セキュリティ通知の連絡先情報の収集に関する検討

斉藤 美織[1,*]　田辺 瑠偉[2]　藤田 彬[2]　吉岡 克成[3]　松本 勉[3]

概要：インターネットに接続した機器の増加に伴い，サイバー攻撃による被害が拡大している．このため，ハニーポット等の観測網を用いて攻撃を観測した際には，攻撃元のシステムの管理者へ通知を行うことが重要である．一般に，通知先の特定には WHOIS などが用いられる．しかし，一部のサービスではプライバシーの観点から連絡先が秘匿されている場合や連絡先情報の更新が滞っている場合があり，正しい情報が取得できない．また，ISP やクラウドサービスプロバイダ等の上位組織へ通知をした際には，実際に攻撃の影響を受ける末端のシステム管理者まで通知が届かない恐れがある．本研究では，ハニーポットで観測した攻撃元の 46,105 IP アドレスに関して，様々な方法でメールアドレスや SNS アカウント等の連絡先情報の収集を試みた．具体的には，通知対象の IP アドレスと関連の深い 2,216 個のドメインをパッシブ DNS から取得し，これらのドメインを用いて Web アクセスを行い，メールアドレス，SNS アカウント，Web 上の連絡フォームを収集すると共に，既存のメールアドレスデータベースを用いて関連メールアドレスを収集した．その結果，対象 IP アドレスの WHOIS から得られる通知先 66,451 件に加えて 4,665 件の通知連絡先を取得することができた．このうち，3,435 件については通知対象との関連性が高い等の理由から通知先として有効と判断される．

キーワード：セキュリティ通知，WHOIS, SNS

On Finding Contact Points for Security Notification

Miori Saito[1,*]　Rui Tanabe[2]　Akira Fujita[2]
Katsunari Yoshioka[3]　Tsutomu Matsumoto[3]

Abstract: With the increase of devices connected to the Internet, the threats caused by cyberattacks are increasing. Besides observation and analysis of these attacks, notification to the affected users is becoming increasingly important. The most common way to identify the contact point for notification is using WHOIS. However, direct contact points may be hidden from the viewpoint of privacy or simply outdated. Moreover, notifications to higher-level organizations such as ISPs and cloud service providers may be ignored or may not be treated properly. In this research, we investigate how to enrich contact information for given IP addresses to be notified. In the experiment, from passive DNS we obtained 2,216 domains corresponding to 46,105 to-be-notified IP addresses observed by honeypot, then crawled the corresponding web pages to obtain possible contact points, namely, e-mail addresses, SNS accounts, and web contact forms. Moreover we used existing email address databases to obtain email addresses related to these domains. As a result, we were able to obtain additional 4,665 contact points besides IP-based WHOIS contacts, among which 3,435 are evaluated to be highly relevant to the IP address and thus effective in security notification.

Keywords: security notification, WHOIS, SNS

1. はじめに

近年増加の一途を辿るサイバー攻撃[1]の分析を行うため，脆弱な機器を模擬したハニーポット等，多くの観測システムが研究開発及び運用されている．また，機器の脆弱性を早期に発見するため，ネットワークスキャン等の能動的観測が行われている[2,3]．これら観測システムにおいて脆弱性および攻撃が観測された際には，脆弱性の悪用や被害の拡大を防ぐため早急にシステム管理者にセキュリティ通知を行う必要がある．

一般的に通知先の特定には WHOIS サービスにより提供される情報が用いられるが，これらの情報においては，プライバシーの観点から連絡先が秘匿されていたり，情報公開代理サービスの連絡先が公開されている場合がある．情報公開代理サービスとは，WHOIS において氏名や住所，メールアドレス等の個人情報をドメイン登録者の代わりに公開するサービスであり，一般にドメイン登録会社が提供している．またこのほかにも，WHOIS データベースの更新が滞っている場合があり，通知時に正しい連絡先情報が取得できない可能性がある．また，情報公開代理サービスや ISP，クラウドサービスプロバイダといった上位組織を経由して通知を行う際には，上位組織が通知内容を誤認識，もしくは見落としてしまう恐れがある．このため，実際に攻撃の影響を受ける末端のシステム管理者まで通知が届かない恐れがある．また，前述の通りセキュリティ通知は早急に行

1 横浜国立大学大学院 環境情報学府
Graduate School of Environment and Information Sciences, Yokohama National University
2 横浜国立大学先端科学高等研究院
Institute of Advanced Sciences, Yokohama National University
3 横浜国立大学大学院環境情報研究院/先端科学高等研究院
Graduate School of Environment and Information Sciences, Yokohama National University / Institute of Advanced Sciences, Yokohama National University
* saito-miori-ck@ynu.jp

うことが求められるが，複数組織を経由して通知を行う場合には末端のシステム管理者に通知が届くまでに時間がかかり，被害の拡大を防止できない恐れもある．このため，観測した事象に関わる複数の対象に通知を行い，通知の効果を向上することが重要である．

本研究では，脆弱な IoT 機器を模したハニーポットにより観測され，マルウェアに感染している恐れがある 46,105 IP アドレスに対して，1) 当該 IP アドレスを基に WHOIS から連絡先を特定する方法 (以降では IP-WHOIS)，2) IP アドレスを用いた Web アクセスによる連絡先情報収集(以降では IP-Web)，3) IP アドレスと関連の深いドメインに基に WHOIS から連絡先を特定する方法(以降では Domain-WHOIS)，4)IP アドレスと関連の深いドメインを基に既存のメールアドレスデータベースから連絡先を特定する方法(以降では Domain-MDB)，5)IP アドレスに関連の深いドメインを用いた Web アクセスによる連絡先情報収集(以降では Domain-Web)により，通知連絡先としてメールアドレス，Web 連絡フォーム， SNS へのリンク(以降，SNS リンク)を合計 71,166 件収集した．さらに，収集した連絡先について，通知対象ホストと関連のある連絡先が取得できているか，当該連絡先が利用可能であるかなどの観点から評価を行った．その結果，対象 IP アドレスの WHOIS から得られる通知先 66,451 件に加えて 4,665 件の通知連絡先を取得することができた．このうち，3,435 件については通知対象との関連性が高く通知先として有効と判断される．

2. 関連研究

近年，ハニーポットやネットワークスキャン等で観測された脆弱性や攻撃について，脆弱性を有するシステムや攻撃元のシステムの管理者等へ通知を行う重要性が高まっている[4]．そのため，どのような通知方法が有効であるかを検証する研究が多数行われてきた[5,6,7]．

文献[5]では通知先の選択や通知内容などの観点から有効性の評価を行なっており，システムの脆弱性の通知に関して WHOIS を用いる通知が，CERT への通知よりも対応率が高いことを明らかにしている．さらに簡潔な通知文よりも詳細に記述した通知文の方が修正率が向上することを示している．

文献[6]では多数の脆弱な Web サーバへの大規模通知を目的とし，セキュリティ通知の有効性と実現可能性について調査している．CERT や ISP を経由して行う間接的な通知と，Web サーバの管理者へ行う直接的な通知による脆弱性の修正率について比較しており，両者とも通知の有効性が確認できるものの直接的な通知の修正率が高いことが示されている．なお，直接的な通知では Domain-WHOIS により取得したメールアドレスと，RFC2142[8]に準拠して生成した通知対象メールアドレスを用いている．

文献[7]では脆弱な権威 DNS サーバを対象として，通知先組織と脆弱性の修正率の関連についてまとめている．サーバ管理者の連絡先は，SOA レコード RNAME フィールドに記述されたメールアドレスとし，ドメイン管理者の連絡先は Domain-WHOIS，ネットワーク管理者の連絡先は IP-WHOIS を用いて取得している．この結果，DNS サーバ管理者へ通知を行うことにより，最も高い修正率を得られることが示されている．

文献[5,6,7]では様々な通知経路を用いた通知実験が行われ，経路によって修正率の差があるものの，いずれの経路でも修正率が依然として低いことが指摘されている．メールのバウンス率も高く，通知の到達率が低いため他の通知連絡先を収集する必要がある．

そこで本研究では，一般に用いられる WHOIS の他に，クローリング，メールアドレスデータベースといったメールアドレス収集手法を用いる．さらに，一般に通知の際にはメールが用いられているが，メール以外の通知経路(Web 連絡フォーム，SNS)についても検討する．

3. 調査手法

3.1 セキュリティ通知の流れ

セキュリティ通知を行う際，一般的には CERT や ISP などの組織へ連絡をするか，WHOIS, GeoIP[9]などのサービスで取得した連絡先を用いる．本研究では受動的に攻撃を観測するハニーポット等により，通知対象が IP アドレスとして与えられる状況を想定し，1) 当該 IP アドレスを基に WHOIS から連絡先を特定する方法 (IP-WHOIS), 2) IP アドレスを用いた Web アクセスによる連絡先情報収集(IP-Web), 3) IP アドレスと関連の深いドメインを基に WHOIS から連絡先を特定する方法(Domain-WHOIS), 4)IP アドレスと関連の深いドメインを基に既存のメールアドレスデータベースから連絡先を特定する方法(Domain-MDB), 5)IP アドレスに関連の深いドメインを用いた Web アクセスによる連絡先情報収集(Domain-Web)の 5 つの方法により，通知先のメールアドレスおよび，Web 連絡フォーム，SNS リンクを収集する(図１)．SNS リンクは，あるアカウントに対し第三者が閲覧不可能なメッセージ(以降，ダイレクトメッセージ)を送る機能を持つ Facebook[10], Twitter[11]の 2 つの SNS へのリンクを収集の対象とする．通知を行う際には収集したリンクが示す SNS アカウントへダイレクトメッセージを送信する．

4 章の実験では，ハニーポットで観測した IP アドレスのうち，マルウェアに感染している恐れがあるものを通知対象とした．また当該 IP アドレスに関連の深いドメインを得るために Passive DNS の 1 つである DNSDB[12]を用いた．なお，本研究では IP アドレスを通知先の基本情報としたが，脆弱な Web サイト等通知先のドメインが事前に分かっ

ている場合には当該ドメインを基本情報として，DNS 名前解決により当該ドメインに対応する IP アドレスを取得し，同様に通知先を収集することも可能である．

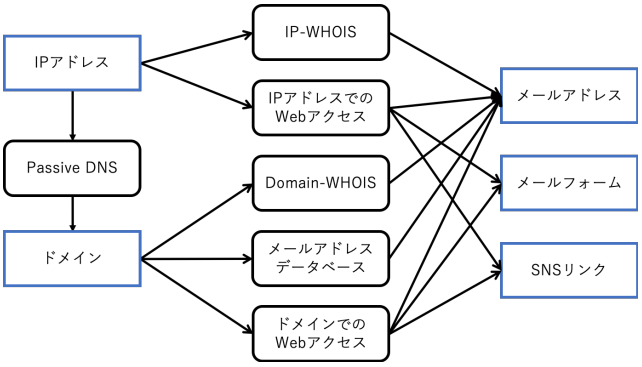

図 1．通知連絡先取得の流れ

3.2 通知連絡先情報の取得の流れ

連絡先の取得方法として，5 種を用いる．以下では， IP アドレスを基にした手法(2 種)とドメインを基にした手法(3 種)に大別して説明する．4 章の実験では IP アドレスを基とした連絡先取得方法は常に実行した．ドメインを基にした連絡先取得方法は，DNSDB を用いて対象 IP アドレスに関連深いドメインが取得できた場合にのみ実行した．

3.2.1 IP アドレスを基とした連絡先取得方法

- **IP アドレスを用いた WHOIS 問い合わせ(IP-WHOIS)**

通知対象となる IP アドレスを入力として WHOIS データベースを検索する．検索結果からメールアドレスを示す正規表現にマッチした文字列を取得する．一般に whois コマンドを用いて検索した際には地域インターネットレジストリの whois データベースに問い合わせをするが，より詳細な情報を取得するために，国別インターネットレジストリの whois 情報も参照する．

- **IP アドレスでの Web アクセス(IP-Web)**

通知対象となる IP アドレスを用いて生成した URL(例: http://192.0.2.0/)にアクセスし HTML コンテンツを取得する．アクセスした Web ページ(以降，Top ページ)内に”問い合わせ”，”contact”等の文字列を含むボタンがあればクリックし，遷移先のページ(以降, Contact ページ)の HTML コンテンツを取得する．さらに以下の手順に従い，メールアドレス，Web 連絡フォーム，SNS リンクを収集する．

メールアドレスの取得：半角@の代わりに全角@，[at]などを含むメールアドレスにマッチする正規表現を用いて Top ページ，Contact ページの HTML コンテンツからメールアドレスを抽出する．

Web 連絡フォームの取得：Contact ページの HTML コンテンツから Form タグを抽出する．サイト内の検索ボックスを除外するため，Form タグの属性に文字列”search”が含まれていない場合に当該サイトを Web 連絡フォームとして抽出する．

SNS リンクの取得：Top ページ，Contact ページの HTML コンテンツから a タグを抽出し，href 属性の属性値に”facebook.com”, “twitter.com”のいずれかの文字列を含む場合，SNS リンクとして抽出する．

3.2.2 ドメインを基とした連絡先取得方法

- **ドメインでの WHOIS 問い合わせ(Domain-WHOIS)**

通知対象となるドメインを入力として WHOIS データベースを検索する．検索結果からメールアドレスを示す正規表現にマッチした文字列を取得する．

- **メールアドレスデータベースの利用(Domain-MDB)**

メールアドレスデータベース[13]はドメインをキーとして当該ドメインのメールアドレスを検索できるサービスである．インターネット上に公開されているメールアドレスがクローリングにより収集され，データベースに格納されている．データベース内のメールアドレスは個人のメールアドレスである Personal と組織内の部署や役割毎のメールアドレスである Generic に分けられている．本研究では通知対象となるドメインを検索し，Generic に分類されているメールアドレスを抽出する．

- **ドメインでの Web アクセス(Domain-Web)**

通知対象となるドメインを用いて http://example.com/, http://www.example.com/等の URL を生成し，生成した URL にアクセスする．以降 3.2.1 節に示した IP-Web と同様に，Top ページと Contact ページの HTML コンテンツを取得し，メールアドレス，Web 連絡フォーム，SNS リンクを収集する．

3.3 取得連絡先の有効性評価方法

収集したメールアドレス，Web 連絡フォーム，SNS リンクに対し，通知連絡先として利用可能であるか評価を行う．具体的には 1)収集した通知連絡先が存在するか(連絡先の有効性)，2)収集した通知連絡先が対象 IP アドレスと関連性が高いものであるか(通知対象との関連性)，3)収集した通知連絡先が不特定多数のユーザから通知や問い合わせを受け取ることを目的として設定された連絡先であるか(連絡先の目的との合致性)という観点から評価した．

3.3.1 メールアドレスの評価方法

メールアドレスについては以下 4 点の評価を行う．

(1) **メールサーバの存在**

取得したメールアドレスが実在するメールアドレスであるかを評価する．メールを送信せずにメールアドレスが実在するかを評価することは困難であるため，本研究ではメールサーバが存在するかを評価する．各メールアドレスについてドメインを抽出し，MX レコードを問い合わせることでメールサーバの FQDN を取得する．さらに，FQDN を名前解決することで，メールサーバの IP アドレスを取得する．メールサーバ

の IP アドレスが取得できた場合，当該サーバが存在したとして対象メールアドレスを有効なメールアドレスとみなす．

(2) **ドメインの一致**

入力としたドメインが組織固有のものである場合，同じドメインのメールアドレスが当該組織で利用されていると考えられる．取得したメールアドレスがドメインを管理する組織と紐づくものであるかを評価するため，メールアドレスのドメインと入力ドメインを比較する．両ドメインが一致した場合，入力ドメイン保有組織のメールアドレスが取得できたとして，有効な通知連絡先とみなす．なお，メールアドレスデータベースはドメインを入力とし，入力ドメインのメールアドレスを出力するためドメインの一致率は 100%となる．

(3-a) **RFC2142 への準拠**

RFC2142[8]では問い合わせや報告を受け取るために設定することが望ましいとされているメールアドレスが定義されている．取得したメールアドレスが RFC2142 に準拠したものになっている場合，有効な通知連絡先であるとする．

(3-b) **通知先メールアドレスに用いられる単語の有無**

RFC2142 で定められた abuse, security 等の他にも，contact, cert といったセキュリティ通知に用いられるメールアドレスに含まれている単語がある．これらの単語がメールアドレスの@以前(以降，メールアカウント)に含まれている場合，有効な通知連絡先であるとする．また，”contact”については 8 つの言語で翻訳し，評価の対象とした．

3.3.2 Web 連絡フォームの評価

(1) **Web 連絡フォームの存在**

収集した Web ページに名前やメールアドレス，問い合わせ内容を記述するフォームが含まれていた場合，通知先が存在すると判断する．

3.3.3 SNS リンクの評価

SNS リンクについては以下の流れで評価を行う．

(1-a) **SNS アカウントの特定**

リンク先の URL にアクセスし，SNS アカウントを一意に特定する．

(1-b) **メッセージの送信**

SNS によっては，非公開のメッセージを受け取らないように設定することができる．(1)で特定したアカウントがメッセージを受信する設定になっている場合，通知先が存在すると判断する．

(2) **管理者の比較**

(1-b)でメッセージを受信する設定となっている SNS アカウントのうち，アカウント管理者が SNS リンクの取得元 Web ページの管理者と同一であると判断できる場合，有効な通知連絡先とする．

4. 実験

4.1 データセット

近年 IoT 機器が普及し，企業への導入も進んでいる[14,15]．そこで本研究では，機器の普及に伴い増加する IoT 機器を狙ったサイバー攻撃に着目する．脆弱な IoT 機器を模したハニーポット[16]において telnet ログインの試行が観測された(ユーザ名を入力し，パスワードの入力を求められた)攻撃元ホストの IP アドレスを通知対象とした．2019/4/24-2019/7/7 の 75 日間観測を行った結果，通知対象として 46,105 IP アドレスを収集した．

対象 IP アドレスと関連の深いドメインの取得には DNSDB[12]を使用した．ハニーポットへの通信観測日の前後 4 週間に対象 IP アドレスへの名前解決が確認された FQDN を取得し，ドメインを抽出する．対象 IP アドレスと関連性が高いドメインを取得するため，名前解決回数が 10 回以下の FQDN は実験対象から除外した．また，末端のシステム管理者に近い通知連絡先を得るため，ISP が管理のために利用していると思われる”IP アドレスを含む FQDN”(例: 192.0.2.0 への名前解決が確認された 192-0-2-0.example.com)も除外した．この結果，747IP に対し 2,216 個の FQDN を取得した．

本研究では通知対象としてドメインではなく FQDN を取得した．取得した FQDN がメールサーバや DNS サーバのものであった場合，MX レコード，NS レコードを検索することで当該サーバが管理するドメインを取得することができる．本研究では当該ドメインも Web アクセスの対象に加えた．

表 1．通知連絡先取得件数

	IP-WHOIS	IP-Web	Domain-WHOIS	Domain-MDB	Domain-Web
メールアドレス	66,451(99.99%)	536(97.01%)	998(98.80%)	2,239(97.14%)	198(65.66%)
メールフォーム		71(100.00%)			124(100.00%)
SNSリンク		279(89.25%)			387(92.25%)

4.2 実験結果

実験の結果，IP アドレスと連絡先の組を 71,166 件収集した．うち 70,336 件がメールアドレス，195 件が Web 連絡フォーム，635 件が SNS リンクであった．表 1 に連絡先取得経路ごとに取得した IP アドレスと通知連絡先の組数を示す．括弧内の数字は当該連絡先取得経路のみで取得できた通知連絡先の割合を示しており，多くが 90%以上と高い割合であった．このことから，連絡先取得経路ごとに固有の連絡先を取得することができ，取得経路を増やすことで，より多くの通知連絡先を取得できることがわかる．

また，取得した通知連絡先について評価を行った．評価の結果を表 2 に示す．有効であると判断された通知連絡先はメールアドレス 47,159 件，Web 連絡フォーム 117 件，SNS リンク 246 件であった．メールアドレスについては複数評価基準があるが，メールサーバが存在し，さらに 3.3.1 節で示した観点 2,3 の 3 種の評価基準のうちどれか 1 基準でも満たしているものについて有効なメールアドレスとした．

5. 考察

5.1 各方法で収集した通知連絡先の特徴

表 1 より当該連絡先取得経路のみで取得できた通知連絡先の割合が高いことから，本研究で取得した通知連絡先は連絡先の取得経路ごとにほぼ固有であると言える．しかし，Domain-Web によって取得されたメールアドレスについては，65.66%と他の取得方法と比べて低い割合となった．これは Domain-MDB で用いたメールアドレスデータベースも同様にクローリングによって収集されたメールアドレスが格納されているためであると考えられる．

また，連絡先取得経路によって，取得できる連絡先に特徴が見られた．例えば，IP-WHOIS を用いた場合には ISP やクラウドサービスプロバイダ等，上位組織のメールアドレスが多く収集されていた．同様に，Domain-WHOIS を用いて取得したメールアドレスには，whoisguard.com のような情報公開代理サービスのドメインのメールアドレスが多く含まれていた．情報公開代理サービスを利用している場合，異なるドメインの WHOIS 問い合わせにより，同一ドメインのメールアドレスが取得できることが予想される．そこで，Domain-WHOIS により取得したメールアドレス 998 件のうち，同一ドメインのメールアドレスが複数取得できている 818 件について情報公開代理サービスであるか調査を行なった．その結果，599 件(73.23%)がドメイン登録会社等の情報公開代理サービスのドメインを持つメールアドレスであった．

このように WHOIS を用いた場合には，末端のシステム管理者への直接的な通知経路を取得することは困難であると言える．一方，Domain-MDB, Domain-Web については特定のドメインが多く占めるケースは確認されず，より末端のシステム管理者に近い連絡先が取得できたものと考えられる．

また，IP-Web では，ルータの Web 管理画面と思われる Web ページから機器の製造元と思われるメールアドレスが取得された．機器製造元のメールアドレスは末端のシステム管理者への通知という観点では有効なメールアドレスであるとは言えない．しかしある特定の企業の製造したルータからの通信を多く観測しているということは，当該ルータが脆弱である可能性がある．このため，製造元企業への通知という観点では有効となり得る．

5.2 各方法で収集した通知連絡先の有効性

各方法で収集した通知連絡先について，実在する連絡先はメールアドレスが 68,631 件，Web 連絡フォームが 117 件，SNS リンクが 272 件であった．

表 2．通知連絡先の評価結果

(2-a) メールアドレスの評価

	IP-WHOIS	IP-Web	Domain-WHOIS	Domain-MDB	Domain-Web
(IP, 連絡先)組数	66,451	536	998	2,239	198
メールサーバ存在	65,028(97.86%)	366(68.28%)	990(99.20%)	2,150(96.03%)	175(88.38%)
RFC2142準拠	27,625(41.57%)	381(71.08%)	403(40.38%)	368(16.44%)	64(32.32%)
メールアカウント	44,713(67.29%)	397(74.07%)	618(61.92%)	567(25.32%)	84(42.42%)
ドメイン一致	6(0.01%)	2(0.37%)	24(2.40%)	2,239(100.00%)	118(59.60%)

(2-b) メールフォームの評価

	IP-Web	Domain-Web
(IP, 連絡先)組数	71	124
メールフォーム存在	52(73.24%)	65(52.42%)

(2-c) SNSリンクの評価

	IP-Web	Domain-Web
(IP, 連絡先)組数	279	387
ユーザの特定	188(67.38%)	323(83.46%)
メッセージ送信可能	119(42.65%)	161(41.60%)
管理者の一致	100(35.84%)	154(39.79%)

メールアドレスについてはIP-Web, Domain-Web以外の方法では96%以上と高い割合でメールサーバが存在していた．WHOIS 情報の登録の際には有効なメールアドレスの登録が必須となっている．また，メールアドレスデータベースについても，データ提供の際にメールアドレスが有効であるかの評価を行っている旨が記されていたため，多くの有効なメールアドレスが収集されたものと考えられる．一方IP-WebおよびDomain-Webで収集されたメールアドレスは，メールサーバが存在しないものがそれぞれ31.72%, 11.62%と他の収集方法と比べ多く収集されていた．メールサーバが存在していなかったメールアドレスを確認してみるとwebmaster@example.com が多く収集されていた．さらに当該メールアドレスが収集された Web ページを確認すると，Web サーバのテストページであることがわかった．当該テストページにはデフォルト設定でメールアドレスwebmaster@example.com が表示されるようになっていた．このようなデフォルト設定のテストページに多くアクセスしたため，メールサーバが存在しないメールアドレスを多く収集したと考えられる．

Web 連絡フォームの収集の際，タグの属性に”search”が含まれないものを対象とすることで検索窓を除外した．しかし，収集した Web サイトの中には Web 連絡フォーム以外の Web サイトが 40%含まれていた. Web 連絡フォームとして誤検知した Web サイトは検索窓を含むものや，ログインのために ID/Password を入力するフォームであった．

Web 連絡フォームを用いたセキュリティ通知を行う際には，自動でメールを送信することが困難であるため，手動で通知を行う．このため，Web 連絡フォームではない Web サイトが収集されていても判断をすることができる．しかし，大規模な通知を行う際には誤検知を減らし，手動の作業を減らすことで，より効率的にセキュリティ通知を行うことができる．

SNS リンクについて詳細な評価結果を表 3 に示す．いずれの SNS に対しても，ユーザを特定することができないリンクが含まれていた．これには大きく分けて 2 つの原因が考えられる．まず 1 点目としては，リンク先の URL が誤っている場合がある. 各 SNS の Web ページではあるものの，”ユーザが存在しない”, ”リンクの期限切れ”等が表示される URL が多く存在した．次に 2 点目としては SNS の Top ページやヘルプページなど，SNS へのリンクではあるが，特定のユーザを示さないページへのリンクである場合が挙げられる．このケースで多く見られたのは，リンクの取得元 Web サイトを SNS 上で共有するためのリンクであった．

また，メッセージの受信設定について，SNS によって大きな差が生じた．Facebook ではユーザを特定できた 228 件のうち，メッセージが送信可能であったアカウントは 196 件(85.96%)であったのに対し，Twitter では 269 件中 76 件(28.25%)という結果となった．この結果には SNS の匿名性が大きく関わっていると考える．Twitter では名前の登録の必要がないのに対し，Facebook では基本的に本名の登録が必要であるため，各アカウントの信頼性が高いと言える．このため，メッセージの送信が可能なアカウントが多いのではないかと考えた．また，アカウントの信頼性はセキュリティ通知を行う際にも重要である．通知の送信元が信頼できる組織や個人でなければ，通知がスパムメールとして扱われてしまう可能性もある．このため，セキュリティ通知を行う際には，信頼できる組織，個人から通知を送信する必要がある．

表 3．SNS リンクの評価結果

(3-a) SNSリンクの評価(Facebook)

	IP-Web	Domain-Web
(IP, 連絡先)組数	141	185
ユーザの特定	90(63.83%)	146(78.92%)
メッセージ送信可能	82(58.16%)	123(66.49%)
管理者の一致	75(53.19%)	121(65.41%)

(3-b) SNSリンクの評価(Twitter)

	IP-Web	Domain-Web
(IP, 連絡先)組数	138	202
ユーザの特定	98(71.01%)	177(87.62%)
メッセージ送信可能	39(28.26%)	38(18.81%)
管理者の一致	25(18.12%)	33(16.34%)

5.3 各方法で収集した通知連絡先の通知対象との関連性

各方法で収集した通知連絡先について, 実在する連絡先のうち通知対象 IP アドレスと関連性が高いと判断されたものはメールアドレスが 2,206 件，SNS リンクが 246 件であった. なお, Web 連絡フォームについてはアクセスした Web サイトの contact ページであるため, 全て関連性が高いとする．

Domain-WHOIS と Domain-Web においてはメールアドレスのドメインと入力ドメインが一致する割合は非常に低い結果となった．Domain-WHOIS については，情報公開代理サービスの連絡先が公開されていることが原因として考えられる．また，Gmail 等 Web メールのメールアドレスも多く確認された．また, Domain-Web においては入力としたドメインが<組織名>.net であるのに対し，取得したメールアドレスが<組織名>.com であるなど，ドメインが部分的に一致している場合も存在した．さらに，ドメインの一致はしていないものの，入力ドメインとメールアドレスのドメインが組織名とサービス名を示しているなど関連のあるケースが確認された．2 つのドメインに全く関連がないと思われるケースは 77 件であった．なお，この 77 件には @example.com 等のメールサーバが存在しないメールアドレスが 22 件，gmail や hotmail 等のフリーメールアドレスが 30 件含まれている．

次に，SNS リンクについて，特定した SNS アカウントが Web サイトの管理者と一致しないものが約 10%存在した．Web サイトの管理者と一致しなかったアカウントの中には企業の公式アカウントから個人のアカウントと思われるものまで様々なアカウントが含まれていた．当該アカウントへのリンクを取得した Web サイトを確認すると，企業の Web サイトと思われるサイトが確認された．当該ページでは関連のある企業や組織のアカウントが多く紹介されていたため，通知対象と関連しないアカウントが多く取得できていた．また，企業の他にも，個人の有名アカウントを多く紹介している Web サイトも存在した．

5.4 各方法で収集した通知連絡先の目的との合致性

各方法で収集した通知連絡先について，実在する連絡先のうち，通知や問い合わせのための連絡先であると判断されたメールアドレスは 45,463 件あった．なお，Web 連絡フォームは全て問い合わせのために用いられるものであるとする．

まず，メールアドレスが通知・問い合わせのためのものであるか，RFC2142 への準拠という観点で評価を行った．RFC2142 に準拠するメールアドレスの割合は連絡先の取得経路ごとに大きな差が生じた．IP-Web によって取得したメールアドレスは約 70%が RFC2142 に準拠しているという結果となった．これは，前述のルータ製造元のメールアドレスや，Web サーバのテストページにおいて取得された webmaster@example.com が RFC2142 に準拠しているためと考えられる．

WHOIS に登録されているメールアドレスは当該 IP アドレスやドメインに対して問い合わせを行うためのものである．このため，同様に問い合わせ用のメールアドレスを定義した RFC2142 に準拠するものが多いと予想したが予想に反し 40%台という結果になった．

次に，abuse, cert 等のセキュリティ通知に用いられるメールアドレスに含まれている単語がメールアカウントに含まれているか調査した．その結果，45,463 件のメールアドレスにこれらの単語が含まれていた．実際にメールアドレスを確認すると，RFC2142 に準拠するメールアドレスの他に abuse-***@domain, support-***@domain などのメールアドレスが確認された．また，”contact”については複数の言語で評価を行ったが，”contacto”, “kontakt”等の文字列により有効であると評価されたメールアドレスも存在した．

このことから RFC2142 に準拠しないメールアドレスも通知先や問い合わせ先として設定されていることがわかった．このため，通知対象としてドメインが取得できた際に RFC2142 に準拠するメールアドレスに通知をするだけでは不十分であると言える．また，メールアカウントは当該メールアドレスを用いるユーザの使用言語と関係していることが確認された．

5.5 広域スキャンシステム

収集したメールアドレスには，internet-census.org[17]や binaryedge.io[18]のような広域スキャンシステムと思われるドメインのメールアドレスが 12 件含まれていた．本研究ではハニーポットにおいてログインの試行が確認されたホストを通知対象としているが，一部のスキャンシステムから送られるリクエストをハニーポットがログイン試行であると誤認したためである．管理者が意図しない通信を行った IP アドレスのみ通知を行う場合には，ハニーポットのログの精査が必要となるがこれは今後の課題とする．

6. まとめと今後の課題

本稿ではハニーポットで観測されたマルウェア感染が疑われるホスト 46,105IP に対し，セキュリティ通知のための連絡先収集方法について調査を行った．通知連絡先の収集の結果，対象 IP アドレスの WHOIS から得られる通知先 66,451 件に加えて 4,665 件の通知連絡先を取得することができた．うち，3,885 件がメールアドレス，195 件が Web 連絡フォーム，635 件が SNS リンクであった．さらに，収集した連絡先について通知連絡先として利用可能であるか評価を行った．

一般的な通知経路として用いられるメールアドレスについてはメールサーバが存在するメールアドレスであるか，通知や問い合わせ等に用いられるメールアドレスであるか等の観点から評価を行った．その結果，通知連絡先として有効であると評価されたメールアドレスは 47,159 件であった．このうち，従来の IP-WHOIS 以外の方法により取得できたメールアドレスは 3,072 件であり，取得したメールアドレスの約 8 割が有効であると評価された．

また，本研究では新たな通知経路として Web 連絡フォームや SNS を提案し，評価を行なった．Web 連絡フォームについては収集した 195 件のうち 117 件が有効であると評価された．SNS リンクについては SNS アカウントへメッセージを送ることが可能であるか，また，リンク収集元の Web サイトの管理者と SNS アカウントの管理者が一致するか，といった観点で評価を行なった結果，246 件の SNS リンクが有効であると評価された．

今後は，収集された連絡先を精査し，実際にセキュリティ通知を行い，その効果を測定することを予定している．

謝辞

本研究の一部は，国立研究開発法人情報通信研究機構（NICT）の委託研究「欧州との連携によるハイパーコネクテッド社会のためのセキュリティ技術の研究開発」により得られた．

参考文献

[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, “NICTER 観測レポート 2018,” http://www.nict.go.jp/cyber/report/NICTER_report_2018.pdf, (参照 2019/08/20).

[2] Shodan, https://www.shodan.io/, (参照 2019/08/20).

[3] Censys, https://censys.io/, (参照 2019/08/20).

[4] Mohammad Hanif Jhaveri, Orcun Cetin, Carlos Ganan, Tryler Moore, and Michel van Eeten, “Abuse Reporting and the Fight Against Cybercrime,” ACM Computing Surveys (CSUR), 2017.

[5] Frank Li, Zakir Durumeric, Jakub Czyz, Mohammad Karami, Michael Bailey, Damon McCoy, Stefan Savage, and Vern Paxson, “Youve got vulnerability: Exploring effective vulnerability notifications,” 25th USENIX Security Symposium, 2016.

[6] Ben Stock, Giancarlo Pellegrino, Christian Rossow, Martin Johns, and Michael Backes, “Hey, You Have a Problem: On the Feasibility of Large-ScaleWeb Vulnerability Notification,” 25th USENIX Security Symposium, 2016.

[7] Orcun Cetin, Carlos Ganan, Maciej Korczynski, and Michel van Eeten, “Make Notifications Great Again: Learning How to Notify in the Age of Large-Scale Vulnerability Scanning,” Workshop on the Economics of Information Security 2017, 2017.

[8] Internet Engineering Task Force, “Rfc 2142,” https://www.ietf.org/rfc/rfc2142.txt, (参照 2019/8/13).

[9] MaxMind, “GeoIP,” https://www.maxmind.com/en/geoip2-city, (参照 2019/08/13).

[10] Facebook, https://www.facebook.com/, (参照 2019/8/20).

[11] Twitter, https://www.twitter.com/, (参照 2019/8/20).

[12] Farsight Security, Inc., “DNSDB,” https://www.dnsdb.info/, (参照 2019/08/15).

[13] 総務省, “平成 30 年版 情報通信白書,” http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h30/html/nd111200.html, (参照 2019/08/20).

[14] 総務省, “日本企業の AI・IoT の導入状況,” http://www.soumu.go.jp/main_content/000610197.pdf, (参照 2019/08/20).

[15] hunter, “domain search,” https://hunter.io/search, (参照 2019/08/13).

[16] Yin Minn Pa Pa, Shogo Suzuki, Katsunari Yoshioka, Tsutomu Matsumoto, Takahiro Kasama, and Christian Rossow, “IoTPOT: Analysing the Rise of IoT Compromises,” USENIX/WOOT’15, 2015.

[17] Internet census group, https://www.internet-census.org/home.html, (参照 2019/08/20).

[18] Binaryedge, https://www.binaryedge.io/, (参照 2019/8/20).