

複数回交渉問題における Boosting を用いた効用関数の推定手法

Estimation of Opponents' Utility Functions using Boosting for Multi-times Negotiations

松根 鷹生*¹ 藤田 桂英*¹
Takaki Matsune Katsuhide Fujita

*¹東京農工大学大学院 工学府 情報工学専攻

Department of Computer and Information Sciences, Graduate School of Engineering, Tokyo University of Agriculture and Technology

Recently, multi-issue closed negotiation problems have attracted attentions in the multi-agent system. Especially, multi-times and multi-lateral negotiation strategies are novel and important studies in multi-issue closed negotiation problems. An automated negotiating agent needs to have strategies for estimating opponent's utility function by learning opponent's behaviors through negotiations since opponent's utility information is not open to others.

In this paper, we propose the method of estimating the opponents' utility functions by combining several existing methods of estimating the opponents' utility functions using Boosting based on the least-squares method and nonlinear programming. Our experimental results demonstrate that the estimation accuracy of utility function is significantly improved under various conditions compared with existing utility function estimation methods without Boosting.

1. はじめに

マルチエージェントシステムに関する研究において、自動交渉エージェントは重要な役割を果たしている。自動交渉エージェントの実現により、複数の独立かつ自律した主体（エージェント）間で発生する衝突に対して、自動的に交渉を行い、協調動作することが可能となる。また、現実に近い交渉設定に対応した自動交渉エージェントの開発及び発展により、人間同士の交渉を支援できるほか、現実の意思決定支援への応用などが期待される [Fatima 14, Kraus 01]。

最新の自動交渉エージェント競技会 (Automated Negotiating Agents Competition (ANAC)) [Aydogan 15] では効用情報非公開下での三者間複数論点交渉問題を対象とし、時間経過による獲得効用の減少や、現実的な交渉シナリオの採用など現実に近い交渉モデルを対象としている。さらに、複数回同じドメイン、エージェントとの交渉を繰り返すことで、過去の交渉履歴を活用できる複数回交渉問題も着目されている。以前に交渉した際の両者で交換した合意案候補 (bid) の情報を保管しておき、次の同じドメインで対戦する際に保存した情報を活用することで、機械学習を用いて、より正確に相手の効用関数を推定することが可能になる。相手の交渉戦略を学習する主な手法として、相手が提案してきた bid を自身の効用空間で評価し、交渉の時間経過により相手が今後どのように譲歩するかを予測する手法がある [Zhou 14]。一方、相手の効用関数を推定する手法も提案されているが [Chen 13, Chen 14]、さらなる推定精度の向上が重要な課題とされている。

本論文では、効用情報非公開下での三者間複数論点交渉問題を対象に、複数回交渉を行った際に、最小二乗法および非線形計画法をベースとした Boosting [Schapire 90] を用いて、複数の既存の効用関数推定手法を組み合わせることで効用関数の推定精度を向上させる手法を提案する。組み合わせる既存の効用関数推定手法は相手が提案した bid の選択肢を数え上げる手法、および、その手法をベースに相手が提案した bid の提案

時の交渉経過時刻を考慮し、重み付けを行った数え上げ手法を用いる [Ikrahi 14]。複数の効用関数推定手法を Boosting により適切な重みを付与して足し合わせることで、対戦相手やドメインに合わせた適切な効用関数の推定を行うことができる。提案した効用関数推定手法を評価するためにシミュレーション実験を行った。効用関数推定手法に関して、多くの場合において既存手法よりも高い推定精度が得られることを示す。

以下に本論文の構成を示す。まず、本研究で対象とする多者間複数回交渉問題について説明する。次に、Boosting を用いた相手効用関数の推定手法を提案する。その後、提案した効用関数推定手法の評価を行い、最後に結論を述べる。

2. 多者間複数回交渉問題

本論文では、効用情報非公開下での三者間複数論点交渉問題を考える。交渉問題 (ドメイン) D は n 個の論点 I_1, \dots, I_n を含み、各論点 I_i は k 個の値 v_i^1, \dots, v_i^k を含む。論点 I_i の値 v_i^c には整数の評価値 $eval(v_i^c) \in [0, +\infty)$ が設定され、各論点には論点の重み $w_i \in \mathbb{R}$ が与えられる。論点の重み w_i は $\sum_{i=1}^n w_i = 1$ かつ $w_i \geq 0$ を満たす。bid は各論点 I_i について選択肢 c を選択したとすると $\vec{b} = [v_1^1, \dots, v_n^1]$ と示され、効用関数 $U(\vec{b})$ は式 1 で示される。また、交渉エージェントはそれぞれ異なる効用関数を持つ。

$$U(\vec{b}) = \sum_{i=1}^n w_i \times \frac{eval(v_i^1)}{\max(eval(I_i))} \quad (1)$$

本論文では $t \in [0, 1]$ の範囲で正規化された時刻を扱う。 $t = 0$ は交渉開始時を意味し、 $t = 1$ は交渉期限時刻を意味する。また、効用情報によっては割引係数または留保価格が設定されることがある。割引係数 d とは交渉の経過時刻 t によって式 2 のように実際に獲得できる効用が減少するものである。留保価格は合意形成に失敗した場合に得られる最低効用値のことであり、合意形成成功時と同様に割引係数が適用される。

$$U_D(\vec{b}) = U(\vec{b}) \times d^t \quad (2)$$

本論文では二者間交渉に使用される交渉プロトコルである Alternative Offers [Rubinstein 82] を三者間以上で使用できる

連絡先: 松根 鷹生, 東京農工大学大学院 工学府 情報工学専攻, 〒184-8588 東京都小金井市中町 2-24-16, matsune@katfujii.lab.tuat.ac.jp

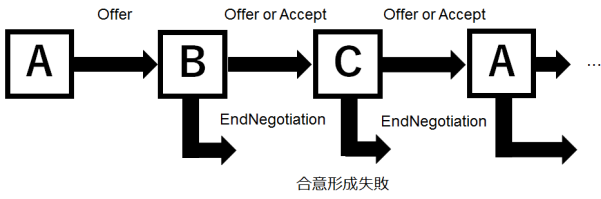


図 1: SAOP の流れ

ように拡張した Stacked Alternating Offers Protocol (SAOP) [Aydogan 16] を対象とする。SAOP の説明図を図 1 に示す。説明のために、エージェント A, B, C が交渉を行うことを考える。SAOP では各エージェントが A, B, C の順に行動する。交渉開始時にエージェント A が B に対し bid の提案を行う。次に、 B は C に対して以下のうちいずれかの行動を行う。

- *Offer*: 相手の bid を拒否し、新たに bid を提案する
- *Accept*: 相手の bid を承諾する
- *EndNegotiation*: 交渉を放棄する

その後は C が A に対し新たに行動するというように順番に行動する。SAOP では、以下に示す条件を満たすまで交渉が継続される。

- 最後に *Offer* したエージェント以外の全員が *Accept* する。
- 合意形成がされないまま交渉期限時刻を超過する
- いずれかのエージェントが *EndNegotiation* を行うことで交渉を放棄する

合意形成に成功するのは *Offer* したエージェント以外の全員が *Accept* した場合のみであり、残りは失敗となる。また、すべての行動はエージェント全員に共有される。

同一の交渉相手、交渉ドメインおよびプロファイルの条件下で複数回交渉を行う場合、以前に提案された bid と経過時間を引き継ぐことができる。本設定により、機械学習等を用いて、相手の戦略や効用関数の推定精度を向上できる可能性がある。

3. Boosting を用いた効用関数推定手法

独自の Boosting アルゴリズムを用いて既存の効用関数推定手法を組み合わせることで推定精度を向上させる手法を提案する。提案手法では同一の交渉相手、交渉ドメインおよびプロファイルの条件下で複数回交渉を行うこととし、2 回目以降の交渉において効用関数の推定精度を改善することを目的とする。

3.1 重み付けによる効用関数の推定手法の改良

Ikarashi らは bid が提案された時点での経過ラウンド数に基づいて重み付けを行い、相手の効用関数を推定している [Ikarashi 14]。しかし、bid 提案時のラウンド数に重み付けをする手法では、交渉の処理速度の変化などにより実際の経過時間とラウンド数の推移が一致しない可能性がある。そこで、重み付けに使用する変数を提案時のラウンド数ではなく、以下に示すように $[0, 1]$ の範囲に正規化した時刻を用いるように改良を加えた。

- (1) 重み付けなし: $V_i = \sum_{k=1}^n v_{i,k}$

(2) 右上がり一次関数: $V_i = \sum_{k=1}^n t_k \cdot v_{i,k}$

(3) 右下がり一次関数: $V_i = \sum_{k=1}^n (1 - t_k) \cdot v_{i,k}$

(4) 下に凸型二次関数: $V_i = \sum_{k=1}^n (t_k - \frac{1}{2})^2 \cdot v_{i,k}$

(5) 上に凸型二次関数: $V_i = \sum_{k=1}^n \{-(t_k - \frac{1}{2})^2 + \frac{1}{4}\} \cdot v_{i,k}$

(6) 指数関数による増大: $V_i = \sum_{k=1}^n e^{\frac{t_k}{\epsilon}} \cdot kv_{i,k}$

t_k はある交渉相手から k 番目に提案された bid の時刻であり、 $v_{i,k}$ は交渉対象のドメインが持つ m 個の Value v_1, \dots, v_m のうち i 番目の Value が提案されたかを $\{0, 1\}$ 変数で表す関数である。先行手法では対数関数による重み付けも含まれていたが、提案手法では評価値が負数になることを防ぐために削除した。

3.2 Boosting を用いた相手効用関数の推定手法

2 回目以降の交渉にて既存の効用関数推定手法 (学習器) に重み付けを行うことにより、効用関数推定精度を向上させる手法を提案する。

初期化: N 個の学習器に対し、学習器の重みベクトル $\mathcal{W} = \{w_1, \dots, w_{N+1}\}$ を用意する。それぞれの要素の値を $\frac{1}{N+1}$ に初期化する。また、学習器が算出した相手の bid の効用値を保存するためのベクトル $\mathcal{B}^{(R)} = \{b_1^{(R)}, \dots, b_N^{(R)}, b_{N+1}^{(R)}\}$ を用意する。

Step1: 各学習器 $\mathcal{L} = \{L_1, \dots, L_N, L_{N+1}\}$ で受信 bid の推定効用値を算出し、それぞれ $\{b_1^{(r)}, \dots, b_N^{(r)}, b_{N+1}^{(r)}\}$ として $\mathcal{B}^{(R)}$ のデータセットに追加する。 r は交渉開始時から現在までに受信した bid の数であり、学習器 L_{N+1} はすべての Value 評価値を 1 として出力するバイアスである。このバイアスは Boosting 過程において以下の役割を持つ。

(1) 学習過程において特定の学習器のみが選択されてしまうことを防止

バイアスを用意しない場合では Step2 において非線形計画問題を解く際に譲歩モデル $\mathcal{C}(t)$ よりもすべての学習器の出力曲線が下回る場合、一つの学習器に重みが偏り過ぎてしまう場合がある。バイアスを導入することにより学習器の重みの偏りを防ぐことが可能になる。

(2) 評価値が低い Value への補正

数え上げをベースとする効用関数推定手法は場合にもよるが、提案頻度が低い Value の評価値が極端に低くなる傾向があり、この傾向は 3.1 節の手法を用いた場合でも同様である。バイアスを加えることにより極端に低い Value の評価値を補正することができる。

Step2: 非線形計画法を用いて式 3 を解き、その解を学習器の重みベクトル $\mathcal{W} = \{w_1, \dots, w_{N+1}\}$ とする。

$$\begin{aligned} \text{minimize } & \sum_{i=1}^r \left(\sum_{j=1}^{N+1} (b_j^{(i)} \cdot w_j) - \mathcal{C}(t) \right)^2 \\ \text{s.t. } & \sum_{j=1}^{N+1} w_j = 1 \\ & w_j \geq 0 (j = 1, \dots, N+1) \\ & w_{N+1} \leq \frac{1}{N+1} \end{aligned} \quad (3)$$

式 3 の 1 行目は w_j により重み付けられた学習器 L_j が推定した効用値 $b_j^{(i)}$ と時刻 t で譲歩モデル $\mathcal{C}(t)$ が示す効用値との二

表 1: 評価実験の使用ドメイン

ドメイン *1	論点	ドメインの大きさ	DF	RV
Clockwork (ANAC2016)	3	120	0.6	0.4
MyAgent (ANAC2016)	5	1280	0.7	0.2
MaxOops (ANAC2016)	7	7200	1.0	0.0
Terra (ANAC2016)	7	35840	0.5	0.4
Politics (ANAC2015)	8	23040	1.0	0.0
University (ANAC2015)	5	2250	1.0	0.0
Party	6	3072	1.0	0.0

乗誤差である。譲歩モデル $C(t)$ は bid を提案した相手が今後自身の効用をどのように譲歩するかを予測した時刻 t の関数である。多くのエージェントの交渉戦略は時間経過により譲歩するため、交渉時間の経過につれて効用値が減少する関数を用意すれば推定誤差の減少につながると考えられる。式 3 の 4 行目は w_{N+1} が大きくなり過ぎることを防ぐための制約条件である。

Step3: 式 4 より、各学習器に重みを掛け、足し合わせることで重み付けられた学習器 L_w を得る。

$$L_w = \sum_{i=1}^{N+1} L_i \cdot w_i \quad (4)$$

4. 評価実験

評価実験には、自動交渉環境 GENIUS[Lin 14] を使用する。Boosting に基づく効用推定手法の評価では、以下のドメインおよびエージェントを使用し総当たり戦を行うことで交渉履歴を記録する。

- 実験方法:総当たり戦
- 交渉エージェント:ANAC2016 出場エージェントの個人効用部門上位 1~7 位 (Caduceus, YXAgent, ParsCat, Farma, MyAgent, Atlas3, Ngent)
- 交渉ドメイン:ANAC で使用されたドメイン等 7 種類 (表 1) エージェントの効用情報は各ドメインのプロファイル 1~3 をそれぞれ使用
- 交渉時間:180 秒
- 割引係数 (DF), 留保価格 (RV):交渉ドメインに依存
- 総交渉回数: 1470

次に記録した各交渉履歴から、各エージェントごとに 3.1 節で述べた重み付け関数を適用することで効用関数を推定する。指数関数のパラメータは $\xi = 0.3$ とした。得られた推定 Value 評価値は重み付け関数により大きさが異なるため、すべての Value が $[0, 1]$ の範囲内に収まるように正規化を行った。推定精度の評価は、Issue の重みについては推定を行っていないため、Value 評価値の推定のみを評価対象とする。はじめに各プロファイルの評価値を正規化し、各 Value の正しい評価値と

*1 ANAC2016 使用ドメインはドメインを提出したエージェントの名前を表記した

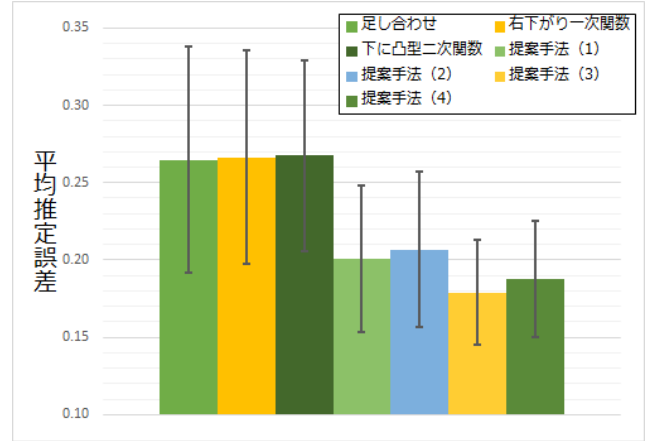


図 2: 平均推定誤差 *2

推定した評価値の差の絶対値を推定誤差として、各重み付け関数による推定誤差をエージェントごと、ドメインごとに算出することで評価する。

その後、提案手法の評価を行うために、記録した交渉履歴を以下の各提案手法に入力し、交渉終了時まで Boosting 処理を行い、重み付けされた効用関数を出力させ推定精度を評価する。Boosting 手法に使用する効用関数推定手法および譲歩モデルは以下の通りそれぞれ 2 種類ずつ用意し、合計 4 通りとした。提案手法の効用関数推定精度の評価は重み付け関数の評価手法と同様にエージェントごと、ドメインごとに推定誤差を算出することで行う。

- (1) 重み付け関数: 足し合わせ, 右上がり一次関数, 右下がり一次関数, 下に凸型二次関数, 上に凸型二次関数, 指数関数
譲歩モデル式: $C(t) = 1.0 - 0.2t$
- (2) 重み付け関数: 足し合わせ, 右上がり一次関数, 右下がり一次関数, 下に凸型二次関数, 上に凸型二次関数, 指数関数
譲歩モデル式: $C(t) = 1.0 - 0.3t$
- (3) 重み付け関数: 足し合わせ, 右下がり一次関数, 下に凸型二次関数
譲歩モデル式: $C(t) = 1.0 - 0.2t$
- (4) 重み付け関数: 足し合わせ, 右下がり一次関数, 下に凸型二次関数
譲歩モデル式: $C(t) = 1.0 - 0.3t$

図 2 および表 2 にドメインごとの平均推定誤差を示す。表 2 より、既存手法同士を比較すると、ドメインの平均で最も推定誤差が低くなったのは足し合わせを用いた場合であることが確認された。提案手法について図 2 より、各ドメインの平均にて、既存手法と比較して推定精度が大幅に向上した。全ドメインの平均 (表 2) の結果により、使用する効用関数推定手法の数が 6 種類の場合に比べ 3 種類の場合の方が全体として精度が良くなり、ドメイン全体での平均推定誤差が約 0.68 倍にまで低下し精度が大幅に向上していることが確認できる。また、提案手法 (1), (2) および (3), (4) 同士の比較により、譲歩モデルによる推定誤差の変化は僅かであることが確認された。

実験結果について、提案手法の精度が既存手法を上回った理由として、提案する Boosting 手法にて使用するバイアスの

*2 エラーバーは標準偏差を示す。

表 2: 平均推定誤差 (ドメイン:各ドメイン平均)

重み付け関数	ドメイン全体	Caduceus	YXAgent	ParsCat	Farma	MyAgent	Atlas3	Ngent
足し合わせ	0.2646	0.3061	0.1698	0.1722	0.3136	0.2670	0.2577	0.3660
右上がり一次関数	0.2688	0.2993	0.1706	0.1667	0.3346	0.2622	0.2615	0.3865
右下がり一次関数	0.2663	0.3136	0.1713	0.1840	0.3111	0.2691	0.2577	0.3577
下に凸型二次関数	0.2676	0.3078	0.1731	0.2012	0.3106	0.2716	0.2632	0.3458
上に凸型二次関数	0.2679	0.3060	0.1703	0.1707	0.3233	0.2640	0.2588	0.3824
指数関数	0.2689	0.2982	0.1710	0.1643	0.3407	0.2628	0.2621	0.3832
提案手法 (1)	0.2009	0.2169	0.1483	0.1306	0.2341	0.2104	0.1998	0.2660
提案手法 (2)	0.2068	0.2336	0.1494	0.1288	0.2461	0.2173	0.2071	0.2655
提案手法 (3)	0.1790	0.1846	0.1495	0.1266	0.2005	0.1892	0.1738	0.2292
提案手法 (4)	0.1875	0.2062	0.1495	0.1254	0.2172	0.1993	0.1861	0.2293

重み付けが適切にされたことが考えられる。しかし、バイアスの重みが大きくなり過ぎることにより、過剰に推定 Value 評価値が高くなる場合があることを確認した。提案手法 (1), (2) および (3), (4) の結果を比較することにより、譲歩モデルによる平均誤差の変化は僅かであることが確認された。この理由は用意した譲歩モデルがどちらも一次関数であり、式 3 の解に大きな影響を与えなかったためと考えられる。

5. 結論

本論文では、同一の交渉ドメインおよび効用情報において複数回交渉を行った際に、複数の既存の効用関数推定手法に最小二乗法をベースとした Boosting アルゴリズムを用いて重み付けを適切に行うことで、推定精度を改善する手法を提案した。提案した Boosting アルゴリズムの評価実験により、複数の既存の効用関数推定手法を組み合わせて、多くの場合にて効用関数の推定精度が向上していることが確認された。また、最良の場合では既存手法に比べ誤差が約 0.62 倍程度まで減少することが示され、本研究の目的である効用関数の推定精度の改善について有効性を示した。

今後の課題として効用関数推定精度のさらなる改善が挙げられる。バイアスの重みが大きくなり過ぎることにより、推定 Value 評価値が高くなり過ぎる場合があるため、精度向上のためにはバイアスの重み付けにより適切な制約を加える必要がある。また、使用した効用関数推定手法は bid の提案時刻を用いて重み付けを行うものであったが、重み付けを行った場合でも推定精度の大幅な改善には至らない場合があった。そのため、他の効用関数推定手法を併せて使用し、推定精度を改善することが今後の重要な課題である。

謝辞

本研究は、JST, CREST および JSPS 科研費 (JP15H01703) の支援を受けたものである。

参考文献

- [Aydogan 15] Aydogan, R., Baarslag, T., Fujita, K., Hindriks, K., Ito, T., and Jonker, C.: The Sixth International Automated Negotiating Agents Competition (2015)
- [Aydogan 16] Aydogan, R., Festen, D., Hindriks, K. V., and Jonker, C. M.: Alternating offers protocols for multilateral negotiation, *Modern Approaches to Agent-based Complex Automated Negotiation*. Springer (2016)

[Chen 13] Chen, S., Ammar, H. B., Tuyls, K., and Weiss, G.: Conditional restricted Boltzmann machines for negotiations in highly competitive and complex domains, in *Proceedings of the 23th International Joint Conference on Artificial Intelligence (IJCAI2013)*, pp. 69–75 (2013)

[Chen 14] Chen, S., Zhou, S., Weiss, G., and Tuyls, K.: Using Transfer Learning to Model Unknown Opponents in Automated Negotiations, in *Recent Advances in Agent-based Complex Automated Negotiation*, pp. 175–192 (2014)

[Fatima 14] Fatima, S., Kraus, S., and Wooldridge, M.: *Principles of Automated Negotiation*, Cambridge University Press (2014)

[Ikrashi 14] Ikrashi, M. and Fujita, K.: Compromising Strategy Using Weighted Counting in Multi-times Negotiations, in *IIAI 3rd International Conference on Advanced Applied Informatics (IIAI AAI2014)*, pp. 453–458 (2014)

[Kraus 01] Kraus, S.: *Strategic Negotiation in Multiagent Environments*, Mit Press (2001)

[Lin 14] Lin, R., Kraus, S., Baarslag, T., Tykhonov, D., Hindriks, K., and Jonker, C. M.: Genius: An Integrated Environment for Supporting the Design of Generic Automated Negotiators, *Computational Intelligence*, Vol. 30, No. 1, pp. 48–70 (2014)

[Rubinstein 82] Rubinstein, A.: Perfect equilibrium in a bargaining model, *Econometrica: Journal of the Econometric Society*, pp. 97–109 (1982)

[Schapire 90] Schapire, R. E.: The strength of weak learnability, *Machine Learning*, Vol. 5, No. 2, pp. 197–227 (1990)

[Zhou 14] Zhou, S., Chen, S., Smirnov, E., Weiss, G., and Tuyls, K.: Automatic Knowledge Transfer for Agent-based Bilateral Negotiation Tasks, in *International Workshop on Agent-based Complex Automated Negotiations (ACAN2014)* (2014)