

# 基于三维地形的路径规划算法研究

周兰凤<sup>1</sup> 钱伟杰<sup>1</sup> 曹国刚<sup>1\*</sup> 章民融<sup>2\*</sup>

<sup>1</sup>(上海应用技术大学计算机科学与信息工程学院 上海 201418)

<sup>2</sup>(上海市计算技术研究所 上海 200040)

**摘要** 研究了三维环境中移动机器人的路径规划问题,对蚁群算法中的启发函数和信息素更新方式进行改进,引入了启发函数中路径距离的影响因子,使算法的搜索具有方向性。同时采用了信息素的一种惩罚机制,结合挥发率的自适应改变,来加快算法的收敛速度。仿真结果表明,改进后的算法规划得到的三维路径长度和搜索效率得到了提高。

**关键词** 蚁群算法 启发函数 信息素 挥发率

**中图分类号** TP3 **文献标识码** A **DOI**:10.3969/j.issn.1000-386x.2018.08.050

## RESEARCH ON PATH PLANNING ALGORITHM BASED ON 3D TERRAIN

Zhou Lanfeng<sup>1</sup> Qian Weijie<sup>1</sup> Cao Guogang<sup>1\*</sup> Zhang Minrong<sup>2\*</sup>

<sup>1</sup>(School of Computer Science and Information Engineering, Shanghai Institute of Technology, Shanghai 201418, China)

<sup>2</sup>(Shanghai Institute of Computing Technology, Shanghai 200040, China)

**Abstract** The path planning problem of mobile robot in three-dimension environment is studied in this paper. The updating methods of heuristic function and pheromone in ant colony algorithm are improved by introducing the influence factor of path distance in heuristic function which directs the algorithm searching. Also, a penalty mechanism of pheromone is adopted, combining with the self-adaptation of volatilization rate, to improve the convergence speed of algorithm. Simulation results show that the improved algorithm improves the length of the 3D path and search efficiency.

**Keywords** Ant colony algorithm Heuristic function Pheromone Volatilization rate

## 0 引言

路径规划技术是移动机器人的关键技术和研究的重要领域之一,移动机器人的路径规划是指在有障碍物的工作环境中,为了给机器人规划一条从给定起始点到目标终点的最短或者最优的行动路径,且使得机器人在运动的过程中能够安全无碰撞地绕过障碍物。目前国内外的研究用于移动机器人的路径规划的算法主要有:A\*算法<sup>[1]</sup>、遗传算法<sup>[2]</sup>、人工势场法<sup>[3]</sup>、神经网络法<sup>[4]</sup>等,这些路径规划的方法或多或少存在一些不足之处,使得规划的效果不尽人意。自蚁群算法提出后,路径规划与蚁群算法相结合,有效地提高了路径规划的效率。然而传统蚁群算法在解决路径规划上存

在收敛速度慢,易陷入局部最优,导致所求的路径并非最好。对原有的算法进行优化处理一直是目前学者们所关心的问题。文献[5]为了解决三维路径规划问题中耗时长,过早失去解的多样性的问题,引入了逆转变异算子和插入变异算子来对蚁群算法进行改进,文献[6]改进了蚁群算法全局信息素更新策略和信息素增量的计算方法,且在启发式函数引入了安全性因素,实现了机器人的高效避障。文献[7]考虑了启发函数中可行节点到目标点的距离因素,提高了搜索方向性和算法搜索的速度。文献[8]将环境中的局部的路径信息加入到信息素的初始化和路径选择概率中,提高了算法收敛速度,同时引入交叉操作来改进蚁群算法。本文基于以上研究基础,对蚁群算法中的启发式进行了设计,同时考虑到信息素的更新策略,全局信息更新

上综合考虑了蚂蚁迭代次数与最优路径间的关系进行了改进。并利用 MATLAB 软件进行仿真实验,通过分析实验结果来验证算法改进后的可行性。

1 环境建模

本文采用空间等分网格法进行空间抽象建模。首先构造一个立方体区域的规划空间  $ABCD-A'B'C'D'$ , 将三维地图包含住。将立方体左下角的顶点  $A$  作为三维空间的原点, 以此建立一个三维坐标系, 其中  $x$  轴定义为沿精度的增加的方向,  $y$  轴定义为沿纬度增加的方向,  $z$  轴定义为垂直  $xoy$  平面的方向, 这样就构成了一个三维路径空间, 接下来采用等分空间的方法从这个三维空间中抽取出三维路径规划中需要的网格点。先沿着边  $AB$  把这个规划空间进行  $n$  等分, 就得到了  $n+1$  个平面  $\prod_i (i = 0, 1, 2, \dots, n)$ , 然后对这  $n+1$  个平面沿着  $AD$  边进行  $m$  等分, 再沿边  $AA'$  进行  $m$  等分, 且求出里面的相关交点。

这样就将整个规划空间离散化成了一个三维点的集合, 设为  $T^*$ , 集合空间中任意一点与两坐标对应, 即  $T(i, j, k) (i = 0, 1, 2, \dots, n; j = 0, 1, 2, \dots, m, k = 0, 1, 2, \dots, m)$  为序号坐标, 其对应的位置坐标为  $T(x_i, y_j, z_k)$ , 其中  $i, j, k$  分别为其对应的划分序号。本文定义任意两点间  $q_i, q_{i+1}$  的欧式距离为路径最短距离。

2 改进的蚁群算法

2.1 信息素

(1) 信息素表示 通常在路径规划的过程中, 会选择把任意两相邻的节点之间的路径存储信息素, 使其作为载体。本文是在三维空间中进行路径规划, 模型中的节点比较多, 如果将它们各节点之间的路径作为载体的话, 对于前面的网格划分的平面, 每一个平面具有  $m^2$  个节点, 相邻的两个平面节点连线就会有  $m^4$  种情况, 平面划分的间隔越小, 导致空间的复杂度就越大。因此, 本文选择节点表示法, 将信息素存在离散点上, 即每个点含有一个信息素值, 信息素值大小与对蚂蚁吸引的程度成正比, 因此相邻的平面间只会有  $2m^2$  种连接的情况, 从而降低了信息存储的空间需求, 并降低了信息存储的空间需求。

(2) 初始信息素的分布 普通蚁群算法初始阶段路径的信息素都是均匀分布的, 这样蚂蚁在任一节点处选择其他可行方向的概率几乎一样, 由此使得初期的搜索分布较广, 具有一定的盲目性。经过较长的时

间才能通过路径信息素的增长来确定可行路径大致分布, 造成了蚁群算法在寻路初期中耗费了大量时间来确定初始可行解。因此, 为了提高算法的初期收敛速度, 需要对算法的初始路径信息素值的分布进行相应的改变, 不能对路径的所有节点使用相等均匀的信息素。本文考虑了寻路过程中是从起点到末点的一条路径, 因此依靠起始点和终点建立三维空间中的一条空间直线方程, 通过求每个可行节点到该直线方程的欧式距离  $l_{ijk}$ , 根据距离值大小, 在初始起点和终点的理想直线路径适当的给予附近可行节点的初始信息素增加, 其中  $\tau'_0$  为一固定信息素常数:

$$\tau_0 = \left(1 + \frac{1}{l_{ijk}}\right)\tau'_0 \tag{1}$$

2.2 路径选择规则的改进

蚂蚁由当前节点移动到下一个节点的概率与启发式函数相关联, 设有  $M$  只蚂蚁, 而且将它们统一放置在起始点, 在路径的选择过程中, 蚂蚁  $k$  按照式(2)决定接下来的移向点:

$$q_{a+1} = \begin{cases} \operatorname{argmax}[\tau_{a+1} \cdot \eta_{a,a+1}] & q \leq q_0 \\ J & q > q_0 \end{cases} \tag{2}$$

式中:  $q$  为  $[0, 1]$  区间的随机数,  $q_0$  为  $[0, 1]$  区间均匀分布的随即数,  $\eta_{a,a+1}$  为两相邻平面之间节点距离,  $J$  为根据式(3)给出的概率分布产生的一个随即变量。对  $M$  只蚂蚁进行初始化后, 所有蚂蚁都位于初始起点的位置。接下来, 对位于平面  $\prod_a$  上的任意位置点  $q_a(i_a, j_a, k_a)$  的蚂蚁  $k$  如何选择下一个相邻平面的  $q_{a+1}(i_{a+1}, j_{a+1}, k_{a+1})$ , 根据节点的信息素值和启发值来进行一定概率的选择, 选择规则如下:

$$q_{a,a+1} = \begin{cases} \frac{\tau_{a+1} \cdot F}{\sum \tau_{a+1} \cdot F} & \text{可行点} \\ 0 & \text{else} \end{cases} \tag{3}$$

式中  $F$  为启发式值(由式(7)决定),  $\tau_{a+1}$  表示某个平面节点  $a+1$  上的信息素值, 它在某种程度上影响着蚂蚁算法在合适的时间找到全局最优解的效率。

较好的启发式函数<sup>[9]</sup>的设计, 在三维路径规划中起着一定的重要性, 蚁群算法的好坏收敛性和其稳定程度都与之相关联。要想规划出一条较优的三维路径, 须构造一个合适的启发函数。而且, 在三维路径规划中要综合考虑各种相关因素对路径规划的影响, 把路径最短距离和避障作为规划中评价函数, 障碍物的约束函数如下:

$$\eta_1 = \begin{cases} 1 & q_{a+1}(i_{a+1}, j_{a+1}, k_{a+1}) \in allowed(i, j, k) \\ 0 & \text{else} \end{cases} \tag{4}$$

蚂蚁在寻路的过程中, 考虑到最短路径原则, 即将

此刻所在平面的节点和其任意相邻平面的节点间的欧式距离作为两个节点的路径最短距离  $d$ , 相邻两点路径的启发函数为  $\eta_2$ :

$$\eta_2 = \frac{1}{d_{q_i q_{i+1}}} \quad (5)$$

尽管蚂蚁在相邻的节点倾向于与它距离最近的点,但是这会导致选择的下一个节点未必与目标点相近。因此为了让算法在搜索上具有一定的方向性,综合考虑当前节点与相邻点以及目标点的关系,定义两个影响因子  $r_1$ 、 $r_2$ ,用它们来衡量  $q_i$ 、 $q_{i+1}$  与终点  $q_n$  之间的分配:

$$\eta_3 = \frac{1}{r_1 d_{q_i q_{i+1}} + r_2 d_{q_{i+1} q_n}} \quad (6)$$

经过多次实验发现  $r_1$ 、 $r_2$  在算法搜索的不同时段影响程度不一样。初始阶段蚂蚁选择的下一节点与目标点的距离对整个算法的好坏起到很大作用,算法前期的  $r_2$  值越大,倾向于最优方向的路径就越多,算法后期  $r_1$  越大,搜索到的相邻点路径间距越小且靠近目标点。将  $r_1$  取为  $d_{q_{i+1} q_n}$  与  $d_{q_i q_n}$  的比值,  $r_2 = 1 - r_1$ , 这样随着后面的待选节点越来越靠近终点,路径搜索在目标点最优与相邻两点最短自适应的改变,使得搜索向着有利方向前进。

因此文中的设计的启发函数为:

$$F = \eta_1 \times \eta_2 \times \eta_3 \quad (7)$$

### 2.3 三维路径的信息素更新策略

(1) 局部信息素更新的设计 本文利用局部信息素更新和全局信息素更新相结合方式对此进行改进。局部信息素更新在路径搜索过程中进行:

$$\tau_{ijk} = (1 - u)\tau_{ijk} + u\tau_0 \quad (8)$$

式中:衰减系数  $u$  是一个介于 (0, 1) 的参数值,  $\tau_0$  是节点的初始信息素值。通过减少经过点的信息素浓度,提高蚂蚁搜索未知点的概率。

(2) 全局信息素更新 为了提高蚂蚁的全局搜索能力,本文在这里利用一种正反馈信息素增量进行改进。每轮结束后统计所有路径的平均值,当下的蚂蚁搜索到的路径长度值如小于之前所有蚂蚁搜索路径的长度的综合最优值,那么就增强这条路径上的信息素分布,从而使得接下来的蚂蚁继续搜索时能在这个基础标准上找到更好的路径。如果当前的蚂蚁搜索路径长度大于之前所有的蚂蚁搜索路径的平均路径值的话,则不保持信息素的增加,同时这里还引入一种惩罚机制,对这条路径的信息素进行有条件的减少,从而使得接下来进行的搜索避免了这些较差的路径。所有蚂蚁每轮搜索全图完成一次迭代后,都会产生一个

最优的解  $L_{\text{best}}$  和一个最差的解  $L_{\text{wrost}}$ , 可以利用这两个解寻找靠近最优解的蚂蚁来对全局搜索过程的信息素浓度公式进行更新:

$$\tau_{ijk} = (1 - \rho)\tau_{ijk} + \rho\Delta\tau'_{ijk} \quad (9)$$

$$\rho = \begin{cases} 0.25 & nc \in [0, 0.5NC_{\text{max}}] \\ k & nc \in (0.5NC_{\text{max}}, NC_{\text{max}}] \end{cases} \quad (10)$$

$$\Delta\tau'_{ijk} = \sum_{l=1}^M \Delta\tau_{ijk} \quad (11)$$

$$\Delta\tau_{ijk} = \begin{cases} \frac{Q}{L^a} \times \frac{L_d - L_{\text{pest}}}{L_{\text{pest}}} & L_a \leq L_{\text{ave}} \\ -\frac{Q}{L^a} \times \frac{L_{\text{wrost}} - L_d}{L_{\text{wrost}}} & L_a > L_{\text{ave}} \end{cases} \quad (12)$$

式中:  $\rho$  为信息素挥发系数;  $Q$  为一个恒定信息素浓度值;  $L^a$  为第  $a$  只蚂蚁找到的路径长度;  $L_d$  为当前循环的最佳路径长度值;  $L_{\text{ave}}$  为最优解和最差解的平均值。这里用来衡量新产生的路径的信息素浓度增减,以往的基本蚁群算法中的挥发系数  $\rho$  都是固定不变的一个值,这会导致算法的寻优能力不足,容易出现局部的最优情况。因此,本文根据搜索次数以及每次搜索的最优路径值、最差路径值,对挥发系数  $\rho$  进行相对应的改变<sup>[11]</sup>,如式(11)引入一个参数  $k$ ,令  $k$  为每次路径的最优路径值与最差路径值的比值,加入到信息素更新公式。前期  $\rho$  较小,能够很好保留路径上信息素,可以确定最优解的大致方向,使得后续蚂蚁沿着这些方向继续寻优,从而缩小了最优路径与最差路径间的信息素浓度差距。中后期当最优路径与最差路径相差不大时,说明搜索的解越来越集中在最优解的周围,此时为了防止算法过早进入收敛状态,  $\rho$  的值会相应提高,使得算法后期搜索能力增强,从而依靠最优解与最差解的关系自适应地更新全局信息素值。这能有效地提高全局搜索的能力,避免了局部收敛和早熟的出现。通过这些信息素的优化更新,提高了对路径搜索的多样性,同时也能提高算法的收敛速度,增进了对最优路径的寻找。

### 2.4 本文的算法流程设计

1) 建立抽象的环境模型,设置在该模型中的起始点位置、目标点位置,确定蚂蚁的搜索方向,将所有蚂蚁置于起始点。初始化蚂蚁群算法的迭代次数  $NC_{\text{max}}$ , 蚂蚁的个数  $M$ , 以及按照式(10)设置  $\rho$  的值。

2) 每只蚂蚁根据启发式信息值和信息素值,由式(2)、式(3)确定其下一个可行路径点。当每只蚂蚁找到一条路径后,按式(8)进行信息素的局部更新。

3) 重复步骤2,直至本轮所有蚂蚁都完成路径的搜索。

4) 本轮循环后,求取此次循环的  $M$  条蚂蚁路径的平均值,并以此平均值为标准对较优路径和较差路径进行区分。

5) 蚂蚁按式(9)、式(11)、式(12)进行全局的信息素更新,判断算法的迭代次数是否达到  $NC_{\max}$ ,若满足条件,则输出最优路径的解,结束算法。若无,返回步骤2,继续执行算法,直至满足最后迭代次数。

2.5 仿真实验

本文采用一台 PC 机(操作系统:Win7,2 GB 内存),基于 MATLAB2013 实验平台,进行仿真,随机生成一个三维地形图。规划空间抽象为  $21 \times 21 \times 21$  的栅格,设置寻找路径的起点和终点以及高度值,种群个数为 21,最大迭代次数为 300,用上述的改进蚁群算法进行寻路,由图 1、图 2 路径曲线图看,本文改进的算法路径规划效果优于基本蚁群算法的路径规划。

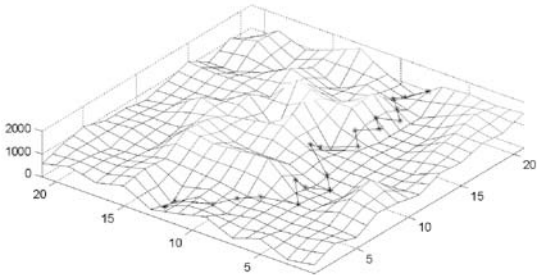


图 1 改进的蚁群算法路径图

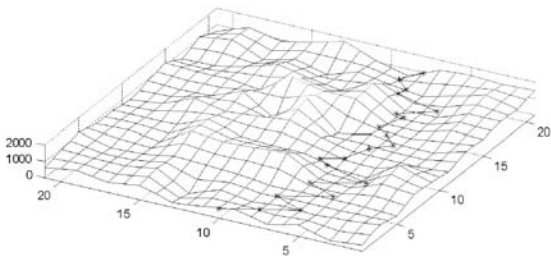


图 2 基本蚁群算法路径图

本文算法最大迭代值设为 300,分别进行多次实验,统计两种算法的迭代次数,并进行比较。随机选取了两组不同算法下的路径长度值与时间作对比,实验结果见表 1,改进的蚁群算法在路径规划所得长度上比基本蚁群算法要短,说明改进的算法中的启发式函数发挥了很大作用,一定程度上降低了蚂蚁不合理地移动到下一相邻节点可能性的选择,图 1 中的路线上的每一个点相比于图 2,离终点的距离都相对较近。在时间效率上(见表 2),本文的改进算法也得到了优化,由于初始信息素的优化分布,以及自适应挥发因子的改变和信息素的惩罚机制,在一定程度上都对蚂蚁搜索的效率进行了提高,使得蚂蚁搜索的更有合理性。

表 1 两种算法路径长度结果比较 km

比较项/km	基本蚁群算法	改进蚁群算法
实验 1	46.7	34.3
实验 2	48.8	36.7
实验 3	50.2	39.4

表 2 两种算法对应时间的比较 s

比较项	基本蚁群算法	改进蚁群算法
实验 1	42.5	35.2
实验 2	46.3	33.6
实验 3	47.5	37.9

比较两种算法寻得最优路径的迭代次数,由表 3 可知,改进算法在迭代速度明显优于基本算法,由于信息素更新规则的改进,使得改进算法在路径搜索的效率得到了极大的提升,不再是对路径信息素盲目地简单改变,而是有条件地对路径上的信息素进行提升或降低。

表 3 两种算法性能比较

比较项(迭代次数)	基本蚁群算法	改进蚁群算法
实验 1	210	163
实验 2	205	160
实验 3	201	164

由图 1 中路径曲线来看,相比于图 2 的路径曲线的跌宕,本文搜索的路径方向大致是一条偏向于终点的且局部路径较平稳的曲线。

3 结 语

本文对三维环境路径规划问题进行了探讨,运用蚁群算法进行路径规划,对算法中的启发式函数进行改进,考虑当前节点到下一节点和目标点之间的关系,引入节点影响因子。在信息素更新策略方向上,利用当前循环的最优路径以及最差路径信息素,引入了信息素增量的惩罚机制,结合信息素挥发因子的自适应取值方式,以提高算法的收敛速度和寻优能力。通过在三维模拟环境中进行实验仿真,改进的算法在搜索得到的路径,时间效率相比基本蚁群算法均得到了改善,表明了本文改进的算法在一定的条件下能达到良好的路径规划效果。随着科技的进步与发展,路径规划所面临的环境充满复杂与不确定性,单靠一种算法的路径规划可能难以取得理想的效果,良好的环境建

## 参 考 文 献

- [1] Shannon C E. A mathematical theory of communication[J]. Bell Labs Technical Journal, 1948, 27(4):379-423.
- [2] Rizzo L. Effective erasure codes for reliable computer communication protocols[J]. ACM SIGCOMM Computer Communication Review, 1997, 27(2):24-36.
- [3] Tang Dan. Research of Methods for Lost Data Reconstruction in Erasure Codes over Binary Fields[J]. Journal of electronic science and technology, 2016, 14(1):43-48.
- [4] Hafner J L, Deenadhayalan V, Rao K K, et al. Matrix methods for lost data reconstruction in erasure codes[C]//Conference on Usenix Conference on File & Storage Technologies. 2005:14-14.
- [5] Hafner J L, Deenadhayalan V, Kanungo T, et al. Performance metrics for erasure codes in storage systems[J]. Chaos An Interdisciplinary Journal of Nonlinear Science, 2004, 18(3):150-156.
- [6] Huang C, Xu L. STAR: An efficient coding scheme for correcting triple storage Technologies[Z]. Berkeley, CA: USE-NIX Association, 2008:97-110.
- [7] Blaum M, Brady J, Bruck J, et al. EVENODD: an optimal scheme for tolerating double disk failures in RAID architectures[J]. IEEE Transactions on Computers, 1995, 44(2):245-254.
- [8] Corbett P, English B, Goel A, et al. Row-diagonal parity for double disk failure correction[C]//Proceedings of the 3rd USENIX conference on File and storage technologies. USE-NIX Association, 2004:1-1.
- (上接第 278 页)
- 模也是路径规划中的重中之重,因此如何构建复杂不稳定的三维环境下的路径规划有待进一步的深入研究。
- ## 参 考 文 献
- [1] 孙炜,吕云峰,唐宏伟,等. 基于一种改进 A\* 算法的机器人路径规划[J]. 湖南大学学报(自然科学版), 2017, 44(4):94-101.
- [2] 张毅,代恩灿,罗元. 基于改进遗传算法的机器人路径规划[J]. 计算机测量与控制, 2016, 24(1):313-316.
- [3] 徐飞. 基于改进人工势场法的机器人避障及路径规划研究[J]. 计算机科学, 2016, 43(12):293-296.
- [4] 钱夔,宋爱国,章华涛,等. 基于自适应模糊神经网络的机器人路径规划方法[J]. 东南大学学报(自然科学版), 2012, 42(4):637-642.
- [5] 李向军,霍艳丽,曾劼炜,等. 三维机器人路径规划的一种变异算子蚁群算法[J]. 计算机仿真, 2015, 32(2):364-368.
- [6] 吴玉香,王超. 一种改进的机器人三维路径规划方法[J]. 华南理工大学学报(自然科学版), 2016, 44(9):53-60.
- [7] 朱颢东,孙振,吴迪. 基于改进蚁群算法的机器人三维路径规划[J]. 华中师范大学学报(自然科学版), 2016, 50(6):812-817.
- [8] 张琦,马家辰,谢玮,等. 基于改进蚁群算法的机器人路径规划[J]. 东北大学学报(自然科学版), 2013, 34(11):1521-1524.
- [9] 赵开新,魏勇,王东署. 改进蚁群算法在机器人路径规划中的研究[J]. 计算机测量与控制, 2014, 22(11):67-70.
- [10] Tiwari P K, Vidyarthi D P. Improved auto control ant colony optimization using lazy ant approach for grid scheduling problem[J]. Future Generation Computer Systems, 2016, 60:78-89.
- [11] 许明乐,游晓明,刘升. 基于统计分析的自适应蚁群算法及应用[J]. 计算机应用与软件, 2017, 34(7):204-211.
- (上接第 294 页)
- [4] 李楠. 基于自然语言处理的医院智能客服系统的设计与研究[J]. 中国数字医学, 2017(8):85-87.
- [5] 肖红,毛明扬. 一种基于 nlp 中文分词系统的中文分词工具:CN 106354714 A[P]. 2017.
- [6] 吴轲. 基于深度学习的中文自然语言处理[D]. 东南大学, 2014.
- [7] 葛彩霞. 基于医学语言处理的面向任务的信息提取框架构建及应用[D]. 浙江大学, 2015.
- [8] 宋晓波. 医疗诊断报告自动信息提取系统设计与开发[D]. 复旦大学, 2009.
- [9] 陈之彦,李晓杰,朱淑华. 基于 Hash 结构词典的双向最大匹配分词法[J]. 计算机科学, 2015, 42(b11):49-54.
- [10] 结巴中文分词项目[EB/OL]. 2012-09-29[2013-01-05]. <https://github.com/fxsjy/jieba>.
- [11] Sun M, T'Sou B K. Ambiguity Resolution in Chinese Word Segmentation[C]//10 Th Pacific Asia Conference on Language. 1995:384-389.
- [12] Ayer T, Alagoz O, Chhatwal J, et al. Breast cancer risk estimation with artificial neural networks revisited[J]. Cancer, 2010, 116(14):3310-3321.
- [13] Zheng X, Chen H, Xu T. Deep learning for Chinese word segmentation and POS tagging[C]//Conference on Empirical Methods in Natural Language Processing. 2013.
- [14] Boonkwan P, Supnithi T. Bidirectional Deep Learning of Context Representation for Joint Word Segmentation and POS Tagging[M]//Advanced Computational Methods for Knowledge Engineering. 2018.
- [15] Gu Y, Yang Z, Xu G, et al. Exploration on efficient similar sentences extraction[J]. World Wide Web-internet & Web Information Systems, 2014, 17(4):595-626.
- [16] 罗燕,赵书良,李晓超. 基于词频统计的文本关键词提取方法[J]. 计算机应用, 2016, 36(3):718-725.
- [17] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation[J]. J Machine Learning Research Archive, 2003, 3:993-1022.