

氏名	山田 太造
学位（専攻分野）	博士（情報学）
学位記番号	総研大甲第 861 号
学位授与の日付	平成 17 年 3 月 24 日
学位授与の要件	複合科学研究科 情報学専攻 学位規則第 6 条第 1 項該当
学位論文題目	適応的ピア発見による非構造 Peer-to-Peer システムの効率的な問い合わせ処理に関する研究
論文審査員	主査 教授 高須 淳宏 教授 丸山 勝巳 教授 山田 茂樹 教授 相澤 彰子 教授 安達 淳（国立情報学研究所） 教授 河野 浩之（南山大学）

論文内容の要旨

blog や Wiki 等に代表されるオープンな協調作業環境下で共有される文書は更新頻度が高く、システムを利用する各ユーザは全文書を必要とするわけではない。これらのシステムのユーザ数およびそこで共有される文書数は増加傾向にあり、今後も注目すべきデータ共有システムの 1 つである。オープンな協調作業環境システムで効率的なデータ共有を行うためにはシステムのスケラビリティの確保およびデータ更新に対して柔軟に対応可能な仕組みを導入する必要がある。そこで自己組織化に優れる自律分散型システムである Peer-to-Peer (P2P) システムを利用することでこれらの問題を解決する。P2P システムと他の集中システムや分散システムと大きく異なる性質として、各ピアの自律性の高さから各ピアが接続対象を動的にかつ適応的に発見することが容易かつ自然に行うことが可能な仕組みを保持している点である。そのためシステムの自己組織化に非常に優れている。

本研究ではピア（ノード）の自律性およびシステムの柔軟性や自己組織化に優れる Gnutella や Freenet に代表される非構造 P2P システムを用いる。非構造 P2P システムでは、各ピアはシステム内の全ピアに対して積極的にコミュニケーションを行うのではなく、同様の嗜好であるピアに対して積極的にコミュニケーションを行い、それらのピアを動的かつ適応的に確保することが可能である。また、各ピアの嗜好や行動に変化があったとしても、その変化に柔軟に適応し、新たに積極的にコミュニケーションを行う対象を確保することも可能である。しかしながら、従来の非構造 P2P システムでは、問い合わせ処理の非効率性が指摘されており、P2P システムでのデータ共有を効率的に行うためには問い合わせ処理の性能を向上させる必要がある。本研究ではシステム上の各ピアの情報および動向に着目し問い合わせ処理の効率性を向上させることを目的とする。

P2P システムでの問い合わせ処理の処理効率を向上させるためには応答時間の向上および帯域幅消費量の低減を達成する必要がある。この問題の解決の手段として問い合わせメッセージのルーティングの方法および文書の配置方法の向上を挙げることができる。そこで本研究では、P2P ネットワーク上の各ピアが適応的にそのピアと同じ嗜好であるピアを発見可能である仕組みを提案する。この仕組みを用いると、同じ嗜好であるピアに対して積極的に問い合わせメッセージを送信するルーティング、およびそのようなピアに対して文書の複製を配置する仕組みを提供することが可能になる。

問い合わせメッセージのルーティングの目的はより少ない問い合わせメッセージで問い合わせを終了させることである。ルーティングの性能向上のためにはその指標を与える必要がある。本研究ではその指標としてピアの有用度を導入し、各ピアはこれを利用して動的に P2P ネットワーク上にリンク（長距離リンク）を形成する。この仕組みにより、ネットワーク上の遠くに位置するピアに対しても直接問い合わせメッセージを送信することが可能になる。ピアの有用度はそのピアが所有する文書の有用度によって決定する。文書の“有用度”は文書の参照回数および文書の更新からの経過時間によって決定する。ピアの情報を維持管理するための仕組みが必要であり、ピア情報を格納するための分散インデックスである Direct Index (DI) を提案する。システム上の各ピアは DI 内のピア情報を伝播させ、インデックス更新メッセージを受信したピアはその内容に応じてその DI を更新させることで他のピアを動的に発見する。問い合わせ処理を行う場合、各ピアは適応的

に長距離リンクを形成し、問い合わせメッセージをルーティングさせる。インデックス更新を行う場合、各ピアはその DI 内の情報を更新し、インデックス更新メッセージを伝播させる。DI を用いたインデックス更新では、文書の追加や更新等の個々の文書の情報の変化はその文書を所有するピア内では即時反映するが、他のピアに対しては敏感に反映させない。インデックス更新メッセージを受け取ったピアはそのメッセージがそのピアの DI 内の情報を変化させない場合、インデックス更新を停止するためである。インデックス更新の頻度を抑えることが可能になるため、インデックス更新のコストを低減することが可能になる。

P2P システムでは各ピアのネットワークへの接続状況は頻繁に変化させる。そのために、発見困難な文書やシステムから喪失してしまう文書が存在してしまう。そこで本研究では文書の複製を行う。文書複製のために、本研究では、複製の数をピアごとに動的に決定し、複製対象となる文書とその所有者であるピアが動的に決定する。各ピアはそのピアが生成できる複製の数をそのピアの有用度とその隣接ピアの有用度から相対的に決定する。複製する文書の決定は文書の有用度とその文書の既存の複製数に従う。有用度が高いものほど複製の対象となりやすいが既存の複製の数によってその文書の複製を抑制する方針である。文書複製を行う場合、複製の配置の対象を決定する必要がある。そのために、本研究では、動的にピアグループを形成し、そのピアグループに属するピアに対して複製を配置する手法を提案する。ピアグループの形成の方針はあるピアと関連するピアを発見し、それらのピア集合でグルーピングを行う。ピア間の関連を決定する方法がそれぞれのピアグループ形成の方針で異なる。

提案する分散インデックスおよび文書複製手法を評価するためにシミュレーションを行い、その効果を示し、提案した各手法の特色を示している。その結果、本研究での問い合わせメッセージのルーティングおよび文書複製の手法は既存の非構造 P2P システムでの手法と比較して、ネットワークへのダイナミクスへ柔軟に対応し、帯域幅消費量の低減だけでなく、問い合わせ処理の応答時間の向上をも達成することが可能であった。本研究でのアプローチは非構造 P2P システムの特徴であるシステムの自己組織化および柔軟性の特徴を失うことがなかった。さらに本研究での手法は従来の非構造 P2P システムでの手法と異なる点は、各ピアが積極的にコミュニケーションを行う対象を単に隣接するピアだけではなくスケラブルに選択することを可能にすることである。そのため、本研究で提案した適応的なピア発見を可能にした分散インデックスおよび複製配置の方法は問い合わせ処理の向上への寄与に有効な手段であることが分かった。

山田太造氏は、新しいタイプの自律分散システムであるPeer-to-Peer (P2P) システム上で効率的な問い合わせ処理を実現するインデキシング法について研究を行った。P2Pシステムは自律性の高い大規模分散システムを構築する上で重要な技術となっているが、P2Pシステムにはいくつかの問題があり、そのひとつに問い合わせ処理のコストの高さがある。具体的には、ネットワークの帯域幅の消費量と問い合わせ処理における応答時間の遅さが重大な欠点となっている。本研究で、山田氏はこの問題を解決するために、新たなインデキシング法であるDirect Indexを提案し、シミュレーションを行ってその有効性を示した。本研究は以下のようにまとめられている。

第1章の序論に続いて、第2章でP2Pシステムの概要および問い合わせ処理に関する研究状況と問題点を整理している。第3章では、P2Pシステムの中でも自律性、柔軟性の高い非構造型P2Pシステムにおいて、ピア間で情報を効率良く伝達するために、P2Pネットワーク上で遠方にあるピアに直接アクセスする方法を提案している。非構造P2Pシステムでは、通常、P2Pネットワーク上の隣接するピアに情報を順次伝えることによって、ネットワーク上の遠方のピアへ情報が伝えられる。そのため、遠方のピアに情報を伝達するためには、多くのピアを経由する必要がある。伝達の遅延と通信量の増大という問題が生じる。本研究では、まず、ピアが持つデータへのアクセス頻度、データの新鮮さ、および、ピアの相対的な位置関係に基づいてピアの有用度という尺度を導入する。そして、この有用度に基づいて非隣接ピアとも直接、情報を伝達することによって非構造P2Pの欠点である応答の遅延と通信量の増大という問題を克服するのが、本研究の基本的なアイデアとなっている。

第4章、第5章では、有用度に基づいた効率の良い問い合わせ処理の実装法を提案している。まず、第4章において、各ピアが有用な隣接および非隣接ピアの情報を維持管理するためのデータ構造であるDirect Index (DI)を導入する。問い合わせ処理においては、その問い合わせに適した情報を所有するピアに優先的に問い合わせを送ることが重要になる。本章では、DIを用いて問い合わせに適したピアに優先的に問い合わせメッセージを送るルーティングアルゴリズムを提案する。P2Pシステムでは、各ピアに含まれるデータは頻繁に更新されるため、データの更新に伴って、各ピアが持つDIを更新する必要がある。そこで、DIの更新方法についても考察し、更新方法を提案している。提案方法では、問い合わせ処理の過程で逐次有用なピアの情報がDIに追加・更新されるため、動的なP2Pシステムに対して、適応的な問い合わせメッセージのルーティングが可能になっている。続いて、問い合わせ処理の効率をさらに向上させるため、データの複製方法について検討を行っている。複製配置先となるピアの発見や選択は問い合わせ処理の性能に深く関わる。本研究では3章で提案した有用度に基づいて適応的に複製配置の対象を選択する仕組みを提案している。

第6章では、シミュレーション実験を通して、応答時間および通信量の観点から、本研究の提案する手法の有効性の評価を行っている。その結果、本研究で提案するインデキシング法は既存の非構造P2Pシステムの手法と比較して、(1) ネットワークへのダイナミクスに柔軟、(2) 帯域幅消費量の低減、および(3) 応答時間の向上、の点で優位であることが示されている。

以上、山田太造氏の研究は、自律性・柔軟性に優れた非構造型 P2P システムにおいて、その欠点である問い合わせ処理の効率化を実現するものであり、P2P システムのスケールビリティの向上に多大な貢献を行う実用性の高い研究である。また、P2P システム上での情報へのアクセス状況に応じて、適応的にネットワークを逐次再構成するアプローチは、P2P システムの動的特性と親和性が高く、この分野の先駆的研究となっている。本研究の成果の一部は、学術雑誌論文 1 篇および国際会議論文 4 篇にまとめられており研究業績も十分である。よって、山田太造氏の提出論文は学位（情報学）を授与するに値すると判断した。