

DDM 中基于历史信息排序的区域匹配算法

王卓¹ 冯晓宁² 刘廷宝²

(哈尔滨工程大学水下机器人技术国防科技重点实验室 哈尔滨 150001)¹

(哈尔滨工程大学计算机科学与技术学院 哈尔滨 150001)²

摘要 DDM 实现的关键是更新区域和订购区域的匹配判断,其匹配算法的设计决定着分布式仿真的效率和性能。针对直接匹配算法需要维护的区域相交索引表和相交信息表的存储空间不断扩大的问题,首先提出区域聚合的基本思想及其实现方式,然后提出将历史信息排序加入到区域匹配算法中,以使算法在有效利用历史信息的基础上,提高区域匹配的运行效率。给出了算法的基本思想和具体步骤,通过应用实例详细解释了算法的具体实现过程,最后通过仿真实验,对比分析了所提算法的优缺点。

关键词 HLA, RTI, 数据分发管理, 区域匹配算法

中图分类号 TP391 **文献标识码** A

Region Matching Algorithm Based on Historical Information Sorting in DDM

WANG Zhuo¹ FENG Xiao-ning² LIU Ting-bao²

(National Key Laboratory of Science and Technology on Autonomous Underwater Vehicle of

Harbin Engineering University, Harbin 150001, China)¹

(College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China)²

Abstract The key point of DDM implementation is to match the update region sets with the subscription region sets. The efficiency and performance of the distributed simulation system is determined by the design of matching algorithm. The matching algorithm must maintain the index form and the information form of region intersection. And the problem is that the forms are constantly extended. Firstly, the idea and realization of region aggregation were provided. Secondly, historical information sorting was added to the region matching algorithm. Therefore, the running efficiency of region matching was improved on the base of using historical information. The basic idea and detail process of the algorithm were provided in the paper. The implementation process of algorithm was amply explained by the application example. Finally, the advantage and disadvantage of the algorithm were analyzed by the simulation data.

Keywords HLA, RTI, Data distribution management, Region matching algorithm

数据分发管理(Data Distribution Management, DDM)是 HLA 的重要管理服务之一,它是在声明管理(DM, Declaration Management)实现基于类的数据过滤后,实现基于值的过滤,其目的是进一步减少大规模仿真中盟员接收冗余数据的可能性和网络中的数据流量。DDM 通过判断更新区域和订购区域是否交叠来决定是否将更新数据发送给订购盟员。DDM 实现的关键是更新区域和订购区域的匹配判断,其匹配算法的设计决定着分布式仿真的效率和性能。为此,实现高效、精确的匹配一直是区域匹配算法追求的目标。现有的区域匹配算法有直接匹配法、网格法^[1,2]、全域空间层次划分法等。其中直接匹配法是在定义好的路径空间上建立多维坐标系,排列发布区域集合与订购区域集合的相同维的坐标值,通过扫描和维护这个排序表来完成区域匹配的。随着仿真规模的不断增加,发布和订购区域也将不断增多,实现经典排序算法时需要维护的区域相交索引表和相交信息表的存储空间

也将不断扩大。为了提高维护区域索引表的效率,从而提高数据传递的速率,本文提出了基于区域聚合和维护一个历史排序表的区域匹配算法。

1 区域匹配的相关定义

在 HLA1516 接口规范中,数据分发管理的关键是区域匹配(区域交叠)的判断^[3]。公布盟员通过“带区域登记对象实例”或“为更新关联区域”等服务来声明自己的更新区域,从而表达向外发送数据的能力;订购盟员通过“带区域订购对象属性”等服务来声明自己感兴趣的区域,从而表达从外部接收数据的兴趣。更新区域集和订购区域集的定义如下。

定义 1(更新区域集) 用于发送交互或更新实例属性的区域实现的集合。盟员利用它来表明发送数据的限制条件。

定义 2(订购区域集) 用于类属性订购或交互类订购的区域实现的集合。盟员利用它来表明接收数据的限制条件。

到稿日期:2010-11-25 返修日期:2011-04-18 本文受黑龙江省青年科学基金项目(2007Q0502-00),哈尔滨市青年科学基金项目(RC2009QN 010039),哈尔滨工程大学校基础研究基金项目(002060260722),水下智能机器人技术国防科技重点实验室开放课题研究基金项目(2010004)资助。

王卓(1977-),女,博士,副教授,E-mail:wangzhuo@hrbeu.edu.cn(通信作者);冯晓宁(1976-),男,博士,副教授。

更新区域集和订购区域集可随着盟员的兴趣或实体在仿真环境中状态的变化而变化。

定义 3(范围交叠) 同一维度上两个范围 $A = \{a_{lower}, a_{upper}\}$ 、 $B = \{b_{lower}, b_{upper}\}$ 交叠, 当且仅当 $a_{lower} = b_{upper}$ 或 $(a_{lower} < b_{upper} \text{ 且 } b_{lower} < a_{upper})$ 。

定义 4(区域交叠) 两区域交叠, 当且仅当这两个区域的所有公共维度上的范围交叠; 如果两个区域没有公共维度, 则不交叠。

定义 5(区域集交叠) 一个更新区域集与一个订购区域集交叠, 当且仅当这两个区域集中至少各存在一个区域, 且这两个区域交叠。

当一个更新区域集与一个订购区域集交叠时, RTI 在这两个区域集所属的盟员之间建立一个连接关系, 从而订购盟员能够收到公布盟员的属性更新。当一个更新区域集与一个订购区域集没有交叠时, RTI 就不会在这两个区域集所属的盟员之间建立数据连接关系, 从而订购盟员也就收不到公布盟员的属性更新。

2 区域聚合

为了使改进后的排序算法更加高效, 本文提出在匹配前先进行一次区域聚合, 使得联邦成员的发布和订购区域数量适当减少, 进而使排序匹配次数适当减少, 以提高效率。

区域聚合的概念是 van Hook^[4] 等人提出的。它是将联邦成员声明的相邻近的发布订购区域合并, 使得区域的数量减少, 进而减少区域匹配的次数。随着仿真的进行, 规模不断增大, 节点不断增多, 发布和订购区域也将不断增加, 这种聚合的方式可以有效地提高 DDM 的执行效率。区域聚合的步骤分为两步: (一) 将某维分成若干段; (二) 在段内对区域进行聚合。

2.1 动态平均分段法

将某维分成若干段是根据仿真系统定义的路径空间, 在各个维上, 分别对区域的范围进行聚合, 即在第 i 维上将整个维分成若干个子段, 也即 $R_i = \{r_1, r_2, r_3, \dots, r_k\}$, 其中应满足以下条件: $R_i = r_1 \cap r_2 \cap r_3 \cap \dots \cap r_k$, 且对于 $\forall i, j (i, j \leq k; i \neq j)$ 有 $r_i \cap r_j \neq \emptyset$ 。段的划分对于区域聚合起到决定性影响, 段划分的粒度过小, 区域范围将被拆得过于琐碎。粒度过大, 则过多的区域将在一个段内聚合, 导致聚合比较次数过多, 导致其效率低下。段内聚合的算法直接影响聚合的效率, 过于复杂的算法给 CPU 带来的负担将降低整个区域匹配的效率。为了在不增加算法开销的前提下得到粒度适中的段划分方法, 提出了动态平均分段和中心比较的聚合方式。所谓动态平均分段, 即在进行分段时动态设定分段个数, 平均划分。个数为 1 时, 整个一维退化为一个段。仿真过程中可根据仿真的规模、区域的个数, 对段进行动态划分。

2.2 中心比较的聚合

在动态平均分段的基础上, 各个子段中的区域范围将按照某种规则进行聚合, 聚合后生成的新区域范围代替原有的区域范围。

所谓中心比较的聚合方式, 即对中心落在段内的区域, 比较两两之间区域的中心距离是否小于设定好的某个值。定义中心比较参数 λ , 小于则视为可聚合区域, 当设置 λ 为无穷大时, 不进行区域聚合运算, 聚合后的区域为聚合前原始区域的

最大区域。中心比较的形式化描述如下:

$Mid(X_K) = (X_{K1} + X_{K2}) / 2, K \in (A, Z)$ 的区域, X 是维数。

当 $Mid(X_M) - Mid(X_N) < \lambda$ 时, M 和 N 区域聚合成功, 聚合后的区域为 $X_{MN} (\min(X_{M1}, X_{N1}), \max(X_{M2}, X_{N2}))$ 。其中 $M, N \in (A, Z)$, λ 为预先设定好的中心比较参数。

以图 1 为例, 在 X 维上, 某划分的段有 A, B, C, D 4 个区域, 对段内的各区域, 计算其 Mid 值, 两两比较, 当差值小于 λ 时, 即可聚合。当图中 $Mid(X_A) - Mid(X_C) < \lambda$ 时, A 和 C 两区域聚合后的结果 $X_{AC} \in (X_{A1}, X_{C2})$ 将代替原有区域。

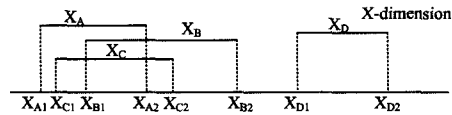


图 1 X 维区域聚合实例

仿真过程中, 中心比较参数 λ 的大小可以动态设定。当 λ 较小时, 聚合成功的区域较少, 精度较高; 当 λ 较大时, 聚合成功的区域多, 但因为聚合后的区域相比原区域将有所扩大, 所以同时会丧失精度。在仿真过程中, 可以根据实际情况, 按需求进行动态调节。

3 基于历史信息排序的区域匹配算法

在 DDM1516 标准中, 区域有以下特点: 在“绝大多数区域的移动都具有连续性的特点”的前提下, 区域的变化不会特别剧烈, 所以根据区域变化修改历史信息表是可行的。基于以上事实, 在排序算法的基础上加入历史信息, 对历史信息用更新排序表来替代排序算法中最耗时的部分——建立排序表, 以提高排序效率。而为了能够分担 RTI 区域匹配的负担, 决定将区域匹配的工作分担至联邦成员本地, 即潜在的发布者处, 每个潜在发布者都维护和自己发布区域相关联的各个维上的订购信息。这样, 需要进行数据分发时, 只需在本地进行区域匹配。对于匹配成功的区域所有者, 进行数据的传送。

本地匹配由于只需匹配本地发布信息 and 所有的订购信息, 因此相比 RTI 匹配所有的发布信息和订购信息, 其效率也将提高。

3.1 区域数据的存储

要实现以上的算法思路, 首先各发布信息的联邦成员需要维护几个数据结构。

1) 订购区域对象:

```
Structure
{
    String subRegionID; // 联邦成员编号, 为 0 表示为本地发布区域
    Float range[ ]; // 各维的范围
} subRegion
```

2) 维边界有序标示表:

```
Structure
{
    String subRegionID; // 联邦成员编号, 为 0 表示为本地发布区域
    Int nodeType; // 标示该点的上下层类型, 0 表示下届, 1 表示上届
    Float Value; // 边界值
} Sequence
Sequence Seq[100]; // 维边界有序标示表
```

3) 维相交向量表:

Structure

```
{
    String subRegionID; //联邦成员编号,为0表示为本地发布区域
    Int overLap; //标示是否相交,0标示未相交,1表示相交。
} Overlapping
Overlapping Olap[100]; //维相交向量表
```

3.2 算法步骤

建立了所需数据结构之后,当非本地成员的订购区域发生改变时,只需从数据结构中删除原区域的相关信息,然后重新插入新的区域信息,并重新进行区域匹配计算即可。因为所采用的是发送方执行的数据过滤方式,所以当订购区域变化时,区域匹配计算分散到了各个发送方成员,起到了一定的负载均衡作用,不会成为系统可伸缩性的瓶颈。下面给出算法的一个具体描述,包括数据结构初始化、区域相交性判断和结果综合3个步骤。

算法1 基于历史信息排序的区域匹配算法

Step1 系统初始化

- 1) 对符合聚合条件的区域进行聚合运算。
- 2) 对每个聚合后的订购区域分配ID标识,根据区域信息构建订购区域对象。
- 3) 根据订购区域对象的范围信息,对每维进行排序,构建维边界有序标示表,并保存。
- 4) 依据改进后的排序算法,同0区域(发布区域)匹配,构建维相交向量表并保存。

Step2 区间相交性判断

- 1) 首先根据历史聚合结果,看是否符合聚合到某一区域。符合则加入,不符合则自成一个区域。
- 2) 在订购区域对象中,将相应订购区域对象各范围表range的上下界值置为新修改的值(或增加、删除)。
- 3) 根据新的区域值,对相应维的边界有序标示表进行更新操作(包括修改、增加或删除)而非重构操作。
- 4) 根据新的边界标示表,同0区域(发布区域)重新匹配,构建维相交向量表并保存。
- 5) 对每一维上的相交向量表进行与操作,即为相交区域。

Step3 结果综合

遍历与本地发布区间相交性发生变化的所有相同维上的区间相交性。如果都相交,则有覆盖关系;否则没有覆盖关系,此时输出新的匹配结果,算法结束。

3.3 算法实例

设置二维路径空间中,发布区域为U,有3个订购区域,分别为S1、S2和S3。在X维上,3个订购区域的范围分别为S1(1,5),S2(11,16),S3(12,17),发布区域为U(2,13);在Y维上,3个订购区域范围分别为S1(4,8),S2(10,17),S3(1,5),发布区域为U(7,15),如图2所示。

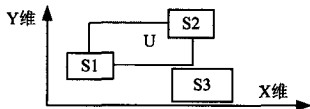


图2 订购和发布区域匹配示例

首先对区域进行聚合。由于只有S2和S3在X维上符合聚合条件,因此聚合后的区域范围为S23(11,17)。根据以上聚合后的范围,初始化订购区域对象。根据区域对象,在X

维和Y维上构建的边界有序标识如表1和表2所列。

表1 X维上的边界有序标识

区域ID	1	0	1	23	0	23
上下界	0	0	1	0	1	1
Value	1	2	5	11	13	17

表2 Y维上的边界有序标识

区域ID	3	1	3	0	1	2	0	2
上下界	0	0	1	0	1	0	1	1
Value	1	4	5	7	8	10	15	17

由以上二维边界有序表可看出区域聚合的效果,经聚合的X维匹配次数将减少。以上两个表将保存,并根据区域的更新而更新,对上表进行扫描并构建X维相交表。

维相交向量表的构建方式:从左向右扫描边界有序表,在遇到0区域的低边界前,对订购区域的低边界,在相交表中置1;对订购区域的高边界,在相交表中置0。遇到0区域的低边界后,对所有遇到的订购区域置1,直到遇到0区域的高边界,扫描结束。

扫描X维有序表,遇到S1的低边界,相交表置1;然后是0区域,S1的高边界,因为之前遇到0区域,S1仍然是1;遇到S23,置1;遇到0区域的高边界,扫描结束,结果如表3所列。同样扫描Y维有序表,得到相交表结果,如表4所列。

表3 X维相交结果

区域ID	1	23
是否相交	1	1

表4 Y维相交表结果

区域ID	1	2	3
是否相交	1	1	0

对各个相交表进行与操作,得出S1和S2区域与发布区域相交。当有发布或订购区域更新时,首先更新订购区域对象,然后对可合并的区域进行合并,再根据新的订购区域,在边界有序标示表中更新区域,并重新扫描,生成相交表,进而得出相交区域。

4 仿真实验及分析

仿真试验假设在 1024×1024 的二维区域中进行,系统分别生成1000、2000、3000、4000、5000个联邦订购成员,每个联邦成员分别在每一维0~1024间随机选择距离小于100的两个值作为订购区域上下界。

由于改进后的排序区域匹配算法,首先需要对订购区域进行聚合,而区域聚合时各个维上分段的个数N以及中心比较参数 λ 将直接影响聚合的结果,因此分别设置每维分段数为0、5、10、15、20,对应设置中心比较参数 λ 为无穷大、20、10、7.5。段数为0, λ 为无穷大时不进行区域聚合。

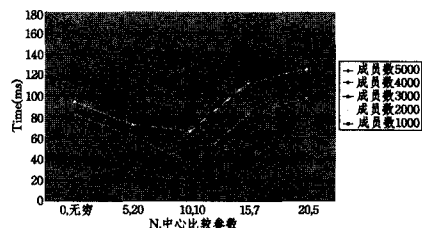


图3 聚合及构建维边界有序表的时间

(下转第214页)

即用 V 系统的基函数作为前馈神经网络中的隐层神经元,构造基于 BP 算法的 V 正交神经网络,在一元函数逼近方面其收敛性能优于传统的 BP 神经网络、小波网络与 Legendre 正交基网络。

参考文献

[1] Hagan M T, Demuth H B, Beale M H. Neural Network Design [M]. 戴葵,译.北京:机械工业出版社,2002

[2] Zhang Qing-hua, Benveniste A. Wavelet Networks [J]. IEEE T Neural Networ, 1992, 3(6): 889-898

[3] Zhang Jun, Walter G G, Miao Yu-bo, et al. Wavelet neural networks for function learning [J]. IEEE T Signal Proces, 1995, 43(6): 1485-1491

[4] 何新贵, 许少华. 过程神经网络[M]. 北京: 科学出版社, 2007

[5] 张雨浓, 钟童科, 李巍, 等. Legendre 正交基前向神经网络及其

其权值直接确定法[J]. 暨南大学学报: 自然科学版, 2008, 29(3): 249-253

[6] 张雨浓, 陈裕隆, 姜孝华, 等. 一种权值直接确定及结构自适应的 Chebyshev 基函数神经网络[J]. 计算机科学, 2009, 36(6): 210-213

[7] 肖秀春, 张雨浓, 姜孝华, 等. 第二类 Chebyshev 前向神经网络权值直接确定及结构自适应确定[J]. 大连海事大学学报, 2009, 35(1): 80-84

[8] Song Rui-Xia, Ma Hui, Wang Tian-Jun, et al. Complete orthogonal V-system and its applications[J]. Commun Pur Appl Anal, 2007, 6(3): 853-871

[9] Feng Yu-yu, Qi Dong-xu. A sequence of piecewise orthogonal polynomials[J]. SIAM J Math Anal, 1984, 15(4): 834-844

[10] 熊刚强. 一类分段多项式正交函数系的构造与应用[D]. 广州: 中山大学, 2010

(上接第 204 页)

在以上条件下得到仿真区域聚合及构建维边界有序标表示的时间变化趋势,如图 3 所示。在 $N=10, \lambda=10$ 时,仿真区域聚合及构建边界有序表时效率最高。

由于本算法为发布成员本地区域匹配,故随机生成 1000 个联邦发布成员,分别同上面生成的联邦订购成员进行区域匹配,取平均时间值作为区域匹配时间,得到平均匹配时间,如图 4 所示。分段数为 0, λ 为无穷时,不进行区域聚合的平均匹配时间较长,分段数少。 λ 值设置较大时,平均匹配效率较高。由于分段数少, λ 值大时,区域聚合度较大,会直接导致区域数量减少,进而匹配次数也随之减少,故效率较高。但由于对原来的区域有所扩大,因此会有部分虚假链接,但这不会出现丢失。

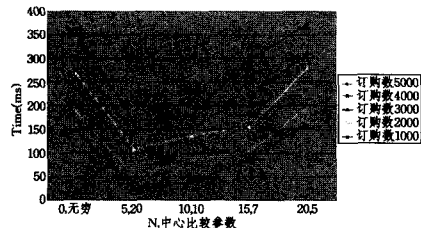


图 4 平均匹配时间

在以上构建好的订购区域中,设定 $N=10, \lambda=10$,随机分别选择 5%、5%~10%、10%~20%、20%~50%、50%~100% 的区域进行更新,得到平均匹配时间的仿真结果如图 5 所示。

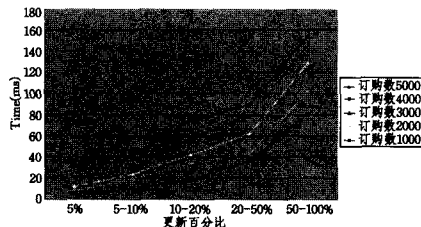


图 5 更新时间

最后比较经典排序区域匹配算法和本文提出的改进排序匹配算法的整体效率。在仿真过程中,设置改进排序匹配算法中每维分段数 $N=10$,中心比较参数 $\lambda=10$,随机在二维路径空间中 1%~100% 区域内,随机生成 5000 个订购区域,生成一个本地发布区域,随机分别选择 5%、5%~10%、10%~

20%、20%~50%、50%~100% 的订购区域进行更新。对于每次情况模拟 100 次,取其平均值为仿真结果,并进行对比。通过图 6 可以看出,改进后的算法在整体效率上相较于经典算法有了较大的提升。

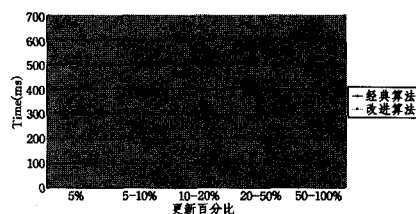


图 6 经典和改进算法的整体效率时间比较

结束语 本文在直接区域匹配算法的基础上,引入区域聚合的思想,同时将历史信息排序加入到区域匹配算法中,通过有效利用历史信息简化区域相交索引表和相交信息表的方式来提高区域匹配的运行效率。在仿真实验中考察了算法的维边界有序表构建时间、平均匹配时间和维边界有序表更新时间,并以这 3 个时间为考察指标,实验结果表明所提算法在时间效率上具有优势。

参考文献

[1] Raczky C, Tan G, Yu J. A Sort-based DDM Matching Algorithm for HLA [J]. ACM Transactions on Modeling and Computer Simulation, 2005, 15(1): 14-38

[2] Morse K L, Petty M D. Data Distribution Management Migration from DoD 1. 3 to IEEE 1516 [C] // Proceedings of the 5th IEEE International Workshop on Distributed Simulation and Real Time Applications. Aug. 2001: 58-65

[3] Van Hook D, Calvin J. Data Distribution Management in RTI [C] // Spring Simulation Interoperability Workshop, No. 98S-SIW-206. 1998: 78-87

[4] 王磊, 张慧慧, 李开生, 等. 基于动态 R-树结构的 DDM 区域匹配算法[J]. 计算机工程, 2008, 34(3): 56-58

[5] 刘胤田, 唐常杰, 曾涛, 等. 基于空间填充曲线的数据分发区域匹配[J]. 系统仿真学报, 2007, 19(4): 63-71

[6] 张家祥. HLA 仿真技术应用教程[M]. 北京: 国防工业出版社, 2007: 86-97

[7] Raczky C, Tan G, Yu J. A Sort-based DMM Matching Algorithm for HLA [J]. ACM Transactions on Modeling and Computer Simulation, 2005, 15(1): 96-110