修士論文 2008 年度 (平成 20 年度)

地理情報を考慮したインターネットトポロジの解析

慶應義塾大学大学院 政策・メディア研究科

空閑 洋平

修士論文要旨　　2008 年度（平成 20 年度）

地理情報を考慮したインターネットトポロジの解析

論文要旨

インターネットの構造や成長を把握するために，インターネットのトポロジを解析することは有効な手段である．ネットワークの接続関係を示すトポロジは，本来地理的な位置とは独立であるが，実際には地理的あるいは物理的な制約が大きい．一方，インターネットのグローバルなトポロジ構造は，アメリカ中心から，地域の接続拠点を相互接続する分散化が進みつつあり，このような構造の変化を理解するためには，地理的な情報を含んだトポロジの把握が必要となる．

本研究は，ヨーロッパやアジアなどの地域ごとのインターネットの構造を把握することを目的に，その地域における AS レベルのトポロジを推測し，地域トポロジ間の比較を行う．ここでは，多地点からの traceroute データを基に，クラスタリングにより特定地域のトポロジを抽出し，その AS レベルのトポロジ構造を AS Core Map を用いて表現する．その結果，地域ごとの AS トポロジの次数分布がべき則に従う共通点や，その地域における Tier-1 AS の役割の違いなどの特色が明らかになった．また，このような地域ごとの比較を行うためには，現状のデータの観測点が，特にアフリカやオセアニアで不足している問題や，地域の範囲をより細かくし，国ごと，都市ごとの比較をするための課題が明らかになった．

キーワード

1．インターネットトポロジ，2．計測，3．地理情報，4．インターネットトポロジ探索，

慶應義塾大学大学院 政策・メディア研究科

空閑 洋平

Abstract of Master's Thesis

Academic Year 2008

Analysis of Regional-level Internet Topology Structures

Summary:

Analysis of AS-level topologies is an effective way to understand the Internet structure and its development. An AS topology represents logical connection relationships of networks, and thus, is independent of geographic locations of connections. However, the actual Internet topologies are constructed with geographical and physical limitations. In the meantime, the global Internet topology is shifting from a U.S. centered star topology to a more distributed topology interconnecting regional hub ASes. Therefore, regional topology views would provide new insights to understanding changes in the Internet structures.

Our goal is to understand differences and similarities a-mongst Internet structures in regional views such as from Europe and from Asia. We examine inference method for regional AS topologies, and compare the resulting regional AS topologies. Our inference technique extracts AS boundaries with traceroute data collected from multiple vantage points. Then, we illustrate the resulting regional AS topologies using AS Core Maps in order to identify the degree structure and AS locations in each regional view. We show that the distributions of AS out-degrees similarly follow power-law but Tier-1 ASes play different roles in regions. We also identify issues such as the lack of vantage points in Africa and Oceania in the current data sets.

Keywords:
1. Network topology, 2. Measurement, 3. Geographical Information,
4. Macroscopic Internet Topology Discovery,

Keio University, Graduate School of Media and Governance

Yohei Kuga

目次

第 1 章　はじめに 1

1.1　背景 1
1.2　本研究の目的 2
1.3　本研究における解析手法の概要 2
1.4　本研究により期待される成果 2
1.5　本論文の構成 3

第 2 章　現在のインターネットトポロジ 4

2.1　インターネットのトポロジ構造 4
2.2　現在の AS トポロジの規模 5
2.3　AS トポロジの地理的な分散化 6
2.4　まとめ 7

第 3 章　インターネットトポロジの解析手法 9

3.1　インターフェースレベルのトポロジ解析 11
3.2　ルータレベルのトポロジ解析 12
3.3　POP レベルのトポロジ解析 13
3.4　AS レベルのトポロジ解析 16
3.4.1　traceroute を用いた AS トポロジの解析 16
3.4.2　BGP の経路情報を用いた AS トポロジの解析 16
3.4.3　IRR を用いた AS トポロジの解析 17
3.4.4　大規模なインターネットトポロジ計測に関するプロジェクト . . . 18
3.5　まとめ 20

第 4 章　地理情報を考慮した AS トポロジ解析手法の提案 21

4.1　インターネットトポロジの解析手法と IP アドレスの地理位置情報 21
4.2　AS 間接続の地理情報の推測 23
4.3　AS 間接続推測の提案手法 24

4.4 まとめ 25

第 5 章 大陸レベルの AS トポロジの解析 26
5.1 データセット 26
5.1.1 トポロジデータ 26
5.1.2 BGP データ 27
5.1.3 ランドマーク 27
5.2 解析結果 28
5.2.1 大陸ごとのクラスタサイズ 28
5.3 大陸ごとの AS トポロジ構造の比較 29
5.3.1 補累積分布関数を用いた比較 29
5.3.2 AS Core Map による AS トポロジの比較 29
5.3.3 まとめ 32
5.4 大陸ごとの AS トポロジの検証 34
5.4.1 特定 AS に注目した解析手法の検証 34
5.4.2 排除されたトポロジ情報の検証 34
5.4.3 まとめ 36

第 6 章 結論 37
6.1 本論文のまとめ 37
6.2 今後の課題と展望 37
6.2.1 トポロジデータの追加 37
6.2.2 AS トポロジの推測手法の検討 38
6.2.3 地域間の AS トポロジの検証 38

参考文献 40

付録 A 大規模なインターネットトポロジ情報収集に関する研究活動 43
A.1 本付録の概要 43
A.2 活動概要 43
A.3 TopHat 44
A.4 トポロジ計測手法 44
A.4.1 DoubleTree アルゴリズム概要 44
A.4.2 DoubleTree によるトポロジ探索 45
A.4.3 現状と今後 47
A.5 StitchRoute 47

A.5.1 目的 47
A.5.2 背景 48
A.5.3 piece 48
A.5.4 アルゴリズム 49
A.5.5 今後の予定 49
A.6 まとめ 51

図目次

2.1 インターネットのトポロジ構造 5
2.2 Cumulative IANA AS block allocations 5
2.3 CAIDA: IPv4 Internet Topology Map (2001 年, 2008 年) 6
2.4 JPIX のトラフィック量の推移 8

3.1 ホスト cpu から www.caida.org に対する traceroute を実行した結果 . . 11
3.2 IP alias resolution を用いた同一ルータの持つインターフェースの推測 . 12
3.3 ルータのホスト名による POP レベルトポロジの解析 14
3.4 Undns: WIDE (AS 2500) のルールセット 15
3.5 traceroute を用いたインタフェース間経路の計測 17
3.6 CAIDA: IPv4 AS Core January 2008 19

4.1 想定環境:多くの AS 間接続 (ピアリングとトランジット) は特定の場所 (IX やデータセンタ) で接続される 23
4.2 AS 境界の IP アドレスを用いた同一の場所に存在するピアリングのクラスタリング 24

5.1 出次数による補累積分布関数を用いた大陸ごとの AS トポロジを比較 . . 29
5.2 AS Core Map (アジア) 30
5.3 AS Core Map (ヨーロッパ) 32
5.4 AS Core Map(アメリカ PST) 33
5.5 AS Core Map(アメリカ EST) 33
5.6 AS Core Map(オセアニア) 33
5.7 AS Core Map(アフリカ) 33

A.1 DoubleTree の扱うトポロジ計測時の問題点 45
A.2 DoubleTree 動作概要 46
A.3 DoubleTree による探索経路の断片化 49

表目次

2.1 大陸ごとの Internet eXchange 数とその増加量 7

3.1 インターネットトポロジの解析手法 9
3.2 大規模なインターネットトポロジ計測プロジェクト 20

4.1 インターネットトポロジの解析手法と IP アドレスの地理位置情報 22

5.1 Skitter と iPlane によって収集されたトポロジ情報 27
5.2 ランドマークデータの詳細 28
5.3 AS 間接続のクラスタサイズ 28
5.4 アジア大陸のハブ AS 31
5.5 ヨーロッパ大陸のハブ AS 31
5.6 IIJ と WIDE のバックボーン位置 34
5.7 大陸トポロジに含まれる AS 間接続数 35
5.8 クラスタ以外のトポロジデータ 35
5.9 クラスタ外データの手動による解析 36

第 1 章

はじめに

1.1 背景

インターネットのグローバルなトポロジ解析は，インターネットの構造と成長の観察，次世代のネットワークアーキテクチャの検討を目的として研究されている．グローバルなインターネットトポロジである AS (Autonomous System) トポロジを知るためには，AS 間の経路制御プロトコルである BGP (Border Gateway Protocol) の経路情報や多地点から計測した traceroute[1] のデータを用いることで把握できる．

一方，インターネットのグローバル化に伴い，国際トラフィックもアメリカ中心から，アメリカ以外の国同士で交換されるトラフィックの割合が増えている [2]．それに従い，インターネットのトポロジは，低遅延やコスト効率のよい伝送を実現するために，アメリカ中心から地域ごとの接続拠点の役割を担う ISP (Internet Service Provider) の台頭や IX (Internet eXchange) を相互接続する分散化が進んでいる．

ネットワークの接続関係を示すトポロジは，本来地理的な位置とは独立している．そのため，インターネットトポロジに関する既存研究の対象は，インターネット全体を対象とした AS レベルのトポロジや単一の AS を対象としたルータレベルのトポロジといった論理的な範囲のトポロジである．しかし，実際のインターネットのトポロジは，地理的あるいは物理的な制限が存在する．例えば，インターネットのバックボーンを構成する ISP の基幹ルータや，物理的に離れた地域間を接続する光ファイバの終端は，地域内の相互接続が集中する都市に配置されている．このような地域の接続拠点や国，大陸のような物理的な範囲のインターネットのトポロジ構造を把握するためには，地理情報を考慮しつつ，トポロジの解析が必要である．地理情報を考慮したインターネットトポロジを調査するには、どの AS とどの AS が直接繋がっているかに加え，どこで繋がっているかという地理情報が重要となる．

1.2　本研究の目的

本研究の目的は，アジアやヨーロッパのような地域ごとのインターネット構造の解析手法の確立を目指す．そのために，本論文では，インターネットのグローバルなトポロジである AS トポロジに関して，特に大陸や国レベルといった物理的な範囲のトポロジを推測する手法を検討し，その類似点と相違点を明らかにする．

このようなトポロジデータは，特定の地域におけるインターネット構造の実態と推移を把握可能にする．インターネットのインフラを重点的に強化すべき地域が明らかになる事で，近隣国間の国際相互接続を効率的に推進するための重要なインプットになると考えられる．

1.3　本研究における解析手法の概要

本論文では，多地点から計測された traceroute データを用いることで，大陸，国レベルの AS トポロジを解析する手法を検討した．本手法は，traceroute データを基に直接接続された AS 間の情報を抽出し，AS 間の基幹ルータの運用ルールや接続拠点の地理情報を用いて，直接接続された AS 間の接続場所を推測する．そして，物理的な範囲ごとの AS 間の接続情報をクラスタリングすることで，国，大陸レベルの AS トポロジを明らかにする．大陸は，アフリカ，アジア，ヨーロッパ，オセアニア，アメリカである．さらに，アメリカは，タイムゾーンが異なる 4 つの地域を分割することで，合計 8 地域の AS トポロジを解析し比較をおこなった．

1.4　本研究により期待される成果

本論文の提案手法を確立することにより，地理情報を考慮したインターネットトポロジの解析が可能になる．このようなトポロジデータは，ユーザが国際トラフィックのニーズに合った ISP を選択するための指標を与え，ISP が POP の配置，ピアリング，トランジット購入を戦略的に進めるための指標を与えると同時に ISP 間のグローバル化競争を促進する．また，本提案手法により地域ごとのトポロジデータと既存のインターネットの計測技術を組み合わせることが可能になるので，物理的な場所を識別子に用いた地域間での回線品質計測への応用が期待できる．

1.5　本論文の構成

本論文は，全 8 章で構成される．第 2 章では，インターネットトポロジの現状とグローバルトポロジの地理的な分散化の様子を述べる．第 3 章では，本研究に関連があるインターネットトポロジの解析手法と，トポロジを解析している既存研究についてまとめる．第 4 章では，地理情報を考慮したインターネットトポロジの解析手法を提案する．第 5 章では，第 4 章で提案した解析手法で用いる大陸ごとのトポロジを解析し，その解析結果と考察を述べる．第 6 章で本論文の結論と今後の課題をまとめる．

第 2 章

現在のインターネットトポロジ

本章では，グローバルなインターネットトポロジである AS トポロジの現状と AS 間の相互接続拠点である IX の現状についてまとめ，AS トポロジの地理的な分散化の様子を示す．

2.1 インターネットのトポロジ構造

現在のインターネットのトポロジ構造は，大きく AS 間と AS 内の 2 つに分類できる．AS とは，商用の ISP (Internet Services Provider) や大学，企業，研究機関といった組織から成る，単一の経路制御ポリシを共有するネットワークである．それぞれの経路制御は，AS 間の経路制御に EGP (Exterior Gateway Protocol)，AS 内の経路制御に IGP (Interior Gateway Protocol) が利用されている．現在のインターネットでは，EGP である BGP (Border Gateway Protocol)[3] が用いられ，IGP には，一般的に OSPF (Open Shortest Path First) や IS-IS (Intermediate System to Intermediate System) が利用されている．

BGP による AS 間の経路情報は，AS 境界に位置する BGP ルータ間で交換される．AS 間の接続には，2 つの接続関係があり [4]，互いのネットワーク間のトラフィックを交換するピアリングと，一方の AS が，もう一方に自らのトラフィックを転送し，インターネット全体へ配送を依頼するトランジットである．インターネット全体の経路をフルルートと呼ぶ．AS は，原則としてインターネット上のフルルートを持つ別 AS からトランジットを購入することで，インターネットとの通信が可能になる．従って，グローバルなインターネットは，AS 間のピアリングやトランジット，相互におこなわれるトランジットで構成される．

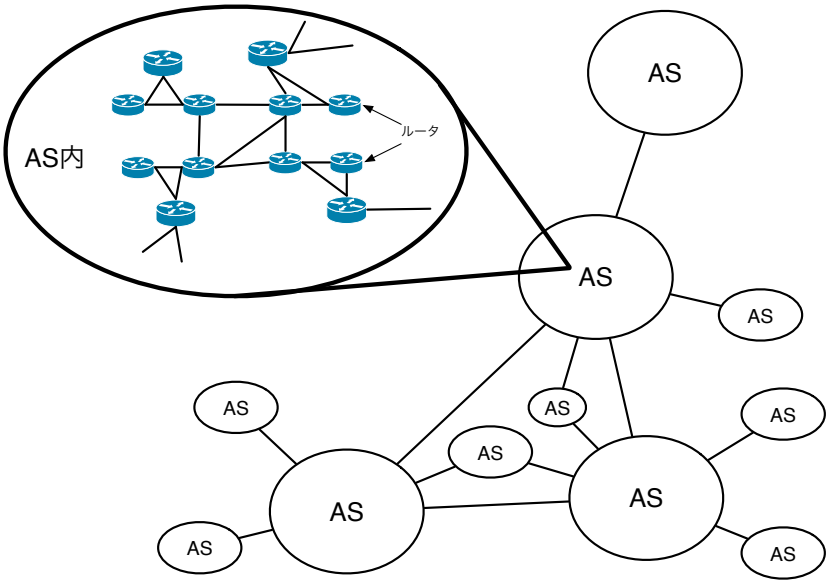

図 2.1　インターネットのトポロジ構造

2.2　現在の AS トポロジの規模

本節では，現在のインターネットのグローバルなトポロジを構成する AS の現状について述べる．

図 2.2 は，1992 年から 2008 年までに IANA (Internet Assigned Number Authority) によって割り当てられた AS 番号数の推移を表している．現在までに，インターネット上の AS 数は増加し続けており，その数は，5 万 AS を超えている．

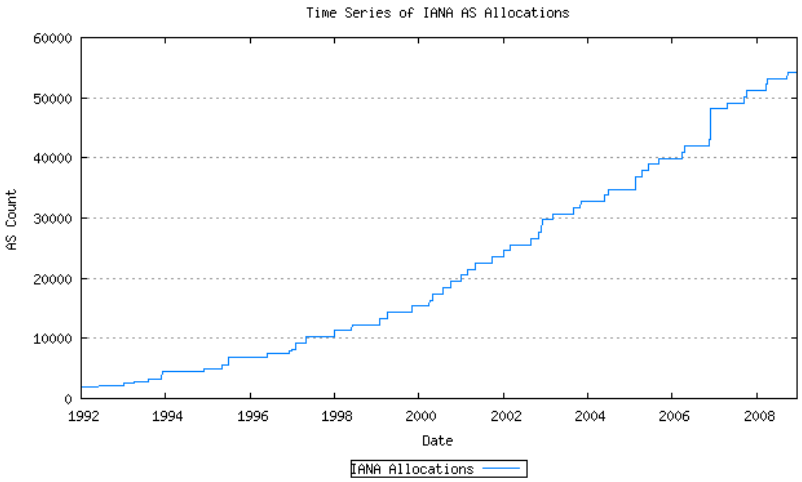

図 2.2　Cumulative IANA AS block allocations

また，図 2.3 は，CAIDA の Skitter プロジェクトが公開している 2001 年と 2008 年のインターネットトポロジである．Skitter は，インターネット上に計測基盤を構築し，トポロジデータの収集と解析をおこなっている．図は，グラフ内のノードが AS を表し，リンクが AS 間接続を表している．AS の配置は，グラフ中心を基点として，角度を AS 組織の住所から変換した経度，半径距離を AS 間接続数の出次数を多い AS を中心近くに表現している．そこで，基点近くの AS がグローバルインターネットにおける多くの接続を持つコア AS であることになる．2001 年のインターネットはアメリカに組織を持つ AS がインターネットの中心メンバであったが，現在のインターネットは，アジアやヨーロッパに拠点を持つ AS 間接続数の割合が増加し，存在感が高まっていることがわかる．また，図中の暖色で描かれるリンクは，接続数が多いコア AS による接続である．暖色の AS が増加していることから，AS 間接続数の増加も確認できる．

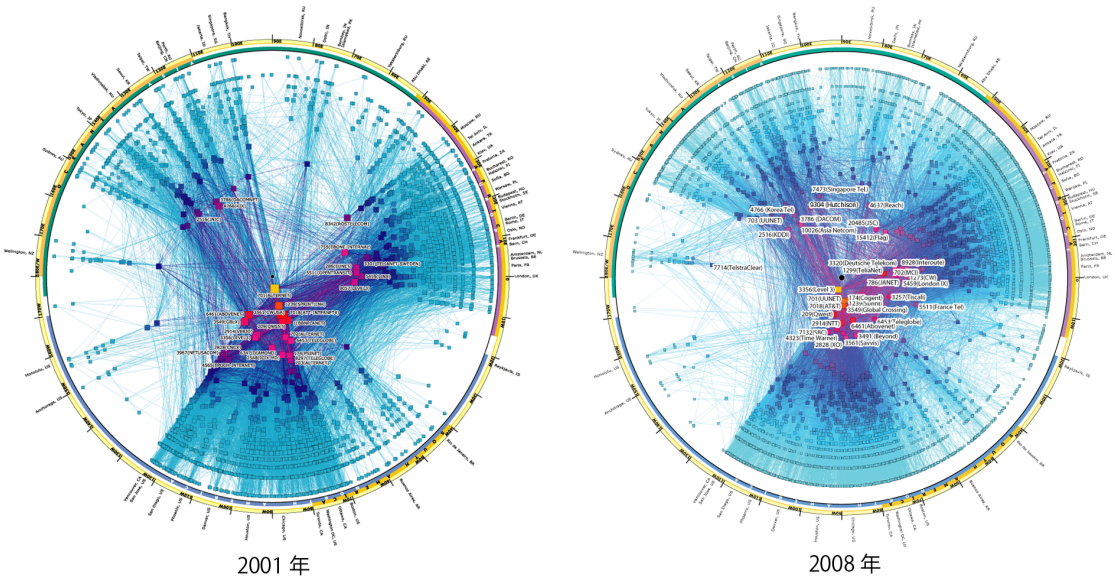

図 2.3　CAIDA: IPv4 Internet Topology Map (2001 年, 2008 年)

現在のインターネットでは，AS 数と AS 間の接続数共に増加を続けている．また，インターネットの全体のネットワークは，今後もその規模を拡大し続けることが考えられる．

2.3　AS トポロジの地理的な分散化

本節では，インターネットトポロジの分散化を調査する目的で，地域ごとの IX 数と IX のトラフィック量を調査する．

インターネットのグローバル化に伴い，国際トラフィックはアメリカ中心から，アメリ

カ以外の国同士で交換されるトラフィックの比重が増えている．それに従い，インターネットのトポロジは，低遅延やコスト効率のよい伝送を実現するために，アメリカ中心から地域ごとの接続拠点である IX (Internet eXchange) を相互接続する分散化が進んでいる．IX とは，AS 間の接続形態の一つである．多くの AS は，通信回線の冗長化やトラフィック分散のために複数 AS とピアリングする．その際に必要となる他 AS との回線コストを抑えるために，接続回線をまとめる相互接続拠点である IX が利用される．

表 2.1　大陸ごとの Internet eXchange 数とその増加量

	Africa	Asia-Pacific	Europe	Latin America	North America	Total
2007	19	66	110	22	89	306
2008	20	73	118	24	92	327
Percent Change	+5%	+11%	+7%	+9%	+3%	+6%

表 2.1 は，Packet Clearing House プロジェクトが解析している大陸ごとの Internet eXchange 数とその増加量を示している．それぞれの大陸には，その地域内での相互接続をまとめるハブである IX が存在していることがわかる．また，過去 1 年の増加量より，AS トポロジ全体に対して，接続拠点がより地理的に分散しつつあることが推測される．特に，アジアでは，その数がより多く増加している．次に図 2.4 は，日本に存在する IX の 1 つである JPIX のトラフィック量の推移である [5]．図中の平均トラフィック量に注目することで，JPIX の設立された 2001 年に比べ，IX を経由するトラフィック量が年々増加していることがわかる．これらの図より，地域内，地域間における情報交換量が増加し続けており，それに従って，今後もさらなる回線コストや冗長化を目的に，トポロジの地域的な分散が促進していくことが考えられる．

2.4　まとめ

グローバルなインターネットトポロジは，現在もその規模を拡大し続けている．それぞれの大陸では，AS 間接続の相互接続拠点である IX が設けられ，IX のトラフィック量は，年々増加している．地理的に分散した相互接続拠点が地域内のトラフィック交換の中心として機能していることが推察される．このようなグローバルなインターネットの地理的な特徴の変化は，本節で述べた IX やグローバル AS の接続状況から推察可能である．

しかし，これらの地域ごとに閉じた変化を客観的に調査する解析は困難である．インターネットで用いられるノードの識別子である IP アドレスや AS 番号は，それらの識別子と機器の位置情報と独立してることから IP アドレスの位置情報や IP アドレス間の接

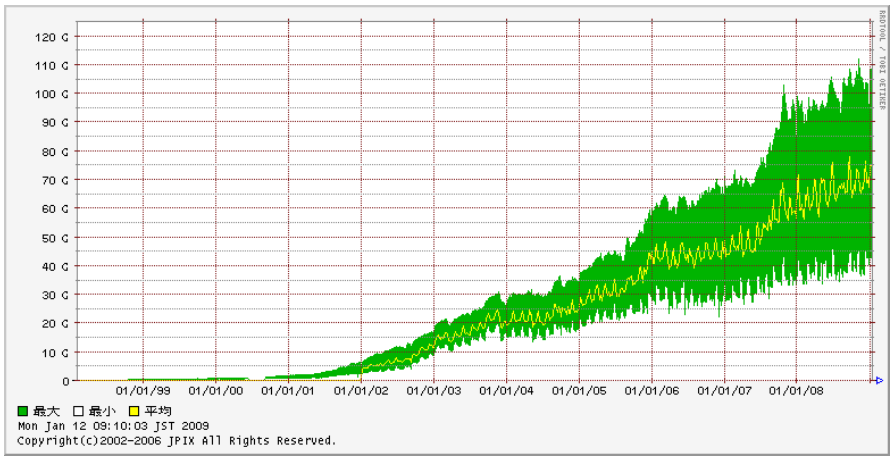

図 2.4　JPIX のトラフィック量の推移

続の把握が困難である．また，多くの ISP では，自 AS の具体的なネットワーク情報や経路制御ポリシを公開していない．外部 AS 内の情報を推測する手法と，IP アドレスの地理位置情報の推測手法を組み合わせ，解析する必要がある．

第 3 章

インターネットトポロジの解析手法

本章では，本論文で提案する地理情報を考慮した AS トポロジの解析手法を検討するために，既存のインターネットトポロジの解析手法をまとめる．

インターネットトポロジの解析手法は，解析対象と範囲の違いで分類できる．図 3.1 は，既存のインターネットトポロジの解析手法をまとめたものである．

表 3.1　インターネットトポロジの解析手法

Level	Method	Related works
Interface	Traceroute	Scamper Paris-traceroute
Router	Traceroute + IP alias resolution	Mercator, Ally, APAR, DisCarte
POP	Traceroute + Reverce DNS	Rocketfuel
AS	Traceroute + BGP, BGP data, IRR	CAIDA Skitter, iPlane, NetDIMES, UCLA Cyclops

インタフェースレベルのトポロジは，traceroute ツールを用いることで，対象ネットワークのトポロジを探索可能である．traceroute は，外部の AS から対象 AS の内部ネットワークを手軽に調査可能なため，インターネット全体の IP アドレスに対して探索を実行することで，インターネット全体のトポロジを推測できる．しかし，多くの ISP では，バックボーンのリンクを冗長化しており，さらに，運用ポリシによっては traceroute の計測パケットに応答しない可能性がある．

次に，ルータレベルのトポロジは，traceroute と IP alias resolution と呼ばれる，同一ルータの持つ複数インタフェースを推測する技術を利用することで，解析可能である．

多くのバックボーンルータは，経路制御のため，複数のインターフェースを持ち，さらに 1 つのインターフェースで複数の IP アドレスを持つ場合が多い．IP alias resolution は，バックボーンルータのインターフェースに対して計測パケットを送信することで，インターフェースが同一のルータのものか調査し，収集した traceroute データと組み合わせてルータレベルのトポロジを明らかにする．しかし，それぞれの IP alias resolution 技術は，すべてのルータのすべてのインタフェースを明らかにすることはできず，特定環境下のルータのみ明らかにできる．通常は，複数の IP alias resolution 技術を組み合わせてルータレベルのトポロジを推測する．

次に，POP レベルのトポロジは，ルータのホスト名からそのルータが存在する地理情報を推測することで，単一の AS 内のトポロジに関して POP ごとのルータトポロジが推測できる．しばしば，ISP は自らが管理するルータのホスト名にそのルータが存在する場所を推測可能にする目的で，ホスト名にカントリーコードや空港コード，省略した都市名といった名前をつける．traceroute によって集めたトポロジ情報と，DNS 逆引きによって集めたホスト名を組み合わせることで，同一のデータセンターにある単一 AS のネットワークを推測する．しかし，AS によっては，バックボーンルータのホスト名に地理情報を埋め込んでいない場合がある．

最後に，グローバルなインターネットトポロジである AS レベルのトポロジ解析手法を述べる．AS トポロジの解析手法は，利用するデータセットの違いで 3 種類ある．

1 つは，インターフェースレベルのトポロジ推測技術と同じく traceroute データを用いる手法である．AS トポロジの解析には，加えてフルルートを持つ BGP ルータの経路表を用いる．BGP の経路表から経路エントリーのプレフィックスとそのプレフィックスを持つ AS 番号を対応づけることができる．その対応表を用いることで，traceroute データのすべてのホップ IP アドレスの所属する AS 番号を明らかにする．traceroute のパスと組み合わせることで，AS 間のつながりであるホップ間 IP アドレスのペアが抽出し，AS 間のつながりを推測できる．

2 つ目に，BGP の経路表を用いた手法である．BGP は，パスベクター型の経路制御プロトコルであり，管理者の経路制御ポリシの反映と経路ループの回避のため，目的に AS 番号までに通過するホップの AS 番号の情報を保持している．この連続した AS 番号を AS パスと呼ぶ．フルルートの経路が設定されたバックボーンルータの AS パスを解析することで，インターネットのグローバルな AS 間のつながりを推測できる．

3 つ目は，IRR の AS の登録情報を用いた手法である．IRR (Internet Routing Registry) は，AS の管理者が自らの経路情報を手動で登録するためのデータベースである．traceroute のみでは発見することの難しい冗長経路が発見できる場合がある．しかし，AS によっては運用ポリシーのために，一部の経路情報を登録しない場合があり，さらには，手動による登録のため，古い情報が登録されたままとなっている可能性がある．

3.1　インターフェースレベルのトポロジ解析

traceroute は，自らの計測機器からあて先ホストまでのパケットが通過する経路を調査するツールである．traceroute は，ICMP (Internet Control Message Protocol) パケットに含まれる IP ヘッダの TTL (Time To Live) を 1 から漸増させながら計測パケットを送信 する．送信された計測パケットにより，あて先ホストに到達する途中で経由するルータが 順番に分かる．また, 送信する計測パケットに対する確認応答パケットが返信されるまでの時間である RTT (Round Trip Time) が計測できる．

図 3.1 にホスト cpu.sfc.wide.ad.jp から www.caida.org に対して計測した traceroute の出力結果を示す．出力結果は，順に TTL の値，ホスト名，IP アドレス，3 つの試行した ICMP パケットの RTT を示す．また，RTT は，小数点以下を切り捨てた値を表す．出力の 3 行目は，ホスト cpu が，0 ミリ秒以下の RTT の値で vlan4.rg-gate.sfc.wide.ad.jp に TTL の値 1 で ICMP パケットが到達したことを表す．traceroute を用いることで，ホスト cpu からあて先ホスト www.caida.org までの経路がわかる．

```
cpu % ¿ traceroute www.caida.org
traceroute to cider.caida.org (192.172.226.123)，64 hops max，40 byte packets
1 vlan4.rg-gate.sfc.wide.ad.jp(203.178.142.129) 0 ms 0 ms 0 ms
2 ve-100.foundry1.fujisawa.wide.ad.jp(203.178.137.91) 0 ms 0 ms 0 ms
3 ve-4.cisco2.notemachi.wide.ad.jp(203.178.138.243) 1 ms 1 ms 1 ms
4 apan-jp.t-lex.net(203.178.133.141) 1 ms 1 ms 1 ms
5 losa-tokyo-tp2.transpac2.net(192.203.116.145) 116 ms 116 ms 131 ms
6 cenichpr-1-lo-jmb-702.lsanca.pacificwave.net(207.231.240.129) 117 ms 117 ms
118 ms
7 riv-hpr–lax-hpr-10ge.cenic.net(137.164.25.5) 123 ms 123 ms 123 ms
8 hpr-sdsc-sdsc2–riv-hpr-ge.cenic.net(137.164.27.54) 125 ms 125 ms 125 ms
9 pinot.sdsc.edu(198.17.46.56) 126 ms 126 ms 125 ms
10 cider.caida.org(192.172.226.123) 126 ms 126 ms 126 ms
```

図 3.1　ホスト cpu から www.caida.org に対する traceroute を実行した結果

Paris-traceroute[6] は，フランスに拠点を持つ CNRS の Brice Augustin らが研究開発

している新しい traceroute ツールである．既存の traceroute 手法では，途中経路に存在する冗長経路構成を検出できない．そこで，既存の traceroute と同じく TTL を漸増させながら計測パケットを送信し，同時に返答 ICMP パケットヘッダの *Sequence* フィールドを用いてパケットのフローを識別し，複数経路を検出する．後述する大規模にインターネットトポロジを解析する Skitter[7] や iPlane[8] では，すでに Paris-traceroute を用いてトポロジデータを収集している．

3.2　ルータレベルのトポロジ解析

ルータレベルのトポロジは，traceroute と IP alias resolution と呼ばれる同一ルータの持つ複数インタフェースを推測する技術を組み合わせて用いることで推測できる [9]．多くのバックボーンルータは，経路制御のため，複数のインターフェースを持ち，さらに 1 つのインターフェースで複数の IP アドレスを持つ．IP alias resolution は，バックボーンルータのインターフェースに対して計測パケットを送信することで，複数の IP アドレスが同一のルータのものか調査し，収集した traceroute データと組み合わせてルータレベルのトポロジを明らかにする．

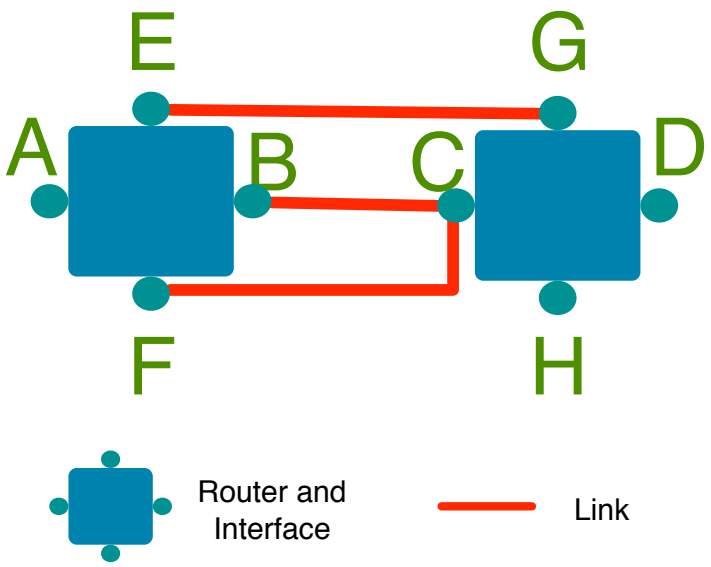

図 3.2　IP alias resolution を用いた同一ルータの持つインターフェースの推測

IP alias resolution 手法の 1 つである Mercator[10] は，対象機器に UDP を用いて，使用されていない短命ポート番号に対して計測パケットを送信する．返答されたパケットの送信元 IP アドレスが送信パケットのあて先 IP アドレスと異なっていた場合，その IP アドレスを Alias アドレスと推測する．

Ally[11] は，2 つの対象 IP アドレスに対して 2 セットの計測パケットを送信し，対象 IP アドレスが同一のルータに所属する IP アドレスか判断する．判断は，返答 IP パケットヘッダの *Identification* フィールドに注目する．通常，ルータは，*Identification* フィールドの ID 番号を，フローごとに 1 ずつ漸増させながら生成する．2 つの IP アドレスが，同一のルータに所属していた場合，連続して計測パケットを送信することで，*Identification* フィールドの ID 番号が似通った値を示すはずである．さらに返答パケットの TTL の値から返答パケットのホップ数を比較する．*Identification* と TTL フィールドの値が閾値以下の場合，2 つの IP アドレスを同一のルータに所属すると推測する．

APAR[12] は，計測パケットを用いず，複数計測した traceroute データを解析することのみで alias アドレスを識別する．

DisCarte[13] は，IP ヘッダのオプションフィールドである *RecordRoute* オプションを用いて Alias アドレスを推測する．traceroute による収集される IP アドレスは，機器の *Incoming* のインターフェースが持つ IP アドレスである．*RecordRoute* オプションによって記録される IP アドレスは，対象機器に対して *Outgoing* インターフェースの IP アドレスが記録されることから，traceroute と逆向きの経路が推測できる．*Incoming* と *Outgoing* の IP アドレスから，同一のルータが持つ 2 アドレスを収集する．

それぞれの IP alias resolution 技術は，すべてのルータが持つすべての IP アドレスを明らかにすることはできず，それぞれ特定環境下のルータのみ明らかにできる．Mercator は，返答パケットの送信元 IP アドレスを選択することができないため，送信パケットの送信先 IP アドレスと同じ IP アドレスの場合が多い．Ally は，2IP アドレスに対して連続して計測パケットを送信することから，IP アドレス数 n に対して $O(n^2)$ 回数計測が必要となるため，インターネット全体を対象とした解析は現実的に不可能である．APAR は，特定トポロジの場合のみ解析可能であり，DisCarte で用いる *RecordRoute* オプションは，計測機器から 9 ホップまでの IP アドレスのみ記録可能であり，インターネット全体を対象とするためには，複数の異なる AS に複数の計測拠点が必要となる．

3.3　POP レベルのトポロジ解析

POP レベルのトポロジ解析は，ルータのホスト名からそのルータが存在する地理情報を推測し，単一の AS 内のトポロジを POP 内のルータトポロジを推測する．

多くのインターネットに接続する機器には，機器の管理者により，ホスト名が設定される．DNS (Domain Name System) は，ホスト名と IP アドレスを対応づけるシステムである．DNS をサービスしているネームサーバに対して問い合わせることで，ホスト名からその機器の IP アドレスを検索する正引きや，逆に IP アドレスからホスト名を検索する逆引きができる．ISP では，自らのネットワーク内のバックボーンルータに対して，ロ

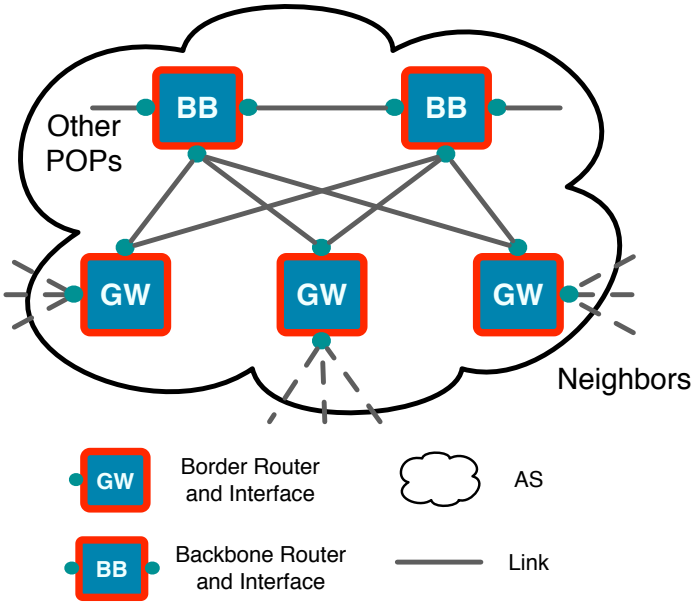

図 3.3　ルータのホスト名による POP レベルトポロジの解析

グの解析や機器の把握といった目的で，ホスト名に機器の地理情報を推測可能なように設定している場合がある．以下に実際の ISP で使用されるホスト名の例を示す．

$$foundry2.otemachi.wide.ad.jp \tag{3.1}$$
$$fa-2-3-4.a10.tokyjp01.jp.ra.gin.ntt.net \tag{3.2}$$

(3.1) のホストは，AS2500 の WIDE Project 内のネットワーク機器である．(3.2) は，AS2914 NTT Communication が管理する機器である．(3.1) のホスト名は，otemachi と jp と いう文字列がホスト名に含まれる．jp は，日本のカントリーコードである．また，otemachi は，日本の東京都千代田区大手町と推測できる．そこで，(3.1) のネットワーク機器は，日本の東京都千代田区大手町に物理的に配置されていると推測できる．同様に，(3.2) の例では，tokyjp という文字列により，このネットワーク機器が日本の東京都に配置されていると推測可能である．ネットワーク機器のホスト名は，それぞれの AS ごとのポリシによって，独自の命名規則 で名前付けが行なわれる．(3.1)，(3.2) では，両ホストとも日本の東京都に存在すると推測可能 であるが，異なる命名規則を持つ．AS ごとのポリシによってホスト名の命名規則が異なる ことから，機械処理によってホスト名から地理情報の推測することは困難である．

Undns は，Rocketfuel[11] で用いられる scriptroute[14] のモジュールである．それぞれの AS ごとにホスト名のルールセットを持ち，ルールに従ってホスト名と機器の地理的な場所を対応づけている．例えば，(3.1) のネットワーク機器は，そのホスト名から

WIDE プロジェクトで管理，運用されている機器とわかる．Undns 内部の WIDE プロジェクトのルールセット は，図 3.3 の通りである．(3.1) の例では，文字列 otemachi が.wide.ad.jp の前にある．Undns を使用することで，(3.1) のホスト名を持つネットワーク機器が，Tokyo，Japan に存在すると推測ができる．

```
2500 \.wide\.ad\.jp$ { \\
  \.([a-z]*)\.wide\.ad\.jp$ loc=1 { \\
    dojima "Osaka, Japan" \\
    fujisawa "Kanagawa, Japan" \\
    nezu "Tokyo, Japan" \\
    notemachi "Tokyo, Japan" /* NTT-otemachi */ \\
    otemachi "Tokyo, Japan" /* otemachi is a building */ \\
    hiroshmia "Hiroshima, Japan" \\
    hachioji "Hachioji, Japan" \\
    tokyo "Tokyo, Japan" \\
  }; \\
}\\
```

図 3.4　Undns: WIDE (AS 2500) のルールセット

Rocketfuel では，複数の計測機器から traceroute を実行し，undns による POP の識別を組み合わせて，特定 AS の内部のルータトポロジを明らかにしている．

1. traceroute をトポロジを明らかにしたい AS を横断するホスト間で実行する．
2. 対象 AS 内の計測ホストから traceroute を実行する．
3. 対象の ISP 内部のノードに対して traceroute を実行する．これらの分類た traceroute データを解析することで，AS 内のネットワークを推測できることを示している．Rocketfuel の解析手法を用いることで，POP のサブネットのサイズやエッジルータ，バックボーンルータとそのリンク数を明らかにできる．

Washington 大学の Scriptroute プロジェクト [14] は，ルータのホスト名から地理情報を推測する undns ツールを公開している [15]．Undns は，271 の AS に対するホスト名の命名規則を記述したルールセットを持ち．Undns を用いることで，ホスト名からルータの地理情報を推測可能となる． しかし，Undns で地理情報を推測できない場合がある．1 つは，小規模な ISP やプライベートピアリングなどの一部のプレフィックスでは，現在

でも機器にホスト名を設定していない．また，Undns は，すべての AS に対するルールセットを持っていない．さらに，ISP によっては，自らのルータにプライベート IP アドレスを用いている．プライベート IP アドレスは，外部の AS からホスト名を知ることが難しい．また，ホスト名と地理情報が必ずしも一致しない場合がある [16].

3.4　AS レベルのトポロジ解析

本節では，既存の AS トポロジ計測手法を 3 つに分類して示す．1 つは，traceroute データを用いた AS トポロジ計測手法である．地理的に分散した複数地点から計測した traceroute データと BGP の経路表を用いることで，AS を横断する AS 間の接続を抽出する．次は，BGP データを用いた AS トポロジの解析手法である．BGP の経路表から AS パス属性を抽出し，AS 間接続状況を解析する．最後に，AS 自らが登録した IRR (Internet Routing Registry) データベースに登録した経路情報を元に AS トポロジを解析する手法である．

3.4.1　traceroute を用いた AS トポロジの解析

それぞれの AS は，独自の経路制御ポリシで運用される．通常，自らの AS 内にあるホストから別 AS 内のトポロジを知ることができない．そこで，既存の AS トポロジに関する研究 [17] では，複数の計測機器から traceroute を用いてトポロジ計測することで，特定の AS 内のインターフェースレベルのトポロジデータを収集する．さらに，インターネット全体に対する経路を持つ BGP ルータの経路表から，それぞれの経路のプレフィックスとそのプレフィックスを管理する Origin AS の対応表を作成する．作成した対応表と収集した traceroute データ組み合わせることで，traceroute データのホップ IP アドレスを管理する AS 番号を明らかにする．最後に，AS 番号を横断する IP アドレス間のデータを抽出することで，AS 間接続が推測できる．

本手法は，インターネット全体の AS を横断するように traceroute データを収集することで AS トポロジを推測する．そのため，計測機器から実行した traceroute の経路上の AS のみトポロジを推測できる．より正確な広域に渡るトポロジを解析するためには，世界各拠点の AS に多くの計測拠点が必要となる．

3.4.2　BGP の経路情報を用いた AS トポロジの解析

BGP の経路情報を用いた AS トポロジの解析手法を示す．

AS 間の経路は，BGP ルータの経路情報を用いることでわかる．BGP の各経路エントリは，あて先 AS までに経由する AS 番号を表した AS パスの情報を含む．そこで, AS レ

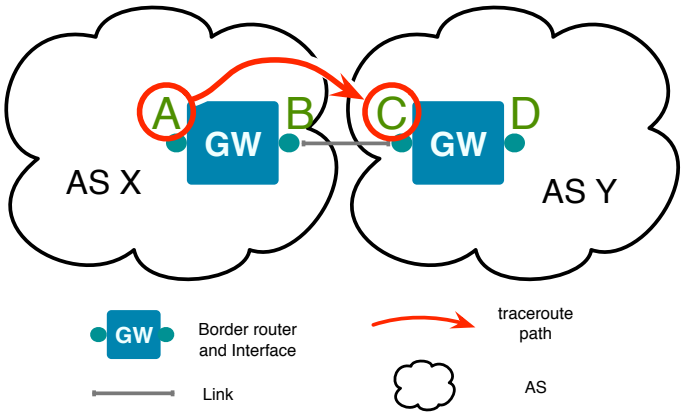

図 3.5　traceroute を用いたインタフェース間経路の計測

ベルのインターネットのトポロジは，フルルートを持つような，多くの AS と接続性が ある BGP ルータの経路情報を用いることで解析できる． Oregon 大学の Route Views[18] では，複数の AS と接続した経路を収集する BGP ルータを運用し，経路情報を公開している．

しかし，2 つの理由から BGP の経路情報を用いた AS レベルのトポロジ計測手法は，広域に渡るインターネットのトポロジ解析が難しいと考える．1 つ目の理由は，ピアリングによる AS 間の接続を知ることが困難な点である．AS 間の接続には，一定の回線品質の確保や接続性の冗長構成のために IX を経由したピアリングとは別に，プライベートなピアリングが存在する．自らの BGP の経路情報のみでは，他 AS 間の冗長経路を知ることが困難となる．もう 1 つは，必ずしも BGP の経路情報では，あて先ネットワークまでのトラフィックがわからない点である．複数の地域 拠点で同一のあて先と AS ピアリングする場合，トラフィックをできるだけ早くあて先 AS まで運ぶ手法が用いられる．*Deflectionrouting* と呼ばれる手法を用いることで，あて先 AS までの行きのトラフィックはピアリング先の AS を多く移動し，帰りのトラフィック が自らの AS 内を多く移動する．それにより，トラフィックの流れが行き帰りで非対称な経路を通る．単一の BGP ルータの経路情報のみでは，トラフィックがわからない．

3.4.3　IRR を用いた AS トポロジの解析

IRR とは，インターネット上のインタードメインの経路制御プロトコルである BGP の経路情報と経路の優先性に関する情報を蓄積するデータベースである．IRR に登録されている AS 間の接続情報は，traceroute や BGP データによる解析では発見が難しいプライベートな AS 間接続を発見できる可能性がある．しかし，IRR は，AS 管理者によって手動で登録されているため，経路情報が古い場合がある．また，ISP による特定企業に

対するプライベートピアリングといった一部の経路情報は，登録されていない可能性がある．

3.4.4　大規模なインターネットトポロジ計測に関するプロジェクト

インターネットトポロジは，インターネットの構造把握や次世代のインターネットアーキテクチャ検討の基礎データとして，多くの研究がされてきた [19, 20, 21]．グローバルなインターネットトポロジである AS トポロジの解析には，traceroute，BGP の経路情報，IRR を用いた手法がある [22]．

CAIDA の Skitter プロジェクト [7] では，traceroute を基にした Scamper ツール [23] を用いて，インターネット全体のトポロジデータを収集，AS トポロジを解析し，可視化を行っている．CAIDA は，California 大学 San Diego Super Computer Center を拠点としたプロジェクトである．グローバルなインターネットを対象にネットワークの特性を解析し，解析ツールを開発している．traceroute と BGP の経路情報を組み合わせることで，ホップごとの IP アドレスが所属する AS を明らかになり，AS 間の接続がわかる．多地点から複数のノードに対して Scamper を実行することで，インターネット全体を対象に，その AS トポロジを推測している．

Skitter では，2009 年 1 月現在，アメリカを中心に世界 31 の拠点間で計測システムを構築し，トポロジデータを収集している．収集したデータは，ネット ワークの接続性の可視化やトポロジ変化の計測に利用される．

AS Core Map は，CAIDA の Skitter プロジェクトによる AS トポロジの可視化手法である．図 3.6 は，CAIDA の公開している 2008 年度の AS Core Map である．Skitter プロジェクトで計測したパス情報を基に，それぞれの AS を拠点の地理情報と出次数を用いて可視化している．出次数とは，AS 間の接続における出て行くリンクの数である．グラフでは，それぞれのノードが AS を表し，グラフの中心を基点にし，AS の拠点の経度を角度，出次数を半径の位置情報で表される．式 3.3 は，グラフ内の AS 位置の算出方法を示している．グラフ中心から，角度を組織の経度，中心からの半径距離を AS 間接続数で表現している．経度は，WHOIS に登録されている AS の拠点の住所から求める．AS Core Map では，それぞれの AS が描かれる角度に変換される．半径は，もっとも出次数が多い AS を中心近くに配置し，その AS の出次数を基準にそれぞれの AS の配置を決定する．出次数が多い AS を円の中心に配置することで，インターネットのコアになっている AS が把握可能になる．

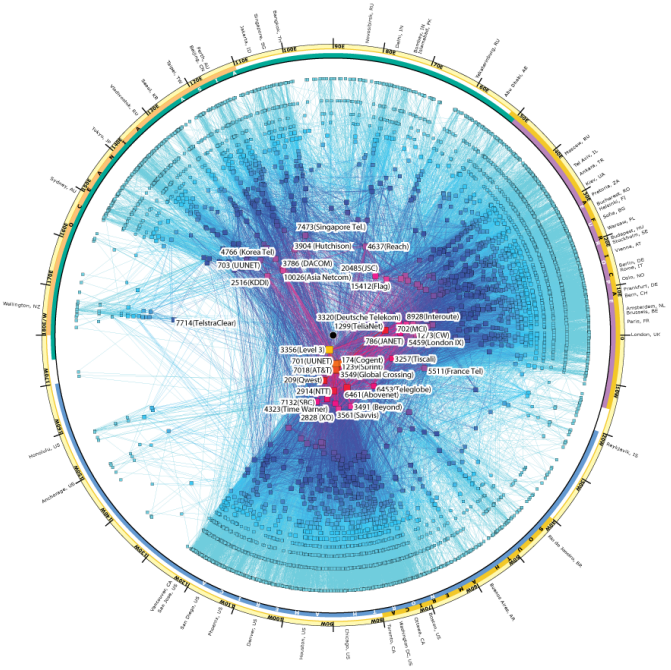

図 3.6　CAIDA: IPv4 AS Core January 2008

$$
\begin{aligned}
radius &= 1 - log\left(\frac{outdegree\,(AS)+1}{max.outdegree+1}\right) \\
angle &= \left(\begin{array}{c}\text{longitude of the AS's}\\ \text{BGP prefixes}\end{array}\right)
\end{aligned}
\tag{3.3}
$$

表 3.2 は，大規模にインターネットトポロジ計測基盤を構築しているプロジェクトと，その計測手法を示している．iPlane は，ワシントン大学に拠点を置くグローバルインターネットを対象とした，トポロジ計測プロジェクトである．計測拠点には，PlanetLab[24] を利用している．計測手法には，Paris-traceroute 手法と UDP パケット，通常の ICMP パケットによる手法を組み合わせて用いている．また，イスラエルのテルアビブ大学に拠点を持つ DIMES プロジェクト [25] では，PC のスクリーンセイバーとして動作する traceroute プログラムを作成し，エンドユーザからのインターネットパス情報を広く収集している．計測手法としては，ICMP パケットによる通常の traceroute を用いている．フランスのパリ第六大学に拠点を持つ OneLab プロジェクト [26] のインターネットトポロジ収集コンポーネントである TopHat プロジェクト [27] では，計測拠点に PlanetLab ノードを利用し，独自の DoubleTree アルゴリズム [28] を用いて計測基盤を構築している．計測手法は，DoubleTree と通常の traceroute を用いており，今後 Paris-traceroute

による計測をおこなう予定である．

表 3.2　大規模なインターネットトポロジ計測プロジェクト

Project	Traceroute Method
CAIDA Skitter	ICMP, ICMP-Paris, UDP
DIMES	ICMP
iPlane	ICMP, ICMP-Paris, UDP
Tophat	ICMP, ICMP-Paris

3.5　まとめ

本章では，インターネットのトポロジ解析についてまとめた．既存のインターネットトポロジ解析は，インターネット全体を対象とした AS の集合や，特定 AS 内のトポロジを解析する．そのため，解析範囲は論理的なネットワークで区切られている．また，IP アドレスや AS 番号，それらの間の接続と地理位置情報は独立しており，識別子のみでは，トポロジの地理情報と直接結びつけることが困難である．しかし，それぞれの解析手法は，IP アドレスの相対的な位置情報が推測できる可能性がある．例えば，ルータレベルのトポロジ解析に用いる $IPaliasresolution$ 技術は，単体ルータに所属する IP アドレスをクラスタリングする．そのため，同一ルータに所属する IP アドレスは，同一の位置に存在すると推測できる．また，POP レベルのトポロジで用いた，$ReverseDNS$ によるルータ位置の推測技術を用いることで，IP アドレスの地理情報が推測可能であり，POP 内のトポロジでは，IP アドレス間の接続の位置も推測可能である．

次章では，本章でまとめた既存のインターネットトポロジ技術から，IP アドレスの地理情報が推測できる手法を考察し，地理情報を考慮した AS トポロジの解析手法を検討する．

第 4 章

地理情報を考慮した AS トポロジ解析手法の提案

本章では，地理情報を考慮した AS トポロジ解析手法を提案する．第 3 章では，インターネットトポロジの解析手法を述べた．本章では，それらのトポロジ解析手法から地理情報推測の可能性を検討し，さらに AS 間接続の地理情報を推測する新しい手法を提案する．

本手法では，多地点から計測した traceroute データを解析することで，AS 間の接続を抽出し，提案手法を用いて地理的に距離の近い AS 間の接続をクラスタリングする．さらにバックボーンルータのホスト名と IX の地理情報を用いてクラスタの地理情報を推測することで，クラスタと地理情報を対応づけ，特定地域の AS 間接続を推測する．

以後の節では，それぞれの手法について背順を説明する．

4.1 インターネットトポロジの解析手法と IP アドレスの地理位置情報

第 3 章では，インターネットのトポロジ解析についてまとめた．IP アドレスや AS 番号やそれらの間の接続と地理位置情報は，独立しており，識別子のみで地理情報とトポロジを直接結びつけることが困難である．しかし，それぞれの解析手法は，IP アドレスの相対的な位置情報が推測できる可能性がある．図 4.1 は，それぞれのインターネットトポロジの解析手法を地理情報の推測可能性をまとめたものである．

traceroute を用いたインターフェースレベルのトポロジ計測のみでは，IP アドレスの位置情報を推測することは困難である．しかし，traceroute で収集した RTT (Round Trip Time) 値によって IP アドレス間の相対的な距離が推測可能である．インターネット上の機器に対して，パケットの到達性確認に用いられるツールである ping や traceroute

表 4.1　インターネットトポロジの解析手法と IP アドレスの地理位置情報

範囲	IP アドレスの位置	IP アドレス間のリンク位置	手法	利用する情報
Interface	-	相対的距離	Traceroute	RTT
Router	-	相対的距離	Traceroute + IP alias resolution	RTT+ Reverse TTL, IP Record Route Option，IP ID
POP	地理位置	地理位置	Traceroute + Reverce DNS	ルータのホスト名
AS	-	相対的距離 †	Traceroute + BGP, BGP data IRR	AS 間接続の IP アドレス

† 本論文で提案

を用いることで，自らのホストからあて先ホストまでの RTT が計測できる．RTT は，あて先ホストに対して ICMP パケットを送信し，確認応答パケットである応答が戻ってくるまでの時間を表す．あて先ホストまでの RTT を計測することで，機器の物理的距離を知る指標となる．図 3.1 では，TTL 値=4 から 5 にかけて，RTT の値が 100 ミリ秒以上増加している．TTL 値=6 以降のホストも RTT の値が 100 ミリ秒以上かかっていることから，TTL 値=4 から 5 にかけてホスト間の物理的距離が離れていると推測できる．さらに TTL 値=1 から 4 までは RTT 値が 1 ミリ秒以下であることから，これらのホップの IP アドレスが地理的に近い可能性が高い．しかし，RTT は，途中ホップの計測機器のパケットまでの速度インターフェースの待ち行列にかかるキューイング遅延によって値が変動しやすい．特に多くのクロストラフィックを処理する AS のバックボーンルータでは，計測のたびにキューイング遅延のために値が大きく変動することが考えられる．

ルータレベルのトポロジ解析に用いる $IPaliasresolution$ 技術は，単体ルータに所属する IP アドレスをクラスタリングする．そのため，同一ルータに所属する IP アドレスは，同一の位置に存在すると推測できる．またインターフェース間の RTT と組み合わせて解析することで，特定環境におけるルータ間の相対的な距離が識別できる可能性がある．

また，POP レベルのトポロジで用いた，$ReverseDNS$ によるルータ位置の推測技術を用いることで，IP アドレスの地理位置が推測可能であり，POP，データセンタ内のトポロジでは，IP アドレス間の接続の位置も推測可能である．

AS レベルのトポロジでは，AS 番号と地理情報が対応づけられていないため，AS の地理情報を識別できない．しかし，多くの AS 間の接続は，ISP のデータセンターや IX といった特定の場所で直接接続されると推測できる．

4.2　AS 間接続の地理情報の推測

AS 間接続の地理情報の推測手法について検討する．図 4.1 は，本提案手法の想定している AS 間接続の形態を示している．第 2 章で述べた通り，AS 間の接続関係には，ピアリングとトランジットがある．通常，AS がトランジットを購入する場合，*Provider* が *Customer* に対してアクセスリンクの接続性とトランジット用の IP アドレスを提供する．*Customer* は，その提供された IP アドレスを自組織のデータセンタに存在するルータに設定する．もしくは，一部のトランジットやピアリングは，*Provider* のデータセンタにルータを設置し，経路制御することが想定される．そのため，AS 間接続で用いられる IP アドレスのペアは，物理的に近い場所に存在している可能性が高い．

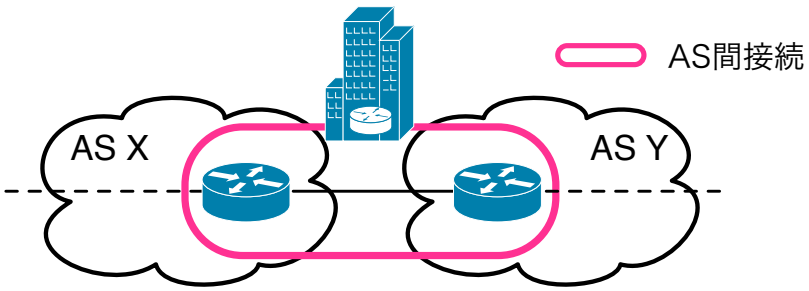

図 4.1　想定環境:多くの AS 間接続 (ピアリングとトランジット) は特定の場所 (IX やデータセンタ) で接続される

しかし，一部の AS 間は，地理的に離れたデータセンタまでレイヤ 2 ネットワークを延ばした専用線や MPLS (Multi-Protocol Label Switching) 上で接続される．分散 IX では，MPLS を用いて複数の拠点間を接続して運用されているため，地理的に離れた IP アドレスペア間で AS 接続されることが考えられる．そのため，地理的に離れた位置でされる AS 間接続を抽出し排除することで，AS 間の接続で用いられる IP アドレスのペアは，相対的な距離が近いと推測可能である．

図 4.2 は，本論文で提案する AS 間接続の地理情報の推測手法である．本手法は，多くの AS 間の接続が IX やデータセンタなどの同一の場所で接続されることから推測する．

本節で述べた通り，多くの AS 間の接続で用いられる IP アドレスのペアは同一の場所に存在すると推測できることから，IP アドレスを用いて，異なる AS 間接続をクラスタリングしている．そうすることで，地理的に同じ場所に存在する AS 間接続ごとのクラスタが作成できる．

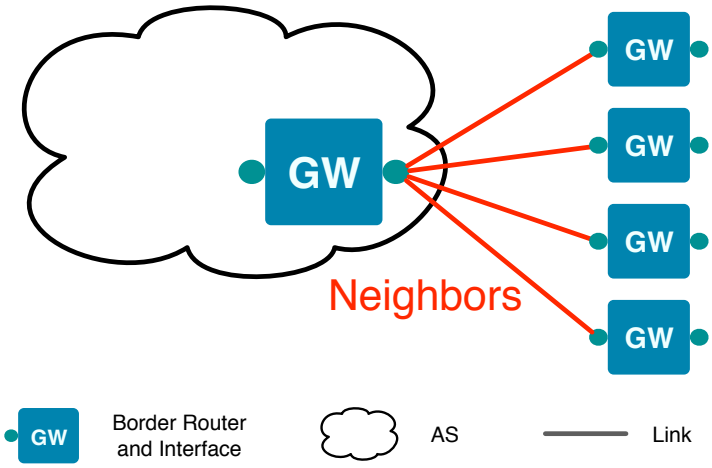

図 4.2　AS 境界の IP アドレスを用いた同一の場所に存在するピアリングのクラスタリング

4.3　AS 間接続推測の提案手法

前節までに，既存のインターネットトポロジの解析手法は，それぞれ特定環境上ではあるが，IP アドレス間の地理情報，もしくは相対的な距離が推測可能なことを示した．本研究は，これらのインターネットトポロジの解析手法を組み合わせることで，地理情報を考慮した AS トポロジの解析手法を提案する．AS トポロジの解析手法には，用いるデータの違いで 3 種類手法がある．本提案手法は，その 1 つである traceroute データと BGP の経路表を組み合わせて AS 間接続を推測手法を用いる．traceroute データを用いることで，Interface レベル，ルータレベル，POP レベルの地理情報が推測可能なトポロジ解析手法を組み合わせることが可能であるためである．

以下に本提案手法を述べる．

1. traceroute データから AS 間の接続である IP アドレスのペアを抽出し，片方の IP アドレスが共通する AS 間の接続なら同じ場所にあると推測し，クラスタリングする．
2. *IPaliasresolution* を用いて同一ルータの持つ IP アドレスを推測し，同一ルータによる AS 間の接続で (1) で作成したクラスタを，さらにクラスタリングする．
3. ルータのホスト名を *ReverseDNS* で調査し，同一の地理情報を持つ AS 間の接続をクラスタリングする．
4. IX のプレフィックスと IX が存在する地理情報の対応表を作成し，IX のプレフィックスにマッチする IP アドレスの AS 間接続の地理情報を推測し，物理的に近い地理情報を持つクラスタをクラスタリングする．

5. 広域に渡るレイヤ 2 ネットワークや MPLS 上で接続される AS 間接続により，本来異なる場所にあるクラスタが結合される可能性がある．$ReverseDNS$ を用いて，異なる地理情報を持つ AS 間接続のクラスタ結合を排除する．
6. (5) までに作成したクラスタは，それぞれ同一の場所に存在する AS 間接続である．$ReverseDNS$ と IX 情報からクラスタ内の IP アドレスの地理情報を調査し，クラスタ全体の位置情報とする．

まず，本手法は，前節で述べた本研究で提案する AS 間接続の地理情報の推測手法を用いることで地理的に近い AS 間接続をクラスタリングする．本手法のみでは，IP アドレスのみを用いていることから，同一 IP アドレスが含まれないクラスタ間を結合できない．そこで，IP アドレスではクラスタリングできない AS 間接続を $IPaliasresolution$ と $ReverseDNS$ による地理情報でさらにクラスタリングしていく．

また，地理的に離れた AS 間接続データは，$ReverseDNS$ によって推測した地理情報によってクラスタリングの結合を排除する．最後に，地理的に近くに存在する AS 間接続であるそれぞれのクラスタの位置情報を，$ReverseDNS$ と IX 情報から調査し，クラスタの位置情報を推測する．多くの地域には，それぞれの地域の接続拠点である IX が存在する．Packet Clearing House[29] や PeeringDB[30] などでは，それぞれの IX の持つプレフィックスと IX の存在する地理情報を収集している．これらのプロジェクトが公開している IX の情報を元に IX の地理情報とプレフィックスの対応表を作成する．それにより，IX のプレフィックスに一致する AS 境界の IP アドレスによる AS 間接続の地理情報が推測可能である．

4.4　まとめ

本章では，第 3 章でまとめた既存のインターネットトポロジ解析手法を，地理情報の推測手法としての視点から整理し，本論文で提案する AS 間接続の地理情報の推測手法と組み合わせることで，地理情報を考慮した AS トポロジの推測手法を提案した．インターネット上のノードの識別子である IP アドレスや AS 番号のみでは，機器の存在する地理情報を知ることができない．本提案手法は，IP アドレス間の相対的な距離の近さによる IP アドレスのクラスタリングによって AS 間接続をクラスタリング可能なことを述べた．本提案手法を用いることで，今まで解析が難しかった物理的な範囲内の AS トポロジが解析可能になる．

第 5 章

大陸レベルの AS トポロジの解析

本章では，第 4 章で提案した地理情報を考慮した AS トポロジの解析手法を用いて，アメリカ，アフリカ，ヨーロッパ，アジア，オセアニアの大陸ごとにおける AS トポロジの構造を解析する．AS トポロジ構造の比較には，AS 間の接続数を用いた補累積分布関数と Skitter プロジェクトの AS Core Map[31] を用いる．

5.1 データセット

本節では，大陸ごとの AS トポロジを解析するために用いた，それぞれのトポロジ情報を述べる．

5.1.1 トポロジデータ

第 3.4 章で述べた通り，グローバルな AS トポロジの解析には，多地点で計測された traceroute 情報が利用可能である．

本研究では，CAIDA の Skitter プロジェクトと Washington 大学の iPlane が収集しているトポロジ情報を用いて解析した．両プロジェクトでは，地理的に分散して存在している複数の計測ノードから ICMP，Paris-traceroute を用いてトポロジ情報を収集している．CAIDA では，プロジェクトに参加している組織と連携し，複数地点間で skitter を運用している．iPlane は，次世代インターネットのテストベット環境構築を目的とした PlanetLab 上の計測環境を構築し，トポロジ情報を収集している．

Skitter では，2009 年 1 月現在，世界 21 カ国，31 の計測機器から広範囲にインターネット上のトポロジ情報を収集している．計測機器は，アメリカに 10 拠点，カナダに 2 拠点，オーストラリア，オーストリア，ブラジル，スイス，チリ，中国，ドイツ，スペイン，フィンランド，ギリシャ，アイルランド，日本，韓国，モロッコ，オランダ，ニュー

ジーランド，フィリピン，台湾，イギリスにそれぞれ 1 拠点ずつ存在する．また，iPlane は PlanetLab 上の計測ノードを用いている．PlanetLab は，2009 年 1 月現在，世界に 943 拠点のノードを所有している．

表 5.1　Skitter と iPlane によって収集されたトポロジ情報

	CAIDA Skitter	iPlane
計測期間	2008/10/18 - 2008/10/20	2008/10/18
計測拠点数	31	943
データ形式	scamper フォーマット (warts)	独自フォーマット
特徴	エンタープライズ	アカデミック
traceroute エントリ数	16,567,960	27,092,756

表 5.1 は，Skitter と iPlane で計測されたデータの詳細を表す．本論文の解析には，2008 年 10 月 18 日から 2008 年 10 月 20 日の期間で計測されたパス情報トポロジ情報を用いた．Skitter の計測拠点数は iPlane に対して少ないが，計測拠点が商用 ISP 多く存在する特徴がある．一方，iPlane の計測拠点の多くは，世界の学術機関の AS に存在する．Skitter のデータフォーマットは，scamper フォーマット (warts) で保存され，iPlane は独自のバイナリフォーマットで保存されている．また，Skitter によって計測された traceroute 情報は計 16,567,960 エントリ，iPlane による traceroute 情報は，計 27,092,756 エントリであった．

5.1.2　BGP データ

AS のプレフィックスと AS 番号の対応表には，Routeviews[18] が公開している BGP の経路表を用いた．BGP の経路表から経路ごとのプレフィックスと対応するオリジンの AS 番号を抽出し，対応表を作成する．Routeviews が公開している BGP の経路表には，AS のプレフィックス数が 280,741 エントリあった．

5.1.3　ランドマーク

本提案手法では，共通する IP アドレスを持たない AS 間の接続をクラスタリングするために，ルータのホスト名と IX の地理情報を用いる．これらの IP アドレスと地理情報を対応づける機器をランドマークと呼ぶ．IX の情報には，Packet Clearing House が公開している exchange-points のデータを用いた．

表 5.2 は，解析したランドマーク数を示す．IX の情報は，228 拠点分のプレフィックスと地理情報を抽出した．Undns で推測できたルータの IP アドレス数は，143,470 エント

リあった．これらの抽出した IP アドレスは，それぞれ物理的に存在する大陸名と対応づけた．

表 5.2　ランドマークデータの詳細

IX 拠点数	228
Undns によって地理情報と対応づけた IP アドレス数	143,470

5.2　解析結果

本節は，第 5.1.1 節で述べた traceroute 情報を解析し，本提案手法により大陸ごとの AS トポロジを解析した結果を示す．

5.2.1　大陸ごとのクラスタサイズ

表 5.3 は，本手法で解析して得られた大陸ごとのクラスタを表す．アメリカのみ大陸ではなくタイムゾーンごとの EST，MST，CST，PST の地域に分割して解析した．*#ofASborderlinks* は，クラスタに含まれる AS 間接続の数であり，*UniqASN* は，クラスタ内に出現する AS 数である．また，*Other* は，クラスタリングできなかったデータであり，*All* は，すべての大陸を合わせたインターネット全体を表している．

表 5.3　AS 間接続のクラスタサイズ

Area	# of AS border links	Uniq ASN
Europe	62,899	6,511
US-EST	24,620	3,341
US-PST	18,865	2,035
US-CST	18,143	2,405
Asia	10,608	1,441
US-MST	9,271	1,005
Oceania	2,597	374
Africa	135	43
Other	70,326	12,660
All	225,792	20,906

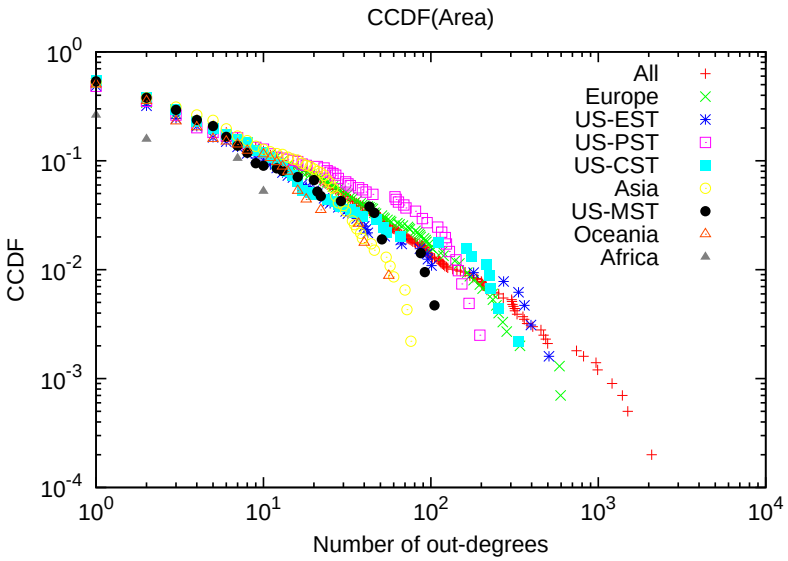

図 5.1　出次数による補累積分布関数を用いた大陸ごとの AS トポロジを比較

5.3　大陸ごとの AS トポロジ構造の比較

本章では，解析した大陸ごとの AS トポロジを検証し，考察を述べる．本手法で推測した AS トポロジのデータを用いて，大陸ごとに AS トポロジを比較する．

5.3.1　補累積分布関数を用いた比較

図 5.1 は，大陸ごとの AS トポロジのクラスタごとに，AS 間の接続の出次数を用いた補累積分布関数 (CCDF) を表す．図中の All は，全大陸を含めた AS トポロジの出次数を表す．大陸ごとの AS による出次数が，それぞれべき乗則に従う共通点が見られ，それぞれの大陸には，接続が集中するハブ AS が存在することを表している．次節では，それぞれの大陸におけるハブ AS について解析する．

5.3.2　AS Core Map による AS トポロジの比較

大陸ごとの AS トポロジを第 3.4.4 節で述べた AS Core Map を用いて可視化する．AS Core Map は，ピアリング数が多い AS を中心に配置することから，それぞれの大陸のハブ AS が把握可能になると期待できる．

図 5.2 は，本手法を用いて推測したアジア大陸の AS トポロジである．図中左下は，出次数が多い AS を 1 位から 5 位まで表している．*rank* は順位，*asn* は AS 番号，*degree*

は，出次数を表している．また，グラフ中の各経度を色相環で色を割り当てている．それぞれのリンクの色は，ピアリングを出している AS の経度に割り当てられた色である．これにより，経度の近い AS 間のピアリングは色相の近い色となり，場所の離れた位置に住所を持つ AS 間のリンク色が色相が遠く表現される．

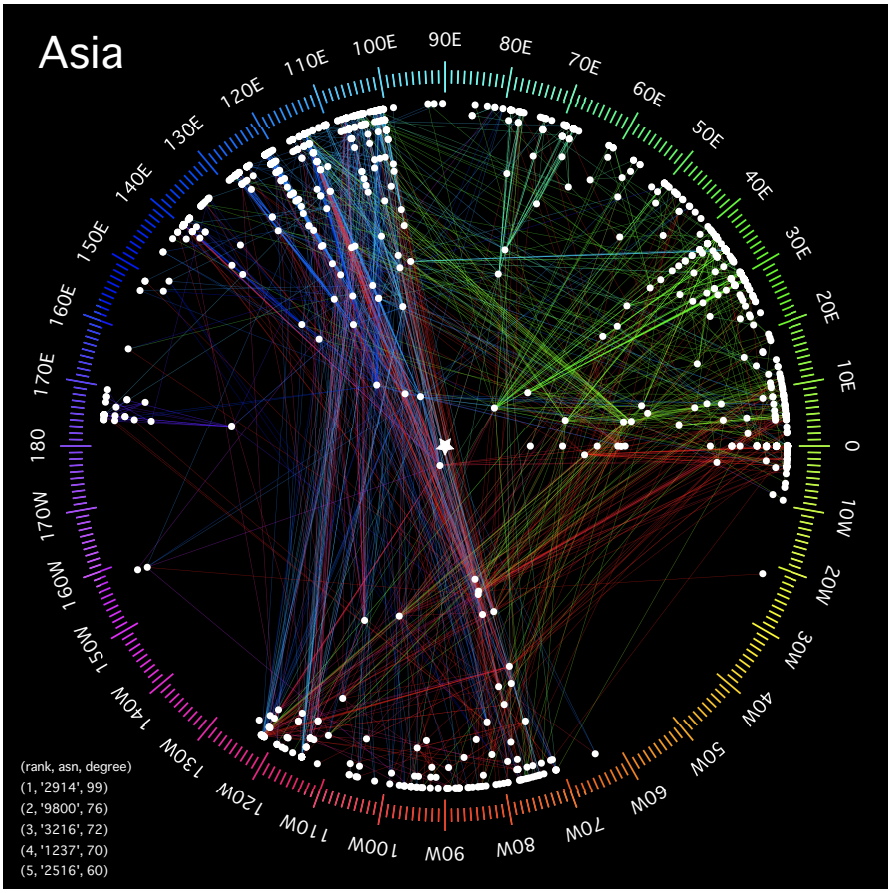

図 5.2　AS Core Map (アジア)

表 5.4 は，アジア大陸の AS トポロジにおける AS 出次数順に並べたものである．アジア大陸では，日本だけでなく，ロシアや中国，韓国の存在感が増していることがわかる．

図 5.3 は，本手法を用いて推測したヨーロッパ大陸の AS トポロジである．アジア大陸に比べ，アメリカの AS が多く中心にあることから，アメリカの AS の存在感の高さがわかる．

表 5.5 は，ヨーロッパ大陸の AS トポロジにおける AS 出次数順に並べたものである．ヨーロッパ大陸では，アジア大陸に比べてアメリカ AS の存在感が大きいことがわかる．

アジアまでネットワークを延ばすような，アメリカの AS のアジアへの進出は，ほとんど観測できなかった．また，経度 40 度のロシア (モスクワ) の AS や経度 175 度に位置するニュージーランドの AS が多い．クラスタリングがうまくできていない懸念はあるが，

表 5.4　アジア大陸のハブ AS

Rank	ASN	AS Name	Country	# of out-degrees
1	2914	NTT America, Inc.	JP	99
2	9800	CHINA UNICOM.	CN	76
3	3216	Golden Telecom	RU	72
4	1237	KISTI	KR	70
5	2516	KDDI	JP	60
6	10026	TISCALI	DE	57
7	9121	TTNet	TR	54
8	7660	APAN-JP	JP	46
9	1299	TELIANET	SE	44
10	1239	SprintLink	US	39

表 5.5　ヨーロッパ大陸のハブ AS

Rank	ASN	AS Name	Country	# of out-degrees
1	174	COGENT	US	736
2	3356	Level3	US	596
3	1299	TELIANET	SE	586
4	3549	Global Crossing	US	341
5	9002	Real Time Network	RU	284
6	3257	TISCALI	DE	269
7	3248	SIL-AT	AU	254
8	20485	TRANSTELECOM	RU	246
9	786	JANET	UK	235
10	6461	ABOVENET	US	234

ロシアの AS がアジアに力をいれている可能性がある．

ヨーロッパでは，イギリスやフランスに住所を持つ AS だけでなく，東ヨーロッパの AS の接続数も多いことがわかる．経度約 105 度のシンガポールや経度 117 度の中国 (北京) の AS の進出が観測された．

図 5.5 と図 5.4 は，それぞれアメリカの東海岸周辺と西海岸周辺の AS トポロジを表す．EST 地域では，アジアの AS の進出があまり観測できないが，PST エリアでは，多くのアジアの AS が存在する．また，EST 地域では，ヨーロッパの AS が多く進出しており，

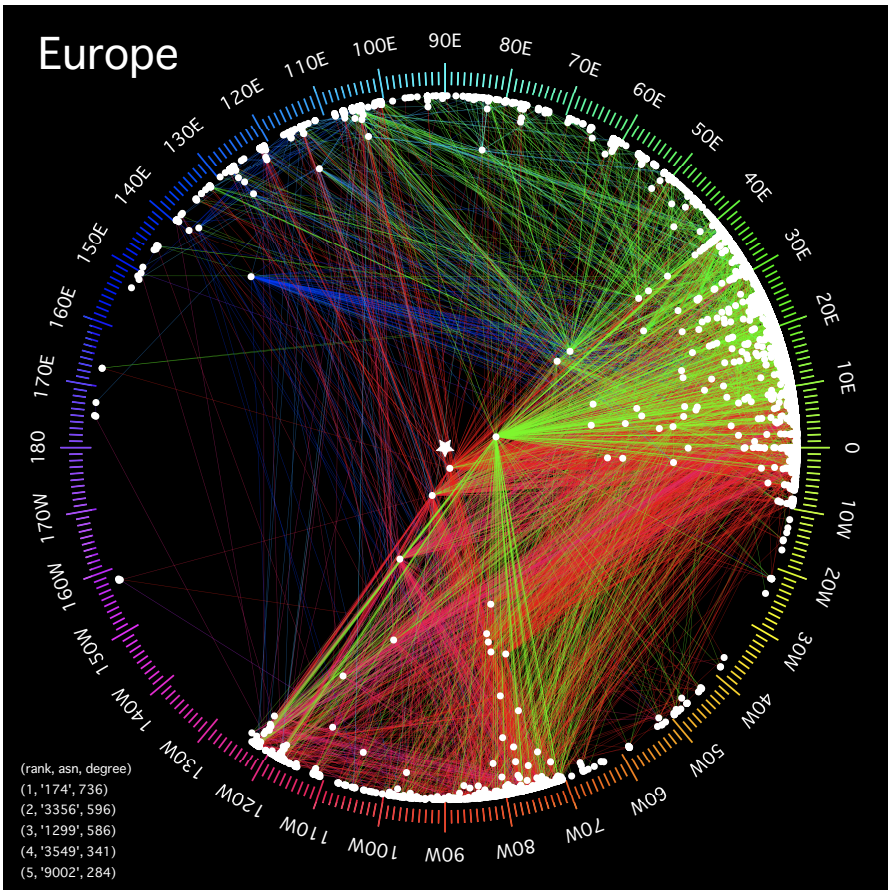

図 5.3　AS Core Map (ヨーロッパ)

PST 地域では少ないことがわかる．

図 5.6 と図 5.7 は，それぞれオセアニア大陸ととアフリカ大陸の AS トポロジを表す．オセアニア大陸のトポロジでは，オーストラリアの AS がハブであることが確認できた．AS 間の接続のほとんどがアメリカの AS 向けであることがわかる．アフリカ大陸のトポロジからは，AS 番号 5713 の SAIX (South African Internet Exchange) を発見できる．しかし，出現 AS 数が少ないことから，本提案手法のみでは，検出できなかった可能性とアフリカ大陸のトポロジデータが不足している可能性が考えられる．

5.3.3　まとめ

大陸ごとの AS トポロジによる比較では，地域に着目することで異なった特色がみえた．それぞれの地域から AS トポロジをみることで，その地域内で接続が集中する拠点がわかる．アメリカの東と西でも AS トポロジが大きく異なっていた．しかし，アフリカの AS トポロジでは，AS 間の接続数が少ない．解析の基になる traceroute データのみでは，

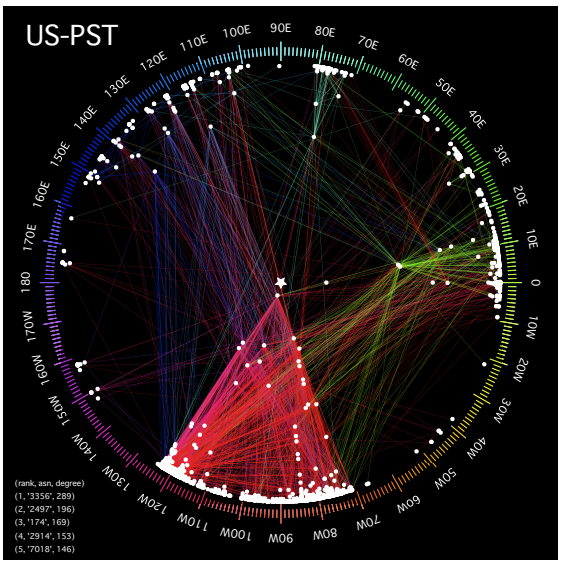

図 5.4　AS Core Map(アメリカ PST)

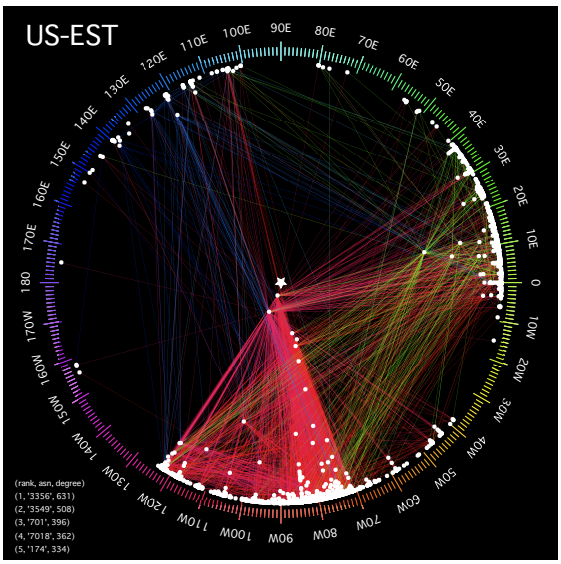

図 5.5　AS Core Map(アメリカ EST)

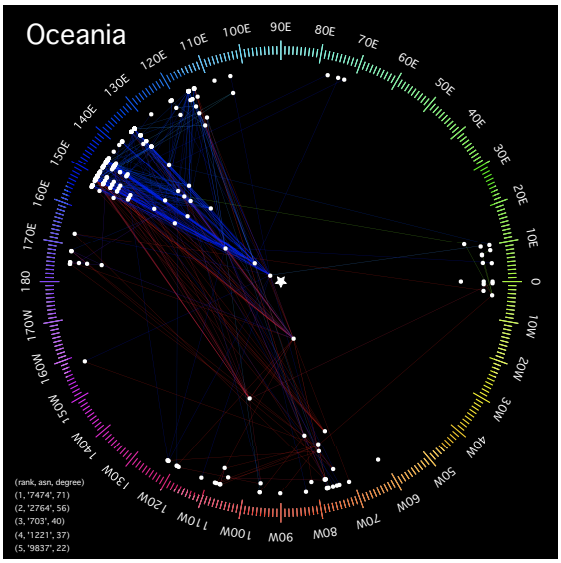

図 5.6　AS Core Map(オセアニア)

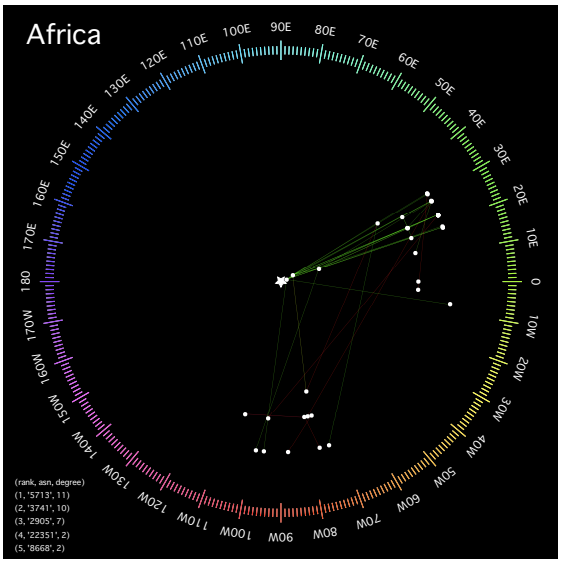

図 5.7　AS Core Map(アフリカ)

アフリカやその他の地域でデータ数が足らない可能性がある．

補累積分布関数による解析では，大陸ごとにおける AS トポロジの出次数による次数分布がべき則に従う共通点があった．また，AS Core Map では，それぞれの大陸からみた AS 間の接続を確認でき，その地域におけるハブ AS が明らかになった．

5.4　大陸ごとの AS トポロジの検証

5.4.1　特定 AS に注目した解析手法の検証

本節では，本論文で提案する地理情報を考慮した AS トポロジの解析手法を検証するために，特定の AS に着目し，その組織のネットワークと本提案手法による解析結果を比較する．AS には，AS 番号 2497 の IIJ (Internet Initiative Japan Inc.) と AS 番号 2500 の WIDE プロジェクトのネットワークを用いる．

IIJ は，日本に拠点を置く ISP であり，WIDE は，日本の学術研究プロジェクトである．表 5.6 は，IIJ と WIDE のバックボーン拠点の地理情報である．IIJ，WIDE 共にヨーロッパ，オセアニア，に接続拠点が存在しないことがわかる．

表 5.6　IIJ と WIDE のバックボーン位置

AS Name	ASN	Backbone
IIJ	2497	JP, US-PST (CA), US-EST (NY, PA, VA)
WIDE	2500	JP, TH, US-PST (CA)

次に，表 5.7 は，それぞれの大陸クラスタに含まれる IIJ (AS2497) と WIDE (AS2500) の AS 間接続数を示す．AS 間の接続関係によっては，*Provider* が *Customer* のデータセンタまでアクセスリンクを延ばし，*Provider* のプレフィックスから切り出した IP アドレスを *Customer* が自らのルータに設定することが考えられる．IIJ がそのような接続状態にあった場合，IIJ のバックボーン拠点が存在せずにヨーロッパやオセアニアに AS 間接続が存在する可能性がある．また，本提案手法では，地理的に離れた場所で接続される一部の AS 間接続が排除できず，クラスタリングされた可能性がある．今後，IIJ，WIDE の詳細なネットワークを検証した上で，さらなる解析をおこなっていく．

5.4.2　排除されたトポロジ情報の検証

本節では，本論文の提案手法の検証のために，地理情報でクラスタリングできなかった AS 間接続情報を解析し，考察を述べる．

表 5.8 は，大陸ごとの AS 間接続をクラスタリングする際，大陸ごとでクラスタされなかったデータを表している．本来，クラスタ外のデータは，レイヤ 2 によるプライベートピアリングや MPLS ネットワーク上で接続された地理的に離れた AS 間接続のみである．しかし，クラスタ外の AS 間接続について，AS ごとの出現回数を調査したところ，本表の結果となった．

表 5.7　大陸トポロジに含まれる AS 間接続数

Area	IIJ (AS2497)	WIDE (AS2500)
All	348	59
Asia	32	38
Europe	32	2
Oceania	6	-
US-PST	215	2
US-MST	1	-
US-CST	9	6
US-EST	13	7

表 5.8　クラスタ以外のトポロジデータ

Rank	ASN	# of links	Rank	ASN	# of links
1	701	926	11	6453	181
2	7018	860	12	4766	179
3	1239	706	13	6461	173
4	174	692	14	3561	155
5	3356	648	15	8220	151
6	209	624	16	3320	150
7	2828	302	17	2914	150
8	7132	251	18	12389	138
9	702	210	19	3549	133
10	3786	195	20	1299	124

表 5.7 は，接続数が多い 5 つの AS について，クラスタリング結合されなかった AS 間接続の IP アドレスに対して DNS の逆引きした結果である．AS701, AS1239, AS3356 については，手作業によるホスト名から地理情報が推測可能であった．Undns のルールセットに追加することで，大陸クラスタに結合可能である．しかし，AS174, AS7018 の一部の経路から，本手法を用いたクラスタリングできない AS が存在することが明らかになった．

本検証の結果より，クラスタを結合できなかった多くのデータは，Undns による AS のホスト名ルールの追加により解決可能なことが明らかになった．ISP によるホスト名のルールは，頻繁に変更されるとは考えにくいことから，一度ルールセットが定義すること

表 5.9　クラスタ外データの手動による解析

ASN	# of links	# of Reverse DNS	手作業による地理情報の推測
701	926	831	可能
7018	860	407	不可能
1239	706	657	可能
174	692	392	不可能
3356	648	294	一部可能

で長期間利用可能であると考えられる．今後の課題としてホスト名のルールセットの更新と追加をおこなっていく．また，一部の AS では，ホスト名による地理情報が推測不可能である．AS 間接続の IP アドレスのペアによるクラスタリングも不可能なため，新たなクラスタリング手法が必要と考えられる．*IPaliasresolution* 技術の 1 つである *APAR* は，複数の traceroute 情報と IP アドレスのサブネットを考慮し，グラフ解析することでルータの alias IP アドレスを明らかにする手法である．現在の本手法では，AS 間接続以外の traceroute 情報を利用していない．そこで，今後の課題として，AS 間接続前後の traceroute 情報を考慮することで，グラフ解析によるクラスタリングする手法を検討する．

5.4.3　まとめ

本節では，大陸ごとの AS トポロジを検証した．特定 AS に注目した解析手法の検証では，特定 AS の AS 間接続が存在しないはずの大陸内トポロジに存在していることが明らかになった，引き続き詳細な検証を続けていく．排除されたトポロジ情報の検証では，Undns の AS ホスト名のルールセットを更新することで，クラスタリング可能な AS 間接続情報を改善できることが明らかになった．また，現在の本提案手法のみでは，クラスタできない AS が存在することが明らかになった．これらの AS 間接続については，今後 *IPaliasresolution* 技術を応用したトポロジを解析手法を検討していく．

第 6 章

結論

本章では，本論文の結論と本研究に関する今後の課題についてまとめる．

6.1 本論文のまとめ

本論文では，今までインターネット全体のみを対象に解析されてきた AS トポロジに関して，地域レベルの AS トポロジの推測する手法を提案し，実際の traceroute 情報を用いて大陸ごとの AS トポロジの推測をおこなった．大陸レベルのトポロジの比較により，大陸ごとのトポロジがベキ則に従う類似点とそれぞれの大陸ごとのハブ AS の役割の相違点を明らかにした．また，現在の手法は，国，大陸ごとのトポロジを対象としているが，将来的には，国，都市ごとの AS トポロジを解析できるよう研究を続けていく．本研究により，それぞれの地域から AS トポロジを把握し，地域間における AS トポロジの比較が可能になる．

本論文により，特定の地域におけるインターネット構造の実態が把握可能なことを示した．今後は，インターネットの重点的に強化すべき国や地域を明らかにし，国間の国際相互接続を効率的に推進するためのインプットとなるように解析を続けていく．

6.2 今後の課題と展望

6.2.1 トポロジデータの追加

大陸ごとの AS トポロジを比較したことで，特にアフリカなどの地域での AS 間の接続のデータ数が足りていないことがわかった．本手法は，traceroute データから地域ごとの AS トポロジが解析可能である．今後，地域の範囲をより細かくし，国ごと，都市ごとの AS トポロジの解析には，本手法の検討とさらなる計測対象の地域におけるトポロジデータが必要である．異なるプロジェクトで計測した traceroute データを加えることで足り

ていないデータ数を補える可能性がある．今後の課題として，別のプロジェクトで計測された traceroute データの利用も検討していく．

6.2.2　AS トポロジの推測手法の検討

本論文では，大陸ごとの AS トポロジを対象に手法を検討している．本論文の手法では，第 5.4.2 節の検証から，一部の AS について地理情報によるクラスタリングができないことが明らかになった．今後は，RTT 情報の導入や，traceroute の経路を用いた解析，そして，新たなランドマークの導入により，国レベルのトポロジ解析の精度向上，都市レベルの AS トポロジを推測する手法を検討していく．

6.2.3　地域間の AS トポロジの検証

本論文では，地域に閉じた AS トポロジについて比較と検証をおこなった．今後の課題は，現在排除している地域間の AS 境界も用いることで，地域間の AS 構造の地域性も明らかにすることである．

謝辞

本研究の作成にあたり，ご指導をいただきました慶應義塾大学環境情報学部教授 村井純博士に感謝します．研究の進む方向性について，貴重なご意見をいただきました．同学部教授 中村修博士に感謝します．氏からは，研究の細部に至るまで本質に踏み込んだご指導をしていただきました．

そして，修士在籍中に最も御世話になった IIJ 技術研究所主幹研究員 長健二朗博士に深く感謝致します．博士は，本論文の研究分野であるインターネットの計測技術に精通されており，学部生から非常に多くのご指導を頂きました．改めて感謝申し上げます．次に，政策メディア研究科 鈴木茂哉氏に感謝致します． 氏には，修士在籍中様々なご指導を頂きました．本論執筆中も，お忙しいところご指導いただき，感謝の念に尽きません．

同大学環境情報学部教授 徳田英幸博士，同学部准教授楠本博之博士，同学部准教授 高汐一紀博士, 同学部准教授 三次仁博士，同学部准教授 植原啓介博士，同学部専任講師 重近範行博士，同学部専任講師 中澤仁博士，同学部専任講師 Rodney D. Van Meter III 博士に感謝致します．Toyota ITC Senior Researcher 湧川隆次博士に感謝致します．

学部生から絶えずご指導を頂いた慶應義塾大学村井研究室 南政樹氏，小原泰弘氏，白畑真氏，小椋康平氏，そして，水谷正慶氏に感謝いたします．また，共に修士論文に取り組んだ仲間である、政策・メディア研究科修士課程 金井瑛氏，奥村佑介氏，神谷尚保氏，町田啓太氏に感謝いたします．そして，モービル広域ネットワーク (MAUI) プロジェクト及び徳田・村井・楠本・中村・高汐・重近・湧川・バンミーター研究室の諸氏に感謝致します．特に岡田耕司氏，石原知洋氏，松谷健史氏，堀場勝広氏，久松剛氏，中村友一氏，遠峰隆史氏，塚田学氏，松園和久氏，佐藤龍氏，六田佳祐氏，波多野敏明氏，上原雄貴氏，永山翔太氏，三部剛義氏には日頃の研究室での活動を大きく支えていただきました．ありがとうございます．

最後に私の研究を影ながら支えてくれた両親とソラ，多くの友人・知人に感謝し，謝辞と致します．

2009 年 2 月

空閑 洋平

参考文献

[1] traceroute. *http://www.traceroute.org/#source%20code*.

[2] Global Traffic Map 2005: Publications: TeleGeography. *http://www.telegeography.com/products/map_traffic/*.

[3] Y. Rekhter and T. Li. A border gateway protocol 4 (BGP-4). RFC 1771, IETF, March 1995.

[4] Lixin Gao. On inferring autonomous system relationships in the internet. Global Telecommunications Conference, 2000. GLOBECOM '00. IEEE, 1:387–396 vol.1, 2000.

[5] JPIX: IX Backplane Maximum/Minimum Traffic Volume. *http://www.jpix.ad.jp/jp/techncal/traffic.html*.

[6] Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien, and Renata Teixeira. Avoiding traceroute anomalies with paris traceroute. In IMC '06: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement, pages 153–158, New York, NY, USA, 2006. ACM.

[7] K. Claffy, Tracie E. Monk, and Daniel McRobb. Internet tomography. Nature, January 1999.

[8] Harsha V. Madhyastha, Tomas Isdal, Michael Piatek, Colin Dixon, Thomas Anderson, Arvind Krishnamurthy, and Arun Venkataramani. iplane: an information plane for distributed services. In OSDI '06: Proceedings of the 7th symposium on Operating systems design and implementation, pages 367–380, Berkeley, CA, USA, 2006. USENIX Association.

[9] Lun Li, David Alderson, Walter Willinger, and John Doyle. A first-principles approach to understanding the internet's router-level topology. In SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications, pages 3–14, New York, NY, USA,

2004. ACM.

[10] R. Govindan and H. Tangmunarunkit. Heuristics for internet map discovery. INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 3:1371–1380 vol.3, Mar 2000.

[11] Neil Spring, Ratul Mahajan, and David Wetherall. Measuring isp topologies with rocketfuel. In SIGCOMM '02: Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications, pages 133–145, New York, NY, USA, 2002. ACM.

[12] M.H. Gunes and K. Sarac. Analytical ip alias resolution. Communications, 2006. ICC '06. IEEE International Conference on, 1:459–464, June 2006.

[13] Adam Bender, Rob Sherwood, and Neil Spring. Fixing ally's growing pains with velocity modeling. In IMC '08: Proceedings of the 8th ACM SIGCOMM conference on Internet measurement, pages 337–342, New York, NY, USA, 2008. ACM.

[14] N. Spring, D. Wetherall, and T. Anderson. Scriptroute: A public internet measurement facility. 4th USENIX Symposium on Internet Technologies and Systems, 2002.

[15] Undns. *http://www.scriptroute.org/source/*.

[16] Ming Zhang, Yaoping Ruan, Vivek Pai, and Jennifer Rexford. How dns misnaming distorts internet topology mapping. In ATEC '06: Proceedings of the annual conference on USENIX '06 Annual Technical Conference, pages 34–34, Berkeley, CA, USA, 2006. USENIX Association.

[17] L. Subramanian, S. Agarwal, J. Rexford, and R.H. Katz. Characterizing the internet hierarchy from multiple vantage points. INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 2:618–627 vol.2, 2002.

[18] University of oregon routeviews project. *http://www.routeviews.org/*.

[19] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On power-law relationships of the internet topology. In SIGCOMM '99: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication, pages 251–262, New York, NY, USA, 1999. ACM.

[20] Sharad Jaiswal, Arnold L. Rosenberg, and Don Towsley. Comparing the structure of power-law graphs and the internet as graph. In ICNP '04: Proceedings of the Network Protocols, 12th IEEE International Conference, pages 294–303, Washington, DC, USA, 2004. IEEE Computer Society.

[21] Shi Zhou and Raul J. Mondragon. Accurately modeling the internet topology, 2004.

[22] Priya Mahadevan, Dmitri Krioukov, Marina Fomenkov, Xenofontas Dimitropoulos, k c claffy, and Amin Vahdat. The internet as-level topology: three data sources and one definitive metric. SIGCOMM Comput. Commun. Rev., 36(1):17–26, 2006.

[23] Scamper. *http://www.wand.net.nz/scamper/*.

[24] PlanetLabEurope. *http://planet-lab.eu/*.

[25] DIMES. The dimes project. *http://www.netdimes.org/new/*.

[26] OneLab. *http://www.onelab.eu/*.

[27] TopHat. *http://top-hat.info/*.

[28] Benoit Donnet, Philippe Raoult, Timur Friedman, and Mark Crovella. Efficient algorithms for large-scale topology discovery. In SIGMETRICS '05: Proceedings of the 2005 ACM SIGMETRICS international confere nce on Measurement and modeling of computer systems, pages 327–338, New York, NY, USA, 2005. ACM.

[29] Pch internet exchange dorectory. *http://www.pch.net/ixpdir/Main.pl*.

[30] Peeringdb. *https://www.peeringdb.com/*.

[31] CAIDA. Cisualizing internet topology at a macroscopic scale. *http://www.caida.org/analysis/topology/as_core_network/*.

[32] Accueil LIP6. *http://www.lip6.fr/fr/*.

[33] The traceroute@home project's home page. *http://trhome.sourceforge.net/*.

[34] Kenji Masui and Youki Kadobayashi. N-tap: A platform of large-scale distributed measurement for overlay network applications. In SAINT-W '07: Proceedings of the 2007 International Symposium on Applications and the Internet Workshops, page 60, Washington, DC, USA, 2007. IEEE Computer Society.

付録 A

大規模なインターネットトポロジ情報収集に関する研究活動

A.1 本付録の概要

本論文では，地理的に広域に分散した計測ノードから収集したトポロジ情報を用いて，地理情報を考慮した AS トポロジの解析手法を提案し，解析をおこなった．本論文で用いたトポロジ情報は，CAIDA の Skitter プロジェクトが収集した情報と，iPlane プロジェクトが収集した情報を組み合わせて用いた．本論文のより，大陸レベルの AS トポロジ解析には，解析対象の地域に計測ノードがあることが望ましい．本論文では，トポロジ情報の解析について考察したが，同時に広域にインターネットのトポロジ情報を収集するフランス国立科学研究センター（CNRS）のプロジェクトに参加し，活動している．本論文の付録として，インターネットのトポロジ情報収集に関する活動について述べる．

A.2 活動概要

WIDE プロジェクトと CNRS の間での研究協力の一環として，両組織間で人的交流・学術的交流を目的とした，学生の交換留学制度を設けている．2008 年 7 月 20 日から 2008 年 10 月 13 日までの約 3 ヶ月間，フランスのパリで現地の研究活動に参加した．受け入れ先は， LIP6 (Laboratoire d’Informatique de Paris 6)[32] の Timur Friedman らの研究室である．滞在中は，Timur らによるインターネットトポロジの計測と解析を目的とした TopHat プロジェクトに参加し，本グループのメンバである Thomas Bourgeau と共に本システムのアーキテクチャについて，議論と実装をおこなった．

TopHat プロジェクト [27] は，2008 年に始まった研究プロジェクトであり，現在も基盤アーキテクチャの議論とシステムの研究開発が続いている．今後は，引き続き Timur

指導の元で Tophat の研究グループに参加していく予定であり，研究プロジェクト間の交流を続けていく．

A.3 TopHat

滞在中，Timur らが活動している OneLab プロジェクト [26] の複数ある研究グループのうち，大規模にインターネットトポロジ情報の収集と解析，そして，収集したトポロジ情報をユーザへ提供するシステムを研究開発している TopHat のグループに参加した．OneLab は, 次世代インターネットのテストベット環境構築を目的として，PlanetLab[24] の普及と高度化を進めているヨーロッパ圏の研究グループである．実際に，TopHat グループでは，PlanetLab 上に計測環境を構築し，トポロジデータの収集を開始している．

滞在時，TopHat グループでは，ネットワーク環境に協調したトポロジ探索手法の検討とユーザへのトポロジ情報の提供システムの開発を開始した状況であった．前者の計測手法については， DoubleTree[28] と呼ばれるトポロジ探索アルゴリズムを提案している．また，計測基盤のアーキテクチャは，Timur らが過去に構築したシステムである traceroute@home[33] のアーキテクチャを基にして構築されている． 本アーキテクチャは，PlaneLab 上に計測ノードを配置し，計測ノード間の情報共有に WIDE プロジェクトのメンバである益井 賢次氏が研究開発している N-TAP[34] を採用している．滞在中は，TopHat のトポロジ探索手法の検討とユーザへのトポロジ情報の提供システムを担当して作業した．

A.4 トポロジ計測手法

本留学では，はじめに TopHat の計測基盤アーキテクチャの理解と議論をおこなった．具体的には，Timur らの論文と実際に動作している TopHat のコードを参照し，Tophat の Problem statement をまとめた．その上で，滞在中の後半では，トポロジ情報の提供システムを担当して，StitchRoute アルゴリズムの提案と実装をおこなった．

本節は，StitchRoute アルゴリズムを説明するために必要である TopHat の計測機能である DoubleTree アルゴリズムを述べる．StitchRoute については，次節で扱う．

A.4.1 DoubleTree アルゴリズム概要

TopHat はインターネット上に分散配置された PlanetLab 上のノードがそれぞれ traceroute を用いてトポロジ情報を収集し，得られた情報を統合することで，インターネット全体のトポロジ情報を探索する．このような，インターネット全体のトポロジを複

数の計測ノードから分散して探索する手法は，他の研究プロジェクトでも採用されている一般的なトポロジ探索手法である．しかし，複数の計測ノードを用いた手法は，探索パケットが重複した経路や同一の宛先ノードを対象とすることで，探索途中のネットワークに対して，高負荷をかける恐れがある．また，重複して経路を探索するため，1 回のトポロジ探索により多くの時間を消費する．

図 A.1 にインターネット上のノードに対して，高負荷をかける状況を示す．左図では，同一ネットワーク上に複数存在する計測ノードからトポロジ情報を探索した結果，イントラドメインの共通する経路を重複探索している．これらの計測ノードから同一のタイミングでトポロジ探索することで，対象ルータに対して必要以上の負荷をかける恐れがある．また，右図では，分散した計測ノードから，同一の宛先ノードをを対象にトポロジ探索することで，宛先ノードとその近くのネットワークを重複探索している．

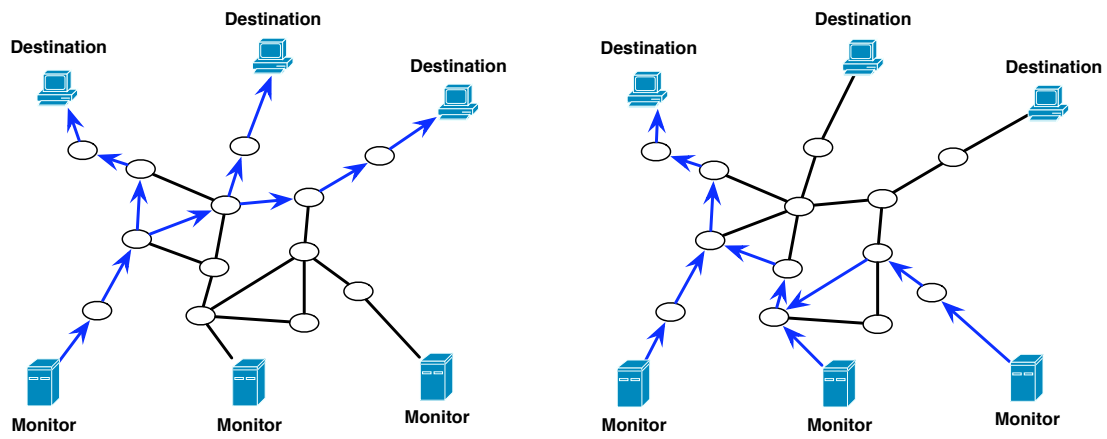

図 A.1　DoubleTree の扱うトポロジ計測時の問題点

DoubleTree は，計測ノード間で Stop set と呼ばれる探索した経路の情報を共有することで，このような重複経路の探索を排除する探索アルゴリズムである．それにより，大規模に展開されるインターネットのトポロジ情報を既存の手法に比べて素早く，また，探索対象のネットワークを構成するノードの負荷を削減できる．

A.4.2　DoubleTree によるトポロジ探索

DoubleTree によるトポロジ探索手法を述べる．DoubleTree では，traceroute と同じように IP パケットの TTL 値を漸増させることで，計測ノードと宛先ノード間のトポロジ情報を収集する．traceroute との違いは，計測開始時の TTL 値である．計測開始時に，TTL 値 h で計測を開始する．探索は，TTL 値を $h+1, h+2, \ldots$ と漸増させながら宛先ノードまでの経路を探索する Forward probing と，TTL 値を $h-1, h-2, \ldots$ と漸減さ

せながら計測ノードまでの逆向きの経路を探索する Backward probing を交互におこなう．それにより，経路の中央近くから末端のノード方向へ探索していく．TTL の初期値 h の選定には，事前に計測した，任意の 2 ノード間における直接疎応答された確率 p を基に決定している．また，計測ノード間で探索済みのトポロジ情報を共有することで，重複した経路を探索しないよう調整する．Forward probing と Backward probing を実行する際，毎試行時に Stop set を参照することで，重複探索を判断する．Forward probing では，Global Stop Set と呼ばれるインタフェース IP アドレスと宛先 IP アドレス，計測ノードから成るデータを用いて判断する．Global Stop Set は，計測ノード間で共有される．探索は，Forward probing の毎試行時に Global Stop Set を参照する．Global Stop Set 内から計測ノード自身の宛先 IP アドレスと直前に発見したインタフェース IP アドレスのペアを発見した場合，探索を停止する．一方，Backward probing では，Local Stop Set と呼ばれるインタフェース IP アドレスのデータから判断する．Local Stop Set は，各計測ノードのみで参照され，共有しない．

図 A.2 に DoubleTree の動作概要を示す．計測ノード (Monitor) A, B, C から宛先ノード (Destination) P に対して DoubleTree を用いてトポロジ情報を収集する．TTL の初期値は $h = 2$ とする．

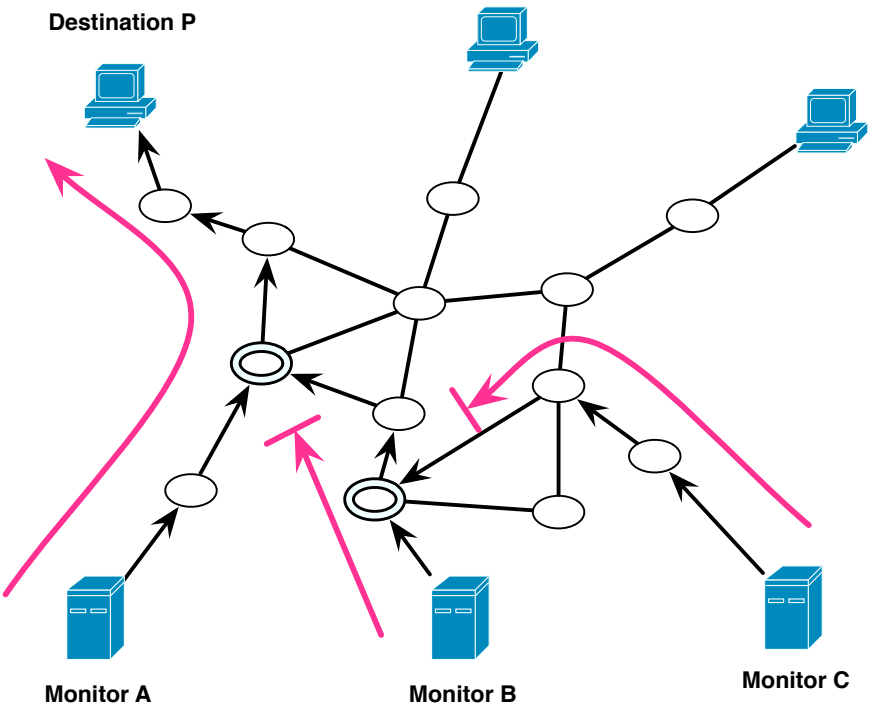

図 A.2　DoubleTree 動作概要

(1) 計測ノード A から P までの経路を探索する．A は，探索した経路情報を Global

Stop Set として保存し，さらに A, B, C 間で共有する．

(2) 計測ノード B が宛先ノード P までの経路を探索する．B は，A と同様に TTL 値 $h = 2$ から探索を開始する．B の Forward probing は，探索途中で Global Stop Set から，すでに A が発見した重複経路部分を発見し，探索を停止される．

(3) 計測ノード C が宛先ノード P までの経路を探索する．C は，B が発見済みの経路までを探索する．

A.4.3 現状と今後

現在，TopHat グループでは，PlanetLab 上で DoubleTree を用いたトポロジの探索を開始している．今後の予定は，インターネット上の経路ループやロードバランスされた経路の検出に対応する目的で，Paris-traceroute による経路探索手法の置き換え作業をおこなう．

また，グループでは，DoubleTree の TTL 初期値 h の決定方法や，計測結果の提示方法についての議論を続けている．DoubleTree で発見されるトポロジの情報量は，TTL の初期値 h で大きく変動するためである．今後は，実際に TopHat システムで収集したトポロジ情報を解析することで TTL 値を検討し，計測手法を改善していく予定である．

A.5 StitchRoute

滞在中おこなった作業は，第 A.4 節で述べた計測手法についての議論に加え，本システムの基幹機能である任意の IP アドレス 2 点間の経路を返答する機能を検討し，実装した．以後，本機能を StitchRoute と呼ぶ．

Tophat グループでは，traceroute@home から続くトポロジ計測基盤アーキテクチャの機能をほぼ実装し，実際にトポロジ計測をはじめている．次のステップとして，TopHat グループでは，収集したトポロジ情報を一般ユーザに提供するアーキテクチャを検討している．本システムは，ユーザの XML-RPC によるリクエストに応答する手法で任意の 2 点間の IP アドレスまたは AS 番号の経路情報を提供することを考えている．

A.5.1 目的

StitchRoute の目的を述べる．本機能は，TopHat が収集した経路の断片 (*piece* と呼ぶ) をつなぎ合わせることで，任意の 2 点間の IP アドレス間の経路を算出するものである．

用語を定義する．計測ノードから宛先ノードまでの完全な経路は，$\mathbf{r} = (r_o, r_1, r_2, \ldots, r_\ell)$ と定義する．r_0 は，ある計測ノードのソース IP アドレスであ

り，それぞれの値である $r_i, i > 0$ は，ホップ i ごとで発見されたインタフェース IP アドレスである．r_ℓ は，宛先 IP アドレスである．次に，piece は，$\mathbf{p} = (p_o, p_1, p_2, \ldots, p_\ell)$ と定義する．p_0 は，*piece* の先頭 IP アドレスであり，p_ℓ は，*piece* の末尾の IP アドレスである．p_0 と p_ℓ は，*Stopset* で経路探索が終了している可能性があるため，必ずしもそれぞれ，計測ノードと宛先ノードの IP アドレスとはかぎならない．この時，*StitchRoute* の目的は，ユーザのリクエスト (S, D) に対して，$\mathbf{p} = (p_0, p_1, p_2, \ldots, p_\ell)$ をつなぎ合わせ，$\mathbf{r} = ((p_{0_0}, \ldots, p_{0_\ell}), (p_{1_0}, \ldots, p_{1_\ell}), \ldots, (p_{n_0}, \ldots, p_{n_\ell}))$ を返答することである．S は *source*，D は destination を表す．

A.5.2 背景

StitchRoute が必要となる背景を述べる．TopHat では，DoubleTree アルゴリズムを用いて，ネットワークに協調したトポロジ探索の計測基盤を構築した．TopHat では，通常の traceroute の結果と異なり，計測した経路を断片化された状態でシステム内部で保持する．*piece* は，DoubleTree による経路探索がすでに発見した重複経路で探索を中止するために発生する．

図 A.3 に経路が断片化される様子を示す．計測ノード A, B, C は，宛先ノード P, Q に対して初期 TTL 値 $h = 3$ で経路を探索する．はじめに P に対して経路を探索する．この時点で，各計測ノードは，Local Stop Set を持つ．次に，宛先ノード Q に対して経路探索する．計測ノード A は，自身の Local Stop Set を参照することから，XQ 間の経路で探索を終了する．計測ノード B は，自身の Local Stop Set と，計測ノード A と共有して得た Global Stop Set を参照し，XY 間の経路のみを探索する．計測ノード C は，すでに CQ 間の経路探索が終了していることから，初期探索のみ実行し，探索を終了する．本図の状況で収集したトポロジ情報から CQ 間の経路を知るには，計測ノード C が探索した CX の経路情報と計測ノード B が探索した XY の経路情報，そして計測ノード A が探索した YQ 間の経路をつなぎ合わせる必要がある．

A.5.3 piece

TopHat では，ユーザからのリクエストに応答するため，事前に計測した *pieces* をさらに細かくし，ホップ間の IP アドレスペアに変換して保持する．細かくしたデータ構造は，$(s_n, d_{q+1}, star)$ であり，本データを *piece* と呼ぶ．s は *source*，d を *destinaton*，*star* は IP アドレスペア間に存在する *noreplay* の数をそれぞれ表す．*piece* を

star の初期値は 0 であり，dst_{q+1} が *noreplay* だった場合，*star* の値を漸増し，dst_{q+2} を *destination* とする．また，ロードバランスされたネットワークでは，経路探索時に，

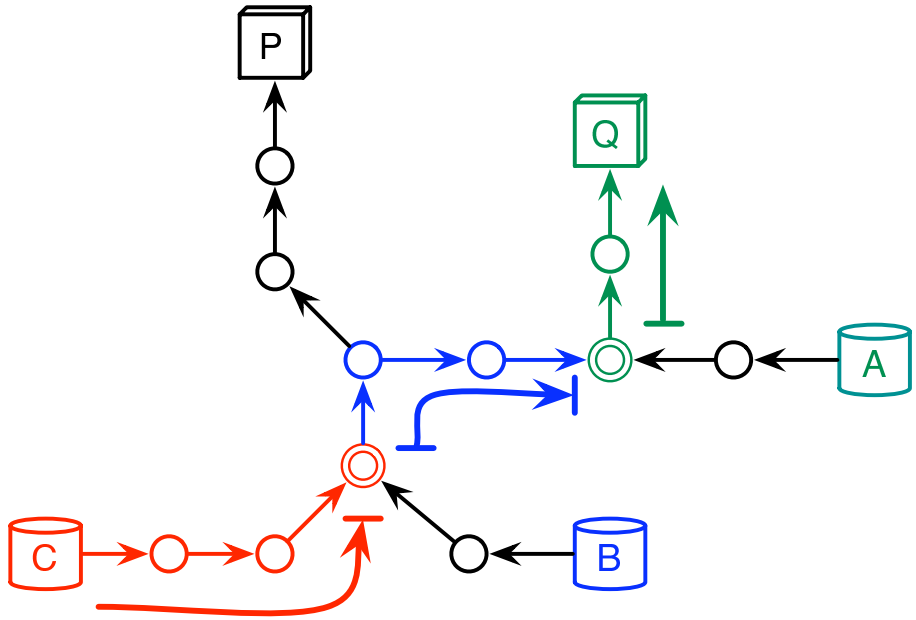

図 A.3　DoubleTree による探索経路の断片化

同一の TTL 値で複数 IP アドレスから返答される場合がある．このようなトポロジデータは，返答された IP アドレスの数だけ *piece* に分解する．

A.5.4　アルゴリズム

図 1 に *StitchRoute* アルゴリズムを示す．*StitchRoute* アルゴリズムは，*piece* をつなぎ合わせることで，経路 **r** を探索する．本手法は，インターネットトポロジを探索が計測ノードを幹とした木構造のように，探索することに注目する．木構造の探索には，反復深化深さ優先探索 (iterative deepening depth-first search: IDSearch) を採用した．単純な深さ優先探索のみでは，経路ループがデータ内に含まる場合に探索が終了しない恐れがある．また幅優先探索では，より多くのメモリを消費することから，探索深度を 1 から漸増させながら深さ優先探索する IDSearch を採用した．

A.5.5　今後の予定

StitchRoute は，滞在中にの実装作業をおこなった，今後，実際に TopHat が収集したトポロジ情報を用いて，実行時間を計測する予定である．

Algorithm 1 Stitchroute algorithm

```
1: procedure StitchRoute(S, D)                          ▷ source, destination
2:     i ← 1
3:     loop
4:         F̂ ← IdSearch(i, S)
5:         if F̂ ∩ D then
6:             response                                  ▷ Output
7:         else if |F̂ ∩ B̂| > 0 then
8:             response                                  ▷ No output
9:         end if
10:        i ← i + 1
11:    end loop
12: end procedure

13: procedure IDSearch(l, P)                            ▷ limit, Path
14:    n ← |P|
15:    m ← P_ℓ
16:    L̂ ← ∅
17:    if n = l then
18:        L̂ ← L̂ ∪ {P_ℓ}
19:    else
20:        for all c ∈ Adjacent(m) do
21:            if c ∩ P = ∅ then
22:                P ← P ∪ {c}
23:                IdSearch(l, P)
24:                P ← P − {P_ℓ}
25:            end if
26:        end for
27:    end if
28: end procedure
```

A.6 まとめ

WIDE プロジェクトと CNRS 間の交換留学生として渡仏し，現地の研究プロジェクトに参加した．滞在中は，Timur 指導の元，インターネットトポロジの計測と解析を目的とした TopHat プロジェクトに参加し，本グループのメンバである Thomas Bourgeau と共に本システムのアーキテクチャについて，議論と実装をおこなった．現地での Tophat グループに参加しての作業は，今後も続けていく予定であり，研究協力関係はこれからも継続される．