

深層学習を用いて次世代無線通信システムにおけるダイナミックスペクトラムアクセスの研究

代表研究者 李 傲寒 慶應義塾大学大学院理工学研究科 後期博士課程 2 年
共同研究者 大槻 知明 慶應義塾大学大学院理工学研究科 教授

1 はじめに

無線デバイスの普及とアプリケーションの急増に伴い、無線トラフィック量が大幅に増加する [1]. そのため、ユーザの通信品質 (QoS: Quality of Service) が低下する可能性がある [2]. また、ネットワーク環境もより複雑になる傾向がある [3]. 更に、無線周波数リソース不足が問題になっている [4]. 多数の無線デバイスに低遅延、高スペクトラム効率な通信品質を保証するのは大きな課題である [5]. 次世代無線通信システムの発展を妨げる大きな要因として、無線通信に使用しやすい周波数の資源不足が考えられる [6]. その解決策として、無線周波数のより効率的な利用が求められており、新たな周波数帯域を開拓、および電波環境や需要に応じて周波数を適応的に割り当てるダイナミックな周波数割り当て技術に注目が集まっている [7, 8]. その中で、ダイナミックスペクトラムアクセスを実現するため、無線デバイスはスペクトラムセンシングで周波数利用状況を検知して、適切に周波数利用状況を把握することで、他の無線デバイスに干渉を与えることなく、より効率的な周波数利用が可能となる [9-12]. しかし、チャンネルの状態を正確にセンシングすることには、センシング時間とエネルギーがかかるため、通信遅延とデバイスのエネルギー消費がより高い [13, 14]. また、他の無線デバイスに干渉を避けるチャンネル選択手法では、従来法には事前情報が必要であるが、現実世界では実現できない可能性がある [15-17]. 深層学習は多層のニューラルネットワークを用いて高い予測性能を発揮することと特徴を抽出することができる機械学習の 1 つであり、階層数をはじめとしたパラメータを次世代無線通信システムに適したものに最適化し、事前情報なくても、学習させることで、高効率なチャンネル割当と高精度なスペクトラム状態の予測が実現できると考えられる [18-20]. そこで、本研究では、将来のスペクトラム状態を予測することでスペクトラムセンシングを行う回数削減技術および事前情報なくても他の無線デバイスに干渉を避けるチャンネル選択技術を検討した.

2 深層学習を用いたスペクトラム状態の予測

2-1 従来研究と本研究の目的

ダイナミックスペクトラムアクセスシステムにおける無線デバイスはセンシングで周波数の利用状況を検知して、他の無線デバイスに干渉を与えることなく、より効率的な周波数を利用する [21, 22]. しかし、スペクトラムセンシングで時間とエネルギーがかかるため、通信遅延とエネルギー消費がより高い [23]. スペクトラム予測技術はスペクトラムセンシングで消費される時間とエネルギーを大幅に削減できるため、期待される技術である [24-26]. ダイナミックスペクトラムアクセスシステムにおけるスペクトラム状態の予測する従来法 [27, 28] では、無線デバイスは自分以外の無線デバイスの情報を知る必要があると想定されている. しかし、現実には他のデバイスの情報を知ることは困難である. そこで、本研究では、深層学習を用いてスペクトラム状態の予測課題に取り組んだ. 深層学習 [29] を用いることで、他のデバイスの情報知らなくても、スペクトラム状態の予測することが可能である. まず、再帰型ニューラルネットワーク (RNN: Recurrent Neural Network) の一種である (LSTM: Long Short-Term Memory) [30] を用いて次のタイムスロットのスペクトラム状態を予測した. 指数オン/オフモデルと隠れマルコフモデルの 2 種類のチャンネルを検討した.

2-2 LSTM を用いたチャンネル予測

(1) LSTM

LSTM は再帰型ニューラルネットワーク (RNN: recurrent neural network) の一種である [31]. RNN の学習において勾配消失は大きな問題である. そして、それを解決するには RNN レイヤのアーキテクチャを根本から変える必要がある. ここで登場するのが「ゲート付き RNN」である. このゲート付き RNN には多くの

アーキテクチャが提案されており、その代表格に LSTM がある。図 1 に、入力層と隠れ層、出力層 4 つの層で構成され、3 つの時間ステップ時間領域で展開された RNN の構造を示す。入力ベクトルは、1 つのタイムステップで 1 つのエレメントを RNN に入力される。この構造は時系列の概念で、フィードフォワードネットワークは固定の入力ベクトルを処理することを示す。RNN ネットワークでは、履歴入力情報をネットワーク内部状態に保存できる。それにより、現在時刻までのすべての利用可能な入力情報を利用可能である。理論的には、RNN は任意の長さの時系列の特性を学習できる。しかし、実験により、RNN ネットワークのパフォーマンスは、勾配消失または勾配爆発により制限される可能性があることが証明されている [32]。RNN ネットワークの勾配の問題を解決するために、LSTM ネットワークは、メモリユニットを導入することによって設計されている。

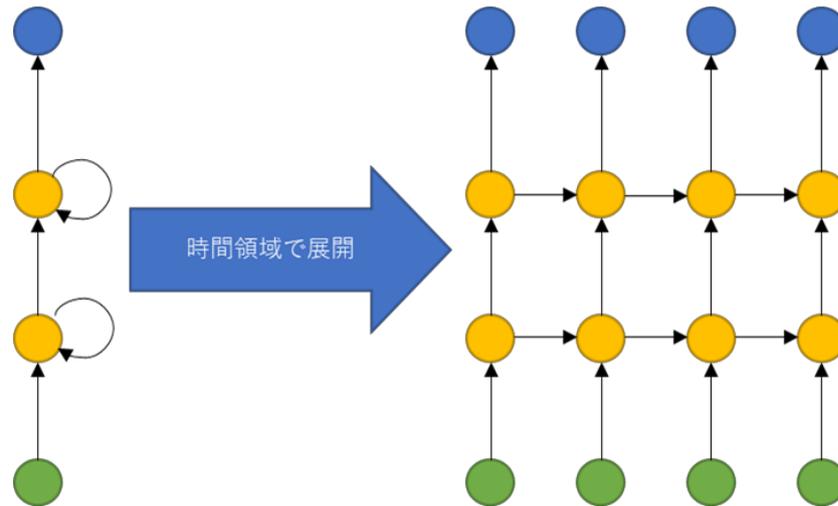


図 1：4 つの層で構成された RNN の構造

図 2 に示すように、メモリユニットはメモリセルと 3 つのゲートを含む。ゲートは、対応する実用的な機能に応じて名前が付けられる。メモリセルには、次のタイムステップで使用されるユニットの現在の状態を記憶する。ゲートは、関連情報の忘却と保存の比率を制御する。 i で示される入力ゲートは、メモリセルに流入する新しい情報の量を制御する。 f で示される忘却ゲートは、現在のメモリセルで破棄されるメモリセルの量を決定する。 o で示される出力ゲートは、メモリユニットの出力アクティベーションを計算し、さらにネットワークの残りの部分に流れ込む情報量を制御する。

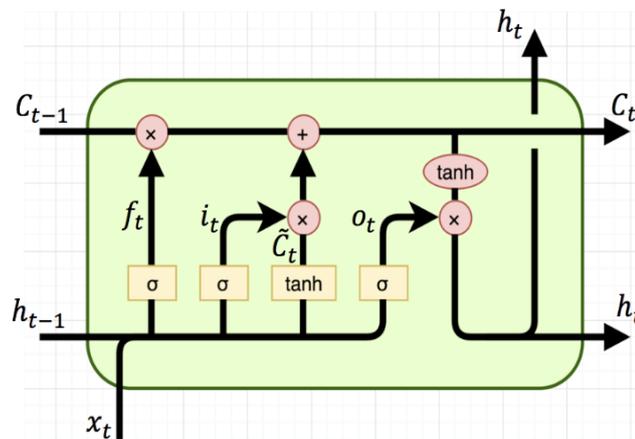


図 2：メモリユニット内の LSTM の構造

(2) LSTM を用いたチャンネル予測法

本研究では、LSTM を用いたダイナミックスペクトラムアクセスシステムにおけるチャンネル状態の予測を検討した。各チャンネルには2つの状態があると考え、つまり、アイドル状態とビジー状態。無線デバイスがチャンネルにアクセスしていない場合、チャンネルはアイドルである。その以外の場合、チャンネルはビジー状態である。2つのモデルに基づいてチャンネル予測を検討した。これらは、指数オン/オフモデル[33]と隠れマルコフモデル[34]である。指数オン/オフモデルでは、オン時間は指数分布からサンプリングされ、次のように表せる。

$$T_{OFF} \sim Exp(\theta) = \theta_0 e^{-\theta_0 t_{OFF}} \quad (1)$$

ここで、 θ_0 は指数オフ分布の平均値であり、平均オフ時間に対応する。図3に隠れマルコフモデルを示す。 $P_{00}, P_{01}, P_{10}, P_{11}$ はチャンネルの遷移確率であり、 P_{00} はチャンネルがアイドル状態からアイドル状態までの遷移確率であり、 P_{01} はチャンネルがアイドル状態からビジー状態までの遷移確率であり、 P_{10} はチャンネルがビジー状態からアイドル状態までの遷移確率であり、 P_{11} はチャンネルがビジー状態からビジー状態までの遷移確率である。

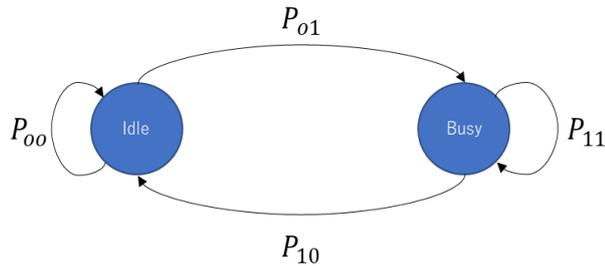


図3: 隠れマルコフモデル

本研究での LSTM を用いたチャンネル予測法では、LSTM に入力はセンシングしたチャンネルの状態であり、出力は予測するチャンネル状態である。

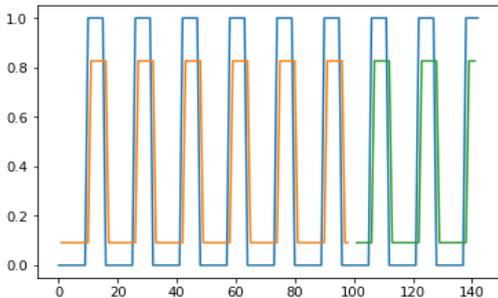
2-3 特性評価

(1) 指数オン/オフモデル

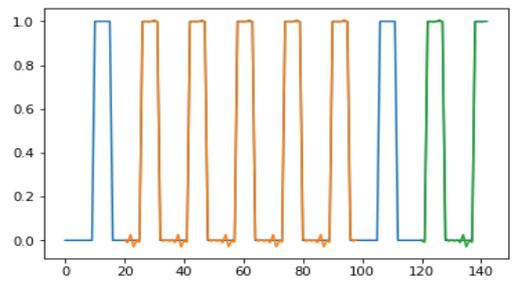
シミュレーションでは、指数オフ分布の平均値 θ_0 は0.1とする。訓練データは98個とし、テストデータは42と設定した。LSTMに入力チャンネル状態の数は1と20に設定した。シミュレーション結果は図3に示す。図4の横軸はチャンネル状態のデータの数であり、縦軸は平均二乗偏差 (RMSE: Root Mean Squared Error)である。青色で示すのは実際のチャンネル状態であり、オレンジ色で示すのは訓練プロセスでの予測結果であり、緑色で示すのはテストプロセスでの予測結果である。RMSEは次式で表すことができる。

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n ((Observed_t - Predicted_t)^2)} \quad (2)$$

ここで、 $Observed_t$ は時刻 t に実際のチャンネル状態であり、 $Predicted_t$ は予測値である。シミュレーション結果により、LSTMに入力チャンネル状態の数は1の時、RMSEは0.33となった。LSTMに入力チャンネル状態の数は20の時、RMSEは0.01となったことがわかる。つまり、LSTMに入力チャンネル状態の数は大きいほど、予測値は実際のチャンネル状態に近くなる。とにかくどちらの場合でも、指数モデルにおけるチャンネル状態が正しく予測できる。



(a) 入力チャンネル状態の数は1

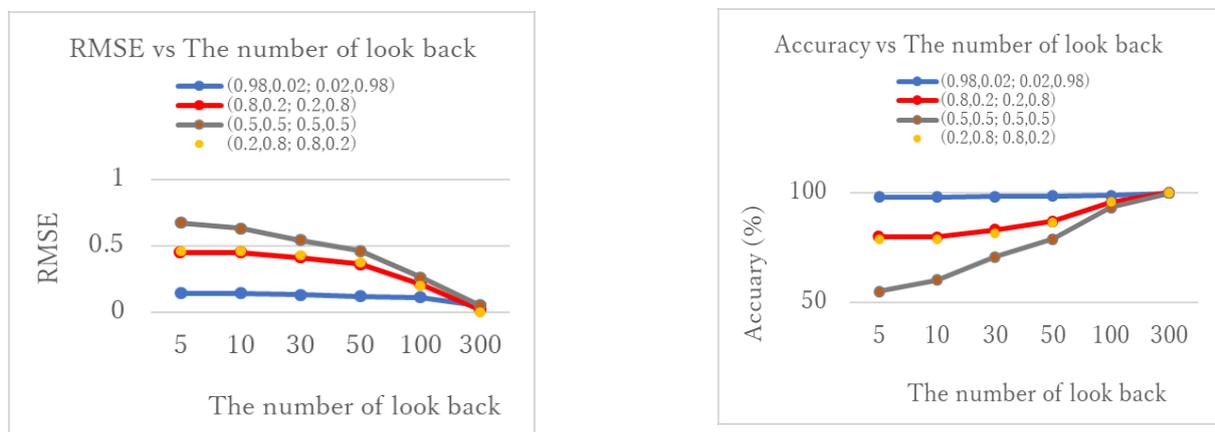


(b) 入力チャンネル状態の数は20

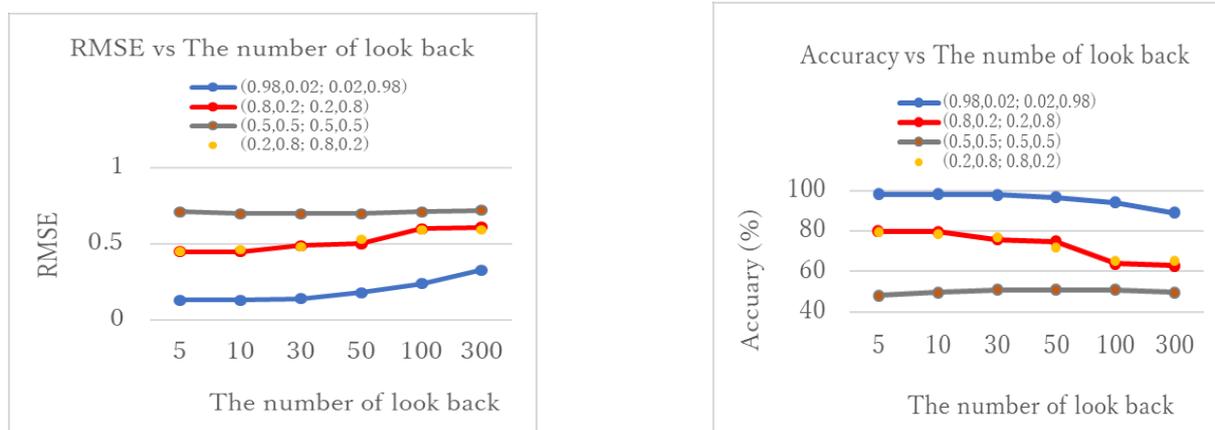
図4: 指数オン/オフモデルにおける予測結果,

(2) 隠れマルコフモデル

シミュレーションでは、隠れマルコフモデルでの遷移確率は4種類に設定される。それぞれは $P_{00}, P_{01}, P_{10}, P_{11} = (0.98, 0.02; 0.02, 0.98), (0.8, 0.2; 0.2, 0.8); (0.5, 0.5; 0.5, 0.5); (0.2, 0.8; 0.8, 0.2)$ である。チャンネル状態を予測するために使用されるチャンネル状態データは5から300まで変化します。予測値が0.5以上の場合、チャンネルはビジーとする。予測値が0.5未満の場合、チャンネルはアイドルとする。図5に隠れマルコフモデルにおける予測結果を示す。図5での「The number of look back」はチャンネル状態を予測するために使用されるチャンネル状態データの数である。シミュレーション結果により、訓練データを使用して予測したチャンネル状態の正確率はlook backの回数が増えると高くなることが分かる。しかし、テストデータを使用して予測したチャンネル状態の正確率はlook backの回数が増えると低くなることが分かる。また、RMSEと予測の正確率はトレーニングデータとテストデータの両方を使用したチャンネル状態の遷移確率に関連があることが分かる。チャンネル状態は不変動に傾向が高いほど、高い予測精度が得られる。チャンネル状態はよく変動に傾向が高い場合、予測精度が低いため、高効率なスペクトラム割当法の検討する必要がある。そこで、深層学習を用いたチャンネル割り当て法を検討していく。



(a) 訓練データを用いた予測結果



(b) テストデータを用いた予測結果

図5： 隠れマルコフモデルにおける予測結果

3 深層学習を用いたチャンネル割り当て

3-1 従来研究と本研究の目的

従来法 [35, 36] では、リソース割り当てを組合せ最適化問題として定式化している。しかし、従来法では、様々な制約条件したで多くの選択肢の中から、ある指標を最も良くする変数の値を求めること困難である。準最適解を求めるため、従来法では定式化された最適化問題をいくつかの準最適化問題に

分解し、それらを別々に解決するので、非効率的で複雑である。また、従来法には事前情報が必要であるが、現実世界では実現できない可能性がある。リソース割り当て問題は、マルコフ決定過程(MDP: Markov Decision Process)としてモデル化できる。ここで、強化学習(RL: Reinforcement Learning)の1つであるQ学習(QL: Q Learning)を用いて、事前情報を必要とせずにMDP問題を効率的に解決可能性がある。しかし、Q学習は状態空間が小さい問題にのみ適用できる[37]。システム規模が大きい場合、システムの状態および行動集合も非常に大きくなり、Q値を保存および更新するために必要な記憶容量が膨大になる。また、必要なQテーブルを作成するために各状態を探索するのに必要な時間は長い。したがって、入力として提供される状態情報を行動集合の各行動に対応するQ値にマッピングするモデルを開発する必要がある。ディープニューラルネットワーク(DNN: Deep Neural Network)は入力状態情報を全ての可能な行動のQ値にマッピングする方法を学習できるため、関数近似として重要である。Q学習とDNNを組み合わせると、深層強化学習(DQL: Deep QL)につながる。そこで、本研究では、ダイナミックスペクトラムアクセスにおける低計算量でシステム容量が大きい通信を実現するため、深層強化学習に基づくリソース割り当て法(DQLRA: Deep Q Learning based Resource Allocation)を検討した。DQLRAでは、無線デバイスが利用する周波数、送信電力、エネルギーハーベスティング時間を自動的に割り当てられる。本研究の主な内容は次のように要約される。(1) 計算量が低いDQLRAを提案した。DQLRAは、時変チャネル状態を考慮し、事前情報を必要とせずに無線デバイスに無線リソースを自動的に割り当て可能である。また、DQLRAはIoTデバイスの数が多い場合も利用可能である。(2) DQLRAは従来法(ECAA: Efficient Channel Allocation Algorithm)[15]と比較して計算量を大幅に削減可能であり、ランダム割り当て法と比較してシステム容量を大幅に改善できる。

3-2 DQLを用いたチャネル割り当て法

(1) システムモデルと問題の定式化

本研究では、図6に示すように、一つのゲートウェイ(GW: GateWay)と N 個のIoTデバイスから成るエネルギーハーベスティングIoTシステムについて上り回線を考える。IoTデバイスは、GWの通信範囲の周囲にランダムに配置される。IoTネットワークは同期しており、各IoTデバイスには容量上限がある充電式バッテリーが装備されると仮定する。充電式バッテリーでIoTデバイスが太陽光、振動、廃熱等のエネルギーから充電できる。GWの帯域幅を B_T とし、 B_T は $\mathbf{S} = \{S_1, S_2, \dots, S_m, \dots, S_M\}$ で示される M ブロックに分割される。各ブロックの帯域幅を B とする。本稿では、長距離、低電力通信向けに設計されたLoRAWAN(Long-RANGE Wide Area Network)ネットワークを考える。LoRAWANでは、各チャネルに拡散率(SF: Spreading Factors)が異なる最大 D 個IoTデバイスが同時にアクセスできる。つまり、異なるSFを使用している限り、ネットワーク内で複数のフレームを同時に交換できる[22]。

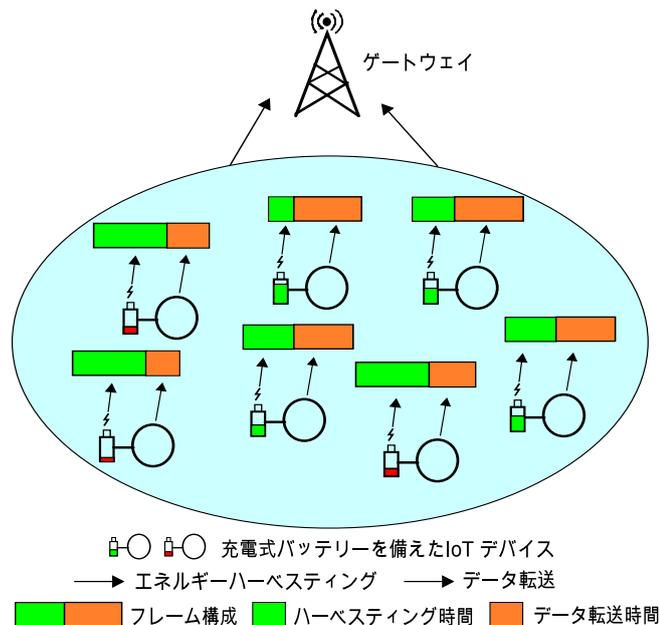


図6: システムモデル

図7に示すように、IoTネットワークでのデータ通信は時分割多重化方式に基づくとする。各フレーム長を F とし、 $\mathbf{K} = \{1, 2, \dots, k, \dots, K\}$ で示される K 個タイムスロットで構成される。IoTデバイスは、各タイムスロットでエネルギーハーベスティングかデータ転送をする。各タイムスロットの最初に、GWは通信範囲の各デバイスに対して、チャンネル状態、エネルギーハーベスティング状態およびIoTデバイスの残りエネルギーに基づいて、各タイムスロットでエネルギーハーベスティング、またはデータ転送を実行するか決定する。

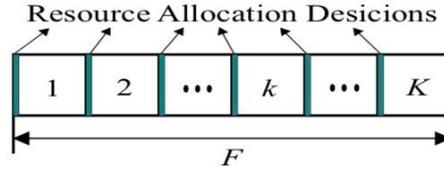


図7: 各エネルギーハーベスティングIoTデバイスのフレーム構造

フレーム j で、IoTデバイス n がデータ転送するタイムスロット数を $T_{n,j}$ と表し、エネルギーハーベスティングするタイムスロット数を $E_{n,j}$ と表す。フレーム j の k 番目のタイムスロットに、IoTデバイス n の充電されたエネルギー及びエネルギー消費を、それぞれ $P_{n,j}^H(k)$ 及び $P_{n,j}^C(k)$ と表す。

フレーム j の k 番目のタイムスロットがデータ転送に使用される場合、 $P_{n,j}^H(k)$ は0になる。同様に、

フレーム j の k 番目のタイムスロットがエネルギーハーベスティングに使用される場合、 $P_{n,j}^C(k)$ は0になる。IoTデバイスのバッテリー容量上限を P_{max}^A とすると、フレーム j の $(k+1)$ 番目のタイムスロットの最初に、IoTデバイス n の利用可能な電力は

$$P_{n,j}^A(k+1) = \min(P_{n,j}^A(k) - P_{n,j}^C(k) + P_{n,j}^H(k), P_{max}^A) \quad (3)$$

フレーム j の k 番目のタイムスロットで、帯域ブロック S_m がIoTデバイス n に割り当てられているかどうかを $\alpha_{m,n,j}(k)$ と表す。

$$\alpha_{m,n,j}(k) = \begin{cases} 1, & S_m \text{ が IoT デバイス } n \text{ に割り当てられる場合} \\ 0, & \text{その他} \end{cases} \quad (4)$$

フレーム j の k 番目のタイムスロットにおけるIoTデバイス n の容量は、

$$R_{n,j}(k) = \sum_{m=1}^M \alpha_{m,n,j}(k) \cdot B \cdot \log_2(1 + \gamma_{n,j}(k)) \quad (5)$$

になる。ここで、 $\gamma_{n,j}(k)$ はフレーム j の k 番目のタイムスロットにおけるIoTデバイス n の上り信号対雑音比 (SNR: Signal-to-Noise Ratio) を表し、次式で与えられる。

$$\gamma_{n,j}(k) = \frac{P_{n,j}(k) \cdot g_{n,j}(k)}{\sigma^2} \quad (6)$$

ここで、 $P_{n,j}(k)$ はフレーム j の k 番目のタイムスロットにおけるIoTデバイス n の送信電力を表し、 $g_{n,j}(k)$ 及び σ^2 は、それぞれフレーム j の k 番目のタイムスロットにおけるIoTデバイス n がアクセスするチャンネルの通信路利得と加法性白色ガウス雑音 (AWGN: Additive White Gaussian Noise) 電力を表す。フレーム j におけるIoTデバイス n の累積容量は次式で与えられる。

$$R_{n,j} = \sum_{k=1}^{T_{n,j}} R_{n,j}(k) \quad (7)$$

フレーム j におけるシステム累積容量は次式で与えられる。

$$R_j = \sum_{n=1}^N R_{n,j} \quad (8)$$

本研究の目標は、エネルギーハーベスティングIoTにおけるリソース割り当て法で累積容量を最大化することである。最適化問題は次のように定式化される。

$$(P1) \quad \max_{\alpha_{m,n,j}(k), P_{n,j}(k), T_{n,j}} R_j \quad (9a)$$

$$\text{s. t. } C_1: 0 \leq P_{n,j}(k) \leq P_{n,j}^A(k) \quad (9b)$$

$$C_2: T_{n,j} \leq F \quad (9c)$$

$$C3: T_{n,j} + E_{n,j} = F \quad (9d)$$

$$C4: \sum_n \alpha_{m,n,j}(k) \leq D \quad \forall m \quad (9e)$$

$$C5: \sum_m \alpha_{m,n,j}(k) \leq 1 \quad \forall n \quad (9f)$$

ここで、 C_1 は各エネルギーハーベスティング IoT デバイスの送信電力がどのタイムスロットでも残り電力を超えられないことを示す。 C_2 と $C3$ は各フレームがエネルギーハーベスティングとデータ転送のために2つの期間に分割されることを示す。 $C4$ は最大 D 個 IoT デバイスが同時に同一チャンネルにアクセスできることを示す。 $C5$ は各 IoT デバイスに同時に割り当てられるチャンネルが1つのみに制限される。最適化問題 $P1$ は従来の最適化法で解ける可能性があるが、特に大規模な IoT システムでは計算量が非常に大きくなる。また、最適化に必要な事前情報を取得するのは困難な場合がよくある。そこで、計算量を削減し、事前情報を不要とする DQLRA を設計する。

(2) DQL

DQL 法は深層強化学習 (DRL : Deep RL) 法の1つである。DRL 法の一般的な構造を図8に示す。各エポック t で、エージェントは方策 π を用いて、DNN によって取得された Q 値に基づいて行動を決定する。DNN は Q 関数の関数近似として使用されており、ベクトルの形で状態情報を入力し、可能な全ての行動の Q 値にマッピングすることを学習する。選択した行動を実行することにより、エージェントが報酬を取得し、環境が新しい状態に変わる。次に、DNN が、取得した報酬、現在の状態に対応する方策 π を用いて選択した行動の Q 値、および次の状態に対応する最大 Q 値に基づいて更新される。図9に DNN モデルの構造を示す。図9では、エポック t で可能な全ての行動の Q 値、つまり、 $Q(S_t, A_1), \dots, Q(S_t, A_l), \dots, Q(S_t, A_L)$ が、エポック t の状態情報 S_t から DNN によってマッピングできる。

ここで、 $A = \{A_1, \dots, A_l, \dots, A_L\}$ は行動集合である。DQL 法の主な目標は、 $Q(s, a; w_t)$ と y_t^{tar} との間の距離を最小化することである。ここで、 y_t^{tar} は TD (Temporal Difference) target と呼ばれ、 $Q(s, a; w_t)$ はエポック t で状態 s と行動 a のペアの Q 値を示す。DQL 法の目標は次式で与えられる y_t^{tar} と $Q(s, a; w_t)$ の二乗損失関数を最小化することである。

$$L_t(w_t) = \mathbb{E}_{s,a,r,s'} [(y_t^{tar} - Q(s, a; w_t))^2]. \quad (10)$$

DQL 法では、

$$y_t^{tar} = [r + \gamma \max_{a'} Q(s', a'; w_t^-) | S_t = s, A(t) = a]. \quad (11)$$

ここで、 S_t と $A(t)$ は、それぞれ現在の状態と行動を示す。 s' と a' は、それぞれ次のエポックでの状態と行動を示す。 r は即時報酬を示す。 γ は将来の報酬の割引率を示す。DQN では、 $Q(s', a'; w_t^-)$ と $Q(s, a; w_t)$ は target network および evaluation network と呼ばれる2つの異なるニューラルネットワークによって別々に推定される。

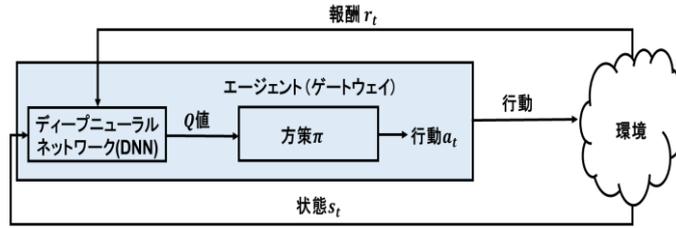


図 8 : DRL 法の構造

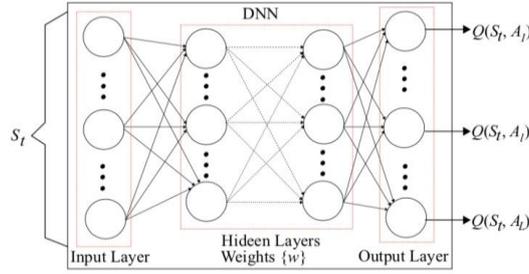


図 9 : DNN モデルの構造

(3) DQLRA

DQLRA では、GW はエージェントであり、エネルギーハーベスティング IoT ネットワークに関連する全てのものを環境とする。行動は ϵ -greedy アルゴリズムを用いて、DNN によって取得された Q 値に基づいて、GW によって選択する。選択した行動を実行することにより、環境が新しい状態に変わり、報酬を取得できる。DQLRA では、IoT デバイス n がフレーム j の k 番目のタイムスロット中の状態は、

$S_k = \{G_j(k), P_j^H(k), P_{n,j}^A(k)\}$ として定義される。ここで、 $G_j(k)$ と $P_j^H(k)$ は、それぞれフレーム j

の k 番目のタイムスロット中のチャンネル状態集合とエネルギーハーベスティング状態集合を示す。各タイムスロット中、チャンネル状態とエネルギーハーベスティング状態は不変だと想定する。

$G_j(k) = \{g_j^1(k), g_j^2(k), \dots, g_j^m(k), \dots, g_j^M(k)\}$ 。ここで、 $g_j^m(k)$ はフレーム j の k 番目のタイムスロット

中でチャンネル m の通信路利得を示す。 $P_j^H(k) = \{P_j^{H,1}(k), P_j^{H,2}(k), \dots, P_j^{H,m}(k), \dots, P_j^{H,M}(k)\}$ 。こ

こで、 $P_j^{H,m}(k)$ は、フレーム j の k 番目のタイムスロット中チャンネル m のエネルギーハーベスティング状態集合を示し、各タイムスロット中にハーベスティングできるエネルギーとして定義される。DQLRA

の行動集合は、チャンネル集合、送信電力レベル集合、エネルギーハーベスティングとデータ転送のインジケータで構成される。エネルギーハーベスティングとデータ転送のインジケータは、エネルギーハーベスティングまたデータ転送のいずれかの行動を示すために使用される。DQLRA では、行動集合

$A = \{A_1, A_2, \dots, A_{2MP_L}\}$ 。ここで、 P_L は送信電力レベル数を示す。行動を実行するための方策が ϵ -

greedy アルゴリズムであり、次式で与えられる。

$$A(k) = \begin{cases} \arg \max_{A(k) \in A} Q(S_k, A(k)), & \text{確率}(1-\epsilon) \text{で} \\ \text{random}(A(k)), & \text{確率}\epsilon \text{で} \end{cases} \quad (12)$$

ϵ -greedy アルゴリズムとは、確率 ϵ でランダムに行動、それ以外の確率 $(1-\epsilon)$ で最も期待値 (Q 値) の高い行動を選択する法である。DQLRA では、DNN の更新は RMSProp (Root Mean Square Propagation)

最適化アルゴリズム[23]を用いて行う。DNN をトレーニングするための損失関数は

$$Loss(w_k) = \mathbb{E} \left[\sum_{S_k, A(k) \in D} (y_k^{tar} - y_k^{eva})^2 \right], \quad (13)$$

となる。ここで、 D は状態と行動ペアを保存するデータセットである。DQLRA では、 y_k^{tar} と y_k^{eva} はそれぞれ式(9)と $Q(s, a; w_k)$ として定義される。フレーム j の k 番目のタイムスロットの報酬関数 r は、次のように定義される。

$$r_j(k) = \begin{cases} R_{n,j}(k), & \text{if } \sum_n \alpha_{m,n,j}(k) \leq D \text{ and } P_{n,j}(k) \leq P_{n,j}^A(k). \\ 0, & \text{Otherwise.} \end{cases} \quad (14)$$

報酬関数は選択した行動に関する。式(12)を示すように、選択された行動には2つのケースがある。

最大 D IoT デバイスが割り当てられたチャネルを占有し、割り当てられた送信電力が残り電力より大きくない場合、IoT デバイス n の報酬はフレーム j の k 番目のタイムスロットでシステム容量として定義される。それ以外の場合、IoT デバイス n の報酬は0とする。

3-3 特性評価

(1) シミュレーション諸元

シミュレーション諸元は文献[15]を参考して設定する。シミュレーションでは、帯域幅 $B = 125$ KHz の $M = 3$ 個 AWGN サブチャネルを考える。長距離、低電力通信向けに設計された LoRAWAN IoT ネットワークを考える。LoRAWAN では、各チャネルに拡散率 (SF) が異なる最大 $D=6$ 個 IoT デバイスが同時にアクセスできる。雑音電力は -123 dBm に設定する。チャネル状態は、通信路利得 $g_j^m = \{0.5 \times 10^{-4}, 1 \times 10^{-4}, 1.5 \times 10^{-4}\}$ にそれぞれ対応する [38]。バッテリー容量の上限は 30dBm とする。各タイムスロットのハーベスティングするエネルギーは、 $[4, 8, 12]$ dBm の1つの値と想定する。送信電力レベルを $[20, 15, 10]$ dBm に設定する。割引率 γ を 0.9 とする。

DNN の隠れ層数を3つに設定する。入力層ニューロン数は、状態集合の要素数に等しい。つまり、7である。出力層ニューロン数は、チャネルと選択可能な送信電力レベルおよびエネルギーハーベスティングとデータ送信のインジケータの組み合わせ数に等しい。つまり、18である。隠れ層ニューロン数は、それぞれ 500, 250, 120 に設定する。DNN は、RMSProp 最適化アルゴリズムに基づいて更新する。表 1 にシミュレーション諸元を示す。

表 1：シミュレーション諸元

シミュレーションエリア & 半径	半径 $r = 1$ km の円形エリア
GW と IoT デバイス 間の距離	GW の周囲にランダムに分布
チャネル数	3
チャネル数ごとの帯域幅	125 KHz
同一チャネルに同時にアクセスできる最大な IoT デバイス数	6

雑音電力	-123dBm
通信路利得	$\{0.5 \times 10^{-4}, 1 \times 10^{-4}, 1.5 \times 10^{-4}\}$
バッテリー容量の上限	30dBm
各タイムスロット中のハーベスティングエネルギー	[4, 8, 12] dBm
送信電力レベル	[20, 15, 10] dBm
ε	0.01
割引率 γ	0.9
DNN 入力層ニューロン数	7
DNN 出力層ニューロン数	18
隠れ層数	3
隠れ層ニューロン数	[500, 250, 120]

(2) シミュレーション結果と考察

図 10 に 6 つのタイムスロットで構成される 1 フレームの平均システム容量を示す。この結果から、DQLRA は、IoT デバイス数に関係なく、ランダム法と比較してシステム容量を大幅に改善できることがわかる。これは、DQLRA が様々な状態でシステム容量が高められるリソース割り当て行動を学習できることが理由として考えられる。また、IoT デバイス数が 18 個以下の場合、ECAA を用いることで、平均システム容量は DQLRA より高いことがわかる。しかし、IoT デバイス数が 18 個より多い場合、ECAA でリソース割り当てることは不可能になる。その理由は以下のとおりである。ECAA でまずマッチング理論[38]に基づいて IoT デバイスにチャンネルを割り当てる。ECAA でのチャンネル割り当て過程の終了条件は、全ての IoT デバイスにチャンネルが割り当てられる。また、各チャンネルは最大 D 個の IoT デバイスに割り当てる可能性がある。DM < N の場合、ECAA を用いて、一部の IoT デバイスにチャンネルを割り当てられないため、マッチング過程は終了できない。

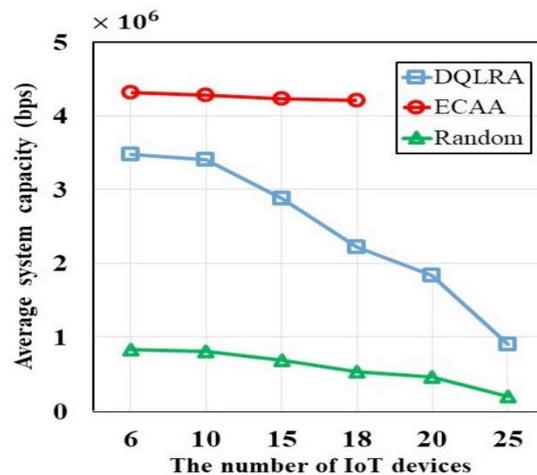


図 10 : IoT デバイスの数の増加に伴う、システム容量の比較

4 まとめ

ダイナミックスペクトラムアクセスを実現するため、無線デバイスはスペクトラムセンシングで周波数

利用状況を検知して、適切に周波数利用状況を把握することで、他の無線デバイスに干渉を与えることなく、より効率的な周波数利用が可能となる。しかし、チャンネルの状態を正確にセンシングすることには、センシング時間とエネルギーがかかるため、通信遅延とデバイスのエネルギー消費がより高い。また、他の無線デバイスに干渉を避けるチャンネル選択手法では、従来法には事前情報が必要であるが、現実世界では実現できない可能性がある。深層学習は多層のニューラルネットワークを用いて高い予測性能を発揮することと特徴を抽出することができる機械学習の1つであり、階層数をはじめとしたパラメータを次世代無線通信システムに適したものに最適化し、事前情報なくても、学習させることで、高効率なチャンネル割当と高精度なスペクトラム状態の予測が実現できると考えられる。そこで、本研究では、将来のスペクトラム状態を予測することでスペクトラムセンシングを行う回数削減技術および事前情報なくて他の無線デバイスに干渉を避けるチャンネル選択技術を検討した。研究結果は以下のようにまとめる。深層学習を用いてスペクトラム状態の予測：指数オン/オフモデルと隠れマルコフモデルの2種類のチャンネルを検討した。指数オン/オフモデルでは、1つのタイムスロットのみスペクトラム状態を使用した場合でも、スペクトラム状態を正確に予測できることを示した。隠れマルコフモデルでは、RMSEと予測精度は、無線デバイスの遷移確率との関係があり、ある定常状態から別の定常状態に遷移する確率が小さいほど、より高い予測精度が取得できることを示した。深層学習を用いてチャンネル選択：無線リソース制限下で、低計算量で事前情報を必要とせずに無線デバイスが利用する周波数、送信電力、エナジーハーベスティング時間を自動的に割り当てられる。

【参考文献】

- [1] J. A. Ansere, G. Han, H. Wang, C. Choi and C. Wu, "A reliable energy efficient dynamic spectrum sensing for cognitive radio IoT networks," *IEEE Internet Things J.*, Vol. 6, No. 4, PP. 6748-6759, Aug. 2019.
- [2] J. Yao and N. Ansari, "Caching in energy harvesting aided internet of things: A game-theoretic approach," *IEEE Internet Things J.*, Vol. 6, No. 2, PP. 3194-3201, Apr. 2019.
- [3] 李傲寒, 大槻知明, "エナジーハーベスティングIoTにおける深層強化学習を用いた無線リソース割り当て法," 電子情報通信学会無線システム研究会, 東京工業大学, 東京, 2020年3月.
- [4] J. Zhu, Y. Song, D. Jiang and H. Song, "A new deep-Q-learning-based transmission scheduling mechanism for the cognitive internet of things," *IEEE Internet Things J.*, Vol. 5, No. 4, PP. 2375-2385, Aug. 2018.
- [5] F. Hussain, S. A. Hassan, R. Hussain and W. Hossain, "Machine learning for resource management in cellular and IoT networks: potentials, solutions, and open challenges," *arXiv:1907.08965*, PP. 1-1, July. 2019.
- [6] C. He, Y. Hu, B. Zeng, "Joint power allocation and channel assignment for NOMA with deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, Vol. 37, No. 10, PP. 2200-2210, Oct. 2019.
- [7] Y. Zhang, Q. Wu, M. Shikh-Bahaei, "On ensemble learning based secure fusion strategy for robust cooperative sensing in full-duplex cognitive radio networks," *IEEE Trans. Commun.*, DOI: 10.1109/TCOMM.2020.3005708, PP. 1-1, June 2020.
- [8] L. You, X. Chen, X. Song, F. Jiang, W. Wang, X. Gao, G. Fettweis, "Network massive MIMO transmission over millimeter-wave and terahertz bands: mobility enhancement and blockage mitigation," *IEEE J. Sel. Areas Commun.*, DOI: 10.1109/JSAC.2020.3005493, PP. 1-1, June 2020.
- [9] S. Kang and C. Joo, "Low-complexity learning for dynamic spectrum access in multi-user multi-channel networks," *IEEE Trans. Mobile Comput.*, DOI: 10.1109/TMC.2020.2999075, PP. 1-1, June 2020.
- [10] B. Hamdaoui, B. Khalfi, and N. Zorba, "Dynamic Spectrum Sharing in the Age of Millimeter Wave Spectrum Access," *IEEE Netw.*, DOI: 10.1109/MNET.011.1900581, PP. 1-7, Apr. 2020.
- [11] L. Li, L. Liu, J. Bai, H. Chang, H. Chen, J. Ashdown, J. Zhang and Y. Yi, "Accelerating Model Free Reinforcement Learning with Imperfect Model Knowledge in Dynamic Spectrum Access," *IEEE Internet of Things J.*, DOI: 10.1109/JIOT.2020.2988268, PP. 1-1, Apr. 2020.

- [12] Y. Zhou, Z. Zhou, Z. Liu, P. Fan, and Y. Guan, "Low-PMEPR Preamble Sequence Design for Dynamic Spectrum Allocation in OFDMA Systems," *IEEE Trans. Commun.*, Vol. 68, No. 5, PP. 2922-2933, Feb. 2020.
- [13] L. Yu, J. Chen, G. Ding, J. Yang and J. Sun, "Spectrum prediction based on taguchi method in deep learning with long shor-term memory," *IEEE Access*, Vol. 6, PP. 45923-45933, Sept. 2018.
- [14] G. Ding, F. Wu, Q. Wu, S. Tang, F. Song, A. Vasilakos and T. Tsiftsis, "Robust Online Spectrum Prediction With Incomplete and Corrupted Historical Observations," *IEEE Trans. Veh. Commun.*, Vol. 66, No. 9, PP. 8022-8036, Sept. 2017.
- [15] X. Liu, Z. Qin, Y. Gao and J. A. McCann, "Resource allocation in wireless powered IoT networks," *IEEE Internet Things J.*, Vol. 6, No. 3, PP. 4935-4945, June 2019.
- [16] Z. Kuang, G. Liu, G. Li and X. Deng, "Energy efficient resource allocation algorithm in energy harvesting-based D2D heterogeneous networks," *IEEE Internet Things J.*, Vol. 6, No. 1, PP. 557-567, Feb. 2019.
- [17] F. Li, H. Jiang, R. Fan and P. Tan, "Cognitive non-orthogonal multiple access with energy harvesting: an optimal resource allocation approach," *IEEE Trans. Veh. Technol.*, Vol. 68, No. 7, PP. 7080-7095, July 2019.
- [18] F. Tang, Y. Zhou and N. Kato, "Deep Reinforcement Learning for Dynamic Uplink/Downlink Resource Allocation in High Mobility 5G HetNet," *IEEE J. Sel. Areas Commun.*, DOI: 10.1109/JSAC.2020.3005495, PP. 1-1, June 2020.
- [19] S. Lee and S. Lee, "Resource Allocation for Vehicular Fog Computing Using Reinforcement Learning Combined with Heuristic Information," *IEEE Internet Things J.*, DOI: 10.1109/JIOT.2020.2996213, PP. 1-1, May 2020.
- [20] X. Xiong, K. Zheng, L. Lei and L Hou, "Resource Allocation Based on Deep Reinforcement Learning in IoT Edge Computing," *IEEE J. Sel. Areas Commun.*, Vol. 38, No. 6, PP. 1133-1146, Apr. 2020.
- [21] M. Wellens, J. Riihijarvi, and P. Mahonen, "Evaluation of adaptive mac-layer sensing in realistic spectrum occupancy scenarios," *IEEE DySPAN*, April 2010, pp. 1–12.
- [22] M. Wellens, J. Riihijarvi, and P. Mahonen, "Evaluation of adaptive mac-layer sensing in realistic spectrum occupancy scenarios," *IEEE DySPAN*, April 2010, pp. 1–12.
- [23] X. Xing, T. Jing, Y. Huo, H. Li and X. Cheng, "Channel Quality Prediction Based on Bayesian Inference in Cognitive Radio Networks," *IEEE INFOCOM*, 2013.
- [24] L. Yu, J. Chen, G. Ding, Y. Tu, J. Yang and J. Sun, "Spectrum Prediction Based on Taguchi Method in Deep Learning with Long Short-Term Memory," *IEEE ACCESS*, Sept. 2018.
- [25] L. Yu, J. Chen and G. Ding, "Spectrum Prediction via Long Short Term Memory," *IEEE ICC*, 2017.
- [26] L. Yu, Q. Eang, Y. Guo and P. Li, "Spectrum Availability Prediction in Cognitive Aerospace Communications: A Deep Learning Perspective," *IEEE CCAA*, Aug. 2017.
- [27] A. Min, K.-H. Kim, J. Singh, and K. Shin, "Opportunistic spectrum access for mobile cognitive radios," *IEEE INFOCOM*, April 2011, pp. 2993–3001.
- [28] V. Raj, I. Dias. T. Tholeti and S. Kalyani, "Spectrum Access in Cognitive Radio Using A Two Stage Reinforcement Learning Approach," *IEEE Journal of Selected in Signal Processing*, 2018.
- [29] Y. Cui, X. Jing, S. Songlin, X. Wang, D. Cheng and H. Huang, "Deep Learning Based Primary User Classification in Cognitive Radios," *IEEE ISIT*, 2015.
- [30] C. Zhang and P. Patras, "Long-Term Mobile Traffic Forecasting Using Deep Spatio-Temporal Neural Networks," *ACM Mobihoc*, 2018.
- [31] Q. Li, Y. Ju, C. Zhao, "Classification of discrete sequential protocol messages based on LSTM network and transfer learning," *IEEE ICCS*, 2020.
- [32] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

- [33] V. Nekoukhou, M. H. Alamatsaz, and H. Bidram, "A discrete analog of the generalized exponential distribution," *Communications in Statistics Theory and Methods*, vol. 41, no. 11, pp. 2000–2013, 2012.
- [34] S. Sohn, S. Jang and J. Kim, "HMM-based Adaptive Frequency-Hopping Cognitive Radio System to Reduce Interference Time and to Improve Throughput," *KSII Trans. Internet and Information System*, Vol. 4, No. 4, Aug. 2010.
- [35] L. Liang, H. Ye and G. Y. Li, "Spectrum sharing in vehicular networks based on multi-agent reinforcement learning," *IEEE J. Sel. Areas Commun.*, Vol. 37, No. 10, PP. 2282-2292, Oct. 2019.
- [36] F. Adelantado, X. Vilajosana, P. T. Perio, B. Martinex, J. Segui and T. Watteyne, "Understanding the limits of LoRaWAN," *IEEE Commun. Mag.*, Vol. 55, No. 9, PP. 34-40, Sept. 2017.
- [37] Geoffrey Hinton, *Neural Networks for Machine Learning*, online course. <https://www.coursera.org/learn/neural-networks/home/welcome>
- [38] S. Bayat, Y. Li, L. Song, and Z. Han, "Matching theory: Applications in wireless communications," *IEEE Signal Process. Mag.*, Vol. 33, No. 6, PP. 103-122, Nov. 2016.

〈 発 表 資 料 〉

題 名	掲載誌・学会名等	発表年月
Deep Reinforcement Learning Based Resource Allocation for Energy Harvesting Internet of Things	論文誌 (掲載雑誌未定)	投稿予定
Convolutional LSTM Based Long-Term Spectrum Prediction in Dynamic Spectrum Access Systems	論文誌 (掲載雑誌未定)	投稿予定
Deep Reinforcement Learning Based Resource Allocation in Wireless Communications	IEEE International Conference on Communications	2020.09.30 投稿予定
Spectrum Prediction Using LSTM in Dynamic Spectrum Access Systems	IEEE International Conference on Communications	2020.09.30 投稿予定
Deep Q-Learning Based Resource Allocation for Energy Harvesting Internet of Things	電子情報通信学会総合大会	2020年03月
エネルギーハーベスティング IoT における深層強化学習を用いた無線リソース割り当て法	電子情報通信学会無線システム研究会	2020年03月