

IoT サマリデータの ICN を用いたキャッシュ法

佐々木 颯[†] 上山 憲昭[†]

[†] 福岡大学工学部電子情報工学科

〒 814-0180 福岡市城南区七隈 8-19-1

E-mail: ††l161203@cis.fukuoka-u.ac.jp, ††kamiyama@fukuoka-u.ac.jp

あらまし 様々な場所に設置されたセンサーデバイスで温度などの様々な環境データを測定し、測定したデータをクラウドデータセンタに集め分析し、有益な知見を得る IoT (Internet of Things) が注目されている。IoT の有力なサービスの1つとして、特定の地域に存在するターゲット種別のセンサ情報の平均値や標準偏差を取得するサービスが考えられる。各センサが送信するデータは少量であるものの、ターゲットエリアに存在する多数のセンサのデータを収集する必要があり、通信のための名前解決やデータ転送のためのルータの転送処理など、ネットワークに与える負荷の増大が懸念される。一方、データの名称を直接用いてデータを転送する ICN (information-centric networking) が、名前解決のオーバーヘッドを回避し、ネットワークが必要なデータを収集する新しいネットワークアーキテクチャとして注目されている。そこで本論文では、IoT データを ICN を用いて転送することを想定し、ネットワーク内に流れるトラフィック量の削減を目的として、平均値や標準偏差の演算結果をルータにてキャッシュして再利用することを提案する。そして計算機シミュレーションにより、キャッシュデータを活用することによって、どの程度、1回の配信における転送データの総経由ホップ数を低減可能かを評価する。その結果、1回の配信における転送データの総経由ホップ数を、キャッシュを用いない場合と比較して、70%~99%程度、低減できることを明らかにする。

キーワード IoT, ICN, cache

Caching Method of IoT Summary Data Using ICN

Sasaki HAYATE[†] and Noriaki KAMIYAMA[†]

[†] Department of Electronic and Information Technology, Fukuoka University

8-19-1, Nanakuma, Jounan-ku, Fukuoka-shi 814-0180

E-mail: ††l161203@cis.fukuoka-u.ac.jp, ††kamiyama@fukuoka-u.ac.jp

Abstract Internet of Things (IoT), which measures various environmental data such as temperature with sensor devices installed in various places, collects the measured data in a cloud data center, analyzes the data, and obtains useful information, is attracting attention. As one of the powerful services of IoT, there is a service that obtains the average value and standard deviation of sensor information of target items existing in a specific area. Although the amount of data transmitted by each sensor is small, it is necessary to collect data of many sensors existing in the target area. There is a concern that the load on the network will increase, such as name resolution for communication and router forwarding for data transfer. On the other hand, information-centric networking (ICN), which transfers data directly using the name of the data, has attracted attention as a new network architecture that avoids the overhead of name resolution and collects necessary data by the network. In this paper, we assume that IoT data is transferred using ICN, and to reduce the amount of traffic flowing in the network, we propose to cache the average and standard deviation calculated in routers and reuse them. Then, by computer simulation, we evaluate how much we can reduce the total number of transfer hops in one transmission by utilizing the cached data summary data. As a result, we show that the total number of route hops of data transferred in one delivery can be reduced by about 70% to 99% compared to the case without using the cached summary data.

Key words IoT, ICN, cache

1. はじめに

Internet of Things (IoT) の普及に伴い、あらゆるものがインターネットにつながり新たなサービスが生まれている。データの取得対象となる IoT デバイスが固定で周期的・継続的に特定の IoT デバイスからデータを収集する場合、場所をロケータとした通信（インターネット）が有効である。しかし、ある条件を満たす全ての IoT デバイスからデータを収集するような場合、場所をロケータとした通信は非効率的である。すなわちインターネットでは事前にすべての各データに対し、存在する場所に応じたアドレスを付与し、アドレスとデータの名称との対応を DNS に登録する必要があるが、IoT では膨大な数の IoT デバイスから継続的にデータが生成されるため、このような処理を行うことは困難である。また IoT デバイスは、1 度設置されるとバッテリーの交換が困難という制約がある。データ通信を行わない時は OFF 状態にすることで IoT デバイスの消費電力を抑制できるが、データを取得したい時に OFF 状態であるとデータの取得ができない。

これらは情報指向ネットワーク (ICN: information-centric networking) を用いることで解決できる。ICN を IoT データの転送に用いることで、IoT アプリが通信先を指定しないで欲しいデータをセンシングデータ名や条件で要求可能となる。また ICN ではルータでキャッシュや配信が可能であり、IoT デバイスのデータをキャッシュしておくことで IoT デバイスが常時ネットワークに接続している必要がなくなる。さらにネットワーク内でデータの演算が可能である。そのため ICN を用いて IoT データを転送することが注目されている [2]。

IoT サービスの 1 種として、特定の地域に存在する特定の種類のセンシングデータの、平均値や標準偏差を取得するサービスが考えられる。しかし本サービスを提供するためには、ユーザの指定した地域に存在する全ての IoT デバイスからセンシングデータを収集する必要があり、指定地域の拡大に伴い、ネットワークに求められる処理負荷が急増する。しかし一方、ICN ではルータでデータをキャッシュし、ルータからデータを配信することが可能である。そのため ICN を用いた IoT データの配信を行う場合、特定の地域に対して算出したデータの総和や二乗和といったサマリデータをルータにてキャッシュしておき、後続するサービス要求に対し再利用することが可能である。そこで本稿では、任意の地域の IoT データの総和、二乗和、サンプル数をサマリデータとしてルータにてキャッシュして再利用することを提案する。本稿の貢献を以下にまとめる。

- ある地域における IoT データの総和、二乗和、サンプル数をサマリデータとしてルータにキャッシュし、より広い地域の IoT データの平均値や標準偏差を算出する際に活用する仕組みを提案する。

- 東京都千代田区を中心とする 1 辺が 16km の正方エリアを対象とした計算機シミュレーション評価により、キャッシュデータを活用することにより、どの程度、1 回の配信における転送データの総経路ホップ数を低減可能かを評価する。

以後、2 節で関連研究についてまとめ、3 節で想定条件につ

いて述べた後、4 節で提案方式について述べる。そして 5 節で計算機シミュレーションによる性能評価結果について示し、最後に 6 節で全体をまとめる。

2. 関連研究

ICN を用いて IoT データを転送するための様々な方式が検討されている。ICN では配信要求 (Interest) の転送元を經由各ルータの PIT(pending interest table) に保存することで、Interest の転送経路を逆向きにデータが配信される。そこで Gundogan らは、経路上のルータの PIT のエントリを長時間、保持することで、該当データが Publish されたときに該当 Subscriber にデータを配信することを提案している [5]。複数のランデブーノード (CP:content proxies) を用意し、各 CP は担当する prefix セットを設定してそれらコンテンツを保持し、Publisher はデータを Publish すると CP に向けて中継ルータが Publish を反復する。一方、Subscriber は Subscribe すると Interest が CP に向けて転送される。また Arshad らは、大学などの階層化された組織に IoT デバイスが配置されている環境において、階層的な名前 (組織名+場所+アクション名) とフラットな名前 (セキュリティのためのハッシュ値) を組み合わせさせた Naming 法を [1]、Bouk らは路車、車車通信を ICN で行う場合の、階層型とフラット型を組み合わせさせたネーミング法を [3]、Dong らは IoT データの要求のための Interest メッセージの新フォーマットや FIB の新フォーマット、FIB エントリの集約機能を各々、提案している [4]。

さらに Kurihara らは、NDN の Pub/Sub である COPSS でコンテンツ名に Z 記法の位置名とデータ名の組を用いたライ木を用いてルータの FIB を構築することで、大規模な Pub/Sub 型の位置名に基づく IoT データの転送方式を提案している [6]。本方式では、宛先の名前に位置を表す Z 記法に基づく 4 進数を採用し、各桁を階層的な名前の要素に対応させることで、位置ベースフォワードリングをライ木で実現し、経路を集約する方法を提案している。さらに Z 記法を用いた位置に基づく転送のための Naming 法の提案や [7]、IoT デバイスの実機を用いて ICN の様々なキャッシュ制御方式の IoT データ転送における性能の比較評価も見られる [9]。また IoT データは鮮度が重要な点に着目し、ノードをルート、中間、エッジの 3 階層に分類し、各階層に対して閾値を設定し、データの生成からの経過時間が閾値未満のときのみデータをキャッシュすることが提案されている [10]。しかしこれまでに、ルータにてキャッシュされた IoT データを活用し、平均値や標準偏差の演算結果をサマリ情報としてキャッシュして再利用することは未検討である。

3. 想定条件

対象となる IoT サービスとして、指定したエリアの指定種類のデータ (気温、気圧、降水量など) の平均値、標準偏差を取得するサービスを想定する。配信方式は、サービス利用者がデータを取得したいタイミングでデータを要求を行う Pull 型配信とする。センシングデータの属性として地域 ID、取得日時、データ種別 (気温、気圧、降水量など) を持たせる。地域 ID には Z

記法を用いる。Z 記法は図 1 に示すように対象となる正方形エリアを 4 分割し、各分割したエリアに 0 から 3 の 4 進数を割当て、さらに各分割した 0 から 3 のエリアで 4 分割を行い 4 進数を割り当てる。例えば 2 回、4 分割を行った各最小エリアは 1/1 や 3/1 のように 2 桁の 4 進数で表現できる。この 4 分割を行う処理を階層的に反復することで各桁が各階層の位置を示すことができ、指定したいエリアを任意の桁数の 4 進数の組み合わせで表現できる。ターゲットエリアを Z 記法で表現することで Interest の Prefix を Z 記法で記述でき、転送テーブル (FIB: forwarding information base) のエントリを 4 進数の組み合わせで表現でき、FIB のサイズの効果的な抑制が可能となる [6].

想定ネットワーク構成は図 2 に示すように、Z 記法で全桁を指定した各最小エリアに、各々 1 つの常時 ON 状態のゲートウェイ (GW) を設置し、GW は最短距離にあるエッジルータに接続される。IoT デバイスは自身の存在する最小エリアの GW と間欠的に接続し、センシングデータを送信し、GW は受信したデータを蓄積する。各 IoT データには有効期限 (TTL: time to live) が設定されており、TTL を超過したデータは削除される。GW には十分な容量のキャッシュが用意されており、GW のキャッシュ容量不足は生じないものとする。GW 間は K 階層のツリー状にネットワークトポロジを構成。サービス利用者は最下層ルータに接続される。

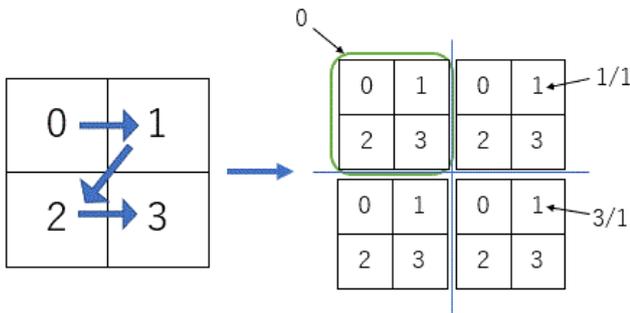


図 1 Z 記法

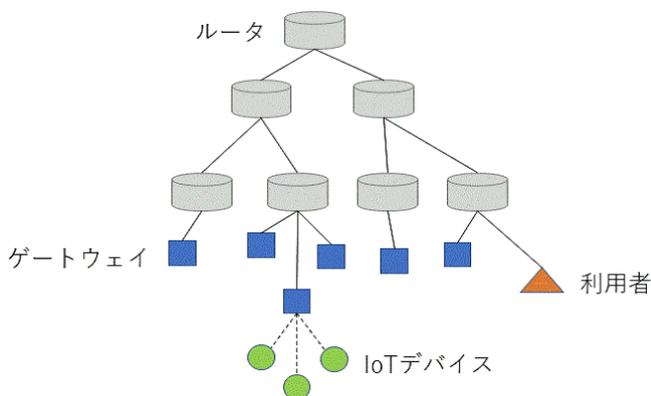


図 2 想定ネットワーク構成

4. 提案方式

各ルータにおいて、Z 記法の任意の領域に対する受信データ

の総和、データの二乗和、サンプルデバイス数の 3 つをキャッシュする。ただしサマリデータのみを転送するためデータサイズは小さいことから、ルータのキャッシュ容量制約は考慮しない。ここで k 桁の任意の Z 記法 $\mathbf{Z}_k = (Z_1, Z_2, \dots, Z_k)$ で表記されるエリアの任意のデータ種別のデータの総和を $S_1(\mathbf{Z}_k)$ 、データの二乗和を $S_2(\mathbf{Z}_k)$ 、サンプルデバイス数を $C(\mathbf{Z}_k)$ とすると、次式が成立する。

$$S_1(\mathbf{Z}_k) = \sum_{j=0}^3 S_1(\mathbf{Z}_k, j) \quad (1)$$

$$S_2(\mathbf{Z}_k) = \sum_{j=0}^3 S_2(\mathbf{Z}_k, j) \quad (2)$$

$$C(\mathbf{Z}_k) = \sum_{j=0}^3 C(\mathbf{Z}_k, j) \quad (3)$$

これらを用いて、エリア \mathbf{Z}_k の平均値 $\mu(\mathbf{Z}_k)$ 、標準偏差 $\sigma(\mathbf{Z}_k)$ は各々、次式で得られる。

$$\mu(\mathbf{Z}_k) = \frac{S_1(\mathbf{Z}_k)}{C(\mathbf{Z}_k)} \quad (4)$$

$$\sigma(\mathbf{Z}_k) = \sqrt{\frac{S_2(\mathbf{Z}_k)}{C(\mathbf{Z}_k)} - \mu(\mathbf{Z}_k)^2} \quad (5)$$

サービス利用者は、場所 Prefix、最古取得時刻、データ種別を指定して Interest を転送する。Interest を受信したルータは FIB を参照し、要求エリアに含まれる全ての GW に Interest が転送されるよう要求 Prefix と合致する全ての隣接ルータに Interest を転送する。ただし各ルータで要求エリアの全部もしくは一部のデータをキャッシュしている場合、その部位のデータについては Interest を転送せずに、そのルータから要求 GW に向かってデータを配信する。

要求エリアに含まれる GW に Interest が到達した際は、GW は時間範囲、データ種別の条件を満たす蓄積済みデータに対し、データの総和、データの二乗和、サンプルデバイス数の 3 つを計算し、その結果 (サマリデータ) を接続エッジルータに転送する。サマリデータを受信したルータは PIT に従い隣接ルータにサマリデータを転送するが、要求エリアの全データが含まれている場合には、同時に受信データをキャッシュする。その結果、任意の Z 記法エリア内の全データがキャッシュされた場合は、それらを集約したサマリデータを新たに計算し、キャッシュする。

Interest の転送例を以下に示す。例えば地域 Prefix が 1/1/2 の要求があった場合、図 3 に示すように 1/1/2 に含まれる全てのサマリデータが集まるルータ R1, R12, R123 にサマリデータ (1/1/2) をキャッシュする。また地域 Prefix が 1/1/0 の要求があった場合、図 4 に示すように、1/1/0 のデータは R112 と R121 が收容する GW に收容された IoT デバイスが取得したデータが対象となる。そのため 1/1/0 に対する Interest は R112 と R121 に転送されるが、1/1/0 の全てのサマリデータが集まるルータ R1 にのみサマリデータ (1/1/0) をキャッシュする。ルータ R11, R12 等では取得データが不完全なのでキャッシュしない。

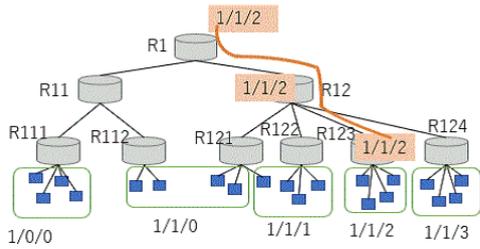


図3 地域 Prefix が 1/1/2 の要求

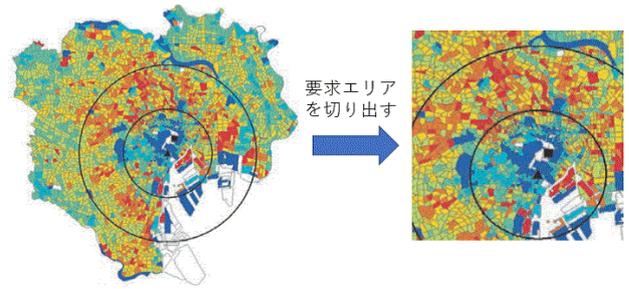


図6 評価に用いた領域の人口密度マップ

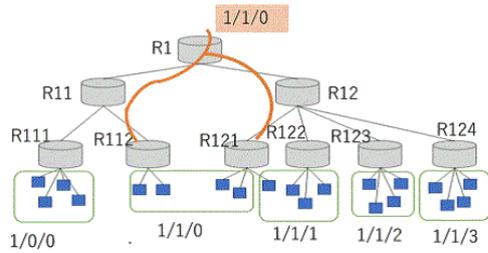


図4 地域 Prefix が 1/1/0 の要求

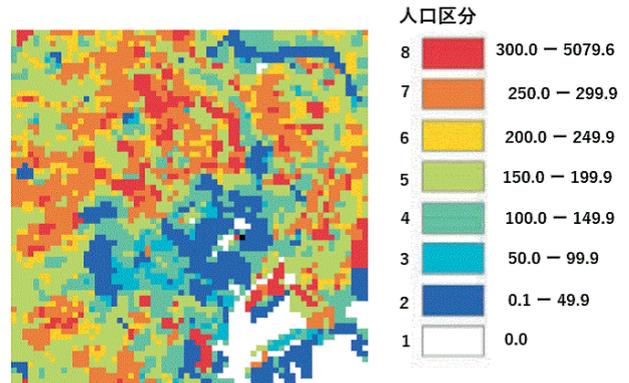


図7 分割した各最小エリアの人口区分

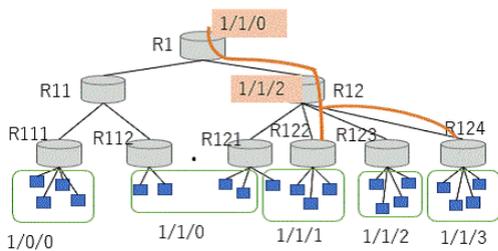


図5 地域 Prefix が 1/1 の要求

さらに図5に示すように、R1がサマリデータ(1/1/0)を、R12がサマリデータ(1/1/2)をキャッシュしている状態で、地域Prefixが1/1の要求があった場合を考える。地域Prefix1/1を構成する地域データのうち、サマリデータ(1/1/0)と(1/1/2)はルータでキャッシュされていることから、サマリデータ(1/1/1)とサマリデータ(1/1/3)のみをGWから取得する必要があり、これらに該当するGWを収容するルータR121, R122, R124にInterestが転送される。そして全てのサマリデータが集まるルータR1にサマリデータ(1/1)をキャッシュする。

5. 性能評価

5.1 シミュレーション条件

図6に示す東京都の人口密度地図[8]より、千代田区を中心とする座標(右上: 35.76, 139.82, 左上: 35.76, 139.65, 右下: 35.62, 139.82, 左下: 35.62, 139.65)の1辺が16kmの正方エリアを評価対象エリアとして切り出した。そして対象エリアを東西・南北方向に4分割する処理を6回反復する。そのため最小エリアは1辺が250mの正方エリアとなり、縦横方向に $2^6 = 64$ 個、全体で 64×64 個の最小エリアに分割される。そして各最小エリアに1つのGWを設置し、各最小エリアの人口

比に比例した数のIoTデバイスを設置する。図7のように切り出した各最小エリアにおいて、ピクセルの色が最も近い各人口区分に各最小エリアを分類する。各人口区分の人口を、タイプ1は0.0, タイプ2~7は中央値(各タイプの番号を*i*としたとき、 $50.0 \times (i - 2) + 25.0$), タイプ8は325.0に設定し、最小エリアの人口比に比例した確率でランダムに選択したGWから平均1個/秒のレートのポアソン過程で要求を発生させる。各配信要求に対し、要求エリアのZ記法の桁数*r*を $1 \leq r \leq 6$ の範囲でZipf分布 $p(z) = c / (7 - r)^\theta$ に従う確率(*c*は正規化定数)でランダムに選択し、さらに各桁の値を0から3の範囲でランダムに選択する。

ルータを評価対象エリア内に含まれる58個のNTT局舎の位置に配置する。ルータのトポロジは深さ3のツリー型を想定し、最上位のレイヤ1のルータは大手町のNTT局舎に割り当て、レイヤ2のルータは評価対象エリアが均等にカバーされる3つの駅の近くに存在する3個のNTT局舎、レイヤ3のルータはやはり評価対象エリアが均等にカバーされる12個のNTT局舎、レイヤ4のルータは残りの42個のNTT局舎に各々設置する。また各ルータを1つ上のレイヤの最寄りのルータに接続する。そして各GWを最も近い位置に存在するレイヤ4のルータに収容する。ルータの転送テーブル(FIB: forwarding information base)とキャッシュのデータ構造には探索処理の時間計算量を減らすため、トライ木を使用する。図8に示すようにトライ木の各ノードは0から3の他のトライ木ノードへのポイントを持たせる。これによりすべてのZ記法で表現できるエリアを指定することができる。FIBでは、Z記法で表された

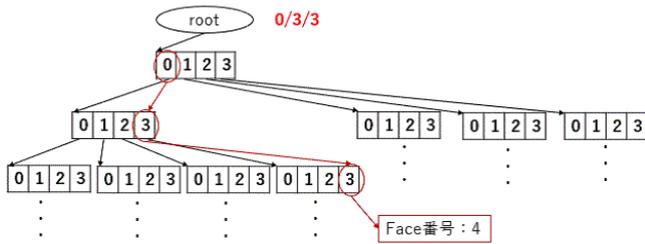


図8 トライ木

地域 ID を用いて各ルータに Interest を転送するため、トライ木の各ノードに該当する隣接ルータの出力 Face 番号を設定する。例えば要求エリアが 0/3/3 の場合、図 8 のように検索を行い、要求と合致したトライ木ノードの出力 Face 番号 4 を得る。キャッシュでは FIB と同じようにトライ木を構成し、トライ木の各ノードには出力 Face 番号ではなく、キャッシュデータを持つかを判断する 2 値変数を持たせる。

5.2 シミュレーション結果

平均総ホップ長 η を、各配信要求に対して配信されたデータを構成する各最小エリアのデータの配信ホップ長の総和の平均値と定義する。図 9 に、時間経過に伴う η をプロットする。ただし TS (Time slot) 長を 100 秒、キャッシュされたデータを保有できる有効期限 (TTL) を ∞ とする。また各配信要求の要求エリアの Z 記法桁数 r を 1 から 6 の各値に固定し、各曲線は r の各値の結果を示している。TTL を ∞ にしているの、時間経過に伴って各ルータのキャッシュにデータが蓄積される。これにより初期段階ではキャッシュデータが少なく、要求エリアが広いほど、すなわち r が小さいほど η が大きい。しかし時間の経過に伴い、要求エリアが広いほど迅速にキャッシュデータが展開されるため、 η は大きく低減する。

図 10 に、要求エリアの範囲の桁数 r を変化させたときの平均総ホップ長 η を、TTL を 5 つの値に設定したときの各々についてプロットする。要求エリアの範囲の桁数 r の増加に伴い、1 回の要求に該当する最小エリア数が減少することにより総転送データ量が減少する効果と、キャッシュにデータが存在する割合が減少することにより総転送データ量が増加する効果の 2 つの相反する効果がある。このため、TTL が小さいときはキャッシュのデータ量が減少するため r の増加に伴う最小エリア数の減少効果が大きくなるため、総ホップ長は r の増加に伴い減少する。また TTL が大きいときは両者の効果が表れるため、 r の増加に伴い総ホップ長は初め増加するが、やがて減少する。

次に r を $1 \leq r \leq 6$ の範囲で Zipf 分布 $p(z) = c/(7-r)^\theta$ に従う確率 (c は正規化定数) でランダムに選択した場合の結果を示す。図 11 に θ の値を変化させたときの平均総ホップ長 η をプロットする。ただし TTL は 100 秒、TS 長は 10,000 秒とした。キャッシュを用いない場合 (No cache) と、最小エリアの単位で経由ルータ上にキャッシュする場合 (Simple cache) と、提案方式 (Partial cache) の結果を各々示す。 θ の増加に伴いエリアが小さい要求の比率が増加し、転送データ量が減少す

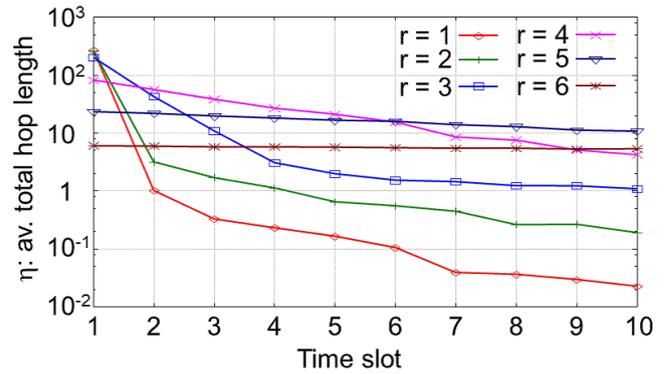


図9 時間経過に伴う平均総ホップ長 η の推移

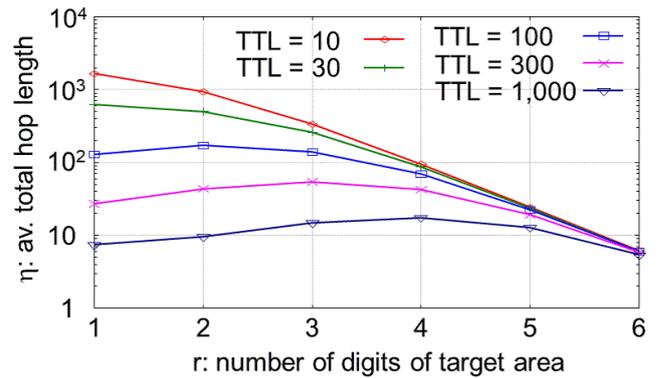


図10 r を変化させたときの平均総ホップ長 η

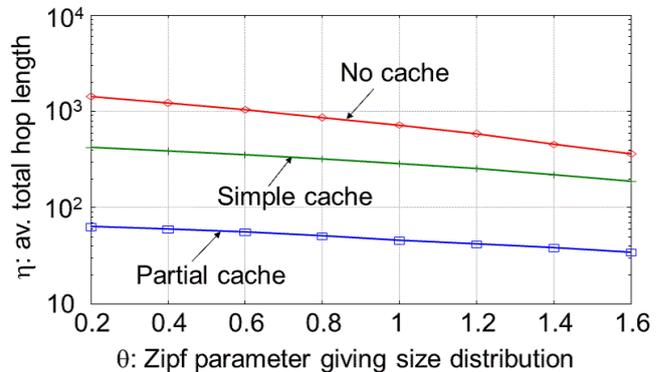


図11 θ を変化させたときの平均総ホップ長 η

るため、総ホップ長は減少するが、提案方式や最小エリア単位キャッシュ法はキャッシュの効果が低下するため、キャッシュなしと比較して総ホップ長の減少程度が小さい。提案方式は他の 2 つの方式と比較して θ の全領域で総ホップ長を大きく低減する。図 12 に $\theta = 0.7$ に設定し、TTL を変化させたときの平均総ホップ長 η をプロットする。ただし TS 長は 10,000 秒とする。TTL の増加に伴い提案方式と単純キャッシュ法はキャッシュデータ量が増加するため、総ホップ長が減少する。提案方式は他の 2 つの方式と比較して TTL の全領域で総ホップ長を大きく低減する。提案方式は 1 回の配信における転送データの総経由ホップ数を、キャッシュを用いない場合と比較して、70%~99%程度、低減できることが確認できる。

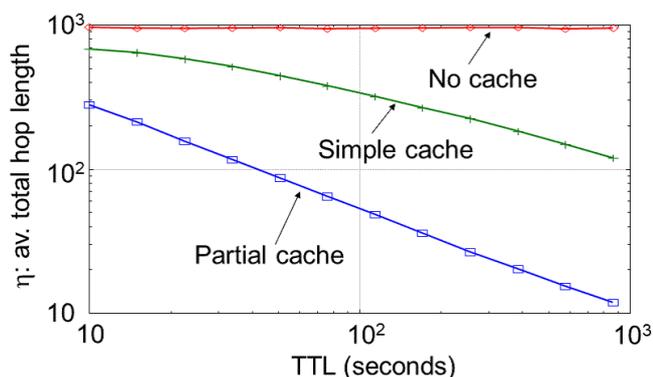


図 12 TTL を変化させたときの平均総ホップ長 η

6. ま と め

IoT の有力なサービスの 1 つとして、特定の地域に存在するターゲット項目のセンサ情報の平均値や標準偏差を取得するサービスが考えられる。各センサが送信するデータは少量であるものの、ターゲットエリアに存在する多数のセンサのデータを収集する必要があるため、通信のための名前解決やデータ転送のためのルータの転送処理など、ネットワークに与える負荷の増大が懸念される。一方、データの名称を直接用いてデータを転送する ICN が、名前解決のオーバーヘッドを回避しネットワークが必要なデータを収集する新しいネットワークアーキテクチャとして注目されている。そこで本稿では、IoT データを ICN を用いて転送することを想定し、ネットワーク内に流れるトラフィック量の削減を目的として、平均値や標準偏差の演算結果をルータにてキャッシュして再利用することを提案した。そして計算機シミュレーションにより、キャッシュデータを活用することによって、どの程度、1 回の配信における転送データの総経由ホップ数を低減可能か評価した。その結果、1 回の配信における転送データの総経由ホップ数を、キャッシュを用いない場合と比較して、70%~99%程度、低減できることを明らかにした。

謝辞 本研究成果は、JSPS 科研費 18K11283 の助成を受けたものである。ここに記して謝意を表す。

文 献

- [1] S. Arshad, et al., Hierarchical and Flat-Based Hybrid Naming Scheme in Content-Centric Networks of Things, IEEE IoTJ, 2018
- [2] S. Arshad, et al., Recent Advances in Information-Centric Networking-Based Internet of Things (ICN-IoT), IEEE Internet of Things Journal, Vol. 6, No. 2, pp. 2128-2158, Apr. 2019
- [3] S. Bouk, et al., Hierarchical and hash based naming with Compact Trie namemanagement scheme for Vehicular Content Centric Networks, Computer Communications, 71, pp.73-83, 2015
- [4] L. Dong, et al., Consumer Oriented IoT Data Discovery and Retrieval in Information Centric Networks, PIMRC 2017
- [5] C. Gundogan, et al., HoPP: Robust and Resilient Publish-Subscribe for an Information-Centric Internet of Things, LCN 2018
- [6] Y. Kurihara, et al., Location-Based Forwarding with Multi-Destinations in NDN Networks, IEICE Trans. Commun., E102-B, 9, pp.1822-1831, Sep. 2019

- [7] Y. Kurihara, et al., Compact Data Structures for Location-based Forwarding in NDN Networks, ICN-SRA 2018
- [8] 草野, 東京都 23 区における高人口密度地帯の人口・世帯・居住住宅の特徴, GIS-理論と応用, Vol. 23, No. 2, pp.17-22, 2015
- [9] J. Pfender, et al., Performance Comparison of Caching Strategies for Information-Centric IoT, ICN 2018
- [10] Z. Zhang, et al., IoT Data Lifetime-Based Cooperative Caching Scheme for ICN-IoT Networks, ICC 2018