悪性ホストの多拠点からの継続的な観測に基づく時系列および地域性の分析

藤井 翔太[1,2]　佐藤 隆行[1]　青木 翔[1]　津田 侑[3]
川口 信隆[1]　重本 倫宏[1]　寺田 真敏[1]

概要：サイバー攻撃において，悪性ホストは C2 サーバをはじめとして，攻撃者にとって重要なインフラと化している．こうした状況においては，悪性ホストを観測し，その実態を明らかにすることで対策に活用することが望ましい．一方で攻撃者もアクセス元の地域に応じたクローキングや短期間の有効化等によって観測の回避を図るため，従前の単拠点・短期間の観測では実態を明らかにするのが困難である．そこで本稿では，悪性ホストを複数拠点から継続的に観測することにより，攻撃者による観測回避を困難にしつつ，悪性ホストの時系列での分析や地域性の分析を可能にするシステムを提案する．本稿では，システムを実装し，13 の拠点から 10,334 件の観測対象に対して，2019 年 11 月から 2021 年 8 月の間にかけて観測を実施することにより，25,148,094 件の観測結果を得た．評価では，複数拠点から継続的に観測を実施することにより，悪性ホストの観測可能性を向上できることを示した．また，観測結果を分析することにより，悪性ホストの時系列変化やクローキング状況の実態を明らかにした．

キーワード：悪性ホスト，継続的観測，多拠点観測，クローキング，ジオフェンシング

Analysis of Time Series and Regional Characteristics of Malicious Hosts based on Continuous Observation from Multiple Regions

Shota Fujii[1,2]　Takayuki Sato[1]　Sho Aoki[1]　Yu Tsuda[3]
Nobutaka Kawaguchi[1]　Tomohiro Shigemoto[1]　Masato Terada[1]

Abstract: In recent cyberattacks, malicious hosts, such as C2 servers, have become a critical infrastructure. Therefore, it is desirable to observe malicious hosts and clarify their whole picture for countermeasures. However, the attackers try to avoid the observation by cloaking according based on geolocation or by activating the host for a short period of time; thus, it is difficult to clarify the whole picture by the conventional single-site and short-term observation. In this paper, we propose a system that enables time-series analysis and regional analysis of malicious hosts by continuously observing malicious hosts from multiple regions, making it difficult for attackers to avoid observation. We conducted a study using thirteen sensors by observing 10,334 malicious hosts from Nov. 2019 to Aug. 2021 and obtained 25,148,094 measurements. The evaluation showed the malicious hosts' observability can be improved. We also report the time series of malicious hosts and the reality of the cloaking situation.

Keywords: malicious host, continuous monitoring, multiregional monitoring, cloaking, geofencing

1. はじめに

サイバー攻撃において，悪性ホストは Command and Control Server（以降，C2 サーバ）をはじめとして，攻撃者にとって重要なインフラと化している．こうした状況においては，悪性ホストを観測し，その変動を検出して迅速に対策を取ると共に全貌を理解することが望ましいと言える．

一方で攻撃者は，そうした観測への対策として様々な策を講じることが知られている．代表的な例として，クローキングが挙げられる [1-3]．クローキングは，アクセス元の IP アドレス等を用いてアクセスしたユーザの環境を判別する手法である．例えば，攻撃対象の地域以外のアクセスには悪性コンテンツを返却しないことにより，検出や解析を回避することが可能となる．また，攻撃を実施するタイミングでのみ有効化することにより，観測そのものを困難にする手法も考えられる．

そこで本稿では，クローキングに耐性を有する悪性ホスト観測システムを提案する．提案システムは，複数の地域に設置したセンサから悪性ホストを観測することにより，特定の地域にしか応答を返さないようなクローキングを困難にする．また，悪性ホストを継続的に観測することにより，短期間の有効化や状態の変化を検出する．

本稿における貢献のまとめは，以下の通りである：

- 悪性ホストを複数拠点から継続的に観測することにより，攻撃者によるクローキングを困難にしつつ，悪性ホストの時系列での分析や地域性の分析を可能にするシステムを設計および実装した．
- 処理時間の測定を行い，提案システムが想定しているユースケースである 10,334 件の観測対象に対する 1 日ごとの観測において，その時間内に問題なく処理を完了できることを実証した．
- 提案システムを用いて，13 の拠点から 10,334 件の観測対象に対して 2019 年 11 月から 2021 年 8 月の間にかけて観測を実施し，25,148,094 件の観測結果を得た．

1 株式会社日立製作所
Hitachi Ltd.
2 岡山大学 大学院自然科学研究科
Graduate School of Natural Science and Technology, Okayama University
3 国立研究開発法人 情報通信研究機構
National Institute of Information and Communications Technology

評価では，複数拠点から継続的に観測を実施することにより，悪性ホストの観測可能性を向上できることを示した．また，観測結果を分析することにより，悪性ホストの時系列変化やクローキング状況の実態を明らかにした．

本稿の構成は次の通りである．まず 2 章で悪性ホストの観測に係る背景を説明する．次に，3 章で提案システムの設計と実装を述べる．さらに，4 章で性能評価を実施し，5 章で観測結果を分析する．その後，6 章で議論事項，7 章で関連研究について述べた後，8 章で本稿のまとめを述べる．

2. 悪性ホストの観測

本章では，悪性ホストの観測に係る背景として，悪性ホストの役割と性質およびその観測における課題について述べる．

2.1 悪性ホストの役割と性質

多くのマルウェア・攻撃において外部と通信・連携して攻撃を達成する事例が報告されている．また，この際の主な通信手法の一つとして HTTP 通信が用いられることが知られている．例えば，HTTP 通信を用いて感染端末上のファイルを外部サーバへアップロードする事例 [3]，攻撃用モジュールのダウンロードや攻撃指令の受信を試行する事例 [5] が報告されている．このように，近年のサイバー攻撃では，悪性ホストはより重要な役割を担うようになっている．こうした状況から，攻撃者が悪性ホストを用いてどのように攻撃を達成するかを明らかにするとともにその知見に基づいて攻撃を検知することが重要である．

これに対し，攻撃者は様々な検知回避手法を利用することが知られている．例えば，アクセス元の国や IP アドレスに基づくクローキングや C2 サーバの短期的な有効化が挙げられる．これらの手法により，攻撃者は単一拠点からの観測や短期的な観測を回避することが可能である．

2.2 悪性ホストの観測における課題

悪性ホストの観測は，サイバー攻撃の実態を明らかにすることや攻撃の検知を達成するための主要な手法の一つであり，多くの研究が行われている．ただし，先述の通り攻撃者も様々な検知回避手法を取ることができるため，そうした回避手法に頑強な観測手法を確立する必要がある．

また，悪性ホストは常に悪性の挙動を有し続けるとは限らず，攻撃を実施するタイミングのみ有効化される場合や一時的な休止状態になる場合，恒久的に破棄される場合等がある．こうした変化を検知することが攻撃手法を明らかにする面，攻撃を検知する面の両面で望ましいが，単一時間点のみでの観測では達成することが困難である．

さらに，休止状態の悪性ホストが有効化された場合においても，必ずしも悪性ホストとして有効化したとは限らない．代表的な例として，シンクホール化された場合が挙げられる．シンクホールとは，法執行機関，セキュリティベンダ，および研究者等が悪意のあるドメインをテイクダウンあるいは再取得したことによって無害化された状態のドメインである．シンクホール化されたドメインにアクセスした際，ドメインを訪れた被害者に感染可能性を示す警告ページを表示する場合 [6] があり，この場合は有効化されたとしても悪意のあるものでは無くなっている．これ以外にも，ユーザがドメインを手放したことによってパークドメイン化している場合や悪意を持たないユーザがドメインを再取得し，運用している場合などが考えられる．これを踏まえると，単純に再稼働を検知するだけでなく，シンクホールか否か判定する等の精査が求められると言える．

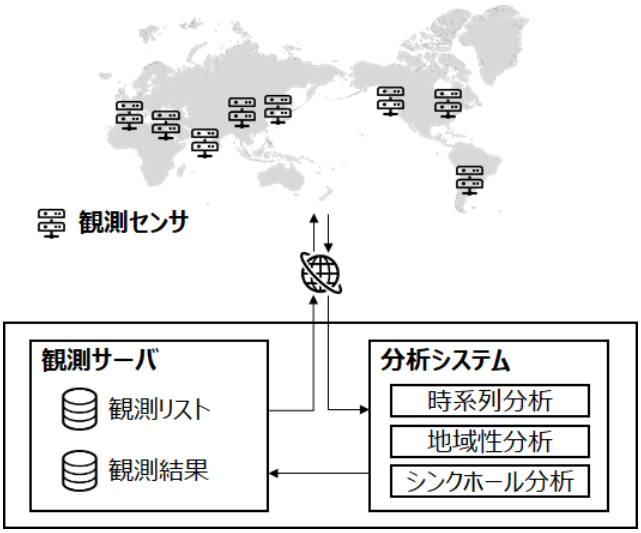

図 1　提案システムの全体像

以上より，悪性ホストの観測においては，以下に示す課題を解決することが望ましいと言える：

（課題1） 継続的な観測の実施
（課題2） 地理的なクローキングへの対応
（課題3） シンクホールの判断

本稿では，上述の課題を解決し，悪性ホストの観測を可能とするシステムを提案する．

3. 設計と実装

本章では，提案システムの設計と実装の詳細を述べる．

3.1 システム概要

まず，（課題 1）に対応するために，悪性ホストに対する観測を一度ではなく定期的かつ継続的に実施する．この際，同一サイトをユニークな ID で管理し，サイトごとに時系列分析が可能な形で保持する．また，（課題 2）への対応として，悪性ホストの観測を実施する観測センサを複数の地域に設置する．これにより，特定の地域からのアクセスにのみ悪性な応答を返すホストの観測可能性を向上する．更に，各課題に対応した観測結果の分析機能を備える．具体的には，時系列での変化を検出する機能（課題 1），地理的

なクローキングを検出する機能（課題 2），およびシンクホールを検出する機能（課題 3）を備える．

上記を踏まえた提案システムの全体像を**図 1** に示す．提案システムは，観測サーバ，観測センサ，および分析システムの 3 コンポーネントから構成され，以下の処理フローによって悪性ホストの観測と分析を実施する．

1. 観測サーバから定期的に観測センサへ観測対象の URL と共に観測指令を送付する．
2. 各観測センサは，受領した観測対象に対して観測を実施する．
3. 観測完了後，観測結果を分析システムで分析する．
4. 観測結果と分析結果を観測サーバに保存する．

以降の節では，各コンポーネントの詳細を述べる．

3.2 観測サーバ

観測サーバは，各地域に設置した観測センサに対して観測対象の URL と共に観測指令を送付する．また，観測結果や分析結果を観測センサや分析システムから受け取り，データベースに保持する．この観測と分析を継続的かつ定期的に実施することにより，悪性ホストの変遷や状態変化の検出を可能とする．観測指令の送信や結果の受領は，REST API で行うこととし，Flask を用いて実装した．

3.3 観測センサ

観測センサは，観測サーバからの観測指令に基づいて悪性ホストの観測を実施する．具体的な観測項目は以下の通りである：

- HTTP GET
- A/AAAA レコードの取得
- スクリーンショットの取得
- ping の送付
- robots.txt の取得とクロール可能性の検証

先述の通り，攻撃に用いられる主な通信手法の一つとして HTTP 通信がある．そこで，各接続先へ HTTP GET を行うことにより，各悪性ホストのステータスコードおよびコンテンツを取得する．この際，User-Agent やポート番号等を設定し，攻撃者の想定するリクエストに可能な限り近づけるようにする．また，HTTP GET に係る項目として，A/AAAA レコードやスクリーンショットを取得する．加えて，ping の送付や robots.txt の存在確認を行い，攻撃者が持つサーバの稼動状態やその構成を確認する．

観測に際しては，正規サーバへの誤観測を抑制するために，robots.txt をまず確認し，クロールを禁止されたページであれば HTTP GET の送付は実施しない．その後，観測した結果を観測サーバに送付する．なお，リダイレクトが検出された際には，リダイレクト先に対しても再帰的に同様の観測を実施する．

観測依頼の受領や観測結果の送付についても，観測サーバと同様に Flask を用いて REST API として実装した．この REST API を用いて JSON として保存した観測結果を観測サーバ側と同期する．また，HTTP GET で得られたコンテンツとスクリーンショットに関しては，rsync によって観測サーバ側と同期することとした．

表 1　観測センサの設置地域とプラットフォーム

#	設置地域	PF
1	米国西部（北カリフォルニア）	AWS
2	米国東部（バージニア北部）	AWS
3	欧州（フランクフルト）	AWS
4	欧州（ロンドン）	AWS
5	欧州（ミラノ）	AWS
6	中東（バーレーン）	AWS
7	南米（サンパウロ）	AWS
8	アジアパシフィック（香港）	AWS
9	アジアパシフィック（ムンバイ）	AWS
10	アジアパシフィック（シンガポール）	AWS
11	アジアパシフィック（シドニー）	AWS
12	アジアパシフィック（東京）	AWS
13	日本	オンプレミス

また，先述の通りクローキングによる観測回避を困難にするために，観測センサを複数地域に設置し，各センサから同時並列的に観測を実施する．この際，広範囲を網羅することをモチベーションに，13 か所に観測センサを設置した．このうち 12 台を Amazon Web Services[a]（AWS）の各リージョンに，残りの 1 台をオンプレミスで日本に設置した．設置地域とプラットフォーム（PF）の内訳は，**表 1** に示すとおりである．

3.4 分析システム

分析システムは，観測データを対象に，時系列，地域性，およびシンクホールの分析を実施する．それぞれの分析項目について，以降の項で述べる．

3.4.1 時系列分析

時系列分析においては，（式 1）を用いて同一観測センサでの時刻 t_1 から時刻 t_n までの変化率を算出し，変化率が閾値よりも高い場合に，時系列での変化があったとして抽出する．（式 1）は，観測結果 S の集合類似度を取り，集合類似度が高い（≒観測結果に差分がない）場合は変化なしとして，集合類似度が低い（≒観測結果に差分がある）場合は変化ありとして検出するものである．なお，観測結果 S は 3.3 節で述べた観測項目から構成されるものである．

$$change\_rate(s_{t1}, s_{t2}, \dots s_{tn}) = 1 - \frac{|\, s_{t1} \cap s_{t2} \dots \cap s_{tn}|}{|\, s_{t1} \cup s_{t2} \dots \cup s_{tn}|} \quad (式\ 1)$$

ここで，接続先の変化を検出した際には，更に観測データを検証する．この際，これまで応答がない状態やエラーのステータスコード（5xx）が返却されていた状態から成功

a Amazon Web Services は，アマゾン テクノロジーズ インコーポレイテッドの商標または登録商標である．

のステータスコード(2xx)が返却されるようになった場合，悪性サイトの活性化，即ち攻撃開始の予兆であるとして，さらに危険な状態として検出する．また，コンテンツの形式がそれ以外のものから実行形式のものに変化した際も同様に予兆として検出する．

3.4.2 地域性分析

地域性分析においては，（式 2）を用いて同一時刻における観測センサ s_1 から s_n までの観測結果の差異を算出し，差異が閾値よりも高い場合に，地域性を有する，即ちアクセス元に応じてクローキングを実施するものとして抽出する．（式 2）は，観測結果 S の集合類似度を取り，集合類似度が高い場合は地域性なしとして，集合類似度が低い場合は地域性ありとして検出するものである．

$$geofenced\_rate(s_{s1}, s_{s2}, \dots s_{sn}) = 1 - \frac{|\, s_{s1} \cap s_{s2} \dots \cap s_{sn}|}{|\, s_{s1} \cup s_{s2} \dots \cup s_{sn}|} \quad \text{(式 2)}$$

3.4.3 シンクホール分析

シンクホールの判定は, 観測データと DNS 情報を利用して実施する．

まず, 観測データを用いる場合の判定フローを**図 2** に示す．観測データを用いるフローでは，コンテンツのハッシュ値やドメイン名，A/AAAA レコードのシンクホールリストを作成し，そのリストとの比較を行う．シンクホールの管理者は, 管理コストの面から自身の管理する 1 つの IP アドレスに複数のドメインを紐づける場合やページに“sinkhole”であることを記した同一のコンテンツを配置する場合があるため，一度リストを作成すると複数のシンクホールを検知することが期待出来る．この判定は観測実施毎に行う．なお，同リストは，提案システムを運用していく中で発見したシンクホールを随時追加することによって構築・更新した．

次に DNS 情報を用いる場合の判定フローを**図 3** に示す．まず, NS レコードがキーワード「sink, hole, dynadot, block, trojan，abuse，virus，malw，hack，black，spam，anti」を含むか否かで判定する．このキーワードリストは，マルウェア Rovnix Downloader がシンクホールを NS レコードベースで判断する際のキーワード [7] であり，これを流用した．なお，同キーワードは，文献 [3] で述べられているドメインをシンクホールするためのネームサーバをカバーしている．また，CNAME レコードや TXT レコードが“sinkhole”を含んでいるかで判断する. 本フローは定期的に DNS 情報を取得し，実施する．

また，シンクホールへのアクセスがあった場合，シンクホールの管理者がアクセス元へマルウェアへの感染可能性等を伝える為にコンタクトを図る場合がある．我々の観測はそうしたものではない為，シンクホール管理者の不要な作業を抑制するためにも，シンクホールであると判断したサイトには以降アクセスしないこととした．

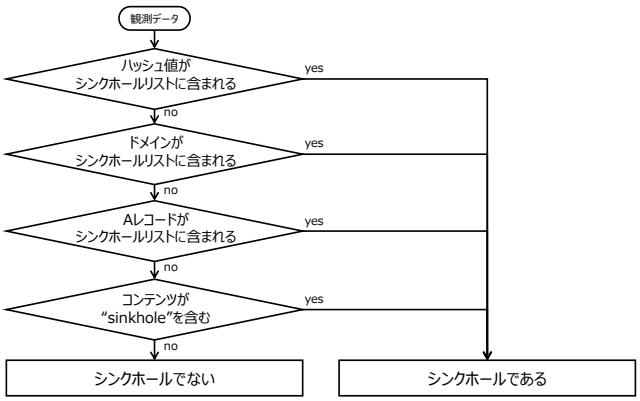

図 2　観測データを用いたシンクホール判定フロー

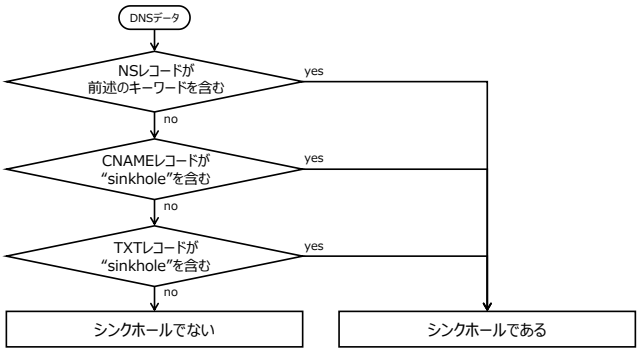

図 3　DNS 情報を用いたシンクホール判定フロー

表 2　評価環境

システム	台数	仕様	
観測サーバ	1	CPU	vCPU 24 cores
		RAM	64 GB
		OS	Ubuntu 18.04 LTS
観測センサ（AWS）	12	CPU	vCPU 2 cores
		RAM	4 GB
		OS	Ubuntu 18.04 LTS
観測センサ（オンプレミス）	1	CPU	vCPU 4 cores
		RAM	8 GB
		OS	Ubuntu 18.04 LTS
VMM	1	CPU	Intel Xeon E5-2670 2.6 GHz×2
		RAM	96 GB
		OS	ESXi 6.7.0

4. 性能評価

本章では，提案システムの観測に係る性能の評価結果を述べる．具体的には，後述する観測対象 10,334 件の悪性ホストの観測を完了するまでに要する時間を処理毎に観測サーバ側で測定した．評価環境は，**表 2** に示す通りである．なお，観測サーバとオンプレミス側の観測センサは，表中の Virtual Machine Monitor（VMM）上で仮想マシンとして動作している．評価結果を**表 3** に示す．表に示す通り，1 回の観測に平均で 17086.4 秒（約 4.7 時間）を要した．この結果から今回の設定においては，例えば 1 日 1 回観測を実施するようなユースケースには充分耐えうると言える．

表 3　処理時間

処理	処理時間（s）			
	最小値	中央値	最大値	平均値
観測指令の送付	9.8	11.9	13.8	12.7
観測の実施	14,341.2	16,681.0	17,240.7	16,740.0
観測結果の分析と登録	321.6	331.7	392.2	333.7
合計	14,672.6	17,024.6	17,646.7	17,086.4

また, 今回の観測センサは, Python[b]言語の concurrent.futures モジュールの ThreadPoolExecutor クラスを用いて 20 スレッドの並行処理として実装している．このスレッド数の増加や計算機資源の増強により，高速化の余地がある．

5. 観測と分析

本章では，提案システムの有効性を検証するために，悪性ホストを多拠点から継続的に観測するとともに，観測結果を分析した結果について述べる．

5.1 観測データ

まず，観測・分析の前段階として観測対象を収集した．ここで，悪性ホストの多くは比較的短命であることが知られている [8]．このため，可能な限り早い段階で悪性ホストを観測対象に加えることが望ましい．そこで，各種情報源の中から，比較的速報性が高いものを選択し，観測対象を収集した．具体的には， URLhaus 等の悪性 URL 共有サイトや Twitter 等の SNS サイトから 10,334 件収集した．左記の観測対象に対して，2019 年 11 月から 2021 年 8 月の間にかけて 1 日 1 回観測を実施し，25,148,094 件の観測結果を得た. なお, 観測対象は運用中に随時追加していたため，期間中の当初から全てのホストを観測してはいない．

分析システムを先述の観測結果に対して適用し, 時系列，地域性，およびシンクホールの分析を実施した．また，提案システムは，複数地域からの継続的な観測により，悪性ホストの観測可能性の向上を図るものである．そこで，観測結果を参照し，観測可能性を定量的に評価した．以降の章では，それぞれの分析結果について述べる．

5.2 時系列変化の分析

本節では，提案システムによって観測した悪性ホスト毎の時系列での変化について述べる．

表 4 に示す通り, 今回の観測対象である 10,334 件の内, 7,879 件が観測期間中に何らかの変化を有していた. なお, ここでは観測毎に（式 1）で算出した変化率が 0 より大きい値を一度以上記録したホスト，即ち 3.3 節で述べた観測項目の内いずれか一つ以上の項目が変化していたホストを時系列変化有として抽出している．

表 4　時系列変化の概要

性質	数	割合（%）
時系列変化無	2,455	23.76
時系列変化有	7,879	76.24
合計	10,334	100.00

表 5　時系列変化の内訳

種別	変化後	数
ステータスコード	2xx	3,285
コンテンツ	executable	2,479

また，時系列変化があったホストからの応答を確認し，どのような手法で地域毎に応答を変更しているか検証した. このうち，3.4.1 項で述べた，提案システムにおいて攻撃開始の予兆であるとして検出するステータスコードの 2xx への変化とコンテンツの実行形式への変化を抽出した結果を **表 5** に示す．なお，今回の観測において実行形式ファイルとしては，DOS-MZ，PE，および ELF 等が含まれており，全てをまとめて executable として表に記載した. 表の通り，ステータスコードが 2xx 以外から 2xx に変化するものもコンテンツが実行形式以外のものから実行形式のものに変化する例も一定数存在することが判明した.

また，今回の観測内で一度ダウンした後一定期間を置いて再度悪性サイトとして活性化したものが散見された．例えば, Dridex の実行ファイルを配布していたホストは, 2020 年 11 月 20 日に Dridex の配布を一旦中断し，HTTP ステータスコード 503 の状態で 1 か月程度ダウンしていたが，その後，2020 年 12 月 11 日よりダウン前と同一のマルウェアを再度配布していた．また，再配布は 2021 年 3 月 25 日まで続いていた．このように，一度ダウンしたように見えた後, 悪性の挙動を再開するような接続先もあったことから，継続的に観測を実施することは有用である．ただし，復活したように見えてもリンク切れのページやシンクホールである場合も散見されたことから、活性化後のサイトが悪性であるか否かの判別を実施する必要あると考えられる．なお，シンクホールに関しては 5.4 節で詳述する．

以上の様に，継続的な観測を実施することにより，変化を伴い，かつ一時的にのみ不審度の高い性質を顕現させるホストであっても観測が可能となる．また，ステータスコードやコンテンツの変化を活用することにより，悪性ホストの再活性化を検出可能なことをケーススタディと共に示した.

5.3 地域性の分析

本節では，提案システムによって検出された地域性を有するホストについて述べる．**表 6** に示す通り，今回の観測対象である 10,334 件の内，1,007 件が地域性を有するホストであった．なお，ここでは観測毎に（式 2）で算出した地域性が 0 より大きい値を一度以上記録したホストを地域

b Python は，パイソン ソフトウェア ファウンデーションの商標または登録商標である．

表 6　地域性の概要

性質	数	割合（%）
地域性無	9,327	90.26
地域性有	1,007	9.74
合計	10,334	100.00

表 7　地域性の内訳

手法	数
地域に応じてコンテンツのハッシュ値を変更	627
地域に応じてコンテンツの形式を変更	323
地域に応じてステータスコードを変更	235
特定の地域にのみ応答	99

性有として抽出している．

また，地域性を有するホストからの応答を確認し，どのような手法で地域毎に応答を変更しているか検証した．この検証結果を**表 7** に示す．今回の観測においては，地域性に係る手法として，大別して 4 つの手法が観測された．最も多かった手法は，地域に応じて異なるハッシュ値のコンテンツを返却するもので，627 件であった．また，地域に応じてコンテンツの形式を変更する手法が 323 件確認された．これは例えば，特定の地域にのみ実行形式のコンテンツを返却し，それ以外の地域にはシンプルな HTML ファイルのような無害なコンテンツを返却するものである．他には，特定の地域にのみ 200 OK のステータスコードを返却するもの（235 件）や特定の地域にのみ応答し，それ以外には応答しないもの（99 件）が確認された．なお，1 つのホストが複数の手法を有している（例：地域に応じてコンテンツとステータスコードの両方を変更する）場合があるため，表 7 に示した全手法の合計数は，地域性を有するホストのユニーク数である 1,007 件とは一致しない．

更に，地域性を有するホストのうち，ドメイン名が割り当てられているもののトップレベルドメイン（TLD）の割合を**図 4** に示す．74 種類の TLD が確認され，.com が最も多い 337 件，.net が 3 番目に多い 28 件と著名な TLD が多く観測された．他方で，.cyou や.club のような比較的新興の TLD も上位に散見された．地域性を有するホストには，短期間で使い捨てられているものが複数見られたが，前述の TLD からなるドメイン名は著名なものよりも比較的安価に取得できることから，攻撃者が使い捨て易いドメイン名として活用しているものと推察される．

以上の様に，複数の観測拠点から観測を実施することにより，地域性を有するホストであっても観測が可能となる．また，観測センサ間の観測結果を比較し，解析回避手法を明らかにした．

5.4　シンクホールの分析

本節では，提案システムによって検出されたシンクホールについて述べる．観測期間中に検出されたシンクホールを**表 8** に示す．

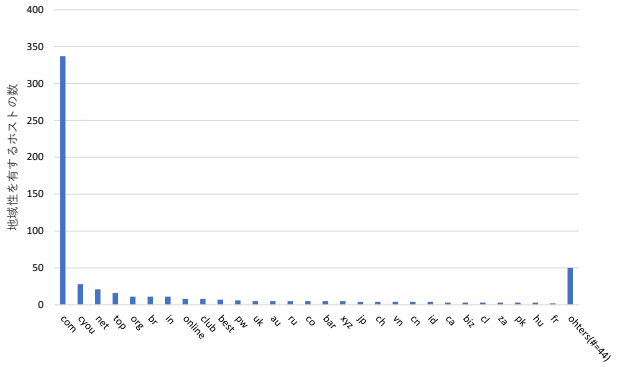

図 4　地域性を有するホストの TLD

表 8　検出されたシンクホール

検出手法	数
コンテンツのハッシュ値	46
NS レコード	23
A レコード	5
CNAME レコード	2
合計	76

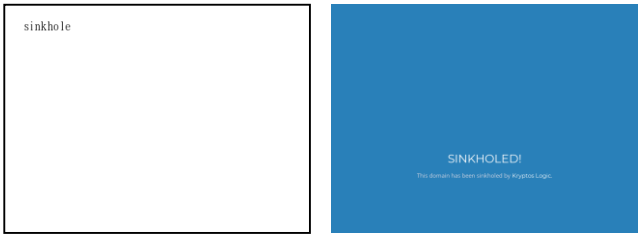

図 5　検出されたシンクホールの例

最も多かったものは，コンテンツのハッシュ値ベースで検知できたシンクホールで，46 件であった．これは，既知のシンクホールにおけるコンテンツのハッシュ値と観測結果のコンテンツのハッシュ値が一致するか否かによって判定されたものである．複数の IP アドレスやドメイン名が同一のコンテンツに紐づいており，同一の管理者によるシンクホールであると推察される．**図 5** は，検出されたシンクホールを示すコンテンツの一例である．

このように，観測データや DNS を利用することにより，想定通りシンクホールが検出可能であることが確認できた．これにより，シンクホール化を悪性サイトの再活性化としてしまう過検知やシンクホールに繰り返し観測を実施してしまうことによる過アクセスを抑制できる．

5.5　観測可能性の検証

本節では，提案システムによる悪性ホストの観測可能性について述べる．提案システムは，先述の通り悪性ホストを継続的に観測することと多拠点から観測することの 2 軸によって観測可能性の向上を図る．観測対象ホストに最初の一度のみ観測を試行したと仮定した場合と拠点が複数ではなかったと仮定した場合の組合せを挙げ，その条件で観測した際に得られるコンテンツのユニーク数を机上で導出

表 9　観測可能性の検証結果

	一度のみ観測	継続的に観測
単一拠点から観測	3,928	49,589
複数拠点から観測	3,970	49,752

した．

導出結果を**表 9** に示す．まず一度のみ観測したと仮定した場合，単一拠点では 3,928 件，複数拠点の場合は 3,970 件のユニークなコンテンツが得られる．これは，単一拠点の場合はクローキングによって得られない可能性があるコンテンツが 42 件存在していたことを意味する．また，観測対象として収集してから実験期間が終了するまでの間に継続的な観測を実施した場合，単一拠点でも 49,589 件，複数拠点の場合はクローキングを行う 163 件を含む 49,752 件のユニークなコンテンツを得られることが分かった．数の面では，観測拠点数に依らず，長期的な観測を実施することで得られるコンテンツが大幅に増加できる．また，観測拠点を増やすことにより，数は多くはないものの，クローキングを行うことから危険性や分析対象としての重要性が高いと思われるコンテンツを最大 163 件増加できることが分かった．さらに，複数の観測拠点から継続的に観測することにより，単一拠点から一度のみ観測する場合に比べて，45,824 件（約 1,167%）得られるコンテンツを増加できることが分かった．

以上の様に，多拠点から継続的に悪性ホストを観測することにより，観測可能性を向上できることを示した．

6. 議論事項

6.1 制限事項

提案システムは，複数の地域に観測センサを設置することにより，アクセス元に応じたクローキングを困難にしている．しかし，別のクローキングとして研究者の IP アドレスを収集後，拒否リストを作成し，その中の IP アドレスからのアクセスに対して無害なコンテンツを返す手法もある [9]．この手法を応用し，観測センサの IP アドレスを拒否リストへ追加することにより，観測を回避できる．また，今回は観測センサを主に AWS に配置したが，AWS の IP レンジは公開されている [10] ため，AWS 上のセンサを拒否リストに登録することは，理論上では可能である．加えて，マルウェアに感染した端末の IP アドレスにのみ悪意のあるコンテンツを返す手法の存在も示唆 [11] されており，提案システムもこの方法によって回避できる．ただし，観測センサの IP アドレスの定期的な変更や AWS 以外への配置により，提案システムの回避をより困難にできると推察される．

また，提案システムは，潜在的に偽陽性を孕んでいる．例えば，5.3 節で述べた地域性に関して，観測センサ間でコンテンツのハッシュ値が異なる場合に地域性があると判定したが，アクセス元に応じてサイトの言語を変更するサイトやサイトに時刻情報のような変動可能性のある情報を含むサイトは，悪意が無いものであっても地域性があると判定してしまう可能性がある．また，時系列分析においても，休止状態のサイトが無害なサイトとして再活性した場合を誤って危険なものと判断してしまう可能性がある．レンタルサーバや良性サイトのファイル共有機能を一時的に悪用されていた場合も，破棄された後に各サイトが 404 Not Found 等を返すため，同様に変化や活性化と過検知してしまう可能性がある．ただし，前者に関しては，今回はハッシュ値が異なるか否かの離散値で判定しているが，ファジーハッシュ等を用いてコンテンツ間の差異を連続値で算出し，差異が大きい場合のみ異なるコンテンツと判断することで，偽陽性を抑制できる．後者に関しても，良性サイトの跡地を表すページ等を許可リストに追加することにより，偽陽性を抑制可能である．

さらに，観測結果は，観測対象の選択バイアスを孕んでいる．例えば，今回実行形式のファイルが多数観測されたが，情報収集先である URLhaus において，収集期間に実行形式ファイルの投稿が集中していたことに起因する面もある．この点に関しては，本質的に解決が困難ではあるものの，本実験においては URLhaus 以外にも Twitter 等の情報源を設けており，緩和を図っている．また，継続的に観測対象を収集し，追加していくことにより，選択バイアスはより緩和されていくと推察される．

6.2 研究倫理

提案システムの観測においては，HTTP GET や ping 等，通常利用で起こり得る通信を実施しており，悪性の通信は実施していない．加えて，HTTP GET の実施前に robots.txt を確認し，クロールが禁止されていた場合はそれ以上の観測を行わないようにしている．

また，悪性ホストへのアクセスを試行する関係上，シンクホールの管理者や IaaS 提供者から注意喚起や各種対応依頼が来る場合がある．我々は，こうした連絡に対応できるよう体制を組んで観測を実施した．なお，観測期間中に 3 度の問い合わせがあり，何れも 24 時間以内に対応を実施した．

7. 関連研究

不審なホストの観測については，多くの研究が行われている．CyberProbe [12] は，アクティブスキャンにより，C2 サーバや Listen 状態のボットを検出するものである．Soska らの手法は，観測データを用いて，サイトが改ざんされるかどうかを予測する手法である [13]．これらの手法は有用であるが，継続的な観測を対象とはしていない．一方で提案システムは，継続的な観測を実施することで，時系列での変化の検出や観測可能性の向上が可能であることを示した．

TARDIS [14] は，CMS を標的とした攻撃の検知手法である．攻撃を検知するために，Web サイトから継続的に取得したコンテンツを活用している．Barron らは，複数の地域にデプロイされたハニーポットを運用し，ハニーポット設置場所の観点から攻撃を分析している [15]．Augur [16] は，Web サイトを複数の地域から継続的に観測することで，検閲の開始や終了を検出している．ICLab [17] や Censored Planet [18] も複数の地域に設置されたセンサを備えた検閲検出システムである．これらのシステムは，継続的かつ複数拠点から観測を行うという点は提案システムと同じであるものの，それぞれ目的が異なっている．提案システムは，継続的かつ複数拠点からの観測結果を活用することにより，クローキングに頑強な形で時系列での変化を検出するとともに，悪性ホストに対する観測可能性の向上を実証した．

8. おわりに

本稿では，悪性ホストの観測システムを提案した．提案システムは，複数拠点から継続的に悪性ホストを観測することにより，攻撃者による観測回避を困難にしつつ，悪性ホストの時系列での分析や地域性の分析を可能にする．

また，提案システムを実装し，10,334 件の観測対象に対して，2019 年 11 月から 2021 年 8 月の間にかけて観測を行うことで，25,148,094 件の観測結果を得た．評価では，観測結果を分析し，提案システムによって悪性ホストの観測可能性が向上できることを示した．また，悪性ホストの時系列変化やクローキング状況の実態を明らかにした．

今後の課題として，観測対象の悪性ホストをより拡充すること，および観測センサの設置地域の拡大やプラットフォームの種類の拡充等によって観測可能性を向上することが挙げられる．観測可能性を向上することが挙げられる．そして，収集したデータの分析を進め，攻撃インフラを明らかにする大規模調査を実施する．

本稿中で使われているシステム・製品名は，各社の商標または登録商標である．

謝辞

本研究は，国立研究開法人情報通信研究機構事業「サイバー攻撃に関する解析作業」で実施したものである．本研究を進めるにあたって有益な助言と協力を頂いた関係各位に深く感謝申し上げる．

参考文献

[1] Wu, B. and Davison, B.D.: Detecting semantic cloaking on the web, *Proc. the 15th international conference on World Wide* (*WWW 2006*), pp. 819-828 (2006).

[2] Invernizzi, L., Thomas, K., Kapravelos, A., et al.: Cloak of Visibility: Detecting When Machines Browse a Different Web, *Proc. 2016 IEEE Symposium on Security and Privacy* (*S&P 2016*), pp. 743-758 (2016).

[3] Zhang, P., Oest, A., Cho, H., et al.: CrawlPhish: Large-Scale Analysis of Client-Side Cloaking Techniques in Phishing, *Proc. 2021 IEEE Symposium on Security and Privacy* (*S&P 2021*), pp. 1109-1124 (2021).

[4] Unit42: Case Study: Emotet Thread Hijacking, an Email Attack Technique, available from <https://unit42.paloaltonetworks.com/emotet-thread-hijacking/> (2021-08-18 accessed).

[5] JPCERT/CC: Malware Used by Lazarus after Network Intrusion, available from <https://blogs.jpcert.or.jp/en/2020/08/Lazarus-malware.html> (2021-08-18 accessed).

[6] Alowaisheq, E., Wang, P., Alrwais, S.A., Liao, X., Wang, X., Alowaisheq, T., Mi, X., Tang, S., and Liu, B.: Cracking the Wall of Confinement: Understanding and Analyzing Malicious Domain Take-downs, *Proc. Network and Distributed Systems Security* (*NDSS*) *Symposium 2019* (2019).

[7] McAFee: Rovnix Downloader Updated with SinkHole and Time Checks, available from <https://www.mcafee.com/blogs/other-blogs/mcafee-labs/rovnix-downloader-sinkhole-time-checks/> (2021-08-18 accessed).

[8] Tanaka, Y., Akiyama, M., and Goto, A.: Analysis of malware download sites by focusing on time series variation of malware, *Journal of Computational Science*, Vol. 22, pp. 301－313 (2017).

[9] Zeeuwen, K., Ripeanu, M., and Beznosov, K.: Improving malicious URL re-evaluation scheduling through an empirical study of malware download centers, *Proc. the 2011 Joint WICOW/AIRWeb Workshop on Web Quality* (*WebQuality 2011*), pp. 42‑49 (2011).

[10] Amazon: AWS IP Address Ranges, available from <http://docs.aws.amazon.com/general/latest/gr/aws-ip-ranges.html> (2021-08-18 accessed).

[11] Mansoori, M., Welch, I., Choo, K.K.R., Maxion, R.A. and Hashemi, S.E.: Real-world IP and network tracking measurement study of malicious websites with HAZOP, *Proc. International Journal of Computers and Applications*, No. 2, Vol. 39, pp. 106–121 (2017).

[12] Nappa, A., Xu, Z., Rafique, M.Z., Caballero, J., and Gu, G.: CyberProbe: Towards Internet-Scale Active Detection of Malicious Servers, *Proc. 2014 Network and Distributed System Security Symposium* (*NDSS 2014*), pp. 1–15 (2014).

[13] Soska, K. and Christin, N.: Automatically detecting vulnerable websites before they turn malicious, *Proc. the 23rd USENIX conference on Security Symposium* (*SEC 2014*), pp. 625–640 (2014).

[14] Kasturi, R.P, Sun, Y., Duan, R., Alrawi, O., Asdar, E., Zhu, V., Kwon, Y., and Saltaformaggio, B.: TARDIS: Rolling Back The Clock On CMS-Targeting Cyber Attacks, *Proc. 2020 IEEE Symposium on Security and Privacy* (*S&P 2020*), pp. 1156–1171 (2020).

[15] Barron, T. and Nikiforakis, N.: Picky Attackers: Quantifying the Role of System Properties on Intruder Behavior, *Proc. the 33rd Annual Computer Security Applications Conference* (*ACSAC 2017*), pp. 387–398 (2017).

[16] Pearce, P., Ensafi, R., Li, F., Feamster, N., and Paxson, V.: Augur: Internet-Wide Detection of Connectivity Disruptions, *Proc. 2017 IEEE Symposium on Security and Privacy* (*S&P 2017*), pp. 427–443 (2017).

[17] Niaki, A.A, Cho, S., Weinberg, X., Hoang, N.P., Razaghpanah, A., Christin, N., and Gill, P.: ICLab: A Global, Longitudinal Internet Censorship Measurement Platform, *Proc. 2020 IEEE Symposium on Security and Privacy* (*S&P 2020*), pp. 135–151 (2020).

[18] Raman, R.S., Shenoy, P., Kohls, K., and Ensafi, R: Censored Planet: An Internet-Wide, Longitudinal Censorship Observatory, *Proc. The ACM Conference on Computer and Communications Security* (*CCS 2020*), pp. 49–66 (2020).