人間による Honeypot の攻撃元ログ調査を支援する User Interface の提案

高田 哲司 †　　小池 英樹 ‡

† 産業技術総合研究所
135-0064 東京都江東区青海 2-41-6
zetaka @ computer . org

‡ 電気通信大学大学院
182-8588 東京都調布市調布が丘 1-5-1
koike @ vogue.is .uec .ac .jp

あらまし　本論文では，複数台の Honeypot により収集された攻撃元ログを対象とし，人間がそのログ情報を調査するためにはどのような User Interface(UI) が必要かについて検討した．我々はその答えとして，大域情報としての傾向情報の提供，情報視覚化による認知支援，段階的な調査方法の提供，そして対話機能による調査作業の簡易化の 4 点が必要であるとし，その検討を基に，プロトタイプシステム MWLT Browser を実装した．またこのプロトタイプシステムを用いた Honeypot ログの調査事例について述べ，提案する UI がログ調査を支援しうることを明確にした．

A Proposal of a User Interface that Assists Human Inspectors in Understanding Attack Source Log in Honeypots

Tetsuji TAKADA †　　Hideki KOIKE ‡

†National Institute of Advanced Industrial Science and Technology
2-41-6, Aomi, Koto-ku, Tokyo, 135-0064, Japan
zetaka @ computer .org
‡Graduate School of Information Systems, University of Electro-Communications
1-5-1, Chofugaoka, Chofu, Tokyo, 182-8588, Japan
koike @ vogue.is.uec.ac.jp

Abstract　In this paper, we consider that what kind of requirements are needed for a user interface for assisting a human inspector in investigating log information from honeypots. And we reach a conclusion that there are four requirements: 1) A system extracts a trend information from a log as a global view. 2) A system represents data using a graph or another visual representation for reducing recognition load. 3) A system provides users with an incremental investigation method from a global view to a detail view. 4) A system provides users with an interactive function to support trial and error inspection. Based on the consideration, we developed a prototype log browsing system. We also explain a log inspection example using the prototype system. Therefore, it makes clear that the prototype system has a potential to assist in log inspection by a human.

1　はじめに

本論文では，Cyber Clean Center[1] から提供されたログのうち，複数台の Honeypot が収集した Malware の活動に関するログ情報「研究用データセット CCC DATAset 2008 の攻撃元データ」を対象とし，そのログ情報を人間が調

査する際にどのような User Interface が望ましいかについて考察を行った．またその考察を基にプロトタイプシステムの実装を行った．

実装したシステムでは，ログ情報から 6 種類の傾向情報を抽出し、それを折れ線グラフと増減インジケータと呼ぶ視覚的表示手法により提示する．これによりログ情報の概要把握を可能にし、かつ要調査と推測される箇所の発見を支援する．また必要に応じて段階的に詳細情報の提示を可能にし，最終的には CCC より提供されたログ情報そのものの閲覧も可能にするにより，ログ情報の調査を User Interface 内でシームレスに実行可能にする．

以降 2 章では，Honeypot が生成するログ情報の調査に必要な User Interface(UI) の要件について行った検討事項について述べ，3 章では，2 章での考察を基に実装したプロトタイプシステム MWLT Browser を紹介する．4 章では MWLT Browser による調査事例について説明し，5 章では考察として残されている課題について述べる．

2 Honeypot のログ調査に関する検討

Honeypot のログ情報は，それを人間が調査する必要がある．それには 2 つの理由がある．1 つは Honeypot はあくまで情報収集の一手段だという点である．Honeypot をセキュリティ対策として活用するためには，人間が Honeypot のログを把握し，そこからなんらかの有益な情報を得て，今後の対策に生かせるよう適切な行動を起こす必要がある．これは Information Security Management System (ISO27001)[2] に明記されている PDCA サイクル[1]を継続実施することに該当する．Honeypot 自体はあくまで “Do” に該当するだけであり，“Check, Action” は人間が介在しなければならないのが現状である．

しかしログ情報は膨大な量の文字記録であり，人間がこれらを一つ一つ調査し，そこから今後の対策に有益と推測される兆候を発見するのは困難である．また Honeypot のログは実運用されているサーバのログとは異なり，全てのログ記録がなんらかの不正行為に基づく情報である．したがって Honeypot のログ情報からセキュリティ対策上有益と推測される兆候を発見するにはなんらかの仕組みが必要となる．そこで我々は，人間による Honeypot のログ調査を支援するための要件として，以下の 4 点を挙げる．

1 つめは，傾向情報の抽出と提示である．これには次の 2 つの目的がある．1 つは大量のログ情報に内包されている兆候の概要把握を支援するためであり，もう 1 つは今後の対策に有益と推測される兆候の発見を支援することである．そのためログ調査を支援するシステムには，ログ情報から傾向情報を抽出する仕組みを持つ必要がある．

2 つめは情報視覚化による提示情報の認知支援である．傾向情報の提示はグラフ表示を利用することになるが，グラフからなんらかの兆候を読み取る作業にも一定の負担がかかり，また読みあやまる可能性もある．そこでグラフ表示認知負担軽減と付加情報の提示という 2 つの役割を実現するため，なんらかの形で視覚的な表現による情報提示を採用する．

これは不正侵入検知の分野でも言われている手法である．完全な不正検知手法が存在しない現状では，それを補完する手法として人間による監査が必要である．しかし前述の通り，何の支援もなしに人間がログ監査を行うのは困難であり，その支援手法として情報視覚化が有望視されている [3, 4]．我々は Honeypot のログ情報調査支援のため，情報視覚化技術を応用する．

3 つめは段階的な調査手段の提供である．ログ情報の調査における最大の問題点は，最も低レベルの情報である文字記録のログ情報を調査しなければならない点にある．しかし人間による調査では，まずはじめに大局的な見地から調査を行い，その中からさらに調査が必要な箇所の発見を試みる．そして発見された該当箇所についてより詳細な調査を行う，という風に段階的に調査を繰り返し，最終的にそれまでの調査を裏付けるべく，最も低レベルだが全ての情報を保持している文字記録のログ情報を調査すべきである．これを UI レベルで実現し，段階的な調査を可能にすべきと考える．

最後の要件は，対話機能による調査作業の支

[1] Plan, Do, Check and Action の略

援である．ログ調査用 User Interface には対話性があることが必須条件であると考える．理由は，調査作業とは断定的な判断により実施できる作業ではなく，本質的に試行錯誤が伴うと考えるからである．よってそのような特徴を持つ作業を支援するためには，可能な限り簡単に異なる条件や別の箇所の調査を実施できることが望ましい．よってログ調査用 UI には，対話機能により異なる条件や箇所の調査が簡単に実施可能であることが望まれる．

3 プロトタイプシステム:MWLT Browser

これまでの検討を基に，プロトタイプシステム MTLW Browser(MalWare Log Trend Browser) を実装した．本章ではこのシステムについて説明する．

3.1 ログ情報の前処理

我々は Honeypot のログ情報把握を可能にするため，大域情報として傾向情報を提供する．本節ではそのために実施したログ情報に対する前処理について述べる．

まずはじめに提供されたログ情報の属性情報を次の 4 つのグループに分類した: 日付時刻 (日付時刻)，攻撃元情報 (IP アドレス，国名)，振る舞い情報 (Protocol，Port 番号, 通信方向)，そして Malware 情報 (検体名称，Hash 値，File 名) である．なお攻撃元情報内の国名情報は，MaxMind 社の GeoIP[5] を利用して取得した．

この上で，傾向情報として次の 6 つの情報を日別に集約して抽出した．

- ログメッセージ数
- ユニークな IP アドレス数
- ユニークな国名数
- ユニークな検体名称数
- ユニークな hash 値数
- ユニークな file 名数

日別のログメッセージ数は，Malware の活動傾向を知るための基本的な情報である．ユニークな IP アドレス数と国名数は「攻撃元」という点でその傾向を知るために抽出した情報であり，ユニークな検体名称数，hash 値数，file 名数は「Malware」に関して，その傾向を知るために抽出した情報である．ログメッセージ数を除く 5 種類の情報を傾向情報として抽出した理由は，日別のログメッセージ数だけでは把握が困難だが，要調査と推測される兆候の発見に有益であろうと考えたためである．

3.2 大域情報の画面表示と対話機能

図 1 は，MWLT Browser における大域情報表示画面の表示例である．この画面では大きく 3 つのグラフ表示領域 (図中 1,2,3) があり，画面上部左側には日付表示部 (図中 4)，画面上部右側には 6 つの色付き四角形による情報提示部 (図中 6) がある．

3 つのグラフ表示部に表示されている情報は前節で述べた日別の傾向情報を次のように割り当てて表示している．上段の領域にはログメッセージ数を，中段には攻撃元に関する情報であるユニークな IP アドレスと国名数の 2 種類の傾向情報を，そして下段には Malware に関する情報であるユニークな検体名称数，hash 値数，file 名数の 3 種類の情報をそれぞれ折れ線グラフで表示している．各グラフの横軸は日付であり，提供された 6ヵ月分のログ情報をすべて表示している．縦軸は各日付別の値である．ただし縦軸の最小値は 0，最大値はそれぞれの傾向情報における最大値と定義してグラフ表示を行った．つまり縦軸の目盛割り当ては個々の傾向情報毎に異なっている．縦軸の目盛を複数の傾向情報間で統一しない理由は，各値の増減傾向を見るのが目的であり，個々の値を比較するのが目的ではないためである．

なお各グラフ表示領域には横軸方向に沿って色つきの四角形により各グラフの増減状況を視覚的に提示する．これは Timeplot プロジェクト [6] を参考に考案した表示法であり，我々はこれを増減インジケータと呼ぶ．これにより各傾向情報の増減状況を離散化し，かつ視覚化表現の応用により直感的に認知可能にする．増減インジケータの色と傾向情報の増減状況との関係は次の通りである．

- 赤: 値の急な増加を示す．該当日の数値が，前日の数値と比較し該当日の値の 20%以上増加したことを示す

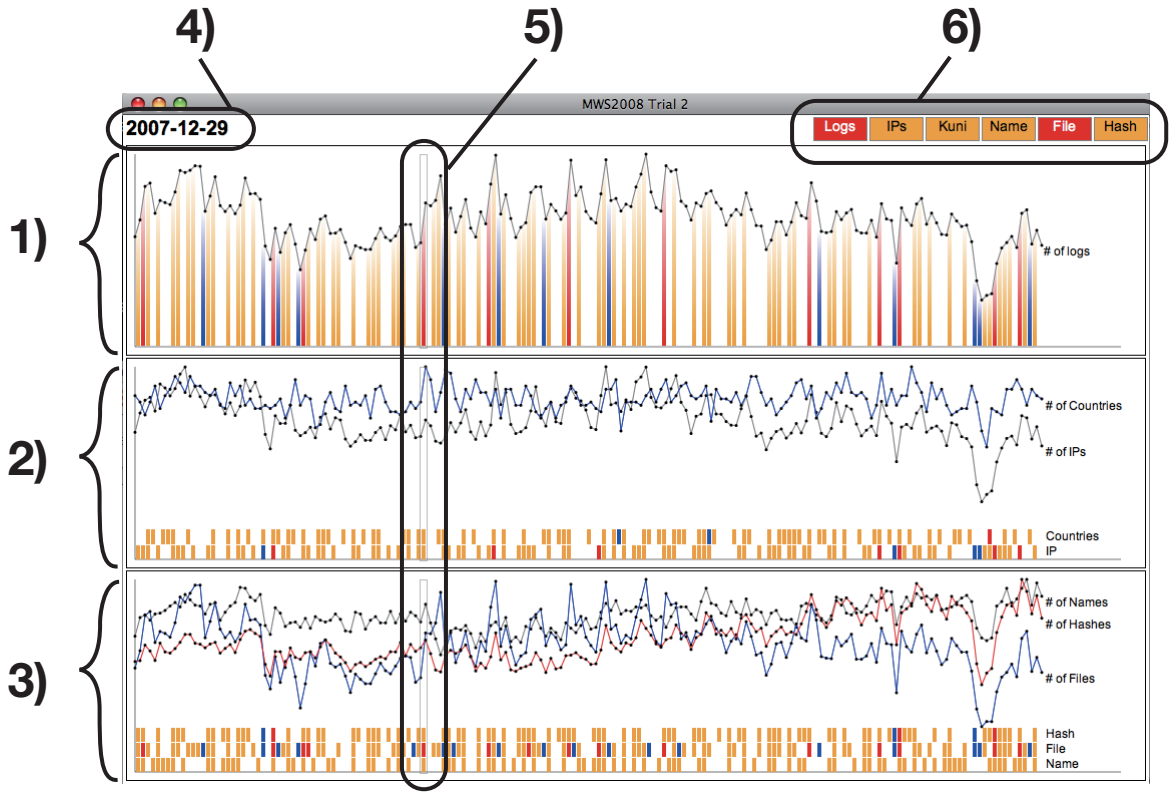

図 1: MWLT Browser 大域情報表示画面

- オレンジ: 値の増加を示す．その日の数値が前日よりも増加していることを示す
- 青: 値の急な減少を示す．該当日の数値が，前日の数値と比較して前日の値の 20%以上減少したことを示す

これにより 6 種類の傾向情報の中から上記の増減状況を示している傾向情報種別とその発生日時の特定をグラフ表示のみの場合よりも容易にする．

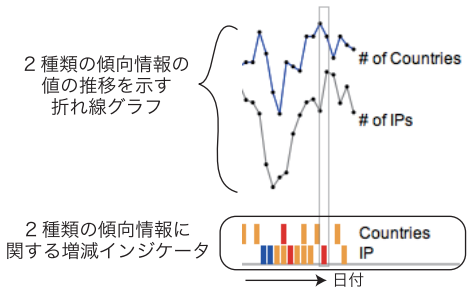

図 2: 複数の傾向情報と増減インジケータ

なお図 1 の中，下段のように複数の傾向情報が 1 つのグラフ内に描画されている場合は，グラフ下部に複数の増減インジケータを層状に提示する (図 2)．どの層がどの傾向情報の増減状況を示すかは，画面右側にラベルで示している．

なお対話機能としては，特定の日付のデータに着目できるようカーソルが用意されており，マウスで操作可能になっている (図 1 の 5)．またこのカーソルが存在する日付を画面上部左側 (図 1 の 4) に，その日の各傾向情報の増減状況を画面上部右側 (図 1 の 6) に表示する仕組みとなっている．また特定の日付にカーソルをあわせた上でマウスをクリックすると，選択した日付のログ情報をより詳細に提示する詳細情報画面に遷移する．

3.3 詳細情報の画面表示と対話機能

詳細情報画面について説明する．本論文では紙面の都合により Malware に関する詳細情報画面についてのみ説明する．詳細情報画面は，大域情報表示画面で注目する日付にカーソルを移動し，各グラフを click することで画面が遷移する．下段のグラフ内で click をした場合は，Malware に関する詳細情報表示画面に移行し，

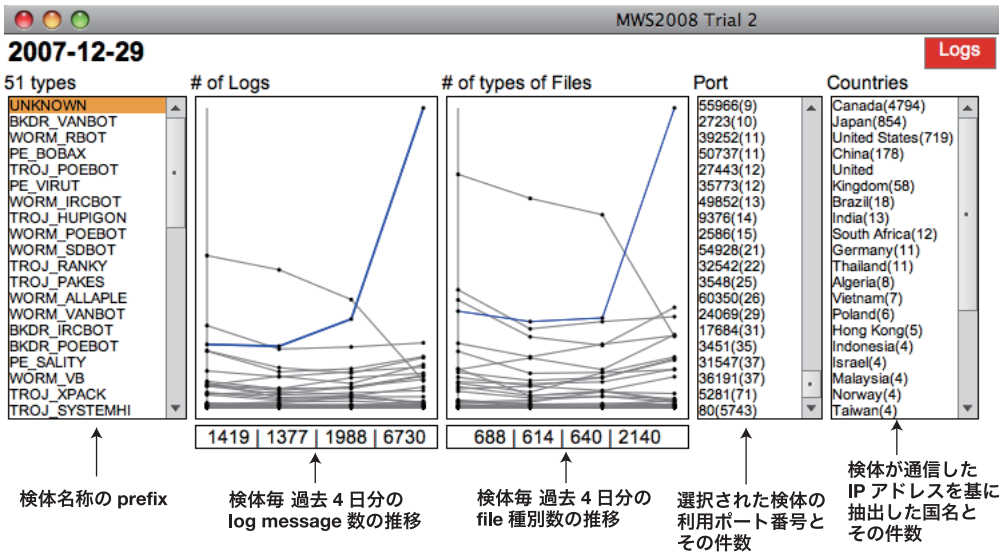

図 3: Malware 情報に関する詳細情報画面

中段のグラフ内の click した場合は，攻撃元に関する詳細情報画面に移行する．図 3 は，Malware 情報に関する詳細情報画面の表示例である．2 つのグラフと 3 つの表示領域が存在するが，そのそれぞれについて説明する．

1 番左の欄は，選択された日付に活動が記録された Malware 検体の名称をリスト表示する．なお検体名は，検体名の prefix 部で集約されている．例えば “WORM_RBOT.GIH” も “WORM_RBOT.CWO” も “WORM_RBOT” として表示される．次に続く 2 つのグラフは，選択日から過去 4 日分の日別の各検体別の数値の推移を折れ線グラフで示している．左側のグラフは検体別ログメッセージ数の推移であり，右側のグラフは検体別の file 名種類数の推移を示している．なお検体名称リストで選択された検体のグラフは青色でハイライトされ，その値はグラフ下部に具体的な数値で表示される．

左から 4 つめと 5 つの表示領域は，最も左側の検体名称リストで選択された検体の振る舞いと攻撃元に関する情報を提供している．図 3 では，検体名称不明を示す “UNKNOWN” が検体名称リスト内で選択されており，該当する検体群が通信に使用しているポート番号と，通信先の国名をリスト形式で提示している．またリスト内の個々の情報には数値が付与されているが，それは選択日付における該当情報の発生件数 (ログメッセージ数) である．

なお最終的には，集約された情報ではなく CCC より提供されたログ情報に MWLT Browser から直接アクセス可能にする予定である．すべてのログを文字表記で閲覧するのは困難なため，これまでの調査で注目した条件を検索条件として設定し，その条件に該当するログメッセージを MWLT Browser 内で閲覧可能にする．これにより概要情報から文字記録によるログ情報まで Browser 内でシームレスに調査可能とする．

4 調査例

MWLT Browser を用いた honeypot ログの調査事例について一例を挙げる．

図 1 にあるように大域情報表示画面でログメッセージ数の推移に注目していたところ，ログメッセージ数が急に増加している日 “2007-12-29” を発見．該当日にカーソルをあてて画面右上に注目すると図 1 にあるように，提示している 6 種類の傾向情報全てが増加傾向を示しており，中でも File 名の種別数も急に増加していることがわかる．そこで Malware に関する詳細情報画面に遷移し，その詳細調査を試みた．

図 3 の左から 3 つめの表示領域にある右側のグラフに注目すると，2008-12-28 から 2008-12-29 にかけて急に File 種別数が増加している Mal-

ware 検体が存在することがわかる．このグラフに対応する検体を探すと，それは検体名 “UNKNOWN” の Malware であった．なおその検体の特徴を見てみると，ポート番号 80 を使用した通信が圧倒的多数であり，かつその通信先の国名は Canada が大多数であることがわかった．検体名称が “UNKNOWN” であるため未知の Malware が活動を開始した可能性があり，適切な対応を取るべき状況であることがわかる．

5 考察

Honeypot のログを人間が調査するにあたり，残されていると考える課題について 3 点ほど述べる．

1 つは膨大な量のログ情報からセキュリティ対策上有用だと推測される兆候を発見する手法の探求についてである．今回の提案では傾向情報を用いたが，これだけに限らず他の方法もあると考える．さらなる考察と，実際にログ調査を実施している方々からの聞き込み調査が必要と考える．

2 つめは Honeypot のログに特化した機能も必要だと考える．例えば検体名称 “UNKNOWN” の動向把握を直接支援する仕組みや，直近数ヵ月内で記録されていない値が発見された場合の通知，そして今回は活用できなかった振る舞い情報 (Port 番号，Protocol そして通信方向) の把握支援も必要であろうと考える．

最後は監視を目的とした User Interface の検討の必要性である．今回のシステムは post-hoc なログ調査を対象としているが，その一方で監視という観点から User Interface を検討する必要もあると考える．一定期間毎の情報を集約して視覚的に提示し，それを過去複数期間分を同一画面に提示することで，時間の経過に伴う Malware の振る舞いの差異発見を支援するなど，今回考察した要件とは別の要件が必要になると考える．

6 おわりに

本論文では，Honeypot のログ情報を人間が調査することを前提にし，その作業を支援する User Interface の要件に関して検討を行った．またその検討に基づき，CCC[1] より提供された CCC Dataset 2008 の攻撃元データを対象としたログ調査支援システム MWLT Browser をプロトタイプとして実装した．これにより 2 章で述べた要件を満たす User Interface の実例を例示でき，またそのプロトタイプシステムを用いた調査事例を示すことで有効性の一端を明確にできたと考える．

このような UI の提供により，はじめから最も低レベルの文字記録によるログ情報の調査を強いることなく，大域的な視点による概要情報の把握から調査をはじめ，そこから要調査と推測される箇所の発見と，該当箇所の詳細調査を段階的に行うことを可能にした．またグラフ表示や視覚的表現に基づく情報提示により，ログ情報の概要と有益と推測される兆候の把握，そして認知負荷の軽減を実現する．そして UI が持つ対話機能により試行錯誤によるログ情報の調査実施を支援する．このようなシステムの提供により，Honeypot のログ情報が人間によって調査され，そこから得られた情報が今後のセキュリティ対策に活用されるようになる: すなわち PDCA サイクルにおける Check の一助となるものと期待する．

参考文献

[1] 総務省・経済産業省 連携プロジェクト Cyber Clean Center, https://www.ccc.go.jp/, accessed 2008-09-10

[2] ISO 27001, Standards Direct International Standards and Documentation, http://17799.standardsdirect.org/iso27001.htm, accessed 2008-09-10

[3] Conti, G.: Security Data Visualization: Graphical Techniques for Network Analysis, No Strach Pr, ISBN: 1593271433, (2007).

[4] Conti, G. and Abdullah, K.: Passive Visual Fingerprinting of Network Attack Tools, Proc. of the VizSec/DMSEC ’04, pp.45-54, (2004).

[5] MaxMind - GeoIP IP Intelligent Solution, http://www.maxmind.com/app/ip-locate, accessed 2008-09-10

[6] SIMILE Timeplot project, http://simile.mit.edu/timeplot/, accessed 2008-09-10

[7] Ragnar, B., Stefan, S. and Silvia, M.: Connecting time-oriented data and information to a coherent interactive visualization, *Conf. on Human Factors in Computing Systems(CHI)*, pp.105-112, (2004).