

<p>研究代表者</p>	<p>所属学系・職名 自然科学・情報学系・准教授 氏名 松本 正晴</p>
<p>研究課題</p>	<p>深層強化学習によるインターネット輻輳制御機構の高性能化 High Performance Internet Congestion Control using Deep Reinforcement Learning</p>
<p>成果の概要</p>	<p><b>【背景】</b> インターネットに代表される通信システムでは、単位時間あたりに伝送可能なデータ量に限りがあり、通信要求過多によって遅延が大きくなったり、データロスが発生する現象を輻輳（ふくそう）と呼ぶ。TCP（Transmission Control Protocol）の性能向上（高スループット、低遅延、公平性）の実現を目的とした輻輳制御アルゴリズム（Congestion Control Algorithm, CCA）に関する研究は、30年以上にわたり活発に行われてきており、これまで膨大な数のCCAが提案されている。一方、近年の機械学習やデータサイエンスの発展とともに、深層強化学習（Deep Reinforcement Learning, DRL）をCCAに応用する試みが進められている。特に、古典的なCC手法と最新のDRL技術を組み合わせたハイブリッドCCA（Abbasloo, et al., 2020）が様々なネットワーク条件下でも一貫して高い性能を達成する成果を挙げている。</p> <p><b>【参考文献】</b> S. Abbasloo, et al., “Classic Meets Modern: a Pragmatic Learning-Based Congestion Control for the Internet,” SIGCOMM’20, July 2020.</p> <p><b>【目的】</b> 本研究では、古典的なCCAと最新のDRL技術を組み合わせたハイブリッドCCAを実装し、さまざまな条件の下で性能を測定することにより、CCAの更なる高性能化を目指すことを目的とする（図1）。</p> <p>DRLのみで構成されるCCAについては近年検討が進められてきているが、DRLベースと従来提案のCCAとの併用は、これまでほとんど検討されてきておらず世界的にも数件程度の例しかない。この研究が進展することにより、CCA設計の新しいデファクトスタンダードを構築できる可能性がある。</p> <p><b>【方法】</b> 本研究は、研究代表者である松本正晴准教授と、研究分担者である内海哲史准教授と共同で行う。主に松本はDRLの実装・学習・評価等を担当し、内海准教授はネットワーク環境のエミュレータ・シミュレータの実装構築や古典的な</p> <div data-bbox="885 1198 1407 1512" data-label="Diagram"> <p>The diagram illustrates the Hybrid CCA architecture. At the center is a cloud containing a server icon. Above the server is a brain icon labeled '深層強化学習' (Deep Reinforcement Learning). Below the server is a speedometer icon labeled '輻輳制御' (Congestion Control). The cloud is connected to several computer icons representing network nodes. A starburst labeled '高性能' (High Performance) is positioned below the cloud. A label 'バケット' (Bucket) points to a yellow bucket-like structure within the cloud. Arrows indicate the flow of data and control signals between the components.</p> </div> <p>図1: ハイブリッド CCA</p>

CC手法の選択・評価を担当する。図2に本研究で構築予定のDRLシステムの概要を示す。本システムは複数台のActorと1台のLearnerから構成されており、まず各Actorはそれぞれに対応する環境下でネットワークの状態 $s_t$ を観測し、方策 $\pi_t$ に基づく行動 $a_t$ を選択することにより、報酬 $r_t$ と次の状態 $s_{t+1}$ を得る。各Actorはそれらの経験 $e = (s_t^i, a_t^i, r_t^i, s_{t+1}^i)$ をローカルバッファに貯めておき、Learnerが持つReplay memory  $D$ が定期的に各Actorが生成する経験 $e$ を収集する。Learnerの実装は、機械学習のフレームワークであるTensorflowを利用して構築し、各Actorのネットワーク環境はMininetやMahimahi等のネットワークエミュレータにより構築する。本研究で最も重要となる深層強化学習のアルゴリズムは、多くの実績を持つDeep Deterministic Policy Gradient (DDPG)アルゴリズムの一種であるTwin Delayed DDPG (TD3)アルゴリズムを用いる。

【成果】

特に2021年度は、CCAに深層強化学習を適用する前段階の古典的なCCAの性能評価についての成果を得た。古典的なCCAとして、2018年にマサチューセッツ工科大学より発表された高性能なCCAであるCopaを選んだ。現在、Copaに深層強化学習を適用する

ため、学習データの収集および、方策や特徴量の選定を実施している。

なお、本研究課題および研究内容は、2021年度科学研究費助成事業（国際共同研究強化(A)）ネットワークの多様性と競合フローとの親和性を考慮した機械学習による輻輳制御の探究（研究代表者：内海 哲史，配分額：15,600,000円）の研究内容にも一部関連しており、その採択にも寄与したと考えられる。

【主な発表論文】

荻野 雅史，岡田 章吾，内海 哲史，“異種輻輳制御機構の競合時における性能評価：CUBIC vs Copa vs BBR”，第14回インターネットと運用技術シンポジウム（IOTS 2021）。（荻野 雅史が学生奨励賞を受賞）

齋藤 千夏奈，岡田 章吾，荻野 雅史，内海 哲史，“ネットワークエミュレータMininetによるマルチプルCopaフローの性能評価”，情報処理学会 第84回全国大会。

【組織】

研究代表者：松本 正晴（役割：実装）

共同研究者：内海 哲史（役割：性能評価）

図 2: DRL システムの概要

