



(12) 发明专利申请

(10) 申请公布号 CN 102622296 A

(43) 申请公布日 2012. 08. 01

(21) 申请号 201210041734. 5

(22) 申请日 2012. 02. 21

(71) 申请人 百度在线网络技术(北京)有限公司  
地址 100085 北京市海淀区上地十街 10 号  
百度大厦三层

(72) 发明人 李卓 金敬亭 肖羽思 赖峥嵘

(74) 专利代理机构 北京清亦华知识产权代理事  
务所(普通合伙) 11201  
代理人 张大威

(51) Int. Cl.

G06F 11/36(2006. 01)

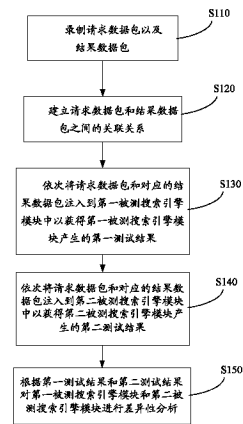
权利要求书 3 页 说明书 10 页 附图 7 页

(54) 发明名称

搜索引擎模块的测试方法、系统及其装置

(57) 摘要

本发明提出一种搜索引擎模块的测试方法,包括:录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包,同时录制在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包;建立请求数据包和结果数据包之间的关联关系;依次将请求数据包和对应的结果数据包注入到第一被测搜索引擎模块中以获得第一测试结果;依次将请求数据包和对应的结果数据包注入到第二被测搜索引擎模块中以获得第二测试结果;根据第一测试结果和第二测试结果对第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。本发明还公开了一种搜索引擎模块的测试系统、一种在线录制装置以及一种测试回放装置。



1. 一种搜索引擎模块的测试方法,其特征在于,包括:

录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包,同时录制在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包;

建立所述请求数据包和所述结果数据包之间的关联关系;

依次将所述请求数据包和对应的所述结果数据包注入到所述第一被测搜索引擎模块中以获得所述第一被测搜索引擎模块产生的第一测试结果;

依次将所述请求数据包和对应的所述结果数据包注入到所述第二被测搜索引擎模块中以获得所述第二被测搜索引擎模块产生的第二测试结果;以及

根据所述第一测试结果和所述第二测试结果对所述第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。

2. 如权利要求 1 所述的搜索引擎模块的测试方法,其特征在于,所述录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包进一步包括:

接收所述在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包;

获取所述请求数据包之中的查询 ID,并将所述请求数据包转发至所述在线搜索引擎模块;以及

根据所述查询 ID 存储所述请求数据包。

3. 如权利要求 1 或 2 所述的搜索引擎模块的测试方法,其特征在于,所述录制在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包进一步包括:

接收所述在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包;

获取所述结果数据包之中的查询 ID,并将所述结果数据包转发至所述在线搜索引擎模块;以及

根据所述查询 ID 存储所述结果数据包。

4. 如权利要求 3 所述的搜索引擎模块的测试方法,其特征在于,依据所述查询 ID 建立所述请求数据包和所述结果数据包之间的关联关系。

5. 如权利要求 1-4 任一项所述的搜索引擎模块的测试方法,其特征在于,所述在线下游搜索引擎模块为多个,以构成在线下游搜索引擎集群。

6. 如权利要求 4 所述的搜索引擎模块的测试方法,其特征在于,所述依次将请求数据包和对应的所述结果数据包注入到所述第一被测搜索引擎模块进一步包括:

将所述请求数据包注入到所述第一被测搜索引擎模块;

根据所述查询 ID 确定与所述请求数据包对应的结果数据包;以及

将所述结果数据包注入到所述第一被测搜索引擎模块。

7. 如权利要求 4 所述的搜索引擎模块的测试方法,其特征在于,所述依次将请求数据包和对应的所述结果数据包注入到所述第二被测搜索引擎模块进一步包括:

将所述请求数据包注入到所述第二被测搜索引擎模块;

根据所述查询 ID 确定与所述请求数据包对应的结果数据包;以及

将所述结果数据包注入到所述第二被测搜索引擎模块。

8. 如权利要求 6 或 7 所述的搜索引擎模块的测试方法,其特征在于,还包括:

如果根据所述查询 ID 未确定与所述请求数据包对应的结果数据包,则返回预设的结果数据包。

9. 一种搜索引擎模块的测试系统,其特征在于,包括:

在线录制装置,用于录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包和在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包;以及

测试回放装置,用于建立所述请求数据包和所述结果数据包之间的关联关系,并依次将所述请求数据包和对应的所述结果数据包注入到所述第一被测搜索引擎模块中以获得所述第一被测搜索引擎模块产生的第一测试结果,以及依次将所述请求数据包和对应的所述结果数据包注入到所述第二被测搜索引擎模块中以获得所述第二被测搜索引擎模块产生的第二测试结果,根据所述第一测试结果和所述第二测试结果对所述第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。

10. 如权利要求 9 所述的搜索引擎模块的测试系统,其特征在于,所述在线录制装置包括:

第一录制桩,所述第一录制桩连接在在线上游搜索引擎模块和在线搜索引擎模块之间,用于录制所述在线上游搜索引擎模块向所述在线搜索引擎模块发送的请求数据包,并将所述请求数据包转发至所述在线搜索引擎模块;

第二录制桩,所述第二录制装置连接在所述在线搜索引擎模块和在线下游搜索引擎模块之间,用于录制所述在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包,并将所述结果数据包转发至所述在线搜索引擎模块。

11. 如权利要求 10 所述的搜索引擎模块的测试系统,其特征在于,所述第一录制桩包括:

第一接收单元,用于接收所述在线上游搜索引擎模块向所述在线搜索引擎模块发送的请求数据包;

第一转发单元,用于获取所述请求数据包中的查询 ID,并将所述请求数据包转发至所述在线搜索引擎模块;

第一存储单元,用于根据所述查询 ID 存储所述请求数据包。

12. 如权利要求 10 所述的搜索引擎模块的测试系统,其特征在于,所述第二录制桩包括:

第二接收单元,用于接收所述在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包;

第二转发单元,用于获取所述结果数据包之中的查询 ID,并将所述结果数据包转发至所述在线搜索引擎模块;以及

第二存储单元,用于根据所述查询 ID 存储所述结果数据包。

13. 如权利要求 9-12 中任一项所述的搜索引擎模块的测试系统,其特征在于,所述测试回放装置包括:

数据包接收模块,用于接收所述在线录制装置发送的请求数据包和结果数据包,并建立所述请求数据包和所述结果数据包之间的关联关系;

请求发送模块,用于向所述第一被测搜索引擎模块和所述第二被测搜索引擎模块发送所述请求数据包;

查询模块,用于根据所述查询 ID 确定与所述请求数据包对应的结果数据包;

结果数据发送模块,用于将所述结果数据包分别发送到所述第一被测搜索引擎模块和

所述第二被测搜索引擎模块以获得所述第一测试结果和所述第二测试结果；

分析模块,用于根据所述第一测试结果和所述第二测试结果对所述第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。

14. 如权利要求 13 所述的搜索引擎模块的测试系统,其特征在于,所述结果数据发送模块还用于在所述查询模块未确定与所述请求对应的结果数据包时,返回预设的结果数据包。

15. 一种在线录制装置,其特征在于,包括:

第一录制桩,所述第一录制桩连接在在在线上游搜索引擎模块和在线搜索引擎模块之间,用于录制所述在在线上游搜索引擎模块向所述在线搜索引擎模块发送的请求数据包;以及

第二录制桩,所述第二录制桩连接在所述在线搜索引擎模块和在线下游搜索引擎模块之间,用于录制所述在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包。

16. 如权利要求 15 所述的在线录制装置,其特征在于,所述第一录制桩包括:

第一接收单元,用于接收所述在在线上游搜索引擎模块向所述在线搜索引擎模块发送的请求数据包;

第一转发单元,用于获取所述请求数据包中的查询 ID,并将所述请求数据包转发至所述在线搜索引擎模块;

第一存储单元,用于根据所述查询 ID 存储所述请求数据包。

17. 如权利要求 15 所述的在线录制装置,其特征在于,所述第二录制桩包括:

第二接收单元,用于接收所述在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包;

第二转发单元,用于获取所述结果数据包之中的查询 ID,并将所述结果数据包转发至所述在线搜索引擎模块;以及

第二存储单元,用于根据所述查询 ID 存储所述结果数据包。

18. 一种测试回放装置,其特征在于,包括:

数据包接收模块,用于接收所述在线录制装置发送的请求数据包和结果数据包,并建立所述请求数据包和所述结果数据包之间的关联关系;

请求发送模块,用于向所述第一被测搜索引擎模块和所述第二被测搜索引擎模块发送所述请求数据包;

查询模块,用于根据所述查询 ID 确定与所述请求数据包对应的结果数据包;

结果数据发送模块,用于将所述结果数据包分别发送到所述第一被测搜索引擎模块和所述第二被测搜索引擎模块以获得所述第一测试结果和所述第二测试结果;以及

分析模块,用于根据所述第一测试结果和所述第二测试结果对所述第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。

19. 如权利要求 18 所述的测试回放装置,其特征在于,所述结果数据发送模块还用于在所述查询模块未确定与所述请求对应的结果数据包时,返回预设的结果数据包。

## 搜索引擎模块的测试方法、系统及其装置

### 技术领域

[0001] 本发明涉及软件测试技术领域，特别涉及一种搜索引擎模块的测试方法、系统及其装置。

### 背景技术

[0002] 差异性 (DIFF) 测试是目前软件测试的一种重要测试手段。特别是对搜索引擎模块来说，随着搜索业务的不断扩展，需要对搜索引擎模块不断地进行升级，因此搜索引擎模块的差异性测试越来越重要。现有对搜索引擎搜索模块进行差异性测试，都是通过搭建整个搜索环境或者被测搜索引擎模块的上下游真实模块来实现的。如图 1 示，首先需要搭建所有相互关联的搜索引擎模块，即建立上下游搜索引擎模块之间的关系。之后将请求分别发送给新版本与对比版本的搜索引擎模块，两个版本的搜索引擎模块各自与后端的搜索引擎模块进行交互，并向上游的搜索引擎模块返回结果。传统的差异性测试就是比较两个版本的搜索引擎模块分别返回的结果。为了提高差异性测试的模块分支覆盖率，通常需要提供大量的后端结果，因此导致需要大量的后端的搜索引擎模块。由于其数量比较多，因此有些模块需要资源单台机器无法满足，包括 CPU、内存、存储空间等资源无法满足，从而导致测试时机器数量需求比较大，不仅造成测试复杂度提高，而且耗费大量的测试时间和测试成本。

[0003] 具体地，现有搜索引擎测试方法有如下缺点：

[0004] (1) 很容易由于测试环境的“噪声”，导致测试结果不准确。其中，以上所述的“噪声”包括网络环境不稳定、上下游模块不稳定等因素。尤其是上下游搜索引擎模块很多时，任何一个搜索引擎模块的不稳定或者网络异常，都有可能会导致测试结果的不准确。而且，由于无法再现，从而导致追查成本很高。

[0005] (2) 容易受到上下游搜索引擎模块性能瓶颈的制约。例如，被测搜索引擎模块本身能承受的压力能达到 1000q/s 以上，但如果后端搜索模块能承受的压力大小只有 300q/s，这就会导致该被测搜索引擎模块无法进一步提高测试速度，从而影响整体的测试进度。

[0006] (3) 机器占用多，测试环境搭建和维护成本高。测试过程中搭建上下游搜索引擎模块，不仅需要较长的时间，而且需要占用大量的测试机。在机器资源有限的情况下，要在测试过程中维护一套专属的测试环境，导致成本很高。

[0007] (4) 差异性测试结果复现困难，成本高。由于搜索引擎中后端库数据会定时更新，测试环境无法长期保存，造成前一次测试发现的差异性测试结果在后一次测试中不能复现，增大了追查的成本。

### 发明内容

[0008] 本发明的目的旨在至少解决上述技术缺陷之一。

[0009] 为此，本发明实施例的目的在于提出一种搜索引擎模块的测试方法，该搜索引擎模块的测试方法能够有效解决上下游模块不稳定的问题，还能减少测试环境资源使用，降

低搭建和维护的成本。

[0010] 本发明实施例的第二个目的在于提供一种搜索引擎模块的测试系统。

[0011] 本发明实施例的第三个目的在于提供一种在线录制装置。

[0012] 本发明实施例的第四个目的在于提供一种测试回放装置。

[0013] 为达到上述目的,本发明第一方面的实施例公开了一种搜索引擎模块的测试方法,包括:录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包,同时录制在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包;建立所述请求数据包和所述结果数据包之间的关联关系;依次将所述请求数据包和对应的所述结果数据包注入到所述第一被测搜索引擎模块中以获得所述第一被测搜索引擎模块产生的第一测试结果;依次将所述请求数据包和对应的所述结果数据包注入到所述第二被测搜索引擎模块中以获得所述第二被测搜索引擎模块产生的第二测试结果;根据所述第一测试结果和所述第二测试结果对所述第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。

[0014] 本发明实施例通过从在线搜索引擎模块获取真实的在线数据,并根据该真实的在线数据对第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性测试,可以避免“噪声”的影响,而且无需搭建真实的上下游搜索引擎模块,只要通过一台测试机进行仿真即可,因此提高了搜索引擎模块进行差异性测试的准确性和测试效率,降低了测试成本。此外,本发明实施例也能够有效地解决上下游模块不稳定、性能瓶颈等环境问题,极大地避免了由于测试环境噪声导致的差异性测试结果不准确的问题。本发明实施例通过减少测试环境资源使用、搭建和维护,极大地节省了机器资源,降低测试成本,提高了测试效率。此外,这种基于数据包录制回放的模块级差异性测试方法,通用性很高,可以使用在搜索端的其它模块,并应用于其它产品线。

[0015] 在本发明的一个实施例中,所述录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包进一步包括:接收所述在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包;获取所述请求数据包之中的查询 ID,并将所述请求数据包转发至所述在线搜索引擎模块;以及根据所述查询 ID 存储所述请求数据包。

[0016] 在本发明的一个实施例中,所述录制在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包进一步包括:接收所述在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包;获取所述结果数据包之中的查询 ID,并将所述结果数据包转发至所述在线搜索引擎模块;以及根据所述查询 ID 存储所述结果数据包。

[0017] 在本发明的一个实施例中,依据所述查询 ID 建立所述请求数据包和所述结果数据包之间的关联关系。通过上述关联关系的建立使得回放的数据稳定,利于复现真实环境下的请求处理。并且,在向第一被测搜索引擎模块和第二被测搜索引擎模块注入请求数据包之后,即可根据该关联关系查找到对应的结果数据包并注入到第一被测搜索引擎模块和第二被测搜索引擎模块中,从而加快第一被测搜索引擎模块和第二被测搜索引擎模块的处理速度,使得测试的整体速度有极大地提升。

[0018] 在本发明的一个实施例中,所述在线下游搜索引擎模块为多个,以构成在线下游搜索引擎集群。

[0019] 在本发明的一个实施例中,还包括:如果根据所述查询 ID 未确定与所述请求数据包对应的结果数据包,则返回预设的结果数据包。通过这种方式,能够极大地避免真实环境

中网络不稳定、关联模块不稳定造成的影响,将问题集中于关注被测搜索模块本身。

[0020] 本发明第二方面的实施例公开了一种搜索引擎模块的测试系统,包括:在线录制装置,用于录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包和在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包;测试回放装置,用于建立所述请求数据包和所述结果数据包之间的关联关系,并依次将所述请求数据包和对应的所述结果数据包注入到所述第一被测搜索引擎模块中以获得所述第一被测搜索引擎模块产生的第一测试结果,以及依次将所述请求数据包和对应的所述结果数据包注入到所述第二被测搜索引擎模块中以获得所述第二被测搜索引擎模块产生的第二测试结果,根据所述第一测试结果和所述第二测试结果对所述第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。

[0021] 本发明提高了搜索引擎模块进行差异性测试的准确性和测试效率,降低了测试成本。有效地解决上下游模块不稳定、性能瓶颈等环境问题,极大地避免了由于测试环境噪声导致的差异性测试结果不准确的问题。通过减少测试环境资源使用、搭建和维护,搜索引擎模块的测试系统极大地节省了机器资源,降低测试成本,提高了测试效率。此外,这种基于数据包录制回放的模块级差异性测试系统,通用性很高,可以使用在搜索端的其它模块,并应用于其它产品线。

[0022] 本发明第三方面的实施例公开了一种在线录制装置,包括:第一录制桩,所述第一录制桩连接在在线上游搜索引擎模块和在线搜索引擎模块之间,用于录制所述在线上游搜索引擎模块向所述在线搜索引擎模块发送的请求数据包;以及第二录制桩,所述第二录制桩连接在所述在线搜索引擎模块和在线下游搜索引擎模块之间,用于录制所述在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包。

[0023] 本发明第四方面的实施例公开了一种测试回放装置,包括:数据包接收模块,用于接收所述在线录制装置发送的请求数据包和结果数据包,并建立所述请求数据包和所述结果数据包之间的关联关系;请求发送模块,用于向所述第一被测搜索引擎模块和所述第二被测搜索引擎模块发送所述请求数据包;查询模块,用于根据所述查询 ID 确定与所述请求数据包对应的结果数据包;结果数据发送模块,用于将所述结果数据包分别发送到所述第一被测搜索引擎模块和所述第二被测搜索引擎模块以获得所述第一测试结果和所述第二测试结果;分析模块,用于根据所述第一测试结果和所述第二测试结果对所述第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。

[0024] 本发明附加的方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

## 附图说明

[0025] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0026] 图 1 为传统搜索引擎模块的差异性测试的示意图;

[0027] 图 2 为根据本发明实施例的搜索引擎模块的测试方法的流程框图;

[0028] 图 3 为根据本发明实施例的搜索引擎模块的测试方法的录制数据包流程图;

[0029] 图 4 为根据本发明实施例的搜索引擎模块的测试方法得到第一测试结果的流程

图；

[0030] 图 5 为根据本发明实施例的搜索引擎模块的测试方法得到第二测试结果的流程图；

[0031] 图 6 为根据本发明实施例的搜索引擎模块的测试系统的结构示意图；

[0032] 图 7 为根据本发明实施例的在线录制装置通信示意图；以及

[0033] 图 8 为根据本发明实施例的测试回放装置测试示意图。

### 具体实施方式

[0034] 下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，仅用于解释本发明，而不能解释为对本发明的限制。

[0035] 下面参考图 2 至图 5 描述根据本发明实施例的搜索引擎模块的测试方法。

[0036] 步骤 S110：录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包，同时录制在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包。在本发明的实施例中，对第一被测搜索引擎模块和第二被测搜索引擎模块的差异性测试所依赖的数据为从在线搜索引擎模块获取的真实数据，从而提高测试的精确性。并且，由于录制在线上进行，不占用测试时间，从而可以极大地提高测试效率。其中，在本发明的一个实施例中，在线下游搜索引擎模块可为多个，以构成在线下游搜索引擎集群。当然在本发明的其他实施例中，在线上游搜索引擎模块也可多个，构成在线上游搜索引擎集群。

[0037] 为了便于描述，在此将在线上游搜索引擎模块记作模块 A，在线搜索引擎模块记作模块 B，在线下游搜索引擎集群记作 C1, C2, C3...Cn。可以理解的是，上述模块 A，模块 B，C1, C2, C3...Cn 仅出于示例的目的，本发明实施例并不限于此。录制过程可以在即有的真实环境中实时进行录制，这样在测试时可以直接使用已录制好的数据，不仅能够节约录制时间，而且获得的数据更接近真实的情况，从而提高了测试的精确度。

[0038] 下面参考图 3 描述请求数据包以及结果数据包的录制过程。

[0039] 录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包进一步包括：

[0040] 步骤 S111：接收在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包，例如，可记录请求数据包为 A2B。在本发明的一个实施例中，该请求数据包中包括查询 ID(QueryID)。可以理解的是，上述 A2B 仅出于示例的目的，本发明实施例并不限于此。

[0041] 步骤 S112：解析并转发该请求数据包具体包括获取该请求数据包之中的查询 ID，并将该请求数据包转发至在线搜索引擎模块。

[0042] 步骤 S113：存储请求数据包，并为该请求数据包建立唯一标识。优选地，在本发明的一个实施例中，可将查询 ID 作为该请求数据包的唯一标识。

[0043] 其中，录制在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包进一步包括：

[0044] 步骤 S114：接收在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包。其中，在线搜索引擎模块在接收到请求数据包之后会进行相应处理，并将处理结果及相应的查询 ID 发送至在线下游搜索引擎模块（或者在线下游搜索引擎集群），在线下游搜索引



擎模块（或者在线下游搜索引擎集群）在完成处理之后，向该在线搜索引擎模块返回结果数据包，该过程为现有技术，因此在此不再赘述。

[0045] 在本发明的一个实施例中，如果为多个在线下游搜索引擎集群的话，可记录结果数据包为 C12B、C22B、C32B... Cn2B。同样可以理解的是，上述 C12B、C22B、C32B... Cn2B 仅出于示例的目的，本发明实施例并不限于此。

[0046] 步骤 S115：解析并转发结果数据包。具体包括获取结果数据包之中的查询 ID，并将结果数据包转发至在线搜索引擎模块。

[0047] 步骤 S116：存储结果数据包。同样，优选地，在本发明的一个实施例中，可将查询 ID 作为该结果数据包的唯一标识。

[0048] 在本发明的实施例中，可将请求数据包和结果数据包集中存储，从而便于后续的测试回放。其中，优选地，上述请求数据包和结果数据包可以采用统一的格式存储，从而便于建立请求数据包和结果数据包之间的关联关系。例如，数据包格式可以如表 1 所示。可以理解的是，上述数据包格式仅出于示例的目的，本发明实施例并不限于此。

[0049]	VALID_NUM	ID	LEN N	PACKAGE_DATA
--------	-----------	----	----------	--------------

[0050] 表 1

[0051] 其中，

[0052] VALID\_NUM：为固定内容，标识数据包是否有效。

[0053] ID：数据包的唯一标示。

[0054] LEN：二进制通信数据包的长度。

[0055] PACKAGE\_DATA：搜索引擎模块之间通信的完整的二进制数据包。

[0056] 步骤 S120：建立请求数据包和结果数据包之间的关联关系；

[0057] 其中，在本发明的实施例中，上述 ID 可以是查询 ID，这样如果请求数据包和结果数据包均采用相同的查询 ID 作为 ID，则就可建立请求数据包和结果数据包之间的关联关系。

[0058] 在本发明的一个实施例中，依据查询 ID 建立请求数据包和结果数据包之间的关联关系。如表 1 所示，ID 这一项表示查询 ID。可以理解的是，上述查询 ID 仅出于示例目的，请求数据包和结果数据包之间的关联关系也可以使用自定义的签名算法等计算，本发明不限于此。

[0059] 步骤 S130：依次将请求数据包和对应的结果数据包注入到第一被测搜索引擎模块（例如新版本）中以获得第一被测搜索引擎模块产生的第一测试结果。在该实施例中，由于已经将在线获得的真实的请求数据包和对应的结果数据包进行了保存，因此对于第一被测搜索引擎模块和第二被测搜索引擎模块的上下游搜索引擎模块来说就可根据请求数据包和对应的结果数据包进行仿真得到，该仿真可在一台测试机上进行即可。具体地，假设以采用一台测试机进行测试为例进行描述，如图 4 所示，步骤 S130 进一步包括：

[0060] 步骤 S131：测试机将请求数据包 A2B 注入到第一被测搜索引擎模块。

[0061] 步骤 S132：测试机提取请求数据包中的查询 ID。

[0062] 步骤 S133：测试机根据查询 ID 确定对应的结果数据包。在本发明的一个实施例

中,在结果数据包 C12B、C22B、C32B... Cn2B 中寻找查询 ID 相同的数据包,例如 C12B 为对应的结果数据包。

[0063] 步骤 S134 :如果根据查询 ID 可以确定与请求数据包对应的结果数据包,则测试机将结果数据包 C12B 注入到第一被测搜索引擎模块。

[0064] 步骤 S135 :如果根据查询 ID 无法或不能确定与请求数据包对应的结果数据包,则向第一被测搜索引擎模块注入预设的结果数据包,从而能够极大地避免真实环境中网络不稳定、关联模块不稳定造成的影响,将问题集中于关注被测搜索模块本身。

[0065] 步骤 S136 :接收第一被测搜索引擎模块产生的第一测试结果,并将返回的第一测试结果记为 B2A\_base。可以理解的是,上述 B2A\_base 仅出于示例的目的,本发明实施例并不限于此。

[0066] 步骤 S140 :依次将请求数据包和对应的结果数据包注入到第二被测搜索引擎模块中以获得第二被测搜索引擎模块产生的第二测试结果,同样以采用一台测试机进行测试为例进行描述,如图 5 所示,进一步包括:

[0067] 步骤 S141 :测试机将请求数据包 A2B 注入到第二被测搜索引擎模块;

[0068] 步骤 S142 :测试机提取请求数据包的查询 ID;

[0069] 步骤 S143 :测试机根据查询 ID 确定对应的结果数据包,例如 C12B 为对应的结果数据包。在本发明的一个实施例中,在结果数据包 C12B、C22B、C32B... Cn2B 中寻找查询 ID 相同的数据包,并将返回的结果数据包记为 C12B。

[0070] 步骤 S144 :如果根据查询 ID 确定与请求数据包对应的结果数据包,将结果数据包 C12B 注入到第二被测搜索引擎模块。

[0071] 步骤 S145 :如果根据查询 ID 未确定与请求数据包对应的结果数据包,则向第二被测搜索引擎模块注入预设的结果数据包,从而能够极大地避免真实环境中网络不稳定、关联模块不稳定造成的影响,避免出现由于未找到结构数据包而出现的测试停顿。

[0072] 步骤 S146 :接收第二被测搜索引擎模块产生的第二测试结果,例如将第二测试结果记为 B2A\_test。可以理解的是,上述 B2A\_test 仅出于示例的目的,本发明实施例并不限于此。

[0073] 步骤 S150 :根据第一测试结果和第二测试结果对第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。在本发明的一个实施例中,比较 B2A\_base 和 B2A\_test 是否存在非预期差异。在本发明的实施例中,由于对第一被测搜索引擎模块和第二被测搜索引擎模块输入的请求数据包和结果数据包都是相同的,且是稳定一致的,因此如果存在非预期的差异性(diff),则即可判断程序出现问题。此外,由于请求数据包和结果数据包完全有测试机仿真获得,因此不存在性能瓶颈的问题。并且,请求数据包和结果数据包都存储在测试机之中,可以在任何情况下完整地进行稳定复现。

[0074] 根据本发明实施例的搜索引擎模块的测试方法提高了搜索引擎模块进行差异性测试的准确性和测试效率,降低了测试成本。有效地解决上下游模块不稳定、性能瓶颈等环境问题,极大地避免了由于测试环境噪声导致的差异性测试结果不准确的问题。通过减少测试环境资源使用、搭建和维护,搜索引擎模块的测试方法极大地节省了机器资源,降低测试成本,提高了测试效率。此外,这种基于数据包录制回放的模块级差异性测试方法,通用性很高,可以使用在搜索端的其它模块,并应用于其它产品线。

[0075] 下面参考图 6 至图 8 述根据本发明实施例的搜索引擎模块的测试系统。

[0076] 如图 6 所示,本发明实施例的搜索引擎模块的测试系统包括在线录制装置 100 和测试回放装置 200。在线录制装置 100 用于录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包和在线下游搜索引擎模块向所述在线搜索引擎模块返回的结果数据包。测试回放装置 200 用于建立请求数据包和结果数据包之间的关联关系,并依次将请求数据包和对应的结果数据包注入到第一被测搜索引擎模块中以获得第一被测搜索引擎模块产生的第一测试结果,以及依次将请求数据包和对应的所述结果数据包注入到第二被测搜索引擎模块中以获得第二被测搜索引擎模块产生的第二测试结果,并根据第一测试结果和所述第二测试结果对第一被测搜索引擎模块和第二被测搜索引擎模块进行差异性分析。在本发明的实施例中,对第一被测搜索引擎模块和第二被测搜索引擎模块的差异性测试所依赖的数据为从在线搜索引擎模块获取的真实数据,从而提高测试的精确性。并且,由于录制在线上进行,不占用测试时间,从而可以极大地提高测试效率。

[0077] 其中,具体地,在线录制装置 100 包括第一录制桩 110 和第二录制桩 120。在本发明的一个实施例中,第一录制桩 110 和第二录制桩 120 作为连接模块之间的通信链路,可以由 C++ 实现。可以理解的是,上述 C++ 实现录制桩仅出于示例的目的,本发明实施例并不限于此,本领域技术人员可采用任何语言实现,这些均应包含在本发明的保护范围之内。

[0078] 优选地,上述请求数据包和结果数据包可以采用统一格式存储。例如,数据包格式可以如表 1 所示。可以理解的是,上述数据包格式仅出于示例的目的,本发明实施例并不限于此。

[0079]	VALID_NUM	ID	LEN	PACKAGE_DATA
--------	-----------	----	-----	--------------

[0080] 表 1

[0081] 其中,

[0082] VALID\_NUM:为固定内容,标识数据包是否有效。

[0083] ID:数据包的唯一标示。其中,在本发明的实施例中,上述 ID 可以是查询 ID,这样如果请求数据包和结果数据包均采用相同的查询 ID 作为 ID,则就可建立请求数据包和结果数据包之间的关联关系。

[0084] LEN:二进制通信数据包的长度。

[0085] PACKAGE\_DATA:模块间通信的完整的二进制数据包。

[0086] 如图 7 所示,第一录制桩 110 连接在在线上游搜索引擎模块 A 和在线搜索引擎模块 B 之间,用于录制在线上游搜索引擎模块 A 向在线搜索引擎模块 B 发送的请求数据包,并将请求数据包转发至在线搜索引擎模块 B。第二录制桩 120 连接在在线搜索引擎模块 B 和在线下游搜索引擎模块 C 之间,用于录制在线下游搜索引擎模块 C 向在线搜索引擎模块 B 返回的结果数据包,并将结果数据包转发至在线搜索引擎模块 B。其中,在本发明的一个实施例中,在线下游搜索引擎模块 C 为多个,以构成在线下游搜索引擎集群。当然在本发明的其他实施例中,在线上游搜索引擎模块也可为多个,构成在线上游搜索引擎集群。可以理解的是,本发明的一个实施例中的第二录制桩 120 数量为 1,在线下游搜索引擎模块 C 为多个仅出于示例目的,本发明不限于此,第二录制桩 120 与在线下游搜索引擎模块 C 的关系可以

为一对一,一对多,多对多等。

[0087] 如图 6 所示,第一录制桩 110 包括第一接收单元 111、第一转发单元 112 和第一存储单元 113。其中,第一接收单元 111 用于接收在线上游搜索引擎模块 A 向在线搜索引擎模块 B 发送的请求数据包。第一转发单元 112 用于获取请求数据包中的查询 ID,并将请求数据包转发至在线搜索引擎模块 B。第一存储单元 113 用于根据查询 ID 存储请求数据包。第二录制桩 120 包括第二接收单元 121、第二转发单元 122 和第二存储单元 123。第二接收单元 121 用于接收在线下游搜索引擎模块 C 向在线搜索引擎模块 B 返回的结果数据包。第二转发单元 122 用于获取结果数据包之中的查询 ID,并将结果数据包转发至在线搜索引擎模块 B。第二存储单元 123 用于根据查询 ID 存储结果数据包。

[0088] 其中,在本发明的实施例中,在线录制装置 100 还用于将第一存储单元 113 和第二存储单元 123 存储的请求数据包和结果数据包进行汇总处理,当然本领域技术人员应当理解的是,该汇总处理为可选步骤。

[0089] 如图 8 所示,测试回放装置 200 用于建立请求数据包和结果数据包之间的关联关系,并依次将请求数据包和对应的结果数据包注入到第一被测搜索引擎模块 D 中以获得第一被测搜索引擎模块 D 产生的第一测试结果,以及依次将请求数据包和对应的结果数据包注入到第二被测搜索引擎模块 E 中以获得第二被测搜索引擎模块 E 产生的第二测试结果,根据第一测试结果和第二测试结果对第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 进行差异性分析。

[0090] 具体地,如图 6 所示,测试回放装置 200 包括数据包接收模块 210、请求发送模块 220、查询模块 230、结果数据发送模块 240 和分析模块 250。其中,数据包接收模块 210 用于接收在线录制装置发送的请求数据包和结果数据包,并建立请求数据包和结果数据包之间的关联关系。请求发送模块 220 用于向第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 发送请求数据包

[0091] 查询模块 230 用于根据查询 ID 确定与请求数据包对应的结果数据包。结果数据发送模块 240 用于将结果数据包分别发送到第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 以获得第一测试结果和第二测试结果。

[0092] 分析模块 250 用于根据第一测试结果和第二测试结果对第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 进行差异性分析。在本发明的实施例中,由于对第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 输入的请求数据包和结果数据包都是相同的,且是稳定一致的,因此如果存在非预期的差异性(diff),则即可判断程序出现问题。此外,由于请求数据包和结果数据包完全有测试机仿真获得,因此不存在性能瓶颈的问题。并且,请求数据包和结果数据包都存储在测试机之中,可以在任何情况下完整地进行稳定复现。

[0093] 优选地,结果数据发送模块 240 还用于在查询模块未确定与请求对应的结果数据包时,返回预设的结果数据包,从而能够极大地避免真实环境中网络不稳定、关联模块不稳定造成的影响,避免出现由于未找到结构数据包而出现的测试停顿。

[0094] 根据本发明实施例的搜索引擎模块的测试系统提高了搜索引擎模块进行差异性测试的准确性和测试效率,降低了测试成本。有效地解决上下游模块不稳定、性能瓶颈等环境问题,极大地避免了由于测试环境噪声导致的差异性测试结果不准确的问题。通过减少测试环境资源使用、搭建和维护,搜索引擎模块的测试系统极大地节省了机器资源,降低测

试成本,提高了测试效率。此外,这种基于数据包录制回放的模块级差异性测试系统,通用性很高,可以使用在搜索端的其它模块,并应用于其它产品线。

[0095] 下面参考图6、图7描述根据本发明实施例的在线录制装置100。在线录制装置100包括第一录制桩110和第二录制桩120,用于录制在线上游搜索引擎模块向在线搜索引擎模块发送的请求数据包和在线下游搜索引擎模块向在线搜索引擎模块返回的结果数据包。在本发明的一个实施例中,第一录制桩110和第二录制桩120作为连接模块之间的通信链路,可以由C++实现。可以理解的是,上述C++实现录制桩仅出于示例的目的,本发明实施例并不限于此。上述请求数据包和结果数据包可以采用统一格式存储。

[0096] 如图7所示,第一录制桩110连接在在线上游搜索引擎模块A和在线搜索引擎模块B之间,用于录制在线上游搜索引擎模块A向在线搜索引擎模块B发送的请求数据包,并将请求数据包转发至在线搜索引擎模块B。第二录制桩120连接在在线搜索引擎模块B和在线下游搜索引擎模块C之间,用于录制在线下游搜索引擎模块C向在线搜索引擎模块B返回的结果数据包,并将结果数据包转发至在线搜索引擎模块B。其中,在本发明的一个实施例中,在线下游搜索引擎模块C为多个,以构成在线下游搜索引擎集群。当然在本发明的其他实施例中,在线上游搜索引擎模块也可为多个,构成在线上游搜索引擎集群。可以理解的是,本发明的一个实施例中的第二录制桩120数量为1,在线下游搜索引擎模块C为多个仅出于示例目的,本发明不限于此,第二录制桩120与在线下游搜索引擎模块C的关系可以为一对一,一对多,多对多等。

[0097] 如图6所示,第一录制桩110包括第一接收单元111、第一转发单元112和第一存储单元113。其中,第一接收单元111用于接收在线上游搜索引擎模块A向在线搜索引擎模块B发送的请求数据包。第一转发单元112用于获取请求数据包中的查询ID,并将请求数据包转发至在线搜索引擎模块B。第一存储单元113用于根据查询ID存储请求数据包。第二录制桩120包括第二接收单元121、第二转发单元122和第二存储单元123。第二接收单元121用于接收在线下游搜索引擎模块C向在线搜索引擎模块B返回的结果数据包。第二转发单元122用于获取结果数据包之中的查询ID,并将结果数据包转发至在线搜索引擎模块B。第二存储单元123用于根据查询ID存储结果数据包。

[0098] 其中,在本发明的实施例中,在线录制装置100还用于将第一存储单元113和第二存储单元123存储的请求数据包和结果数据包进行汇总处理,当然本领域技术人员应当理解的是,该汇总处理为可选步骤。

[0099] 根据本发明实施例的在线录制装置将录制的请求数据包与结果数据包一一对应,在数据包回放时根据请求返回真实的结果数据包,达到复现真实环境下的请求处理的目的,具有稳定且准确的特点。

[0100] 下面参考图6、图8描述根据本发明实施例的测试回放装置200,测试回放装置200用于建立请求数据包和结果数据包之间的关联关系,并依次将请求数据包和对应的结果数据包注入到第一被测搜索引擎模块D中以获得第一被测搜索引擎模块D产生的第一测试结果,以及依次将请求数据包和对应的结果数据包注入到第二被测搜索引擎模块E中以获得第二被测搜索引擎模块E产生的第二测试结果,根据第一测试结果和第二测试结果对第一被测搜索引擎模块D和第二被测搜索引擎模块E进行差异性分析。

[0101] 具体地,如图6所示,测试回放装置200包括数据包接收模块210、请求发送模块

220、查询模块 230、结果数据发送模块 240 和分析模块 250。其中,数据包接收模块 210 用于接收在线录制装置发送的请求数据包和结果数据包,并建立请求数据包和结果数据包之间的关联关系。请求发送模块 220 用于向第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 发送请求数据包。

[0102] 查询模块 230 用于根据查询 ID 确定与请求数据包对应的结果数据包。

[0103] 结果数据发送模块 240 用于将结果数据包分别发送到第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 以获得第一测试结果和第二测试结果。

[0104] 分析模块 250 用于根据第一测试结果和第二测试结果对第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 进行差异性分析。在本发明的实施例中,由于对第一被测搜索引擎模块 D 和第二被测搜索引擎模块 E 输入的请求数据包和结果数据包都是相同的,且是稳定一致的,因此如果存在非预期的差异性 (diff),则即可判断程序出现问题。此外,由于请求数据包和结果数据包完全有测试机仿真获得,因此不存在性能瓶颈的问题。并且,请求数据包和结果数据包都存储在测试机之中,可以在任何情况下完整地进行稳定复现。

[0105] 优选地,结果数据发送模块 240 还用于在查询模块未确定与请求对应的结果数据包时,返回预设的结果数据包,从而能够极大地避免真实环境中网络不稳定、关联模块不稳定造成的影响,避免出现由于未找到结构数据包而出现的测试停顿。

[0106] 根据本发明实施例的测试回放装置为模拟后端,因此不存在性能瓶颈。并且测试回放装置能够有效解决上下游模块不稳定的问题,还能减少测试环境资源使用,降低搭建和维护的成本。

[0107] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任何一个或多个实施例或示例中以合适的方式结合。

[0108] 尽管已经示出和描述了本发明的实施例,对于本领域的普通技术人员而言,可以理解在不脱离本发明的原理和精神的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由所附权利要求及其等同限定。

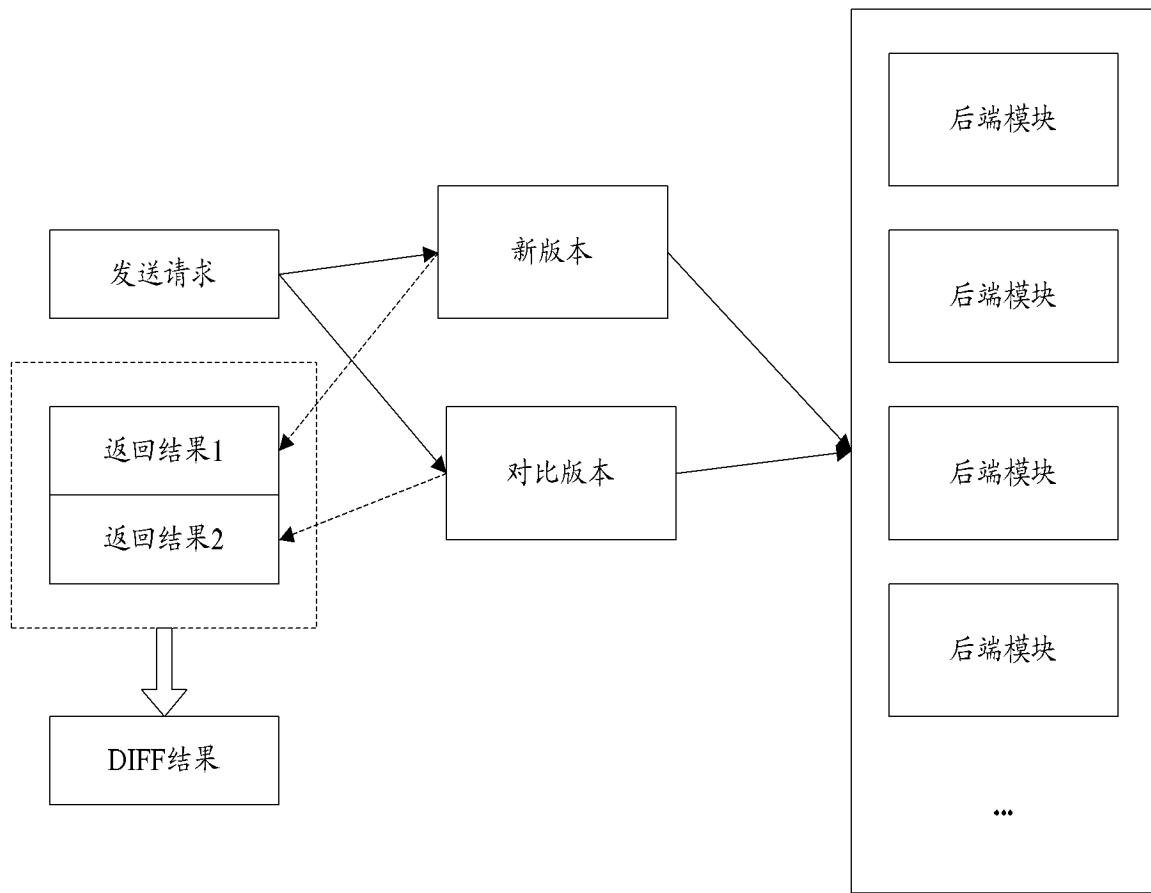


图 1

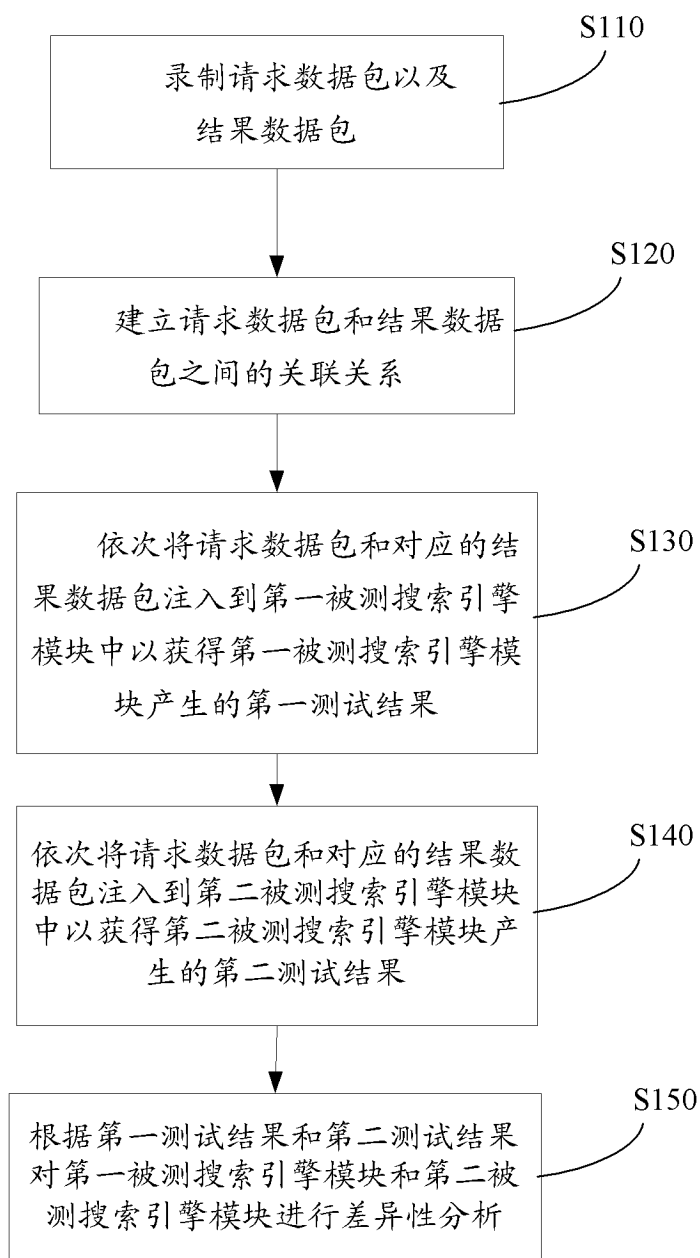


图 2



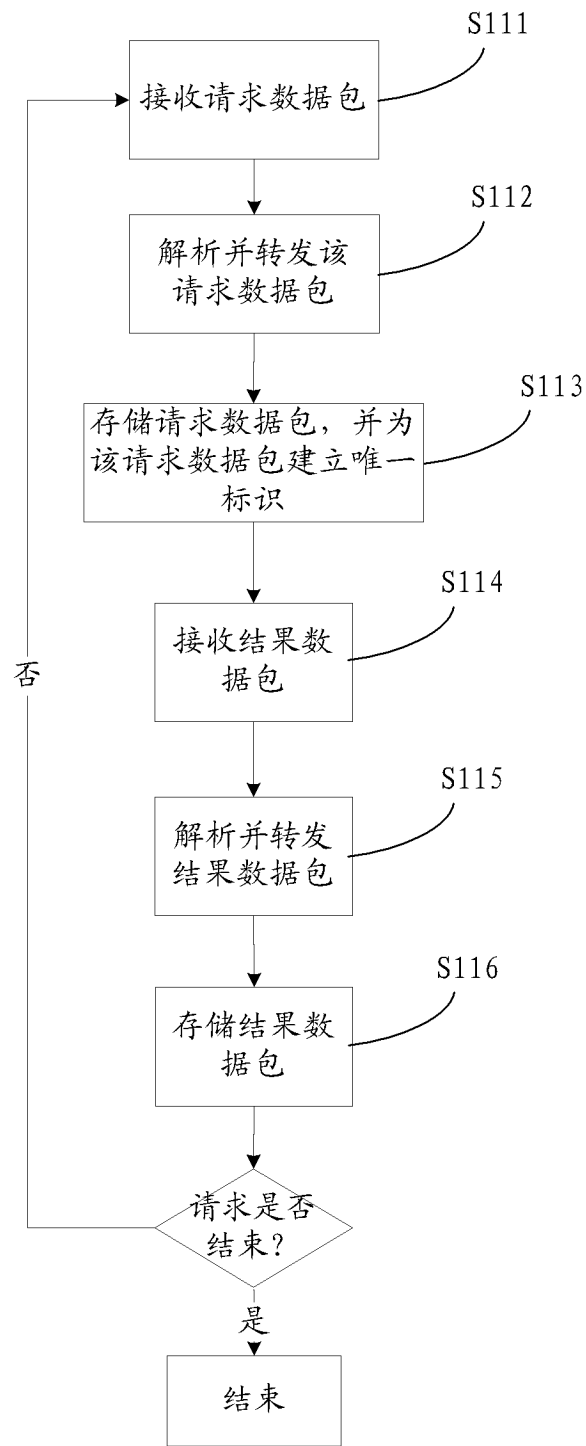


图 3

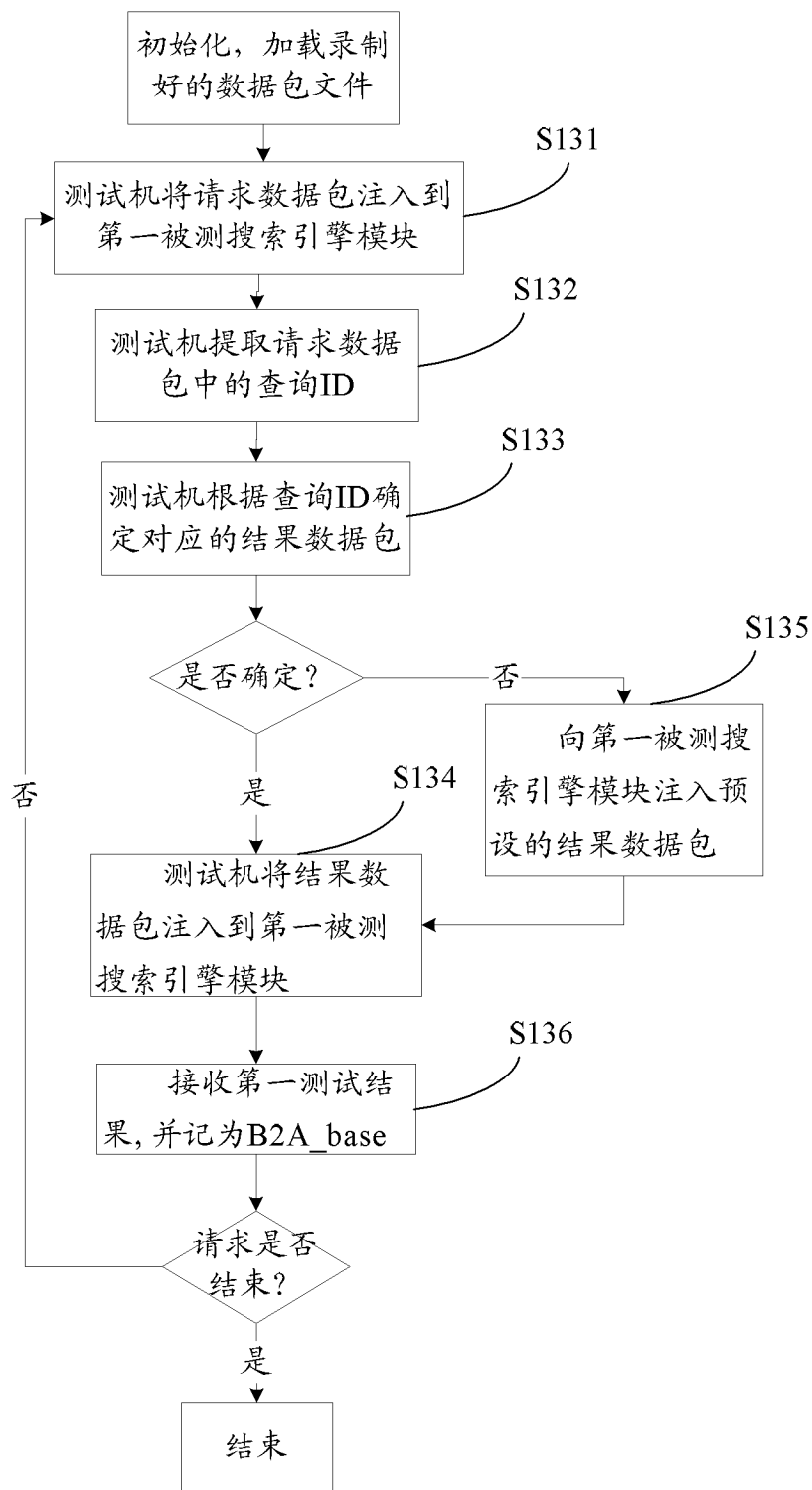


图 4

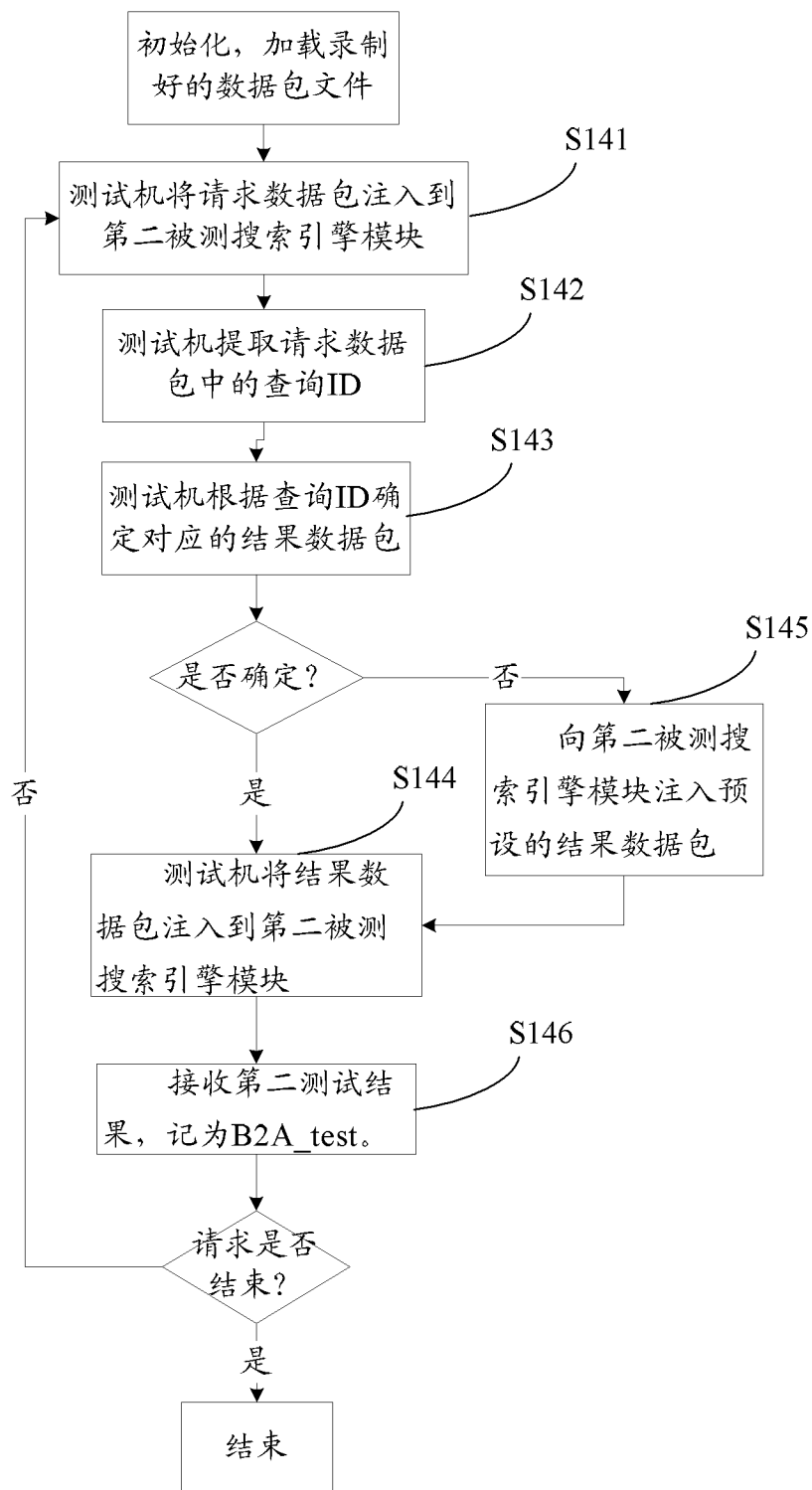


图 5

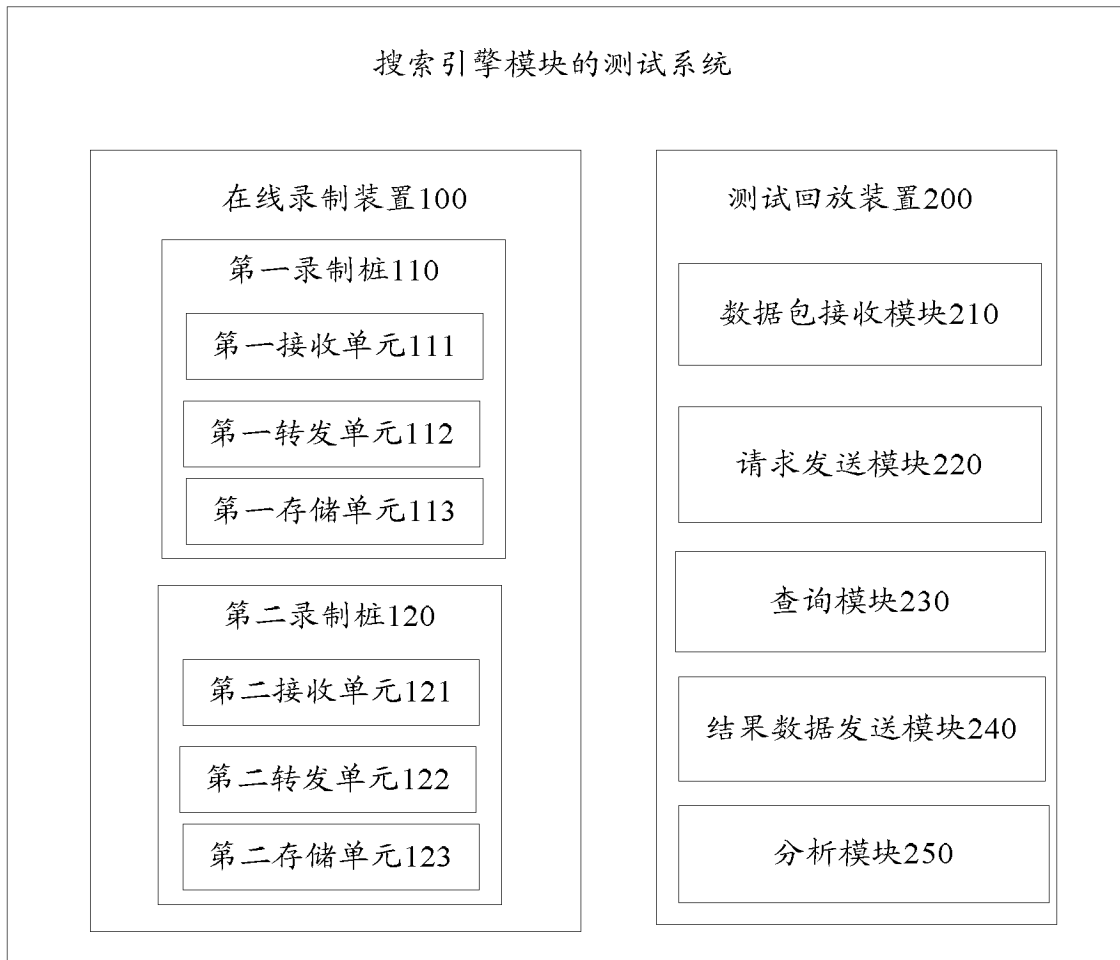


图6

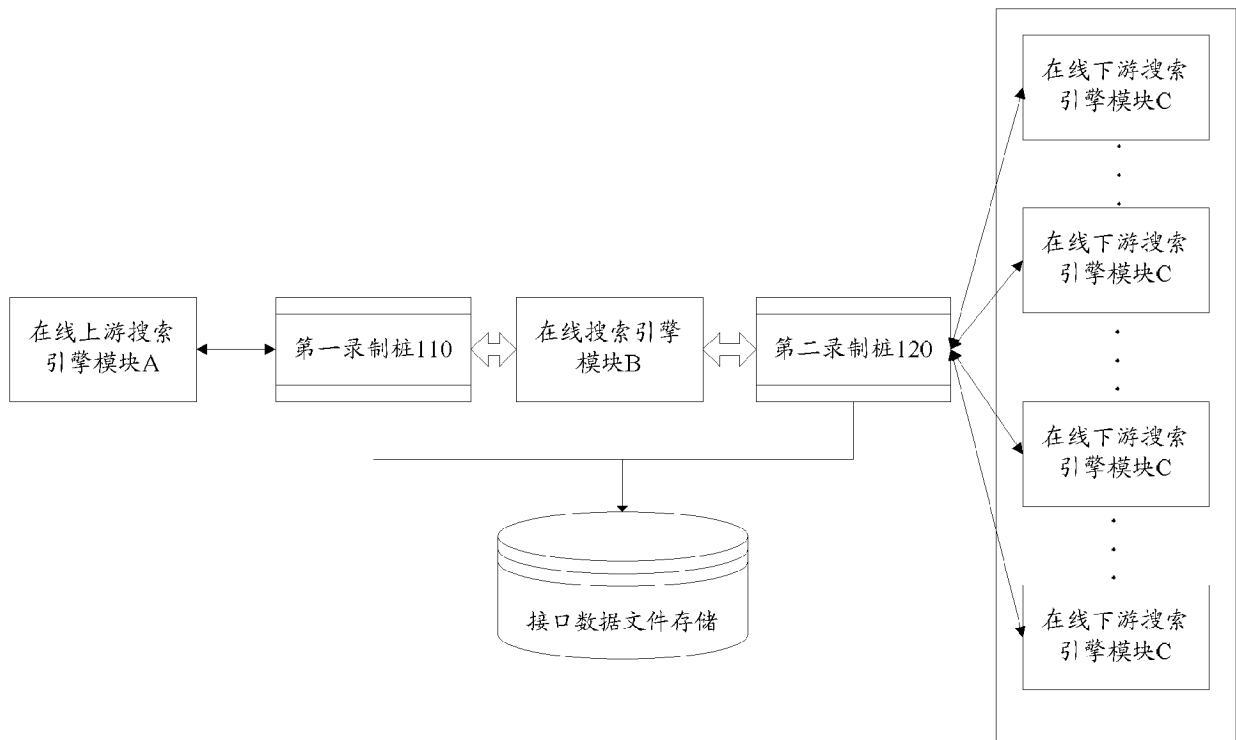


图 7

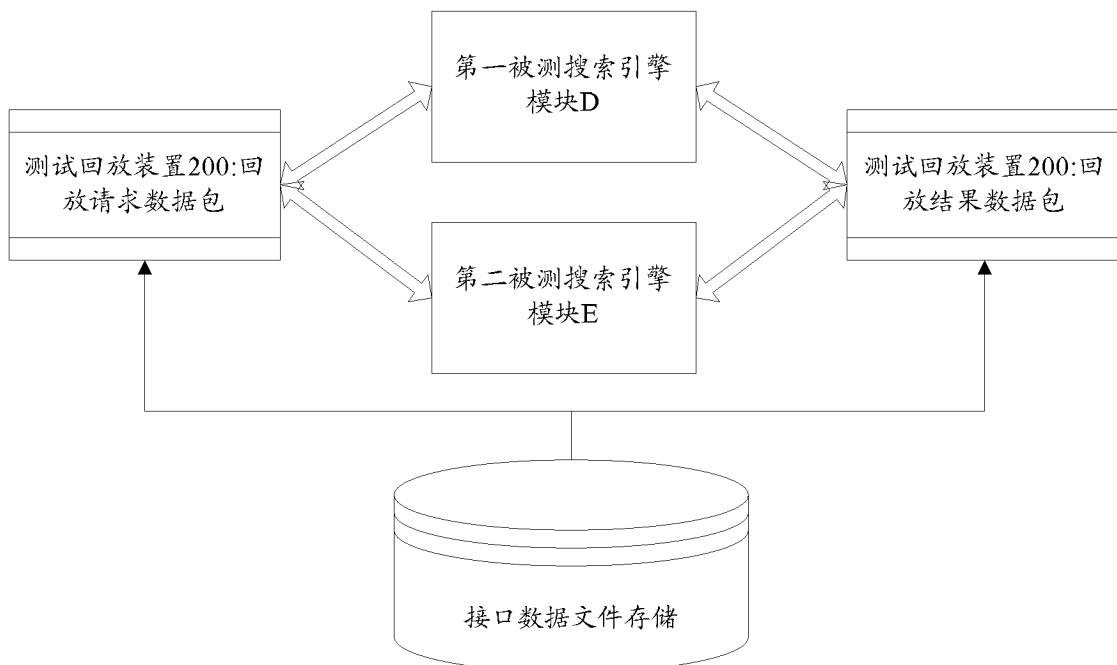


图 8