

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 24 日現在

機関番号：32689

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24500098

研究課題名(和文) ネットワークコーディングにおける符号化ノードの最適配置に関する研究

研究課題名(英文) Optimal Encoding Node Placement for Network Coding

## 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) 3,400,000円

研究成果の概要(和文)：ネットワーク内のすべてのルータで符号化を実施すると、データ転送効率を最大化することができる。しかし、符号化を実施するためには処理能力が必要であり、ルータの負荷となる。そこで、本研究課題では、一部のノードでのみ符号化を行いながらも、すべてのノードで行う場合と同等程度の転送効率を実現するための、符号化ノード配置方法および転送方式について検討を行った。符号化によってデータ転送効率を向上できる原因が、データの冗長化による同一データの重複転送回避にあることを明らかにし、その上でネットワークトポロジの性質を表す指標を元にした符号化ノードを配置方法を提案し、効率よくデータ転送効率を向上することに成功した。

研究成果の概要(英文)：Encoding data at all nodes in a network can maximize data transmission efficiency. Encoding consumes processing power of routers, however. In this research, we investigated the placement of encoding nodes in a part of network nodes and data transmission strategies in such a network while achieving the comparable data transmission efficiency to the one in encoding at every node. We first identified that the reason of data-transfer efficiency improvement resides in the redundancy which help us to avoid duplicated data transmission. We also proposed a encoding-node placement algorithms exploiting indices that expresses features of network topologies and succeeded in improving data transmission efficiency effectively.

研究分野：総合領域

キーワード：トラヒックエンジニアリング ネットワークコーディング BitTorrent

### 1. 研究開始当初の背景

ネットワークコーディングでは、転送ノードから送出されるパケットは流入した複数のパケットを符号化(コーディング)して構築する。これにより、ネットワークのトポロジとノード間を結ぶリンクの容量による制限が許容する範囲で最大の転送レートでデータを転送することができる。

しかし、符号化に掛かる計算コストは小さくない。によれば、3.6GHzのXeonプロセッサを二つ搭載したコンピュータでも、256個のパケットから符号化を行う場合、高々280KBps程度の処理しか行うことができないことが報告されている。

我々は平成21年度から平成23年度まで、科学研究費補助金 基盤研究(C)の課題「ピア・ツー・ピア・グリッド用並列データ転送に関する研究」において、ピア・ツー・ピアネットワークにおける大容量データ高速転送に関する研究を行ってきており、ネットワークコーディングについてもその利用の可能性について検討をした。その中で、一般のネットワークコーディングで考えられているようにすべてのノードで符号化を行わない、情報源符号化(ソースコーディング)でも転送効率を大きく改善できることを発見した。このことから、少ないノードのみで符号化を行うだけでも転送効率を改善できる可能性があることがわかった。

### 2. 研究の目的

図1のようにいかに符号化の実施回数を少なくしながら、すなわち最小の数の符号化ノードによって、データ転送レートをネットワークの許容する最大帯域に近づけることを目標とした。

### 3. 研究の方法

上記目標を達成するために以下の小目標を設定し、検討を行った。

- (1) リンク帯域の制約が存在するところで、符号化を使うことによってなぜデータ転送を高速化できるのかをトポロジ的な側面から解明する。

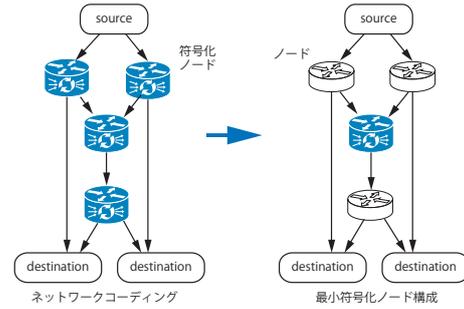


図1: 目標ネットワーク構成

- (2) リンク帯域の狭いボトルネックリンクと、当該ボトルネックリンクに対して転送帯域上有効に機能する一つの符号化ノードとのトポロジー上の位置関係を特定するとともに、最適な符号化パラメータを求め、その選定の有効性を評価する。符号化パラメータとしては、どの程度の冗長符号を発生させるかといった指標が考えられる。
- (3) 大規模なネットワークにおいて、最適な符号化ノードの配置を求めるアルゴリズムおよびその符号化パラメータ設定を求める。
- (4) 最適なコーディングノード配置問題は計算量が膨大である可能性が高いので、準最適配置を求める経験則アルゴリズムを提案し評価する。

### 4. 研究成果

#### (1) 符号化によるデータ配送の高速化

符号化によって、同一の情報を冗長に転送することを避けることができ、このことによって、ネットワークを効率よく活用できることになる。

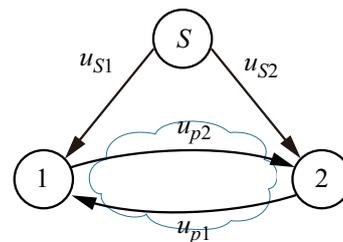


図2: ソースコーディング

図2のような、2つのピア1と2それぞれが、直接情報源ピアSと帯域  $u_{S1}$ 、 $u_{S2}$

のリンクで接続され、またピア1とピア2は他のピアを経由して相互に帯域  $u_{p1}$  と  $u_{p2}$  の経路で接続されているネットワークを仮定する。ここで情報源ピア  $S$  にある  $K$  個のブロックをすべてピア1と2に転送するとする。この場合、 $K$  個のブロックをピア  $S$  からピア1に転送するのに要する時間は、2つの経路で転送するブロックに重複が無いものと仮定すれば、

$$\frac{K}{u_{S1} + u_{p1}}$$

になる。しかし、ピア  $S$  で符号化を行わない場合、両経路で転送されるブロックに重複が発生する可能性が有り、その遅延を考慮した転送時間は

$$\frac{K}{u_{S1} + u_{p1} - D(S, 1)} \quad (1)$$

となる。ここで、 $D(S, 1)$  はピア  $S$  からピア1にブロックを送る場合に、ブロックが重複する単位時間当たりの割合である。

図2のピア  $S$  とピア1の間に他のピアがあり(このときピア1をピア  $i$  と記述することにする) またピア  $S$  とピア  $i$  を結ぶ経路が3本以上有る場合、重複するブロックによる遅延を考慮した転送時間は、式(1)と類似した以下の式によって求めることができる。

$$\frac{K}{F(S, i) - D(S, i)} \quad (2)$$

ここで  $F(S, i)$  は、ピア  $S$  からピア  $i$  までのすべての経路が許容する帯域の総和である。式(2)は、ピア  $S$  が必ずしも情報源でなく、上流からブロックを受信し、下流へ送信するピアでも同様に成り立つ。

## (2) 符号化器の配置

式(2)より、 $D(S, i)$  をできるだけ小さくすることによって、ブロック受信完了までの時間を短縮できることがわかった。

符号化を行わない場合、ピア  $S$  からピア  $i$  への経路の数が増える程、重複したブロックを受信する確率は高まり、その結果、 $D(S, j)$  は増大し、遅延も増大する。図2の場合、重複するブロックを減少し、

$D(S, i)$  を減少させるには、ピア  $S$  においてブロックを符号化してブロック数を増加させ、ピア  $S$  からピア1に送るブロックと、ピア  $S$  からピア2に送るピアに重複が無いようにすればよい。以上から、多くの経路が通過するピアにおいて符号化を実施し、異なる経路に重複しないブロックを送出すれば、データ配送時間を短縮できることがわかる。

情報転送を高速化するために、符号化器を配置するピアを決定するためには、多くのピアを結ぶ経路上にあるピアを見つけなければならない。あるピアが、他のピア間を結ぶ経路上にある度合いを示す指標として、媒介中心性 (Betweenness Centrality) とフロー中心性 (Flow Centrality) が提案されている。媒介中心性は、あるピアが、他のピア間を結ぶ何本の経路上にあるかを示す。またフロー中心性は、あるピアを経由する、他のピア間を結ぶ経路の帯域の総和を表す。

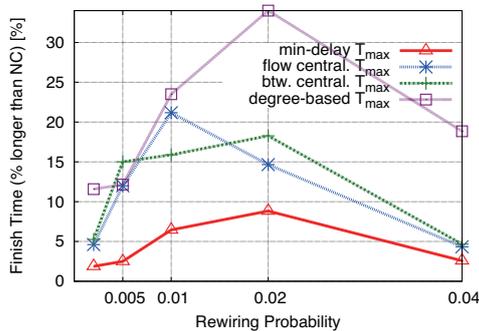
媒介中心性の大きいピアを、符号化を行うピアとして選択すれば、符号化を行うピアを多くの経路が通過することになり、重複したブロックを受信する確率を減少できることが期待できる。またフロー中心性の大きなピアを、符号化を行うピアとして選択すると、多くの経路が通過するばかりでなく、多くのブロックがそのピアを通過することになり、さらに多くの重複の削減が期待できる。

このような経験則による、アルゴリズムが考えられる一方で、式(2)から、各ピアに起因する受信完了の遅延を直接求め、大きな遅延の原因となるピアで符号化を実施することにより、遅延を抑えることも可能である。式(2)から、ピア  $i$  に起因してピア  $j$  で発生する遅延  $d(i, j)$  は

$$d(i, j) = \frac{K}{F(j) - D(i, j)} - \frac{K}{F(j)} \quad (3)$$

となり、各ピア  $i$  において  $d(i, j)$  を求め、 $d(i, j)$  の大きなピア  $i$  を符号化を行うピアとして選択するのである。なお、ここで  $F(j) = F(S, j)$ 、すなわち情報源ピア  $S$  からピア  $j$  までのすべての経路が許容する帯域の総和である。

これら三つのアルゴリズム、媒介中心性とフロー中心性を使って符号化を行うピアを選択するアルゴリズム、および大きな遅延を発生させるピアを符号化ピアとして選択するアルゴリズムの性能を比較すると、遅延に基づくアルゴリズムがこれらアルゴリズムの中で、すべてのピアで符号化を行う場合に最も近い性能を示した。すべてのピアで符号化をすれば、当然、最も短い時間ですべてのブロックの転送を完了することができる。5,000ピア程度のスモールワールドネットワークで評価した場合、遅延に基づくアルゴリズムを使って、符号化を行うピアを250だけ選んだ場合でも、ブロックの転送完了に掛かる時間の増加は、すべてのピアで符号化を行う場合と比べても10%未満で済んでおり、その有効性が示されているとともに、すべてのピアで符号化することが処理の無駄であることも示すことができた(図3)。



符号化器数 : 250

図3: 全ピアで符号化した場合との転送時間の比較

媒介中心性を使ったアルゴリズムとフロー中心性を使ったアルゴリズムを比較した場合、フロー中心性を使う方が効果が高く、条件によっては媒介中心性のアルゴリズムに比べて10%程度、遅延を短縮することができる。また、フロー中心性のアルゴリズムと遅延に基づくアルゴリズムを比較した場合も、条件によっては、その間に10%程度のブロック転送時間の差が見られた。

以上から、ブロック転送時間による性能についてまとめると、遅延に基づくア

ルゴリズムの性能が最も良く、フロー中心性を使ったアルゴリズム、媒介中心性を使ったアルゴリズムの順に性能が落ちていく。一方で、計算複雑度は、遅延に基づくアルゴリズムが  $O(VE^3)$ 、フロー中心性を使ったアルゴリズムが  $O(V^2E^2)$ 、媒介中心性を使ったアルゴリズムが  $O(E)$  となり、媒介中心性を使ったものがいへん軽量なアルゴリズムとなっている。

### (3) ブロック選択問題

BitTorrentでは、隣接ピアからダウンロードするブロックを選択する基準として、“Rarest-First”という基準が適用される。この基準では、隣接ピアが保持するブロックの中で最も希少なブロックを、ダウンロードするブロックとして選択する。もし、同程度に希少なブロックが存在する場合、その中からランダムに選択される。しかし、ピアで符号化を行う場合、この基準だけでは不十分である。

同程度に希少なブロックの中での選択基準として、

最も近くのピアによって符号化されたブロックを選ぶ

同程度の近さのピアで符号化されたブロックがある場合には、その中で最も新しく符号化されたブロックを選ぶ

という基準を採用する必要がある。

最も近くのピアによって符号化されたブロックは、符号化を行ったピアから選択を行うピアまでの経路数が小さいことが期待され、重複を防げる可能性が高い。また最も新しく符号化されたブロックには、それ以前に符号化されたブロックには含まれていない新しいブロックの情報が含まれている可能性が高く、重複を防げる可能性が高くなる。

### (4) 符号化率の最適化

図2において、ピアSからピア1に送るブロックとピア2に送るブロックに重複が無く、かつ合計K個のブロックがピア1に到着する時刻に、ピアSからの直接経路とピア2を経由する経路の二つの

経路で同時に転送を完了する場合が最短時間での転送完了となる。この条件を満たす最小の符号化後のブロック数は、

$$K \left( \frac{u_{S1}}{u_{S1} + u_{p1}} + \frac{u_{S2}}{u_{S2} + u_{p2}} \right)$$

となる。この符号化後ブロック数より以上のブロックを符号化により生成しても、転送時間をさらに短縮することはできない。

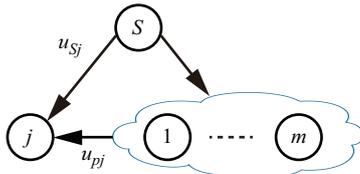


図 4: 複数ピアが直接情報源に接続する場合

図 4 のように、情報源ピア  $S$  が  $m$  個のピアに直接接続している場合、符号化前には情報源ピア  $S$  に  $K$  個のブロックがあったとすると、ブロックの重複を起こさない最小の符号化後のブロック数は、

$$\sum_{j=1}^m \frac{u_{Sj}K}{u_{Sj} + u_{pj}} \quad (4)$$

となる。 $F(j)$  は、情報源ピア  $S$  からピア  $j$  までのすべての経路が許容する帯域の総和であるから、

$$F(j) = u_{Sj} + u_{pj}$$

であり、式 (4) は以下のように書き換えることができる。

$$\sum_{j=1}^m \frac{u_{Sj}K}{F(j)} \quad (5)$$

一方、情報源ではない中間のピアが符号化を行う場合、情報源で符号化を行うのと条件が異なるのは、中間ピアでは、必ずしも情報源にあるすべてのブロックを符号化する必要がないため、符号化を行うブロックの総量を確定できない。そこで、中間ピアについては、符号化後のブロック数ではなく、符号化前後のブロック数の比率(符号化率)の最適値を求めることとする。中間ピア  $i$  での符号化率が

$e_i$  であるということは、単位時間当たりの入力ブロック数が  $B$  であるとする、単位時間当たりの出力ブロック数は  $Be_i$  となる。

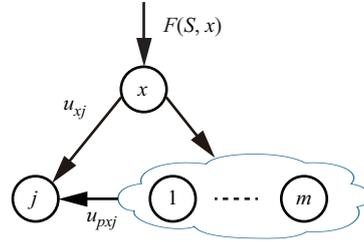


図 5: 中間ピア  $x$  による符号化

図 5 のような構成を仮定する。符号化を行う中間ピア  $x$  から、その隣接ピア  $j$  までには、二つの経路が存在する。一つはピア  $x$  が直接ピア  $j$  に接続する経路であり、その帯域は  $u_{xj}$  である。もう一つの経路は、ピア  $x$  から、ピア  $j$  以外の隣接ピアを経由してからピア  $j$  に到達する経路であり、その帯域は  $u_{pxj}$  である。

ピア  $j$  は、この二つの経路を使って、 $K_j$  個のブロックを受信する。残りの  $K - K_j$  個のブロックは、ピア  $x$  を経路しない経路を通してピア  $j$  は受信することになる。ピア  $x$  を経由する上記二つの経路を使って、ピア  $j$  が  $K_j$  個のブロックを受信するのに要する時間を  $T_j$  と表すことにする。ここで、 $T_1 \leq T_2 \leq \dots \leq T_m$  と仮定しても一般性を失わないので、その仮定をとることとする。

ピア  $j$  は二つの経路を同時に使用し、ちょうど  $T_j$  の時間で二つの経路で同時に、合計  $K_j$  個のブロックを受信完了する場合が最短となり、その時間が  $T_j$  である。 $T_j$  で受信を停止することによって、無駄な処理や帯域の使用を回避し、最適なネットワークの利用ができることになる。

下流の  $T_j$  で受信を停止することになるので、最適なネットワーク利用をするためには、ピア  $x$  での符号化率も変動させなければならないことになる。ピア 1 から  $m$  が、ピア  $x$  から受信するブロックの量は、時間とともに、すなわち、ピア  $j$  が受信を完了するかどうかによって、以下

のように変化する:

$$\text{enc}(x, t) = \begin{cases} u_{x_1}t + u_{x_2}t + \cdots + u_{x_m}t, & \text{if } 0 < t < T_1; \\ u_{x_1}T_1 + u_{x_2}t + \cdots + u_{x_m}t, & \text{if } T_1 \leq t < T_2; \\ \vdots & \\ u_{x_1}T_1 + u_{x_2}T_2 + \cdots + u_{x_m}T_m, & \text{if } t \geq T_m. \end{cases} \quad (6)$$

一方、ピア  $x$  が受信するブロック量は以下のようになる:

$$\text{recv}(x, t) = \begin{cases} F(x)t, & \text{if } 0 < t < T_x; \\ K, & \text{if } t \geq T_x. \end{cases} \quad (7)$$

式 (6) と (7) から、ピア  $x$  の符号化率  $e_x$  は以下の様に求めることができる:

$$e_x(t) = \begin{cases} (u_{x_1}t + u_{x_2}t + \cdots + u_{x_m}t)/(F(x)t), & \text{if } 0 < t < T_1; \\ (u_{x_1}T_1 + u_{x_2}t + \cdots + u_{x_m}t)/(F(x)t), & \text{if } T_1 < t < T_2; \\ \vdots & \\ (u_{x_1}T_1 + u_{x_2}T_2 + \cdots + u_{x_{j-1}}T_{j-1} + u_{x_j}t + \cdots + u_{x_m}t)/(F(x)t), & \text{if } T_i < t < T_j; \\ \vdots & \\ (u_{x_1}T_1 + u_{x_2}T_2 + \cdots + u_{x_m}T_m)/K, & \text{if } t \geq T_m; \end{cases} \quad (8)$$

ピア  $x$  において、その符号化率  $e_x$  を式 (8) のように変化させれば、重複するブロックによる転送時間の遅延を発生させず、かつ無駄な符号化を行わない最低限の符号化を行うことができる。

#### < 引用文献 >

- M. Wang and Baochun Li. How practical is network coding? In *IWQoS 2006*, pp. 274–278, June 2006.
- C. Gkantsidis and P.R. Rodriguez. Network coding for large scale content distribution. In *INFOCOM 2005*, pp. 2235–2245, March 2005.
- L.C. Freeman, “A set of measures of centrality based on betweenness,” *Sociometry*, vol.40, no.1, pp.35–41, 1977.
- L.C. Freeman, S.P. Borgatti, and D.R. White, “Centrality in valued

graphs: a measure of betweenness based on network flow,” *Social Networks*, vol.13, pp.141–154, 1991.

#### 5. 主な発表論文等

〔雑誌論文〕(計3件)

- Dinh Nguyen and Hidenori Nakazato. Network coder placement for peer-to-peer content distribution. *IEICE Trans. on Comm.*, 査読有, E96-B(7):1661–1669, July 2013, DOI:10.1587/transcom.E96.B.1661.
- Dinh Nguyen and Hidenorii Nakazato. Hybrid network coding peer-to-peer content distribution. *Journal of Computing*, 査読有, 5(4):8–17, April 2013, <https://sites.google.com/site/journalofcomputing>.
- Dinh Nguyen and Hidenorii Nakazato. Centrality-based network coder placement for peer-to-peer content distribution. *International Journal of Computer Networks and Communications*, 査読有, 5(3):157–174, May 2013, DOI:10.5121/ijcnc.2013.5312.

〔学会発表〕(計2件)

- Dinh Nguyen and Hidenorii Nakazato. Minimal network coding redundancy for peer-to-peer content distribution. In *IEEE ICC 2014*, 査読有, pp. 3112–3117, June 2014.
- Dinh Nguyen and Hidenorii Nakazato. Rarest-first and coding are not enough. In *Globecom 2012*, 査読有, pp. 2707–2712, Dec. 2012.

#### 6. 研究組織

(1) 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