室蘭工業大学
学術資源アーカイブ
Muroran Institute of Technology Academic Resources Archive

関係性マイニングを用いた楽曲推薦手法の提案

メタデータ	言語: Japanese 出版者: 計測自動制御学会 公開日: 2016-01-27 キーワード (Ja): ラフ集合, 関係性マイニング, 楽曲推薦 キーワード (En): 作成者: 比田, 啓貴, 工藤, 康生, 岡田, 吉史, 村井, 哲也 メールアドレス: 所属:
URL	http://hdl.handle.net/10258/3845

室蘭工業大学
学術資源アーカイブ
Muroran Institute of Technology Academic Resources Archive

関係性マイニングを用いた楽曲推薦手法の提案

著者	比田　啓貴，工藤　康生，岡田　吉史，村井　哲也
雑誌名	計測自動制御学会システム・情報部門学術講演会講演論文集
巻	2015
ページ	681-682
発行年	2015-11-18
URL	http://hdl.handle.net/10258/3845

SS4-5

関係性マイニングを用いた楽曲推薦手法の提案

○比田啓貴　工藤康生　岡田吉史（室蘭工業大学）　村井哲也（北海道大学）

研究背景と目的

通常のラフ集合[1)]では、求められる決定ルールは個々の属性とその属性値との組み合わせによる規則に基づいている

2つの属性間の相互関係を新たな属性として決定表に組み込み、決定ルールを求める（関係性マイニング[2)]）

問題点　関係性マイニングは表形式のデータのみを対象としている

目的　関係性マイニングによる系列データからの推薦システムの構築

ラフ集合

対象と属性データを用いた決定表を用いて以下を導き出すデータマイニング手法

- 対象を識別するのに必要最低限の属性の集合(縮約)
- 対象が属するクラスを識別する簡潔なルール（決定ルール）
- 個々の属性とその値の組み合わせに基づく特徴

（例）患者一人の頭痛、鼻水、熱、筋肉痛、流感の決定表$(U, C \cup D, V, \rho)$

◆条件属性 C：頭痛，鼻水，熱，筋肉痛　◆決定属性 D：流感

患者	頭痛	鼻水	体温	筋肉痛	流感
P_1	no	no	high	yes	yes

上記の決定表より患者P_1に対して

[頭痛=no]∧[鼻水=no]∧[熱=high]∧[筋肉痛=yes] ⇒ [流感=yes]

という決定ルールが得られる

ラフ集合による関係性マイニング

1. 決定表を用意する

条件属性　決定属性

対象	c_1	c_2	c_3	d
x_1	1	3	4	1
x_2	2	0	3	1
x_3	4	3	0	0

2. 値の集合上の二項関係を定める

（例）$\{>, <, =, \cdots\}$

3. 相互関係条件属性を作成する

対象	c_1	c_2	$c_1 > c_2$
x_1	3	1	1

4. 相互関係決定表を作成する

条件属性　相互関係条件属性　決定属性

対象	c_1	c_2	c_3	$c_1 > c_2$	$c_1 > c_3$	$\cdots$	d
x_1	3	1	4	1	0		1
x_2	1	2	3	0	0		1
x_3	4	3	1	1	1		0

5. 相互関係決定表に対し縮約・ルール抽出を行い結果を解釈する

（得られる決定ルールの例）

$[c_1 = 3] \wedge [(c_1 > c_2) = 1] \rightarrow [d = 1]$

本研究の手法

● 楽曲データからの決定表作成

楽曲データで決定表を作成するには、条件属性と決定属性を設定する必要がある

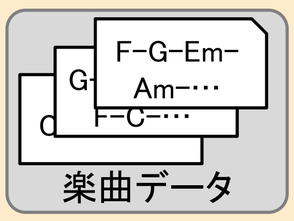
◆ 条件属性の設定

楽曲データの部分系列を条件属性に設定

◆ 決定属性の設定

1. クエリに現れる部分系列の出現頻度が多い順に3種類の属性 a，b，c を選択

2. 選択した属性を全て含むかを表す属性を決定属性として新たに作成



対象	F|G|Em (a)	G|Em|Am (b)	Em|F|C (c)	C|F|G	d
クエリ	5	3	2	1	1
x_1	0	2	4	3	0
x_2	1	2	0	4	0
x_3	4	0	6	0	0
x_4	7	4	1	0	1

決定表

条件属性の値は楽曲中に現れる部分系列の出現頻度

決定属性の値は

$$\begin{cases} 1, & a \neq 0 \text{かつ} b \neq 0 \text{かつ} c \neq 0 \\ 0, & \text{それ以外} \end{cases}$$

● 楽曲推薦手法

1. ユーザがクエリとする楽曲を1曲選択し、本手法を用いて決定表を作成し、更に相互関係決定表を作成する

対象	F|G|Em	G|Em|Am	Em|F|C	C|F|G	F|G|Em > G|Em|Am	F|G|Em > Em|F|C	…	d
クエリ	5	3	2	1	1	1		1
x_1	0	2	4	3	0	0		0
x_2	3	2	0	4	1	0		0
x_3	0	0	6	2	0	0		0
x_4	7	4	1	0	1	1		1

相互関係決定表

2. 並列分散化を用いて縮約計算を行いできるだけ多数の縮約を抽出する

縮約3 評価値：0.05

縮約2 評価値：0.12

縮約1 評価値：0.66667

縮約属性：F|G|Em > G|Em|Am, F|G|Em > Em|F|C

縮約結果

3. 得た縮約で評価値が最も良い縮約を選択

対象	F|G|Em > G|Em|Am	F|G|Em > Em|F|C
クエリ	1	1
x_1	0	0
x_2	1	0
x_3	0	0
x_4	1	1

4. 選択した縮約に含まれる属性を参照し、相互関係決定表から同値類を求める

- クエリ，x_4
- x_1，x_3
- x_2

x_4を推薦

5. 求めた同値類で、クエリを含んだ同値類からクエリ以外の楽曲をユーザに推薦する

実験

- 被験者：1名の大学生
- 対象楽曲：　全270曲(J-POP)
- 楽曲の評価値：5段階評価（1:好まない～5:好む）
- クエリ楽曲：被験者が高評価（評価値5）とした26曲
- 相互関係条件属性に用いる二項関係：{>，<，=}
- 条件属性とするための頻度閾値：25回

実験評価

◆ 推薦楽曲数が10曲未満と10曲以上の場合における、推薦楽曲数・楽曲評価値・縮約評価値の平均値の比較を行った

◆ 本手法による楽曲評価値と楽曲データ全体での楽曲評価値との比較

実験結果

	クエリ	推薦楽曲数	楽曲評価値	縮約評価値
10曲未満	14	4.42	3.40	0.14
10曲以上	4	35.50	3.22	0.39
推薦不能	8	ー	ー	ー

楽曲データ全体での平均評価値
3.05

考察

- 楽曲評価値：10曲未満＞10曲以上
  → 10曲未満の方がユーザの好みの楽曲を推薦する可能性が高い
- 縮約評価値：10曲以上＞10曲未満
  →縮約の評価値が高くなるにつれて、縮約内の属性の種類が減少し、同値類のパターン数が少なくなるため、推薦楽曲数が増加しやすい
- いずれの場合も楽曲データ全体の楽曲評価値より高い値
  →本手法による推薦結果は、ユーザの好みをある程度反映できている
- 推薦不能となった結果の内訳

◆ クエリの同値類にクエリ以外の楽曲が含まれないため推薦を行えない

◆ 決定属性を選択する際に、クエリの頻出部分系列数が3種類未満であったため、決定属性を作成できず推薦を行えない

今後の課題

- 推薦不能を回避する手法の検討
- 他手法との比較実験による本手法の検証

参考文献

1) Pawlak, Z.: Rough Sets - Theoretical Aspects of Reasoning about Data, Kluwer Academic Pubilishers (1991).
2) 工藤康生, 村井哲也: ラフ集合による関係性マイニング の構想, 第29 回ファジィシステムシンポジウム講演会論 文集, 33/36 (2013).