

2022年度 独創的研究助成費 実績報告書

2023年 3月16日

報告者	学科名	情報通信工学科	職名	教授	氏名	榊原勝己
研究課題	Q学習を用いたランダム・アクセス方式の設計とその性能評価					
研究組織	氏名	所属・職		専門分野	役割分担	
	代表	榊原 勝己	情報通信工学科・教授	通信工学	総括, 理論解析, 検証	
	分担者	高林 健人 原 佑輔 山内 彰悟	情報通信工学科・助教 大学院学生 大学院学生 大学院学生	通信工学	シミュレータ構築, 実装実験 プログラミング, データ整理, 成果発表	
研究実績の概要	<p>人工知能の進展により, 無線通信システムにおいても, 機械学習, 強化学習の適用が研究されている. 不特定多数端末からの散発的なデータを收容するためのランダム・アクセス方式に対しても, 強化学習の一手法である, Q 学習を, 無線 LAN で利用されている CSMA/CA に適用した RL-CSMA/CA が提案されている(T.-W. Kim and G.-H. Hwang, J. Inform. Commun. Convergence Eng., Vol. 19, Issue 1, pp. 1-7, Mar. 2021). また, 動的フレーム付アロハに対して, フレーム内での送信結果(アイドル, 成功, 衝突)に応じてデータを保持している端末数を推定するアルゴリズムを組み込んだ LD-DFSA が提案されている(Z. Jiang, B. Li, M. Yang and Z. Yan, MDPI Sensors, Vol. 20, Issue 1, Article-ID 228, Dec. 2019).</p> <p>本研究では, これの2種の方式を組み合わせ, Q 学習と端末数推定を装備した方式を提案し, その性能を計算機シミュレーションにより評価する.</p> <p><math>N</math>個の単一バッファ端末が, 固定長パケットを共通の受信局に送信するランダム・アクセス方式を考える. 時間軸はパケットを送信するために十分なタイムスロットに分割され, パケットは1つのタイムスロット中で送信される. 以下では, 複数のタイムスロットでフレームを構成する方式を考える. ここで, <math>m</math>番目に定められたフレーム長を <math>L_m</math> とする(初期値は <math>L_1</math>).</p> <p>提案方式の各端末は, 1 フレーム内のスロット数 <math>L_m</math> と等しい長さの実数ベクトル(Q テーブル)を持つ. <math>L_m</math>が定められてからの第 <math>t</math> フレーム開始時において端末 <math>n</math> が持つ Q テーブルを <math>\vec{Q}_t(n, m) = [Q_t(n, m, 1), Q_t(n, m, 2), \dots, Q_t(n, m, L_m)]</math> をとする (<math>t = 1, 2, \dots, n = 1, 2, \dots, N, m = 1, 2, \dots</math>). ただし, 初期値は <math>\vec{Q}_t(n, m) = \vec{0}</math> とする. 第 <math>t</math> フレーム開始時にパケットを有している端末は, フレーム内で最大の Q 値 <math>Q_t(n, m, k)</math> を持つタイムスロット <math>k</math> でパケットを送信する. 最大の Q 値を持つタイムスロットが複数ある場合は, その中からランダムで送信スロットを選択する. その後, 送信の成否に応じて Q 値 <math>Q_t(n, m, k)</math> を</p> $Q_{t+1}(n, m, k) = \gamma Q_t(n, m, k) + R$					

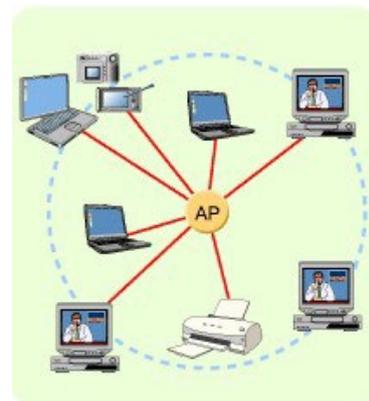


図1. ネットワーク構成

により更新する. ここで,  $\gamma$ は学習率( $0 < \gamma \leq 1$ ),  $R$ は報酬であり, パケット送信成功時に+3, 衝突時に-1をとる. なお, パケットを送信した以外のタイムスロットではQ値 $Q_t(n, m, k)$ は更新されない. また, パケットの成功あるいは衝突が予め定められた閾値を超えた場合, 観測した成功スロット数(あるいは衝突スロット数)に応じて端末数  $N_{exp}$ を推定し, フレーム長を  $L_m$  から  $L_{m+1} = N_{exp}$  に更新する.

各端末は, バッファが空であればスロット時間当たり確率 $P_{gen}$ でパケットを発生し( $0 < P_{gen} \leq 1$ ), 端末と受信局間での伝搬遅延時間及び通信路誤りは無視できるものとする. 端末数 $N = 100$ , 初期フレーム長 $L_1 = 70$  に対し, 提案方式と従来方式のスループットを, 計算機シミュレーションにより求めた結果を図2に示す. 図2より, 提案方式(図2のRL-LC)従来方式に比べ, 高いスループットを示すことが明らかになった. 今後は, Q 学習における学習速度の高速化, 端末数推定の精度向上を図る.

研究実績  
の概要

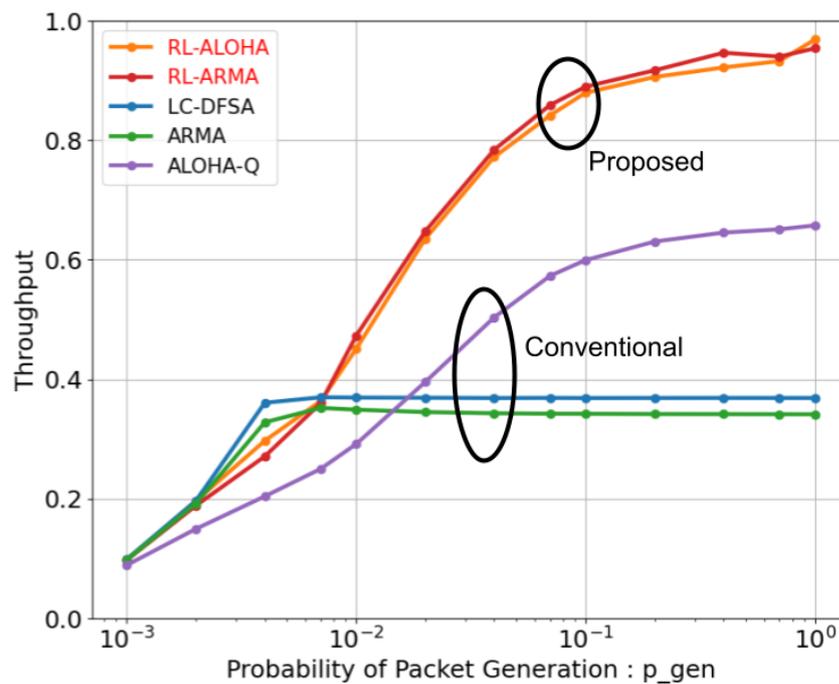


図2. 提案方式(RL-LC)のスループット特性

成果資料目録

山本龍樹, 榊原勝己, 高林健人, "Q 学習と ARMA フィルタを用いた DFSA プロトコルの性能評価," 令和4年度(第73回)電気・情報関連学会中国支部連合大会, オンライン, no.R22-20-05, 2022年10月.