

# 共起度を用いた Web キャッシュ制御方式

Web cache control method using co-occurrence degree

他力 誠人<sup>1</sup> 三角 真<sup>1</sup> 上山 憲昭<sup>1</sup>  
Masato Tariki Makoto Misumi Noriaki Kamiyama

福岡大学 工学部 電子情報工学科<sup>1</sup>  
Faculty of Engineering, Fukuoka University

## 1 はじめに

Web ページの待ち時間を削減する方法として、Web ページを構成する複数のオブジェクトを並列取得することで、Web ページ表示の待ち時間を削減する HTTP/2 や HTTP/3 が標準化され広く普及している。HTTP/2 や HTTP/3 の Web ページの表示待ち時間削減効果は、同一の配信サーバから取得するオブジェクト集合に対して有効な反面、多数の配信サーバからのオブジェクトの取得については各配信サーバにコネクションを作成しなければならぬため効果が低い。これに対して、[1] では、任意の個数のオブジェクト組が複数の WEB ページ内で出現することを示す共起という概念を提案し、オブジェクト組が出現する WEB ページの数を共起度と定義し、約 8,000 の Web ページ URL から取得したオブジェクトについて共起度の分析と、共起度のキャッシュ制御法への適用可能性を示している。

そこで本稿では、HTTP/2 や HTTP/3 の並列取得効果の向上を目的とし、複数の HTTP/2 や HTTP/3 に対応したキャッシュサーバが存在するとき、共起を考慮したキャッシュ制御を行うことで、共起するオブジェクト組を構成するオブジェクトを、単一のキャッシュサーバに可能な限り集約する手法を提案する。本稿では、提案手法による効果をオブジェクトの集約度合いを指標として評価する。

## 2 提案方式

提案方式では、挿入時、排出時に共起を考慮したキャッシュ制御方式を提案する。

**挿入ポリシー** 複数のキャッシュサーバが存在するとき、オブジェクトを挿入するキャッシュサーバを選択する必要があるが、HTTP/2 や HTTP/3 の並列取得効果を得るためには、共起オブジェクト組は同一のキャッシュサーバに配置されていることが望ましい。そこで、オブジェクトをキャッシュサーバに新たにキャッシュするキャッシュサーバを選択するとき、既にキャッシュされているオブジェクトに、新たにキャッシュするオブジェクトが加わることによって、新たに発生する共起オブジェクト組の数をキャッシュサーバ毎に求め、その数が最も多いキャッシュサーバを新たなオブジェクトをキャッシュするキャッシュサーバとして選択する。

**排出ポリシー** キャッシュサーバのキャッシュがあふれたとき、キャッシュされたオブジェクトを排出する必要があるが、提案手法では、共起するオブジェクト組を可能な限りキャッシュ上に残す手法を提案する。提案手法において、キャッシュから排出するオブジェクトは、共起を考慮した LRU を提案する。

提案手法では、単一のオブジェクトに加えて共起オブジェクト組単位でアクセス順のリストを管理する。LRU を共起オブジェクト組で管理すると、キャッシュ内のあるオブジェクトは、LRU リストの複数の要素から参照されるため、キャッシュされているオブジェクト毎に参照数を管理する。共起オブジェクト組を構成するいずれかのオブジェクトのリクエストが発生したとき、共起オブジェクトの組を構成するオブジェクト全てのリクエストが発生しなかった場合も、その共起オブジェクト組のアクセス順を更新する。

あるオブジェクトあるいは共起オブジェクト組のリクエストが発生したとき、他の共起オブジェクト組によってそれらのオブジェクトは既にキャッシュ内には存在し

ているものの、LRU のリストにはリクエストのあったオブジェクトあるいは共起オブジェクト組が存在しない状況が発生する。そのようなとき、LRU のリストにリクエストされたオブジェクトあるいは共起オブジェクト組を追加し、キャッシュされているオブジェクトの参照数をインクリメントする。

排出するオブジェクトを決定するとき、LRU に従い最も使われていないオブジェクトまたは共起オブジェクト組を取り出し、そこに含まれるオブジェクトについて、キャッシュの参照数をデクリメントする。参照数が 0 になったオブジェクトは、キャッシュから削除する。これを、必要な数の空き領域が得られるまで繰り返す。

## 3 評価

キャッシュ排出時の共起考慮の効果を確認するために集約度合いの平均値で提案手法の有効性を評価する。HTTP/2 や HTTP/3 の Web ページの表示待ち時間削減効果は、同一の配信サーバから取得するオブジェクト集合に対して有効であることから、 $r$  回目のリクエストで Web ページ  $w$  の URL をリクエストしたときの、キャッシュヒットしたオブジェクトの特定のキャッシュサーバへの偏り度合いを表す、集約度合い  $C_r$  を  $C_r = \sum_{s \in S} O_{w,s}^2 / \sum_{s \in S} O_{w,s}$  と定義する。ここで、 $S$  は全キャッシュサーバ集合を、 $s$  は各キャッシュサーバを、 $O_{w,s}$  はすでにキャッシュ  $s$  に存在する Web ページ  $w$  を構成するオブジェクトの数を表す。本稿では、4 台のキャッシュサーバで、Web ページのアクセス数のランキングを公開している Alexa の Web ページから取得した、16 のカテゴリ毎の上位 500 の Web ページに含まれるオブジェクトを対象として、Alexa の人気順に従い Zipf 分布のパラメータ  $\theta$  に従いリクエストを発生させ、集約度合い  $C_r$  の平均値を単純な LRU と比較した。図 1 に示す通り、提案手法は単純な LRU 比較し共起するオブジェクトの集約度合いが向上することが確認された。

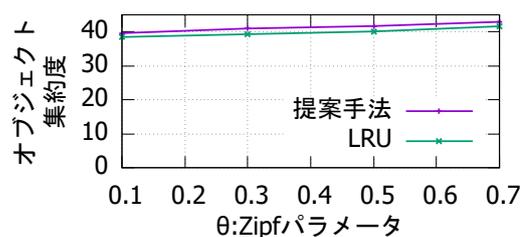


図 1 Zipf 分布のパラメータ  $\theta$  に対する平均集約度合い

## 4 まとめ

本稿では、HTTP/2 や HTTP/3 対応キャッシュサーバアクセス時の Web ページの表示待ち時間削減を目的とし、複数のキャッシュサーバで共起を考慮したキャッシュ制御法を提案し、その集約度合いを評価した。今後は、共起を考慮したキャッシュ挿入ポリシーの適用を行い、オブジェクトの集約度合いの評価を行う。

## 謝辞

本研究成果は、SCAT 研究費助成 180047 の援助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] 桜井洗輝, 上山憲昭, 中尾彰宏, “Web オブジェクトの共起度分析,” 電子情報通信学会技術研究報告 (ICM2018-63), pp.77–82, March 2019.