

2023年度

修 士 論 文

キャッシュを考慮した
コンテンツ推薦システム

指導教員: 上山 憲昭

立命館大学大学院 情報理工学研究科
情報理工学専攻 博士課程前期課程
計算機科学コース

学生証番号: 6611220041-0

氏名: 村上 悠斗

概要

現在のインターネットの主流サービスは、コンテンツ配信である。音楽や動画などの大容量のコンテンツのニーズが高まり、インターネットの通信量に拍車をかけている。Netflixなどのコンテンツ配信サービスにおいて、ユーザの嗜好に沿ったコンテンツを推薦するシステムがコンテンツ要求の大部分を占める重要な存在となっている。また、配信遅延や低ビットレートがユーザのサービス離脱率に影響する。そこで、従来の研究では、ネットワークの負荷軽減や推薦システムの課題であるコールドスタート問題を解決するアプローチがそれぞれ提案されている。しかし、従来の推薦方式ではネットワーク負荷の軽減とコールドスタート問題の両方については考慮していない。また、ユーザの観点では個々のユーザの好みに合ったコンテンツの推薦が望ましく、一方でネットワークの観点では配信コスト削減のために高人気コンテンツの推薦が望ましい。そこで本論では、ユーザの満足度および、キャッシュの効果向上を目的とした推薦と、協調フィルタリングのコールドスタート問題を改善するコンテンツ推薦方式を提案する。提案するキャッシュサーバ内のキャッシュされているコンテンツに限定した協調フィルタリングでは、ユーザから物理的に近いキャッシュサーバでキャッシュされているコンテンツを推薦し、その推薦されたコンテンツを視聴することで、ネットワーク負荷軽減とアクセス時間の短縮を図る。また、協調フィルタリングの課題である、コールドスタート問題への対応として、 ϵ -greedy 法と呼ばれる強化学習の推薦方式にキャッシュに限定した協調フィルタリングを組み合わせる。提案方式の有効性を計算機シミュレータにより評価する。

目次

第1章 序論	3
1.1 研究の背景	3
1.2 情報指向ネットワーク	4
1.3 協調フィルタリング	5
1.4 研究の目的	6
第2章 関連研究	7
2.1 コールドスタート問題	7
2.2 キャッシュを利用したコンテンツ推薦	7
第3章 提案方式	8
3.1 提案方式の概要	8
3.2 キャッシュに限定した協調フィルタリング	8
3.3 ϵ -greedy 法	9
第4章 ユーザの視聴行動モデル	11
4.1 ユーザの視聴	11
4.2 ユーザのコンテンツに対する嗜好度のモデル化	11
4.3 視聴行動モデル	12
第5章 性能評価	14
5.1 評価条件	14
5.2 評価尺度	15
5.3 CDN における評価結果	17
5.4 ICN における評価結果	21
第6章 まとめ	25
第7章 謝辞	26
外部発表文献	28

第1章 序論

1.1 研究の背景

Netflix や Amazon Prime などのコンテンツ配信サービスが急速に普及している現代において、サービス内のユーザ数の増加とともに、さらなる今後の需要の拡大が予測されている。これらのサービス内では、膨大な数のコンテンツが提供されており、ユーザが自分の興味・関心に合致するコンテンツを検索などによって見つけることは一層困難となっている。そこで、この問題を解決するため、多くのサービス内では、関心度が高いコンテンツをユーザに提案するシステムが導入されている。特に、コンテンツ提供事業者である CP (Content Provider) では、推薦システムがユーザのコンテンツ要求をサポートする重要な存在となっている。実際に Netflix では、推薦システムによって発生するコンテンツ要求数が全体の 8 割を占めることがわかっている [1]。そのため、ユーザにとって魅力的なコンテンツを推薦することで、ユーザのサービスに対するエンゲージメントが向上し、ユーザ数やコンテンツの視聴時間の増加が CP の収益につながる [2]。

特に、ビデオストリーミングサービスでは、低ビットレートや配信遅延 はサービスの離脱率の上昇につながる事が示されている。そのため、ユーザから高い QoE (Quality of Experience) を求められることが予想される。そこで、多数の CP は CDN (Content Delivery Network) と呼ばれる Web コンテンツをユーザに効率的に配信するネットワークを使用している。Akamai が CDN を提供する代表的な例である [3]。CDN は地理的に分散配置された複数のキャッシュサーバで構成されており、コンテンツがキャッシュされる。CDN の動作を図 1.1 に示す。

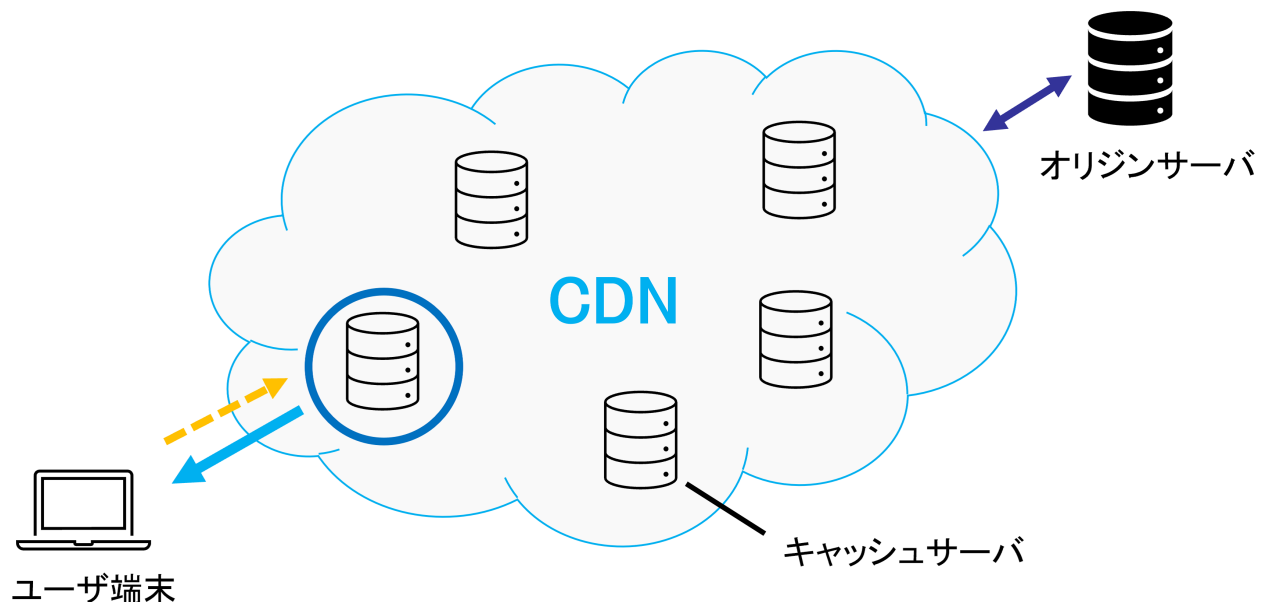


図 1.1: CDN のコンテンツ配信システム

ユーザからの要求に対してコンテンツの配信を行ったとき、ユーザから物理的に近いサーバでコンテンツがキャッシュされる。ユーザからのコンテンツ要求に対して、要求ユーザから物理的に近いキャッシュサーバからコンテンツを配信することで、ネットワークの状態やユーザとオリジンサーバ間の距離にかかわらず、安定したコンテンツ配信を実現する。このCDNの仕組みにより、コンテンツへのアクセスを高速化し、ユーザの体感する配信遅延とオリジンサーバの負荷を軽減することで、CPは高品質なサービスを生み出している。

1.2 情報指向ネットワーク

現在、音楽や動画などの大容量コンテンツの流通や、タブレットやスマートフォンなどの携帯端末、SNS（ソーシャルネットワーキングサービス）に代表されるユーザ主体の情報発信サービスが、インターネット上を流れるトラフィックの大部分を占有している。これに伴って、大量のデータがネットワークに提供され、頻繁にデータが更新される環境が一般化することが予想される。そこで、新しいネットワークアーキテクチャとして、情報指向ネットワーク（ICN: information-centric networking）が注目を浴びている。従来のIPアドレスを用いた場所に基づく通信を行うインターネットは、DNS（Domain Name System）によって通信相手となるホストのIPアドレスを検索し、該当のIPアドレスを持つサーバにアクセスし、コンテンツを取得する。ICNはIPアドレスに基づく通信ではなく、コンテンツ名を識別子として通信を行う。配信経路上のルータなどのノードに要求されたコンテンツをキャッシュし、配信要求が発生すると近くに存在するノードに該当のコンテンツがキャッシュされている場合、そのノードから配信される。そのため応答時間の短縮やコンテンツの配信の品質、効率の向上が期待される。ICNの動作を1.2に示す。

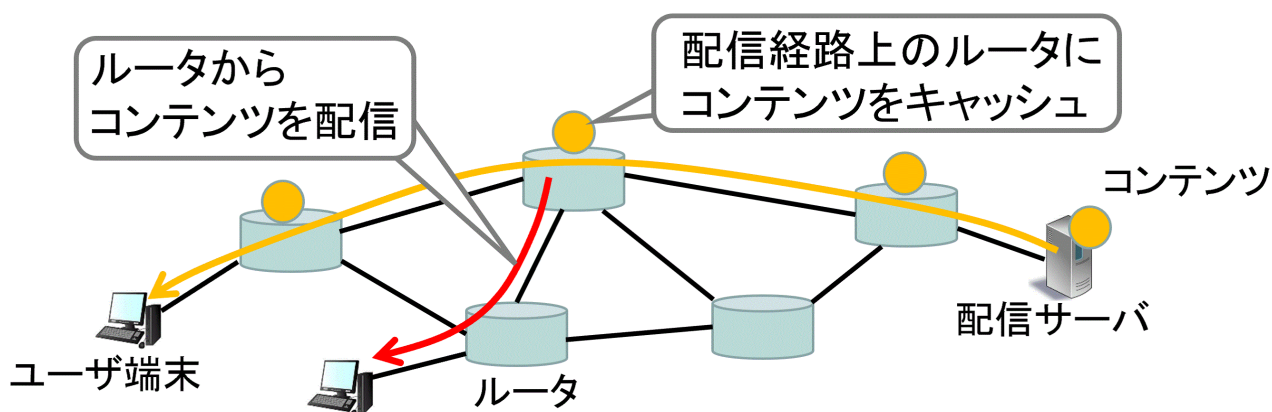


図 1.2: 情報指向ネットワーク

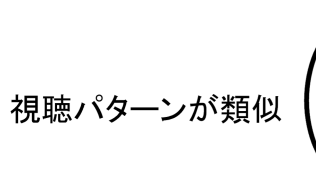
1.3 協調フィルタリング

従来の推薦方式は大きく2つに分類される。1つは、コンテンツベースフィルタリングと呼ばれる、コンテンツの特徴をもとに、ユーザが過去に評価したコンテンツの内容（例：映画のジャンル、キャストなど）を分析し、それらの特徴に基づいて、コンテンツを推薦する方式である。もう一つは、協調フィルタリングと呼ばれる、ユーザ間の嗜好度の類似性に基づいた推薦方式である。2つの方式のうち、協調フィルタリングが広く採用されている。コンテンツベースフィルタリングでは、あらかじめ、コンテンツの特徴を正確に把握するための分析技術やコンテンツの内容からユーザの嗜好が読み取れるような特徴を抽出するプロセスが必要となる。更には、ユーザが過去に好んだコンテンツと類似するコンテンツが推薦されることで、セレンディピティに欠ける可能性がある。例えば、特定のジャンルのコンテンツに限定された推薦である。このように個人の過去の行動履歴のみを利用すると、多様なコンテンツが推薦されない。一方、協調フィルタリングは、あるユーザの行動履歴（コンテンツの視聴や購入）をもとに、訪問ユーザが利用する可能性が高いコンテンツを推薦するため、コンテンツの内容や特徴を知っておく必要はない。また、嗜好の類似するユーザを利用することで、ユーザが知らなかった意外なコンテンツを推薦できる可能性がある。

協調フィルタリングには、ユーザベースフィルタリングとアイテムベースフィルタリングの2種類が存在する。

- ユーザベースフィルタリング

ユーザベースフィルタリングは、ユーザ間の関連性を考慮して推薦を行う手法である。この手法では、訪問ユーザと行動や好みが類似している他のユーザが好むコンテンツは、訪問ユーザも好む可能性が高いと仮定されている。アプローチとしては、類似する嗜好度や評価パターンを持つユーザを見つけだし、そのユーザが好むコンテンツを推薦する。図1.3に動作例を示す。例では、推薦対象のユーザCが映画Aと映画Dを視聴しているとする。システムはCと似た評価パターンを持つ他のユーザAを探し出す。次に、ユーザAが視聴したがユーザCがまだ見ていない映画BをユーザCに推薦する。



	映画 A	映画 B	映画 C	映画 D
ユーザ A	○	-	○	○
ユーザ B	○	○	×	-
ユーザ C	○	×	✓	○
ユーザ D	×	○	-	×

図 1.3: ユーザベースフィルタリング

- アイテムベースフィルタリング

アイテムベースフィルタリングは、アイテム間の関連性を考慮して推薦を行う手法である。この手法では、ユーザは過去に好んだコンテンツと似たコンテンツを好むと仮定されている。アプローチとしては、特定のコンテンツに対するユーザの評価パターンを分析し、似たような評価パターンを持つコンテンツを見つけ出して推薦する。図1.4に動作例を示す。例では、推薦対象のユーザが映画Aを既に視聴している場合、映画Aと類似度が高い映画Dを、システムはこれらの映画が似ていると判断し、映画Aを好むユーザに映画Dを推薦する。

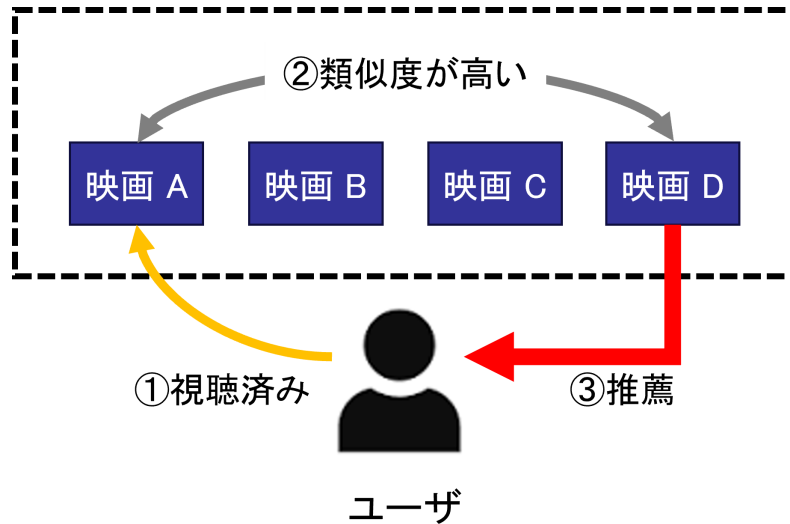


図 1.4: アイテムベースフィルタリング

しかし、これらの協調フィルタリングのアプローチでは、ユーザの行動履歴に関するデータを使用してコンテンツを推薦するため、データが少ない場合で精度が低下するコールドスタート問題と呼ばれる課題が存在する。ユーザ間で共通に評価しているコンテンツが存在しない場合に、嗜好の類似するユーザがを見つけられず、推薦ができない状態である。新規コンテンツや新規ユーザの増加は、今後の配信サービスの普及していく過程で、必ず発生するものである。そのため、協調フィルタリングは優れた推薦方式である一方で、適切に機能するための多くの行動データの取得が必要となる。

1.4 研究の目的

従来の研究では、ユーザの嗜好に沿ったコンテンツ推薦を目的とした方式だけでなく、ネットワークの負荷軽減やデータのコールドスタート問題を解決するアプローチも提案されている。しかし、従来の推薦方式ではネットワーク負荷の軽減とコールドスタート問題の両方については考慮していない。また、ユーザの観点では個々のユーザの好みに合ったコンテンツの推薦が望ましく、一方でネットワークの観点では配信コスト削減のために高人気コンテンツの推薦が望ましい。そこで本論では、ユーザの満足度および、CDN や ICN などのネットワークにおけるキャッシュの効果向上を目的とした推薦と、協調フィルタリングのコールドスタート問題を改善するコンテンツ推薦方式を提案する。さらに、計算機シミュレータにより、提案方式の有効性を確認する。

第2章 関連研究

2.1 コールドスタート問題

[4] で, Yuan らが提案したシステムは, 新規コンテンツを既存の行動データを持つコンテンツ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. 提案されたペアリングプロセスでは, 新しいコンテンツを既存のコンテンツと関連付けるための深層学習アルゴリズムである, DLM (Deep Learning Matcher) が使用されている. このペアリングにより, ユーザの新規コンテンツに関する行動データが無くても, 新規コンテンツを協調フィルタリングで考慮することが可能となっている. [5] では, IoT データの一般的なコールドスタート問題を提起し, 新しいハイブリッドキャッシュ戦略を提案している. キャッシュスペース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つに分割し, プロアクティブキャッシュでは深層学習を用いてデータの人気を予測し, リアクティブキャッシュでは改良された LRU アルゴリズムを使用している. この2つの組み合わせにより, パフォーマンスを確保しつつ, データのコールドスタート問題を解決している.

2.2 キャッシュを利用したコンテンツ推薦

[6] の研究では, 推薦システムがユーザの需要を形成し, キャッシュの利得を高めることに着目し, 基地局においてコンテンツのキャッシュと推薦を共同で最適化する手法を提案している. 需要形成の基本的なアイデアは, 推薦システムが個々のユーザの嗜好に最も合致したコンテンツを必ずしも推薦するのではなく, ユーザの嗜好にある程度まで合致し, 他のユーザにも魅力的なコンテンツを推薦することである. ユーザがそのような推薦を受け入れることで, ユーザの需要はより均質になり, コンテンツの人気度も偏る. キャッシュの効果を向上させるためのアプローチとして, CDN や無線ネットワークにおいてこの需要形成のアイデアが活用されている. [7] では, YouTube におけるコンテンツ推薦システムにおいて, すでにキャッシュされているコンテンツを関連リストの上部に配置することでキャッシュヒット率が増加したことが観察された.

[8] の研究では, キャッシュされたコンテンツの推薦に多様性を考慮することで, ユーザの満足度向上を図っている. 多様なカテゴリのコンテンツをユーザに推薦することで, より多くの選択肢を提供し, キャッシュに限定することによってユーザの嗜好に十分に対応できない問題点を補っている.

第3章 提案方式

3.1 提案方式の概要

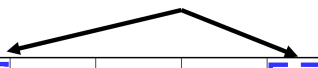
一般的にキャッシュされているコンテンツは配信回数が多く、人気度が高い。そこで、個々のユーザの嗜好に沿った推薦を提供するユーザベースの協調フィルタリングにおいて、推薦対象をキャッシュされているコンテンツに限定する。キャッシュに限定することで、他のユーザからの人気度が高く、かつ訪問ユーザの嗜好に沿ったコンテンツを推薦することが可能である。提案するキャッシュサーバ内のキャッシュされているコンテンツに限定した協調フィルタリングでは、ユーザから物理的に近いキャッシュサーバでキャッシュされているコンテンツを推薦し、その推薦されたコンテンツを視聴することで、コンテンツのアクセス時間が短縮され、パフォーマンスが向上する。また、オリジンサーバにかかる負荷が分散する。しかし、推薦候補が存在するキャッシュサーバをユーザの最寄りのサーバに限定することで、推薦候補となるコンテンツの数も限定され、各ユーザの嗜好を十分に考慮できない可能性がある。そこで、最寄りのキャッシュサーバに限定せず、ユーザからサーバまでの範囲を拡大させながら推薦の自由度向上を図る。また、協調フィルタリングの課題である、コールドスタート問題への対応として、 ϵ -greedy 法と呼ばれる強化学習の推薦方式にキャッシュに限定した協調フィルタリングを組み合わせる。確率 ϵ でランダムにコンテンツの推薦 (探索) を行い、確率 $1-\epsilon$ で現時点で収集済みの嗜好データのみを用いた協調フィルタリングによりコンテンツを推薦する (活用)。コールドスタートの問題により、推薦の対象となる機会が無いコンテンツは、ユーザへ露出する手段を失うことになる。しかし、探索において、推薦コンテンツのユーザの視聴の有無をデータとして収集することで、ユーザの行動履歴などに関するデータが少ないコンテンツに対して視聴データを収集することが可能である。また類似度の高いユーザの行動履歴にないコンテンツも推薦するため、協調フィルタリングにおける推薦するコンテンツの偏りをより緩和する。

3.2 キャッシュに限定した協調フィルタリング

提案方式の具体的な動作を図 3.1 を用いて説明する。評価点に基づく協調フィルタリングの動作を以下に示す。図 3.1 ではユーザの各コンテンツに対する 5 段階評価を嗜好データとして用いており、"-/-" は未評価を表す。提案方式ではキャッシュに存在するコンテンツのみを対象に評価度を比較する。

類似した嗜好を持つユーザ間の類似性をピアソン相関係数を用いて測定する。この相関係数は、二つの変数間の線形関係の強さを示し、協調フィルタリングにおいては特にユーザの評価パターンの類似性を評価するのに有効である。相関係数は、 $-1 \sim 1$ の範囲の実数を取り、正のときに同じような傾向 (1 に近いほど同じ傾向が強い)、負のときに逆の傾向 (-1 に近いほど逆の傾向が強い)

Cached content



Content \	1	2	3	4	5	6	7	8	9	10	Correlation coefficient
User 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				4.50	3.33	2.00				

図 3.1: Collaborative filtering for only cached contents

1. 訪問ユーザ X と各ユーザ間で類似度（相関係数） P_{xy} を計算 (3.1)
2. 相関係数が高い上位 k 人 ($k = 3$, C,D,E) の、キャッシュ中の各コンテンツに対する評価度の平均を計算
3. 平均値が最大のコンテンツ (= 5) をユーザ X に推薦

$$P_{xy} = \frac{\sum_{m \in C_{xy}} (r_{xm} - R_x)(r_{ym} - R_y)}{\sqrt{\sum_{m \in C_{xy}} (r_{xm} - R_x)^2} \sqrt{\sum_{m \in C_{xy}} (r_{ym} - R_y)^2}} \quad (3.1)$$

C_{xy} : ユーザ x と y の両方が評価値をつけているコンテンツの集合

r_{nm} : ユーザ n のコンテンツ m に対する評価値

R_n : ユーザ n の C_{xy} に対する評価値の平均値

3.3 ϵ -greedy 法

ϵ -greedy 法の動作を 3.2 に示す。ユーザの訪問時に確率 ϵ で全コンテンツの中からランダムにコンテンツを推薦する。また、確率 $1-\epsilon$ で協調フィルタリングによりコンテンツを推薦する。各推薦後、ユーザは視聴するか否かを判断する。

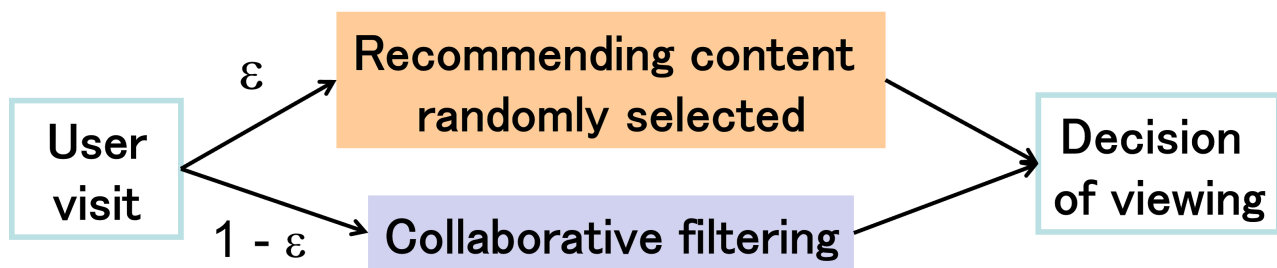


图 3.2: ϵ -greedy method

第4章 ユーザの視聴行動モデル

4.1 ユーザの視聴

ユーザは検索等による独自の検索行動と推薦システムによって、コンテンツを視聴する。検索行動では、ユーザは zipf 分布に従って視聴するコンテンツを選択する。zipf の法則は、出現頻度が k 番目に大きい要素の頻度が、1 位の要素の頻度と比較して、 $1/k$ に比例するという経験則である。主に、単語の出現頻度や web ページのアクセス頻度、都市の人口などの現象で適用される。zipf 分布は次式であらわされる

$$f(k; s, N) = \frac{1/k^s}{\sum_{n=1}^N (1/n^s)} \quad (4.1)$$

ただし f は出現頻度、 k は人気順位、 s は分布の偏りを調整するパラメタ、 N は全要素の数である。

各コンテンツの人気度順位は、全ユーザからの評価数をもとにして決定する、人気のあるコンテンツは、一定の品質や信頼性が保証され、話題性がある。そこで、推薦によらないコンテンツの視聴を行う場合、ユーザは膨大な数のコンテンツの中から視聴するコンテンツを選択する場合、ユーザ自身で選択する労力を減らすため、ユーザ全体で人気度が高いコンテンツを視聴すると想定する。

推薦によるコンテンツの視聴では、ユーザに推薦されたコンテンツへの嗜好度がある閾値を超えている場合に、その推薦コンテンツを視聴する。閾値は、推薦を受け入れるための厳しさや基準を表し、閾値が高いほど、ユーザはより限られた数の推薦を受け入れる。一方で閾値が低い場合は、推薦されるコンテンツを受け入れる基準が緩和され、多くの推薦を受け入れる。ユーザのコンテンツに対する嗜好度の推測には、コンテンツとカテゴリの関連度、およびユーザのカテゴリに対する関心度を考慮する [3]。

4.2 ユーザのコンテンツに対する嗜好度のモデル化

訪問ユーザの嗜好に近いコンテンツを推薦することで、ユーザはそのコンテンツを要求する確率が高くなる。[9]より、ユーザの各コンテンツに対する嗜好度を、コンテンツとカテゴリ間の関連度とユーザのカテゴリに対する関心度から推測する。

- コンテンツとカテゴリ間の関連度

各コンテンツは、1 つまたは複数のカテゴリに関連している。

(例 "Toy Story", Adventure|Animation|Children|Comedy|Fantasy)

あるコンテンツが m 個の異なるカテゴリで記述されている場合、カテゴリとコンテンツ間の関連度 $f^i(j)$ を $1/m$ に設定し、コンテンツ i がカテゴリ j にどれだけ関連しているかを示す。関連度は、0 から 1 の範囲で、全てのスコアの合計が 1 になるように正規化される。

例ではコンテンツ "Toy Story" を構成しているカテゴリは、Adventure から Fantasy までの 5 つで構成されているため ($m = 5$)、各カテゴリとの関連度は $f_i(j) = 1/5 = 0.2$ となる。

- ユーザのカテゴリに対する関心度

ユーザ u のカテゴリ j に対する関心度 $f_u(j)$ を (4.2) で導出する．関心度も，0 から 1 の範囲で，全てのスコアの合計が 1 になるように正規化される．

$$f_u(j) = \frac{\sum_{i \in C_r(u)} f^i(j)}{\sum_{j=1}^M \sum_{i \in C_r(u)} f^i(j)} \quad (4.2)$$

ただし $C_r(u)$ はユーザ u が評価したコンテンツの集合， M は全カテゴリの総数である．

- ユーザの各コンテンツに対する嗜好度

各ユーザ u は，各コンテンツに対する嗜好度分布 (P_u^{pref}) を持つ．ユーザ u がコンテンツ i に対する嗜好度 ($P_u^{pref}(i)$) は， u と i の特徴ベクトルから推定できる．2つのベクトルのコサイン類似度 a_{ui} を使用して，嗜好度を計算する．類似した嗜好を持つユーザ間の類似性の計算には，協調フィルタリングで一般的に使用されるコサイン類似度 (cosine similarity) を使用する．コサイン類似度は，2つのベクトル間の類似性を測定し，テキスト分析やクラスタリングのグループ分けにおいて使用される．

ユーザ u とコンテンツ i の類似度 a_{ui} を f_u と $f^i(j)$ から次式で得る．

$$a_{ui} = \frac{\sum_{j=1}^M f_u(j) f^i(j)}{\sqrt{\sum_{j=1}^M f_u(j)} \sqrt{\sum_{j=1}^M f^i(j)}} \quad (4.3)$$

$a_{u,i}$ を正規化することで，ユーザ u のコンテンツ i に対する嗜好度 $P_u^{pref}(i)$ を導出する．

ユーザ u のコンテンツ i に対する嗜好度は，

$$P_u^{pref}(i) = \frac{a_{ui}}{\sqrt{\sum_{i \in I} a_{ui}}} \quad (4.4)$$

となる．この方法により，ユーザの嗜好を特徴付け，モデル化する．

4.3 視聴行動モデル

視聴行動モデルを図 4.1 に示す．本システムの推薦対象とする訪問ユーザ X を決定し，推薦確率 p_r で訪問ユーザ X に対してコンテンツの推薦を行う．推薦を行う場合， ϵ -greedy 法に移行し，確率 ϵ で訪問ユーザ X が未視聴のコンテンツ集合の中からランダムに推薦する．確率 $1-\epsilon$ でキャッシュに限定した協調フィルタリングによる推薦を行う．推薦が受け入れられなかった場合とコールドスタートの場合は，zipf 分布に従って未視聴のコンテンツの集合の中からランダムに視聴する．コールドスタートの状態として，訪問ユーザ X と他の Y の両方が評価値をつけているコンテンツの集合 C_{XY} が 0 の場合とキャッシュに推薦可能なコンテンツが存在しない場合を想定する．

ε - greedy法

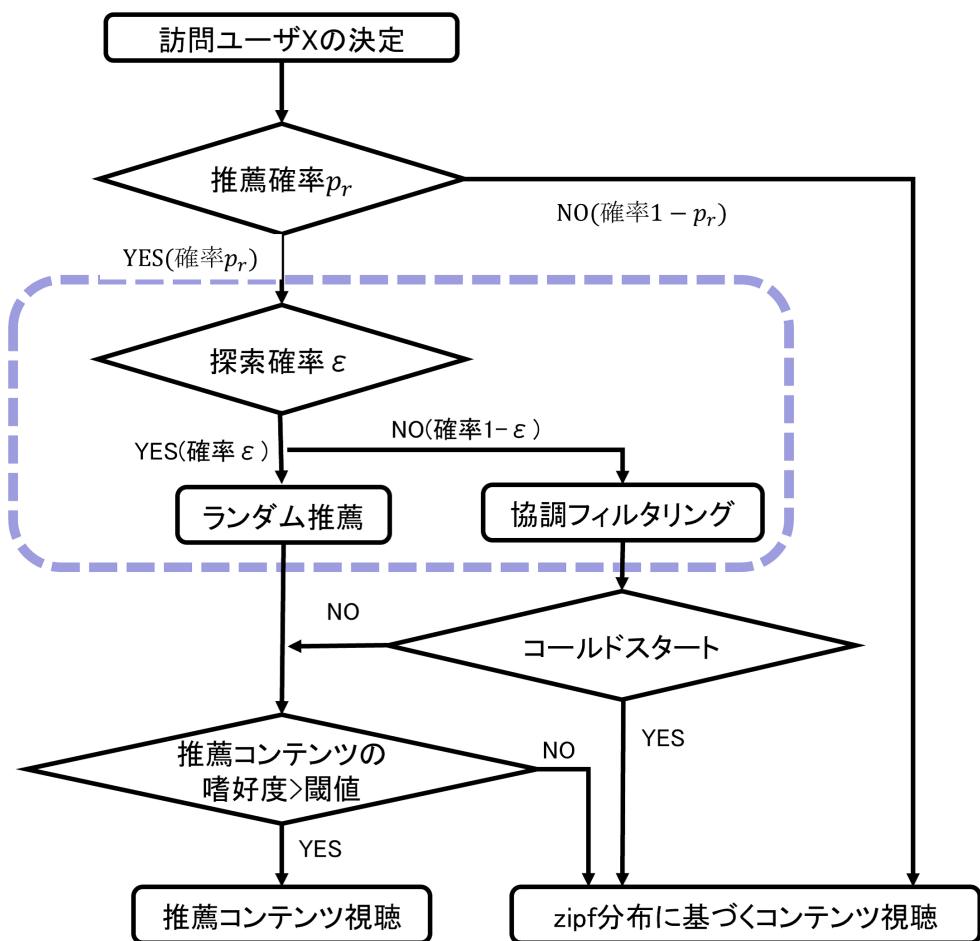


図 4.1: 視聴行動モデル

第5章 性能評価

5.1 評価条件

- 想定するネットワーク

評価に使用したネットワークトポロジを図 5.1 に示す. 各コンテンツのオリジンサーバは, 各ノードの人口比に比例して配置する. 本論では ICN と CDN について, 評価を行う.

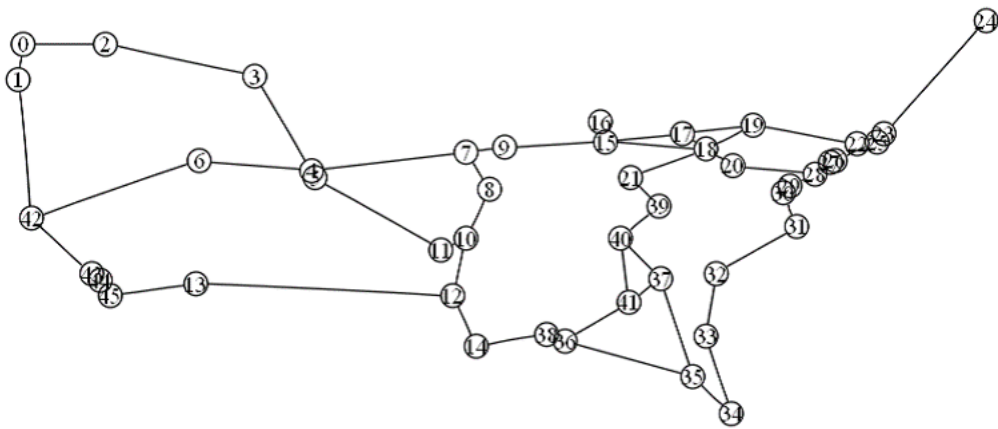


図 5.1: At Home Network

- データセット

シミュレーションで使用する各映画の評価やタイトル, ジャンルのデータを Movielens のデータセットから用いる [10]. Movielens は, 幅広い映画に関する豊富な情報を提供する公開データベースであり, ユーザによる映画の評価や映画のメタデータを含んでいる. 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, リアルなユーザ行動の分析に適している. 総コンテンツ数 $N_c = 9742$, 全ユーザ数 $N_u = 610$, 総ジャンル数 $N_j = 20$ である.

- システム訪問ユーザの選択

シミュレーションでは, Movielens のデータセット内にある各ユーザのコンテンツに対する評価回数を, そのユーザのコンテンツ配信サービスにおける利用頻度として考える. これにより, 実際のコンテンツ消費パターンを模倣し, 頻繁にサービスを利用して, コンテンツを視聴しているかを示す指標となる. 評価数から得られた累積分布より, 各シミュレーションにおける, システムの訪問ユーザを決定する. 各ユーザは各ノードの人口比に基づいて, コンテンツ配信サービスを利用する位置をランダムに割り当てる.

- ユーザによる評価

ユーザに推薦されたコンテンツへの嗜好度がある閾値を超えている場合にその推薦コンテンツが視聴される。[2] より、推薦コンテンツは全要求コンテンツの 8 割を占めている。ユーザ数が $N_u = 610$ 、コンテンツ数が $N_c = 9742$ である場合、ユーザのコンテンツに対する嗜好度は $N_{pref} = N_u \times N_c = 5942620$ 個存在する。そこで、 N_{pref} が約 8 割 ($5942620 \times 0.8 = 4754096$) の場合のコンテンツ嗜好度を閾値として設定する。ユーザがコンテンツの視聴後に付与する評価値は、5 段階評価を想定し、視聴したコンテンツの嗜好度に 5.0 を掛けたものとする。

- コンテンツ配信

ユーザの存在するノードで要求コンテンツのキャッシュの有無を確認し、キャッシュに存在すれば直ちにユーザに配信する。もしキャッシュになければ、要求コンテンツのオリジンサーバまで要求パケットを送信し、コンテンツの配信経路上のキャッシュサーバで要求コンテンツのキャッシュを行う。コンテンツの配信経路は、ダイクストラ法で設定する。キャッシュ置換方式は、LRU (Least Recently Used) を使用し、最も長い間要求されていないコンテンツをキャッシュから削除する。ICN では、コンテンツの配信経路上の全ルータでコンテンツをキャッシュする。CDN では、配信要求を送信したユーザの最寄りのサーバでキャッシュする。

5.2 評価尺度

提案方式を計算機シミュレーションにより以下の 3 つの評価尺度を評価し、有効性を確認する。

- H : 配信経路上のキャッシュヒット率

CDN の場合、要求パケットが発生したノードで要求コンテンツがキャッシュされていた場合に、オリジンサーバまで要求が行かず、そのノードからコンテンツが配信される。その場合のヒット率を次式で定義する。

$$H_{CDN} = \frac{n_{r_{CDN}}}{a_r} \quad (5.1)$$

ただし $n_{r_{CDN}}$ は要求パケットが発生したノードで要求コンテンツがキャッシュされてた回数であり、 a_r はネットワーク全体の要求の発生回数である。

ICN の場合、要求パケットがオリジンサーバまで転送される経路上の任意のノードで要求コンテンツがキャッシュされていた場合、オリジンサーバまで要求が行かず、途中のノードからコンテンツが配信される。その場合のヒット率を次式で定義する。

$$H_{ICN} = \frac{n_{r_{ICN}}}{a_r} \quad (5.2)$$

ただし $n_{r_{ICN}}$ は要求パケットの転送経路上の任意のノードで要求コンテンツがキャッシュされてた回数である。

- HL : 配信ホップ長の平均

$$HL = \sum_{n=1}^N \frac{a_n}{h_n} \quad (5.3)$$

ただし N は全ノード数、 a_n はノード n に到着した要求数、 h_n はノード n でキャッシュヒットした回数である。

- U_{cf} : 協調フィルタリングで推薦されたコンテンツに対するユーザ満足度

$$U_{cf} = \frac{\sum_{i \in C_{cf}} P_i \times 5.0}{a_{cf}} \quad (5.4)$$

ただし C_{cf} は協調フィルタリングで推薦された要求の集合, P_i は要求 i におけるユーザの満足度, a_{cf} は協調フィルタリングで推薦された要求数である. ユーザ満足度を 0~5 の範囲でスケーリングするため 5.0 を乗じる.

- U_v : 全視聴要求に対するユーザ満足度

$$U_v = \frac{\sum_{i \in R} P_i \times 5.0}{a_r} \quad (5.5)$$

ただし R は全要求の集合である. シミュレーション時間 (T_s) の中で, ランダムなコンテンツ推薦を行う探索時間の割合を P_{ex} とする. 探索期間の終了後, 時間が経過するにつれて探索の確率は減少し, 協調フィルタリングを用いた推薦 (活用) の割合が増加する. これにより, 初期段階ではランダムな探索が主となり, その後徐々にユーザの好みに基づいた推薦へと移行する.

CDN と ICN のそれぞれにおいて, 各探索時間の割合 P_{ex} を変化させた場合のキャッシュヒット率, ユーザの満足度を評価する.

5.3 CDN における評価結果

CDN では要求ユーザが存在するノードのキャッシュのみが使用される。シミュレーション時間 (T_s) の中でランダムなコンテンツ推薦を行う探索時間の割合 P_{ex} における、キャッシュヒット率と平均ホップ長、ユーザ満足度の性能評価の結果を示す。

- キャッシュヒット率

キャッシュヒット率 (5.1) の結果を図 5.2 に示す。

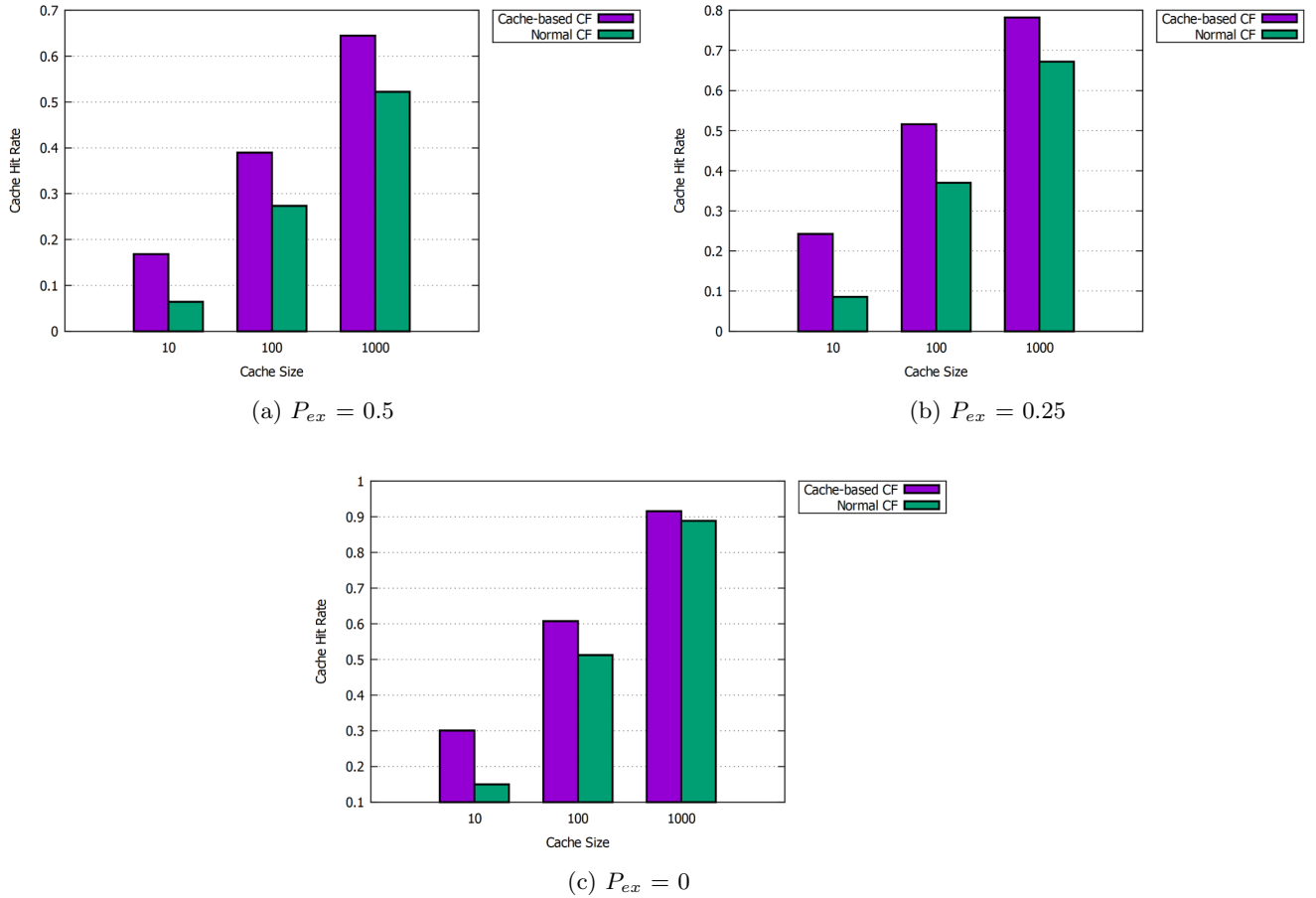
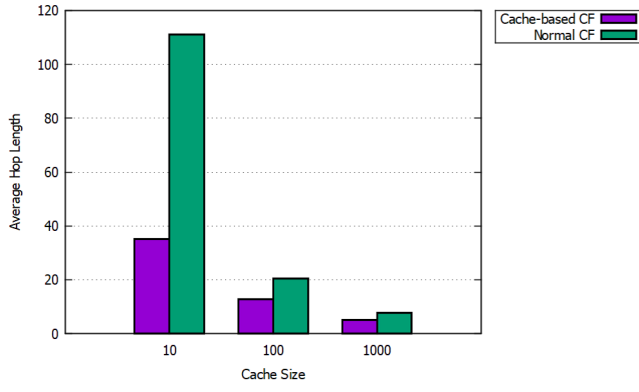


図 5.2: CDN におけるキャッシュヒット率

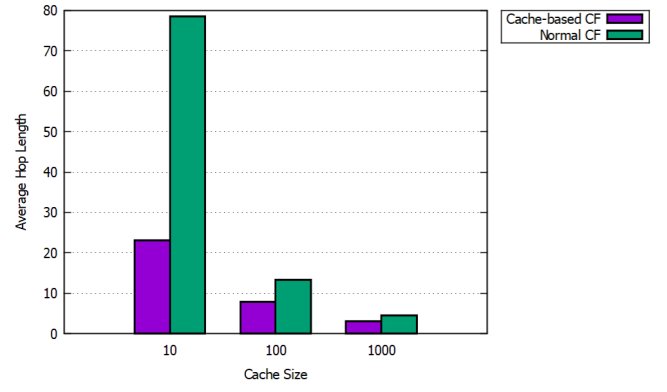
全ての探索時間と各キャッシュサイズにおいて、キャッシュを考慮した協調フィルタリングが従来の協調フィルタリングと比較して、キャッシュヒット率が高い。

- 平均ホップ長

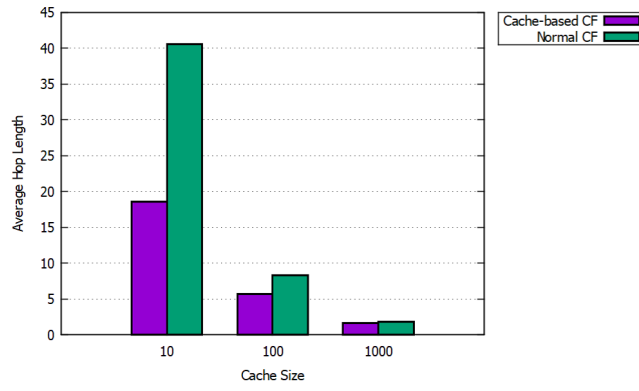
平均ホップ長 (5.3) の結果を図 5.3 に示す。



(a) $P_{ex} = 0.5$



(b) $P_{ex} = 0.25$



(c) $P_{ex} = 0$

図 5.3: CDN における平均ホップ長の評価

全てのキャッシュサイズと P_{ex} において、従来方式より、提案方式の平均ホップ長が短く、キャッシュサイズがより小さい場合に高い性能を発揮していることがわかる。

- ユーザ満足度

協調フィルタリングによるユーザの満足度 (U_{cf}) を図 5.4 に示す。

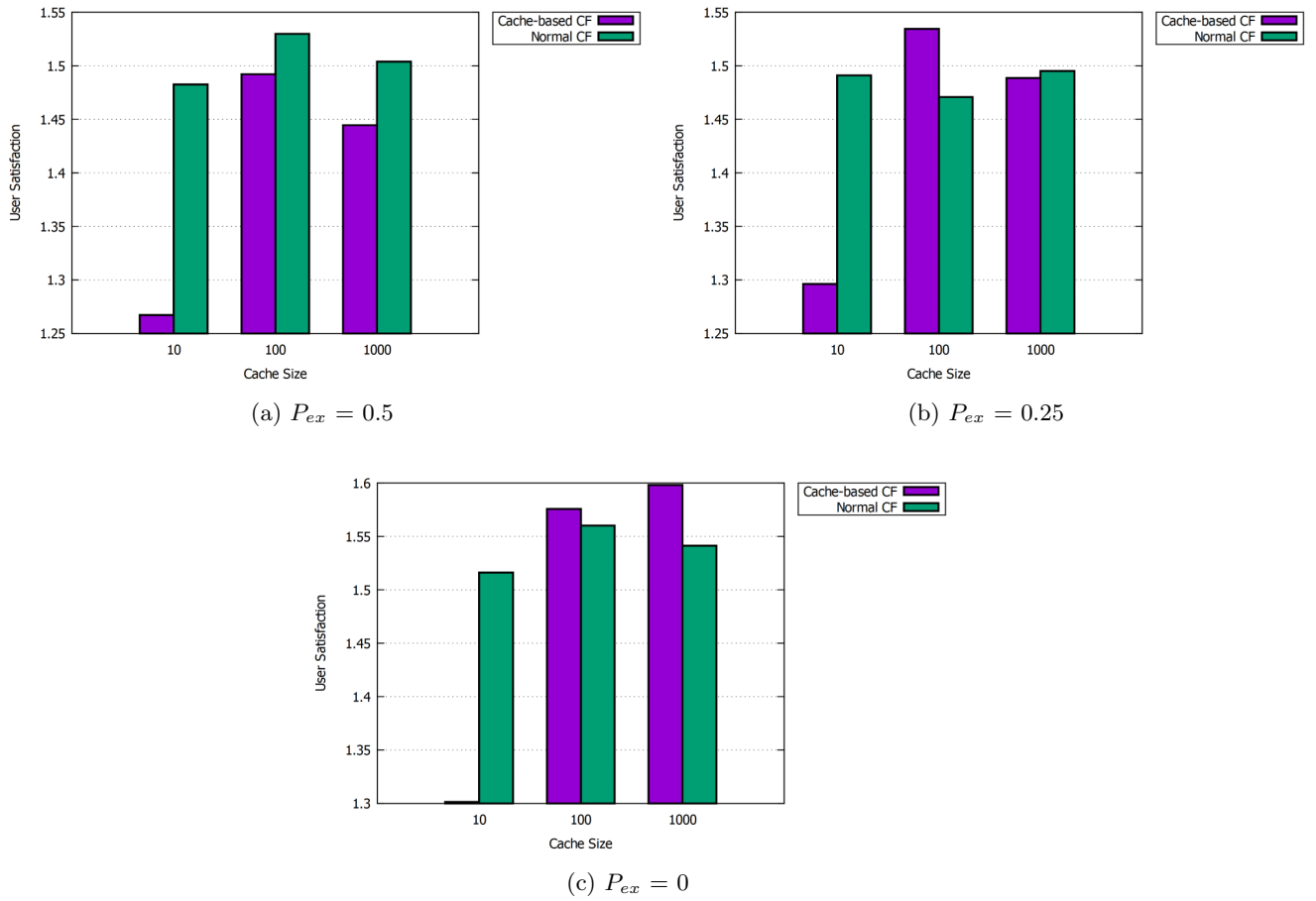


図 5.4: CDN における協調フィルタリングのユーザ満足度評価

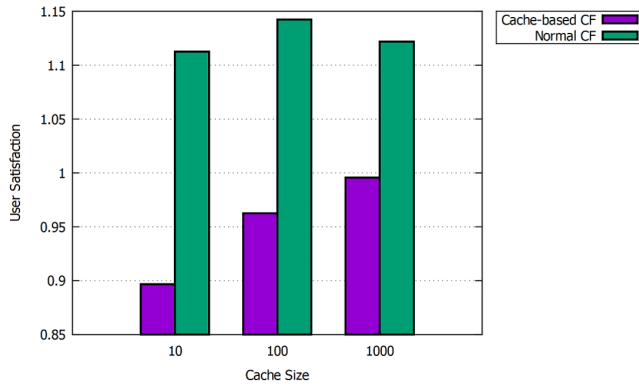
探索期間の短縮と協調フィルタリングに基づく推薦の頻度増加により、ユーザ満足度が向上してユーザの満足度が向上していることがわかる。また、以下にキャッシュサイズが 1000, $P_{ex} = 0$ ときの提案方式と従来方式における視聴確率を示す。

従来方式による視聴回数 = 22869 / 総視聴回数 = 95823 / 視聴確率 = 0.238659

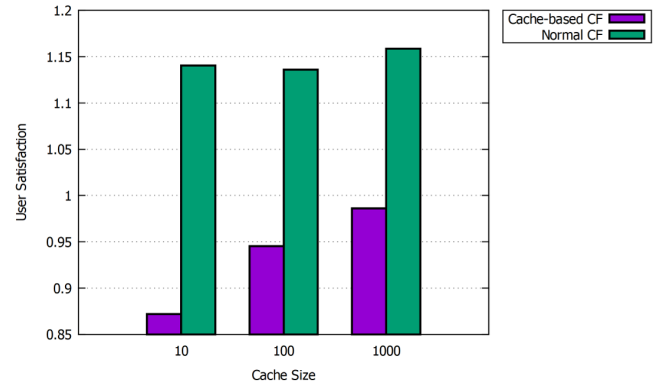
提案方式による視聴回数 = 18435 / 総視聴回数 = 95284 / 視聴確率 = 0.193474

従来方式と比較して、提案方式の視聴確率が低い。推薦したコンテンツは高い確率でユーザに視聴されるため、キャッシュに推薦できるコンテンツが存在せず、従来方式より推薦回数が少なくなったと考えられる。図 5.4c(c) に示される結果では、提案方式が推薦したコンテンツに対するユーザの満足度が従来方式を上回っており、提案方式の推薦コンテンツの質が高いことがわかる。

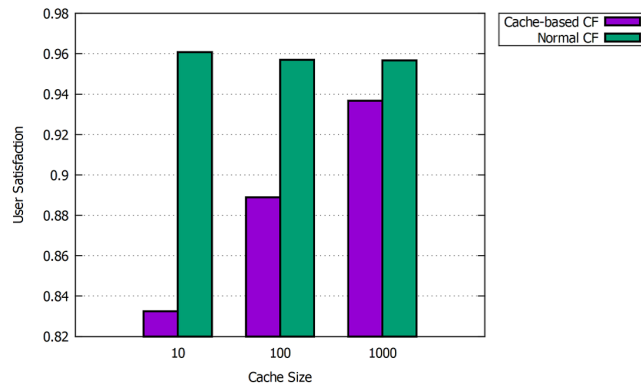
視聴した全コンテンツに対するユーザ満足度 (5.5) の結果を図 5.5 に示す。



(a) $P_{ex} = 0.5$



(b) $P_{ex} = 0.25$



(c) $P_{ex} = 0$

図 5.5: CDN における視聴した全コンテンツに対するユーザ満足度評価

キャッシュヒット率とユーザ満足度の評価結果から、キャッシュに推薦するコンテンツが限定されることから、キャッシュサイズが小さい場合にキャッシュに限定しない従来の協調フィルタリングと比較して、ユーザの満足度は低下するが、キャッシュ効果が向上した。5.5c では、ユーザの視聴した全てのコンテンツに対する満足度は僅かに低下している。

5.4 ICN における評価結果

シミュレーション時間 (T_s) の中でランダムなコンテンツ推薦を行う探索時間の割合 P_{ex} における, キャッシュヒット率と平均ホップ長, ユーザ満足度の性能評価の結果を示す.

- キャッシュヒット率

キャッシュヒット率 (5.2) の結果を図 5.6 に示す.

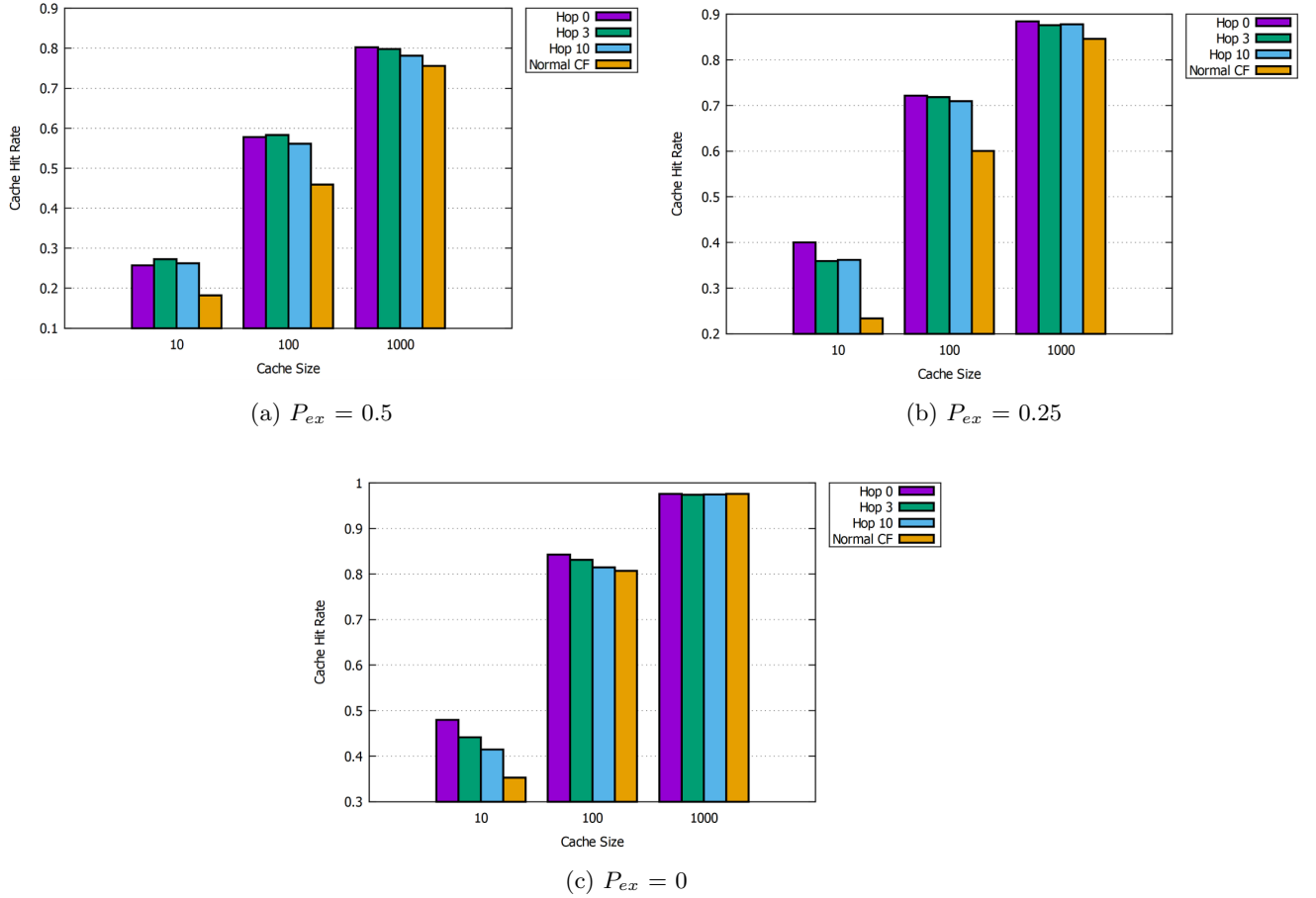


図 5.6: ICN におけるキャッシュヒット率評価

5.6 の結果から, キャッシュに限定しない従来の協調フィルタリングと比較して, 提案方式のキャッシュヒット率が最も高く, キャッシュ効率を向上させた.

- 平均ホップ長

平均ホップ長 (5.3) の結果を図 5.7 に示す.

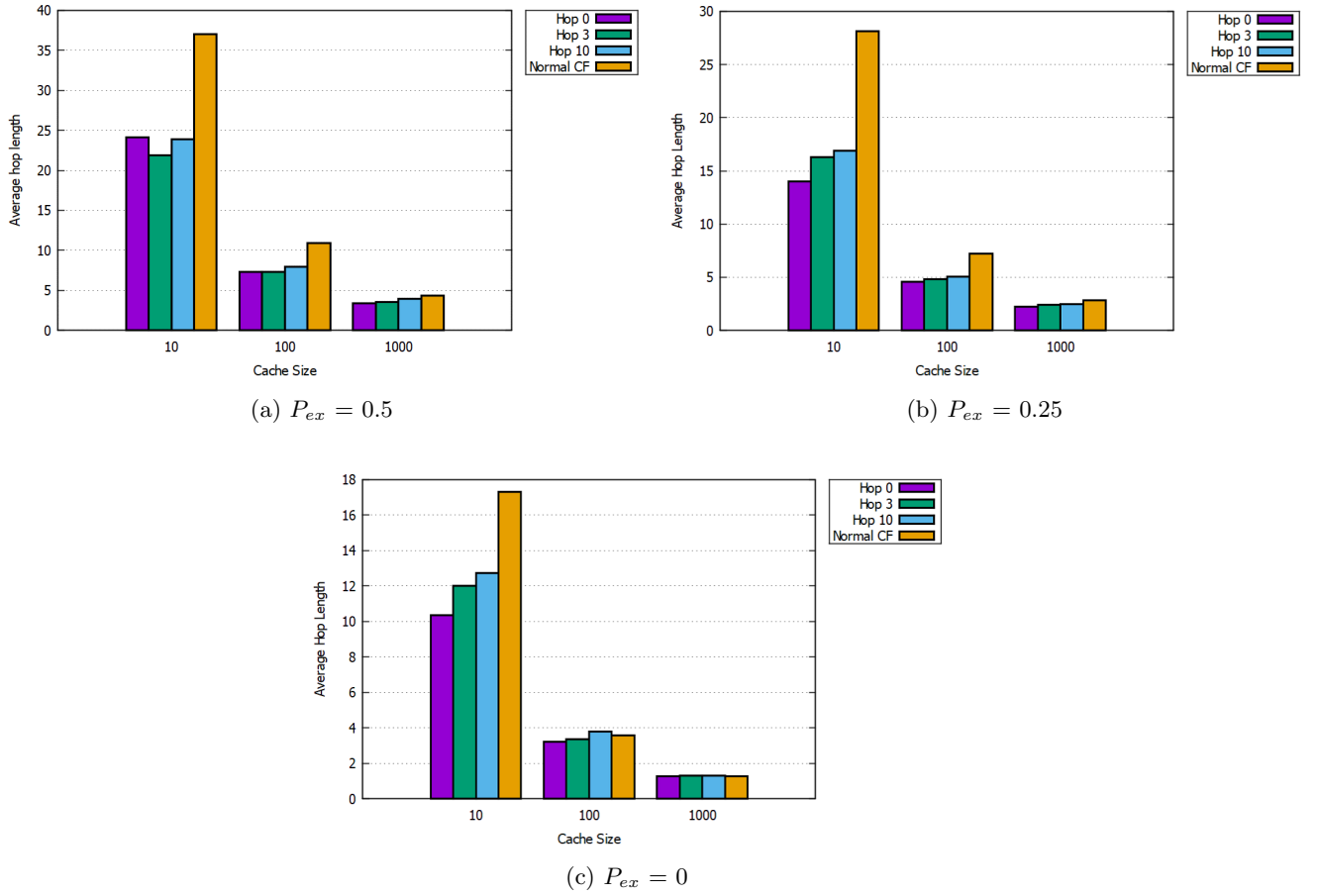
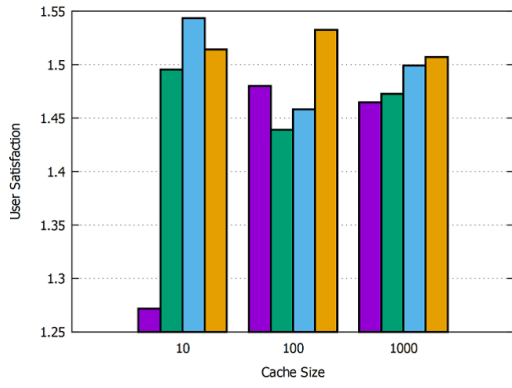


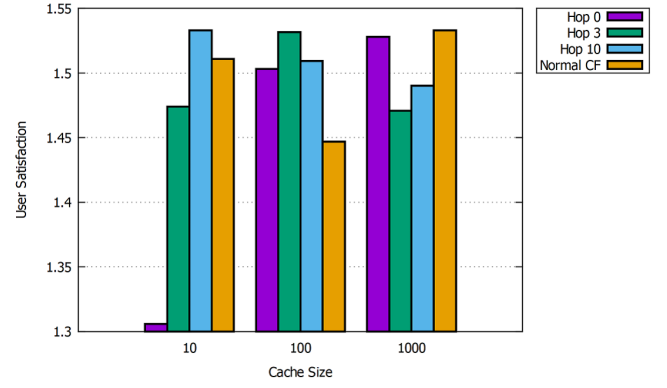
図 5.7: ICN における平均ホップ長の評価

CDN の結果と比較して、要求パケットの配信経路上のルータにコンテンツがキャッシュされていると、そのルータから配信されるため、CDN の場合の結果 (5.3) と比較して、各キャッシュサイズと P_{ex} において平均ホップ長が短くなっている。ICN におけるキャッシュ方式として、配信経路上の全てのサーバに要求コンテンツをキャッシュするオールキャッシュを想定しているため、各サーバには共通のコンテンツがキャッシュされている確率が高い。そのため、推薦に利用するキャッシュサーバ数の拡大より、キャッシュサイズの増加が性能向上に効果的となったと考えられる。また、推薦自由度を拡大することで、推薦コンテンツの視聴確率とともに、平均ホップ長も増加している。ランダム推薦の割合である P_{ex} については、 P_{ex} が小さい程、ホップ長が短くなっていることがわかる。 P_{ex} を大きくすると、ランダムにコンテンツを推薦する頻度が高くなる。ランダムに推薦され、視聴されたコンテンツは人気度やキャッシュに基づいていないため、キャッシュされている可能性が低く、再び要求される期待値は小さい。そしてランダム推薦の度に、要求パケットはオリジンサーバまで送信され、配信経路上の全てのサーバでキャッシュされる。そのため、ランダムに推薦する頻度が少なく、協調フィルタリングによる推薦頻度が多くなるにつれて、ホップ長が短くなったと考えられる。また、ユーザから近いサーバに要求コンテンツがキャッシュされており、配信要求時の経路ノード数が少なかったと考えられる。

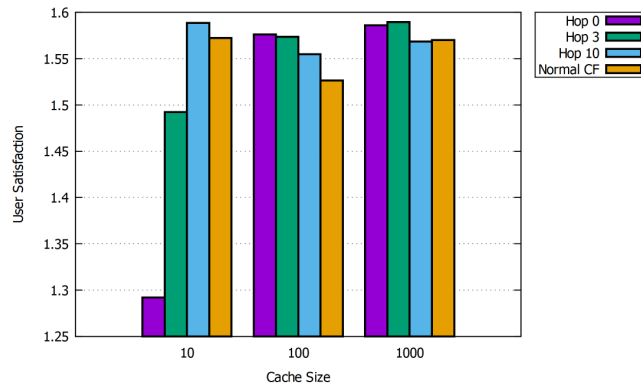
協調フィルタリングによるユーザの満足度 (5.4) の評価結果を図 5.8 に示す。



(a) $P_{ex} = 0.5$



(b) $P_{ex} = 0.25$

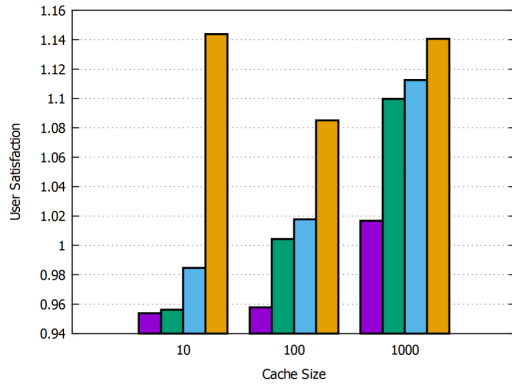


(c) $P_{ex} = 0$

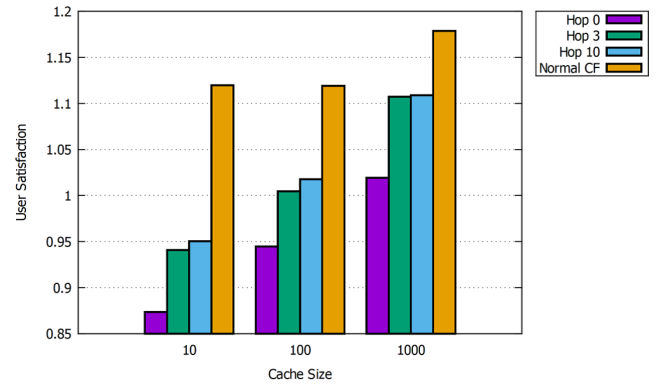
図 5.8: ICN における協調フィルタリングのユーザ満足度評価

推薦候補がキャッシュに限定されるため、キャッシュサーバ内のコンテンツ数が少ないと、ユーザの嗜好を十分に考慮できない。そのため、ユーザから最寄りのサーバに限定され、キャッシュサイズが10のとき、CDNの場合と同様にユーザ満足度が最も低い結果となっている。協調フィルタリングによる推薦の頻度が多い場合、コールドスタートの頻度が多くなり、ユーザはzipf分布に従ってコンテンツを視聴する。そのため、ランダムに推薦する場合と比較して、人気度が高いコンテンツが要求され、協調フィルタリングによる推薦精度が高くなったと考えられる。

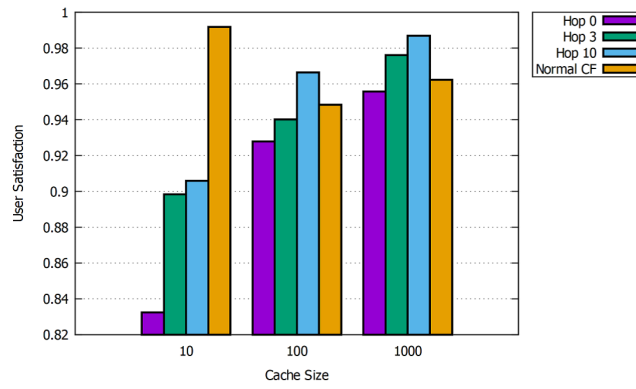
ユーザが視聴した全コンテンツに対する満足度 (5.5) の結果を図 5.9 に示す。



(a) $P_{ex} = 0.5$



(b) $P_{ex} = 0.25$



(c) $P_{ex} = 0$

図 5.9: ICN におけるユーザが視聴した全コンテンツに対する満足度評価

キャッシュに限定しない従来の協調フィルタリングと比較して、図 5.9a, 図 5.9b では提案方式の方がユーザ満足度が低い。しかし、図 5.9c では、キャッシュサイズが 100, 1000 の場合に従来方式のユーザ満足度を上回っている。 $P_{ex} = 0$ の場合、ランダム推薦により、視聴されるコンテンツ数は 0 である。そのため、zipf 分布に従った人気度コンテンツの視聴、または協調フィルタリングによる視聴のどちらかとなる。データが十分に溜まっていない初期の段階では協調フィルタリングは機能しないため、zipf 分布に基づくコンテンツの視聴により、データが蓄積され、サーバにコンテンツがキャッシュされる。キャッシュされたコンテンツは人気度が高いため、キャッシュを活用した提案方式によるユーザの満足度が従来の方式と比較して高くなったと考えられる。

第6章 まとめ

Netflix などのコンテンツ配信サービスにおいて、ユーザの嗜好に沿ったコンテンツを推薦するシステムがコンテンツ要求の大部分を占める重要な存在となっている。また、配信遅延や低ビットレートがユーザのサービス離脱率に影響する。そこで、従来の研究では、ネットワークの負荷軽減や推薦システムの課題であるコールドスタート問題を解決するアプローチがそれぞれ提案されている。しかし、従来の推薦方式ではネットワーク負荷の軽減とコールドスタート問題の両方については考慮していない。そこで本論では、ユーザの満足度および、キャッシュの効果向上を目的とした推薦と、協調フィルタリングのコールドスタート問題を改善するコンテンツ推薦方式を提案した。提案するキャッシュサーバ内のキャッシュされているコンテンツに限定した協調フィルタリングでは、ユーザから物理的に近いキャッシュサーバでキャッシュされているコンテンツを推薦し、その推薦されたコンテンツを視聴することで、ネットワーク負荷軽減とアクセス時間の短縮を図った。また、協調フィルタリングの課題である、コールドスタート問題への対応として、 ϵ -greedy 法と呼ばれる強化学習の推薦方式にキャッシュに限定した協調フィルタリングを組み合わせた。提案方式の有効性を計算機シミュレータにより評価した結果、キャッシュに限定しない協調フィルタリングと比較して、推薦候補のコンテンツをキャッシュに限定したことによって、キャッシュサイズが小さく、最寄りのサーバに限定した場合にユーザの満足度が低下した一方で、キャッシュヒット率が向上した。提案方式により、キャッシュに限定することで平均ホップ長も従来方式と比較して短くなった。ICN では要求パケットの配信経路上のルータにコンテンツがキャッシュされていると、そのルータから配信されるため、CDN と比較して平均ホップ長が短くなっている。データが溜まっていない初期の段階から、協調フィルタリングによる推薦の頻度を増加させることで、コールドスタートが発生し、zipf 分布に従ったコンテンツの視聴によって、人気度が高いコンテンツがキャッシュされている状況下では、ユーザの視聴した全てのコンテンツに対する満足度は僅かに低下しているが、提案方式によって推薦されたコンテンツに対するユーザの満足度は向上し、推薦するコンテンツの質も高いことがわかった。今後の展望として、本論でランダムにコンテンツを推薦した探索において、ユーザからの人気度が高いコンテンツを予測し、そのようなコンテンツを推薦することで、キャッシュに限定することによる提案方式の性能向上が期待できる。また、ICN におけるキャッシュ方式として、本論ではオールキャッシュを使用したのが、キャッシュするサーバを限定することで推薦自由度を拡大した場合の性能向上が図れると考える。

第7章 謝辞

本研究を行うにあたり，ご指導をいただいた上山教授に感謝します．また日常，有益な議論をしていただいた研究室の皆様に感謝します．

参考文献

- [1] C. A. Gomez-Urbe and N. Hunt, “ The Netflix recommender system: Algorithms, business value, and innovation, ” ACM Transactions on Management Information Systems (TMIS), vol. 6, no. 4, pp. 1-19, Dec. 2015.
- [2] Z. Tufekci, “ Youtube, the great radicalizer, ” The New York Times, 2018.
- [3] E. Gourdin, P. Maill´e, G. Simon, B. Tuffin, ”The Economics of CDNs and Their Impact on on Service Fairness”, IEEE Transactions on Network and Service Management, IEEE, 2017, 14 (1), pp.22-33.
- [4] Y. Jianbo, ”Solving Cold-Start Problem in Large-scale Recommendation Engines: A Deep Learning Approach”, IEEE International Conference on Big Data (Big Data) 2016.
- [5] 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.
- [6] D. Liu, and C. Yang, ”A Learning-based Approach to Joint Content Caching and Recommendation at Base Stations”, IEEE Global Communications Conference (GLOBECOM) 2018.
- [7] D. K. Krishnappa, M. Zink, C. Griwodz, and P. Halvorsen, “Cachecentric video recommendation: An approach to improve the efficiency of youtube caches,” ACM Trans. Multimedia Comput. Commun. Appl., vol. 11, no. 4, pp. 48:1–48:20, Jun. 2015.
- [8] H. Zhang, C. Li, Y. Zhao, and C. Zhang, ”Caching Videos with Category Diversity based Recommendation: Model and Algorithm”, IEEE 6th International Conference on Computer and Communications 2020.
- [9] L. E. Chatzileftheriou, M. Karaliopoulos and I. Koutsopoulos, ”Caching-aware Recommendations: Nudging User Preferences towards better Caching Performance ”, IEEE. INFOCOM 2017.
- [10] Movielens, <https://movielens.org/>.

外部発表文献

- [1] 村上悠斗, 上山憲昭, キャッシュを考慮したコンテンツ推薦システム, 電子情報通信学会ソサイエティ大会, 2023 年 9 月
- [2] 村上悠斗, 上山憲昭, キャッシュを考慮したコンテンツ推薦システムのキャッシュヒット率評価, 電子情報通信学会総合大会, 2024 年 3 月
- [3] 村上悠斗, 上山憲昭, キャッシュ状態を考慮したコンテンツ推薦システム, 電子情報通信学会ネットワークシステム (NS) 研究会, 2024 年 3 月