

(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
2018 年 6 月 7 日 (07.06.2018)



(10) 国际公布号
WO 2018/099219 A1

(51) 国际专利分类号:
H04L 29/06 (2006.01)

(21) 国际申请号: PCT/CN2017/107865

(22) 国际申请日: 2017 年 10 月 26 日 (26.10.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201611076447.2 2016 年 11 月 29 日 (29.11.2016) CN

(71) 申请人: 中国银联股份有限公司 (CHINA UNIONPAY CO., LTD.) [CN/CN]; 中国上海市浦东新区含笑路 36 号银联大厦, Shanghai 200135 (CN)。

(72) 发明人: 杨阳(YANG, Yang); 中国上海市浦东新区含笑路 36 号银联大厦, Shanghai 200135 (CN)。胡景秀(HU, Jingxiu); 中国上海市浦东新区含笑路 36 号银联大厦, Shanghai 200135 (CN)。陈舟(CHEN, Zhou); 中国上海市浦东新区含笑路 36 号银联大厦, Shanghai 200135 (CN)。尹亚伟(YIN, Yawei); 中国上海市浦东新区含笑路 36 号银联大厦, Shanghai 200135 (CN)。

(74) 代理人: 北京同达信恒知识产权代理有限公司 (TDIP & PARTNERS); 中国北京市海淀区宝盛南路 1 号院 20 号楼 8 层 101-01, Beijing 100192 (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,

(54) Title: METHOD AND DEVICE FOR DETECTING PHISHING WEBSITE

(54) 发明名称: 一种钓鱼网站检测方法和装置

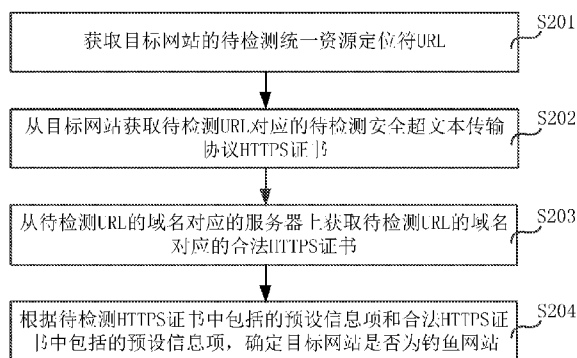


图 2

- S201 Obtain a uniform resource locator (URL) to be determined of a target website
- S202 Obtain, from the target website, a hypertext transfer protocol secure (HTTPS) certificate to be determined and corresponding to the URL to be determined
- S203 Obtain, from a server corresponding to a domain name of the URL to be determined, a valid HTTPS certificate corresponding to the domain name of the URL to be determined
- S204 Determine, according to a preconfigured information item in the HTTPS certificate to be determined and a preconfigured information item in the valid HTTPS certificate, whether the target website is a phishing website

(57) Abstract: The invention relates to the field of communications, and specifically, to a method and device for detecting a phishing website, effectively determining, when a URL domain name is a valid domain name, whether a target website being accessed by a user is a phishing website. The method comprises: obtaining a uniform resource locator (URL) to be determined of a target website; determining that the URL to be determined is a valid domain name; obtaining, from the target website, a hypertext transfer protocol secure (HTTPS) certificate to be determined and corresponding to the URL to be determined; obtaining, from a server corresponding to the domain

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, 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))。

name of the URL to be determined, a valid HTTPS certificate corresponding to the domain name of the URL to be determined; and determining, according to a preconfigured information item in the HTTPS certificate to be determined and a preconfigured information item in the valid HTTPS certificate, whether the target website is a phishing website. As a result, the embodiment is utilized to effectively determine, when a URL domain name is a valid domain name, whether a target website being accessed by a user is a phishing website.

(57) 摘要: 本发明实施例涉及通信技术领域, 尤其涉及一种钓鱼网站检测方法和装置, 用于有效检测待检测URL的域名为合法域名时用户正在访问的目标网站是否是钓鱼网站。获取目标网站的待检测统一资源定位符URL; 待检测URL对应的域名为合法域名; 从目标网站获取待检测URL对应的待检测HTTPS证书; 从待检测URL的域名对应的服务器上获取待检测URL的域名对应的合法HTTPS证书; 根据待检测HTTPS证书中包括的预设信息项和合法HTTPS证书中包括的预设信息项, 确定目标网站是否为钓鱼网站; 进而有效检测待检测URL的域名为合法域名时用户正在访问的目标网站是否是钓鱼网站。

一种钓鱼网站检测方法和装置

本发明要求在 2016 年 11 月 29 日提交中华人民共和国知识产权局、申请号为 201611076447.2、发明名称为“一种钓鱼网站检测方法和装置”的中国专利申请的优先权，其全部内容通过引用结合在本发明中。

技术领域

本发明涉及通信技术领域，尤其涉及一种钓鱼网站检测方法和装置。

背景技术

钓鱼网站通常伪装成银行及电子商务网站，窃取用户提交的银行帐号、密码等私密信息。恶意的攻击者通过伪基站、钓鱼邮件等方式，给受害者提供一个看似正常的银行或金融类网站域名，诱骗用户打开后在钓鱼网站内输入用户名、密码、卡号、身份证号等敏感信息。

现有的检测钓鱼网站的工具主要使用黑名单、白名单、启发式分析等多种技术识别钓鱼网站的网址。对于待检测统一资源定位符 (Uniform Resource Locator, 简称 URL) 的域名错误的情况下，主要是采用黑名单、白名单的钓鱼网站检测技术进行检测：基于黑名单的钓鱼网站检测技术是根据已知的错误的 URL 构建的数据库来判断目标网站是否为钓鱼网站，但由于钓鱼网站往往存在时间较短、黑名单的及时更新十分困难，使得该技术不能有效防止钓鱼；基于白名单的钓鱼网站检测技术是根据可信网站地址构建的数据库来判断目标网站是否为合法网站，由于可信网站地址数量庞大，若用户访问的合法网址未被加入数据库，则可能导致将合法网站误判为钓鱼网站。

对于待检测 URL 的域名正确的情况下，主要是采用启发式分析的钓鱼网站检测技术进行检测：根据目标网站和合法网站进行网页页面布局相似度进行比对，确定目标网站是否为钓鱼网站，但这种检测方法容易被一些欺骗手段绕过，导致对钓鱼网站存在一定程度的漏检。

现有技术中只根据待检测 URL 或根据目标网站和合法网站的页面布局相似度检测目标网站是否为钓鱼网站，对于域名系统（Domain Name System，简称 DNS）劫持、流量劫持、安全超文本传输协议（Secure Hypertext Transfer Protocol，简称 HTTPS）中间人等复杂的钓鱼攻击手段，获取的待检测 URL 对应的域名为合法域名，且目标网站和合法网站的相似度极高，现有技术中的方法无法有效确定目标网站是否为钓鱼网站。因此，亟需一种钓鱼网站检测方法，有效检测待检测 URL 的域名为合法域名时用户正在访问的目标网站是否是钓鱼网站。

发明内容

本发明实施例提供一种钓鱼网站检测方法和装置，用以有效检测待检测 URL 的域名为合法域名时用户正在访问的目标网站是否是钓鱼网站。

第一方面，本发明实施例中提供一种钓鱼网站检测方法，包括：获取目标网站的待检测统一资源定位符 URL；所述待检测 URL 对应的域名为合法域名；从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书；从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书；根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站。

可选地，所述根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站，包括：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项全部匹配的情况下，确定所述目标网站为正常网站。

可选地，所述根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站，包括：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：在确定所述待检

测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配、根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻有效、且根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书未被注销，则确定所述目标网站为正常网站。

可选地，所述根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站，包括：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：确定所述待检测 HTTPS 证书和所述待检测 HTTPS 证书满足第一预设条件中的至少一项，则确定所述目标网站为钓鱼网站；其中，所述第一预设条件包括：在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配；根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻失效；根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书被注销。

可选地，所述获取目标网站的待检测 URL 之后，从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前，还包括：确定所述待检测 URL 使用 HTTPS。

可选地，所述获取目标网站的待检测 URL 之后，从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前，还包括：在确定所述待检测 URL 未使用 HTTPS 的情况下，上报告警信息，所述告警信息用于指示所述目标网站存在危险。

可选地，所述获取目标网站的待检测 URL 之后，从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前，还包括：确定所述待检测 URL 满足第二预设条件中的任一项或任多项，其中，所述第二预设条件包括：确定所述待检测 URL 的域名与至少一个预设的 URL 域名匹配；确定所获取的所述待检测 URL 对应的网页源代码中的属性类关键词中存在至少一个与预设属

性类关键词匹配；确定所获取的所述待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配。

第二方面，本发明实施例提供的一种钓鱼网站检测装置，包括：获取单元，用于获取目标网站的待检测统一资源定位符 URL；所述待检测 URL 对应的域名为合法域名；从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书；从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书；处理单元，用于根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站。

可选地，所述处理单元，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项全部匹配的情况下，确定所述目标网站为正常网站。

可选地，所述处理单元，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配、根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻有效、且根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书未被注销，则确定所述目标网站为正常网站。

可选地，所述处理单元，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：确定所述待检测 HTTPS 证书和所述待检测 HTTPS 证书满足第一预设条件中的至少一项，则确定所述目标网站为钓鱼网站；其中，所述第一预设条件包括：在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配；根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻失效；根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测

HTTPS 证书被注销。

可选地，所述处理单元，还用于：确定所述待检测 URL 使用 HTTPS。

可选地，所述处理单元，还用于：在确定所述待检测 URL 未使用 HTTPS 的情况下，上报告警信息，所述告警信息用于指示所述目标网站存在危险。

可选地，所述处理单元，还用于：确定所述待检测 URL 满足第二预设条件中的任一项或任多项，其中，所述第二预设条件包括：确定所述待检测 URL 的域名与至少一个预设的 URL 域名匹配；确定所获取的所述待检测 URL 对应的网页源代码中的属性类关键词中存在至少一个与预设属性类关键词匹配；确定所获取的所述待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配。

第三方面，本发明实施例提供一种钓鱼网站检测装置，包括处理器和存储器；

所述存储器，用于存储可执行程序；

所述处理器用于读取所述存储器中的可执行程序，执行：

获取目标网站的待检测统一资源定位符 URL；所述待检测 URL 对应的域名为合法域名；从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书；从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书；根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站。

第四方面，本发明实施例提供一种非暂态计算机可读存储介质，所述非暂态计算机可读存储介质存储计算机指令，所述计算机指令用于使所述计算机执行第一方面或第一方面的任意可能的实现方式中的方法。

第五方面，本发明实施例提供一种计算机程序产品，所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，使所述计算机执行第一方面或第一方面的任意可能的实现方式中的方法。

本发明实施例中从目标网站获取待检测 URL 对应的待检测 HTTPS 证书，待检测 URL 的域名正确，因此才能获取待检测 URL 对应的合法 HTTPS 证书；一个 URL 的域名对应一个合法 HTTPS 证书，从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书，根据待检测 HTTPS 证书和合法 HTTPS 证书，进而可有效检测用户正在访问的目标网站是否为钓鱼网站。

附图说明

为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简要介绍。

图 1 为本发明实施例提供的一种钓鱼网站检测的系统架构示意图；

图 2 为本发明实施例提供的一种钓鱼网站检测方法流程示意图；

图 3 为本发明实施例提供的在另一种钓鱼网站检测方法流程示意图；

图 4 为本发明实施例提供的一种钓鱼网站检测装置的结构示意图；

图 5 为本发明实施例提供的另一种钓鱼网站检测装置的结构示意图。

具体实施方式

为了使本发明的目的、技术方案及有益效果更加清楚明白，以下结合附图及实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。

图 1 示例性示出了本发明实施例适用的钓鱼网站检测的系统架构示意图，如图 1 所示，本发明实施例适用的钓鱼网站检测系统架构 100 包括服务器 101、服务器 102、终端 103、钓鱼网站检测装置 104；服务器 101、服务器 102 和终端 103 可以通过无线连接或有线连接或其它方式连接；服务器 101、服务器 102 和钓鱼网站检测装置 104 也可通过无线连接或有线连接或其它方式连接；其中，服务器 101 为目标网站对应的服务器，服务器 102 为待检测 URL 的域名对应的服务器，钓鱼网站检测装置 104 安装于终端 103 中。

终端 103 向服务器 101 发送访问目标网站的 URL，服务器 101 接收到该 URL 后，向终端 103 发送该 URL 对应的网页内容，例如网页源代码等。钓鱼网站检测装置 104 可从服务器 102 上获得待检测 HTTPS 证书；钓鱼网站检测装置 104 可从服务器 102 上获得 URL 的域名对应的合法 HTTPS 证书。例如，目标网站为百度，则百度对应的 URL 为 <https://www.baidu.com/>，URL 的域名为 www.baidu.com。

可选地，终端 103 可以为手机、平板电脑、计算机等；可选地，钓鱼网站检测装置 104 可安装于终端 103 中，用于检测终端访问的目标网站是否为钓鱼网站；可选地，钓鱼网站检测装置 104 可以为银联程序插件，也可以为安全卫士应用。

图 2 示例性示出了本发明实施例提供的一种钓鱼网站检测方法流程示意图。

基于图 1 所示的系统架构，如图 2 所示，本发明实施例提供的一种钓鱼网站检测方法，包括以下步骤：

步骤 S201：钓鱼网站检测装置获取目标网站的待检测统一资源定位符 URL；所述待检测 URL 对应的域名为合法域名；

步骤 S202：钓鱼网站检测装置从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书；

步骤 S203：钓鱼网站检测装置从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书；

步骤 S204：钓鱼网站检测装置根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站。

基于上述实施例，步骤 S201 中，合法域名为用户预期访问的目标网站的正确域名；例如，用户当前访问预期访问的目标网站为中国工商银行，用户输入的待检测 URL 为 <http://www.icbc.com.cn/icbc/>，待检测 URL 的域名为 www.icbc.com.cn，这个域名为中国工商银行的正确域名，即为合法域名。在

正常情况下，当用户访问 <http://www.icbc.com.cn/icbc/> 时，进入的目标网站为用户预期访问的网站：中国工商银行；在 DNS 劫持、流量劫持、HTTPS 中间人等复杂的钓鱼攻击手段存在的情况下，用户当前访问 <http://www.icbc.com.cn/icbc/> 时，其对应的域名仍为合法域名 www.icbc.com.cn，但是用户进入的目标网站为钓鱼网站，现有技术中无法有效检测出用户当前访问的待检测 URL 的域名为合法域名时对应的目标网站是否为钓鱼网站，本发明实施例提供的方法可以有效检测用户当前访问的待检测 URL 的域名为合法域名时对应的目标网站是否为钓鱼网站。

本发明实施例中，钓鱼网站检测装置以银联程序插件为例，银联程序插件使用安全传输层协议(Transport Layer Security Protocol，简称 TLS)安全连接银联后台服务器；银联后台服务器根据待检测 URL 的域名，若银联后台服务器的缓存数据库中存在待检测 URL 的域名对应的合法 HTTPS 证书，则直接将合法 HTTPS 证书发送至银联程序插件，进行验证待检测 HTTPS 证书的合法性；若银联后台服务器的自身缓存数据库中不存在待检测 URL 的域名对应的合法 HTTPS 证书，银联后台服务器主动访问待检测 URL，从待检测 URL 的域名对应的服务器上获取待检测 URL 对应的合法 HTTPS 证书，将合法 HTTPS 证书存放于自身缓存数据库，并将待检测 URL 对应的合法 HTTPS 证书发送至银联程序插件。为保证自身缓存数据库的时效性，银联后台服务器定期验证自身缓存数据库中的 HTTPS 证书是否被注销或不在有效期内，若存在 HTTPS 证书被注销或不在有效期内的情况，将该 HTTPS 证书标记为不合法 HTTPS 证书。

本发明实施例中，待检测 HTTPS 证书和合法 HTTPS 证书中的任一个 HTTPS 证书包括证书签发者标识、证书有效期、证书序列号、证书签名等多个关键因素；可选地，预设信息项可包括多个关键因素中的任一个，例如将证书签发者标识作为预设信息项；预设信息项也可包括多个关键因素中的任 N 个，N 为大于 1 的整数，例如将证书签发者标识、证书有效期作为预设信息项，又例如将证书签发者标识、证书序列号作为预设信息项，再例如将证书

签发者标识、证书序列号、证书签名作为预设信息项。

由于本发明实施例中从目标网站获取待检测 URL 对应的待检测 HTTPS 证书,待检测 URL 的域名正确,因此才能获得待检测 URL 对应的合法 HTTPS 证书;一个 URL 的域名对应一个合法 HTTPS 证书,从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书,根据待检测 HTTPS 证书和合法 HTTPS 证书,进而可有效检测用户正在访问的目标网站是否为钓鱼网站。

可选地,所述根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项,确定所述目标网站是否为钓鱼网站,包括:在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项全部匹配的情况下,确定所述目标网站为正常网站。

可选地,预设信息项包括一个关键因素时,例如预设信息项为证书签发者标识,在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配的情况下,确定所述目标网站为正常网站;可选地,预设信息项包括 N 个关键因素时,例如,N 等于 2,以预设信息项包括的两个关键因素为证书签发者标识、证书序列号为例,在确定所述待检测 HTTPS 证书中包括的证书签发者标识、证书序列号和所述合法 HTTPS 证书中包括的证书签发者标识、证书序列号全部匹配的情况下,确定所述目标网站为正常网站;例如,N 等于 4,以预设信息项包括的四个关键因素为证书签发者标识、证书有效期、证书序列号、证书签名为例,在确定所述待检测 HTTPS 证书中包括的证书签发者标识、证书有效期、证书序列号、证书签名和所述合法 HTTPS 证书中包括的证书签发者标识、证书有效期、证书序列号、证书签名全部匹配的情况下,确定所述目标网站为正常网站。

可选地,在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中至少一项不匹配的情况下,确定所述目标网站为钓鱼网站。例如,N 等于 2,以预设信息项包括的四个关键因素为证书签发者标识、证书序列号为例,在确定所述待检测 HTTPS 证书中包括的证书

签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配的情况下，确定所述目标网站为钓鱼网站；或者，在确定所述待检测 HTTPS 证书中包括的证书序列号和所述合法 HTTPS 证书中包括的证书序列号不匹配的情况下，确定所述目标网站为钓鱼网站；或者，在确定所述待检测 HTTPS 证书中包括的证书签发者标识、证书序列号和所述合法 HTTPS 证书中包括的证书签发者标识、证书序列号均不匹配的情况下，确定所述目标网站为钓鱼网站。如此，钓鱼网站检测装置可根据待检测 HTTPS 证书中和合法 HTTPS 证书中包括的预设信息项是否完全匹配，确定待检测 HTTPS 证书是否合法证书，进而有效确定目标网站是否为正常网站；而且，在确定 HTTPS 证书是合法证书情况下，确定目标网站为正常网站，终端可以正常浏览待检测 URL 对应的网页；在确定 HTTPS 证书不是合法证书情况下，确定目标网站为钓鱼网站，上报告警信息，并提醒用户不要继续访问待检测 URL 对应的网页，以避免造成个人信息的泄露。本发明实施例中提供的方法不需要进行待检测 URL 的黑白名单查询，因此检测结果不受黑白名单数据库更新频率和覆盖率的影响。

可选地，所述根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站，包括：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：在同时满足以下三个条件的情况下，确定所述目标网站为正常网站：条件一，确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配；条件二，根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻有效；条件三，根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书未被注销。如此，在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下，钓鱼网站检测装置可根据以上三个条件，确定目标网站是否为正常网站。

在实际应用场景下，由于内容分发网络(Content Delivery Network，简称

CDN)或负载均衡的存在,不同的终端访问目标网站时看到的待检测 HTTPS 证书可能不一致;例如华东地区的终端连接华东地区的目标网站服务器和华南地区的终端连接华南地区的服务器,待检测 HTTPS 证书包括的证书签发者标识一样,证书序列号有可能不一样。本发明实施例提供的方案充分考虑了可能由于 CDN 或负载均衡的存在,导致不同终端访问目标网站时可能得到的 HTTPS 证书不一致的情况,避免了将待检测 HTTPS 证书和合法 HTTPS 证书的预设信息项中存在至少一项不匹配的正常网站确定为钓鱼网站的情况。

可选地,所述根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项,确定所述目标网站是否为钓鱼网站,包括:在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下:确定所述待检测 HTTPS 证书和所述待检测 HTTPS 证书满足第一预设条件中的至少一项,则确定所述目标网站为钓鱼网站;其中,所述第一预设条件包括:在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配;根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻失效;根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书被注销。如此,本发明实施例提供的方案,在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下,充分考虑了可能由于 CDN 或负载均衡的存在,导致不同终端访问目标网站时可能得到的 HTTPS 证书不一致的情况,更准确的将待检测 HTTPS 证书满足上述第一预设条件中的至少一项的目标网站确定为钓鱼网站。本发明实施例中提供的方法不需要进行待检测 URL 的黑白名单查询,因此检测结果不受黑白名单数据库更新频率和覆盖率的影响;同时,也可避免大规模搜索所有网站的 URL,并验证所有网站的 URL 是否为钓鱼网站的资源浪费。

可选地,所述获取目标网站的待检测 URL 之后,从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前,还包括:确定所述待检测 URL 使

用安全超文本传输协议 HTTPS。例如,待检测 URL 为 <https://zhidao.baidu.com/>, 该待检测 URL 的协议为 HTTPS; 如此, 钓鱼检测装置可根据待检测 URL 使用的 HTTPS, 继续获取待检测 HTTPS 证书, 继续判断待检测 HTTPS 证书的合法性, 进而判断目标网站是否为钓鱼网站。

可选地, 所述获取目标网站的待检测 URL 之后, 从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前, 还包括: 在确定所述待检测 URL 未使用 HTTPS 的情况下, 上报告警信息, 所述告警信息用于指示所述目标网站存在危险。例如, 待检测 URL 为 <http://abc.com/>, 该待检测 URL 的协议为 http, 未使用 HTTPS; 如此, 钓鱼检测装置不需要继续获取待检测 HTTPS 证书, 可根据待检测 URL 未使用 HTTPS, 直接上报目标网站存在危险的告警信息, 避免用户在不知情的情况下进入存在危险的目标网站。

可选地, 所述获取目标网站的待检测 URL 之后, 从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前, 还包括: 确定所述待检测 URL 满足第二预设条件中的任一项或任多项, 其中, 所述第二预设条件包括: 条件一, 确定所述待检测 URL 的域名与至少一个预设的 URL 域名匹配; 条件二, 确定所获取的所述待检测 URL 对应的网页源代码中的属性类关键词中存在至少一个与预设属性类关键词匹配; 条件三, 确定所获取的所述待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配。可选地, 所述获取目标网站的待检测 URL 之后, 在确定所述待检测 URL 满足第二预设条件中的任一项的情况下, 确定目标网站为正常网站, 可正常在终端上访问目标网站。

本发明实施例中, 预设的 URL 域名可以根据实际需要进行设定, 可选地, 可以将预设的 URL 域名设为涉及到银行账户信息以及个人信息的网站的域名; 可选地, 预设的 URL 域名可以为银行网站 URL 域名, 例如, 中国工商银行的 URL 域名为 www.icbc.com.cn; 例如, 中国农业银行的 URL 域名为 www.abchina.com; 预设的 URL 域名也可以为社交网站 URL 域名, 例如, 新浪微博的 URL 域名为 weibo.com。预设属性类关键词可根据实际需求设定,

可选地，终端访问的目标网站为银行网站，可将预设属性类关键词设置为包括银行、网银、支付、金融等；预设输入类关键词可根据实际需求设定，可选地，可将预设输入类关键词设置为登录账户、卡号、密码、身份证号等。

可选地，待检测 URL 对应的网页源代码中属性类关键词可从待检测 URL 对应的网页源代码中的 <TITLE> 标记、<META name="Keywords"/> "Description"/> "Copyright"> 标记中的 CONTENT 属性、<body> 标记中的提取；可选地，待检测 URL 对应的网页源代码中输入类关键词可从待检测 URL 对应的网页源代码中的 <input> 标记等标记中提取。同时考虑到部分钓鱼网站会使用截图来布局，上述属性类关键词和输入类关键词在提取时会同步提取通过光学字符识别 (Optical Character Recognition, 简称 OCR) 识别到的图片中的文字。

例如，待检测 URL 的域名为 www.icbc.com.cn，则与预设的中国工商银行的域名匹配，则继续从目标网站获取待检测 HTTPS 证书，以进行验证待检测 HTTPS 证书是否合法；例如，待检测 URL 的域名为 www.abc.com.cn，与预设的中国工商银行的域名不匹配，则可以继续判断待检测 URL 是否满足第二预设条件中的条件二和条件三，即待检测 URL 对应的网页源代码中的属性类关键词是否与至少一个预设属性类关键词匹配，或输入类关键词中存在至少一个与预设输入类关键词匹配，或属性类关键词是否与至少一个预设属性类关键词匹配和输入类关键词中存在至少一个与预设输入类关键词匹配。如此，可根据待检测 URL 是否满足第二预设条件，确定目标网站是否需要进行检测，这种判断目标网站是否需要进行检测的方法简单易行、效率高、针对性强，可有效的防御敏感信息泄露，同时避免影响在终端上访问正常网站时的用户体验。

为了更清楚的介绍上述方法流程，本发明实施例提供以下示例。

图 3 示例性示出了本发明实施例提供的另一种钓鱼网站检测方法流程示意图，基于图 1 所示的系统架构，如图 3 所示，该方法包括以下步骤：

步骤 S301：钓鱼网站检测装置获取目标网站的待检测统一资源定位符

URL;

步骤 S302: 待检测 URL 对应的域名是否为合法域名; 若是, 则执行步骤 S303; 若否, 则执行步骤 S312;

步骤 S303: 待检测 URL 的域名是否与至少一个预设的 URL 域名匹配; 若是, 则执行步骤 S306; 若否, 则执行步骤 S304;

步骤 S304: 所获取的待检测 URL 对应的网页源代码中的属性类关键词中是否存在至少一个与预设属性类关键词匹配; 若是, 则执行步骤 S305; 若否, 则执行步骤 S313;

步骤 S305: 确定所获取的待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配; 若是, 则执行步骤 S306; 若否, 则执行步骤 S313;

步骤 S306: 确定待检测 URL 是否使用安全超文本传输协议 HTTPS; 若是, 则执行步骤 S307; 若否, 则执行步骤 S314;

步骤 S307: 从目标网站获取待检测 URL 对应的待检测 HTTPS 证书, 并从待检测 URL 的域名对应的服务器上获取待检测 URL 的域名对应的合法 HTTPS 证书;

步骤 S308: 确定待检测 HTTPS 证书中包括的预设信息项和合法 HTTPS 证书中包括的预设信息项是否全部匹配; 若是, 则执行步骤 S313; 若否, 则执行步骤 S309;

步骤 S309: 确定待检测 HTTPS 证书中包括的证书签发者标识和合法 HTTPS 证书中包括的证书签发者标识是否匹配; 若是, 则执行步骤 S310; 若否, 则执行步骤 S312;

步骤 S310: 根据待检测 HTTPS 证书中的证书有效期确定待检测 HTTPS 证书在当前时刻是否有效; 若是, 则执行步骤 S311; 若否, 则执行步骤 S312;

步骤 S311: 根据待检测 HTTPS 证书中的证书是否被注销的信息确定待检测 HTTPS 证书是否未被注销; 若是, 则执行步骤 S313; 若否, 则执行步骤 S312;

步骤 S312: 确定目标网站为钓鱼网站;

步骤 S313: 确定目标网站为正常网站;

步骤 S314: 上报告警信息, 所述告警信息用于指示所述目标网站存在危险。

从上述内容可以看出: 本发明实施例中提供了一种钓鱼网站检测方法的方法, 根据待检测 URL 是否满足第二预设条件判断目标网站是否需要进行钓鱼网站检测, 这种判断方法简单易行、效率高、针对性强, 可有效的防御敏感信息泄露, 同时避免影响用户在终端上访问正常网站时的用户体验。由于本发明实施例中从目标网站获取待检测 URL 对应的待检测 HTTPS 证书, 待检测 URL 的域名正确, 因此才能获取待检测 URL 对应的合法 HTTPS 证书; 一个 URL 的域名对应一个合法 HTTPS 证书, 从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书, 根据待检测 HTTPS 证书和合法 HTTPS 证书, 进而可有效检测用户正在访问的目标网站是否为钓鱼网站。本发明实施例中充分考虑了可能由于 CDN 或负载均衡的存在, 导致不同终端访问目标网站时可能得到的 HTTPS 证书不一致的情况, 更准确、有效的检测用户正在访问的目标网站是否为钓鱼网站。进一步, 本发明实施例中提供的方法不需要进行待检测 URL 的黑白名单查询, 因此检测结果不受黑白名单数据库更新频率和覆盖率的影响, 也可避免大规模搜索所有网站的 URL, 并验证所有网站的 URL 是否为钓鱼网站的资源浪费。

图 4 示例性示出了本发明实施例提供的一种钓鱼网站检测装置的结构示意图。

基于相同构思, 本发明实施例提供的一种钓鱼网站检测装置, 用于执行上述方法流程, 如图 4 所示, 该钓鱼网站检测装置 400 包括获取单元 401、处理单元 402; 其中:

获取单元 401, 用于获取目标网站的待检测统一资源定位符 URL; 所述待检测 URL 对应的域名为合法域名; 从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书; 从所述待检测 URL 的域名对应

的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书；

处理单元 402，用于根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站。

可选地，所述处理单元 402，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项全部匹配的情况下，确定所述目标网站为正常网站。

可选地，所述处理单元 402，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配、根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻有效、且根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书未被注销，则确定所述目标网站为正常网站。

可选地，所述处理单元 402，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：确定所述待检测 HTTPS 证书和所述待检测 HTTPS 证书满足第一预设条件中的至少一项，则确定所述目标网站为钓鱼网站；其中，所述第一预设条件包括：在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配；根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻失效；根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书被注销。

可选地，所述处理单元 402，还用于：确定所述待检测 URL 使用 HTTPS。

可选地，所述处理单元 402，还用于：在确定所述待检测 URL 未使用 HTTPS 的情况下，上报告警信息，所述告警信息用于指示所述目标网站存在危险。

可选地，所述处理单元 402，还用于：确定所述待检测 URL 满足第二预设条件中的任一项或任多项，其中，所述第二预设条件包括：确定所述待检测 URL 的域名与至少一个预设的 URL 域名匹配；确定所获取的所述待检测 URL 对应的网页源代码中的属性类关键词中存在至少一个与预设属性类关键词匹配；确定所获取的所述待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配。

从上述内容可以看出：本发明实施例中提供了一种钓鱼网站检测方法的装置，根据待检测 URL 是否满足第二预设条件判断目标网站是否需要进行钓鱼网站检测，这种判断方法简单易行、效率高、针对性强，可有效的防御敏感信息泄露，同时避免影响用户在终端上访问正常网站时的用户体验。由于本发明实施例中从目标网站获取待检测 URL 对应的待检测 HTTPS 证书，待检测 URL 的域名正确，因此才能获取待检测 URL 对应的合法 HTTPS 证书；一个 URL 的域名对应一个合法 HTTPS 证书，从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书，根据待检测 HTTPS 证书和合法 HTTPS 证书，进而可有效检测用户正在访问的目标网站是否为钓鱼网站。本发明实施例中充分考虑了可能由于 CDN 或负载均衡的存在，导致不同终端访问目标网站时可能得到的 HTTPS 证书不一致的情况，更准确、有效的检测用户正在访问的目标网站是否为钓鱼网站。进一步，本发明实施例中提供的方法不需要进行待检测 URL 的黑白名单查询，因此检测结果不受黑白名单数据库更新频率和覆盖率的影响，也可避免大规模搜索所有网站的 URL，并验证所有网站的 URL 是否为钓鱼网站的资源浪费。

基于相同构思，本发明实施例提供另一种钓鱼网站检测装置，可用于执行上述钓鱼网站检测方法流程。图 5 为本发明实施例提供的另一种钓鱼网站检测装置的结构示意图。如图 5 所示，钓鱼网站检测装置 500 包括处理器 501、存储器 502。可选的，还可以包括通信接口 503，所述处理器 501、所述存储器 502 与所述通信接口 503 之间通过总线 504 连接。

总线 504 可以是外设部件互连标准（peripheral component interconnect，

简称 PCI) 总线或扩展工业标准结构 (extended industry standard architecture, 简称 EISA) 总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示, 图 5 中仅用一条粗线表示, 但并不表示仅有一根总线或一种类型的总线。

存储器 502 可以包括易失性存储器 (volatile memory), 例如随机存取存储器 (random-access memory, 简称 RAM); 存储器也可以包括非易失性存储器 (non-volatile memory), 例如快闪存储器 (flash memory), 硬盘 (hard disk drive, 简称 HDD) 或固态硬盘 (solid-state drive, 简称 SSD); 存储器 502 还可以包括上述种类的存储器的组合。

通信接口 503 可以为有线通信接入口, 无线通信接口或其组合, 其中, 有线通信接口例如可以为以太网接口。以太网接口可以是光接口, 电接口或其组合。无线通信接口可以为 WLAN 接口。

处理器 501 可以是中央处理器 (central processing unit, 简称 CPU), 网络处理器 (network processor, 简称 NP) 或者 CPU 和 NP 的组合。

处理器 501 还可以进一步包括硬件芯片。上述硬件芯片可以是专用集成电路 (application-specific integrated circuit, 简称 ASIC), 可编程逻辑器件 (programmable logic device, 简称 PLD) 或其组合。上述 PLD 可以是复杂可编程逻辑器件 (complex programmable logic device, 简称 CPLD), 现场可编程逻辑门阵列 (field-programmable gate array, 简称 FPGA), 通用阵列逻辑 (generic array logic, 简称 GAL) 或其任意组合。

所述处理器 501, 用于读取所述存储器 502 中的程序, 执行下列方法:

获取目标网站的待检测统一资源定位符 URL; 所述待检测 URL 对应的域名为合法域名; 从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书; 从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书; 根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项, 确定所述目标网站是否为钓鱼网站。

所述存储器 502，用于存储一个或多个可执行程序，可以存储所述处理器 501 在执行操作时所使用的数据。

可选地，所述处理器 501，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项全部匹配的情况下，确定所述目标网站为正常网站。

可选地，所述处理器 501，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配、根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻有效、且根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书未被注销，则确定所述目标网站为正常网站。

可选地，所述处理器 501，用于：在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：确定所述待检测 HTTPS 证书和所述待检测 HTTPS 证书满足第一预设条件中的至少一项，则确定所述目标网站为钓鱼网站；其中，所述第一预设条件包括：在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配；根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻失效；根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书被注销。

可选地，所述处理器 501，还用于：确定所述待检测 URL 使用 HTTPS。

可选地，所述处理器 501，还用于：在确定所述待检测 URL 未使用 HTTPS 的情况下，上报告警信息，所述告警信息用于指示所述目标网站存在危险。

可选地，所述处理器 501，还用于：确定所述待检测 URL 满足第二预设条件中的任一项或任多项，其中，所述第二预设条件包括：确定所述待检测 URL 的域名与至少一个预设的 URL 域名匹配；确定所获取的所述待检测 URL

对应的网页源代码中的属性类关键词中存在至少一个与预设属性类关键词匹配；确定所获取的所述待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配。

应理解，以上各个单元的划分仅仅是一种逻辑功能的划分，实际实现时可以全部或部分集成到一个物理实体上，也可以物理上分开。本发明实施例中，获取单元 401 和处理单元 402 可以由处理器 501 实现。如图 5 所示，钓鱼网站检测装置 500 可以包括处理器 501、存储器 502，可选的，还可以包括通信接口 503。其中，存储器 502 可以用于存储处理器 501 执行方案时的代码，该代码可为钓鱼网站检测装置 500 出厂时预装的程序/代码。

从上述内容可以看出：本发明实施例中提供一种钓鱼网站检测方法和装置，由于本发明实施例中从目标网站获取待检测 URL 对应的待检测 HTTPS 证书，待检测 URL 的域名正确，因此才能获取待检测 URL 对应的合法 HTTPS 证书；一个 URL 的域名对应一个合法 HTTPS 证书，从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书，根据待检测 HTTPS 证书和合法 HTTPS 证书，进而可有效检测用户正在访问的目标网站是否为钓鱼网站。

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、当使用软件程序实现时，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申请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。所述指令可以存储在计算机存储介质中，或者从一个计算机存储介质向另一个计算机存储介质传输，例如，所述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、磁光盘（MO）等）、光介质（例如，CD、DVD、BD、HVD 等）、或者半导体介质（例如 ROM、EPROM、EEPROM、非易失性存储器（NAND FLASH）、固态硬盘（Solid State Disk，SSD））等。

本领域内的技术人员应明白，本申请实施例可提供为方法、系统、或计算机程序产品。因此，本申请实施例可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本申请实施例是参照根据本申请实施例的方法、设备（系统）、和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由指令实现流程图和 / 或方框图中的每一流程和 / 或方框、以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

这些指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

这些指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

显然，本领域的技术人员可以对本申请实施例进行各种改动和变型而不脱离本申请的精神和范围。这样，倘若本申请实施例的这些修改和变型属于本申

请权利要求及其等同技术的范围之内，则本申请也意图包含这些改动和变型在内。

权 利 要 求

1、一种钓鱼网站检测方法，其特征在于，包括：

获取目标网站的待检测统一资源定位符 URL；所述待检测 URL 对应的域名为合法域名；

从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书；

从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书；

根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站。

2、如权利要求 1 所述的方法，其特征在于，所述根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站，包括：

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项全部匹配的情况下，确定所述目标网站为正常网站。

3、如权利要求 1 或 2 所述的方法，其特征在于，所述根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站，包括：

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：

在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配、根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻有效、且根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书未被注销，则确定所述目标网站为正常网站。

4、如权利要求 1 或 2 所述的方法，其特征在于，所述根据所述待检测

HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项, 确定所述目标网站是否为钓鱼网站, 包括:

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下:

确定所述待检测 HTTPS 证书和所述待检测 HTTPS 证书满足第一预设条件中的至少一项, 则确定所述目标网站为钓鱼网站; 其中, 所述第一预设条件包括:

在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配;

根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻失效;

根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书被注销。

5、如权利要求 1 所述的方法, 其特征在于, 所述获取目标网站的待检测 URL 之后, 从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前, 还包括:

确定所述待检测 URL 使用 HTTPS。

6、如权利要求 5 所述的方法, 其特征在于, 所述获取目标网站的待检测 URL 之后, 从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前, 还包括:

在确定所述待检测 URL 未使用 HTTPS 的情况下, 上报告警信息, 所述告警信息用于指示所述目标网站存在危险。

7、如权利要求 1 所述的方法, 其特征在于, 所述获取目标网站的待检测 URL 之后, 从所述目标网站获取待检测 URL 对应的待检测 HTTPS 证书之前, 还包括:

确定所述待检测 URL 满足第二预设条件中的任一项或任多项, 其中, 所述第二预设条件包括:

确定所述待检测 URL 的域名与至少一个预设的 URL 域名匹配;

确定所获取的所述待检测 URL 对应的网页源代码中的属性类关键词中存在至少一个与预设属性类关键词匹配;

确定所获取的所述待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配。

8、一种钓鱼网站检测装置，其特征在于，包括：

获取单元，用于获取目标网站的待检测统一资源定位符 URL；所述待检测 URL 对应的域名为合法域名；从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书；从所述待检测 URL 的域名对应的服务器上获取所述待检测 URL 的域名对应的合法 HTTPS 证书；

处理单元，用于根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站。

9、如权利要求 8 所述的装置，其特征在于，所述处理单元，用于：

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项全部匹配的情况下，确定所述目标网站为正常网站。

10、如权利要求 8 或 9 所述的装置，其特征在于，所述处理单元，用于：

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配、根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻有效、且根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书未被注销，则确定所述目标网站为正常网站。

11、如权利要求 8 或 9 所述的装置，其特征在于，所述处理单元，用于：

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：

确定所述待检测 HTTPS 证书和所述待检测 HTTPS 证书满足第一预设条

件中的至少一项，则确定所述目标网站为钓鱼网站；其中，所述第一预设条件包括：

在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配；

根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻失效；

根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书被注销。

12、如权利要求 8 所述的装置，其特征在于，所述处理单元，还用于：
确定所述待检测 URL 使用 HTTPS。

13、如权利要求 12 所述的装置，其特征在于，所述处理单元，还用于：
在确定所述待检测 URL 未使用 HTTPS 的情况下，上报告警信息，所述告警信息用于指示所述目标网站存在危险。

14、如权利要求 8 所述的装置，其特征在于，所述处理单元，还用于：
确定所述待检测 URL 满足第二预设条件中的任一项或任多项，其中，所述第二预设条件包括：

确定所述待检测 URL 的域名与至少一个预设的 URL 域名匹配；

确定所获取的所述待检测 URL 对应的网页源代码中的属性类关键词中存在至少一个与预设属性类关键词匹配；

确定所获取的所述待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配。

15、一种钓鱼网站检测装置，其特征在于，包括处理器和存储器；
所述存储器，用于存储可执行程序；

所述处理器用于读取所述存储器中的可执行程序，执行：

获取目标网站的待检测统一资源定位符 URL；所述待检测 URL 对应的域名为合法域名；从所述目标网站获取待检测 URL 对应的待检测安全超文本传输协议 HTTPS 证书；从所述待检测 URL 的域名对应的服务器上获取所述待

检测 URL 的域名对应的合法 HTTPS 证书；根据所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项，确定所述目标网站是否为钓鱼网站。

16、如权利要求 15 所述的装置，其特征在于，所述处理器，用于：

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项全部匹配的情况下，确定所述目标网站为正常网站。

17、如权利要求 15 或 16 所述的装置，其特征在于，所述处理器，用于：

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：

在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识匹配、根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻有效、且根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书未被注销，则确定所述目标网站为正常网站。

18、如权利要求 15 或 16 所述的装置，其特征在于，所述处理器，用于：

在确定所述待检测 HTTPS 证书中包括的预设信息项和所述合法 HTTPS 证书中包括的预设信息项中存在至少一项不匹配的情况下：

确定所述待检测 HTTPS 证书和所述待检测 HTTPS 证书满足第一预设条件中的至少一项，则确定所述目标网站为钓鱼网站；其中，所述第一预设条件包括：

在确定所述待检测 HTTPS 证书中包括的证书签发者标识和所述合法 HTTPS 证书中包括的证书签发者标识不匹配；

根据所述待检测 HTTPS 证书中的证书有效期确定所述待检测 HTTPS 证书在当前时刻失效；

根据所述待检测 HTTPS 证书中的证书是否被注销的信息确定所述待检测 HTTPS 证书被注销。

19、如权利要求 15 所述的装置，其特征在于，所述处理器，还用于：

确定所述待检测 URL 使用 HTTPS。

20、如权利要求 19 所述的装置，其特征在于，所述处理器，还用于：

在确定所述待检测 URL 未使用 HTTPS 的情况下，上报告警信息，所述告警信息用于指示所述目标网站存在危险。

21、如权利要求 15 所述的装置，其特征在于，所述处理器，还用于：

确定所述待检测 URL 满足第二预设条件中的任一项或任多项，其中，所述第二预设条件包括：

确定所述待检测 URL 的域名与至少一个预设的 URL 域名匹配；

确定所获取的所述待检测 URL 对应的网页源代码中的属性类关键词中存在至少一个与预设属性类关键词匹配；

确定所获取的所述待检测 URL 对应的网页源代码中输入类关键词中存在至少一个与预设输入类关键词匹配。

22、一种非暂态计算机可读存储介质，其特征在于，所述非暂态计算机可读存储介质存储计算机指令，所述计算机指令用于使计算机执行权利要求 1~7 任一权利要求所述方法。

23、一种计算机程序产品，其特征在于，所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，使所述计算机执行权利要求 1~7 任一权利要求所述方法。

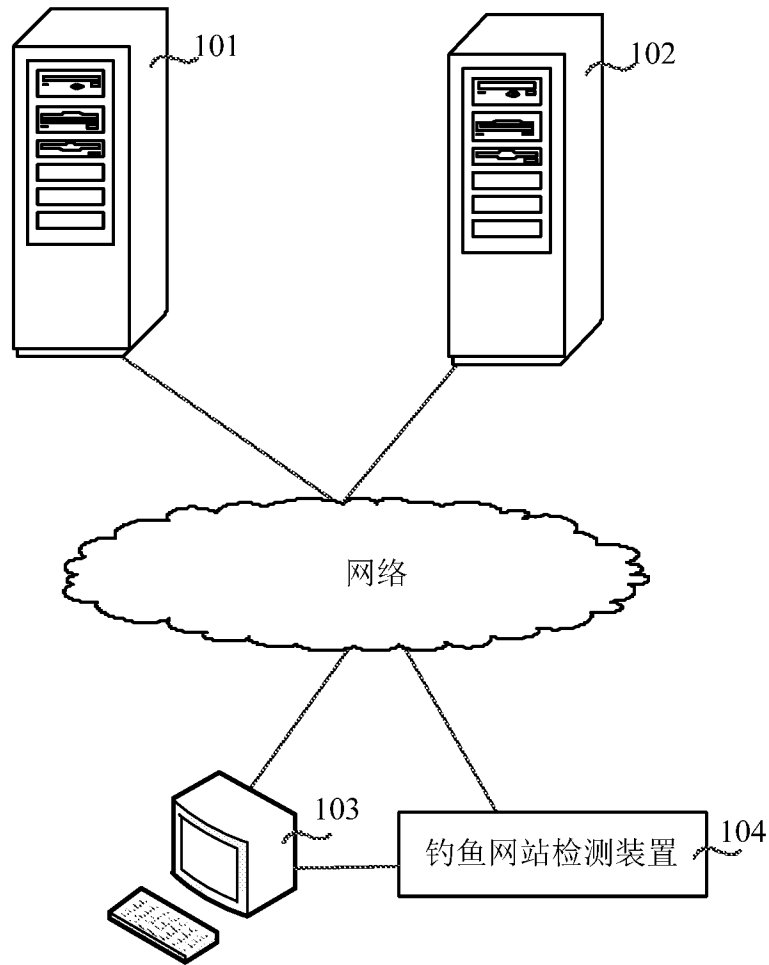


图 1

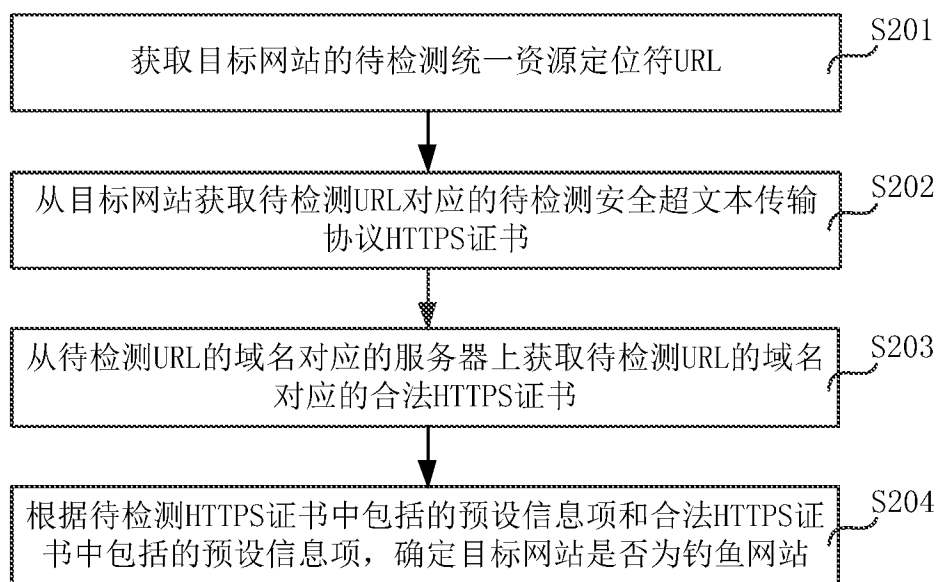


图 2

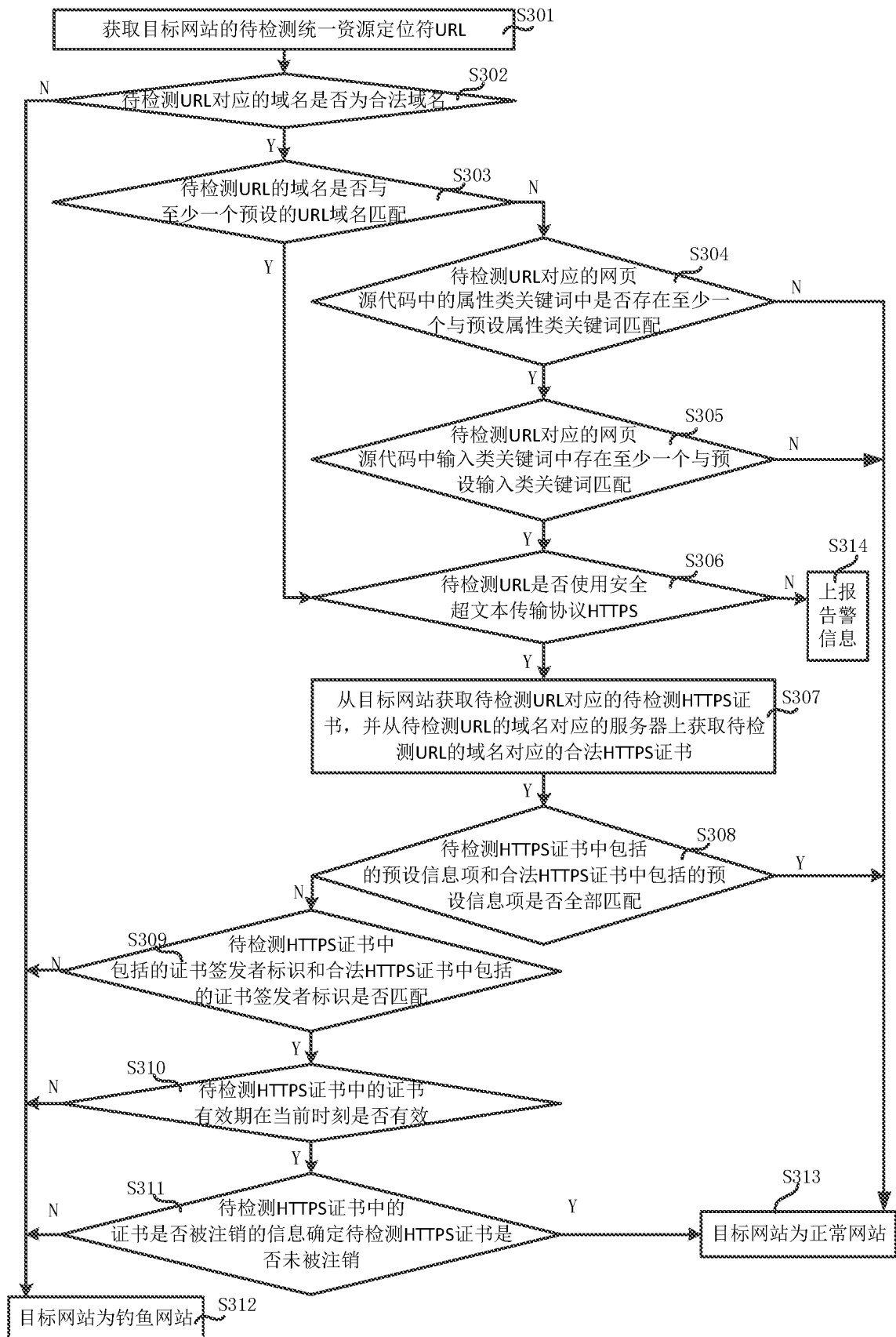


图 3

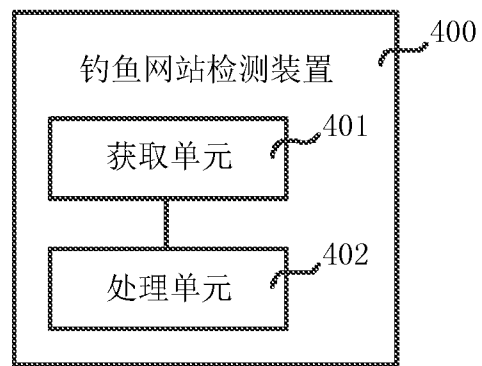


图 4

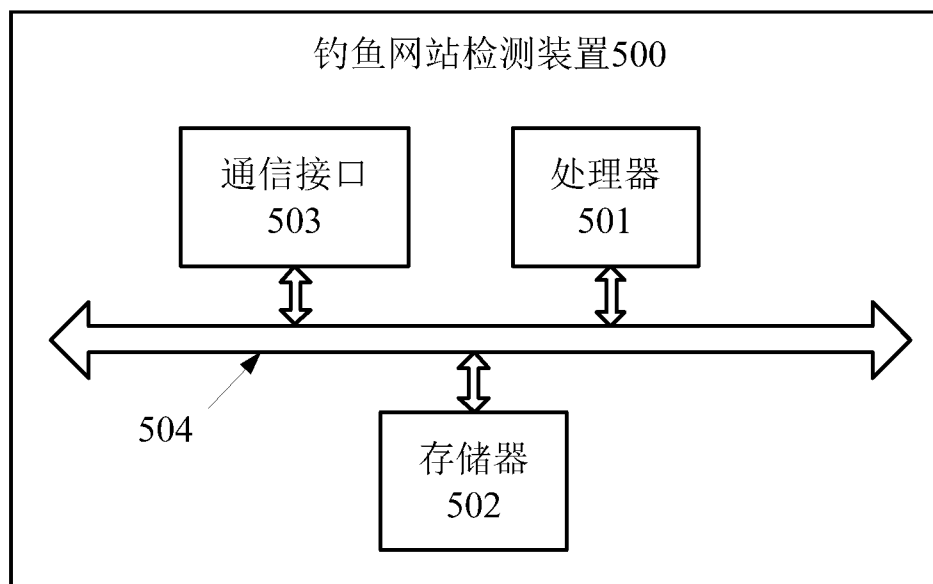


图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/107865

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; 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)

CNPAT, CNKI, WPI, EPODOC: 钓鱼, 恶意, 黑客, 网址, 统一资源定位符, 证书, 安全超文本传输协议, 域名系统; malice, hacker, address, URL, certificate, HTTPS, domain system

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 106789939 A (CHINA UNIONPAY CO., LTD.), 31 May 2017 (31.05.2017), claims 1-14, and description, paragraphs 0082-0084	1-23
A	CN 105516169 A (BEIJING QIHOO TECHNOLOGY CO., LTD. et al.), 20 April 2016 (20.04.2016), description, paragraphs 0027-0064	1-23
A	CN 101977235 A (BEIJING VRV SOFTWARE CO., LTD.), 16 February 2011 (16.02.2011), entire document	1-23
A	CN 101141447 A (FORTINET, INC.), 12 March 2008 (12.03.2008), entire document	1-23
A	US 7698442 B1 (VOLTAGE SECURITY, INC.), 13 April 2010 (13.04.2010), entire document	1-23

☐ 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 29 December 2017	Date of mailing of the international search report 25 January 2018
Name and mailing address of the ISA 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 PENG, Liang Telephone No. (86-10) 62413350

International application No.
PCT/CN2017/107865

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2017/107865

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; 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)) CNPAT, CNKI, WPI, EPODOC: 钓鱼, 恶意, 黑客, 网址, 统一资源定位符, 证书, 安全超文本传输协议, 域名系统; malice, hacker, address, URL, certificate, HTTPS, domain system																				
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 106789939 A (中国银联股份有限公司) 2017年 5月 31日 (2017 - 05 - 31) 权利要求第1-14项, 说明书第0082-0084段</td> <td>1-23</td> </tr> <tr> <td>A</td> <td>CN 105516169 A (北京奇虎科技有限公司 等) 2016年 4月 20日 (2016 - 04 - 20) 说明书第0027-0064段</td> <td>1-23</td> </tr> <tr> <td>A</td> <td>CN 101977235 A (北京北信源软件股份有限公司) 2011年 2月 16日 (2011 - 02 - 16) 全文</td> <td>1-23</td> </tr> <tr> <td>A</td> <td>CN 101141447 A (飞塔信息科技北京有限公司) 2008年 3月 12日 (2008 - 03 - 12) 全文</td> <td>1-23</td> </tr> <tr> <td>A</td> <td>US 7698442 B1 (VOLTAGE SECURITY, INC.) 2010年 4月 13日 (2010 - 04 - 13) 全文</td> <td>1-23</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 106789939 A (中国银联股份有限公司) 2017年 5月 31日 (2017 - 05 - 31) 权利要求第1-14项, 说明书第0082-0084段	1-23	A	CN 105516169 A (北京奇虎科技有限公司 等) 2016年 4月 20日 (2016 - 04 - 20) 说明书第0027-0064段	1-23	A	CN 101977235 A (北京北信源软件股份有限公司) 2011年 2月 16日 (2011 - 02 - 16) 全文	1-23	A	CN 101141447 A (飞塔信息科技北京有限公司) 2008年 3月 12日 (2008 - 03 - 12) 全文	1-23	A	US 7698442 B1 (VOLTAGE SECURITY, INC.) 2010年 4月 13日 (2010 - 04 - 13) 全文	1-23
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 106789939 A (中国银联股份有限公司) 2017年 5月 31日 (2017 - 05 - 31) 权利要求第1-14项, 说明书第0082-0084段	1-23																		
A	CN 105516169 A (北京奇虎科技有限公司 等) 2016年 4月 20日 (2016 - 04 - 20) 说明书第0027-0064段	1-23																		
A	CN 101977235 A (北京北信源软件股份有限公司) 2011年 2月 16日 (2011 - 02 - 16) 全文	1-23																		
A	CN 101141447 A (飞塔信息科技北京有限公司) 2008年 3月 12日 (2008 - 03 - 12) 全文	1-23																		
A	US 7698442 B1 (VOLTAGE SECURITY, INC.) 2010年 4月 13日 (2010 - 04 - 13) 全文	1-23																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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																
* 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																		
2017年 12月 29日		2018年 1月 25日																		
ISA/CN的名称和邮寄地址		受权官员																		
中华人民共和国国家知识产权局 (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		彭亮 电话号码 (86-10)62413350																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/107865

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	106789939	A	2017年 5月 31日	无	
CN	105516169	A	2016年 4月 20日	无	
CN	101977235	A	2011年 2月 16日	无	
CN	101141447	A	2008年 3月 12日	无	
US	7698442	B1	2010年 4月 13日	无	