 报文（简写为 CER），这三个握手报文往往在规则集合中有对应的规则子集。下文描述的例子中，除上述三个握手报文有各自明确标号的规则子集外，对应其他握手报文的规则也可以有各自明确标号的规则子集。监控设备中保存着上述三个握手报文各自的多个规则子集，以及包括至少一个该规则子集的规则集合，以及规则集合与应用的映射关系。例如可以使用链表来存储规则集合，其中 CH、SH 和 CER 的规则子集可以各存储在一级链表中。但对一应用，其规则集合中往往包括后续的其他多种握手报文某几种握手报文的规则子集，故其他多种握手报文的规则子集可管理在同一级链表中，或者将该其他多种握手报文的规则子集用另一种数据结构进行管理。例如使用数组。

这样，接收到一个流的上述握手报文，就可以与规则子集进行匹配，返回命中的规则子集的编号。每匹配一个流的握手报文，检查该流握手报文命中的规则子集的并集是否能唯一对应一个应用，若能，则对该流识别成功，结束匹配流程；若对应一个应用集合，则获得该流的其他握手报文后继续匹配；若没有对应的应用，则匹配失败或者在这个握手报文为 CH。SH 以及 CER 包之后的握手报文的情况下，继续匹配该流的其他握手报文，直到匹配出对应的应用或者匹配尽该流的所有报文。具体的，可以是匹配出后续的单个握手报文的规则子集后，查询监控设备中保存的包括该规则子集的规则集合，再确定这些规则集合中是否有能匹配该流的多个报文的，如果有，则匹配成功，如果没有则说明需要更换该后续报文继续匹配。

以及，监控设备中使用 SeqInfo 数据结构管理规则集合，也就是通过链表，按照握手过程中握手报文的交互顺序排列一个规则集合中的规则子集。例如链表 CH_SH，表示规则集合中包括对应 CH 的规则子集和 SH 的规则子集但不包括对应 CER 的规则子集，链表 CH_CER 表示规则集合中包括对应 CH 的规则子集和对应 CER 的规则子集但不包括对应 SH 的规则子集，以及 链表 CH_SH_CER 表示规则集合中包括对应 CH 的规则子集、对应 CER 的规则子集以及对应 SH 的规则子集上述的规则子集，这个例子中，对不同的 CH 对应的规则子集的标号，其后挂载的 CH_SH, CH_CER,和 CH_SH_CER 链表是不同的，可以是其中的至少一个，当然，CH_SH, CH_CER,和 CH_SH_CER 链表也可以是该规则子集后挂载链表的一部分，其后还可以带有对应其他握手报文或者其他规则，例如，该其他规则可以是用于限定一个流的多个握手报文顺序的规则等。也可以不挂载任何链表，还可以是包括对应其他握手报文的链表，由于应用多种多样，其对应的规则集合的链表形式也很多样，本申请对规则集合可能的链表形式此不做限制。这样，当解析出一个 CH 对应的规则子集后，再接收到其他的报文，就可以根据该握手报文的名称只取该规则子集挂载的链表中的一部分进行匹配，可以缩小查询范围，更快速匹配出最终对应的应用。

下面描述在使用上述的数据结构维护规则集合和应用的映射关系的情况下，匹配一个数据流的过程。其中，对应 CH 的规则子集具有以 1 开头的编号，1 后的数字为该规则子集在对应 CH 的规则子集中的编号，应理解，这个编号可以是连续或者不连续的。类似的，对应 SH 的规则子集具有以 2 开头的编号，对应 CER 的规则子集具

有以 3 为开头的编号。例如，本例子中，监控设备保存如表 1 所示的规则集合与应用的映射关系，其中表 1 中只列举了包括规则子集 10+20+30 的规则集合，一列表示一种握手报文的规则子集，表中列名表示该列对应的握手报文，一行表示一个规则集合，表中以 SeqRule 表示，其后的数字为该行表示的规则集合的索引值（index），NA 表示该行的规则集合中不包括此列的握手报文对应的规则子集。表 1 只是为了便于理解，实际实现中，规则集合与规则子集之间的关系并不一定以表的形式管理，本申请对规则集合与规则子集的映射关系的管理形式不做限定。

表 1

	CH	SH	CER	SKE	SHD	CKE
SeqRule0	10	20	30	40	50	60
SeqRule1	10	20	30	41	50	61
SeqRule2	10	20	30	41	51	61
SeqRule3	10	20	30	41	50	NA
SeqRule4	10	20	31	NA	NA	NA

图 6 中示意了一些匹配过程中可能涉及的规则子集，其中箭头表示从一矩形框表示的节点指向下一步可能匹配到的节点（也由一矩形框表示）。以及表中的 1,2,3,7，8，9 表示一规则集合的索引值。

如图 6 所示，经过查询，一个流的 CH 包匹配上了规则子集 10，图中表示为 CH=10，那么找到 CH=10 所在的 SeqInfo 数据结构，其上挂载着三个链表，CH_SH, CH_CER, CH_SH_CER，显然这三个链表所表示的规则集合包括相同的对应 CH 的规则子集。监控设备收到该流的 SH 包后进行解析，匹配 CH_SH 和 CH_SH_CER 两个链表中 SH 对应的部分，得到该 SH 包匹配上了规则子集 20，图中表示为 SH=20，实际上此时匹配出链表上对应 CH=10 且 SH=20（对应表示该节点满足 CH=10 且 SH=20 的条件）的节点；监控设备又收到该流的 CER 包，则解析 CER 包后匹配该节点后的 CH_SH_CER 链表，得到该 CER 包匹配上规则子集 30，图中表示为 CER=30，则此时匹配出 CH_SH_CER 链表上对应 CH=10、 SH=20 且 CER=30 的节点。那么应理解，同时维护有类似 10+20+30 和 10+20+30+40 的两个规则集合是不可取的，这样显然满足 10+20+30 的流可能匹配到两个应用。以及，图中也表示出了 CH=10 后带的 CH_SH 链表中，可以携带多个 SH 的节点，这样可以在表示 SH 的节点中查询，例如图中描述的从代表 SH=20 的节点通过 NEXT 指示的 SH=21 的节点，再例如 CH_SH_CER 链表中，SH=20 且 CER=30 的节点后，通过 NEXT 指示的 SH=20 且 CER=31 的节点。

这个例子中，其余类型的握手报文对应的规则子集以数组形式存储，例如该数组可称为 ExtSeqInfo 数组。其中，上述对应 CH=10、SH=20 且 CER=30 的节点还对应

有数组结构，也就是说包括 CH=10、SH=20 且 CER=30 这三个规则子集的规则集合中还包括其他规则或者规则子集。那么该流的匹配过程没有结束，监控设备接收到该流的握手报文还要解析以及匹配，确定出一符合条件的规则子集，再确定是否有与该多个规则子集的并集相同的规则集合，这里要说明的是，由于握手报文很多，而一个规则集合中往往只包括后续部分握手报文的子集，故在查询出 CH、SH 以及 CER 之后的单个握手报文对应的规则子集后，如果监控设备中并不存在与该流匹配的规则子集组成的并集相同的规则集合，可以舍弃这个握手报文的规则子集，继续匹配后续握手报文。

例如，接收到该流的 ServerKeyExchange 包（简称 SKE），匹配出该 SKE 命中规则子集 40，遍历 ExtSeqInfo 数组，查询到包含规则子集 40 的规则集合，得到索引值（index）为 0 和 2 的规则集合中包括该规则子集，然而也发现这两个规则集合都没有还包括规则子集 10+20+30，故该流没有匹配上规则集合 0 和 2，还要继续匹配，可先记录下规则集合 0 和 2 的序号。应理解，这个例子中，匹配 CER 包后匹配 SKE 包，是该流的握手流程中传输完 CER 包后是 SKE 包。

类似的，再解析接收到的该流的 ServerHelloDone 包（简称 SHD），匹配出该 SKE 命中规则子集 50，遍历 ExtSeqInfo 数组，查询到包含规则子集 50 的规则集合，得到索引值（index）为 0 和 1 的规则集合中包括该规则子集，然而也发现这两个规则集合都没有还包括规则子集 10+20+30，故该流没有匹配上规则集合 0 和 1，还要继续匹配，可先记录下规则集合 0 和 1 的序号。应理解，确认的是规则集合中是否还包括规则子集 10+20+30，而不是 10+20+30+40，正如前文所说，由于匹配 SKE 时没有匹配到包括 10+20+30+40 的规则集合，在匹配 SHD 时忽略 SKE 命中的规则子集。当然，如果该例子中使用的规则集合中，不包括某些握手报文对应的规则子集，那么匹配过程可以跳过这个握手报文。

接着，又解析接收到的该流的 ClientKeyExchange 包（简称 CKE），匹配出该 CKE 命中规则子集 60，遍历 ExtSeqInfo 数组，查询到包含规则子集 60 的规则集合，得到索引值（index）为 0 和 3 的规则集合中包括该规则子集，其中规则集合 3 还包括规则子集 10+20+30，故该流匹配上规则集合 3。

由于现有技术中，可以以一条 TLS 流的握手报文传输的先后顺序作为规则来识别流，故本申请描述的规则集合还可以包括这类指示多个握手报文的传输顺序的规则，例如称为 MsgSeq 规则，图中也表示了一编号为 122334 的 MsgSeq 规则。

因此，如果上述规则集合 3 中包括这类指示多个握手报文的传输顺序的规则，则应继续检验该流的握手报文的顺序是否满足规则集合 3 中包括 MsgSeq 规则，如果满足，则该 TLS 流匹配规则集合 3，规则集合 3 对应的应用就是该 TLSL 流对应的应用。如果不满足，则为命中规则集合 3，应继续解析接收到的该流的其他握手报文。当然，如果上述规则集合 3 中不包括 MsgSeq 规则，那么该 TLS 流匹配规则集合 3，规则集合 3 对应的应用就是该 TLSL 流对应的应用。

图 7 描述了一种识别加密数据流的装置 700 的示意图，该装置包括解析模块 701，匹配模块 702 和确定模块 703。该识别加密数据流的装置 700 可以实现通过加密数据流的一个或者多个握手报文识别该加密数据流对应的应用。也就是说，该识别加密数

据流的装置 700 可以实现本申请上述的任一种方法。以及，图 3 对应的监控设备也可以看做是该装置的一种实现，其中，包处理器对应解析模块 701，流量识别引擎对应匹配模块 702 和确定模块 703。

一方面，解析模块 701 用于根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的多个握手报文，以获得所述多个握手报文包含的多个字段；匹配模块 702 用于根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个握手报文中多个字段的顺序，所述多个握手报文中的字段满足所述匹配的规则集合中的规则；确定模块 703 用于根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用。

该方面的各种名词解释、实现方式，以及有益效果，请参见前文相应的方法。例如图 5a 所示的方法。其中，匹配模块 702 可以用于执行前文中描述的“根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合”这一步骤对应的各种实现方式。确定模块 703 可以用于执行前文中描述的与“根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用”这一步骤对应的各种实现方式。

另一方面，解析模块 701 用于根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的一握手报文，以获得所述握手报文包含的多个字段；匹配模块 702 用于根据所述多个字段，从多个规则集合中确定出所述握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个报文中多个字段的顺序，所述握手报文中的字段满足所述匹配的规则集合中的规则；确定模块 703 用于根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用。

该方面的各种名词解释、实现方式，以及有益效果，请参见前文相应的方法。例如图 5b 所示的方法。其中，匹配模块 702 可以用于执行前文中描述的“根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合”这一步骤对应的各种实现方式。确定模块 703 可以用于执行前文中描述的与“根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用”这一步骤对应的各种实现方式。

以及，上述两方面的装置中，都可以还包括一个训练模块，该训练模块未体现在图 7 中，该训练模块用于通过机器学习算法训练多个样本，以得到目标应用对应的至少一个规则集合，其中，所述多个样本已知是否为所述目标应用对应的加密数据流的握手报文，所述多个样本中包括所述目标应用对应的加密数据流的握手报文。

以及，上述两方面的装置中，都可以保存有前文描述的多个规则集合，以及规则集合和应用的映射关系。例如，前文描述的识别信息库。

如图 8 所示，本发明实施例提供一种物理机，该物理机可以用于执行上述的任一种方法，例如图 5a、图 5b 对应的方法等。该物理机包括处理电路 40、接口电路 41、存储介质 42 和系统总线 43。以及，图 7 对应的装置也可以布局在该物理机上，处理

电路 40 通过运行存储介质 42 中的指令实现图 7 对应的装置 700 的各个模块。当然，一种实现方式下，存储介质 42 中还保存有上文提及识别信息库。以及，该物理机可以是如图 3 所示的监控设备，则图 3 中的输入接口和输出接口可由图 8 中的接口电路 41 实现，图 3 中的包处理器、流处理器、流量识别引擎以及识别信息库，可以由处理电路 40 通过运行存储介质 42 中的指令实现。

存储介质 42 用于存储计算机执行指令，处理电路 40、接口电路 41 和存储介质 42 通过系统总线 43 相互连接，当该物理机运行时，处理电路 40 执行存储介质 42 存储的计算机执行指令，以使该物理机执行本发明实施例提供的方法。

处理电路 40 可以通过一个或多个处理器实现，图 8 仅以一个处理器为例进行示例性的说明。处理电路 40 可以为中央处理器（英文：central processing unit，缩写：CPU）。处理电路 40 还可以为其他通用处理器、数字信号处理器（英文：digital signal processing，缩写：DSP）、专用集成电路（英文：application specific integrated circuit，缩写：ASIC）、现场可编程门阵列（英文：field-programmable gate array，缩写：FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

接口电路 41 具体可以是物理机上的通信接口。该通信接口可以为无线通信接口。例如，无线通信接口可以是物理机的无线模块等。处理电路 40 通过接口电路 41 与其他设备，例如其他物理机之间进行数据的收发。

存储介质 42 可以包括易失性存储器（英文：volatile memory），例如随机存取存储器（英文：random-access memory，缩写：RAM）；存储介质 42 也可以包括非易失性存储器（英文：non-volatile memory），例如只读存储器（英文：read-only memory，缩写：ROM），快闪存储器（英文：flash memory），硬盘（英文：hard disk drive，缩写：HDD）或固态硬盘（英文：solid-state drive，缩写：SSD）；存储介质 42 还可以包括上述种类的存储器的组合。

存储介质 42 可以包括底层存储介质和内存。其中，内存耦合至底层存储介质，用于作为底层存储介质的缓存。

系统总线 43 可以包括数据总线、电源总线、控制总线和信号状态总线等。本实施例中为了清楚说明，在图 8 中将各种总线都示意为系统总线 43。

可选的，本实施例还提供一种可读存储介质，该可读存储介质包括计算机执行指令，当物理机运行时，物理机的处理器执行该计算机执行指令，以使物理机执行本发明实施例提供的任一种方法。

可选的，本实施例中的可读存储介质可以为上述如图 8 所示的存储介质 42。

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，为描述的方便和简洁，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将装置的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。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。

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，装置和方法，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，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。比如，匹配模块 702 和确定模块 703 可以是一个模块，如上文图 3 中提及的流量识别引擎。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接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本发明实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，(例如，软盘、硬盘、磁带)、光介质(例如，DVD)、或者半导体介质（例如固态硬盘 Solid State Disk (SSD)，相变存储器）等。

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，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将装置的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。上述描述的系统，装置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，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，所揭露的系统，装置和方法，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，所述模块或模块的划分，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系统，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接口，装置或模块的间接耦合或通信连接。

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模块，即可以位于一个地方，或者也可以分布到多个网络模块上。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。上述集成的模块可以采用软件功能模块的形式实现。

所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，该技术方案的全部或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一个存储介质中，

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）或处理器执行本发明各个实施例所述方法的全部或部分步骤。所述存储介质是非短暂性（英文：non-transitory）介质，包括：快闪存储器、移动硬盘、只读存储器、随机存取存储器、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述，仅为本申请记载的发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应所述以权利要求的保护范围为准。

权 利 要 求 书

1.一种识别加密数据流的方法，其特征在于，所述方法包括:

根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的多个握手报文，以获得所述多个握手报文包含的多个字段;

根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个握手报文中多个字段的顺序，所述多个握手报文中的字段满足所述匹配的规则集合中的规则;

根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用。

2.根据权利要求 1 所述的方法，对一条字段规则，所指示的字段的特征为字段的长度，或者字段的类型，或者字段的长度和类型，或者字段的类型和值，或者字段的长度、类型和值。

3.根据权利要求 1 或 2 任一所述的方法，其特征在于，所述多个字段包括多个分组，所述多个分组中的每个分组对应一个握手报文，所述根据所述多个字段，从多个规则集合中确定所述多个握手报文所匹配的规则集合，包括:

按照所述多个握手报文的接收顺序，将所述多个分组与所述多个规则集合中的规则匹配，以从所述多个规则集合中得到所述多个握手报文所匹配的规则集合。

4.根据权利要求 1 到 3 任一所述的方法，其特征在于，所述多个规则集合中包括字段规则和顺序规则，所述根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合包括:

将所述多个字段与多个规则集合中的字段规则和顺序规则分别进行匹配，以从多个规则集合中确定出所述多个握手报文所匹配的规则集合。

5.根据权利要求 1 到 4 任一权利要求所述的方法，其特征在于，所述匹配的规则集合包括多个子集，所述多个子集中的每个子集对应至少一个应用，所述根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用，包括:

根据所述多个子集中每个子集与应用的映射关系，得到所述多个子集对应的多个应用集合;

求所述多个应用集合的交集，以得到所述规则集合对应的唯一的应用，所述应用为所述加密数据流对应的应用。

6. 根据权利要求 1 到 5 任一权利要求所述的方法，其特征在于，所述多个规则集合中包括以链表形式保存的多条规则，其中，所述链表中的每一个节点保存对应一个握手报文的名称的规则。

7.根据权利要求 1 到 6 任一权利要求所述的方法，其特征在于，所述解析一加密数据流的多个握手报文，包括:

对所述加密数据流的多个握手报文中的每一个握手报文，执行以下操作:

解析所述每一个握手报文的名称;

根据所述每一个握手报文的名称，从所述多个规则集合中，确定出对应于所述每一个握手报文的名称的至少一条规则;

解析所述每一个握手报文中，所述至少一条规则所指示的字段。

8.根据权利要求 1 到 7 任一权利要求所述的方法，所述多个握手报文为安全套接字

层 SSL 握手报文、传输层安全 TLS 握手报文或者数据报传输层安全 DTLS 握手报文。

9.根据权利要求 1 到 8 任一权利要求所述的方法，其特征在于，所述方法还包括:

通过机器学习算法训练多条加密数据流的多个样本，以得到目标应用对应的至少一个规则集合，其中，所述多个样本已知是否为所述目标应用对应的加密数据流的握手报文，所述多个样本中包括所述目标应用对应的加密数据流的握手报文。

10.一种识别加密数据流的方法，其特征在于，所述方法包括:

根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的一握手报文，以获得所述握手报文包含的多个字段;

根据所述多个字段，从多个规则集合中确定出所述握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个数据包中多个字段的顺序，所述握手报文中的字段满足所述匹配的规则集合中的规则;

根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用。

11.根据权利要求 10 所述的方法，对一条字段规则，所指示的字段的特征为字段的长度，或者字段的类型，或者字段的长度和类型，或者字段的类型和值，或者字段的长度、类型和值。

12.根据权利要求 10 或 11 所述的方法，其特征在于，所述多个规则集合是对应于所述握手报文的名称的规则集合。

13.根据权利要求 10 到 12 任一所述的方法，其特征在于，所述多个规则集合中包括字段规则和顺序规则，所述根据所述多个字段，从多个规则集合中确定出所述握手报文所匹配的规则集合，包括:

将所述多个字段与多个规则集合中的字段规则和顺序规则分别进行匹配，以从多个规则集合中确定出所述握手报文所匹配的规则集合。

14.根据权利要求 10 到 13 任一权利要求所述的方法，所述握手报文为安全套接字层 SSL 握手报文、传输层安全 TLS 握手报文或者数据报传输层安全 DTLS 握手报文。

15.根据权利要求 10 到 14 任一权利要求所述的方法，其特征在于，所述方法还包括:

通过机器学习算法训练多个样本，以得到目标应用对应的至少一个规则集合，其中，所述多个样本已知是否为所述目标应用对应的加密数据流的握手报文，所述多个样本中包括所述目标应用对应的加密数据流的握手报文。

16.一种识别加密数据流的装置，其特征在于，所述装置包括:

解析模块，所述解析模块用于根据一加密数据流对应的安全加密传输协议，解析所述加密数据流的多个握手报文，以获得所述多个握手报文包含的多个字段;

匹配模块，所述匹配模块用于根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个握手报文中多个字段的顺序，所述多个握手报文中的字段满足所述匹配的规则集合中的规则;

确定模块，所述确定模块用于根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用。

17.根据权利要求 16 所述的装置，对一条字段规则，所指示的字段的特征为字段的长度，或者字段的类型，或者字段的长度和类型，或者字段的类型和值，或者字段的长度、类型和值。

18.根据权利要求 16 或 17 任一所述的装置，其特征在于，所述多个字段包括多个分组，所述多个分组中的每个分组对应一个握手报文，在根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合的方面，所述匹配模块用于：

按照所述多个握手报文的接收顺序，将所述多个分组与所述多个规则集合中的规则匹配，以从所述多个规则集合中得到所述多个握手报文所匹配的规则集合。

19.根据权利要求 16 到 18 任一所述的装置，其特征在于，所述多个规则集合中包括字段规则和顺序规则，在根据所述多个字段，从多个规则集合中确定出所述多个握手报文所匹配的规则集合的方面，所述匹配模块用于：

将所述多个字段与多个规则集合中的字段规则和顺序规则分别进行匹配，以从多个规则集合中确定出所述多个握手报文所匹配的规则集合。

20.根据权利要求 16 到 19 任一权利要求所述的装置，其特征在于，所述匹配的规则集合包括多个子集，所述多个子集中的每个子集对应至少一个应用，在根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用的方面，所述确定模块用于：

根据所述多个子集中每个子集与应用的映射关系，得到所述多个子集对应的多个应用集合；求所述多个应用集合的交集，以得到所述规则集合对应的唯一的应用，所述应用为所述加密数据流对应的应用。

21. 根据权利要求 16 到 20 任一权利要求所述的装置，其特征在于，所述多个规则集合中包括以链表形式保存的多条规则，其中，所述链表中的每一个节点保存对应一个握手报文的名称的规则。

22.根据权利要求 16 到 21 任一权利要求所述的装置，其特征在于，在解析一加密数据流的多个握手报文的方面，所述解析模块用于：

对所述加密数据流的多个握手报文中的每一个握手报文，执行以下操作：

解析所述每一个握手报文的名称；

根据所述每一个握手报文的名称，从所述多个规则集合中，确定出对应于所述每一个握手报文的名称的至少一条规则；

解析所述每一个握手报文中，所述至少一条规则所指示的字段。

23.根据权利要求 16 到 22 任一权利要求所述的装置，所述多个握手报文为安全套接字层 SSL 握手报文、传输层安全 TLS 握手报文或者数据报传输层安全 DTLS 握手报文。

24.根据权利要求 1 到 8 任一权利要求所述的装置，其特征在于，所述装置还包括训练模块，所述训练模块用于通过机器学习算法训练多条加密数据流的多个样本，以得到目标应用对应的至少一个规则集合，其中，所述多个样本已知是否为所述目标应用对应的加密数据流的握手报文，所述多个样本中包括所述目标应用对应的加密数据流的握手报文。

25.一种识别加密数据流的装置，其特征在于，所述装置包括：

解析模块，所述解析模块用于根据一加密数据流对应的安全加密传输协议，解析所

述加密数据流的一握手报文，以获得所述握手报文包含的多个字段；

匹配模块，所述匹配模块用于根据所述多个字段，从多个规则集合中确定出所述握手报文所匹配的规则集合，其中，所述多个规则集合中的每个规则集合包括字段规则和顺序规则中的至少一种规则，所述字段规则用于指示一个字段的特征，所述顺序规则用于指示一个数据包中多个字段的顺序，所述握手报文中的字段满足所述匹配的规则集合中的规则；

确定模块，所述确定模块用于根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用。

26.根据权利要求 25 所述的装置，对一条字段规则，所指示的字段的特征为字段的长度，或者字段的类型，或者字段的长度和类型，或者字段的类型和值，或者字段的长度、类型和值。

27.根据权利要求 25 或 26 所述的装置，其特征在于，所述多个规则集合是对应于所述握手报文的名称的规则集合。

28.根据权利要求 25 到 27 任一所述的装置，其特征在于，所述多个规则集合中包括字段规则和顺序规则，在根据所述多个字段，从多个规则集合中确定出所述握手报文所匹配的规则集合的方面，所述匹配模块用于将所述多个字段与多个规则集合中的字段规则和顺序规则分别进行匹配，以从多个规则集合中确定出所述握手报文所匹配的规则集合。

29.根据权利要求 25 到 28 任一权利要求所述的装置，所述握手报文为安全套接字层 SSL 握手报文、传输层安全 TLS 握手报文或者数据报传输层安全 DTLS 握手报文。

30.根据权利要求 25 到 29 任一权利要求所述的装置，其特征在于，所述装置还包括训练模块，所述训练模块用于通过机器学习算法训练多个样本，以得到目标应用对应的至少一个规则集合，其中，所述多个样本已知是否为所述目标应用对应的加密数据流的握手报文，所述多个样本中包括所述目标应用对应的加密数据流的握手报文。

31.一种识别加密数据流的设备，其特征在于，所述设备包括处理电路、接口电路和存储介质，所述接口电路用于通过所述存储介质中的指令与其他设备收发数据包，所述处理电路用于运行所述存储介质中的指令控制所述接口电路，以实现如权利要求 1 到 9 所述的方法。

32.一种识别加密数据流的设备，其特征在于，所述设备包括处理电路、接口电路和存储介质，所述接口电路用于通过所述存储介质中的指令与其他设备收发数据包，所述处理电路用于运行所述存储介质中的指令控制所述接口电路，以实现如权利要求 10 到 15 所述的方法。

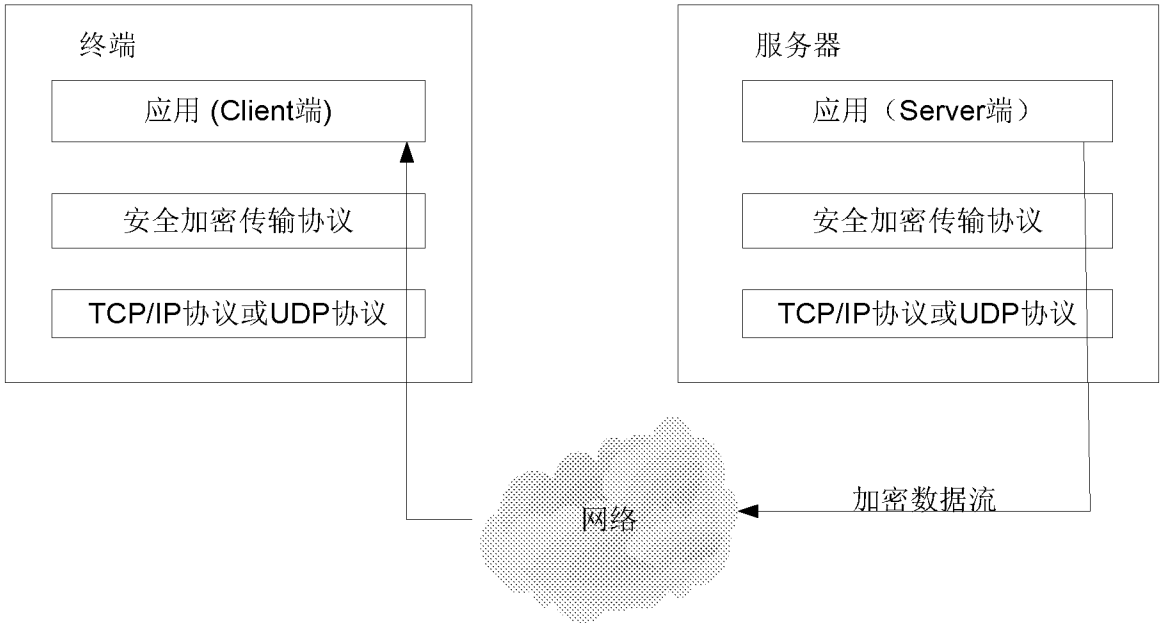
终端
应用 (Client端)
安全加密传输协议
TCP/IP协议或UDP协议
服务器
应用（Server端）
安全加密传输协议
TCP/IP协议或UDP协议
网络
加密数据流

图 1

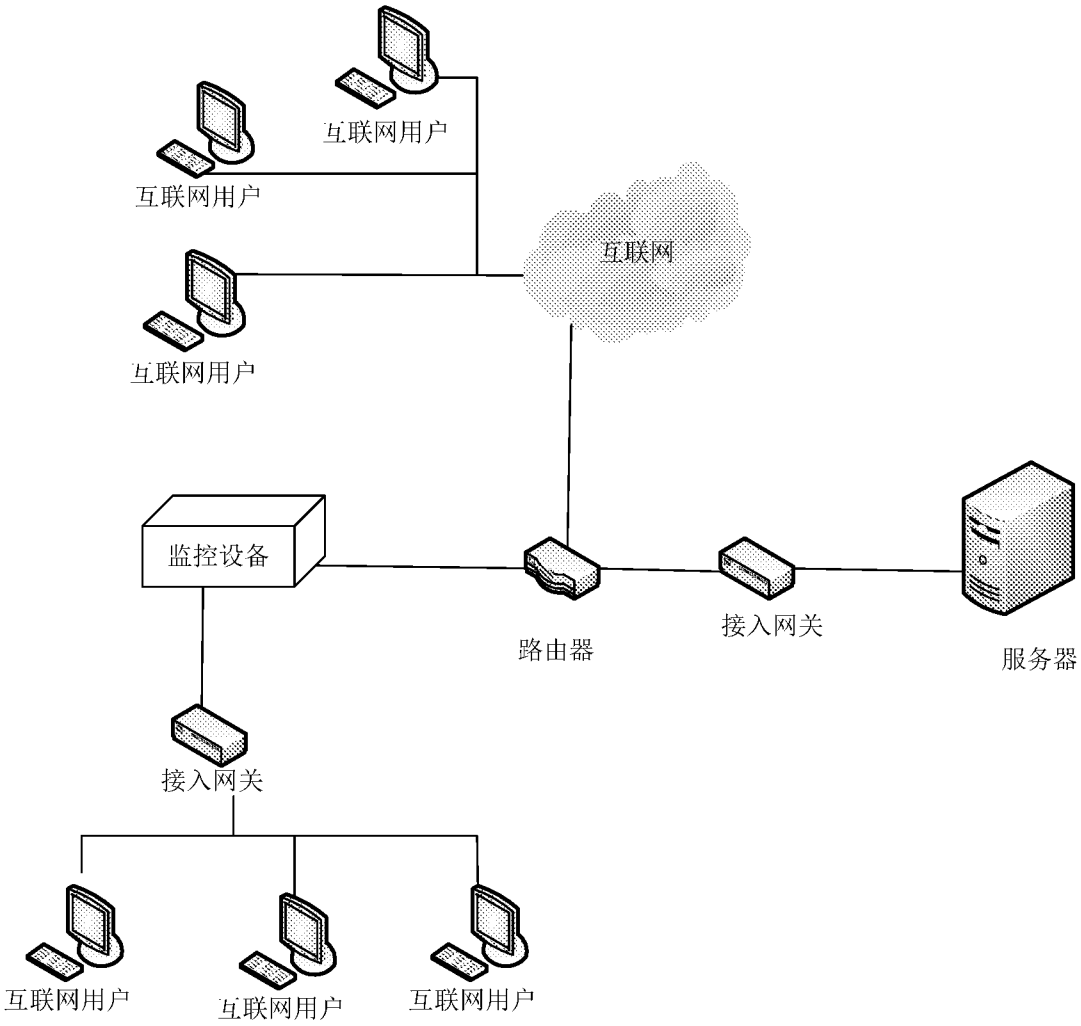
互联网用户
互联网用户
互联网
互联网用户
监控设备
路由器
接入网关
服务器
接入网关
互联网用户
互联网用户
互联网用户

图 2

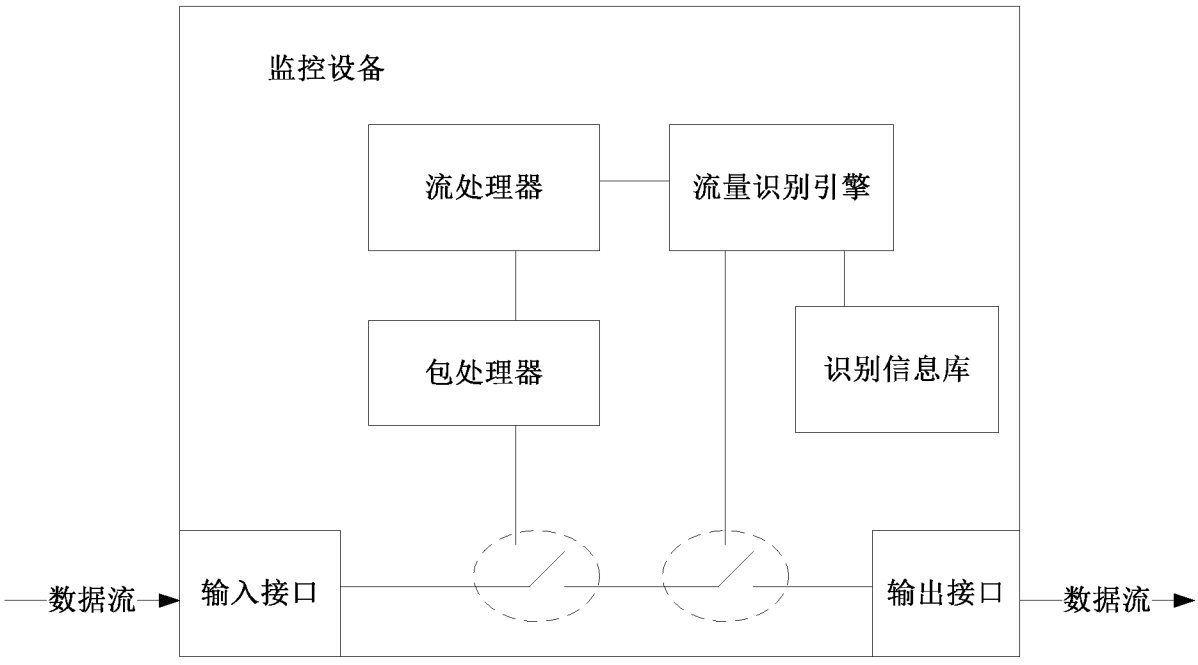
监控设备
流处理器
流量识别引擎
包处理器
识别信息库
数据流
输入接口
输出接口
数据流

图 3

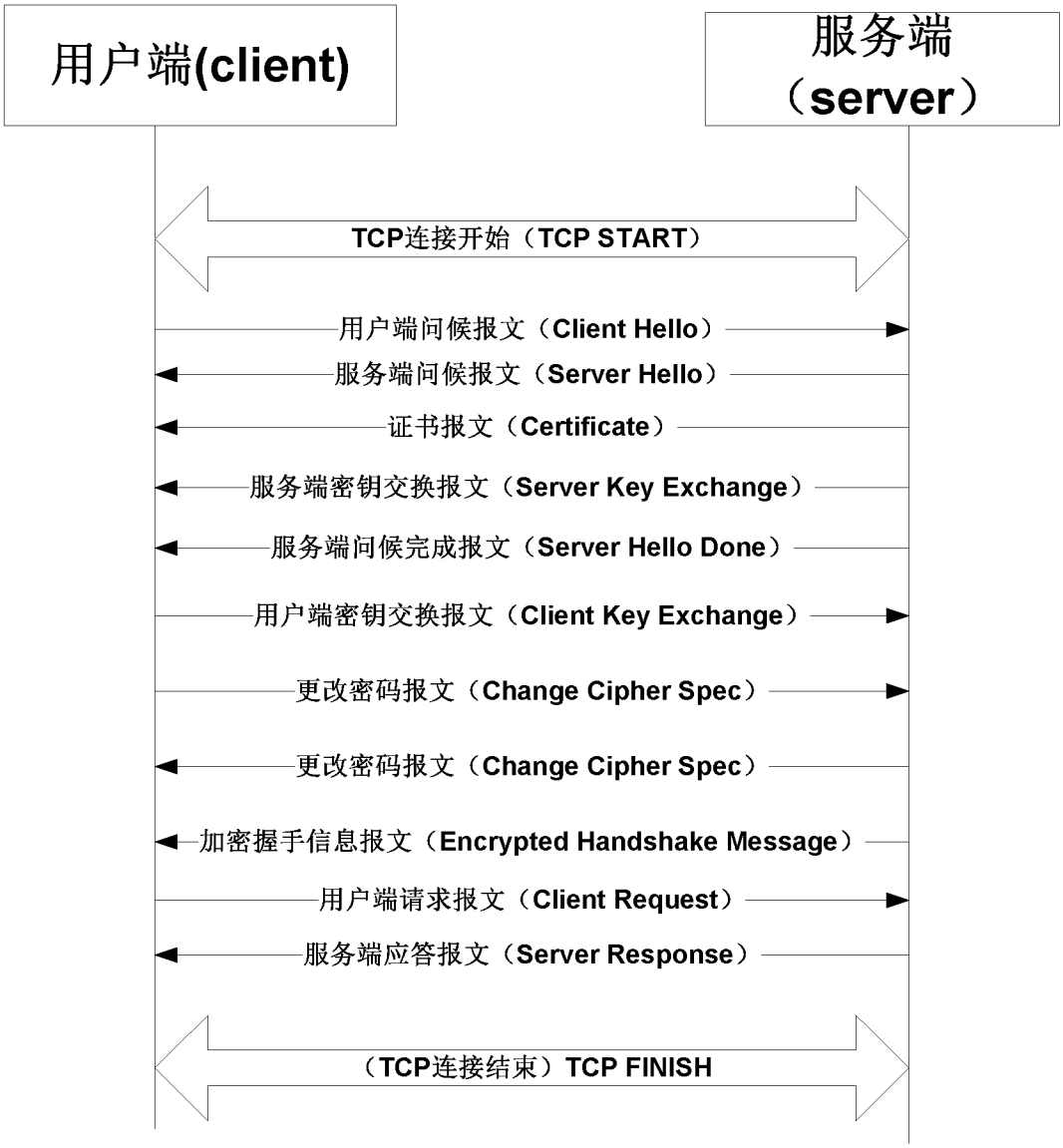
用户端(client)
服务端
（server）
TCP连接开始（TCP START）
用户端问候报文（Client Hello）
服务端问候报文（Server Hello）
证书报文（Certificate）
服务端密钥交换报文（Server Key Exchange）
服务端问候完成报文（Server Hello Done）
用户端密钥交换报文（Client Key Exchange）
更改密码报文（Change Cipher Spec）
更改密码报文（Change Cipher Spec）
加密握手信息报文（Encrypted Handshake Message）
用户端请求报文（Client Request）
服务端应答报文（Server Response）
（TCP连接结束）TCP FINISH

图 4a

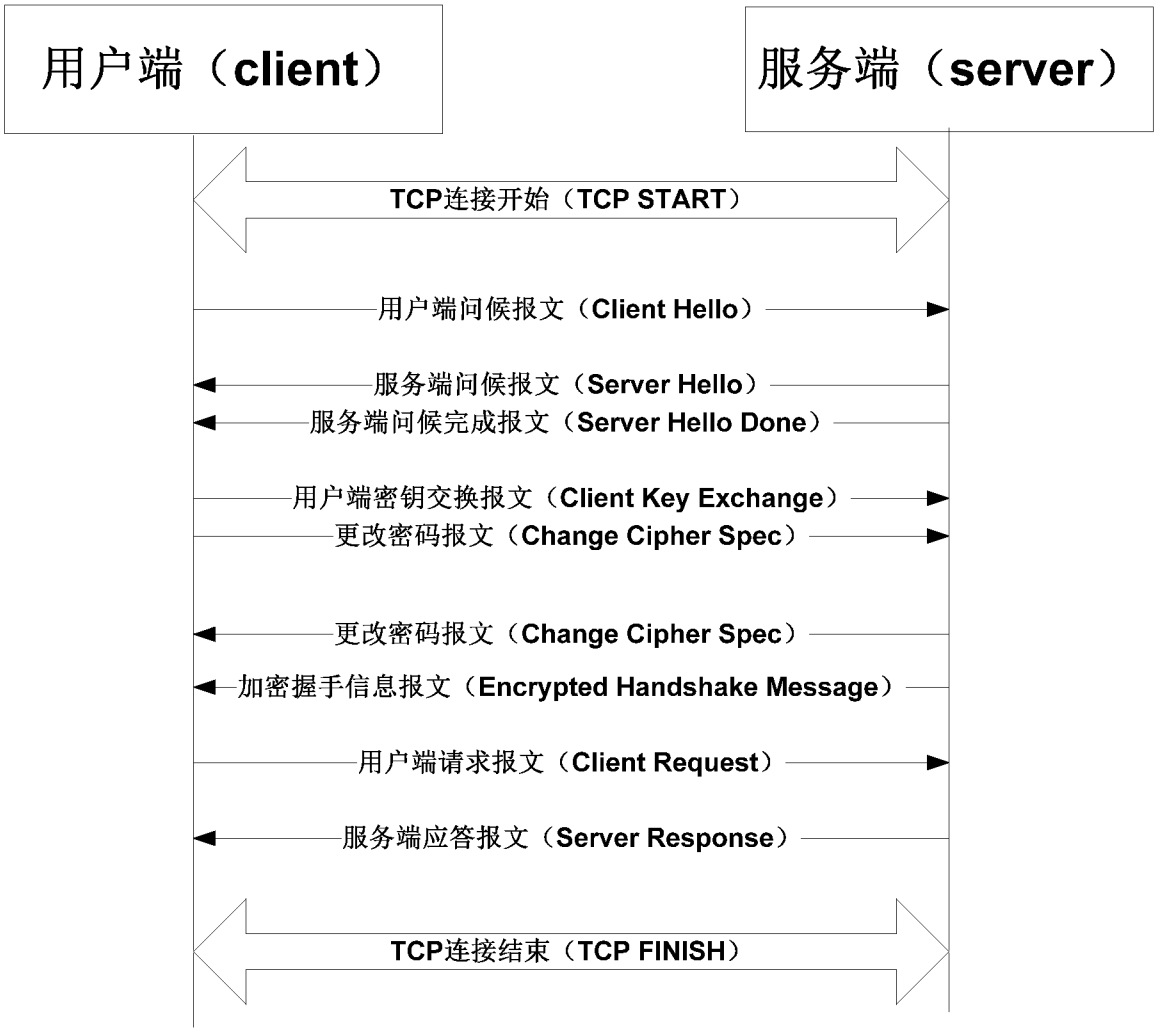
用户端（client）
服务端（server）
TCP连接开始（TCP START）
用户端问候报文（Client Hello）
服务端问候报文（Server Hello）
服务端问候完成报文（Server Hello Done）
用户端密钥交换报文（Client Key Exchange）
更改密码报文（Change Cipher Spec）
更改密码报文（Change Cipher Spec）
加密握手信息报文（Encrypted Handshake Message）
用户端请求报文（Client Request）
服务端应答报文（Server Response）
TCP连接结束（TCP FINISH）

图 4b

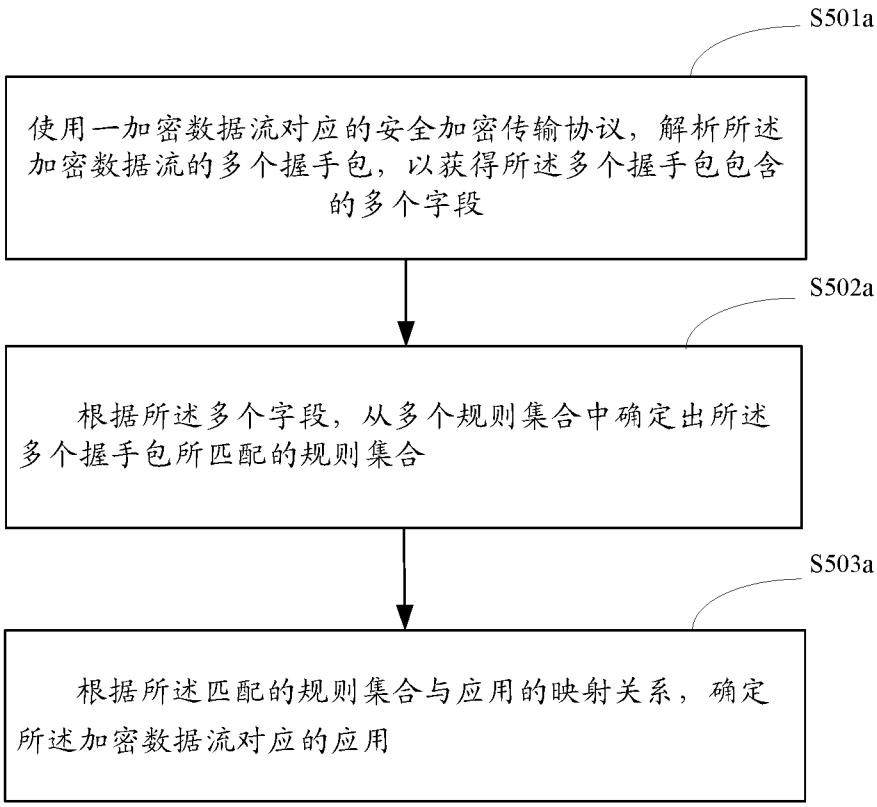
S501a
使用一加密数据流对应的安全加密传输协议，解析所述加密数据流的多个握手包，以获得所述多个握手包包含的多个字段
S502a
根据所述多个字段，从多个规则集合中确定出所述多个握手包所匹配的规则集合
S503a
根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用

图 5a

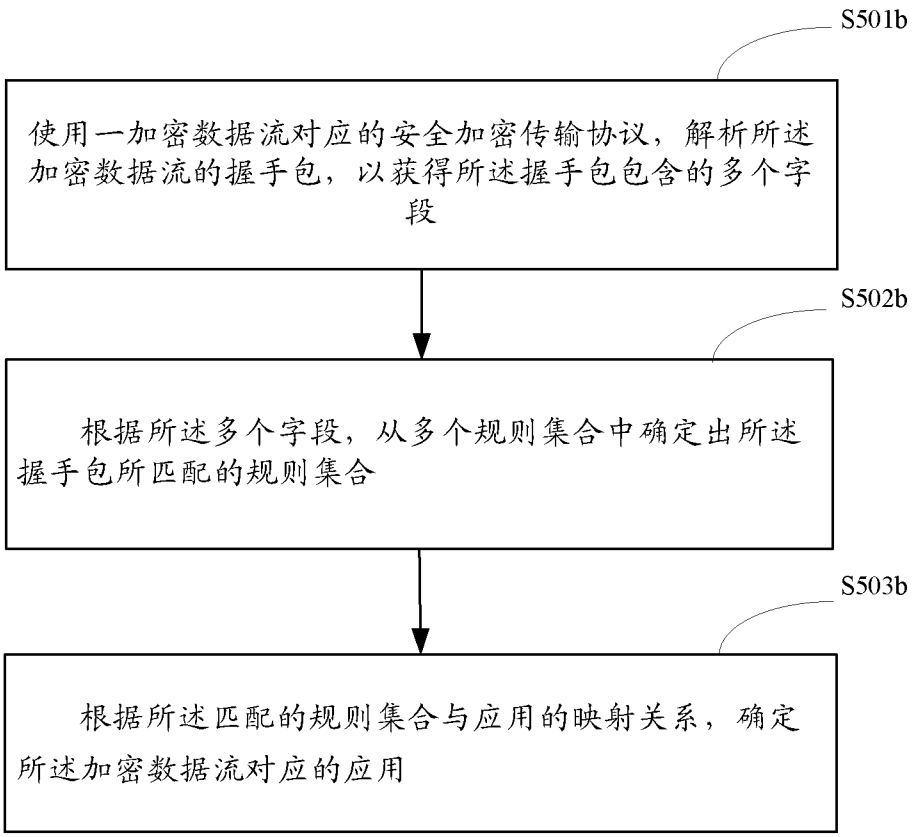
S501b
使用一加密数据流对应的安全加密传输协议，解析所述加密数据流的握手包，以获得所述握手包包含的多个字段
S502b
根据所述多个字段，从多个规则集合中确定出所述握手包所匹配的规则集合
S503b
根据所述匹配的规则集合与应用的映射关系，确定所述加密数据流对应的应用

图 5b

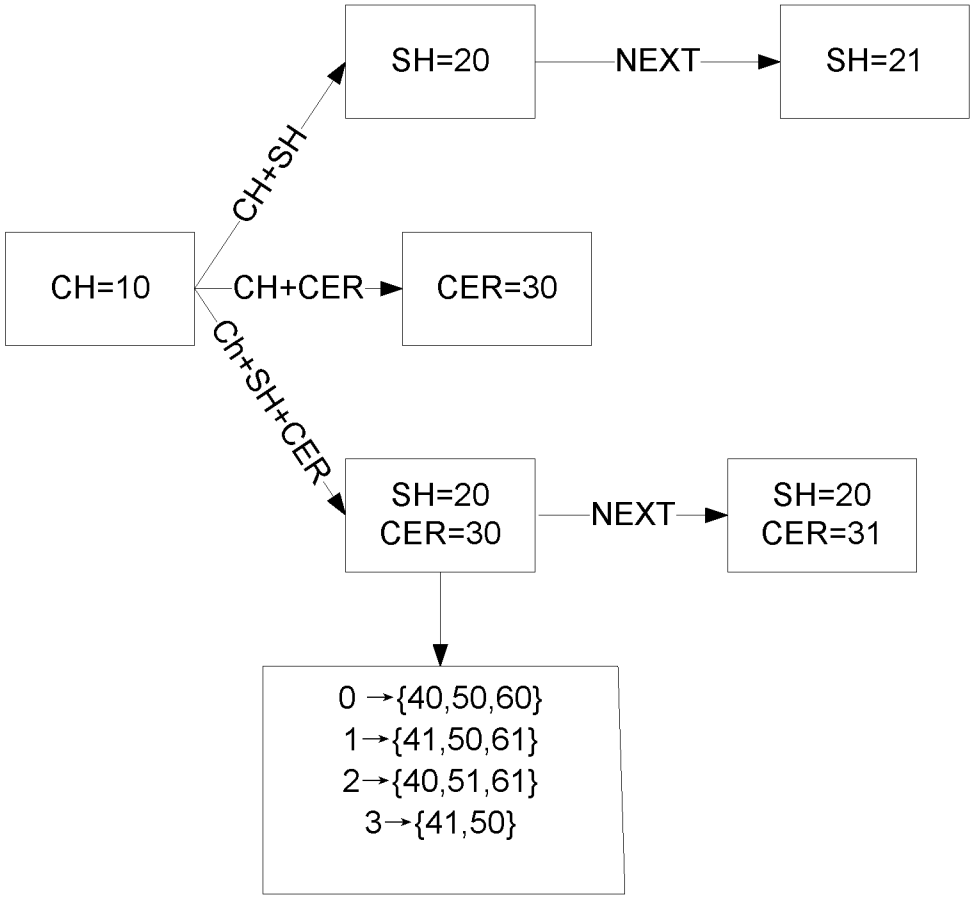
SH=20
NEXT
SH=21
CH+SH
CH=10
CH+CER
CER=30
Ch+SH+CER
SH=20
CER=30
NEXT
SH=20
CER=31
0 →{40,50,60}
1→{41,50,61}
2→{40,51,61}
3→{41,50}

图 6

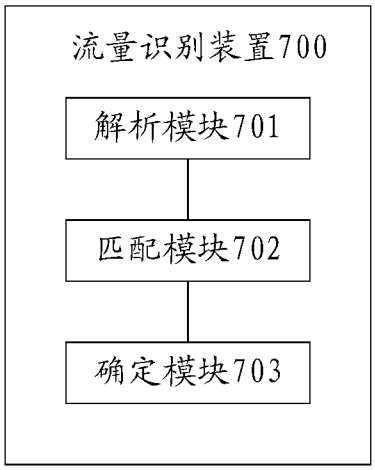
流量识别装置700
解析模块701
匹配模块702
确定模块703

图 7

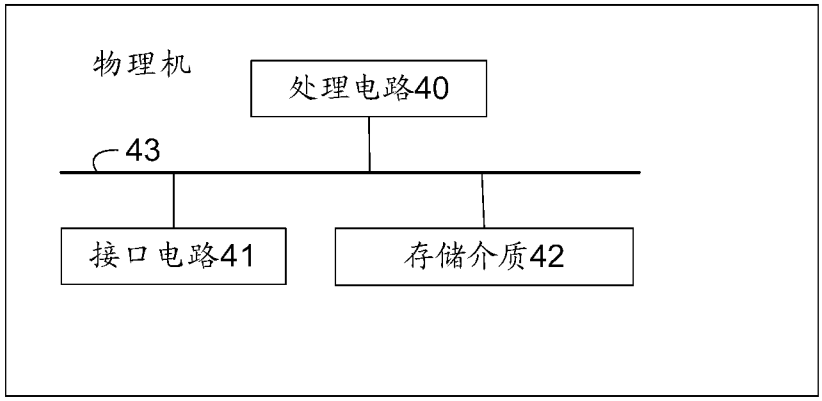

图 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/116207

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)

H04L

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)

CNABS; CNTXT; CNKI; VEN; WOTXT; EPTXT; USTXT: 辨别, 识别, 检测, 分辨, 握手, 安全套接字, 应用, 业务类型, identify, verify, detect, handshake, TLS, SSL, certificatetatus, certificateverify, certificate, clienthello, serverhello, finished, application, service type

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 103618726 A (BEIJING ZHONGCHUANG TELECOM TEST CO., LTD.) 05 March 2014 (2014-03-05) description, paragraphs 11-17 and 37-123, and figures 4-17	1-32
Y	扶佩佩 (FU, Peipei). "针对SSL协议的网络应用精细化分类技术研究 (Research on Fine-Grained Classification Technology of Network Application Based on SSL Protocol)" *中国优秀硕士学位论文全文数据库 (China Master's Theses Full-Text Database),* No. vol. 11, 15 November 2013 (2013-11-15), entire document	1-32
A	CN 105871832 A (BEIJING INSTITUTE OF TECHNOLOGY) 17 August 2016 (2016-08-17) entire document	1-32
A	WO 2017079980 A1 (HUAWEI TECHNOLOGIES CO., LTD.) 18 May 2017 (2017-05-18) entire document	1-32
A	GB 2450897 B (TIDEWAY SYSTEMS LTD.) 23 September 2009 (2009-09-23) entire document	1-32
A	US 9118484 B1 (SYMANTEC CORP.) 25 August 2015 (2015-08-25) entire document	1-32

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

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"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)

"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

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
04 January 2019	**31 January 2019**

Name and mailing address of the ISA/CN	Authorized officer
State Intellectual Property Office of the P. R. China **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing** **100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/116207

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	103618726	A	05 March 2014	None			
CN	105871832	A	17 August 2016	CN	105871832	B	02 November 2018
WO	2017079980	A1	18 May 2017	None			
GB	2450897	B	23 September 2009	GB	2450897	A	14 January 2009
				GB	0713414	D0	22 August 2007
US	9118484	B1	25 August 2015	US	8707027	B1	22 April 2014

国际检索报告

国际申请号
PCT/CN2018/116207

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)

H04L

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））

CNABS;CNTXT;CNKI;VEN;WOTXT;EPTXT;USTXT:辨别，识别，检测，分辨，握手，安全套接字，应用，业务类型，identify，verify，detect，handshake，TLS,SSL,certificatetatus，certificateverify，certificate，clienthello，serverhello，finished，application，service type

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
Y	CN 103618726 A (北京中创信测科技股份有限公司) 2014年 3月 5日 (2014-03-05) 说明书第11-17段，第37段至第123段，图4至图17	1-32
Y	扶佩佩. "针对SSL协议的网络应用精细化分类技术研究" 中国优秀硕士学位论文全文数据库，第11期，2013年 11月 15日 (2013-11-15)， 全文	1-32
A	CN 105871832 A (北京理工大学) 2016年 8月 17日 (2016-08-17) 全文	1-32
A	WO 2017079980 A1 (华为技术有限公司) 2017年 5月 18日 (2017-05-18) 全文	1-32
A	GB 2450897 B (TIDEWAY SYSTEMS LTD) 2009年 9月 23日 (2009-09-23) 全文	1-32
A	US 9118484 B1 (SYMANTEC CORP) 2015年 8月 25日 (2015-08-25) 全文	1-32

☐ 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	国际检索报告邮寄日期
2019年 1月 4日	2019年 1月 31日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 薛乐梅 电话号码 86-(20)-28950448

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/116207

检索报告引用的专利文件			公布日（年/月/日）	同族专利			公布日（年/月/日）
CN	103618726	A	2014年 3月 5日	无			
CN	105871832	A	2016年 8月 17日	CN	105871832	B	2018年 11月 2日
WO	2017079980	A1	2017年 5月 18日	无			
GB	2450897	B	2009年 9月 23日	GB	2450897	A	2009年 1月 14日
				GB	0713414	D0	2007年 8月 22日
US	9118484	B1	2015年 8月 25日	US	8707027	B1	2014年 4月 22日