

## 嗜好にあったインターネット動画を提供する エージェントシステムに関する検討

松田 康佑 加藤 誠巳  
(上智大学 理工学部)

### 1. まえがき

近年、インターネット利用者のうち、「動画配信サイト（無料コンテンツ）」の存在を知っている人は 88%、利用経験のある人は 70%を占めている<sup>[1]</sup>など、インターネット動画は普及している。しかし、数多くの動画が存在する中、本当に見たい動画が直ちに検索できるかは疑問である。ユーザの嗜好を把握し、他のユーザの視聴行動をもとにおすすめの動画を提供するシステムがあれば、動画を検索する手間が省けると考えられる。

本稿ではユーザごとに嗜好にあった動画を提供するシステムについて検討を行った。

### 2. システムの概要

#### 2.1 Yahoo!動画

Yahoo!動画は、スポーツや映画、アニメなど様々なカテゴリの動画番組をパソコンで視聴できるサービスであり、10万本以上の動画を提供している。ユーザはその中から、カテゴリ選択やキーワード検索により見たい動画を探すことができる。

このYahoo!動画にも、ユーザの嗜好に合わせたおすすめ動画を提供するサービスはあるが、提供される動画数は少ない。

#### 2.2 RSS データ

RSS とは Rich Site Summary または RDF Site Summary の略で、記事やサイトに対する情報が記述されているデータである。

本システムで使用した RSS データは、Yahoo!動画が提供しているもので、動画のタイトル、URL、公開日、動画の内容詳細、カテゴリが記述されており、URL 以外の情報を一覧にまとめて表示した。ユーザは、一覧の中から見たい動画を選択して視聴する。

### 2.3 システムの流れ

システムの流れを以下の図 1 に示す。

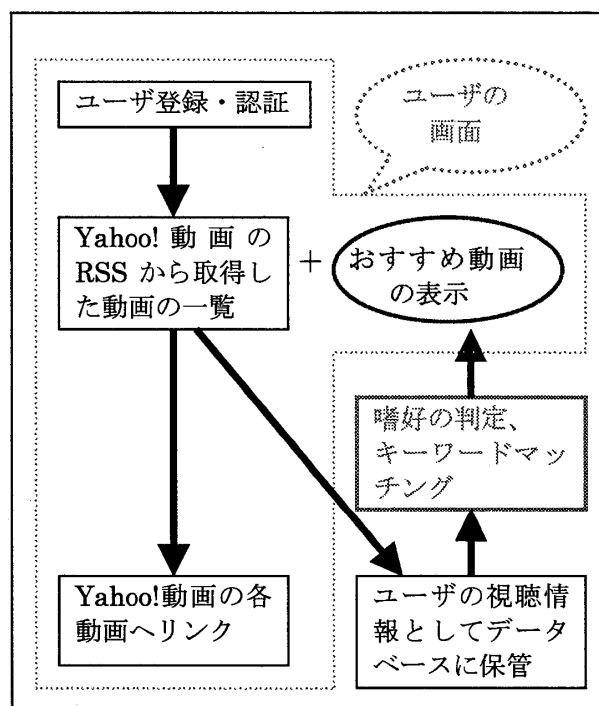


図 1 システムの流れ

ユーザは、ユーザ認証を行ってログインし、Yahoo!動画の RSS から取得した動画の一覧から見たい動画を選択する。これにより、ユーザは選択した動画を視聴できると共に、サーバ側に各ユーザが視聴した番組の情報が送られ、データベースに保管される。

データベースには 3 つのテーブルを用意した。Yahoo!動画の動画一覧を格納したテーブル、ユーザの視聴履歴を格納したテーブル、ユーザの視聴した各カテゴリの割合を数値化したテーブルである。

ユーザの視聴した各カテゴリの割合を数値化したテーブルの結果から嗜好の似たユーザを割り出し、そのユーザが視聴した動画を、動画一覧を格納したテーブルから取り出しておすすめ動画として提供する。

An Agent System for Providing Suitable Internet Video According to One's Preference  
Kosuke MATSUDA, Masami KATO  
Sophia University

## 2.4 嗜好の評価方法<sup>[2]</sup>

あるユーザが見た動画を、嗜好の似た別のユーザにも提供するためには、嗜好が似ているという評価を行う必要がある。評価の方法は、動画のカテゴリごとに行う。それぞれのユーザがどのカテゴリをどれくらいの割合で見ているかを数値化し、嗜好の比較を行う。例として3人のユーザに関する嗜好の評価を以下の図2に示す。

ユーザ	映画	ニュース	お笑い	スポーツ	合計
A	30	10	15	7	62
B	60	13	30	11	114
C	60	6	30	2	98

図2 各ユーザのカテゴリごとの視聴数

ユーザごとに、それぞれのカテゴリにおける視聴数を全体の視聴数で割る。

上の例では、Aさんの場合、映画： $30/62=0.483$   
ニュース： $10/62=0.161$  となる。

この作業をYahoo!動画の全カテゴリ（計10）に関して行い、各ユーザの嗜好を把握する。嗜好の似たユーザが他にいれば、他のユーザが見ている動画をおすすめとして提供する。

視聴数の上位3つのカテゴリが、順番を含めて全て一致し、各カテゴリの視聴数の割合の差がいずれも閾値未満であれば嗜好が似ていると判断する。

上の例の場合、Aさんの上位3つのカテゴリとCさんの上位3つのカテゴリは映画、お笑い、ニュースの順で一致している。さらに、視聴数の割合の差を求めると、映画、お笑い、ニュースの順に0.128、0.101、0.064となる。閾値を0.15に設定すると、いずれも閾値を下回っているため、嗜好が似ていると判断できる。AさんにはCさんが視聴した動画を、CさんにはAさんが視聴した動画を提供する。

## 2.5 キーワードマッチング<sup>[3]</sup>

上記の嗜好の判定だけではユーザの见たい動画を提供できるとは考えにくい。そこで、gooの行っている検索ランキングのデータを用いた。データには、検索ランキングの上位20個のキーワードが記載されており、タイトルや内容詳細にそのキーワードが含まれている動画を提供するようにしている。

データの用語の種類は、女優名、俳優名、流行語といったものがある。

## 2.6 システムの実行例

システムの実行例を以下の図3に示す。

**動画一覧のページ**

**kmatsuさんへのおすすめ動画**

kmatsuさんに似た嗜好の人はこんな動画も見ています

1位「藤井かすみ/チャージャー高沖 1」(秘伝!プロの技!)  
コース:愛知・南愛知カントリークラブ

8位「競輪 KEIRINグランプリ07選考\_決定!」GPレーサー9人、  
競輪界最速のレース「KEIRINグランプリ」! 選ばれしトップレーサー9人が優勝

10位「博多雅丸・不吉」(よしもと滝沢セクシオン)  
博多弁を巧みに操り地元ネタで客席を沸かせる九州の雄!

11位「競輪」(よしもと滝沢セクシオン)  
ご存知「M-1」の常連、「競輪です」!

**他のおすすめ動画は**

1位「女王様 VS キャンパクラ嬢\_第1回」(にしおかすみこ...  
にしおかすみこが、キャンパクラ嬢の素顔に迫る!

5位「「キャンパクラダンゴと乗客」」(島崎みゆきの人物...  
島崎みゆきが箱内で大暴れ!

図3 システムの実行例

## 3. 検討

ユーザの嗜好を把握し、他のユーザの視聴行動をもとにおすすめの動画を提供するシステムについて検討した。しかし、本システムではカテゴリごとの解析、キーワードマッチングのみであり、それだけではユーザの嗜好を捉えにくい。例えば、スポーツの中にも様々なジャンルがあるが、本システムでは区別できていない。ユーザの視聴した番組のタイトルや内容詳細に対して形態素解析を用いるなど、今後は、別の方法も加えて検討する必要がある。

## 4. むすび

Web上に数多くの動画が存在するなかで、ユーザが見たい動画を検索する手間を省くことに注目が集まっている。そのシステムの一つとして、本稿でのシステムの概念が有効に利用されることが期待される。

最後に、有益な御討論を戴いた本学 e-LAB/マルチメディア・ラボの諸氏に謝意を表す。

## 参考文献

- [1] macromill: “インターネット動画サイトの利用実態調査”。
- [2] 倉島, 他: “比較評価情報に基づくランキング手法,” 日本データベース学会 Letters, Vol.6, No.1(2007.6).
- [3] 中原, 加藤: “ナビゲーションシステムでの利用を意図した飲食店に対する評判情報のWebマイニングに関する検討,” 情報処理学会第68回全国大会, 3Q-8 (2006.3).