Expert Systems with Applications 41 (2014) 7630–7640

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A reinforcement learning optimized negotiation method based on mediator agent

Lihong Chen [a], Hongbin Dong [a,*], Yang Zhou [b]

[a]Department of Computer Science and Technology, Harbin Engineering University, Harbin, China
[b]Department of Humanities, Harbin Engineering University, Harbin, China

ARTICLE INFO

Article history:
Available online 12 June 2014

Keywords:
Multi-agent system
Reinforcement learning
Optimized negotiation
Mediator agent
Negotiation strategy
Adaptive learning

ABSTRACT

This paper firstly proposes a bilateral optimized negotiation model based on reinforcement learning. This model negotiates on the issue price and the quantity, introducing a mediator agent as the mediation mechanism, and uses the improved reinforcement learning negotiation strategy to produce the optimal proposal. In order to further improve the performance of negotiation, this paper then proposes a negotiation method based on the adaptive learning of mediator agent. The simulation results show that the proposed negotiation methods make the efficiency and the performance of the negotiation get improved.

1. Introduction

Negotiation is an important means to realize online e-commerce and also an important goal of designing software agent (Jennings, Faratin, & Lomuscio, 2001). How to improve the automatic negotiation ability of agent has always been one of the important problems multi-agent system need to resolve urgently.

Now there have been a lot of methods researching on agent automated negotiation. Such as, Costin Bădică et al. proposed a monotonous concession protocol based on mediator agent (Bădică, 2010), and verified its effectiveness, but its applicability in different environment still needs further research. Nabila Hadidi et al. proposed a alternate biding negotiation protocol based on argument (Hadidi, Dimopoulos, & Moraitis, 2011), which promoted both sides to reach an agreement, but did not research on selecting the optimal proposed strategy, which affected the quality of the negotiation. Tianhao Sun et al. proposed a reinforcement learning negotiation strategy based on Bayesian classification (Sun, Chen, Zhu, & Cao, 2011), experimental results verified the effectiveness and usability of the strategy, but the strategy could not guarantee the equality of both sides, and both agents compromised too fast, which did not conform to the real life. Xin Sui et al. proposed an multi-agent negotiation strategy based on Q-Learning (Sui, Cai, & Shi, 2010), which shorted the negotiation time and improved the efficiency of negotiation, but the multiple issues here are independent of each other. Linlan Zhang et al. put forward a multi-round sealed-bid negotiation mechanism based on mediator agent (Zhang, Song, Chen, & Hong, 2010), this method was simple, stable, and fair, but had limitations and inflexibility in the form of the utility function, and had more negotiation times and poor efficiency in negotiation strategy. Ahmad Esmaeili et al. used multi-objective particle swarm optimization algorithm to improve agent negotiation (Esmaeili & Mozayani, 2010), this method was fast and efficient but ignored the competitive relationship between agents. Jan Richter et al. proposed multi-stage fuzzy decision-making method to simulate the adaptive negotiation strategy (Richter, 2010), the negotiation agents had limited and uncertain information of the competitor, this method was adaptive to different opponents behavior. Amine Chohra et al. proposed social and cognitive negotiation behavior according to the automated negotiation with incomplete information (Chohra, Bahrammirzaee, & Madani, 2011), this method could improve the negotiation effectiveness, reduce the number of negotiation, but still needed to further improve the equality of the negotiation.

According to the above works, it can be seen that the developed methods always just consider the improvement of the single negotiation performance such as the equality of negotiation, the utility of negotiation, or the negotiation time and so on, which would always affect the other performance accordingly under improving some negotiation performance; or always ignore the correlation between issues; or the negotiation model is much idealistic, which does not conform to the negotiation in real life.

Thus, the objective of this article is to improve the negotiation method and to optimize the negotiation performance of agent from the negotiation model, negotiation strategy, the adaptive learning ability of mediator agent and so on.

* Corresponding author.
E-mail address: donghongbin@hrbeu.edu.cn (H. Dong).

http://dx.doi.org/10.1016/j.eswa.2014.06.003
0957-4174/

The main contributions of this paper include the following four parts:

– Considering the equality of negotiation and that both negotiators refuse to disclose more information for their own interests, it introduces a mediator agent as the mediation mechanism. The buyer agent and seller agent submit their proposals to the mediator agent simultaneously, then the mediator agent need to judge whether there is a trading opportunity and determine the final negotiation agreement.
– Reinforcement learning (Gao, Chen, & Lu, 2004) is an efficient machine learning method, learning from environment state to action mapping to maximize the cumulative reward obtained from environment, finding the optimal behavioral strategy by trial and error. Because it need not to learn the Markov decision model (*T* function and *R* function), Q-Learning is one of the most commonly used algorithm in the reinforcement learning. Therefore, to improve the efficiency of negotiation, this paper uses Q-Learning algorithm to generate the optimal negotiation strategy dynamically. In the process of negotiation, Q-Learning algorithm will calculate the cumulative reward values (that is the negotiation utility) of all negotiation strategies according to the current environment, and ultimately choose the negotiation strategy with the maximal reward value.
– To avoid the negotiation agents making too much concession in the beginning, it introduces the parameter expected reduction rate to restore the original expected utility.
– We also introduce a benchmark concessions utility function. The mediator agent adjusts the buyer and the seller agent's negotiation strategy through the benchmark function, which is to coordinate the negotiation between agents.

Finally, we have carried out two set of comparative experiments to verify the effectiveness and the efficiency of the optimization methods proposed.

2. Bilateral optimized negotiation model based on RL

This model negotiates on the issue price and the issue quantity. It also introduces a mediator agent as mediation mechanism, using utility function to evaluate the proposals.

2.1. Formal definition of the model

This section mainly defines some basic symbols, using the general negotiation model of agent, described as: $\text{Neg} = \langle G, A, D, U, T\rangle$, where *G* is the set of negotiating agents, including the buyer agent b, the seller agent s and the mediator agent m. A represents the set of negotiation issues, including the issue price *p* and the issue quantity *q*. *D* is the value interval of negotiation issues, $[IPb, RPb]$ and $[RPs, IPs]$ respectively represent the price interval of the buyer agent and the seller agent. *U* is the utility of negotiation agents. $U = \{U_b, U_s, UR_b, UR_s\}$, respectively represent the utility of the buyer and the seller, the reservation utility of the buyer and the seller. T represents the deadline of negotiation, $T = \{Tb, Ts\}$, respectively represent the deadline of the buyer and the seller.

There are some symbols not listed, in order to understand easily we will introduce the corresponding symbols when give the related definitions.

2.2. Utility evaluation

The utility function is used to evaluate a proposal and a standard judging whether an agent receives a proposal or not. We assume each agent incurs a fixed cost, *c*. The buyer and the seller's expected utility are defined as in Zhang et al. (2010):

(1) Seller agent

$$U_s(p,q) = (p - RPs)\cdot q - c \tag{1}$$

(2) Buyer agent

The buyer determines its purchase quantity q according to the market demand *x*. We assume x is a uniform probabilistic distribution on $[a, b]$. Let p_0 denote the sale price. If $q > x$, the buyer will sell the additional products with a low price αp_0. Assuming that αp_0 is lower than purchase price *p*, and p_0 is higher than *p*. Since $p \in [IPb, RPb]$, accordingly, $\alpha p_0 < IPb$ and $p_0 > RPb$. Then for a given proposal $\langle p, q\rangle$, the buyer agent's utility function f is defined as follows:

$$f(p,q,x) = \begin{cases} (p_0 - p)\cdot q & \text{if } q \leqslant x \leqslant b, \\ (p_0 - p)\cdot x + (p - \alpha p_0)\cdot(q - x) & \text{if } a \leqslant x \leqslant q. \end{cases}$$

Note that *x* is a variable, so the buyer's expected utility is defined as follows:

$$\begin{aligned} U_b(p,q) &= \int_a^b f(p,q,x)\cdot\frac{1}{b-a}dx - c = \frac{1}{b-a}\left(\int_a^q (p_0 - p)\cdot x \right.\\ &\quad \left. + \left(\alpha p_0 - p\right)\cdot(q-x)dx + \int_q^b (p_0 - p)\cdot q\,dx\right) - c \\ &= \frac{1}{b-a}\cdot\left(-\frac{(1-\alpha)p_0}{2}\cdot q^2 + (p_0(b-\alpha a) - p(b-a))\cdot q \right.\\ &\quad \left. - \frac{(1-\alpha)p_0}{2}\cdot a^2\right) - c. \end{aligned} \tag{2}$$

2.3. The improved reinforcement learning negotiation strategy

This paper uses Q-Learning algorithm to generate the optimal proposed strategy. In the process of generating proposals, we need consider the time belief, issue belief (i.e., probability distributions knowledge), etc. Therefore, following the reference (Sun et al., 2011) here gives three definitions:

Definition 1. Time belief refers to the probability that negotiation agent thinks the opponent accept its offer, which is correlated with time. *bfs* (t), *sfb* (t) respectively represent the time belief of the buyer and the seller, which generally falls into three categories: increasing function (e.g., $bfs = t/Tb$), decreasing function (e.g., $bfs = 1 - t/Tb$), and constant function (e.g., $bfs = 0.5$). Here we have $bfs = 1 - t/Tb$ and $sfb = 1 - t/Ts$.

Definition 2. Price belief refers to the cognition of probability distribution that negotiation agent thinks the issue p in its value interval. *bps* (t), *spb* (t) respectively represent the price belief of the buyer and the seller, which $bps\ (t) = 1/(p_0 - IPb)$, $spb\ (t) = 1/(IPs - RPs)$.

Definition 3. Quantity belief refers to the cognition of probability distribution that negotiation agent thinks the issue q in its value interval. *bqs* (t), *sqb* (t) respectively represent the quantity belief of the buyer and the seller. Assuming that the value interval of *q* of both agents is $[a, b]$, then $bqs\ (t) = sqb\ (t) = 1/(b - a)$.

Both agents submit their offers to the mediator agent simultaneously. Only after the negotiation achieves success, it can obtain the corresponding reward value *r*. Expected utility is the reward value acquired after success. Assuming when the negotiation achieves success, the agreement price is p^T, and the agreement quantity is q^T, then the reward value of each agent is as follows: