

P2P ネットワークにおけるノードの離脱を考慮した複製配置手法

谷口貴柁 † 中村健二 ‡ 小柳滋 †

† 立命館大学情報理工学部

‡ 大阪経済大学情報社会学部

1 はじめに

データの分散管理の手法として P2P ネットワークが以前から注目され、多くの研究がなされている。P2P ネットワークにて情報検索をする場合はサーバが存在しないため、各ピアが協調して検索する仕組みが必要である。

P2P ネットワークの構造には構造型と非構造型に大別でき、前者では DHT 検索、後者では flooding 検索が用いられる。DHT 検索ではネットワークへの参加に数学的規則を用いデータ管理するピアを明確化させ、検索時に分散ハッシュテーブルを用いる。そのため、検索効率が高い利点がある一方で、ピアの参加離脱によって検索性能が低下する。

一方 flooding は隣接するピア全てにクエリを送信する方法であり、目的のデータを保持するピアがない場合クエリ数が増大するという課題がある。

これらの特性を踏まえ、本研究ではピアの離脱が発生する状況下の非構造ネットワークにおいて、検索クエリ数が増加する課題の解決手法を提案する。

2 関連研究

flooding 検索において検索クエリ数が増大する問題を解決する手法として、複製を配置する手法、人気度に応じて index を配置する手法 [1] がある。複製を配置する手法としては Owner Replication[2]、Path Replication[2] がある。Owner Replication では検索が成功した際に検索を要求したノードにのみ複製を配置する手法である。複製数が少ないため負荷が小さくなるが検索クエリの削減の効果を出す即効性に欠ける。Path Replication では検索が成功した際に、検索要求ノードからデータ保持ノードに至るまでの経路上の全てのノードに複製を配置する手法である。検索クエリの削減の即効性が見込まれるが負荷が増大する。

また、川崎 [1] らの手法では検索パスの中央にコンテンツの保有ノードを示すインデックスを配置し、参照数が閾値を越えた場合、複製を配置する事で検索の効率化を図っている。

複製配置手法は検索クエリ数が増大する問題を解決する事は可能であるが、検索が成功する度に複製を配置するため、ネットワーク負荷が増大する。またこれらの手法はノードの参加離脱を考慮しておらず、実際の

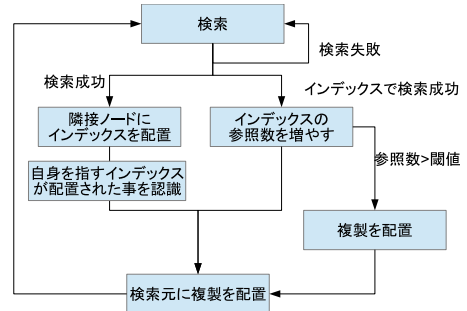


図 1: 処理の流れ

P2P ネットワークでの実用性が担保されていない状況である。

3 提案手法

3.1 提案手法の概要

複製配置手法では不要な複製が配置されることでネットワーク負荷が増大するという課題がある。そこで、本研究では既存研究 [1] と同様にインデックスを配置することに加え、インデックスの誘導先のノードがインデックスを配置しているノードの離脱（以下、離脱ノード）を監視する手法を提案する。離脱を確認した場合は、アクセスが多いアイテムのインデックス及び誘導先のノードを示していたインデックス別のノードに配置することで、無駄な複製及び離脱による検索効率の低下を抑制することを目的とする。

3.2 提案手法の流れ

提案手法ではネットワーク上でアイテム検索を行い、検索が成功した際、Owner Replication と同様に複製を配置する。一定回数毎にノードを離脱させる。提案手法の処理の流れを図 1 に示す。図 1 における検索成功はインデックスを参照せずアイテムを発見したことであり、インデックスで検索成功はインデックスを用いてアイテムを発見したことである。

3.2.1 インデックスと複製配置処理

検索が成功した場合、検索元の隣接ノードにインデックスを配置する。インデックスには検索に成功したアイテムとその所持ノードとを記録する。その際に記録したことを所持ノードに通知し、所持ノードはインデックスを配置したノードを記憶する。記憶する事で離脱の監視が可能になる。検索時にインデックスは参照回数を増加させ、参照回数が閾値を越えた場合、人気アイテムと判断し複製を配置する。

Replication Method Considering Disconnection on P2Pnetwork
 †Takamasa TANIGUCHI †Kenji NAKAMURA †Shigeru OY-ANAGI
 †College of Information Science and Engennering Ritsumeikan University
 ‡Faculty of Imformation Technology and Social Science Osaka University of Economics

表 1: 実験条件

| 項目 | 値 |
|----------------|---------|
| アイテム数 | 3,952 個 |
| ノード数 | 6,000 個 |
| 最大リンク数 | 5 本 |
| TTL | 5 |
| 検索回数 | 50 万回 |
| ノードが保有可能な複製の上限 | 50 アイテム |
| 離脱したノードの数 | 1,000 |

表 2: 実験結果

| | baseline | Owner | Path | 提案手法 |
|-----------|----------|-------|-------|-------|
| 平均ホップ | 3.83 | 3.28 | 2.16 | 2.55 |
| 平均クエリ | 34.88 | 27.68 | 17.48 | 14.95 |
| 検索成功率 (%) | 79.8 | 83.10 | 86.0 | 88.8 |
| 複製数 | 0 | 390k | 931k | 395k |

3.3 ノードの離脱処理

ノードの離脱は、一定回数の検索を行った際に発生させる。インデックスに記録されたノードは、記憶したインデックスを保有するノードの離脱の有無を定期的に確認する。その際、ノードの中で最もアクセスされたアイテムと、そのアイテムにアクセスする際に最も多く経由されるノード（以下、経由ノード）とを記録する。

ノードが離脱済みの場合には離脱ノードが保有していたインデックスのうち自身を指すインデックス、及び離脱ノードの中で最も多くアクセスがあったアイテムを周辺ノードから探索する。探索したアイテムのインデックスを経由ノードに配置し、ノードの離脱による人気があるアイテムの消失を回避する。

4 実験

4.1 実験内容

各ノードは検索アイテムと所持アイテム群の2種類のアイテムを保有する。検索はランダムに選ばれたノードが任意のアイテムを1つ選択し、隣接する全てのノードに検索クエリを送信する。

クエリを受信したノードは、自身の所持アイテム群、複製アイテム群、インデックスの中に検索アイテムがあるかを確認する。アイテム群に存在する場合は検索成功とする。またインデックスに存在する場合はクエリをインデックスが示すノードに誘導する。検索アイテム及びインデックス内に検索アイテムを持つノードが存在しない場合は、TTLが0になるまで検索を続行する。検索アイテムを発見できなかった場合検索失敗とする。

本実験ではbaseline(flooding検索), Owner Replication, Path Replication, 提案手法を用いて50万回の検索を行い、平

均のクエリ数、ホップ数、検索成功率、複製配置数を評価する。本実験ではアイテム群としてMOVIELENS[3]のデータセットを用いる。ユーザをノードとし、ユーザが見た映画を半分に分け、検索アイテムと所持アイテムとする。実験条件を表1に示す。表1においてアイテム数はMOVIELENSのタイトル数の事であり、ノード数はMOVIELENSのユーザ数である。最大リンク数は各ノードのリンク数が5以下であることを示している。

4.2 実験評価

実験結果を表2に示す。表2より提案手法をbaselineと比較し平均ホップ数が約33.5%、平均クエリ数が約57.3%減少し、検索成功率は約9%向上している。Owner Replicationと比較するとホップ数は同等だが、クエリ数は30.3%、検索成功率は約14%向上している。複製数は12%増大しているがこれはアクセスされた回数が多いアイテムを複製しているためである。Path Replicationと比較すると平均ホップ数が17.8%増大しているが、平均クエリ数が約14.5%減少し、検索成功率は約2.8%向上し、複製数は約42.5%減少した。平均ホップ数が増大した理由としてはインデックスで誘導する際に1ホップ多くなることが考えられる。一方で平均クエリ数や複製数は減少しているため無駄な複製が少なくなっている。これらの結果から本提案手法はノードの離脱が発生する環境下における複製配置数の抑制及び離脱による検索効率の低下の抑制を実現できていることがわかる。

4.3 おわりに

本研究では非構造型ネットワークにおいてノードの離脱を考慮した複製配置手法を提案した。

本実験の結果より本提案手法はノードの離脱が発生する環境下での検索効率の向上、及び既存の複製配置手法と比較して複製数の抑制に成功した。今後の予定としては、提案手法は離脱が発生した際の周辺への人気アイテムの探索がネットワークに与える負荷を考慮していないため、負荷を把握する実験が必要である。また負荷が大きい場合は処理コストを低減する手法を考案する予定である。

参考文献

- [1] 川崎 陽平, 佐藤 崇, 吉田 紀彦: P2P ネットワークにおけるコンテンツの人気度を反映した複製配置, インターネットコンファレンス 2005 論文集, pp.106-113, 東京, Oct. 2005.
- [2] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker: Search and Replication in Unstructured Peer-to-Peer Networks, Proc. 16th ACM Int'l Conf. on Supercomputing, 2002.
- [3] movieleens - movie recommendations, 入手先 <<http://movieleens.umn.edu/>>
- [4] 森一馬, 中村健二, 小柳滋: Flooding 検索におけるピアの嗜好を用いた P2P ネットワーク再構築手法の提案, 第 74 回全国大会講演論文集