

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5083760号
(P5083760)

(45) 発行日 平成24年11月28日(2012.11.28)

(24) 登録日 平成24年9月14日(2012.9.14)

(51) Int.Cl. F 1
G 0 6 F 21/00 (2006.01)
 G 0 6 F 21/00 1 5 6 G
 G 0 6 F 21/00 1 5 6 E

請求項の数 20 (全 23 頁)

(21) 出願番号	特願2007-203281 (P2007-203281)	(73) 特許権者	301022471
(22) 出願日	平成19年8月3日(2007.8.3)		独立行政法人情報通信研究機構
(65) 公開番号	特開2009-37545 (P2009-37545A)		東京都小金井市貫井北町4-2-1
(43) 公開日	平成21年2月19日(2009.2.19)	(74) 代理人	100130111
審査請求日	平成22年8月2日(2010.8.2)		弁理士 新保 斉
		(72) 発明者	中尾 康二
			東京都小金井市貫井北町4-2-1 独立 行政法人情報通信研究機構内
		(72) 発明者	吉岡 克成
			東京都小金井市貫井北町4-2-1 独立 行政法人情報通信研究機構内
		(72) 発明者	井上 大介
			東京都小金井市貫井北町4-2-1 独立 行政法人情報通信研究機構内

最終頁に続く

(54) 【発明の名称】 マルウェアの類似性検査方法及び装置

(57) 【特許請求の範囲】

【請求項1】

ネットワーク上で他のコンピュータに対して不正処理を行う第1のソフトウェアの処理結果から得られる第1の挙動情報と、検査対象の第2のソフトウェアの処理結果から得られる第2の挙動情報とを比較して両者の類似性を検査するマルウェアの類似性検査方法であって、

コンピュータの第1アドレス走査情報検出手段が、該第1の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第1アドレス走査情報検出工程、

コンピュータの第1マルウェア分析手段が、該第1の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第1マルウェア分析工程、

コンピュータの第2アドレス走査情報検出手段が、該第2の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第2アドレス走査情報検出工程、

コンピュータの第2マルウェア分析手段が、該第2の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第2マルウェア分析工程、

コンピュータの挙動比較手段が、上記第1の挙動情報に関する各工程による情報と、上記第2の挙動情報に関する各工程による情報との相互の同一性又は類似性を所定の比較式

10

20

に基づいて比較処理する挙動比較工程

を有することを特徴とするマルウェアの類似性検査方法。

【請求項 2】

ネットワーク上で他のコンピュータに対して不正処理を行う第 1 のソフトウェアの処理結果から得られる第 1 の挙動情報と、検査対象の第 2 のソフトウェアの処理結果から得られる第 2 の挙動情報とを比較して両者の類似性を検査するマルウェアの類似性検査方法であって、

コンピュータの第 1 アドレス走査情報検出手段が、該第 1 の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第 1 アドレス走査情報検出工程、

コンピュータの第 1 脆弱性攻撃コード検出手段が、該第 1 の挙動情報のうち脆弱性攻撃コードを検出して記憶手段に格納する第 1 脆弱性攻撃コード検出工程、

コンピュータの第 1 マルウェア分析手段が、該第 1 の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第 1 マルウェア分析工程、

コンピュータの第 2 アドレス走査情報検出手段が、該第 2 の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第 2 アドレス走査情報検出工程、

コンピュータの第 2 脆弱性攻撃コード検出手段が、該第 2 の挙動情報のうち脆弱性攻撃コードを検出して記憶手段に格納する第 2 脆弱性攻撃コード検出工程、

コンピュータの第 2 マルウェア分析手段が、該第 2 の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第 2 マルウェア分析工程、

コンピュータの挙動比較手段が、上記第 1 の挙動情報に関する各工程による情報と、上記第 2 の挙動情報に関する各工程による情報との相互の同一性又は類似性を所定の比較式に基づいて比較処理する挙動比較工程

を有することを特徴とするマルウェアの類似性検査方法。

【請求項 3】

前記第 1 アドレス走査情報検出工程において前記第 1 アドレス走査情報検出手段が、及び前記第 2 アドレス走査情報検出工程において前記第 2 アドレス走査情報検出手段が、それぞれ、

各宛先ポートのアクセス回数の比率か、連続するパケットのソースポート番号間の差異の平均値か、用いられるプロトコルの比率か、TCP (Transmission Control Protocol) における TCP フラグの比率か、単位時間あたりのパケットの平均個数かの少なくともいずれかの統計情報を検出する

請求項 1 又は 2 に記載のマルウェアの類似性検査方法。

【請求項 4】

前記第 1 脆弱性攻撃コード検出工程において前記第 1 脆弱性攻撃コード検出手段が、及び前記第 2 脆弱性攻撃コード検出工程において前記第 2 脆弱性攻撃コード検出手段が、それぞれ、

脆弱性攻撃コードに含まれるシェルコード内の命令のシーケンスを検出する

請求項 2 に記載のマルウェアの類似性検査方法。

【請求項 5】

前記第 1 マルウェア分析工程において前記第 1 マルウェア分析手段が、及び前記第 2 マルウェア分析工程において前記第 2 マルウェア分析手段が、それぞれ、

マルウェア本体を逆アセンブリし、そのアセンブリコードを取得する

請求項 1 ないし 4 のいずれかに記載のマルウェアの類似性検査方法。

【請求項 6】

前記第 1 マルウェア分析工程において前記第 1 マルウェア分析手段が、及び前記第 2 マルウェア分析工程において前記第 2 マルウェア分析手段が、それぞれ、

10

20

30

40

50

マルウェアの実行によるファイル又はレジストリのアクセスログか、プラットフォーム上のAPI (Application Program Interface) ログか、複数のサーバへのアクセスログか、マルウェアにより生成されるパケットログかの少なくともいずれかのログを取得する請求項1ないし5のいずれかに記載のマルウェアの類似性検査方法。

【請求項7】

前記挙動比較工程においてコンピュータの挙動比較手段が、
前記第1アドレス走査情報検出手段で検出された走査シグネチャと、前記第2アドレス走査情報検出手段で検出された走査シグネチャとの同一性を照合する
請求項1ないし6のいずれかに記載のマルウェアの類似性検査方法。

【請求項8】

前記挙動比較工程においてコンピュータの挙動比較手段が、
前記第1アドレス走査情報検出手段で検出された宛先ポートのポート番号の比率と、前記第2アドレス走査情報検出手段で検出された宛先ポートのポート番号の比率との類似性を相関関係式により算出する
請求項1ないし7のいずれかに記載のマルウェアの類似性検査方法。

【請求項9】

前記挙動比較工程においてコンピュータの挙動比較手段が、
前記第1脆弱性攻撃コード検出手段で検出した脆弱性攻撃コードと、前記第2脆弱性攻撃コード検出手段で検出した脆弱性攻撃コードとを、それぞれ要約処理し両者の同一性を照合する
請求項2に記載のマルウェアの類似性検査方法。

【請求項10】

前記挙動比較工程においてコンピュータの挙動比較手段が、
前記第1マルウェア分析手段で検出したマルウェアのコードと、前記第2マルウェア分析手段で検出したマルウェアのコードとを、それぞれ要約処理し両者の同一性を照合する
請求項1ないし9のいずれかに記載のマルウェアの類似性検査方法。

【請求項11】

ネットワーク上で他のコンピュータに対して不正処理を行う第1のソフトウェアの処理結果から得られる第1の挙動情報と、検査対象の第2のソフトウェアの処理結果から得られる第2の挙動情報とを比較して両者の類似性を検査するマルウェアの類似性検査システムであって、

該第1の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第1アドレス走査情報検出手段と、

該第1の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第1マルウェア分析手段と、

該第2の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第2アドレス走査情報検出手段と、

該第2の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第2マルウェア分析手段と、

上記第1の挙動情報に関する各工程による情報と、上記第2の挙動情報に関する各工程による情報との相互の同一性又は類似性を所定の比較式に基づいて比較処理する挙動比較手段と

を少なくとも備えたことを特徴とするマルウェアの類似性検査システム。

【請求項12】

ネットワーク上で他のコンピュータに対して不正処理を行う第1のソフトウェアの処理結果から得られる第1の挙動情報と、検査対象の第2のソフトウェアの処理結果から得られる第2の挙動情報とを比較して両者の類似性を検査するマルウェアの類似性検査システムであって、

該第1の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第1アドレス走査情報検出手段と、

該第 1 の挙動情報のうち脆弱性攻撃コードを検出して記憶手段に格納する第 1 脆弱性攻撃コード検出手段と、

該第 1 の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第 1 マルウェア分析手段と、

該第 2 の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第 2 アドレス走査情報検出手段と、

該第 2 の挙動情報のうち脆弱性攻撃コードを検出して記憶手段に格納する第 2 脆弱性攻撃コード検出手段と、

第 2 の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第 2 マルウェア分析手段と、

上記第 1 の挙動情報に関する各工程による情報と、上記第 2 の挙動情報に関する各工程による情報との相互の同一性又は類似性を所定の比較式に基づいて比較処理する挙動比較手段と

を少なくとも備えたことを特徴とするマルウェアの類似性検査システム。

【請求項 1 3】

前記第 1 アドレス走査情報検出手段及び前記第 2 アドレス走査情報検出手段が、それぞれ、

各宛先ポートのアクセス回数の比率か、連続するパケットのソースポート番号間の差異の平均値か、用いられるプロトコルの比率か、TCP (Transmission Control Protocol) における TCP フラグの比率か、単位時間あたりのパケットの平均個数かの少なくともいずれかの統計情報を検出する

請求項 1 1 又は 1 2 に記載のマルウェアの類似性検査システム。

【請求項 1 4】

前記第 1 脆弱性攻撃コード検出手段及び前記第 2 脆弱性攻撃コード検出手段が、それぞれ、

脆弱性攻撃コードに含まれるシェルコード内の命令のシーケンスを検出する

請求項 1 2 に記載のマルウェアの類似性検査システム。

【請求項 1 5】

前記第 1 マルウェア分析手段及び前記第 2 マルウェア分析手段が、それぞれ、

マルウェア本体を逆アセンブリし、そのアセンブリコードを取得する

請求項 1 1 ないし 1 4 のいずれかに記載のマルウェアの類似性検査システム。

【請求項 1 6】

前記第 1 マルウェア分析手段及び前記第 2 マルウェア分析手段が、それぞれ、

マルウェアの実行によるファイル又はレジストリのアクセスログか、プラットフォーム上の API (Application Program Interface) ログか、複数のサーバへのアクセスログか、マルウェアにより生成されるパケットログかの少なくともいずれかのログを取得する

請求項 1 1 ないし 1 5 のいずれかに記載のマルウェアの類似性検査システム。

【請求項 1 7】

前記挙動比較手段が、

前記第 1 アドレス走査情報検出手段で検出された走査シグネチャと、前記第 2 アドレス走査情報検出手段で検出された走査シグネチャとの同一性を照合する

請求項 1 1 ないし 1 6 のいずれかに記載のマルウェアの類似性検査システム。

【請求項 1 8】

前記挙動比較手段が、

前記第 1 アドレス走査情報検出手段で検出された宛先ポートのポート番号の比率と、前記第 2 アドレス走査情報検出手段で検出された宛先ポートのポート番号の比率との類似性を相関関係式により算出する

請求項 1 1 ないし 1 7 のいずれかに記載のマルウェアの類似性検査システム。

【請求項 1 9】

前記挙動比較手段が、

10

20

30

40

50

前記第1脆弱性攻撃コード検出手段で検出した脆弱性攻撃コードと、前記第2脆弱性攻撃コード検出手段で検出した脆弱性攻撃コードとを、それぞれ要約処理し両者の同一性を照合する

請求項12に記載のマルウェアの類似性検査システム。

【請求項20】

前記挙動比較手段が、

前記第1マルウェア分析手段で検出したマルウェアのコードと、前記第2マルウェア分析手段で検出したマルウェアのコードとを、それぞれ要約処理し両者の同一性を照合する請求項11ないし19のいずれかに記載のマルウェアの類似性検査システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はマルウェアの類似性を検査する方法と装置に関し、特に2種類以上の異なる程度の検査を組み合わせて効率的にマルウェアの特定を行う技術に関する。

【背景技術】

【0002】

ボットのような高度に組織化され、精巧になったマルウェアの最近の異常発生により、それらを検出し、解析し、対応する技法の必要性が高まっている。多様な商業的なプロジェクト、学問的なプロジェクト、または政府支援によるプロジェクトが研究を進めている。(非特許文献1~13参照)。

【0003】

従前のステップとして、これらのプロジェクトの多くは、ネットワーク・イベント監視に基づいて、特定のポート番号でのアクセスの急激な増加等、統計的なデータの提供に集中している。これらの活動の過程で、世界的に公表されている未使用のIPアドレス(非特許文献1、2、14参照)の集合であるダークアドレス空間を監視することが一般的な手法である。

【0004】

これらのアドレス空間では、多様なマルウェアをひき付けるためにハニーポットを設置したり、あるいはマルウェアが感染先を探索するために行う走査、DDoS攻撃のバックスキュタ等を含む入信パケットを監視する(ブラックホール監視)センサを用意する(非特許文献15~18参照)。

【0005】

別の一般的な方法は、実際のネットワークに設置されたIDSログとFWログを分析することである。ネットワーク・イベント監視に基づいたこれらの巨視的観察をマクロ分析と呼ぶ。マクロ分析は、世界的に分散されたセンサによって、インターネット上のマルウェア動作の過程で(走査等の)巨視的な挙動を把握するために使用する。しかし、それは巨視的なレベルの遠隔観察に基づいて、攻撃者の挙動及び攻撃者とセンサ間の環境に関する明示的な情報なしに実行されるため、多くの場合、結果にある程度の不確実性を残す問題がある。

【0006】

他方、実際のマルウェア実行ファイルを分析することには別の課題がある。マルウェアの構造を解析する際、マルウェア実行ファイルを逆アセンブルするためにリバース・エンジニアリング技法が適用される(非特許文献19、20参照)。

また、マルウェア・コードが実際には閉じられた(アクセスが制御された)実験環境で実行されるサンドボックス分析はその挙動を観察することができる(非特許文献19、21~23参照)。

【0007】

マルウェア・コード自体を目標としたこれらの微視的分析をミクロ分析と呼ぶ。ミクロ分析は、閉じられた実験環境において実行されるため、実際のネットワークでのマルウェアの活動を観察することはできないが、マルウェアの詳細な構造及び挙動を明らかにする

10

20

30

40

50

。 前述したマクロ分析とミクロ分析が研究され、多様な分析システムに配備されているが、これらの活動から得られた知識は効果的かつ効率的にリンクされておらず、セキュリティ・インシデントの根本原因の特定をさらに難しくしている。

【 0 0 0 8 】

これらの課題を解決するため、本件発明者は戦術的緊急対応のためのネットワーク・インシデント分析センタ (n i c t e r) を開発されている (非特許文献 3、19、23、24 参照)。 n i c t e r は、マクロ分析による実ネットワーク上の攻撃の観察及びミクロ分析による「実験室内の」マルウェア分析を関連付けることで、観察された攻撃について、その考えられる根本原因、すなわちマルウェアとを結び付けることを可能とする、マクロ - ミクロ相関分析を実現している。

10

【 0 0 0 9 】

しかしながら、相関分析はおもにブラックホール監視によって観察されるマルウェアの走査挙動の関連付けによってだけ行われてきたため、マクロ分析とミクロ分析の間のリンクは詳細な攻撃の挙動の正確な識別を保証するほど強力ではなかった。

なお、このような 2 つ以上のマルウェアによる挙動の類似性を検査する技術として、非特許文献 24、25 が開示されており、本発明出願時未公開の特許文献 1、2 において開示されている。

【 0 0 1 0 】

【非特許文献 1】Bailey, M., Cooke, E., Jahanian, F., Nazario, J., Watson, D.: The internet motion sensor: a distributed blackhole monitoring system, The 12th Annual Network and Distributed System Security Symposium (NDSS05), 2005年

20

【非特許文献 2】Moore, D.: Network telescopes: tracking Denial-of-Service attacks and internet worms around the globe, 17th Large Installation Systems Administration Conference (LISA '03), USENIX, 2003年

【非特許文献 3】Nakao, K., Yoshioka, K., Inoue, D., Eto, M., Rikitake, K.: nictet: an incident analysis system using correlation between network monitoring and malware analysis, The 1st Joint Workshop on Information Security (JWIS06), pp. 363 - 377, 2006年

【非特許文献 4】Yegneswaran, V., Barford, P., Plonka, D.: On the design and use of Internet sinks for network abuse monitoring. Recent Advances in Intrusion Detection (RAID 2004), LNCS 3224, pp146 - 165, 2004年

30

【非特許文献 5】HoneyPot project, <http://www.leurrecom.org/>

【非特許文献 6】Internet Motion Sensor, <http://ims.eecs.umich.edu/>

【非特許文献 7】IT Security Center, Information-Technology Promotion Agency, Japan, <https://www.ipa.go.jp/security/index-e.html>

【非特許文献 8】Japan Computer Emergency Response Team Coordination Center, <http://jpcert.jp/isdas/index-en.html>

【非特許文献 9】National Cyber Security Center, Korea, <http://www.ncsc.go.kr/eng/>

【非特許文献 10】REN-ISAC: Research and Education Networking Information Sharing and Analysis Center, <http://www.ren-isac.net/>

40

【非特許文献 11】SANS Internet Storm Center, <http://isc.sans.org/>

【非特許文献 12】Telecom Information Sharing and Analysis Center, Japan, <https://www.telecom-isac.jp/>

【非特許文献 13】@police, http://www.cyberpolice.go.jp/english/obs_e.html

【非特許文献 14】Bailey, M., Cooke, E., Jahanian, F., Myrick, A., Sinha, S.: Practical darknet measurement, 2006 Conference on Information Sciences and Systems (CISS '06), pp. 1496 - 1501, 2006年

【非特許文献 15】Alata, E., Nicomette, V., Kaaniche, M., Dacier, M.: Lessons learned from the deployment of a high-interaction honeypot, 6th European Dependable

50

Computing Conference(EDCC-6), pp. 39 - 44, 2006年

【非特許文献16】Leita, C., Dacier, M., Massicotte, F.: Automatic handling of protocol dependencies and reaction to 0-day attacks with ScriptGen based honeypots, 9th International Symposium on Recent Advances in Intrusion Detection (RAID2006), pp. 185 - 205, 2006年

【非特許文献17】Provos, N.: Honeyd A virtual honeypot daemon. 10th DFN-CERT Workshop, 2003年

【非特許文献18】Provos, N.: A virtual honeypot framework. 13th USENIX Security Symposium, pp 1 - 14, 2004年

【非特許文献19】Nakao, K., Matsumoto, F., Inoue, D., Baba, S., Suzuki, K., Eto, M., Yoshioka, K., Rikitake, K., Hori, Y.: Visualization technologies of nictex incident analysis system, IEICE Technical Report, vol.106, no. ISEC-176 pp.83 - 89, 2006年

【非特許文献20】Isawa, R., Ichikawa, S., Shiraishi, Y., Mohri, M., Morii, M.: A virus analysis supporting system; for automatic grasping virus behavior by code-analysis result, The Computer Security Symposium 2005 (CSS2005), vol. 1, pp. 169 - 174, 2006年

【非特許文献21】Hoshizawa, Y., Morii, M., Nakao, K.: A proposal of automated malware behavior analysis system, Information and Communication System Security, IEICE, ICSS2006-07, pp. 41 - 46, 2006年

【非特許文献22】C. Willems, T. Holz, and F. Freiling, "Toward Automated Dynamic Malware Analysis Using CWSandbox" Security & Privacy Magazine, IEEE, Volume 5, Issue 2, pp. 32 - 39, 2007年

【非特許文献23】Nakao, K., Matsumoto, F., Inoue, D., Baba, S., Suzuki, K., Eto, M., Yoshioka, K., Rikitake, K., Hori, Y.: Visualization technologies of nictex incident analysis system, IEICE Technical Report, vol.106, no. ISEC-176 pp.83 - 89, 2006年

【非特許文献24】Yoshioka, K., Eto, M., Inoue, D., Nakao, K.: Macro-micro correlation analysis for binding darknet traffic and malwares, The 2007 Symposium on Cryptography and Information Security(SCIS2007), 2F2-2, 2007年

【非特許文献25】衛藤 将史、園田 光太郎、吉岡 克成、井上 大介、竹内 純一、中尾 康二、「スペクトラム解析を用いたマルウェアの類似性検査手法の提案」 IEICE、SCIS2007、2007年

【特許文献1】日本特許出願2007年12070号

【特許文献2】日本特許出願2007年12071号

【発明の開示】

【発明が解決しようとする課題】

【0011】

本発明は上記従来技術が有する問題点に鑑みて創出されたものであり、その目的はマルウェアのマイクロ分析及びマクロ分析の相関に基づくマルウェアの検査精度を向上させると共に、効率よくマルウェアの検査を行う技術を提供することである。

【課題を解決するための手段】

【0012】

本発明は、上記の課題を解決するために、次のようなマルウェア類似性検査方法を提供する。

請求項1に記載の発明によれば、ネットワーク上で他のコンピュータに対して不正処理を行う第1のソフトウェアの処理結果から得られる第1の挙動情報と、検査対象の第2のソフトウェアの処理結果から得られる第2の挙動情報とを比較して両者の類似性を検査するマルウェアの類似性検査方法を提供する。

【0013】

本方法はコンピュータ上に実装されるものであって次の各工程からなる。

(S1) コンピュータの第1アドレス走査情報検出手段が、第1の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第1アドレス走査情報検出工程

(S2) コンピュータの第1マルウェア分析手段が、第1の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第1マルウェア分析工程、

(S3) コンピュータの第2アドレス走査情報検出手段が、第2の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第2アドレス走査情報検出工程、

(S4) コンピュータの第2マルウェア分析手段が、第2の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第2マルウェア分析工程、

(S5) コンピュータの挙動比較手段が、上記第1の挙動情報に関する各工程による情報と、上記第2の挙動情報に関する各工程による情報との相互の同一性又は類似性を所定の比較式に基づいて比較処理する挙動比較工程。

【0014】

請求項2に記載の発明によれば、上記請求項1の各工程に加え、脆弱性攻撃コードに係る類似性検査方法を提供する。すなわち、次の各工程からなる。

(S1) コンピュータの第1アドレス走査情報検出手段が、第1の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第1アドレス走査情報検出工程、

(S02) コンピュータの第1脆弱性攻撃コード検出手段が、第1の挙動情報のうち脆弱性攻撃コードを検出して記憶手段に格納する第1脆弱性攻撃コード検出工程、

(S2) コンピュータの第1マルウェア分析手段が、第1の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第1マルウェア分析工程、

(S3) コンピュータの第2アドレス走査情報検出手段が、第2の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第2アドレス走査情報検出工程、

(S04) コンピュータの第2脆弱性攻撃コード検出手段が、第2の挙動情報のうち脆弱性攻撃コードを検出して記憶手段に格納する第2脆弱性攻撃コード検出工程、

(S4) コンピュータの第2マルウェア分析手段が、第2の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第2マルウェア分析工程、

(S5) コンピュータの挙動比較手段が、上記第1の挙動情報に関する各工程による情報と、上記第2の挙動情報に関する各工程による情報との相互の同一性又は類似性を所定の比較式に基づいて比較処理する挙動比較工程。

【0015】

請求項3に記載の発明によれば、上記の第1アドレス走査情報検出工程において第1アドレス走査情報検出手段が、さらに第2アドレス走査情報検出工程において第2アドレス走査情報検出手段が、それぞれ、各宛先ポートのアクセス回数の比率か、連続するパケットのソースポート番号間の差異の平均値か、用いられるプロトコルの比率か、TCP (Transmission Control Protocol) におけるTCPフラグの比率か、単位時間あたりのパケットの平均個数かの少なくともいずれかの統計情報を検出することを特徴とする。

【0016】

請求項4に記載の発明によれば、上記の第1脆弱性攻撃コード検出工程において第1脆弱性攻撃コード検出手段が、さらに第2脆弱性攻撃コード検出工程において第2脆弱性攻撃コード検出手段が、それぞれ、脆弱性攻撃コードに含まれるシェルコード内の命令のシーケンスを検出することを特徴とする。

10

20

30

40

50

【 0 0 1 7 】

請求項 5 に記載の発明によれば、第 1 マルウェア分析工程において第 1 マルウェア分析手段が、さらに第 2 マルウェア分析工程において第 2 マルウェア分析手段が、それぞれ、マルウェア本体を逆アセンブリし、そのアセンブリコードを取得することを特徴とする。

【 0 0 1 8 】

請求項 6 に記載の発明によれば、第 1 マルウェア分析工程において第 1 マルウェア分析手段が、第 2 マルウェア分析工程において第 2 マルウェア分析手段が、それぞれ、マルウェアの実行によるファイル又はレジストリのアクセスログか、プラットフォーム上の A P I (Application Program Interface) ログか、複数のサーバへのアクセスログか、マルウェアにより生成されるパケットログかの少なくともいずれかのログを取得することを特徴とする。

10

【 0 0 1 9 】

請求項 7 に記載の発明によれば、上記の挙動比較工程においてコンピュータの挙動比較手段が、第 1 アドレス走査情報検出手段で検出された走査シグネチャと、第 2 アドレス走査情報検出手段で検出された走査シグネチャとの同一性を照合することを特徴とする。

【 0 0 2 0 】

請求項 8 に記載の発明によれば、上記の挙動比較工程においてコンピュータの挙動比較手段が、前記第 1 アドレス走査情報検出手段で検出された宛先ポートのポート番号の比率と、前記第 2 アドレス走査情報検出手段で検出された宛先ポートのポート番号の比率との類似性を相関関係式により算出することを特徴とする。

20

【 0 0 2 1 】

請求項 9 に記載の発明によれば、挙動比較工程においてコンピュータの挙動比較手段が、第 1 脆弱性攻撃コード検出手段で検出した脆弱性攻撃コードと、第 2 脆弱性攻撃コード検出手段で検出した脆弱性攻撃コードとを、それぞれ要約処理し両者の同一性を照合することを特徴とする。

【 0 0 2 2 】

請求項 1 0 に記載の発明によれば、挙動比較工程においてコンピュータの挙動比較手段が、第 1 マルウェア分析手段で検出したマルウェア本体のコードと、第 2 マルウェア分析手段で検出したマルウェア本体のコードとを、それぞれ要約処理し両者の同一性を照合することを特徴とする。

30

【 0 0 2 3 】

また、本発明は、次のようなマルウェアの類似性検査システムを提供することもできる。

すなわち、請求項 1 1 に記載の発明によれば、ネットワーク上で他のコンピュータに対して不正処理を行う第 1 のソフトウェアの処理結果から得られる第 1 の挙動情報と、検査対象の第 2 のソフトウェアの処理結果から得られる第 2 の挙動情報とを比較して両者の類似性を検査するマルウェアの類似性検査システムを提供することができる。

【 0 0 2 4 】

本システムには、第 1 の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第 1 アドレス走査情報検出手段と、第 1 の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第 1 マルウェア分析手段と、第 2 の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第 2 アドレス走査情報検出手段と、第 2 の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第 2 マルウェア分析手段と、上記第 1 の挙動情報に関する各工程による情報と、上記第 2 の挙動情報に関する各工程による情報との相互の同一性は類似性を所定の比較式に基づいて比較処理する挙動比較手段とを少なくとも備えたことを特徴とする。

40

【 0 0 2 5 】

請求項 1 2 に記載の発明によれば、ネットワーク上で他のコンピュータに対して不正処

50

理を行う第1のソフトウェアの処理結果から得られる第1の挙動情報と、検査対象の第2のソフトウェアの処理結果から得られる第2の挙動情報とを比較して両者の類似性を検査するマルウェアの類似性検査システムを提供することができる。

【0026】

該システムにおいて、第1の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第1アドレス走査情報検出手段と、第1の挙動情報のうち脆弱性攻撃コードを検出して記憶手段に格納する第1脆弱性攻撃コード検出手段と、第1の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第1マルウェア分析手段と、第2の挙動情報のうちネットワークアドレスの走査に係る挙動の統計情報を検出して記憶手段に格納する第2アドレス走査情報検出手段と、第2の挙動情報のうち脆弱性攻撃コードを検出して記憶手段に格納する第2脆弱性攻撃コード検出手段と、第2の挙動情報のうちマルウェア本体のコード及び挙動の少なくともいずれかを分析しその結果を記憶手段に格納する第2マルウェア分析手段と、上記第1の挙動情報に関する各工程による情報と、上記第2の挙動情報に関する各工程による情報との相互の同一性又は類似性を所定の比較式に基づいて比較処理する挙動比較手段とを少なくとも備えたことを特徴とする。

10

【0027】

請求項13に記載の発明によれば、上記の第1アドレス走査情報検出手段と第2アドレス走査情報検出手段が、それぞれ、各宛先ポートのアクセス回数の比率か、連続するパケットのソースポート番号間の差異の平均値か、用いられるプロトコルの比率か、TCP (Transmission Control Protocol) におけるTCPフラグの比率か、単位時間あたりのパケットの平均個数かの少なくともいずれかの統計情報を検出することを特徴とする。

20

【0028】

請求項14に記載の発明によれば、上記の第1脆弱性攻撃コード検出手段と第2脆弱性攻撃コード検出手段が、それぞれ、脆弱性攻撃コードに含まれるシェルコード内の命令のシーケンスを検出することを特徴とする。

【0029】

請求項15に記載の発明によれば、上記の第1マルウェア分析手段と第2マルウェア分析手段が、それぞれ、マルウェア本体を逆アセンブリし、そのアセンブリコードを取得することを特徴とする。

30

【0030】

請求項16に記載の発明によれば、上記の第1マルウェア分析手段と第2マルウェア分析手段が、それぞれ、マルウェアの実行によるファイル又はレジストリのアクセスログか、プラットフォーム上のAPI (Application Program Interface) ログか、複数のサーバへのアクセスログか、マルウェアにより生成されるパケットログかの少なくともいずれかのログを取得することを特徴とする。

【0031】

請求項17に記載の発明によれば、上記挙動比較手段が、第1アドレス走査情報検出手段で検出された走査シグネチャと、第2アドレス走査情報検出手段で検出された走査シグネチャとの同一性を照合することを特徴とする。

40

【0032】

請求項18に記載の発明によれば、上記挙動比較手段が、第1アドレス走査情報検出手段で検出された宛先ポートのポート番号の比率と、第2アドレス走査情報検出手段で検出された宛先ポートのポート番号の比率との類似性を相関関係式により算出することを特徴とする。

【0033】

請求項19に記載の発明によれば、上記挙動比較手段が、第1脆弱性攻撃コード検出手段で検出した脆弱性攻撃コードと、前記第2脆弱性攻撃コード検出手段で検出した脆弱性攻撃コードとを、それぞれ要約処理し両者の同一性を照合することを特徴とする。

【0034】

50

請求項20に記載の発明によれば、上記挙動比較手段が、第1マルウェア分析手段で検出したマルウェア本体のコードと、第2マルウェア分析手段で検出したマルウェア本体のコードとを、それぞれ要約処理し両者の同一性を照合することを特徴とする。

【発明の効果】

【0035】

本発明は、以上の構成を備えることにより、次の効果を奏する。

すなわち、本発明の特徴であるマルウェアのミクロ分析とマクロ分析の結果から相関度を求める際に従来のように走査のプロファイルと比較するだけでなく、脆弱性攻撃コード(Exploit code:脆弱性実証コードとも言う。)やマルウェア自体のコード、挙動をも比較対象とすることにより高精度の相関分析が実現できる。例えば、走査の挙動が類似した2つの異なるマルウェアに対しても、脆弱性攻撃コードやマルウェア自体のコード、挙動の相関度を求めることで正確な識別をおこなうことができる。

10

本発明では、走査に係る挙動(走査層)とマルウェア本体のコード及び挙動(マルウェア層)とを組み合わせ、走査層における検出にはより簡易なホストを用いてネットワーク上の広域に配備することを可能にすると共に、マルウェア層の検出には高精度なホストを設置し、さらに正確な検出に寄与することができる。

【0036】

特に、マルウェア層の検出用ホストは、実際にマルウェアに感染させて挙動を調べるために、十分に管理していないと他のホストを攻撃する恐れがある。そのため、極めて慎重な運用が必要であり、多くを配備することはできない。一方、走査層の検出用ホストは他のホストに攻撃または感染するような高度の能力を持たないため、感染する恐れがなく、大量に配備することができる。

20

【0037】

そして、走査層の検出用ホストを広域に配置することで、ネットワーク上の全体的なマルウェアの走査の傾向や分布を把握した上で、さらに、それぞれの走査が具体的にどのマルウェアによってもたらされたものであるかを高い精度で推定することができる。これによりネットワーク上で実際に活動中のマルウェアの種類や比率を推定することができる。

【0038】

また、脆弱性攻撃コードを比較する工程(脆弱性攻撃コード層)を設けることで、マルウェア層の検出ほど高度なホストを要さず、また感染の危険性を減らしながら、走査層のみによる相関分析に比して精度の高い検出を行うことができる。

30

このように各層の検出用ホストをネットワーク上に段階的に配備することで、高精度なマルウェアの検出に寄与することができる。

【発明を実施するための最良の形態】

【0039】

以下、本発明の実施形態を、図面に示す実施例を基に説明する。なお、実施形態は下記に限定されるものではない。

まず、本発明の概要を説明する。本発明は、正確かつ実践的な分析をすでに提案している技術に加え、マクロ・ミクロ相関分析の精度を大幅に高めるためにマルチレイヤの観察に基づいた新規の分析方法を提案するものである。

40

【0040】

本発明の概念では、マルウェアの活動はマルウェアの根本的な伝搬ステップに従って複数の層で監視、分析する。すなわち、マルウェアは最初にネットワーク上で走査することによってそのターゲットを探す。攻撃されやすいホストを見つけると、マルウェアはそのターゲット・ホストの権限を握るために脆弱性攻撃コードを送信する。その後、マルウェアはターゲット・システム上で自らをコピーし、次の伝搬に備える。

【0041】

このような特性に着目し、本発明ではこれらの3つのステップを対応する3つの層、すなわち走査層、脆弱性攻撃コード層、及びマルウェア層で観察、分析する。

その後、その3つの層におけるマクロ分析とミクロ分析両方からのすべての観察と分析

50

の結果が相関アナライザによって効果的に収集され、相互に関連付けられ、その結果ネットワーク攻撃とその原因を高い精度で効果的に特定することができる。

これによって取得されるマルウェアの内部の挙動及びネットワーク挙動は、マルウェア駆除ツールやIDS (Intrusion Detection System) のためのシグネチャを生成する等の、マルウェアのさらなる活動を軽減し、防止するための処置を講じるために活用できる。

【0042】

次にマルウェア活動のマルチレイヤ観察に基づいた新規の分析概念、マクロ-ミクロ相関分析について説述する。マクロ-ミクロ相関分析の基本的な考え方は、ネットワーク上で実際に起こっているマルウェア活動の観測をベースにした広域的な観察(マクロ分析)を、隔離された「実験室」で実行される詳細なマルウェア分析(ミクロ分析)の結果に結び付けることである。

10

相互関連の精度を高めるために従来システムを高度化するものとして、本発明のマルチレイヤ観察を提案する。

【0043】

すなわち、マルウェアの基本的な伝搬ステップに従って、マルウェア活動を走査層、脆弱性攻撃コード層、及びマルウェア層という3つの層で観察、分析する。その3つの層におけるマクロ分析とミクロ分析両方からのすべての観察と分析の結果は相関アナライザによって効果的に相互に関連付けられ、その結果ネットワーク攻撃とその根本的な原因ははるかに高い確実性をもって効果的に結び付けられる。

【0044】

20

ここで、図1には開発中のシステムの概要を説明する。本システムは、インフラストラクチャとエンドユーザの両方に多大な損害を生じさせることのあるワーム、ウイルス、及びボット等の伝搬するマルウェアを検出、分析することを目的として開発されている。

図1に示すように4つのサブシステム、つまりマクロ分析システム(MacS)(10)、ミクロ分析システム(MicS)(11)、ネットワーク・マルウェア関連付けシステム(Nemesis)(12)、及びインシデント処理システム(IHS)(13)から成り立っている。

【0045】

MacS(10)はインターネット(20)でのマルウェア活動の過程で走査等の巨視的な挙動を把握するためにネットワーク上で分散されたセンサ(21)・・・を配備する。本発明ではこのようなネットワークから検出される検査対象のマルウェアを第2のソフトウェアと呼んでいる。

30

本実施例では観察のために複数の/16ダークネットと/24ダークネットがあり、その中で入信パケットだけを検知する幅広い範囲のブラックホール・センサ、TCP SYNパケットとICMPエコー要求等の特定の入信パケットに対応する低対話型センサと、マルウェア・サンプル自体を含む多様な情報を引き出すために攻撃者と多種多様な対話が可能な高対話型センサ(ハニーポット)を配備している。

【0046】

図2では、2007年3月に発明者らによって/16ブラックホール・センサの1つによって観察された入信パケット数(実線)と一意のIPアドレスの数(点線)を示している。平均では、1日あたりほぼ350万個のパケット、及び15万以上の一意のIPアドレスが観察される。

40

これらのセキュリティ・イベントは追加分析のためにMacS(10)内の後述する多様なアナライザに送信される。MacS(10)のアナライザは、重大なセキュリティ・インシデントの初期的または予兆的事象であるインシデント候補(IC)を検出する。

例えば、変化点アナライザ(非特許文献26参照)は、特定のポート番号での走査頻度データ等の時系列データの急速な変化を検出し、IC警報(14)を発行する。IC警報(14)は、人間のオペレータが手動による詳細な分析を開始するようにIHS(13)に送信してもよい。

【非特許文献26】Takeuchi, J., Yamanishi, K.: Aunifying framework for detecting

50

outliers and change points from non-stationary time series data, IEEE Transactions on Knowledge and Data Engineering, Vol. 18, No. 4, pp.482 - 489, 2006年

【0047】

このIC警報(14)はNemesis(11)にも送信され、本発明に係る相関分析をトリガする。

一方、すべての分析結果(15)は相関分析のためにNemesis(11)に送信される。さらに、本システムは、人間のオペレータが観察されたトラフィックを直感的に理解するための複数のトラフィック・ビジュアライザを配備することもできる(トラフィック・ビジュアライザについては非特許文献23参照)。

【0048】

他方、MicS(12)は、微視的にマルウェアを分析する。MicS(12)では、後述する高対話型センサによって取り込まれる、あるいは入力として他のソース(22)から受け取られるマルウェア実行ファイル(23)を採取する。

予め分析しておくマルウェアを本発明では第1のソフトウェアと呼んでいる。従って、第1のソフトウェアは他のソース(22)から取得してもよいし、インターネット(20)から収集してもよい。

【0049】

次に、MicS(12)は、マルウェアの詳説された構造及び動作を明らかにするために、マルウェアの挙動分析とマルウェアのコードの分析によって該実行ファイル(23)を分析する。

MicS(12)のアナライザもIC警報(16)を発行することができる。例えば、アナライザは、未知の挙動のマルウェア検出時にそれを発行する。すべての分析結果(17)は相関分析のためにNemesis(11)に送信される。

【0050】

Nemesis(11)は、観察された攻撃をさらに正確なレベルで特定するために、MacS(10)とMicS(12)からの結果を相互に関連させる。Nemesis(11)は、MacS(10)とMicS(12)の両方から、分析結果(15)(17)を受け取る。それがMacS(10)またはMicS(12)からIC警報(14)(15)を受け取ると、MacS(10)で観察された第2のソフトウェアによる攻撃を、MicSで分析されたマルウェア(第1のソフトウェア)と結び付けるために相関分析を開始する。

【0051】

IHS(13)は、多様な分析結果とオペレータ(30)の間のインタフェースとなる。オペレータはIHS(13)を介してセキュリティ・イベントと分析結果を取り出すことができる。想定されるどのインシデントが検出され、より深く、あるいは手作業で分析されるべきかをオペレータが理解できるようにMacS(10)とMicS(12)の中のアナライザによって発行されるIC警報(14)(16)も管理する。

これは、オペレータ(30)に対してシステムの最終的な出力として発行されるインシデント・レポート(31)を作成するためのシステムでもある。

【0052】

セキュリティ・イベントの獲得、MacS(10)とMicS(12)それぞれにおける分析、及びNemesis(11)による相関分析を含むすべての手順が、想定されるセキュリティ・インシデントが迅速に検出、分析できるように、完全に自動化され、リアルタイムで処理される点は本発明の特徴である。

【0053】

前述したように、マルウェアにはその活動を特徴付ける複数の典型的な伝搬ステップがある。最初に、マルウェアは攻撃されやすいホストを求めてネットワーク上で走査する。攻撃されやすいホストを検出した後、マルウェアはそのターゲット・ホストの制御を奪取するために脆弱性攻撃コードを送信する。最後に、マルウェアは自らをターゲット・ホスト上でコピーし、さらなる伝搬に進む。

10

20

30

40

50

【0054】

本発明は、このようなマルウェアの特徴を捕らえて、実際のネットワークにおける活動をより高精度に検出できるように、マルウェア本体に加えて、マルウェアによる走査及び脆弱性攻撃コードの間のリンクを検出する。

そのために本発明ではマルウェア活動が3つの層、すなわち走査層、脆弱性実行コード層、マルウェア層で観察する新しい概念を提案する。

【0055】

図3には本発明に係るマルウェアの類似性検査システムの構成図を示す。図1に示した従来提案済みのシステムの各構成を基礎としているが、新たに各層ごとの分析技術を加えるための構成である。そのため図面中では符号を改めている。

MacS(40)では、走査及び脆弱性攻撃コードは、広範囲のネットワーク(例えばインターネット)(50)で分散されているブラックホール・センサ(51)と低対話型センサ(52)によってそれぞれ観察される。

【0056】

さらに、マルウェア・サンプルであるマルウェア実行ファイル(54)は高対話型センサ(53)によって取得される。その結果、それらに対応するアナライザ、つまり走査アナライザ(41)、脆弱性攻撃コード・アナライザ(42)、及びマルウェア・アナライザ(43)によって分析される。

分析結果はNemesis(60)の相関アナライザ(61)に通知される。

【0057】

なお、本実施例では、マルウェア・サンプルを第2のソフトウェアとしてマルウェア・アナライザ(43)に入力すると共に、第1のソフトウェアとして後述する設定可能サンドボックス等に送信している。(図中の通知線55)

【0058】

一方、MicS(70)では、設定可能サンドボックス(80)と呼ばれる新規の構成を用いることを特徴としている。サンドボックスとは、マルウェアの実行ファイルについて、その挙動の分析のために観察できるように実行される実験環境である。

設定可能サンドボックス(80)は、そのネットワーク環境が3つの焦点を当てられているネットワーク挙動、つまりブラックホール・ネットワークでの走査(ブラックホールモード)(81)、脆弱性攻撃コードの送信(低対話モード)(82)、それら自体のターゲット・ホストでのコピーを撤回するために適する高対話モード(83)の3つの実行モードを用いる。

【0059】

MacS(40)と同様に、取得された走査及び脆弱性攻撃コード、マルウェア本体は対応するアナライザ(71)(72)(73)によって分析され、分析結果はNemesis(60)の相関アナライザ(61)に送信される。

【0060】

Nemesis(60)では、相関アナライザは3つの層のMacS(40)とMicS(70)の両方からすべての分析結果を受け取る。受け取った分析結果はマルウェア知識プール(MNOP)と呼ばれる図示しない記憶手段のデータベースに記憶される。MNOP内で分析結果をリンクするために、相関アナライザ(61)が走査プロファイル(90)(91)、脆弱性攻撃コード・サンプル(92)(93)、及びマルウェア・サンプル(94)(95)のために後述する方法で相関関係を求める。

【0061】

このような相関関係の導出によって、相関アナライザ(61)はMacS(40)で観察された攻撃するマルウェアを、MicS(70)で分析されたマルウェアと結び付けることができる。

【0062】

以下、図3及び図4を用いて各層の分析手法を詳述する。
(走査層)

10

20

30

40

50

走査層の役割は、広範囲のブラックホール・ネットワークでマルウェアのネットワーク走査を分析することである。マルウェアによるネットワーク走査は広範囲のブラックホール・センサによって見られるときに特徴的である。

それらがどの宛先ポート上を走査するのか、それらがソースポート番号をどのようにして選ぶのか、それらがどれほど速く走査できるのか、それらがどのようにして宛先IPアドレスを選ぶのか等の走査挙動は、観察された走査の発生元を特定するために有用な情報である。したがって、実際のネットワークの監視によって観察されるものと比較するためにサンドボックス内のそれぞれの取り込まれたマルウェアの走査を観察し、要約する。

【0063】

本発明に係る第1アドレス走査情報検出手段は、設定可能サンドボックスのブラックホールモード(81)と走査アナライザ(71)とからなる。第2アドレス走査情報検出手段は、ブラックホール・センサ(51)と走査アナライザ(41)とからなる。

走査アナライザ(41)(71)は未処理のパケット・データを採取し、ソースIPアドレスでそれをスライスし、ソースIPアドレスごとにパケットの基本的な統計を計算する。

【0064】

ここでいう統計には、各宛先ポートのアクセス比率、連続するパケットのソースポート番号の間の差異の平均値、プロトコル(TCP/UDP/ICMP)とTCPフラグ(TCPを使用している場合)の比率、時間単位あたりのパケット平均数等を含むことができる。これらのいずれか、又は任意の組み合わせとしてもよい。

【0065】

さらに、公知の方法で、所定の分類規則のセットに基づいて走査挙動を複数のタイプに分類する。(非特許文献27参照)。

次に、走査タイプと宛先ポート・セットの文字列連結の要約計算方法(MD5:MessageDigest 5)である走査シグネチャが、走査を特定するために計算される。前記の基本的な統計と走査シグネチャの集合を走査プロファイルと呼ぶ。(非特許文献24に記載)

本発明のアドレス走査情報である走査プロファイルはXMLフォーマットで出力され、記憶手段に格納される。

【非特許文献27】Suzuki, K., Baba, S., Takakura, H.: Analyzing traffic directed to unused IP address blocks, IEICE Technical Report, vol.105, no.530, IA2005-23, pp.25 - 30, 2006年1月

【0066】

MacS(40)の走査層では、多様なブラックホール・センサ(51)を複数の/16ネットワークと/24ネットワークを含む広範囲のダークアドレス空間に展開する。入信トラフィックは走査アナライザ(41)に輸入され、各ソースIPアドレスの走査プロファイルがNemesis(60)に出力され、挙動比較手段である相関アナライザ(61)で分析される。

【0067】

MicS(70)の走査層では、マルウェア・サンプルはブラックホール・サンドボックス(81)で実行され、取り込まれたトラフィックは走査アナライザ(71)に輸入される。ブラックホール・サンドボックス(81)は、MacS(40)のブラックホール・センサ(51)と同じ挙動となる動作モードを実現している。

【0068】

すなわち実行されるマルウェアによって生成されるあらゆる走査にตอบสนองしないように構成され、その結果、結果として生じる走査プロファイルはMacS(40)で観察される走査プロファイルと比較できる。走査アナライザ(71)はNemesisの相関アナライザ(61)に対し、マルウェア実行ファイルごとの走査プロファイルを出力する。

【0069】

(脆弱性攻撃コード層)

脆弱性攻撃コード層の役割は、各マルウェアがどのような種類の脆弱性攻撃コードを使

10

20

30

40

50

用するのかを探査することである。脆弱性攻撃コードは、ターゲット脆弱性が悪用された後に実行されるシェルコードを含んでいる。

そこで各分析システム(40)(70)の脆弱性攻撃コード・アナライザ(42)(72)は、ターゲット脆弱性を活用するために必須であるシェルコード内の命令のシーケンスを検出する。

【0070】

これらの命令はシェルコードの識別子として抽出され、使用される。脆弱性攻撃コード層では、実際のネットワーク監視によって観察されるものと比較するためにサンドボックス内の各マルウェアによって生成される攻撃パケットのペイロードでこれらの命令を観察し、抽出する。

10

【0071】

本発明に係る第1脆弱性攻撃コード検出手段は、設定可能サンドボックスの低対話モード(82)と脆弱性攻撃コード・アナライザ(72)とからなる。第2脆弱性攻撃コード検出手段は、低対話型センサ(52)と脆弱性攻撃コード・アナライザ(42)とからなる。

【0072】

脆弱性攻撃コード・アナライザ(42)(72)における脆弱性攻撃コード検出技法は、IDSを使用することについて公知の技術がある。最近の研究では、自己書き換えをおこなうポリモーフィック型の脆弱性攻撃コードの検出及び分類を可能にしている(非特許文献28参照)。

20

【非特許文献28】Payer, U., Teufel, P., Lamberger, M.: Hybrid engine for polymorphic shellcode detection, Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA2005), LNCS 3548, pp 19 - 31, 2005年

【0073】

さらに本実施例では高速処理を達成するために、ASHULAと呼ばれている最先端の脆弱性攻撃コード検出技術を活用することもできる。(非特許文献29参照)。

ASHULAはパケットのペイロードから遠隔システムの権限を握るために必須である前述された命令のシーケンスを検出し、抽出する。この命令は脆弱性攻撃コードの実際の一部であるため、脆弱性攻撃コードを検出するためのパターン・マッチングのためのIDSシグネチャとして使用することもできる。

30

【非特許文献29】Nogawa, H.: Shellcode detection: fight against polymorphism, The Joint Information Security Workshop on Internet Monitoring and Analysis (ISWIMA), 2006年

【0074】

MacS(40)の脆弱性攻撃コード層では、複数の/24ネットワークのダークアドレス空間に低対話型センサ(52)を配備する。多くのマルウェアはそのターゲットの権限を握るために脆弱性攻撃コードを送信する前に、走査し、TCP接続を確立する等の特定のステップに従う。

【0075】

そこで本発明の低対話型センサ(52)はマルウェアから脆弱性攻撃コードを引き出すために入信攻撃に正しく反応するように構成されている。低対話型センサ(52)への入信トラフィックはソースIPアドレスでスライスされ、その後スライスされたトラフィックのそれぞれが脆弱性攻撃コード・アナライザ(42)に入力される。アナライザ(42)はソースIPアドレスごとに抽出された命令のリストを出力する。

40

【0076】

一方、MiCS(70)の脆弱性攻撃コード層では、マルウェア実行ファイルは低対話モードサンドボックス(82)で実行され、取り込まれたトラフィックは脆弱性攻撃コード・アナライザ(72)に入力される。

低対話モードサンドボックス(82)は、実行されるマルウェアからの要求に適切に答えるように構成され、それはMacS(40)内の低対話型センサ(52)と同じ挙動で

50

ある。脆弱性攻撃コード・アナライザ(72)はマルウェア実行ファイルごとに抽出された命令のリストを出力する。

【0077】

(マルウェア層)

マルウェア伝搬の最終段階では、マルウェアのコピーをターゲット・ホスト上でダウンロードし、ホストがリブートされる時にコピーを確実に実行するようにしている。コピーは、FTP及びHTTPのようなプロトコルを使用して特定のサーバから、あるいは以前に感染したホストからダウンロードされることもある。マルウェア層ではマルウェア・アナライザ(43)(73)によってサンプル自体を分析する。

【0078】

本発明に係る第1マルウェア分析手段は、設定可能サンドボックスの高対話モード(83)とマルウェア・アナライザ(73)とからなる。第2マルウェア分析手段は、高対話型センサ(53)とマルウェア・アナライザ(43)とからなる。

【0079】

本実施例では、マルウェア本体のコードを分析するマルウェア・コード・アナライザと、マルウェアによる挙動を分析するマルウェア挙動アナライザの2つのアナライザをマルウェア・アナライザ(43)(73)に備えている。

【0080】

マルウェア・コード・アナライザは実行ファイルを逆アセンブルし、その内部の特徴と構造を明らかにするために使用される。もっとも多くのマルウェアはASPack、FSG、Petite、UPX等の多様なパッキング技術を使用して難読化されている。

難読化のために、多くのマルウェアは従来の方法では逆アセンブルできない。難読化を克服するために、マルウェア・コード・アナライザはまず、隔離されたホストでマルウェア・サンプルを実行する。

【0081】

大部分のマルウェアはアクティブとなるためにDLLをロードまたはアンロードするときにメモリ上で自らを解凍(難読化の解除)するため、マルウェア・コード・アナライザは、サンプルが動的リンク・ライブラリ(DLL)をロードまたはアンロードするたびにホストのメモリをダンプする。

【0082】

次に、マルウェア・コード・アナライザはダンプされたコードのPEファイルヘッダのいくつかのフィールド(例えばPointerToRawData、SizeOfData等)を適切に上書きする。

マルウェア・コード・アナライザは、最終的にはダンプされたコードを逆アセンブルし、アセンブリコードを取得できる。アセンブリコードを読み取ることによって、Windows(登録商標)APIの呼出シーケンスを取得する。

【0083】

同時に、マルウェア・コード・アナライザは、作成・修正されたURLまたはIPアドレスであるファイルとレジストリを抽出してもよい。

最後に、すべての情報はコード分析レポートとして要約、XMLフォーマットで出力される。

【0084】

他方、マルウェア・アナライザ(43)(73)のうち、マルウェア挙動アナライザは実際のインターネットに配備されている高対話型センサ(53)におけるマルウェア実行から、あるいは実験室での設定可能サンドボックス(83)から取得される多様なログを分析する。

【0085】

ログは、ファイル/レジストリ・アクセス・ログ及びAPIログを被害ホストの中に、サーバ・アクセス・ログを複数のタイプのサーバの中に、あるいはマルウェア・サンプルによって生成されるパケットログを含む。

10

20

30

40

50

これらのログに従って、マルウェア挙動アナライザは、所定の挙動定義に基づいたマルウェア・サンプルの顕著な挙動を抽出する。抽出された挙動は挙動分析レポートとして要約され、XMLフォーマットで出力される。

【0086】

MacS(40)のマルウェア層では、高対話型センサ(53)がダウンロードされた実行ファイルを取り込むために配備されている。高対話型センサ(53)には、それらが攻撃ホストから容易に発見され、感染し、最終的にダウンロードされるサンプルを取り込むように、脆弱性が回復されていないホストや、あるいは仮想のWindows(登録商標)を備えたシステムを用いることができる。

【0087】

マルウェア・サンプルの収集のためには、公知のnepenthes(非特許文献29)及びArgos(非特許文献30)のようなツールも使用できる。

【非特許文献29】P. Baecher, M. Koetter, T. Holz, M. Dornseif, F. C. Freiling, "The Nepenthes Platform: An Efficient Approach to Collect Malware," Proceedings of Recent Advances in Intrusion Detection, 9th International Symposium, RAID 2006, pp. 165-184, 2006年

【非特許文献30】G. Portokalidis, A. Slowinska, and Herbert Bos, "Argos: an emulator for fingerprinting zero-day attacks for advertised honeypots with automatic signature generation," Proceedings of the 2006 EuroSys conference, pp. 15 - 27, 2006年

【0088】

MacS(40)は、マルウェア・サンプルを取り込むために脆弱なWindows(登録商標)システムの複数の高対話型センサ(53)だけではなく、/24ネットワークにもnepenthesを配備することができる。MD5について新しいサンプルがセンサ(53)によって取り込まれると、それは追加の分析の為にマルウェア・コード・アナライザとマルウェア挙動アナライザ(あわせて43)に送信される。

【0089】

一方、サンプルは、MicS(70)が3つの層でサンプルの分析を開始できるようにMicS(70)の設定可能なサンドボックス(80)に送信される。

最後に、各マルウェアのコード分析レポート及び挙動分析レポートからなるマルウェア・レポート(94)が相関アナライザ(61)に送信される。

【0090】

MicSのマルウェア層では、マルウェア・サンプルは高対話モードのサンドボックス(83)で実行される。高対話型サンドボックス(83)は、新しいサンプルをダウンロードできる別の被害ホストを提供する。ダウンロードされるサンプルがMD5について新しい場合、それはマルウェア・コード・アナライザとマルウェア挙動アナライザ(あわせて73)に送信される。最後に、各マルウェアのコード分析レポート及び挙動分析レポート(あわせてマルウェア・レポート95)が相関アナライザ(61)に送信される。

【0091】

設定可能サンドボックス(80)についてさらに詳述する。本発明で創出した設定可能サンドボックス(80)は、マルウェア活動の多様な観察結果を提供する重要な技術である。

マルウェア・サンプルが入力されると、ファイル・アクセス、レジストリ・アクセス、及び通信を含むそれらの挙動を観察するために実験環境での入力サンプルを実行する。

【0092】

最新のマルウェアは、それらが実験環境で実行されないことを確認するために様々な手法をとっている。これらの分析は多くの場合、便宜上仮想ホストで行われるためにホストOSとゲストOS間のチャンネルを検出することによって仮想OSを検出する他のマルウェアがある一方、インターネットの到達可能性をチェックするために特定のウェブページにアクセスしようとするものもある。

10

20

30

40

50

【 0 0 9 3 】

したがって、本発明の設定可能サンドボックス（80）はAPIフッキング機構及びファイル/レジストリ監視機構を備える被害ホストのために実際のホストを使用する。

設定可能サンドボックス（80）は、マルウェアを欺くためにインターネット・エミュレータと呼ばれる環境も使用する。インターネット・エミュレータは被害ホストに仮想インターネット・アクセスを提供する。

【 0 0 9 4 】

該インターネット・エミュレータは、入力マルウェアが実際のインターネット環境であると信じ、伝搬を開始するように、DNS、FTP、TF TP、HTTP、SMTP及びIRC等の多様なサービスを実装している。

マルウェアによって生成されるすべてのパケットはインターネット・エミュレータに送られ、次にインターネット・エミュレータはそのパケットをそのポート番号の代わりにそのペイロードの観察に従って適切なサービスに転送する。

【 0 0 9 5 】

これは、マルウェアが特にHTTPサービス及びIRCサービス低対話モードサンドボックス（82）に対して不規則なポート番号を使用することがあるためである。設定可能サンドボックス（80）は、3つのモード、つまりブラックホールモード（81）、低対話モード（82）、及び高対話モード（83）で動作する。

【 0 0 9 6 】

ブラックホール・サンドボックスとも呼んでいるブラックホール・モード（81）では、サンドボックス（80）は、Mac S（40）でのブラックホール監視に類似した状況を生じさせるために最小のインターネット・サービスを提供する。

【 0 0 9 7 】

被害ホストから感染の結果生じるトラフィックはグローバル・アドレスでの走査、ローカル・アドレスでの走査、特定のサーバに対するアクセス等に分けられる。

被害ホストが実際のインターネット環境にあると、ブラックホール・センサによってグローバル・アドレスでの走査だけが監視できる。そのため、グローバル走査のスライスされたトラフィックが走査プロファイルを取得するために走査アナライザに送信される。

【 0 0 9 8 】

低対話型サンドボックスとも呼ぶ低対話モード（82）では、サンドボックス（80）は、脆弱性攻撃を観察するために低相互作用監視に類似した状況を作成し、実行されている被害ホストの要求に適切に答えるように構成される。

被害ホストからの結果として生じるトラフィックは、どの脆弱性攻撃コードが入力マルウェアから送達されるのかをチェックするために脆弱性攻撃コード・アナライザ（72）に入力される。

【 0 0 9 9 】

高対話型サンドボックスとも呼ぶ高対話モード（83）では、サンドボックス（80）は、感染しやすいホストが存在する環境をエミュレートするように構成されている。

例えば、上述したArgosまたはnepenthesのようなツールを、攻撃されやすいホストをエミュレートするために設定可能サンドボックス（80）にインストールすることができる。

【 0 1 0 0 】

サンプルをダウンロードするために、設定可能サンドボックス（80）は実際のインターネットに接続される必要がある。したがって、サンプルをダウンロードするための必要なトラフィックだけがサンドボックス（80）から確実に出るようにするためにアクセス制御は重要である。

【 0 1 0 1 】

MD5について新しいマルウェアが取り込まれるときには、それはMic S（70）のマルウェア・アナライザ（73）に出力されるだけでなく、設定可能サンドボックス（80）にも再帰的に入力される。

10

20

30

40

50

【0102】

次に、本発明の挙動比較手段である相関アナライザ(61)について説述する。

相関アナライザ(61)は、MacS(40)とMicS(70)のすべてのアナライザから分析結果を受け取り、それらをNemSys(60)の図示しない外部記憶手段に格納されたマルウェア知識プール(MNOP)と呼ばれるデータベースに記憶する。相関アナライザ(61)の役割は、それらの間でリンクを検出することによりそれらを統合し、充実させるために種々の分析結果をリンクすることである。

【0103】

図4に示すように、各アナライザからの分析結果の主要な属性を要約する。この図は、さまざまなアナライザからの分析結果が、それら自体を相互に関連付けるために使用できる多くの共通の属性を有していることを示している。以下では、これらの共通の属性における相関分析の手法を説明する。

10

【0104】

(走査層における相関分析)

走査を相互に関連付けるための最も簡単な方法は、プロファイルの中の走査シグネチャを比較することである。走査シグネチャは非特許文献27に開示されるような規則によって示される走査タイプと宛先ポートのセットの連結のMD5を用いるのが簡便である。

したがって、2つの走査プロファイルのシグネチャが正確に一致する場合、それは2つの走査が同じタイプであり、それらが同じポート番号に向けられていることを意味する。

【0105】

20

さらに厳密な方法は、2つの走査間の類似性を定義することである。走査プロファイルは多様な静的パラメータを含むため、それらを使用してその類似性を定義することができる。例えば、走査プロファイルは各宛先ポートに対するアクセスの比率を含む。2つのプロファイルのこれらの比率は、相関係数を使用することによって比較できる。単純な統計を使用して走査を相互に関連付けるため技術は非特許文献24のように公知である。

【0106】

大きなデータセットから検索する必要があるときには、シグネチャ・マッチングは非常に高速かつ有用であるが、それは「類似する」走査を検出できない。実際に、3つの宛先ポート139、445及び1025での走査及び2つのポート139と1025上で走査のためのシグネチャは、たとえそれらがいくらか類似し、関連付けられていても、別のシグネチャを有する。

30

【0107】

ブラックホール・センサ(51)は多くの場合単一の攻撃者から走査パケット全体の一部だけを観察するため、前者の方法では正確に同一性を照合できない場合がある。他方、類似性の定義による相関分析は、それらに対する検索の複雑度が定義に応じて高い場合にも類似した走査を検出できる可能性が高い。

【0108】

(脆弱性攻撃コード層における相関分析)

脆弱性攻撃コードの相関分析は、抽出される命令を比較することによって行うことができる。走査の相互関連のケースに関しては、要約マッチングが最も容易な方法である。命令の2つのシーケンス間の類似性を測定するための方法は任意であるが、例えば上記のようにMD5を用いても同一性を照合してもよい。

40

【0109】

(マルウェア層の相関分析)

2つのマルウェア・サンプルを比較する最も簡単な方法は、その要約(例えばMD5)によるものである。別の簡単な方法は、既知のマルウェアにとって効果的な方法であるが、周知のアンチウィルス・ソフトウェアを用いる方法である。特に簡便で有効な方法は、それらの分析レポートを比較することである。

【0110】

マルウェア・コード・アナライザ及びマルウェア挙動アナライザは、API、その出現

50

順、変更されたレジストリ・キーとファイル、それらのミューテックスの名前、それらがアクセスを試みるサーバ名等を含むマルウェアについての情報を検出可能である。

上述した走査層における相互関連と同様に、マルウェア・アナライザ(43)(73)の分析結果を使用してそれらの間の類似性を算出することができる。

【図面の簡単な説明】

【0111】

【図1】従来提案されているマクロ分析・ミクロ分析の相関分析を行うシステムの構成図である。

【図2】ブラックホールセンサによる測定結果例のグラフである。

【図3】本発明に係るマルウェアの類似性検査システムの構成図である。

10

【図4】本発明に係るマルウェアの類似性検査システムのリンクを示す図である。

【符号の説明】

【0112】

40 マクロ分析システム

41 走査アナライザ

42 脆弱性攻撃コード・アナライザ

43 マルウェア・アナライザ

50 インターネット

51 ブラックホール・センサ

52 低対話型センサ

20

53 高対話型センサ

60 ネットワーク及びマルウェア関連付けシステム

61 相関アナライザ

70 ミクロ分析システム

71 走査アナライザ

72 脆弱性攻撃コード・アナライザ

73 マルウェア・アナライザ

80 設定可能サンドボックス

81 同、ブラックホールモード

82 同、低対話モード

30

83 同、高対話モード

90 走査プロファイル

91 走査プロファイル

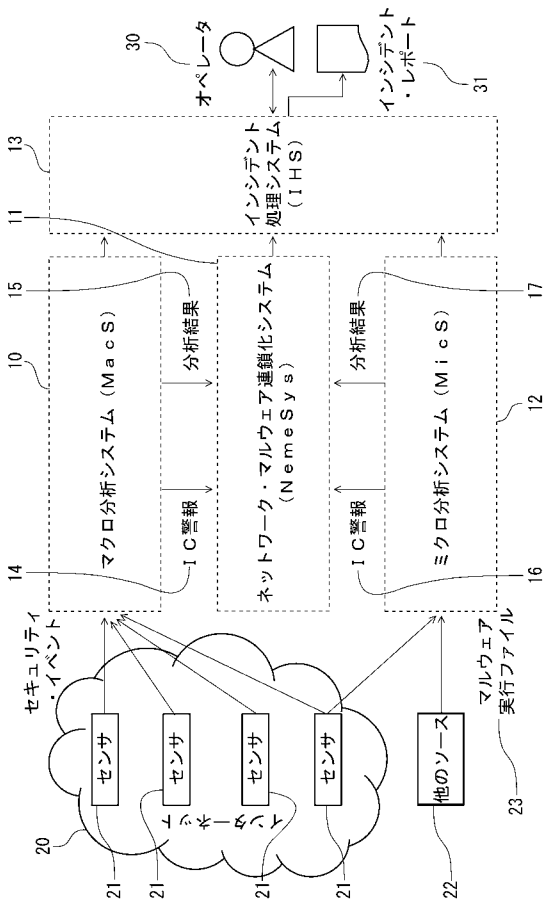
92 脆弱性攻撃コード

93 脆弱性攻撃コード

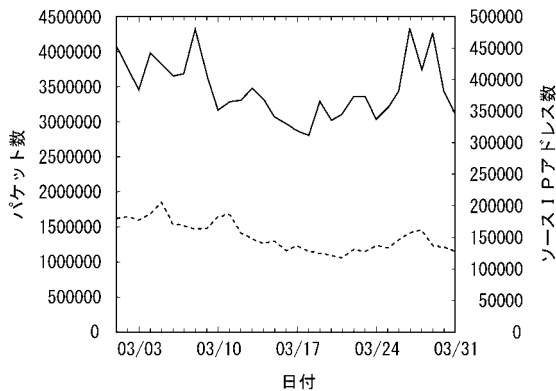
94 マルウェア・レポート

95 マルウェア・レポート

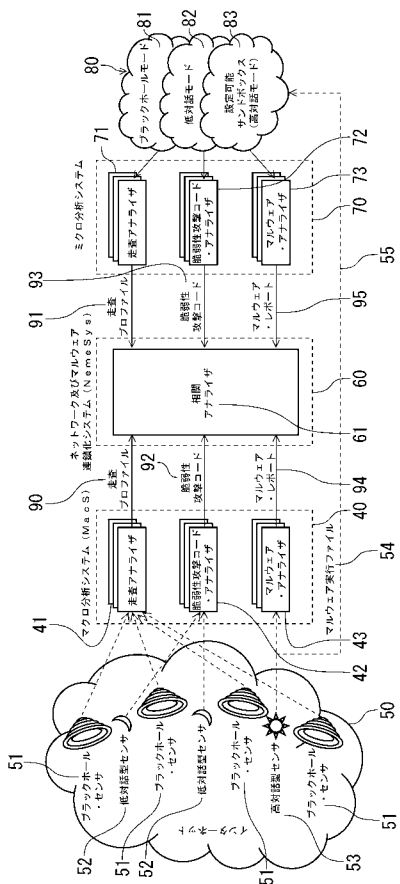
【図1】



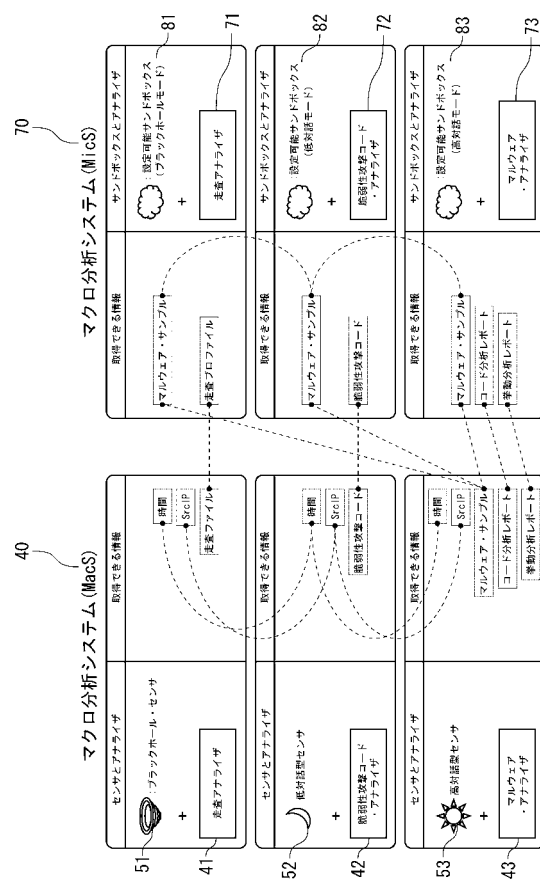
【図2】



【図3】



【図4】



フロントページの続き

(72)発明者 衛藤 将史

東京都小金井市貫井北町4 - 2 - 1 独立行政法人情報通信研究機構内

審査官 市川 武宜

(56)参考文献 国際公開第2004/075060(WO, A1)

特開2005-109847(JP, A)

特開2007-058514(JP, A)

特開2005-236863(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 21/00