

構造化オーバーレイにおけるデータ間の関連性による マルチクエリに対する応答の高速化

Speeding Up of Response to a Multi-query Using Relation Estimation of Data in Structured Overlay Networks

小泉悠介
Yusuke Koizumi

渡部康平
Kohei Watabe

中川健治
Kenji Nakagawa

長岡技術科学大学 大学院 工学研究科
Graduate School of Engineering, Nagaoka University of Technology

1 概要

構造化オーバーレイネットワーク [1] では、負荷分散を行うためにデータがネットワーク上の各ノードへと分散して保存される。このネットワークにおいて、特定の条件に該当するデータを複数取得するという要求 (マルチクエリ) を行ったとする。その場合、分散して保存したデータを取得するため複数のノードを探索することが必要となるため、多くの時間を要するという問題点がある。そこで本研究では、分散して保存されている各データ間の関連性に基づきデータの配置方法を工夫することで、1回のクエリで複数のデータを取得できるようにする。これにより、マルチクエリに対する応答を高速化する手法を提案する。

2 提案法について

提案法では同時に要求されたデータすべてを同じノードへ複製して保存することによってマルチクエリに対する応答を高速化する。

データの複製および保存はキャッシュによって実現する。各ノードにおいて複数のデータを同時に取得した際に、どのデータが同時に要求されたかを確認し、その後、同時に要求されたデータを取得元のノードへキャッシュとして保存する。キャッシュは参照されてからの経過時間が長いものから削除することによって、関連性が低く不必要なキャッシュは自動的に削除することができる。

以下に提案法の処理手順を示す。

1. 任意のノードにおいて、マルチクエリを出した結果としてデータを取得した際に、取得したデータの内容と取得元ノードの所在全てを一時的に保持する。
2. データを取得したノードが取得元ノード全てに対して、取得したデータ全てを送信してキャッシュを作る指示を出す。
3. 指示を受けたノードは受け取ったデータそれぞれが自ノードの保存領域内に存在するかどうかを確認し、存在しないデータをキャッシュとして保存し、キャッシュ管理用リストに追加する。保存されているキャッシュの個数がキャッシュの容量を上回った場合、参照されてから最も時間が経過したキャッシュを削除する。

3 シミュレーション

構造化オーバーレイネットワークを構成するノード数を変化させたとき、全データ取得までのクエリの平均所要ホップ数がどのように変化するかを、提案法を用いない場合と用いた場合それぞれについてシミュレーションを行い測定した。各都道府県の気象情報を1日ごとにデータとして記録し、1週間分のデータを保存している構造化オーバーレイネットワークを考える。このネットワークに対するクエリにおいて、1つの県における1週間分のデータをまとめて要求するマルチクエリに着目して全データ取得までのクエリの平均所要ホップ数の変化を測定した。ネットワーク全体で保存してある合計データ数を329[個]として、ノード数を10[個]から100[個]まで変化させた。このときマルチクエリ処理を100[回]行い平均所要ホップ数を算出した。また、初期条件の影響をなくすために、測定前にあらかじめマルチクエリ処理を1000[回]行った。また、データの取得要求は形状パラメータ $s = 1.4$ のジップ分布に従うと仮定する。

図1にシミュレーション結果を示す。図1より提案法を用いない場合に比べ、提案法を用いた場合の方が平均所要ホップ数が減少していることがわかる。

4 今後の方針

データ間の関連性をバイズ推定や最尤推定といった方法を用いて推定を行い、データの配置方法へと反映することでさらなる応答の高速化を目指す。

参考文献

- [1] Ion Stoica *et al.*, “Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications”, IEEE-ACM Trans. Netw., VOL.11, NO.1, 2003.

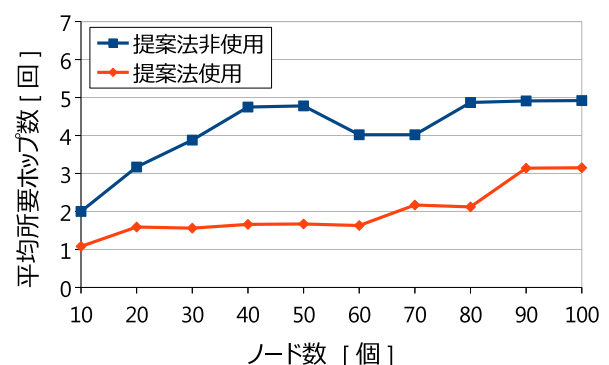


図1 シミュレーション結果