

融合词、句层级信息的抽取式摘要优化框架

林心宜^{1,2} 严睿^{1,†} 赵东岩¹

1. 北京大学计算机科学技术研究所, 北京 100080; 2. 北京大学信息科学技术学院, 北京 100871;

† 通信作者, E-mail: ruiyan@pku.edu.cn

摘要 提出一个混合的抽取式摘要优化框架, 在优化单词层级信息的同时, 将句子层级信息作为优化约束。在约束条件下, 该优化框架迭代地进行摘要文本中单元的替换, 得到不断逼近目标函数的最优解。与传统方法对比, 该框架在 DUC 数据集上获得 ROUGE 评测的高分, 证明了该框架的有效性。

关键词 抽取式摘要生成; 词层级信息; 句层级信息; 混合迭代优化框架

中图分类号 TP391

A Hybrid Optimization Framework Fusing Word- and Sentence-Level Information for Extractive Summarization

LIN Xinyi^{1,2}, YAN Rui^{1,†}, ZHAO Dongyan¹

1. Institute of Computer Science and Technology, Peking University, Beijing 100080; 2. School of Electronic Engineering and Computer Science, Peking University, Beijing 100871; † Corresponding author, E-mail: ruiyan@pku.edu.cn

Abstract In order to fuse word-level and sentence-level information from different semantic spaces, the authors propose a hybrid optimization framework to optimize word-level information while simultaneously incorporate sentence-level information as constraints. The optimization is conducted by iterative unit substitutions. The performance on DUC benchmark datasets demonstrates the effectiveness of proposed framework in terms of ROUGE evaluation.

Key words extractive summarization; word-level information; sentence-level information; hybrid optimization framework

多文档摘要生成工作是自然语言处理领域的核心任务之一, 十几年来吸引了众多的研究者投入该任务中, 其目标是利用计算机, 自动为一组文档生成一段限定字数的精简、浓缩的摘要, 辅助读者快速领悟文档集合的要点, 捕捉重要信息。

完成这个任务的一个很自然的想法, 是将文档拆解成一组语义单元, 然后根据预设的标准, 从中抽取单元集合的子集进行重组, 形成摘要。组成文档最常用的语义单元是句子和单词, 所以摘要工作通常也是基于这两个粒度。目前, 大多数研究主要关注句子层级, 例如质心法(centroid-based method)

(MEAD^[1]和 NeATS^[2])、基于图排序法(graph-based method) (PageRank^[3-4], HITS^[5]和 ClusterCMRW^[6])和子模方法(submodular functions)^[7]。基于单词级别的方法主要是整数线性规划法(integer linear programming process)^[8]。Gillick等^[8]将单词视为赋有不同权重的语义单元, 将摘要工作抽象为一个整数线性规划问题, 在长度限制内, 抽取句子集合形成摘要, 使之包含的单词总权重最大。然而, 研究者很少将句子级别和单词级别的信息进行融合, 主要是因为单词和句子来自于不同的语义空间, 在大部分情况下, 这两个空间非线性相关。Xie 等^[9]在融

合句子和单词信息时,对句子和单词的分数分别进行归一化,然后将两个分数进行线性叠加。尽管这个方法效果欠佳,但不失为融合词句信息的一个简单直接的解决思路。

本文提出一个混合词层级和句层级信息的迭代优化摘要框架。之所以要融合词层级和句层级的信息,是因为词层级和句层级从不同角度满足了摘要的需求,并且功能存在互补。句子拥有完整的语法结构,其中大部分字词的作用是连接,使之形成语法上可读的句子,而与语义内容无关。相比之下,单词粒度小,表达的含义更凝练、精准,以单词为粒度能够更加充分地利用有限的摘要长度,包含更多的重要文档信息。虽然单词体积小,以它为粒度能够提高摘要的有效信息率,但摘要还需考虑其可读性、连续性和逻辑完整性。简单地将一个个单词实体拼接起来,无法形成可读的文段。句子的优势在于,它是更完整的语义单元,包含必需的语法结构、语义连接以及实体间的逻辑关系,保证了可读性,并且保留了一定的连续性、时间的顺序性以及逻辑的完整性。因此,更优的策略应该是结合两个层级的优势,在抽取重要句子的同时,使之包含尽可能多的重要单词。

相比于直接将句子和单词的权重进行线性组合的简单方法,我们考虑到两个语义空间的数学关系不明确,为避免直接刻画句子层级和单词层级信息的关系,使用类似双任务的主任务和副任务框架来实现句子层级和单词层级信息的融合。本文将任务抽象为带依赖关系的背包问题,用迭代框架对该优化问题进行求解。一方面,由于重要程度不同,每个单词有不同的权重得分。任务的主要目标是在长度限制内,迭代更新选择的句子集合,使之包含单词的总得分最高。另一方面,句子作为更高层次的语义单元,也有其句子层级的得分,其特征包括句子的重要性、句子组合的多样性和信息覆盖率等,将这些特征的得分作为迭代替换句子单元时的约束条件。该框架在寻求单词层级信息的最优解的同时,保证了句子层级信息的不断优化,最终得到一个句子层级得分尽可能高的单词层级上的近似最优解。具体而言,我们通过最大化单词级别的分数获得一个初始摘要,这个初始摘要虽然是单词层级上的近似最优解,但它在句子层级上的分数可能较低,所以在满足长度限制的约束下,迭代地进行句子集合的替换,使之满足句子层级评分的约束条件,并

且保证该替换得到的摘要在单词级别上表现出与最优解相差不超过一个阈值,最终获得一个单词级别和句子级别表现双高的平衡点。实验结果表明,该框架在标准数据集 DUC 上使用 ROUGE^[10] 进行评测,获得明显的效果提升。

1 融合词句层级信息的迭代优化框架

1.1 词层级优化

如果将单词视为原子单元,句子则是由多个单词构成的有序元组。首先将任务刻画为单词层级上得分最大化,的一个最优问题的求解,最终目标是选取句子集构成摘要,使得摘要中的单词总分最高。我们虽然以句子为单元进行抽取,但只保证了摘要的可读性,优化目标还是在于使得单词得分最大化,即句子层面信息依旧未考虑进来。

为了获取单词的得分,我们选用最经典的单词评分策略 tf-isf 方法,为每个单词进行打分^[11]。tf 表示单词在句子中的词频; isf 表示逆向句子频率,是出现该单词的句子占有所有句子的比例的倒数再取对数值。给定一个多次出现的单词 w_i (实验中,筛去出现次数小于 5 的单词,以省去不必要的处理),将其对应所有句子的 tf-isf 分数进行求和,并取平均值作为该单词的得分,记做 $\pi(w_i)$ 。

通常,基于单词的摘要有一个基本的假设,即单词间是相互独立的,所以我们可以简单地将单词的得分直接相加,用来评价一个摘要在单词级别上的质量。

$$O(S) = \sum_{w_i \in S} \pi(w_i), \quad (1)$$

其中, S 是摘要, $O(S)$ 是该摘要在单词级别上的得分。为了避免冗余,摘要中重复出现的单词不进行重复计分。我们的目标是抽取包含尽可能多重要单词的句子,使得 $O(S)$ 最大化。我们通过贪心法获得 $\arg\max O(S)$ 的近似解,将其作为迭代优化框架的初始摘要。

1.2 句层级优化

通过贪心法求得的近似解只在单词层级上获得近似最优,却没有考虑句子层级的信息。句子的特征对于形成一篇逻辑关系完整合理、信息覆盖率高、冗余度低的摘要十分关键^[7]。为了加强摘要在句子层级上的表现,需要为摘要设定句子层级上优化的方向。

首先,我们考虑句子层级上最突出的几个特征,

包括信息覆盖度、信息冗余度和代表性。通过对这些特征的刻画,能够近似地评价摘要在句子层级上的表现。

1.2.1 句子集合的信息覆盖度

我们引入 Lin 等^[7]刻画句子集合信息覆盖度的公式(式(2))。将文档集合视为一个图结构,图结构的顶点表示文档集合中的句子,顶点间的边表示句子之间的相似度。我们用 $C(S)$ 刻画摘要 S 和文档剩余句子集合的相似度,表示摘要对原始文档集合信息的覆盖程度,该指标可以表明摘要的内容是否全面。

$$C(S) = \sum_{s_i \in D \setminus S} \min \left\{ \sum_{s_j \in S} \text{sim}(s_i, s_j), \alpha \cdot \sum_{s_k \in D \setminus S_i} \text{sim}(s_i, s_k) \right\}, \quad (2)$$

其中, s_i 和 s_j 分别表示句子 i 和句子 j ; $\text{sim}(s_i, s_j)$ 表示句子 i 与 j 的相似度; $C(S)$ 表示摘要 S 的信息覆盖度,其值等于摘要对剩余句子覆盖程度的叠加。求取最小值是为了避免摘要对不重要的句子过分重视。两个句子的相似度可以用余弦相似度表示。我们将每个句子表示成一个向量,该向量的维数等于所有被计数的单词总数,向量的第 i 维是第 i 个单词在该句子中的 tf-idf 值。通过求解两个句子向量的余弦相似度来表示这两个句子的相似程度。

$$\text{sim}(s, s') = \frac{\vec{s} \cdot \vec{s}'}{|\vec{s}| \times |\vec{s}'|}. \quad (3)$$

1.2.2 句子集合的信息多样性

由于自动摘要生成有长度限制,所以如何减少冗余信息,使得在限定长度内,摘要能尽可能地表达更多信息,提高摘要的有效信息率,成为自动摘要生成的关键问题。本文使用如下公式,定义一个句子集合的多样性:

$$D(S) = - \sum_{s_i \in S} \max_{s_j \in S \setminus \{s_i\}} \text{sim}(s_i, s_j). \quad (4)$$

我们叠加摘要 S 中句子间的相似性,取其负数作为摘要多样性的评分。 $D(S)$ 的评分越高,表明该摘要 S 包含的句子相似度越低,进一步说明摘要 S 的冗余信息少,空间利用率得到提升。

1.2.3 句子集合的代表性

虽然我们将句子视为单词的有序集合,但句子并不是简单的单词拼接,单词间的组合顺序赋予句

子高于单词层级的语义信息。所以,作为更高层级的完整语义单元,句子对文档主旨的贡献力度有另外的评测方法。

PageRank 方法^[3-4]广泛应用于多文档摘要工作对句子重要性的评价。目前对句子重要性进行评估的图方法中,大多数是 PageRank 方法的变种和演化。我们借用 PageRank 方法的思路,为给定的文档集合构建一副无向图,图的顶点是文档中的句子,顶点间的边是句子对的相似性。用 PageRank 算法处理该无向图,并对获得的节点得分进行归一化处理,便可获得每个句子 s 的重要性得分 $\varphi(s)$ 。再对摘要 S 中包含的 $\varphi(s)$ 进行叠加,获得该摘要的总体代表性得分 $I(S)$:

$$I(S) = \sum_{s \in S} \varphi(s). \quad (5)$$

$I(S)$ 越高,说明摘要中包含的“高质量”句子越多,对文档的重要信息刻画得越好。

1.2.4 句子级别的总体得分

为了综合刻画摘要在句子级别上的表现,我们定义公式(6),该公式通过线性组合的方式对句子集合在 3 个特征上的得分进行叠加,将获得的总和作为对句子集合的综合打分(所有的项都进行了归一化处理)。

$$U(S) = \lambda_c \cdot C(S) + \lambda_i \cdot I(S) + \lambda_d \cdot D(S). \quad (6)$$

为简化说明,设置 $\lambda_c + \lambda_i + \lambda_d = 1$ 。

1.3 约束条件下的迭代优化框架

考虑到词、句层级信息的非线性关系,我们避免直接描述其数学关系的方法,而采用以单词级别优化作为主优化目标,以句子级别优化作为迭代求解主优化过程中约束条件的混合迭代优化框架。首先,我们通过贪心法获得单词级别优化的近似解,将该解作为迭代优化的初始值。由于该解既非单词级别上能获取的最优解,又非句子级别上的最优解,所以我们从该解起步,进行迭代的双方优化。在每一轮迭代过程中,根据句子层级的约束条件,进行摘要句子的更换,一方面保证了句子层级上的优化,另一方面,在目标函数的指导下,在多轮迭代中保证了单词级别的近似最优。通过这种方式,达到对词、句层级信息的融合,获得一个在词、句层级上表现双优的摘要。更具体地, $S^{(t-1)}$ 是第 $(t-1)$ 轮迭代结束时获得的摘要解, $D \setminus S^{(t-1)}$ 是第 $(t-1)$ 轮迭代结

束时文档剩余句子的集合, $S^{(0)}$ 是由贪心法获得的初始化句子集合, 在第 t 轮迭代过程中, 根据一定的规则, 从 $S^{(t-1)}$ 中选择部分句子, 与 $D \setminus S^{(t-1)}$ 中的部分句子进行交换, 获得 $S^{(t)}$, 即该轮迭代对摘要解更新。

替换句子组首先需要满足句子级别上的约束条件。我们将满足句子级别上约束条件的替换句子组放入一个替换候补集合中, 该集合中的替换句子组都会带来 $S^{(t)}$ 在句子级别上表现的提升。 Ω_t 是替换句子组集合。

$$\Omega_t = \{ \langle X^{(t)}, Y^{(t)} \rangle \mid X^{(t)} \subseteq S^{(t-1)}, Y^{(t)} \subseteq D \setminus S^{(t-1)} \}。$$

获得替换候补集合后, 从中选择一个替换句子组, 在该候补集合中, 选中的替换组获得的摘要在单词层级上是候补集合的所有替换中表现最优的, 并且该替换仍然使得摘要在单词集合上满足近似最优。换言之, 每次迭代在保证单词级别表现的近似最优的同时, 还带来句子级别上的提升。

迭代框架的公式如下:

$$\operatorname{argmax}_{\langle X^{(t)}, Y^{(t)} \rangle \in \Omega_t} O((S^{(t-1)} - X^{(t)}) \cup Y^{(t)}) - O(S^{(t-1)}), \quad (7)$$

使得

- 1) $L((S^{(t-1)} - X^{(t)}) \cup Y^{(t)}) < L,$
- 2) $U((S^{(t-1)} - X^{(t)}) \cup Y^{(t)}) - U(S^{(t-1)}) > 0,$
- 3) $O(S^{(\max)}) - O((S^{(t-1)} - X^{(t)}) \cup Y^{(t)}) < \varepsilon。$

约束条件 1 表示每一轮替换都必须满足摘要的长度限制。约束条件 2 表示每一轮替换都必须带来句子级别上的优化。将迭代过程中产生的在单词级别上最优的解记做 $S^{(\max)}$ 。在迭代过程中, 允许摘要在单词层级的表现相对于 $S^{(\max)}$ 有一定阈值内的损失。一方面是因为优化的目标函数非凸, 可能存在多个局部最优解, 所以允许一定幅度的临时下降是为了跳过部分的局部最优解, 以增加达到全局最优解的可能性; 另一方面, 约束 3 的限制使得在单词级别表现上, 相对于最优解, 最终解仅有小幅度损失, 这对摘要整体表现的影响是微小的, 而这小部分的损失能够换取摘要在句子级别上性能的提升。该框架在多轮迭代后将收敛, 并且最终获得在单词级别和句子级别上表现双优的摘要。

2 实验

2.1 实验环境设置

实验使用的数据集是自动摘要生成工作的标准

数据集 DUC2003 和 DUC2004。DUC2003 和 DUC2004 都包含 5 个任务, 其中任务 2 是多文档摘要生成任务。DUC2003 的任务 2 提供 30 个文档集合, DUC2004 的任务 2 提供 50 个文档集合。DUC2003 和 DUC2004 的每个文档集合都包含 10 篇长度为 200~500 个词的文章, 任务要求为每个文档集生成 100 个词的摘要。本实验分别抽取 DUC2003 和 DUC2004 的 1/2 文档集合作为训练集, 剩余部分作为测试集合。

本实验搭建在 PKUSUMSUM 平台^[12]上, 该平台提供多个非监督学习摘要方法的实现, 并且这些方法都是目前表现最优的传统方法。这些方法中, 一部分是基于句子级别的信息, 比如 MEAD^[1]、ClusterCMRW^[6]、LexPageRank^[3]和子模方法^[7]; 另一部分基于单词层级信息, 比如 ILP^[8]。为了说明融合词句层级信息的优势, 我们列出这些方法在对应数据集上公布的最优结果, 并与本文框架的表现进行对比。为了测试带约束的迭代优化框架的有效性, 我们实现了同样是融合词句层级信息的线性组合算法(Linear combination)^[9](抽取句子来最大化句子和单词的线性加和评分), 并与本文框架进行对比。为保证实验的公平性, 对所有的算法进行相同的文本预处理。

我们使用 ROUGE^[10]作为实验的评测标准。ROUGE 是 DUC 数据集官方规定的评测标准, 也是摘要工作最常用的评价指标。ROUGE 通过计算生成摘要和标准摘要的单元的重合度, 对摘要的质量进行打分。表 1 公布了 ROUGE 评测结果, 分别为 ROUGE1(R-1)、ROUGE2(R-2), ROUGE-W-1.2(R-W-1.2) 和 ROUGE-SU4(R-SU4) 的 Recall 分数(R)和 F -measure 分数(F)。

2.2 实验结果分析

从表 1 可以看出, 在 DUC2003 和 DUC2004 数据集上, 本文的框架几乎所有的得分都超过其他方法。

首先, 对比本文方法(融合词句信息)与其他方法的实验结果。融合词、句层级信息能够带来效果提升, 说明词、句层级的信息对摘要的贡献互补, 能够互相提高。子模方法最具挑战难度, 这是因为它在对句子特征进行评分时, 采用更加复杂有效的方法, 从而获得对句子级别信息更准确的评价。由于本文的重点在于研究融合词、句信息的有效性以及如何有效地融合词、句信息, 所以对句子信息及

表 1 于 DUC2003 和 DUC2004 数据集上多文档摘要任务的 ROUGE 评分
Table 1 ROUGE scores for multi-document summarization task on DUC2003 and DUC2004 dataset

DUC2003	R				F			
	R-1	R-2	R-W-1.2	R-SU4	R-1	R-2	R-W-1.2	R-SU4
MEAD	0.3075	0.0695	0.1020	0.0848	0.3343	0.0759	0.1353	0.0990
ILP	0.3515	0.0840	0.1197	0.1115	0.3526	0.0844	0.1496	0.1124
LexPageRank	0.3268	0.0705	0.1078	0.0982	0.3582	0.0774	0.1439	0.1165
ClusterCMRW	0.3784	0.0947	0.1255	0.1263	0.3830	0.0958	0.1579	0.1291
Submodular	0.3717	0.0893	0.1226	0.1219	0.3765	0.0904	0.1543	0.1249
Linear combination	0.3540	0.0853	0.1200	0.1141	0.3552	0.0856	0.1500	0.1149
本文	0.3824	0.0952	0.1260	0.1299	0.3861	0.0960	0.1583	0.1323

DUC2004	R				F			
	R-1	R-2	R-W-1.2	R-SU4	R-1	R-2	R-W-1.2	R-SU4
MEAD	0.3316	0.0783	0.1004	0.0979	0.3668	0.0876	0.1357	0.1268
ILP	0.3619	0.0743	0.1121	0.1154	0.3601	0.0743	0.1438	0.1185
LexPageRank	0.3830	0.0920	0.1128	0.1048	0.3607	0.0755	0.1407	0.1202
ClusterCMRW	0.3829	0.0899	0.1179	0.1326	0.3760	0.0908	0.1517	0.1308
Submodular	0.3935	0.0940	0.1190	0.1333	0.3890	0.0943	0.1527	0.1340
Linear combination	0.3599	0.0742	0.1122	0.1150	0.3607	0.0743	0.1439	0.1155
本文	0.3867	0.0946	0.1191	0.1347	0.3875	0.0948	0.1528	0.1352

说明：粗体数据表示在该列得分的最大值。

单词信息的刻画都只采用最直接的特征，因此分别在词、句的信息抽取性能上会略输于专注抽取其一的方法。然而，在采用最简单直接的特征分别刻画词、句的情况下，本文框架仍然在绝大部分指标上超过其他方法，进一步说明融合词、句层级信息这一思路的有效性。

然后，对比融合词、句信息的不同方法，即对比本文框架与线性组合框架的实验结果。可以看出，迭代优化框架明显优于线性组合框架。这主要是因为单词的语义空间与句子的语义空间之间不是简单的线性关系，所以直接通过线性组合来融合词、句信息是低效甚至有损性能的。更合理的方式是，避开确切的词、句信息的关系刻画，将其分别处理。本文将单词信息作为优化目标，句子信息作为优化约束，在迭代优化过程中达到词、句层级都得到优化的效果。实验结果表明，分开处理、迭代优化的框架在处理词、句信息融合上具备有效性。

2.3 参数作用分析

实验中调整参数时，每次只改变一个参数，其

他参数保持固定，记录实验结果。参数调整的结果如图 1 所示。

第一组参数是约束条件 2 中的 $\lambda_c, \lambda_d, \lambda_i$ ，结果如图 1(b)~(d)所示。从图 1 可知，句子集合的多样性、代表性都对文档摘要的效果有较大影响，而句子集合覆盖度的影响力度较小。从 3 个参数的走势可知，句子集合的信息覆盖度、多样性、代表性都是构成句子级别表现的重要因素，所以参数值过高或过低都会影响其他两个特征发挥作用，但它们会在一个中间点达到平衡。本文实验中设置 $\lambda_c = 0.3$ ， $\lambda_d = 0.4$ ， $\lambda_i = 0.3$ 。

另一个关键参数是约束条件 3 中的 ε ，其结果如图 1(a)所示。当 ε 增大时，表明迭代过程中对单词性能的临时性降低的容忍阈增大，该容忍程度的增大有助于跳出局部最优点，并且为句子性能的提升提供更大的空间。该参数在一定区间内显示出不稳定的升降幅度，当 ε 到达 0.17 后，随着 ε 的增大，性能降低，这是因为该降低幅度过大损害了单词级别的表现。本文实验设置 $\varepsilon = 0.17$ 。

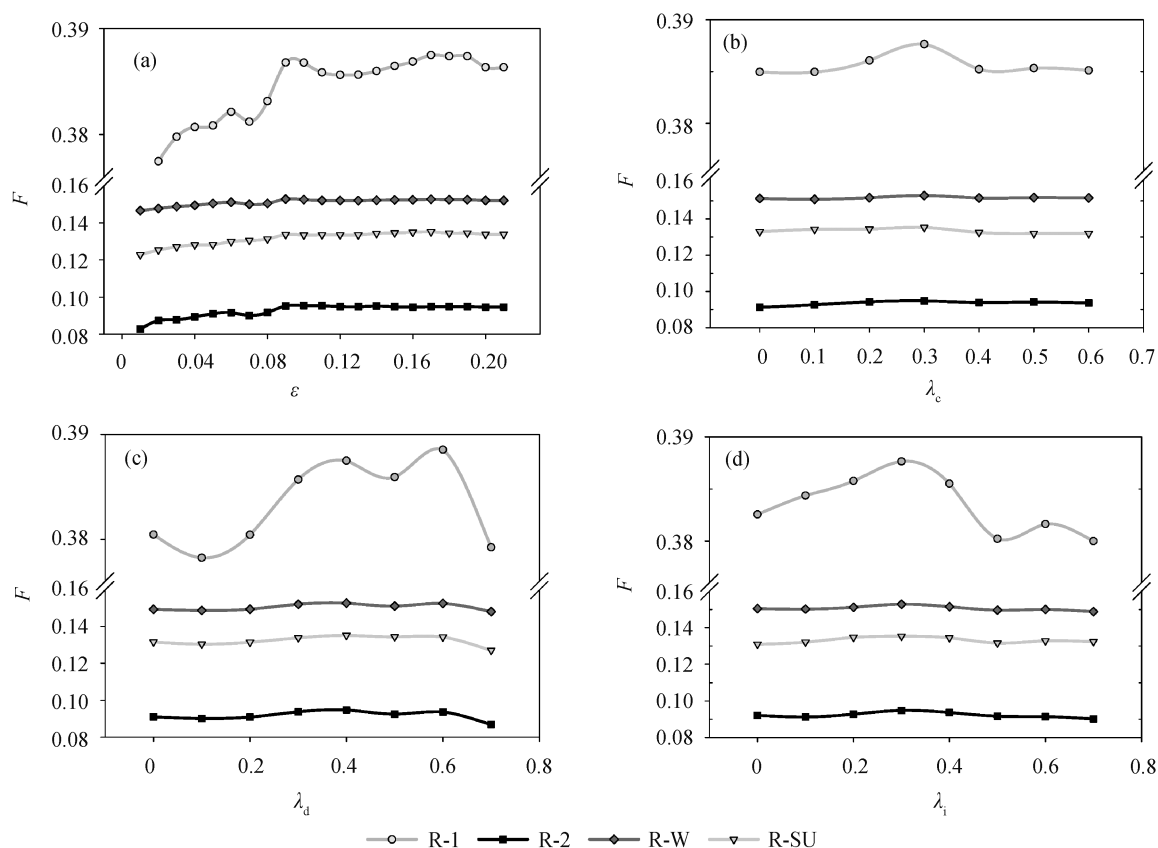


图 1 DUC2004 数据集上的调参结果
Fig. 1 Parameter tuning on DUC2004 dataset

3 总结

本文考虑词、句层级信息对摘要贡献的互补, 提出一个创新的融合词、句层级信息的抽取式摘要生成框架。该方法将单词信息作为优化目标, 句子信息作为优化求解过程中的约束, 以迭代替换的框架进行优化的求解, 是一个有效的融合词、句层级信息的模型。实验结果也显示了本文优化框架的优越性。

由于本文的重点在于提出融合词、句层级信息的重要性及融合的具体策略, 实验中使用的词、句特征都较为肤浅, 以便比较并突显框架本身的效果。如果采用更加有效、蕴含信息更加丰富的词、句特征(比如词向量和句向量表示), 或者采用神经网络的方法对框架进行改进, 则该框架还有很大的优化空间。

致谢 感谢 PKUSUMSUM 平台对本研究项目的帮助。

参考文献

- [1] Radev D, Allison T, Blair-Goldensohn S, et al. MEAD — a platform for multidocument multilingual text summarization // Proceedings of the 4th International Conference on Language Resources and Evaluation. Lisbon, 2004: 699–702
- [2] Lin C Y, Hovy E. From single to multi-document summarization: a prototype system and its evaluation // Proceedings of the 40th Annual Meeting of Association for Computational Linguistics. Philadelphia, 2002: 457–464
- [3] Erkan G, Radev D. LexRank: graph-based lexical centrality as salience in text summarization. Journal of Artificial Intelligence Research, 2004, 22: 457–479
- [4] Mihalcea R. Graph-based ranking algorithms for sentence extraction, applied to text summarization //

- Proceedings of the ACL Interactive Poster and Demonstration Sessions. Barcelona, 2004: 170–173
- [5] Wan Xiaojun. Document-based HITS model for multi-document summarization // Proceedings of the 10th Pacific Rim International Conference on Artificial Intelligence: Trends in Artificial Intelligence. Hanoi, 2008: 454–465
- [6] Wan Xiaojun, Yang Jianwu. Multi-document summarization using cluster-based link analysis // Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Singapore, 2008: 299–306
- [7] Lin Hui, Bilmes J. Multi-document summarization via budgeted maximization of submodular functions // Proceedings of Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Los Angeles, 2010: 912–920
- [8] Gillick D, Favre B. A scalable global model for summarization // Proceedings of the Workshop on Integer Linear Programming for Natural Language Processing. Boulder, 2009: 10–18
- [9] Xie S S, Favre B, Hakkani-Tür D, et al. Leveraging sentence weights in a concept-based optimization framework for extractive meeting summarization // Proceedings of the 10th Annual Conference of the International Speech Communication Association. Brighton, 2009: 1503–1506
- [10] Lin C Y. ROUGE: a package for automatic evaluation of summaries // Proceedings of Workshop on Text Summarization Branches Out, Post-Conference Workshop of the Annual Meeting of the Association for Computational Linguistics. Barcelona, 2004: 74–81
- [11] Neto J L, Santos A D, Kaestner C A A, et al. Generating text summaries through the relative importance of topics // Advances in Artificial Intelligence. Atibaia, SP, 2000: 300–309
- [12] Zhang Jianmin, Wang Tianming, Wan Xiaojun. PKUSUMSUM: a java platform for multilingual document summarization // Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations. Osaka, 2016: 287–291