

グラフ構造深層強化学習による電力系統制御

Reinforcement Learning with Graph Attention Networks and Its Applications to Power System Control

伊藤 秀将*

Hidemasa Ito

釜谷 幸男*

Yukio Kamatani

花井 克之*

Katsuyuki Hanai

1 はじめに

社会インフラシステムの中には、電力系統網、道路網、コンピュータ網、水道網などのようにその構造がグラフ構造で表現できるものがある。また、こういった社会インフラシステムの制御は、複数の要因を許容範囲に収めつつコストを最小化する最適化問題として考えることができる。時系列的に環境が変動する中、将来にわたる累積コストを最小化する最適制御を網羅的に探索する課題に対しては、強化学習が有効と考えられる。本論文では、グラフ構造をもった社会インフラシステムへの強化学習適用例として、電力系統網を最適制御するための強化学習手法を提案する。

2 問題設定

安定した送配電のためには、電力需要が時系列的に変化する中、設備の電圧値を一定範囲に収める必要がある。本論文では電圧値を安定化する制御手段の一つとして、変圧器のタップスイッチやシャントを切り替える方法を扱う。ただし、タップスイッチの切り替えはその回数によって点検規定が定められており、切り替え回数の増加はコスト要因になる [1]。そのため、設備の電圧値を制約範囲に収めつつも、将来に渡ってタップスイッチの切り替え回数を最小化する制御が求められる。これは将来報酬の最適化であり、強化学習の問題として解くことができる。

電力系統の制御を強化学習するに際して、システム構成変更を考慮することが重要である。通常、インフラ自体は運用期間が長く、その間に機器の保守、改良、置換が行われる。しかし、Q 関数表や全結合型のニューラルネットでは対象のシステム構成の変更を想定していないため、変更後の状態で再学習しなければならず、運用コストがかかる。

本論文では、システム変更においても再学習を要しない手法を提案する。

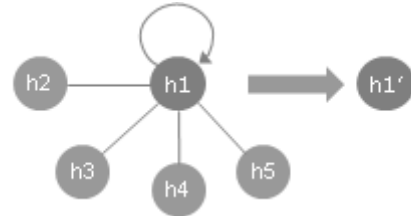


図 1: グラフアテンションネットワークの基本構造

3 関連技術

3.1 電力系統強化学習

従来の電力系統制御への強化学習適用の例として、IEEE case14 のタップスイッチ制御に Q 学習を適用した研究がある [2]。Q 関数の状態空間が固定であるため、設備の追加や削除を行えば再学習は必須である。深層強化学習によって電力系統制御を行う研究も存在するが、モデルは時系列認識用の CNN に全結合ニューラルネットを加えた構成であり、系統構造が変化すると再学習が必要になる [3]。

3.2 Graph Attention Network

グラフ構造システムを扱う Graph Convolutional Network(GCN) は、大きく二種類に分類される [4]。グラフフーリエ変換を利用する間接法と、グラフの隣接関係を直接記述する直接法である。そのうち今回は設備の構造変化に対応するため、GCN の中でも直接法の Graph Attention Network(GAT)[5] を採用している。図 1 と式 (1) を用いて処理概要を説明する。図 1 は、ノード 1 の特徴量 h_1 に隣接ノードの特徴量 h_2 から h_5 が足しこまれ、 h'_1 が算出される様子を表現している。式 (1) はあるノード i の特徴量 h_i と隣接するエッジの特徴量 h_j , $j \in N_i$ から特徴量 h'_i を求める式である。 $\langle h_i, h_j \rangle$ はベクトルを結合する処理を示す。 α_{ij} はノード i に対する接続エッジ j の重要度である。 σ は活性化関数である。

$$w_{ij} = \sigma(W \langle h_i, h_j \rangle) \quad \alpha_{ij} = \frac{e^{w_{ij}}}{\sum_{k \in N(i)} e^{w_{ik}}} \quad (1)$$

$$h'_i = \sum_j \alpha_{ij} W h_j \quad j \in N_i$$

*東芝デジタルソリューションズ株式会社

4 提案モデル

電力系統制御を対象とした強化学習モデルを提案する。システム構成変更への対応力の課題を解決するため、電力系統をグラフ構造として扱い、それを扱うニューラルネットワークモデルを定義する。こうすることで、システム構成変更への対応が可能となる。

電力系統をグラフとして次のようにモデル化する。電力系統に内包される設備のグラフ的構造は大きく三種類に分類される。一つ目は母線である。母線同士は後述する送電線や変圧器で接続されるほか、発電機やシャント等が接続しており、接続数は一定ではない。よって、母線はグラフ上のノードの役割を果たしている。二つ目は、送電線や変圧器である。これらは2つの母線をつなぐ構造を持っており、グラフではエッジの役割を持つ。最後は発電機やシャントである。これらは一つの母線に接続するため、ノードの属性情報として扱うことができる。

これを扱うモデルはグラフ構造を入力可能で、かつシステムの構造変化に耐えうるものでなければならない。そこで、GCN を利用した強化学習モデルを作成した (図 2)。モデル構造は大きく 5 つに分かれる。Graph Attention モジュール、Edge Convolution モジュール、時系列認識モジュール、状態価値推定モジュール、行動確率算出モジュールである。全体構造は図 2 に示した通りで、Graph Attention モジュールと Edge Convolution モジュールを交互に複数回かけて receptive field を広げた後、時系列認識モジュールを経て状態価値推定モジュールと行動確率算出モジュールに繋がる。

4.1 Graph Attention モジュール

GAT はノード (母線) とそれに接続するノードの特徴量から一層深いノード特徴量を入力するが、電力系統は母線に接続する送電線や変圧器の特徴量も扱いたい。そのため、母線とそれに接続する送電線や変圧器の特徴量を入力とし、一層深い母線特徴量を入力するモジュールを作成した (図 3)。

送電線や変圧器を始めとした多様な設備との接続を扱うため、設備の種類毎に個別の重み、アテンションパラメータを学習させている。式 (2) はあるノード i の特徴量 h_i と隣接するエッジの特徴量 h_j , $j \in N_i$ から特徴量 \hat{h}_i を求める式である。結合重み計算用パラメータ $W_{t(i), t(j)}^{g0}$ の数はノードとエッジの組み合わせ数にのみ依存し、エッジ特徴量の変換パラメータ $W_{t(j)}^{g1}$ の数はエッジ数にのみ依存するため、同様の設備が追加や削除されてシステム構造が変化したとしても、パラメータを共有することができる。なお、ここで t はノードやエッジの種類を参照する関数である。

$$w_{ij} = \sigma(W_{t(i), t(j)}^{g0} \langle h_i, h_j \rangle)$$

$$\alpha_{ij} = \frac{e^{w_{ij}}}{\sum_{k \in N(i)} e^{w_{ik}}} \quad (2)$$

$$\hat{h}_i = \sum_j \alpha_{ij} W_{t(j)}^{g1} h_j \quad j \in N_i$$

4.2 Edge Convolution モジュール

二つの母線を結ぶ送電線や変圧器の特徴量と、それに隣接する母線の特徴量を入力とし、一層深い送電線や変圧器の特徴量を入力するモジュールである (図 3)。送電線や変圧器に接続する母線の本数は常に二つであるため、入力は一層深いベクトルになる。このモジュールは式 3 のように、エッジとそれに隣接する母線を入力とした全結合ニューラルネットワークで実装される。このパラメータはエッジ種類と隣接するノード種類の組み合わせによって個別に学習される。

$$\hat{h}_i = \sigma(W_{t(i), t(j), t(k)}^e \langle h_i, h_j, h_k \rangle) \quad j, k \in N_i \quad (3)$$

4.3 時系列認識モジュール

環境から観測される状態は必ずしもマルコフ性を持つとは限らない。例えば、電力負荷は季節や時刻によって時系列的に変化していくが、観測は現在の電力負荷のみを入力する。そのため、LSTM を利用して時系列情報を記憶させ、内部状態にマルコフ性を持たせる (式 (4))。LSTM のパラメータはノード i の機種 $t(i)$ 毎に重みを共有する。内部状態は設備毎に記憶し、更新する。

$$\hat{h}_i = LSTM_{t(i)}(h_i) \quad (4)$$

4.4 状態価値推定モジュール

将来にわたる累積集計を推定するモジュールである。これは全結合ニューラルネットワークで実装されるものであり、そのノードが将来得る推定状態価値 $V(h_i)$ を出力する (式 (5))。ここで G_r は報酬が与えられるノード、エッジの集合である。今回は母線の電圧安定を目的とするため、母線の集合に相当する。システム全体に対するスカラの報酬値を得たい場合は、これらの報酬を総和ないしは平均する。

$$V(h_i) = MLP_{t(i)}(h_i) \quad i \in G_r \quad (5)$$

4.5 行動確率算出モジュール

全結合ニューラルネットワークによって、各アクションの選択確率を算出するモジュールである。式 (6) はノード i の行動確率分布 $\pi(a_i|h_i)$ を求める式である。ここで G_a

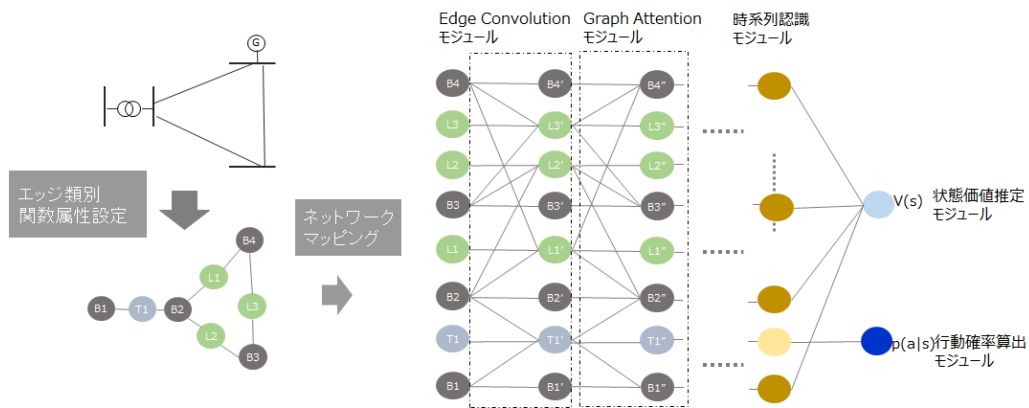


図 2: 電力システムのグラフ構造化とニューラルネットへのマッピング

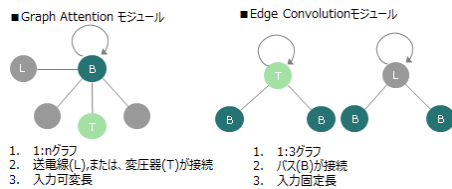


図 3: Graph Attention, Edge Convolution モジュールの基本構造

はアクションを行うノード、エッジの集合である。問題が電力システムのタップスイッチ制御及びシャント制御である場合は、変圧器を示すエッジとシャントの付属する母線のノードがそれにあたる。また、ここでは各アクションは特徴量を共有するが、それぞれ独立の分布からサンプルされるものとして扱っている。これは課題としていたシステム構成変更への対策であり、同種のアクションが増えても再学習の必要がない他、計算時間も線形になる。

$$\pi(a_i|h_i) = MLP_{t(i)}(h_i) \quad i \in G_a \quad (6)$$

5 学習アルゴリズム

強化学習アルゴリズムは価値ベースと方策ベースに二分されるが、今回は報酬を算出するグラフ構造が可変であり、報酬のスケールが不明であることから、方策ベースの手法を採用した。学習アルゴリズムには A3C を採用した [6]。A3C は Asynchronous Advantage Actor-Critic の略称であり、方策ベースの強化学習手法である。分散非同期でロールアウトとパラメータ更新を行う、また、連続値の行動を離散化することなく陽に扱うことができる。

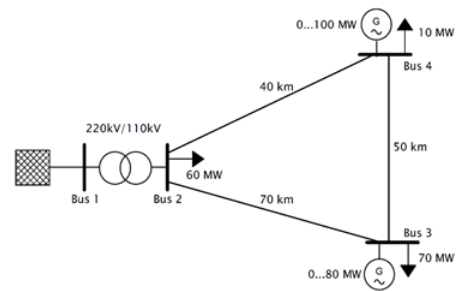


図 4: 4 母線電力系統 [7]

6 予備実験

深層強化学習による電力系統制御の挙動を検証するため、提案モデルの検証を行う前に簡易な系統制御を行った。実験に用いたのは図 4 に示した。潮流シミュレータ pandapower [7] のテスト系統で、4 母線構成であり、母線 1 と母線 2 の間に変圧器が存在し、そのタップスイッチが制御対象となる。

最小のタップスイッチ制御で電圧を基準値の範囲に収めることを目標とする。

6.1 実験設定

6.1.1 シミュレータ

系統の潮流シミュレーションには pandapower を利用する。負荷値はステップ t に対して $300,000 * \sin(\pi * t / 9) V$ で計算される。エピソードの最大ステップ数は 9 である。

6.1.2 報酬関数

報酬関数は二項から構成される。一つは母線 2 番の電圧安定性コストであり、もう一項はタップスイッチを動かすことで生じる設備の劣化コストである。これらは

トレードオフの関係にある。電圧安定性コストは母線に設定された基準電圧からの相対値で算出する。1.0pu から離れるほどコストが上がるほか、0.98pu から 1.02pu の範囲を超えると大きなペナルティを与える。

6.1.3 制御項目

母線1と母線2を繋ぐ変圧器のタップスイッチを制御する。タップスイッチは19段階に制御できる。

6.2 モデル

全結合ニューラルネット一層とLSTM一層で構成した。出力は-9から+9の19段階の離散値で、行動の事後確率はsoftmax関数を通して出力される。

6.3 実験結果

1,000エピソード学習し、結果を図5、図6、図7に示した。初期値であり制御対象外である時刻0は母線の電圧比が0.98puを下回っているが、以後は0.98puから1.02pu内に収まっている。消費電力の上昇に伴ってタップスイッチを下げ、下降に伴ってタップスイッチを上げており、タップスイッチ制御合計3回で対処できている。深層強化学習によって、目標通り母線電圧を基準に収める最小のタップスイッチ制御が可能であることを確認した。

7 実験

電力系統ネットワークIEEE case14を題材に制御実験を行った[8]。IEEE case14(図8)は14母線構成の電力系統であり、シャント一器と変圧器五器を持つ。変圧器五器のうちタップスイッチを持つのは三器である。

7.1 実験設定

7.1.1 シミュレータ

系統の潮流シミュレーションにはpandapowerを用いた。IEEE case14の電力消費には時系列性がないが、実際の電力系統は時系列的に変化するものである。そのため、東京電力の需要曲線データ[9]を、IEEE case14の各母線の消費電力基準値でスケールし、母線の消費電力とした。なお、上流系と下流系の需要バランスの変化を表現するため、時系列変化を起こすのは9-14番の下位ノードのみとした。また、2016年の電力消費をトレーニング、2017年の電力消費をテストに用いた。

7.1.2 報酬関数

報酬関数は二項から構成される。一項は各母線の電圧安定性コストであり、もう一項はタップスイッチを動かすことで生じる、設備の劣化コストである。これら

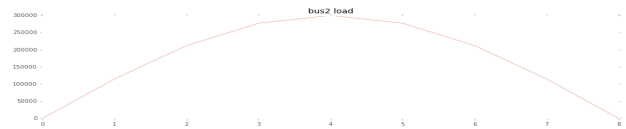


図5: 4母線電力消費

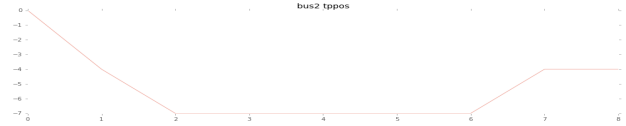


図6: 4母線タップスイッチ

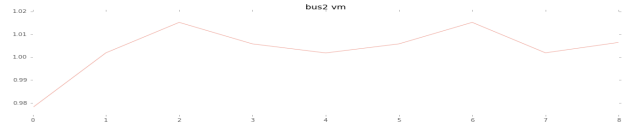


図7: 4母線電圧

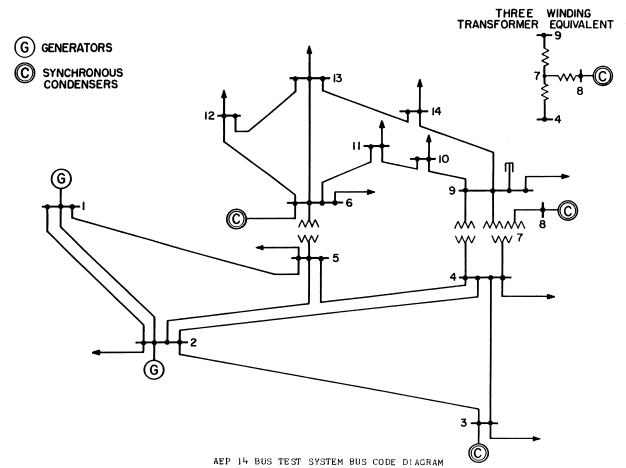


図8: IEEE case14 電力系統 [8]

は6.1.2と同様の構成である。ただし、母線電圧の下限は0.94vmpu、上限は1.06vmpuに設定している。

7.1.3 制御項目

変圧器のうちタップスイッチを持つ三箇所とシャントを制御項目とした。タップスイッチは7段階に別れており、シャントもまた7段階である。そのため、行動空間は $7^4 = 2401$ 通りになる。

7.2 全結合ニューラルネットとの比較実験

提案モデルの有効性を確認するため、全結合ニューラルネット+LSTMを用いたモデルとの比較を行った。全結合ニューラルネット+LSTMの平均報酬は339.38、提案モデルの平均報酬は349.98となった。Wilcoxonの符号順位検定と比較したところ、提案モデルは全結合ニューラルネットを使ったモデルに対し、それを上回る精度が確認された。

提案方式による精度向上の要因は二つ挙げられる。一

つは、同種の設備に対しパラメータを共通化したことにより、電圧等の認識が容易になったことである。もう一つは、母線電圧は近傍の変圧器や発電機の影響を受けやすいため、グラフ上の近傍関係が反映されたモデルの方が有利になったと考えられる。

7.3 システム構成変更への対応力の確認実験

課題としていたシステム構成変更に対し、提案モデルの対応力を検証する。特定の系統集合で学習した提案モデルが、それに近い未学習の系統集合を制御し得るかを実験し、確認する。

7.3.1 部分系統の作成

IEEE case14 構成を基準とし、そのうちの一母線とそれに接続する設備を撤去した部分グラフを作成する。こうして作成された 15 種類の部分グラフ集合のうち、負荷を初期値に設定した際に潮流計算が収束しなかった 2 種類を取り除き、13 種類のグラフ集合を作成する (表 1)。この集合を用いて、3fold 検定を行う。

表 1: 系統グループ

学習/テストグループ	撤去した母線番号
A	無切除, 4, 10, 13
B	2, 5, 8, 11, 14
C	3, 6, 9, 12

7.3.2 実験結果

図 9, 図 10 は母線 8 番を切除した系統の同じ一日の電圧変化とスイッチ位置を示している。母線 8 番は発電機を含む母線であり、それが失われたことで 8 番の近傍に位置する母線 7 番, 9 番の電圧が低くなっている。時刻 9, 10 では母線 7 番, 9 番の電圧が下限値の 0.94pu を下回る。時刻 11 でモデルがシャントのスイッチを切り替えて電圧を上昇させ、以降は一定範囲に留めている。図 11 図 12 は母線 3 番を切除した系統の電圧変化とスイッチ位置である。図 9 の系統と同様のモデル、同様の需要曲線を用いているが、その振る舞いは異なる。このように、系統構造の違いによって効率的な挙動をとるモデルが得られたことから、学習は正常に行われたと考えられる。

ここでは 1,000 回実行した平均報酬で比較する。実験結果を表 2 に示す。まず、いずれの組み合わせで学習したモデルも、ランダム行動モデルに比べ高い報酬が得られていることがわかる。表 2 の縦軸に着目すると、学習グループの違いによる報酬の差異が小さいことが

わかる。よって、提案モデルは設備増設や削除に耐えうる系統制御を獲得できたと言える。

7.3.3 考察

平均報酬が総じて高い B グループの系統は、母線 2 番, 5 番, 8 番, 11 番, 14 番を切除した 5 系統であった。発電機付きの母線が切除されたのは 2 系統であり、うち母線 8 番は末端に位置する母線である。母線 8 番 11 番 14 番は送電線の接続数が 2 以下であり、切除しても系統全体への影響は小さかったものと思われる。一方、平均報酬の低い C グループは母線 3 番, 6 番, 9 番, 12 番を切除した 4 系統であり、うち 2 系統が発電機を持っている。また、母線 6 番及び 9 番は、系統の上流と下流を繋ぐハブの役割を果たしている。一方のハブが失われた状態で電力需要を満たすには、他方のハブに負荷をかける他ない。また、制御対象の変圧器のうち 1 台は 6 番, 2 台は 9 番に接続しているため、それらの母線が切除された系統では、タップスイッチを切り替えても母線電圧を制御できない。これらの要因により、結果として制御が困難になったものと考えられる。図 14 の学習曲線を参照すると、学習対象に C グループの含まれるモデルは含まれないモデルに比べて学習初期の報酬値が低い傾向にあることが読み取れる。

提案法とは逆に、ランダムモデルは C グループのみ平均報酬が高い。これは、変圧器が系統から切り離されることで、制御失敗のリスクが下がったからであると考えられる。

なお、事前検証で得られた平均報酬に比べ、いずれの組み合わせの報酬も低くなっている。これは事前検証が母線切除のない IEEE case 14 でのみ行われたことに起因する。

表 2: 学習グループと平均報酬

train group	test group		
	A	B	C
A & B	271.42	279.26	269.53
A & C	273.79	281.73	264.74
C & B	272.92	278.52	263.17
random action	116.23	114.91	126.27

8 おわりに

可変性のある電力系統を最適制御する強化学習モデルを提案した。提案モデルは全結合型ニューラルネットワークと LSTM を組み合わせたモデルよりも同条件でより高い報酬を得られたほか、設備構造の変更にも耐えうるという結果が得られた。今後はさらに複雑な制御や多様な系統構造への対応を目標とする。

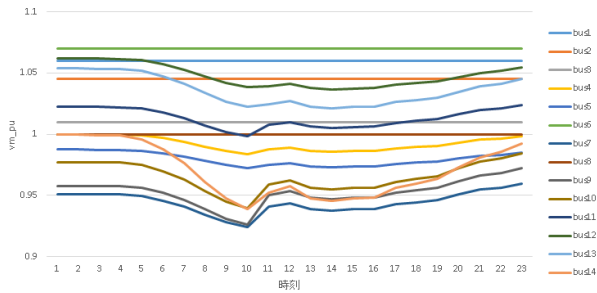


図 9: IEEE case 14 母線 8 切除系統 母線電圧

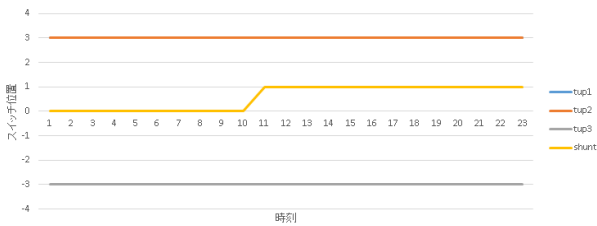


図 10: IEEE case 14 母線 8 切除系統 タップスイッチ

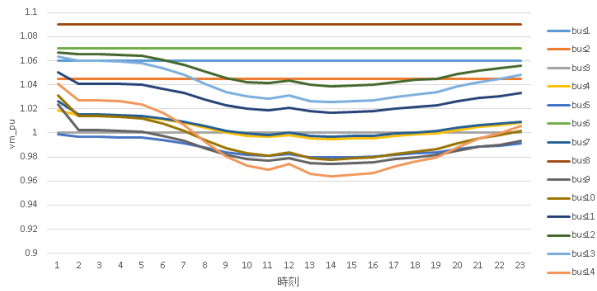


図 11: IEEE case 14 母線 3 切除系統 母線電圧



図 12: IEEE case 14 母線 3 切除系統 タップスイッチ

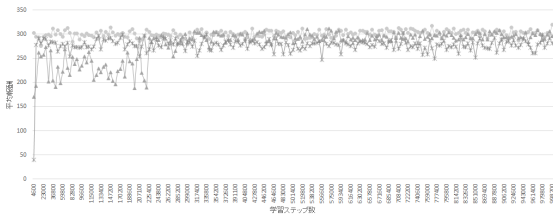


図 13: 各モデルの学習曲線. 報酬値は学習時と同じ系統グループを用いて算出

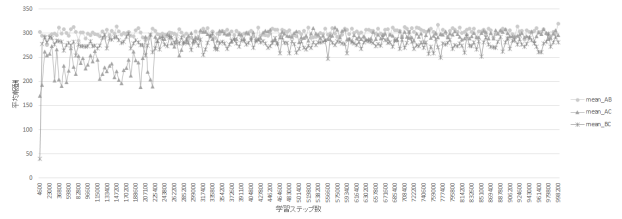


図 14: 各モデルの学習曲線. 報酬値は学習時と同じ系統グループを用いて算出

参考文献

- [1] 瀧口幸延, 谷口嘉信, 蝦名定信, ”油入変圧器用真空バルブ式負荷時タップ切換器”, 東芝レビュー (Toshiba Review) 56号 2001
- [2] J. G. Vlachogiannis, N. D. Hatziaargyriou, ”Reinforcement Learning for Reactive Power Control”, IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 19, NO. 3, 2004
- [3] Francois-Lavet, V., Taralla, D., Ernst, D., and Fonteneau, R, Deep reinforcement learning solutions for energy microgrids management. In European Workshop on Reinforcement Learning.
- [4] Zhang, Z., Cui, P., and Zhu, W. ”Deep learning on graphs: A survey”, arXiv preprint arXiv:1812.04202, 2018.
- [5] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, Yoshua Bengio, ”Graph Attention Networks”, ICLR 2018
- [6] V. Mnih, ”Asynchronous Methods for Deep Reinforcement Learning”, Proceedings of ICML, 2016.
- [7] pandapower – Convenient Power System Modeling and Analysis based on PYPOWER and pandas –, technical report, Fraunhofer IWES/U. of Kassel, Version1.4. 0, 2017
- [8] ”IEEE 14 Bus Power Flow Test Case”, Power Systems Test Case Archive, Electrical Engineering, U. of Washington, https://labs.ece.uw.edu/pstca/pf14/pg_tca14bus.htm
- [9] 「数表でみる東京電力」, 東京電力ホームページ, http://www.tepco.co.jp/about/fact_database/