

配信品質と嗜好を考慮したコンテンツ推薦システム

Content Recommendation System Considering Delivery Quality and Preferences

落合 翔大

上山 憲昭

Shota Ochiai

Noriaki Kamiyama

立命館大学 情報理工学部 情報理工学科

College of Information Science and Engineering, Ritsumeikan University

1. はじめに

動画配信サービスやオンラインショッピングなど、扱うコンテンツ数が膨大なサービスが普及していくにつれ、コンテンツ推薦はサービス運営側とユーザ側の双方に大きな利点がある。一方で配信遅延の増加や低ビットレートはサービス離脱率の上昇につながるため、ネットワークの負荷の軽減が必要となる。そこで多くのコンテンツ事業者が、CDN (Content Delivery Network) の仕組みを利用することで配信遅延やオリジンサーバ負荷の軽減を測っている。しかしユーザの観点ではよりユーザの好みに合ったコンテンツの推薦が望ましいのに対し、ネットワークの観点では配信コスト削減のため高人気コンテンツの推薦が望ましい。そのためこれらの両方を考慮する必要がある。

そこで筆者らは、推薦対象をキャッシュされているコンテンツに限定し、ユーザ間やアイテム間の類似度から推薦するコンテンツを決定する協調フィルタリング (CF) により推薦する、ネットワークの配信サーバの負荷軽減と、ユーザの嗜好の両方を考慮した推薦方式を提案した [1]。しかし推薦対象をキャッシュされているコンテンツに限定することで推薦対象が減少し、推薦の精度が低下する。また静的な環境を想定しているが、実際には動的にユーザやコンテンツが生成され、コンテンツの人気度が増える。そこで本稿では、動的な環境において推薦精度とキャッシュヒット率の両方を向上する強化学習を用いたコンテンツ推薦方式を提案する。そして提案方式の有効性を確認するための計算機シミュレーションによる数値評価結果を示す。

2. 提案方式

動的にユーザやコンテンツが生成され、コンテンツの人気度が増えるような環境では、単一時点での定式化が困難である。そのため、試行錯誤によって最適解を求める強化学習が有効である。また、コールドスタート問題への対策や、強化学習による学習の偏りを少なくするために常に最適な選択を取るのではなく、一定確率でランダムな行動を行う ϵ -greedy 法を用いる。強化学習の手法としては Q 学習を用いる。Q 学習では選択した行動で得られる報酬を最大化させるために価値関数である Q テーブルを更新するアルゴリズムである。行動に対する報酬 R は、推薦したコンテンツの視聴の有無を表す 2 値変数 D 、コンテンツの現在の人気度 P 、人気度の変化量 T 、キャッシュの有無を表す 2 値変数 C 、各々の重み w_i を用いて次式で定義する。

$$R = w_1 D + w_2 P T + w_3 C \quad (1)$$

シミュレーション前に事前学習を行う場合は、ひとりのユーザにひとつのコンテンツを推薦することを 1 エピソードとして、任意の回数学習を行う。

3. 性能評価

提案方式の有効性を計算機シミュレーションにより評価する。シミュレーションで使用するデータセットとして Movielens の各映画の評価やタイトル、ジャンルのデータを用いる [2]。ユーザやコンテンツはシミュレーション中に動的に生成され、コンテンツの人気度はユーザによって生成されるレビューによって変化し続ける環境を想定する。単一のキャッシュサーバのみを考慮して、キャッシュの置換方式としては、キャッシュされているコンテンツの中で最後に要求されてからの経過時間が長いものから置換される LRU (Least Recently Used) を想定する。強化学習を行わず CF のみを利用した推薦システム (CF)、強化学習を行うが事前学習は行わない推薦システム (RL0)、強化学習を行い、100, 000 エピソードの事前学習を行う推薦システム (RL1) の 3 パターンで推薦精度とキャッシュヒット率の比較を行う。こ

で推薦精度とはユーザが生成したレビューの平均値とする。レビューはユーザがコンテンツを視聴した後に、ユーザの好みと現在のコンテンツの人気度より 0 から 5 の値で生成する。

CF と RL0, RL1 のそれぞれの時間経過による平均推薦精度の推移を図 1(a) に示す。ユーザが推薦されたコンテンツに対して生成したレビューの時間経過による推移を示したものである。シミュレーション時間を 10 分割したそれぞれの期間の ARA (Average Recommendation Accuracy) がプロットされている。ARA は分割された期間ごとのレビューの平均値を示す。RL1, RL0, CF の順に推薦精度が高く、提案手法により推薦精度が向上する。また事前学習を行った RL1 では CF, RL0 に比べシミュレーション開始時から高い推薦精度が得られることが確認できる。

強化学習の時間経過における学習レベルの推移を図 1(b) に示す。Q テーブルが更新される回数をシミュレーション時間を 10 分割してプロットしたものである。時間経過とともに減少していく傾向が確認できる。このことから強化学習にかかる時間の影響は時間経過とともに減少すると考えられる。しかし、動的な環境ではコンテンツの人気度が変化するため一定数は残り続けると考えられる。

表 1 は各方式で 10 回シミュレーションを行ったとき、推薦したコンテンツがキャッシュの中にあつた割合を示している。強化学習なしの CF ではキャッシュヒット率が 3.2% 程度であったのに対して、RL0 と RL1 はそれぞれ 11.5%, 15.5% に上昇していることが確認できる。

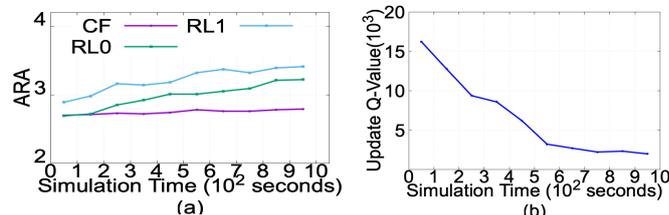


図 1: (a) 推薦精度 (b) 学習レベルの推移

表 1: キャッシュヒット率

CF	RL0	RL1
0.0325	0.1153	0.1545

4. まとめ

動的な環境において、強化学習を用いた推薦精度とキャッシュヒット率の両方の向上させる提案手法の有効性を確認するための計算機シミュレーションによる数値評価結果を示した。今後の研究方針として、本稿では単一のキャッシュサーバのみを考慮したシングルエージェント強化学習での評価を行ったが、複数のキャッシュサーバを考慮したマルチエージェント強化学習での評価を行う予定である。

謝辞 本研究成果は電気通信普及財団研究調査助成を受けたものである。ここに記して謝意を表す。

参考文献

- [1] Yuto Murakami, Yuma Fukagawa, and Noriaki Kamiyama, Content Recommendation Considering Cache State, IEEE HPSR 2024, July 2024
- [2] Movielens, <https://movielens.org/>