

## 第 XXVI 部

### ネットワーク情報の視覚化



## 第 26 部

### ネットワーク情報の視覚化

#### 第 1 章 netviz ワーキンググループについて

netviz ワーキンググループの目的は、ネットワークの視覚化およびその他の表現、そのためのツールやノウハウに関する議論と情報の共有を行うことである。

現在、WIDE プロジェクトには、ビジュアライゼーション自体を研究テーマにしている者はいないが、目に見えないインターネット技術を説明するため、また、説得力のあるプレゼンテーションのために、道具としてビジュアライゼーションを使いこなすことはすべての研究者に必要なになっている。

従来は、mawi ワーキンググループなどで計測データの視覚化などの議論をしてきたが、研究目的を持ったワーキンググループの中では、どうしても研究内容に話が集中するため、参加者の間口が狭くなる。そこで、独立したワーキンググループとして、データ表現にフォーカスした netviz が設立された。

netviz ワーキンググループでは、ネットワークと表現に関わるさまざまな話題を取り扱う。現在のところ、参加者の関心は以下の 3 つの領域に分けられる。

##### 1. データ表現、データ抽出

ネットワークトポロジやシステムログの視覚化、音声や光などアンビエントメディアを用いたネットワーク状態の表現など、データ表現の手法に関する領域。

##### 2. ツール学

グラフィックツール、文章作成やプレゼンテーション作成ツール、AV 機器の使い方のノウハウに関する領域。ここには、色や音声の有効な使い方、アクセシビリティやユーザビリティも含まれる。

##### 3. 社会学、文化学

ネットワークとデータ表現を広くとらえ、社会的あるいは文化学的な議論を行うことも、データ表現に対する意識を高めるのに役立つと考えている。

前述したように、現状はこれらを主研究テーマとする者はいないため、当面は、情報交換とノウハウの蓄積を行っていく予定である。将来的には、WIDE プロジェクト内にもこのような分野の研究者が出てくることを期待したい。

#### 第 2 章 netviz ワーキンググループ 2005 年度の活動概要

2005 年 6 月 22 日に第 1 回 netviz ワークショップを開催し、ネットワークとデータ表現に特化して議論する試みを行った。ここでは、情報の可視化を中心に関連研究の調査報告、WIDE プロジェクト内にある可視化すべき情報についての項目の洗い出し、ワーキンググループの方向性や今後のアクションアイテムについて話し合った。

ワーキンググループの方向性として主に 2 つが挙げられる。

1 つ目は一般論としてのプレゼンテーションに関するもので、専門の研究者がいなくても、データ表現の質の向上、あるいは、その重要性の継続的な認識のため、WIDE プロジェクト内に議論の場を提供することである。ネットワーク研究におけるデータ表現の重要性の認識は浸透しつつあるものの、個別には予算やその他のリソースの割り当てがし辛い現状があるため、体制を作ってそのような研究を支援する。

2 つ目は、データ表現技術の中に、インターネットらしさ、WIDE プロジェクトらしさをアピールするようネットワーク技術者の視点を取り入れることである。たとえば、スケーラビリティはインターネット技術の主要なテーマなので、グローバルなスケールのインターネットをうまく伝えるようなデータ表現や、膨大な時系列データをインタラクティブに扱えるようなデータ表現は、我々が主体となって作っていかなければならない。

具体的なアクションアイテムとしては、以下ののようなものがある。

## ●第26部 ネットワーク情報の視覚化

- デザイナーとのコラボレーション  
「国内インターネットのトラフィックマトリクス」を題材にして、計測データの可視化の試みを行う。これに関しては次章で詳細を報告する。
- 球面ディスプレイ、3D表現などを使い、地球規模でインターネットを表現する方法を検討する。
- データ表現コンポーネントの自作  
コンピュータ制御で光る、あるいは、動くような工作が比較的簡単にできるようになってきた。これらのノウハウを蓄積し、新しいデータ表現への応用を検討する。

また、ワーキンググループの体制に関しては、前チェアの吉田薫の留学にともない、2005年10月より和田英一がチェアとなっている。

### 第3章 トラフィックローカリティの可視化——試行錯誤の過程——

#### 3.1 背景

可視化とは、見えないものを見えるようにすることである。しかし、そこから言いたいことを伝達するためには、どのような見せ方があるか、どのような手法がよいのかを検討するべき、との考えから「国内インターネットのトラフィックマトリクス」をテーマにトラフィックローカリティの可視化を試みた。

6月のnetvizワークショップ、9月のWIDE合宿などでの議論を含め、サンプルデータのみがあるという状態から最終的な形に整えるまでの試行錯誤の過程を報告する。

#### 3.2 可視化とは

何を見せるか = 何を言いたいのか

これには差、同様、特異などいろいろ存在する。これは情報を可視化する上で最も重要なポイントであり、すなわち可視化とは「表現を最適化する」ということだと考える。表現を最適化するためには、いくつかの段階を踏み、必要に応じて戻りながら進めていくべきである。

Phase1：何が見えるか

Phase2：何を見せるか

Phase3：どんな手法にするか（グラフ、3DCG、アニメーションなど）

Phase4：どのように見せるか（色、質感、テキストなど）

また、表現の傾向として大きく2つに分けることができる。

#### ○具体的な傾向

正確さを重視し、わかりやすく簡潔にまとめる（数字の羅列、目盛など）。

#### ○抽象的な傾向

見え方を重視し、グラフィカルな要素を強くする（格好よく、きれいになど）。

今回は後者の抽象的な傾向に注力して進める。

#### 3.3 最適化までの試行錯誤——サンプルデータ

まずは、phase1：どのように見えるのか、である。各都道府県の総トラフィック量のサンプルデータ（図3.1）を元にボリュームをつかむ。図3.1の通りに北海道から順に上から並べ、INとOUTをそれぞれ左右に配置し棒グラフを作成した。（図3.2）

|    |      |           | IN          | OUT         |
|----|------|-----------|-------------|-------------|
| 1  | 北海道  | hokkaido  | 833.3133541 | 703.328458  |
| 2  | 青森県  | aomori    | 118.5739343 | 97.39717566 |
| 3  | 岩手県  | iwate     | 137.1537408 | 94.25967554 |
| 4  | 宮城県  | miyagi    | 278.709743  | 228.8963367 |
| 5  | 秋田県  | akita     | 119.7014571 | 104.6289502 |
| 6  | 山形県  | yamagata  | 93.64432145 | 77.75200704 |
| 7  | 福島県  | fukushima | 211.3554335 | 158.3476694 |
| 8  | 茨城県  | ibaraki   | 167.7697585 | 159.3878722 |
| 9  | 栃木県  | tochigi   | 178.882625  | 174.3752537 |
| 10 | 群馬県  | gunma     | 221.9218553 | 191.0498037 |
| 11 | 埼玉県  | saitama   | 868.223687  | 914.8335375 |
| 12 | 千葉県  | chiba     | 732.2393839 | 722.7037926 |
| 13 | 東京都  | tokyo     | 2025.111421 | 1950.629774 |
| 14 | 神奈川県 | kanagawa  | 1153.189676 | 1286.570861 |

図3.1. 各都道府県の総トラフィック量のサンプルデータ

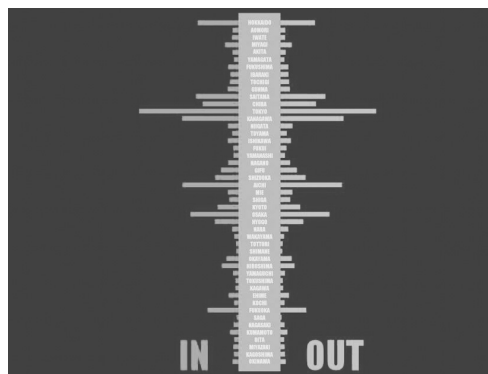


図3.2. 図3.1を棒グラフに表す

## 要素

- 47 都道府県
- IN：各県に入ってくる総トラフィック量
- OUT：各県から出ていく総トラフィック量

これらから、東京都の数値が圧倒的に大きいことがわかる。その他、神奈川県、愛知県などの主要都市の数値が高い。現時点では目盛や単位がないため、各県の微妙な差がわかる程度で、数値を把握することができない。

つぎに、「見せたいものが何か」を探るために、どんな見せ方があるか、いくつか作成した（Phase2、3、4の反復）。

A から E の 5 種類を作成し、このうちダッシュがついているものは、議論の内容を反映させて更新している。

## 【タイプ A】図 3.3

地理情報として日本地図を使用することによってシンボリックなイメージになり、感覚で地理情報を把握できる。IN（入ってくる）と OUT（出ていく）をイメージできるように、それぞれ IN を下方に、OUT を上方に高さを与え、3 次元で表現した。

## 要素

- 各県の総トラフィック量（IN/OUT）
- 都道府県（日本地図を使用）

## 評価

視点を固定させると見えない部分が出てくる。また、回転させると見慣れない地図の景色になるため、わかりづらい。IN と OUT については、天地をつけるとマイナスのイメージにつながり、嫌がられる場合がある。

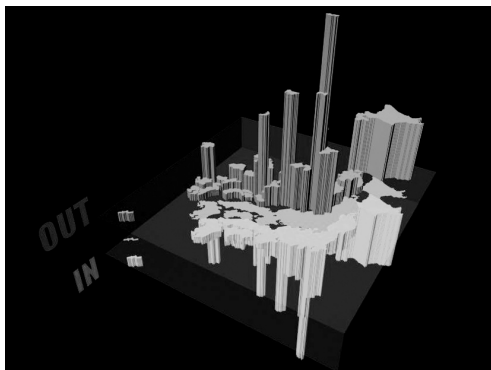


図 3.3. 図 3.2 の棒グラフを日本地図に表す

## 【タイプ B】図 3.4

軸を 1 つ加えることによって興味深いデータが生まれるのではないかと思います、1 人当たりのトラフィック量を各県の人口から算出し、都道府県、各県の総トラフィック量に加えて 3 つ目の軸として 3 次元で表現してみる。3 つの軸が交わる場所に IN と OUT のポイントをそれぞれ打った。

## 要素

- X：都道府県
- Y：各県 1 人当たりのトラフィック量
- Z：各県の総トラフィック量
- ポイント：IN/OUT

## 評価

単位や目盛がないため、何を示しているかがわからない。また、ポイントが重なって見えにくく、各県の比較がしにくい。

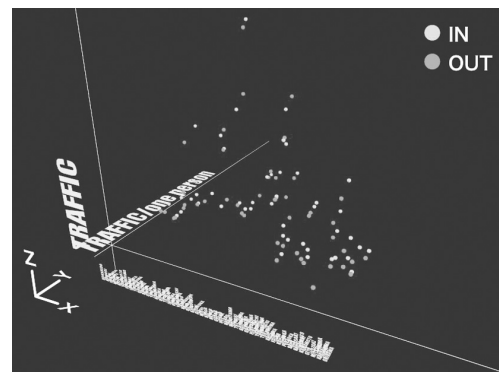


図 3.4. XYZ に要素を与え 3 次元で表す

## 【タイプ B'】図 3.5

## 改善点

- 各県の数値の差をわかりやすくするために、スケールを 4 倍にする。
- 軸の方向を変えて読み取りやすくする。
- 背景を明るくし、暗い印象を払拭する。

## 要素

- X：各県の総トラフィック量
- Y：各県 1 人当たりのトラフィック量
- Z：ソース—都道府県
- ポイント：IN/OUT

## 評価

スケールを調整してもまだポイントが重なって効果がない。スケールの調整よりもそれぞれのサイズや形状を変える方が、効果が出るのではないかと。静止画としてみるには、都道府県が細かすぎる。また、

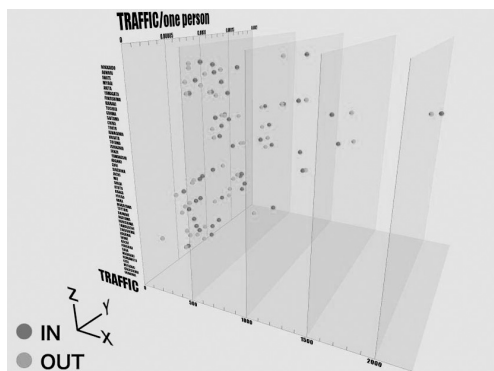


図 3.5. 図 3.4 を改善

新たに軸を加えると趣旨が混同するため、加える必要はない。

#### 【タイプ C】図 3.6

たくさんの要素を一度に見せたらどうなるか試してみる。図 3.1 を元に各県の総トラフィック量に比例して配分されるとして県間トラフィックデータを算出し、47 都道府県では細かすぎるため、東北・関東・北陸・中部・近畿・中国・四国・九州の 8 エリアに区分した。ソースを X、デスティネーションを Y にとってマトリクスに示し、各エリアに配分される値を Z として頂点に面を張った。効果的なライティングを施せば、僅かな値の差や全体のボリュームがわかりやすい。

##### 要素

- 都道府県を 8 エリアに区分
- 3 つの軸  
X：ソース（8 エリア）  
Y：デスティネーション（8 エリア）  
Z：トラフィック量

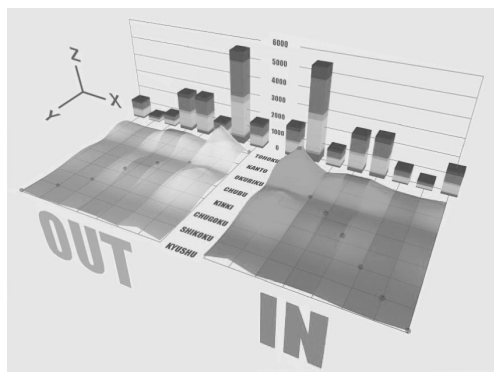


図 3.6. 要素が多すぎてわかりづらい例

- 棒グラフ：各エリアのトラフィック（IN/OUT）
- ポイント：各エリアに閉じているトラフィック評価

見る方向によってさまざまな情報を得ることができることを狙ったが、要素を盛り込みすぎてどこを見ればよいかわからなくなり、「何を見せたいのか」が明確でない悪い例。

#### 【タイプ D】図 3.7

モザイク画のようなイメージで、行（OUT）と列（IN）の幅を等しくしたマトリクスを着色し、彩度と明度で表現した。

##### 要素

- 各県の総トラフィック量 IN（横軸）
- 各県の総トラフィック量 OUT（縦軸）

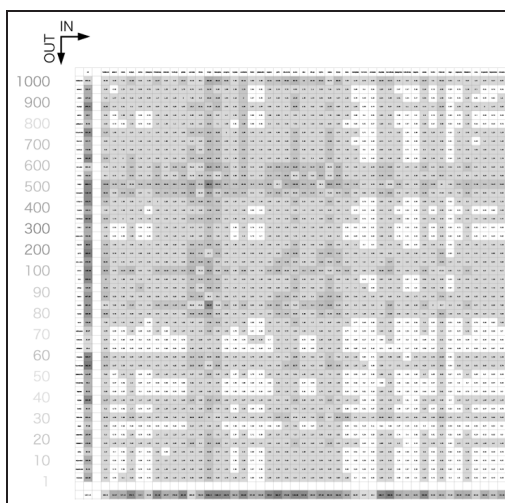


図 3.7. 数値で色分けしマトリクスを着色

##### 評価

微妙な変化がわかって興味深いのが、0～90、100～1000 の 2 つに分けて配色したために割り当て方が不適切で誤解を招く恐れがある。0～1000 の間でグラデーションの配色をするべきである。

#### 【タイプ E】図 3.8

これまで IN と OUT のトラフィック総量が異なるためにそれぞれ作図していたが、量ではなく地域間の分布の割合に的を絞り、要素をまとめてみる。8 エリア間のトラフィック分布を縦軸にソース、横軸にデスティネーションを取り、通信相手の分布を行と列で示す。ソース、デスティネーションそれぞれの各エリアに正規化した割合を線の太さ、トラフィッ

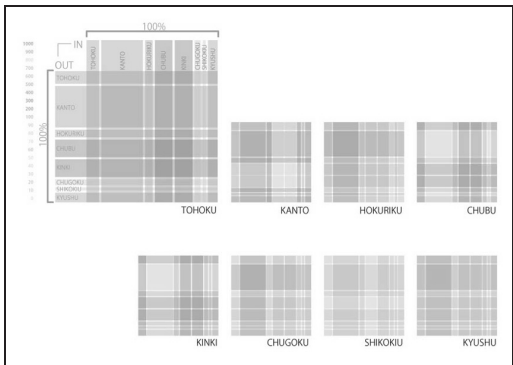


図 3.8. トラフィック量と地域間の分布の割合を線の太さと色で表す

量を色で表現した。微妙な変化が見えるのではないかと期待した。

要素

- 都道府県（8 エリア）
- トラフィック量（色）
- ソース（縦軸）
- デスティネーション（横軸）
- ソース / デスティネーションの各エリアに正規化した割合（線の太さ）

評価

各エリアによって柄がちがって面白いので、47 都道府県をやってみてはどうだろうか。チェック柄を意識しすぎたため、色を半透明にすると縦軸と横軸が重なった部分の色に対して説明がつかない。

- ここまでのまとめ
  - 各県の総トラフィック量を正規化  
各県のトラフィック量の値の差が大きすぎるため、総トラフィック量 (IN/OUT) を一度に見せるのは難しい。全体を表示している状態での最大値と最小値の比較は困難であるため、合計値を正規化してスケールを均一にし、割合を見る方がよいのではないか。

- 47 都道府県を 8 エリアに  
47 都道府県を並べただけではわかりづらい。日本地図は都道府県を表記する必要はなくなるが、配置換えができず、バリエーションには限りがある。8 エリア（東北・関東・北陸・中部・近畿・中国・四国・九州）にわけた方がよいのではないか。
- 色への配慮  
背景を黒にすると、コントラストがはっきりするが暗い印象になる。一方、明るい配色のみで構成すると区別がつきにくい。
- 何をさせるか
  - 各県に閉じているトラフィック量
  - 「各県の地域性がない」ということ
- 成果物  
ポスターを作成する。

3.4 最適化までの試行錯誤——正式データ

2005 年 7 月 4 日から 7 月 10 日までの 1 週間、ある国内 ISP の全国ブロードバンドユーザを対象に、アクセス網を収容しているルータ群で、Sampled NetFlow 機能を用いて収集されたデータが届いた時点でこれを正式データとし、サンプルデータで作成したものをベースに差し替えて進めた。前の項目で述べたとおり、何をさせたいか、どう表現するか要素がまとまってきたが、Phase2、3、4 を反復しながら最終形へ整えていく。

正式データ

図 3.9 に例を挙げた。ソースを縦軸、デスティネーションを横軸に取り、トラフィック量全体を 100 として各県に分布する割合を示す。A-D の部分に 47 都道府県もしくは 8 エリアが入り、srcB と dstB が交わる“18”が B に閉じているトラフィックの割合である。

- データ 1——全体に対する割合

| dst \ src | ALL | A  | B  | C  | D  |
|-----------|-----|----|----|----|----|
| ALL       | 100 | 20 | 40 | 20 | 20 |
| A         | 20  | 4  | 8  | 4  | 4  |
| B         | 50  | 10 | 18 | 12 | 10 |
| C         | 10  | 2  | 4  | 2  | 2  |
| D         | 20  | 4  | 10 | 2  | 4  |

データ 1

| dst \ src | ALL | A  | B  | C  | D  |
|-----------|-----|----|----|----|----|
| ALL       | 100 | 20 | 40 | 20 | 20 |
| A         | 100 | 20 | 40 | 20 | 20 |
| B         | 100 | 20 | 36 | 24 | 20 |
| C         | 100 | 20 | 40 | 20 | 20 |
| D         | 100 | 20 | 50 | 10 | 20 |

データ 2

| dst \ src | ALL | A   | B   | C   | D   |
|-----------|-----|-----|-----|-----|-----|
| ALL       | 100 | 100 | 100 | 100 | 100 |
| A         | 20  | 20  | 20  | 20  | 20  |
| B         | 50  | 45  | 60  | 50  | 50  |
| C         | 10  | 10  | 10  | 10  | 10  |
| D         | 20  | 25  | 10  | 20  | 20  |

データ 3

図 3.9. 使用するデータ例

- データ2—ソースの各県に正規化した割合
- データ3—デスティネーションの各県に正規化した割合

【タイプE'】図3.10

47都道府県におけるデータ2を横軸(左上) データ3を縦軸(上中央)に取り、それぞれの数値を太さで表現し、これらを重ねて、ひとつの図にまとめた。

改善点

- 説明がつくように、色を2色に限定

評価

2色にしたことで何を見せたいかが明確になった。各県で試したが、細かすぎて変化がよくわからない。

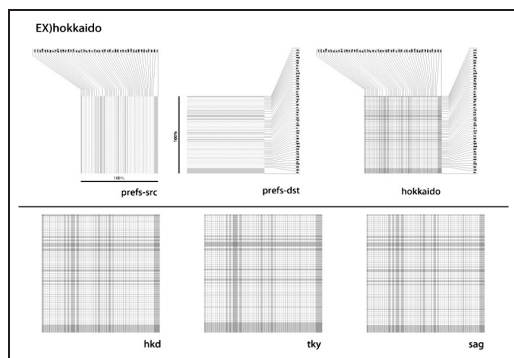


図3.10. データ例2・3を線の太さで2次元化

【タイプF】図3.11

図右のように、まずは原点を中央にして左右にそれぞれ47都道府県におけるデータ2とデータ3を棒グラフで表した。各県に閉じたトラフィックは別の色で示している。棒グラフを立体にして日本地図の県庁所在地の位置に配置した。日本地図は棒グラフの原点を意味する。

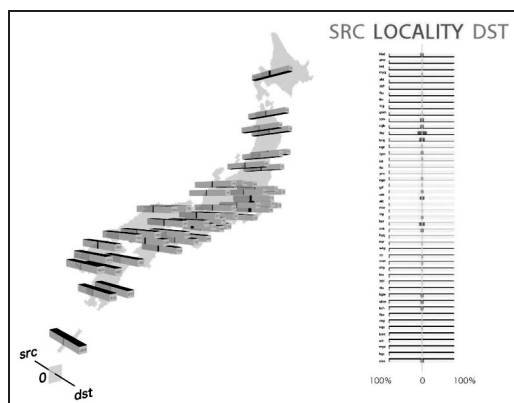


図3.11. データ例2・3を棒グラフで表し3次元化

要素

- 都道府県(日本地図を使用)
- ローカリティ:各県に閉じたトラフィック
- ソース:src/左・奥
- デスティネーション:dst/右・手前

評価

閉じているトラフィックの割合が少ないため、この部分のみを着色しても変化がわからず効果はない。重なった部分が見えない。

【タイプG】図3.12

各エリアの情報を日本地図に配置。上面、正面、側面の3方向にそれぞれ異なる情報を表している。見る方向によって違う情報を得られる。

要素

- 都道府県(8エリア)
- 3方向
  - X:ソースの各エリアに正規化した割合
  - Y:デスティネーションの各エリアに正規化した割合
  - Z:ローカリティ(各エリアに閉じたトラフィック)

評価

X軸、Y軸から見ると、ボールが重なって見えない。上面に視点を固定してボール自体を見たい方向に回転させてはどうか。ひと目でわかるような見せ方を考えるべきである。

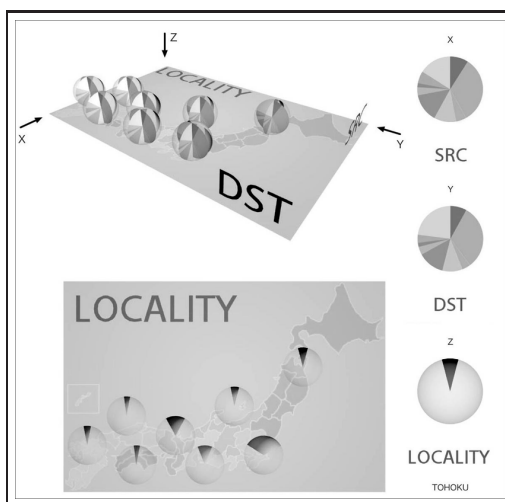


図3.12. 3種類の円グラフで構成された球を日本地図に配置

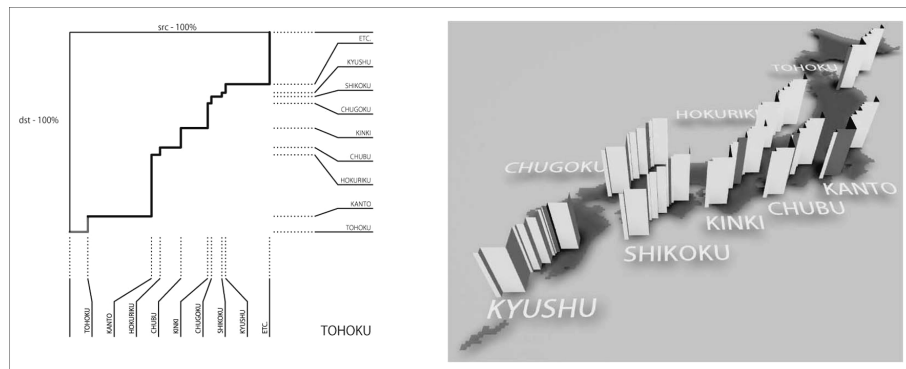


図 3.13. 地域間の割合を線の長さで表す

## 【タイプ H】図 3.13

まず 2 次元で、データ 2 と 3 の数値をそれぞれ横線と縦線の折れ線で表し、エリア内に閉じたトラフィックを別の色で表現（図 3.13 左、東北を例）した。この要領で各エリアの折れ線を作り、それぞれ日本地図の各エリアの中央に配置し（図 3.13 右）各エリアに閉じたトラフィックをわかりやすくするために折れ線に高さを与えた（高さ自体は意味をもたない）。

## 要素

- 都道府県（8 エリア）
- ソースの各エリアに正規化した割合（縦軸）
- デスティネーションの各エリアに正規化した割合（横軸）
- エリア内に閉じているトラフィック

## 評価

要素が整理されていてわかりやすいが、高さに定義がないのであれば 3 次元で表現する必要はないのではない。3 次元ではなく、すべてをまとめて 2 次元で表したらどうか。また、上が九州ではなく東北になる方が見やすい。

## 【タイプ H'】図 3.14

## 改善点

- 各エリアの折れ線ひとつの図にまとめて 2 次元にし、エリアごとに色分けする
- 各エリアに閉じたトラフィックを矩形の領域で表現

## 要素

- 都道府県（8 エリア）
- ソースの各エリアに正規化した割合（縦軸）
- デスティネーションの各エリアに正規化した割合（横軸）

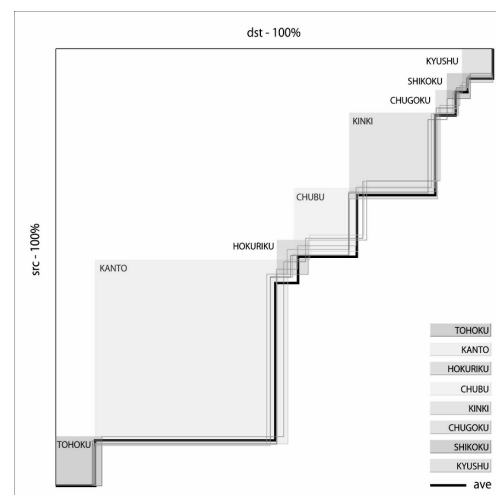


図 3.14. 図 3.13 を改善

- エリア内に閉じているトラフィック（矩形の領域）

## 評価

上が九州、下が東北というのはわかりづらいので、東北が上にくるように反転させた方がよいのではない。また、どこかの県を取り出して、エリアではなく各県の分布を表したらどうか。

## 【タイプ I】図 3.15

各県の分布の割合に図の左にデータ 2（src 縦軸/dst 横軸）右にデータ 3（src 横軸/dst 縦軸）の数値を直線で表し、各県ごとに線で結んだ。

## 要素

- 都道府県（上から順に北海道→沖縄）
- 8 エリアを色分け
- ソースの各県に正規化した割合
- デスティネーションの各県に正規化した割合

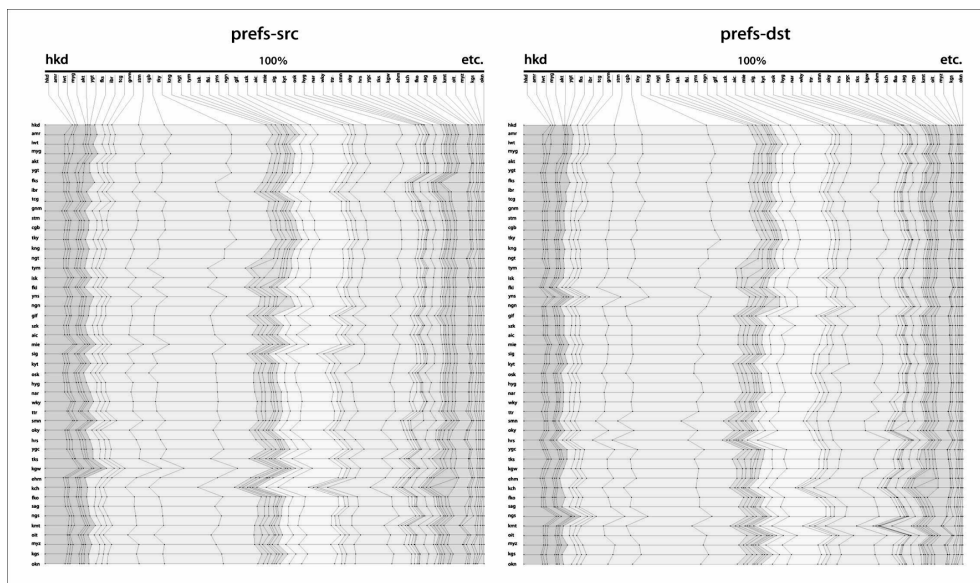


図 3.15. 各県の数値を線で結びエリアで色分け

#### 評価

微妙な変化がよく見えてわかりやすいが、色が淡すぎるのではないか。

#### ●ここまでのまとめ

##### ○都道府県の表現は8エリアに区分

日本地図を使用して3次元で表現することは、重なる部分が見えない場合が多く、適切ではない。3次元で作成したものを2次元に収めるのはかなり困難である。

また、47都道府県を並べると要素が多くなるため、8エリアに区分した方がわかりやすい。

##### ○何をさせるか

- 各エリアとも地域性がないこと
- 各エリアに閉じたトラフィック量の割合
- 各エリアの割合

##### ○2次元で表現

時間の経過にともなって劇的に変化するデータではないことから、アニメーションや3次元の表現ではなく2次元の表現にする。

したがって、「各エリアとも地域性がない」という傾向をひと目で表現するのに最適なのは、【タイプH'】であるといえる。

### 3.5 トラフィックローカリティの可視化——完成へ

タイプH'を整え、完成させる。

東北が上にくるように配置を変更することによって日本地図のような配置になり、感覚的にわかりや

すくなった。都道府県の表現についてはこれからの課題でもあるが、要素をできるだけ少なくまとめて8エリアに区分した。凡例は日本地図を使用し、地理情報と配色を一度に見せることができた。タイプH'では、各エリアに閉じているトラフィックを各エリアの数値を取って矩形で表していたが、重なってわかりづらいため、平均の数値を取った。また、図がどのように構成されているのかをマトリクスを添えて説明している。図を大きくみせることを最優先し、説明文を加え、配色やレイアウト、フォントやサイズなどを見やすく仕上げて完成とした(図3.16)。

### 3.6 まとめ

「何を言いたいのか」は、可視化するために最も重要なポイントである。今回は、これを見つけるために試行錯誤を重ねて多くの時間を費やした。「何を言いたいのか」が明確であれば、「phase3 どのような手法にするか」、「phase4 どのように見せるか」に早い段階でたどり着くことができ、より質の高い効果的な可視化、さらにはプレゼンテーションが可能になる。そのためには「何が見えるのか」に留まらず、そこから「何を言いたいのか」「どう見せたいのか」を積極的に考えていくべきだろう。日頃から何を言おうとしているのか意識して物事を見ることが、可視化やプレゼンテーションをする際に大きな助けとなってくれるはずである。

今後も継続して取り組み、「可視化マニュアル」を

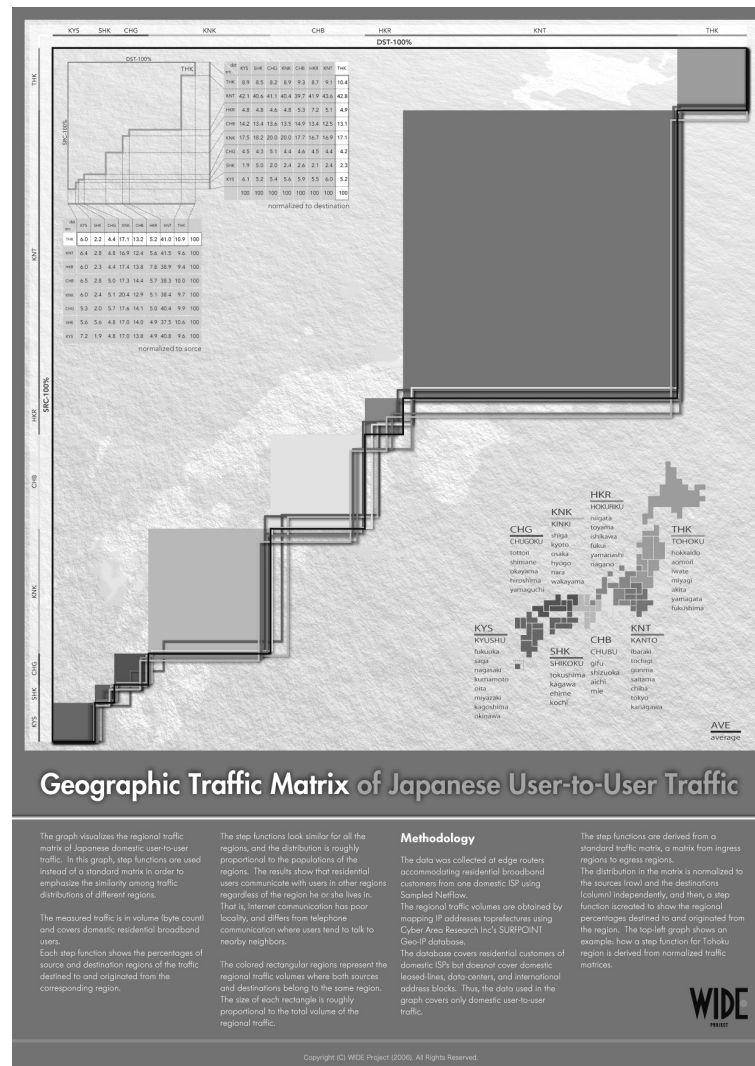


图 3.16. 完成图

作ることが可能になるくらいに事例を積み重ねて、プレゼンテーション能力の向上につなげていきたい。

目に見える成果を出すことを目標にしている。

## 第4章 まとめ

インターネット研究において、計測データをいかに表現するかは重要なテーマとなってきた。一方で、WIDE プロジェクトにはデータ表現を主テーマとするような研究者がいなかったこともあり、なかなかワーキンググループとしてアウトプットが出ない状態が続いている。

そこで来年度は、具体的にテーマを絞り、まさに

