

クラスタ形成手法とピア数がドメイン間 P2P ファイル転送に及ぼす影響

1130331 木村 紀夫 【植田研究室】

1 はじめに

近年、インターネットの普及に伴い、ネットワーク上を流れるトラフィック量が増加した。中でも P2P (Peer-to-Peer) ファイル共有ソフト利用者によるトラフィックが多い問題がある。ネットワークに流れるトラフィックを抑制するため本研究では、複数のクラスタ形成手法でシミュレーションを行い、ピア数の違いによるコンテンツ取得距離への影響を明らかにする。

2 従来手法の問題点

従来の P2P ファイル共有ソフトのクラスタリングは、所持しているコンテンツの傾向が同様なピア同士を同じクラスタ内に配置することで、コンテンツ検索の効率を上げている。しかし従来手法では実ネットワークを考慮していないため、実ネットワーク上で近くのピアが目的のコンテンツを所持していた場合でも、所属するクラスタが異なるとそのピアと比べて離れたピアから取得する可能性がある。この場合、トラフィックの経路が長くなり、結果としてネットワークへの負荷が増加する。

3 問題解決の方法

従来のクラスタ形成手法のみでは上記問題が考えられるため、従来のコンテンツ傾向を考慮したクラスタに加えてネットワーク距離を考慮したクラスタも形成し切り替えを行う。多くのピアが持つコンテンツはこのネットワーク距離のクラスタで検索を行うことで、従来のコンテンツ傾向のクラスタと比べてネットワーク的に近いピアからコンテンツを取得できるため、トラフィックの経路長の短縮が可能となる。

4 検索結果に及ぼす影響の検証

今回形成手法の異なるクラスタを比較しどれほどコンテンツ取得距離に違いが出るのかシミュレーションを行った。シミュレーションは、上記の 2 種類のクラスタ形成手法を使った。ネットワーク距離のクラスタの構築にはピア間距離を考慮したクラスタ形成手法 [1] を用いて、コンテンツ傾向のクラスタは Winny 等で用いられている従来の手法を参考に構築したものをを用いた。ピアは日本の 8 地域区分の人口分布比率に従いランダムに配置した。コンテンツは 10 種類用意し、それぞれ人気度 (参加ピア数に占める割合) を Zipf の法則に従い決めた。ピア数の与える影響を検証するために、参加ピア数を 100 台から 1000 台までの 100 台刻みで変化させて行った。コンテンツ検索を、それぞれのネットワーク規模で 50 回行い、選択されたピアまでの距離を計測した。

コンテンツの人気度が 20% の時に検索を行い、見つかったピアまでの平均距離を図 1 に示す。図 1 を見る

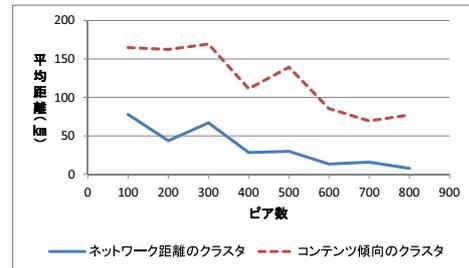


図 1 ピア数の変化による検索距離への影響

と、すべてのネットワーク規模でコンテンツ傾向のクラスタに比べてネットワーク距離のクラスタが近くのピアから取得できていることが分かる。双方のクラスタともにネットワーク規模が大きいかほど近くのピアからコンテンツを取得できる傾向にあることがわかる。ネットワーク規模が小さい時に検索が困難なコンテンツであったとしても、ネットワーク規模が大きくなるにつれて検索が容易になると考えられる。しかし、検索が困難なコンテンツをネットワーク距離のクラスタで検索すると、検索により不要なトラフィックを発生させてしまうため、ネットワーク規模や探索するコンテンツの人気度により、検索するクラスタ手法を適宜選択する必要があることが考えられる。今回のシミュレーションでは最大ピア数が 1000 台で行ったが、実際のネットワークではさらに多くのピアが P2P ネットワークに参加することが考えられるので、より多くのピア数でシミュレーションを行う必要がある。同様にコンテンツもより希少な人気度を用意してシミュレーションを行う必要があるほかに、8 つだった地域区分も増やしてより実際のネットワークに近いシミュレーション環境にする必要がある。

5 まとめ

ネットワーク距離を考慮したクラスタを用いることで、距離の近いピアからもコンテンツを取得することが可能となり、ネットワークのトラフィック範囲を抑制する事ができると考えられる。

シミュレーションにより検証した結果、双方のクラスタともにネットワークの規模が大きくなるにつれて近くのピアから取得できる傾向にある事が分かった。

参考文献

- [1] 上田達也, 安倍広多, 石橋勇人, 松浦敏雄, “P2P 手法によるインターネットピアの階層的クラスタリング”, 情報処理学会論文誌, vol.47, no.4, pp.1063-1076, Apr. 2006