

D-013

展示会におけるブース推薦のための特徴ベクトルと満足度関数を用いた興味推定

Estimation of User Interest to Recommend Booths
with Impress Vectors and Utility Function中川 陽介[†]
Yosuke Nakagawa原田 史子[‡]
Fumiko Harada島川 博光[‡]
Hiromitsu Shimakawa

1. はじめに

近年、東京モーターショーや CEATEC JAPAN[1] などの、来場者数が数十万人規模の大規模イベントが多く開催されている。具体的規模は、CEATEC JAPAN 2007 を例にとると、来場者数 205,859 人で出展企業数は 895 社である [1]。大規模イベントの来場者は、イベント会場内に設置されているブースを見て回る。しかし、大規模イベントでは、ブースの数は数百に及び、会場も広い。そのため、数百もあるブースの中から興味をもてるブースを探す来場者の負担は大きい。来場者の負担軽減のために、来場者の興味に沿ったブースの推薦が必要になる。ブースの推薦を行うためには、来場者の興味を推定しなければならない。

イベント会場には、多種多様な展示内容のブースが設置されている。来場者は興味のありそうな展示内容だと判断したらそのブースを訪れ、展示内容を見終わったら次のブースへ向かう。この来場者の一連の流れの中から、興味推定に必要な情報を取得しなければならない。

本稿では、興味推定に必要な情報を取得し、それらを分析することにより、利用者の興味を推定する手法を提案する。本研究では、興味推定に必要な情報を得るために携帯端末と RFID タグを利用する。興味推定に必要な情報として、ブースの展示内容とそれに対する興味の度合いを用いる。ブースの展示内容は多次元情報として得られるので、本システムでは多次元ベクトルとして表現する。そして、多次元ベクトルと興味の度合いを収集し、分析することで未訪問のブースの展示内容に対する興味の度合いを推定する。推定された興味の度合いの高いブースを推薦することで、来場者の興味に沿ったブースの推薦となる。また、本手法を実装したシステムを用いた実験の結果、50%の割合で興味に沿ったブースの推薦ができた。

2. 大規模イベントの現状

2.1 ブースの推薦

来場者がブースを探す負担を軽減するためには、来場者の興味のあるブースを自動的に推薦する案内役が必要である。来場者の興味のあるブースを自動的に推薦することにより、来場者自身で探すより多くの興味のあるブースに訪れることが可能となる。一方、来場者は自らの興味に沿っていないと思っていても、実際に展示内容を見てみると実は興味のある展示内容である場合がある。このような来場者自身が自覚していない潜在的興味を知ることができれば、より多くの興味あるブースを来場者に推薦することができる。

2.2 既存研究

展示会におけるガイドシステムについては多くの研究がなされている [2]。既存研究において、利用者の興味のある展示内容を推薦しているものもある。たとえば、文献 [2] では、展示ブースに関連キーワードを設定しておき、利用者が訪れたブースのキーワードを基にブースを推薦するシステムが提案されている。また、そのキーワードをもって推薦理由を利用者に提示している。しかし、推薦理由のキーワードが利用者にとって馴染みの深いものでなければ、利用者は利用者自身にとってブースに関心があるか把握できない。つまり、利用者の興味に沿っていたとしても、そのブースに利用者が訪れない可能性がある。また、既存研究では、利用者の潜在的興味に沿った展示物を推薦していない。

3. 興味推定に基づく自動推薦

3.1 ブースの特徴と満足度からの興味推定

本手法では、利用者の興味推定のために、利用者が訪れたブースの展示内容と利用者の展示内容に対する評価の2つの情報を利用する。それらの情報を取得するため、RFID タグと RFID タグリーダを搭載した携帯端末を用いる。本研究では、各ブースにブース固有の ID が書かれた RFID タグを設置し、各利用者はすべてのブースの情報が登録されている携帯端末を持っているという環境を想定する。利用者は、ブースを訪れたときに携帯端末でブースに設置されている RFID タグを読み込む。こうすることにより、利用者が訪れたブースを携帯端末に記録する。そして、利用者は展示内容を見終わったさいに、展示内容に対してどの程度興味があったかという評価を行う。利用者から展示内容に対する評価を取得することで、携帯端末は利用者がそのブースの展示内容にどの程度興味があるのかを把握できる。携帯端末は、利用者が訪問済みのブースの展示内容とそれらへの評価を履歴として保持する。この履歴を分析することにより、利用者の興味を推定する。携帯端末は、推定した興味を基に利用者の興味に沿ったブースを未訪問のブースの中から選出し、自動的に推薦する。

3.2 訪問履歴からの特徴ベクトルの収集

ブースの展示内容は各ブースごとに異なる。利用者が見た展示内容はどのようなものかという情報を得るには、ブースの展示内容をデータ化しなければならない。本研究では、各ブースの展示内容を多次元ベクトルとしてデータ化する。このベクトルを特徴ベクトルと呼ぶことにする。特徴ベクトルの各軸には展示会のテーマに沿った分野が設定されており、各軸の数値は各軸に対応する分野への関連の程度を表す。設定される分野は、すべてのブースで共通である。また、すべてのブースの特徴ベクトルは、利用者がもっている携帯端末にあらかじめ保持されている。利用者はブースを訪れたときに、携帯端

[†]立命館大学大学院 理工学研究科 情報理工学専攻[‡]立命館大学 情報理工学部 情報システム学科

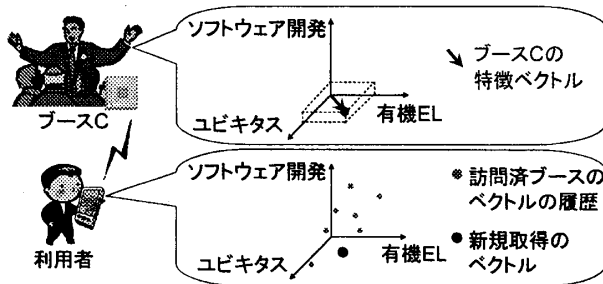


図1: 履歴の取得と特徴ベクトル

末でブースに設置されているRFIDタグを読み込むことで、どのブースを訪れたのかという情報が携帯端末に履歴として残る。

履歴取得の例を図1を用いて説明する。あるIT関係のイベントにおいて、いくつかのブースを訪問済みの利用者は、ブースCを訪れようとしている。ブースCには、ソフトウェア開発、ユビキタス、有機ELの3つの分野を用いた特徴ベクトルが設定されている。ブースCを訪れた利用者は、ブースCに設置されているRFIDタグを携帯端末で読み込む。利用者の今まで訪れたブースの履歴に、ブースCを訪れたという情報を加えて新たな履歴とする。履歴をもとに、収集した訪問済みのブースの特徴ベクトルから、利用者がどのような分野の展示内容を見たかを携帯端末が把握できる。

3.3 満足度の取得

利用者は、興味の深そうなブースを訪れる。しかし、利用者がブースを訪れたからといって、必ずしもそのブースの展示内容が利用者の興味に沿っているとは限らない。そのため本研究では、利用者の興味推定のために、利用者のブースの展示内容に対する興味の度合いを取得する。

展示内容に対する興味の度合いを満足度と呼ぶことにする。満足度は、展示内容を見た後に利用者に入力してもらう。ブースに設置されたRFIDタグを携帯端末で読み込むと、携帯端末にアンケート画面が表示される。そして、利用者が答えたアンケート結果が携帯端末により数値化され、その数値を満足度とする。満足度を用いることで、利用者はどのような展示内容に興味があるのかを携帯端末が把握できる。

3.4 主成分分析による特徴ベクトルの低次元化

ブースの展示内容に対する満足度の分布を得ることができれば、未訪問のブースの満足度を推定できると考えられる。しかし、ブースの特徴ベクトルは多次元ベクトルである。多次元ベクトルのような情報量が大きいデータを、処理能力の低い携帯端末で扱うと、携帯端末に大きい負荷がかかる。そこで、情報量の欠損をなるべく低く抑えつつ、多次元ベクトルを低次元化する手法として、主成分分析[3]を用いる。主成分分析を行うことで、特徴ベクトルをより少ないデータサイズで表現できる。

訪問済みのブースの特徴ベクトルを主成分分析し、低次元化するまでの過程を図2に示す。訪問済みの特徴ベクトルに対して主成分分析を行うと、特徴ベクトルをより低い次元で表現するための新しい軸が求まる。この新しい軸を主成分という。また、元の各特徴ベクトルを主成分上にマッピングした値を主成分値と呼ぶ。この主成分値が、特徴ベクトルを低次元化したデータとなる。主

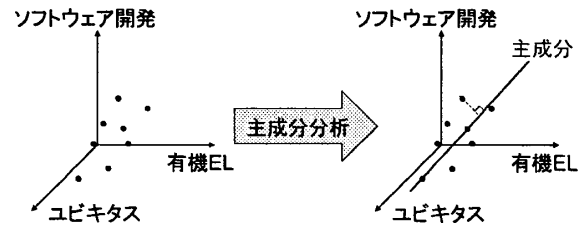


図2: 主成分の導出

成分分析を用いて特徴ベクトルを低次元化することにより、利用者の興味推定のための計算を容易にすることができる。

3.5 低次元化した特徴ベクトルと満足度による興味推定

特徴ベクトルの主成分値と利用者から取得した満足度から、携帯端末は利用者の興味を推定する。まず、特徴ベクトルの主成分値と、満足度を軸とする、新しい空間を構成する。そして、訪問済みブースの特徴ベクトルの主成分値と満足度の組を、新しい空間上に点としてプロットする。訪問済みのブースの点の分布から、未訪問のブースの点も含めた全体の点の分布がどのような傾向になるかを分析する。その分析結果として、求められた全体の点の分布の傾向を満足度関数と呼ぶ。満足度関数を用いることで、未訪問のブースの満足度を推定することができる。つまり、求めた満足度関数が、推定した利用者の興味である。

特徴ベクトルと満足度の分布の傾向の分析から、利用者の興味推定までを図3に示す。特徴ベクトルの主成分値と満足度による二次元平面上の分布において、点間の補完をすることにより、全体の分布を曲線として求める。この曲線が満足度関数となる。未訪問のブースにおいて特徴ベクトルの主成分値さえ計算すれば、求めた満足度関数を用いることで未訪問のブースの満足度を容易に推定することができる。推定された満足度の高いブースは、利用者の興味に沿ったブースである。本システムは、すべての未訪問のブースにおいて満足度を推定し、推定された満足度の高い順にブースを並べ、上位いくつかのブースを利用者に推薦する。

4. 評価

4.1 実験目的と内容

本システムが、実際に利用者の興味に沿ったブースを推薦できるかを検証するための実験を行った。なお、本実験では仮想的なイベント会場としてホテル一室の広さの部屋を用意し、その会場で実験を行うこととした。仮

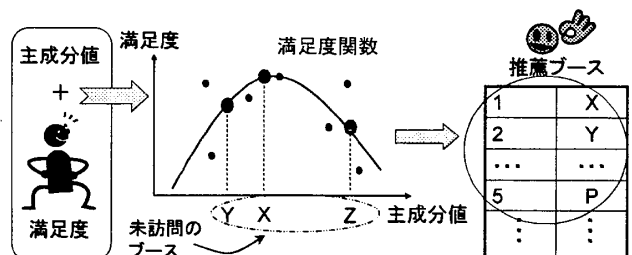


図3: 満足度関数による満足度の推定

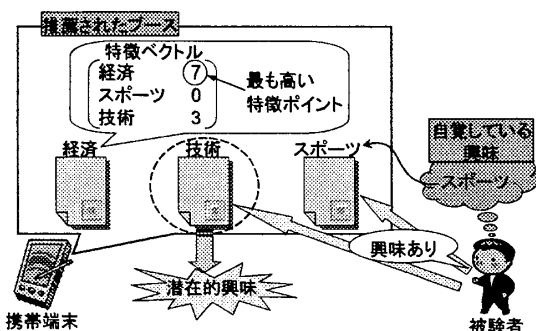


図4: 潜在的興味の抽出

想的なイベント会場では、出展内容を展示するブースの代わりとして新聞記事を用いる。新聞記事の内容がブースの展示内容となる。すべての新聞記事には実際のブースと同様にRFIDタグを貼り付け、それらを仮想イベント会場の机や壁に設置する。一方、携帯端末にはすべての新聞記事の特徴ベクトルがあらかじめ保持されている。

今回の実験の被験者は6人である。まず、実験部屋を興味推定用の新聞記事があるエリアAと、推薦対象の新聞記事があるエリアBの二つに分ける。各エリアには、新聞記事が30本ずつ存在する。各被験者には、RFIDリーダライタがついた携帯端末を持ってもらう。始めに、被験者にエリアAの新聞記事から10本の新聞記事を読んでもらう。被験者は訪れた興味推定用の新聞記事に対して、携帯端末に満足度を入力する。携帯端末は、読んでもらった10本の新聞記事の履歴から、被験者の満足度関数を求め、エリアBに存在する推薦対象の新聞記事から5本の新聞記事を推薦する。次に、被験者にはエリアBにおいて新聞記事30本すべてを読んでもらう。エリアBの新聞記事の被験者の満足度を入力してもらうことで、各ブースに対する実際の満足度がわかる。エリアBにおいて推薦された新聞記事が被験者の興味に沿っているかを、事後アンケートによって調査する。これにより、携帯端末が利用者の興味を正しく推定できているか検証する。

4.2 結果と考察

被験者6人に対しアンケートを実施し、システムによって推薦された5つの新聞記事に、実際に興味があるかを答えてもらった。実験では、被験者1人あたり5本の新聞記事を推薦している。被験者全体で30本の新聞記事を推薦したことになり、そのうち15本の新聞記事に興味があると答えてもらった。つまり、平均して50%の割合で、利用者の興味に沿ったブースを推薦できた。

また、今回のアンケートで潜在的興味が推定できているかの評価も行った。具体的方法を、図4示す。アンケートでは、特徴ベクトルに設定された分野の中から、被験者の興味ある分野を挙げてもらった。そして、推薦された新聞記事の最も高い特徴ポイントをもつ分野を、その記事が最も関連する分野とする。被験者が興味を示した推薦記事の最も関連する分野のうち、被験者が興味ある分野以外の分野を潜在的興味の分野とする。今回は、被験者一人あたり一つでも潜在的興味の新聞記事があった場合、潜在的興味が推定できたとする。その結果、6人中3人において、潜在的興味が推定できた。潜在的興味が推定できた3人において、被験者自身自ら挙げた興味のある分野以外の分野に最も関連している新聞記事に興

表1: 潜在的興味の記事に対する満足度

	興味がある 記事の満足度	潜在的興味分野の 記事全体の満足度の平均
被験者 X	4	2.33
被験者 Y	4	3
被験者 Z	5	3

味があると答えた。

表1には、潜在的興味が抽出できた被験者が興味があると答えた新聞記事の実際の満足度と、その被験者における潜在的興味の分野に最も関連する新聞記事全体の実際の満足度の平均を表した。被験者が興味があると答えた新聞記事の実際の満足度は、被験者の潜在的興味の分野の新聞記事の実際の満足度の平均よりも高い。つまり、潜在的興味の分野に最も関連がある新聞記事の中でも、特に興味のある新聞記事を推薦できたといえる。したがって、本システムは利用者にとって有益な推薦を行っている。しかし、すべての被験者において、潜在的興味が推定できたわけではない。自覚している興味の分野を4つ以上と多く挙げていた被験者がいたため、その被験者に対しては潜在的興味が推定されなかったと考えられる。自覚している興味が少なければ、潜在的興味の分野である可能性のブースは多くなる。したがって、本システムは、自覚している興味が少ない利用者に対して、特に有益であると考えられる。また、本手法において、利用者の潜在的興味が検出できたことにより、推薦されるブースの範囲が広まる。推薦されるブースの範囲が広まるということは、利用者はより多くの興味あるブースを訪れることができ、興味のあるブースを見逃す可能性が低くなる。このことから、本システムは利用者にとって有益であると考えられる。

5. おわりに

本稿では、大規模イベントにおいて、利用者の興味に沿ったブースを推薦するために、利用者の興味を推定する手法を提案した。本手法による推薦を受けることで、利用者は自ら興味あるブースを探す負担を軽減できる。また、利用者の興味に沿ったブースを推薦できているかを検証するため、実験を行った。その結果、50%の割合で興味に沿った推薦ができた。そして、被験者6人中3人において、潜在的興味が推定できた。よって、本システムを用いることで、利用者は自ら探すより多くの興味あるブースを訪れることが可能である。

今後は、特徴ベクトルの設定規則の作成と興味の的中率の向上を目指す。

参考文献

- [1] CEATEC JAPAN 運営事務局: CEATEC JAPAN 2008 開催のご案内, http://www.ceatec.com/2008_pre/ja/(参照 2007-01-22).
- [2] 高橋徹, 益岡あや, 深谷拓吾, 伊藤禎宣, 片桐恭弘: ubiNEXT: 自由選択学習を支援する展示ガイドシステム, 第19回人工知能学会全国大会(2005).
- [3] 田中豊, 脇本和昌: 多変量統計解析法, p.53-99, 現代数学社(1983).