

NTT DOCOMO

# テクニカル・ジャーナル

Technical Journal

Vol.29 No.4 | Jan. 2022

## DOCOMO Today

- グリーンキャリアを目指して

## Technology Reports (特集)

### AI特集 —AI技術のさまざまな業界への展開とAI活用を支える基盤技術の開発—

- DXを実現する自然言語処理技術
- 要約作業を効率化する多機能なニュース記事自動要約AIシステム
- ユーザ行動の時系列予測モデルを利用したレコメンドエンジンの開発
- ドコモ画像認識プラットフォームを用いた画像認識AI提供
- スポット画像特徴量を用いた観光スポットマイニング手法「ジェネリック観光地」抽出技術
- エッジAI対応5Gデバイスを活用したパーソナルモビリティ向け自動運転支援システム

## Technology Reports

- 5G MBHにおけるSegment Routing対応ルータ装置の開発
- ネットワーク仮想化基盤におけるETSI NFV Stage3仕様に準拠したマルチベンダ対応MANOへの移行

## Standardization

- 3GPPにおける産業用アプリケーション連携に向けたフレームワークの標準化



## グリーンキャリアを目指して

移動通信キャリアは、日本全体で自家用車約1,000万台／年相当のCO<sub>2</sub>を排出しています<sup>\*</sup>。端末の充電をはじめ、基地局の稼働や端末の製造・配送などに関連する排出量は非常に大きく、排出CO<sub>2</sub>削減という観点から、私たちは大きな社会的責任を抱えています。世界中で脱炭素社会への動きが加速するなか、ドコモは気候変動問題への対応を企業の重要な課題と位置付け、2021年9月に「2030年カーボンニュートラル宣言」を発表しました [1]。

この宣言においてドコモは、自社、バリューチェーン、お客様・パートナー企業の3つの軸で取り組んでいきます。まず自社では、ネットワークを省電力化する技術開発・設備導入や、自社専用の太陽光発電所などからの再生可能エネルギー（以下、再エネ）の調達などにより、自社のカーボンニュートラルを達成します。そして、バリューチェーンであるドコモショップへの太陽光パネル設置による再エネ化などを推進します。さらに、お客様・パートナー企業では、実質再エネで提供する「グリーン5G」の展開や、「ドコモでんき」によるグリーン電力の提供など、脱炭素に貢献する新たなサービスを提供するとともに、ドコモのサービスと連携したパートナー企業の環境配慮型商品・サービスを提供しお客様にご利用いただくことで、社会全体のカーボンニュートラルに貢献します。このような取組みにより、ドコモはグリーンキャリアになることを目指します。

この宣言の下、ドコモR&DにおいてはネットワークとサービスによるCO<sub>2</sub>排出量削減に向けた技術開発を推進します。ネットワークに関しては、基地局のスリープ機能の高度化や第5世代移動通信システム（5G）装置の省電力化、そしてIOWN（Innovative Optical and Wireless Network）ほか次世代ネットワーク・情報処理基盤などにおけるCO<sub>2</sub>排出量削減に寄与する技術など多岐にわたります。一方、サービスに関しては、例えば人の低炭素型への行動変容を促すようなサービスに関する技術など、さまざまな分野における先進ソリューションに貢献していきます。CO<sub>2</sub>排出量は、通信量の増加に伴い成行きでは右肩上がりに増加します。脱炭素社会の実現に向け、このようなネットワークやサービスにおけるさらなるイノベーション開発へのチャレンジが不可欠であり、ドコモR&Dへの期待はますます大きくなっていると感じます。

さて、私の所属するR&D総務部は、横須賀リサーチパークにあるドコモR&Dセンタにおいて、研究開発の環境作りと安心・安全なセンタ運営に取り組んでいます。コロナ禍においては、各種感染拡大防止策の強化やワクチンの職域接種などを行いました。最近のトピックスはセンタにおける脱炭素化に向けた取組みですが、省エネ



R&D総務部 部長  
くろかわ よしふみ  
黒川 善文

ルギー性能の高い空調設備導入などによる省電力化を段階的に進めています。また、大規模な太陽光パネル設置などによる再エネ化の拡大や、非化石証書活用による購入電力の実質再エネ化も計画しており、「研究開発段階からのカーボンニュートラル」を実現し、グリーンキャリア化に貢献したいと考えています。

最後に展示ホールのご紹介です。ドコモR&Dセンタでは、未来の移動通信とスマートライフを実感できる展示ホール「WHARF」を一般公開しています。これまで約14万人と大変多くの皆様にご来館いただいています。5G・第6世代移動通信システム（6G）の要素技術のほか、ヒューマノイドロボット、XR、eスポーツなどさまざまな分野における5G活用事例などを一堂に集め、お客様が最新技術を体感していただける展示となっております。今後はカーボンニュートラルに関連するイノベーション事例なども追加していく予定です。なお、受付では自律走行型ロボットがお客様をご案内していますが、このような新しい技術の実証実験の場としてもR&Dセンタを積極的に活用していきたいと思っています。

### 文 献

- [1] NTTドコモ報道発表資料：“2030年までに温室効果ガス排出量を実質ゼロにする「2030年カーボンニュートラル宣言」を発表—お客様・パートナー企業とともに社会全体のカーボンニュートラルに貢献する取り組みを開始—” Sep. 2021.

<sup>\*</sup>国内の移動通信キャリア（ドコモ・KDDI・ソフトバンク・楽天）が対外的に発表しているCO<sub>2</sub>排出量（2019年度）を足し合わせ、それを自家用車1台当りのCO<sub>2</sub>排出量で割り、台数を算出。

## [ Contents ]



### DOCOMO Today

グリーンキャリアを目指して 黒川 善文 1



### 特別寄稿

IoTとSociety 5.0への期待 安本 慶一 4

## Technology Reports (特集)

### AI特集 —AI技術のさまざまな業界への展開と AI活用を支える基盤技術の開発—

DXを実現する自然言語処理技術 7

DX

自然言語処理

文書分類

要約作業を効率化する多機能なニュース記事自動要約AIシステム 15

AI

自動要約

自然言語処理技術

ユーザ行動の時系列予測モデルを利用したレコメンドエンジンの開発 26

行動予測

レコメンド

RNN

ドコモ画像認識プラットフォームを用いた画像認識AI提供 33

AI

画像認識

Deep Learning

スポット画像特徴量を用いた観光スポットマイニング手法  
「ジェネリック観光地」抽出技術 41

スポット推薦システム

観光情報学

ニューラルネットワーク

エッジAI対応5Gデバイスを活用したパーソナルモビリティ向け  
自動運転支援システム 48

5G

エッジAI

自動運転



(P.33)



(P.48)

# Technology Reports

5G MBHにおけるSegment Routing 対応ルータ装置の開発

55



(P.65)

MBH

SR-MPLS

TI-LFA

ネットワーク仮想化基盤におけるETSI NFV Stage3仕様に準拠した  
マルチベンダ対応MANOへの移行

65

ネットワーク仮想化

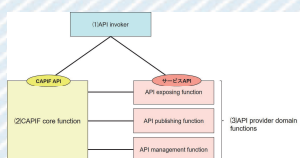
オーケストレーション

オンラインアップグレード

## Standardization

3GPPにおける産業用アプリケーション連携に向けた  
フレームワークの標準化

76



(P.76)

CAPIF

SEAL

API



(P.83)

## News

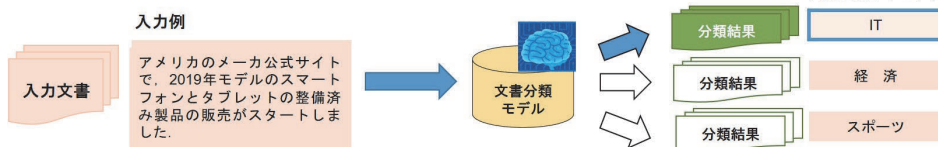
2021年「日本ITU協会創立50周年 各賞」受賞

83

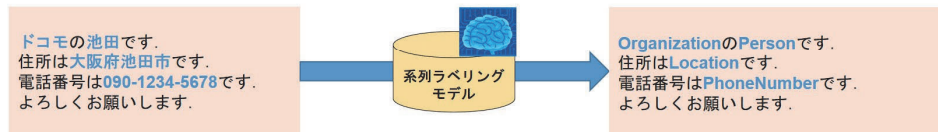
2021年「日本ITU協会賞」受賞

84

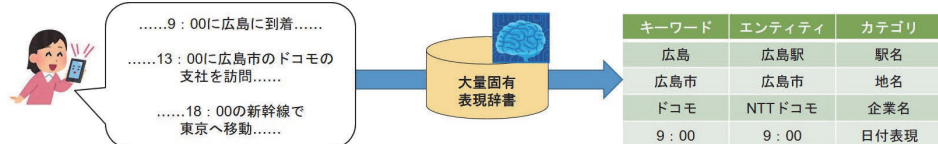
①文書分類：事前学習した文書分類器に基づき大量文書を高速に分類



②系列ラベリング：文中に含まれる個人情報や誹謗中傷表現などを特定してラベリング



③エンティティリンキング：Wikipediaから生成した大量固有表現辞書により文中から固有名詞などを抽出



Technology Reports (特集) DXを実現する自然言語処理技術 (P.7)  
自然言語処理AIの機能一覧

# IoTとSociety 5.0への期待

奈良先端科学技術大学院大学 先端科学技術研究科 情報科学領域 教授 やすもと けいいち 安本 慶一さん

2021年3月に閣議決定された第6期科学技術・イノベーション基本計画には、第5期計画で掲げられたSociety 5.0、すなわち、「サイバー空間とフィジカル空間を高度に融合させたシステムにより、経済発展と社会的課題の解決を両立する人間中心の社会」を具体的に実現していくことが盛り込まれた。Society 5.0を実現するためには、IoT (Internet of Things) やセンサの技術を使ってフィジカル空間の情報をサイバー空間に取り込み、AIによりデータを高速に分析し、分析結果として得られた知見を現実世界にフィードバックするサイバーフィジカルシステム (CPS: Cyber-Physical System) をあらゆる場所に浸透させなければならない。そのために、家、仕事場・学校、街、自然環境など、世界中のあらゆる場所でデジタルデータを収集できるようにするIoTの役割は大きい [1]。なお、世界中のIoTデバイス数は2020年時点で253億台に達しているが、2023年には340億台と急激に増加 (特に、医療、産業、コンシューマ、自動車・宇宙航空用途において) すると予想されている [2]。

現在、IoT化された家、すなわち、スマートホームに関する研究が世界中で実施されている。私の研究室では、スマート家電や各種センサを設置したスマートホームを学内に設置し、居住者の行動に伴い生成されるセンサデータ (人感センサやドアの反応、家電の消費電力や環境センサデータの変化など) を機械学習で分析することにより、居住者がどのような生活行動を行っているのかをリアルタイムに推定するシステムを開発している。また、一般住宅に後付可能なスマートホーム化キットSALON [3] を開発している。SALONは、屋内光で発電し動作する人感センサとドアセンサ、ボタン電池で長期間動作する環境センサ、センサデータ収集用ゲートウェイ、ラベリングツールから構成されている。本システムを一般の高齢者宅に設置し生活行動データを取得した結果、70%弱の精度で生活行動を推定可能であることが分かった。研究用途のスマートホームは国内外でさかんに開発されている。例えば、横浜市、ドコモ、and factoryにより創設された「未来の家プロジェクト」では、実証実験用にトレーラーハウス型IoTスマートホームが開発されており、居住者の行動推定に加え、体重や睡眠時間、摂取カロリーなどの健康関連情報も取得可能にするIoTデバイスを設置している。他方、住宅メーカーはHEMS (Home

Energy Management System)\*<sup>1</sup>を搭載しエネルギーの効率的な利用に注力した「スマートハウス」を販売しているが、行動認識機能を備えるものは未だ開発されていない。今後の行動認識精度の向上による実用化・市販住宅への搭載が切望される。

ただし、行動認識結果のフィードバック方法については、さらなる研究の進展が必要である。将来的には、行動を認識・予測するだけでなく、居住者の健康やQoL (Quality of Life) を把握・予測し、QoLを改善するための行動を誘導する技術が求められる。一例として、IoT化した箸を使って食行動をセンシングし、食べる間隔の速さ、選んだ食材によって、ピースに分かれた絵に色が塗られる仕組み (ゆっくりとバランスよく食べるほど、色が重ならず鮮明な絵になる) を通して、より健康的な食行動を誘発するシステムeat2pic [3] が開発されている。今後あらゆる場面で人の行動変容を促す情報技術が求められると考えている。

スマートシティは「IoT化された街」ととらえることができる。ただし、街の至る所にIoTデバイスを設置するのは、導入・運用コストなどさまざまな理由から難しい。そこで、誰もが所持するスマートフォンを活用し、人を移動型IoTデバイスとみなす「ユーザ参加型センシング」が活用されている。参加型センシングでは、一般ユーザに貢献してもらうことが重要であり、インセンティブの提供が必須である。文献 [4] では、航空会社のフリークエント・フライヤーズ・プログラム (マイレージサービスなど) にヒントを得て、センシングタスクをこなして一定以上のポイントを貯めることで上級ステータスに達したユーザが、同じセンシングタスクでもより高いポイント (ボーナスマイル) を取得できる仕組みを設定することで、上級ユーザになろうとするインセンティブを与え、センシングタスク受諾率を向上させている。

参加型センシングにおけるユーザの役割は、システムより依頼されたタスクに対し、データ (あるスポットの写真や、状況を説明するテキストなど) をアップロードすることである。さらに、センシングタスクを自動化することで、ユーザの負担を軽減することができる。桜センサ [5] は、スマートフォンを自動車のダッシュボードに設置し、運転中の前景をドライブレコーダと同様にビデオ録画する。並行して、フラクタル解析や色ヒストグラム解析といっ



## Profile

1991年大阪大学基礎工学部情報工学科卒業。1995年同大学大学院基礎工学研究科博士後期課程退学。博士（工学）。滋賀大学経済学部助手、講師、助教授を経て、2001年奈良先端科学技術大学院大学情報科学研究科助教授。2011年より同研究科教授。1997年モンテリオール大、2004年イリノイ大客員研究員。IoT、ユビキタスコンピューティング、分散システムなどの研究に従事。情報処理学会理事・関西支部長を歴任。同学会山下記念研究賞、論文賞など受賞。

た簡易な画像解析をビデオフレームに実時間適用することで、開花している桜およびその開花度合いを検出し、位置情報とともにサーバにアップロードする。これにより、桜のリアルタイム開花情報が自動的に収集される。同様の方法で、ドライブレコーダ映像から道路上の違法駐車車両をリアルタイムに検出・記録する研究も実施されている。

スマートシティでは、各スポットや交通機関における混雑度をリアルタイムに把握できることは必須の機能である。しかしながら、カメラやLiDAR (Light Detection And Ranging)<sup>\*2</sup>などを使った混雑度計測システムを街全体に導入するのはコスト的にも、プライバシー的にも難しい。そこで、電波を用いて低コストに混雑度をセンシングする技術が注目を集めている。私の研究室では、Wi-Fi CSI (Channel State Information)<sup>\*3</sup>を用いた群衆センシングや、乗客のスマートフォンが発するBLE (Bluetooth Low Energy) 信号<sup>\*4</sup>を使ったバス車内混雑度センシングなどを開発している。奈良交通株式会社の協力の下、複数バス路線に適用し、246人の平均絶対誤差<sup>\*5</sup>でバス車内の人数カウントができるという結果が得られた。

ここまで、街全域のさまざまな情報を収集する方法について述べたが、これらの情報を人々にフィードバックし、行動変容を促す技術が必要である。ISO-Tour [3] は、観光スポットの訪問時間帯別の満足度（混雑や景観の時間変化により満足度は変化）の情報を使って、より適時に観光スポットを訪問できるように観光客を誘導する。一方で、センシング参加者から収集した動画や写真は膨大なデータとなるため、必要な情報のみを必要なユーザに効率良く届けるには、データを分析し、分かりやすい情報にまとめる「キュレーション」が必須となる。私の研究室では、複数の観光スポットを順に観光する際の移動経路のビデオと、観光スポットでの景観のスライドショーを含む1分程度の短編ビデオを自動生成する手法を開発している。さらに、移動中のビデオにおいて、単調なシーンを高速に再生し変化のあるシーンを低速に再生することで、ビデオ全体を短縮しても移動経路中の雰囲気が変わらないような工夫がなされている。この手法を応用し、観光用レンタカーのドライブレコーダが記録した走行中ビデオから、観光的に特徴のあるシーンを特定・抽出し、旅の思い出を短編ビデオに自動的にまとめるdash-cam

[3] を開発している。

以上、IoT技術のさまざまな応用について述べた。Society 5.0の実現には、IoTデバイスのさらなる浸透と、より高度なAI技術・フィードバック技術の開発が待たれる。高度化したIoTデバイスが浸透し、新しいサービスを生むためには、第5世代移動通信システム（5G）やBeyond 5Gの高速・多数同時接続・低遅延の通信が必須である。今後のIoT・CPSの研究の発展とSociety 5.0の早期の実現を願っている。

## 文 献

- [1] 安本 慶一, 荒川 豊, 松田 裕貴: “IoTの本,” 電気書院, 2021.
- [2] 総務省: “令和3年版情報通信白書,” 2021.  
<https://www.soumu.go.jp/johotsusintokei/whitepaper/r03.html>
- [3] 奈良先端科学技術大学院大学ユビキタスコンピューティングシステム研究室 (NAIST-UBI): “Ubi Research Project Catalogue.”  
<http://research.ubi-lab.com/>
- [4] Y. Ueyama, M. Tamai, Y. Arakawa and K. Yasumoto: “Gamification-based incentive mechanism for participatory sensing,” IEEE PerCom Workshops, 2014.
- [5] S. Morishita, S. Maenaka, D. Nagata, M. Tamai, K. Yasumoto, T. Fukukura and K. Sato: “SakuraSensor: quasi-realtime cherry-lined roads detection through participatory video sensing by cars,” ACM UbiComp 2015, Sep. 2015.

\*1 HEMS: 経済産業省によれば、「快適な室内環境」と、「年間で消費する住宅のエネルギー量が正味でおおむねゼロ以下」を同時に実現する住宅のことをいう。

\*2 LiDAR: 近赤外光や可視光を使って対象物に光を照射し、その反射光を光センサでとらえ距離を測定するリモートセンシング方式。

\*3 Wi-Fi CSI: 通信リンクのチャネル特性の1つであり、Wi-Fiの信号が送信機から受信機まで伝搬する際の、散乱・フェージング・信号強度の減衰などの複合的な効果を表す。

\*4 BLE信号: 低電力消費・低コスト化の図られた近距離無線通信技術規格。

\*5 平均絶対誤差: 各データに対して、推定値と正解値の差（誤差）の絶対値を計算し、その総和をデータの数で割った値。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

# AI

# 特集

## AI 技術のさまざまな業界への展開と AI 活用を支える基盤技術の開発

---

スマートフォンが日常生活に溶け込み、また 5G のサービス開始に伴い高速・低遅延の通信環境が現実のものとなった現在、世界中でスマートフォンや通信サービスを通じた生活の利便性向上を目的として AI の研究開発がさかに行われており、ドコモグループでも活発に取り組んでいる。

AI の基本は機械学習や統計分析にあり、根源的なタスクはデータの分類とそれに伴う示唆となる一方、その応用は幅広いものとなっている。

NTT DOCOMO テクニカル・ジャーナルでは定期的に AI を特集しているが、今回の特集では、分類に関しては言語処理と画像認識に関する取組みを、また行動支援に関してはレコメンドエンジンや自動運転の研究開発について紹介する。

これらの事例では単純な技術の適用にとどまらないさまざまな挑戦が行われており、AI の可能性を拡げる応用例が多数提示されている。これらが読者の AI への理解をさらに深め、従来の枠にとどまらない新たな AI の創造につながることを期待している。

(編集部)

---

# DXを実現する自然言語処理技術

サービスイノベーション部  
 しらみず ゆうたろう おか けいすけ  
 白水 優太郎 岡 慶介  
 たつみ しゅうすけ  
 辰巳 守祐

近年、AIを利用してDXの推進に資する技術開発がさかんである。特に、企業に蓄積される大量のテキストデータを高速かつ正確に処理するAI技術として、自然言語処理と呼ばれる分野に注目が集まっている。本稿では、ドコモが開発している自然言語処理AIとその技術的特徴、およびAIの利用をサポートするGUIツールを紹介するとともに、実際の導入事例やRPAツールとの相乗効果、今後の展望について解説する。

## 1. まえがき

近年、「働き方改革」が政府主導で進められており、RPA（Robotic Process Automation）ツール<sup>\*1</sup>やチャットボット<sup>\*2</sup>をはじめとするデジタル技術を活用した、ビジネスの加速、生産性の向上、稼働の削減などが強く求められている。特にRPAツールは、アンケートやユーザからの問合せをはじめとする大量のテキストデータの処理業務において改善効果が大きく、フロントオフィスやバックオフィスを中心に普及が進んでいる。さらに、テキストのカテゴリ分類やタグ付けのように、知的な判断を必要とする作業のデジタルトランスフォーメーション（DX：Digital Transformation）<sup>\*3</sup>はRPAツールだけでは実

現が困難であることから、人間の言語理解を代替できる「自然言語処理」と呼ばれる分野のAIによる自動処理も需要が増大している。

しかし、実際の現場では利用できる計算資源や求められる性能が案件ごとに異なるためAIの多様性が必要であること、SaaS（Software as a Service）<sup>\*4</sup>で社内のテキストデータを処理するのはセキュリティ面で抵抗感があること、システムエンジニアを抱えていない部門が独力でAIを導入して運用を担うのは負担が大きいことなどの諸問題が立ちはだかり、AIの導入は決して容易ではない。

そこでドコモは、計算資源を存分に活用した高性能な自然言語処理アルゴリズムを開発し、加えて、省リソースなPCでも動作する軽量高速なアルゴリ

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

\*1 RPAツール：パソコン操作をシナリオとして記録させ、パソコン操作を自動化するソフトウェア型ロボット。

\*2 チャットボット：音声やテキストチャットを介して、人との会話を自動的に行うプログラム。



ズムも開発した。これにより、クラウドやローカルの処理環境、PCやサーバのスペック、利用者のセキュリティポリシーなど、ユーザの要件に合わせた柔軟な自然言語処理AIの選択・利用が可能となった。さらに運用者の負担を軽減するために、これらのAIを手軽に利用したりチューニングしたりできるGUI (Graphical User Interface)\*5ツールも併せて開発した。

本稿では、テキストのカテゴリ分類や個人情報の秘匿化など、現場の多様なニーズに広く対応した自然言語処理AIやGUIの技術開発、ならびに社内外での活用事例について解説する。

## 2. DXに推進に資する自然言語処理AI

### 2.1 概要

近年、自然言語処理の学術研究は世界中で活発に

行われており、学術研究における整備された実験設定下では、高い性能を誇るアルゴリズムが数多く提案されている。また、これらのアルゴリズムはOSS (Open Source Software)\*6として公開されることが多く、最新のアルゴリズムを誰でも手軽に利用できる。一方、ビジネス現場では、利用できる計算資源やデータに制約が多く、アルゴリズムをそのまま適用してもニーズを満たせない。そこでドコモでは、最新のOSSも活用しつつ、前述したような現場の多様なニーズに対応するため、追加機能の開発や展開方法の工夫に取り組んできた。

### 2.2 機能

#### (1) 文書分類

文書分類機能は、文書にラベルを自動付与する機能である (図1①)。ドコモでは、ニーズに応じて、軽量な分類器と高性能な分類器の2種類の分類器を

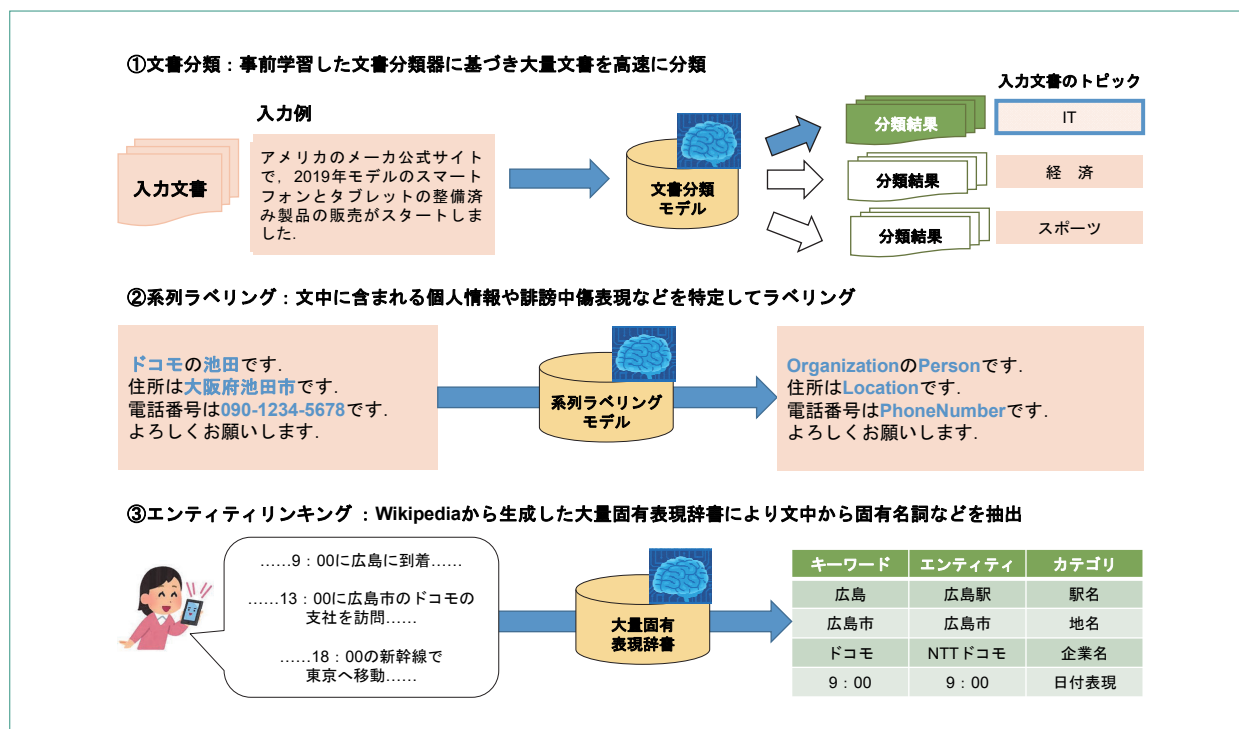


図1 自然言語処理AIの機能一覧

- \*3 デジタルトランスフォーメーション (DX)：IT技術を活用してサービスやビジネスモデルを変革させ、事業を促進するとともに人々の生活をあらゆる面で良い方向に変化させること。
- \*4 SaaS：インターネットなどのネットワーク経由で利用するソフトウェアのこと。
- \*5 GUI：ボタンやアイコンの組合せからなり、直感的な操作性や

- 優れた視認性を特長とするインタフェース。
- \*6 OSS：ソースコードが無償で公開されており、誰でも再利用や改変が行えるソフトウェア。

提供している。

#### (a)軽量な分類器

軽量な分類器は、コストやセキュリティの観点からクラウドが使えない場合など、通常業務用PCのローカル環境での利用ニーズに応えたものである。アルゴリズムには、実装が単純で動作が軽く、省リソースなPCでも実行可能な多層パーセプトロン<sup>\*7</sup>を利用している。また、特徴量<sup>\*8</sup>として、一般的に利用される形態素<sup>\*9</sup>に加えて、文字そのものや文字N-gram<sup>\*10</sup>も利用している。この工夫により、少量のコーパス<sup>\*11</sup>でも高い精度が得られ、さらには誤字脱字にも頑強な分類器を構築することができる。これらの特徴量が有効に機能する事例としてFAQ分類を図2に示す。FAQ分類では、入力された質問文を、用意されたFAQに自動分類する。この際、入力文に誤字脱字がある場合、形態素のみでは正しく分類できない。一方、文字や文字N-gramを利用すれば、部分文字列（図中のサンプルでは「b, l, u, e, t, o, …」）をヒントに推論することで正しく分類できる。

#### (b)高性能な分類器

高性能な分類器は、GPU（Graphics Processing Unit）<sup>\*12</sup>が利用できるクラウドサーバかオンプレミス<sup>\*13</sup>サーバを保有している利用者の、潤沢な計算資源を最大限活用して高い性能の分類器を構築したいというニーズに応えたものである。アルゴリズムは、近年の言語処理のデファクトスタンダードであるBERT（Bidirectional Encoder Representations from Transformers）[1]を採用した。BERTは大量の学習データと計算資源を必要とするが、性能が非常に高い。BERTの事前学習モデル<sup>\*14</sup>には、NTT人間情報研究所の成果技術であるNTT版BERT<sup>\*15</sup>を利用している。

また、マルチラベル分類<sup>\*16</sup>への対応をドコモで独自で行った。NTT版BERTはシングルラベル分類<sup>\*17</sup>しか対応していなかったが、ビジネス現場ではマルチラベル分類のニーズが高い。そこで、ドコモでは、NTT版BERTのラベル出力部分の関数を工夫することで、出力をマルチラベル化し提供している。

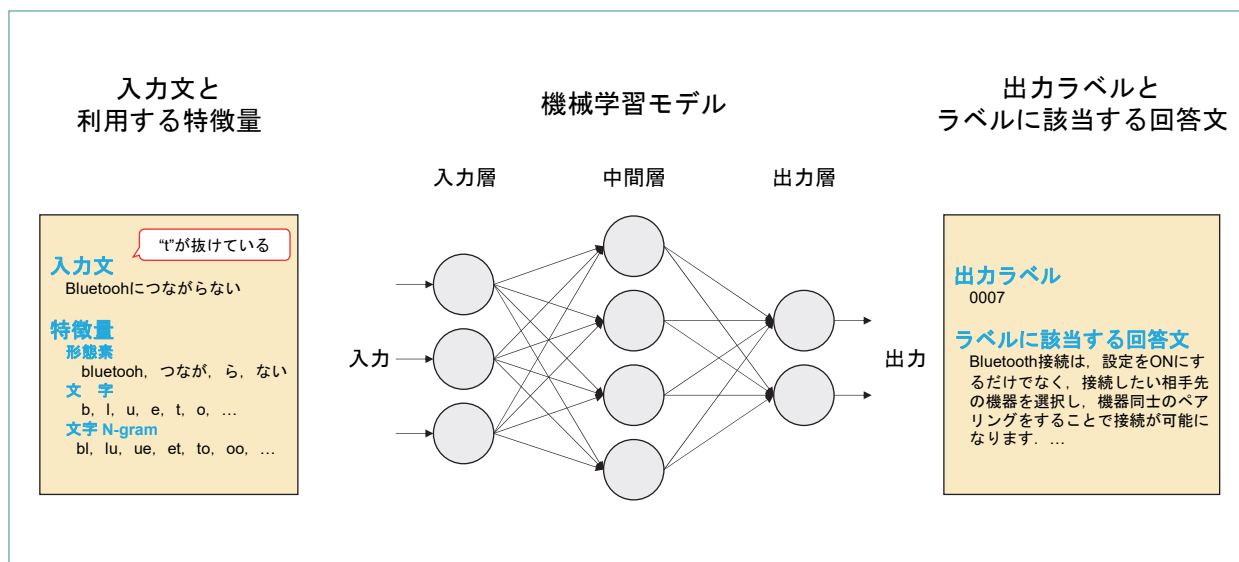


図2 軽量な分類器の特徴量が有効に機能する事例

- \*7 多層パーセプトロン：機械学習アルゴリズムであるニューラルネットワークの一種。
- \*8 特徴量：データから抽出される、そのデータを特徴付ける量（数値）のこと。
- \*9 形態素：言語表現の最小単位。
- \*10 N-gram：任意の連続した $n$ 個の要素。要素の例としては、単語

や文字などが挙げられる。

- \*11 コーパス：テキストを大量に収集してデータベース化した言語資料。
- \*12 GPU：並列的な計算処理に優れたプロセッサ。並列計算を要する深層学習と相性が良い。

## (2)系列ラベリング

系列ラベリング機能は、系列データ<sup>\*18</sup>にラベルを自動付与する機能である。ドコモは、この機能を個人情報や誹謗中傷表現をはじめとする、テキストのマスキングに利用している（図1②）。当機能の開発においてドコモは、軽量なアルゴリズムを扱えて、かつ開発者コミュニティの大きいOSSとして知られるFlair [2]を採用した。OSSの選定において、アルゴリズムの軽量を重視した理由は、機密性の高いデータの処理では、セキュリティの観点からクラウドの利用を避けたいというニーズが多く、ローカル環境でも動作する、軽量なものが求められるためである。また、開発コミュニティの大きさを重視した理由は、多様な利用ユーザのフィードバックに基づくインターフェースの改善や継続的なバグフィックスが期待できるためである。コミュニティの小さいOSSでは、利用マニュアルやノウハウの貯蓄が少なく、運用が属人化してしまう傾向があるため、これの利用を避けた。

## (3)エンティティリンキング

エンティティリンキング機能は、文中からキーワードを抽出し、さらにそのキーワードをエンティティと紐づける機能である。例えば、「ドコモ」「エヌティーティー・ドコモ」「DOCOMO」といった語は、一般的な単語（例えば、「机」や「電車」など）とは違い、固有名詞である。さらに、これらの語は文字列こそ異なるものの、いずれも「NTTドコモ」という同一の概念（エンティティ）を指している。このように、「文中の何がキーワードか（固有の意味をもっているか）」「そのキーワードがどのエンティティを指すか」を機械的に抽出・推定する機能がエンティティリンキングである（図1③）。

## (a)Wikipediaからの抽出データに基づく辞書の構築

ドコモでは、特にWikipedia [3]から抽出したデータをエンティティリンキングに活用している。抽出された記事本文、アンカーテキスト、

ページ閲覧数などの各種データは統計的に処理され、キーワードの重み付けやエンティティ・キーワード辞書の構築に利用される。Wikipediaを利用する利点として、①記事の新規作成・更新が頻繁であり新語や流行語などへの対応が早いこと、②毎日新たにデータベース・ダンプ<sup>\*19</sup>が公開されておりデータを利用しやすいこと、③データ抽出から辞書構築までを一気通貫で自動処理できるため、辞書のメンテナンスコストが低いことが挙げられる。

## (b)カテゴリの付与

さらにドコモでは、「拡張固有表現階層<sup>\*20</sup>」[4]に基づくカテゴリを上位概念としてエンティティに付与しており、抽出した結果を200種類程度のカテゴリに分類した上で、それらを利用する。例えば、対話システムでの活用を考えた場合、「ラーメンを食べたい」「りんごを食べたい」のような、想定されるすべてのユーザ発話に対して1つひとつシステム応答を用意するのは現実的ではない。しかし、「ラーメン」「りんご」に付与されるカテゴリが「食べ物名」であることを利用すれば、「[[食べ物名]を食べたい」だけに対してシステム応答を用意すればよいので、応答シナリオ作成の負担を大きく低減することができる。

## (c)活用事例

エンティティリンキングの想定活用事例として、ニュースからのキーワード抽出が挙げられる。Wikipediaは時勢を反映して記事内容が更新されたり特定のページ閲覧数が伸びたりするため、Wikipediaの統計データを利用したエンティティリンキングは、ニュース本文からのキーワード抽出と相性がよい。加えて、単純な文字列抽出のみではなく「表記ゆれの吸収」も同時に行うため、抽出・リンキングされたエンティティあるいはエンティティに付与されたカ

\*13 オンプレミス：企業がシステムを構成するハードウェアを自社で保有し、自社で保守運用すること。

\*14 事前学習モデル：目的タスクの教師あり学習実施前に、大量のコーパスで教師なし学習したモデル。

\*15 NTT版BERT：NTT人間情報研究所が独自で学習データを収集し、そのデータで事前学習させたBERTモデル。

\*16 マルチラベル分類：1つのデータに複数のラベルを付与する分類方式。

\*17 シングルラベル分類：1つのデータに1つのラベルを付与する分類方式。

\*18 系列データ：文字列や音声波形、購入履歴などのように、要素が直列に並んでいるデータのこと。

テゴリを記事のタグとして活用しやすいというメリットがある。

### 3. 自然言語処理AIの展開を加速する GUIツールの開発

#### 3.1 概要

自然言語処理AIの導入や活用には、AIの精度を維持・向上していくためのチューニングが欠かせない。チューニングとは、AIの精度検証、アノテーション\*21作業による学習データの修正など、AIモデルの構築と継続的な更新に発生する一連の作業を指し、運用者による目視の作業なども伴うことから、一定の人的な運用コストが発生する。

このような作業にかかる運用コストを軽減するために、ドコモはAIのチューニングをサポートするGUIツールも併せて開発している。このGUIツールは、構築およびチューニングのような繰り返し作業が必要となるAIの学習・評価を、シンプルな画面操作で手早く実行することができる。また、ユーザが学習させたAIの予測結果をアノテーションに活用して、人手の作業量を調整したり作業の優先度を決めたりできるので、効率的な作業が可能である。さらに、構築したAIとデータはリビジョンという単位でシステムにより管理されるので、どのデータを利用してどのAIをいつ構築し、その精度がどれくらいだったかといった煩瑣な情報管理をユーザが意識する必要がなくなる。導入の観点からは、ローカル/クラウドサーバといった作業環境やホスト側のOSに依存しないよう、ツールのバックエンド\*22エンジンをコンテナ仮想化技術\*23の上で動作させている。また、GUIツールも併せてコンテナ化することで、バックエンドからフロントエンド\*24までを簡便に導入できる。

このように、容易に導入できるGUIツールを利用することで、画面に沿った操作のみで手軽にAI構

築、チューニング、管理を行うことが可能となり、自然言語処理AIの導入・運用のハードルを軽減させる効果が期待できる。

#### 3.2 GUIツールの開発とその機能

本ツールでは、AIの構築に必要なデータセットのアップロード、アノテーションデータ作成、学習、評価など一連の機能をGUIとして提供しており、AIの初期構築からメンテナンスに至るまでのすべてをツール画面上で実行可能である(図3)。

主な機能詳細を以下に示す。

##### (1)学習・評価

利用ユーザが、あらかじめ用意したアノテーション済みのテキストデータをツール経由でバックエンドサーバへアップロードし、簡易な画面操作を行うことで学習が行われ、独自のAIを作成できる。また作成したAIが、入力に対して自動付与するラベルの精度を判定するため、正解として用意したアノテーション済みデータをGUIツールへアップロードすると、本ツールが正答率を自動で算出し、精度評価することも可能である。

このように、利用ユーザはアノテーション済みのテキストデータを準備するだけで、容易にAIの構築と精度評価が可能となる。

##### (2)チューニング

学習済みのAIを用いて、テキストデータに自動でラベルを付与し、ユーザがラベルの確認・修正を行うことが可能な機能を提供している。一般にAIは、人によってラベル付けがなされたテキストを学習することで、人の判断を再現することが可能となり、学習テキストを増やしていくことで精度が向上する。ある程度学習を繰り返したAIにテキストのアノテーションを任せ、間違った部分を人間の判断で修正することでアノテーションの作業の効率化が可能となる。そして、アノテーションしたテキストをさらにAIに学習させることによる精度向上も可

\*19 データベース・ダンプ：データベースの内容をそのままファイルに出力したもの。

\*20 拡張固有表現階層：語を「人名」「市区町村名」「国名」のような200種類程度のカテゴリに分類したもの。カテゴリは「地名>天体名>惑星名」のように、最大3階層の構造になっている。

\*21 アノテーション：テキストや画像などのデータに対して人手で

ラベルを付与すること。

\*22 バックエンド：GUIを動作させるためのシステム部分。主にエンジンおよびエンジンとGUIの動作をつなぐシステム部分のこと。

\*23 コンテナ仮想化技術：アプリケーション本体や、アプリケーションに必要なファイル群を「コンテナ」としてパッケージングし、コンテナエンジンというプロセス上で動かす技術の1つ。



図3 チューニングフローとGUIツールの画面イメージ

能である。

このように、本機能を利用することで、ラベル付きテキストデータの作成コストを軽減しつつ、AIの精度向上も実現できる。

また、学習時のハイパーパラメータ<sup>\*25</sup>の変更、学習実行、精度評価などが実行できる画面も提供しており、チューニング時のパラメータ調整<sup>\*26</sup>が容易である。

### (3)リビジョン管理

チューニングを行う際には、過去のデータセットや評価結果の管理の煩雑さが往々にして運用上の課題となる。本ツールでは、学習したAIや学習に利用したデータセット、評価サマリなどを、「リビジョン」という作業単位ごとに管理できる機能を提供しており、前述したような煩雑な管理をユーザが意識することなく、過去のデータセットや評価サマ

りをシステム側で個別に管理し、各データセットのチューニングの効果やデータの差分を分析することが可能である。

## 4. 活用事例の紹介

### 4.1 概要

ドコモでは、自然言語処理AIやGUIツールの開発のみならず、社内外で自然言語処理AIによるDXを推進している。

社内においては、ユーザや従業員からの意見・要望がテキストデータとして集約されるCS (Customer Satisfaction)<sup>\*27</sup>部門と相互に協力し、自然言語処理AIの構築や導入サポート、蓄積してきた人手による分類結果の共有などを通して、ユーザ満足度向上のためのオペレーション改善に努めている。

<sup>\*24</sup> フロントエンド：ユーザが目にして操作することが可能なシステム部分。主にGUI部分のこと。

<sup>\*25</sup> ハイパーパラメータ：学習時の設定値のこと。設定値により性能が変化するため、最も性能が良くなるように最適化することが必要。

<sup>\*26</sup> パラメータ調整：性能が最も良くなるように関連する設定値を

選定すること。

<sup>\*27</sup> CS：顧客満足のこと。

また社外においても、法人営業部門と共同でRPA - AI連携ソリューション展開を進めており、自然言語処理のみならずOCR（Optical Character Recognition）\*28や音声認識\*29技術をも活用した実証実験を広く実施している。

## 4.2 社内CS部門との取組み

ドコモでは、ドコモショップやコールセンタなどでユーザや従業員からいただいた意見・要望を、特定の個人を識別することができないよう加工した上で分析し、ユーザ満足度・従業員満足度の向上に繋がるようなサービスの開発や業務の改善に活用している。例えば、ユーザより受領したアンケートは専門のスタッフが目を通し、「要望」「賞賛の声」などの単位でまとめられて関連部署へ共有される。しかし、日々届くアンケートは大量であることから精読には時間がかかる。また、人の目で確認する以上、見落としや分類ミスなどによる精度の低下などは避けられない。さらに、分類作業はスタッフの知識や経験に依存するため、作業内容の質の均一化が難しいという問題がある。

そこでドコモは、ユーザ要望の速やかかつ適切な実現やオペレーションの効率化を目指し、自然言語処理AIの導入によるアンケート自動分類化を社内CS部門と進めてきた。収集したアンケートは社内閉じたネットワークの中で利用・分析されており、外部の高スペックなサーバによる自動処理が難しかったことから、前述した軽量な分類アルゴリズムをアンケート分類のAIとして開発・提供した。また、CS部門にはこれまで人手で分類してきたデータの蓄積があったことから、これらをAIの学習データとして活用した。

自然言語処理AIを作業に導入することで、これまで長時間人手をかけて実施していた内容の大半が自動化され、大きな稼働時間の削減効果があった。今後はアンケートの自動分類のみならず、ユーザが

誤って入力してしまった氏名や住所などの個人情報、人の手を介さずに自動で秘匿するAIを導入することで、ユーザにより一層満足いただける価値の提供を目指す。

## 4.3 RPAツールとAIの連携ソリューション

近年、自然言語処理やOCR用のAIを「頭脳」に見立ててRPAツールと組み合わせることで、より複雑な業務を自動化するソリューション（Cognitive Automation\*30）が模索されている。ドコモにおいても、RPAツール「WinActor」の導入・拡大およびRPA - AI連携ソリューションの一環として、「既存システムへの電子カルテ自動投入」の実証実験を、奈良県総合医療センターと共同で実施した。

医療現場には、診察や検査といった「人にしかできない業務」のほかに、患者ごとのカルテ作成や電子カルテシステムへのデータ入力など、ITによる効率化が可能な業務も多くある。しかし、特にシステムへのデータ投入にあたっては、データ内のテキストを「理解」した上で適切な項目に転記する必要があるため、RPAツールのみでの対応は容易ではない。また、個人情報を扱うことから、外部ネットワークとのデータ通信が発生するSaaSの利用よりも、作業用PC間のローカルネットワークに閉じた処理が比較的好まれる。加えて、院内ではデータベースやシステムがすでに導入済み・稼働中である場合が多く、これらを改修しないでそのまま活用できることが求められる。

以上の点を踏まえ、ローカルでも動作する軽量な自然言語処理AIと外部サーバの不要なRPAツールを、院内のローカルネットワークに接続しているPC上だけに構築し、それらを連携させて、これまで人間が実施していたシステムへの転記作業をそのまま自動化することで、後続のフローに大きな変更を加えることなく医療従事者の稼働削減による超過勤務解消を図った。さらに音声認識ソフトを新たに

\*28 OCR：文字認識技術のこと。

\*29 音声認識：人間の発話音声解析し、テキスト化したり感情を推定したりする技術のこと。

\*30 Cognitive Automation：自然言語処理や画像認識、音声認識などのAIとRPAツールとを組み合わせることで、人間が行っている「判断に基づく作業」を自動化すること。

導入し、これまで紙に控えていた患者情報を音声発話から直接電子テキスト化することで、データ作成の効率化も目指した。

実証実験の結果、電子カルテシステムへの投入時間削減といった定量的な効果のみならず、「記入ミスの削減に繋がった」「患者と向き合う時間が増えた」という担当者からの定性的な評価もあった。このように、RPA-AI連携ソリューションの導入によって本来の業務に集中して取り組めるようになったこと、ひいてはお客様の新たな価値創造に繋がる結果が得られたことは、DXが単純なコスト削減だけにとどまらず、業務そのものの質的向上にも寄与できる可能性を強く示唆している。

## 5. あとがき

本稿では、DXを実現する自然言語処理アルゴリズムの開発、操作性・利便性を向上させるためのGUIツールの開発、および社内外への展開と今後の展望について解説した。複数種類の手法やアルゴリズムをサポートした多様な自然言語処理AIを開発することによって、利用者は、用途や要件、実行環境などに合わせた適切なAIの選択が可能になった。

並行して、運用部門からの機能要望を盛り込んだGUIツールを開発することで、現場でのAI運用（学習、性能評価、精度改善）の負担軽減を目指した。さらに、開発した自然言語処理AIを導入・活用することで、稼働を定量的に削減できるだけでなく、RPAツールとの親和性の高さや、業務フロー改善に伴って作業の質が向上するといった効果も示せた。

今後は、新規アルゴリズムの追加などバックエンドの機能拡張と並行して、社内外での活用事例創出や利用組織拡大に向けてフロントエンドのGUIツールを訴求していきたい。

### 文 献

- [1] J. Devlin, M.-W. Chang, K. L. and K. Toutanova : “BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding,” Proc. of NAACL, May 2019.
- [2] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter and R. Vollgraf : “FLAIR : An easy-to-use framework for state-of-the-art NLP,” Proc. of NAACL, Jun. 2019.
- [3] ウィキペディアホームページ。  
<https://ja.wikipedia.org/wiki/メインページ>
- [4] 森羅プロジェクト：“拡張固有表現。”  
<http://ene-project.info>

# 要約作業を効率化する多機能な ニュース記事自動要約AIシステム

ドコモ北京研究所

ぐお  
郭しゅーほん  
塚宏なかむら  
中村いっせい  
一成り  
李あんしん  
安新

サービスイノベーション部

ふじもと  
藤本ひろし  
拓

現在、多くのメディアのニュースでは、記事の要約を表示することにより読者に内容を分かりやすく伝えている。しかし、人手による要約作業は分量が多い上に専門的なスキルを必要とし、人材確保の困難さの観点で課題となっている。そこでドコモは、要約作業を効率化できるよう、ユーザの意向に沿った要約を自動で作成する多機能な自動要約AIシステムを開発した。これにより、要約作業時間を短縮し、人手不足を解消することが可能となる。

## 1. まえがき

現在、多くのメディアのニュースでは、記事原稿の主要な内容を端的にまとめた要約を表示することにより、読者に内容を分かりやすく伝えている。しかし要約には、配信するメディアによって字数制限が存在することが多く、人手で必要な情報量を適切な長さに要約する作業は、分量が多い上に専門的なスキルを必要とする。そのため、スタッフの育成に一定期間の研修が必要な場合もあり、人材確保の観点で課題となっている。

このような課題を解決するために、近年、要約作業を効率化できるAIを用いたさまざまな自動要約システムがリリースされている。これらのシステムの多くは、原文から主要な文を抽出する抽出式要約や、原文から一部の単語やフレーズを削除、または追加することで新たな文を作成する生成式要約を提供している。抽出式要約と生成式要約の例を図1に示す。

生成式要約については、生成する要約の長さに合わせて一部の単語やフレーズを削除、または生成するため、既存技術では文法に誤りが生じることや、

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。



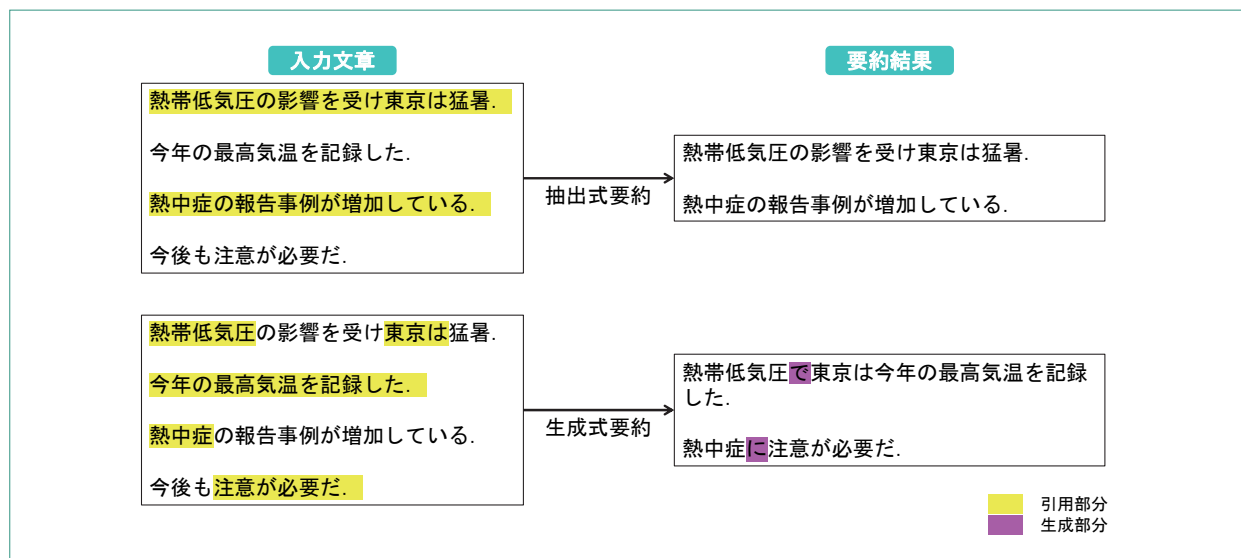


図1 抽出式要約と生成式要約の例

文法は正しい一方で指定の長さから逸脱してしまうことが課題である。また、既存の要約システムの多くは、ユーザの含めたいキーワードなどを考慮する機能や、要約内容が入力文章のどの位置に由来するかを特定し確認する機能がないため、実際の運用時では、要約内容に対して事前に制約を加えることができない点や、要約後に文法などに誤りがある場合に効率的に人手で修正することができない点も課題である。

そこでドコモは、上記の課題を解決できるユーザの意向に沿った要約を作成する多機能な自動要約AIシステムを開発した。本システムは特に3つの特長をもつ。1つ目は、要約の文字数制御機能についてユーザが指定した文字数の70~100%の長さの要約を作成できるようチューニングし、既存技術と比較して文字数制御の品質と要約の文法などの品質を大幅に向上させた点である。2つ目は、ユーザが目的の要約を作成しやすいように、要約に含める、もしくは含めないキーワードが設定できるヒント機能や、タイトルの内容を含む要約が生成できるタイトル機能を具備している点である。3つ目は、要約の

結果が正しいか確認、または修正といったユーザの作業を効率化できるよう、原文の抽出箇所と要約後の文章の生成箇所を可視化することができる位置特定機能を具備している点である。

本システムを利用することで、文字数や内容の観点でユーザの求める最適な要約を自動で生成し、生成後に人手による修正が必要な場合でも効率的に修正作業ができるため、人手による要約作成や従来の要約システムを利用した場合と比較して要約作業に要する時間を短縮し、人手不足を解消することが可能となる。

本稿では、ドコモの自動要約AIシステムの各機能の内容、生成式要約の技術的な性能改善方法、要約性能について解説する。

## 2. ドコモの自動要約AIシステム

### 2.1 システム全体像

ドコモの自動要約AIシステムには、深層学習<sup>\*1</sup>を用いた抽出式要約と生成式要約の2つのシステムがある(図2)。各システムに実装されている機能

\*1 深層学習：多層のニューラルネットワークを用いた機械学習の一種。

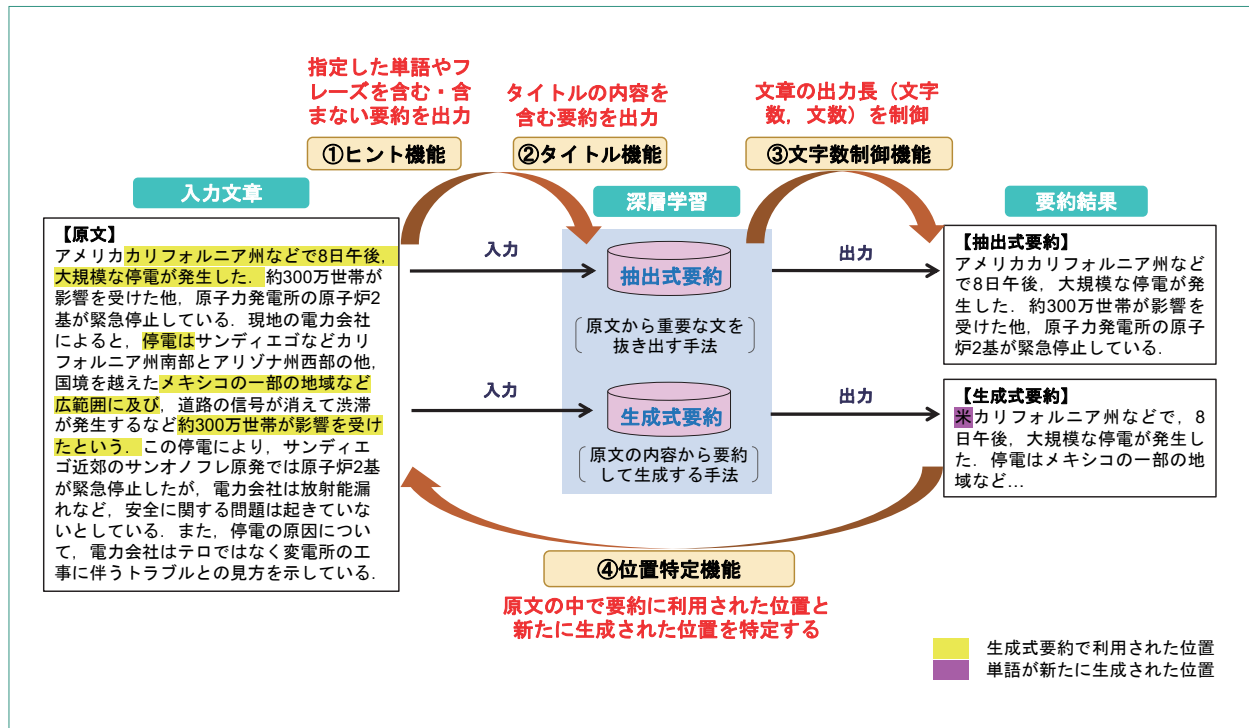


図2 ドコモの自動要約AIシステムの全体像

を以下に示す。

- ① ヒントとなるキーワードやフレーズを指定して要約内容に制約を加えるヒント機能
- ② タイトルの内容を含む要約を生成するタイトル機能
- ③ 要約の文字数や文の数を制御する文字数制御機能
- ④ 原文上の引用位置や新たに生成された文字の位置を可視化する位置特定機能

## 2.2 ヒント機能

ヒント機能は、ユーザが要約に含めたい、もしくは含めたくないキーワードやフレーズを指定することで、要約内容に制約を加えることができる機能である。ニュース記事において、登場人物、地域名、企業名、などさまざまな情報の中からどの情報を要約に含めたいかはユーザごとに異なる可能性があるため、各ユーザに対して最適な要約を生成できる機

能を開発した。

本システムでは、複数のキーワードやフレーズを入力することが可能である。図3は要約に含めたいキーワードをユーザが指定することで、ヒントが無い場合には要約に含まれない内容に含まれるように変化させた例である。図4は要約に含めたくないキーワード（図中ではマイナスヒントと記載）をユーザが指定することで、ヒントがない場合に要約に含まれていた内容が、含まれないように変化した例である。このようにユーザが入力するヒントによって、ユーザの意向により近い要約を出力することが可能である。

## 2.3 タイトル機能

タイトル機能は、ユーザが原文のタイトルを提供することにより、要約の質を向上させることができる機能である。通常、深層学習による要約の場合、

原文	指定文字数	ヒント指定なし 要約結果
<p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。約300万世帯が影響を受けた他、原子力発電所の原子炉2基が緊急停止している。現地の電力会社によると、停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他、国境を越えたメキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は放射能漏れなど、安全に関する問題は起きていないとしている。また、停電の原因について、電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している。(294文字)</p>	120	<p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。停電はメキシコの一部の地域など広範囲に及び、約300万世帯が影響を受けたという。また、停電の原因について、電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している。(118文字)</p>
		<p>ヒント指定“緊急停止” 要約結果</p> <p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は放射能漏れなど、安全に関する問題は起きていないとしている。(107文字)</p>

■ ヒント

図3 ヒント機能を利用したヒントを含む生成式要約結果の例

原文	指定文字数	マイナスヒント指定なし 要約結果
<p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。約300万世帯が影響を受けた他、原子力発電所の原子炉2基が緊急停止している。現地の電力会社によると、停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他、国境を越えたメキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は放射能漏れなど、安全に関する問題は起きていないとしている。また、停電の原因について、電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している。(294文字)</p>	100	<p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。停電はメキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという。(90文字)</p>
		<p>マイナスヒント指定“渋滞” 要約結果</p> <p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は、安全に関する問題は起きていないとしている。(100文字)</p>

■ マイナスヒント

図4 ヒント機能を利用したマイナスヒントを含まない生成式要約結果の例

要約の品質は学習データの量と質に依存し、学習データに含まれないような内容の文章が入力された場合に、要約の品質が低下することがある。そこで、タイトルに含まれるキーワードは原文の中でも特に重要な内容を抽出したものであると考え、タイトルをヒントとして与え、要約の質を高めるタイトル機能を開発した。図5に示すように、タイトルを指定することで、タイトル内のキーワードが含まれる要

約を生成することが可能である。

## 2.4 文字数制御機能

文字数制御機能は、ユーザが要約の文字数を指定できる機能である。ニュースサイトやソーシャルメディアには、配信メディアの要約表示枠のサイズの制約から要約文字数の最大値が厳しく制限されることが多いため、本システムではユーザが設定した文

字数の70～100%の範囲で要約を出力するようにチューニングをしている。図6に示すように同じ原文に対して異なる文字数を設定すると、それに合った文字数の要約をそれぞれ得ることができる。この機能を用いることで、さまざまな長さの要約を生成

することが可能である。なお、文字数だけでなく文の数を指定することも可能である。

## 2.5 位置特定機能

位置特定機能は、要約文言の原文上の引用位置や

入 力	指定文字数	タイトル指定なし 要約結果
<p>【タイトル】 米国内などで<b>大規模停電</b>、<b>原子炉2基緊急停止</b></p> <p>【原文】 アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。約300万世帯が影響を受けた他、原子力発電所の原子炉2基が緊急停止している。現地の電力会社によると、停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他、国境を越えたメキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は放射能漏れなど、安全に関する問題は起きていないとしている。また、停電の原因について、電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している。(294文字)</p>	100	<p>アメリカカリフォルニア州などで8日午後、<b>大規模な停電</b>が発生した。停電はメキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという。(90文字)</p>
		タイトル指定あり 要約結果
		<p>アメリカカリフォルニア州などで8日午後、<b>大規模な停電</b>が発生した。この停電により、サンディエゴ近郊のサンオノフレ原発では<b>原子炉2基が緊急停止</b>したが、電力会社は、安全に関する問題は起きていないとしている。(100文字)</p>

■ ■ タイトル関連キーワード

図5 タイトル機能を利用した生成式要約結果の例

原 文	指定文字数	要約結果
<p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。約300万世帯が影響を受けた他、原子力発電所の原子炉2基が緊急停止している。現地の電力会社によると、停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他、国境を越えた<b>メキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けた</b>という。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は放射能漏れなど、安全に関する問題は起きていないとしている。また、停電の原因について、電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している。(294文字)</p>	80	<p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。停電はメキシコの一部の地域など広範囲に及び、約300万世帯が影響を受けたという。(70文字)</p>
<p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。約300万世帯が影響を受けた他、原子力発電所の原子炉2基が緊急停止している。現地の電力会社によると、停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他、国境を越えた<b>メキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けた</b>という。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は放射能漏れなど、安全に関する問題は起きていないとしている。また、<b>停電の原因について、電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している。</b>(294文字)</p>	160	<p>アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。停電はメキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという。また、停電の原因について、電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している。(136文字)</p>

■ 引用部分

図6 文字数制御機能を利用した生成式要約結果の例

新たに生成された文字の位置を可視化する機能である。既存の自動要約AIシステムでは、出力する要約が原文のどこに位置するか表示されないため、ユーザが原文と要約結果を比較して、要約が原文の中の重要な内容を含み、文法的に正しいかどうかを確認する作業に時間を要する。そこで、ユーザが効率的に原文と要約結果を比較できるように、原文の中で要約として引用された文言の位置と、要約の中で新たに生成された文言の位置を特定できる位置特定機能を開発した。図7に示すように、「原子炉2基が緊急停止」という同じフレーズが原文に複数ある場合や、単語が新たに生成された場合にも正しく参照箇所をマッピングすることが可能である。

### 3. 生成式要約の品質改善手法

#### 3.1 概要

通常、深層学習を用いた要約モデルの学習では、原文と人手で作成した正しい要約のペアのデータを用いて学習を行うが、人手で要約を作成するのは時間的、また金銭的なコストがかかるため、大量のデータが得られず性能の向上が図れないことが課題である。

そこで、ドコモの自動要約AIシステムでは、大

量のデータを用意する代わりに、原文と要約のペアデータから機械的に生成された誤りのある要約や、正解の要約がない大量の原文データ、一文の中から不要な情報が除かれた圧縮文を活用して要約モデルの性能向上を行った。具体的には、図8と以下に示すように、主に文法、非冗長性、流暢性、文字数制御の観点で、独自の技術を導入した。その結果、既存技術と比較して高い性能を達成した。

- ①文法誤りのある文を機械的に生成し、文法誤りを低減する強化学習を導入。
- ②冗長な内容が含まれる要約を機械的に生成し、冗長性を低減する対照学習\*2を導入。
- ③大量の原文データを用いて流暢性を改善する事前学習を実施。
- ④文字数制御のために圧縮文を用いて学習した文圧縮モデルを導入。

#### 3.2 文法

文法に関しては、特に学習データセットに含まれないような単語の組合せのパターンなどを要約する場合に誤りが起こりやすい。その原因として、文章を生成する深層学習のモデルを学習する際に広く使われている方法である「Teacher forcing」アルゴリズム [1] が挙げられる。深層学習を用いたテキ

原文	要約結果
アメリカカリフォルニア州などで8日午後、大規模な停電が発生した。約300万世帯が影響を受けた他、原子力発電所の原子炉2基が緊急停止している。現地の電力会社によると、停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他、国境を越えたメキシコの一部の地域など広範囲に及び、道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は放射能漏れなど、安全に関する問題は起きていないとしている。また、停電の原因について、電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している。	*カリフォルニア州などで8日午後、大規模な停電が発生した。この停電により、サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが、電力会社は放射能漏れなど、安全に関する問題は起きていないとしている。

生成式要約で引用された位置  
 単語が新たに生成された位置

図7 要約の引用位置や生成位置の可視化例

\*2 対照学習：類似データ間の特徴距離が非類似データ間の特徴距離よりも近くなるように学習することでモデルの高精度化を行う手法。

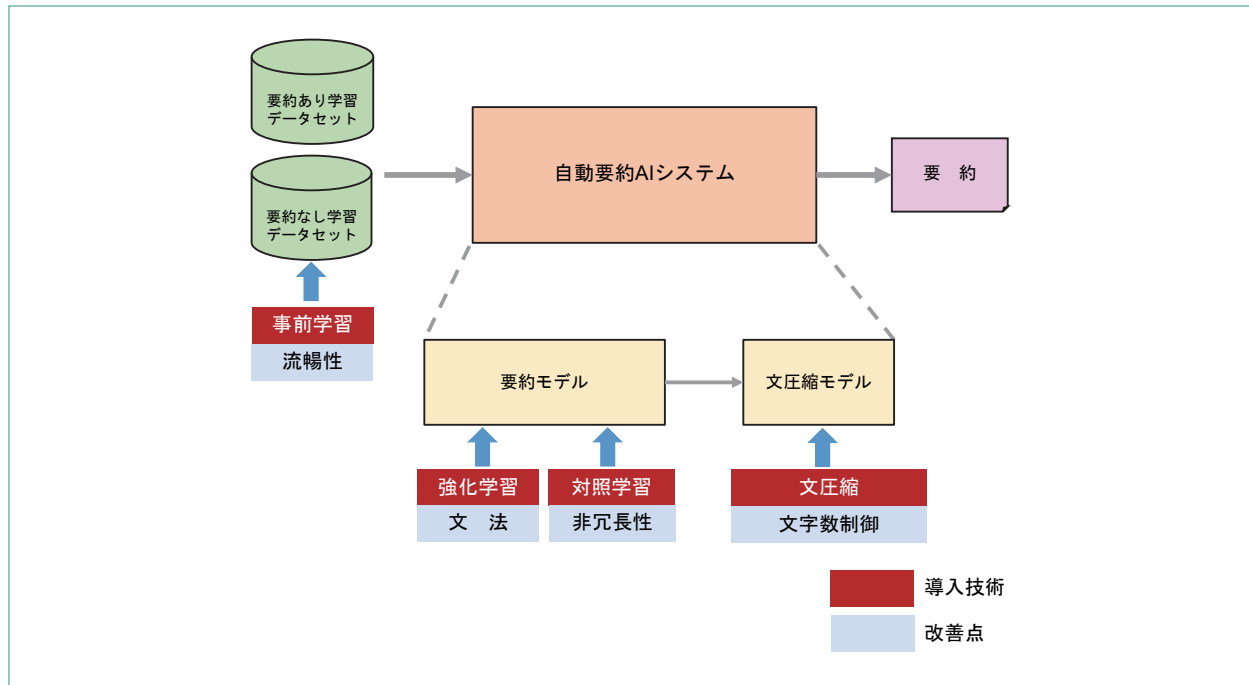


図8 要約システムの改善点

スト生成では、直前に生成した単語を用いて次の単語を生成するが、学習の初期段階では不正解の単語の生成が多いため、Teacher forcingを用いる場合、直前に生成した単語ではなく正解の単語を深層学習のDecoder<sup>\*3</sup>側に入力することで、学習の効率化が行われる。このような学習では、特定の単語に対して次に生成される単語を学習するため、文全体ではなく単語レベルでの局所的な学習が行われる。

課題は、学習時は正解の単語が与えられるにも関わらず、実際に未知の文章を用いてテキスト生成を行う際には、正解の単語ではなく直前の深層学習によって生成された単語を用いてテキストが生成されるため、一度誤りとなる単語が生成されると、その単語の次に生成されるテキストが意味的にも文法的にも破綻しやすいことである。

そこで、ドコモの自動要約AIシステムでは、強化学習の方法を取り入れた。強化学習はエージェントとそれに報酬を返す環境の間のフィードバックを

通して、エージェントを学習させる手法である。一般的にロボット制御などの分野で利用される学習手法であるが、近年自然言語処理の分野でも利用される例がある [2]。本システムでは、強化学習の各要素が下記のように構成される。

- ・ エージェント (Agent)：要約モデル。
- ・ 環境 (Environment)：文を与えたときに文法的に正しいか正しくないかを判断する識別器。
- ・ 状態 (State)：要約結果。
- ・ 動作 (Action)：次の単語の生成。
- ・ 報酬 (Reward)：文法の正しさのスコア (識別器の判断結果)。

要約モデルが要約を生成する行動を通じて、報酬が最も多くなるように、つまり文法の誤りが少なくなるように学習を行うことで、既存手法と比較して文法誤りの少ない要約モデルを構築した。

<sup>\*3</sup> Decoder：テキストを入力して新たなテキストを出力する多層のニューラルネットワークのうち、テキストを入力に受け取り特徴量化する部分をEncoder、Encoderの出力する特徴を用いて新たなテキストを生成する部分をDecoderと呼ぶ。

### 3.3 非冗長性

本稿では、要約内容の中で、同一文内や複数の文間で意味的に同一の内容が繰り返し言及されることを「冗長」と表現する。深層学習を用いたテキスト生成モデルでは、冗長な内容を含む要約の生成が課題として指摘されている [3]。

そこで、ドコモの自動要約AIシステムでは、対照学習を導入することにより本課題の改善を行った。対照学習では、Anchor、正例、負例の3つを用意し、学習時にAnchorと正例の特徴間の距離が、Anchorと負例の特徴間の距離よりも近接するように学習することでモデルの精度を向上する手法である。ドコモの自動要約AIシステムでは、Anchorは要約モデルによって生成された要約、正例は人手で作成した正しい要約であり、負例は正例に対して機械的に同一の単語、フレーズ、文を繰り返した、冗長性の観点で誤りのある要約である。このように冗長性の観点で誤りのある負例を用いた学習を行うことで、既存手法と比較して冗長な文を生成しにくい要約モデルを構築した。

### 3.4 流暢性

本稿では、単語やフレーズ間の組合せが正しいことを「流暢」と表現する。実際のシステム利用時に、学習データセットの中で見られない単語の組合せを含む原文が入力される場合がある。このとき既存手法を用いると、文法的に正しくても使い方の誤った単語の組合せで要約が出力されることがある。

このような課題を解決するために、ドコモの自動要約AIシステムでは、大量の文章を用いて単語の組合せに関する事前学習を行い、要約モデルの流暢性の向上を図った。

### 3.5 文字数制御

文字数制御で最も一般的な手法は、要約を生成する過程でBeam Search<sup>\*4</sup>を行う際に、ルールベース

でより適切な長さとなる出力結果を選択する手法である。このようなルールベースの手法の場合、適切な長さとなる出力結果を選択するために、文法的な正しさと主題の抽出度合いについて考慮されない場合があり、要約結果がユーザの指定した長さに近い一方で、文法的に誤りのある要約や主題から逸脱した要約が出力されやすくなる [4]。

そこでドコモの文字数制御機能は、要約モデル自体に出力する要約の長さの情報を特徴量<sup>\*5</sup>として入力し、要約モデルが要約を出力する際に文法、主題、長さなどを同時に考慮するよう学習させることで、要約モデルの最適化を行った。また要約モデルの後処理として、文字数を削減するための文圧縮モデルを導入することにより、要約モデルより出力された文がユーザの設定する文字数を上回る場合は、文を適切な長さに圧縮し要約を生成できるように改善した。

## 4. 生成式要約の性能評価

### 4.1 データセット

日本テレビ放送網株式会社が提供する約18万件のニュース記事とそれらを人手で要約したデータセットを利用し、本システムの生成式要約の学習と評価を行った。

### 4.2 ROUGEによる評価

正解要約 (Reference) に対する要約結果の網羅性を評価する Recall について各 ROUGE (Recall-Oriented Understudy for Gisting Evaluation) の計算方法を図9に示す。ROUGEは、正解とされるテキストとモデルが作成したテキストの類似度を比較する要約の指標の中で最も広く利用されるものである [5]。ROUGE-1、ROUGE-2は、それぞれテキスト間の unigram<sup>\*6</sup>と bigram<sup>\*7</sup>の重なり度合いを表し、ROUGE-Lは、一致する最長の単語の長さを用いてテキスト間の重なり度合いを測定する。ROUGE-1、

<sup>\*4</sup> Beam Search: 本稿では、ニューラルネットワークが出力する単語の候補をスコアに基づいて複数選定し、いくつかの要約結果の候補を得ることを指す。

<sup>\*5</sup> 特徴量: データから抽出される、そのデータの特徴付ける量 (数値) のこと。

<sup>\*6</sup> unigram:  $n$  単語連続して続く文字列を  $n$ -gram と呼び、 $n$  が 1 の場合の 1 単語だけの文字列を指す。

<sup>\*7</sup> bigram: 2 単語連続して続く文字列を bigram と呼ぶ。

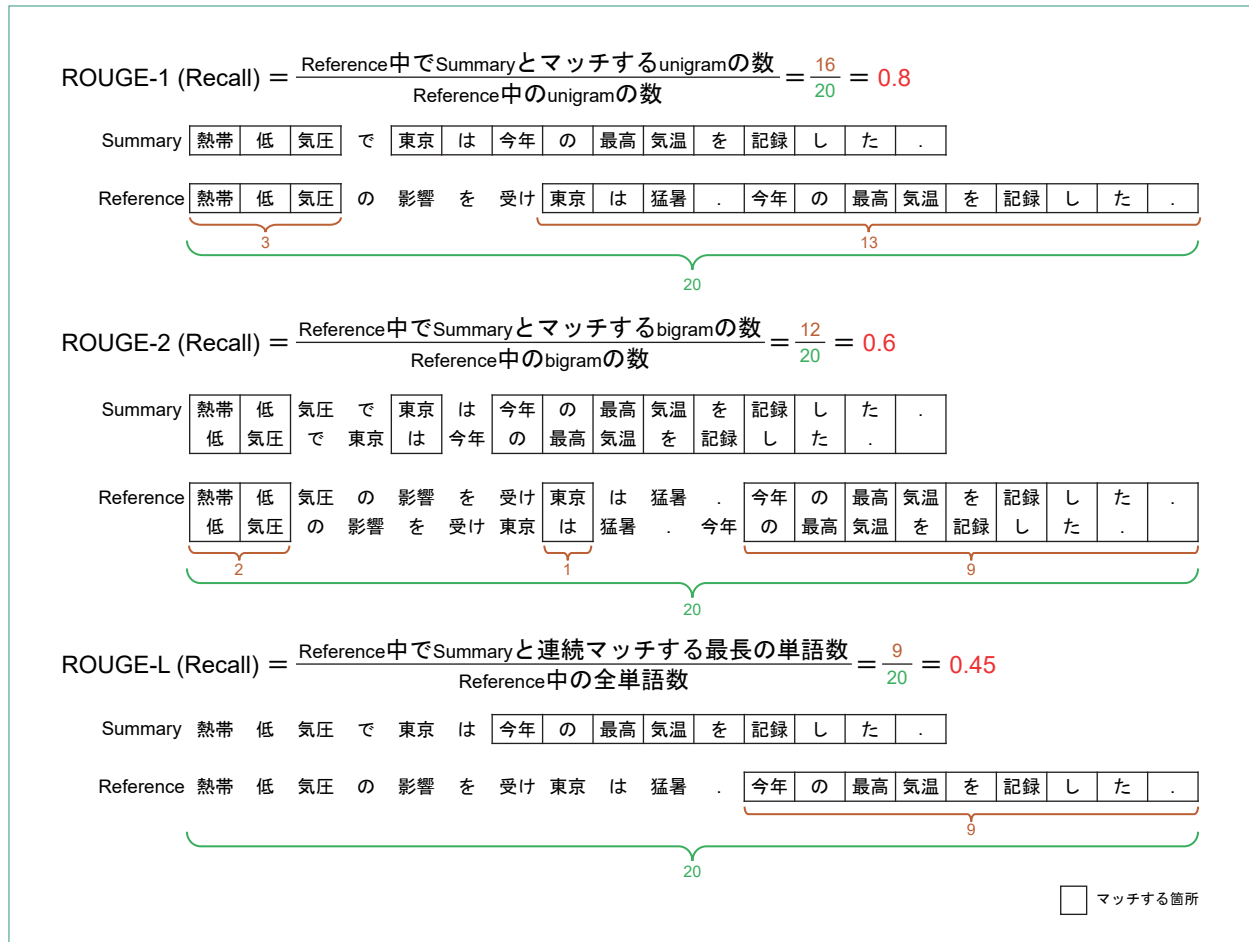


図9 各ROUGEの計算方法

ROUGE-2, ROUGE-Lすべてにおいて、値が大きいほどテキスト間の重なりが大きく、テキスト生成モデルの性能が高いことを示す。

表1は、日本テレビ放送網株式会社が提供するデータセットのうち3,000件を用いて、本システムの評価を行った結果である。比較対象として原文の先頭3つの文を要約結果とするLead-3、また、TextRank [6]、SumBasic [7]、LSA (Latent Semantic Analysis) [8]、Submodular [9]、PGN (Pointer Generator Network) [3] を利用した。TextRank、SumBasic、LSA、Submodularについては抽出式要約であり、抽出する文の中で重要度の高い3つの文

を要約結果とした。PGNについては生成式要約であり、Beam Search結果の中で最も正解の要約の文字数に近い要約を、最終的な要約結果として評価した。これらの既存手法と比較して、ドコモの自動要約AIシステムはROUGEの数値が大きく、人手で作成した正解に、より近い要約を生成することが示された。

### 4.3 人手による評価

前述のROUGEの評価による欠点として、文法の誤りや意味的な冗長性などを評価できない点が挙げられる。そこで、日本語を母国語とする者が以下の



4つの観点で評価を行った。その際スコアを、4を最高点とする4段階評価とした。

- ①文法：要約に文法的な誤りが少ないこと。
- ②主題性：生成された要約が原文の主要な内容をカバーしていること。
- ③非冗長性：要約の中に意味的に同一の単語、フレーズ、文の繰返しが無いこと。
- ④流暢性：生成された要約が単語間、文間で流暢であること。

表2は、評価データセットのうち100件を用いてドコモの自動要約AIシステムとPGNの出力結果を人手で評価した評価値の平均値である。また表3に示すように、指定した要約の文字数に対して生成された要約の文字数の割合を算出し、4段階のスコアに割り当てた。この要約文字数のスコアを評価データごとに算出し、平均値を表2の「要約文字数」列に記載した。このような人手による評価についても既存手法のPGNと比較して、ドコモの自動要約AI

システムは各評価指標の数値が大きく、より品質の高い要約が生成されることが示された。なお、自動要約の処理時間は、抽出式要約が約1秒、生成式要約が10秒程度のため、数分から数十分かかる手動の要約作業と比較して速度が大幅に向上する。

## 5. あとがき

本稿では、ドコモの自動要約AIシステムの各機能や技術、性能について解説した。ドコモは、ユーザがヒントを指定するヒント機能、タイトルを使用するタイトル機能、要約の文字数や文の数を制御する文字数制御機能、原文上の引用位置や新たに生成された文字の位置を可視化する位置特定機能を開発し、ユーザの意向に沿った要約を出力しやすいシステムを実現した。また、文法、非冗長性、流暢性、要約文字数の制約の課題を解決するために、強化学習、対照学習、事前学習などを導入し、性能向上を行った。性能評価結果が示すように、ドコモの自動

表1 ROUGEによる評価結果

		ROUGE-1	ROUGE-2	ROUGE-L
抽出式要約	Lead-3	74.46	63.89	72.48
	TextRank	64.06	50.07	60.16
	SumBasic	64.49	49.18	58.38
	LSA	62.28	46.48	56.85
	Submodular	55.41	36.91	47.41
生成式要約	PGN	79.25	70.36	77.45
	本システム	84.49	76.47	81.80

表2 人手による評価結果

	文法	主題性	非冗長性	流暢性	要約文字数
PGN	2.82	2.40	3.89	3.06	3.25
本システム	3.85	3.53	3.92	3.84	3.89

表3 要約の文字数に関するスコア

要約の文字数範囲	スコア
$0.7 \times L \leq S \leq 1.0 \times L$	4
$0.6 \times L \leq S < 0.7 \times L$ or $1.0 \times L < S \leq 1.1 \times L$	3
$0.5 \times L \leq S < 0.6 \times L$ or $1.1 \times L < S \leq 1.2 \times L$	2
$S < 0.5 \times L$ or $1.2 \times L < S$	1

$L$ : 指定した要約の文字数

$S$ : 要約モデルが作成した要約の文字数

要約AIシステムは既存の技術と比較し、ROUGEによる評価と人手による評価の両方の指標の値を大幅に改善した。この自動要約AIシステムにより、要約作業時間を短縮し人手不足を解消することが可能となる。今後は、実サービスを通して得られた課題を基に既存機能の性能向上や新機能の開発を行い、さらに高性能な自動要約AIシステムを実現していく。

## 文献

- [1] R. J. Williams and D. Zipser : "A Learning Algorithm for Continually Running Fully Recurrent Neural Networks," *Neural Comput.*, Vol.1, No.2, pp.270–280, 1989 (doi : 10.1162/neco.1989.1.2.270).
- [2] L. Yu, W. Zhang, J. Wang and Y. Yu : "SeqGAN : Sequence generative adversarial nets with policy gradient," 31st AAAI Conf. Artif. on Intell., pp.2852–2858, 2017.
- [3] A. See, P. J. Liu, and C. D. Manning : "Get To The Point : Summarization with Pointer-Generator Networks," *ACL 2017 - 55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap.)*, Vol.1, pp.1073–1083, 2017 (doi : 10.18653/v1/P17-1099).
- [4] B. Eikema and W. Aziz : "Is MAP Decoding All You Need? The Inadequacy of the Mode in Neural Machine Translation," *Proc. of the 28th International Conference on Computational Linguistics*, pp.4506–4520, 2021 (doi : 10.18653/v1/2020.coling-main.398).
- [5] C.-Y. Lin : "Looking for a Few Good Metrics : ROUGE and its Evaluation," *NTCIR Work.*, pp.1–8, Jun. 2004.
- [6] R. Mihalcea and P. Tarau : "TextRank : Bringing Order into Text," *Proc. of 2004 Conf. Empir. methods Nat. Lang. Process.*, pp. 404–411, 2004.
- [7] L. Vanderwende, H. Suzuki, C. Brockett and A. Nenkova : "Beyond SumBasic : Task-focused summarization with sentence simplification and lexical expansion," *Inf. Process. Manag.*, Vol.43, No.6, pp.1606–1618, NoV. 2007 (doi : 10.1016/j.ipm.2007.01.023).
- [8] M. G. Ozsoy, F. N. Alpaslan and I. Cicekli : "Text summarization using latent semantic analysis," *J. Inf. Sci.*, Vol.37, No.4, pp.405–417, Aug. 2011 (doi : 10.1177/0165551511408848).
- [9] H. Lin and J. Bilmes : "A class of submodular functions for document summarization," *Proc. of 49th Annu. Meet. Assoc. Comput. Linguist. Hum. Lang. Technol.*, Vol.1, pp.510–520, Jun. 2011.

# ユーザ行動の時系列予測モデルを利用したレコメンドエンジンの開発

サービスイノベーション部 いとう伊藤 たく拓 ささき佐々木 ゆうり祐理  
たなか田中 しげき茂樹

近年、WebにおけるBtoCサービスにおいて商品・コンテンツのレコメンド機能の提供が主流となっている。レコメンドエンジンはさまざまな種類があるが、多くは人気ランキングに基づいたものであり、これらはユーザのサービス利用の文脈を正しく理解できていなかった。ユーザの興味を惹くレコメンドを提示するためには、レコメンドエンジンがユーザ文脈を理解し、次に興味をもつコンテンツを予測する必要がある。そこでドコモでは、購買してもらう確率を上げることを目的に、行動を時系列で解釈して予測を行う深層学習アルゴリズムを活用したレコメンドエンジンを開発し、社内サービスに適用した。これにより、高精度なレコメンド提供が可能となった。

## 1. まえがき

近年、EC (Electronic Commerce) サービス<sup>\*1</sup>や動画配信などのBtoC (Business to Consumer) サービスでは、商品のレコメンド<sup>\*2</sup>を行うことが主流となっている。ドコモサービスの多くについても、「ユーザへのおすすめ」としてレコメンド機能が提供されている。レコメンド手法においては、人気ラン

キングを基にした単純な手法から、機械学習<sup>\*3</sup>を用いた手法まで幅広く提案されており、特にここ数年では市中で簡便に利用できる機械学習ライブラリ<sup>\*4</sup>が配布されているほか、クラウドサービス<sup>\*5</sup>を用いたレコメンドサービスが提供され始めている。

レコメンド機能は数年前より提供されているが、既存では、人気ランキング<sup>\*6</sup>をベースとしたレコメンドが多く行われていた。人気ランキングベースは、

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

\*1 ECサービス：ネットを通じて物の売買を行うサービスのこと。

\*2 レコメンド：ユーザに合わせて、商品やコンテンツを推薦すること。

\*3 機械学習：入力されたデータを基にパターンを学習し、何らかのタスクを実行するコンピュータアルゴリズムのこと。

大多数のユーザの嗜好を捉えているため一定のクリック率を稼げるが、異なるユーザ間で同一の商品がレコメンドされるため、人気コンテンツを好まないユーザには響かず、加えて出ているコンテンツが固定化されるため、他の多くの商品はレコメンドされず、ユーザが新しい商品を発見しにくいという課題があった。また、ユーザは通常、一連のWeb回遊の中で欲しい商品をクリックするという手順を踏むため、レコメンドサービスにおいては、趣味嗜好よりも直前にどのようなコンテンツをクリックしたかという時系列のデータが重要になる。例えば、ポータルサイトで食料品を探しているユーザには、普段の趣味嗜好に関わらず、Web回遊をしている場面では食料品が提示されることが望ましい。生鮮食品を直前にクリックしているユーザには、続けて生鮮食品を提示することでユーザ所望のコンテンツを推薦することができるようになる。

そこでドコモでは、時系列予測のアルゴリズムであるRNN (Recurrent Neural Network) をレコメンドに適用し、ユーザの長期・短期での興味の推移を

捉えることを可能とした。本技術では、直近の興味のあるジャンルを主にレコメンドするほか、定常的に興味があるジャンルも共にレコメンドすることで、パーソナライズしつつ状況に合わせたレコメンドの提示が可能となり、クリック率向上に寄与した。本稿では、実際に適用したレコメンドアルゴリズムとサービスでの効果検証について解説する。

## 2. RNNのレコメンドへの活用

### 2.1 RNN概要

RNNは、再帰型ニューラルネットワーク<sup>\*7</sup>と呼ばれ、時系列データ<sup>\*8</sup>などの連続的なデータのパターンを認識するように設計されたニューラルネットワークのモデルである [1]。従来の一般的なニューラルネットワークモデルでは、固定長の入力層、出力層とその間に中間層が与えられる (図1)。例えば画像認識タスクでは、実際の画像を画素単位で分割し、画素値を入力層として与え、出力層では分類対象のカテゴリを出力する。この場合、入力層

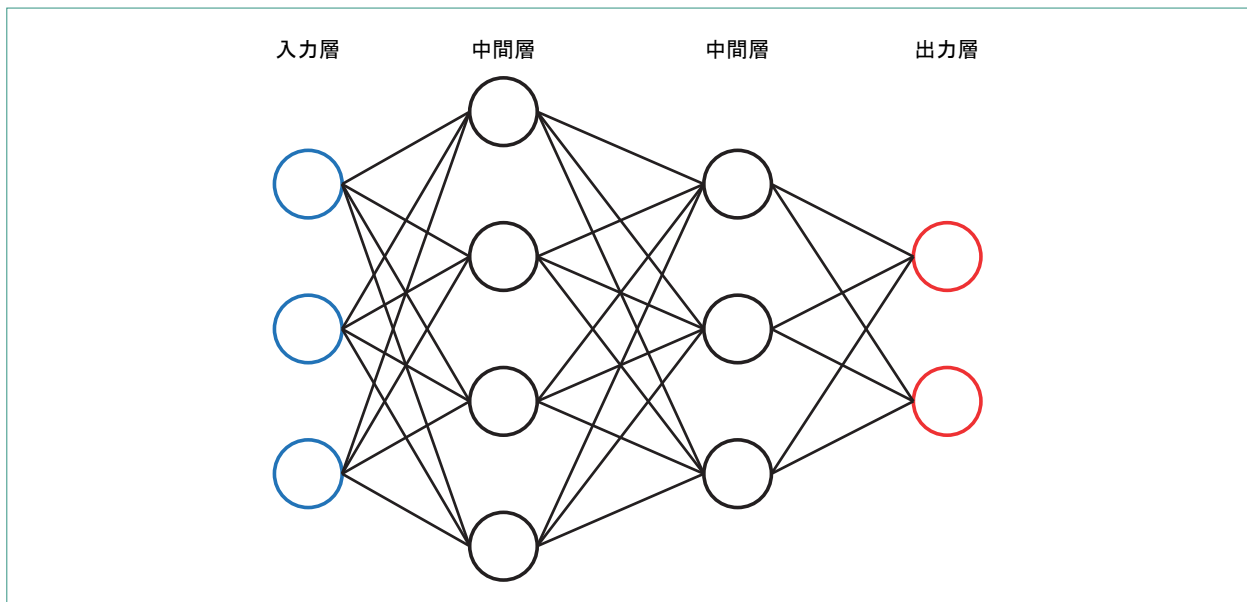


図1 一般的なニューラルネットワーク

\*4 ライブラリ：汎用性の高い複数のプログラムを、再利用可能な形でひとまとまりにしたもの。  
 \*5 クラウドサービス：利用者の手元にあるコンピュータ上ではなく、遠隔のサーバ上で動いているサービスをネットワーク経由で利用者に提供するもの。  
 \*6 ランキング：サービス全体でクリック・購買が多い順に商品・

コンテンツを並べたもの。最も基本的なレコメンドアルゴリズムの1つ。  
 \*7 ニューラルネットワーク：人間の脳内の神経回路網を数式モデルで表したもの。入力層、中間層、出力層から構成される。  
 \*8 時系列データ：値が時間的に変化した情報をもつデータのこと。

は画素数、出力層は分類パターン数となり、それぞれ1つの値となるため、各ノード\*9は固定長として与える必要がある。

一方で、現実のデータは必ずしも固定長で与えられるわけではない。例えば、文書予測などのタスクの場合、入力としては文章が与えられることになるが、画像のような固定長でないため、上記のようなモデルでは表現することができない。そこで、可変長のデータを取り扱えるように改良されたモデルがRNNである。RNNでは、可変長データの各地点を時系列として扱い、内部（中間層）に状態という変数を保持する（図2）。入力データに応じて状態が変化していき、その状態が次の時系列ステップに伝搬していく。そして、状態に応じた値が出力層に提示される。このように、状態が時系列で伝搬していくというモデルとすることで、可変長データを入力とする対象に対してニューラルネットワークを適用することができる。

## 2.2 RNNのレコメンドへの適用

上記RNNは、文章認識\*10や機械翻訳\*11などの文

章をターゲットとするタスク、および音声認識などの音声ターゲットとするタスクに利用されることが多い。与えられた入力値に対してどのような出力値が与えられるべきか、というデータをネットワークに学習させることで、新しい文章・音声に対して所望の出力値を得ることができる。このような特性から、RNNは時系列予測にも用いることができる。入力と出力の値を同一にして学習させることで、特定の入力が与えられたとき、次にどのような値が来るかを予測させるモデルを構築することができる [2]。

近年では、RNNの時系列予測タスクをレコメンドに適用する事例が増えている。1つ当りの入力値を商品クリック・動画視聴といったユーザ行動として定義し、出力値を同様にユーザ行動として与えることで、ある商品・動画に興味をもったユーザが、次にどのような商品・動画に興味をもつのかユーザの興味（状態）の遷移を予測することができる（図3）。加えて、その時のユーザ状態に合わせたレコメンドが可能となるため、最新のユーザ行動をモデルに入力することにより、常に最新化された

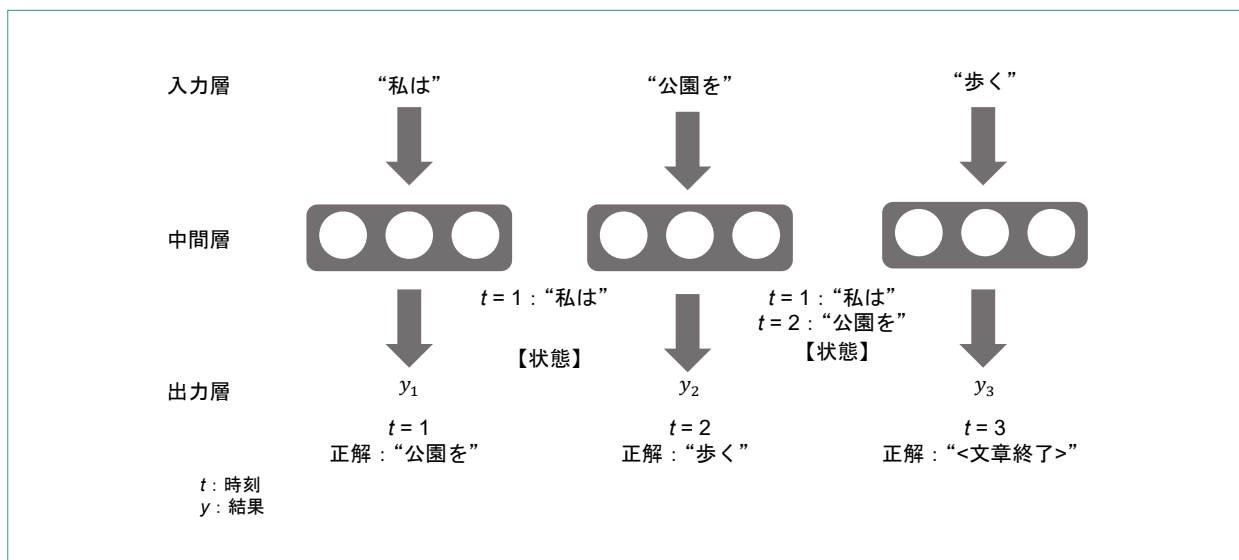


図2 RNNの模式図

\*9 ノード：ネットワークにおいて、入力から受け取った値を伝搬する点のこと。  
 \*10 文章認識：ある文章がどのようなカテゴリーの文章かを判定するタスク。  
 \*11 機械翻訳：ある言語の文章を、別の言語の文章に、コンピュータを用いて自動的に翻訳するタスク。

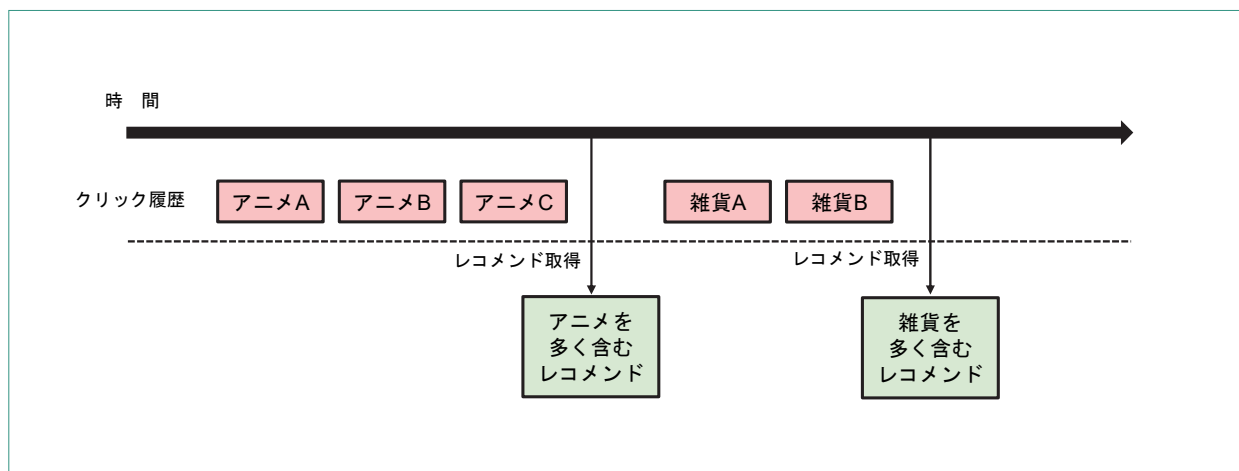


図3 RNNのレコメンド利用概要

レコメンドをリアルタイムに提供することができるようになる。

### 3. 提案手法

#### 3.1 概要

RNNをレコメンドに適用する際、何点かサービスマインに合わせた改良を行い、さらにユーザの直近の興味に適したレコメンドが出力されるような工夫を行っている。具体的には、①階層型RNNを用いた長期予測モデルと短期予測モデルを組み合わせたレコメンド、②異なる情報を用いた複数のRNNを組み合わせたレコメンド、の構築を行った。以下、各モデルについて解説する。

#### 3.2 階層型RNNを用いた長期短期予測モデル

ユーザの興味は一定の周期で移り変わるわけではなく、長期的な興味遷移と短期的な興味遷移があると考えられる [3]。例えば、インストール履歴ベースでユーザにスマートフォンアプリケーション（以下、アプリ）をレコメンドするシステムにおいて、ユーザがX日にSNSアプリを検索しているが、X+1日には動画配信アプリを検索しているという場合

が考えられる。その場合、X日にはSNSアプリをレコメンドし、X+1日には動画配信アプリをレコメンドすることが望ましい。しかし、単純なRNNではX日とX+1日の違いを考慮することができないため、X日の履歴を基にX+1日にもSNSアプリをレコメンドしてしまう可能性がある。そのため、X日の中では当日のインストール履歴を基にSNSアプリをレコメンドし、日をまたいだ場合はより長期的な興味遷移を予測してレコメンドする必要がある。なお、長期的な興味遷移とは、SNSアプリを検索した次の日には動画配信アプリを検索するなどの例を指している。

上記の長期・短期でのユーザ興味を捉えたレコメンドを実現するために、階層型RNNを提案する (図4)。前日までの履歴からレコメンドする内容を決定する日別RNNと、前日と当日の履歴を基に決定する同日RNNの二階層に分けることで、前述の課題を解決することができる。

#### 3.3 異なる情報を用いた複数RNNを組み合わせたレコメンドモデル

商品・動画をレコメンドする際には、それらの内容を踏まえて考える必要がある。一方で、RNNで

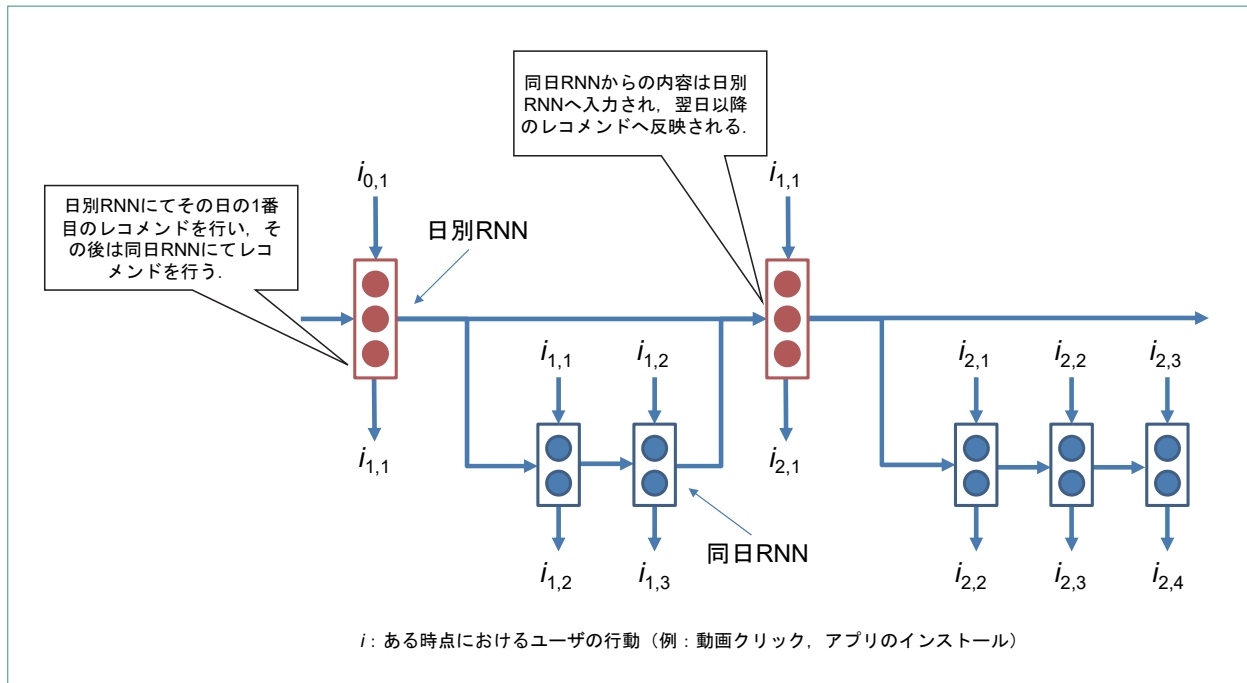


図4 階層型RNNの模式図

扱うデータは数値であり、商品や動画を数値データとして扱う必要がある。一般的に、商品や動画にはコンテンツIDが割り振られており、RNNではこのIDという数値データの時系列遷移を予測する。この手法では、コンテンツID単位での予測となるが、同一のコンテンツが過去の履歴に存在していない場合、予測を行うことができない。しかし、コンテンツIDが異なる場合でも、類似のコンテンツであれば同じような時系列推移になると考えられるため、商品や動画の内容（カテゴリ）の考慮を行う必要がある。例えば、飲食カテゴリの商品をクリックしたユーザは続けて飲食カテゴリの商品をクリックする可能性が高いと考えられるため、コンテンツID単位での時系列だけでなく、カテゴリ単位の時系列も併せて予測することが望ましい。

上記のレコメンドを実現するために、複数RNNを組み合わせたレコメンドモデルを提案する（図5）。異なる階層のデータを独立して時系列予測するRNN

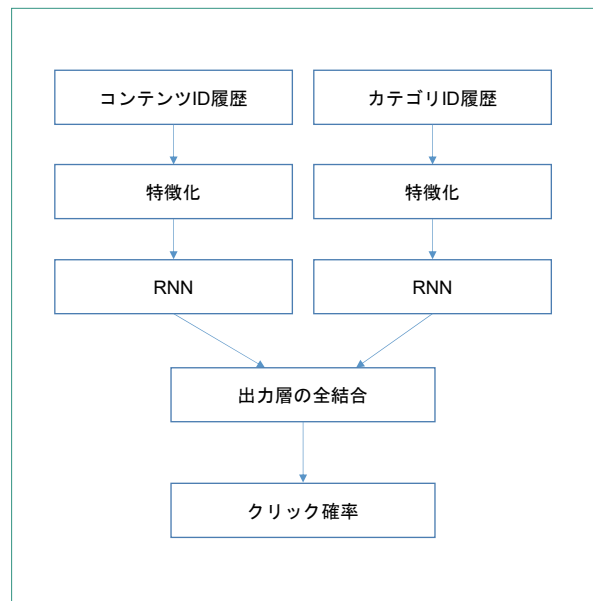


図5 複数RNNの模式図

を組み合わせることにより、コンテンツ単位での細かなレコメンドと、カテゴリ単位での大まかなレコ

レコメンドを両方実現できることになり、精度とカバレッジ\*12の向上を両方実現できるようになる。

## 4. ドコモサービスへの適用

上記技術を、実際のドコモサービスに適用して効果検証を行った。適用したサービスは、ユーザに対してアプリをレコメンドする「おすすめアプリ」と、ドコモが運営する、BtoCサービスの商品をレコメンドするポータルサイト「dマーケット」の2つである。おすすめアプリには「階層型RNNを用いたレコメンド」、dマーケットには「複数RNNを用いたレコメンド」をそれぞれ適用している。それぞれのサービス概要と、効果検証結果について解説する。

### 4.1 おすすめアプリ

サービス画面を図6に示す。ユーザのアプリインストール履歴を基に、ユーザにとって有用と思われるアプリを提示するサービスである。提示したアプリをより多くインストールしてもらうことを目的と

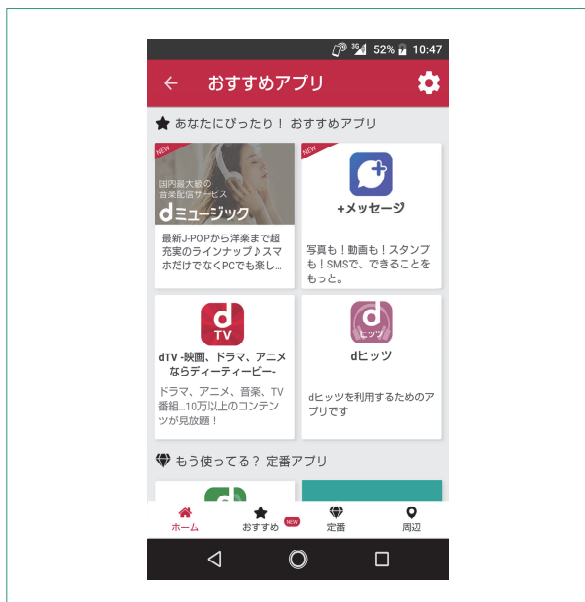


図6 おすすめアプリのサービス画面

し、前述の階層型RNNを用いたレコメンドの効果検証を、実際にユーザに対して行った。2020年9月より本技術と、履歴によるユーザ特徴量\*13を用いた既存のレコメンドアルゴリズムとのA/Bテスト\*14を実施し、既存のアルゴリズムと比較してインストール率\*15 2.7倍を達成した。高いインストール率が出た理由としては、時系列を考慮することにより、例えばTwitterとカメラアプリを連続でインストールしたユーザにはInstagramをレコメンドするなど、直前のユーザの行動からユーザが求めるアプリの特性を推論することができたためであると考えられる。

### 4.2 dマーケット

サービス画面を図7に示す。図に示すとおり、「お買い物」「グルメ」など16種類のタブで分類された



図7 dマーケットのサービス画面

\*12 カバレッジ：全ユーザにレコメンドされる延べコンテンツ数が、全体のコンテンツ数に占める割合。カバレッジが高いということは、幅広いコンテンツがレコメンドされていることを示す。  
\*13 特徴量：データから抽出される、そのデータの特徴付けする量(数値)のこと。

\*14 A/Bテスト：2つのアルゴリズムについて、どちらの方が効果を発揮するか比較するためのテストのこと。  
\*15 インストール率：レコメンド機能が提示したアプリの中で、実際にユーザがインストールに至ったアプリの割合。



領域において、ユーザへのおすすめコンテンツをサービス横断で表示するサービスである。複数のサービスにまたがって、ユーザがコンテンツに興味をもち、クリックしてもらうことを目的としてレコメンドを表示しており、また、Web回遊の中で、常にユーザに興味をもってもらえるよう、表示するたびにレコメンド内容を変化させている。前述の複数RNNを組み合わせたレコメンドの効果検証を実ユーザに対して実施した。本技術と既存の人気ランキングベースでのレコメンドアルゴリズムとのA/Bテストを行い、2.0%のCTR (Click Through Rate)<sup>\*16</sup>向上を達成した。CTRが向上した理由としては、カテゴリIDを考慮したことにより、異なるサービス間のコンテンツの共通要素を抽出できたためであると考えられる。例えば、グルメタブの人気ランキングでは食料品が多くレコメンドされるが、dマガジン<sup>\*17</sup>を多く利用しているユーザにはグルメ雑誌を多くレコメンドするなど、同じグルメであってもそのユーザに適したサービスのコンテンツをレコメンドすることができていた。

## 5. あとがき

本稿では、RNNを拡張したレコメンドアルゴリ

ズムについて解説した。サービスドメインの特性に合わせ、長期・短期双方でのユーザ興味推移をとらえるための階層型RNNを用いたレコメンドと、コンテンツの複数の特徴を活用した複数RNNを組み合わせたレコメンドについて述べた。両方のアルゴリズムについて実際のドコモサービス上で効果検証を行い、既存のアルゴリズムと比べての精度向上に繋がることを確認した。ドコモは、レコメンドアルゴリズムについて、今後も最新技術を追求し、サービスの提供価値向上に向けた取組みを進めていきたい。

### 文献

- [1] J. Li, L. Deng, R. H. Umbach and Y. Gong : "Robust Automatic Speech Recognition : A Bridge to Practical Applications," Academic Press, Oct. 2015.
- [2] D. Wierstra, J. Schmidhuber, and F. Gomez : "Evolino : Hybrid Neuroevolution/Optimal Linear Search for Sequence Learning," Proc. of the 19th International Joint Conference on Artificial Intelligence (IJCAI), Edinburgh, PP.853-858, Jul. 2005.
- [3] B. Hidasi, A. Karatzoglou, L. Baltrunas and D. Tikk : "Session-based recommendations with recurrent neural networks," arXiv preprint arXiv : 1511.06939, Nov. 2015.

<sup>\*16</sup> CTR : レコメンド機能がコンテンツを表示した回数に対して、ユーザが実際にコンテンツをクリックした回数の割合。

<sup>\*17</sup> dマガジン : ドコモが提供するサブスクリプション型の雑誌配信サービスのこと。

# ドコモ画像認識プラットフォームを用いた画像認識AI提供

サービスイノベーション部 さかい としき 酒井 俊樹 いわた もとき 岩田 基暉

近年、AI技術が普及し、画像認識の分野においてAIを用いた、人の作業の代替、自動化、省人化が図られているが、その導入には、高性能なサーバ環境の準備、ソフトウェアの導入といったAI開発環境の整備が必要であった。そこでドコモでは、画像認識AI開発用のデータを準備すれば、さまざまな種類の画像認識AIの開発、デプロイが可能となる「ドコモ画像認識プラットフォーム」を開発した。これにより、AIの利用企業は自前のAI開発・運用環境を用意しなくても、画像認識AI技術を導入することが可能となった。

## 1. まえがき

近年、Deep Learning<sup>\*1</sup>を用いたAI技術の発展、普及は目覚ましく、特に画像認識AI<sup>\*2</sup>の産業応用が進んでいる。例えば、工場における検品の補助、防犯カメラ映像での人や車両の検知、医療における画像診断の補助など、今までは人が行っていた作業の補助、代替として活用されている。しかし、画像認識AIを用いて、既存の作業の効率化、自動化、DX (Digital Transformation)<sup>\*3</sup>を行う際、各企業が解決したい課題をそのまま解決できる画像認識AIが提供されていることは少ない。このため各企業が、保持している画像／動画データと、当該デー

タに対して画像認識AIにどのような認識処理を行わせたいか、その答えを記録したアノテーションデータを準備した上で、企業独自の画像認識AIを開発する「学習」と、学習した画像認識AIを使える状態にする「デプロイ<sup>\*4</sup>」の作業が必要となる。デプロイにおいては、作成した画像認識AIをサーバ上に展開し、画像の入力を受け付け、結果をテキストなどで返却するAPI (Application Programming Interface)<sup>\*5</sup>を構築する。しかし、学習とデプロイをAIの利用企業が自身で行うには、Deep Learningを扱うためのフレームワーク<sup>\*6</sup>の理解、高速な学習、認識を行うためのハードウェアの準備などの課題がある。

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

\*1 Deep Learning：人の神経処理メカニズムを模して作られたニューラルネットワークをより複雑にすることによって、より複雑な概念の学習や、判断、推定の模倣を可能とした機械学習手法。

\*2 画像認識AI：画像入力に対して、機械が判断、推定などを行い、結果を返却するAI。

このような課題に対処すべく、画像認識の学習、デプロイをクラウド上で手軽に行えるようにするサービスが提供されている。例えば、Amazon Rekognition Custom Labels [1]、AutoML Vision [2] などがある。ドコモでも、自社で開発してきた画像認識技術\*7を、より多くのユーザに手軽に使っていただくべく、ドコモ画像認識プラットフォーム [3] を2020年に提供開始した。

画像認識AIの学習・デプロイ機能の提供にあたっては、ユーザごとに必要とする画像認識機能が異なるため、それらをより効率的に提供する仕組みが求められる。本稿では、ドコモ画像認識プラットフォームにおいて、複数の画像認識機能を提供するための仕組みについて解説する。また、ドコモ画像認識プラットフォームでは、ドコモの閉域網内で画像認識を行うことで、よりセキュアな画像認識環境を実現しており、このメリットについても解説する。

## 2. ドコモ画像認識プラットフォームの概要

ドコモ画像認識プラットフォームは、クラウド上で画像認識AIの学習および画像認識AIを用いた認

識（以下、「推論」）を行える状態にするデプロイの機能を提供する。

### (1)提供される画像認識機能

ドコモ画像認識プラットフォームが提供する画像認識機能の一覧を図1に示す。①物体検出は、画像の中から特定の物体を見つけ、画像内でのその物体の座標を推定する機能である。②一般物体認識は、画像に写っている物、風景、シーン、画像全体の特徴から、カテゴリに分類する機能である。③文字認識とは、画像内に写っている文字を認識する機能である。④類似画像検索は、画像と似た画像を探す機能である。⑤特定物体認識は、あらかじめ用意した多数の物体の画像と認識したい物体の画像とを照合することで、認識したい物体が何かを特定する機能である。⑥姿勢推定とは、画像内に写っている人の骨格／関節の位置を推定する機能である。

### (2)想定ユースケース

各画像認識機能の想定ユースケースを図2に示す。物体検出は、ドローン映像からの雑草検出や、監視カメラ映像からの人や車の検出に活用できる。検出した植物や、車の画像領域に対して、一般物体認識を組み合わせることで、植物の生育度や車種の判別が可能となる。また、物体検出と文字認識を組み合

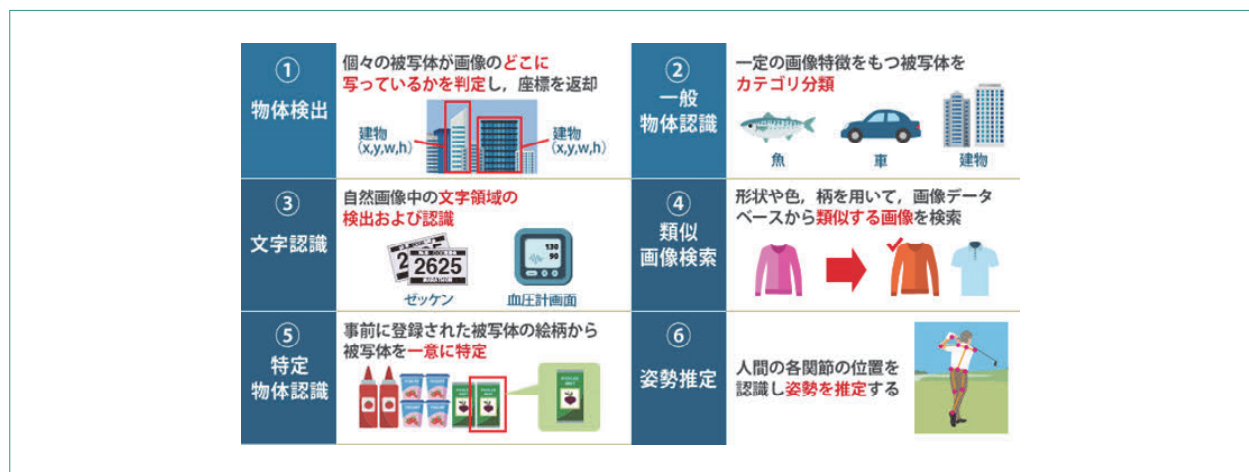


図1 ドコモ画像認識プラットフォームで提供する画像認識機能

\*3 DX：ITの浸透が人々の生活をあらゆる面で良い方向に変化させること。  
\*4 デプロイ：アプリケーションをそれらの実行環境に配置して展開すること。  
\*5 API：ソフトウェアの機能を他のプログラムから利用できるように切り出したインタフェース。

\*6 フレームワーク：ある領域のソフトウェアに必要なとされる汎用的な機能や基本的な制御構造をまとめたもの。ライブラリでは、開発者が個別の機能呼び出す形となるが、フレームワークでは、全体を制御するのはフレームワーク側のコードで、そこから開発者が個別に追加した機能呼び出す形となる。



図2 画像認識の想定ユースケース

わせた伝票や書類から文字を読み込む利用方法や、物体検出と特定物体認識を組み合わせた商品棚の画像を撮影するだけで、どこに何の商品が置いてあるかデータ化する利用方法などが想定される。

さらに、類似画像検索では、似たファッションアイテムを画像ベースで探す、姿勢推定ではスポーツの動作のフォームを確認するなどの用途で活用できる。

### (3)機能構成

ドコモ画像認識プラットフォームの仕組みを図3に示す。ドコモ画像認識プラットフォームでは、「物体検出」「一般物体認識（カテゴリ分類）」においては、ユーザのデータで独自のAIを提供できる学習機能と推論機能の両方をカスタム学習モデルとして提供している。また、「類似画像検索」「特定物体認識」は、ユーザ独自の画像データを基に検索・認識ができる「ユーザ辞書の作成機能」を提供している。「物体検出」「一般物体認識（カテゴリ分類）」「文字認識」「姿勢推定」においては、ユーザが独自に学習用のデータを用意しなくても推論が可能なよ

う、ドコモにて独自に学習した画像認識AI（共通学習済み学習モデル）も提供している。

### (4)Webサイト／Webコンソールの提供

ドコモ画像認識プラットフォームでは、ユーザ向けにWebサイト／Webコンソールを提供している。本コンソールにアクセスすることで、学習、学習結果の確認（「評価」）、学習して作成した画像認識AIのデプロイまでが実施できる。デプロイされたAIにはWebAPIを通して、認識リクエスト（推論リクエスト）を行うことができる。画像データをWebAPIに入力すると、認識結果がテキスト形式で返却される。

## 3. ドコモ画像認識プラットフォームにおける2つの取組み

ドコモ画像認識プラットフォームでは、2つの取組みでユーザの利便性の向上を図っている。1つ目が、画像認識機能をコンテナ化することであり、こ

\*7 画像認識技術：画像処理技術や機械学習（\*8参照）技術を用いて、画像を機械に理解させ、意味を取り出す技術。

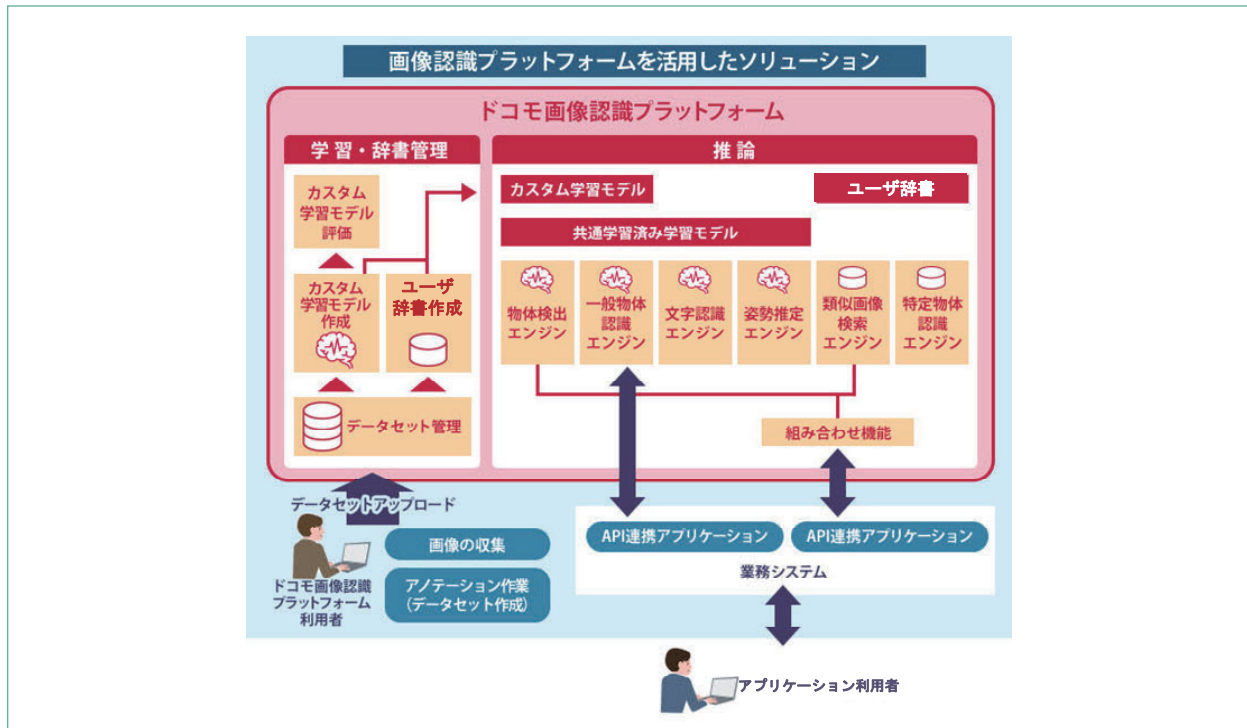


図3 ドコモ画像認識プラットフォームの仕組み

れにより画像認識機能の提供／開発／更新スピードが増加する。充実した画像認識機能のラインナップが可能となり、ユーザの課題を解決できる画像認識機能を提供できる。2つ目は、ドコモ画像認識プラットフォームをドコモの閉域網内のデータセンタ上に構築し、インターネットを介さずに接続できるようにすることである。これにより、セキュアな画像認識AIの提供を実現している。

### 3.1 コンテナを活用した画像認識AIのWebAPI化

前述のとおり、ドコモ画像認識プラットフォームでは、複数の画像認識機能を提供している。ユーザが必要とする画像認識機能は、ユーザの解決したい課題によって異なり、今後も増加することが予想される。また、画像認識AI技術は日々向上しており、より高精度／高速なものをユーザに提供するために

は、すでに提供している画像認識機能の逐次的なアップデートが必要となる。そこで、ドコモ画像認識プラットフォームでは、各画像認識機能の提供／開発／更新スピードの向上のため、コンテナを用いた機能の仮想化と、コンテナ内での機能のモジュール化／共通化を実施している。概念図を図4に示す。

コンテナを用いた仮想化とは、画像認識機能のようなアプリケーションとその起動に必要なライブラリなどを「コンテナ」と呼ばれるパッケージにまとめ、サーバのOS／環境に可能な限り依存せずに、アプリケーションを実行できるようにする技術である。ドコモ画像認識プラットフォームでは、各画像認識機能のコアプログラムと、画像認識機能をWebAPIとして提供するためのインタフェースを1つのコンテナとしてまとめている。これにより、ドコモ画像認識プラットフォーム内では、コンテナを追加／削除することで、新しい画像認識機能の追加

や、既存の画像認識機能のアップデート、認識リソースの増強（より多くの画像を学習／推論できるようにするための変更）が可能となっている。

また、画像認識のコンテナ内では、前述のインタフェースを司るプログラムと、画像認識のコアプログラムは独立であり、インタフェースとなるプログラムを共通化することで、新規の画像認識機能を開発する際の開発範囲を削減している。インタフェースを司る前者のプログラムと、コアプログラムとの間のインタフェースの入出力形式を、表1のとおり各コアプログラム側で柔軟に変更できるようにし、画像認識機能の拡張のための柔軟性を確保している。

加えて、画像認識AIのうちユーザごとに学習した画像認識モデルをコンテナ外に独立して保存している。これにより、ユーザの作成した特定の画像認識AIのリソースのみを増強することや、コンテナの更新時にコンテナのみを入れ替え、モデルを引き継ぐことを可能としている。

### 3.2 ドコモ・オープンイノベーションクラウドを活用したセキュアな画像認識

画像認識機能を利用するには、画像というセンシティブな情報の入力が必要となるため、セキュリティの確保も課題となる。画像認識の過程で、画像

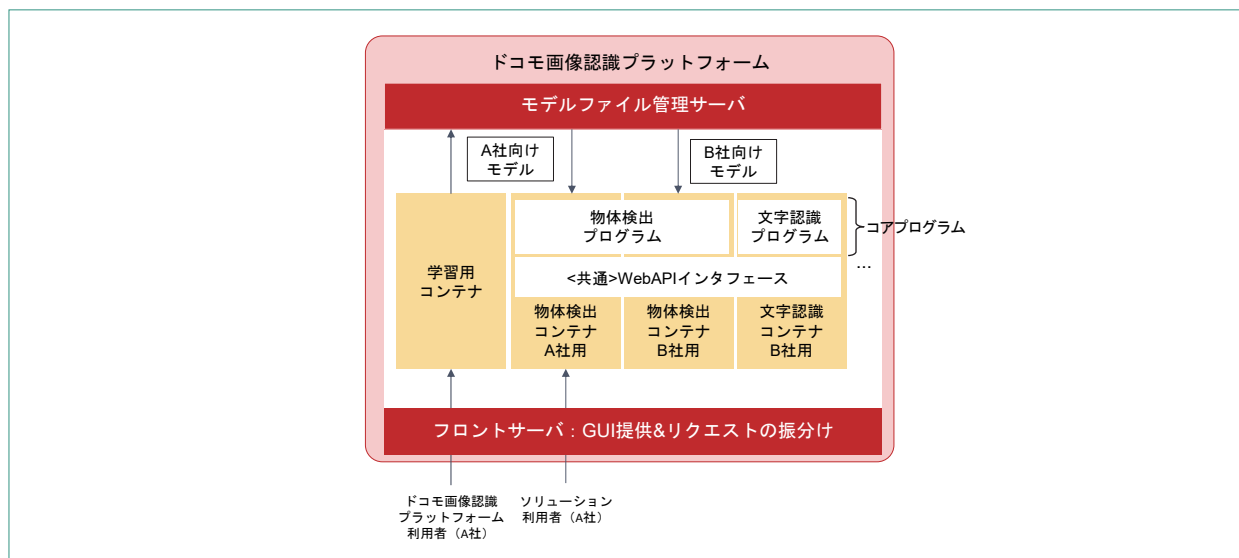


図4 コンテナ化を用いた画像認識機能提供

表1 画像認識のインタフェースプログラムとコアプログラムの役割分担

	共通化	各コアプログラムにて定義
コンテナの起動時	<ul style="list-style-type: none"> <li>学習した画像認識モデル</li> <li>ファイルの入力方法</li> </ul>	<ul style="list-style-type: none"> <li>学習した画像認識モデルのロード方法</li> </ul>
推論時の入力	WebAPIの入力形式 <ul style="list-style-type: none"> <li>JSON形式</li> <li>multipart/form-data形式</li> </ul>	WebAPIの入力の詳細 <ul style="list-style-type: none"> <li>画像をどのようにWebAPI入力のリクエスト内に格納するかなど</li> </ul>
推論時の出力	WebAPIの出力形式 <ul style="list-style-type: none"> <li>JSON形式</li> </ul>	WebAPIの出力の詳細 <ul style="list-style-type: none"> <li>画像認識結果をどのようにwebAPI出力のJSON形式内に格納するか</li> </ul>

情報が外部に出ることがないように、画像認識機能の利用にあたっては、画像の伝送経路、および画像認識を行うサーバ側で留意が必要となる。

画像認識を行う際には、画像認識システムをクラウド上や、自前のデータセンタ上で構築することにより、処理を集約して画像認識処理を行うことができる。このため、クラウド／データセンタ上の計算リソースの利用率を上げることができ、計算リソースを効率的に利用して画像認識処理が可能となる。その際システムのセキュリティ対策を適切に実施することでリスクを低減することが考えられる。しかし、この場合も画像をクラウド／データセンタに送信する経路のセキュア化が別途必要であり、インターネットを介さずセキュア化するためには専用線を利用するなどの対策が必要となる。

この課題を解決するために、画像認識処理をローカルのPCや端末、エッジ側のデバイスで行う方法がある。これにより、画像のネットワーク上での伝送を無くすことができ、画像はデバイス内にとどまるため、セキュリティを確保することができる。一方で、エッジ側のデバイス、PCや端末は画像認識を行うには十分な性能が無く、認識処理に時間がかかるといった問題がある。また性能が十分な場合でも、画像認識を個別のデバイスで行う際には、クラ

ウド／データセンタで処理を行う場合と比べ、画像認識処理を集約してリソースの利用効率を上げることができないという課題がある。

そこでドコモ画像認識プラットフォームでは、両者の中間として、ドコモの通信網内にデータセンタ（ドコモ・オープンイノベーションクラウド）を設置し、そこで画像認識処理を行う構成をとることで、これらの問題を解決した。概念図を図5に示す。画像認識処理は、ドコモの通信網内のデータセンタで行われるため、ドコモ回線（第5世代移動通信システム（5G）もしくは第4世代移動通信システム（4G））を通じたアクセスの場合、インターネットに出ることなく通信することが可能となる。これにより、画像情報が外部に流出するリスクを減らすことができる。一方で、画像認識機能はデータセンタ側に集約されるため、計算コストの最適化も図ることが可能となる。

ドコモ画像認識プラットフォームでは、この閉域網を使った接続をクラウドダイレクト [4] 接続オプションとして提供している。この閉域網を使ったアクセスには、インターネットを経由しないため接続までの遅延が少ないというメリットもある。また、接続に5Gを用いることで、さらなる遅延の低下を図ることができる。

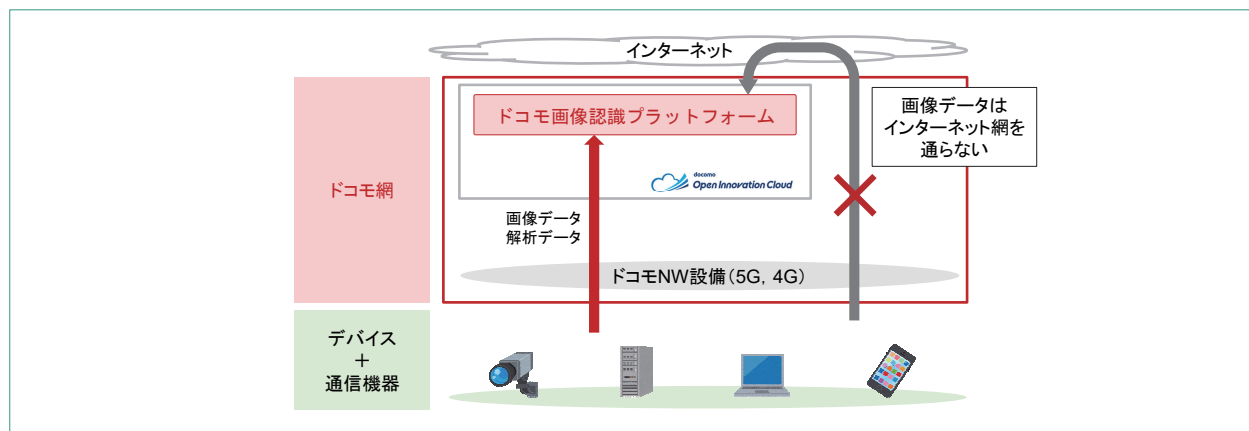


図5 閉域網接続を用いた画像認識のセキュア化

## 4. ドコモ画像認識プラットフォーム上の画像認識APIの構築・利用方法

ドコモ画像認識プラットフォームでは、APIを利用するまでに、大きく分けて学習とデプロイという2つの工程を経る。これらの作業をWebブラウザ上（クラウド上）で、シンプルなユーザインタフェースで可能としたのが本サービスである（図6）。学習をすることで個々のタスクにあったモデルの作成を実施し、デプロイをすることで学習したモデルや汎用的に提供されているモデルが利用可能となる。また、類似画像検索、特定物体認識では、ユーザ独自の辞書の作成ができる。ドコモ・オープンイノベーションクラウドの開発者ポータル内にて各工程の詳細なマニュアルが公開されている。

### 4.1 「学習」の実施

現在のドコモ画像認識プラットフォームでは、物体検出と一般物体認識の機能において学習を実施す

ることができる。具体的な学習の手順は、開発者ポータル内のチュートリアルに記載されている。

学習には画像データに加えて、アノテーションデータが必要である。アノテーションデータとは、本画像認識技術を含めた一般的な機械学習\*8の学習および評価の工程にて使用される、必要な情報をタグ付けされたデータのことである。学習の工程では、アノテーションデータを正解の情報としてモデルの学習を行い、評価の工程では、アノテーションデータと推論結果を比較して、作成したモデルの評価を実施する。アノテーションデータの作成のためのツールは、ドコモからサンプルツールとして提供されている。

ユーザは、ローカル環境にて、画像データとそれに対応するアノテーションデータを仕様に従ってまとめたデータセットを作成し、ドコモ画像認識プラットフォームにアップロードする。同様に検証用のデータセットも作成・アップロードを実施し、学習用データセット、検証用データセット、そして学



図6 サービス利用画面一例

\*8 機械学習：サンプルデータから統計処理により、有用な判断基準をコンピュータに学習させる技術。



習アルゴリズム（物体検出or一般物体認識）を設定して学習を実施する。

学習したモデルはその精度評価をプラットフォーム上で実施することが可能である。学習・検証用のデータセットと同様に、評価用のデータセットを作成・アップロードしたのち、サービス利用画面の「学習モデル評価」から学習したモデルと評価用データセットを選択し評価を実施する。評価結果はプラットフォームからダウンロードでき、画像分類・物体検出ともに、Accuracy（正解率）\*<sup>9</sup>、Precision（適合率）\*<sup>10</sup>、Recall（再現性）\*<sup>11</sup>、およびF1スコア\*<sup>12</sup>の精度を示す各指標を確認することができる。

## 4.2 「デプロイ」の実施

デプロイを実施することで、学習して作成したモデルおよびドコモにて独自に学習し提供しているモデルが、Web APIを通じて利用（推論）可能となる。デプロイした画像認識AIは、APIキーを払い出して割り当てることで利用可能となる。APIの方式としてREST API (REpresentational State Transfer API) \*<sup>13</sup>を採用し、シンプルな設計でユーザー側システムとの連携が可能である。

デプロイした画像認識AIは、サービス利用画面の「API管理」および「API認証キー設定」にて管理の実施ができる。API管理では、学習したモデルの管理および画像認識AIの増設・削除を実施することができ、簡単にスケールさせることができる。API認証キー設定では、デプロイした画像認識AIに対する認証キーの割当て・変更が可能である。

## 4.3 「辞書の作成」の実施

ドコモ画像認識プラットフォームの類似画像検索と特定物体認識では、独自の辞書を作成することが可能である。類似画像検索では、リクエストされた

画像が、事前に作成した辞書画像内のどの画像に類似しているかを判定できる。また、特定物体認識では、リクエストされた画像を、作成した辞書内の画像と照らし合わせることで、リクエストされた画像が何なのかを特定できる。

どちらの辞書作成においても、画像の収集およびアノテーションの実施が必要となる。辞書作成後は、他の画像認識機能同様にデプロイを実施することで各機能を利用することができる。

## 5. あとがき

本稿では、画像認識サービスを手軽に利用できるようにするにあたっての背景や課題と、ドコモ画像認識プラットフォームにおいて提供している画像認識機能について解説した。また、ドコモ画像認識プラットフォームにおいて、さまざまな画像認識機能を提供可能とするシステムの特性と、画像認識においてセキュリティ向上を図るための閉域網内での認識処理について述べた。また、実際にドコモ画像認識プラットフォームを利用する際のイメージについても解説した。今後もドコモでは、ユーザの必要とする画像認識機能を提供し続けるべく、継続的な機能改善、アップデートを行っていく予定である。

### 文献

- [1] AWS：“Amazon Rekognition カスタムラベル。”  
<https://aws.amazon.com/jp/rekognition/custom-labels-features/>
- [2] Google Cloud：“AutoML Vision.”  
<https://cloud.google.com/vision/automl/docs>
- [3] NTTドコモ：“ドコモ画像認識プラットフォーム。”  
<https://www.nttdocomo.co.jp/biz/service/dirp/>
- [4] NTTドコモ：“クラウドダイレクトとは。”  
<https://developer.dev-portal.d-oic.com/document/docs/cloud-direct/concepts/overview.html>

\*<sup>9</sup> Accuracy（正解率）：推論したデータのうち、正しく分類・検出できた割合。

\*<sup>10</sup> Precision（適合率）：物体検出においては、あるラベルにおいて、検出した物体のうち、確かにそれが該当の物体である割合。一般物体認識においては、あるラベルと推論したデータのうち、その推論結果が正しい割合。

\*<sup>11</sup> Recall（再現性）：あるラベルの全データのうち、正しくそのラベルと分類（物体検出）できた割合。

\*<sup>12</sup> F1スコア：PrecisionとRecallの調和平均。

\*<sup>13</sup> REST API：RESTの制約に従ったAPI。RESTはRoy Fielding氏が2000年に提唱した設計原則を基に発展した、ソフトウェアアーキテクチャのスタイル。

# スポット画像特徴量を用いた 観光スポットマイニング手法 「ジェネリック観光地」抽出技術

クロスステック開発部

かつみ ひさお やまだ わたる  
 勝見 久央 山田 渉  
 おちあい けいいち  
 落合 桂一

AI技術を用いて各地域から新たな観光スポットをマイニングすることができれば、観光客の一極集中や、難航する地方都市への観光客の誘致などの社会的課題に対する新たな解決策となり得る。本稿では有名な観光スポットに類似し、これらの代替となり得るスポットを「ジェネリック観光地」と呼称し、AI技術によってWeb上のデータからジェネリック観光地をマイニングするアイデアとその実装案を紹介する。また、実際にWeb上のデータからジェネリック観光地をマイニングし、その結果を評価した内容について解説する。本稿で紹介する取組みが今後AI技術を活用した観光産業への新たな一歩となることが期待される。

## 1. まえがき

移動手段やメディアの急速な発達により、我々はさまざまな場所を旅行することが可能になった。一方で、特定の観光地に観光客が一極集中するオーバーツーリズムが問題視されている [1]。オーバーツーリズムは、旅行者と当該観光地の住民の双方に悪影響を及ぼす深刻な問題である。これに対する有効な取組みとして、①既存の観光スポットの観光客受入れ容量を増やす、②既存の観光スポット間で観光客を時間的、空間的に分散させる、③新たに観光

スポットを創出し、観光客をそちらに誘導するという3つが挙げられる [2]。

これらのうち①では、観光施設の駐車場を整備するなどのさまざまな取組みが進んでいる。②の取組みの代表例としては、リアルタイムの混雑状況を観光スポットへの推薦に反映して混雑回避を図る手法 [3] や、Web上の画像投稿サイトにおいて、観光スポット画像ごとの投稿ユーザの属性などを分析し、観光客よりも地元住人に好まれているような穴場スポットをマイニングする手法 [4] などがある。また、③のように、これまで観光スポットとみなされ

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

ていなかったようなスポットを観光資源とみなし、都市内の観光スポットの上限数を増やすような取り組みも重要である。

ドコモは、③の新たな観光スポットを創り出す取り組みとして、すでに有名なスポットに見た目が似ていることが話題となり、さらに人気になったスポットが存在することに着目した。例えば、日本国内で有名な「竹田城跡」では、ペルーのマチュ・ピチュに見た目が似ていることがメディアなどで話題となり、その結果、観光客が増加した。このように、これまで観光スポットとみなされていなかったスポットに対して、有名スポットに類似しているという観点で観光価値を見いだすことができれば、新たな観光スポットをマイニングすることができる。ドコモは、こうした既存の有名なスポットに類似し、これらの代替となり得るスポットを「ジェネリック観光地」と名付けた。

本稿では、画像処理AI技術などによって、既存の有名観光スポットに見た目が類似したスポットを

ジェネリック観光地として提示する手法を解説する。さらに、Web上のスポット画像を用いて提案手法の評価実験を行った内容について解説する。

## 2. ジェネリック観光地抽出技術

提案手法の概要を図1に示す。本手法では、ジェネリック観光地のマイニング対象となる候補スポットおよび既存の有名観光スポットの画像をWeb上から収集し、これらすべての組合せについて画像類似度を算出する。そして、画像類似度が上位の組合せに含まれる候補スポットをジェネリック観光地として出力する。

本手法とそのアイデアは、スポット画像のみを用いて新たな観光スポットをマイニングするものであり、一般的な観光スポットマイニング技術と比較して、事前のスポットの口コミや投稿画像などの収集量に依存せずに新たな観光スポットをマイニングできる可能性がある。さらに、地元住民すらもこれま

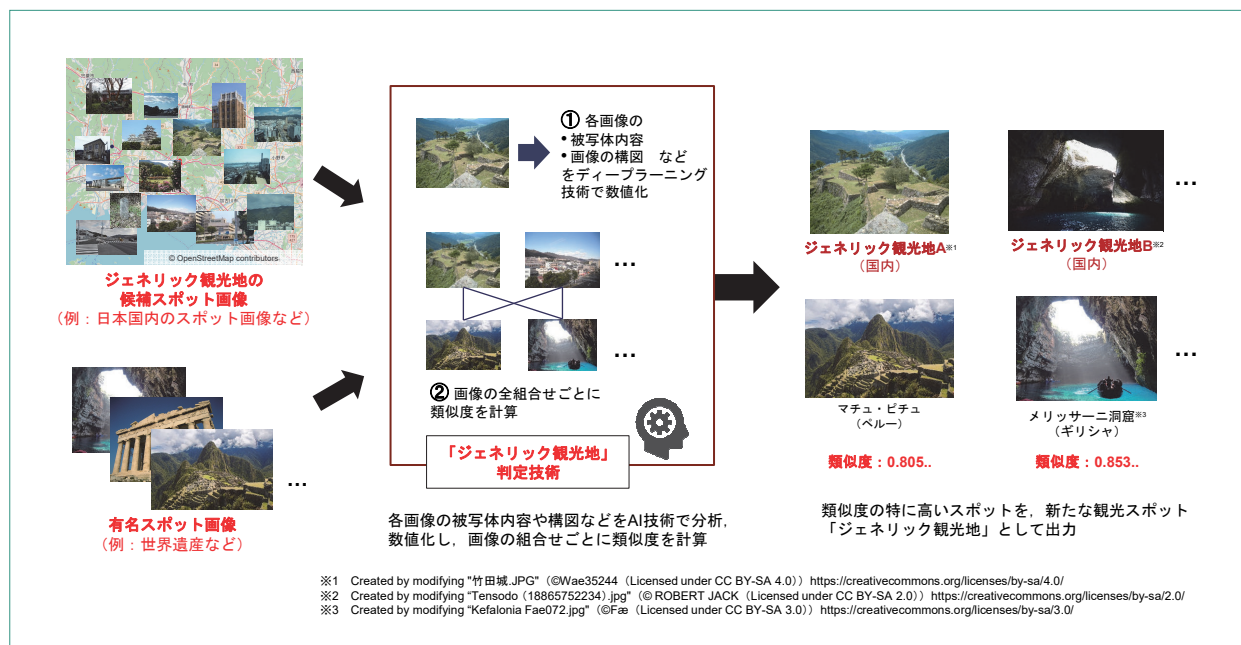


図1 提案手法の概要

で観光スポットと認識していなかったようなスポットに対しても新たな観光価値を見いだすことができる。

### 3. ジェネリック観光地抽出技術

本手法は主に4つの処理から構成される（図1）。

手順①：Web上から、任意の範囲内で撮影された候補スポット画像を収集する。

手順②：Web上から、名称を基に有名スポット画像を収集する。

手順③：収集された各候補スポット画像、有名スポット画像から、すべての組合せの画像類似度を算出する。

手順④：画像類似度の上位 $N$ 件の組合せに含まれる候補スポットをジェネリック観光地として出力する。

手順①、②では、Web上から得られたさまざまなスポット画像を利用することが想定される。また、本手法の評価実験では、写真投稿サイトであるFlickr<sup>\*1</sup>上に投稿されたスポット画像を利用した。

手順③では、すべての候補スポット画像と有名スポット画像の組合せの画像類似度を算出した。画像類似度の算出にあたっては、各候補スポット、有名スポットの画像を特徴量ベクトル<sup>\*2</sup>に変換し、それらのコサイン類似度<sup>\*3</sup>を各スポット画像同士の類似度とした。スポット画像の特徴量ベクトルとしては、各画像の内容などの意味的特徴量や、構図などの構造的特徴量を抽出する手法が有用であると考えられる。前者としては、各画像に山や河川などの365種類の風景カテゴリの成分がどの程度含まれているかを数値化したVisual Concept [5]がある。後者としては、各画像の構図や明暗箇所の分布などの特徴を抽出するGIST特徴量 [6] などがある。

最後に、手順④ですべての組合せのうち、類似度

上位の組合せを基にジェネリック観光地を出力する。本手法では、類似度上位の $N$ 件までに含まれる候補スポットをジェネリック観光地と判定して出力した。

### 4. 評価実験

提案手法の妥当性を検証するため、評価用データを作成した。さらに、提案手法を用いて評価用データから実際にジェネリック観光地をマイニングし、その結果が正解例とどの程度一致しているかを誤差として算出し、評価した。

#### 4.1 評価用データの作成

評価用データの作成方法の概要を図2に示す。まず、すでに海外の有名スポットに似ていることで知られている日本国内の観光スポット5カ所（図2の左の表）を、ジェネリック観光地の正解例とした。

次に、これらの正解例を用いて、ジェネリック観光地の正解例を含む候補スポット画像をFlickr上から収集した。具体的には、5つの正解例が属する標準地域メッシュ<sup>\*4</sup>内で撮影された画像を、候補スポット画像として収集した。標準地域メッシュとして、日本国内を約10km四方のメッシュに分割する2次メッシュを用いた。図2右側は、該当の2次メッシュ内に評価データとして利用したジェネリック観光地の正解例を★印、候補スポット画像の撮影地点を●印でプロットしたものである。このような方法で2,353枚の候補スポット画像が得られた。

また、5つの正解例に対応する有名スポット画像については、正解例に含まれる5つの有名スポット名の英名をそれぞれクエリ<sup>\*5</sup>としてFlickr上から得られた画像5枚ずつ、合計25枚を利用した。ただし、実際に有名スポット名の検索結果として得られる画像には、スポットそのものを撮影した画像だけでなく、人物や料理などを撮影したようなスポット画像以外の画像がノイズとして含まれる。評価用データ

\*1 Flickr：米国Oath Inc. の商標または登録商標。

\*2 特徴量ベクトル：データのパターンや特徴を数値の列であるベクトルで表現し、プログラムなどで扱えるようにしたもの。

\*3 コサイン類似度：2つのベクトルの向きがどの程度近いかを数値化したもの。

\*4 標準地域メッシュ：日本の各地域データを集計する用途などのために、日本国内を緯度経度に従ってメッシュ状に分割したもの。

\*5 クエリ：データベースに対する問合せ（処理要求）。

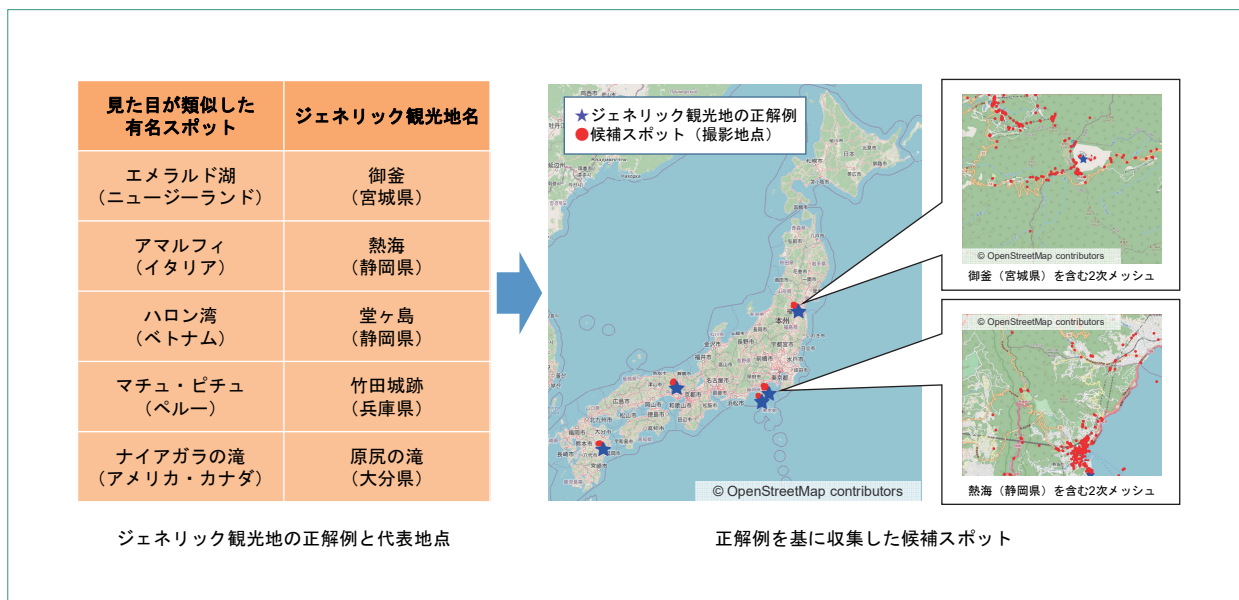


図2 評価用データの作成

からはこのようなノイズとなる画像を手で除去した。

なお、上記の収集画像は、評価用データの作成段階で投稿者がFlickr上で設定していたクリエイティブ・コモンズ・ライセンス\*6表記を基に、研究用途での利用が可能な画像のみを対象として得られたものである。

## 4.2 スポット画像の特徴量ベクトル化

本稿で提案するジェネリック観光地のマイニング手法では、各スポット画像を特徴量ベクトル化してからスポット画像間の類似度を算出する。評価時には特徴量ベクトル化の方法として次の3種類の方法を用いて、それぞれ得られた出力結果を比較した。

手法①：Visual Concept

この手法では、各スポット画像が山、河川、海など365種類の風景カテゴリのそれぞれにどの程度当てはまるかをニューラルネットワーク\*7で判定し、スポット画像をそれぞれ365次元の特徴量ベクトルに変換した。なお、ニューラルネットワークは風景画像データセットである

Places365を用いて学習されたResNet18を用いた [7]。

手法②：GIST特徴量

各スポット画像をGISTにより特徴量ベクトル化した。GIST特徴量は、各スポット画像内の大まかな明暗の分布などの構造的な特徴を抽出したものである。

手法③：埋込みベクトル\*8

ニューラルネットワークが段階的に特徴量の抽出を行う性質を考慮し、手法①で用いたニューラルネットワークが最終的にVisual Conceptとして365次元の特徴量ベクトルを算出する前段階の512次元のベクトルを特徴量ベクトルとして利用した。このようにして得られる特徴量ベクトルには、ニューラルネットワークが手法①で最終的にVisual Conceptとして365種類の風景カテゴリを推論する直前の、より抽象度の高い特徴が埋め込まれており、365種類のカテゴリ上では異なるカテゴリに分類されてしまうような構造物などに対しても類似し

\*6 クリエイティブ・コモンズ・ライセンス：クリエイティブ・コモンズが定めるルールに従って、著作物の利用ルールなどについて意思表示したもの。

\*7 ニューラルネットワーク：生物の脳の働きを数理的モデルによって模倣し、数値的なパターンの認識や推論ができるようなモデル。

\*8 埋込みベクトル：高次元の特徴量から必要な情報を低次元に絞り込んでベクトル化したもの。

ているものとして扱い易くなる。

### 4.3 評価結果と考察

上記の方法で作成した評価用データから、提案手法を用いて $N$ カ所 ( $N=10, 20, 30$ ) のジェネリック観光地をマイニングした。手法③を用いて $N=30$

としてマイニングされたジェネリック観光地のスポット画像を図3に、それらの位置を図4に示す。マイニングされたジェネリック観光地と正解例の誤差については、図4右側に示すように、各メッシュ内でジェネリック観光地としてマイニングされた合計 $N$ カ所のスポットと正解例の位置との平均距離

ジェネリック観光地と判定されたスポット	竹田城跡 (兵庫県) "Takeda-jo ruins" © cotaro70s (CC BY-ND 2.0)	熱海周辺 (静岡県) "Atami Port in Atami City, Shizuoka, Japan" © Izu navi (CC BY 2.0)	原尻の滝 (大分県) "原尻の滝 1" © wa_me (CC BY-NC-ND 2.0)	堂ヶ島 (静岡県) "APC_0064_iPhone.jpg" © Misako Kuniya (CC BY-NC 2.0)	御釜 (宮城県) "御釜" © NISSAN MOTOR CO., LTD. (CC BY-NC-ND 2.0)
類似している有名スポット	マチュ・ピチュ (ペルー) "Machu Picchu" © F Delventhal (CC-BY 2.0)	アマルフィ (イタリア) "Amalfi cartolina" © Antonio Salsano (Salsano2016 - SIAE) (CC BY-NC-ND 2.0)	ナイアガラの滝 (アメリカ・カナダ) "Niagara Falls" © Peter Miller (CC BY-NC-ND 2.0)	ハロン湾 (ベトナム) "Quảng Ninh VN - Hạ Long Bay 03" © Daniel Mennerich (CC BY-NC-ND 2.0)	エメラルド湖 (ニュージーランド) "Emerald Lake" © Andrea Schaffer (CC BY 2.0)

※画像は各画像の下部に示す作品名および著作者の作品から複製または変更したものを、その下部に示すCreative Commonsライセンスに記載された条件に従って使用している。

図3 手法③ ( $N=30$ ) によってジェネリック観光地としてマイニングされたスポット例

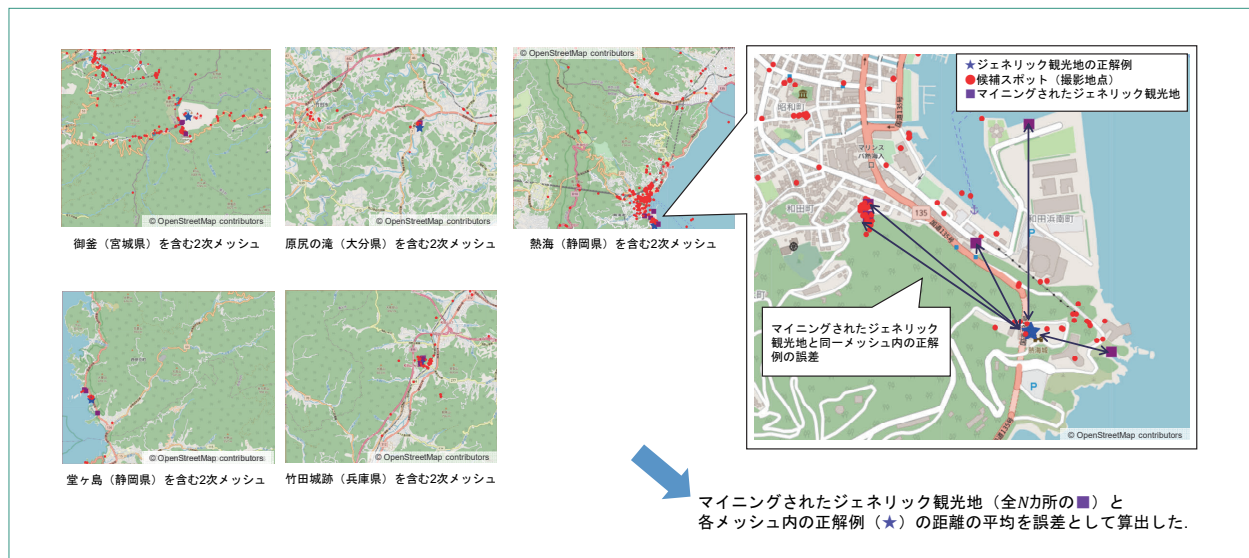


図4 手法③ ( $N=30$ ) によってジェネリック観光地としてマイニングされたスポットの位置と誤差の算出方法

(km)として算出した。ジェネリック観光地として、より正解例に近い候補スポットをマイニングできているほど、この誤差は小さくなり、正解例と位置が完全に一致するスポットをマイニングすると誤差は0となる。ただし、Web上から収集されるスポット画像には同じ対象を撮影したものでも撮影地点が異なるものが含まれるため、正解例を正しくジェネリック観光地としてマイニングしていても誤差は必ずしも0になるとは限らない。例えば図4右側に示すように、熱海周辺でのマイニング結果では、正解例に含まれるアマルフィに類似した熱海のスポット画像を撮影することができる複数カ所の撮影スポットがジェネリック観光地としてマイニングされていた。最終的に得られた各特徴量化手法とNの組合せごとの誤差の評価結果を表1に示す。

図3、4および表1の結果から、おおむね正解例の近傍で有名スポットに類似した見た目の候補スポットをジェネリック観光地としてマイニングできていると考えられる。また、表1に示すように、すべての場合で手法③がその他の手法を上回る結果となった。さらに、手法①と手法②を比較すると、すべての場合で前者が後者を上回る結果となった。これらの結果から、スポット画像の被写体の内容などの意味的な特徴が、構図などの構造的な特徴よりジェネリック観光地のマイニングに有用であると考えられる。

本手法は、候補スポット画像に対してクラスタリング\*9などの処理を行わないため、入力される候補スポット画像や有名スポット画像の内容や枚数に

関わらず手法を適用することができると考えられる。今後、評価データの規模を拡大した場合も同様の結果が得られるかどうかを評価する予定である。

## 5. あとがき

本稿では、既存の有名スポットに類似したスポットをジェネリック観光地と呼称し、新たな観光スポットとしてマイニングするアイデアについて解説した。また、ニューラルネットワークなどを用いたマイニング手法を提案し、Web上のスポット画像を用いて評価を行った内容について述べた。ただし、本稿で解説した提案手法はWeb上から収集した5カ所の正解例を基に評価を行ったものであるため、今後、提案手法が大量の候補スポット画像や有名スポット画像の中からも正しくジェネリック観光地をマイニングできるかどうかや、マイニングされたジェネリック観光地を観光客の一極集中の緩和や観光客の誘致に活用できるかどうかを検証する予定である。

本稿で解説したアイデアや提案手法が、AI技術を活用してこれまで観光スポットと見なされていなかったスポットから新たな観光価値を見いだし、オーバーツーリズムだけでなく地方創生などへも活用されるような取組みの第一歩となることを期待している [8]。また、昨今の新型コロナウイルス感染防止を意識した新たなライフスタイルの受容とも相まって、海外や国内の遠く離れた場所への旅行だけでなく、これまで観光価値に気付かなかった近場の

表1 各特徴量化手法とNの組合せごとの誤差

	上位10スポットの平均誤差 (km)	上位20スポットの平均誤差 (km)	上位30スポットの平均誤差 (km)
手法①Visual Concept	2.18	2.20	2.22
手法②GIST	2.56	2.35	2.45
手法③埋込みベクトル	0.33	0.27	0.34

\*9 クラスタリング：大量のデータを類似した特徴をもつデータ同士でグループにして分割すること。

さまざまなスポットを新たな視点で捉え、その隠れた魅力を再認識するようなニューノーマル\*10な観光スタイル確立への一歩となれば幸いである。

### 文 献

- [1] M. Duignan : “Overtourism”? Understanding and Managing Urban Tourism Growth beyond Perceptions : Case Studies,” United Nations World Tourism Organisation (UNWTO), pp.34-39, Mar. 2019.
- [2] T. Mainil, E. Eijgelaar, J. Klijs, J. Nawijn, and P. Peeters : “Research for TRAN Committee-Health tourism in the EU : a general investigation,” European Parliament, Directorate General for Internal Policies, 2017.
- [3] M. Hidaka, Y. Kanaya, S. Kawanaka, Y. Matsuda, Y. Nakamura, H. Suwa, M. Fujimoto, Y. Arakawa and K. Yasumoto : “On-site Trip Planning Support System Based on Dynamic Information on Tourism Spots,” Smart Cities, Vol.3, No.2, pp.212-231, Apr. 2020.
- [4] C. Zhuang, Q. Ma, X. Liang and M. Yoshikawa : “Anaba : An obscure sightseeing spots discovering system,” 2014 IEEE International Conference on Multimedia and Expo (ICME), pp.1-6, Sep. 2014.
- [5] C. Peters, T. Deselaers, N. Ferro, J. Gonzalo, G. J. F. Jones, M. Kurimo, T. Mandl, A. Penas and V. Petras : “Evaluating Systems for Multilingual and Multimodal Information Access,” 9th Workshop of the Cross-Language Evaluation Forum, CLEF 2008, Aarhus, Denmark, Revised Selected Papers, p.527, 2008.
- [6] A. Oliva and A. Torralba : “Modeling the Shape of the Scene : A Holistic Representation of the Spatial Envelope,” International journal of computer vision, Vol.42, No.3, pp.145-175, May 2001.
- [7] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva and A. Torralba : “Places : A 10 Million Image Database for Scene Recognition,” IEEE Transactionson on Pattern Analysis and Machine Intelligence, Vol.40, Issue 6, pp.1452-1464, Jul. 2017.
- [8] NTTドコモ報道発表資料：“日本初！海外の有名スポットにそっくりな四国の景観を発見するフォトコンテストを開催 ～遠出が難しい今、海外旅行気分を味わえる場所を発見！風景画像の類似度をAIで判定～,” Jul. 2021.

---

\*10 ニューノーマル：社会の環境や情勢の変化に伴い、不可逆的に新たな常識が定着した状態。



# エッジAI対応5Gデバイスを活用したパーソナルモビリティ向け自動運転支援システム

移動機開発部

たかはし 高橋	まこと 誠	なかざわ 中澤	ゆうき 祐希
もりべ 森部	ともや 智也	いけだ 池田	けんや 健矢

高齢化や労働力不足に伴う社会課題の解決に向けて、さまざまなパーソナルモビリティやサービスロボットの自動運転サービスへの期待が高まっている。これらのサービスの社会実装には、安全性確保が最重要課題となっており、障害物回避性能向上や遠隔操縦サービスへの対応が求められている。そこでドコモでは、リアルタイム性に優れたデバイス上での障害物検知や、閉域環境での5G通信を用いたセキュアな低遅延遠隔操縦の技術開発を行い、社会実装に資する安全な自動運転支援を可能とした。

## 1. まえがき

日本の総人口は2020年10月1日時点で1億2,571万人となった。65歳以上の人口は3,619万人であり、高齢化率は28.8%にのぼる。また、2065年には、約2.6人に1人が65歳以上となる統計データが内閣府より示されている [1]。このような高齢化社会に対して、高齢者が手軽に移動できるような自動運転 車いすや、自宅までのラストワンマイル区間の自動運転ロボットによる配送などの省人化サービスの社会実装\*1が期待されている。自動運転サービスの社会実

装に向けては安全性の確保が最重要課題となり、従来のLiDAR (Light Detection And Ranging)\*2などの距離センサに加えて、デバイス上での画像認識技術\*3を用いた即時対応可能な障害物検知機能、また自律走行\*4が困難な場合での自動運転を可能とする遠隔操縦モードへの対応、などの安全性向上に向けた取組みが求められている。

本稿では、さまざまなパーソナルモビリティ\*5やサービスロボットの安全な自動運転に向けて、ドコモのMEC (Multi-access Edge Computing)\*6環境上に構築した遠隔操縦システムとエッジAI\*7対応

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

\*1 社会実装：本稿では、社会課題を解決する仕組みとして社会生活に浸透し、継続的に利用されていく状態を示す。

\*2 LiDAR：近赤外光などの光を周囲の物体に照射し、反射光により物体検知を可能とするセンサ。

\*3 画像認識技術：画像処理技術や機械学習技術を用いて、画像を機械に理解させ、意味を取り出す技術。

5Gデバイスを組み合わせて活用することで実現した、パーソナルモビリティ向け自動運転支援システムについて解説する。

## 2. システム概要

ドコモが開発したパーソナルモビリティ向け自動運転支援システムの構成図を示す（図1）。本システムは、①さまざまなパーソナルモビリティやサービスロボットにアドオンして用いる「エッジAI対応5Gデバイス」、②リアルタイム性が要求される処理を当該デバイス上で行う「エッジAIアプリケーション」、③パーソナルモビリティ自動運転の監視・走行サポートを可能とする「遠隔操縦システム」の3つの機能ブロックからなる。なお、本システムでは車いすなどのパーソナルモビリティ自体を開発するのではなく、パートナー企業の開発機体にアドオンすることで利用可能な構成を取っており、さまざまなタイプのパーソナルモビリティやサービスロボットに対して安全性を高める自動運転支援が可能であ

る。

具体的には、本システムを用いることで、パーソナルモビリティ周辺の、ある障害物が人物であることを認識した上で障害物回避を行うなどの適切な機体制御が可能となり、LiDARなどの距離センサのみを用いた自動運転システムの安全性をさらに高めることができる。また、自動運転サービスの社会実装に向けては、パーソナルモビリティの自律走行技術の安全性向上とともに、自律走行中の予期せぬトラブルに対して遠隔から自動運転をサポートする機能が求められる。本システムは現地映像を遠隔サポートセンタ側へ伝送し、遠隔サポートセンタ側から機体状況をモニタリングする機能を具備しており、周囲の人混みや障害物などによって、パーソナルモビリティが自律走行中に自己位置を見失った場合、または自律走行に必要なデジタルマップが存在しない場所への移動を行う場合に、遠隔サポートセンタ側の管理者による遠隔操縦を可能とする。本システムは、さらに現地映像に映り込む人たちへのプライバシー配慮として、検知した人物の顔に対してモザ

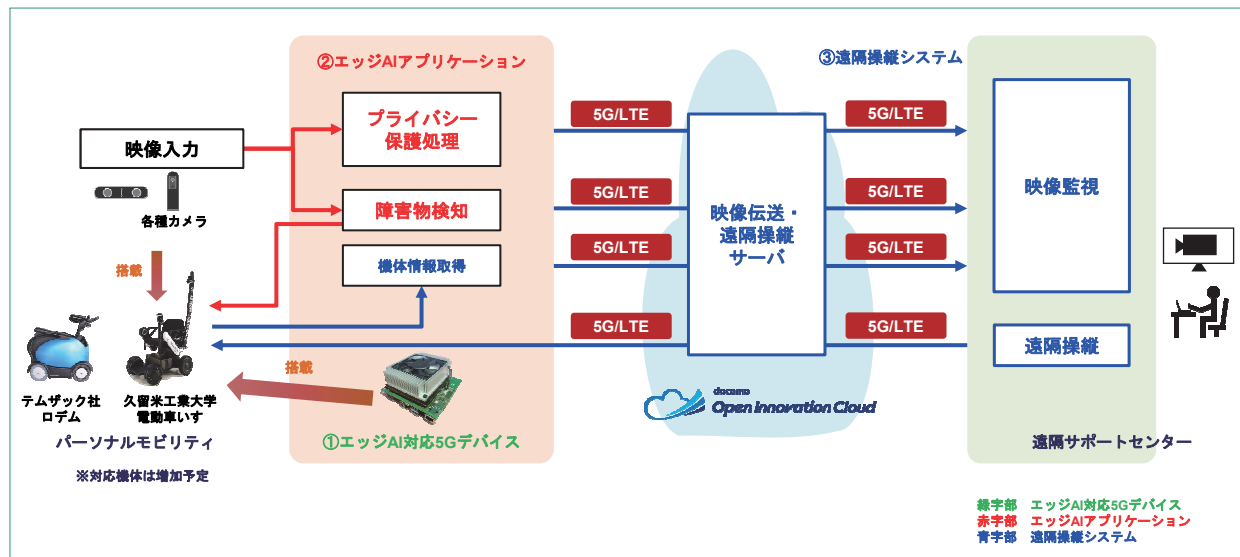


図1 「パーソナルモビリティ向け自動運転支援システム」の全体構成図

- \*4 自律走行：デジタルマップおよびパーソナルモビリティ（\*5参照）に搭載された距離センサ、モーションセンサなどを用いて、自己位置推定しながら目的地へ走行する機能。
- \*5 パーソナルモビリティ：本稿では、電動タイプの個人用の乗り物を示す。
- \*6 MEC：移動通信網において、お客様により近い位置にサーバやストレージを配備する仕組み。

- \*7 エッジAI：本稿では、MEC環境上で処理するAIではなく、デバイス上で処理するAIを示す。

イク加工を行うマスク処理を施した映像を伝送することも可能となっている。

また、本システムを構築するクラウド基盤として、ドコモが提供するMEC環境であるドコモオープンイノベーションクラウド<sup>®</sup> [2] を活用し、第5世代移动通信システム（5G）を用いたクラウドダイレクト接続を行うことにより、5Gならではの高精細かつ低遅延な映像伝送、およびインターネットを経由しないことによる高セキュリティ通信が可能である。

### 3. エッジAI対応5Gデバイス

#### 3.1 デバイス概要

本デバイスは、後述の高性能プロセッサ、5Gモジュール、各種インタフェースを具備したキャリアボード<sup>\*8</sup>から構成される。パーソナルモビリティでの利用にあたっては、エッジデバイス上での画像認識などの高負荷処理、高精細映像などの大容量・低遅延通信、機体への組み込みやすさなどが求められる。本デバイスは以下の各構成要素を組み合わせることで、パーソナルモビリティ自動運転支援への活用を可能とした。

#### 3.2 エッジAI処理可能な高性能プロセッサ搭載

パーソナルモビリティの自律走行では、リアルタイム性に優れたエッジデバイス上での複数のAI処理が求められている。一般的に画像認識などのAI処理はマシンパワーの優位性からクラウド上で実行されることが多いが、本デバイスはNVIDIA社製のエッジコンピューティング<sup>\*9</sup>用プロセッサ「Jetson AGX Xavier [3]」を搭載することで、エッジデバイス上でのAI処理を可能としている。また、より省電力・廉価な「Jetson Xavier NX [4]」にも対応しており、ユーザの用途に合わせてプロセッサの

選択が可能となっている。

#### 3.3 高速・大容量・低遅延通信を実現する5Gモジュール搭載

5G対応通信モジュールについては、ドコモネットワークとの相互接続性試験を完了したTelit社製モジュール「FN980 [5]」を採用した。遠隔操縦によるパーソナルモビリティ自動運転支援においては、高精細映像や360度カメラ映像など複数カメラを用いて安全性を高めているが、5Gの導入によりそれらの映像伝送を実現可能としている。

本モジュールは外部アンテナにも対応しており、パーソナルモビリティの機体内部に本デバイスを組み込んでも、機体外部のアンテナによって安定的な通信が可能である。また、通信モジュールとの接続インタフェースとして、汎用的なM.2コネクタ<sup>\*10</sup>を採用することにより、多様な機器との接続性を確保した。

#### 3.4 外部機器接続用インタフェースを具備したキャリアボード搭載

前述の高性能プロセッサや、5Gモジュールを組み込むためのキャリアボードを開発した。本ボードでは、シリアル通信用インタフェースや、HDMI (High-Definition Multimedia Interface) 入力（「Jetson AGX Xavier」のみ）など、豊富なインタフェースを具備している（表1）。これにより、スマートフォンや一般的なルータでは困難な、外部カメラをはじめとするセンサ類との接続が可能となっている。また、本デバイスはパーソナルモビリティにアドオンする形で使用が想定されているため、組み込みやすさを重視した形状となっている。特定のパーソナルモビリティだけでなく、さまざまな機体への組込みに対応可能な汎用性をもっている。

\*8 キャリアボード：必要な入出力インタフェースを持つハードウェア。

\*9 エッジコンピューティング：ユーザの近くにエッジサーバを分散させ、距離を短縮することで通信遅延を短縮する技術。

\*10 M.2コネクタ：薄型で高性能なデバイスに適した接続端子。

表1 キャリアボードの製品仕様

サイズ	120×120mm
動作温度	-20～+80℃
動作湿度	10～90%
インタフェース	<ul style="list-style-type: none"> <li>・ 1×HDMI Type A (出力)</li> <li>・ 1×USB3.1 Type C</li> <li>・ 2×USB3.1 Type A</li> <li>・ 1×USB2.0 Micro B (OTG)</li> <li>・ 1×RJ-45 for GbE</li> <li>・ 1×Micro SDスロット</li> <li>・ 1×DC-in 9～19V (6pinユーロブロック)</li> <li>・ 1×M.2 B-Key 3052 (Telit FN980m 5G/LTE)</li> <li>・ 1×M.2 M-Key 2280 (SC710N1 M2 HDMI) (1×HDMI Type A入力)</li> <li>・ 1×M.2 M-Key 2280 (M.2外部ストレージ追加可能)</li> <li>・ 1×nano SIMスロット</li> </ul>
拡張ピンヘッダー	<ul style="list-style-type: none"> <li>・ 1×CANBUS</li> <li>・ 1×UART 3.3V/5V TTL (JST-GH 6pin)</li> <li>・ 1×UART 3.3V/5V TTL (DF-13-6pin)</li> <li>・ 1×RS-232/1×I2C/5×GPIO</li> <li>・ 1×MIPI CSIコネクタ (120pin)</li> <li>・ 1×フロントパネル (リセット/リカバリ/電源ON)</li> </ul>

## 4. エッジAIアプリケーション

### 4.1 アプリケーション概要

ドコモは、パーソナルモビリティ自動運転の安全性向上を目的としたリアルタイム性が要求される障害物との衝突回避のための周辺映像認識や、遠隔操縦のための伝送映像に写り込んだ人物に対するプライバシー保護処理などの、デバイス上で動作するアプリケーションを開発した。

### 4.2 パーソナルモビリティ周囲の障害物検知

自律走行時には、パーソナルモビリティの周囲に存在する障害物を精度よく把握するため、ステレオカメラ映像を用いた画像認識・解析を行うことで、障害物種別の判別や障害物までの距離を算出可能とした。また、障害物が高などの移動物体である場合には今後の移動方向・移動量も予測する。これらの映像解析による出力情報について、クラウドを介さ

ずにデバイスから即時にパーソナルモビリティへ通知することで、安全な自動運転に繋がる速やかな機体制御を可能としている。

遠隔操縦時には、遠隔操縦者の操作性向上およびシステムの安全性向上のため、パーソナルモビリティの周囲360度の画像認識・解析を行い、それらを遠隔地の画面に表示することで、付近のどの位置に人が存在しているか遠隔操縦者が視覚的に確認可能とした。

### 4.3 プライバシー情報の保護

医療施設などで映像伝送を行う際、伝送映像に移り込んだ人物に対するプライバシー保護が要求される場合がある。本システムでは当該ケースへの対応のため、遠隔操縦者に映像伝送する際にデバイス上で検知した人物へのマスク処理を可能とした。ただし、人物全体にマスク処理を行うと遠隔操縦者の視認性が損なわれることから、検知した人物の顔のみ

にマスク処理を行っている。なお、人の顔自体を画像認識対象とすると、認識対象部分が小さくプライバシー保護精度が低くなる傾向が見られたため、人物検出と顔検出を段階的に用いた。

#### 4.4 映像・制御情報伝送の低遅延化

パーソナルモビリティの自律走行や遠隔操縦では、衝突防止や安全性確保のため、高負荷処理に対応できるだけでなく、低遅延であることも非常に重要となる。また、エッジデバイス上で複数AI処理を並列に行うため、限られたエッジデバイスの計算処理リソースで効率的に処理することも求められる。本システムでは安価かつ迅速なサービス開発のため、オープンソースAIフレームワーク<sup>\*11</sup>であるYolo [6]を活用し、エッジAIアプリケーションの処理を高速化するための工夫を図っている。

本システムでは、推論処理時に単精度浮動小数点演算<sup>\*12</sup>ではなく半精度浮動小数点演算<sup>\*13</sup>を採用している。例えば、人のようにある程度の大きさをもつ認識対象の場合、単精度浮動小数点演算により計算精度を向上させても推論の精度が大きく向上しないことが知られている。そこで単精度浮動小数点演算よりも高速な推論処理が可能である半精度浮動小数点演算を活用し、また、エッジAIアプリケーションの各処理パラメータ最適化に加え、デコード用のバッファサイズの最適化や、入力画像のダウンサンプリング<sup>\*14</sup>などによってエッジデバイスにかかる計算処理リソース全般の節約を図り、低遅延化

を実現した。

## 5. 5G遠隔操縦システム

### 5.1 システム概要

本システムはエッジAI対応5Gデバイス（図1①）、ドコモオープンイノベーションクラウド上に構築された映像伝送・遠隔操縦サーバ、監視用Web/iPad<sup>\*15</sup>アプリケーション、操縦用コントローラー（図1③）から構成される。映像伝送では、デバイスからエッジAIアプリケーションが生成した映像や各種センサ情報がサーバを経由して送信され、遠隔サポートセンタの監視用Web/iPadアプリケーション上に表示される。遠隔操縦では、コントローラーに入力された操縦情報がサーバを経由してデバイスからパーソナルモビリティへと送信される。通信は5G/LTEに対応し、ドコモオープンイノベーションクラウド上のサーバへのクラウドダイレクト接続を採用している。本システムの検討にあたって抽出した、遠隔操縦の実用化に向けた課題を示す（表2）。

### 5.2 ドコモオープンイノベーションクラウドを活用した低遅延性・高セキュリティ性

低遅延性の観点では、遠隔サポートセンタ側の管理者が遠隔操縦を安全に行うために、映像の入力から操縦情報の入力・反映までを500ms以内に完結させることが求められている。この課題に対し、前述したエッジデバイス上の処理の工夫に加え、ドコモ

表2 遠隔操縦システム実用化に向けた課題

①低遅延性	状況確認から操作の反映までの低遅延性が必要。
②全周囲の安全確認	横方向または後方から接近する人・物体の視認が必要。
③機体情報の確認	機体の残バッテリー情報など、機体情報の監視が必要。
④機体に応じた柔軟なコントローラー対応	システムの拡張性を高めるために、機体に応じた指定コントローラーへの対応が必要。

\*11 オープンソースAIフレームワーク：AI開発を容易にする、さまざまなプログラムがまとまった骨子をAIフレームワークと呼ぶ。現在さまざまなAIフレームワークがオープンソースとして無償で利用可能になっている。

\*12 単精度浮動小数点演算：32bit演算。FP (Floating Point) 32と省略される。

\*13 半精度浮動小数点演算：16bit演算。FP16と省略される。

\*14 ダウンサンプリング：計算量削減のために画像の解像度を低くすること。

\*15 iPad：Apple、Appleのロゴ、iPadは、米国および他の国々で登録されたApple Inc.の商標。TM and © 2020 Apple, Inc. All rights reserved.

独自のMEC基盤であるドコモオープンイノベーションクラウドと、専用接続回線クラウドダイレクトを組み合わせることで、要求された500ms以内の低遅延を実現した。ドコモオープンイノベーションクラウド上にシステム構築することで、インターネット上に存在するパブリッククラウド\*16を使用する場合に比べ低遅延性を向上させている。遅延量は、エッジAIアプリケーションによるプライバシー保護処理を加えた映像を伝送し、遠隔地で表示するまでに300ms程度、操縦情報の伝送が100ms程度であり、トータルで遅延500ms以下の安全な遠隔操縦の達成をベンダとの実証実験により確認している。さらに低遅延以外のメリットとして、通信経路がドコモ網内に閉じている閉域構成であることにより外部から分離されるため、不正アクセスのリスクを下げることができ、高いセキュリティ性も実現している。

### 5.3 映像・センサ情報の可視化

遠隔操縦する際の安全性確保の観点では、前方の視界確保だけではなく、横方向または後方から接近

する人・物体の視認が求められている。また、機体情報の確認として、パーソナルモビリティのバッテリー情報やセンサにより取得した姿勢情報などを監視することも、安全な走行のために求められている。これらの課題に対しては、監視用Web/iPadアプリケーション上でWebカメラの前方映像、360度カメラによる全周囲映像の両方を表示可能とすること、また、デバイスから任意のテキスト、数値情報をサーバに送信し、Webアプリケーションにて可視化する機能で対応している(図2)。360度カメラ映像についてはドラッグ操作による視点操作が可能であり、操縦者は全周囲の着目したい部分に視点をフォーカスすることが可能である。さらに、可視化情報として360度カメラ映像からAIアプリケーションが検出した人物の位置情報を、レーダー表示する機能も備えている。また、任意のデータ可視化機能では、各パーソナルモビリティベンダの制御アプリケーションが、可視化したい情報を、指定のAPI(Application Programming Interface)\*17を用いて送信すればよい。可視化したい情報や使用するセン



図2 映像伝送・遠隔操縦システムの表示画面イメージ図

\*16 パブリッククラウド：インターネットを介して誰でも利用ができるクラウドコンピューティングサービス。

\*17 API：OSやミドルウェアなどが提供する機能を、他のアプリケーションソフトウェアが利用するためのインタフェース。

サ類などについては、各ベンダが自由に選定することが可能である。

#### 5.4 複数モビリティ・コントローラーへの汎用性

最後に本システムの操作性の観点では、さまざまなパーソナルモビリティに適した、多種多様なコントローラーの利用が求められている。この課題に対して、本システムではコントローラーからの操縦情報を独自フォーマットなどに変換することなく、透過的にサーバに伝送する仕様としている。操縦情報はドコモオープンイノベーションクラウドを介し低遅延にデータ伝送され、各パーソナルモビリティのベンダは、自社で開発する制御アプリケーションに操縦情報を取り込むだけで、遠隔操縦が実現できる。ベンダは、制御アプリケーションに即したコントローラー機器を選定するだけでよいため、各パーソナルモビリティに適したコントローラーに柔軟に対応できる。

## 6. あとがき

本稿では、高齢化や労働力不足に伴う社会課題に対する解決手段の1つとして期待されている、パーソナルモビリティやサービスロボットの安全な自動運転に寄与するパーソナルモビリティ向け自動運転支援システムについて解説した。

具体的には、Webカメラや360度カメラなどを接続して、画像認識処理や5G通信による遠隔監視用の高精細映像の低遅延映像伝送を実現する「エッジAI対応5Gデバイス」、リアルタイムな機体制御に活用可能な障害物検知や、伝送映像に対するプライバシー保護処理をデバイス上で可能とする「エッジAIアプリケーション」、セキュアかつ低遅延な遠隔

監視・遠隔操縦を可能とする、ドコモオープンイノベーションクラウドを活用した「遠隔操縦システム」についての特徴を説明した。

本領域についてはさまざまな実証実験が行われており、一部では有料での商用サービスが提供され始めているが、ハードウェアやソフトウェアを含めたサービス全体にかかる提供コストが高価であり、今後の普及に向けてはさらなる安全性の向上とともにサービス提供価格の低減が求められている。今後は、ドコモのもつ5GネットワークやMEC基盤に加え、安価かつ市場競争力のある市中の機器や技術を融合した汎用的なシステム開発によるコスト低減がより重要になると考えられる。種々の社会課題の解決に向けて、協創パートナーとともに新たな価値創造および社会実装を進めていきたい。

#### 文 献

- [1] 内閣府：“令和3年版高齢社会白書（概要版）第1節 高齢化の状況。”  
<https://www8.cao.go.jp/kourei/whitepaper/w-2021/gaiyou/pdf/1s1s.pdf>
- [2] NTTドコモ：“ドコモオープンイノベーションクラウド。”  
<https://www.nttdocomo.co.jp/biz/service/doic/>
- [3] NVIDIA：“Jetson AGX Xavier.”  
<https://www.nvidia.com/ja-jp/autonomous-machines/embedded-systems/jetson-agx-xavier/>
- [4] NVIDIA：“Jetson AGX Xavier NX.”  
<https://www.nvidia.com/ja-jp/autonomous-machines/embedded-systems/jetson-xavier-nx/>
- [5] Telit：“FN980.”  
[https://y1cj3stn5fbwhv73k0ipk1eg-wpengine.netdna-ssl.com/wp-content/uploads/2021/02/Telit\\_FN980-FN980m\\_Datasheet.pdf](https://y1cj3stn5fbwhv73k0ipk1eg-wpengine.netdna-ssl.com/wp-content/uploads/2021/02/Telit_FN980-FN980m_Datasheet.pdf)
- [6] J. Redmon, S. Divvala, R. Girshick and A. Farhadi：“You Only Look Once：Unified, Real-Time Object Detection,” Cornell University, May 2016.  
<https://arxiv.org/abs/1506.02640>

# 5G MBHにおけるSegment Routing対応ルータ装置の開発

無線アクセス開発部 伊賀上 義夫 原田 拓弥  
まつだ ゆうき おがわ ひろあき  
松田 雄大 小川 弘顕

5G時代のモバイルNWでは、従来に加えて新たな要件が求められており、基地局からコアノード間を接続するNWであるMBHにおいてもさらなる高度化が急務となっている。NW構築／運用の効率化、信頼性のさらなる向上、5G時代のトラフィック増加に合わせた大容量NWの提供、および多様なサービス（低遅延／帯域保証など）に柔軟に対応可能なNWの提供というような要件に対し、これまでMBHで採用してきた接続方式だけでは、対応が難しくなってきた。そこでドコモは、解決策の1つとしてSR技術を用いた新たな接続方式によるNW設計を検討してきた。本稿では、新たに商用導入したSR対応ルータ装置、およびその要素技術であるSR技術の概要について解説する。

## 1. まえがき

モバイルNWにおいて、基地局とコアノード<sup>\*1</sup>を接続する区間をモバイルバックホール（MBH：Mobile BackHaul）と呼ぶ。MBHは、ユーザの利用エリア拡大のために広範囲に設置される基地局を効率よくコアノードまで収容する必要がある。従来のMBHは、トラフィック収容の効率性や障害時における信頼性に優れたNW設計となっていたが、第5世代移動通信システム（5G）時代の到来により、MBHにおいても収容トラフィックの増加や5G要件である

高速大容量（eMBB：enhanced Mobile BroadBand）、低遅延（URLLC：Ultra Reliable and Low Latency Communications）、多数端末同時接続（mMTC：massive Machine Type Communications）を実現するための新たな設計が必要となってきた。新たな要件に対して効率的にMBHに基地局を収容していくためには、これまでドコモが運用してきたイーサネット方式 [1] の装置では、以下の課題に対応することが困難になることが想定される。

①5G時代のトラフィック増加に合わせた大容量NWの提供（既存装置の上限は10GbE（Gigabit

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

\*1 コアノード：交換機、加入者情報管理装置などの上位ノード。



Ethernet))

- ②NW構築／運用の効率化およびNW信頼性のさらなる向上
- ③多様なSLA (Service Level Agreement)<sup>\*2</sup> (低遅延／帯域保証など) に、柔軟に対応可能な経路制御の実現

課題①の解決には装置自体のインターフェースの拡張が求められるため、装置に具備する物理インターフェースを、現行の上限10Gから100G以上へ拡張することを検討した。課題②に対応するために、SR (Segment Routing) 方式への対応を検討した。課題①②を共に解消するために、SR対応ルータ装置で構成される新方式でのMBHを導入した。さらに、課題③を解消することを目的として、上位コントローラ<sup>\*3</sup>との連携によりモバイルNW全体での各種サービスに適したSLAを満たす経路制御およびNW／オペレーションの効率化の実現を目指した検討を進

めている。

本稿では、2020年より商用導入を開始したSR対応ルータ装置の概要、SR対応ルータ装置を用いた新たなMBHであるアクセスルータ網 (以下、AR (Access Router) 網) のNW構成、さらにAR網を構成する上で要素技術となるSRの機能概要とその優位性、および装置に具備する障害検知機能などを紹介し、最後に今後のMBHの機能拡張について解説する。

## 2. AR網を構成する装置とNW設計

### 2.1 AR網の概要

AR網は、AGR (Area aGgregation Router) と ACR (Area aCcess Router) と呼ばれる2種類のNE (Network Element)<sup>\*4</sup>、およびNEの監視・制御・設定などの機能を担うNE-OpS (NE Operation System)<sup>\*5</sup>で構成される。AR網の構成概要を図1に

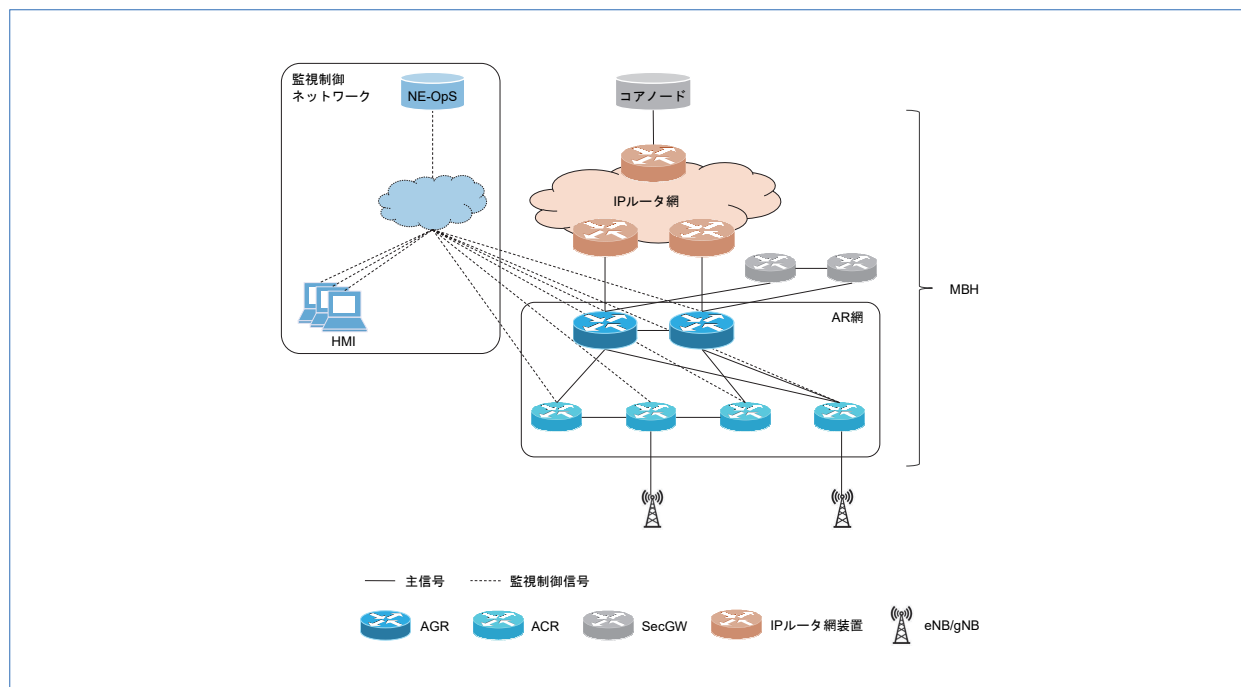


図1 AR網のネットワーク構成図

\*2 SLA：提供するサービスの品質保証。

\*3 コントローラ：NE（\*4参照）を制御する役割をもつ装置。

\*4 NE：NWを構成する装置の総称。

\*5 NE-OpS：NEを監視制御するシステムの総称。

示す。

AGR/ACRは、100GbEおよび10GbE、1GbEの物理インタフェースをもったラインカード<sup>\*6</sup>を複数收容でき、LAG (Link Aggregation Group)<sup>\*7</sup>を用いることにより、NWの必要帯域に応じて伝送容量を確保するための設備構築を可能にする (表1)。

ACRは基地局のすぐ上位に配置され、複数の基地局エリアとその通信を集約し、AGRと接続する役割を担う。ACRがAGRと接続する際のNWトポロジはスター型<sup>\*8</sup>だけではなく、隣接のACRと、AGRを介さず直接接続してビルをまたぐような長距離伝送路構築を回避することで、伝送路のコスト削減や新規ACRの追加時に柔軟かつ容易にNWトポロジを拡張することができる。また、AGRとの接続に際しては物理回線も冗長化させることができ、信頼性の観点からも障害に強いNWを構成することが可能である。

AGRは、各エリアを集約して上位のIPルータ網 [2] と接続する装置であり、高い信頼性が求められるため装置を冗長化している。これにより、片方の装置が故障してしまっても、サービスに影響なく通信を継続させることが可能となる。また、通信集約の役割以外にも、バックホール区間の通信を暗号化

するためのSecGW (Security GateWay)<sup>\*9</sup>や監視制御ネットワークとの接続を担う。

NE-OpSは監視制御ネットワークを介してNEと接続しており、これにより日本全国のNEに対して遠隔で監視・制御・設定を行うことができる。

## 2.2 AR網の特徴

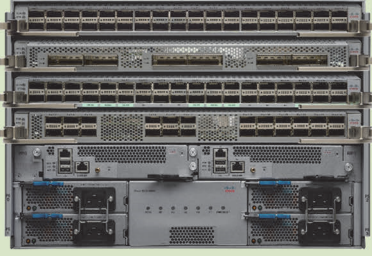



### (1)多様なユーザ・システム收容

AR網ではVPN (Virtual Private Network)<sup>\*10</sup>方式を、5Gなどのセルラシステムや無線LANシステムなどにシステム単位で適用することにより、同一の物理的な網に多様なシステムが收容可能である。これにより、ユーザごとのルーティング、QoS (Quality of Service)<sup>\*11</sup>、セキュリティ、遅延などを最適化し、さらにユーザ間の運用の独立性を担保することができる。

### (2)経路集約効果

セルラシステムでは、AR網内と上位NWであるIPルータ網との間で莫大な経路情報が流通するため、上位NWとの接続装置であるAGRにおいて経路集約を行った上で上位 (IPルータ網) にbroadcastすることで、AGRやIPルータ網が保持する経路情報数を削減している。これにより、NW全体の負荷を下

表1 AR網装置

	AGR	ACR
シャーシ		
ラインカード (100Gモジュール)		

\*6 ラインカード：物理インタフェースを收容するシャーシ型ルータ装置に具備するモジュール。

\*7 LAG：複数の物理回線を仮想的に1つの回線として扱う機能。

\*8 スター型：NWトポロジの一種。複数の通信機器が、1つの中心となる通信機器に接続するNW構成。

\*9 SecGW：基地局との暗号化通信を終端する通信機器。

\*10 VPN：サービスごとに論理的に構成する仮想NW。

\*11 QoS：使用帯域の制御、遅延量や廃棄率などの制御などに用いられるNW上の品質規定。

げられるため障害発生時にも安定的なNW運用が可能となる。

### (3)AR網内でのShaping機能<sup>\*12</sup>

AR網と接続する基地局側では十分な帯域を確保できない(1GbE区間が多い)ため、IPルータ網から流入してくる下りトラフィックに対して、AR網の入口であるAGRにおいてフラグを立て、AR網の出口であるACRでshapingを適用し、下位の伝送路に合わせて送信帯域を調整している。これにより、帯域が十分に確保できない区間に対しても、AR網側で負荷を軽減することで、輻輳<sup>\*13</sup>とそれによるパケット破棄を減らすことが可能となる。

### (4)micro-BFD (Bidirectional Forwarding Detection)

およびBGP-PIC (Border Gateway Protocol-Prefix Independent Convergence) 機能

NW障害の際に前述のように装置を冗長化しておくことと合わせて、以下に説明するmicro-BFDやBGP-PICのような技術を用いてAR網を構成することで、障害検知から経路切替までを高速に完了させることができる。

#### (a)micro-BFD

AGR～AGR区間、AGR～ACR区間、AGR～IPルータ網区間では回線冗長化のため複数の100GbE回線で接続し、LAGを用いて論理的に1つの回線として運用している。本機能はLAGを構成する各物理回線で死活監視パケットを高頻度に送受信するものであり、障害を速やかに検知することで、高速な経路切替を実現する。

#### (b)BGP-PIC

AR網で収容するセルラシステムでは、万単位の経路が流通するため、障害発生時にバックアップ経路を見つけ出す(Lookup)までに秒単位の時間を要してしまう。本機能はバックアップ経路Lookupの演算効率を向上させる機能であり、障害検知から経路切替までのタイムラグを短縮することができる。

<sup>\*12</sup> Shaping機能：帯域制限の方式の一種。データ量が装置に設定するしきい値を超えた際には、バッファして送信する技術。

<sup>\*13</sup> 輻輳：通信の要求が短期間に集中して通信制御サーバの処理能力を超え、通信サービスの提供に支障が発生した状態。

## 2.3 AR網の保守・運用

NE-OpSはNEの管理・監視および制御・設定を行う。多数の拠点に配置するACRやAGRを一括で管理することにより、保守・運用の効率化を実現している。NE-OpSまでの通信に関してもAR網内を通るため、NW区間で障害が発生しても上記のような信頼性で設計され、AGRとACRの管理・監視に問題が生じないような接続構成となっている。

## 3. SRの機能

AR網を構成する上で、主要な機能概要を解説する。特に、AR網において柔軟な転送制御を実現するSRについて、NW設計で用いた機能を中心に解説する。

### 3.1 SRの概要

一般的に、ルータ装置はパケットを転送する際に宛先IPアドレスを用いて転送経路を決定するが、SRでは「Segment」をパケットに付与する事により、網内でのパケットの転送制御を行う。従来の宛先IPアドレスを用いた転送制御よりも、シンプルで柔軟な制御を実現することが可能なルーティング方式である。Segmentには複数の種類があるが、AR網では、主にIGP (Interior Gateway Protocol) -Prefix SegmentとIGP-Adjacency Segmentを使用する。SRの転送動作概要を図2に示す。

#### (1)Prefix-SID (IGP-Prefix Segment ID)

AR網内でユニークな値であり、基本的な転送制御ではこのPrefix-SIDを使用する。単一のノードで使用されるIGP-Prefix (Loopbackアドレス<sup>\*14</sup>など)に対応するPrefix-SIDはNode (IGP-Node Segment) -SIDと呼ばれ、SR対応ルータ装置は、これを用いてドメイン内で一意にノードを指定することができる。

<sup>\*14</sup> Loopbackアドレス：ノードなどに割り当てられる自身を示すIPアドレス。

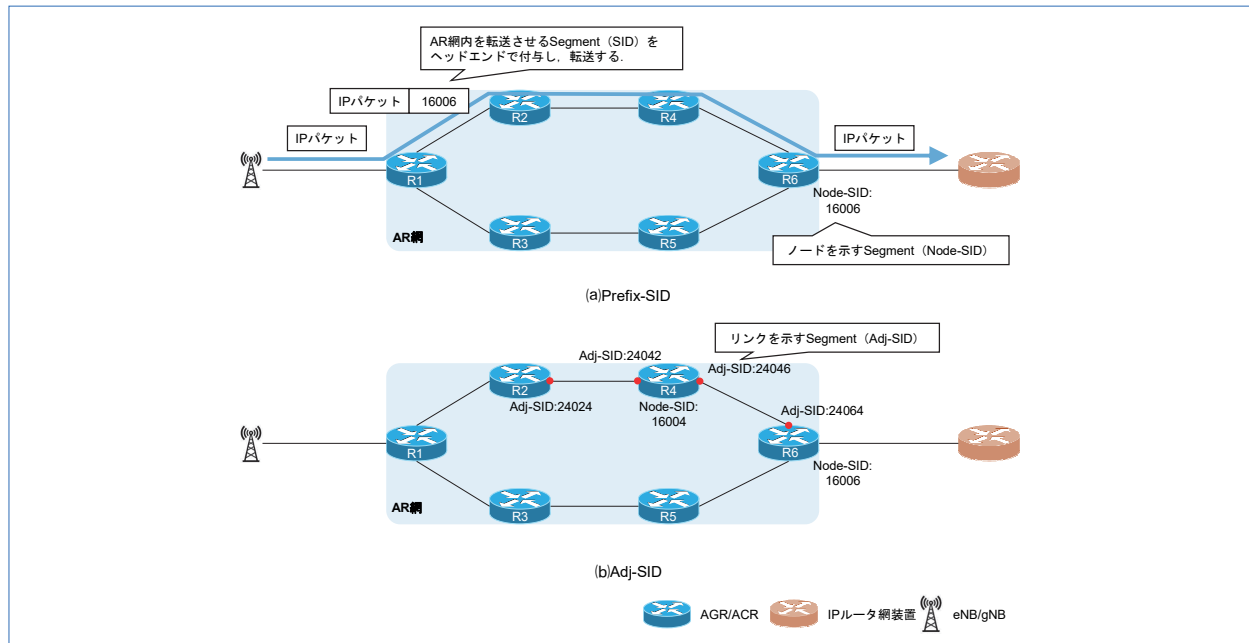


図2 SRの転送動作概要

## (2) Adj (IGP-Adjacency Segment) -SID

各SRノード内でユニークな値であり、隣接IGPノードへのリンクごとに割り当てられ、特殊な転送制御で使用される。Adj-SIDを受け取ったノードは、宛先ノードへの最短経路とは関係なく、指定されたリンク向けにパケットを転送する（送信先インタフェースを指定する）。

## 3.2 SRの特徴

AR網では、SRをMPLS (Multi-Protocol Label Switching)\*<sup>15</sup>データプレーン\*<sup>16</sup>に適用するSR-MPLS方式を採用しており、以下のような特徴がある。

### (1) データプレーン

本方式では、従来のMPLSデータプレーンをそのまま利用可能である。また、MPLSヘッダのLabelフィールドにSIDをエンコードすることでSegmentを表現する。

### (2) コントロールプレーン\*<sup>17</sup>

以下の3種類の実装方法から選択可能である。

- ・ distributed (分散型) は、ルーティングプロトコル (OSPF (Open Shortest Path First)\*<sup>18</sup>など) の機能を拡張することで対応する方式。ラベル配布専用のプロトコル (LDP (Label Distribution Protocol)\*<sup>19</sup>, RSVP (Resource reSerVation Protocol)\*<sup>20</sup>) が不要であり、従来のMPLSと比較してシンプル (AR網では本方式をOSPFで採用)。
- ・ centralized (集中型) は、上位のコントローラがSegment生成、割当てを含むすべてのコントロールプレーンを担う方式。
- ・ hybrid (ハイブリッド型) は、ルーティングプロトコル (OSPFなど) を用いた分散型を基本とし、コントローラを用いた集中型を併用する方式。

### (3) ソースルーティング

本方式では、SR-policy\*<sup>21</sup>をSR-MPLS網の入口で適用することで、網の出口までの経路をすべて指定可能である。また、SDN (Software Defined

\*<sup>15</sup> MPLS: IPアドレスの代わりにラベルを用いたパケット転送技術。

\*<sup>16</sup> データプレーン: ルータでトラフィック転送する際に使用するルーティングテーブルやそのロジック。

\*<sup>17</sup> コントロールプレーン: ルータでトラフィック転送を制御するためのルーティングプロトコル。

\*<sup>18</sup> OSPF: ルータが隣接している接続情報を基に経路を選択するプロトコル。

\*<sup>19</sup> LDP: ラベル交換プロトコルの一種。ラベル情報交換により、NE間のセッションを確立することが可能となる。

Network)<sup>\*22</sup>との親和性が高く（コントローラによる集中型制御をしやすい）、低遅延や帯域保証などユーザごとに最適化した経路制御を提供できるアーキテクチャとなっている。

#### (4)高速切替機能

AR網は、セルラ通信の制御プレーン<sup>\*23</sup>や音声通

話サービスを収容するため、高水準の可用性が要求される。AR網では、あらゆる障害に対しても速やかに経路切替を行い、トラフィック影響を最小限に留められるような設計が求められる。図3に示すように、従来のMPLSでは、トポロジによってはパケットが迂回した後も障害発生箇所に向かってしま

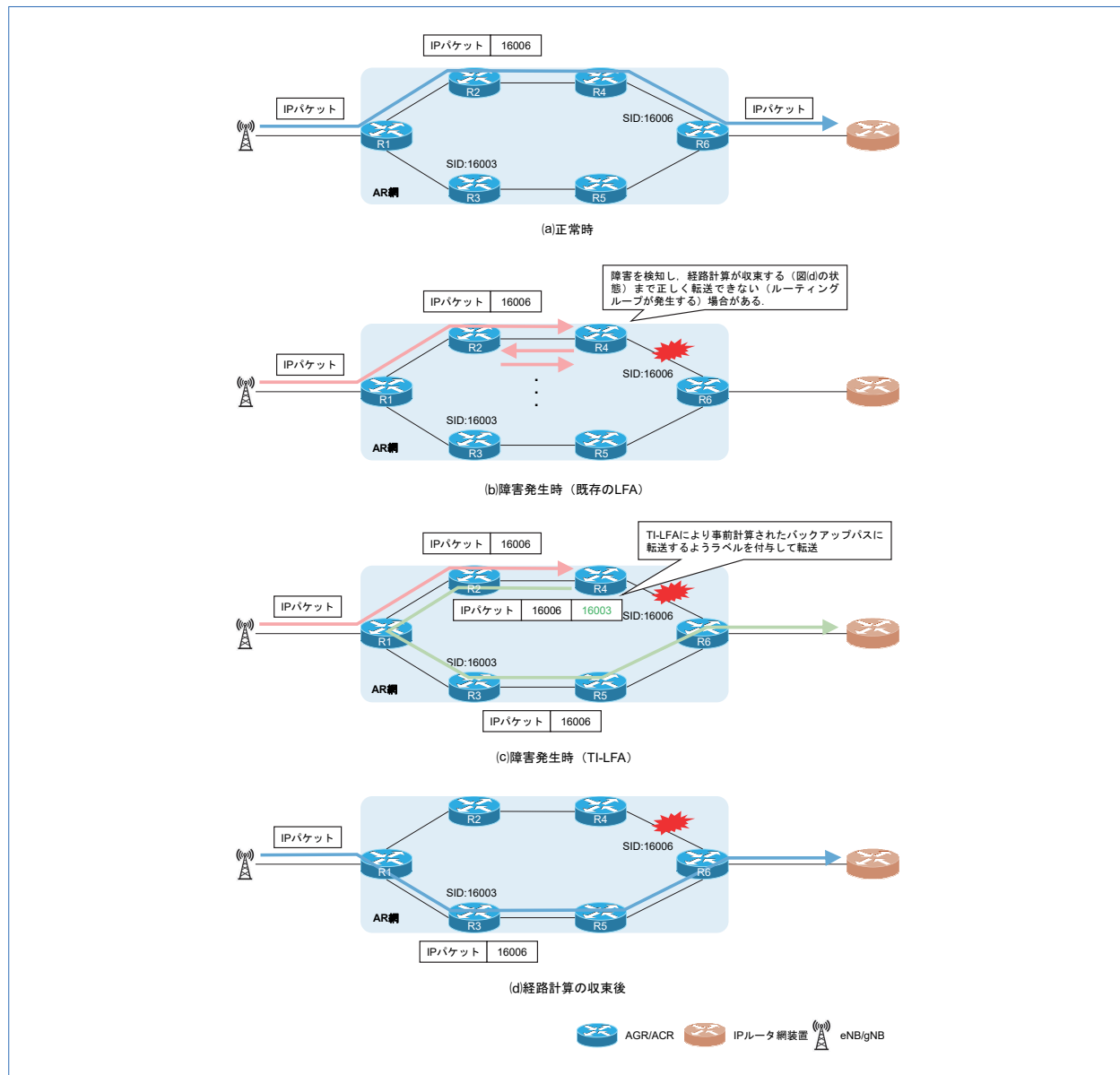


図3 AR網内の障害時動作とTI-LFAの動作概要

\*20 RSVP：IP-NW上でデータの送信先までの帯域幅をあらかじめ予約することによって、即時性が要求されるデータ通信などで必要な通信品質を確保するためのプロトコル。

\*21 SR-policy：SR網内の転送経路を指定するポリシー。

\*22 SDN：NWを構成する機器の設定・動作をソフトウェアで制御すること。

\*23 制御プレーン：セルラ通信の確立や切断などをするための制御信号を転送するためのプロトコル。

いループが発生してしまうことが懸念されていたが(図3(b)), SR-MPLSでは転送バックアップの経路をMPLSラベルスタックによって厳密に指定することで、ループを回避可能である(図3(c)(d)). これはTI-LFA (Topology Independent Loop Free Alternate) と呼ばれ、IGP (OSPF) 網内の障害に対しては、本機能によって高速な経路切替を実現する。従来のIP高速迂回技術であるLFAやrLFA (remote LFA) は、トポロジによってはループフリーな迂回経路を準備できないという課題があったが、SRを用いることでループフリーな迂回経路を提供することが可能となる。

#### (5)OAM (Operations And Maintenance)\*24機能

SR対応ルータ装置は保守機能として、いくつかのOAM機能を有する。

まずは、基本的なOAM機能として、Ping\*25, traceroute\*26をIGP網内などの疎通確認に使用することが可能であり、隣接ノードへの到達性や転送経

路を確認できる。

次に、AR網内でVPN収容しているユーザのVPN内の疎通確認では、VRF (Virtual Routing and Forwarding) -ping/tracerouteを使用することができる。本コマンドは、AR網に開通されるVPN端点をsource (送信元) として、VPN内の任意のdestination (宛先) への到達性や転送経路を確認する際に使用される。destinationはAR網内でなくても良く、基地局やコアノードまでの到達性を確認することもできる。

さらに、MPLS-Pingを用いてIGP網内などの疎通確認をMPLSデータプレーン上で実施することもできる。SID情報の交換に失敗しているなど、特殊なトラブルが発生していないかを確認するために活用する。

SR対応ルータ装置の特徴として、サイレント障害\*27を検知する機能であるDPM (Data Plane Monitoring) が挙げられる。図4に示すように、DPMにより、NWの正常性を定期的に確認することができる。また、自ノード経由のデータプレーン通信に対する

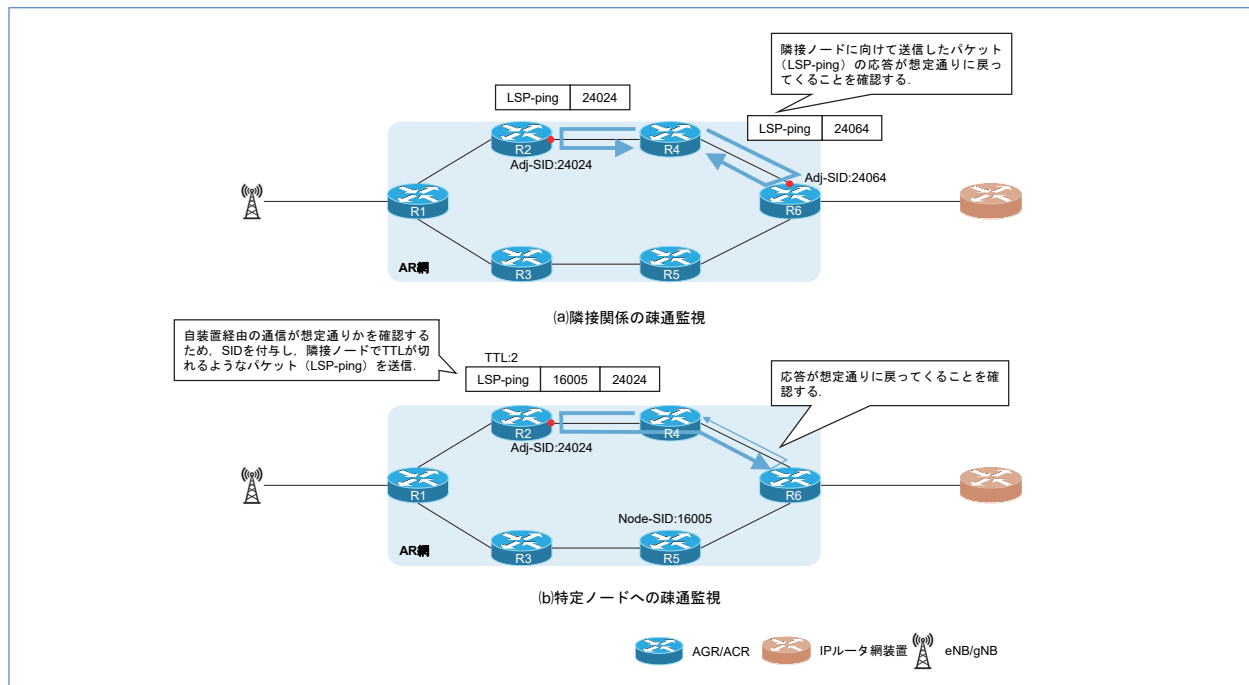


図4 DPMの動作概要

\*24 OAM : NWにおける保守運用管理機能.

\*25 Ping : IP-NW上でノードへの到達性を確認するコマンド.

\*26 traceroute : IP-NW上でノードへのネットワーク経路を調べるコマンド.

\*27 サイレント障害 : NE-OpSで検知できない障害のこと.

正常性も確認することができるため、サイレント障害発生時の検知、切分け時に効果が期待できる。

## 4. 5G時代におけるMBHの進化

### 4.1 SR-TEの概要

AR網内の経路を明示的に指定することのできるTE (Traffic Engineering) にSR技術を適用することにより、従来よりもサービスに応じて柔軟な経路制御を実現することができる。MBHの間でも同様に、SRを用いたTEにより、きめ細やかなNW制御がユーザに提供可能となる。なお、モバイルNWでは、ユーザ端末からコアノードまでの通信がTEの対象となる。

5Gの導入により、これまでよりも低遅延、大容量、多数端末同時接続といった多様なサービスを収容しなければならない。これらのサービスを提供す

る各NWの特長を考慮した柔軟なNW設計が、従来よりも求められる。

SRによるTEの制御概要を図5に示す。SR-TEでは前述のソースルーティングによりパケットの転送経路が決定されるため、AR網の入口でMBH区間を通る経路を明示的に指定することができる。MBH内にパケットが流入したタイミングで、サービスごとに付与された識別子を用いて転送する経路を振り分けし、論理的に分けられたNWとして運用することが可能となる。サービスを振り分ける識別子として、IPアドレスやポート番号、DSCP (Differentiated Services Code Point) 値<sup>\*28</sup>などを使い、接続するノードに合わせた制御方式を実現することができる。

例えば、低遅延サービスを提供する場合は、ノード間の遅延量を用いて転送経路を計算し、各NW装置へ計算結果である最短経路を設定することが可能となる。OSPFなどのパスコスト<sup>\*29</sup>を用いた経路計

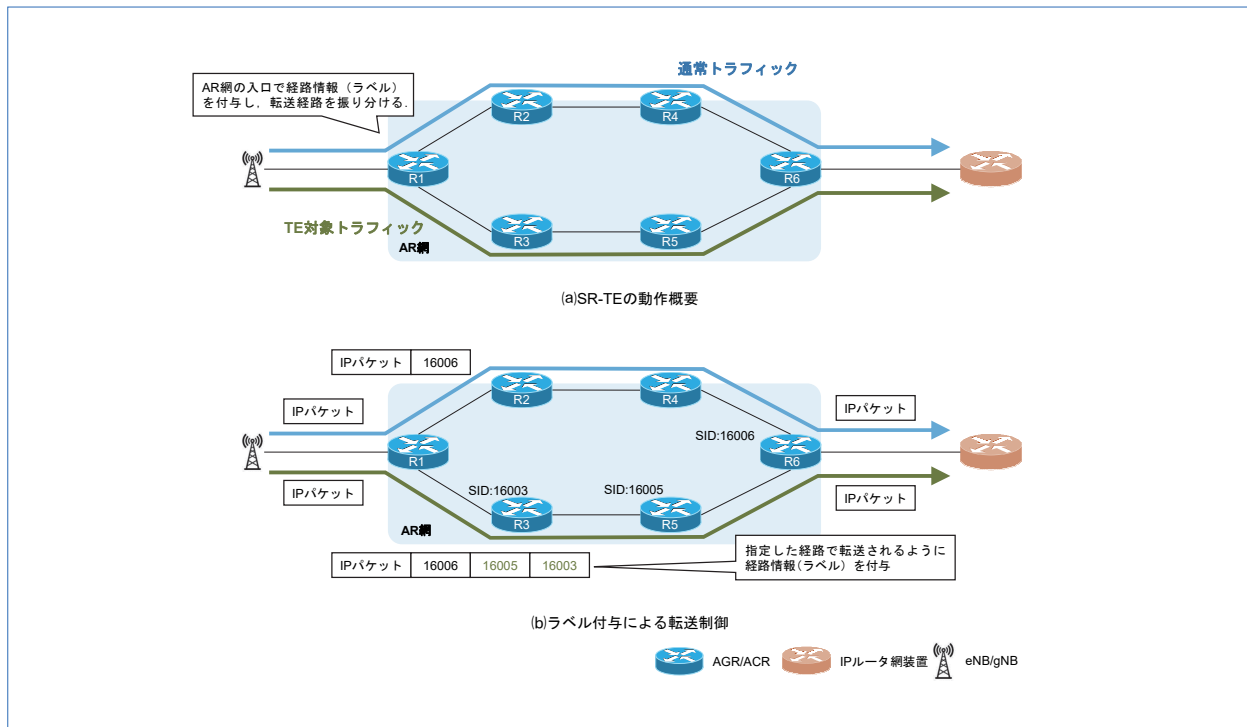


図5 SR-TEの制御概要

\*28 DSCP値：パケットの優先制御を行う際の、優先度を表す値。  
\*29 パスコスト：宛先までの経路上の距離（重みづけ）を累積した値。

算で決められた経路と、遅延量を用いた経路計算で決められた経路を使い分けることで従来よりも柔軟なMBH区間を構築できる。

## 4.2 E2EでのNWスライシングの実現に向けて

上位コントローラを介したNW設計にすることで、MBH区間だけでなくMBHに接続するNWを含めた経路最適化が可能となり、E2E（End to End）での経路制御の実現が見込まれる。E2Eでのスライシングにより、これまでよりもサービス単位での要求条件に沿ったNWをユーザに提供できることが期待される。

上位コントローラを介したNW構成を図6に示す。SR対応ルータ装置だけではAR網区間以外のNW構成を把握することができないため、接続先のNW構成を接続先NW用上位コントローラ経由で取得することが必要となる。経路計算についても、AR網区間のルータおよびAR網用上位コントローラだけで実施するのではなく、接続先NW用上位コントローラ

側と連携して経路制御に必要な他NW構成情報を互いに受け渡し、経路計算を実施する。

このように各コントローラ間を連携させ、各コントローラ配下のルータにそれぞれ設定することで、E2Eで一貫した経路制御ポリシーを適用することが可能となり、多様なサービス要件に応じて論理的なNWを構成できる。その際にも今回導入したSR対応ルータ装置を用いることで、コントローラとも連携した経路制御が容易となり、多様なサービスに対して柔軟なNW設計を提供することが可能となる。

## 5. あとがき

本稿では、5G時代に対応し得るSR対応ルータ装置の概要、そのNW構成、今後実現するMBHの進化について解説した。

これまで述べてきたように、SR対応ルータ装置の導入により、NW容量の大幅な増加、効率的なNWの帯域利用が可能となり、また障害時において

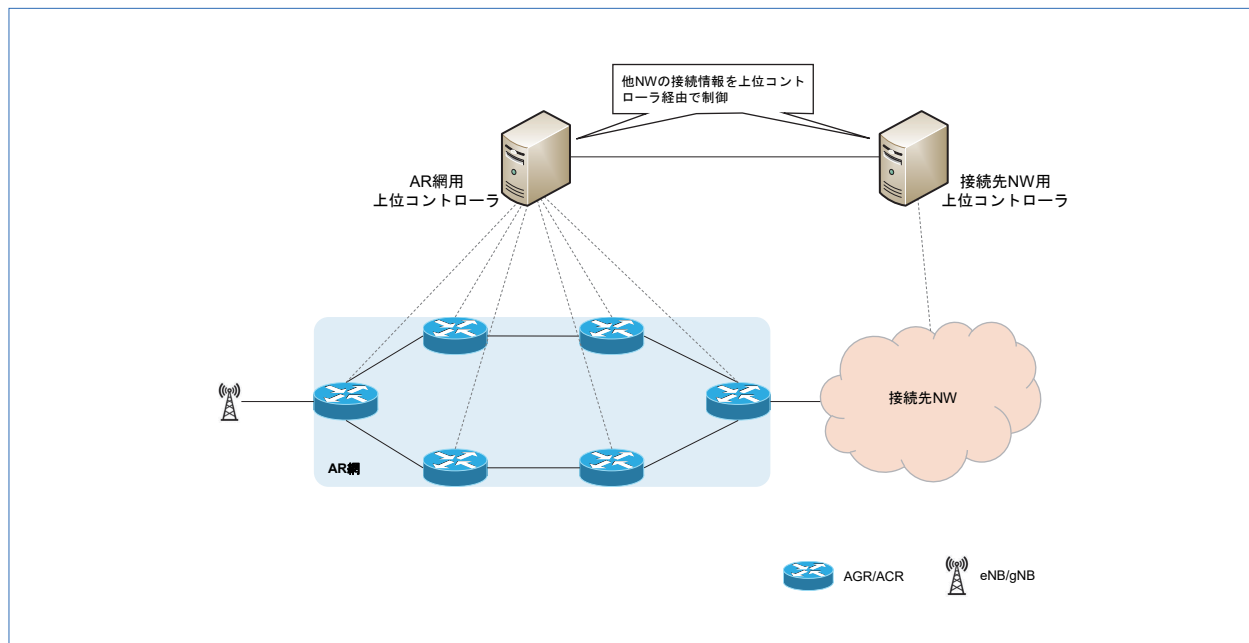


図6 上位コントローラを介したNW構成



も高速切替機能や障害検知の向上によりNWの信頼性を高めることができる。

今後は上位コントローラとの連携により、多種多様な5Gサービスに柔軟に対応するための機能拡張を引き続き検討していく。

## 文 献

- [1] 森田, ほか: “伝送路のAll-IP化に向けたイーサネット伝送装置 (ERP-SW) の開発,” 本誌, Vol.18, No.3, pp.52-56, Oct. 2010.
- [2] 大崎, ほか: “「IPルータ網」の構成技術 - L2/L3統合VPNバックボーンの概要 -,” 本誌, Vol.12, No.1, pp.57-63, Apr. 2004.

# ネットワーク仮想化基盤における ETSI NFV Stage3仕様に準拠した マルチベンダ対応MANOへの移行

ネットワーク開発部 ネットワーク仮想化  
ドコモ・テクノロジー株式会社

たむら ひろなお く の ゆうや  
田村 宏直 久野 友也  
すずき よしゆき  
鈴木 吉行

ドコモは、2015年度末にコアネットワークを対象とした仮想化基盤システムを商用導入し、現在まで安定した運用を続けてきている。この仮想化基盤システムはマルチベンダで構成されており、また多くの仮想化されたコアネットワーク装置を収容している。このたび、仮想化基盤システムを構成するNFVOとVNFMのEoLを迎えるにあたり、最新のETSI NFV Stage3仕様を取り込んだマルチベンダ構成の仮想化基盤システムを実現した。この際、通信サービスを停止することのない無中断アップグレードによるシステム移行を完遂した。本稿では、その取組みについて解説する。

## 1. まえがき

ドコモは、2015年度末にコアネットワーク<sup>\*1</sup>を収容する仮想化基盤システムを、世界で初めてマルチベンダで構成し、通信ネットワークのCAPEX (CAPital EXpenditure)<sup>\*2</sup>/OPEX (OPerating EX-pense)<sup>\*3</sup>の低減を実現してきた [1] [2]。その後も多くのコアネットワーク装置の仮想化を推進するとともに、安定運用を続けてきている。コアネットワーク装置の仮想化適用率は2020年度末で約50%を

超え、400以上の仮想化されたコアネットワーク装置が運用されている。これらの仮想化されたコアネットワーク装置 (VNF: Virtual Network Function<sup>\*4</sup>) のライフサイクル制御を担うのがMANO (Management ANd Orchestration)<sup>\*5</sup>である。ライフサイクル制御には、ハードウェア障害や仮想マシン (VM: Virtual Machine)<sup>\*6</sup>障害が発生した際に、正常なハードウェア上にVMを移動するヒーリング制御、ハードウェアやVMの負荷状況に応じて、通信ソフトウェアとVMを増減することで処理能力

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

\*1 コアネットワーク：交換機、加入者情報管理装置などで構成されるネットワーク。移動端末は無線アクセスネットワークを経由してコアネットワークとの通信を行う。

\*2 CAPEX：ハードウェアと、ハードウェア設置工事にかかる費用。

\*3 OPEX：設備を維持し運用するために発生する費用。

\*4 VNF：仮想化された通信機能（通信システム）。

\*5 MANO：NFVO（\*16参照）、VNFM（\*17参照）、VIM（\*19参照）の機能ブロックをまとめた総称。

を最適化するスケーリング制御，サービス追加や不具合対処のための通信ソフトとVMのソフトウェアを更新するVNFアップグレード制御などがある。安定した通信サービスを提供するために，本仮想化基盤システムはミッションクリティカル<sup>\*7</sup>なシステムである（図1）。

さて，ドコモの仮想化基盤システムは，前述のとおり，複数ベンダのプログラムの組合せで構成されているが，これらプログラムはEoL（End of Life）<sup>\*8</sup>のタイミングやサポートする標準仕様のバージョン，さらには利用するAPI（Application Programming Interface）<sup>\*9</sup>やそのInput/Output Parameter<sup>\*10</sup>，Attribute<sup>\*11</sup>が異なる。これらについては，標準化動向やベンダのリリース計画，ベンダが製品提供している各オペレータからの要件などから，各ベンダが自ら決定している。よって，複数ベンダのプログラムの組合せで実現されているシステムのバージョンアップにおいては，プログラム間のインタフェース（IF）仕様の設計を始めとするインテグレーション<sup>\*12</sup>

の複雑化が課題として挙げられる。一方で，ドコモはETSI（European Telecommunications Standards Institute）<sup>\*13</sup> NFV（Network Functions Virtualisation）<sup>\*14</sup> [3]における標準化活動を積極的に展開し，Stage3仕様<sup>\*15</sup> [4]の策定に貢献してきた。現在もマルチベンダ製品間のインテグレーション，信頼性，運用許容性を考慮した標準化活動を行っている。

本稿では，プログラムのEoLを契機に，プログラム間のIF仕様をETSI NFV Stage3に準拠すべく，システム間IF仕様を調整した取組みについて解説する。また，通信オペレータならではの要件である，通信サービスを停止することのないオンラインアップグレードの手法について解説する。なお，アップグレードするプログラムはMANOを構成するNFVO（NFV Orchestrator）<sup>\*16</sup>とVNFM（VNF Manager）<sup>\*17</sup>であり，クラウド<sup>\*18</sup>環境として仮想リソースを管理するVIM（Virtual Infrastructure Manager）<sup>\*19</sup>は対象外のため，VIMの運用継続を条件とした。

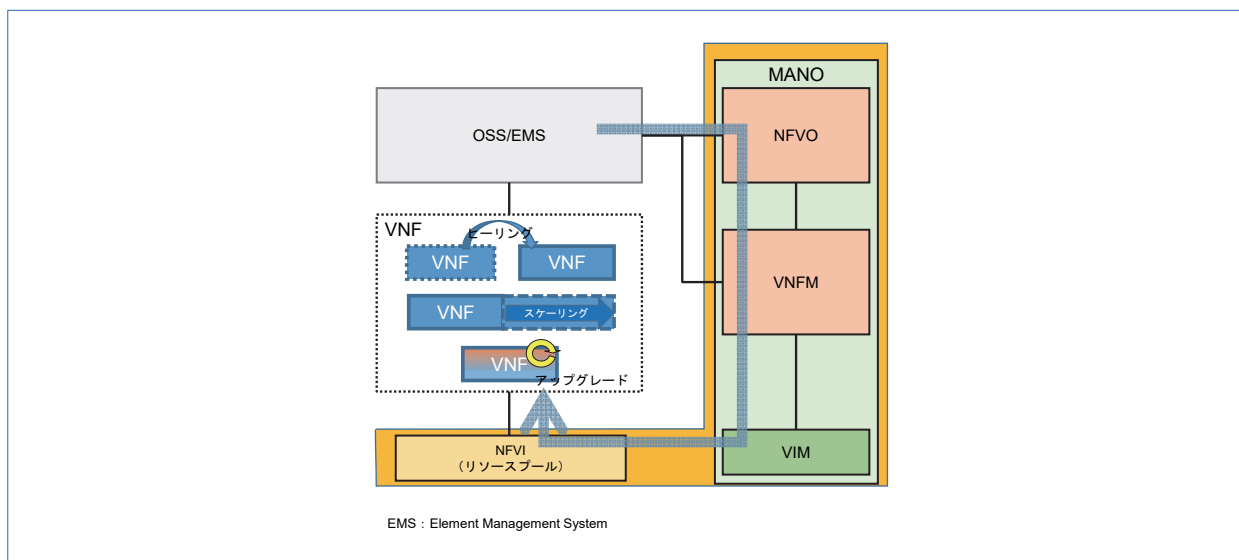


図1 仮想化基盤システムにおけるヒーリング，スケーリング，VNFアップグレード制御

<sup>\*6</sup> 仮想マシン（VM）：ソフトウェアによって仮想的に構築されたコンピュータ（仮想マシン）。  
<sup>\*7</sup> ミッションクリティカル：業務やサービスを遂行するために必要不可欠であり，その停止や中断が許容されないような重要な要素のこと。  
<sup>\*8</sup> EoL：製品の生産／販売終了，ソフトウェア製品のサポートサービス終了，バグや機能改善の修正・更新プログラムの提供終了を指す。  
<sup>\*9</sup> API：ソフトウェアコンポーネントやプログラム同士を接続するために情報をやり取りするための仕様。

<sup>\*10</sup> Parameter：ETSI NFVにおけるシステム間でやり取りするデータ群の総称。APIを利用するときに，APIを呼び出すシステムが渡すデータ群をInput Parameter，APIを開放しAPIを実行したシステムがその実行結果を返すデータ群をOutput Parameterと呼ぶ。  
<sup>\*11</sup> Attribute：ETSI NFVにおける各システムが保持しているリソース情報の総称。ETSI NFVの仕様はREST（<sup>\*36</sup>参照）ベースのため，各装置が保持しているリソースをAttributeと呼び，APIにおけるInput/Output Parameterと区別している。

## 2. 仮想化基盤システムを取り巻く状況

### 2.1 仮想化基盤システムの動向

ドコモは、これまでETSI NFVにおける標準化活動を積極的に展開し、NFVの普及に貢献してきた。NFVが世界的に普及するとともに、仮想化基盤システムは多くの通信オペレータに導入されてきているが、ドコモのようなマルチベンダ構成による導入や、単一ベンダ構成による導入など、オペレータによって導入形態は異なる。さらには、収容するVNFのアーキテクチャも異なるため、サポートするVNFのライフサイクルのユースケースや実現方式も異なるのが現状である。これは、オペレータによるETSI NFV仕様の適用方法や、仮想化基盤システムが準拠すべきIFの範囲が異なることを意味する。

### 2.2 マルチベンダ環境におけるグローバルプロダクトの活用

近年、インフラ装置へのクラウド技術の適用によって、オペレータの開発手法は大きく変化している。クラウド技術の進化は早く、プロダクトのライフサイクルも早くなっているため、すべてを自社で開発することは非効率となってきている。一方で、仮想化基盤システムのような保守系装置は、国ごとに遵守すべき法制度や考慮すべき地理的条件、各オペレータの組織構造や運用フローなどが異なるため、自社で開発したシステム以外を利用することは困難であった。

そのため、図2のとおり、自社の業務分析をベースとした一般的な開発工程（ウォーターフォール型開発の例）から、(1)背景や運用方法が異なるオペレータそれぞれの業務分析をベースとしたオペレー

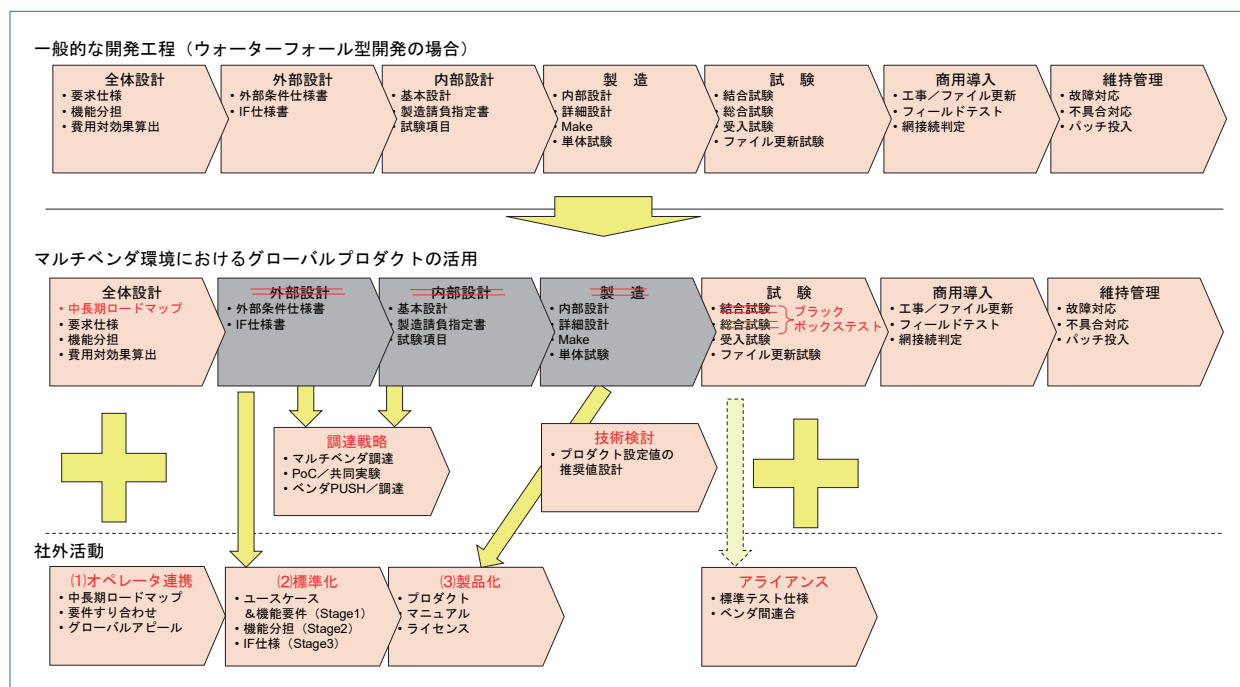


図2 マルチベンダ環境におけるグローバルプロダクトの活用

- \*12 インテグレーション：装置、またはシステムを、オペレータが運用しているネットワークに組み込むこと。
- \*13 ETSI：欧州電気通信標準化機構。欧州の電気通信技術に関する標準化団体。
- \*14 NFV：ネットワーク仮想化。通信キャリアのネットワークを仮想化技術により汎用ハードウェア上で実現すること。
- \*15 Stage3仕様：ITU-T（International Telecommunication Union-Telecommunication sector）I.130で規定されたテレコムにおける標準化手法のうち、信号のプロトコルを規定した仕様の総称。

- \*16 NFVO：複数のVIM（\*19参照）をまたがる仮想リソースの統合的な管理システム。
- \*17 VNFM：VNFのライフサイクル制御として、起動や停止などVNFの制御を担うシステム。
- \*18 クラウド：ネットワーク経由でサービスを利用する形態、仕組み。利用状況に応じたサーバリソースの分配が可能のため、スケラビリティが高い。
- \*19 VIM：仮想化基盤の物理コンピュータ、物理ストレージ、物理ネットワークの各資源を管理するシステム。

タ連携や、(2)標準化による社外活動を通じた要求条件の共通化、共通的な仕様策定が極めて重要になり、さらに(3)その要求条件に対するオペレータのニーズを可視化し、多くのベンダによる標準仕様に準拠した製品化を推進していく必要がある。

これらの活動により、多くのオペレータが同一仕様のプロダクト（グローバルプロダクト）を採用していくことで、マルチベンダ環境における相互接続性を維持しつつ、低コスト化を狙うことが可能となる。

#### (1)オペレータ連携

仕様の共通化のため、オペレータ間で要件の共通化が必要になる。そのためにも、オペレータ各社が自社の中長期のロードマップを作成し、要件をすり合わせることで、要件の共通化とその発信（グローバルアピール）が重要となる。

#### (2)標準化

オペレータの共通要件を基に、共同でユースケースの作成が行われる。ユースケースを明確化することで、機能要件（Stage1）が明確になる。その後、機能分担とインフォメーションフロー（Stage2）、IF仕様（Stage3）と続く。オペレータとベンダで共通的な標準仕様を作成することで、マルチベンダ環境でもグローバルプロダクトが接続できるようになる。

#### (3)標準仕様に準拠した製品化の促進

オペレータによる標準仕様に準拠した製品の調達と、それに伴うベンダの製品化（プロダクト製造、マニュアル作成、ライセンス化）が行われる。

このように、グローバルプロダクトを利用した仮想化基盤の開発においても、一般的な開発と同じく、機能分担やIF仕様書制定のような設計、プロダクトを作る製造、そして要件通りの動作を確認する試験というような必要となる工程に大きな差は無い。

しかし、共通的な要件と仕様を利用するために、社外活動が必要となる。

## 3. ドコモのMANOアップグレードへのアプローチ

仮想化基盤システムは、前述のとおりクラウド技術を採用した各プロダクトを接続させるための活動が必要となるため、ドコモはETSI NFVを中心とした標準化活動と標準仕様に準拠した製品化の促進に取り組みつつ、マルチベンダMANOのIF仕様調整を行い、さらにはオンラインでのアップグレードを実現した。

### 3.1 デファクトバージョンの見極めと製品化の促進

通信系装置は、10年という単位でサービス提供し続ける必要があるが、前述のとおり仮想化基盤はクラウドの技術を取り入れたことで、プロダクトのライフサイクルが短くなる傾向があり、また機能分担が頻繁に変更されるようになった。ETSI NFVは、クラウド上に通信系装置を載せることを目標としているためクラウドの文化に近づいてきており、その標準仕様は半年に一度マイナーバージョン、2年に一度メジャーバージョンが上がり、そのたびに機能追加や機能分担の変更が起きる。

そのため、複数のプロダクトを接続するときには、前述の(1)オペレータ連携や(2)標準化で、自社のユースケース、要求条件を標準化している前提で、どの仕様のどのバージョンをサポートしているかと、そのバージョンのサポート期間を見極める必要がある。また、各バージョンによってサポートされている機能や、セキュリティ関連仕様が変わっていることもあるため、自社のユースケースや環境に合致した

バージョンをベンダに要求していく必要もある。

これらを踏まえてデファクトとなるバージョンを見極めつつ、前述の(3)標準化に準拠した製品化の促進を考え、オープンな場でデファクトとなるバージョンを啓発したり、各プロダクトベンダに要求したり、オペレータ間で要求機能の足並みを揃えたりしていくことが重要である。これにより、各プロダクトベンダにいつまでにどのバージョンのどのような製品が必要かというニーズを理解して貰い、プロダクト化のインセンティブを高めることが重要である。

### 3.2 バージョンアップ開発の着手に向けた実現手法の検討

ETSI NFV Stage3に準拠するIFを図3に示す。ベンダのプロダクトは、基本的にマーケットニーズに基づいて製品化されるため、ローミング\*20をサポートする通信装置のように標準仕様が実装されないと通信サービスが成立しない装置と異なり、OSS (Operating Support System)\*21/EM (Element

Manager)\*22やMANOのような保守系装置は、標準仕様を必ずしもすべて実装しない。保守系装置の場合、オペレータの運用方法によって必要となる機能が全く異なるため、多くのオペレータが要望する共通的な機能やIF仕様のみを、ベンダは自プロダクトに実装しようとする。その結果、MANOの製品間を接続しようとする、標準仕様として規定されているIF仕様であっても、各製品がサポートしているAPIの種類やその中のAttribute、さらにはそのAttributeが取り得るValue\*23の種類が異なるため、MANO製品間を接続するための検討や作業が必要となる。

標準仕様で確実に接続するためには、以下(1)~(3)のとおり、自社のユースケースを早急に完成させ、新バージョンの装置導入のために最低限必要な標準仕様に準じた機能を明確化してベンダに要求する必要がある。また、自社のユースケースを実現するために、安易にプロダクトにおいて標準仕様以外の機能やExtension\*24の実装や利用を認めると、安定的

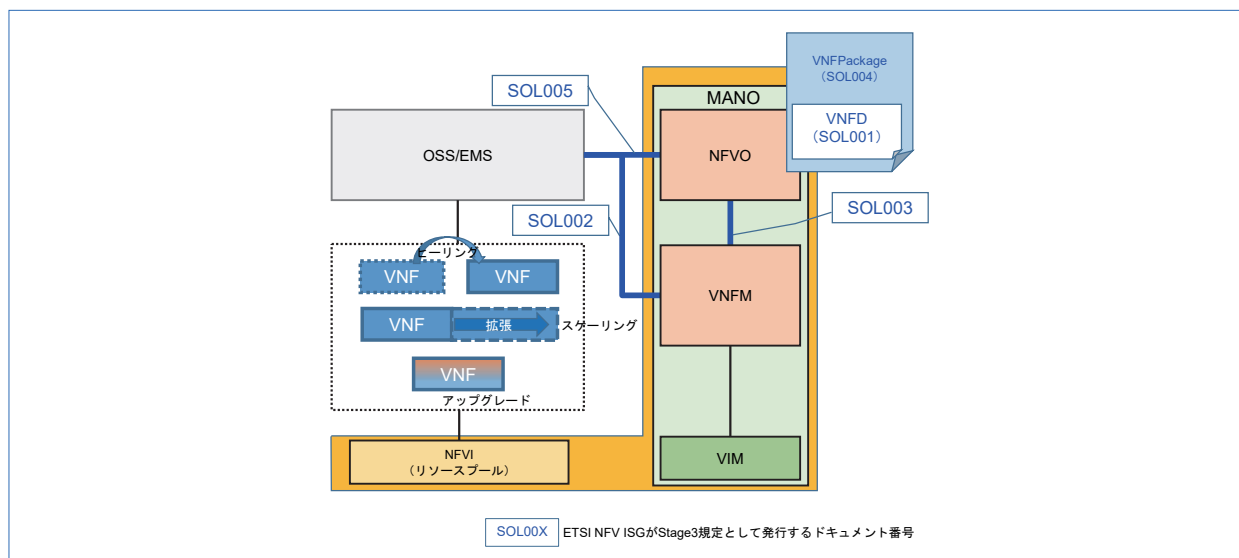


図3 ETSI NFV Stage3 IF

\*20 ローミング：利用者が契約している通信事業者のサービスエリア外でも、提携事業者のサービスエリア内であれば、契約している事業者と同様のサービスを利用できる仕組み。

\*21 OSS：移動通信網で発生している故障や輻輳の発見と、それに対する制御・措置を行っている事業者の運用支援システム。通信事業者の場合、提供しているサービスを利用するために、ネットワークやシステムの「障害管理」「構成管理」「課金管理」「性能管理」「セキュリティ管理」のすべて、もしくは一部を行う。

\*22 EM：個々の通信装置に対するFCAPS (Fault, Configuration, Accounting, Performance and Security) の管理・監視を担うシステム。

\*23 Value：Attributeによって受け渡される具体的な値。REST (\*36参照) においてはAttributeとValueのKey Value Pairによって受け渡される。

\*24 Extension：プロダクトが独自の機能を利用するために拡張されたAPIやVNFD (\*35参照) のパラメータ。

なアップグレードや接続可能なバージョンの選定を阻害したり、試験工程や商用運用時の問題発生の際に解決が困難になったりするため、プロダクト間の接続条件において標準仕様以外での接続は認めないという強い方針が大切である。バージョンアップ開発に向けて、あらかじめ実現手法として検討すべきことは以下のとおりである。

#### (1)新バージョンにおけるプロダクト間シーケンス<sup>\*25</sup>の作成と機能分担の決定

自社のユースケースを基に、標準仕様を利用したシーケンスの検討が必要になる。シーケンスの種類は大きく分けると①デプロイ<sup>\*26</sup>（オンボーディング<sup>\*27</sup>とインスタンス化）、②更新（ヒーリング制御・スケーリング・ファイル更新）、③撤去（ターミネーション<sup>\*28</sup>）が必要となる。これらを実現するために、VNF Instance<sup>\*29</sup>、Virtual Resource<sup>\*30</sup>、VNF Package<sup>\*31</sup>のリソースオーナー<sup>\*32</sup>を担うプロダクトを決定し、シーケンスの作成が必要となる。

#### (2)機能分担に基づいたError-handling<sup>\*33</sup>方式と手動復旧手順の決定

ETSI GS NFV-SOL 003の仕様にはError-handlingの方針と仕様が記載されているが、サポートしている機能と実装依存のError Caseはしっかりまとめていく必要がある。特に、通信NG時の検知方法と復旧方法、ユースケースごとにRetry方式で復旧するかRollback方式で元に戻すか、VNF LCM (LifeCycle Management)<sup>\*34</sup>が失敗したときのError-codeと復旧方式の確認は必要である。

#### (3)利用するParameter値の決定

ETSI NFVの標準化では、さまざまなユースケースに対応できるように多数のParameterとAttributeが定義されており、インフォメーションフローのオプションも多数用意されている。そのため、(1)のシーケンスおよび(2)の採用したError-handling方式に基

づいて、主要Parameterがどのように複数システムで交換されていくのか、VNFD (VNF Descriptor)<sup>\*35</sup>に何を記載すべきか、またユースケース実現のために必要なParameterとAttributeはどれかを1つずつ確認する。特に、VNFとネットワークの接続情報、信頼性維持のための冗長構成を組むVMの配置方針、仮想リソース制御で利用するVIMやVNFの識別子、これらが各装置で一意に判断できるためのネーミングルールなどを決めていく必要がある。

これらの検討により、シーケンス、Error-handling、各Parameter値などをしっかりドキュメント化し、各製品の接続条件としてそれぞれのベンダに要求することが大切である。なお、ドコモは次回以降のアップグレードが円滑に実現できるように、シーケンス (Procedure)、装置間のParameter交換 (Key information exchange)、REST (REpresentational State Transfer)<sup>\*36</sup>環境における競合条件 (dependent and non-dependent side procedure)、主要なエラーケース (Error-case) をETSI GS NFV-SOL 016として標準化することに貢献した。

### 3.3 オンラインアップグレード

ドコモの仮想化基盤システムでは、400以上のコアネットワーク装置のVNFが運用されており、これらの装置への悪影響は、ユーザの通信が繋がりにくくなったり、停止したりといったユーザサービスの低下に直結する。そのため、NFVO、VNFMのアップグレードにあたっては、運用中のVNFが提供している通信サービスへ影響を与えないことが求められる。

通信サービスへ影響を与えずに、NFVO、VNFMのアップグレードを行うためには、以下の要件が存在する。

\*25 シーケンス：システム間や機能ブロック間などで取り決める処理の順序。

\*26 デプロイ：アプリケーションをそれらの実行環境に配置して展開すること。

\*27 オンボーディング：VNF PackageをMANOシステムへ登録すること。

\*28 ターミネーション：VNFを停止し、仮想化基盤上から削除すること。

\*29 VNF Instance：MANOシステムによって仮想化基盤上に構築（インスタンス化）されたVNFの実体。

\*30 Virtual Resource：仮想化されたハードウェアリソース（CPU、メモリ、ハードディスクなど）。

\*31 VNF Package：VNFの機能、動作を定義したVNFD（\*35参照）、付帯情報、VNFイメージをパッケージ化したもの。

\*32 リソースオーナー：VNF Instance、Virtual Resource、Packageのリソースの管理元システム。

\*33 Error-handling：エラー発生時の処理方法。

\*34 VNF LCM：VNFのライフサイクル処理。

\*35 VNFD：仮想化基盤上に構築するVNFの定義ファイル。

## ①VNFの再構築が不要であること

アップグレードの1つの方法として、旧システムとは別に構築した新システム上で、新たなVNFを構築（インスタンスエーション）し、VNFを切り替える方法が考えられる。しかし、コアネットワーク装置によっては、VNFの切替え時に通信サービス影響が発生すること、400以上のVNFの再構築、切替えには、膨大な期間とコストが必要となることから、VNFの再構築を不要とする必要がある。

## ②運用中のVNFに対するライフサイクルが実行可能であること（一定時間の停止は許容）

仮想化基盤システムでは、ライフサイクルの1つであるヒーリング制御によって、ハードウェア障害時の信頼性を向上させている。具体的には、ハードウェア障害時にVNFを構成するVMを別の正常なハードウェア上に再生成し、冗長性を維持する。ヒーリング制御ができない場合、VNFの冗長性が維持できず、次のハードウェア障害発生によりVNFの通信サービスに影響が出るリスクが高くなる。そのため、ヒーリング制御については継続して実行可能とする必要がある。

## ③アップグレード作業は切戻しが可能であること

万一に備え、異常時にはアップグレードを切戻し、正常運用に復旧できることが必要となる。

これらの要件を満たし、通信サービスへの影響を与えずに、NFVO、VNFMのアップグレードを実現した。その仕組みを以下に解説する。

## (1)VNF再構築を不要とする仕組み

NFVOとVNFMは、VNFのライフサイクル機能を提供するため、収容するVNFの仮想リソースに関する管理データ（以下、VNFデータ）を保持し

ている。アップグレードに伴うVNFの再構築を不要とするためには、旧システムで管理するVNFデータを新システムへ引き継ぎ、新システムで継続してVNFを運用可能とする必要がある。

一般的に、データを引き継いだシステムのアップグレードの方法としては、OSアップデートのように同一サーバ上で旧システムを新システムに更新する方法や、新システムを旧システムとは別サーバに構築し、データ移行後に切り替える方法が考えられる。本アップグレードでは、対象となるNFVO、VNFMのハードウェア要件が変更となることから、別サーバで新システムを構築し、データ移行後に切り替える方式とした。また、NFVOは、アップグレード対象外のVNFMも含め、複数のVNFMを収容しているため、アップグレードは、図4に示すとおり、NFVO、VNFMの2段階で行うこととした。

データ移行については、2段階目のVNFMのアップグレード時にNFVOとVNFMの双方で新形式のデータへの変換が必要となる。アップグレードするNFVO、VNFMは、ETSI NFV Stage3仕様に準拠した製品となるため、現行システムとは機能分担の一部変更が伴い、管理するVNFデータの構造やシステム間IFのParameterが大幅に異なる。そのため、現行システムから新システムへデータを移行するには、NFVO、VNFMをまたがった複雑かつ大量のデータ変換、整合が必要となることが分かった（図5）。これらの処理の開発には、製造ベンダが異なるシステム間でのデータ変換を行う必要があること、また、さまざまなデータパターンを考慮する必要があることから、移行データに誤りが発生するリスクがある。そこで、移行データの信頼性を確保しつつ、効率的にデータ移行する仕組みとして、データ変換、整合処理を開発するのではなく、NFVO、VNFMの基本的なライフサイクル機能であるイン

\*36 REST：APIの1つで、各リソース（URL）に対してGET、POST、PUT、DELETEでリクエストを送信し、レスポンスをXMLやjsonなどで受け取る形式のこと（レスポンスのフォーマット形式は指定されていない）。



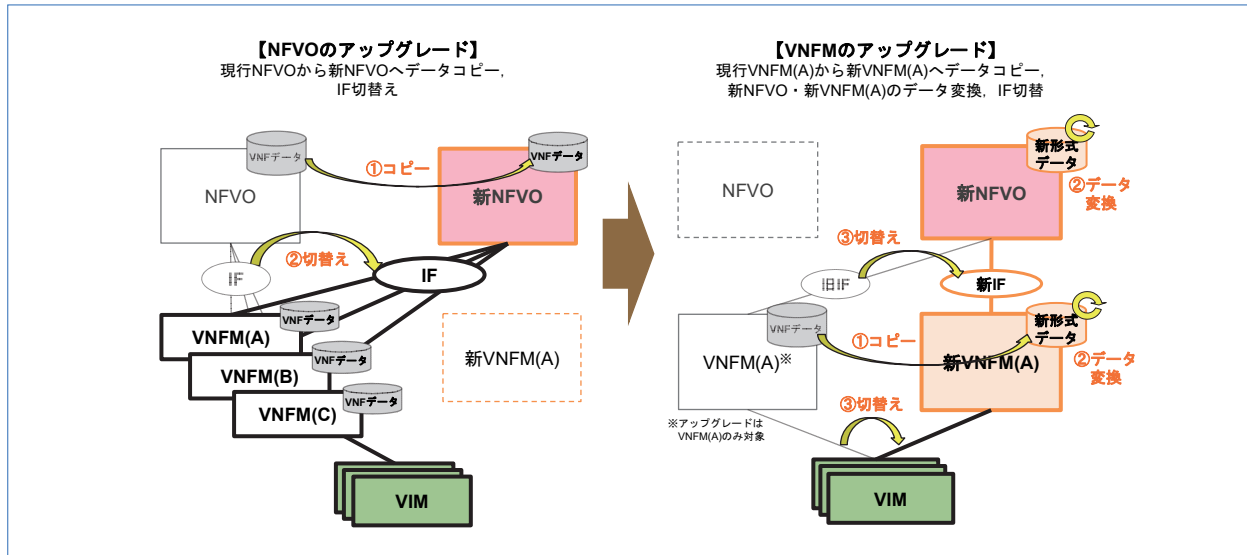


図4 アップグレード方法

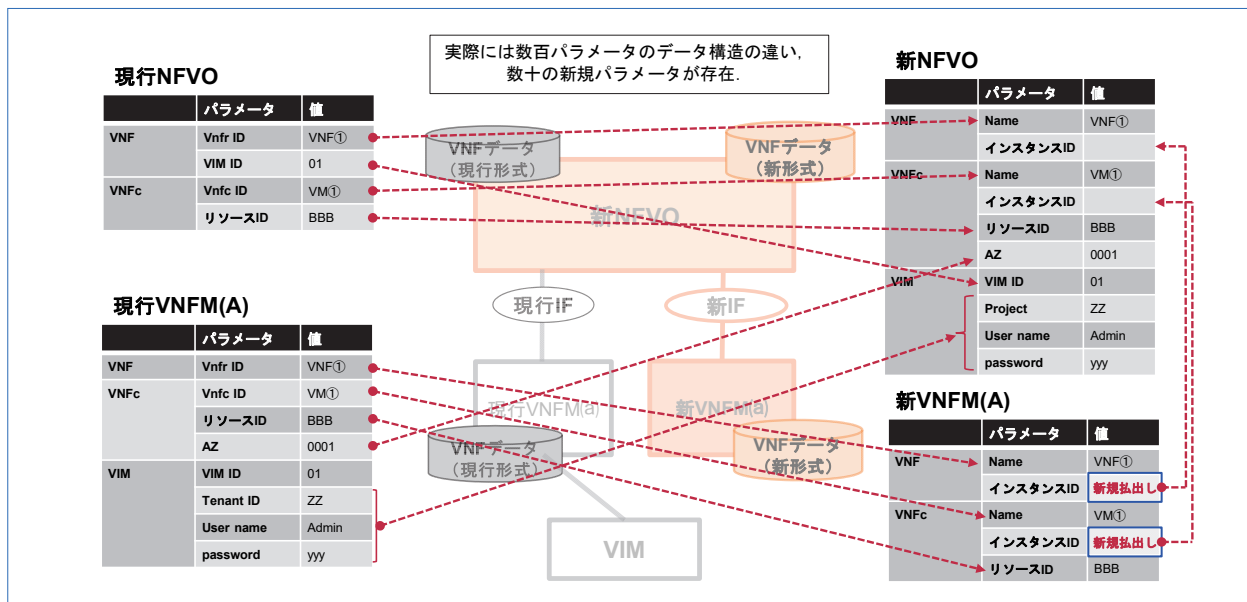


図5 データ変換・移行イメージ

スタンシエーション機能を利用する方式を検討した。インスタシエーションは、VNFの機能、動作を定義したVNFD、VMイメージを含むVNF Pack-

ageを基に、NFVO、VNFM、VIMが連携し、NFVI (NFV Infrastructure)\*37上に新たなVNFを構築する機能である。このインスタシエーションを、

\*37 NFVI：仮想化基盤を構成する汎用サーバ、ストレージ、ネットワーク機器の総称。

NFVI上に新たなVNFを構築することなく、新システム（NFVO, VNFM）間でのみ疑似的に実行させることで新形式のデータを生成する。これを疑似インスタンスーションと呼ぶ（図6）。現行システムで運用中のVNFを、新システム上で疑似インスタンスーションすることにより、新システム（NFVO, VNFM）上に新形式のVNFデータが生成されるため、新たなデータ変換処理を開発することなく、運用中VNFのデータを移行することを可能とした。

疑似インスタンスーションによるデータ移行を実現するため、新NFVO, 新VNFMでは、移行機能として、それぞれ以下の対応を行っている。

(a)疑似インスタンスーションの自動化（新NFVO）

従来、インスタンスーションは、VNF Packageを登録した上で、VNFごとに保守者がNFVOを操作する必要がある。しかし、400以上の

VNFデータを移行するため、保守者の稼働の増大や、設定ミスが懸念される。そこで、インスタンスーション時に設定が必要な情報を自動で抽出する仕組みとした（図6(a)）。

(b)インスタンスーション処理の終端（新VNFM）

通常のインスタンスーションでは、VNFMがVIMへの仮想リソース制御を行うが、疑似インスタンスーションでは新VNFMで処理を終端し、VIMへの仮想リソース制御を行わない。本来は、VIMへの仮想リソース制御の結果、運用中VNFの仮想リソースID情報（生成されたVMなどの仮想リソースのID情報）をVIMから収集することになるが、この制御の代わりとして、事前に仮想リソースID情報を現行VNFMから新VNFMへ移行しておき、疑似インスタンスーション処理でこの情報を取

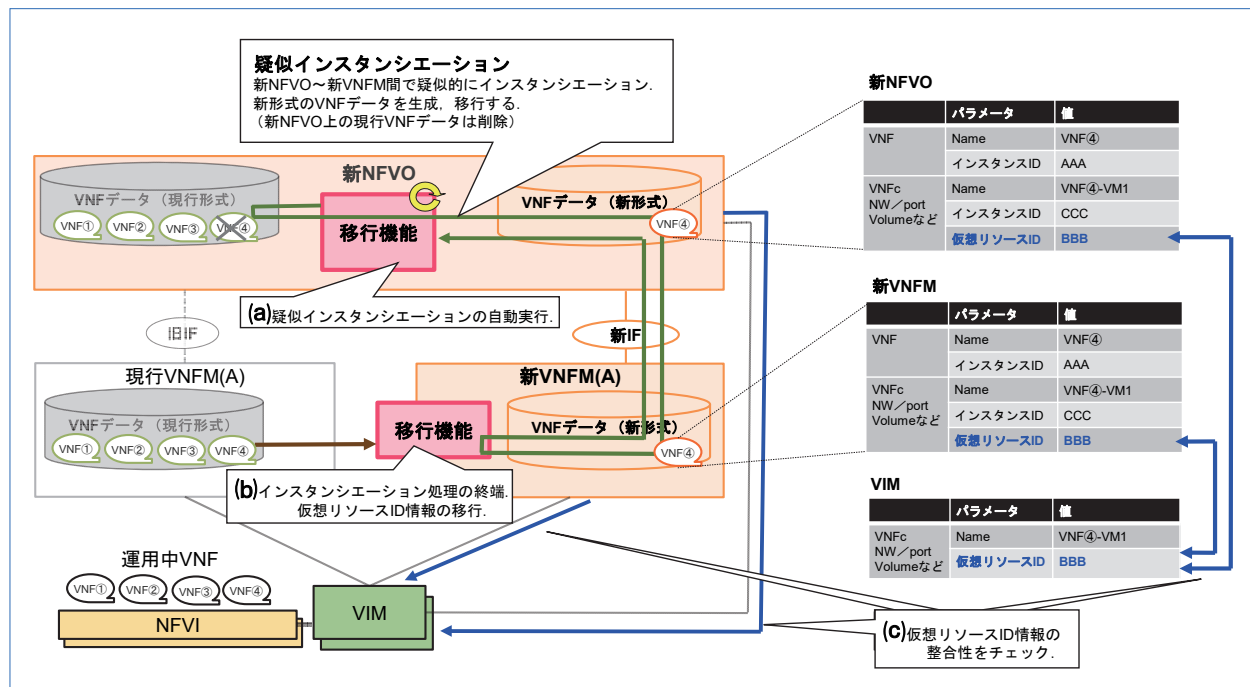


図6 疑似インスタンスーションの仕組み

り込む。これにより、VIMへの仮想リソース制御を実施せずに、インスタンスエーション処理を可能とした(図6(b))。

(c)システム間の移行データ整合(新NFVO, 新VNFM)

新NFVO, 新VNFMに、疑似インスタンスエーションの結果生成されたVNFデータ内の仮想リソースID情報を、VIMで管理される仮想リソースID情報と突合チェックする仕組みとし、新NFVO, 新VNFM, VIMの各システムでデータ整合性を確保した(図6(c))。

(2)ヒーリング制御継続の仕組み

疑似インスタンスエーションでデータ移行済みのVNFは、新システムへの切替えが完了するまで、ライフサイクルの実行ができなくなる。また、疑似インスタンスエーションにより、移行データの信頼性を向上させた一方で、400以上のVNFを処理するため、全VNFのデータ移行には時間が必要となり、ヒーリング制御の実行不可期間が許容時間を超えて

しまう(図7)。

そこで、データ移行期間中は、復旧手段を限定することで、VNFM単独でヒーリング制御を実行可能とする方式とした。NFVOの仮想リソース管理機能を利用しない限定的なヒーリング制御機能をシミュレートする仕組み(移行機能)を新VNFMに追加し、データ移行期間中はVNFM~新NFVOの接続を、VNFM~新VNFMに切り替えることで、ヒーリング制御継続を実現した(図8)。なお、NFVOの仮想リソース管理機能を利用できないため、データ移行期間中のヒーリング制御は、保守者が仮想リソース状況を確認した上で実行可能なマニュアルヒーリング制御に限定した。

(3)切戻し

問題発生時の迅速な復旧、リトライを可能とするため、アップグレード作業における切戻しは、可能な限り細かい単位で実行できる考慮が重要である。本アップグレードでは、データ移行および新システムへの切替え工程ごとに切戻しを考慮した。

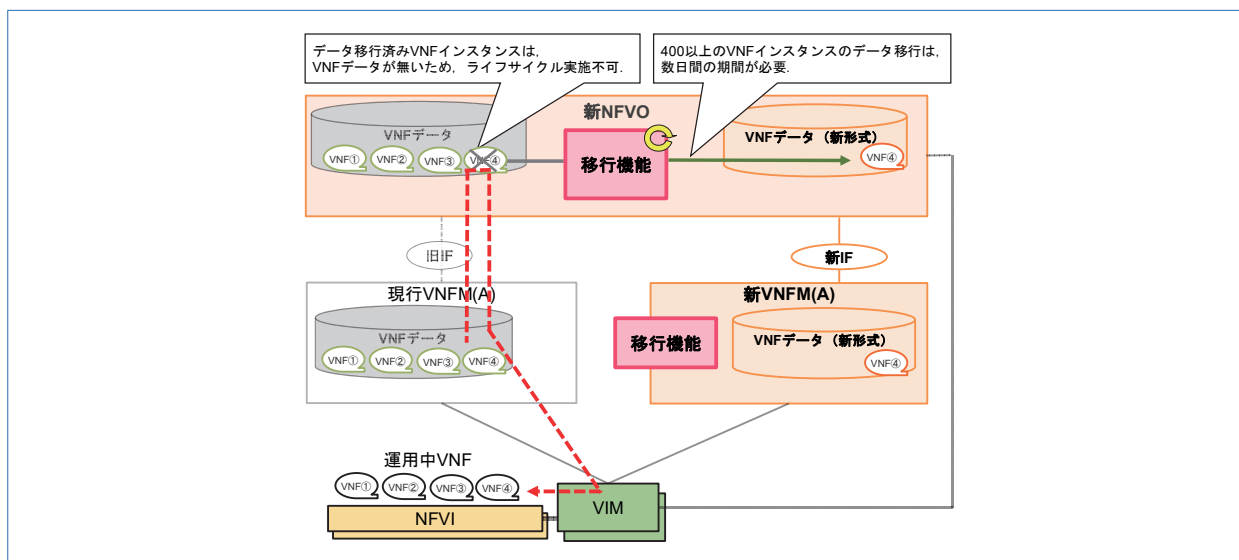


図7 データ移行中のVNFライフサイクル継続の課題

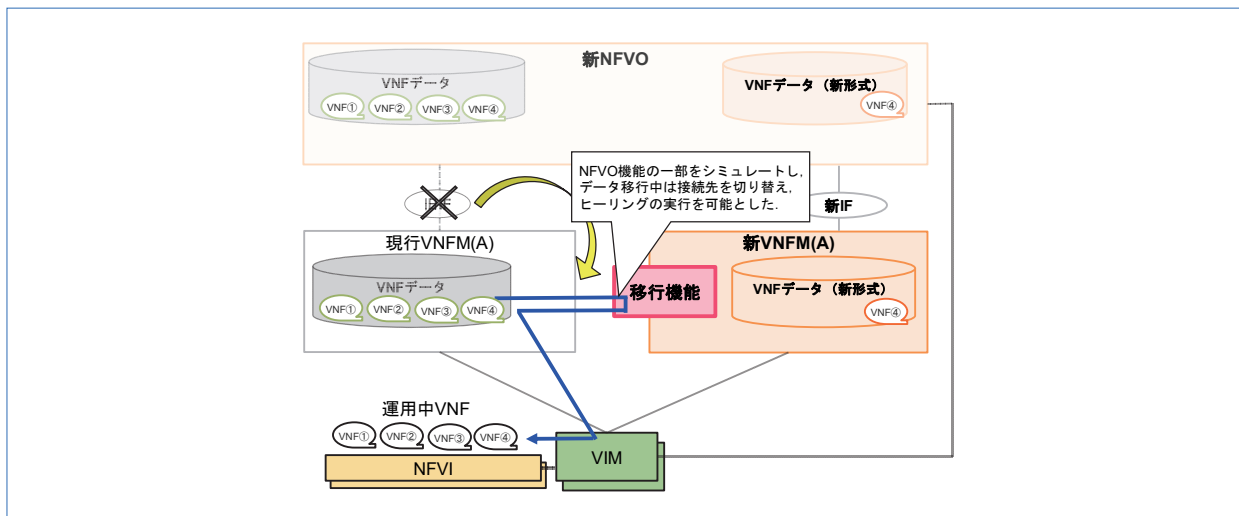


図8 データ移行中のヒーリング継続の仕組み

疑似インスタンスエーションによるデータ移行は、保守者が選択したVNF単位に実行可能とし、エラー時のロールバック機能を備えた。これにより、VNF Packageの作成ミスやデータ不具合による問題が発生した場合でも、何度でも切戻し、リトライを可能としている。また、データ移行完了後の新システムへの切替作業では、ネットワークを切り替えることで簡易に旧システムに切り戻すことを可能とした。これらは、検証における問題発生時のデータ修正などの対処、迅速なリトライのためにも活用され、短期間での検証完了の実現にも繋がった。

## 4. あとがき

本稿では、ドコモ仮想化基盤システムにおけるNFVO, VNFMのアップグレードの取組みと通信サービス無停止によるオンラインアップグレードを実現した手法を解説した。本取組みにより、NFVO, VNFMのアップグレード手法を確立するとともに、ETSI NFV Stage3準拠の製品が相互運用されるマ

ルチベンダ環境を実現した。

今後は、より効率的なアップグレード実現のため、引続きETSI NFVへの積極的な貢献、オープンな場におけるオペレータ間の連携や、ベンダへの製品化の働きかけを行っていく。また、マルチベンダ製品の相互接続性確保のためのPlugtest<sup>\*38</sup> [5] の推進と活用も進めていく予定である。

## 文献

- [1] 音, ほか: “通信ネットワークにおける仮想化技術の適用,” 本誌, Vol.24, No.1, pp.6-11, Apr. 2016.
- [2] 鎌田, ほか: “ドコモネットワークにおける仮想化基盤システムの実用化,” 本誌, Vol.24, No.1, pp.20-27, Apr. 2016.
- [3] ETSI: “Network Functions Virtualisation (NFV).” <https://www.etsi.org/technologies/nfv>
- [4] ETSI: “Directory Listing/ISG/NFV/Open/Publications\_pdf/Specs-Reports.” [https://docbox.etsi.org/ISG/NFV/Open/Publications\\_pdf/Specs-Reports](https://docbox.etsi.org/ISG/NFV/Open/Publications_pdf/Specs-Reports)
- [5] ETSI: “NFV&MEC IOP Plugtests 2021.” <https://www.etsi.org/events/1935-nfv-mec-iop-plugtests-2021>

\*38 Plugtest: NFVの導入と相互運用性を加速させるため、ETSI NFVで実行されているさまざまな組織間（プロダクトベンダなど）での共同テストプログラム。

# 3GPPにおける産業用アプリケーション連携に向けたフレームワークの標準化

ネットワーク開発部 **すずき ゆうじ**  
鈴木 悠司

近年、自動運転やドローン、スマートファクトリーなど、さまざまな産業におけるモバイルネットワークの活用が注目されている。3GPP TSG SA WG6では、多様な産業用アプリケーションが3GPPネットワークを活用するためのソリューションを検討している。特に、Release 15で導入されたCAPIFとRelease 16で導入されたSEALは、多くの産業で共通的に利用できるサービスフレームワークとして期待されている。本稿では3GPP TSG SA WG6の活動概要と、CAPIF、SEALの技術仕様を解説する。

## 1. まえがき

高速・大容量、低遅延、多数端末同時接続を掲げる第5世代移動通信システム（5G）の導入や、IoT（Internet of Things）技術の発展、各種産業分野のデジタル化などにより、通信に限らない多種多様な産業分野でモバイルネットワークを活用することが期待されている。3GPP（3rd Generation Partnership Project）では、SCEF（Service Capability Exposure Function）<sup>\*1</sup>やNEF（Network Exposure

Function）<sup>\*2</sup>のAPI（Application Programming Interface）<sup>\*3</sup>によるネットワーク機能の提供に代表されるように、3GPPドメイン外のアプリケーションとの連携も視野に入れながら技術仕様を拡張してきた。

3GPP TSG SA WG6（以下、SA6）では、従来、ミッションクリティカル<sup>\*4</sup>な通信に関するソリューションを中心に議論してきたが、Release 15以降はこれに加え、前述のような産業連携を念頭に置き、3GPPドメイン外のアプリケーションによる3GPP

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

<sup>\*1</sup> SCEF：3GPPシステムがもつ機能の一部を3GPPドメイン外に提供するための論理ノード。主に4G向けのコアネットワークで利用される。

<sup>\*2</sup> NEF：SCEFと同様に、3GPPシステムがもつ機能の一部を3GPPドメイン外に提供するための5Gコアネットワークの機能部。

<sup>\*3</sup> API：アプリケーションから特定のサービスを利用するためのインタフェース。本稿では、特にRESTful APIを指す。

ネットワーク活用を支援する機能の仕様策定も視野に含めて活動している。その例として、Release 15では3GPPが提供するノースバウンドAPI<sup>\*5</sup>に対する統一的なフレームワークであるCAPIF (Common API Framework)<sup>\*6</sup>を策定し、Release 16では各種の産業用アプリケーションが共通的に利用できる機能を切り出したSEAL (Service Enabler Architecture Layer for Verticals)<sup>\*7</sup>を導入した。

本稿ではSA6の活動概要を紹介し、その中で特に産業用アプリケーション連携の基礎となるCAPIF、SEALという2種類のサービスフレームワークを解説する。

## 2. 3GPP SA6の活動

3GPPでは、最上位のグループであるPCG (Project Coordination Group)<sup>\*8</sup>の下、無線関連技術を検討するRAN (Radio Access Network)<sup>\*9</sup>、サービス機能やシステム全体のアーキテクチャを検討するSA (Service and System Aspects)<sup>\*10</sup>、コアネットワーク<sup>\*11</sup>や端末のインタフェースや機能を検討するCT (Core Network and Terminals)<sup>\*12</sup>という3つのTSG (Technical Specification Group)<sup>\*13</sup>と呼ばれるグループで、仕様策定を分担している。各TSGはそれぞれ4~6のワーキンググループに細分化され、SA6はTSG SAの中で、特にミッションクリティカル関連のアプリケーションにかかわる機能やアーキテクチャを検討するワーキンググループとして、2014年に発足した。

SA6発足当初は、パブリックセーフティ<sup>\*14</sup>分野での活用を念頭に置いたミッションクリティカルサービス (プッシュ・トゥ・トーク<sup>\*15</sup>通信、ビデオ通信、データ通信) を対象として仕様検討を進めていた。その後、5Gの初期仕様の策定と同じRelease

15にて、3GPPが提供するノースバウンドAPIに対する統一的なフレームワークとしてCAPIFを策定し、Release 16以降もミッションクリティカルサービスに限らない各種サービスのフレームワークや、V2X (Vehicle-to-Everything)<sup>\*16</sup>・ドローン<sup>\*17</sup>・スマートファクトリー<sup>\*18</sup>に代表される種々の産業分野におけるアプリケーションの活用について検討を開始するなど、対象を拡大してきた。現在では「Application Enablement and Critical Communication Applications」を議論対象として、3GPPネットワークを活用するアプリケーション向けのアーキテクチャを中心に仕様検討している [1]。

## 3. CAPIF

### 3.1 CAPIF導入の目的

CAPIFは、3GPPで提供する種々のAPIに対する統一的な枠組みを提供する目的で、Release 15にて導入された。3GPPでは前述のSCEF、NEFやMBMS (Multimedia Broadcast/Multicast Service)<sup>\*19</sup>向けの機能部などが、サービスAPIを提供している。一方、これらのAPIの開発・利用に際しては、提供するAPIの発行や管理、セキュリティ関連の機能など、共通して考慮すべき課題があり、CAPIFはこのような共通的な課題を解決するためのフレームワークとしての役割を担っている。

### 3.2 CAPIFのアーキテクチャ

CAPIFの代表的なアーキテクチャを図1に示す。

#### (1)API invoker

API invokerは、CAPIF APIおよびサービスAPIの呼出し元である。ネットワークオペレータがもつアプリケーション、ネットワークオペレータ以外の他事業者がもつアプリケーションのいずれもAPI

\*4 ミッションクリティカル：サービスを継続的に提供できることが極めて重要であり、障害などによる中断が許されない、あるいは非常に大きな損害になり得るシステムを指す。

\*5 ノースバウンドAPI：APIを提供している装置から見て、上位のアプリケーションに対して提供しているAPI。

\*6 CAPIF：ノースバウンドAPIを提供するための共通的な機能をまとめた3GPPのフレームワーク。

\*7 SEAL：3GPPネットワークを利用する複数の産業用アプリケーションで共通的に利用する機能をまとめた層。

\*8 PCG：3GPPの最高意思決定機関。3GPP活動全体の計画や進捗

管理などを行う。

\*9 RAN：3GPPにおいて、コアネットワークと端末の間に位置する。無線レイヤの制御を行う基地局などで構成されるネットワークに関する仕様化を行っているグループ。

\*10 SA：3GPPにおいて、サービス要求条件、アーキテクチャ、セキュリティ、コーデック、ネットワーク管理に関する仕様化を行っているグループ。

\*11 コアネットワーク：交換機、加入者情報管理装置などで構成されるネットワーク。移動端末は無線アクセスネットワークを経由してコアネットワークとの通信を行う。

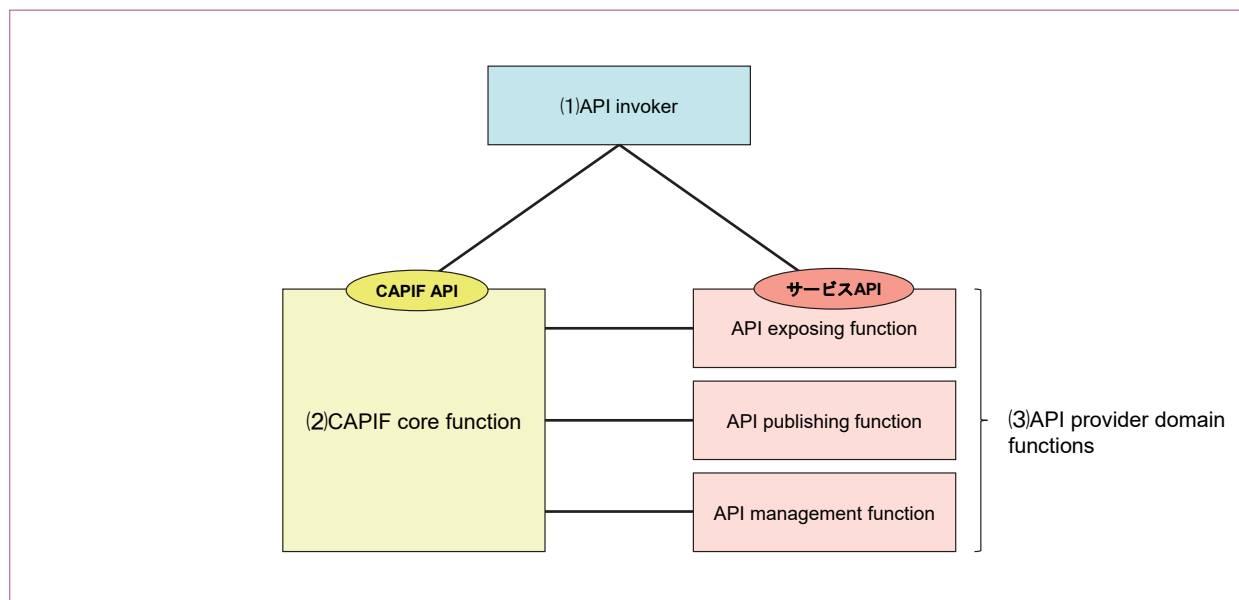


図1 CAPIFのアーキテクチャ

invokerになり得る。

#### (2)CAPIF core function

CAPIF core functionは、API invokerの認証・認可やAPIの登録、ポリシーの管理など、CAPIFで提供する各種機能において中心的な役割を担う。これらの機能はAPIとして提供され、API invokerをはじめとする各機能部はこのAPIを呼び出すことで機能を利用できる。CAPIF core functionは、モバイルネットワークオペレータが信頼できるドメイン内に配置される。

#### (3)API provider domain functions

「API exposing function」「API publishing function」「API management function」の3機能部は、総称してAPI provider domain functionsと呼ばれる。CAPIFでは、CAPIF core functionの提供者をCAPIFプロバイダ、API provider domain functionsの提供者をAPIプロバイダと定義しており、両者は別々の事業者でも同一の事業者でもよい。

API exposing functionは、API invokerからのサービスAPI呼出しを受け入れる機能部である。

API publishing functionは、サービスAPIをAPI invokerが利用できるようにするために、CAPIF core function宛にサービスAPI情報を発行 (publish) する役割を担う。最後に、API management functionは、発行されたサービスAPIの管理を担い、サービスAPI呼出しログの監査やサービスAPIの状態の監視などの機能をもつ。

### 3.3 CAPIFの主な機能

ここではCAPIFの代表的な機能として、API invokerのオンボーディング、サービスAPIの発見、API invokerの認証・認可の3機能に絞って解説する。これらの機能は、API invokerがサービスAPIを利用する前段階として重要な機能である。その他の詳細な機能については文献 [2] を参照されたい。

\*12 CT：3GPPにおいて、コアネットワーク内、および移動端末とコアネットワーク間のプロトコルの仕様化を行っているグループ。

\*13 TSG：3GPPにおいて、技術仕様策定を担うグループ。

\*14 パブリックセーフティ：警察、消防、救急のような公共の安全のためのサービス。

\*15 プッシュ・トゥ・トーク：トランシーバのように、ボタンを押している間だけ発信ができる音声通信の方式。

\*16 V2X：車両と車両、車両と路上の物体（信号機など）のように、車両と周囲の環境の間で通信できるようにする技術。

\*17 ドローン：人が搭乗して操縦しない航空機のこと。UAV

(Uncrewed Aerial Vehicle) とも言う。3GPPではドローン関連の機能を含むシステムをUAS (Uncrewed Aerial System) と呼ぶ。

\*18 スマートファクトリー：IoTなどの通信技術を活用した工場システム。3GPPでは特にFF (Factories of the Future) と呼ぶ。

\*19 MBMS：3GPPシステムで提供する1対多型（ブロードキャスト・マルチキャスト）の通信サービス。

### (1)API invokerのオンボーディング

API invokerは、サービスAPIの呼出しを要求するにあたり、事前に自身の情報をCAPIF core functionに提供し、承認してもらわなければならない。この手順をオンボーディングと言う。オンボーディングに成功すると、API invokerは以後の認証・認可に必要な情報を受け取ることができる。

オンボーディングによりCAPIF core functionはAPI invokerを認識し、この手順で取得した情報を使って、API invokerの認証・認可ができるようになる。また、CAPIF core functionはこのオンボーディングの中で、API invokerが呼び出せるサービスAPIの情報を送ることもできる。

### (2)サービスAPIの発見

API invokerが呼び出せるサービスAPIの情報については、前述のオンボーディング手順の中でCAPIF core functionから提供されるのを待つだけでなく、API invoker自身がCAPIF core functionに問い合わせることもできる。この問合せによりサービスAPI情報を取得する手順を、サービスAPIの発見 (discovery) と言う。

API invokerが自身の識別情報とともに、発見したいAPIの条件をCAPIF core functionに送信すると、CAPIF core functionは保管しているAPI情報の中から条件に一致するAPIを取得する。CAPIF core functionはまた、特定のカテゴリのAPIを発見の対象から外すなど、取得したAPIの情報を自身の発見ポリシー (discovery policy) に従ってさらに選別することができる。このようにして得られたサービスAPIの一覧をAPI invokerに送信することで、API invokerは目的のサービスAPIの情報を得ることができる。

### (3)API invokerの認証・認可

目的のサービスAPIの情報を取得したAPI invok-

erは、そのサービスAPIの呼出しにあたって認証・認可のプロセスを経る必要がある。API invokerとAPI exposing functionの間でどのような認証・認可の手法 (以下、セキュリティメソッド) を採用するかを、両者がどのセキュリティメソッドをサポートするかなどの情報に基づいて、事前にAPI invokerとCAPIF core functionの間で決定する。セキュリティメソッドの種類としては、「TLS-PSK (Pre-Shared Key Ciphersuites for Transport Layer Security)\*<sup>20</sup>の利用」、「PKI (Public Key Infrastructure)\*<sup>21</sup>の利用」、「TLSとOAuth\*<sup>22</sup>トークン」の3種類が規定されている [3]。なお、このようなセキュリティ関連の詳細な仕様策定は3GPP TSG SA WG3 (SA3) が担っている。

利用するセキュリティメソッドの決定後、API invokerはサービスAPIを呼び出す前段階として、あるいはサービスAPIの呼出しと同時にAPI exposing functionに認証・認可を要求する。API exposing functionは、必要に応じてCAPIF core functionと連携しながら、あらかじめ決定したセキュリティメソッドに基づいてAPI invokerの認証・認可を行う。このようなセキュリティの仕組みを統一的に提供することもCAPIFの重要な役割の1つである。

## 4. SEAL

### 4.1 SEAL導入の目的

SEALは、3GPPネットワークを利用する複数の産業用アプリケーションで共通的に利用する機能をまとめた層である。3GPPではV2X・ドローン・スマートファクトリーなど、ある特定分野のサービス・製品を提供する業界や企業の集合をパーティカルドメイン、あるいは単にパーティカルと呼ぶ。各パーティカルアプリケーションはそれぞれの要求条

\*20 TLS-PSK：事前に共有した鍵 (PSK) を利用して、通信を暗号化するTLSコネクションを確立する方法。

\*21 PKI：暗号技術に用いる公開鍵の登録者を証明する仕組み、公開鍵基盤とも言う。本稿では、特にPKIを利用してTLSコネクションを確立する方法を指す。

\*22 OAuth：アクセス権限の認可を行うための標準仕様。本稿では、特にIETF RFC (Internet Engineering Task Force Request for Comments) 6749で規定されたOAuth 2.0を指す。トークンと呼ばれるデータを発行してアクセス権限を制御する。



件に応じて必要な機能を実装するが、中には複数のパーティカルアプリケーション間で共通して必要となる機能もある。そのような機能をSEALとしてまとめて提供することで、パーティカルアプリケーションごとに個別に実装する必要がなくなり、効率的なシステム開発につながる。

## 4.2 SEALのアーキテクチャ

SEALの代表的なアーキテクチャを図2に示す。

SEALは、クライアント側の機能を担うSEALクライアントと、サーバ側の機能を担うSEALサーバからなる。両者は互いに通信することで位置情報管理やグループ管理などの機能を実現する。SEALサーバは3GPPネットワークシステムが提供する機能（例えば、NEFが提供するAPI）を利用することができる。なお、特定のSEALの機能に着目する場合、SEALクライアント・SEALサーバはその機能

に従った名前と呼ばれることがある（例えば、位置情報管理のためのSEALサーバは特にLocation management serverと呼ばれる）。

SEALの上位にはVAL（Vertical Application Layer）が位置づけられる。このVALが、各パーティカルに特有のアプリケーション（例えば、V2Xアプリケーション）が乗る層であり、VALクライアント・VALサーバは、下位層のSEALクライアント・SEALサーバと通信することで、SEALが提供する機能を利用することができる。

## 4.3 SEALの主な機能

ここではSEALの代表的な機能として、位置情報管理、グループ管理、ネットワークリソース管理の3機能に絞って解説する。SA6では各パーティカル向けに個別のソリューションも検討しているが、それらの検討の中でも上記3機能の活用は特にさかん

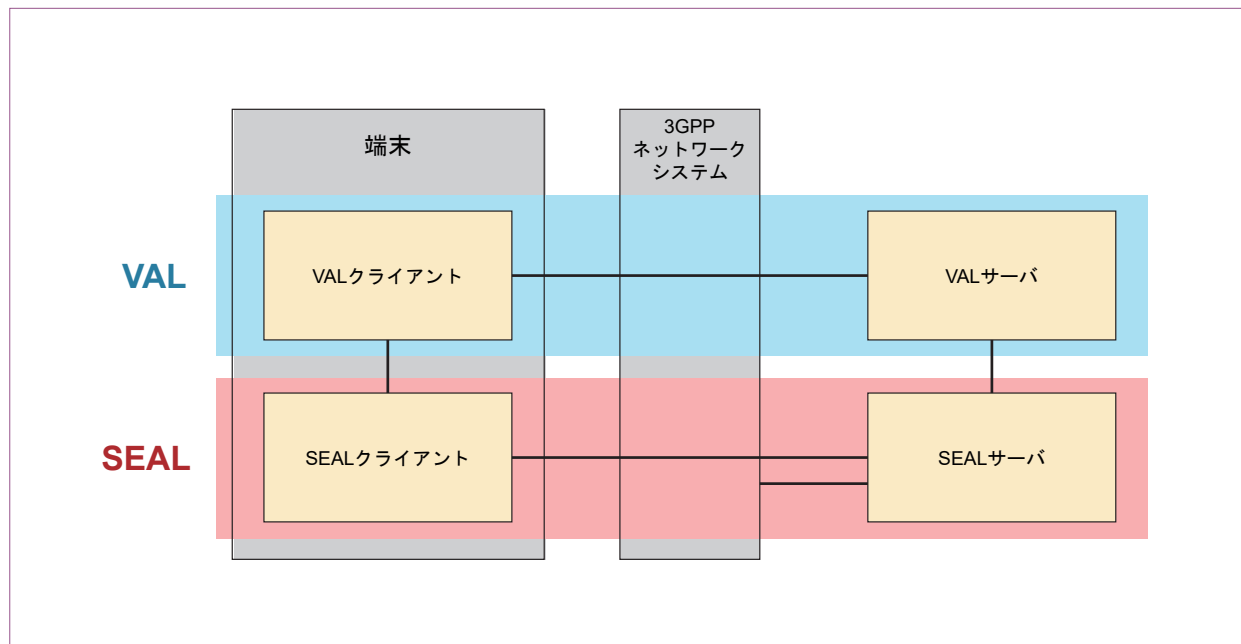


図2 SEALのアーキテクチャ

に議論されている。その他の詳細な機能については文献 [4] を参照されたい。

#### (1)位置情報管理

位置情報管理を担うSEALクライアント・サーバはそれぞれLocation management client/serverと呼ばれる。位置情報管理機能では、VALサービス利用者の位置に関する情報を取得できる。

例えば、VALサーバが特定のVALサービス利用者の位置情報を特定したいとき、VALサーバはLocation management serverに対して位置情報報告の動作を開始するための信号として、位置情報報告トリガ (Location report trigger) を送信する。この信号を受け取ったLocation management serverは、Location management clientに位置情報を問い合わせ、受け取った位置情報をVALサーバに送信する。

VALサーバからだけでなく、ある端末から別の端末の位置情報を特定することも可能である。例えば、Location management client Aが実装されている端末の位置情報を知りたいとき、その端末とは別の端末上に実装されているLocation management client Bは、Location management serverに対して位置情報報告トリガを送信することができる。前述の例と同様に、位置情報報告トリガを受け取ったLocation management serverは、Location management client Aに位置情報を問い合わせ、受け取った位置情報をLocation management client Bに送信する。これにより、目的の端末の位置情報を取得することができる。

位置情報を取得するタイミングについても、目的に合わせて変更できる。具体的には、VALサーバまたはLocation management clientが特定のVALサービス利用者の位置情報を取得するとき、位置情報報告トリガを送信した直後に位置情報を受け取る

ことも、あるいは特定の条件を設定して、その条件が満たされたタイミングで位置情報を受け取る (例えば、一定時間おきに位置情報を受け取る) ことも可能である。また、Location management serverは、Location management clientから直接位置情報を受け取るだけでなく、3GPPネットワークから端末の位置情報を受け取り、その情報をVALサーバなどに送信することも可能である。

#### (2)グループ管理

グループ管理を担うSEALクライアント・サーバはそれぞれGroup management client/serverと呼ばれる。グループ管理機能では、複数のVALサービス利用者からなるグループの作成や、そのグループ内のメンバの管理などができる。

例えば、グループ作成権限をもつVALサービス利用者が新しくグループを作成したいとき、そのVALサービス利用者のGroup management clientはGroup management server宛にグループ作成要求 (Group creation request) を送信する。このとき、同一グループ内に含めたいユーザの識別情報も一緒に送信する。この要求に基づいてGroup management serverは新規グループを作成する。グループ作成後、Group management clientからGroup management serverにグループ情報問合せ要求 (Group information query request) を送信することにより、作成したグループに関する情報を取得することができる。また、Group management clientはGroup management serverに対して、グループメンバーシップ更新要求 (Group membership update request) を送信することで、グループのメンバを追加・削除することもできる。

前述した位置情報管理機能と組み合わせて使うことで、位置情報ベースのグループ作成も可能である。このとき、Group management serverは、特定の

位置情報の中に存在するユーザのリストをLocation management serverに要求して取得し、そのリスト内のユーザでグループを作成する。このように、SEALの複数の機能を利用してサービスを提供することもできる。

### (3) ネットワークリソース管理

ネットワークリソース管理を担うSEALクライアント・サーバは、それぞれNetwork resource management client/serverと呼ばれる。ネットワークリソース管理の具体的なユースケースとして、VALサービスの要求条件に応じたQoS (Quality of Service)<sup>\*23</sup>の適用などができる。

目的のQoSを得たい場合、VALサーバはNetwork resource management server宛に特定の端末（もしくは複数端末からなるグループ）に対するネットワークリソース適応要求（Network resource adaptation request）を送信する。これに基づき、Network resource management serverは、目的の端末ないし端末グループにネットワークリソースを割り当てる。Network resource management serverは、PCF (Policy Control Function)<sup>\*24</sup>などの3GPPシステムとも接続し、要求に基づいてPCC (Policy and Charging Control) プロシージャ<sup>\*25</sup>を開始する。これにより、各VALサービスの要件に応じたQoSを適用することが可能になる。

## 5. あとがき

本稿では、3GPP SA6の活動概要と、そこで標準

化され、産業用アプリケーションと3GPPシステムとの連携における役割が期待されるCAPIFとSEALという2種類のフレームワークについて解説した。CAPIFはRelease 15で、SEALはRelease 16で導入されて以降、今も継続的に拡張が議論されており、ドコモも機能拡張の検討に貢献している。また、Release 17以降では、エッジコンピューティング<sup>\*26</sup>実現のためのアーキテクチャや、V2X・ドローン・スマートファクトリーといった個々の産業向けのサービスに特有な技術仕様も検討されている。現在、SA6ではRelease 18の技術検討を開始した段階であり、ドコモはさらなる産業用アプリケーションのユースケースを見据えながら、3GPP SA6での標準化活動を推進していく。

### 文献

- [1] 3GPP SP-210265 : "Terms of Reference (ToR) for 3GPP TSG SA WG6 (SA6)," Mar. 2021.
- [2] 3GPP TS23.222 V17.5.0 : "Functional architecture and information flows to support Common API Framework for 3GPP Northbound APIs; Stage 2," Jun. 2021.
- [3] 3GPP TS33.122 V16.3.0 : "Security aspects of Common API Framework (CAPIF) for 3GPP northbound APIs," Jul. 2020.
- [4] 3GPP TS23.434 V17.3.0 : "Service Enabler Architecture Layer for Verticals (SEAL); Functional architecture and information flows," Sep. 2021.

<sup>\*23</sup> QoS : 通信サービスの品質、帯域幅や遅延などが代表的な指標となる。

<sup>\*24</sup> PCF : QoSなどのポリシー制御や、課金制御を担う5Gコアネットワークの機能部。

<sup>\*25</sup> PCCプロシージャ : ポリシー制御や課金制御に関連する一連の処理。Network resource management serverからの要求は、PCCプロシージャによって3GPPシステム内部にも反映される。

<sup>\*26</sup> エッジコンピューティング : 端末に近い場所で計算処理を行うサービスの形態。通信における遅延の短縮や、ネットワーク負荷の分散などの効果が期待される。

## 2021年「日本ITU協会創立50周年 各賞」受賞

日本ITU協会は2021年9月1日に創立50周年を迎え、「日本ITU協会創立50周年記念式典」を完全リモート開催しました。この式典において、電波企画室の橋本 明が総務大臣賞、日本ITU協会創立50周年特別賞および感謝状を、ドコモ・テクノロジーの尾上 誠蔵が日本ITU協会創立50周年記念賞を受賞しました。

橋本は、1980年より継続的にITU-R (International Telecommunication Union-Radio communication sector) における標準化活動に参加し、研究委員会、無線通信総会などの議長職を歴任するとともに、さまざまな地上無線通信システムの国際標準化活動で多大な貢献をしました。特にITU-R SG5 (Study Group 5) 議長時代には、第4世代移動通信システム無線インタフェース勧告 (勧告 ITURM.2012) の議論を主導し、タイムリーに完成させました。この勧告は、全世界で45億人以上が利用するLTE-Advancedシステムの技術規格を含んでおり、現在



(右上) 橋本 明



尾上 誠蔵

に至る世界の無線通信市場の拡大と利便性向上に大きく貢献しています。さらに、2018年にはITU無線通信規則委員会 (RRB : Radio Regulations Board) の委員に選出され、現在も無線通信に関する世界的な規律の維持向上に多大な貢献をしていることが認められ、総務大臣賞および日本ITU協会創立50周年特別賞を受賞しました。

尾上は、移動通信分野の研究開発において、卓越した技術力・指導力により、第3世代から第4世代に至る移動通信システムの実用化および国際標準化に多大な貢献をしました。第4世代移動通信システムにおいては、第3世代からの円滑な世代移行が求められる中、Super3Gコンセプト (後のLTE (Long Term Evolution)) をいち早く提唱し、3GPP (3rd Generation Partnership Project) におけるLTE/LTE-Advanced標準仕様を完成させ、最終的に第4世代移動通信システム無線インタフェース勧告 (勧告 M.2012) に盛り込まれました。本技術は今日の情報社会に欠かすことのできないものであり、これをなした多大な貢献が認められ日本ITU協会創立50周年記念賞を受賞しました。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。



## 2021年「日本ITU協会賞」受賞

2021年11月30日に開催された「第53回世界情報社会・電気通信日のつどい」において、ネットワーク部の田村 基が日本ITU協会賞「功績賞」を、ネットワーク開発部の青柳 健一郎、6G-IOWN推進部の栗田 大輔、電波企画室の坂本 信樹、無線アクセス開発部の手島 邦彦が日本ITU協会賞「奨励賞」を受賞しました。

日本ITU協会賞は、電気通信／ICTと放送分野に関する国際標準化や国際協力の諸活動において、これまでに優れた功績を遂げられた者ならびに今後の貢献が期待される者に贈呈されるものです。中でも、功績賞は、世界情報社会サミットにおける基本宣言および行動計画の実現および国際標準化、国際協力に関するITU（International Telecommunication Union）などの活動または我が国におけるITUなどに関連する諸活動に貢献し、その他情報通信および放送の発展に寄与し、その功績が著しい者に贈られます。また、奨励賞は、功績賞に該当する諸活動にすでに参加し、今後これらの領域において継続して寄与することが期待される者に贈られます。

田村は、第2世代から第5世代にわたる移動通信網のネットワークアーキテクチャ標準化に一貫して携わり、ITU、3GPP（3rd Generation Partnership Project）その他の関連標準化機関において、各世代における移動通信網技術の国際標準仕様策定に貢献し、移動通信ネットワークの実用化を推進しました。また途上国における技術支援に尽力し、移動通信の新技術導入に寄与した功績が認められ、功績賞を受賞しました。

青柳は、移動通信の標準技術仕様検討団体である3GPPにて、無線アクセス、およびコアネットワー

ク、双方の分野で技術仕様策定に貢献しました。無線アクセス分野では、携帯へ緊急地震速報を実現する技術仕様の策定を主導しており、コアネットワーク分野においては、サービス初期段階では限定的な5G利用エリアを4Gで補完する技術仕様の策定に貢献するなど、今後移動通信に関する標準化活動において、ますますの活躍が期待され、奨励賞受賞となりました。

栗田は、3GPP標準化において、UMTS（Universal Mobile Telecommunications System）のアンテナ要求性能規定や、LTE/LTE-AdvancedのMIMO（Multiple Input Multiple Output）アンテナ評価技術の策定、5G NR（New Radio）のバックホールリンク適用技術、アンライセンス周波数利用技術、通信エリア拡張技術、機能制約UE（User Equipment）の技術検討・仕様策定などにおいて技術議論を主導しました。技術仕様策定にも貢献してきており、今後ますますの活躍が期待され、奨励賞受賞となりました。

坂本は、IMT（International Mobile Telecommunications）用周波数に関する国際標準化活動に従事しており、ITU-R（International Telecommunication Union-Radio communication sector）WP5D（Working Party 5D）会合における関係各国との調整、ARIB（Association of Radio Industries and Businesses）内における対処方針の取り纏め、CJK IMT-WG（Working Group）における中国、韓国との連携強化などを通じて議論を促進しました。WRC（World Radiocommunication Conference）-19におけるIMT用周波数の追加特定や、その他周波数共用検討に関する研究に貢献するなどしており、今後のさらなる活動を期待され、奨励賞受賞となりました。

手島は、移動通信分野の中心的な役割を果たしている標準化団体である3GPPにおいて、複数の無線技術案件で取り纏め役を務めるなど、4Gおよび5Gの標準仕様の策定に大きく貢献し、また、無線基地局のオープン化を目的とした団体であるORAN（Open Radio Access Network）アライアンスにおいて、ワーキンググループの共同議長を務め、活動を牽引するさらなる活躍が期待され、奨励賞受賞となりました。



（左から）坂本、田村、手島、栗田

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。  
※所属は推薦時のもの。

**NTT DOCOMO**  
**テクニカル・ジャーナル Vol.29 No.4**

2022年1月発行

企画編集 株式会社NTTドコモ R&D戦略部  
〒100-6150  
東京都千代田区永田町 2-11-1  
山王パークタワー39階

発行 一般社団法人 電気通信協会  
〒101-0003  
東京都千代田区一ツ橋 2-1-1  
如水会ビルディング6階

本誌掲載内容についてのご意見は  
e-mail: dtj@nttdocomo.com 宛

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標です。本誌掲載記事の無断転載を禁じます。

© 2022 NTT DOCOMO, INC.