国立研究開発法人
情報通信研究機構

プレスリリース
2022年6月14日
国立研究開発法人情報通信研究機構

セキュリティ情報融合基盤“CURE”のデータエンリッチメント機能を開発

～『Enricher』(エンリッチャ)によるセキュリティ情報分析の更なる高度化を実現～

【ポイント】

- サイバーセキュリティ関連情報を大規模集約・横断分析する「CURE」に新たな情報を融合
- 新機能『Enricher』(エンリッチャ)を開発し、CUREのセキュリティ情報分析能力を更に強化
- CUREの情報融合促進と分析能力強化により、質の高い国産の脅威情報の生成を加速

国立研究開発法人情報通信研究機構(NICT、理事長: 徳田 英幸)サイバーセキュリティ研究室は、サイバーセキュリティ関連情報を大規模集約するセキュリティ情報融合基盤「CURE[*1]」(キュア)の新機能として、蓄積したデータに様々な付加情報を与えることで、データの有用性を向上させるデータエンリッチメント機能を開発するとともに、新たに複数の情報源をCUREに融合しました。これにより、CUREを用いたサイバー攻撃の実態把握の精度が向上し、より質の高い国産脅威情報の生成が期待できます。

機能強化したCUREは、2022年6月15日(水)～17日(金)に幕張メッセで開催される「Interop Tokyo 2022」で動態展示を行います。

【背景】

サイバー攻撃の実態を把握するためには、サイバー攻撃の観測情報や外部機関が公開するセキュリティレポートなど、多種多様なサイバーセキュリティ関連情報を集約し、多角的に分析する必要があります。そのため、NICTはセキュリティ情報融合基盤「CURE」を開発し、サイバーセキュリティ関連情報の大規模集約と横断分析の研究を行ってきました。

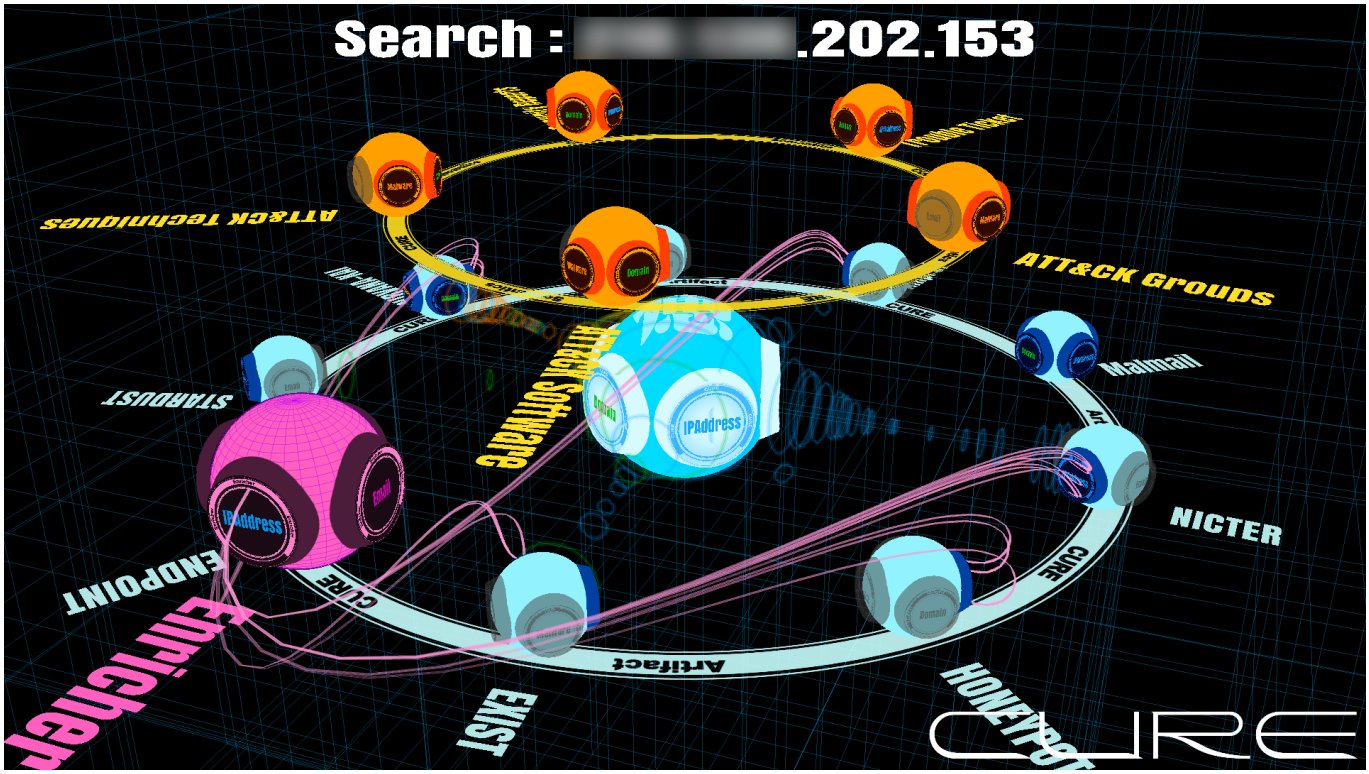

図1 CURE全体図(『Enricher』による類似IPアドレスの検出)

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観測情報(Artifact)と分析情報(Semantics)を格納するデータベース(DB)群。桃色の球体は付加情報を与えるEnricherで、類似の挙動を示すIPアドレス群を検出。

これまで CURE は様々な観測情報（Artifact）や分析情報（Semantics）を融合してきましたが、サイバー攻撃の実態把握の精度を向上するために、CURE に格納するサイバーセキュリティ関連情報の「量」と「質」を継続的に高めていくことが課題でした。データの量を増やすためには、新たな情報源の融合が効果的であり、データの質を向上させるためには、データに付加的な情報を与えることで、その有用性や信頼性を向上させるデータエンリッチメントの仕組みが必要でした。

【今回の成果】

今回、CURE に新たに 3 つの情報源（AmpPot[*2]、Malmail[*3]、Trouble Ticket[*4]）を融合しました。AmpPot は DDoS 攻撃の一種であるリフレクション攻撃の観測情報、Malmail はメールに添付されたマルウェアの動的解析情報、Trouble Ticket は組織内のインシデント対応に関する管理情報です。CURE は観測情報を格納する Artifact レイヤと分析情報を格納する Semantics レイヤの 2 層から構成されますが、AmpPot と Malmail は Artifact レイヤに、Trouble Ticket は Semantics レイヤにそれぞれ追加しました（図 2、図 3 参照）。

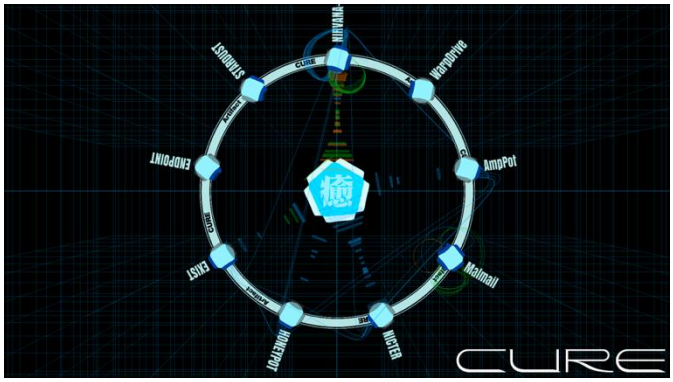

図 2 Artifact レイヤ
観測情報を格納する Artifact レイヤ。新たな情報源として AmpPot と Malmail を融合（2 時～4 時方向）

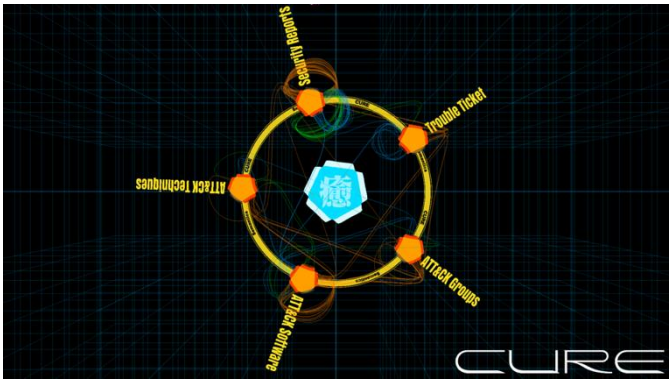

図 3 Semantics レイヤ
分析情報を格納する Semantics レイヤ。新たな情報源として Trouble Ticket を融合（2 時方向）

また、サイバー攻撃の痕跡情報（IoC: Indicator of Compromise）として、これまで IP アドレス、ドメイン名、マルウェア情報を用いていましたが、攻撃に悪用されたメールアドレスも IoC として扱えるように機能追加し、メールアドレスによるデータの関連付けや検索が可能になりました。

さらに、CURE に格納されたデータに対して付加的な情報を与えるデータエンリッチメントの仕組みとして、今回初めて『Enricher』（エンリッチャ）を開発しました。データエンリッチメントのコンセプトは広く、様々な応用が考えられますが、今回は概念実証として、CURE 内のデータに Doc2Vec[*5] を用いて類似度スコアを付与することで、類似の挙動を示す IP アドレス群を検出する Enricher を実装しました（図 1、図 4 参照）。

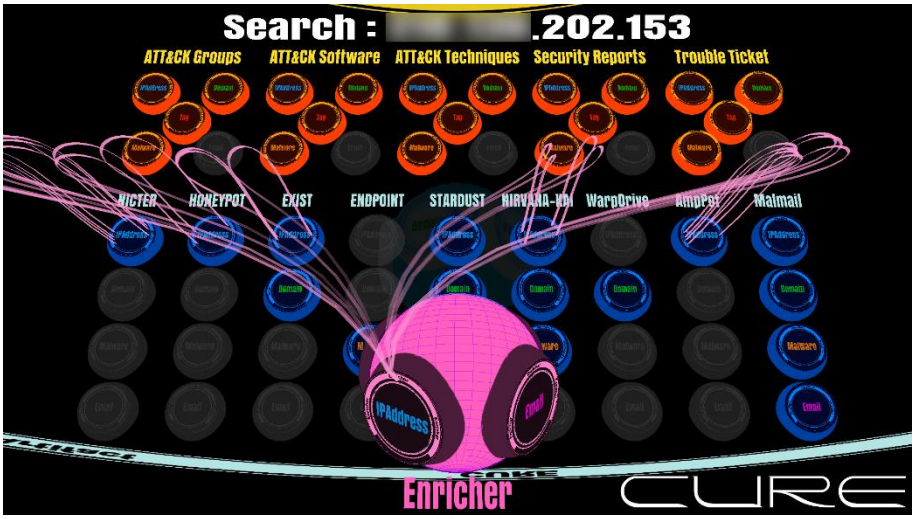

図 4 詳細情報表示（Enricher による類似 IP アドレスの検出）

これまで CURE は、完全一致する IoC でデータの関連付けを行っていましたが、Enricher によって「この IP アドレスと似た活動をしている IP アドレス」のように、より柔軟な関連付けができるようになりました。これにより、大規模スキャナによる調査活動に用いられている IP アドレス群の網羅的な把握や、暗号資産を不正に採掘するマルウェアの接続先 IP アドレス群の抽出に成功するなど、データの完全一致だけでは実現できなかった柔軟な横断分析が可能になりました。

データエンリッチメントはデータの質を向上させるために付加情報を与えるという広いコンセプトであり、今回実装した類似 IP アドレス検出だけでなく、Enricher によって IP アドレスやドメイン名に悪性度スコアを付与することや、解析者によるデータの信頼性の評価などを付与することなど様々な応用が可能であり、CURE に集約されたサイバーセキュリティ関連情報の質を向上させ、情報分析能力を更に強化することができます。

【今後の展望】

データエンリッチメント機能の追加により、CURE を用いたサイバー攻撃の実態把握の精度が向上し、より質の高い国産脅威情報の生成が期待できます。機能強化した CURE は、2022 年 6 月 15 日（水）～17 日（金）に幕張メッセで開催される「Interop Tokyo 2022」で動態展示を行います。

＜用語解説＞

***1 CURE（キュア）**

CURE（Cybersecurity Universal REpository）は、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セキュリティ情報融合基盤。個別に散在していた情報同士を自動的につなぎ合わせ、サイバー攻撃の隠れた構造を解明し、リアルタイムに可視化する。

報道発表「セキュリティ情報融合基盤“CURE”を開発」
2019 年 6 月 6 日
https://www.nict.go.jp/press/2019/06/06-1.html

報道発表「セキュリティ情報融合基盤 “CURE” を機能強化！」
2020 年 10 月 27 日
https://www.nict.go.jp/press/2020/10/27-1.html

***2 AmpPot**

リフレクション攻撃（DRDoS: Distributed Reflection Denial-of-Service）を観測するハニーポット。横浜国立大学吉岡研究室と NICT サイバーセキュリティ研究室との共同研究・共同運用を行っている。

***3 Malmail**

NICT に届いた悪性メールに関する情報と、添付されたマルウェアをサンドボックス（箱庭環境）で動的解析した結果を集約するシステム。

***4 Trouble Ticket**

NICT 内のインシデント対応に関する情報を管理するためのチケット管理システム。インシデント対応の進捗状況や詳細解析情報を集約するとともに、解析者が登録したタグ（インシデントに関連した単語）を CURE にフィードバックする機能を有する。

***5 Doc2Vec**

教師なし学習により、任意の長さの文書から固定長の文書ベクトルを得るアルゴリズム。文書ベクトル化することにより、内積計算で類似度が算出できるなど、文書の処理が容易になる。

＜ 本件に関する問合せ先 ＞
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研究室
井上 大介、津田 侑、鈴木 宏栄
E-mail: nicter@ml.nict.go.jp

＜ 広報（取材受付）＞
広報部 報道室
E-mail: publicity@nict.go.jp