

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7301181号
(P7301181)

(45)発行日 令和5年6月30日(2023.6.30)

(24)登録日 令和5年6月22日(2023.6.22)

(51)国際特許分類

FI

G 0 6 F 21/55 (2013.01)

G 0 6 F 21/55 3 2 0

G 0 6 F 21/56 (2013.01)

G 0 6 F 21/56

請求項の数 8 (全21頁)

(21)出願番号	特願2022-4041(P2022-4041)	(73)特許権者	517356106
(22)出願日	令和4年1月14日(2022.1.14)		村上 真梨
(62)分割の表示	特願2017-197520(P2017-197520)の分割		神奈川県横浜市磯子区汐見台 2-2-2 -427
原出願日	平成29年10月11日(2017.10.11)	(74)代理人	110002815
(65)公開番号	特開2022-50617(P2022-50617A)		I P T e c h 弁理士法人
(43)公開日	令和4年3月30日(2022.3.30)	(72)発明者	アパリシ ロマン
審査請求日	令和4年2月7日(2022.2.7)		神奈川県横浜市磯子区汐見台 2-2-2 -427
		審査官	打出 義尚

最終頁に続く

(54)【発明の名称】 検出方法、検出プログラムおよび検出装置

(57)【特許請求の範囲】

【請求項1】

第1のコンピュータと、他のコンピュータと通信する際に前記第1のコンピュータを経由して通信を行う第2のコンピュータと、を含んで構成されるネットワーク内において、前記ネットワークに含まれる第3のコンピュータが、

前記第1のコンピュータが中継した、前記第2のコンピュータが通信元又は通信先となった通信の通信記録に含まれる、通信元の識別情報および通信先の識別情報の組ごとに、通信が行われた時間間隔を算出し、

前記組ごとに算出した各時間間隔の個数を計数し、

前記組ごと、前記計数した各時間間隔の個数の総和に対する時間間隔 t の個数の割合 $p(t)$ に基づいて、時間間隔のエントロピーを求め、

前記組ごとに算出した時間間隔のエントロピーの平均及び標準偏差を用いた閾値と、前記組ごとに算出した時間間隔のエントロピーの大きさを比較することにより、前記第2のコンピュータにおいてマルウェアによる通信が行われたかどうかを検出する

ことを含み、

前記通信記録は、前記第3のコンピュータが前記第1のコンピュータから取得することにより、前記第3のコンピュータに保持されたものである

検出方法。

【請求項2】

前記エントロピーの大きさとして、 $-p(t) \log b p(t)$ についてのすべての時

間隔 t に関する和に基づく値を算出する（ただし、 b は対数関数 $\log$ の底である）、
請求項 1 に記載の検出方法。

【請求項 3】

第 1 のコンピュータと、他のコンピュータと通信する際に前記第 1 のコンピュータを経由して通信を行う第 2 のコンピュータと、を含んで構成されるネットワーク内において、前記ネットワークに含まれる第 3 のコンピュータに、

前記第 1 のコンピュータが中継した、前記第 2 のコンピュータが通信元又は通信先となった通信の通信記録に含まれる、通信元の識別情報および通信先の識別情報の組ごとに、通信が行われた時間間隔を算出するステップと、

前記組ごとに算出した各時間間隔の個数を計数するステップと、

前記組ごと、前記計数した各時間間隔の個数の総和に対する時間間隔 t の個数の割合 $p(t)$ に基づいて、時間間隔のエントロピーを求めるステップと、

前記組ごとに算出した時間間隔のエントロピーの平均及び標準偏差を用いた閾値と、前記組ごとに算出した時間間隔のエントロピーの大きさを比較することにより、前記第 2 のコンピュータにおいてマルウェアによる通信が行われたかどうかを検出するステップと、
を実行させ、

前記通信記録は、前記第 3 のコンピュータが前記第 1 のコンピュータから取得することにより、前記第 3 のコンピュータに保持されたものである、

検出プログラム。

【請求項 4】

前記エントロピーの大きさとして、 $-p(t) \log b p(t)$ についてのすべての時間間隔 t に関する和に基づく値を算出するステップ（ただし、 b は対数関数 $\log$ の底である）

を実行させる、請求項 3 に記載の検出プログラム。

【請求項 5】

第 1 のコンピュータと、他のコンピュータと通信する際に前記第 1 のコンピュータを経由して通信を行う第 2 のコンピュータとを含んで構成されるネットワークに含まれる検出装置であって、

記憶部を含み、

前記第 1 のコンピュータが中継した、前記第 2 のコンピュータが通信元又は通信先となった通信の通信記録に含まれる、通信元の識別情報および通信先の識別情報の組ごとに、通信が行われた時間間隔を算出するステップと、

前記組ごとに算出した各時間間隔の個数を計数するステップと、

前記組ごと、前記計数した各時間間隔の個数の総和に対する時間間隔 t の個数の割合 $p(t)$ に基づいて、時間間隔のエントロピーを求めるステップと、

前記組ごとに算出した時間間隔のエントロピーの平均及び標準偏差を用いた閾値と、前記組ごとに算出した時間間隔のエントロピーの大きさを比較することにより、前記第 2 のコンピュータにおいてマルウェアによる通信が行われたかどうかを検出するステップと、
を有し、

前記通信記録は、前記第 1 のコンピュータから取得することにより、前記記憶部に保持されたものである、

検出装置。

【請求項 6】

前記エントロピーの大きさとして、 $-p(t) \log b p(t)$ についてのすべての時間間隔 t に関する和に基づく値を算出するステップ（ただし、 b は対数関数 $\log$ の底である）

を実行させる、請求項 5 に記載の検出装置。

【請求項 7】

前記時間間隔を算出するステップにおいて、ホワイトリストに含まれる識別情報を含む通信記録を除外し、前記時間間隔を算出する、

10

20

30

40

50

請求項 5 又は請求項 6 に記載の検出装置。

【請求項 8】

前記時間間隔を算出するステップにおいて、通信元の識別情報および通信先の識別情報の組ごとに通信の時間間隔が所定の値以下の時間間隔を除外し、前記時間間隔を算出する、請求項 5～請求項 7 の何れかに記載の検出装置。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、セキュリティ脅威の検出方法、検出プログラムおよび検出装置に関する。

【背景技術】

【0002】

コンピュータシステムのセキュリティ脅威の一つとしてマルウェアが知られている（例えば、特許文献 1 参照。）。マルウェアがユーザのコンピュータ内で動作すると、例えばユーザのコンピュータが当該マルウェアを介して外部の不正攻撃者により操作自在となり、ユーザのコンピュータに記憶されているデータを流出させたり、ユーザのコンピュータを介して他のコンピュータを攻撃させたりするなど、ユーザのコンピュータに不正な動作を行なわせてしまう。

【0003】

また、未知のマルウェアが日々出現しているため、ユーザのコンピュータに格納されているファイルをスキャンしても、未知のマルウェアの検出がされるまで数か月かかってしまう場合があり、手遅れとなってしまう。そもそも最近では、ファイルレス（fileless）のマルウェアも出現しており、ファイルのスキャンによってマルウェアを検出することは困難になってきている。

【先行技術文献】

【特許文献】

【0004】

【文献】米国特許公報 8, 578, 493 号明細書

【発明の概要】

【発明が解決しようとする課題】

【0005】

以上に例示した課題を解決することを目的の一つとして、本開示では、未知マルウェアであっても検出を行なうことができる検出方法、検出プログラムおよび検出装置を提供する。

【課題を解決するための手段】

【0006】

本開示の一態様に係る検出方法は、通信記録に含まれる通信元の識別情報および通信先の識別情報の組ごとに通信が行われた時間間隔を算出し、組ごとに算出した時間間隔の規則性に基づいて、マルウェアによる通信が行われたかどうかを検出することを含む。

【0007】

本開示の一態様に係る検出プログラムは、コンピュータに、通信記録に含まれる通信元の識別情報および通信先の識別情報の組ごとに通信が行われた時間間隔を算出し、組ごとに算出した時間間隔の規則性に基づいて、マルウェアによる通信が行われたかどうかを検出させるための、検出プログラムである。

【0008】

本開示の一態様に係る検出装置は、通信記録に含まれる通信元の識別情報および通信先の識別情報の組ごとに通信が行われた時間間隔を算出する間隔算出部と、間隔算出部により組ごとに算出された時間間隔の規則性に基づいて、マルウェアによる通信が行われたかどうかを検出する検出部と、を備える。

【発明の効果】

【0009】

本開示によれば、未知マルウェアであっても検出を行なうことができる。

【図面の簡単な説明】

【0010】

【図1】本開示の一実施形態に係る検出装置が使用される全体システムの構成の一例図である。

【図2】本開示の一実施形態に係る検出装置が読み込む通信ログの一例図である。

【図3】本開示の一実施形態に係る検出装置の機能ブロック図の一例図である。

【図4】本開示の一実施形態に係る検出装置が、時間間隔を算出するために使用するデータ構造の一例と、通信ログの読み込みにより当該データ構造が変化する様子の一例と、を示す図である。

【図5】本開示の一実施形態に係る検出装置の処理のフローチャートの一例図である。

【図6】本開示の一実施形態に係る検出装置が、時間間隔の規則性を検出するために使用するデータ構造の一例図である。

【図7】本開示の一実施形態に係る検出装置の処理のフローチャートの一例図である。

【図8】本開示の一実施形態に係る検出装置の全体処理のフローチャートの一例図である。

【図9】本開示の一実施形態に係る検出装置を実現する計算機の構成の一例図である。

【図10】本開示の一実施形態に係る検出装置の機能ブロック図の一例図である。

【図11】本開示の一実施形態に係る検出装置の機能ブロック図の一例図である。

【図12】本開示の一実施形態に係る検出装置の処理のフローチャートの一例図である。

【図13】実際の通信ログに基づいて作成したエントロピー分布のヒストグラムである。

【図14】実際の通信ログに基づいて、通信の時間間隔のエントロピーと通信回数とをプロットしたグラフである。

【発明を実施するための形態】

【0011】

(全体構成)

図1は、本開示の一実施形態に係る検出装置を含む全体システムの構成の一例である。全体システム1は、ネットワークNWと、クラウドネットワークCLと、ユーザのコンピュータA(2)と、ユーザのコンピュータB(3)と、ユーザのコンピュータC(4)攻撃者41のコンピュータ5を含む。なお、コンピュータA(2)、コンピュータB(3)およびコンピュータC(4)のユーザは同じであってもよいし、異なってもよい。

【0012】

ネットワークNWとクラウドネットワークCLとは、一体のネットワークであってもよく、例えば、ネットワークNWとクラウドネットワークCLとによりインターネットを構成していてもよい。あるいは、クラウドネットワークCLがインターネットを構成していてもよく、この場合には、ネットワークNWは、コンピュータA(2)およびコンピュータB(3)が設置されている組織内のLAN(Local Area Network)であってもよい。

【0013】

ネットワークNWとクラウドネットワークCLとは、多数のコンピュータが接続される。図1においては、コンピュータA(2)は、ネットワークNWに接続され、コンピュータB(3)は、コンピュータA(2)を介してネットワークNWに接続される。コンピュータC(4)は、コンピュータB(3)に接続され、コンピュータC(4)は、コンピュータB(3)およびコンピュータA(2)を介してネットワークNWに接続される。また、不正なプログラムを含むマルウェアを、コンピュータA(2)やコンピュータB(3)に実行させ不正な動作をさせようとする攻撃者41のコンピュータ5が、図1においてはクラウドネットワークCLに接続されている。なお、攻撃者41が、コンピュータA(2)またはコンピュータB(3)が設置されている組織内に存在する場合には、攻撃者41のコンピュータ5は、ネットワークNWに直接接続されている場合もあり得る。また、コンピュータC(4)がコンピュータB(3)に接続される前にマルウェアに感染し、コンピュータB(3)に接続される場合もあり得る。

【0014】

図1において、攻撃者41がコンピュータA(2)およびコンピュータB(3)の少なくとも一方にマルウェアを実行させた場合、また、コンピュータC(4)にすでにマルウェアに感染しマルウェアが実行される場合、そのマルウェアはコンピュータ5と通信を行なうことを想定する。別言すれば、コンピュータA(2)やコンピュータB(3)、コンピュータC(4)がマルウェアを実行すると、マルウェアは、その存在を攻撃者41のコンピュータ5に知らせ、あるいは、不正な指令を受信するために、攻撃者41のコンピュータ5に所定のデータを送信することを想定する。このような想定下において、マルウェアによるデータの送信を「ビーコン」という場合がある。

【0015】

コンピュータA(2)およびコンピュータB(3)の少なくとも一台には、開示の一実施形態に係る検出装置100-1および100-2が備えられていてもよい。図示はされていないが、コンピュータC(4)にも検出装置が備えられていてもよい。あるいは、コンピュータA(2)およびコンピュータB(3)の外部に検出装置100-3が備えられていてもよい(以後、検出装置100-1、検出装置100-2および検出装置100-3を総称して検出装置100と記す場合がある。)

【0016】

検出装置100は、コンピュータA(2)およびコンピュータB(3)の少なくとも1台に備えられている場合には、コンピュータA(2)およびコンピュータB(3)の少なくとも1台の内部で動作するプログラムとして実現することが可能である。あるいは、検出装置100は、コンピュータA(2)およびコンピュータB(3)の少なくとも1台のハードウェアの一部となってもよい。検出装置100が、コンピュータA(2)、コンピュータB(3)およびコンピュータC(4)の外部に備えられている場合には、独立した装置として実現することが可能であり、また、コンピュータを検出装置100として動作させるためのプログラムを当該コンピュータに動作させることによっても実現可能である。

【0017】

また、コンピュータA(2)およびコンピュータB(3)のそれぞれは、通信ログL1およびL2を記憶することができる。ここでいう通信ログは、通信記録の集合である。通信記録は、通信日時、当該通信の通信元(送信元)の識別情報および通信先(送信先)の識別情報を含む情報である。通信ログの通信記録は、通信日時により順序付けることができる。また、短時間に同じ通信元から通信先に複数回の通信が行われた場合、同じ内容の通信記録が通信ログに含まれる場合もあり得る。

【0018】

コンピュータB(3)が、ネットワークNWまたはクラウドネットワークCLに接続されているコンピュータやコンピュータC(4)へデータを送信すると、通信ログL2には、データの送信日時を通信日時とし、コンピュータB(3)の識別情報を通信元の識別情報とし、ネットワークNWまたはクラウドネットワークCLに接続されているコンピュータやコンピュータC(4)の識別情報であって、コンピュータB(3)のデータを受信するコンピュータの識別情報を通信先として含む通信記録が通信ログL2に追加される。また、逆に、ネットワークNWまたはクラウドネットワークCLに接続されているコンピュータやコンピュータC(4)が送信したデータをコンピュータB(3)が受信すると、データの受信日時を通信日時とし、コンピュータB(3)が受信したデータを送信したネットワークNWまたはクラウドネットワークCLに接続されているコンピュータやコンピュータC(4)の識別情報を通信元の識別情報とし、コンピュータB(3)の識別情報を通信先の識別情報として含む通信記録が通信ログL2に追加される。

【0019】

同様に、コンピュータA(2)を通信元または通信先とする通信の通信記録が通信ログL1に追加される。また、コンピュータB(3)やコンピュータC(4)とネットワークNWまたはクラウドネットワークCLに接続されているコンピュータとがコンピュータA

10

20

30

40

50

(2) を経由してデータの送受信を行なうとすると、当該送受信の通信記録が通信ログ L1 に記録される。例えば、コンピュータ B (3) が、ネットワーク NW またはクラウドネットワーク CL に接続されているコンピュータへデータを送信すると、通信ログ L1 には、コンピュータ B (3) の識別情報を通信元の識別情報とし、ネットワーク NW またはクラウドネットワーク CL に接続されたコンピュータの識別情報であって、コンピュータ B (3) のデータを受信するコンピュータの識別情報を通信先とする通信記録が通信ログ L1 に追加される。この場合、通信日時は、コンピュータ A (2) がコンピュータ B (3) からデータを受信した日時およびコンピュータ A (2) がコンピュータ B (3) から受信したデータを、ネットワーク NW またはクラウドネットワーク CL に接続されているコンピュータへデータを送信した日時の少なくとも一方となる（なお、コンピュータ A (2) がコンピュータ B (3) からデータを受信した日時と、コンピュータ A (2) がコンピュータ B (3) から受信したデータをネットワーク NW またはクラウドネットワーク CL に接続されているコンピュータにデータを送信した日時との両方を通信日時として追加する場合には、2 つの通信記録が追加され、すなわち、コンピュータ A (2) がコンピュータ B (3) からデータを受信した日時を通信日時とする通信記録とコンピュータ A (2) がコンピュータ B (3) から受信したデータを、ネットワーク NW またはクラウドネットワーク CL に接続されているコンピュータにデータを送信した日時を通信日時とする通信記録が追加される。）。 10

【0020】

コンピュータ B (3) とネットワーク NW またはクラウドネットワーク CL に接続されているコンピュータとがコンピュータ A (2) を経由してデータの送受信を行なうとすると、コンピュータ A (2) は、ファイアウォールやプロキシサーバとして機能してもよい。 20

【0021】

コンピュータ A (2) およびコンピュータ B (3) それぞれに検出装置 100 が備えられている場合には、それぞれの検出装置 100 は、適宜な時間において、通信ログ L1 および通信ログ L2 をそれぞれ読み出し（図 1 の Read 1 および Read 3）、処理を行ない、マルウェアによる通信の通信記録の有無を検出する。

【0022】

また、コンピュータ B (3) が、ネットワーク NW またはクラウドネットワーク CL に接続されているコンピュータと通信する場合、コンピュータ A (2) を介して通信が行われる場合には、コンピュータ B (3) が検出装置 100-2 を備えなくても、コンピュータ A (2) に検出装置 100-1 が備えられていれば、検出装置 100-1 が通信ログ L1 を読み出し（Read 1）、処理を行なうことで、コンピュータ A (2) 内で動作するマルウェアによる通信が行われたかどうかのみならず、コンピュータ B (3) 内で動作するマルウェアによる通信が行われたかどうかの検出も可能である。 30

【0023】

また、検出装置 100 は、コンピュータ A (2) およびコンピュータ B (3) の外部に存在する場合には、通信ログ L1 および通信ログ L2 をコンピュータ A (2) およびコンピュータ B (3) の外部に位置する検出装置 100 に転送し（図 1 の Read 2 および Read 4）、検出装置 100 は、通信ログ L1 および通信ログ L2 の処理を行ない、ビーコンの送信が行われたかどうかを検出する。 40

【0024】

あるいは、検出装置 100 は、通信ログ L1 および L2 を読み出さず、通信記録が生成される都度、通信記録が検出装置 100 に生成された通信記録が入力されるようになっていてもよい。

【0025】

また、コンピュータ C (4) 内で動作するマルウェアがコンピュータ 5 に向けてビーコンを送信すると、その送信のログは、通信ログ L2 および通信ログ L1 に記録され、通信ログ L2 および通信ログ L1 を検出装置 100 が解析することにより、コンピュータ C (4) 内で動作するマルウェアが検出することも可能である。 50

【0026】

あるいは、コンピュータC(4)が乗っ取られ、攻撃者41の支配下になり、コンピュータA(2)またはコンピュータB(3)にマルウェアが感染し、コンピュータC(4)にビーコンを送信する場合も、通信ログL2または通信ログL1を検出装置100が解析することにより、コンピュータA(2)あるいはコンピュータB(3)にマルウェアが感染したことを検出することが可能である。

【0027】

図2は、通信ログに含まれる通信記録の例を示す。図2の例では、一行それぞれが一つの通信記録を形成しており、通信日時、通信先の識別情報および通信先の識別情報がカンマにより区切られた形式となっている。通信日時は、西暦年、月、日、時、分および秒により表わされているが、所定の日時(例えば、グリニッジ標準時の1970年1月1日0時0分0秒)からの経過時間が秒単位の整数により表わされていてもよい。また、秒単位よりも小さな単位(例えばマイクロ秒単位)あるいは大きな単位(例えば10秒単位)で経過時間が表わされていてもよい。また、通信元の識別情報および通信先の識別情報は、IPv4アドレスの形式であっても、FQDN(Fully Qualified Domain Name)であってもよいし、あるいは、IPv6アドレスの形式であってもよい。また、他のコンピュータの通信を中継しないコンピュータであって、通信を行なうプログラムを実行した利用者が判明するようなコンピュータであれば、通信元の識別情報は、当該利用者名となってもよい。また、他のコンピュータの通信を中継しない末端のコンピュータの場合には、当該コンピュータが通信元となるので、通信ログに通信元の識別情報の記録を省略してもよい。

【0028】

また、さらに、通信先のポート番号および通信プロトコルのいずれか一以上が通信記録に含まれていてもよい。また、通信プロトコルがFTP(File Transfer Protocol)やHTTP(HyperText Transfer Protocol)の場合には、通信により要求されるファイル名が通信記録に含まれていてもよい。

【0029】

(実施形態1)

図3は、実施形態1に係る検出装置100の機能ブロック図の一例を示す。検出装置100は、間隔算出部301と、検出部302とを備える。

【0030】

間隔算出部301は、通信ログの通信記録に含まれる通信元の識別情報および通信先の識別情報の組ごとに、通信が行われた時間間隔を算出する。例えば、通信元の識別情報をIPアドレス1.2.3.4とし、通信先の識別情報をIPアドレス5.6.4.8とする組(1.2.3.4;5.6.4.8)について、2017年8月8日14時20分38秒に通信が行われ、次に、2017年8月8日14時20分50秒に通信が行われた場合、時間間隔として12秒を算出する。また、組(1.2.3.4;5.6.4.8)についてさらに、2017年8月8日14時21分3秒に通信が行われた場合、通信間隔として13秒を別の時間間隔として算出する。

【0031】

この場合、例えば2017年8月8日14時20分50秒と2017年8月8日14時21分3秒との間に、別の組、例えば、(2.2.3.4;5.6.7.8)を通信元の識別情報および通信先の識別情報の組とする通信が行われていても、組(1.2.3.4;5.6.4.8)の通信の時間間隔の算出には影響が及ばない。

【0032】

通信ログの通信記録に含まれる通信元の識別情報および通信先の識別情報の組ごとに算出される時間間隔は、複数個あり得る。そこで、ある組について算出される時間間隔が複数個ある場合には、時間間隔の列と呼ぶ場合がある。また、一般に、ある組について算出される時間間隔の列に同じ値の時間間隔が存在しても、一つの時間間隔にまとめることはしない方が好ましい。したがって、時間間隔の列は、多重集合(マルチセット)になり得

10

20

30

40

50

る。

【0033】

図4は、間隔算出部301が、時間間隔を算出するために使用するデータ構造と、当該データ構造が通信ログの読み込みにより変化する様子との一例を示す。間隔算出部301は、例えば通信ログを順に読み込み、図4(A)に示すような構造のデータを生成してもよい。

【0034】

図4(A)は、間隔算出部301が、図2に示される通信ログの先頭から3つの具体的な数値で例示されている通信記録のうち、最後の「2017/8/8-ORIGIN₃, DESTINATION₃」までを読み込んだ時点におけるデータ構造を示す。

10

【0035】

ノード410、420および430は、図2に示される3つの具体的な数値で例示される通信記録それぞれの通信元の識別情報および通信先の識別情報の組を表わすノードである。このようなノードを「ヘッドノード」と呼ぶ。ヘッドノード410は、最初の通信記録「2017/8/8-15:24:00, ORIGIN₁, DESTINATION₁」の通信元の識別情報および通信先の識別情報の組(ORIGIN₁; DESTINATION₁)に対応する。ヘッドノード420およびヘッドノード430についても同様であり、図2において具体的な数値で例示されている2番目および3番目の通信記録に対応するヘッドノードとなっている。

【0036】

20

ヘッドノード410、420および430それぞれに接続する次のノード411、421および431は、順に通信記録を読み込んでいる時点において、ヘッドノード410、420および430それぞれに対応する通信元の識別情報および通信先の識別情報の組について最後に読み取った通信記録の通信日時を格納する。したがって、ノード411、421および431に格納される日時は、通信記録の読み込みにより変化し得る。図4(A)は、間隔算出部301が、図2に示される3つの具体的な数値により示される通信記録(i)「2017/8/8-15:24:00, ORIGIN₁, DESTINATION₁」、(ii)「2017/8/8-15:27:14, ORIGIN₂, DESTINATION₂」および(iii)「2017/8/8-15:27:15, ORIGIN₃, DESTINATION₃」までを読み込んだ時点での状態を例示している。したがって、ノード411には、(ORIGIN₁; DESTINATION₁)の組の通信日時である2017/8/8-15:24:00が格納されている。ノード421およびノード431についても同様である。

30

【0037】

通信日時を格納するノード411、421および431には、通信が行われた時間間隔のリストが続く。例えば、ヘッドノード410に続くノード411には、空リスト499が続いている。これは、図2に示す通信記録(i)「2017/8/8-15:24:00, ORIGIN₁, DESTINATION₁」を間隔算出部301が読み込むまでは、通信元の識別情報をORIGIN₁とし、通信先の識別情報をDESTINATION₁とする通信が行われた通信記録が存在しなかったことを意味する。すなわち、通信記録(i)「2017/8/8-15:24:00, ORIGIN₁, DESTINATION₁」を間隔算出部301が読み込んだ時点では、ヘッドノード410が作られていなかったため、検出装置100がヘッドノード410およびノード411を作成したことを意味する。

40

【0038】

ヘッドノード420に続くノード421には、3が格納されているノード422と、1が格納されているノード423とにより構成されるリストがつながっている。これは、通信記録(ii)「2017/8/8-15:27:14, ORIGIN₂, DESTINATION₂」を読み込む前に、通信元の識別情報をORIGIN₂とし、通信先の識別情報をDESTINATION₂とする通信記録が2つ存在し、それぞれをR1およびR

50

2とすると、R1とR2との通信の時間間隔が1秒であり、R2と(i i)「2017/8/8-15:27:14, ORIGIN₂, DESTINATION₂」との通信の時間間隔が3秒であることを示している。

【0039】

同様に、ヘッドノード430に続くノード431には、少なくとも、100が格納されているノード432と、89が格納されているノード433と、4が格納されているノード434とにより構成されるリストがつながっている。これは、通信記録「2017/8/8-15:27:15, ORIGIN₃, DESTINATION₃」を読み込む前に、通信元の識別情報をORIGIN₃とし、通信先の識別情報をDESTINATION₃とする通信が、少なくとも3つ以上存在していたことを示し、通信間隔が、4秒、89秒、100秒であることを示している。

10

【0040】

図4(A)に示すデータ構造が検出装置100により作られたのちに、図4(B)に示す通信記録440が読み込まれたときのデータ構造の変化を図4(C)に示す。図4(B)に示す通信記録は、通信元の識別情報をORIGIN₁とし、通信先の識別情報をDESTINATION₁とする通信があったことを示すので、ヘッドノード410が特定される。次に、ヘッドノード410に続く411に格納されている通信日時2017/8/8-15:24:00が読み込まれ、通信記録440の通信日時である2017/8/8-15:27:16との差が196秒と算出されるので、196という数値を格納するノード412が作られ、ノード411に接続される。また、ノード411に格納される値が、通信記録440の2017/8/8-15:27:16に変更される。

20

【0041】

図5は、以上において具体的な値を用いて説明した処理、すなわち、検出装置100が間隔算出部301を用いて、通信元の識別情報および通信先の識別情報の組ごとに通信が行われた時間間隔を算出して図4に示したデータ構造を作成する処理のフローチャートの一例を示す。

【0042】

ステップS501において、通信ログの全ての通信記録を読み込んだかどうかを判断する。例えば、通信ログを読み込もうとしてREADシステムコールなどを発行した際、ファイルの最後(EOF(End of File))に到達したことが検出されたかどうかを判断する。もし、EOFに到達した場合には、データ構造を作成する処理が終了する。

30

【0043】

EOFに到達していない場合には、ステップS502に進み、通信ログから次の通信記録を取得し読み込む。

【0044】

ステップS502の後には、ステップS502にて取得した通信記録から、通信元の識別情報および通信先の識別情報の組に対応するヘッドノードが存在しているかどうかを判断する。例えば、通信元の識別情報と通信先の識別情報との組をキー値として保持し、キー値の組に対応するヘッドノードのアドレスを値とする索引(インデックス)を図4に示すデータ構造とともに作っておき、そのインデックスを参照し、ヘッドノードが存在するかどうかを判断することができる。そのようなインデックスが無い場合は、ヘッドノードの内容を順にチェックして判断してもよい。

40

【0045】

組に対応するヘッドノードが存在すれば、ステップS504に進み、最終通信日時、すなわち、ヘッドノードに続くノードに格納されている通信日時と、通信記録の通信日時との時間間隔を算出し、当該時間間隔を新たなノードに格納し、当該ノードを、最終通信日時を格納しているノードに続くリストに追加する。

【0046】

ステップS504の後には、ステップS506に進み、ヘッドノードに続く最終通信日時を格納するノードの値を、ステップS502で取得した通信記録の通信日時に更新する

50

。その後は、ステップ S 5 0 1 へ戻る。

【 0 0 4 7 】

ステップ S 5 0 3 において、組に対応するヘッドノードが存在しなければ、ステップ S 5 0 5 に進み、ステップ S 5 0 2 で取得した通信記録の通信元の識別情報と通信先の識別情報とに対応するヘッドノードを追加し、最終通信日時を通信記録の通信日時とするノードを追加し、さらに空リストをつなぐ。その後は、ステップ S 5 0 1 へ戻る。

【 0 0 4 8 】

なお、短時間、例えば一秒未満の間に、通信元の識別情報と通信先の識別情報とが同一である通信が行われる場合には、マルウェアによる通信ではないと判断して、ステップ S 5 0 4 をスキップしてもよい。この理由は、通信元の識別情報と通信先の識別情報とを同じにする通信が短時間に行われていると、コンピュータの管理者等により異状が検出されやすくなり、マルウェアとして存在できる可能性がなくなるからである。また、一つのウェブページに同一ホストに格納されている多数の画像へのリンクが存在すると、当該ウェブページをブラウザが読み込むと、ブラウザは、その同じホストに画像のリクエストを多数送信し、これはマルウェアでなくても通常のソフトウェアにより起こり得ることである。

【 0 0 4 9 】

図 3 に戻り、検出装置 1 0 0 の説明を続ける。検出部 3 0 2 は、間隔算出部 3 0 1 により算出された時間間隔の規則性に基づいて、マルウェアによる通信が行われたかどうかを検出する。例えば、検出部 3 0 2 は、間隔算出部 3 0 1 により算出された時間間隔に規則性があると判断すれば、マルウェアによる通信が行われたと検出する。

【 0 0 5 0 】

間隔算出部 3 0 1 により算出された時間間隔に規則性があるかどうかの判断の基準には、種々の基準を挙げることができる。そのような基準の一例を挙げると、通信元の識別情報および通信先の識別情報の組について算出された時間間隔を生成する（例えば形式言語学における）文法規則が存在するかどうかを挙げることができる。文法規則の一例としては、正規表現による文法を挙げることができる。また、別の文法規則としては、文脈依存文法や文脈自由文法、帰納的可算言語を生成する文法も挙げることができる。

【 0 0 5 1 】

例えば、時間間隔が実質的に同一である場合には、「I+」により例示される正規表現による文法が存在する。ここに「I」は、実質的に同一である時間間隔の値であり、「+」は、「+」の前の値が 1 個以上繰り返して現れることを意味する。例えば、時間間隔が、4、4、4、4・・・のように、4 秒が続く場合には、「4+」として、正規表現による文法による表現が可能である。したがって、規則性があると判断できる。

【 0 0 5 2 】

なお、「実質的に同一」と記したのは、通信記録に含まれる通信日時の測定が離散的に例えば 1 秒間隔で行われたり、通信日時の測定の基準となるコンピュータのクロックなどの揺らぎが存在したりすることなどが配慮されていることを意味する。

【 0 0 5 3 】

あるいは、時間間隔の規則性の程度を数値化し、数値化した結果の値を、定められた値と比較することにより、規則性の有無を判断してもよい。

【 0 0 5 4 】

例えば、上述した文法規則を表現する文字の数が少ないほど、規則性の程度が高いと判断してもよい。より具体的には、文法規則を表現する文字の数が例えば 5 1 2 文字より少なければ規則性があると判断する。

【 0 0 5 5 】

時間間隔の規則性の程度を数値化する別の例としては、例えば、時間間隔の分散、すなわち、時間間隔の平均からの偏差の 2 乗の平均値の算出を挙げることができる。分散が小さいほど、実質的に同一、あるいは、ほぼ同一の時間間隔で通信を行なっていることになり、人の操作が介入した通信ではなく、なんらかのプログラムの実行により通信が行われている可能性が高く、当該プログラムはマルウェアである可能性が高いといえる。

【0056】

また、別の時間間隔の規則性の程度を数値化するさらに別の例としては、各時間間隔の頻度を算出し、算出した頻度を、標準分布やポアソン分布、 χ 二乗分布などの統計分布に基づく処理を行なって数値化することを挙げることができる。一例として、上述した分散を算出した例においては、さらに平均を求めることにより想定される標準分布と、算出された時間間隔の分布の適合度を規則性の程度として数値化することができる。また、マルウェアの理論的な通信の時間間隔をあらかじめ算出しておき、間隔算出部が算出した時間間隔と理論的な通信の時間間隔とに基づいて検定統計量を算出し、有意水準に対応する χ 二乗分布の値と検定統計量との差分などにより、規則性の程度を数値化することも可能である。

10

【0057】

マルウェアの理論的な通信の規則性のある時間間隔の別の例としては、実質的に一定の範囲内で変動する時間間隔を挙げることができる。別言すれば、時間間隔の上限 U と下限 L を有する時間間隔の列である。また、さらに上限 U と下限 L との間で、時間間隔が一樣に分布している、すなわち、時間間隔が、上限 U と下限 L との間の疑似乱数と判断できる場合を挙げることができる。疑似乱数と判断できれば、疑似乱数を生成して通信の時間間隔を制御しているマルウェアの存在が検出可能である。

【0058】

図6は、検出部302が、時間間隔の規則性の程度を数値化する際に使用するデータ構造の一例を示す。図6において、ノード601は、図4に示すヘッドノード410に対応し、 $ORIGIN_1$ を通信元の識別情報とし、 $DESTINATION_1$ を通信先の識別情報とする通信を表わし、ノード601から、各時間間隔の頻度を格納するノード602、603、604、605・・・のリストが伸びている。ノード602の「2:3」は、 $ORIGIN_1$ を通信元の識別情報とし、 $DESTINATION_1$ を通信先の識別情報とする通信において、2秒の時間間隔が3回発生したことを表わす。別言すれば、2秒の時間間隔の発生頻度が3であることを表わす。以下同様に、 $ORIGIN_1$ を通信元の識別情報とし、 $DESTINATION_1$ を通信先の識別情報とする通信において、ノード603は、10秒の時間間隔が95回発生し、ノード604は、153秒の時間間隔が1回発生し、ノード605は、196秒の時間間隔が2回発生したことを表わす。

20

【0059】

検出部302は、ノード601から伸びているリストをたどり、例えば、時間間隔の平均や分散を求め、時間間隔の規則性の程度を数値化する。

30

【0060】

図7は、図4(A)などに示した通信元の識別情報と通信先の識別情報との組のヘッドノードに続く通信時間ノードにつながっている通信の時間間隔のリストから、図6に示したデータ構造を生成するフローチャートの一例を示す。ステップS701の処理として、通信の時間間隔のリストを、通信の時間間隔の昇順または降順にソートする。時間間隔が、図4(A)に示したようなリスト構造となっていれば、例えばクイックソートのアルゴリズムを好適に使用することができる。ソートにより、通信の時間間隔が降順または昇順に並ぶリストを得ることができるので、ステップS702の処理として、当該リストを順に走査し、同じ時間間隔 I が連続する個数 C を計数し、 $I:C$ を格納したノードを、例えば図6のノード601から伸びているリストに追加する。

40

【0061】

図8は、以上説明した検出装置の処理のフローチャートの一例を示す。ステップS801として、時間間隔を算出する。別言すれば、通信ログの通信記録に含まれる通信元の識別情報および通信先の識別情報の組ごとに、通信が行われた時間間隔を算出する。

【0062】

ステップS802として、時間間隔の規則性に基づいてマルウェアによる通信が行われたかどうかを検出する。

【0063】

50

図 9 は、一実施形態に係る検出装置 100 を計算機により実現する態様を示す。計算機 900 は、CPU (Central Processing Unit) 901 と、主記憶装置 902 と、補助記憶装置 903 と、インタフェース 904 とを備える。

【0064】

検出装置 100 は、例えば検出装置 100 を実現するためのプログラムが補助記憶装置 903 に記憶された状態から主記憶装置 902 にロードされ、当該プログラムを CPU 901 により実行することにより実現される。当該プログラムの実行により、例えばインタフェース 904 を介して通信記録の読み取りが行われ、主記憶装置 902 内、また、必要であれば補助記憶装置 903 内に、図 4 (A) に例示されるデータ構造が作られ、通信の時間間隔の算出が行われる。そして、例えば図 6 に例示されるデータ構造が作られ時間間隔の規則性の数値化などにより、時間間隔の規則性に基づいてマルウェアによる通信が行われたかどうかを検出する。

【0065】

なお、補助記憶装置 903 は、一時的でない有形の記録媒体の一例である。一時的でない有形の記録媒体の代表的な例としては、磁気ディスク、光磁気ディスク、CD-ROM、DVD-ROM、半導体不揮発性メモリなどが挙げられる。

【0066】

また、CPU 901 は、検出装置 100 を実現するためのプログラムのみを実行することに限定はされず、例えば、図 1 のコンピュータ A (2) およびコンピュータ B (3) の機能を実現するためのプログラムを実行することも可能である。

【0067】

(主な効果の例示)

以上のように、本実施形態においては、通信の時間間隔の規則性に基づいてマルウェアによる通信の有無を検出するので、未知マルウェアであっても検出することが可能である。また、通信の時間間隔の規則性に基づいてマルウェアによる通信の有無を検出するので、正常なサイトが攻撃者などにより乗っ取られてしまった場合でも、当該サイトにビーコンを送信するマルウェアを検出することができる。

【0068】

(実施形態 2)

図 10 は、実施形態 2 に係る検出装置の検出装置の機能ブロック図の別の例を示す。検出装置 100 は、間隔算出部 301 と、検出部 302 と、ホワイトリスト判定部 1001 とを備える。したがって、図 10 に機能ブロック図が示される検出装置は、実施形態 1 に係る検出装置がさらにホワイトリスト判定部 1001 を備える構成を有する。

【0069】

ホワイトリスト判定部 1001 は、間隔算出部 301 が通信記録を読み取るごとに、読み取った通信記録が安全な通信によるものであるかどうかを判定する。当該判定の一態様としては、ホワイトリスト判定部 1001 は、マルウェアと無関係であるコンピュータの識別情報のリスト、すなわちホワイトリスト、にアクセス可能であり、読み取られた通信記録に含まれる通信元の識別情報および通信先の識別情報が、当該ホワイトリストに含まれているかどうかを判定する。このようなホワイトリストに含まれるコンピュータの識別情報としては、ウィルス対策ソフトウェアのダウンロード先のコンピュータの識別情報や、著名な検索エンジンのコンピュータの識別情報がある。

【0070】

また、通信記録に通信により要求されるファイル名が含まれる場合には、当該ファイル名が、ブラウザのブックマークに格納するアイコンが格納されているファイル名であったり、ウェブページのスタイルシートを格納しているファイル名であったりする場合には、ホワイトリスト判定部 1001 は、間隔算出部 301 が読み取った通信記録が安全な通信によるものであると判定してもよい。

【0071】

本実施形態においては、間隔算出部 301 は、通信ログより通信記録を読み出すごとに

10

20

30

40

50

、当該通信記録をホワイトリスト判定部 1001 に判定させ、安全な通信によるものであると判定されれば、当該通信記録をスキップし、次の通信記録を読み出す。

【0072】

(主な効果の例示)

本実施形態によれば、ホワイトリスト判定部 1001 により安全な通信による通信記録であると判定されれば、当該通信記録がスキップされるので、図 4 (A) などに示されるようなデータ構造に使用される記憶容量を小さくすることができる。また、安全な通信をマルウェアによる通信として誤検出する場合を少なくしたり、無くしたりすることができる。

【0073】

(実施形態 3)

図 11 は、実施形態 3 に係る検出装置の機能ブロック図のさらに別の例を示す。検出装置 100 は、間隔算出部 301 と、検出部 302 と、を備え、検出部 302 は、エントロピー算出部 1101 を有する。また、検出装置 100 は、実施形態 2 のように、ホワイトリスト判定部を備えていてもよい。したがって、実施形態 3 に係る検出装置は、実施形態 1 または実施形態 2 に係る検出装置の検出部が、エントロピー算出部を備えている形態である。

【0074】

エントロピー算出部 1101 は、通信元の識別情報および通信先の識別情報の組ごとに算出された時間間隔に基づいて、通信元の識別情報および通信先の識別情報の組ごとにエントロピーを算出する。ここでいう「エントロピー」とは、時間間隔の情報量の期待値であると定義することができる。

【0075】

ここで、時間間隔の情報量とは、時間間隔の総個数 N に対する特定の時間間隔 t の個数 n の割合により定まる値と定義することができる。具体例としては、時間間隔の情報量は、 N/n の対数値である（別言すれば、割合 n/N の対数値に負号を付した値である）と定義することができる。

【0076】

例えば、図 6 に示したノード 601 から伸びるリストが、ノード 602、ノード 603、ノード 604 およびノード 605 のみを有していると仮定した場合には、 $(ORIGIN_1; DESTINATION_1)$ の組ごとに算出された時間間隔の総個数 N は、 $3 + 95 + 1 + 2 = 101$ であるので、 $t = 2$ の情報量は、 $\log(101/3)$ となり、 $t = 10$ の情報量は、 $\log(101/95)$ となり、 $t = 153$ の情報量は、 $\log(101/1)$ となり、 $t = 196$ の情報量は、 $\log(101/2)$ となる。なお、情報量を算出する際の対数関数 $\log$ の底は、2 とする場合が多い。

【0077】

次に、時間間隔の情報量の期待値、すなわち「エントロピー」とは、各時間間隔の情報量に重みづけをした平均値であり、当該重みづけは、時間間隔の総個数 N に対する特定の時間間隔 t の個数 n の割合であると定義することができる。

【0078】

例えば、図 6 に示したノード 601 から伸びるリストが、ノード 602、ノード 603、ノード 604 およびノード 605 のみを有していると仮定した場合の情報量は、 $(3/101) \times \log(101/3) + (95/101) \times \log(101/95) + (1/101) \times \log(101/1) + (2/101) \times \log(101/2)$ となる。

【0079】

なお、通信元の識別情報 S および通信先の識別情報 D の組 $(S; D)$ ごとに算出された時間間隔の総個数 N に対する特定の時間間隔 t の個数 n の割合 n/N を、特定の時間間隔 t の生起した確率 $P(S; D)(t)$ と解釈した場合、上述したエントロピーは、時間間隔の分布の下限を L 、上限を U として、

10

20

30

40

50

【数 1】

$$\sum_{t=L}^U P_{(S;D)}(t) \log_b \left(\frac{1}{P_{(S;D)}(t)} \right) = - \sum_{t=L}^U P_{(S;D)}(t) \log_b (P_{(S;D)}(t))$$

となり、クロード・エルウッド・シャノン (C l a u d e E l w o o d S h a n n o n) が導入したエントロピーと同じ定義となる。なお、対数関数 $\log$ の底 b は、上述したように 2 として情報量を算出する場合が多い。底が 2 の場合、情報量の単位は「ビット」と呼ばれる。また、底をネイピア数とする場合、情報量の単位は「ナット」と呼ばれ、底を 10 とする場合、情報量の単位は「ディット」と呼ばれることが多い。

10

【0080】

なお、 $P(S;D)(t)$ は、離散的に分布する t に対する関数であると考えることができ、確率質量関数と呼ばれる場合がある。

【0081】

したがって、対数の底 b が 1 より大である場合、通信元の識別情報 S および通信先の識別情報 D の組 $(S;D)$ ごとに算出された時間間隔のエントロピーが大きいほど、時間間隔の不規則性の程度が大きいといえ、したがって、エントロピーが小さいほど、時間間隔の規則性の程度が大きいといえる。よって、エントロピーが小さいほど、マルウェアが D で識別される通信先にビーコンを送っている可能性が高いといえる。

20

【0082】

なお、対数関数の底 b が 1 より小さい場合には、対数関数の底が 1 より大きい場合とエントロピーの符号が正負逆になるので、エントロピーの絶対値が小さいほど、マルウェアが D で識別される通信先にビーコンを送っている可能性が高いといえることになる。

【0083】

エントロピー（あるいはエントロピーの絶対値）が小さいか、大きいかは、各組 $(S;D)$ に対して算出された時間間隔のエントロピーの平均値 M と標準偏差 σ とを求め、例えば $M - 2\sigma$ を超えるエントロピーを大きなエントロピーとし、 $M - 2\sigma$ 以下のエントロピーを小さなエントロピーとすることができる。

30

【0084】

具体的な例を示す。時間間隔が、4、4、4、4・・・のように、4 秒が続く場合には、時間間隔として 4 秒が生じた確率が 1 となり、1 の対数の値は 0 なので、エントロピーも 0 となる。

【0085】

また、通信の遅延により、平均して 10 個の時間間隔のうち 5 個が 4 秒であり、残りの 5 個が 5 秒である場合には、 $b = 2$ とすると、エントロピーは、 $(5/10) \times \log(10/5) + (5/10) \times \log(10/5) = (1/2) \times 1 + (1/2) \times 1 = 1$ となる。

40

【0086】

一方、平均して 10 個の時間間隔のうち 5 個が 4 秒であり、残りの 5 個が 5 秒である場合について、実施形態 1 で述べたように、分散を計算すると、0.25 となり、エントロピーの値と単純に比較すると極めて小さいように思える。

【0087】

しかしながら、時間間隔のばらつきが大きくなり、平均して 10 個の時間間隔のうち 5 個が 4 秒であり、残りの 5 個が 46 秒であるとする、分散は 441 となるが、エントロピーは 1 のままである。

【0088】

したがって、実施形態 1 で述べた規則性の検出手法と本実施形態で述べたエントロピー

50

の算出による規則性の程度の数値化を組み合わせたのが好ましいともいえる。

【0089】

図12は、本実施形態に係る検出装置100の処理のフローチャートの一例である。図12のフローチャートは、通信元の識別情報および通信先の識別情報の組を定めてから、時間間隔のエントロピーを算出してマルウェアによる通信を検出する処理である。

【0090】

間隔算出部301は、通信の時間間隔を求めた後、ステップS1201において、時間間隔の個数の総和を算出する。

【0091】

エントロピー算出部1101は、時間間隔のエントロピーを算出するために、ステップS1202において、時間間隔の個数を総和で除し（割り算を行ない）、除した値に基づいてステップS1203においてエントロピーを算出する。

【0092】

ステップS1204において、検出部302は、算出されたエントロピーの値が定められた値以下であるかどうかを判断する。定められた値の具体例としては、後に示すように2.5ビットという値がある。ただし、状況に応じて、定められた値を、例えば時間間隔の個数の総和の大きさに応じて、変化させてもよく、総和が大きければ、大数の法則により正確な時間間隔の分布が得られたとして、定められた値を小さくし、総和が小さければ、時間間隔の分布に誤差が含まれるとして、定められた値を大きくしてもよい。

【0093】

ステップS1204において、算出されたエントロピーの値が定められた値以下であれば、ステップS1205に処理を進め、マルウェアを検出したとする。また、ステップS1204において、算出されたエントロピーの値が定められた値以下でなければ、ステップS1205をスキップする。

【0094】

（具体例）

図13は、実際にインターネットに接続されたコンピュータの通信ログを用いて、通信元の識別情報および通信先の識別情報の組ごとにエントロピーを算出し、算出されたエントロピーの区間ごとに、組の数を計数して作成したヒストグラムを示す。エントロピーが2以下の組は、2個存在し、エントロピーが2を超え、2.75以下の組は28個存在する。2.75を超え、3.5以下の組は405組存在し、3.5を超え4.25以下の組は、343個存在し、4.25を超え5以下の組は114個存在し、5を超える組は、17個存在したことが示されている。したがって、2.75の大きさを持つエントロピーの前後で分布が大きく異なっていることがわかる。

【0095】

図14は、図13の作成に用いた通信ログを用いて、組が通信ログに現れた回数とエントロピーとにしたがってプロットしたグラフを示す。言い換えると、正方形がグラフの上に位置するほど、エントロピーの高い組を表わし、正方形がグラフの右側に位置するほど、通信ログに現れた回数が多い（通信回数が多い）ことを示す。図13においては、辺が図13の左右上下方向に対して斜め45度方向に傾いている正方形（菱形）と、辺が図13の左右上下方向と平行になっている正方形とを用いてプロットされている。この違いは、上述のように時間間隔のエントロピーの平均値 M と標準偏差 σ とを求め、 $M - 2\sigma$ を超えるエントロピーが算出された組を、辺が図13の左右上下方向に対して斜め45度方向に傾いている正方形（菱形）とし、 $M - 2\sigma$ 以下のエントロピーが算出された組を辺が図13の左右上下方向と平行になっている正方形として表した結果である。したがって、辺が図13の左右上下方向と平行になっている正方形は、グラフの下方に現れている。また、辺が図13の左右上下方向と平行になっている正方形はグラフの左側に位置しているので、通信回数が少ないことも理解され、通信回数を少なくして通信回数により検出されないようになっていることも理解される。

【0096】

10

20

30

40

50

図 13 においては、 M は、約 3.6 となり σ は、約 0.58 である。したがって、 $M - 2\sigma$ は、2.44 となった。このため、エントロピーが 2.44 以下の組が、図 13 の左右上下方向と平行になっている正方形として表されており、これらの正方形に対応して、マルウェアによる通信が行われたと判断することができる。

【0097】

(主な効果の例示)

以上のように、本実施形態においては、時間間隔の規則性の程度を通信時間の間隔の個数の割合やエントロピーの値により算出するので、マルウェアによる通信の有無を検出するので、未知マルウェアであっても検出することが可能である。また、例えば、日中の営業時間に動作し、夜間や営業時間外は停止するパーソナルコンピュータにマルウェアが感染した場合であっても、あるいは、営業時間にビーコンを送信し営業時間外はビーコンを停止するマルウェアが常時動作するコンピュータに感染した場合であっても、マルウェアのビーコンの時間間隔の規則性によりマルウェアを検出することができる。

【0098】

以上、開示に係る実施形態について説明した。なお、以上に説明した本実施形態は、特許請求の範囲に記載されている本開示に係る発明の内容を不当に限定するものではない。また本実施形態で説明される構成の全てが、本開示に係る発明の必須構成要件であるとは限らない。

【符号の説明】

【0099】

1・・・全体システム 2、3、4、5・・・コンピュータ 100・・・検出装置 301・・・間隔算出部 302・・・検出部 1001・・・ホワイトリスト判定部 1101・・・エントロピー算出部

10

20

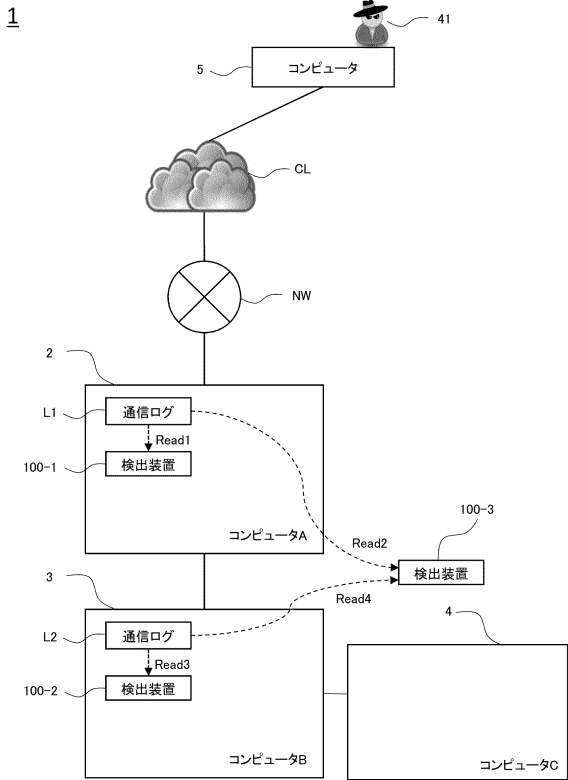
30

40

50

【図面】

【図 1】



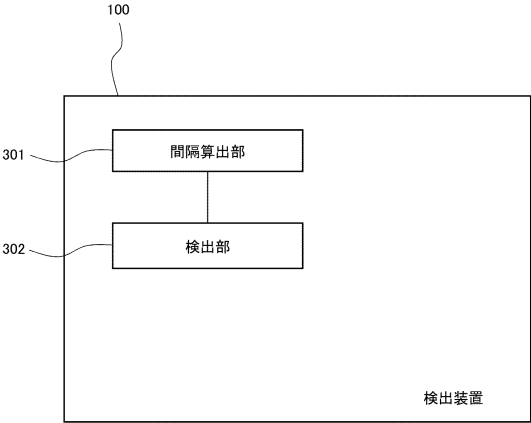
【図 2】

⋮	⋮	⋮
2017/8/8-15:24:00,	ORIGIN ₁ ,	DESTINATION ₁
2017/8/8-15:27:14,	ORIGIN ₂ ,	DESTINATION ₂
2017/8/8-15:27:15,	ORIGIN ₃ ,	DESTINATION ₃
⋮	⋮	⋮

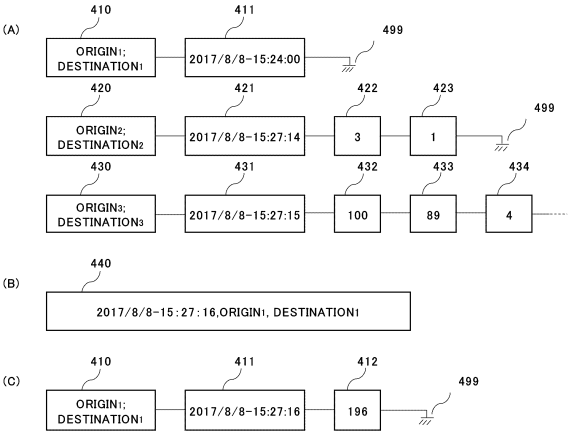
10

20

【図 3】



【図 4】

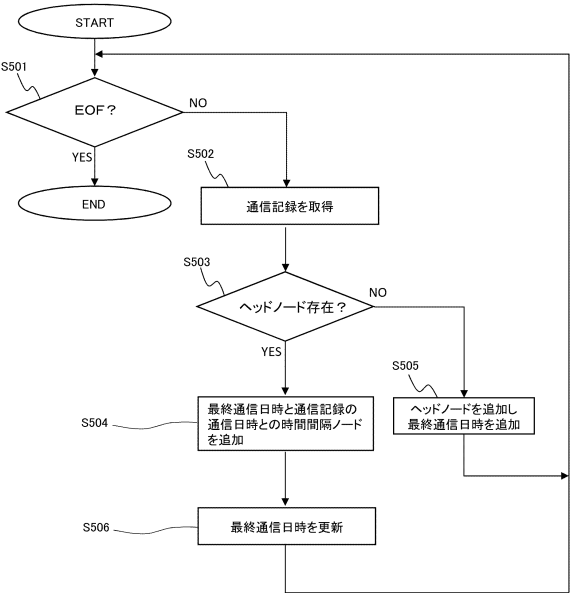


30

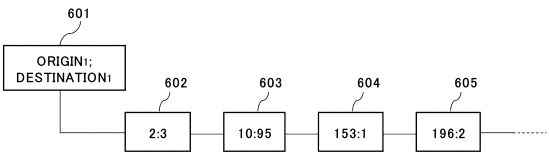
40

50

【図 5】

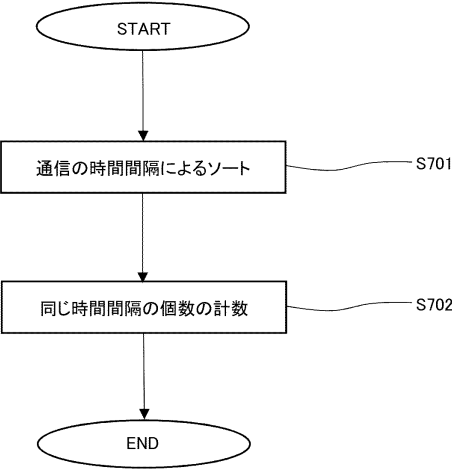


【図 6】

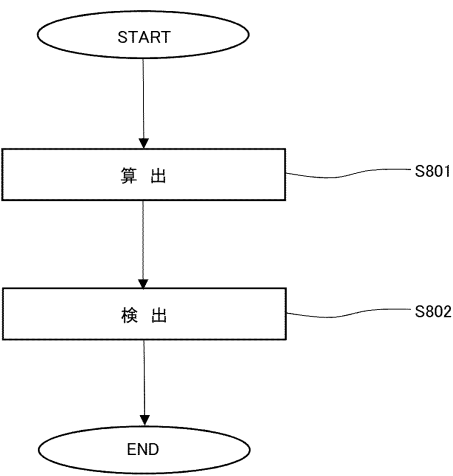


10

【図 7】



【図 8】



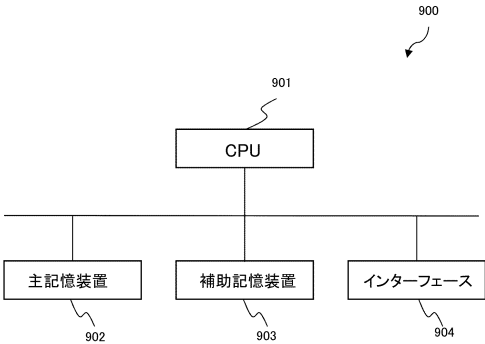
20

30

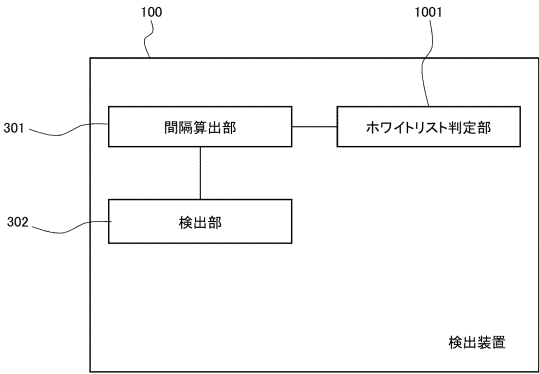
40

50

【図 9】

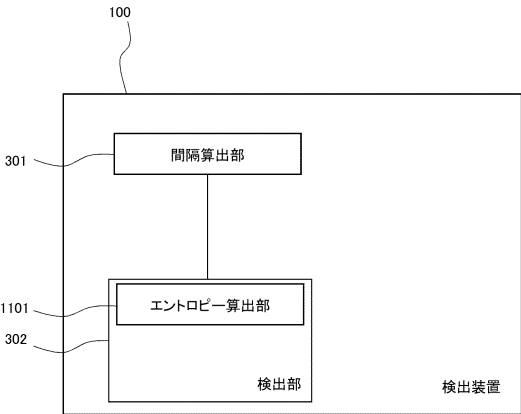


【図 10】

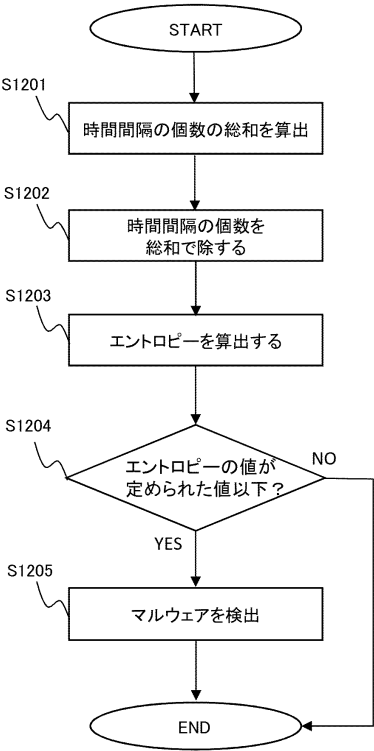


10

【図 11】



【図 12】



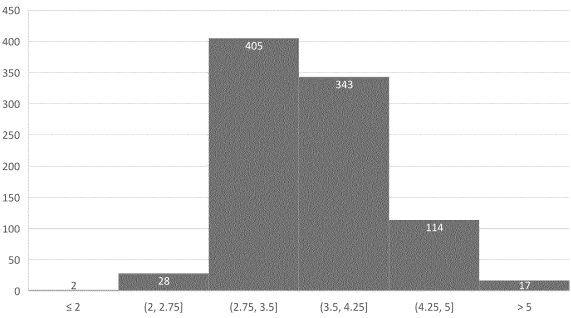
20

30

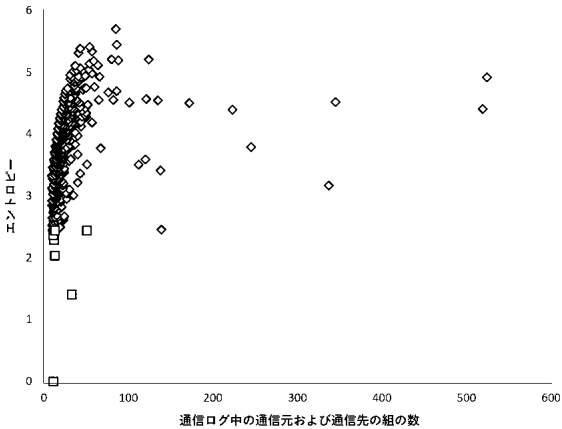
40

50

【図 13】



【図 14】



10

20

30

40

50

フロントページの続き

- (56)参考文献 米国特許出願公開第2015/0304350 (US, A1)
特開2006-319633 (JP, A)
特開2014-057307 (JP, A)
特開2005-151289 (JP, A)
- (58)調査した分野 (Int.Cl., DB名)
G06F 21/55
G06F 21/56