基于追尾行为的改进型人工萤火虫群算法

李咏梅[1] 周永权[2] 姚祥光[1]

(广西大学计算机与电子信息学院 南宁 530004)[1] (广西民族大学数学与计算机科学学院 南宁 530006)[2]

摘 要 利用人工鱼群算法的追尾思想并在过程中加入拥挤度因子,对人工萤火虫群算法进行了改进,提出了一种改进型人工萤火虫群算法,并将该算法用于多峰函数的优化问题。通过实验仿真及与其他算法进行的对比分析表明,改进后的人工萤火虫群算法在种群规模较小、迭代次数较少的情况下也可以精确捕获函数定义域内的所有峰值。

关键词 人工萤火虫群算法,追尾行为,拥挤度因子,多峰函数优化

中图法分类号 TP18 **文献标识码** A

Improved Glowworm Swarm Optimization Based on the Behavior of Follow

LI Yong-mei[1] ZHOU Yong-quan[2] YAO Xiang-guang[1]

(School of Computer and Electron Information, Guangxi University, Nanning 530004, China)[1]

(College of Mathematics and Computer Science, Guangxi University for Nationalities, Nanning 530006, China)[2]

Abstract To improve the glowworm swarm optimization, the behavior of follow of the artificial fish school algorithm and a swarm degree were used. The improved algorithm was used to optimise multi-modal functions. The experiment results show that the algorithm, with the smaller populations and the fewer number of iterations, can simultaneous capture multiple optima of several standard multimodal test function.

Keywords Glowworm swarm optimization, Behavior of follow, Swarm degree, Multimodal function optimization

1 引言

在科学计算与工程技术领域,许多求解问题在本质上都是一些多峰值函数优化问题,需要在可行域内找出多峰问题的多个最优解,从而为决策者提供多种选择或多方面的信息[1,2]。在处理复杂多峰优化问题时,传统的单点搜索方法往往一次只能非确定地搜索到一个极值点,因此对这类问题求解基本无效。群智能算法研究的不断深入,为求解多峰值函数优化问题提供了一些有效的新途径。特别是人工萤火虫群智能算法(glowworm swarm optimization, GSO)的出现,为解决多峰问题的寻优提供了一种新的具有竞争力的求解算法。但是,通过分析发现,基本GSO算法存在着由于人工萤火虫群漫无目标的随机移动而导致的收敛速度不够快以及解的精度不够高的问题。

本文针对基本GSO算法存在的不足,引入人工鱼群算法(artificial fish-swarm algorithm, AFSA)的追尾行为概念,并在过程中加入了拥挤度因子,对GSO算法进行了改进,提出了一种基于追尾行为的改进型人工萤火虫算法。通过对不同的多峰函数进行的仿真实验结果表明,本文所提出的算法可在种群规模较小、迭代次数较少的情况下搜索到多峰函数的所有局部最优解,且每个最优解精度都达到了理想值。

2 人工萤火虫群算法

人工萤火虫群智能算法是由Krishnanand和Ghose在2005年提出来的一种群体智能算法[3-6],该算法思想来源于自然界的萤火虫群中的萤火虫用发光来吸引伴侣或者猎物,并且越亮的萤火虫吸引力越大,最后形成大多数的萤火虫聚集在多个位置的现象。其中萤火虫发光的强度是依靠其携带的发光元素的质量来决定的,我们将每只萤火虫携带的能够发光的元素称为荧光素(luciferin)。

在GSO中,我们把算法中的个体(agent)看成是人工萤火虫(以下简称萤火虫),它们的亮度与自己所在位置上的目标值有关,愈亮的萤火虫表示它所在的位置愈好,即有较佳的目标值。算法中的萤火虫依靠动态决策域(dynamic decision domain)来确认它的邻居并且计算它的移动。每只萤火虫都选择比自己亮即荧光素值比自己高的萤火虫做自己的邻居,然后利用概率机制向其移动。也就是说,它们总是被比自己亮的邻居所吸引。这些仅仅由本地信息决定的移动使得人工萤火虫群被分成不相交的子群,最终大多数的萤火虫会聚集于多个位置上。

GSO算法主要有4个阶段:萤火虫部署阶段、亮度更新阶段、移动阶段以及动态决策域更新阶段。

1)部署阶段:使萤火虫均匀散布在解空间中,同时每只萤

到稿日期:2010-04-06 返修日期:2010-07-19 本文受国家自然科学基金项目(60461001),广西自然科学基金项目(0832082,0991086),国家民委科研项目基金(08GX01)资助。

李咏梅(1986—),女,硕士生,主要研究方向为计算智能及应用,E-mail:liyongmei180@163.com;**周永权**(1962—),男,博士,教授,主要研究方向为计算智能理论与方法、神经网络及应用;**姚祥光**(1985—),男,硕士生,主要研究方向为计算智能及应用。

火虫携带相同的初始亮度和感应半径。

2)亮度更新阶段:算法中萤火虫的荧光亮度与解品质有关,目标函数值愈好,其亮度在同一代中也愈亮。在每一代萤火虫都移动完毕之后(或初始部署时),下一轮开始,所有萤火虫的亮度都会更新。

3)移动阶段:在移动阶段,每只萤火虫会选择另外的萤火虫作为其邻居,然后利用概率机制朝其方向移动。萤火虫 A 要成为另一只萤火虫 B 的邻居,要满足两个条件:一是萤火虫 A 要在萤火虫 B 的决策区域内,二是萤火虫 A 要比萤火虫 B 亮。

4)决策区域半径更新阶段:GSO 算法中各个萤火虫采取自适应的区域决策范围。也就是说,在移动位置之后,萤火虫会根据邻居密度来动态更新其区域决策半径。区域决策半径大小由目前邻居密度(即决策区域范围内所涵盖的邻居数量比率)的高低决定。若邻居密度小,半径会加大,以搜索邻居,反之区域决策半径会缩小。若密度没有变化,半径也保持不变。

3 人工萤火虫群算法改进机理

人工鱼群算法 ASFA 是 2002 年我国学者李晓磊等人提出的一种新型自适应寻优算法[7]。人工鱼群算法是一种基于模拟鱼群行为的随机搜索优化算法,主要利用了鱼的觅食、聚群和追尾行为,从构造单条鱼的底层行为做起,通过鱼群中各个体的局部寻优来达到全局最优值在群体中突现出来的目的。在鱼群算法中,人工鱼具有 3 种行为:觅食、聚群和追尾。其中追尾行为是指人工鱼向最优伙伴方向前进一步。人工鱼搜索其视野内的所有伙伴中函数值最优的伙伴,如果该伙伴状态较优且其周围不太拥挤,则朝它的方向前进一步。否则,人工鱼在搜索范围内尽量随机搜索一个比当前状态更优的状态向其前进一步。这种尾随邻近的最优点移动的行为使人工鱼能快速地找到食物源。

萤火虫群算法和鱼群算法都属于群优化算法。它们的共同特点是,对于单个个体而言(萤火虫或人工鱼)不存在智能行为,只是遵循某种规律而运动。但当个体数量达到一定程度后,整个种群将会表现出某种智能行为。基本 GSO 单纯地以机率挑选邻居,较亮的邻居有较高的机率被挑中。本文算法引入鱼群算法中的追尾概念,使萤火虫直接朝区域决策范围内最亮的邻居飞行,提高了收敛速度。但是,单纯地跟着最亮的萤火虫飞很容易出现碰撞现象。为了避免这种现象发生,引入了鱼群算法的拥挤度因子。鱼群算法中人工鱼游动方向除了决定于状态的最优值以外,还决定于该位置的拥挤度。如果该位置过于拥挤,即使该区域具有较优的函数值,人工鱼也可能不向该区域游动。本文算法中萤火虫在自己的区域决策范围内搜索到目前荧光素最高的邻居后,计算该邻居当时的拥挤度。如果拥挤度小于此时刻的拥挤度阈值,则表示该邻居附近不太拥挤,萤火虫可以直接向该邻居飞行。否则,表示该邻居周围过于拥挤,萤火虫则在自己的区域决策范围内尽量随机选择一个荧光素较高的萤火虫作为其邻居并向其飞行。

鱼群算法中,拥挤度在算法的寻优过程中始终起作用。引入拥挤度因子,在算法的初期可以避免算法个体过早地集结到最优值附近,从而可避免发生早熟现象,提高算法的全局寻优能力。但在算法后期,拥挤度将会对算法的收敛性以及收敛速度造成影响。也就是说,拥挤度在寻优初期可改善算法的寻优性能,在寻优后期则对寻优性能产生一定的负面影响,因此拥挤度阈值的选择十分重要。在本文算法中,拥挤度阈值是随着迭代次数而动态更新的。在算法初期,拥挤度阈值选择接近于 0 的较小值,这样大多数萤火虫都可以自主随机地选择移动方向,可避免萤火虫过早地集中在某位置而造成算法早熟现象的出现。随着迭代次数的增加,拥挤度阈值逐渐增大,也就是荧光素的浓度在指导萤火虫选择邻居的作用逐渐增强,最后可演变为邻居选择与拥挤度无关,完全由荧光素的浓度来决定,从而保证算法能够快速地收敛。

人工萤火虫群算法的移动步长对算法的收敛速度有较大影响:步长大,萤火虫移动的速度就快,算法初期收敛速度也就快。但随着萤火虫个体逐渐靠近最优点,步长太大,萤火虫可能会跨过最优点,使得算法不稳定,收敛速度变慢,求解精度不高;步长小,虽然求解精度可以得到提高,但算法收敛缓慢。针对这一问题,为了提高算法的收敛速度和求解精度,本文在算法的初期赋予每只萤火虫较大的步长初值,随着算法的进行,步长自适应地逐渐减小,从而加快了算法的收敛速度。

4 改进型人工萤火虫群算法

基于人工萤火虫群算法改进机理型,我们引入人工鱼群算法的追尾概念和拥挤度概念,将二者有机地结合起来。改进型人工萤火虫群算法完整算法流程如图 1 所示。

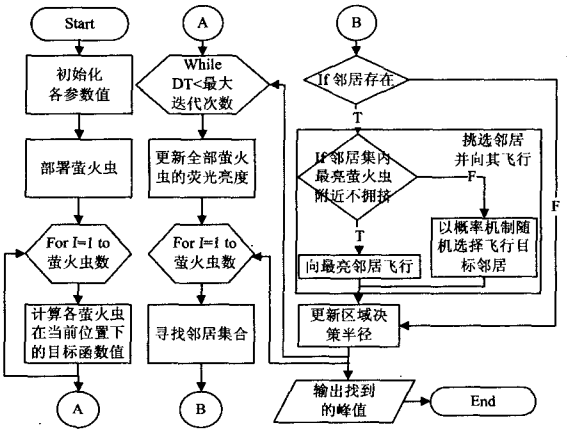

图 1 改进型人工萤火虫群算法流程图

算法具体实施步骤如下。

步骤 1(部署萤火虫) 将 n 只萤火虫随机地散布在搜索空间中,同时为每只萤火虫赋以相同的初始亮度 l_0 和感应半径 r_0,形成初始萤火虫群。同时设定萤火虫的移动步长初始值 s_0,初始迭代次数 DT 为 0,最大迭代次数为 maxDT。

步骤 2(更新全部萤火虫的荧光亮度) 各萤火虫按照文献[8]中的荧光素更新公式更新荧光素。

步骤 3(搜寻邻居集合) 萤火虫在各自的决策范围内选择比自己亮的萤火虫作为自己的邻居,同时计算邻居集内各邻居的转移概率。如果邻居存在,执行步骤 4,否则转到步骤 6 执行。

步骤 4(挑选移动目标) 从邻居集内挑选出某个邻居作为移动目标。挑选时先计算邻居集中荧光素最高即最亮萤火虫当时的拥挤度 $q_i(t)$。本文在萤火虫的邻居密度公式[8]的

基础上进行修改，从而形成拥挤度的计算公式，以计算当时某只萤火虫周围是否已经聚集过多萤火虫而造成拥挤。拥挤度计算公式如下：

$$q_i(t)=\frac{|N_i(t)|}{\pi(r_d^i(t))^2}$$

式中，$|N_i(t)|$表示萤火虫 i 在 t 时刻的邻居个数，$r_d^i(t)$表示萤火虫 i 在 t 时刻的区域决策半径。

如果 $q_i(t)<\delta(t)$，则表示该邻居附近不太拥挤，萤火虫可以直接向该邻居飞行。否则，萤火虫在自己的区域决策范围内以概率机制重新随机选择一个萤火虫作为其邻居并向其飞行。其中，$\delta(t)$表示 t 时刻的拥挤度阈值，并按下式进行更新：

$$\delta(t)=1-e^{-c}$$

式中，c 为阈值变化系数。

步骤 5　向挑选好的目标邻居按照移动公式[8]移动。其中移动步长 S 随着迭代次数按下式更新：

$$s=0.96^t s_0$$

步骤 6(更新区域决策半径)　在移动位置之后，萤火虫会根据邻居密度来动态更新其区域决策半径。

步骤 7(终止条件判断)　判断 DT 是否已经达到 maxDT，若不满足，则 DT 加 1 并转到步骤 2 执行，进行下一代萤火虫群优化过程，否则转到步骤 8 执行。

步骤 8(算法终止)　输出找到的峰值。

5　仿真实验与性能分析

为验证本文算法的有效性，选取了几个典型多峰函数(求解极大值)进行仿真测试。实验中大部分参数的设置参考文献[8]中各参数的取值。另外，通过大量实验得出阈值变化系数 c 取 0.1768 时效果最好。本文算法用 Matlab 7.0 编程实现，实验均在 Pentium 4，2.99GHz，512 MB 内存的微机上进行。实验将小生境微粒群算法(Niche-PSO)[9]、基本 GSO 以及本文算法在以下 3 个函数中的捕峰能力进行了比较，在其他参数一致的情况下，改变群体内萤火虫的个体数，记录不同萤火虫个体数下不同算法能捕获的函数峰个数。此外，为了验证本文算法在捕峰速度上的优势，还将基本 GSO 与本文算法在其他参数一致的情况下捕获一定数量峰所需的迭代次数及运行时间进行了比较。为了保证实验的有效性，每次实验都重复进行了 30 次以上。

1)Himmelblau's 函数[9]

$$J_1(x,y)=200-(x^2+y-11)^2-(x+y^2-7)^2$$

该函数为典型线性不可分的等高、非等距的二维多峰函数，在$[-5,5]\times[-5,5]$的搜索空间内有 4 个非均匀分布且比较平坦的相等峰值，理论值均为 200.0，性能不佳的算法很难快速精确地搜索到全部 4 个峰。从图 2 中可以看出，基本 GSO 算法和本文算法随着萤火虫个体数的增加，能捕获的峰的个数也在增加，并且最终都能捕获全部 4 个峰。但是，显然本文算法捕峰个数增长比较快，并且在萤火虫个体数 n 为 60 时就能基本捕获全部 4 个峰。

图 3 展示了基本 GSO 算法与本文算法在不同迭代次数下捕获峰的平均值。从图 3 中可以看出，本文算法捕获全部 4 个峰所需的迭代次数大大少于基本 GSO 算法。实验表明，本文算法捕获全部 4 个峰所需运行时间平均为 1.828s，而基本 GSO 算法需要 5.215s。

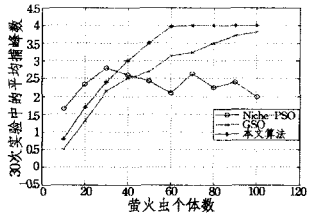

图 2　不同算法在不同萤火虫个体数下对函数 J_1 的平均捕峰数

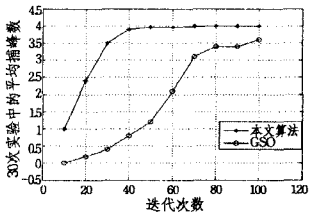

图 3　不同算法在不同迭代次数下对函数 J_1 的平均捕峰数

2)Equal-peaks-B 函数[10]

$$J_2(x,y)=\cos^2(x)+\sin^2(y)$$

该函数在$[-5,5]\times[-5,5]$的搜索空间内有 12 个函数值相等的峰。3 种算法在不同萤火虫个体数下的平均捕峰数如图 4 所示。从图中可以看出，本文算法在萤火虫个体数为 350 时就能基本捕获搜索空间内所有的 12 个峰，捕峰能力比另外两个算法要强。通过图 5 可以看出，本文算法在 40 代时即可捕获函数在搜索空间内所有的 12 个峰，而基本 GSO 算法则至少需要迭代 100 次。

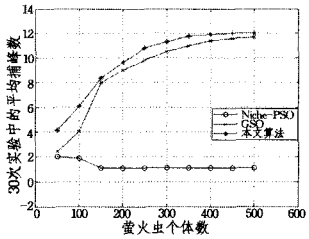

图 4　不同算法在不同萤火虫个体数下对函数 J_2 的平均捕峰数

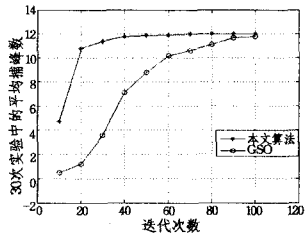

图 5　不同算法在不同迭代次数下对函数 J_2 的平均捕峰数

3)Rastrigin's 函数

$$J_3(x,y)=20+(x^2-10\cos(2\pi x)+y^2-10\cos(2\pi y))$$

该函数由于其广泛的搜索空间及大量的局部极大值而成为一个相当棘手的问题，它在定义域内有 100 个峰。因此 Rastrigin's 函数经常用来作为测试函数优化算法性能的基准函数[11,12]。本文实验中取的搜索范围为$[-2,2]\times[-2,2]$，该函数在该范围中有 16 个函数值不等的峰。3 种算法在不同萤火虫个体数下的平均捕峰数如图 6 所示。从图中可以看出，本文算法在萤火虫个体数为 400 时就能基本捕获搜索空间内所有的 16 个峰，捕峰能力比另外两种算法要强。同时，从图 7 可以看出，本文算法捕获所有峰所需的迭代次数要比基本 GSO 算法少，在 40 代时就能基本捕获所有的 16 个峰。

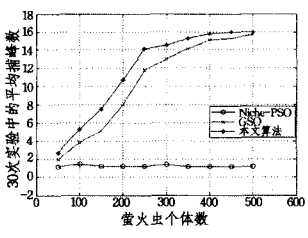

图 6　不同算法在不同萤火虫个体数下对函数 J_3 的平均捕峰数

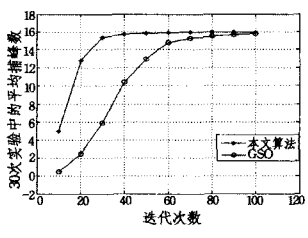

图 7　不同算法在不同迭代次数下对函数 J_3 的平均捕峰数

结束语　在人工萤火虫群智能算法中引入追尾概念，并在过程中加入了拥挤度因子，大大地改善了 GSO 的性能。将

改进后的GSO算法用在多峰函数问题中，仿真实验结果表明，改进后的GSO算法具有在种群规模较小、迭代次数较少的情况下也可以精确捕获函数定义域内所有峰的特点，同时消除了人工萤火虫群漫无目标的随机移动的局限。本算法不仅保持了基本GSO算法的特点，而且提高了算法的运行效率。由于GSO理论和应用研究还处于不断完善阶段，因此许多问题有待于人们不断地探索和解决。比如，GSO对于其参数的选择有较高的敏感度，参数选择的好坏直接影响结果的精度。因此，如何选择恰当的参数和对GSO算法进行收敛性分析，以及与其它智能算法有机地融合，设计出一些高效的群混合智能算法并用之求解实际问题，将是我们今后所要做的研究工作。

参 考 文 献

[1] Wei L Y, Zhao M. A niche hybrid genetic algorithm for global optimization of continuous multimodal functions[J]. Applied Mathematics and Computation, 2005, 160(3): 649-661

[2] 李敏强，寇纪淞. 多模态函数优化的协同多群体遗传算法[J]. 自动化学报，2002，28(4)：497-504

[3] Krishnanand K N, Amruth P, Guruprasad M H, et al. Glowworm-inspired robot swarm for simultaneous taxis towards multiple radiation sources[C]// IEEE International Conference on Robotics and Automation. Orlando, Florida, May 2006: 958-963

[4] Krishnanand K N, Ghose D. Detection of multiple source locations using a glowworm metaphor with applications to collective robotics[C]// Swarm Intelligence Symposium. June 2005: 84-91

[5] Krishnanand K N, Ghose D. Theoretical foundations for multiple rendezvous of glowworm-inspired mobile agents with variable local-decision domains[C]// American Control Conference. June 2006: 14-16

[6] Krishnanand K N, Ghose D. Theoretical foundations for rendezvous of glowworm-inspired agent swarms at multiple locations [J]. Robotics and Autonomous Systems, 2008, 56(7): 549-569

[7] 李晓磊，邵之江，钱积新. 一种基于动物自治体的寻优模式：鱼群算法[J]. 系统工程理论与实践，2002，22(11)：32-38

[8] Krishnanand K N, Ghose D. Glowworm swarm optimisation: a new method for optimising multi-modal functions[J]. Int. J. Computational Intelligence Studies, 2009, 1(1): 93-119

[9] Brits R, Engelbrecht A P, van den Bergh F. A niching particle swarm optimizer[C]// The 4th Asia-Pacific Conference on Simulated Evolution and Learning. 2002: 692-696

[10] Parsopoulos K, Vrahatis M N. On the computation of all global minimizers through particle swarm optimization[J]. IEEE Transactions on Evolutionary Computation, 2004, 8(3): 211-224

[11] Fevrier V, Patricia M. Parallel evolutionary computing using a cluster for mathematical function optimization[C]// The Annual Meeting of the North American Fuzzy Information Processing Society. 2007: 598-603

[12] Muller S D, Marchetto J, Koumoutsakos S A P. Optimization based on bacterial chemotaxis[J]. IEEE Transactions on Evolutionary Computation, 2002, 6(6): 16-29

(上接第217页)

最终语义关系图如图7所示，带箭头的虚线表示预测的语义关系，百分数表示该语义关系存在的可能性。

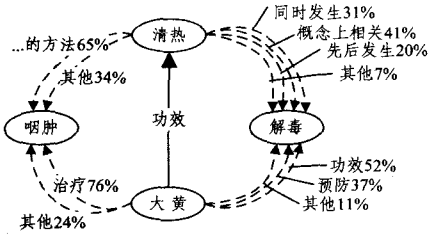

图7 语义关系图

结束语 本文提出了一种利用中医药本体知识库对中医药文献进行语义关系图发现的方法。本文的关键词提取采用中医药本体的概念名称作为字典，关键词提取无需依赖语法分析程序。分词后采用类似关联规则的算法进行高频关键词计算。然后利用中医药本体知识库对高频关键词组进行语义关系图的识别和预测。最后本文通过实验对提出的算法进行了验证，结果证实了算法的可行性。

利用语义关系图对中医药文献进行知识发现，具有表达直观性，知识获取全面性及可推理性等优点，能够有效帮助相关研究人员对大量文献进行知识获取、分类、推理等。

本文提出的语义关系图发现方法，并不仅限于中医药领域，理论上说如果有对应的领域本体知识库，就可以对该领域文献进行知识发现。

本文提出的语义关系图的发现方法是一种初步的探索，还有很多方面有待继续探索。比如在关键词的提取部分，结合基于统计的方法和一些额外的词库对新词进行发现；在语义关系预测部分，结合语法进行语义关系类型的预测。

参 考 文 献

[1] Berners-Lee T, Hendler J, Lassila O. The Semantic Web[J]. Scientific American, 2001

[2] Shadbolt N, Berners-Lee T, Hall W. The Semantic Web Revised [J]. IEEE Intelligent Systems, 2006, 21(3): 96-101

[3] Takaaki H, Sekine S, Grishman R. Discovering relations among named entities from large corpora [C]// Proceeding of Conference ACL 2004. Barcelona. Spain: Association for Computational Linguistics, 2004: 415-422

[4] Agrawal R, Imielinski T, Swami A. Mining Association Rules Between Sets of Items in Large Databases[C]// SIGMOD Conference. 1993: 207-216

[5] Tsai C H. MMSEG: A word identification system for Mandarin Chinese text based on two variants of the maximum matching algorithm[R]. 2000

[6] Chen H J, Wu Z H. Semantic Web Model, Methodology and Applications[M]. Springer-Verlag, GmbH, 2008

[7] Salton G, Wong A. A vector space model for automatic indexing [J]. Communications of the ACM, 1975, 18(11): 613-620

[8] Leskovec J, Grobelnik M, Milic-Frayling N. Learning sub-structures of document semantic graphs for document summarization [C]// KDD 2004 Workshop on Link Analysis and Group Detection(LinkKDD). Seattle, Washington

[9] Zhang Xiaogang, Chen Huajun. Ontology Based Semantic Relation Verification for TCM Semantic Grid[C]// ChinaGrid Annual Conference. ChinaGrid '09. Fourth, 2009: 185-191

[10] 何前锋，尹爱宁，刘静，等. 中医药同异名现象与标准研究[J]. 中国中医药信息杂志，2008(S1)