Technology Reports　AIエージェント　先読み　パーソナライズ
新AIエージェント特集

ユーザの行動を能動的にサポートする先読みエンジン

サービスイノベーション部　山田 直治（やまだ なおはる）　山田 将人（やまだ まさと）　勝丸 徳浩（かつまる のりひろ）　服部 響（はっとり きょう）
人事部　井上 正裕（いのうえ まさひろ）
R＆D戦略部　竹花 幸伸（たけはな ゆきのぶ）

ユーザの日々の生活を能動的にサポートするエージェントサービスを実現する上で，ユーザの行動や特性を正確に理解することは重要である．スマートフォン上には，位置情報や，予定情報，メール，画像など，多くのパーソナルデータが存在する．これらを分析することで，ユーザの行動を正確に把握することができる．そこでドコモでは，さまざまなパーソナルデータを統合的に分析する基盤である先読みエンジンを開発した．本稿では，本エンジンの概要とそれぞれの機能について解説する．

1. まえがき

近年，多くの企業がアシスタントサービスを提供している．例えば，Googleアシスタント™*1や，Apple社のSiri®*2，しゃべってコンシェル，Microsoft社のCortana*3などがある．

アシスタントサービスは現在，ユーザからの問いかけに答える受動型から，アシスタントからユーザに呼びかける能動型のサポートへと進化しつつある．従来の受動型のアシスタントサービスは，ユーザ側が明示的に問いかけない限り，適切なサポートができない．一方で能動型は，突然のゲリラ豪雨や電車遅延など，ユーザがまだ知らないトラブルをアシスタントが察知し，その情報をPush通知などでユーザに伝えるといった，より能動的なサポートが可能である．

このような能動型のアシスタントサービスを実現する上では，スマートフォン上のパーソナルデータを用いてユーザのプロファイルや次の行動を予測することが重要である．いまやほとんどの人が常にスマートフォンを携帯し，スケジュールの管理，家族や友人とのコミュニケーション，ネットショッピン

©2018 NTT DOCOMO, INC.
本誌掲載記事の無断転載を禁じます．

*1　Googleアシスタント™：米国Google LLCの商標または登録商標．
*2　Siri®：米国および他の国々で登録されたApple Inc.の商標．
*3　Cortana：米国Microsoft Corporationの米国およびその他の国々における登録商標または商標．

グ，写真や動画の撮影，ルート検索など，さまざまな用途で利用している中，その過程で取得・蓄積される位置情報や，予定，メールや画像のデータを活用すれば，ユーザの行動や特性を正確に理解し，それに基づいたサポートが可能になる．

ユーザの行動や特性を把握するにあたって，これまでにも，スマートフォン上の位置情報から，ユーザの自宅・職場の推定や，次の目的地の予測を行う研究は多く存在した．しかし，位置情報を用いた目的地予測手法は，過去の移動履歴に基づく予測であり，よく行く場所を予測することはできるものの，美容院のようにたまにしか行かない場所や，初めて行く場所を予測することは本質的に不可能である．

そこでドコモでは，位置情報に加えて，予定やメール，画像といった複数のパーソナルデータを利用することで，ユーザの行動を正確に推定する先読みエンジンを開発した．これにより，推定できるサポートのバリエーションを拡げることができる．例えば，位置情報を用いた目的地予測では，基本的には，自分が過去に行ったことのある場所しか予測できない．しかしながら，スケジュールデータやメールデータを解析することで，旅行など非日常的で，初めて行く場所に関しても特定することができる．

本稿では，先読みエンジンの概要とエンジンを構成するそれぞれの機能について解説する．

2. 先読みエンジン

先読みエンジンとは，ユーザのスマートフォン上のパーソナルデータを分析・活用し，ユーザ1人ひとりに合わせた情報やサービスを，最適なタイミングで提供するエンジンである．

パーソナルデータには，位置情報，予定情報，メール，画像などがある．先読みエンジンでは，これらを用いて，現在の行動推定，未来の行動先読み，プロファイル（静的特性）推定を行う．さらに，パーソナライズPushでは，これらの推定結果に基づいて最適なコンテンツを抽出し，プッシュ配信を行う（**図1**）．

以降では，位置解析機能，予定解析機能，メール解析機能，画像解析機能，パーソナライズPush機能について詳述する．

2.1 位置解析機能

ユーザが常に携帯するスマートフォンで定期的に位置情報を取得することにより，いつ，どこに滞在し，どこからどこへ移動したかを把握することが可能である．先読みエンジンでは，定期的に取得した位置情報から，現在の行動推定，未来の行動予測，プロファイル推定の3つを行う．

⑴現在の行動推定

定期的に取得した位置情報に基づいて，ユーザが滞在中か移動中かを判定する．

ユーザが滞在中の場合は，滞在位置からユーザの滞在施設を特定する．しかし，ゴルフ場など，施設の空間的範囲が大きい場合，単純に施設中心位置がユーザの滞在位置から最も近い施設を抽出する手法では，正しい施設を抽出できない．そこで，本システムでは，敷地や建物の境界（シェープ）との位置関係などを特徴量[*4]とする機械学習[*5]モデルにより，訪問施設の推定を行う．訪問施設の推定イメージを**図2**に示す．この例では，施設中心位置がユーザの滞在位置から最も近いものは施設Dとなるが，施設境界を考慮することで，施設Aに訪問したと推定することができる．

ユーザが移動中の場合は，取得した位置情報系列から，ユーザの移動経路を特定する．しかし，スマートフォンの消費電力を抑えるために位置情報の取得間隔を長くすると，移動中の位置情報がまばらになり，移動経路が不正確になる課題がある．そこで本システムでは，同じ移動経路を繰返し通過した際に取得した位置情報を利用することで，移動経路

＊4　特徴量：データから抽出される，そのデータを特徴づける値のこと．

＊5　機械学習：事例を基にした統計処理により，計算機に入力と出力の関係を学習させる枠組み．

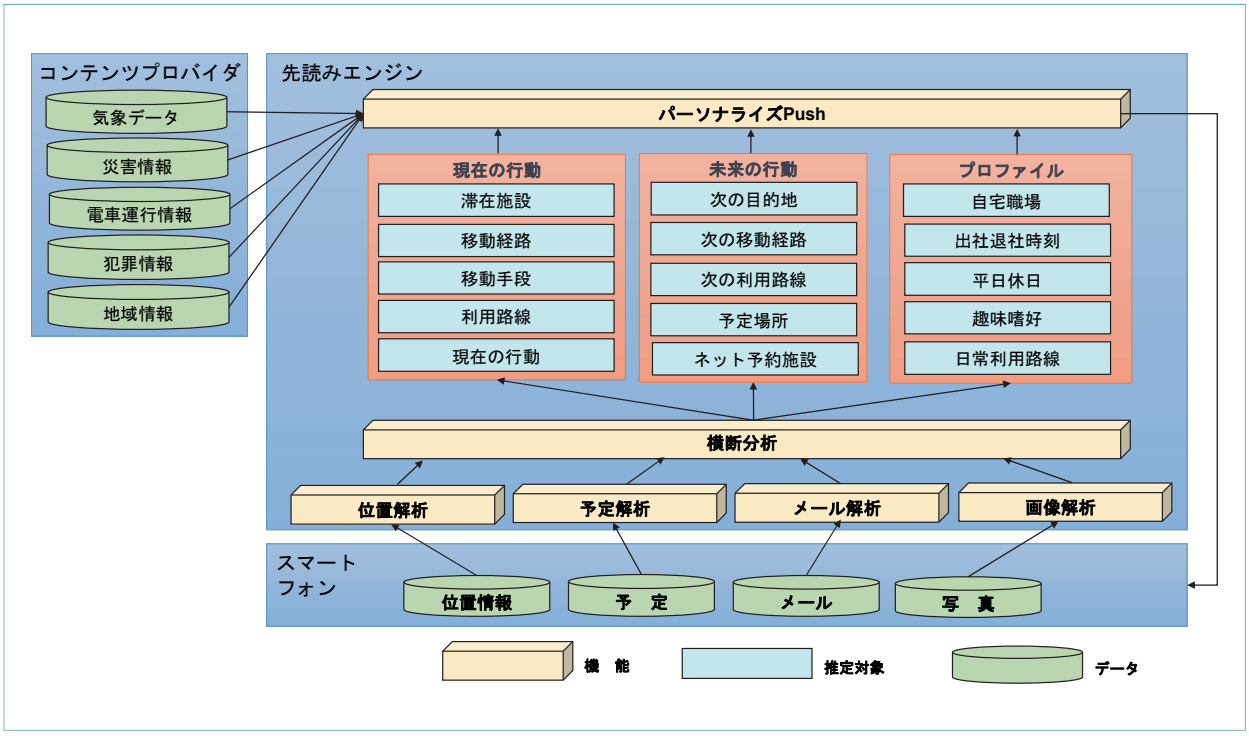

図1　先読みエンジンの機能概要

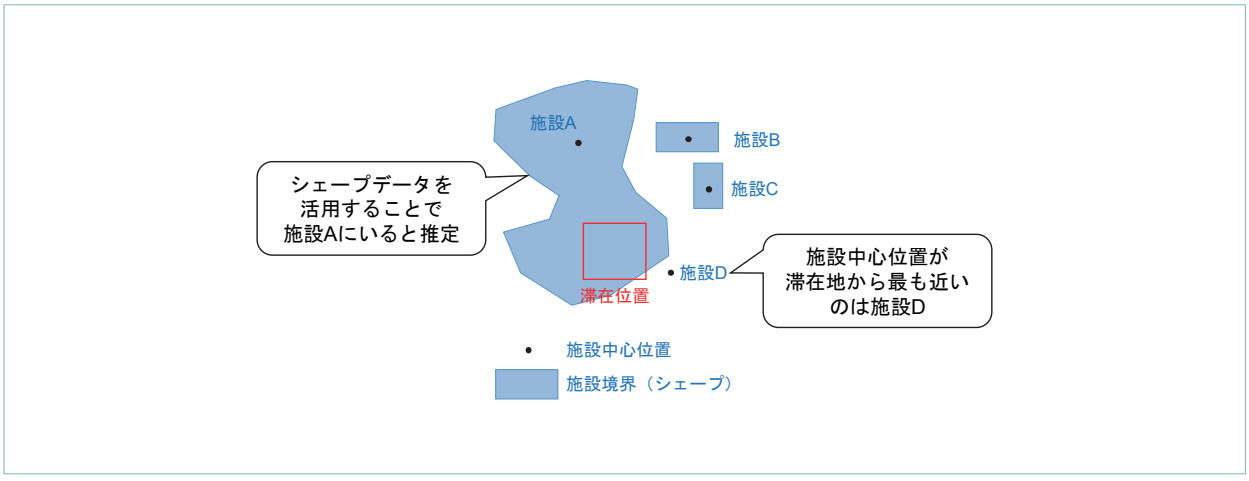

図2　訪問施設の推定イメージ

の精度を向上させる．具体的には，まず，移動中の位置情報から移動経路を推定する．次に，過去に推定した移動経路と比較することで，過去に同一の移動経路が存在するかを判定する．そして，存在する場合はその位置情報をユーザの移動経路に統合することで，精度を向上させている．そして，推定した移動経路と，電車の駅位置情報，移動速度などから，移動手段として，電車，車，徒歩の判定を行う．電

車で移動中の場合には，利用している路線名，乗車駅名，下車駅名，乗換駅名を特定する．

⑵未来の行動予測

先読みエンジンは，ユーザの過去の移動履歴として，曜日・時間帯別の移動経路の通過回数を保持している．そこで現在時刻と現在地に基づいて，過去の同一時間帯・同一場所における移動履歴情報から，次の訪問先と，そこへ至る移動経路，電車や車，徒歩などの移動手段，および，山手線などの利用路線を予測する．

⑶プロファイル推定

日々の位置情報を基に，滞在日数や滞在時間長などを算出し，それらを特徴量とする機械学習により，自宅と職場の位置を推定する．次に，日々の出社のタイミングに基づき，出社する曜日や曜日ごとの自宅を出る時刻，会社に着く時刻，会社を出る時刻，自宅に着く時刻の推定を行う．一方，ユーザの滞在履歴や移動経路履歴から，よく滞在する場所や，よく利用する路線名を特定する．さらに，ユーザが滞在した施設履歴に基づいて，例えば，美術館によく滞在していることから美術館めぐりが趣味といった，趣味嗜好も推定する．

2.2　予定解析機能

一般にユーザは，旅行や美容院予約，通院など，不定期なイベント情報を，備忘のためにスケジュールアプリに登録することが多い．これらのデータには，予定の場所を示す文字が含まれていることが多く，それを特定することでユーザの行動を予測することができる．

予定の場所を示す文字として，住所が記入されている場合と，「東京タワー」といったPOI（Point Of Interest）*6情報が記入されている場合がある．住所表記については，正規表現*7を用いて抽出するが，POI情報については，単純な形態素解析*8では，POI名を正確に取得することができず，例えば，「東京タワー」の場合，「東京」のみを地名として抽出してしまう．

そこで，POI名を正確に取得するために，先読みエンジンでは，系列ラベリング*9手法であるCRF（Conditional Random Fields）*10を用いて抽出している．具体的には，予定データのうち，地名を示す形態素にラベル付けを行う．ラベル付けのイメージを**表1**に示す．「東京タワーでランチ」という予定に対して，地名を示す「東京タワー」に地名ラベル「B-LOC」「I-LOC」を人手で付与する．このような人手によるラベル付けを大量の予定データに対して実施し，CRFで学習することで，地名を高精度に抽出している．これらの技術により，実データを用いた評価では，90％の精度でPOI名を正しく検出できている．

2.3　メール解析機能

最近では，ホテルや新幹線，美容院，レッスンなどの予約や，さまざまな商品の購入をインターネットのサイトで行う機会が増えてきている．この場合，ユーザには予約確認や購入確認のメールが送信される．これを解析することで，ユーザがいつ，どこへ行くのかを特定したり，ユーザが何を，いくらで購入したのかを特定することが可能となる．

メール解析は，ルールベースと機械学習を組み合わせて実施している．

⑴ルールベースによる情報抽出

予約／購入確認メールには，各社独自のフォーマットがあり，それを用いてユーザへ機械的にメールを送信している．そこで，本システムでは，送信元メールアドレスごとに，メールから予約情報や購入情報をあらかじめ定めたルールに基づいて抽出する．ルールは，各メールに記載された文字列のうち，どの文字列を抽出したいかを記載したものである．例えば，ある会社の予約確認メールでは，「予約番号：ABC123」というフォーマットだと仮定する．

＊6　POI：店舗や施設のこと．
＊7　正規表現：文字列の集合を1つの文字列で表現する方法の1つ．
＊8　形態素解析：文を構成する最小の意味単位である形態素の列に分割する作業．
＊9　系列ラベリング：事例の統計処理により得られた判断基準に基づいて，文字列などの系列の各要素に対し，適切なラベルを自動的に付与する手法．
＊10　CRF：条件付確率場．入力された要素が連なったもの（系列）に対して，その特徴量に基づいてあらかじめ定められたラベルを付与する手法の一種．

表1　地名のラベル付けイメージ

形態素	品　詞	ラベル
東京	固有名詞，地名	B-LOC
タワー	一般名詞	I-LOC
で	接続詞	その他
ランチ	一般名詞	その他

先読みエンジンでは，このメールに対して，「予約番号：」の後に出てくる文字列である「ABC123」を抽出するというルールを記述しておくことで，予約番号を正確に抽出することができる．

⑵機械学習による情報抽出

ルールベースによる情報抽出手法の課題として，ルールを人手で作成する必要があり，すべてのインターネットサイトに対応することは不可能である．そこで本システムでは，ルールベースに基づいて収集した予約メールを学習データとした機械学習のアプローチを適用する．

具体的には，まず，メールをカテゴリ分類する．カテゴリは，ネット通販の購入確認，映画の予約，航空券の予約，クーポンなど10種類程度存在する．メールに記載された文章を形態素解析により，単語集合を作成し，多クラスロジスティック回帰[＊11]によりカテゴリ分類を行う．

次に，各カテゴリについて，必要な項目（例えば，映画の予約の確認メールでは，タイトル，上映日時，映画館の名称，予約した座席番号など）を定義し，各項目に関する情報を，CRFを用いて抽出する．具体的には，形態素解析された文字列に対して，抽出したい項目名に「K」，抽出したい値に「V」，これらを分離する記号に「S」を付与する．一例として，「予約受付日：2014-12-17（火）」という文字列に対するCRFのラベル付けのイメージを**表2**に示す．この例では，「予約受付日」に「K」を，「：」に「S」を，「2014-12-17（火）」に「V」をラベル付けしている．このようなラベル付けを大量のメールデータに対して行い，CRFを用いて学習することで，メールからの情報抽出を可能にしている．これらの技術により，実データを用いた評価では，*F*値[＊12]で90％の精度で情報抽出できている．

表2　機械学習による情報抽出のラベル付けイメージ

形態素	品　詞	ラベル
・	記　号	O
予約	名　詞	K
受付	名　詞	K
日	名　詞	K
：	記　号	S
2014	名　詞	V
－	記　号	V
12	名　詞	V
－	記　号	V
17	名　詞	V
（	記　号	V
火	名　詞	V
）	記　号	V

2.4　画像認識機能

ユーザのもつ画像には，趣味・嗜好に関する手がかりが多く含まれる．例えば犬を飼っていれば犬の写真を多く撮影しており，海外旅行好きであれば海

＊11　多クラスロジスティック回帰：ロジスティック回帰は，ベルヌーイ分布に従う変数の統計的回帰モデルの一種である．ロジスティック回帰は，二値分類であるが，多クラスロジスティック回帰は，それを多値分類に拡張した手法である．

＊12　*F*値：正確性と網羅性の総合的な評価の際に利用される尺度であり，適合率と再現率の調和平均によって算出される．

外で撮影した写真を多くもっている可能性が高い．画像に写っているシーン・イベントを推定することがユーザの趣味・嗜好の推定にそのままつながる．

画像のシーン・イベント推定として，先読みエンジンでは，2つのアプローチを取っている．1つは，単一実画像を用いた推定であり，もう1つは，複数の画像を組み合わせた推定である．

(1)単一実画像を用いたシーン・イベント推定

単一実画像を用いたシーン・イベント推定として，深層学習の一手法である畳込みニューラルネットワーク＊13を用いている．本手法ではデータを学習する過程が必要であり，いかに多く質の良い学習データを用意するかが精度に大きく影響する．しかし，一般的に収集可能なWeb上の画像データはプロカメラマンが撮影した写真などが多く，一般ユーザが撮影する写真とは画像の傾向が異なることが多い．そこでドコモは，学習データとして，ユーザが撮影した画像を利用することで精度を上げている．

学習データにWeb画像のみとユーザデータを用いた場合の精度を**図3**に示す．グラフより，ユーザデータを学習データに用いた条件の方が適合率＊14と再現率＊15共に高くなっていることが分かる．

(2)複数実画像を用いたシーン・イベント推定

単一実画像では把握が難しいシーン・イベントも，複数枚の実画像を用いることで，推定できることがある．例えば，ディズニーランドで複数枚の画像を撮影したとする．食事風景を撮影した画像には，「食事」というシーンしか付与することができないが，ディズニーランドで撮影した複数枚の画像をまとめて解析することで，「ディズニーランド」というイベントも付与することができる．

複数実画像の分析を行うために，以下の2つの手順で行う．まず，撮影日時や撮影場所といった画像メタデータを用いて，画像のグルーピングを行う．そして，各画像グループに対して，シーン・イベント推定を行う．

①グルーピングでは，ユーザの一連の写真のうち，撮影場所と撮影日時が総合的に近いものを同一イベントの写真としてグループにまとめる．グループ化してから推定することで，1枚では何

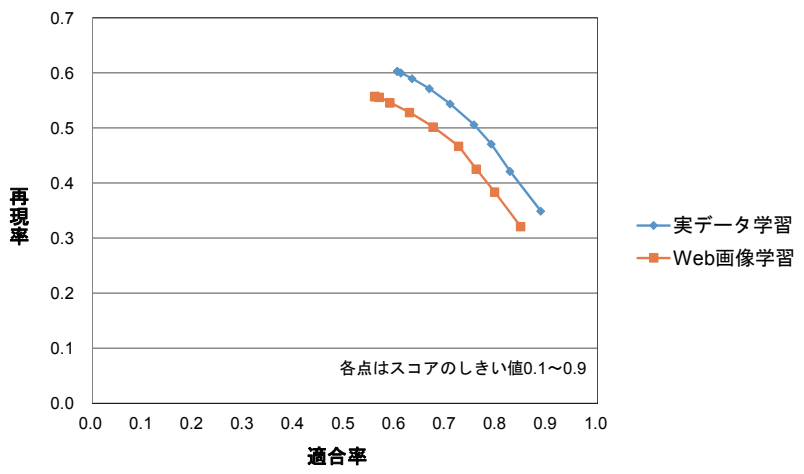

図3　Web画像のみとユーザデータ（実データ）を用いた場合の精度

＊13 畳込みニューラルネットワーク：深層学習の手法の一種である．何段もの深い層から構成されるニューラルネットワークにおいて，畳み込み層やプーリング層などの特徴的な機能を持った層を入れることにより，特に画像認識の分野で優れた性能を発揮している．

＊14 適合率：推定結果の正確さを表現することができる指標のこと．しかし，推定結果の網羅性を表現することができない．

＊15 再現率：推定結果の漏れの少なさを表現することができ，網羅性を意味する．しかし，推定結果の正確さを表現することができない．

の写真か不明でも同一グループの他の写真を手がかりにイベントの推定を可能にする．

②イベント推定では，グルーピングされた複数枚の写真に対して，それぞれのもつ画像認識結果，撮影位置・日時情報を用いて，グループ全体として何のイベントかを推定する．例えば，動物の写真が多く撮影位置が動物園周辺の場合，動物園に行ったと推定する．

2.5 パーソナライズPush機能

パーソナライズPush機能は，これまでに述べた推定結果と，天気・運行・イベントなどのコンテンツ情報を組み合わせ，それぞれのユーザにとって最適なコンテンツを，最適なタイミングでPush配信する機能である．

パーソナライズPushの実現に向けては，まず，Push配信の条件を記載したECAルール[*16]を作成する．ECAは，Event，Condition，Actionの頭文字であり，EventはPush配信を「いつ行うのか」，Conditionは「誰に行うのか？」，Actionは「何を配信するのか？」を示している．例えば，雨が降ることを教えるユースケースでは，Event：降水確率が50％を超えたら，Condition：降水確率が50％を超えたエリアにいる人，Action：「雨が降りそうですよ」となる．これにより，新たなPush配信を，ECAルールの追加のみで対応することが可能である．

次に，ECAルールのEventが条件に合致しているかを定期的に判定する．そして，条件に合致していれば，Conditionに該当するユーザを抽出し，Actionに記載されたメッセージをPush配信する．

3. あとがき

本稿では，パーソナルデータを解析する基盤である，先読みエンジンについて解説した．ユーザを能動的にサポートするアシスタントサービスを実現する上で，パーソナルデータからユーザの行動やプロファイルを把握することは必要不可欠である．

先読みエンジンでは，スマートフォン上のデータとして，位置情報や，予定，メール，画像データを解析することで，ユーザの現在の行動や，未来の行動，プロファイルの推定を可能にしている．

スマートフォン上には，アプリケーション利用ログや，端末センサデータなど，上記以外にも，まだ多くのパーソナルデータが存在する．解析対象のパーソナルデータを拡張することで，ユーザの行動・プロファイルの推定精度を向上するとともに，ライフイベントなどの新たなプロファイルの推定にも取り組んでいく予定である．

＊16 ECAルール：イベント，条件，アクションの組合せによって構成される処理方法の規定．