

修士論文

IEEE 802.11 ネットワークにおける
強化学習に基づくキャリアセンス
閾値制御に関する研究

令和5年度修了

三重大学大学院工学研究科

電気電子工学専攻

通信工学研究室

東山 祐斗



目次

第1章 序論	1
1.1 研究背景	1
1.2 本研究の目的	3
1.3 論文構成	4
第2章 従来研究	5
2.1 IEEE 802.11 無線 LAN	5
2.1.1 無線 LAN の技術的動向	5
2.1.2 IEEE 802.11 DCF	7
2.1.3 隠れ端末および晒し端末問題	9
2.1.4 数理モデルに基づくネットワークシステムの最適化	11
2.2 機械学習の通信プロトコル設計への応用	12
2.2.1 機械学習	13
2.2.2 強化学習	15
2.2.2.1 Q学習	16
2.2.3 強化学習の MAC プロトコル設計への応用	18
第3章 提案手法	22
3.1 システムモデル	22

3.1.1	ネットワークモデル	22
3.1.1.1	隠れ端末を想定したネットワーク	22
3.1.1.2	晒し端末を想定したネットワーク	23
3.1.1.3	隠れ端末および晒し端末の混在想定したネットワーク	24
3.1.2	トラヒックモデル	24
3.1.3	チャネルモデル	25
3.1.4	アクセスモデル	25
3.2	提案 CS 閾値制御	26
3.2.1	本提案手法の概要	26
3.2.2	本提案手法におけるチャネルアクセス	26
3.2.3	学習モデル	27
3.2.3.1	状態: S	27
3.2.3.2	行動: A	27
3.2.3.3	報酬	28
第 4 章	提案手法の性能評価	32
4.1	評価条件	32
4.1.1	シミュレーション条件	32
4.1.2	評価項目	34
4.1.3	性能比較対象	34
4.2	隠れ端末環境下における評価	35
4.2.1	ネットワークトラヒック負荷に対する特性評価	35
4.2.2	任意の端末台数に対する特性評価	37
4.3	晒し端末環境下における評価	40

4.3.1	ネットワークトラヒック負荷に対する特性評価	40
4.4	隠れ端末および晒し端末の混在環境下における評価	41
4.4.1	ネットワークトラヒック負荷に対する特性評価	41
4.4.2	任意の端末台数に対する特性評価	44
第5章	まとめと今後の課題	46
	参考文献	48
	謝辞	53
	研究業績	54

目次

2.1	DCFにおけるチャネルアクセス	8
2.2	隠れ端末問題	10
2.3	晒し端末問題	11
2.4	強化学習の基本モデル	16
2.5	強化学習における行動選択	17
2.6	CWの状態遷移モデル [14]	20
2.7	強化学習適用チャネルアクセス例 [16]	20
3.1	1つの基地局を持つ無線LAN (リング配置)	23
3.2	1つの基地局を持つ無線LAN (ランダム配置)	24
3.3	2つの基地局を持つ無線LAN	25
3.4	2つの基地局を持つ無線LAN (ランダム配置)	26
3.5	キャリアセンス閾値の状態遷移図	28
4.1	端末台数15台におけるスループット (リング配置)	35
4.2	端末台数15台における衝突率 (リング配置)	36
4.3	平均CS閾値の時系列変化	37
4.4	選択CWの時系列変化	37
4.5	端末台数15台におけるスループット (ランダム配置)	38

4.6	端末台数 15 台における公平性 (ランダム配置)	38
4.7	任意の端末台数におけるスループット (リング配置)	39
4.8	任意の端末台数におけるスループット (ランダム配置)	39
4.9	任意の端末台数における平均 CS 閾値の収束値	40
4.10	晒し端末のみの環境におけるスループット	41
4.11	平均 CS 閾値の時系列変化	41
4.12	隠れ晒し端末混在環境下におけるスループット (端末台数 15 台) . . .	43
4.13	隠れ晒し端末混在環境下における公平性 (端末台数 15 台)	43
4.14	CS 閾値及び晒し端末によるフリーズ回数 $N'_F(\text{STA1})$	44
4.15	CS 閾値及び晒し端末によるフリーズ回数 $N'_F(\text{STA10})$	44
4.16	隠れ晒し端末混在環境下におけるスループット (任意の端末台数) . . .	45

表 目 次

3.1 CS 閾値に対する基本報酬値 R_C	29
4.1 シミュレーション諸元	33
4.2 学習パラメータ諸元	33

第1章

序論

1.1 研究背景

近年のスマートフォンなどの移動通信端末の普及や様々なネットワークアプリケーションの登場によりネットワークに接続する端末の台数は急増している。またアプリケーションの利用形態は多様化しており、幅広い分野での無線通信の需要が高まっている。無線通信に対する需要の急増に対応するため、第5世代(5G)移動通信システムが標準化、実用化されている [1]。さらに、次世代の通信システムとして第6世代(6G)の標準化に向けた研究が盛んに行われている [2], [3]。特に無線 Local Area Network (LAN) はオフィスや家庭内を中心に幅広く普及している。無線 LAN では複数の端末が同時にデータを送信することによる通信の衝突がネットワーク全体の性能劣化を招く。

現在の IEEE 802.11 規格の無線 LAN システムで広く採用されている Medium Access Control (MAC) プロトコルである IEEE 802.11 Distributed Coordination Function (DCF) では、通信の衝突を回避するために Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) を採用している [4]。CSMA/CA では、送信を行う端末が事前に近隣端末のチャンネルの使用状況を確認する「キャリアセンス (Carrier Sense: CS)」と送信開始前にランダムな待機時間を設定することにより端末ごとの送信タイミン

グを分散させる「バックオフ」により通信の衝突を回避するよう設計されている。一方で、端末間の距離や障害物の影響により、CSが機能せず通信の衝突を招く「隠れ端末問題」や、お互いのチャンネルの利用に影響の無い送信をCSすることで送信の機会が減少しチャンネルの利用率が低下する「晒し端末問題」により、通信性能が劣化する。これらの問題はCS範囲に起因しており、CS範囲が狭い場合には隠れ端末、広い場合には晒し端末の影響が大きくなるトレードオフの関係にある。これらの性能劣化に対する有効な対策としてチャンネルを事前に予約し、周囲に通知するRTS/CTS [5] や各端末を識別して通信が競合するかを判断するBSS Coloring [6] がある。しかし、これらの技術を適用するためには無線LANの仕様を変更する必要があり、容易に適用することはできない。一方で、システムを制御するパラメータを適切に設定することで、無線LANの仕様を変更せずに性能劣化を軽減する手法が検討されている。隠れ端末および晒し端末の影響を把握し、性能劣化を軽減できるよう制御パラメータを設定することで効率的な通信を行うことが可能である。

制御パラメータの設定に関する研究はこれまで主に、数理モデルを用いた手法が多く提案されてきた [7]- [13]。プロトコルの制御パラメータがシステム性能に与える影響を数式として得ることで、システムの振る舞いの把握だけでなく、システムの最適化への応用も可能である。文献 [7] では、DCFにおいて解析モデルを用いてCS範囲を調整する手法を提案している。隠れ端末および晒し端末の影響を数理的に導出することで、隠れ端末による通信衝突を回避しつつ晒し端末によるチャンネル利用率の低下を軽減するようなCS範囲を設定し、スループットを向上している。一方、数理モデルにはいくつかの理想化が含まれるため、一般的に想定されるネットワーク環境あるいは特定の状況下での対応が主であり、あらゆるネットワーク環境に対し柔軟に対応することは困難である。

一方、環境に依らず柔軟にシステムの最適化を行う手法として、近年強化学習が注

目されている。強化学習では、学習エージェントが与えられた環境(状態)において最適な方策を見つけるために、行動を選択実行する。選択した行動の試行結果の統計から傾向を見出し最適な方策を決定する。この強化学習の概念を IEEE 802.11DCF の制御に応用した手法が提案されている。DCF おける制御パラメータの最適化手法の一つとして、Contention Window (CW) の最適化が提案されている [14]- [16]。CW はバックオフにおける待機期間の長さを決定するパラメータであり、CW が小さい場合、送信タイミングが十分に分散されず、送信タイミングが重なり通信衝突が発生する。一方、CW が大きい場合、待機期間が長くなり送信機会が減少する。このように、CW にもトレードオフの関係がある。文献 [14]- [16] では、学習エージェントを搭載した個々の端末が、送信の成功・失敗の結果から、通信が競合する端末台数に応じて、通信衝突の回避と送信機会の確保を両立するような CW を学習に基づいて自律的に制御することで、ネットワーク全体の性能向上が得られることを示している。しかし、これらの既存研究は CS により周辺端末の送信が全て検知可能なネットワークを想定しており、CS による検知に起因して生じる隠れ端末や晒し端末が存在するネットワーク環境下では、十分にその性能を発揮することができないと考えられる。

1.2 本研究の目的

本研究では、IEEE 802.11 無線 LAN における隠れ端末による通信衝突および晒し端末によるチャネル利用率の低下を改善し、ネットワークスループットを向上することを目的とする。この目的を達成するために、本論文では、IEEE 802.11 ネットワークにおける強化学習に基づく CS 閾値制御を提案する。個々の端末が周囲の隠れ端末および晒し端末の影響に応じて自律的に CS 閾値を強化学習に基づいて探索して設定する。その際に、CS 閾値の状態遷移モデルと学習のための報酬付与方針を新たに提

案する. 計算機シミュレーションの結果より, 各端末が適切に CS 閾値を学習により設定することで, 個々の端末における隠れ端末および晒し端末の影響を軽減し, ネットワークスループット性能の向上が得られることを示して本提案手法の有効性を確認する.

1.3 論文構成

本論文の構成として, 第2章では, 無線 LAN の技術的動向について触れた後, IEEE 802.11 の基本的な Medium Access Control (MAC) プロトコルである Distributed Coordination Function (DCF) の動作とその課題点について論述する. さらに, 近年注目を集めている機械学習の概要と通信プロトコル設計への応用例を示し, 本提案手法で用いる強化学習の概要と具体的な学習手順及び通信プロトコル設計への応用の関連研究を示す. 第3章では, 本研究で想定するシステムモデルと本提案手法の動作を詳説する. 第4章では, 本提案手法を用いた結果と, 従来手法との比較からその有効性の検討を計算機シミュレーションによる結果から行う. 第5章では, 本論文のまとめと今後の課題について記す.

第2章

従来研究

本章では, 2.1 節に IEEE 802.11 無線 LAN における技術的動向と MAC 層の基礎技術であり現在も広く採用されている IEEE 802.11 DCF の動作および課題点について記述する. 2.2 では, 通信分野への技術的応用が注目される機械学習の概要およびその分類の一つである強化学習を用いた MAC プロトコル設計に関する研究と課題について記述する.

2.1 IEEE 802.11 無線 LAN

2.1.1 無線 LAN の技術的動向

無線 LAN とは, 有線のケーブルを用いずに無線信号によりネットワークに接続する技術である. 1997 年に無線 LAN の最初の標準規格となる IEEE 802.11 [4] が技術標準化機関 Institute of Electrical and Electronics Engineers (IEEE) により策定された. 初期規格の策定後, IEEE 802.11 は多数の追加規格により, 高速化・高機能化を実現している. IEEE 802.11 規格の拡張は物理層と MAC 層について行われており, 物理層ではそれぞれの規格ごとに様々な技術が用いられている. 一般的に普及している通信規格の内, 11a/b/g では, 相補符号変調方式や直行周波数分割多重 (Orthogonal Frequency Division Multiplexing: OFDM) 方式 [17] を採用することで, 使用するそ

それぞれの周波数帯域における伝送速度の高速化を実現している。続く 11n では、複数のアンテナを用いて送受信を行う Multiple-Input and Multiple-Output (MIMO) [18] や複数のチャネルを結合するチャネルボンディングが用いられた。11ac では、周波数帯域の拡張、変調方式の改善に加えダウンリンクにおけるマルチユーザ対応の Multi User MIMO (MU-MIMO) [19] を導入している。さらに 11ax では、アップリンク、ダウンリンク共に MU-MIMO における同時接続が可能となり、伝送効率を大幅に改善している。さらに、端末が所属する AP を識別する Basic Service Set Coloring (BSS Coloring) により、異なる AP に所属する近隣端末の通信の影響を低減し、周波数リソースの空間的利用効率を向上している。そして現在標準化中である。11be では、変調方式の更なる改善等が行われ、周波数の利用率向上、広帯域化が行われている。11be の標準化により、従来は困難であった伝送遅延や尊号伝送時間のズレによる映像・音声等の乱れの改善が見込まれ、拡張現実 (Augmented Reality : AR)/仮想現実 (Virtual Reality : VR) や産業向けアプリケーションといったリアルタイム性が要求される利用状況への展開が見込まれる。また、現在の 11be の標準化と並行して 11be に次ぐ規格となる 11bn の規格策定に向けた議論が進められている。11be で規定されたマルチリンク伝送が単一の AP 内の実装を想定していたのに対し、複数の AP が協調・連携を行う複数 AP 協調機能や 11be で規定される Multiple Resource Unit (MRU) を拡張させた distributed RU (dRU) 等の様々な技術が標準化に向け議論されている。

物理層における様々な技術的拡張に対して、MAC 層の技術的特徴として、有線のイーサネットで使用されているアクセス制御方式である衝突検出機能付きキャリア感知多重アクセス (Carrier Sense Multiple Access with Collision Detection: CSMA/CD) [20] を踏襲した衝突回避機能付きキャリア感知多重アクセス (Carrier Sense Multiple Access with Collision Avoidance: CSMA/CA) 方式が採用されている点が挙げられる。CSMA/CA 方式は各端末が自律的にチャネルの使用状況を検知し

て通信を行うため、制御が簡単かつ移動通信システムにおいて効率的なアクセス制御方式であり、この後の一連の無線LAN発展の基礎をなす技術である。CSMA/CAの基本動作については2.1.2節にて詳しく解説する。初期規格から11gまではMAC層の技術に関する拡張は特に行われず、続く11nで伝送効率を改善するMAC層の技術として、複数のフレームを連結させ、フレーム伝送の効率を向上するフレームアグリゲーションや送信側からの複数のフレームに対し1つの確認応答(ACK)を返すブロックアックが用いられ、それまでの規格に比べてMAC層の改善が行われている。11axでは、無線基地局が端末のアップリンク送信を制御するために、事前にチャンネルを予約するRequest-to-Send(RTS)を含むいくつかのトリガーフレームが導入されている。そして11beでは、Multi Link Device(MLD)と呼ばれる一つの筐体の実装された複数の無線機能を協調・連携させ、MLD間で伝送路を複数確立し高速・高信頼な伝送を可能にするマルチリンク伝送技術が規定されている。これらMAC層の改善により、通信効率は向上している。一方で、基礎技術であるCSMA/CA方式に関しては初期規格から大きな改善はされておらず、無線LANの需要や利用形態が変化している現在において完全に対応することは困難である。

2.1.2 IEEE 802.11 DCF

IEEE 802.11無線LANでは、同一の無線チャンネルを複数端末で共有するためのアクセス制御方式として、CSMA/CAを採用したIEEE 802.11 DCF [4]が使用されている。DCFではフレームを保持している端末はフレーム送信前にキャリアセンス(Carrier Sense: CS)を行うことでチャンネルの使用状況を確認する。CSを行うにあたり、受信信号の電力レベルを用いてチャンネルの使用状況を判断するCS閾値が設定される。CS閾値未満の電力レベルであれば、チャンネルは未使用(Idle: アイドル状態)で

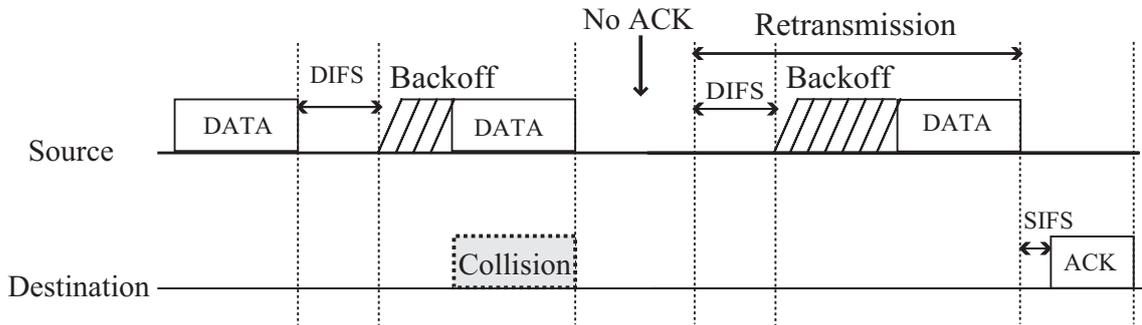


図 2.1: DCF におけるチャネルアクセス

あると判断し、送信待機期間 (バックオフタイム) のカウントを開始する。一方、キャリアセンス閾値以上の電力レベルが検出された場合、端末はチャネルが使用中 (Busy: ビジー状態) であると判断し、送信を延期して送信を控えることでフレーム衝突を回避する。

図 2.1 に、IEEE 802.11 DCF におけるチャネルアクセスの例を示す。チャネルが DIFS 期間だけアイドル状態になった後、フレーム送信を行う端末は、規定の乱数発生範囲 (Contention Window: CW) 内で乱数を発生させ、乱数値をもとにしたランダムなバックオフタイムを選択する。チャネルがアイドル状態のときにバックオフタイムを 1 ずつ減算していき、0 となった端末が送信を行う。データ送信端末は送信後に ACK を受信することで送信成功と判断する。同時にバックオフタイムのカウントを 0 にする端末が存在する場合、同時に送信が行われ、受信側で信号を復調できず送信フレームの衝突となり再送信を行う。DCF では再送信におけるバックオフ制御に 2 進指数バックオフが用いられている。フレーム衝突が発生した際に、再送ごとに CW を 2 倍に増加させ、 $[0, W_{N_{retx}}]$ の範囲からバックオフタイムを選択し直すことで、再送信のフレーム再衝突確率を低減させる。ここで、 $W_{N_{retx}}$ を N_{retx} 回目の再

送時における CW 値とすると

$$W_{N_{retx}} = \min(2^{N_{retx}} CW_{min}, CW_{max}), N_{retx} \in [0, m] \quad (2.1)$$

と表せる. ここで m は最大再送回数, CW_{min} および CW_{max} はそれぞれ CW 値の最小値, 最大値である. フレームの衝突による再送ごとに CW 値は指数関数 (2 進指数) で増加する. CW 値が CW_{max} に達した場合は, CW 値は CW_{max} を維持しつつ最大再送回数 m となるまで送信を試行し, m 回再送に失敗したフレームは破棄される. 送信が成功した場合は, CW 値を初期化し, 回目の送信フレームに対して CW_{min} の範囲からランダムにバックオフタイムを選択する.

2.1.3 隠れ端末および晒し端末問題

DCF における通信性能の劣化を引き起こす要因である「隠れ端末問題」と「晒し端末問題」について記述する. DCF において, 端末間の距離が離れている場合や電波を通さない障害物の影響によってお互いの無線信号が到達しない (CS による送信検知が機能しない) 状態が発生する可能性がある. これらの関係にある 2 つの端末を互いに隠れ端末と呼ぶ. 図 2.2 に 1 台の基地局 (Access Point: AP) と互いに隠れ端末の関係にある端末を含めた 3 台の STA で構成するネットワークを示す.

また, 複数の AP およびそれらに接続された端末が高密度に配置される環境において, 同一チャンネルで異なる送信宛先に対する送信を CS することでチャンネルがビジー状態となり, 必要のない送信待機をしてしまう可能性がある. これらの関係にある 2 つの端末を互いに晒し端末と呼ぶ. 図 2.3 に 2 台の AP と互いに晒し端末の関係にある端末を含めた 3 台の STA で構成するネットワークを示す. STA0 と STA1 は AP1 に接続されており, STA2 は AP2 に接続されている. STA1 は 2 つの隣接する AP の中間付近に配置する. 図 2.3 より, STA1 と STA2 はそれぞれ別の宛先にフレーム送

STA0 と STA1 は互いに CS により、送信を検知可能であるのに対し、STA0 と STA2 は互いの隠れ端末の関係にある。そのため、STA2 が送信を開始しても STA0 はバックオフタイマを停止せずカウントを続ける。STA2 のフレーム送信中に STA0 のバックオフタイマが 0 になり、送信を開始することで通信の衝突が発生する。これらの隠れ端末による通信の衝突によってスループット性能が劣化する問題を隠れ端末問題と呼ぶ。

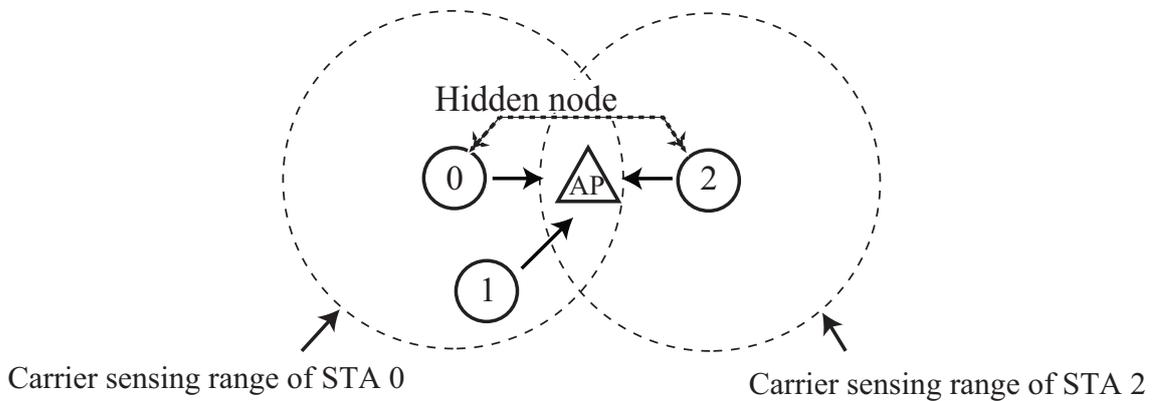


図 2.2: 隠れ端末問題

信するが、互いの送信を検知する晒し端末の関係にある。そのため、STA2 のバックオフタイマが 0 になり、フレーム送信を開始すると STA1 は STA2 の送信を CS して自身のバックオフタイマのカウントを停止し、送信待機することで STA2 とのフレーム衝突を回避しようとする。実際には STA1 と STA2 の送信宛先は異なるため送信待機する必要は無く、同時に送信可能である。このように、余分な送信待機によりチャネルの空間利用率が低下する問題を晒し端末問題と呼ぶ。

DCF において隠れ端末による性能劣化を防ぐには、CW を大きく設定して通信衝突を回避する、CS 閾値を小さく設定して検知範囲を広げ隠れ端末を無くすといった方法がある。また、晒し端末による性能劣化を防ぐには、CS 閾値を大きく設定し

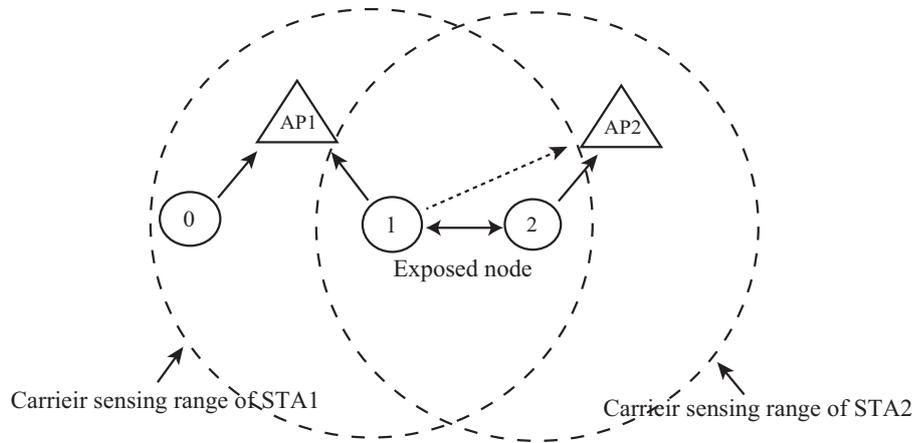


図 2.3: 晒し端末問題

て検知範囲を狭め晒し端末を無くすといった方法がある。このように、システムを制御するパラメータを適切に設定することで通信性能の劣化を防ぐことが可能である。しかし、隠れ端末や晒し端末の影響はネットワークごとに異なり、適切な制御パラメータの設定値も異なる。そのため、ネットワーク環境に応じて制御パラメータを適切に調整・最適化することが望ましいと考えられる。

2.1.4 数理モデルに基づくネットワークシステムの最適化

制御パラメータの最適化に関する研究としてこれまで主に、システムの動作を数理モデルに表現したうえで、その性能を評価する研究が多く提案されている [7]- [13].

文献 [9] では、マルコフ連鎖モデルを用いて IEEE 802.11 ネットワークにおける CSMA/CA の動作を理論的に表現している。マルコフ連鎖とは、未来の挙動は現在の値でのみ決定され過去の挙動は無関係である確率過程 (マルコフ過程) のうち、とり得る状態が離散的なものをいう。この理論解析モデルでは、解析の仮定としてネットワーク内の各端末は全ての他端末の送信を検知可能である (隠れ端末無し)

とし、全ての端末が送信バッファ内に常にフレームを保持している飽和状態であるとしたうえで、1台のAPと n 台の端末で構成されるネットワークにおいてアップリンクにおける最大スループットを理論的に導出している。文献 [7] では、ネットワークトポロジ、受信電力、データレートを考慮し、適切なCS閾値を導出するための解析モデルを提案している。導出されたCS閾値により、端末が静的なネットワークにおいて隠れ端末、晒し端末による性能劣化を改善している。また、文献 [8] では、多数の端末が高密度に配置される環境を想定し、APが配下の端末数を推定し、端末数に応じたCWを数理モデルに基づいて導出している。

数理モデルを構築する利点として、プロトコルの制御パラメータとシステム性能に与える影響を数式として表現することにより、その動作とシステムの振舞いの直感的な理解だけでなく、システムの最適化への応用も可能である。一方、数理モデルにはいくつかの理想化が含まれるため、一般的に想定されるネットワーク環境あるいは特定の状況下での対応が主であり、あらゆるネットワーク環境に対して柔軟に対応することは困難である。

2.2 機械学習の通信プロトコル設計への応用

本節では、機械学習に基づくMACプロトコル設計について記述する。2.2.1節で機械学習の学習技法における分類とそれぞれの学習技法の概要を紹介する。2.2.2節では、機械学習の学習技法の一つである強化学習の学習手順について詳説し、2.2.3節でDCFにおける強化学習に基づくシステムの最適化に関する研究を紹介したうえで既存研究の課題点について述べる。

2.2.1 機械学習

機械学習は Machine learning (ML) とも呼ばれる人工知能 (Artificial Intelligence: AI) を支える技術の1つである。機械やコンピュータに対象とするデータを読み込んで学習させ、規則や法則を導くことにより、適した行動を選択させる。AI や機械学習の研究は数多く行われており、近年、情報通信技術 (Information and Communication Technology: ICT) の実用化やビッグデータの環境が整いつつあり、機械学習は実用化されつつある。無線通信の分野においても機械学習の応用は幅広く研究され目的に応じた様々な手法が検討されている。機械学習は学習方法の違いにより、「教師あり学習」、「教師なし」、「強化学習」に大きく分類できる。

教師あり学習は、あらかじめ正解となる結果をデータとして学習させることで、入力データに対して正解となる出力を行う学習方法である。そのため、正解や最適となる答えが明確である問題に対して非常に有効であり、主にビッグデータを利用した予測やシステム異常の検知、特定に利用される。文献 [21] では、教師あり学習を用いて、ソーシャルネットワークにおける将来的なリンク予測を行っている。教師あり学習にて科学論文データの特徴セットを特定する。文献 [21] は、教師あり学習によるクラス予測が一般的な分類アルゴリズムと同等以上に可能であることを示している。教師あり学習は、人間が正解となるデータを与えるため、学習精度が高く、学習速度も早いという利点を持つ。また、学習に用いることのできる過去データが多いほど学習精度が高くなる傾向にある。一方、正解となるデータを与える必要があるため、正解となるデータの質が学習精度に大きく影響する。加えて、学習精度を高めるために多数のデータセットを準備する必要があり、教師あり学習を行うまでに多くの時間や手間がかかる。

教師なし学習は、正解となるデータが存在せず、入力されたデータを利用して正解

となる出力を導き出す学習方法である。教師なし学習では予測や判定の対象となる正解が存在しないため、大量のデータを学習させることでデータの特徴や法則を記憶していき、それが正解か否かを自身が判断する。文献 [22] では、教師なし学習に基づくネットワーク脅威状況評価モデルを提案している。IoT システムに対する攻撃的な脅威による異常閾値を教師なし学習により計算し、ネットワーク脅威をより直観的かつ効果的に評価している。教師なし学習は、データに正解となるデータを用意する必要が無いため、教師あり学習と比べて学習の開始が容易である。そのため、教師あり学習に比べて、より効率的で時間やコストを消費せずに最適な学習結果を得られる可能性がある。一方、正解となる学習データが無いため、学習結果の精度は低くなる傾向にあり、人間が想定できない新たなパターンを見つけ出せる反面、学習結果が求められるものと異なる場合が考えられる。

正解・不正解のラベル付けの有無に関わらず、過去のデータを参照して学習を行う。対して強化学習は、出力される結果に報酬を与えることで、最も望ましい結果を出すための行動を学習する。行動に対する報酬のみを考慮し、報酬をいかに最大化するかを判断して自身の行動を選択する。そのため、教師あり学習や教師なし学習とは異なり、強化学習は事前に準備するデータセットに依存せず、収集したデータを試行錯誤し学習を続けることにより、環境に応じた出力を行う。このことから、強化学習は2.1.4節で紹介した数理モデルに基づく手法における課題であった様々なネットワーク環境における柔軟な対応が可能である。また、従来の強化学習にディープラーニングを応用した深層強化学習を用いた研究も盛んに行われている [23]- [25]。ディープラーニングは機械学習の一手法である。従来の機械学習では、学習データのどの部分が出力にどの程度の影響を与えているかを示す「特徴量」を人間が自ら調整する「特徴量抽出」を経た上で、機械学習により予測モデルを構築する。対して、ディープラーニングでは、学習端末自身が学習データから特徴量を抽出し、予測モデルを構

築する「表現学習」が可能である。入力に対して重み付けして出力を返すニューロンという層を何層にも重ねたニューラルネットワークにより、データが各層で処理されていく中で、データの持つ特徴を判断する。ニューラルネットワークにより得たデータの特徴に対し、報酬を最大化する行動を選択することにより、精度の高い学習を行うことが可能である。深層強化学習は、大量のデータを用いて精度の高い学習を行うことが可能である。一方で、従来の強化学習に比べて、学習端末は非常に大量のデータを蓄積する必要がある、高い処理能力が求められる。そこで本研究は、学習端末の処理能力に比較的依存せず、あらゆるネットワーク環境においても適用可能な強化学習に基づく制御パラメータの制御に注目する。強化学習の具体的な学習手順および強化学習を応用した関連研究については2.2.2節、2.2.3節にて詳説する。

2.2.2 強化学習

図2.4に強化学習の基本モデルを示す。強化学習において、実際に学習を行う学習エージェントには環境から入力される状態(State)とそれぞれの状態で行う行動(Action)のペアの集合が与えられる。図2.4より、学習エージェントは最初に環境からの情報に基づき現在の状態を固定する。そして、その状態に対して行動を選択し、実際に環境に対して実行する。この行動選択は方策(Policy)と呼ばれる行動知識によって決定され、価値が最も高いものを選択する、それぞれの選択肢の価値に比例した確率で選択するなど様々な方法がある。行動を実行した結果に応じて、環境から報酬(Reward)が得られる。得られた報酬と実行された行動などを基にして、それぞれの選択の価値を更新していくことで学習を行う[26]。学習エージェントは得られる報酬を最大化するような行動を学習することにより最適な選択を行うようになる。2.2.2.1節では、本研究で採用する強化学習の代表的なアルゴリズムであるQ学習の

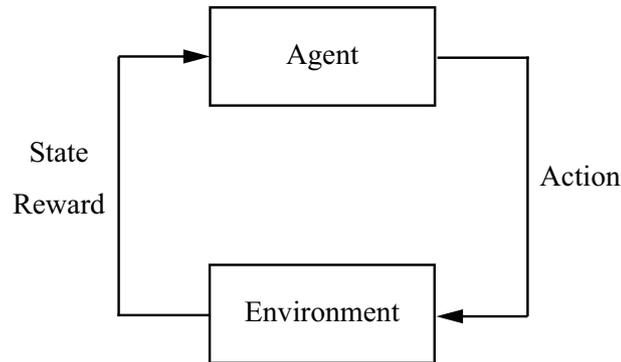


図 2.4: 強化学習の基本モデル

基本的な学習手順について記述する.

2.2.2.1 Q 学習

Q 学習では, 行動選択における各選択の評価を Q 値と呼ばれる評価値に基づいて決定する. ある状態 s において次の行動 a を採用することによる評価値 $Q(s, a)$ を設定し, 設定した Q 値に従い行動を選択することで報酬 r を得る. Q 値を適切に求めることで, 各状態において Q 値の最大値を与える行動 a を選択し, 適切な行動を可能とする. そのため, エージェントが選択した行動に対して得た報酬に基づいて Q 値の更新作業を行うことで学習を進める.

Q 値の更新式はエージェント自身の状態 s , 選択する行動 a , 行動により得られる報酬 r , 及び学習のスピードを決定する学習係数 α を用いて

$$Q(s, a) \leftarrow Q(s, a) + \alpha r \quad (2.2)$$

で表す. しかし, 式 (2.2) をそのまま適用し続けると, 学習を繰り返し Q 値が無限大に発散してしまう. この発散問題に対して一般的に減算項を追加することで発散を

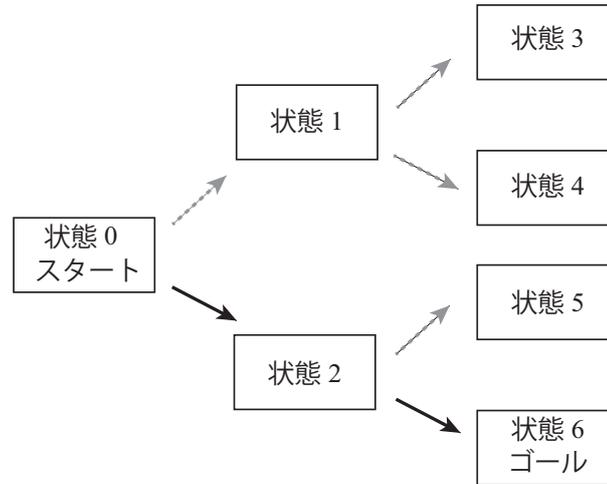


図 2.5: 強化学習における行動選択

回避する. このとき, 報酬を得た場合の Q 値の更新式は

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r - Q(s, a)) \quad (2.3)$$

と表される. また, 強化学習では, 行動選択に対して報酬が遅れて得られる場合がある. このときの例として強化学習における行動選択を図 2.5 に示す. 学習エージェントは状態 0 から行動選択を開始し, 状態 6 に到達した場合に報酬が与えられるとする. このとき, 状態 6 に向かうための行動選択として状態 0→2 を選択してもすぐに報酬を得ることが出来ず, その後状態 2→6 の選択に対して報酬を得る. このように, すぐに報酬が得られない状況では, 報酬を得られるゴールに到達した場合のみではなく, ゴールに続く行動を選択すれば Q 値が増加し, それ以外の行動を選択すると Q 値が増加しないようにする必要がある. このとき, Q 値の更新式は

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma \max Q(s_{next}, a_{next}) - Q(s, a)) \quad (2.4)$$

と表せる. ここで, γ は割引率であり, ゴールに続く選択にどの程度の価値を与えるかを決定する. また, $\max Q(s_{next}, a_{next})$ は次の状態で取りうる行動に対する Q 値の

最大値である。学習エージェントは式 (2.4) より、Q 値を更新・改善することで最適な行動を学習する。Q 学習のアルゴリズムを以下にまとめる。

1. 行動の初期状態に戻る
2. 次の状態に至る行動を Q 値に基づいて選択
3. 式 (2.4) に基づいて Q 値を更新
4. 選択した行動によって次の状態に遷移
5. 目標状態に至るか、あらかじめ決められた回数の行動選択を終えると初期状態に戻る
6. 5 の条件に当てはまらない場合、2 に戻る

2.2.3 強化学習の MAC プロトコル設計への応用

強化学習では、学習エージェントがその環境 (状態) において最適な方策を見つけるために、行動を選択実行する。選択した行動の試行結果の統計から傾向を見出し最適な方策を実行する。近年、この強化学習の概念を MAC プロトコル設計に応用した様々な研究が行われている。

文献 [27] では、無線センサーネットワークにおいて強化学習に基づいた適応型 MAC プロトコルである RL-MAC を提案している。各端末が各フレームの最初にキューに入れられたパケット数に応じて、データの送受信を行うアクティブ期間と送信を控えるスリープ期間の割合を自律的に制御することで、様々なトラフィック条件に応じて高いスループットと低消費電力を実現している。文献 [28] では、継続的に変化し続ける水中環境において、各端末がフレーム内でデータ送信を行うスロットを強化

学習に基づいて適応的に選択することで、通信衝突を回避し、チャンネル効率を向上している。強化学習の MAC プロトコル設計への適用はネットワークやアクセス方式の種類を問わず幅広く行われ、DCF に対する適用も検討されている。

文献 [29] では、IEEE 802.11DCF において AP に搭載された学習エージェントがネットワークの衝突率やスループットから CW の最適値を学習し周囲の端末に通知することでスループットの向上を図っている。AP が、ネットワーク全体の衝突率の時系列情報を記録し、それらから AP がネットワークにおける最適な CWmin を選択し、配下の端末へ通知することで、衝突を回避し、スループット性能の向上を得ている。しかし、個々の端末は、周囲の端末台数、トラヒック状況によって通信環境が異なる。そのため、全ての端末で同じ CWmin の値を共有するよう制御する手法では、端末ごとに衝突率が異なるようなネットワークに対して、適切な効果は得られないと考えられ、適用可能なネットワークは限定的である。

文献 [14]- [16] では、それぞれの端末に学習エージェントを搭載させ、各端末が送信成功、失敗の試行結果から自律的に CW の学習を行い、個々に適切な値を探索・設定する手法を提案している。文献 [14] では、車車間通信を想定した IEEE 802.11p を用いた車両アドホックネットワーク (Vehicular Ad Hoc Network: VANET) を想定しており、図 2.6 に示すような CW の現在の値を強化学習における「状態」、CW を 2 倍する (Increase: I), 0.5 倍にする (Reduce: R), 維持する (Keep: K) を各状態における「行動」とした状態遷移モデルを用いている。このモデルでは、ACK の受信の有無をもとに現在 CW の値において選択した行動に対し、報酬を与える。報酬は ACK を正常に受信した場合には正、それ以外の場合は負の報酬として、更新式に基づいて、Q 値を更新していく。これにより、各状態において一番報酬を多く得ることができる行動、すなわち ACK を高確率で受信できる CW を端末が自律的に設定できるようにできる。このとき、状態、すなわち現在の CW に応じた報酬値を設定する。

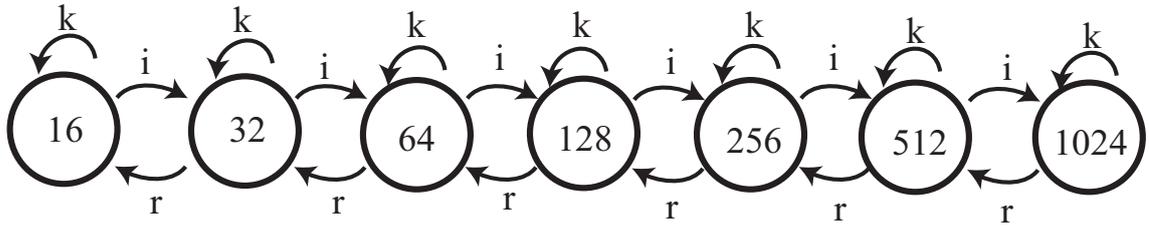


図 2.6: CW の状態遷移モデル [14]

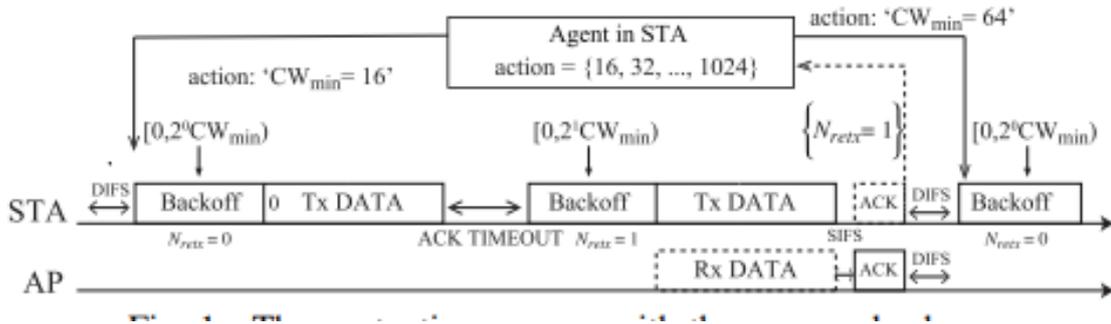


図 2.7: 強化学習適用チャネルアクセス例 [16]

文献 [14] では、CW の値が小さいほど報酬値の絶対値を大きくすることにより、それぞれの端末は正の報酬かつ可能な限り大きい報酬を得るために、CW を過度に大きくしないような選択を行う。

文献 [15], [16] では、強化学習における状態遷移は考慮せず、CWmin の集合の中から適切な CWmin を学習を用いて直接設定している。図 2.7 に文献 [16] で提案されている学習制御手法におけるチャネルアクセス例を示す。図 2.7 より、STA は再送回数 N_{retx} に従い、バックオフタイム W_i は

$$W_i \in [0, 2^{N_{retx}} CW_{min}] \quad (2.5)$$

から決定される。送信に対して ACK フレームの返答がない場合、 N_{retx} を 1 加算することで、 W_i の選択範囲を広域化し、再送信を行う。ACK フレームを受信すると

学習エージェントはACK フレーム受信するまでの再送回数 N_{retx} と選択している CWmin をもとに報酬値を決定する。報酬値は以下のように決定される。

$$R(a_{cur}) = \begin{cases} -\frac{\min(\mathcal{A})}{a_{cur}}, & \text{if } N_{retx} \text{ is larger than } \theta \\ \frac{\min(\mathcal{A})}{a_{cur}}, & \text{otherwise} \end{cases}. \quad (2.6)$$

ここで、 $\min(\mathcal{A})$ は CWmin の集合の最小値であり、 a_{cur} は現在の CWmin である。再送回数が閾値 θ を越えた場合、現在の CWmin をもとにした負の報酬を与え、選択確率を下げることで最適な CWmin を探索する。このようにすることで、強化学習における状態数を削減し、文献 [14] と同等の性能が得られることを示している。

これらの研究はいずれも CS により周辺端末の送信が検知可能なネットワークを想定し、DCF のバックオフに関する制御パラメータである CW を強化学習に基づいて適切に調整・最適化することでネットワーク全体の性能向上が得られることを示している。一方で、CS による検知に起因して生じる隠れ端末や晒し端末が存在するネットワーク環境下では、十分にその性能を発揮することができないと考えられる。

第3章

提案手法

本章では, IEEE 802.11 WLAN における強化学習を用いた CS 閾値制御の学習フレームワークについて記述する. 3.1 節では本論文で想定するシステムモデル, 3.2, ??節ではそれぞれ, 隠れ端末問題, 晒し端末問題改善のための学習方針の設定および学習フレームワークの一連の流れについて述べる.

3.1 システムモデル

3.1.1 ネットワークモデル

本論文では, 隠れ端末, 晒し端末による無線通信に対する影響に対して提案手法を評価する. そのため, 隠れ端末, 晒し端末それぞれの影響が確認できるような異なるネットワークを想定し, それぞれの問題を考慮した CS 閾値の学習制御手法の評価を行う.

3.1.1.1 隠れ端末を想定したネットワーク

全ての端末に等しく隠れ端末の影響がある環境での評価を行うために図 3.1 に 1 つの AP を持つリング型のネットワーク例を示す. AP を中心とする半径 r [m] の円

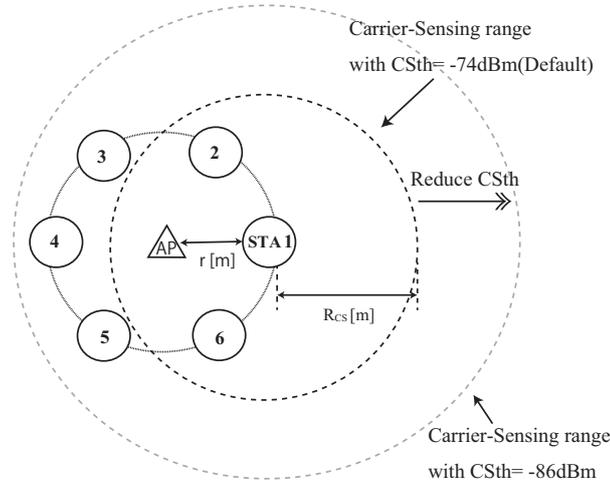


図 3.1: 1つの基地局を持つ無線 LAN (リング配置)

周上に等間隔に n 台の端末を配置する. 各端末の CS 範囲をそれぞれ R_{CS} とし, R_{CS} は選択する CS 閾値によって変化する. 各端末は R_{CS} の範囲内の端末の送信を検知できる.

次に, 端末毎に隠れ端末の影響が異なる環境での評価を行うために図 3.2 に 1つの AP を持つランダム型のネットワーク例を示す. AP を中心とする半径 r [m] の円内に n 台の端末を一様ランダムに配置する.

3.1.1.2 晒し端末を想定したネットワーク

晒し端末がシステム性能に与える影響のみを評価するために図 3.3 に想定した 2つの AP を持つネットワーク例を示す. 1台の AP に対して1台の端末のみが接続する基本サービスセット (Basic Service Set: BSS) が2つ隣接する. AP 間の距離は r_{AP} [m] だけ離れており, 各端末は2台の AP 間の直線上かつ接続する AP から r [m] の距離に配置する. BSS 内での競合端末が存在しないため通信衝突は発生せず, 晒し

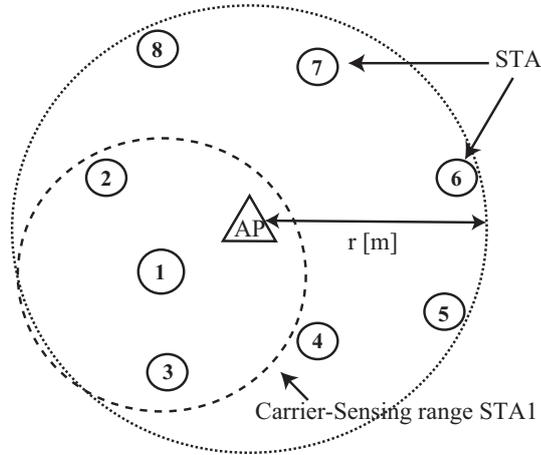


図 3.2: 1つの基地局を持つ無線 LAN (ランダム配置)

端末の影響でのみ性能が劣化する.

3.1.1.3 隠れ端末および晒し端末の混在想定したネットワーク

隠れ端末および晒し端末が混在する環境での評価を行うために図 3.4 に 2 つの AP を持つランダム型ネットワーク例を示す. ネットワーク全体の端末数を n 台とし, それぞれの AP の通信可能範囲内に各 AP に接続する端末をランダムに配置する. 配置のランダム性により, 端末ごとの競合端末, 隠れ端末, 晒し端末の数は異なる. また, 一部の端末では, 複数の隠れ端末と晒し端末が混在する状況が想定される

3.1.2 トラフィックモデル

ネットワークにおける全端末は, 固定長ペイロード P のデータを生起する. ネットワークのオフアドロードを O [Mbps] とし, 各端末の packets 生起率は等しいと

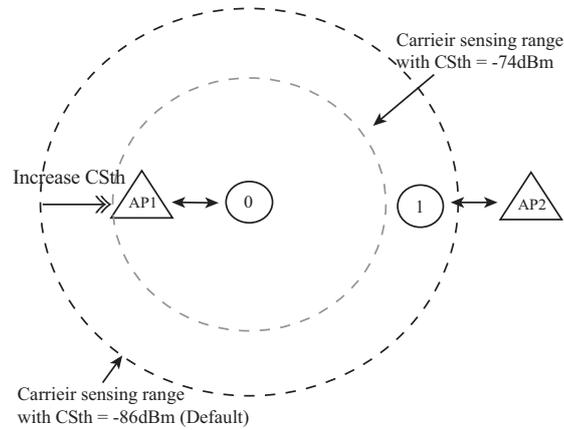


図 3.3: 2つの基地局を持つ無線 LAN

する. このとき, 端末の生起率は $\frac{O}{nP}$ である. また, 本研究では端末から AP へのアップリンク通信のみを想定する.

3.1.3 チャネルモデル

本研究では, 伝搬路における距離減衰のみを考慮し, フェージングによるチャネル変動は考慮しない.

3.1.4 アクセスモデル

本研究では, それぞれの端末は CSMA/CA に基づいてチャネルアクセスを行う. CSMA/CA におけるバックオフ制御は, IEEE 802.11 DCF に基づき, 再送信時に CW を 2 倍にする指数バックオフに基づいて待機時間を決定する.

3.2 提案CS閾値制御

3.2.1 本提案手法の概要

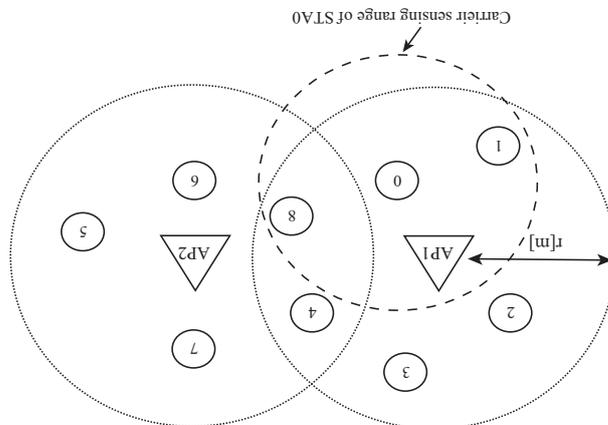
本提案手法では、通信に参加する各端末に強化学習の学習エージェントを搭載し、個々の端末が自身の動作を最適化することでネットワーク全体の最適化を行う。強化学習を用いてCS閾値を適切に探索・設定することで、あらゆるネットワーク環境における隠れ端末による通信衝突および晒し端末によるチャネル利用率の低下への対応による性能改善を可能とすることを目標とする。

3.2.2 本提案手法におけるチャネルアクセス

本提案手法におけるチャネルアクセスは2.3節で紹介した文献 [16] を参考にす。以下にチャネルアクセス中の学習手順を示す。

1. 全ての学習エージェントはCS閾値を初期値に設定し、情報通信を開始する。

図 3.4: 2つの基地局を持つ無線LAN (ランダム配置)



2. 送信宛先 (AP) からの ACK フレームが受信できない場合, 再送回数が最大値に達していなければ, 再送回数 N_{retx} を 1 加算し, 指数バックオフに基づき CW を 2 倍にして再送信を行う.
3. 送信宛先 (AP) からの ACK フレーム受信後, 報酬付与方針に従って決定された報酬値を獲得し, 学習エージェントに与える. このとき, 再送回数を 0 に戻し, CW を初期値に設定する.
4. 学習エージェントは獲得した報酬から方策の改善 (Q 値の更新) を行い, 再度行動を選択する.
5. 2~4 を繰り返すことで累積報酬が最大となるような CS 閾値を導出する.

3.2.3 学習モデル

本提案手法における学習制御の構成要素の設定を示す.

3.2.3.1 状態: S

本提案手法では, CS 閾値の選択候補を $C = \{-74, -78, -82, -86\}$ とし, 学習モデルにおける状態を $\{S(C_{cur})\}, (C_{cur} \in C)$ として定義する. C_{cur} は行動選択によって選ばれたの CS 閾値の現在値である.

3.2.3.2 行動: A

本提案手法では, それぞれの状態において, (1) CS 閾値を 4[dBm] 増加 (i_C), (2) CS 閾値を維持 (k), (3) CS 閾値を 4[dBm] 減少 (r_C) の 3 つの行動からいずれかを選択

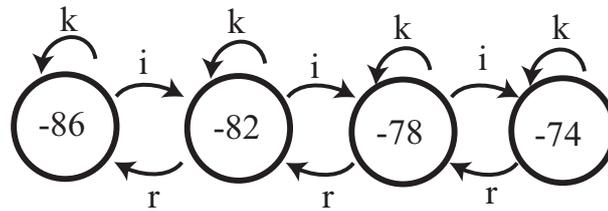


図 3.5: キャリアセンス閾値の状態遷移図

する. 本提案手法では, '状態' として定義した現在の CS 閾値を選択した'行動' に従って遷移させることで, 周囲の通信状況に応じて適切な CS 閾値を探索・設定する. 図 3.5 に本提案手法におけるキャリアセンス閾値の状態遷移図を示す.

3.2.3.3 報酬

提案学習手法における, 報酬付与について詳説する. 本提案手法では, 行動選択に対する報酬値は「現在の CS 閾値」, 「ACK フレーム受信時の再送回数」, 「CS した他端末の宛先割合」によって決定する. ここで, CS した他端末の宛先割合とは ACK フレームが正常に受信されるまでに, CS によって他端末の送信を検知し, バックオフタイマを停止した回数 N_F の内, 異なる宛先への送信に対して CS を行った回数 N'_F の割合 N'_F/N_F と定義する. 各端末において N'_F/N_F は周囲の晒し端末の影響の度合いを示し, N'_F/N_F が割合閾値 θ 以上であれば晒し端末による性能劣化が生じていると判断する. はじめに, 現在の CS 閾値に応じた基本報酬値 R_C を表 3.1 の通りに設定する. ここで, 現在の CS 閾値が高いほど R_C は大きく, 現在の CS 閾値が低いほど R_C は小さくなる. これにより, CS 範囲を狭めることで競合端末を減らし, 送信機会を確保する選択ほど高い報酬が得られる.

しかし, 各端末が高報酬を得るために CS 閾値を増加させ続ければ隠れ端末が増加

表 3.1: CS 閾値に対する基本報酬値 R_C

C_{cur}	-74	-78	-82	-86
R_C	1	0.75	0.5	0.25

し、通信衝突が多発する。そのため本提案手法では、想定する状況に応じて ACK フレーム受信時の再送回数 N_{retr} 、CS した他端末の宛先割合 N'_F/N_F から R_C の正負を決定する。想定する状況は通信衝突の有無、晒し端末の影響から下記に示す [i.]-[iv.] のパターンに分類する。

- パターン [i.] では、 $N_{\text{retr}} = 0$ かつ $N'_F/N_F > \theta$ すなわち、通信衝突は無く、晒し端末による性能劣化が生じている状況を想定する。この時、端末は N'_F/N_F が増加するような行動には負の報酬、それ以外の行動には正の報酬を獲得する。これにより、CS 範囲を狭めて晒し端末の影響を軽減するような行動を選択することが可能である。
- パターン [ii.] では、 $N_{\text{retr}} > 0$ かつ $N'_F/N_F > \theta$ すなわち、通信衝突が起きており、晒し端末による性能劣化も生じている状況を想定する。この時、端末は負の報酬を獲得する。これは、本提案手法における方針として晒し端末の影響がある場合においても通信衝突の回避を優先するためである。
- パターン [iii.] では、 $N_{\text{retr}} = 0$ かつ $N'_F/N_F \leq \theta$ すなわち、通信衝突は無く、晒し端末による性能劣化も生じていない状況を想定する。この時、端末は正の報酬を獲得する。パターン [iii.] は通信を行う上で最も望ましい状況であり、各端末はパターン [iii.] になるように学習を進め、CS 閾値を遷移させる。

- パターン [iv.] では, $N_{retx} > 0$ かつ $N'_F/N_F \leq \theta$ すなわち, 通信衝突が起きており, 晒し端末による性能劣化は生じていない状況を想定する. この時, 端末は負の報酬を獲得する. これにより, 通信衝突を回避するために CS 範囲を広げるような行動を選択する.

それぞれの状況 [i.]-[iv.] において獲得した報酬を基に選択した行動の Q 値を更新し, 最適な CS 閾値を探索・設定する. また, 学習を進めていくごとに Q 値による選択が主になるように, 本提案手法では, ε -decay アルゴリズムを用いる [?]. 以下に, 学習の際に想定される状況と, その状況下での報酬値の設定のまとめを示す.

$$i. \frac{N'_F}{N_F} > \theta \cap N_{retx} = 0$$

$$R = \begin{cases} -R_C, & \frac{N'_F}{N_F} \text{ increase} \\ R_C, & \text{otherwise} \end{cases} . \quad (3.1)$$

$$ii. \frac{N'_F}{N_F} > \theta \cap N_{retx} > 0$$

$$R = -R_C \quad (3.2)$$

$$iii. \frac{N'_F}{N_F} \leq \theta \cap N_{retx} = 0$$

$$R = R_C \quad (3.3)$$

$$iv. \frac{N'_F}{N_F} \leq \theta \cap N_{retx} > 0$$

$$R = -R_C \tag{3.4}$$

第4章

提案手法の性能評価

本章では, 計算機シミュレーションの結果から本提案手法の有効性を確認する. 4.1 節では, シミュレーションにおける評価条件を示す. 4.2 節, 4.3 節, 4.4 節では, それぞれ隠れ端末のみ, 晒し端末のみ, 隠れ端末および晒し端末の混在のそれぞれのネットワーク環境における特性評価を行う.

4.1 評価条件

4.1.1 シミュレーション条件

本研究における性能評価として, 十分に学習が行われた状態における性能を評価するために, シミュレーションを 20[sec] 実施し, 開始から 15[sec] から 20[sec] までのデータを採用した. 表 4.1 及び表 4.2 にシミュレーション諸元及び学習パラメータ諸元をそれぞれ示す.

表 4.1: シミュレーション諸元

データフレームの伝送レート	18 Mbps
ACK フレームの伝送レート	12 Mbps
ACK フレームの送信時間	32 μ sec
SIFS 期間	16 μ sec
DIFS 期間	34 μ sec
1 スロットタイム	9 μ sec
ペイロードサイズ	200 bytes
CWmin	16
CWmax	1024
最大再送回数	7
シミュレーション時間	20 sec
送信電力	16.02 dBm
送信可能範囲	45 m

表 4.2: 学習パラメータ諸元

探索率 ε	[0.99, 0.001]
探索係数 c_ε	0.998
学習率 α	0.1
割引率 γ	0.5
CS 閾値の集合	[-74, -78, -82, -86]
各 CS 閾値における R_{CS}	[27m, 37m, 51m, 70m]

4.1.2 評価項目

本提案手法の特性評価の指標として、システムスループット、衝突率、公平性の各特性を用いた。各評価指標の定義を以下に示す。

- システムスループット S 本研究では、全体のシミュレーション時間に対して、APが受信成功したデータ量をシステムスループットとして定義し、以下の式で算出する。

$$S = \sum_{k=1}^n \frac{Rx_{suc}P \times 8}{T_{sim}} [Mbps] \quad (4.1)$$

ここで、 Rx_{suc} はシミュレーション時間内に受信成功したパケット数、 $P[bytes]$ はペイロードサイズ、 T_{sim} はシミュレーション時間を示す。

- 衝突率 p_{col} データの送信試行回数に対する衝突による送信失敗回数の割合を示し。以下の式で算出する

$$p_{col} = \frac{N_{col}}{N_{tx}} \quad (4.2)$$

ここで、 N_{tx} はシミュレーション時間内にデータの送信試行を行った回数を示す。

- 公平性ネットワーク内の各端末がどの程度均等に通信に参加できているかの公平性を表す指標としてフェアネスインデックス FI を想定した。

$$FI = \frac{(\sum_{n=1}^N S_n)^2}{N \cdot \sum_{n=1}^N S_n^2} \quad (4.3)$$

ここで、 N はネットワーク内の全端末台数である。

4.1.3 性能比較対象

本提案手法との性能比較対象として、CS 閾値を固定値として CW のみを学習制御した場合 (CWRL) [14], CWmin, CS 閾値を共に固定値として学習制御を行わない場

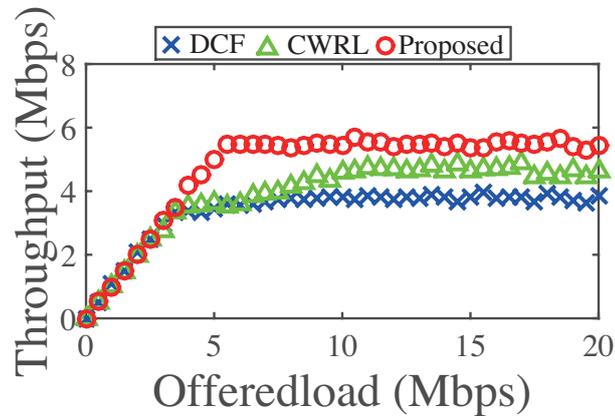


図 4.1: 端末台数 15 台におけるスループット (リング配置)

合 (DCF) を示す.

4.2 隠れ端末環境下における評価

4.2.1 ネットワークトラヒック負荷に対する特性評価

図 4.1, 4.2 に 3.1.1.1 節において想定する端末台数が 15 台をリング型に配置したネットワークでのオフロードに対するネットワークスループット, 衝突率を示す. 図 4.1 より, オフロードの増加とともに, ネットワークスループットが増加し, ある負荷において最大値を得る. また, オフロードをさらに増加させるとスループットは一定値をとる. 本提案手法は, CS 閾値を固定値とする DCF, CWRL と比較して高いスループットを得ていることが分かる. これは, 図 4.2 から衝突率の改善によるものと分かる. また, 図 4.3, 4.4 に飽和状態における平均 CS 閾値及び選択 CW の時系列変化を示す. 図 4.3 より, 時間が経過し, 十分学習が行われると, 隠れ端末の影響を軽減するために CS 範囲を広げるよう CS 閾値は $-86[\text{dBm}]$ に収束し

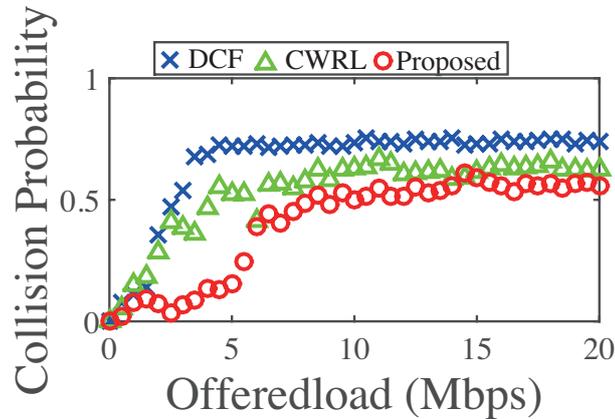


図 4.2: 端末台数 15 台における衝突率 (リング配置)

ている。これにより、隠れ端末に起因する通信衝突を回避できる。そのため、図 4.4 より、CWRL が通信衝突を回避するために CW を大きくして送信頻度が減少しているのに対し、本提案手法は CS 閾値の制御により隠れ端末の影響を軽減し、小さい CW で送信が成功する。これらの要因によりスループットが向上し、本提案手法の有効性が確認できる。

また、図 4.5, 4.6 に 3.1.1.1 節において想定する端末台数 15 台をランダムに配置したネットワークでのオフロードに対するスループット、公平性を示す。リング配置と同様にオフロードの増加とともにスループットが増加している。図 4.5 より、本提案手法が最も高いスループットを得ていることが確認できる。これにより、端末ごとに隠れ端末の影響が異なる場合においても各端末が隠れ端末を無くすよう CS 範囲を広げることで、ネットワーク全体の性能が向上することが分かる。また、図 4.6 より、本提案手法は DCF や CWRL に比べて FI の低下を軽減している。これは、CS 閾値を固定値とした DCF, CWRL では、周囲の隠れ端末数の差によって不公平性が生じるのに対し、本提案手法では各端末が CS 範囲を広げることにより、隠れ端末数の差が小さくなるためである。

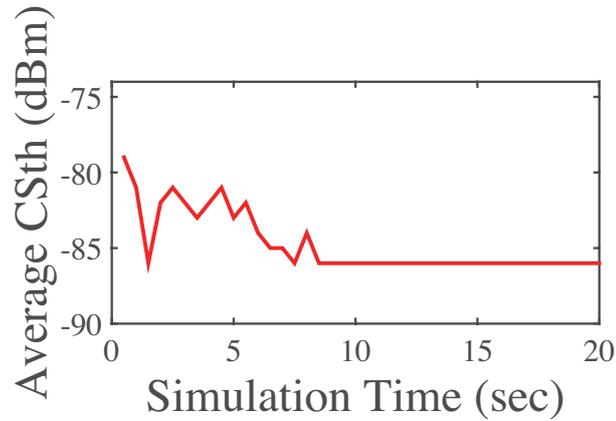


図 4.3: 平均 CS 閾値の時系列変化

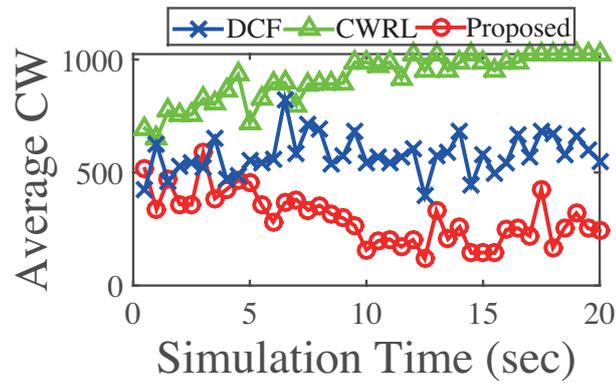


図 4.4: 選択 CW の時系列変化

4.2.2 任意の端末台数に対する特性評価

図 4.7, 4.8 にリング型, ランダム型配置における端末台数に対する最大スループット特性を示す. 本研究では, オファードロードを変化させたときに, 最も高い値のスループットを最大スループットとする. また, 図 4.9 に任意の端末台数におけるそれぞれ配置方法における平均 CS 閾値を示す. 図 4.9 における CS 閾値は学習による収束後の値としてシミュレーション開始から 15[sec] 以降の CS 閾値の平均値としている. これらの結果より, 本提案手法はリング型, ランダム型配置のいずれにおいても端末

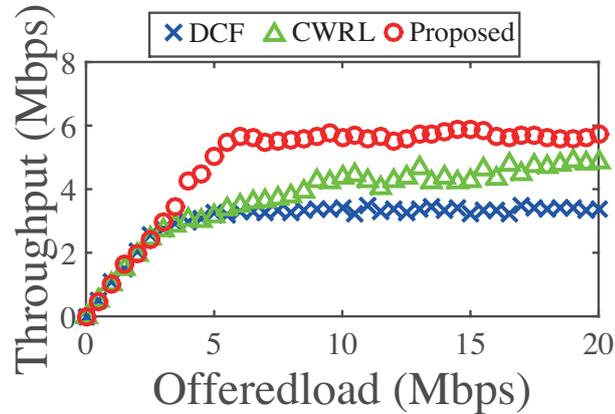


図 4.5: 端末台数 15 台におけるスループット (ランダム配置)

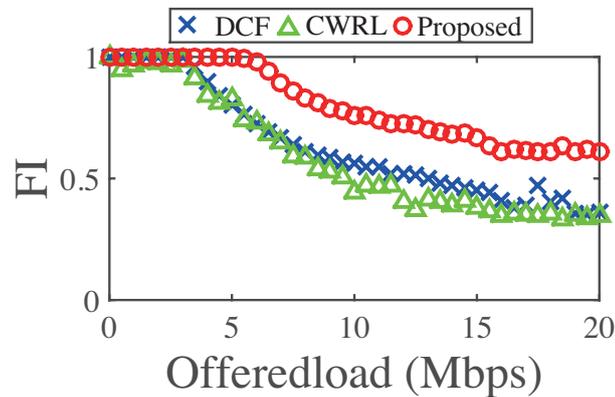


図 4.6: 端末台数 15 台における公平性 (ランダム配置)

台数 15, 30 台において, CS 閾値は $-86[\text{dBm}]$ に収束し, CS 範囲を広げたことで隠れ端末の影響を軽減しスループットが向上している. また, 端末台数 5 台においてランダム配置における CS 閾値がリング配置に比べて高くなっている. これは, 端末台数が少ない場合でもリング配置では常に端末間の距離が一定以上離れており, CS 閾値を広げる必要があるのに対し, ランダム配置では CS 閾値が高い状態でも隠れ端末がなくなるような配置が考えられるためである.

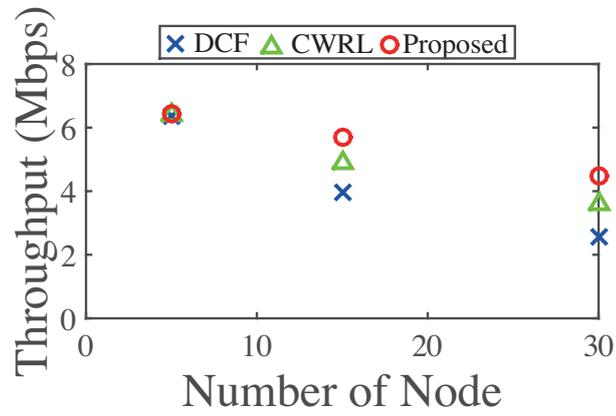


図 4.7: 任意の端末台数におけるスループット (リング配置)

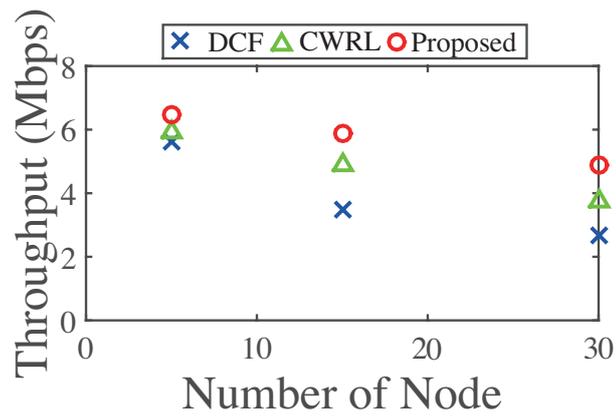


図 4.8: 任意の端末台数におけるスループット (ランダム配置)

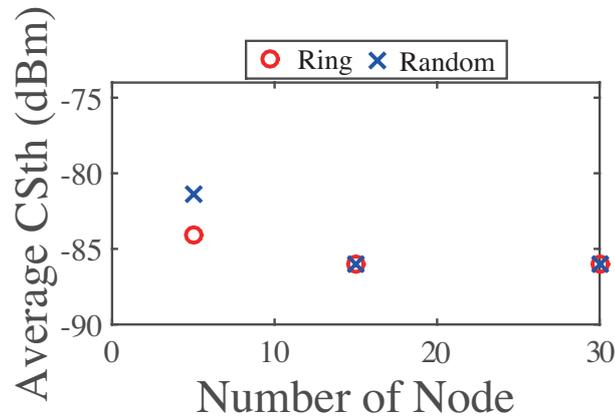


図 4.9: 任意の端末台数における平均 CS 閾値の収束値

4.3 晒し端末環境下における評価

4.3.1 ネットワークトラヒック負荷に対する特性評価

図??に 3.1.1.2 節において想定する 2 台の AP で構成されるネットワークにおけるオフロードに対するスループットを示す。図 4.11 より、本提案手法が最も高いスループットを達成していることが分かる。また、図??に飽和状態での本提案手法における各 STA の CS 閾値の時系列変化を示す。結果より、それぞれの STA は時間経過により学習が進むと CS 閾値は -74 [dBm] に収束している。想定するネットワークでは通信衝突は発生せず、通信性能の劣化要因は晒し端末によるチャネル利用率の低下となっている。そのため、本提案手法は CS 範囲を狭めるような行動選択により晒し端末の影響を抑えることにより、スループットが向上したと考えられる。

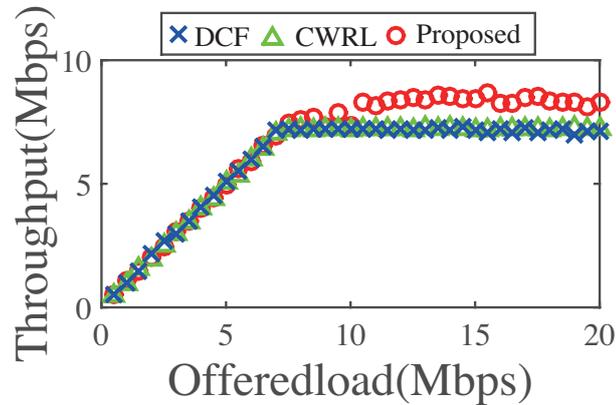


図 4.10: 晒し端末のみの環境におけるスループット

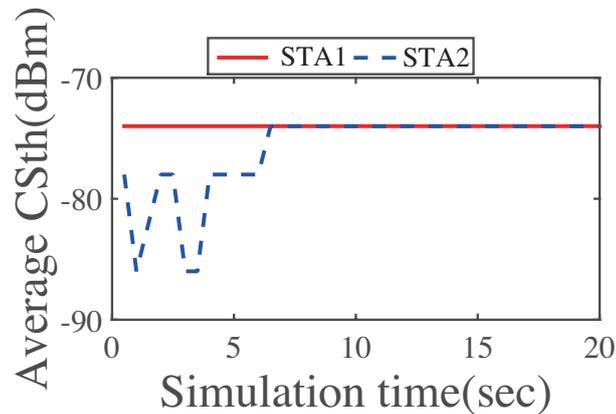


図 4.11: 平均 CS 閾値の時系列変化

4.4 隠れ端末および晒し端末の混在環境下における評価

4.4.1 ネットワークトラヒック負荷に対する特性評価

図 4.12, 4.13 に 3.1.1.3 節において想定する端末台数 15 台をランダムに配置したネットワークにおけるオフロードに対するスループット, 公平性を示す. 隠れ端末および晒し端末の混在環境下での評価対象として DCF, CWRL はそれぞれ CS 閾値を $-74[\text{dBm}]$, $-86[\text{dBm}]$ で固定とした場合の結果を示している. これにより,

CS 閾値が固定値である DCF, CWRL において CS 範囲が狭く隠れ端末の影響が大きくなる場合と CS 範囲が広く晒し端末の影響が大きくなる場合における振る舞いを確認する。図 4.12, 4.13 の結果より、隠れ端末および晒し端末の混在環境において本提案手法は DCF, CWRL と比較し、公平性を損なうことなく高いスループットを達成している、また、DCF, CWRL において CS 閾値を-74[dBm] とした場合に大きくスループットが劣化している。これにより混在環境において隠れ端末の影響が晒し端末の影響に比べて支配的であることが分かる。また、図 4.14, 4.15 に全端末から無作為に選択した 2 台の STA(STA1 と STA10) における CS 閾値および晒し端末によりタイマがフリーズした回数の割合 N'_F/N_F の時系列変化を示す。図 4.14 より、STA1 では時間経過により学習が進むと CS 閾値は-74[dBm] に収束しており、 N'_F/N_F が DCF, CWRL に比べて小さくなっている。これより、STA1 が CS 範囲を狭めることで周囲の晒し端末の影響を抑制出来ていることが分かる。一方、STA10 では CS 閾値は-82[dBm] の付近を移動しており、 N'_F/N_F は DCF, CWRL と同等の値となっている。これより、STA1 に比べて STA10 では晒し端末の影響が小さく、CS 範囲を広げて隠れ端末の影響を抑制するような行動選択をしていることが分かる。これらの結果より、個々の端末が周囲の状況に応じて CS 閾値を制御し、隠れ端末および晒し端末の影響を軽減することでネットワーク全体のスループットが向上していることが確認できる。

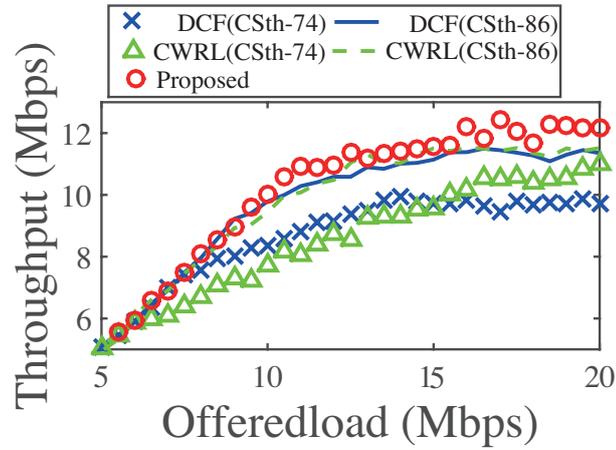


図 4.12: 隠れ晒し端末混在環境下におけるスループット (端末台数 15 台)

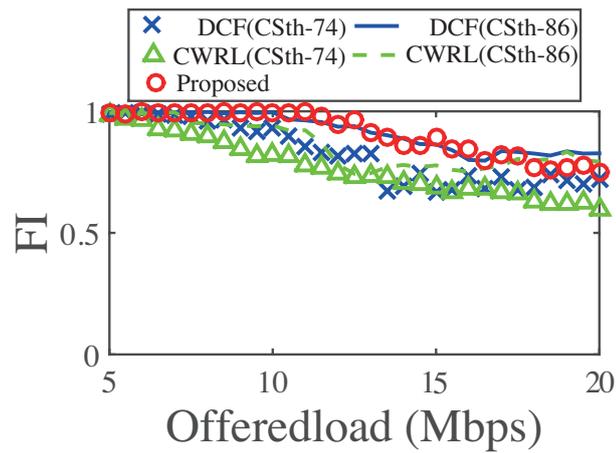


図 4.13: 隠れ晒し端末混在環境下における公平性 (端末台数 15 台)

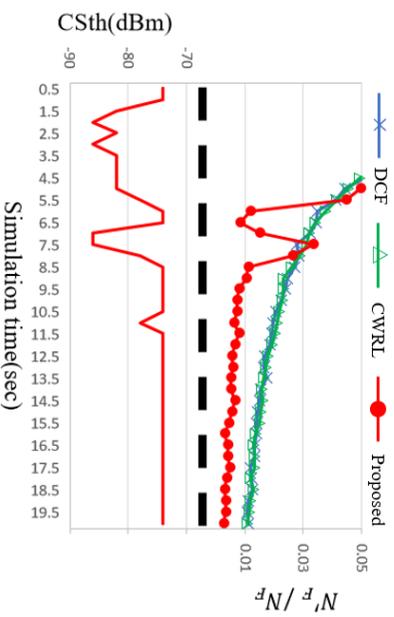


図 4.14: CS 閾値及び晒し端末によるフリーズ回数 N'_F (STA1)

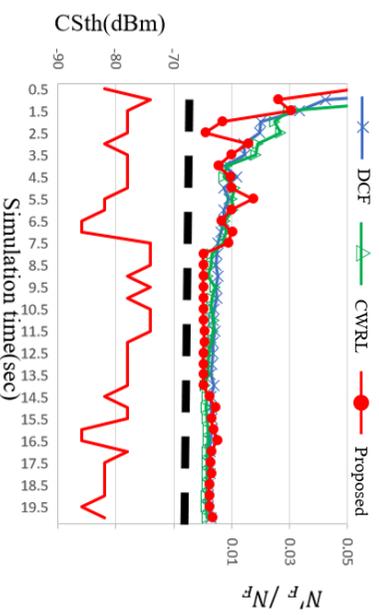


図 4.15: CS 閾値及び晒し端末によるフリーズ回数 N'_F (STA10)

4.4.2 任意の端末台数に対する特性評価

図 4.16 に端末台数に対する最大スループットを示す．図 4.16 より，隠れ端末および晒し端末の混在環境において任意の端末台数で本提案手法が高いスループットを達成していることが分かる．また，端末台数の増加するのに従い，CWRL と本提案手法の性能差が大きくなっていく．これは，端末台数が増加し，通信衝突が多発すると CWRL では CW を極端に大きく設定する必要があり，晒し端末の影響と合わせて極

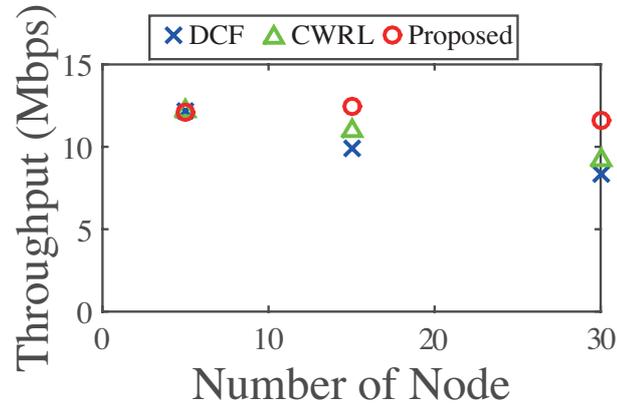


図 4.16: 隠れ晒し端末混在環境下におけるスループット (任意の端末台数)

端に送信機会が減少するためであると考えられる。一方, 本提案手法では CS 閾値を制御し, 隠れ端末および晒し端末の数を減らすことで, 端末台数が増加しても通信衝突の回避と送信機会の確保のいずれにも対応できている。これらの結果より, 隠れ端末および晒し端末の混在環境における本提案手法の有効性が確認できる。

第5章

まとめと今後の課題

本研究では、あらゆるネットワーク環境において隠れ端末および晒し端末による通信性能を改善することを目的とし、強化学習に基づくCS閾値の制御手法を提案した。これまで、システムの制御パラメータを調整する手法は数理モデルを用いた手法が多く研究されてきた。しかし、数理モデルを用いた手法にはいくつかの理想化が含まれ、一般的に想定するネットワーク環境あるいは特定の状況下での対応が主であり、あらゆるネットワーク環境に柔軟に対応することは困難であった。この課題に対し、本研究では、与えられた環境に柔軟に対応することが可能な強化学習に基づいてCSの制御パラメータであるCS閾値を周囲の状況に応じて適切に制御することで隠れ端末および晒し端末による通信性能の劣化を軽減出来ることを計算機シミュレーションにおいて示した。

特性評価では、それぞれ隠れ端末のみ、晒し端末のみ、隠れ端末および晒し端末の混在するネットワーク環境を想定し、学習を用いない従来のDCFおよびCS閾値を固定値としてCWのみを学習制御するCWRLと本提案手法の性能比較を行った。それぞれのネットワーク環境において、スループット、衝突率、公平性の観点から本提案手法の有効性を示した。

今後の課題として、各端末が移動し、隠れ端末および晒し端末の関係が時系列で変化する状況や新たな通信端末の新規参入や離脱といった端末台数が変化するような

状況, また, マルチホップネットワーク等の異なるネットワーク環境において本提案手法の有効性を示すことが考えられる.

参考文献

- [1] WANG, Cheng-Xiang, et al. Cellular architecture and key technologies for 5G wireless communication networks. *IEEE communications magazine*, 2014, 52.2: 122-130.
- [2] LETAIEF, Khaled B., et al. The roadmap to 6G: AI empowered wireless networks. *IEEE communications magazine*, 2019, 57.8: 84-90.
- [3] DOGRA, Anutusha; JHA, Rakesh Kumar; JAIN, Shubha. A survey on beyond 5G network with the advent of 6G: Architecture and emerging technologies. *IEEE Access*, 2020, 9: 67512-67547.
- [4] IEEE Computer Society LAN MAN Standards Committee, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-1997*.
- [5] Yumen, Yosuke, et al. "Throughput analysis for wireless multi-hop networks with synchronous full-duplex MAC." 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC). IEEE, 2020.
- [6] Kim, Hyerin, and Jungmin So. "Improving Spatial Reuse of Wireless LAN Uplink Using BSS Color and Proximity Information." *Applied Sciences* 11.22 (2021): 11074.
- [7] Zhu, Jing, et al. "Adapting physical carrier sensing to maximize spatial reuse in 802.11 mesh networks." *wireless communications and mobile computing* 4.8 (2004): 933-946.

- [8] Morino, Yoshiaki, et al. "A novel contention window control scheme based on a Markov chain model in dense WLAN environment." 2015 3rd International Conference on Artificial Intelligence, Modelling and Simulation (AIMS). IEEE, 2015.
- [9] Bianchi, Giuseppe. "Performance analysis of the IEEE 802.11 distributed coordination function." IEEE Journal on selected areas in communications 18.3 (2000): 535-547.
- [10] Shi, Zhefu, Cory Beard, and Ken Mitchell. "Analytical models for understanding space, backoff, and flow correlation in CSMA wireless networks." Wireless networks 19 (2013): 393-409.
- [11] Xia, Qiuyan, and Mounir Hamdi. "Contention window adjustment for IEEE 802.11 WLANs: a control-theoretic approach." 2006 IEEE international conference on communications. Vol. 9. IEEE, 2006.
- [12] Ni, Qiang, et al. "Modeling and analysis of slow CW decrease IEEE 802.11 WLAN." 14th IEEE Proceedings on Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003.. Vol. 2. IEEE, 2003.
- [13] Ma, Xiaomin, and Xianbo Chen. "Performance analysis of IEEE 802.11 broadcast scheme in ad hoc wireless LANs." IEEE Transactions on Vehicular Technology 57.6 (2008): 3757-3768.
- [14] Wu, Celimuge, et al. "A MAC protocol for delay-sensitive VANET applications with self-learning contention scheme." 2014 IEEE 11th Consumer Communications and Networking Conference (CCNC). IEEE, 2014.

- [15] 眞田耕輔, “IEEE 802.11 ネットワークにおけるコンテンションウィンドウの強化学習に基づく適切な CWmin 選択手法の提案,” 信学技報, vol. 121, no. CCS-442, pp. 90–96, Mar. 2022.
- [16] K. Sanada, H. Hatano and K. Mori, ”Simple Reinforcement Learning based Contention Windows Adjustment for IEEE 802.11 Networks,” 2023 IEEE 20th Consumer Communications Networking Conference (CCNC), Las Vegas, NV, USA, 2023, pp. 692-693, doi: 10.1109/CCNC51644.2023.10059960.
- [17] Minn, Hlaing, and Vijay K. Bhargava. ”An investigation into time-domain approach for OFDM channel estimation.” IEEE Transactions on broadcasting 46.4 (2000): 240-248.
- [18] Lu, Lu, et al. ”An overview of massive MIMO: Benefits and challenges.” IEEE journal of selected topics in signal processing 8.5 (2014): 742-758.
- [19] Tataria, Harsh, et al. ”Channel correlation diversity in MU-MIMO systems—analysis and measurements.” 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE, 2019.
- [20] Liu, Yih-Chiao, and G. Wise. ”Performance of a CSMA/CD protocol for local area networks.” IEEE Journal on Selected Areas in Communications 5.6 (1987): 948-955.
- [21] Al Hasan, Mohammad, et al. ”Link prediction using supervised learning.” SDM06: workshop on link analysis, counter-terrorism and security. Vol. 30. 2006.

- [22] Yang, Hongyu, et al. "An unsupervised learning-based network threat situation assessment model for internet of things." *Security and Communication Networks* 2020 (2020): 1-11.
- [23] Kanda, Takamochi, et al. "ACK-Less Rate Adaptation Using Distributional Reinforcement Learning for Reliable IEEE 802.11 bc Broadcast WLANs." *IEEE Access* 10 (2022): 58858-58868.
- [24] Chen, Chen, et al. "Contention resolution in Wi-Fi 6-enabled Internet of Things based on deep learning." *IEEE Internet of Things Journal* 8.7 (2020): 5309-5320.
- [25] Zhang, Lyutianyang, et al. "Multiaccess point coordination for next-gen Wi-Fi networks aided by deep reinforcement learning." *IEEE Systems Journal* 17.1 (2022): 904-915.
- [26] Kulin, Merima, et al. "A survey on machine learning-based performance improvement of wireless networks: PHY, MAC and network layer." *Electronics* 10.3 (2021): 318.
- [27] Liu, Zhenzhen, and Itamar Elhanany. "RL-MAC: a reinforcement learning based MAC protocol for wireless sensor networks." *International Journal of Sensor Networks* 1.3-4 (2006): 117-124.
- [28] Park, Sung Hyun, Paul Daniel Mitchell, and David Grace. "Reinforcement learning based MAC protocol (UW-ALOHA-QM) for mobile underwater acoustic sensor networks." *IEEE Access* 9 (2020): 5906-5919.

- [29] Wydmański, Witold, and Szymon Szott. "Contention window optimization in IEEE 802.11 ax networks with deep reinforcement learning." 2021 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2021.
- [30] Ekici, Ozgur, and Abbas Yongacoglu. "IEEE 802.11 a throughput performance with hidden nodes." IEEE Communications Letters 12.6 (2008): 465-467.

謝辞

本研究の遂行及び論文作成に当たり、ご多忙の中多大なご指導と助言をしていただいた眞田 耕輔助教, 羽多野 裕之准教授, 森 香津夫教授に深く感謝申し上げます。

最後に、研究室の整備等で陰ながら支えて下さった山本 好弘技術長補佐, ならびに本研究のため御助力を注いでくださいました大学院生, 学部生の皆様方に深く感謝いたします。

令和4年 東山 祐斗

研究業績

- Yuto HIGASHIYAMA, Kosuke SANADA, Hiroyuki HATANO, Kazuo MORI, “Q-learning based Carrier Sense Threshold adjustment for Multi-hop Network,” Proceedings of the 12th International Symposium for Sustainability by Engineering at Mie University (Research Area C), September 2022.
- 東山祐斗, 眞田耕輔, 羽多野裕之, 森香津夫, “IEEE 802.11 WLANにおける強化学習に基づくCWminおよびキャリアセンス閾値の統合制御,” 信学技報, CCS2023-20, vol. 123, no. 147, pp. 19-24, 2023年8月.
- Yuto HIGASHIYAMA, Kosuke SANADA, Hiroyuki HATANO, Kazuo MORI, “Reinforcement learning-based control of CWmin and Carrier Sense Threshold for IEEE 802.11 WLAN” Proceedings of the 13th International Symposium for Sustainability by Engineering at Mie University (Research Area C), September 2023.