

基于机器学习的自动协商决策模型^{*}

程 昱¹⁺, 高 济¹, 古华茂², 傅朝阳¹

¹(浙江大学 计算机科学与技术学院, 浙江 杭州 310027)

²(浙江工商大学 信息学院, 浙江 杭州 310035)

Automated Negotiation Decision Model Based on Machine Learning

CHENG Yu¹⁺, GAO Ji¹, GU Hua-Mao², FU Zhao-Yang¹

¹(School of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China)

²(School of Information, Zhejiang Gongshang University, Hangzhou 310027, China)

+ Corresponding author: E-mail: cyskyvein@126.com

Cheng Y, Gao J, Gu HM, Fu ZY. Automated negotiation decision model based on machine learning. *Journal of Software*, 2009,20(8):2160-2169. <http://www.jos.org.cn/1000-9825/3362.htm>

Abstract: The proposed model labels the negotiation history data automatically by making full use of the implicit information in negotiation history. Then, the labeled data become the training samples of least-squares support vector machine that outputs the estimation of opponent's utility function. After that, the self's utility function and the estimation of opponent's utility function constitute a constraint optimization problem that will be further figured out by genetic algorithm. The optimal solution is the counter-offer of oneself. Experimental results show that the proposed model is effective and efficient in environments where information is private and the prior knowledge is not available.

Key words: e-commerce; automated negotiation; negotiation decision model; support vector regression; genetic algorithm

摘 要: 所提出的模型利用协商历史中隐含的信息自动对数据进行标注以形成训练样本,用最小二乘支持向量回归机器学习此样本得到对手效用函数的估计,然后结合自己和对手的效用函数构成一个约束优化问题,用遗传算法求解此优化问题,得到的最优解就是己方的反建议.实验结果表明,在信息保密和没有先验知识的条件下,此模型仍然表现出较高的效率和效用.

关键词: 电子商务;自动协商;协商决策模型;支持向量回归;遗传算法

中图法分类号: TP18 文献标识码: A

随着多 Agent 技术的快速发展,电子商务领域中 Agent 技术的应用日趋广泛.Agent 能够代替人类用户更加高效地处理复杂的商务活动.在电子商务活动中,自动协商是最为核心的部分,因此它成为基于 Agent 的电子商务研究的重点内容.自动协商的研究主要包括 3 个方面^[1]:① 协商协议^[2]:协商的参与者在交互中必须遵守的规

* Supported by the National Natural Science Foundation of China under Grant Nos.60775029, 60773177 (国家自然科学基金); the National Basic Research Program of China under Grant No.2003CB317005 (国家重点基础研究发展计划(973))

Received 2007-10-22; Revised 2008-02-01; Accepted 2008-04-15

则集.② 协商内容^[3]:所需协商议题的组成和各议题的取值区间等.③ 协商策略^[4-6]:在特定的协商协议下,协商 Agent 如何根据目前的状态信息来选择的自己所要执行的协商动作.本文主要侧重于双边多议题协议^[7,8]下协商策略的研究.

协商策略的研究已经取得了很多有价值的研究成果.特别是近年来把机器学习的方法引入到协商决策模型中成为这个领域一个新的研究热点.Bayesian learning^[9]可能是在协商决策模型中使用最广泛的学习方法.在这种方法中,Agent 首先需要建立一个对手的偏好或是保留值的假设集(set of hypotheses),并设定这些假设的初始概率;在协商过程中根据对手的建议,利用 Bayesian 规则不断地更新这些假设的概率;最后利用学习到的这些假设的概率分布作出协商决策.进化计算(evolutionary computing)方法把协商策略编码成基因^[8],这些基因在协商过程中不断地通过变异、交叉、选择等操作来进化,以得出最优的协商策略.核密度估计(kernel density estimation,简称 KDE)方法^[10]首先估计对手每个协商项的权重,然后根据相似度函数和对手的协商项权重计算出与对手建议最相似的反建议.国内在协商策略的研究上也取得了不少成果.比如,王黎明等人把销售 Agent 的选择问题转变成 K 臂赌博机问题(K-armed bandit problem)^[11],提出了信任和声誉的度量模型,结合 K 臂赌博机问题的求解技术,采用学习机制,提出了几个确定奖励分布的改进算法,并以模拟协商过程为基础,将改进算法、信任和声誉有机地结合起来选择销售 Agent.高坚等人给出了一个 Agent 间多议题的协商模型(MN)^[12],并在此基础上提出了双边多议题协商的一种加速混沌进化算法(ACEA).ACEA 算法将混沌机制引入进化计算,并采用了压缩技术对算法进行加速.

上述协商决策模型普遍存在一个缺陷:需要应用领域的先验知识.核密度估计方法^[10]需要为每个协商项设计相似度函数,比如颜色相似度函数、形状相似度函数等,这需要色彩和形状领域的相关知识.Bayesian 方法^[9]和进化计算的方法^[8,12]要求事先定义出一个假设集和先验概率的信息.而这些假设与应用领域先验知识是密切相关的.如文献[9]中假设了供货者的保留价格为\$117 或是\$100,这需要特定货品定价的相关知识.Raymond 等人改进了传统的 Bayesian 方法,提出利用数据挖掘的方面来获得相关的先验知识^[13].但此方法只能处理协商项的取值是离散的情况.而且此方法需要对协商空间所有建议进行排序,当协商项增加时,协商空间将会成指数级增长,协商的计算时间也成指数级增长.本文实验部分与文献[13]中的方法从效用与计算时间上进行了比较.

本文提出了一个基于学习的双边多议题的协商决策模型(learning decision model,简称 LDM).LDM 通过协商中隐含的信息自动对协商历史数据标注对手的近似效用值,被标注的历史数据成为支持向量回归(support vector regression,简称 SVR)算法的训练样本,SVR 的学习结果为对手效用函数的估计.综合己方的效用函数和对手效用函数的估计构建一个约束优化问题.最后利用遗传算法求解此优化问题,得到的最优解就是己方的反建议.LDM 的优点主要有以下 3 个方面:① 无须应用领域先验知识,具有较强的通用性,能够应用于各种不同的协商任务和环境.② 无须预先获得对手的私有信息.在电子商务环境中,协商双方大多数情况下都不会透漏自己的私有信息,因此这一点在电子商务的环境中非常重要.③ 能够得到较优的协商性能,主要表现在协商双方的联合效用有较大的提高,协商回合数有明显的减少.

1 协商框架

1.1 框架组成

本文所基于的协商框架/framework of negotiation,简称 FN)定义为如下的一个 7 元组:

$$FN = \langle A, T, S, D, M, U, P \rangle,$$

其中, A 为参与协商的 Agent 集合,本文讨论的是双边协商,定义 $A = \{Agent_1, Agent_2\}$. 以其中某个 Agent 的角度来看,另一个 Agent 是它的对手,因此 A 也可以定义为 $\{Player, Opponent\}$. T 为协商的时间期限,可以表示为具体的时间或协商回合数. S 为协商项(协商议题)的集合, $S = \{s_1, s_2, \dots, s_n\}$. 例如在购买显示器的协商中,协商项可能是显示器的{价格,数量,品牌,保质期,分辨率,尺寸}. $D = \{D_1, D_2, \dots, D_n\}$, D_i 为协商项 s_i 有效的取值范围. M 协商双方互发送的建议(proposal)分为两类:一类为 offer,另一类为接受建议(accept)和拒绝建议(reject). offer o 定义为

$o = (b_1, b_2, \dots, b_n)$, 其中 b_i 为协商项 s_i 的具体取值. 如 s_i 为显示器的尺寸, 若 b_i 取 17 则表示 17 寸的显示器. 接受建议(accept)表示接受对手最后发来的 offer, 协商成功. 拒绝建议(reject)表示拒绝与对手继续协商, 协商失败. U 为协商 Agent 的效用函数, $U(\bullet): O \rightarrow R$, 其中 O 为所有 offer 的集合, R 为实数集. $U_i(o)$ 为 offer o 对于 Agent_{*i*} 的效用值, $U_i(\bullet) \in [u_i^{\min}, u_i^{\max}]$, u_i^{\min}, u_i^{\max} 为 $U_i(\bullet)$ 取值区间的最小值与最大值. P 为协商协议. 本文考虑的协议为交替提议协议(alternating offers protocol, 简称 AOP)^[14]. 根据 AOP 协商过程, 按图 1 所示的步骤进行.

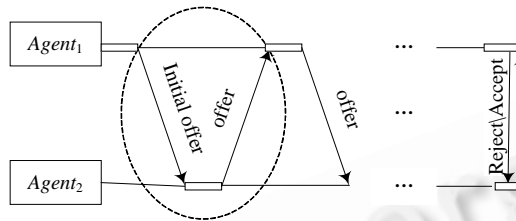


Fig.1 Negotiation process

图 1 协商过程

1.2 基本定义

定义 1(协商回合). Agent 接到对方的建议然后发送建议给对方, 称为一个协商回合, 简称回合, 记为 Round. 例如, $(o^{1 \rightarrow 2}, o^{2 \rightarrow 1})$ 是回合, 其中 $o^{i \rightarrow j}$ 表示 Agent_{*i*} 发给 Agent_{*j*} 的 offer. 图 1 中虚线圈标出了协商的第一个回合.

定义 2(协商过程). 协商中所有回合按时间顺序排列, 表示为

$$(o_1^{1 \rightarrow 2}, o_1^{2 \rightarrow 1}; o_2^{1 \rightarrow 2}, o_2^{2 \rightarrow 1}; \dots, Accept_k^{i \rightarrow j} | Reject_k^{i \rightarrow j}) \quad (i \neq j, \text{且 } i, j \in \{1, 2\}).$$

其中, $o_h^{i \rightarrow j}$ 表示第 $h(1 \leq h \leq k)$ 回合 Agent_{*i*} 发给 Agent_{*j*} 的 offer, $Accept_k^{i \rightarrow j} | Reject_k^{i \rightarrow j}$ 为最后回合 Agent_{*i*} 发给 Agent_{*j*} 的接受或拒绝的消息.

定义 3(可接受区间). u_i^{accept} 为可被 Agent_{*i*} 接受的最小效用值, 即 offer o 能够被 Agent_{*i*} 接受 $\Leftrightarrow U_i(o) \geq u_i^{accept}$. $[u_i^{accept}, u_i^{\max}]$ 称为 Agent_{*i*} 的可接受区间, 相应地, $[u_i^{\min}, u_i^{accept}]$ 称为 Agent_{*i*} 的不可接受区间.

定义 4(可接受域). $\{o | o \in D_1 \times D_2 \times \dots \times D_n, u_i^{accept} \leq U_i(o) \leq u_i^{\max}\}$ 为 Agent_{*i*} 的可接受域, 记为 T_i . offer o 能够被 Agent_{*i*} 接受 $\Leftrightarrow o \in T_i$

定义 5(协商可行域). Agent_{*1*} 与 Agent_{*2*} 的可接受域的交集 $T_1 \cap T_2$ 称为协商可行域. 显然, 最后达成的合同肯定落在可行域中. 当可行域为空时, 协商双方不可能达成合同. 图 2 显示了可接受域(acceptable region)与可行域(feasible region)的关系.

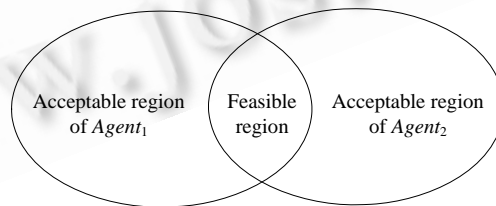


Fig.2 Acceptable and feasible region

图2 可接受域与可行域

定义 6(可行域比重). 可行域比重为可行域体积占 Agent_{*1*} 与 Agent_{*2*} 的可接受域体积的百分比, 记为 E . 即

$$E = \frac{H(T_1 \cap T_2)}{H(T_1 \cup T_2)}, \text{ 其中 } H(\bullet) \text{ 为体积算子.}$$

2 协商决策模型

2.1 模型总体结构

协商决策模型 LDM 由 6 个部分组成:消息接收器(receiver)、协商历史数据库(negotiation history database)、效用值估计模块(estimate utility)、支持向量回归(SVR)、遗传算法(genetic algorithm,简称 GA)、评估机制(evaluate).图 3 展示了这 6 个部分和及其连接方法.

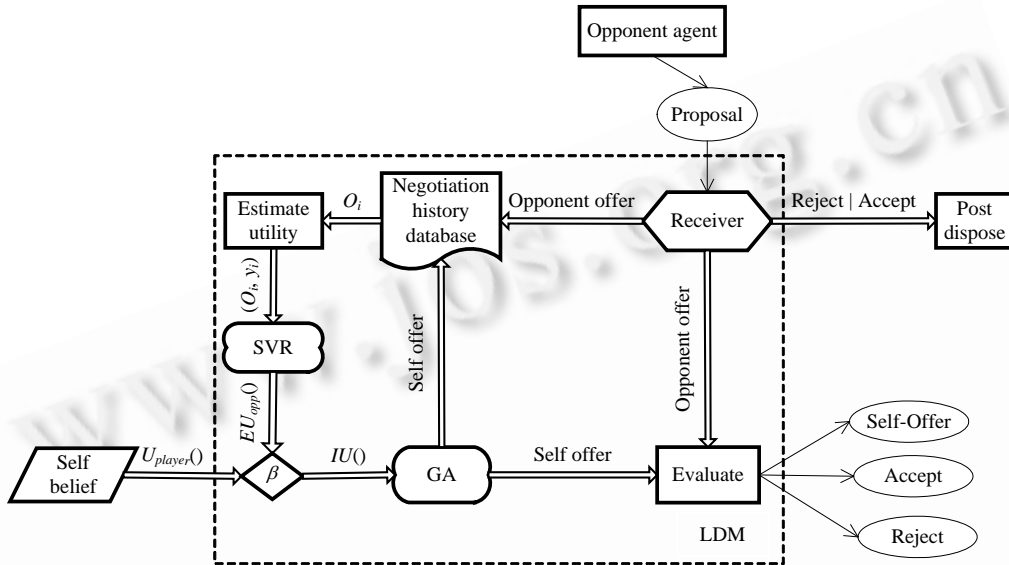


Fig.3 Negotiation decision model

图 3 协商决策模型

消息接收器(receiver):接收对手发来的协商建议(proposal),根据对手的建议类型分成两种情况进行处理。
 ① 当收到的建议是 offer 时,把它复制成两份,一份存入协商历史数据库,另一份交给评估机制模块。
 ② 当收到的建议是 Accept 或 Reject 时,则交给协商后续处理模块(post dispose)来进行协商结束的后续处理.因为后续处理不属于本文讨论的范围,这里不作介绍.

协商历史数据库(negotiation history database):存储协商的历史信息,即以前与对手协商的信息.本文中所用到的历史数据主要为建议序列(对手的建议和己方的反建议),包括协商双方发送的 offer、建议的发送方、建议所属的协商实例 ID 和建议所在的回合序号,具体数据库表结构见表 1.

Table 1 Negotiation history data

表 1 协商历史数据

Negotiation instance ID	Round	Sender	Did the receiver accept the offer	Offer
00031	1	Opponent	yes	$b_{11} b_{12} \dots b_{1n}$
00031	2	Self	No	$b_{21} b_{22} \dots b_{2n}$
...

估计对手效用值(estimate utility):这个模块的作用是生成支持向量回归机(SVR)的训练样本.本文的训练样本由 offer o_i 及其对应的对手效用值估计 y_i 构成.通常训练样本都是由专家对原始数据进行标注而生成.而本文训练样本的标注是由效用估计模块利用协商历史数据中的隐含信息自动产生.具体如何估计对手效用值将在第 3.2 节中详细描述.

支持向量回归机(SVR):主要任务是学习由效用值估计模块产生的训练样本,并由此得到对手效用函数的估计(estimation of $U_{opp}(\bullet)$,记为 $EU_{opp}(\bullet)$).考虑到学习算法的时间效率,本文采用最小二乘支持向量回归机,具

体的学习过程在第 3.3 节中描述。

遗传算法(GA):通过支持向量机对样本的学习得到对手效用函数的估计 $EU_{opp}(\bullet)$,综合考虑 $EU_{opp}(\bullet)$ 与己方的效用函数 $U_{player}(\bullet)$ 在协商中的重要性,组合成一个综合效用函数(integrative utility function,简记为 $IU(\bullet)$):

$$IU(o) = \beta \times U_{player}(o) + (1 - \beta) \times EU_{opp}(o), \beta \in [0,1] \tag{1}$$

β 为己方与对手效用的平衡因子. β 越大,表示越重视己方的效用,反之则表示越重视对手的效用.可以在协商过程中不断改变 β 的值来调整这种平衡关系.例如,初始时 β 可以设置得比较大,在协商过程中逐渐降低 β 的值.模型计算出的反建议应在满足协商约束的条件下使 $IU(\bullet)$ 值最大,即可得到如下的优化问题:

$$\begin{cases} \max_o IU(o) = \beta \times U_{player}(o) + (1 - \beta) \times EU_{opp}(o) \\ \text{s.t. } U_{player}(o) \geq u_{player}^{accept} \end{cases} \tag{2}$$

用遗传算法来求解此优化问题,所求得的近似最优解就是综合考虑了己方效用和对手效用的具体反建议值.具体的求解过程在第 3.4 节中描述。

评估机制(evaluate):设 o^{opp} 为对手建议, o^{self} 为此回合己方反建议.① 当 $U_{player}(o^{opp}) \geq U_{player}(o^{self})$ 时分两种情况来处理:a) $U_{player}(o^{opp}) \geq u_{player}^{accept}$ 表示双方的建议找到了交点, o^{opp} 可以被接受;b) $U_{player}(o^{opp}) < u_{player}^{accept}$,此时必有 $U_{player}(o^{self}) < u_{player}^{accept}$,无法与对手达成一致,拒绝与对手继续协商.② 当 $U_{player}(o^{opp}) < U_{player}(o^{self})$ 时也分两种情况来处理:a) 当 $U_{player}(o^{self}) \geq u_{player}^{accept}$ 时,表示仍然有一定的让步空间,继续与对手协商,发出 o^{self} ;b) 当 $U_{player}(o^{self}) < u_{player}^{accept}$ 时,表示超过了让步的极限,不可能与对手找到互相接受的合同,拒绝与对手继续协商.以上的 4 种情况可以形式化为以下 4 条规则:

$$\begin{aligned} (U_{player}(o^{opp}) \geq U_{player}(o^{self})) \wedge (U_{player}(o^{opp}) \geq u_{player}^{accept}) &\rightarrow \text{Accept}, \\ (U_{player}(o^{opp}) \geq U_{player}(o^{self})) \wedge (U_{player}(o^{opp}) < u_{player}^{accept}) &\rightarrow \text{Reject}, \\ (U_{player}(o^{opp}) < U_{player}(o^{self})) \wedge (U_{player}(o^{self}) \geq u_{player}^{accept}) &\rightarrow o^{self}, \\ (U_{player}(o^{opp}) < U_{player}(o^{self})) \wedge (U_{player}(o^{self}) < u_{player}^{accept}) &\rightarrow \text{Reject}. \end{aligned}$$

2.2 估计对手效用值

在电子商务环境中,Agent 的效用值不是公开信息,无法直接获得.但是,经分析发现,协商过程在一定程度上隐含了对手的效用值信息.① 对手发出的 offer 都应位于它的可接受域之内,且协商过程中 Agent 通常会逐渐让步,因此对手发出的 offer 效用值是逐步递减的.② 己方发出的 offer 如果被对手接受,则它也应落在对手的可接受域;若不被对手接受则 offer 应落在对手的不可接受域。

根据上述分析把 offer 被分成两类:① 落在对手可接受域之内的 offer,包括对手发出的 offer 和己方发出被对手接受的 offer,这一类 offer 的对手效用值会随着回合数的增加而逐渐减少.② 落在对手可接受域之外的 offer,这一类 offer 是己方发出且未被对手接受的 offer.这一类 offer 的对手效用值落在对手的不可接受区间中,文中取对手不可接受区间的中点,即 $(u_{opp}^{accept} + u_{opp}^{min})/2$.根据上面所分析的隐含信息, o_i 对于对手的效用值估计 y_i 可以通过公式(3)计算得出。

$$y_i = \begin{cases} u_{opp}^{max} - \frac{R(o_i) - 1}{WR(o_i)} \times (u_{opp}^{max} - u_{opp}^{accept}), & o_i \text{ is accepted by opponent} \\ \frac{u_{opp}^{accept} + u_{opp}^{min}}{2}, & o_i \text{ is not accepted by opponent} \end{cases} \tag{3}$$

$R(o_i)$ 表示 o_i 所处的协商回合, $WR(o_i)$ 为包含 o_i 的协商实例的回合总数,这两个是已知信息.但 $u_{opp}^{min}, u_{opp}^{max}, u_{opp}^{accept}$ 是对手的私有信息,对于己方来说是未知的,因此模型中无法直接利用公式(3)计算出 y_i .但可以对对手的效用函数 $U_{opp}(\bullet)$ 进行线性变换,使对手的可接受区间 $[u_{opp}^{accept}, u_{opp}^{max}]$ 和不可接受区间 $[u_{opp}^{accept}, u_{opp}^{max}]$ 映射到己方的可接受区间 $[u_{player}^{accept}, u_{player}^{max}]$ 和不可接受区间 $[u_{player}^{accept}, u_{player}^{max}]$.因此,可以设对手的可接受区间与不可接受区间与己方相同.由

此,式(3)中的 $u_{opp}^{\min}, u_{opp}^{\max}, u_{opp}^{accept}$ 可以用 $u_{player}^{\min}, u_{player}^{\max}, u_{player}^{accept}$ 代替,得到式(4).

$$y_i = \begin{cases} u_{player}^{\max} - \frac{R(o_i) - 1}{WR(o_i)} \times (u_{player}^{\max} - u_{player}^{accept}), & o_i \text{ is accepted by opponent} \\ \frac{u_{player}^{accept} + u_{player}^{\min}}{2}, & o_i \text{ is not accepted by opponent} \end{cases} \quad (4)$$

效用值估计模块通过式(4)算出 offer o_i 对于对手效用值的估计 y_i 后,即完成了原始数据的自动标注过程,形成了 SVR 的训练样本.

2.3 估计对手效用函数

通过第 3.2 节中的方法得到了训练样本 $\{(o_i, y_i), i=1, 2, \dots, l\}$, 利用高维特征线性空间函数:

$$f(o) = W \cdot \phi(o) + b \quad (5)$$

来拟合训练样本.其中, $\phi(\cdot)$ 将训练数据从输入空间映射到高维特征空间,由此将输入空间的非线性拟合问题转化为高维特征空间的线性拟合问题.考虑到协商问题的实时性要求,本文模型选用了训练速度较快的最小二乘支持向量回归机(least-squares support vector regression, 简称 LSSVR)^[15].LSSVR 的具体改进是用等式约束代替了标准支持向量机中的不等式约束,经验风险也由偏差的二次方变成了一次方.根据 LSSVR 原理,回归问题可以表示为约束优化问题:

$$\min_{w, \xi} \frac{1}{2} \|W\|^2 + \frac{C}{2} \sum_{i=1}^l \xi_i^2,$$

$$\text{s.t. } y_i = W \cdot \phi(o_i) + b + \xi_i, \quad i = 1, 2, \dots, l.$$

用 Lagrange 法把约束优化问题变成无约束优化问题,即

$$\min_{w, b, \xi, \alpha} L = \frac{1}{2} \|W\|^2 + \frac{C}{2} \sum_{i=1}^l \xi_i^2 - \sum_{i=1}^l \alpha_i (W \cdot \phi(o_i) + b + \xi_i - y_i) \quad (6)$$

根据 KKT 条件有:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^l \alpha_i \phi(o_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^l \alpha_i = 0 \\ \frac{\partial L}{\partial \xi_i} = 0 \rightarrow \alpha_i = C \xi_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \rightarrow W \cdot \phi(o_i) + b + \xi_i - y_i = 0 \end{cases} \quad (7)$$

从方程组(7)中消去 w, ξ_i 后得到:

$$\begin{bmatrix} 0 & I^T \\ I & Q + C^{-1}E \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (8)$$

其中, $Y = (y_1, y_2, \dots, y_l)$, $I = (1, 1, \dots, 1)$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)$, $Q = (\phi(o) \times \phi(o)) = K(o, o)$, $i, j = 1, 2, \dots, l$.

由式(8)可以解出 α , 把 α 代入方程组(7)可得 W 和 b , 把 W 和 b 代入式(5)得出 LSSVR 的输出,即拟合函数:

$$f(o) = \sum_{i=1}^l \alpha_i K(o, o_i) + b \quad (9)$$

此拟合函数(9)就是对手效用函数的估计 $EU_{opp}(\bullet)$.

2.4 求解最优 offer

LSSVR 得到对手的效用函数估计 $EU_{opp}(\bullet)$, 根据式(1)把它与己方的效用函数 $U_{player}(\bullet)$ 加权组合成一个综合效用函数 $IU(\bullet)$. 为了使得协商得到双赢(即 $IU(\bullet)$ 最大), 同时满足 offer 落在己方的可接受区域, 可得到约束优化问题(2). 用遗传算法来求解此约束优化问题, 其基本过程如下:

```

t := 0;
初始化种群 pop(t) = {o1, o2, ..., on}
repeat
    计算 pop(t) 中个体的适应度;
    for i ← 1 to m do
        选择操作;
        对选出的个体进行交叉操作产生新的个体;
        对新产生的个体进行变异操作;
        把变异后的个体加入种群 pop(t) 中;
    end for
    按适应度降序排列 pop(t);
    pop(t) 前 n 个个体加入 pop(t+1)
    t ← t + 1;

```

Until (满足终止条件)

下面具体说明上述算法的每个步骤.

- 1) 编码表示: 本文采用了实数编码, 这样编码方式和 offer 的表示方式能够完全一致, 即编码的染色体结构为 (b_1, b_2, \dots, b_n) , b_i 为染色体中的一个基因, 代表协商项 s_i 的值. 通过模拟实验的对比, 本文的编码方式得到的计算结果优于经典的二进制编码方式.
- 2) 初始化种群: 初始种群中个体的产生方法为: 先为每个协商项设置一个取值区间 $[L_i, H_i]$, 每个基因 b_i 都在相应的区域随机取值, 然后判断新生成的个体是否属于可接受域, 如果满足, 则把个体加入到初始种群, 重复上面的步骤直至种群规模达到设定的大小. 设定种群的规模为协商项个数的 10 倍.
- 3) 适应度计算: 适应度函数为优化问题(2)的目标函数: $IU(o) = \beta \times U_{player}(o) + (1 - \beta) \times EU_{opp}(o)$, 把个体直接带入到适应度函数中就可以得到个体的适应度.
- 4) 选择操作: 本算法采用适应度比例法(fitness proportional model), 也称轮盘赌法或蒙特卡罗选择方法. 个体被选中的概率与它的适应度成正比, 适应度越高, 选中的概率就越大. 其中, 群体中第 i 个个体被选中的概率 sp_i 为 $sp_i = \text{Fitness}(i) / \sum_{j=1}^{\text{pop_size}} \text{Fitness}(j)$.
- 5) 交叉操作: 对选出的两个个体进行交叉操作, 使用以下两种交叉策略: ① 当协商项少于 5 个时, 采用单点交叉; ② 当协商项多于 5 个时, 采用多点交叉, 且每增加 5 个协商项相应地增加 1 个交叉点.
- 6) 变异操作: 因为本文采用了实数编码方式, 所以相应地使用实值变异策略. 实值变异中最重要的是选择变异的步长. 算法采用的变异算子为 $b'_i = b_i \pm 0.5(H_i - L_i)\Delta$. 其中, $\Delta = \sum_{i=0}^m a(i)/2^i$, $a(i)$ 以概率 $1/m$ 取值为 1, 以概率 $(1-1/m)$ 取值为 0, $m=20$. b_i 为变异前的基因值; b'_i 为变异后的基因值; L_i, H_i 为基因 b_i 取值的下限和上限. 变异的概率设定为 0.15. 当变异后的染色体不在自己的可接受域范围内时放弃变异.
- 7) 算法终止条件

终止条件控制算法何时结束. 算法的终止条件要综合考虑计算时间的限制和解的质量. 本算法定义了两个结束条件, 满足其中的一个条件则算法终止: ① 迭代次数超过 300; ② 迭代时间超过计算时间的上限.

3 实验

在实验室开发的服务协同式自治 VO 系统中实现了本文提出的协商决策模型. 为了分析模型的性能, 选取一个在线视频服务的协商实验用例. 此协商的参与者为一个服务提供者 $Agent_1$ 和一个服务需求者 $Agent_2$. 协商包含 5 个协商项(服务的价格、质量、响应时间等). 协商双方的效用函数采用线性加权和函数^[10]. 每个协商项 s_i 定义一个子效用函数 $v_i(b_i)$, 子效用函数定义如下:

$$v_i(b_i) = \begin{cases} -1, & \frac{b_i - tol_i}{des_i - tol_i} < 0 \\ \frac{b_i - tol_i}{des_i - tol_i}, & 0 \leq \frac{b_i - tol_i}{des_i - tol_i} \leq 1, \\ 1, & \frac{b_i - tol_i}{des_i - tol_i} > 1 \end{cases}$$

b_i 为 offer o 中协商项 s_i 的值, des_i 为 Agent 渴望得到的协商项 s_i 的值(渴望值), tol_i 为 Agent 对协商项 s_i 能够被接受的极限值(容忍值).总的效用函数 $U(o)$ 是子效用函数 $v_i(b_i)$ 的加权和.

$$U(o) = \begin{cases} \sum_{i=1}^n w_i v_i(b_i), & \forall v_i(b_i) \geq 0 \\ -1, & \exists v_i(b_i) < 0 \end{cases}$$

$\sum_{i=1}^n w_i = 1$, w 为协商项在效用函数中的权重. $Agent_1$ 与 $Agent_2$ 的 5 个协商项权重分别为(0.5,0.1,0.2,0.15,0.05)

与(0.1,0.6,0.05,0.2,0.05).取 $u_i^{\min} = -1$ $u_i^{accept} = 0$ $u_i^{\max} = 1$ ($i=1,2$),支持向量机采用线性核.协商决策模型的算法用 C# 语言实现,两个协商 Agent 分别运行在两台通过局域网相连配置为 P4 2.0G 处理器,512M 内存,Windows XP 操作系统的 PC 机上.

我们从两个方面来分析模型的性能:① 协商回合数.在电子商务的环境中,这是协商的一个重要性能指标.为了加快协商进程,协商双方希望用尽量少的协商回合达成交易.② 协商双方的联合效用.联合效用越大,协商性能越优;如果协商失败,则协商双方的效用都为 0.

首先分析可行域比重 E 不变时(E 设为 0.41%),协商双方采用不同协商决策模型,协商的性能变化.第 1 个阶段,协商的双方($Agent_1, Agent_2$)都采用随机决策模型(random decision model,简称 RDM)^[14]来进行协商.在随机决策模型中,协商 Agent 每个回合都会作出适当的让步,即在每个回合都降低一定的 offer 效用值的条件下,随机选择各个协商项的值.当每次协商完成时,相应的协商纪录存入协商历史数据库中.第 2 个阶段,协商双方都采用本文提出的协商决策模型(LDM),利用前一阶段所存储的协商数据作为模型的历史数据.实验的第 3 个阶段为协商的一方采用随机决策模型 RDM,另一方采用本文提出的学习模型 LDM.其中,学习模型的参数保持与第 2 阶段相同.每个阶段都单独进行 1 000 次协商,取平均值作为每个阶段的实验结果.表 2 显示了 3 个阶段的实验结果.实验结果表明如果协商双方都采用 RDM,则协商的联合效用最小,而且协商所需回合数最多.当一方采用 LDM,另一方采用 RDM 时,这两个性能指标有明显的改善;当双方都采用 LDM 时,这两个指标将达到最优.这是因为 RDM 只是在随机地摸索双方都能接受的 offer,所以需要更多的回合数,且找到使双方联合效用最大的 offer 的概率较小.而 LDM 每一步都尽可能地使得双方的联合效用最大,这样容易达到双赢的结果,而且能够较快地达成一致的 offer.

Table 2 Performance in three experiment phases

表 2 3 个实验阶段的性能

	Joint utility	Number of negotiation round
RDM vs. RDM	0.393	29.2
RDM vs. LDM	0.512	21.3
LDM vs. LDM	0.546	15.3

在进一步分析可行域比重 E 变化的情况下,协商双方采用协商决策模型 RDM^[14]、Probabilistic negotiation decision making mechanism(PDM)^[13]、LDM 时,协商性能的变化情况.PDM 方法只能处理离散型的协商变量,使用 PDM 进行模拟实验前,需要对协商项进行离散化处理.具体的处理方法是把协商项的取值区间均匀地划分成 10 等份.从图 4 与图 5 可以看出,在不同可行域比重 E 下,LDM 所需协商回合数最少,且协商双方获得的联合效用最高.这是因为 PDM 的方法只是从历史数据中挖掘出什么样的 offer 能够被对手接受,而本文提出的方法不仅考虑了 offer 能否被对手接受,而且还进一步估计了每个 offer 的对手效用值.这样,对手的效用信息更加准确,

使得协商过程更容易达到双赢的结果.

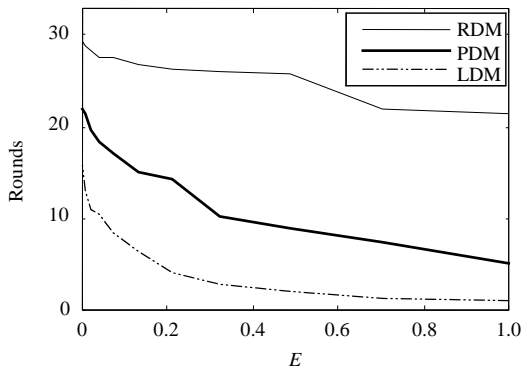


Fig.4 Number of negotiation rounds at given E

图 4 给定 E 值下的协商回合数

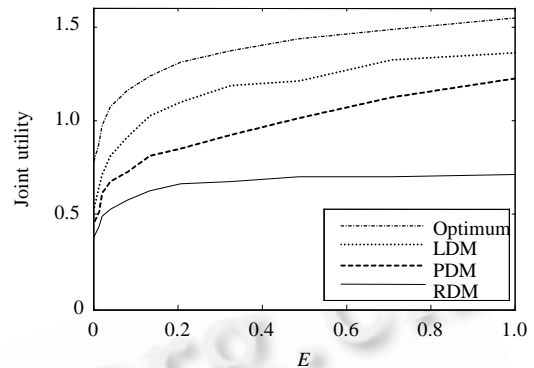


Fig.5 Joint utility at given E

图 5 给定 E 值下的联合效用

最后,我们分析了单个回合计算反建议的时间,它是整个协商所消耗时间的重要组成部分.图 6 显示了当协商项增加时,LDM 与 PDM 计算反建议的时间.当协商项增加时,两种方法的计算时间都呈上升趋势.但 PDM 的上升速度明显高于 LDM.这是因为 PDM 的计算复杂度与协商项个数成指数级增长关系.当协商项的数目较多时,在反建议计算时间方面,LDM 与 PDM 相比将体现出较大的优势,且协商项数目越多,其优势越明显.

4 结束语

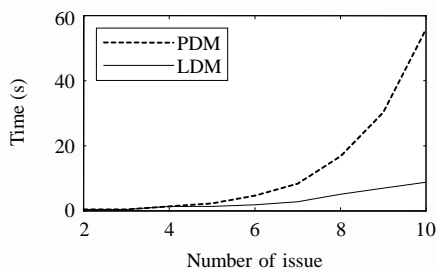


Fig.6 Time of compute counter-proposal at given number of issues

图 6 给定协商项个数下反建议计算时间

针对电子商务环境中对手信息不透明的情况,本文提出了一个基于学习的协商决策模型.虽然在以往很多的自动协商研究中用到机器学习的方法来增强 Agent 的协商能力,但它们或多或少地需要对手的私有信息或应用领域的先验知识.这往往成为了这些方法实用性的最大障碍.本文所提出的模型不需要利用任何应用领域的先验知识,只通过学习历史的协商数据来获得较优的协商策略.

本文中考虑的是一对一的协商,从协商开始到结束 Agent 面对的协商对手只有 1 个,在以后的工作中将发展到同时与多个对手协商.在面对多个协商对手的情况下,如何利用学习的方法提高协商性能将是下一步研究的方向.

致谢 在此,我们向对本文的工作给予支持和建议的实验室成员郭航、艾解清、彭艳斌同志表示感谢.

References:

- [1] Jennings NR, Faratin P, Wooldridge M. Automated negotiation: prospects, methods and challenges. *Int'l Journal of Group Decision and Negotiation*, 2001,10(2):199-215.
- [2] Du TC, Chen HL. Building a multiple-criteria negotiation support system. *IEEE Trans. on Knowledge and Data Engineering*, 2007, 19(6):804-817.
- [3] Ragone A, di Noia T, di Sciascio E, Donini FM. Description logics for multi-issue bilateral negotiation with incomplete information. In: *Proc. of the 22nd AAAI Conf. on Artificial Intelligence (AAAI 2007)*. Menlo Park: AAAI Press, 2007. 477-482.
- [4] Fatima SS, Wooldridge M, Jennings NR. Approximate and online multi-issue negotiation. In: *Proc. of the 6th Int'l Conf. on Autonomous Agents and Multi-Agent Systems*. Heidelberg: Springer-Verlag, 2007. 947-954.

- [5] Mok WWH, Sundarraj RP. Learning algorithms for single-instance electronic negotiations using the time-dependent behavioral tactic. *ACM Trans. on Internet Technology*, 2005,5(1):195–230.
- [6] Rahwan I, Sonenberg L, Jennings NR, McBurney P. STRATUM: A methodology for designing heuristic agent negotiation strategies. *Int'l Journal of Applied Artificial Intelligence*, 2007,21(6):489–527.
- [7] Li CH, Giampapa J, Sycara K. Bilateral negotiation decisions with uncertain dynamic outside options. *IEEE Trans. on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, 2006,36(1):31–44.
- [8] Gerding EH, Poutré HL. Bilateral bargaining with multiple opportunities: Knowing your opponent's bargaining position. *IEEE Trans. on Systems, Man and Cybernetics, Part C: Applications and Reviews*, 2006,36(1):45–55.
- [9] Zeng DJ, Sycara K. Bayesian learning in negotiation. *Int'l Journal of Human-Computer Studies*, 1998,48(1):125–141.
- [10] Coehoorn RM, Jennings NR. Learning an opponent's preferences to make effective multi-issue negotiation trade-offs. In: *Proc. of the 6th Int'l Conf. on E-Commerce*. New York: ACM Press, 2004. 59–68.
- [11] Wang LM, Huang HK, Chai YM. Choosing multi-issue negotiating object based on trust and K-armed bandit problem. *Journal of Software*, 2006,17(12):2537–2546 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/17/2537.htm>
- [12] Gao J, Zhang W. An accelerating chaos evolution algorithm of bilateral multi-issue automated negotiation in MAS. *Journal of Computer Research and Development*, 2006,43(6):1104–1108 (in Chinese with English abstract).
- [13] Lau RYK, Wong O. Mining negotiation knowledge for adaptive negotiation agents in e-marketplaces. In: *Proc. of the 40th Hawaii Int'l Conf. on System Sciences*. Piscataway: IEEE Computer Society Press, 2007. 130–139.
- [14] Faratin P, Sierra C, Jennings NR. Using similarity criteria to make issue tradeoffs in automated negotiations. *Artificial Intelligence*, 2002,142(2):205–237.
- [15] Suykens JAK, Vandewale J. Least squares support vector machine classifiers. *Neural Processing Letters*, 1999,9(3):293–300.

附中文参考文献:

- [11] 王黎明,黄厚宽,柴玉梅.基于信任和 K 臂赌博机问题选择多问题协商对象. *软件学报*,2006,17(12):2537–2546. <http://www.jos.org.cn/1000-9825/17/2537.htm>
- [12] 高坚,张伟.多 Agent 系统中双边多指标自动协商的 ACEA 算法. *计算机研究与发展*,2006,43(6):1104–1108.



程昱(1977—),男,湖南湘潭人,博士,讲师,主要研究领域为分布式人工智能,机器学习,自动协商.



古华茂(1975—),男,博士,讲师,主要研究领域为人工智能,本体论.



高济(1946—),男,教授,博士生导师,主要研究领域为人工智能,网络计算,软件工程,信息智能.



傅朝阳(1977—),男,博士生,讲师,CCF 学生会员,主要研究领域为 MAS,ES,SOC.