

基于空间聚类 and 蚁群算法的车辆路径问题的研究

朱锦新

(盐城工学院 实验教学部, 江苏 盐城 224051)

摘要:针对蚁群算法在解决车辆路径问题(VRP)上易陷入局部最优解的缺陷,首先利用加权K-means算法对客户进行区域划分,再利用蚁群算法对每个区域进行求解,实验结果表明方法具有良好的性能。

关键词:车辆路径问题;蚁群算法;聚类分析;改进K-means算法

中图分类号:TP301.6 **文献标识码:**A **文章编号:**1671-5322(2009)04-0044-04

车辆路径问题(Vehicle Routing Problem, VRP)^[1]自1959年由Danting和Ramser提出以来,一直是网络优化问题中最基本的问题之一,由于其应用的广泛性和经济上的重大价值,一直受到国内外学者的广泛关注。VRP是一个NP-hard问题,针对这个难题国内外学者建立了各种数学模型,提出了许多启发式和非启发式算法。Clarke和Wright首先提出一种启发式节省法来建立车队配送路线。Willard首先利用禁忌搜索算法用来求解VRP,随后亦有许多位学者也发表了求解VRP的TS算法。西南交通大学的袁庆达等设计了考虑时间窗口和不同车辆类型的禁忌算法。Osman对VRP的模拟退火算法进行了研究,提出的模拟退火方法主要适合于解决路线分组。现在多数学者采用混合策略,分别采用两种人工智能方法进行路线分组和路线优化^[2-4]。

蚁群算法在解决VRP上易陷入局部最优解,算法开始时收敛速度比较快,但在算法进行过程之中,当迭代到一定次数后,蚂蚁也可能在某个或某些局部最优解的邻域附近发生停滞现象,与其他求解组合优化问题的算法相比,蚁群算法在迭代次数和解的质量上都有一定的优势,但搜索时间上明显落后于其它算法。因此,本文为了克服蚁群算法的缺陷,首先利用加权K-means算法对客户进行区域划分,之后再利用蚁群算法对每个区域进行求解。

1 空间聚类

在进行物流配送时,每天的配送量都很大,一般都需要多车辆对客户进行配送。即先根据车辆载重量等各种约束条件进行配送区域划分,然后在各小区域内设计最优配送路线,而且不同配送区域的划分方法对最后结果影响很大,所以确定一种比较合理的划分算法对整个系统是比较关键的,该算法应达到以下的效果:

(1)被划分的区域不会重叠,相互之间没有覆盖;

(2)同一区域之间的个体相对集中,这样每一辆车辆运送物资时,不会消耗相对过多的时间在路程上;

(3)每一区域的运货量大致相同,不应该出现某一车辆运送的物资特别多或特别少的现象;

(4)如果可能,应该是使每一辆车所走过的总路程大致相等;

划分聚类中K-means法是最著名、最为常用的,相对易于理解。于是,本文采用一种改进K-means聚类算法来完成物流配送区域的划分。由于在基本k-means算法中只考虑了两维坐标向量,即仅仅是坐标聚类,而忽略了纬度的参考信息,该信息表现在物流中,就是有着至关重要的汽车载重量信息完全被忽视了。所以,在实践中,为了保证每一聚类中的客户需求量不会超过一辆车的载货量,在每个客户加入时候都会检查总的载

收稿日期:2009-07-07

作者简介:朱锦新(1978-),男,江苏盐城人,助教,南京理工大学在读硕士,主要研究方向为模式识别和计算机网络。

货量,如果超出要求则会弹出一个牺牲客户。改进后的加权 K-means 算法考虑到了节点的权重,一方面,在聚类过程中每次都计算了节点所加入的聚类的总重量,并通过弹出节点的方法对总重量进行控制,使得其保持在一个相对合理的范围;另一方面,又为了不让某个弹出点总是被弹出,设置了一个标记,防止聚类的抖动。

2 蚁群算法

蚂蚁利用信息素来寻找路径。当碰到一个没走过的路口时,就随机挑一条,同时释放与路径长度有关的信息素。蚂蚁走的路径越长,释放的信息量就越少,反之亦然。随后的蚂蚁在经过该路口时,选择信息量较大路径的概率就比较大,这样就形成一个正反馈机制。于是,最优路径上的信息量就越来越大,而其他路径上的信息素则会逐渐削减,最终整个蚁群会找出最优路径。同时,蚁群也能适应环境的变化,当出现障碍物时,蚂蚁也能很快的找到新的最优路径^[5]。

2.1 蚁群算法的数学模型

为了便于说明和理解,本文以 TSP 问题为例,引入以下符号:设 m 是蚂蚁的数量, n 是城市的数量, $d_{ij}(i, j=1, 2, \dots, n)$ 表示城市 i 和 j 之间的距离, $b_i(t)$ 表示 t 时刻位于城市 i 的蚂蚁个数 $m = \sum_{i=1}^n b_i(t)$, $\tau_{ij}(t)$, 表示 t 时刻在 ij 连线上残留的信息量。初始时刻,各条路径上的信息量相等,设为 $\tau_{ij}(0) = C$ (C 为常数)。蚂蚁 $k(k=1, 2, \dots, m)$ 在运动过程中,根据各条路上的信息量决定转移方向, $p_{ij}^k(t)$ 表示在 t 时刻蚂蚁 k 从 i 转移到 j 的概率:

$$p_{ij}^k(t) = \begin{cases} \frac{\tau_{ij}^\alpha(t) \eta_{ij}^\beta(t)}{\sum_{s \in allowed_k} \tau_{is}^\alpha(t) \eta_{is}^\beta(t)} & j \in allowed_k, \\ 0 & otherwise \end{cases} \quad (1)$$

其中, $allowed_k = \{1, 2, \dots, m\} - tabu_k$ 表示蚂蚁 k 下一步允许选择的的城市。与实际蚁群不同,人工蚁群系统具有记忆功能, $tabu_k(k=1, 2, \dots, m)$ 用以记录蚂蚁 k 当前所走过的城市,集合 $tabu_k$ 随着进化过程做动态调整。 α 为信息启发式因子,表示轨迹信息量的相对重要性,它反映了蚂蚁在运动过程中所积累的信息素在蚂蚁运动时所起的作用,其值越大,则该蚂蚁越倾向于选择其他蚂蚁经过的路径,表明蚂蚁之间协作性越强; β 为期望启发式因子,表示能见度的相对重要性,反映了蚂蚁在运动过程中启发信息在蚂蚁选择路径中的受重视

程度,其值越大,则该状态转移概率越接近于贪婪规则; $\eta_{ij}(t)$ 为启发函数, $\eta_{ij}(t) = 1/d_{ij}$ 表示由城市 i 转移到城市 j 的期望程度。对于蚂蚁 k , d_{ij} 越小, $\eta_{ij}(t)$ 越大, $p_{ij}^k(t)$ 也就越大。

2.1.1 状态转移规则

一只位于城市 i 的蚂蚁通过公式(2)给出的规则来选择下一个将要移动到的城市 s 。

$$s = \begin{cases} \arg \max \{ [\tau_{is}(t)] [\eta_{is}(t)]^\beta \} & q \leq q_0 \\ S & otherwise \end{cases} \quad (2)$$

其中, q 为均匀分布在 $[0, 1]$ 上的一个随机变量, q_0 为在 $[0, 1]$ 上的参数, S 是根据式(1)计算出来的概率分布来进行选择,这种规则被称为伪随机比例规则,类似于遗传算法中的“轮盘选择”。

2.1.2 信息素更新规则

为了避免残留信息素过多引起残留信息覆盖启发信息,在每只蚂蚁走完一步或者完成对所有 n 个城市的遍历(即一次循环结束)后,要对残留信息进行更新处理。这种更新策略模仿了人类大脑记忆的特点,在新信息不断存入大脑的同时,存储在脑海中的旧信息随着时间的推移逐渐淡化,甚至忘记。

(1) 全局更新规则

在蚁群算法中,只有最优蚂蚁才被允许释放信息素。这种选择,以及伪随机比例的规则的适用,其目的都是为了使搜索过程中更具有指导性:蚂蚁的搜索主要集中在当前循环为止所找出的最好路径的领域内。全局更新根据公式(3)在所有的蚂蚁都完成它们的路径之后执行:

$$\begin{aligned} \tau_{ij}(t+1) &= (1-\rho) \cdot \tau_{ij}(t) + \rho \cdot \Delta \tau_{ij}(t) \\ \Delta \tau_{ij}(t) &= \begin{cases} 1/L_{gb} & \text{若 } (i, j) \in \text{最优路径} \\ 0 & \text{否则} \end{cases} \end{aligned} \quad (3)$$

其中, ρ 为全局信息素挥发因子, $0 < \rho < 1$; L_{gb} 表示到目前为止找到的全局最优解或者本次迭代中蚂蚁找到的最优解。

(2) 局部更新规则

在每只蚂蚁构建路径的过程中,蚂蚁每经过一条边 (i, j) ,都将立刻调用公式(4)对自己所走的每一步进行信息素更新:

$$\tau_{ij}(t+1) = (1-\xi) \tau_{ij}(t) + \xi \cdot \tau_0$$

其中, ξ 是局部信息素挥发因子, $0 < \xi < 1$; τ_0 是信息素的初始值,为常数。局部更新的作用在于:蚂蚁每一次经过边 (i, j) ,该边的信息素 τ_{ij} 就会减少,从而使得其它蚂蚁选中该边的概率相对减少。也就是说,这将增加探索未使用过的边的机会,使

得算法不会陷入停滞状态。

2.2 蚁群算法的实现步骤

步骤 1 参数初始化。令时间 $t=0$ 和循环次数 $N_c=0$, 设置最大循环次数 $N_{\max}$, 将 m 只蚂蚁置于 n 个城市上, 令图中每条边 (i, j) 的初始化信息量 $\tau_{ij}(t)$ 为常数, 且初始时刻 $\Delta\tau_{ij}(0)=0$ 。

步骤 2 根据公式(1)计算转移概率;

步骤 3 根据计算出的转移概率和随机产生的 q 值, 按照公式(2)为每一只蚂蚁选择下一条移动的路径;

步骤 4 当每一只蚂蚁都走过一条边到达下一城市后, 就按公式(4)的局部更新规则对这条边进行一次信息素的局部更新;

步骤 5 对每一只蚂蚁重复以上循环执行步骤 2 到步骤 4, 直到每一只蚂蚁都生成一条包含全部城市的路径;

步骤 6 在生成的全部路径中找出最短的一条路径, 则走过该路径的蚂蚁就是最优蚂蚁;

步骤 7 对最优蚂蚁所经过的每一条边, 按公式(3)的全局更新规则对这条路径进行一次信息素的全局更新;

步骤 8 重复执行步骤 2 到步骤 7, 直到执行次数达到指定的最大迭代次数或连续若干代内没有更好的解出现为止。

3 VRP 实例结果和实验分析

本文先利用加权 K-means 算法对客户进行区域划分, 之后再利用蚁群算法对每个区域进行求解, 来解决 VRP 问题。以 VRPLIB 标准数据库中的 Eil51 作为基本仿真数据。经大量运行测试, 并与其它算法作了比较, 得到了较好的结果。Eil51 的描述如下: 0 号为仓库和 1—50 号为客户点, 每个点的坐标及需求量如表 1 所示, 仓库共有 5 辆车, 每辆车的载重量均为 160 kg。

运用加权 K-means 算法与蚁群算法进行路线优化, 其中, 蚁群算法的参数设置如下: $\alpha=1$; $\beta=5$; $m=25$; $N_{\max}=200$; $\rho=0.6$; $q_0=0.8$; $\xi=0.1$ 。将程序运行了 10 次, 每次的结果如表 2 所示。

从表 2 的数据中可以看出: 第 5 次运行已经得到了最优解, 10 次的平均值为 541.072, 其中最优解的路线图如图 1 所示, 每次运行的最短总距离均接近于最优解, 证明了此算法的有效性。

对同样的问题, 利用单纯的蚁群算法和文献[6]所提供的遗传算法, 各运行了 10 次, 所得结果分别如表 3 和表 4 所示。

由表 2、表 3 和表 4, 及图 2 可以看出, 利用本

表 1 Eil51 的点坐标及各点需求量

Table 1 Point coordinate of Eil51 and requirements of point coordinate

点	横坐标	纵坐标	需求量
0	30	40	0
1	37	52	7
2	49	49	30
3	52	64	16
4	20	26	9
5	40	30	21
6	21	47	15
7	17	63	19
8	31	62	23
9	52	33	11
10	51	21	5
11	42	41	19
12	31	32	29
13	5	25	23
14	12	42	21
15	36	16	10
16	52	41	15
17	27	23	3
18	17	33	41
19	13	13	9
20	57	58	28
21	62	42	8
22	42	57	8
23	16	57	16
24	8	52	10
25	7	38	28
26	27	68	7
27	30	48	15
28	43	67	14
29	58	48	6
30	58	27	19
31	37	69	11
32	38	46	12
33	46	10	23
34	61	33	26
35	62	63	17
36	63	69	6
37	32	22	9
38	45	35	15
39	59	15	14
40	5	6	7
41	10	17	27
42	21	10	13
43	5	64	11
44	30	15	16
45	39	10	10
46	32	39	5
47	25	32	25
48	25	55	17
49	48	28	18
50	56	37	10

文算法 10 次运行得到的结果均优于使用单纯蚁群算法和遗传算法所得的结果, 而且第 5 次还得到了最优解, 再有, 利用本文算法的平均值要比利用其他两种算法得到的平均值要低。所以, 利用

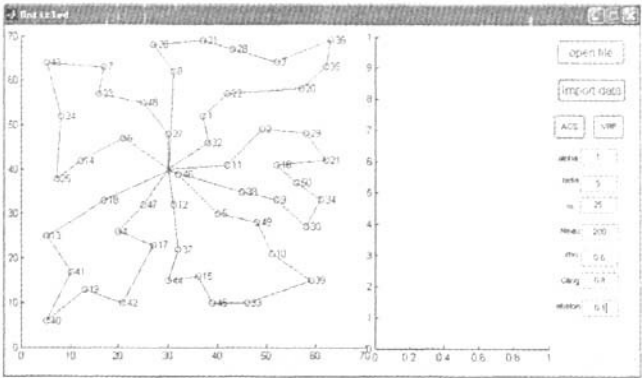


图 1 加权 k-means 与蚁群算法相结合解决 VRP 问题效果图

Fig.1 View of VRP Based on the improved K-means and Ant Colony Algorithm

表 2 运行结果

Table 2 Experiment Results

计算次序	1	2	3	4	5	6	7	8	9	10	平均值
最短总距离	534.023	556.127	545.683	554.639	524.611	528.588	527.146	530.670	557.273	551.959	541.072

表 3 单纯蚁群算法运行结果

Table 3 Experiment Results of Ant Colony Algorithm

计算次序	1	2	3	4	5	6	7	8	9	10	平均值
最短总距离	564.822	557.651	566.579	568.572	530.670	544.591	549.876	534.023	570.088	566.037	555.291

表 4 遗传算法运行结果

Table 4 Experiment Results of GA

计算次序	1	2	3	4	5	6	7	8	9	10	平均值
最短总距离	547.766	567.348	552.150	557.651	544.591	530.670	534.023	549.876	572.076	562.428	551.858

本文所提出的算法——加权 K-means 和蚁群算法相结合,可以方便有效地、更好地求得问题的最优解或近似最优解。

4 结论

首先应用聚类算法对配送区域进行划分,并对所采用的 K-means 聚类算法进行了改进和实现。在此基础上,将路径规划问题抽象为旅行商问题(TSP)的一种模式,并在抽象模式上再引入改进蚁群算法。并通过仿真实验,研究分析了蚁群算法在求解中小规模 TSP 问题上,各参数的最佳设置。并且对实验结果进行分析,证明该实验的有效性和正确性。

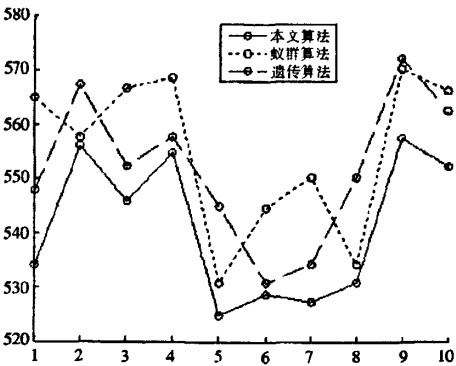


图 2 本文算法和遗传算法的运行结果比较图

Fig.2 Result's View of Compared with GA

(下转第 59 页)

- [5] Xiaogang Wang, Eric Grimson. Spatial Latent Dirichlet Allocation. Proceedings of Neural Information Processing Systems (NIPS2007). 2007[EB/OL]. Http://books.nips.cc/papers/files/nips20/NIPS2007_0964.pdf.
- [6] McCallum A, Corrada - Emmanuel A, Wang X. Topic and role discovery in social networks[A]. Proceedings of 19th Joint conference on artificial intelligence. 2005.
- [7] Thorsten Brants, Francine Chen, Ioannis Tsochantaris. Topic - based document segmentation with probabilistic latent semantic analysis[A]. Proceedings of the eleventh international Conference on Information and knowledge management McLean, Virginia, USA. 2002. 211 - 218.
- [8] Thomas Minka, John Lafferty. Expectation - Propagation for the Generative Aspect Model[A]. Uncertainty in Artificial Intelligence, 2002.

Web Text Classification based on LDA Model

MENG Hai-tao¹, CHEN Si², ZHOU Rui²

(1. School of Information Technology Yancheng of Institute Technology, Jiangsu Yancheng 224051, China;
2. International Department Beijing Daxing No. 1 Middle School, Beijing 102600, China)

Abstract: A kind of web text classification is put forward on the basis of LDA model. Latent Dirichlet Allocation (LDA) is an unsupervised topic learning model which extracts latent topics from text data. Parameters are estimated with Gibbs sampling of MCMC and the word probability is represented. Thus different latent topics are associated with observable words. Contrasting to SVM and Bayesian Network, the result in the experiment shows that LDA has the better performance than any other algorithm.

Keywords: Latent Dirichlet Allocation(LDA); topic model; WEB classification

(责任编辑:沈建新;校对:张英健)

(上接第47页)

参考文献:

- [1] Paolo Toth, Daniele Vigo. The Vehicle Routin Problem[A]. Society for Industrial and Applied Mathematics philadelphia. 2002.
- [2] Clarke G, Wright J. Scheduling of vehicles from central depot to a number of delivery points[J]. Operations Research, 1964, 12: 568 - 581.
- [3] 袁庆达, 闫昱, 周再玲. Tabu Search 算法在优化配送路线问题中的应用[J]. 计算机工程, 2001(11): 86 - 89.
- [4] Osman I H. Meta. strategy simulated annealing an Tabu search algorithms for the vehicle routin problem[J]. Annu Oper Res, 1993, 41: 77 - 86.
- [5] 段海滨. 蚁群算法原理及其应用[M]. 北京: 科学出版社, 2007.
- [6] 弓晋丽, 程志敏. 基于 matlab 物流配送路径优化问题遗传算法实现的实现[J]. 物流科技, 2005, 29(131): 103 - 105.

Research and Practice of VRP Based on the Improved K - means and Ant Colony Algorithm

ZHU Jin-xin

(School of Department of Experiment Teaching, Yancheng Institute of technology, Jiangsu Yancheng 224051, China)

Abstract: A study is made on VRP. To avoid long - time searching, precocity and stagnation and tendency to local optimization of traditional ant colony algorithm. First, improved K - means is applied to divide regions of customs, ant colony algorithm is applied to solve the problem in each region. Experiments indicate that the proposed method has good performance.

Keywords: VRP; Ant Colony Algorithm; Clustering Analysis; improved K - means

(责任编辑:张英健;校对:沈建新)