6　セキュリティアーキテクチャ技術
6-1　セキュリティ情報のディスカバリ技術とそれを用いた知識ベースの構築

高橋健志

本稿ではサイバーセキュリティに関する各種情報の共有を実現すべく、インターネット上の構造化情報のディスカバリ技術と同時にそれに基づく知識ベースを構築する方式を提案する。提案方式は、ネットワーク上に分散している各種サイバーセキュリティ情報をリンクし、横断検索を実現するものである。本方式はその情報構造に特徴があり、柔軟性と拡張性を兼ね備えている。また、本稿ではプロトタイプ実装についても報告する。

1　はじめに

サイバーセキュリティを担保すべく、組織・国の壁を越えた情報共有が求められている。それを実現すべく、各種機関がインターネット上で各種サイバーセキュリティ情報を公開している。その代表的なものとして、National Vulnerability Database（NVD）[1] や Japan Vulnerability Notes（JVN）[2] などが存在し、今後、より多くの情報が世界中の組織から提供されてくることが期待されている。しかしながら、現時点においてはこれらの情報の存在をすべて把握すること、そして大量に提供される情報の中から自分に必要な情報のみを抽出して活用することは困難である。情報共有を促進するためには、ユーザはこれらの情報の所在を把握し、そこから必要な情報を発見・特定し、取得できるディスカバリ技術が求められている。

本稿では、ネットワーク上に存在する各種サイバーセキュリティ情報を特定し、かつその検索、交換を可能とするディスカバリ技術を提案する。提案方式は、検索に用いるメタ情報の構造に特徴があり、カテゴリとフォーマットを分けて定義することにより柔軟性と拡張性を実現しており、そのカテゴリには著者の従来研究であるサイバーセキュリティ情報オントロジにて定義された情報カテゴリ [3] を利用し、フォーマットについては各種団体により規格化されたスキーマを用いている。

また、本稿では提案方式のプロトタイプ実装も紹介する。本実装は、インターネット上の各種レポジトリを横断検索することが可能である。なお本稿は著者の文献 [4] の要約であり、詳細については当該論文を参照いただきたい。

2　関連研究

サイバーセキュリティに関する情報を組織間で交換・共有するためには、共通のフォーマットが必要である。既に各種団体がサイバーセキュリティ情報の構造化記述手法を検討しており、例えば、脆弱性情報の識別子とその XML 記述手法を定義した CVE[5] をはじめ、ARF[6]、CAPEC[7]、CCE[8]、CEE[9]、CPE[10]、CRF[11]、CVRF[12]、CVSS[13]、CWE[14]、CWSS[15]、IODEF[16]、MAEC[17]、OCIL[18]、OVAL[19]、そして XCCDF[20] などが存在する。

また、ネットワーク上の情報を特定する手法としては、その代表的なものとして、RDF[21] が存在する。RDF は、リソースの情報を記述し、インターネット上でそのリソースを特定・検索可能にする W3C 規格であり、これを用いて任意のエンティティを記述可能である。RDF の検索エンジン機能を実現するものとして、SPARQL[22] が定められており、各種の実装も存在している。

上述の技術を用い、本稿では各種セキュリティ情報を特定し、かつそれらを検索できる方式を提案する。

3　提案方式の設計理念

提案方式は、ネットワーク上に存在する各種サイバーセキュリティ情報を構造化し、それを特定、検索、交換することができ、下記の基本方針に基づいて構築されている。

a）検索対象は XML 形式の情報のみ：既に CVE や CAPEC のように XML にて記述される各種情報が存在しており、提案方式は、これら XML 形式

で提供される情報のみを扱う。また、フリーキーワード検索に加えてタグベースの検索を提供する。

b）拡張性を維持：XML 形式のサイバーセキュリティ情報のスキーマ数は現時点では限られているものの、今後増加することが予想される。よって、将来のスキーマもサポート可能な拡張性を維持する必要がある。

c）スケーラビリティを担保：情報量は現時点では限定的であるものの、今後増大が見込まれている。よって、情報量によらず、必要な情報を発見できる必要がある。

d）既存の XML 形式の情報をそのまま活用：情報を適切に検索するためには新たなプロトコルが必要であるが、オンラインでの情報提供元が現在保持する情報自体には何の修正も要さないことが、実装展開を考える際に重要となる。

e）提案方式自体のセキュリティ担保技術についてはスコープ外：暗号や認証技術など、多種多様な技術が既に存在しており、提案方式を実運用する際にはそれらの技術を組み込む必要があるが、本稿のスコープ外とする。

4　アーキテクチャ

提案方式には図 1 のとおり、D-Client、D-Server、Registry、そして InfoSource という 4 つのロールが存在する。各ロールについて、下記に詳述する。なお、1 つのエンティティが複数のロールを兼ねるケースも存在することに留意されたい。

- D-Client: 本ロールは D-Server と通信をしてサイバーセキュリティ情報を検索する。なお、必要に応じて 1 つ以上の D-Server と通信する。
- D-Server: 本ロールは、D-Client に対するインターフェースであり、D-Client からの要求に従い適切な InfoSource の URI を検索する。その過程で D-Server は 1 つ以上の Registry と通信し、それらからの返信を集約する。
- Registry: 本ロールは InfoSource に対するインターフェースであり、InfoSource からの登録要求に対して、その InfoSouce に関するメタデータを収集・蓄積している。本メタデータにより、InfoSource 内部にある情報を特定できる。
- InfoSource: 本ロールは、XML 形式にて記述されたサイバーセキュリティ情報を保持する。そして、本情報をネットワーク上で発見可能にすべく、1 つ以上の Registry に自身の情報を登録する。

5　情報の構造

本節では、**4** にて定義したアーキテクチャ内にて利用するデータ構造を規定する。提案方式はサイバーセキュリティ情報を発見する。それをネットワーク上で実現するには、その情報が機械可読である必要がある。そのため、提案方式は各種の既存情報フォーマットを利用する。しかしながら、様々な情報フォーマットが存在し、それらは今後変更され、更なる発展を遂げることが予想されている。

そのフォーマットの拡張性を担保すべく、提案方式では、データ構造を情報のカテゴリとフォーマットの 2 段階に分けて定義する手法を提案する。情報カテゴリは抽象度が高く、長期間変更を要さないものを設定すべく、文献 [3] のオントロジにて規定した情報カテゴリを利用する。また、フォーマットは、実際に利用しやすいように明示的に定義されたスキーマを利用すべく、各種業界規格で定義されたスキーマを活用する。

図 2 に、提案方式のデータ構造を示す。InfoSource に関するすべての情報は、URI をキーとして保存されており、その URI は、表現フォーマット別に整理され、そのフォーマットは情報カテゴリ別に整理される形になっている。また、各 URI には、少なくとも InfoSource の情報登録の日時を示すタイムスタンプ（timestamp）と検索に必要なメタデータ（metadata）が紐づけられる形になっている。

提案方式は、この 2 段階の構造により、そして、RDF にて実装することにより、将来拡張性を担保している。将来的に新たなスキーマを活用したい際には、そのスキーマを上記の情報カテゴリのいずれかにリンクするのみでよく、そのリンクは RDF で管理されている情報構造では実現も容易である。

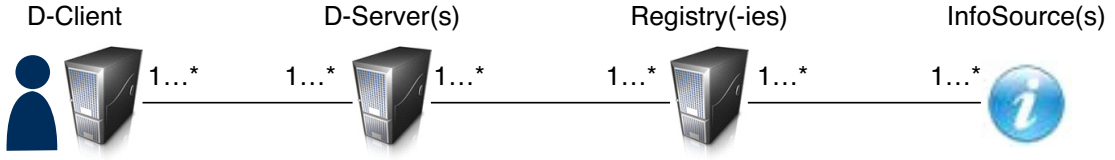

図 1　提案方式を構成するロール群

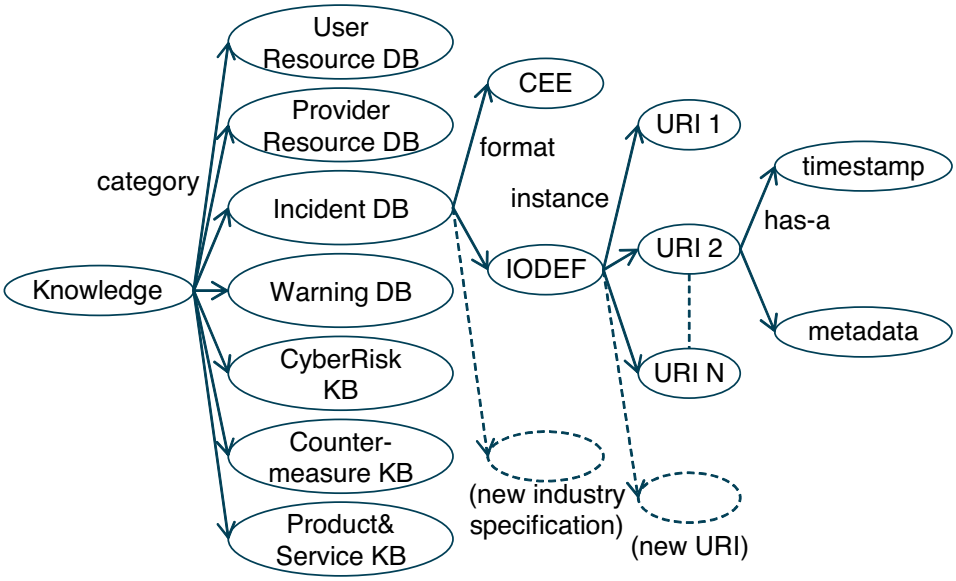

図 2　カテゴリとフォーマットから構成される 2 段階のデータ構造

6　プロトコル

本節では、提案方式が情報を発見するのに必要な 3 種類の手続き、すなわち情報の登録、サーバの登録とキャンセル、そして情報の検索の手続きを定義する。

6.1　情報の登録

本手続きは、InfoSource が情報を公開する際に必須となる。InfoSource は Registry に対し registration メッセージを送る。本メッセージは InfoSource の URI と情報カテゴリに関する情報を保持している。本メッセージを受信すると、Registry はその URI にアクセスし、その InfoSource 内にある XML 形式の情報を獲得する。Registry はその情報に対して XSLT を実施し、その InfoSource に関する情報のメタデータを RDF 形式にて生成し、保存・更新する。

メタデータが保存・更新されると、Registry は notification メッセージを D-Server へ送る。本メッセージには、メタデータの更新情報（その InfoSource の URI を含む）が含まれている。D-Server は、登録されている D-Client に対し、その notification メッセージを送信することも可能であり、それにより、D-Client はセキュリティ情報の更新をより早く獲得することが可能になる。

6.2　サーバの登録と登録削除

サーバ登録手続きは、Registry が D-Server を選択する際に必要となる。Registry は、利用したい D-Server に対し join メッセージを送信する。それを受信した D-Server はサポートしている情報フォーマットと、それが属する情報カテゴリに関する情報を result メッセージに埋め込んで返信する。なお、本稿では、文献 [3] のオントロジに従う単一のカテゴリを提案しているが、この result メッセージで返信するカテゴリと情報フォーマットに任意のものを指定することも可能である。

Registry がそのサーバの利用を中止したい際には、leave メッセージを当該サーバに送るか、D-Server 内のタイムアウト時間の経過を待つことにより、サーバの登録を削除する必要がある。当該サーバは、どちらのケースでも、result メッセージを Registry に送信し、当該 Registry に対するサービス中止の連絡を実施する。

また、本手続きは D-Client も実施することができ、事前にサーバがサポートする情報カテゴリとフォーマットを知ることが可能である。

6.3　情報の検索

本手続きでは、D-Client が必要な情報を保持する InfoSource の URI を取得し、情報を受信する手続きを定義する。InfoSource は D-Server に query メッセージを送信し、D-Server は登録されているすべての Registry にそのメッセージを転送する。それぞれの Registry はそのメッセージを受けて、自身が保持する InfoSource に関するメタデータを検索し、候補となる InfoSource を順位付けし、そのランキング情報を result メッセージに乗せて D-Server に返信する。すべての Registry から result メッセージを受信した D-Server は、それらをひとつに集約し、ひとつの result メッセージに載せ、D-Client へと送信する。D-Client は受信した result メッセージに記載されている InfoSource 候補の中からひとつ選択し、その

InfoSource の URI にアクセスし、所望する XML 形式の情報を取得する。

7　プロトタイプ実装

本節では、提案方式のプロトタイプ実装について紹介する。D-Client、D-Server、Registry、そして InfoSource のすべてについて、Java を用いて実装され、CentOS 上にて動作している。また、SPARQL エンジンの実装のひとつである Sesame[23] を活用することにより、RDF データの検索操作を実施している。下記に、本実装の概要を紹介する。

図 3 は、D-Server を用いて D-Client が検索する際のインターフェースを Web に準備したものである。本来、RESTful な実装のため、この Web インターフェースがなくても動作するが、使い勝手向上のため、本インターフェースも用意した。

本画面は、4 種類の情報収集手段を提供している。第 1 に、自由テキスト検索機能を提供する。本機能により、ユーザは自由な文字列を入力して検索が可能である。第 2 に、タグ指定検索機能を提供する。自由テキスト検索同様、自由な文字列を入力して検索可能であるが、検索すべき対象となる XML のタグを絞り込むことが可能である。また、複数の検索条件を組み合わせることも可能となっている。第 3 に、カテゴリ検索機能を提供する。ユーザは、取得したい情報のカテゴリ、そしてフォーマットを指定することにより、ほしい情報の一覧を得ることができる。最後に、最新情報検索機能を提供する。これは、最近登録された情報のみを一覧表示するものである。なお、上述のとおり、本システムは RESTful に実装されているため、上述の Web インターフェースにとらわれない利活用ができる点にも留意されたい。

8　考察

本節では、提案方式の拡張性と実装面を考察すると同時に、本技術を社会で活用するための国際標準化活動について紹介する。

8.1　提案方式の拡張性と実装面の考察

提案方式は将来的に登場する各種情報スキーマに対応すべく、柔軟な情報構造を実現している。既存規格の情報構造が不便になり、利用されなくなってきた際には、新たな情報スキーマを定義・規格化し、それを本方式のデータ構造内にて定義されているいずれかの情報カテゴリに紐付けるのみでよい。また、情報カテゴリは近い将来には変更が不要となるようにオントロジで定義されたカテゴリを活用しているものの、万が一変更が必要な際には、上述のとおり別のカテゴリを活用することも可能となっている。以上より、提案方式は拡張性を保持していると考えられる。

次に、提案方式を実際のインターネットに展開する際の考慮点について考察する。本稿に書かれている内容は、サイバーセキュリティ情報を交換することを主目的にデザインされているのみであり、その技術の安全利用については触れていない。しかしながら、サイバーセキュリティ情報という、悪用されると大きな被

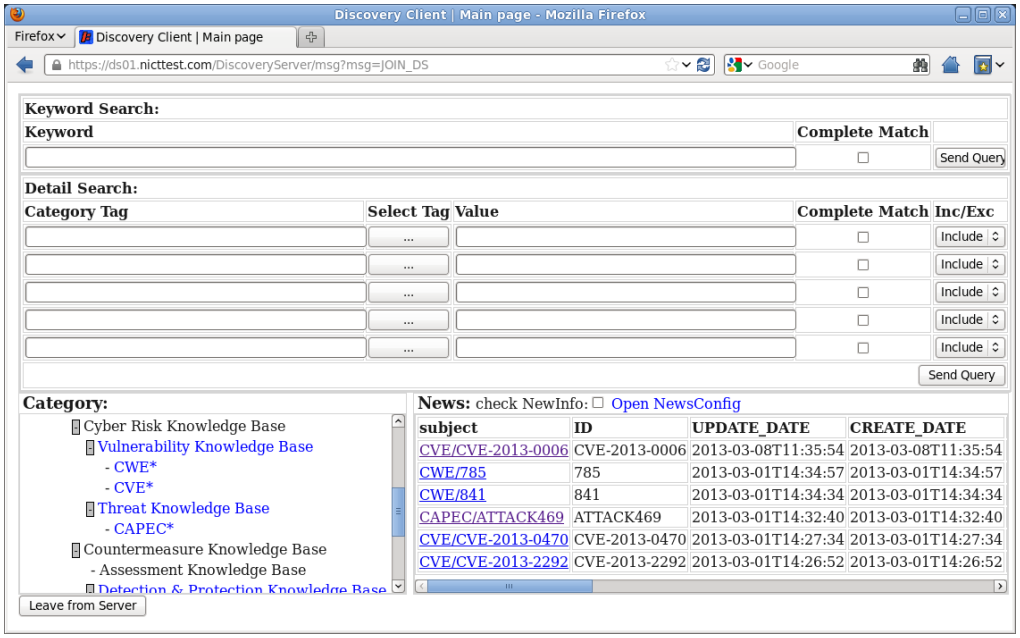

図 3　D-Server を用いた検索画面

害を生じかねない情報を扱っているため、本技術の実装・活用時には、十分なセキュリティ面での考慮が必要不可欠である。例えば、情報登録者や利用者を認証したり、通信を暗号化するなど、各種既存技術を活用することで対策を実施することが可能である。

また、プロトタイプはその処理速度にいまだ課題がある。4つのロールのうち、その処理のボトルネックとなるのはRegistryであり、必要に応じてRegistryは負荷分散を実現できる形で実装されることが望ましい。

8.2　標準化活動

提案方式は、複数の組織の間で情報交換を実施することが目的である。そのためには、最低限の共通インターフェースが定義される必要がある。そのひとつが従来研究でも紹介した、各種情報スキーマである。例えば、情報スキーマが国際標準として定義されることにより、組織間の情報交換は正規化され、効率化されることが期待できる。著者も国際標準化活動には積極的に従事してきており、ITU-T 及びIETFにおいて、各種技術の規格化に貢献してきた。特に、情報交換のフレームワークを定めたITU-T Recommendation X.1500[24]、インシデント情報のスキーマ拡張を定めたIODEF-SCI[25]、そしてディスカバリ技術のフレームワークを定めたITU-T Recommendation X.1570[26]については、提案方式と深く絡むものであり、参考にしていただきたい。

9　結論

提案方式は、ネットワーク上に存在する各種サイバーセキュリティ情報を構造化し、それらの情報の特定・検索・交換を実現した。提案方式はその情報構造に特徴があり、カテゴリとスキーマを分けて定義することにより、柔軟性と将来拡張性を実現している。また、プロトタイプ実装により、提案方式が現実に動作することを示した。本方式は、組織・国を超えたサイバーセキュリティ情報の交換を促進するとともに、グローバルサイバーセキュリティを後押ししていくものであると考えており、今後、実運用に耐え得るシステムの構築をしていきたい。その際には、システムの悪用などのセキュリティ面の対策を実施していく必要があり、これらは今後の課題として対処していく所存である。

謝辞

本研究を実施するにあたり、様々なご支援を頂いた中尾康二主管研究員、平和昌研究所長及びパンタ・ボーラ技術員に深く感謝する。

【参考文献】

1 National Institute of Standards and Technology, "National Vulnerability Database Version 2.2," 2014. [Online]. Available: http://nvd.nist.gov/.
2 JPCERT/CC and IPA, "Japan Vulnerability Notes," 2014. [Online]. Available: http://jvn.jp.
3 T. Takahashi, Y. Kadobayashi, "Reference Ontology for Cybersecurity Operational Information," The Computer Journal, 2015.
4 T. Takahashi, Y. Kadobayashi, "Mechanism for Linking and Discovering Structured Cybersecurity Information over Networks," IEEE International Conference on Semantic Computing, 2014.
5 International Telecommunications Union, "Common vulnerabilities and exposures," ITU-T Recommendation X.1520, 2014.
6 National Institute of Standards and Technology, "Specification for the Asset Reporting Format 1.1," NIST Interagency Report 7694, 2011.
7 International Telecommunications Union, "Common attack pattern enumeration and classification," ITU-T Recommendation X.1544, 2013.
8 National Institute of Standards and Technology, "Common Configuration Enumeration (CCE)," [Online]. Available: http://nvd.nist.gov/cce/index.cfm. [Last Access: 2014].
9 The MITRE Corporation, "Common Event Expression," [Online]. Available: http://cee.mitre.org/. [Last Access: Jan. 2014].
10 International Telecommunications Union, "Common platform enumeration," ITU-T Recommendation X.1528, 2012.
11 The MITRE Corporation, "Common Result Format Specification Version 0.3," [Online]. Available: http://crf.mitre.org/. [Last Access: Jan. 2014].
12 Industry Consortium For Advancement of Security on the Internet, "The Common Vulnerability Reporting Framework v1.1," [Online]. Available: http://www.icasi.org/cvrf-1.1. [Last Access: Jan. 2014].
13 International Telecommunications Union, "Common vulnerability scoring system," ITU-T Recommendation X.1521, 2011.
14 International Telecommunications Union, "Common weakness enumeration," ITU-T Recommendation X.1524, 2012.
15 International Telecommunications Union, "Common Weakness Scoring System," ITU-T Recommendation X.1525, 2015.
16 The Internet Engineering Task Force, "The Incident Object Description Exchange Format," RFC 5070, dec 2007.
17 International Telecommunications Union, "Malware attribute enumeration and characterization," ITU-T Recommendation X.1546, 2014.
18 National Institute of Standards and Technology, "Specification for the Open Checklist Interactive Language (OCIL) Version 2.0," NIST Interagency Report 7692, 2011.
19 International Telecommunications Union, "Language for the open definition of vulnerabilities and for the assessment of a system state," ITU-T Recommendation X.1526, 2014.
20 International Organization for Standardization/International Electrotechnical Commission, "Specification for the Extensible Configuration Checklist Description Format (XCCDF) Version 1.2," ISO/IEC 18180:2013, 2013.
21 The World Wide Web Consortium, "Resource Description Framework (RDF): Concepts and Abstract Syntax," http://www.w3.org/TR/2004/REC-rdf-concepts -20040210/, 2004.
22 The World Wide Web Consortium, "SPARQL query language for RDF," http://www.w3.org/TR/2013/REC-sparql11-overview-20130321/ .
23 openRDF.org, "SESAME," [Online]. Available: http://www.openrdf.org/. [Last Access: March 2012].
24 International Telecommunications Union, "Overview of Cybersecurity information exchange (CYBEX)," ITU-T Recommendation X.1500, 2011.
25 The Internet Engineering Task Force, "An Incident Object Description Exchange Format (IODEF) Extension for Structured Cybersecurity Information," RFC 7203, April 2014.
26 International Telecommunications Union, "Discovery mechanisms in the exchange of cybersecurity information," ITU-T Recommendation X.1570, 2011.

高橋健志　（たかはし　たけし）
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
主任研究員
博士（国際情報通信学）
サイバーセキュリティ、通信プロトコル