

科学研究費助成事業 研究成果報告書

令和元年6月19日現在

機関番号：33910

研究種目：基盤研究(C) (一般)

研究期間：2015～2018

課題番号：15K00321

研究課題名(和文)人と機械学習の協調学習によるスマートグリッド制御システムの学習加速

研究課題名(英文) Collaboration technique of human and machine-learning for speed up of problem solving and its application for smart grid operations

研究代表者

山内 康一郎 (YAMAUCHI, Koichiro)

中部大学・工学部・教授

研究者番号：00262949

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：bigデータが無い未知の領域ではAIは学習ができない。そこで人間と学習機械が協調学習をすることによって、未知の問題に対する解を高速に導き出し、人と人工知能がともに成長する仕組み Collaborative Bagging(ColBagging)を考案した。これは複数のワーカーが未知の問題に対する解の候補を出す一方で、ワーカーの横には学習機械があり、対応するワーカーの振る舞いを模倣するように学習する。解候補は信頼性に応じて重みづけされたうえで平均され、ワーカーに教示する。そこから学びを得たワーカーはさらに良質の解候補を出すことを期待できる。学習後、学習機械は単独でも解候補を出せるようになる。

研究成果の学術的意義や社会的意義

実用に供される人工知能は、学習に非常に多くの学習データを必要とする。だが未知の領域では学習データそのものが無いため、人工知能は学習ができない。本研究成果は、この問題を解決する一手段を示した。この手法の最もユニーク且つ重要なポイントは、「人」と「人工知能」の両方の能力を高めるという点である。今後人が、人工知能に頼る場面が増えるならば、人はその能力を失う恐れがある。本手法はこれを防ぐ意味で重要である。

研究成果の概要(英文)：Artificial Intelligence generally needs a big data set for its learning. However, in unknown environments, such big data do not exist. To solve this difficulty, we proposed a “collaborative bagging” which is a variation of cloud sourcing technique to make workers solve unknown problems by collaborating with learning machines. For each worker, there is a learning machine that imitates the worker’s behaviors. The solution to candidates from the worker are evaluated according their confidence ratios and translate into a single solution by weighted average. The single solution is used for the feedback-teaching for each worker. As a result, each worker learns the solution to yield better solution candidate in the next time. Therefore, the worker and the learning machine improve their ability gradually by repeating this cycle. Finally, the learning machines become to work as a very excellent problem solver after the latter steps of the learning.

研究分野：知能情報学

キーワード：機械学習 人工知能 協調学習 未知の問題領域 クラウドソーシング

様式 C-19、F-19-1、Z-19、CK-19 (共通)

1. 研究開始当初の背景

人工知能が急激に実用されはじめるも、実用に供するためのネットワークは巨大であり、決定しなければならないパラメータ数は1億個を超えるケースもある。通常、学習に必要なサンプルのサイズは、パラメータ数以上は最低限必要であることを考えると、そのデータ収集にかかるコストが大きすぎることは明白である。

特に、誰も最適解を知らない問題領域においては、学習データそのものが存在しない。このような領域では現状の人工知能の学習は不可能であった。

2. 研究の目的

誰も最適解を知らない問題領域において、人間と学習装置が共に学習を進めることで、素早く最適解を発見・学習が完了するシステムを考案することが目的である。

3. 研究の方法

人間の洞察力の力を借りて、未知の問題領域の回答を集め、機械学習を進める。このためにクラウドソーシングに似たアプローチをとった。クラウドソーシングでは不特定多数のワーカーに、最適解が不明な問題を提示し、回答案を提出させる。そして信頼性の高い回答をピックアップすることで解を得るものである。ただ、従来のクラウドソーシングはワーカーの誰かが正確な回答をすることが前提であり、まったく未知の問題に対する解は答えようがない。そこで、ここでは、

- (1) 多数の回答案の中から暫定的に最も良いと思われる解をワーカーにフィードバックし、ワーカーに自らの回答案を修正させる。これを繰り返して解を洗練させる。
- (2) 学習器にも同時にこれを学習させる。具体的には、それぞれのワーカーの回答案を模倣するように学習させ、学習後は学習器だけでも回答できるようにする。

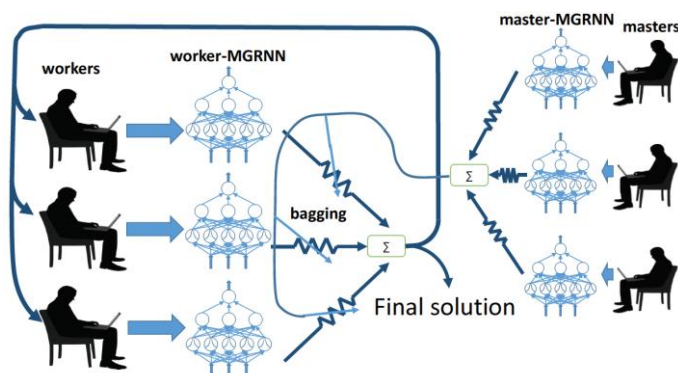


Fig 1 ColBagging2

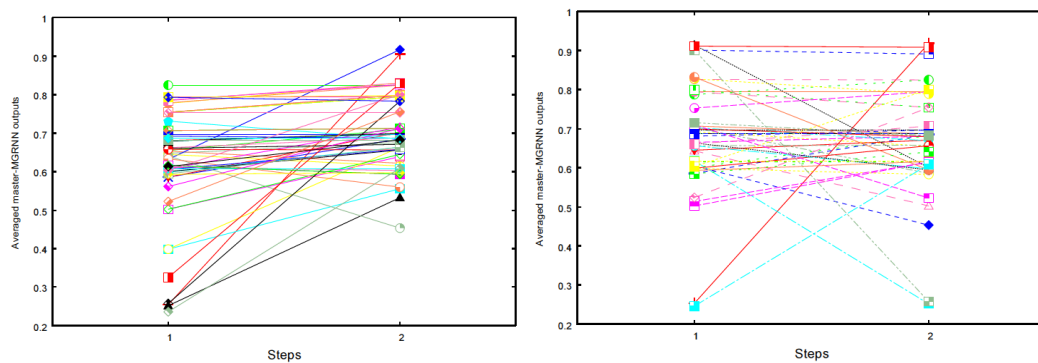


Fig 2 回答案提出->評価->統合->フィードバックを1stepとしてこれを2回繰り返した場合の提案法でカラーコーディネートした被験者の回答案に対する Master の評価値の平均値(左)とコントロール (feedback teaching を受けない被験者) 被験者の結果に対する Master の評価値の平均値 (右)。左図の方が2回目の評価値が優位に高かった。

4. 研究成果

(1) 手法開発

Collaborative Bagging (ColBagging) システムを提案した[雑誌論文①, 学会発表⑦] (Fig1 参照)。これは複数のワーカーが未知の問題に対する解の候補を出す一方で、ワーカーの横には学習機械があり、対応するワーカーの振る舞いを模倣するように学習する。解候補は信頼性に応じて重みづけされ、それらの重み付き平均され、ワーカーに教示する。そこから学びを得たワーカーはさらに良質の解候補を出すことを期待できる。学習後、学習機械は単独でも解候補を出せるようになる。実験として、カラーコーディネートに関する問題を複数人の被験者に提示して実験した結果、本手法を使った方が多くの被験者が良い配色を選べるようになった。

た(Fig 2 参照)。ただし、重み付き平均から良い配色を選択できるようになるまでの時間は、従来のクラウドソーシング手法とくらべて有意な差が得られなかった。この時点では、心理実験の規模が小さく、学習加速の証拠はまだ得られない状態であった。

(2) 大規模心理実験の試み

ColBagging による「学習加速」を確認するためにはさらに大規模な心理実験が必要であった。そこで対象とする課題としてスマートグリッドの手法の一つである電力卸売取引への応用を例にとってこれを確かめようとした。これは入札を伴う取り引きで、それぞれのワーカーは売電業者もしくは、買電業者に相当する。売電と買電ではそれぞれ対象の異なる課題を解くことになるため、ColBagging は二対必要となる(Fig 3 参照)。

まずは「買電業に限って」大規模な実験環境を整えた[学会発表⑤](Fig 4)。価格は単純に天候から計算式で求める形を取った。数十名の被験者に依頼してこれを行った。ところが、この単純化した取引は、他の人の助けを得るまでもなく取引スキルを得ることが容易であったためか、初回から最高収益を得る被験者が多く、全体として学習加速があったかどうか、はっきりとしたデータは得られなかった[学会発表③⑤]。

[明らかになった課題]

- 学習の加速を観測するにあたり、もっと複雑で難しい課題：すなわち、他の人の意見をヒントにしながら問題に取り組まないと解けないほど難しい課題であるべきであった。
 - そこで当初の計画通り、買電も組み合わせて、本当に入札取引を行わせることが望まれる。しかし入札を伴う環境は、売電・買電業者双方の振る舞いによって評価関数が時間変動するため格段に難易度の高い最適化問題である。この問題に対して ColBagging が適用可能なのか？この点の見通しがつきにくい。例えば、Game ベースの取引課題に適用する場合、初期段階でのフィードバック教示は逆効果となる恐れがある。
 - 学習の加速を観測する実験を行う場合、被験者はこの実験に関する事前知識を持ってはいけい。つまり一度実験に参加した被験者は二度とこの実験に参加させてはいけいのである。その結果、実験を繰り返し実行しようとする、たちまち被験者の確保が困難となる。
- 以上の課題を踏まえ、以下に示すように計算機シミュレーションによって実験的に確かめることにした。

(3) PSO による近似的シミュレーション[学会発表②]

電力卸売取引は、売り手と買い手の相互作用によって望ましい需給状態が平衡状態になるように環境を整えようとするものである。この環境下において ColBagging を使った場合、Nash 平衡状態を発見できるのか？そしてその収束速度はどのようになるのか？を確認する。

具体的には、Particle Swarm Optimization を使って近似的に Colbagging の振る舞い全体を把握する。ColBagging は複数のワーカーの協力によって未知の問題に対する回答を効率よく発見するための仕組みであり、それぞれのワーカーは最も利益の大きかったワーカーの取引方法に自らの取引手法を近づけるように学習する。このような振る舞いは、Particle Swarm Optimization(PSO)の各々の粒子の振る舞いに似ている。

そこで PSO それぞれの粒子が入札希望価格や買電価格などを決定するようにした。各々の粒子が入札する額と電力量は各々の売電ワーカーの意思「高い価格で売却したい」が反映されるように次のように決定した。例えば第 i 売電 worker の入札する電気の量は次の式で決めるものとした。

$$\text{Amount}_i = 0.2 \left(0.8 \text{Battery}_i w_i^1 + \text{Weather}_{\text{today}} w_i^2 + \text{Weather}_{\text{tomorrow}} w_i^3 \right)^2 w_i^4$$

ここに Battery_i , $\text{Weather}_{\text{today}}$, $\text{Weather}_{\text{tomorrow}}$ は第 i ワーカーのバッテリー残量、今日と明日の天候を数値化したものである。 $(w_i^1 \ w_i^2 \ w_i^3 \ w_i^4)$ は PSO の枠組みに従って、最も

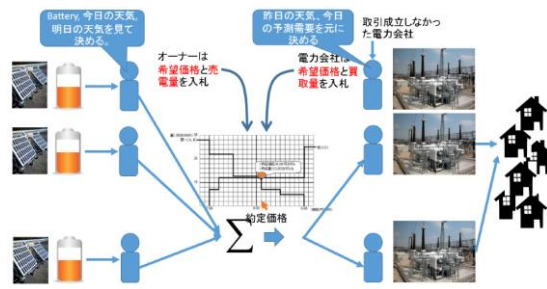


Fig 3 ブラインドシングルブライソオークションによる電力卸売りのモデル。売電業者は太陽電池と蓄電池のオーナーであり、今日の天気とバッテリー残量、明日の天気などの情報を基に売電量と希望価格を入札する。一方買電業者は需要を基に買電量と買電希望価格とを入札する。



Fig 4 取り引き用インターフェース

利益の大きかったワーカー j^* のパラメータ(w_j^1 w_j^2 w_j^3 w_j^4)に近づくように更新される。このようにしてモデル化したワーカーについて、重み(w_i^1 w_i^2 w_i^3)をプロットして、これがどのように推移するかを確かめた (Fig5 参照)。

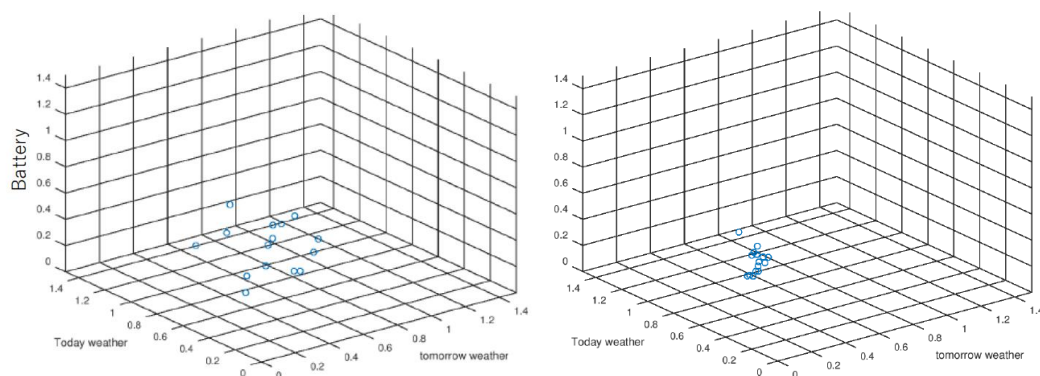


Fig 5 発電業者の入札する電力量を決める重みの分布 (左: 初回、右: 20取引後)

すると20取引後にはある特定の位置に収束している。同様に買電ワーカーについては本報告書では割愛するが、同様にパラメータが特定の値に収束した。約定価格もある一定値に落ち着いている (Fig 6)。これは初期状態を変えても同じように収束をし、再現性のある結果であった。

[考察] Game 理論の一例として有名なものにタカ派ハト派ゲームがある。このゲームではタカ派的に行動した者が一時的に利益を得ても、すぐにその負の影響がでて改善を迫られ、結果としてハト派的な状況に落ち着く。このゲームは一時的にタカ派が得をする。このような領域に ColBagging を適用したとすると、一時的に全員がタカ派になる恐れがある。一方電力卸売取引の場合、ブライントシングルプライスオークションという仕組みのため、各々のワーカーの利己的な入札は即、取引不成立を招く結果となる。そのため、売り手と買い手双方の納得のいく入札価格を設定する者が高い収入を得る結果となる。この価格は初期状態では売り手と買い手の思惑が入り乱れて大きく変動するが、すぐにある一定値に収束する。

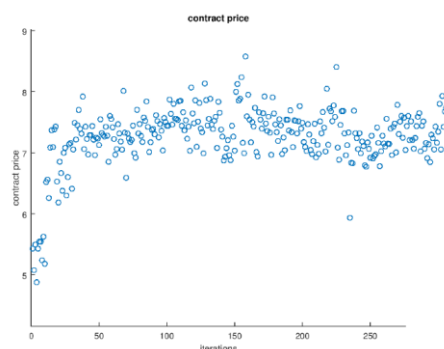


Fig 6 約定価格(Yen/KWh)の推移

ColBagging ではこの適切な価格をいち早く見つけられるものと考えられる。今後もこの点の検証を続ける。

(4) 強化学習エンジンを使った半自動実行[学会発表①]

本システムはクラウドソーシングに似ており、複数のワーカーに協力を仰ぐ必要がある。だがオリジナルの枠組みではワーカーは常にコンソールに張り付き操作を繰り返すことが強要された。これは決して小さくはない負担となる。そこで半自動実行ができるように、supervised actor critic を導入した (Fig 7 参照)。

ここで扱う問題は DinoGame (<https://chromedino.com/>) においてサボテンを何回飛び越えられるかを競うもので、本研究ではこれに提案手法をインプリメントして実行できるようにした。(Fig 9)。ゲームプレイヤーは、最初は手動で恐竜を操作してサボテンを乗り越える。このとき、システムはゲームプレイヤーを真似るように supervised learning が実行される。プレイヤーが途中で自動モードに切り替えると恐竜は自律的にサボテンを乗り越える方法を探すが、すでに実行した supervised learning の結果から学習を再開する。こうすることで事前知識無しの状態から強化学習を実行するよりは、はるかに高速に学習を収束させることができる。但しこちらの場合においても、被験者は事前知識ゼロの状態から実験に参加する必要がある。そのため、効果があるのか否かを明確にするために、PSOをつかって疑似的に実行して効果を確認した。具体的には PSO において、最大化する評価関数の「勾配方向に進む」ことを supervisor の行動とみなし、各粒子の移動手続きに、ある一定の割合で勾配方向に進む操作を加えた。するとこの supervisor を加えない場合に比べて supervisor を加えた方が、素早く全ての粒子が最適解に到達する (Fig 8)。

実際にこれをゲームプレイヤーに実験させた。フィードバック教示は統合された Actor に DinoGame を自動実行させたものを動画として作成し、それを各々のプレイヤーに見せて教示し

た。ただし比較のためフィードバック教示を受けないプレイヤーも配置し、これをコントロールプレイヤーとした。

その結果、

○ゲームプレイヤー自身は多くの場合その能力を高めることに成功したものの、コントロールプレイヤーと大きな差が無い。

○ゲームプレイヤーが教示の結果スコアを大きく伸ばしても、対応する Actor の見まね学習がうまく進まず、結局プレイヤーが不在の時には高いスコアは得られない。

の二点が明らかになっている。教示の方法が、単に動画を視聴するだけであることが不十分である可能性と、Actor として使用した学習器の追記学習能力が低い点が課題である[学会発表①]。今後リスクを基に学習を行う強化学習[学会発表④]を導入してゲームプレイヤーの修正を反映しやすくした強化学習エンジンの開発を継続する。

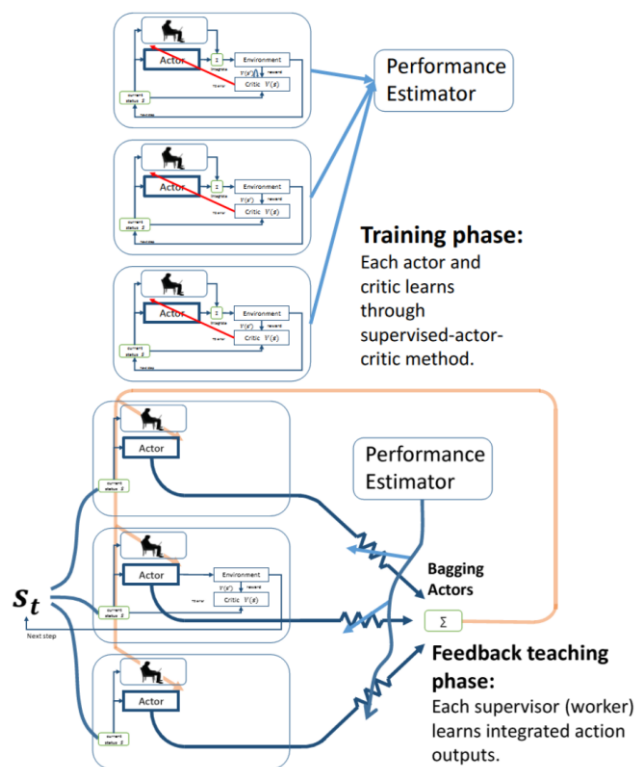


Fig 7 Supervised Actor Critic を使った ColBagging

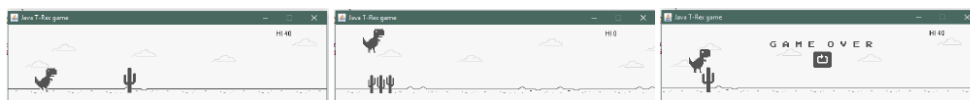


Fig 9 DinoGame

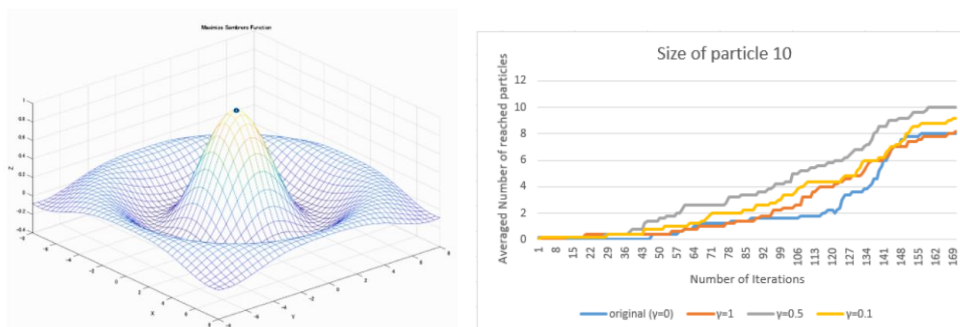


Fig 8 PSO を使った Supervised Actor Critic をベースとする ColBagging のシミュレーション。左の評価関数のピーク点に到達した粒子の個数と PSO 繰り返し回数との関係。教師データを 50%程度取り入れたときにもっとも多くの粒子がピーク点を発見している。

5. 主な発表論文等

[雑誌論文] (計1件)

① Takaya Ogiso, Koichiro Yamauchi, Norio Ishii, Yuri Suzuki . "Co-learning system for humans and machines using a weighted majority-based method ", International Journal of Hybrid Intelligent Systems, No.13, pp. 63--76, DOI: 10.3233/HIS-160225 (2016). 査読あり

[学会発表] (計7件)

- ① Ashwin Devanga, Koichiro Yamauchi. "Collaborative Learning of Human and Computer: Supervised Actor-Critic based Collaboration Scheme", in Maria De Marsico, Gabriella Sanniti di Baja, Ana Fred, editors, Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2019), SCITEPRESS Science and Technology Publications, Lda., pp. 794--801, February, (2019). (short paper) 査読あり
- ② 若原 涼, 山内康一郎. "人間とコンピュータによる共学習システムのモデル化と計算機シミュレーション", in 平成30年度 電気・電子・情報関係学会東海支部連合大会, vol.K1-5, September, (2018)., 査読無し
- ③ 小川僚也, 小木曾 寛太, 山田 知賢, 郭 宏文, 山内康一郎, 澤野弘明, 石井成郎, 鈴木裕利. "人間と機械学習による共学習を利用した卸電力取引への応用", in 平成29年度 電気・電子・情報関係学会東海支部連合大会, vol.Po2-18, September, (2017). 査読無し
- ④ Kazuyoshi Kato, Koichiro Yamauchi. "Risk Sensitive Reinforcement Learning Scheme Is Suitable for Learning on a Budget", in Akira Hirose, Seiichi Ozawa, Kenji Doya, Kazushi Ikeda, Minh Lee, Derong Liu, editors, Neural Information Processing --23rd International Conference, ICONIP 2016--, Springer-Verlag, vol.LNCS9949, pp. 202--210, (2016). 査読あり
- ⑤ 安宅 司, 岩瀬 貴則, 山内康一郎, 石井 成郎, 鈴木 裕利. "MGRNNを使った人間とコンピュータの共学習システムに関する研究", in 平成28年度 電気・電子・情報関係学会東海支部連合大会, E3-1, September, (2016)., , 査読無し
- ⑥ 加藤 和義, 山内 康一郎. "一般回帰ニューラルネットワークによる群知能統合", in 第25回日本神経回路学会全国大会, 日本神経回路学会, vol.P-46, pp. 36--37, 9月, (2015)., , 査読無し
- ⑦ 小木曾貴也, 山内康一郎, 石井成郎, 鈴木裕利. "アンサンブル法によるユーザスキル収集と教育への応用", in 日本デザイン学会誌 第62回研究発表大会概要集, pp. 150--151, June, (2015). 査読無し

6. 研究組織

研究分担者氏名：石井 成郎
ローマ字氏名：(ISHII Norio)
所属研究機関：一宮研伸大学看護学部
職名：准教授
研究者番号：80399237

研究分担者氏名：鈴木 裕利
ローマ字氏名：(SUZUKI Yuri)
所属研究機関：中部大学工学部
職名：教授
研究者番号：20340200

研究分担者氏名：澤野 弘明
ローマ字氏名：(SAWANO Hiroaki)
所属研究機関：愛知工業大学
職名：准教授
研究者番号：10609431

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。