

コンテキストウェアネスとその応用

松平 正樹 星川 恭子
瀧 浩平

OKIが目指す「差別のない『個』別化」を特長としたユビキタスサービスを実現するために、「個」の多様な行動ログを利用して、「個」の状況に応じた（コンテキストウェアな）サービスを構築するためのUS-PALA^{TM*}1) エンジンを開発している。US-PALAはUbiquitous Service by Personal Activity Log Analysisの略であり、行動ログを分析し、行動予測、コンテキスト推定技術によってコンテキストウェアなサービスを実現するものである。図1にUS-PALAエンジンのコンセプトを示す。

US-PALAエンジンを使えば、たとえば、旅行サイトにアクセスした際に、過去に購入した商品、閲覧したWebページなどから興味のある観光スポットを抽出し、ツアーを個別化して提案するサービスや、従業員の位置情報、PC操作履歴、スケジュールをもとに、PCに対する操作（行動）に対して必要となるであろう情報を先読みして収集・提供するようなサービスを構築することが可能になる。

本稿では、US-PALAエンジンの概要と技術、および、US-PALAエンジンを活用したサービスプロトタイプについて紹介する。

行動ログとコンテキスト

まず、行動ログとコンテキストについて、整理してみたい。

行動ログとは、過去から現時点までの個人のアクティビティの履歴であり、移動履歴、購買履歴、メール履歴、PC操作、TV視聴、家電利用履歴、あるいは、心拍・血圧などの生体情報、などを含む。行動ログのデータは、連続値・離散値、値の範囲や粒度、データ更新頻度など多種多様な形式が混在する。また、取得方法も、センサ機器、PCや携帯電話などの情報機器や家電、ホストコンピュータやWebサービスなど多岐にわたる。これら多種多様なデータを、いかに統合して処理できるかが重要な課題である。

一方、コンテキストは、Georgia工科大学のDeyらが次のように定義している¹⁾。

Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and

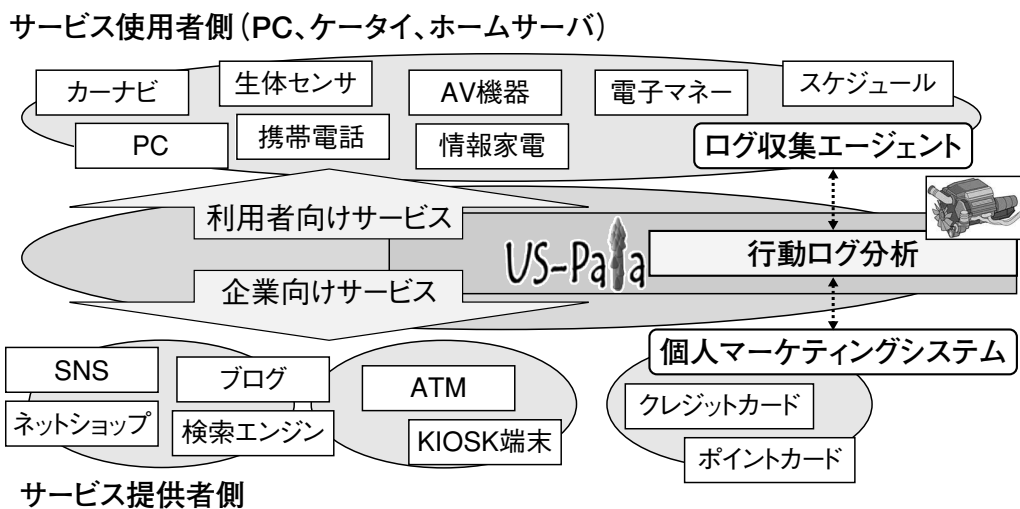


図1 US-PALAエンジンのコンセプト

*1) US-PALAは沖電気工業株式会社の商標です。

an application, including the user and applications themselves.

コンテキストとは、実体（エンティティ）の状況を特長付けるのに用いられる情報である。実体は、人や場所、モノであり、ユーザーとアプリケーションの相互作用、あるいはユーザーとアプリケーションそのものに関連していると考えられる。

我々は、Deyらの定義を踏まえ、コンテキストとは、行動ログデータから得られる個人や環境の状況であり、ログデータをもとに意味づけしたものであると考える。たとえば、「静止している」「歩いている」「電車に乗っている」といった行動状態、業務タスク・ステップや「個」の嗜好などがこれに相当する。

さらにDeyらは、コンテキストウェアなシステムを次のように定義している。

A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task.

システムが、ユーザーに関連した情報やサービスを提供するためにコンテキストを利用し、その関連がユーザーのタスクに依存している場合に、コンテキストウェアである。

我々はこの定義に基づき、「個」別化サービスを単なるパーソナライズではなく、「個」のコンテキストを利用して必要なときに必要な情報や機能を提供するコンテキストウェアなサービスと位置づけている。

US-PALAエンジン概要

US-PALAエンジンは、ペイジアンネットワーク技術と自然言語処理技術により、行動ログから行動モデルを構築する機能と、現在の条件から行動予測・コンテキスト推定を行う機能を持っている。図2に、US-PALAエンジンの構成を示す。

行動モデル構築機能は、携帯機器、PCなどに組み込まれたエージェント、空間に配置されたセンサおよびセンサネットワーク、環境データなどを配信するWebサービスなどを利用して収集された「個」の行動ログを読み込み、確率的な行動モデルを構築する機能である。たとえば、行動ログとしての退社時刻と位置情報（あるいは書

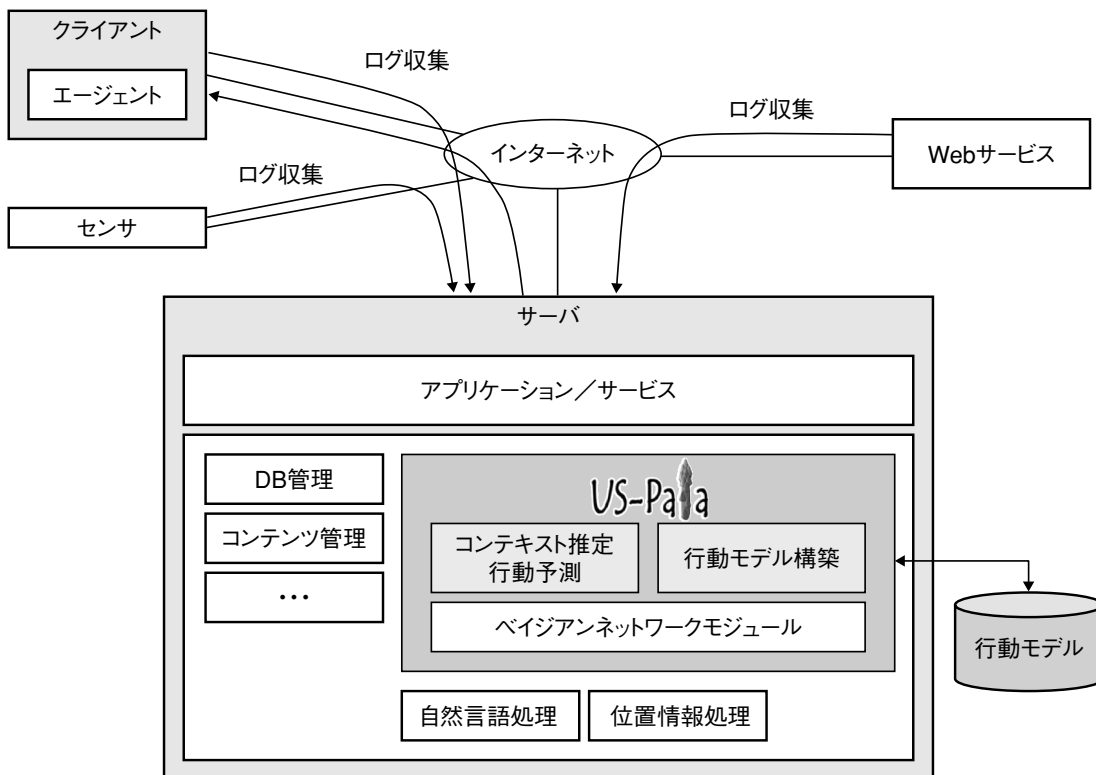


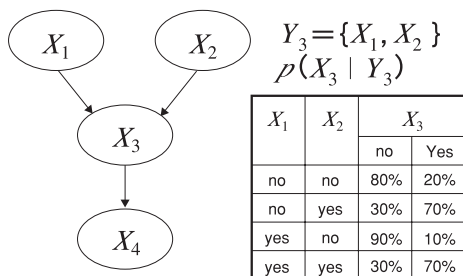
図2 US-PALAエンジンの全体構成

店での購買履歴)との相関から、確率的な因果関係を学習する。これにより、「会社の退出時刻が通常より早い場合は、書店に寄る確率が高い」といった行動モデルが構築される。

行動予測・コンテキスト推定機能は、「個」の行動モデルをもとに、「個」の位置や時刻、直前の行動など観測できる条件を入力することで、行動予測・コンテキスト推定を行う機能である。この予測・推定に従って、行動支援サービスや通知サービス、推薦サービスを構築することが可能である。以下で、ベイジアンネットワーク、自然言語処理技術について解説する。

ベイジアンネットワーク

ベイジアンネットワークは複数の変数間の依存関係を検出することで、変数間の同時確率分布を効率よく表現するための技術である^{2) 3)}。ベイジアンネットワークはたとえば図3に示したように、変数間の依存関係を表した有向グラフと、変数ごとの条件付き確率分布表からなる。有向グラフ中の1つのノードが1つの変数に対応する。有向辺はそれが結ぶノードに対応する変数間に依存関係があることを表す。ノード X_1 からノード X_3 に向かう有向辺は、対応する変数 X_3 が変数 X_1 に依存することを表す。それぞれの変数は、その変数が依存する変数を条件とする条件付き確率表を持つ。



$$p(X_1, X_2, X_3, X_4) = p(X_1)p(X_2) \cdot p(X_3 | X_1, X_2) \cdot p(X_4 | X_3)$$

図3 ベイジアンネットワークの例

ベイジアンネットワークでは有向グラフで表される以外の依存関係がまったくないことを仮定して、同時確率を以下のように近似する。

$$p(X_1, \dots, X_n) = \prod_i p(X_i | Y_i)$$

ただし Y_i は X_i が依存する変数の集合を表す。

US-PALAエンジンでは、このベイジアンネットワーク

を用いて「個」の行動モデルを表現し、行動予測、コンテキスト推定をおこなっている。すなわち、「個」の多様な行動ログから、ベイジアンネットワークの変数および条件付き確率を学習し、たとえば、「個」の位置や時刻、直前の行動など観測できる条件を入力することで、行動予測・コンテキスト推定を行う。

ベイジアンネットワークは確率という枠組みによって、人の不確実な行動を柔軟に表現することが可能である。

自然言語処理・知識処理

US-PALAエンジンのもうひとつの技術的な特長は、OKIが長年培ってきた自然言語処理および知識処理技術を統合している点である。

行動ログの中で、特に「個」の嗜好はテキスト中に表現されていることが多い。たとえば、ジャズコンサートのチケットを購入した人は、ジャズピアノが大好きかもしれないし、60年代のジャズファンかもしれない。そのような嗜好は、個人のブログや商品に関する紹介文のテキストに表出していることが多く、そこからキーワードを抽出することで、より「個」の嗜好を表現できる。

テキストからキーワードを抽出する方式として、TF/IDF (Term Frequency / Inverse Document Frequency) 法が知られている。ある文書中で出現頻度の多い単語はその文書の特徴づけ、また同時に、他の多くの文書でも出現する単語は特徴的ではないという考えに基づき、単語 t の文書 d における重要度 $w(t, d)$ は、次式のように定義される。

$$w(t, d) = tf(t, d) \cdot idf(t)$$

ここで、 $tf(t, d)$ は文書 d における単語 t の出現頻度であり、 $idf(t)$ は文書数 N 、および単語 t を含む文書数 $df(t)$ によって、次式で与えられる。

$$idf(t) = \log \frac{N}{df(t)}$$

TF/IDF法は、論文や新聞記事などある程度長い文書では有効であるが、商品説明文のような短い文章では単語の出現頻度TFが1になる場合が多く、改良が必要である。我々は、単語間の関連を記述したシソーラス辞書やオントロジー辞書⁴⁾を構築し、キーワードを拡張することを試みている。

また、キーワードをベイジアンネットワークで処理するための工夫も必要である。単純にキーワードを変数としてベイジアンネットワークに与えると、キーワードの数だけ

変数を処理しなければならず、処理時間を膨大に費やしてしまう。それを避けるため、シソーラス辞書、オントロジー辞書の利用や、統計処理で相互情報量を考慮することで、処理の削減を行っている。

サービスプロトタイプ

US-PALAエンジンの適用領域として下記のような「個」別化およびコンテキストウェアなサービスが考えられる。

- (金融・購買分野) 探していたモノを必要な時に通知したり、状況に応じて店舗や商品を推薦する
- (交通・旅客分野) 目的地や経路、旅行スケジュール、状況に応じてドライバや旅行者を支援する
- (ヘルスケア分野) 患者の状況を見守り、必要に応じて医師に通知する

金融・購買分野での「個」別化サービスとしては、書籍のレコメンドが有名であろう。購入した書籍に応じて、関連する書籍を推薦してくれるサービスである。ただし、推薦の方式上、サイト内の購買履歴のみから「個」別化を行っているため、その他の行動が反映されるわけではない。

交通・旅客分野についても、「個」別化サービスが非常に強く求められている。近年、個人観光旅行が増加してきており、2005年の国内旅行の市場構造では、個人観光旅行が77.8%を占めているという調査結果もある⁵⁾。インターネットの普及により、24時間、検索や予約が可能な環境が普及していることや、消費行動の多様な価値判断が当たり前になってきた社会要因もあるだろう。現状のサービスが、そのような利用者のニーズを満足している

とは言い難く、「個」別化サービスに対するニーズが強くなってきている。

我々は、このような市場環境を考慮し、個人の嗜好や状況に応じて、店舗や観光地を推薦するサービスを開発している。図4に表示例を示す。

今後の取り組み

今後、エンジンをOKIや他社のシステムと連携させて「個」別化サービスを構築していく予定である。たとえば、「個」に対して推薦された観光スポットおよびレストラン情報の中から、行きたいところをユーザーが選択し、その情報をトラベルアシストサービスと共有することで、「個」別化された旅のスケジュールリングを容易に行うことが可能となる^{6) 7)}。

参考文献

- 1) Dey, A. K. and Abowd, G. D. "Toward a Better Understanding of Context and Context-Awareness", In Proceedings of the CHI2000 Workshop on The What, Who, Where, and How of Context-Awareness, 2000.
- 2) 本村：ベイジアンネットソフトウェア、人工知能学会誌 Vol.17 No.5, 2002年
- 3) 本村, 岩崎：ベイジアンネットワーク技術 ユーザ・顧客のモデル化と不確実性推論, 東京電機大学出版局, 2006年
- 4) 松平, 上田, 大沼, 森田：Webコンテンツの分析に基づくオントロジー構築および情報整理の試み, 人工知能学会 第4回セマンティックウェブとオントロジー研究会資料, 2003年
- 5) 東レ経営研究所：旅行業界の現状と課題 一足元の需要は堅調, 主役は若年層から中高年層へ, TBR調査レポート, 2006年
- 6) 河原：ユビキタスサービスを実現するプラットフォーム, OKIテクニカルレビュー210号, Vol.74 No.2, pp.4-9, 2007年
- 7) 荒牧, 顔, 中川, 榎本：リアル/バーチャル融合「トラベルアシストサービス」, OKIテクニカルレビュー210号, Vol.74 No.2, pp.70-75, 2007年

筆者紹介

松平正樹：Masaki Matsudaira. ユビキタスサービスプラットフォームカンパニー ユビキタス技術二部 開発第四チーム
 星川恭子：Kyoko Hoshikawa. ユビキタスサービスプラットフォームカンパニー ユビキタス技術二部 開発第四チーム
 瀧浩平：Kohei Taki. ユビキタスサービスプラットフォームカンパニー ユビキタス技術二部 開発第四チーム



図4 サービスプロトタイプの表示例