低信頼なエッジキャッシュを用いた高信頼キャッシュ法

Reliable Caching Method for Unreliable Edge Caches

吉田　開　　　　上山　憲昭

Kai Yoshida　　　　Noriaki Kamiyama

立命館大学　情報理工学部

College of Information Science and Engineering, Ritsumeikan University

1. はじめに

5G では, スマートフォン等の移動端末で動画を視聴する機会が増加しており, モバイルネットワークのバックホールの負荷が急増し, 通信品質の劣化やネットワークコストの増大が課題となっている. そこでバックホールのトラフィックを削減するために, 無線の基地局にキャッシュサーバ (ES: edge server) を設置し, ユーザに近い ES からコンテンツを配信する Mobile Edge Cache（MEC）が用いられている. しかし,ES の設置数は膨大であるため, コストの低い低信頼な ES が用いられ, 障害による不稼働率の増加が課題となっている. また 従来のキャッシュ制御では, 全ての ES が常時稼働し, 想定したキャッシュ容量が常に利用できることを想定しているが, 低信頼な ES を想定したキャッシュ制御が必要である. そこで低信頼な ES を想定した既存研究 [1] では, エッジリソースの信頼性の低さに対処するために, erasure coding を用いて ES の可用性を高める方式を提案しているが, コンテンツを事前に ES に配置することを想定している. しかし実際の ES では LRU 等を用いた置換による動的制御が一般的である. そこで本稿では, 低信頼な ES において LRU を用いた置換制御を想定した ES の不稼働を考慮した上で, ES からのコンテンツ取得可能性を向上させる erasure coding を用いた ES 挿入法を提案する.

2. 低信頼なキャッシュ資源を持つ分散協調キャッシングシステム

[1] で Li らは, 低信頼なキャッシュ資源を持つ分散協調キャッシングシステム（DEC）を提案した. DEC では 各ファイルの長さは全て M ビットと仮定し [2], 各ファイルを K 個のチャンクに分割し, RS 符号などの MDS 符号 [3] によって, K 個のチャンクから x 個の冗長チャンクを生成し, これらを N 個の ES に分散配置する. そして ES から K 以上のチャンクが取得できれば, 元のファイルを復元可能である. 各ファイルのチャンクを異なる ES にキャッシュすることで, ファイル取得の待ち時間, コアネットワークの輻輳, リモートサーバの負荷を軽減する. [1] では ES から取得できるデータの総量が最大化するよう, 各コンテンツの静的な要求頻度に対し, 各 ES に配置するチャンクを最適設計する.

3. 提案方式

本節では ES へのチャンク挿入アルゴリズムを提案する. あるエリアに存在する N 個の ES 上へのチャンクキャッシュを想定し, 置換方式は LRU を使用する. 各コンテンツ m の, N 個の ES 上にキャッシュされるチャンク数の目標値 T_m が [1] の設計値に一致するよう, コンテンツ m のキャッシュへの挿入確率 f_m を事前に設定する. 各 ES の稼働率を r とすると, 各コンテンツ m のチャンクが目標値 T_m の個数, ES 上にキャッシュされているための, ES の目標ヒット率 $\hat{h}_m$ は (1) 式で与えられる.

$$T_m = Nr\hat{h}_m \tag{1}$$

次に目標ヒット率 $\hat{h}_m$ と実際のヒット率 h_m の差が最小となるようキャッシュ挿入確率 f_m を最適設計する。各コンテンツ m の要求比率を q_m とすると, 最小化を目的とする目的関数は (2) 式となる.

$$min \sum q_m(h_m - \hat{h}_m)^2 \tag{2}$$

以下の遺伝的アルゴリズムで本最適化問題の近似解を得る.

1. キャッシュ挿入確率 f_m として離散的な 10 個の区間を考え, 全コンテンツに対して, 各コンテンツ m に対して設定する f_m の区間の値を要素としてもつベクトルを各遺伝子をして考え, 初期の複数の遺伝子をランダムに生成
2. 各遺伝子の適応度を (2) 式の逆数で算出し, 適応度の高いものをトーナメント選択で選択
3. 選択した遺伝子をもとに次世代の遺伝子を生成
4. 一定の確率で, 遺伝子内の f_m に対して交叉または突然変異を適用
5. step2〜step4 を G 世代まで反復
6. step5 で最後の世代まで反復し, 生成遺伝子の中で最も適応度の高い遺伝子をキャッシュ挿入確率 f_m として決定

4. 性能評価

ES の個数 $N = 5$, 各 ES の容量 $s = 5$, 各 ES の稼働率 $r = 0.8$, 最低必要チャンク数 $K = 2$, コンテンツ数 $m = 30$ としパラメタ $\theta = 0.8$ の Zipf 分布に従いコンテンツの要求を発生させた. また遺伝的アルゴリズムのパラメタを個体数 200, 世代数 200 で設定した.

図 1(a) に各コンテンツ m の GA で算出された最適キャッシュ挿入確率 f_m を, また図 1(b) に GA で算出された f_m を用いたときのコンテンツ m のヒット率 h_m と目標ヒット率 $\hat{h}_m$ を, 各々 m に対してプロットする. キャッシュ挿入確率 f_m を提案方式で最適化することで, 実際のヒット率 h_m と目標ヒット率 $\hat{h}_m$ の差異は小さく, 適切に f_m を設定できることが確認できる. また図 2 に N の 2 つの場合について各々, 全コンテンツの平均取得成功率 (ES からのチャンク取得のみでコンテンツを復元できた割合) を, 提案方式を用いて f_m を設定した場合 ($f_m(GA)$) と, キャッシュミス時に常にコンテンツをキャッシュした場合 ($f_m(1.0)$) について, 各々, 各 ES の稼働率 r に対してプロットする. 提案方式を用いることで, 単にコンテンツをすべて ES にキャッシュする場合と比較して, 平均取得成功率を大きく改善できることが確認できる.

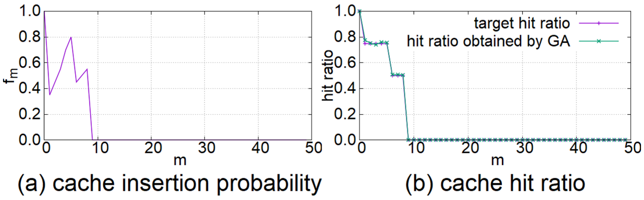

図 1: (a) Cache insertion probability f_m, against m, (b) comparison between cache hit ratio obtained by proposed GA and target cache hit ratio

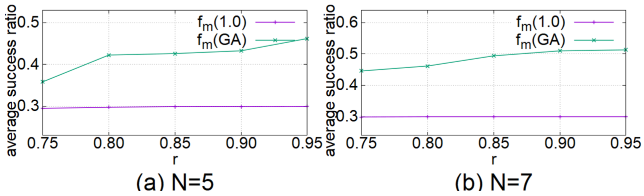

図 2: Average success ratio against reliability of each ES

謝辞

本研究成果は JSPS 科研費 21H03436 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] Y. Liu, et al., Distributed Cooperative Caching in Unreliable Edge Environments, IEEE INFOCOM 2022

[2] Y. Zeng, et al, “Privacy-preserving distributed edge caching for mobile data offloading in 5g networks,” in 2020 IEEE ICDCS 2020

[3]A. Soro and J. Lacan, “Fnt-based reed-solomon erasure codes,” IEEE CCNC 2010