



CSBSE研讨会

基于搜索变换的软件工程

Transformed Search Based Software Engineering

江 贺

jianghe@dlut.edu.cn

大连理工大学 软件学院

2015年6月13日



汇报提纲

- 基于搜索的软件工程：起源及内容
- 基于搜索的软件工程：发展现状
- 基于搜索变换的软件工程
- NRP问题的骨架导向多级归约算法
- 基于空间平滑的Memetic算法
- 相关工作

基于搜索的软件工程：起源及内容



◆ 2001年英国Mark Harman 教授等正式提出基于搜索的软件工程的概念(IST, 2001, 43(14): 833 – 839)

基于搜索的软件工程是指将软件工程问题转化为组合优化问题, 并采用以演化算法、模拟退火算法、禁忌搜索算法等为代表的现代启发式搜索算法来求解。

• 面向软件工程全生命周期

• 灵活应用各种启发式算法

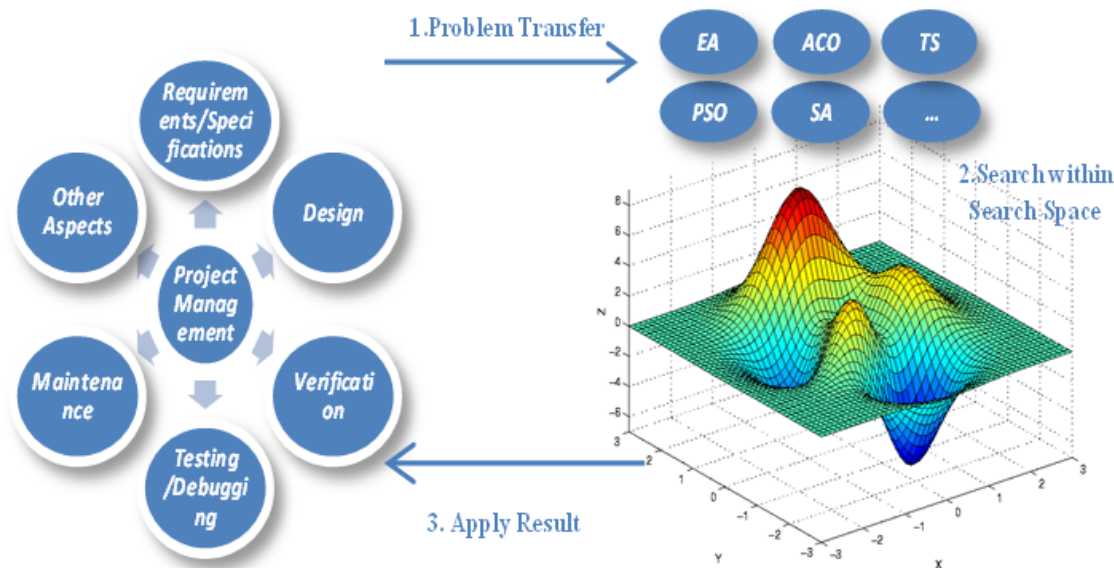


图1. 基于搜索的软件工程的研究流程示意图



汇报提纲

- 基于搜索的软件工程：起源及内容
- 基于搜索的软件工程：发展现状
- 基于搜索变换的软件工程
- NRP问题的骨架导向多级归约算法
- 基于空间平滑的Memetic算法
- 相关工作



基于搜索的软件工程：发展现状

■ 基于搜索的软件工程国内外研究现状

• 专刊

– JSME 2007, IEEE TSE 2010, IST 2010, SPE 2011, EMSE 2011, JSS 2013,

• 报告、讲座

– ICST 2015, SPLC 2014, WCRE2013, GECCO 2013, ASE 2012, ESEM 2012, TASE 2012,

• 国际工业界

– 微软 (ICTSS 2010, 142–157)、爱立信 (GECCO 2010, 1357–1364)、IBM (UCL Tech Rep. RN/11/07)、谷歌 (STVR 2012, 22 (2) : 67–120)、华为

• 国内学者工作

– 北京大学梅宏、张路教授课题组 (ISSTA, 2009, 213–224; ASE, 2014, 701–712)
– 南京大学聂长海教授课题组 (ACM Computing Survey)
– 中山大学张军教授课题组 (TSE, 2013, 39 (1) : 1–17)
– 国防科学技术大学毛晓光教授课题组 (ICSE, 2014, 254–265)
– 北京化工大学李征教授课题组 (TSE, 2007, 33 (4) : 225–237)
– 中国矿业大学巩敦卫教授课题组 (ICSE2014, ASEJ)



汇报提纲

- 基于搜索的软件工程：起源及内容
- 基于搜索的软件工程：发展现状
- 基于搜索变换的软件工程
- NRP问题的骨架导向多级归约算法
- 基于空间平滑的Memetic算法
- 相关工作

基于搜索空间变换的软件工程

基于搜索空间变换的软件工程首先针对软件工程问题解形成的搜索空间进行变换，包括：1) 空间归约，以获得规模更小、全局最优解更易定位的新搜索空间；2) 空间平滑，得到一系列逐次光滑的搜索空间，其次再进行高效搜索获得高质量的解。

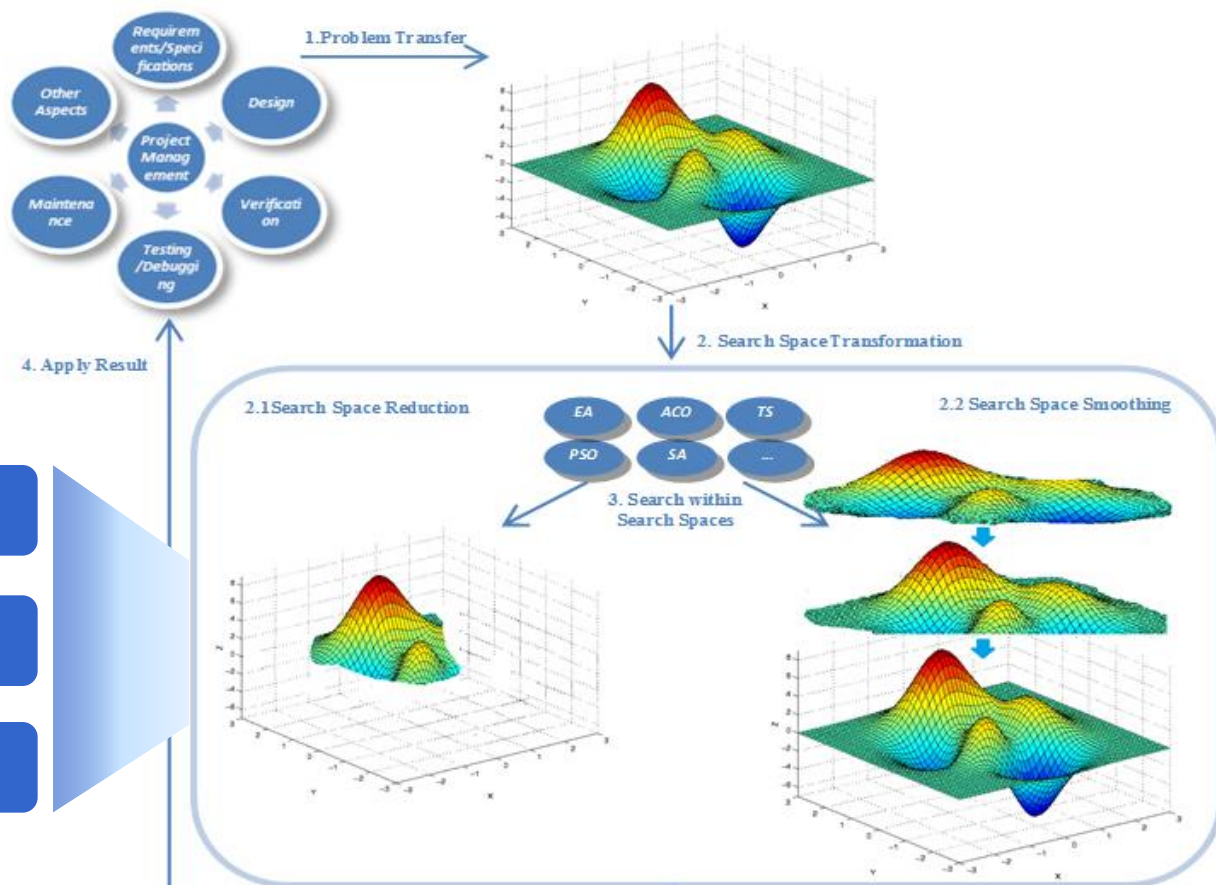
由“大海捞针”改为
“池塘捞针”

以骨架为手段，设计
搜索空间归约算法。

1. 骨架计算复杂性分析

2. 骨架获取

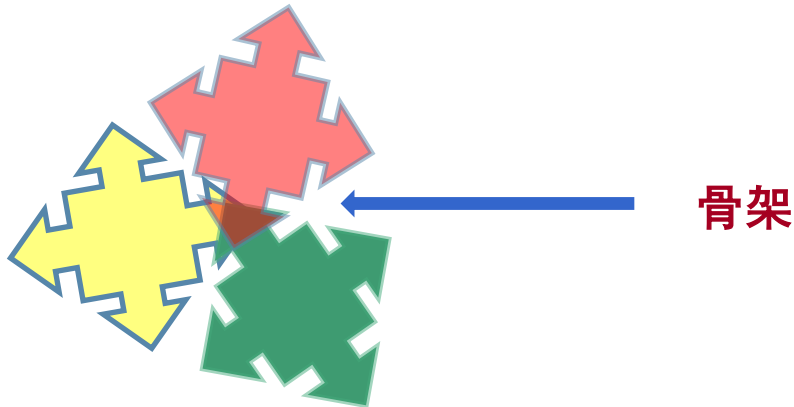
3. 搜索空间归约算法设计



搜索空间归约导向的SBSE

以骨架为手段，设计搜索空间归约算法。

骨架：搜索空间中所有的全局最优解的共同部分。



相关科学问题

•骨架计算复杂度

– 是否存在多项式时间算法可以获得骨架？

•骨架逼近

– 如何获得近似骨架？

•基于骨架的搜索空间归约

– 如何利用骨架（近似骨架）对搜索空间进行归约？

搜索空间归约导向的SBSE

搜索空间变换: 搜索空间归约框架

Algorithm 1: Search Space Reduction

Input: search space Π , search algorithms A , maximum number α of reduction levels, a set of solutions Γ

Output: best solution

```
1 begin
2   for  $k = 1$  to  $\alpha$  do
3     Obtain a set of solutions  $\Gamma_k$  by  $A$  in  $\Pi_k$ 
4     Calculate high quality part  $\Gamma_k'$  of  $\Gamma_k$ 
5     Reduce the search space to  $\Pi_{k+1}$  by  $\Gamma_k'$ 
6   end
7   Obtain a local optimal solution  $\Gamma_{\alpha+1}'$  in the final reduced search space  $\Pi_{\alpha+1}$  by  $A$ 
8   for  $k = \alpha$  to 1 do
9     Refine the solution with  $\Gamma_{k+1}'$  and  $\Gamma_k'$  in level  $k$ 
10  end
11  return the best solution achieved
12 end
```

获得骨架

搜索空间归约

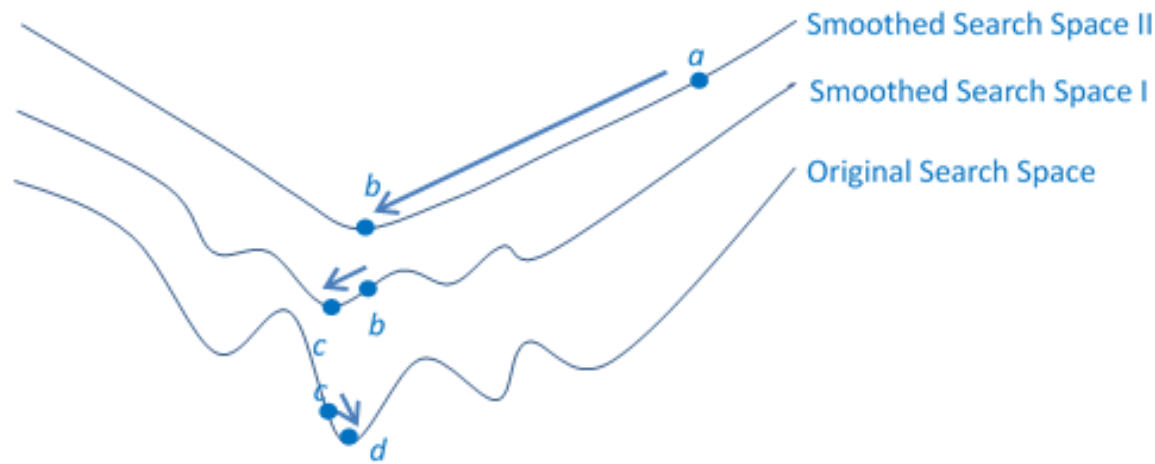
骨架整合

生成最优解

搜索平滑导向的SBSE

以平滑策略为手段，设计空间平滑算法。

平滑：对搜索空间中崎岖的部分进行归一化。



相关科学问题

•平滑策略设计

— 如何针对软件工程的问题空间设计平滑策略？

•平滑时间分析

— 如何权衡平滑时间与解的质量？

搜索平滑导向的SBSE

搜索空间变换: 搜索空间平滑框架

Algorithm 2: Search Space Smoothing

Input: search space Π , search algorithms A , maximum number β of smoothing levels, a set of solutions Γ

Output: best solution

```
1 begin
2   Generate a smoothed search space  $\Pi_0$ 
2   Generate initial solutions  $\Gamma_0$  in  $\Pi_0$ 
3   for  $k = 1$  to  $\beta$  do
4     Tune the search space to  $\Pi_k$ , towards the original, rugged space.
5     Assign the current best solutions  $\Gamma_{k-1}$  as the initial solution
6     Apply  $A$  with  $\Gamma_{k-1}$  in  $\Pi_k$  to get the current best solutions  $\Gamma_k$ 
7   end
8   return the best solution achieved
9 end
```

获取初始解与解空间

迭代进行空间平滑

执行搜索算法

生成最优解



汇报提纲

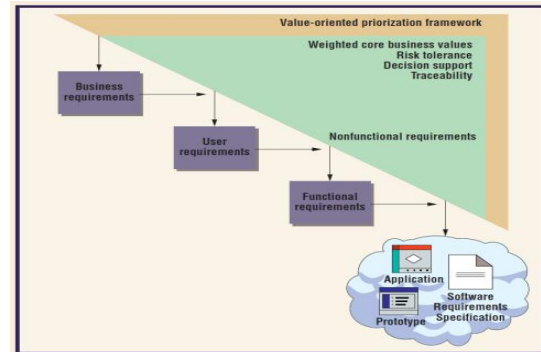
- 基于搜索的软件工程：起源及内容
- 基于搜索的软件工程：发展现状
- 基于搜索变换的软件工程
- NRP问题的骨架导向多级归约算法
- 基于空间平滑的Memetic算法
- 相关工作

NRP问题的骨架导向多级归约算法

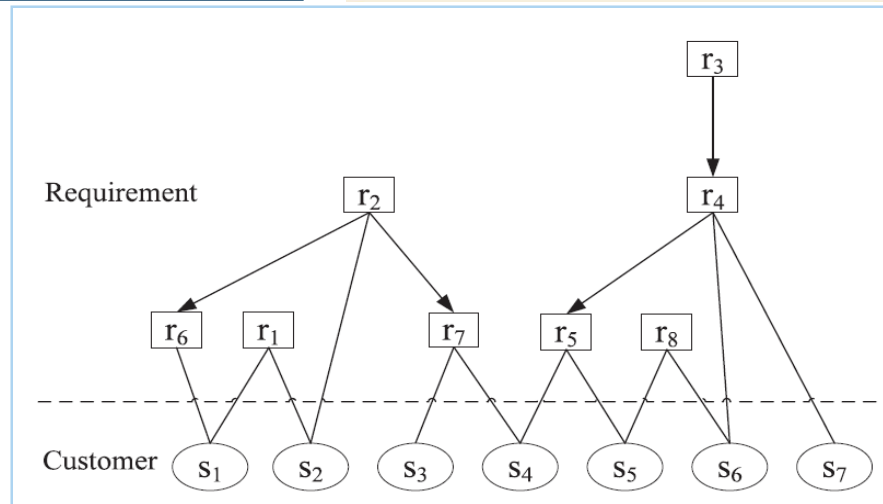
软件下一版本问题

Given a directed acyclic requirements dependency graph $G = (R, E)$, each customer $s_i \in S$ directly requests a set of requirements R_i . The profit of s_i is $w_i \in W$ and the cost of requirement $r_j \in R$ is $c_j \in C$. A predefined budget bound is b .

The goal of the NRP is to find an optimal solution X^* , to maximize $\omega(X)$, subject to $\text{cost}(X) \leq b$.



- 在发布下一版本软件之前
 - ◆ 每个候选需求具有开发代价
 - ◆ 需求间可能存在依赖关系
 - ◆ 每个用户可以通过实现需求来满足
- 如何通过实现需求子集，在保证开发成本约束前提下最大化用户收益？

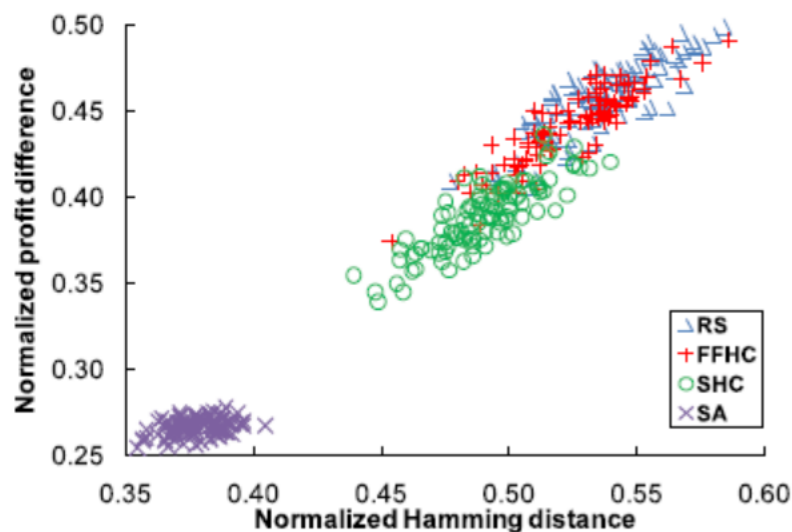
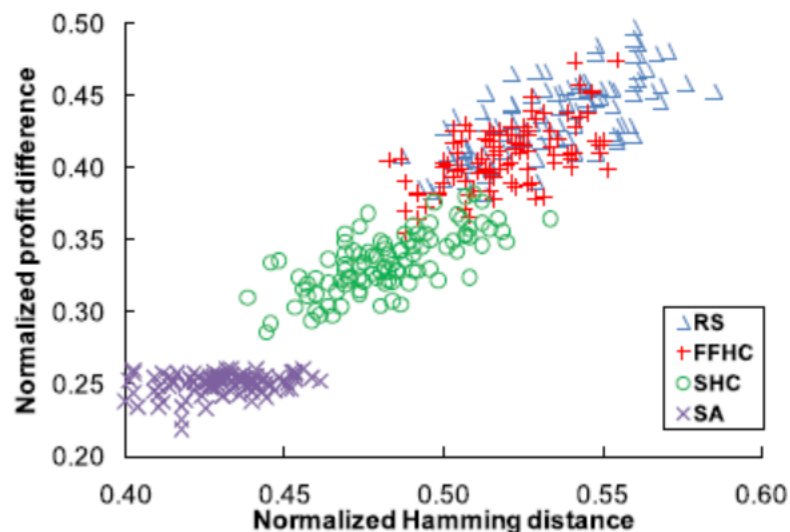


NRP问题的骨架导向多级归约算法

软件下一版本问题: **骨架计算复杂性。**

获取骨架的计算复杂性为NP-Hard 的。

软件下一版本问题: **骨架逼近。**

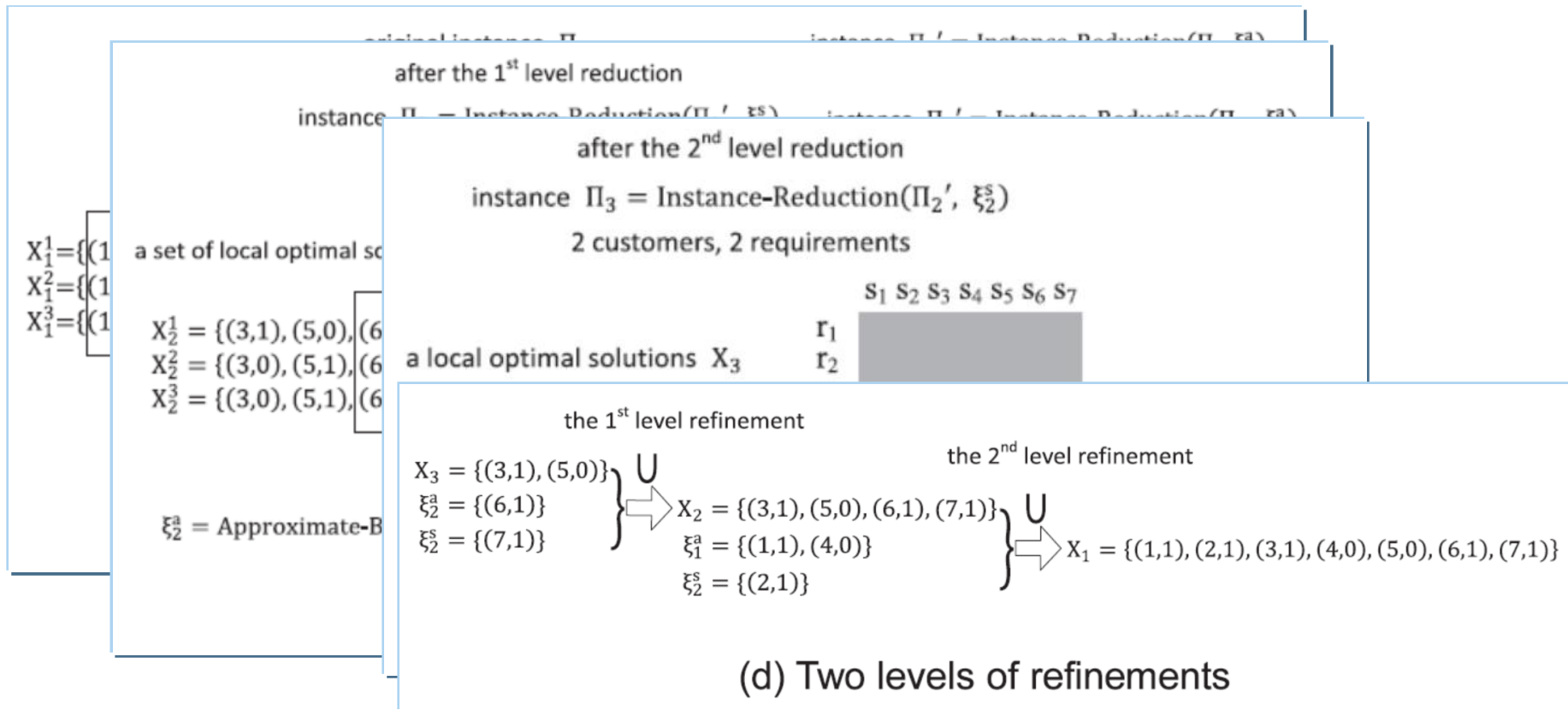


局部最优解可以模拟全局最优解，用局部最优解交集逼近骨架。



NRP问题的骨架导向多级归约算法

软件下一版本问题: 基于骨架的搜索空间归约





NRP问题的骨架导向多级归约算法

软件下一版本问题: 基于骨架的搜索空间归约

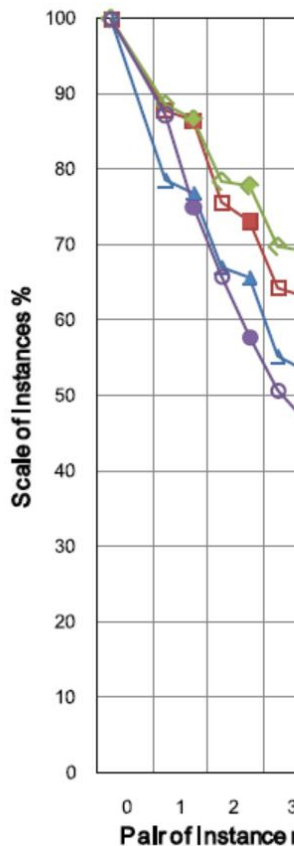
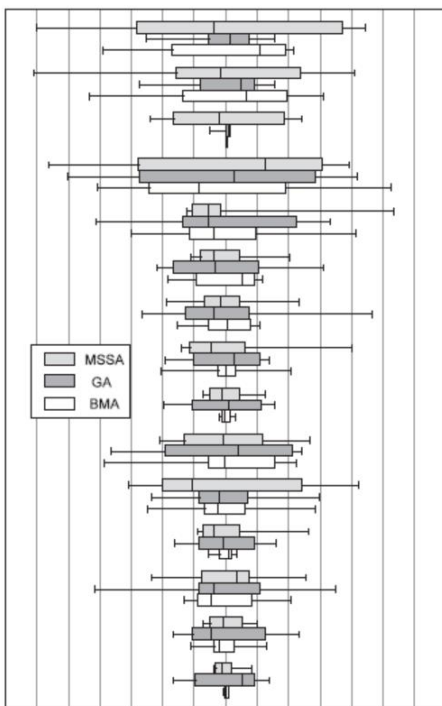


TABLE 4 Total Scales Reduced by 12 Instance Reductions				
Instance name	nrp-2-0.5	nrp-3-0.5	nrp-4-0.5	nrp-5-0.5
Original scale	500	500	750	1000
Scale reduced by approximate backbone%	88.4	88.6	79.9	60.8
Scale reduced by soft backbone%	11.2	7.6	7.2	32.4
Sum of scale reduced%	99.6	96.2	87.1	93.2

NRP问题的骨架导向多级归约算法

软件下一版本问题: 基于骨架的搜索空间归约

TABLE 8
Performance for MSSA, GA, and BMA on 15 Classic Instances

Instance		MSSA			GA			BMA					Profit distribution %												
Name	Bound	Best	Average	Time	Best	Average	Time	Best	Average	Time	MSSA%	GA%													
nrp-1-0.3	257	998	976.5	108.65	1187	1178.1	85.63	1201	1188.3	52.68	21.69	0.87													
nrp-1-0.5	429	1536	1505.2	98.93	1820	1806.1	99.22	1824	1796.2	55.91	19.33	-0.55													
nrp-1-0.7	600	2301	2273.6	91.70	2507	2505.4	79.19	2507	2507.0	34.51	10.27	0.06													
nrp-2-0.3	1514	3220	3158.3	320.76	2794	2737.0	654.23	4726	4605.6	246.14	45.83	68.27													
nrp-2-0.5	2524	5229	5094.1	288.70	5363	5276.4	891.55	7566	7414.1	280.87	45.54	40.51													
nrp-2-0.7	3534	8002	7922.6	255.52	9018	8881.1	911.55	10987	10924.7	277.47	37.89	23.01													
nrp-3-0.3	2661	5147	5088.8	461.21	5851	5719.0	910.99	7123	7086.3	436.90	39.25	23.91													
nrp-3-0.5	4435	8725	8553.4	420.66	9639	9574.2	542.22	10897	10787.2	438.80	26.12	12.67													
nrp-3-0.7	6209	13600	13518.2	489.79	12454	12360.7	265.23	14180	14159.2	215.90	4.74	14.55													
nrp-4-0.3	6648	6797	6708.4	1153.34	6675	6595.7	1849.15	9818	9710.5	854.48	44.75	47.22													
nrp-4-0.5	11081	11355	11120.6	1017.39	12781	12595.4	1587.22	15025	14815.5	907.03	33.23	17.63													
nrp-4-0.7	15513	19077	18830.1	1104.26	17327	17189.9	549.71	20853	20819.7	672.60	10.57	21.12													
nrp-5-0.3	1198	11421	11279.2	502.87	10689	10507.0	3069.26	17200	17026.9	475.85	50.96	62.05													
nrp-5-0.5	1996	17843	17756.6	472.48	18950	18732.9	1696.38	24240	24087.5	459.05	35.65	28.58													
nrp-5-0.7	2794	28347	28232.5	628.25	22174	22026.5	376.57	28909	28894.2	171.70	2.34	31.18													



汇报提纲

- 基于搜索的软件工程：起源及内容
- 基于搜索的软件工程：发展现状
- 基于搜索变换的软件工程
- NRP问题的骨架导向多级归约算法
- 基于空间平滑的Memetic算法
- 相关工作

基于空间平滑的Memetic算法

软件下一版本问题: **空间平滑基本策略。**

幂指数平滑, 倒数平滑, S形曲线平滑 (Sigmoidal Smoothing) 等。

软件下一版本问题: **平滑方案。**

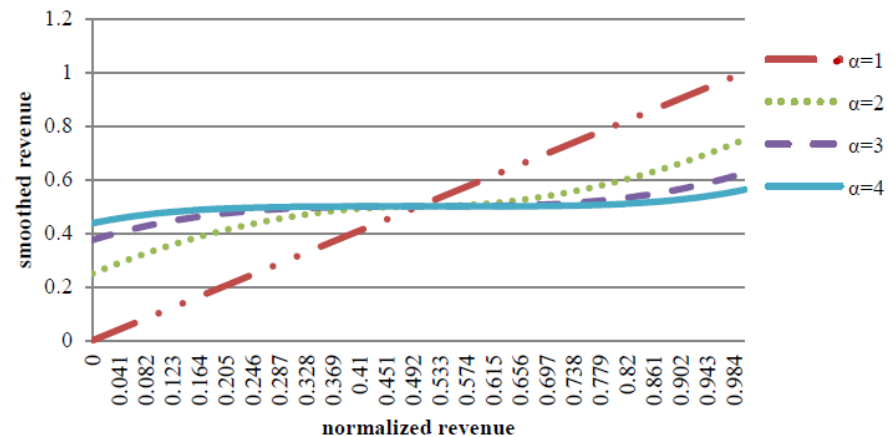
平滑公式^[1]

$$w_i(\alpha) = \begin{cases} \bar{w} + (w'_i - \bar{w})^\alpha, & w'_i \geq \bar{w} \\ \bar{w} - (\bar{w} - w'_i)^\alpha, & w'_i < \bar{w} \end{cases} \quad (1)$$

■ 其中

- ◆ w'_i 为第 i 个客户收益
- ◆ \bar{w} 为所有客户平均收益
- ◆ α 为平滑因子

平滑因子影响



通过控制平滑因子细粒度的调整搜索空间地貌, 优化初始解。

基于空间平滑的Memetic算法

软件下一版本问题：面向NRP的搜索空间平滑

Algorithm: Search Space Smoothing based Memetic Algorithm (SSS-MA)

Input: maximum iterator $nIter$; population size $nPop$, elitism rate $eRate$, mutation rate $mRate$

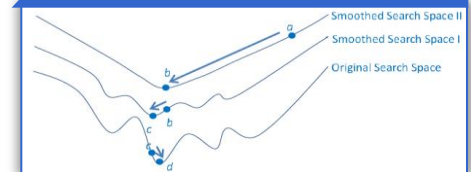
Output: best solution achieved

```

1 begin
2   initialization
3   for  $i \leftarrow 1$  to  $nIter$  do
4      $\alpha \leftarrow \left\lfloor 6 - \frac{5 \times i}{nIter} \right\rfloor$ 
5     Modify instance variables with Eq. 1
6     for  $nPop \times (1 - eRate)$  do
7       Randomly select two individuals as parents
8       Apply uniform crossover
9       Apply bit-flipping mutation over the offspring
10      Apply hill climbing over the offspring
11    end
12    Apply elitism selection
13  end
14  return best solution achieved
15 end
  
```

获取初始解

搜索空间平滑



Memetic算法 (MA)

生成最优解

基于空间平滑的Memetic算法

软件下一版本问题：面向NRP的搜索空间平滑

•参数设置

•小规模NRP实例集

•大规模NRP实例集

实验结果：空间平滑能够提升搜索算法的性能，并在6个NRP实例集上超过已知最好结果

Instance	BMA		MA		SSS-MA		
	Best		Best		Best		
	Average		Average		Average		
nrp-e1-0.3	7572	7396	7344.5	12.95	7539	7460.8	20.63
nrp-e1-0.5	10664	10607	10555	15.16	10740	10676.3	22.90

Instance	BMA		MA		SSS-MA		
	Best		Best		Best		
	Average		Average		Average		
nrp1-0.3	1201	1204	1191.1	1.22	1200	1189.2	2.59
nrp1-0.5	1824	1836	1812.8	1.38	1834	1784.2	2.62
nrp1-0.7	2507	2507	2507	1.15	2507	2507	2.42
nrp2-0.3	4726	4007	3927.7	5.57	4365	4179.8	13.53
nrp2-0.5	7566	7034	6840.7	7.16	7353	7202.2	15.79
nrp2-0.7	10987	10585	10419	7.85	10683	10589.5	16.56
nrp3-0.3	7123	6846	6756	7.25	7001	6894.2	14.70
nrp3-0.5	10897	10566	10522.2	7.95	10758	10644.6	15.75
nrp3-0.7	14180	13867	13819.5	7.78	13990	13953	15.58
nrp4-0.3	9818	8950	8841.6	17.96	9164	9003.8	29.72
nrp4-0.5	15025	14609	14457.6	20.22	14794	14613.6	32.95
nrp4-0.7	20853	19996	19906.6	22.60	20205	20117.4	35.76
nrp5-0.3	17200	14873	14564.3	19.33	15417	15165.7	40.68
nrp5-0.5	24240	22409	22204.5	14.95	22785	22616.3	34.89
nrp5-0.7	28909	27494	27283.6	10.41	27854	27761.8	28.75



汇报提纲

- 基于搜索的软件工程：起源及内容
- 基于搜索的软件工程：发展现状
- 基于搜索变换的软件工程
- NRP问题的骨架导向多级归约算法
- 基于空间平滑的Memetic算法
- 相关工作



相关工作

Jifeng Xuan, He Jiang(*), Yan Hu, Zhilei Ren, Weiqin Zhou, Towards Effective Bug Triage with Software Data Reduction Techniques, IEEE Transactions on Knowledge and Data Engineering, 2015, 27 (1) : 264-280

Zhilei Ren, He Jiang(*), Jifeng Xuan, Yan Hu, Zhongxuan Luo, New Insights Into Diversification of Hyper-Heuristics, IEEE Transactions on Cybernetics, 2014, 44 (10) : 1747-1761

Xuan, Jifeng, Jiang, He(*), Ren, Zhilei, Luo, zhongxuan, Solving the Large Scale Next Release Problem with a Backbone Based Multilevel Algorithm, IEEE Transactions on Software Engineering, 2012, 38(5): 1195-1212

Ren, Zhilei, Jiang, He(*), Xuan, Jifeng, Luo, Zhongxuan, An Accelerated-Limit-Crossing-Based Multilevel Algorithm for the p-Median Problem, IEEE Transactions on Systems, Man, and Cybernetics. Part B-Cybernetics, 2012, 42 (4) : 1187-1202

Xuan, Jifeng, Jiang, He(*), Ren, Zhilei, Zou, Weiqin, Developer prioritization in bug repositories, 34th International Conference on Software Engineering, 2012.7.2-2012.7.9

Jiang, He(*), Ren, Zhilei, Li, Xiaochen, Lai, Xiaochen, Transformed Search Based Software Engineering: A New Paradigm of SBSE, SSBSE, 2015

基于搜索变换的软件工程网址: oscar-lab.org/tsbse/

Thanks!

