(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 106295334 B
(45)授权公告日 2019.07.26

(21)申请号 201510307070.6

(22)申请日 2015.06.05

(65)同一申请的已公布的文献号
申请公布号 CN 106295334 A

(43)申请公布日 2017.01.04

(73)专利权人 阿里巴巴集团控股有限公司
地址 英属开曼群岛大开曼资本大厦一座四层847号邮箱

(72)发明人 郑瀚

(74)专利代理机构 北京博浩百睿知识产权代理有限责任公司 11134
代理人 宋子良

(51)Int.Cl.
G06F 21/56(2013.01)

(56)对比文件
CN 102622556 A,2012.08.01,
CN 104182689 A,2014.12.03,
CN 103955449 A,2014.07.30,
CN 104199925 A,2014.12.10,
CN 103886258 A,2014.06.25,

审查员 张曼

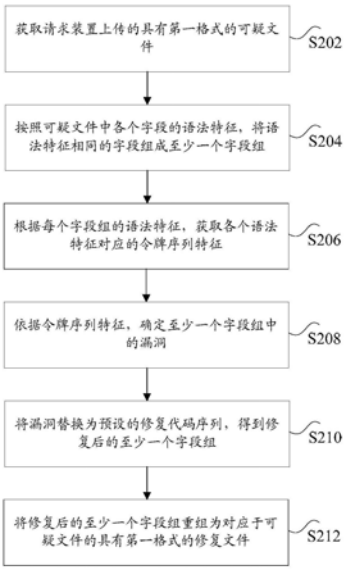
权利要求书4页 说明书18页 附图10页

(54)发明名称

文件修复方法及装置

(57)摘要

本申请公开了一种文件修复方法及装置。其中,该方法包括:获取请求装置上传的具有第一格式的可疑文件;按照所述可疑文件中各个字段的语法特征,将语法特征相同的字段组成至少一个字段组;根据每个字段组的语法特征,获取各个语法特征对应的令牌序列特征;依据所述令牌序列特征,确定所述至少一个字段组中的漏洞;将所述漏洞替换为预设的修复代码序列,得到修复后的所述至少一个字段组;将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件。本申请解决了由于现有技术采用同一套Patch代码实行造成的漏洞修复准确性较差的技术问题。



CN 106295334 B

1.一种文件修复方法,其特征在于,包括:

获取请求装置上传的具有第一格式的可疑文件;

按照所述可疑文件中各个字段的语法特征,将语法特征相同的字段组成至少一个字段组;

根据每个字段组的语法特征,获取各个语法特征对应的令牌序列特征;

依据所述令牌序列特征,确定所述至少一个字段组中的漏洞;

将所述漏洞替换为预设的修复代码序列,得到修复后的所述至少一个字段组;

将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件。

2.根据权利要求1所述的方法,其特征在于,在所述可疑文件的数量为至少两个的情况下,在所述获取请求装置上传的具有第一格式的可疑文件之后,所述按照所述可疑文件中各个字段的语法特征,将语法特征相同的字段组成至少一个字段组之前,所述方法还包括:

提取各个可疑文件的元信息,其中,所述元信息包括对应可疑文件的文件路径及消息摘要算法MD5,所述文件路径是指所述可疑文件在所述请求装置所在的用户设备上的路径;

根据所述文件路径及所述MD5,对所述各个可疑文件进行去重处理。

3.根据权利要求2所述的方法,其特征在于,所述根据所述文件路径及所述MD5,对所述各个可疑文件进行去重处理包括:

确定所述各个可疑文件中重复的可疑文件,其中,所述重复的可疑文件是指所述文件路径相同且所述MD5相同的文件;

保留所述重复的可疑文件中的第一文件,将除所述第一文件以外的所述重复的可疑文件删除,其中,所述第一文件为所述重复的可疑文件中的任意一个。

4.根据权利要求1所述的方法,其特征在于,所述依据所述令牌序列特征,确定所述至少一个字段组中的漏洞包括:

在所述至少一个字段组中查找与所述令牌序列特征相同的字段,将所述与所述令牌序列特征相同的字段确定为所述漏洞。

5.根据权利要求1所述的方法,其特征在于,所述将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件包括:

按照所述修复后的所述至少一个字段组中各个字段的父子关系,将所述修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件。

6.根据权利要求1所述的方法,其特征在于,在所述将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件之后,所述方法还包括:

接收所述请求装置发送的第一请求,所述第一请求用于下载更新后的漏洞修复程序;

向所述请求装置发送所述更新后的漏洞修复程序,其中,所述更新后的漏洞修复程序中包含所述修复文件,以便所述请求装置根据所述修复文件对所述可疑文件进行修复。

7.根据权利要求6所述的方法,其特征在于,在所述接收所述请求装置发送的第一请求之前,所述方法还包括:

接收来自所述请求装置发送的第二请求,所述第二请求用于获取所述更新后的漏洞修复程序的标识信息;

向所述请求装置发送所述更新后的漏洞修复程序的MD5,其中,由所述请求装置根据所

述更新后的漏洞修复程序的MD5，判断所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序是否相同，若相同，所述请求装置则发起所述第一请求。

8.根据权利要求7所述的方法，其特征在于，所述请求装置根据所述更新后的漏洞修复程序的MD5，判断所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序是否相同包括：

所述请求装置获取所述已存储的漏洞修复程序的MD5；

所述请求装置判断所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5是否相同；

若所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5相同，则确定所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序相同；

若所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5不相同，则确定所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序不相同。

9.一种文件修复方法，其特征在于，包括：

将具有第一格式的可疑文件发送至文件修复装置，其中，由所述文件修复装置按照所述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组，根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征，依据所述令牌序列特征，确定所述至少一个字段组中的漏洞，将所述漏洞替换为预设的修复代码序列，得到修复后的所述至少一个字段组，将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件；

从所述文件修复装置中获取对应于所述可疑文件的具有所述第一格式的修复文件。

10.根据权利要求9所述的方法，其特征在于，所述从所述文件修复装置中获取对应于所述可疑文件的具有所述第一格式的修复文件包括：

向所述文件修复装置发送第一请求，所述第一请求用于下载更新后的漏洞修复程序；

接收所述文件修复装置返回的所述更新后的漏洞修复程序，其中，所述更新后的漏洞修复程序中包含所述修复文件。

11.根据权利要求10所述的方法，其特征在于，在向所述文件修复装置发送第一请求之前，所述方法还包括：

向所述文件修复装置发送第二请求，所述第二请求用于获取所述更新后的漏洞修复程序的标识信息；

接收所述文件修复装置返回的所述更新后的漏洞修复程序的MD5；

根据所述更新后的漏洞修复程序的MD5，判断所述更新后的漏洞修复程序与请求装置已存储的漏洞修复程序是否相同，若相同，所述请求装置则发起所述第一请求。

12.一种文件修复装置，其特征在于，包括：

获取单元，用于获取请求装置上传的具有第一格式的可疑文件；

处理单元，用于按照所述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组；

第一确定单元，用于根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征；

第二确定单元，用于依据所述令牌序列特征，确定所述至少一个字段组中的漏洞；

修复单元，用于将所述漏洞替换为预先配置的修复代码序列，得到修复后的所述至少一个字段组；

重组单元，用于将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件。

13.根据权利要求12所述的装置，其特征在于，还包括：

提取单元，用于提取各个可疑文件的元信息，其中，所述元信息包括对应可疑文件的文件路径及消息摘要算法MD5，所述文件路径是指所述可疑文件在所述请求装置所在的用户设备上的路径；

去重单元，用于根据所述文件路径及所述MD5，对所述各个可疑文件进行去重处理。

14.根据权利要求13所述的装置，其特征在于，所述去重单元包括：

确定模块，用于确定所述各个可疑文件中重复的可疑文件，其中，所述重复的可疑文件是指所述文件路径相同且所述MD5相同的文件；

处理模块，用于保留所述重复的可疑文件中的第一文件，将除所述第一文件以外的所述重复的可疑文件删除，其中，所述第一文件为所述重复的可疑文件中的任意一个。

15.根据权利要求12所述的装置，其特征在于，所述第一确定单元用于执行以下步骤依据所述令牌序列特征，确定所述至少一个字段组中的漏洞：

在所述至少一个字段组中查找与所述令牌序列特征相同的字段，将所述与所述令牌序列特征相同的字段确定为所述漏洞。

16.根据权利要求12所述的装置，其特征在于，所述重组单元用于执行以下步骤将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件：

按照所述修复后的所述至少一个字段组中各个字段的父子关系，将所述修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件。

17.根据权利要求12所述的装置，其特征在于，还包括：

第一接收单元，用于接收所述请求装置发送的第一请求，所述第一请求用于下载更新后的漏洞修复程序；

第一发送单元，用于向所述请求装置发送所述更新后的漏洞修复程序，其中，所述更新后的漏洞修复程序中包含所述修复文件，以便所述请求装置根据所述修复文件对所述可疑文件进行修复。

18.根据权利要求17所述的装置，其特征在于，还包括：

第二接收单元，用于接收来自所述请求装置发送的第二请求，所述第二请求用于获取所述更新后的漏洞修复程序的标识信息；

第二发送单元，用于向所述请求装置发送所述更新后的漏洞修复程序的MD5，其中，由所述请求装置根据所述更新后的漏洞修复程序的MD5，判断所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序是否相同，若相同，所述请求装置则发起所述第一请求。

19.一种请求装置，其特征在于，包括：

第三发送单元，用于将具有第一格式的可疑文件发送至文件修复装置，其中，由所述文件修复装置按照所述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组，根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征，依据所述

令牌序列特征，确定所述至少一个字段组中的漏洞，将所述漏洞替换为预设的修复代码序列，得到修复后的所述至少一个字段组，将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件；

控制单元，用于从所述文件修复装置中获取对应于所述可疑文件的具有所述第一格式的修复文件。

20.根据权利要求19所述的装置，其特征在于，所述控制单元用于执行以下步骤从文件修复装置中获取对应于所述可疑文件的具有所述第一格式的修复文件：

向所述文件修复装置发送第一请求，所述第一请求用于下载更新后的漏洞修复程序；

接收所述文件修复装置返回的所述更新后的漏洞修复程序，其中，所述更新后的漏洞修复程序中包含所述修复文件。

21.根据权利要求20所述的装置，其特征在于，所述控制单元，还用于向所述文件修复装置发送第二请求，所述第二请求用于获取所述更新后的漏洞修复程序的标识信息；接收所述文件修复装置返回的所述更新后的漏洞修复程序的MD5；根据所述更新后的漏洞修复程序的MD5，判断所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序是否相同，若相同，所述请求装置则发起所述第一请求。

文件修复方法及装置

技术领域

[0001] 本申请涉及互联网领域，具体而言，涉及一种文件修复方法及装置。

背景技术

[0002] 随着互联网的快速发展，各个运营商的服务器运行着大量的web网站，出于技术上的考虑，这些web网站大多数直接采用目前市场上开源或者收费的CMS(Content Management System，内容管理系统)进行网站系统的搭建。而这些CMS由于源代码容易获得，常常遭到黑客的漏洞挖掘，在发现了可利用的漏洞之后，黑客会利用自己手上的ODAY(零日)漏洞对web网站进行攻击，进而获取大量的商业资料。

[0003] 在web攻防的对抗中，对于传统的二进制漏洞文件、或者是网站文本文件Patch(补丁)技术中，现有技术一般采用的是统一的替换模版的方法，即不管漏洞文件或漏洞代码有多少种变种，在服务器端都采用同一套Patch代码实行漏洞修复，这就导致漏洞修复准确性较差、系统安全性较低的问题。

[0004] 针对上述的问题，目前尚未提出有效的解决方案。

发明内容

[0005] 本申请实施例提供了一种文件修复方法及装置，以至少解决由于现有技术采用同一套Patch代码实行造成的漏洞修复准确性较差的技术问题。

[0006] 根据本申请实施例的一个方面，提供了一种文件修复方法，包括：获取请求装置上传的具有第一格式的可疑文件；按照上述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组；根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征；依据上述令牌序列特征，确定上述至少一个字段组中的漏洞；将上述漏洞替换为预设的修复代码序列，得到修复后的上述至少一个字段组；将修复后的上述至少一个字段组重组为对应于上述可疑文件的具有上述第一格式的修复文件。

[0007] 根据本申请实施例的另一方面，还提供了一种文件修复方法，包括：将具有第一格式的可疑文件发送至文件修复装置，其中，由上述文件修复装置按照上述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组，根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征，依据上述令牌序列特征，确定上述至少一个字段组中的漏洞，将上述漏洞替换为预设的修复代码序列，得到修复后的上述至少一个字段组，将修复后的上述至少一个字段组重组为对应于上述可疑文件的具有上述第一格式的修复文件；从上述文件修复装置中获取对应于上述可疑文件的具有上述第一格式的修复文件。

[0008] 根据本申请实施例的另一方面，还提供了一种文件修复装置，包括：获取单元，用于获取请求装置上传的具有第一格式的可疑文件；处理单元，用于按照上述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组；第一确定单元，用于根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征；第二确定单元，用于依据

上述令牌序列特征，确定上述至少一个字段组中的漏洞；修复单元，用于将上述漏洞替换为预先配置的修复代码序列，得到修复后的上述至少一个字段组；重组单元，用于将修复后的上述至少一个字段组重组为对应于上述可疑文件的具有上述第一格式的修复文件。

[0009] 根据本申请实施例的另一方面，还提供了一种请求装置，包括：第三发送单元，用于将具有第一格式的可疑文件发送至文件修复装置，其中，由上述文件修复装置按照上述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组，根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征，依据上述令牌序列特征，确定上述至少一个字段组中的漏洞，将上述漏洞替换为预设的修复代码序列，得到修复后的上述至少一个字段组，将修复后的上述至少一个字段组重组为对应于上述可疑文件的具有上述第一格式的修复文件；控制单元，用于从上述文件修复装置中获取对应于上述可疑文件的具有上述第一格式的修复文件。

[0010] 在本申请实施例中，采用获取请求装置上传的具有第一格式的可疑文件；按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组；根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征；依据令牌序列特征，确定至少一个字段组中的漏洞；将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组；将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件的方式，通过按照语法特征对可疑文件进行拆分得到至少一个字段组之后，依据各个语法特征对应令牌序列特征及修复代码序列进行漏洞的查找及修复，达到了对漏洞进行1:1的自动修复的目的，从而实现了提高漏洞修复的准确性及系统安全性的技术效果，进而解决了由于现有技术采用同一套Patch代码实行造成的漏洞修复准确性较差的技术问题。

附图说明

[0011] 此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

[0012] 图1是根据本申请实施例的一种运行文件检测方法的计算机终端的硬件结构框图；

[0013] 图2是根据本申请实施例的一种可选的文件修复方法的流程示意图；

[0014] 图3是根据本申请实施例的另一种可选的文件修复方法的流程示意图；

[0015] 图4是根据本申请实施例的又一种可选的文件修复方法的流程示意图；

[0016] 图5是根据本申请实施例的又一种可选的文件修复方法的流程示意图；

[0017] 图6是根据本申请实施例的又一种可选的文件修复方法的流程示意图；

[0018] 图7是根据本申请实施例的又一种可选的文件修复方法的流程示意图；

[0019] 图8是根据本申请实施例的一种可选的文件修复装置的结构示意图；

[0020] 图9是根据本申请实施例的另一种可选的文件修复装置的结构示意图；

[0021] 图10是根据本申请实施例的一种可选的去重单元的结构示意图；

[0022] 图11是根据本申请实施例的又一种可选的文件修复装置的结构示意图；

[0023] 图12是根据本申请实施例的又一种可选的文件修复装置的结构示意图；

[0024] 图13是根据本申请实施例的一种可选的请求装置的结构示意图。

具体实施方式

[0025] 为了使本技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分的实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本申请保护的范围。

[0026] 需要说明的是，本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，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、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0027] 实施例1

[0028] 根据本申请实施例，还提供了一种文件修复方法的方法实施例，需要说明的是，在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，虽然在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤。

[0029] 本申请实施例一所提供的方法实施例可以在移动终端、计算机终端或者类似的运算装置中执行。以运行在计算机终端上为例，图1是本申请实施例的一种文件修复方法的计算机终端的硬件结构框图。如图1所示，计算机终端10可以包括一个或多个(图中仅示出一个)处理器102(处理器102可以包括但不限于微处理器MCU或可编程逻辑器件FPGA等的处理装置)、用于存储数据的存储器104、以及用于通信功能的传输装置106。本领域普通技术人员可以理解，图1所示的结构仅为示意，其并不对上述电子装置的结构造成限定。例如，计算机终端10还可包括比图1中所示更多或者更少的组件，或者具有与图1所示不同的配置。

[0030] 存储器104可用于存储应用软件的软件程序以及模块，如本申请实施例中的文件修复方法对应的程序指令/模块，处理器102通过运行存储在存储器104内的软件程序以及模块，从而执行各种功能应用以及数据处理，即实现上述的应用程序的漏洞检测方法。存储器104可包括高速随机存储器，还可包括非易失性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器104可进一步包括相对于处理器102远程设置的存储器，这些远程存储器可以通过网络连接至计算机终端10。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0031] 传输装置106用于经由一个网络接收或者发送数据。上述的网络具体实例可包括计算机终端10的通信供应商提供的无线网络。在一个实例中，传输装置106包括一个网络适配器(Network Interface Controller，NIC)，其可通过基站与其他网络设备相连从而可与互联网进行通讯。在一个实例中，传输装置106可以为射频(Radio Frequency，RF)模块，其用于通过无线方式与互联网进行通讯。

[0032] 在上述运行环境下，本申请提供了如图2所示的文件修复方法。图2是根据本申请实施例一的文件修复方法的流程图。

[0033] 如图2所示，该文件修复方法可以包括如下实现步骤：

[0034] 步骤S202，获取请求装置上传的具有第一格式的可疑文件。

[0035] 本申请上述步骤S202中，请求装置可以是指安装在用户设备上的一个应用程序(例如客户端)，该请求装置可以根据服务器端的配置，定时启动对用户设备进行扫描，当发现可疑文件时，通过预先配置的通道上传该可疑文件至文件修复装置，请求装置也可以是计算机终端或者类似的运算装置，定时启动对自身磁盘文件进行扫描，当发现可疑文件时，通过预先配置的通道将可疑文件发送给文件修复装置。

[0036] 其中，文件修复装置接收到来自请求装置的具有第一格式的可疑文件之后，可以对该可疑文件进行解析，以实现可疑文件的修复，文件修复装置例如可以为AliYunDun服务器。

[0037] 以请求装置为AliYunDun(阿里云盾)为例，AliYunDun基于C/S(Client/Server，客户端/服务器)的网络架构，每台用户设备上都部署有一个独立的AliYunDun程序，并通过加密通道和AliYunDun服务器保持长连接。根据服务器的配置，AliYunDun定时从服务器下载最新AliVulfix(云盾附属组件)漏洞修复程序，替换本地的漏洞修复程序。根据服务器的配置，AliYunDun可以控制AliVulfix漏洞修复程序定时启动，对用户设备上的所有磁盘目录进行全盘遍历扫描，并根据服务器下发的配置信息收取指定目录的文件，例如配置项有一条为:/plus/mytag_js.php，则AliVulfix会扫描并通过AliYunDun的加密通过上报这个可疑文件。

[0038] 步骤S204，按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组。

[0039] 本申请上述步骤S204中，文件修复装置在获取该可疑文件之后，为了对漏洞采取一对一针对性修复并且提高修复效率，以免修复时间过长导致黑客利用该时间间隙入侵用户设备，文件修复装置可以按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组。

[0040] 以文件修复装置为AliYunDun服务器为例，因为阿里云上用户大部分是电商、网站用户，用户常常有对自己网站代码文件做定制化修改的情况，为了保证漏洞修复不影响这些用户网站的正常运行，AliYunDun服务器采取了一对一针对性修复的策略，然而，一对一修复会产生一个问题，采用1:1修复后，在阿里云数十万体量的用户数下，在服务器端会收到大量的存在漏洞的文件，即使进行过滤，文件数量还是很多，因此，AliYunDun服务器可以将上述可疑文件解析成语法树结构。

[0041] 本申请实施例中，文件修复装置首先要对可疑文件进行语法分析，确定可疑文件中各个字段的语法特征，其中，语法分析指的是将代码扫描到一个容器中，然后对该容器中的字符在词法分析的基础上将字段组合成各类语法短语。在确定了各个字段的语法特征之后，文件修复装置可以将语法特征相同的字段组成至少一个字段组，进一步地，文件修复装置可以依据各个字段之间的父子关系，将至少一个字段组解析为语法树结构。

[0042] 步骤S206，根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征。

[0043] 本申请上述步骤S206中，令牌序列特征可以是由安全运营人员预先配置的，文件修复装置利用对应于各个语法特征的令牌序列特征，实现漏洞的查找。

[0044] 步骤S208，依据令牌序列特征，确定至少一个字段组中的漏洞。

[0045] 本申请上述步骤S208中，文件修复装置在获取到各个语法特征对应的令牌序列特征之后，可以依据令牌序列特征在至少一个字段组中进行漏洞的定位。

[0046] 此处以通俗的例子来说明，假如有一百个人，目的是确定眼球颜色蓝色的人，即令牌序列特征是“眼球-蓝色”，那么按照本申请实施例提供的文件修复方法的思路，我们先将这一百个人的各个部分进行分组，例如，将一百个人的眼睛分为一组、鼻子分为一组、嘴巴分为一组等等，那么，在根据令牌序列特征进行漏洞查找时，只需要在眼睛这一组中进行查找，无需再搜索其他部分，达到提高漏洞查找效率的目的。

[0047] 可选地，依据令牌序列特征，确定至少一个字段组中的漏洞包括：在至少一个字段组中查找与令牌序列特征相同的字段，将与令牌序列特征相同的字段确定为漏洞。

[0048] 本申请实施例中，安全运营人员只需配置一次上述令牌序列特征，文件修复装置则可以自动化进行漏洞的定位，提高了安全运营的时间成本。

[0049] 步骤S210，将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组。

[0050] 本申请上述步骤S210中，预设的修复代码序列也可以是安全运营人员预先配置的。文件修复装置在依据令牌序列特征，确定至少一个字段组中的漏洞之后，可以获取该预设的修复代码序列，进而，将漏洞替换为该预设的修复代码序列，实现可疑文件的修复。

[0051] 本申请实施例的文件修复装置采用的1:1的精确定点修复技术，即所有的可疑文件都直接基于请求装置原始文件进行针对性的修改，这样可以最大程度地避免因为漏洞修复造成代码逻辑的不一致从而导致用户的正常业务逻辑受到影响，甚至不可用。同时，安全运营人员只需要配置一个修复代码序列，在文件修复装置确定漏洞时，将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组。

[0052] 步骤S212，将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件。

[0053] 本申请上述步骤S212中，由于上述文件修复过程中，文件修复装置将具有第一格式的可疑文件按照各个字段的语法特征拆分为了至少一个字段组，因此在得到修复后的至少一个字段组之后，文件修复装置可以按照其逆过程，将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件。

[0054] 可选地，将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件包括：按照修复后的至少一个字段组中各个字段的父子关系，将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件。

[0055] 以文件修复装置为AliYunDun服务器为例，AliYunDun服务器可以将上述可疑文件解析成语法树结构之后，依据token序列特征(令牌序列特征)定位漏洞，进而在语法树中插入了预设的修复代码序列之后，然后重建可疑文件，完成本次可疑文件的漏洞修复，得到上述的修复文件。

[0056] 本申请实施例提供的文件修复方法，由于可疑文件的数量可能成千上万，因此采用上述的文件修复方法整个过程都可以通过程序代码实现自动化，实现整个漏洞修复。从发现，到审核、到修复的全过程自动化，理论上可以实现阿里云全网的用户设备在20分钟内全部修复某一个漏洞。

[0057] 本申请实施例的文件修复方法至少具有以下技术效果：

[0058] 1、精确地针对每个用户的特定业务修复其中的漏洞；

[0059] 2、漏洞修复全过程自动化，安全运营人员只需要专注于漏洞的研究，并进行一次配置，之后的全网修复动作完全是全自动化；

[0060] 3、实现阿里云全网的机器在20分钟内全部修复某一个漏洞；

[0061] 4、采用语法树的方式进行漏洞的定位和修复，可以避免正则、字符串搜索带来的误报，具有很高的准确率和很低的误报率；

[0062] 5、不影响请求装置待修复程序的正常运行。

[0063] 由上可知，本申请上述实施例一所提供的方案，通过按照语法特征对可疑文件进行拆分得到至少一个字段组之后，依据各个语法特征对应令牌序列特征及修复代码序列进行漏洞的查找及修复，达到了对漏洞进行1:1的自动修复的目的，从而实现了提高漏洞修复的准确性及系统安全性的技术效果，进而解决了由于现有技术采用同一套Patch代码实行造成的漏洞修复准确性较差的技术问题。

[0064] 本申请上述实施例提供的一种可选方案中，如图3所示，在可疑文件的数量为至少两个的情况下，上述步骤S202获取请求装置上传的具有第一格式的可疑文件之后，上述步骤S204按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组之前，文件修复方法还可以包括：

[0065] 步骤S302，提取各个可疑文件的元信息，其中，元信息包括对应可疑文件的文件路径及消息摘要算法MD5，文件路径是指可疑文件在请求装置所在的用户设备上的路径。

[0066] 本申请上述步骤S302中，文件修复装置获取到的可疑文件的数量可能成千上万，那么其中可能存在一些重复的文件，因此，在可疑文件的数量为至少两个的情况下，文件修复装置可以对可疑文件进行去重处理。

[0067] 那么，在可疑文件的数量为至少两个的情况下，文件修复装置可以提取各个可疑文件的元信息，并将元信息保存到分布式的指纹数据库中。其中，元信息可以包括对应可疑文件的文件路径及MD5(Message Digest Algorithm，消息摘要算法)，文件路径是指可疑文件在请求装置所在的用户设备上的路径。

[0068] 步骤S304，根据文件路径及MD5，对各个可疑文件进行去重处理。

[0069] 本申请上述步骤S304中，去重处理是指去掉重复的可疑文件。文件修复装置在提取了各个可疑文件的文件路径和MD5之后，可以根据文件路径及MD5判断该各个可疑文件之中是否存在重复的文件，进而对各个可疑文件进行去重处理。

[0070] 本申请上述实施例提供的一种可选方案中，如图4所示，上述步骤S304，根据文件路径及MD5，对各个可疑文件进行去重处理可以包括：

[0071] 步骤S402，确定各个可疑文件中重复的可疑文件，其中，重复的可疑文件是指文件路径相同且MD5相同的文件。

[0072] 本申请上述步骤S402中，针对如何根据文件路径及MD5，对各个可疑文件进行去重处理，本申请实施例提供的方法为文件修复装置通过判断各个可疑文件的文件路径是否相同，以及各个可疑文件的MD5是否相同，若文件路径相同且MD5相同，则确定为重复的可疑文件。

[0073] 例如，文件修复装置在提取各个可疑文件的文件路径和MD5之后，判断出可疑文件1、可疑文件7、可疑文件24及可疑文件30的文件路径相同且MD5相同，那么文件修复装置确定该可疑文件1、可疑文件7、可疑文件24及可疑文件30为重复的可疑文件。

[0074] 步骤S404,保留重复的可疑文件中的第一文件,将除第一文件以外的重复的可疑文件删除,其中,第一文件为重复的可疑文件中的任意一个。

[0075] 本申请上述步骤S404中,文件修复装置在确定了各个可疑文件中重复的可疑文件之后,需要只保留其中一个可疑文件,删除其他重复的可疑文件以实现去重处理。

[0076] 例如,文件修复装置在确定了可疑文件1、可疑文件7、可疑文件24及可疑文件30为重复的可疑文件之后,可以保留可疑文件1、可疑文件7、可疑文件24及可疑文件30中的任意一个文件,例如保留可疑文件1,进而将除可疑文件1以外的重复的可疑文件删除,即删除可疑文件7、可疑文件24及可疑文件30,实现对可疑文件的去重处理。

[0077] 本申请上述实施例提供的一种可选方案中,如图5所示,上述步骤S212,将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件之后,文件修复方法还可以包括:

[0078] 步骤S502,接收请求装置发送的第一请求,第一请求用于下载更新后的漏洞修复程序。

[0079] 本申请上述步骤S502中,请求装置根据配置可以定时下载最新的漏洞修复重新,进而,文件修复装置可以接收到请求装置发送的用于下载更新后的漏洞修复程序的第一请求。

[0080] 以请求装置为AliYunDun、文件修复装置为AliYunDun服务器为例,AliYunDun基于C/S网络架构,通过加密通道和AliYunDun服务器保持长连接,根据AliYunDun服务器的配置,AliYunDun定时从AliYunDun服务器下载最新AliVulfix漏洞修复程序。

[0081] 步骤S504,向请求装置发送更新后的漏洞修复程序,其中,更新后的漏洞修复程序中包含修复文件,以便请求装置根据修复文件对可疑文件进行修复。

[0082] 本申请上述步骤S504中,文件修复装置在完成对可疑文件的修复之后,会得到一份最新的漏洞修复程序,例如,该漏洞修复程序的形式可以为:{"filename":"/webapps/manager/WEB-INF/web.xml","md5":"bc5ef661b746b0c55462353583a7eb34","action":"delete"};{"filename":"/webapps/manager/WEB-INF/web.xml","md5":"bc5ef661b746b0c55462353583a7eb55","action":"replace"}。这个最新的漏洞修复程序准确定位到上述的可疑文件的文件路径下,对应MD5的可疑文件,请求装置在下载了该更新后的漏洞修复程序之后,可以根据更新后的漏洞修复程序中包含的上述修复文件,对可疑文件进行修复,例如将可疑文件为上述修复文件。

[0083] 本申请上述实施例提供的一种可选方案中,如图6所示,上述步骤S502,接收请求装置发送的第一请求,第一请求用于下载更新后的漏洞修复程序之前,文件修复方法还可以包括:

[0084] 步骤S602,接收来自请求装置发送的第二请求,第二请求用于获取更新后的漏洞修复程序的标识信息。

[0085] 本申请上述步骤S602中,由于请求装置是根据配置定时下载漏洞修复程序,因此请求装置可以先判断该更新后的漏洞修复程序是否为最新的。

[0086] 仍旧以请求装置为AliYunDun、文件修复装置为AliYunDun服务器为例,根据AliYunDun服务器的配置,AliYunDun定时从AliYunDun服务器下载最新AliVulfix漏洞修复程序,判断是否为最新的标准为,检查保存在AliYunDun服务器的AliVulfix漏洞修复程序

的MD5是否和本机之前下载过的AliVulfix漏洞修复程序的MD5是否相同，如果不同，说明AliYunDun服务器进行了修改变动，则重新下载一次最新的AliVulfix漏洞修复程序，替换本地的AliVulfix漏洞修复程序。

[0087] 步骤S604，向请求装置发送更新后的漏洞修复程序的MD5，其中，由请求装置根据更新后的漏洞修复程序的MD5，判断更新后的漏洞修复程序与请求装置已存储的漏洞修复程序是否相同，若相同，请求装置则发起第一请求。

[0088] 可选地，请求装置根据更新后的漏洞修复程序的MD5，判断更新后的漏洞修复程序与请求装置已存储的漏洞修复程序是否相同可以包括：请求装置获取已存储的漏洞修复程序的MD5；请求装置判断更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5是否相同；若更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5相同，则确定更新后的漏洞修复程序与请求装置已存储的漏洞修复程序相同；若更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5不相同，则确定更新后的漏洞修复程序与请求装置已存储的漏洞修复程序不相同。

[0089] 在上述运行环境下，本申请还提供了一种文件修复方法。该文件修复方法可以包括如下实现步骤：

[0090] 步骤S10，将具有第一格式的可疑文件发送至文件修复装置，其中，由文件修复装置按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组，根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征，依据令牌序列特征，确定至少一个字段组中的漏洞，将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组，将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件。

[0091] 本申请上述步骤S10中，请求装置可以是指安装在用户设备上的一个应用程序(例如客户端)，该请求装置可以根据服务器端的配置，定时启动对用户设备进行扫描，当发现可疑文件时，通过预先配置的通道上传该可疑文件至文件修复装置，请求装置也可以是计算机终端或者类似的运算装置，定时启动对自身磁盘文件进行扫描，当发现可疑文件时，通过预先配置的通道将可疑文件发送给文件修复装置。

[0092] 步骤S12，从文件修复装置中获取对应于可疑文件的具有第一格式的修复文件。

[0093] 可选地，步骤S12从文件修复装置中获取对应于可疑文件的具有第一格式的修复文件包括：

[0094] 步骤S20，向文件修复装置发送第一请求，第一请求用于下载更新后的漏洞修复程序。

[0095] 本申请上述步骤S20中，请求装置根据配置可以定时下载最新的漏洞修复重新，进而，文件修复装置可以接收到请求装置发送的用于下载更新后的漏洞修复程序的第一请求。

[0096] 以请求装置为AliYunDun、文件修复装置为AliYunDun服务器为例，AliYunDun基于C/S网络架构，通过加密通道和AliYunDun服务器保持长连接，根据AliYunDun服务器的配置，AliYunDun定时从AliYunDun服务器下载最新AliVulfix漏洞修复程序。

[0097] 步骤S22，接收文件修复装置返回的更新后的漏洞修复程序，其中，更新后的漏洞修复程序中包含修复文件。

[0098] 本申请上述步骤S22中，文件修复装置在完成对可疑文件的修复之后，会得到一份

最新的漏洞修复程序，例如，该漏洞修复程序的形式可以为：{"filename":"/webapps/manager/WEB-INF/web.xml","md5":"bc5ef661b746b0c55462353583a7eb34","action":"delete"};{"filename":"/webapps/manager/WEB-INF/web.xml","md5":"bc5ef661b746b0c55462353583a7eb55","action":"replace"}。这个最新的漏洞修复程序准确定位到上述的可疑文件的文件路径下，对应MD5的可疑文件，请求装置在下载了该更新后的漏洞修复程序之后，可以根据更新后的漏洞修复程序中包含的上述修复文件，对可疑文件进行修复，例如将可疑文件为上述修复文件。

[0099] 可选地，在步骤S20，向文件修复装置发送第一请求之前，文件修复方法还包括：

[0100] 步骤S30，向文件修复装置发送第二请求，第二请求用于获取更新后的漏洞修复程序的标识信息。

[0101] 本申请上述步骤S30中，由于请求装置是根据配置定时下载漏洞修复程序，因此请求装置可以先判断该更新后的漏洞修复程序是否为最新的。

[0102] 仍旧以请求装置为AliYunDun、文件修复装置为AliYunDun服务器为例，根据AliYunDun服务器的配置，AliYunDun定时从AliYunDun服务器下载最新AliVulfix漏洞修复程序，判断是否为最新的标准为，检查保存在AliYunDun服务器的AliVulfix漏洞修复程序的MD5是否和本机之前下载过的AliVulfix漏洞修复程序的MD5是否相同，如果不同，说明AliYunDun服务器进行了修改变动，则重新下载一次最新的AliVulfix漏洞修复程序，替换本地的AliVulfix漏洞修复程序。

[0103] 步骤S32，接收文件修复装置返回的更新后的漏洞修复程序的MD5。

[0104] 步骤S34，根据更新后的漏洞修复程序的MD5，判断更新后的漏洞修复程序与请求装置已存储的漏洞修复程序是否相同，若相同，请求装置则发起第一请求。

[0105] 可选地，请求装置根据更新后的漏洞修复程序的MD5，判断更新后的漏洞修复程序与请求装置已存储的漏洞修复程序是否相同可以包括：请求装置获取已存储的漏洞修复程序的MD5；请求装置判断更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5是否相同；若更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5相同，则确定更新后的漏洞修复程序与请求装置已存储的漏洞修复程序相同；若更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5不相同，则确定更新后的漏洞修复程序与请求装置已存储的漏洞修复程序不相同。

[0106] 在一种可选的方案中，以AliYunDun为例，结合图7，对本申请上述文件修复方法进行描述：

[0107] 步骤A，AliYunDun基于C/S的网络架构，每台用户设备上都部署有一个独立的AliYunDun程序，并通过加密通道和AliYunDun服务器保持长连接。

[0108] 本申请上述步骤A中，AliYunDun相当于上述的请求装置，AliYunDun服务器相当于上述的文件修复装置，请求装置可以根据服务器端的配置，定时启动对用户设备进行扫描，当发现可疑文件时，通过预先配置的通道上传该可疑文件。

[0109] 步骤B，根据AliYunDun服务器的配置，AliYunDun定时从AliYunDun服务器下载最新的AliVulfix漏洞修复程序。

[0110] 本申请上述步骤B中，判断是否为最新的标准为，检查保存在AliYunDun服务器的AliVulfix漏洞修复程序的MD5是否和本机之前下载过的AliVulfix漏洞修复程序的MD5是

否相同，如果不同，说明AliYunDun服务器进行了修改变动，则重新下载一次最新的AliVulfix漏洞修复程序，替换本地的AliVulfix漏洞修复程序。

[0111] 由于请求装置是根据配置定时下载漏洞修复程序，因此请求装置可以先判断该更新后的漏洞修复程序是否为最新的，请求装置根据更新后的漏洞修复程序的MD5，判断更新后的漏洞修复程序与请求装置已存储的漏洞修复程序是否相同可以包括：请求装置获取已存储的漏洞修复程序的MD5；请求装置判断更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5是否相同；若更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5相同，则确定更新后的漏洞修复程序与请求装置已存储的漏洞修复程序相同；若更新后的漏洞修复程序的MD5与已存储的漏洞修复程序的MD5不相同，则确定更新后的漏洞修复程序与请求装置已存储的漏洞修复程序不相同。

[0112] 在确定更新后的漏洞修复程序与请求装置已存储的漏洞修复程序不相同时，文件修复装置可以接收到请求装置发送的用于下载更新后的漏洞修复程序的第一请求，并向请求装置发送更新后的漏洞修复程序，其中，更新后的漏洞修复程序中包含修复文件，以便请求装置根据修复文件对可疑文件进行修复。

[0113] 步骤C，根据AliYunDun服务器的配置，用户设备上的AliVulfix漏洞修复程序会定时启动执行。

[0114] 本申请上述步骤C中，AliYunDun可以根据服务器端的配置，定时启动对用户设备进行扫描，当发现可疑文件时，通过预先配置的通道上传该可疑文件。

[0115] 步骤D，AliVulfix对用户机器上的所有磁盘目录进行全盘遍历扫描。

[0116] 本申请上述步骤D中，根据AliYunDun服务器下发的配置信息收取指定目录的文件，例如配置项有一条为：/plus/mytag_js.php，则AliVulfix会扫描并通过AliYunDun的加密通过上报这个可疑文件。

[0117] 步骤E，AliYunDun服务器收到AliYunDun上报的可疑文件后，提取可疑文件的元信息。

[0118] 本申请上述步骤E中，文件修复装置获取到的可疑文件的数量可能成千上万，那么其中可能存在一些重复的文件，因此，在可疑文件的数量为至少两个的情况下，文件修复装置可以对可疑文件进行去重处理。那么，在可疑文件的数量为至少两个的情况下，文件修复装置可以提取各个可疑文件的元信息，并将元信息保存到分布式的指纹数据库中。其中，元信息可以包括对应可疑文件的文件路径及MD5，文件路径是指可疑文件在请求装置所在的用户设备上的路径。

[0119] 去重处理是指去掉重复的可疑文件。文件修复装置在提取了各个可疑文件的文件路径和MD5之后，可以根据文件路径及MD5判断该各个可疑文件之中是否存在重复的文件，进而对各个可疑文件进行去重处理。针对如何根据文件路径及MD5，对各个可疑文件进行去重处理，本申请实施例提供的方法为文件修复装置通过判断各个可疑文件的文件路径是否相同，以及各个可疑文件的MD5是否相同，若文件路径相同且MD5相同，则确定为重复的可疑文件。

[0120] 例如，文件修复装置在提取各个可疑文件的文件路径和MD5之后，判断出可疑文件1、可疑文件7、可疑文件24及可疑文件30的文件路径相同且MD5相同，那么文件修复装置确定该可疑文件1、可疑文件7、可疑文件24及可疑文件30为重复的可疑文件。

[0121] 可选地，文件修复装置在确定了各个可疑文件中重复的可疑文件之后，需要只保留其中一个可疑文件，删除其他重复的可疑文件以实现去重处理。

[0122] 例如，文件修复装置在确定了可疑文件1、可疑文件7、可疑文件24及可疑文件30为重复的可疑文件之后，可以保留可疑文件1、可疑文件7、可疑文件24及可疑文件30中的任意一个文件，例如保留可疑文件1，进而将除可疑文件1以外的重复的可疑文件删除，即删除可疑文件7、可疑文件24及可疑文件30，实现对可疑文件的去重处理。

[0123] 可选地，AliYunDun服务器在提取了可疑文件的元信息之后，可以将元信息保存到分布式的指纹数据库中。

[0124] 步骤F，把可疑文件解析成token语法树结构。

[0125] 本申请上述步骤F中，因为阿里云上用户大部分是电商、网站用户，用户常常有对自己网站代码文件做定制化修改的情况，为了保证漏洞修复不影响这些用户网站的正常运行，AliYunDun服务器采取了一对一针对性修复的策略，然而，一对一修复会产生一个问题，采用1:1修复后，在阿里云数十万体量的用户数下，在服务器端会收到大量的存在漏洞的文件，即使进行过滤，文件数量还是很多，因此，AliYunDun服务器可以将上述可疑文件解析成语法树结构。

[0126] 本申请实施例中，文件修复装置首先要对可疑文件进行语法分析，确定可疑文件中各个字段的语法特征，其中，语法分析指的是将代码扫描到一个容器中，然后对该容器中的字符在词法分析的基础上将字段组合成各类语法短语。在确定了各个字段的语法特征之后，文件修复装置可以将语法特征相同的字段组成至少一个字段组，进一步地，文件修复装置可以依据各个字段之间的父子关系，将至少一个字段组解析为语法树结构。

[0127] 步骤G，安全运营人员只需要配置一次token序列特征，用于定位文件中存在的漏洞。

[0128] 本申请上述步骤G中，token序列特征相当于上述的令牌序列特征，令牌序列特征可以是由安全运营人员预先配置的，文件修复装置利用对应于各个语法特征的令牌序列特征，实现漏洞的查找。

[0129] 文件修复装置在获取到各个语法特征对应的令牌序列特征之后，可以依据令牌序列特征在至少一个字段组中进行漏洞的定位。

[0130] 此处以通俗的例子来说明，假如有一百个人，目的是确定眼球颜色蓝色的人，即令牌序列特征是“眼球-蓝色”，那么按照本申请实施例提供的文件修复方法的思路，我们先将这一百个人的各个部分进行分组，例如，将一百个人的眼睛分为一组、鼻子分为一组、嘴巴分为一组等等，那么，在根据令牌序列特征进行漏洞查找时，只需要在眼睛这一组中进行查找，无需再搜索其他部分，达到提高漏洞查找效率的目的。

[0131] 可选地，依据令牌序列特征，确定至少一个字段组中的漏洞包括：在至少一个字段组中查找与令牌序列特征相同的字段，将与令牌序列特征相同的字段确定为漏洞。

[0132] 本申请实施例中，安全运营人员只需配置一次上述令牌序列特征，文件修复装置则可以自动化进行漏洞的定位，提高了安全运营的时间成本。

[0133] 步骤H，安全运营人员同时只需要配置一个patch token序列，用于在定位漏洞后插入patch token序列。

[0134] 本申请上述步骤H中，patch token序列相当于上述的预设的修复代码序列。预设

的修复代码序列也可以是安全运营人员预先配置的。文件修复装置在依据令牌序列特征，确定至少一个字段组中的漏洞之后，可以获取该预设的修复代码序列，进而，将漏洞替换为该预设的修复代码序列，实现可疑文件的修复。

[0135] 本申请实施例的文件修复装置采用的1:1的精确定点修复技术，即所有的可疑文件都直接基于请求装置原始文件进行针对性的修改，这样可以最大程度地避免因为漏洞修复造成代码逻辑的不一致从而导致用户的正常业务逻辑受到影响，甚至不可用。同时，安全运营人员只需要配置一个修复代码序列，在文件修复装置确定漏洞时，将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组。

[0136] 步骤I，在token语法树中插入了patch token序列之后，然后重建可疑文件，完成本次文件的漏洞修复。

[0137] 本申请上述步骤I中，AliYunDun服务器可以将上述可疑文件解析成语法树结构之后，依据token序列特征(令牌序列特征)定位漏洞，进而在语法树中插入了预设的修复代码序列之后，然后重建可疑文件，完成本次可疑文件的漏洞修复，得到上述的修复文件。

[0138] 由于上述文件修复过程中，文件修复装置将具有第一格式的可疑文件按照各个字段的语法特征拆分为了至少一个字段组，因此在得到修复后的至少一个字段组之后，文件修复装置可以按照其逆过程，将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件。

[0139] 可选地，将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件包括：按照修复后的至少一个字段组中各个字段的父子关系，将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件。

[0140] 步骤J，完成可疑文件审核之后，在AliYunDun服务器就会得到一份最新的漏洞修复程序。

[0141] 本申请上述步骤J中，文件修复装置在完成对可疑文件的修复之后，会得到一份最新的漏洞修复程序，例如，该漏洞修复程序的形式可以为：{"filename":"/webapps/manager/WEB-INF/web.xml","md5":"bc5ef661b746b0c55462353583a7eb34","action":"delete"};{"filename":"/webapps/manager/WEB-INF/web.xml","md5":"bc5ef661b746b0c55462353583a7eb55","action":"replace"}。这个最新的漏洞修复程序准确定位到上述的可疑文件的文件路径下，对应MD5的可疑文件，请求装置在下载了该更新后的漏洞修复程序之后，可以根据更新后的漏洞修复程序中包含的上述修复文件，对可疑文件进行修复，例如将可疑文件为上述修复文件。

[0142] 步骤K，AliVulfix的每次定时启动运行的时候，会从AliYunDun服务器拉取最新的漏洞规则库文件，并在全盘扫描的时候对本机上对应匹配到的文件进行修复操作(包括删除、替换操作)。

[0143] 步骤L，完成对阿里云用户网站脚本的漏洞修复。

[0144] 本申请上述步骤L中，AliYunDun服务器从AliYunDun收取用户设备上的真实CMS漏洞文件(即可疑文件)，并进行一对一的针对性修复，最大限度的保证了修复后和修复前的业务兼容性，不至于发生因为漏洞修复导致网站业务不可用；在AliYunDun服务器采取了自动化审核、以及基于Token语法树自动文件修复方式，极大提高的处理速度；AliYunDun基于文件路径、MD5定位可疑文件，避免因为出现误操作，对用户的网站造成影响。

[0145] 本申请实施例的文件修复方法至少具有以下技术效果：

[0146] 1、精确地针对每个用户的特定业务修复其中的漏洞；

[0147] 2、漏洞修复全过程自动化，安全运营人员只需要专注于漏洞的研究，并进行一次配置，之后的全网修复动作完全是全自动化；

[0148] 3、实现阿里云全网的机器在20分钟内全部修复某一个漏洞；

[0149] 4、采用语法树的方式进行漏洞的定位和修复，可以避免正则、字符串搜索带来的误报，具有很高的准确率和很低的误报率。

[0150] 由上可知，本申请上述实施例一所提供的方案，通过按照语法特征对可疑文件进行拆分得到至少一个字段组之后，依据各个语法特征对应令牌序列特征及修复代码序列进行漏洞的查找及修复，达到了对漏洞进行1:1的自动修复的目的，从而实现了提高漏洞修复的准确性及系统安全性的技术效果，进而解决了由于现有技术采用同一套Patch代码实行造成的漏洞修复准确性较差的技术问题。

[0151] 由此可知，现有技术存在的采用同一套Patch代码实行造成的漏洞修复准确性较差的问题，本申请提出一种基于语法特征对可疑文件的字段进行拆分的方法，进而依据各个语法特征对应令牌序列特征及修复代码序列进行漏洞的查找及修复，达到了对漏洞进行1:1的自动修复的目的，从而实现了提高漏洞修复的准确性及系统安全性的技术效果。

[0152] 需要说明的是，对于前述的各方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请并不受所描述的动作顺序的限制，因为依据本申请，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定是本申请所必须的。

[0153] 通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的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(如ROM/RAM、磁碟、光盘)中，包括若干指令用以使得一台终端设备(可以是手机，计算机，服务器，或者网络设备等)执行本申请各个实施例所述的方法。

[0154] 实施例2

[0155] 根据本申请实施例，还提供了一种用于实施上述方法实施例的装置实施例，如图8所示，该装置包括：

[0156] 图8是根据本申请实施例的文件修复装置的结构示意图。

[0157] 如图8所示，该文件修复装置可以包括获取单元802、处理单元804、第一确定单元806、第二确定单元808、修复单元810以及重组单元812。

[0158] 其中，获取单元802，用于获取请求装置上传的具有第一格式的可疑文件；处理单元804，用于按照所述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组；第一确定单元806，用于根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征；第二确定单元808，用于依据所述令牌序列特征，确定所述至少一个字段组中的漏洞；修复单元810，用于将所述漏洞替换为预先配置的修复代码序列，得到修复后的所述至少一个字段组；重组单元812，用于将修复后的所述至少一个字段组重组为对应于所

述可疑文件的具有所述第一格式的修复文件。

[0159] 由上可知，本申请上述实施例一所提供的方案，通过按照语法特征对可疑文件进行拆分得到至少一个字段组之后，依据各个语法特征对应令牌序列特征及修复代码序列进行漏洞的查找及修复，达到了对漏洞进行1:1的自动修复的目的，从而实现了提高漏洞修复的准确性及系统安全性的技术效果，进而解决了由于现有技术采用同一套Patch代码实行造成的漏洞修复准确性较差的技术问题。

[0160] 此处需要说明的是，上述获取单元802、处理单元804、第一确定单元806、第二确定单元808、修复单元810以及重组单元812对应于实施例一中的步骤S202至步骤S212，六个模块与对应的步骤所实现的示例和应用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中，可以通过软件实现，也可以通过硬件实现。

[0161] 可选地，如图9所示，文件修复装置还可以包括：提取单元902和去重单元904。

[0162] 其中，提取单元902，用于提取各个可疑文件的元信息，其中，所述元信息包括对应可疑文件的文件路径及消息摘要算法MD5，所述文件路径是指所述可疑文件在所述请求装置所在的用户设备上的路径；去重单元904，用于根据所述文件路径及所述MD5，对所述各个可疑文件进行去重处理。

[0163] 此处需要说明的是，上述提取单元902和去重单元904对应于实施例一中的步骤S302至步骤S304，该模块与对应的步骤所实现的示例和应用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中，可以通过软件实现，也可以通过硬件实现。

[0164] 可选地，如图10所示，去重单元904可以包括：确定模块1002和处理模块1004。

[0165] 其中，确定模块1002，用于确定所述各个可疑文件中重复的可疑文件，其中，所述重复的可疑文件是指所述文件路径相同且所述MD5相同的文件；处理模块1004，用于保留所述重复的可疑文件中的第一文件，将除所述第一文件以外的所述重复的可疑文件删除，其中，所述第一文件为所述重复的可疑文件中的任意一个。

[0166] 此处需要说明的是，上述确定模块1002和处理模块1004对应于实施例一中的步骤S402至步骤S404，该模块与对应的步骤所实现的示例和应用场景相同，但不限于上述实施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中，可以通过软件实现，也可以通过硬件实现。

[0167] 可选地，所述第一确定单元806用于执行以下步骤依据所述令牌序列特征，确定所述至少一个字段组中的漏洞：在所述至少一个字段组中查找与所述令牌序列特征相同的字段，将所述与所述令牌序列特征相同的字段确定为所述漏洞。

[0168] 可选地，所述重组单元812用于执行以下步骤将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件：按照所述修复后的所述至少一个字段组中各个字段的父子关系，将所述修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件。

[0169] 可选地，如图11所示，文件修复装置还可以包括：第一接收单元1102和第一发送单元1104。

[0170] 其中，第一接收单元1102，用于接收所述请求装置发送的第一请求，所述第一请求

用于下载更新后的漏洞修复程序;第一发送单元1104,用于向所述请求装置发送所述更新后的漏洞修复程序,其中,所述更新后的漏洞修复程序中包含所述修复文件,以便所述请求装置根据所述修复文件对所述可疑文件进行修复。

[0171] 此处需要说明的是,上述第一接收单元1102和第一发送单元1104对应于实施例一中的步骤S502至步骤S504,该模块与对应的步骤所实现的示例和应用场景相同,但不限于上述实施例一所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中,可以通过软件实现,也可以通过硬件实现。

[0172] 可选地,如图12所示,文件修复装置还可以包括:第二接收单元1202和第二发送单元1204。

[0173] 其中,第二接收单元1202,用于接收来自所述请求装置发送的第二请求,所述第二请求用于获取所述更新后的漏洞修复程序的标识信息;第二发送单元1204,用于向所述请求装置发送所述更新后的漏洞修复程序的MD5,其中,由所述请求装置根据所述更新后的漏洞修复程序的MD5,判断所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序是否相同,若相同,所述请求装置则发起所述第一请求。

[0174] 此处需要说明的是,上述第二接收单元1202和第二发送单元1204对应于实施例一中的步骤S602至步骤S604,该模块与对应的步骤所实现的示例和应用场景相同,但不限于上述实施例一所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中,可以通过软件实现,也可以通过硬件实现。

[0175] 可选地,所述请求装置根据所述更新后的漏洞修复程序的MD5,判断所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序是否相同可以包括:所述请求装置获取所述已存储的漏洞修复程序的MD5;所述请求装置判断所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5是否相同;若所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5相同,则确定所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序相同;若所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5不相同,则确定所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序不相同。

[0176] 根据本申请实施例,还提供了一种用于实施上述方法实施例的装置实施例,如图13所示,该装置包括:

[0177] 图13是根据本申请实施例的请求装置的结构示意图。

[0178] 如图13所示,该请求装置可以包括第三发送单元1302和控制单元1304。

[0179] 其中,第三发送单元1302,用于将具有第一格式的可疑文件发送至文件修复装置,其中,由所述文件修复装置按照所述可疑文件中各个字段的语法特征,将语法特征相同的字段组成至少一个字段组,根据每个字段组的语法特征,获取各个语法特征对应的令牌序列特征,依据所述令牌序列特征,确定所述至少一个字段组中的漏洞,将所述漏洞替换为预设的修复代码序列,得到修复后的所述至少一个字段组,将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件;控制单元1304,用于从所述文件修复装置中获取对应于所述可疑文件的具有所述第一格式的修复文件。

[0180] 此处需要说明的是,上述第三发送单元1302和控制单元1304对应于实施例一中的步骤S10至步骤S12,该模块与对应的步骤所实现的示例和应用场景相同,但不限于上述实

施例一所公开的内容。需要说明的是，上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中，可以通过软件实现，也可以通过硬件实现。

[0181] 可选地，所述控制单元1304用于执行以下步骤从文件修复装置中获取对应于所述可疑文件的具有所述第一格式的修复文件：向所述文件修复装置发送第一请求，所述第一请求用于下载更新后的漏洞修复程序；接收所述文件修复装置返回的所述更新后的漏洞修复程序，其中，所述更新后的漏洞修复程序中包含所述修复文件。

[0182] 可选地，所述控制单元1304，还用于向所述文件修复装置发送第二请求，所述第二请求用于获取所述更新后的漏洞修复程序的标识信息；接收所述文件修复装置返回的所述更新后的漏洞修复程序的MD5；根据所述更新后的漏洞修复程序的MD5，判断所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序是否相同，若相同，所述请求装置则发起所述第一请求。

[0183] 由此可知，现有技术存在的采用同一套Patch代码实行造成的漏洞修复准确性较差的问题，本申请提出一种基于语法特征对可疑文件的字段进行拆分的方法，进而依据各个语法特征对应令牌序列特征及修复代码序列进行漏洞的查找及修复，达到了对漏洞进行1:1的自动修复的目的，从而实现了提高漏洞修复的准确性及系统安全性的技术效果。

[0184] 实施例3

[0185] 本申请的实施例还提供了一种存储介质。可选地，在本实施例中，上述存储介质可以用于保存上述实施例一所提供的文件修复方法所执行的程序代码。

[0186] 可选地，在本实施例中，上述存储介质可以位于计算机网络中计算机终端群中的任意一个计算机终端中，或者位于移动终端群中的任意一个移动终端中。

[0187] 可选地，在本实施例中，存储介质被设置为存储用于执行以下步骤的程序代码：获取请求装置上传的具有第一格式的可疑文件；按照所述可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组；根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征；依据所述令牌序列特征，确定所述至少一个字段组中的漏洞；将所述漏洞替换为预设的修复代码序列，得到修复后的所述至少一个字段组；将修复后的所述至少一个字段组重组为对应于所述可疑文件的具有所述第一格式的修复文件。

[0188] 可选地，存储介质还被设置为存储用于执行以下步骤的程序代码：提取各个可疑文件的元信息，其中，所述元信息包括对应可疑文件的文件路径及消息摘要算法MD5，所述文件路径是指所述可疑文件在所述请求装置所在的用户设备上的路径；根据所述文件路径及所述MD5，对所述各个可疑文件进行去重处理。

[0189] 可选地，存储介质还被设置为存储用于执行以下步骤的程序代码：确定所述各个可疑文件中重复的可疑文件，其中，所述重复的可疑文件是指所述文件路径相同且所述MD5相同的文件；保留所述重复的可疑文件中的第一文件，将除所述第一文件以外的所述重复的可疑文件删除，其中，所述第一文件为所述重复的可疑文件中的任意一个。

[0190] 可选地，存储介质还被设置为存储用于执行以下步骤的程序代码：在所述至少一个字段组中查找与所述令牌序列特征相同的字段，将所述与所述令牌序列特征相同的字段确定为所述漏洞。

[0191] 可选地，存储介质还被设置为存储用于执行以下步骤的程序代码：按照所述修复后的所述至少一个字段组中各个字段的父子关系，将所述修复后的所述至少一个字段组重

组为对应于所述可疑文件的具有所述第一格式的修复文件。

[0192] 可选地，存储介质还被设置为存储用于执行以下步骤的程序代码：接收所述请求装置发送的第一请求，所述第一请求用于下载更新后的漏洞修复程序；向所述请求装置发送所述更新后的漏洞修复程序，其中，所述更新后的漏洞修复程序中包含所述修复文件，以便所述请求装置根据所述修复文件对所述可疑文件进行修复。

[0193] 可选地，存储介质还被设置为存储用于执行以下步骤的程序代码：接收来自所述请求装置发送的第二请求，所述第二请求用于获取所述更新后的漏洞修复程序的标识信息；向所述请求装置发送所述更新后的漏洞修复程序的MD5，其中，由所述请求装置根据所述更新后的漏洞修复程序的MD5，判断所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序是否相同，若相同，所述请求装置则发起所述第一请求。

[0194] 可选地，存储介质还被设置为存储用于执行以下步骤的程序代码：所述请求装置获取所述已存储的漏洞修复程序的MD5；所述请求装置判断所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5是否相同；若所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5相同，则确定所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序相同；若所述更新后的漏洞修复程序的MD5与所述已存储的漏洞修复程序的MD5不相同，则确定所述更新后的漏洞修复程序与所述请求装置已存储的漏洞修复程序不相同。

[0195] 上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。

[0196] 在本申请的上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述的部分，可以参见其他实施例的相关描述。

[0197] 在本申请所提供的几个实施例中，应该理解到，所揭露的订单信息的处理装置，可通过其它的方式实现。其中，以上所描述的装置实施例仅仅是示意性的，例如所述单元的划分，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。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，单元或模块的间接耦合或通信连接，可以是电性或其它的形式。

[0198] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0199] 另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

[0200] 所述集成的单元如果以软件功能单元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，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者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(可为个人计算机、服务器或者网络设备等)执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U盘、只读存储器(ROM，Read-Only Memory)、随机存取存

储器(RAM,Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0201] 以上所述仅是本申请的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本申请原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本申请的保护范围。

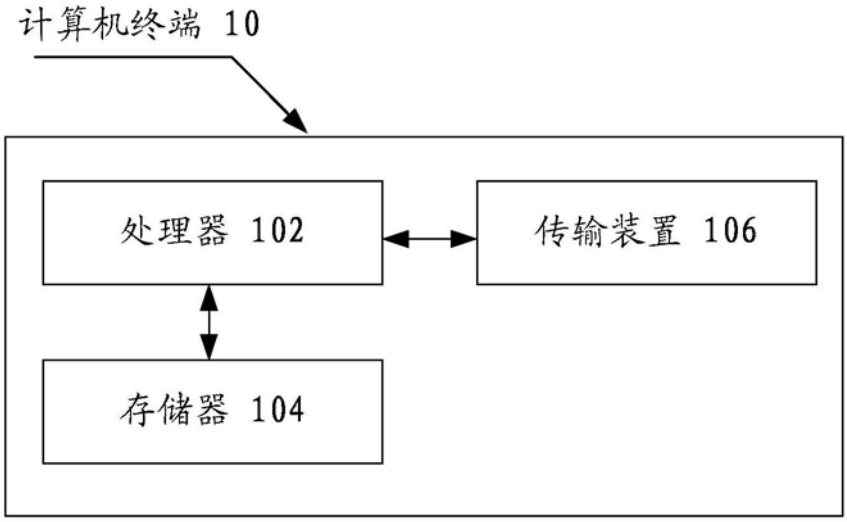


图1

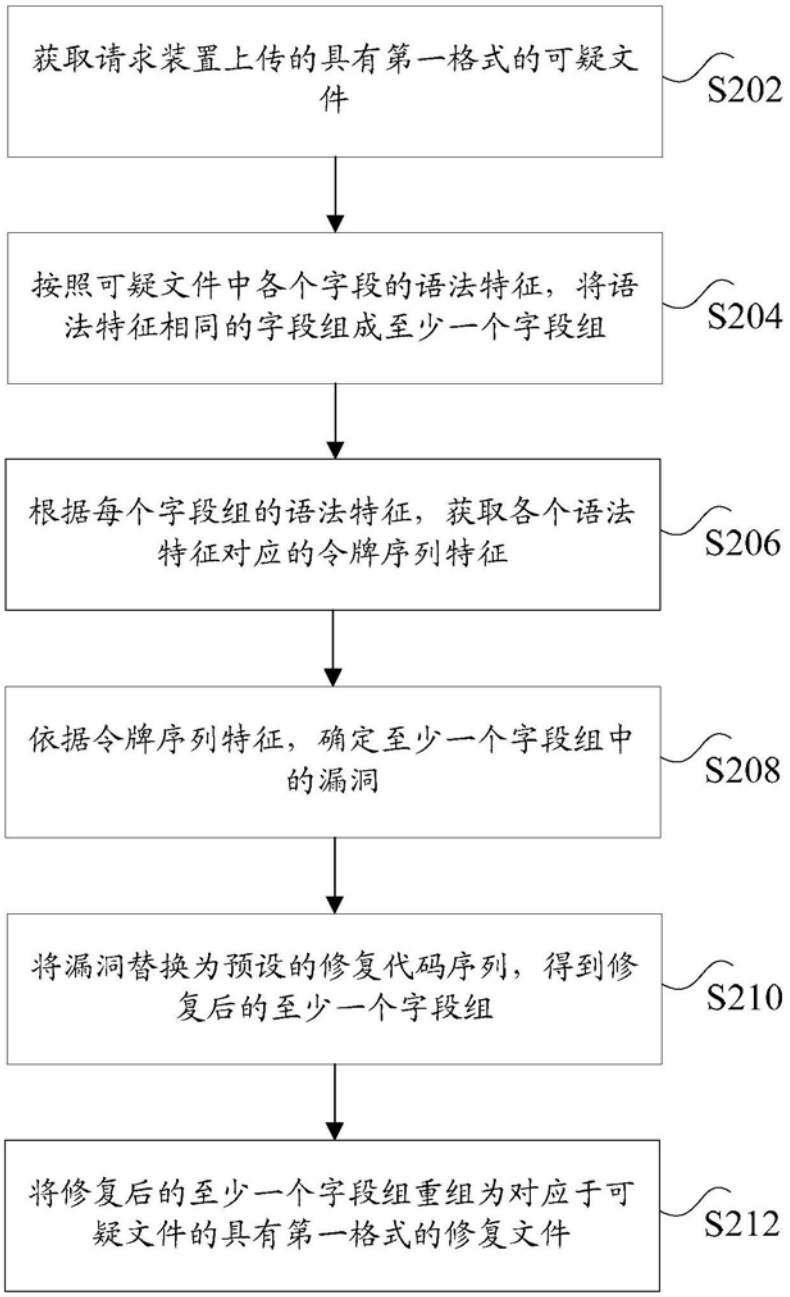

获取请求装置上传的具有第一格式的可疑文件
S202
按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组
S204
根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征
S206
依据令牌序列特征，确定至少一个字段组中的漏洞
S208
将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组
S210
将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件
S212

图2

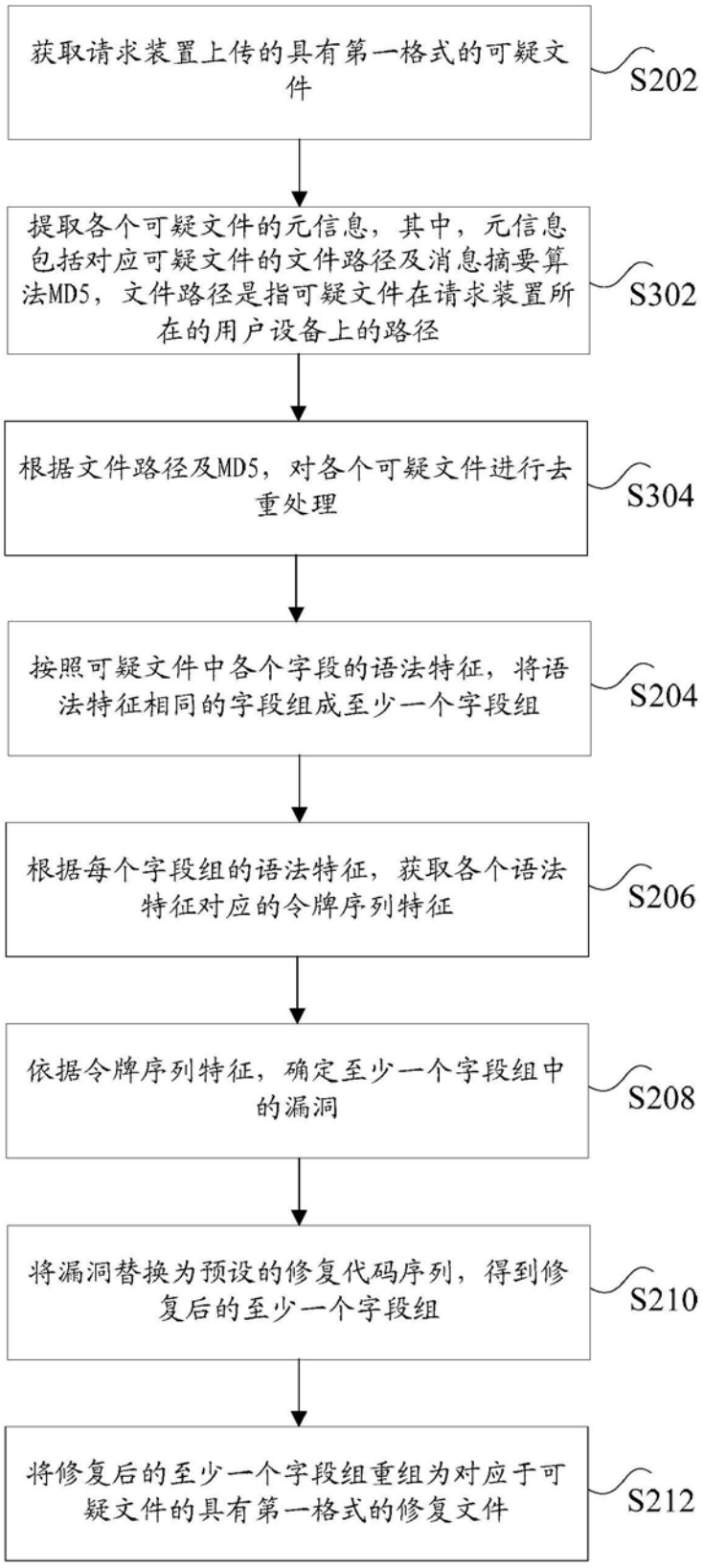

获取请求装置上传的具有第一格式的可疑文件
S202
提取各个可疑文件的元信息，其中，元信息包括对应可疑文件的文件路径及消息摘要算法MD5，文件路径是指可疑文件在请求装置所在的用户设备上的路径
S302
根据文件路径及MD5，对各个可疑文件进行去重处理
S304
按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组
S204
根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征
S206
依据令牌序列特征，确定至少一个字段组中的漏洞
S208
将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组
S210
将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件
S212

图3

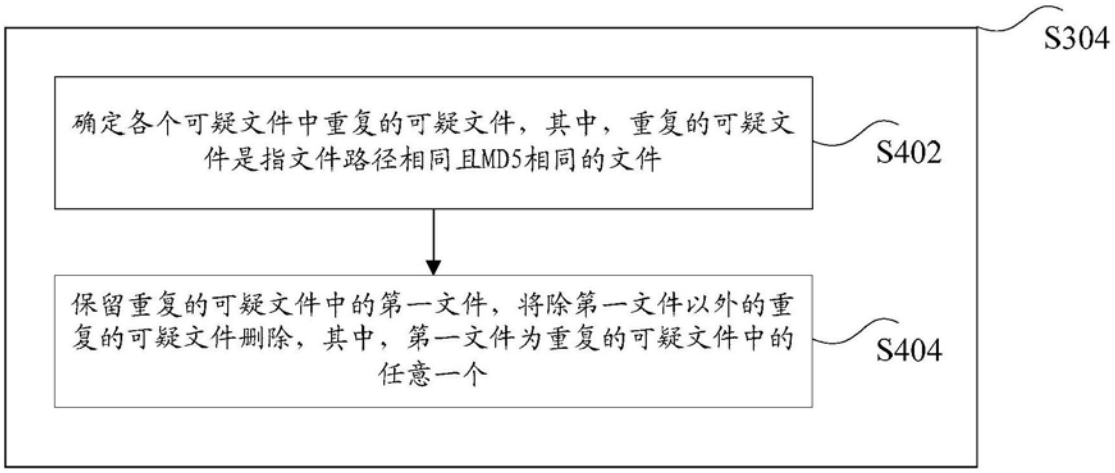

S304
确定各个可疑文件中重复的可疑文件，其中，重复的可疑文件是指文件路径相同且MD5相同的文件
S402
保留重复的可疑文件中的第一文件，将除第一文件以外的重复的可疑文件删除，其中，第一文件为重复的可疑文件中的任意一个
S404

图4

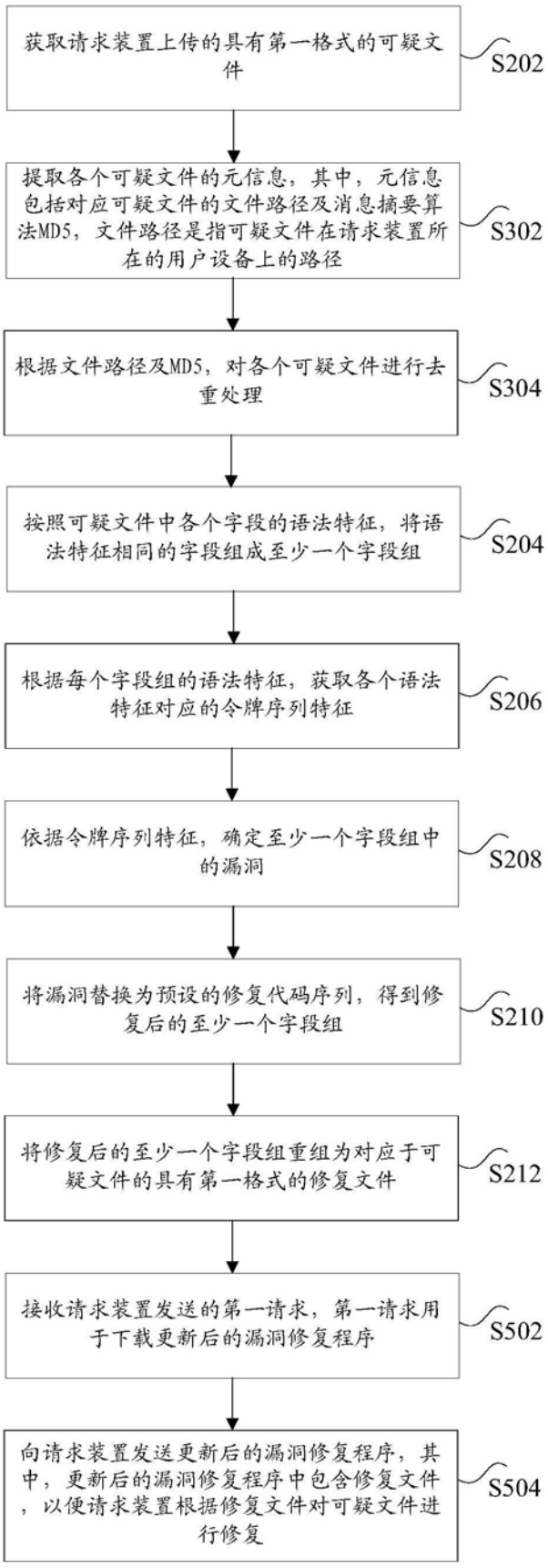

获取请求装置上传的具有第一格式的可疑文件
S202
提取各个可疑文件的元信息，其中，元信息包括对应可疑文件的文件路径及消息摘要算法MD5，文件路径是指可疑文件在请求装置所在的用户设备上的路径
S302
根据文件路径及MD5，对各个可疑文件进行去重处理
S304
按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组
S204
根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征
S206
依据令牌序列特征，确定至少一个字段组中的漏洞
S208
将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组
S210
将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件
S212
接收请求装置发送的第一请求，第一请求用于下载更新后的漏洞修复程序
S502
向请求装置发送更新后的漏洞修复程序，其中，更新后的漏洞修复程序中包含修复文件，以便请求装置根据修复文件对可疑文件进行修复
S504

图5


获取请求装置上传的具有第一格式的可疑文件
S202
提取各个可疑文件的元信息，其中，元信息包括对应可疑文件的文件路径及消息摘要算法MD5，文件路径是指可疑文件在请求装置所在的用户设备上的路径
S302
根据文件路径及MD5，对各个可疑文件进行去重处理
S304
按照可疑文件中各个字段的语法特征，将语法特征相同的字段组成至少一个字段组
S204
根据每个字段组的语法特征，获取各个语法特征对应的令牌序列特征
S206
依据令牌序列特征，确定至少一个字段组中的漏洞
S208
将漏洞替换为预设的修复代码序列，得到修复后的至少一个字段组
S210
将修复后的至少一个字段组重组为对应于可疑文件的具有第一格式的修复文件
S212
接收来自请求装置发送的第二请求，第二请求用于获取更新后的漏洞修复程序的标识信息
S602
向请求装置发送更新后的漏洞修复程序的MD5，其中，由请求装置根据更新后的漏洞修复程序的MD5，判断更新后的漏洞修复程序与请求装置已存储的漏洞修复程序是否相同，若相同，请求装置则发起第一请求
S604
接收请求装置发送的第一请求，第一请求用于下载更新后的漏洞修复程序
S502
向请求装置发送更新后的漏洞修复程序，其中，更新后的漏洞修复程序中包含修复文件，以便请求装置根据修复文件对可疑文件进行修复
S504

图6


AliYunDun
AliYunDun 服务器
步骤A，AliYunDun基于C / S的网络架构，每台用户设备上都部署有一个独立的AliYunDun程序，并通过加密通道和AliYunDun服务器保持长连接
步骤B，根据AliYunDun服务器的配置，AliYunDun定时从AliYunDun服务器下载最新的AliVulfix漏洞修复程序
步骤C，根据AliYunDun服务器的配置，用户设备上的AliVulfix漏洞修复程序会定时启动执行
步骤D，AliVulfix对用户机器上的所有磁盘目录进行全盘遍历扫描
步骤E，AliYunDun服务器收到AliYunDun上报的可疑文件后，提取可疑文件的元信息
步骤F，把可疑文件解析成token语法树结构
步骤G，安全运营人员只需要配置一次token序列特征，用于定位文件中存在的漏洞
步骤H，安全运营人员同时只需要配置一个patch token序列，用于在定位漏洞后插入patch token序列
步骤I，在token语法树中插入了patchtoken序列之后，然后重建可疑文件，完成本次文件的漏洞修复
步骤J，完成可疑文件审核之后，在AliYunDun服务器就会得到一份最新的漏洞修复程序
步骤K，AliVulfix的每次定时启动运行的时候，会从AliYunDun服务器拉取最新的漏洞规则库文件，并在全盘扫描的时候对本机上对应匹配到的文件进行修复操作(包括删除、替换操作)
步骤L，完成对阿里云用户网站脚本的漏洞修复

图7


获取单元802
处理单元804
第一确定单元806
第二确定单元808
修复单元810
重组单元812

图8

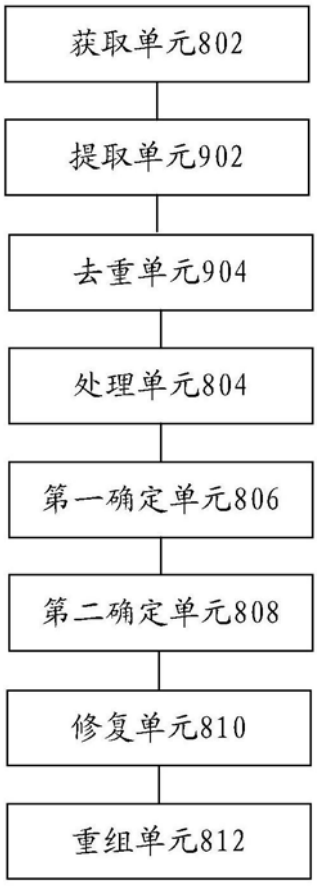

获取单元802
提取单元902
去重单元904
处理单元804
第一确定单元806
第二确定单元808
修复单元810
重组单元812

图9

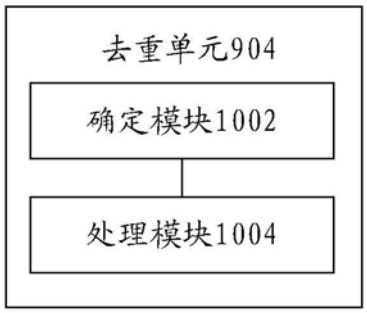

去重单元904
确定模块1002
处理模块1004

图10


获取单元802
处理单元804
第一确定单元806
第二确定单元808
修复单元810
重组单元812
第一接收单元1102
第一发送单元1104

图11


获取单元802
处理单元804
第一确定单元806
第二确定单元808
修复单元810
重组单元812
第二接收单元1202
第二发送单元1204
第一接收单元1102
第一发送单元1104

图12

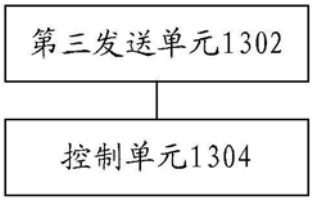

第三发送单元1302
控制单元1304

图13