

最適化手法を用いた大規模通信ネットワークの可視化・視覚分析の研究

研究代表者	中尾 彰宏	東京大学大学院情報学環	教授
共同研究者	高橋 成雄	会津大学コンピュータ理工学部	教授
共同研究者	宮村 浩子	日本原子力研究開発機構	副主幹研究員
共同研究者	吳 湘筠	慶應大学理工学部	特任助教

1 はじめに

インターネットの規模は年々拡大しており、今後もますますその規模は大きくなっていき、さらにすべてのものがインターネットに接続するようになっていくことでより複雑化していくことが考えられる。大規模化するネットワークを適切に管理、運用、制御するためには、まず、ネットワーク上を流れるデータの状況を測定し、分析することによって把握する必要がある。ネットワークの状況把握、分析には、従来ネットワーク構造特徴を表すメトリックを統計的な処理によって数値に置き換えて分析することが一般的であった。しかし、統計的な処理ではネットワークデータ全体としての特徴は取得できるが、提示できる情報が統計分析をする際に用いるメトリックに集約されて他の情報を失ってしまうため、構造のような直観的特徴の解析はできない。そこで、ネットワークデータそのものを、ノードを点、リンクを線で表現するグラフとして可視化することにより直感的にネットワークの持つ特徴を把握しやすくする手法が利用されてきた。しかし、大規模データを対象とした場合にはノードやリンクが重なって提示されてそのグラフが煩雑になる。そのため、特徴の視覚的な把握が困難になり、意味ある知見を抽出することが難しくなる問題がある。大規模データの視認性の低下を排除するために、膨大なリンクを束ねて可視化するエッジバンドリング[1]や、ノードの類似性からクラスタリングを用いる手法[2]、ネットワークグラフの中心構造(セントラリティ)を抽出する手法[3]などが提案されてきたが、視認性を高めながら特徴を提示することは難しく、大規模なネットワークにおけるデータの流れの構造を把握するには至っていない。

一方で、通信ネットワークはそれぞれの送信元からの通信を集約したのちにそれぞれの受信先へ届けられるため一般に階層構造をもち、この階層構造を把握することがネットワークの管理、運用、

制御に役立つことは広く知られている。この階層構造を把握できるようにネットワークグラフのノードやリンクを配置することで視覚的認識を助ける可視化手法の開発もなされてきた。例えば、ノードの次数や、Tier 情報を階層として可視化する手法が提案された[4]。しかしこの手法では、正確な階層を作成できなかつたり、正確な階層を得るためには階層を示す正確な情報をあらかじめ取得する必要がある。そのため、対象にすることができるネットワークデータに制約がかかることになる。視覚的認識を助けるために可視化されたネットワークグラフに対して対話的な操作をすることによって階層構造を把握する手法[5]も提案されたが、観察者が選択したノードをルートノードとしたときの階層構造を示すため、ルートノードが既知であるネットワークデータにしか適用できず、対象データに制約がかかる。

そこで本研究では、対象データに制約を設けず、大規模ネットワークデータから階層構造を抽出できるように、データの流れ情報から階層構造を求めることを最適化問題に置き換え、定式を確立する。また実際にその定式から得られる階層構造を最適化し、大規模ネットワークデータの階層構造を抽出し、可視化に利用する。本研究の特筆すべき特色は、有向リンクを持つネットワークデータから、階層構造を抽出するための最適化問題を定式化し、階層構造を推定することである。また、ここで推定した階層構造を利用したネットワークグラフのレイアウトを実現した可視化システムを構築することである。可視化対象となるデータセットとして BitTorrent を用いてインターネットから測定したログデータを用い、インターネットの中での位置関係とインターネットが物理的に存在している位置の関係が直感的にわかるようにすることで通信の状況を把握しやすくする可視化手法を検討した。これらによって、従来のネットワークデータの可視化手法では視認による分析が不可能とされてきた大規模ネットワークデータの分析に大きな飛躍をもたらすことが期待される。

2 ネットワークデータの取得

本研究では分析するネットワークとして、P2P ファイル共有ネットワークである BitTorrent を用い、インターネットから BitTorrent のピア間のネットワークトポロジ情報を測定した。測定方

法は、文献[6]の BitTorrent クローラを用いて行った。BitTorrent による P2P ファイル共有ネットワークに参加しているピアの情報を取得するために、まず測定用ピアを用意し、トラッカーから同一のコンテンツをダウンロードしている Swarmに参加しているピアの情報を取得する。さらに、新しく参加するピアに、既に参加しているピアの IP アドレスを教える役割を担う Tracker を仲介にして、用意した測定用ピアの IP アドレスを Swarmに参加しているピアに配布する。これによって、Swarmに参加しているピアから測定用ピアへの接続が生じ、その際に接続してきたピアの情報を取得できる。ここで得られたデータは、BitTorrent の Swarmに参加しているピアの IP アドレスとなる。これを繰り返すことにより、Swarmに参加するピアの IP アドレスを次々と入手することができ、Swarmを構成するすべてのピアの IP アドレスを取得する。この Swarm の IP アドレス群を測定日の CAIDA データベース[7]を参照して AS (Autonomous system) 番号に変換し、BGP のルーティング情報のログから AS 間の接続関係を取得し、ネットワークのトポロジを測定する。このトポロジ情報から AS 間の最短経路を求め、Swarmに参加しているピアの AS だけを残したものを可視化対象のデータとする。なお、AS は一般に ISP (Internet Service Provider) であるため、AS 内に複数のルータで構成される AS 内のネットワークがあるが AS 内のネットワークも 1つのノードとして扱い、さらにピアもネットワーク構造上のノードとした。つまり、AS とピアをノードとして構成されたネットワークの経路をリンクとしてグラフ化を行う。なお、今回の可視化に用いたデータセットでは、1つの Swarm あたりのノードの最大数は 1,000 程度であった。

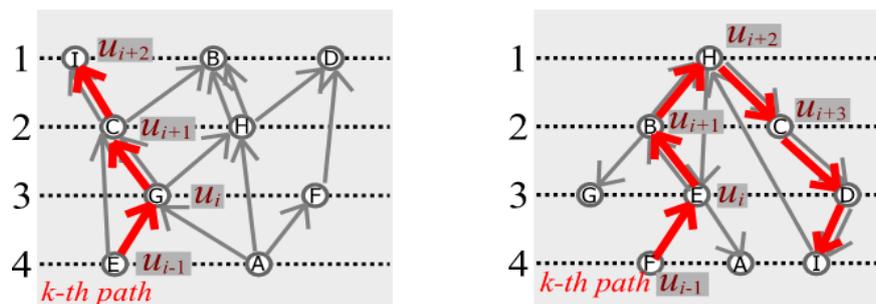
3 ネットワーク階層構造の取得

ネットワーク構造の分析には、ネットワークそのものを、点 (ノード) と線 (リンク) で表現されるグラフとして可視化する手法が効果的である。しかしながら、データの流れに関する情報を直接グラフ上に可視化しても、通信ネットワークに潜在する階層構造を視覚的に把握することが難しい。そこで本研究では、同じネットワークポリシーを共有する AS がインターネット上で階層構造をなす知見を用い[8]、その階層構造の推定を最適化問題として定式化し、P2P ネットワークにおけるピアのクライアント・カスタマ関係の視覚分析を実現する。さらに、一般に、インターネッ

トの経路は始点から上位の階層の AS 上がっていき、その後、下位の AS に下がっていくことで、終点にたどり着く。この一度下ってから上という経路はないという経験則を Valley-free という。定式化するには、インターネットの経路が Valley-free 構造であることも考慮に入れる。

本手法の基本的なアイデアは、ネットワーク上のデータの流れをパスとして入力にとり、パス上にあるノードの半順序関係 (Valley-free) を、整数計画法による最適化を用いて類推し、ノードを適切な階層にそれぞれ配置することにある。ここでは、説明の便宜上ノード間の片道パスと往復パスに話を分けて説明する。

片道パスとは、図 1 (a) のようにパス上のノードがネットワーク階層を順々に上って行くものをさすとする。パス上のノード列は、最初と最後のノードを u_a と u_b としたときに $u_a \dots, u_i, u_{i+1}, \dots, u_b$ となることに注意する。



(a) 片道パス

(b) 往復パス

図 1 片道パスと往復パスの定義

このとき、上下に隣り合うノード間の階層について以下の式が成り立つ。

$$p(u_i) - p(u_{i+1}) \geq 1$$

ここで、図 1 のように上位にある階層ほど振り分けた階層 ID は小さくなるとし、 $p(u_i)$ はノード u_i の階層 ID を表す正の整数を表すとする。しかしながら、実際のネットワークにおけるデータの流れにおいては、同じネットワーク階層にあるノード間でデータの相互通信が生じる場合がある。そのため、実際には前述した式は以下のように表記する必要がある。

$$p(u_i) - p(u_{i+1}) \geq 1 - v_{i,i+1}$$

ここで、 $-v_{i,i+1}$ はペナルティ値を示す。このペナルティ値を導入することで、相互通信が生じた場合に対応できるようにする。このペナルティ値の和が全体として小さくなるように最適化計算を行なうことで、各ノードに最適な階層 ID を振り分ける。

次に、一般的な P2P ネットワークにおけるピア間のデータ流れのパスは、図 1 (b) のように、低い階層のノードから高い階層に移動し、再度低い階層のノードにアクセスする、山形の形をなす往復パスとなることに注目する。ここで、リンク $\overline{u_i, u_{i+1}}$ が上りと下りの場合、それぞれ 0, 1 をとる論理変数 $\alpha_{i,i+1}$, $\beta_{i,i+1}$ は $\alpha_{i,i+1} + \beta_{i,i+1} = 1$ が成り立ち、以下の式を満たす。

$$p(u_i) - p(u_{i+1}) \geq +\alpha_{i,i+1} - M\beta_{i,i+1} - v_{i,i+1},$$

$$p(u_i) - p(u_{i+1}) \leq -\beta_{i,i+1} + M\alpha_{i,i+1} + v_{i,i+1},$$

ここで、ノード u_k においてパスの上りと下りが変わるとすると、 $\alpha_{k-1,k} XOR \alpha_{k,k+1}$ は 1 をとり、さらに、 $\alpha_{i-1,i} = 1 (i \leq k)$, $\beta_{i,i+1} = 1 (i \geq k)$ となる。

図 2 に、P2P ネットワーク階層を計算した例を示す。赤線は、ある特定のパスを示しており、ピアが始点、および、終点となっている。提案手法では、図 2 に示すように階層内の往復パスが AS 番号 286 を頂点として山形に表現できており、ネットワーク階層の視覚分析を実現していることがわかる。なるべく多くの経路に対して Valley-free が成り立つようにすべてのノードを配置しているため、それぞれのノードが適切な階層（階層 ID）に配置されることになる。

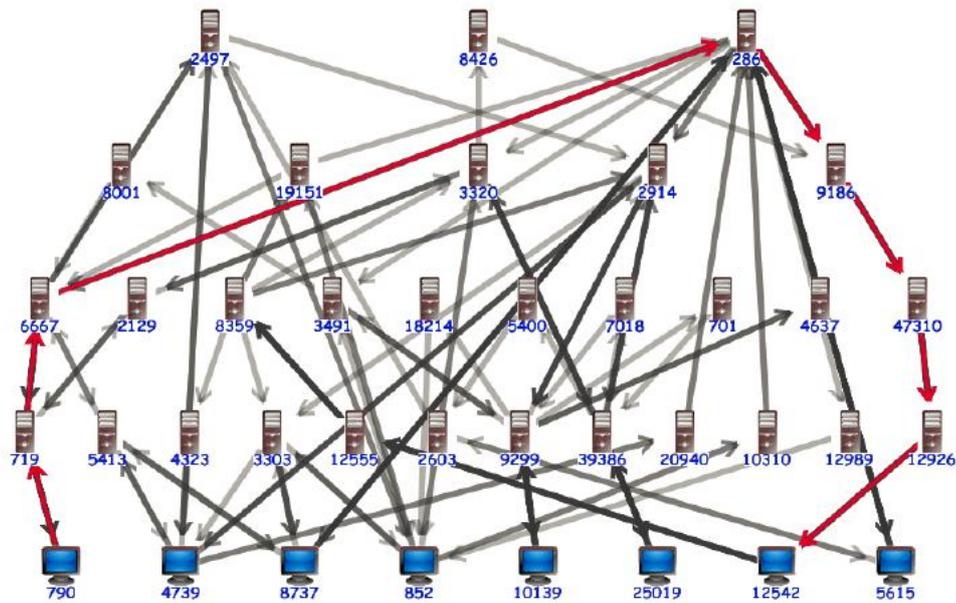


図2 ネットワーク階層計算結果例

4 階層データを用いた可視化システム

前章で求めた階層データをグラフィックレイアウトに利用した可視化システムを構築する。本システムを構築するにあたり、観察したい対象となる項目を、グラフのレイアウトや対話的操作によって取得できるようにした。

まず、グラフのレイアウトには、既にノードが持っている、AS 番号、緯度経度によって表現される地理情報、発信元・到達先・経由点の分類、Tire 階層、最適化問題によって求めた階層 ID などの属性を用いることとし、それらの情報を可視化要素に対応付けて可視化する（図 3）。

また、対話的操作については、Shneiderman が提唱した Visual Information Seeking Mantra[9] に沿ったユーザインタフェースを実装することで、観察者が注目する属性に特化した表現形式で対話的に観察できるようにした。これは可視化対象となるデータが同じであっても、分析したい属性、属性間の関係が異なる場合や、分析作業を進める過程で、興味ある対象やその詳細度が変化した場合に対応できるようにするためである。具体的には、Overview, Zoom_and_Filter, Details_on_Demand に沿ったユーザインタフェースを構築する。

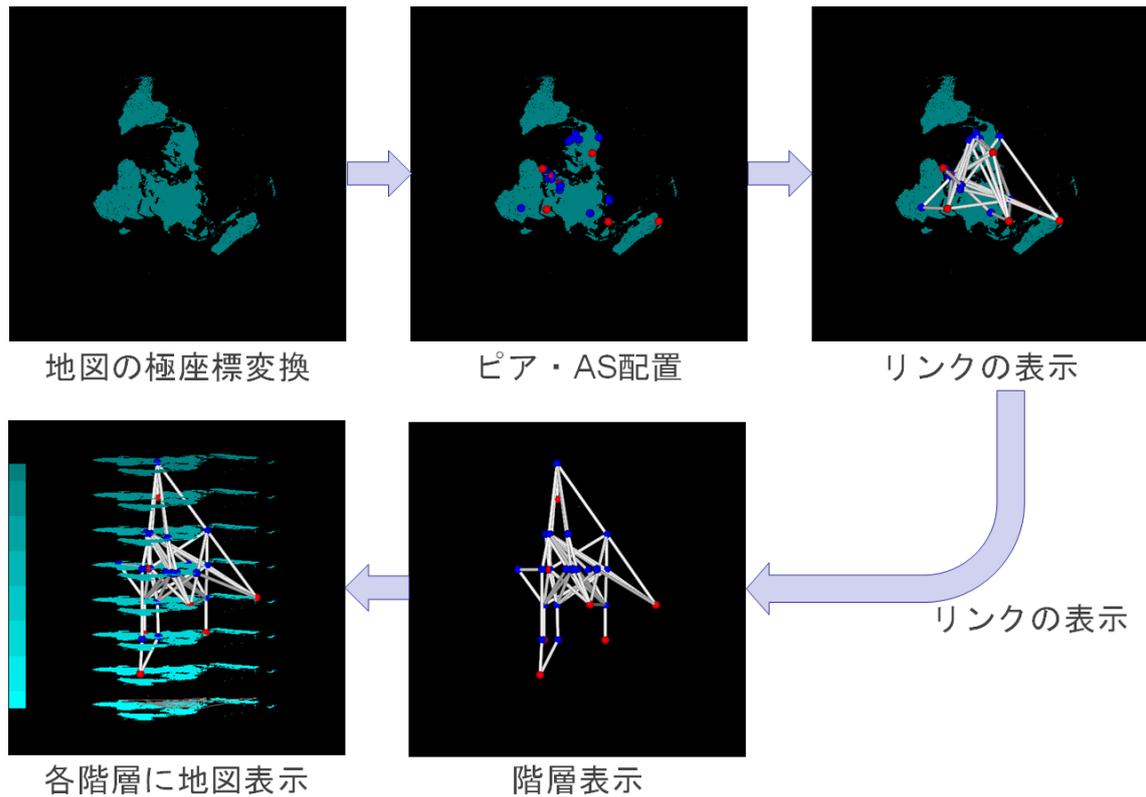


図3 グラフレイアウト

4.1 Overview

Shneiderman が提唱した Visual Information Seeking Mantra での Overview を実現するために、以下のようにグラフレイアウトをデザインする。

2次元表現：

AS 番号はネットワーク上での位置情報であるが、可視化する際にその番号を軸にしてノードを配置したとしても、人が見たときに直感的に情報を与えることができない。そこで、ピア、AS の地理情報を x , y 座標で表現し、地図も一緒に表示することで実世界でのネットワークを介したデータの流れの概要を直感的に理解できるように表示する。このとき、日本を中心に配置したメルカトル図法を用いた地図では、実際の物理的なルートとは異なり、太平洋を横断するリンクが大西洋を横断するように遠回りに描画される問題がある。そこで、地図は北極を中心に極座標系を使用することで、直感的な位置関係の情報を提示しつつ、本来の物理的経路で描画できるようにした。

色情報：

発信元と到達先であるノードと経由点であるノードの分類を色で表現する。発信元と到達先であるノードを赤，経由点であるノードを青の点で表現することで視覚的に識別可能とする。

高さ：

階層最適化計算によって得られた階層情報を z 座標（高さ方向）に割り当てることで，通信経路を山形に表現する。これによって上位階層を経由する経路の把握を容易にする。さらにそれぞれの階層の平面に地図を表示させることで，どの階層のノードでも地理的な位置情報も同時に確認することができるようにする。

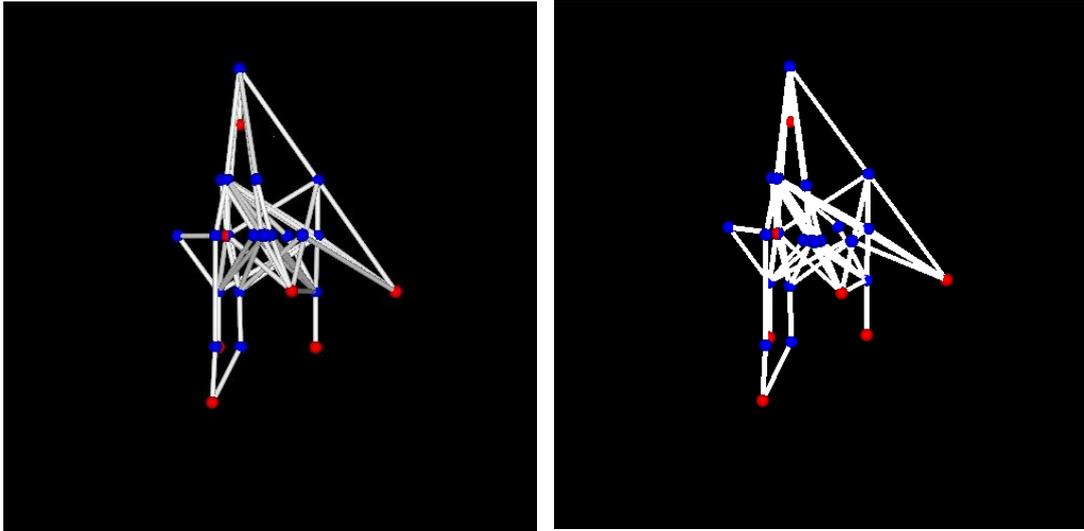
またネットワーク構造の概要をより認識できるように，観察者がユーザインタフェースを使ってグラフの描画形式の変更や表示したい情報のオン・オフを切り替えられるようにしている。以下に対話的操作機能を紹介する。

リンク，ノードの3次元表現：

混雑するリンクの視覚的な追跡を容易にするため，陰影付けた3次元表現を使用する。陰影をつけることにより，容易に奥行きを表現でき，ノード間の接続関係を把握しやすくする（図4）。これにより，遠近感が生まれ，階層に着目した場合は，隣接しているノードが異なる階層に属しているといったことや，地理的な位置関係に着目した場合は，隣接しているノード間で，どの地理的な位置にあるのかが，直感的に把握しやすくなる。

リンク曲線表示：

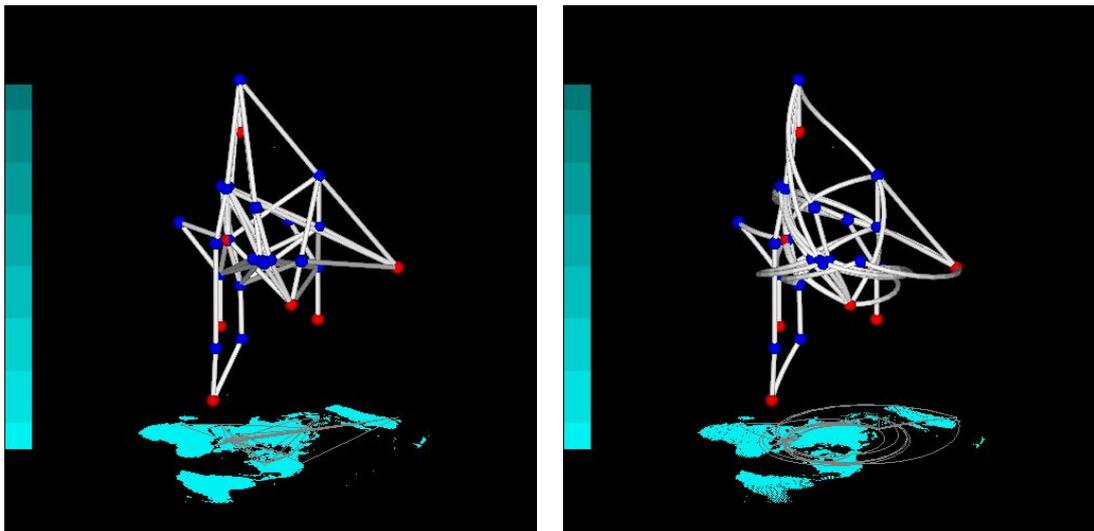
リンクは直線，曲線を切り替えられる（図5）。つまり，最短の距離を把握したい場合は，直線表示をすることで，距離を直感的に把握しやすくなる。そして，直線表示ではリンクが重なり合い，視認性が低下する場合には，リンクを曲線表示することでリンクの重なりを回避することで視認性を高め，ノード間の接続関係を把握しやすくする。



(a) 3次元表現（陰影あり）

(b) 2次元表現（陰影なし）

図4 リンクの3次元表現効果の比較



(a) 直線表現

(b) 曲線表現

図5 リンクの直線・曲線表現の比較，影付き表現

xy 平面投影：

xy 平面にノードとリンクを投影した図（影）を作成することで、 z 軸方向には階層情報を残した経路を表示しながら、影が投影された平面から階層情報を破棄した地理的なデータの経路を観察可能とする（図5）。これまで同時に把握することが難しかったネットワーク内の位置情報とネッ

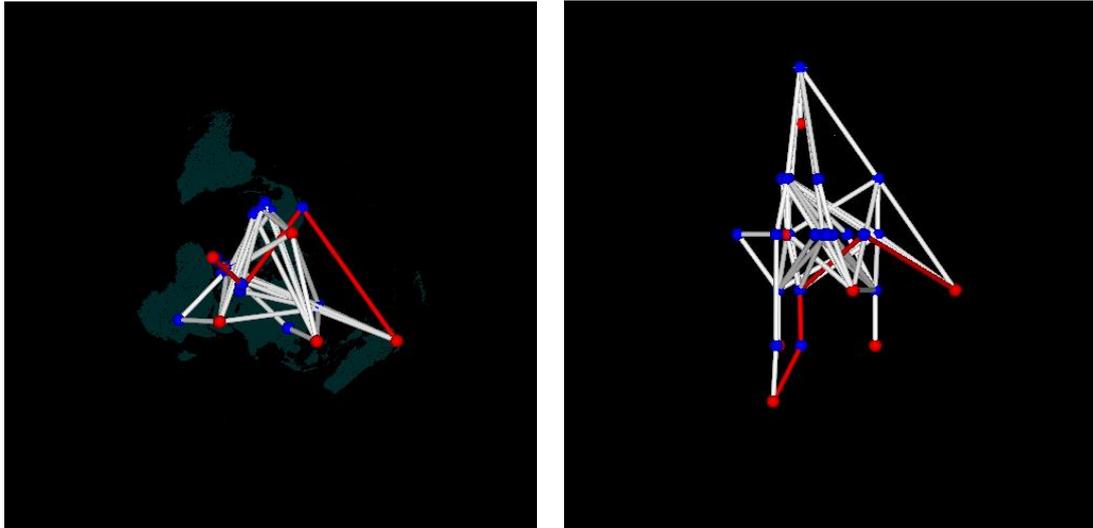
トワークが存在する物理的な地理情報を把握することが可能になり，必要に応じて着目しない方の情報を落とすことにより，必要な情報の視認性を高められる。

4.2 Zoom and Filter

Zoom and Filter を実現するために，本システムでは回転・移動・拡大縮小などの変換を対話的にできる。大規模データを対象とする場合は，対話的な操作に時間を要することがあるため，リンクの表示を線で表現し，陰影を付けない，地図情報を一時的に非表示にするなどの工夫をするとよい。また，ズームに併せてノードの情報を表示したり非表示にしたりすることもできる。Filter のために，以下の機能を備えている。

選択リンクのハイライト表示：

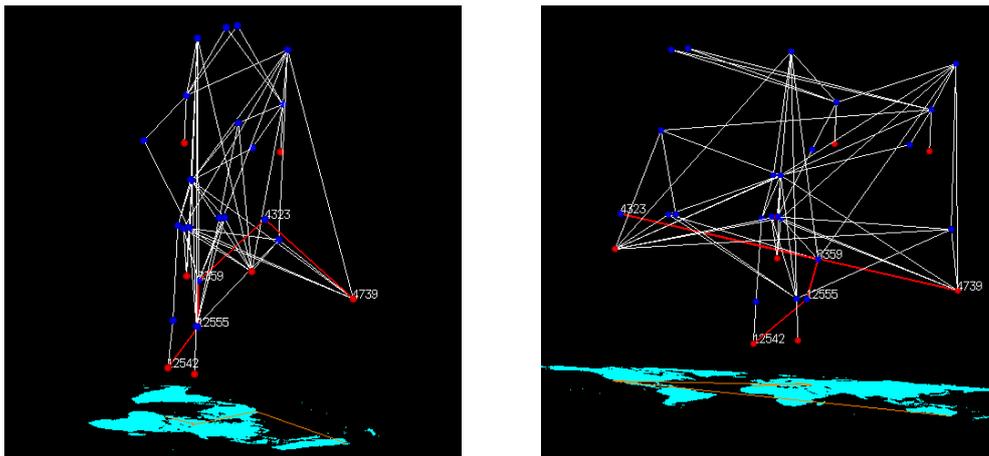
選択した経路を赤でハイライト表示することで，注目する経路情報に焦点を当てた観察を可能とする（図 6）。また，地図上の位置情報とリンクの関係を観察したい場合には図 6(a)のように改装情報を破棄した 2 次元投影図を観察し，地理情報よりも階層とリンクの関係を観察したいときは，図 6(b)に示すように地図を非表示にして階層の把握に注目できるよう 3 次元表現を利用する。その他にも，選択したノードに接続するリンクだけをハイライト表示したり，ある値より次数の高いノードだけを表示したりする等，条件を与えて Filtering を実施することもできる。さらに，あるリンクにだけ特に注目して地図上の位置情報とリンクの関係を観察することもできる。例えば階層情報も含めたリンクの関係を把握するための 3 次元表示では，全リンクを表示しつつ選択したリンクをハイライト表示することで，全体のネットワーク内でのある特定のリンクの状態を把握できる。同時に表示する地図上のリンク可視化では，選択したリンクだけを表示することで地理情報の認識を高められる。このように，投影情報ごとに表示する情報を取捨選択する Filtering も効果的である（図 7）。



(a) 上から観察

(b) 横から観察

図6 ある経路のハイライト表示



(a) 極座標

(b) メルカトル図法

図7 階層と地理情報の関係の可視化

4.3 Details_on_Demand

Zoom and Filter に併せて AS 番号を表示させたり (図 7), 注目する階層周辺の状態を詳細に把握できるようにするために, 階層軸の階層間隔を変化させたり (図 8) するなどの, Details_on_Demand 機能を実装している. これらの機能は観察者の要望に応じて実施される.

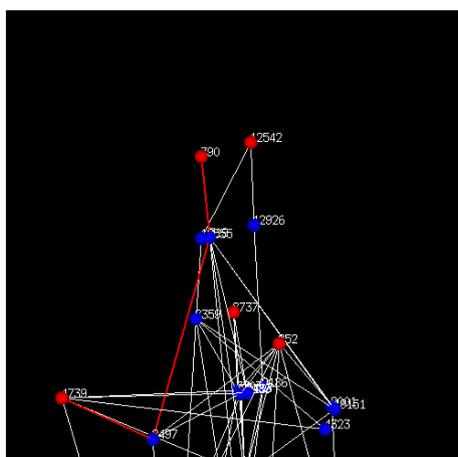
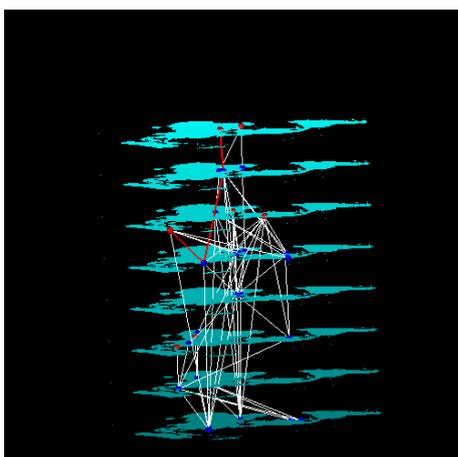
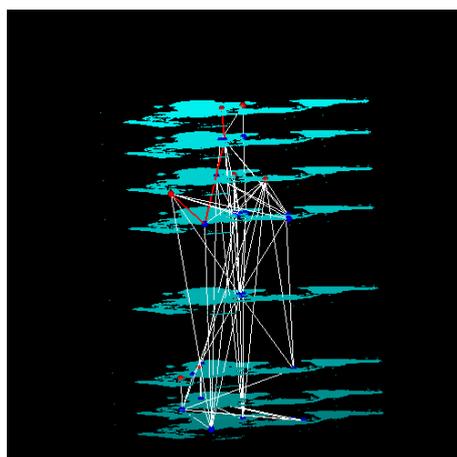


図 7 AS 番号の表示



(a) 全階層等間隔配置



(b) 5 階層目に注目した階層間隔の変更

図 8 階層間隔の不均等化

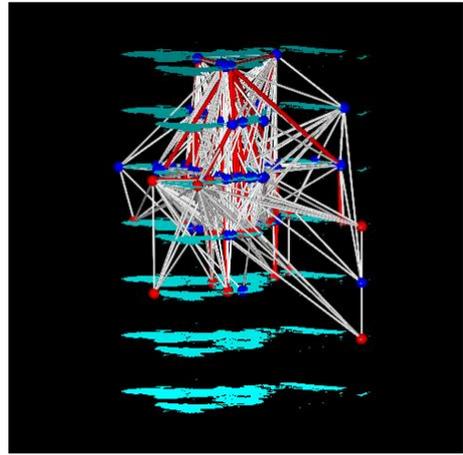
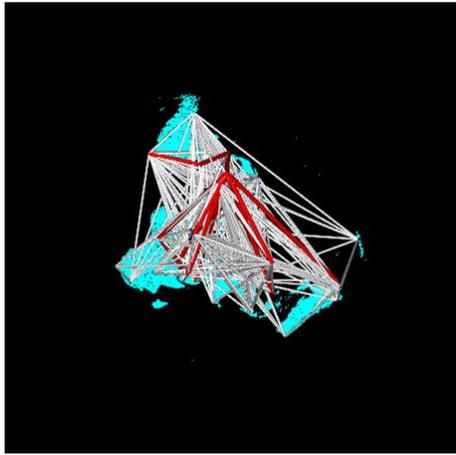
5 適用実験

P2P ネットワークにピア間の経路上における AS 間のクライアント・カスタマ関係のログデータを可視化した結果を図 7 示す。ネットワークグラフとして、Swarm を構成するピア間の IP ネットワークが可視化されている。ここでは、Tier1 の 3 つのノードに注目した。(a)アメリカ大陸にある UUNET, (b)ヨーロッパにあるドイツテレコム, (c)日本の NTT コミュニケーションズそれぞれに着目し、それらを経由する経路をハイライト表示した。すべての結果から、ネットワークグラフ

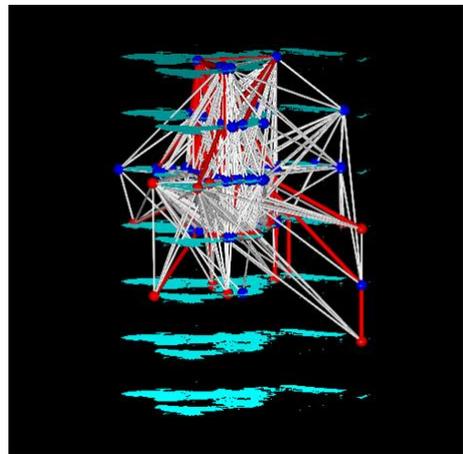
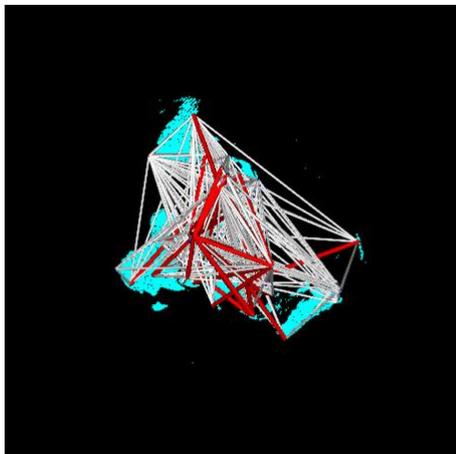
は6層として可視化し、世界中のピアから Swarm が構成してり、偏りなくさまざまな大陸からの経路に利用されていることがわかる。また、横からグラフを観察すると、縦に延びるリンクが多くあることから、地理的に近いノード同士が IP ネットワークで接続していることがわかる。つまり、IP ネットワークの構成では、大陸間のルーティングなどの地理的に離れたノード間の経路は、上位の層で行われていることがわかる。

しかし、物理的な位置がそれぞれ異なる UUNET、ドイツテレコム、NTT コミュニケーションのどの Tier 1 ノードでも、世界中のピア間の通信経路となっている。BitTorrent では、コンテンツをピースに分割し、ピア間にピースを分散配布した後に、ピア間で所持していないピースをギブアンドテイク (tit-for-tat 戦略) で交換しながら、コンテンツを完成させて入手する通信方式である。このため、ピースは、世界中のピアに分散して配布され、世界中のピアとピースを交換していることになる。さらに、それぞれの Tier 1 のノードがそれぞれ独立に世界中のピア間の通信経路となっていることから、同一のコンテンツを入手しようとしているピアが同一の大陸の中で、ピースを交換し合い、ネットワークの利用効率を考慮するような仕組みが既存の tit-for-tat 戦略では弱いことがわかる。

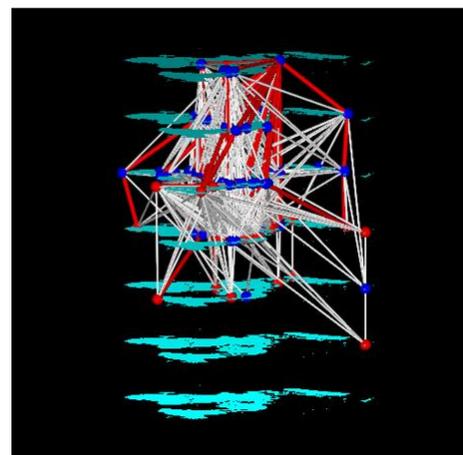
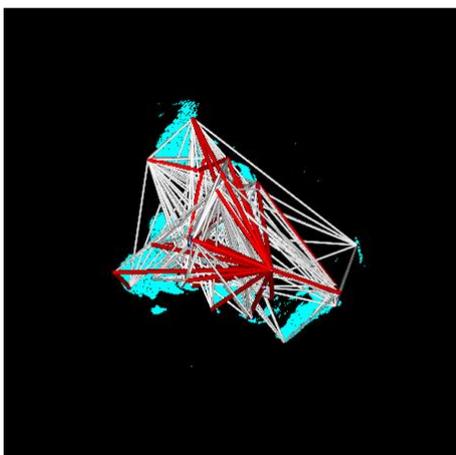
リンクを多く持つノードは、上位の層に見られる。上位の層のノードは、物理的に近い下位層のノードからの経路を集約し、物理的に離れた同等の役割を持つノードに対して IP ルーティングするためにリンクが多くなっている。より下位のノードは物理的に遠いノードに対してルーティングをする必要がなく、ノードが持つリンクは、ルートの冗長性を持たせるものだけであり、その数は少なくなっている。



(a) UUNET



(b) ドイツテレコム



(c) NTT コミュニケーションズ

図9 各地域の Tier1 ノードを経由する経路のハイライト表示，上からと横からの観察

6 まとめ

本研究では、ネットワークの経路が Valley-free であることに着目して最適化問題に定式化することにより、ネットワークの階層構造を推測した。さらに、推測された構造を用いて大規模ネットワークでも視認しやすくする可視化する手法を提案し、実装により有用性を確認した。今後の課題として、さらに大規模なネットワークに対する適用や、インターネット以外のネットワークに対する適用が考えられる。

【参考文献】

- [1] E. R. Gansner, Y. Hu, S. North, and C. Scheidegger, Multilevel Agglomerative Edge Bundling for Visualizing Large Graphs, In Proc. of *IEEE Pacific Visualization Symposium*, pp.187-194, 2011.
- [2] P. Kumar and K. Zhang, Visualization of Clustered Directed Acyclic Graphs with Node Interleaving, In Proc. of *ACM symposium on Applied Computing*, pp. 1800-1805, 2009.
- [3] F. V. Ham and M. Wattenberg, Centrality Based Visualization of Small World Graphs, *Computer Graphics Forum*, Vol. 27, No. 3, pp. 975-982, 2008.
- [4] 宮村(中村) 浩子, Hsiang-Yun Wu, 吉田 雅裕, 大坐畠 智, 中尾 彰宏, 高橋 成雄, AS 間接続関係を用いたオーバレイネットワークの可視化, 信学技報, NS2012-264, pp. 577-582, 2013.
- [5] 小池 英樹, 吉原 大敬, 対話型システムにおける大規模階層構造視覚化へのフラクタルの応用, 情報処理学会論文誌, Vol.35, No.12, pp.2703-2711, 1994.
- [6] M. Yoshida and A. Nakao, Measuring BitTorrent Swarms Beyond Reach, Proc. of IEEE International Conference on Peer-to-Peer Computing, pp. 220–229, 2011.
- [7] Caida: Topology Research, <https://www.caida.org/research/topology/>
<平成 28 年 6 月 4 日参照>
- [8] L. Gao. On Inferring Autonomous System Relationships in the Internet, *IEEE/ACM Transactions on Networking*, Vol. 9, No. 6, pp.733-745, 2001.
- [9] B., Shneiderman, Designing the User Interface Strategies for Eective Human-Computer Interaction, *Addison-Wesley*, 3rd edition, 1998, Chapter15.

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
視覚分析のためのP2Pネットワーク階層最適化	電子情報通信学会総合大会	平成 28 年 3 月
階層最適化を用いたネットワークグラフィックレイアウト	電子情報通信学会総合大会	平成 28 年 3 月