

基于智能 Agent 的供应商选择策略^{*}

韩伟¹ 王云² 王成道¹

(华东师范大学计算机系 上海200062)¹ (临沂师范学院计算机系 山东临沂276000)²

摘要 面向实际应用研究了电子市场中买方 Agent 的供应商选择问题,基于 Gittins 的方法,给出了全面考虑价格、质量和购买概率等因素的情况下计算每个供应商吸引指数的方法。对于新的供应商,可以用平均参数计算吸引指数。

关键词 电子市场,智能 Agent,购买概率,吸引指数

An Agent-Based Supplier Choosing Strategies

HAN Wei¹ WANG Yun² WANG Cheng-Dao¹

(Department of Computer, East China Normal University, Shanghai 200062)¹

(Department of Computer, Linyi Normal University, Linyi, Shandong 276000)²

Abstract The paper discusses the problem of suppliers choosing for buyer agent in electronic market. It puts forward a method for computing the attractive index of suppliers, which takes consideration of factors such as price, quality and buy probability. To improve correctness of agent's decision making, it proposes a credit-based mechanism.

Keywords Electronic marketplaces, Intelligent agent, Buy probability, Attractive index

1 引言

在电子市场上,购买某种特定商品的买方 Agent 需要从提供该商品的卖方 Agent 中选择最合适的供应商。各个供应商提供的商品在价格、质量上都有所不同。在购买之前,购买者只知道商品的价格,无法确切知道商品的质量,但购买者可以根据之前与该供应商的交易记录推测其提供的商品的质量。因此,对于比较了解的供应商,购买者总是可以根据与之交往的经验选择能提供“质优价廉”的商品的供应商。对于不了解的供应商,通过交易购买者可以学习到关于其商品质量的知识,这对于今后购买同类商品是有益的。因此,每个买方 Agent 都面临一个供应商选择问题,即如何给定一个评价标准,用来评价所有新老供应商。

Gittins 曾讨论了 Agent 在利用现有知识和学习新知识间如何做出选择的两难困境问题^[1]。Gittins 对每一个可能的选择都计算一个吸引指数,并且证明了吸引指数最高的选择为最佳选择。但是 Gittins 的方法应用于供应商选择问题需要根据实际情况作一定程度的扩展。本文面向实际应用,基于 Gittins 的方法,讨论供应商选择问题。

2 Gittins 吸引子

Gittins 的方法来源于多手柄老虎机问题:一个

带有多个手柄的老虎机,每个手柄被拉动时释放一定数量的金币,手柄 i 产生的金币数量服从已知分布 θ_i ,每次只能拉动一个手柄,要求设计一个策略,给定次数内使老虎机释放的金币数量最多。很明显,每次拉动 k 个手柄中的哪一个取决于当前决策者的经验和知识。

Gittins 的方法是为每一个选择(手柄) i 计算一个吸引指数 λ_i , λ_i 代表在当前知识下选择 i 的回报, Gittins 进一步证明选择吸引指数最高的手柄获得的期望收益最高。Gittins 将每一个手柄与假设的固定收益 λ 做比较,避免了各个手柄之间的两两比较。考虑收益的时效性,假定收益随时间以指数函数的形式递减,折扣因子为 a 。

定义1 若折扣因子为 a , \bar{x} 和 \hat{s} 为手柄 i 产生金币数量的均值和标准差。则在手柄 i 和固定收益 λ 间作出最佳选择时的期望收益函数为:

$$R(\lambda, \bar{x}, \hat{s}, n) = \max \left\{ \frac{\lambda}{1-a}, \bar{x} + a \int (R(\lambda, \chi(\bar{x}, x, n+1), \sigma(\hat{s}, \bar{x}, x, n+1), n+1) f(x|\bar{x}, \hat{s}, n) dx) \right\}$$

$n \geq 2$ 。其中 χ 表示第 $n+1$ 次的值对前 n 次均值 \bar{x} 的调整, $\chi(\bar{x}, x, n+1) = \frac{n * \bar{x} + x}{n+1}$; σ 表示第 $n+1$ 次的值对前 n 次标准差的调整, $\sigma(\hat{s}, \bar{x}, x, n+1) =$

$$\sqrt{\frac{(n-1)\hat{s}^2 + (x-\bar{x})^2}{n+1}}$$

^{*} 本文得到国家863计划项目(2002AA134020-04)资助。韩伟 博士研究生,主要研究方向:多 Agent 系统仿真与建模研究,电子商务。王云 讲师,研究方向:电子商务。王成道 教授,博导,主要研究方向:人工智能与模式识别。

定义1中 max 函数的右半部分表示选择手柄 i 后总的期望收益。若手柄 i 被选中,选择之前的期望收益为 \bar{x} ,实际金币数量为 x ,那么下次选择时需要根据更新了关于手柄 i 的均值 χ 和标准差 σ 重新计算。 χ 和 σ 由手柄产生的金币数量 x 决定, x 的分布函数为 $f(x|\bar{x},\hat{s},n)$ 。Max 函数的左半部分表示选择固定收益 λ 的总的期望收益,本次收益为 λ ,下次为 $a\lambda, \dots$,总的收益为 $\lambda+a\lambda+a^2\lambda+\dots=\frac{\lambda}{1-a}$ 。使 max 函数左右两部分相等的 λ 的值称作该手柄的吸引指数,Gittins 用 $v(\bar{x},\hat{s},n)$ 表示。他在文[1]中证明: $v(\bar{x},\hat{s},n)=\bar{x}+\hat{s}\cdot v(0,1,n)$,并给出了不同折扣率下 $v(0,1,n)$ 查询用表。显然,吸引指数随均值 \bar{x} 和标准差 \hat{s} 增大而增大, \hat{s} 实际上代表决策风险,说明风险越大越需要进行尝试以获取更多的关于该手柄的信息来消除其不确定性^[2]。

3 供应商选择问题

3.1 问题描述

假定市场上销售某种特定商品的供应商有 N 个,供应商 i 提供的商品的质量均值为 μ_i ,标准差 σ_i ,价格 $price_i$,买方 Agent 分阶段购买商品,在每一给定阶段 t 购买商品的概率为 p 。买方 Agent 并不知道 μ_i 和 σ_i ,只知道价格 $price_i$,但买方 Agent 和供应商 i 之间有 n_i 次交易记录,据此可以推算供应商 i 提供的商品的质量为 \bar{x}_i ,标准差为 \hat{s}_i 。若购买的商品的质量为 x_i ,价格为 $price_i$,买方 Agent 获得的效用为 $x_i - price_i$ 。因此,买方 Agent 从供应商 i 处购买商品的期望效用为 $\bar{x}_i - price_i$ 。买方 Agent 在每个阶段的收益折扣率为 δ ,也就是买方 Agent 在下次的收益为 μ 相当于本次收益为 $\delta\mu$,折扣率代表了对于未来收益的不确定或者利率等因素^[3]。

我们首先假定所有供应商提供商品的价格相等,区别只在于质量。应用 Gittins 的方法,可以得到每个供应商的吸引指数,然后根据吸引指数选择供应商。但实际问题却要复杂得多:首先购买行为并不一定发生在预定阶段,在不同阶段都有可能发生;其次,不同供应商提供商品的价格也不同;再者,对于不熟悉的供应商没有足够的交易记录用以学习。

3.2 引入购买概率

买方 Agent 并不能精确预测在哪一个阶段实施购买,只能给出在阶段 t 的购买概率。例如,若每周从某处买两本书,则每天从该处购书的概率为 $2/7$ 。引入购书概率后,我们重新考虑 Gittins 的方法^[4]。

引理1 若买方 Agent 在每个阶段购买概率为 p ,折扣率为 δ ,则选择固定收益 λ 的总收益 $\frac{p\lambda}{1-\delta}$ 。

引理2 若买方 Agent 在每个阶段购买概率为 p ,折扣率为 δ ,则从阶段一到最后一次购买,折扣率

的期望 $E(a)=p\delta+p(1-p)\delta^2+\dots+p(1-p)^i\delta+\dots=\frac{p\delta}{1-\delta+p\delta}$ 。

证明:最后一次购买发生在第一阶段的概率为 p ,折扣率为 δ ;最后一次购买发生在第二阶段的概率为 $p(1-p)$,折扣率为 δ^2 ;...

定理1 设 $R(\lambda,\bar{x},\hat{s},n)$ 为购买概率 p 下的期望收益函数, ω 为 $a=\frac{p\delta}{1-\delta+p\delta}$ 时的 Gittins 吸引指数,则购买概率 p 下吸引指数 $\lambda=\omega\cdot\frac{1-\delta+p\delta}{p}$ 。

证明:由定义1和引理1, $\frac{p\lambda}{1-\delta+p\delta}=\frac{\omega}{1-a}$,代入 $a=\frac{p\delta}{1-\delta+p\delta}$,得到 $\lambda=\omega\cdot\frac{1-\delta+p\delta}{p}$ 。

3.3 商品价格不同的情形

假定各个供应商所提供的商品的价格有所不同, $R(\lambda,\bar{x}_i,\hat{s}_i,n_i,price_i)$ 表示在购买价格 $price_i$ 下的期望收益函数, $v(\bar{x}_i,\hat{s}_i,n_i,price_i)$ 表示供应商 i 的吸引指数。

定理2 若商品价格 $price_i$ 为常数,有:(1) $R(\lambda,\bar{x}_i,\hat{s}_i,n_i,price_i)=R(\lambda,\bar{x}_i-price_i,\hat{s}_i,n_i)$;(2) $v(\bar{x}_i,\hat{s}_i,n_i,price_i)=\bar{x}_i-price_i+\hat{s}_i\cdot v(0,1,n_i)$ 。

证明:考虑价格因素后,选择供应商 i 的收益均值为 $x_i - price_i$,标准差保持不变,(1)式成立。由 Gittins 的结论,(2)式成立。

3.4 引入信用学习机制

实际生活中我们选择供应商时,除了要求最大化期望效益,也希望对方是诚实可信的。为此,我们引入一个信用学习机制,每个买方 Agent 基于强化学习维护和动态更新由若干卖方 Agent 组成的可信集合。

定义2 买方 Agent b 的可信集合定义为: $S_b^t=\{i|v(\bar{x}_i,\hat{s}_i,n_i,price_i)\geq\theta_1 \text{ 且 } r^b(i)\geq\theta_2\}$ 其中 θ_1, θ_2 为阈值, $r^b(i):S\rightarrow(-1,1)$ 为 b 对 r 的信任函数, S 为卖方 Agent 组成的集合。 $r^b(i)$ 初始化为 $v(\bar{x}_i,\hat{s}_i,n_i,price_i)$ 的增函数。

根据可信集合, b 选择供应商 $\hat{i}=\operatorname{argmax}_{i\in S_b^t} R(\lambda,\bar{x}_i,\hat{s}_i,n_i,price_i)$ 。每次与 \hat{i} 交易完成后,实际效益 $R(x_i,price_i)=x_i - price_i$ 。可知,从而可以利用 $R_r(x_i,price_i)$ 信任函数。记 $\Delta=R_r(x_i,price_i)-R(\lambda,\bar{x}_i,\hat{s}_i,n_i,price_i)$,若 $\Delta\geq 0$,说明 \hat{i} 比预料的要好,应该对其增加授信程度。现实生活中,对一个人的信任总是逐步增加的,而对一个人产生不信任感却很容易,若信任增长率为 u ,信任减少率为 v ,则有 $|u|<|v|$ 。信用函数更新如下:

若 $\Delta\geq 0, r^b(i)=\begin{cases} r^b(i)+u[1-r^b(i)] & r^b(i)\geq 0 \\ r^b(i)+u[1+r^b(i)] & r^b(i)< 0 \end{cases}$
若 $\Delta< 0, r^b(i)=$

$$\begin{cases} r^b(i) + v[1 - r^b(i)] & r^b(i) \geq 0 \\ r^b(i) + v[1 + r^b(i)] & r^b(i) < 0 \end{cases}$$

4 方法及算例

4.1 方法

根据前面的讨论, 供应商选择问题首要的是计算吸引指数 $v(\bar{x}_i, \hat{s}_i, n_i, price_i)$, 然后根据吸引指数选择供应商。大体步骤如下:

1. 计算 a. 给定 p 和 δ , $a = \frac{p\delta}{1 - \delta + p\delta}$.
2. 根据 a, 查 Gittins 给出的吸引指数表, 求得
亚马逊
新风雨
中国书城

到货时间	供货质量
10天	33.9
7天	52.1

到货时间	供货质量
1天	91.1
1天	91.1
4天	68.9
2天	83.0

到货时间	供货质量
2天	83.0
2天	83.0
1天	91.1
5天	62.7
4天	68.9

三处书店价格、质量比较

	价格	质量均值	质量标准差
亚马逊	40	43.00	12.87
新风雨	75	83.53	18.13
中国书城	70	77.74	22.85

若仅从价格和质量考虑, 消费者应选择中国书城。下面我们会看到, 若全面考虑购书概率、标准差等因素, 最佳选择应为亚马逊书店。

$$\begin{aligned} 1. \text{ 计算 } a. a &= \frac{p\delta}{1 - \delta + p\delta} \\ &= \frac{0.1 \times 0.911}{1 - 0.911 + 0.1 \times 0.911} = 0.5. \end{aligned}$$

2. 查表。

n	w
2	0.726587142
4	0.094863017
5	0.070580465

$$\begin{aligned} 3. \lambda &= \omega \cdot \frac{1 - \delta + p\delta}{p} = \omega \cdot \frac{1 - 0.911 + 0.1 \times 0.911}{0.1} \\ &= \omega \cdot 1.819 \end{aligned}$$

$$4. v = \bar{x}_i - price_i + \hat{s}_i \cdot \lambda.$$

	$price_i$	\bar{x}_i	\hat{s}_i	λ	v_i
亚马逊	40	43.0	12.87	1.321	20.00
新风雨	75	83.53	18.13	0.173	11.69
中国书城	70	77.74	22.85	0.128	10.66

$$\omega = (0, 1, n).$$

$$3. \text{ 计算实际吸引指数 } \lambda. \text{ 给定 } p \text{ 和 } \delta, \lambda = \omega \cdot \frac{1 - \delta + p\delta}{p}.$$

$$4. \text{ 计算 } v. v = \bar{x}_i - price_i + \hat{s}_i \cdot \lambda.$$

4.2 算例

某消费者欲通过网上书店购买一本价值100元的专业书, 假设他(她)有三个供应商可供选择, 并且希望尽快获得图书。此书的价值折扣率 $\delta = 0.911$, 意即 t 天后书的价值为 $100 \cdot 0.911^t$ 。消费者从三处书店获得的购书历史如下表。

结束语 在电子市场中, 每个供应商提供商品的价格、质量都有所差异, 买方 Agent 对于商品需求的迫切程度、对各个供应商的了解程度(交易记录)也各不相同。对于熟悉的供应商, 买方 Agent 可以从交易记录中推测其产品质量^[5], 基于 Gittins 的方法, 本文给出了全面考虑价格、质量和购买概率等因素的情况下计算每个供应商吸引指数的方法, 并引入了基于强化学习的信用机制。进一步的研究应根据实际应用考虑更多的评价因素, 如购买数量、运输费用等。

参考文献

- 1 Gittins J C. Multi-armed Bandit Allocation Indices. John Wiley & Sons, 1989
- 2 Eric C, Hann K, Hitt I. The nature of competition in electronic markets. WP, The Wharton School of the Univ. of Pennsylvania, 1998
- 3 Salganicoff M, Ungar L H. Active exploration and learning in real-valued spaces using multi-armed bandit allocation indices. In: Proc. of the 12th Intl. Conf. on Machine Learning, San Francisco, CA, 1995. 480~487
- 4 Azoulay-Schwartz R, Kraus S. Exploitation vs. Exploration: Choosing a Supplier in an Environment of Incomplete Information. International Journal on Decision Support Systems and Electronic Commerce, 2003
- 5 Tran T, Cohen R. A learning Algorithm for Buying and Selling Agents in Electronic Marketplaces. Advances in Artificial Intelligence, Canada, 2002