

基于 Pairwise 算法的组合测试应用研究

郑根让¹ 林正春²

(1. 中山职业技术学院 广东省中山市 528404 2. 广东技术师范大学 广东省广州市 510665)

摘要: 本文对基于 Pairwise 算法的组合测试常用实现方法进行了综述和比较分析。利用 Pairwise 算法思想对电商平台系统测试进行了需求分析及测试设计, 使用 PICT 工具对电商平台进行组合测试实验, 根据电商平台运行需求利用贪心算法对约束关系进行了优化, 进一步提高组合测试效能。为验证 PICT 在电商平台组合测试的有效性, 对电商平台相关关系进行了仿真实验和实际测试, 其在测试用例覆盖、测试生成速度和测试成本等方面具有一定优势。

关键词: Pairwise; 组合测试; PICT 工具

1 引言

随着软件技术的不断发展, 软件开发的复杂性、集成度和可靠性要求越来越高, 对软件质量保证的软件测试提出了更高的要求。同时, 随着软件在各行业的广泛应用, 软件系统功能不断强大, 系统复杂度也在持续提高, 软件系统需要支持多平台和多场景, 这就需要使用组合测试方法对软件系统进行测试。

组合测试是一种重要的软件测试方法。它以覆盖表作为测试用例集, 能够以少量的测试用例有效地检测由软件各参数之间的交互作用所触发的失效。覆盖表生成是组合测试要解决的关键问题, 也是软件测试重点研究的领域。组合测试覆盖表生成主要基于三种框架: 测试用例演化 (Evolve Test Case)、参数演化 (Evolve Parameter) 和整表演化 (Evolve Test Set)。组合测试可以通过三种框架和不同算法有效结合形成不同的覆盖表生成算法。

配对测试 (Pairwise) 可以根据已设定的条件, 自动生成在测试效率和测试覆盖率之间做出平衡的组合, 由于两因素组合测试在测试用例个数和错误检测能力上达到了较好的平衡, 因此它是目前主流的组合测试方法。

组合测试的研究与发展一直是由应用驱动的。电商平台是目前应用广泛, 功能庞杂的应用系统, 因其功能模块多, 用户需求多, 应用终端结构多样, 使得测试工作复杂而困难。

PICT 是一个现有的公共可用工具, 构建在一个灵活的组合测试用例生成引擎之上, 通过设计许多组合策略来帮助测试人员选择能最大限度地发现缺陷的输入组合子集。

本研究利用 Pairwise 算法思想对电商平台系统测试, 使用 PICT 工具对电商平台进行组合测试实验, 根据电商平台运行需求利用贪心算法对约束关系进行了优化, 进一步提高组合测试效能。

2 组合测试

2.1 组合测试及其发展

组合测试是一种充分考虑了各种因素及其组合相互作用的软件测试方法, 实践证明组合测试在多种因素及其相互作用下的复杂软件系统测试中应用极为广泛也最为有效。

组合测试 CT (combinatorial testing) 采用系统的抽样机制对参数间的交互作用进行有针对性的覆盖, 从而减少测试用例的规模^[1,2]。已有研究表明: 软件系统中大约 70% 的故障是由两个参数间的交互作用引起的; 同时, 与故障相关的参数个数一般不超过 6 个^[3]。因此, 组合测试是一种科学有效的软件测试方法。

组合测试具有广泛的应用场景和发展空间, 已发展成为一种独具特色和影响力的科学实用的软件测试方法。从上世纪末, 美国和日本的研究者将实验设计思想引入计算机软硬件测试, 并逐渐发展成组合测试方法。组合测试已在输入参数测试、配置测试、软件产品线测试、事件驱动软件测试、并发程序测试、移动应用测试和安全测试等方面得到广泛应用^[4]。

2.2 组合测试实现原理

组合测试是一种测试用例生成方法, 其对传统测试用例方法进行了综合。组合测试理论可由一个数学模型和一个假设表示:

如果影响待测软件系统的因素有 n 个^[5], 记为 $F = \{f_1, f_2, \dots, f_n\}$, 其中, 参数 f_i 有 l_i 个可能取值, 每个参数取值为 $w_i = \{0, 1, \dots, l_i - 1\}$, $1 \leq i \leq n$ 。

设 m 元组 $(w_1, w_2, \dots, w_n) (w_i \in W_i)$ 为 SUT (software under testing) 的一条测试用例。组合测试的测试用例集组成覆盖表。

SUT 的 t 维覆盖表 $CA(N; t, n, (w_1, w_2, \dots, w_n))$ 是一个 $N \times n$ 的数组, w_i 是第 i 列对应的第 i 个参数, 任意 t 个参数形成的 $N \times t$ 的子数组中包含了该 t 个参数的所有 t 元组, t 表示组合覆盖测试的强度^[6]。

以电商平台为例, 它有多种应用操作系统 (f_1), 可以使用不同的浏览器 (f_2), 有不同的网络连接方式 (f_3) 以及不同

● 基金项目: 2020 年度广东省教育厅普通高校认定类科研项目“基于人工智能 LDA 的软件测试平台关键技术研究与应用”(项目编号: 2020KTSCX333)。

的内存配置 (f4), 电商平台配置见表 1, 共有 4 个参数, 每个参数都有 3 个取值, 可以形成二维覆盖表 CA(9;2,34)。

表 1: 电商平台配置

Operating System	Web Browser	Connection Type	Memory
Windows	IE	LAN	528MB
Linux	Safari	WLAN	2GB
Macintosh	Chrome	Internet	4GB

2.3 基于Pairwise算法的组合测试

Pairwise 可以根据设定的条件, 自动生成在测试效率和测试覆盖率之间做出平衡的组合, 由于两因素组合测试在测试用例个数和错误检测能力上达到了较好的平衡, 它是目前主流的组合测试方法。当然, Pairwise 也有不足, 主要问题是 Pairwise 对于维度的分解来说, 需要对业务很熟悉, 需要正交测试法的理论支持。需要中等专业的测试人员才能完成。其次, Pairwise 还是有一定的遗漏。相比于全正交设计法来说, Pairwise 算法对于多于 2 个因素相互作用所产生的 bug, 没有覆盖到。

2.4 基于Pairwise算法的组合测试工具

现很多程序都围绕 Pairwise 算法产生, 最著名的就是 ReduceArray, SmartDesgin 和微软的 PICT 工具。

ReduceArray 按照 Pairwise 覆盖要求^[6], 输出剔除无效用例后的一个用例子集可以支持约束条件, 支持 2 维或 3 维覆盖可以给出每个用例覆盖的 Pairwise 对的数量, 可以给出输入用例集和输出用例集有哪些 Pairwise 对没有覆盖到, 给出的用例集合可以达到最优。速度很快, 效率高, 但随着用例规模的增加, 效率下降明显。在表格中, 编写的脚本来执行, 操作较简单。

SmartDesgin 符合 N 维 Pairwise 全覆盖的一组用例, 支持约束条件^[7], 支持 2 维的 Pairwise 覆盖, 给出的用例集合不是最优。速度很快, 但效率较低。在表格中进行操作, 稍有一慎可能导致结果错误。运行过程中有时会导致异常关闭。

PICT (The Pairwise Independent Combinatorial Testing tool) 是一款由微软公司出品的命令行用例生成工具, 生成的测试集可以覆盖任意 N 个变量的取值组合。基于组合测试的理论, 该测试用例集可以发现 N 个因素共同作用引发的缺陷。PICT 接受纯文本模式的文件作为输入, 输入文件中包括参数类别、取值集与约束关系, 输出是一个用来组成组合覆盖测试用例的矩阵, 每一行表示一个测试用例, 每一列代表系统的一个参数, 每一项代表测试用例对应的参数取值。PICT 可帮助测试人员产生比手动生成更高效的测试用例。

2.5 Pairwise算法在PICT中的实现思路

Pairwise 是 L. L. Thurstone(29 May1887 - 30 September 1955)^[8] 在 1927 年首先提出来的。他是美国的一位心理统计学家。Pairwise 也正是基于数学统计和对传统的正交分析法进行优化后得到的产物。

Pairwise 基于如下 2 个假设:

(1) 每一个维度都是正交的, 即每一个维度互相都没有交集。

(2) 根据数学统计分析, 73% 的缺陷 (单因子是 35%, 双因子是 38%) 是由单因子或 2 个因子相互作用产生的。19% 的缺陷是由 3 个因子相互作用产生的。因此, Pairwise 基于覆盖所有 2 因子的交互作用产生的用例集合性价比最高而产生的。

```
# assume test cases r1,...,ri-1 are already produced[9]
# slots in P reflecting combinations selected by r1,...,ri-1 are
set to covered
if there are any unused seed combinations not violating any
exclusions
add a seed combination to ri
mark all slots in P covered by the seed combination as
covered
while there are parameters with no values in ri
if ri is empty
choose a parameter interaction p from P with most
uncovered slots
pick the first uncovered combination from p
else
# assume values l1,...,lk-1 have already been chosen and
added to ri
look at subset Q of P that covers at least one parameter
with no
representation in l1,...,lk-1
look at slots in Q which values are consistent with already
chosen values in l1,...,lk-1
if there exist uncovered combinations
pick a slot with values which when added to ri would
cover the most uncovered
combinations with l1,...,lk-1 and the resulting partial test
case ri would not contain
an excluded combination
else
pick randomly a covered combination which when added
to l1,...,lk-1 would not
contain an excluded combination
add values of this combination to ri
mark the chosen combination in P as covered
```

3 电商平台测试设计

3.1 电商平台适合组合测试

电子商务平台简称电商平台, 是一个为企业或个人提供网上交易洽谈的平台。企业电子商务平台是建立在 Internet 网上进行商务活动的虚拟网络空间和保障商务顺利运营的管理环境; 是协调、整合信息流、货物流、资金流有序、关联、



图 1: 电商平台支付系统页面

表 2: 电商平台输入参数及取值

支付方式	配送方式	优惠券	礼品卡
货到付款	自营配送	商品优惠券	有
在线支付	自选快递	运费券	无
朋友代付	上门自提	优惠兑换码	申请

表 3: PICT 生成的电商平台测试用例

支付方式	配送方式	优惠券	礼品卡
在线支付	上门自提	商品优惠券	申请
朋友代付	自营配送	运费券	有
在线支付	自营配送	优惠兑换码	无
朋友代付	自选快递	商品优惠券	无
货到付款	自选快递	优惠兑换码	有
货到付款	上门自提	商品优惠券	有
货到付款	自选快递	运费券	申请
朋友代付	上门自提	优惠兑换码	无
在线支付	上门自提	运费券	有
货到付款	自营配送	商品优惠券	申请
货到付款	上门自提	运费券	无
在线支付	自选快递	优惠兑换码	申请
朋友代付	自选快递	商品优惠券	申请

高效流动的重要场所。企业、商家可充分利用电子商务平台提供的网络基础设施、支付平台、安全平台、管理平台等共享资源有效地、低成本地开展自己的商业活动。

电商平台具有终端多样化、交易复杂化、信息多元化等特点，以终端设备为例，终端可以是 PC，也可以是 SP (Smartphone)、PAD 还可以是 PDA，再比如购物的付款方式有货到付款、银行卡付款、微信支付、支付宝付款及白条等，电商平台支付方式页面如图 1 所示。因此，电商平台测试适用组合测试来解决多因素组合的测试问题。

3.2 电商平台组合测试设计

电商交易平台是一个复杂的网络交易系统，系统功能模块多，输入参数多且一般有约束关系。以某在线的电子商务平台的商品结算模块为例，商品支付模块包括支付方式、配送方式和使用优惠等 3 个功能选项。支付方式有货到付款、在线支付、朋友代付，配送方式有自营配送、自选快递和上门自提，优惠券有商品优惠券、运费券和优惠码兑换，礼品卡有使用、无和添加 3 种状态。

构建模型

通过分析电商平台可以构建支付方式、配送方式、优惠券、礼品卡等有 4 个输入参数的模型，每个参数有 3 个可选值，如表 2 所示。

生成测试用例：

将构建模型的参数及取值，输入 PICT 生成测试用例，如表 3 所示。

3.3 约束 (constraint) 处理

表 4: 具有约束条件的 PICT 生成用例

支付方式	配送方式	优惠券	礼品卡
在线支付	自营配送	商品优惠券	申请
货到付款	自选快递	优惠兑换码	无
朋友代付	上门自提	优惠兑换码	申请
朋友代付	自选快递	商品优惠券	有
朋友代付	自营配送	运费券	无
在线支付	上门自提	商品优惠券	无
在线支付	自营配送	优惠兑换码	有
货到付款	自选快递	运费券	有
在线支付	上门自提	商品优惠券	有
在线支付	自选快递	运费券	申请
货到付款	自营配送	商品优惠券	申请

表 5: 用例生成规模对比实验

Task	IPO	CTS	DDA	PICT
34	9	9	9	9
312	17	15	18	18
415317220	34	39	35	37
41320235	26	29	27	27
2100	15	10	15	15
1020	212	210	201	210

在组合测试的基础理论中, 各个因素的取值是相互独立的, 即因素 A 的取值不会影响因素 B 的取值。但是, 大多数被测试应用的因素之间存在约束关系。

PICT 默认使用两变量组合方法, 生成的测试集可以覆盖任意两个变量的取值组合。并且由于组合测试的数学模型没有描述变量之间的约束关系, 生成的用例集存在一定的冗余, 利用 PICT 可通过添加输入的约束条件来进行改进。

在电商平台中, 当支付方式选择“朋友代付”时, 不能选择所有优惠券和礼品卡。当支付方式选择“货到付款”时, 配送方式不能选择“上门自提”; 当配送方式选择“上门自提”时, 优惠券不能领取“运费券”。根据电调平台中关于支付方式、配送方式、优惠券及礼品卡之间的关联关系编写约束条件代码:

```
IF [支付方式]="朋友代付" THEN ([优惠券] <> "商品优惠券" OR [优惠券] <> "运费券" OR [优惠券] <> "优惠兑换码");
```

```
IF [支付方式]="货到付款" THEN [配送方式] in {"自营配送","自选快递"};
```

```
IF [配送方式]="上门自提" THEN [优惠券] in {"商品优惠券","优惠兑换码"};
```

将约束条件代码输入 PICT, 完成具有约束条件的用例生成, 如表 4 所示。

4 测试结果及分析

为了进一步分析 PICT 在组合测试中的效率, 以某电商平台为测试对象, 设计了 4 个覆盖数组 CA (34)、CA (312)、CA (2100) 和 CA (1020) (4 个 3 值, 12 个 3 值, 200 个 2 值, 20 个 10 值) 和 2 个混合覆盖数组 MCA (415317220)

(表示待测实例有 52 个参数, 其中, 15 个 4 值, 17 个 3 值, 20 个 2 值) 和 MCA (41320235) (表示待测实例有 56 个参数, 其中, 1 个 4 值, 20 个 3 值, 35 个 2 值), 分别使用 AETG、IPO、CTS、DDA 等方法与 PICT 进行用例覆盖表的生成规模生成对比实验, 结果如表 5 所示。

从用例覆盖表的生成规模情况可以看出, 用 PICT 生成的用例覆盖表的生成规模总体上是处于中等的, 但 PICT 核心算法的效率与其他工具相比还是占有优势的, 在工具的易用性和核心引擎的可扩展性等方面 PICT 具有较好的表现。

5 结语

总之, PICT 是一款开放实用的组合测试工具, 它是构建在一个灵活的组合测试用例生成引擎之上, 具有简单易用、低成本、高效率的特点, 特别是在测试生成的速度上明显优于其他方法。组合测试很难一次性就把工作做到很完美, 在组合测试中应该把握测试的灵活性使组合测试有更好的自适应性和可拓展性。

参考文献

- [1] 戚荣志, 王志坚, 黄宜华, 李水艳. 基于 Spark 的并行化组合测试用例集生成方法 [J]. 计算机学报, 2018, 41(6): 1284-1299.
- [2] Qi RZ, Wang ZJ, Huang YH, Li SY. Generating combinatorial test suite with spark based parallel approach. Chinese Journal of Computers, 2018, 41(6): 1284-1299 (in Chinese with English abstract).
- [3] Zamli KZ, Alkazem BY, Kendall GA. Tabu search hyper-heuristic strategy for t-way test suite generation. Applied Soft Computing, 2016, 44: 57-74.
- [4] 王燕, 聂长海, 钮鑫涛, 吴化尧, 徐家喜等. 覆盖表生成的禁忌搜索算法 [J]. 软件学报, 2012, 29(2): 192-194, 236.
- [5] 聂长海, 蒋静. 覆盖表生成的可配置贪心算法优化, 软件学报 [J]. 2013, 24(7): 1469-1483.
- [6] 黄沛. 基于 RESTful 架构的科技信息共享接口系统的设计 [J]. 软件, 2018, 39(7): 170-172.
- [7] 聂长海. 组合测试 [M] 北京: 科学出版社, 2015.
- [8] 吴化尧. 组合测试方法及其有效性研究 [D]. 南京: 南京大学. 2018
- [9] Nie CH, Leung H. A survey of combinatorial testing. ACM Computing Surveys, 2011, 43(2): 1-29.

作者简介

郑根让 (1970-), 男, 陕西省宝鸡市人。工学硕士, 副教授。研究方向为软件开发与测试。
林正春 (1981-), 男, 江西省人。博士, 副教授。研究方向为图像与自然语言处理。