

ニューラルネットワークに基づくQ学習を用いた 無線リソース割り当て手法

相原 直紀[†] 安達 宏一[†] 田久 修^{††} 太田 真衣^{†††} 藤井 威生[†]

[†] 電気通信大学 先端ワイヤレス・コミュニケーション研究センター

〒182-8585 東京都調布市調布ヶ丘1-5-1

^{††} 信州大学 電子情報システム工学科

〒380-8553 長野県長野市若里4-17-1

^{†††} 福岡大学 電子情報工学科

〒814-0180 福岡県福岡市城南区七隈8-19-1

E-mail: [†]{aihara,adachi,fujii}@awcc.uec.ac.jp, ^{††}takyu@shinshu-u.ac.jp, ^{†††}maiohta@fukuoka-u.ac.jp

あらまし モノのインターネット (IoT: Internet of Things) 機器の増加に伴い、周波数資源の更なる逼迫が懸念されている。多数の無線ノードがランダムアクセスを用いて周波数を共有する環境で高信頼通信を実現するためには、ノード間で生じる干渉を低減もしくは回避する必要がある。例えば、衝突回避機能付きキャリアセンス多元接続 (CSMA/CA : Carrier Sense Multiple Access/Collision Avoidance) をランダムアクセス方式として用いたときに、互いに送信を検知できない隠れノードが存在した場合、受信ノードにおいて送信パケットが衝突し、パケット損失が発生する可能性がある。本稿では、パケット衝突により引き起こされるパケット配信率 (PDR: Packet Delivery Rate) の低減を回避するためにQ学習を用いることで多数の無線ノードに無線資源を割り当てるための検討を行う。検討システムでは、受信ノードが観測可能な受信成功パケット数をQ学習における報酬関数として強化学習を行うことにより、各無線ノードに無線リソースを割り当てる。お互いに隠れノード関係にあるような無線ノードには異なる無線リソースを割り当てることによりパケット衝突を回避し、PDRの向上を実現する。LoRaWANネットワークを仮定した環境において、計算機シミュレーションにより提案法を用いることでランダムに無線リソースを割り当てた場合と比較してPDRの10%値を22%改善できることを示す。

キーワード 周波数共有, 機械学習, 無線リソース割り当て

Wireless Resource Allocation Using Q-Learning Based on Neural Network

Naoki AIHARA[†], Koichi ADACHI[†], Osamu TAKYU^{††}, Mai OHTA^{†††}, and Takeo FUJII[†]

[†] Advanced Wireless and Communication Research Center, The University of Electro-Communications
1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585

^{††} Dept. of Electrical and Computer Engineering, Shinshu University
4-17-1, Wakasato, Nagano, 380-8553

^{†††} Dept. of Electronics and Computer Science, Fukuoka University
8-19-1, Nanakuma, Jonan, Fukuoka 814-0180

E-mail: [†]{aihara,adachi,fujii}@awcc.uec.ac.jp, ^{††}takyu@shinshu-u.ac.jp, ^{†††}maiohta@fukuoka-u.ac.jp

Abstract Depletion of wireless resource becomes more serious due to the rapid increase of Internet-of-Things (IoT) devices. The management of mutual interference among the wireless nodes is essential in the densely deployed wireless network. If carrier-sense multiple access/collision avoidance (CSMA/CA) is adopted as a random access scheme, some wireless nodes may not be able to carrier sense each other. The packet collision due to the hidden terminal incurs the packet loss. In this paper, a wireless resource allocation scheme based on Q-learning is proposed for a large number of wireless nodes in order to avoid the packet delivery rate (PDR) degradation. The proposed scheme allocates the wireless resources to wireless nodes by adopting the number of successfully received packets, which is observable at the network controller, as a reward of Q-learning. By learning and resource allocation, different resources are allocated to the wireless nodes which are in the hidden terminal relation. Under the LoRaWAN network, the numerical results elucidate that the proposed scheme can improve 22% PDR's 10% value compared to random allocation scheme.

Key words Frequency Sharing, Machine Learning, Wireless Resource Allocation

1. はじめに

近年、スマートフォンの普及に伴いマルチメディア通信やソーシャルネットワークサービス (SNS: Social Networking Service) のように高速・大容量な通信が増加している。さらには、スマートメータやセンサネットワークといったモノのインターネット (IoT: Internet-of-Things) 機器の増加により、周波数資源の更なる逼迫が懸念されている。多数の IoT 無線ノードが周波数を共有するような環境では、送受信ノード間の無線伝播路で生じる信号の減衰や歪みに加えて、無線ノード間で発生する相互干渉が問題となる。LoRaWAN に代表される LPWA (Low Power Wide Area) では、無線ノードの低消費電力化・バッテリーの長寿命化のために、集中制御ではなく各無線ノードが自律分散的ランダムアクセスを行うことで周波数の共有を行う。LoRaWAN において、各無線ノードはアップリンクパケットの送信前にチャンネルが使用中かどうかをキャリアセンス (CS: Carrier Sense) することによってパケットの衝突を回避する。しかしながら、広カバレッジエリアを考えた場合、無線ノード間の距離減衰や障害物の影響により通信中の無線ノードがいるのにも関わらずチャンネルが空いていると誤検知してしまう隠れ端末問題が生じる可能性がある。無線ノードが自身の周辺無線環境を観測し、その情報をフィードバックすることによりパケット衝突を回避することができると考えられる。しかしながら、周辺環境の観測やフィードバック情報送信のために生じる余計な電力消費やネットワーク負荷の増大といった問題が生じる。

筆者らは以前、衝突回避機能付きキャリアセンス多元接続 (CSMA/CA : Carrier Sense Multiple Access/Collision Avoidance) [1] を用いる無線 LAN 環境において、アクセスポイント (AP: Access Point) が観測することのできる AP と無線ノード間のチャンネル (可観測情報: Observable information) と無線ノード間のチャンネル (不可観測情報: Unobservable information) との関係性を教師あり学習を用いることで学習し、その学習結果に基づいてパケット衝突を回避するように無線リソースの割り当てを行う方法を提案した [2]。しかしながら、教師あり学習では、予め教師データを取得する必要があるため、実システムに導入するには問題がある。

そこで本稿では、教師データを必要とせず学習を行うことのできる Q 学習を用いた無線リソース割り当て法を提案する。Q 学習に用いる報酬関数としては、受信ノードが観測することのできる各端末に対する受信成功パケット数を用いる。これにより無線ノードからの明示的なフィードバックを必要とせず、無線リソースの割り当てが可能となる。本研究では Q 学習を用いて各端末からの受信成功パケット数を最大化できるような無線リソース割当を探索する。受信成功パケット数と PDR に強い相関があるような環境では、このリソース割り当ての探索により PDR 特性を向上させることが可能である。

LoRaWAN に準拠した計算機シミュレーションにより、提案法を用いることにより、各無線ノードにランダムに周波数リソースを割り当てた場合と比較して、パケット配信率 (PDR:

Packet Delivery Rate) の 10% 値を 22% 向上できることを示す。

本論文は以下の構成になっている。第 2 章で本研究で用いた既存技術について説明する。第 3 章で提案手法である Q 学習を用いた無線リソース割り当てについて説明する。第 4 章で計算機シミュレーションの結果について説明し、第 5 章で本論文のまとめを行う。

2. 既存技術

2.1 強化学習

強化学習はある与えられた状況において、主体 (エージェント) の最適な行動はどれかを求める学習の一形式である。エージェントは教師あり学習のように特定の状況においての最適な行動の組は与えられないが、特定の状況及び行動に対する報酬が与えられる。エージェントはこの各行動の報酬の総和である報酬関数を元に、ある状況に対する最適な行動を行う。しかし、この報酬関数は環境に依存し、一般的に解析的に求めることは困難である。この問題に対し、強化学習を用いて実際の行動を元に報酬関数を近似的に求めることが提案されている。この学習形式は特にマルコフ連鎖のような、行動が次時間での状況に変化を与えるような状況や、一連の行動の連鎖に対して結果が与えられるような状況に対して有効である。本研究では強化学習の一つである Q 学習を用いる。

2.1.1 Q 学習

ある時刻 t において Q 学習では式 (1) で与えられる更新式を元に報酬関数を近似する。

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)] \quad (1)$$

ここで、状態集合及び行動集合を S, A とした時に $Q(S_t, A_t)$ は状態 $S_t \in S$ において行動 $A_t \in A$ を行うことにより得られる報酬の期待値、 R_{t+1} は時刻 $t+1$ で行動 A_t に対し与えられる即時の報酬、 γ は割引率、 α は Q 学習率である。この更新式に従う近似は確率 1 で実際の報酬関数 Q^* に収束することが知られている [4]。

2.1.2 ニューラルネットワークを用いた Q 学習

純粋な Q 学習では、ある状況およびある行動に対する Q 値を表の形式で記録する。この形式では、状態と行動の組み合わせ数 ($|S| \times |A|$) が非常に大きくなる場合に Q 値の保存に膨大な記憶容量を必要とする。これを回避するために、Q 値の表現にニューラルネットワーク (NN: Neural Network) を用いる方法が提案されている [5]。NN は機械学習の一手法であり、入力と出力の組を与えることで、その入出力関係の近似モデルを得ることができる。これを用いて、図 1 に示すように、NN に現在の状況 $S_t \in S$ を入力すると、それぞれの行動 $A_t \in A$ に対する Q 値の近似値が出力されるように、NN 内の重みを学習する。

2.2 LoRaWAN

LPWA の一規格である LoRaWAN では、物理層の規格としてチャープ・スペクトラム拡散 (CSS: Chirp Spread Spectrum) 変調と MAC 層規格としてランダムアクセス方式を用いる。ク

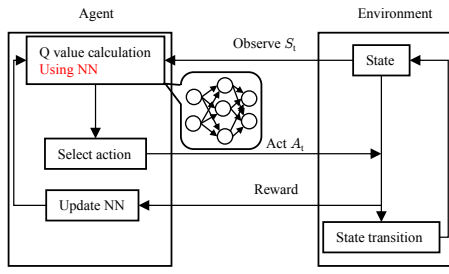


図1 深層強化学習のシステムモデル

ラス A では、純 ALOHA 方式に基づくアップリンク送信の直後に 2 つの受信ウィンドウを設けることによりダウンリンク受信を行う。クラス B ではこれに加え定期的な間隔でのピーコン受信ウィンドウ、クラス C では常に受信ウィンドウを開放することでより多くのダウンリンク受信を可能としている。そのため、クラス C の消費電力がこの中では最も大きくなるが、低遅延での通信が可能であり、加えて CSMA/CA などの複雑な多元接続方式を用いることが可能となっている。[3]

3. 提案方式

3.1 学習モデル

3.1.1 Q 学習器の設計

本研究では Q 学習器を以下のように定義する。なお、一回のチャンネル割当、Q 報酬値の観測及び学習処理のサイクルを 1 エポックと定義する。各無線ノードに対して異なる Q 学習器が用意されているものとする。各エポックでの状態をそれぞれの Q 学習器に入力し、得られた出力からそれぞれ独立に割当チャンネルを決定する。ここでは、一般性を失うことなく無線ノード $i \in \mathcal{I}$ に関する Q 学習器に着目して説明を行う。

無線ノード i の Q 学習に必要な状態集合 S 、行動集合 A_i 、Q 報酬値 Q_i を以下のように定義する。

- 状況 S : エポック t において各端末に割り当てられているチャンネルのインデックスの組み合わせ
- 行動 A_i : 次エポック $t+1$ において無線リソース無線ノード i に割り当てられるリソースインデックス k_i
- Q 報酬値 Q_i : エポック t におけるリソース割当て状況 S_t に対し、次エポック $t+1$ で無線リソース m_{k_i} を割り当てた際のエポック長の中に無線ノード i から受信成功したパケット数。なお、ここで m_k は無線リソースであり、本研究では Q_i に関して、そのエポックまでに受信に成功した最大パケット数で正規化したものを用いるものとした。これは NN を用いるため報酬関数を 0 から 1 へ正規化する必要があるためである。

今回は行動決定に ϵ -greedy 法を用いる。本研究では確率 ϵ で全無線ノードがランダムに行動を決定し、 $1-\epsilon$ で全無線ノードがそれぞれ最大の Q 値を持つ行動を選択するものとする。ソフトマックスを用いた行動決定を行わない理由としては、各

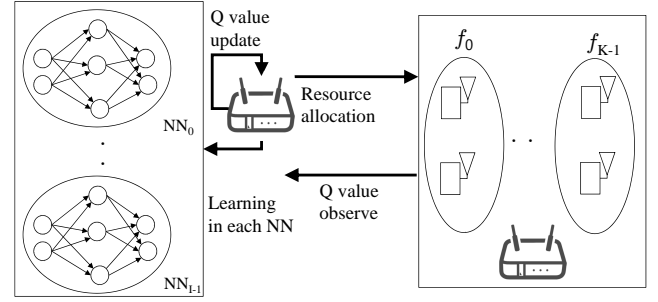


図2 提案モデル

無線ノードが Q 値の低いリソースを出力する確率があり、システム全体で最適リソースを割り当てられないためである。

3.1.2 NN の設計

本研究では Q 値の近似に用いる NN として L 層パーセプトロンを用いる。無線ノード i に対し、第 l 層におけるニューロン n の出力モデル $z_i^{(j-1)}$ は式 (2) 及び式 (3) で与えられる。

$$y_{i,n}^{(l)} \left(\mathbf{z}_i^{(l-1)}, \mathbf{w}_{i,n}^{(l-1)} \right) = \sum_{j=0}^J w_{i,n,j}^{(l-1)} \phi(z_{i,j}^{(l-1)}) \quad (2)$$

$$z_{i,j}^{(l-1)} = f(y_{i,n}^{(l-1)}) \quad (3)$$

ここで $z_{i,j}^{(l-1)}$ は第 $l-1$ 層におけるニューロン j からの出力、 $w_{i,n,j}^{(l-1)}$ は第 $l-1$ 層におけるニューロン j から第 l 層におけるニューロン n までの結合重みを示している。また、 $\phi(x)$ 及び $f(x)$ は入力の基底関数及びニューラルの活性化関数を示しており、本研究ではそれぞれ恒等関数及び式 (4) に示すシグモイド関数を用いた。

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

3.2 Q 学習を用いた無線リソース割り当て

無線リソース割当を行う際、すべての無線ノードに対して同時に Q 学習を適用すると探索すべき行動空間のサイズが非常に大きくなるため、無線ノード全体をいくつかのクラスタに分解し、そのクラスタに属する無線ノードにのみ Q 学習を適用し最適リソース割り当てを行う。クラスタに属していない無線ノードに関しては以前のエポックと同じ無線リソースを用いるものとする。あるクラスタ内の無線ノードに対してリソース割り当ての探索が終了した後は別のクラスタに対してリソース割り当てを探索し、すべてのクラスタに対してリソース割り当ての探索が終了後、その時点でのリソース割り当てを最終的な最適化結果とする。

情報収集局 (FC: Fusion Center) は各無線ノードに対して Q 学習器からの出力に基づき無線リソースを割り当てる。本研究での提案モデルを図 2 に示す。エポック t における割り当てアルゴリズムの動作を示す。ここで、エポック t において学習を行うノードの集合 (クラスタ) を $C_t \in \mathcal{I}$ とする。

ステップ 1 現在のリソース割当て状況 ($S_{t-1} \in S$) をクラスタ C_t に属する無線ノード $i \in C_t$ に関する NN へと入力し、

各 NN からの出力 $Q_{i,t}$ を得る.

ステップ2 クラスタ C_t に属する無線ノードに対して確率 ϵ でランダムに無線リソースを割り当て、確率 $1 - \epsilon$ で $m_i^* = \arg \max_{k \in \mathcal{K}} Q_i(S_t, m_k)$ の無線リソースを割り当てる. ここで $\mathcal{K} = \{0, 1, 2, \dots, K - 1\}$ は無線リソースの集合である. 割り当て状況を $S_t = (m_0^*, m_1^*, \dots, m_{I-1}^*)$ と更新する.

ステップ3 S_t における受信成功パケット数を記録し、それまでのエポックにおける最大受信成功パケット数で正規化する. これにより、PDR と相関を持つ受信成功パケット数が最大になるような無線リソースを各無線ノードに対して割り当てることができる.

無線ノードが FC から離れた場所に位置している場合、受信信号強度 (RSSI: Received Signal Strength Indicator) が非常に小さくなり、PDR 特性が著しく劣化する可能性が存在する. 一方で、FC 付近の無線ノードは RSSI が大きいいため、キャプチャ効果によりパケット配信に成功する可能性がある. そのため、無線ノード間のチャネル状態を考慮せずに、クラスタ化して学習を行うと、無線ノードの PDR 特性が不均衡になることが懸念される. そこで、本稿では学習を行う前に RSSI 値に基づいて無線ノードをソートしクラスタリングする方法について検討を行う. 具体的には、RSSI 値に基づいて無線ノードを降順もしくは昇順にソートし、ソート後の無線ノードをクラスタ化し Q 学習を適用する. これにより、FC へのチャネル状態が劣悪な環境にいる無線ノードに優先的に周波数リソースを割り当てることが可能になると考えられる.

4. 数値結果

4.1 シミュレーションモデル

4.1.1 システムモデル

本稿では、シミュレーションエリア内にランダムに配置された I 個の無線ノードが固定配置された FC にパケットを送信する上りリンクを仮定する. 無線ノードに割り当てることができる周波数リソース数は K 個あるものとする. 通信路モデルとしては、式 (5) で示す伝搬距離に依存するパス損失と空間相関を持つシャドウィング損失を考慮する [6].

FC における無線ノード $i \in \mathcal{I}$ における受信信号電力は式 (5) で与えられる.

$$P_{r,i} = P_t - 10a \log_{10} d_i - b - 10c \log_{10} f_c - \eta(x_i, y_i) \quad (5)$$

P_t は送信電力、 d_i は FC と無線ノード i 間の距離 [km]、 f_c は搬送波周波数 [MHz] である. また、 $\eta(x_i, y_i)$ は無線ノード i が存在する位置における空間相関を持つ対数正規分布に従うシャドウィング損失である [6]. 伝搬パラメータ a 、 b 及び c はそれぞれ距離、オフセット及び周波数により変動する伝搬減衰に対するパラメータであり、文献 [7] のものを用いた.

各無線ノードはパケット生成強度 λ [packets/sec] に従うポアソン到着過程でパケットを生成するものとする. 複数の無線ノードが同一の周波数リソースを用いて FC にパケットを同時に送信した場合、各無線ノードの受信信号電力対干渉電力+

雑音電力比 (SINR: Signal-to-Interference plus Noise Ratio) を計算し、もし SINR が閾値 Γ を上回っていた場合には正しく受信されるものとする. 正しく受信されなかった場合、無線ノードは二進バックオフ [1] に従いパケットを再送するものとする. 再送時のバックオフ幅は $[0, CW]$ から一様分布で決定し、CW は式 (6) で与えられる.

$$CW = CW_{\min} * 2^{N_r} \quad (6)$$

ここで、 CW_{\min} は最小バックオフ幅であり、 N_r は再送回数である.

PDR は式 (7) で与えられる.

$$PDR = D/S \quad (7)$$

ここで、 D を各ノードに対し最低一つの FC で復調できたパケット数、 S を全ノードが生起したパケット数とする. なお、提案手法を用いた場合は学習後のリソース割り当てを用いて PDR を計測するものとする.

LoRaWAN のシステムパラメータは文献 [8] の日本におけるパラメータ設定 AS923 を用いた.

4.1.2 学習モデル

本稿では無線ノードを P 個のクラスタに分割し、エポック数に対応したクラスタ p に対し前述したアルゴリズムで無線チャネルの割当を行う. ϵ は式 (8) で与えられるものとする.

$$\epsilon(t_p) = \frac{T - t_p}{T} \quad (8)$$

ただし、 $T = N/P$ とし、 t_p はクラスタ p に対する最適化の過程でのエポック数とする.

NN に関しては、本稿では $L = 3$ 層の単純パーセプトロンを用いた. 入力層及び出力層を除いた第 2 層を隠れ層と定義し、隠れ層は J 個のニューロンを持つものとする. チャネルインデックス $a_i \in \mathcal{A}_i$ を割り当てられた無線ノード i に対し、隠れ層におけるニューロン j -出力層ニューロン a_i 間の重み $w_{i,a_i,j}^{(2)}$ の更新式は式 (9) で、隠れ層におけるニューロン r -入力層におけるニューロン q 間の重み $w_{i,r,q}^{(1)}$ の更新式は式 (10) で表される [5].

$$w_{i,a_i,j}^{(2)} \leftarrow w_{i,a_i,j}^{(2)} + h \times E_i \times o_i \times (1 - o_i) * z_{i,a_i}^{(2)} \quad (9)$$

$$w_{i,r,q}^{(1)} \leftarrow w_{i,r,q}^{(1)} + h \times z_{i,r}^{(1)} \times \delta_r \quad (10)$$

ここで、式 (9) に関して、 h は NN 学習率、 E_i は実際の報酬関数 o_i と NN から出力された報酬関数の近似値との誤差であり、式 (11) で与えられる. また、式 (10) に関して、 δ_r は式 (12) で与えられる.

$$E_i = o_i - z_{i,a_i}^{(2)} \quad (11)$$

$$\delta_r = z_{i,r}^{(1)} \times (1 - z_{i,r}^{(1)}) \times w_{i,r} \times E_i \times o_i \times (1 - o_i) \quad (12)$$

4.2 受信成功数と PDR の相関

ここではまず、PDR 特性の向上を目的とした Q 学習におい

表 1 無線システムパラメータ

シミュレーションエリア	3×3 km ²
送信電力 P_t	16 dBm
搬送波周波数 f_c	923 MHz
帯域幅	125 kHz
無線ノード数 I	100
伝搬係数 a	2.0
伝搬損失オフセット b	32.45
伝搬周波数係数 c	2.0
シャドウイング標準偏差	3.48 dB
雑音電力スペクトラム密度	-174 dBm/Hz
無線ノードクラスタ分割数 P	4
パケット長	5 symbol
パケット生成強度 λ	0.005packets/sec
キャリアセンス閾値	-80 dB [9]
受信 SINR 閾値 γ	-8.0 dB [10]
リソース数 K	16

表 2 学習パラメータ

エポック長	5000 symbol
エポック数 T	10000
Q 学習率 α	0.4
NN 層数 L	3
NN 隠れ層ニューロン数 J	15
NN 学習率 h	20

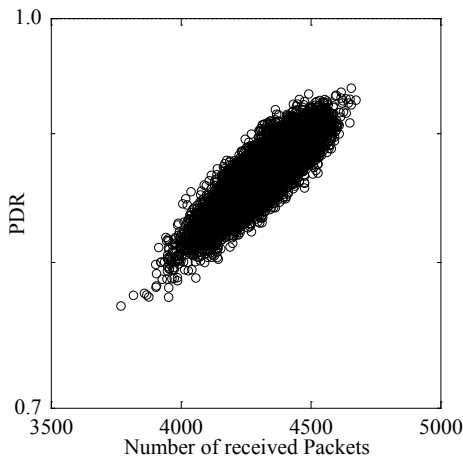


図 3 あるエポックにおける受信パケット数と PDR の相関関係

て Q 報酬値として受信成功パケット数を用いることの妥当性を検証する。図 3 に受信成功パケット数と PDR の関係を示す。また、受信成功パケット数と PDR 間の相関値を定量的に評価するため、式 (13) で定義されるピアソンの積率相関係数を計算する。

$$r = \frac{\sum_{t=0}^{T-1} (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=0}^{T-1} (x_t - \bar{x})^2 \sum_{t=0}^{T-1} (y_t - \bar{y})^2}} \quad (13)$$

ここで、 x_t と y_t はそれぞれエポック t における PDR 値と受信成功パケット数であり、 \bar{x} と \bar{y} はそれらの平均値である。式 (13) に基づいて計算された r を用いて得られる決定係数 r^2 はおよそ 0.73 となり、このモデルにおける受信成功パケット数

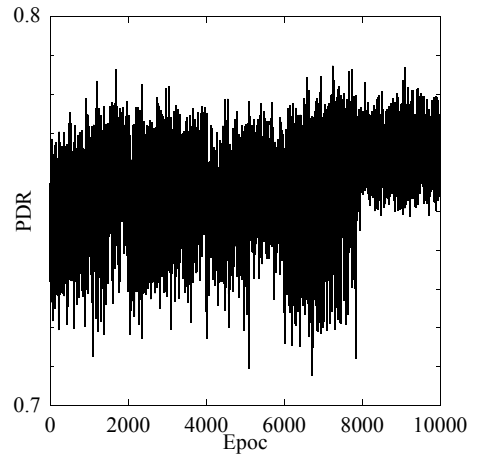


図 4 あるチャンネルにおけるリソース最適化の過程 (再送を考慮しない場合)

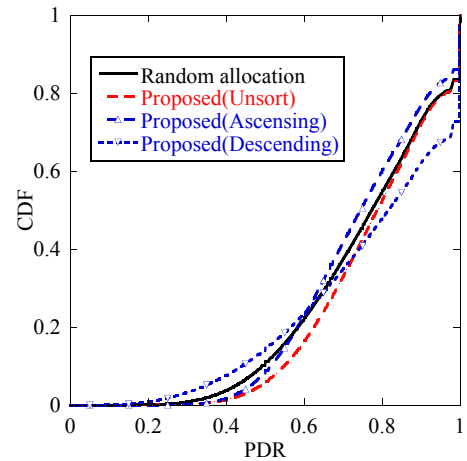


図 5 PDR の累積確率分布 (再送を考慮しない場合)

と PDR には強い相関があることが分かる。

4.3 PDR の評価

4.3.1 再送を考慮しない場合

図 4 にあるチャンネル実現に対しての Q 学習の過程を示す。この結果から、エポックが進むにつれて PDR が向上していることが確認できる。これは提案手法により適切なリソース割り当てを行うことができたためである。次に、図 5 にランダム割当てと各提案手法における PDR の累積確率分布を示す。この結果から、提案手法を用いることで、PDR の 10% 値が 5% 程度向上し、また平均 PDR が 3% 程度向上できていることがわかる。次に RSSI に基づくノードソートを行った場合について考える。昇順ソートを行うことで、平均 PDR は 1% 程度劣化するが、ノード間の不均衡が改善されていることがわかる。一方で、降順ソートを行うことでノード間の不均衡が悪化していることがわかる。これは、RSSI の弱い、もしくは強いノードを早い最適化のクラスタに配置したことで、そのノードに優先的なチャンネルを割り当てる結果となったことが理由として考えられる。

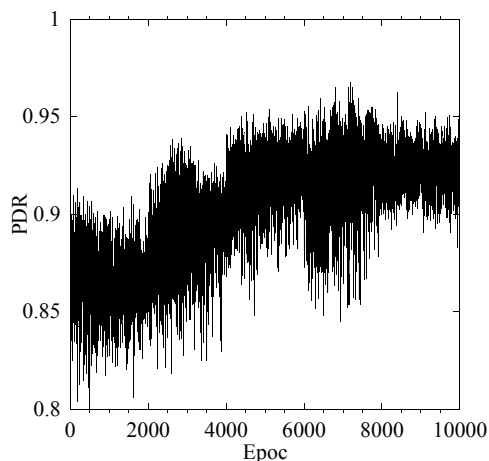


図 6 あるチャネルにおけるリソース最適化の過程 (再送を考慮した場合)

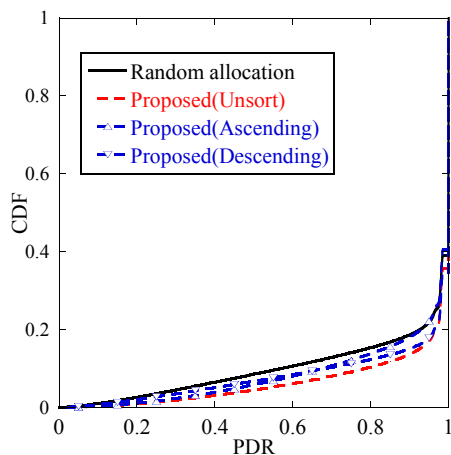


図 7 PDR の累積確率分布 (再送を考慮した場合)

4.3.2 再送を考慮した場合

図 6 にあるチャネル実現に対しての Q 学習の過程を示す。この結果から、再送を考慮しない場合と同様に、エポックが進むにつれて PDR が向上していることが確認できる。次に、図 7 から、提案手法を用いることで PDR の 10% 値が 22% 程度向上し、また平均 PDR が 5% 程度向上できていることがわかる。一方で、RSSI に基づくノードソートを行った場合、全体的な PDR 特性はソートを行わない場合より劣化した。これは、再送を考慮することでキャリアセンスの関係が受信成功パケット数に表れづらくなり、全体として最適なりソース配分を決定することが難しくなったためだと考えられる。RSSI に基づくノードソートを行った場合は、再送を考慮しない場合と同様の不均衡の改善および悪化が発生する結果が得られたが、再送を考慮しない場合よりも影響は小さくなった。これは再送によりノードからのパケットを受信することが簡単になったためだと考えられる。

5. 終わりに

本稿では、CSMA/CA において隠れノードの存在により生

じる干渉回避のために Q 学習を用いた無線リソース割り当てを提案し、計算機シミュレーションによる評価を行った。本手法は、Q 学習を用いて各ノードの受信パケット数が最大となるようなリソース割り当てを探索することで、明示的なオーバーヘッドなしにノード間のパケット衝突で回避することができるものである。シミュレーション結果から、提案手法はランダムな割当法と比較して PDR10% 値を 22% 程度向上できることを示した。今後の研究課題としては、ノード間不公平性改善のための学習モデル設計や、効率的な学習法の実装などが挙げられる。

謝辞本研究開発は総務省 SCOPE(受付番号 175104004) の委託を受けて行われたものである。

文 献

- [1] C. D. M. Cordeiro, D. P. Agrawal, "Ad Hoc Sensor Networks: Theory and Applications", World Scientific, 2006
- [2] Naoki Aihara, Koichi Adachi, "Orthogonal Resourced Allocation Using SVM for CSMA/CA", in *Proc. 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pp.779-784, Hawaii, Nov. 2018
- [3] J. Ortin, M. Cesana, A. Redondi, "How do ALOHA and Listen Before Talk Coexist in LoRaWAN?", in *Proc. IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp.1-7, Italy, Dec. 2018
- [4] R. S. Sutton, A. G. Barto, F. Bach, "Reinforcement Learning: An Introduction", A Bradford Book, 1998
- [5] 小高 知宏, "強化学習と深層学習 C 言語によるシミュレーション", オーム社, 2017
- [6] H. Claussen, "Efficient modelling of channel maps with correlated shadow fading in mobile radio systems", in *Proc. IEEE 16th Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 512-516, Germany, Sep. 2005
- [7] ITU-R, "Propagation data and prediction methods for the planning of short-range outdoor radiocommunication systems and radio local area networks in the frequency range 300 MHz to 100 GHz", [available] https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.1411-9-201706-I!!PDF-E.pdf (2018/12/21Accessed)
- [8] LoRa Alliance, "LoRaWAN Regional Parameters v1.1rB", [Available] https://lora-alliance.org/sites/default/files/2018-04/lorawantm_regional_parameters_v1.1rb_-_final.pdf (2018/12/18Accessed)
- [9] 総務省, "920MHz 帯省電力無線システムの高度化に係る技術基準の見直し(案)", [Available] http://www.soumu.go.jp/main_content/000452569.pdf (2018/12/18 Accessed)
- [10] H. Mroue, A. Nasser, B. Parrein, S. Hamrioui, E. Mona-Cruz, G. Rouyer, "Analytical and Simulation study for LoRa Modulation", in *Proc. 25th International Conference on Telecommunications*, pp. 655-659, France, Jun, 2018