

令和 4 年 6 月 7 日現在

機関番号：32689

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K11952

研究課題名（和文）大規模情報指向ネットワークにおける情報取得経路探索に関する研究

研究課題名（英文）Information Retrieval Path Search on Large-scale Information Centric Network

研究代表者

中里 秀則（Nakazato, Hidenori）

早稲田大学・理工学術院・教授

研究者番号：30329156

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究では以下の二つの仕組みを提案した。一つは、各ルータにCache Table (CT) というデータ構造を置き、経路とコンテンツを一時記憶する。もう一つは、定期的にコンテンツ名が類似するコンテンツを収集するHash-Collectorを使ったHash-Collecting Systemという機構である。これにより、実質的に物理的な隣接関係にあるルータに、コンテンツ名が類似するコンテンツを記憶することができるようになり、また経路も冗長化されコンテンツ発見を高速化できた。この仕組みをシミュレータに実装し、ネットワークトポロジ、アクセス頻度などを変化させて評価を行い、有効性を示すことができた。

研究成果の学術的意義や社会的意義

現在のインターネットは、ビデオコンテンツなどの増大するデータ量、セキュリティなどの課題を抱えている。これらの課題を解決する一つの手法として現在検討が進められているのが情報指向ネットワーク（ICN）である。しかし、ICNでは、大量のコンテンツを直接指定して通信を行うため、そのコンテンツの在処を如何に素早く見つけ出すかが課題となる。本研究は、この課題を解決することにより、ICNを実用化に近づけ、ネットワークの負荷やセキュリティなどのインターネットの課題の解決の一助となることを目指すものである。

研究成果の概要（英文）：We proposed two mechanisms in this research. One is a data structure called “Cache Table (CT)” which each router is equipped with and stores both route information and content itself. Another mechanism is “Hash-Collecting System” which is equipped with “Hash-Collector” collecting and storing contents with similar names. These mechanisms make routers in neighborhood store contents with similar names and redundant routes, and consequently speed up content search. The mechanisms are implemented in a simulator and the simulation showed the effectiveness of the proposed mechanisms in various network topologies, access frequencies, and so on.

研究分野：情報通信ネットワーク

キーワード：コンテンツ指向ネットワーク 情報指向ネットワーク ICN キャッシュ 経路探索 経路キャッシュ ハッシュの利用

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

ビデオ配信などにより、インターネットで流通するパケット量が増加し、それとともにサーバ負荷、ネットワーク負荷が増大している。また、インターネットには元々セキュリティ機能は含まれておらず、様々なセキュリティ上の問題が発生している。これらの問題に対処する一つの手法として情報指向ネットワーク (ICN) が提案、研究されている。

ICN では、コンテンツの名前によってパケットの配送が行われ、コンテンツの取得を行うが、コンテンツの数とそれに呼応するコンテンツ名の数は、インターネットで用いられるホスト数とそれに対応する IP アドレス数に比べても 10^5 倍以上と数が多く、その大量のコンテンツの中から要求されたコンテンツを取得するためのパケット配送経路を決めることが課題であった。

2. 研究の目的

そこで本研究では、ICN において、その特徴であるルータでのコンテンツ一時記憶も活用し、名前指定されたコンテンツに到達するパケット配送経路を効率良く探索する手法を開発することを目的とした。

3. 研究の方法

以下の点について検討し、コンテンツ探索を効率化する構成とプロトコルを開発する。

- (1) ICN ルータがもつ、コンテンツの一時記憶機能を活用する。
 - (2) コンテンツ間に何らかの関係性をもたせ、関係性の高いコンテンツの情報を論理的な近傍に集めることにより、発見しやすくする。
 - (3) 論理的な隣接ノード間の通信を可能とする仕組みを考案する。
 - (4) 関係性の高いコンテンツを収集する仕組みを組み込む。
- 開発した構成、プロトコルについてシミュレーションにより性能を評価する。

4. 研究成果

本研究で提案したコンテンツ経路探索システムを Hash-Collecting System と呼ぶ。Hash-Collecting System は、ICN の一方式である Named-Data Networking (NDN) において、コンテンツ探索を効率化する仕組みとなっている。

Hash-Collecting System でのパケット転送の流れを図 1 に示す。NDN では、パケットの転送方向を決定する戦略を Strategy と呼ぶが、Hash-Collecting System では、NDN におけるコンテンツ要求である Interest パケットの転送は二つのフェーズに分けて行われ、それぞれ異なる Strategy が適用される。

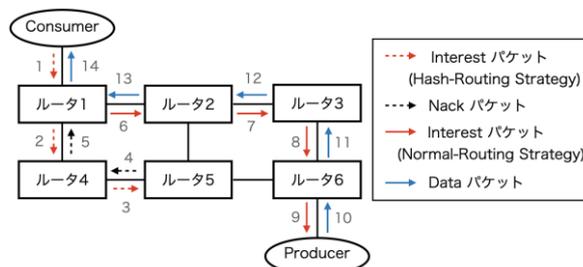


図 1 Hash-Collecting System でのパケット転送

第一フェーズでは、コンテンツ名にハッシュ関数を適用して求めたハッシュ値を用いた転送戦略である Hash-Routing Strategy が適用される。Hash-Routing Strategy で転送される Interest パケットには、転送できる転送回数 (ホップ数) の限界を定める TTL があらかじめ定める。TTL はノードを経るごとに減少し、TTL が 0 になるまでの間に Interest パケットの要求するコンテンツが見つからなかった場合には、その Interest パケットを発信した Consumer に隣接したルータまで Nack パケットが送信される。Consumer に隣接したルータが、Nack パケットを受け取ると、Interest パケットの転送は第二フェーズに移行する。

第二フェーズでは、再び同じ Interest パケットを転送するが、このフェーズでは通常の NDN ルータで行われる、Forwarding Information Base (FIB) という経路表を用いた転送を行う。ここでは、この通常の NDN の転送戦略を “Normal-Routing Strategy” と呼ぶことにする。Normal-Routing Strategy では TTL の制限はない。

Hash-Routing Strategy では、コンテンツ名のハッシュ値を用いて Interest パケットの転送方向を決定する。Hash-Routing Strategy では、各ルータに設置する Cache Table (CT) というデータ構造を参照して経路決定 (ルーティング) を行う。CT の構造を表 1 に示す。CT はコンテン

ツ名、コンテンツそのもの、コンテンツ名をハッシュ関数にかけて得られるハッシュ値と、このコンテンツを受信したインタフェースの ID (Face ID) とを記録している。Interest パケットがやってくると、ルータは、CT の中で Interest パケットのコンテンツ名のハッシュ値と「最も近い」ハッシュ値を持つ CT のエントリを探し、そのコンテンツを受信した Face の方向へと Interest パケットを転送する。

表 1 CT の構造

コンテンツ名	コンテンツ	ハッシュ値	Face ID
/prefix/A/02	00110...	3bc21a84f...	3
/prefix/B/05	1010101...	4c3ecf0a5...	1
...

二つのハッシュ値の距離は、二つのハッシュ値に対して各ビットごとに XOR を取り、1 となるビットの合計数により定義される。CT 内の全てのエントリに対し、エントリに格納されたハッシュ値を Interest パケットの運ぶコンテンツ名のハッシュ値と比較し、XOR を取った後に 1 となるビットの合計数が最も小さくなるエントリに格納されたコンテンツを、Interest パケットと「ハッシュ値が最も近い」コンテンツであると定義する。

Hash-Routing Strategy において、Interest パケットがあるノードに到達した場合、到達したノードが要求されたコンテンツを提供するサーバであれば、Interest パケットのコンテンツ名に対応したコンテンツを応答パケット (Data パケット) に格納し、Data パケットを要求元に向けて返送する。受信した Interest パケットは破棄される。到達したノードがルータである場合は、ルータはまず、CT を確認し、Interest パケットの要求するコンテンツがキャッシュされているか否かを確認する。キャッシュされている場合には、Data パケットを作成して返送し、Interest パケットは破棄される。キャッシュされていない場合には、既に同じコンテンツに対する Interest パケットを転送済みかどうかを判断し、転送済みの場合には、Interest パケットを破棄し、転送済みでない場合には Interest パケットの転送方向を次のように決定する。もし CT が空でなければ、ルータは、Interest パケットのコンテンツ名のハッシュ値と最も近いハッシュ値を持つ CT エントリを探し、そこに記録されているインタフェースへと Interest パケットを転送する。CT が空であれば、ルータは、Interest パケットの送信が可能なインタフェースのうち一つをランダムに選び、その Face の方向へ Interest パケットを転送する。

Hash-Collecting System において、CT エントリを蓄積させるために、コンテンツを積極的に収集する役割を持った、ユーザを擬似する “Hash-Collector” をネットワーク内のいくつかのノードに設置する。Hash-Collector は、コンテンツ名のハッシュ値が一定の「近さ」に収まるようなコンテンツを要求する。Hash-Collector に関わるパケットの流れを図 2 に示す。

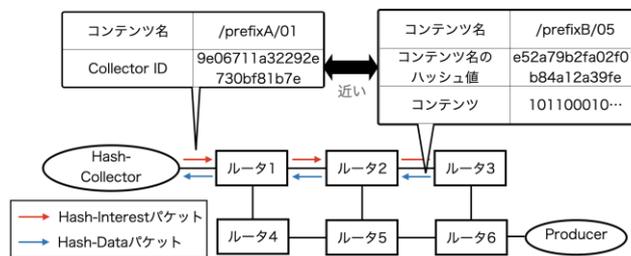


図 2 Hash-Collector に係わるパケットの流れ

それぞれの Hash-Collector は、固有の ID である “Collector ID” を持っており、その ID を Interest パケットに持たせて送信する。Collector ID は、Hash-Routing Strategy で使用したハッシュ値の取りうる値の範囲から選ばれる。Hash-Collector から送信される Interest パケットを「Hash-Interest パケット」と呼ぶ。Hash-Interest パケットは、通常の Interest パケットと異なり、自らの運ぶコンテンツ名を持ったコンテンツを要求するのではなく、自らの運ぶ Collector ID と「近い」ハッシュ値を持つ CT 内のコンテンツを要求する。通常の Interest パケットは CT もしくはサーバからコンテンツを取得するが、Hash-Interest パケットは CT からのみコンテンツを取得する。もし Hash-Interest パケットが「ハッシュ値の近い」コンテンツを発見する前にサーバに到達してしまった場合には、その場で Nack パケットが送信される。Hash-Interest パケットにはあらかじめ TTL が設定されており、この TTL が 0 になるまでの間にハッシュ値の近いコンテンツを発見できなかった場合にも Nack パケットが送信される。

ここで、Hash-Interest パケットと「ハッシュ値の近い」コンテンツの定義を説明する。前節で述べた「ハッシュ値が最も近い」コンテンツとは定義が異なるので注意が必要である。CT 内のあるエントリに対し、Hash-Interest パケットの持つ Collector ID と CT エントリに格納されたハッシュ値との XOR を取り、1 となるビットの合計数が一定の閾値に収まる場合に、このエントリに格納されたコンテンツは Hash-Interest パケットと「ハッシュ値の近い」コンテンツであるとみなされる。

Hash-Interest パケットを受け取ったルータは、CT 内に Hash-Interest パケットの運ぶ Collector ID と近いハッシュ値を持つコンテンツが一つ以上あれば、そのうち一つを、特定の判断基準を満たした場合にのみ、Hash-Collector へと返送する。この判断行程を “Collection-Judgement” と呼ぶ。Collection-Judgement の処理手順は以下の通りである。Hash-Interest パケットがノードに到着したら、ルータはまず CT を調べ、Hash-Interest パケットとハッシュ値の近いコンテンツ群を全てピックアップする。ハッシュ値の近いコンテンツが一つもない場合、Hash-Interest パケットは Hash-Routing Strategy によって次のノードへと転送される。ただし、通常の Interest パケットと異なり、Hash-Routing Strategy での転送方向の決定には、コンテンツ名のハッシュ値ではなく Collector ID が用いられる。CT 内に Hash-Interest パケットとハッシュ値の近いコンテンツが複数見つかった場合は、その中から返送するコンテンツを一つ選ぶ。

ルータは次に、先ほど選択したコンテンツを、最終的に Data パケットとして返送するか否かを、一定の確率ルールに従って決定する。この確率ルールは、あるコンテンツを返送する確率が、そのコンテンツを要求した Hash-Collector に近いノード上ほど低く、遠いノード上ほど高くなるように設定する。このように設定する理由は、Hash-Interest パケットが Hash-Collector の近くのノードでばかり探索してしまうのを防ぎ、より多様なコンテンツを収集できるようにするためである。この確率ルールの判定により、選択されたコンテンツが返送されない場合には、ルータは Hash-Routing Strategy を用いて Hash-Interest パケットを次のノードへ転送する。選択されたコンテンツが返送される場合、ルータは、「Hash-Data パケット」という Data パケットを作成し、返送する。

また、Hash-Collector に隣接したノードが、Hash-Collector の持つ Collector ID と近いハッシュ値を持つコンテンツ群を保持しておくように、“Hash-LRU” というキャッシュポリシーも提案した。Hash-LRU では、Hash-Collector に隣接したノード以外では Least Recently Used (LRU) のキャッシュ置換ポリシーが適用される。Hash-Collector に隣接したノードでは、事前に Hash-Collector の Collector ID が伝えられ、CT に空きがない場合に Collector ID と最も近いハッシュ値を持つコンテンツから置換をおこなう。

以上の提案手法をシミュレータに実装し、Data パケットが生成されてからユーザに到達するまでの平均ホップ数を計測して、通常の NDN の転送での Data パケット平均ホップ数と比較した。また、提案手法での実験パラメータを変動させ、より平均ホップ数の短縮に繋がる条件についても検討を行った。

実験に利用するネットワークトポロジとして、欧州のネットワークテストベッドである GEANT トポロジを用いた。本実験における GEANT トポロジ上のユーザ (Consumer)、Hash-Collector、サーバ (Producer) の配置設定を図 3 に示す。

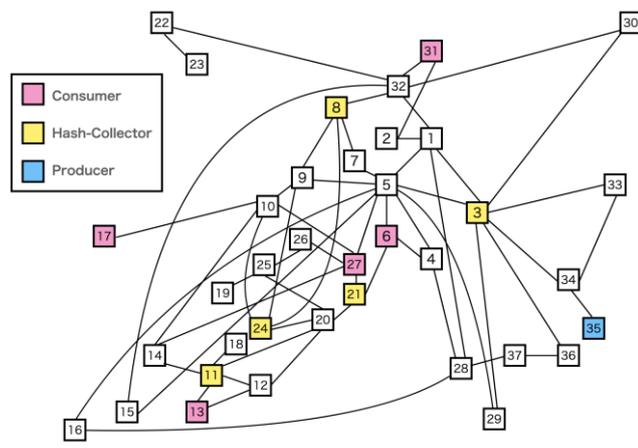


図 3 GEANT トポロジでのノード配置

Data パケットの平均ホップ数は、配置した五つの Consumer のすべてにおいて計測した。Hash-Routing Strategy で使用するハッシュ関数は、任意の文字列から 160 ビットのハッシュ値を生成する SHA-1 を用いた。Hash-Collecting System において、Interest パケットが Hash-Routing Strategy で転送される TTL、Hash-Interest パケットが転送される TTL をそれぞれ 4 ホップとした。Interest パケット及び Hash-Interest パケットは、1 秒間に 100 個のペースで 10 秒間送信される。

本実験においては、Consumer が生成するコンテンツの種類、コンテンツ名の発生確率の分布を変化させて実験を行った。以下の 4 つの設定で実験を行った。

設定 A: Hash-Collecting System を使用せず、通常の NDN の転送を行い、各 Consumer に到達した Data パケットの平均ホップ数を計測する。この場合、図 3 における Hash-Collector は使用しない。Consumer が生成するコンテンツの種類は 1000、コンテンツ名の発生確率は一様分布とする。

設定 B: Hash-Collecting System を適用し、コンテンツの種類を 1000、コンテンツ名の発生確率を一様分布に設定して計測を行う。

設定 C: 設定 B から、コンテンツの種類を 10 に変更して計測を行う。

設定 D: 設定 C からコンテンツ名の発生確率分布をジップ分布に変更して計測を行う。

本実験では、Hash-Collector の持つ Collector ID は、SHA-1 の取りうる値の範囲 ($0 \sim 2^{160} - 1$) からランダムに選択する。CT 内で「ハッシュ値の近い」コンテンツとは、Collector ID と各ビットごとに XOR を取った場合に 1 となるビットの合計数が一定の閾値に収まるハッシュ値を持ったコンテンツであると定義した。本実験では、この閾値を 48、64、80、96、112 に設定した。ハッシュ値の近いコンテンツを CT からすべてピックアップした後、その中から一つを選ぶ方法は以下の二通りとし、それぞれ条件 a、b と名付ける。

条件 a: Collector ID と最も近いコンテンツを選ぶ。

条件 b: ランダムに一つのコンテンツを選ぶ。

上記の方法でコンテンツを選択した後、コンテンツを返送するか否かを決定する確率ルールは、Hash-Interest パケットがあるノードに到達した時点での残りの TTL を t とすると、そのノードでコンテンツが返送される確率 p は以下の式で与える。

$$p = 1 - 0.2 \times t$$

設定 A~D における Data パケットの平均ホップ数の計測結果をエラー! 参照元が見つかりません。に示す。表中のノード番号は、図 3 に示された番号を示している。また、設定 D/条件 b/閾値 48 に設定した際の Data パケットの平均ホップ数の時間変化を図 4 に示す。

設定 A と B とを比較すると、通常の NDN の転送では全くキャッシュヒットが起こっていないが、Hash-Collecting System を導入したことによりキャッシュヒットが起こり、ホップ数が改善されていることが分かる。設定 B において、一部の設定では大幅に平均ホップ数が減っているが、各設定条件とホップ数の減少効果との間に明確な相関関係は見られない。次に設定 B と C とを比較すると、設定 C において閾値を 80~112 に設定した場合に B よりも大きくホップ数が減少することがわかる。このことから、提案手法は、コンテンツの種類が限られている場合により大きな効果を発揮すると言える。最後に設定 C と D とを比較すると、全体的に C よりも D の方が平均ホップ数が少ない。これにより、提案手法は、コンテンツに人気度がついている場合により効果を発揮すると言える。以上を総括すると、提案手法にはコンテンツの種類数や分布に限らずホップ数を短縮する効果、すなわち効率良くコンテンツを探索する効果があるが、特にコンテンツの種類が限られていて、人気度がついている場合により効果を発揮すると言える。

表 2 Data パケットの平均ホップ数

実験パラメータ			Consumer を設置したノード番号				
設定	条件	閾値	6	13	17	27	31
A	-	-	4.00	7.00	6.00	4.00	5.00
B	a	48	4.00	6.997	6.00	3.997	5.00
		64	3.017	6.928	6.00	3.836	4.999
		80	4.00	6.987	5.997	3.997	5.00
		96	3.996	6.967	5.996	3.996	5.00
		112	3.992	6.948	5.598	3.992	5.00
	b	48	4.00	6.935	6.00	4.00	4.996
		64	4.00	6.974	5.958	3.982	4.991
		80	3.989	6.85	5.949	2.888	4.988
		96	3.994	6.957	5.991	3.994	4.999
		112	3.968	5.435	5.967	3.966	4.721
C	a	48	4.00	7.00	6.00	4.00	4.998
		64	4.00	7.00	6.00	4.00	5.00
		80	3.991	5.896	5.991	3.991	5.00
		96	3.633	5.011	5.625	3.629	4.999
		112	3.973	5.144	5.973	3.973	5.00
	b	48	4.00	7.00	6.00	4.00	5.00
		64	4.00	7.00	6.00	4.00	5.00
		80	3.993	6.290	5.991	3.991	5.00
		96	3.615	4.840	5.395	3.615	5.00
		112	3.768	5.054	5.761	3.768	4.998
D	a	48	2.528	4.716	3.857	2.675	3.497
		64	2.745	5.016	4.170	2.852	3.718
		80	2.840	4.576	4.118	2.941	3.795
		96	2.776	4.531	4.081	2.885	3.677
		112	2.777	4.680	4.131	2.815	3.837
	b	48	2.766	5.033	3.944	2.667	3.770
		64	2.671	5.125	3.933	2.756	3.831
		80	2.679	4.640	3.899	2.739	3.747
		96	2.751	4.547	3.916	2.927	3.802
		112	2.885	4.565	4.003	2.806	3.748

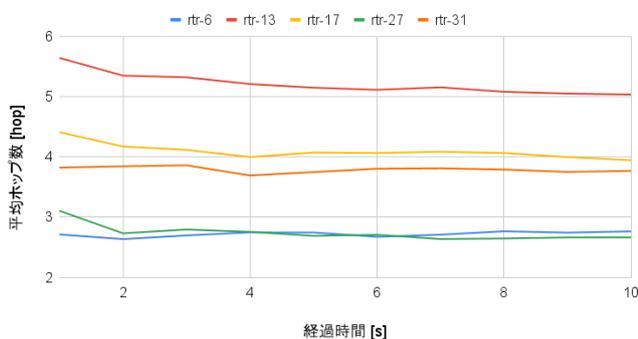


図 4 設定 D/条件 b/閾値 48 の場合の Data パケットの平均ホップ数の時間変化

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計4件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 福田 奈津子、中里 秀則
2. 発表標題 Named Data Networking への Freenet 転送手法の適用
3. 学会等名 電子情報通信学会 通信方式研究会
4. 発表年 2020年

1. 発表者名 福田奈津子、中里秀則
2. 発表標題 Hash-Collecting System: Named-Data NetworkingにおけるFreenet転送手法の導入およびコンテンツ収集の活発化
3. 学会等名 電子情報通信学会 通信方式研究会
4. 発表年 2022年

1. 発表者名 Natsuko Fukuda and Hidenori Nakazato
2. 発表標題 Hash-Collecting System : Applying Freenet Routing Method to Named Data Networking
3. 学会等名 25th Conference on Innovation in Clouds, Internet and Networks (ICIN) (国際学会)
4. 発表年 2021年

1. 発表者名 御法川 凌太、中里 秀則
2. 発表標題 NDNにおけるNLSRによるキャッシュコンテンツ広告のトラフィックへの影響
3. 学会等名 電子情報通信学会 通信方式研究会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------