3E-4

標的型サイバー攻撃の感染範囲特定方式に関する提案

松本 光弘† 高橋 洋一† 白木 宏明† 大松 史生†
三菱電機株式会社 情報技術総合研究所†

1 はじめに

近年、標的型サイバー攻撃によって、組織や企業は多大な損害を出している。攻撃を受けた場合は、システムを迅速に復旧させ、対策を講じることで、被害を最小限に抑えることができる。そのため、攻撃内容を把握し、端末の感染範囲を特定する必要がある。

そこで本論文では、攻撃者の端末内での操作を追跡することで、感染範囲を特定する方法を提案する。攻撃が検知された後に、攻撃に関する通信ログ（【表 2】攻撃通信ログ）から、攻撃に関する端末ログ（【表 4】攻撃端末ログ）を抽出し、攻撃者を特定する。感染端末を操作する攻撃者が他端末へアクセスしたことを端末ログから特定することで、感染範囲を特定する。また、攻撃通信ログに対応した攻撃端末ログが存在しない場合は、端末ログが改ざんされている、つまり、端末が感染しているとみなし、感染範囲を特定する。

2 関連研究

研究用データセット CCC DATAset2008 を時系列分析することによって、連鎖感染の可視化を行った研究がある [1]。この研究でも、本論文と同様に通信ログを用いて連鎖感染マルウェアによる脅威の全体像の把握を試みており、未知検体が多数の既知検体と関連していることや、未知検体からの連鎖はポート 80 番を使用する割合が高いといった知見を得ている。

上記論文[1]では、検証用環境にてマルウェアの振る舞いを検証しているが、本論文では業務環境にてマルウェアの感染範囲を特定する。上記研究 [1]の知見を利用することで、より精度良く感染範囲を特定することができる可能性がある。

3 攻撃シナリオ

標的型サイバー攻撃は、企業や組織の機密情報を搾取するために行われる。そのため、攻撃者は組織内の一端末を感染させ、組織内のコンピュータシステムを調査し、目的となる機密情報の在り処を特定する。IPA [2]では、以下のような攻撃シナリオを紹介している。

① 計画立案段階
② 攻撃準備段階
③ 初期潜入段階
④ 基盤構築段階
⑤ 内部潜入段階
⑥ 目的遂行段階
⑦ 再侵入

本論文は、上記の攻撃を検知した後に、マルウェアの感染範囲を特定する方法を提案する。

4 感染範囲特定方式

まず、本論文が提案する感染範囲特定の処理フローを記載する。処理の工程は以下の５つである。

処理① 攻撃通信ログを取得する。
処理② 攻撃通信ログから、攻撃端末ログを取得する。
処理③ 端末ログの改ざんを検知する。
処理④ 攻撃端末ログから、攻撃ユーザを特定する。
処理⑤ 攻撃ユーザの他端末への感染活動を特定する。

4.1 処理①攻撃通信ログ取得

感染範囲を特定するには、攻撃者の活動内容を把握する必要がある。そこで、本論文では、IPS/IDS や FW（ファイアウォール）、Proxy 等の通信装置から出力される通信ログとサーバ・端末から取得される端末ログを用いて感染範囲を特定する。

表 1に想定する通信ログを示す。

表 1 通信ログ

ID	日付	時刻	サービス	アクセス元ホスト	アクセス先ホスト	アクセス元ポート	アクセス先ポート
1	2013/5/2	13:54:43	pop3	Server_C	PC_A	25	110
2	2013/5/4	15:21:23	smtp	PC_A	Server_C	2431	25
3	2013/5/5	12:00:00	http	PC_A	cracker.com	15421	80
4	2013/7/31	20:30:02	ftp	PC_A	PC_B	50001	20
5	2013/8/2	21:15:22	ftp	PC_A	PC_D	50003	20
6	2013/8/2	21:43:54	rtmp	PC_A	movie.com	42213	554

通信ログは攻撃検知装置によって分析され、攻撃通信ログが抽出される。攻撃通信ログでは表 2のように攻撃ステップの項目が付与される。

表 2 攻撃通信ログ

ID	日付	時刻	攻撃ステップ	サービス	アクセス元ホスト	アクセス先ホスト	アクセス元ポート	アクセス先ポート
1	2013/5/2	13:54:43	3	pop3	Server_C	PC_A	25	110
2	2013/5/5	12:00:00	4	http	PC_A	cracker.com	15421	80
3	2013/7/31	20:30:02	5	ftp	PC_A	PC_B	50001	20

Detection of Computers Suspected Infected with Virus against Targeted Attacks
Mitushiro Matsumoto†, Yoichi Takahashi†, Hiroaki Shiraki†, Fumio Ohmatsu†
Information Technology R&D Center, Mitsubishi Electric Corporation. †

攻撃ステップは3章に記載した各攻撃段階を表しており、表 2の第 1 レコードは攻撃ステップ 3 の初期潜入段階（メールや Web からマルウェアを端末に潜入させる）に関する通信ログレコードであることを示している。

4.2 処理②，③攻撃端末ログ特定＋端末ログ改ざん検知

表 2の第 1～3 レコードから、PC_A は攻撃ステップ 3～5 の一連の攻撃を受けたことが分かるため、PC_A の端末ログを分析して、攻撃者の端末内での活動内容を把握する。サーバ・端末からは以下のようなログを取得できる。

表 3　端末ログ

ID	日付	時刻	アクセス元ホスト	アクセス先ホスト	アクセス元ユーザ	アクセス先ユーザ	アクセスファイル	イベント
1	2013/5/2	13:54:44	Server_C	PC_A	System	user_A1	0502.eml	move
2	2013/5/5	12:00:02	PC_A	cracker.com	user_A1	-	rat.exe	move
3	2013/7/31	20:30:02	PC_A	PC_B	user_A1	user_B1	rat.exe	move
4	2013/8/2	20:30:10	PC_B	-	user_B1	-	rat.exe	read
5	2013/8/2	21:15:23	PC_A	PC_D	user_A1	user_D1	rat.exe	move
6	2013/8/3	5:21:41	PC_A	Server_E	user_A1	user_E1	rat.exe	move

表 2と表 3の日付・時刻やアクセス元ホストおよびアクセス先ホストから攻撃通信ログと端末ログの対応付けを行い、攻撃に関する端末ログ（攻撃端末ログ）を取得する。時刻については、通信ログと端末ログとで、ログを取得する装置が異なるため、多少のズレが生じる可能性があるため、ある程度のマージン（1,2 秒程度）を考慮する必要がある。

表 2の第 1～3 レコードと表 3の第 1～3 レコードから、表 4の攻撃端末ログを取得できる。

表 4　攻撃端末ログ

ID	日付	時刻	攻撃ステップ	アクセス元ホスト	アクセス先ホスト	アクセス元ユーザ	アクセス先ユーザ	アクセスファイル	イベント
1	2013/5/2	13:54:44	3	Server_C	PC_A	System	user_A1	0502.eml	move
2	2013/5/5	12:00:02	4	PC_A	cracker.com	user_A1	-	rat.exe	move
3	2013/7/31	20:30:02	5	PC_A	PC_B	user_A1	user_B1	rat.exe	move

攻撃通信ログの各レコードに対応した端末ログのレコードが存在しない場合は、端末ログが改ざんされている可能性が高いため、端末が感染していると判断する。

4.3 処理④攻撃ユーザ特定

表 4の第 1～3 レコードから、端末 PC_A のユーザ user_A1 は攻撃ステップ 3～5 の攻撃段階に関わっている。このような一連の攻撃に加担しているユーザを特定することで、攻撃ユーザを特定する。

4.4 処理⑤他端末への感染活動を特定

処理④にて攻撃ユーザが特定されているため、攻撃ユーザの他端末へのアクセスログを特定することで、他端末への感染活動を特定する。端末 PC_A の攻撃ユーザ User_A1 は表 3の第 3,5,6 レコードより、PC_B, PC_D, Server_E にアクセスしていることが分かる。よって、PC_B, PC_D, Server_E はマルウェアに感染している可能性が高い。

一方、処理③で端末ログに改ざんが特定された場合は、端末ログから感染活動を特定することができない。そこで、通信ログを用いて感染活動を特定する。表 1の第 4,5,7 レコードより、端末 PC_A からアクセスされている端末は、PC_B, PC_D, Server_E であるため、PC_B, PC_D, Server_E はマルウェアに感染している可能性がある。

端末ログが改ざんされている場合は、攻撃ユーザを特定することができないため、感染端末から他端末への全てのアクセスを特定する必要があり、誤検知が増えてしまう。一方、端末ログが改ざんされていない場合は、攻撃ユーザが特定できるため、攻撃ユーザの他端末へのアクセスのみを特定することで、感染端末を特定することができ、感染端末特定精度を高くすることができる。

4.5 再帰的処理による感染範囲特定

処理⑤で、他端末への感染活動が検出された場合は、感染の可能性のある端末（4.4節の場合、PC_B, PC_D, Server_E）について、処理②～⑤を繰り返す。これにより、攻撃が検知された端末から再帰的に感染範囲を特定できる。

5 まとめ

本論文は、標的型サイバー攻撃において、マルウェアに感染した端末を特定する感染範囲特定方式を提案した。

本方式は、通信ログと端末ログを組み合わせることで、端末ログの改ざんを発見できると共に、改ざんがない場合は、攻撃ユーザを特定し、感染範囲を精度よく特定することができる。

しかしながら、組織内のシステムに多くの端末が接続されていた場合は、膨大な端末ログを処理する必要があり、運用や処理速度に問題を生じる可能性がある。そのため、今後は通信ログのみで精度よく感染範囲を特定できる方式を考える必要がある。

参考文献

1. **松木隆宏.** 時系列分析による連鎖感染の可視化と検体種別の推測. マルウェア対策研究人材育成ワークショップ 2008 (MWS 2008)

2. **IPA.** 「標的型メール攻撃」対策に向けたシステム設計ガイド. 2013 年.
http://www.ipa.go.jp/security/vuln/newattack.html (2013/12/25 アクセス)