DAMBALLA

サイバー犯罪者の手に渡った DGA

最新のマルウェア防御回避テクニックの検証

ここ数年の間でドメイン名生成アルゴリズム（DGA: Domain Generation Algorithm）は、POC 段階の技術から大きな進化を遂げ、従来の固定的なドメインブラックリスト方式のシステムをすり抜けることができるフル機能を備えたステルスモジュールとして、激増する高度な防御回避向け商用クライムウェアツールキットで採用されるようになっています。DGA は「ドメイン・フラクシング (domain fluxing)」としても知られています。

クライムウェアのエージェントと共に送り付けられるコマンド・アンド・コントロール (C&C) ドメインに関する固定的なリストや、新しいドメイン名ファイルの更新を待って C&C サーバーを追加する形態ではなく、動的に C&C になり得るドメインを探索するよう設計された DGA ベースのシステムを採用したクライムウェア群が増加しています。

DAMBALLA Labs では、過去 12 ヶ月の間に 6 つのクライムウェア群で高度な回避テクニックが採用されていることを確認しました。このようなテクニックは、既に数十ものサイバー犯罪組織で利用されており、これらの犯罪組織の多くは、一般的なホストやネットワークベースの防御手段を回避し続けています。商用のクライムウェアツールキットに DGA モジュールが含まれるようになったことで、サイバー犯罪者は、DGA アルゴリズムのチューニングやパーソナライズを行いボットネット毎に DGA 機能を実装することで、固定的なブラックリスト方式に頼る防御システムへの攻撃力を高めています。

DAMBALLA Labs、ジョージア工科大学、ジョージア大学の共同研究によって開発された新しい検知テクニックによって、このような新しい脅威の拡大が明らかになりました。新しく確認された DGA ベースのクライムウェア群の中には、少なくても 2011 年 11 月から運用が開始され、ネットワークの防御を回避し続けてきたものもあります。この新しい検知技術を使用することで、DAMBALLA Labs は、2012 年 2 月時点でコミュニティが捕捉したクライムウェアのバイナリでは見つかっていなかった 6 つの DGA ベースのクライムウェア群を特定することができました。

DAMBALLA Labs が新たに発見した複数の DGA には 6 つ既知の（但しこれまで詳細に分析されていなかった）クライムウェア群が含まれていました。これは、Shiz、Bamital、BankPatch、Expiro.Z、Bonnana そして Zeus の亜種でした。

調査報告書の要点：

- 防御を回避する目的で DGA を使用するクライムウェア群を新たに 6 つ特定
- これまで知られていなかった、又はサンプルが発見されていなかったクライムウェアの DGA を新たに 6 つ確認
- 最新の DGA クライムウェアをサポートする C&C サーバーの大部分が東欧諸国に集中
- 最も悪用頻度の高いトップレベルドメイン (TLD) は、「.com」、「.ru 」そして「 .org」
- サイバー犯罪者は、DGA の候補ドメインとなる直前の 1 時間以内に C&C ドメインを登録して 24 時間以内には廃棄

DGA とは?

ボットネットやリモート制御が可能なクライムウェアは、新たな指示を受信して盗み出したデータを素早く転送するためにリモートサーバーを配置して接続できる必要があります。従来のマルウェアは、被害者のコンピュータへのインストールが成功すると、ハードコードされたリストに記載された C&C サーバー用のドメイン名や IP アドレスを使用して接続を試みていました。ほとんどの場合マルウェアは「有効な」C&C サーバーが見つかるまでリストを繰り返しチェックします。このため C&C サーバーになり得る候補が多いほど、削除処理やインターネットフィルタリング技術に対する耐性が増すことになります。

マルウェアがその内部に C&C の候補リストを持っている場合、セキュリティベンダーやアナリストによって、いずれ不審なバイナリが発見・分析されるため、サイバー犯罪者にとってはこれが障害となります。このような弱点を克服するため、最新のマルウェア（つまりクライムウェア）の多くが、ハードコードされたリストの使用する方法ではなく、C&C 候補情報を定期的に更新する方法へ転換を図っています。

クライムウェアは、その耐久性を増していますが、依然として情報の更新に依存しているため、分析にかけられたり C&C が削除されたりするという弱点が残ったままです。サイバー犯罪者は、このような事態に対抗するため、日付や時間、そしてシード値（乱数発生の初期設定値）を指定することで、複数の候補ドメイン名を生成してテストを行い、C&C サーバーとして機能するかどうか決定するアルゴリズムを開発しました。初期に出現し最も分析が進んだ DGA ベースのクライムウェア群の 1 つに「Conficker（Downup、Downadup または Kido とも呼ばれる）」があります。

2008 年に発見された Conficker は、疑似乱数方式の DGA を採用し、5 つの利用可能なトップレベルドメイン (TLD) から 250 の候補ドメインを選定します。2009 年初に発見された、4 番目の亜種である Conficker.D は、110 の TLD をまたいで 5 万のドメイン名を生成し、500 の候補ドメインをランダムに選定してテストします。それ以降、様々な DGA ベースのクライムウェアによって使用されるアルゴリズムは多様性を高め、隠れた C&C サーバーをより高い効率で配置できるようになっています。

ドメイン生成アルゴリズムの目的には、以下のような内容があります：

- 固定的な C&C ドメインリストを持つレピュテーションシステムが正確性を維持できないようにする
- サイバー犯罪者が境界防御ベースのネットワークフィルタリング技術をすり抜けられるようにする
- 短期間だけ設定し稼動するだけで良い小規模で小回りの利く物理 C&C インフラストラクチャーを確保する
- 対抗策や法の網にかからないよう「ぎりぎりのタイミング」でドメイン名を登録する
- C&C インフラストラクチャーが暴かれずにクライムウェアエージェントを増殖させ大きな範囲に影響を与える

DGA の感染方法を理解する

DAMBALLA Labs の研究員は、ジョージア工科大学、ジョージア大学と共に DGA ベースのクライムウェアを検知する多数のテクノロジーを発見し、サイバー犯罪者の C&C ホスティングに関係するインフラストラクチャーを特定することに成功しました。さらに、これらの技術を駆使しながら新しい範疇に属する脅威がどのように拡大しているかを確認しました。

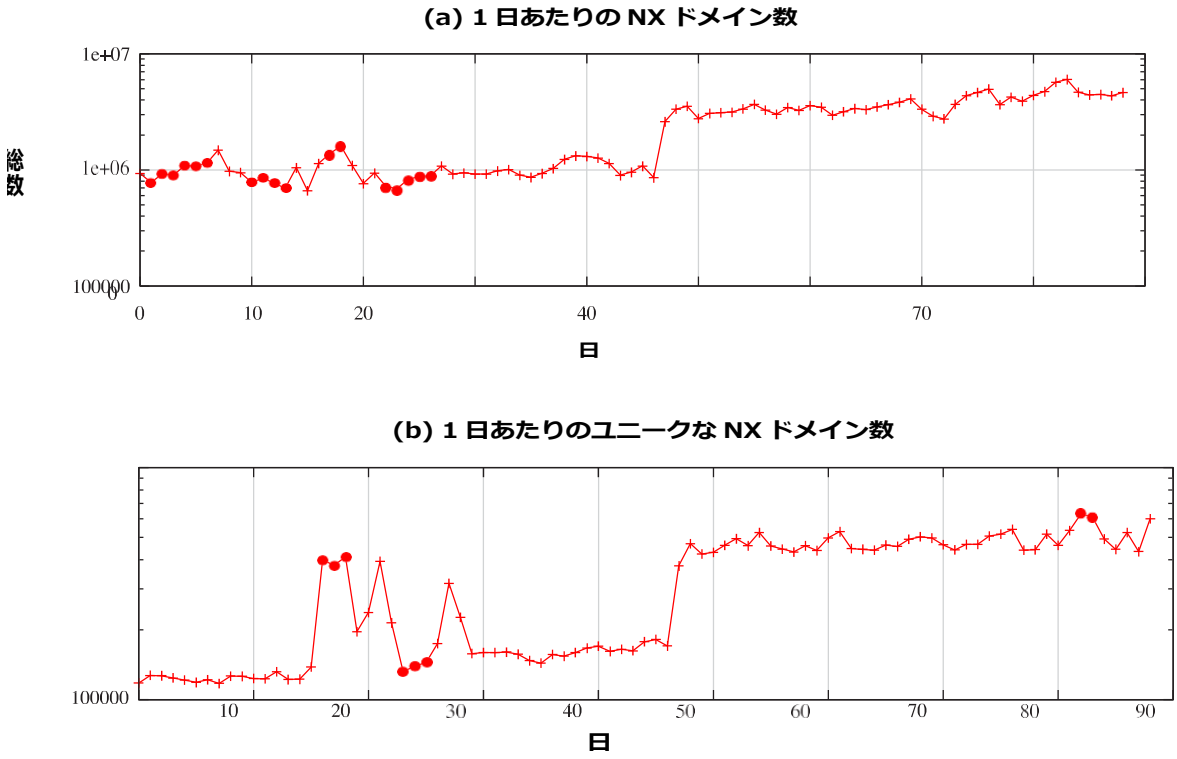
DAMBALLA Labs では、グローバル規模で DNS トラフィックを確認できる DAMBALLA の機能を最大限に活用し、DGA ベースのクライムウェアにおける感染の主要な特徴を特定しました。DGA を使って有効な C&C サーバーを探し出すクライムウェアを特定するための重要な鍵は、その失敗結果にありました。特に、毎日のように発生する、DNS を使い存在しないドメイン名の IP アドレス解決を試みた失敗結果がこれに該当します。存在しないドメイン名（以下本書では「NX ドメイン」とします）のレスポンス内容こそが DGA を使用していることを決定付ける証になります。

DGA の検知と大規模ネットワーク上の NX ドメイン

DNS トラフィック内には、ほぼ毎日非常に高い確率で NX ドメインのレスポンスが含まれていますが、そのほとんどは、故意ではなく良性のものとなっています。存在しないドメインへのアクセスは、タイプミスやドメイン名のプリフェッチ、古いドメインレコードの参照、ネットワーク設定のミスなど日常的に発生する現象と言えます。

下記の図は、ある大規模なインターネットサービスプロバイダー (ISP) の特定地域（約 250 万のサブスクライバを含む）における NX ドメイントラフィックを 3 ヶ月にわたって観察した結果を示しています。グラフ (a) は、重複を排除した日々の NX ドメインの総数（つまりレスポンスの全体数）であり、グラフ (b) は、NX ドメインからの日々のユニークなレスポンス数を示しています。

図 1：
大規模 ISP における 3 ヶ月間の NX ドメイントラフィック



相当量の不規則な NX ドメイントラフィックによって、シグニチャまたはルールベースの検知システムが過去に学習して隔離した DGA クライムウェアの通信を特定しようとして失敗に終わり、非常に高い率で誤判定のアラートを出していることが裏付けられています。

DAMBALLA では、脅威の本質を探るため、最初に 250 万のサブスクライバが存在するネットワークセグメントで、Murofet（以前から存在して広く知られ、大規模な研究が行なわれた初期の DGA を使用したクライムウェア群）に感染した数千のデバイスに注目して調査を行いました。DAMBALLA Labs は、Murofet の被害者の 90%が 1 日に 10 以上の NX ドメイン名を発生させる一方で、当該ネットワーク上にある全デバイスの 92%において、1 日に発生させる NX ドメイン名が 10 未満であることを突き止めました。

本質的に DGA では、統計的な判定やドメインのブラックリストを使用した検知システムを回避するため、非常に多くのドメイン名を疑似乱数的に発生させる必要があります。また、クライムウェアの配置と接続を試みるために利用可能なドメイン名がいくつあったとしても、サイバー犯罪者が実際にコントロールできる数には限界があります。

DAMBALLA Labs は、教師なし/教師あり機械学習の両方の技術を駆使し、2 つの主要な特性である NX ドメイン名の構造特性と一連の NX ドメイン名を生成しているデバイスの「群れ」の関連性に基づいて対象となる DGA を動的にトラッキングしました。さらに DAMBALLA Labs は、実績ある技術を使って、クライムウェアのサンプル取得や DGA が利用している特定のアルゴリズムを理解する必要なしに、自動的に発見された新たな DGA クラスタを説明するための新しい統計モデルの学習を行いました。

DGA の使用とその影響

メインあるいは予備的な検知回避戦略として DGA ベースのテクノロジーを採用するクライムウェア群が増え続けています。これまで十分な調査が行なわれてきた Conficker、Murofet、Bobax あるいは Sinowal については、メインのネットワーク検知回避テクノロジーとして、DGA を採用していることが分かっています。一方で、Zeus の亜種は、企業のネットワーク防御テクノロジーによって C&C サーバーの固定リストを使った配置や接続の試みが阻止された場合の予備的手段として DGA を使用しています。

DAMBALLA Labs が新たに発見した複数の DGA には、6 つ既知の（但しこれまで詳細に分析されていなかった）クライムウェア群が含まれていました。これは、Shiz、Bamital、BankPatch、Expiro.Z、Bonnana そして Zeus の亜種です。

DGA による回避テクノロジーを採用しているクライムウェアは、大規模な「プロの」サイバー犯罪組織に利用される傾向にあり、一般的に現在インターネットユーザーが遭遇する平均的なマルウェアよりも高度なものとなっています。しかし、その使用範囲が限定されていたり、規模が限定的であったりするものではありません。

2011 年における DGA ベースのクライムウェア トップ 5

最も流行した DGA ベースクライムウェア群トップ 5	
1	Conficker（全バージョン）
2	Murofet
3	BankPatch
4	Bonnana
5	Bobax

さらに DAMBALLA Labs は、Zeus の亜種（最も流行している DGA ベースのクライムウェア群 トップ 5 には無い）が、メインの接続テクニックが失敗した場合の予備的なバックアップ手段として、DGA ベースの回避テクニックを使用していることを突き止めました。さらに、数百のデバイスが 2012 年 2 月の前半までに感染し増加を続けていることが分かりました。感染率は、様々な感染活動に応じて異なります。新しいボットネットは、DAMBALLA Labs のもう 1 つの DGA ケーススタディの対象となりました。

C&C インフラストラクチャー

サイバー犯罪者は、クライムウェアに DGA 機能を追加することで、特定の日時まで（つまりクライムウェアにデバイスを感染させ、C&C から接続して盗み出したデータをアップロードし、新しいバッチコマンドを受け取るまで）C&C インフラストラクチャーをセキュリティ調査や法の適用から確実に隠蔽できるようにします。これらが終了すると、サイバー犯罪者が C&C をインターネットから解放してしまうため、サーバーにアクセスして法的措置を取ろうとしても不可能な状態になります。

新たに発見された DGA ベースのクライムウェアを経験した増加し続ける被害者を通じて、DAMBALLA Labs は、6 つのボットネットの C&C インフラストラクチャーを追跡調査することができました。DGA システムのドメイン名登録は短時間の内に行われますが、DAMBALLA Labs は、IP をホスティングするインフラストラクチャーを特定することができました。

以下の表は、6 つの DGA を使用するボットネットで使用されている C&C サーバーをホストしている上位の 5 ヶ国を示しています。

DGA ベースのクライムウェア C&C をホストしている上位 5 ヶ国	
UA – ウクライナ	22 サーバー
RO – ルーマニア	17 サーバー
RU – ロシア	16 サーバー
HU – ハンガリー	14 サーバー
TR – トルコ	12 サーバー

最も不正利用の対象となっている TLD

最新のクライムウェアで利用されている疑似乱数ドメイン名発生アルゴリズムは、特別高度である必要はなく、単に「十分であって」予測不能なものであれば良いのです。生成されたドメイン名はランダムなものですが、既存の DNS 構造とは互換性があり、正式な候補（例えば .com、.net、.co.uk など）リストから TDL を割り当てれば良いのです。

サイバー犯罪者が C&C として利用するためには、これらの TDL は正式なもので同時にドメイン名自体も正しく登録される必要があります。このため DGA が使用する TLD は、サイバー犯罪者が容易に入手できるものになる傾向があります。

DAMBALLA Labs が、新しいクライムウェアの亜種に関連する NX ドメインのレスポンスを観察したところ、悪用された TLD のトップ 5 は以下のようになりました。

悪用されている TLD トップ 5	
.com	3,2974,31
.ru	460,567
.net	38,382
.biz	37,783
.org	36,573

短期間でのドメイン登録

DAMBALLA Labs では、DGA を利用している一部のサイバー犯罪者が、DGA がドメイン名を登録する直前の 1 時間以内に C&C ドメインの登録と DNS の設定を実施する傾向にあることを突き止めました。彼らは DGA がドメイン名を生成してから 24 時間以内にドメインを閉鎖して別の C&C サーバーに乗り換えているのです。

まとめ

DGA は、ブラックリストやシグニチャフィルター、そして統計レピュテーションシステムなどから確実に逃れる手段として犯罪者に採用され、その数も増加し続けています。またその悪用ケースは、ステルス性の攻撃で顕著に見られます。犯罪者が DGA ベースのクライムウェアネットワークを維持するためには、より多くの「作業」が必要になりますが、完全に犯罪に気付かれず C&C インフラストラクチャーが残存するチャンスはいっそう広がります。

Zeus のソースコードが暴かれたことで、DGA を接続のためのメイン手段あるいは補助的な手段として利用する Zeus のような、より多くの派生形の出現に対応することが必要になりました。DAMBALLA Labs が明らかにした DGA ベースの脅威は、DAMBALLA Labs の DGA ケーススタディで解説されています。DAMBALLA Labs は、自分達が発見した DGA ベースの脅威のほとんどが最終的には DGA に関連するマルウェアであり、セキュリティコミュニティでは効果的に分析できなかった（あるいは部分的にしか分析できなかった）ものであることを突き止めました。DGA は、メインとなる接続手段（例えばリストの使用やピア・ツー・ピア）が失敗した場合、または誤った C&C インフラストラクチャーにドメイン接続してしまった場合の補助的な接続手段として使用されるため、DGA 機能を全て見落としているケースがほとんどです。

興味深いことに本報告書で掲載した様々な亜種は、類似した情報を盗み出すよう設計されています。ドメイン名は短命であるため、マルウェア分析、特にパッシブ DNS データやその分析機能が無い状態ではその特定が困難です。このような状況により、犯罪主体を特定することがより困難になっています。さらに攻撃側が攻撃対象をセグメント化し、タイムゾーンの違いによってドメイン名を変えているような場合には、犯罪主体の特定がより困難なものになります。

本書に記載した様々な要因により、防御する側の企業顧客はもちろんのこと、セキュリティコミュニティにとっても DGA を使用した攻撃の検知はいっそう困難なものになっています。DAMBALLA Labs では、このような傾向を長年にわたり観察し続けることで、セキュリティコミュニティがマルウェアを発見・分析する以前の初期の段階で脅威を特定するテクノロジー（特許出願中）を開発しました。この発明によって、お客様に対し DGA ベースのマルウェアに感染したデバイスを迅速に（かつマルウェアや影響を調査することなく）検知し特定する自動検知機能や、マルウェア群を自動的に分類し脅威の特性情報を提供できるようになりました。

DAMBALLA について

高度な脅威に対する防御と抑制に関するエキスパートである DAMBALLA は、全てのセキュリティ防御層をすり抜けるアクティブな脅威を検知します。また、悪意あるネットワークトラフィックの証跡をリアルタイムに検知し、業務に対する大きなリスクとなる感染デバイスを迅速に特定します。
特許取得済みのソリューションでは、業界最大の顧客やエンタープライズのネットワークトラフィックデータから構成されるビッグデータを機械学習機能と併せて使用することで、犯罪活動を自動的に検知・抑制し、データ漏洩を防ぎ、業務妨害を最小限に抑え、対応と対策に必要な時間を短縮することができます。また、PC、Mac、iOS、Android、組み込みシステムなど、デバイスや OS を問わず防御を行うことができます。DAMBALLA は、主な業界の企業や世界最大規模の ISP、テレコムプロバイダーなどが所有する 4 億 4 千万台以上のエンドポイントを日々防御しています。

DAMBALLA Failsafe の詳細については、DAMBALLA の Web サイト www.DAMBALLA.com をご覧いただくか、japan@DAMBALLA.com までご連絡ください。Twitter は、@DAMBALLAInc でフォローいただけます。

翻訳版制作にあたって：
本書は 2012 年に英語によって作成された資料の翻訳版です。脅威の主体者による DGA の活用と有用性については 2016 年現在でも依然非常に高いと考えられ、その喚起の意味も含めて本書の翻訳が重要と考え作成いたしました。掲載の各種統計情報は、2012 年調査当時の情報となりますのでご留意ください。