

シーダを保証する Location Based Clustering による P2P 動画配信

Location Based P2P Clustering with Guaranteed Seeders

高田 和也[†]後藤 滋樹[†][†]早稲田大学 基幹理工学研究科 情報理工学専攻

概要

P2P の配信方式に関する研究が活発に行われている。本研究は P2P の通信コストを抑えることを重視する。単純な P2P プロトコルではピア同士の近接性を考慮することなくピアを選択するために遠距離のトラフィックが発生することがある。このようなトラフィックの問題を解決するためには地理情報クラスタリングを用いるのが有効である。ただし地理情報に頼るだけではシーダが存在しないクラスタが生じる場合がある。本研究は従来の方法を改善するために、クラスタリング手法にソフトクラスタリングを導入する。さらに提案手法を実証するために PlanetLab を用いて検証する。

1 提案手法

本手法の目的は先行研究 [1] において、クラスタリングの際にシーダを含まないクラスタが生成される場合があるという課題を解決することである。先行研究で提案されたインテリジェント・トラッカーのピアリスト生成アルゴリズムに変更を加えることにより目的を達成する。インテリジェント・トラッカーとは BitTorrent トラッカーがピアリストを生成する際に地理情報（緯度・経度）を利用して地理的に近いピア同士が通信することを狙うものである。動作例を図 1 に示す。

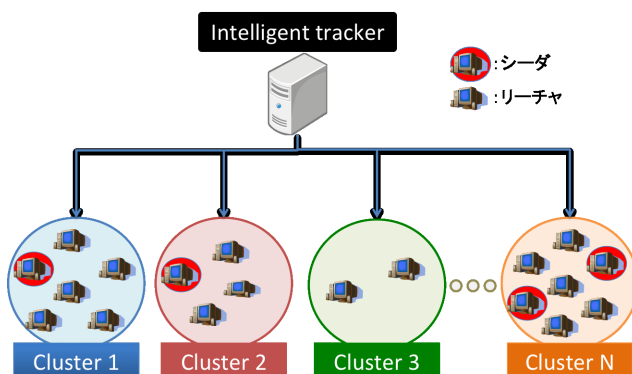


図1 インテリジェント・トラッカー動作例

トラッカー側で地理的に近いピア同士をグループ化するために、先行研究ではハードクラスタリングを用いている。ハードクラスタリングでは、要素が一意にクラスタに属するため、地理的な距離が近いノード同士を一意にグループ化できる。しかし、クラスタ内にシーダが含まれない場合はそのクラスタ内のノードは完全なデータを取得できない。そこで、提案手法ではクラスタ内に必ずシーダが含まれるようにするため、クラスタリング手法にソフトクラスタリングである Fuzzy c-means 法を用いる。ソフトクラスタリングでは、要素が各クラスタへ属する度合（帰属度）が出力となる。以下ではソフトクラスタリングを用いたシーダを保証する地理情報クラスタリング手法について述べる。

シーダを保証するクラスタリング手法 ピアの群（スウォーム）に属する各ピアがシーダか否かの判定はトラッカーのアクセスログ情報に含まれる left の値を用いる。パラメータ left は各ピアのダウンロード完了までの残りピース数を示しており、シーダの場合は left=0 となる。また、地理情報を得るには MaxMind 社の提供している GeoLiteCountry [3] を利用する。このライブラリを利用することにより IP アドレスから緯度・経度情報の取得が可能となる。本手法ではこの緯度・経度情報を利用し、Fuzzy c-means 法でクラスタリングを行う。具体的には以下のようにクラスタへの割り当てを行う。

1. ピアの IP アドレスから緯度・経度の値を取得。
2. 緯度・経度をデータとしてクラスタリングを行う。
3. 各ピアを帰属度が最も高いクラスタへ割り当てる。
4. 生成されたクラスタにシーダが含まれるかを判定。
5. シーダが含まれる場合は割り当て完了。
含まれない場合は次に帰属度の高いクラスタへ割り当て、4へ戻る。

2 実証実験

2.1 クラスタリング実験

この実験の目的は、提案手法によってシーダが保証されているかを確認することである。実験環境には世界規

模で運用されているテストベッドである PlanetLab を利用する。クラスタリングを行う際に参加ノード数に応じて適切なクラスタ数を設定する必要があるが、今回は、先行研究 [1] で示された線形関数を利用する。また、適切なシードの数は、実環境では状況によって異なるため今回の実験では評価のために全ノード数の 20% に設定して行う。表 1 に実験を行ったノード数、クラスタ数、シード数を示す。

表 1 参加ノード数に対するクラスタ数・シード数。

	ノード数	シード数	クラスタ数
Case 1	30	6	4
Case 2	200	40	7
Case 3	400	80	10
Case 4	600	120	14

実験を行った結果 Case1-Case4 すべての状況において提案手法を用いることでシードが保証できていることを確認した。ここでは Case1 ノード数 30 の場合のクラスタリング結果を図 2 に示す。比較対象として既存手法であるハードクラスタリングを用いた場合の結果も図 2 に示す。シードに着目すると、オセアニアにシードが存在しない。しかし、ハードクラスタリングではオセアニアのノードのみでクラスタを形成している。一方、提案手法では、オセアニアのノードはシードの存在するアジアのクラスタと同一クラスタを形成しているため、完全なデータをダウンロードすることが可能である。

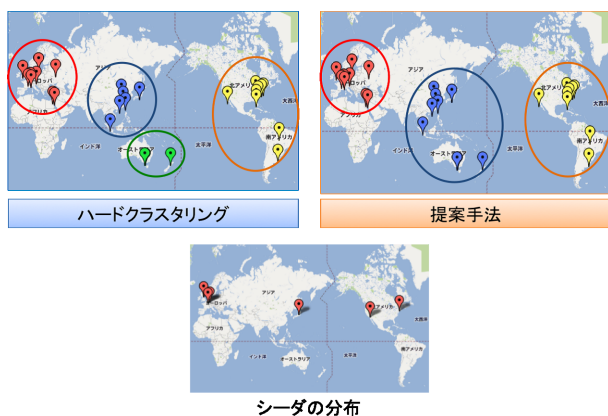


図 2 クラスタリング結果 ノード数 30

2.2 動画データ配信実験

提案手法によって生成されたクラスタ内でダウンロード実験を行う。本研究は、P2P 動画配信を想定しているためピアのピース選択アルゴリズムに Earliest First [1] を採用する。通常の BitTorrent クライアントはピース選択アルゴリズムとして Rarest First を代表とする、いくつかのピース選択アルゴリズムを採用している。しかし、それらはストリーミング再生を考慮したものではなく、そのままでは動画再生中にピース不在のため動画が再生できなくなる可能性がある。そこで先行研究 [1] で

提案されたものが Earliest First である。Earliest First はピースを先頭から順番に取得することでスムーズな動画再生を可能としたアルゴリズムである。

本実験では、BitTorrent においてファイル転送効率を最適化するピース選択アルゴリズムを変更するため、ファイルのダウンロード速度低下が考えられる。さらに、クラスタリングによって通信できるピアの選択肢が減少するため、これによる速度低下も考えられる。そこで、本実験ではクラスタ内 P2P 動画配信においてダウンロード速度が配信レートを上回るか検証する。ここで配信レートは 2Mbps を想定する。この値は YouTube の HD 画質モードの最高画質配信レートである。最も遅いノードが条件を満たしていれば動画配信には問題のないダウンロード速度であることが示せる。本実験では実験 2.1 において検証したノード数 30 の場合のノードを用いてダウンロード実験を行う。用いるコンテンツのサイズは 200MB で、ピースサイズは 256kB に設定。表 2 に各クラスタ内で最も平均ダウンロード速度が遅かったノードの結果を示す。

表 2 各クラスタ内で最も遅いノードの平均ダウンロード速度

	アジア・オセアニア	南北アメリカ	欧州
ダウンロード速度 (Mbps)	2.30	7.84	3.60

表 2 の結果より、すべてのクラスタで最も平均ダウンロード速度が遅いピアでも条件として設定した配信レート 2Mbps を上回ることが示せた。

3 まとめ

本研究では、地理情報クラスタリングによる P2P 動画配信においてクラスタ内にシードを保証する手法を提案した。ノード数を 30~600 まで変動させて実験した結果、いかなる場合でもシードを保証できることを示した。また、提案手法によって生成されたクラスタ内での P2P 動画配信実験においても十分なダウンロード速度を確保できることを示した。今後の課題は、提案手法を実サービスで動作させ、検証することである。

参考文献

- [1] 大村淳己, 高田和也, 後藤滋樹, Location Based Clustering を用いた P2P ストリーミング, 電子情報通信学会技術研究報告, vol. 110, no. 373, IN2010-124, pp.37-42, 2011 年 1 月.
- [2] Bittorrent Protocol Specification v1.0
<http://wiki.theory.org/BitTorrentSpecification>
- [3] MaxMind - GeoIP — IP Address Location Technology
<http://www.maxmind.com/app/ip-location>
- [4] 新納浩幸 著, R で学ぶクラスタ解析, オーム社, 2007.