

Jh180059-NWH

## 端末・エッジ・クラウド連携の三位一体による「考えるネットワーク」の研究

中尾彰宏 (東京大学)

概要 本研究は、端末・エッジ・クラウドの三位一体協調システムアーキテクチャにおいて、(1)アプリケーション毎のスライス技術・エッジコンピューティング技術、(2)クラウド連携によるオンライン機械学習により有用な情報を抽出しトラフィック予測や、その応用について研究する。今年度は、モバイルネットワークのトラフィックをキャプチャーし機械学習によりトラフィックを分類し、ネットワーク・スライシングにより低遅延処理を可能とするエッジコンピューティングの研究を行う。同時に、キャプチャーしたトラフィックをオフラインで分析し、高度なネットワーク運用やトラフィック予測実現に向けて、その可能性を評価する。

### 1. 共同研究に関する情報

#### (1) 共同研究を実施した拠点名

お茶の水女子大学、工学院大学、九州大学、北海道大学、東北大学

#### (2) 共同研究分野

□ 超大容量ネットワーク技術分野

#### (3) 参加研究者の役割分担

- (1) アプリケーション毎のスライス技術・エッジコンピューティング技術 (東京大学)
- (2) クラウド連携によるオンライン機械学習
  - (2-1) クラウド連携 (東京大学・北海道大学・東北大学・九州大学)
  - (2-2) 機械学習・トラフィック解析 (東京大学・お茶の水女子大学・工学院大学)

### 2. 研究の目的と意義

モバイルネットワークトラフィックは増加の一途を辿り、2017 年の CISCO の Visual Network Index (VNI)によれば、2021 年にはモバイルネットワークの全体量は現代の 7 倍、月間で 48.3EB(エグザバイト)となると予想されている。同時に、5G モバイルネットワークのサービスインが 2020 年の東京オリンピック・パラリンピックで実現すれば、オリン

ピック会場の実況中継などの 4K/8K ビデオ配信だけではなく、スポーツ選手と同期した VR/AR などなど超低遅延のサービスも注目される。一般には、mMTC (massive Machine Type Communication), URLLC (Ultra Reliable and Low Latency Communication), eMBB (enhanced Mobile BroadBand)と呼ばれるカテゴリーに分類された多様なアプリケーションの通信が台頭することが予想される。

特に、協調運転等の車の自動制御など、超低遅延のサービスを実現するためには、センサーからアクチュエータまで、1msec 以内のフィードバックの通信を実現する必要がある。それだけではなく、特に低遅延のフィードバックが必要なアプリケーションはクラウドからのフィードバックではなく端末 (UE: User Equipment)の近傍においてデータ処理をするエッジコンピューティングの機構が必要となる。特定のアプリケーションに対して、超多数、超低遅延、超大容量などの多様な要件を満たすための柔軟かつ動的なアジャイルなインフラ技術が必要とされる。そこで 2 つの課題を設定した。第一の課題は、多様アプリケーションを分類し、互いに異なる性質を持つアプリケーションを「非干渉に

収容する」エンドツーエンドのネットワークスライス技術の実現、第二に、アプリケーション毎に収容されたスライスにおいて、低遅延でフィードバックを実現するためのエッジコンピューティングの基盤技術の開発である。

本研究では、これら 2つの課題に対応し、「端末、エッジ、クラウドが三位一体で連携し機械学習を駆使した最適ネットワーク制御を行う「考えるネットワーク」の実証実験」を目的とする。本研究提案は、前年度の研究の継続であり、前年度の研究結果を踏まえ、上記課題に対し、それぞれ以下の実証実験を行う。

第一の課題に対し、アプリケーション毎のスライス基盤技術を実証する。

具体的には、ネットワークスライス技術においては、

(a) 教師データを準備して暗号化パケットデータにも適用可能なアプリケーションを機械学習によりスライスに分類する機構

(Machine Learning Classification)

(b) スライスの資源の「効率」(performance)を他のスライスと干渉しないように収容する機構(Performance Isolation)をネットワーク機器の仮想化により実現、およびスライスに収容されたアプリケーションが他のスライスの他のアプリケーションから不可侵である機構(Security Isolation)を実現する基盤技術

(c) アプリケーション毎に適用可能なエッジコンピューティングの基盤技術を実証する。第二の課題に対し、複数のクラウドを連携し、ネットワークで接続された大規模な計算資源を確保し、クラウドで実行する機械学習とエッジにて結果を適用する連携アーキテクチャをアプリケーションにより実証する。

低遅延のフィードバックを必要とするエッジコンピューティングは、端末の近傍で実行されることが必須であるが、機械学習に

よるアプリケーション毎のトラフィックの分類や、トラフィックの詳細な分析など、大規模な計算資源やストレージ資源を必要とするデータ処理は、必ずしも端末の近傍で実行される必要はなく、ある程度の遅延を許容するクラウド上で実行するのが合理的と考えられる(図1)。

つまり、本研究は、端末、エッジコンピューティング、クラウドコンピューティングがそれぞれの特徴を活かしたデータ処理を適材適所で行い、低遅延のフィードバックを三位一体で実現する協調システムアーキテクチャを実証するという意義を持つ。

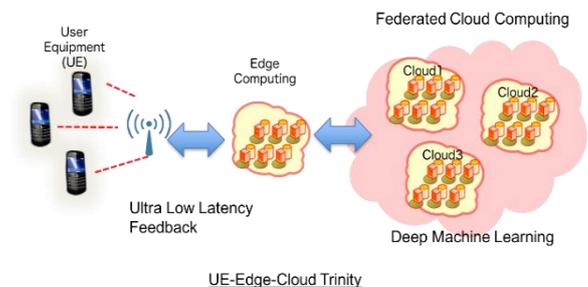


図1 端末・エッジ・クラウドが三位一体で連携し機械学習を駆使する最適ネットワーク制御を行う「考えるネットワーク」概念図

具体的には、提案する端末・エッジ・クラウドの三位一体協調システムアーキテクチャにおいて、(1)アプリケーション毎のスライス技術・エッジコンピューティング技術、(2)クラウド連携によるオンライン機械学習を実施する。

これらの2つのサブ研究テーマにおいて長期(年間)にわたり常時発生する大容量ネットワーク・トラフィックに対し、アプリケーションスライス技術とオンラインで機械学習を可能とするクラウド連携技術において、ネットワーク帯域が増加した際のスケラビリティの検証、超大規模データ共有のためのネットワーク品質の確保や制御、超大容量ネットワーク自体の構築と運用に必要な

監視や管理, そのようなネットワークの安全性の評価と維持確保、などの観点から、提案する端末・エッジ・クラウドの三位一体の協調システムアーキテクチャの有用性を検証する意義を持つ。

また、平成 30 年度は、平成 29 年度の学術的な共同研究結果を踏まえて、一步、社会実装に踏み込む研究を進めた。つまり、実際の商用ネットワークの MVNO からのトラフィックを本研究で構築する通信基盤に誘導し、実トラフィックを用いた実証実験を展開する予定である。既に、平成 29 年度の準備により、MVNO 事業者 2 社、および、端末企業 1 社からの協力が得られることがわかっており、また、商用化に向けた技術要件の議論も進んでいる。これらのことから、今年度の本継続提案においては、学術研究のみならず、社会実装に向けた実証実験であるという社会的インパクトの大きな研究提案であるという意義を持つ。

### 3. 当拠点公募型共同研究として実施した意義

本研究では、地理的に分散する異なる計算資源とストレージ資源をネットワーク接続により連携させ、常時入力される大容量トラフィックに対するオンライン機械学習を行うことにより、アプリケーション毎にトラフィックをスライスに分類し、エッジコンピューティングを実施可能とするシステムアーキテクチャの実証を行う。

このため、連携が得られる北海道大学、東北大学、九州大学、東京大学の各拠点に分散する計算資源とストレージ資源を SINET5 にて接続して活用する拠点公募型共同研究が必須となる。またこの連携により構築するインフラを利用し評価するアプリケーションの研究として、東京大学・お茶の水女子大学・工学院大学が進めるトラフィックの機械学習による分析とエッジコンピューティングの適用の共同研究を実施する。

### 4. 前年度までに得られた研究成果の概要

ディープラーニングを使用し、携帯トラフィックのパケットに対しユーザデータを監視しないでアプリケーション情報取得する技術を使用し、をエッジクラウド連携アーキテクチャにより 90%以上の精度で推定する技術を確立した。本技術により、端末・エッジ・クラウド連携の三位一体のオンライン学習のアーキテクチャを提案した。本技術実現ではネットワーク内でインテリジェンスを持ったノードを想定しているため、分散的にプログラマブルに計算処理する機能が必要になる。昨年度は機械学習や SNS データ解析情報によりトラフィック制御を API を設け実現するアーキテクチャとし、各種ネットワークノード要素技術について研究開発を行った。

### 5. 今年度の研究成果の詳細

エンドエンドでスライシングする技術の研究開発を進めた。モバイル端末では、無線区間のリソース・アイソレーションの実現が重要になる。

前年度はモバイル端末のトラフィックに対しアプリケーション識別を P-GW 接続のエッジで行い、機械学習で得られた情報によりアプリケーションの推定を行い、アプリケーション毎のトラフィック帯域を制御する方式を実現した。今年度は、アプリケーション識別の情報を eNodeB のダウンリンク側に載せ替え eNodeB でトラフィックを識別し、無線リソース・ブロック (RB) スケジューラーを制御する図 2 の構成を考案し、ソフトウェア LTE 無線基地局により実験実証を行った。

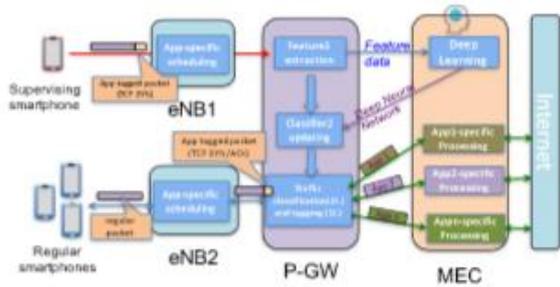


図 2 アプリケーション識別を eNB のダウンリンクで実現するアーキテクチャ

本技術によりアプリケーションに応じて RB をポリシーに従って割り当てることが可能になり RAN スライシングが実現できる。図 3 に同一アプリケーションとしての iperf に異なるポリシーを適用した場合図 3(a)、および iperf と SpeedTest の 2 つのアプリケーションを識別し、異なるポリシー適用したとき図 3(b)のスループット測定結果を示す。

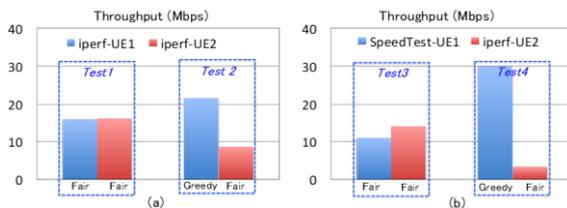


図 3 アプリケーション識別による RB スケジューラのポリシー制御

この結果は、同一および異なるアプリケーションに対して、端末間、端末内で、それぞれに対応したポリシーにより無線リソース割当てが実現できたことを意味する。

次に無線帯域制御に第 2 の課題であるエッジコンピューティング技術による無線リソース・スライシングで低遅延化実現技術の開発を行った。

低遅延処理は、遠隔のクラウドの計算リソースではなく、エッジの計算リソースを利用した計算処理により実現される。さらに遅延を最小とするには、無線基地局でのパケット遅延の低減も必要になる。

そこで、図 4 に示す eNB、RPC 連携の階層型スライシング方式を提案し、ソフトウェア基地局により実験実証を行った。端末からのパケットに対し、EPC にて端末トラフィックを収容するスライス ID を識別し、識別情報に基づきキュー割り当てを行う。キューに蓄積されたパケットに対し、階層型トークン・バケット制御 (Hierarchical Token Bucket, HTB) ではタイムスロット割当てにより帯域制御を行えるが遅延制御が行えない。

端末からのパケットに対し、EPC にて端末トラフィックを収容するスライス ID を識別し、識別情報に基づきキュー割り当てを行う。キューに蓄積されたパケットに対し、階層型トークン・バケット制御 (Hierarchical Token Bucket, HTB) ではタイムスロット割当てにより帯域制御を行えるが遅延制御が行えない。

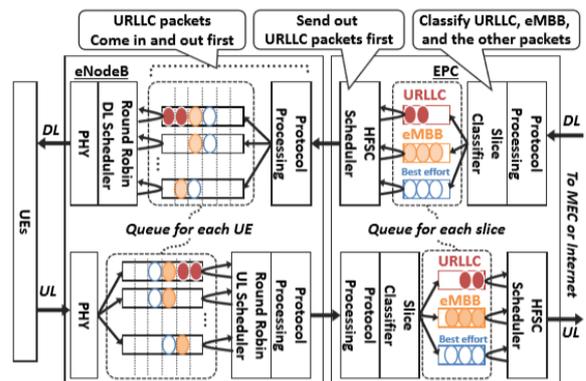


図 4 階層型低遅延スケジューラ構成

低遅延が要求される MQTT パケットに対して HFSC による低遅延化の効果を実験評価した。図 5 に HTB と HFSC による遅延量の統計測定データを示す。HFSC により、低遅延が実現できることを確認した。

クライアントとサーバ間が透過的になるように番号の付け替え機能の実装を行った。

図 5 低遅延スライスの遅延累積分布測定

提案の MEC による低遅延化方式について、各種プロトコルに対し、クラウドアクセス遅延時間について実験的に評価した。図 6 は、遅延特性が分かりやすい ICMP パケットの RTT の累積確率分布の測定結果である。クラウドからの応答に対し、応答時間が改善されるが、分布は均一な変化にはならなかった。これは eNB のスケジューラに起因するもので、RB スケジューラは 1ms ごとに更新するが、UE からの RB 要求が 10ms 間隔であり、さらにサブフレームに反映される時間が 4ms 程度であるため、遅延時間が均一に部分しない要因となっていることがわかった。超低遅延化を目指すためには、スケジューラ手順の最適化の検討が必要であることが分かった。

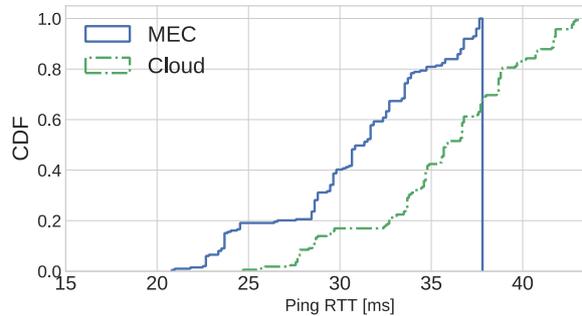


図 6 ICMP パケットに対する遅延時間特性

次にエッジコンピューティングの応用として、サーバからの応答の高速化のため、ネットワーク内のエッジでサーバ応答を返す TCP マイグレーションの技術の研究を行った。図 6 に示すように、ソフトウェア実装によるスイッチをエッジに配置し、クライアントからサーバへのクエリーをスイッチで識別し、アプリケーションレイヤの応答をスイッチで実行することで応答時間短縮を図った。図 6 のスイッチを実現するため、TCP セッションのシーケンス番号、ACK 番号を管理し、ク

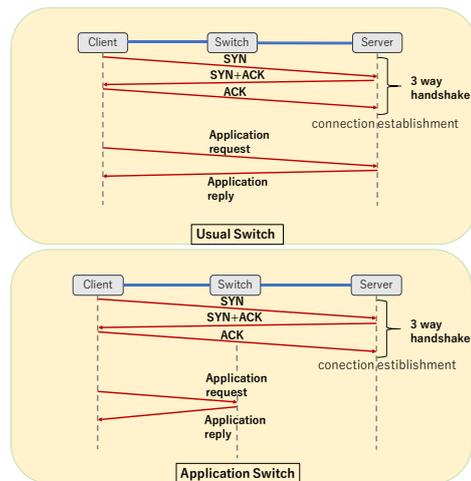


図 7 アプリケーション・スイッチ手順

アプリケーションとして Cassandra データベースを使用し、スイッチによるデータアクセス時間について実験評価を行った。図 7 に測定結果を示す。最初の接続以降では、データがスイッチにキャッシュされるため、平均接続時間が短縮されていくことを確認した。

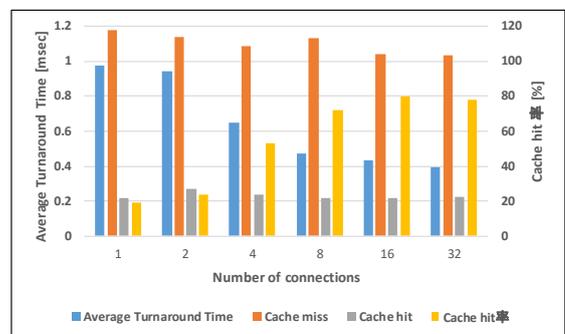


図 8 スイッチによるデータベースアクセス時間短縮効果の評価

本研究プロジェクト全体として、ネットワーク・トラフィックデータを教師データとして、深層学習によりトラフィック特徴量を抽出する技術の研究を行っている。例えば、大災害時等で発生するネットワーク輻輳による異常トラフィックを深層学習により予測検知して、ネットワーク制御により輻輳を回避することを目的とする。

ネットワークのトラフィックのような時系列データの予測に適した機械学習手法として RNN(Recurrent Neural Network)の適用を行った。しかし、RNN のような誤差逆伝搬のアルゴリズムでは、時系列が長くなると予測勾配の消失や発散などの問題が生じてします。そこで、RNN を改善した LSTM(Long Short-Term Memory)を使用し、RNN では扱うことができなかった長期依存のネットワーク・トラフィック予測の可能性を検討した。

深層学習のプラットフォームとして Chainer を使用し、2 台の PC から生成されるデータ・トラフィックをキャプチャし、教師データとし、LSTM による UDP および TCP のトラフィックの予測精度を評価した。評価は教師データで学習後、学習系に再度教師データを入力することにより評価した。

図 8 は、TCP 輻輳制御によるトラフィック変動を予測するためネットワーク内に遅延を入れて評価した結果である。図 8 は、入力データは教師データではなく、評価用データを用いた場合であるが、TCP 輻輳制御のトラフィック・パターンが正確に予測することができた。

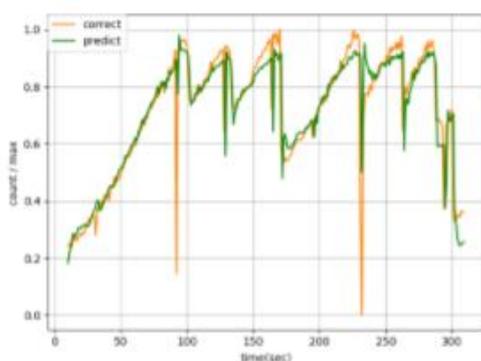


図 9 RNN による TCP 輻輳制御トラフィックの予測

ソーシャル・ビッグデータとして Twitter のテキスト・データを用い、WEB やニュースにないローカルなリアルタイム情報を抽出する技術について研究を行った。昨年度は

Twitter から災害などによるローカルな通信障害情報をリアルタイムで抽出し、いち早くネットワーク障害検知し、SDN ネットワーク制御と組合せたネットワーク制御の研究を行った。今年度は、前年度まで開発した技術を利用して、ユーザの時間、位置情報と、Twitter の日時、場所とその場所で行われる予定のイベントを抽出して、情報提供するサービスへの応用を進めた。

図 8 に前記サービスを実現するためのシステム構成を示す。Twitter からのイベント情報抽出には、SVM (Support Vector Machine) と Random Forest による機械学習を使用した。図 9 に示すようにイベントカテゴリーに対する正解率として F 値で評価した。その結果 Random Forest の方が比較的高い F 値が得られた。

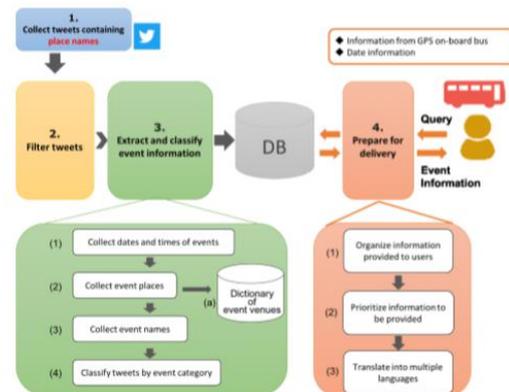


図 10 時空間イベント情報抽出システム

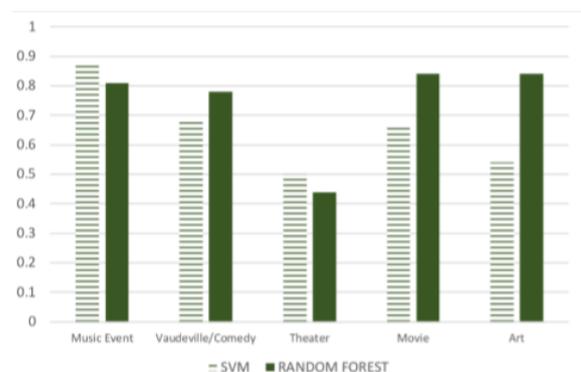


図 11 SVM と Random Forest の F 値

本開発システムを翻訳機能と組合せ、例えば

海外からのローカルな情報に不慣れな旅行者に対し、旅行者のロケーションやスケジュールに合わせてイベント情報を適切に配信する仕組みが提供できると考えている。

## 6. 今年度の進捗状況と今後の展望

2 で述べた 3 つの技術課題を実現するため 5 に示すように研究を進めている。これに加え、2 の後半に記述したように研究成果の実用性を高めるため社会実装に踏み込む研究を進めるため、実際の商用ネットワークの MVNO からのトラフィックを本研究で構築する通信基盤に誘導し、実トラフィックを用いた実証実験を行うためネットワーク構築およびシステム開発を進めている。

構築を進めているネットワーク構成の概要を図 10 に示す。研究室で管理している AS とアドレスブロックを使用して SINET とピアリングにより、MVNO 企業が接続している P-GW からのトラフィックを VPN 接続により研究室で開発したアプリケーション識別トラフィック収集システムと収集したデータを教師データとする深層機会学習エンジンの開発を進め、実トラフィックのアプリケーションを機械学習するための設備の構築を行った。図 8 は事業者 1 社であるが、現在 3 社についてネットワーク構築を行った。

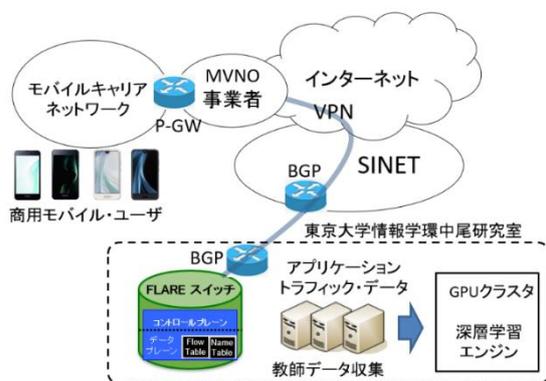


図 12 MVNO 事業者連携によるシステム構築

以上研究開発のみならず実ネットワークで

実用性を確認するためのシステム構築を行っている。

## 7. 研究成果リスト

### (1) 学術論文

### (2) 国際会議プロシーディングス

### (3) 国際会議発表

[1] P. Du, A. Nakao, “Deep Learnig-based Application Specific RAN Slicing for Mobile Networks,” IEEE International Conf. on Cloud Networking, Oct. 23, 2018

[2] K. Amemiya, Y. Akiyama, K. Kobayashi, Y. Inoue, S. Yamamoto, A. Nakao, “Ron-site Evaluation of a Software Cellular based MEC system with Downlink Slicing Technology” IEEE International Conf. on Cloud Networking, Oct. 22, 2018

[3] T. Kanaya, H. Yamauchi, A. Nakao, M. Oguchi, S. Yamamoto, and S. Yamaguchi, “Intelligent Application Switch Supporting TCP,” IEEE International Conf. on Cloud Networking, Oct. 23, 2018

[4] R. Kudo, M. Enoki, A. Nakao, S. Yamamoto, S. Yamaguchi, M. Oguchi, “Proposal and Evaluation of Event Search Method Based on SNS Data Analysis Focusing on Place and Time,” 5th IEEE/ACM International Conf. on Big Data computing, application and Technology (BDCAT2018), Dec. 17-20, 2018

### (4) 国内会議発表

[5] 雨宮宏一郎、秋山裕子、小林一成、井上義雄、山本周、中尾彰宏、“MEC を備えたソフトウェアモバイルネットワークシステムとスライシング技術のフィールド検証および評価”、電子情報通信学会総合大会 NS 研究会、NS2018-85, 2018 年 7 月

[6] 金谷知明、山内啓彰、葺澤慎之介、中尾彰宏、山本周、小口正人、山口実靖、“インテリジェントアプリケーションスイッチにおける TCP マイグレーションを伴う動的キャッシングによる KVS 性能の向上”、電子情報通信学会総合大会 NS 研究会、

NS2018-85, 2018 年 11 月

[7] 小山内遥香、中尾彰宏、山本周、山口実靖、小口正人、“RNN を用いたネットワークトラフィック変動の予測”、電子情報通信学会総合大会 NS 研究会、NS2018-85, 2018 年 11 月

[8] P. Du, A. Nakao, “Context-aware UE Slicing Leveraging Deep Learning and Software-Defined RAN,” 2019 年電子情報通信学会総合大会, BS-4-1, March 19, 2019

(5) その他（特許, プレス発表, 著書等）