「マルチメディア，分散，協調とモバイル
(DICOMO2021) シンポジウム」 令和3年6月-7月

データドリブンなサイバーセキュリティ研究の最前線

井上大介 [1]

概要：東京 2020 オリンピック・パラリンピック競技大会を間近に控え，サイバー攻撃への対処能力の向上が喫緊の課題となっている．情報通信研究機構（NICT）は日本のセキュリティ向上を目指し，サイバーセキュリティ分野の研究開発に取り組んでいる．この分野の研究開発ではサイバー攻撃の実データを大規模観測・蓄積・分析し，対策導出を行うことが肝要であり，データ駆動型の研究分野といえる．本稿では NICT におけるデータドリブンなサイバーセキュリティ研究の最前線を概説する．

1. はじめに

サイバーセキュリティ分野では，ネットワークで発生する実際のサイバー攻撃について，1) 様々な観測技術を用いて大規模に観測し，2) その観測データを蓄積するとともに，3) リアルタイムに分析して，4) 対策に繋げていく，というデータ駆動型のオペレーションが今も昔も変わらぬ王道的手法となっている．したがって，サイバーセキュリティ分野の研究開発においても，サイバー攻撃の実データをいかに大規模に収集・蓄積するかが国際的な研究開発競争の第一歩となっている．

サイバー攻撃は，ネットワークに広く拡散を試みる無差別型攻撃と，ターゲットとなる組織を絞った標的型攻撃に大別でき，観測手法も大きく異なる．情報通信研究機構（NICT）では無差別型攻撃対策のための Global 観測（全域的観測）と，標的型攻撃対策のための Local 観測（局所的観測）の両方の研究開発に取り組んでいる（図 1 横軸）．

また，サイバー攻撃を Passive（受動的）に待ち受けて観測する手法から，攻撃元とインタラクションを重ねる Active（能動的）な手法まで広くカバーしている（図 1 縦軸）．次章以降で代表的なシステムについて概説する．

2. インシデント分析センタ NICTER

NICTER（ニクター）は，ダークネットと呼ばれる未使用の IP アドレス空間の大規模観測（本稿執筆時点で約 30 万アドレス）を行っている[1]．未使用の IP アドレスに対しパケットが送信されることは，通常のインターネット利用の範囲においては稀であるが，実際にダークネットを観測してみると，相当数のパケットが到着することが分かる．これらのパケットの多くは，ワーム型マルウェアが感染拡大する際に，攻撃対象を自動探索するためのスキャンと呼ばれる通信である．そのため，ダークネットに到着するパケットを大規模に観測・分析することで，マルウェアによる無差別型攻撃の大局的な活動傾向の把握が可能になる．

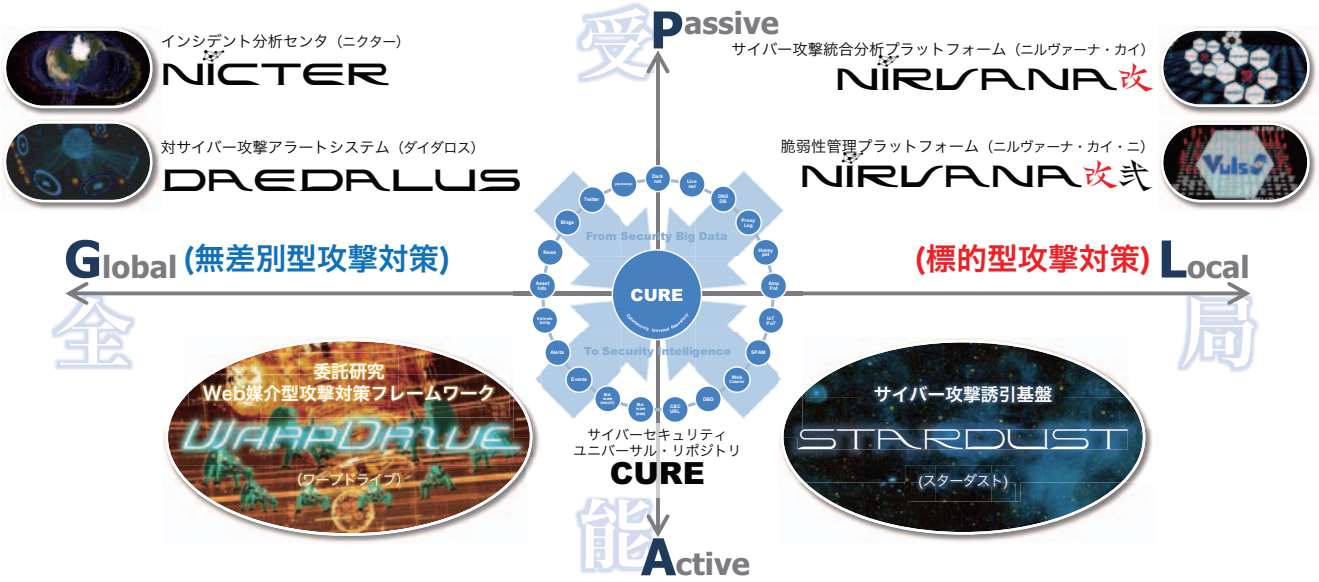

図 1 NICT におけるサイバーセキュリティ研究の全体像

1 国立研究開発法人情報通信研究機構 サイバーセキュリティ研究所
サイバーセキュリティネクサス

図 2 は NICTER のダークネット観測網に届くパケットを可視化したものである．

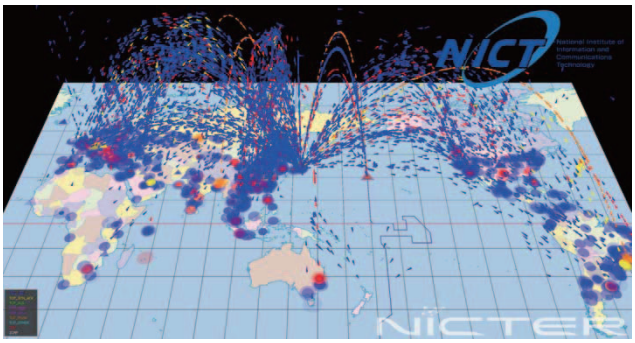

図 2 NICTER

表 1 は過去 10 年間の NICTER のダークネット観測統計である．2020 年は通年で約 5,001 億パケットの攻撃関連通信が観測された[2]．平均すると，1 年間に 1 つの IP アドレスに対して約 182 万パケットの攻撃関連通信が届いた計算となる．この 1 IP アドレス当たりの年間総観測パケット数は，2013 年頃より増加傾向が続いている．

表 1 ダークネット観測統計（過去 10 年）

年	年間総観測パケット数	ダークネットIPアドレス数	1 IPアドレス当たりの年間総観測パケット数
2011	約45.4億	約12万	40,654
2012	約77.8億	約19万	53,085
2013	約128.8億	約21万	63,655
2014	約256.6億	約24万	115,323
2015	約545.1億	約28万	213,523
2016	約1,281億	約30万	469,104
2017	約1,504億	約30万	559,125
2018	約2,121億	約30万	789,876
2019	約3,220億	約30万	1,187,935
2020	約5,001億	約30万	1,820,722

図 3 は 2020 年に NICTER で観測したパケットを，宛先ポート番号（インターネット上でサービスを表す番号）別に分類したものである．ただし，調査目的のスキャンと判定されたパケットは除いている．

攻撃対象として最も多かった宛先ポート番号は 23/TCP であり，これは Telnet と呼ばれるリモート接続のために使われているポート番号である．Web カメラやブロードバンドルータ等の多くの IoT 機器が，Telnet サービスを不用意に（弱い ID とパスワードで）インターネット側に公開しており，IoT 機器がマルウェア感染する要因の一つとなっている．

2 番目に多く観測された 445/TCP は，Conficker[3]や WannaCry[4]が感染拡大に利用する Windows のファイル共有等で使われる Server Message Block（SMB）のポート番号である．445/TCP は全体の 5%程度に留まっており，無差別型サイバー攻撃の対象が Windows から IoT 機器に大きく変わって来ていることが分かる．

3 番目以降も 2019 年とほぼ同様の傾向であり，IoT 機器の管理用インターフェイスを提供する Web サーバが動作する 80/TCP や 8080/TCP，81/TCP，サーバ等の遠隔操作で使用される SSH（Secure Shell）の 22/TCP など，IoT 機器を狙った攻撃活動に関連するポート番号が引き続き上位に観測された．

図 3 において青色で塗られたポート番号は過去に IoT 機器で使用されたことがあるものであり，上位 30 ポートのうち 36.9%を占めていた．2020 年の無差別型攻撃の対象は IoT 機器が支配的であり，その傾向は 2021 年も続いている．

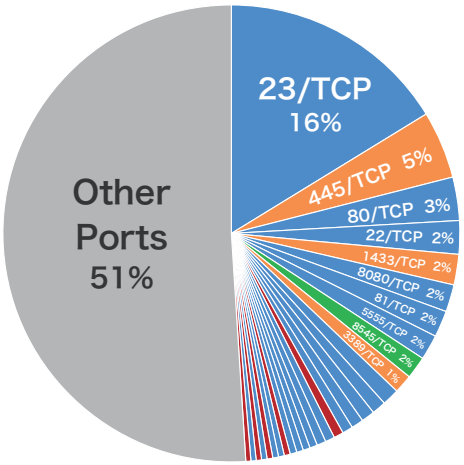

図 3 宛先ポート番号別パケット数（2020 年）

3. 対サイバー攻撃アラートシステム DAEDALUS

DAEDALUS（ダイダロス）は，NICTER の大規模ダークネット観測網を応用したアラートシステムである[5]．その動作原理は非常にシンプルで，特定の組織内でサーバやホストが使用している IP アドレス空間（以下，ライブネット）から NICTER のダークネット観測網にパケットが送信されると，当該組織に向けてアラートを発報するというものである．図 4 はある組織（画像中央下部のリング状のオブジェクト）から大量のスキャン（黄色）が送出され，DAEDALUS アラートが発報されている様子である．2021 年 4 月現在，700 を超える地方自治体に対し DAEDALUS アラートを無償提供している．

図 4 DAEDALUS

4. サイバー攻撃統合分析プラットフォーム NIRVANA 改

2011 年に国内の防衛関連企業への侵入と情報窃取を契機に明るみになった標的型攻撃は，特定の組織をターゲットに絞って攻撃を行うため，NICTER のような大規模観測網では捉えられないサイバー攻撃であり，現在も重大な脅威となっている．NIRVANA 改（ニルヴァーナ・カイ）は，このような標的型攻撃に対抗するため，組織内部にセンサを設置し，ライブネットを流れるパケットをリアルタイムに観測するとともに，各種のセキュリティ機器から発報されるアラートを集約し，分析・可視化・トリアージを可能にする統合分析プラットフォームである．図 5 は IPv4 アドレス空間上に各種セキュリティ機器のアラート（六角形）を可視化したものである．

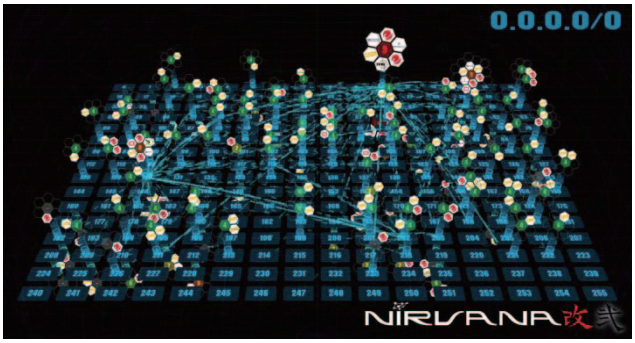

図 5 NIRVANA 改

5. Web 媒介型攻撃対策フレームワーク WarpDrive

WarpDrive（ワープドライブ）は，Web サイトを閲覧するだけでマルウェアに感染するような Web 媒介型攻撃の実態把握と対策技術の向上を目指したユーザ参加型プロジェクトである．本稿執筆時点で延べ 14,000 名以上のユーザがプロジェクトに参画しており，1 日あたり 500〜1000 万の URL アクセスを収集している．図 6 は WarpDrive のポータルサイト[6]である．

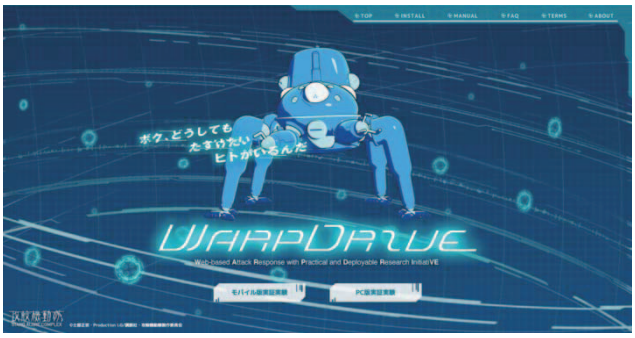

図 6 WarpDrive ポータルサイト

©士郎正宗・Production I.G/講談社・攻殻機動隊製作委員会

WarpDrive ではアニメ「攻殻機動隊 S.A.C.」シリーズに登場するキャラクター「タチコマ」をモチーフに，Web 媒介型攻撃対策ソフトウェア「タチコマ・セキュリティ・エージェント」（以下，タチコマ SA）を開発した．タチコマ SA は，1) ユーザの Web ブラウザの中で Web 媒介型攻撃の観測・分析を行い，2) 攻撃検知時には悪性 Web サイトの閲覧をブロックし，3) ユーザに警告やアドバイスを行う．さらに，インターネット上に分散したタチコマ SA たちが，4) 並列化（情報集約・横断分析・新機能展開等）を繰り返し，最新の Web 媒介型攻撃に対応する．

6. サイバー攻撃誘引基盤 STARDUST

STARDUST（スターダスト）は，標的型攻撃を行う人間の攻撃者の挙動を観測・分析するためのサイバー攻撃誘引基盤である[7]．攻撃者を外部から誘引し長期分析するために，STARDUST は企業サイズの精巧な模擬環境“並行ネットワーク”を自動構築する．そして，標的型攻撃で使用されるマルウェアを並行ネットワーク内で実行し，外部の攻撃者を誘き寄せることで，ステルス性の高いリアルタイム観測・分析を行い，標的型攻撃に関する貴重な実データを得ることを可能にする．図 7 は STARDUST の概念図である．

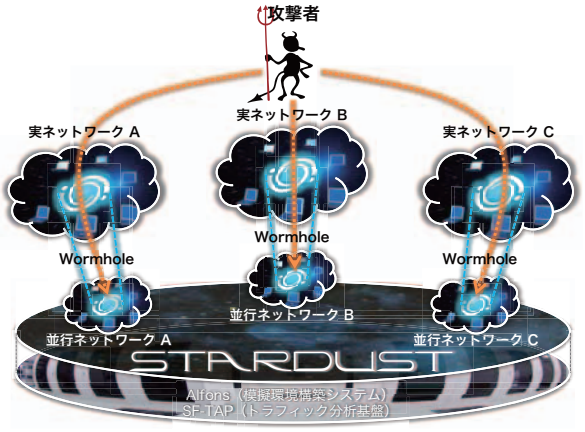

図 7 STARDUST 概念図

7. セキュリティ情報融合基盤 CURE

NICT では無差別型攻撃や標的型攻撃の観測，組織内のアラートやエンドポイント情報の収集，さらには様々な情報源からの脅威情報の取得など，多種多様なサイバーセキュリティ関連情報の収集を行ってきた（表 2）．セキュリティ情報融合基盤 CURE（キュア）は，これらサイバーセキュリティ関連情報を一元的に集約し，異種情報間の横断分析を可能にするシステムである[8]．CURE によって，個別に散在していた情報同士を自動的につなぎ合わせることが可能となり，これまで把握が困難であったサイバー攻撃の隠れた構造の解明につながる．

表 2 NICT が収集するサイバーセキュリティ関連情報

カテゴリ	蓄積データの具体例
ダークネット関連情報	未使用IPアドレス空間で観測したパケット、その統計情報、など
ライブネット関連情報	NICT内部のトラフィックやフロー情報、など
アラート情報	NICT内部のセキュリティ機器群のアラート情報、など
エンドポイント情報	NICT内部のPC端末内のプロセス情報、通信履歴、感染情報、など
マルウェア関連情報	マルウェア検体、静的解析結果、動的解析結果、など
スパム関連情報	スパム（ダブルバウンス）メール情報、など
Android関連情報	Android APK、カテゴリや説明文などアプリのメタデータ、など
ブログ・SNS情報	セキュリティベンダブログ、ツイート、など
Webクローラ取集情報	URLリストやWebコンテンツ、それらの評価結果、など
Webアクセス情報	ユーザのブラウザからのWebアクセス情報、など
ハニーポット収集情報	高対話型/低対話型/DRDoSハニーポット観測情報、など
脅威情報	有償/無償の脅威情報、IP/URLレピュテーション、C&C情報、など

図 8 は CURE の可視化エンジンであり，中央水色の球体が CURE 本体, 外周青色と橙色の小球体はそれぞれ Artifact（観測情報）と Semantics（分析情報）を格納するデータベース群である．CURE 本体では IP アドレス，ドメイン，マルウェア，自然言語のタグにより横断分析を行い，同一の情報が見つかるとデータベース間にリンクを描画する(青: IP アドレス，緑：ドメイン，橙：マルウェア，赤：タグ)．

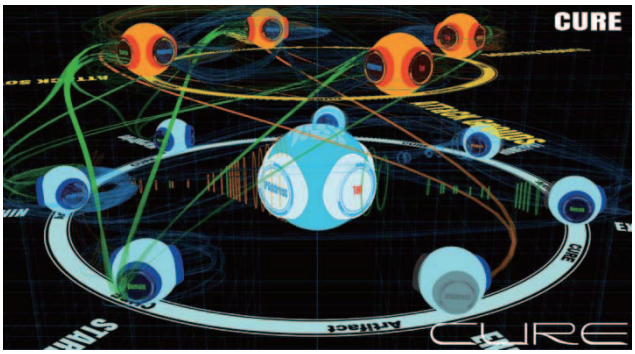

図 8 CURE

8. 機械学習とサイバーセキュリティの融合

NICT では機械学習をサイバーセキュリティに応用し，各種の自動分析技術やセキュリティ・オペレーションの自動化について融合研究を行っている．図 9 に示す通り，研究課題として，1）インシデントの優先順位判定，2）マルウェア機能分析自動化，3）攻撃の検知・脅威予測の 3 つのカテゴリに取り組んでいる．以下，これまでの研究成果を一部紹介する．

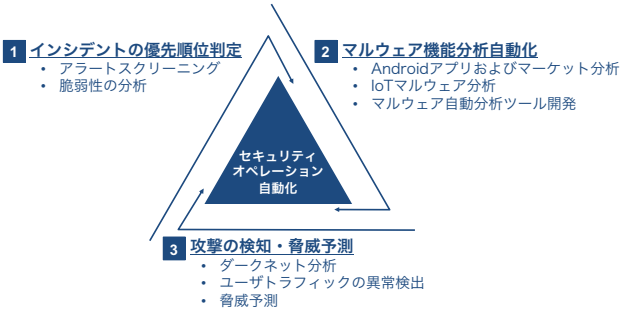

図 9 機械学習とサイバーセキュリティの融合研究

インシデントの優先順位判定では，Isolation Forest と呼ばれる手法を用いてセキュリティ機器から出されるアラートを分類し，87.4%のアラートの削減を達成した[9]．マルウェア機能分析自動化では，多層パーセプトロン（MLP）を用いて，Android マルウェアを 99.8%の精度で検出することに成功した[10]．攻撃の検知・脅威予測では，マルウェアの感染活動の同期性を Graphical Lasso（GLASSO）と呼ばれる手法を用いて 97.1%の精度でリアルタイムに検出し，感染拡大の早期検知の可能性を示した[11].

9. おわりに：データ負けからの脱却に向けて

サイバーセキュリティの研究開発には実データの大規模な収集・蓄積が必須であるが，それが高い参入障壁となり，日本の多くの組織が「データ負け」の状況に陥っている．その結果，研究開発が停滞し，サイバーセキュリティ自給率の低迷[12]を招いている．この状況を打破するため，NICT では 2021 年 4 月に CYNEX（サイネックス：サイバーセキュリティネクサス）という新組織を立ち上げ，サイバーセキュリティ情報を国内で収集・蓄積・分析・提供するとともに，社会全体でサイバーセキュリティ人材を育成するための共通基盤を構築し，産学官の結節点として開放する活動を開始した（図 10）．今後，CYNEX を通じて国産のサイバーセキュリティ技術の創出およびセキュリティ人材育成のさらなる発展に寄与していく．

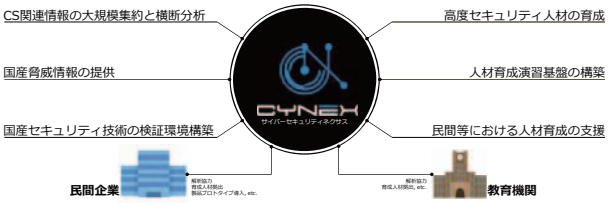

図 10 CYNEX の概要

参考文献

[1] NICTER Web. https://www.nicter.jp
[2] 情報通信研究機構サイバーセキュリティ研究室, “NICTER 観測レポート 2020,” 2021. https://www.nict.go.jp/cyber/report/NICTER_report_2020.pdf
[3] 井上大介, “大規模感染型マルウェア Conficker の動向調査とダークネット観測事例,” IPA 情報セキュリティ技術動向調査（2009 年上期），2009.
[4] piyokango, “世界各地で発生したランサムウェア WannaCry の感染事案についてまとめてみた,” piyolog, 2017. https://piyolog.hatenadiary.jp/entry/20170513/1494700355
[5] Daisuke Inoue, Masashi Eto, Koei Suzuki, Mio Suzuki, Koji Nakao, “DAEDALUS-VIZ: Novel real-time 3D visualization for darknet monitoring-based alert system,” 9th International Symposium on Visualization for Cyber Security (VizSec 2012), Oct 2012.
[6] WarpDrive. https://warpdrive-project.jp

[7] 津田 侑, 遠峰 隆史, 金谷 延幸, 牧田 大佑, 丑丸 逸人, 高野 祐輝, 安田 真悟, 三浦 良介, 太田 悟史, 宮地 利幸, 神薗 雅紀, 衛藤 将史, 井上 大介, 中尾 康二, “サイバー攻撃誘引基盤 STARDUST,” マルウェア対策研究人材育成ワークショップ 2017 (MWS2017), 2A2-1, 2017.
[8] 津田 侑, 井上 大介, 鈴木 宏栄, 高木 彌一郎, 田中 秀一, 金谷 延幸, 竹本 亜希, 古本 啓祐, “セキュリティ情報融合基盤 CURE,” マルウェア対策研究人材育成ワークショップ 2020 (MWS2020), 2C5-1, 2020.
[9] Muhamad Erza Aminanto, Tao Ban, Ryoichi Isawa, Takeshi Takahashi, Daisuke Inoue, “Threat Alert Prioritization Using Isolation Forest and Stacked Auto Encoder With Day-Forward-Chaining Analysis,” IEEE Access, Volume 8, pp. 217977-217986, 2020.
[10] Bo Sun, Tao Ban, Shun-Chieh Chang, Yeali S. Sun, Takeshi Takahashi, Daisuke Inoue, “A Scalable and Accurate Feature Representation Method for Identifying Malicious Mobile Applications,” 34th ACM/SIGAPP Symposium (SAC 2019), 2019.
[11] Chansu Han, Jumpei Shimamura, Takeshi Takahashi, Daisuke Inoue, Jun'ichi Takeuch, Koji Nakao, “Real-Time Detection of Global Cyberthreat Based on Darknet by Estimating Anomalous Synchronization Using Graphical Lasso,” IEICE Transactions on Information and Systems, Volume E103.D, Issue 10, pp. 2113-2124, 2020.
[12] サイバーセキュリティ戦略本部 研究開発戦略専門調査会, “サイバーセキュリティ研究・技術開発取組方針,” 2019. https://www.nisc.go.jp/conference/cs/kenkyu/dai12/pdf/kenkyu_torikumi.pdf