

Technology Reports

通信と放送を融合した新しいサービス — NOTTV —

NOTTVの利用を促進するレコメンデーション技術

2012年4月にスマートフォン向け放送局「NOTTV^{*1}」が開局し、放送と通信を融合した新サービスがスタートした。NOTTVでは利用促進を目的として、ユーザの視聴状況や端末操作状況を基に、コンテンツ情報をおすすめするレコメンド機能を提供している。通信を利用し、端末とレコメンドシステムの間で、情報を送受信することで、おすすめコンテンツ、自動予約制御、ランキングなどの情報や機能を提供可能とするレコメンドシステムを開発した。これによりユーザごとの嗜好に基づいたサービス提供を可能とし、NOTTVの利用促進を実現した。

サービス&ソリューション
開発部

いがらし けい たちばな ゆうや
五十嵐 圭 立花 優也
みやがわ さとし くのち かつひろ
宮川 聡 野口 勝広

1. まえがき

NOTTVでは多種多様なコンテンツが放送されており、多彩なジャンルや、リアルタイム視聴 (RT)^{*2}やシフトタイム視聴 (ST)^{*3}といった放送種別が存在する。しかしながら、このように充実したコンテンツに対しては、ユーザの検索のみによる選択には限界があり、コンテンツ視聴機会の損失につながってしまうというおそれがあるため、この問題を解決するためにNOTTV向けの高性能なレコメンドシステムを開発した。レコメンドシステムは、モバキャスTM^{*4}の双方向による通信を利用することで実現している。

レコメンドサービスの導入によ

り、ユーザに対して興味のあるコンテンツへの気付きを与え、NOTTVの利用促進やコンテンツ視聴機会の

増大を図ることが可能となる。端末は本システムが出力したレコメンドリストを取得し、図1のようにおす



図1 おすすめコンテンツ表示

© 2012 NTT DOCOMO, INC.
本誌掲載記事の無断転載を禁じます。

† 現在、スマートコミュニケーションサービス部

*1 NOTTVTM：NOTTVおよびNOTTVのロゴは、(株)mmbiの商標または登録商標。
*2 リアルタイム視聴 (RT)：従来のテレビと同様に、放送波で配信されている番組をリアルタイムに視聴すること。

すめコンテンツとして画面に表示，あるいは，自動蓄積対象コンテンツとすることによって，ユーザに対してコンテンツ視聴への導線を提供する。

本稿では，上記実現に向けた課題やシステム構成，個別の技術や機能について解説する。

2. レコメンドシステム開発における課題

レコメンドシステムで考慮すべき本サービスの特徴としては，①多種多様なコンテンツ形態が存在すること，②コンテンツが放送波で提供されるため頻繁に入れ替わること，③新規利用ユーザには操作履歴が蓄積されていないこと，が挙げられる。これらの特徴に起因する本開発の課題は以下の3点である。

・NOTTV向けレコメンドアルゴリズムの確立

数多くのコンテンツの中からユーザが興味をひかれるものを精度よくレコメンドすることが重要となる。特に従来技術[1][2]では，サービス開始直後のように操作履歴が少ない状況において，レコメンド可能なコンテンツが限定されるというコールドスタートと呼ばれる問題が発生する。NOTTVでは，これらの問題を解決するレコメンドアルゴリズムが必須となる。

・柔軟なレコメンドオペレーションの実現

多種多様なコンテンツ形態に

対しては，さまざまなレコメンドルールやパラメータなどを設定できる必要がある。また，事前に適用したレコメンド設定の効果を定量的に測定し，その評価を踏まえて改善に向けたチューニングが可能なオペレーション環境が求められる。

・高速データ処理技術の実装

本サービスでは計算対象とすべきコンテンツが頻繁に入れ替わるため，短時間でレコメンドを更新することが不可欠である。

3. レコメンドシステムに実装されている技術

3.1 レコメンドアルゴリズム

本節では，NOTTV向けレコメンドシステムに実装されているアルゴリズムとして，コンテンツベースフィルタリング，Hybrid協調フィルタリングについて概要と特徴を説明する。このレコメンドアルゴリズムにおいては，コンテンツベースフィルタリングにより，各ユーザの嗜好にあったコンテンツの選定が実現され，Hybrid協調フィルタリングにより，レコメンドするコンテンツバリエーションの広がりをもたせることが可能となる。

(1)コンテンツベースフィルタリング

コンテンツベースフィルタリングには，ユーザ情報ベースのレコメンド方式とアイテムベースのレコメンド方式が存在し，本システムでは両方式を実装している。

前者のユーザ情報ベースの方式は，コンテンツの特徴とユーザの嗜好をもとにレコメンドを生成する手法である。

本アルゴリズムにおけるユーザ情報ベースのレコメンド方式は，図2のようにコンテンツやユーザの嗜好を特徴ベクトル^{*5}として表現する[3]。コンテンツ特徴ベクトルは，特定のワードがメタデータ内に出現する頻度から計算される。ユーザ特徴ベクトルは，各ユーザの操作履歴をもとにユーザの嗜好をベクトル表現する[4]。ユーザ特徴ベクトルとコンテンツ特徴ベクトルの類似度^{*6}を計算し，各ユーザの嗜好に合わせたコンテンツがレコメンドされる。

次にアイテムベースのレコメンド方式について述べる。本アルゴリズムは，各コンテンツに対して類似度の高いコンテンツを選定する手法である。

あるコンテンツについて，他コンテンツの特徴ベクトルとの類似度を計算し，その値の高いものを関連コンテンツとしてレコメンドする。

端末上で所定のコンテンツが表示される際に，関連コンテンツもユーザへ提示することにより，他コンテンツへの気付きや導線を与えることができる。

(2)Hybrid協調フィルタリング

Hybrid協調フィルタリングは，嗜好の近い他ユーザの特徴ベクトルとコンテンツ特徴ベクトルをもとにレコメンドを生成する。

本アルゴリズムでは各ユーザについて，本人のユーザ特徴ベクトルと

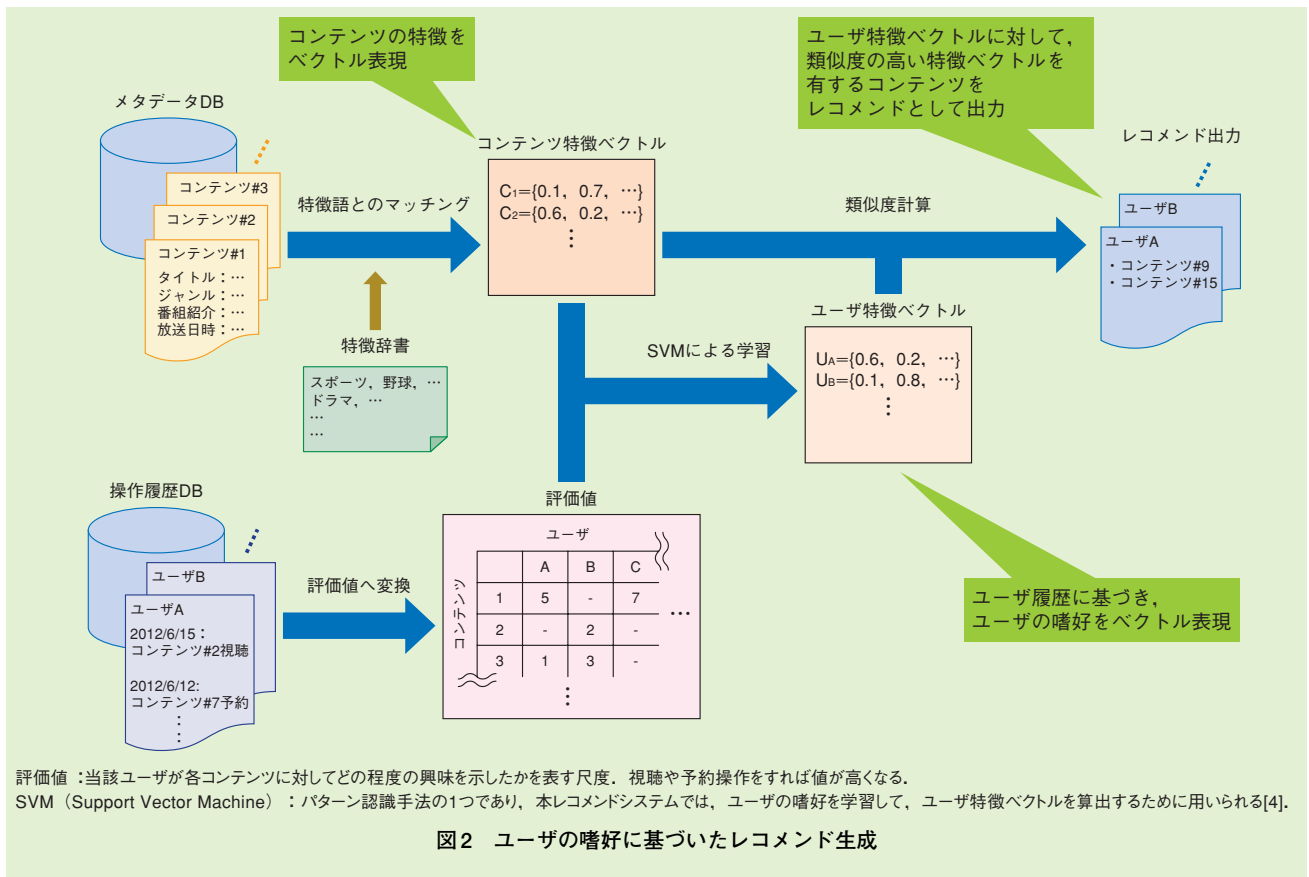
*3 シフトタイム視聴 (ST) : あらかじめ登録し，自動でモバキャスト端末内に蓄積されたコンテンツをユーザが好きなときに視聴すること。なお，NOTTVではRTで流されるものを「番組」，STで流されるものを「コンテンツ」と使い分けている

が，本稿では「コンテンツ」という呼び方で統一している。

*4 モバキャスト™ : 「モバキャスト」および「モバキャスト」のロゴは(株)ジャパン・モバイルキャストの商標。

*5 特徴ベクトル : 特徴語の出現頻度を数値化してベクトル表現したもの。

*6 ベクトルの類似度 : 所定のベクトルと他のベクトルがどの程度似ているかを表すための尺度。この値が大きければ互いが似ているということになる。



類似度の高い他ユーザの特徴ベクトルから、新たなユーザ特徴ベクトルを生成する[5]。これにより、嗜好の近い他ユーザの特徴を取り込んだユーザ特徴ベクトルが得られる。その後、ユーザ特徴ベクトルとコンテンツ特徴ベクトルとの類似度が算出され、レコメンドコンテンツが決定される。

本手法は、従来技術のように類似ユーザが視聴したコンテンツをそのままレコメンドするのではなく、ユーザ特徴ベクトルを更新してからレコメンドコンテンツを決定する。これによりユーザ全体として未視聴のコンテンツもレコメンド対象とする

ことができるため、コールドスタートの問題を解決できる。

3.2 レコメンドオペレーション

レコメンドシステムでは提供するサービスに応じた細やかな設定値のチューニングが重要であり、確かな効果を上げるためには「仮説→設定→評価→改善」のサイクルを繰り返す。適切な設定値を見極める必要がある。どのようなレコメンドが効果的であるか「仮説」をたて、実際に「設定」してレコメンドを提供した後、その結果を「評価」して、必要に応じて設定値変更などの「改善」を行う。NOTTV向けレコメンドシ

ステムでは、上記サイクルを1つのオペレータ画面でサポート可能である。オペレータ画面を利用したレコメンドオペレーションのイメージを図3に示す。

(1)レコメンドの各種設定

本システムではオペレータがユーザに提供したいレコメンドの単位であるレコメンド種別の新規作成や削除、編集が可能なオペレーション環境を実現している。作成したレコメンド種別には前述のアルゴリズムと任意のルールを組み合わせて設定することができる。レコメンドルールとしては、コンテンツのジャンルやタイトルなどのメタデータ、ユーザ

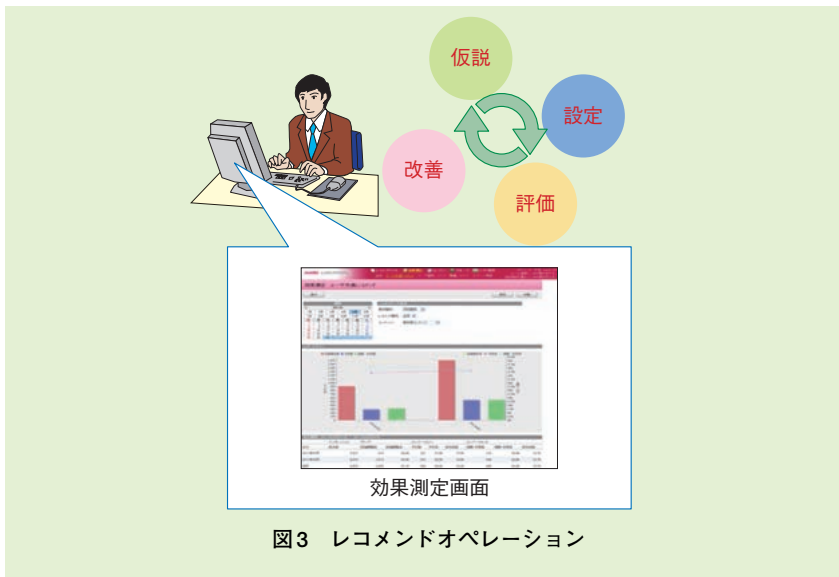


図3 レコメンドオペレーション

の属性、および、操作履歴を基に、コンテンツの追加や削除、ソートを行うことが可能である。また、作成されたレコメンドリストはリアルタイムに確認や編集が可能であり、特定のコンテンツの追加や並び順の変更など、コンテンツ単位での編集を行うことができる。作成されたレコメンドリストはオペレータが事前確認した後に配信許可をするという手動配信や、事前確認不要で自動配信するといった制御も可能となっている。

また、本システムでは特徴辞書の更新や前述のレコメンドアルゴリズム、レコメンドエンジンの内部処理で利用するパラメータの設定変更をサービス停止させずに実現することが可能である。以上のように「設定」されたパラメータは即時にレコメンドエンジンに反映され、次回以降のレコメンド生成処理に利用される。

(2)効果測定

NOTTV向けレコメンドシステムでは、レコメンドの効果をインプレッション^{*7}、クリック^{*8}、コンバージョン^{*9}の指標値を利用して定量的に評価する。上記指標値は、端末から収集した操作履歴をもとに算出し、月/日/時間ごとの単位で任意の期間を指定してグラフや表としてオペレータ画面に表示することができる。そのため、レコメンドがどの程度ユーザの視聴機会を創出につながっているかを定量的、視覚的につかむことができる。また、ユーザを複数のグループに分けて異なるレコメンド設定をすることができ、どちらのグループの設定がより効果的であったかを同時期に比較するABテストも可能である。効果測定はオンラインで実施することができ、日々のレコメンドサービス運用の中でレコメンド効果を確認し、最適な設定を「評価」、「改善」していくことが

できる。

3.3 分散処理技術 (Hadoop)

NOTTVでは操作履歴やメタデータなどの大量データ処理が必要となる。従来のようなDBを主体とした処理方法では、データ量の増加に伴い、DBアクセス負荷や処理時間が急激に増大してしまう。そこで本システムではHadoopTM^{*10}[6]を採用し、並列分散処理によるレコメンド計算の高速化を図っている。

本システムでは、特徴ベクトル生成や類似度計算などのタスクを、ユーザによって異なるリソースに割り当てることにより並列分散処理を実現している。

上記により高速なレコメンド処理が実現され、鮮度の高いレコメンド情報の担保が可能となる。

4. システム構成

レコメンドシステム構成を図4に示す。

内部構成は、端末やオペレータ端末からのアクセスや顧客認証などのオンライン処理を担うフロントサーバ、操作履歴やメタデータ^{*11}、レコメンドリストを管理するDBサーバ、レコメンド計算処理を行うバッチサーバおよびHadoop[6]により実現される。

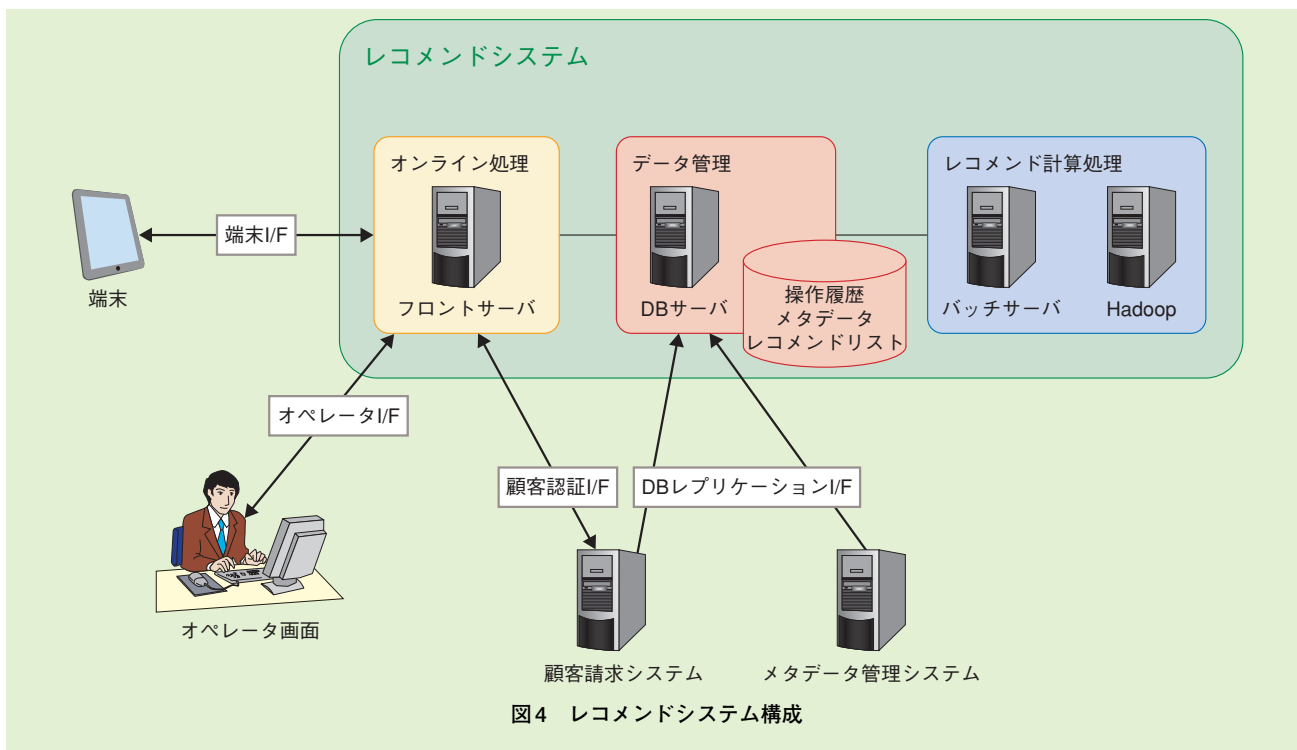
外部インターフェース（以下、I/F）としては、端末やオペレータ画面、顧客の属性や契約状態を管理する顧客請求システム、コンテンツ情報を管理するメタデータ管理システムと

*7 インプレッション：レコメンドされたコンテンツが表示された回数を示す指標値。

*8 クリック：レコメンドされたコンテンツのうち、ユーザがコンテンツの詳細情報を閲覧した回数を示す指標値。

*9 コンバージョン：レコメンドされたコンテンツのうち、ユーザが予約、あるいは視聴を実施した回数を示す指標値。

*10 HadoopTM：並列分散処理を実現するためのプラットフォーム。HadoopはApache Software Foundationの登録商標。



接続している。端末向けI/Fは、レコメンドリストや操作履歴の送受信機能のみならず、ユーザがキーワードを登録し、関連したコンテンツのリストを取得する機能や、システム内に蓄積されたユーザの嗜好情報をリセットする機能などを提供する。オペレータ向けI/Fとしては、前述の通り、オペレータ画面およびAPI (Application Program Interface)^{*12}を開発した。また、顧客情報システムやメタデータ管理システムに対して、DBレプリケーション^{*13}により顧客情報やメタデータをレコメンドシステムへ取り込む機能も具備している。

5. あとがき

本稿ではNOTTV向けレコメンド

システムについて解説した。本システムは、ユーザの嗜好に基づいたレコメンドアルゴリズム、柔軟なオペレーション機能、高速なレコメンド処理により、NOTTVの利用促進を実現する。

今後は実際のサービス運用からのフィードバックをもとに、レコメンドのさらなる改善をめざす。

文献

- [1] G. Linden, B. Smith and J. York : “Amazon.com Recommendations Item-to-Item Collaborative Filtering,” IEEE Internet Computing, Vol.7, pp.76-80, Jan. 2003.
- [2] B. Sarwar, G. Karypis, J. Konstan and J. Riedl : “Item-based Collaborative Filtering Recommendation Algorithms,” Proc. of WWW10, pp.285-295, May 2001.

- [3] Y. Fukazawa, M. Hara and H. Ueno : “Automatic Cell Phone Menu Customization Based on User Operation History,” Transactions of the Japanese Society for Artificial Intelligence, Vol.25, No.1, pp.68-77, Mar. 2010.
- [4] T. Joachims : “Optimizing Search Engines Using Click-through Data,” Proc. of KDD2002, pp.133-142, Jul. 2002.
- [5] 脊戸柳 昌宏, 深澤 祐介, 野口 勝広, 宮川 聡, 高橋 良之 : “ソーシャルゲームにおけるアイテム形状の類似性に着目したレコメンデーション,” IEICE SIG Notes WI-2011-41, pp.69-74, Sep. 2011.
- [6] Apache Hadoop Project : “Welcome to Apache™ Hadoop™ !” <http://hadoop.apache.org/>

*11 メタデータ：コンテンツのジャンルや紹介文などが記述されたデータ。
 *12 API：レコメンドシステムが提供する機能を、他システムや端末が利用するためのインタフェース。
 *13 DBレプリケーション：所定のDBから他

のDBに対して複製をつくる技術。