

移動先の予測需要を用いた D2D キャッシュ制御方式の数値評価

Numerical Evaluation of D2D Cache Control Method Using Predicted Demand in Moved Location

常清 睦与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で、セルラネットワーク (CN: cellular network) のバックホールのトラフィ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。また 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、移動後のコンテンツ要求分布で配信要求した際のキャッシュヒット率を計算機シミュレーションにより評価した。評価結果より、LRU よりも提案方式に基づくキャッシュ置換方式の方が高いキャッシュヒット率を達成可能であることを確認した。しかし [2] の評価に用いたコンテンツの数は 400 と少数であり、また人気の偏りを考慮していない。そこで本稿では、シミュレーションに用いるコンテンツの作成方法を Zipf 分布を考慮した方法に改良することで人気の偏りを考慮し、またコンテンツ数を 8,000 個に増加させた時のキャッシュヒット率を評価する。そして提案方式の有効性を確認する。

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、 $\tau < t < T$ 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 タイトルの中から欠損データの特に少ない 4 タイトル使用して、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。データファイルはタイムスロット (TS: time slot) 1 つを 8 分とし、 $(1,440 \text{ 分} (24 \text{ 時間}) / 8 \text{ 分}) \times 100 \text{ 日} = 18,000$ 個の時系列データが格納されている。

これまでのコンテンツの作成方法は 1 日 24 時間の周期で視聴回数は変化することから、時系列データを k 個シフト (k を 180 の整数倍に設定) させることでコンテンツを作成した。データファイルの末尾まで達した後は、最終日 (TS17,820~18,000) 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した。したがって、CA (California) と NY (New York) の 4 タイトルの視聴回数から、それぞれ 400 個のコンテンツを作成した。これに対して、本稿では Zipf 分布に基づきコンテンツを作成することを考える。まず (1) 式より全ての TS に渡る各コンテンツ k の総要求数 Q_k を与える。

$$Q_k = \frac{\alpha \cdot k^{-\theta}}{\sum_{j=1}^{8,000} j^{-\theta}} \quad (1)$$

ここで (1) 式の α を 1 番人気のコンテンツが任意に与えた数になるように設定し、 θ を 0.8 とした。400 個のコンテンツを 8,000 個に増加させることを想定しているため、 $Q_1 \sim Q_{8,000}$ の計 8,000 の Q_k を与える。また (2) 式より、 k 番人気のコンテンツ k の全ての TS に渡る総要求数が Q_k となるよう、コンテンツ k の各 TS t における要求数 $q_k(t)$ を決定する。

$$q_k(t) = Q_k(t) \cdot \frac{r_k(t)}{\sum_{t=1}^{18,000} r_k(t)} \quad (2)$$

ここで (2) 式の $r_k(t)$ は 4 つの映画の時系列データから作成したコンテンツ k の TS t の要求数を表している。これら 8,000 個のコンテンツ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

MT が TS9,000 より前時刻は CA に滞在し、TS9,000 以降は NY に移動することを想定する。即ち TS9,000 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 TS9,000 のみと TS9,000~18,000 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、TS9,000 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。[2] と同様の評価基準でシミュレーションを実施し、2 通りのコンテンツ作成方法によるキャッシュヒット率を比較し、提案方式の有効性を評価する。

図 1 は CA でキャッシュを作成し、移動先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 1(a) が本稿で構築した 8,000 個のコンテンツを用いた場合の、図 1(b) は [2] で得られた 400 個の結果を示している。図 1 より、提案方式 (Predicted) は、要求数の実測値を用いてキャッシュ判断を行った理想的な場合 (Measured) とほぼ同じヒット率を達成している。また、図 1(a) は図 1(b) と比較してグラフの立ち上がり急になっている。これは、Zipf 分布でコンテンツ毎に人気の偏りを与えることで、特定のコンテンツへの需要が高まることに起因している。図 1(a) においても提案方式 (Predicted) は LRU と比較して高いヒット率を示しており、コンテンツの人気の偏りがある場合においても、提案方式の有効性が確認できる。

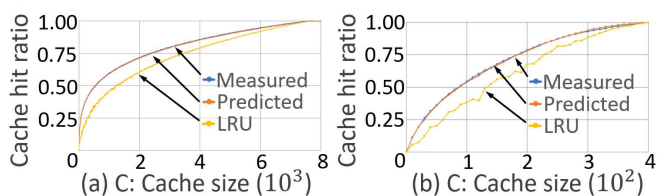


図 1: コンテンツ数の違いによるキャッシュヒット率の比較

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 睦与, 上山 憲昭, 深層学習による移動先の予測需要を用いた D2D キャッシュ制御方式, 信学会 CQ 研究会, CQ2022-54, 2022 年 11 月