基于Meta平衡的多Agent Q学习算法研究

王万良[1] 濮约庆[1] 赵燕伟[2]
(浙江工业大学计算机科学与技术学院 杭州 310023)[1]
(浙江工业大学特种装备制造与先进加工技术教育部重点实验室 杭州 310012)[2]

摘 要 多Agent强化学习算法的研究一直以来大多都是针对于合作策略,而NashQ算法的提出对非合作策略的研究无疑是一个重要贡献。针对在多Agent系统中,Nash平衡无法确保求得的解是Pareto最优解及其计算复杂度较高的问题,提出了基于Meta平衡的MetaQ算法。与NashQ算法不同,MetaQ算法通过对自身行为的预处理以及对其它Agent行为的预测来获取共同行为的最优策略。最后通过研究及气候合作策略游戏实验,证明了MetaQ算法在解决非合作策略的问题中有着很好的理论解释和实验性能。

关键词 强化学习,Meta平衡,NashQ,多Agent系统
中图法分类号 TP181 **文献标识码** A

Research on Multi-agent Q Learning Algorithm Based on Meta Equilibrium

WANG Wan-liang[1] PU Yue-qing[1] ZHAO Yan-wei[2]
(College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310023, China)[1]
(Key Laboratory of Special Equipment and Advanced Processing Technology Ministry of Education, Zhejiang University of Technology, Hangzhou 310012, China)[2]

Abstract Multi-agent reinforcement learning algorithms aim at cooperation strategy, while NashQ is frequently mentioned as a pivotal algorithm to the study of non-cooperative strategies. In multi-agent systems, Nash equilibrium can not ensure the solutions obtained Pareto optimal, besides, the algorithm with high computation complexity. MetaQ algorithm was proposed in this paper. It is different from NashQ that MetaQ finds out the optimal solution by the pre-treatment of its own behavior and the prediction of the others behavior. In the end, a game-climate cooperation strategy was used in this paper, and the results shows that MetaQ algorithm, with impressive performance, is fit for non-cooperative problem.

Keywords Reinforcement learning, Meta equilibrium, NashQ, Multi-agent system

1 引言

多Agent Q学习是Q学习算法在多Agent环境中的应用推广,但多Agent强化学习算法的研究一直以来大多数是针对合作策略进行研究,故Hu和Wellman[1]基于一般和随机博弈(general sum stochastic game)框架提出的Nash Q-Learning(NashQ)算法无疑是一个重要的贡献。在Nash Q算法中,多Agent环境被看作是一场一般和随机博弈,而每个Agent都是这场博弈的局中人,它们的即时奖赏总是与当前的状态及它们的联合动作相关的。依据单Agent Q学习的流程,每个Agent都会根据当前状态的Q矩阵选择它的动作,然后用下一状态的Q值来更新当前状态的Q矩阵[2]。NashQ算法提出,动作的选择和Q值更新的依据应当是多Agent系统中每个Agent之间某个特定的Nash平衡[1]。NashQ算法的研究一直没有停止,Littmann在一些特殊策略博弈中通过FFQ(Friend-and-Foe Q)算法取代NashQ算法[3]。Greenwald等人提出了均衡概念,通过联合NashQ算法和FFQ算法提出了CEQ算法[4]。近年来,Pateto最优概念的多目标演化算法成为当前演化计算的研究热点[5],宋梅萍等人则使用Pareto占优解代替非合作的Nash平衡进行学习[6],使得算法偏向于合作策略的求解,但其并不适用于非合作策略的研究。

尽管Hu和Wellman从理论上证明了NashQ算法的收敛性,并以实验证实了NashQ算法具有优于单Agent Q学习的性能[7],但我们发现采用Nash平衡指导多Agent的学习依然存在以下两个缺陷:

(1) Nash平衡求得的解很可能不是Pareto最优解[8]。这意味着Agent更愿意采用某个非Nash平衡的联合策略,使它获得更多的收益。

(2) Nash平衡的计算复杂度较高。Nash平衡的计算通

本文受国家自然科学基金项目(60874074),浙江省重大科技专项(2009C11039)资助。
王万良(1957-),男,博士,教授,博士生导师,主要研究方向为智能自动化、数字媒体、无线传感器网络等,E-mail:wwl@zjut.edu.cn;濮约庆(1987-),男,硕士生,主要研究方向为智能算法;赵燕伟(1959-),女,博士,教授,博士生导师,主要研究方向为物流系统智能配送与优化调度、数字化产品现代设计等。

常可归结为线性互补问题(1inear complementary problem, LCP)[9]，而求解 LCP 的经典 Lemke-Howson 算法一次只能得出一个解，并且在最坏情形下算法复杂度下界达到指数级[10]。

上述缺陷是由于 Agent 的 Nash 理性造成的，Nash 理性要求平衡策略是相互间的最佳反应策略[3]。这种严格最优的要求使 Agent 显得很自私，往往无法达到一个纯策略的平衡，更不要说达到 Pareto 最优。

本文采用 Meta 平衡来替代 Nash 平衡，并且证明了Meta平衡能够满足以下条件：①Meta 平衡总是存在的；②Meta 平衡应当是一个联合纯策略；③博弈至少存在一个 Pareto 最优解；④求解平衡的算法复杂度应相对较低。

2 多 Agent Q 学习算法

一般和博弈 G：具有 n 个 Agent 的一般和博弈是一个元组$(N,A^1,\cdots,A^n,r^1,\cdots,r^n)$，其中 N 是n 个 Agent 的集合；A^i 是 Agent_i 的动作集合，如果 Agent_i 拥有m 种不同的动作，那么 $A^i=\{a^1,a^2,\cdots,a^m\}$；$r^i:A^1\times A^2\times\cdots\times A^n\rightarrow R$ 是 Agent_i 的盈利函数。

随机博弈：具有 n 个 Agent 的随机博弈 Γ 是一个元组$(S,A^1,\cdots,A^n,r^1,\cdots,r^n,p)$，其中 S 为状态空间；A^i 是 Agent_i 的动作空间；$r^i:S\times A^1\times\cdots\times A^n\rightarrow R$ 是 Agent_i 的即时奖赏函数；$p:S\times A^1\times\cdots\times A^n\rightarrow\Delta(S)$是状态转移函数；而 $\Delta(S)$是状态空间 S 上的概率分布。

在一个有折扣的随机博弈 Γ 中，Agent_i 的目的是最大化其带折扣的期望奖赏总和：

$$v(s,\pi^1,\cdots,\pi^n)=\sum_{t=0}^{\infty}\beta^t E(r_t\mid\pi^1,\cdots,\pi^n,s_0=s) \tag{1}$$

式中，$\pi^i:S\rightarrow A^i$ 是 Agent_i 的策略，$\beta\in(0,1]$是折扣因子。式(1)的标准解决方法是通过迭代搜索法即由 Bellman 公式搜索一个固定点，如下：

$$v(s,\pi_*^1,\cdots,\pi_*^n)=\max_{a^1,\cdots,a^n}\{r(s,a^1,\cdots,a^n)+\beta\sum_{s'\in S}p(s'\mid s,a^1,\cdots,a^n)v(s',\pi_*^1,\cdots,\pi_*^n)\} \tag{2}$$

Q 学习算法的基本思想是对 Q 函数的定义，而常规 Q 函数如下：

$$Q_*(s,a^1,\cdots,a^n)=r(s,a^1,\cdots,a^n)+\beta\sum_{s'\in S}p(s'\mid s,a^1,\cdots,a^n)v(s',\pi_*^1,\cdots,\pi_*^n) \tag{3}$$

通过这个定义，则 $Q_*(s,a^1,\cdots,a^n)$是在状态 s 下执行动作序列$(a^1,\cdots,a^n)$，然后按照最优策略得到的总折扣奖赏。由式(2)和式(3)可以得到：

$$v^i(s,\pi_*^1,\cdots\pi_*^n)=\max_{a^1,\cdots,a^n}Q_*^i(s,a^1,\cdots,a^n) \tag{4}$$

如果能够计算出 $Q_*^i(s,a^1,\cdots,a^n)$，则最优策略$(\pi_*^1,\cdots,\pi_*^n)$可以认为是在状态 s 下最大化$Q_*^i(s,a^1,\cdots,a^n)$的动作序列。问题可以缩小为计算 $Q_*^i(s,a^1,\cdots,a^n)$的函数值来取代对 $v^i(s,\pi_*^1,\cdots\pi_*^n)$的最优值的计算。

Q 学习提供了一个简单的更新函数，其中 Agent 的初始 Q 值可以为任意值，而对于随机博弈的 Q 值更新函数如下：

$$Q_{t+1}^i(s,a^1,\cdots,a^n)=(1-\alpha_t)Q_t^i(s,a^1,\cdots,a^n)+\alpha_t[r_t^i+\beta\text{Optimal}Q_t^i(s')] \tag{5}$$

式中，学习率 $\alpha_t\in(0,1)$，这里的 Optimal 表示在博弈的某种平衡情况下求得的最优值。

3 多 Agent MetaQ 算法

3.1 Meta 平衡

Meta 博弈是由 Howard 首先提出的[11]，它允许每一个 Agent 的策略是在其它 Agent 策略的假定基础上做出的反应策略。

Meta 博弈：Meta 博弈是一种假定博弈，它是由其中一个 Agent 假定其它 Agent 的行动策略而得到自己的策略的一种博弈。

Meta 博弈可以作为一种扩展的战略形式，当其它 Agent 扩展 Agent_i 的策略，则得到博弈 G 的 Meta 博弈 iG。显然，Meta 博弈是可递归的。在具有 n 个 Agent 的多 Agent 系统中，Meta 博弈 12…nG，$n(n-1)$…21G，或者任一以 1，2，…，n 其中之一排列为前缀的 Meta 博弈都是由博弈 G 衍化而来的。

Meta 平衡：如果每个 Agent 满足

$$\min_{\pi_{P_i}}\max_{\pi_i}\min_{\pi_{F_i}}(R_i(\pi_{P_i},\pi_i,\pi_{F_i})\leqslant R_i(\pi_{P_i,m},\pi_{i,m},\pi_{F_i,m}) \tag{6}$$

则联合策略 $\pi^m=(\pi^{1,m},\cdots,\pi^{n,m})$是博弈 12…$n$G 的一个 Meta 平衡。其中 P_i 是 i 之前的策略集合序列，F_i 是 i 之后的策略集合序列。

对称 Meta 平衡：博弈 G 的一个 Meta 平衡$(\pi^1,\cdots,\pi^n)$，如果它既能满足式(6)，同时又能满足

$$\min_{\pi_{F_i}}\max_{\pi_i}\min_{\pi_{P_i}}(R_i(\pi_{F_i},\pi_i,\pi_{P_i})\leqslant R_i(\pi_{F_i,m},\pi_{i,m},\pi_{P_i,m}) \tag{7}$$

那么就称 Meta 平衡$(\pi^1,\cdots,\pi^n)$为对称 Meta 平衡。

3.2 Meta 平衡集的计算

文献[12]给出了一般情形下各平衡集的表达式，将其结论重写为如下：

设 $G=(a_{ij},b_{ij})_{m\times n}$为一个 $m\times n$ 博弈，$E(G)$表示 Meta 平衡集，$SE(G)$表示对称 Meta 平衡集，则其各平衡集可表达为

$$E(G)=\{(s,t)\mid a_{st}=\max_k a_{kt}\wedge b_{st}=\max_l b_{sl}\}$$

$$E(iG)=\{(s,t)\mid a_{st}=\max_k a_{kt}\wedge b_{st}\geqslant\max_j\min_i b_{ij}\}$$

$$E(jG)=\{(s,t)\mid a_{st}\geqslant\max_i\min_j a_{ij}\wedge b_{st}=\max_k b_{sk}\}$$

$$E(jiG)=\{(s,t)\mid a_{st}\geqslant\min_j\max_i a_{ij}\wedge b_{st}\geqslant\max_j\min_i b_{ij}\}$$

$$E(ijG)=\{(s,t)\mid a_{st}\geqslant\max_i\min_j a_{ij}\wedge b_{st}\geqslant\min_i\max_j b_{ij}\}$$

$$SE(G)=\{(s,t)\mid a_{st}\geqslant\min_j\max_i a_{ij}\wedge b_{st}\geqslant\min_i\max_j b_{ij}\}$$

3.3 MetaQ 算法

MetaQ 算法保留了 NashQ 的基本框架，不同之处在于动作选择和 Q 值更新这两个关键步骤：NashQ 算法根据 Nash 平衡策略选择动作、更新 Q 值；而 MetaQ 算法根据 Meta 平衡策略选择下一个动作，并更新 Q 值。其中 MetaQ 算法描述如下：

Step 1　初始化：令 $t=0$；选择初始状态 s_0；

Step 2　对所有的 $s\in S,a^1\in A^1,\cdots,a^n\in A^n$，设定 Q 的初始值 $Q_t^i(s,a^1,\cdots,a^n)=0$；

Step 3　从状态 s_t 的 Q 矩阵求解出 *Meta* 平衡，并根据这一平衡采取相应的动作 a^i；

Step 4　观察 $a^1,\cdots,a^n$ 和 $r^1,\cdots,r^n$ 以及新状态 s_{t+1}；

Step 5　按如下迭代公式更新 Q 矩阵：

$$Q_{t+1}^i(s_t,a^1,\cdots,a^n)=(1-\alpha_t)Q_t^i(s_t,a^1,\cdots,a^n)+\alpha_t[r_t^i+\beta MetaQ_t^i(s_{t+1})] \tag{8}$$

其中

$$MetaQ_t^i(s_{t+1}) = \pi^1(s_{t+1})\cdots\pi^n(s_{t+1})\cdot Q_t^i(s_{t+1}) \quad (9)$$

Step 6　调整学习率 α_t。令 $t=t+1$，转 Step 3 直至结束。

3.4　MetaQ 算法性能分析

根据以上 MetaQ 算法的描述，每个 Agent 在博弈中需要更新 n 个 Q 函数。假设在算法中任一 Agent 能预测其它 Agent 的行动和观察到它们的即时奖赏情况，则它依次不断更新 Q 函数。由于 MetaQ 算法在空间上主要由 Agent 更新 Q 函数并使用，因此 MetaQ 算法的空间复杂度与 NashQ 算法的相同。设 $|S|$ 为博弈的状态空间大小，$|A^i|$ 为 Agent_i 的动作空间 A^i 的大小，假设 $|A^1|=\cdots=|A^n|=|A|$，则算法的空间复杂度为 $n|S|\cdot|A|$。即算法空间复杂度相对状态的数目是线性的，但随着 Agent 的数量增加其呈指数增长[7]。

相比较于 NashQ 算法的运行时间主要是在 Q 函数更新中使用的 Nash 平衡的计算，MetaQ 算法的运行时间同样在 Q 函数更新中使用的 Meta 平衡的计算。但 NashQ 算法中找到一个 Nash 平衡的复杂度是未知的，作为一个 n-Agent 博弈中具有代表性的 2-Agent 的博弈常用算法，在最坏情况下有着指数级的计算复杂度[7]。而对于 MetaQ 算法，其 Meta 平衡的计算由 Meta 平衡集的计算过程可知，在一般和博弈的情形求解 Meta 平衡可归结为求解矩阵的 min-max 或 max-min 值，其复杂度仅为 n^2。与求解 LCP 相比，它要简单得多，甚至可以很方便地求解出全部的 Meta 平衡点。

文献[13,14]中给出了 Meta 平衡的 3 个重要性质：

性质 1　任意一般和博弈 G，必有纯策略 Nash 平衡，从而 G 的 Meta 平衡必存在。

性质 2　任意一般和博弈 G，它的纯策略 Nash 平衡必是对称的 Meta 平衡。

性质 3　任意一般和博弈 G，若 (c,d) 是 G 的 Meta 平衡，(a,b) Pareto 优于 (c,d)，则 (a,b) 也是 G 的 Meta 平衡。

性质 1 保证了 Meta 平衡的存在性。在此基础上，性质 2 表明当只考虑纯策略时，Meta 平衡比 Nash 平衡更为宽泛。性质 3 则保证了 Meta 平衡不会像 Nash 平衡那样陷入局部困境，将能够获得 Pareto 最优解。其中 (a,b) Pareto 优于 (c,d) 是指 $a\geqslant c$ 并且 $b\geqslant d$。

由以上 3 个性质可知，在一般和随机博弈中，MetaQ 算法能够满足以下 3 个要求：①Meta 平衡总是存在的；②Meta 平衡应当是一个联合纯策略；③博弈至少存在一个 Pareto 最优解。同时又由于 MetaQ 算法的计算复杂度相对 NashQ 算法要低，因此在一般和随机博弈中由 MetaQ 算法取代 NashQ 算法能够取得更好的效果。

3　气候合作策略游戏实验及分析

采用文献[15]中描述的气候合作策略游戏来测试 MetaQ 算法各方面的性能，并与 NashQ 算法比较。气候合作策略游戏初始设定参与玩家为 6，即 $N=6$ 的多 Agent 系统。开始的时候每个 Agent 被提供€ 40，并被要求匿名地向“气候账户”进行 10 次的投资，每次投资的金额仅可以为€ 0，€ 2 或 € 4。同时每个 Agent 对目标是确定感知的，即如果在 10 次投资结束之后，投资总额能够达到或者超过€ 120，那么所有的 Agent 都将会得到他们账户中剩余的所有的钱；反之游戏没有达到目标，作为惩罚，他们将以一定惩罚概率（本文设置惩罚概率为 0.9）失去他们所有的钱。由于 Agent 的个体自私性，所有 Agent 都试图使得自身在游戏结束时能获得更多的钱，如此必然要求自己在每次投资的时候尽可能少地投入，由此引发了个人利益与集体利益的博弈。

对 Agent 的训练以“轮”为单位进行。在任意一轮训练中，只要进行了 10 次投资，无论是否达到目标训练都结束，同时所有 Agent 置初始状态，然后再开始新一轮的训练。

显然在这个气候合作策略游戏中，对每个 Agent 利益及集体利益综合考虑，最好的投资方式是每次每个 Agent 仅投资€ 2。取学习率的初始值 $\alpha=0.1$ 进行实验，使用 MetaQ 及 NashQ 2 种算法，分别对游戏进行仿真并比较分析，得到两种算法的每轮投资目标和的收敛图，如图 1 所示。

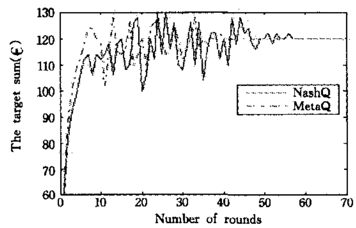

图 1　MetaQ&NashQ 目标收敛图

由图 1 可知 MetaQ 算法在第 8 轮已经达到投资 120 的目标，而 NashQ 算法在第 19 轮才达到目标。同时 MetaQ 算法在 41 轮就已经收敛，而 NashQ 算法在 56 轮才开始收敛。

图 2 记录了在上述实验中的一个 Agent 在两种不同算法中的每轮投资金额情况。可以看到，两种算法的学习过程同样都是收敛的，但明显 MetaQ 算法要早于 NashQ 算法达到基本投资标准 20，并且 MetaQ 算法相对于 NashQ 算法收敛于 20 要早 17 轮。

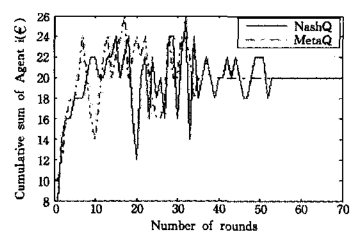

图 2　MetaQ&NashQ 对 Agent_i 投资收敛图

结束语　本文针对 NashQ 算法存在的问题及不足，在 NashQ 算法的框架基础上提出了 MetaQ 算法。与 NashQ 算法不同，MetaQ 学习算法通过对自身行为的预处理以及其他对象的行为预测来获取共同行为的最优策略。最后通过研究及气候合作策略游戏实验证明，MetaQ 算法有着很好的理论解释和实验性能，其收敛速度优于 NashQ 算法。

虽然在理论分析及实验中显示 MetaQ 算法优于 NashQ 算法，但对于 MetaQ 算法的收敛性还没有给出理论证明，这将是我们进一步的主要研究任务。

参 考 文 献

[1]　Hu Jun-ling, Wellman M P. Multiagent reinforcement learning: theoretical framework and an algorithm[C]//Proceedings of the Fifteenth International Conference on Machine Learning. 1998: 242-250

[2]　Wang Hao, Gao Yang, Chen Xing-guo. RL-DOT: A Reinforcement Learning NPC Team for Playing Domination Games[J].

IEEE Transactions on Computational Intelligence and AI in Games,2010,2(1):17-26
[3] Littman M L. Friend-or-foe q-learning in general-sum games[C]//Proceedings of the Eighteenth International Conference on Machine Learning. Williams College:Morgan Kaufman,2001:322-328
[4] Greenwald A,Hall K,Serrano R. Correlated-q learning[C]//Proceedings of the Twentieth International Conference on. Washington DC,2003:242-249
[5] 赵凤强,徐毅,李广强. 基于岛屿群体模型的多目标演化算法研究[J]. 计算机科学,2010,37(12):190-192
[6] 宋梅萍,顾国昌,张国印,等. 一般和博弈中的合作多 agent 学习[J]. 控制理论与应用,2007,24(2):317-321
[7] Hu Jun-ling,Wellman M P. Nash Q-learning for general-sum stochastic games[J]. Journal of Machine Learning Research,2003,4(11):1039-1069
[8] Vassiliades V,Cleanthous A,Christodoulou C. Mutltiagent Reinforcement Learning:Spiking and Nonspiking Agents in the iterated Prisoner's Dilemma[J]. IEEE Transactions on Neural Networks,2011,22(4):639-653
[9] Aumann R J, Hart S. Computing equilibria for two-person games. Handbook of Game Theory with Economic Applications [R]. Amsterdam:Elsevier,2002
[10] Murty K G. Computational complexity of complementary pivot methods[C]//Mathematical Programming Study 7 Complementarily and Fixed Point Problems. Amsterdam: North-Holland Publishing Co,1978:61-73
[11] Howard N. Paradoxes of Rationality[M]. Theory of Metagames and Political Behavior, MIT Press, Massachusetts: Cambridge, 1971
[12] Thomas L C. Games, Theory and Application [M]. Halsted Press,Chichester,1984:129-149
[13] Cao Yong, Li Ren-hui. Conflict Analysis between Tacit Knowledge Sharing and its Exclusivity Based on Meta-game Theory[C]//2009 International Symposium on Information Engineering and Electronic Commerce. IEEC,2009:31-35
[14] Sharma R,Gopal M. Hybrid Game Strategy in Fuzzy Markov-Game-Based Control[J]. IEEE Transactions on Fuzzy Systems, 2008,1(16):1315-1327
[15] Milinski M,Sommerfeld R D,Krambeck H J,et al. The Collective-risk Social Dilemma and the Prevention of Simulated Dangerous Climate Change[J]. Proceedings of the National Academy of Sciences,2008,105(7):2291-2294

(上接第 231 页)

3 应用实例

基于本文的思路和涉及到的关键技术,在一系列 CyberSIGStudio 服务器[6-8]的支撑下,设计并实现了面向网络的空间信息提取原型系统。下面以"地震事件"为例来阐述本系统的功能。

用户对话框中输入"地震事件",系统将所提取的地震实例在页面的左侧显示,并在页面右侧的地图中标出事发的地点(见图 4)。在图中,左侧列表显示的结果是系统从网络中获取到的最近发生的 3 项地震实例。

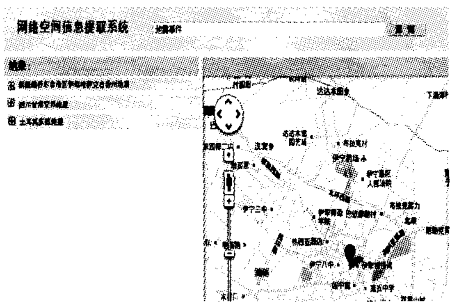

图 4 空间信息提取结果示意图

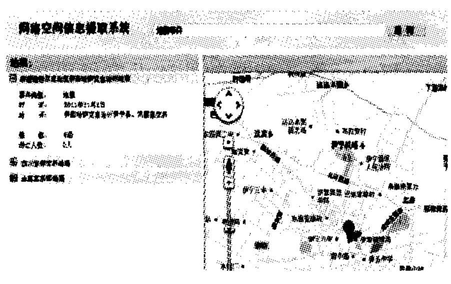

图 5 空间信息提取结果示意图

当用户点击页面左侧的具体事例时,系统将显示该事件的具体侧面,包括时间、伤亡人数、震级等信息。图 5 展示是当用户点击"新疆维吾尔族自治区伊犁哈萨克自治州地震"时,该事件的详细信息被展示出来。

结束语 本系统充分挖掘了网络中存在的大量空间信息,增加了空间数据获取的有效途径,能够有效地降低空间数据获取的成本,为空间数据服务系统提供技术与数据的支持。针对空间信息在网络中存在的特点,本系统分别提出了空间位置信息识别算法与属性信息提取算法,利用统计模型提高了地址识别的准确度,并进一步利用地址地理编码技术将其空间化。

在地址地名库的构建当中,如何有效地对复杂多样的地址地名进行建模与索引是需要进一步研究的课题。由于地址地名库中的标准地址都是短语文本,因此研究更加高效的匹配方法十分必要。此外,研究如何适应不同比例尺的地址地理编码技术也很有意义。

参考文献

[1] McCurley S. Geographic Mapping and Navigation of the Web [C]//The 10th www Conference. Hong Kong,2001
[2] http://p2pfoundation.net/Michael_Goodchild_on_Volunteer_Mapping's_Role_in_Geospatial_Science
[3] 朱建伟,王泽民. 地理编码原理及其本地化解决方案[J]. 北京测绘,2002,12(2):24-27
[4] 马皓明. 中文地址地理编码研究与原型系统实现[D]. 北京:北京大学,2007
[5] Ricardo B. Modern Information Retrieval[M]. New York:Addison Wesley,1999:110-112
[6] 林绍福,李琦,董宝青. 数字城市应用服务平台体系结构研究[J]. 计算机科学,2002,29(12):98-102
[7] 李琦,甘杰夫. 数字城市空间信息与服务集成交换平台系统分析与设计[J]. 计算机科学,2005,32(9):123-126
[8] 史文勇,李琦,林宇. 数字城市核心系统平台的服务总线设计[J]. 计算机科学,2006,33(3):279-282