社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

みんなで Order/Degree 問題を解いて究極の低遅延相互結合網をつくろう

藤原　一毅[†]　藤田　聡[††]　中野　浩嗣[††]　井上　武[†††]　鯉渕　道紘[†]

† 国立情報学研究所　〒 101–8430 東京都千代田区一ツ橋 2–1–2
†† 広島大学大学院工学研究院　〒 739–8527 広島県東広島市鏡山 1–4–1
††† NTT 未来ねっと研究所　〒 239–0847 神奈川県横須賀市光の丘 1–1
E-mail: †{ikki,koibuchi}@nii.ac.jp, ††{fujita@se,nakano@cs}.hiroshima-u.ac.jp,
† † †inoue.takeru@lab.ntt.co.jp

あらまし　大規模並列計算機の相互結合網を低遅延化するには、トポロジとして直径・平均距離の小さいグラフを採用することが効果的である。与えられたノード数と次数をもつグラフの直径・平均距離を最小化する問題は Order/Degree 問題と呼ばれるが、その効率的な解法については現在まで研究が進んでいない。我々は、低遅延な相互結合網の設計に貢献するために Order/Degree 問題の追究を加速する。その手段として、直径・平均距離の小さいグラフとその構成法を広く一般から募るコンペを開催している。本報告では、開催中の小直径グラフ探索コンペ “Graph Golf” を紹介するとともに、そこで発見されたグラフを相互結合網に応用する上での課題を検討する。

キーワード　ネットワークトポロジ, 相互結合網, Order/Degree 問題

Let’s Solve the Order/Degree Problem to Make the Lowest-latency Interconnections

Ikki FUJIWARA[†], Satoshi FUJITA[††], Koji NAKANO[††],
Takeru INOUE[†††], and Michihiro KOIBUCHI[†]

† National Institute of Informatics　2–1–2 Hitotsubashi, Chiyoda-ku, Tokyo, 101–8430 Japan
†† Graduate School of Engineering, Hiroshima University　1–4–1 Kagamiyama, Higashihiroshima-shi, Hiroshima, 739–8527 Japan
††† NTT Network Innovation Laboratories　1–1 Hikarinooka, Yokosuka-shi, Kanagawa, 239–0847 Japan
E-mail: †{ikki,koibuchi}@nii.ac.jp, ††{fujita@se,nakano@cs}.hiroshima-u.ac.jp,
† † †inoue.takeru@lab.ntt.co.jp

1. はじめに

プロセッサのメニーコア化、計算機システムの大規模並列化（チップ内：数十コア、チップ間：10 万ノード規模）が進むにつれて、計算機システムのメモリ、ストレージ、プロセッサコアなどの構成要素間を接続するネットワークである相互結合網 (interconnection networks) の通信遅延が計算システム全体に与える性能の割合が大きくなってきている。例えば、次世代の高性能システムにおける多くのマルチコア並列アプリケーションは、数百ナノ秒〜1 マイクロ秒の MPI 通信遅延を必要とすることが予測されている [1,2]。また、そのメニーコアプロセッサチップ内のパケットネットワークに関しては、プロセッサコア間通信遅延および L2 キャッシュとの通信遅延が数サイクルであることが求められており、チップ内のコア数の増大がその実現をより困難なものとしている。これらの設計では、コア間をどのように効率的に相互接続すべきか？というネットワーク構成（以下、ネットワークトポロジと呼ぶ）の設計の問題に直面している。相互結合網の通信遅延を削減する有効な 1 つの方法は、直径・平均距離の小さいネットワークトポロジを採用することである。実際に、直径・平均距離の小さなネットワークトポロジを採用することにより、多くの並列アプリケーションの性能が向上することが報告されている [3]。

ネットワークトポロジは無向グラフとしてモデル化される。与えられた頂点数と次数をもつ無向グラフの中で、直径・平均距離が小さいのはランダムグラフであることが経験的に知られている。この性質に着目し、ランダムグラフをトポロジとして

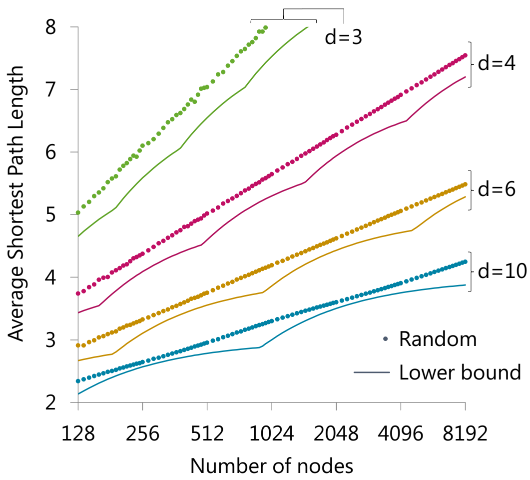

図 1　次数 d のランダムグラフの平均距離と、同じ次数のグラフの平均距離の下界

Fig. 1　The average shortest path length of random graph with degree d, and the lower bound of ASPL of the same-degree graph

採用することで相互結合網の低遅延化を図ろうとする研究が近年注目されている [4,5]。ところが、適当なランダムグラフを生成して直径・平均距離を計算してみると、その値は 4. 2 節で定義される理論的な下界よりも大きいことに気づく。図 1 は、次数 $d \in \{3,4,6,10\}$ のランダムグラフの平均距離 l と、同じ次数のグラフの平均距離の下界 L をノード数 n に対してプロットしたものである。ランダムグラフは乱数の種を変えて生成した 10 個体の中から平均距離が最小の個体を選んだ。プロットした範囲では、どの次数でも平均距離と下界のギャップ $(l-L)/L$ が最大 12%前後ある。直径についても同様に、理論的な下界とのギャップが最大 57%〜67%ある（プロットは省略）。すなわち、ランダムグラフよりもさらに平均距離の小さい「究極のグラフ」が存在する可能性が否定できない。

与えられた次数・直径を満たす中で頂点数がなるべく大きいグラフを見つける問題は Degree/Diameter 問題と呼ばれ、これまでに多くのグラフ理論家が取り組んできた成果が蓄積されている [6]。しかし、並列計算機のノード数は予算や電力などの外部要因で決まるため、Degree/Diameter 問題の解をそのままトポロジとして採用できる可能性は低い。相互結合網を設計する上では、与えられた頂点数・次数を満たす中で直径がなるべく小さい無向グラフを見つけることが求められる。これは Order/Degree 問題と呼ばれる。Order/Degree 問題は、有向グラフについては効率的な解法が知られており、それを Peer-to-Peer ネットワークのトポロジとして応用しようとする研究も存在する [7]。しかしながら、大規模並列計算機の相互結合網は双方向リンクであり、無向グラフとしてモデル化される。我々の知る限り、無向グラフの Order/Degree 問題の効率的な解法は現在までに研究されていない。

そこで我々は、グラフ理論研究者による研究の進展を待つのではなく、オープンサイエンスの手法によって Order/Degree 問題の追究を加速することを目指して、小直径グラフ探索コンペ “Graph Golf” を開催している。このコンテストは、直径・平均距離の小さいグラフとその構成法を広く一般から募集し、優れたグラフの作者を国際会議で表彰するものである。

本報告は、はじめに低遅延相互結合網の実現に向けた技術的背景を概説する（2. 章）。次に、理論的背景としてグラフの Degree/Diameter 問題と Order/Degree 問題を定義する（3. 章・4. 章）。そして、小直径グラフ探索コンペ “Graph Golf” の詳細を紹介する（5. 章）。最後に、Order/Degree 問題の解を相互結合網の設計に応用する上での課題を検討する（6. 章）。

2.　技術的背景

2. 1　スパコン・データセンターの相互結合網

Top500 ランキングに登場するスパコン 1 台を構成するプロセッサコア数が近年急増している [8]。1993 年〜2003 年当時の 1 位のスパコンのプロセッサコア数は 1 千個〜5 千個程度だったが、2014 年 11 月の記録では 3 百万個を越えている。つまり、超並列処理によりアプリケーションの実効性能を向上することが必要となっている。スパコンの相互結合網は従来、バイセクションバンド幅あるいは規則的な経路によるデータ転送のバンド幅を重要視して k-ary n-cube や fat tree が主に用いられてきたが、数百万並列時代の到来にともなって、小規模なコアによる細粒度（例：3kB 程度 [1]）の通信に適した、低遅延性を重視した設計への転換点が訪れている。

現在、InfiniBand QDR スイッチ 1 台の遅延が 100 ナノ秒程度であり、ネットワーク内ではスイッチ遅延が支配的であるため、直径 (diameter)・平均距離 (average shortest path length, ASPL) の小さいトポロジを用いることがネットワーク内の低遅延化につながる [4]。

現在、従来の fat tree や k-ary n-cubes（トーラス・メッシュ）に代わるスパコン向けネットワークトポロジについての議論が再燃している。配線のフロアプランを意識したメタトポロジである Dragonfly [9] や Skywalk [10] などが提案されており、特に高次元ネットワークを対象として盛んに議論されている。現在のところ、知られている中で最も直径・平均距離が小さいのはランダムトポロジである [4]. ランダムトポロジは多様なアクセスパターン対して低遅延・高スループットな通信を可能とすることが報告されている [4,11]。また、ランダムトポロジは任意のノード数で構築可能であり、データセンターやスパコンで強く求められる拡張性・耐故障性にも優れている。

2. 2　メニーコアプロセッサの相互結合網

メニーコアプロセッサでは、二次キャッシュおよびプロセッサコア間通信の遅延が直接的に性能に影響する。そのため、Order/Degree 問題の解に基づくグラフを用いたネットワークトポロジの実装が望まれる。3 次元積層したチップマルチプロセッサにおいて、1 枚のチップのみを（擬似的な）ランダムトポロジとしてコア間を接続するのみで、十分に平均ホップ数の削減ができることが報告されている [3]。したがって、任意のトポロジに対して、チップの動作周波数に応じてレイアウタブルにする最適化技術が有益である。我々は、これまでにこれらの最適化手法を開発しており、Order/Degree 問題の解をプロ

セッサ内ネットワークに利用可能である [12]。最近、長いリンクを直接的に実装する技術として、1 サイクルで信号伝搬可能な距離までパケット転送を実現する方法 [13] が登場している。また、研究段階ではあるが、チップ内で 60GHz 指向性電波あるいは光無線 (Free Space Optics, FSO) を使うことも提案されている。このような技術が実用化された場合、Order/Degree 問題の解をそのままプロセッサ内ネットワークに利用できるようになる。

3. 理論的背景

3.1 Degree/Diameter 問題

相互結合網をグラフとしてモデル化するにあたって重要なパラメータは、頂点数 n（ノード数を表す）、次数 d（ノードあたりのポート数を表す）、直径 k（最大通信遅延に関係する）の 3 つである。このうち 2 つのパラメータを固定し、残り 1 つのパラメータを最適化する問題として、次の 3 通りが考えられる。これらは相互に関連しているが、等価ではない（ひとつが解ければ他も解けるというものではない）。

- Degree/Diameter 問題：与えられた次数 d と直径 k を満たすグラフのうち、頂点数 n が最大であるグラフを探す問題
- Order/Degree 問題：与えられた頂点数 n と次数 d を満たすグラフのうち、直径 k が最小であるグラフを探す問題
- Order/Diameter 問題：与えられた頂点数 n と直径 k を満たすグラフのうち、次数 d が最小であるグラフを探す問題

3 つの問題の中で、グラフ理論研究者の関心は大部分が Degree/Diameter 問題に注がれてきた [14, 15]。多数の探索者によって発見された解が Wiki にまとめられ、新たな解が発見されるたびに更新されている [6]。こうして発見される頂点数 n は d, k の組合せに応じた飛び飛びの値をとる。一方、相互結合網のノード数 n は予算や電力といった外部的要因によって定められる値であり、自由に設計できるものではない。したがって、Degree/Diameter 問題の解をそのまま相互結合網のトポロジとして利用できる可能性は低い。

3.2 Moore Bound

次数 d と直径 k を満たすグラフ G の頂点数 n には Moore Bound と呼ばれる自明な上界 $N_{d,k}$ が存在する。$d = 1$ の場合は明らかに $k = 1$ であり $N_{1,1} = 2$ である。以下、本報告では $d \geqq 2$ を仮定する。G のひとつの頂点を起点とし、そこからの距離が i であるような頂点の数を n_i とする。$n_0 = 1$ であり、かつ、$1 \leqq i \leqq k$ において $n_i \leqq d(d-1)^{i-1}$ であることから、

$$
\begin{aligned}
N_{d,k} &= \sum_{i=0}^{k} n_i \\
&= 1 + d\sum_{i=1}^{k}(d-1)^{i-1} \\
&= \begin{cases} 1+2k & \text{if } d = 2 \\ 1 + d\frac{(d-1)^k - 1}{d-2} & \text{if } d > 2 \end{cases}
\end{aligned} \tag{1}
$$

が導かれる。これを Moore Bound といい、Moore Bound に等しい頂点数 $n = N_{d,k}$ を持つグラフを Moore Graph という。$k = 1$ の Moore Graph は完全グラフである。$k \geqq 2$ において、Moore Graph はいくつかの特別な d, k にしか存在しないことが知られている [16]。そのため、グラフ理論研究者は Moore Bound にできるだけ近い頂点数 n を持つグラフを探すことに多大な努力を払ってきた [15]。しかし、そうして発見される n は飛び飛びの値をとるため、Moore Bound に近いグラフを探す努力の成果は相互結合網の設計に直接応用できない。

なお、有向グラフにおいては、Kautz Graph が Moore Bound に近い頂点数をもつことが知られている。

4. Order/Degree 問題

4.1 定　　義

前節で述べた通り、相互結合網の設計では、デバイス技術的・コスト的な側面から次数と頂点数が定まる。次数と頂点数の任意の組合せに対して最小直径のネットワークを構築するためには、伝統的な Degree/Diameter 問題ではなく、Order/Degree 問題を解く必要がある。Order/Degree 問題は、与えられた次数 d と頂点数 n を満たすグラフの集合から、直径 k が最小であるグラフ G を選択する最適化問題として、次のように定義できる。

$$
\begin{aligned}
\min_{G=(V,E)} \quad & k(G), \\
\text{s.t.} \quad & |V| = n, \\
& |\delta(v)| = d \ \forall v \in V.
\end{aligned}
$$

ここで、V と E はそれぞれ G の頂点集合と辺集合であり、$k(G)$ は G の直径を表し、$\delta(v) = E(\{v\}, V \setminus \{v\})$ は v に接続する辺集合を表す。直径の等しいグラフが複数あるときは、さらに平均パス長が最小であるグラフを探索する。

Order/Degree 問題の定義は簡単であるが、解くのは難しい。トポロジの変化に対する直径の変化は不規則であり、膨大な数のグラフから最適なグラフを効率的に求める方法は知られていない。有名な巡回セールスマン問題と同様に、Order/Degree 問題も NP 困難と呼ばれる計算量クラスに属していると考えられる。つまり、n が大きくなると最適なグラフは見つからない可能性が高く、なるべく最適に近いグラフを得ることが現実的な目標となる。

制約を満たすグラフの数を大雑把に見積もってみよう。制約を満たすグラフは $m = dn/2$ 本の辺を持つ。また、頂点数 n のグラフには $m' = n(n-1)/2$ 種類の辺（頂点ペアの選び方）がありうる。よって、頂点数 n、辺数 m のグラフは

$$
\binom{m'}{m} = \frac{\frac{n(n-1)}{2}!}{\left(\frac{n(n-1)}{2} - \frac{dn}{2}\right)!\,\frac{dn}{2}!}
$$

個も存在する。この中には、個々の頂点の次数が d にならないグラフも含まれる。そこで、次数の制約を満たすグラフの数を Graphillion [17] を用いて正確に数えた結果を表 1 に示す。このように、頂点数 n に対して可能なグラフ数が急激に増加するため、探索によって最適解を発見することは絶望的である。

4.2 直径・平均距離の下界

次数 d と頂点数 n を満たすグラフ G の直径 k の自明な下界

表 1　頂点数 n、次数 d を満たすグラフの数（対称解を含む）

Table 1　Number of graphs that satisfies n and d

n	#graphs ($d=4$)	#graphs ($d=6$)
8	19355	105
9	1024380	30016
10	66462606	11180820
11	5188453830	5188453830
12	480413921130	2977635137862

$K_{n,d}$ は、Moore Bound を満たす最小の k であり、

$$K_{n,d} = \begin{cases} \left\lceil \frac{(n-1)}{2} \right\rceil & \text{if } d = 2 \\ \left\lceil \log_{d-1}\left(\frac{(n-1)(d-2)}{d} + 1 \right) \right\rceil & \text{if } d > 2 \end{cases} \quad (2)$$

となる。また、G の平均距離 l の自明な下界 $L_{n,d}$ は、G のひとつの頂点から他の $n-1$ 個の頂点までの距離の下界の平均であり、

$$L_{n,d} = \begin{cases} 1 & \text{if } K_{n,d} = 1 \\ \dfrac{S_{n,d} + K_{n,d}R_{n,d}}{n-1} & \text{if } K_{n,d} \geqq 2 \end{cases} \quad (3)$$

$$S_{n,d} = \sum_{i=1}^{K_{n,d}-1} id(d-1)^{i-1}$$

$$R_{n,d} = n - 1 - \sum_{i=1}^{K_{n,d}-1} d(d-1)^{i-1}$$

となる [18]。

なお、有向グラフにおいては、今瀬・伊藤グラフ [19] が下界より高々1 だけ大きい直径をもつことが知られているほか、いくつかの優れた小直径グラフ構成法が存在する [20]。しかし、大規模計算機の相互結合網は双方向チャネルを用いた全二重通信が普通であり、我々の関心はあくまで無向グラフにある(注1)。Degree/Diameter 問題および Order/Degree 問題の既知の解法をまとめると表 2 のようになる。このミッシングピースである「無向グラフにおける Order/Degree 問題の解」をこそ我々は欲しているのである。

5. 小直径グラフ探索コンペ “Graph Golf”

5.1 概　　要

無向グラフにおける Order/Degree 問題の追究を加速するため、我々はこの問題をオープンサイエンスの俎上に載せようと考えた。究極の目標は、あらゆる頂点数・次数の組合せに対し最小の直径・平均距離をもつグラフのカタログを作り、相互結合網の設計者に提供することである。この遠大な目標に少しでも近づくには群衆の英知を結集することが近道である。そこで我々は、Order/Degree 問題の解を一般から募集し、その直径・平均距離の小ささを競うコンペ “Graph Golf” を開催している [21]（図 2）。この名称は Code Golf (注2) になぞらえたもの

(注1)：有向グラフの辺を単に無向化すれば無向グラフになるが、その直径・平均距離は無向ランダムグラフよりも大きくなる場合がほとんどである。

(注2)：特定の動作をするプログラムのソースコードを可能な限り短いバイト数で記述する競技。ショートコーディングともいう。

表 2　Degree/Diameter 問題および Order/Degree 問題の既知の解法

Table 2　Known solutions for the Degree/Diameter Problem and the Order/Degree Problem

Graph	Degree/Diameter	Order/Degree
Undirected	Ref. [15]	Unknown
Directed	Ref. [15]	Ref. [20]

である。コンペは 2015 年以降毎年開催する予定であり、主催者は年ごとにいくつかの頂点数と次数の組合せを出題する。参加者は与えられた条件の中で直径・平均距離が小さいと思われるグラフを見つけて投稿する。投稿締切までの間、主催者は投稿された解を毎週 1 回公開する。参加者は他の参加者の解と比較して自分の順位を確認したり、他の参加者の解をベースとしてさらなる改良に取り組んだりできる(注3)。2015 年のコンペは 10 月 15 日まで投稿を受け付けており、優れた解の作者は 12 月 8 日から札幌で開催される国際会議 CANDAR’15 で表彰する予定である。参加資格は「オンライン投稿システムを利用できること」だけであり、年齢や職種による差別はない。読者もぜひ挑戦していただきたい。

5.2 ルール設計

コンテストを成功させるには、参加者に適切なインセンティブを与え、なおかつ高い公平性を確保するルール設計が肝要である。“Graph Golf” の目標は低遅延相互結合網の設計に貢献することであるが、グラフ理論研究者、アルゴリズム研究者、競技プログラミング愛好家、さらには高校生・中学生など、可能な限り幅広い参加者を集めたい。そのため、ルール設計にあたっては慎重な検討を行った。

まず、出題する頂点数 n と次数 d を検討した。計算機システム研究者の立場からは、計算機として作りやすい数（2 の累乗など）を出題するのが妥当である。一方、アルゴリズム研究者の立場からは、最適な構成法が自明でない数（素数など）が出題された方が興味を引く。そこで、2015 年のコンペでは

$$n \in \{16, 64, 256, 4096, 10000\}$$
$$d \in \{3, 4, 16, 23, 60, 64\}$$

の組合せ（ただし $d < n$）を出題した(注4)。

次に、表彰規定を検討した。計算機システム研究者の立場からは、なるべく多くの頂点数 n・次数 d の組合せに対し最善のグラフを知りたい。一方、グラフ理論研究者の立場からは、Moore Bound から導かれる平均距離の下界 $L_{n,d}$ を満たすグラフがどれだけ存在するのかに興味がある。そこで、2015 年のコンペでは次の 2 つの賞を設けた。

- Widest Improvement Award: 最多数の最善解を発見した参加者に授与される

(注3)：他の参加者がグラフを公開している場合。参加者は投稿時にグラフの公開／非公開を選択でき、非公開の場合は直径・平均距離などの統計情報のみが公開される。

(注4)：このうち $(n, d) = (64, 23), (64, 60), (256, 60), (256, 64)$ については、最初にベースラインとして提供したランダムグラフが平均距離の下界を満たしていたため、表彰対象から外した。

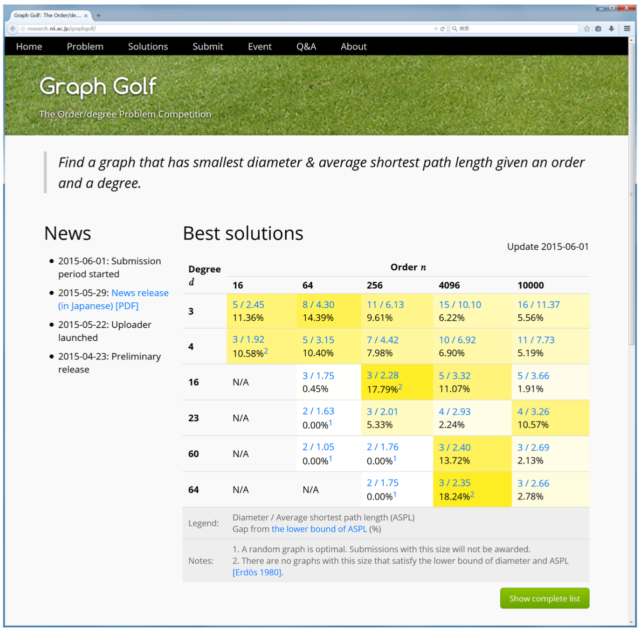

図 2　投稿受付開始前の Graph Golf ウェブサイト。初期値としてランダムグラフの直径・平均距離・平均距離ギャップ（%）が表示されている。現在の状況は http://research.nii.ac.jp/graphgolf/ を参照のこと

Fig. 2　The website of Graph Golf displaying the diameter, the average shortest path length (ASPL) and the ASPL gap of random graphs. See the above URL for the current status

- Deepest Improvement Award: すべての n, d にわたって最小の平均距離ギャップを達成した参加者に授与される

ここで最善解とは、それぞれの n, d において最小の直径 k をもつグラフ（複数ある場合は、それらの中で最小の平均距離 l をもつグラフ）である。平均距離ギャップとは、グラフの平均距離 l がその下界 $L_{n,d}$ より何パーセント大きいかを表す値であり、$p = (l - L_{n,d})/L_{n,d}$ として求められる。

同一の k, l をもつグラフが複数の作者から投稿された場合は、原則としてすべての作者を表彰する。ただし、投稿されたグラフは毎週公開されるため、その公開されたグラフをベースとして、それと同じ直径・平均距離をもつグラフを翌週以降に投稿することは容易である。そのような二番煎じをも表彰すると先行者のインセンティブを損なうため、同一スコアの解については最も早い週に投稿されたものだけを表彰対象とした。

5.3　システム設計

参加者からの投稿を受け付け、各グラフの直径・平均距離を計算し、その結果を集計・公開するためのコンテスト運営システムを内製した。このシステムは、投稿システムをパブリッククラウドに、集計システムとウェブサイトをオンプレミスサーバーに、それぞれ配備するハイブリッド構成とした。

投稿システムは、多数の参加者が殺到したとき容易に拡張できるスケーラビリティをもつことが望ましい。反面、コンテストの投稿期間が 6 月から 10 月までと長いため、平時には負荷が低いことが予想される。そこで、投稿システムは伸縮性に優れるパブリッククラウド (Amazon EC2) 上に構築することとした。このようなシステムでは RDB サーバーが性能上・拡張上のボトルネックとなる場合が多い。投稿されるグラフの実体はテキストファイルであり、投稿システムは本質的には単なるファイルアップローダである。そこで、本システムでは RDB を使用せず、投稿された 1 個のグラフファイルごとに 1 個の JSON ファイルを生成し、そこに作者名などのメタ情報を格納する設計とした。これによりボトルネックが排除され、きわめて容易に拡張可能な構成となっている。投稿システムの実装には小規模向け Web アプリケーションフレームワークである Sinatra を用いた。

グラフの直径・平均距離は主催者側で計算しなければならない。できれば投稿システムの画面上で投稿確定前に計算結果を表示したいが、全点間最短経路の計算量が大きく（$n = 10000, d = 64$ のグラフの直径・平均距離の計算は Xeon 2.9GHz のサーバーで 1 分以上かかる)、これをパブリッククラウド上で行うとユーザビリティ低下とコスト増加が懸念される。そこで、投稿システム上では簡単な書式チェックとノード数・次数のチェックだけを行い、直径・平均距離の計算はオンプレミスサーバー上の集計システムで行うこととした。集計システムは、投稿されたグラフファイルごとに直径・平均距離の計算と可視化（画像ファイル生成）を行ったあと、すべての解をスコア順に並べたデータファイルを生成する。このデータファイルを元に HTML ファイル群を生成し、ウェブサイトとして公開している。集計システムの実装には Python のグラフ処理ライブラリである NetworkX と igraph を用いた[注5]。ウェブサイトの実装には静的サイトジェネレータである Middleman を用いた。

5.4　広　　報

幅広い分野から参加者を集めるには、幅広い分野への広報が必要である。我々は、2015 年 6 月 1 日の投稿受付開始に合わせ、ウェブサイトでの告知、関連研究者が多く登録しているメーリングリストでの告知、ならびに国立情報学研究所からプレスリリースの発行を実施した。プレスリリースは日本語で書かれ、専門家以外でも理想のスパコン設計に貢献できる機会であることを強調している。このリリース内容は日刊工業新聞（6 月 9 日）と文教速報（6 月 3 日）に記事として掲載された。このほかにも、国立情報学研究所オープンハウスなど、一般の方の目に触れる機会を捉えて広報活動に取り組んでいる。

6.　応用上の課題

Order/Degree 問題の解を相互結合網の設計に応用するにあたっての課題を検討する。

6.1　ルーティング

InfiniBand、イーサネットなどのスイッチではルーティングテーブルを持っているため、同じ次数・ノード数のネットワークトポロジであれば、ルーティングアルゴリズムに関わらず、

(注5)：NetworkX は Python で書かれていてインストールが容易である。一方、igraph は C で書かれていて高速な反面、インストール時にコンパイルが必要である。参加者の利便を考え、Graph Golf ではグラフファイルの書式を「NetworkX の read_edgelist 関数で読み込めるもの」と規定した。そのため、集計システムはグラフの読み込みを NetworkX で行い、直径・平均距離の計算を igraph で行っている。

スイッチのコストはほぼ同じとなる。そのため、ルーティング、トポロジの実装という面で、Order/Degree 問題の解に基づくネットワークトポロジが、トーラス、Fat ツリーといった既存のスパコンのトポロジと比べて、大幅に困難となることはないと考えられる。

6.2 アプリケーションとの相性

ステンシル計算を行う科学技術並列処理では、隣接通信に極めて高い性能をもつトーラストポロジなどの規則網が適している。また、特定のネットワークトポロジに合わせて設計された並列アルゴリズムについても同様である。このような並列アプリケーションを対象とする場合、Order/Degree 問題の解に基づくネットワークトポロジは性能面で最善とはいえない。つまり、並列アプリケーションとネットワーク構成のコデザインによる性能チューニングを想定した場合、Order/Degree 問題の解を利用することは難しいと考えられる。

一方で、大規模な集合通信を頻繁に用いる並列アプリケーションや、データサーチ、グラフ解析などのビッグデータ処理などでメッセージ交換パターンが不規則なアプリケーションでは、計算機システムのトポロジに適した並列アルゴリズムを事前に設計することが難しいため、直径・平均距離の小さいトポロジが有利となる。我々は、今後このような「チューナブルでない」並列アプリケーションのニーズが増加し、Order/Degree 問題の解の有益性が一層増すと考えている。

Chip Multiprocessor (CMP) では、プロセッサコア間の通信遅延を削減することが性能向上の鍵となる。一方で、コヒーレンスプロトコルとネットワークトポロジのコデザインは、あまり行われていないため、並列アプリケーションの性能チューニングという面での不利益は大きくないと考えられる。

7. まとめ

グラフの Order/Degree 問題を解くことは、そのまま究極の低遅延相互結合網の設計につながる。我々はオープンサイエンスの潮流に乗って Order/Degree 問題の追究を加速するべく、小直径グラフ探索コンペ “Graph Golf” を開催し、直径・平均距離の小さいグラフとその構成法を広く一般から募集している。本報告では、Order/Degree 問題の理論的背景を述べるとともに、一般参加のコンテストである “Graph Golf” を企画・運営する上での検討事項をまとめた。また、Order/Degree 問題の解を相互結合網の設計に応用する上で課題を考察した。主著者はこのようなコンテストの運営に関する経験が浅く、改善の余地が多々あると考えている。読者諸賢からのご意見を賜り、今後の運営に役立てていきたい。

謝辞　議論を通じて有益な示唆をくださった間野暢氏に感謝する。本研究の一部は科研費#15K00144 の支援による。

文　献

[1] K.S. Hemmert, J.S. Vetter, K. Bergman, C. Das, A. Emami, C. Janssen, D.K. Panda, C. Stunkel, K. Underwood, and S. Yalamanchili, “Report on IAA Interconnection Networks Workshop 2008,” Technical report, Future Technologies Group, Oak Ridge National Laboratory, 2008.
[2] J. Tomkins, “Interconnects: A Buyers Point of View,” Technical report, ACS Workshop, Baltimore, 2007.
[3] H. Matsutani, M. Koibuchi, I. Fujiwara, T. Kagami, Y. Take, T. Kuroda, P. Bogdan, R. Marculescu, and H. Amano, “Low-latency wireless 3D NoCs via randomized shortcut chips,” Design, Automation & Test in Europe Conference & Exhibition (DATE), 2014, pp.1–6, March 2014.
[4] M. Koibuchi, H. Matsutani, H. Amano, D.F. Hsu, and H. Casanova, “A case for random shortcut topologies for HPC interconnects,” International Symposium on Computer Architecture (ISCA), pp.177–188, June 2012.
[5] A. Singla, C.-Y. Hong, L. Popa, and P.B. Godfrey, “Jellyfish: networking data centers randomly,” USENIX Conference on Networked Systems Design and Implementation (NSDI), p.17, April 2012.
[6] “The Degree/Diameter Problem - Combinatorics Wiki,” http://combinatoricswiki.org/wiki/The_Degree/Diameter_Problem.
[7] Deke Guo, Jie Wu, Yunhao Liu, Hai Jin, Hanhua Chen, and Tao Chen, “Quasi-Kautz Digraphs for Peer-to-Peer Networks,” IEEE Transactions on Parallel and Distributed Systems, vol.22, no.6, pp.1042–1055, June 2011.
[8] “TOP500 Supercomputer Sites,” http://www.top500.org/.
[9] J. Kim, W.J. Dally, S. Scott, and D. Abts, “Technology-Driven, Highly-Scalable Dragonfly Topology,” International Symposium on Computer Architecture (ISCA), pp.77–88, June 2008.
[10] I. Fujiwara, M. Koibuchi, H. Matsutani, and H. Casanova, “Skywalk: A Topology for HPC Networks with Low-Delay Switches,” International Parallel and Distributed Processing Symposium (IPDPS), pp.263–272, May 2014.
[11] X. Yuan, S. Mahapatra, W. Nienaber, S. Pakin, and M. Lang, “A new routing scheme for Jellyfish and its performance with HPC workloads,” International Conference for High Performance Computing, Networking, Storage and Analysis (SC13), pp.1–11, Nov. 2013.
[12] 藤原一毅，鯉渕道紘，“高次元トポロジ NoC の配線長最小化手法，” 並列／分散／協調処理に関する『北九州』サマー・ワークショップ (SWoPP 北九州 2013), 情報処理学会研究報告, vol.2013-ARC-206, no.12, pp.1–6, July 2013.
[13] T. Krishna, C.O. Chen, W. Kwon, and L. Peh, “Smart: Single-cycle multihop traversals over a shared network on chip,” IEEE Micro, vol.34, no.3, pp.43–56, 2014.
[14] M. Miller, “An Overview of the Degree/Diameter Problem for Directed, Undirected and Mixed Graphs,” International Workshop on Optimal Networks Topologies, pp.1–11, 2011.
[15] M. Miller and J. Širán, “Moore Graphs and Beyond: A survey of the Degree/Diameter Problem,” Electronic Journal of Combinatorics, vol.20, no.2, #DS14v2, pp.1–92, 2013.
[16] R.M. Damerell, “On Moore graphs,” Mathematical Proceedings of the Cambridge Philosophical Society, vol.74, no.02, p.227, Oct. 1973.
[17] T. Inoue, H. Iwashita, J. Kawahara, and S. Minato, “Graphillion: software library for very large sets of labeled graphs,” International Journal on Software Tools for Technology Transfer, pp.1–10, Oct. 2014.
[18] V.G. Cerf, D.D. Cowan, R.C. Mullin, and R.G. Stanton, “A lower bound on the average shortest path length in regular graphs,” Networks, vol.4, no.4, pp.335–342, 1974.
[19] 伊藤正樹，今瀬真，吉田靖之，“準最小な直径を持つ正則グラフ構成法，” 電子通信学会論文誌 A，vol.66，no.1，pp.48–55, 1983.
[20] M. Miller, Slamin, J. Ryan, and E.T. Baskoro, “Construction Techniques for Digraphs with Minimum Diameter,” Combinatorial Algorithms, eds. by T. Lecroq and L. Mouchard, vol.8288, pp.327–336, Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2013.
[21] “Graph Golf,” http://research.nii.ac.jp/graphgolf/.