第２部

特集2 NECOMAプロジェクト：日欧協調によるマルチレイヤ脅威分析およびサイバー防御の研究開発

田崎 創、岡田 和也、長 健二郎、福田 健介、加藤 朗、関谷 勇司、門林 雄基

第１章　はじめに

NECOMA(Nippon-European Cyberdefense-Oriented Multilayer threat Analysis)プロジェクト[*1]は，WIDEプロジェクトに参加している5組織(奈良先端科学技術大学院大学，IIJ-II，国立情報学研究所，慶應義塾大学，東京大学)が2013年6月より総務省から委託を受けている「日欧協調によるマルチレイヤ脅威分析およびサイバー防御の研究開発」である．本プロジェクトは，総務省と欧州委員会のFP7(Framework Program 7)による日欧協調の研究開発プロジェクトであり，欧州側からはフランスのIMTを中心とする5つの大学・研究機関・民間企業が参画している(表1.1)．

本プロジェクトは，多種多様なサイバー脅威を複数のデータを分析しその影響を軽減する手法を研究開発することを目的としている．これまでにサイバー脅威データ収集，サイバー攻撃の検知，防御手法は，独立した研究・開発が多く行われてきた．NECOMAでは，データ収集，解析，防御の個別の研究開発に加え，これらを繋ぎ迅速にサイバー脅威へ対応できる仕組みを研究開発する．

インターネット上におけるサイバー攻撃は，規模もその手法も巧妙になってきている．そのため，迅速で効果的なサイバー脅威の検知とその防御が求められており，他者や国家に頼ることなく自らのネットワーク，サーバ，ユーザ端末を防御し，安全に運用することが肝要である．自律的なサイバー防御の実現には，自らが攻撃を検知，防御する手段，方法を確立しなければならない．また，異なる組織間でのサイバー脅威情報の共有，連携が重要である．

WIDEプロジェクトは，独自の広域インターネット網を保有し，ネットワークやクラウドサービス(WIDE Cloud)の運用をWIDEメンバ自らが行っている．また，日々発生するサイバー攻撃に対するインシデントレスポンスを行っている．従って，WIDEプロジェクトは自律的なサイバー防御手法を研究するには最適なインフラとサービスを有しており，NECOMAプロジェクトが目標とする技術を研究開発するにあたって十分な基盤と技術を有している．また，商用サービス・プロバイダでは，法律とプライバシ問題が深く関係するためトラフィックデータ，各種ログの収集・解析を容易に実施できない．WIDEプロジェクトがこれらに先駆けて研究開発を行うことは，今後のインターネットにおけるサイバー防御に資することである．

表1.1　NECOMA プロジェクト参画組織一覧

日本側	欧州側
奈良先端科学技術大学院大学	Institut Mines-Télécom（フランス）
IIJ イノベーションインスティテュート	Atos Spain S.A.（スペイン）
国立情報学研究所	Foundation for Research Technology - Hellas（ギリシャ）
慶應義塾大学	Research and Academic Computer Network (ポーランド)
東京大学	6cure（フランス）

＊1　http://www.necoma-project.jp

本報告では，NECOMAプロジェクトで取り組む課題の概要と今年度の研究成果を述べる．

第２章　プロジェクトの研究課題

NECOMAプロジェクトでは，4つの研究課題に取り組む．以下，それぞれの研究課題について概要を説明する．

課題1:サイバー攻撃の多層的な観測に関する研究

課題1では，サイバー防御に必要なデータの収集，データ形式の変換，組織間での共有を行うことを目的とする．これまでにもWIDEプロジェクトでは，MAWIWGを中心にネットワークトラフィックデータの収集を行ってきた．NECOMAプロジェクトでは，バックボーンネットワークに限らず，ユーザ端末を含むエンドポイントも含むデータ収集を行う．また，データの種別もトラフィックに限らずサイバー攻撃に関わる多層的なデータ収集を目的とする．本課題では，データを収集・保存するシステムの設計，実際のデータ収集及び収集手法の検討を行う．

課題2:サイバー攻撃に対する回復性のあるデータ解析

課題2では，課題1で収集した各種データを元にサイバー攻撃の検知をする複数の手法を開発することを目的とする．また，攻撃検知のためのデータ解析を行うための基盤構築も行う．データ解析では，バックボーントラフィック，DNSトラフィック解析といった個別のデータ解析に加えて，複数の異なるデータを用いたマルチレイヤデータの解析も行う．

課題3:サイバー防御に関する研究

課題3では，課題2の解析やCERTなどから報告されたサイバー脅威情報(DoS/DDoS，フィッシング，APTなど)を元に，ユーザ端末，ネットワーク機器で防御する手法，方式を研究開発する．

課題4:実証実験

課題4では，課題2及び課題3で開発された解析・防御手法を実環境及びエミュレーション環境において検証することを目的とする．

第３章　活動記録

●2013年6月30日 NECOMA日本側コンソーシアム・キックオフミーティング(慶應義塾大学 日吉キャンパス)
日本国内参加機関一同で，各自のこれまでの活動紹介と，今後の研究開発に向けての課題整理を実施した．

●2013年9月5日-6日 NECOMA日欧全体キックオフミーティング(フランス　パリ)
EU・日本双方の参加機関で今後の共同研究活動にむけて，各課題のリーダ，マイルストン・成果物のレビュー，各参加機関の活動紹介，プロジェクトマネジメント上の課題を洗いだし，議論・検討した．

●2013年7月24日 ハッカソン(2週間おきに開催)
NECOMAプロジェクト内での研究開発作業を，集中的に行うために，隔週でハッカソン・勉強会を実施している．ここでの成果を日々蓄積される観測データの解析，脅威分析へとフィードバックする．

●2013年9月30日-10月4日 ハッカソン(北陸StarBED技術センター)
課題1，課題2で利用するデータセットフォーマット，Hadoop基盤の作り込み・検証と，StarBED新規ノードであるグループNの予備検証を実施した．

●2013年11月6日 Gregg Schudel氏(Cisco)講演:LISP—A Next-Generation Networking Archi-tecture
米国CiscoのGregg Schudel氏に東京大学でLISP(Locator/ID Separation Protocol)について講演をしていただいた．講演では，LISPのプロトコル概要から具体的なユースケース，今後の展望についてお話を頂いた．

第４章　今年度の現状

4.1　マルチレイヤデータの収集(課題1)

課題1では，各参加機関の提供可能なデータセットの一覧を作成する所から研究開発を開始した．それを元に，データセット間で共通な属性・項目を洗いだしをする事で，解

析のためのインターフェース(API)の設計を開始した.

また，大量に蓄積される計測データを効率的に読み出し，解析できる仕組みとしてHadoopを元にした運用基盤を構築した.

●各参加機関の提供可能なデータセットとして，ダークネットトラフィック，ボット監視記録，トラフィック計測(netflow/sflow/pcap)，DNSクエリ記録などを整理し，重複測定項目や共通的に扱える属性の洗い出しを行った.

●データアクセスAPIの検討
当面のAPI設計方針として，既にポーランドNASKで研究開発が進んでいるn6 API[*2]をベースに拡張する事を決定した．n6 APIはRESTとJSONをベースで提供されているAPIで，蓄積した脅威分析データ(network security incident exchange)を取り出し可能とするものである.

●Hadoopクラスタ構築
NECOMAでは，膨大なトラフィックデータ，DNSクエリ記録などを解析するためにApache Hadoopをベースとした様々なデータを分散環境で並列処理可能な仕組みを構築・運用している．分散されるノードの構築，データの蓄積・変換，ソフトウェアの改善[*3]，解析モジュールの作成などを随時行っている．これまでに，各種トラフィック，DNSクエリログ，メールといったデータを収集し蓄積している(表4.1).

表4.1　Hadoop hdfs上に蓄積している観測データと件数
（2013/11/26時点）

	#records	Duration
sFlow	977,233,577	2013/10-
netFlow	349,152,099	2013/10-
DNS	3,374,450,688	2013/9 -
Spam	8,920	2012/6 -
Phishing	約 80,000(テストデータ)	

4.2　マルチレイヤデータ解析基盤の構築(課題2)

●ドメイン名生成アルゴリズム(ZeuS DGA)を利用したボット感染端末の観測
Domain Generation Algorithm(DGA)と呼ばれる，ボットの利用する自動生成されたドメイン名への問い合わせを観測データより検知し，これを起点としたボット感染後の活動を追跡する事で，感染端末の検知，悪性サイトのtake down等を目的とする．計測データの解析より，大学内のDNSキャッシュサーバにて，1日あたり数件のDGA(ZeuSアルゴリズム)による問い合わせが観測された．今後はこのクエリを起点とした感染後の活動を分析していく.

●DNSオープンリゾルバとの通信パケット量観測
DNSオープンリゾルバを利用したDDoS攻撃を検知することを目的として，sFlowデータとオープンリゾルバ(IPv4)のリストを用いて，オープンリゾルバとの通信パケット量を観測した．今後は，継続的にこのトラフィク量を分析し，大規模なDDoS攻撃が発生した際の攻撃活動を明らかにしていく.

●ネットワーク異常検知器の性能改善
バックボーンネットワークトラフィック内のアノマリ(異常)を検知において，複数検知器間での検知結果の重複を視覚的に分析するために，Chord図により可視化を試みた．過去4年間に計測したバックボーントラフィックを対象として解析した結果，19%の検知精度向上を実現した．また結果については国際学会SAC 2014にて発表予定である[3].

●大規模DNS分析
バックボーンネットワークで採取したDNSトラフィックに対し，NX domain error内のドメイン名にランダムテストを実施する事で，不正ドメイン名の検出を試みた．4000個の日本のSNS類似のドメインを乱雑さ試験により発見し，実態としては2台のサーバに誘導されるドメイン(うち1台は既知のブラックリストにも掲載されているもの)であることを確認した．また，これらはスパム送信に用いられる使い捨てドメインとして利用されている事も

＊2　http://n6.cert.pl/
＊3　https://github.com/necoma

確認した．この分析については，国際学会AINTEC 2013 [1]にて対外発表を実施した．

●悪性サイト閲覧時の視線追跡情報測定

フィッシングサイトからの利用者保護を目的として、Webサイト利用者が，正規サイトと悪性サイトにおける視線追跡の違いを観測する為に視線追跡装置を導入した．これらの観測情報を用いて，今後、利用者の特徴をとらえるフィッシング対策支援技術や、ブラウザのセキュリティ評価への応用を予定している．

4.3　レジリエントなサイバー防御手法の研究開発(課題3)

●LISPにおけるEID，RLOCを対応づけるMapサーバを改変し，DDoS/DoS等のサイバー攻撃のみを囮サーバに転送することにより，正規サーバへの影響を緩和する仕組みを提案，実装した．本研究成果は，IEEEのICC2014 CIISに投稿中である．

第５章　おわりに

NECOMAプロジェクトは，2013年度から2015年度までの3年間のプロジェクトである．2年目にあたる2014年は，課題2における複数のデータ解析に取り組むとともに，課題3のネットワークインフラ，エンドユーザ端末における防御技術の研究開発に取り組んでいく．本プロジェクトの成果は，各課題毎に研究論文を発表するとともに，外部の産業組織，学術組織がその知見を活かせるようにソースコードや運用方法を報告書やブログ，ステークホルダーを交えた報告会を通じて公開していく予定である．