

# 道路网匹配的蚁群算法求解模型

巩现勇<sup>1,2</sup> 武 芳<sup>1</sup> 姬存伟<sup>1</sup> 翟仁健<sup>1,2</sup>

1 信息工程大学地理空间信息学院,河南 郑州,450052  
2 地理信息工程国家重点实验室,陕西 西安,710054

**摘 要:**利用蚁群算法的群体优势,寻找全局最优的道路网同名实体匹配方案。首先从几何矢量误差和结构特征两方面建立了匹配问题的数学约束模型;然后阐述了蚁群算法求解匹配问题的基本原理,设计了问题求解模型,并引入自适应和局部搜索策略提高了算法效率;最后给出了求解的关键步骤。实验证明,利用蚁群算法进行道路网匹配是有效、可行的,为求解匹配问题提供了新思路。

**关键词:**数据更新;同名实体匹配;蚁群算法;组合优化

**中图法分类号:**P208;P283.1      **文献标志码:**A

如何进行空间数据集成、融合、更新和共享,进行有效的空间分析利用和决策,已成为当前亟待解决的问题。同名实体匹配,作为空间数据集成和融合的关键技术之一,已经成为数据更新、影像配准、航位与姿态估计等诸多应用研究的一个必然需求和瓶颈问题,具有十分重要的科学和应用价值<sup>[1]</sup>。

道路网匹配是当前最为活跃的一个研究方向。文献[2-4]采用缓冲区增长法确定候选弧段集合,再进行匹配判断。文献[5-7]首先对道路弧段的节点进行初始匹配,再迭代匹配节点关联的弧段。冯申等基于版本数据库增量更新的目的分析了道路网变化情况和匹配方法<sup>[8]</sup>。陈军、胡云岗等通过分析多尺度、多时态道路网的分解、基本和抽象等3个匹配层次,提出了基于网眼密度的道路网选取和顾及层次结构的匹配方法,并应用于缩编更新<sup>[9-10]</sup>。以上研究虽然能够融合多个匹配条件,但很难有效地合理平衡多个实体与多个候选对象之间交叉匹配的矛盾,从而无法寻求出一个全局近似最优的匹配方案。道路网匹配实质上是一种全局寻优问题,而国内外鲜有人利用智能算法研究道路网匹配问题。

Dorigo 等通过研究自然界中蚁群搜索路径的行为,提出了一种寻找优化路径的机率型群体智能优化算法——蚁群算法<sup>[11]</sup>(ant colony opti-

mization, ACO),并成功应用于空间数据挖掘、遥感影像处理、组合导航等诸多领域<sup>[12]</sup>。

## 1 匹配问题的数学模型约束

空间实体的相似性描述主要有几何、拓扑、语义等方面。目前匹配算法可大致归为两类:①几何特征的单实体级判别法<sup>[6,13]</sup>,常见几何信息包括距离<sup>[4,13]</sup>、形调函数、方向<sup>[4,14]</sup>等;②结构特征辅助几何信息的特征级判别法<sup>[4,7-9,15]</sup>,例如节点的拓扑连通度<sup>[7]</sup>、具有良好连通性的Stroke<sup>[15]</sup>、层次特征<sup>[8-9]</sup>、尺度影响<sup>[16]</sup>。特征级匹配充分考虑了地理学第一定理,顾及具有空间关联的不同实体之间的相互影响和作用,相对于单实体匹配方法而言,是一种十分健壮的全局协同寻优匹配方法,它具有令人满意的统计特性,灵活性好、鲁棒性强。因此,本文从几何矢量误差约束和结构特征约束两个方面来描述空间实体之间的相似性。

### 1.1 几何矢量误差约束

几何图形本身是道路网中最基础的、质量最有保障的信息。本文采用长度、匹配节点之间的差距和匹配弧段的方向3个参数来约束几何矢量误差。

假设存在一对潜在匹配关系(SL,TL),相邻节点构成实体的一段弧段 $S_i^o S_{i+1}^o$ 或 $T_i^o T_{i+1}^o$ ,设m

$>n, \|S_i^o S_{i+1}^o\| > \|T_i^o T_{i+1}^o\|$ 。假设 $S_i^o$ 和 $T_i^o$ 已确定为匹配关系,从 $S_i^o$ 出发,在弧段 $S_i^o S_{i+1}^o$ 上行走长度为 $\|T_i^o T_{i+1}^o\|$ 的距离,在当前位置插入虚拟节点 $S_{i+1}$ ,如图1 空心点所示。依次交替对 $(SL, TL)$ 进行上述"加密"处理,可得新的节点序列: $SL = \langle S_1, S_2, \dots, S_i, \dots, S_p \rangle, TL = \langle T_1, T_2, \dots, T_i, \dots, T_q \rangle$ ,如图1 所示。 $(SL, TL)$ 的几何矢量误差约束定义如下<sup>[6-7,14]</sup>:

$$d(SL, TL) = \frac{1}{R} \sum_{i=1}^R \|S_i T_i\| \tag{1}$$

$$len(SL, TL) = 1 - \min \left\{ \frac{\|SL\|}{\|TL\|}, \frac{\|TL\|}{\|SL\|} \right\} \tag{2}$$

$$agl(SL, TL) = \sum_{i=1}^{R-1} w_i \times \Delta agl(S_i S_{i+1}, T_i T_{i+1}) \tag{3}$$

式中, $w_i$ 为方向约束的权重因子,其定义为:

$$w_i = \frac{\|S_i S_{i+1}\| + \|T_i T_{i+1}\|}{\|SL\| + \|TL\|} \tag{4}$$

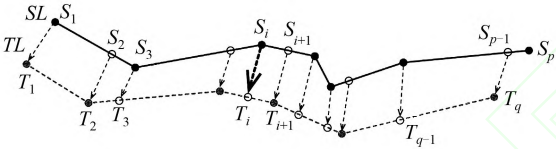


图1 几何矢量误差相似性计算

Fig.1 Calculation of Geometric Error Similarity

1.2 结构特征约束

参考社会网络分析中角色重要性程度的评价方法,定义道路的节点连通度为:与道路网中一个节点直接关联的弧段的数目。用节点连通度来描述 $(SL, TL)$ 的拓扑结构的差异,即不一致度,如下式:

$$tp(SL, TL) = 1 - \frac{1}{R} \sum_{i=1}^{R-1} \frac{\text{Min\_TP}(S_i, T_i)}{\text{Max\_TP}(S_i, T_i)} \tag{5}$$

式中, $\text{Min\_TP}()$ 和 $\text{Max\_TP}()$ 分别表示节点对 $(S_i, T_i)$ 的最小和最大连通度。如图2 中,1 号、2 号和3 号节点连通度分别为2、3、4。则1 号与2 号节点的拓扑结构不一致度为 $1 - \text{Min\_TP}(2, 3) / \text{Max\_TP}(2, 3) = 1/3$ ,1 号与3 号的拓扑结构不一致度为 $1 - \text{Min\_TP}(2, 4) / \text{Max\_TP}(2, 4) = 1/2$ 。

2 基于蚁群算法的匹配求解模型

基于蚁群算法的匹配问题求解模型的基本思想为将当前图层的 $n$  个实体看作城市节点。相邻节点间有路径通行,路径是待匹配图层中当前节

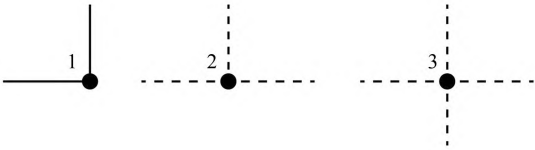


图2 节点连通度

Fig.2 Vertex Connection Degree

点的候选匹配实体集,其"代价"是同名实体的相似性目标函数。其值越大,则蚂蚁选择此路径的概率越大。蚂蚁的每一次跳转,构成一次匹配关系。利用蚁群算法,寻找使得匹配相似度函数全局最大的一条全局最优路径,此时蚂蚁走过的路径序列即为所求解的匹配关系。

2.1 候选列表和匹配函数

位于节点 $i$  的蚂蚁 $k$  能直接访问节点的集合,由待匹配实体的候选列表确定。其定义为实体 $i$  在一定半径的缓冲区内所查询到的待匹配图层的实体集 $list_i^k$ 。对 $list_i^k$  中每个实体的4 个匹配约束参数进行纵向归一化处理为 $N_d$ 、 $N_{len}$ 、 $N_{angle}$ 和 $N_{tp}$ ,并加权形成相似度函数 $F(SL, TL)$ ,也即匹配函数:

$$F(SL, TL) = 1 - a \times N_d - b \times N_{len} - c \times N_{angle} - d \times N_{tp} \tag{6}$$

式中, $a$ 、 $b$ 、 $c$ 、 $d$ 为权重参数。 $F(SL, TL)$ 越大,表明 $SL$ 和 $TL$ 越相似。为避免局部最优,降低时间复杂度,在数据预处理时,采用禁忌表策略,将相似性小于阈值 $\delta$ 的实体屏蔽,禁忌其访问性。

2.2 信息素

信息素 $\tau_{ij}$ 值越大,算法正反馈性和自组织性越强,但全局搜索能力和鲁棒性降低。信息素的初始值为:

$$\tau_{ij} = \begin{cases} 1 / \sqrt{\text{Count}(\text{list}_i^k)}, & \text{其他} \\ 0, & \text{Count}(\text{list}_i^k) = 0 \end{cases} \tag{7}$$

其更新包括蒸发和释放两部分。基于排列规则思想更新信息素,在每次迭代过程中,只有生成了目前为止最优路径的 $m$  只蚂蚁和排列在最前的 $m-1$  只蚂蚁才能释放信息素:

$$\tau_{ij} \leftarrow (1 - \rho) \tau_{ij} + \sum_{r=1}^m (m-r) \times \Delta \tau_{ij}^r + m \times \Delta \tau_{ij}^{\text{best}} \tag{8}$$

式中, $\rho$  ( $0 \leq \rho \leq 1$ ) 定义为挥发系数,可以采用自适应策略来控制。设置一个阈值 $\rho_{\min}$ ,则可以令 $\rho = \max(\exp(-0.05t), \rho_{\min})$ , $t$ 为当前迭代次数; $\Delta \tau_{ij}^{\text{best}}$ 是生成当前最优路径的蚂蚁释放的信息素; $\Delta \tau_{ij}^r$ 为本次迭代中相似度函数值的顺序统计量排列在第 $r$  的蚂蚁释放的信息素:

$$\Delta \tau_{ij}^r = 1 - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{[F(X,Y)_k]^2}{2}\right),$$
$$(i,j) \in T^k \tag{9}$$

式中, $T^k$  为蚂蚁 $k$  的路径; $F(X,Y)_k$  为蚂蚁 $k$  在本次循环中访问的路径所对应的相似度函数的总和。

### 2.3 启发式信息

启发式信息  $\eta_{ij}$  是关于匹配关系的一个局部信息,与目标函数成正比。 $\eta_{ij}$  的值越大,其对应的节点被选择作为匹配对象的概率越大。其定义为:

$$\eta_{ij} = \frac{F(SL,TL)_k - \delta}{\delta} \tag{10}$$

### 2.4 转移概率

在蚂蚁访问路径过程中,位于节点 $i$  的蚂蚁 $k$ ,根据伪随机原则选择所访问路径的下一个节点,其转移概率为:

$$P_{ij}^k = \frac{\xi \tau_{ij} + (1 - \xi) \eta_{ij}}{\sum_{j \in list_i^k} [\xi \tau_{ij} + (1 - \xi) \eta_{ij}]} \quad j \in list_i^k \tag{11}$$

式中, $\xi(0 \leq \xi \leq 1)$  为固定参数,表示  $\tau_{ij}$  和  $\eta_{ij}$  的相对影响。

## 3 算法步骤

根据以上求解模型,利用蚁群算法求解同名道路匹配关系的步骤如下:

- 1) 数据预处理。构建实验数据的拓扑关系,统一空间基准。通过缓冲分析获取每个实体的候选列表,确定相似度函数  $F(SL,TL)$ 。求出匹配关系的禁忌表。
- 2) 蚁群算法初始化。设定最大搜索次数  $T$ 、蚂蚁总数目  $M$ 、信息素强度  $Q$ 、信息素和启发式信息的相对影响系数  $\xi$ 、信息素挥发系数  $\rho$  等参数的初始值。将  $M$  只蚂蚁随机置放于待搜索路径的节点。
- 3)  $M$  只蚂蚁遍历所有路径节点,获得  $M$  个解,得出潜在的匹配关系。
- 4) 计算相似度函数值,更新信息素。
- 5) 返回步骤3),直到最大搜索次数  $T$  或最优解已平稳收敛。
- 6) 输出最优解路径  $(i,j)$  序列,整理得出实体的匹配关系。

## 4 实验与结果分析

为检验蚁群算法求解匹配问题的有效性,本

文基于 ArcEngine9.3 设计开发了道路网匹配试验系统。试验数据为2000 年和2010 年洛杉矶部分道路网数据,来源是美国人口普查局使用的地图数据 TIGER (topologically integrated geographic encoding and referendng)。借鉴前人解决组合优化问题的经验<sup>[12]</sup>,试验中部分参数的设置范围如表1。

表1 算法部分参数取值

Tab.1 Partial Parameters of the Algorithm

记号	参数名或意义	取值范围	参考取值
$T$	最大迭代次数	$T \geq 20$	60
$M$	蚂蚁总数目	$M \geq 1$	$\sqrt{n}$
$\xi$	信息素和启发式信息的相对影响系数	$0 \leq \xi \leq 1$	0.4
$\rho$	信息素蒸发率	$0.1 \leq \rho \leq 0.5$	0.2
$\delta$	相似度阈值	$0 \leq \delta \leq 1$	0.6

试验数据区域内,2000 年共有4 679 条路段,2010 年共有5 976 条路段。叠加对比可得出:不同时间的部分道路网在空间位置上存在明显的不一致性;2010 年时部分道路发生了变更、新增和消失等空间变化;2010 年的数据更加详实准确,拓扑合理,现势性更强。按照运行算法,a b c d 权重因子分别取0.3 0.2 0.2 0.3。匹配关系结果、约束参数及相似度函数结果见表2 (限于篇幅,仅随机列出5 条)。部分匹配结果见图3。

表2 部分匹配结果

Tab.2 Partial Matching Results

匹配对 $(i,j)(j,j$ 为道路编号)	$N_d$	$N_{len}$	$N_{angle}$	$N_{tp}$	$F(i,j)$
(23 439.58 886)	0.27	0.13	0.07	0.33	0.78
(23 487.67 318)	0.16	0.04	0.11	0	0.922
(23 372.58 749)	0.21	0.09	0.03	0.12	0.877
(24 338.13 721)	0.07	0.03	0.03	0	0.967
(26 027.48 972)	0.14	0.22	0.12	0.05	0.875

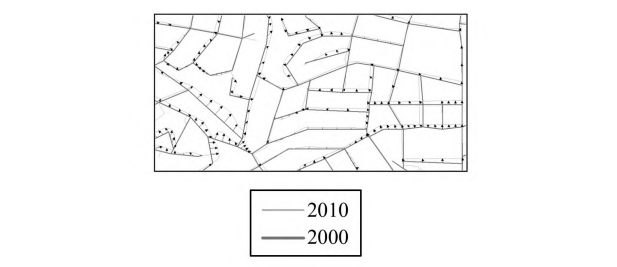


图3 不同时相的数据匹配结果(局部)

Fig.3 Matching Results of Networks at Different Time (Part)

为了检查算法的可靠性,本文通过人眼判读匹配对比的方法,采用随机抽样法评价匹配结果。判读人员为2 名制图生产单位工程师、3 名

助理工程师、1 名地图专业讲师、3 名地图专业学生。每人随机抽取50 对匹配关系进行评价,统计得出本文方法的平均判对率、判错率、漏判率分别为91.38%、5.71%、2.91%。

通过统计分析算法运行结果和判读人员调查数据,可以得出如下两点:

1) 对典型的道路网模式结构,匹配结果相对较为优效(图4)。这是由于典型的道路网模式形态明显,结构规范,相同的模式具有一致的形态结构,在空间认知上遵循视觉分组的原则。例如较为规则的网格模式(图4(a))遵循完整和闭合法则,呈辐射状的星形模式(图4(b))遵循简洁性法则,具有良好连通性的Stroke 模式(图4(c))遵循连续法则等。模式的图形特征与背景的差异较大,易于区分。

2) 对立交桥、多线道路口等复杂结构道路网的匹配结果不够理想。如图5 中,对于图5(a)、

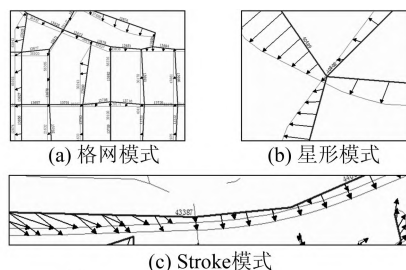


图4 典型道路网结构模式匹配结果

Fig. 4 Matching Results of Typical Road Network Pattern

图5(b)等组成结构变化不大的道路口能够正确识别匹配;对于图5(c)结构变化大、拓扑变形严重的道路口,其匹配关系错误较为明显。究其原因,主要是人类对复杂道路网结构的科学理解和认知有限,这直接导致所选参数不能全面、正确、合理地描述道路网的形态和空间结构,其改进有待于空间认知和计算几何等相关学科的更深层次的应用。

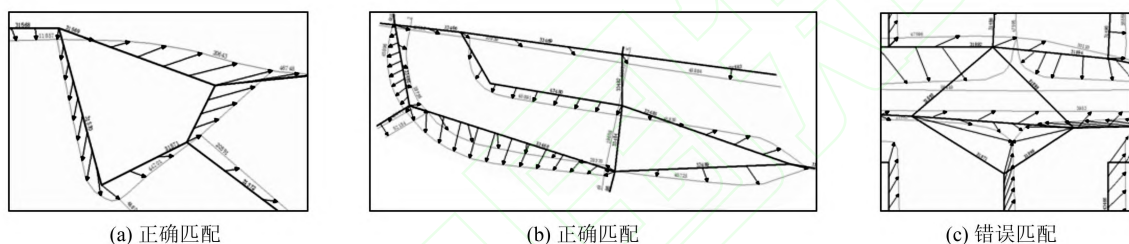


图5 部分道路口匹配结果

Fig. 5 Partial Matching Results of Road Network Crossing

本文利用蚁群算法求解道路网匹配问题,引入智能算法对空间数据匹配进行了有效的尝试。不足之处在于:① 对于立交桥、多边路等复杂道路结构,需要特定的专业知识进一步的研究;② 蚁群算法中的信息素挥发系数、最大搜索次数和蚁群规模等也需要根据经验和多次试验确定。

## 参 考 文 献

- [1] Ruiz J J, Ariza F J, Urena M A, et al. Digital Map Conflation: A Review of the Process and a Proposal for Classification[J]. International Journal of Geographical Information Science, 2011, 25 (9): 1 439-1 466
- [2] Walter V, Fritsch D. Matching Spatial Data Sets: A Statistical Approach[J]. International Journal of Geographical Information Systems, 1999, 13 (5):445-473
- [3] Zhang M, Meng L. An Iterative Road-Matching Approach for the Integration of Postal Data[J]. Computers, Environment and Urban Systems, 2007, 31 (5):597-615
- [4] Tong Xiaohua, Deng Susu, Shi Wenzhong. A Probabilistic Theory-based Matching Method[J]. Acta Geodetica et Cartographica Sinica, 2007, 36 (2): 210-217 (童小华, 邓懋懋, 史文中. 基于概率的地图实体匹配方法[J]. 测绘学报, 2007, 36 (2):210-217)
- [5] Safra E. Location-based Algorithms for Finding sets of Corresponding Objects over Several Geo-spatial Data Sets[J]. International Journal of Geographical Information Science, 2010, 24 (1):69-106
- [6] Song W, James M K, Timothy L H, et al. Relaxation-based Point Feature Matching for Vector Map Conflation[J]. Transactions in GIS, 2011, 15 (1): 43-60
- [7] Zhao Dongbao, Sheng Yehua. Research on Automatic Matching of Vector Road Networks Based on Global Optimization[J]. Acta Geodetica et Cartographica Sinica, 2010, 39 (4):416-421 (赵东保, 盛业华. 全局寻优的矢量道路网自动匹配方法研究[J]. 测绘学报, 2010, 39 (4):416-421)
- [8] Ying Shen, Li Lin, Liu Wanzeng, et al. Change-only Updating Based on Object Matching in Version



- Databases [J]. *Geomatics and Information Science of Wuhan University*, 2009, 34 (6): 752-755 (应申, 李霖, 刘万增, 等. 版本数据库中基于目标匹配的变化信息提取与数据更新[J]. 武汉大学学报·信息科学版, 2009, 34 (6): 752-755)
- [9] Hu Yungang, Chen Jun, Zhao Renliang, et al. Matching of Roads under Different Scales for Updating Map Data [J]. *Geomatics and Information Science of Wuhan University*, 2010, 35 (4): 451-456 (胡云岗, 陈军, 赵仁亮, 等. 地图数据缩编更新中道路数据匹配方法[J]. 武汉大学学报·信息科学版, 2010, 35 (4): 451-456)
- [10] Chen Jun, Hu Yungang, Zhao Renliang, et al. Road Data Updating Based on Map Generalization [J]. *Geomatics and Information Science of Wuhan University*, 2007, 32 (11): 1 023-1 027 (陈军, 胡云岗, 赵仁亮, 等. 道路数据缩编更新的自动综合方法研究[J]. 武汉大学学报·信息科学版, 2007, 32 (11): 1 023-1 027)
- [11] Dorigo M, Bratani M, Stutzle T. Ant Colony Optimization—Artificial Ants as a Computational Intelligence Technique [J]. *IEEE Computational Intelligence Magazine*, 2006, (1): 28-33
- [12] Qin Kun. *Intelligent Spatial Information Processing* [M]. Wuhan: Wuhan University Press, 2009: 165-196 (秦昆. 智能空间信息处理[M]. 武汉: 武汉大学出版社, 2009: 165-196)
- [13] Chen Yumin, Gong Jianya, Shi Wenzhong. A Distance-based Matching Algorithm for Multi-scale Road Networks [J]. *Acta Geodetica et Cartographica Sinica*, 2007, 36 (1): 84-90 (陈玉敏, 龚建雅, 史文中. 多尺度道路网的距离匹配算法研究[J]. 测绘学报, 2007, 36 (1): 84-90)
- [14] Yang B, Zhang Y, Luan X. A Probabilistic Relaxation Approach for Matching Road Networks [J]. *International Journal of Geographical Information Science*, DOI: 10.1080/13658816.2012.683486
- [15] Zhang M, Meng L. Delimited Stroke Oriented Algorithm Working Principle and Implementation for the Matching of Road Networks [J]. *Annals of GIS*, 2008, 14: 44-53
- [16] Mustière S, Devogele T. Matching Networks with Different Levels of Detail [J]. *Geoinformatica*, 2008, 12 (4): 435-453

## Ant Colony Optimization Approach to Road Network Matching

GONG Xianyong<sup>1,2</sup> WU Fang<sup>1</sup> JI Cunwei<sup>1</sup> ZHAI Renjian<sup>1,2</sup>

<sup>1</sup> Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450052, China

<sup>2</sup> State Key Laboratory of Geo-information Engineering, Xi'an 710054, China

**Abstract:** Corresponding feature matching, essentially as a matter of global combinatorial optimization, is one of the key technologies for geospatial data integration, fusion and update. In this paper, a global optimum matching solution is achieved taking the advantages of ant colony optimization groups and random search, without the centralized control and global model. The basic principle of ant colony optimization for road network matching is explained first, with a mathematical constraint model considering both geometric error and structural characteristics. Then, the matching problem solution model is designed, with a self-adaptation and local search strategy employed to improve efficiency. Finally, the key steps are given. Experiments show that the ant colony optimization approach is effective, feasible and practical, providing a new idea for road network matching.

**Key words:** spatial data update; corresponding feature matching; ant colony optimization; combinatorial optimization

First author: GONG Xianyong, post graduate, specializes in pattern recognition, automated cartography generalization and spatio-temporal data analysis. E-mail: gongxygs@whu.edu.cn

Foundation support: The National Natural Science Foundation of China, Nos. 41171354, 41101362, 41171305; the State Key Laboratory of Geo-information Engineering Foundation, No. SKLGE2013-M-4-6.