

ラフ集合を利用したイベント推薦システムの構築

兵藤 允彦^{1,a)} 奥野 拓^{2,b)}

概要：近年、インターネットの普及に伴い Web を活用したイベント情報の発信が積極的に行われている。しかし、イベント情報の数が多いとその中から興味のあるイベント情報を探すことは困難である。そこで本研究では、主に感性工学で使われているラフ集合を利用してイベント情報を推薦するシステムの構築を行う。ラフ集合とは、様々な属性の組み合わせを対象の特徴としてモデル化する手法である。これにより、推薦対象のイベント情報に対して興味の有無を選択しシステムに学習させることで、ユーザ嗜好の特定が可能となる。多様なジャンルのイベント情報がある中から興味のあるイベント情報を推薦するために、イベント情報のジャンルに適した属性および属性値をそれぞれ決定する。その後、ラフ集合を適用し最も重要度の高い属性の組み合わせを導出する。この結果を利用し、ユーザの興味に合ったイベント情報を推薦する。

Construction of Events Recommendation System Using the Rough Set

HYODO MASAHICO^{1,a)} OKUNO TAKU^{2,b)}

1. はじめに

現在、全国の地方自治体では様々な方法でイベント情報の発信を行っている。近年は、インターネットの普及に伴い、Web を活用したイベント情報の発信も積極的に行われている。Web を活用したイベント情報発信の例として、札幌市で行われているイベント情報を網羅的に発信している「あなた情報マガジンびもー」がある [1]。また、函館市では、「イベント検索サービス HakoEve」という市民向けのイベント情報を発信している Web サイトがある [2]。「イベント検索サービス HakoEve」では、毎月約 120 件の多様な種類のイベント情報が更新されており、函館市近郊の多くのイベント情報を掲載している。このようなイベント情報発信サイトを活用して、ユーザは自分の興味のあるイベントを探すことになる。しかし、イベント情報発信サイトではイベント情報が一覧で表示されていることが多く、イベント情報の数が多いとそこから興味のあるイベントを探すのが困難になる。

さらに、多くのイベント情報発信サイトでは前述したよ

うに多数のイベント情報が一覧として表示され、そこから興味のあるイベント情報を選択し、詳細情報を閲覧する。ユーザは興味のありそうなイベント情報の詳細情報をひとつひとつ確認し、そこから興味があるのかを判断するのは困難である。

そこで本研究では、ユーザの興味のあるイベント情報との出会いを促すことを目的とする。興味のあるイベント情報をユーザに提示するためにユーザ嗜好に合わせた推薦を行う。本研究では、主に感性工学で用いられるラフ集合を用いてイベント情報を推薦するシステムを構築する。

2. 情報推薦手法

ユーザの興味にあったアイテムを提示するための手法として様々なアルゴリズムが提案されている。ユーザの嗜好を予測するアルゴリズムの最も一般的な手法として、協調フィルタリングとコンテンツベースフィルタリングが挙げられる [3]。以降でそれぞれの手法の先行研究について述べる。

2.1 協調フィルタリング

一般に協調フィルタリングは、多くのユーザ数を確保できなければアイテムを推薦することは難しいとされている。

¹ 公立はこだて未来大学大学院

² 公立はこだて未来大学

^{a)} g211525@fun.ac.jp

^{b)} okuno@fun.ac.jp

澤井らは SNS を利用してユーザ同士の関係や他のユーザの書込みを取得することで協調フィルタリングを適用し、テレビ番組を推薦する手法を提案している [4]。SNS を利用することでユーザ数の問題は解決できるが、推薦アイテムに関連する書込みを必ずしも充分に取得できるとは限らない。本研究で対象としている市民向けのイベント情報に関する書込みを SNS から取得するのは難しい。

2.2 コンテンツベースフィルタリング

吉田耕陽らは、イベント情報の特徴を表すタグとユーザの嗜好を表すタグをそれぞれ定義し、それらのタグの一致度に基づいてイベント情報を推薦する手法を提案している [5]。また、吉田大我らは、各映像に付与されているタグの重要度に基づいてランキングを作成し、そのタグランキングとユーザ嗜好の類似度が高い映像を推薦する手法を提案している [6]。

本研究でもこれらの手法と同様にイベント情報に含まれる詳細情報を利用し、ラフ集合という手法に基づいて推薦を行うシステムの構築を行う。ここでの詳細情報とは、どのような内容のイベントなのかが記述されている部分のことである。

3. ラフ集合を用いた推薦

3.1 ラフ集合

ラフ集合とは、対象の集合を特定できる範囲で情報を荒くすることでその集合にほどよい記述を求める手法である [7]。一般にラフ集合を実行する場合、対象の属性を決め、それらから縮約ルールを用いて特徴を表現する。この方法は、対象の属性からその対象についての知見を得たり、推論を行うために用いられることがある。ラフ集合は一般に属性の種類を増すことで、これまで識別不可能なものが識別可能になる可能性が高くなる。

3.2 ラフ集合の応用

本研究では、ラフ集合を利用してイベント情報の推薦を行う。これまで、ラフ集合を応用し、情報推薦を行った研究がいくつかある。湯元は、ラフ集合を用いて学生向け賃貸物件推薦システムの構築を行っている [8]。この研究では、賃貸物件の家賃や駅から距離、学校からの距離などを属性としてラフ集合を適用している。

また、小見らは、ラフ集合を用いたユーザ嗜好に基づく楽曲推薦手法を提案している [9]。この研究では、楽曲の音圧や振幅の割合、周波数の割合などを属性としてラフ集合を適用している。

これらのように、ラフ集合を応用し、ユーザの嗜好に合ったアイテムを推薦する研究が積極的に行われている。

表 1 決定表

サンプル	属性 1	属性 2	属性 3	属性 4	決定属性
S_1	A	C	E	H	Good
S_2	A	C	F	G	Bad
S_3	B	C	E	H	Bad
S_4	B	D	E	H	Bad
S_5	A	D	F	H	Good

表 2 識別行列

サンプル	S_2	S_3	S_4
S_1	E H	A	A C
S_5	D H	A D F	A F

3.3 縮約ルールの導出

ラフ集合で用いられる縮約ルール作成について説明する。縮約ルールとは、対象を識別するために必要な最小の属性値の部分集合のことである [8]。表 1 の様な決定表があるとすると、サンプル S ($S_1 \sim S_5$) はユーザによって「Good」または「Bad」と評価されたデータ群である。属性 1 は A および B、属性 2 は C および D、属性 3 は E および F、属性 4 は G および H をそれぞれ属性値としてもつものとする。このとき「Bad」と評価されている S_2 、 S_3 および S_4 に対する「Good」と評価されている S_1 および S_5 の独自性をそれぞれ属性値で表現すると表 2 のようになる。

表 2 より、 S_2 、 S_3 および S_4 それぞれに対する S_1 および S_5 の独自性は、表の縦方向に論理和を計算すればよい。さらに、 S_2 、 S_3 および S_4 に対する S_1 かつ S_5 の独自性は、先ほど求めた結果を横方向に論理積を計算することで導くことができる。 S_2 、 S_3 および S_4 に対する S_1 かつ S_5 の独自性を表すのが式 (1) で、これを縮約ルールと呼ぶ。このとき、縮約ルールの属性値の組を極小条件と言う。ただし、式 (1) において表 1 の属性における排他性を考慮して $AB=CD=EF=GH=0$ とする。

$$\begin{aligned}
 & (E + H + D + H)(A + A + D + F)(A + C + A + F) \\
 & = (D + E + H)(A + D + F)(A + C + F) \\
 & = AD + DF + AE + AH + FH
 \end{aligned} \tag{1}$$

3.4 C.I. 値の算出

ラフ集合では、3.3 で導出された縮約ルールから、「Good」と判断したサンプルデータ内に含まれる属性値の割合を C.I 値とする。C.I. 値はどの属性の組み合わせが重要であるかを表す指標である。例えば、式 (1) における AH は S_1 および S_5 のどちらにも一致しているため C.I. 値は 1.0 となる。また、その他の縮約ルールは S_1 または S_5 のどちらかにしか一致していないため C.I. 値は 0.5 となる。情報推

表 3 算出された C.I. 値

極小条件	AD	DF	AE	AH	FH
C.I. 値	0.5	0.5	0.5	1.0	0.5

薦に適用する場合、この C.I. 値が大きいほどユーザの嗜好にあった極小条件と言える。表 3 は、式 (1) から算出された C.I. 値をまとめたものである。

3.5 イベント情報推薦への適用

本研究では、ラフ集合を利用してユーザに興味のあるイベント情報を推薦するシステムの構築を行う。イベント情報を推薦する際にも、関連研究と同様にアイテムの属性と属性値を決定し、縮約ルールの導出および C.I. 値の算出を行う。しかし、前述した賃貸物件や楽曲とは異なり、イベント情報から属性と属性値を決定することは容易ではない。その理由は、イベント情報は様々な種類のジャンルが存在し、一意に属性や属性値を定めるのは困難であるからである。例えば、イベント情報のジャンルとして、「音楽」や「歴史」などが考えられる。「音楽」のイベント情報と「歴史」のイベント情報では、全く異なる内容のイベントであることが多い。一部共通で決定できる属性も存在するが、そのほとんどはジャンルごとに属性を決定しなければユーザの嗜好の予測を行うことは難しいと考えられる。

4. イベント推薦システム

4.1 イベント推薦システムの概要

本研究では、3.5 のような課題を解決するため、推薦対象のイベント情報のジャンルに対してそれぞれ属性とその属性値を定める。システムがイベント情報を推薦するアルゴリズムを以下に示す。

- (1) ユーザに、興味のあるジャンルをいくつか選んでもらう。
- (2) システムは、ユーザが選択したジャンルに関連するイベント情報をランダムに 10 件表示する。
- (3) ユーザは、提示されたイベント情報それぞれに対して「Good」もしくは「Bad」を選択する。
- (4) システムは、選択されたイベント情報をサンプルデータとして縮約ルールの導出する。また、導出された縮約ルールについて C.I. 値を求める。
- (5) システムは、推薦対象のイベント情報で極小条件を満たすものに、その極小条件の C.I. 値を加点していく。
- (6) システムは、C.I. 値が高い極小条件を含むイベント情報から順に推薦する。

4.2 属性と属性値の決定

ユーザが最初に選択したジャンルによって、付与する属性および属性値を変化させる。例えば、ユーザが「音楽」を選択したときに付与する属性とその属性値を表 4 のよう

表 4 「音楽」を例にした場合の属性と属性値

属性	属性値	定性値
参加費	無料	A1
	有料	A2
出演者	プロ	B1
	アマチュア	B2
形態	演奏	C1
	体験	C2

に定義する。また、表 4 のような属性値を論理計算するため全ての属性値に定性値を割り振る。

4.3 対象イベント情報の提示方法

4.1 のアルゴリズムより縮約ルールおよびその C.I. 値を求めることで、イベント情報をどのような順序で推薦するか決定することができる。推薦対象のイベント情報で極小条件を満たすものに C.I. 値を加点していく。その結果、合計の点数が高いイベント情報から順に推薦する。

5. 実験

提案システムの課題を発見するため被験者 1 名に対して実験を行った。実験では、「イベント情報検索サービス HakoEve」のデータを利用する。この Web サイトには表 5 のようなジャンルのイベント情報が登録されている。

5.1 実験方法

本研究の実験は、以下の手順で行う。

- (1) 被験者に好きなジャンルを 1 つ選択してもらう。
- (2) 被験者が選択したジャンルのイベント情報を 5 件提示する。
- (3) 提示されたイベント情報について、被験者はそのイベント情報に興味があれば「Good」、そうでなければ「Bad」を入力する。
- (4) (3) の結果を用いて、縮約ルールと極小条件の C.I. 値を求める。

表 5 イベント情報のジャンル

ジャンル
展示会
趣味・教養
まちづくり
健康
自然・環境
福祉
祭り
福祉
音楽
スポーツ
講演
歴史

表 6 「歴史」の属性と属性値

属性	属性値	定性値
参加費	無料	A1
	有料	A2
歴史舞台	函館	B1
	北海道	B2
	全国	B3
形態	まちあるき	C1
	体験	C2
	講演	C3
	展示	C4
	演劇	C5

表 7 実験で得られた決定表

イベント	属性 1	属性 2	属性 3	決定属性
イベント 1	A1	B2	C3	Bad
イベント 2	A1	B1	C1	Good
イベント 3	A2	B1	C2	Bad
イベント 4	A2	B3	C4	Good
イベント 5	A2	B3	C3	Bad

- (5) 推薦対象のイベント情報で極小条件を満たすものに、その極小条件の C.I. 値を加点していく。
- (6) 最終的に点数が高いイベント情報を 5 件とランダムに選択したイベント情報 5 件を被験者に提示する。
- (7) 提示されたイベント情報に対して被験者にアンケートを行う。

実験で使用するアンケートでは、「提示されたイベント情報についてどのくらい興味がありますか。」という質問をし、興味の度合いを 4 段階の選択肢から該当するものを選んでもらう。また、それぞれの選択肢に対して 4 点, 3 点, 2 点, 1 点の点数を付与する。

5.2 結果

被験者は、好きなジャンルとして「歴史」を選択した。選択された「歴史」の属性と属性値は表 6 のように定義した。表 6 の属性値「北海道」は函館以外の北海道が歴史舞台となっているイベント情報、属性値「全国」は北海道以外の都道府県が歴史舞台となっているイベント情報のことを指す。被験者に「歴史」のジャンルに含まれるイベント情報を 5 件提示した結果、5 件中 2 件を「興味ある」と回答した。これを決定表にまとめると表 7 のようになる。

表 7 の決定表を用いてラフ集合を適用し、得られた縮約ルールの極小条件から C.I. 値を求めた結果、表 8 のようになった。また、極小条件の定性値と属性値の関係を表 9 に示す。これらの結果を用いて、推薦対象のイベント情報で極小条件を満たすものを探し、満たすものにその極小条件の C.I. 値を加点した。その後、被験者に合計の点数が高いイベント情報を 5 件、ランダムに得られたイベント情報 5 件をそれぞれ提示し、アンケートを実施した。アンケート

表 8 実験で算出された C.I. 値

極小条件	A1B1	B1C5	C1	A2C5	A1B3	B3C5
C.I. 値	0.5	0.25	0.5	0.5	0.25	0.5

表 9 極小条件の関係

定性値	属性値
A1B1	有料 函館
B1C4	函館 展示
C1	まちあるき
A2C4	無料 展示
A1B3	有料 講演
B3C5	全国 展示

の選択肢に割り振った配点を合計した結果、提案システムでは「14 点」、ランダムでは「12 点」という結果になった。

5.3 考察

5.2 の結果より、ランダムに提示された場合より提案システムで提示したイベント情報の方が興味を惹いていることがわかった。しかし、属性の数が少なかったため、ユーザ嗜好の分析が不十分であったと考えられる。より多くの属性をイベント情報に対して付与することで精度の高いユーザ嗜好を抽出できると考えられる。しかし、現状の手法では、各ジャンルに対して網羅的に属性を決定することは困難である。今後は、各ジャンルごとに属性を網羅的に決定することができる手法を検討する。

また、今回の実験で得られた極小条件に「無料かつ展示」、「有料かつ講演」という結果があったが、これについても課題がある。「無料かつ展示」とは、言い換えると「無料であれば展示会に興味がある」ということである。さらに、「有料かつ講演」とは、言い換えると「有料であれば講演会に興味がある」ということである。前者の場合は、ユーザの嗜好として充分考えられるが、後者はユーザの嗜好としては明らかに不適切である。「講演会」に興味があるのだから、当然参加費が「無料」であっても興味を持つはずである。この解決策として、極小条件として「A2B3」を追加するか、「A1B3」を「B3」に置き換える必要がある。今後は、この問題についても解決できるような手法の改善を行う。

6. おわりに

本稿では、ラフ集合を用いたイベント情報推薦システムについて説明した。ラフ集合を利用することでユーザの嗜好をラフに分析することができるので、潜在的に興味を持っているイベント情報についても推薦することが可能になる。属性と属性値をイベント情報のジャンルごとに決定することで、イベント情報の推薦にラフ集合を適用することを可能にした。今後は、提案システムを改良し、本研究の有効性を示すため、被験者を増やし再度実験を行う。

参考文献

- [1] 株式会社調和技研：あなた情報マガジンびもーる (online), <http://bemall.jp/sapporo/> (2016.02.22).
- [2] 函館市地域交流まちづくりセンター：イベント検索サービス HakoEve(online), <http://hakomachi.com/hakoeve/events> (2016.02.22).
- [3] 神島敏弘：推薦システムのアルゴリズム, (2014), 入手先 <http://www.kamishima.net/archive/recsysdoc.pdf> (2016.02.22).
- [4] 澤井里枝, 有安香子, 藤沢寛, 金次保明：SNS を利用した協調フィルタリングによる番組推薦手法, 情報処理学会研究報告, Vol. 2010-DBS-151, No. 43, pp. 1-8, 2010
- [5] 吉田耕陽, 遠藤聡志, 當間愛晃, 赤嶺有平, 山田孝治：タグを用いた嗜好抽出による観光イベント推薦システム, 情報処理学会研究報告, Vol. 2010-MPS-80, No. 15, pp. 1-6, 2010
- [6] 吉田大我, 入江豪, 佐藤隆, 小島明：タグランキングに基づく映像推薦, 情報科学技術フォーラム講演論文集, 10, 2, pp. 1-6, 2011
- [7] 森典彦 田中英夫 井上勝雄：ラフ集合と感性一データからの知識獲得と推論, 海文堂出版 (2013).
- [8] 湯本 真樹：条件検索とラフ集合の縮約ルールを利用した学生向け賃貸物件推薦システムの開発, 電学論 C, 133, 4, pp. 740-748, 2013
- [9] 小見裕明, 津谷篤, 田中敦：ラフ集合理論を用いたユーザの嗜好に基づく楽曲推薦手法, 第 75 回全国大会講演論文集, pp. 519-520, 2013