

P2P環境におけるロバストな索引管理方式

由利 誠[†] 能登谷 淳一^{††} 草薙 良至^{††} 笠井 雅夫^{††}[†] 秋田県立大学システム科学技術研究所 ^{††} 秋田県立大学システム科学技術学部

1 はじめに

近年、計算機の高性能化・低価格化と高速情報通信インフラの発達により P2P 技術が注目されている。必要なサービスを検索するための索引を集中管理するサーバを持たないピア P2P ネットワークでは、サービスを少ない通信コストで高速に発見する機構の提供が重要である。高速なサービス検索手法として、DHT(Distributed Hash Table) を利用した構造化 P2P が提案されているが、構造化 P2P はノードの頻繁な参加離脱を伴う応用には適さないことが知られている [2][3]。本研究では、ネットワーク構造やデータを冗長化することにより、参加離脱耐性の高い構造化 P2P を得ることを目的とする。

2 DHT を利用したサービスの検索

DHT を利用したサービス検索手法のひとつに、Chord[1] がある。Chord では、 2^m の大きさをもつ仮想的な 1 次元円環状ハッシュ空間を索引として利用する。ネットワークに存在する全てのノードとサービスはノードの IP アドレス、サービスのキーの m bit ハッシュ値にしたがって、円環状ハッシュ空間上に配置される。Chord においては、 m bit のハッシュ値をノードやサービスの ID と呼ぶ¹。各サービスは、ハッシュ空間を時計回りに辿り、最初に出会うノードに割り当てられる (図 1)。各ノードは、割り当てられたサービスに関する情報、ハッシュ空間上で前後に位置するノードの情報、ルーティング情報を保持する。ルーティング情報には、サービス検索要求を受け取り、対応するサービスを保持していなかった場合に検索要求を転送すべきノードの情報が含まれている。ルーティング情報は各ノードに m 個ずつ存在し、ハッシュ値 n を持つノードにおけるルーティング情報は、次式で定義される区間列を基に決定される。

$$[(n + 2^{k-1}) \pmod{2^m}, (n + 2^k) \pmod{2^m}) \quad 1 \leq k \leq m$$

ここで、各区間に対応するノードは、それぞれ $(n + 2^{k-1}) \pmod{2^m}$ からハッシュ空間を時計回りに辿り、最初に出会うノードとなる。区間列はハッシュ空間全体の分割となっており、各区間は $2^k - 1$ の大きさをもつ。よって、検索要求が k 番目の区間に対応するノードに転送されたとき、検索対象となる区間の大きさは $2^k - 1$ に絞られる。ノードの個数 N を 2^m とした場合、検索要求が転送されるたび検索対象となる区間が少なくとも $2^m - 1$ の大きさに絞られることから、サービスの検索は $O(\log N)$ 回の転送回数で行うことができる。

¹実際には複数のサービス、ノードが同一のハッシュ値を持つ可能性があるため、本稿では ID ではなく単にハッシュ値と呼ぶ。

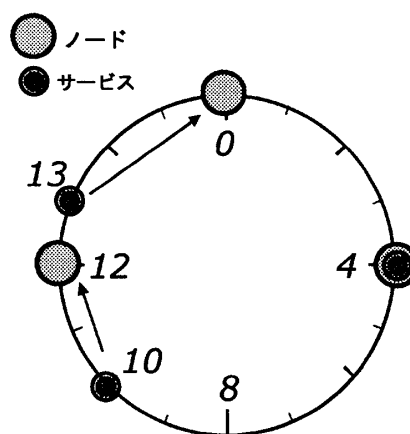


図 1: $m = 4$ におけるハッシュ空間とサービスエントリの配置例

Chord のルーティング情報管理方式では、ノードの参加離脱が頻繁に行われる応用において正確なルーティング情報を常に保持しておくことは難しく、サービスの検索を行うための転送回数が増大してしまう。また、各ノードが離脱する際には、割り当てられたサービスを他のノードに引き継ぐ必要がある [3]。この作業が行われることなくノードが離脱してしまうと、ノードが保持していたサービスを参照することは不可能になる。また、ルーティング情報の修正コストも増加してしまう。

3 冗長化 DHT の管理手法

本研究では、Chord の索引管理手法を拡張し、ネットワークの構造やデータの配置を冗長化することで、ノードの参加離脱に対してロバストな DHT 検索を行うことを目的とする。本研究で用いる冗長化の方法として、同一のハッシュ値を持つ複数のノードの作成を考える。そのために、ノードの IP アドレスのハッシュ値ではなく、ノードが所属するグループを表すキー (グループ ID) を与え、グループ ID のハッシュ値にしたがってノードをハッシュ空間上に配置する。同一のグループ ID を持つ複数のノードは、ハッシュ空間上の同一の点に配置される。同一のグループ ID をもつ複数のノードからなる集合を、ノードグループと呼ぶ。ノードグループ k に含まれるそれぞれのノードが同一の前後ノード情報、ルーティング情報、サービスの複製を所持し、さらにノード情報とルーティング情報が示すノードグループに含まれる複数のノードのアドレスを保持しておくことで、ノードグループ中の全てのノードが離脱しない限り、正

しくサービスの参照、検索要求の転送を行うことができる。

各ノードグループ内の平均ノード数は大きくなるほど冗長度が高まるが、ノードグループ間の区間長は大きくなり、各ノードのサービス保持数は増大する。ノードの平均サービス保持数、最大サービス保持数はP2Pシステム全体の性能に影響を与える。サービス保持数と耐障害性はトレードオフの関係にあり、対象となる応用のサービスの形態やノードの参加・離脱率、システムを構成する機器の故障率などから、適切に設定する必要がある。本研究では、各ノードグループに所属できる最大のノード数を指定することにより、DHTシステム全体の冗長度を制御する手法を提案する。各ノードグループに所属することができる最大のノード数を最大多重度 r とする。最大多重度 r 、ノード数 N 、サービス数 S の場合を考えると、ノードの平均サービス保持数は rS/N となる。ルーティング情報が、ノードグループ中の r 個全てのノードに関する情報をもつとすると、ルーティング情報の維持コストは $O(r \log N)$ となる。さらに、各ノードの離脱率を p とすると、各サービスは $(1-p)^r$ の確率でいずれかのノードに存在しつづけることができる。

本手法では、新しくノードをネットワークに挿入する際、ネットワークに参加する前にいくつかのノードグループを調査し、所属しているノード数が最大多重度 r に満たないグループがあれば、そのグループ番号を挿入するノードに設定する。もしそのようなノードグループが見付からなければ、サービス保持数が最大のノードグループを選び、そのノードグループが持つサービスをハッシュ値でソートし、中間に位置するサービスのキーを挿入するノードのグループ番号に設定する。これにより各ノードグループにおける保持サービス数を分散することが期待できる。

4 評価

シミュレーションに基づく本提案手法の評価実験を行った。実験では、ノードの離脱による検索失敗率の変化と、ノードの最大サービス保持数を測定した。ノードの離脱による検索失敗率の変化では全てのノードがネットワークに配置され、全てのサービスに対する検索が正常に行われている状態から任意の個数のノードを離脱させ、各ノードのルーティング情報を更新した後、ランダムに選んだサービスの検索を行い、その失敗率を計測した。この時に用いた値は、最大多重度 $r = 1, 2, 3$ 、全ノード数 $N = 10^3$ 、全サービス数 $S = 10^5$ である。次に、最大多重度 $r = 1, 2, 3$ 、全ノード数 $N = 10^3$ 、全サービス数 $S = 10^4$ の条件の下でノードの最大サービス保持数を調べた。

シミュレーション結果を図2, 3に示す。ここで、 $r = 1$ の場合は、まったく冗長度されていない状態とみなすことができる。図2について、 r の値が増加していくにつれ、全ての範囲で検索の失敗率を低下させることが確認できる。また、いずれの場合にも、検索の失敗率は $1-p^r$ 程度となっていることが確認できた。図3について、 $r = 1, 3$ においてはほぼ同じ値となっており、 $r = 2$ の場合には最大サービス保持数が減少している事がわかる。よって、本手法では冗長度するノードグループを適切に選ぶことで、最大サービス保持数を低下させることが確認できた。

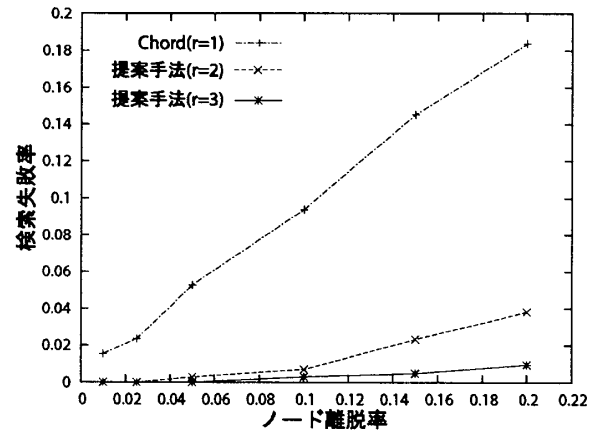


図2: サービス検索の失敗率

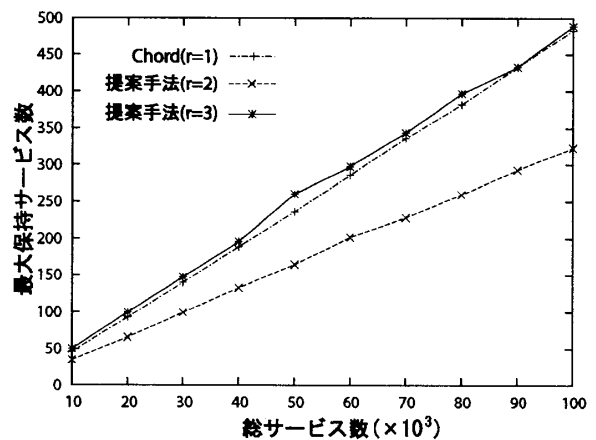


図3: 最大サービス保持数

5 まとめ

ノードの参加離脱が頻繁に行われる状況に対応させるために、ネットワークの構造やデータの配置を冗長度化する手法を提案した。また、ネットワークの構造やデータの配置を冗長度化することで検索の失敗率を低下させることを確認した。また、最大多重度を変化させることで、システムの冗長度を設定できることを確認した。

参考文献

- [1] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, H. Balakrishnan "Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications" August 2001.
- [2] Sean Rhea, Dennis Geels, Timothy Roscoe, John Kubiatowicz "Handling Churn in a DHT" June 2004.
- [3] Anne-Marie Bosneag, Monica Brockmeyer "A Unified Formal Specification for a Multi-Consistency Replication System for DHTs" 2005.