

ユーザの嗜好に基づくキャッシュ制御の一検討

北村 司[†] 上山 憲昭^{††} 山本 幹^{†††}

[†] 関西大学大学院理工学研究科 〒564-8680 大阪府吹田市山手町 3-3-35

^{††} 福岡大学工学部 〒814-0180 福岡県福岡市城南区七隈 8 丁目 19-1

^{†††} 関西大学システム理工学部 〒564-8680 大阪府吹田市山手町 3-3-35

E-mail: [†]k709267@kansai-u.ac.jp, ^{††}kamiyama@fukuoka-u.ac.jp, ^{†††}yama-m@kansai-u.ac.jp

あらまし 近年、インターネットトラフィックの大部分は YouTube や NetFlix などの動画トラフィックが占めている。そこで、オリジナルサーバに存在する動画の複製をユーザ近隣の複製サーバに配置することは、ネットワーク内のトラフィック負荷を軽減するための一般的な解決策である。従来の方式では、複製するコンテンツを決定する際、コンテンツの人気度の時間推移や地域の特徴に着目していた。本稿では、ユーザの動画コンテンツに対する嗜好に着目し、ユーザの嗜好に基づいたキャッシュ決定ポリシーを提案する。類似した嗜好を持つユーザは、類似した動画コンテンツを要求する傾向があるため、我々はユーザ間の嗜好の関係性を利用することによって、近い将来に要求される動画コンテンツの予測を行う。MovieLens の動画評価データを用いた数値評価により、ユーザを嗜好の傾向に基づきクラスターに分類できることと、提案する嗜好の傾向に基づくキャッシュ制御法により LRU と比較してキャッシュヒット率を改善することの 2 つを明らかにする。

キーワード キャッシュ判断法, ユーザ嗜好, 機械学習

A Study on Cache Decision Policy using User Tastes

Tsukasa KITAMURA[†], Noriaki KAMIYAMA^{††}, and Miki YAMAMOTO^{†††}

[†] Graduate School of Science and Engineering, Kansai University 3-3-35 Yamate-cho, Suita-shi, Osaka, 564-8680, Japan

^{††} Faculty of Engineering, Fukuoka University 8-19-1 Nanakuma, Jonan-ku, Fukuoka, 814-0180, Japan

^{†††} Faculty of Engineering Science, Kansai University 3-3-35 Yamate-cho, Suita-shi, Osaka, 564-8680, Japan

E-mail: [†]k709267@kansai-u.ac.jp, ^{††}kamiyama@fukuoka-u.ac.jp, ^{†††}yama-m@kansai-u.ac.jp

Abstract Today, the majority of internet traffic is video traffic such as YouTube and NetFlix. Placing copy of original video contents to replication servers nearby users is a general solution to reduce traffic load in networks. In existing research, they focus on the spatial locality and the transition of video contents. In this paper, we focus on user tastes for video contents and propose a new cache decision policy based on measured user tastes. Users which have similar tastes tend to require similar video contents, thus we utilize the measured relation among users' requests to predict near future video content requests. Our evaluation results using MovieLens dataset show that our proposal can make clusters of users who have similar tastes and improve cache hit ratio in comparison to LRU.

Key words Cache Decision Policy, User Tastes, Machine Learning

1. ま え が き

近年、ネットワーク内で流通するトラフィック量は増加の一途を辿り、全世界におけるトラフィックは 2016 年の時点で年間 1.2 ゼタバイトに達し、2021 年までに 3.3 ゼタバイトに達すると予想されている [1]。また 2016 年の時点で、動画コンテンツがインターネット内に流通する全トラフィックの 7 割を占め、2021 年までに 8 割以上を占めると予想されている。この背景として、YouTube [2] や NetFlix [3] といったコンテンツプロバイダのコ

ンテンツ配信におけるトラフィックの増加がある。インターネット内に流通するトラフィック量の増加に対処するために、オリジナルサーバのコンテンツ（オリジナルコンテンツ）を複製して保持・提供を行うサーバ（キャッシュサーバ）を配備する方法が存在する。キャッシュサーバの配備は、ネットワークとユーザの双方にとって利点がある [4]。まず、ネットワークの観点では、ユーザのトラフィック需要がユーザ近隣のキャッシュサーバで満足されることにより、ネットワーク内に流通するトラフィックが減少し、輻輳の発生が軽減される。また、ユーザの観点で

は、オリジナルサーバより、近隣に存在するキャッシュサーバから応答が返信されることにより、コンテンツ取得にかかる時間が低減する。

現在、インターネットにおいて実現されている代表的なコンテンツ配信技術として、コンテンツデリバリーネットワーク (CDN: Content Delivery Network) が存在する。CDN は能動的なキャッシュ制御を採っており、オリジナルサーバが提供するコンテンツを複製して保持するサーバ (複製サーバ) に分散配置する。複製サーバはコンテンツプロバイダによって管理され、ユーザからサーバまでの平均的なネットワーク距離を短縮するように配置される。そのため、コンテンツプロバイダは複製サーバをネットワーク内に複数配備することが一般的である。結果、ユーザの要求が近隣の複製サーバによって満足され得るため、オリジナルサーバの負荷を分散しつつ、ユーザの要求に対する応答速度を向上させることができる。

複製サーバのキャッシュストレージ容量には限界があるため、複製サーバはオリジナルサーバの全コンテンツの中から選択的にキャッシュ配置を行う [5] [6] [7]。例えば文献 [5] では、コンテンツの人気度は Zipf 則 [8] に従うことが経験的に判明している。そのため、Zipf 則を基に各コンテンツの人気度を確認し、人気なコンテンツから順にキャッシュ配置を行う。この手法を用いることにより、人気コンテンツをキャッシュサーバに配置することが可能となるため、ユーザの要求がキャッシュサーバで満たされる確率、すなわちキャッシュヒット率が向上することが確認されている。次に文献 [6] では、コンテンツをいくつかのカテゴリに分類し、各カテゴリの人気度の時間的な推移を確認している。この観測結果に基づき、今後も人気が続くコンテンツや、現在は人気ではないが今後人気上昇するコンテンツを優先的にキャッシュ配置することにより、キャッシュヒット率が向上することが確認されている。最後に文献 [7] では、コンテンツの人気度は地域性があること [9] に着目したキャッシュ配置を行っている。ユーザがコンテンツに対して自身の好みと一致しているかを基準に五段階評価した情報から、その地域におけるコンテンツの人気度を予測している。その予測結果を基に地域ごとに人気なコンテンツをキャッシュ配置することによって、その地域におけるキャッシュヒット率が向上することが確認されている。

上述の既存研究 [5] [6] [7] はコンテンツ人気度が時間に応じて推移すること、もしくは、地域によって異なることに着目したキャッシュ配置を行っている。すなわち、各コンテンツ単独の人気の特性に着目したキャッシュ配置を行っている。しかし、コンテンツを要求するのはユーザであるため、ユーザのコンテンツ要求傾向についても着目したキャッシュ配置をする必要がある。本稿では、既存の研究のように各コンテンツ単独の人気の特性に着目したキャッシュ配置ではなく、複数のユーザ間のコンテンツの嗜好の類似性に着目したキャッシュ判断法を提案する。ユーザ同士の嗜好の類似性により、現在は要求されていないが今後要求される可能性の高いコンテンツを予測することが可能であり、そのコンテンツをキャッシュ配置することによってキャッシュヒット率が向上することが見込まれる。性

能評価では、本提案方式により今後要求されると予測されたコンテンツをキャッシュサーバにキャッシュ配置し、そのキャッシュヒット率を確認する。

2. 背景技術

人気コンテンツを提供するウェブサーバが一つしか存在しない場合、ユーザのアクセスはそのサーバに集中する。この状況は、サーバとユーザの双方にとって深刻な問題を引き起こす。まず、サーバの観点では、あるサーバにアクセスが集中することでサーバに対する負荷が上昇するだけではなく、サーバ付近のリンクにおいて輻輳が発生する。加えて、アクセス集中による負荷や故障等によりサーバの機能が停止すると、コンテンツプロバイダはコンテンツを一切提供できなくなるため、耐故障性の点で問題が発生する。ユーザの観点では、オリジナルサーバが一つしか存在しないため、あるユーザにとってはコンテンツ取得先が遠隔地に位置する場合がある。遠隔地に位置するオリジナルサーバへ到達するまでに多くのリンクを経由するため、途中で輻輳したリンクが存在する可能性が高まる。輻輳したリンクを介してコンテンツをダウンロードする場合、取得完了までの時間が増加する。そこで、オリジナルコンテンツの複製を保持するサーバ (複製サーバ) をネットワーク内に複数配置し、ユーザのコンテンツ要求をユーザ近辺の複製サーバにより満足させるコンテンツデリバリーネットワーク (CDN: Content Delivery Network) が考案された [4]。

CDN ではユーザの要求が複数の複製サーバに分散されるため、オリジナルサーバの負荷を低減することができる。さらに、一つのサーバが機能停止に陥った場合においても、他の複製サーバによるコンテンツ提供が可能であるため、ウェブシステムとしての耐故障性が上昇し、サービスの可用性が向上する。またユーザと物理的に距離が近いサーバからコンテンツが提供されることで、ユーザの要求に対する遅延性能が向上する。

複製サーバのキャッシュストレージ容量の観点から、オリジナルサーバの全コンテンツを複製サーバが保持することは現実的ではない。そこで、複製サーバはオリジナルコンテンツのうち、特定のコンテンツを選択的にキャッシュしている [5] [6] [7]。ここでは、キャッシュするコンテンツをどのように選択しているのかについて紹介する。

コンテンツの要求傾向として、コンテンツの要求回数の分布が Zipf 則に従うことがわかっている [8]。Zipf 則とは、要求発生が k 番目に多いコンテンツの要求発生確率が、要求発生が一番多いコンテンツの要求発生確率の $1/k$ になるという経験則である。そこで、文献 [5] は Zipf 則を基にコンテンツの人気度を確認し、人気順にキャッシュするコンテンツを決定している。この手法を用いることにより、人気コンテンツがキャッシュ配置され、ユーザのコンテンツ要求がキャッシュによって満たされる率、すなわちキャッシュヒット率が向上することが確認されている。このコンテンツ人気度は時間的に、もしくは空間的に変化することが知られており、上述した方法はこれに対処するものとはなっていない。

コンテンツの人気度の時間推移について、ニュースなどのコ

コンテンツを例に説明する。事件や事故が発生した当日には、そのニュースに関連する動画がよく視聴されるが、ある程度時間が経つと殆ど視聴されなくなるのが一般的である。つまり、事件や事故が発生した直後では、それに関連したニュースなどのコンテンツの人気は高いが、時間経過に伴い人気は低下する。そのため、事件や事故の発生当初の人気度のみをもとにキャッシュするコンテンツを決定すると、今後要求されないコンテンツをキャッシュサーバに配置する可能性があり、効率的なキャッシュ配置を行えない。文献[6]では、コンテンツを映画、テレビ番組、ミュージックビデオ、スポーツ、ニュースといった5つのカテゴリに分け、各カテゴリの人気度の推移を確認している。その観測結果に基づき、今後人気度が上昇するコンテンツや、人気が長続きするコンテンツを識別したうえで、これらを優先してキャッシュ配置することで、時間的な人気度の変化を考慮したキャッシュ配置法を提案している。しかし、コンテンツの人気度はどの場所においても同一であるというわけではなく、コンテンツの人気度には地域性が存在することが報告されている[9]。ここでいう地域性とは、地域ごとにコンテンツの人気度分布が異なるということである。文献[6]はコンテンツ人気度の地域性を考慮しておらず、その効果が限定的となる可能性がある。

コンテンツの人気度には地域性が存在することについて、ニュースなどのコンテンツを例に挙げて説明する。これは、事件や事故が発生した地域においては、そのニュースに関連する動画がよく視聴されるが、遠く離れた場所においては、そのニュースは殆ど視聴されなくなるということである。そのため、地域性を考慮せずにキャッシュするコンテンツを決定する事は、今後要求されないコンテンツをキャッシュサーバに配置する可能性があり、効率的なキャッシュ配置とは言えない。文献[7]では、ユーザがコンテンツを視聴した後に自身の好みと一致しているかを基準に五段階でコンテンツを評価した情報から、各地域におけるコンテンツの人気度を推測している。ユーザが全コンテンツを視聴して評価することは現実的ではないため、ユーザの嗜好とコンテンツの特徴からユーザのコンテンツに対する評価値を推測している。例えば、アクションが好きなユーザは、アクションの特徴を持つコンテンツに対して高評価を付けると考え、評価値を高く推測する。そして、それぞれのコンテンツの評価値情報を基に、各地域において人気なコンテンツを複製サーバにキャッシュする方法が提案されている。すなわち、ユーザの嗜好とコンテンツの特徴の関係性から今後要求されるコンテンツを推測し、キャッシュ配置を行っている。しかし、嗜好という観点で言えば、ユーザの嗜好とコンテンツの特徴の関係性だけではなく、ユーザ間の嗜好の類似性にも着目することが出来る。すなわち、文献[7]は前者に関しては考慮しているが、後者に関しては一切考慮していない。

3. ユーザ間の嗜好の類似性を利用したキャッシュ判断法

本節では、動画に対するユーザ間の嗜好の類似性を利用したキャッシュ判断法の提案を行う。2章で述べたように、文献[7]

はユーザの嗜好とコンテンツの特徴の関係性から今後要求されるコンテンツを推測しキャッシュ配置を行っているが、ユーザ間の嗜好の類似性は考慮していない。そこで、本研究はユーザを嗜好の似たグループに分類することで、各グループで未視聴だが将来要求される可能性の高いコンテンツをより効果的に特定し、キャッシュ配置することによるキャッシュヒット率の向上を狙う。

3.1節では、嗜好の類似したユーザグループを作成する方法についての説明を行う。3.2節では、作成したユーザグループにおける動画要求傾向をもとに、キャッシュストレージに動画コンテンツを配置していく方法についての説明を行う。

3.1 K-means 法

異なる嗜好を持つユーザが混在する集合から、互いに類似した嗜好を持つユーザを集めてユーザクラスタを作る手法として、K-means 法[10]が存在する。K-means 法では、入力パラメータ k を与えることにより、 k 個のユーザクラスタを作成することが可能である。動作説明を簡単化するために、各ユーザの嗜好は2次元で表記することができるとする。

- step1. ユーザクラスタの中心値となるユーザを k 個選択
 - step2. 各ユーザは最も近隣の中心値のユーザクラスタに属する
 - step3. 中心値を各ユーザクラスタの中心値に更新
 - step4. 各ユーザは再度、最も近隣の中心値のユーザクラスタに属する
- step3 と step4 を繰り返し実行し、各ユーザクラスタに属するユーザ数に変化が無くなった時点で K-means 法の動作は終了する。

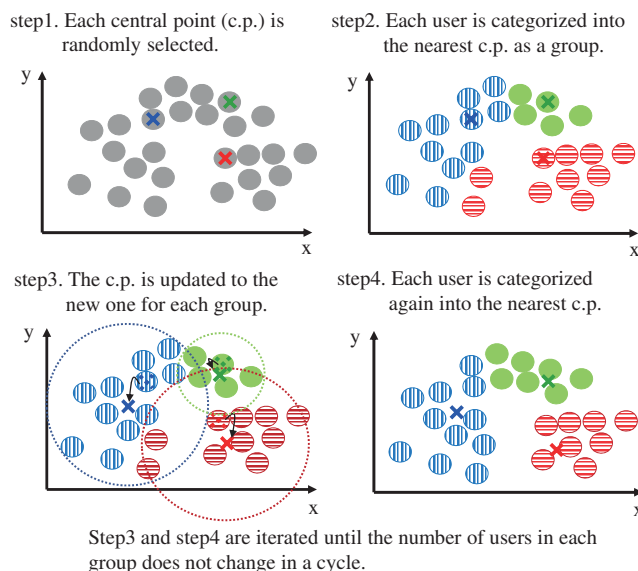


図1 K-means 法 ($k=3$, 嗜好の次元数: 2)

本稿では、ユーザの動画要求傾向を表すトレースデータとして MovieLens [11] を使用する。MovieLens は表1に示す18個のジャンルを用意しており、各動画は18個のジャンルのうち1つ以上のジャンルの属性を有している。ここで、使用するトレースデータにおける各ジャンルの要求数を図2に示す。Drama, Action, Comedy の要求数は他ジャンルよりも多いことから、全てのユーザから要求されるジャンルであると考えら

れる．嗜好の類似したユーザグループを作成する際には特定のユーザが要求しているようなニッチなジャンルが重要となるため，上述の3つのジャンル以外のジャンルを対象とする．

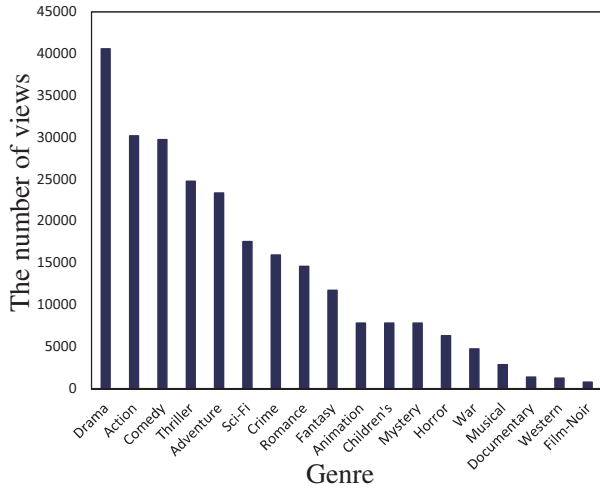


図2 ジャンルごとの要求数

ユーザ u のジャンル g に対する視聴回数を I_u^g とすると，ユーザ u のジャンル g に対する嗜好 T_u^g は式 (1) を用いて表す．

$$T_u^g = \frac{I_u^g}{\max_{g \in G} \{I_u^g\}} \quad (1)$$

クラスタリングを行う際にはユーザ間の類似性，すなわち距離を導出する必要がある．そこで，式 (2) を用いてユーザ p とユーザ q のユークリッド距離 $d(p, q)$ を定義する．

$$d(p, q) = \sqrt{\sum_{g \in G} (T_p^g - T_q^g)^2} \quad (2)$$

距離の値が小さいほどユーザ同士の嗜好は類似しており，逆に距離の値が大きいほどユーザ同士の嗜好は類似していないことを示す．すなわち，距離の値が小さいほど，同一のユーザクラスタに属する可能性が高くなる．

前述の通り，K-means 法は任意の入力パラメータである k 個のユーザクラスタを作成することができる．入力パラメータ k の値を決定する際に，SSE (Sum of Squared Errors) の値を使用する．SSE 値はユーザとユーザが属するユーザクラスタの中心値との距離の総和を取ったものであり，ユーザクラスタ数 k の値と反比例する値である．そこで，本稿では，ユーザクラスタ数 k の値を増加させることによる SSE 値の減少量が大きい領域からある程度小さくなった時点ユーザクラスタ数として選択する．ユーザクラスタ数 k の値を増加させることによる SSE 値の減少量を図 3 に示す．グラフより，ユーザクラスタ数を $k=6$ 以降， k の値を増加させたとしても SSE 値の減少量が小さいことがわかる．

3.2 キャッシュ配置方法

3.1 節において作成した各ユーザクラスタ内のユーザは嗜好が類似しているため，今後要求する動画も類似する可能性が高い．そこで，各ユーザクラスタにおける動画の要求傾向から今後要求される可能性が高いと判断される動画をキャッシュストレージに配置することにより，キャッシュヒット率の向上が見

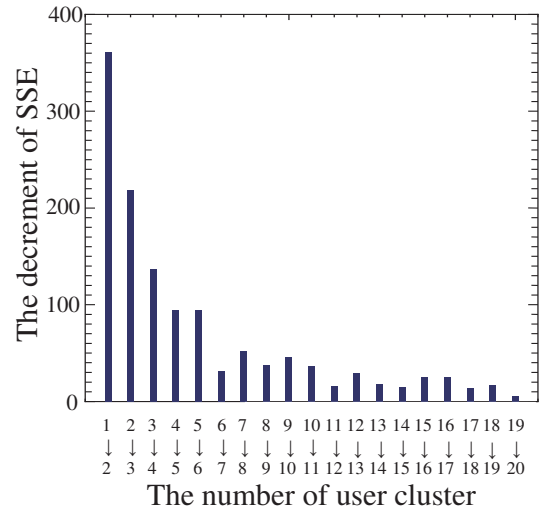


図3 ユーザクラスタ数の増加による SSE 値の減少量

表1 ジャンル

| | | |
|------------|-------------|----------|
| Action | Documentary | Mystery |
| Adventure | Drama | Romance |
| Animation | Fantasy | Sci-Fi |
| Children's | Film-Noir | Thriller |
| Comedy | Horror | War |
| Crime | Musical | Western |

込まれる．本稿では，全ユーザに対して配置するキャッシュストレージは1つであることを想定しているため，それぞれのユーザクラスタに対して割り当てるキャッシュストレージ容量を，ユーザクラスタの総視聴回数を用いて決定する．ユーザクラスタ c の総視聴回数を W_c とすると，ユーザクラスタ c に割り当てるキャッシュストレージ容量 CS_c を式 (3) を用いて決定する．

$$CS_c = CS * \frac{W_c}{\sum_{c \in C} W_c} \quad (3)$$

式 (3) により，ユーザクラスタに属するユーザの人数や要求回数がユーザクラスタごとに異なることを考慮したキャッシュ配置を行う．

アルゴリズム 1 において，キャッシュ配置の手順を説明する．各種パラメータに関しては表 2 の通りである．ユーザクラスタ内で多くのユーザが動画 v を視聴している場合，視聴していないユーザが今後動画 v を視聴する可能性は高い．ただし，本稿では同じユーザが同一の動画を複数回視聴することは稀であると仮定する．つまり，ユーザクラスタ内の大半のユーザが視聴した動画 v は今後視聴される可能性が低いので，動画 v をキャッシュするという判断は好ましくない．そこで，ユーザクラスタにおける動画 v のキャッシュ優先度 d_c^v を，ユーザクラスタにおいて最も人気な動画の視聴回数の半分と動画 v のユーザクラスタ内における視聴回数 w_c^v の差分の逆数と定義する．ユーザクラスタごとに動画のリストが形成されるので，キャッシュ優先度が高い順になるようにソートを実行した後，ユーザクラスタごとに割り当てられたキャッシュストレージ CS_c を満

たすまでリストの先頭にある動画を配置していく。ただし、対象とした動画が既にキャッシュストレージ CS に配置されている際は、重複した配置を行わずにリストからの削除のみを行う。

Algorithm 1 キャッシュ配置 アルゴリズム

```

Input:
   $C, V, N, M_c (\forall c \in C), CS_c (\forall c \in C), w_c^v (\forall c \in C, \forall v \in V)$ 
Initialize:
   $S \leftarrow \phi$ 
   $n = 1, 2, \dots, |C|$ 
   $d_c^v = \frac{1}{\left| \frac{M_c}{2} - w_c^v \right|} (\forall c \in C, \forall v \in V)$ 
Make List:
   $L_c = \{d_c^1, d_c^2, \dots, d_c^N\} (\forall c \in C)$ 
Ascending sort the List:
   $L_c (\forall c \in C)$ 
Place video into cache:
  while ( $|S| \leq CS$ )
    for ( $1 \leq |n| \leq C$ )
      Dequeue the first movie  $v$  from the list  $L_n$ 
      if ( $CS_n \leq S_n$  or  $|S| = CS$ )
        do nothing
      else if  $v$  is in  $S$ 
        do not cache  $v$ 
      else
        cache  $v$  into  $S_n$ 
      end if
    end for
  end while

```

表 2 list of symbols

| Symbol | Definition |
|-----------|--|
| T_u^g | The taste of user u to genre g |
| I_u^g | The number of views of user u to genre g |
| CT_c^g | The taste of cluster c to genre g |
| $d(p, q)$ | The distance between user p and user q in genre dimensions as taste metric |
| CS | Cache capacity |
| CS_c | Cache capacity for cluster c |
| W_c | The number of views in cluster c |
| C | Set of clusters |
| V | Set of movies |
| G | Set of genre |
| S | Set of movies in the cache storage |
| S_c | Set of movies of cluster c in the cache storage |
| N | The number of movies |
| M_c | The number of views to the most popular movie in cluster c |
| w_c^v | The number of views to movie v in cluster c |
| d_c^v | The cache priority of movie v in cluster c |
| L_c | The list of movies following the cache priority in cluster c |

4. 性能評価

本章では、複数のユーザ間のコンテンツの嗜好の類似性に着目したキャッシュ判断法の評価を行う。ユーザを嗜好の似たグ

ループに分類することで、各グループで未視聴だが将来要求される可能性の高いコンテンツをより効果的に特定し、キャッシュ配置することによるキャッシュヒット率の向上が期待される。

4.1 評価モデル

本節では評価モデルについて説明する。提案方式の性能を評価するため、ユーザの動画要求傾向を表すトレースデータとして MovieLens [11] を使用する。MovieLens は映画のレビューのためのウェブサイト及びデータセットであり、ユーザは映画に対して 10 段階の評価を行っている。本稿においては、”評価した”ということをも”視聴した”と解釈し、上述のトレースデータを使用する。評価に使用するトレースデータの内容を表 3 にまとめる。

表 3 トレースデータ

| 観測期間 | 2015 年 2 月 16 日～3 月 16 日 |
|--------|--------------------------|
| 要求総数 | 90591 [回] |
| ユーザ数 | 2017 [人] |
| コンテンツ数 | 8642 [個] |

期間として 1 ヶ月間のトレースデータの内、前半の 2 週間の要求を学習データとして使用する。学習データを用いてクラスタリングを実行することにより、嗜好の類似したユーザクラスを作成できていることの評価を行う。ユーザクラスごとのジャンルに対する関心度を評価指標として、式 (4) で定義する。

$$CT_c^g = \frac{\sum_{u \in c} I_u^g}{\max_{g \in G} \left\{ \sum_{u \in c} I_u^g \right\}} \quad (4)$$

また、トレースデータ後半の 2 週間の要求を使用し、提案方式の有効性の評価を行う。評価モデルとして、ユーザノードとコンテンツ提供サーバの間に介在するキャッシュサーバにコンテンツを配置し、そのコンテンツをユーザノードが要求した場合にキャッシュヒットと判定する。本稿におけるキャッシュサーバのキャッシュストレージにはコンテンツ単位でコンテンツを配置し、キャッシュストレージの容量をパラメータとして変動させる。

キャッシュヒット率を評価指標とし、式 (5) で定義する。

$$\frac{\text{ユーザの要求がキャッシュヒットした回数}}{\text{ユーザの動画要求総数}} \quad (5)$$

比較方式はキャッシュ方式として一般的に用いられる LRU (least recently used) を使用する。提案方式は、トレースデータ後半 2 週間の評価期間において配置コンテンツのリプレースメントを一切行わない静的な動作をするが、LRU は要求されてから最も長い期間が経ったコンテンツのリプレースメントを行う動的な動作をする。すなわち、LRU は要求されたコンテンツを受動的にキャッシュストレージに配置するため、要求頻度の高く全体的に人気なコンテンツがキャッシュストレージに配置される。

4.2 評価結果

本節では、複数のユーザ間のコンテンツの嗜好の類似性に着目したキャッシュ判断法の評価結果について述べる。まず、

4.2.1節ではクラスタリングを実行することで、嗜好の類似したユーザクラスタを作成できていることを示す。次に、4.2.2節では提案方式とLRUのキャッシュヒット率について示す。

4.2.1 各ユーザクラスタのジャンル関心度

図4にユーザクラスタごとのジャンルに対する関心度を示す。それぞれのユーザクラスタが関心をもつジャンルは異なっており、要求数が多いジャンルであるThrillerやAdventureに関心を持たず中人気のジャンルに対して一番の関心を持っているユーザクラスタも形成できていることがわかる。すなわち、本提案方式では嗜好の類似したユーザによってユーザクラスタを形成できていると考えられる。従って、そのユーザクラスタ内の動画要求傾向より将来要求される可能性の高いコンテンツの予測が可能となり、そのコンテンツをキャッシュストレージに配置することによって、キャッシュヒット率の向上を見込むことが可能である。

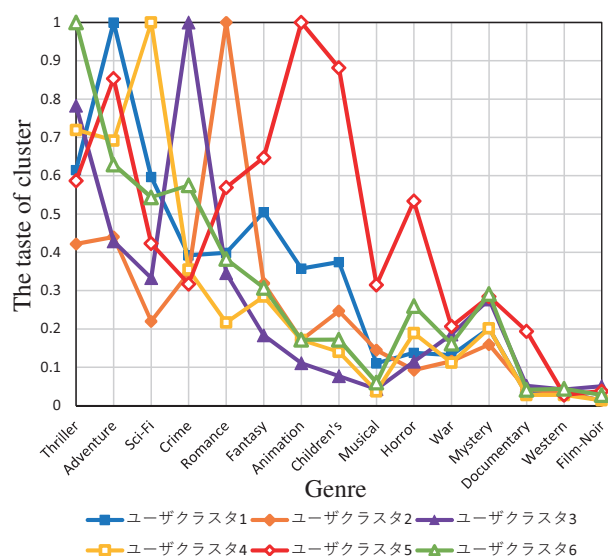


図4 ユーザクラスタのジャンルに対する関心度

4.2.2 キャッシュヒット率

提案方式とLRUのキャッシュヒット率について評価を行う。図5に各方式のキャッシュヒット率を示す。提案方式はLRUよりも高いキャッシュヒット率を示していることがわかる。LRUは低人気のコンテンツも一旦はキャッシュストレージにキャッシュするため、今後要求される可能性が高いコンテンツをリプレースメントすることが考えられる。それに対して、提案方式では複数のユーザ間のコンテンツの嗜好の類似性より将来要求される可能性の高いコンテンツをキャッシュ配置し続けるため、LRUと比較して高いキャッシュヒット率を示していると考えられる。特にキャッシュストレージ容量が小さい際における性能改善率が高いことから、キャッシュストレージ容量が小さい際により有用な方式とわかる。

5. まとめ

本稿では、キャッシュサーバのキャッシュヒット率を向上させることを目的としている既存方式を複数紹介した。また、既存方式はコンテンツの人気の時間的もしくは空間的な特徴に着

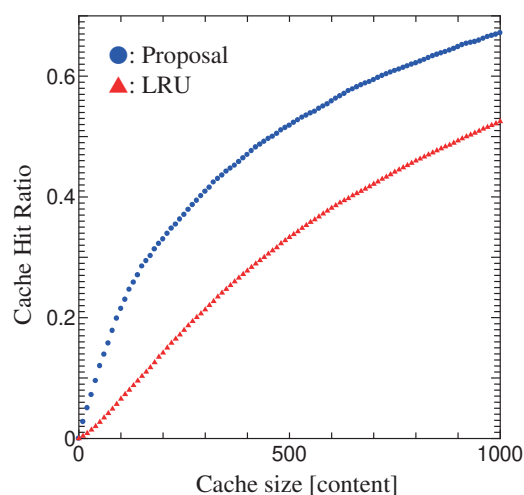


図5 キャッシュヒット率

目していることや、空間的なユーザの要求傾向に着目していることを述べた。そこで、複数のユーザ間のコンテンツの嗜好の類似性に着目し、既存研究とは異なるアプローチのキャッシュ判断法を提案した。性能評価において、一般的なキャッシュ方式であるLRUよりも、提案方式が高い性能を有することを示した。今後の課題として、MovieLens以外のトレースデータにおいても同様の結果が得られるのかを確認する。

謝辞 本研究の一部は、JSPS 科研費 17H01740 および 18K11283 により得られたものである。

文献

- [1] Cisco white paper, "Cisco Visual Networking Index: Forecast and Methodology, 2016-2021," <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>
- [2] "YouTube," <https://www.youtube.com/>
- [3] "Netflix," <https://www.netflix.com/>
- [4] Jim Kurose and Keith Ross, *Computer Networking A Top-Down Approach six edition*. Pearson, 2012.
- [5] Gerhard Hasslinger and Oliver Hohlfeld, "Efficiency of Caches for Content Distribution on the Internet," in *Proc. ITC 22*, September, 2010.
- [6] Yipeng Zhou, Liang Chen, Chunfeng Yang and Dah Ming Chiu, "Video Popularity Dynamics and Its Implication for Replication," *IEEE Transactions on Multimedia*, vol.17, no.8, August, 2015.
- [7] Stefan Dernbach, Nina Taft, Jim Kurose, Udi Weinsberg, Christophe Diot and Azin Ashkan, "Cache Content-Selection Policies for Streaming Video Services," in *Proc. IEEE INFOCOM*, April, 2016.
- [8] Lee Breslau, Pei Cao, Li Fan, Graham Phillips and Scott Shenker, "Web Caching and Zipf-like Distributions: Evidence and Implications," in *Proc. IEEE INFOCOM*, March, 1999.
- [9] Anders Brodersen, Salvatore Scellato and Mirjam Wattenhofer, "YouTube Around the World: Geographic Popularity of Videos," in *Proc. WWW12*, April, 2012.
- [10] David Arthur and Sergei Vassilvitskii, "k-means++: the advantages of careful seeding," in *Proc. ACM SODA*, January, 2007.
- [11] F. Maxwell Harper and Joseph A. Konstan, "The MovieLens Datasets: History and Context," *ACM Transactions on Interactive Intelligent Systems*, vol.5, no.4, December, 2015.