

特集

モバイル時代の サービスを支える技術



編集に あたって

携帯電話会社の将来予測によると、2020年の移動通信のトラフィックは、2010年比で、500倍～1,000倍になるとされている。スマートフォンやウェアラブル端末に加え、家電、家具、建築物、薬など、一見ITとは無関係なものもインターネットにつながっていく。ヒトとモノが生成するトラフィックは、さまざまな相互作用をもたらすことが予想できる。IT調査会社の予測によると、2020年にはインターネット上の全情報の3分の1がクラウドに置かれるか、クラウドを通過すると見積もられている。クラウドで把握できるビッグデータを活かすことができれば、今日の社会が抱える問題を解決したり、明日のニーズを予測したりすることができるかもしれない。さらに、ヘルスケア、交通／輸送、エネルギーなどの分野においても、全人類にメリットをもたらす新しい道筋を切り開くことができるかもしれない。

渡邊 晃 ■ 名城大学

このような状況を見据えて、ビッグデータを扱うためのさまざまな技術が生まれており、これらの技



術を整理すべく今回の特集記事を企画した。通信会社は膨大なトラフィックをさばくために、ネットワークをどのように変革しようとしているのだろうか。人とモノから発生するトラフィックは、どのような性質を持ち、どのように処理されるべきなのだろうか。収集したビッグデータとは、どのように付き合ったらよいのだろうか。さらに、複数種類のビッグデータを処理するためには、新たな発想が必要になるのではないだろうか。このような疑問に答えるため、それぞれの専門分野の方々に執筆を依頼した。各記事における内容は以下のとおりであり、特集1～4は5につながるための伏線となっている。

特集1「次世代のモバイルネットワークはどうなるのか」 奥村, 浅井, 岩科, 清水 (NTTドコモ): 次世代の移動通信ネットワークの実現に向け、どのような検討が進められているのかを概観する。

特集2「コンシューマデバイスとパーソナルデータの利活用および保護の動向」 石川, 松前 (駒澤大): 人が身につけるコンシューマデバイスを利用したパ

ーソナルデータ収集の現状, 解決すべき課題, 今後の展望について示す。

特集3「M2Mの情報流」 猿渡 (静大), 森川 (東大): 膨大な量のM2M (Machine-to-Machine) データが発生するため、これをさばくために考慮しなければいけない点は何なのかを明らかにする。

特集4「ビッグデータとのつきあい方」 原 (阪大): 人とモノから生み出された多量のデータの扱いについて、クラウド上での行うべき処理・蓄積技術を述べ、その応用基盤として“データ銀行”について述べる。

特集5「多数のデータストリームをリアルタイムで融合・編纂し利活用するための次世代『情報流』技術」 安本 (奈良先端大), 山口 (阪大): データを「流れ」として捉え、ビッグデータの組合せにおける複合的処理の必要性を述べる。また、新しい研究分野の立ち上げを根付かせる研究コミュニティ形成の展望について述べる。

(2014年7月1日)

1 次世代のモバイルネットワークは どうなるのか

— 5G 無線アクセスとコアネットワーク —

奥村幸彦 浅井孝浩 岩科 滋 清水敬司

■ (株) NTT ドコモ 先進技術研究所

モバイルネットワークの発展経緯

モバイルネットワークはこれまで「高速・大容量化」を繰り返す形で継続的に発展し、おおよそ10年ごとに新しい世代へと進化を遂げてきている(図-1)。第1世代のアナログ方式、第2世代のデジタル方式までは各国、地域で異なる方式が導入され、相互に互換性のない複数の方式が運用されていた。第3世代では、新しいデジタル方式により高速のデータサービスを提供することに加えて、世界で統一された標準方式の導入が実現し、ユーザは1つのモバイル端末(携帯電話機)を持ち歩いて複数の国や地域において使用できるようになった。現在は、第4世代/4Gに向けてLTE(Long Term Evolution)^{☆1}の普及が急速に進んでいる。

次世代モバイルネットワークの要求条件

2020年以降の実現が期待される次世代(第5世代/5Gに相当)のモバイルネットワークは、2020~2030年頃までを見据えた将来のモバイルサービス、システムに対する要求条件を満足する必要がある。要求条件の主な側面として、以下の3つが挙げられる。

★ サービス要求

今後、通信サービスへのユーザ要求の高度化・多様化を背景としたよりリッチなコンテンツを扱うサービス・端末の出現や、すべての「もの」が無線で接続されることによる各種情報の収集・監視と各種デバイスの制御・管理等を行う新サービスの出現が考えられ、その具体例を次に示す：

- パーソナル端末：個人の生活スタイルに密着し、生活の各場面に対応した多種・多様な機能・サービスを提供。
- 移動体搭載用通信モジュール：車、バス、電車等に搭載し、交通渋滞、車両コンディションなどの情報の収集、表示機能を提供。
- 家庭/家屋用通信モジュール：家電製品、家具、屋内設備等の遠隔制御機能、監視・セキュリティ機能等を提供。
- ウェアラブル端末：時計、装身具、衣服などに装着し、各種ヘルスケアサービス等を提供。

☆1 LTEは3GPP(3rd Generation Partnership Project, <http://www.3gpp.org/>)において国際標準仕様化されており、Release 8が最初のLTE仕様バージョンである。

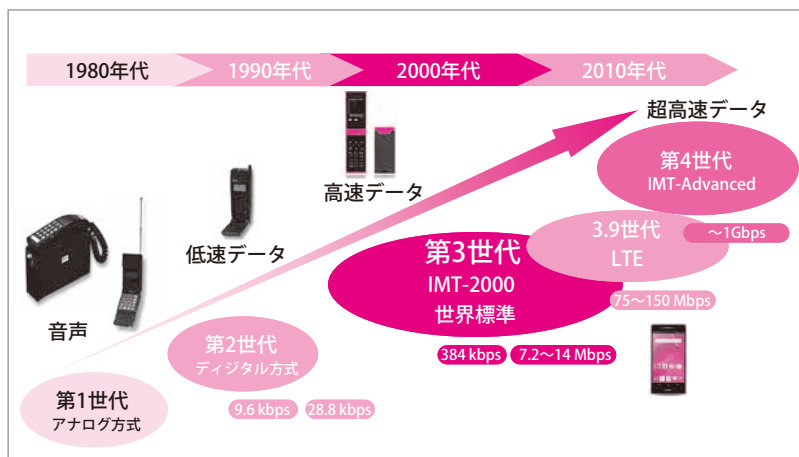


図-1 モバイルネットワーク発展経緯

- センサ搭載通信モジュール：工場、農場等におけるさまざまな管理機能、制御機能等を提供。
- 新型ディスプレイ／ヒューマンインタフェース搭載端末：高精細動画（4k/8k）視聴，ヘッドマウント型表示機能，感触通信，遠隔医療サービス等を提供。
- 通信教育システム，災害・事故時の救命・サポートシステム，各種クラウド処理サービス。

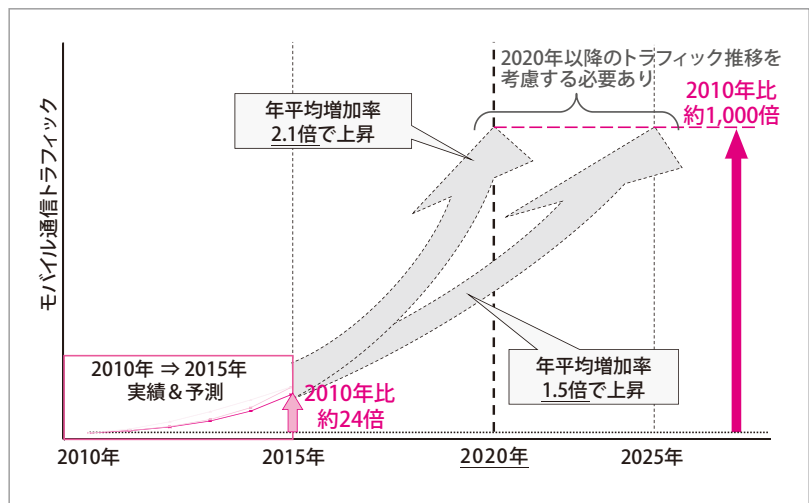


図-2 モバイル通信トラフィック予測（例）

★ システム能力要求

上述した新サービスの追加も背景とするモバイル通信トラフィックの増加は，2020年代に2010年比で約1,000倍に達するものと予測され（図-2），これを収容するシステム容量を確保していく必要がある。また，よりリッチなコンテンツを扱うサービスと端末の爆発的増加に対応するためには，LTE-Advanced（4G無線アクセス）^{☆2}のセル（基地局）あたりのスループットが標準仕様では1 Gbpsクラスであるのに対して，さらに1桁上の10 Gbpsクラスの超高速スループットをサポートすることや，100倍近い数の端末接続能力，1 ms以下の伝送遅延の実現などが求められるようになるものと考えられる。これら要求値に関して，国内においては，総務省の電波政策ビジョン懇談会等で示されるとともに，海外の主要ベンダや研究機関等においても，おおむね同様の値が議論されている。

★ コスト要求

モバイルネットワークの飛躍的な能力向上にあたっては，同時に無線アクセスおよびコアネットワークの各ノード装置と伝送路を含むネットワーク全体の設備コスト・運用コストを十分に抑えることが重要であり，そのための新しいネットワークアーキテクチャと諸技術の検討が不可欠である。

以下では，上述したサービス，システム能力，コストの各要求を満足すべく検討している次世代モバイルネットワークについて，無線アクセスネットワ

ークおよびコアネットワークに分けて紹介する。

次世代無線アクセスネットワーク

以下では，現在検討中のLTE-Advancedに後続する無線アクセス（以下，5G無線アクセス）と，5G無線アクセスを支えるモバイル光ネットワークからなる次世代の無線アクセスネットワークについて述べる。

★ 5G 無線アクセス

無線アクセスの能力向上アプローチと発展の方向

5G無線アクセスに向けたアクセス能力向上にあたっては，図-3に示す多様なターゲットと，それらを実現する3つのアプローチが考えられる。より進化した無線アクセス技術やMIMO伝送^{☆3}技術の採用による周波数利用効率向上，より高い周波数帯の採用による周波数帯域幅の拡張，およびより多くの基地局配置による高密度ネットワーク対応があり，これら複数のアプローチを併用しながら各ターゲットを達成していく。

☆2 LTEの発展形であるLTE-Advancedが，3GPP Release 10/11として標準仕様化され，現在さらに，後継のRelease 12/13の仕様化検討が進んでいる。

☆3 MIMO（Multiple-Input Multiple-Output）伝送は，送信と受信にそれぞれ複数のアンテナ（またはアンテナ素子）を用いて，無線信号を空間的に多重して伝送するもので，一般にアンテナ数に比例して伝送レートを増やすことができる。最近MIMO伝送は，モバイル通信のほか，無線LAN等にも採用されている。

また、今後の無線アクセスの発展においては、図-4に示す2つの方向がある。第1はLTEとのバックワードコンパチビリティを維持しながら継続的に発展させる方向、第2はLTEとのコンパチビリティを維持せず新しい無線アクセス技術の採用して大幅に性能向上させる方向であり、後者で使用する無線周波数帯には、モバイル用として新たな開拓が望まれる6 GHz以上の周波数帯も想定する。

基本アーキテクチャ

5G無線アクセスのセル構成に関する基本的なアーキテクチャは、ファントムセルコンセプト¹⁾をベースにセルレイヤを多層化した構造を考える。ファントムセルコンセプトは、図-5に示すように、従来のマクロセルに新たにスモールセルをオーバーレイ配置し、マクロセルとスモールセルで使用する周波数を変えることを特徴としており、マクロセルにおいてはより低い周波数（同一電力においてより遠くまで電波が届く）を用いることでカバレッジやモビリティを確保し、スモールセルにおいてはより高い周波数（1つの無線信号が占有する周波数帯域幅をより広く確保しやすい）を用いることで広帯域高速無線伝送を可能とする。また、マクロセルにおいてC (Control)-planeの接続リンクを確立し、同リンクを用いて呼制御やモビリティにかかわる制御を行う一方、スモールセルにおいてはデータに特化したU (User)-planeの接続リンクを確立し、各ユーザが環境に応じたベストエフォートのデータ通信を行う。

ここで、高い周波数を用いたスモールセルの導入による容量増大効果をより多く得るには、トラフィックが混雑している場所へ適切にスモールセルを配

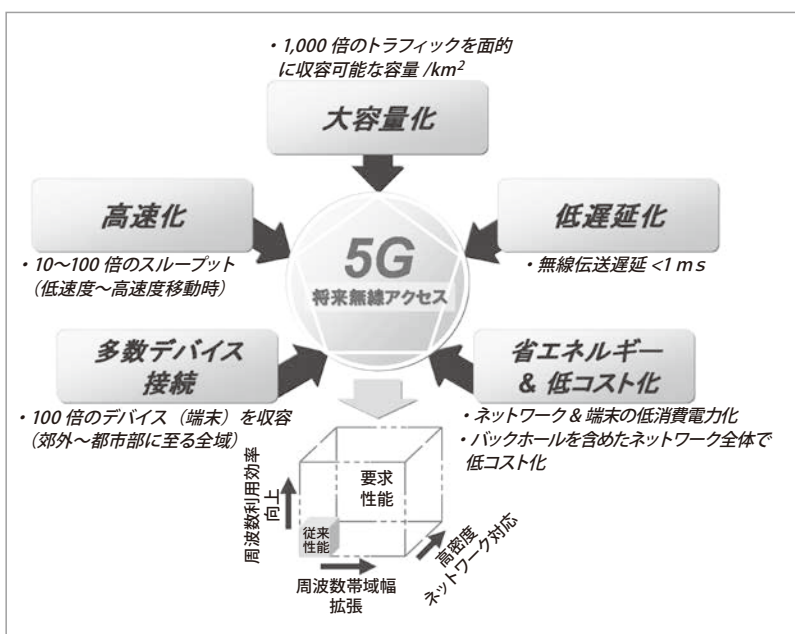


図-3 無線アクセスの能力向上アプローチ

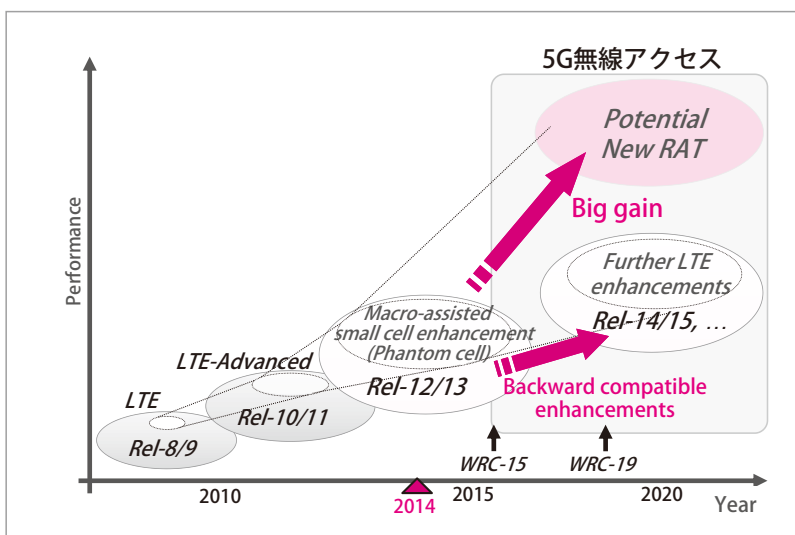


図-4 無線アクセスの発展の方向

置した上で、トラフィックのスモールセルへの積極的な誘導を行うことが必要となる。しかしながら、このようなスモールセルへの誘導は、システム全体の容量増大とユーザスループット・品質の向上に対して常に有利に働くとは限らない。たとえば、高速移動中のユーザがマクロセルから接続切換えしたスモールセルにきわめて短時間しか接続せず、再度マクロセルへ接続切換えとなった場合、かえってユーザスループットが低下し、かつ無駄なU-planeパス切換えによる制御負荷の増大を招く。ユーザのスループット要求の違いやセルの混雑度の大小に応じて

接続先セル種別を適切に選択しなかった場合も、システム全体の容量やユーザーグループの低下を招く。

このような問題を回避しつつ、既存もしくは将来追加される各種セルを柔軟かつ適切に収容するため、5G 無線アクセスにおいては、ファントムセルコンセプトを拡張して、サービスエリアを確保することを主たる目的としたカバレッジセルと、容量・スループットを増大させることを主たる目的としたキャパシティセルの2つのセル区別を導入する²⁾。図-6のように多くのセル種別が存在する5Gでは、マクロセルに加えて面的に高密度配置された一部の低SHF帯^{☆4}スモールセルもカバレッジセルに分類し、さらに、高い周波数を用いるセル（高SHF帯^{☆4}スモールセルやEHF帯^{☆5}スポット）、電車・バス等の移動体に設置

し車内の複数ユーザを同時収容するムービングセル、および、Wi-Fiスポット等のセルラ以外のスモールセルをキャパシティセルに分類する。その上で、C-planeはカバレッジセルのみで提供し、U-planeはカバレッジセルもしくはキャパシティセル内の各種セルの中から、ユーザの移動状況やサービス要求

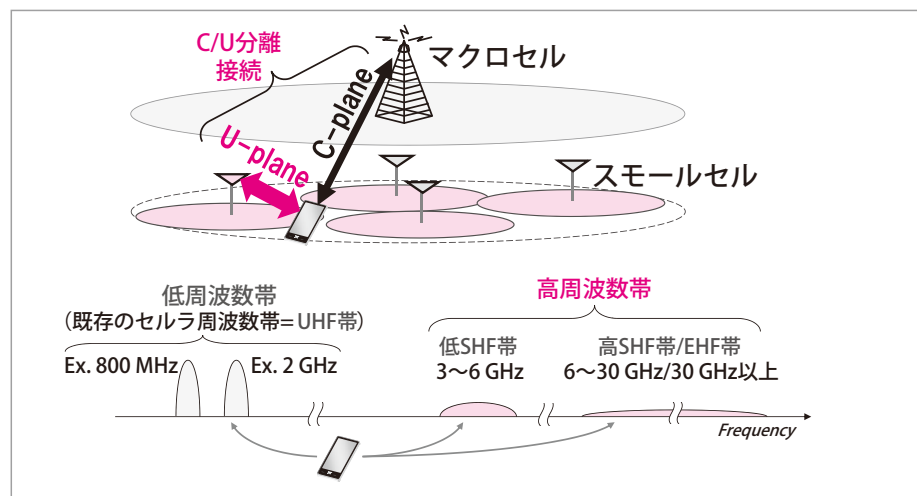


図-5 ファントムセルコンセプト

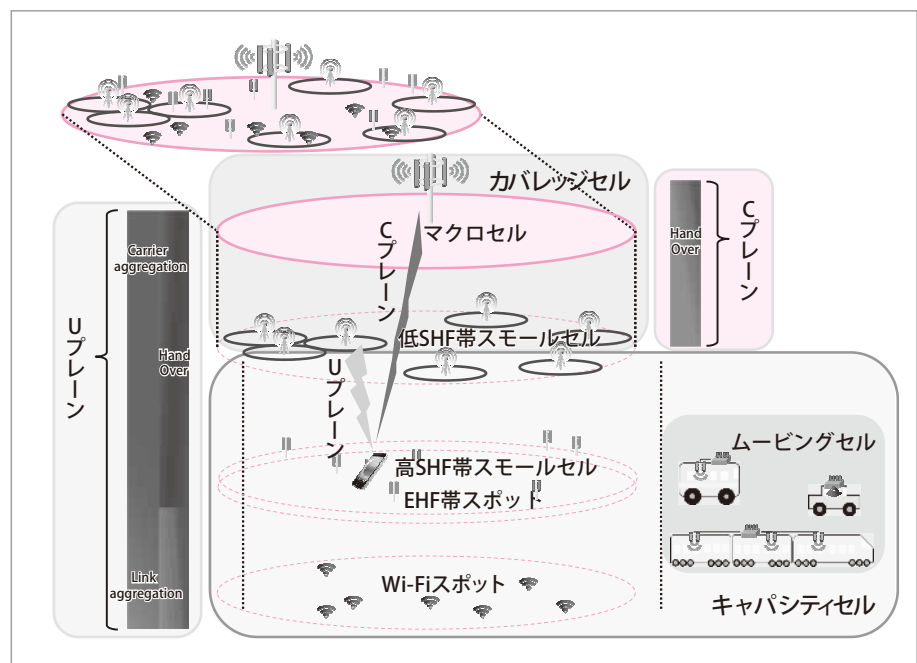


図-6 拡張ファントムセルコンセプト

に応じてあらかじめ設定されたポリシーに従って選択（リソース割り当て）されたセルで提供する。

マルチドメイン無線リソースマネジメント

ユーザが複数のセル種別配下にいるときは、ユーザ状況（移動状態／サービス状態）に応じて接続セル・リンクの切り換え（Hand over）、帯域集約（Carrier aggregation）、同時利用（Link aggregation）等の接続方法を、たとえば図-7に示すポリシーに基づいてマルチドメイン無線リソースマネージャが決定する。ここで、高速移動中のユーザに対しては、高SHF帯スモールセルが検出できたとしても、すぐに

☆4 周波数が3 GHzから30 GHzまでの電波をSHF (Super High Frequency) と呼び、波長は1～10 cmである。本稿では、SHF帯のうち3 GHz以上～6 GHz未満を低SHF帯、6 GHz以上～30 GHz未満を高SHF帯と呼ぶ。

☆5 周波数が30 GHzから300 GHzまでの電波をEHF (Extra High Frequency) と呼び、波長が1～10 mmであるため、ミリ波とも呼ばれる。

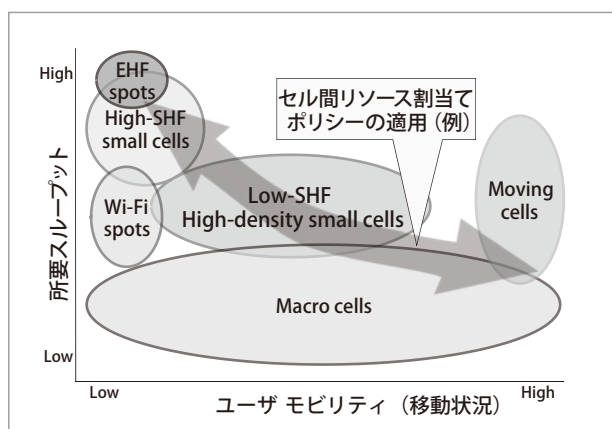


図-7 無線リソース割当てポリシー (例)

スモールセルの圏域外となることを考慮して、セル切換えを行わない。また、移動速度が遅く大容量ダウンロードを行うユーザに対しては、スループットが期待できる高 SHF 帯スモールセルへ切り換えるか、低 SHF 帯リンクを束ねて帯域集約を行うことで所要スループットを確保する。Wi-Fi 等の混雑状況が推定できないセルへの接続では、セルラへの接続も維持しながら同時利用することで、ユーザビリティの劣化を回避する。さらに、ユーザのムービングセルへの在圏/離圏を制御することで、多くのユーザの同時移動に伴う制御のバースト性やトラフィック集中等のグループモビリティ特有の問題を回避することができる。以上のようなマルチドメイン無線リソースマネジメントにより統括的に無線リソースを割当て・管理することで、ユーザの QoE (Quality of Experience) を満足させつつ、設置されたセルを最大限有効に活用する。

高周波数帯スモールセルの導入

LTE-Advanced の無線伝送帯域幅は最大 100 MHz であり、このとき、仕様上 1 Gbps クラスのスループットを達成できるのに対して、5G 無線アクセスにおいて 10 Gbps クラスの超高速スループットを実現するためには、数 100 MHz を超える帯域幅が必要となる。従来マクロセルにおける周波数は、800 MHz 帯や 2 GHz 帯等の UHF (Ultra High Frequency) 帯が主に用いられているが、数 100 MHz 超えの広帯域幅を確保するためには、従来セルラシ

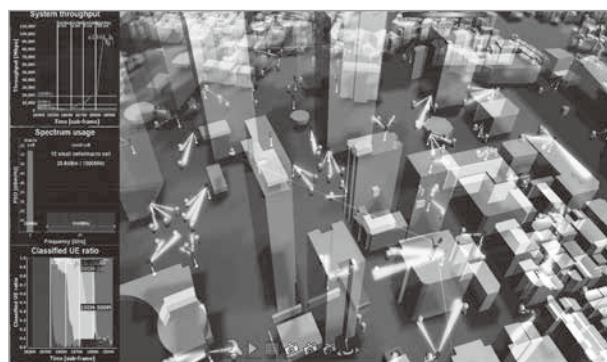


図-8 Massive MIMO 高周波数帯スモールセル導入

ステムへの適用は困難とされていた 6 GHz を超える周波数帯 (高 SHF 帯～EHF 帯) を新たに使用可能となるよう開拓していくことが求められる。ところが、周波数が高くなるにつれて電波の伝搬損失が大きくなり、十分なセルサイズを確保できなくなるため、その補償を行う Massive MIMO^{☆6}等の技術が重要となる。

高周波数帯スモールセルの導入効果については、たとえば図-8 に示す都市部へのセル展開を仮定したシミュレーションにより、2 GHz 帯で 20 MHz 帯域幅の伝送を行うマクロセルに対して 20 GHz 帯で 1,000 MHz 帯域幅の Massive MIMO 伝送を行うスモールセルをマクロセルのセクタ当たり 12 個オーバーレイ配置することで、システムスループットを 1,400 倍程度に向上可能であることが確認できる。

★5G 無線アクセスを支えるモバイル光ネットワーク

多数のスモールセルをオーバーレイ配置することを特徴とするファントムセルコンセプトにおいては、スモールセルを集約する基地局と多数スモールセル間を接続するネットワークを低コストで実現する必要がある。特に、スモールセル高密度化による容量増大効果を十分得るためには、スモールセル間の緊密な干渉制御が必要となることから、ネットワーク

☆6 Massive MIMO は超多数素子のアンテナを用いた MIMO 伝送方式である。高周波数帯ではアンテナ素子も小型化できるため、同一面積あたりの素子数を大幅に増やせるとともに、それらを用いてより鋭いビームを生成し、伝搬損失を補償することが可能である。

に対する遅延要件も厳しくなる。そのため、5G 無線アクセスがターゲットとする無線容量拡大を実現するためには、低遅延で大容量伝送が可能な光ネットワークを低コストで実現することが不可欠である。

スモールセルと上位局間の有線ネットワーク接続形態については、既存の基地局—上位局間と同様に IP (Internet Protocol) 伝送に基づくバックホールにより接続される形態が考えられる (図-9)。これに加え、既存の光張出し局と同様にスモールセルが無線送受信機能のみを備え、集約する基地局とスモールセルが光ファイバ接続される形態が考えられる (ここで、既存の光張出し局—集約基地局間で用いられる CPRI (Common Public Radio Interface) 仕様^{☆7}に基づき光ファイバ接続される形態を、以下ではフロントホールと称する)。さらに、スモールセルと上位局を無線により接続する形態も考えられる。

スモールセルと集約基地局が、CPRI 仕様に基づくフロントホールにより接続される構成は、IP 伝送に基づくバックホール接続構成と比較して低遅延での処理が可能のため、スモールセル間の緊密な干渉制御が可能になると考えられる。一方、フロントホール接続では、無線ベースバンド信号を量子化した後のデジタル信号が伝送されるため、バックホール接続と比較して多くのデジタル伝送帯域が必要となる。

CPRI 仕様に基づいて無線ベースバンド信号を光デジタル伝送する方法では、たとえば 20 MHz 帯域、2 × 2 MIMO の LTE 信号をサポートするために、2457.6 Mbps の伝送速度で光デジタル伝送が行われる。一方、より高速な無線信号伝送をサポートすることを目的に、最大 10137.6 Mbps の伝送速度が CPRI 仕様では規定されているが、本稿が対象とする 5G 無線アクセスでは、さらなる高速無線伝送をターゲットとしているため、それらをサポート可能なモバイル光ネットワークの実現が不可欠であり、

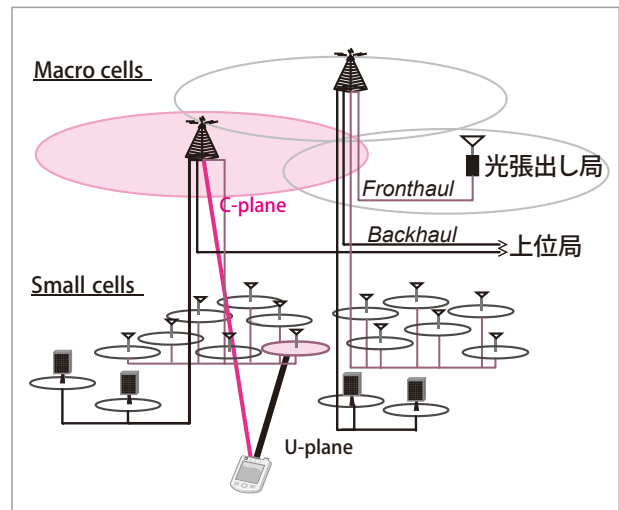


図-9 モバイル光ネットワーク

各種の検討が進められている³⁾。より高速な無線伝送をサポートするためには、波長多重技術 /WDM (Wavelength Division Multiplexing) の適用が最も有望と考えられるが、これに加え、スモールセルを集約する基地局と光張出しされた無線装置 (光張出し局) との間の機能分担を、従来とは異なる構成に変更することにより光伝送区間における所要帯域を削減する方法等が検討されている。

また、低コストでのモバイル光ネットワークの実現に向け、FTTH (Fiber to the Home) で利用されている PON (Passive Optical Network) 技術を活用する方法が検討されている。特に、広帯域な伝送容量を確保することを目的として、WDM を活用する PON 技術が検討されているが、波長チャネルの数だけレーザ光源などを用意する必要があるため、フロントホールにおいて必要となる帯域を削減し、必要最小限の波長チャネルで運用することが、コスト削減の観点からは望ましいと考えられる。

次世代コアネットワーク

上述した次世代の無線アクセスネットワークを効率的に収容する次世代のコアネットワークの要点とそれを実現する技術について述べる。

★ 次世代コアネットワークの要点

☆7 CPRI は無線基地局の Radio Equipment Control (REC) と Radio Equipment (RE) 間インタフェースのオープンな仕様化を目的として設立された産業協業団体 (<http://www.cpri.info/jp/>) であり、LTE-Advanced にも対応する ver.6.0 が最新仕様バージョンである。

設備・運用コストの低減

第1の要点は、増大していくトラフィックを設備・運用コストを十分に抑えて収容することである。課題となる側面をいくつか述べる。

現在のコアネットワークを構成するネットワーク機能は、キャリアグレードと呼ばれる高性能かつ高信頼な仕様を満たすため、テレコム向け専用ハードウェアで動作するソフトウェアにより実現されている。テレコム向け専用ハードウェアは特別な仕様のため、市場規模が小さく価格が高止まりする。また、技術進歩や市場競争により高性能かつ低コストのハードウェアが利用可能となっても、ハードウェアのみを交換することが困難であるためこのような利点を享受しにくい。IT系サービスでは、市場規模の大きい汎用ハードウェアを短いライフサイクルで更新していくことで、サービスを継続しながら設備コストを低減していく取り組みが進んでいる。長期間にわたって利用される移動通信サービスにおいても同様な利点を享受し設備コストを低減していくことが求められている。

次世代モバイルネットワークにおいては、多種多様なネットワーク機能を具備する装置が必要となる。一方、多様な機能を持つ装置は、それぞれ異なる運用方法が必要となる傾向にあり、運用コストの増大につながる。多種多様なネットワーク機能を効率的に運用するには、運用方法の定型化や集中管理、運用作業の一部自動化など運用稼働を削減する取り組みが必要となる。

高い信頼性を必要とされるコアネットワーク装置は、通常アクティブとスタンバイの2つの系を同時運用し、障害を検出すると即座にスタンバイに切り替えることでサービスへの影響を回避している。しかし、切り替わった状態のままでは、引き続き障害がサービスに影響を与えてしまう。そのため、障害を起こした系を迅速に修理するべく24時間対応可能な保全体制をとっているが、これは運用コストの相応の部分の占めており、運用コスト低減に向けた主要課題の1つとなっている。

予測不可能な需要変動への対応

第2の要点は、予測できない急激な需要変動への

対応である。

大規模な災害の発生により安否確認のための通信需要が急激に増大する、あるいは、ユーザ数が非常に多いスマートフォンアプリケーションの動作に起因して通信混雑が発生するといった事象が報告されている。移動通信サービスは、多くのユーザにとって欠くことのできない社会インフラであると認識されており、このような予測ができない事象が起こった際でも、通信サービスを提供し続けることが期待されている。しかし、予測困難な事象に備えるために、多大な設備をあらかじめ保有することは経済的合理性の面で問題となる。そのため、通信設備を有効に活用し予測困難な事象に事後的に対応できる柔軟性を持つアーキテクチャが望まれる。

★ ネットワーク仮想化技術の活用

ネットワーク仮想化とは

ネットワーク仮想化とは、ネットワークの資源と機能を組み合わせて迅速にサービスを構築する環境である。仮想化技術を用いた処理基盤は、この環境を実現する技術であり、データセンタ向けシステムの構成技術として研究開発され、昨今はIT系商用サービスの提供に広く活用されている。

仮想化技術を用いた処理基盤はハードウェアを直接駆動するオペレーティングシステムとアプリケーションソフトウェアを動作させる処理基盤としてのオペレーティングシステムとの中間に存在し、ハードウェアとアプリケーションソフトウェアの依存性を低減するための抽象化レイヤを提供する(図-10)。この仮想化技術を用いた処理基盤上に設定される仮想マシンは、ハードウェアで構成される純粋なマシンと異なり、設定の変更により各種能力の変更や生成・消去等の構成変更を行うことが可能である。本技術をモバイルコアネットワークの装置に適用するには、まず、専用のハードウェア上で動作するソフトウェアによる実現から、汎用ハードウェアで動作するソフトウェアによる実現に変えていくことが必要である。これにより、ハードウェアのみの更新が可能なアーキテクチャとなり、サービス

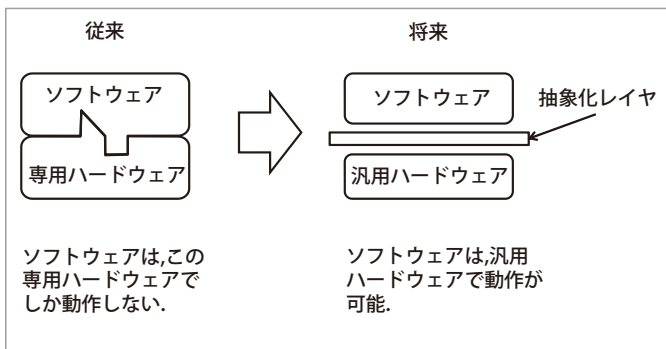


図-10 仮想化技術導入による効用

のライフタイムの間の技術進歩や市場競争による利点を享受することが可能となる。

仮想化技術による処理基盤は、コアネットワーク機能をソフトウェアにより管理制御する機能を具備している。この機能を拡張していくことにより、運用方法の定型化や集中管理、運用作業の一部自動化が可能となり運用稼働削減が期待できる。

仮想化技術による処理基盤を導入することにより、コアネットワークに柔軟性を実現する2つの特徴的な機能を活用できるようになる。ヒーリングとスケールリングである。

ヒーリングは、稼働しているコアネットワーク機能に障害が起こったときに、障害が起こった機能をシステムから切り離すと同時に、プールされているリソースから新しい仮想マシンを割り当て、障害が起こったコアネットワーク機能の代替を自動的に立ち上げる機能を言う。アクティブ系とスタンバイ系に二重化し、障害を検出すると即座にスタンバイ系に切り替える方式に、ヒーリングを組み合わせる方式が考えられる(図-11)。新しく割り当てた仮想マシンに障害が起こったコアネットワーク機能を代替させ、それを新たなスタンバイ系としてシステムに組み込むことで、保全要員の介在なく、アクティブ系とスタンバイ系が稼働する二重化の状態に復帰させることができる。この機能が具現化されれば、引き続き障害がサービスに影響を与える状態をオンラインで迅速に回避することが可能となるため、保全要員の体制を緩和し、運用コストの低減化が実現できると期待されている。

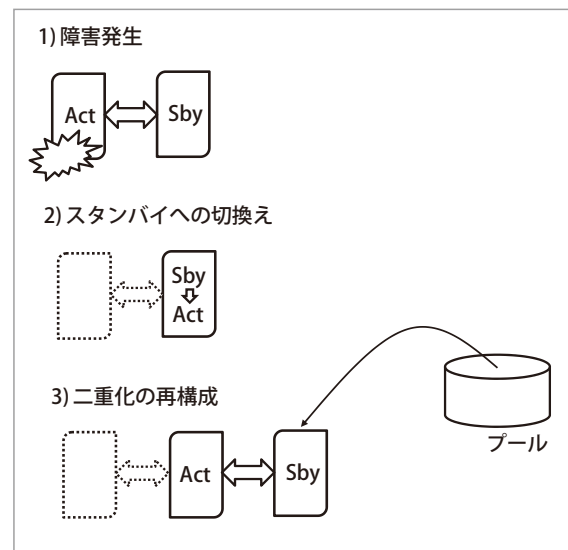


図-11 ヒーリング手順概要

スケールリングは、仮想化基盤においてある機能に割り当てられるリソースを増減させることである。あるサービスの需要が増加したとき、プールされているリソースから新しい仮想マシンを立ち上げ、そのサービスを実現する機能の容量を増やす。これをスケールアウトと呼ぶ。これは従来のネットワーク装置では困難なことである。また、逆にサービスの需要が減少したとき、割り当てられている一部のリソースを切り離し、容量を減少させる。これをスケールインと呼ぶ。この機能により、予想困難な需要の変動に対して柔軟に対応することが可能になる。

以上のように、ネットワーク仮想化は次世代コアネットワークの要点で述べた各種課題を解決できる主要な技術となる。

ETSI ISG NFV：産業界のイニシアティブ

このようなモバイルコアネットワークにおけるネットワーク仮想化の潜在性は、2011年ごろから多くの通信事業者の注目を集め、研究開発の中心課題となっていた。そこで、産業界の関係機関が集まり包括的に議論を行うことを目的として、欧州電気通信標準化機構 ETSI (European Telecommunications Standards Institute) のなかに、ネットワーク機能の仮想化(以降、NFV: Network Functions Virtualisation)を検討するグループである ISG NFV (Industry Specification Group for NFV) ^{☆8} が、2012年11月

に設立された。

世界の主要通信事業者7社の提案により発足したISG NFVは、今では世界中から220以上の企業が参加するイニシアティブとなっており、通信業界の次世代に向けた大きな潮流の1つとなっている。ISG NFVの対象は広くネットワーク機能としているが、モバイルコアネットワークへの活用は最優先課題となっており、世界中の移動通信事業者や移動通信機器ベンダが検討を牽引している。設立から1年後の2013年10月に

はNFVを定義する4つの基本文書（ユースケース、要求条件、アーキテクチャ、用語集）^{☆8}が完成した。NFVを実現する基本となるこれらの文書は、その具現化に必要となる課題と、その解決に向けたアクションを通信業界やIT業界に対して提示し、実用化開発の促進を図るものである。

図-12は、4つの文書の1つETSI GS NFV 002 V1.1.1が定義するNFVのアーキテクチャである。主な構成要素は以下の3つである。

- 中央左側に並ぶVNF (Virtualised Network Function) は、ネットワーク機能を提供するアプリケーションソフトウェアである。既存のネットワーク機能と同様に運用管理のためのエレメント・マネジメント (EM) を伴う。
- 下段のNFVI (Network Functions Virtualisation Infrastructure) は、VNFを動作させるのに必要な仮想化された処理基盤 (インフラ) である。汎用サーバやネットワーク機器などのハードウェア資源と、これらを仮想化する仮想化ソフトウェア (Virtualisation Layer) から構成される。

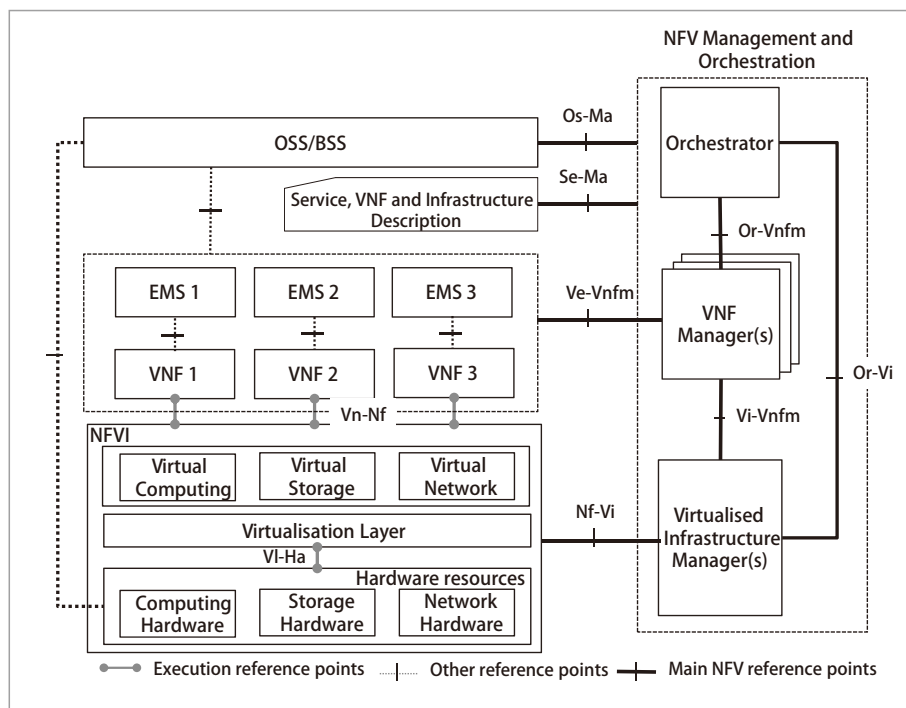


図-12 NFVアーキテクチャ

- NFV Management and Orchestration は、NFVIとVNFを管理制御する機能部である。3つの機能ブロックが定義されている。下から、NFVIを管理する仮想インフラ管理 (Virtualised Infrastructure Manager), VNFを管理する仮想ネットワーク機能管理 (VNF Manager), VNFを組み合わせて構築されるサービスの管理を行う Orchestrator である。

Orchestratorは、現行の運用管理システムであるOSS/BSS (Operations Support System / Business Support System) と連携し、仮想化されたネットワーク機能の管理制御を有機的に実施する。

このアーキテクチャは、NFVの実現に必要な機能とその関係を表現している。その設計指針には、さまざまな企業が参入しNFV準拠の製品が相互運用されるようなマルチベンダによるエコシステムの実現がある。

★ 産業界のイニシアティブを支える研究開発

ETSI ISG NFVにおける検討は、先導的な研究開発の成果により牽引されている。特にモバイルコアネットワークの分野では、これまで、多くの研究開発プ

^{☆8} ETSI ISG NFVの活動内容、メンバ、基本文書 (Group Specification) 等の詳細は <http://www.etsi.org/nfv> を参照。

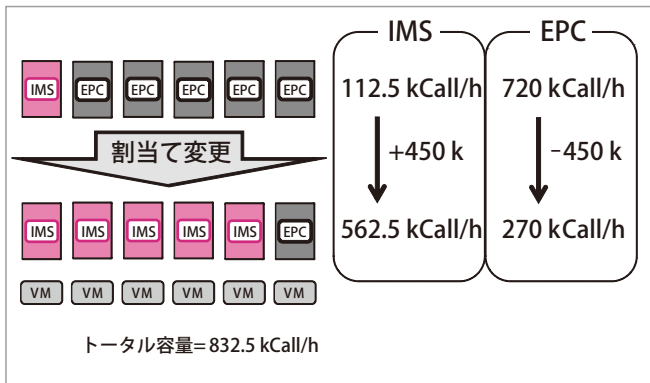


図-13 災害時のリソース融通例

プロジェクトが推進されてきた。以下では、そのような先導的な研究開発の取り組みをいくつか紹介する。

黎明期の取り組み

ネットワーク仮想化に至る取り組みの原点は、スタンフォード大学の Clean Slate Project で研究開発された SDN/OpenFlow 技術である。SDN/OpenFlow 技術は、対象をパケットの転送を行うスイッチやルータとし、制御ソフトウェアと転送ハードウェアの分離を図り、サービスの要件にあったネットワーク機能を迅速に構築するコンセプトを提唱した。この技術とデータセンタシステム向けに研究が進んでいた仮想化技術を組み合わせ、モバイルコアネットワーク機能をサービスとして提供する Network As a service を実現する取り組みが提案されている⁴⁾。

スケーリングによる通信混雑回避の研究開発

ネットワーク仮想化技術をモバイル網に適用した取り組みも行われている。音声通話の通信混雑を回避することを目的として、IMS システムを仮想化し、SDN (Software Defined Networking) /Openflow 技術でネットワーク制御を行うものである。

この研究開発では、特に東日本大震災等を想定した災害時をユースケースにして、IMS (Ip Multimedia Subsystem) および、EPC (Evolved Packet Core) の双方を仮想化し、互いのハードウェアリソースの融通（割当て変更）を可能にしている（図-13）。その仮想化された実験システムを利用し、EPC が提供するパケット通信向けのリソースを減少させ、災害時に必要とされる IMS が提供する音声

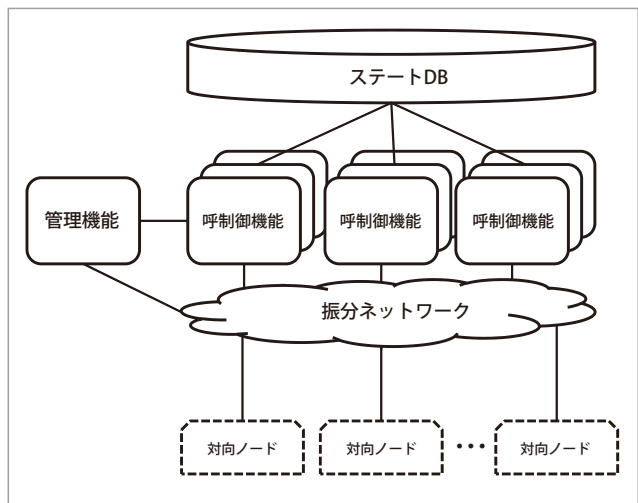


図-14 エラスティックコアアーキテクチャ

通信の容量を 5 倍 (112.5 kCall/h から 562.5 kCall/h) に増強する実験を行い、災害時の通信混雑を回避できることを確認している^{☆9)}。

エラスティックコア：効果的なスケーリング・ヒーリングの実現

これまでの取り組みは、既存の IMS や EPC を仮想化基盤に搭載するものであるが、既存のソフトウェアは仮想化を前提としておらずステートフルであるため、スケーリングやヒーリングによる効果を十分に享受できない。このため、仮想化技術の適用を前提として設計されるデータセンタ向けアプリケーションの考えを導入し、従来のテレコムアプリケーションのアーキテクチャに変更を提唱しているのがエラスティックコアの研究開発である。

エラスティックコア方式のアーキテクチャは、図-14 に示すようにステート DB、振分ネットワーク、呼制御機能、管理機能からなる。

- ステート DB：呼制御機能で使用するステートを共有するためのバックエンド DB。
- 振分ネットワーク：信号内の ID を用いて呼制御信号を振り分けるネットワークまたは装置。同一ユーザ／端末に関する継続信号は、收容するサーバへ振り分ける。

☆9 総務省による委託を受けて実施した「情報通信ネットワークの耐災害性強化のための研究開発」(平成 23 年度一般会計補正予算 (第 3 号)) の成果を含みます。

- 呼制御機能：信号処理部であり，ステートマシン制御を行う。
- 管理機能：呼制御サーバの障害・負荷を監視し，振分ルールや，サーバの増減，切換えを管理制御する。

これらのコンポーネントにより，ユーザのステート情報を外部に出し，呼制御機能部分をステートレス化することで，ユーザ呼処理を呼制御サーバ間で動的に振り替えることができる。そのため，スケールイン・アウト時の負荷のリバランスや，障害時や災害時にはサーバを切換える冗長処理が可能で，仮想マシンの増減をソフト処理で実施できる仮想化に適したアーキテクチャとなっている。

エラスティックコアのように仮想化基盤の機能を前提とし，容易にスケールアップやヒーリングの効果を享受するための取り組みはいくつかある。いずれもユーザのステート情報（セッション情報を含む）を仮想マシン間で共有または移動する機構を持つことで呼処理の継続性を保証することを目的としている。

N-ACT 構成のアプリケーションにおいて，ステート情報のバックアップを分散保持することで，スケールアップ時の呼処理継続を可能とする取り組み⁵⁾やセッション情報を効率よくバックアップし，障害・過負荷時に別サーバで復旧することで耐障害耐性を向上させる取り組み⁶⁾がある。また通信機器ベンダでも同様の取り組みがなされている。

次世代モバイルネットワークの実現に向けて

2020年以降のサービス要求，システム能力要求，コスト要求を満足する次世代モバイルネットワークの実現に向けて，無線アクセスネットワークにおいては，多種の周波数帯を用いる基地局セルをオーバーレイ配置可能とする高密度多層セル構造や，高周波

数帯を用いて広帯域無線伝送を可能とする Massive MIMO 伝送技術，より多くの基地局セルを経済的かつ柔軟に収容可能とするモバイル光ネットワーク等の検討が進められる一方，コアネットワークにおいては，ネットワーク仮想化を中心とした新しいアーキテクチャと技術の検討が進められていることを述べたが，これらは，既存のモバイルネットワークからのスムーズな移行も意識しながら段階的に実用化していくことによって，次世代のモバイルネットワークの着実かつ大きな進化を遂げられることが期待される。

参考文献

- 1) Ishii, H. et al. : A Novel Architecture for LTE-B, C-plane/ U-plane Split and Phantom Cell Concept, IEEE Globecom 2012 Workshop (Dec. 2012).
 - 2) 奥村幸彦, 他：将来無線アクセス・モバイル光ネットワーク～その1～/～その2～, 信学技報, RCS2013-231/232 (Dec. 2013).
 - 3) 寺田 純, 他：将来無線アクセス・モバイル光ネットワークにおける光アクセスシステムの検討課題, 信学技報, CS2013-49 (Nov. 2013).
 - 4) 今井和雄, 他：コンピューティング連携型ネットワーク仮想化とその実現技術, 信学論 B, Vol. J95-B, No.6, pp.707-716 (June 2012).
 - 5) 入江道生, 他：スケールアウトと柔軟な構成変更を実現するセッション制御サーバのクラスタモデル, 信学総大, B-6-11 (Mar. 2011).
 - 6) 北辻佳憲, 他：省電力とサービス継続を可能とする IMS の構成に関する考察, 信学総大, BS-2-2 (Mar. 2013).
- (2014年7月25日受付)

奥村幸彦 (正会員) okumurray@nttdocomo.com

NTT ドコモ先進技術研究所 主幹研究員。博士 (工学)。ワイヤレス方式の研究に従事。電子情報通信学会, IEEE 各シニア会員。

浅井孝浩 asaitaka@nttdocomo.com

NTT ドコモ先進技術研究所 主幹研究員。博士 (情報学)。ワイヤレス方式の研究に従事。電子情報通信学会, IEEE 各会員。

岩科 滋 iwashina@nttdocomo.com

NTT ドコモ先進技術研究所 主幹研究員。通信プロトコル, 移動通信ノードの研究開発に従事。

清水敬司 takashi.shimizu.yg@nttdocomo.com

NTT ドコモ先進技術研究所 主幹研究員。博士 (工学)。IP トラフィック制御, ネットワーク仮想化の研究に従事。電子情報通信学会, IEEE, ACM 各会員。

2 コンシューマデバイスとパーソナルデータの利活用および保護の動向

石川憲洋 松前恵環

■ 駒澤大学グローバル・メディア・スタディーズ学部

コンシューマデバイスの動向

コンシューマデバイスの動向として、急速に普及が進むスマートフォンに代表されるスマートデバイスと、近年、大きな注目を集めているウェアラブルデバイスの動向について述べる。一方で、個人が所有するさまざまなデバイスから大量のパーソナルデータがインターネット上で収集され、ビッグデータ解析技術の進展により、さまざまな分野で利活用を図る動きが急速に進展している。このような状況を受け、パーソナルデータの保護にかかわるルールや指針の見直しが国際的に進みつつある。個人情報保護法の現状とその改正に関する議論動向について、諸外国の動向も含めて概説する。

★スマートデバイスの動向

文献1) で述べた通り、2012年から2014年にかけては、コンシューマデバイスおよびコンシューマサービスにとって非常に大きな変革期であった。

スマートデバイスの代表であるスマートフォンの世界での売り上げ台数は、2013年に初めて、フィーチャーフォンと呼ばれる従来型の携帯電話の売り上げ台数を上回った。2013年の世界の携帯電話の売り上げ台数は約18億台であり、そのうち、約10億台がスマートフォンであった。米国、欧州、日本などの先進国を中心に普及していたスマートフォンは、2013年以降、ラテンアメリカ、中国、インドなどの市場で、急速に普及している。世界的に見ても、携帯電話のフィーチャーフォンからスマートフォンへの移行はほぼ決定的であり、この事実は、サ

ービス事業者、携帯電話事業者、携帯端末メーカーのビジネス戦略に大きな影響を及ぼしている。

まず、ビジネスモデルにおいては、日本のiモードサービスに代表されるサービス、移動通信ネットワーク、携帯電話を一体化した垂直統合型のビジネスモデルから、サービス事業者、携帯電話事業者、携帯電話メーカーが独立してビジネスを展開する水平分業型のビジネスモデルへと移行することは決定的であると考えられる。

サービス事業者の観点から見た場合、キャリアフリー、デバイスフリーなサービスを世界規模で展開することが可能となるため、OTT (Over The Top) と呼ばれるサービス事業者にとって、大きなビジネスチャンスが生まれている。その代表は、ソーシャルメディアではFacebook、無料電話・メッセージングなどのコミュニケーションサービスではLINEである。これらのサービスでは、「ネットワーク外部性」の効果が大きく働くため、勝ち組となったサービス事業者が急速にユーザ数を増やしている。特に、LINEは、2011年11月のサービス開始以来、約4年間で世界で5億人以上のユーザを獲得したことは、驚異的である。LINEのコミュニケーションサービスは、携帯電話事業者のサービスと競合するため、両者の関係に注目が集まったが、近年では、携帯電話事業者はLINEなどのサービス事業者を、ネットワークのトラフィックを増やす存在として捉え、両者の連携する動きが目立つようになりつつある。

次に、携帯電話事業者の動向であるが、サービス事業者、携帯電話メーカーが独立したビジネス展開、

サービス展開を行うようになった結果、サービス事業者および携帯電話事業者に対する影響力は確実に弱まりつつあり、携帯電話事業者間の競争は、料金競争、ネットワーク競争が中心になりつつある。料金競争に関しては、サービス事業者の台頭により、音声収入の減少、パケット収入の増加が顕著になり、国内では、音声通話料の定額化、海外では、パケット通信料の従量制課金の動きが見られるようになった。ネットワークに関しては、スマートフォンの普及によりパケットトラフィックが急速に増加したため、LTE (Long Term Evolution)、LTE-Advanced、VoLTE (Voice Over LTE) などの新技術が導入され、いかに繋がりやすい高速なネットワークを経済的に構築するかが競争のポイントになると考える。

残された大きな課題は、SIM (Subscriber Identity Module) ロックの扱いであるが、長期的には、SIM ロックフリーとすることにより、ユーザの携帯電話および携帯電話事業者の選択の自由度を高めるとともに、携帯電話の販売形態が多様化することで、携帯電話メーカー間の競争を促進させ、スマートフォンの低価格化を招く可能性がある。また、仮想移動体通信事業者 (MVNO : Mobile Virtual Network Operator) を含めた携帯電話事業者間の競争も促進され、結果として、ユーザに携帯電話および利用料金の低価格化も含めたさまざまなメリットが生じるのではないかと考える。

第3に、携帯電話メーカーの動向であるが、メーカー間の競争の激化、スマートフォンのコモディティ化により、各社ともに厳しい状況にある。事実、iPhone でスマートフォンのマーケットを開拓し、iPad などによるマルチデバイス化、iTunes Store などのコンテンツ配信ビジネスで先行するアップルと、アンドロイドスマートフォンで成功したサムスンを除くとほとんど各社が赤字の状況にある。特に、アップルを除くと、各社ともにアンドロイド OS を提

供するグーグルの支配下にあり、高速化、低価格化などの要因以外に差別化が困難な状況にある。したがって今後、メーカー各社は、アンドロイド OS から独立し、新たなスマートフォン向け OS の利用、開発などに着手して端末の差別化を図る可能性が高い。

最後に、もう1つの代表的なスマートデバイスであるタブレット端末の動向について述べる。

2012年から68%増え、2013年のタブレット端末の世界での販売台数は、約2億台となった。アンドロイドタブレット端末のシェアが60%を超えた一方で、アップルのiPadシェアは36%となり、アンドロイドタブレット端末がアップルのiPadのシェアを大きく超えたが、アンドロイドタブレット端末には、中国、台湾などの多数のベンダが低価格のタブレット端末で参入しているため、1社のシェアでは、アップルが他社を大きく引き離している。一方で、2013年のパーソナルコンピュータ (PC) の世界での販売台数は、約3億台であり、前年比で約10%の減少となっている。この傾向が続くと、2、3年でタブレット端末の販売台数がPCの販売台数を抜くことが予想される。まだ、シェアは数%であるが、マイクロソフトのWindowsタブレット端末は、Windows PCと互換性が高く、ビジネス用途などを考えた場合、今後、大きくシェアを伸ばす可能性もあると考える。

「ポストPCの時代」などと呼ばれるようになったが、PCがなくなることがなく、むしろ、ノートPCとタブレット端末の境界が非常に曖昧になることが予想される。タブレット端末のキラーアプリケーションとしては、ゲーム、電子書籍などが考えられるが、今後、キーボード付きのタブレット端末も多数販売される見込みであり、性能の向上とともにビジネス用途にも幅広く活用されることが予想される。その結果、個人ユーザの用途の大部分はタブレット端末でカバーできることになり、多機能型のスマー



トフォンのマーケットが減少し、電話、キャリアメールを含むメッセージング、カメラ機能、Felicaを含むNFC（Near Field Communication）機能などに機能を限定した従来型のフィーチャーフォンに近い携帯電話がシェアを伸ばす可能性も考えられる。

★ ウェアラブルデバイスの動向

グーグルのGoogle Glassの発表以来、ウェアラブルデバイスが非常に大きな注目を集めている。Google Glassは、すでにサービスを終了している「セカイカメラ」などで注目を集めた拡張現実（AR：Augmented Reality）を実現したメガネ型デバイスであり、2014年内に一般ユーザ向けの販売が予定されている。ナビゲーション機能などが主な用途として想定されているが、スマートフォン（アンドロイド、iPhoneなど）とBluetoothを利用して連携することにより、さまざまなアプリケーションが開発されることが期待されている。一方で、カメラ機能を持つため、プライバシー問題に関する懸念も表明されている。

すでに市販されている目に装着するウェアラブルデバイスとしては、ヘッドマウントディスプレイ（HMD：Head Mount Display）がある。主に非透過型のデバイスで、3D対応映画の視聴、3D対応ゲームの画面などが主要な用途として利用されている。

スマートウォッチと呼ばれる腕時計型のウェアラブルデバイスも、数社から販売されている。スマートフォンと連携して、通話機能、メール・ソーシャルメディアの到着情報の表示、カメラ機能などを持つデバイスが多いが、腕時計はファッション性が高い歴史のある伝統的なデバイスであり、画面サイズも小さいことから、キラアプリケーションがない現状では、コアなユーザ層以外には、急速な普及は難しいと考える。腕時計とスマートフォンなどとの機能分担を見直し、たとえば、腕時計には、FelicaなどのNFC機能のみを搭載し、その他の残高表示機能、チャージ機能などをスマートフォン側で実現すれば、一見、普通の腕時計がSuicaなどと同等の機能を持つことになり、一般ユーザにも魅力的なデ

バイスになる可能性があるのではないかと考える。

現在、ウェアラブルデバイスの代表的な応用分野として考えられているのは、ヘルスケア分野である。ヘルスケアのためのウェアラブルデバイスは、大部分がリストバンド型で、歩数、消費カロリー、移動距離、睡眠時間・睡眠状態などが測定可能で、Bluetoothなどを利用して、スマートフォンに情報を転送することで、スマートフォン上で、日々の活動状況を把握することが可能となっている。

しかしながら、ウェアラブルデバイスの日本における2014年度の市場規模は、100万台程度と予想され、まだ、大きな市場に育っているとは言いがたい。2014年内にアップルからもウェアラブルデバイスの発表が予想されており、本格的な高齢化社会を迎えヘルスケア市場に大きな期待が寄せられている。ウェアラブルデバイス、スマートフォン、ヘルスケアサービスがユーザに負担をかけることなくシームレスに連携してユーザビリティを向上させ、ユーザの健康管理により的確なアドバイスをするサービスを提供することにより、今後、大きな市場へと発展することが期待されている。

コンシューマデバイスからのパーソナルデータ^{☆1} 収集

現在、個人が所有するさまざまなデバイスから大量のパーソナルデータがインターネット上で収集され、ビッグデータ解析技術の進展により、さまざまな分野で利活用が進められている。

スマートフォン、タブレット端末などのスマートデバイスだけでなく、カード型デバイス（磁気ストライプカード、非接触型ICカードなど）、PC、最近注目を集めているウェアラブルデバイスまで、幅広いデバイスを利用して、下記の一般的な会社員の1日の行動を例として、どのようなパーソナルデー

☆1 「パーソナルデータ」とは、「個人情報保護に関する法律」2条の「個人識別性」の要件を必ずしも満たさない情報をも含む、「個人に関する情報」の総称として用いられる用語である。詳しくは、次章を参照。

タが収集可能かについて述べる。

朝起きるとヘルスケアデバイスを装着した後、スマートフォン、PCなどで、メール、ニュースなどのWebサイトをチェックする。必要なメールに返信を送った後、TVを観ながら朝食を取る。スマートフォンで1日のスケジュールを確認した後、電車を利用して、会社に出勤する。途中で、コンビニなどに立ち寄り、新聞などを購入する。会社での仕事の後、スマートフォンで友人などに連絡を取り、居酒屋などで食事を楽しんだ後、帰宅する。帰宅後は、PCなどで、オンラインショッピングなどを行う。

このような典型的な1日の生活から収集可能なパーソナルデータを表-1にまとめる。表-1に示すように、さまざまなデバイスを利用して、個人の1日の大部分の行動履歴を把握することが技術的に可能となっている。収集されたパーソナルデータの利用目的は、個人利用、ターゲティング広告を含むマーケティングなどのビジネス利用に大別できる。

たとえば、ヘルスケアデバイスから収集されたパーソナルデータは、現状では健康管理などの個人利用が大部分であり、プライバシー保護の観点からは大きな問題はないと考えられる。一方で、ビジネス利用に関しては、パーソナルデータの第三者提供など、さまざまな課題を抱えている。たとえば、A社とB社が提携して、A社が収集したパーソナルデータとB社が収集したパーソナルデータを相互利用するようなケースも想定される。セキュリティと比較しても、パーソナルデータを含むプライバシーを技術的に保護することは困難な面もあるため、法律によるプライバシー保護も重要である。

次章では、コンシューマデバイスからのパーソナルデータ収集および利用を巡る法の現状と近時の議論動向について、諸外国の動向をも含めて概観する。

パーソナルデータに関する法の現状と近時の議論動向

パーソナルデータを含む大量のデータが処理される現状を受け、近時、個人情報の保護にかかわるル

デバイス	個人識別情報	パーソナルデータ
ヘルスケアデバイス	ユーザID	1日の消費カロリー、歩数、睡眠時間など
TV	製造番号など	TVの視聴履歴
PC	ユーザID、クッキーなど	Webサイトのアクセス履歴、メールの送受信履歴、オンラインショッピングの履歴など
スマートフォン	ユーザID、端末IDなど	1日のスケジュール、メール送受信履歴、通話履歴、位置情報の履歴
非接触型ICカード(定期券など)	カードID	電車の乗降車履歴
磁気カード(会員カードなど)	カードID	コンビニなどでの購買履歴、飲食店での飲食履歴など

表-1 収集可能なパーソナルデータ

ールや指針の見直しが国際的に進みつつある。これまで、プライバシー・個人情報の保護に関する国際的な指針として中心的な位置を占めてきたのは、1980年に経済開発協力機構(OECD: Organization for Economic Co-Operation and Development)が公表した「OECD プライバシーガイドライン」(OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data)であり、この中で提示されたいわゆる「OECD 8原則」は、各国のプライバシー・個人情報保護法制の基礎となってきた。本ガイドラインは、スマートフォンをはじめとする技術の進歩、国際的なデータ流通、組織や個人の活動の変化等の現代的局面に即して、2013年に改正されている。また、これまで1995年のいわゆる「EU個人データ保護指令」(Council Directive 95/46/EC of the European Parliament and of the Council of 24 Oct. 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data)を中心に個人情報の保護のための取り組みを進めてきた欧州連合(EU: European Union)でも、2014年に個人データ保護規則提案が欧州議会において可決され、違反の際の厳しい制裁金制度の導入や、データ主体が自らに関するデータの消去を管理者や第三者に求めることができる消去権等の定めを含む、規則の制定に向けた動きが加速している²⁾。米国においても、連邦取引委員会(FTC: Federal Trade Commission)

を中心に、消費者のプライバシー保護のための取り組みが進んでいるほか、2012年にはホワイトハウスにより、消費者のプライバシー保護のための7つの権利を謳う、「消費者プライバシー権利章典」(Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy) が公表されるに至っている。

こうした諸外国の動向やパーソナルデータ収集を巡る状況の変化を受け、日本においても、パーソナルデータに関する法制度の整備に向けた検討が進められている³⁾。2013年6月14日に閣議決定された「世界最先端IT国家創造宣言」を受けて開催された「パーソナルデータに関する検討会」では、パーソナルデータの利活用に関するルールの検討が行われ、2013年12月20日には、「パーソナルデータの利活用に関する制度見直し方針」(以下、「見直し方針」という)が決定された。これを踏まえ今般、2014年6月24日には、「パーソナルデータの利活用に関する制度改正大綱」(以下、「大綱」という)が取りまとめられ、2003年に制定された「個人情報保護に関する法律」(以下、「個人情報保護法」という)の改正を視野に入れた具体的な方向性が示されるに至っている。

以下、パーソナルデータに関する日本法の現状と近時の議論動向について、主要な論点を中心に、海外の動向と比較しながら概観する。

★ 保護すべきパーソナルデータの範囲

まず重要なのが、保護すべきパーソナルデータの範囲の問題である。そもそも個人情報保護法における保護の対象となる「個人情報」とは、個人情報保護法2条において、「生存する個人に関する情報であって、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの(他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む)」と定義されている。しかし近時、かかる個人識別性の要件を必ずしも満たさないと解

される可能性のある情報をも含む、いわゆるパーソナルデータ、すなわち、個人の行動・状態等に代表される個人に関する情報⁴⁾が収集・利用される場面が増大している。こうした情報は、仮に「個人情報」に該当しないとしても、Web上での行動履歴や位置情報など、使い方によってはプライバシーをはじめとする個人の権利利益に脅威を及ぼす危険性がある。そのため、情報処理技術の発展したビッグデータ時代においては、必ずしも特定の個人を識別しないが、個人に結びつく可能性のあるパーソナルデータを広く保護の対象に含める必要性が指摘されている。

この点、EUでは、適用対象となるパーソナルデータの範囲を実質的に判断する傾向が強く、たとえば位置情報についても、通常パーソナルデータに該当すると解されている⁵⁾。保護すべきパーソナルデータの範囲について柔軟な解釈をとる傾向は、近時米国においても見られるようになってきている。前出の「消費者プライバシー権利章典」においては、権利章典が適用される範囲に関して、特定の個人と結び付き得るデータであればよいとされ、「特定のコンピュータその他のデバイスと結び付けられるデータ」をも含むとされている。

日本でも、総務省「パーソナルデータの利用・流通に関する研究会」が2013年6月に公表した報告書において、『「実質的個人識別性」(プライバシーの保護というパーソナルデータの利活用の基本理念を踏まえて実質的に判断される個人識別性)』を判断基準として、保護すべき情報の範囲を決すべきであるという考え方が示された。これを受けた「見直し方針」においてもこの考えが踏襲されており、保護されるパーソナルデータの範囲を実質的な個人識別性を基準として判断することが明確にされている。

★ パーソナルデータの利活用のための措置

かかるパーソナルデータの保護の必要性が高まる一方で、パーソナルデータを有効に利活用することによって、新たなサービスや産業の発展が見込まれるという側面もある。しかし、そもそも個人識別性

は程度の問題であり、保護されるべきパーソナルデータの範囲を一義的に画定することは難しい。このため、実務上、事業者等がパーソナルデータの利活用について委縮してしまうという、「利活用の壁」が存することが指摘されている。2013年に世間の注目を集めた、JR東日本のSuicaの乗降履歴販売の問題は、こうしたパーソナルデータを巡るその保護と利活用の2つの要請の衝突が顕著に表れた例であったとも言える。JR東日本は、氏名や連絡先を除く情報は個人情報保護法にいう「個人情報」に当たらないとして、乗客の乗降履歴等を販売していたものであるが、これに対しては、多くの苦情や批判が寄せられている⁶⁾。

そこで、こうしたパーソナルデータを円滑に利活用するためのさまざまな措置が検討され始めている。近時とりわけ注目を集めているのが、パーソナルデータの非識別化措置であり、中でも米国FTCが示したデータの非識別化のための3要件は、日本の議論にも影響を及ぼした。これは、事業者が、①データの非識別化のための合理的な措置を講じていること、②非識別化したデータを再識別化しないことを公に約束すること、③非識別化したデータを第三者に提供する場合には、提供先が再識別化を行うことを契約で禁ずること、という3つの要件を満たした場合には、個人に結び付けられ得る情報に該当せず、プライバシー保護ルールの適用対象外となるという考え方である⁷⁾。

こうした動きを受け日本でも、「パーソナルデータ検討会」の下に設置された「技術検討ワーキンググループ」を中心に、非識別化措置についての議論が重ねられてきた。2014年に公表された「大綱」においては、「パーソナルデータの利活用を促進するための枠組み」として、『個人データ』を特定の個人が識別される可能性を低減したデータに加工したものである⁸⁾については、本人の同意を得ずに第三者提供や目的外利用を行うことを可能とするなど、パーソナルデータの円滑な利活用のための措置が提示されている。

もっとも、非識別化措置にもさまざまなレベルが

あり、場合によっては再識別化が可能なケースもあり得ることに照らせば⁸⁾、どのような条件の下で「特定の個人が識別される可能性を低減したデータに加工した」と評価されるのか、今後、具体的な加工方法や要件などを慎重に検討する必要がある。

★ 第三者機関の設置

いま一つ重要な論点が、パーソナルデータの保護についての監視・監督、および紛争処理等を行う独立の第三者機関の設置を巡るものである。前出「EU個人データ保護指令」は加盟国にその設置を求めており、EUでは「プライバシー・コミッショナー」等と呼称される第三者機関が、個人データ保護法の執行やパーソナルデータ保護に関する新たな課題への統一的な対応等の場面で、重要な役割を担っている。また、2013年の「OECDプライバシーガイドライン」においても、プライバシー執行機関の設置が規定されている。これに対し、民間・公的部門双方を規律する包括的な個人情報保護法を持たず、分野ごとの個別の法規制と自主規制との組合せによってプライバシー・個人情報の保護を図ってきた米国には、EUにおける同義の第三者機関は存在しない。もっとも、FTCが、FTC法5条（「不公正又は欺瞞的な（unfair or deceptive）取引行為・慣行は違法である」）⁹⁾に違反する民間事業者を調査・訴追し得るという仕組みが用意されている。

他方、日本の個人情報保護法ではかかる独立の第三者機関の設置は規定されておらず、個人情報が取り扱われる分野ごとに、各省庁の主務大臣が法の執行を行うという仕組みが採用されている。これについてはかねてから、パーソナルデータの取扱いに関する問題全般に機動的に対応する必要性や、諸外国の制度との整合性といった観点から、第三者機関設置の必要性が指摘されてきた。2013年に成立した「行政手続における特定の個人を識別するための番号の利用等に関する法律」において、ようやく職権行使の独立性が謳われた特定個人情報保護委員会の規定が置かれるに至ったが、同委員会の権限はあくまで番号制度に関するものに限られている。そこで

今般の「大綱」では、同委員会の機能や権限を拡張し、パーソナルデータの取扱い一般に関する監督や苦情処理、基本方針の策定、国際協力等の業務を行う独立の第三者機関の設置・体制整備を行うことが提案された。

「大綱」では、先に挙げたパーソナルデータの範囲や非識別化措置に関して、民間事業者による自主的なルールを策定を重視しているが、こうしたルールの実効性の担保や、次項で述べる越境データ保護のための国際執行協力の観点からは、第三者機関に期待される役割がきわめて大きい。その機能や権限、体制および人材確保等さまざまな面において、実効性のある機関設計が求められる。

★ 越境データの保護

グローバル化が進み、パーソナルデータを含む大量のデータが国境を越えて流通する現代においては、かかる越境データの保護のあり方も喫緊の課題である。越境データの保護のための代表的な方策としては、「EU 個人データ保護指令」で採用されている、一定の場合を除き、個人情報の保護について十分なレベルに達していない国への移転を禁ずるといったものが挙げられる。今般の「大綱」ではかかる仕組みも念頭に、パーソナルデータが外国事業者へ移転される場合に対応すべく、個人情報保護法の域外適用のほか、他国への越境データの移転を制限する枠組みの採用等が提案されている。なお、「EU 個人データ保護指令」における充分性の基準との関係では、日本の個人情報保護法制はいまだ十分であるとの認定を受けていないところ、EUをはじめとする海外から日本へとパーソナルデータが移転されることに障害が出ぬよう、国際的な整合性を図ることが望ましい。

また、国境を越えて法執行を行うための、国際的な執行協力の仕組みの構築も進んでいる。OECDでは2007年に、「プライバシーを保護する法の執行における越境協力に関する理事会勧告」(OECD Recommendation on Cross-border Co-operation in the Enforcement of Laws Protecting Privacy)が

採択され、2010年には「国際プライバシー執行ネットワーク」(GPEN: Global Privacy Enforcement Network)が設置された。また2010年には、アジア太平洋経済協力(APEC: Asia-Pacific Economic Cooperation)の「APEC 越境プライバシー執行のための協力取決め」(APEC Cross-border Privacy Enforcement Arrangement)も発効している。国内法の整備だけでなく、前項で言及した第三者機関の設置を前提に、こうした国際的な体制への積極的な参画も求められる。

★ 法改正に向けた課題

以上、主要な論点に焦点を当ててパーソナルデータを巡る議論動向を概観した。今後法改正に向けた具体的な検討を行う過程でとりわけ留意すべきは、パーソナルデータの保護と利活用とのバランスの確保である。産業界からは利活用の促進を求める声も強いが、あくまでプライバシー・個人情報の保護を尊重しつつ利活用を進めるための制度整備、および実務慣行の確立が望まれる。

また、諸外国の制度との整合性を視野に入れることも肝要である。グローバル化が進み、パーソナルデータが国境を越えて流通する現代において、プライバシー・個人情報保護の問題はもはや一国内に止まるものではなくなっている。かかる状況下では、プライバシー・個人情報の保護のための国際的な協力が不可欠であり、国際的な動向に常に目を配りながら制度形成を行っていくことが、プライバシー保護の観点からも、また、ビジネスの展開という観点からも、望ましい。今後も引き続き、海外の動向を注視しながら、また同時に、プライバシーという概念が有する文化的・社会的背景に照らし、日本固有の状況にも配慮しつつ、政策検討を行っていくことが求められる。

今後の展望

コンシューマデバイスがコモディティ化し、一般ユーザにも広く普及することで、生活の利便性が向

上する一方で、コンシューマデバイスから収集される大量なパーソナルデータがインターネット上で収集され、さまざまな分野で利活用を図る動きが急速に活発化している。また、収集されたパーソナルデータが漏洩し、売買される事例も数多く報告されている。

ユビキタスコンピューティングなどの研究分野では、ユーザの位置情報取得などの研究が数多く発表されているが、今後は、悪用されるリスクも考慮し、ユーザのプライバシー保護も視野に入れて研究開発を進める必要がある。

顧客情報などのプライバシーデータを扱う情報システムに関しても、“Privacy By Design”の原則に則り、システムの設計、開発から運用に至るまで、プライバシー管理を徹底する必要があると考える。

現在、個人情報保護法の改正の検討が進められているが、設置が想定されているEUの「プライバシー・コミッショナー」に相当する第三者機関が、起こり得るさまざまな事例に対して有効に機能できるように、長期的な視野に立って、プライバシー保護の原理・原則を明確化した、骨太な法改正を期待したい。

参考文献

- 1) 石川憲洋：コンシューマサービスの今後の展望，情報処理，Vol.53, No.10, pp.1047-1050 (Oct. 2012)。
- 2) European Commission, Progress on EU Data Protection Reform Now Irreversible Following European Parliament Vote (2014), http://europa.eu/rapid/press-release_MEMO-14-186_en.htm

- 3) 宇賀克也：「パーソナルデータの利活用に関する制度見直し方針」について，ジュリスト，No.1464, pp.12-17 (2014)。
- 4) 総務省：パーソナルデータの利用・流通に関する研究会報告書—パーソナルデータの適正な利用・流通の促進に向けた方策 (2013), http://www.soumu.go.jp/main_content/000231357.pdf
- 5) Article 29 Data Protection Working Party, Opinion 13/2011 on Geolocation Services on Smart Mobile Devices (2011), http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2011/wp185_en.pdf
- 6) 朝日新聞「JR東、スイカ履歴販売」, 2013年7月26日朝刊38頁。
- 7) Federal Trade Commission, Protecting Consumer Privacy in an Era of Rapid Change : Recommendations for Business and Policymakers (2012), <http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>
- 8) 松前恵環：個人識別性／識別可能性といわゆる「FTC3要件」：個人識別性を巡る米国の議論動向を踏まえて，堀部政男情報法研究会第9会シンポジウム (2013年12月22日), http://www.horibemasao.org/horibe9_Matsumae2.pdf
- 9) 15 U.S.C. § 45(a) (1).

(2014年7月14日受付)

石川憲洋 (正会員) isic@komazawa-u.ac.jp

NTT, NTT ドコモを経て, 2010年より駒澤大学グローバル・メディア・スタディーズ学部教授。博士(情報学)。2003年度本会業績賞, 電子情報通信学会, ACM, IEEE 各会員。

松前恵環 matsumae@komazawa-u.ac.jp

東京大学法学部卒業, 同大学院学際情報学府修了, 同大学院情報学環, 同先端科学技術研究センターを経て, 2013年より駒澤大学グローバル・メディア・スタディーズ学部, 修士(社会情報学)。

3 M2M の情報流

猿渡俊介 ■ 静岡大学大学院情報学研究科

森川博之 ■ 東京大学先端科学技術研究センター

オートメーション

「墮落する前のアダムは、エデンの園で瞑想と動物の命名という仕事を課せられた。いまやオートメーションの場合がそうである。工程なり製品なりを命名してプログラミングを行いさえすれば、それが実現する」McLuhanの1964年の著作「メディア論」¹⁾の33章「オートメーション」において、機械の時代から電気の時代への移り変わりについて述べられた言葉である。

Machine-to-Machine (M2M) の本質はオートメーションによる生産性の向上と付加価値の創出にある。スマートシティ、スマートコミュニティ、スマートグリッド、スマートウォーターなどに代表されるように、持続可能な社会の構築が求められるようになってきたことが、M2Mに対して新たな視点を与えている。スマート化を図るためには、消費電力、

物流、流通、位置、環境、部品などの多様なデータをセンサや産業設備から収集し、これらデータを有機的に結合し、社会基盤の効率化および高度化を目指さなければならないからである。

極論すれば、M2Mの最終目的は、持続的な成長が可能な新たな国づくりを行っていくことにある(図-1)。地球規模で解決しなければならない人口爆発、食糧枯渇、資源枯渇、大規模自然災害、環境などの課題や、国内で解決しなければならない人口減少、少子高齢化、社会資本ストックの劣化などといった課題に対してM2Mが一助となり得る。

M2Mが社会に与えるインパクトは1960年代後半にPLC(Programmable Logic Controller)が製造プロセスに与えた影響がヒントとなろう。製造プロセスにおけるセンサとアクチュエータの関係が小型のコンピュータであるPLCによってプログラマブルになった。PLC以前の自動車の製造プロセスでは

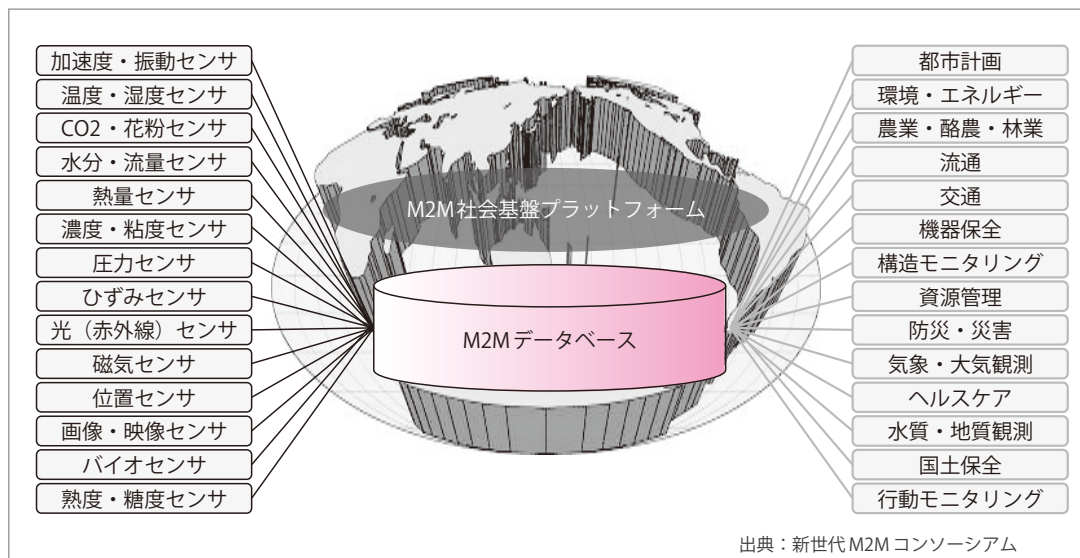


図-1 センサからの情報流を活用することで、農業、都市、環境、流通、資源、医療等の生産性を高め、新サービス創出に資することができる

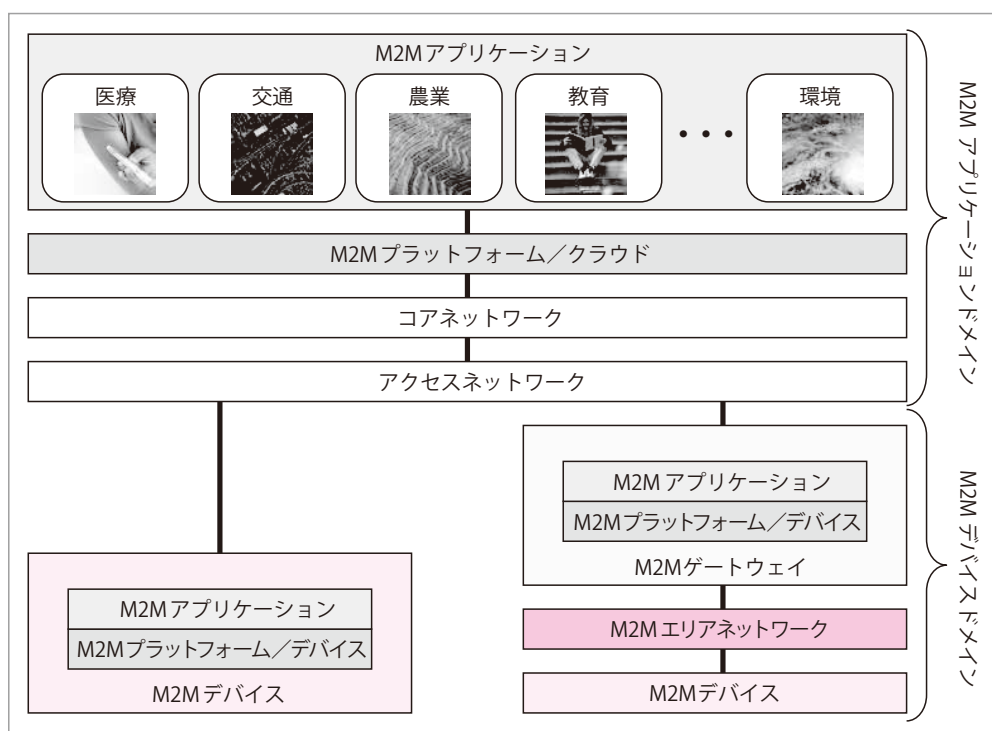


図-2 M2M の構成要素

人が関与せざるを得なかったが、PLCによって製造プロセスのオートメーションが可能となり、生産効率を大幅に向上させることになった。

このように M2M のビジネスチャンスは、生産性向上の余地がある分野を見つけ出すことにある。対象はすべての産業分野にまたがる。これまで人が行ってきたプロセスを置き換えることで、新たな価値を創出できる。

たとえば、水道管の漏水検知作業では熟練工の作業割合が非常に大きい。熟練工の技能の一部を M2M によって代替することができれば、国内においては高齢化による熟練工の減少への対策となる。海外においては 30% とも言われる非常に高い漏水率を低減することが可能となる。

そもそも、M2M の起源は、Wiener のサイバネティクス²⁾にまでさかのぼる。サイバネティクスとは、センサによる入力を通信と計算によってアクチュエータにフィードバックするオートメーションの根幹となる概念である。Wiener のサイバネティクスが扱っている領域は、制御・計算・通信の理論的側面だけでなく、機械としてのヒトの理解にまで踏み込んでおり、人工知能や認知科学などに多大な影響を

与えた。

しかしながら、今まで制御は制御分野に閉じて発展してきた。PLC の製造プロセスへの導入、SCADA (Supervisory Control And Data Acquisition) によるシステム監視やプロセス制御、FieldBus、Modbus、産業用イーサネットなどの制御向け通信プロトコルの開発などである。

情報処理や通信の技術が進展し低コスト化したことで、制御・計算・通信が融合しつつある。GE の Industrial Internet、ドイツの国家プロジェクト Industrie4.0 などは、まさに制御・計算・通信の融合を目指している。Wiener がサイバネティクスで示した世界の実現に近づきつつある。

M2M ネットワーク

M2M は、大きく分けると M2M アプリケーションドメインと M2M デバイスドメインとに分けられる (図-2)。

M2M アプリケーションドメイン (図-2 上) では、多種多様な M2M デバイスを有機的に連携させて M2M アプリケーションを構築するための機能が

M2M プラットフォーム／クラウドとして提供されている。

M2M プラットフォーム／クラウドの目的の1つが、膨大な数の M2M デバイスの運用管理にかかる時間とコストを低減することである。M2M デバイスのモニタリング、故障検知、課金、認証、アクティベーション、SIM カード管理などとともに、ローミングやプロバイダの切り替えなどのサポート機能を担う。数百万というオーダーのデバイスを1つの顧客が所有している形態であることから、従来の携帯端末管理とは異なる機能が必要となる。

これらの機能と膨大に存在する既存のソフトウェア資産とを有機的に連携させるために、M2M プラットフォーム／クラウドは Web サービスとの融合を意識した設計がなされている。たとえば標準化団体 oneM2M において RESTful インタフェースや SOAP によって API を統一することが検討されている。RESTful とは、現在の Web のような、URI を用いてリソースにアクセスするステートレスの API の設計思想を意味する。JavaScript を用いたマッシュアップ等との親和性が高いことから、Yahoo!, Google, Facebook, Twitter など多くの Web サービスで採用されている。マッシュアップにより、複数の異なるサービスの部品を組み合わせる新たなサービスを簡単に生み出すことができる。

M2M デバイスドメイン (図-2下) では、センサ等のデバイスをネットワークに接続するための仕組みが提供されており、M2M デバイス、M2M エリアネットワーク、M2M ゲートウェイから構成される。M2M デバイスとしては、自販機、自動車、医療デバイス、スマートメータ、工作機械など膨大な種類のデバイスが存在する。M2M デバイスをアクセスネットワークに接続する方法としては、直接接続する方法 (図-2左下) と、M2M ゲートウェイを介した方法 (図-2右下) の2種類が存在する。携帯電話モジュールをデバイスに組み込むことで、直接アクセスネットワークに接続される。コマツの遠隔車両監視、富士アイティによるコインロッカー管理、コカコーラによる自動販売機の遠隔在庫管

理などがこれに相当する。携帯電話モジュールに M2M プラットフォーム／デバイスが組み込まれているため、アプリケーションに応じたカスタマイズが可能である。

M2M ゲートウェイを介した方法では、M2M デバイスは、Bluetooth や IEEE 802.15.4 などの M2M エリアネットワークによって M2M ゲートウェイに接続される。ユーザは M2M ゲートウェイで提供されている M2M プラットフォーム／デバイスを用いて個別のアプリケーションや、M2M アプリケーションドメインと連携したアプリケーションを構築できる。たとえば、スマートフォンをゲートウェイとして、活動量センサからの情報を Bluetooth を介して収集してランニングサポートを提供する Nike+ などがこれに相当する。

IETF では、M2M エリアネットワークと M2M プラットフォーム／クラウドとの連携がしやすい仕組みが標準化されている。CoRE ワーキンググループが提供する CoAP (Constrained Application Protocol) は UDP で動作する RESTful インタフェースを標準化している。CoAP は簡易版 HTTP として設計されており、プロキシを介して通常の HTTP と透過的に扱うことができる。6LoWPAN (IPv6 over Low power Wireless Personal Area Networks) ワーキンググループでは、IEEE 802.15.4 を前提としてネットワーク層で IPv6 を利用する仕組みが提供されている。IEEE 802.15.4 では最大パケット長が 127 バイトであるのに対して IPv6 ではヘッダサイズだけで最低 40 バイト必要であるため、RFC 6282 ではヘッダの圧縮方法やパケットのフラグメンテーションのための仕組みが策定されている。Roll (Routing Over Low power and Lossy networks) ワーキンググループでは、低消費電力かつ不安定な通信リンクでもマルチホップネットワークを構築して効率的にセンサデータを収集できる RPL (IPv6 Routing Protocol for Low power and Lossy Networks) が提供されている。

IEEE でも、M2M エリアネットワークを想定した IEEE802.11ah の議論が始まっている。IEEE802.11ah



図-3 Nest のサーモスタットと火災報知機

のユースケースの要求条件として、伝送速度が 100kbit/s ~ 20Mbit/s, 許容パケット誤り率が 1 ~ 10%, アクセスポイント 1 台あたりの端末数は 50 ~ 6,000 台, トラフィックパターンは連続, 定期的, バースト的, イベント駆動など多様な条件が挙げられている。また, 上り回線の輻輳対策や省電力化の議論が行われている。

情報流の価値

2014 年 1 月, Google はスマートなサーモスタットや火災報知機 (図-3) を作っている Nest を 32 億ドルで買収した。サーモスタットとは, ヒーターやエアコンのコントローラであって, 家の温度管理のハブとなる機器である。Nest のサーモスタットは学習機能を備えており, 家に人がいる時間を推定して制御することで電気代やガス代を 20% ほど節約できる。すでに 100 万台近く売られているとも言われている。

また, Nest Developer Program を用いてサーモスタットと火災報知機の連動 (一酸化炭素を検知するとガスの元栓を閉めるなど) や, 冷蔵庫, 洗濯機, 照明, 自動車などといったサードパーティ機器との連動 (自宅に近づくと到着予定時刻に合わせて室内を設定温度に調整するなど) も可能である。

しかしながら, サーモスタットや火災報知機といった機器だけでは 32 億ドルもの価値にはならない。家の中に存在するあらゆる機器のハブとなることで, 将来的には家の中のすべてのデータを収集できることに対する期待が大きい。もちろん, Nest の人材



図-4 Insure The Box のボックス

の魅力も価値に含まれてはいるものの, Nest の機器から生み出される情報流自体に大きな価値を有することが 32 億ドルにつながっていると考えられる。

英国の保険会社の Insure The Box は顧客の運転状況を把握し, 運転が滑らかなドライバーの保険料を安価に設定する「テレマティクス型」と呼ばれる保険を提供している。

クルマの位置, 速度, 走行距離, 燃費などを計測する「ボックス」(図-4) を設置し, 一人一人の運転の安全性を把握している。英国では個人向け自動車保険料が上昇しているため, 特に若い世代の顧客に訴求している会社である。

しかしながら, Insure The Box の本領は, クルマからの情報流で発揮される。クルマからは位置, 速度, 距離, 燃費, 事故, 故障, 運転時間, 運転距離, 運転頻度などのデータを得ることができる。すでに, 6 億マイル, 1.6 億旅行分のデータを収集しているとのことである。今後これらのデータを活用したサービスが登場してくる可能性が高い。

そもそも Google, Amazon, Facebook 等の企業は膨大な量のデータを集めている点に強みがある。集めたデータ自身がプラットフォームを構成しており, 多様なサードパーティがプラットフォーム上にサービスを展開するエコシステムが巷を席卷している。

M2M に関しても同様である。将来は, 多様な M2M 情報流を手中に収める企業が覇者となる。こ

【ゼンメルワイスの悲劇】

ウィーン総合病院第一産科の医師ゼンメルワイスは、産褥熱による死亡率が第二産科（2.0%）に対して第一産科（13.1%）の方が高いことに注目し、解剖にあたった医師が手指を消毒しないで診療にあたることによるものであるとの説を唱えた。手指を消毒することにより死亡率は2.4%にまで低下したとされる。手指の消毒さえ行わないということは現在の通念からいえば驚くべきものであるが、病原菌の存在を知らなかった当時の医師会は、ゼンメルワイスの説を科学的ではないとして受け入れることはしなかった。ゼンメルワイスの死後、感染は病原菌によって起こることが発見され、今では消毒法と院内感染予防の父として認識されている。

の企業は、Google や Amazon に匹敵する企業となる可能性もある。M2M データは、現在の Web 企業が対象としている Web データとは異なる新しいデータであり、現時点で膨大な量の M2M データを収集している企業は存在しない。M2M においては、大企業でさえこれから、というフェーズにあり、誰にでもチャンスがある世界である

OECD においても、「新たな成長源：知識ベース資産（New Sources of Growth：Knowledge-Based Capital）」と題するプロジェクトが始まっており、データ自身が新たな知見を産み出し、経済成長に資する源となることを謳い始めている。囲み記事のゼンメルワイスの悲劇を持ち出すまでもなく、データはあらゆる分野において価値を産み出す。データ駆動型経済の到来を踏まえ、データ自身の価値をあらためて認識することが必要となる。

さまざまな分野での情報流

★ モバイルヘルス

モバイルヘルスとは、健康、医療、介護といった広義のヘルスケア領域に対して、無線通信技術を用いたサービスのことである。患者が自宅で体温、心拍、血圧などのバイタルサインを測定し、医療機関に送信することで通院負担を軽減させたりすることが可能になりつつある。糖尿病患者を対象とした血糖値データの医療機関への送信といったサービスも登場している。

このような病院「外」のデータは、医療従事者にとって貴重なものである。活動量が指示通りであるかなどを把握することができるとともに、疾病の兆候をも把握することができ、病院の内と外での一体的な医療提供が可能となり医療や健康管理のあり方も変わる。

Scanadu 社の開発した Scout は、小型円状のデバイスをこめかみに当てるだけでさまざまなヘルスケアデータを同時に取得する。遠赤外線センサで体温を、マイクで心拍数と呼吸数を、可視光 LED と近赤外線 LED で血中酸素飽和度を、心電図センサで不整脈と血圧を測定する。人間の体内を覗いてどんな病気でも診断できる「スター・トレック」のトリコーダの現代版である。

多様なデバイスが開発され、さまざまな情報流が得られることになる。医療・ヘルスケアに関するデータを容易にかつ多量に収集し、かつパーソナルヘルスレコードと連携させることで、患者に対して新しい価値を提供することができる。

★ 農業

農業も M2M で大きな効率改善が見込まれる分野である。生産、流通、販売、消費のサプライチェーンにおいては膨大な情報流が生まれている。たとえば、環境データ（気温、湿度、CO2 濃度、光強度、光質、日長、培地温度、培養液温度・濃度・組成・pH、給液頻度など）、エネルギー管理データ（エネルギー消費量、空調制御データなど）、生育データ（品質、熟度・糖度、樹勢、病虫害など）、収量・出荷量データ、労務管理データ（収穫等の作業時間や作業量など）、物流データ（輸送経路、保管時間、積み下ろし回数など）、市場価格データ、販売時点情報管理データ、フードログデータ、健康データなどである。

これらを活用することで、栽培管理、労働管理、収量管理、収量予測、市場価格とバランスをとった環境管理、統合的エネルギー管理によるコスト削減、安全安心や品質などの高付加価値化、ソーシャルとの連携などによる新たな食文化の創造な

どに資することが期待できる³⁾。

★ 予防保全

M2Mの本流は、施設、プラント、工場、インフラなどの故障や異常の予兆監視による予防保全である。センサからの情報流を分析し、健全な運用状態と比較し、不健全な状態を把握して予防保全を実現することで、安全性の向上や効率的な運用を実現できる。

NECは、島根原子力発電所に設置された振動計、圧力計、温度計、加速度計といった約2,500種類、約3,500におよぶセンサからの情報流を独自のインバリエント分析技術でもって解析し、「いつもと違う」挙動を検出するシステムを納入している。

また、富士通はオムロンの草津工場において、はんだ印刷機、高速マウンター、多機能マウンター、リフロー炉の4工程からの情報流を製品ごとに紐づけることで、予防保全のみならずコストに直結するライン全体の生産性を把握できるようにしている。

社会インフラの予防保全に関しては、笹子トンネルの事故を受け、国土交通省において2013年10月から「社会インフラのモニタリング技術活用推進検討委員会」を開催している。適切な維持管理手法の導入により長寿命化、維持管理、更新のトータルコストの縮減・平準化を図ることが目的であるが、社会インフラにおいてはデータが蓄積されていないことが大きな問題である。橋、トンネル、建物などの構造物の崩壊現象は現象自体が稀であるためである。

そのため、まずはデータを収集するところから始めなければならない。このような観点から、筆者らの一人の猿渡は、東京都市大学、筑波技術大学と共同して、加速度センサ、音声センサ、カメラなどでの軍艦島モニタリングプロジェクトを始めている。崩壊が進んでいる軍艦島から得られるデータは、予防保全分野の進展に資する貴重なデータになると考えている。

情報流と社会イノベーション

M2Mと社会的課題の解決との親和性はきわめて高い。M2Mで社会イノベーションを実現するためには、「新たな技術の開発」に加えて「社会的課題への気づき」と「課題と技術のマッチング」が重要となる。すなわち、社会のデザインが必要である。

社会的課題に対して敏感になるとともに、対象とするフィールドに自ら出向き、他分野の研究パートナーとの深い議論を通して、フィールドごとに存在する要求や制約を抽出しなければならない。

すでに多くの企業がフィールドに出向いていることで、さまざまな分野でM2Mの活用事例が蓄積されつつある。しかし、全体の産業セグメントから見ればまだまだ一部でしかない。

M2Mの情報流は、「スマートゴミ箱（ゴミの量をセンシングし、回収タイミングを最適化するシステム）」など、我々の手の届くエリアにも転がっている。多くの研究者や技術者がフィールドに出向き、M2Mの情報流を見出し、活用事例を今まで以上に蓄積していくことで、真の意味での社会イノベーションが実現されることを期待したい。

参考文献

- 1) McLuhan, M.: Understanding Media: The Extensions of Man, McGraw-Hill (1964). 栗原, 河本 (訳): メディア論.
- 2) Weiner, N.: Cybernetics: or Control and Communication in the Animal and the Machine, The MIT Press (1948). 池原, 彌永, 室賀, 戸田 (訳): サイバネティックス—動物と機械における制御と通信.
- 3) 東京大学アンビエント社会基盤研究会, 農林環境ワーキンググループ報告書 (2012).

(2014年7月9日受付)

猿渡俊介 (正会員) saru@inf.shizuoka.ac.jp

2007年東京大学大学院博士課程修了。現在、静岡大学大学院情報学研究科助教(テニュアトラック)。専門はモバイルシステム、センサネットワークなど。電子情報通信学会論文賞、本会山下記念研究賞など受賞。

森川博之 (正会員) mori@mlab.t.u-tokyo.ac.jp

1992年東京大学大学院博士課程修了。現在、東京大学先端科学技術研究センター教授。本会論文賞、電子情報通信学会論文賞(3回)、ドコモモバイルサイエンス賞、総務大臣表彰など受賞。新世代M2Mコンソーシアム会長、OECD/CDEP副議長等。

4 ビッグデータとのつきあい方

—ビッグデータ活用のための技術と応用基盤—

原 隆浩 ■ 大阪大学

ビッグデータ時代の到来

人やモノから多様な種類の大量のデータが日々生成されており、「ビッグデータ」時代がまさに到来している。ビッグデータを解析することで、さまざまな有益な情報を抽出・生成できる可能性があるため、産業界および学術界において非常に注目されており、活発に研究開発が行われている。

ビッグデータに関する研究開発は、その重要性から、多くの国において政府主導で戦略的に推進されている。たとえば、米国政府は2012年3月29日に“Big Data Research and Development Initiative”を発表し、研究開発を強力に支援している。さらに、欧州委員会の第7次研究枠組計画(FP7)においてビッグデータに関する多数の研究プロジェクトが推進されている。また、中国でも国家戦略として、IoT(Internet of Things)を中心とするビッグデータの研究開発に注力している。我が国でも、総務省が平成25年版の情報通信白書において、「『スマートICT』の進展による新たな価値の創造」を課題に挙げ、ビッグデータ活用の重要性について記載している。さらに、日本政府は、ビッグデータの活用のために、個人情報保護法を中心とした法整備にとりかかっている。

★ ビッグデータの応用例

2012年の米国大統領選挙でObama大統領がビッグデータを活用して、勝利を取めたことはよく知られている。大量のデータを収集し、選挙の情勢や各地域の住民の嗜好・特性を正確に把握し、有効な選挙戦略を立てたことが勝因の1つといわれている。このように、大量のデータを解析することで、これまでは不可能であった、対象となる事象・物の詳細なモデ

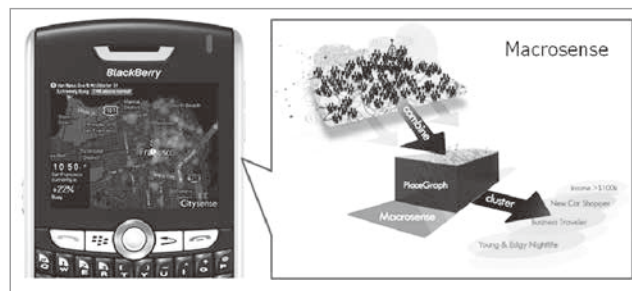


図-1 Citysense

ル化、現状の詳細な把握・追跡、さらには将来の予測が可能となる。そのため、ビッグデータの解析を用いたさまざまな新たな応用が研究開発されている。

最も典型的な例として、Webサービスにおけるユーザの利用履歴(購買履歴など)やクリックスルー(操作・閲覧履歴)、レビューの大規模な解析によるマーケティング戦略、サービスの改善などが一般的に行われている。また、スマートフォンなどに搭載されているセンサから生成されるデータを大規模に収集・共有する「参加型センシング」が近年注目されており、その解析、サービス応用もビッグデータ応用例の1つである。Citysense^{☆1}は、Sensor Networks社(2014年1月にYP社が買収)が2008年に開始した実サービスであり、参加型センシングによるビッグデータ解析を最初期に商用利用したものの1つである。CitySenseではスマートフォンを持つユーザを対象とし、San Franciscoのユーザ分布をユーザの端末に表示する(図-1)。たとえば、「今、どこに人が集まっていて、どこに行こうとしているか?」などの問合せが可能となる。このために、同社のMacroSense(位置データを用いたユーザ動作解析ツール)を利用し、現在の数万単位のユーザの位置情報と、過去の数10

☆1 <http://www.sensenetworks.com/>

億のデータを利用して、現状をリアルタイムに予測している。

★ビッグデータ処理のための技術

それでは、上記のような応用を実現するためには、どのような技術が必要となるだろうか？

まず、M2M (Machine to Machine) などでは各種センサや監視機器などから連続的に発生するデータ（データストリーム）を、アプリケーションなどの要求に応じて取捨選択し、加工するなどしてデータベースに格納する。さらに、異常検出などのモニタリングアプリケーションでは、データの到着時にリアルタイムで、特定の値やパターンに合致するデータを発見する必要がある。そのために、データストリームを効率的に処理する技術が必要となる。

さらに、ユーザや車両が絶え間なく生成するセンサデータの解析や、Web サービス等におけるユーザの利用履歴、操作情報の解析は、いずれも大規模になるため、単一のサーバで処理を行うのは現実的ではない。そのため、効率的な大規模分散処理の仕組みが必要となる。

生成される大規模データに対する処理だけではなく、データを効率的に蓄積・管理するデータベース技術も重要となる。これまで、大規模データの管理には、関係データベースとSQLを用いることが一般的であったが、データやその発生源、用途が多様化した現在では、十分な機能・性能を得られない場合がある。

上記の3つはビッグデータ処理のための基盤技術であるが、最近では、これまで各機関・企業・個人が個別に収集・管理していたビッグデータを公開し共有することの重要性も謳われている。そのため、ビッグデータの共有を可能とする応用基盤の構築が急務である。

次章以降では、上記の4つの技術について、代表的なものを紹介する。

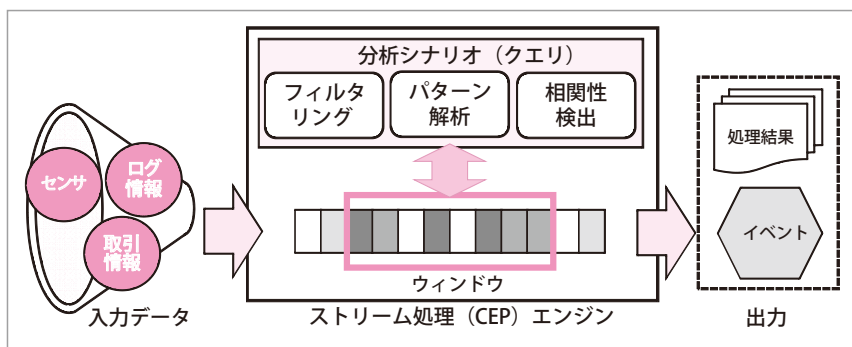


図-2 ストリーム処理 (CEP) エンジンの構成

ストリーム処理技術

ストリーム処理に関しては、2000 年前後からデータベース分野を中心に、単一もしくは複数のデータストリームに対して指定した条件に合致したデータを効率的に発見するための問合せ処理や、特定のパターンに類似した時系列データをストリームから発見するための類似検索、類似性に基づいて時系列データをグループ化するクラスタリングなど、さまざまな研究開発が行われている。

ビッグデータへの注目に伴い、ストリーム処理の重要性も高まっており、複合イベント処理 (CEP: Complex Event Processing) と称して、各種ベンダがさまざまな製品を開発している。たとえば、オラクルの Oracle CEP、サイベースの SAP Sybase Event Stream Processor、日立製作所の uConsminexus Stream Data Platform などがある。

図-2 は、ストリーム処理 (CEP) エンジンの一般的な構成を表している。入力データとしては、ユーザの位置情報や各種センサデータ、クリックスルーやクエリログ、システムログ、さらに金融情報の場合は入出金・取引データ、株価などさまざまなものが、短い時間間隔で定期的もしくは不定期にシステムに投入される。

データストリームは通常、絶えず (半永久的に) 到着するため、すべてのデータを蓄積するのではなく、到着するたびに処理が行われる。この際、最近の一定数もしくは一定時間に到着したデータ (ウィンドウ) を単位として処理を行うウィンドウ演算の機能を持つものが多い。さらに、ユーザは、到着したデータに対して、どのように処理・分析を行うか (分析シナリオ)

を、問合せ（クエリ）としてシステムに指定する。これは、従来のデータベースにおける単発的な問合せとは異なり、データストリームに対して連続的に（絶えず）実行されるものであるため、「連続問合せ」と呼ばれる。一般的には、データストリームのすべてではなく、問合せで指定した条件に合致するもののみがデータベースに蓄積される場合が多い。

上記の連続問合せには、到着したデータの取捨選択（フィルタリング）や、特徴パターン解析、時系列データの類似性判定、相関性抽出など、さまざまなものが含まれる。問合せの記述言語としては、Stanford 大学で開発された CQL (Continuous Query Language)¹⁾ が有名である。CQL は、SQL に似た宣言的な記法で、問合せを記述できる。

大規模分散処理技術：Hadoop

ビッグデータを効率的に分散処理するフレームワークとしては、Apache ソフトウェア財団がオープンソースとして公開・開発している Hadoop^{☆2} が有名である。Yahoo! や Facebook、楽天など多くの IT サービスプロバイダが、Hadoop を用いてログの集計・解析などさまざまな業務上のタスクを実行している。Hadoop は主に、Hadoop 分散ファイルシステム (HDFS) と MapReduce エンジンから構成される。MapReduce は、Google によって 2004 年に提案された分散処理のためのプログラミングモデルである。

Hadoop の MapReduce は、HDFS に格納された大量のデータに対して、複数のサーバを用いて一括処理（バッチ処理）を行い、その結果をファイルとして HDFS に書き込む。HDFS はこのような処理に対して、データ転送のスループットを高めるように設計されている。

★ HDFS

HDFS のシステム構成の概要を図-3 に示す。HDFS は基本的に、ネームノードと呼ばれる管理サーバと、複数のデータノードと呼ばれるデータサーバから構成

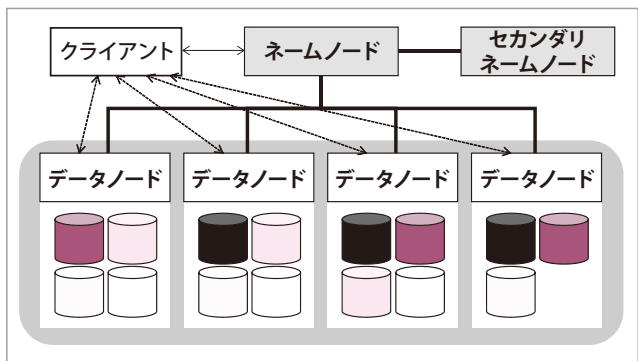


図-3 HDFSの構成

される。データは、指定されたサイズ（デフォルトは 64MB）のデータブロックに分割され、障害時に備えて指定数（デフォルトは 3）に複製されてデータノードに配置される。この配置場所はネームノードが決定する。データノードの追加・離脱の際も、ネームノードがデータの再配置を決定する。ネームノードはクライアント（たとえば MapReduce プロセス）からの処理依頼に応じて、データの読み書きをどのデータノードに行うかを決定し、クライアントに伝える。実際のデータのやりとりは、クライアントと該当するデータノード間で直接行われる。

HDFS では、データをブロック単位でハードディスクの連続領域に書き込み、一度書き込んだデータの上書きを許可しない（追加は可能）ことで、データのまとまった読出しを高速化している。

HDFS では、ネームノードが単一障害点となるため、障害時に備えて、バックアップのためにセカンダリネームノードが配置される。

★ MapReduce

MapReduce では、HDFS に格納されているデータに対する並列処理を、Map 処理と Reduce 処理の 2 段階で実行する。Map 処理では、データを小さな単位に分割し、大規模数のサーバにそのデータに対するタスクを割り当てる。この際、無駄なデータ転送を削減するために、各データを所持するサーバ（データノード）に該当するタスクを割り当てるのが一般的である。各サーバは担当のタスクを実行し、中間結果を出力する。この中間結果は、Reduce 処理の各タスクを実行するサーバに転送される。Reduce 処理では、

☆2 <http://hadoop.apache.org/>

Map 処理の結果（中間結果）を集約し、最終的な結果を得る。最終結果は、HDFS に書き込まれる。

MapReduce の処理の例を図-4 に示す。この例では、大量の Twitter やアンケート情報の中から地名に関する単語を抽出し、各地名の出現回数をカウントしている。まず、Map 処理において、各サーバ (A, B, …) は割り当てられた全データから地名を抽出する。これが Map 処理の出力結果であり、中間結果となる。そして、抽出した地名のリスト内の各項目（地名と出現回数の組）を、Reduce 処理のタスクを実行するサーバ (Q, R, …) へ送信する。この際、地名をキーとして、五十音順で一定の範囲のキーが、各サーバに割り当てられている。Reduce 処理では、各サーバは自身に割り当てられた各キーの出現回数を集約し、最終結果としている。

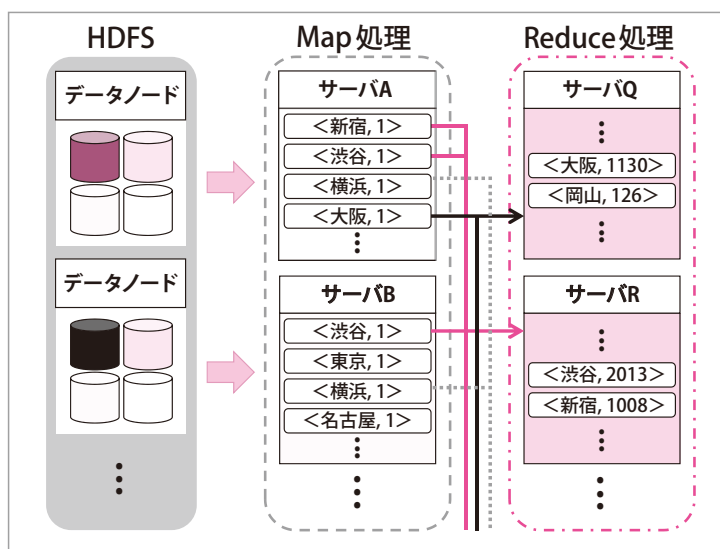


図-4 MapReduce 処理の実行例

NoSQL

HDFS は、MapReduce によるバッチ（分散）処理を効率的に行うためのファイルシステムであり、基本的にその処理に一度だけ発生するデータ読み出し・転送の性能を最大化することを目的としている。一方、蓄積されたビッグデータに対して、（分散処理だけではない）さまざまなデータ処理を効率的に実行することが求められている。この際、従来の関係データベースだけでは十分な性能が出ない場合が多いことから、“Not Only SQL (NoSQL)” と称して、多種多様なデータベースが開発されている。つまり、NoSQL はある特定のデータベースのことを指すものではない。NoSQL では、蓄積されたデータベースに対する頻繁なデータ読み出し・書き込みや、その整合性などが考慮されている。なお、HDFS や MapReduce 自体は NoSQL とは呼べないが、HDFS 上に NoSQL を構築することは可能である。実際に、HBase がその代表例である。また、MapReduce は、HDFS 上に構築された NoSQL に対して効率的な分散処理の機能を提供するための、一種のイネーブラとして捉えることも可能である。以下

では、NoSQL が登場した背景と特徴について紹介する。

★ NoSQL が登場した背景

関係データベースは長年にわたり商用システムの根幹を支えてきたが、その長所である「スキーマ定義がしっかりしている」「正規化と結合処理によって無駄や情報欠損がなく表形式で対象を表現できる」「処理の整合性や複製間の一貫性が保障される」といった特徴が、ビッグデータの処理では弊害となる場合がある。その例を以下に示す。

問題①：さまざまな情報源からデータが生成される環境では、あらかじめスキーマを定義するのが困難な場合が多く、柔軟性・拡張性の面で問題がある。

問題②：正規化によって、対象を複数の表に分けて表現し、結合処理によって再現することは、ビッグデータでは負荷が非常に大きい。さらに、ビッグデータの分散処理のために、データを水平分割（タプル・レコード単位で分割）して大規模数のサーバに分散配置することが多いが、まず正規化によって分割された表を結合する必要がある、このための負荷が膨大となる。

問題③：大規模な種類のデータに対してその複製を大規模数のサーバに分散配置することが多いため、それらの一貫性（バージョン）を完全に保障すること自体の負荷が非常に大きく、システム全体の性能を低下させる場合がある。

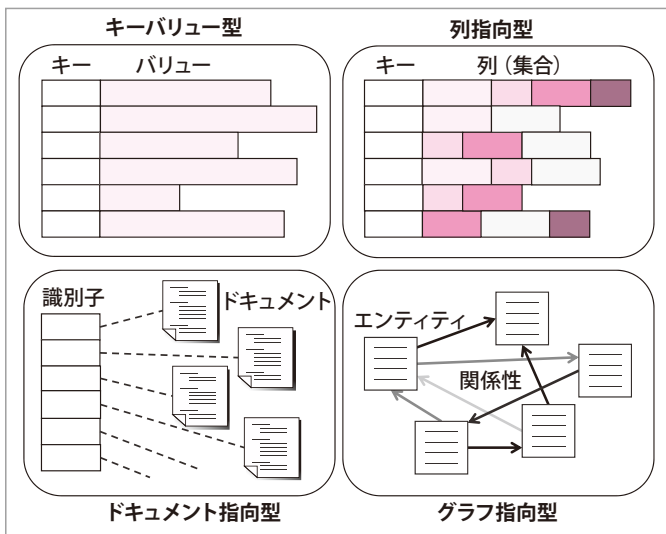


図-5 NoSQL の分類と構造のイメージ

以上のような問題点を解決するために、さまざまなNoSQLが開発されている。

★ NoSQL の分類と特徴

NoSQL の分類にはいくつか異なる方法があるが、キーバリュ型、列指向型、ドキュメント指向型、グラフ指向型に分類するものが多い²⁾。各型のデータ構造のイメージを図-5に、代表的な商用およびオープンソースのNoSQLを表-1に示す。以下では、各型の特徴について、上記の問題と対応付けて議論する。なお、問題③については、各NoSQLによって対応がさまざまであり、想定するアプリケーションに応じて、一貫性を軽減したりしている。

キーバリュ型

キーバリュ型は、データをキーとバリュ(値)の組という単純な構造で表す。バリュがデータ本体で、キーがその識別子のようなイメージである。バリュで扱うデータは文字列などさまざまであるが、構造を持つデータもバリュとしてひとまとめになるため、構造を意識した処理には対応できない。そのため、キーを用いてデータを参照し処理を行う、キーに従ってデータをサーバに配置するなどといった単純な処理を対象とする場合に用いられる。構造を意識しないという特徴から、上記の問題①と②を解消している。

列指向型

列指向型は、キーバリュ型を拡張して、バリュ

分類	代表例
キーバリュ型	Amazon Dynamo, Redis, ROMA
列指向型	HBase, Cassandra, Hypertable
ドキュメント指向型	MongoDB, CouchDB, Terrastore
グラフ指向型	Neo4j, AllegroGraph, Sones

表-1 NoSQL の分類と代表例

に複数の列(カラム)を持てるようにしたものである。ただし、関係データベースとは異なり、データごとに列の数や属性(列名)を動的に決定できる。キーバリュ型の特徴をおおむね継承しており(そのため問題①と②を解消)、さらに、カラムの属性を指定したデータの読み書きや検索、処理が可能となる。

ドキュメント指向型

ドキュメント指向型は、XMLやJSONなどの記述言語によって記述されたドキュメントをデータとして扱うことが可能であり、各データには一意の識別子が付与される。前述の2つの型と同様に、上記の問題①および②を基本的に解消している。該当する記述言語の仕様に従い、各データへの読み書き・処理等を行うことが可能である。

グラフ指向型

グラフ指向型は、対象を複数のノード(エンティティ)とノード間の枝(関係性)としてグラフ構造で表現する。エンティティと関係性は属性を持ち、それぞれ具体的にどのようなものかを表している。グラフ指向型のNoSQLはたとえば、ソーシャルネットワークサービス(SNS)におけるユーザ同士の関係(グラフ構造)の解析などに用いられる。関係データベースでも、データ間の関係を複数の表の共通属性などで表現できるが、結合処理等によって再現する必要がある。グラフ指向型は、このような関係を直接的に表現できるため、上記の問題②を解消できる(問題①を解消できるかはノードの表現方式に依存する)。

情報銀行

ここまでで、ビッグデータを扱うための基盤技術について紹介したが、ビッグデータの恩恵を最大限に活

用するためには、さまざまな企業・機関および個人が個別に収集・管理(独占)している大規模なデータを共有可能な応用基盤の確立や法整備が急務と考えられる。法整備については、各国において、プライバシーを保護しつつ、匿名化した個人情報であれば第三者に提供可能にするといった法改正等が検討されている。

応用基盤については、情報通信研究機構のK-L Gridなどセンサ、科学、言語、音声、文書データを共有するためのプラットフォームが多数構築されている。しかし、近年のスマートフォンユーザが生成する位置情報やセンサデータなどを大規模に収集・管理し、プライバシーを考慮した共有を可能とする応用基盤はほとんど存在しない。

ここで、「情報銀行」と呼ばれる興味深いコンセプトが東京大学・柴崎亮介教授らによって提唱されている³⁾。これは、個人が自身の管理下でさまざまな活動情報(行動履歴、購買履歴、予定、既往歴など)を統合し、承認の上で「情報銀行」に預け、運用を委託するというものである(図-6)。これにより、個人の管理・制御の下でプライバシーを保護しながら、大規模な個人情報を社会・産業のために高次利用することが可能となり、価値創造のサイクルを実現できる。つまり、ビッグデータの恩恵を最大限に活用することが可能となる。具体的には、以下のような利点が生まれる。

- 個人が、自身の情報がどのように使われているのか、意図に反して使われていないかを監視できる。さらに、使われ方を制御できる。
- 事業者としては、これまで自身の事業に関する断片的な情報のみしか利用できなかったが、複数サービスに関する情報をユーザの管理下で横断的に利用できるようになる。
- 個人としては、サービス品質の向上が期待できる。社会・産業的には、有益な情報・新たなサービスの創生が期待できる。

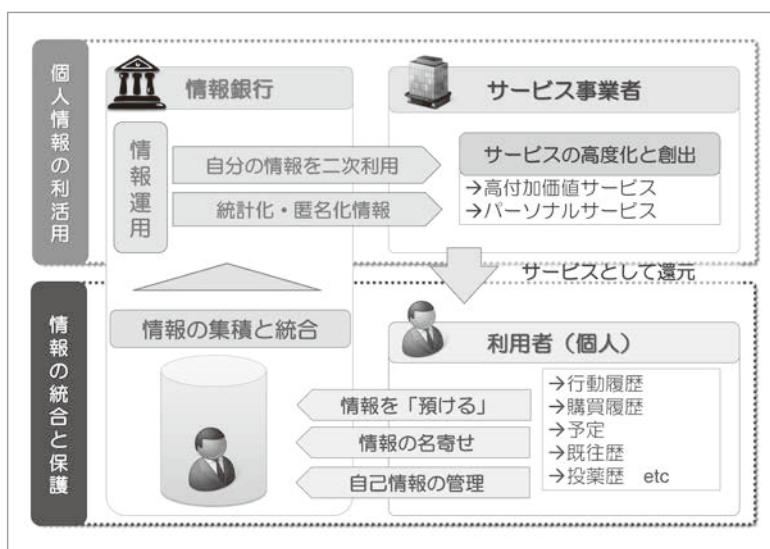


図-6 情報銀行における価値創造のサイクル (<http://shiba.iis.u-tokyo.ac.jp/research/ibank/pdf/2011.iBank-Poster.pdf> より引用)

今後の展望

本稿では、ビッグデータを活用するための、ストリーム処理、大規模分散処理、データベースについて、それぞれ代表的な技術を紹介した。後者の2つについては、たとえばHadoopを簡単に利用可能な機能を有するNoSQLデータベースもいくつかあり、すでにある程度の連携が行われている。しかし、ストリーム処理エンジンと後者2つの連携や、さらにそれらを多段に連携させて、大規模データストリームをリアルタイムに分散処理する基盤や、その結果等をクラウド上で新たなサービスとして共有可能とする基盤については、その有効性が認識されているものの、実用的なフレームワークはまだないのが現状である。このようなフレームワークの構築が今後の重要課題の1つと考えられる。

参考文献

- 1) Arasu, A., Babu, S. and Widom, J.: The CQL Continuous Query Language: Semantic Foundations and Query Execution, The VLDB Journal, Vol.15, No.2, pp.121-142 (2006).
- 2) 太田 洋 監修, 本橋信也, 河野達也, 鶴見利幸 著: NoSQLの基礎知識—ビッグデータを活かすデータベース技術, リックテレコム (2012).
- 3) 情報銀行コンソーシアム(仮称), <http://www.information-bank.net/> (2014年6月18日受付)

原 隆浩 (正会員) hara@ist.osaka-u.ac.jp

1997年大阪大学大学院前期課程修了。2004年より同大学院情報科学研究科准教授。工学博士。2008年、2009年本会論文賞受賞。データベースシステム、モバイルコンピューティングに関する研究に従事。IEEE、ACM、電子情報通信学会、日本データベース学会各会員。

5 多数のデータストリームを実時間で融合・編纂し利活用するための次世代「情報流」技術

—情報流キュレーション基盤実現に向けた課題抽出と取り組み—

安本慶一 ■奈良先端科学技術大学院大学

山口弘純 ■大阪大学

超大量データ時代の「情報流」

近年、センサの小型化・デジタル化・ネットワーク化が進み、人、車、機械、ビルなど、さまざまな場所または物体に埋設または付帯された多種多様なセンサから、実世界の状況を示すセンサデータを収集できるようになりつつある。そういったセンサデータはストレージに蓄積され、膨大な時間と計算パワーを費やして緻密な分析を行うことで、特定のサービス向けの知識やパターンを抽出することが従来の活用方法であった。しかし、実世界を鑑みたとき、人や車、モノ、環境の時々刻々とした変化により生成されるセンサデータはより複雑で雑多であり、さまざまな場所や地域から同時並行的に(あたかも湧出するように)生成される。

そういった状況を考慮すると、データは静的解析のために蓄積するのではなく、より高度な二次利用、三次利用(高次利用)のために積極的にリアルタイム流通させるべきである。なぜなら、雑多で多様かつ膨大な複数のデータストリームをある価値判断や必要性に基づいて実時間で選別する仕組みや、選別されたデータストリームを高次利用して新しい「知」を創出するための仕組みがあれば、社会の変革に寄与する次世代サービスを提供できる可能性が高いためである。たとえば、渋滞情報提供に利用されるプローブカー位置情報も、物資の搬送情報やさまざまな地点の天候・路面状況情報と実時間で組み合わせることで、都市空間の配送業務や除雪作業の効率化にも活用できる。また、2020年の東京オリンピック開催を控え、訪日観光客が大幅に増加することが予想され

ているが、Google Glassのような一人称視点カメラで時々刻々と撮影されるオリンピック会場や日本の観光地映像をリアルタイム選別し、産官民が発信する日本の関連情報やSNS(ソーシャル・ネットワーキング・サービス)から得られる情報とマッシュアップして世界に発信すれば、体感的で草の根視点である高臨場感のリアルタイム映像やライブ状況情報が提供でき、我が国への理解促進や魅力向上にもつながるだろう。

筆者らは、近年ニュース配信サービスなどで浸透しつつあるキュレーション(curation)という概念で、そういったサービスの実現に必要な仕組みが整理できるのではないかと考えている。キュレーションでは、キュレータと呼ばれる情報の編纂者が、さまざまな情報を独自の判断で収集・整理し、新しい価値を持たせて共有し、ユーザは整理された価値の高い情報を受け取ることができる。我々はこれに倣い、「情報流」とよぶ多数の実時間データストリームを、リアルタイムに融合・編纂・配信する実時間キュレーションを実現するプラットフォーム「情報流キュレーション基盤」の実現を目指している。キュレータは知識エキスパートやユーザといったシステムへの参加者であるが、それらエキスパートの知識を反映したイベントフィルタやオンライン機械学習を行った判定器であってもよい。そういった「キュレータ」たちが情報流を操作し活用する基盤を構築することで、蓄積では間に合わない、雑多で超大量の実時間データストリームの高度活用を促進できる。

筆者らは有志による情報流準備検討委員会¹⁾を組織し、シンポジウムの開催などを通して、研究コミュニティを形成するための活動を展開している。ここで

は、同委員会で検討している情報流キュレーション基盤について、その活用シナリオと実現に向けた課題を述べるとともに、関連技術の現状と基盤の基本設計を紹介する。

情報流基盤の活用シナリオ

現在、情報流キュレーション基盤のシナリオとして、マルチソースストリーミングキュレーション、および実時間クラウドソーシングによる都市災害時のセンシング情報集約を検討している。それらの概要を紹介する。

2014年はFIFAワールドカップが開催され、深夜や早朝からサッカーのライブ中継に見入った人も多かったのではないだろうか。テレビのライブ中継では、スタジアムの複数個所にカメラとカメラマンが配置され、ディレクターがカメラマンと協調して映像を取捨・選択することで、放送する映像をリアルタイムに作成し配信する。専用の機材、経験豊富なディレクター、カメラマンにより大多数の視聴者が満足する洗練された映像中継が実現されるものの、2つの意味で今後求められるリアルタイム映像ストリームのトレンドを満たしていないと思われる。第1の課題は、それらの映像にはいわゆる群衆(crowd、ここでは多数の視聴者やファン)の叡智や嗜好が反映されていない点である。人気ブロガーが人をひきつける記事を配信するように、視聴者やファンは独自の価値観に基づきSNSなど複数の情報ソースを組み合わせて新しい価値を生み出すことができるはずである。たとえばあるプレーヤのビッグファンたちがその知識を駆使し、そのプレーヤのゲーム中の動きや活躍度をまとめてファンたちのコメントや付加データとともに中継し、リアルタイムでニッチな解説付き映像にすることも可能となる。このように書くとニコニコ動画のコメントや、Ustream・Twitter連携等で十分と思われるかもしれない。しかしそれらは、リアルタイムな複数ストリーム統合とそれらを整理するキュレータの介在を考慮しておらず、

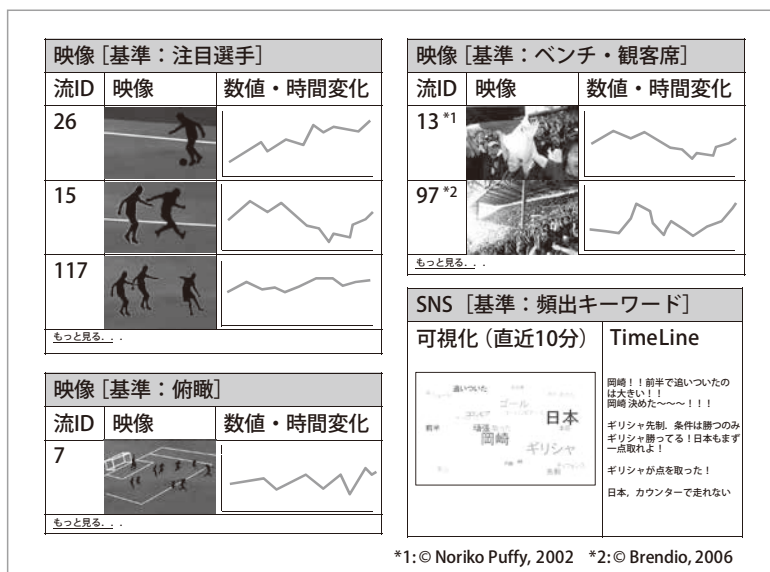


図-1 キュレータ支援画面のイメージ

それが第2の課題となる。情報流キュレーション基盤は、映像やSNSなどの雑多な情報の中から、情報流の高次利用を促すキュレータ支援機能を提供することで、同じ流の集合から多数の異なるキュレーション結果を生み出すことを狙っている。

ここでは観客がさまざまなアングルやズームレベルで撮影・配信しているフィールド上のプレイ映像や観客席等の映像、会場外のユーザが撮影・配信している試合に関するさまざまな映像、SNS上に流れている試合に関するメッセージを「情報流」として扱い、キュレータがリアルタイムに編纂して配信する状況を想定する。顔認識やシーン認識、音声解析などにより映像やSNSストリームの情報流にリアルタイムでタグが付与されることで、キュレータはタグ情報に基づき各情報流の重要度やランキングを判断できる。キュレータの負荷を軽減するためには、たとえば特定の選手が長時間あるいは大きく映っている映像情報流、盛り上がっている観客席が映っている映像情報流などを探し出せるといったような、いわゆる価値判断の半自動化とそのための具体的な要求指定方法も検討すべきである。

この議論に基づくマルチソースストリーミングキュレーション支援ツールのイメージを図-1に示した。キュレータは、映像情報流に関して3つの価値関数(注目選手、俯瞰、ベンチと観客席)を設定したものとし、それらの価値の時間変化が表示される。半自動化の

ためには、それらの映像情報流を、指定した時間のクリップ（たとえば、1分間中に10秒間のクリップ）で選択するフィルタなどがあってもよい。SNS情報流に関しては、直近の頻出キーワードをもとにした動的な価値判断を行うものとし、頻出キーワードとその出現頻度が可視化される。情報流キュレーション基盤では、このような支援ツールをもとに、各時点での映像情報流を出力として用いるか、映像に合成して表示するSNSメッセージとしてどれを選択するかを決定していくことで、図-2に示すようなコンテンツがリアルタイムに作成され配信されるシナリオを想定している。

群衆による情報流生成と情報流キュレーションによる取捨選択が有効なもう1つのシナリオとして都市型大規模災害時の被災状況把握が考えられる。都市型災害では、ビル倒壊や窓ガラス散乱、道路陥没などに伴う通行可能状況の大きな変化、帰宅者や緊急車両でごった返す市街地の状況、交通機関の情報や携帯網の通信可能状況など多種多様な状況が複雑に入りまじり、自治体関係者、警察消防などの救助サイドのみならず、被災者である人々も正確で適切な状況を把握することがきわめて困難であることは、これまでの多数の研究報告に加えて、東日本大震災発生時の東京などの状況を振り返れば言うに及ばない。東日本大震災時では、音声通話が繋がらない環境でSNSの救助要請拡散が実際の救助に結び付いた例もあるものの、個人が発信するSNS情報には（故意か否かにかかわらず）流説や推測が混在するため、信頼性において課題を残している。また、帰宅難民が避難あるいは帰宅行動を判断するのに十分な交通情報や混雑情報を収集したい場合、Twitter画像や映像ストリームの状況把握は有用な情報源となる。しかしそれらは特定の駅の混雑把握といったスポット的なリアルタイム情報としてはきわめて有効であるものの、街全体の状況把握を総合的に行うには散在的であり、その一方でリアルタイム情報として刻一刻と変化していくため、蓄積してはまったく意味をなさない。個々

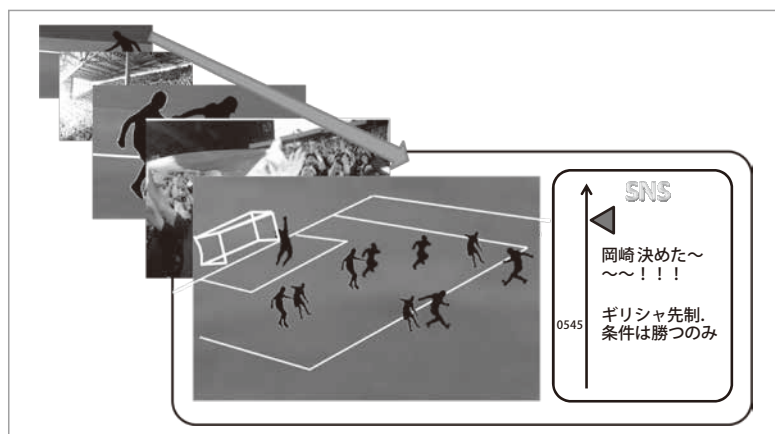


図-2 キュレーション結果のイメージ

人がボランティアとして知識や経験を駆使し、キュレータの役割を担い、そういったスポット的な情報を紡ぎ合わせることで、帰ろうとする道は避難者による過密で身動きが取れないのか、同じように困っている人はどのぐらいいるのか（それを知れば帰宅の途につくべきかの判断材料になる）、列車の運休状況はどうか、代替輸送で待っている人はどれぐらいいるのか、といった面的な情報がリアルタイムで共有できる可能性がある。時々刻々と移動するプローブカーによる通行可能道路マップにスマートフォンユーザからの混雑状況写真の投稿を整理・集約し、被災者からの道路状況のSNSレポートを重畳することで、街全体の被災者の分布や困難の程度を推定したり、警察が交通信号整理や群衆誘導のためにどの現場に人を配分するかなどを判断することも不可能ではない。

そういった群衆（ボランティア）によるデータ処理の概念を広く災害時の状況把握や状況センシングに拡大することで、情報流プラットフォームは災害時の情報共有プラットフォームに発展する可能性も秘めている。サッカーゲームのようにコンテキストが推定しやすい例とは異なり、雑多な画像や不正確な位置情報が入りまじる中でのリアルタイム情報抽出には特に人間の知が必要であり、それも集合知あるいは集合力に頼るクラウドソーシングの概念の導入が不可欠である。被災当事者側から発信されるリアルタイム情報を整理するボランティアキュレータたちの存在は適切なレスキュー活動や早期の混乱収拾につながり、マルチソースストリーミングキュレーションとともに期待さ

れるシナリオである。

情報流に関連する技術

あらゆるモノをネットワーク化しインターネットに接続する、モノのインターネット (IoT, Internet of Things) が登場し、発展・拡大を続けている。IoT では、従来の通信プロトコルにおけるメッセージ交換に加え、モノの周辺の環境やモノ自体の移動・変化に伴い発生する実世界情報流が取得・伝送される。ビッグデータ解析もそういった「超」大量の情報を対象とすることで、実世界のパターンや法則を抽出し役立つ方向へとシフトしつつある。ビッグデータ解析では、3つのV: Variety (多様性), Velocity (頻度), Volume (量) を考慮したデータの解析が行われており、Velocity (頻度) として実時間/準実時間処理を目指しているものもある。しかしその多くは Hadoop 上の大規模並列処理で処理時間を短縮する等の試みであり、多数の情報流の実時間キュレーションには適用できない。IoT によりインターネット接続された人・車・機械などから連続的に生成される位置トレースデータや定点カメラからの映像ストリーム、実時間ソーシャルコンテンツなど、個々の情報は我々の想定を遙かに超えて拡大しつつある。多数の「情報流」を蓄積することなく即時的かつ複合的に活用するための概念や基盤は IoT の観点からも非常に重要である。

筆者らが提唱する情報流基盤は、Complex Event Processing (CEP) もしくはその関連技術とつながりが強い。CEP は多数の実時間情報ストリームを処理するためのデータ処理基盤であり、最近では SNS のリアルタイム解析基盤としても活用されている。たとえば Apache Samza プロジェクトでは分散 Pub/Sub 型のメッセージングプラットフォーム Apache Kafka と Hadoop2 (YARN) のような次世代 MapReduce の分散データ基盤を併用し、リアルタイムメッセージングサービスをスケーラブルに実現できる。オンライン機械学習を用いてデータのリアルタイム高度利用を実現する基盤も現れており、Jubatus や Samoa などが知られている。そのほかにも Apache S4²⁾ や Aurora³⁾

などがオープンソースソフトウェアとして公開されている。商用では IBM の InfoSphere Streams⁴⁾ などが知られており、情報ストリームの分岐・合流、抽出・集約、加工、簡易分析のためのオペレータセットを提供している。しかしこれらでは整形後の情報流の低位処理と流通に着目しており、入力情報流の発見、獲得およびタグ付け、情報流の解析と理解、複数情報流の高度な融合やキュレータ処理といった高次編纂の包括的機能は提供されていない。情報流キュレーション基盤では、これらのスキームを従来の CEP 技術やプラットフォームの上位層、もしくは複数機能を包括する機能プレーンとして実現することを目指している。

日欧共同プロジェクトである ClouT⁵⁾ では、IoT や参加型センシングにより生成される情報ストリームを取得し組み合わせることで新たなサービスを実現するための基盤を提案・開発している。情報ストリーム間のインターオペラビリティを実現するため、センサやデバイスを仮想化してサービス化する技術や、Web などの情報を共通センサ化する技術などが開発されている。筆者らが提唱する情報流キュレーション基盤は、ClouT が実現するセンサ仮想化や共通化といった情報ストリームの準リアルタイム処理に必要な技術を取り入れながら、高次処理やコンテキスト依存型ストリーム編纂の実現、それに伴うユーザやデータホルダのプライバシー確保、ユーザ参加型基盤の実現など、より包括的な概念の実現に注力した新しいプラットフォームを目指す点で ClouT とは異なっている。

情報流キュレーション基盤の基本設計

情報流キュレーション基盤を実現するためには、(1) 情報流を生成する技術やそれを容易に検索するための技術が不可欠である。次に、(2) キュレータによる情報流編纂を支援するための、情報流のコンテキストやコンテンツ理解の支援技術が必要である。また、(3) 選別した情報流をキュレータの判断や資源の制約に従って加工(要約、抜粋など)する技術も求められる。さらに、(4) 複数の情報流を1つのコンテンツに融合・編纂するための技術が必要になる。多数

の情報流を扱うには、複数のキュレーション（選別・加工・融合・編纂の組合せ）を階層的に接続し、多数の情報流を実時間で分散処理する機能が必要である。加えて、情報流には、個人を特定する位置情報や個人情報そのものである生体情報などが含まれる。そのため、(5) プライバシーの漏洩を防止するセキュリティ技術が必須となる。

情報流準備検討委員会では、前述の各技術課題を解決するためのリアレンジアーキテクチャ「情報流クラウド(Cloud)」を検討している。その概念図を図-3に示す。

情報流クラウドは、クラウド内に、A. 情報流獲得機能、B. 解析・理解機能、C. 加工機能、D. 編纂・提示機能、E. 分散処理機能、F. セキュリティ機能がサービスとして実装されることを想定している。以下は、それぞれの機能について概要を述べる。

Aは、情報流の選別・配送・生成依頼からなる。多数の情報流から必要な情報流を選別する機能では、地理的・種別的に分散・並行して生成される多数・多様な情報流に対し、キーワードや地理情報といった従来の検索の仕方を超えて、より抽象度の高いレベルで検索できる仕組み、たとえば、キュレータの興味、価値判断、必要性に基づき、関連する情報流を発見できる機能を実現する。前述のサッカーの例(図-1)では、注目選手、俯瞰、ベンチ・監督といった興味・価値判断に基づき、価値関数値の高い映像が自動的に検索されキュレータ支援画面に表示されている。情報流の配送機能では、発見した情報流をキュレータにリアルタイム配送するため、情報流発生源(モバイルノードやセンサノード)とキュレーション実行マシンとの間に必要な通信帯域を確保する。加えて、利用可能な通信帯域を計測する機能、利用可能通信帯域範囲内に情報流の粒度を調整する機能を実現する。情報流の生成依頼機能では、必要な情報流を獲得できない際に、参加型センシング(広範囲あるいは複数の興味地

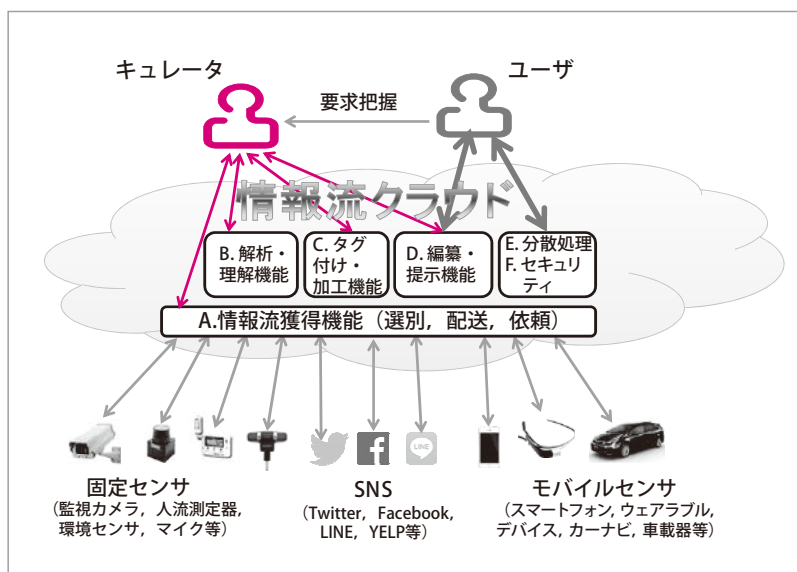


図-3 情報流キュレーション基盤の概念図

点(PoI)におけるセンシングをボランティアユーザへの依頼により行う仕組み)に基づき、キュレータが必要な情報流の生成を依頼する機能を実現する。

Bは、情報流の可視化・変化予測からなる。情報流の可視化機能では、情報流の動的な性質を考慮し、各情報流について過去の時間変化を一目で把握できる時系列可視化機能や位置情報を含む情報流については、空間的な分布およびその時間変化の把握を容易化する時空間可視化機能を実現する。情報流の変化予測機能では、キュレータが、複数の情報流から各時点で適切なものを選別し、コンテンツを編纂できるようにするために、現時点ではコンテンツの編纂に必要でないが、未来のある時点で必要になる可能性がある情報流を事前に把握できるようにする。そのため、情報流における変化の傾向を解析により求め、未来の変化パターンの予測・可視化や、あらかじめ指定した条件を満たす時刻を予測しキュレータに提示する機能を実現する。例として、図-1のキュレータ支援画面には、検索された各映像について価値関数値の時間変化が表示されている。

Cは、情報流へのタギング・加工からなる。情報流へのタギング機能では、センサが生成する一次的な情報流(一次情報流)が構造化されておらず容量も大きいため、そのままでは、キュレーションするのは難しいことに配慮し、情報流の任意の区間に対し、メタ情

報(タグ)を付与する機能を実現する。一次情報流へのタグgingは人が手動で行う方式とデータ解析(画像解析含む)により自動で行う方式が考えられる。たとえば、図-1のキュレータ支援画面における検索を実現するためには、注目選手やベンチ・観客席が映っているか、プレイ映像はアップか俯瞰かなどを示すタグが、映像中の各シーンに付与されている必要がある。情報流の要約・切出機能では、複数の情報流をキュレーションする際、キュレータが、それぞれの情報流に対しある価値判断のもとで決定した重要度から、編纂するコンテンツに含める各情報流の粒度を決める必要がある。数値データからなる情報流の要約は、既存のストリームコンピューティング処理系で利用可能な統計処理等をベースに実現可能だと考えられる。映像や音声の情報流に対しては、付与したタグの情報をもとに、指定した条件に従って一部を切り出す機能を実現する。

Dは、情報流の融合・編纂・提示のためのインタフェースである。選別・加工した情報流を融合しコンテンツとして容易に編纂・提示できるようにするため、情報流に適した新しいインタフェースを考案・実現しなければならない。インタフェースとして、複数の情報流を融合し1つのコンテンツとして編纂するため、一貫した1つのビューに複数情報流をマッピングする機能を設計・実現する。その際、情報流の種類(数値、音声、映像など)に応じた臨場感が高いインタフェースや複数の情報流を複合的かつ直感的に表現するインタフェースを考案する。図-2のキュレーション結果のイメージは、複数の映像流とSNS流を、1つのビュー(映像枠とSNS枠が各1つ)にマッピングしているが、複数の映像枠やSNS枠を持つビューを定義・利用することもできる。また、キュレーション自体をクラウドソーシングにより進化させるため、キュレータが自身で新たにインタフェースを設計・適用できる機能や、過去の編纂により作成されたインタフェースを蓄積し共有、再利用できる仕組みを実現する。

Eは、編纂後コンテンツの情報流としての再配信とキュレーションの登録・共有からなる。編纂後コンテンツを新たな情報流として生成・配信する機能では、情報流の選別・加工はキュレータが行うため、同時に



図-4 第1回情報流シンポジウムの様子

扱える情報流の数は人間の処理能力に依存して抑えられてしまうことを考慮し、スケーラビリティを高めるため、編纂したコンテンツを新たな情報流として再配信し、さらなるキュレーションが適用できるようにする。そのため、複数の情報流から構成される高次情報流を定義可能な情報流の共通フォーマットを定義する。キュレーションの登録・共有機能では、過去のキュレーションにおいて評価の高いものをフィルタとして登録・共有し、情報流の自動キュレーションを行える機能を実現する。

Fは、情報流の改ざん防止、情報流提供者のプライバシー保全を実現する。キュレーションにより一次情報流は加工され情報量が減る。情報流の正当性証明のため、編纂後の高次の情報流がオリジナルの情報流から生成されたものであることを保証する機能を実現する。また、安心して情報流を提供できるようにするため、K-匿名化や追跡の抑制、複数情報流間の関連付けの抑制を行う「プライバシーフィルタ」を実現する。

第1回情報流シンポジウム報告と今後の活動

情報流検討準備委員会は本会 MBL/UBI 研究会/電子情報通信学会 MoNA 研究会の協力を得て、2014年3月15日土曜に慶應義塾大学日吉キャンパスにおいて「第1回情報流シンポジウム～溢れ出る情報を「流」として捉える新研究トレンド～」を実施し、70名超の聴衆が集まる大盛況の会となった(図-4)。「情

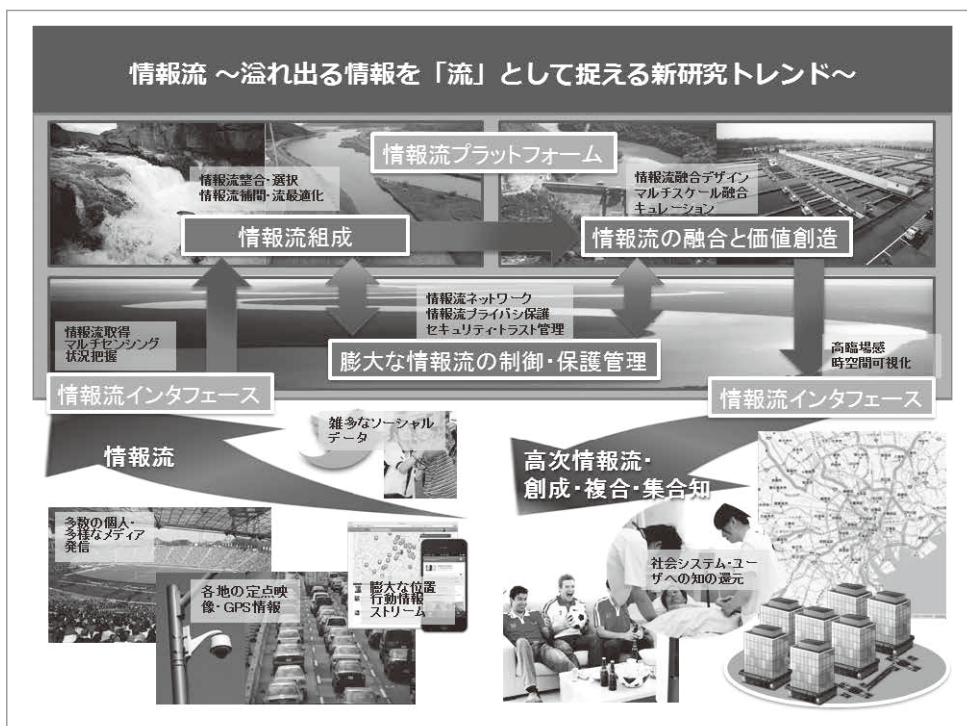


図-5 情報流の俯瞰

「情報流」とは何かを俯瞰し、講演者、参加者で概念や目的を共有するとともに、第1次コミュニティの形成に成功する非常に有意義なキックオフとなった。引き続きコミュニティの拡大を図るとともに、異分野研究者への働きかけや連携も図っていきたいと考えている。なお、2014年9月には本会と電子情報通信学会共催のFIT 情報科学技術フォーラム、および電子情報通信学会主催のソサイエティ大会でそれぞれ関連イベント企画が予定されている。

まとめ

我々にとって従来の「情報」は、インターネット上などに「静的」に整理・蓄積され、人々の用途に応じて加工され、必要とする人々に伝送・活用されるものであった。近年のビッグデータ活用もこの延長線上のトレンドとみなすことができる。一方で、人・車・機械などから時々刻々と連続的に生成される位置トレースデータや映像ストリーム、ソーシャルコンテンツといった実時間情報を、「リアルタイム」で「複合的に」処理することで社会的に大きな価値を生み出せる。情報システムは、従来の「情報」の想定を遙かに超え

る量で湧出する「情報流」を活用すべき時代へ移行しつつある。そのような時代には「情報流」を対象とした技術の確立が急務である(図-5)。

本稿が「情報流」研究コミュニティの拡大、さらには新たな研究トレンドの確立につながれば望外の喜びである。

参考文献

- 1) InfoFlow (情報流) プロジェクト, <http://www.infoflow.org/>
- 2) Neumeyer, L., et al. : S4 : Distributed Stream Computing Platform, 2010 IEEE Int'l. Conf. on Data Mining Workshops (ICDMW), pp.170-177 (2010).
- 3) CHERNIACK, Mitch, et al. : Scalable Distributed Stream Processing, Proc. of First Biennial Conf. on Innovative Data Systems Research (CIDR. 2003), pp.257-268 (2013).
- 4) IBM : InfoSphere Streams, <http://www-03.ibm.com/software/products/ja/infosphere-streams>
- 5) ClouT : Cloud of Things for Empowering the Citizen Clout in Smart Cities, <http://clout-project.eu/jp/>

(2014年7月1日受付)

安本慶一 (正会員) yasumoto@is.naist.jp

奈良先端科学技術大学院大学情報科学研究科教授。モバイルコンピューティング、ユビキタスコンピューティングに関する研究に従事。

山口弘純 (正会員) h-yamagu@ist.osaka-u.ac.jp

大阪大学大学院情報科学研究科准教授。モバイルコンピューティングとネットワーク・アプリケーションに関する研究に従事。