社団法人　情報処理学会　研究報告
IPSJ SIG Technical Reports

2007－ＤＳＭ－45（18）
2007／5／11

トラフィックマイニングと可視化によるPeer-to-Peerファイル共有検出支援システムの構築

戸川　聡[†]　金西　計英[††]　矢野　米雄[†††]

† 四国大学経営情報学部
〒771–1192 徳島市応神町古川123–1
†† 徳島大学高度情報化基盤センター
〒770-8506 徳島市南常三島2–1
††† 徳島大学大学院ソシオテクノサイエンス研究部
〒770-8506 徳島市南常三島2–1
E-mail: †doors@shikoku-u.ac.jp, ††marukin@cue.tokushima-u.ac.jp, †††yano@is.tokushima-u.ac.jp

あらまし　Peer-to-Peer（P2P）型ファイル共有コミュニティへの情報漏洩が社会問題になっている．Winnyなどのファイル共有ソフトウエアに感染する暴露ウイルスにより，コンピュータ内部のファイルがファイル共有ネットワークへ漏洩する．これらの状況から大学や企業のキャンパスネットワークでは，P2Pファイル共有ソフトウエアの利用を禁止している．しかし現実にはP2Pファイル共有が行われている場合がある．これを制限する場合，既存のフィルタリング技術では実現が困難であり，結果として管理者はトラフィックを常時監視し，P2Pファイル共有通信の存在を認識しなければならない．本稿では，ネットワーク管理者が行うP2Pファイル共有通信の検出作業を，トラフィックマイニングと可視化により支援するシステムを構築した．そして，実際にP2Pファイル共有プログラムが発するトラフィックを含む全体のトラフィックを可視化し，有効性を検証した．

キーワード　管理者支援，ネットワーク危機管理，トラフィックマイニング，トラフィック可視化，自己組織化マップ

Peer-to-Peer File Sharing Commnication Detection System Using Network Traffic Mining and Traffic Visualization

Satoshi TOGAWA[†], Kazuhide KANENISHI[††], and Yoneo YANO[†††]

† Faculty of Management and Information Science, Shikoku University
123–1 Furukawa, Ojin-cho, Tokushima, 771–1192 Japan
†† Center for Advanced Information Technology, University of Tokushima
2–1 Minami-Josanjima, Tokushima, 771–8506 Japan
††† Institute of Technology and Science, University of Tokushima
2–1 Minami-Josanjima, Tokushima, 771–8506 Japan
E-mail: †doors@shikoku-u.ac.jp, ††marukin@cue.tokushima-u.ac.jp, †††yano@is.tokushima-u.ac.jp

Abstract　In this research, we have proposed the assistance system for peer-to-peer traffic detection. Recently, an illegal file has been exchanged with peer-to-peer file exchange software. These files are extracted from music CD and DVD. Most files do not obtain the copyright person's approval and are open to the public. Neither enterprise nor the Campus Network user of the university must acquire these files from the problem in morality. However, the illegal file is actually acquired via Campus Network. The network administrator should observe the users' peer–to-peer communication. In this paper, first of all, We explain a problem of peer-to-peer file sharing system. Next, we explain the assistance system for peer-to-peer file sharing traffic detection using traffic mining and visualization. Finally, we conclude it.

Key words　Administrator Assistance, Network Risk Managemnt, Traffic Mining, Traffic Visualization, Self-Organizing Maps

1. はじめに

Peer-to-Peer（P2P）形式によるファイル共有が問題になっている．これを実現するソフトウエアとして，従来から WinMX [1] や Winny [2]，BitTorrent [3]，Share などが存在する．最近では Perfect Dark [4] など，新たな改良により，さらに匿名性を高めたファイル共有ソフトウエアも開発されつつある．

インターネット上には，これらファイル共有ソフトウエアによるノードが相互接続されたファイル共有ネットワークが存在する．ファイル共有ネットワークでは，音楽データや映像データ，画像や文書データに至るまで，あらゆる種類のコンテンツ共有されている．また，ファイル共有ネットワークを流通するコンテンツは，著作権法で保護される著作物を抽出し，違法に流通させているものが多い．利用者は，ファイル共有ソフトウエアを用いてファイル共有ネットワークに接続し，そこで共有されるリソースを検索，取得する．

従来，ファイル共有ソフトウエアの利用が好ましくないとされる第一要因は，著作権者の許諾を得ず流通するコンテンツを入手することにあった．これは，ファイル共有ソフトウエアの*利用*が著作権法違反に加担する行為と見なされる可能性を持つからである．この部分に関して，現在でも著作権法違反の懸念が払拭されている訳ではない．

しかし現在，より深刻な問題は，ファイル共有ソフトウエアを使用することによる，深刻な情報漏洩リスクにある．Antinny などの暴露型コンピュータウイルスは，感染したコンピュータ内のリソースをアーカイブしファイル共有ネットワークへ公開する．この結果，使用するコンピュータのローカルストレージに保存されている個人情報や商取引上の機密情報，個人のプライベート情報などを含む機密情報が，ファイル共有ネットワークへ流出する．事実，自衛隊組織の機密情報 [5] や警察組織の操作資料 [6] など，国家レベルでの高度な機密情報までファイル共有ネットワークに流出し，これらの情報漏洩事件は社会問題となっている．

ファイル共有ネットワークへの情報漏洩は，ファイル共有ソフトウエアの使用行為が直接的原因ではない．しかし，暴露ウイルスによる影響が原因とはいえ，ファイル共有ソフトウエアの使用は，個人が管理するあらゆる情報の漏洩リスクを間違いなく高めている．

大学や企業のキャンパスネットワークでは，ファイル共有ネットワークの帯域占拠や著作権法違反への加担を避けるため，ファイル共有ソフトウエア使用を規制する事例が多い．しかし現状は，一部利用者によるキャンパスネットワークからのファイル共有行為が行われている．

これは，著作権法違反を助長するだけでなく，組織からの情報漏洩に対する危機管理の観点からも好ましいものではない．ネットワーク管理者は，利用者のモラルにのみ頼るのではなく，キャンパスネットワークを介して行われる違法行為や機密情報流出の兆候を可能な限り把握し，適切な対策を講じなければならない．

管理者が P2P ファイル共有の制限を試みる場合，パケットフィルタの適用を検討できる．しかし，インターネット上の P2P ノードは，不特定多数かつ可変であるため，IP アドレスによるフィルタリングは困難である．さらに Winny や Share など，自律分散ノード集合で共有ネットワークを構成するものは，標準的な待機ポートを持たないことが多い．この結果，P2P ノードはランダムな TCP ポート番号で接続を待ち受けるため，ポート番号によるフィルタリングも困難である．管理者が，利用者による P2P ファイル共有を制限しようと試みるならば，キャンパスネットワーク内の P2P トラフィックに気づき，個別に対処しなければならない．

そこで本稿では，キャンパスネットワークを対象としたトラフィックマイニングと可視化による P2P ファイル共有通信検出支援システムを提案する．本システムは，キャンパスネットワーク内から行われる P2P ファイル共有通信の検出を支援する．特に，自律分散ノードで構成される Pure 型 P2P ファイル共有の検出支援を目的とする．

全クライアントから送出されるトラフィックを対象に，特定の特徴を持つトラフィックの特性を強調する．その後，トラフィック可視化により全体傾向を俯瞰可能な特徴マップを生成し，管理者に提示する．特徴マップにより異変に気づいた管理者は，対象を絞った調査を行うなど，次段階の作業に取り掛かることができる．

以下本稿では，2. で P2P ファイル共有通信の現状について述べ，3. でこれらファイル共有通信の検出支援モデルについて述べる．4. で本研究で使用するトラフィックモデルの構成と可視化手法を述べ，5. で試作システムの概要を述べ，その後実証実験の概要と考察を述べる．

2. P2P ファイル共有通信の現状

2.1 P2P ファイル共有の通信モデル

P2P ファイル共有の通信モデルは，次の 2 つに分類できる [7]．

Hybrid 型： 図 1 に Hybrid 型通信モデルを示す．これは，ファイル所在情報である索引を保持するインデックスサーバと，実体ファイルを保持するノード群から構成される．あるノードがファイル入手を試みる場合，目的ファイルの所在をインデックスサーバに問い合わせる．インデックスサーバは当該ファイルの所在情報を探索元ノードに返す．探索元ノードはファイルを所有するノードとコネクションを確立し，目的ファイルを入手する．

Pure 型： 図 2 に Pure 型通信モデルを示す．Pure 型は索引情報を保持するインデックスサーバを持たない．ファイルやノードの探索機能はノード自体に実装される．ノードがファイルを探索する場合，近接するノードに探索要求を発行する．これを繰り返すことで探索要求は P2P コミュニティ内に伝搬する．あるノードが要求に合致するファイルを所有していた場合，そのノードはファイル所在情報を返す．所在情報を受信した探索元ノードは，ファイルを所有するノードとコネクションを確立し，実体ファイルを入手する．

Hybrid 型アプリケーション例として WinMX，BitTorrent

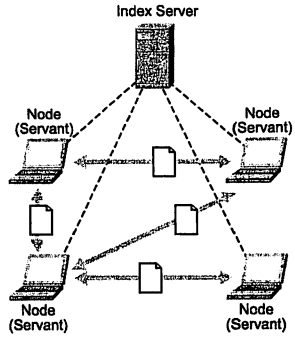

図 1 Hybrid 型 P2P ファイル共有モデル
Fig. 1 Hybrid P2P File Sharing Model

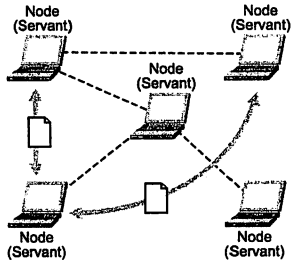

図 2 Pure 型 P2P ファイル共有モデル
Fig. 2 Pure P2P File Sharing Model

表 1 予備実験結果 (Share)
Table 1 Pilot Study for Share

IP パケット送信数	16,009
TCP PUSH フラグ付与件数	5,056
宛先 IP アドレス数	299
宛先 TCP ポート数	254
宛先 DNS 名前解決率	0.1%

があり，Pure 型アプリケーション例として Winny，Share，Perfect Dark が挙げられる．

2.2 Pure 型 P2P ファイル共有の通信特性

Pure 型 P2P ファイル共有の通信特性を明らかにするため，予備実験にてトラフィック解析を行った．本実験の目的は P2P ファイル共有プロトコルを明らかにすることではない．送信元から見た Pure 型ファイル共有トラフィックの表層的振る舞いを明らかにする．

実験機に Share を導入しファイル共有トラフィックを発生させた．比較対象として，人手による Web 閲覧を行った．これは，一般的なキャンパスネットワークにおいて，利用の大部分は Web アクセスが占めると想定できるためである．

実験時間はそれぞれ 15 分間とし，実験機から送信される IP パケットを収集し解析した．表 1 に Share での計測結果を，表 2 に Web 閲覧における計測結果を示す．

まず，Share の IP パケット送信数は 16,009 件であり，Web 閲覧の約 3 倍である．IP パケット送信数に占める TCP PUSH フラグ付与率は，Share が約 31.6&，Web 閲覧が約 8.8%である．Share から見た宛先 TCP ポート番号は 1 番から 65535 番まで一様分布していた．なお，Web 閲覧の宛先 TCP ポート番号は 80 番と 443 番のみであった．これらから，Share は通常の Web 閲覧に比べ大量の IP パケットを送信し，TCP PUSH フラグの付与率が高い．これは，広範囲な宛先 IP アドレス，および一様分布する宛先 TCP ポート番号に対しコネクションを確立すると言える．

表 2 予備実験結果 (Web 閲覧)
Table 2 Pilot Study for Web Browsing

IP パケット送信数	5,322
TCP PUSH フラグ付与件数	469
宛先 IP アドレス数	34
宛先 TCP ポート数	2
宛先 DNS 名前解決率	92.6%

また，宛先ホストに関する名前解決の観点からも解析をおこなった．この結果，Web 閲覧における宛先ホストの名前解決率が 92.6%だったのに対し，Share の場合は 0.1%であることが明らかになった．Web 閲覧時において宛先ホストを指定するとき，ほとんどの場合 FQDN を含む URL が指定される．このため宛先 IP アドレスを入手するため，DNS 名前解決が発生する．一方，P2P ファイル共有の場合，宛先ホストを直接 IP アドレスで指定する場合がほとんどであるため，DNS 名前解決が発生しないと言える．

2.3 フィルタリングによる利用制限の検討

P2P ファイル共有通信制限のため，フィルタリングによる制限を検討する．

Hybrid 型 P2P ファイル共有ではインデックスサーバが単一障害点となる．インデックスサーバへの経路を遮断すれば，理論上容易にリソース検索機能を遮断できる．このため，既存のフィルタリング技術は，Hybrid 型 P2P ファイル共有通信の制限に関しては一定の効果が期待できる．

一方，Pure 型 P2P ファイル共有の制限は困難を伴う．前節で述べたように，Pure 型 P2P ファイル共有の通信では，ランダムかつ広範囲な宛先 IP アドレスかつ宛先 TCP ポートに対しコネクションを生成する．このため既存のフィルタリング技術による利用制限は困難と言える．

2.4 フロー情報解析による異常トラフィック自動検出

NetFlow [8] や sFlow [9] によりフロー情報を収集，解析し，P2P トラフィックなどを異常事象として検出を試みる手法が存在する [10].

フロー情報の収集粒度は，フローレコードのサンプリングレートで決定される．フローを収集しようとするルータ，もしくは L3 スイッチの性能にもよるが，一般的には 1:100 から 1:1000 程度のサンプリングレートを設定する．フロー情報をもれなく収集するには，サンプリングレートを 1:1 に近づければよいが，当該機器への負荷を考慮すれば非現実的である．このためフローによるトラフィック収集は，サンプルを得られるだけであり，解析しようとするトラフィックを完全な状態で収集することは難しい．

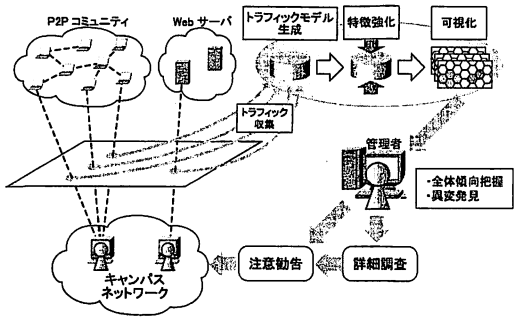

図 3 検出支援モデル
Fig. 3 Assistance Model for P2P Detection

また，現時点において NetFlow，sFlow いずれのフロー収集技術も，キャリアクラスや大規模エンタープライズクラスの L3 スイッチ製品に実装される．このためコスト的要因から，キャンパスネットワークにおける任意の計測点にてフロー収集を実現することは難しい．

このため，フロー情報解析による異常トラフィック検出は一定の効果が期待できるものの，推定や予測の誤差による誤検出は避けられない．

3. P2P ファイル共有のための検出支援

3.1 トラフィックマイニングと可視化による検出支援モデル

管理者による P2P ファイル共有の検出支援を実現するため，図 3 に示す検出支援モデルを定義する．このモデルは「トラフィック生成」「特徴強化」「可視化」機能から構成される．

本稿では，これら一連の流れから実現されるトラフィック特徴強化と可視化の流れをトラフィックマイニングと定義する．

3.1.1 トラフィックモデル生成

キャンパスネットワークからの送信トラフィックを収集しモデル化する．

P2P ファイル共有検出のため把握すべきことは，広範囲なコネクションを持ち，TCP PUSH フラグ付与率が高いトラフィックが存在するか否かと言える．さらに，宛先 IP アドレスが名前解決の結果得られたものではなく，直接得られたものが多ければ，そのトラフィックを発生しているクライアントは P2P ファイル共有を行っている可能性が高い．

このため本研究では，ネットワーク内部から外部に対するコネクション生成状況が把握可能な情報抽出を行う．また，コネクションごとの TCP PUSH フラグ付与状況，および宛先 IP アドレスにおける名前解決の試行状況を抽出する．抽出した特徴量を送信元 IP アドレスを要素とし，単位時間ごとにモデル化する．本研究ではこれをトラフィックモデルと呼ぶ．

3.1.2 特 徴 強 化

トラフィックモデルの特徴量を強化する．分散したコネクション状態を持ち，TCP PUSH フラグが付与された要素に重み付けを行う．さらに宛先 IP アドレスに関する名前解決が試みられていない要素についても重み付けを行う．これにより，P2P ファイル共有を行っている可能性を持つ要素を強調できる．この結果，他のトラフィックに埋没する P2P トラフィックを浮上させ，その存在を管理者に気づかせることができる．

3.1.3 可 視 化

単位時間で集積されたトラフィックモデルを可視化し，管理者に提示する．本研究で扱う監視は，全体傾向把握とその変化による異変発見の支援である．このため管理者への情報提示は一目で全体状況が把握できることが望ましい．単位時間の状況が把握できればトラフィック全体の俯瞰が可能となり，変化の追跡も容易となる．

4. モデル構成と可視化

4.1 トラフィックモデルの構成

トラフィックモデルは単位時間におけるトラフィック特性を定量的に集積しなければならない．このため，モデル生成にはベクトル空間モデル（Vector Space Model:VSM）を適用する．モデルを構成する特徴ベクトルには送信元 IP アドレスが対応し，特徴量として宛先 IP アドレスとその出現量を集積する．各要素の特徴量は，TCP PUSH フラグの出現量，および DNS 名前解決の有無により重みを付け，その特徴を強化する．

特徴ベクトルを $\boldsymbol{x}$，宛先 IP アドレスごとの出現量を a_1〜a_n とすると，特徴ベクトルは次式で表わされる．

$$\boldsymbol{x} = \{a_1, a_2, \ldots, a_n\}$$

トラフィックモデルは，生成されたすべての特徴ベクトルを集めたものである．トラフィックモデルを $\boldsymbol{D}$，特徴ベクトルを $\boldsymbol{x}_1$〜$\boldsymbol{x}_m$ とするとトラフィックモデルは次式で表わされる．

$$\boldsymbol{D} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\}^T$$

これにより，ネットワーク内ノードから送信されるトラフィック特性を，特徴ベクトル $\boldsymbol{x}$ のベクトル集合で表現できる．結果，ノード間類似度を特徴ベクトル間の余弦類似尺度のみで距離関係を算出でき，コネクション特性の類似性をベクトル間類似度で置き換えることができる．

4.2 自己組織化マップによる可視化

生成されたトラフィックモデルは多次元ベクトル集合として構成されている．これは送信元 IP アドレスと宛先 IP アドレスの関係が，多次元空間上の分布として表現できることを意味する．人間は基本的に三次元までの空間は直感的に把握可能だが，それ以上の多次元空間の把握には困難を伴う．

自己組織化マップ (Self-Organizing Map:SOM) は，2 層のニューラルネットワークで構成される教師なし競合学習モデルである．SOM はデータ間の幾何学的構造を可能な限り保った状態で二次元平面に写像する．同時にクラスタリングをおこなう．この結果，管理者は平易な二次元平面にて管理対象組織のトラフィック傾向の俯瞰が可能となる．

5. 試作システムの概要

本章では，実証実験のために構築した試作システムについて述べる．図 4 に試作システムの構成を示す．本システムは「ト

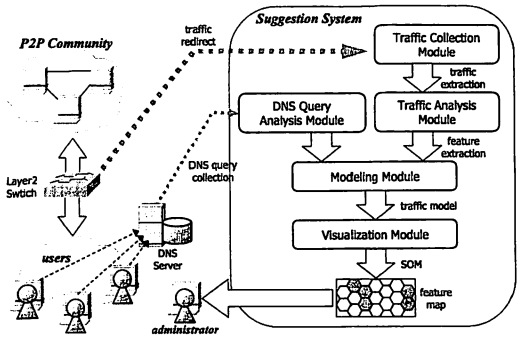

図 4 システム構成

ラフィック収集部」「トラフィック解析部」「DNS クエリ解析部」「モデル化部」「可視化部」から構成される．以下，各部の概要を述べる．

5.1 トラフィック収集部

トラフィック収集部では，監視対象ネットワークが受発信するすべての IP パケットを収集・蓄積する．L2 スイッチのポートミラーリング機能により，獲得する IP パケットを本システムにリダイレクトする．トラフィック収集部は，導入システムの Ethernet カードを promiscuous mode に設定し，リダイレクトされた IP パケットを収集する．

5.2 トラフィック解析部

収集された IP パケット群を解析し，送信元 IP アドレス，送信元ポート番号，宛先 IP アドレス，宛先ポート番号，パケットサイズ，フラグを抽出する．

5.3 DNS クエリ解析部

キャンパスネットワーク内のクライアントが利用する DNS サーバが，インターネット上に散在する DNS サーバへ問い合わせる DNS クエリの結果を収集する．

この結果，キャンパスネットワーク内からインターネット上の DNS に対して，名前解決要求が発行された FQDN とその IP アドレスが対となったクエリ情報が得られる．

5.4 モデル化部

トラフィックモデルを生成する．送信元 IP1 つに対し，宛先 IP・ポート番号数が次元となる多次元ベクトルを生成する．モデル全体では特徴ベクトルが全送信元 IP 数分集積されたベクトル集合となる．ベクトルの各要素には宛先 IP アドレス・宛先ポート番号別にパケット出現回数とパケットサイズを集積する．

また，モデル化部では重み付けを行う．Share などの P2P プログラムやストリーミングプログラムではパケット送信時に PUSH フラグが設定される．このため，PUSH フラグが設定されたパケットは P2P やストリーミングによるトラフィックである可能性が高い．さらに P2P アプリケーションは，接続する相手ノードの IP アドレスを直接指定しコネクションを生成する．このため DNS による名前解決が発生しない．これらの特徴に合致する特徴ベクトルに重み付けし，特徴を強化する．

5.5 可視化部

得られたトラフィックモデルを SOM アルゴリズムを用いて可視化する．SOM アルゴリズムにより抽出されたパケット群が自己組織化され，似た特性を有する特徴ベクトルが集約された特徴マップが生成される．PUSH フラグが設定されたパケットなど，特に特徴を持つ特徴ベクトルはクラスタとして表出する．このため管理者に対し，管理対象ネットワークに発生した特異トラフィックへの気づきを支援できる．

表 3 実験環境

Table 3 Experimental Environment

CPU	Intel Pentium4 3.2GHz
Memory	1 Gbytes
HD	300 Gbytes
OS	Linux (kernel 2.4.18)

表 4 実験データ件数

Table 4 Experimental Data Amount

種別	件数
実験データ件数	1,423,592 件
特徴ベクトル生成数	16,356 件

6. 実験と考察

6.1 実験環境

試作システムに実験データを入力し特徴マップ生成を行った．表 3 に実験環境を示す．

ある組織に協力を得て，2006 年 11 月 20 日にその組織内の端末が受発信したすべての IP パケットを収集し，実験データとした．なお，実験期間中 1 台の端末にて意図的に「Share」を動作させ，適当なデータファイルをダウンロードした．表 4 に実験データ件数および処理過程で生成された特徴ベクトル数を示す．

6.2 考　　察

図 5 に実験で生成した特徴マップを示す．

1 つの特徴マップは 20 × 16 の 320 要素を持つ．それぞれの要素には比較的多く出現した特徴ベクトルが表出する．今回の実験で生成された特徴ベクトル総数は 16,356 件であるため，約 2%の大規模通信が表出することになる．特に広範囲の宛先 IP および宛先ポートに対して通信を行っている送信元 IP のベクトルは自己組織化されクラスタとして表出している．

図 5 は，前述のトラフィックモデルに対して IP アドレスごとの通信量および PUSH フラグ出現率，および，DNS による名前解決情報を重みとして付加した後，可視化した結果を示している．また，処理結果において特徴を持つノードにラベルを表記している．

見かけ上単一 IP アドレスを持つ装置が対外的に複数の宛先 IP アドレスと通信を行う形態を有する NATBOX については，クラスタとして認識されておらず，他の一般ホストと同様に認識することができる．

一方，実験的に発生させた Share に関連するトラフィック，および，実験中にある利用者が使用していた Skype に関連するトラフィックについて，クラスタとして表出していることがわかる．今回生成した特徴マップは，DNS による名前解決情報

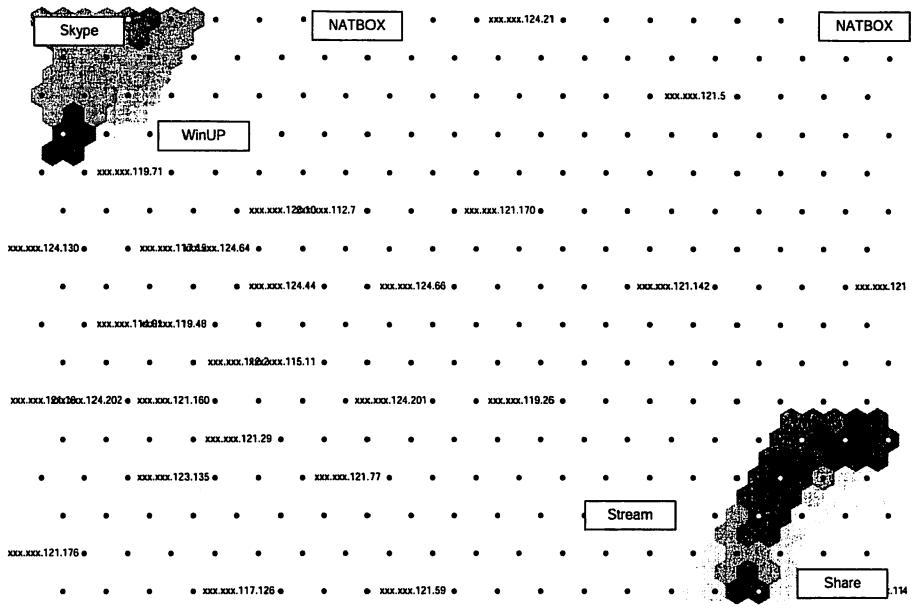

図 5 特徴マップ
Fig. 5 Feature Map

を重みとして加えたものである．P2P クライアントは相手ノードに対し直接 IP アドレスを指定してコネクションを生成するため，DNS による名前解決が発生しない．この特徴を収集し，重み付けに利用することで Share および Skype に関係するトラフィックを，明確なクラスタとして表出させることができた．

Skype は，キャンパスネットワークにおける情報漏洩リスクとなる，ファイル共有通信とは異なる性質のアプリケーションであるが，トラフィック特性としては Pure 型 P2P ファイル共有通信とほぼ同一の特性を持つ．このため，現段階においては特徴マップにおいてクラスタとして表出する可能性が高い．

7. ま と め

本稿では，企業や大学のキャンパスネットワークで行われる P2P ファイル共有通信の問題について述べ，これらの P2P トラフィックを既存のフィルタリング技術で制限することの困難性について述べた．その上でキャンパスネットワーク内から受発信される P2P トラフィックを検出する手法を検討し，管理者がおこなう P2P トラフィック検出支援のために，トラフィックマイニングと可視化による監視支援モデルについて述べた．さらに支援モデルを実現するために必要なトラフィックのモデル化手法について述べ，多次元モデルの認識限界を下げトラフィック傾向の俯瞰を可能にするため行う可視化手法について述べた．また，本提案の有効性を検証するために実装した試作システムについて述べ，実証実験の結果である特徴マップを示し考察をおこなった．

今回の実験においては，P2P ファイル共有トラフィックのみを分離し提示することは実現できなかったものの，Pure 型 P2P トラフィックと非 P2P トラフィックとを比較的明確に分離表示することができた．今後は重み付け手法の改良などにより，特徴マップ上での P2P ファイル共有トラフィックの明確な提示を試みる．

文 献

[1] WinMX Web Site, http://www.winmx.com/
[2] Winny Web Site, http://www.geocities.co.jp/SiliconValley/2949/
[3] BitTorrent Web Site, http://bittorrent.com/
[4] Perfect Dark@ウィキ, http://www21.atwiki.jp/botubotubotubotu/
[5] Internet Watch, “海上自衛隊の「秘」情報が Winny で流出”, http://internet.watch.impress.co.jp/cda/news/2006/02/23/10993.html
[6] 毎日新聞, “愛媛県警：4400 人分の情報流出か ウィニー介して”, http://www.mainichi-msn.co.jp/shakai/jiken/news/20060320k0000m040096000c.html
[7] 石川博, “次世代データベースとデータマイニング”, CQ 出版社, 2005.
[8] NetFlow, http://www.cisco.com/warp/public/732/Tech/nmp/netflow/index.shtml
[9] sFlow.org, http://www.sflow.org/
[10] 藤井聖，中村豊，藤川和利，砂原秀樹，“通信先ホスト数の変化に注目した異常トラヒック自動検出手法の提案と評価”，信学論，Vol.J88-B，No.10，pp.1922-1933，2005.