パケットキャプチャによる画像の収集と表示手法の研究

窪田 幸*　　小池 英樹**　　安村 通晃*

* 慶應義塾大学
大学院政策・メディア研究科
252-8520 神奈川県藤沢市遠藤 5322
{koo, yasumura}@sfc.keio.ac.jp

** 電気通信大学
大学院情報システム学研究科
182-8585 東京都調布市調布ヶ丘 1-5-1
koike@is.uec.ac.jp

あらまし　インターネットの急速な普及とともに，インターネットコンテンツの健全性に関する問題が発生している．仮に不適切な通信を一般ユーザがチェックしようとしても通信技術に関する一定の知識を要するため，現状では困難である．そこで，本稿では通信内容の中でも特に画像に着目し，キャプチャした通信の情報と画像とを関連づけることにより，テキストログよりも通信の概要把握が容易になる表示手法を検討し，プロトタイプシステムを実装した．近年のインターネットコンテンツは，表現を豊かにするために多くの画像が用いられている．このプロトタイプシステムを利用することによって画像のみを抽出して閲覧するだけでも通信内容が大観できるとともに，短時間で通信内容を確認できる可能性が示された．

Research on Image Gathering and Displaying by Packet Capturing

Ko Kubota*　　Hideki Koike**　　Michiaki Yasumura*

* Graduate School of Media and Governance,
Keio University
5322 Endo Fujisawa-Shi Kanagawa
252-8520 Japan
{koo, yasumura}@sfc.keio.ac.jp

** Graduate School of Information Systems,
University of Electro-Communications
1-5-1 Chofugaoka Chofu-Shi Tokyo
182-8585 Japan
koike@is.uec.ac.jp

Abstract　With the rapid spread of internet, it has been a social issue to preserving the contents of internet sanity. In general, average users cannot check the appropriate use of their network(e.g. home network), because they will be required expert knowledge of network. Therefore, we studied the method that general users can review outline of network use and implemented a prototype system for checking inappropriate use of internet. It focuses on the images in the network communications. In recent years, a number of websites are made up of many graphics data as rich expression. Therefore, we can scan the contents of network communications quickly by extractiong images with our system.

1　はじめに

インターネットは生活インフラと呼べるにまで広く普及してきているが，あらゆる情報がどこでも誰でも入手できることから，コンテンツの健全性に関する問題が発生している．特に，近年モバイルを含むインターネットの有害サイトによる青少年の被害が多発しており，社会的に適切な対応をとることが急務となっている．

そこで，内閣官房情報通信技術担当室では『インターネット上の違法・有害情報対策』[1] についてウェブ上で広く情報提供しているほか，また 2009 年 9 月にはインターネットコンテンツ審査監視機構 (I-ROI) [2] がインターネットコンテンツの健全性を認定する活動を開始する予定である (2009 年 8 月 31 日執筆時)．このように安心して利用できるインターネット環境の構築のための組織が設立され，活動が活発になってきている．

一方で，ネットワークユーザが自発的にインターネット利用の健全性を維持する手段は，ポータルサイトや携帯電話事業者，セキュリティベンダー等が提供するフィルタリングサービスやフィルタリングソフトを導入することである．これらの利点は専門的な知識と技術を必要とするフィルタリングを，サービスを利用するだけで導入できるという点である．また，2009 年 3 月に東京都青少年・治安対策本部が都内の小・中学生の子どもに携帯電話等を持たせている保護者を対象にした調査を行い,「フィルタリングに関する実態調査報告書」[3] を公開している．これによるとフィルタリングサービスに現在加入している保護者は 57 %にとどまり,「フィルタリングサービスへの加入が進んでいないと思う理由」でもっとも多いものが,「ネット上の有害情報やフィルタリングについてよくわからない (45 %)」となっている．このことから，フィルタリングに関する知識だけでなく，保護者がインターネットコンテンツの実態についての理解が不十分であることが指摘できる．モバイルを含めたインターネットの健全な利用にはサイトの健全性を第三者が認定するだけでなく，ネットワークユーザ自身のリテラシーの向上も同時に必要であることが挙げられる．

そこで，本稿ではネットワークに関する専門的な知識を持たないネットワークユーザ (以下，一般ユーザと呼ぶ) でも通信内容を確認し概要を把握できる環境を提供することで，インターネットの健全な利用を促すシステムを検討し，その実装を行った．特に，近年のインターネットコンテンツは表現を豊かにするために画像を多く取り入れたものが多い．そこに着目し，画像を中心とした表示手法をとることで，専門的な知識がなくても通信の概要を把握することが可能となる表示手法をとっている．

次章ではシステムの検討内容，3 章ではシステムの構築，第 4 章でシステムの実行例について述べる．そして，第 5 章で考察を行い，第 6 章で本稿をまとめる．

2　システムの検討

システムを検討するにあたり，以下の 2 点を留意点とした．

- 本システムを利用する想定シナリオは，一般家庭や SOHO 等小規模なネットワークにおいて，適切・健全な利用がされているかを一般ユーザが把握することであり，そのためにどのような画像が通信されているか短時間で把握することができるシステムを設計する．
- 本システムを，専門的な知識を極力用いることなく利用できること．

一般ユーザによる利用を前提としたため，テキストログや専門用語以外の情報でネットワークの内容を伝えるためには，画像データが適しているのではないかと我々は考えた．そこで，ネットワーク通信のうち画像データのみを抽出し，表示することによって通信内容の概要を把握できないか，予備実験を行った．予備実験を行うにあたり，画像を収集する手段としてパケットキャプチャを行い，必要な情報を取得した．

2.1　パケットキャプチャツール

パケットキャプチャを行い画像を収集する手段として，Driftnet [4] を用いた．Driftnet は Chris Lightfoot により開発されたイメージキャプチャツールである．これは UNIX システム上で動作し，TCP ストリームをキャプチャする．取得できる画像の種類は JPEG と GIF のみであるが，オプションにより MPEG も可能である．また，これと同種のツールとして EtherPEG [5] があるが，これは PowerPC 上の MacOS でのみ動作する．いずれのツールもリアルタイムで

キャプチャした画像をディスプレイ上に表示するが，アプリケーション終了と同時にキャプチャした画像も削除される．また，過去の画像にさかのぼって閲覧することができず，再現性がない．今回は Driftnet を MacOSX(Ver.10.5.8) に導入してキャプチャを行った．Driftnet を実行することにより得られる情報は，画像の送信元 IP アドレス，送信先 IP アドレスと，パケット中の画像である．画像情報としてはタイムスタンプ，サイズ，pixel 値が取得できる．

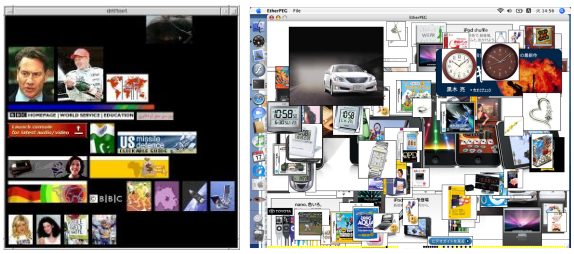

図 1: キャプチャの様子 (左:Driftnet, 右:EtherPEG)

2.2 予備実験

筆者個人の通信をキャプチャし，2009 年 5 月 27 日に 17:08 から 18:08 までの 1 時間，予備実験を実施した．その結果得られた画像の一部を図 2 に示す．

また，この 1 時間で取得した画像数は JPEG,GIF あわせて 1339 であった．これらの結果を精査したところ，予備実験から以下の 3 点の知見が得られた．

1. 1 時間の 1 ユーザのキャプチャだけでも 1000 を超える大量の画像を取得する．
2. 写真やイラストなどとは異なり，単色塗りつぶしの画像や，矢印，丸印など通信内容を推測しづらい画像 (以下，無意味画像と呼ぶ) が多い．
3. 画像のみを閲覧するだけでも，Web サイトのタイトル画像や写真，イラスト等からどのようなサイトを見ていたか推測が可能である．

図 2: キャプチャした画像の例．塗りつぶしの画像，バナーなど様々．

このことから，多くの画像を短時間で閲覧するためにすばやく見ることのできるインタフェースが重要であるとともに，本システムの目的に沿わない無意味画像については，事前にフィルタリングすることによって，閲覧数そのものを削減することが必要であることがわかった．

2.3 画像のフィルタリング

パケットキャプチャにより，無意味画像も多く含まれることを述べたが，これらの画像の特徴としては，以下の 2 点が挙げられる．

1. 矢印や丸印，単色塗りつぶしの画像は pixel 数が小さい
2. ウェブデザインのためのフレームとしての画像はアスペクト比が極端である．

これをふまえ，画像の pixel サイズを取得し，pixel 値の小さいものとアスペクト比が極端に大きなものをフィルタリングした．pixel 値は大きい程フィルタリングの対象となる画像が多い傾向にあるが，収集した画像のうち無意味画像以外のものまでフィルタしすぎてしまわないよう配慮するため，一辺が 10,12,14,16,18,20pixel という条件でそれぞれフィルタリング実験を行い，その結果 16pixel が最も適当であると判断

した．また，アスペクト比も同様に複数条件で実験を行って最適値を判断し，10:1 よりも大きいものをフィルタリングすることとした．その結果を，表 1 に示す．

表 1: フィルタリングの結果

pixel 値とアスペクト比によるフィルタリング	画像の枚数
フィルタリング前	1339
フィルタリング後	865

pixel 値とアスペクト比の基準によりフィルタリング対象となった画像数は 474 となり，約 35 %の無意味画像を取り除くことができた．

2.4 画像の表示手法

本システムは，一般家庭や小規模なオフィス等での利用を想定しており，通信が適切・健全に利用されているかを把握するため，通信されている画像を短時間に把握できることをインタフェース設計上での要件としている．そこで，画像と送信先 IP アドレス (各ユーザ端末の IP アドレス) とを Driftnet のログから関連づけている．画像のレイアウトは Y 軸に IP アドレス，X 軸に時間軸をおいて，各送信先 IP アドレスごとの画像を左からタイムスタンプの古い順に並べる．さらに，IP アドレスの左側にも画像表示領域を設け，10FPS の速さで IP アドレスごとのすべての画像を高速に見ることができるようになっている．また，表示全体を Y 軸を中心として左に 45 度回転させることにより，表示全体としては左端が手前に迫るように表示され，閲覧者の視線がまず左に向くように促し，左から右へ視線を動かすように図っている．これによって，まず初めに各 IP アドレスごとの画像を高速に閲覧し，不適切な画像があった場合 IP アドレスを確認，さらに右にある表示領域で不適切な画像の前後も含めて確認することで，通信のコンテクストを把握するようにした．

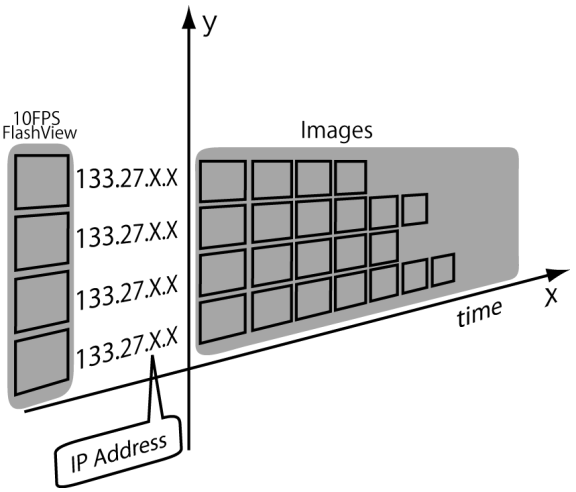

図 3: 画像のレイアウト

3 システムの構築

システムは大きく分けて「画像収集部」と「画像表示部」に大別して構築した．

3.1 画像収集部

画像を収集する手段として予備実験で用いた Driftnet を導入している．まず画像収集部でパケットキャプチャによる画像の収集と前述した簡易なフィルタリングを行い，さらに送信先 IP アドレスごとにディレクトリを作成し，対応する画像を IP アドレスごとにソートする．処理は perl により実現した．

3.2 画像表示部

画像を表示する手段として，Quartz Composer を使用した．Quartz Composer は MacOS X 10.4 から開発環境の一つとして標準で付属しているものであり，ビジュアルプログラミングの一種である．検討したインタフェースデザインを構築し，X 軸方向にはマウススクロール，Y 軸方向にはポインタの Y 座標と同期させることで，スクロール表示ができる．また，画像の高速表示部はキーボード操作によりスタート，ストップ，リセット操作を可能にしている．

4　システムの実行例

本システムを実行し，収集した画像の表示を試みた．その条件を表 2 に示す．

表 2: システムの実行条件と画像収集数

実行環境	使用人数 10 人未満の小規模ネットワーク
実施日時	2009 年 7 月 29 日 15:40〜16:40
取得画像数	1605
フィルタ後画像数	1234

また，実行の表示例を図 4 に示す．

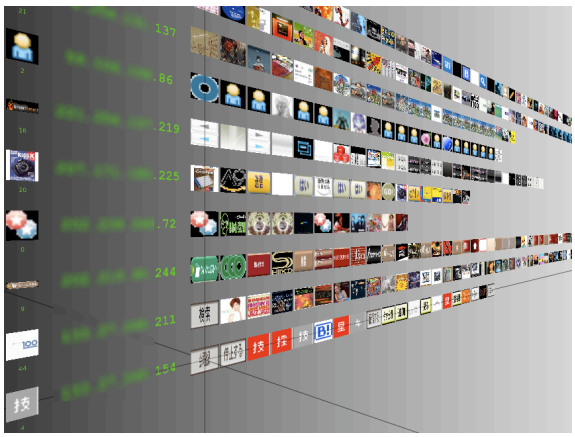

図 4: システムを実行した様子

なお，今後本格的に長時間運用し，データを収集する予定である．

5　考察

本システムの利点と問題点について考察する．

5.1　利点

本システムの利点として以下の 2 点が挙げられる．

5.1.1　閲覧の容易性

パケットキャプチャにより画像のみを抽出して見るという手法で，通信内容の健全性を把握する上で概要を推測できることが確認できた．テキストログのように，読み進めるもしくは解析して理解するという手法より閲覧の容易性は優位であるといえる．また，IP アドレスと画像とを同一の Y 軸上に表示するというシンプルなレイアウトをとっているため，事前の説明なくとも直感的に理解することが可能である．そして，不健全・不適切な画像があった場合は，高速表示，一列表示のいずれにしても同種の画像が連続することが多いため，その存在を確認することは容易である．

5.1.2　必要な情報の選択とフィルタリング

通常，通信内容を見ようとすれば，無味乾燥なテキストベースのログを解析することになるが，本システムでは IP アドレス，画像，時間に情報を絞っているため，通信を把握する上での負荷を大幅に軽減しているといえる．また，本システムの目的にはそわない無意味画像を自動でフィルタリングすることにより，閲覧する画像の絶対数を削減することができた．これも本システム利用者への負荷の軽減と所要時間の削減につながっている．さらに，表示内容を画像に特化したことで，パスワードなど秘匿性の高いものを取得・閲覧できてしまうこともないため，一般的な監視ツールのように必要以上に他人の通信を監視する状況も生起しないと考えられる．

5.2　問題点・課題

本システムの現状の問題点，改善が必要となる点は次の 2 点がある．

5.2.1　表示手法の改善

大量の情報をより早く正確に人間が認識するために，情報視覚化の研究が数多く行われて行きている．情報視覚化の研究では，着目しているミクロな視点の情報と全体の構造といったマクロな視点の情報を同時に見せる技術が研究されており，Focus+context 手法と呼ばれている．代表的なものに *Generalized Fisheye Views* [6]

や *Fractal Views* [7] などがあり，これらは大量の情報を効率的に見るための視覚化手法として先駆的なものである．

本システムでは，1時間という短い間隔のキャプチャしか行っていないが，実際の利用においては数時間もしくは数十時間という時間の履歴を閲覧する機会のほうが多いと予想され，扱う情報量が膨大になることが考えられる．現段階のインタフェースでは，より長いタイムラインの情報表示と，ある一点における情報表示の混在が難しい．逆に，Focus+context 手法を取り入れることによって，IP アドレス別の把握に加えて，時間帯別の傾向の把握といった，新たな視点での活用が考えられる．

5.2.2 対話性

現行のインタフェースでは一画面で複数の IP アドレスの情報が閲覧できるが，例えば，本システムの利用者が特定の IP アドレスに絞って情報を精査したり，時間帯を指定して表示するなどの対話性に乏しい．先述の Focus+context 手法にも関連することであるが，全体の情報から個別の情報へ絞り込みを行ったり，また個別の情報と全体との関連性を調べるという利用が想定される．そのような本システム利用者の要求に即座に応えることができるような対話性をもたせることが，今後の課題である．

6 おわりに

本稿では，インターネット利用の健全性をいかに維持していくかという問題に対し，一般ユーザ自らが通信内容を把握できる環境を提供することで健全な利用が促されるという提案を行った．そこで，画像に特化することにより通信内容が把握しやすくなるシステムを検討し，実装を行った．パケットキャプチャにより得られた画像は数も多く，通信内容を推測しづらい無意味画像が多くあることからフィルタリングを行い，画像を送信先 IP アドレスごとに表示するインタフェースを試作した．これにより収集した画像を通信ログのように閲覧することが可能となり，その容易性を確認した．そして，本システムを導入することによるインターネットの健全な利用に対する効果について議論した．このようなシステムを導入することで最終的に期待できることは，インターネットコンテンツの健全性に対する関心を向上させることである．また，通信内容を一般ユーザでも見ることができるようにすることで，ネットワークの公共性が増し，ネットワークユーザの自発的で健全なインターネットの利用を促すことが期待できる．

謝辞

本研究を行うにあたり，産業技術総合研究所高田哲司氏，慶應義塾大学 SFC 研究所渡邊恵太氏にアドバイスをいただいたことを感謝いたします．

参考文献

[1] 内閣官房『インターネット上の違法・有害情報対策』. http://www.it-anshin.go.jp

[2] インターネットコンテンツ審査監視機構 (I-ROI). http://www.i-roi.jp

[3] 東京都　青少年・治安対策本部総合対策部青少年課．「フィルタリングに関する実態調査」の結果について http://www.metro.tokyo.jp/INET/CHOUSA/2009/03/60j3i100.htm

[4] Driftnet. http://www.ex-parrot.com/~chris/driftnet/

[5] EtherPEG. http://www.etherpeg.org/

[6] George W. Furnas. *Generalized Fisheye Views*. Human Factors in Computing Systems CHI ‘ 86 Conference Proceedings, 16-23. 1986.

[7] Hideki Koike. *Fractal Views: A Fractal-Based Method for Controlling Information Display*. ACM Transaction on Information Systems, Vol. 13, No. 3, July, pp.305-323, ACM, 1995.