

通信事業者の業務を下支えする ビッグデータ分析基盤

武田 喜美子 山下 孝行

要旨

通信事業者は、契約者一人当たりの売上向上のため、一人ひとりの趣味・嗜好に合ったサービスの迅速・タイムリーな提供を求められています。そのため、従来の分析専門家に加え、サービス提供者自身がその知識・知見をもとに顧客データや各サービスの分析を直観的に素早く行い、サービスの創出・改善に努める必要があります（セルフサービス型BI。以下、セルフBI）。本稿では、分析専門家が使う既存の分析環境に追加した、セルフBI環境の導入から定着までの一連の手法を紹介します。



影響分析/メタデータ/セルフサービス型BI/セルフBI/データソース/データマート/DM

1. はじめに

通信事業者において、スマートフォンなどモバイル機器によるデータ通信の収益が事業の大きな柱となっています。通信料収入を増やすためには、通信インフラの改善（ネットワークや基地局配置の改善）に日々取り組むだけでなく、それと並行して、付加価値の高いサービス提供にも取り組む必要が通信事業者にはあります。

この高付加価値サービスとは、スマートフォン上で展開する各種コンテンツ配信サービスがその一例で、顧客の楽しみをより充実させることで、通信料収入増収につなげることを目指しています。つまり、顧客一人ひとりの趣味・嗜好に合った魅力あるサービスの迅速かつタイムリーな提供が求められています。

一方、サービスを顧客一人ひとりによりフィットさせるためには、これら膨大なデータのなかから必要かつ正確なデータを取捨選択し、目的にあった分析を行い、その結果を読み解く必要があります。

このような状況において、より迅速なサービス提供のため、サービスの知識・知見を持つサービス提供者が、素早く自分の見たい視点で分析を行うセルフサービス型BI（Business Intelligence。以下、セルフBI）が必要と

なっています。

本稿では、このセルフBIの導入から定着までの一連の手法を紹介します。

2. 通信事業者におけるビッグデータ分析業務の 現状と課題

2.1 現状

(1) 分析業務

ビッグデータ分析業務は、定型分析業務と随時分析業務に大別できます。定型・随時分析業務は、更に高度分析と通常分析に分類できます。これらの分析は、通常、通信事業者内のサービス提供者から分析専門家に依頼され、分析レポートとして納品されます。そのため、サービス提供者には、待ち時間が発生します。また、分析レポートが固定フォーマットのため、自分の見たい視点で自由自在にいろいろな角度からデータ検証を行うことが困難となっています（図1）。

(2) 定型分析

この分析は、日次、週次、月次といった時間軸のデータを決まった手法で集計し、定型レポートを作成します。主な定型分析は、以下となります。そのうち1)

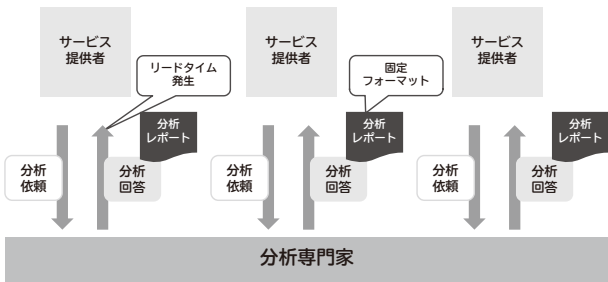


図1 分析業務の一例

が定型分析案件数の大部分を占めています。

- 1) 各種サービスの利用実績集計（通常分析）
- 2) 各種サービスへの入会・退会実績集計（通常分析）
- 3) KPI (Key Performance Indicator)、KGI (Key Global Indicator) の実績集計（通常分析）
- 4) サービス継続期間、サービス利用頻度、利用コンテンツなどをクラスタ分析することで得られる同じニーズを持つ顧客の塊の抽出（顧客セグメンテーション）（高度分析）
- 5) 直帰率/離脱率/CVR^{*1}/カート破棄率などによるWebサイトの動線分析（高度分析）

(3) 随時分析

この分析では、キャンペーンなどのある特定ニーズに対する、その場限りの限定的なレポートを作成します。定型分析で得られた結果の詳細分析や、キャンペーンなどの各種施策を講じるための対象者抽出といったものがその一例です。主な随時分析は次のとおりで、7) が随時分析案件数のなかで大きな割合を占めています。

- 1) 各種サービスへの入会・利用・退会状況の問題点抽出と施策の示唆（高度分析）
- 2) 優良顧客の入会・退会経路の見える化（カスタマージャーニー）（高度分析）
- 3) 他サービスへの入会確率の予測（入会予兆モデル）（高度分析）
- 4) サービスからの退会確率の予測（退会予兆モデル）（高度分析）

- 5) 各種コンテンツのレコメンド抽出（レコメンドモデル）（高度分析）
- 6) 優良顧客育成に向けてのコンタクト方法・コンタクトタイミングの予測（高度分析）
- 7) キャンペーンなどの各種施策の対象者抽出（顧客セグメンテーション）（高度分析）

(4) データソース^{*2}

これら定型、随時分析で用いられる主なデータソースは、顧客属性データと顧客行動データ（視聴履歴、検索履歴、ダウンロードログなど）、各種コンテンツのログデータです。これらデータソースの大きな特徴は、データが数百万件から数十億件の規模に及ぶ場合があること、さまざまな形式のデータを扱う場合があることです^{*3}。

2.2 現状の課題と改善策

次に、図1の分析業務の依頼者であるサービス提供者に着目し、分析業務の課題とその要因、改善策を図2に整理しました。

従来の分析業務では、サービス提供者が分析専門家に分析レポートの作成を依頼し、分析専門家がレポートを作成・提供するやり方でした。このやり方だと、どうしてもレポート作成の待ち時間が発生し、分析PDCAサイクルを高速に回すことができません。その具体的な要因は、次

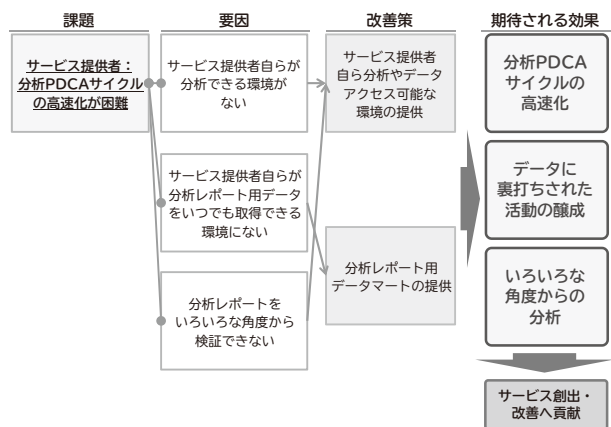


図2 課題とその改善策

^{*1} CVR (Conversion Rate:コンバージョン率) : 広告などのWebサイトから得られる成果の割合。成果は、資料請求やサービス購入など。

^{*2} データの入手元。各種データベースやファイルなど。

^{*3} データ使用に関しては、各国法令などの制約がある場合がある。

のとおりです。

- 1) サービス提供者自らが分析できる環境がない
- 2) サービス提供者自らが分析レポート用データをいつでも取得できる環境にない
- 3) 分析レポートをいろいろな角度から検証できない

これらの改善策としては、サービス提供者自らが分析を行うセルフBI (1) と3) の改善策) とセルフBIのデータソースとなるデータマート (DM) (2) の改善策) の導入・定着があります。

3. DMとセルフBIの導入・定着

3.1 DMの導入

DMは、分析PDCAを高速に回すために遅延なく提供できるように、次の流れで導入します。

(1) DMのデザイン

DMは、マスター系のデータと実績系のデータに大別できます。マスター系のデータは、顧客名やサービス名といった基本情報の集まりです。一方、実績系のデータは利用回数やダウンロード回数など、やりとりのデータが該当し、大部分が時系列データとなります。そのため、このような大量の時系列データは、週月・4半期といった単位でパーティショニングを行い、アクセス性能の高速化を図ります。また、サービスごとにデータはアクセスされるので、個々のサービスごとにデータを固めて配置し、アクセスの局所化を図ります。

(2) データソースの選定

DMを生成するには、複数のデータソース (マスター系と実績系) からデータを取得する必要があります。(3) のDM作成処理の実装にかかわるのですが、多数のデータソースを一度に結合してDMを作成すると、CPU/メモリといったリソースの大量消費につながります。この状況を回避するため、一度に結合するデータソースの数を2から3に抑え、数回にわたって処理を実施する仕組みを実装します。

(3) DM作成処理の実装

実績系データはマスター系データに比べ、大幅にデータ件数が多くなります。このようなデータを処理する場合、結合の仕方に処理性能が大きく依存するので、環境にあった結合方法を選択します。

また、処理作成及び保守効率化のため、メタデータを活用したGUIツールも導入できます。ツールにより、さまざまなデータソースのメタデータ (定義・属性情報) や処理情報を採取し、データの所在・定義やデータ・処理の流れを可視化できます。これにより、瞬時にデータ・処理フローを把握 (影響分析) でき、業務上正しくないデータがあった場合に、迅速な原因究明が期待できます (図3)。更に、必要に応じてデータソース用アダプターをGUIで追加・変更すれば、さまざまなデータに対する収配信を効率的かつ柔軟に実装できます。

(4) 処理のスケジューラへの登録

DMは、夜間の日次バッチ処理で毎日大量に作成・更新されます。DMは、サービスの種類によりデータ量

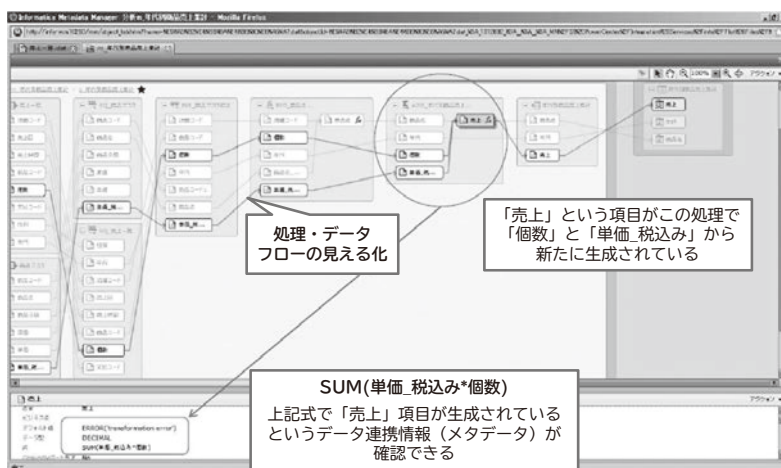


図3 影響分析の一例

や数にばらつきがあります。また、システム障害を起こし得るデータ（例えば、改行コード、カンマなど）が含まれている場合もあるため、これらを踏まえ、処理の並列化や待ち合わせ、チェックポイント・リスタートといった方法をバッチ処理に吸収し、既定時間内に完了するようスケジューリングします。

(5) データのマスキング

このDM作成の際に必要なに応じてデータをマスクします。ハッシュ化などのマスク処理に予想以上の処理時間を要し、所定スケジュールでのデータ分析の実施が難しくなる場合があります。よって、処理方式は、データのマスキング処理時間も考慮した実装が必要となります。

3.2 セルフBIの導入

セルフBIは、サーバとクライアントで構成されます。サーバは公開された分析環境とデータセキュリティを、クライアントは主に個人用の非公開分析環境を提供します。セルフBIは、次の観点で導入の検討を行います。

(1) セルフBIツールの選定

クライアントとしては、エンドユーザーが使いやすい、直感的に使える分析ツールを選定します。

サーバとしては、以下の(2)から(7)の観点で検討を行い、選定します。

(2) 可用性

分析PDCAを高速に回すため、サービス担当者がいついかなるときにもデータ分析が可能となる環境が必要です。そのため、HWのみならず、ツールも可用性の構成をとることができるサーバの導入を検討します。

(3) 性能

レスポンスが遅いと使われなくなってしまうため、性能を意識した環境を用意します。セルフBIクライアントから直接巨大なDMをつと分析するのは非効率なため、そのサブセットとなるデータセット（抽出ファイル）をクライアント側に作成する場合があります。

(4) 監視

可用性とも通じますが、セルフBI環境で不測の事態が発生したときの素早い対応のため、監視の仕組みを実装します。

(5) 認証

セルフBIのクライアントから直接データソースをアク

セスする実装も可能ですが、データへのセキュリティを担保するため、セルフBIサーバを経由したデータのアクセス制御を実装します。併せて、全社の認証基盤との連携も検討、ユーザー運用を定義します。

(6) 権限

認証と同様に、データへの読み書きや分析結果の公開をセルフBIサーバ側で制御します。

(7) Version up

BIツールは、ライフサイクルが短い傾向にあるため、Version upが頻繁に発生しやすい製品です。複数Versionが混在した場合、データ共有などが困難になる可能性があるため、SW構成管理の仕組みを実装し、スムーズなVersion upを実現します。

3.3 定着

セルフBIとDM導入後、定着化を行います。定着化の検討に必要な主なタスクは以下となります。これらの実施により、セルフBIの浸透とサービス提供者のスキル向上が期待できます。

(1) テンプレート、ガイドの提供

分析テンプレート、ガイドを、定期的にエンドユーザーであるサービス提供者へ提供します。テンプレート作成側は、エンドユーザーの意見を取りまとめ、最大公約数のものを提供します。

(2) 定期的なトレーニングの提供

定期的なトレーニングや新規ユーザー向けのハンズオンを開催します。セルフBIだけでなく、データリテラシー向上のため、一般的な数学や統計、データベース、SQLの知識習得の場も提供します。

(3) ヘルプデスクの導入

サービス提供者の近くにヘルプデスクを設けます。これにより、face to faceでの迅速な問い合わせ対応を可能とします。

(4) BIツール資格取得の奨励

BIツールベンダーが実施しているBIツール資格の取得を奨励し、分析力向上の見える化とモチベーションアップを図ります。

(5) データ・処理の見える化

分析対象のデータを効率的に見つけられるように、また、データ不正をできるだけ早く検出できるように、データの見える化の仕組みを導入します。

4. おわりに

以上、通信事業者の業務を支えるビッグデータを活用した分析基盤であるセルフBIとDMの導入から定着までを紹介しました。ビッグデータは、現状、「データから情報へ」のBusiness Intelligenceの段階にあります。今後は「情報から知識、知性」への段階であるBusiness Analyticsや知識発見的データマイニング（セルフBI）、Artificial Intelligenceが注目されてくると予想しています。

執筆者プロフィール

武田 喜美子

通信業ソリューション事業部
シニアエキスパート

山下 孝行

クラウドプラットフォーム事業部
主任

関連 URL

データ処理フロー可視化ツール

<http://www.nec.co.jp/pfsoft/powercenter/>

セルフBIツール

<http://www.nec-solutioninnovators.co.jp/sl/tableau/>

NEC 技報のご案内

NEC 技報の論文をご覧いただきありがとうございます。
ご興味がありましたら、関連する他の論文もご一読ください。

NEC技報WEBサイトはこちら

NEC技報(日本語)

NEC Technical Journal(英語)

Vol.68 No.3 新たな価値創造を支えるテレコムキャリアソリューション特集

新たな価値創造を支えるテレコムキャリアソリューション特集によせて
変革期を迎えたテレコム産業に向けた NEC のソリューション

◇ 特集論文

ネットワークに新たな価値を提供する SDN/NFV ソリューション

SDN/NFV ソリューション技術体系
ネットワークのインテリジェントな運用管理を実現する MANO 技術
vEPC におけるユーザープレーン制御の実現
付加価値の高い MVNO ビジネスを支援する vMVNO-GW
通信事業者向け仮想化 IMS ソリューションへの取り組み
NFV で実現する IoT ネットワーク
通信事業者向けトランスポート SDN ソリューション
通信事業者の収益向上を実現するトラフィック制御ソリューション (TMS)
トラフィック制御ソリューション (TMS) の要素技術

トラフィックの増大に対応するトランスポートシステム

大規模データセンター向け OpenFlow イーサネットファブリック
増大するトラフィック対応に向けた 10G-EPON の開発
大容量基幹ネットワークを支える要素技術とマルチレイヤ統合トランスポート装置
光デジタルコヒーレント通信技術の開発
光海底ケーブルシステムを支える大容量光伝送技術

無線アクセスの高度化に対応するワイヤレスソリューション

ロシアでの通信事業者向けネットワーク最適化プロジェクト
サウジアラビアモバイル通信事業者向け大容量無線伝送システムを実現する iPASOLINK ソリューション提案
世界最高の周波数利用効率を実現する超多値変調方式用位相雑音補償方式の開発
モバイル通信の高度化を支える高密度 BDE

通信事業者向け ICT ソリューション

NEC Cloud System の競争力強化と OSS モデル構築 SI 技術への取り組み
会話解析ソリューションの通信事業者への適用
止まらないキャリアシステム開発への取り組み
通信事業者の業務を下支えするビッグデータ分析基盤

◇ 普通論文

セキュアな重複排除型マルチクラウドストレージ「Fortress」

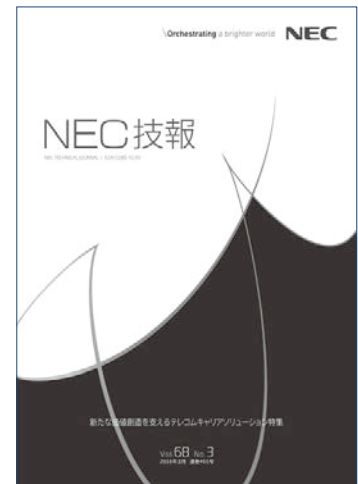
◇ NEC Information

C&C ユーザーフォーラム & iEXPO2015 Orchestrating a brighter world

基調講演
展示会報告

NEWS

2015 年度 C&C 賞表彰式開催



Vol.68 No.3
(2016年3月)

特集TOP