

文章编号: 1001-2486(2002)04-0048-05

一种基于多层次网格的高效相关过滤方法*

徐春蕾, 曾亮, 李思昆

(国防科技大学计算机学院, 湖南长沙 410073)

摘要: 在分布式仿真系统中, 网上流动的大量冗余数据严重影响了系统的可伸缩性。基于均匀网格的相关过滤法减少了冗余数据, 但具有匹配不精确, 格子尺寸单一, 难以适应所有实体的缺点。提出了一种基于多层次网格的相关过滤方法, 克服了均匀网格法的上述缺点, 并继承了均匀网格法处理速度快的特点, 具有快速灵活的特点。

关键词: 相关过滤; 发布区域; 定购区域; HLA; 数据分发管理

中图分类号: TD391.9 文献标识码: A

An Efficient Multi-level Grids Based Relevance Filtering Method

XU Chun-lei, ZENG Liang, LI Si-kun

(College of Computer, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: In the distributed simulation system, lots of redundant data flowing on the net seriously restricts the system scalability. The relevance filtering base on even-grid reduces the redundant data while it is inaccurate in region matching. Because the sizes of all cells in even-grid are identical, it is difficult to fit all entities in the simulation. A relevance filtering method based on a multi-level grid is put forward which overcomes the above shortcomings and is characterized by agility and speediness.

Key words: relevance filtering; publishing region; subscribing region; HLA; data distribution management

分布式交互仿真集成了计算机、网络、图形图像、数据管理等多方面的技术, 通过模拟的方法对各种已存在或待开发的系统进行研究开发或验证, 具有方便、快速、低成本、危险性小等特点。1994年, 北约在全世界范围内成功的组织了一次基于分布式交互仿真的军事演练, 并因此引起了世界范围的分布式交互仿真热。现在, 分布式交互仿真在军事、决策系统、娱乐等众多方面都有应用。

以往构建分布式交互仿真通常遵循 1993 年提出的分布式交互仿真协议 (DIS^[1])。但随着分布式交互仿真规模的发展, DIS 逐渐不能满足需求。美国国防部 (DoD) 为了克服 DIS 的缺陷, 并保障新仿真与已有仿真之间的互操作性, 于 1995 年开始开发一种新的适用于分布式交互仿真的高层体系结构 (HLA^[2])。并在 1996 年把 HLA 作为国防部分布式交互仿真开发的标准体系结构。自从 HLA 标准制定以来, 各种遵循 HLA 标准的运行时支撑结构 (RTI^[2]) 已经问世。HLA 提出了一些相关过滤机制, 但如何在大型分布式交互仿真中高效的实现这些机制还是个难题。

1 分布式交互仿真的可伸缩性与相关过滤

实际需求的不断增长使得分布式交互仿真的规模日益扩大, 大量的信息在网络上流动。以前使用的 DIS 采用广播方式, 每个结点接收来自其它结点的所有信息。但实际上, 由于分布式交互仿真的局部性, 各结点接收的很多信息是无关信息。广播方式极大的浪费了网络资源和结点的处理资源, 严重影响了分布交互仿真的可伸缩性。

为了减少网上传输以及结点接收的冗余数据, 使数据尽量只发给需要它的结点而进行的数据过滤称为相关过滤。HLA 规定了两种相关过滤方法, 一种是实体可以定购它所需的类、类的属性以及交互, 另一种是针对每个特定的属性或交互, 实体可以在路径空间中定义定购区域或发布区域。路径空间可以由仿真实体定义, 只有当一实体的发布区域与另一实体的定购区域在同一路径空间并且相交

* 收稿日期: 2002-01-10

基金项目: 国家 863 高技术资助项目 (2001AA115132)

作者简介: 徐春蕾 (1976-), 女, 博士生。

时,前者才可能向后者发送信息。在实际仿真中,常把实际系统中信息的可到达区域作为相应的发布区域,而把信息的可感知区域作为相应的订购区域。这种机制在 RTI 中的数据分发管理 (DDM^[3]) 模块中实现。

判断订购区域和发布区域是否相交的方法主要有均匀网格法^[4]和分层过滤法^[5]。分层过滤法将一个路径空间中所有区域组成的集合分成几个子集,分别用规则体包围起来;对来自两个不同子集的区域,只有两个相应规则体相交时,才需要对它们进行相交判断。采用这种方法,结果精确,但如何合理划分子集是个比较困难的问题;每当一个区域发生变化时,都要重新计算所属子集的规则体,重新判断该规则体与其它规则体的相交,必要时还要进一步判断来自该规则体的区域是否与来自其它规则体的区域相交。

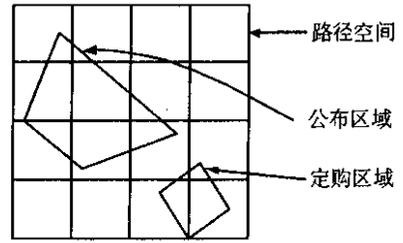


图1 路径空间中的均匀网络

Fig.1 Even-grid in the routing space

均匀网格法将整个路径空间划分成格子大小相等的网格,若某个发布区域与某个订购区域有共同相交格子,则认为它们相交(见图1)。每个格子对订购区域与之相交的仿真实体进行登记,在计算一个仿真实体的信息发送对象时,首先计算与该实体发布区域相交的格子,并认为这些格子登记的仿真实体就是发送对象。网格法不需要直接判断订购区域和发布区域是否相交,在区域发生变化时也很容易重新计算与它相交的格子,速度快。但是,由于存在两个区域与同一格子相交但两区域不相交的情况,所以网格法存在匹配不精确的缺点。提高网格法精确性的方法是缩小格子尺寸。但缩小尺寸之后,与每个发布区域相交的格子数增加,这些格子登记的订购区域的重复性增强,使得所需存储资源大量增加,查找相应订购区域的速度降低。

2 基于多层次网格的过滤

2.1 基于多层次网格的过滤方法

在应用均匀网格法进行区域相交判断时,如果每个区域都完全覆盖与它相交的格子,该方法就是完全精确的。一般情况下,区域对相交格子的覆盖面越大,相交判断结果越精确。实际应用中,与区域中心地带相交的格子往往被完全覆盖,而与区域边缘地带相交的格子往往被部分覆盖。对同一精度要求,若仅考虑较大区域的中心地带,采用较大的格子尺寸就可满足;在考虑到较大区域的边缘地带及较小区域时,则需要较小的格子尺寸。针对这种情况,本文提出了以多层次网格取代平均网格的相关过滤方法。本文对多层次网格的说明与分析均在二维路径空间进行;对方法作适当、相应的修改后,也可用于一维或多维路径空间。

多层次网格包含三个具有不同格子尺寸的均匀网格,以满足不同大小区域以及区域内不同地带对格子尺寸的不同要求。每个均匀网格均覆盖了整个路径空间且三个网格间具有一定关系。在格子尺寸方面,第 i 层网格的格子尺寸是第 $i+1$ 层网格的四倍。在格子的空间位置方面,第 i 层网格每个格子的空间位置恰与第 $i+1$ 层网格中四个相邻格子的空间位置完全重合。在已知第 i 层网格时,将它的每个格子划分成大小均匀的四格子,就得到了第 $i+1$ 层网格。

定义1 在多层次网格中,对于第 i 层网格的每个格子,认为它是来自第 $i+1$ 层网格且空间位置与它完全重合的四个格子的父亲;将它与它所有后代构成的深度为2的树称为格子树。

定义2 一个区域占有一个格子当且仅当:

- i. 若是第1层网格的格子,该区域与该格子的所有孙子相交
- ii. 若是第2层网格的格子,该区域与该格子的所有儿子相交且没有占有该格子的父亲
- iii. 若是第3层网格的格子,该区域与该格子相交且没有占有该格子的父亲或祖父

多层次网格的相关过滤法用格子树代替了均匀网格中的格子。对每个区域,计算所有格子树中被

它占有的格子。一方面,若一区域占有一个格子,那么该区域也和这个格子的后代及祖先相交。另一方面,与一区域相交的格子只可能是它占有的格子,以及它占有格子的祖先与后代。利用多层次网格与占有概念来描述区域的空间位置,与具有相同精度的均匀网格比较,只需使用较少的格子。图2描述了三个区域以及这些区域占有的格子。在(a)中,区域占有了第1层网格中的一个格子。在(b)中,区域占有了第2层网格中的两个格子和第三层网格中的两个格子。在(c)中,区域占有了第1层网格中的一个格子,第2层网格中的八个格子和第3层网格中的四个格子。

每个格子对占有它的订购区域进行登记。在计算实体信息的发送对象时,首先计算它的发布区域占有的格子;每个被占有的格子,它的祖先及后代所登记的实体就是所求发送对象。用多层次网格描述区域位置并判断区域是否相交的精确度与利用格子尺寸与第3层网格相同的均匀网格是相同的;但多层次网格大大减少了订购区域的重复登记,降低了存储资源需求量,提高了查找相应订购区域的速度。

2.2 基于多层次网格的过滤在数据分发管理中的应用

HLA中的数据分发管理机制由RTI中的数据分发管理模块实现。RTI主要有两种实现方式,集中式与分布式,见图3。集中式RTI便于管理;但单个RTI由于处理能力有限,容易成为分布式交互仿真的瓶颈。基于多层次网格的相关过滤很容易在集中式RTI的数据分发管理模块中实现。分布式RTI有较强的处理能力,但需要考虑各个RTI之间的协调和联络,实现起来比集中式RTI困难。

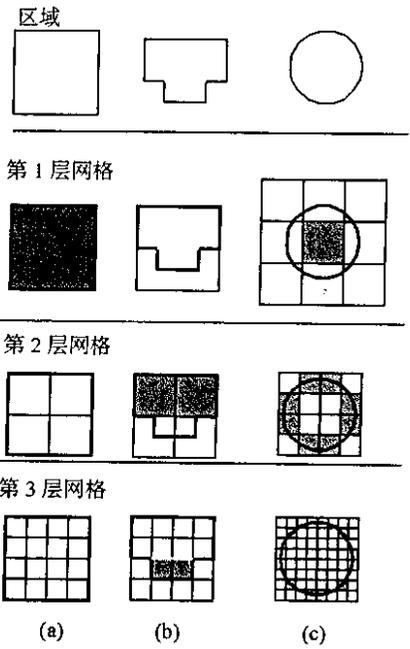


图2 区域及其占有的格子

Fig.2 Regions and the cells occupied by them

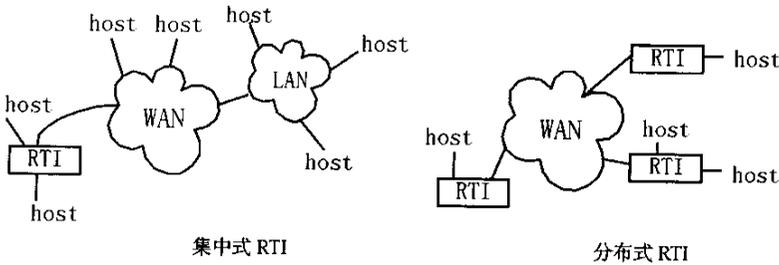


图3 集中式 RTI 与分布式 RTI

Fig.3 Centralized RTI and distributed RTI

基于多层次网格的相关过滤用于分布式 RTI 时,若每个 RTI 都记录所有区域与多层次网格的详细占有关系,势必浪费大量存储资源,且当区域改变时,将有大量数据在 RTIs 间传递。把每个 RTI 与它所管理的各个结点称为一个站点,对于外部站点实体与内部站点实体,RTI 采用不同的方式记录区域与多层次网格间的关系;三层网格的每个格子都对订购区域或发布区域占有它的内部实体进行登记,而只有第1层网格的格子登记订购区域与它相交的外部实体。若一外部区域与第1层网格的一个格子相交,则该区域可能与占有该格子或该格子后代的内部区域相交。若 RTI 发现一内部发布区域可能与一外部订购区域相交,则将该发布区域占有格子的情况发送到订购区域所在 RTI 以做进一步判断。

用于一个分布式 RTI 的基于多层次网格的相关过滤算法描述如下。

if 一个内部实体的发布区域 U 发生了改变

```

清空它的相应发送队列
计算 U 占有的格子
for 每个被占有的格子 C
    将订购区域占有 C, 或 C 的祖先, 或 C 的后代的内部实体加入发送队列
查找订购区域可能与 U 相交的外部实体
将 U 占有格子的情况发送到这些外部实体所在 RTIs
elseif 一个实体的订购区域 U 发生了改变
    从所有内部实体的发送队列中将该实体移出
    计算 U 占有的格子
    for 每个被占有的格子 C
        将实体加入到发布区域占有 C, 或 C 的祖先, 或 C 的后代的内部实体的发送队列中
    计算 U 与第一层网格的相交情况, 将结果发送给其它 RTIs
elseif 接收到外部实体 E 的发布区域 U 对多层次网格的占有情况
    for U 占有的每个格子 C
        将订购区域占有 C, 或 C 的祖先, 或 C 的后代的内部实体加入到 E 的发送队列中
elseif 接收到外部实体 E 的订购区域 U 与第一层网格的相交情况
    从每个内部实体发送队列中删去 E
    寻找可能与 U 相交的内部发布区域, 并将这些区域对多层次网格的占有情况发送到 E 所在 RTI

```

3 基于多层次网格相关过滤的性能分析及比较

基于网格的相关过滤不必直接计算两个区域是否相交，所以它在速度方面优于层次过滤法。下面对多层次网格法与均匀网格法在精确度和实体冗余登记方面进行比较。

利用网格法进行相交判断时，多层次网格与格子尺寸同它第 3 层网格相同的均匀网格具有相同的精确度。因为多层次网格包含三个格子尺寸各不相同的网格，且实体登记方式也与均匀网格不同，它登记的冗余实体比均匀网格少得多。例如，区域占有第 1 层网格一个格子的实体，多层网格仅在该格子中登记实体，而具有相同精度的均匀网格需要在十六个格子中登记实体。

为了比较两种网格法实体登记的冗余程度，模仿一个地域范围为 $350\text{km} \times 350\text{km}$ ，有 500 个仿真实体参加，基于集中式 RTI 的分布式交互仿真。每个实体有一个圆形发布区域和一个圆形订购区域。均匀网格的格子尺寸设为 5km ，与多层次网格中第 3 层网格的格子尺寸相同。以区域半径范围为参数，随机产生 1000 个区域的半径和位置。比较均匀网格与多层次网格登记的实体总次数（见表 1），可以发现，采用多层次网格可以大大减少登记次数，从而减少存储量，并使得查找发送对象的工作更简单。在需要较大精度而必须把均匀网格划分的足够细致时，使用多层次网格的优点会更明显。

表 1 均匀网格与多层次网格实体登记数的比较

Tab.1 Compare the times entities register in even - grid and multi - level grid

区域半径 (km)	均匀网格登记的实体次数	多层次网格登记的实体次数	登记次数减少的百分比 (%)
10 - 20	16018	10021	37.4
20 - 30	61012	43015	29.5
30 - 40	135037	78028	42.2
40 - 50	237127	129067	45.6
10 - 50	16056	10038	37.5
20 - 50	61402	43165	29.7
30 - 50	135071	84038	37.8

4 结论

基于均匀网格的相关过滤利用均匀网格近似表示区域的空间位置,从而快速判断区域间是否相交;但由于均匀网格的格子尺寸单一,无法同时恰当的表示区域中心与边缘,造成了存储资源浪费,也影响了相交判断的速度。针对这一问题,提出了基于多层次网格的相关过滤,用格子尺寸不同的多层次网格代替均匀网格,克服了均匀网格法的上述缺点。

随着路径空间维数的增加,相关过滤变得更复杂。路径空间维数越大,多层次网格中的每个父格子就拥有越多的子格子,存储量和相交判断计算量大增。所以,基于多层次网格的相关过滤比较适合低维路径空间,随着维数的增加,该方法复杂性的增长幅度较大。

参考文献:

- [1] Richard M Fujimoto, Richard M. Weatherly. Time Management in the DoD High Level Architecture [EB]. citeseer. nj. nec. com/fujimoto96time. html.
- [2] Paul T Braham, Randall E Barker, Joanne L Metzger, et al. HLA as a distributed virtual worlds architecture[EB]. www. rbd. com/publications. html.
- [3] Petty M D, Morse K L Computational complexity of HLA data distribution management [C]. Proceeding of the Fall 2000 Simulation Interoperability Workshop, Orlando FL, September 2000.
- [4] Ng Wee Ngee. A study of HLA use in non - military installations, and the grid - based approach to data distribution management in High - Level Architecture [EB]. www. comp. nus. edu. sg/~ ngweenge/urop/final-report. htm.
- [5] 史扬,金世尧,张晨曦. 分布交互仿真中基于 Server 的层次过滤机制 [J]. 软件学报, 10 (10).

