Website Protection Schemes Based on Behavior Analysis of Malware Attackers

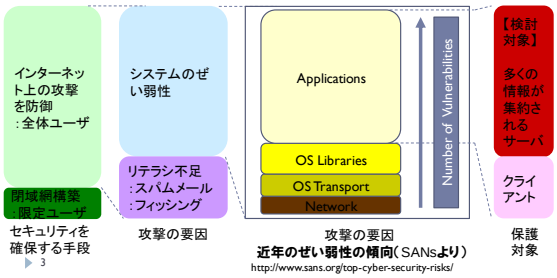
2012年12月11日

情報ネットワーク学専攻　先進ネットワークアーキテクチャ講座
八木毅

もくじ

背景

- ネットワーク技術の普及に伴いサイバー攻撃が増加
- 技術で対応する必要があり、影響が大きい「サーバ上のアプリケーションのぜい弱性に起因する攻撃」への対策を検討



研究目標

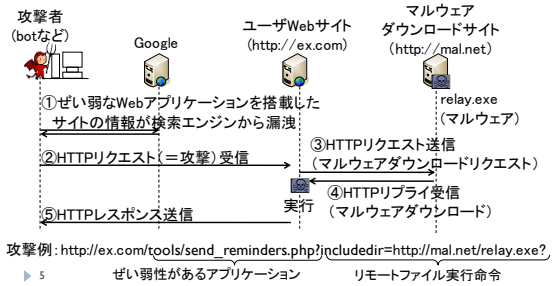
- クラウドコンピューティング/ホスティングサービスの普及
  - Webサイト運用経験のないユーザがフリーソフトでサイトを構築
- ぜい弱なWebアプリケーションを悪用した攻撃が多発
- Webサイト不正制御の原因となるマルウェア感染が脅威
  - 不正制御されたWebサイトは新たな攻撃の送信元として悪用
- プロバイダがWebサイトのマルウェア感染を保護する必要有り
  - スキルの無いユーザの増加
  - Webサイト環境提供サービスとしてセキュリティがデフォルト化

研究目標　Webサイト群をマルウェア感染から保護するための Webサイト防衛方式の確立

マルウェア（Malware）：コンピュータウィルスに代表される、悪意あるソフトウェア

典型的な攻撃

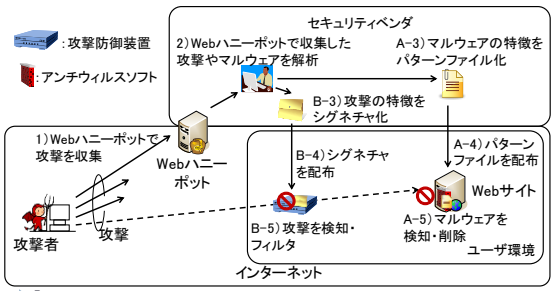
- 「指定された場所のファイルを読み込む」Webアプリケーションプログラムに「外部ファイルを読み込むことが可能」なぜい弱性がある場合、マルウェアダウンロードを強要可能



- 背景と研究目標
- 従来方式
  - 攻撃の収集と防御
  - 攻撃の収集（ハニーポット（おとりシステム））
  - 攻撃の防御（攻撃防御装置IDS/IPS/WAF）
- 目標達成のための要求条件と従来方式の課題
- 提案方式
  - マルウェア配布URLのブラックリスト化/アクセスフィルタ
  - 攻撃の宛先に応じたハニーポットでの受信制御
  - 攻撃者の行動を考慮したブラックリストURLの最適監視
- まとめ

攻撃の収集と防御

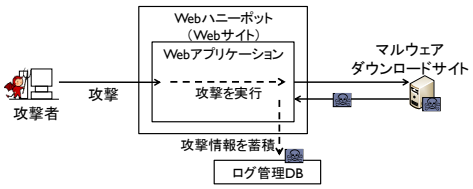
- ぜい弱なWebサイトを模擬するおとり（Webハニーポット）を用いて収集した攻撃/マルウェアの情報に基づいて感染を検知



7

攻撃の収集（ハニーポット（おとりシステム））

- おとりシステム。収集したい攻撃に応じた種類が存在。
- Webサイトへの攻撃の収集にはWebハニーポットを適用
- マルウェア感染攻撃の収集には、ぜい弱なシステムを用いて攻撃を実行して詳細な情報を収集するハニーポットが必要
- ドメインやIPアドレスを付与し、通常Webサイトと同様に運用（攻撃者に感知される可能性低）



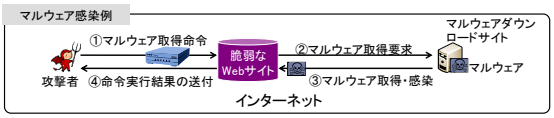
8

攻撃の防御（攻撃防御装置IDS/IPS/WAF）

- サーバ用アンチウィルスソフトでは、正常ツールとマルウェアの識別が困難なため、攻撃を防御困難

⇒攻撃の情報に基づいて感染を検知する攻撃防御装置が必須

- 攻撃防御装置では、受信アクセスのパターンがシグネチャ※と一致するか検査（下図①を検査）



※シグネチャ例
・ぜい弱なプログラムパス：vulne/login.php?dir=
・過去攻撃を実施した送信元IPアドレス
・異常なuser agent、メッセージサイズ　etc

：攻撃防御装置

9

10

目標達成のための要求条件

- ①Webサイト群を②多くの攻撃から③高精度に保護するためには、下記の要求条件が存在

①多様な攻撃の検知：4章に相当
- Webサイト群上の多数Webアプリケーションへの攻撃の検知が必須

⇒Webアプリケーション毎に異なる多様な攻撃を検知できる必要有

②収集可能な攻撃情報量の最大化：5章に相当
- 対策用情報の生成に必要な攻撃の収集は、ハニーポット配置のコストや、検索エンジンや攻撃者の動作などに依存するため、難易度高

⇒1件の攻撃から可能な限り情報を抽出する必要有

③攻撃検知の高精度化：6章に相当
- 攻撃者は攻撃検知を回避するよう攻撃手法を制御

⇒正常サービスSLA劣化やインシデント発生の元凶となる誤検知/検知漏れを抑制する必要有

11

効率性/高精度化に向けた攻撃防御装置の課題

- Webサイト用マルウェア感染を検知可能な機器は稀少
  - 独自環境で収集した2週間分のマルウェア感染攻撃をtcp replayで再現してOSS仮想IPS（suricata）と市販WAF（SecureSphere）と市販IPS（Proventia）が検知可能か調査⇒マルウェア感染攻撃を網羅的に観測して防御するための情報を生成できている機器は存在せず

⇒攻撃コードが急激に増加するため、全攻撃への対処が困難

評価1：攻撃検知率の比較

	suricata	SecureSphere	Proventia
攻撃検知率	59.4%(139/234)	46.5%(109/234)	5.9% (14/234)

攻撃コード例　　攻撃コード
http://ex.com/tools/send_reminders.php?noSet=0&includedir=http://mal.net/relay.exe

攻撃コード増加要因
・ぜい弱なアプリケーション数
・意味をなさない文字の挿入パターン

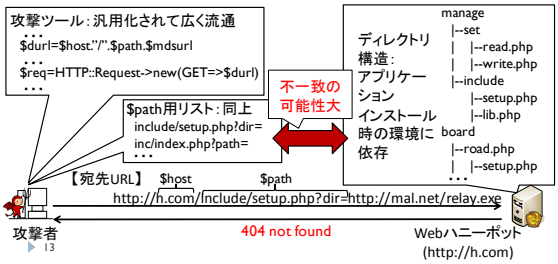
⇒ 時間に対して急増するため、全攻撃への対処が困難

12

収集情報量の最大化に向けたハニーポットの課題

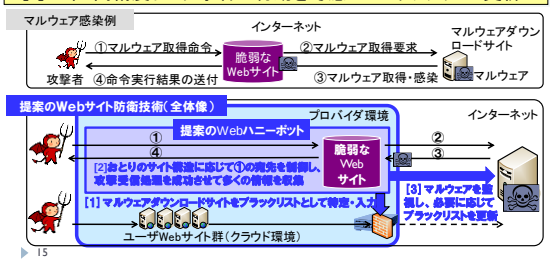
- 攻撃者は宛先パス候補を予めリスト化して攻撃ツールに設定
- 攻撃者は攻撃ツールを用いて自動的に攻撃を実施
- 攻撃の宛先URLのパス名の実在性が低く、攻撃検知が困難

⇒収集情報量が制限（攻撃収集力低下・マルウェア収集不可）



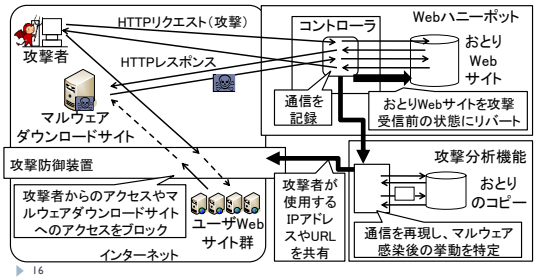
提案方式の全体像

[1]の攻撃防御方式と、[1]の効果を高める[2][3]の方式を提案
[1]：4章：多様な攻撃の防御：ブラックリストを用いた防御
[2]：5章：収集情報量の改善：ハニーポットでの攻撃宛先URL制御
[3]：6章：高精度化：攻撃者の行動を考慮したブラックリスト更新



4章：マルウェア配布用URLのブラックリスト化

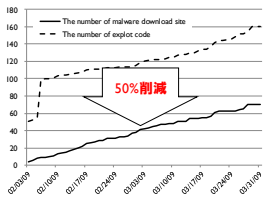
- Webハニーポットでおとり Webサイトとインターネット間の通信を収集して分析⇒マルウェア感染の原因となる宛先URLを特定
- ユーザWebサイトから上記宛先URLへのアクセスをフィルタ



4章：マルウェア配布用URLのブラックリスト化

- 3ヵ月間、実態調査を実施し、攻撃コード数とマルウェアダウンロードサイト数の増加率や絶対値を比較したところ、マルウェアダウンロードサイト数が増加率/絶対値ともに少数

⇒マルウェアダウンロードサイトへのアクセスをフィルタする提案方式が効率的



期間	2009/1/30 ～2009/4/1
攻撃コードがユニークである確率	28.85%
攻撃元IPアドレスがユニークである確率	16.74%
マルウェアダウンロードサイトがユニークである確率	11.94%

4章：アンチウイルスソフトとブラックリスト方式の比較

- 5ヵ月間収集したマルウェア検体を1ヵ月後に6種類のアンチウイルスソフトにかけた場合、4種類の検知率は50%以下
- 上記期間で収集した攻撃元IPアドレスやマルウェアダウンロードサイトの約98%は再現性有り

⇒ファイル検知よりアクセス検知であるブラックリスト方式が有効

収集したマルウェアを用いたアンチウィルスソフトの検知率調査結果

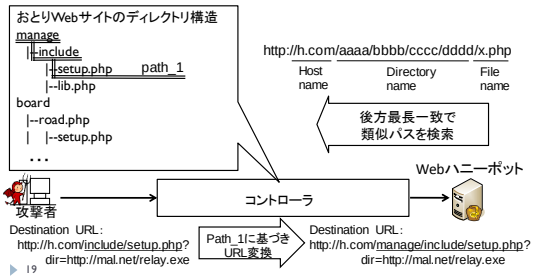
アンチウィルスソフトウェア	検知率 [%]
Software A	41
Software B	57
Software C	34
Software D	74
Software E	38
Software F	35

収集した攻撃を用いたブラックリスト評価結果

	Total	再観測率 [%]
攻撃元IPアドレス	4,621	97.9
マルウェアダウンロードサイト	2,666	98.3

5章：攻撃の宛先に応じたハニーポットでの受信制御

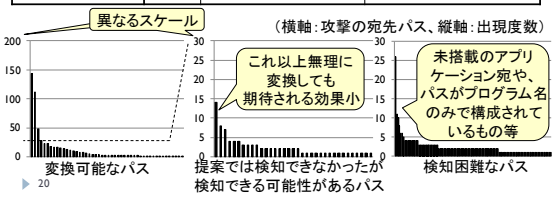
- 宛先URLのパスがおとりWebサイト上に存在しない場合、おとりWebサイトのパス群に対して宛先URLのパスの後方最長一致を検索し、ヒットしたパス宛へ宛先URLを変換



5章：攻撃の宛先に応じたハニーポットでの受信制御

- 半年間、実態調査を実施し、効果を確認（表参照）
- 宛先パスの情報が不足している攻撃等は提案方式での宛先パス変換が困難だが、数量的には限定的（グラフ参照）

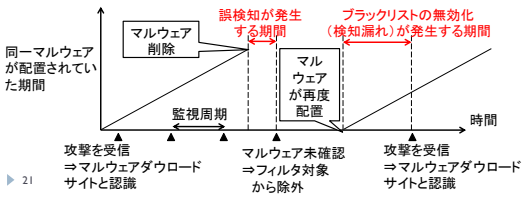
	総数	従来方式での検知数	提案方式での検知数
攻撃件数	1035	33（3.2%）	516（49.9%）
収集マルウェア数	63	14（22.2%）	45（71.4%）



6章：マルウェア配布用URLのブラックリスト化に伴う課題

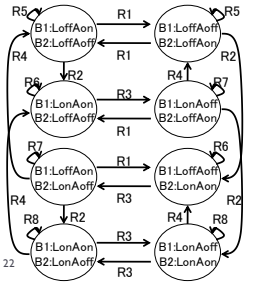
- 多くのマルウェアダウンロードサイトは乗っ取られたWebサイト
  ⇒マルウェアが削除された際はブラックリストから除外する必要有り
  - 確認周期が短い⇒監視パケット送信数増加⇒攻撃者が検知してマルウェアダウンロードサイトを変更⇒ブラックリストのフィルタが無効化
  - 確認周期が長い⇒マルウェアダウンロードサイト上のマルウェアが駆除されてからも当該サイトへのアクセスをフィルタ⇒誤検知が増加

⇒最適な監視パケット（：プローブ）送信周期の決定が重要



6章：攻撃者の行動を考慮したブラックリストの更新

- 各botのブラックリスト掲載状況（Lon/off）とマルウェアダウンロードサイト化されているか否か（Aon/off）で状態を規定
- マルコフ解析で誤検知/検知漏れを抑えるプローブ間隔を算出

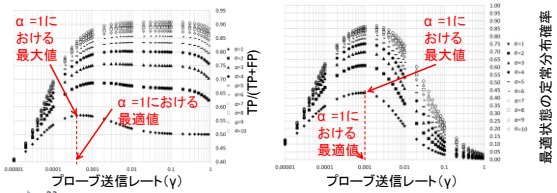


R1：攻撃者の意図によるbot変更
　レート：L（実測値から算出）
R2：防御者による攻撃bot検知
　レート：α（実測値から算出）
R3：攻撃者の意図によるbot変更または攻撃者によるプローブ検知
　レート：L+D（D=α/γ）
R4：防御者による攻撃bot遷移検知
　レート：γ
R5：exp{−(R1+R2)}
R6：exp{−R3}
R7：exp{−(R1+R2+R4)}
R8：exp{−(R3+R4)}
※α：単位時間当たりの攻撃数
※γ：単位時間当たりのプローブ数

6章：攻撃者の行動を考慮したブラックリストの更新

- 定常確率を導出後、攻撃/プローブ数を変化させて下記を算出
  - 各状態の各botがTP（マルウェアダウンロードサイトをリスト化）かTN（正常サイトを非リスト化）かFP（正常サイトをリスト化）かFN（マルウェアダウンロードサイトを非リスト化）かを特定、TP/（TP+FP)を算出
  - 全botがTPかTNである最適状態の定常分布確率を算出

⇒最適なプローブ送信レートが存在、本解析で特定可能



まとめ

- **Webサイト群をマルウェア感染から保護するためのWebサイト防衛方式の確立に向けて、以下の技術を提案**
  - マルウェアダウンロードサイトのURLをブラックリスト化する防御技術

  ⇒多様な攻撃に対する防御を実現
  - ハニーポット上で攻撃を検知するよう宛先を制御する技術

  ⇒攻撃の特徴やマルウェアを最大限収集
  - 攻撃者の行動を考慮して最適なブラックリストURL監視を実施する技術

  ⇒マルウェアダウンロードサイトのみを高精度にリスト化する監視を実現
- **本方式をWebサイト提供環境（クラウド環境/ホスティング環境 etc）などに適用することで、サービスプロバイダは安心・安全なWebサイト構築環境をユーザに提供可能**