

分布式包装实时数据库 ARS 算法应用

李同英^{1,2}, 朱洪波²

(1.南通理工学院, 南通 226002; 2.南京邮电大学, 南京 210003)

摘要: **目的** 研究具有连续状态空间的复杂包装产品信息在分布式网络实时数据库中的查询方式。**方法** 通过结合增强学习(EL)和自适应共振结构神经网络(ARS), 给出一种基于增强学习的自适应共振结构神经网络算法——ELARS2。在 ARS2 算法中引入增强学习的选择和评估方式, 解决在 ARS2 算法中分类模式的查询问题。设计在存储空间中使用分布式网络实时数据库查询目标的仿真试验, 并用 2 种 ELARS2 算法(TDARS2 和 QARS2 算法)来实现, 并与经典的 EL 算法进行对比。**结果** 2 种 ELARS2 算法完成查询目标的平均时间显著小于经典的 EL 算法。**结论** 在 2 种 ELARS2 算法中, TDARS2 比 QARS2 效果更好。

关键词: 增强学习; ARS2 神经网络; TD 神经网络算法; Q 学习

中图分类号: TB487 **文献标识码:** A **文章编号:** 1001-3563(2017)11-0088-04

ARS Algorithm Application of Real-time Database of Distributed Packaging

LI Tong-ying^{1,2}, ZHU Hong-bo²

(1.Nantong Institute of Technology, Nantong 226002, China;

2.Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

ABSTRACT: The work aims to study the way of query in a distributed network real-time database for complex packaging product information with continuous state space. Combined with the enhanced learning (EL) and the adaptive response structure (ARS) neural network, the ELARS2 algorithm, ARS neural network based on enhanced learning was given. The way of selection and assessment methods for EL were introduced in ARS2 algorithm, so as to solve the query problems of sort mode in ARS2 algorithm. The simulation test in which the distributed network real-time database in the memory space was used to inquire targets was designed and it was achieved with two ELARS2 algorithms (TDARS2 and QARS2). Such algorithms were compared with the classical EL algorithm. The average time to complete target query in two ELARS2 algorithms was significantly less than the classic EL algorithm. In two ELARS2 algorithms, the effect of TDARS2 is better than QARS2.

KEY WORDS: enhanced learning; ARS2 neural network; TD neural network algorithm; Q learning

随着计算机辅助设计与制造在包装工程中的广泛使用, 包装产品数据信息都存储在相关数据库中。商业关系数据库提供的简单 SQL 查询只能满足普通管理的数据需求。随着大数据时代的到来, 目前在大型包装工程中, 设计与制造一体化对实时控制和包装设计的大数据处理提出了强烈需求。商业关系数据库已经不能满足这些处理需求, 只有分布式网络实时数据库才能满足目前包装敏捷协同制造过程中设计与制造一体化对大数据处理的需求。分布式网络实时数据库具有许多连续状态空间的复杂信息, 经典策略查

询表已不能满足要求^[1-3]。分布式网络实时数据库的数据查询是当前高新技术领域中的研究热点, 已受到业内学者广泛的关注^[4-6]。文献[7]提出的松弛响应关联维分析算法, 依据数据初始路径的关联维规则, 完成分布式网络实时数据库的数据查询, 该方法简单, 但是局限性较大。文献[8]将禁忌搜索启发式与遗传算法相结合构建数据查询路径方案, 完成分布式网络实时数据库的数据查询路径优化选择, 该算法适应性强, 但是路径选取误差大。文献[9]通过计算数据流量和路径阻碍, 运用改变伪随机查询状态及最优路径信

收稿日期: 2016-11-10

基金项目: 国家高技术研究发展计划(2014AA01A705)

作者简介: 李同英(1969—), 女, 博士, 南通理工学院高级工程师, 主要研究方向为数据挖掘、信息物理系统。

息更新,完成分布式网络实时数据库中的数据查询路径优化,该方法精确度较高,但是求解时间过长。文献[10]将改进模拟退火算法融合于 Dijkstra 算法中,完成分布式网络实时数据库中数据查询最优路径选取,避免了局部最优的问题,但是计算复杂、迭代时间长。针对上述问题,文中提出一种基于增强学习及自适应共振结构神经网络算法的分布式网络实时数据库数据查询方法,并通过模拟仿真实验对其相关性能进行验证。

1 ELARS2 神经网络原理及其算法设计

1.1 增强学习有关算法

1) 时间差神经网络。最简单的时间差(TD)神经网络算法称为一步 TD 神经网络算法,即 TD(0)。在该算法中,代理获得即时值时只向后返回一步,只迭代修改相邻的值函数^[11-12]。TD(0)收敛速度比较慢,若代理获得即时值时能够向后返回任意步,则算法的收敛速度可以提高很多,即 TD(λ)。

2) Q Learning 神经网络。Q Learning 是基于时差策略的增强学习方法。累积回报函数称为 $Q(r, z)$,指在状态 r 条件下完成动作 z 后系统期望状态所得到的累积回报。 $Q(r, z)$ 的状态-动作对应的 Q 值全部存储在二维的表里,在每个时步周期中, $Q(r, z)$ 值修改 1 次^[13-14]。

3) 自适应共振结构神经网络。自适应共振结构神经网络(ARS)含有输入层(比较层)及输出层(识别层)。算法既有输入到输出的前馈连接权,也有从输出到输入的反馈连接权^[15-16]。算法通过不断地将信号从输入层向输出层传输来进行识别、通过输出层向输入层进行比较来完成学习和工作过程。在识别与比较发生共振情况下,算法就完成了对 1 个输入学习的记忆及分类。目前,ARS 主要有 ARS1, ARS2, ARS3 3 种模型,其中 ARS1 用于表达二进制信息;ARS2 用于表达二进制信息及模拟信息;ARS3 用于进行分级搜索。文中以应用 ARS2 为主。

1.2 ELARS2 神经网络算法设计

ARS2 含有注意及定向 2 个子系统。注意子系统主要由输入层、输出层构成,定向子系统主要由重整模块和增益控制模块构成。输入预处理完成后,通过竞争方式,注意子系统选择与输入最合适的信号。定向子系统对该信号进行相似度测试,若通过则 ARS2 进行共振并调整当前单元的增益权值,若没有通过,则舍弃当前单元,继续在注意子系统中寻找其他单元以寻求合适的模式。如所有单元的模式均不符合要求,则给出新的输出单元。对 ARS2 来讲,相似度测试提供了模式信号学习和分类的稳定性需求。

在分布式网络实时数据库的具有连续状态空间数据查询过程中,有大量的状态和行为值需要查询。如果使用简单的查询表存储具有连续状态空间的状态和行为值,需要的存储空间将会出现几何级数的扩张,因此查询这些状态和行为值会变得极其困难;在状态空间存在大量的维度或广泛空间的情况下,分布式网络实时数据库的数据查询过程需要大量的内存。以上情况均可采用基于增强学习的函数逼近方法来解决。基于以上的需求,文中提出一种采用增强学习的 ARS2 算法。在算法中,ARS2 用来存储大量的连续状态空间和分类模式;此外,在算法中还引入基于增强学习的评估和选择机制,用来处理评估和选择已存储在 ARS2 中的分类模式,解决还需要人工处理评估和选择分类模式的问题。基于增强学习的 ARS2 (ELARS2)的具体原理:如果当前分类模式运行效果良好,就在 ARS2 中增大相应的增益权重;反之,在 ARS2 中减小相应的增益权重。当学习达到规定条件时就选取其他分类模式。ELARS2 算法控制的学习流程见图 1。

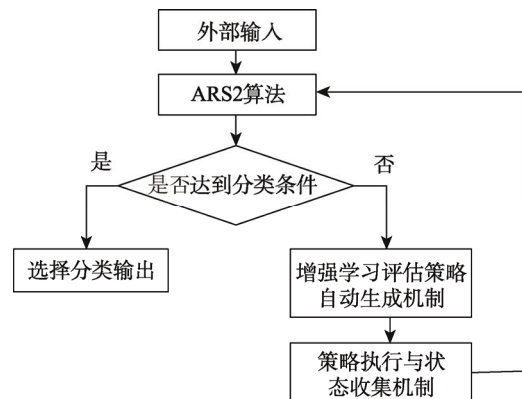


图 1 ELARS2 算法控制的学习流程

Fig.1 The learning flowchart of ELARS2 algorithm

ELARS2 算法如下所述。

- 1) 在 ELARS2 中对相关参数与增益赋予初始值。
- 2) 在 ELARS2 中输入初始信号。
- 3) 更改比较层的特征值,计算识别层的特征值,找出识别层中特征值最大的单元,统计当前循环次数;若达到规定的循环次数则退出循环,进入步骤 5。
- 4) 阈值测试,阈值一般介于 0.7~1,若测试值小于阈值,则放弃当前单元,然后返回步骤 3;若测试值大于阈值,则保留当前单元,退出循环,进入步骤 5。
- 5) 计算当前分类模式的评价值,改变单元的增益权值。
- 6) 改变比较层的特征值,计算增益权值次数,若达到规定的增益权值次数则退出循环;统计当前循环次数,若达到规定的循环次数则退出循环,否则返回步骤 5。

2 仿真设计

2.1 环境设置

包装产品分布式网络实时数据库的数据查找过程见图2。设置8个数据子库的存储空间,这8个数据子库的存储空间通过8扇门相互连通。现设d为目标数据子库的存储空间,初始时,查询起点可能位于这8个数据子库的存储空间中的任一位置。现要求它从任一初始位置找到一条最优路径到达目标数据子库的存储空间。

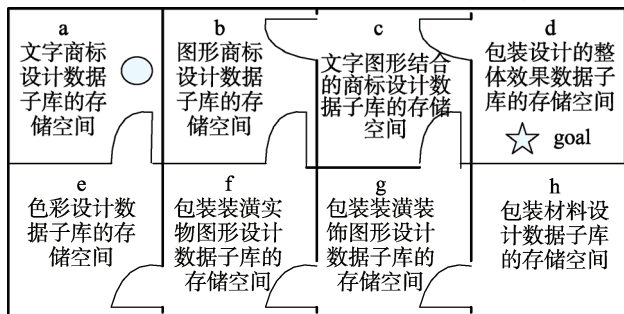


图2 包装产品分布式网络实时数据库的数据查询仿真环境
Fig.2 Data query simulation environment of packaging distributed network real-time database

2.2 算法

考虑到具体问题,采用二进制代码编写输入模式,采用ARS1代替ARS2,建立的ELARS可简化算法结构,提高学习速度。具体步骤如下所述。

1) 输入层有5个输入神经元,用来输入a, b, c, d, e, f, g, h 8个分布式网络实时数据库数据子

库存储空间的5位二进制编码。

2) 输出层有8个输出神经元,神经元a, b, c, d, e, f, g, h 分别代表相应的分布式网络实时数据库数据子库的存储空间。

3) ARS中的前馈权值和反馈权值存储着数据库查询当前的查询规则。

4) 当输入查询所处的当前数据子库的存储空间代码时,ARS存储的规则将通过竞争选出一个优胜状态,作为数据查询的下一个数据子库的存储空间。

5) 通过评估模块评估输出结果,若效果较好,则加强相应的增益权值,反之就减小相应的增益权值并返回进行步骤4—5。

3 仿真结果与分析

仿真试验在传统的增强学习策略查询表和ARS基础上实现了Q Learning及TD学习。采用2种ELARS算法:基于Q Learning的QARS和基于TD的TDARS,并和经典的增强学习方法进行对比;采用Q Learning和TD作为基线,试验中的各算法与其进行对比。完成时间以产生相同的最优策略为准。Q Learning参数、TD参数、QARS参数、TDARS参数的变化情况见表1,可知TDARS相对QARS来说效果更好。为了追求输入状态上的泛化,ARS被设计为具有隐层的结构——输入神经元和输出神经元的数目可变。当算法具有隐层,并且引入了不同的输入状态表示时,为了收敛,学习过程需要更多的迭代次数。通过分析表1,

表1 不同算法参数变化情况
Tab.1 Different algorithm parameters table

算法	学习率	折扣因子	平均完成时间 (10次)/s	最佳完成时间 (10次)/s	完成时间标准差 (10次)/s	最佳完成时间下 所需迭代次数
Q Learning	0.1	0.9	21.718	21.331	0.5615	112 080
	0.5	0.9	4.857	4.737	0.1114	24 772
	0.9	0.9	2.994	2.834	0.1402	14 576
	0.7	0.3	0.644	0.620	0.0214	2965
	0.7	0.7	1.285	1.242	0.0376	6113
TD	0.1	0.9	17.582	17.144	0.3816	91 053
	0.5	0.9	4.063	4.026	0.0351	21 232
	0.9	0.9	2.344	2.304	0.0400	11 872
	0.7	0.3	0.547	0.521	0.0228	2575
	0.7	0.7	1.228	1.022	0.1814	5097
QARS	0.1	0.9	13.032	10.585	3.2150	52 958
	0.5	0.9	3.021	2.954	0.1248	13 783
	0.9	0.9	1.735	1.352	0.3627	5737
	0.7	0.3	0.654	0.591	0.0568	1772
	0.7	0.7	1.052	0.992	0.0529	4078
TDARS	0.1	0.9	9.614	8.913	0.6646	50 905
	0.5	0.9	2.504	2.233	0.2568	12 722
	0.9	0.9	1.619	1.502	0.1201	8411
	0.7	0.3	0.404	0.360	0.0409	1905
	0.7	0.7	0.812	0.741	0.0661	3958

可看出 ELARS 算法完成学习的平均时间显著小于经典增强学习算法, TDARS2 比 QARS2 平均查询时间更短些。

4 结语

包装产品分布式网络实时数据库作为包装产品系统的核心, 其数据查询性能的高低对包装产品系统具有重要作用, 因此, 研究分布式网络实时数据库的数据查询问题具有深远的发展意义。包装产品分布式网络实时数据库具有许多连续状态空间的复杂包装产品信息, 使用经典策略查询表存储这些行为值和状态值, 需要的存储空间将会呈现几何级数的扩展, 那么寻找这些行为值和状态值会变得相当困难。为此, 文中提出了一种基于增强学习的自适应共振结构神经网络 ELARS2 算法, 仿真结果表明, 2 种 ELARS2 算法完成查询目标的平均时间显著小于经典的 EL 算法, 其中 TDARS2 比 QARS2 的平均查询时间更短。

参考文献:

- [1] WANG Y, DE SILVA C W. A Machine-learning Approach to Multi-robot Coordination[J]. Engineering Applications of Artificial Intelligence, 2008, 21(3): 470—484.
- [2] YEN G, HICKEY T. Reinforcement Learning Algorithms for Robotic Navigation in Dynamic Environments[J]. ISA Transactions, 2004, 43(2): 217—230.
- [3] BUSONI L, BABUSKA R, SCHUTTER B. A Comprehensive Survey of Multi-agent Reinforcement Learning[J]. IEEE Trans on Systems, Man, and Cybernetics: Applications and Reviews, 2008, 38(2): 156—172.
- [4] LIU F, GEOK S N, CHAI Q. RLDDE: a Novel Reinforcement Learning-based Dimension and Delay Estimator for Neural Networks in Time Series Prediction[J]. Neurocomputing, 2007, 70(7): 1331—1341.
- [5] TOUSSAINT K, NPOULIOTS S, MONTAMBAUL T. Transmission Line Maintenance Robots Capable of Crossing Obstacles: State-of-the-ARS Review and Challenges Ahead[J]. Journal of Field Robotics, 2009, 26(5): 477—499.
- [6] 欧阳相成. 变异遗传散布控制执行算子的数据库分类查询[J]. 科技通报, 2014(4): 143—145.
OUYANG Bo-cheng. Query Algorithm of Database Classification Based on Genetic Variation with Distribution Control of Executive Operator[J]. Bulletin of Science and Technology, 2014(4): 143—145.
- [7] 龙琼, 胡列格, 张谨帆, 等. 不确定性环境下个性化路径选择的多属性决策方法[J]. 中国公路学报, 2014, 27(9): 105—110.
LONG Qiong, HU Lie-ge, ZHANG Jin-fan, et al. Multi-attribute Decision Making for Personalized Route Selection under Uncertain Condition[J]. China Journal of Highway and Transport, 2014, 27(9): 105—110.
- [8] 冯洪玉, 李艳翠. 机器人路径选择优化方法研究与仿真[J]. 计算机仿真, 2013, 30(8): 372—375.
FENG Hong-yu, LI Yan-cui. Research and Simulation of Robot Path Optimization Methods[J]. Computer Simulation, 2013, 30(8): 372—375.
- [9] 安毅生, 袁绍欣, 赵祥模, 等. 基于蚁群算法的动态路径选择优化方法[J]. 交通运输系统工程与信息, 2014, 14(3): 97—103.
AN Yi-sheng, YUAN Shao-xin, ZHAO Xiang-mo, et al. Optimization of Dynamic Route Choice Based on Ant Colony Algorithm[J]. Journal of Transportation Systems Engineering and Information Technology, 2014, 14(3): 97—103.
- [10] 杨兆升, 于尧, 孙文飞. 城市路网下应急救援路径选择方法[J]. 哈尔滨工程大学学报, 2013(12): 1566—1572.
YANG Zhao-sheng, YU Yao, SUN Wen-fei. Routing Problem for Emergency Rescue Vehicles in Urban Road Networks[J]. Journal of Harbin Engineering University, 2013(12): 1566—1572.
- [11] 苏莹莹, 王宛山, 王建荣, 等. 基于神经网络增强学习算法的工艺任务分配方法[J]. 东北大学学报(自然科学版), 2009, 30(2): 279—282.
SU Ying-ying, WANG Wan-shan, WANG Jian-rong, et al. Research on Task Allocation of Process Planning Based on Reinforcement Learning and Neural Network[J]. Journal of Northeastern University (Natural Science), 2009, 30(2): 279—282.
- [12] 柴江松, 王琪, 刘洪豪. 印刷网点微观图像阈值分割算法研究[J]. 包装工程, 2015, 36(13): 115—121.
CHAI Jiang-song, WANG Qi, LIU Hong-hao. Threshold Segmentation Algorithm for Printing Dot Microscopic Image[J]. Packaging Engineering, 2015, 36(13): 115—121.
- [13] 魏书宁, 王耀南, 印峰, 等. 基于 k-最近邻分类增强学习的除冰机器人抓线控制[J]. 控制理论与应用, 2012, 29(4): 470—476.
WEI Shu-ning, WANG Yao-nan, YIN Feng, et al. Line-grasping Control of de-icing Robot Based on k-nearest Neighbor Reinforcement Learning[J]. Control Theory & Applications, 2012, 29(4): 470—476.
- [14] 刘岩. 增强学习的 PID 控制参数优化快速整定算法[J]. 计算机测量与控制, 2014, 22(2): 467—470.
LIU Yan. Enhanced Learning of PID Controller Parameters Optimized for Fast Tuning Algorithm[J]. Computer Measurement & Control, 2014, 22(2): 467—470.
- [15] 吴庆涛, 曹再辉, 施进发. 傅里叶变换耦合曲率尺度空间的图形检索算法[J]. 包装工程, 2016, 37(13): 157—164.
WU Qing-tao, CAO Zai-hui, SHI Jin-fa. A Graphical Shape Retrieval Algorithm Based on 2D Fourier Transform Coupled with Curvature Scale Space[J]. Packaging Engineering, 2016, 37(13): 157—164.
- [16] 段勇, 崔宝侠, 徐心和. 进化强化学习及其在机器人路径跟踪中的应用[J]. 控制与决策, 2009, 24(4): 532—536.
DUAN Yong, CUI Bao-xia, XU Xin-he. Evolutionary Reinforcement Learning and Its Application in Robot Path Tracking[J]. Control and Decision, 2009, 24(4): 532—536.