

情報指向ネットワークのための自己組織型名前解決システム

Suyong EUM 地引昌弘 村田正幸 朝枝 仁 西永 望

本稿は、情報指向ネットワーク (ICN) のための名前解決システムのトポロジーを、特にシステムの低平均ディグリーに力を注ぎつつ、自己組織型で構築するためのアルゴリズムを提案するものである。自己組織型の運用及び低平均ディグリーの維持により、管理面でのオーバーヘッドは大きく低減され、スケーラブルな展開が可能となっている。さらに、提案するアルゴリズムを理論的に解析し、この提案について、より深い理解を提供する。

1 はじめに

ネットワークのパラダイムは、人々のあいだのコミュニケーションからコンテンツの検索へと移行してきた。シスコが報告しているように、2015年までに世界全体でのコンシューマー用トラフィックの最大90%がマルチメディアコンテンツで占められるようになるだろう^[1]。このパラダイムシフトに対応するために、近年、ネットワーク研究コミュニティにおいて、情報指向ネットワーク (ICN)^{[2][7]} が大きな関心を集めるようになった^[8]。ICN 研究における主要な課題の1つは、スケーラビリティを備えたルーティングメカニズムを設計することである。ICN においては、ユーザーのリクエストはその名前に基づいて希望するデータオブジェクトにルーティングされる。いわゆる名前ベースのルーティングである。この問題に対する共通のアプローチは、分散ハッシュテーブル (DHT) に基づく名前解決システム (NRS) を採用することで、これによって、データオブジェクトの識別子をロケータに翻訳し、ユーザーのリクエストを当該のデータオブジェクトにルーティングする際にそのロケータを使えるようにする、というものである。たとえば CCNx^[4] では、データオブジェクトのルーティング用エントリをプレフィックス (たとえば IP におけるネットワーク・プレフィックス) に集約する名前集約メカニズムを採用している。このアプローチを現実的なものにするためには、データオブジェクトのプレフィックスのもとである程度の集約を実現するため、データオブジェクト識別子のプレフィックスを、ルーティングのための他の識別子 (たとえばデータオブジェクトの発行者・所有者を示す識別子など) に紐付けるべきである。こうした運用は暗黙のうちに、このようなアプローチではデータオブジェクトの識別子をそのプレフィックスに紐付けるような名前解決システムが必要になるとい

うことを意味している。他方、NetInf^[9] のように現在展開されている IP インフラを活用するような、一群の ICN アーキテクチャーが存在する。前述のアプローチとは異なり、ここでは、データオブジェクトの識別子を1つの IP アドレスにマッピングするような名前解決システムを明示的に要求している。

ICN の名前解決システム的设计においては、いくつかの課題が存在する。第1に、名前解決システムには、ユーザーからの名前解決クエリ、ネットワーク内キャッシュからのマッピングレコード更新クエリを含む、多数のクエリをサポートすることが期待される。たとえば、ユーザーが複数のデータオブジェクトを含むウェブページを閲覧するために、同時に多数の名前解決クエリを発生させる可能性がある。加えて、個々のネットワーク内キャッシュの容量には限界があるため、キャッシュ内のコピーの一部が新たに発生したコピーに不規則に置き換えられ、ICN の名前解決システムのマッピングレコードに頻繁な更新が生じる^[9]。第2に、ICN の名前解決システムには、そのサイズを動的に変化させることが期待される。ICN の名前解決システムでは、 $10^{15} \sim 10^{22}$ ものマッピングレコードをサポートすることが期待される^[10]。さらに、時間帯の違いによる個々のデータオブジェクトの人気の変動により、名前解決システムのある部分がダイナミックに成長する、あるいはむしろ縮小する可能性がある。こうしたノード参加の動的な振る舞いは、P2P の文脈における「チェーン (攪拌)」と呼ばれている^[11]。以上のような問題の理解を踏まえて、我々は ICN のための自己組織型 DHT 名前解決システムのトポロジーを構築する ES (“Efficient yet Simple” [効率的だがシンプル]) モデルを提唱する。その主要な設計目標は、マッピングレコードのルックアップに関する低レイテンシーと、非常に動的な環境における低メンテナンスコストを実現することである。これらの設計目標は、トポロジー

的にダイアメーターは低いが平均ディグリーは小さい ICN 名前解決システムのトポロジー構造を構築することに匹敵する。ダイアメーターが低いため、各クエリに対する応答時間は短くなり、これは最終的にネットワーク資源の節約にもつながる。さらに、平均ディグリーの小ささにより、名前解決システムの保守が簡潔になり、「チェーン」シナリオにも効率的に対応できるようになる。

本稿の構成は以下のようになる。2 では、このプロジェクトの動機を説明し、関連研究を検証する。3 では、我々が提案するモデルを、その理論的解析まで含めて記述する。4 では、膨大なシミュレーションと理論的解析を通じて我々の提案を評価する。最後の 5 は本稿のまとめである。

2 背景

2.1 動機

本研究は、ICN のための DHT 名前解決システムに関して最も効率的なトポロジーを設計するために開始された。というのも、トポロジー設計こそ、名前解決システムの設計の確信であると我々は信じているからである。

まず我々はカウツグラフ (kautz graph) に注目した^{[12][13]}。これは、個々のノードのディグリーを最小限にしたうえでの理論最適なダイアメーターをほぼまとめたものである。たとえば、カウツグラフにおける平均距離は、同一条件下での Chord^[14] によるものの半分以下である。またカウツグラフは、発信ノード・宛先ノード間の経路が重複していないため、最適のフォルトトレランスであることが知られている。この理由から、Fissione^[15]、Moore^[16]、BAKE^[17]、KCube^[18]などでは、カウツグラフの変種が DHT のトポロジーとして提案されてきた。しかし、カウツグラフはそのサイズを拡張していくためのコストが高い。たとえば、グラフにノードを 1 個追加すると、リンクのほぼ半分以上を再配置することになる^[19]。この問題に対処するために部分ライングラフ (PLG)^[20] が提案されたが、それでも、サイズを拡張するためには個々のノードの識別子を更新する必要がある。この拡張に伴う問題ゆえに、動的なネットワーク環境におけるカウツグラフの活用は妨げられてきた。要するに、任意のノード数及びディグリーにおけるネットワークの距離及び堅牢性という点で効率的なトポロジーを構築し得る理論的枠組みは複数ある。

しかしながら、その複雑さ及び拡張に伴う問題ゆえに、動的なネットワーク環境で採用されるには至らなかったのである。

複雑さ及び拡張の問題を生じさせずに効率性を実現する有望なアプローチとしては、Symphony^[21] がある。これは「スモールワールド」^[22] 概念を利用したもので、少数のリンクを用いた「小」距離のトポロジーを構築するものである。たとえば、ノードの数とネットワークの距離が条件として与えられた場合、Symphony は Chord のわずか半分のリンク数で条件を満たすトポロジーを構築できる。しかし、Symphony の特性で 1 つ残念なのは、システム変数を推定するためにシステムのバックグラウンドで動いている推定プロトコルである。こうした推定プロトコルを走らせるためのコストは、特にスケーラブルな名前解決システムの設計においては、システムのサイズが拡大するにつれて、許容できないほど大きくなってしまふ。我々は、この Symphony の残念な特性を克服するだけでなく、いっそうシンプルな運用でもさらに優れたパフォーマンスを実現したいという動機を持っている。

2.2 関連研究

名前解決システムとして最も広く知られているのは、DNS (ドメインネームシステム) である。DNS はオブジェクト識別子の階層構造に依拠しており、フラット識別子をサポートするには不向きである。また DNS では広範囲なキャッシュを多用していることから、膨大なバインディング情報の迅速な更新に対応できない。この理由から、CoDoNS^[23]、SFR^[24]、LISP-DHT^[25] といった DHT ベースの複数のアプローチが見られる。だがこれらは、バインディング情報がグローバルに分散している場合、レイテンシーの高さに悩むことになる^[26]。

ICN の文脈では、NetInf^[6] が 2 つの解決スキームを提示している。多層 DHT (MDHT)^[9] 及び階層的スキップネット^[27] である。PSIRP^[28] では、利用可能な多くのオーバレイ設計のなかから、階層的 Chord^[29] を選んだ。一般的な ICN アーキテクチャーに関して、Kostas et al. は、階層的 Pastry^[30] を提唱した。こうしたアプローチのねらいは、ネットワークの物理的な近接性、管理上の境界、ドメイン内のルーティングポリシー、あるいはこれらの組み合わせを見逃してしまうという現在の DHT 設計の問題に対処することである。だが、これらのアプローチは、最適化の余地を残していた Chord^[14] など従来の DHT に関して、階層的又はモジュラー化した構造を作るだけである。すでに述べたように、カウツグラフの平均ディスタンスは、任意のノード数及びディグリーにおいて、Chord^[14] に比べて半分よりも小さい。

我々は、単に従来のシステムを継ぎ合わせるだけでなく、「クリーンスレート」の原則に基づき、ICN 名前解決システムの設計により多くの関心を注ぐ価値があ

ると主張している。というのも、効率的な名前解決システムは、ICNのグローバルレベルでの展開を成功させるための鍵になると考えるからである。

3 ESモデル

図1は、ESモデルにおける1つのノードの参加プロセスを示している。各ポイントが何らかの暗号化ハッシュ機能から得られたハッシュ値に対応しているようなリングを想像していただきたい。すると、ここに参加する各ノードは、そのMACアドレスなどの一意的な識別子を使って自らのハッシュ値を生成し、リング上の識別子空間における自らの位置を見いだすことになる。新規参加ノードは自らの位置を見つけたうえで、自らの識別子から時計回りにすぐ隣のノードへとたどることで、識別子空間を把握する。

新規参加ノードは2つの異なるタイプのリンクを生成する。すなわち、短距離リンク(S)及び長距離リンク(L)である。短距離リンクは、参加ノードと、その直接隣り合うノードとのあいだに生成される。長距離リンクはルールに従って生成される。

1) 長距離リンク：新規参加ノードは、たとえばハッシュ[MACアドレス+1]、……ハッシュ[MACアドレス+L]のように、L個のハッシュ値を生成する。そして、参加ノードは、対応するハッシュ値を管理する個々のノードを発見する。相手となる管理ノードが各々見つければ、管理ノードは近隣ノードの1つを無作為に選択し、そのノードの識別子を新規参加ノードに渡す。最後に、新規参加ノードは与えられた識別子を用いてそのノードと接続する。各ノードに許容される最大リンク数によっては、参加リクエストが拒絶されるケースもある。この場合、最初に選択された管理ノードが、新規参加ノードにとってのエンドポイントとなる。

たとえば、図1における参加ノードAは、自らの位置から最も近いノードを発見してリング・トポロ

ジーに参加する。このトポロジーは様々な方法で実装可能であり、限定されたエリアにおけるフラッディング・メカニズムや、ブートストラップ・サーバを採用することができる。それから、参加ノードと時計回りに直接隣り合うノードとのあいだにショートリンクが生成される。各ロングリンクについては、参加ノードAは、自らのMACアドレスのハッシュ値をベースに一意的な識別子を生成し、その識別子を管理するノード(この例ではノードB)を見つける。ハッシュ機能のランダム化のおかげで、ノードBは、リング・トポロジーの中で無作為に選択されたノードと見なすことができる。ノードAとノードBが接続されると、ノードBは近隣ノードから無作為に1つを選び(この例ではノードC)、その識別子を参加ノードに渡す。最後に、参加ノードは受け取った識別子を使ってノードCと接続する。ノードCは、許容される最大リンク数を理由に参加リクエストを拒否することができる。この場合、参加ノードAは、ノードCの代わりに、当初選択したノードBを選択する。3)における理論的解析では、参加リクエストが拒否された場合には参加プロセスが最初から繰り返されるものと想定している。

2) ルーティング：ESモデルは、堅牢で信頼性の高い運用を実現するため、双方向性のリンクを用いる。各ノードは相互に数値を比較可能な独自の識別子を持っているため、我々の提案においては、グリーディルーティングのいくつかの変種が潜在的な候補となり得る。何らかのリクエストメッセージがあると、各ノードは、その宛先ポイントまでの絶対距離が減少するような形で、それを単純に近隣ノードの1つに転送する。しかしながら、本論文の眼目はルーティングではなく、名前解決システムのトポロジー構築である。ルーティング問題に関する取組は、我々にとって今後の研究課題の1つとなっている。

3) 理論的解析：ここで、我々が提案するモデルのパラメータを紹介し、そのパラメータ相互の関係と、モデルによって構築されるトポロジーの平均距離について解析を行う。ESモデルのパラメータは以下のとおりである。

- S：ショートリンクの数
- L：ロングリンクの数
- m：平均ディグリー(S+L)
- T：カットオフ最大ディグリー
- AD：平均距離

まず、解析はこのモデルによるトポロジーのディグリー分布を導出することから始まる。この解析は、システムの各状態の変異は、システムをその状態に導こうとする影響と、システムをその状態から引き離そうとする影響力の差異として描写できる。図2は、ネッ

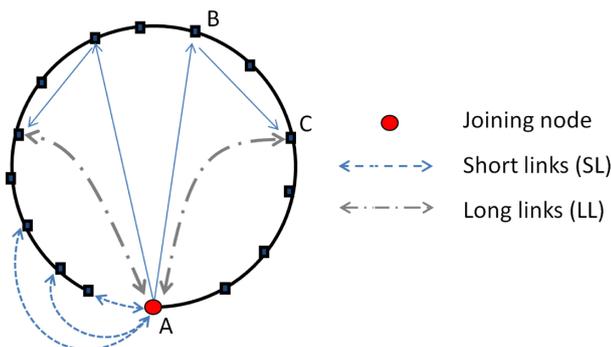


図1 ESモデル

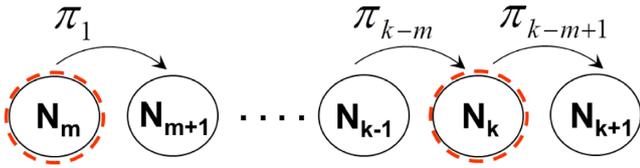


図2 ノード数の変化とそのディグリーを示す状態変化率ダイアグラム

トワーク内のノードの数が增大するにつれて、ノードのディグリー数 N_k がどのように変化するかを示すものである。ここで、 N_k はディグリー k であるようなノードの数である。どの時点をとっても、これはまず N_m から始まる。というのも、ネットワーク内には常にディグリー m であるようなノードが少なくとも1個は存在するからだ。構築されたネットワークに N 個のノードがあるとしよう。そして、ディグリーが m であるような新規ノードがネットワークに参加しようとしている。新規ノードが付加されることによって N_{k-1} が N_k になる確率を、式(1)に示すように、 Π_{k-m} と定義する。

$$\Pi_{k-m} = \frac{S}{L+S} \frac{mN_{k-1}}{n} + \frac{L}{L+S} (k-1) \frac{N_{k-1}}{2n} \quad (1)$$

この式の右辺の第1項は、パラメーター $S/(L+S)$ で決まる無作為な追加によって N_{k-1} が N_k になるプロセスを示している。というのも、ESモデルにおけるショートリンクは、無作為な追加に基づいて生成されるからである。ネットワーク内から無作為に選ばれたノードのディグリーが $k-1$ である確率は N_{k-1}/N に等しく、この確率はパラメーター m に比例する。なぜなら、各々の新規ノードは既存のネットワークと m 個の接続を作るからである。 n は時間に似た整数パラメーターである。というのは、参加ノードは各タイムフレーム n のなかでネットワークに導入されるからである。 n は、 n が無限に増大する場合のみ漸近的に N に等しくなる。第2項は、 $(1-S/(L+S))$ によって決定される優先的 (preferential) 追加によって N_{k-1} が N_k に変化するプロセスを示している。無作為に選択されたノードの近隣ノードがディグリー $k-1$ となっている確率は、無作為に選択されたリンクがディグリー $k-1$ であるようなノードに属している確率に等しい。したがってこの確率は、ディグリー $k-1$ であるようなノードに接続されたすべてのリンクを、ネットワーク内のリンクの総数で除すことで得られ、 $(k-1)N_{k-1}/mN$ となる。この確率もやはりパラメーター m に比例する。なぜなら、各々の新規ノードは既存のネットワークと m 個の接続を作るからである。したがって、分子と分母の双方から m が消去され、第2項には m が現われない。式の表現を簡単にするため、以下の関係を定義する。

$$\alpha = \frac{S}{L+S} \quad (2)$$

図2において、新規ノードの追加による N_k の変化は、以下のように、 Π_{k-m} 及び Π_{k-m+1} の差として定義できる。

$$\frac{dN_k}{dn} = \frac{N_{k-1}}{n} (\alpha m + (1-\alpha)(k-1)/2) - \frac{N_k}{n} (\alpha m + (1-\alpha)k/2) \quad (3)$$

経過時間 n の合計は、ネットワーク内のノードの総数に等しいため、 Nk/n は、無作為に選択したノードのディグリーが k である確率を示す $P(k)$ に等しくなる。さらに、平衡状態においては dNk/dn も $P(k)$ に等しくなる。というのは、 N_k/n が $nP(k)$ に等しく、 $dnP(k)/dn$ が $P(k)$ になるからである。したがって、 N_k/n 及び dN_k/dn を $P(k)$ で、また N_{k-1}/n を $P(k-1)$ で置き換えると、式(3)は次のようになる。

$$P(k) = \frac{2\alpha m + (1-\alpha)(k-1)}{2 + 2\alpha m + (1-\alpha)k} P(k-1) \quad (k \geq m+1) \quad (4)$$

この解析を完了するために、初期条件を $P(m)$ と定義する。図2において、新たなピアの追加による N_m の変動は、ディグリーが m であるような個々のノードがリンクを増すことなくディグリーを維持する確率を計算することで得られる。したがって、

$$\frac{dN_m}{dn} = 1 - \Pi_1 \quad (5)$$

ここで Π_1 は、 $N(m)$ が $N(m+1)$ に変化する確率を示す。したがって、 $1 - \Pi_1$ は、新規ノードの参加によって $N(m)$ が $N(m+1)$ に変化しない確率を示す。式(1)において、 k を $m+1$ に置き換えれば、次のようになる。

$$\Pi_1 = \alpha \frac{mN_m}{n} + (1-\alpha) \frac{mN_m}{n} \quad (6)$$

すると、式(6)を式(5)に代入すれば、初期条件 $P(m)$ は次のようになる。

$$P(m) = \frac{2}{\alpha m + m + 2} \quad (7)$$

したがって、構築されるトポロジーのディグリー分布は、パラメーター S, L, m が与えられたときに、式(4)及び(7)から再帰的に導くことができる。

ここまでの部分で導いたディグリー分布に基づいて、平均距離 AD をディグリー分布の関数として定義する必要がある。ここでは、任意のトポロジーのディグリー分布を示す次のような生成関数 $G_0(x)$ を用いる。

$$G_0(x) = \sum_{k=0}^{\infty} P(k)x^k \quad (8)$$

ここで $P(k)$ はあるトポロジーで無作為に選択したノードがディグリー k である確率である。したがって、無作為に選択したノードの平均ディグリーは、以下で与えられる。

$$\langle k \rangle = \sum_{k=0}^{\infty} kP(k) = G'_0 \quad (9)$$

ここで $G'_0(1)$ は、 $d/dx G_0(x)|_{x=1}$ を意味する。

以下の導出の理解を高めるために、図3を用意した。ここで我々が興味を惹かれるのは、無作為に選択したノード(たとえば図3におけるノード a) から無作為に選択したリンク(たとえば ℓ または ℓ') に従うことで到達するノード(たとえばノード b 又は c) のディグリー分布である。なぜなら、それによって、無作為に選択したノードから一定のホップ数以内にあるノードの数が推測できるからである。この分布は、無作為に選択したリンクの端にあるノードのディグリーによって影響される(たとえば、ディグリーの高いノードはディグリーの低いノードよりも、無作為に選択されたリンクによって到達される傾向が強い)。したがって、正規化された分布は次のように与えられる。

$$\frac{\sum_{k=0}^{\infty} kP(k)x^k}{\sum_{k=0}^{\infty} P(k)x^k} = x \frac{G'_0(x)}{G_0(x)} \quad (10)$$

ここで、無作為に選択したリンクに従って到達されたノード(図3ではノード b) のディグリーのみを考慮し、特に無作為に選択したリンクそのもの(すなわち、リンク ℓ) を排除するとすれば、 x の1乗で除する必要がある。すると、次のようになる。

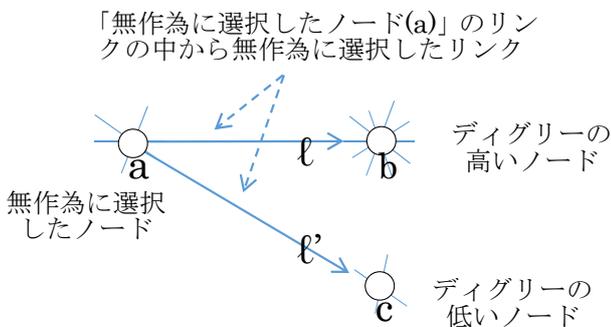


図3 無作為に選択したノード a 及びその隣接ノード b 又は c との関係。ノード a は隣接する b 及び c に到達するためにリンク (ℓ または ℓ') を無作為に選択する。たとえば、ディグリーが高いノード b は、ディグリーが低いノード c に比べて、無作為に選択されたノードによって到達される傾向が強い。

$$G_1(x) = \frac{G'_0(x)}{G'_0(1)} = \frac{G'_0(x)}{\langle k \rangle} \quad (11)$$

生成関数のプロパティを用いると、無作為に選択した元のノードの2番目に隣り合うノードの分布は、次の式で与えられる。

$$\sum_{k=0}^{\infty} P(k)[G_1(x)]^k = G_0(G_1(x)) \quad (12)$$

拡張して、 $G_0(G_1(\dots G_1(x)\dots))$ により、 M 個隣のノードの分布を生成することができる。無作為に選択したノードから M ホップ離れた隣のノードの数の分布を示す生成関数 $G^{(M)}(x)$ を次のように定義する。

$$G^{(M)}(x) = \begin{cases} G_0(x) & M=1 \\ G^{(M-1)}(G_1(x)) & M \geq 2 \end{cases} \quad (13)$$

すると、無作為に選択したノードからホップ数 M に位置する近隣ノードの平均個数は、たとえば(式(9)を参照すると) 次の式で与えられる。

$$G^{(M)'}(1) = G'_1(1)G^{(M-1)'}(1) = [G'_1(1)]^{(M-1)}G'_0(1) \quad (14)$$

式(4)及び(7)で ($a=0$) とした場合に見られるディグリー分布から、以下の式が得られる。

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)} \quad (m \leq k \leq T) \quad (15)$$

すると、このトポロジーに関する生成関数は、以下のように与えられる。

$$G_0(x) = \sum_{k=m}^T \frac{2m(m+1)}{k(k+1)(k+2)} x^k \quad (16)$$

式(14)を活用するには、式(16)を用いて $G'_0(1)$ 及び $G'_1(1)$ を定義する必要がある、式(11)より、 $G'_1(1)$ の近似値は、以下のように定義される。

$$G'_1(x) = \frac{1}{G'_0(1)} \frac{2m(T-m+1)}{(T+2)} \quad (17)$$

したがって、ES トポロジーにおいて、無作為に選択したノードからホップ数 M の位置にあるノードの平均数は、以下の式で与えられる。

$$G^{(M)'}(1) = m \left[\frac{2(T-m+1)}{(T+2)} \right]^{M-1} \quad (18)$$

ここで我々は $\langle k \rangle = m$ を用いた。さて、式(18)を使えば、ES トポロジーにおける平均距離 AD を次のように定義することができる。

$$N = \sum_{M=1}^{AD} G^{(M)}(1) \tag{19}$$

平均距離 AD は式 (18) から導くことができ、以下の式で与えられる。

$$AD = \frac{\log \left[\frac{N}{m} \left(\frac{2m(T-m+1)}{T+2} - 1 \right) + 1 \right]}{\log \left[\frac{2m(T-m+1)}{T+2} \right]} \tag{20}$$

ES モデルから得られる最適化された平均距離 AD_{opt} は、 $T \rightarrow \infty$ とすることで見いだすことができる。

$$AD_{opt} = \frac{\log \left[\frac{N}{m} (2m-1) + 1 \right]}{\log [2m]} \tag{21}$$

4 ES モデルとそのパフォーマンスの評価

本節では、先に提示した理論的解析とともにシミュレーションの結果を述べる。

4.1 平均距離 (AD) とカットオフ最大ディグリー (T)

図 4 は Symphony^[21] 及び ES によるトポロジーにおいて 2 つの異なるルーティングメカニズム、すなわち グリーディルーティング (GD ルーティング) と最短経路ルーティング (SP ルーティング) を用いたときの平均距離 (AD) のばらつきを示している。GD ルーティ

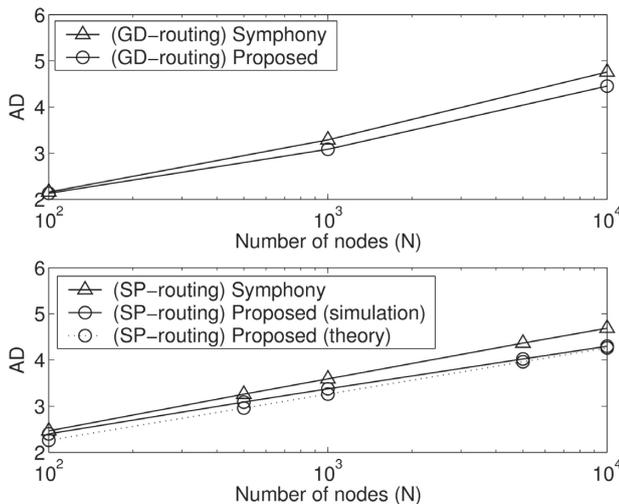


図 4 平均距離 (AD)。上: ($H=3, S=3, L=3$) としたときのグリーディルーティング (GD ルーティング)。下: ($T = \infty, S=1, L=3$) としたときの最短経路ルーティング (SP ルーティング)。我々が提案する ES モデルと Symphony の平均ディグリーはどちらのケースでも同じである (6 (GD-routing の場合) 及び 4 (SP-routing の場合) = $S+L$)。シミュレーション結果の各ポイントは 10 回の実行の平均を示している。

ングは DHT でよく用いられており、地理的な距離を減らしていくことでクエリメッセージを宛先ノードへと転送する。我々は GD ルーティングに H という追加パラメーターを導入した。これは、ホップ数 H だけ離れた近隣ノードの到達可能性情報を示すものである。結果は、パラメーター S, L, H について合理的な範囲の値において、我々が提案するアプローチのパフォーマンスは、Symphony に比べて優れたものとなった。さらに、我々の ES トポロジーに類似した、より不均一なディグリー分布を伴うトポロジーにおける GD ルーティングに関しても、複数の研究が存在する。

他方、SP ルーティングは、所与のトポロジーにおける最適化されたルーティングパフォーマンスを示すものであり、これによって、何ら偏見を伴わずに、我々の提案と Symphony を比較することができる。こうした理由から、これ以降は、我々の提案を評価するために、SP ルーティングを用いて平均距離 (AD) を取得する。また、ES モデルの理論的解析は SP ルーティングを想定して実施された。第 1 に、この結果からは、我々の提案するアプローチでは、トポロジーの平均距離が Symphony に比べ 10% 前後減少していることが分かる。レイテンシーが減少しているのは、各ノードが他のノードに到達するためのハブとなるノードとしてディグリーの高いノードを用いていることによるものである。また、理論上の結果と数値的な結果はよく対応しており、前述の式 (21) に示した理論的解析の正確さを裏付けている。

図 5 は、Symphony 及び我々の提案によって構築されたトポロジーにおける平均距離を、カットオフ最大ディグリー T の関数としてプロットしたものである。すでに説明したように、どちらのケースでも、カット

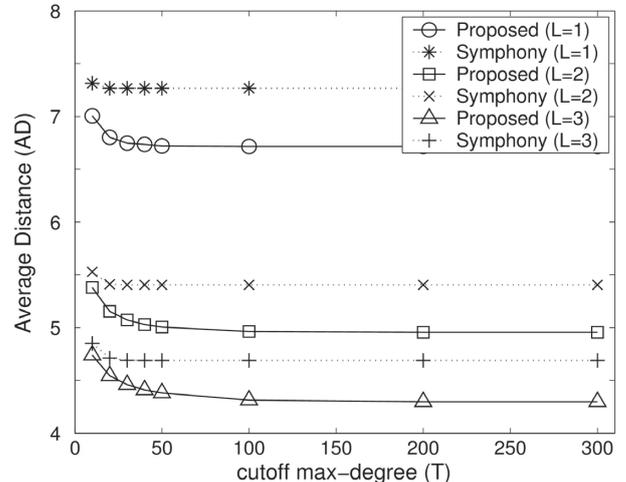


図 5 ES モデル ($N=10^4, S=1$) 及び Symphony によるトポロジーにおける AD

オフ最大ディグリー T が増大するにつれて、平均距離は減少する。パラメータ L に関しては、それがトポロジーの平均ディグリーを示すため、平均距離はパラメータ L に反比例することが見られる。ここでもう1つ興味深い点は、長距離リンクの数 L が1から2に増大すると、平均距離が大きく減少することである。しかしながら、さらに L を大きくしてもレイテンシーの減少は目立たない。

4.2 カットオフ最大ディグリー T の決定

Symphony ではシステムのパラメータを決定するために推定プロトコルを必要とするが、ES モデルでは、そうした推定プロトコルをバックグラウンドで走らせるといった運用は意図していない。そのため、ES モデルにおいて1つの非常に重要なシステム・パラ

メータとなるのがカットオフ最大ディグリー T である。このパラメータは各ノードについて設定され、そのノードの最大ディグリーを制限する。

図6は、ES モデルによるトポロジーの平均距離をプロットしたものである。様々なサイズのトポロジーにおいて、平均距離がカットオフ最大ディグリー T の関数として変動することは、理論的解析から十分に予想できる。この図をもとに許容可能なカットオフ最大ディグリー T を見出すには、式(20)の微分をある値に設定することで可能である(たとえば $\partial AD/\partial T = -1$)。

図7は、式(20)に示した発展解析モデルに基づいて、どのようにカットオフ最大ディグリー T が決定されるかを示す事例である。上の図は、最適平均距離($T=1$)を10%増大させるようなカットオフ最大ディグリーを示している。下の図は、上の図におけるカットオフ最大ディグリーに対応する平均距離を示したものである。この図から、平均距離 AD はカットオフ最大ディグリー T 及び長距離リンク L に対する感受性が低いことが分かる。したがって、我々はこれらの値の上限を決定論的に設定することができる。たとえば、カットオフ最大ディグリー T 及び長距離リンク L を、各々100前後、24に設定すると、 10^{22} 個のノードを持つESトポロジーの予想最大平均距離は、30ホップ以下になる。これに対して、Chord^[14]の場合は、 10^{22} 個のノードを持つトポロジーをサポートするために、各ノードにつき、 $(\log 2N) \approx 73$ 個の外向きのリンクを必要とし、平均距離は37(73/2)という結果となる。

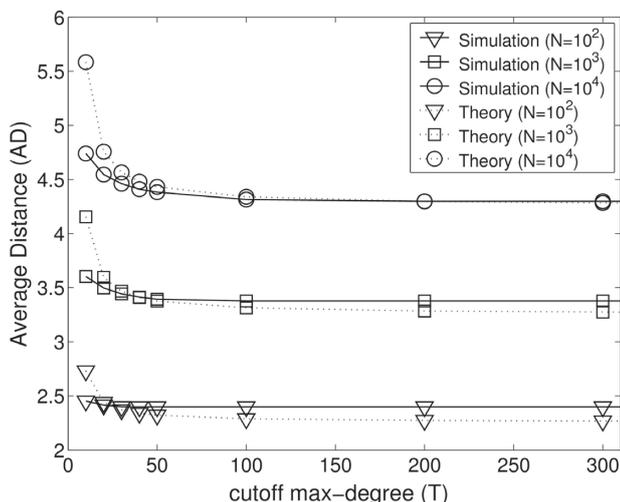


図6 シミュレーション及び理論的解析 ($S=1$, $L=3$) に基づく ES モデルによる AD トポロジー

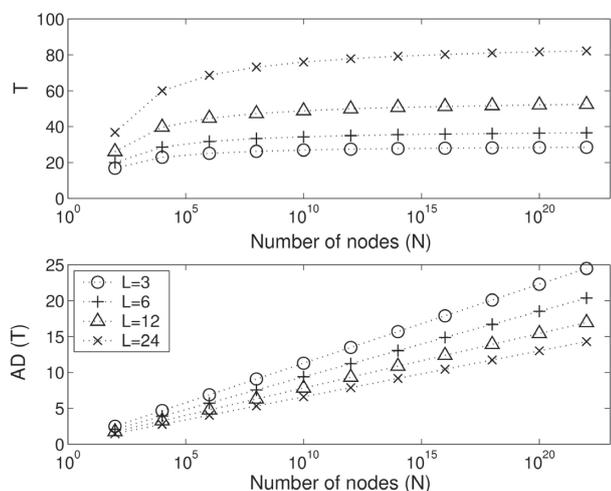


図7 平均距離 (AD)。上: 最適平均距離 ($T=無限$) を 10% 増大させるカットオフ最大ディグリー。下: 上の図のカットオフ最大ディグリーに対応する平均距離。

5 結論

本稿では、ICN のための DHT 名前解決システムのトポロジーを構築する ES モデルを紹介した。我々は、提示したモデルにより、Symphony の平均距離を約 10% 減少させることができ、その一方で、推定プロトコルをバックグラウンドで走らせる必要がないことから運用はシンプルになることを示した。さらに、理論的解析によって、非常に大規模なシナリオにおいても、ES モデルの詳細な振る舞いを把握することが可能となった。

【参考文献】

- 1 Cisco Visual Networking Index: Forecast and Methodology, 2010-2015. [Online]. Available: <http://www.cisco.com/>
- 2 D. R. Cheriton and M. Gritter, "TRIAD: a Scalable Deployable NAT-based Internet Architecture," Distributed Systems Group, Stanford University, Technical Report, Jan. 2000.
- 3 T. Koponen, A. Ermolinskiy, M. Chawla, K. H. Kim, I. Stoica, B. gon Chun, and S. Shenker, "A data-oriented (and beyond) network architecture," in Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications (SIGCOMM '07), Kyoto, Japan, Aug. 2007.

- 4 V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking Named Content," in Proceedings of the 5th international conference on Emerging networking experiments and technologies (ACM CoNEXT '09), Rome, Italy, Dec. 2009.
- 5 K. Visala, D. Lagutin, and S. Tarkoma, "LANES: An Inter-Domain Data-Oriented Routing Architecture," in Proceedings of the 2009 workshop on Re-architecting the Internet, Rome, Italy, Dec. 2009.
- 6 C. Dannewitz, "NetInf: An Information-Centric Design for the Future Internet," in Proc. 3rd GI/ITG KuVS Workshop on The Future Internet, Munich, Germany, May 2009.
- 7 S. Eum, K. Nakauchi, M. Murata, Y. Shoji, and N. Nishinaga, "CATT: Potential Based Routing with Content Caching for ICN," in Proceedings of the SIGCOMM 2012 ICN workshop, Helsinki, Finland, Aug. 2012.
- 8 ICNRG: Information Centric Networking Research Group. [Online]. Available: <http://irtf.org/icnrg/>
- 9 M. D'Ambrosio, C. Dannewitz, H. Karl, and V. Vercellone, "MDHT: a hierarchical name resolution service for Information-Centric Networks," in Proceedings of the ACM SIGCOMM workshop on Information-centric networking, Toronto, Canada, Aug. 2011.
- 10 ICN Research Challenges. [Online]. Available: <http://tools.ietf.org/html/draft-kutscher-icnrg-challenges-02>
- 11 D. Stutzbach and R. Rejaie, "Understanding Churn in Peer-to-peer Networks," in Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement, Rio de Janeiro, Brazil, Aug. 2006.
- 12 W. H. Kautz, "Bounds on directed (d, k) graphs," Theory of cellular logic networks and machines, AFCRL-68-0668 Final report, pp.20-28, 1968.
- 13 W. H. Kautz, "The design of optimum interconnection networks for multiprocessors," Architecture and Design of Digital Computer, NATO advances summer Institute, pp.249-277, 1969.
- 14 I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in Proceedings of the SIGCOMM 2001, San Diego, California, United States, Aug. 2001.
- 15 D. Li, X. Lu, and J. Wu, "FISSIONE: a scalable constant degree and low congestion DHT scheme based on Kautz graphs," in Proceedings of 24th Annual Joint Conference of the IEEE Computer and Communications Societies, Miami, FL, USA, March 2005.
- 16 D. Guo, J. Wu, H. Chen, and X. Luo, "Moore: An Extendable Peer-to-Peer Network Based on Incomplete Kautz Digraph with Constant Degree," in Proceedings of 26th IEEE International Conference on Computer Communications, Anchorage, AK, USA, May 2007.
- 17 D. Guo, Y. Liu, and X. Li, "BAKE: A Balanced Kautz Tree Structure for Peer-to-Peer Networks," in Proceedings of 27th IEEE International Conference on Computer Communications, Phoenix, AZ, USA, March 2008.
- 18 D. Guo, H. Chen, Y. He, H. Jin, C. Chen, H. Chen, Z. Shu, and G. Huang, "KCube: A novel architecture for interconnection networks," Inf. Process. Lett., pp.821-825, 2010.
- 19 P. Tvrđik, "Necklaces and scalability of Kautz digraphs," Oct. 1994.
- 20 M. A. Fiol and A. S. Lladó, "The Partial Line Digraph Technique in the Design of Large Interconnection Networks," IEEE Trans Comput., pp.848-857, 1992.
- 21 G. S. Manku, M. Bawa, and P. Raghavan, "Symphony: Distributed Hashing in a Small World," in Proceedings of the 4th usenix symposium on Internet technologies and systems, San Antonio, TEXAS, USA, June 2003.
- 22 J. Kleinberg, "Navigation in a small world." Nature, Vol.406, p.845, Aug. 2000.
- 23 V. Ramasubramanian and E. G. Sirer, "The design and implementation of a next generation name service for the Internet," SIGCOMM Comput. Commun. Rev., pp.331-342, 2004.
- 24 M. Walfish, H. Balakrishnan, and S. Shenker, "Untangling the web from DNS," in Proceedings of the 1st conference on Symposium on Networked Systems Design and Implementation - Volume 1, San Francisco, California, USA, March 2004.
- 25 L. Mathy and L. Iannone, "LISP-DHT: towards a DHT to map identifiers onto locators," in Proceedings of the 2008 ACM CoNEXT Conference, December, Madrid, Spain 2008.
- 26 R. Cox, A. Muthitacharoen, and R. Morris, "Serving DNS Using a Peer-to-Peer Lookup Service," in the First International Workshop on Peer-to-Peer Systems, IPTPS, Cambridge, MA, USA, March 2002.
- 27 C. Dannewitz, M. D'Ambrosio, and V. Vercellone, "Hierarchical DHT-based name resolution for Information Centric Networks," Elsevier Computer Communications, pp.736-749, 2013.
- 28 M. Ain et al, "D2.3 - Architecture Definition, component Descriptions, and Requirements," Deliverable, PSIRP 7th FP EU-funded project, Feb, 2009.
- 29 J. Rajahalme, M. S'arel'a, K. Visala, and J. Riihij'arvi, "On name-based inter-domain routing," Elsevier Computer Networks, pp.975-986, 2011.
- 30 K. V. Katsaros, N. Fotiou, X. Vasilakos, C. N. Ververidis, C. Tsilopoulos, G. Xylomenos, and G. C. Polyzos, "On inter-domain name resolution for Information-Centric Networks," in Proceedings of the 11th international IFIP TC 6 conference on Networking - Volume Part I, Prague, Czech Republic, May 2012.



Suyong EUM

ネットワーク研究本部ネットワークシステム
総合研究室主任研究員
Ph.D.
情報指向ネットワーク



地引昌弘 (じびき まさひろ)

ネットワーク研究本部ネットワークシステム
総合研究室専門研究員
博士(システムズ・マネジメント)
新世代ネットワーク、情報指向ネットワーク、
超大規模情報流通ネットワーク



村田正幸 (むらた まさゆき)

ネットワーク研究本部ネットワークシステム
総合研究室 R&D アドバイザー
工学博士
ネットワークアーキテクチャ



朝枝 仁 (あさえだ ひとし)

ネットワーク研究本部ネットワークシステム
総合研究室プランニングマネージャー
博士(政策・メディア)
情報指向ネットワーク



西永 望 (にしなが のぞむ)

ネットワーク研究本部ネットワークシステム
総合研究室室長
博士(工学)
新世代ネットワーク