CL-003

継続的かつ複数拠点からの観測に基づく悪性サイトのクローキング調査
Long-term and Multiregional Measurement for Prevalence of Cloaking

藤井 翔太[1)2)] 佐藤 隆行[1)] 青木 翔[1)] 津田 侑[3)]
Shota Fujii Takayuki Sato Sho Aoki Yu Tsuda

川口 信隆[1)] 重本 倫宏[1)] 寺田 真敏[1)]
Nobutaka Kawaguchi Tomohiro Shigemoto Masato Terada

概要

サイバー攻撃で悪用されるホストの中には，潜在的な攻撃対象にのみ悪性コンテンツを返却するクローキングを備えるものが存在し，セキュリティ機構の検知を回避し得ることから深刻な脅威となっている．そこで本研究では，悪性ホストを能動的に観測してクローキングを検出する Stargazer を実装し，約 2 年間の観測と分析を実施した．分析の結果，クローキングが広く悪用されており，その手法が大別して 7 種類あることを明らかにした．また，クローキングを行う悪性ホストには，比較的長期間生存するものやレピュテーションサイトに未登録のコンテンツを含むものが存在することを明らかにした．本研究の結果が各種クローキングの実態の理解や検出手法の開発の一助となることを期待する．

1 はじめに

今日のサイバー攻撃では，攻撃の実行に様々なタイプの悪性ホストが利用されている．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 [1]，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 [2] が報告されている．また，他の HTTP 通信を用いた事例として，マルウェアの配布 [3, 4]，フィッシング [5, 6, 7]，Fraudulent Services [8, 9]，偽 AV スキャナの配布 [10, 11] 等があり，その数は年々増加している [12]．このような状況において，サイバー攻撃に利用される不審サーバへの通信を遮断することが被害を抑制するうえで重要となっている．

これに対し，攻撃者は悪性ホストの検出技術に対する回避技術を実装している．このような回避技術はクローキングと呼ばれている．クローキングは，攻撃対象とする組織や地域からのアクセスか否かを判定し，そうでなければ，即ちセキュリティ研究者を含む攻撃対象外からのアクセスに対しては良性のコンテンツを返却するものである．こうした技術によって，セキュリティ研究者による調査や悪性ホストの自動検知システムを回避することが可能になり，結果として脅威情報が共有されず，対処が遅れることにつながる．

先行研究では，特にフィッシング，広告ネットワークの悪用，偽 AV スキャナの配布等に係るクローキングの特徴が明らかにされている [5, 6, 7, 13, 14]．一方で，これらの手法とは異なる，人手を介さないものに対するクローキングについては，それほど多くは調査されていない．前述の攻撃に係る回避技術を使用しているウェブサイトは，主に自動化されたクローラの訪問と潜在的な人間の被害者を区別することで，攻撃対象かそれ以外を判別し，クローキングを実現している [13]．しかし，それ以外，例えば C2 サーバとの通信や攻撃モジュールのダウンロード等の汎用的な攻撃のフローにおける通信は人手が介在しないことから，攻撃対象かそれ以外かの判別が人とクローラの区別とは必ずしも一致しない．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．

そこで我々は，世界中に能動的に悪性ホストを観測するセンサを設置し，時間的・地域的なクローキング（以降，geofencing）の実態を調査する．この調査のために，世界各地に配置した複数センサから同時並列的かつ長期的に観測可能なプラットフォーム Stargazer [15] を改良・活用し，時間的・地域的なクローキングを回避した上で観測を実施した．具体的には，Stargazer を用いて，13 の拠点から 18,397 件の観測対象に対して 2019 年 11 月から 2022 年 2 月の間にかけて観測を実施し，30,359,410 件の観測結果を得た．この観測結果を分析し，約 15%の不審ホストが geofencing によるクローキングを，約 17%の不審ホストが時間的なクローキングを図っている可能性を示した．また，フィッシング等の特定の脅威だけでなく，クローキング技術は広く悪用されていることが確認された．この中には，単一拠点や単一時間点の観測では検出が困難な地理情報を用いたクローキングや時間ベースのクローキングも含まれている．加えて，geofencing としては 3 種類，時系列のクローキングとしては 4 種類，合計 7 種類の手法が存在することを体系的に明らかにした．さらに，クローキングを行う悪性ホストは，比較的長期的に生存するものやコンテンツが過去に VirusTotal 上で報告されていないものを含んでいることが確認された．

これらの分析により，これまで定性的に語られていた悪性ホストの地理的，および時系列でのクローキングについて，その理解を深めた．Stargazer は，geofencing や時系列でのクローキングを自動で明らかにし，攻撃者による検知回避を困難にするだけでなく，今後のクローキング検知手法の開発にも貢献するものであると考えている．本研究の主な貢献は以下の通りである．

- Stargazer を用いて，13 の拠点から 18,397 件の観測対象に対して 2019 年 11 月から 2022 年 2 月の間にかけて観測を実施し，30,359,410 件の観測結果を得た．観測結果を定量的に分析し，クローキングは脅威の種別に関係なく普遍的に悪用されていることを確認した．また，クローキングの種別が時系列の 3 種類と geofencing の 4 種類，合計 7 種類に大別できることを明らかにした．さらに，クローキングさ

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.

Copyright © 2022 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

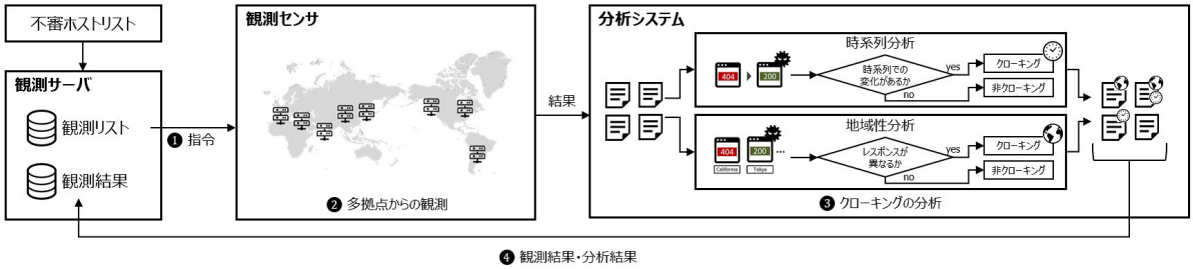


図 1　Stargazer の全体像

れているコンテンツは，そうでないものに比べて VirusTotal のようなレピュテーションサイトへの投稿数が有意に少ないことが分かった（§4）．

- 観測結果の中からより特徴的なものをケーススタディとして定性的に深堀した．具体的には，一度非稼働状態になった後再稼働することによりトータルでは長期的に生存していたホストや短期間で頻繁に IP アドレスやコンテンツを変更することで検出回避を図るホストを示した．また，同一キャンペーンのクローキングサイトの中でも，検知されやすい検体のダウンロードサイトは比較的生存期間が短く使い捨てであることに対し，後段の C2 サーバは生存期間が長く使いまわされているという風に，その役割によって異なる特性を有することを示した（§5）．

2　背景

2.1　サイバー攻撃におけるクローキング技術

従来より，インターネットサービスにおいては，サイトの訪問者にあわせて提供するコンテンツを変更するものがある．例えば，訪問者の端末やブラウザを識別してモバイル用のコンテンツと PC 用のコンテンツを返すものや訪問者の地域を識別してそれに合わせてコンテンツの言語を変更するものがある．これらはユーザの利用性の向上を考慮した良性のものである．他方で，こうした識別技術は，前述のように，サイバー攻撃ではクローキングが悪用されている．

典型的なクローキングとしては，訪問者がボットか否かを識別し，ボットには良性コンテンツを，人には悪性コンテンツを返すものがある．他には，攻撃対象の地域や活用する脆弱性を含みうる対象にのみ悪性コンテンツを返すものが存在する．これにより，的確に攻撃対象にのみ悪性コンテンツを配布するとともに，自動的な検知システムやセキュリティ研究者からの発見を遅らせる，あるいは回避することが可能となる．

2.2　クローキングの観測に係る課題

上記の背景を受け，これまでに多くのクローキング検出手法の提案やクローキング手法の調査が行われている．特に，フィッシングサイトに関しては，より多くの攻撃対象にリーチするために，Search Engine Optimization によって検索エンジンの上位に表示させたい一方で，研究者には分析されたくないという相反する要求を有することから，クローキングが実施されがちであることが知られており，研究も盛んである [13, 14, 16, 17]．ただし，既存研究は，多くの場合特定のカテゴリの攻撃，特に上述のフィッシングに焦点を当てている．また，これらの研究は，web ブラウザでアクセスした際のスクリーンショットの差分や JavaSript の構造等，クローキングサイトをはじめとした Web ページを有するホスト特有の特徴量を利用した手法である．一方で，これらの攻撃に限らない悪性ホストにおいてもクローキングが行われているという報告は存在する．例えば，文献 [4] の調査では，一部のファミリや種別のマルウェアが地域性を有していることが手動の分析によって導出されている．このように，幅広いカテゴリの攻撃において普遍的にクローキングは利用されているという証拠はあるものの，先行研究では大規模な調査はほとんど実施されていない．

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考えられる．こうした状況から，継続的な観測を行っている研究もある．ただし，配布される検体やコンテンツの変遷・生存期間の分析等はされているものの，時系列でのクローキングの観点での大規模な調査はほとんど実施されておらず，その実態は明らかになっていないと言える．

3　手法

3.1　全体像

本章では，時間的・地域的なクローキングの実態の調査を可能にするシステムである Stargazer を紹介する．

まず，時間的なクローキングの検出を可能にするために，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，geofencing の検出を可能にするために，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．

図 1 に，Stargazer の全体像を示す．Stargazer は，観測サーバ，観測センサ，および分析システムの 3 つから構成されており，悪性ホストに対する観測と分析を以下の手順で実施する．

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説明する．

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

Copyright © 2022 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

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	US West (N. California)	AWS
2	US East (N. Virginia)	AWS
3	Europe (Frankfurt)	AWS
4	Europe (London)	AWS
5	Europe (Milan)	AWS
6	Middle East (Bahrain)	AWS
7	South America (Sao Paulo)	AWS
8	Asia Pacific (Hong Kong)	AWS
9	Asia Pacific (Mumbai)	AWS
10	Asia Pacific (Singapore)	AWS
11	Asia Pacific (Sydney)	AWS
12	Asia Pacific (Tokyo)	AWS
13	Japan	on-premise

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．

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．具体的には，観測によって以下を取得する．

- HTTP GET によって得られるコンテンツ
- A/AAAA レコード
- スクリーンショット
- ping の応答結果

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．また，HTTP GET に係る項目として，A/AAAA レコードを取得すると共に，別途ヘッドレスブラウザを用いてスクリーンショットを取得する．加えて，ping を送付し，攻撃者が持つサーバの稼動状態を確認する．さらに，後述のシンクホール判定のため，不定期に各種 DNS レコードを取得する．

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．

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（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．なお，表 1 に示した各センサが正しく想定した地域に紐づいていることは，事前に GeoIP2[1] を用いて確認している．

3.4　分析システム

分析システムは，観測データを対象に，時系列と地域性の観点での分析を実施する．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* 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のカテゴリデータから構成される集合である．

ここで，コンテンツに関する変化には，アクセス元に応じて言語を変えるサイトや時刻情報などの可変情報を含むサイト等の良性な範囲での小規模な変化と無害なファイルからマルウェアに変化するような悪性のものが存在する．両者の変化を同じものとして扱うと，前者の良性な範囲での変化を悪性として過検知する可能性がある．そこで，コンテンツについては，ハッシュ値が異なるかどうかという離散的な値を用いるのではなく，ファジーハッシュを用いてコンテンツ間の類似度を連続値で計算することにより，微小な良性の変化に対する誤検知の抑制を図っている．なお，今回は，*ssdeep*[2] をファジーハッシュとして活用した．

$$change\_rate(s_{t-1}, s_t) = 1 - \left(\frac{s_{t-1} \wedge s_t}{s_{t-1} \vee s_t} * Sim(Content(s_{t-1}), Content(s_t))\right) \quad (1)$$

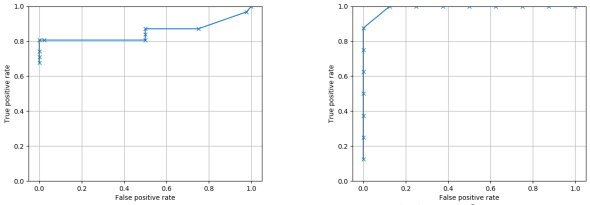
3.4.2　地域性分析

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 geofencing 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．コンテンツについては，時系列変化の場合と同様に，類似度は *ssdeep* を用いて連続値として算出する．地域性分析では，アクセス元によって言語が変わるサイトは，その変化が軽微で悪意がなくてもクローキングと判定されることがある．そこで，誤検知を抑制するために，ファジーハッシュとして *ssdeep* を活用した．

$$geofenced\_rate(s_1, ..., s_n) = 1 - \left(\frac{s_1 \wedge ... \wedge s_n}{s_1 \vee ... \vee s_n} * Sim(Content(s_1), ..., Content(s_n))\right) \quad (2)$$

3.5　予備評価

本節では，3.4 節で述べた式（1）と式（2）の精度を評価するともに各式の閾値の決定を図る．評価のために，Stargazer の観測結果の一部から手動で geofencing や時系列のクローキングを実施している悪性ホストのも



(A) 時系列のクローキング　　(B) geofencing

図 2　クローキング検出における ROC 曲線

1）https://www.maxmind.com/en/geoip2-databases

2）https://ssdeep-project.github.io/ssdeep/

Copyright © 2022 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

表 2　時系列変化の種別と観測数

種別	観測数
応答なしからの復活	1,860　(10.11%)
ステータスコード 2xx 以外から 2xx への変更	2,419　(13.14%)
コンテンツの変更	2,535　(13.78%)
合計（ユニーク数）	3,128　(17.00%)

のを抽出し，データセットを構築した．その後，式（1）と式（2）をデータセットに適用し，時系列での変化と geofencing のスコアを算出し，ROC 曲線を作成した．この結果を図 2 に示す．時系列のクローキングに関しては閾値 0.35（TPR: 81.2%，FPR: 5.0%），geofencing に関しては閾値 0.23（TPR: 92.1%，FPR: 2.2%）を選択した．本観測は，クローキング手法の実態を明らかにすることが目的であるため，偽陽性をある程度許容しても見落としが少なくなるように閾値を選択した．以降では，上記の閾値を活用して分析を進める．

4　観測と分析

本章では，Stargazer の有効性を検証するために，悪性ホストの多拠点からの継続的な観測結果とその分析結果を述べる．

4.1　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 [18]．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[3] 等の悪性 URL 共有サイトや Twitter[4] 等のセキュリティ分析者・ベンダ等が IOC を共有することのある SNS サイトから 18,397 件収集した．左記の観測対象に対して，2019 年 11 月から 2022 年 2 月の間にかけて凡そ 1 日 1 回観測を実施し，30,359,410 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，時系列と地域性に係る分析を実施した．また，Stargazer 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．

なお，観測結果のうち，シンクホールの可能性が高いものは除外した．シンクホールか否かの判定は，文献 [19] 等を参考に，観測データと DNS 情報を利用して実施した．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ことでシンクホールか否かを判定した．同リストは，Stargazer 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．また，定期的に DNS 情報を取得し，NS レコード，CNAME レコード，および TXT レコードが *sinkhole* 等のキーワードを含んでいる場合，シンクホールと判定した．

4.2　時系列分析

本節では，Stargazer を用いて観測した各悪性ホストの時系列変化について述べる．

表 2 に示すように, 3,128 件のホストにおいて，時系列

3）https://urlhaus.abuse.ch/
4）https://twitter.com/

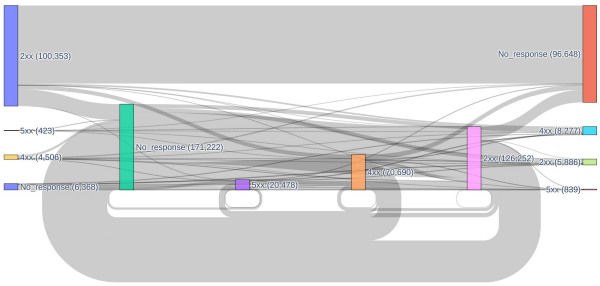


図 3　ステータスコードの時系列変化

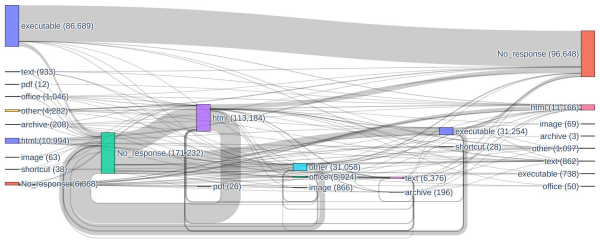


図 4　コンテンツの時系列変化

でのクローキングの可能性がある変化が観測された．また，時間での変化を確認したところ，大きく 3 種類に分けられることが判明した．具体的には，応答なしからの復活（1,860 件），ステータスコードの *2xx* 以外から *2xx* への変更（2,419 件），およびコンテンツの変更（2,535 件）である．なお，1 つのホストが複数の手法を用いている場合があるため，それぞれの合計数とクローキングサイトのユニーク数（3,128 件）は一致しない．

また，時系列変化をステータスコードとコンテンツの両観点から，より詳細に検証した．時系列におけるステータスコードの遷移を図 3 に，コンテンツの遷移を図 4 に示す．各図において，左端が最初に観測した際の状態，中央が中間での状態，右端が最終的に観測した際の状態である．なお，中間状態から別の中間状態への遷移も含まれている．また，コンテンツに関しては，同じ種類のものは丸めて記載している．例えば，今回の観測において実行ファイルとしては，DOS-MZ，PE，および ELF が含まれており，図ではすべて *executable* としてまとめて記載している．各項目の内訳は，表 3 の通りである．ステータスコードで最も多かったのは，*200 OK* から無応答の状態への遷移であった．これは主に，攻撃終了に伴い，ホストを放棄した場合に発生すると考えられる．ただし，無応答状態からホストが復活するようなクローキングの可能性が高いケースも確認されている．中間状態内での遷移は，*200 OK*，*404 Not Found*，および無応答を行き来するものが多い．また，それ以外から *2xx* への変化も一定数存在し，一旦応答がなくても再活性化する可能性を考慮する必要があることが示唆された．

コンテンツに関しても，ステータスコードと同様に，無応答状態への遷移が最も多かった．これに関しても，ホストを放棄した際の攻撃終了が主な原因であると考えられる．その他の変化としては，HTML への変遷が多く見られた．また，観測回避の一種として，実行ファイルを一度配布し，2 回目以降は無害な HTML に置き換えるケースもあった．このような，より危険な変更を検知することで，攻撃に対する対応速度や網羅率を向上することが期待できる．

Copyright © 2022 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

表 3　コンテンツの種別と各カテゴリのコンポーネント

コンテンツの種別	コンポーネント
executable	DOS-MZ, PE, ELF, shell script
html	HTML
image	JPEG, PNG
office	Excel, Word
archive	Zip, ACE, JAR
text	Text
pdf	PDF
shortcut	MS Windows shortcut
other	empty, data, very short file (no magic)

表 4　Geofencing の種別と観測数

種別	観測数
地域に応じてコンテンツのハッシュ値を変更	2,708　(14.71%)
地域に応じてコンテンツの形式を変更	1,467　(7.97%)
地域に応じてステータスコードを変更	1,553　(8.44%)
特定の地域にのみ応答	1,102　(5.99%)
合計（ユニーク数）	2,716　(14.76%)

以上のように，時系列でのクローキング手法がいくつか観測され，大別して 3 つのカテゴリに分類できることが分かった．また，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．さらに，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である可能性を示した．

4.3　地域性分析

本節では，Stargazer を用いて検出した geofencing を行う悪性ホストについて述べる．

今回観察された geofencing を表 4 に示す．今回の観測では，geofencing にかかる手法として，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，2,708 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 1,467 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は 1,553 件，特定の地域にのみ応答し，それ以外には応答しないものが 1,102 件確認された．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 4 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 2,716 件とは一致しない．

更に，geofencing を行うホストのうち，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 5 に示す．74 種類の TLD が確認され，*.com* が最も多い 1,001 件，*.net* が 3 番目に多い 33 件と著名な TLD が多く観測された．他方で，*.cyou* や *.club* のような比較的新しい TLD も上位に散見された．これらの TLD は，著名な TLD に比べて比較的安価に取得できるため，攻撃者が使い捨て易いドメイン名として利用していることが推測される．

Geofencing によって対象とされた地域の数を表 5 に示す．83.98%は 1 つの地域を対象としているが，2 つ以上の地域を対象としているものもある．また，オンプレミスのセンサが配置されているのは日本のみであり，こ

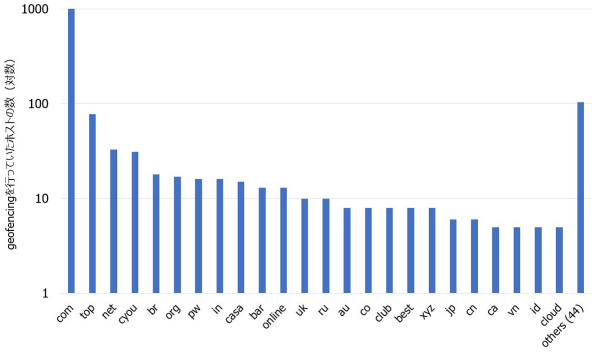


図 5　Geofencing を行うホストの TLD

表 5　同一 geofencing における対象地域の数

対象地域数	観測数
1	2,281　(83.98%)
2	222　(8.17%)
3	93　(3.43%)
4 以上	120　(4.42%)
合計	2,716　(100.00%)

のことが関係している可能性がある。また，各地域の相関係数を図 6 に示す．前述のとおり，ほとんどが単一地域を対象としているが，フランクフルト，ミラノ，およびバーレーンの間には緩やかな相関が見られる．これらを同時に標的とした攻撃を調査したところ，少なくともドイツ語話者とイタリア語話者を標的とした *TA551/Shathak* グループに属するものが多く報告されていた [20]．我々の調べた範囲ではこれに関する情報は見つからなかったものの，バーレーンも同攻撃キャンペーンに乗じて狙われていた可能性があると推察される．

以上の様に，複数の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．

4.4　観測可能性

本節では，悪性コンテンツの観測可能性を検証する．具体的には，観測の結果得られたコンテンツのハッシュ値を VirusTotal[5] に問い合わせ，存在するか否かを検証する．実験のために，2,208 件のコンテンツをランダムに抽出し，VirusTotal に存在するかとクローキングされていたものか否かの 2 軸でカウントした．なお，HTML のコンテンツはその多くが良性なため，今回は除外した．

検証結果を表 6 に示す．まず，地域性を有さない 1,224 のコンテンツ（81.06%）が VirusTotal で確認された．これは，コンテンツがクローキングされておらず，地域・時間帯に依らずアクセス可能であったために，高い割合で VirusTotal に投稿されたことが理由であると推察される．一方で，クローキングされていたコンテンツのうち，60.32%が VirusTotal では確認できなかった．このように，クローキングしているホストから配信されたコンテンツは，クローキングしていないホストから配信されたコンテンツよりも VirusTotal での存在率が低くなっている．また，クローキングの有無と VirusTotal での存在有無に関連がないことを帰無仮説とし，カイ 2 乗検定を実施したところ，有意水準 0.05 で仮説が棄却された ($p \fallingdotseq 3.22E\text{-}83 < 0.05$)．これらの結果から，Stargazer

5）https://www.virustotal.com/

Copyright © 2022 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

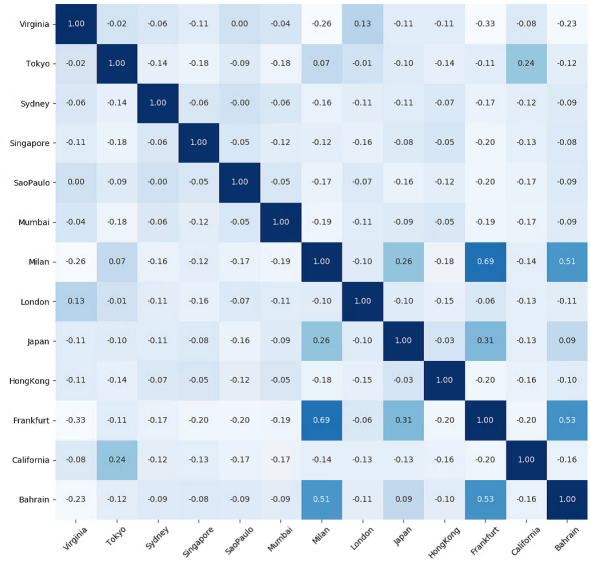


図 6　対象地域同士の相関係数

表 6　geofence されたコンテンツの VirusTotal での存在性

	存在		非存在		合計	
Geofencing 有	277	(39.68%)	421	(60.32%)	698	(100%)
Geofencing 無	1,224	(81.06%)	286	(18.94%)	1,510	(100%)
合計					2,208	-

によって観測可能性が向上できていると考えられる．

5　ケーススタディ

本章では，悪性ホストの時系列変化やクローキングに係るケーススタディについて述べる．典型的な例としては，先述した活動状態・休止状態の間で遷移するような観測回避や geofencing による観測回避もある．ここでは，それらの中から，より特徴的なものをピックアップし，ケーススタディとして示す．

ケース 1：時系列のクローキング．一つ目のケースは，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である．本サイトは，pdf 形式に偽装した *Dridex* の実行ファイルを配布するものとして 2020 年 11 月 19 日ごろに報告されていた．観測開始後，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，最終的に *404 Not Found* になるまで，再配布は 2021 年 3 月 25 日頃まで続いていた．結果として，一度は *503* になったものの．再活性化することによって実質的に 4 か月ほど稼働していた．これは，生存期間が短いと言われる悪性ホストの中では，比較的長い生存期間である．攻撃者側の観点では，数週間停止していたサイトを再稼働させるという単純な操作で，1 つのドメインに関連する攻撃サイトの寿命を延ばし，攻撃に対する投資対効果を向上させることができると言える．

ケース 2：Geofencing．2 つ目のケースは，ミラノとバーレーンからのアクセスには 2 次検体（*Ursnif*）をダウンロードするスクリプト（図 7 右）を，それ以外からのアクセスには *403 Forbidden*（図 7 左）を返すものである．同様の特徴を有するホストが同時期に複

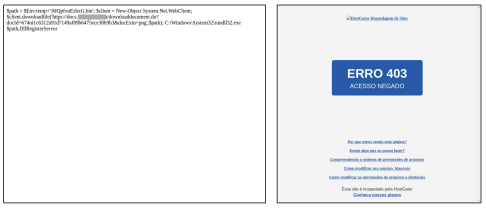


図 7　クローキングされていた悪性ホストの例

数観測された．類似の特徴を有する悪性ホストとしては，攻撃対象以外には，*403 Forbidden* ではなく，*404 Not Found* を返すものや *200 OK* と共に空のファイルを返すようなものも散見された．また，攻撃対象以外には *'Connection aborted.', RemoteDisconnected('Remote end closed connection without response',)* 等のエラーを出し，応答を返さないものも散見された．また，中には攻撃対象に対して最初の一度のみ実行可能ファイル等の悪意ある応答を返却し，同一センサからの 2 回目以降のアクセスには他の攻撃対象外の地域からのアクセスと同様に無害なものを返すものもあった．何れのケースでも，攻撃対象に対してはファイル共有サービス上に配置した DLL をファイル名をランダム化したうえでダウンロードする共通点が見られた．

また，上記のキャンペーンに関連すると思われるサイト群に関しては，最初に実行される検体のダウンロードサイト・C2 サーバの両方ともクローキングがされていた．ただし，検体のダウンロードサイトは 1 週間前後の生存期間なのに対し，C2 サーバの生存期間は 2 週間以上でかつ各検体の C2 サーバとして使いまわされていた．これは，アクセス性が必要なために比較的検知されやすい最初の検体配布サイトは生存期間を短くし，後段の C2 サーバは生存期間を長くしつつ複数検体で使いまわすことにより，攻撃にかかる費用を抑制しているものと推察される．

ケース 3：複数の時系列クローキングと AWS 検知の組み合わせ．3 つ目のケースは，複数の time-based cloaking を組み合わせるとともに，AWS の検知を実施していたものである．同じドメイン下の異なるパスで 2 件類似例が確認されたため，その両方を本節にて紹介する．図 8 に，それぞれの観測タイミングごとのステータスコード，IP アドレス，コンテンツの種類と実行可能ファイルについてはハッシュ値（SHA256/ssdeep）を示す．これらは，いずれも *RedLine Stealer* の配布ホストとして 2021 年 5 月 20 日ごろに URLhaus にて報告されていたものであり，Stargazer では 2021 年 6 月 16 日より観測を開始した．まず，図から短いスパンで IP アドレスが変遷すると共に，IP アドレスに応じたアクセス先によってステータスコード（*200 OK*，*404 Not Found*，*503 Service Temporarily Unavailable*）が変わり，かつ *200 OK* のときも異なるサーバからは異なる実行可能ファイルが降ってくることが確認できる．これは，攻撃者が随時サーバ上のファイルを操作していることを示唆している．また，同じ IP アドレスの場合でもタイミングによって異なる実行可能ファイルがダウンロードされたり，ステータスコードが異なるものとなっている．これらの検体はいずれもパックされており，単純に ssdeep を用いて各検体間の類似性を計算しても結果は 0 であった．ただし，オン

Copyright © 2022 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

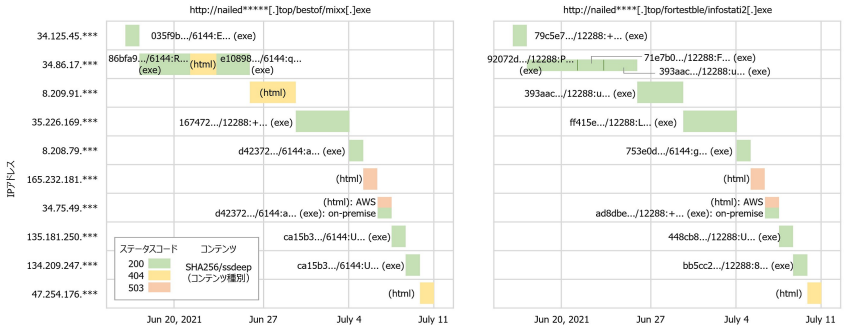


図 8　同一ドメインに紐づけられた IP アドレスとステータスコードの時系列での変化

ラインサンドボックス（*JoeSandbox*[6]）で各検体の動作レポートを確認したところ，いずれも *RedLine* と判定されており，かつ同一の C2 サーバに接続していたため，同等のあるいは類似した検体であると推察される．

ここで，各実行ファイルの VirusTotal での first submission を確認したところ，ほとんどが初観測日時よりも後になっていた．これは，観測時点では VirusTotal にない新規検体を常に配布することによって，ハッシュ値ベースの検出を逃れる意図があったものと推察される．加えて，34[.]75[.]49[.]***に紐づいた際に，AWS を検知するような挙動が確認された．具体的には，AWS からアクセスした際には *503 Service Temporarily Unavailable* が返却されてきたのに対し，オンプレミスの環境からアクセスした際には，*200 OK* とともに実行可能ファイルが返却された．同じホストの別のタイミングでは，AWS の検知を行うような動作は見受けられなかったため，設定ミス等による意図しない判別の可能性もあるものの，AWS を検知し，場合によってはその如何に応じた挙動の変更が可能なことが確認された．

さらに，スクリーンショットを確認したところ，アクセスのタイミングや実際のステータスコードに関わらず，*404 Not Found* とのみ記載された html が返却されるようになっていた．これは，ブラウザからアクセスした際に *404* に見せかけることによって，検知を回避する狙いがあるとみられる．本ホストは，2021 年 7 月 11 日まで稼働しており，報告日から約 2 か月弱の間生存していた計算になる．

本例は，IP アドレスを短期間で付け替える Fast-fluxing [21] として知られる手法に加えて，コンテンツ，ステータスコード，AWS の検知，およびブラウザの検知など，様々な手法を駆使して検知の回避を試みている．結果として，2 か月弱生存しており，攻撃者の投資対効果を高めるものになっている．

本章では，前の章で述べた各クローキングのうち，より特徴的なものをケーススタディとして紹介した．今回は Stargazer で観測できたものの，何れも単純な観測では検知をすり抜ける可能性が高いものであり，事実として比較的長期間生存していたものが大半であった．こうしたものについても，Stargazer を活用し，早期検出を行っていくことが重要であると考えられる．

6　議論

6.1　制限事項

Stargazer 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 [22]．また，ケーススタディに示したように AWS の IP レンジからのアクセスを拒否するものもある．さらに，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 [23] されており，この方法によっても Stargazer は回避できる．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，Stargazer の回避をより困難にできると推察される．

また，Stargazer は，潜在的に偽陽性を孕んでいる．例えば，時系列分析において，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，こうした偽陽性は，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で抑制できる．

6.2　研究倫理

Stargazer 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．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

クローキングの検出に関する研究は，以前から多く行われている．CrawlPhish [13] は，Phishing に係るクローキングを検出するとともに，大規模な分析を実施している．文献 [14] は，クローキング検出手法を提案し，検索・広告に紐づく URL に係るクローキングの大規模な調査と分析を実施している．Wu らは，Web クローラーと Web ブラウザによって取得したコンテンツの差分を用いることで，セマンティッククローキングページを検出する手法を提案している [16]．Mansoori らは，6 つの地域に位置するクライアントから悪意のあるホストに同時アクセスすることで geofencing の検出を試み，TLD と対象地域の相関を明らかにした [24]．しかし，これらのうち多くの手法は，スクリーンショットの差分や JavaSript の構造等，クローキングサイトをはじめとした Web ページを有するホスト特有の特徴量を利用した

6）https://www.joesandbox.com/

Copyright © 2022 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

テーラードな手法である．本研究では，特定の種別の悪性ホストには限定せず，多拠点から長期的に観測し，その差分を利用するという汎用的な手法を示した．また，一部クローキングを特定の種別の攻撃に限定せずに観測した研究もある [24] ものの，クローキングに対する継続的な観測・解析は行われていない．本研究では，より広い観測範囲を持つ観測センサを用いた継続的な観測により，時系列でのクローキングを含める形で観測と分析を図った．

TARDIS [25] 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 [26]．Augur [27] 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 [28] や Censored Planet [29] 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り，HTML の構造や DNS 情報を用いて検閲されているか否かを判定している．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 Stargazer と同じであるものの，それぞれ目的が異なっている．Stargazer 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

本研究では，世界中に配置したセンサから能動的に悪性ホストを観測し，時間的・地域的なクローキングを検出する Stargazer を用いて，のべ 18,397 件の悪性ホストを約 2 年間観測し，その実態の解明を図った．観測の結果，従来より確認されていたようなフィッシングに限らず，クローキング技術は悪性ホストに偏在することが確認された．これらの中には，時系列でのクローキング，geofencing，およびその組み合わせによって検知回避を行う悪性ホストが含まれており，これらは Stargazer のように，多拠点かつ長期的な観測を行わなければ発見が難しいと推察される．

また，クローキングを行う悪性ホストは，比較的長期的に生存するものやコンテンツが VirusTotal に存在しないものを含んでいることが確認された．これらの結果は，クローキング技術の実態を明らかにするとともに，既存技術では観測するのが容易でないクローキングサイトが存在することを示唆している．我々の発見は，今後の観測システムの設計の指針を示すとともに，クローキング検出手法の開発にも貢献するものと考えている．

参考文献

[1] Unit42: Case Study: Emotet Thread Hijacking, an Email Attack Technique. https://unit42.paloaltonetworks.com/emotet-thread-hijacking/.

[2] JPCERT/CC: Malware Used by Lazarus after Network Intrusion. https://blogs.jpcert.or.jp/en/2020/08/Lazarus-malware.html.

[3] Invernizzi, et al. Nazca: Detecting Malware Distribution in Large-Scale Networks. In *21st Symposium on Network and Distributed System Security Symposium*, 2014.

[4] Juan Caballero, et al. Measuring pay-per-install: The commoditization of malware distribution. In *20th USENIX Security Symposium*, 2011.

[5] Adam Oest, et al. Inside a phisher's mind: Understanding the anti-phishing ecosystem through phishing kit analysis. In *2018 APWG Symposium on Electronic Crime Researchx*, pp. 1–12, 2018.

[6] Adam Oest, et al. Phishtime: Continuous longitudinal measurement of the effectiveness of anti-phishing blacklists. In *29th USENIX Security Symposium*, pp. 379–396, 2020.

[7] Adam Oest, et al. Sunrise to sunset: Analyzing the end-to-end life cycle and effectiveness of phishing attacks at scale. In *29th USENIX Security Symposium*, pp. 361–377, 2020.

[8] Emiliano De Cristofaro, et al. Paying for likes? understanding facebook like fraud using honeypots. In *2014 Conference on Internet Measurement Conference*, pp. 129–136, 2014.

[9] Shehroze Farooqi, et al. Characterizing key stakeholders in an online black-hat marketplace. In *2017 APWG Symposium on Electronic Crime Research*, pp. 17–27, 2017.

[10] Marco others Cova. An analysis of rogue av campaigns. In *Recent Advances in Intrusion Detection*, pp. 442–463, 2010.

[11] Takashi Koide, et al. It never rains but it pours: Analyzing and detecting fake removal information advertisement sites. In *Detection of Intrusions and Malware, and Vulnerability Assessment*, pp. 171–191, 2020.

[12] Benjamin Zi Hao Zhao, et al. A decade of mal-activity reporting: A retrospective analysis of internet malicious activity blacklists. In *2019 ACM Asia Conference on Computer and Communications Security*, pp. 193–205, 2019.

[13] Penghui Zhang, et al. Crawlphish: Large-scale analysis of client-side cloaking techniques in phishing. *2021 IEEE Symposium on Security and Privacy*, pp. 1109–1124, 2021.

[14] Luca Invernizzi, et al. Cloak of visibility: Detecting when machines browse a different web. In *2016 IEEE Symposium on Security and Privacy*, pp. 743–758, 2016.

[15] 藤井翔太ほか. 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. コンピュータセキュリティシンポジウム 2021 論文集, pp. 357–364, 2021.

[16] Baoning Wu, et al. Detecting semantic cloaking on the web. In *15th International Conference on World Wide Web*, pp. 819–828, 2006.

[17] Nayanamana Samarasinghe, et al. On cloaking behaviors of malicious websites. *Computers & Security*, Vol. 101, , 2021.

[18] Mitsuaki Akiyama, et al. Design and implementation of high interaction client honeypot for drive-by-download attacks. *IEICE Transactions on Communications*, Vol. E93-B, No. 5, pp. 1131–1139, 2010.

[19] Eihal Alowaisheq, et al. Cracking the wall of confinement: Understanding and analyzing malicious domain take-downs. In *26th Annual Network and Distributed System Security Symposium*, 2019.

[20] MITRE ATT&CK: TA551, GOLD CABIN, Shathak, Group G0127. https://attack.mitre.org/groups/G0127/.

[21] Ziji Guo, et al. Active Probing-based Schemes and Data Analytics for Investigating Malicious Fast-Flux Web-Cloaking based Domains. In *International Conference on Computer Communications and Networks*, pp. 1–9, 2018.

[22] Kyle Zeeuwen, et al. Improving malicious url re-evaluation scheduling through an empirical study of malware download centers. In *2011 Joint WICOW/AIRWeb Workshop on Web Quality*, pp. 42–49, 2011.

[23] Masood Mansoori, et al. Real-world ip and network tracking measurement study of malicious websites with hazop. *International Journal of Computers and Applications*, Vol. 39, No. 2, pp. 106–121, 2017.

[24] Masood Mansoori, et al. Geolocation tracking and cloaking of malicious web sites. In *2019 IEEE 44th Conference on Local Computer Networks*, pp. 274–281, 2019.

[25] Pai Kasturi, et al. Tardis: Rolling back the clock on cms-targeting cyber attacks. In *2020 IEEE Symposium on Security and Privacy*, pp. 1156–1171, 2020.

[26] Timothy Barron, et al. Picky attackers: Quantifying the role of system properties on intruder behavior. In *33rd Annual Computer Security Applications Conference*, pp. 387–398, 2017.

[27] Paul Pearce, et al. Augur: Internet-wide detection of connectivity disruptions. In *2017 IEEE Symposium on Security and Privacy*, pp. 427–443, 2017.

[28] Arian Akhavan Niaki, et al. ICLab: A Global, Longitudinal Internet Censorship Measurement Platform. In *2020 IEEE Symposium on Security and Privacy*, pp. 214–230, 2020.

[29] Ram Sundara Raman, et al. Censored planet: An internet-wide, longitudinal censorship observatory. In *2020 ACM SIGSAC Conference on Computer and Communications Security*, pp. 49–66, 2020.

Copyright © 2022 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.