

アイテム混合方式に基づく偶発性を引き起こす 情報推薦システムの提案

岡田 裕規[†] 奥 健太[†] 服部 文夫[†]

[†]立命館大学情報理工学部 〒525-8577 滋賀県草津市野路東1丁目1-1

E-mail: [†]{cc002078@ed, oku@fc, fhattori@is}.ritsumeai.ac.jp

あらまし 意外性や発見性を重視した推薦システムが重要視されてきている。しかし、利用者が予期できないほどの意外なアイテムが推薦されても受け入れられにくい。本研究では、ある程度予測可能で、ある程度予測不可能な性質を偶発性とよび、その偶発性を引き起こす推薦システムについて検討する。具体的には任意の二つのアイテムを混ぜ合わせることで、利用者にとってある程度は予測可能であるが、新しく価値のあるアイテムを導き出す方式を提案する。

キーワード 情報推薦, 偶発性, アイテム混合

Recommender System Leading to Contingency based on Item-Fusion Approach

Yuki OKADA[†] Kenta OKU[†] and Fumio HATTORI[†]

[†] College of Information Science and Engineering, Ritsumeikan University

1-1-1 Nojihigashi, Kusatsu-city, Shiga, 525-8577 Japan

E-mail: [†]{cc002078@ed, oku@fc, fhattori@is}.ritsumeai.ac.jp

1. はじめに

情報推薦技術は、1990年代後半から注目を浴び始め、利用者の嗜好を正確に抽出する嗜好抽出技術などに関する研究が主流に行われてきた[1][2]。しかし、昨今、世の中には多種多様なモノが溢れるようになり、モノに対する人々の価値表現も多様化してきた。その中で、情報推薦システムの性能として、正確性だけでなく、意外性や発見性、セレンディピティなどの観点から評価することも重要視されるようになってきた[3]。

利用者の嗜好に対し過度に正確な推薦がされた場合、利用者にとっては新鮮味がなく推薦システムに対し飽きを感じるであろう。例えば、『ハリーポッター』の第1巻から第6巻を購入した利用者に対し、『ハリーポッター』第7巻を推薦されたとしても利用者にとっては容易に予測可能であり新鮮味がない。一方で、意外性や発見性を重要視するとはいえ、利用者にとって予測不可能なアイテムが何の脈絡もなく推薦されたとしても、利用者は戸惑いを感じる。利用者にとって、ある程度予測可能な範囲内であるが、利用者自身では予測が困難であるようなアイテムを推薦することが求められる。

本研究では、ある程度予測可能で、ある程度予測不

可能な性質のことを偶発性とよび、その偶発性を引き起こす推薦システムについて検討する。利用者にとってある程度予測可能な状態であるためには、推薦システムから自動的に推薦アイテムが決定されるよりも、利用者が積極的にシステムに関与していくことが必要であると考え、利用者とは推薦システムとがインタラクションすることを前提とする。インタラクションする方法としては様々なものが挙げられるが、本研究では偶発性を引き起こすための方法として、任意の二つのアイテムを「混ぜ合わせる」という行為に着目する。

「混ぜ合わせる」という行為は、色を混ぜ合わせる、料理の食材を混ぜ合わせる、音を混ぜ合わせるなど人間にとっても身近で直感的な行為である。また、混ぜ合わせた後の結果が予測しやすいという特徴と、ときには思いがけない結果を引き起こすという特徴を併せ持つ。身近な行為の一つに、例として絵具を混ぜ合わせる事が挙げられる。既に用意されている絵具の中に、使いたい色がない場合は、既存の色を混ぜ合わせて新しい色を作り出すことができる。また、例えば青色に白色を混ぜ合わせると、薄い青色、すなわち水色になることは予測しやすい。さらには、「混ぜ合わせる」という行為は、人間の興味を誘発するものでもある。

例えば、「ある色とまた別の色とを混ぜ合わせるとどのような色になるのか」といったように興味がそられる。そういった行為を繰り返していくうちに思いがけない結果に巡り合うことも期待できる。

本研究では、このような「混ぜ合わせる」という行為に着目し、二つのアイテムを「混ぜ合わせる」ことで人間にとって新しく価値のあるアイテムを見つけ出すことを可能にする情報推薦システムを開発することを目標とする。技術的課題としては、二つのアイテムの表現方法、およびそれらを混ぜ合わせることで新たなアイテムを導き出す方式を確立することが挙げられる。

本原稿の構成は以下のとおりである。第 2 章では、情報推薦分野の関連研究について述べる。第 3 章では、二つのアイテムの表現方式およびそれらの混合方式について定式化する。第 4 章では、実験、評価方法について記す。第 5 章では、実験結果についての考察を行う。第 6 章では本研究をまとめ、今後の課題を示す。

2. 関連研究

2.1. 情報推薦の基本方式

情報推薦の基本方式として、(a) コンテンツに基づくフィルタリングと、(b) 協調フィルタリングの 2 種類がある[1]。

(a) コンテンツに基づくフィルタリングでは、利用者のプロフィールとアイテムのプロファイルが一致もしくは類似するアイテムを推薦するものである。利用者のプロフィールは利用者による明示的な嗜好情報の入力や、SVM や決定木などの機械学習手法により利用者の過去のアイテム購入履歴に基づき作成されるものがある[4][5]。基本的には、利用者のプロフィールに忠実なアイテムが推薦されるため、意外性や発見性のある推薦はされにくい。

(b) 協調フィルタリングでは、嗜好が類似する他の利用者が好むアイテムを対象ユーザに推薦するものである。推薦システムを採用している有名なサイトとして Amazon.com[6][7]が挙げられる。他の利用者の嗜好情報を参照することで、ときには意外なアイテムが推薦されることも期待されるが、推薦精度が高すぎると類似アイテムばかりが推薦されるという問題が生じる。例えば、『ハリーポッター』の DVD の情報を閲覧しているときには、推薦リスト内のほとんどが『ハリーポッター』シリーズの DVD が提示され、利用者にとって自明な推薦となってしまう。

2.2. 発見性・意外性を考慮した情報推薦方式

清水ら[8]は、推薦リスト内の知らないアイテムの割合を測定するために発見性という指標を提案し、発見

性を考慮した情報推薦システムを提案している。利用者の既知・不既知プロフィールを用い、協調フィルタリングにより不既知のアイテムを予測することで、利用者にとって知らないアイテムを推薦することで、発見性のある推薦を可能としている。

村上ら[9]は、「意外性の向上が利用者満足度の向上に結びつく」と仮定し、推薦の意外性を考慮した推薦システムを提案している。利用者の嗜好モデルとは別途、習慣モデルを提案し、嗜好モデルと習慣モデル、それぞれによる予測結果の差異を考慮することで、推薦アイテムの意外性を推定している。

これらは、利用者が知っているか否か、習慣的に利用しているか否かという観点から発見性や意外性を定義している。本研究では、これらの観点とは異なり、利用者がある程度予測可能であり、ある程度予測不可能という観点から偶発性を定義し、その偶発性を引き起こす推薦システムの実現を目指している。

3. アイテム混合方式に基づく情報推薦システム

本研究では、二つのアイテムを「混ぜ合わせる」ことで人間にとって新しく価値のあるアイテムを見つけ出すことを可能にする情報推薦システムを開発する。そのためには、アイテムの特徴表現方式および混合方式について定義する必要がある。本章では、それぞれについて記述する。

3.1. アイテムの特徴表現方式

基本的なアイテムの特徴表現方式としては、(a) 特徴ベクトル表現、(b) ビット表現、(c) グラフ表現などが挙げられる。

(a) 特徴ベクトル表現

従来の情報検索や情報推薦においてよく用いられている。対象とするアイテムの種類によって扱う属性も異なるが、基本的にはアイテムに付随するテキストから抽出された特徴語句や、メタデータなどからアイテムの特徴ベクトル化を行う方法が取られている。

(b) ビット表現

単純にアイテムに関連するキーワードやタグの有無を{0,1}で表現するビット表現にして扱う。

(c) グラフ表現

ユーザがアイテムに対し、評価値を与えているとすると、評価値の与えられ方からアイテム間の類似関係を算出できる。その関係から、アイテムをノード、アイテム間の関係をエッジとして表現したグラフ構造を形成することができる。このグラフ構造を用いてアイテムを表現するという方法も考えられる。

本稿では、(b) のビット表現を用い、アイテムの特徴表現を行う。3.2 節では、(b) のビット表現を採用したときのアイテムの混合方式について検討する。他の (a) および (c) を採用したときの混合方式については別途検討課題とする。

まず、形態素解析器を用いてアイテムのタイトルや付随するテキスト情報などからキーワードを抽出する。キーワード k_i が抽出されたとき、 $\text{bit}(k_i)=1$ として扱い、抽出されなかったキーワード k_j については、 $\text{bit}(k_j)=0$ として扱う。

3.2. アイテムの混合方式

二つのアイテムの混合方式について検討する。本研究では、混合パターンとして、和・差・積の概念に着目する。先述した特徴表現方式ごとに和・差・積による混合方式を定義する。それぞれの混合方式により算出された結果をクエリとして新たなアイテムを検索する。そして、そのクエリに最も合致するアイテムをユーザに提示する。

ここでは、先述した特徴表現のうち、(b)のビット表現を採用したときの混合方式を定義する。この場合の和・差・積の演算は基本的な論理演算に相当する。以下、和・差・積による混合方式の定義について説明する。説明の例として、二つのアイテム a および b を考える。また、キーワード群として $\{k_1, k_2, k_3, k_4, k_5\}$ を考え、アイテム I のキーワード有無に基づくビット表現を $I = \{\text{bit}(k_1), \text{bit}(k_2), \text{bit}(k_3), \text{bit}(k_4), \text{bit}(k_5)\}$ とする。ここでは、 $a = \{1, 1, 1, 0, 0\}$ 、 $b = \{0, 0, 1, 1, 1\}$ とする。

a) 和混合方式 : $a + b$

アイテム a と b を足し合わせることで、新たなクエリ Q^+ を算出する。和混合方式はビット表現を採用としたとき、基本的な論理和 (OR) に相当する。すなわち、和混合方式は次式により定義される。

$$Q^+ = a + b = \{1, 1, 1, 0, 0\} \text{ OR } \{0, 0, 1, 1, 1\} \\ = \{1, 1, 1, 1, 1\}$$

そして、クエリ Q^+ を用いてアイテムを検索する。すなわち、この例ではすべてのキーワード $\{k_1, k_2, k_3, k_4, k_5\}$ を含むアイテムが検索される。検索結果中の N 件のアイテムを推薦アイテムとする。

b) 差混合方式 : $a - b$

アイテム a から b を差し引くことで、新たなクエリ Q^- を算出する。差混合方式はビット表現を採用としたとき、基本的な論理差 (NOT) に相当する。すなわち、差混合方式は次式により定義される。

$$Q^- = a - b = \{1, 1, 1, 0, 0\} \text{ NOT } \{0, 0, 1, 1, 1\} \\ = \{1, 1, 0, 0, 0\}$$

そして、クエリ Q^- を用いてアイテムを検索する。すなわち、この例ではキーワード $\{k_1, k_2\}$ を含むアイテムが検索される。検索結果中の N 件のアイテムを推薦アイテムとする。

c) 積混合方式 : $a \times b$

アイテム a と b の共通部分を取ることで、新たなクエリ Q^\times を算出する。積混合方式はビット表現を採用としたとき、基本的な論理積 (AND) に相当する。すなわち、積混合方式は次式により定義される。

$$Q^\times = a \times b = \{1, 1, 1, 0, 0\} \text{ AND } \{0, 0, 1, 1, 1\} \\ = \{0, 0, 1, 0, 0\}$$

そして、クエリ Q^\times を用いてアイテムを検索する。すなわち、この例ではキーワード k_3 を含むアイテムが検索される。検索結果中の N 件のアイテムを推薦アイテムとする。

4. 評価実験

実データセットを用いて、提案手法の評価を行う。特徴ベクトル空間法により推薦結果のアイテムと推薦元との類似度を計算し、その結果を用いて手法の有用性を評価する。

4.1. データセット

本実験では、データセットとして、楽天株式会社から提供されているデータセット「楽天トラベル:施設、お客様の声情報」[10]を利用した。本データセットは、全国の宿泊施設に対して利用者から付けられたユーザ評価とユーザレビューがまとめられている。ユーザ評価は「総合」、「立地」、「食事」、「サービス」などの7項目に対し、それぞれ0-5の6段階でつけられている。加えて、施設に対するレビューや、存在する場合は施設からの返答などが付与されている。ユーザ評価データは338,045件、ユーザレビューは348,564件、施設データは11,468件、登録されている。

なお本実験では処理時間を短縮するため、全施設データの中から無作為に選択した1000件の施設データを用いて実験を行った。

4.2. 実験方法

ランダムに選んだ二つの施設データを選ぶ（ここでは、アイテム $a \in A$ とアイテム $b \in B$ とよぶ）。アイテム a と b について、3.2 節で述べた、和・差・積の各混合方式によりクエリを作成し、そのクエリに基づいて推薦アイテム $c \in C$ を選択する。具体的な実験手順は以下のとおりである。

1) アイテム a と b それぞれに関するレビュー文からそれぞれの特徴表現を行う。具体的には、各レビュー

文から TF-IDF により特徴語句を抽出し、TF-IDF 値が最も高い上位 M 個の語句をその施設のキーワードとして扱う。そのキーワードに基づき、3.1 節 (b) のビット表現により施設データの特徴表現を行う。

2) ビット表現されたアイテム a と b に対して、3.2 節で述べた、和・差・積の各混合方式により新たな検索クエリ Q^+ , Q^- , Q^* を作成する。

3) 1,000 件の施設データのレビュー文を対象に、各クエリによる検索を行う。検索結果中の N 件のアイテムを推薦アイテム $c \in C$ とする。

以上の手順により、検索結果として得られた施設データ集合の内容から、定量分析および定性分析を行う。

4.3. 定量分析結果

推薦されたアイテム c が、混合元のアイテム a と b それぞれにどの程度類似しているかという観点から、和・差・積の各混合方式によるアイテム推薦の傾向の違いについて分析する。類似度算出にはベクトル間のコサイン類似度を用いる。各アイテムを 4.2 節で抽出したキーワードに基づいた特徴ベクトルとして表現する。その特徴ベクトル間のコサイン類似度を求める。

まず、各混合方式について期待される結果の分布を図 1 に示した。推薦された結果であるアイテム c に着目したとき、横軸をアイテム a との類似度、縦軸をアイテム b との類似度としてプロットしたものである。

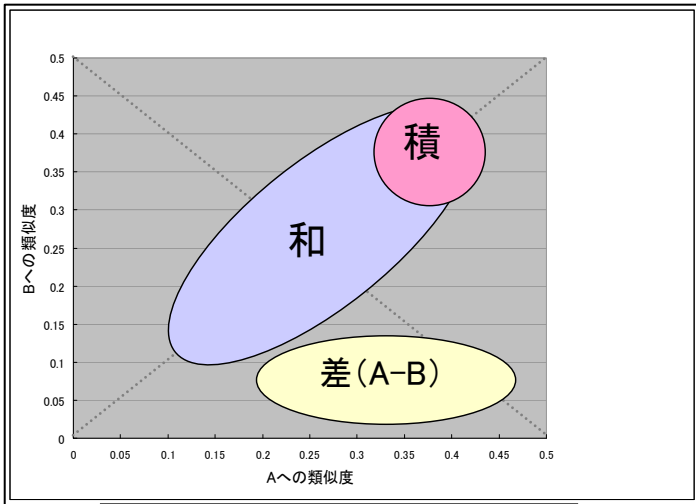


図 1 期待される推薦結果の分布

和混合方式は、 $a \cdot b$ の要素を持った多様な結果が出るのが期待される。すなわちグラフ上では結果が、広い範囲で偏らず分布するのが理想である。

差混合方式は、 a の要素を多く含み、 b の要素をあまり含まない結果が出るのが期待される。すなわちグラフ上では結果が、 a への類似度が高く b への類似度は低いエリアに分布するのが理想である。

積混合方式は、 $a \cdot b$ の要素をどちらも強く含む結果が

出ることが期待される。すなわちグラフ上では結果が、 $a \cdot b$ ともに類似度が高めのエリアに集中して分布するのが理想である。

そして図 2 は、複数の混合パターンについて実際に推薦された結果を和・差・積の混合方式ごとに記号を変えて示している。また、全体的な傾向をみるため、各推薦データ集合の重心および近似直線を図示している。以下、この結果に基づいて、各混合方式について考察する。

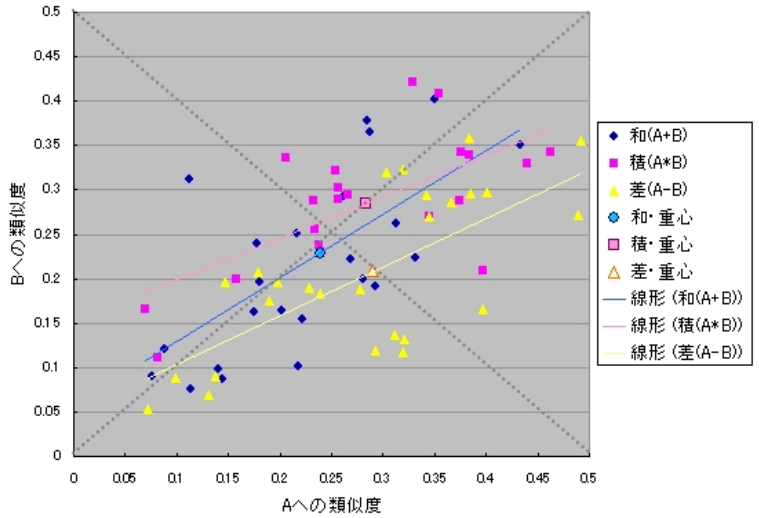


図 2 実際に推薦された結果の分布

a) 和混合方式による推薦結果

アイテム a およびアイテム b との類似度共に、0~0.5 の範囲内に広く分布した。和混合方式では、アイテム a , アイテム b のいずれかに現れるキーワードに基づいて新たなアイテムを推薦するため、アイテム a に類似するがアイテム b には類似しないアイテム、またはアイテム b に類似するがアイテム a に類似しないアイテムが推薦結果に含まれる。この傾向を踏まえると、アイテム a , アイテム b いずれかには関連するが、あまり候補を絞り込まず広くアイテムを探したいときに和混合方式が用いられることが考えられる。

b) 積混合方式による推薦結果

積混合方式による推薦結果では、その推薦データ集合の重心が和・差混合方式のそれと比べ図中の右上方向（すなわち、アイテム a, b との類似度が共に 0.2~0.5 の範囲）に現れた。積混合方式では、アイテム a とアイテム b の両方に現れるキーワードに基づいて新たなアイテムを推薦するため、よりの絞った推薦結果が含まれる。積混合方式により、ユーザは二つのアイテムを提示することで、それらに共通する特徴を持つアイテムを探したいという意図を示すことができる。

c) 差混合方式による推薦結果

差混合方式による推薦結果では、和混合方式と同様に分布の広がりが見られたが、若干アイテム a に類似するアイテムが推薦結果に多く含まれる傾向が見られた。差混合方式では、アイテム a に現れるが、アイテム b には現れないようなキーワードにより新たなアイテムを推薦するため、この結果は自明であるといえる。差混合方式により、ユーザは否定的なアイテムとして b を指定しながら、肯定的なアイテム a に関連するアイテムを探したいときに用いることができる。

4.4. 定性分析結果

提案手法による推薦結果に対し、具体的な事例を取り上げ、定性的な分析を行う。

以下の二つの施設データをそれぞれ混合元アイテム a, b としたときの推薦結果をみる。

a 「伊勢かぐらリゾート 千の杜」

b 「箱根湯本温泉 おかだ」

a は三重県にあるリゾート施設であり、露天風呂と伊勢海老が名物である。「露天風呂」「伊勢」「朝食」「松坂牛」といった語句の TF-IDF 値が高くなったため、これらをキーワードとした。

b は箱根の旅館であり、12の温泉と海の幸、山の幸を使った食事が名物である。「温泉」「舟盛り」「風呂」といった語句をキーワードとした。どちらにも共通するのは「風呂」が魅力的ということであり、またどちらも山と海が近くにあるという点が特徴である。これらを混合したとき推薦されたアイテム集合を $C=\{c_1, c_2, \dots\}$ とし、各結果について考察する。

a) 和混合方式による推薦結果

c_1 「うぐいす谷温泉 竹の葉 (旧:竹葉亭)」

温泉を売りにしている点で、 b との類似性が見られる。また地域は違うが山と海が近くにあるロケーションではある。

c_2 「伊豆高原 海一望 アクアガーデン ラグドール」

a の特徴である露天風呂付き客室がここにもある。また三重県ではないが、伊勢海老が食べられるという点も類似している。

c_3 「花の小宿 重兵衛」

a と同じく三重県に所在しており、伊勢海老を名物にしている。また b の特徴である温泉で有名である点も持つ。

b) 積混合方式による推薦結果

c_1 「伊豆高原海一望 アクアガーデン ラグドール」

和混合方式による推薦結果にも含まれる施設が推薦

された。これは a の「露天風呂」「伊勢海老」という語句、 b の「温泉」「舟盛り」という語句を全て含んでいたのが要因である。

c-1) 差混合方式による推薦結果 ($a-b$)

c_1 「アパヴィラホテル<仙台駅五橋>」

チェーンのシティホテルではあるが天然温泉を利用した大浴場がついている。

c_2 「下田温泉 下田セントラルホテル」

温泉があり、伊豆ならではの新鮮な海の幸がポイントである。

c_3 「湯の川温泉 平成館 しおさい亭」

函館の温泉施設である。

共通するのが、全て規模が大きい近代的な施設である。 b が古風な温泉施設ということを見ると、 b にみられる特徴が排除されたといえる。

c-2) 差混合方式による推薦結果 ($b-a$)

c_1 「あわら温泉 政竜閣」

露天風呂付き客室と地元産のカニが人気である。

c_2 「湯野浜温泉 福宝館 漁師の宿」

漁師から直接仕入れる海の幸がポイントの温泉民宿である。

c_3 「網元の宿 徳兵衛」

新鮮な舟盛りが魅力の伊豆の宿。温泉の湯を利用した浴場がある。

c_4 「弥彦温泉 ホテルヴァイス」

千年の歴史を持つ弥彦神社に併設されている。やはり温泉に関するレビューが多い。

全て海鮮が人気の温泉宿である。規模は小さめの、古風な施設が多い。 b の「舟盛り」「カニ」という語句に適合する施設が多い。しかし、($a-b$) において b の特徴である「温泉」という語句をマイナスに指定したにもかかわらず、温泉旅館が推薦されている。これは温泉に類似する語句として「浴場」などの語句が排除されなかったことが原因だと思われる。対策としては、例えば、排除しようとする「温泉」との関連度を調べ、それに基づき、関連する語句も排除する方針などが挙げられる。

和混合方式による推薦結果では、温泉・露天風呂を基本とした上で、伊勢海老や海鮮など a または b の特徴を含んだ結果が出ている。積混合方式による推薦結果は1件しか得られなかったが、 a, b の両方が持つ特徴を含んでいるアイテムが推薦されたといえる。

第1章での述べた、「ある程度予測可能であり、

ある程度予測不可能」という偶発性の観点からみると、和・積混合方式による推薦結果では、温泉旅館が中心となっている点から予測可能ではあるといえる。予測不可能性に関しては、今回の実験からでは十分に分析できていないため、今後多様なデータセットを対象に被験者実験などで分析を行っていく必要がある。

5. 考察

定量分析の結果から、大きな違いは見られなかったものの、和・差・積の各混合方式の特徴が推薦結果に表れる傾向がみられた。今後は、多様なジャンルのデータセットを用い、詳細な実験を行うことで、この傾向についてより深く分析していく。

この傾向は、特に和・積・差の混合方式いずれにおいても、アイテム a とアイテム b の類似度が高い場合にみられた。似たようなアイテムを混合したほうが、期待通りの混合結果が出るということである。混合元のアイテム間の類似度が推薦結果に及ぼす影響に関しては、今後の検証課題である。

定性分析の結果から、差混合方式については、その特徴が表れるような結果とはいえなかった。差混合方式 $a - b$ の結果としては、 a との類似度が高く、 a との類似度が低いアイテムが抽出されることが理想的ではあるが、本実験では a にも b にも類似度が高いアイテムが多く推薦された。これは $a \leftrightarrow b$ の類似度が低い場合に起こっているように見える。原因としては、TF-IDF方式によって特徴的なキーワードを抽出して用いていることであると考えられる。 $(a - b)$ を行っても、そもそも a と b の要素がかぶっていないので意味がなく、またその上で c と b の「特徴的な語」でない語に類似性が見られたため、 b の類似度が高くなってしまっている。この特徴語句の抽出方式についても検討が必要である。

6. まとめ

本稿では、ある程度予測可能である程度予測不可能な性質のことを「偶発性」とよび、その偶発性を引き起こす推薦システムについて「アイテムを混ぜ合わせる」という行為に着目した推薦システムについて検討を行った。アイテム混合方式としては、基本的な論理演算に基づいた和・差・積混合方式を提案した。

評価実験では、楽天トラベルの宿泊施設レビューを用いた。レビューに基づき施設データを特徴表現し、和・差・積混合方式により新たなアイテムを推薦することを行った。評価実験の結果、和・差・積の各混合方式による推薦結果について、大きな違いはみられなかったものの、各混合方式の特徴が推薦結果に表れる傾向がみられた。

今回はアイテムの特徴表現方式としてキーワードの

有無によるビット表現を用いたが、今後は、他の特徴表現方式である、特徴ベクトルやグラフ表現などについても検証を行う。また、各混合方式の特徴を考慮しながら、アイテム混合の考え方に基づく新しい情報推薦システムの構築を目指し、主に以下の点について検討していく。

- ・混合元のアイテム同士の類似度に応じてどのように推薦結果が変化するか傾向を分析する。

- ・他の様々なジャンルのデータセットや、巨大なデータセットにおける推薦結果について分析する。

- ・他の特徴表現方式を採用した際の変化について分析する。

- ・混合方式として、和・差・積の概念だけでなく、遺伝的アルゴリズムの交叉の考え方なども取り入れながら検討を行う。

- ・被験者実験による推薦システムの評価を行う。

謝辞

本研究では楽天株式会社から提供された「楽天トラベル施設情報、お客様の声」を利用した。ここに記して謝意を表します。

参考文献

- [1] Resnick, P., N. Iakovou, M. Sushak, P. Bergstrom, and J. Riedl. GroupLens: An open architecture for collaborative filtering of netnews, In Proceedings of the 1994 Computer Supported Cooperative Work Conference, 1994.
- [2] 土方 嘉徳: 嗜好抽出と情報推薦技術, 情報処理学会誌, Vol.48, No.9, pp.957-965, 2007.
- [3] 村上 知子, 森 紘一郎, 折原 良平: 推薦の意外性向上のための手法とその評価, 人工知能学会論文誌, Vol.24, pp.428-436, 2009.
- [4] 奥 健太, 中島 伸介, 宮崎 純, 植村 俊亮: 状況依存型ユーザ嗜好モデリングに基づく Context-Aware 情報推薦システム, 情報処理学会論文誌: データベース, Vol.48, No.SIG 11 (TOD34), pp.162-176, 2007.
- [5] 岩濱 数宏, 土方 嘉徳, 西田 正吾: 決定木を用いた音楽情報フィルタリングシステムとその有効性の検証, 電子情報通信学会論文誌 D1, Vol.J88-D-I, No.3, pp.642-656, 2005.
- [6] Amazon.com: <http://www.amazon.com/>
- [7] Linden, G., B. Smith, and J. York. Amazon.com Recommendations: Item-to-Item Collaborative Filtering, IEEE Internet Computing, Jan.-Feb. 2003.
- [8] 清水 拓也, 土方 嘉徳, 西田 正吾: 発見性を考慮した協調フィルタリングアルゴリズム, 電子情報通信学会論文誌 D, Vol. J91-D, pp.538-550, 2008.
- [9] 村上 知子, 森 紘一郎, 折原 良平: 推薦の意外性向上のための手法とその評価, 人工知能学会論文誌, Vol.24, pp.428-436, 2009.
- [10] 楽天データ公開, <http://rit.rakuten.co.jp/rdr/index.html>