(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
2013 年 8 月 15 日 (15.08.2013)

WIPO | PCT

(10) 国际公布号
WO 2013/117151 A1

(51) 国际专利分类号:
G06F 21/00 (2013.01)

(21) 国际申请号: PCT/CN2013/071383

(22) 国际申请日: 2013 年 2 月 5 日 (05.02.2013)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201210026760.0 2012 年 2 月 7 日 (07.02.2012) CN
201210026759.8 2012 年 2 月 7 日 (07.02.2012) CN
201210026585.5 2012 年 2 月 7 日 (07.02.2012) CN

(71) 申请人: **北京奇虎科技有限公司 (BEIJING QIHOO TECHNOLOGY COMPANY LIMITED)** [CN/CN]; 中国北京市西城区新街口外大街 28 号 D 座 112 室（德胜园区）, Beijing 100088 (CN)。 **奇智软件（北京）有限公司 (QIZHI SOFTWARE (BEIJING) COMPANY LIMITED)** [CN/CN]; 中国北京市朝阳区酒仙桥路 14 号兆维大厦 4 层东侧单元, Beijing 100016 (CN)。

(72) 发明人: **邹贵强 (ZOU, Guiqiang)**; 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。 **付旻 (FU, Min)**; 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。

(74) 代理人: **北京智汇东方知识产权代理事务所（普通合伙） (WISEAST INTELLECTUAL PROPERTY LAW FIRM)**; 中国北京市海淀区成府路 28 号优盛大厦 D-1111 室, Beijing 100083 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明，要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA,

[见续页]

(54) **Title**: METHOD AND SYSTEM FOR RAPIDLY SCANNING FILES

(54) **发明名称** : 文件快速扫描方法和系统

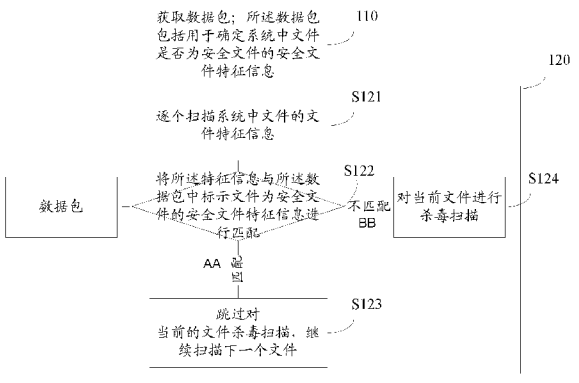

图 1 /Fig.1

110 OBTAINING A DATA PACKET, THE DATA PACKET COMPRISING SECURITY-FILE CHARACTERISTIC INFORMATION USED FOR DETERMINING WHETHER FILES IN A SYSTEM ARE SECURITY FILES
120 DATA PACKET
S121 SCANNING FILE CHARACTERISTIC INFORMATION ABOUT THE FILES IN THE SYSTEM ONE BY ONE
S122 MATCHING THE CHARACTERISTIC INFORMATION WITH THE SECURITY-FILE CHARACTERISTIC INFORMATION FOR INDICATING FILES AS SECURITY FILES IN THE DATA PACKET
S123 SKIPPING ANTI-VIRUS SCANNING FOR THE CURRENT FILE, AND CONTINUING TO SCAN THE NEXT FILE
S124 CONDUCTING ANTI-VIRUS SCANNING ON THE CURRENT FILE
AA MATCHING
BB MISMATCHING

(57) **Abstract**: A method and system for rapidly scanning files relate to the technical field of networks. The method comprises: obtaining a data packet, the data packet comprising security-file characteristic information used for determining whether files in a system are security files; and scanning file characteristic information about the files in the system one by one, if the currently scanned file characteristic information is matched with the security-file characteristic information for marking files as security files in the data packet, skipping anti-virus scanning for the current file, and continuing to scan the next file. In the method, using a data packet which comprises security-file characteristic information used for determining whether files in a system are security files, when a new user performs a first scanning, if scanning a file with characteristic information the same as that in the data packet, a secure file with a long time which is represented by the file can be skipped, and the first scanning time can be reduced.

(57) **摘要**: 一种文件快速扫描方法和系统，涉及网络技术领域。该方法包括：获取数据包；数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息；逐个扫描系统中文件的文件特征信息，若当前扫描的文件特征信息与数据包中标示文件为安全文件的安全文件特征信息匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件。该方法通过利用包括用于确定系统中文件是否为安全文件的安全文件特征信息的数据包，新的用户在第一次扫描时，如果扫描到与数据包中特征信息相同的文件时就可以跳过其代表的时间长、并且安全的文件，可减少首次扫描的时间。

WO 2013/117151 A1

RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

文件快速扫描方法和系统

技术领域

本发明涉及网络技术领域，特别是涉及一种文件快速扫描方法和系统。

背景技术

随着计算机的普及，用户端基本上都需要安装杀毒软件对计算机内的文件进行扫描。在杀毒软件进行扫描时，需要大量的CPU运算和磁盘操作，使得扫描过程漫长而且影响系统速度。而对于计算机里的文件，有很多文件都是相同的，比如Windows的文件，很多软件的安装包文件，帮助文件，压缩文件等。

现有技术中，用户端的杀毒软件第一次扫描时会全盘扫描计算机内的所有文件，并且会扫描文件中的各种内容，如果对于包括内容比较大的文件，其扫描的时间耗费的相当多，从而导致第一次扫描的时间非常的长。比如对于压缩包，现有技术会按照正常的扫描流程将压缩包中的内容解压逐个进行扫描，以保证文件是安全的，这样对于一个压缩包来说，现有技术的扫描时间比较长，从而延长了整个扫描的时间，而对于用户来说，由于长时间的扫描而长时间占用的CPU运算和磁盘操作影响了系统的速度，从而影响的用户对计算机的使用。

另外，通常杀毒软件会对上次扫描的安全文件记录进入白名单，然后在下一次扫描时，逐个对当前文件与白名单中文件进行匹配；如果匹配上，则跳过当前文件，进入下一个文件扫描过程；如果未匹配上，则对该未匹配上的文件的全部内容进行完整扫描，扫描完毕后再进入下一个文件扫描过程。在这个过程中，需要对上一个文件进行匹配或者完整扫描后才能对下一个文件进行匹配和扫描过程，下一个文件的等待时间比较长，导致整个杀毒软件对整个系统文件的扫描时间比较长的问题。

发明内容

鉴于上述问题，提出了本发明以便提供一种克服上述问题或者至少部分地解决或者减缓上述问题的文件快速扫描方法和系统。

根据本发明的一个方面，提供了一种文件快速扫描方法，其包括获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全

文件特征信息；逐个扫描系统中文件的文件特征信息，若当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件。

根据本发明的另一个方面，提供了一种文件快速扫描系统，其包括：第一用户端，所述第一用户端包括：

获取模块，用于获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息；

扫描匹配模块，用于逐个扫描系统中文件的文件特征信息，若当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件。

根据本发明的又一个方面，提供了一种计算机程序，其包括计算机可读代码，当所述计算机可读代码在服务器上运行时，导致所述服务器执行根据权利要求 1-11 中的任一个所述的方法。

根据本发明的再一个方面，提供了一种计算机可读介质，其中存储了如权利要求 24 所述的计算机程序。

本发明的有益效果为：

首先，本公开内容通过利用包括用于确定系统中文件是否为安全文件的安全文件特征信息的数据包，新的用户在第一次扫描时，如果扫描到与所述数据包中特征信息相同的文件时就可以跳过其代表的时间长、并且安全的文件，可减少首次扫描的时间。

其次，本公开内容通过设置两个扫描模块，其中特征信息扫描模块用于将第一用户端系统的文件的特征信息与数据包中安全文件的特征信息进行匹配，如果第一用户端系统的文件的特征信息未与数据包中安全文件的特征信息匹配上，则通知安全扫描模块对该文件的进行完整扫描，而特征信息扫描模块继续扫描下一个文件。通过上述方式，使第一用户端对于文件的特征信息的匹配过程和完整扫描过程相分离，上一个未匹配上的文件的扫描不会影响下一个文件的匹配，这样节省了其中的等待时间，从而加快了文件的扫描速度。

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

附图说明

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

图 1 是依据本发明一个实施例的文件快速扫描方法的流程示意图；

图 2 是依据本发明一个实施例的数据包生成方法的流程示意图；

图 3 是依据本发明一个实施例的压缩包头部数据的示例；

图 4 是依据本发明一个实施例的文件快速扫描系统的结构示意图；

图 5 是依据本发明一个实施例的文件快速扫描系统的结构示意图；

图 6 是依据本发明一个实施例的文件快速扫描系统的结构示意图；

图 7 是依据本发明一个实施例的文件快速扫描系统的结构示意图；

图 8 是依据本发明一个实施例的文件快速扫描方法的流程示意图；

图 9 是依据本发明一个实施例的数据包生成方法的流程示意图；

图 10 示意性地示出了用于执行根据本发明的方法的服务器的框图；以及

图 11 示意性地示出了用于保持或者携带实现根据本发明的方法的程序代码的存储单元。

具体实施例

下面结合附图和具体的实施方式对本发明作进一步的描述。

在实际中，对于处于一定数量级（比如 105）以上的用户端来说，如果在该数量级以上的用户端都对具有同样的特征（比如包括文件名、文件大小、文件修改时间、文件内容描述信息、内容信息等）的文件进行扫描时，如果该种文件是安全的，那么其他用户在使用具有同样特征的文件基本上也是安全的。本申请即根据该种特性统计巨量用户对完整扫描耗时比较长，并且安全的文件的特信息，然后基于所述特征信息生成待匹配的数据包（升级补丁等），用户在使用所述数据包后，可扫描文件的特征信息与所述数据包中的特征信息进行匹配，如果匹配上，那么即可跳过这些正常扫描时间比较长的文件，从而可达到节省扫描时间的目的。

参照图 1，其示出了本申请一种文件快速扫描方法的流程示意图，包括：

步骤 110，获取数据包；所述数据包包括用于确定系统中文件是否为

安全文件的安全文件特征信息。

新的用户端首先获取包括安全特征信息的数据包，然后新的用户端则可根据该数据包进行扫描，以节省第一次扫描的时间。其中新的用户端在本申请中采用第一用户端进行描述，所述第一用户端主要包括未安装杀毒软件的用户端和安装过杀毒软件但未进行过全盘文件扫描的用户端，还可以包括安装过杀毒软件并且进行过全盘文件扫描的用户端，但在最近一次进行全盘文件扫描后用户端中出现了新的未扫描过的文件的用户端。

第一用户端可通过安装存入所述数据包的安装包，或者将已安装的杀毒软件升级获取所述的数据包。在实际中，第一用户端可能完全没有进行过扫描，也可能在使用过程中可能会获取到很多新的文件，而这些文件用户端也可能没有进行完整扫描，当该用户端想进行快速扫描时，则可通过本申请的数据包进行。

其中，所述数据包中的确定系统中文件为安全文件的特征信息可通过统计各用户端的完整扫描结果获得。比如，对于多个用户进行完整扫描的结果，对于在各用户端中具有相同特征信息（比如文件大小、文件修改时间、文件名 CRC 值和内容匹配信息等特征信息）的文件，如果各用户端对该文件的扫描结果均安全，那么即可将该特征信息存入所述数据包，用于确定系统中具有该特征信息的文件为安全文件。

任选地，所述特征信息通过统计各用户端发送到处理中心服务器中的安全文件特征信息获得。即对于各用户端进行完整扫描得到的确认安全的文件的特征信息，处理中心服务器自动对其进行统计和分析，可将大于一定数量阈值的特征信息生成相应的数据包。在本申请中，首先需要生成数据包，参照图 2，示出了依据本发明一个实施例的数据包生成方法的流程示意图，包括:

步骤 210，接收各用户端发送的安全文件的特征信息；所述特征信息包括用户端完整扫描文件时，确定安全的文件的特征信息。

在实际中，存在第二用户端群，其中包括了大量用户端，这些用户端都都可将对自己系统中的文件进行完整扫描后的符合条件的特征信息发送至处理中心服务器中。即当第二用户端群的各用户端选择完整扫描其计算机中的文件时，会将扫描耗时大于或大于等于一定时间阈值的安全文件的特征信息发送到处理中心服务器。

在实际中，本申请通过对巨量的用户端扫描正常的扫描耗时较长的

文件的特征信息进行统计，然后基于大于或大于等于阈值的对应文件的特征信息生成待匹配的数据包，然后新的用户在第一次扫描时，如果扫描到与所述数据包中特征信息相同的文件时就可以跳过其代表的时间长、并且安全的文件，可减少首次扫描的时间。

任选地，在用户端完整扫描文件时：

所述的完整扫描为用户端系统中各文件的全部内容进行杀毒扫描。

步骤 S11，当所述扫描的文件安全时，获取文件的特征信息，所述特征信息包括文件名、文件大小、文件修改时间和文件内容描述信息。

用户端选择完整扫描文件时，扫描引擎逐个扫描用户端计算机中的各个文件，对于很多文件，比如对于压缩包，扫描引擎需要按照压缩包的要求在引擎中先解压，再对压缩包中的包括的各个文件进行杀毒扫描，一般情况下，这种扫描方式需要的时间相当的多；又比如对于软件的安装包，扫描引擎也需要将安装包中的信息解压出来再进行扫描，同样需要耗费相当多的时间。那么当用户端扫描整个文件花费的时间大于或者大于等于阈值时，则可获取当前文件的特征信息所述特征信息包括文件名、文件大小、文件修改时间和文件内容描述信息，所述特征信息包括文件名、文件大小、文件修改时间和文件内容描述信息。

当所述扫描时间大于或大于等于时间阈值并安全时，获取文件的特征信息，所述特征信息包括文件名、文件大小、文件修改时间和文件内容描述信息。其中，计算用户端对于一个文件的扫描时间时，可根据开始扫描文件时和结束扫描文件时的时间点，调用系统 API 函数 GetTickCount 计算从开始扫描文件和结束扫描文件时的扫描时间。比如开始扫描文件时间点为，2012 年 1 月 17 日星期二 21 时 50 分 30 秒和 00 毫秒，结束扫描文件时时间点为 2012 年 1 月 17 日星期二 21 时 51 分 30 秒和 00 毫秒，那么用户端对于当前文件的扫描时间即为 1 分钟。如果设置时间阈值为 30 秒，那么对于该文件来说，用户端即获取该文件的特征信息，包括文件名、文件大小、文件修改时间和文件内容描述信息。

其中文件内容描述信息只占整个文件的很小一部分，当用户端扫描文件内容描述信息部分时，所耗费时间远远小于扫描整个文件的时间。比如对于压缩文件，其文件内容描述信息在压缩包头部数据中，扫描时可根据头部数据告知的当前文件内容描述信息占用的字节数，用户端只需要扫描当前文件头部数据所在字节地址空间即可，参照图 3，其为一种压缩包的头部数据示例，在压缩包的头部数据中描述了压缩包内的文件

信息（文件名、大小、数据校验值等），只要这些关键数据没有变化，就说明压缩包内容没有变化。又比如对于软件的安装包，内容描述信息在安装包的尾部数据中，扫描时可根据文件告知的尾部数据占用的字节数，扫描相应字节地址空间即可。

步骤 S12，将文件名进行循环冗余码校验计算，得到文件名 CRC 值。

由于文件名涉及用户的隐私，本申请将文件名作一个循环冗余码校验（CRC，Cyclical Redundancy Check）计算，得到一个无可读性的文件名 CRC 值。

步骤 S13，将文件内容描述信息进行消息摘要算法计算，得到内容匹配信息。

对于文件内容描述信息，相对整个文件来说，文件内容描述信息虽然很小，但是如果以文件内容描述信息本身去进行匹配，则可能由于文件内容描述信息庞大而导致时间相对较多，本申请则将文件内容描述信息进行消息摘要算法（Message-Digest Algorithm）计算，得到内容匹配信息，一般本申请进行 MD5（Message-Digest Algorithm 5，消息摘要算法第五版）计算，得到的内容匹配信息为 MD5 值，通过 MD5 值进行匹配时能大大减少匹配时间，保证快速的对比匹配，同时也保证了文件的安全性。

步骤 S14，将包括文件大小、文件修改时间、文件名 CRC 值和内容匹配信息的特征信息进行发送。

在得到上述的包括文件大小、文件修改时间、文件名 CRC 值和内容匹配信息的特征信息之后，即可将所述特征信息发送至处理中心服务器，等待处理中心服务器处理。

本申请的应用环境包括了提供杀毒软件相关安装包，升级包等数据的处理中心服务器，大量的通过网络语处理中心服务器连接的用户端，因此对于在线用户对文件进行的完整扫描工作，均可进行上述步骤。

步骤 220，对于相同的特征信息，统计其重复次数。

在本步骤中，处理中心服务器会对收到的数据进行去重运算，即将相同客户端发送的多个具有相同特征信息的数据去重，使其次数为 1.

任选地，对于相同的特征信息，统计其重复次数:

S21，将接收的特征信息按总重复次数进行排序。

首先将接收到的特征信息按总重复次数进行排序，比如对于对应某些用户端文件的特征信息（m,100kb，2012/1/11/21:50:30:10,n）其中 m 为

文件名 CRC 值，n 为文件内容描述信息的内容匹配信息即 MD5 值，相应的每条特征信息对应一个发送该信息的用户端，比如以对用户端 A，其发送到处理中心服务器的信息可以 A-(m,100kb,2012/1/11/21:50:30:10,n)的形式表明。将处理中心服务器接收到的所有条具有相同（A,100kb,2012/1/11/21:50:30:10,B）统计其重复总次数，然后根据统计总次数进行排序。在该次统计排序的过程中，能方便的找出相同用户端发送的具有相同特征信息的数据。

S22，对于各相同特征信息，将属于同一个用户端发送的相同特征消息进行消重运算。

对于同一个客户端发送的多个具有相同特征信息的数据，比如客户端 A 发送特征信息为（m,100kb，2012/1/11/21:50:30:10,n）有 10 条，那么将其去重，使处理中心将该用户端重复发送的（m,100kb，2012/1/11/21:50:30:10,n）特征信息记为 1 次。如此可以保证对于某个特征信息进行统计时用户数量的准确性，保证本申请的有效性。

S23，针对消重后的各相同特征信息，统计各特征消息的重复次数。

再去重后，可统计各特征消息的重复次数，该次数与扫描得到该特征信息的用户端数量一致。

本申请还可通过其他方法进行消重运算，统计各个文件特征信息的重复次数，对此本申请不对其加以限制。

步骤 230，提取重复次数大于或大于等于数量阈值的特征信息。

在得到特征信息的重复次数以后，将重复次数与数量阈值进行比较，如果所述重复次数大于阈值，则将其存入数据文件。比如对于前述（m,100kb，2012/1/11/21:50:30:10,n）特征信息的重复次数为 15 万，而阈值设置为 10 万，那么可将该特征信息数据提取。

步骤 240，将所述提取出来的特征信息存入数据文件，并依据所述数据文件生成数据包。

在本申请中数据文件可以中的数据可以列表的形式存在，每条特征信息包括多个维度，比如对于（m,100kb，2012/1/11/21:50:30:10,n），则列表中该条特征信息包括四个维度，分别为文件大小、文件修改时间、文件名 CRC 值和内容匹配信息。然后在后续扫描匹配时以此四个维度进行匹配。

对于生成的数据文件，可以打包到杀毒软件的安装包，也可以生成升级补丁的形式。

另外，在所述数据包生成之前还包括:

步骤 S210，接收用户端发送的不安全文件的特征信息，不将该特征信息存入数据文件或者将数据包中与该特征信息相同的特征信息进行删除。

用户端对于完整扫描时间大于阈值的出现了病毒的文件，也会将该不安全文件的特征信息标记为不安全，再发送至处理中心服务器中，处理中心服务器对于具有该特征信息的所有条数据可均不进行处理，也不存入数据文件。比如有 5 万个用户端发送了前述（m,100kb，2012/1/11/21:50:30:10,n）特征信息的数据，但其中一个用户端发送的该特征信息的数据标识为不安全，即发现病毒，则处理中心服务器则可不对具有（m,100kb，2012/1/11/21:50:30:10,n）特征信息的数据进行处理，不将包括（m,100kb，2012/1/11/21:50:30:10,n）的特征信息存入数据包中。

另外，在实际中，处理中心服务器在对各用户端发送的特征信息进行统计时，一般以一定时间长度为周期进行统计，并在对该周期的统计结果生成安全文件的特性信息的数据包。比如处理中心服务器以天为时间单位对各用户端发送的特征信息数据进行统计，然后生成数据包，可提供给下一周期的第一用户端使用。

步骤 120，逐个扫描系统中文件的文件特征信息，若当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件。

比如对文件内容中包括的恶意代码或者病毒等进行杀毒扫描。

在获取到所述数据后，则可进行快速扫描:

步骤 S121，第一用户端的系统逐个扫描系统中文件的特征信息;

步骤 S122，将当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息进行匹配；若匹配时，则转入步骤 S123，跳过当前的文件，继续扫描下一个文件；若不匹配时，则转入步骤 S124，对当前文件进行完整的杀毒扫描，即扫描当前文件的所有内容。

在实际中，在第一用户端进行扫描之前还包括：确认是否选择快速扫描，如果是，则扫描文件的特征信息，并调用所述数据包进行匹配。

即第一用户端可以选择进行快速扫描或者是对文件的完整扫描，如果选择快速扫描，则可扫描文件的特征信息，并调用所述数据包进行匹配进行特征信息的匹配。当所述文件的特征信息与所述数据包中的特征信息匹配时，则跳过当前文件，继续扫描下一个文件。

在本申请中，当第一次扫描时，可提示第一用户端是否选择快速扫描，如果选择则扫描文件的特征信息，并调用所述数据包进行匹配。

如果第一用户端选择进行快速扫描，那么第一用户端在扫描时则首先获取第一用户端文件的特征信息进行匹配，而不用扫描文件的全部内容。

在第一用户端将所述文件的特征信息与所述数据包中的特征信息进行匹配时：

在缓存中从匹配效率最高的特征信息开始进行匹配。

由于数据包中特征信息是多维的，在判定是否匹配时，可以从效率最高的数据开始比较，比如不需要进行而外计算的维度，例如文件大小和文件修改时间，因为文件大小和修改时间是在遍历文件时即可获得的系统数据，不需要额外的运算，那么如此操作可以提高预置缓存的比较匹配的效率。

任选地，在匹配时：

步骤 S31，当文件大小和文件修改时间与数据库中的一条特征信息匹配时，则将文件名进行所述循环冗余码校验计算，得到文件名 CRC 值，并将所述文件名 CRC 值与该条特征信息的文件名 CRC 值进行匹配。其中，文件名的 CRC 运算是内存操作，数据量很小。

在匹配时，首先从匹配效率较高的维度进行匹配，对于数据包中一条特征信息维度中，文件大小和文件修改时间不需要进行额外计算，可直接进行匹配，那么当用户端扫描时，对于获取到的特征信息的文件大小和文件修改时间，比如可首先比较文件大小，在比较文件修改时间。若当前扫描的文件的文件大小和文件修改时间相同时，则比较计算量相对较小的维度，比如通过 CRC 运算计算文件名 CRC 值，然后将所述文件名 CRC 值与该条特征信息的文件名 CRC 值进行匹配，如果未匹配上，则进入完整扫描，如果匹配上，则进入计算量相对较多的维度进行匹配，比如转入步骤 S32。

步骤 S32，当所述文件名 CRC 值与该条特征信息的文件名 CRC 值匹配时，则将文件内容描述信息进行所述消息摘要算法计算，得到内容匹配信息，并将所述内容匹配信息与该条特征信息的内容匹配信息进行匹配。

当文件大小，文件修改时间和文件名 CRC 值都匹配上时，则将文件内容描述信息进行消息摘要算法计算，一般是进行 MD5 计算，得到文件

内容匹配信息，则将所述内容匹配信息与该条特征信息的内容匹配信息进行匹配，当匹配上，则跳过当前的文件，转入扫描下一个文件。

在本申请中，匹配时，如果对于特征信息中，有一个维度未匹配上，即表示该文件未匹配上，即可将该文件进行杀毒扫描。比如前述的特征信息的四个维度：文件大小，文件修改时间，文件名 CRC 值和内容匹配信息，该四个维度的匹配顺序为：1、文件大小，2、文件修改时间，3、文件名 CRC 值 4、内容匹配信息；那么如果一个将文件 A 的文件大小与数据包中的特征信息的第一维度，即文件大小进行匹配时未匹配上时，可不用进行 2、3 和 4 维度的匹配，即可文件 A 需要进行杀毒扫描；如果文件 A 的文件大小匹配上，再将文件 A 的文件修改时间与数据包中的该条特征信息的第 2 维度，即文件修改时间进行匹配时未匹配上，那么可不用进行 3 和 4 维度的匹配，即可将文件 A 进行杀毒扫描。其他情况可以依此类推。

另外，在每次扫描时，任选地，在第一用户端进行扫描时，将当次扫描结果中安全文件的特征信息存入所述数据包的安全文件信息列表；当第一用户端下一次进行扫描时根据上一次扫描后的记录进行扫描。

在第一用户端结合当前的数据包进行扫描时，对于特征信息未包含在数据包中的文件，如果当次检测安全，则可将其该安全文件的特征信息存入数据包的安全文件信息列表中，当第一用户端进行下一次扫描时，可根据上一次的扫描结果更为快速的进行扫描。另外，当用户端对各文件进行完整扫描时，如果检测到某个文件不安全，但该文件的特征信息又在数据包中，则可将其从数据包中的安全文件信息列表中删除。

相应的，参照图 4，本申请还公开了一种文件快速扫描系统的结构示意图，包括：

第一用户端 410，所述第一用户端包括：

获取模块 S411，用于获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息；

扫描匹配模块 S412，用于逐个扫描系统中文件的文件特征信息，若当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件。

参照图 5，其示出了依据本发明一个实施例的文件快速扫描系统的结构示意图，包括：

第一用户端 510，第二用户端群 520 和处理中心服务器 530;

所述第一用户端包括 510:

获取模块 S511，用于获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息;

扫描匹配模块 S512，用于逐个扫描系统中文件的文件特征信息，若当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件;

所述处理中心服务器 530 用于统计各用户端发送到处理中心服务器中的安全文件特征信息获得所述数据包中的特征信息;

所述第二用户端群 520 用于发送的安全文件的特征信息。

任选地，所述的处理中心服务器包括:

特征信息模块，用于接收各用户端发送的安全文件的特征信息；所述特征信息包括用户端完整扫描文件时，确定安全的文件的特征信息;

统计模块，用于对于相同的特征信息，统计其重复次数;

提取模块，用于提取重复次数大于或大于等于数量阈值的特征信息。

任选地，所述第二用户端群的各用户端包括:

特征信息获取单元，用于当扫描的文件安全时，获取文件的特征信息，所述特征信息包括文件名、文件大小、文件修改时间和文件内容描述信息;

文件名计算单元，用于将文件名进行循环冗余码校验计算，得到文件名 CRC 值;

内容描述信息计算单元，用于将文件内容描述信息进行消息摘要算法计算，得到内容匹配信息;

特征信息发送单元，用于将包括文件大小、文件修改时间、文件名 CRC 值和内容匹配信息的特征信息进行发送。

任选地，所述统计模块包括：排序单元，用于将接收的特征信息按总重复次数进行排序;

消重单元，用于对于各相同特征信息，将属于同一个用户端发送的相同特征消息进行消重运算;

统计单元，用于针对消重后的各相同特征信息，统计各特征消息的重复次数。

任选地，在用户端完整扫描文件时:

根据开始扫描文件时和结束扫描文件时的时间点，调用系统 API 函数 GetTickCount 计算从开始扫描文件和结束扫描文件时的扫描时间。

任选地，在提取模块之后还包括:

生成单元，用于将所述提取出来的特征信息存入数据文件，并依据所述数据文件生成数据包。

任选地，在所述数据包生成之前还包括:

去除单元，用于接收用户端发送的不安全文件的特征信息，不将该特征信息存入数据文件或者将数据包中与该特征信息相同的特征信息进行删除。

任选地，通过如下方法将所述文件的特征信息与所述数据包中的特征信息进行匹配:

在缓存中从匹配效率最高的特征信息开始进行匹配。

任选地，当文件大小和文件修改时间与数据库中的一条特征信息匹配时，则将文件名进行所述循环冗余码校验计算，得到文件名 CRC 值，并将所述文件名 CRC 值与该条特征信息的文件名 CRC 值进行匹配;

当所述文件名 CRC 值与该条特征信息的文件名 CRC 值匹配时，则将文件内容描述信息进行所述消息摘要算法计算，得到内容匹配信息，并将所述内容匹配信息与该条特征信息的内容匹配信息进行匹配。

任选地，还包括:

提醒模块，用于当第一次扫描时，提示第一用户端是否选择快速扫描，如果选择则扫描文件的特征信息，并调用所述数据包进行匹配。

任选地，还包括：更新记录模块，用于在第一用户端进行扫描时，将当次扫描结果中安全文件的特征信息存入所述数据包的安全文件信息列表；当第一用户端下一次进行扫描时根据上一次扫描后的安全记录结果进行扫描。

对于系统实施例而言，由于其与方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

参照图 6，其示出了依据本发明的一个实施例的文件快速扫描系统的结构示意图，包括：第一用户端 610，所述第一用户端包括:

获取模块 S611，用于获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息;

特征信息扫描模块 S612，用于逐个扫描系统中文件的特征信息，并将所述特征信息与所述数据包中的特征信息进行匹配；当匹配上时，跳过当前文件,继续扫描下一个文件；当未匹配上时，第跳过当前文件,继续

扫描下一个文件，并通知安全扫描模块对该文件进行完整扫描。

安全扫描模块 S613，用于对于未匹配上的文件，对该文件进行完整扫描。

参照图 7，其示出了依据本发明一个实施例的文件快速扫描系统的结构示意图，包括：

第一用户端 710，第二用户端群 720 和处理中心服务器 730；

所述第一用户端包括 710：

所述第一用户端包括：

获取模块 S711，用于获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息。

特征信息扫描模块 S712，用于特征信息扫描模块扫描系统中文件的特征信息，并将所述特征信息与所述数据包中标示文件为安全文件的安全文件特征信息进行匹配；对于匹配上的文件，不对该文件进行杀毒扫描，对于未匹配上的文件，通知安全扫描模块对该文件进行的杀毒扫描。

安全扫描模块 S713，用于用于对于未匹配上的文件，对该文件进行杀毒扫描。

所述处理中心服务器 730 用于统计各用户端发送到处理中心服务器中的安全文件特征信息获得所述安全文件的特征信息；

所述第二用户端群 720 用于发送的安全文件的特征信息。

任选地，所述的处理中心服务器包括：

特征信息模块，用于接收各用户端发送的安全文件的特征信息；所述特征信息包括用户端完整扫描文件时，确定安全的文件的特征信息；

统计模块，用于对于相同的特征信息，统计其重复次数；

提取模块，用于提取重复次数大于或大于等于数量阈值的特征信息。

任选地，所述第二用户端群的各用户端包括：

特征信息获取单元，用于当所述扫描的文件安全时，获取文件的特征信息，所述特征信息包括文件名、文件大小、文件修改时间和文件内容描述信息；

文件名计算单元，用于将文件名进行循环冗余码校验计算，得到文件名 CRC 值；

内容描述信息计算单元，用于将文件内容描述信息进行消息摘要算法计算，得到内容匹配信息；

特征信息发送单元，用于将包括文件大小、文件修改时间、文件名 CRC 值和内容匹配信息的特征信息进行发送。

任选地，所述统计模块包括：排序单元，用于将接收的特征信息按总重复次数进行排序；

消重单元，用于对于各相同特征信息，将属于同一个用户端发送的相同特征消息进行消重运算；

统计单元，用于针对消重后的各相同特征信息，统计各特征消息的重复次数。

任选地，在用户端完整扫描文件时：

根据开始扫描文件时和结束扫描文件时的时间点，调用系统 API 函数 GetTickCount 计算从开始扫描文件和结束扫描文件时的扫描时间。

任选地，在提取模块之后还包括：

生成单元，用于将所述提取出来的特征信息存入数据文件，并依据所述数据文件生成数据包。

任选地，在所述数据包生成之前还包括：

去除单元，用于接收用户端发送的不安全文件的特征信息，不将该特征信息存入数据文件或者将数据包中与该特征信息相同的特征信息进行删除。

任选地，通过如下方法将所述文件的特征信息与所述数据包中的特征信息进行匹配：

在缓存中从匹配效率最高的特征信息开始进行匹配。

任选地，当文件大小和文件修改时间与数据库中的一条特征信息匹配时，则将文件名进行所述循环冗余码校验计算，得到文件名 CRC 值，并将所述文件名 CRC 值与该条特征信息的文件名 CRC 值进行匹配；

当所述文件名 CRC 值与该条特征信息的文件名 CRC 值匹配时，则将文件内容描述信息进行所述消息摘要算法计算，得到内容匹配信息，并将所述内容匹配信息与该条特征信息的内容匹配信息进行匹配。

任选地，还包括：

提醒模块，用于当第一次扫描时，提示第一用户端是否选择快速扫描，如果选择则扫描文件的特征信息，并调用所述数据包进行匹配。

参照图 8，其示出了依据本发明的一个实施例的文件快速扫描方法的流程示意图，包括：

步骤 810，获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息。

步骤 820，逐个扫描系统中文件的文件特征信息，将当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息进行匹配；若匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件；若未匹配时，对当前文件进行杀毒扫描，当扫描时间大于或大于等于时间阈值并且安全时，将该文件的安全文件特征信息更新至数据包。

比如对文件内容中包括的恶意代码或者病毒等进行杀毒扫描。

在获取到所述数据后，则可进行快速扫描：

步骤 S821，第一用户端的系统逐个扫描系统中文件的特征信息；

步骤 S822，将当前扫描的文件特征信息与所述数据包中标示文件为

安全文件的安全文件特征信息进行匹配；若匹配时，则转入步骤 S823，跳过当前的文件，继续扫描下一个文件；若不匹配时，则转入步骤 S824，对当前文件进行完整的杀毒扫描， 当扫描时间大于或大于等于时间阈值并且安全时，转入步骤 S825，将该文件的安全文件特征信息更新至数据包。

在实际中，步骤 S825 中的特征信息可通过如下步骤获得:

步骤 M1，获取文件的特征信息，所述特征信息包括文件名、文件大小、文件修改时间和文件内容描述信息;

步骤 M2，将文件名进行循环冗余码校验计算，得到文件名 CRC 值;

步骤 M3，将文件内容描述信息进行消息摘要算法计算，得到内容匹配信息;

步骤 M4，将包括文件大小、文件修改时间、文件名 CRC 值和内容匹配信息的作为安全文件特征信息。步骤 M1 至步骤 M4 具体方法与前述步骤 S11 至 S14 相似。

其中，更新数据包的过程中，可直接将所述安全文件的特征信息更新至本地系统中的数据包中，也可将所述安全文件的特征信息发送至处理中心服务器中，由处理中心服务器根据其规则更新至对应各应用端的数据包中。

在实际中，在第一用户端进行扫描之前还包括：确认是否选择快速扫描，如果是，则扫描文件的特征信息，并调用所述数据包进行匹配。

即第一用户端可以选择进行快速扫描或者是对文件的完整扫描，如果选择快速扫描，则可扫描文件的特征信息，并调用所述数据包进行匹配进行特征信息的匹配。当所述文件的特征信息与所述数据包中的特征信息匹配时，则跳过当前文件，继续扫描下一个文件。

在本申请中，当第一次扫描时，可提示第一用户端是否选择快速扫描，如果选择则扫描文件的特征信息，并调用所述数据包进行匹配。

如果第一用户端选择进行快速扫描，那么第一用户端在扫描时则首先获取第一用户端文件的特征信息进行匹配，而不用扫描文件的全部内容。

参照图 9，示出了依据本发明一个实施例的数据包生成方法的流程示意图，包括:

步骤 910，接收各用户端发送的安全文件的特征信息；所述特征信息包括用户端完整扫描文件时，文件的扫描时间大于或大于等于时间阈值

的安全的文件的特征信息。

任选的，在用户端完整扫描文件时，当所述扫描时间大于或大于等于时间阈值并安全时：所述的完整扫描为用户端系统中各文件的全部内容进行杀毒扫描。

步骤 920，对于相同的特征信息，统计其重复次数。

步骤 930，提取重复次数大于或大于等于数量阈值的特征信息。

步骤 940，将所述提取出来的特征信息存入数据文件，并依据所述数据文件生成数据包。

对于与其它实施例类似的步骤，在这里不再详述。

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

例如，图 10 示出了可以实现根据本发明的文件快速扫描方法的服务器，例如应用服务器。该服务器传统上包括处理器 1100 和以存储器 1200 形式的计算机程序产品或者计算机可读介质。存储器 1200 可以是诸如闪存、EEPROM（电可擦除可编程只读存储器）、EPROM、硬盘或者 ROM 之类的电子存储器。存储器 1200 具有用于执行上述方法中的任何方法步骤的程序代码 1310 的存储空间 1300。例如，用于程序代码的存储空间 1300 可以包括分别用于实现上面的方法中的各种步骤的各个程序代码 1310。这些程序代码可以从一个或者多个计算机程序产品中读出或者写入到这一个或者多个计算机程序产品中。这些计算机程序产品包括诸如硬盘，紧致盘（CD）、存储卡或者软盘之类的程序代码载体。这样的计算机程序产品通常为如参考图 11 所述的便携式或者固定存储单元。该存储单元可以具有与图 10 的服务器中的存储器 1200 类似布置的存储段、存储空间等。程序代码可以例如以适当形式进行压缩。通常，存储单元包括计算机可读代码 1310'，即可以由例如诸如 1100 之类的处理器读取的代码，这些代码当由服务器运行时，导致该服务器执行上面所描述的

方法中的各个步骤。

本文中所称的“一个实施例”、“实施例”或者“一个或者多个实施例”意味着，结合实施例描述的特定特征、结构或者特性包括在本发明的至少一个实施例中。此外，请注意，这里“在一个实施例中”的词语例子不一定全指同一个实施例。

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

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

此外，还应当注意，本说明书中使用的语言主要是为了可读性和教导的目的而选择的，而不是为了解释或者限定本发明的主题而选择的。因此，在不偏离所附权利要求书的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。对于本发明的范围，对本发明所做的公开是说明性的，而非限制性的，本发明的范围由所附权利要求书限定。

权 利 要 求

1、一种文件快速扫描方法，其包括:

获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息;

逐个扫描系统中文件的文件特征信息，若当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件。

2、根据权利要求 1 所述的方法，其中:

若未匹配时，对当前文件进行杀毒扫描，当扫描时间大于或大于等于时间阈值并且安全时，将该文件的安全文件特征信息更新至数据包。

3、根据权利要求 1 或 2 所述的方法，其中:

所述特征信息通过统计各用户端发送到处理中心服务器中的安全文件特征信息获得。

4、根据权利要求 3 所述的方法，其中，通过如下步骤统计各用户端发送到处理中心服务器中的安全文件特征信息获得所述特征信息:

接收各用户端发送的安全文件的特征信息；所述特征信息包括用户端完整扫描文件时，确定安全的文件的特征信息;

对于相同的特征信息，统计其重复次数;

提取重复次数大于或大于等于数量阈值的特征信息。

5、根据权利要求 4 所述的方法，其中，在用户端完整扫描文件时:

当所述扫描的文件安全时，获取文件的特征信息，所述特征信息包括文件名、文件大小、文件修改时间和文件内容描述信息;

将文件名进行循环冗余码校验计算，得到文件名 CRC 值;

将文件内容描述信息进行消息摘要算法计算，得到内容匹配信息;

将包括文件大小、文件修改时间、文件名 CRC 值和内容匹配信息的特征信息进行发送。

6、根据权利要求 3 或 5 所述的方法，其中，对于相同的特征信息，统计其重复次数:

将接收的特征信息按总重复次数进行排序;

对于各相同特征信息，将属于同一个用户端发送的相同特征消息进行消重运算;

针对消重后的各相同特征信息，统计各特征消息的重复次数。

7、根据权利要求 4 所述的方法，其中，在提取重复次数大于或大于等于数量阈值的特征信息之后包括:

将所述提取出来的特征信息存入数据文件，并依据所述数据文件生成数据包。

8、根据权利要求 3 所述的方法，其中，在所述数据包生成之前还包

括:

接收用户端发送的不安全文件的特征信息，不将该特征信息存入数据文件或者将数据包中与该特征信息相同的特征信息进行删除。

9、根据权利要求 3 所述的方法，其中:

在生成上一个数据包后，当统计到新的安全文件的特征信息后，更新至上一个数据包。

10、根据权利要求 1 或 2 所述的方法，其中，通过如下方法将所述文件的特征信息与所述数据包中的特征信息进行匹配:

在缓存中从匹配效率最高的特征信息开始进行匹配。

11、根据权利要求 10 所述的方法，其中:

当文件大小和文件修改时间与数据库中的一条特征信息匹配时，则将文件名进行所述循环冗余码校验计算，得到文件名 CRC 值，并将所述文件名 CRC 值与该条特征信息的文件名 CRC 值进行匹配;

当所述文件名 CRC 值与该条特征信息的文件名 CRC 值匹配时，则将文件内容描述信息进行所述消息摘要算法计算，得到内容匹配信息，并将所述内容匹配信息与该条特征信息的内容匹配信息进行匹配。

12、一种文件快速扫描系统，其包括:

第一用户端，所述第一用户端包括:

获取模块，用于获取数据包; 所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息;

扫描匹配模块，用于逐个扫描系统中文件的文件特征信息，若当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息匹配时，则跳过对当前文件的杀毒扫描，继续扫描下一个文件。

13、根据权利要求 12 所述的系统，其中，还包括:

第二用户端群和处理中心服务器;

所述处理中心服务器用于统计各用户端发送到处理中心服务器中的安全文件特征信息获得所述数据包中的特征信息;

所述第二用户端群用于发送的安全文件的特征信息。

14、根据权利要求 13 所述的系统，其中，所述的处理中心服务器包括:

特征信息模块，用于接收各用户端发送的安全文件的特征信息; 所述特征信息包括用户端完整扫描文件时，确定安全的文件的特征信息;

统计模块，用于对于相同的特征信息，统计其重复次数;

提取模块，用于提取重复次数大于或大于等于数量阈值的特征信息。

15、根据权利要求 13 所述的系统，其中，所述第二用户端群的各用户端包括:

特征信息获取单元，用于当扫描的文件安全时，获取文件的特征信息，所述特征信息包括文件名、文件大小、文件修改时间和文件内容描

述信息;

文件名计算单元，用于将文件名进行循环冗余码校验计算，得到文件名 CRC 值;

内容描述信息计算单元，用于将文件内容描述信息进行消息摘要算法计算，得到内容匹配信息;

特征信息发送单元，用于将包括文件大小、文件修改时间、文件名 CRC 值和内容匹配信息的特征信息进行发送。

16、根据权利要求 14 所述的系统，其中，所述统计模块包括：排序单元，用于将接收的特征信息按总重复次数进行排序;

消重单元，用于对于各相同特征信息，将属于同一个用户端发送的相同特征消息进行消重运算;

统计单元，用于针对消重后的各相同特征信息，统计各特征消息的重复次数。

17、根据权利要求 14 所述的系统，其中，在提取模块之后还包括:

生成单元，用于将所述提取出来的特征信息存入数据文件，并依据所述数据文件生成数据包。

18、根据权利要求 14 所述的系统，其中，在所述数据包生成之前还包括:

去除单元，用于接收用户端发送的不安全文件的特征信息，不将该特征信息存入数据文件或者将数据包中与该特征信息相同的特征信息进行删除。

19、根据权利要求 14 所述的系统，其中，还包括:

更新模块，用于在生成上一个数据包后，当统计到新的安全文件的特征信息后，更新至上一个数据包。

20、根据权利要求 15 所述的系统，其中，通过如下方法将所述文件的特征信息与所述数据包中的特征信息进行匹配:

在缓存中从匹配效率最高的特征信息开始进行匹配。

21、根据权利要求 20 所述的系统，其中:

当文件大小和文件修改时间与数据库中的一条特征信息匹配时，则将文件名进行所述循环冗余码校验计算，得到文件名 CRC 值，并将所述文件名 CRC 值与该条特征信息的文件名 CRC 值进行匹配;

当所述文件名 CRC 值与该条特征信息的文件名 CRC 值匹配时，则将文件内容描述信息进行所述消息摘要算法计算，得到内容匹配信息，并将所述内容匹配信息与该条特征信息的内容匹配信息进行匹配。

22、根据权利要求 12 所述的系统，其中，还包括:

提醒模块，用于当第一次扫描时，提示第一用户端是否选择快速扫描，如果选择则扫描文件的特征信息，并调用所述数据包进行匹配。

23、根据权利要求 12 所述的方法，其中，还包括:

更新记录模块，用于在第一用户端进行扫描时，将当次扫描结果中安全文件的特征信息存入所述数据包的安全文件信息列表；当第一用户端下一次进行扫描时根据上一次扫描后的安全记录结果进行扫描。

24、一种计算机程序，包括计算机可读代码，当所述计算机可读代码在服务器上运行时，导致所述服务器执行根据权利要求 1-11 中的任一个所述的文件快速扫描方法。

25、一种计算机可读介质，其中存储了如权利要求 24 所述的计算机程序。

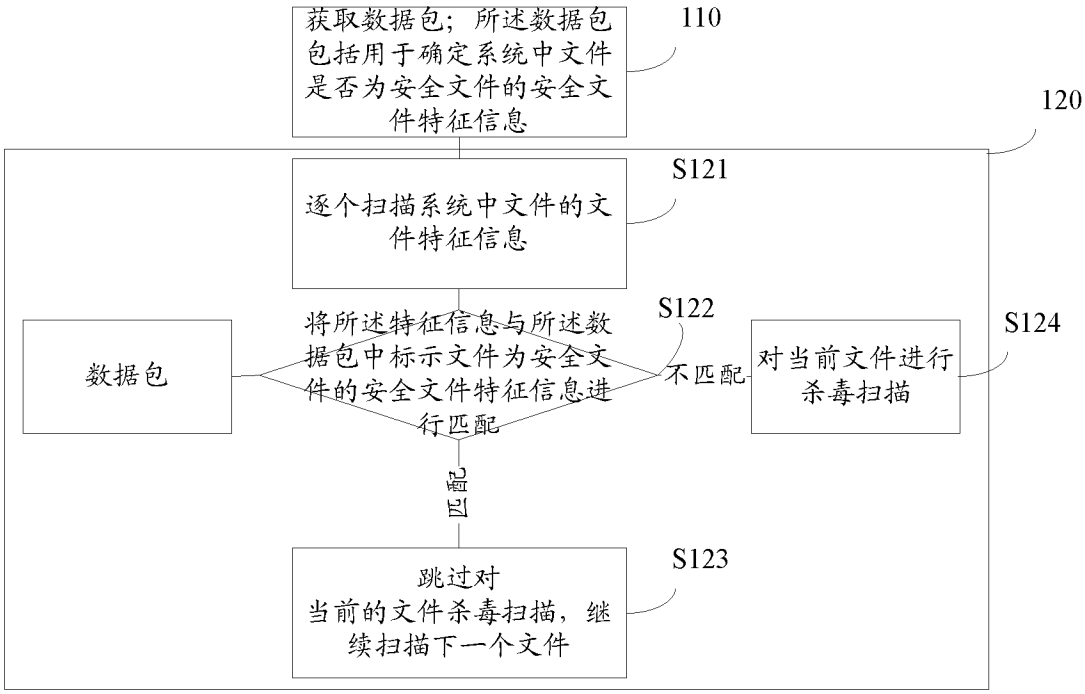
获取数据包；所述数据包包括用于确定系统中文件是否为安全文件的安全文件特征信息
110
120
S121
逐个扫描系统中文件的文件特征信息
S122
数据包
将所述特征信息与所述数据包中标示文件为安全文件的安全文件特征信息进行匹配
不匹配
对当前文件进行杀毒扫描
S124
匹配
S123
跳过对当前的文件杀毒扫描，继续扫描下一个文件

图 1

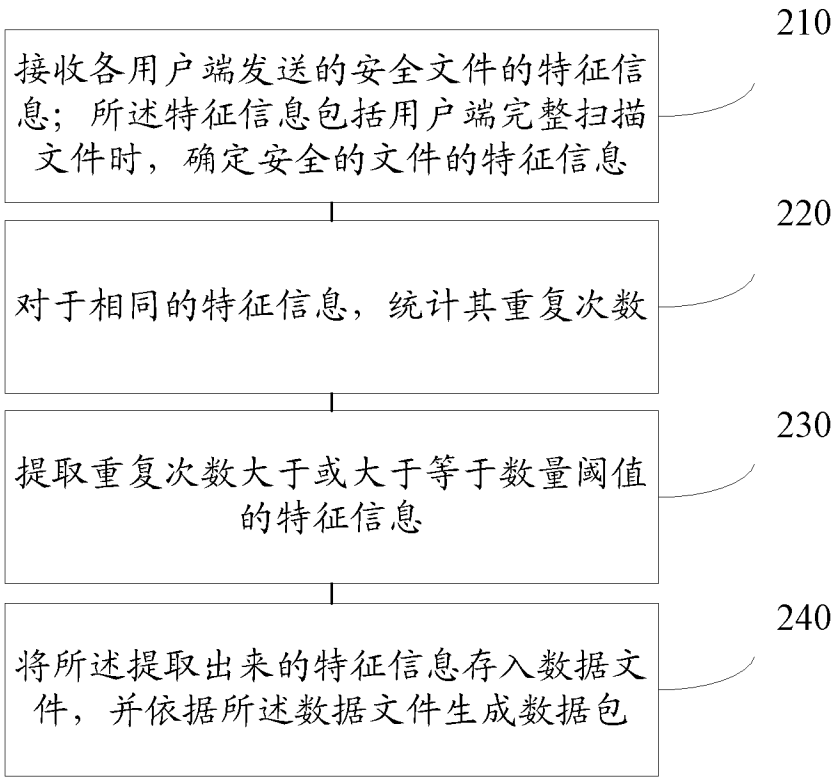
210
接收各用户端发送的安全文件的特征信息；所述特征信息包括用户端完整扫描文件时，确定安全的文件的特征信息
220
对于相同的特征信息，统计其重复次数
230
提取重复次数大于或大于等于数量阈值的特征信息
240
将所述提取出来的特征信息存入数据文件，并依据所述数据文件生成数据包

图 2

```
000000  52 61 72 21 1A 07 00 CF  90 73 00 00 0D 00 00 00  Rar!.....s......
000010  00 00 00 00 30 3D 74 C0  90 2E 00 06 E9 34 01 58  ....0=t.....4.X
000020  54 D2 85 02 F7 84 E2 66  74 5B 2C 40 1D 33 09 00  T.......ft[,@.3..
000030  20 20 00 00 33 36 30 73  64 2E 44 4D 50 00 B0 EE    ..360sd.DMP...
000040  C5 10 00 21 91 11 0C 89  26 66 14 02 6F 64 00 40  ...!....&f..od.@
000050  20 9A 00 00 A0 6E AA 00  20 06 6C 64 21 02 30 76   ....n.. .ld!.0v
000060  13 C0 90 27 86 92 43 C5  62 CD 9E 21 20 4F 0D 27  ...'..C.b..! O.'
000070  89 E0 A8 60 A2 B0 51 4D  1E 4E 63 C0 AC 7B 9F 00  ...j..QM.Nc..{..
000080  7C A7 00 20 00 AE AF 2F  70 F3 CF 31 79 5E 9F 22  |.. .../p..1y^."
000090  01 F3 1F 11 D7 C3 84 04  09 F0 FD 00 AB C0 12 40  ...............@
0000a0  80 BC CD F1 92 61 36 41  EF E2 FE FF 8C 8E 6E 66  .....a6A......nf
0000b0  5D D5 D5 D5 D5 D5 D5 D5  DE 5E 61 9B 9B 0F FC 6E  ].........^a...n
0000c0  56 70 95 9F 65 67 08 59  F9 B5 57 95 F5 66 5D EE  Vp..eg.Y..W..f].
0000d0  FE 5E FD 93 45 9A 10 25  19 24 C7 F7 78 39 F6 50  .^..E..%.$..x9.P
0000e0  02 B2 17 18 57 E1 E0 65  B9 09 4E 77 08 20 E1 F4  ....W..e..Nw. ..
0000f0  26 74 DF 18 15 81 1C 89  B7 F5 06 0E 34 25 93 83  &t..........4%..
000100  D8 61 4D B6 F6 58 E4 0B  32 E1 3B 3C 39 22 F6 71  .aM..X..2.;<9".q
000110  34 A6 1F ED 2C 2D 36 FD  77 86 11 3F 16 3B D0 4C  4...,-6.w..?.;.L
000120  38 4D 68 EF E3 81 1E E0  70 02 4C 1E 66 2C 4F 66  8Mh.....p.L.f,Of
000130  97 9B 52 F9 37 E5 97 92  AA 5E 4D 28 7C 38 9F 3A  ..R.7....^M(|8.:
000140  99 5F B2 53 D4 75 14 14  73 D2 14 74 F3 D4 3A 8B  ._.S.u..s..t..:.
000150  CF 0F 4B 1F CF EA F3 BD  58 2D CB 38 11 26 42 D8  ..K.....X-.8.&B.
000160  1A 43 D7 63 B8 8F 88 49  C1 F2 D7 B7 5A AC 11 28  .C.c...I....Z..(
```

图 3

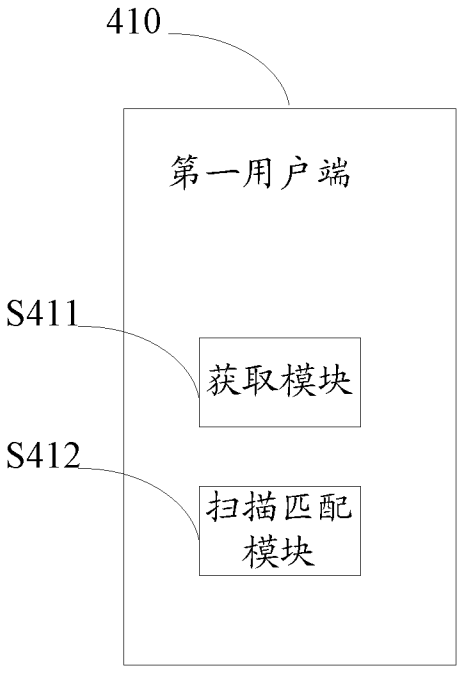

图 4

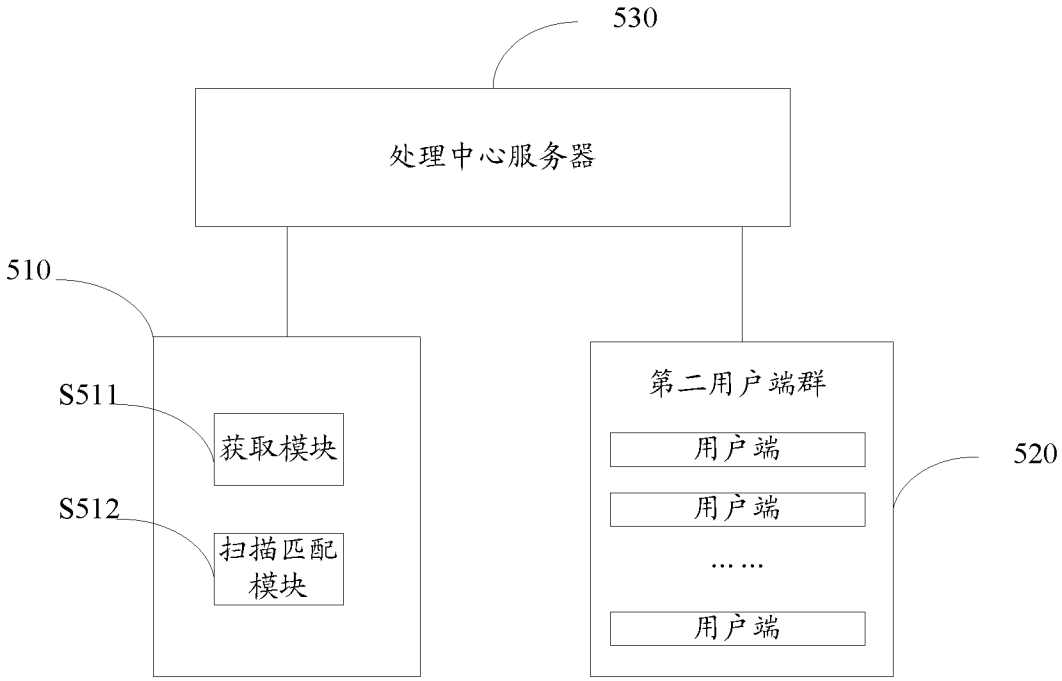
530
处理中心服务器
510
S511
获取模块
S512
扫描匹配模块
第二用户端群
用户端
用户端
……
用户端
520

图 5

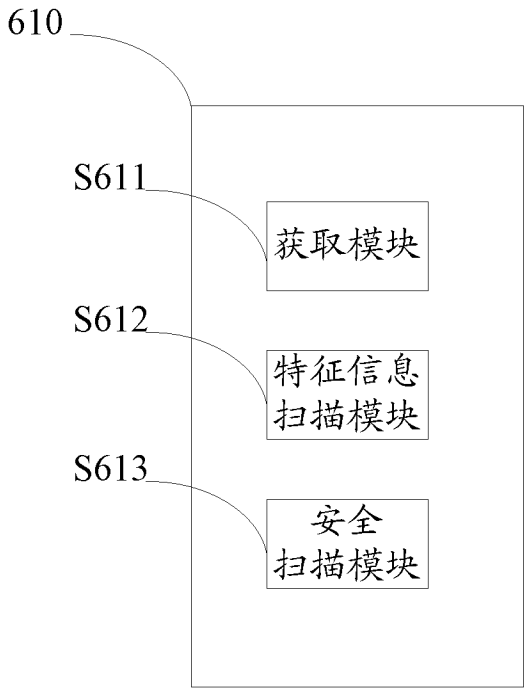
610
S611
获取模块
S612
特征信息扫描模块
S613
安全扫描模块

图 6

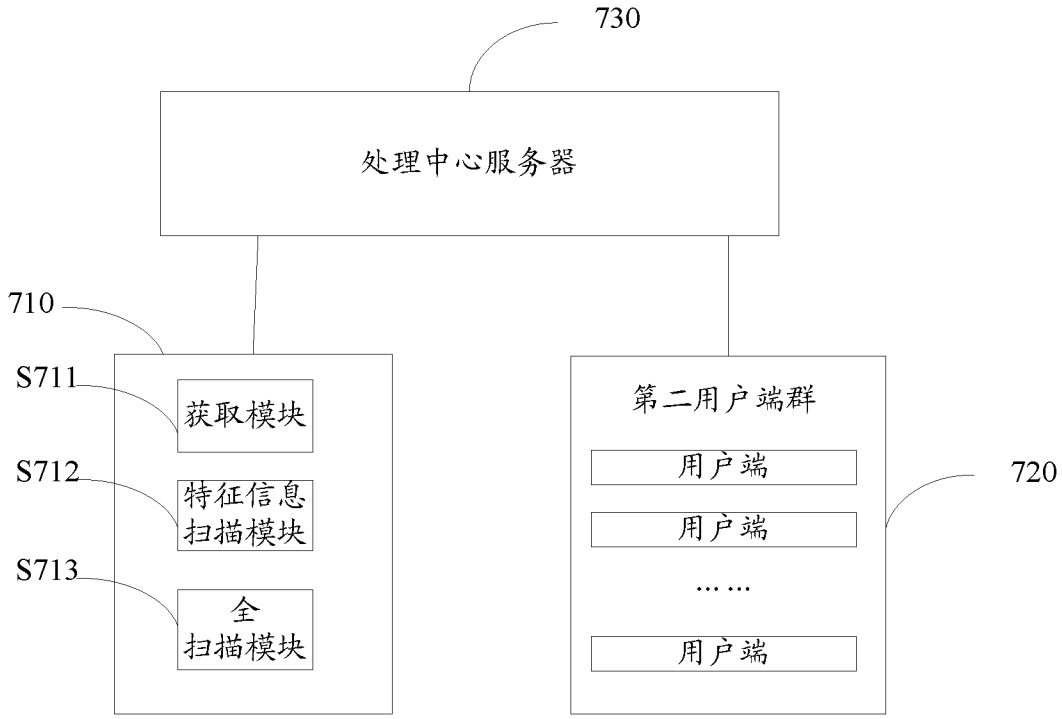
730
处理中心服务器
710
S711
获取模块
S712
特征信息
扫描模块
S713
全
扫描模块
第二用户端群
用户端
用户端
……
用户端
720

图 7

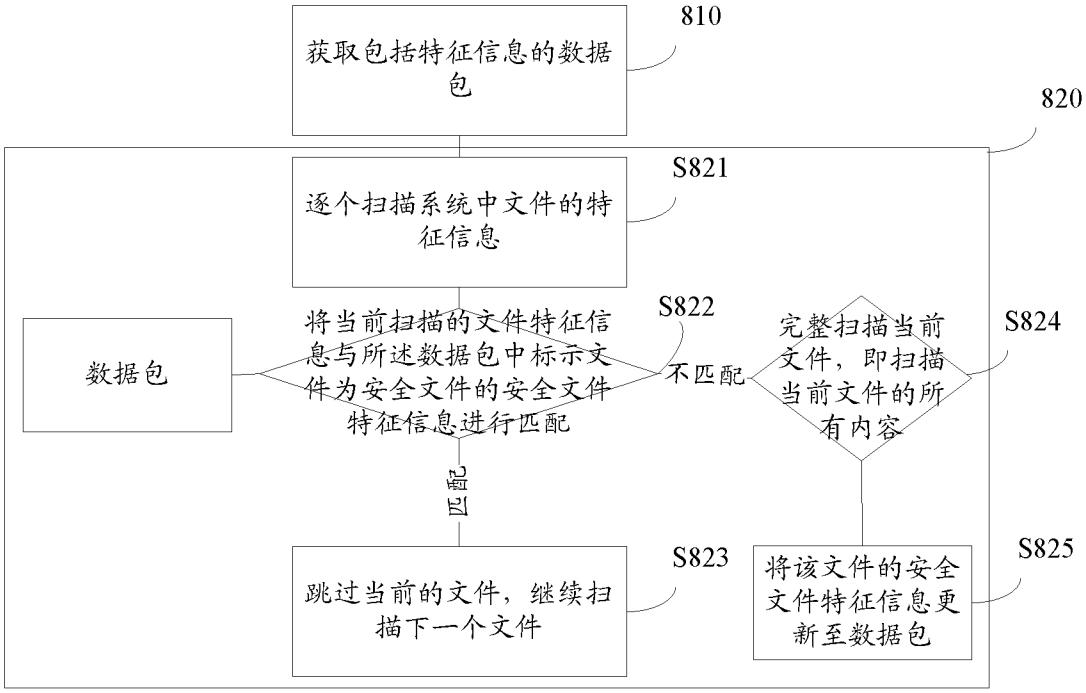
获取包括特征信息的数据包
810
820
逐个扫描系统中文件的特征信息
S821
数据包
将当前扫描的文件特征信息与所述数据包中标示文件为安全文件的安全文件特征信息进行匹配
S822
不匹配
完整扫描当前文件，即扫描当前文件的所有内容
S824
匹配
跳过当前的文件，继续扫描下一个文件
S823
将该文件的安全文件特征信息更新至数据包
S825

图 8

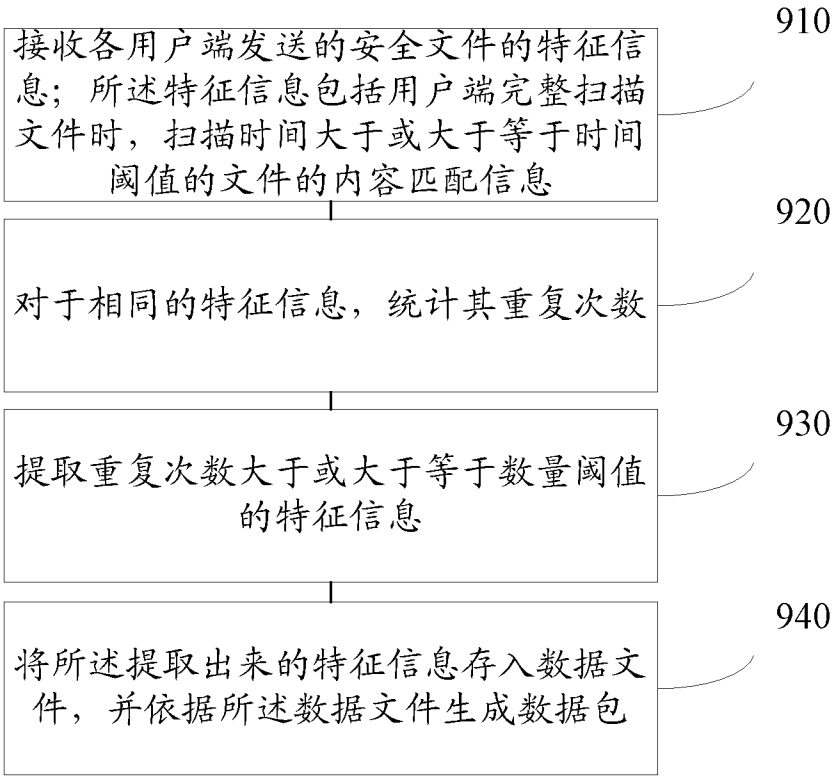
接收各用户端发送的安全文件的特征信息；所述特征信息包括用户端完整扫描文件时，扫描时间大于或大于等于时间阈值的文件的内容匹配信息
910
对于相同的特征信息，统计其重复次数
920
提取重复次数大于或大于等于数量阈值的特征信息
930
将所述提取出来的特征信息存入数据文件，并依据所述数据文件生成数据包
940

图 9

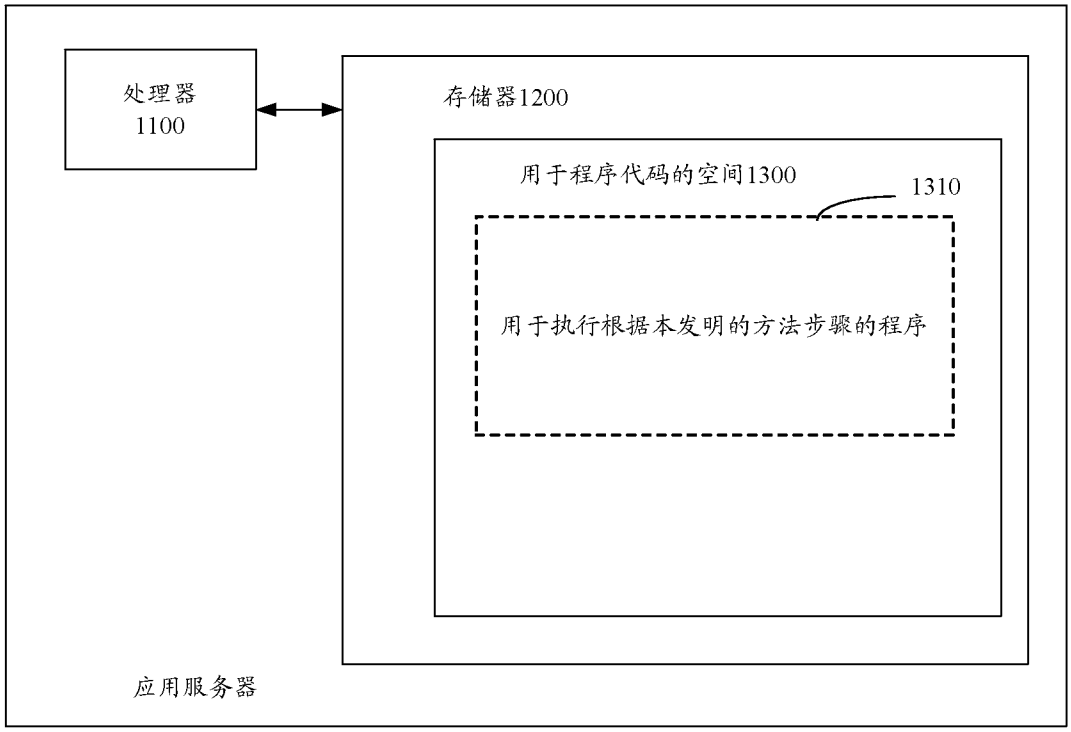
处理器
1100
存储器1200
用于程序代码的空间1300
1310
用于执行根据本发明的方法步骤的程序
应用服务器

图 10

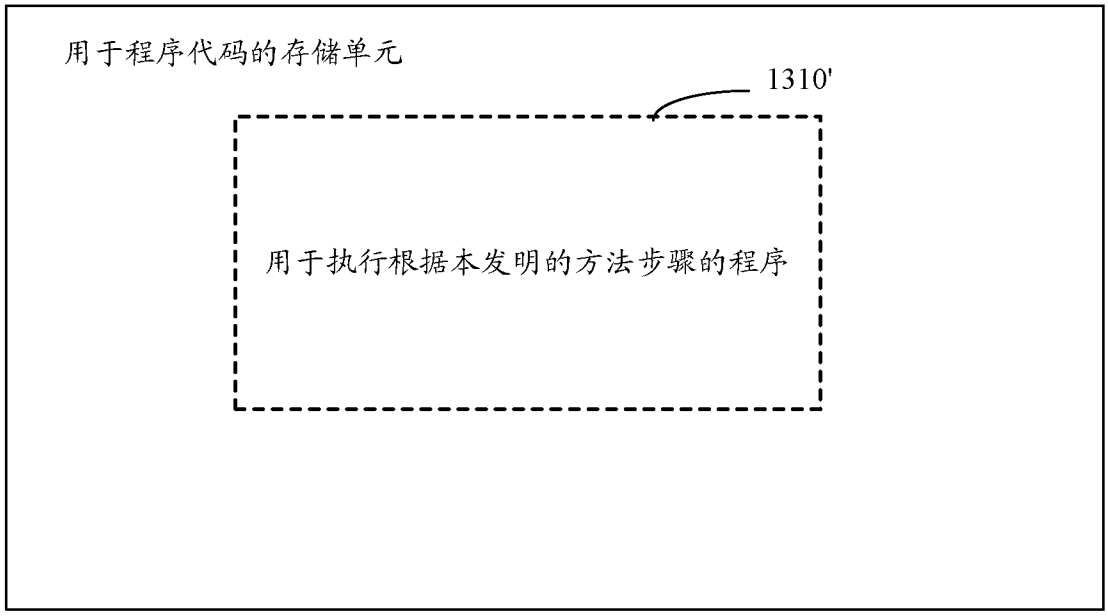
用于程序代码的存储单元
1310'
用于执行根据本发明的方法步骤的程序

图 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/071383

A. CLASSIFICATION OF SUBJECT MATTER

G06F 21/00 (2013.01) i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F 21/-

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: file, scan, feature, match, antivirus, data, packet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 1964357 A (BEIJING KINGSOFT SOFTWARE CO LTD) 16 May 2007 (16.05.2007) see the abstract, description page 7 to page 8	1, 3, 4, 7-10, 12, 24, 25
A	CN 101639880 A (HUAWEI TECHNOLOGIES CO LTD) 03 February 2010 (03.02.2010) see the whole document	1-25
A	CN 101308533 A (HUAWEI TECHNOLOGIES CO LTD) 19 November 2008 (19.11.2008) see the whole document	1-25
P, X	CN 102609515 A (QIZHI SOFTWARE BEIJING CO LTD) 25 July 2012 (25.07.2012) see claims 1 to 24, description page 4 to page 11	1-25
P, X	CN 102609653 A (QIZHI SOFTWARE BEIJING CO LTD) 25 July 2012 (25.07.2012) see claims 1 to 24, description page 4 to page 11	1-25
P, X	CN 102594809 A (QIZHI SOFTWARE BEIJING CO LTD) 18 July 2012 (18. 07.2012) see claims 1 to 24, description page 4 to page 11	1-25

☐ 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
12 April 2013 (12.04.2013)	**02 May 2013 (02.05.2013)**
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 MING, Mei Telephone No. (86-10) **62411851**

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/071383

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 1964357 A	16.05.2007	None	
CN 101639880 A	03.02.2010	WO 2010012175 A1	04.02.2010
CN 101308533 A	19.11.2008	None	
CN 102609515 A	25.07.2012	None	
CN 102609653 A	25.07.2012	None	
CN 102594809 A	18. 07.2012	None	

国际检索报告

国际申请号
PCT/CN2013/071383

A. 主题的分类

G06F 21/00 (2013.01) i

按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC：G06F 21/-

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

文件，扫描，特征，匹配，杀毒，数据包；file, scan, feature, match, antivirus, data, packet

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
X	CN1964357A （北京金山软件有限公司） 16.5 月 2007（16.05.2007） 参见参见摘要，说明书第 7－8 页	1, 3, 4, 7-10, 12, 24, 25
A	CN101639880A （华为技术有限公司） 03.2 月 2010（03.02.2010） 参见全文	1-25
A	CN101308533A （华为技术有限公司） 19.11 月 2008（19.11.2008） 参见全文	1-25
P，X	CN102609515A （奇智软件（北京）有限公司） 25.7 月 2012（25.07.2012） 参见权利要求 1-24，说明书第 4－11 页	1-25
P，X	CN102609653A （奇智软件（北京）有限公司） 25.7 月 2012（25.07.2012） 参见权利要求 1-24，说明书第 4－11 页	1-25
P，X	CN102594809A （奇智软件（北京）有限公司） 18.7 月 2012（18.07.2012） 参见权利要求 1-24，说明书第 4－11 页	1-25

☐ 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
12.4 月 2013(12.04.2013)	**02.5 月 2013 (02.05.2013)**
ISA/CN 的名称和邮寄地址： 中华人民共和国国家知识产权局 中国北京市海淀区蓟门桥西土城路 6 号 100088 传真号：　(86-10)62019451	受权官员 明媚 电话号码：(86-10) **62411851**

国际检索报告

关于同族专利的信息

国际申请号
PCT/CN2013/071383

检索报告中引用的专利文件	公布日期	同族专利	公布日期
CN1964357A	16.05.2007	无	
CN101639880A	03.02.2010	WO2010012175A1	04.02.2010
CN101308533A	19.11.2008	无	
CN102609515A	25.07.2012	无	
CN102609653A	25.07.2012	无	
CN102594809A	18.07.2012	无	