

文章编号:1001-9081(2006)07-1724-03

基于博弈的企业供应链管理模型与仿真

王全明,崔杜武

(西安理工大学 计算机科学与工程学院,陕西 西安 710048)

(wangqm0629@sohu.com)

摘要:设计并实现了一种企业供应链管理模型。描述了模型的设计,详细介绍了模型的实现算法,并对模型进行了仿真分析。仿真结果表明:本模型能够在企业外部需求不断变化的条件下,简便地解决其最优策略的确定问题,即此时企业可以获得最大赢利,同时也可使整个供应链达到全局最优。

关键词:博弈;供应链管理;仿真

中图分类号: TP391.9 **文献标识码:** A

Enterprise supply chain management model and simulation based on game

WANG Quan-ming, CUI Du-wu

(School of Computer Science and Engineering, Xi'an University of Technology, Xi'an Shaanxi 710048, China)

Abstract: An enterprise Supply Chain Management(SCM) model was designed and implemented. The design of the model and the implementation algorithm of models were described. The simulation result shows that the model can simply solve the problem of the optimum strategy in the condition of changing external demand, at this time the largest profit can be obtained, meanwhile, the state of whole SCM is optimum.

Key words: game; Supply Chain Management(SCM); simulation

0 引言

供应链管理越来越成为一种普遍的企业运营模式,其目标是以一种协同性的战略,整合整个企业资源来追求利润的最大化。供应链管理的设计与建模实现企业行为模拟,对于成功实施供应链管理至关重要。然而在目前对供应链管理模型的研究中,却存在以下不足:

- 1) 供应链模型缺乏足够的智能性,难以适应不断变化的供应链需求;
- 2) 目前提出的大多数供应链管理模型实际上就是一个管理软件,建立的模型难以模拟实际的企业行为。

在具有核心企业的企业组织中,一切商业活动都围绕着核心企业来进行^[1],例如原材料采购、生产、组装和销售都是在核心企业控制下完成,核心企业的行为将很大程度上影响整个供应链。

本文在企业行为策略不断变化的条件下,以不断提高赢利为目的,寻求最优博弈策略的模型,并在此基础之上,对构建的模型进行了仿真分析。分析结果表明,该模型有效解决了企业最优博弈策略的确定问题,同时可使整个供应链状态趋于稳定,达到全局最优。

1 企业供应链管理模型

1.1 模型描述

我们根据企业的行为策略,对供应链中的企业进行类型划分:

类型 1 企业第一次博弈采取欺骗(Defeat)策略,然后在随后的博弈过程中,采取“针锋相对”战略,即选择上一次博弈时,对手采取的行动。

类型 2 企业第一次博弈采取合作(Cooperate)策略,然后在随后的博弈过程中,采取“针锋相对”战略,即选择上一次博弈时,对手采取的行动。

类型 3 在所有的博弈过程中均采取中立(Neutral)策略,即不与任何企业进行博弈。

不同类型的企业随机分布在一个 N 个单元格组成的圆环面(R)上,每个单元格最多由一个企业所占据,空单元格意味着没有被企业所占据。 R 上企业的分布密度为 p ,设三种不同类型的企业在 R 上所占的比例为 p_1, p_2, p_3 ,则有 $p_1 + p_2 + p_3 = 1$ 。

在任何时刻,一个企业可以跟 R 上的任何一个单元格的其它企业进行博弈。由定义的企业类型所知,企业可采取的策略有合作(C)、欺骗(D)和中立(N)三种,在博弈过程中,采用一种奖励和惩罚机制,即根据采取的行为策略企业获得不同的赢利值。本模型中, $S(C) > S(N) > S(D)$,其中 $S(C), S(N), S(D)$ 是企业采取合作(C)、欺骗(D)和中立(N)行为时的赢利值,这些赢利值可由企业赢利矩阵来表示,如表 1 所示。

表 1 企业赢利矩阵表

策略	策略		
	Cooperate	Defeat	Neutral
Cooperate	{ b, b }	{ d, a }	{ e, e }
Defeat	{ a, d }	{ c, c }	{ e, e }
Neutral	{ e, e }	{ e, e }	{ e, e }

表 1 中,二元组 $\{x, x\}$ 表示采用不同策略的赢利,例如二元组 (a, d) ,我们根据它在表中的位置,可以知道它表示企业采用欺骗(D)策略的赢利值为 a ,企业采用合作(C)的赢利值为 d 。

根据经典博弈理论,只要满足以下条件:1) $a > b > c >$

收稿日期:2006-01-16;修订日期:2006-03-11 基金项目:陕西省自然科学基金资助项目(04JK263)

作者简介:王全明(1981-),男,四川仪陇人,硕士,主要研究方向:电子商务、计算机网络应用技术; 崔杜武(1945-),男,陕西西安人,教授,博士生导师,主要研究方向:Internet应用、多媒体技术。

$d; 2) 2b > a + d$, 则存在唯一的 Nash 均衡^[2-3]。在本模型中, 我们假定 $a = 7, b = 4, c = 3, d = 1, e = 5$ (e 为企业采取中立行为时的赢利)。

1.2 模型算法

参与博弈的企业在圆环面 (R) 上随机分布, 企业可以用五元组 $\{name, type, location, history, interest\}$ 表示。其中, $name$ 表示企业的名字, 在本模型中, 企业名字以企业编号代替, 即企业从 1 开始编号, 依次加 1, 一直编到最后一个企业 K, K 为企业总数; $type$ 表示企业的类型, $type \in \{1, 2, 3\}$; $location$ 表示企业在圆环面 R 上的位置, 即为 R 上单元格的位置编号, 单元格的位置从第一象限, 以 1 开始逆序编号, $location \in \{1, 2, \dots, N\}, N$ 为 R 上的单元格总数; $history$ 表示企业博弈的历史数据; $interest$ 表示企业的赢利值, 它是企业进行每次博弈的赢利值之和。

从 t 次到 $t + 1$ 次博弈, 企业的博弈算法可分为交往博弈、类型自学习和企业位置重新调整三个步骤:

1) 交往博弈。从编号为 1 的企业开始, 随机与圆环面 (R) 上的其他企业进行博弈, 博弈过程中, 根据双方的策略和赢利矩阵得到本次的博弈赢利, 将本次所得赢利与企业已有的赢利值累加, 同时更新企业的博弈历史数据, 直至编号 1 的企业与 R 上的其余所有企业博弈完成。

编号 2 的企业开始与圆环面 (R) 上的企业开始博弈, 随机选择 t 次彼此没有进行博弈的企业进行博弈。例如, 由于编号 2 的企业在 t 次已经与编号 1 的企业博弈过, 故编号 2 的企业不再与编号为 1 的企业进行博弈, 直至编号 2 的企业完成本次博弈。

然后是编号 3 的企业开始博弈, 直至圆环面 (R) 所有的企业都完成类似的博弈。

2) 类型自学习。在博弈过程中, 不同类型的企业的赢利值也不同, 赢利值低的企业会学习其他企业, 改变自身的类型, 以期增大自身的赢利值。在本模型中, 采用一种强化学习 (Reinforcement Learning, RL) 算法, RL 是求解随机的、序贯的 Markov 决策问题的有效方法^[4], 可以帮助企业在博弈过程中做出最好的决策。设第 t 次博弈结束时企业 i 获得博弈赢利值为 $N(i, t)$, 期望值为 $\hat{N}(i, t)$ 。将期望值 $\hat{N}(i, t)$ 对应于过去的博弈赢利, 并作为当前赢利的一个基准, 这样来估计它是进步了还是没有。如果 $N(i, t) > \hat{N}(i, t)$, 那么将维持其企业类型不变; 否则修改企业类型。企业类型自学习的详细步骤如下:

步骤 1 企业类型初始化

设 s_1, s_2, s_3 为企业选择类型 1、类型 2、类型 3 的概率, 则有 $s_1 + s_2 + s_3 = 1$ 。以三元组 $S_i = \{s_1, s_2, s_3\}$ 表示 i 类型企业的类型选择概率。当 $t = 0$ 时, 设 $S_1 = \{1, 0, 0\}, S_2 = \{0, 1, 0\}, S_3 = \{0, 0, 1\}$, 即初始类型为 i 的企业选择类型 i 的概率为 100%。

步骤 2 计算第 t 次博弈的期望赢利值 $\hat{N}(i, t)$ 。

计算类型为 i 的企业第 t 次博弈结束的平均赢利值 $N(i, t)$, 并计算 $\hat{N}(i, t)$ 的值。 $\hat{N}(i, t)$ 的计算公式如下:

$$\hat{N}(i, t) = \lambda \hat{N}(i, t-1) + (1 - \lambda) \times N(i, t-1)$$

其中 $0 < \lambda < 1$ 为一恒量, 它度量过去的平均赢利对当前赢利的影响。

当 $t = 1$ 时, $N(i, t-1) = 0$, 即企业在首次博弈前并没有赢利。

步骤 3 计算赢利变化率 $r(i, t)$ 的值:

$$r(i, t) = \frac{N(i, t) - \hat{N}(i, t)}{\hat{N}(i, t)}$$

步骤 4 计算企业类型修正值 $S_a(i, t)$:

$$S_a(i, t) = S_a(i, t) + a \times r(i, t)$$

其中 a 是一个决定学习速度的正恒量。此模型中取 $a = 0.1$ 。

步骤 5 类型修正值 $S_a(i, t)$ 正规化:

$$\sum_{a=\{1, 2, 3\}} S_a(i, t) = 1$$

得到正规化后的 $S_a(i, t)$ 值。

步骤 6 根据 $S_a(i, t)$ 选择企业类型。

如果满足 $S_a(i, t) = \text{Max}(S_1(i, t), S_2(i, t), S_3(i, t))$

则选择企业类型 $S_a(i, t)$ 对应的企业类型。

在 $t + 1$ 次博弈过程中, 企业 i 从步骤 2 重新开始博弈。

3) 企业位置重新调整。企业完成自学习后, 所有企业再一次随机分布在圆环面 (R) 上, 为下一次博弈做准备。

2 模型仿真

2.1 仿真环境

为了验证模型设计的正确性, 同时考虑到仿真环境的可重用性, 采用面向对象语言进行了实现, 开发环境如下:

开发工具 JBuilder 9, Oracle 9.0;

开发平台 Windows 2000 server + SP4。

仿真环境是 40 个不同类型的企业随机分布在由 100 个单元格组成的圆环面上, 企业进行 100 次博弈、自学习和位置调整。实验数据表明, 这样的分布密度和运行次数能较好地模拟企业进行博弈的行为。

2.2 仿真分析

为了对企业博弈行为进行量化, 提出了两个度量指标, 即企业类型个数和企业平均赢利值。

企业类型数 它是指在一次博弈过程中某种类型企业的个数, 该指标可以反映企业自学习的程度和企业行为策略对企业类型的影响;

企业赢利值 它是指某种类型的企业在博弈过程中的平均赢利值, 该指标可以反映某种类型的企业在博弈过程中的赢利大小。

我们设定进行仿真的 3 种类型的企业比例为 p_1, p_2, p_3 。为了能够反映不同的企业分布情形, 我们假定 4 种分布, 即分布比例为 $\{0.3, 0.3, 0.3\}, \{0.8, 0.1, 0.1\}, \{0.1, 0.8, 0.1\}, \{0.1, 0.1, 0.8\}$ 4 种情形。我们利用 Java 语言对企业类型个数变化和企业赢利变化进行了绘图。

从图 1 可以看出, 在企业类型分布比例存在相当差异的条件下, 经过多次博弈和自学习, 类型 2 的企业会变得越来越, 其次是类型 3 的企业, 类型 1 的企业最少。还可以看出 3 种类型企业个数在运行的初期存在一定的波动, 这是由于 RL 学习算法是基于一个试探性过程的算法^[5], 但经过一定次数的试探学习后, 企业行为选择会更加理性, 即企业类型会逐渐趋于一个稳定值。

从图 2 可以看出, 在企业类型分布比例存在相当差异的条件下, 随着博弈次数的增多, 企业赢利呈一个逐步增大的趋势, 最后趋于一个稳定值, 且类型 2 的企业赢利值最大, 其次是类型 3 的企业, 类型 1 的企业赢利值最少。

对两种曲线进行综合分析,可以得出以下结论:在企业进行多次博弈后,由于对企业行为策略予以适当的奖惩,企业的

合作行为得到了加强,而中立和欺骗行为得到了削弱,对某特定企业而言,也确定了其最优博弈策略和最大赢利值。

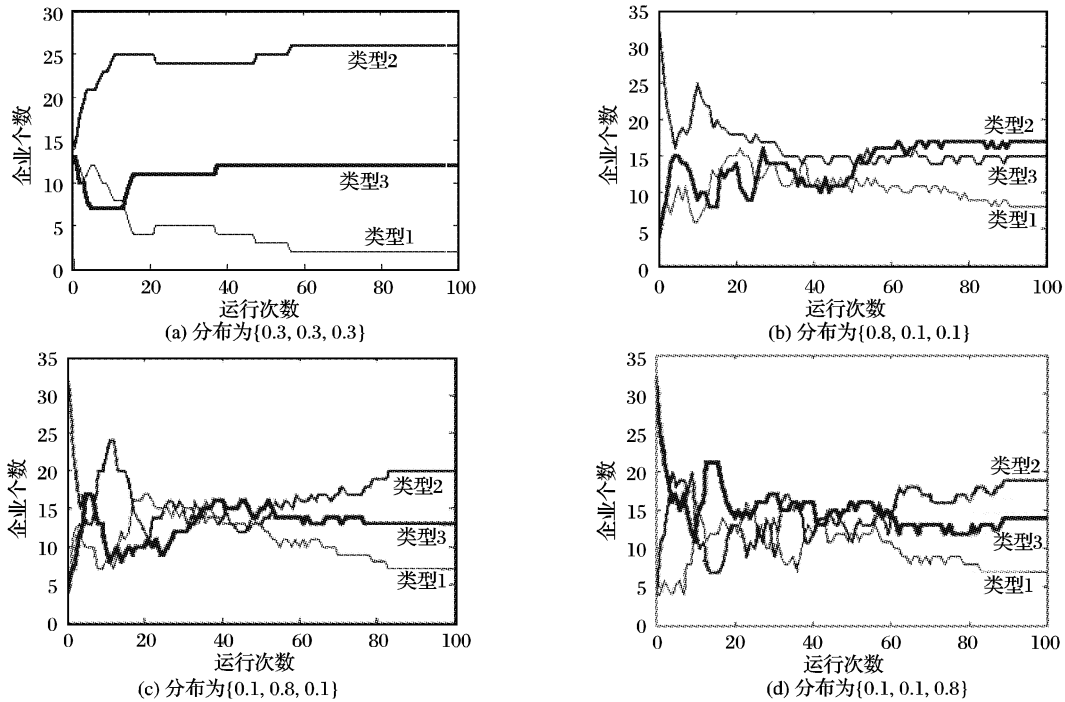


图1 企业类型变化曲线

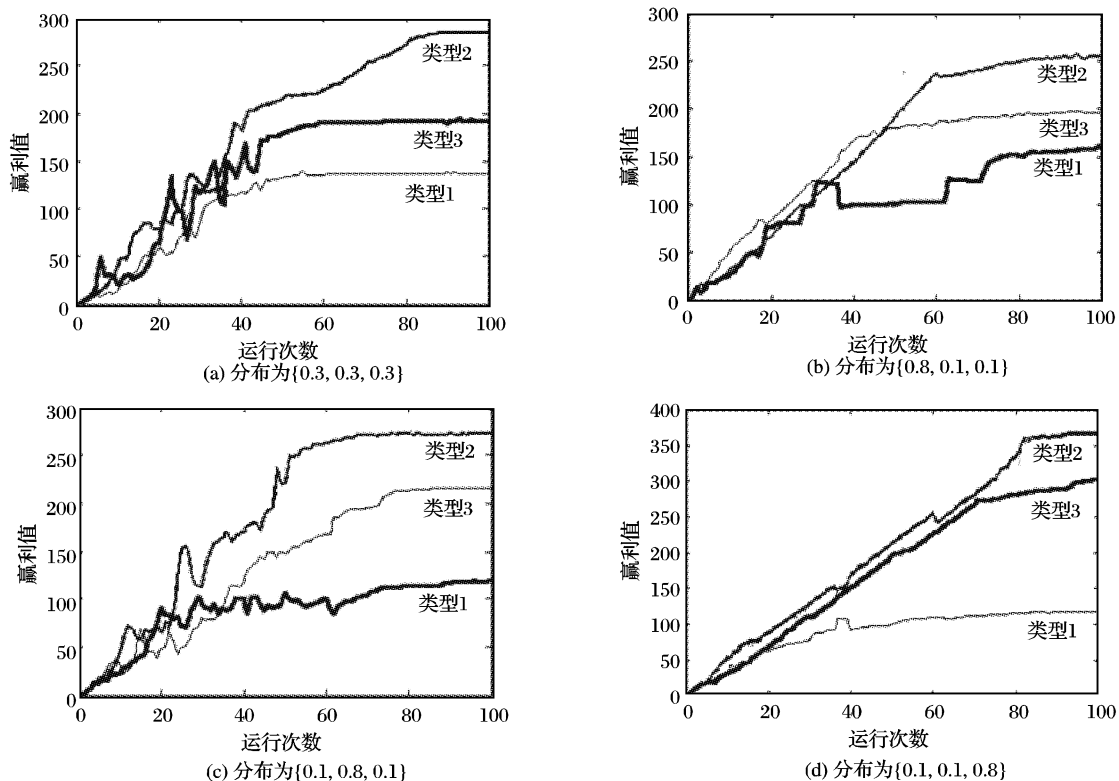


图2 企业赢利变化曲线

3 结语

本文设计了供应链管理中的企业合作行为模型,提出了企业博弈算法,对模型进行了模拟和仿真,并设计了反映企业博弈程度的指标。通过对指标进行分析可以看出,采取行为策略的奖惩机制可以有效地促进合作,减少欺骗和中性行为,同时也确定了每个企业的最优博弈策略。

参考文献:

- [1] SUTTON RS, BARTO AG. Reinforcement Learning: An Introduction[M]. MIT Press, 1998.
- [2] 侯定丕. 博弈论导论[M]. 合肥: 中国科学技术大学出版社, 2004.
- [3] 王醒策, 张汝波, 顾国昌. 多机器人动态编队的强化学习算法研究[J]. 计算机研究与发展, 2003, 40(10): 1444 - 1450.
- [4] 马继辉, 张仲义, 任远. 基于 Multi-Agent 的企业间供应链管理系统模型[J]. 北京交通大学学报, 2004, 28(10): 108 - 110.
- [5] FUDENBERG D, TRIOLE J. Game Theory[M]. MIT Press, 1991.