第 10 巻、第 3 号

削ぎ落とされる収益

Web スクレイパーが E コマースに与える影響

Akamai

インターネットの現状／セキュリティ

目次

Web トラフィック全体の約半分が、ボットによるものだとご存じですか？特に、利益を生む Web アプリケーションやアセットに依存しているコマース業界は、リスクの高いボットトラフィックの影響を最も受けています（図 1）。また、ボットは進化しているとよく言われますが、中でも Web スクレイパーボット（リンク先ブログは英語のみ）は、その経済的影響（表面化しないことが多いです）が他の種類のボットとは異なり、現在、E コマースを生業とする企業に警戒されています。スクレイパーボットの検知も、ますます困難になっています。人工知能（AI）ボットネットやヘッドレスブラウザー技術の台頭により、かなりの確率で検知を回避できるようになっているためです。たとえば、Akamai の顧客である E コマース企業の 1 社で、リスクの高いトラフィックの 99% を抑止したことがありましたが、彼らはこの原因がスクレイパーボットであると認識できていませんでした。

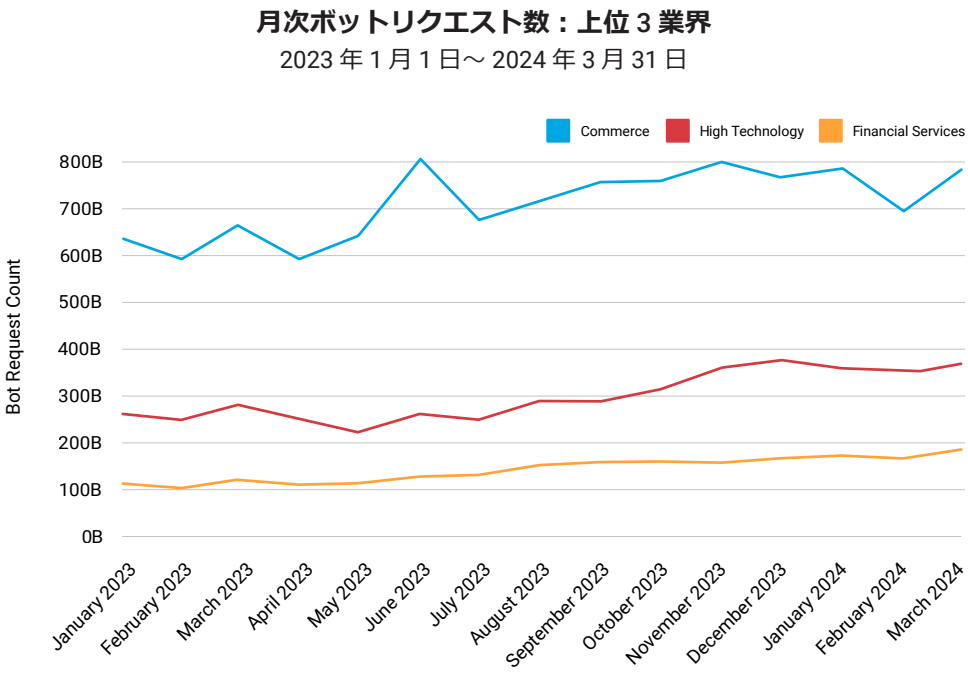

図 1：ボットリクエストが最も多いのはコマース業界。2023 年初頭から 2024 年第 1 四半期にかけて、この業界でボットトラフィックが世界的に増加している

そのため、今回のインターネットの現状（SOTI）レポートでは、この種のボットの進化と特殊性、およびその運用者に焦点を当てます。ボットは目新しいものではありませんが、依然としてさまざまなグループが犯罪的な攻撃や、詐欺行為、競合分析に利用しています。最近は、あらゆるボットの使用が増加傾向にあり、スクレイパーボットによるビジネスへの悪影響も増加しています。このレポートは、テクノロジーに関する知見と攻撃手法の両方を共有し、この拡大する問題の認識をコマース企業全体で高めることを目的としています。

ボット：良性、悪性、および迷惑なもの

目的に応じて進化し続け、ますます特殊になっているボットに、E コマースに重点を置く主要企業各社が頭を悩ませています。コマース業界内には、さまざまなタスクを実行する多種多様なボットが存在します。それらをわかりやすくするために、「良性ボット」、「悪性ボット」、「グレーボット」という 3 つのグループに分類します。「良性ボット」は、顧客に企業のサイトを見つけやすくする効果があります。「悪性ボット」は悪用を目的としたもので、企業のサイトをスクレイプします。「グレーボット」は、正規でありながらノイズとなる傾向にあるのですが、実際は良性ボットのサブカテゴリーです（たとえば、常に ping を実行するパートナーボット、その他の頻繁に API を呼び出すプログラムなど。

便利なチャットボットや検索エンジンボットについては、ユーザーからの基本的な質問に回答したり、より正確な検索結果を返す Web サイトコンテンツを提供するなど有益な効果をもたらすことから、IT コストを抑えつつ最適化して活用したいと考えられています。顧客のアカウントに不正アクセスして乗っ取ろうとする Credential Stuffing ボットなどの悪性ボットについては、顧客体験全般に影響を与えないかたちで予防措置を講じたいと考えられています。最近登場したボットの一種で、収益の減少、ロイヤルティの弱体化、コストの増加をもたらし、特に問題になっているのが、Web スクレイパーボットです。

スクレイパーボットは、インターネット上の Web サイトからデータやコンテンツを直接抽出する、特殊なボットネットです。運用方法や、ビジネスに与える影響、検知方法が他のボットと異なることから、注意する必要があります。また、Web スクレイパーを通じて収集した情報を収益化する方法が組織や運用者によって異なり、そのユースケースも多面的です。スクレイパーは、どのような目的であっても、収益を奪い、IT コストを増加させ、顧客体験全体を劣化させるものなのです。

この SOTI レポートでは、E コマースにおけるスクレイピングの影響を分析し、ビジネスオーナー（デジタル、マーケティング、ブランド、金融、リスク、セキュリティなど）が悪質なスクレイパーの阻止に共通の関心を持つべき理由を考えます。スクレイパーの影響に対する理解を深めるには、Web スクレイパーが進化する理由や、用途、動作、影響、また、コマース企業において可能な対処について、全体像を把握することが重要です。

本レポートの主要な知見

Web スクレイピングは、単なる不正行為やセキュリティ上の問題ではありません。ビジネス上の問題でもあります。スクレイパーボットは、組織の多くの面、たとえば、収益、競争力、ブランドのアイデンティティ、顧客体験、インフラコスト、デジタル体験などに悪影響をもたらします。

Akamai の調査によると、全トラフィック活動の 42.1% がボット由来で、そのボットトラフィックの 65.3% が悪性ボットから発生しています。また、悪性ボットトラフィック全体の 63.1% が高度な技術を使用しています。

スクレイパーの状況を変えたのは、ヘッドレスブラウザー技術です。この種のボットの活動を管理するためには、他の JavaScript ベースの緩和策よりも高度なアプローチが必要です。

スクレイピングが悪意を持って行われたか、有益な意図を持って行われたかにかかわらず、スクレイピングを受けた結果として組織が受ける技術的な影響には、Web サイトのパフォーマンス低下、サイトの指標の汚染、フィッシングサイトからの不正な認証情報を使用した攻撃、コンピューティングコストの増加などが挙げられます。

Web サイトで発生しているのが人間、基本的なボット、高度なボットのいずれのトラフィックなのかを判断するためには、さまざまなトラフィックパターンを観察し、把握することが重要です。パターンは、24 時間周期から、断続、連続まで、多岐にわたります。

良性ボットと悪性ボット

まずは基本から入りましょう。ボットは、「ロボット」の省略形で、自動化されたタスクを人間よりも高速かつ正確に実行できるコンピュータープログラムです。ボットにはさまざまな役割や種類があり、主に、良性ボットと悪性ボットの 2 つのカテゴリーに分類されます（図 2）。グレーボットは良性ボットのサブカテゴリーですが、ここでは、シンプルに比較できるよう良性ボットに統合します。

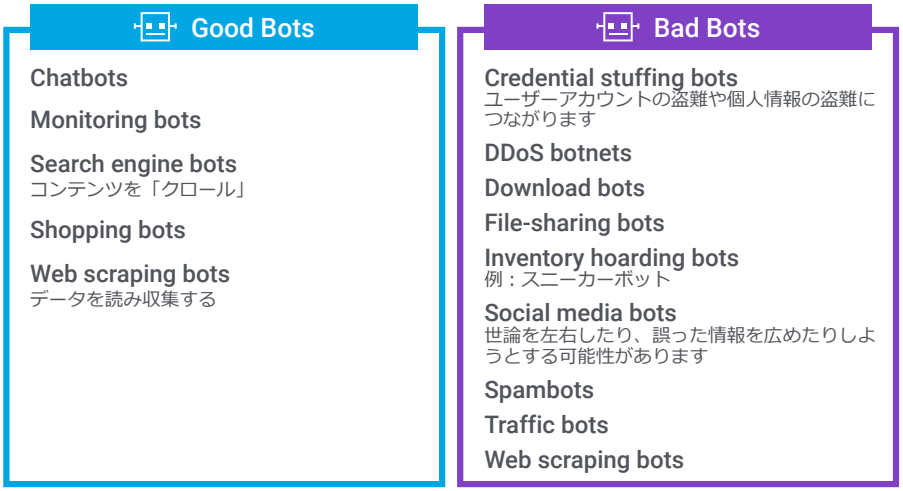

図 2：例を挙げた良性ボットと悪性ボットの対照比較

良性ボットはツールやサービスの提供に役立つ便利なボットです。それに対し、悪性ボットはサイバー犯罪者や詐欺師に悪意を持って使用されることがよくあります。この種の悪意の例として、Web サイトのクリック数やトラフィックを増やすために人間のふるまいを模倣するトラフィックボット（つまり、広告詐欺）が挙げられます。

Web スクレイパーボットは、良性ボットと悪性ボットの両方のカテゴリーに含まれています。この区別は、ボットで収集する情報を組織がどのように使用するかで決まります。ここでは、世界最大クラスの小売企業や E コマースブランドが直面している、スクレイパーボットの良い面と悪い面の両方に関するさまざまなユースケースを細かく見ていきます。

スクレイピングの基礎

Web スクレイピングは、E コマース企業で一般的に使用されています。旅行＆ホテル業界では、たとえば、旅行情報収集サイトがパートナーのホテルや航空会社からコンテンツを動的にスクレイピングし、最新の空き情報や料金を提示しています。この種のスクレイピングは予期されており、企業は、実際のユーザーが予約しようとしている時間帯には、スクレイパーを抑制しています。また、組織では、データ抽出サービスプロバイダーも使用し、競合他社からリードやその他の関連情報を収集しています。さらに、スクレイパーボットはデータの分析や動向の見極めにも使用されています。スクレイピングは、サイトのレビューでオンライン サービスや製品を改善したり、検索エンジンなどを通じて潜在的な消費者が企業の製品をより簡単に見つけられるようにしたりするのに有益です。これらすべての機能は、企業が競争力を発揮するのに役立ちます。その一方で、多くの組織が、あまり称賛できない理由でスクレイパーを使用していることも否めません。

スクレイピングの裏面 ― 消費者が注意すべき点

残念ながら、フィッシング詐欺の被害に遭った消費者の話をよく耳にします。この場合、スクレイパーボットは、製品画像、説明、価格情報を取り込んで、認証情報やクレジットカード情報を盗むことを目的とした偽造品店やフィッシングサイトを作成するために使用されます。このようなフィッシング／偽造サイトはブランドのなりすましの一種で、被害者組織の知的財産が潜在的な顧客との信頼関係を築くために盗用されます。

世界最大規模の E コマースブランドの一部が、ブランドのなりすましキャンペーンの一環として、偽造サイトや、フィッシングキャンペーン、企業 Web データの盗難の影響を受けています（図 3）。不本意ですが、フィッシングサイトが成功すると、正規のブランドがその影響を被り、顧客の信頼やロイヤルティを失うことになります。

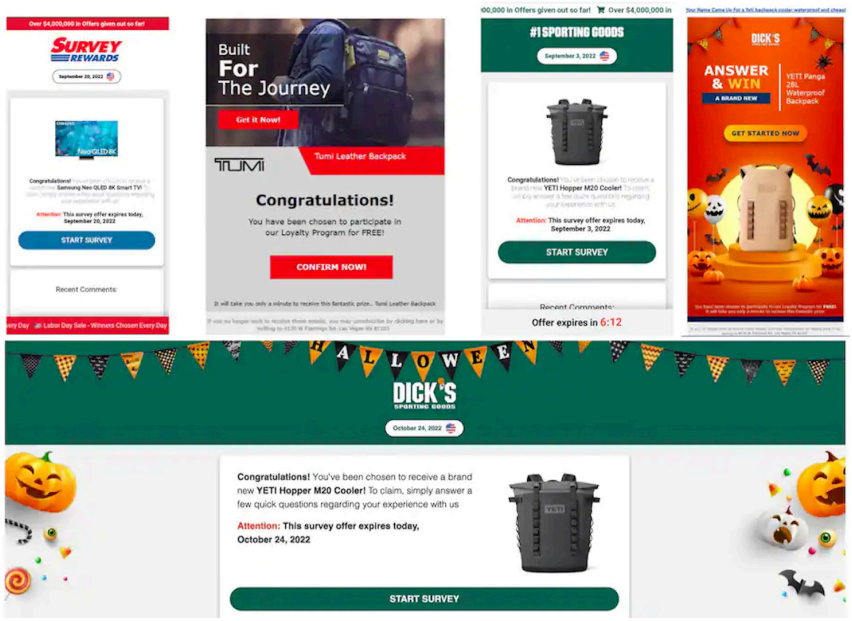

図 3：ブランドのなりすましの被害に遭った大手 E コマース企業の例

転売も Web スクレイピングに起因する場合があります。転売業者が、正当な消費者が購入する前に、商品のサイトをスクレイピングし、買い占めてしまうことがあるためです（図 4）。

スクレイパーのユースケース

コンテンツをスクレイピングすることで、収益を得られるケース

競合

競合他社は、サイトの情報を利用して、同じ商品の価格を下げたり、オファーを変更したり、新しい機会や脅威を把握したりする。

ダフ屋

ダフ屋は、商品が入手可能になるのを待ってサイトに絶えずアクセスし、カートに追加して、顧客がそれらの商品を入手できないようにしている。

偽造者

偽造業者は正規サイトのコンテンツを利用して偽のサイトや製品カタログを作成し、ユーザーを騙して偽造品ではなく正規の商品を購入していると信じ込ませる。

図 4：スクレイパーのユースケース

このような種類の悪質なスクレイピング活動を行う攻撃者は、その悪性行為が被害者に与える影響を認識しています。たとえば、競合他社の情報収集／スパイ行為や、在庫買い占め／転売、サイト／グッズの偽造や詐欺、メディアサイトのスクレイピングや再投稿による悪影響などです（表 1）。現在のところ、スクレイパーボットの使用を明示的に禁止する法律はありません。

影響	説明
競合他社の情報収集／スパイ行為	競合他社が、他社組織のサイトから得た情報を使用して価格を下げ、サービスを変更し、新たなチャンスや脅威を察知します。
在庫買い占め／転売	転売業者は常に標的サイトをチェックし、入手可能になった商品を見つけるとカートに入れ、実際の消費者がそれらの商品を入手できないようにします。
サイト／グッズの偽造や詐欺	偽造者はスクレイプしたコンテンツを使用して、偽のサイトや商品カタログを作り、偽造品ではなく正規の商品を購入していると消費者に思い込ませます。
メディアサイトのスクレイピングと再投稿	ニュース記事やブログなどのコンテンツをスクレイプして、自分のサイトに投稿することにより、本来のサイトの訪問者数を減少させたり、見込まれる広告収入に損失をもたらしたりします。 広告料はサイトの訪問者数／視聴回数で決まることが多いため、訪問者数が減少すると、より多くの広告料を得られるはずだったメディアサイトが収益を失うことになります。

表 1：Web スクレイパーによる意図的な悪影響

Web スクレイピングの一般的な副作用

Web スクレイピングの目的に関わらず、組織はその副作用がもたらす出費に対処する必要があります。有益な有料のスクレイピングサービスを利用している企業もありますが、スクレイピングされる企業も自らコストを負担しています。そのコストには、ボット対策ソリューションの費用や、サイトパフォーマンスの低下や主要指標の評価の悪化による経済的な悪影響が含まれます（表 2）。

影響	説明
ボットトラフィックへの対応のため、サーバー、CDN、クラウドのコストが増加	これは収益に影響を与えます。また、競合他社、攻撃者、偽造者によるコンテンツの使用が、評判の低下につながります。
サイトパフォーマンスの低下	スクレイパーボットは止まるまで稼働し続けるため、サーバーや配信のコストを増大させます。望まないボットトラフィックの費用を負担することになり、しかもサイトやアプリのパフォーマンスが低下するなど、ユーザー体験の悪化に悩まされます。
主要指標への悪影響	ボット活動を検知できないと、サイトのコンバージョン率などの主要指標に深刻な影響が生じます。経営陣はこれらの主要指標に基づいて、製品のポジショニング戦略やマーケティングキャンペーンなどの投資判断を行います。

表 2：Web スクレイパーによる意図しない悪影響

有料のスクレイピング：サードパーティの Web スクレイピングサービス

前述のとおり、Web スクレイパーボットには良性と悪性があります。Credential Stuffing 攻撃に使用するボットは、悪性ボットとして知られ、当然のことながらブロックされますが、このようなボットとは異なる正当な Web スクレイピングボットを提供する企業もあります。多くの組織がこのようなサードパーティの Web スクレイピングサービスを使用し、データを抽出して、組織内に提供しています。これは、競争の激しいマーケティング業界では特に有益です。

さまざまな種類の Web スクレイピング／データ抽出サービスを提供している企業は多数あり、サービスを宣伝するカンファレンスも開催されています。たとえば、Bright Data は ScrapeCon というカンファレンスを開催しています。ここでは、ボット検知の回避を得意とするエキスパートを集め、企業がデータのスクレイピング方法を学べるようにしています。表 3 は、サードパーティの Web スクレイピング企業が提供しているサービスレベルの例です。

サービスレベル 1	プロキシサービスはスクレイピングの一部です。データセンターのモバイル IP や家庭のアドレスなどのインフラを提供します。
サービスレベル 2	このレベル 2 には、自動化されたデータ抽出が含まれることもあります。データ抽出では、貴重なインテリジェンスを抽出してビジネス上の意思決定を導く顧客データ・サイエンス・チームのメンバーが簡単に使用できるように、データを整備して構造化します。
サービスレベル 3	最高レベルのサービスには、実際のビジネスインテリジェンス自体の抽出が追加されることがあります。これにより、ビジネス上の意思決定プロセスがさらに強化されます。「AI ボットネット」とも呼ばれます。

表 3：サードパーティの Web スクレイピング
企業が提供しているさまざまなサービスレベル

利用者は、最も基本的なレベルから高度なレベルまで、サービスレベルを選択できるほか、データ収集の頻度の選択や、ターゲットの指定を行えます。多くの場合、提供されるサービスのレベルや、選択するボットネットは、克服すべきサイト側の保護レベルによって異なります。より基本的なボットネットは、トラフィック負荷の分散を目的にデータセンターに設置された数千台のプロキシサーバーを使用し、高度なスクリプトを介してデータを収集できます。保護が緩いと、ボットネットがこの手法で、ボット管理による防御とセキュリティインフラの Web アプリケーションファイアウォールをすり抜ける可能性があります。

その一方、保護が高度な場合は、ヘッドレスブラウザー攻撃などのより巧妙なスクレイピングアプローチが必要になります。これは、スクレイピングの意図の善悪に関わらずあてはまります。また、高度なインフラほど、一般に基本的なレベルのサービスよりも高価なため、企業が負うコストは安価でありません。高度な防御には、チャレンジテクノロジー（CAPTCHA、プルーフ・オブ・ワークなど）、クライアント側のフィンガープリント認証のために設計された複数の検知レイヤ ー、Hypertext Transfer Protocol（HTTP） や Transport Layer Security（TLS）の特性の分析が含まれます。

AI ボットネットのスクレイピングプロセス

スクレイピング手法は、基本的な Web スクレイパーの方がより一貫していることがあります。それに対し、AI ボットネットは、フォーマットや場所に一貫性のない非構造化データやコンテンツを探索してスクレイプすることができます。さらに、実際のビジネスインテリジェンスを使用して意思決定プロセスを強化することもできます。高度な AI ボットネット（表 3 のサービスレベル 3）は、3 段階のプロセスでデータをスクレイピングします。データを収集し、抽出して、処理します（図 5）。

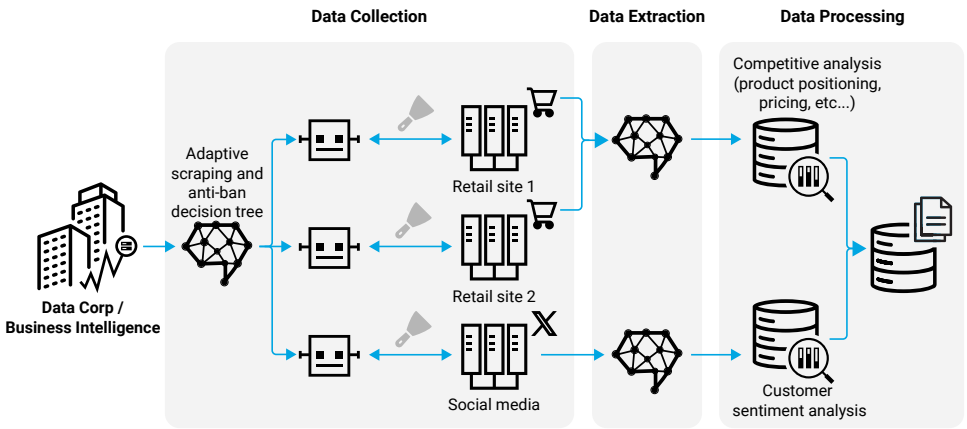

図 5：AI ボットネットとその 3 段階のプロセス

この 3 段階について理解を深められるよう、さらに詳しくご説明します。

データ収集

Web スクレイピングは、Web サイトから抽出したデータを整理し、組織が必要に応じて適用および分析できる新たなデータセットを作成できるようにします。その最初に行うのがデータ収集です。

データ収集は、適応型スクレイピングと「BAN 対策」または「ボット検知対策」テクノロジーを組み合わせて構成し、迅速かつスムーズに機能できるようにする必要があります。このようなテクノロジーは、採用されている可能性のある防御策のさまざまな側面を検知する意思決定ツリーとして設定されます。ここで重要なのは耐障害性です。ボット防御策には、JavaScript のフィンガープリント、HTTP および TLS のフィンガープリント（HTTP ヘッダーと TLS ハンドシェイクを評価）、インターネットプロトコル（IP）レピュテーション検知などがあります（図 6）。このようなワークフローの一部には、機械学習（ML）が含まれていることがあります。特に、成功率に関する統計を収集する場合や、cookie 戦略、HTTP ヘッダー、TLS パラメーターを調整する場合、JavaScript のフィンガープリントコードを評価する場合は、ML が含まれます。ヘッドレスブラウザーが役立つ場合もあります。

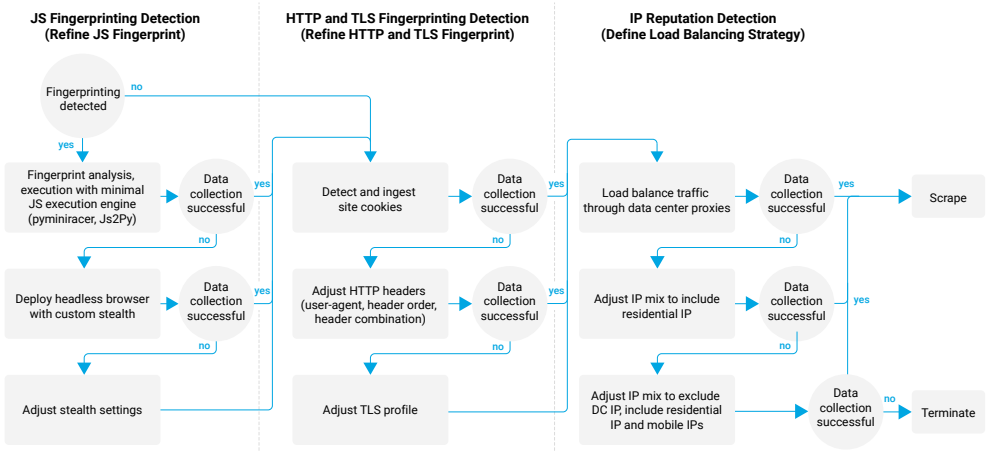

図 6：データを収集しようとすると、このボット検知対策の意思決定ツリーに沿って、JavaScript のフィンガープリント、HTTP および TLS のフィンガープリント、IP レピュテーション検知の回避を試みる

ヘッドレスブラウザー

ヘッドレスブラウザーは、グラフィカル・ユーザー・インターフェース（GUI）のない Web ブラウザーです。つまり、人間はヘッドレスブラウザーに表示される Web ページを直接操作できません。代わりにコマンドラインインターフェース（CLI）やネットワーク通信を介して操作します。たとえば、人気のあるオープンソースのヘッドレスブラウザー「Selenium」は、ブラウザーが自動化されており、Web スクレイピングに広く使用されています。これは、コンテンツを動的にスクレイピングしようとしているデータ検索者にとって非常に便利です。

ヘッドレスブラウザーでは、スクリーンショットや Web サイトのコードを効率よくコピーすることも、選択したデータをページ全体のレンダリングをせずに抽出することもできます。ただし、ヘッドレスブラウザー攻撃は実行コストが高く、攻撃で残したフィンガープリントから検知されることも依然としてあります。ただし、他の高度なインフラにかかる費用も、ヘッドレスブラウザーにかかる費用と同じく、一般的に高額になります。

データ抽出とデータ処理

抽出した情報は、通常、HTML コンテンツや JSON コンテンツで構成されます。抽出したすべてのデータのうち、分析に役立つのはほんの一部です。たとえば、競合分析で使用されるのは、一般に、価格、割引、在庫、商品の SKU 番号、カテゴリー、説明文です。分析に不可欠な情報は、それを認識するためにさまざまな構造やデータ形式を使用してトレーニングされた ML モデルによって、自動的に抽出されます。これにより、手動のデータ抽出では欠かせなかった余分な処理をすべてなくすことができるほか、HTML コンテンツや JSON コンテンツのコード構造を調査する必要もなくなります。さらに、サイトのデザインの進化につれて、コンテンツコードの構造が変わる可能性もあります。分析範囲に複数の Web サイトが関わる場合は、処理するために ML ロジックを追加する必要もあります。

ケーススタディ：Web スクレイピング検知ソリューションのメリット

Akamai の研究者が、スクレイピング活動を検知する Web スクレイピングソリューションで保護された E コマース利用者を観察し、1 週間のトラフィック活動を調査しました。このサンプルサイズは、リクエストおよそ 69 億件分に達しました。分析で考慮されたのは、HTML リクエストと AJAX リクエストのみです。静的コンテンツ（画像、JavaScript、スタイルシート）は、ほとんどのボットがリクエストしないため、分析から除外されました。これにより、データの不必要な増加を防ぐことができました。

活動全体を Akamai Content Protector が分類した結果、低リスクの人間のトラフィックが 49.3%、ボットトラフィックが 42.1%（高リスクの悪性ボットが 27.5%、良性ボットが 14.6%）、分類できない中リスクのトラフィックが 8.7% でした（図 7）。

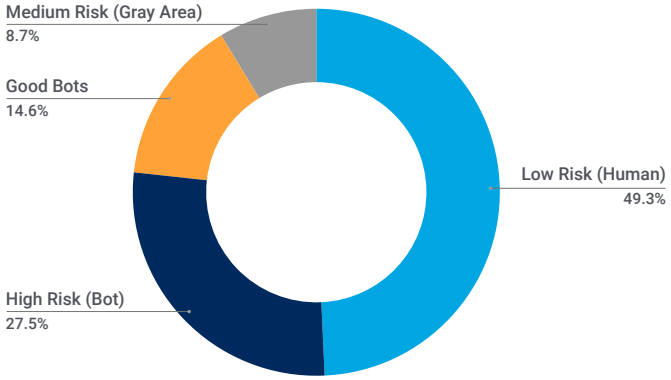

図 7：トラフィック活動の内訳

図 8 は、42.1% のボットトラフィックのうち、65.3% が悪性ボットと思われるスクレイパーから発生したもので、残りの 34.7% が良性ボットに分類されるスクレイパー（Web 検索エンジン、SEO、ソーシャルメディア、オンライン広告など）から発生したものであることを示しています。

図 8：良性ボットトラフィックと悪性ボットトラフィックの割合

ボットトラフィック全体の 65.3% を占める高リスクの悪性ボットの巧妙さレベルも評価しました。その結果、37% が基本的なスクリプトを使用したボットネット（シンプルなステートレス方式を通じて簡単に検知可能）のトラフィック、47.6% が高度なスクリプトを使用したボットネット（ML を使用する高度なステートフル検知方式が必要）のトラフィック、15.5% がヘッドレスブラウザー（高度な JavaScript フィンガープリントとステートフル検知方式が必要）のトラフィックでした（図 9）。

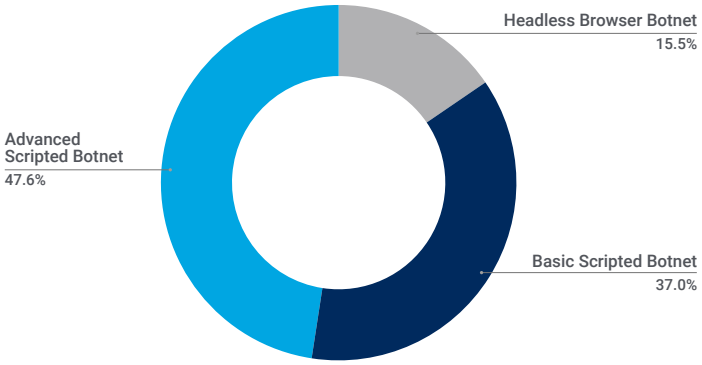

図 9：巧妙さレベルに基づく悪性ボットの分布
（合計が 100% にならないのは、数値を四捨五入したため）

このデータから、悪性ボットスクレイパーの数が良性ボットスクレイパーを大幅に上回ること、トラフィック全体のおよそ半分がボットで構成されていること、悪性ボットトラフィックの大半が高度なスクリプトを使用したボットネットであること（47.6%）がわかります。

このようなボットに対して防御策が講じられ、スクレイパーが排除されると、Web サイトはより高速かつ効率的に機能するようになります。その結果、ユーザー／顧客体験が向上します。図 10 に示すとおり、緩和策を講じたところ、高リスクボットからのリクエスト件数が大幅に減少しました。

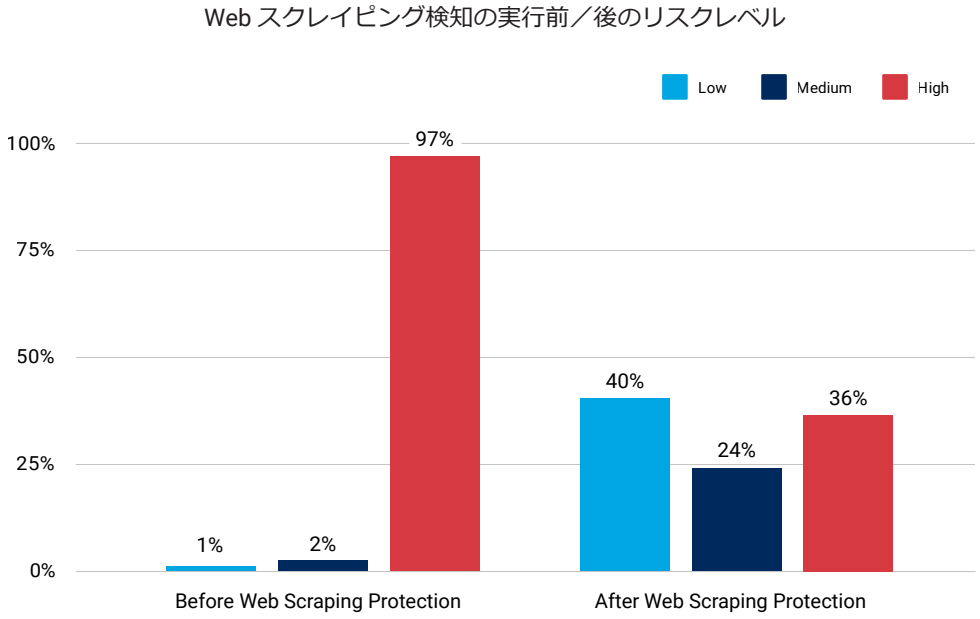

図 10：Content Protector で緩和策を実行する前／実行した後のリスクレベル

保護と緩和

このセクションでは、Web スクレイパーの検知における重要な指標と、防御策になるツールに関する情報をご紹介します。

基本的なスクレイパーを検知する

高度なスクレイパーの検知は困難かもしれませんが、ボット管理ソリューションは、あらゆる種類の侵入型スクレイパーによるデータ収集を防ぐことができます。また、特に次の特徴に注目し、シンプルな Web スクレイパーボットを検知することもできます。

- 古いブラウザーや OS バージョンに偽装するリクエスト
- HTTP ヘッダーシグネチャに潜む異常
- より一般的な HTTP v2 や新しい HTTP v3 ではなく、古いバージョンの HTTP（v1.1 など）を使用
- 数千ものクラウドサービス／データセンターからのリクエスト

より高度なスクレイパーを検知する

より高度なスクレイパーには、前述の特徴がどれも該当しません。代わりに、次のような特徴が見られます。

- 最新のブラウザーや OS バージョンからのリクエスト
- 正規のブラウザーと同一に見える HTTP ヘッダー設定
- HTTP v2 を使用
- 数十万ものレジデンシャル IP アドレスやモバイル IP アドレスからのリクエスト

トラフィックパターンを特定する

Web サイトで発生しているトラフィックが、人間（図 11）、基本的なボット（図 12）、高度なボット（図 13）のいずれによるものかを特定できる重要な指標があります。

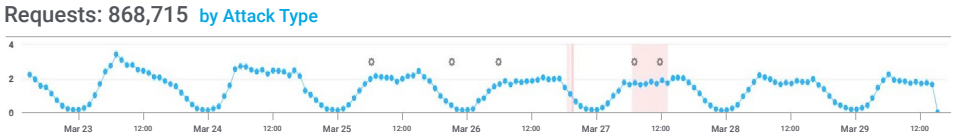

図 11：正規ユーザーのトラフィックは、通常 24 時間周期の活動

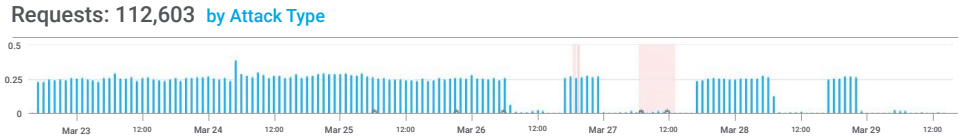

図 12：一般的なボットのトラフィックは、一時的な中断はあるものの、定期的な活動

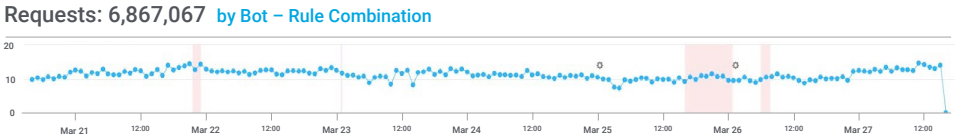

図 13：高度なボットのトラフィックは、昼夜を問わず連続

ほかにも、負荷分散戦略が弱いがフィンガープリント戦略は高度（または、その逆）という、基本的でも高度でもないボットネットもよく見かけます。一方、より高度なボットネットは非常に巧妙なため、完璧なフィンガープリントを持っていると見せかけたり、正規ユーザーのトラフィックパターンを再現したりすることもできます。

このようなスクレイパーボットに警戒するほか、Web スクレイピングに対抗するツール（たとえば Content Protector など）を使用すると、スクレイパーだらけの荒々しい環境の中でも特別なメリットが得られ、スムーズな運用ができます。たとえば次のようなメリットが得られます。

- コンバージョン率の向上と IT コストの削減
- 正確な指標が、的確な投資判断を可能にし、収益が増加する
- 価格に関する圧力が減少し、それが、競合他社の値下げから売上を守ることにつながる
- 消費者が欲しい商品にアクセスできるようになり、お気に入りの商品を確保した消費者は他の商品もカートに入れることがあるため、アップセルによる売上増にもつながる
- 消費者が保護され、低品質の偽物を正規の販売者からの正規品と勘違いして購入することがなくなり、ブランドの評判が保たれる
- 製品収益を保ち、顧客ロイヤルティを維持する
- 広告収益の増加／保護
- 視聴者やサイト訪問者の維持

コンプライアンスの考慮事項

Payment Card Industry Data Security Standard（PCI DSS）v4.0 が施行されました。変更事項の多くは、依然として企業に影響を与えている脅威の傾向に起因しています。このような攻撃に対処する上で鍵となるのが、可視性です。従来の JavaScript 環境にある脅威でも、変革を促進するために使用する API にある脅威でも、迅速に検知し、修正することが重要です。

また、ガバナンス機能が追加された新しい NIST Cybersecurity Framework バージョン 2.0 では、コンプライアンスに関する新たな傾向が見られます。NIST は、多くの政府規制の基盤になっており、多くの商用サイバーセキュリティフレームワークに浸透しています。そのため、この機会に、新しいガイダンスを確認し、それを基にポリシーを更新するか、現在のドキュメント作成工程を明確化し、ガイダンスと一致しない部分を把握することをお勧めします。

上場企業や一般に普及している会計原則（GAAP）を採用している企業には、もう 1 つ、サイバーセキュリティの重要性というコンプライアンス領域もあります。重大なリスクや脅威を定義するためには、経営陣全体が連携する必要があります。重大な脅威（ランサムウェアなど）を特定したら、緩和策（マイクロセグメンテーションなど）を明確にする必要があります。危機管理計画が開示スケジュールに対応していることを確認し、証券取引委員会のサイバー・インシデント・フォーム 8-K を提出する必要がある最悪のシナリオを想定したプレイブックを用意しておきましょう。

結論

このレポートから、組織に経済的悪影響をもたらしかねない分野に関する知見を得られたでしょうか？ボットが Web サイトに与える影響はますます大きくなっています。そのため、有益なボットを最適化し、悪性ボットの影響を緩和して、顧客体験全体のフリクションを抑えることが重要です。これはビジネスに影響を及ぼすセキュリティ問題です。他のすべてのセキュリティ問題と同じく、まずは可視化しましょう。次に影響を分析し、最後に、リスクと収益の ROI を決定して、適切なセキュリティ制御を実現できるようにします。

把握できなければ防御もできません。この機会に、可視化が不十分な部分を特定してください。そのためには、Web サイトでの Web スクレイピング活動のレベルとその目的を特定する必要があります。ボットには良性と悪性があり、スクレイパーボットは、その用途により、両方のカテゴリーに存在します。良性と悪性の境界線はあいまいですが、ボットの高度化（ヘッドレスブラウザー攻撃を仕掛ける Web スクレイパーなど）は進んでいます。いずれの Web スクレイパーボットも、E コマース企業の IT コストと顧客体験の両方に多大な影響を与えかねません。よって、ボット活動と Web サイトへの影響を分析するツールを用意することが重要になります。

誰もが回避したいのは、犯罪につながるビジネスモデルを企業の Web サイトで実行し、ロイヤルティポイントの現金化、不正な注文、返品詐欺など、さまざまな悪事を行う攻撃者です。限定イベントのチケットを買い占めるボットや、人気商品を購入するボットも困ります。ボットは、特別セールを巧みに利用して不正な新規アカウントの開設を促すこともあり、このような場合はキャンペーンの分析やコストに影響を与えてしまいます。大規模な分散型サービス妨害（DDoS）ボットネットは、Web アプリケーションに過剰な負荷をかけることがあり、ユーザー体験の質の低下、または注文や予約ができない原因を作り、損失や顧客のフリクションにつながります。ボットは、オンラインでの人間のふるまいを模倣することもでき、Web サイトのクリック数やトラフィックを増やし、入念に作り上げたデジタル体験のマーケティング分析とパフォーマンス分析の両方をゆがめてしまいます。このようなことは、誰も望んでいないはずです。

最初に述べたとおり、世界のコマース Web トラフィックの半分以上がボットによるもので、ボットトラフィックの割合は上昇し続けています。本レポートの知見やアドバイスは、Web スクレイピングに対する防御機能を備えた Content Protector などの Akamai のセキュリティプラットフォームに基づいています。Akamai は多くの大手 E コマース企業と提携しているため、そのような企業が顧客を最大限に保護するために使用できる保護対策や緩和策を共有したいと考えました。当社の予測では、Web スクレイパーボットの使用、サービスレベルの選択肢、利用可能なボットの種類は増加します。そのため、会社のリスク対策を継続的に評価し、現在のセキュリティ制御が上層部のリスク選好と一致しているかどうかを判断することが必要になります。

最新の Akamai リサーチ情報をセキュリティ・リサーチ・ハブでご確認いただけます。

手法

Content Protector のデータ

このデータサンプルは、Akamai の Content Protector ツールが監視しているトラフィックに割り当てるリスクレベルの分類を示します。この分類は、良性ボットと悪性ボットのスクレイピング活動を検知するためと、対処しているボットが良性と悪性のどちらなのかを判断するために使用されます。ほとんどのボットは静的コンテンツをリクエストしないため、HTML リクエストと AJAX リクエストのみを考慮し、不要なデータの増加を防ぎます。

このサンプルは、2024 年 4 月 12 ～ 19 日の 1 週間のデータを使用しています。サンプルの合計サイズは、65 億件を超えるリクエストで構成されています。

ボット攻撃

このデータは、Akamai の Web アプリケーションファイアウォール（WAF）とボット管理ツールを通じて観測されたトラフィックに関するアプリケーションレイヤーのアラートです。保護されている Web サイト、アプリケーション、API へのリクエスト内にボットのペイロードを検知した場合に、ボットアラートが作動します。このボットアラートは、悪性ボットと良性ボットのいずれによっても作動されます。このアラートは、攻撃が成功したことを意味するものではありません。この製品では高度なカスタマイズが可能ですが、このレポートで提示されているデータは、保護対象のプロパティのカスタム設定を考慮せずに収集されています。データは、Akamai Connected Cloud で検知されたセキュリティイベントを分析するための内部ツールから抽出されました。Akamai Connected Cloud とは、130 か国以上、1,300 近くのネットワーク上の 4,000 か所以上に配置された約 34 万台のサーバーからなるネットワークです。このデータは、ペタバイト／月の単位で測定され、Akamai セキュリティチームによる攻撃のリサーチ、悪性のふるまいの警告、Akamai ソリューションへのインテリジェンスの追加のために使用されます。

2023 年 1 月 1 日～ 2024 年 3 月 31 日までの 15 か月間のデータを使用しています。

クレジット

編集長

Lance Rhodes

共同執筆者

David Senecal
Maria Vlasak

校閲およびテーマ別寄稿者

Mitch Mayne
Susan McReynolds
Christine Ross
Badette Tribbey
Steve Winterfeld

データ分析

Chelsea Tuttle

販促資料

Annie Brunholzl

マーケティング・出版

Georgina Morales
Emily Spinks

その他の「インターネットの現状／セキュリティ」レポート

高い評価を受けている Akamai の「インターネットの現状／セキュリティ」レポートのバックナンバーおよび今後のリリースについては、**akamai.com/soti** をご覧ください。

その他の Akamai 脅威リサーチ

akamai.com/security-research では、最新の脅威インテリジェンス分析、セキュリティレポート、サイバーセキュリティリサーチを通じ、常に最新情報を把握できます。

このレポートに掲載されているデータ

このレポートに引用されているグラフや図のハイクオリティバージョンを以下のリンクからご覧いただけます。これらの画像は、出典元として Akamai を明記し、Akamai のロゴをそのまま残すことを条件に、利用および引用が可能です：**akamai.com/sotidata**

Akamai ソリューションの詳細

Web スクレイパーの検知および保護を目的とした Akamai ソリューションの詳細については、Akamai の **Content Protector ページ**をご覧ください。

Akamai は、お客様が生み出すものすべてにセキュリティを組み込む取り組みを支援することで、どこで構築しどこへ提供しようとも、顧客体験、従業員、システム、データを守ります。グローバルな脅威に対する可視性を備えた Akamai のプラットフォームは、セキュリティ体制の適応と進化を後押しして、ゼロトラストの実現、ランサムウェアの阻止、アプリケーションと API のセキュリティの確保、DDoS 攻撃の撃退を支援します。これにより、お客様は自信を持って、継続して、イノベーションを起こし、可能性を広げ、新たな可能性を生み出すことができます。Akamai のクラウドコンピューティング、セキュリティ、コンテンツデリバリーの各ソリューションの詳細については、akamai.com および akamai.com/blog をご覧いただくか、X（旧 Twitter）と LinkedIn で Akamai Technologies をフォローしてください。
公開日：2024 年 6 月。