

キャッシュ状態と嗜好を考慮したコンテンツ推薦システム

落合 翔大[†] 上山 憲昭[†]

[†] 立命館大学 情報理工学部

〒567-8570 大阪府茨木市岩倉町 2-150

E-mail: †is0609fk@ed.ritsumei.ac.jp, ††kamiaki@fc.ritsumei.ac.jp

あらまし 近年、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するにつれ、コンテンツ推薦システムの重要性が高まっている。同時にコンテンツの配信遅延の増大や低スループットはサービス離脱率の上昇につながるため、多くのコンテンツ事業者は CDN (Content Delivery Network) を利用している。しかし CDN を併用している場合、ユーザとネットワークの観点からは、望ましい推薦対象コンテンツが異なる。そのためコンテンツ推薦において、ユーザの嗜好だけでなく、CDN のキャッシュの状態をも同時に考慮しない場合、配信品質が低下する可能性がある。従来の研究として、これらの両方を考慮したものが存在するが、多くが静的な環境を想定している。そこで本稿では、動的にユーザのコンテンツの嗜好などが変化する環境において、強化学習を用いることで推薦精度とキャッシュヒット率の両方の向上を目指す推薦方式を提案する。また動的な環境においては、静的な環境と比較してコールドスタート問題の影響が大きくなる。そこで、 ϵ -greedy 法を用いることでコールドスタート問題の解消を目指す。そして、提案方式の推薦精度とキャッシュヒット率向上についての有効性を計算機シミュレータで評価する。また、強化学習の学習の重みを変更したときの数値評価と、時間経過における強化学習の学習レベルの推移について評価する。

キーワード 強化学習, キャッシュ, 嗜好, CDN, 協調フィルタリング, ϵ -greedy 法

Content Recommendation System Considering Cache State and Preference

Shota OCHIAI[†] and Noriaki KAMIYAMA[†]

[†] College of Information Science and Engineering, Ritsumeikan University

2-150, Iwakura, Ibaraki, Osaka 567-8570

E-mail: †is0609fk@ed.ritsumei.ac.jp, ††kamiaki@fc.ritsumei.ac.jp

Abstract In recent years, the importance of content recommendation systems has increased as services that handle a huge amount of content, such as video distribution services and online shopping, have become more widespread. At the same time, many content providers use CDNs (Content Delivery Networks) because increased content delivery delays and low throughput lead to an increase in the service dropout rate. However, when using a CDN in combination, the desired content to be recommended differs from the perspective of the user and the network. Therefore, if content recommendation does not take into account not only user preferences but also the state of the CDN cache at the same time, the delivery quality may decrease. There are previous studies that take both of these into account, but most of them assume a static environment. In this paper, we propose a recommendation method that aims to improve both recommendation accuracy and cache hit rate by using reinforcement learning in an environment where users' content preferences change dynamically. In addition, the cold start problem has a greater impact in a dynamic environment than in a static environment. Therefore, we aim to solve the cold start problem by using the ϵ -greedy method. We then evaluate the effectiveness of the proposed method in improving recommendation accuracy and cache hit rate using a computer simulator. We also evaluate the numerical results when changing the learning weights of reinforcement learning, and the transition of the learning level of reinforcement learning over time.

Key words Reinforcement learning, Cache, Preference, CDN, Collaborative filtering, ϵ -greedy method

1. はじめに

近年, YouTube や Netflix などの動画配信サービスや, Amazon などの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急速に普及している. これらのサービスにおいて, サービスを利用するユーザは自分の嗜好に合ったコンテンツを自ら探し出すことは困難になっている. そのため, ユーザの嗜好を推測して, ユーザの嗜好に合ったコンテンツを推薦するコンテンツ推薦システムの重要性が高まっている. ユーザにとっては, 観たい動画や欲しい商品にアクセスでき, サービス運営者にとっては売上や広告収入やテナント料などの収益増加につながる [1]. そのため, コンテンツ推薦システムはユーザ側とサービス運営側の双方に大きな利点がある. しかし, コンテンツの配信遅延の増加や低スループットはサービスの離脱率上昇につながる. そこで, 多くのコンテンツ事業者が, CDN (Content Delivery Network) を利用している. CDN の動作を図 1 に示す.

CDN は地理的に分散した場所に多数のキャッシュサーバを配置し, オリジンサーバに収容されたコンテンツのオリジナルのコピーをキャッシュする. そしてユーザからの要求に対して, 最も要求ユーザに近いと推定されるキャッシュサーバからコンテンツを配信する. そのため, 要求を行うユーザとオリジンサーバの距離に関わらず, 安定したコンテンツ配信が可能になる. またトラフィックが複数のキャッシュサーバに分散されることで, ネットワーク負荷の分散にもつながる. 一方で, CDN の仕組みを利用したコンテンツ推薦システムにおいては, ユーザの観点からは, より嗜好に合ったコンテンツの推薦が望ましいのに対して, ネットワークの観点からは, 配信コスト削減のために, より高人気なコンテンツの推薦が望ましいという違いがある. そのため, これらの両方を考慮しないでコンテンツを推薦すると, キャッシュや推薦システムの性能が低下する可能性がある.

そこで筆者らは, 推薦対象をキャッシュサーバに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により推薦を行うことで, 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 [2]. 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ため, ユーザの嗜好に合ったコンテンツがキャッシュされていない場合には, 推薦精度が低下する. また, サービスを利用するユーザの数やコンテンツの数, コンテンツの人気度が変化しない静的な環境が想定されている.

そこで本稿では, ユーザの数やコンテンツの数, コンテンツの人気度が変化し続けるような動的な環境において, 推薦対象をキャッシュコンテンツに限定することなく, ユーザの嗜好とキャッシュの状態を考慮することで, 推薦精度とキャッシュヒット率の両方の向上を目指す推薦方式を提案する. しかし, 動的な環境では単一時点での最適化が困難であるため, 変化し続ける環境に対して, 試行錯誤による学習で最適な解を見つけていく強化学習を適用する.

また, コンテンツ推薦システムにおいて新しいシステムやサービスが十分なデータを持っていないために, 適切な予測や推薦ができないコールドスタート問題がある. 動的な環境においては, ユーザやコンテンツの数, コンテンツ人気度が変化し続ける. そのため動的な環境においては静的な環境と比較して, コールドスタート問題の影響が大きくなる. そのため, ϵ -greedy 法を用いて確率 ϵ でランダムなコンテンツを推薦することにより, 推薦されるコンテンツの偏りを低減し, コールドスタート問題の解消を図る.

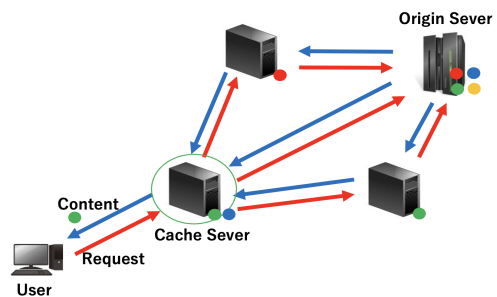


図 1 CDN の仕組みを利用したコンテンツ配信

2. 関連研究

2.1 キャッシュ状態を考慮したコンテンツ推薦システム

CDN の仕組みを利用したコンテンツ推薦システムでは, ユーザやネットワークの観点では推薦対象として望ましいコンテンツが異なる. そのため, これらの両方を考慮しない場合, キャッシュや推薦システムの性能が低下する可能性がある. [2] では, 推薦対象をキャッシュサーバにキャッシュされているコンテンツに限定した推薦方式が提案された. キャッシュされているコンテンツは一般的に配信回数が多く, 人気度が高いため多くのユーザから高評価を得ることが期待できる. また, 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キャッシュヒット率の低下を回避することが可能になる. [3] では, キャッシュをプロアクティブキャッシュとリアクティブキャッシュの 2 つに分割することで改良した LRU アルゴリズムを提案している. [4] では, 静的な環境で最適化問題としてキャッシュヒット率の向上を確認した.

2.2 動的な環境におけるコンテンツ推薦システム

動的な環境での推薦システムでは, コンテンツ人気度やユーザの嗜好の変化を考慮する必要がある. [5] では, 時間経過によるユーザの嗜好の変化点を見つけることでユーザの嗜好を考慮したコンテンツ推薦システムを提案している. 対象ユーザの過去データのエン트로ピーを計算し, 現在の好みと類似する期間のみのデータを推薦に用いる. その結果, より現在の好みに合ったコンテンツの推薦と, 推薦に要する計算量の低減を可能にしている. [6] では, 動的な環境において DQN (Deep Q-Network) を用いた強化学習により推薦精度の向上を確認した. また [7] では, [6] の拡張を行い, 複数種類のデータから学習を行うことを可能にし, 推薦精度の向上や現実世界への適用が可能になる方式を提案している.

2.3 コールドスタート問題への対策

コンテンツ推薦システムにおいて、ユーザに対してコンテンツの推薦を行う際に、ユーザやコンテンツの過去のデータが少なく、ユーザの嗜好やコンテンツの特徴に基づく推測が困難になり、推薦の精度が低下するコールドスタート問題が課題となる。[8]では、新しく追加されたコンテンツを既存のアイテムとペアリングすることでコールドスタート問題を解決している。深層学習を用いたコンテンツのペアリングによって、新しいコンテンツと既存のアイテムをひとつのコンテンツとみなし、既存のコンテンツが推薦されると新しいコンテンツも推薦が行われるようになる。[9]では協調フィルタリングを用いた推薦に時間的重みやk-means法を適用することでコールドスタート問題を解決している。

3. 提案方式

3.1 提案方式の概要

ユーザやコンテンツが動的に生成され、コンテンツ人気度が変化し続けるような環境では、試行錯誤による学習で最適な解を見つける強化学習が有効的と考えられる。そのため、ユーザに対してコンテンツとなる映画を推薦するような推薦システムを想定して、強化学習を用いることで推薦精度とキャッシュヒット率の両方を目指す。強化学習の手法として、逐次的に行動に対する報酬を付与して学習を行うQ学習ベースのものをを用いる。また、学習はユーザにコンテンツを推薦したときにユーザがコンテンツに対して生成する評価値の予測値のテーブルを作成し、更新していくことで行う。

3.2 ϵ -greedy 法によるコンテンツ選択方法の決定

訪問ユーザ (コンテンツ要求ユーザ) に対して、 ϵ -greedy 法を用いて、最適なコンテンツを推薦するか、ランダムなコンテンツを推薦するかを決定する。図2に ϵ -greedy 法の動作を示す。

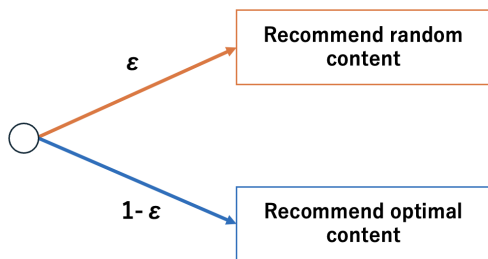


図2 ϵ -greedy 法

ϵ -greedy 法を用いることで、確率 ϵ でランダムなコンテンツが推薦される。そのため、強化学習の偏りを減少させることが可能になる。また、生成されてすぐのコンテンツでも推薦対象となるため、コールドスタート問題の解消にもつながる。

3.3 協調フィルタリングによるユーザの嗜好の推測

最適なコンテンツの推薦を行うためにユーザの嗜好の推測

が必要となる。提案手法では、協調フィルタリングの仕組みを利用することでユーザの嗜好を推測する。協調フィルタリングの動作例を図3に示す。図3は、訪問ユーザであるユーザXに対してコンテンツ1から10の中で最もユーザの嗜好に合っているものを、ユーザXの過去データと他のユーザA, B, C, D, Eの5人分の過去データとの類似度を相関係数を計算することで選択する動作の例となる。「-」は未評価を表し、各コンテンツに対する評価値は0から5の値で評価を行う。

Target user	content	1	2	3	4	5	6	7	8	9	10	Correlation coefficient
X		-	4	2	-	-	-	-	1	1	5	1.000
A		5	3	4	-	-	-	-	1	2	1	0.063
B		-	-	-	1	1	2	4	5	4	0	-0.982
C		2	3	2	1	-	4	3	2	1	4	0.941
D		2	-	-	2	4	5	1	2	5	5	0.500
E		-	5	0	-	5	1	-	1	1	2	0.615
Degree of recommend		2.00	4.00	1.00	1.50	4.50	3.33	2.00	1.66	2.33	3.66	

図3 協調フィルタリング

1. 訪問ユーザ x と A から E の各ユーザ y の類似度を表す相関係数 P_{xy} を計算する。相関係数は-1から1の範囲で値を取り、1に近いほど正の相関がありユーザ x と y の類似度が高いことを示し、反対に-1に近いほど負の相関がありユーザ x と y の類似度が低いことを示す。相関係数を求める計算式を(1)に示す。 C_{xy} はユーザ x と y の両方が評価値を生成しているコンテンツの集合、 r_{nm} はユーザ n がコンテンツ m に対して生成した評価値、 R_n はユーザ n が C_{xy} に対して生成した評価値の平均値を表す。

$$P_{xy} = \frac{\sum_{m \in C_{xy}} (r_{xm} - R_x)(r_{ym} - R_y)}{\sqrt{\sum_{m \in C_{xy}} (r_{xm} - R_x)^2} \sqrt{\sum_{m \in C_{xy}} (r_{ym} - R_y)^2}} \quad (1)$$

2. 1で計算した相関係数が高い上位 k 人を選択する。 k は任意の数をとるが、今回の動作例では $k = 3$ とする。選択した相関係数が高い上位 k 人が生成した、各コンテンツに対する評価値の平均を、推薦するコンテンツの指標となる推薦度として計算する。

3. 計算した推薦度が最大のコンテンツを推薦する。

3.4 最適なコンテンツの推薦

強化学習を用いて最適なコンテンツを推薦するために学習テーブルを作成する。作成するテーブルでは、コンテンツ i を訪問ユーザ x に対して推薦したときに、ユーザ x により生成される i の評価値を予測した値を管理する。

協調フィルタリングで得られたユーザ間の類似度を表す相関係数 P_{xy} を用いて、訪問ユーザ x と類似ユーザ y との類似度を S_{xy} とする。協調フィルタリングで得られた訪問ユーザ x と正の相関をもつユーザを類似ユーザとして利用する。類似ユーザ y のコンテンツ i に対する評価値を E_{yi} とすると、加重平均を用いることでユーザ x がコンテンツ i に対する評価値の予測値が導出できる。しかし、コンテンツ人気度が変化し続ける環境においては、現在の人気度を考慮する必要がある。

ある．動的な環境に適用するために，推薦するコンテンツの現在の人気度と，コンテンツ人気度の変化量を用いて拡張する．コンテンツ i の人気度の変化量 T_i を，現在の人気度 P_i ，1 つ前の人気度 $P^{pre1}(i)$ ，2 つ前の人気度 $P^{pre2}(i)$ を用いて式 (2) で定義する．

$$T_i = (P^{pre1}(i) - P^{pre2}(i)) + (P(i) - P^{pre1}(i)) \quad (2)$$

これらの変数を用いて，コンテンツ人気度の変化を考慮した，訪問ユーザ x に対してコンテンツ i を推薦したときの，ユーザ x のコンテンツ i に対する評価値の予測値 $V(x, i)$ を式 (3) で与える．

$$V(x, i) = \frac{\sum_k S_{ki} \cdot E_{ki}}{\sum_k S_{ki}} + P_i + T_i \quad (3)$$

推薦したコンテンツをユーザが実際に視聴したか否か，またキャッシュの中にあったコンテンツか否か，そして得られたレビューの値によって，強化学習の報酬を付与する．すなわち報酬 R を，推薦したコンテンツの視聴の有無を表す 2 値変数 D ，コンテンツに対する評価値として得られたレビューの値 E ，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，式 (4) で与える．

$$R = w_1 D + w_2 E + w_3 C \quad (4)$$

重みの基準値としては，推薦したコンテンツが視聴されるかどうかを最も重要視し，ユーザから得られる評価値と，キャッシュ状態に対する学習を同等程度の割合で学習することを想定する．2 値変数である D と C は，0 または 1 をとり，評価値 E は 0 以上 5 以下の範囲で生成される．そのため，それぞれの重みの値は， $w_1 = 10$ ， $w_2 = 1$ ， $w_3 = 2.5$ とする．

学習テーブルの更新式を報酬 R ，予想される Q 値の最大値 Q_{max} と 2 つのパラメタ学習率 α と割引率 γ を用いて，以下の式 (5) で定義する．学習率 α は新しい情報をどの程度反映するかを決定するパラメタで，割引率 γ は将来の報酬を現在の報酬と比較する際に，どの程度重視するかを決定するパラメタである．2 つのパラメタは 0 から 1 の値をとり，学習率が 1 に近いほど新しい情報を反映し，割引率が 1 に近いほど長期的な利益を重視する．

$$Q_{xy} = Q_{xy} + \alpha(R + \gamma Q_{max} - Q_{xy}) \quad (5)$$

以上の動作を繰り返し，価値関数を最大化させる行動を選択することで学習レベルを高め，推薦精度とキャッシュヒット率の両方の向上を目指す．しかし，推薦するユーザが過去に視聴したコンテンツの数が少ない場合は，嗜好の推測ができない．そのため，協調フィルタリングで類似するユーザが選択できない場合は， ϵ -greedy 法の結果に関わらずランダムなコンテンツを推薦する．その後，協調フィルタリングで類似するユーザを選択して，最も類似度が高いユーザの学習テーブルを引き継ぐ．

3.5 事前学習

初期段階でのユーザの嗜好の予測をできるだけ速やかに行

うため，事前学習を行う．事前学習では，推薦システムがシミュレーション開始時に初期段階で持っているコンテンツに対しての学習テーブルを作成する．学習に用いるユーザとしては，シミュレーションとは別のユーザを利用する必要がある．

4. 性能評価

4.1 性能評価条件

提案方式の有効性を検証するために，計算機シミュレーションを用いた数値評価を行う．評価に用いたパラメタ設定値を表 1 に示す．

表 1 評価に用いたパラメタ設定値

記号	定義	値
C	キャッシュ容量	50
$N_{content}$	総コンテンツ数	9,742
N_{user}	総ユーザ数	610
$n_{content}$	シミュレーション開始時のコンテンツ数	300
n_{user}	シミュレーション開始時のユーザ数	100
T	シミュレーション時間	1,000 (秒)
$I_{request}$	要求発生間隔	10.7 (ミリ秒)

簡略化のため単一のキャッシュサーバのみを考慮する．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最も長いものから削除する LRU (Least Recently Used) を想定する．データセットとしては MovieLens の映画のタイトル，ジャンル，各映画に対する評価をデータセットとして使用する [10]．データセットは，映画推薦システムの研究や開発に広く用いられており，リアルなユーザの分析に適している．

従来手法として協調フィルタリングのみを用いた推薦システム (CF) と，シミュレーション開始時から強化学習を用いて推薦精度とキャッシュヒット率の向上のための学習を行う RL0，RL0 に加えてシミュレーション開始前に事前学習として 10,000 エピソードの学習を行った RL1 の 3 つのシナリオで比較する．ユーザは推薦されたコンテンツを視聴するかどうかをユーザの嗜好とコンテンツの人気度，人気度の変化量から判断する．ユーザの嗜好のモデルは，[2] で提案したユーザのコンテンツに対する嗜好度のモデルを，コンテンツ人気度と人気度の変化量を用いてコンテンツの人気度が変化する環境に対応するように拡張する．ユーザ x とコンテンツ i の類似度を a_{xi} としたとき，ユーザ x のコンテンツ i に対する嗜好度 $P_x^{pref}(i)$ はコンテンツ i の現在の人気度 P_i ，コンテンツ i の人気度の変化量 T_i を用いて式 (6) で与える．

$$P_x^{pref}(i) = \frac{a_{xi}}{\sqrt{\sum_{i \in I} a_{xi}}} + P_i + T_i \quad (6)$$

ユーザは推薦されたコンテンツに対する嗜好度が閾値を超えていた場合，コンテンツを視聴する．本稿における視聴の判断を行う閾値は，推薦されたコンテンツが，全てのコンテンツの中でユーザの嗜好度が高い上位 40% 以上のコンテン

ツであれば視聴するように設定する。また、ユーザは推薦されたコンテンツに対する嗜好度を 0 から 5 の値でレビューとして生成する。生成されたレビューの値によってコンテンツ人気度を変化させる。コンテンツ人気度はコンテンツに対して生成されたレビューの平均値とする。コンテンツ i の現在の人気度 P_i の計算式はコンテンツ x に対して生成された n 個のレビュー R_x を用いて式 (7) で与える。

$$P_i = \frac{\sum_{k=1}^n R_k}{5n} \quad (7)$$

生成されるレビューは 0 から 5 の値で評価を行うのに対して、コンテンツ人気度は 0 から 1 の値をとるため分母に 5 を掛けて数値を調整する。

4.2 推薦精度とキャッシュヒット率の推移

本稿における推薦精度は、コンテンツの推薦を受けたユーザが、推薦されたコンテンツに対して生成した評価値と定義する。またキャッシュヒット率は、推薦したコンテンツがキャッシュの中にあった割合と定義する。従来手法である CF、提案手法である RL0 と RL1 の推薦精度の比較を図 4 に、キャッシュヒット率の比較を表 2 に示す。

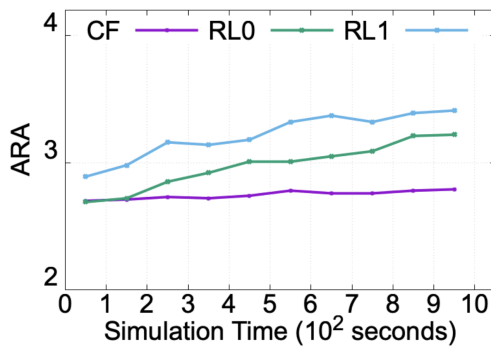


図 4 時間経過による推薦精度の推移

表 2 キャッシュヒット率

CF	RL0	RL1
0.0325	0.1153	0.1545

図 4 は、横軸をシミュレーション時間とし、縦軸を ARA (Average Recommendation Accuracy) である。ARA は、シミュレーション時間を 10 等分したときの各期間における推薦精度の平均値として定義する。比較対象である協調フィルタリングのみの推薦システム CF は、時間経過とともに僅かに推薦精度が向上するが、精度の向上は小さい。これは、協調フィルタリングは過去のデータから推薦するコンテンツを決定するため、コンテンツ人気度が動的に変化するような環境での現在の人気度が反映できないからだと思う。それに対して、提案手法である RL0 と RL1 は時間経過とともに推薦精度が向上していくことが確認できる。事前学習を行う RL1 は事前学習を行わない RL0 に比べてシミュレーション開始時から推薦精度の値が高くなっている。これは事前学習

によりシミュレーション開始時のコールドスタート問題が解消し、ユーザの好みの推測ができていていることを示す。しかし、1,000 秒のシミュレーション期間で RL0 は、RL1 よりもシミュレーション開始時からの向上の変化量が大きいことが確認できる。そのため、シミュレーション時間を長くしたときにどのような値となるのか検証を今後行う予定である。

表 2 から提案手法である RL0 と RL1 のキャッシュヒット率はそれぞれ約 12%, 約 15% となり、どちらも従来手法である CF の約 3% と比較して、キャッシュヒット率が高いことが確認できる。

4.3 学習の重みを変更した時の性能評価

提案手法の動作で示した学習における報酬の式 (4) の重みの値である w_1 , w_2 , w_3 を変更することで強化学習での学習の比率を変更することが可能になる。比較対象としては、推薦精度をキャッシュヒット率の 2 倍程度学習を行う A2C1 と、1 対 1 程度の割合で行う A1C1 と、キャッシュヒット率を推薦精度の 2 倍程度学習を行う A1C2 の 3 つを考える。それぞれのシナリオにおける重みの設定値を表 3 に示す。

表 3 重みの設定値

シナリオ	w_1	w_2	w_3
A1C2	10	1	5
A1C1	10	1	2.5
A2C1	10	2	2.5

重みを変更したときの推薦制度の推移を図 5、キャッシュヒット率の結果を表 4 に示す。

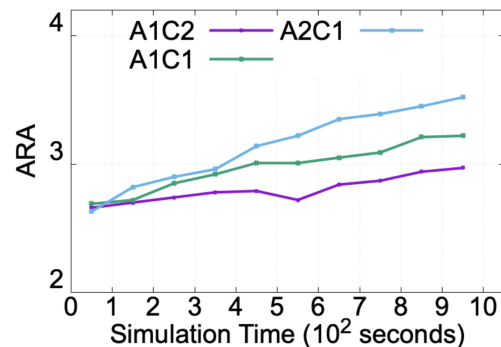


図 5 重みを変更したときの推薦精度の推移

表 4 重みを変更したときのキャッシュヒット率

A1C2	A1C1	A2C1
0.1723	0.1153	0.0576

図 5 と表 4 の結果から、重みを変更することで向上させる割合を調節できることが確認できる。また、学習比率が低いときでも従来手法の協調フィルタリングのみを用いた推薦システムと比較して、推薦精度とキャッシュヒット率の両方が

向上していることが確認できる。これは、動的に変化している環境のなかで推薦を行う時点での各ユーザの評価とコンテンツ人気度変化の予測が可能になっているからと考えられる。

4.4 時間経過による学習レベルの推移

強化学習を用いる提案手法では、報酬の最大化のために価値関数を更新する必要がある。そのため、価値関数の更新にかかる時間によってコンテンツ推薦に遅延が発生する可能性も考えられる。そこで提案手法において価値関数が更新される回数の時間経過による推移を図6で示す。横軸をシミュレーションの経過時間とし、縦軸はシミュレーション時間を10等分したときのそれぞれの期間で、価値関数が更新された回数を表す。

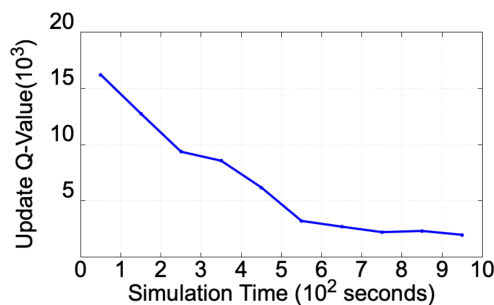


図6 時間経過による強化学習の学習レベル

シミュレーション開始時は試行錯誤しながら学習を行うために価値関数の更新回数が多いことが確認できる。しかし、時間経過とともに強化学習の学習レベルが高くなり、価値関数が更新される回数が減少することが確認できる。そのため価値関数の更新による遅延の影響は時間経過とともに少なくなることが考えられる。一方で、コンテンツ人気度が変化し続ける環境においては価値関数の更新回数は一定数残り続けると考えられる。

5. ま と め

コンテンツ数が膨大なサービスの普及に伴い、コンテンツ推薦システムの重要性が高まっている。一方で、多くのコンテンツ事業者は配信品質の向上のため、CDNを用いてコンテンツを配信している。そのためコンテンツ推薦においてキャッシュ状態を考慮しない場合、キャッシュヒット率が低下する可能性がある。また、既存のコンテンツ推薦システムの研究の多くは静的な環境が想定されている。そこで本稿では、動的な環境で強化学習を用いることで、推薦精度とキャッシュヒット率の両方の向上を目指すコンテンツ推薦方式を提案した。そして計算機シミュレーションの数値評価により協調フィルタリングのみの推薦システムと比較して、推薦精度とキャッシュヒット率の両方が向上することを確認した。また、強化学習の学習レベルの向上により、強化学習で発生する遅延の影響は時間経過とともに低減することを確認した。

今後は、複数のキャッシュサーバを、自律的に推薦コンテ

ンツを決定するエージェントとみなしたマルチエージェント強化学習を用いた推薦方式を検討する。そして複数のキャッシュサーバを考慮して動的にコンテンツの需要が変化する場合の推薦精度とキャッシュヒット率の数値評価を行う予定である。また強化学習の性能向上のため、Q学習からニューラルネットワークを用いるDQN (Deep Q-Network) への拡張を行う予定である。

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。

文 献

- [1] Z.Tufekci, "Youtube,the great radicalizer, The New York Times,2018.
- [2] Yuto Murakami, Yuma Fukagawa, and Noriaki Kamiyama, Content Recommendation Considering Cache State, IEEE HPSR 2024, July 2024
- [3] X.Yang, H.Zhang, H.Ji, and X.Li, Hybrid Cooperative Caching Based IoT Network Considering the Data Cold Start, IEEE Wireless Communications and Networking Conference (WCNC) 2021
- [4] Livia Elena Chatzieleftheriou, Merkouris Karaliopoulos, and Iordanis Koutsopoulos, Jointly Optimizing Content Caching and Recommendations in Small Cell Networks, IEEE Transactions on Mobile Computing 2019
- [5] Thitiporn Neammanee, and Saranya Maneeroj, Time-Aware Recommendation Based on User Preference Driven, 2018 42nd IEEE International Conference on Computer Software & Applications
- [6] Guanjie Zheng, Fuzheng Zhang, Zihan Zheng, Yang Xiang, Nicholas Jing Yuan, Xing Xie, and Zhenhui Li, DRN: A Deep Reinforcement Learning Framework for News Recommendation, WWW 2018, April 23-27, 2018, Lyon, France
- [7] Nan Guo, Zhangpeng Fu, and Qihui Zhao, Multimodal News Recommendation Based on Deep Reinforcement Learning, 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)
- [8] Jianbo Yuan, Walid Shalaby, Mohammed Korayem, David Lin, Khalifeh Aljadda, and Jiebo Luo, Solving Cold-Start Problem in Large-scale Recommendation Engines: A Deep Learning Approach, 2016 IEEE International Conference on Big Data (Big Data)
- [9] Yecheng Lao, and Kaixuan Lin, Hybrid recommendation algorithm based on K-means++ and dynamic time weights, 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE)
- [10] MovieLens, <https://movielens.org/>