情報ネットワーク特論
ネットワーク・フォレンジックス

名古屋大学 情報基盤センター
情報基盤ネットワーク研究部門
嶋田 創

ネットワーク・フォレンジックスとは

- フォレンジックス(forensics)
  - forensic: 法廷の、法廷で用いる、弁論の
  - 法廷で使える証拠と考えてもらえばOK
- ネットワーク・フォレンジックス:ネットワークを介して得られる法廷で使える証拠
  - サーバ内部の情報なども含む
- 様々なフォレンジックス
  - デジタル・フォレンジックス: デジタルデータ(主にファイルシステム内)による証拠
  - コンピュータ・フォレンジックス: 計算機内部からの証拠

→境界はけっこうあいまい

ネットワーク・フォレンジックスで重要なこと

- 複数の機器に分散して証拠は残る
- 揮発性の証拠が多いので素早く動く必要がある
  - 場合によっては、証拠集めの順序を決める
- 機器の操作が証拠を汚染することもある
- 止めることができない機器で作業することもある
  - 作業による影響は最小限に留めなければならない
- 書き込み保護されたコピーで行うことが望ましい
  - ただし、正確なコピーかを保証できるかどうかが難しい
- 単に証拠集めだけでなく、何らかのアクションにつなげる必要も起こりうる
  - 例: マルウェアに感染した端末の通信を遮断
  - 例: 行方不明になった端末の保護

フォレンジックス全体で重要なこと

- 対象は何も攻撃者のみではない
  - セキュリティ意識の薄い者とか内部情報を売ろうとする労働者とか
- 証拠を提示する先がそれを証拠として認識してくれるか?
  - 組織の上司とか幹部とかは?
  - 警察とか裁判官とか外部は?
  - 2つのエビデンスを意識する
    - 意見や判断が妥当であるかどうかを示す情報や痕跡のこと
    - 法的捜査における事実を立証するのに用いられる情報、および、法廷で供述として適格である情報

フォレンジックスで重要な点

OSCAR

- 情報の収集(Obtain Information)
- 戦略(Strategize)
- 証拠の収集(Collect Evidence)
- 解析(Analyze)
- 報告(Report)

情報の収集(Obtain Information)

- 何が起こったかの描写
- インシデントが発見された日付、時間
- 関わった人物
- 関わったシステム、データ
- 発見以降取られた対応

戦略(Strategize)

- 揮発性の証拠を確保するために素早く決定
- 調査の概算時間を理解する
- 持っている資源のリスト化(人員、時間、設備を含む)
- もっともらしい証拠を持つ機器等を見極める
- 機器から証拠を得るためのコストを推定する
- 入手する証拠の優先順位をつける
- 第一段階の入手/解析計画を立てる
- 初期の解析を終えたあと、立ち戻ってさらに証拠を入手

証拠の収集(Collect Evidence)

- 優先順位に従って証拠の収集
  - できるだけ早く取得する(合法的に)
  - 信頼出来るコピー作り、暗号化する
  - オリジナルを隔離し、管理とアクセスを制限する
- コピーのみを解析する
- 解析には評判の良い信頼のできるツールを使う

- 証拠入手時の注意点
  - 証拠を集めるにあたってシステムへのアクセスや取った行動の全てに対して注意深く記録を取る
  - 証拠それ自体を保存する
  - 確実に証拠がセキュアに保存され、管理の連鎖が保存される

解析(Analyze)

- 相関性: 複数のソースからの情報の相関を確認(タイムスタンプなど)
- 時系列の構築
- 重要イベントの確定
- 確証: 複数の証拠により信頼度を高める
- 解釈: 仮説を立て、証拠の意味の評価、追加の潜在的な証拠を探索
- 収集→解析の繰り返しによる証拠の強化

報告(Report)

- 誰に?
  - 上司、管理者、幹部、警察、裁判官
- 調査結果を専門でない人に向けて、(自然科学的な厳密さを備えた上で)明確に説明する必要がある

証拠のカテゴリ

- 伝統的な物: 直接、目撃、状況、業務記録
  - 直接: USBデバイス、HDD、他のPC構成要素の中身
  - 目撃: デバイスの持ち出し、デバイスの接続
  - 状況: 関連を匂わせるチャットのログ
  - 業務記録: 電子メールなどを含む業務文書
- 電子システム上の物: デジタル、ネットワーク経由のデジタル
  - デジタル: 通信のセッション、ログ
  - ネットワーク経由のログ: IDS、FW、サーバのログ
    - プライバシーの問題が大きい
- 「伝聞」は証拠の補強にはなるが単体では証拠にならない

フォレンジックスから解決までの例(1/2)

- 大学内で〇〇さんのノートPCが盗まれた(とする)
- ノートPCがなくなったのは正確にはいつだろうか?
  - 〇〇さんに最後に使った時間となくなったのに気づいた時間を尋ねる
  - 無線LANアクセスポイント、ネットワークスイッチのARPテーブル、認証サーバのログ、などを探る
- ノートPCを見つけ出して回収することができるだろうか?
  - 盗まれたあとに学内ネットワークに接続されていたかどうか確認
    - もし接続履歴があった場合は接続された場所を特定
- ノートPC上には重要データが保存されているだろうか?
  - 成績情報などの重要データは保存していないか?
  - ノートPCに保存されている可能性が高いメールはあるか?
    - メールサーバ側のデータと突き合わせ

フォレンジックスから解決までの例(2/2)

- 盗人が学内ネットワークでさらなるアクセス権を得るために〇〇さんの情報を利用していないか？
  - 情報基盤センターのサービスに何か操作はされていないか？
  - アクセスがあった場合はログが残る
    - ログを利用した犯人の絞り込み
    - ただし、さらなる情報漏えいが起きた可能性も
    - 単なる物品目的の窃盗ではなく情報も目当てだと考えられる
- 解決例
  - 無線LAN認証ログからどのアクセスポイント経由で認証されたかを特定
  - 〇〇さんに無線LANアクセスポイント周りでの行動を確認して忘れた場所を思い出してもらった

フォレンジックスから解決までの例2

- 大学内での著作権侵害の検知
- 侵入検知システム(IDS)がアラートを発した
  - P2Pファイル共有の検知
  - 学内はサブネット化されていてIPアドレスのみが分かる
- P2Pトラフィック源はどこか
  - 有線LAN? 無線LAN?
  - 有線LAN: 情報コンセントはどこに? その部屋の管理者は? 情報コンセントの先のネットワークスイッチ内のMACアドレステーブルは?
  - 無線LAN: 無線LANの認証サーバ、アクセスポイントに残ったMACアドレスは?
- 解決例
  - 有線LANであり、部屋とMACアドレスを特定
  - 部屋を利用したユーザの1人のノートPCとMACアドレスが一致

ネットワーク機器からのフォレンジックス

- ネットワークの証拠はたくさんのところにある
  - 侵入検知/防御システム(IDS/IPS)
  - ファイアウォール
  - DHCPサーバ
  - ウェブプロキシ
  - ログイン認証サーバ
  - メールサーバ
  - ネットワークスイッチ
- 揮発性が高いものが多いため、システムが動いている間に証拠を集める
- オンラインでデバイスと対話しなくてはならないことも
  - オンラインでの情報収集は環境を変更するので、影響を最小限に抑える

ネットワーク機器からの収集の戦略(1/

- システム時間を記録する
  - 常にデバイスと信頼出来るソースの時間差をチェック
  - 補正しなければ、証拠を関連づけることが難しくなる
  - 多くのツールに時間の歪みを補正する簡単な方法はない
  - 手動でログを比較するか、スクリプトを使用して補正
  - 長時間デバイスを利用していると時間のずれが変わるので、定期的に時間を補正するのがよい
- 揮発性のレベルに応じて証拠を集める
  - ちゃんと順序をつけた方が全ての証拠を集めやすい
  - 揮発性の高い証拠は集めるのが難しいものが多いので、不必要なときはこの限りではない

ネットワーク機器からの収集の戦略(2/

- 環境全体に残す足跡を最小限に抑えながら証拠を集めなければならない
- デバイスの再起動やシャットダウンを控える
  - ネットワークベースの証拠は揮発性のメモリの中に存在していることが多い
    - MACアドレスやARPのテーブルはダイナミックに変化し、リブート時に保存されない
  - ディスク領域が限られていて上書きされる設定がされている場合、ログが変更される可能性ある
- ネットワークからではなく(シリアル)コンソールから接続する
  - ネットワークからだとトラフィックを生成し状態を変更してしまう
  - 攻撃者がいる場合、攻撃者に見つかってしまう可能性あり

ネットワーク機器からの収集の戦略(3/

- 調査活動を記録する
  - CLIの場合、scriptやscreenで記録できる
  - 他の調査の時にすぐ参照できるようにする
  - 別途、すべての活動を記録するのが良い
  - GUIでは記録は難しいが、可能な限り画面のキャプチャ、写真、グラフィカルな接続記録を取る
- 調査には常に足跡を残すことになる
  - 証拠を得れば得るほど足跡は増えていく

ファイアウォール(FW)からのフォレンジックス(+IDS/IPS)

- 通信元IPアドレス/ポート、通信先IPアドレス/ポート、プロトコルは最低でも記録されている
- パケットの冒頭を利用したアプリケーションの判別
  - URLなどの特定の文字列とのマッチングもあり
- FWのログは多くの情報を保持している
  - 接続試行, データ転送量, プロトコル
    使用したアプリケーション, パケットの内容…
- FWの設定はサービスやデータが晒されたかどうかを明らかにするから
- 調査員が証拠を集めたり、システムにアクセスするためにFWの設定を変更する必要があるから
- FW自身が破損している可能性があるから

DNS(キャッシュ)サーバ

- Time to Liveで指示されている時間だけ名前解決のログを保持
- ブラックリスト入りしているドメイン名の名前解決はあるか?
- 変なドメイン名の名前解決をしていないか?
  - やたらと長いドメイン名は無いか?
  - ドメイン名のエントロピーがある?

無線LANアクセスポイントからのフォレンジックス

- 設定やログの性能は製品によって様々
  - ローエンドモデルでは
    - ウェブ管理インターフェース
    - ログ機能
    - MACアドレスフィルタリング
    - DHCPサーバ
  - ハイエンドモデルでは上記に加え
    - ルータ機能
    - システムログ
    - SNMP

802.1X認証サーバ

- 802.1XはLANの拡張認証フレームワークの規格
  - 有線/無線を問わない
  - アクセスログが認証システム内に保存されている
- バックエンド認証システムは監査場所を作りやすい
  - RADIUSシステム、Active Directory、LDAPサーバ、など

他のサーバ

- DHCPサーバ
  - IPアドレスの払い出しとMACアドレス
- ウェブプロキシ
  - アクセスしたURL
- メールサーバ
  - メールの送受信時間
  - メール到着の確認(POP, IMAP)時間

ネットワークスイッチ

- ARPテーブル
  - MACアドレスとIPアドレスの関係
- ルーティングテーブル
  - どのサブネットに通信をしたか

ログ取得(1/2)

- 調査員はデバイスログを次のように考える
  - 事件に関係のある証拠を宝庫
  - 新しい証拠を集めるもの
  - 自分自身の活動の記録を残すもの
- Local Logging
  - デフォルトで様々なログをローカルメディアに保存
  - しかし,普通は高揮発性・限られた容量
  - イベントが起こるたびにコンソールを確認しない限り情報は失われてしまうだろう
  - 無関係なログを生成して攻撃を紛らわすのは昔ながらの攻撃方法
- コンソールログ
  - キャプチャーするには”カメラ”を使うのがベスト

ログ取得(2/2)

- ターミナルログ
  - ”script”等のツールでログを保存
- SNMP
  - trapを使ってイベントを出力できる
  - ネットワーク層にアクセスできればキャプチャ可能
    - 通常は非暗号化UDPで送信するので簡単
  - リアルタイムデータを検索ならtrapの盗聴がよい
- syslog
  - 最も古く,広く利用されているログシステムの1つ
  - ログのローテートの設定は適切に
- 認証, 承認, アカウントロギング

ユーザインタフェースと記録方法

- ウェブインタフェース
  - 画面をキャプチャ
- CLI(SSHなどを含む)
  - GNU screenのセッションログ保存機能
- 面倒な商用のインタフェース
  - Apple AirPort Express
    - Mac標準搭載、Win無料、Linux無し
  - Cisco VMS, CiscoWorks, ASDM
    - Javaベース、要JVMインストール

機器へのアクセス無しでのフォレンジックス

- ポートスキャン
  - 空いているポートやポートを開けているソフトウェアのバージョン調べる効率のいい方法
  - nmap等のツールを利用
  - トラフィックを生成するので、ターゲットのデバイスの状態を変更してしまう可能性
- ターゲットのシステムに様々な既存の脆弱性がないかをテスト
  - トラフィックを作りデバイスの状態を変更
  - デバイスがクラッシュすることもあるので注意

パケット解析によるフォレンジックス

- トラフィックをキャプチャしてどうする?
  - 調査の性質に応じて文字列を探したり、ファイルを切り開いたり
- パケット解析は次のような事象で役立つ
  - 疑わしいトラフィックに関するIDSからの警告を受け、その原因を特定したい
  - 組織の人員が機密データをエクスポートしているので、外向き通信を特定のキーワードでサーチしたい
  - 原因が特定できない謎のトラフィックを発見した
- パケットを見る時の注意
  - オクテット単位で転送されます
  - 基本はビッグエンディアンです
- Wireshark: 有名なツール

パケットからのプロトコル解析

- 理想ではきちんと仕様通りプロトコルが実装
  - 現実はそんなわけなかった
  - 知的財産の保護、競争の阻止、セキュリティの目的のため故意に機密にされている
  - 他のプロトコルはシンプルだけど、文章化されてない(MSとか)
- いくつかのプロトコルはIETF指定規格で文章化
  - それでもベンダーが適切に実装するとは限らない(MSとか)
  - 製造者はプロトコルが標準化される前に実装したり、部分的にしか実装しなかったり
- 公開規格を完璧に尊守していることは稀
- この事実は日常的に攻撃者に利用
  - 侵入検知システムやFirewallを回避
  - データの密輸
  - 騒ぎを引き起こす

プロトコルごとの解析ツール

- smtpdump, findsmtpinfo
  - SMTPパケットより認証情報、本文、添付ファイルなどを抽出
- docxtract
  - docxよりファイルを抽出するスクリプト

マルウェアのフォレンジックス

- 発見が比較的簡単なもの
  - IRCプロトコルベース
  - 他のP2Pプロトコルベース
- 最近は他の通信に紛れるようにするので発見は難しい
  - 例: Waledac
    - 他ノードとの通信にHTTP POST/GETメッセージを使用
    - Referer, User-Agent文字列としてMozilla
  - ドメイン名も短めのものを利用するように発展
  - 通信をSNSを経由とするもの