モバイルエージェントを用いた分散型インターネット観測システムの提案

葛野弘樹[†] 中井優志[††] 渡邊集[††]
川原卓也[††] 加藤貴司[††]
ベッド バハドゥール ビスタ[††] 高田豊雄[††]

近年，OS やソフトウェアの脆弱性を利用し感染を拡大するワームやウィルス等による被害が増加している．被害の拡大を未然に防ぐための有効な手段として，各地のネットワークにおいて攻撃やトラフィックを観測する観測点を設置し，観測点上で収集される情報を解析することがあげられる．本論文では，インターネット上で起こる世界的なトラフィック情報等を把握でき，利用者による様々な解析を可能とする，固定観測点を持たない個人ユーザ参加型のインターネット観測システム ABLA（Agent Based Log Analyzing System）を提案する．ABLA では複数の観測点により Peer-to-Peer ネットワークを構築し，モバイルエージェントを用いて，観測点で収集したログの解析を行う．加えて本論文では，ABLA を実装し，その有効性の評価を行う．

A Proposal of Distributed Internet Monitoring System Using Mobile Agent

Hiroki Kuzuno,[†] Yushi Nakai,[††] Atsumu Watanabe,[††]
Takuya Kawahara,[††] Takashi Katoh,[††] Bhed Bahadur Bista[††]
and Toyoo Takata[††]

Damages caused by internet worms and virus, which use vulnerabilities of software and operating system, are spreading and rapidly increasing. One of the effective means of detecting the damages caused by the worms and virus in early stage is to analyze the packets or network communication logs stored in network sensor that are spread in wide area. In this paper, we propose an agent based log analyzing system by integrating concepts of P2P network and mobile agents to realize detection and protection from the damages which may be caused by the worms and virus in early stage. We also show results of experiments using our prototype system. The results show that our system can analyze useful information from a large number of network sensors with less network traffic.

1. はじめに

近年，ネットワークや計算機の多くが，マルウェアと呼ばれるプログラムによる脅威にさらされている．マルウェアとは，ワームやウィルス等の不正プログラムの総称であり，特に OS やソフトウェアの脆弱性を利用し感染活動を行い，被害を拡大させていくものが増加している．

† 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute of Science and Technology
†† 岩手県立大学大学院ソフトウェア情報学研究科
Graduate School of Software and Information Sciences, Iwate Prefectural University

マルウェアには，ネットワークを経由して感染活動を行い，被害を拡大させるという特徴を持つものが多い．これらマルウェアによる被害は，コンピュータネットワークの普及にともない拡大している．これらの被害を防ぐための対策としては，計算機へのウィルス対策ソフトウェアやパーソナルファイアウォールの導入，ネットワーク全体としてネットワーク型侵入検知システムを導入する等の措置をとることがあげられる．これにより，マルウェアによる攻撃を検出，防御することが可能となり，攻撃による被害の拡大をある程度防ぐことができる．しかし，ソフトウェアの脆弱性は日々新たに報告されており，脆弱性の修正パッチが提供される前やウィルスパターンファイルが更新される前に行われる攻撃，すなわち，ゼロデイアタックの被害も

増加しており，このような新しい脆弱性を利用し感染活動を行うマルウェアが開発されることは十分に考えられる．そのため，つねに最新の攻撃動向やインターネットのトラフィック状況等を把握し，得られた情報をもとに対策を行っていかなければ，マルウェアによる攻撃を十分防ぐことができるとはいえない．

これを受け，マルウェアによる攻撃やトラフィック増加を早期に発見し，被害の拡大を未然に防ぐための試みとして，インターネット上で発生している攻撃の動向やトラフィックの発生状況観測をリアルタイムで行うことを目的とした，インターネット観測システム[1)~3)]の運用が行われている．これらの観測システムでは，攻撃の動向やトラフィックの発生状況をある一定の期間ごとにまとめて公開している．ネットワーク管理者や一般ユーザ等の利用者は，これらの情報をもとに，攻撃に対する対策を事前に行うことができる．

しかし，利用者が自身の環境に合致した独自の対策を行うためには，設置者によって解析が行われた後の情報ではなく，利用者が自身の環境に合わせた解析を行い情報収集を行わなければならない場合がある．にもかかわらず，既存の観測システムにおいて公開されている情報は，システムの設置者によって解析が行われた後の情報のみであることがほとんどである．

また，利用者は対策を実施する際に参照する観測システムの観測情報がその時点ではすでに古く，適切な処置を行えない可能性がある．これは，既存の観測システムは利用者からの要求を受け付けていないため，利用者が観測結果の更新を要求することは困難なためである．さらに最近の研究では観測システムの観測点を抽出する手法も提案されており[4),5)]，この手法を利用し観測点を回避しながら感染活動を行うマルウェアが出現することも考えられる．

そこで我々は，これらの問題点を考慮した分散型インターネット観測システム ABLA（Agent Based Log Analyzing System）の概念を提案してきた[6),7)]．

インターネット観測システムは多数の観測点により構成されるため，各観測点において収集したネットワーク上のパケットやシステムのログや各観測点における解析結果のすべてを 1 カ所に集積すると，膨大なトラフィックや計算コストが発生する可能性がある．そこで，ABLA では，トラフィック量・処理量を抑えるために，モバイルエージェントが各観測点を移動し，観測点上のログ情報を解析，結果を収集する．また，インターネット上の広域のトラフィック情報を取得するために，Peer-to-Peer（P2P）によりネットワークを構築することで観測システムへ多数の観測点の参加を可能とする．

本論文では，ABLA の実現方法を提案するとともに，その詳細について述べ，実験により，1) ABLA による利便性の向上，2) モバイルエージェントによるトラフィックの軽減，について示す．

以下，2 章では既存のインターネット観測システムとその問題点について述べる．3 章では ABLA の詳細について述べ，4 章で ABLA の有効性を検証するための実験と結果の考察を行う．5 章では実験結果の分析と，それに基づいた ABLA の効果，観測にかかるコストについて評価を行う．6 章ではまとめと今後の課題について述べる．

2. 既存手法の問題点と解決方法の提案

インターネット観測システムには，大きく分けて，1) インターネット全体のトラフィックや攻撃動向の統計情報を提供することを目的とし，固定的な観測点によるインターネット定点観測システム，2) インターネットの観測によってワームやウィルス活動の検出を行い，その活動を抑制することを目的とするワーム観測システム，がある．

2.1 インターネット定点観測システム

インターネット定点観測システムには，SANS の Internet Storm Center[1)]，JPCERT/CC の ISDAS[2)] や警察庁の@Police[3)] 等があり，定期的に観測結果や動向情報が提供されている．これらの定点観測システムは，多数の固定された観測点を運用しており，観測点において受信したパケットのログや侵入検知システムのアラート情報を収集し解析を行っている．

このようなインターネット定点観測システムの問題点として，利用者は観測システムの設置者によって観測結果の解析が行われた後の情報しか利用できないことがあげられる．そのため，利用者のネットワークや計算機環境に合致した独自の対策を行うために，利用者が観測システムによって収集された情報に対し別の角度から解析を行うことはできない．

また，利用者は対策を実施する際に参照する観測システムの観測情報がその時点ではすでに古く，適切な処置を行えない可能性がある．これは，観測情報の更新を既存の観測システムに対して要求することが実質的に困難なためである．加えて，利用者は独自の観測点として観測システムの観測点を補完できる可能性があるが，既存の観測システムにおいては，事実上，特定の利用者は観測システムに協力することができず，利用者がインターネットを観測した観測結果を観測システムに提供することは困難といえる．

2.2 ワーム観測システム

ワーム観測システムは，インターネットの観測によってワームを検出し，その抑制を目的とする．

ワーム観測システムに関する研究として，Zou らは，ワームの特性のうち，拡散する際に送出する特徴的なパケットに着目することで，ワーム検出を行う手法を提案している[8]．Singh らは，侵入検知システムが不正アクセスかどうかの検知に利用するシグネチャを自動的に生成することで，ワームの検出，特定を行う Earlybird を提案している[9]．しかし，これらの提案では，観測点で収集した観測情報をすべて 1 カ所に集積した後に解析を行うため，トラフィック量・処理量が膨大になる恐れがある．

Cai らは，観測点により P2P ネットワークを構成し，各観測点で侵入検知システムで利用するシグネチャや検知結果の共有を行い，その検知結果を収集することによりワームを早期検知する WormShield を提案している[10]．しかし，観測点間において観測点の運用上公開できない検知結果の共有や提供が起こり，結果として，機密情報の漏洩や，攻撃者に対して攻撃の足がかりとなる情報を与えてしまう可能性がある．

2.3 提案する解決方法

我々はこれまでに述べた問題の解決方法として，P2P ネットワークとモバイルエージェントを利用した新しいインターネット観測システム ABLA（Agent Based Log Analyzing System）を提案する．

ABLA により構築されるインターネット観測システムは，1) 個人ユーザを含む，一般的なインターネット利用者がインターネット上の攻撃動向を把握し，恒久的なセキュリティの確保・維持を行うための利便性の向上，2) 多数の観測点を確保し，かつ悪意あるプログラムによる観測点のスキャンに対応するために観測点の固定的な設置から解放し，観測範囲を既存の観測システムよりも拡大させることによる観測精度の向上，を目的としている．

ABLA ではシステムに参加するすべての計算機を観測点とするインターネット観測システムを実現するために，P2P によりネットワークを構築する．これにより多数の観測点による観測システムが容易に構築でき，既存の観測システムよりも多くの観測情報を解析対象とすることが可能となり，観測精度の向上が期待される．

また，ABLA では，多数の観測点が存在しうるため，すべてのログ情報を集めると，それだけで膨大なトラフィックとなる可能性がある．そこで，ABLA ではモバイルエージェントが観測点を移動し，観測点上のログ情報を解析，結果を収集する．このことにより，多数の観測点に散在するログ情報すべてを共有し 1 カ所に集積した後に解析を行うよりもトラフィック量・処理量を抑えることが可能である．

3. ABLA: Agent Based Log Analyzing System

ABLA では，複数の計算機を観測点として P2P ネットワークを構築し，モバイルエージェントを用いて，それらの観測点を移動しながら，観測点上で収集した観測情報の解析を行う．ABLA は，1) P2P ネットワークを利用することで，特定のサーバを持つことなく攻撃動向やトラフィックの発生状況等の情報を共有することが可能となる，2) モバイルエージェントを用いることで，移動先の観測点上のログ情報のうち，公開が許可された部分のみを解析し，結果を収集することで，観測点となる計算機の運用上公開できない情報を隠蔽しながらログの解析を行うことができる，3) モバイルエージェントを用いることで，散在するログ情報すべてを 1 カ所に集積した後に解析を行う手法や，各観測点に解析要求を発行し，観測点上においてログ情報を解析，得られた解析結果を受け取る手法よりもトラフィック量・処理量を抑えることができ，かつログ情報の提供者の匿名性を確保することができる，といった特徴を持っている．

利用者は ABLA による観測システムに自由に参加，離脱，観測結果の要求を行うことが可能である．これにより動的かつ多数の観測点を確保することが可能となり，これら多数の観測点で観測された情報から最新の観測結果を得ることができる．また，このように動的かつ多数の観測点を確保することは，観測点をスキャンし回避する種類のマルウェアへの対抗策となることが期待できる．すなわち，観測点が動的に変化することにより観測点の回避を困難にし，また，仮に回避されたとしても，そのこと自体が ABLA に参加している利用者への被害も回避されることとなる．

ABLA が広く利用されることにより，個人ユーザが参加可能なインターネット観測システムを構築することができる．その結果，従来のインターネット観測システムよりも広域的かつ，多数の観測点によるログの情報を取得でき，早期対策を講じる手段の 1 つとなる．これは，大規模な組織を必要としないセキュリティ情報収集の手段となることが期待される．

3.1 ABLA の概要

ABLA の概要を**図 1** に示す．ABLA は複数の観測点より P2P ネットワークを構築し，観測点において

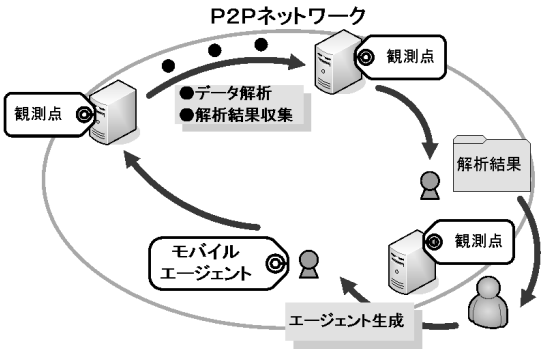

図 1 ABLA 概要
Fig. 1 ABLA overview.

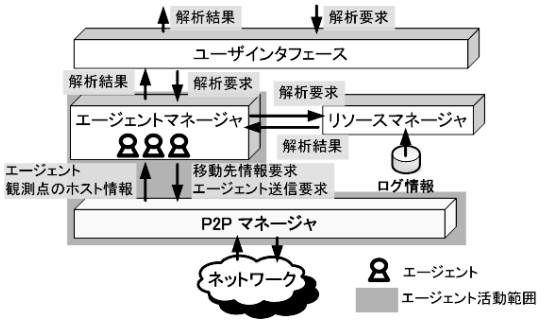

図 2 ABLA アーキテクチャ図
Fig. 2 ABLA architecture.

収集したネットワーク上のパケットやシステムのログを，モバイルエージェントにより解析，収集する．

利用者がログの解析要求を発行し，観測結果を受け取るまでのシステムの動作は次のとおりである．

(1) 利用者は ABLA による観測システムに対する解析要求として，以下の項目を指定．
- 解析対象とする時刻の範囲
- 頻度を計算する解析を行う時間間隔
- 解析の終了条件
- 解析対象とする情報

(2) 観測システムは利用者からの解析要求を受け取り，モバイルエージェントを生成．

(3) モバイルエージェントは移動先候補となる観測点とその情報を要求（詳細は 3.2.1 項）．

(4) モバイルエージェントは利用者の解析要求と，移動先候補のホスト情報から移動先の観測点を決定．

(5) モバイルエージェントは決定した観測点に移動，観測情報の解析を要求し，解析結果を受け取り，自身の持つ観測情報に統合．

(6) 帰還条件を満たしていれば，モバイルエージェントは利用者のもとに帰還．帰還条件を満たしていない場合，手順 3 より繰り返す．

(7) 利用者は帰還したモバイルエージェントより，観測情報を取得．

3.2 ABLA の構成要素

ABLA のアーキテクチャを**図 2** に示す．ABLA は，4 つのコンポーネント，1) P2P マネージャ，2) リソースマネージャ，3) エージェントマネージャ，4) ユーザインタフェース，と 5) モバイルエージェントから構成される．

P2P マネージャは，モバイルエージェントに対して移動先候補の情報，転送機能を提供する．エージェントマネージャは，ユーザインタフェースより解析要求を受け取り，エージェントを生成する．加えて，エージェントに対し P2P マネージャとの仲介機能を提供する．リソースマネージャは，ログを解析し，エージェントに解析結果を提供する．ユーザインタフェースは，利用者に対し解析要求の受け取りとエージェントより受け取った解析結果を出力する．解析要求はエージェントマネージャへ渡される．モバイルエージェントは，リソースマネージャへ解析要求を発行することでログを収集し，P2P マネージャへ次の観測点への移動の要求を発行することで，他の観測点へ移動する．

3.2.1 P2P マネージャ

ABLA を複数の計算機で動作させ，それらを接続することで P2P ネットワークを構築する．P2P マネージャは直列化されたモバイルエージェントの転送を行う．さらに，エージェントが次の移動先観測点を決定できるように，観測点間で観測点のホスト情報としてログの種類や接続先情報等のやりとりを行う．

観測点において公開されるログ情報には観測点固有の情報が含まれている可能性があり，エージェントが訪れた観測点の IP アドレスを入手することで，観測点の IP アドレスと収集した観測結果を対応付けされることは防がなければならない．そこで，観測点は自身の IP アドレスのハッシュ値を持ち，P2P ネットワーク上においては，このハッシュ値を観測点を示す識別子とする．P2P マネージャのみが接続中の観測点の IP アドレス，その観測点のハッシュ値とホスト情報を管理し，他のコンポーネントには IP アドレスを通知せず，観測点のハッシュ値のみを通知することで観測点が特定されることを困難とする．

エージェントは自身の転送要求の際に，P2P マネージャに移動先の観測点のハッシュ値と直列化したエージェント自身を渡すだけでよく，エージェントに対し IP アドレスを隠蔽したまま，ハッシュ値のみによっ

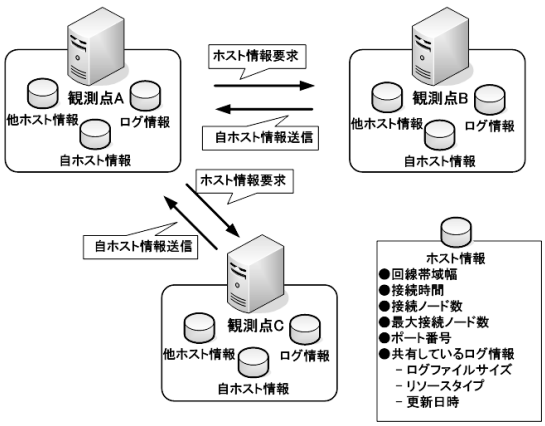

図 3 ホスト情報の受渡し
Fig. 3 Exchange of host information.

て観測点間を移動することが可能となる．

P2P ネットワークの構築手順は次のとおりである．

(1) 接続要求を行う観測点は，P2P ネットワークに参加するための最初の接続先として，すでに接続待機状態となっている観測点の IP アドレスを利用者より受け取り，その観測点に対して接続要求を発行する．

(2) 接続要求を受け取った観測点は，接続応答とともに自身のホスト情報（後述）を返す．

(3) 接続要求を行う観測点は，接続応答と接続先のホスト情報を確認後，自身のホスト情報を送り，接続先のハッシュ値を得て，接続中の観測点とする．

(4) 接続要求を受け取った観測点は，接続要求元のホスト情報を確認後，その観測点のハッシュ値を得て，接続中の観測点とする．

P2P マネージャが他の観測点と交換するデータは，接続回線の帯域幅，ABLA によって構築された P2P ネットワークへの接続時間，現在の接続ノード数，公開しているログリソースの情報，サービスを提供しているポート番号であり，これらをホスト情報と呼ぶ．観測点間でのホスト情報の交換について概要を**図 3** に示す．これらのデータは，エージェントが移動を行う際に，利用者から受け取った要求と合致するログ情報を持つ観測点を探索するのに利用される．なお，これらのデータは接続時およびホスト情報要求を受け取った際に送信される．

3.2.2 リソースマネージャ

リソースマネージャは，解析対象となるログ情報を管理し，モバイルエージェントから解析要求を受けた際は，解析を実行し，その結果をエージェントに渡す．

モバイルエージェントはリソースマネージャの仲介なしには，解析対象のログ情報を含むあらゆる観測点上のリソースにアクセスすることはできない．これは，モバイルエージェントを利用するシステムにおいては，悪意を持った利用者が改竄したモバイルエージェントにより移動先の観測点上で不正な命令の実行による被害等を防ぐことが重要なためであり（詳細は 3.4 節），エージェントが観測点となる計算機上のログ情報に直接アクセスすることを許可した場合，運用上公開できない情報（観測点の IP アドレスやシステムファイル等）にアクセスされる可能性がある．リソースマネージャのみが解析対象のログ情報へ直接アクセスすることで，エージェントはリソースマネージャに対し，解析要求を出さない限り，ログ情報の解析は行えない．これにより，観測点の提供者により公開が許可された部分のみを解析対象として，運用上公開できない情報を非公開とすることを可能としている．

3.2.3 エージェントマネージャ

エージェントマネージャは，ユーザインタフェースより利用者からの解析要求を受け取った場合，モバイルエージェントを生成し，生成したエージェントに利用者からの解析要求を渡す．また，P2P マネージャとエージェントとの間でやりとりされる移動先候補のホスト情報の仲介を行い，エージェントが移動先の観測点を決定した場合，エージェントマネージャはエージェントを直列化し，エージェントより受け取った移動先の観測点の情報とともに P2P マネージャにエージェント送信要求として送る．さらに，P2P マネージャより他の観測点から移動してきた直列化されたエージェントを受け取った場合，直列化されたエージェントを直列化復元し，動作させる．

3.2.4 ユーザインタフェース

ユーザインタフェースは，エージェントへの解析要求の設定や解析の開始，エージェントが持ち帰った解析結果の出力等，利用者とシステム間の橋渡しを行う．なお，実装では CUI，GUI の 2 種類のインタフェースを利用することができ，このうちの GUI を**図 4** に示す．GUI では，エージェントの持ち帰った解析結果を，グラフとして可視化することができる．解析結果をグラフとして可視化した例を**図 5** に示す．

3.2.5 モバイルエージェント

ABLA におけるエージェントは，エージェントマネージャより生成されるモバイルエージェントであり，観測システムの P2P ネットワークに接続された観測点を自律的に移動し，リソースマネージャと解析要求，解析結果のやりとりを行う．エージェントが持つ解析

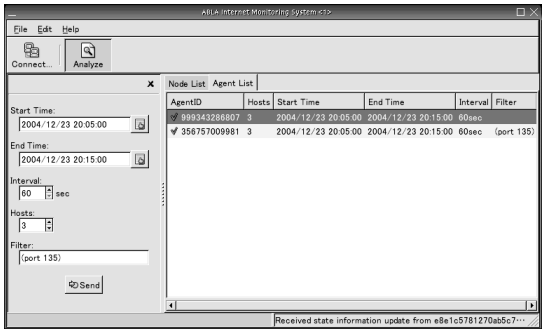

図 4 ABLA の GUI
Fig. 4 GUI interface of ABLA.

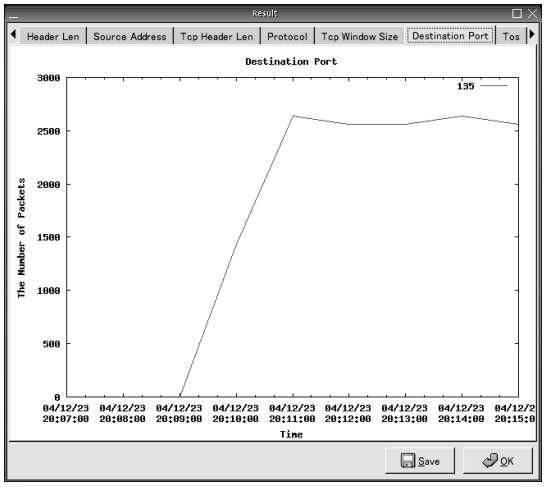

図 5 解析結果の例
Fig. 5 Example of analysis result.

表 1 エージェントが持つ解析要求の例
Table 1 Example of analysis request on agent.

時刻範囲	2005/8/19 20:00〜2005/8/19 20:30
時間間隔	5 分
解析情報の種類	tcpdump
解析対象の情報	(port 135)

要求の例を**表 1** に示す.

モバイルエージェントは，次の順序でログの解析，移動を行う.

(1) 利用者からの解析要求を受け取り，観測点間の移動を開始.
(2) 他の観測点に移動後，解析要求を移動先観測点のリソースマネージャに渡し，共有が許可されたログの解析結果を取得.
(3) 観測点の移動を繰り返し，利用者から与えられた観測点の巡回数を満たした場合，解析終了と見なし，解析結果とともに利用者のもとに帰還.

利用者から発行され，エージェントマネージャより生成されたモバイルエージェントへ渡される解析要求は，1) 観測点上のログ情報の解析に使用される項目，2) エージェントの移動先選択に使用される項目，に分けられる．観測点上のログ情報の解析に使用される項目は，解析対象とする時刻の範囲，解析を行う時間間隔，解析対象とする情報，とし，エージェントは移動先の観測点において，これらの項目を解析要求としてリソースマネージャに渡し，リソースマネージャがログ情報を解析し，その解析結果を受け取る．エージェントの移動先選択に使用される項目は，解析の終了条件，解析対象とする情報，とし，エージェントは移動先とする観測点を，解析対象とする時刻の範囲と P2P マネージャより受け取る観測点のホスト情報を比較し決定する．ただし，解析の終了条件が成立していた場合は，利用者のもとへ帰還する.

モバイルエージェントが観測点を移動する際の移動アルゴリズムは次のとおりである.

(1) P2P マネージャより他の観測点のホスト情報を取得.
(2) 取得した観測点のホスト情報から，次の条件 (a)〜(c) を順に満たす観測点を移動先として選択し (3a) の処理を実行. (a) と (b) をともに満たす選択先がない場合は (3b) の処理を実行.
  (a) 自身が持つ解析要求の条件を満たしていること.
  (b) まだ訪れていない観測点であること.
  (c) 最も大量のログ情報を持つこと.
(3a) 選択先の観測点に移動.
(3b) いったん最も接続ノード数の多い観測点に移動し，移動アルゴリズムを (1) より繰り返す.

モバイルエージェントは観測点の移動を行った結果，エージェントが持つ観測要求に含まれる解析の終了条件を満たしていない限り，現在の観測点から次の観測点に移動できなくなることは避けなければならない．そのため，条件 (a)，(b) を満たす観測点が存在しない場合には，移動アルゴリズムの (3b) において，いったん最も接続ノード数の多い観測点に移動し，移動アルゴリズムを (1) より繰り返す．このことで，移動アルゴリズムの条件 (a)〜(c) を満たす移動先候補を発見する可能性を向上させる.

モバイルエージェントがこの移動アルゴリズムを利用することで，以下のような利点を確保しながら，ネットワークに散在するログ情報の中からユーザの要求に合致したものだけを収集する.

- ログ情報を 1 カ所に集積する手法や，ログ情報の解析結果を受け取る手法により起こる膨大なトラ

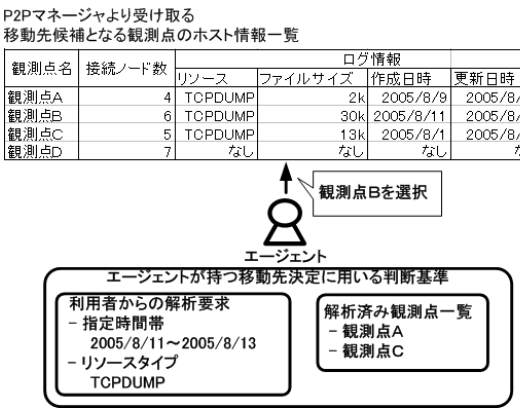

図 6 エージェントの経路選択
Fig. 6 Route selection of agent.

フィックを大幅に削減できる．このアルゴリズムを用いた場合，トラフィックはエージェントが，1) 解析を行うために観測点を移動，2) 解析終了後，利用者のもとに帰還，するときにのみ発生する．

- 1 カ所に集積された膨大なログ情報を一括で解析するのではなく，エージェントが移動先となる観測点上でそれぞれ解析を行うため，処理量を大幅に抑えることができる．
- P2P ネットワーク上をエージェントが観測点の IP アドレスのハッシュ値のみを利用し，移動しながら解析を行うことで，どのような IP アドレスを持つ観測点のログ情報を解析したかを隠蔽することができ，ログ情報提供者の匿名性を確保することが可能となる．

移動先選択の一例として，**図 6** では，1) エージェントは，P2P マネージャより移動先候補のホスト情報一覧を受け取る．2) エージェントが保持する利用者からの解析要求と，エージェントがすでに移動し，ログ情報を解析済みの観測点一覧をもとに，条件 (a) 解析要求の条件を満たしている観測点かつ (b) まだ訪れてない観測点のうち，(c) 最も大量のログ情報を持つ，という条件に該当する観測点を次の移動先として決定する．(3a) 移動先の観測点を決定したエージェントは，エージェントマネージャにより直列化され，移動先の情報とともに P2P マネージャに送られ，エージェントは指定した移動先へと送信される．

3.3 ログ情報の解析方法

観測点において，公開されるログ情報の解析を直接行うのはリソースマネージャである．リソースマネージャは観測点で公開するログ情報を管理しており，移動してきたエージェントより解析要求を受け取り，要求に該当するログ中の公開が許可されている部分に対して解析を行い，その結果をエージェントへ渡す．

なお，現在の実装では解析手法を頻度解析としている．すなわち，解析の対象となるログに含まれる各パケットのヘッダフィールドの内容について，出現頻度をカウントする．解析時には，観測点上で指定されている公開許可の条件，移動してきたエージェントより受け取る解析要求に含まれる，解析対象とする時刻の範囲，頻度をカウントする時間間隔，解析対象と見なす条件を利用する．

また，ログ中に観測点の運用上公開できない情報を含む場合は，その様な情報を選択的に非公開とすることができなければならない．そのため ABLA においては，利用者が公開するログに含まれる情報の公開/非公開の選択を自由に行えるようにしている．具体的には，パケットキャプチャのログの場合，宛先アドレスや送信元アドレス，ポート番号等，TCP/IP における各ヘッダフィールドごとに公開か非公開かを設定することが可能である．また，公開できない情報を隠蔽したうえで，可能な範囲でログの一部を公開するといったことを可能とするためには，公開するログの範囲を柔軟に設定できることが望ましい．そのため，パケットキャプチャデータに対する一般的なフィルタ形式である BSD Packet Filter（BPF）形式の条件式[11] を利用し，BPF で記述可能な特定の要素を含むパケットを非公開とすることを可能としている．

3.4 ABLA のセキュリティについて

ABLA では，モバイルエージェントを利用しているため，悪意を持った利用者がエージェントを利用し，ABLA に参加している観測点に対し攻撃を行うことが考えられる．

モバイルエージェントを利用した観測点への攻撃としては，エージェントが移動先において不正な命令を実行するように変更を加えられたことで起こる，エージェントの改竄による攻撃，故意にサイズが大きく設定されたエージェントによる Denial of Service（DoS）攻撃が考えられる．また，悪意を持つ利用者が複数の観測点を設置し，捏造したログ情報を提供することによる観測情報の操作，ABLA のコンポーネントの 1 つである P2P マネージャの改竄による IP アドレスとハッシュ値の漏洩等が考えられる．

一般に，モバイルエージェントの移動は，1) 強い移送，2) 弱い移送，に分けられる[12]．強い移送では，ソースコード，インスタンス変数，実行環境を転送し，エージェントは移動前と同じ環境で実行を継続することが可能である．弱い移送では，ソースコード，インスタンス変数のみが転送され，エージェントは移動後，

移動前の実行を継続することはできない．

ABLA では，観測点となる計算機を保護するため，エージェントの改竄による攻撃を防がなければならない．そこで ABLA では，弱い移送で転送する情報のうち，インスタンス変数のみを送ることでエージェントの移動を行う．ソースコード自体は転送せずに，エージェントの移動先におけるソースコードを利用することで，不正な命令がソースコードに含まれることを回避し，エージェントの改竄による攻撃を困難としている．また，移動先の観測点のソースコードが改竄されていた場合，移動先においてエージェントに対する攻撃が行われる可能性があるが，エージェントは巡回した観測点のハッシュ値のみしか保持しておらず，エージェントが攻撃された場合においても，攻撃者がエージェント送信元，巡回した観測点の IP アドレスを知ることは困難である．

故意にサイズが大きく設定されたエージェントによる観測点に対する DoS 攻撃に対しては，エージェントが移動する際，エージェントのサイズに対し制限を設定することである程度回避が可能である．そこでこの攻撃の回避策として，エージェントのサイズで観測点間のエージェントの転送の可否を判断することを検討中である．

ABLA では，ABLA を利用する誰もが観測システムに参加することが可能である．そのため，悪意を持った利用者が複数の観測点を設置し，捏造したログ情報の提供やエージェントが保持する観測情報の改竄により，異常なトラフィックの隠蔽（自ら開発したマルウェアの活動の隠蔽等）を試みる恐れがある．ABLA におけるエージェントは，移動先の観測点においてログ情報の解析結果を受け取り，すでに保持している観測情報に加えるため，巡回する観測点のうち 1 つでも異常なトラフィックが観測されていれば，その結果が反映される．したがって，異常なトラフィックの隠蔽が可能となるのは，エージェントが巡回する観測点のうち，利用者のもとに帰還する直前の観測点において，捏造したログ情報を提供またはエージェントが保持する観測情報を改竄された場合である．しかし，直接利用者の観測点と接続しなければ，利用者のもとにエージェントが帰還する直前の観測点とはなりえないため，観測結果を操作することは困難である．よって，悪意ある利用者は多くの利用者と接続するために観測システムの観測点の大半を自身の観測点で占める必要があるため，観測結果から異常なトラフィックを隠蔽することは事実上困難である．一方，異常なトラフィックを捏造しログ情報として提供された場合，その結果が観測結果に反映される恐れがある．しかし，ABLA において利用者は詳細な解析要求を設定することにより，巡回する観測点数の変更や，特定のポート番号，IP アドレス等に関する観測情報の選択が可能である．したがって，利用者は ABLA により異常なトラフィックを検出した際，巡回する観測点数を変更する等さらに詳細な解析要求を発行することで，このトラフィックが捏造されたものであるかどうか確認できると考えられる．また，悪意ある利用者の問題は社会調査の分野では，非標本誤差[13] の問題として知られており，様々な対処法が議論されている[13),14)]．さらに除外のための判定方法（スミルノフ・グラプス検定等）についても研究されている[15)]．ABLA においても，こうした手法を考慮した，改竄された観測情報を除外するための判定方法を検討中である．

また，悪意を持った利用者が ABLA のコンポーネントの 1 つである P2P マネージャを改竄し，ABLA に参加することで，改竄が行われた観測点に隣接ノードとして接続する他の観測点の IP アドレスとハッシュ値が漏洩する恐れがある．ABLA の P2P ネットワークにおいて観測点間は TCP/IP を用いて通信しており，改竄された観測点の隣接ノードとして接続した観測点については IP アドレス漏洩を防止することは困難である．しかしながら，ABLA の利用者にとっても最も脅威なのは，観測システムの観測結果と自身の観測点の IP アドレスが対応付けされることにより，その観測点において提供しているサービスや脆弱性情報が漏洩することであるが，ABLA では IP アドレスのハッシュ値を観測点の識別子とし，改竄されたエージェントによって訪れた観測点の IP アドレスと観測結果が対応付けされることを防いでいる．したがって，悪意を持つ利用者に観測点の IP アドレスとそのハッシュ値を入手された場合であっても，観測点の提供しているサービスや脆弱性情報は漏洩しないため，観測点が攻撃を受け脅威にさらされる可能性は低いといえる．

4. 実　　験

我々は，ABLA による観測システムの有効性を検証するために，1) ABLA の利用による利便性の向上，2) モバイルエージェントによるトラフィックの軽減，に関する 2 つの実験を行った．

4.1 実験 1：ABLA による観測実験

ABLA による観測を実際に行い，ABLA を利用することによってどのような効果が得られるかについて実験を行った．

なお，実験環境として，ABLA を起動した計算機 5

台が相互に接続し構築された P2P ネットワークを利用した．システム管理者はこのネットワークに参加し，解析を行うものとした．これら 5 台の計算機は，(a) ファイアウォールの外にあるグローバル IP アドレスを持つ計算機 3 台，(b) ポートフォワーディングによって外部からの接続が可能となっているファイアウォール内の計算機 1 台，(c) 以下のシナリオにおけるシステム管理者が ABLA を利用する計算機 1 台，で構成される．なお，(a) のうちの 1 台はゲートウェイとして動作しており，この計算機以下のネットワークにはセキュリティパッチ未適用の WindowsXP がインストールされた計算機が接続されている．

これら 5 台の計算機では，tcpdump により，実際のネットワークトラフィックデータを収集している．これらの実際のネットワークトラフィックデータを ABLA で公開する観測情報とし，その内公開するヘッダフィールドは送信元 IP アドレス，宛先ポート番号の 2 つとした．

実験 1 におけるシナリオは以下のとおりである．

(1) システム管理者は管理下にあるネットワークで異常なトラフィックを検知した．この異常なトラフィックの原因を知るため，システム管理者は ABLA を利用し解析を行った．この際，システム管理者はエージェントに対し，以下のような解析要求を発行した．
開始時間 2006/01/04 19:15:00
終了時間 2006/01/04 20:15:00
時間間隔 60 秒
巡回ホスト数 4
フィルタ なし

(2) 帰還したエージェントから結果を受け取りグラフ化したところ，**図 7** を得た．これより，ポート 135 に対するパケットの数が異常な数値を示していることが分かった．

(3) システム管理者は，このポート 135 に対するパケットを送出しているホストを知るために，さらに詳細な解析を行うため以下のような解析要求を発行した．
開始時間 2006/01/04 19:15:00
終了時間 2006/01/04 20:15:00
時間間隔 10 秒
巡回ホスト数 4
フィルタ (port 135)

(4) 再び帰還したエージェントから結果を受け取りグラフ化したところ，**図 8** を得た．これより，ポート 135 に対するパケットは，特定のホスト (x.156.43.39) から大量に送出されていることが分かった．ここで，IP アドレスを特定することができたのは，各観測点が持つログ情報のうち，送信元 IP アドレスを公開しているためである．
なお，これらの送信元 IP アドレスは各観測点上で観測したパケットの送信元 IP アドレスであり，エージェントが巡回した観測点の IP アドレスではない．したがって，エージェントが巡回した観測点のハッシュ値を観測点の IP アドレスと対応付けすることはできない．

(5) この情報をもとに，システム管理者は x.156.43.39 からのポート 135 に対するアクセスを拒否することで，異常なトラフィックを管理下のネットワークから排除することができる．

本実験では，ABLA に参加している各ホストにおいて，送信元 IP アドレスを公開していることを想定したが，各観測点において送信元 IP アドレスを公開

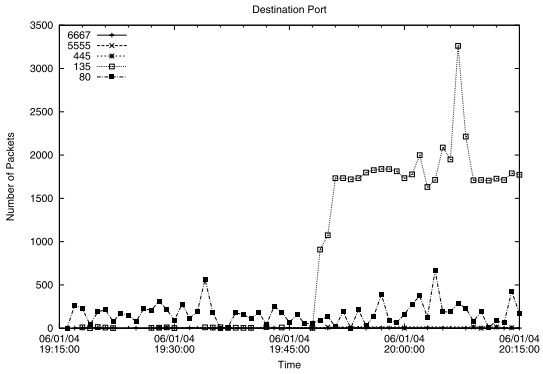

図 7 ABLA を利用して解析されたパケットの数（宛先ポート）
Fig. 7 Number of packets analyzed using ABLA (Destination Port).

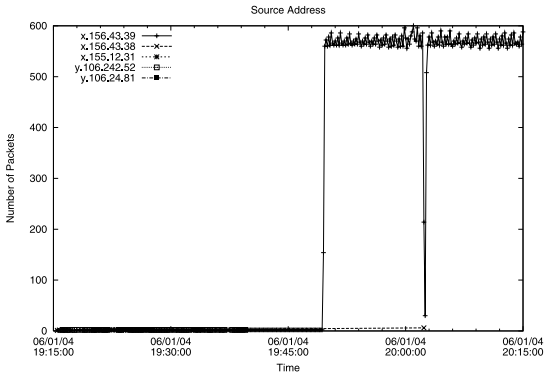

図 8 ABLA を利用して解析されたパケットの数（送信元ホスト）
Fig. 8 Number of packets analyzed using ABLA (Source Host).

していない場合でも，(2) でポート 135 への異常なパケットの送出が起こっていることを把握できる．したがって，ポート 135 に対するアクセスを全面的に拒否することで対処することが可能である．

提案手法と従来のインターネット観測システムとを比較すると，従来システムでは受動的な観測しか行うことができず，ポート 135 に対するパケットの数が異常値を示していても，その結果を受けてさらに詳細な解析を行うこと等が不可能であったが，ABLA を利用することで，利用者は能動的な観測を行うことができる．すなわち，ある解析結果を受け，さらに詳細な解析を行い，その解析結果をもとにセキュリティ対策を行うことが利用者側において可能となる．

また，管理者自身が解析要求を発行できるため，従来のインターネット観測システムでは得られなかった，よりリアルタイム性の高い情報を得ることが可能である．

本実験では ABLA による観測ネットワークに接続されたホストの数は 5 であったが，さらに多くのホストが接続された場合でも同様の解析が可能であると考えられる．より多くのホストが接続された環境での実験は今後の課題である．

4.2 実験 2：モバイルエージェントによるトラフィックの軽減

実験 2 では，ABLA による観測システムのトラフィック削減の有効性を検証するために，解析情報を収集するモバイルエージェントのサイズについて評価実験を行った．解析要求元に解析結果が集まるまでのネットワークトラフィックのサイズ変化の比較には，散在するログ情報すべてを 1 カ所に集積した後に解析を行う手法，および，各観測点に解析要求を発行し，観測点上においてログ情報を解析，得られた解析結果を受け取る手法を利用した．

まず，13 個の観測点からなる観測システムを構築した．それぞれの観測点において解析対象となるログ情報を用意し，すべての観測点のログ情報の総計は 122,293,543 バイトとした．その後，観測点の 1 つにおいて，エージェントに対しすべての観測点を移動するように解析要求を出す．各観測点のログ情報を解析，その結果を収集した際のエージェントのサイズ（バイト）の変化を計測した．また，ログ情報を 1 カ所に集積する手法では，すべての観測点よりログ情報を集積した際の集積点におけるログ情報の総計サイズの変化を計測した．各観測点上においてログ情報の解析を行い，その解析結果を受け取る手法では，すべての観測点より受け取った解析結果の総計サイズの変化を計測した．

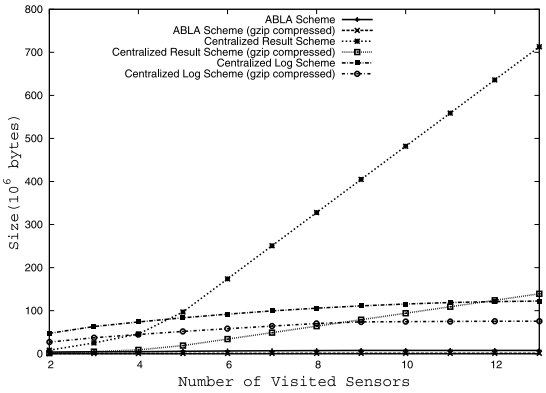

図 9 ネットワークを流れるデータサイズの変化
Fig. 9 Network traffic of proposed method and centralized method.

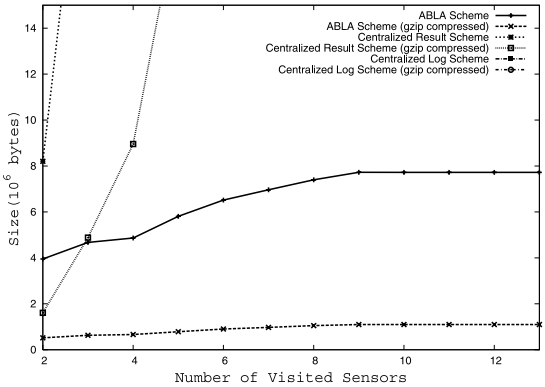

図 10 ネットワークを流れるデータサイズの変化（拡大）
Fig. 10 Network traffic of proposed method and centralized method (magnified).

結果は**図 9** のとおりである．なお，**図 10** は図 9 の下部を拡大したものである．図の横軸はログ情報を提供している観測点の数，縦軸はこれらの観測点において解析を行った際にネットワーク上を流れるデータ量を示している．また，“ABLA Scheme” は ABLA を利用した場合，“ABLA Scheme（gzip compressed）” はエージェント移動の際にエージェントの持つデータを gzip 圧縮した場合であり，“Centralized Log Scheme” は各観測点上におけるログ情報すべてを 1 カ所に集積する場合，“Centralized Result Scheme” は各観測点上におけるログ情報の解析を行い，その解析結果を受け取る場合である．また，“Centralized Log Scheme (gzip compressed)”，“Centralized Result Scheme (gzip compressed)” はそれぞれ，各観測点上におけるログ情報およびその解析結果を gzip 圧縮した場合である．

この図から，Centralized Log Scheme や Central-

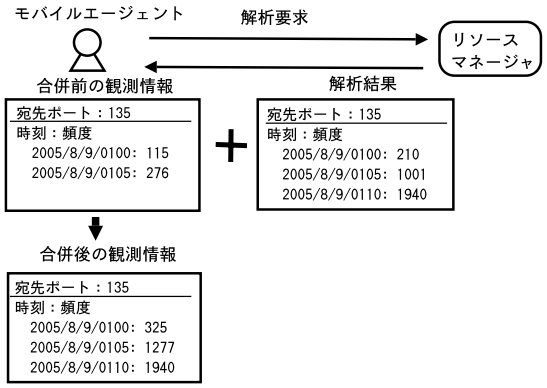

図 11 ABLA によるエージェントにおける解析結果の合併
Fig. 11 Merge of analyzed results.

ized Result Scheme よりも，ABLA によるエージェントが観測点間を移動しログ情報を解析，収集することで，観測システムとしてのトラフィック軽減を実現していることが分かる．これは，ABLA によるエージェントは，各観測点で得た解析結果を個々に保持したまま収集するのではなく，**図 11** のように各観測点で得た解析結果に含まれる内容を，エージェントが元々保持していた観測情報に加えるためである．さらに，解析結果を圧縮することで，トラフィックを軽減することが可能である．

また図 9 において，Centralized Result Scheme のトラフィックが他の手法のトラフィックよりも大きくなっているのは，実験において利用したログ情報が圧縮されたバイナリデータであり，ログ情報の解析結果は非圧縮のテキストデータのためである．したがって，Centralized Result Scheme においても解析結果をバイナリデータとすることでトラフィックの削減は可能であるが，観測点の増加にともないトラフィックが増大することに変わりはない．すなわち，巡回する観測点が増加すると，トラフィックの点において，本提案手法が優位である．

5. 評　　価

ABLA はシステムに参加するすべての計算機を観測点とする個人ユーザが参加可能なインターネット観測システムである．従来のインターネット観測システムでは受動的な観測しか行えず，また，ある観測結果を受けてさらに詳細な解析を行うこと等ができなかったのに対し，ABLA ではある観測結果を受けてさらに詳細な解析を行い，その結果をもとにセキュリティ対策を行うことが可能となることを実験により示した．

ABLA による観測システムには，多数の観測点が存在しうることから，各観測点において収集した観測情報や解析結果のすべてを利用者のもとに集めると，膨大なトラフィック・処理量が発生する可能性がある．実験より，モバイルエージェントを用いて各観測点上のログ情報を解析，結果を収集することで，観測システムとしてのネットワークトラフィックを軽減しているという結果を得た．今後は，ABLA の実環境での運用を行い，既存の観測システムとの観測精度の比較等，さらなる評価が必要である．

ABLA に参加している観測点の 1 つが，ウィルスやワームの活動によるトラフィックを観測し，さらに観測情報の多くを公開していた場合を想定すると，観測結果を見ることにより，ABLA の利用者は観測システムの観測網のどこかにおいて疑わしいトラフィックが発生していることを知ることができる．さらに，ABLA では利用者が自身の環境に合わせた柔軟な解析を行うことが可能なため，その観測結果より，ウィルスやワームの感染元である計算機を特定し，それらからの攻撃による被害を防ぐための適切な対策をとることが可能となる．

また，ログ情報として送信元アドレスが公開されていない場合でも，4.1 節における実験結果のように，一部のポート番号に対するパケット数の異常な増加から，疑わしいトラフィックが発生していることを知ること等は十分に可能である．したがって，感染元である計算機の特定は困難としても，ABLA を利用することで得た解析結果は，利用者にとって適切なセキュリティ対策をとるための重要な指針となる．

計算コストについては，ABLA Scheme および Centralized Result Scheme では，Centralized Log Scheme における解析コストに加え，各観測点から受け取った解析結果の集計（加算）のコストが必要となる．しかしこの加算は，利用者の解析要求に依存するが，（巡回した観測点数）×（公開情報に含まれる識別子（ポート番号，送信元 IP アドレス等）の総数）×（解析対象期間/時間間隔）[回] 程度であり，それほど大きなコスト増ではないといえる．さらに，ABLA Scheme 以外では，ログ情報や解析情報を提供する観測点が増加するにともない大量のデータが利用者のもとに集積し，その解析・結果集計に膨大な処理が発生する可能性がある．一方，ABLA Scheme では，ログ情報の解析，解析結果集計はともに各観測点に分散して行われる．そのため，各観測点はそれほど高い計算能力を備える必要はなく，比較的性能の低い計算機でも，ABLA による観測システムに参加することが可能である．

6. おわりに

本論文では，既存のインターネット観測システムの 1) 利用者は観測情報を自身の環境に合わせた解析を行えない，2) 利用者から更新要求を発行できない，3) 利用者がインターネットを観測した観測結果を観測システムに提供することができない，といった問題点を解決する，モバイルエージェントと P2P ネットワークを利用した分散型インターネット観測システム ABLA（Agent Based Log Analyzing System）を提案し，その有効性を検証するための評価実験を行った．

ABLA によって，固定観測点を持たない個人ユーザ参加型のインターネット観測システムが実現可能であり，インターネット上で発生する世界的なトラフィック情報等の把握と，利用者による様々な角度からの解析が可能となる．その結果，マルウェアによる攻撃やトラフィックの増加を早期に発見し，被害の拡大を未然に防ぐための対策を行うことが可能となる．

今後の課題として，ABLA による観測システムにおいて，より利用者の解析要求に合致した観測点を巡回するために，モバイルエージェントの移動アルゴリズムの改善があげられる．また，観測点から構築される P2P ネットワークのスケーラビリティに関する評価を行うことがあげられる．

謝辞 本研究の一部は情報処理推進機構（IPA）による 2005 年度未踏ソフトウェア創造事業（未踏ユース）の支援を受けて研究開発を行っている．

参考文献

1) SANS, Internet Storm Center. http://isc.sans.org/
2) JPCERT/CC，インターネット定点観測システム．http://www.jpcert.or.jp/isdas/
3) @Police，インターネット定点観測. http://www.cyberpolice.go.jp/detect/observation.html
4) Bethencourt, J., Franklin, J. and Vernon, M.: Mapping Internet Sensors With Probe Response Attacks, *14th USENIX Security Symposium* (*SEC'05*), pp.193–208 (2005).
5) Shinoda, S., Ikai, K. and Itoh, M.: Vulnerabilities of Passive Internet Threat Monitors, *14th USENIX Security Symposium* (*SEC'05*), pp.209–224 (2005).
6) 川原卓也，渡邊 集，葛野弘樹，中井優志，加藤貴司，Bhed Bahadur Bista，高田豊雄：Peer-to-Peer ネットワークにおけるエージェントを用いたログ解析ソフトウェア，コンピュータセキュリティシンポジウム（CSS2004），pp.97–102 (2004).
7) 葛野弘樹，川原卓也，渡邊 集，中井優志，加藤貴司，Bhed Bahadur Bista，高田豊雄：暗号と情報セキュリティシンポジウム 2005（SCIS2005），pp.1729–1734 (2005).
8) Zou, C.C., Gao, L., Gong, W. and Towsley D.: Monitoring and Early Warning for Internet Worms, *10th ACM Conference on Computer and Communication Security* (*CCS'03*), pp.27–31 (Oct. 2003).
9) Singh, S., Estan, C., Varghese, G. and Savage, S.: Automated Worm Fingerprinting, *Proc. Usenix Symposium on Operating Systems Design and Implementation* (*OSDI'04*), pp.45–60 (2004).
10) Cai, M., Hwang, K., Kwok, Y., Song, S. and Chen, Y.: Collaborative Internet Worm Containment, *IEEE Security and Privacy Magazine*, May/June, pp.25–33 (2005).
11) McCanne, S. and Jacobson, V.: The BSD Packet Filter: A New Architecture for User-level Packet Capture, *USENIX Winter*, pp.259–270 (1993).
12) Cugola, G., Ghezzi, C., Picco, G.P. and Vigna, G.: Analyzing Mobile Code Languages, *Mobile Object Systems*, Vol.1222 of LNCS, pp.94–109 (1997).
13) 新 睦人：社会調査の基礎理論，川島書店 (2005).
14) 飽人 弘：社会調査ハンドブック，日本経済新聞社出版局 (1987).
15) 佐伯 胖，松原 望：実践としての統計学，東京大学出版会 (2000).

(平成 17 年 9 月 30 日受付)
(平成 18 年 3 月 2 日採録)

葛野 弘樹

1982 年生．2005 年岩手県立大学ソフトウェア情報学部ソフトウェア情報学科卒業．同年奈良先端科学技術大学院大学情報科学研究科に入学．侵入検知システムに関する研究に従事．

中井 優志

1982 年生．2005 年岩手県立大学ソフトウェア情報学部ソフトウェア情報学科卒業．同年同大学大学院ソフトウェア情報学研究科に入学．VR，Peer-to-Peer オーバレイネットワークに関する研究に従事．

渡邊 集

1981 年生．2004 年岩手県立大学ソフトウェア情報学部ソフトウェア情報学科卒業．同年同大学大学院ソフトウェア情報学研究科に入学．情報隠蔽とモバイルエージェントに関する研究に従事．

川原 卓也

1981 年生．2004 年岩手県立大学ソフトウェア情報学部ソフトウェア情報学科卒業．同年同大学大学院ソフトウェア情報学研究科に入学．RFID におけるプライバシ保護に関する研究に従事．

加藤 貴司

1971 年生．2001 年東北大学大学院情報科学研究科博士後期課程修了．現在，岩手県立大学ソフトウェア情報学研究科講師．博士（情報科学）．マルチエージェントシステムにおけるエージェントの協調に関する研究に従事．人工知能学会，電子情報通信学会各会員．

ベッド バハドゥール ビスタ（正会員）

1967 年生．1991 年 York 大学電子工学科卒業．1997 年東北大学大学院情報科学研究科博士課程修了．1997 年から 1998 年宮城大学勤務．現在，岩手県立大学ソフトウェア情報学研究科助教授．博士（情報科学）．プロトコルの仕様記述と合成，モバイル通信に関する研究に従事．

高田 豊雄（正会員）

1962 年生．1989 年大阪大学大学院基礎工学研究科博士後期課程修了．現在，岩手県立大学ソフトウェア情報学研究科教授．工学博士．セキュリティと誤り制御通信に関する研究に従事．電子情報通信学会，情報理論とその応用学会，IEEE，ACM 各会員．