

潜在的興味分析に基づく詳細な興味グループ判定方式の提案

澤田 礼我[†] Panote Siriaraya[†] 堀 篤史[†] 山口 由莉子^{††} 森下 民平^{†††}
 稲垣 陽一^{†††} 中本 レン^{†††} 張 建偉^{††††} 中島 伸介[†]

[†] 京都産業大学 コンピュータ理工学部 〒 603-8555 京都府京都市北区上賀茂本山

^{††} 京都産業大学大学院 先端情報学研究科 〒 603-8555 京都府京都市北区上賀茂本山

^{†††} 株式会社きざしカンパニー 〒 103-0015 東京都中央区日本橋箱崎町 20-14 日本橋巴ビル 6F

^{††††} 岩手大学 理工学部 〒 020-8550 岩手県盛岡市上田 3 丁目 18-8

E-mail: [†]{g1544719,k6180,g1445223,nakajima}@cc.kyoto-su.ac.jp, ^{††}i1658164@cse.kyoto-su.ac.jp,

^{†††}{mimpei,inagaki,reyn}@kizasi.jp, ^{††††}zhang@iwate-u.ac.jp

あらまし 近年、企業や法人では、自社の製品・サービスの購買者、使用者など、自社の製品・サービスに興味を持つユーザグループ（以下、興味グループ）を判定し、自社の製品・サービスの分析、評価が行われている。しかしながら、従来の興味グループ判定は、一般的に、単純な興味グループ判定によって行われている状況であり、潜在的な興味を考慮した詳細な興味グループを判定するには限界がある。そこで、本研究では、Web 空間におけるユーザの行動履歴から、潜在的興味を考慮した詳細な興味グループ判定方式の開発を目指す。具体的には、特定のものに興味のあるユーザの閲覧履歴を取得し、それらの閲覧履歴情報を興味グループ判定システムに入力することで特徴を持つ興味グループを判定する。この興味グループ判定システムを用いる事による潜在的興味分析方式を提案する。

キーワード web 広告、ユーザプロファイリング、アクセスログ分析、潜在的興味分析

1. はじめに

近年、企業や法人では、自社の製品・サービスの購買者、使用者など、自社の製品・サービスに興味を持つ興味グループを判定し、自社の製品・サービスの分析、評価が行われている。しかしながら、従来の興味グループ判定は、一般的に、単純な興味グループ判定によって行われている状況であり、潜在的な興味を考慮した詳細な興味グループを判定するには限界がある。そこで我々は、ユーザの潜在的興味を考慮した詳細な興味グループ判定方式を用いた Web 広告推薦手法の開発や潜在的興味分析に基づく詳細な興味グループ判定方式を用いたマーケティングでの活用を最終目標としている。この実現を目指し、本研究では、潜在的興味分析に基づく詳細な興味グループの判定方式の提案を目的としている。

ここで、我々が提案する潜在的興味分析に基づく詳細な興味グループ判定方式を用いたマーケティング場面を想定した具体例（図 1 参照）を示す。X 社の車に関する興味グループ判定を行った場合、従来（図 1 左図）は、車好きグループという一つのグループ以外、ターゲットとなる興味グループを判定することが困難な状態であり、マーケティング戦略を取る場合もこの単純な興味グループに対してマーケティングを行うこと以外、他の方法が少ない状態である。しかし、車好きグループという一つのグループの中にも様々なユーザが存在し、いくつかのグループに分けることができると考えられる。そこで、提案する詳細な興味グループ判定システム（図 1 右図）を用いた場合、例えば、車好きグループには、図 1 の右図のような 3 つの興味グループを判定をすることができると考えられる。そし

て、車好きグループ中に旅行好きグループとアウトドア好きグループという興味グループの存在を認知していた場合、これまで認知されていない、映画好きグループを新たに発見する事が可能となり、例えば、映画館や映画雑誌に広告を出すなど、映画好きグループに対して、効果的なマーケティング戦略を取ることが可能になると考えられる。以上のような、ユーザの潜在的興味を考慮した Web 広告推薦手法の開発や潜在的興味分析に基づく詳細な興味グループ判定方式を用いたマーケティングでの活用を目指し、本研究では詳細な興味グループ判定システムを構築し、潜在的興味分析に基づく詳細な興味グループの判定方式を提案する。

以下、2 章にて、関連研究との比較を述べ、3 章にて、具体的な提案手法の流れを説明する。4 章では、潜在的興味分析に基づく詳細な興味グループの判定の妥当性を示すための評価実験について述べる。最後に 5 章にて、まとめと今後の課題を述べる。

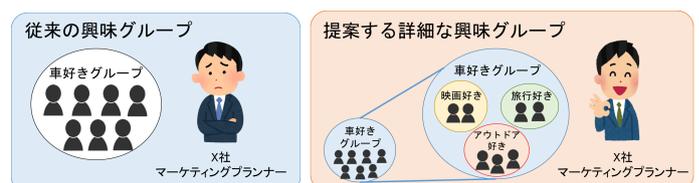


図 1 従来の興味グループ判定と提案する興味グループ判定の比較

2. 関連研究

以下に、Web 広告および詳細な興味グループ判定に関連した

研究について述べ、我々の提案手法との差異を示す。

鈴木らは Web サイトのアクセスとログとデータを用いて任意のトピック消費者の購買行動を明らかにするするため、購買行動に混合分布を当てはめて、購買サイクルを推定し購買の前後の行動の特徴を分析している [1]。また、生田目らは EC サイトのアクセスとログとデータを用いてサイト会員の日常の閲覧行動を考慮した購買予兆の発見モデルの提案 [2] をしており、また久松らはその購買予兆を発見するモデルをロジット・モデルをもとに作成している [3]。以上はユーザの購買予兆を発見し広告を表示するという研究を行っているが、本研究では購買する直前を発見するのではなく、潜在的興味に基づく詳細な興味グループを考慮して推薦する広告を決めることを目的としている。

内野らはユーザが次に見たい情報を予測し、それに関する広告を配信する Web 広告配信システムに kMER およびマルコフモデルを応用した研究を行っている [4]。本研究ではユーザらが次に見たい情報を予測するのではなく、潜在的興味に基づく詳細な興味グループ判定を用いた広告の推薦を目的としている。

小河らは消費者の購買行動における段階を考慮し、消費者が求めている商品情報とデザインおよびメッセージを個人に合わせて推薦する、インターネット広告の構成手法を提案している [5]。本研究ではメッセージやデザインをユーザに合わせるのではなく、潜在的興味に基づく詳細な興味グループ判定を考慮して広告を推薦することを目的としている。

潜在的興味分析を考慮した Web 広告推薦手法に関する研究は、我々の先行研究でも行われている。閲覧 Web ページのカテゴリ情報を用いた潜在的興味分析手法に基づく Web 広告推薦方式を提案している [6] [7]。

以上より、潜在的興味分析に基づく詳細な興味グループ判定を考慮した Web 広告推薦やマーケティング戦略は行われていない。

3. 潜在的興味分析に基づく詳細な興味グループの判定方式

本節では、潜在的興味分析に基づく詳細な興味グループ判定方式について述べる。3.1 節で提案する詳細な興味判定グループ判定方式の概要について述べ、3.2 節で詳細な興味グループ判定システムの処理手順について述べる。

3.1 潜在的興味分析に基づく詳細な興味判定グループ判定方式の概要

興味グループ判定は、単純な興味グループ判定しか行われておらず、広告主やマーケティングプランナーにとって、新しい購買者、購買層を獲得することが難しい。そこで、提案する潜在的興味分析に基づく、詳細な興味グループ判定方式について図 2 に沿って説明する。

本研究では、閲覧履歴情報を用いたユーザのクラスタリングにより詳細な興味グループ判定を行う手法を提案する。事前にターゲットサイトにアクセスしたことがあるユーザらの閲覧履歴を取得する。それらの閲覧履歴情報を、興味グループ判定システムに入力する。システムは特徴を持つ興味グループを判定し、判定結果を詳細な興味グループとして出力する流れである

(図 2 参照)。また、本稿で行った実験では閲覧履歴よりユーザ毎のあるターゲットサイトに対するアクセス回数とカテゴリを特徴量として用いる。これらの特徴量については、4.1 節にて説明する。ここで、具体的な詳細な興味判定方式の例を挙げる。X 社の車に関連するサイトにアクセスしたことがあるユーザ (X 社の車好きユーザとみなす) の閲覧履歴を取得し、その閲覧履歴情報を興味グループ判定システムに入力する。システムにより出力された各クラスタを興味グループとみなし、どのような特徴をもつ興味グループであるのか分析する。判定した結果、X 社の車好きグループの詳細な興味グループは、アウトドア好きグループ、旅行好きグループ、映画好きグループのように特徴を持ったグループを出力する。実際に、Web 広告推薦に用いる場合は、この興味判定グループ判定結果を広告主に提示し、アウトドア好きグループには、SUV のようなアウトドア車の広告を推薦するように、広告主自身がどのような広告を推薦するのか考えることが可能となる。以上が本研究の提案手法である。

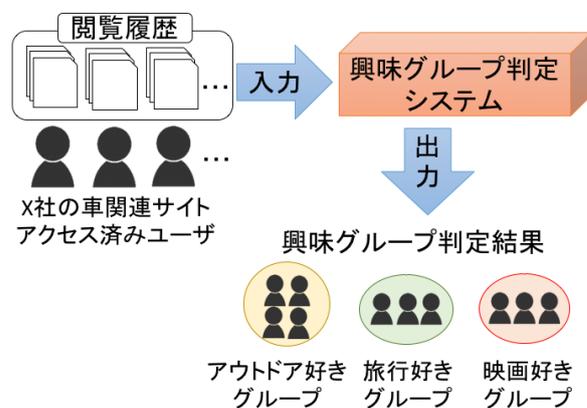


図 2 提案する詳細な興味グループ判定方式の流れ

3.2 詳細な興味グループ判定システムの処理手順

図 3 における、黒字 ((1) ~ (3)) で記載した処理手順である。以下に処理手順を示す。

- (1) 入力された閲覧履歴情報を用いてユーザのクラスタリング。
- (2) クラスタリング結果より各クラスタの潜在的興味分析。
- (3) 潜在的興味を考慮した各クラスタの特徴を判別し、詳細な興味グループを判定。

各手順の詳細を以下に述べる。(1) のユーザのクラスタ分類では、クラスタリング手法の 1 つである k-means 法を用いる。計算量も少なくできるため、ユーザの閲覧履歴情報といった膨大な量のデータを分析するのに有効である。

(2) では、クラスタリング結果より、各クラスタ内でどのような興味を持つユーザが存在するのか分析を行う。分析方法は、TF-IDF やカテゴリ出現率などを用いて、クラスタ毎の特徴を抽出する。

(3) では、各クラスタの潜在的興味を分析した結果を考慮

し、どのような特徴を持つ興味グループであるのか判定する。判定結果を詳細な興味グループとして出力する。

以上が詳細な興味グループ判定を行うシステムの処理手順である。

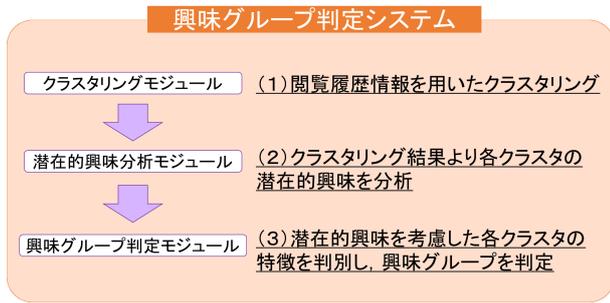


図 3 詳細な興味グループ判定システムの処理手順

4. 潜在的興味分析に基づく詳細な興味グループ判定方式に関する実験的考察

本節では、潜在的興味分析に基づく詳細な興味グループ判定方式に関して、実験に基づく検討を行ったので説明する。

本実験では、あるターゲットサイトにアクセスしたことのああるユーザら一つの興味グループとみなし、その興味グループの詳細な興味グループ判定の妥当性を検証することを目的とする。4.1 節では、実験条件やクラスタリング結果について、4.2 節では、実験の結果及び考察について述べる。

4.1 実験概要

本実験では、ターゲットサイト毎にアクセスしたことのああるユーザの閲覧履歴を取得し、それらの閲覧履歴情報を用いて、ユーザのクラスタリングを行う。そして、クラスタリング結果より、各クラスターがどのような特徴を持つクラスターであるか分析し、判定を行うことで、その判定結果をその興味グループの詳細な興味グループとして扱う。今回の実験では、クラスタリング結果より、各クラスターがどのような特徴を持つクラスターであるのかを特定するための分析方法について検討する。

本実験で用いたターゲットサイトは、以下のサイト (a) とサイト (b) の 2 サイトである。

サイト (a) 遊んでポイントが貯まるお得で楽しいポイントサイト

サイト (b) 東証一部上場企業が運営する安心・便利な共通ポイントサイト

サイト (a) とサイト (b) はどちらもポイントサイトである。また、閲覧履歴のデータには、ユーザ ID、ページカテゴリへのアクセス回数、カテゴリ ID が含まれる。ベクトルはユーザ毎のアクセス回数を、特徴量はカテゴリ ID を用いた。実験に使用したデータの全ユーザ総数、各クラスターのユーザ総数は以下の通りである。

サイト (a) :

ユーザ総数 : 12117

クラスター 1 のユーザ数 : 5689

クラスター 2 のユーザ数 : 3207

クラスター 3 のユーザ数 : 3221

データ取得期間: 1 日間

サイト (b) :

ユーザ総数 : 902

クラスター 1 のユーザ数 : 413

クラスター 2 のユーザ数 : 231

クラスター 3 のユーザ数 : 258

データ取得期間: 1 日間

以下に実験手順を示す。

手順 1 あるターゲットサイトにアクセスしたことのあるユーザの閲覧履歴情報を用いて、クラスタリング手法の k-means 法によりユーザのクラスタリングを行う。

手順 2 クラスタリング結果より、各クラスターを TF-IDF を用いて特徴を抽出する。TF-IDF を用いた結果をタグクラウドで表示する。その結果から人手で、クラスター毎にどのような特徴を持つクラスターであるのか分析を行う。

(例)あるサイトのクラスタリング結果 (クラスター数 '2')

クラスター1の結果



クラスター2の結果



図 4 ターゲットサイトにアクセスしたユーザの閲覧履歴でクラスタリングを行った結果のタグクラウド表示例

上記手順より、手順 2 で行ったタグクラウド表示の例を図 4 に示す。図 4 は、あるポイントサイトでクラスター数を 2 に設定し、クラスタリングを行った結果を、さらに TF-IDF を用いて、クラスター 1 とクラスター 2 にカテゴリの TF-IDF 値が高いほど大きいフォントサイズでタグクラウド表示したものである。今回は、このタグクラウドで表示された結果を比較することで、各クラ

スタの特徴を判定するための分析を行う。



図5 サイトaにアクセスしたユーザの閲覧履歴でクラスタリングを行った結果のタグクラウド表示例

4.2 実験結果および考察

本節では、前節で述べた2個のターゲットサイト(サイト(a)とサイト(b))に関する実験結果について考察する。

クラスタ数を3に設定し、クラスタリング結果をTF-IDFを用いて、ターゲットサイト毎にタグクラウドで表示した実験結果を図5、図6に示す。

まず、サイト(a)(図5参照)に注目すると、クラスタ1の結果より、映画に関するカテゴリが強く現れている。また、クラスタ2の結果では、アウトレット、クラスタ3の結果では、自動車産業に関するカテゴリが強く現れている。これらの結果より、サイト(a)では、映画、アウトレット、自動車産業の3つの特徴が現れており、3つのユーザグループが存在すると考えられる。このように、自社のサイトを閲覧している詳細なユーザグループを認知することができると考えられる。また、今まで未知であったユーザグループを認知することによって、新規

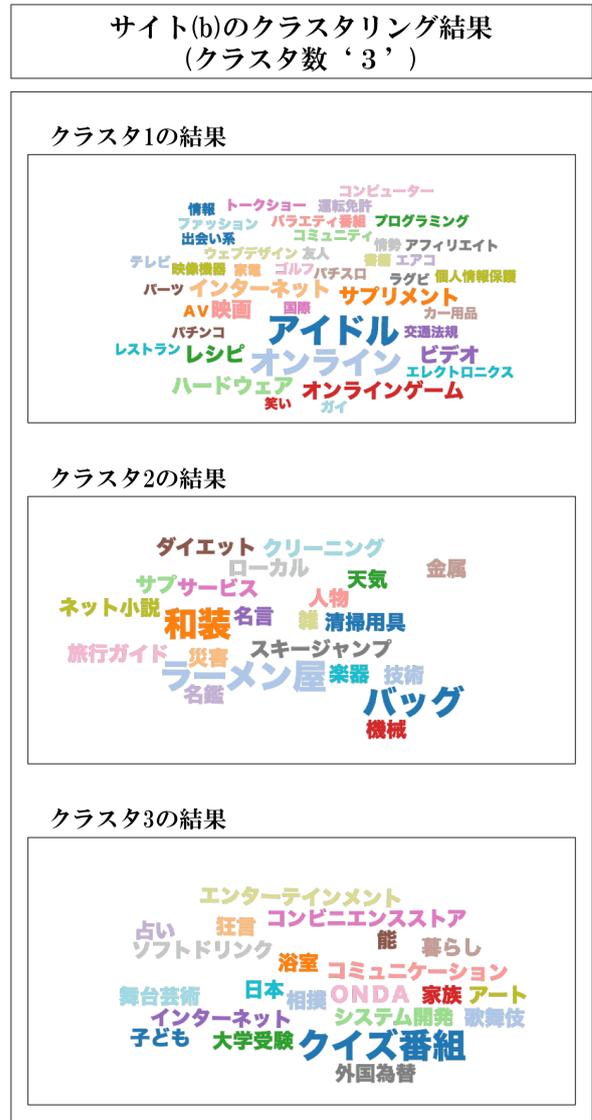


図6 サイトbにアクセスしたユーザの閲覧履歴でクラスタリングを行った結果のタグクラウド表示例

ユーザ獲得が可能になると期待できる。

次に、サイト(a)(図5参照)とサイト(b)(図6参照)を比較する。サイト(a)では前述で述べた通り、映画やアウトレットや自動車産業カテゴリ関連のユーザグループが存在していると考えられる。一方、サイト(b)ではアイドルやラーメンやクイズ関連のユーザグループが存在していると考えられ、サイト(a)とは違うユーザグループが存在していると推測できる。このように、サイト(a)とサイト(b)は、ポイントサイトという同じカテゴリサイトであるが、サイト毎に異なる特徴を持つユーザグループが存在すると考えられ、同じカテゴリサイトであっても区別することができる。また、詳細なユーザグループを特定することで、同カテゴリサイトであっても、それぞれのサイト毎に合わせたマーケティング戦略をとることやWeb広告推薦が可能になると考えられる。

以上より、閲覧履歴のクラスタリングおよびタグクラウド表示を行うことで、各サイトの閲覧ユーザのグループ分けを行い、

その特徴の違いの把握が可能であることを確認した。今後は、ポイントサイト以外のカテゴリサイトで実験を行うことやポイントサイトに関するターゲットサイト数を増やし、ターゲットサイト毎の新たな特徴を見つけ出すことを目指して、引き続き実験を進めていく予定である。

5. ま と め

Web 広告推薦やマーケティングにおいて、潜在的興味分析に基づく、詳細な興味グループ判定方式の提案を行った。さらに、ターゲットサイト毎の詳細な興味グループ判定の妥当性を検証するための実験概要についても説明を行った。今後は、まず、ターゲットサイト毎に詳細な興味グループ判定の妥当性を検証するための評価実験を継続し、自動判定を行う詳細な興味グループ判定システムの構築を考えている。また、実験に関しては、最適なクラスタ数の検討、潜在的興味分析方法や興味グループ判定方法などの実験条件を様々なものに変更し、比較することでターゲットサイト毎の詳細な興味グループの特徴や傾向を捉えることができると考えている。特に最適クラスタ数については、x-means 法を利用して自動的に求められる。将来的には、潜在的興味分析に基づく詳細な興味グループを考慮した Web 広告推薦やマーケティングでの活用を目指す。

謝 辞

本研究の一部は、岩手大学平成 29 年度研究人材育成支援経費、JSPS 科研費 26330351, 17H01822 による。ここに記して謝意を表します。

文 献

- [1] 鈴木元也, 生田目崇, 購買前後のアクセスを考慮した Web サイトの顧客行動分析, 日本オペレーションズ・リサーチ学会 2012 年秋季研究発表会 (2-F-3), 2012.
- [2] 生田目崇, 朝日真弓, 久松俊道, 外川隆, 顧客の閲覧行動を考慮した購買予兆発見モデル, 日本オペレーションズ・リサーチ学会 2012 年秋季研究発表会 (2-F-2), 2012.
- [3] 久松俊道, 外川隆, 朝日真弓, 生田目崇, EC サイトにおける購買予兆発見モデル, 日本オペレーションズ・リサーチ学会 2012 年秋季研究発表会 (2-F-2), 2013.
- [4] 内野英治, 森田博彦, 下野雅芳, Web 広告動的配信システムへのマルコフモデルと kMER の応用, 22nd Fuzzy System Symposium (Sapporo, Sept. 6-8, 2006) 6B1-1, 2006.
- [5] 小河真久, 原田史子, 島川博光, 消費者の情報探索行動に着目した広告の内容と表示の個別化, 情報処理学会研究報告 Vol.2010-DBS-150 No.17 Vol.2010-IFAT-99 No.17, 2010.
- [6] 山口由莉子, 森下民平, 稲垣陽一, 中本レン, 張建偉, 青井順一, 中島伸介, ユーザの潜在的興味に基づく Web 広告推薦方式の検討, DEIM Forum 2016 B1-2, 2016.
- [7] 山口由莉子, 森下民平, 稲垣陽一, 中本レン, 張建偉, 青井順一, 中島伸介, Web 広告推薦のための閲覧カテゴリ情報を用いたユーザの潜在的興味分析, DEIM Forum 2017 B2-3, 2017.