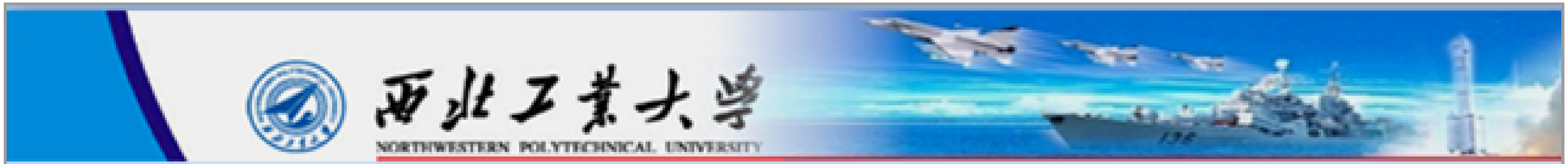
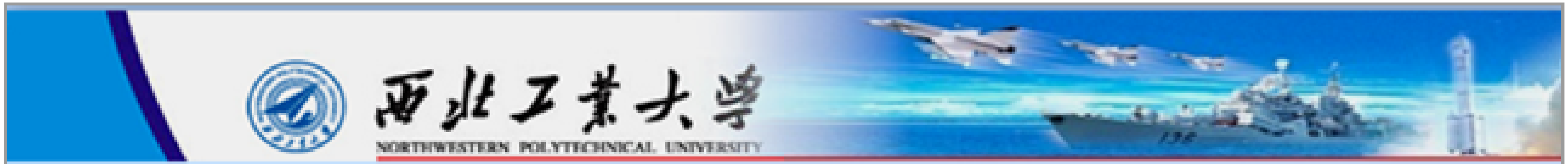


多目标优化方法在软件产品线和软件测试上的研究

西北工业大学 郑炜

2019.6.29



一、软件产品线

- TOSEM(2016) FSE(2016) SIP: Optimal Product Selection from Feature Models Using Multi-Objective Evolutionary Optimization[J]

二、软件测试

- Information Sciences(SCI 二区) Multi-objective optimisation for regression testing[J]
- IET SOFTWARE(2018) (CCF-B刊) A Multi-objective Optimization-based Mutant Reduction Approach for Mutation Testing [J]
- JSS(2019) (CCF-B刊) Towards Understanding Bugs in An Open Source Cloud Management Stack: An Empirical Study of OpenStack Software Bugs [J]
- ICSR(2019) Automatically Extracting Bug Reproducing Steps from Android Bug Reports[C]
- ICSE(2019) ReCDroid: Automatically Reproducing Android Application Crashes from Bug Reports [C]
- FSE(2019) Event Trace Reduction for Effective Bug Replay of Android Apps via Differential GUI State Analysis[C]

三、正在研究的内容

- SGX, 机器学习，深度学习，智能合约

TOSEM论文

内容提要

1.摘要

软件产品线是一组具有共同体系构架和可复用组件的软件系统，通过特征建模呈现组件模块之间依赖、互斥等关系。合理的配置特征模型直接影响和决定了整个软件产品的质量和成败，然而产品线配置问题是一个多目标选择难题，其解决方案域不唯一，因此如何获取最优配置方案解尤为重要。

西北工业大学

1.摘要

本文中我们使用了以下两种方法提高搜索效率。

1）提出一种新型的编码方式减少描述字段长度；

2）将某个优化目标作为限制条件考虑，只有该目标满足过后才考虑优化其他目标。

5

2.研究背景

近年来，软件产品线SPL领域已经引起了很多的关注。特征模型在许多大型公司中被使用，例如波音公司[Sharp 1998]，和东芝公司[Matsumoto 2007]、西门子公司[Hofman et al. 2012]。

2.研究背景

1. 一部分研究关注于如何自动分析特征模型，主要是来判断特征模型（一个或者多个）是否有效或者是否存在有效产品。

2. 另一部分研究主要关注于通过用户的偏好或者其他信息来自动寻找最佳解。

2.研究背景

本文采用特征模型

1.四个传统模型

BerkeleyDB

ERS

E-Shopping

WebPortal

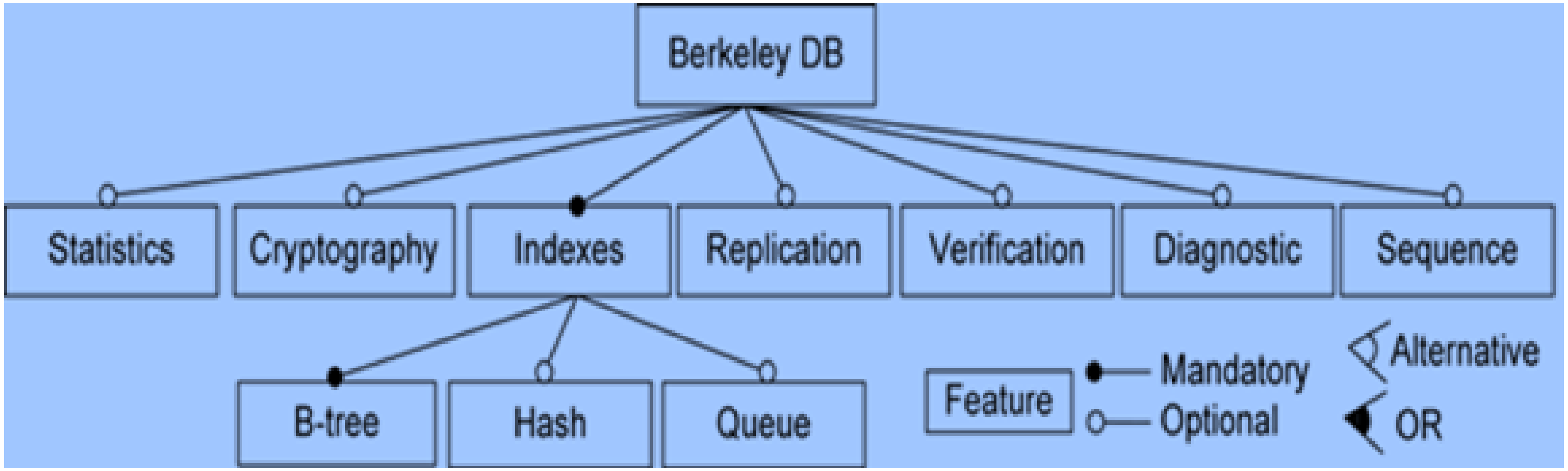
2.两个真实模型

AmazonEC2

Drupal

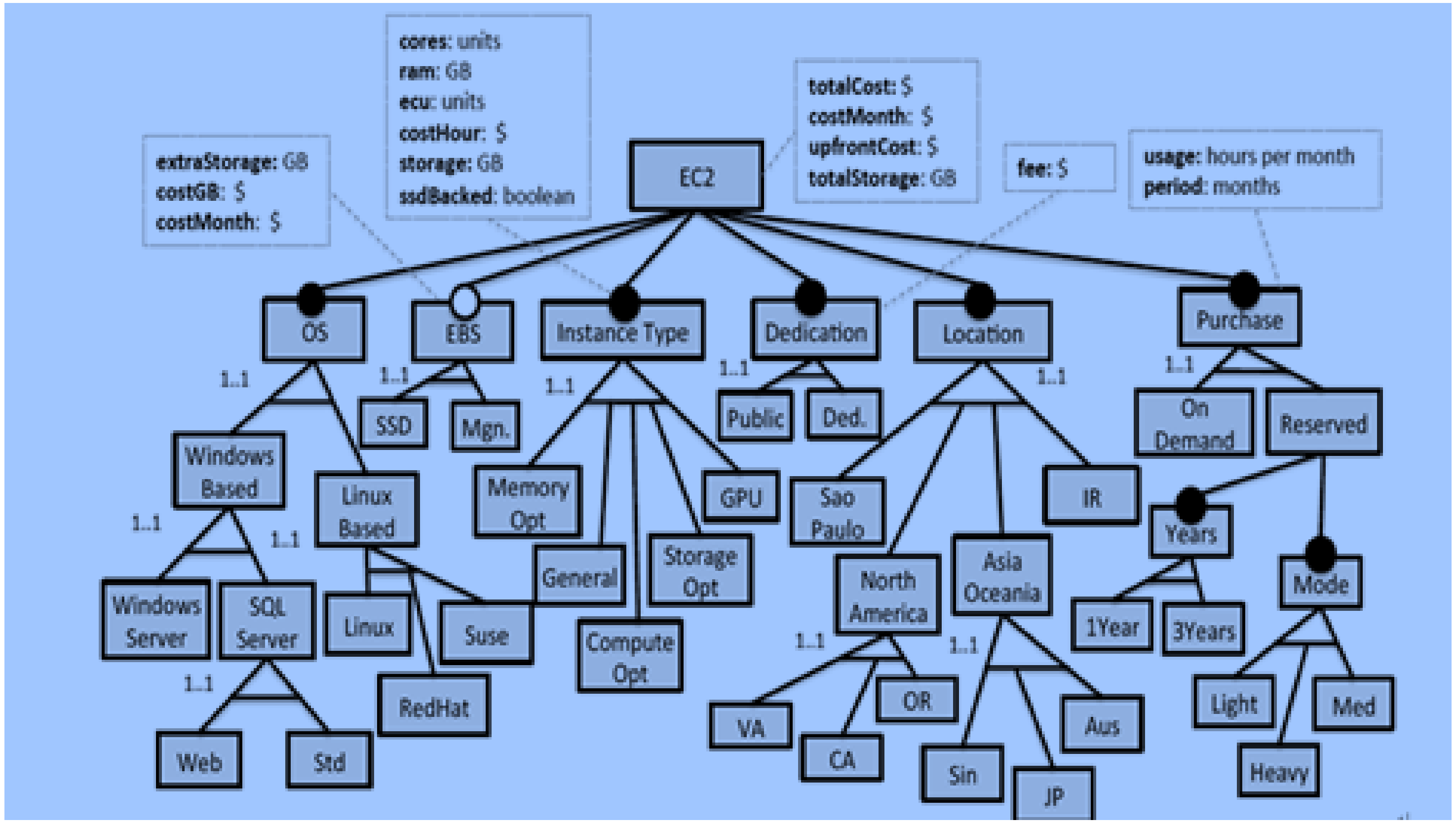
3.两个随机大特征项模型

2.研究背景



BerkeleyDB（文件数据库模型）

2.研究背景



AmazonEC2（亚马逊弹性计算服务模型）

10

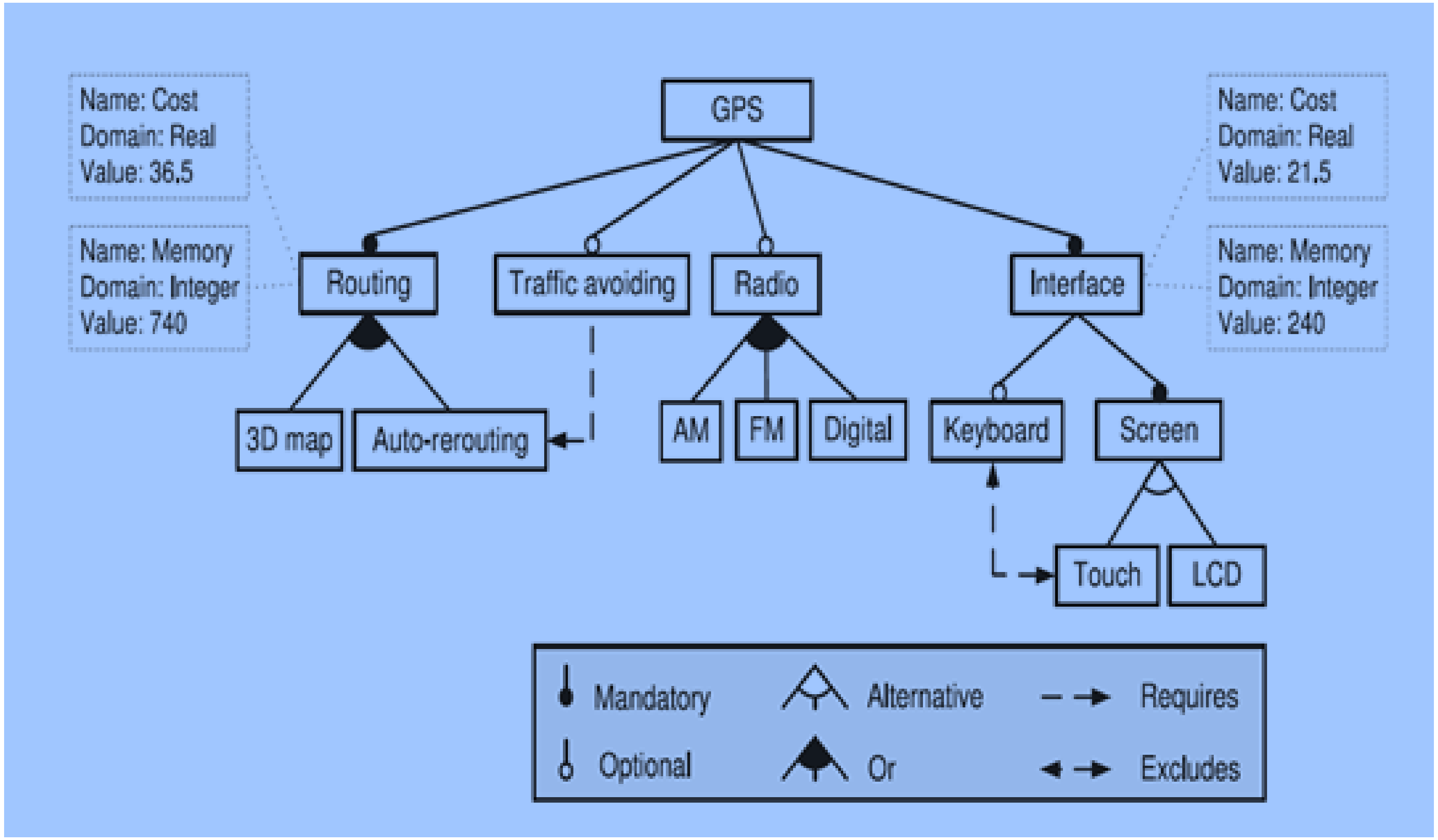
2.研究背景

软件产品	特征个数	Mandatory 节点个数	Optional 节点个数	Group 节点个数
BerkeleyDB	13	2	8	2
ERS	36	10	0	25
E-Shopping	290	66	74	149
WebPortal	43	8	17	17
AmazonEC2	79	6	1	71
Drupal	48	8	39	0
Random-5000	5000	1655	1685	1659
Random-10000	10000	3209	3343	1087

SPL	#Features	#CTC	#Attributes per Feature	#Products
BerkeleyDB	13	0	4	512
ERS	36	0	7	6,912
WebPortal	43	6	4	$2.1 \cdot 10^6$
E-shop	290	21	4	$5.02 \cdot 10^{49}$
Drupal	48	21	22	$2.09 \cdot 10^9$
AmazonEC2	79	0	17	66,528
Randomly generated	10,000	0	4	$\leq 2^{10000} - 1$

特征模型基本数据

3.研究内容

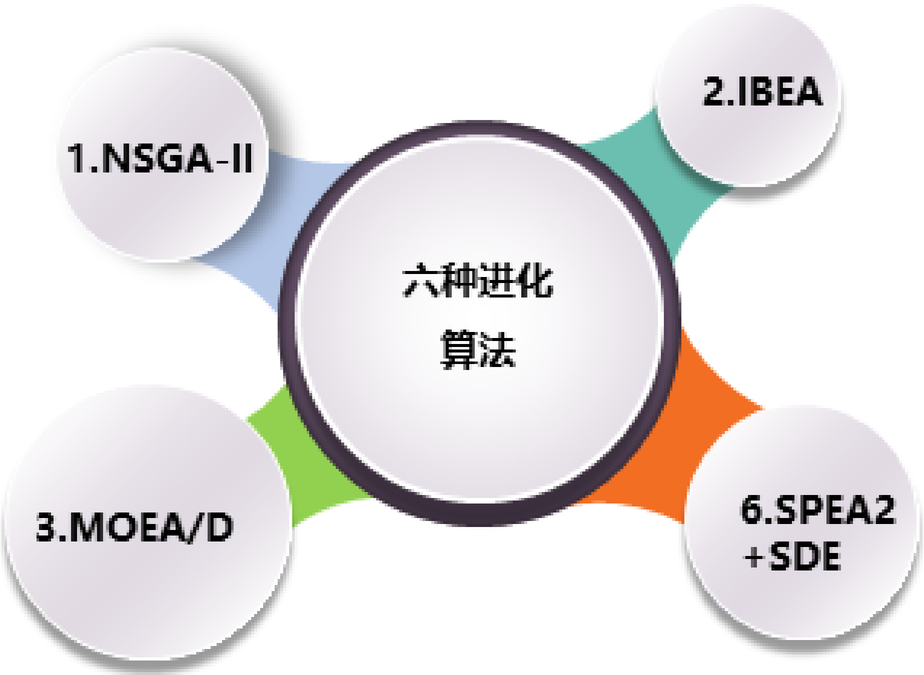


A sample feature model

13

3.研究内容

论文中考虑了六种进化算法



3.MOEA/D-WS
4.MOEA/D-TCH
5.MOEA/D-PBI

其中只有NSGAII和IBEA出现在SPL选择领域的文章中。

3.研究内容

一个特征模型具有许多限制条件，具体限制条件包括：

1 根节点规则

2 子父节点规则

3 Mandatory节点规则

4 Group节点规则

5 等价互斥规则

3.研究内容

根据应用要求不同，其他适应度目标函数还可能包括：

1）特征个数
2）特征被用性
3）已知缺陷个数
4）开销
5）代码行数
6）圈复杂度
7）测试断言个数
8）特征已配置次数
9）开发人员数

3.研究内容

以前的SPL特征模型选择研究中，所有适应度目标被同等对待。

APPROACH 1. (n + 1) approach

APPROACH 2. (1 + n) approach

3.研究内容

Table I. Features Included in Each Encoding on the GPS SPL

Feature	Encoding: Direct	Core	Hierarchical	Novel
GPS	✓		✓	
Routing	✓		✓	
3D map	✓	✓	✓	✓
Auto-rerouting	✓	✓	✓	✓
Traffic avoiding	✓	✓	✓	✓
Radio	✓	✓		
FM	✓	✓	✓	✓
AM	✓	✓	✓	✓
Digital	✓	✓	✓	✓
Interface	✓		✓	
Keyboard	✓	✓	✓	✓
Screen	✓		✓	
Touch	✓	✓	✓	✓
LCD	✓	✓	✓	✓

4.实验设计

通用五目标设计Sayyad et al. [2013d] and Henard et al. [2015]:

(1)Correctness （minimized）
(2) Richness of features （maximized）
(3) Features that were used before （maximized）
(4) Known defects （minimized）
(5) Cost（minimized）

4.实验设计

■ 对应特征模型Drupal,使用八目标的真实属性值

(1)Correctness （minimized）
(2) Richness of features （maximized）
(3) Number of lines of code （minimized）
(4) Cyclomatic complexity （minimized）
(5) Test assertions （maximized）
(6) Number of installations that contain the feature （maximized）
(7) Number of developers（minimized）
(8) Number of changes （minimized）

4.实验设计

- 对应特征模型型**Amazon,**使用特殊的八目标

(1) Correctness（minimized）
(2) Richness of features （maximized）
(3) EC2.costMonth （minimized）
(4) Instance.cores （maximized）
(5) Instance.ecu （maximized）
(6) Instance.ram （maximized）
(7) Instance.costHour: （minimized）
(8) Instance.ssdBacked （maximized）

4.实验设计

评价指标

(1)HV

(2)VN

(3)VR

西北工业大学

5.实验结果

SIP方法中的所有代码可以从以下网址获得

https://dx.doi.org/10.17633/rd.brunel.2115802

所有实验的数据从以下网址中获得

https://dx.doi.org/10.17633/rd.brunel.2115490.v1

23

5.实验结果

■ 统计测试方法（SPSS）

1) 采用post-hoc Kruskal-Wallis test(事后KW检测方法，是一种事后非参数检验方法)方法对六个运算结果进行两两比较，并通过Bonferroni校正获取两个算法之间的p-value。

2) 通过使用Mann-Whitney U test(曼-惠特尼秩和检验)方法来求出算法两两之间的效应量ES。

5.实验结果

RESEARCH QUESTION 1.

How does the encoding affect the performance of the search techniques?

The Importance of the Encoding.

5.实验结果

RESEARCH QUESTION 2.

Does the use of the 1 + n approach make it easier for the search techniques to find valid products?

The Value of Using the 1 + n Approach

26

5.实验结果

RESEARCH QUESTION 3.

Are particular EMO algorithms more effective than others?

RESEARCH QUESTION 4.

Does the relative performance differ when realistic attribute values are used?

5.实验结果

RESEARCH QUESTION 5.

Is the relative performance similar when using a larger randomly generated feature model?

RESEARCH QUESTION 6.

How does the number of valid products returned by search differ between approaches?

RESEARCH QUESTION 7.

How does the execution time differ between approaches?

5.实验结果

Table III. E-Shop, 50,000 Evaluations, Direct Encoding

Algorithm	(n + 1) Approach			1 + n Approach		
	HV	VN (/30)	VR	HV	VN (/30)	VR
NSGA-II	0.000000	0	0.00%	0.136545	13	100%
IBEA	0.000000	0	0.0%	0.169191	16	100%
MOEA/D-WS	0.016184	26	21.42%	0.184810	5	100%
MOEA/D-TCH	0.000000	0	0.00%	0.199697	1	100%
MOEA/D-PBI	0.018815	4	10.50%	0.166157	5	100%
SPEA2+SDE	0.000000	0	0.00%	0.144341	15	100%

Table IV. E-Shop, 50,000 Evaluations, Hierarchical Encoding

Algorithm	(n + 1) Approach			1 + n Approach		
	HV	VN (/30)	VR	HV	VN (/30)	VR
NSGA-II	0.003545	26	2.50%	0.202163	3	100%
IBEA	0.236265	7	42.20%	0.212066	3	100%
MOEA/D-WS	0.021644	24	27.64%	0.242515	2	100%
MOEA/D-TCH	0.000000	0	0.00%	0.230922	1	100%
MOEA/D-PBI	0.000000	0	0.00%	0.196956	2	100%
SPEA2+SDE	0.000000	0	0.00%	0.187481	3	100%

Table V. E-Shop, 50,000 Evaluations, Core Encoding

Algorithm	(n + 1) Approach			1 + n Approach		
	HV	VN (/30)	VR	HV	VN (/30)	VR
NSGA-II	0.003343	28	2.07%	0.149158	30	100%
IBEA	0.267410	30	33.91%	0.175422	30	100%
MOEA/D-WS	0.074223	30	26.74%	0.207303	30	100%
MOEA/D-TCH	0.000000	0	0.00%	0.190666	30	100%
MOEA/D-PBI	0.070765	30	30.62%	0.151068	30	100%
SPEA2+SDE	0.000000	0	0.00%	0.152678	30	100%

5.实验结果

Table XXIV. Best-Performing EMO Algorithm, Novel Encoding, and 1 + n

Subject Model	Best-Performing MOEA	Superior to All Except
E-Shop 50,000	MOEA/D-TCH	MOEA/D-WS, MOEA/D-PBI
E-Shop 500,000	MOEA/D-WS	MOEA/D-TCH
WebPortal	IBEA	MOEA/D-PBI
Amazon	SPEA2+SDE	NSGA-II, IBEA, MOEA/D-WS, MOEA/D-TCH
Berkeley	SPEA2+SDE	IBEA
Drupal	IBEA	MOEA/D-PBI
ERS	NSGA-II	SPEA2+SDE, MOEA/D-TCH, MOEA/D-WS
Drupal, real	SPEA2+SDE	IBEA
Amazon, realistic	SPEA2+SDE	MOEA/D-WS
Larger model	SPEA2+SDE	NSGA-II, IBEA

32

6.结论与展望

01 更多模型的实验与验证

02 工业界推广

33