

# 表現の他者文脈への伸延による流通促進の試み

Recommender system based on extending expressions along context of other expressions

田中 克明\*<sup>1</sup>      堀 浩一\*<sup>1</sup>      山本 真人\*<sup>2</sup>  
Katsuaki Tanaka      Koichi Horii      Masato Yamamoto

\*<sup>1</sup> 東京大学先端科学技術研究センター

Research Center for Advanced Science and Technology, The University of Tokyo

\*<sup>2</sup> 株式会社 エヌ・ティ・ティ・ドコモ

NTT DoCoMo, Inc.

The authors have developed a recommender system that captures personal context from history of personal online and offline activity, treats information on web sites as a large set of context, and finds out and extend overlap of them, then recommends information located there. The aim of the system is that a person can enjoy waves of information again. The system worked as a part of "My-life Assist Service." It is a service for mobile phones provided by NTT DoCoMo, Inc. as a field experiment from Dec. 2007 to Feb. 2008.

## 1. はじめに

10年とすこしほど前, WWWが普及し始めの頃, 新しい表現手段と伝達手段を手に入れたことに気づいた人々が, 自身のWebサイトを開設し, 同時に興味を持った他の人のWebサイトにリンクを張り, 「Web」の形を作っていた. さらに多くの人は, それらのサイトからサイトへとネットをサーフィンするように見て回った.

しかしながら, WWWがメディアの主流に近づくにつれその構成も変化し, あまりにも多くの表現がその中にあふれかえるようになり, 個人が直接そこに, 自由気ままに接することは, ほぼ不可能となった. 莫大な量の表現蓄積場所となったネットにおいて, 独りの個人ができることは, 強い意図を忘れずに検索エンジンを頼りに何かを探ること, あるいは他者の目的の下に作り込まれたサイトの中に安住すること, 辺りが限界である.

この限界は, ひとつのWebサイト内部, たとえば多くの商品を扱うECサイトでも発生している. これを解決するために, 同じ商品を購入した他者が他にどのような商品を購入しているかを表示する協調リコメンデーション手法により, 利用者が商品を見て回る手助けをする仕組みなどが用いられている. このように, 計算機による情報推薦は, 膨大な情報と受け手となる個人のバランスをとる方法のひとつである [Resnick 97].

計算機による情報推薦を, いっきに日常生活において現実のものとする試みとして, 携帯電話を用いて個人の日常生活における行動を入力とし, それを基に情報の推薦を行うシステムを作成し, 経済産業省情報大航海プロジェクトのマイ・ライフ・アシストサービスにおいて個人へ推薦する情報の選択を行う「歴史型推論エンジン」として運用した. マイ・ライフ・アシストサービスのシステム概要を図1に, 歴史型推論エンジンの構成を図2に示す.

マイ・ライフ・アシストサービスは, サービスプラットフォーム [河又 08] を中心として, 携帯電話電話上で動作する待ち受けアプリケーションを通して利用者情報を収集し, これに基づ

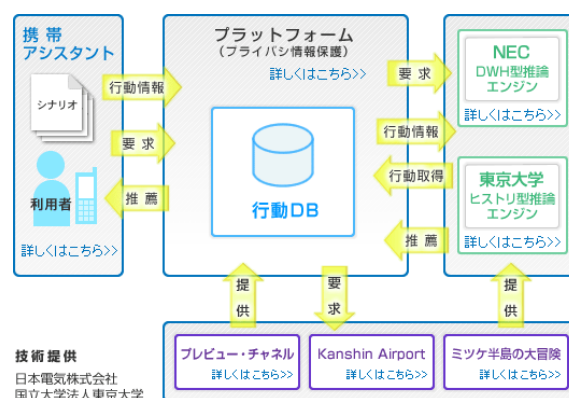


図 1: マイ・ライフ・アシストサービスシステム概要 (http://pr.mylifeassist.jp/systeminfo.html より)

いて利用者へコンテンツを提供するサービスである. 歴史型推論エンジンでは, 利用者情報として, 利用者の実世界における位置情報, 提供したコンテンツの閲覧履歴, サービス内の投稿サイトへ投稿されたテキストを用いた.

提供対象コンテンツとして, 日常生活に沿った情報を提供することを鑑みて, プレビュー・チャンネルでは (株) NTT レゾナントの教えて! goo\*<sup>1</sup> の約 200 万組の Q&A と (株) 角川クロスメディアの Walkerplus\*<sup>2</sup> の約 1 万件のグルメ店舗情報などを, Kanshin Airport では (株) 関心空間\*<sup>3</sup> の約 4 万件のユーザ投稿情報を利用した. プレビュー・チャンネルの画面例を図3に示す.

## 2. 表現の他者文脈への伸延による推薦

歴史型推論エンジンでは, サービス利用者の実世界における位置, Web の閲覧および投稿を利用者自身を示す表現としてとらえ, 提供対象コンテンツに含まれる表現と類似す

連絡先: 田中克明, 東京大学先端科学技術研究センター 知能工学研究室, 〒153-8904 東京都目黒区駒場 4-6-1, (03)5452-5289, jsai2008@katsuaki-tanaka.net

\*<sup>1</sup> <http://oshiete.goo.ne.jp/>

\*<sup>2</sup> <http://www.walkerplus.com/>

\*<sup>3</sup> <http://www.kanshin.com/>

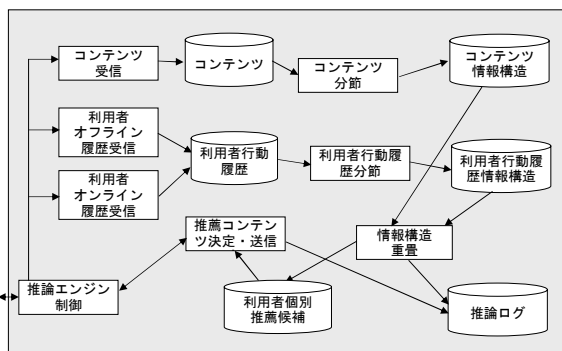


図 2: ヒストリ型推論エンジン構成

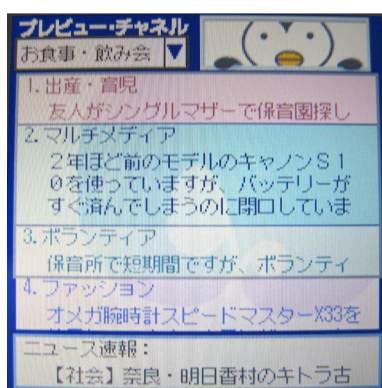


図 3: プレビュー・チャンネル画面例

るものを探し、さらにコンテンツ表現が含む文脈に基づいて、類似はせずとも関係するコンテンツを含めて推薦対象とする。そのために、サービス利用者の情報収集、情報の構造化、および構造化した情報の重ね合わせを行う。

たとえば、ひとつのグルメ店舗情報は、店舗の理念・メニュー・営業時間・場所・特長など、文書内にいくつかの内容を含む。これらの文書の集合をばらばらにし再構成することにより、店舗の理念の集合、メニューの集合、などを形作ることができれば、あるメニューへの興味に対し「店舗の理念」を通して別のメニューを推薦するといったことができる。同様の処理は「文書」として処理できるその他の表現についても可能である。このように、ヒストリ型推論エンジンは、ある表現を出発点とし、大規模な表現集合に含まれる文脈を用いて別の表現の選択を行うことにより、個人の表現を拡大し新たな表現との接点を探索する。

### 2.1 利用者情報の収集

サービス利用者情報として、利用者の同意と制御の下、携帯電話の待ち受けアプリケーションにて、GPS などを用いて取得した位置情報、Web アクセス機能による閲覧記録、および個人が作成したコンテンツを収集した。これらはサービスプラットフォームへ送信され、プラットフォームからヒストリ型推論エンジンが情報を取得し行動履歴を解析して推薦するコンテンツを選択し、プラットフォームを経由し利用者の携帯電話へと送る。

なお、位置をその他の情報と同様に処理するためには、文書として表現する必要がある。そこで、ある位置に対してその周囲にある提供対象コンテンツを、その場所の特徴を記述した情

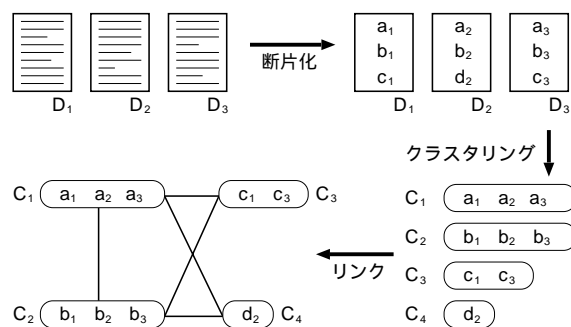


図 4: 情報の構造化

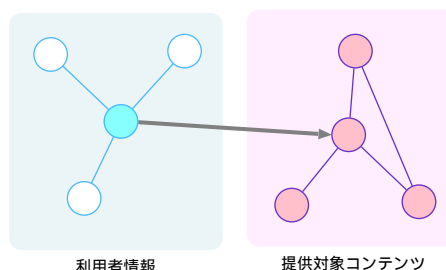


図 5: 情報の重ね合わせ

報であると見なし、位置をその周囲の提供対象コンテンツからなる文書集合へと変換した。

### 2.2 情報の構造化

情報の構造化は、情報をより小さな断片に分ける断片化、断片から情報の固まりを形成するクラスタリング、クラスタ間のリンクの設定という3つの段階により行う(図4) [Tanaka 04]。断片化は一定長の断片に文書を変換することにより行い、これを形態素解析して単語ベクトルとして表現、階層的ベイズクラスタリング [岩山 98] によりクラスタリング、クラスタに含まれる断片が同一文書であった割合に基づきリンクの設定を行う。

### 2.3 情報の重ね合わせ

構造化した利用者情報と提供対象コンテンツから、相互に関連するノード(クラスタ)を探し、提供対象コンテンツ側のノードからリンクをたどって推薦対象の情報を選択する(図5)。相互の構造から関連するノードを探索する際、利用者の好みに近い情報を推薦し続けるのでは、早期に飽きられてしまう可能性が高いこと [McNee 06]、相互に近いノードを全探索すると計算量が多くなることから、利用者情報の生成時刻を用いて、比較的最近生成された情報からなるノード、生成時刻分布の標準偏差が小さいノードなどを優先的に選択した。

## 3. 運用

上記のシステムの運用を2007年12月から2008年2月にかけて行った。最大約500人の利用者がおり、30分に1回の位置情報の自動取得、約2万件の閲覧記録、約200件の情報投稿を受けた。運用期間中のアンケートでは、利用者の約25%から、推薦情報が興味関心を広げる、興味関心に合っている、という評価を得た。

## 4. おわりに

本稿では、個人の情報表現を提供対象の情報表現に含まれる文脈に基づいて拡大することにより情報推薦を行う仕組みについて述べた。これにより、情報を新たな読み手へと届けることができた。

今回実現したシステムでは、利用者の表現を投稿する部分がサービスメニュー中の奥の方へ位置づけられたため、表現の流通を相互に促進するには至らなかったが、情報を受けた利用者が実世界において行動を起こすきっかけを作ることができたものと考えている。

## 参考文献

- [McNee 06] McNee, S. M., Riedl, J., and Konstan, J. A.: Being accurate is not enough: how accuracy metrics have hurt recommender systems, in *Proc. SIGCHI '06 Conference on Human Factors in Computing Systems*, pp. 1097–1101 (2006)
- [Resnick 97] Resnick, P. and Varian, H. R.: Recommender systems, *Communications of the ACM*, Vol. 40, No. 3, pp. 56–58 (1997)
- [Tanaka 04] Tanaka, K. and Takasu, A.: Topic Change Extraction from Problem Solving Records, in *Proc. of the 8th World Multi-Conference on Systemics, Cybernetics and Informatics* (2004)
- [河又 08] 河又 恒久, 村上 千央, 永井 洋一, 今野 清孝, 松川 淑子, 木内 直人, 山田 洋志, 亀井 真一郎: 行動情報を利用した携帯端末への情報配信システムのアーキテクチャ, 情報処理学会第 70 回全国大会 (2008)
- [岩山 98] 岩山 真, 徳永 健伸: 確率的クラスタリングを用いた文書連想検索, *自然言語処理*, Vol. 5, No. 1, pp. 101–118 (1998)