ハイパーオブジェクトによるネットワークトラフィック可視化システムの実装と評価

上田 達巳† 高井 昌彰‡

北海道大学大学院情報科学研究科† 北海道大学情報基盤センター‡

1. はじめに

近年，インターネットの利用形態は，Web，メール，チャット，ゲーム，ストリーミング，P2P など極めて多岐にわたっている他，DDoS など大規模なネットワーク障害を引き起こす要因も増え続けている．そのため，ネットワークトラフィックをわかりやすく可視化し，時々刻々と変化するネットワーク状況やトラフィック分布の遷移現象を，リアルタイムで直観的に把握したいというネットワーク管理者からのニーズは高い．

本稿では，ネットワークトラフィックの解析結果をもとに，ネットワーク利用状況を 3 次元球体状のハイパーオブジェクトを用いて可視化するネットワークトラフィック可視化システムについて述べ，プロトタイプの開発および評価実験の概要について示す．

トラフィックのモデル化および可視化に関する先行研究としては，ネットワークトラフィックの相関関係を利用しインシデントの検知をおこなうもの[1]，IDS のログデータから学習を行い，既存の攻撃パターンに対して特異な攻撃を検出および可視化するもの [2] や，自己組織化マップ[3]を用いることで球体表面上にノードのクラスタリングを行いトポロジーの可視化を行う研究などがある[4, 5]．

本研究ではインシデントに限定しない一般的なユーザのネットワーク利用動向の変化も含めたトラフィックの傾向とその変遷をモデル化し，直観的な 3 次元形状により可視化を行うことを目指す．

2. ネットワークトラフィックの可視化

トラフィック可視化システムの概略図を図 1

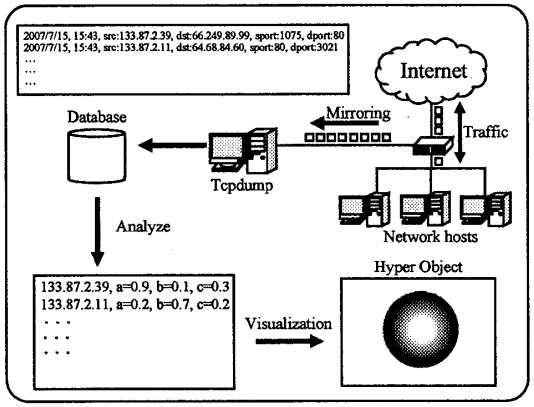

図 1: トラフィック可視化システムの概略図

に示す．本システムはネットワークからの情報の取得，取得された情報からの解析，およびハイパーオブジェクトによる可視化という手順から成っている．

2.1 情報の取得

ネットワークに設置したスイッチやルータでポートミラーリングを行い，パケットのキャプチャを実行する．パケットのヘッダ情報に含まれる送信元，受信先 IP アドレス，リモート，ローカルポート番号，パケット長と時刻などの情報を取得しパケットごとにデータベースに格納する．

2.2 解析

データベースを参照し，一定時間前から現在までの間に送受信されたパケットのリストを作成する． IP アドレスごとに入出力別でのポート利用率，ネットワーク全体でのポート入出力別でのポート利用率を求める．クラスタリング対象として，ネットワーク全体におけるリモートおよびローカルポート利用率のそれぞれ上位 10 位までを選択し，可視化対象のデータセットとする．

球面自己組織化マップを使用し，クラスタリングを行う．マップ上の全てのノードを 20 次元のランダムな値で初期化し，IP アドレスごとのポート利用率をデータセットとして，マップの

Implementation and Evaluation of Network Traffic Rendering by a Hyper-Object

† Tatsumi Ueda, Graduate school of Information Science and Technology, Hokkaido University

‡ Yoshiaki Takai, Information Initiative Center, Hokkaido University

学習を行う．次に，データセットに含まれるポート利用率のデータを順番に取り出して，データとノードの持つ参照ベクトルの間のユークリッド距離を計算し，最も距離の近いノードを中心として，マップ上での位置関係における近傍ノードの参照ベクトルを更新する．更新時にはガウス関数による近傍関数を用いることで，マップ上での距離に応じて更新の影響が現れるようにした．学習後，学習に用いたデータを再度投入し，球面上における IP アドレスの位置を決定する．また，トラフィックの変化に追従しながら球面上の状態を変更する．

2.3 可視化

3 次元のハイパーオブジェクトを用いて，図 2 のように解析結果の出力を行う．

テクスチャ上の各 IP アドレスに対応する位置にマーク(点)を配置する．このとき IP アドレスが属するクラスターごとに別の色で塗り分ける．

データセットの履歴を参照し，各 IP アドレスごとにそれぞれ現在と過去のポート利用率ベクトル間のユークリッド距離を算出する．距離は全 IP アドレス内で正規化され，グレイスケールで表現される頂点テクスチャにおいて，それぞれの IP アドレスと対応づけられる座標を中心とする明るい円によって表現される．

球体のプリミティブに対しテクスチャを貼り付け，頂点テクスチャの明度に応じて球体の表面を隆起させることによって，形状変化による可視化を行う．球面上のマーク(点)の配置と配色を見ることにより，ネットワーク全体におけるトラフィック使用傾向を俯瞰することができる．

また，情報のズーミング機能としてユーザがハイパーオブジェクト上の表面をクリックすることで，クリックした位置に対応する IP アドレスの利用状況やクラスターの情報を見ることができる．

3 実装と評価

情報取得部および解析部のプロトタイプの実装を行った．情報取得部は Gentoo Linux 2007.0 上で，ruby と mysql，解析部は Windows プラットホーム上で，C++/CLI と mysql を用いて実装している．また，可視化部は Windows プラットホーム上で，C++と OpenGL を用いており，現在実装を行っている段階である．

ハイパーオブジェクトによる可視化例を図 3 に示す．ポート使用傾向が変化した IP アドレスに対応する位置が隆起している．

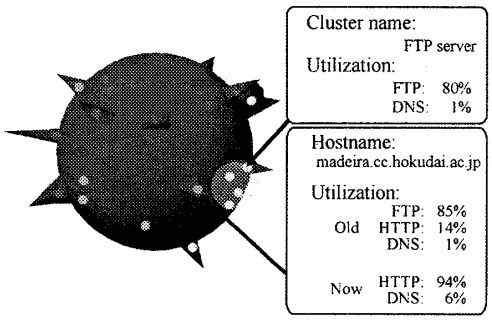

図 2: 可視化イメージ図

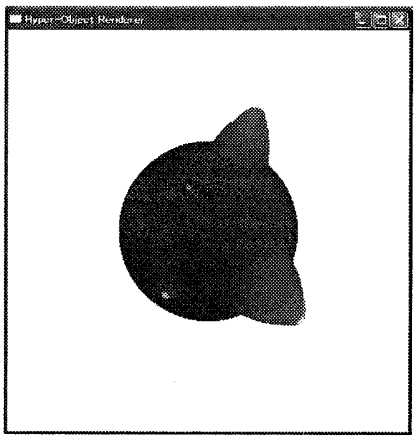

図 3: ハイパーオブジェクトによるトラフィックの可視化例

4 まとめ

ハイパーオブジェクトによるトラフィックの可視化手法と，そのプロトタイプ実装について述べた．自己組織化マップの学習時間が大きなオーバーヘッドになると考えられるため，再整合ノードの探索や参照ベクトルの更新時の処理を改良することにより，動作の高速化を行う．今後，情報取得方法としてパケットキャプチャによる全取得以外に，sFlow[6]によるパケットのサンプリングを利用することで，より大規模なネットワークでの可視化を可能にする．

参考文献

[1] 和泉勇治，廣瀬淳一，角田裕，根本義章，“相関係数発生確率行列を利用したネットワーク状態評価方式”，信学論(B)，vol.J90-B，no.7，pp.660-669，2007

[2] 大庭隼人，宋中錫，高倉弘喜，岡部寿男，“機械学習によるネットワーク IDS ログデータの解析および可視化”，情報処理学会研究報告，2007-QAI-22，pp.31-36，2007

[3] T.Kohonen，“自己組織化マップ”改訂版，シュプリンガー・ジャパン，2005

[4] 徳高平蔵，大北正昭，藤村喜久朗，“自己組織化マップとその応用”，シュプリンガー・ジャパン，2007

[5] Y.Wu and M.Takatsuka，“Visualizing Multivariate Network on the Surface of a Sphere”，Proceedings of Asia-Pacific Symposium on Information Visualization 2006，pp.107-114，2006

[6] sFlow.org，http://www.sflow.org/