

アプリケーションレベルの情報をを用いた EC サイトに適したキャッシング方式に関する研究

21T324 東原慶一良（最所研究室）

本研究では、従来のキャッシング方式に加え EC サイト特有の情報を制御に用いることで、キャッシュを活用しつつ古いコンテンツの提供を抑えるキャッシング方式を提案している。本稿では、提案方式を用いたシステムの開発及び評価について述べる。

1. はじめに

近年、Web サービスを用いた商品販売サービス (EC サイト) が普及している。EC サイトはセールやプロモーションイベント時に突発的にアクセス集中が発生する場合がある。アクセス集中が発生し、レスポンスタイムが長くなった場合、EC サイトでは売り上げの低下に繋がる。このレスポンスタイムの遅れへの対策として、Web コンテンツを提供する Web サーバの前段に設置したリバースプロキシを用いて Web コンテンツをキャッシュし、Web サーバの負荷を低減する方法がある。しかし、動的に生成される EC サイトでは、コンテンツの内容が頻繁に更新される場合がある。そのため、単純にコンテンツをキャッシュしてしまうと、実態と合っていないコンテンツを提供するケースが起き、ユーザへ不満を与えることとなる。本研究ではこのような問題を解決できる EC サイトに適したキャッシング方式を提案し、提案方式を用いたシステムを開発した。

本稿では、提案方式を用いたシステムの実装、評価実験について述べる。

2. 課題

従来のキャッシング方式は、データの整合性チェックに基づく制御が行われている¹⁾。この制御にはリクエスト毎に整合性チェックを行う制御や有効期限を設定し、一定期間整合性チェックを行わない制御がある。キャッシュの有効期限を長く設定する場合には、Web サーバへのリクエストが減少し、負荷軽減の効果が高まる。しかし、同時にデータの不整合が起きるリスクが高まる。一方で、有効期限を短く設定する場合には、データの不整合を防ぎやすくなるが、整合性チェックのためのリクエストが増加することで、有効期限を長く設定する場合と比較して Web サーバの負荷が増大する。このように、キャッシュの更新頻度や有効期限を決める際には、Web サーバの負荷低減とデータの不整合の防止を両立させることは難しく、トレードオフになってしまうという課題が存在する。

また、動的に生成される EC サイトでは、コンテンツの内容が頻繁に更新される場合があり、EC サイトの商品が売り切れたタイミングでユーザがキャッシュされた販売ページを参照すると、実際には在庫がない商品が「在庫あり」と表示される。この時、ユーザにとっては、在庫の有無、購入の可否という情報が重要であるが、その情報を提供すること

が出来ていないという問題が課題となる。

3. 提案方式の概要

EC サイトにおける在庫情報のようなアプリケーションレベルの情報も用いるキャッシュ制御方式を提案する。例えば、EC サイト内の特定の商品が売り切れるタイミングでは Web ページのコンテンツが在庫有りの表示から売り切れの表示へと更新される。その際にキャッシュが有効期限切れになっていないことが要因でユーザに更新前の古いコンテンツを提供し続けることを防止する必要がある。そこで、売り切れが発生するタイミングを予測し、予測したタイミングに合わせてキャッシュが有効期限切れになるように、キャッシュの有効期限を短くする。また、有効期限を短くしたにも関わらず、売り切れ表示へと更新される直前に在庫有りの状態のコンテンツが再度キャッシュされた場合、キャッシュが有効期限切れになるまでの間はコンテンツの不整合が発生してしまう。そのため、不整合が発生する期間をより短くする場合、売り切れが予測されるタイミングの一定数秒前までにキャッシュを無効化する制御も行う。

4. 設計と実装

提案方式を用いたシステムの概要を図 1 に示す。ユーザが Web サイトにアクセスする際、リバースプロキシにリクエストを送信する。Web サイトに対して来たリクエストは、リバースプロキシがキャッシュを保持していれば、それを用いてレスポンスが返される。本提案のキャッシュ制御に必要な機能を以下に示す。

1. EC サイトの状態監視機能
2. 監視した EC サイト情報の処理・送信機能
3. キャッシュ制御判断機能
4. キャッシュ制御機能

機能 1, 2 を Web サーバに、3, 4 をリバースプロキシに組み込む。1 は、EC サイト内の情報を定期的

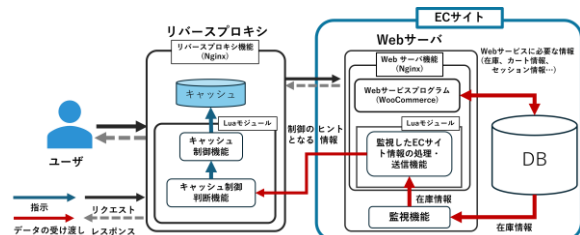


図1 システムの概要

に取得する機能である。Web サーバ上のエージェントとして EC サイトが持つデータベースに対して定期的にクエリを送信し、提案方式で用いる情報として商品の在庫情報を定期的に取得する。2 では、監視機能で取得した EC サイト情報を、3 で行う制御判断に用いる形式に変換する処理を行い、Web サーバからリバースプロキシに送信する。3 では 1, 2 を用いて取得した EC サイトの状態と、現在のキャッシュの状態を用いて実際にキャッシュ制御を行う必要があるかどうか判断し、キャッシュの有効期限を変更する必要がある場合は有効期限の変更処理を、キャッシュを無効化する必要がある場合はキャッシュしないようにする処理を4に依頼し実行する。

実装においては1ではECサイト内の特定の商品の在庫数を定期的に取得し、2で最小二乗法を用いて在庫の変動を算出している。その後、3では現在のキャッシュの状態と2で得た情報を照らし合わせ、売り切れが予測されるタイミングに合わせてキャッシュの有効期限を調整する。この時、現在のキャッシュが有効期限切れになってから売り切れが発生するまでにかかる時間を元に、次にセットされるキャッシュの有効期限を決定してある。また、売り切れが予測されるタイミングの一定数秒前までにキャッシュを無効化する制御を行う場合、「予測した売り切れ時点の何秒前か」を指定した正の値を閾値として設定する。逆に、売り切れ後にキャッシュが有効期限切れになることを許容する場合は、閾値として負の値を設定する。そして、予測した売り切れ時点から閾値を差し引いた時刻までにキャッシュが有効期限切れになるように4にて制御を行う。さらに閾値が正の場合は実際の売り切れ時刻までキャッシュの無効化も4にて行う。

5. 評価実験

EC サイトへのアクセス及び購入処理を複数行い、①実際には在庫がない商品が「在庫あり」と表示される割合を抑えることが出来ているか、②キャッシュを活用することが出来ているかを確認する。①を確認するために交差検証を用い、キャッシュ上での在庫の有無を予測値、実際の在庫の有無を正解値として真陽性と偽陽性から算出する適合率を用いた。②を確認するために、キャッシュヒット率と、Web

サーバの CPU 使用率がキャッシュを用いていない場合と比較して低減出来ているかの2点を用いた。また、提案方式を用いたシステムが適切に動作しているかを確認するための機能評価を行った。

5.1 機能評価

図 2(左)は売り切れが予測されるタイミングと同時にキャッシュが有効期限切れになるように制御を行った場合の動作結果である。101 秒時点、109 秒時点では、売り切れまでの予測時間を元にキャッシュの有効期限が更新されている。そして、112 秒の時点で売り切れが発生し、実際にキャッシュが有効期限切れになったのが113秒の時点であったため、不整合が発生したのは1秒以内であった。以上の結果から、提案方式を用いたシステムが設計通り適切に動作していることが確認できる。

5.2 有効性評価

提案方式の有効性を確認する。図 2(右)は、提案方式を用いなかった従来方式の場合と用いた場合を比較した評価実験結果となっている。閾値0の場合、従来方式の有効期限が20秒の場合のキャッシュヒット率とほぼ等しい結果となっているが、適合率は約8%高くなっている。閾値10の場合、不整合が発生していないため適合率が100%となっている。

6. 考察

評価実験の結果から提案方式を用いることで、キャッシュヒット率と適合率のトレードオフを緩和できること及び、コンテンツの不整合が発生する期間をより短くする場合にも対応できることが確認できた。また、提案方式でのリバースプロキシ及びWebサーバのCPU使用率は提案方式を用いない従来方式の場合と比較してほとんど変わらないことから、提案方式がECサイトを構成するシステムに与える影響がほとんどないことが確認できた。

7. おわりに

本研究で提案した方式は、動的に生成されるECサイトに適したキャッシングに寄与していると言える。今後は、大規模なECサイトにおける評価が課題となる。

参考文献

- 1) Mozilla Developer Network (MDN), HTTP Caching, <https://developer.mozilla.org/ja/docs/Web/HTTP/Caching>

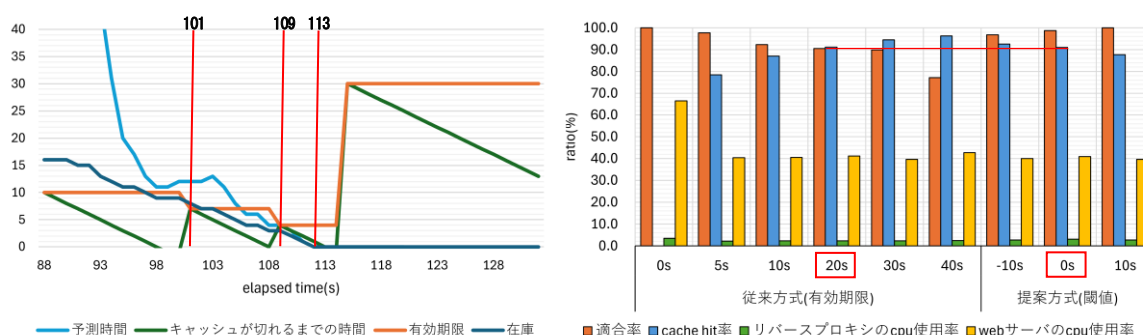


図2 実験結果