

# 5G-8 履歴データベースを用いたTV番組選択支援エージェントの評価について

中島 太郎\* 渡辺 尚\* 樽口 秀昭\*\*

\* 静岡大学, \*\* ヤマハ株式会社

## 1 はじめに

BS デジタル放送や地上波のデジタル化が始まろうとしている。放送電波がデジタル化されると、デジタル圧縮による多チャンネル化などが生じる [1]。多チャンネル化によってユーザは多くの番組を提供され、その選択肢が広がる。しかし、その一方ユーザは膨大な情報から自分の見たい番組を選択する労力を求められる。そのためにユーザは各ジャンルに点数を付けてエージェントに好みを教えたり [2]、あるいはユーザの手を煩わせず、視聴時間によって番組やその出演者などの評価を決めたり [3] する方式が検討されている。しかしこれらの方式ではユーザが自分の好みが変わるたびにそれをエージェントに教える必要があったり、好みの変化を反映するのが難しかったりする。そこで本稿では選択されなかったと言う情報も加味してユーザに番組を推薦する方法を考える。TV 番組選択支援エージェント (Testa) はユーザが取った行動を履歴として記録し、それをもとにユーザが番組検索のために入力したキーワードの意味を推測して、それを置換や補完して番組表の検索を行なう。このようにしてユーザが見たい番組を探すのを助ける。

## 2 TV 番組選択支援エージェント (Testa)

図 1 に本システムの概観図を示す。

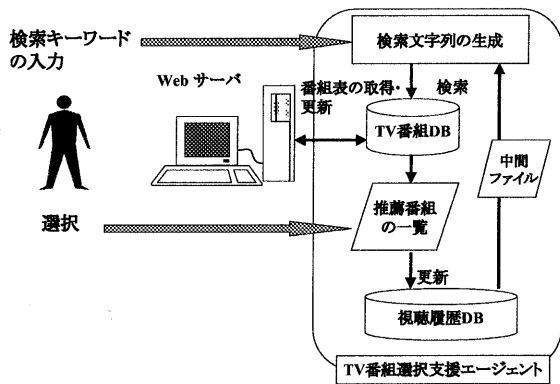


図 1: TV 番組選択支援エージェント (Testa) の概観

ユーザは見たい番組をキーワードを入力して探す。Testa は、

- ユーザが番組検索用に入力したキーワード

An Evaluation of TV Program Selection Support Agent with History Database

\*Taro Nakajima, Takashi Watanabe and \*\*Hideaki Taruguchi  
\*Shizuoka University, \*\*Yamaha Corporation

- ユーザが選択した番組の情報とその推薦順位、その番組評価値
- ユーザが選択しなかった番組のうち、選択した番組より上位で推薦した番組の情報とその推薦順位、番組評価値

を視聴履歴データベースに記録している。ここで、番組の情報というのは日付、曜日、放送開始時間、放送局、ジャンル名、番組タイトル、番組内容 (出演者等) といった、web サーバから得られる情報である。

中間ファイルはユーザの行動を時系列に並べた視聴履歴データベースからユーザが検索用に入力したキーワードに注目し、MBR (Memory-Based Reasoning) [4] を利用して、その時視聴した番組データを検索候補文字列とし、出現頻度などをもとに履歴評価値を算出したものである。これを利用して現在ユーザが入力したキーワード以外に、それらを置換や補完して検索文字列を生成する。

Web から得られた番組表からユーザが見たいと予想される番組を探すために、検索文字列で番組表を検索し、番組評価値が高い順にソートしてその番組の情報をユーザに提示する。ユーザが Testa が提示した番組の中からある番組を選択し、視聴すると、それらの情報を視聴履歴データベースに反映し、更新する。

## 3 システムの評価実験

ユーザが検索用に入力したキーワードに注目すると、現在ユーザが入力したキーワードは図 2 に示すように場合分けできる。

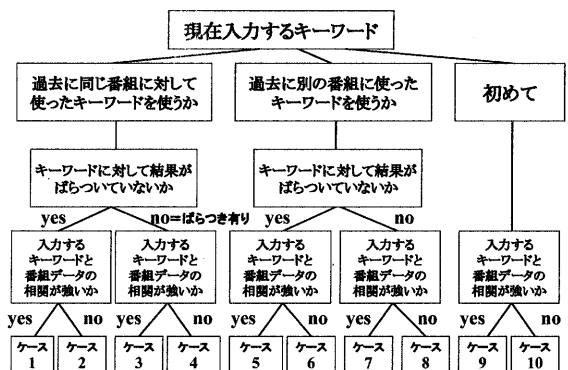


図 2: 現在ユーザが入力したキーワードの場合分け

現在ユーザが入力したキーワードは、まずそのキーワードを過去に使用したことがあるかないかで分類される。さらに過去に使用したことがあるキーワードは、今回探したい番組を探すのに使用したことがあるかないかで分類される。これらは履歴が目的の番組を検索するのに有効に働くか、かえって害を及ぼすか、あるいは役に立たないかの違いによる。

次に過去に使用したキーワードの特徴により分類できる。そのキーワードを入力した際にユーザが視聴した番組の種類が1つであるならば、そのキーワードと番組には強い相関があるといえる。逆にそのキーワードを入力した際に視聴した番組の種類が多ければ多い程そのキーワードと番組を結びつけて考えるのが難しい。

また現在ユーザが入力したキーワードが目的の番組を探すのに有効に働くかどうかで分類できる。

以上の組合せにより現在ユーザが入力したキーワードは10パターンに分類できる。

### 3.1 評価実験

まず3人の被験者に対して3週間分の視聴履歴を収集する実験を行なった。次に収集した視聴履歴を用いて以下のように本方式の評価実験を行なった。

- 最初の2週間分の視聴履歴を視聴履歴データベースの初期データとし、それらは全て1位で推薦され、選択したものと仮定する。
- 最初の2週間分の視聴履歴の各番組に対する番組評価値(Past)を0.3,1.0,3.0,6.0とする。
- 現在ユーザが入力したキーワードの評価値(Now)を中間ファイルの評価値(履歴評価値)に基づいて、
  - 中間ファイル中の各キーワードに対する最大の履歴評価値の平均値(Ave), その1/2倍(Ave/2)及び2倍(Ave\*2)
  - 中間ファイル中の各キーワードに対する最大の履歴評価値の最大値(Max), その1/2倍(Max/2)及び2倍(Max\*2)
  - 定数1.0

の各場合を比較する。

- 現在ユーザが入力したキーワードの評価値と過去の番組評価値の組み合わせにより、最後の1週に視聴した番組が、実際に視聴した日の何番目に推薦されるかを調べる。

#### 3.1.1 実験結果

- ケース1, ケース2  
履歴において既に、現在ユーザが入力したキーワードと目的の番組との間に強い相関がある場合に当たる。現在ユーザが入力したキーワードが有効なケース1の場合はもちろん、ケース2の場合でも希望の順位で推薦できた。
- ケース3  
ケース1と比較すると履歴において、現在ユーザが入力したキーワードと目的の番組との間の相関が弱い場合であるが、出現頻度の高い番組については希望の順位で推薦できた。
- ケース4  
ケース3より目的の番組を検索するのがさらに困難な場合である。現在ユーザが入力したキーワードが当てにならず、履歴に頼る場合であるが、現在ユーザが入力したキーワードと目的の番組の相関もさほど強くない場合になる。この場合はNowを重視し

ないAve/2, Max/2が他に比べて上位で推薦できた。

- ケース5, ケース7  
履歴が目的の番組を検索するのに邪魔になる場合がある。この場合はNowを重視したAve\*2, Max\*2が他に比べて上位で推薦できた。
- ケース6, ケース8, ケース10  
これらのケースでは現在ユーザが入力したキーワードからも過去の履歴からも目的の番組を推測しにくいいため、推薦することが難しい。
- ケース9  
このケースは現在ユーザが入力したキーワードのみによる検索になりやすい。そのため、推薦順位は現在ユーザが入力したキーワードがどれだけ目的の番組を特徴付けられるかによった。

#### 3.1.2 実験のまとめ

以上のように全て1位で視聴したと仮定して初期データを与え、その番組評価値(Past)と現在ユーザが入力したキーワードの評価値(Now)を変えて推薦順位を調べた。すべてのケースについて共通の最適な組み合わせはなかったが、Pastを1.0とし、Nowを中間ファイル中の最大値の平均値(Ave)とした場合が適切であった。

### 4 おわりに

TV番組選択支援エージェント(Testa)は多くの番組の中からユーザが希望する番組を推測し、提示する。そのためにユーザが入力したキーワードや視聴した番組に関するデータだけでなく、選択した番組より上位に推薦した番組に関するデータも視聴履歴データベースに保存しておき、この情報も使ってユーザが入力したキーワードを推測する。

今回の実験では履歴データベースに初期データを与えた後の1回目の推薦順位を調べた。今後は、繰り返し番組を選択する場合の検討を行なう。

### 参考文献

- [1] 金 淵培, “エージェント技術の放送への応用” 映像情報メディア学会誌 Vol.52, No.4, pp.447-451, 1998
- [2] 村崎 康博, 柴田 正啓, 金 淵培, 山田 一郎, “エージェントテレビのための自動番組選択手法の検討” 1998年映像情報メディア学会冬季大会, p83
- [3] 矢川 雄一, 鶴飼 ひろみ, 田中 哲雄, 柴田 正啓, 金 淵培, “個人の嗜好に合ったテレビ番組を自動編成するエージェントの検討” 信学技報 AI98-55, pp9-16, 1998
- [4] C.Stanfill, D.Waltz, “Toward Memory-Based Reasoning” Communication of ACM, vol.29, no.12, pp.1213-1223, 1986