

基于遗传算法的物流配送车调度优化算法*

An Optimal Algorithm of Generic Algorithm Based on Logistics Delivery in E-business

吴璟莉^{1,2},李陶深¹

Wu Jingli^{1,2},Li Taoshen¹

(1. 广西大学计算机与电子信息学院,广西南宁 530004;2. 广西师范大学数学与计算机科学学院,广西桂林 541004)

(1. School of Comp. and Elec. Info., Guangxi Univ., Nanning, Guangxi, 530004, China; 2. School of Math. and Comp. Sci., Guangxi Normal Univ., Guilin, Guangxi, 541004, China)

摘要:介绍电子商务物流信息平台中配送车调度问题的解决方法,阐述遗传算法在物流配送问题上的应用缺点,并提出一种基于遗传算法的物流配送车调度的优化算法。

关键词:电子商务 物流配送 车调度 遗传算法

中图分类号:TP393 文献标识码:A 文章编号:1002-7378(2005)04-0209-03

Abstract: The solution of scheduling transportation problems about logistics delivery in the E-business information platform is introduced, and the application of the generic algorithm on logistics delivery and the algorithm's design and implement is discussed.

Key words: e-business, logistics delivery, scheduling transportation, genetic algorithm

电子商务物流配送是指物流配送企业采用计算机网络技术和现代化的硬件设备、软件系统及先进的管理手段,针对社会需求,严格地、守信用地按用户的订货要求,进行一系列分类、编配、整理、分工、配货等理货工作,定时、定点、定量地交给没有范围限度的各类用户,满足其对商品的需求。在电子商务物流配送中,建立可靠而高效的物流配送系统是简化业务流程,降低企业运作成本,增强市场竞争力的有利保障^[1]。开发物流配送软件,通过计算机实现快速、合理地安排运输路线和运输车次具有重要的意义。

电子商务物流配送系统通过物流信息平台实现货物配送。物流信息平台是信息传入载体和用户接入手段,包括各种各样的物理传送介质和传递方式,它将物流平台信息中心、网上银行、商家、客户和各个配送网点的通信介质集成在一起。该平台的业务流程^[1,2]是:(1)客户通过 Internet 访问物流信息平台的网站,填写送货单;(2)物流信息平台认证客户

身份;(3)客户通过输入银行信用卡账号和网上专用密码进行支付;(4)物流信息平台将支付信息提交网上银行;(5)网上银行返回扣款结果,网上交易结束;(6)物流信息平台中心服务器进行订单派送调度;(7)订单配送至相应配送网点,各配送网点安排运输车次与运输路线,配送货物,这是物流配送中的配送车调度问题。本文介绍电子商务物流信息平台中配送车的调度问题的解决方法,阐述遗传算法在物流配送问题上的应用,及其具体的算法设计与实现。

1 物流配送车的调度问题

1.1 概述

物流配送车的调度问题可以这样描述^[3]:一个配送网点,对于物流中心分派下来的 n 张订单,要用 m ($1 \leq m \leq M$, M 为最大允许的车辆数) 辆车把订单上的货物送到指定的送货点,车辆从配送网点出发,分别旅行一条路线,使得每个送货点有且仅有一辆车作一次访问,最后回到配送网点。问题的关键是如何合理安排车辆数目和车辆路线,使得总的旅行路程最短。

以便于讨论,对问题做几点假设:(1)被配送的是已知的同一种物资;(2)各用户的所在地已知;

收稿日期:2005-06-16

作者简介:吴璟莉(1976-),女,广西人,助教,主要从事电子商务、遗传算法的研究。

* 广西自然科学基金项目(桂科自 0229008)和广西教育厅科技项目(桂教科研[2001]401号)联合资助。

(3) 各用户的需求量已知; (4) 从配送网点到各用户之间的运输距离已知; (5) 配送网点有足够的资源可以供应配送, 并且拥有足够的配送能力。另外, 配送计划中的最优发车路线, 必须符合下列约束条件: (1) 配送必须满足所有用户的需求; (2) 对每一辆发送车的装载量有一定的限制, 不允许超载运行; (3) 对发送车每天的总运行时间(或总运行距离)有预定的上限; (4) 必须满足用户提出的到货时间要求。对一个具体的问题, 上述约束条件可能全部存在, 也可能只存在一部分。解决配送问题就是在以上约束条件下应如何派送车辆, 给出车辆数、型号和各车辆的具体行车路线。订单上的货物全部送到即可完成当日的运输任务, 又可以使总运输公里数最小。

1.2 问题的分析

配送车调度问题可以说是对巡回旅行商问题(TSP, Traveling Salesman Problem)加以一定的限制而形成的, 属于约束性的多重 TSP 问题(CMTSP, Constrained Multiple Traveling Salesman Problem)^[3,4]。因为有多辆车可用, 除去实际的配送网点, 我们假设几个虚拟的配送网点, 使之与配送车一一对应, 然后把全部发送路线划分为几条大的路线, 每一辆发送车只沿一条大路线送货, 即只访问配送网点一次。例如设有配送网点(编号 0)需运货物到其余编号为 1, 2, ..., N 的 N 个送货点, 有 M 辆汽车可用, 设置 $M - 1$ 个虚拟配送网点(实际就是同一个配送网点, 编号分别为 $N + 1, N + 2, \dots, N + M - 1$), 并根据实际意义设置距离矩阵。这样就把 CMTSP 问题转化为 $(N + M)$ 个点的 TSP 问题, 这里我们只讨论对称 TSP 问题。

下面首先给出几个路径定义:

定义 1 合法子路径: 从配送网点出发, 经由其他送货点后回到配送网点(不包括)的地点序列或单独一个配送网点。

定义 2 合法遍历路径: 不包括重复地点(配送网点除外)的几条合法子路径组成的序列。

定义 3 合法遍历方案: 子路径满足有关约束条件的合法遍历路径。

设 $d(x, y)$ 和 $R(x)$ 分别表示 x, y 间的距离和配送网点需运送到送货点 x 的货物量, Q 表示汽车运送货物容量, D 表示汽车运送货物允许的最大行程距离。这样, 配送车调度问题就是要设置若干条(小于或等于 M) 子路径, 使得遍历路径总长度最短, 要求每条子路径上的货物量之和不超过 Q , 同时每条子路径的长度不大于 D 。设 X 是一个 $N + M$ 维

向量, 用于表示合法遍历路径, 则 X 可用下式表示:

$$X = (x_0, x_1, \dots, x_{N+M-1}),$$

其中, $\forall k \in [0, N + M - 1], \exists i \in [0, N + M - 1], x_i = k, x_0 = 0$, 则 CMTSP 问题可用数学模型描述如下:

$$\begin{aligned} \min F(x) &= \sum_{i=0}^{N+M-2} d(x_i, x_{i+1}) + d(x_{N+M-1}, x_0) \\ \text{s. t } &\sum_{i=\rho l+1}^{\rho l+1-1} R(x_i) \leq Q, \\ &\begin{cases} x_{\rho l} > N, p_0 = 0, p_{M+1} = 0, \\ p_l < p_{l+1}, (l = 0, 1, \dots, M - 1), \end{cases} \\ &\sum_{i=\rho l}^{\rho l+1} d(x_i, x_{i+1}) \leq D. \text{ 同上。} \end{aligned}$$

2 基于遗传算法的物流配送车调度的优化算法

目前解决 TSP 问题的优化方法很多, 如模拟退火算法^[5]、神经网络方法^[3]、节约法^[3]和遗传算法^[5]等。遗传算法是一种能处理复杂问题的、鲁棒性强的启发式随机搜索方法。但是传统简单的遗传算法存在着收敛速度慢、易陷入次全局最优点等缺点。文献^[6]提出一种知识提高遗传算法, 它将启发式方法和遗传算法结合起来, 用小规模种群来解决大规模的 TSP 问题, 有效地避免了局部最优解的问题。我们改造文献^[6]中的 k -最近邻接子图的构造方法以及种群初始化的方法, 使其更适于由 CMTSP 问题转化而成的 TSP 问题, 而提出一种基于遗传算法的物流配送车的优化调度的算法。

在算法的设计上, 为了提高编码效率, 我们采用标准化路径表示, 让路径中的第一个地点总是 0。例如: 序列(0, 1, 2, 3)表示一条从地点 0 到 1, 2, 3, 最后返回 0 的路径, 那么不同的序列(0, 1, 2, 3), (2, 3, 0, 1), (1, 2, 3, 0)都代表相同的路径(0, 1, 2, 3)。具体算法如下:

步骤 1: 遗传代数 $gen = 0$, 初始化种群 $P = \{S_1, S_2, \dots, S_n\}$, S_i 是第 i 个个体, 种群规模为 n 。

步骤 2: 运用交叉算子来产生后代: $O_c = \{S_{n+1}, \dots, S_{n+j}\}$ 。

步骤 3: 运用变异算子来产生后代(每个变异算子采用它当前的变异率):

$$O_m = \{S_{n+j+1}, \dots, S_{n+j+k}\}.$$

步骤 4: 计算每个个体 S_i 的适应值 $f(S_i)$ 。

步骤 5: 计算每个变异算子变异率的平均进步值 $Progress(M_l), l = 1$ 到 m (m 是变异算子的个数), 假

设一个父代 p 通过变异产生一个子代 a , 那么这个变异算子的的进步值计算如下:

$$Progress = \text{Max}(f(p), f(a)) - f(p)$$

假设这一代中 r 个父代被这个变异算子选中, 这个变异算子 M_1 的平均进步值 $Progress(M_1)$ 为这 r 个进步值的平均值。

步骤 6: 根据适应值选择 n 个优良的个体。

步骤 7: 根据各参选变异算子的平均进步值调整它们的变异率。

步骤 8: 如果不满足结束条件返回步骤 4, $gen = gen + 1$, 否则停止计算。

在算法实现时, 重点考虑以下几个方面。

(1) 染色体编码。通过增加虚拟配送网点, 采用合法遍历方案的遍历地点次序排列作为编码, 假设 7 是虚拟配送网点, 则码串: 01237456 (一个配送网点, 6 个送货点) 表示从配送网点 0 开始, 依次经送货点 1, 2, 3, 回到配送网点 0, 再依次访问 4, 5, 6, 最后回到配送网点 0 的遍历路径。

(2) 局部初始化。对于大规模 TSP 问题, 如路径序列数为 n , 如果我们采用传统的初始化方法, 即用单纯的随机产生方法来产生种群, 则需要种群的大小在 n 和 $2n$ 之间, 那么种群的规模将非常大, 需要大量的计算时间。我们根据文献[6]提出的方案, 生成一个小且具有代表性的初始种群。文献[6]中, 在包含点 c_i 的边集中选取 k 个较短的边来构成 c_i 点的 k -最近邻接子图, 这是一个边的集合。本文的问题中, 有 $M-1$ 个点是虚拟的, 它们实际上和点 0 是同一个点, 如果不加约束地沿用文献[6]中的做法, 会导致网点的 k -最近邻接子图中有许多点是网点的虚拟点。为此, 我们增加了约束: k -最近邻接子图中的边长必须大于零; 网点的 k -最近邻接子图中只包含从网点到送货点的边; 送货点的 k -最近邻接子图只包含送货点到送货点的边。这样局部初始化时, k -最近邻接子图中的边有较高的优先级。例如, 从点 c 出发, 首先在它的 k -最近邻接子图中随机选出下一个点, 如果 k -最近邻接子图中的所有点都已经被用过了, 才在所有没用过的点中随机选一个。初始化染色体工作流程为: ① 如果还有未被配送的送货点, 则随机选一个网点, 根据其 k -最近邻接子图找到它的第一个送货点, 否则退出, 染色体初始化结束; ② 检验此时是否到达或超过网点的配送能力或最大行程数, 如果是则返回 ①, 否则到 ③; ③ 根据 k -最近邻接子图找到它的下一个送货点, 返回 ②。

(3) 交叉算子。交叉操作通过交换选定的 2 个个体的一些基因来产生后代。这里给定两个标准化路径表示的父代, P_1 和 P_2 , 第一个后代的构造方法是: 从一个随机产生的地点 c 出发, 然后检查起点为 c 的边或者终点为 c 的边是否在 P_1 和 P_2 中, 如果在则选取这条公共边, 否则将 P_1 和 P_2 中以 c 为起点的边进行比较, 选短的那条, 如果短的那条会产生一条回路, 则选择长的那条, 如果长的那条也产生回路, 则沿用局部初始化中阐述的方法, 在没有选过的点中随机选择 1 个点。第 2 个后代的构造方法与此相似, 不同的是比较边长时, 将 P_1 和 P_2 中以 c 为终点的边进行比较。

(4) 变异算子。变异算子通过在个体中随机改变一个或多个基因来生成后代, 使后代具有一些与父代不同的特征。变异的目的在于使搜索方向多样化, 有效抵制局部最优解, 提高找到全局最有解的可能性。我们采用文献[7]提出的动态变异遗传算法(DMGA)的有关思想, 根据不同的遗传阶段自动地选择适当的变异算子来产生后代。具体思路是: 同时采用几种变异算子, 初始化时, 几种变异算子的变异率是相同的, 经过一代或几代后, 每个算子的变异率自动地、动态地调整。这样, 适宜的变异算子(通过适应值来衡量)的变异率提高, 不适宜的变异算子的变异率降低, 通过选择变异率大的变异算子来产生后代。我们选取 2-opt 和 3-opt 启发式变异算子作为 DMGA 中的参选变异算子, 将 2-opt 和 3-opt 启发式变异结合起来使用, 将比单独使用其中任何一个要好得多, 而且不会增加运算时间[7]。

(5) 选择机制。选择算子是遗传算法实现进化的操作, 其目的是使适应度高的个体具有更大的生存概率, 并有效地控制种群规模。在算法中我们采用锦标赛选择方法。

(6) 适应值函数。适应值函数是用来区分种群中个体好坏的标准, 是算法的演化过程的驱动力, 是自然选择的唯一依据。设遍历路径长度为 T_d , 不满足约束条件的子路径段数为 N , 如果某条子路径超过约束条件很多, 可算作 2~3 段。我们在算法中将适应值函数设计为:

$$f = \frac{c}{T_d + \alpha * N},$$

其中, c 为调整系数, α 为惩罚系数。

(下转第 218 页)

$$f(x) = \langle w, \varphi(x) \rangle + b, \tag{24}$$

$$\max - \frac{1}{2} \sum_{i,j=1}^r (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle \varphi(x_i), \varphi(x_j) \rangle - \epsilon \sum_{i=1}^r (\alpha_i + \alpha_i^*) + \sum_{i=1}^r y_i (\alpha_i - \alpha_i^*). \tag{25}$$

约束条件不变,从而得到

$$w = \sum_{i=1}^r (\alpha_i - \alpha_i^*) \varphi(x_i). \tag{26}$$

在支持向量机中,引入核函数 Kernel Function^[3]来简化非线性逼近.核函数 $k(x, x')$ 满足

$$k(x, x') = \langle \varphi(x), \varphi(x') \rangle,$$

这样(24)式变为

$$\max - \frac{1}{2} \sum_{i,j=1}^r (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k(x_i, x_j) - \epsilon \sum_{i=1}^r (\alpha_i^* + \alpha_i) + \sum_{i=1}^r y_i (\alpha_i^* - \alpha_i), \tag{27}$$

而(22)式变为

$$b = \frac{1}{N} \left\{ \sum_{0 < \alpha_i < C} [y_i - \sum_{x_j \in SV} (\alpha_j - \alpha_j^*) k(x_j \cdot x_i) - \epsilon] + \sum_{0 < \alpha_i^* < C} [y_i - \sum_{x_j \in SV} (\alpha_j - \alpha_j^*) k(x_j \cdot x_i) + \epsilon] \right\}. \tag{28}$$

由(24)、(26)、(28)式可得计算回归估计函数为

$$\hat{f}(x) = \sum_{x_i \in SV_s} (\alpha_i - \alpha_i^*) k(x_i \cdot x) + b. \tag{29}$$

由(27)式可知,尽管通过非线性函数将样本数据映射到具有高维甚至为无穷维的特征空间,但在计算回归估计函数时并不需要显式计算该非线性函数,而只需计算核函数,从而避免高维特征空间引起的维数灾难问题.核函数 $k(x, x')$ 是对称正实数函

数,且必须满足 Merce 条件.

4 讨论

由于统计学习理论和支持向量机建立了一套较好的有限样本机器学习的理论框架和通用方法,具有严格的理论基础,能够较好地解决小样本、非线性、高维数和局部最小点等问题,因此成为 20 世纪 90 年代末发展最快的研究方向之一.本文深入推导了用于解决回归估计问题的SVM方法,与其它方法相比,SVM具有泛化性强、效率高等特点.但由于 SVM 是一种新技术,其发展仅有十多年时间,还有一些问题需要深入研究,如受核矩阵存储空间约束的支持向量机适应大样本的情况、核函数类型及其相应参数的选择、输入样本映射到高维空间后是否为线性可分的判定等问题.

参考文献:

[1] Vapnik V N. The Nature of Statistical Learning Theory[M]. New York: Springer-Verlag, 1995.
 [2] Cherkassky V, Mulier F. Learning Form Data: Concepts, Theory and Methods[M]. New York: John Wiley & Sons, 1997.
 [3] Nello Cristianinni, John Shawe-Taylor[英]. 支持向量机导论[M]. 李国正,王 蒙,曾华军译.北京:电子工业出版社,2005. 100-103, 70-80, 24-45.
 [4] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32-42.
 [5] 边肇祺, 张学工. 模式识别[M]. 第 2 版. 北京: 清华大学出版社, 2000. 284-303.
 [6] 杜树新, 吴铁军. 用于回归估计的支持向量机方法[J]. 系统仿真学报, 2003, 15(11): 1580-1585.

(责任编辑:黎贞崇)

(上接第 211 页)

3 结束语

目前,我国的电子商务正在蓬勃发展,正是建立和发展电子商务物流体系的黄金时期,可靠的、高效的物流配送系统是电子商务应用研究中的重要组成部分.本文提出的一种基于遗传算法的物流配送车的优化调度算法,在研制物流配送软件过程中,通过将遗传算法应用于配送车的调度问题,可以实现快速、合理地安排运输路线和运输车次,能够取得较好的应用效果.

参考文献:

[1] 文 岗. 电子商务时期的第三方物流管理[M]. 北京: 中国商业出版社, 2000.

[2] 王海龙,王行愚. 一种基于配送体系物流信息平台的研究[J]. 计算机应用研究, 2001, 18(7): 32-34.
 [3] 蔡希贤,夏士智编译. 物流合理化的数量方法[M]. 武汉:华中工学院出版社, 1985.
 [4] 陈 龙. 基于遗产算法的约束性多 TSP 问题及其应用[J]. 重庆邮电学院学报, 2000, 12(2): 67-69, 74.
 [5] 李陶深. 人工智能[M]. 重庆:重庆大学出版社, 2002.
 [6] Rong Yang. Solving large travelling salesman problems with small populations [EB/OL]. <http://citeseer.nj.nec.com/145347.html>.
 [7] Hong Tuangpei, Wang Hongshuang, Chen Weichou. Simultaneously applying multiple mutation operators in genetic algorithms[J]. Journal of Heuristics, 2000, 6: 439-455.

(责任编辑:邓大玉)