

移動先の予測需要を用いたD2Dキャッシュ制御方式

Device to Device Cache Control Method Using Predicted Demand in Moved Location

常清 瞳与¹ 上山 憲昭²

Makoto Tsunekiyo Noriaki Kamiyama

福岡大学大学院 工学研究科 電子情報工学専攻¹

Graduate School of Engineering, Fukuoka University

立命館大学 情報理工学部²

College of Information Science and Engineering, Ritsumeikan University

1.はじめに

移動端末(MT: mobile terminal)で動画を視聴する形態が一般化したこと、セルラネットワーク(CN: cellular network)のバックホールのトラヒック負荷の急激な増大が懸念されている。バックホールの負荷を低減するため、基地局に設けられたキャッシュから動画コンテンツを配信するモバイルエッジコンピューティングが注目されているが、さらに負荷を軽減する方式として、MTでコンテンツをキャッシュしてD2D(device-to-device)通信で配信することが有効である。しかしMTのキャッシュ容量は有限であるためMTの移動経路上で高い需要が見込めるコンテンツを優先的にキャッシュすることが有効である。

そこで筆者らは、[1]のコンテンツの需要推定方式をD2Dキャッシュ配信に応用し、深層学習のアルゴリズムの一つである長短期記憶(LSTM: long short-term memory)ニューラルネットワークを用いて、移動経路上の他のMTが要求する可能性の高いコンテンツを推測し、MTにキャッシュするコンテンツを選択する方式を提案した[2]。[2]では著名映画10タイトルに関するキーワード検索回数を視聴回数とみなして時系列データを作成し、提案方式の需要推定部分の有効性と、学習モデルの汎用性を確認した。本稿では、MTが移動前に移動後の地点におけるコンテンツの需要予測値の大きなものを優先的にキャッシュに残し、移動後はキャッシュを置き換えないでどの程度のキャッシュヒット率が達成できるかを計算機シミュレーションにより評価する。そして、アメリカのカリフォルニア州(CA: California)とニューヨーク州(NY: new york)の各10タイトルの検査回数を用いた評価を行い、提案方式を用いることで、LRUでキャッシュ置換を行う場合と比較して、移動先でのキャッシュヒット率を向上できることを確認する。

2. 移動先の予測需要を用いたキャッシュ制御

現在時刻 τ に対し、 $\tau < t \leq T$ の任意の時点 t におけるMTの存在場所を c_t とするとき、任意個数 K 個の集合 $(t(k), c_{t(k)}), k = 1, \dots, K$ が与えられ、各時点 $t(k)$ における地点 $c_{t(k)}$ でのコンテンツ m に対する予測需要量 $r_{t(k), c(k)}$ に対し、 $R_m = \sum_{k=1}^K r_{t(k), c(k)}$ を各コンテンツ m に対して算出する。

MTが任意の τ において、任意のコンテンツ m を受信した時、キャッシュに存在するコンテンツ m' の $R_{m'}$ の最小値と R_m を比較し $R_{m'} < R_m$ の場合、 $R_{m'}$ が最小のコンテンツをキャッシュアウトし、 m を新たにキャッシュする。

3. 評価条件

著名映画10タイトルの中から特に欠損データの特に少ない4タイトルを使用して、要求分布を作成し計算機シミュレーションを実行する。欠損部分については、前後のデータの平均値を取ることで補完した。データファイルはTime Slot(TS)1つを8分とし、 $(1440\text{分}(24\text{時間})/8\text{分}) \times 100\text{日} = 18000$ 個の時系列データが格納されている。

また、1日24時間の周期で視聴回数は変化していくため、時系列データを k 個シフト、即ち180の整数倍でシフト(k を180の整数倍に設定)させることでコンテンツ数を増加させる。データファイルの末尾まで達した後は、最終日(TS17820~18000)を巡回させ、1タイトルから合計100個のコンテンツを生成する。したがって、CAとNYの4タイトルの視聴回数から、それぞれ400個(コンテンツ1つ当たりTS0~18000の時系列データ)のコンテンツを作成し、これらを用いて計算機シミュレーションを実施する。

4. 評価結果

CA・NYにおける全TSに渡るコンテンツ毎の要求数を足し合わせ、人気の高い順に並べ実測値と予測値の差を比較することでLSTMの予測精度を評価する。図1はCAおよびNYの実測値と予測値の全TSに渡るコンテンツ毎の総要求数を人気

の高い順に並べ、縦軸を総要求数、横軸をコンテンツのIDでプロットしたグラフである。CA・NY共に実測値と予測値とで小さな差は散見されるものの、グラフ全体を見た時に概ね一致していることから、高精度に予測できていることが確認できる。

また、MTがTS9000より前時刻はCAに滞在し、TS9000以降はNYに移動することを想定する。即ちTS9000まではCAの要求分布に基づき配信要求を行い、キャッシュを作成する。この時、キャッシュ容量以上にキャッシュを挿入する場合に置換処理を行うが、その際にNYのTS9000のみとTS9000~18000の推定要求分布を参照し、NYで高需要となるコンテンツを優先的にキャッシュに挿入する。CAで作成したキャッシュを用いて、TS9000以降はNYで配信要求を行い、要求されたコンテンツがキャッシュに存在した場合ヒット数をカウントし、ヒット率を算出する。この時、NYでは置換処理は行わない。また、提案方式に基づく置換方法と既存のキャッシュ置換アルゴリズムのLRUとを比較することで、提案方式の有用性を評価する。LRUでは、TS0~18000を通して置換処理を行う。

図2はCAでキャッシュを作成し、移動先(NY)で要求されたコンテンツとキャッシュ内のコンテンツとのキャッシュヒット率を提案方式に基づく置換処理とLRUとで比較したグラフである。縦軸はキャッシュヒット率、横軸はキャッシュ容量を表している。図中の青線はCAでの置換処理の際にNYのTS9000以降の要求分布の実測値を参照した場合のヒット率である。橙線と緑線は、CAでの置換処理で参照する要求分布をNYのTS9000以降およびTS9000のLSTMによる予測値とした場合のヒット率である。まず、橙線と緑線を比較すると、橙線が緑線よりもヒット率が高い。次に、青線と橙線を比較すると、Cが100~200の間で、橙線の方がやや精度が落ちているがグラフ全体は概ね一致していることから、実測値と予測値の差は極めて小さく提案方式のLSTMの予測に基づくキャッシュ制御方式の有効性が確認できる。また提案方式はLRUと比較してヒット率が大きくすることが確認できる。以上のことから、提案方式の有用性が確認できる。

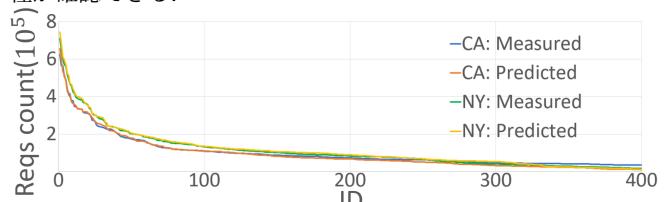


図1: 全TSに渡るコンテンツ毎の総要求数の比較

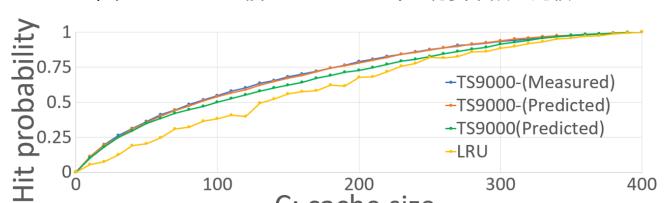


図2: LRUと提案方式のキャッシュヒット率の比較

謝辞本研究成果はJSPS科研費18K11283と21H03437の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] A.Narayanan, et al., Making Content Caching Policies ‘Smart’ using the DeepCache Framework, ACM CCR, 48(5), Oct. 2018
- [2] 常清瞳与, 上山憲昭, 深層学習を用いたD2Dキャッシュ制御方式, 電子情報通信学会ソサイエティ大会, 2021年9月