

Web 広告動的配信システムへのマルコフモデルと kMER の応用

Application of Markov Model and kMER to Dynamic Delivery System in Web Advertising

内野英治[†]森田博彦[†]下野雅芳^{††}

Eiji Uchino

Hirohiko Morita

Masayoshi Shimono

[†]財団法人ファジィシステム研究所^{††}株式会社キューブス

Fuzzy Logic Systems Institute

Cue s Co., Ltd

Abstract: In web advertisement, commercial messages are automatically posted on the monitor of a user. If they meet with the user's interest, not only the user can get to the necessary information immediately but also the effectiveness of the advertisement will increase. The authors have developed a dynamic delivery system in web advertisement. Any prior registration of the user's preference is not necessary. The system we propose here not only grasps the change of user's preference in the long term, but also predicts the user's current preference in the short time. A Markov model and kMER are used to analyze the user's site transfer history. Application to the actual web advertisement system is reported.

1 はじめに

インターネットの普及により、膨大な量の情報が容易に手に入るようになった反面、利用者が必要な情報に辿り着くのに時間がかかるようになった。Web 広告は、閲覧したホームページ上に自動的に掲示されるもので、利用者のニーズと一致すれば、短時間で目的の情報に辿り着くことができる。しかし多くの広告は、利用者の興味を全く無視して強制的に掲示しているだけである。筆者らは、利用者が次に見たい情報を予測し、それに関する広告を配信する Web 広告配信システムを、kMER およびマルコフモデルを応用して実現した。その効果を確認したのでここに報告する。

2 Web 広告配信モデル

ポータルサイトの利用者は、見たいカテゴリのボタンをクリックして情報を検索する。ポータルサーバーはカテゴリのクリック履歴をデータベースに保存し、利用者の興味の変化を解析できる。

2.1 マルコフモデルによる記述

ユーザが次に見るカテゴリは、それまでにユーザが見たカテゴリ、すなわちユーザがそれまでに閲覧したカテゴリの履歴に影響を受ける。カテゴリ間の遷移は式 (1) に示される状態遷移確率行列 P で記述することが出来る

$$P = \{a_{i,j}\} \quad (1)$$

$$a_{i,j} \geq 0, \quad i, j = 1, 2, \dots, m \quad (2)$$

$$\sum_{j=1}^m a_{i,j} = 1, \quad i = 1, 2, \dots, m \quad (3)$$

ここで、 $a_{i,j}$ はカテゴリ i からカテゴリ j へ遷移する確率を表している。 m は全カテゴリ数である。状態遷移確率 $a_{i,j}$ は確率の条件式 (2) と (3) を満たしている。

2.2 マルコフモデルの作成

マルコフモデル [1] を作成するには、カテゴリの閲覧履歴データから状態遷移確率を推定しなければならない。いま、時刻 0 から時刻 T までの状態遷移の履歴を s_0, s_1, \dots, s_T とする。 s_t は時刻 t における状態である。 m を全カテゴリ数、状態 $s_{t-1} = i$ から状態 $s_t = j$ への遷移の回数を $n_{i,j}$ とすると、履歴が s_0, s_1, \dots, s_T となる確率は、

$$a_{s_0, s_1} a_{s_1, s_2} \cdots a_{s_{T-1}, s_T} = \prod_{i=1}^m \prod_{j=1}^m (a_{i,j})^{n_{i,j}} \quad (4)$$

である。ただし、初期状態は与えられているものとする。 $a_{i,j}$ の推定値は最尤推定法により、次のように求めることができる。

$$a_{i,j} = \frac{n_{i,j}}{n_i} \quad (5)$$

2.3 マルコフモデルを用いた広告配信方法

具体的な広告配信方法としては、静的配信モデルと確率的配信モデルが考えられる。

静的配信モデルは、過去に利用者がそれぞれのカテゴリを訪問した回数の累積値に応じた分布により、広告を配信する方法である。確率的配信モデルは 2 つに分けられ、状態遷移確率に基づき、遷移確率が最も高いカテゴリの広告を配信する方法と、遷移確率に基づいて乱数を発生させ、次に遷移すると予想されるカテゴリの広告を配信する方法に分かれる。

3 重み付きマルコフモデル

マルコフモデルを用いた広告配信システムの欠点は、モデルの作成に多くの遷移履歴データを必要とすることである。そこで、モデル作成に要する時間を削減する方法として、重み付きマルコフモデルを用いた広告配信システムを提案する。

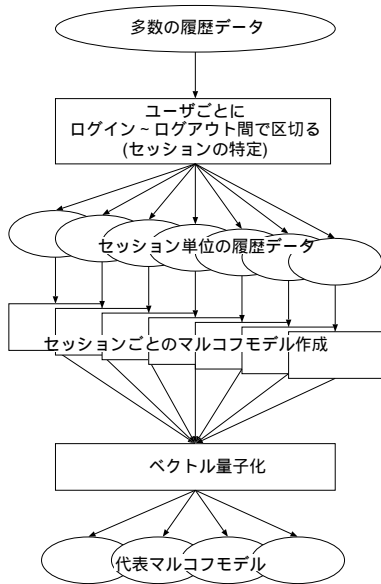


図 1: 代表マルコフモデルの作成手順

重み付きマルコフモデルは、あらかじめ用意しておいた数個のマルコフモデル (代表マルコフモデル) を線形加算することにより、新たなマルコフモデルを作成する手法である。

まず、多数の履歴データ (各ユーザのカテゴリの遷移履歴データ) から各ユーザ毎のマルコフモデルを作成する。作成したモデルの遷移確率行列をベクトル量子化する。代表マルコフモデルの作成手順を図 1 に示す。本研究においては、代表マルコフモデルの作成に用いるベクトル量子化として、kMER [2] を用いた。

ユーザモデルは、ログイン開始からの少ない履歴データで作成された仮のユーザーのマルコフモデルと、各代表マルコフモデルとの類似度を計算し、その類似度に応じた重みで線形加算することにより推定する。

4 広告配信モデルの評価

ユーザが時刻 k において閲覧した Web ページのカテゴリを $C(k)$ 、代表マルコフモデルが予測して配信した広告のカテゴリを $\hat{C}(k)$ として、正解率を以下のように定義する。

$$D = \frac{\sum_{k=1}^N \delta(C(k), \hat{C}(k))}{N} \quad (6)$$

$$\delta(C(k), \hat{C}(k)) = \frac{a_{C(k-1), C(k)}}{a_{C(k-1), \hat{C}(k)}} \quad (7)$$

ここで、 N は広告を配信した最終時刻である。また、 $a_{C(k-1), C(k)}$ は、 $k-1$ 時刻でカテゴリ $C(k-1)$ を閲覧し、 k 時刻でカテゴリ $C(k)$ を閲覧する確率を示す。

個々の評価例として、ユーザ ID57 を例に、正解率が経時変化する様子を図 2 に示す。また、代表マルコフモ

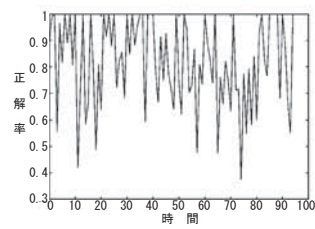


図 2: 正解率の経時変化

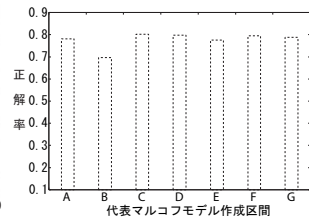


図 3: 代表マルコフモデル作成区間別予測正解率

デル作成時期の影響を調べるため、それぞれ異なる過去の 7 つの区間で代表マルコフモデルを作成し、データの存在する最新の区間で予測正解率を評価した結果を図 3 に示す。図 2 および 3 より、本システムが、70-80% の正解率で、ユーザが次に移るカテゴリを予測しているのがわかる。

5 おわりに

Web 利用者のカテゴリ遷移過程が、マルコフ過程で示されると仮定し、これを Web 広告配信システムに応用した。これにより従来の静的な配信モデルに比べ、ユーザの動きをダイナミックに予測する動的広告配信システムが実現できた。

システム実装時に予想されるサーバーの負荷を軽減するため、個々のユーザのマルコフモデルを、あらかじめ求めた代表マルコフモデルから推定する方法を提案し、その有効性を確かめた。

今後は実システムでの検証データを蓄積し、代表マルコフモデルの数の最適化や、モデルの更新頻度などを検討していく予定である。

6 謝辞

本研究を進めるにあたり、ご協力いただきました石垣忠平氏、玉井康博氏に感謝致します。また本研究は、財団法人福岡県産炭地域振興センターの委託事業として実施されたものである。

参考文献

- [1] Ronald A. Howard, Dynamic Probabilistic Systems Volume I: Markov Models, John Wiley & Sons, 1971.
- [2] Marc M. Van Hulle, Faithful Representations and Topographic Maps, John Wiley & Sons, 2000

連絡先

〒820-0067 飯塚市川津 680-41
財団法人ファジィシステム研究所
内野英治
TEL:0948-24-2771 FAX:0948-24-3002
E-mail: uchino@yamaguchi-u.ac.jp