2009－OS－110 (11)
2009／1／29

IM-VIS: 情報統合機能を用いたネットワーク状態可視化システムの提案

明石 修† 寺内 敦† 福田 健介‡
{akashi,terauchi}@core.ntt.co.jp kensuke@nii.ac.jp
NTT 未来ねっと研究所†，国立情報学研究所‡

概要

インターネットは複数の自律システム (AS) から構成される巨大分散システムであり，ネットワークの運用・障害解析・制御においては経路情報を含めた系全体の振舞を理解することは重要である．しかしながら，経路情報は空間・時間的に変化しながら伝搬し，なおかつ他のプロトコルや様々なイベントと相互作用をするため，その振舞を直接理解することは困難である．この問題に対応するため，BGP の時系列情報から構成したトポロジ情報上に，さまざまなネットワーク状態を表す情報を統合することにより系全体のビューを明確にし，更に状況に応じて柔軟かつ動的に特定範囲の情報をフィルタ，あるいは統合・拡張していくことにより，詳細な情報把握を行うことを可能とするナビゲーションシステム IM-VIS システムを提案し，そのアーキテクチャに関して述べる．

IM-VIS: Inter-domain network-status visualization system based on information integration

Osamu Akashi† , Atsushi terauchi† , Kensuke Fukuda‡
{akashi,terauchi}@core.ntt.co.jp, kensuke@nii.ac.jp
NTT Network Innovation Laboratories†, National Institute of Informatics‡

abstract

The Internet consists of more than 28,000 autonomous systems (ASes). It is important to realize the behavior of inter-AS routing information to appropriately manage, diagnose, and control the Internet. However, the routing information mutates and spreads through the Internet interacting with various events and this situation makes the complete understanding difficult. To cope with this problem, we propose the IM-VIS system that can flexibly map various network-related information over the BGP topologies at designated time. This information used in flexible integration has been continuously observed at multiple ASes and statistically analyzed. IM-VIS can support operators effectively recognize the current and previous network status on a on-demand basis.

1 はじめに

インターネットは複数の自律システム (Autonomous Systems: AS) から構成される巨大分散システムであり，BGP-4[1] により運用される．この AS 間経路制御を含めたネットワークの運用・障害解析・制御を適切に行うためには，経路情報を含めた系全体の振舞を理解することは重要である．しかしながら，経路情報は AS 上を hop-by-hop 方式で空間・時間的に変化しながら伝搬するため，その振舞自体を把握することも難しく，多くの障害事例も報告されている [2]．更に，経路の到達性は，

他のプロトコルとの相互作用や，DNS などの外部システムなどの影響など様々な外部イベントと関連することも，その振舞を把握することを困難にしている．

これら全体の振舞を理解するためには，グローバルな視点からの状況把握が必要であるが，各 AS は独立した組織がそれぞれの管理ポリシに従って運用するため，中央管理方式によるシステム構築は困難であり，分散協調問題解決による手法が望まれている．

本稿では，複数の視点で観測した BGP の時系列情報から構成した，指定した時点の BGP トポロジ情報に，さまざまなネットワーク状態を表す情報を柔軟に統合することにより系全体のビューを明確にすることが可能な可視化システム IM-VIS を提案し，そのアーキテクチャに関して述べる．IM-VIS は，状況に応じて柔軟かつ動的に特定範囲の情報をフィルタ，あるいは統合・拡張していくことにより，詳細な情報把握のためのナビゲーションを行うことが可能である．

2 背景と要求条件

前章で述べたように，AS 間の経路情報は，空間的・時間的に変動しながら AS 間で伝播していくため，その把握自体が困難である．グローバルな完全ビューを構成することは困難であるため，我々はその一部分を推論することによって，グローバルビューを代用する手法を取ってきた．

広報した BGP 経路の到達性確認に基づく検証に関しては，複数 AS に分散配置したエージェントによる経路障害診断システム ENCORE[3] を提案し，経路障害が BGP 上の静的な状態として観測可能な症例に対応した．その動作モデルを図 1 に示す．本モデルは，自分が広報した BGP 経路は外部 AS でのみ観測可能であるため，外部 AS に配置したエージェントと協調して，自分が広報した BGP 経路を観測し，検証，障害発生時の診断を行う．ENCORE は仮説とその検証ルールからなる診断知識を持ち，定常観測動作からスレッシュホールド値を越えた値を発見した場合や，外部 AS からの観測による経路広報異常の可能性を発見した場合，診断動作を行う．

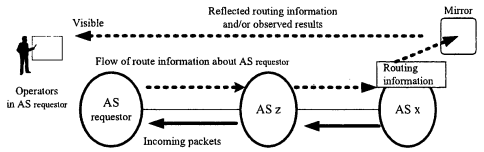

図 1: BGP 経路情報の到達性確認による検証

実際に管理する AS の障害対応記録によれば，このような症例はオペレーション対応数の大半を占め [4]，ENCORE の商用 ISP への適用や，実ネットワークでの経路ハイジャック監視を通じて，その有効性は検証されてきた．しかしながら，より詳細な原因の切り分けや，障害箇所特定のためには，BGP アップデートレベルの時系列解析，他のプロトコルとの相関を含めた包括的な解析が必要となる．そのため，ENCORE の拡張形であり，BGP の時系列解析，他のプロトコルとの相関を含めた包括的な IP サービス診断システム IP-MIND[5] も提案してきた．図 2 にマルチエージェントを用いたネットワークの観測・診断・制御のためのシステム全体像を示す．本図では，診断系システムの他にも，トラッフィクなどのネットワーク状況に応じたフィードバックベースの経路制御システム AISLE/VR[6]，ポリシの集約による動的な ISP 上流でのトラフィックフィルタリングシステム Chorus[7] などの，その他のマルチエージェントベースのシステムも示す．

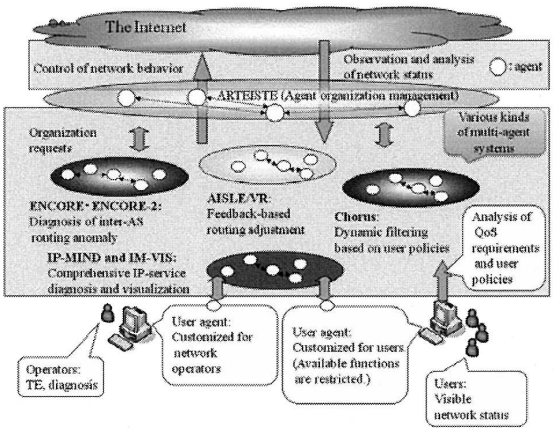

図 2: マルチエージェントによる観測・診断・制御

一方，IP-MIND の観測・診断動作において，場合の数が増えるにつれ，システムの持つ仮説・検証の枠組みでの知識が複雑になり，適用したローカルな環境に応じた知識や，ユーザ定義のマクロや診断ルールの維持管理自体が難しくなってきた．また，状態の把握や記述自体も複雑になり，システムの使用・診断結果の提示やその理解自体も高度な知識を要求するようになってきた．

そこで再度，詳細レベルの情報を包含した上で系全体のビューを明確にし，オペレータがネットワークの状況を把握するのをサポートするため，IP-MIND システムの持つデータを他のシステムからアクセス可能な DB へ格納した上で，全体の情報を可視化するアプローチを取る．

具体的には，複数の視点で観測した BGP の時系列情報から構成したトポロジ情報上に，さまざまなネットワーク状態を表す情報を柔軟に統合する可視化システム IM-VIS を提案する．IM-VIS では，状況に応じて柔軟かつ動的に特定範囲の情報をフィルタ，拡張，統合していくことにより，詳細な情報把握のためのナビゲーションを行う．オペレータをサポートするための，本システムへの要求条件は以下の通りである．

- 可視化情報は，オペレータとのインタラクションにおいて，柔軟にフィルタ処理を行ったり，あるいは逆に詳細情報を拡大表示したりすることが可能．
- 表示する情報は，さまざまな種類のネットワーク状態が表示できるように，複数のデータソースから統合可能とする．
- 表示する情報には，直交した概念として時間の表現を入れ，時間を指定した表示，あるいは遡って表示する機能を導入する
- 複数の観測ポイントの情報は並行して表示可能であるが，同時にこれらの異なる観測ポイントの統合表示を行うことを可能とし，直交した概念として扱う．
- これらの操作は，オペレータとのインタラクションにおいて，柔軟かつ統一した枠組みで制御できるように，

  (“データに対する統合・絞込・拡張条件”，“表示アクション”)

  の組からなる記述で全て行う．これを，インタラクションフィルタと呼ぶ．

3 IM-VIS 基本設計

オペレータに対して明確なネットワークのビューを提供するため，ナビゲーションモデルを定義する．ベースとしては，BGP の AS パス属性から構成した，ノードを AS，エッジを接続関係とするグラフを用いる．

なお，ある AS の受信した BGP 情報から構成した AS グラフは，必ずしもスパニングツリーとはならない．例えば，ある $AS0$ が受信したプレフィックス P_1 の AS パスが ($AS1$ $AS2$ $AS3$)，プレフィックス P_2 の AS パスが ($AS4$, $AS2$ $AS5$) であれば，内部にループ構造をもつグラフ，すなわち ($AS0$ $AS1$ $AS2$ $AS4$ $AS0$) が構成される．実際に，上流のトランジット AS が複数あれば，複数に広報するのが通常であるし，異なるオリジン AS の広報した BGP エントリが，様々な要因で別経路をたどって受信されるケースは少なくない．

IM-VIS システムの表示基本モデルの柱は，図 3 に示すように，以下の 3 つの軸から構成する．

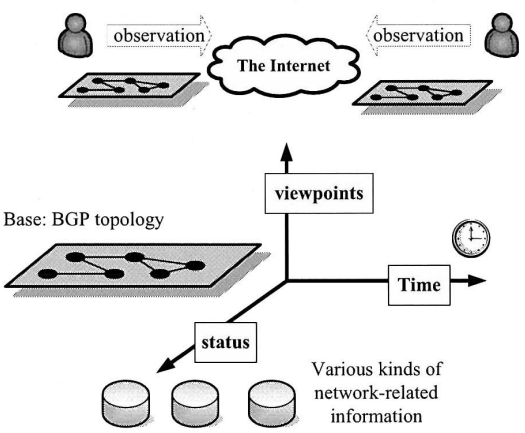

図 3: IM-VIS コンセプト

- ベースとして，ある時点，ある観測ポイントで作成した BGP トポロジー図を用いる

- BGP トポロジー情報をベースに，様々なネットワークステータス情報を重ね合わせる．
- 直交する操作軸として，複数の視点からの軸 (観測箇所の違い，空間軸) を置く．
- 直交する操作軸として，異なる時間を表現する時間軸を置く．

また，これらを統合して操作するための操作基本モデルとして，アクティブフィルタを導入し，これらのフィルタを結合する動作により，操作モデルを定義する．

- 表示する対象は，重ね合わせ/絞り込み/検索/拡張条件からなる操作対象設定条件で決定し，それらを表示アクションにより指定した表示を行う．これをアクティブフィルタと呼ぶ．
- 各アクティブフィルタを動的に組合わせることにより，表示対象を変化させ，ナビゲーション機能を実現する．

なお，情報の重ね合わせ，空間軸の概念は，アクティブフィルタモデルに straightforward にマップできるが，時間の表現 (時間軸) も，時系列データからなる BGP トポロジやネットワーク関連情報を，与えられたある時点という条件で絞り込む操作と考え，統一して扱うこととする．

IM-VIS のアクティブフィルタを柔軟かつ統一した簡潔な記述に構成するため，図 4 に示すように，対象設定条件とそれに対する表示アクションを直列に連結するモデルを採用する．すなわち，アクティブフィルタを適用した対象を，更に次のアクティブフィルタの入力として，対象設定条件の候補とする．これは，UNIX シェルのパイプと同様な構造である．IM-VIS では，全ての絞り込みの概念をこのフィルタの重ね合わせモデルで表現する．また，よく使われる表現に関しては，GUI 部分からセレクトできる形にする．

```
アクティブフィルタ列 ::=
 {アクティブフィルタ}{'|' アクティブフィルタ}*
アクティブフィルタ ::=
 {重ね合わせ/絞り込み/検索/拡張条件}
 {表示アクション}+
```

図 4: アクティブフィルタ基本モデル

図 5 に，IM-VIS で想定する様々な状態把握のためのなナビゲーション例を示す．

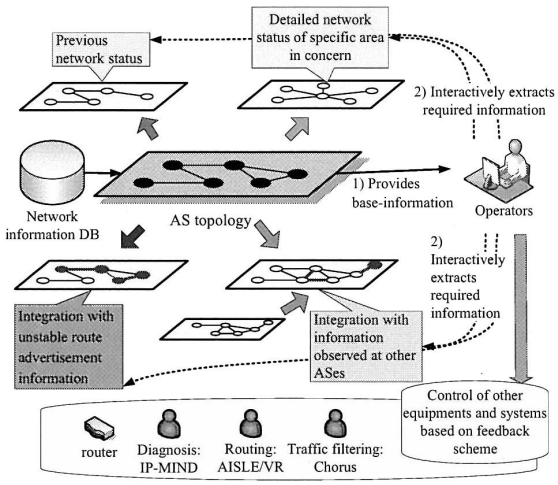

図 5: IM-VIS システムによる様々な状態把握例

この例では，ベースとなるある時点の BGP トポロジー情報に対し，過去のトポロジー情報を表示する例，不安定な経路を重ね合わせて異なる色で表示する例，注目点を詳細表示する例，他の AS で観測した BGP 経路と重ね合わせて表示する例を示す．これらはオペレータが IM-VIS が表示する情報を見ながら，インタラクティブに必要な情報を取り出すイメージである．

なお将来的には，IM-VIS 単体での使用の他に，IM-VIS で把握した情報に基づいて，IP-MIND 診断システム自体への情報のフィードバックを行うことや，更にはフィードバックベースで他の機器やシステムを制御することを想定している．具体的な制御対象は，ルータなどのネットワーク機器に加えて，図 2 で紹介した，トラッフィクなどのネットワーク状況に応じたフィードバックベースの経路制御システム AISLE/VR[6]，ポリシの集約による動的な ISP 上流でのトラフィックフィルタリングシステム Chorus[7] などである．

同時に，IM-VIS を使用して，オペレータがパフォーマンスの問題などの原因を切り分ける過程で，異常な攻撃トラフィックや不安定な経路を抽出した場合，その抽出ルール表現を利用して加工・拡

張を加えることにより，ルータの設定変更に適用したり，トラフィックのフィルタや経路の迂回などの設定ルールに適用することを想定している．

4 IP-MIND/IM-VIS システム

4.1 システム構造

図 6 に，IP-MIND システムの構造を示す．IM-VIS システム自体は，ENCORE と同様，Allegro Common Lisp 上に構築されたエージェントシステムであり，BGP peer としてボーダルータと BGP 接続をし，その update メッセージを記録する．

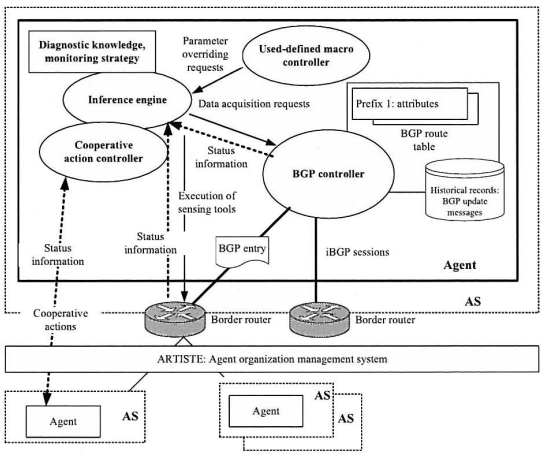

図 6: IP-MIND システム構造

IM-VIS システムは，当面 IP-MIND と独立して動作させるため，また BGP 接続をする IP-MIND システムに影響を及ぼさないため，独立したプロセスとして動作させる．すなわち，IP-MIND システムは従来通り，取得した BGP update レベルの詳細情報をログファイルに書き込み，IM-VIS システムの方で，それらを読み込み，データベース (DB) に格納するデーモンを動作させる．IM-VIS の本体機能である表示・解析系モジュールは，それらの DB の存在を仮定して実装する．

なお，DB は表示・解析系と連動して，インターラクティブに使える範囲の速度が要求される．現在，日本 (2 箇所)，US の観測ポイントで IP-MIND システムは動作中であり，現在の実装におけるデータ使用量概要に関しては，表 1 に示す．なお，DB の設計と実装の詳細に関しては別途 [8] にて報告する．

AS	上流 transit AS 数	経路数	DB データ量 (KB)
AS1(JP)	1	277866	869
AS2(JP)	3	269589	3501
AS3(US)	1	267221	2207

表 1: DB データ量 (2008/12/1-12/20)

AS1, AS3 は上流トランジット AS として tier-1 ISP と接続する AS であり，IP-MIND システムは AS 内のボーダルータと iBGP 接続をする．AS2 においては，上流トランジット AS は 3 箇所であり，IP-MIND システムは AS2 内のルートリフレクタと iBGP 接続する形態のため，受信する BGP update メッセージ数は最も多い．AS3 は AS1 に比べて受信メッセージ数が多いが，期間中にメンテナンス関係で何度か BGP セッションのリセットを起こし，BGP フルルートの受信を行っているため，その影響も受けている．

図 7 に，IM-VIS システムの構造を示す．表示系および DB への書き込みを行うデーモンは，IP-MIND システムと同様に Allegro Common Lisp で記述し，DB は MySQL(Linux/FreeBSD) を使用する．

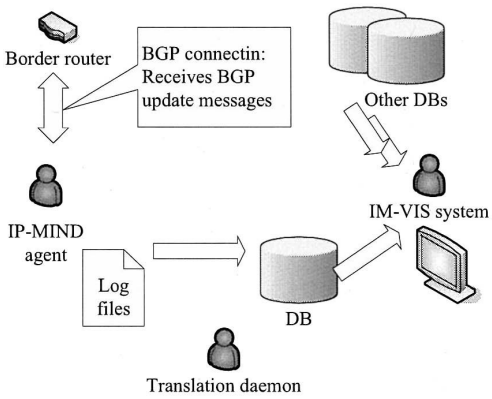

図 7: IM-VIS システム構造

4.2 画面デザインと基本操作

IM-VIS 立ち上げの手順は以下の通りである.

- 読み込み先の DB，対象データの時間を指定.
- 該当 DB から BGP フルルートテーブルを構成し，その時点で存在する BGP エントリの AS パス属性から AS の隣接情報を取り出す.
- 表示初期状態の条件を指定
  - 中央に配置する AS
  - 初期フィルタとして，最大表示 AS 数，表示する AS の connectivity 値[1]の最低値を指定

図 8 に，IM-VIS システムの画面表示例を示す. 基本表示は，AS トポロジーを表すための部分がメインであり，ノード (AS) を数字付の円，エッジ (接続関係) を直線として表示してある. ノード配置は，ばねモデルの一種である Kamada-Kawai model[9] を採用した. 基本機能は，カスタマイズウィンドウにて，関連 Lisp 関数のパラメータを変えたり，修正・拡張して関数を評価することにより，動的なカスタマイズは可能であるが，使用頻度の高そうなものは，画面右側の GUI 操作可能なパネル，画面上段のプルダウンメニューから操作可能とした. また，ノード，エッジ上にカーソルを置くことにより，随時属性情報の一部が画面下に表示されるが，詳細情報はクリックすることによるポップアップ画面にて表示される.

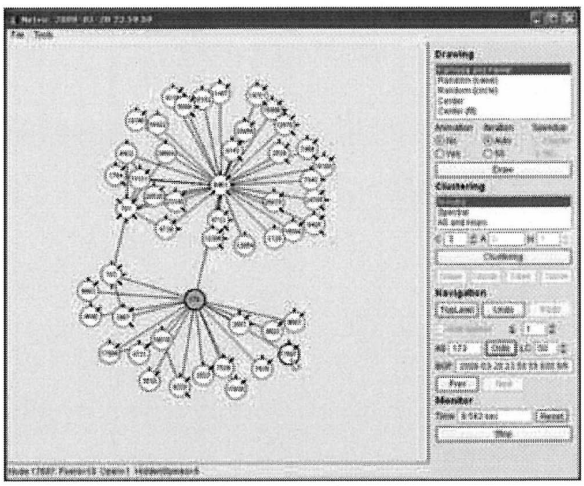

図 8: IM-VIS 画面表示例

5 IM-VIS システムの適用

以下に，IM-VIS で想定する典型的な重ね合わせの例を列挙する.

- 複数の観測地点から見たトポロジー自体の合成 (OR 合成)
  - トポロジ表示自体に対するフィルタ ... AS 数，AS の connectivity 値，中心 AS からの hop 数
- ネットワークの状態を表すデータの重ね合わせ
  - エッジを流れる BGP エントリ数. 更に複数地点の観測ポイントの BGP エントリ数 (AND/OR 合成)
  - 不安定な経路の情報

また，現時点で未実装であるが，重ね合わせるネット関連の状態を表す情報としては，以下のような項目を想定している.

- DNS データ
- トラフィック，フローの情報

なお，選択された対象はそのままに，直交する軸である，表示する情報の時間を動かすことが可能である.

5.1 IM-VIS システムの適用例

本節では，IM-VIS を用いた状態把握と絞り込みの使用例を示す.

まず最初に，読み込む対象の BGP データの入っている DB(AS2) と時刻 (2008-12-20, 23:59:59) を指定し，BGP トポロジデータを読み込む. 最初は，表示 AS 数 200，表示 connectivity>=1 の状態で立ち上げる. (図 9)

次に，スライドバーを動かし，上で表示された状態の 40 分過去の時点でのトポロジ表示をしたの

[1]BGP トポロジー上の peering 数

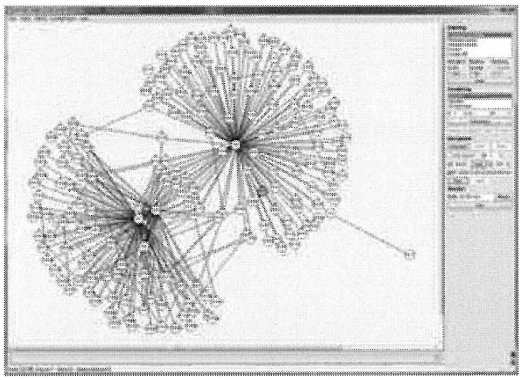

図 9: IM-VIS 初期立ち上げ画面

が，図 10 である．

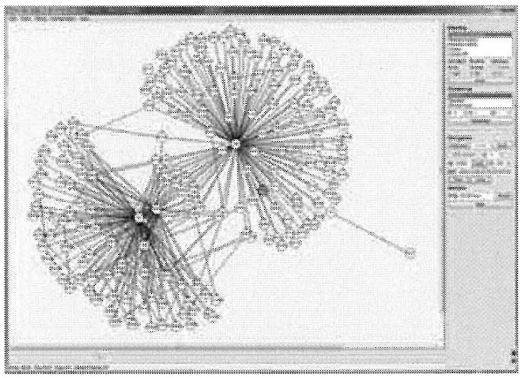

図 10: IM-VIS の時系列を指定した表示

次に，不安定な経路を選択し，赤く表示するアクションで表示されたものが，図 11 である．ここで，不安定な経路とは，過去 15 分以内に `withdrawn`，`advertise` を 3 回以上広告した prefix に対して `t` を返す関数を用いた．これは，Cisco ルータのデフォルト値である．

IM-VIS では，オペレータの望む状態把握を行うため，表示されるデータを見ながら，表示情報を動的に絞り混み，あるいは展開していくことが可能である．すなわち，クリックすることで，対象 AS の属性 (peer 情報，connectivity など) を確認することが可能であり，また 1 AS-hop づつトポロジ情報を展開して表示することが可能である．次

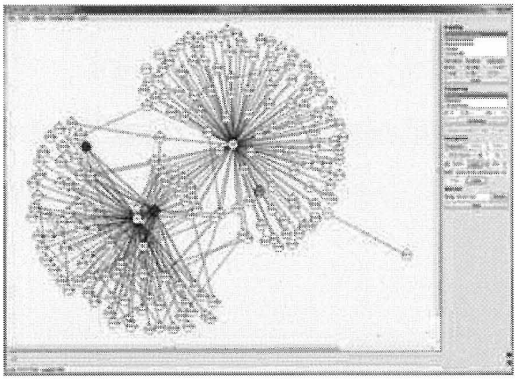

図 11: BGP トポロジーに不安定な経路をマップした図

の例では (図 12)，更に注目点のトポロジ情報を展開した様子を示す．

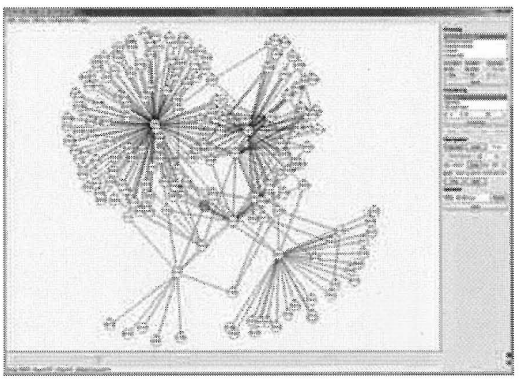

図 12: 更に注視点を展開した表示

6 おわりに

オペレータがインターネットの状態を把握することをサポートする可視化システム IM-VIS を提案し，そのプロトタイプを作成した．IM-VIS では，BGP の情報から構成したトポロジ情報上に，さまざまなネットワーク状態を表す情報を柔軟に統合することにより系全体のビューを明確にし，時間軸，複数の視点からなる異なる空間軸でのビューの変更を可能とする．操作においては，状況に応

じて柔軟かつ動的に特定範囲の情報をフィルタ，あるいは統合・拡張していくアクティブフィルタの枠組みを導入し，オペレータのナビゲーション機能を統一した．今後は，統合する情報の種類を増やし，実データを使った検証実験を継続してその有効性を検証することと共に，IP-MIND/IM-VIS観測ポイントの観測ポイント自体も増やしていく予定である．

参考文献

[1] Y. Rekhter and T. Li. "A Border Gateway Protocol 4 (BGP-4)", 1995. RFC1771.

[2] The North American Network Operators' Group. NANOG mailing list. http://www.nanog.org.

[3] O. Akashi, T. Sugawara, K. Murakami, M. Maruyama, and K. Koyanagi. "Agent System for Inter-AS Routing Error Diagnosis". *IEEE Internet Computing*, Vol. 6, No. 3, pp. 78 – 82, 2002.

[4] O. Akashi, A. Terauchi, K. Fukuda, T. Hirotsu, M. Maruyama, and T. Sugawara. "Detection and Diagnosis of Inter-AS Routing Anomalies by Cooperative Intelligent Agents". In *16th IFIP/IEEE Int'l Workshop on Distributed Systems: Operations and Management (DSOM/MANWEEK)*, pp. 181–192. Springer-Verlag, Oct 2005. LNCS 3775.

[5] O. Akashi and A. Terauchi. "Diagnosis of IP-Service Anomalies based on BGP-Update Temporal Analysis". In *8th IEEE Int'l Workshop on IP Operations and Management (IPOM/MANWEEK)*, pp. 28–40. Springer-Verlag, Sep 2008. LNCS5275.

[6] O. Akashi, K. Fukuda, T. Hirotsu, and T. Sugawara. "Policy-based BGP-control Architecture for Inter-AS Routing Adjustment". *Journal on Computer Communications, Elsevier*, Vol. 31, pp. 2996–3002, 2008.

[7] 瀬林, 明石, 丸山. "集約したユーザポリシを用いた攻撃防御方式の提案". 電子情報通信学会報告 (IN), Vol. 106, No. 578, pp. 113–118, Mar 2007.

[8] 寺内, 明石. "ネットワーク解析診断のための情報蓄積システムの提案". 電子情報通信学会研究報告 (IN), Feb 2009. (To appear).

[9] T. Kamada and S. Kawai. "An Algorithm for Drawing General Undirected Graphs". *Information processing Letters*, Vol. 31, No. 1, pp. 7 – 15, 1989. Elsevier.