

基于改进的 LDA 模型的中文主观题自动评分研究

罗海蛟 柯晓华

(广东外语外贸大学思科信息学院 广州 510006)

(广东外语外贸大学语言工程与计算实验室 广州 510006)

摘要 主观题自动评分(Automated Scoring Subjective Responses, ASSR)在语言学习与语言测试领域的诊断信息及信度方面具有重要的应用前景。将主题模型中的隐含狄利克雷分配(Latent Dirichlet Allocation, LDA)引入到中文主观题自动评分中,提出了一种结合专家知识的改进的 LDA 模型,并采用了一种综合文档-隐含主题概率向量及隐含主题-核心词项概率向量的文本特征表示方式。实验对比了改进的 LDA 与潜在语义分析(Latent Semantic Analysis, LSA)的自动评分效果,结果表明改进的 LDA 模型在中文主观题自动评分中切实有效。

关键词 主观题自动评分,潜在语义分析(LSA),隐含狄利克雷分配(LDA),绝对一致性,相邻一致性

中图法分类号 TP18 **文献标识码** A

Automated Scoring Chinese Subjective Responses Based on Improved-LDA

LUO Hai-jiao KE Xiao-hua

(Cisco School of Informatics, Guangdong University of Foreign Studies, Guangzhou 510006, China)

(Laboratory of Language Engineering and Computing, Guangdong University of Foreign Studies, Guangzhou 510006, China)

Abstract Automated scoring subjective responses (ASSR) have great promise for providing diagnostic information and reliability to aid language learning and testing. In the present study, we introduced the latent Dirichlet allocation (LDA) into an automated scoring task with Chinese subjective responses, and an improved LDA model with experts' knowledge was proposed. In the novel model, we proposed a text feature representation approach integrating document-latent topic probability vector and latent topic-core terms probability vector. Experiment results show that the improved-LDA is better than LSA in terms of the autoscoring performances. The findings of this study highlight the model selection in application of automated scoring Chinese responses with language testing.

Keywords Automated subjective question scoring, Latent semantic analysis (LSA), Latent Dirichlet allocation (LDA), Absolute accuracy rate, Adjacent accuracy rate

在国内外的各类语言测试中,简答、作文等是常见的主观题形式,目前大多使用传统人工批改,但该方法效率低下,评分质量具有很强的主观性,特别是在大型考试中很难保证阅卷的信度和效度。而采用计算机来评阅主观题可作为人工评阅的有效补充,帮助检查人工评分的错漏,有效提高评分工作的信度。该方法不但高效、快捷,而且可以保障大型考试阅卷的公正性和效率^[1]。

主观题自动评分的研究与实践早在 20 世纪 60 年代就已开始,目前国外已有多个成型的主观题自动评分系统,如 Project Essay Grade (PEG), Intelligent Essay Assessor (IEA) 和 E-rater 等,并已应用于 GRE, TOFEL 和 GMAT 等大型考试^[2-3]。在国内该领域的研究起步较晚,梁茂成主持了中国学生英语作文自动评分模型的构建^[4];任春艳进行了汉语水平考试(HSK)高等作文评分的客观化研究^[5];柯晓华开展了中文主观题的文本质量分析研究^[6]等。

早期的主观题自动评分方法多采用多元回归及相关统计技术,通过提取一系列表层形式特征(如文本长度、词语密度

等)来进行拟合,但这些指标多是在词层面简单地统计特征;为了发现文本内容的潜在语义,Landauer 等借鉴信息检索领域的成果,首次将潜语义分析(Latent Semantic Analysis, LSA)用于作文自动评分^[7-8],该方法得到了应用语言学界的认可;桂诗春论证了“潜语义分析可以用来比较不同语篇的语义相似性”^[9]。目前大部分的自动评分系统都结合 LSA 模型来改善评分效果,但 LSA 也存在不足。LSA 最初被用于信息检索领域,其实质属于广义上的主题模型,采用的奇异值分解(SVD)算法的复杂度较高,不利于增量更新模型;且其模型在本质上假设各个 topic 是互相垂直的向量,该假设不太合理,可能导致在某些维度上是负分量。Hofmann 等^[10]放弃使用 LSA 中的矩阵分解转换方法,进而采用概率主题的生成模型思路,提出了概率潜语义分析(Probabilistic Latent Semantic Analysis, PLSA)。相对于 LSA,该分析有了坚实的统计理论支撑,模型中每个变量以及相应的概率分布和条件概率分布都有明确的物理解释。但 PLSA 模型中每一个文档集都对应着模型的一系列参数,难以处理数据集外的新文本,且待估

本文受广东省自然科学基金资助项目(2015A030313575)资助。

罗海蛟 男,硕士,讲师,CCF 会员,主要研究方向为自然语言处理、语言测试,E-mail:luohaijiao@gdufs.edu.cn;柯晓华(1976—),女,博士,CCF 高级会员,主要研究方向为计算语言学、语言测试,E-mail:carriek@gdufs.edu.cn(通信作者)。

参数随着训练文本的增加而线性增加,其过拟合的问题也随之凸显。Blei 等于 2003 年提出了隐含狄利克雷分配(Latent Dirichlet Allocation, LDA)^[11],其可看作是 PLSA 的贝叶斯化版本,在数学上是更完美的完全主题模型。LDA 模型参数空间的规模与数据集规模无关,与 PLSA 相比更不容易出现过拟合问题;LDA 不仅对已有的数据集文本进行估计,还能引入的新数据中的相似文本给出一个较高的概率值,从而解决 PLSA 难以处理新数据的问题。LDA 也有主题模型难以弥补的一些缺点,例如 topics 的数量不易确定,尽管相关研究提出了一些自动确定 topics 参数的方法^[12],但其模型和计算都非常复杂,效果也不理想。

LDA 大量应用于机器学习的各领域,但在主观题的文本自动评分中还未得到充分的重视和研究。中文主观题不同于一般的命题作文,其有标准参考答案,有相对固定的主题构成,LDA 模型能够更有效地从答题文本中抽取隐含主题。本文引入 LDA 模型并改进其用法,通过实验来对比 LDA 和 LSA 在以中文简答题为代表的客观题的评分效果,结果表明改进的 LDA 模型适用于简答题类型的中文主观题自动评分。

1 相关工作

1.1 LSA

LSA 在文档离散化预处理时采用词袋模型(PLSA 和 LDA 的文本预处理过程也采用词袋模型),预处理后对特征项加权(常用的加权方式有 TF-IDF、信息熵等),从而使文档数据集转换为加权词项-文档矩阵,LSA 处理的核心是对该加权词项-文本矩阵进行截断奇异值分解(Truncated Singular Value Decomposition, TSVD),如图 1 所示。

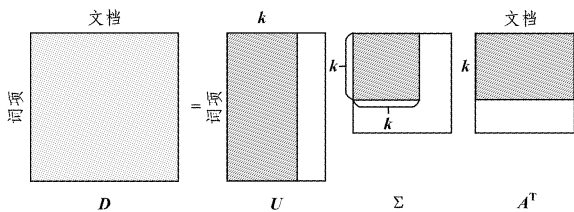


图 1 TSVD 说明

$D_{m \times n}$ ($m \geq n$, n 为文档数)为加权词项-文档矩阵,SVD 分解为:

$$D = U \Sigma V^T \quad (1)$$

其中, U 和 V 分别为 $m \times n$ 和 $n \times n$ 的正交矩阵,称为左、右奇异值向量, Σ 是 $m \times n$ 的对角矩阵,对角元素为 D 的奇异值。 D 的 k 秩近似矩阵 D_k 为:

$$D_k = U_k \Sigma_k V_k^T \quad (2)$$

其中, k 表示从原对角矩阵“截取” k 个奇异值, Σ_k 为 $k \times k$ 的对角矩阵, V_k^T 为生成的低维潜语义空间(此处为 k 维),用于描述语言元素间的语义特征结构。

1.2 PLSA 与 DSA

不同于 LSA 只考虑词项-文档间的关系,PLSA 是一种文档-隐含主题-词项的三级两层的概率生成模型。该模型假定文档由若干个满足多项分布的隐含主题混合而成,而每个主题又是由若干个满足多项分布的词项构成。设文档集合 $D = \{d_1, d_2, \dots, d_M\}$,词项集合 $W = \{w_1, w_2, \dots, w_V\}$,隐含主题变量 $Z = \{z_1, z_2, \dots, z_T\}$,其概率图模型(阴影圆圈表示观察变

量节点,空白圆圈表示隐藏变量节点)如图 2 所示。其中, M 表示文档总数, T 表示隐含主题数, N_{d_i} 表示文档 d_i 的长度(即单词总数), $\vec{\theta}_{d_i}$ 表示某文档 d_i 内的隐含主题分布, $\vec{\phi}_t$ 表示某隐含主题 z 上的词项分布。

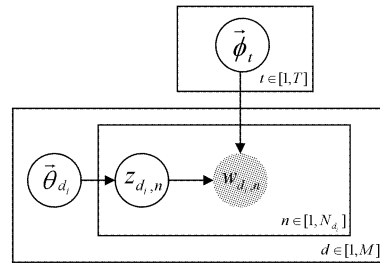


图 2 PLSA 概率图模型

$P(d_i)$ 表示从文档集中选择文档 d_i 的概率, $P(w_j | z_t)$ 表示某个词 w_j 在给定主题 z_t 下出现的概率, $P(z_t | d_i)$ 表示某个主题 z_t 在给定文档 d_i 下出现的概率,则 PLSA 文档生成的模型为:

$$P(d_i, w_j) = P(d_i) P(w_j | d_i) \\ = P(d_i) \sum_{t=1}^T P(w_j | z_t) P(z_t | d_i) \quad (3)$$

其中,隐含主题-词项分布 $P(w_j | z_t)$ 和文本-隐含主题分布 $P(z_t | d_i)$ 为需要获取的未知模型参数,通常采用极大似然估计策略,使用 EM 算法来估算^[13]。

相比于 LSA 隐含假定的高斯分布假设,PLSA 假定的多项分布假设更符合文本特性,但随着文档集合的增大,该模型需要学习的模型参数也线性增加,而且也不好处理文档集合外的新文本。LDA 可以看作是贝叶斯化版本的 PLSA,是一种三层产生式模型,其概率图模型如图 3 所示。

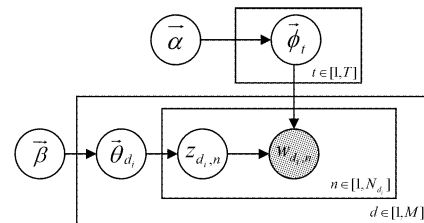


图 3 LDA 概率图模型

其文本-隐含主题分布不再从文档中直接产生,而是服从一个参数为 $\vec{\alpha}$ 的先验 Dirichlet 分布,隐含主题-词项分布也服从一个参数为 $\vec{\beta}$ 的先验 Dirichlet 分布(见图 3)。则在 LDA 模型下一篇文章的生成过程为:

- 1) 从超参数为 $\vec{\alpha}$ 的 Dirichlet 分布中取样生成文档 d_i 的文本-隐含主题分布 $\vec{\theta}_{d_i}$;
- 2) 从文本-隐含主题分布 $\vec{\theta}_{d_i}$ 中取样生成文档 d_i 的第 n 个词的隐含主题 $z_{d_i,n}$;
- 3) 从超参数为 $\vec{\beta}$ 的 Dirichlet 分布中取样生成隐含主题 $z_{d_i,n}$ 的词项分布 $\vec{\phi}_{z_{d_i,n}}$;
- 4) 从词项分布 $\vec{\phi}_{z_{d_i,n}}$ 中取样最终生成文档 d_i 的第 n 个词 $w_{d_i,n}$;
- 5) 重复 2)–4) N_{d_i} 次,最终生成文档 d_i 的全部的 N_{d_i} 个单词。

按照概率图模型,生成一个文档的联合概率分布为:

$$P(\vec{w}_{d_i}, \vec{z}_{d_i}, \vec{\theta}_{d_i}, \vec{\varphi}_i | \vec{\alpha}, \vec{\beta}) = \prod_{n=1}^{N_{d_i}} P(w_{d_i,n} | \vec{\varphi}_{z_{d_i,n}}, z_{d_i,n}) P(z_{d_i,n} | \vec{\theta}_{d_i}) \times P(\vec{\theta}_{d_i} | \vec{\alpha}) \times \prod_{i=1}^T P(\vec{\phi}_i | \vec{\beta}) \quad (4)$$

除了 $w_{d_i,n}$ 是观察变量, $\vec{\alpha}, \vec{\beta}$ 是根据经验给定的先验超参数, 其余都是隐含变量, 反向推出未知的隐含模型参数的策略一般是采用变分 EM 算法^[11] 或者 Gibbs 采样算法^[14] 进行估算, 本文采用后者。算法的最终目的是获得每个文档中的 T 维文档-隐含主题概率向量 $\vec{\theta}$, 及对应的 $T \times V$ 维隐含主题-词项概率向量 $\vec{\phi}$ 。

2 基于改进 LDA 模型的主观题自动评分

在主观题自动评分领域, LSA 模型使用得最为广泛, 笔者也曾引入 PLSA 模型代替 LSA 模型^[15], 但在评分效果上差别不显著; LDA 模型广泛用于文本分类及相关领域, 本文将 LDA 模型引入到以简答题型为代表的中文主观题的自动评分中, 代替了常用的 LSA 模型, 并针对 LDA 模型的特性进行了一定的改进。

2.1 改进的 LDA 模型

2.1.1 基于专家知识的隐含主题数的确定

对于 LDA 模型, 其隐含主题数 T 的确定是个难点, 目前并没有高效的方法, 常见的方法有基于密度的最优模型选择最优主题数^[16]、根据数据自动确定的 HDP 模型^[12] 或者通过穷举主题数来择优。但无论采用哪种方法, 其训练效率都非常低, 而且效果也不能得到保证。但对于中文简答题这种主观题型, 其包含的主题能够确定, 可以通过专家知识来确定构建的 LDA 模型的主题数的大致范围, 例如对于一道马克思主义原理课程的简答题, 专家完全可以根据标准答案给出其包含的主题范围, 从而确定一个主题数范围, 再在该范围内使用 HDP 等模型算法择优, 可在很大程度上提高效率。

2.1.2 改进的概率向量模型

LDA 采用词袋模型表示降维后抽取的文本特征, 当其用于文本分类时, 特征表示为各个文档-隐含主题的概率向量, 进而通过分类算法来比较各个文档的概率向量相似度。该方法未利用主题-词项的概率向量, 也未考虑领域背景中各词项的权重。本文提出一种改进方法, 即融入专家知识来修改原始词袋模型, 并创新地提出一种综合文档-隐含主题概率向量及主题-核心词项概率向量的表示模型, 具体的实现方式如下。

(1) 由于主题的分配向高频词所属主题倾斜, 因此首先对原始简答题答案文本集去除停用词并分词, 由专家根据标准答案提取 m 个核心词, 在扫描文本集后进行核心词的同义词合并 (将所有与核心词同意的词都合并为核心词)^[17]; 然后对每篇答案文本构建词袋模型, 并对其中出现的核心词给予权重提升, 本文将其词频乘以 1.5 倍。

(2) 对每篇文本集经过 LDA 降维降噪后的特征表示方式进行改进, 构造如下概率向量模型:

$$(\theta_1 \times (\phi_{11}, \dots, \phi_{1m}), \dots, \theta_z \times (\phi_{z1}, \dots, \phi_{zm}), \dots, \theta_T \times (\phi_{T1}, \dots, \phi_{Tm})) \quad (5)$$

其中, $\theta_1, \theta_2, \dots, \theta_T$ 是 T 维文档-隐含主题概率向量 $\vec{\theta}$ 的值, $\phi_{z1}, \phi_{z2}, \dots, \phi_{zm}$ 表示对应第 z 个隐含主题的词项分布中的 m 个核心词的概率分布, 因此构造的概率向量模型为 $T \times m$ 维。这种文档的特征提取方式既保留了 LDA 降维降噪的效果, 也融入了专家的领域知识, 具有更好的语义代表性。

2.2 评分模型整体框架

对于主观题的自动评分, 可以将分数看作是离散的“分数段类别”, 从而转换为文本分类问题。将改进的 LDA 模型用于文档的特征抽取/降维降噪, 以形成文档的主题-词项综合概率向量模型, 再使用 KNN、SVM、随机森林等分类算法建立评分模型, 将各个文档“分类”到各档成绩中。本文提出的评分模型的整体框架如图 4 所示。

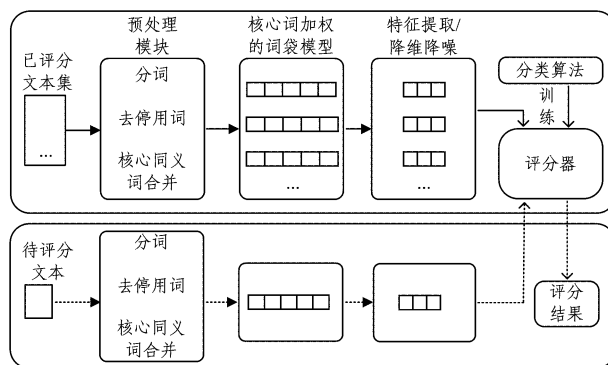


图 4 评分模型的整体框架

3 实验

3.1 实验数据

实验使用的语料文本为某大学的毛泽东思想和中国特色社会主义理论体系概论课期末考题 (满分 6 分), 题目的内容以及答案如下。

题目: 经济发展是实现人口、资源、环境与经济协调发展的根本出路, 同时经济的发展也离不开人口、资源和环境的支持。

答案: 实现可持续发展是我国现代化建设的一个重大战略。要把控制人口、节约资源、保护环境放到重要位置, 使人口增长与社会生产力的发展相适应, 使经济发展与资源、环境相协调, 实现良性循环。中国解决所有问题的关键是要靠自己的发展, 根本是发展经济。所以, 实现可持续发展首先在于发展。(观点正确的依据) 经济发展是实现人口、资源、环境与经济协调发展的根本出路。但同时, 经济发展又离不开人口、资源和环境的支持。

由于此考题并非出自于大规模考试, 其样本数量相对较小, 因此样本质量尤其重要。样本采自两届学生的真实作答, 为避免教师改卷的主观偏差, 每一份答案都另请 3 名专家独立评分, 剔除无效作答后, 从中选择 3 分、4 分、5 分、6 分答案文本各 160 篇 (共 640 篇) 作为原始文本集合 D , 用于生成训练集及测试集。

3.2 实验设置

(1) 训练集及测试集的生成。常用的生成方法有留出法 (hold-out)、 k 折交叉验证法 (k -fold cross validation) 等。但本研究的样本数据获取方式的成本较高, 数量相对较小, 因此本实验采用改进的自助法 (bootstrapping), 具体方式如下:

1) 对原始文本集 D 的 4 个分数段分别进行随机采样, 各抽取一个样本放入数据集 P 中, 然后将此 4 个样本放回 D ;

2) 重复操作 1) 105 次, 则 P 集合中有 420 个样本可作为训练集使用, 其余的 220 个样本作为测试集使用;

3) 重复操作 1) 和操作 2) 5 次, 生成 5 对训练集/测试集, 用于对每一种算法进行 5 轮实验验证并取结果的平均值。

(2)LDA 模型的参数选择。LDA 的隐含主题数 T 由 3 位专家根据题干和标准答案给出,本实验中 T 取值为 6,超参数 $\vec{\alpha}$ 统一取值为 $1/T$,超参数 $\vec{\beta}$ 统一取值为 0.01。

(3)分类算法选择。为了更全面地评估本研究的工作成果,本文采用 K 近邻(KNN)算法和支持向量机(SVM)算法分别加以验证。KNN 中 K 取值为 5;SVM 分类器使用 LibSVM,核函数使用线性核,其他参数使用默认值。

(4)评分效果验证。在语言测试领域中,常采用绝对一致性(exact agreement)和相邻一致性(adjacent agreement)百分比来检验机器的自动评分效果^[18],某些文献也将以上指标称为绝对准确率和相邻准确率,本文采用这两个指标来对比基于 LSA 及改进的 LDA 模型的评分效果。

3.3 主要实验结果

将 bootstrapping 方法生成的 5 对训练集/测试集分别用于评分,因此图 5—图 8 中都分别给出了 5 个实验结果及 1 个平均结果。图 5 给出了采用 KNN 分类算法得到的评分绝对一致性,图 6 给出了采用 KNN 分类算得到的评分相邻一致性,图 7 给出了采用 SVM 分类算法得到的评分绝对一致性,图 8 给出了采用 SVM 分类算法得到的评分相邻一致性。

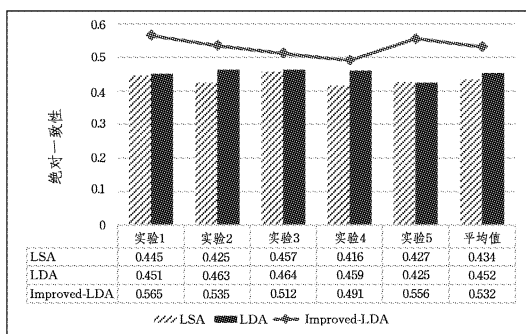


图 5 绝对一致性对比图(KNN)

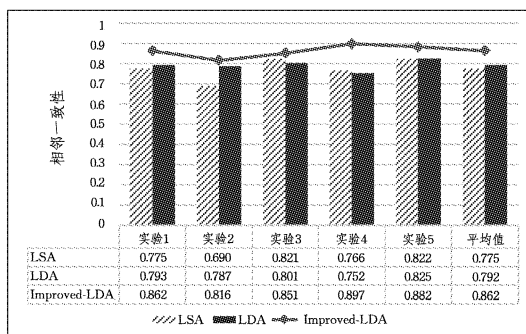


图 6 相邻一致性对比图(KNN)

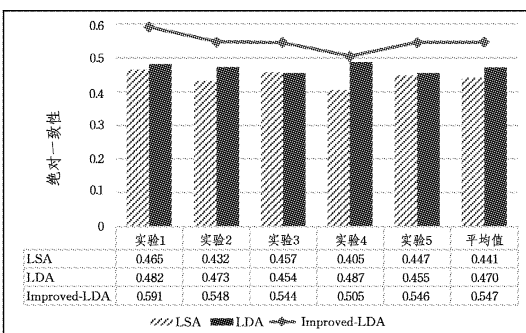


图 7 绝对一致性对比图(SVM)

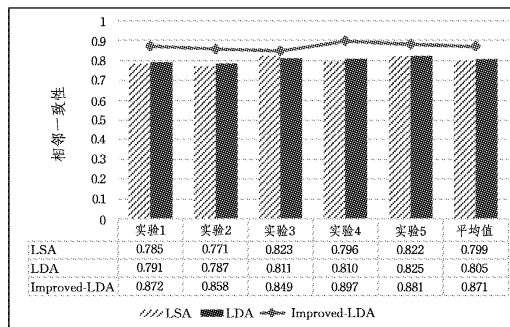


图 8 相邻一致性对比图(SVM)

由实验结果可知,基于 LDA 的自动评分效果整体优于基于 LSA 的自动评分效果,而基于改进的 LDA 模型与标准 LDA 相比在评分效果上有进一步的提高。使用 KNN 分类器时,LDA 模型相对 LSA 模型在绝对一致性方面有约 4 个百分点的提升,在相对一致性方面有约 2 个百分点的提升,而改进的 LDA 模型相对 LDA 模型在绝对一致性方面有约 15 个百分点的提升,在相对一致性方面有约 8 个百分点的提升;当使用 SVM 分类器时,提升效果略好于 KNN,LDA 模型相对于 LSA 模型在绝对一致性方面有约 6 个百分点的提升,相对一致性有约 3 个百分点的提升,而改进的 LDA 模型相对于 LDA 模型在绝对一致性方面有约 16 个百分点的提升,相对一致性有约 8 个百分点的提升。

结束语 本文将隐含主题模型引入中文主观题自动评分领域,并对 LDA 算法进行改进,融入专家知识,提出一种综合文档-隐含主题概率向量及主题-核心词项概率向量的文本特征表示模型。该模型在一定程度上弥补了词袋模型固有的缺点,实验结果也表明其相对于 LSA 方法的效果有一定提升。本文提出的改进模型切实有效。

改进的 LDA 模型在本质上是基于词袋模型来提取文本特征,对语句、段落和篇章的内在语义挖掘不够深入,我们后续的工作中将研究文本集表示为基于依存句法的复杂网络,结合复杂网络与 LDA 来提高自动评分的效果。

参考文献

- [1] 徐昌火,陈东,吴倩,等. 汉语作为第二语言作文自动评分研究初探[J]. 国际汉语教学研究,2015(1):83-89.
- [2] DIKLI S. An overview of automated scoring of essays[J]. Journal of Technology, Learning, and Assessment,2006(1):3-35.
- [3] QUELLMALZ, PELLEGRINO E S, W J. Technology and testing[J]. Science,2009(2):75-79.
- [4] 梁茂成. 中国学生英语作文自动评分模型的构建[M]. 北京:外语教学与研究出版社,2011.
- [5] 任春艳. HSK 作文评分客观化探讨[J]. 汉语学习,2004(6):58-67.
- [6] XIAHUA K, YONGQIANG Z, MA Q, et al. Complex dynamics of text analysis[J]. Physica A: Statistical Mechanics and its Applications, 2009, 415C: 307-314.
- [7] LANDAUER T K. Automatic essay assessment[J]. Assessment in Education: Principles, Policy and Practice, 2003(3):295-309.
- [8] LANDAUER, K T, DUMAIS, et al. A solution to Plato's problem; The latent semantic analysis theory of acquisition, induction, and representation of knowledge[J]. Psychological Review, 1997, 104(2):211-240.

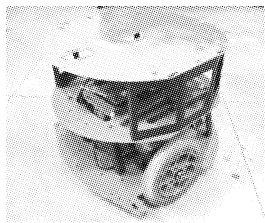


图8 差速驱动移动机器人平台的机构

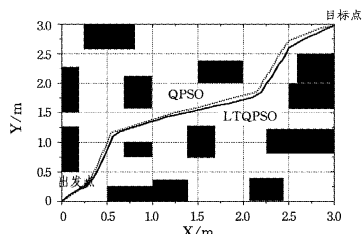


图9 LTPQSO算法与PQSO算法的最优路径对比

(1)最优路径长度的比较

从图9可以看出,LTPQSO算法搜索的路径优于PQSO算法。LTPQSO的路径长度为4.0252m,而PQSO规划的路径长度为4.3812m。PQSO算法在搜索的过程中陷入局部最优,而LTPQSO算法规划选择的路径是全局最优的。

(2)收敛速度的比较

从图10、图11可以看出,PQSO算法与LTPQSO算法的收敛速度曲线均在149次左右就达到了局部的最优值,在接下来的迭代次数中LTPQSO算法大概在200次时跳出局部最优值,又经过50次左右的迭代搜索到了全局的最优值,而PQSO算法仍然陷入局部的最优值。显而易见,LTPQSO算法跳出局部最优解的能力强于PQSO算法。

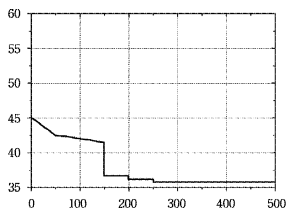


图10 LTPQSO的收敛过程

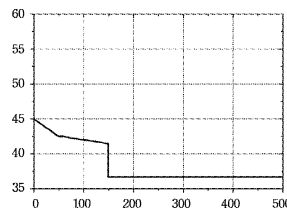


图11 PQSO的收敛过程

结束语 本文提出了一种改进的LTQPSO算法。将LTQPSO算法应用于移动机器人的路径规划中,初始分布取均匀分布和正态分布时,随着粒子维度的变化,分析了LTQPSO算法的粒子数与迭代次数的关系,提出了一元线性回归基本参数评估方程。在不同粒子维度环境下比较了LTQPSO算法与QPSO算法的性能。仿真实验结果表明,与QPSO算法的路径规划方法相比,基于LTQPSO算法和基本参数评估方程的路径规划方法具有更强的搜索能力、更快的收敛速度以及更好的稳定性。

参考文献

[1] CLERC M, KENNEDY J. The Particle Swarm-Explosion, Stability, and Convergence in a Multidimensional Complex Space[J]. IEEE Transaction on Evolutionary Computation, 2002, 6(1): 58-73.

[2] TAN G Z, BAO K, RICHARD M R. A composite particle swarm algorithm for global optimization of multimodal functions[J]. Journal of Central South University, 2014, 21(5): 1871-1880.

[3] LUNG X L, LI W F, ZHANG Y, et al. An adaptive particle swarm optimization method based on clustering[J]. Soft Comput, 2014, 14: 1262-1280.

[4] SHIKHA A, SANJAY S. FRPSO: Fletcher-Reeves based particle swarm optimization for multimodal function optimization[J]. Soft Computing, 2014, 18: 2227-2243.

[5] 周游. 基于粒子群优化的动态优化研究[D]. 杭州: 浙江大学, 2014.

[6] 徐小俊, 雷秀娟, 郭玲. 基于SWGPSO算法的多序列比对[J]. 计算机工程, 2011(6): 184-186.

[7] 黄仁全, 靳聪, 等. 自适应局部增强微分进化改进算法[J]. 空军工程大学学报(自然科学版), 2011(3): 84-89.

[8] 武鲁晓. 改进粒子群算法在电力系统无功优化中的应用[D]. 济南: 山东大学, 2012.

[9] 李仁府, 独孤明哲, 胡麟. 基于PSO算法的路径规划收敛性与参数分析[J]. 华中科技大学学报(自然科学版), 2013(1): 271-275.

[10] 方伟, 孙俊, 谢振平, 等. 量子粒子群优化算法的收敛性分析与控制参数研究[J]. 物理学报, 2010, 59(6): 3686-3694.

(上接第105页)

[9] 桂诗春. 潜伏语义分析的理论及其应用[J]. 现代外语, 2003(1): 76-84.

[10] HOFMANN, THOMAS. Unsupervised Learning by Probabilistic Latent Semantic Analysis[J]. Machine Learning, 2001, 42(1/2): 177-196.

[11] BLEI D M, NG A Y, JORDAN M I. Latent Dirichlet Allocation [J]. The Journal of Machine Learning Research, 2003, 3: 993-1022.

[12] TEH Y W, JORDAN M I, BEAL M J, et al. Hierarchical Dirichlet Processes[J]. Journal of the American Statistical Association, 2006, 101(476): 1566-1581.

[13] MEI Q, ZHAI C. A note on EM algorithm for probabilistic latent semantic analysis[C]// Proceedings of the International Conference on Information and Knowledge Management, CIKM, 2001.

[14] PORTEOUS I, NEWMAN D, IHLER A, et al. Fast collapsed gibbs sampling for latent dirichlet allocation[C]// Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2008.

[15] KE X H, LUO H J. Using LSA and PLSA for Text Quality Analysis[C]// International Conference on Electronic Science and Automation Control. Atlantis Press, 2015.

[16] 曹娟, 张勇东, 李锦涛, 等. 一种基于密度的自适应最优LDA模型选择方法[J]. 计算机学报, 2008, 31(10).

[17] 哈工大社会计算与信息检索研究中心-同义词词林[EB/OL]. <http://www.ltp-cloud.com/download>.

[18] SHERMIS M D, BURSTEIN J C. Automated essay scoring: A cross-disciplinary perspective[M]. Routledge, 2003.