(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6348656号
(P6348656)
(45)発行日　平成30年6月27日(2018.6.27)　(24)登録日　平成30年6月8日(2018.6.8)

(51)Int.Cl.　F I
G06F 21/56 (2013.01)　G06F 21/56 360

請求項の数 8 (全 31 頁)

(21)出願番号　特願2017-506466(P2017-506466)
(86)(22)出願日　平成28年3月8日(2016.3.8)
(86)国際出願番号　PCT/JP2016/057119
(87)国際公開番号　WO2016/147944
(87)国際公開日　平成28年9月22日(2016.9.22)
審査請求日　平成29年4月21日(2017.4.21)
(31)優先権主張番号　特願2015-55281(P2015-55281)
(32)優先日　平成27年3月18日(2015.3.18)
(33)優先権主張国　日本国(JP)

(73)特許権者　000004226
日本電信電話株式会社
東京都千代田区大手町一丁目５番１号
(74)代理人　110002147
特許業務法人酒井国際特許事務所
(72)発明者　青木　一史
東京都千代田区大手町一丁目５番１号　日本電信電話株式会社内
(72)発明者　神谷　和憲
東京都千代田区大手町一丁目５番１号　日本電信電話株式会社内

審査官　青木　重徳

最終頁に続く

(54)【発明の名称】マルウェア感染端末の検出装置、マルウェア感染端末の検出システム、マルウェア感染端末の検出方法およびマルウェア感染端末の検出プログラム

(57)【特許請求の範囲】
【請求項１】
監視対象ネットワークの通信およびマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する系列生成部と、
マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の前記共通イベント系列同士で類似する前記共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する検知用系列抽出部と、
前記系列生成部によって生成された監視対象ネットワークの通信に基づくイベント系列と、前記検知用系列抽出部によって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する検知部と、
を有することを特徴とするマルウェア感染端末の検出装置。
【請求項２】
前記検知用系列抽出部は、
前記共通イベント系列から、前記イベントをノード、前記イベント間の発生順序をエッ

ジ、前記イベントの前後関係の出現回数を前記エッジの重みとする有向グラフを生成し、前記有向グラフの単純道ごとに前記重みの総和を計算し、前記重みの総和が最大となる前記単純道を前記代表イベント系列とすることを特徴とする請求項１に記載のマルウェア感染端末の検出装置。

【請求項３】

前記検知用系列抽出部は、

前記有向グラフの前記単純道に含まれる前記エッジのうち、前記重みが所定の閾値以上であるエッジについて前記重みの総和を計算し、前記重みの総和が最大となる前記単純道を前記代表イベント系列とすることを特徴とする請求項２に記載のマルウェア感染端末の検出装置。

【請求項４】

前記検知用系列抽出部は、

前記同一クラスタにおける最長の前記共通イベント系列の長さに対する前記代表イベント系列の長さの割合が、所定の値よりも小さい場合、前記同一クラスタに含まれる全ての前記共通イベント系列を前記代表イベント系列とすることを特徴とする請求項１に記載のマルウェア感染端末の検出装置。

【請求項５】

前記検知部は、

前記監視対象ネットワークの通信に基づくイベント系列である判定対象イベント系列と前記検知用イベント系列との合致部分の長さの、前記検知用イベント系列の長さに対する割合である第一の合致率と、前記検知用イベント系列の長さの、前記検知用イベント系列が属するクラスタにおける最長の共通イベント系列の長さに対する割合である第二の合致率と、を乗じた値が所定の閾値以上である場合は、前記判定対象イベント系列と前記検知用イベント系列とが合致していると判定し、前記監視対象ネットワークにマルウェア感染端末が存在していることを検知することを特徴とする請求項１に記載のマルウェア感染端末の検出装置。

【請求項６】

マルウェア実行環境と、

監視対象ネットワークと、

マルウェア感染端末の検出装置と、を備えるマルウェア感染端末の検出システムであって、

前記マルウェア感染端末の検出装置は、

前記監視対象ネットワークの通信および前記マルウェア実行環境で実行されるマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する系列生成部と、

マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の前記共通イベント系列同士で類似する前記共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する検知用系列抽出部と、

前記系列生成部によって生成された監視対象ネットワークの通信に基づくイベント系列と、前記検知用系列抽出部によって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する検知部と、

を有することを特徴とするマルウェア感染端末の検出システム。

【請求項７】

マルウェア実行環境と、

監視対象ネットワークと、
マルウェア感染端末の検出装置と、を有するマルウェア感染端末の検出システムで実行されるマルウェア感染端末の検出方法であって、
前記監視対象ネットワークの通信および前記マルウェア実行環境で実行されるマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する系列生成工程と、
マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の前記共通イベント系列同士で類似する前記共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する検知用系列抽出工程と、
前記系列生成工程によって生成された監視対象ネットワークの通信に基づくイベント系列と、前記検知用系列抽出工程によって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する検知工程と、
を含んだことを特徴とするマルウェア感染端末の検出方法。
【請求項８】
監視対象ネットワークの通信およびマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する系列生成ステップと、
マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の前記共通イベント系列同士で類似する前記共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する検知用系列抽出ステップと、
前記系列生成ステップによって生成された監視対象ネットワークの通信に基づくイベント系列と、前記検知用系列抽出ステップによって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する検知ステップと、をコンピュータに実行させるためのマルウェア感染端末の検出プログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、マルウェア感染端末の検出装置、マルウェア感染端末の検出システム、マルウェア感染端末の検出方法およびマルウェア感染端末の検出プログラムに関する。
【背景技術】
【０００２】
近年、情報漏えいや不正アクセス等の脅威をもたらす不正プログラム（以下、「マルウェア」と呼ぶ）が猛威を振るっている。マルウェアは、感染後に攻撃者からサーバ等を介して指令を受け取り、攻撃や情報漏えいなどの脅威をもたらす。昨今のマルウェアは攻撃者との通信を正規の通信に偽装する手法を取る（例えば、非特許文献１参照）。
【０００３】
発見されるマルウェアの数の増加も著しく、数秒に１つの新たなマルウェアが出現しているということが報告されている（例えば、非特許文献２参照）。そのため、アンチウィルスソフト等のホスト側での対策だけではマルウェアによる脅威を防ぎきれない。そこで

、通信データを分析し、マルウェアに感染した端末を特定することでマルウェアの脅威を低減させる手法が注目されている（例えば、非特許文献３参照）。

【０００４】

マルウェアに感染した端末を検知する手法として、マルウェアに感染した端末に見られる通信の特徴をパターン化し、マルウェアに感染した端末を検知する手法が知られている（例えば、特許文献１参照）。マルウェアに感染した端末を検知する手法の一例は、通信データを分析対象とし、マルウェア解析で得られた通信データをパターン化し、監視対象ネットワーク（ＮＷ）の通信に同様のパターンが現れるかを突き合わせることでマルウェアに感染した端末を検知する手法である。

【先行技術文献】

【特許文献】

【０００５】

【特許文献１】特許第５００９２４４号公報

【非特許文献】

【０００６】

【非特許文献１】高度なサイバー攻撃の動向、[online]、[平成26年9月4日検索]、インターネット<URL：http://www.fireeye.com/jp/ja/resources/pdfs/fireeye-advanced-cyber-attack-landscape.pdf>

【非特許文献２】Annual Report Pandalabs 2013 summary、[online]、[平成26年9月3日検索]、インターネット<URL：http://press.pandasecurity.com/wp-content/uploads/2010/05/PandaLabs-Annual-Report_2013.pdf>

【非特許文献３】Sebastian Garcia他、Survey on network-based botnet detection methods、Security and communication networks 2013、[online]、[平成26年3月13日検索]、インターネット<URL：http://onlinelibrary.wiley.com/doi/10.1002/sec.800/full>

【発明の概要】

【発明が解決しようとする課題】

【０００７】

しかしながら、上記従来の技術には以下のような問題があった。すなわち、前述の通り昨今はマルウェアの数が膨大であるため、全マルウェアの全通信をパターン化してしまうとパターン数が膨大となり、監視対象のネットワークの通信に当該パターンが存在するか否かを判定するのに長い時間を要する。また、前述の従来技術では、通信ペイロードごとに状態を定義し、状態の遷移をパターンとするため、マルウェアが異なる通信ペイロードで通信するだけで新たなパターンが生成されてしまう。さらに、マルウェアの通信には感染していない端末の通信と類似の通信も確認されるため、全ての通信パターンを監視対象のネットワークの通信での検知に使うと、誤検知を誘発してしまう。

【０００８】

そこで、本発明は、上述の課題を解決し、マルウェア感染端末を検出する装置、方法およびプログラムを提供することを目的とする。

【課題を解決するための手段】

【０００９】

上述した課題を解決し、目的を達成するために、本発明のマルウェア感染端末の検出装置は、監視対象ネットワークの通信およびマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する系列生成部と、マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の前記共通イベント系列同士で類似する前記共通イベ

ント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する検知用系列抽出部と、前記系列生成部によって生成された監視対象ネットワークの通信に基づくイベント系列と、前記検知用系列抽出部によって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する検知部と、を有することを特徴とする。

【００１０】

また、本発明のマルウェア感染端末の検出システムは、マルウェア実行環境と、監視対象ネットワークと、マルウェア感染端末の検出装置と、を備えるマルウェア感染端末の検出システムであって、前記マルウェア感染端末の検出装置は、前記監視対象ネットワークの通信および前記マルウェア実行環境で実行されるマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する系列生成部と、マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の前記共通イベント系列同士で類似する前記共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する検知用系列抽出部と、前記系列生成部によって生成された監視対象ネットワークの通信に基づくイベント系列と、前記検知用系列抽出部によって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する検知部と、を有することを特徴とする。

【００１１】

また、本発明のマルウェア感染端末の検出方法は、マルウェア実行環境と、監視対象ネットワークと、マルウェア感染端末の検出装置と、を有するマルウェア感染端末の検出システムで実行されるマルウェア感染端末の検出方法であって、前記監視対象ネットワークの通信および前記マルウェア実行環境で実行されるマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する系列生成工程と、マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の前記共通イベント系列同士で類似する前記共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する検知用系列抽出工程と、前記系列生成工程によって生成された監視対象ネットワークの通信に基づくイベント系列と、前記検知用系列抽出工程によって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する検知工程と、を含んだことを特徴とする。

【００１２】

また、本発明のマルウェア感染端末の検出プログラムは、監視対象ネットワークの通信およびマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する系列生成ステップと、マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベン

トを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の前記共通イベント系列同士で類似する前記共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する検知用系列抽出ステップと、前記系列生成ステップによって生成された監視対象ネットワークの通信に基づくイベント系列と、前記検知用系列抽出ステップによって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する検知ステップと、をコンピュータに実行させる。

【発明の効果】

【００１３】

本発明によれば、監視対象ＮＷで照合すべきパターンを削減し、照合にかかる時間を削減することができるとともに、監視対象ＮＷで通常発生する通信を誤って検知する事態を削減することができる。

【図面の簡単な説明】

【００１４】

【図１】図１は、実施の形態に係る検出装置の概要を示す構成図である。

【図２】図２は、実施の形態に係る監視対象ＮＷ分析結果の一例を示す図である。

【図３】図３は、実施の形態に係るマルウェア通信分析結果の一例を示す図である。

【図４】図４は、実施の形態に係る検出装置の構成例を示す図である。

【図５】図５は、実施の形態に係る共通イベント系列の一例を示す図である。

【図６】図６は、実施の形態に係る共通イベント系列から生成されるイベント系列グラフの一例を示す図である。

【図７】図７は、実施の形態に係るイベント系列グラフから生成される単純道の一例を示す図である。

【図８】図８は、実施の形態に係る共通イベント系列および代表イベント系列の一例を示す図である。

【図９】図９は、実施の形態に係るイベント照合部における処理の一例を示す図である。

【図１０】図１０は、除外イベント抽出部による除外イベント抽出処理手順を示すフローチャートである。

【図１１】図１１は、イベント系列生成部によるイベント系列生成処理手順を示すフローチャートである。

【図１２】図１２は、共通イベント系列抽出部による共通イベント系列抽出処理手順を示すフローチャートである。

【図１３】図１３は、代表イベント系列抽出部による代表イベント系列抽出処理手順を示すフローチャートである。

【図１４】図１４は、イベント照合部および候補判定部による候補判定処理手順を示すフローチャートである。

【図１５】図１５は、検知部による検知処理手順を示すフローチャートである。

【図１６】図１６は、イベント照合部による照合処理手順を示すフローチャートである。

【図１７】図１７は、マルウェア感染端末の検出プログラムを実行するコンピュータを示す図である。

【発明を実施するための形態】

【００１５】

以下に、本願に係るマルウェア感染端末の検出装置、マルウェア感染端末の検出システム、マルウェア感染端末の検出方法およびマルウェア感染端末の検出プログラムの実施形態を図面に基づいて詳細に説明する。なお、この実施形態により本願に係るマルウェア感染端末の検出装置、マルウェア感染端末の検出システム、マルウェア感染端末の検出方法、マルウェア感染端末の検出プログラムが限定されるものではない。

【００１６】

［概要］

まず、図１を用いて、マルウェア感染端末の検出装置である検出装置１００が行う処理の概要を説明する。図１は、実施の形態に係る検出装置１００の概要を示す構成図である。図１に示すように、検出装置１００による処理は、検出装置１００が有する系列生成部１３０と検知用系列抽出部１４０と検知部１５０とによって実行される。検出装置１００は、検出を行う前にあらかじめ収集しておいた監視対象ＮＷ（Network、ネットワーク）分析結果（系列抽出用）とマルウェア通信分析結果から検知用イベント系列を生成し、監視対象ＮＷ分析結果（検知用）から生成したイベント系列と検知用イベント系列を照合することにより、監視対象ＮＷにおいてマルウェアに感染している端末（ホスト）を検出する。

【００１７】

ここで、監視対象ＮＷ分析結果（系列抽出用および検知用）には、監視対象ＮＷ内のホストを識別する識別子、イベント、イベント発生時刻のフィールドを有するデータが格納されている。なお、イベントとは、通信に一定の特徴が確認できた際の、各々の特徴を捉えた事象を意味する。例えば、イベントは、ＦｉｒｅｗａｌｌやＷｅｂＰｒｏｘｙなどに記録される装置ログの分析により特定の通信先との通信が含まれていたという事象、あらかじめ決められた時間内に一定回数以上の通信が行われたという事象、ＩＤＳ（Intrusion Detection System）によって悪性なデータ送信が検知されたという事象などが該当する。すなわち、イベントは、監視対象ＮＷにおける通信のうち、悪性の通信の蓋然性が高いものと特徴付けられるルールに合致する事象が該当する。検出装置１００は、例えば、所定の外部装置によって通信を特徴付けるルールに合致するか否かの分析が行われ、ルールに合致したと判定されたイベントを監視対象ＮＷ分析結果として取得する。なお、マルウェア通信分析結果は、サンドボックス等のマルウェア実行環境等においてマルウェアを実際に動作させた際の通信データを、前述の監視対象ＮＷ分析結果を取得した際と同様の観点で分析した結果である。また、イベント系列とは、監視対象ＮＷ分析結果を監視対象ＮＷのホストごとに時系列に沿って並べたもの、またはマルウェア通信分析結果をマルウェア検体ごとに時系列に沿って並べたものである。

【００１８】

ここで、図２を用いて、監視対象ＮＷ分析結果の一例を示す。図２は、実施の形態に係る監視対象ＮＷ分析結果の一例を示す図である。図２に示すように、監視対象ＮＷ内ホストの識別子ごとに検出されたイベントは、イベントの種類と、イベント発生時刻とが対応づけられて格納される。例えば、図２では、「１９２．１６８．１０．１１」により識別されるホストで、「特定の通信先との通信検知」というイベントが「２０１４年１０月１５日１２時２０分１２秒」に発生した例を示している。次に、図３を用いて、マルウェア通信分析結果の一例を示す。図３は、実施の形態に係るマルウェア通信分析結果の一例を示す図である。図３に示すように、マルウェア識別子ごとに検出されたイベントについても監視対象ＮＷ分析結果と同様に、イベントの種類と、イベント発生時刻とが対応づけられて格納される。

【００１９】

以下に、検出装置１００が行う処理について、流れに沿って説明する。検出装置１００に係る系列生成部１３０は、除外イベント抽出部１３１と、イベント系列生成部１３２とを備え、監視対象ＮＷ分析結果と、マルウェア通信分析結果とを入力として、各々についてイベント系列を生成する。系列生成部１３０は、監視対象ネットワークの通信およびマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ネットワークの端末またはマルウェアを区別する識別子ごとに取得されたイベントから、当該イベントの発生順序を踏まえて形成されるイベント系列を生成する。

【００２０】

具体的には、除外イベント抽出部１３１は、監視対象ＮＷ分析結果（系列抽出用）が入力された場合に、当該分析結果において多くの監視対象ＮＷ内のホストで確認されているイベントを除外イベントと設定する。なお、除外イベント抽出部１３１を設けることで誤

検知率を低減させることができるが、検出装置１００は、除外イベント抽出部１３１を設けない構成としてもよい。

【００２１】

イベント系列生成部１３２は、監視対象ＮＷ分析結果とマルウェア通信分析結果とのうち、除外イベントに該当しないイベントからなるイベント系列を生成する。一般に、監視対象ＮＷには感染している端末は少ないため、多くのホストで確認されるイベントはマルウェアによる通信の特徴を捉えたものではないと判断できる。そのため、イベント系列生成部１３２は、除外イベントを除くことにより、マルウェアに感染していない端末で確認されるイベントを除いてイベント系列を生成することができる。すなわち、イベント系列生成部１３２によれば、感染端末の検出における誤検知を低減させることが可能になる。

【００２２】

また、イベント系列生成部１３２は、同一ホストまたは同一マルウェアのイベントについて、イベントの発生間隔が一定時間以内であるものからひとつのイベント系列を生成する。すなわち、イベント系列生成部１３２は、マルウェアの動作に関連する一連の事象を区切ることでイベント系列を生成する。さらに、イベント系列生成部１３２は、同一ホストまたは同一マルウェアのイベントのうち、重複したイベントを除外してイベント系列を生成する。

【００２３】

イベント系列生成部１３２の処理について、具体例を挙げて説明する。例えば、あるホストの分析結果として、イベントＡ、イベントＢ、イベントＣが「ＡＢＣＡＢＣＡＡ」の順で確認されたとする。例えば、イベントＡは、ある特定のサーバへのアクセスを示すイベントであり、イベントＢは、ある特定のサーバからファイルをダウンロードすることを示すイベントであり、イベントＣは、イベントＢでダウンロードしたファイルに基づき所定のサーバにアクセスしたことを示すイベントである。このとき、イベント系列生成部１３２は、「ＡＢＣＡＢＣＡＡ」という一連のイベントから、重複したイベントを除外する。すなわち、イベント系列生成部１３２は、「ＡＢＣＡＢＣＡＡ」から、イベント系列として「ＡＢＣ」を生成する。つまり、イベント系列生成部１３２は、あるホストのイベントを発生時刻が早いものから順にイベント系列の要素として追加し、二回目以降に確認されたイベントはイベント系列に追加しない。

【００２４】

これにより、イベント系列生成部１３２は、マルウェアの実行タイミングやＣ＆Ｃ（Command and Control）サーバからの指令などに起因して繰り返しの通信が発生している場合であっても、繰り返しの回数の差を吸収したイベント系列を生成できる。すなわち、イベント系列生成部１３２によれば、後述する検知処理における精度の向上を図ることが可能となる。

【００２５】

続いて、検出装置１００に係る検知用系列抽出部１４０の処理について説明する。検知用系列抽出部１４０は、共通イベント系列抽出部１４１と、代表イベント系列抽出部１４２と、イベント照合部１４３と候補判定部１４４とを備え、系列生成部１３０が生成したイベント系列に基づいて、検知用イベント系列を抽出する。

【００２６】

検知用系列抽出部１４０は、マルウェアが発生させる通信に基づくイベント系列間での類似度が一定以上のイベント系列同士により形成されるクラスタにおいて、同一クラスタに属するイベント系列間で共通して出現するイベントを取り出し、取り出したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出し、複数の共通イベント系列同士で類似する共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する。

【００２７】

具体的には、共通イベント系列抽出部１４１は、マルウェア通信分析結果から抽出されたイベント系列間の類似度を算出した上でクラスタリング(clustering)を行う。その後、

共通イベント系列抽出部１４１は、一定以上の類似度を有するイベント系列同士において、各イベント系列間で共通的に確認されるイベントについて、順序を加味して抽出し、共通イベント系列とする。

【００２８】

代表イベント系列抽出部１４２では、共通イベント系列抽出部１４１で抽出された共通イベント系列のうち、類似する共通イベント系列から出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列候補として抽出する。なお、詳細な代表イベント系列の抽出方法については後述する。

【００２９】

イベント照合部１４３は、監視対象ＮＷ分析結果（系列抽出用）のイベント系列と検知用イベント系列候補とを照合し、各検知用イベント系列候補が監視対象ＮＷ内のホストをどの程度検出しうるかを算出する。イベント照合部１４３は、監視対象ネットワークの通信に基づくイベント系列である判定対象イベント系列と検知用イベント系列候補との合致部分の長さの、検知用イベント系列候補の長さに対する割合である第一の合致率と、検知用イベント系列候補の長さの、検知用イベント系列候補が属するクラスタにおける最長の共通イベント系列の長さに対する割合である第二の合致率と、を乗じた値が所定の閾値以上である場合は、判定対象イベント系列と検知用イベント系列候補とが合致していると判定する。

【００３０】

候補判定部１４４は、イベント照合部１４３で算出した検知用イベント系列候補ごとの検知ホスト数に基づき、監視対象ＮＷの総ホスト数に対する検知ホスト数の割合が一定以下である場合、検知用イベント系列候補を検知用イベント系列として出力する。

【００３１】

続いて、検出装置１００に係る検知部１５０の処理について説明する。検知部１５０は、イベント照合部１５１と検知結果出力部１５２とを備え、監視対象ＮＷ内のマルウェア感染端末を検知する。検知部１５０は、系列生成部１３０によって生成された監視対象ネットワークの通信に基づくイベント系列と、検知用系列抽出部１４０によって抽出された検知用イベント系列と、が合致していると判定された場合に、当該監視対象ネットワークにマルウェア感染端末が存在していることを検知する。

【００３２】

具体的には、イベント照合部１５１は、検知用系列抽出部１４０のイベント照合部１４３と同様に、監視対象ＮＷ分析結果（検知用）から生成されたイベント系列と検知用イベント系列とが合致するかどうかを照合する。イベント照合部１５１は、監視対象ネットワークの通信に基づくイベント系列である判定対象イベント系列と検知用イベント系列との合致部分の長さの、検知用イベント系列の長さに対する割合である第一の合致率と、検知用イベント系列の長さの、検知用イベント系列が属するクラスタにおける最長の共通イベント系列の長さに対する割合である第二の合致率と、を乗じた値が所定の閾値以上である場合は、判定対象イベント系列と検知用イベント系列とが合致していると判定する。

【００３３】

検知結果出力部１５２は、イベント照合部１５１での照合の結果、検知用イベント系列と合致したと判定されるホスト情報を出力する。ホスト情報とは、例えば、監視対象ＮＷ内の端末のＩＰ（Internet　Protocol）アドレスなどである。

【００３４】

このように、検出装置１００は、監視対象ＮＷ分析結果（系列抽出用）とマルウェア通信分析結果から検知用イベント系列を生成し、監視対象ＮＷ分析結果（検知用）から生成したイベント系列と検知用イベント系列とを照合することにより、監視対象ＮＷにおいてマルウェアに感染している端末を検出する。

【００３５】

上記のように、実施の形態に係る検出装置１００は、複数のマルウェアの通信のうち、マルウェアを特徴付ける共通的な特徴の時系列パターンの中からさらに代表的なものとし

て選ばれた時系列パターンである検知用イベント系列のみを用いて感染端末の検出を行う。このため、検出装置１００によれば、監視対象ＮＷで照合すべきパターンを削減し、照合にかかる時間を削減することができる。さらに、検出装置１００は、あらかじめ監視対象ＮＷで観測されうるイベントやイベントの時系列が除外された検知用イベント系列を処理に用いるため、監視対象ＮＷで通常発生する通信を誤って検知する事態を削減することができる。

【００３６】

なお、検出装置１００は、検知用イベント系列の生成では、監視対象ＮＷ分析結果（系列抽出用）を用いず、マルウェア通信分析結果のみを用いてもよい。また、検出装置１００に係る処理についての詳細は、フローチャートを用いて後述する。

【００３７】

［検出装置の構成］

次に、図４を用いて、実施の形態に係る検出装置１００について説明する。図４は、実施の形態に係る検出装置１００の構成例を示す図である。

【００３８】

図４に例示するように、実施の形態に係る検出装置１００は、ＩＦ（interface）部１１０と、イベント系列記憶部１２０と、検知用イベント系列記憶部１２１と、系列生成部１３０と、検知用系列抽出部１４０と、検知部１５０とを有する。

【００３９】

ＩＦ部１１０は、例えば、ＮＩＣ（Network Interface Card）等であり、外部装置との間で各種データを送受信する。例えば、ＩＦ部１１０は、監視対象ＮＷ分析結果として、監視対象ＮＷに設置されたＦｉｒｅｗａｌｌやＷｅｂＰｒｏｘｙの装置ログ等を分析した結果を受信する。

【００４０】

イベント系列記憶部１２０および検知用イベント系列記憶部１２１は、例えば、ＲＡＭ（Random Access Memory)、フラッシュメモリ（Flash Memory）等の半導体メモリ素子、または、ハードディスク、光ディスク等によって実現される。イベント系列記憶部１２０および検知用イベント系列記憶部１２１は、系列生成部１３０や、検知用系列抽出部１４０や、検知部１５０が扱う情報を適宜記憶する。

【００４１】

例えば、イベント系列記憶部１２０は、系列生成部１３０が生成したイベント系列を記憶する。また、検知用イベント系列記憶部１２１は、検知用系列抽出部１４０が抽出した検知用イベント系列を記憶する。なお、検出装置１００は、イベント系列記憶部１２０または検知用イベント系列記憶部１２１を構成要素とすることを要しない。例えば、検出装置１００は、イベント系列記憶部１２０または検知用イベント系列記憶部１２１と同様の処理を行う外部記憶装置を利用してもよい。

【００４２】

系列生成部１３０、検知用系列抽出部１４０および検知部１５０は、例えば、ＡＳＩＣ（Application Specific Integrated Circuit）やＦＰＧＡ（Field Programmable Gate Array）等の集積回路により実現される。また、系列生成部１３０、検知用系列抽出部１４０および検知部１５０は、例えば、ＣＰＵ（Central Processing Unit）やＭＰＵ（Micro Processing Unit）等によって、図示しない記憶装置に記憶されているプログラムがＲＡＭを作業領域として実行されることにより実現される。

【００４３】

系列生成部１３０は、除外イベント抽出部１３１と、イベント系列生成部１３２とを備え、監視対象ＮＷ分析結果と、マルウェア通信分析結果とを入力として、各々についてイベント系列を生成する。除外イベント抽出部１３１は、監視対象ＮＷ分析結果（系列抽出用）が入力された場合に、当該分析結果において多くの監視対象ＮＷ内のホストで確認されているイベントを除外イベントと設定する。具体的には、除外イベント抽出部１３１は、入力された監視対象ＮＷ分析結果（系列抽出用）に含まれる監視対象ＮＷ内の全ホスト

数と、所定のイベントが含まれるホスト数とを取得する。続いて、除外イベント抽出部１３１は、所定のイベントが含まれるホスト数と全ホスト数との割合に基づき、所定のイベントが含まれるホストが一定の割合を超える場合には、所定のイベントを除外イベントと設定する。これにより、除外イベント抽出部１３１は、多くのホストで行われている一般的な処理を排除したイベントのみでイベント系列を生成させることを可能とする。

【００４４】

イベント系列生成部１３２は、監視対象ＮＷ分析結果とマルウェア通信分析結果とのうち、除外イベントに該当しないイベントからなるイベント系列を生成する。具体的には、イベント系列生成部１３２は、監視対象ＮＷ分析結果またはマルウェア通信分析結果のうち、除外イベントに該当しないイベントを入力として取得する。このとき、イベント系列生成部１３２は、読み込んだイベントのイベント発生時刻を記録する。そして、イベント系列生成部１３２は、記録したイベント発生時刻が直前に読み込んだイベント発生時刻から一定時間以上離れているかを判定する。そして、イベント系列生成部１３２は、イベント発生時刻が直前イベント時刻から一定時間以上離れていない場合、当該イベントはその前のイベントと同一のイベント系列の要素であると推定し、判定対象となったイベント同士をイベント系列として生成する。このように、イベント系列生成部１３２は、一般的な処理を除外したイベントによりイベント系列を生成するので、感染端末の検出における誤検知を低減させることが可能になる。

【００４５】

検知用系列抽出部１４０は、共通イベント系列抽出部１４１と、代表イベント系列抽出部１４２と、イベント照合部１４３と候補判定部１４４とを備え、系列生成部１３０が生成したイベント系列に基づいて、検知用イベント系列を抽出する。

【００４６】

共通イベント系列抽出部１４１は、系列生成部１３０が生成したイベント系列から、共通イベント系列を抽出する。具体的には、共通イベント系列抽出部１４１は、マルウェア通信分析結果から抽出されたイベント系列間の類似度を算出した上でクラスタリングを行い、一定以上の類似度を有するイベント系列同士において、各イベント系列間で共通的に確認されるイベントについて、順序を加味して抽出する。そして、共通イベント系列抽出部１４１は、イベントを時系列順に並べた場合に、共通的に確認できたイベントの長さがあらかじめ決められた長さよりも長いときには、共通的に確認されたイベントからなるイベント系列を共通イベント系列とする。

【００４７】

代表イベント系列抽出部１４２は、共通イベント系列抽出部１４１で抽出された共通イベント系列のうち、類似する共通イベント系列から出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列候補として抽出する。具体的には、代表イベント系列抽出部１４２は、共通イベント系列から、イベントをノード、イベント間の発生順序をエッジ、イベントの前後関係の出現回数をエッジの重みとする有向グラフを生成し、有向グラフの単純道（Simple path）ごとに重みの総和を計算し、最大の重みを示す単純道を代表イベント系列とする。

【００４８】

まず、図５を用いて、代表イベント系列を抽出する共通イベント系列の例について説明する。図５は、それぞれ一つのクラスタから抽出された複数の共通イベント系列を示している。また、図５のラベルはイベントを示している。さらに、図５のイベント間の矢印は、イベントが発生する順序を示している。また、図５の共通イベント系列１、４および５は全く同じイベント系列であるため、一つのイベント系列にユニーク化した上で代表イベント系列の抽出を行ってもよい。

【００４９】

ここで、図５に示すように、ある共通イベント系列が、他の共通イベント系列の一部として表れる場合がある。例えば、図５の共通イベント系列１の「ＡＢＣＤ」という共通イベント系列は、共通イベント系列２および３の一部として表れている。また、ある共通イ

ベント系列の一部のイベントを置き換えたものが他の共通イベント系列となる場合もある。例えば、「ＡＢＣ１ＥＦ」という共通イベント系列のうち「１」を、「２」または「３」に置き換えた「ＡＢＣ２ＥＦ」や「ＡＢＣ３ＥＦ」が存在する場合がある。

【００５０】

同一クラスタ内の共通イベント系列には、上述のように類似した系列が存在するということを前提として、代表イベント系列抽出部１４２における代表イベント系列の抽出方法について説明する。まず、代表イベント系列抽出部１４２は、抽出対象となるクラスタを指定し、指定したクラスタに含まれる共通イベント系列からイベント系列グラフを生成する。ここで、イベント系列グラフとは、イベントをノード、イベントの前後関係をエッジ、イベントの前後関係の出現回数をエッジの重みとしてもつ有向グラフである。代表イベント系列抽出部１４２は、イベント系列グラフの先頭および末尾には、それらを意味するノードを付与する。

【００５１】

図６を用いて、イベント系列グラフの具体的な例について説明する。図６は、実施の形態に係る共通イベント系列から生成されるイベント系列グラフの一例を示す図である。図６に示すように、代表イベント系列抽出部１４２は、イベントをノード、イベントが発生する順序を示す矢印をエッジ、イベントの前後関係の出現回数をエッジの重みとして表している。また、「ＳＴＡＲＴ」というラベルを持つノード１１がイベント系列グラフの先頭を示している。そして、「ＥＮＤ」というラベルを持つノード１８がイベント系列グラフの末尾を示している。

【００５２】

まず、図７に示すように、代表イベント系列抽出部１４２は、イベント系列グラフを生成した後、イベント系列グラフの始点から終点までの単純道を取り出す。図７は、実施の形態に係るイベント系列グラフから生成される単純道の一例を示す図である。次に、代表イベント系列抽出部１４２は、取り出した各単純道について、代表イベント系列を取り出し得るかを判定する。以降、代表イベント系列抽出部１４２が代表イベント系列を抽出する手順について詳細に説明する。

【００５３】

この時、代表イベント系列抽出部１４２は、イベントの前後関係の出現回数が一定数以上であるイベントを有向グラフのノードとしてもよい。例えば、図７の例において、一定回数を2と設定した場合、ノード１７の前後関係の出現回数は1であり一定回数以上ではないため、イベント系列グラフのノードには含まれないことになる。

【００５４】

まず、代表イベント系列抽出部１４２は、各単純道について、グラフの先頭を「イベント（前）」、次のイベントを「イベント（後）」とする。次に、代表イベント系列抽出部１４２は、スキップフラグを偽に設定し、代表イベント系列の重みを0に初期化する。そして、代表イベント系列抽出部１４２は、「イベント（前）」と「イベント（後）」の間の発生回数をイベント系列グラフの基となっている共通イベント系列の個数で除算した値があらかじめ設定された閾値以上であるか否かを判定する。

【００５５】

第一に、「イベント（前）」と「イベント（後）」の間の発生回数をイベント系列グラフの基となっている共通イベント系列の個数で除算した値があらかじめ設定された閾値以上である場合について説明する。この場合、代表イベント系列抽出部１４２はスキップフラグが真であるかどうかを判定する。スキップフラグが真である場合は、代表イベント系列抽出部１４２は、「イベント（前）」を代表系列の要素に選定し、スキップフラグを偽に設定する。そして、「イベント（後）」を代表イベント系列の要素に選定し、「イベント（前）」と「イベント（後）」の発生回数を当該代表イベント系列の重みに追加する。なお、初回の場合は、スキップフラグは必ず偽であるが、これらの処理は各イベントに対して繰り返し行われるので、２回目以降はスキップフラグが偽である場合と真である場合の両方が考えられる。

【００５６】
第二に、「イベント（前）」と「イベント（後）」の間の発生回数をイベント系列グラフの基となっている共通イベント系列の個数で除算した値があらかじめ設定された閾値よりも小さい場合について説明する。この場合、代表イベント系列抽出部１４２は、スキップフラグを真に設定する。
【００５７】
その後、「イベント（後）」がグラフの終点と一致していた場合には、代表イベント系列抽出部１４２は、次の単純道の判定を実施する。終点と一致していなかった場合、代表イベント系列抽出部１４２は、「イベント（後）」を「イベント（前）」にし、さらに「イベント（後）」の次のイベントを新たな「イベント（後）」に設定しなおし、処理を継続する。そして、代表イベント系列抽出部１４２は、全ての単純道の判定を実施した後、各代表イベント系列のうち、重みが最大となるものを当該クラスタの代表イベント系列として出力する。なお、代表イベント系列抽出部１４２は、重みが最大となるものが複数存在する場合は、その全てを代表イベント系列として出力してもよいし、代表イベント系列に含まれるイベントの数がもっとも多いもの、もしくはもっとも少ないものを出力してもよい。さらに、クラスタに属する共通イベント系列長の最大値に対する代表イベント系列長の割合が、あらかじめ決められた値よりも小さい場合には、代表イベント系列抽出部１４２は、共通イベント系列を代表イベント系列としてもよい。
【００５８】
図７を用いて、代表イベント系列抽出部１４２が代表イベント系列を抽出する手順を具体的な例を挙げて説明する。最初に、単純道１０ａに対して判定を行う例を示す。まず、代表イベント系列抽出部１４２は、ノード１１ａを「イベント（前）」、ノード１２ａを「イベント（後）」とする。次に、代表イベント系列抽出部１４２は、スキップフラグを偽、閾値を0.3に設定し、代表イベント系列の重みを0に初期化する。そして、ノード１１ａとノード１２ａの間の発生回数を示すエッジ２１ａの重みが5であり、これをイベント系列グラフの基となっている共通イベント系列の個数である5で除算した値が1であるため、代表イベント系列抽出部１４２は、算出した値が閾値以上であると判定する。
【００５９】
そして、スキップフラグが偽であるため、代表イベント系列抽出部１４２はノード１１ａを代表系列の要素に選定せず、スキップフラグを偽から変更しない。そして、代表イベント系列抽出部１４２はノード１２ａを代表イベント系列の要素に選定し、ノード１１ａとノード１２ａの発生回数を示すエッジ２１ａの重み5を当該代表イベント系列の重みに追加する。
【００６０】
続いて、代表イベント系列抽出部１４２は、ノード１２ａとノード１３ａ、ノード１３ａとノード１４ａ、ノード１４ａとノード１５ａについても同様の処理を繰り返す。そして、代表イベント系列抽出部１４２は、ノード１５ａを「イベント（前）」、ノード１８ａを「イベント（後）」とする。次に、ノード１５ａとノード１８ａの間の発生回数を示すエッジ２５ａの重みが3であり、これをイベント系列グラフの基となっている共通イベント系列の個数である5で除算した値が0.6であるため、代表イベント系列抽出部１４２は、算出した値が閾値以上であると判定する。
【００６１】
そして、「イベント（後）」であるノード１８ａを代表イベント系列の要素に選定し、「イベント（前）」であるノード１５ａと「イベント（後）」であるノード１８ａの発生回数を示すエッジ２５ａの重みである3を当該代表イベント系列の重みに追加する。そして、「イベント（後）」がグラフの終点であるノード１８ａと一致しているため、代表イベント系列抽出部１４２は、単純道１０ａの判定を終了し、次の単純道の判定を実施する。この時、代表イベント系列の重みにはエッジ２１ａ、２２ａ、２３ａ、２４ａおよび２５ａの重みが追加され、単純道１０ａの重みの総和は23となる。このとき、単純道１０ａから抽出された代表イベント系列は「ＡＢＣＤ」である。

【００６２】
次に、単純道１０ｂに対して判定を行う例を示す。まず、代表イベント系列抽出部１４２は、ノード１１ｂを「イベント（前）」、ノード１２ｂを「イベント（後）」とする。次に、代表イベント系列抽出部１４２は、スキップフラグを偽、閾値を0.3に設定し、代表イベント系列の重みを0に初期化する。そして、ノード１１ｂとノード１２ｂの間の発生回数を示すエッジ２１ｂの重みが5であり、これをイベント系列グラフの基となっている共通イベント系列の個数である5で除算した値が1であるため、代表イベント系列抽出部１４２は、算出した値が閾値以上であると判定する。
【００６３】
そして、スキップフラグが偽であるため、代表イベント系列抽出部１４２はノード１１ｂを代表系列の要素に選定せず、スキップフラグを偽から変更しない。そして、代表イベント系列抽出部１４２はノード１２ｂを代表イベント系列の要素に選定し、ノード１１ｂとノード１２ｂの発生回数を示すエッジ２１ｂの重み5を当該代表イベント系列の重みに追加する。
【００６４】
続いて、代表イベント系列抽出部１４２は、ノード１２ｂとノード１３ｂ、ノード１３ｂとノード１４ｂ、ノード１４ｂとノード１５ｂ、ノード１５ｂとノード１６ｂについても同様の処理を繰り返す。そして、代表イベント系列抽出部１４２は、ノード１６ｂを「イベント（前）」、ノード１８ｂを「イベント（後）」とする。次に、ノード１６ｂとノード１８ｂの間の発生回数を示すエッジ２７ｂの重みが1であり、これをイベント系列グラフの基となっている共通イベント系列の個数である5で除算した値が0.2であるため、代表イベント系列抽出部１４２は、算出した値が閾値より小さいと判定し、スキップフラグを真に変更する。この場合、代表イベント系列抽出部１４２は、「イベント（前）」であるノード１６ｂと「イベント（後）」であるノード１８ｂの発生回数を示すエッジ２７ｂの重みである1を当該代表イベント系列の重みに追加しない。
【００６５】
ここで、「イベント（後）」がグラフの終点であるノード１８ｂと一致していないため、代表イベント系列抽出部１４２は、「イベント（後）」であるノード１６ｂを「イベント（前）」にし、さらに「イベント（後）」の次のイベントであるノード１８ｂを新たな「イベント（後）」に設定しなおし、処理を継続する。
【００６６】
そして、「イベント（後）」がグラフの終点であるノード１８ｂと一致しているため、代表イベント系列抽出部１４２は、単純道１０ｂの判定を終了し、次の単純道の判定を実施する。この時、代表イベント系列の重みにはエッジ２１ｂ、２２ｂ、２３ｂ、２４ｂおよび２６ｂの重みが追加され、単純道１０ｂの重みの総和は22となる。このとき、単純道１０ｂから抽出された代表イベント系列は「ＡＢＣＤ１」である。
【００６７】
次に、単純道１０ｃに対して判定を行う例を示す。まず、代表イベント系列抽出部１４２は、ノード１１ｃを「イベント（前）」、ノード１２ｃを「イベント（後）」とする。次に、代表イベント系列抽出部１４２は、スキップフラグを偽、閾値を0.3に設定し、代表イベント系列の重みを0に初期化する。そして、ノード１１ｃとノード１２ｃの間の発生回数を示すエッジ２１ｃの重みが5であり、これをイベント系列グラフの基となっている共通イベント系列の個数である5で除算した値が1であるため、代表イベント系列抽出部１４２は、算出した値が閾値以上であると判定する。
【００６８】
そして、スキップフラグが偽であるため、代表イベント系列抽出部１４２はノード１１ｃを代表系列の要素に選定せず、スキップフラグを偽から変更しない。そして、代表イベント系列抽出部１４２はノード１２ｃを代表イベント系列の要素に選定し、ノード１１ｃとノード１２ｃの発生回数を示すエッジ２１ｃの重み5を当該代表イベント系列の重みに追加する。

【００６９】
続いて、代表イベント系列抽出部１４２は、ノード１２ｃとノード１３ｃ、ノード１３ｃとノード１４ｃ、ノード１４ｃとノード１５ｃ、ノード１５ｃとノード１６ｃについても同様の処理を繰り返す。そして、代表イベント系列抽出部１４２は、ノード１６ｃを「イベント（前）」、ノード１７ｃを「イベント（後）」とする。次に、ノード１６ｃとノード１７ｃの間の発生回数を示すエッジ２８ｃの重みが1であり、これをイベント系列グラフの基となっている共通イベント系列の個数である5で除算した値が0.2であるため、代表イベント系列抽出部１４２は、算出した値が閾値より小さいと判定し、スキップフラグを真に変更する。この場合、代表イベント系列抽出部１４２は、「イベント（前）」であるノード１６ｃと「イベント（後）」であるノード１７ｃの発生回数を示すエッジ２８ｃの重みである1を当該代表イベント系列の重みに追加しない。
【００７０】
ここで、「イベント（後）」がグラフの終点であるノード１８ｃと一致していないため、代表イベント系列抽出部１４２は、「イベント（後）」であるノード１７ｃを「イベント（前）」にし、さらに「イベント（後）」の次のイベントであるノード１８ｃを新たな「イベント（後）」に設定しなおし、処理を継続する。
【００７１】
そして、「イベント（前）」であるノード１７ｃと、「イベント（後）」であるノード１８ｃとの間の発生回数を示すエッジ２９ｃの重みが1であり、これをイベント系列グラフの基となっている共通イベント系列の個数である5で除算した値が0.2であるため、代表イベント系列抽出部１４２は、算出した値が閾値より小さいと判定し、スキップフラグを真から変更しない。この場合、代表イベント系列抽出部１４２は、「イベント（前）」であるノード１７ｃと「イベント（後）」であるノード１８ｃの発生回数を示すエッジ２９ｃの重みである1を当該代表イベント系列の重みに追加しない。
【００７２】
そして、「イベント（後）」がグラフの終点であるノード１８ｃと一致しているため、代表イベント系列抽出部１４２は、単純道１０ｃの判定を終了する。この時、代表イベント系列の重みにはエッジ２１ｃ、２２ｃ、２３ｃ、２４ｃおよび２６ｃの重みが追加され、単純道１０ｃの重みの総和は22となる。このとき、単純道１０ｃから抽出された代表イベント系列は「ＡＢＣＤ１」である。
【００７３】
そして、代表イベント系列抽出部１４２は、単純道１０ａ、１０ｂおよび１０ｃから抽出された代表イベント系列の重みの総和を比較し、単純道１０ａから抽出された代表イベント系列の重みの総和が23、単純道１０ｂおよび１０ｃから抽出された代表イベント系列の重みの総和が22であることを得る。これより、代表イベント系列抽出部１４２は、重みの総和が最も大きい単純道１０ａから抽出された代表イベント系列「ＡＢＣＤ」を当該クラスタの代表イベント系列として選定する。
【００７４】
なお、代表イベント系列抽出部１４２は、同一クラスタにおける最長の共通イベント系列の長さに対する代表イベント系列の長さの割合が、所定の値よりも小さい場合、当該クラスタの全ての共通イベント系列を代表イベント系列としてもよい。例えば、所定の値を0.5と設定して、代表イベント系列が「ＡＢ」であり、同一クラスタに「ＡＢＣＤＥ」という最長の共通イベント系列が存在している場合、最長の共通イベント系列の長さに対する代表イベント系列の長さの割合は0.4であり、所定の値未満であるため、共通イベント系列「ＡＢＣＤＥ」を含む全ての共通イベント系列を代表イベント系列とする。すなわち、共通イベント系列の最大長に対する代表イベント系列の長さが所定の値より小さい場合には、当該クラスタの共通イベント系列を代表するイベント系列を作成することができないと判断し、共通イベント系列を代表イベント系列とする。
【００７５】
このように、代表イベント系列抽出部１４２によれば、イベント系列をクラスタリング

し、共通的なイベントを代表イベント系列の要素とすることで、類似の動作を行うマルウェアの亜種が発生した場合にも、通信に共通の特徴を代表する特徴が見られる場合には同一のイベント系列で検知部１５０による判定を実施することが可能となる。すなわち、代表イベント系列抽出部１４２によれば、マルウェアの亜種が頻繁に発生する状況下にあっても、検知に用いるイベント系列を多数用意する必要がないため、検知処理の効率化とマルウェアの亜種に幅広く対応することが可能となる。さらに、検出装置１００は、共通的なイベント系列を代表するイベント系列のみを用いることで、照合の判定を行うイベント系列の数を削減し、処理時間の削減を可能にする。

【００７６】

イベント照合部１４３は、監視対象ＮＷ分析結果（系列抽出用）のイベント系列と検知用イベント系列候補とを照合し、各検知用イベント系列候補が監視対象ＮＷ内のホストをどの程度検出しうるかを算出する。具体的には、イベント照合部１４３は、系列生成部１３０により生成された監視対象ＮＷ分析結果（系列抽出用）のイベント系列と、代表イベント系列抽出部１４２が抽出した検知用イベント系列候補とを入力として取得する。そして、イベント照合部１４３は、両者のイベント系列を照合し、合致すると判定された監視対象ＮＷ分析結果（系列抽出用）に対応するホスト数を算出する。そして、イベント照合部１４３は、算出されたホスト数を出力として候補判定部１４４に出力する。

【００７７】

ここで、イベント照合部１４３における検知用イベント系列候補と監視対象ＮＷの系列との合致の判定方法を図８および図９を用いて説明する。図８は、実施の形態に係る共通イベント系列および代表イベント系列の一例を示す図である。図９は、実施の形態に係るイベント照合部における処理の一例を示す図である。ここでは、図８に示すように、代表イベント系列抽出部１４２において、３つの共通イベント系列から代表イベント系列「ＡＢＤ」を抽出した場合を例に挙げて説明する。なお、前述の通り、以降の説明において検知用イベント系列候補とは、代表イベント系列抽出部１４２で抽出された代表イベント系列を示すものである。

【００７８】

ここで、イベント照合部１４３は、監視対象ＮＷ分析結果の系列と検知用イベント系列候補との最長共通部分列（ＬＣＳ：Longest Common Subsequence）の、検知用イベント系列候補の長さに対する割合である合致率の合致と判定するための閾値を、例えば0.8と定めるものとする。図９の（ａ）は、監視対象ＮＷを分析して得られる系列である。まず、図９の（ｂ）に示すように、仮に、共通イベント系列を検知用イベント系列候補として判定を行った場合の監視対象ＮＷの系列との合致率を見ると、いずれの場合も3/4、すなわち0.75である。前述の通り、合致していると判定する合致率は0.8であるため、イベント照合部１４３は、共通イベント系列は監視対象ＮＷの系列と合致していると判定しない。

【００７９】

次に、図９の（ｃ）に示すように、検知用イベント系列候補を用いて検知を行った場合、監視対象ＮＷの系列との合致率が3/3、すなわち1.0であるため、イベント照合部１４３は、代表イベント系列は監視対象ＮＷと合致していると判定する。この時、「ＡＢＤ」というパターン自体はマルウェア感染に特徴的ではなく、「Ｂ」と「Ｄ」の間に「１」「２」「３」等を含むとマルウェア感染に特徴的である場合は、イベント照合部１４３の代表イベント系列を用いた判定は誤判定につながることになる。

【００８０】

前述のような誤判定を防止するため、イベント照合部１４３は、下記の式に基づいて合致率を再計算する。

第一の合致率＝検知用イベント系列候補と監視対象ＮＷの系列とのＬＣＳ長／検知用イベント系列候補の系列長

第二の合致率＝検知用イベント系列候補の系列長／検知用イベント系列候補を選定したクラスタ内の最長の共通イベント系列長

合致率＝第一の合致率×第二の合致率

【００８１】

図９の例の場合は、第一の合致率が3/3、第二の合致率が3/4であるため、合致率が3/4すなわち0.75となり、イベント照合部１４３は、検知用イベント系列候補は監視対象ＮＷの系列と合致していないと判定する。なお、イベント照合部１４３における判定方法は、ここで説明した方法に限定されず、第二の合致率を用いず、第一の合致率を合致率とする方法であってもよい。

【００８２】

候補判定部１４４は、イベント照合部１４３で算出した検知用イベント系列候補ごとの検知ホスト数に基づき、監視対象ＮＷの総ホスト数に対する検知ホスト数の割合が一定以下である場合、検知用イベント系列候補を検知用イベント系列として出力する。具体的には、候補判定部１４４は、イベント照合部１４３によってイベント系列が合致したと判定された検知用イベント系列候補の検知ホスト数を、監視対象ＮＷの全ホスト数で除算し、イベント系列ごとの検知ホスト割合を算出する。そして、候補判定部１４４は、検知用イベント系列候補の中から、検知ホスト割合が一定以下であるイベント系列を検知用イベント系列として出力する。これにより、候補判定部１４４は、除外イベント抽出部１３１の処理と同様、一般に監視対象ＮＷにはマルウェアに感染している端末が少ないことを踏まえ、検知用イベント系列からあらかじめ誤検知につながるものを除外することができる。このため、候補判定部１４４の処理によれば、監視対象ＮＷにおいて感染端末を検知する際の誤検知を低減することが可能になる。

【００８３】

検知部１５０は、イベント照合部１５１と検知結果出力部１５２とを備え、監視対象ＮＷ内のマルウェア感染端末を検知する。具体的には、イベント照合部１５１は、監視対象ＮＷ分析結果（検知用）と検知用イベント系列のイベント系列同士が合致するかどうかを照合する。そして、検知結果出力部１５２は、イベント照合部１５１での照合の結果、検知用イベント系列と合致したと判定されるホスト情報を出力する。言い換えれば、検知結果出力部１５２は、シグネチャである検知用イベント系列と合致したと判定されるホストについては、マルウェア感染端末である可能性が高いものとして、検知用イベント系列と合致したと判定されるホストを識別することのできる情報を出力することにより、マルウェア感染端末を検出する。なお、イベント照合部１５１においては、イベント照合部１４３と同様の判定方法を用いてもよい。その場合、イベント照合部１４３における検知用イベント系列候補は、イベント照合部１５１においては検知用イベント系列に置き換えられる。

【００８４】

［処理手順］

次に、上述した検出装置１００による検出処理の手順について詳細に説明する。

【００８５】

（除外イベント抽出処理）

まず、図１０を用いて、除外イベント抽出部１３１が実行する除外イベント抽出処理について説明する。図１０は、実施の形態に係る除外イベント抽出部１３１による除外イベント抽出処理手順を示すフローチャートである。

【００８６】

図１０に示すように、除外イベント抽出部１３１は、監視対象ＮＷ分析結果（系列抽出用）を入力として読み込む（ステップＳ１０１）。そして、除外イベント抽出部１３１は、監視対象ＮＷ内のホスト数を取得する（ステップＳ１０２）。ここで、監視対象ＮＷのホスト数は、監視対象ＮＷに存在するホスト数があらかじめわかっている場合はその数としてもよいし、監視対象ＮＷ分析結果（系列抽出用）に出現するホスト数を監視対象ＮＷのホスト数とみなしてもよい。言い換えれば、監視対象ＮＷ内のホスト数とは、監視対象ＮＷで観測可能な総ホスト数であり、あらかじめ存在する総ホスト数が観測されている場合には当該総ホスト数が適用され、総ホスト数が不明の場合には、監視対象ＮＷ分析結果

（系列抽出用）により観測可能なホストの総数が適用される。
【００８７】
続いて、除外イベント抽出部１３１は、読み込んだ監視対象ＮＷ分析結果（系列抽出用）に含まれる全てのイベントについて、除外イベントとするかどうかを判定する処理を実行したか否かを判定する（ステップＳ１０３）。全てのイベントに対して処理を実行したと判定した場合、除外イベント抽出処理は終了する（ステップＳ１０３肯定）。
【００８８】
一方、全てのイベントに対して処理を実行していないと判定した場合（ステップＳ１０３否定）、除外イベント抽出部１３１は、除外イベント抽出処理を続行する。このとき、除外イベント抽出部１３１は、あるイベントで検出されたホスト数を監視対象ＮＷのホスト数で除算し、イベントの検出割合を取得する（ステップＳ１０４）。
【００８９】
そして、除外イベント抽出部１３１は、検出割合があらかじめ指定された値よりも大きいか否かを判定する（ステップＳ１０５）。検出割合があらかじめ指定された値よりも大きいと判定した場合（ステップＳ１０５肯定）、除外イベント抽出部１３１は、判定対象である当該イベントを除外イベントに設定する（ステップＳ１０６）。一方、検出割合があらかじめ指定された値よりも大きくないと判定した場合（ステップＳ１０５否定）、除外イベント抽出部１３１は、当該イベントを除外イベントとは設定せずに、異なるイベントについての処理を続行する（ステップＳ１０３へ移行）。
【００９０】
このように、除外イベント抽出部１３１は、多くのホストで確認されるイベントはマルウェアによる通信の特徴のみを捉えたものではないと判定し、当該イベントを抽出し、除外イベントに設定する。これにより、除外イベント抽出部１３１は、感染端末の検出処理における誤検知を低減することが可能になる。
【００９１】
（イベント系列生成処理）
次に、図１１を用いて、イベント系列生成部１３２が実行するイベント系列生成処理について説明する。図１１は、イベント系列生成部１３２によるイベント系列生成処理手順を示すフローチャートである。
【００９２】
図１１に示すように、イベント系列生成部１３２は、監視対象ＮＷ分析結果（系列抽出用および検知用）とマルウェア通信分析結果とのイベント系列生成処理について、全てのホストまたはマルウェアの分析結果を処理し終えたか否かについて判定する（ステップＳ２０１）。全てに対して処理を実行したと判定した場合、イベント系列生成処理は終了する（ステップＳ２０１肯定）。
【００９３】
一方、全てのホストまたはマルウェアの分析結果を処理し終えていないと判定した場合（ステップＳ２０１否定）、イベント系列生成部１３２は、分析結果を読み込むホストまたはマルウェアを指定する（ステップＳ２０２）。イベント系列生成部１３２は、監視対象ＮＷ分析結果からイベント系列を抽出する際には、監視対象ＮＷ内のホストごとにイベント系列の生成を行う。ここで、ホストの識別には、例えば、ホストのＩＰアドレスを用いる。また、イベント系列生成部１３２は、マルウェア通信分析結果からイベント系列を生成する際には、マルウェアごとにイベント系列の生成を行う。ここで、マルウェアの識別には、例えば、マルウェアのハッシュ値を用いる。なお、監視対象ＮＷ分析結果およびマルウェア通信分析結果は、いずれもイベントの確認された時刻でソートされているものとする。
【００９４】
そして、イベント系列生成部１３２は、以下に説明する処理に先立ち、直前イベント時刻およびイベント系列（処理中）を初期化する（ステップＳ２０３）。
【００９５】

まず、イベント系列生成部１３２は、指定されたホストまたはマルウェアの分析結果を処理し終えたか否かについて判定する（ステップＳ２０４）。分析結果を処理し終えたと判定した場合（ステップＳ２０４肯定）、イベント系列生成部１３２は、イベント系列として出力していないイベント系列（すなわち、生成処理中のイベント系列）が存在するか否かを判定する（ステップＳ２０５）。イベント系列として出力していないイベント系列が存在する場合には（ステップＳ２０５肯定）、イベント系列生成部１３２は、処理中のイベント系列をイベント系列として出力する（ステップＳ２０６）。

【００９６】

一方、イベント系列として出力していない処理中のイベント系列が存在しない場合には（ステップＳ２０５否定）、イベント系列生成部１３２は、処理をステップＳ２０１に移行させる。

【００９７】

ステップＳ２０４において、分析結果を処理し終えていないと判定した場合（ステップＳ２０４否定）、イベント系列生成部１３２は、指定されたホストまたはマルウェアのイベントとイベント発生時刻を読み込む（ステップＳ２０７）。そして、イベント系列生成部１３２は、読み込んだイベントが除外イベントに該当するか否かを判定する（ステップＳ２０８）。除外イベントに該当する場合（ステップＳ２０８肯定）、イベント系列生成部１３２は、読み込んだイベントについてはイベント系列には加えずに、処理をステップＳ２０４に移行する。なお、除外イベントの抽出および除外処理により誤検知率を低減させることができるが、イベント系列生成処理においては、除外イベントの抽出および除外処理を行わないようにしてもよい。

【００９８】

一方、読み込んだイベントが除外イベントに該当しない場合（ステップＳ２０８否定）、イベント系列生成部１３２は、読み込んだイベントのイベント発生時刻を記録する（ステップＳ２０９）。そして、イベント系列生成部１３２は、記録したイベント発生時刻が直前イベント時刻から一定時間以上離れているかを判定する（ステップＳ２１０）。

【００９９】

イベント発生時刻が直前イベント時刻から一定時間以上離れている場合（ステップＳ２１０肯定）、読み込んだイベントは、処理中のイベント系列とは異なるイベント系列に追加されることになるため、イベント系列生成部１３２は、処理中のイベント系列をイベント系列として出力する（ステップＳ２１１）。この場合、イベント系列生成部１３２は、出力したイベント系列（処理中）を初期化する（ステップＳ２１２）。

【０１００】

ステップＳ２１０において、イベント発生時刻が直前イベント時刻から一定時間以上離れていない場合（ステップＳ２１０否定）、イベント系列生成部１３２は、読み込んだイベントのイベント発生時刻を直前イベント時刻に設定する（ステップＳ２１３）。言い換えれば、イベント系列生成部１３２は、イベント発生時刻が直前イベント時刻から一定時間以上離れていない場合、そのイベントはその前のイベントと同一のイベント系列の要素であると推定し、イベント系列（処理中）に追加するかどうかを判定する（後述するステップＳ２１４）。

【０１０１】

そして、イベント系列生成部１３２は、読み込んだイベントがイベント系列（処理中）に含まれているか否かを判定する（ステップＳ２１４）。読み込んだイベントがイベント系列（処理中）に含まれている場合（ステップＳ２１４肯定）、イベント系列生成部１３２は、重複するイベントをイベント系列（処理中）に追加しないため、処理をステップＳ２０４に移行させる。

【０１０２】

一方、読み込んだイベントがイベント系列（処理中）に含まれていない場合（ステップＳ２１４否定）、イベント系列生成部１３２は、当該イベントをイベント系列（処理中）に追加する（ステップＳ２１５）。その後、イベント系列生成部１３２は、処理をステッ

プＳ２０４に移行させる。
【０１０３】
このように、イベント系列生成部１３２は、読み込んだイベントが除外イベントに該当している場合はそのイベントをイベント系列に組み込まない。また、イベント系列生成部１３２は、イベントが発生した時刻を記録し、そのイベントが発生した時刻がその前のイベントが発生した時刻と比較し、一定時間以上離れているかどうかを判定する。これにより、イベント系列生成部１３２は、イベント間の発生間隔が短いイベントによって一つのイベント系列が形成されるように、イベント系列を生成する。さらに、イベント系列生成部１３２は、イベント系列（処理中）に処理対象であるイベントが含まれているかどうかを判断し、含まれている場合には、当該イベントはイベント系列に追加しない。すなわち、生成されたイベント系列には、重複するイベントが存在しない。なお、イベント系列（処理中）に重複するイベントを追加するか否かの判定（ステップＳ２１４）は、監視対象ＮＷ分析結果およびマルウェア通信分析結果の特徴を踏まえ、行わなくてもよい。例えば、監視対象ＮＷ分析結果およびマルウェア通信分析結果で確認されているイベントの種類が少ない場合（例えば、１種類のみなどの場合）には、イベント系列（処理中）に重複するイベントを追加するか否かの判定を行わず、除外イベントに該当しないイベントを全てイベント系列（処理中）に追加してもよい。
【０１０４】
（共通イベント系列抽出処理）
次に、図１２を用いて、共通イベント系列抽出部１４１が実行する共通イベント系列抽出処理について説明する。図１２は、共通イベント系列抽出部１４１による共通イベント系列抽出処理手順を示すフローチャートである。
【０１０５】
図１２に示すように、共通イベント系列抽出部１４１は、マルウェア通信分析結果から抽出したイベント系列を処理対象として読み込む（ステップＳ３０１）。そして、共通イベント系列抽出部１４１は、イベント系列間の類似度行列を生成し、階層的クラスタリングを実施する（ステップＳ３０２）。ここで、類似度行列の生成では、例えば、各イベントに対して一意に識別可能な文字を割り当て、イベント系列を文字列とみなした上で、イベント系列間のレーベンシュタイン距離を計算し、イベント系列の類似度を求める。
【０１０６】
そして、共通イベント系列抽出部１４１は、実施する階層的クラスタリングにおいて、あらかじめ設定された類似度以上のイベント系列同士を同一のクラスタに設定する（ステップＳ３０３）。
【０１０７】
ここで、共通イベント系列抽出部１４１は、全てのクラスタから共通イベント系列を抽出する処理を実行したか否かを判定する（ステップＳ３０４）。全てのクラスタから共通イベント系列を抽出する処理を実行したと判定した場合（ステップＳ３０４肯定）、共通イベント系列抽出部１４１による共通イベント系列抽出処理は終了する。
【０１０８】
一方、全てのクラスタから共通イベント系列を抽出する処理を実行していないと判定した場合（ステップＳ３０４否定）、共通イベント系列抽出部１４１は、共通イベント系列を抽出するクラスタを指定する（ステップＳ３０５）。
【０１０９】
そして、共通イベント系列抽出部１４１は、同一クラスタ内のイベント系列同士で共通する部分列のうち、最長共通部分列を抽出する（ステップＳ３０６）。そして、共通イベント系列抽出部１４１は、あらかじめ決められた長さよりも長い最長共通部分列を共通イベント系列として出力する（ステップＳ３０７）。
【０１１０】
このように、共通イベント系列抽出部１４１は、マルウェア通信分析結果から抽出されたイベント系列間の類似度を算出した上でクラスタリングを行う。その後、共通イベント

系列抽出部１４１は、一定以上の類似度を有するイベント系列同士において、各イベント系列間で共通的に確認されるイベント系列を抽出し、共通イベント系列とする。なお、同一のクラスタ内に単一のイベント系列しか存在しない場合には、共通イベント系列抽出部１４１は、そのイベント系列の長さが一定以上であった場合には当該イベント系列を共通イベント系列として出力する。また、共通イベント系列抽出部１４１は、共通イベント系列とするイベント列の長さを任意に設定することができる。例えば、共通イベント系列抽出部１４１は、最少のイベント系列の長さとして、２以上のイベントが系列に含まれている場合を設定するようにしてもよい。

【０１１１】

（代表イベント系列抽出処理）

次に、図１３を用いて、代表イベント系列抽出部１４２が実行する代表イベント系列抽出処理について説明する。図１３は、代表イベント系列抽出部による代表イベント系列抽出処理手順を示すフローチャートである。

【０１１２】

図１３に示すように、代表イベント系列抽出部１４２は、全てのクラスタから代表イベント系列を抽出したか否かについて判定する（ステップＳ４０１）。全てのクラスタから代表イベント系列を抽出したと判定した場合、代表イベント系列抽出処理は終了する（ステップＳ４０１肯定）。

【０１１３】

一方、全てのクラスタから代表イベント系列を抽出していないと判定した場合（ステップＳ４０１否定）、代表イベント系列抽出部１４２は、代表イベント系列を取り出すクラスタを指定する（ステップＳ４０２）。そして、代表イベント系列抽出部１４２は、指定されたクラスタの共通イベント系列からイベント系列グラフを生成する（ステップＳ４０３）。さらに、代表イベント系列抽出部１４２は、イベント系列グラフの始点から終点までの単純道を抽出する（ステップＳ４０４）。

【０１１４】

次に、代表イベント系列抽出部１４２は、全ての単純道の判定を実施したか否かについて判定する（ステップＳ４０５）。そして、代表イベント系列抽出部１４２は、全ての単純道の判定を実施したと判定した場合（ステップＳ４０５肯定）、代表イベント系列のうち、重みが最大のものを当該クラスタの代表イベント系列として出力する（ステップＳ４０６）。また、代表イベント系列抽出部１４２は、全ての単純道の判定を実施していないと判定した場合（ステップＳ４０５否定）、特定対象の単純道を選択する（ステップＳ４０７）。

【０１１５】

代表イベント系列抽出部１４２は、特定対象のパスを選択すると、グラフの先頭のイベントを「イベント（前）」に設定し、次のイベントを「イベント（後）」に設定し、スキップフラグを偽に設定し、さらに代表イベント系列の重みを0に初期化する（ステップＳ４０８）。そして、代表イベント系列抽出部１４２は、「イベント（前）」と「イベント（後）」の間は発生頻度が閾値以上であるか否かを判定する（ステップＳ４０９）。ここで、発生頻度とは、ある「イベント（前）」と「イベント（後）」の関係がイベント系列グラフにおいて出現する回数を、イベント系列グラフを形成する共通イベント系列の個数で除算した値である。さらに、代表イベント系列抽出部１４２は、「イベント（前）」と「イベント（後）」の間は発生頻度が閾値以上であると判定された場合（ステップＳ４０９肯定）、スキップフラグが真であるか否かを判定する（ステップＳ４１０）。そして、スキップフラグが真であると判定された場合（ステップＳ４１０肯定）、代表イベント系列抽出部１４２は、「イベント（前）」を代表イベント系列の要素に選定し、スキップフラグを偽に設定する（ステップＳ４１１）。

【０１１６】

次に、代表イベント系列抽出部１４２は、「イベント（後）」を代表イベント系列の要素に選定し、「イベント（前）」と「イベント（後）」の間の発生回数を代表イベント系

列の重みに追加する（ステップＳ４１２）。また、「イベント（前）」と「イベント（後）」の間は発生頻度が閾値以上でないと判定された場合（ステップＳ４０９否定）、代表イベント系列抽出部１４２は、スキップフラグを真に設定する（ステップＳ４１３）。

【０１１７】

そして、「イベント（後）」がグラフの終点でなかった場合（ステップＳ４１４否定）、代表イベント系列抽出部１４２は、「イベント（後）」を「イベント（前）」に設定し、「イベント（後）」の次のイベントを「イベント（後）」に設定する（ステップＳ４１５）。また、「イベント（後）」がグラフの終点であった場合（ステップＳ４１４肯定）、代表イベント系列抽出部１４２は、全てのパスの判定を実施したか否かの判定を行う（ステップＳ４０５）。なお、代表イベント系列抽出部１４２で最終的に出力された代表イベント系列を、以降の処理における検知用イベント系列候補とする。

【０１１８】

このように、代表イベント系列抽出部１４２によれば、イベント系列をクラスタリングし、共通的なイベントを代表イベント系列の要素とすることで、類似の動作を行うマルウェアの亜種が発生した場合にも、通信に共通の特徴を代表する特徴が見られる場合には同一のイベント系列で検知部１５０による判定を実施することが可能となる。すなわち、代表イベント系列抽出部１４２によれば、マルウェアの亜種が頻繁に発生する状況下にあっても、検知に用いるイベント系列を多数用意する必要がないため、検知処理の効率化とマルウェアの亜種に幅広く対応することが可能となる。さらに、検出装置１００は、代表的なイベント系列のみを用いることで、照合の判定を行うイベント系列の数を削減し、処理時間の削減を可能にする。

【０１１９】

（候補判定処理）

次に、図１４を用いて、イベント照合部１４３および候補判定部１４４が実行する候補判定処理について説明する。図１４は、イベント照合部１４３および候補判定部１４４による候補判定処理手順を示すフローチャートである。

【０１２０】

図１４に示すように、イベント照合部１４３は、監視対象ＮＷ分析結果（系列抽出用）のイベント系列を検知対象イベント系列として取得する（ステップＳ５０１）。また、イベント照合部１４３は、共通イベント系列抽出部１４１によって抽出された検知用イベント系列候補をシグネチャ系列として取得する（ステップＳ５０２）。

【０１２１】

そして、イベント照合部１４３は、取得した検知対象イベント系列とシグネチャ系列とについて、イベント照合処理を実行する（ステップＳ５０３）。なお、イベント照合部１４３によるイベント照合処理は、検知部１５０に係るイベント照合処理と同様であるため、詳細については後述する。

【０１２２】

続いて、候補判定部１４４は、イベント照合処理によって合致したと判定された照合用イベント系列ごとの検知ホスト数を、監視対象ＮＷのホスト数で除算し、照合用イベント系列ごとの検知ホスト割合を算出する（ステップＳ５０４）。ここで、照合用イベント系列とは、シグネチャ系列のうちから選択された一のイベント系列のことをいう。すなわち、候補判定部１４４は、シグネチャ系列に含まれるイベント系列ごとに検知ホスト割合を算出する。そして、候補判定部１４４は、検知ホスト割合が一定以下である照合用イベント系列を検知用イベント系列として出力する（ステップＳ５０５）。これにより、イベント照合部１４３および候補判定部１４４が実行する候補判定処理は終了する。

【０１２３】

このように、候補判定部１４４は、イベント照合部１４３によって照合された検知用イベント系列候補ごとの検知ホスト数に基づき、監視対象ＮＷの総ホスト数に対する検知ホスト数の割合が一定以下である場合に、当該検知用イベント系列候補を検知用イベント系列として出力する。これは、除外イベント抽出部１３１の処理と同様、一般に、監視対象

ＮＷにはマルウェアに感染している端末が少ないことを踏まえ、検知用イベント系列からあらかじめ誤検知につながるものを除外するための処理である。

【０１２４】

すなわち、監視対象ＮＷにはマルウェアに感染している端末が少ないと仮定すると、本処理で合致したと判定された検知用イベント系列候補はマルウェアの通信のみならず、一般の通信でも確認できるイベント系列であるとみなせるため、検知に用いたときには誤検知を誘発しやすいイベント系列であると判断できる。このため、候補判定部１４４の処理により、一般の通信と区別が難しいマルウェアの通信のイベント系列をあらかじめ除外することで、検知部での誤検知を低減することが可能である。検出装置１００は、例えば、イベント照合処理によって合致したと判定された照合用イベント系列ごとの検知ホスト数を監視対象ＮＷのホスト数で除算した結果が０、すなわち検知用イベント系列候補で監視対象ＮＷ分析結果（系列抽出用）のイベント系列を検知しなかったイベント系列のみを検知用イベント系列として出力してもよい。これにより、検出装置１００は、検知用イベント系列に誤検知を発生させうるものが混入するのを抑制できる。

【０１２５】

（検知処理）

次に、図１５を用いて、検知部１５０が実行する検知処理について説明する。図１５は、検知部１５０による検知処理手順を示すフローチャートである。

【０１２６】

図１５に示すように、検知部１５０に係るイベント照合部１５１は、監視対象ＮＷ分析結果（検知用）のイベント系列を検知対象イベント系列として取得する（ステップＳ６０１）。また、イベント照合部１５１は、検知用系列抽出部１４０によって抽出された検知用イベント系列をシグネチャ系列として取得する（ステップＳ６０２）。そして、イベント照合部１５１は、取得した検知対象イベント系列とシグネチャ系列とについて、イベント照合処理を実行する（ステップＳ６０３）。

【０１２７】

続いて、検知部１５０に係る検知結果出力部１５２は、イベント照合処理によって合致したと判定されたホストをマルウェア感染ホストと判定し、その結果を検知結果として出力する（ステップＳ６０４）。これにより、検知部１５０が実行する検知処理は終了する。

【０１２８】

このように、検知部１５０は、系列生成部１３０によって生成された監視対象ＮＷ分析結果（検知用）のイベント系列と、検知用系列抽出部１４０によって抽出された検知用イベント系列とを照合する。これにより、検知部１５０は、あらかじめ監視対象ＮＷで観測されうるイベントやイベントの時系列が除外されたイベント系列同士を照合することができるので、監視対象ＮＷで通常発生する通信を誤って検知する事態を削減させ、マルウェア感染端末を検知することができる。

【０１２９】

（照合処理）

次に、図１６を用いて、検知部１５０に係るイベント照合部１５１が実行する照合処理について説明する。図１６は、イベント照合部１５１による照合処理手順を示すフローチャートである。なお、検知用系列抽出部１４０に係るイベント照合部１４３も、以下に説明する処理と同様の処理を実行する。

【０１３０】

図１６に示すように、イベント照合部１５１は、監視対象ＮＷ分析結果（検知用）のイベント系列を検知対象イベント系列として取得する（ステップＳ７０１）。また、イベント照合部１５１は、検知用系列抽出部１４０によって抽出された検知用イベント系列をシグネチャ系列として取得する（ステップＳ７０２）。

【０１３１】

そして、イベント照合部１５１は、全ての検知対象イベント系列を判定したか否かを判

定する（ステップＳ７０３）。全ての検知対象イベント系列を判定した場合には（ステップＳ７０３肯定）、イベント照合部１５１による照合処理は終了する。

【０１３２】

一方、全ての検知対象イベント系列を判定していない場合（ステップＳ７０３否定）、イベント照合部１５１は、判定対象イベント系列およびホスト情報を検知対象イベント系列から取得する（ステップＳ７０４）。そして、イベント照合部１５１は、取得したホスト情報に基づいて、検知対象となるホストについて、全てのシグネチャ系列と判定を行ったか否かを判定する（ステップＳ７０５）。全てのシグネチャ系列と判定を行った場合には（ステップＳ７０５肯定）、イベント照合部１５１は、処理をステップＳ７０３に移行させる。

【０１３３】

一方、全てのシグネチャ系列と判定を行っていない場合には（ステップＳ７０５否定）、イベント照合部１５１は、シグネチャ系列から照合用イベント系列を取得する（ステップＳ７０６）。そして、イベント照合部１５１は、判定対象イベント系列と、照合用イベント系列との、最長共通部分列長を取得する（ステップＳ７０７）。

【０１３４】

続いて、イベント照合部１５１は、最長共通部分列長を照合用イベント系列長で除算した値が、あらかじめ指定した値よりも大きいか否かを判定する（ステップＳ７０８）。あらかじめ指定した値よりも大きい場合には（ステップＳ７０８肯定）、イベント照合部１５１は、判定対象イベント系列と照合用イベント系列とが合致したと判定する（ステップＳ７０９）。

【０１３５】

一方、あらかじめ指定した値よりも大きくない場合には（ステップＳ７０８否定）、イベント照合部１５１は、判定対象イベント系列と照合用イベント系列とが合致しなかったと判定する（ステップＳ７１０）。

【０１３６】

そして、イベント照合部１５１は、照合用イベント系列、判定対象イベント系列のホスト情報、判定結果を出力する（ステップＳ７１１）。そして、イベント照合部１５１は、処理をステップＳ７０５へ移行させる。

【０１３７】

このように、イベント照合部１５１は、マルウェアの特徴的な通信をもとに抽出されたイベント系列をシグネチャ系列として、検知対象となるイベント系列との照合処理を実行する。これにより、検出装置１００は、類似の通信パターンを有するマルウェアに感染した端末を少ない誤検知で検知することができる。

【０１３８】

なお、実施形態と同様の処理は、監視対象ＮＷ内の端末装置と検出装置１００とを備える検出システムによって実現されてもよい。この場合、端末装置は、監視対象ＮＷにおいて所定のイベントを発生させ、検出装置１００は、当該端末装置ごとにイベントを取得する。また、マルウェア感染端末の検出システムには、仮想的にマルウェアの通信を発生させる情報処理装置が含まれてもよい。この場合、検出システムに含まれる検出装置１００は、情報処理装置が発生させたイベントをマルウェア分析結果として取得する。

【０１３９】

［効果］

上述してきたように、実施形態に係る検出装置１００は、監視対象ＮＷの通信およびマルウェアが発生させる通信のうち、通信を特徴付けるルールに合致する事象であるイベントであって、監視対象ＮＷの端末およびマルウェアを区別する識別子ごとに取得されたイベントから、イベントの発生順序を踏まえて時系列に基づいて形成されるイベント系列を生成する。そして、検出装置１００は、マルウェアが発生させる通信に基づくイベント系列間での類似度を算出し、類似度が一定以上のイベント系列同士を同一クラスタに設定し、同一クラスタに属するイベント系列間で共通的に出現するイベントを取り出し、取り出

したイベントを時系列順に結合した一定の長さ以上のイベント系列を共通イベント系列として抽出する。さらに、検出装置１００は、複数の共通イベント系列同士で類似する共通イベント系列から、出現頻度の多いイベント間の関係からなる代表イベント系列を検知用イベント系列として抽出する。そして、検出装置１００は、監視対象ＮＷの通信に基づくイベント系列である判定対象イベント系列と、検知用イベント系列との合致部分の長さの、検知用イベント系列の長さに対する割合である第一の合致率によって、イベント系列同士が合致していると判定された場合に、監視対象ＮＷにマルウェア感染端末が存在していることを検知する。

【０１４０】

これにより、実施形態に係る検出装置１００は、監視対象ＮＷで照合すべきパターンとなるシグネチャを削減し、照合にかかる時間を削減することができる。また、検出装置１００は、単一のマルウェアの通信をシグネチャとするのではなく、マルウェア通信分析結果に基づいてクラスタリングされたイベント系列のうち共通したイベント系列を代表するイベント系列の集合を検知用イベント系列（シグネチャ）とする。これにより、検出装置１００は、既知のマルウェアだけでなく、既知のマルウェアに類似する通信を行う亜種のマルウェアについても検出することが可能となる。

【０１４１】

また、検出装置１００は、共通イベント系列から、イベントをノード、イベント間の発生順序をエッジ、イベントの前後関係の出現回数をエッジの重みとする有向グラフを生成する。そして、検出装置１００は、有向グラフの単純道ごとに重みの総和を計算し、最大の重みを示す単純道を代表イベント系列とする。これにより、検出装置１００は、類似の共通イベント系列から最も効率的に検知を行える代表イベント系列を抽出することが可能となる。

【０１４２】

また、検出装置１００は、有向グラフの単純道に含まれるエッジのうち、重みが所定の閾値以上であるエッジについて重みの総和を計算し、重みの総和が最大となる単純道を代表イベント系列とする。これにより、検出装置１００は、代表イベント系列抽出対象となるイベントを予め削減することができるため、処理を削減することが可能となる。

【０１４３】

また、検出装置１００は、同一クラスタにおける最長の共通イベント系列の長さに対する代表イベント系列の長さの割合が、所定の値よりも小さい場合、同一クラスタにおける共通イベント系列を代表イベント系列とする。これにより、クラスタを代表するイベント系列の抽出が困難な場合には、極端に短い代表イベント系列が生成されることを防止でき、誤検知を低減することが可能となる。

【０１４４】

また、検出装置１００は、監視対象ネットワークの通信に基づくイベント系列である判定対象イベント系列と検知用イベント系列との合致部分の長さの、検知用イベント系列の長さに対する割合である第一の合致率と、検知用イベント系列の長さの、検知用イベント系列が属するクラスタにおける最長の共通イベント系列の長さに対する割合である第二の合致率と、を乗じた値が所定の閾値以上である場合は、判定対象イベント系列と検知用イベント系列とが合致していると判定し、監視対象ネットワークにマルウェア感染端末が存在していることを検知する。これによって、誤検知を低減することができる。

【０１４５】

（構成等）

なお、図示した各装置の各構成要素は機能概念的なものであり、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、各装置の分散・統合の具体的形態は図示のものに限られず、その全部または一部を、各種の負荷や使用状況などに応じて、任意の単位で機能的または物理的に分散・統合して構成することができる。さらに、各装置にて行なわれる各処理機能は、その全部または任意の一部が、ＣＰＵおよび当該ＣＰＵにて解析実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウ

ェアとして実現され得る。

【０１４６】

また、本実施形態において説明した各処理のうち、自動的におこなわれるものとして説明した処理の全部または一部を手動的におこなうこともでき、あるいは、手動的におこなわれるものとして説明した処理の全部または一部を公知の方法で自動的におこなうこともできる。この他、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、特記する場合を除いて任意に変更することができる。

【０１４７】

（プログラム）

また、上記実施形態に係る検出装置１００が実行する処理をコンピュータが実行可能な言語で記述したプログラムを作成することもできる。この場合、コンピュータがプログラムを実行することにより、上記実施形態と同様の効果を得ることができる。さらに、かかるプログラムをコンピュータに読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータに読み込ませて実行することにより上記実施形態と同様の処理を実現してもよい。以下に、検出装置１００と同様の機能を実現する検出プログラムを実行するコンピュータの一例を説明する。

【０１４８】

図１７は、マルウェア感染端末の検出プログラムを実行するコンピュータを示す図である。図１７に示すように、コンピュータ１０００は、例えば、メモリ１０１０と、ＣＰＵ（Central Processing Unit）１０２０と、ハードディスクドライブインタフェース１０３０と、ディスクドライブインタフェース１０４０と、シリアルポートインタフェース１０５０と、ビデオアダプタ１０６０と、ネットワークインタフェース１０７０とを有する。これらの各部は、バス１０８０によって接続される。

【０１４９】

メモリ１０１０は、ＲＯＭ（Read Only Memory）１０１１およびＲＡＭ（Random Access Memory）１０１２を含む。ＲＯＭ１０１１は、例えば、ＢＩＯＳ（Basic Input Output System）等のブートプログラムを記憶する。ハードディスクドライブインタフェース１０３０は、ハードディスクドライブ１０９０に接続される。ディスクドライブインタフェース１０４０は、ディスクドライブ１０４１に接続される。ディスクドライブ１０４１には、例えば、磁気ディスクや光ディスク等の着脱可能な記憶媒体が挿入される。シリアルポートインタフェース１０５０には、例えば、マウス１１１０およびキーボード１１２０が接続される。ビデオアダプタ１０６０には、例えば、ディスプレイ１１３０が接続される。

【０１５０】

ここで、図１７に示すように、ハードディスクドライブ１０９０は、例えば、ＯＳ１０９１、アプリケーションプログラム１０９２、プログラムモジュール１０９３およびプログラムデータ１０９４を記憶する。上記実施形態で説明した各情報は、例えばハードディスクドライブ１０９０やメモリ１０１０に記憶される。

【０１５１】

また、検出プログラムは、例えば、コンピュータ１０００によって実行される指令が記述されたプログラムモジュールとして、ハードディスクドライブ１０９０に記憶される。具体的には、上記実施形態で説明した検出装置１００が実行する各処理が記述されたプログラムモジュールが、ハードディスクドライブ１０９０に記憶される。

【０１５２】

また、検出プログラムによる情報処理に用いられるデータは、プログラムデータとして、例えば、ハードディスクドライブ１０９０に記憶される。そして、ＣＰＵ１０２０が、ハードディスクドライブ１０９０に記憶されたプログラムモジュール１０９３やプログラムデータ１０９４を必要に応じてＲＡＭ１０１２に読み出して、上述した各手順を実行する。

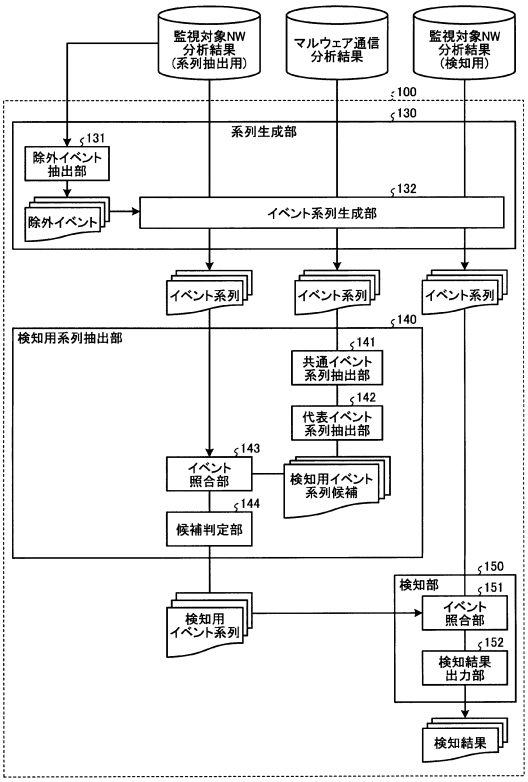
【０１５３】
なお、検出プログラムに係るプログラムモジュール１０９３やプログラムデータ１０９４は、ハードディスクドライブ１０９０に記憶される場合に限られず、例えば、着脱可能な記憶媒体に記憶されて、ディスクドライブ１０４１等を介してＣＰＵ１０２０によって読み出されてもよい。あるいは、検出プログラムに係るプログラムモジュール１０９３やプログラムデータ１０９４は、ＬＡＮ（Local Area Network）やＷＡＮ（Wide Area Network）等のネットワークを介して接続された他のコンピュータに記憶され、ネットワークインタフェース１０７０を介してＣＰＵ１０２０によって読み出されてもよい。

【符号の説明】

【０１５４】

１００　検出装置
１３０　系列生成部
１３１　除外イベント抽出部
１３２　イベント系列生成部
１４０　検知用系列抽出部
１４１　共通イベント系列抽出部
１４２　代表イベント系列抽出部
１４３　イベント照合部
１４４　候補判定部
１５０　検知部
１５１　イベント照合部
１５２　検知結果出力部

【図１】



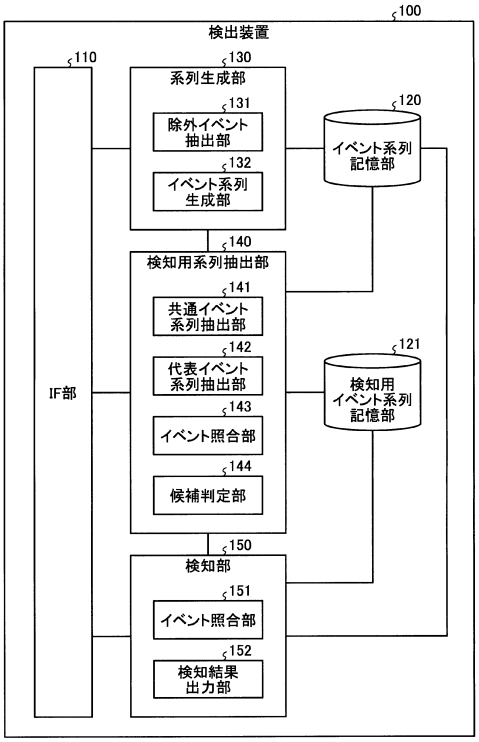
【図２】

監視対象NW内ホストの識別子	イベント	イベント発生時刻	…
…	…	…	…
192.168.10.11	特定の通信先との通信検知	2014/10/15 12:20:12	…
192.168.10.11	X時間内にy回の通信検知	2014/10/15 12:20:15	…
192.168.10.11	悪性なデータの送信検知	2014/10/15 12:20:20	…
…	…	…	…

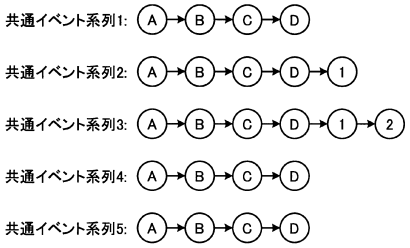
【図３】

マルウェア識別子	イベント	イベント発生時刻	…
…	…	…	…
e5fa44f2b31c1fb553b6021e7360d07d5d91ff5e	特定の通信先との通信検知	2014/9/16 22:25:14	…
e5fa44f2b31c1fb553b6021e7360d07d5d91ff5e	X時間内にY回の通信検知	2014/9/16 22:25:16	…
e5fa44f2b31c1fb553b6021e7360d07d5d91ff5e	悪性なデータの送信検知	2014/9/16 22:25:21	…
…	…	…	…

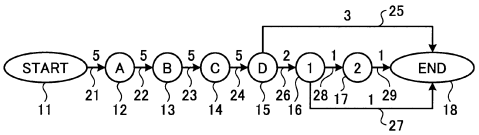
【図４】



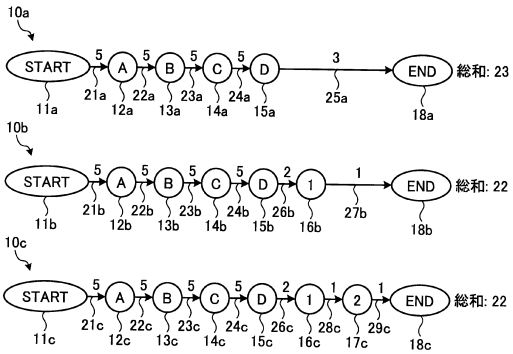
【図５】



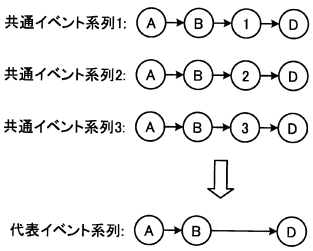
【図６】



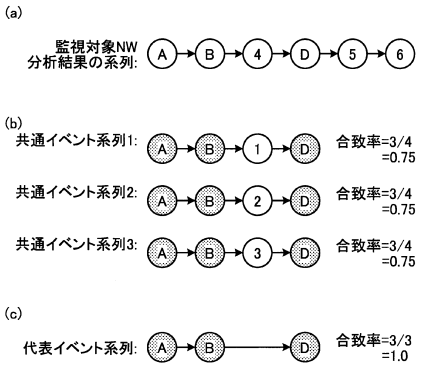
【図７】



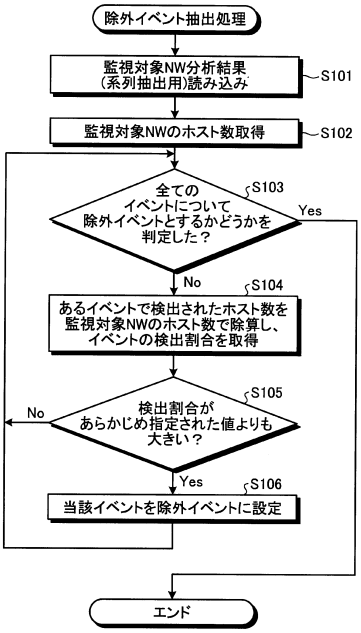
【図８】



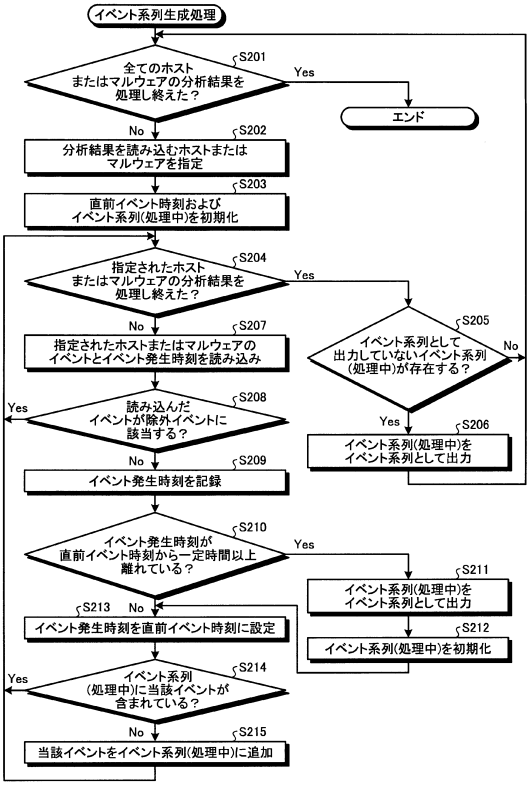
【図９】



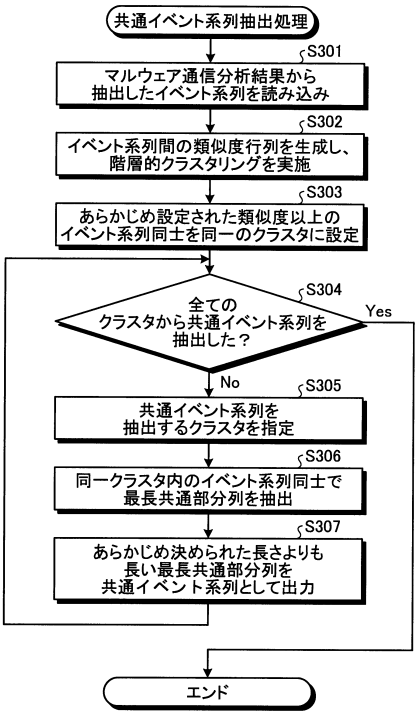
【図１０】



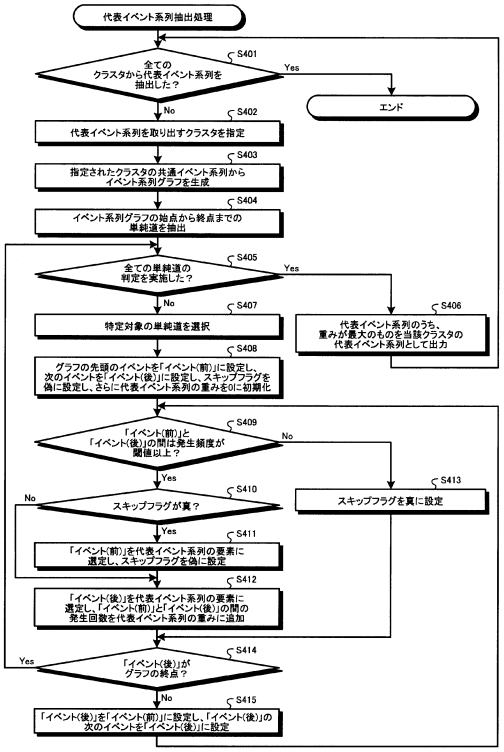
【図１１】



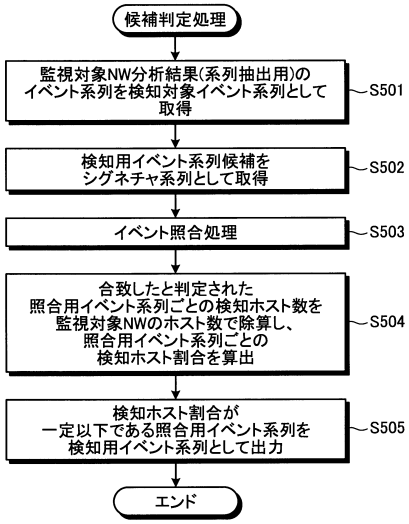
【図１２】



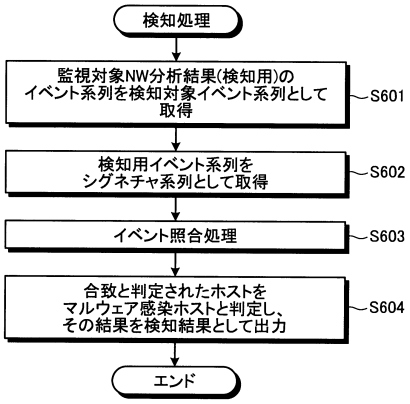
【図１３】



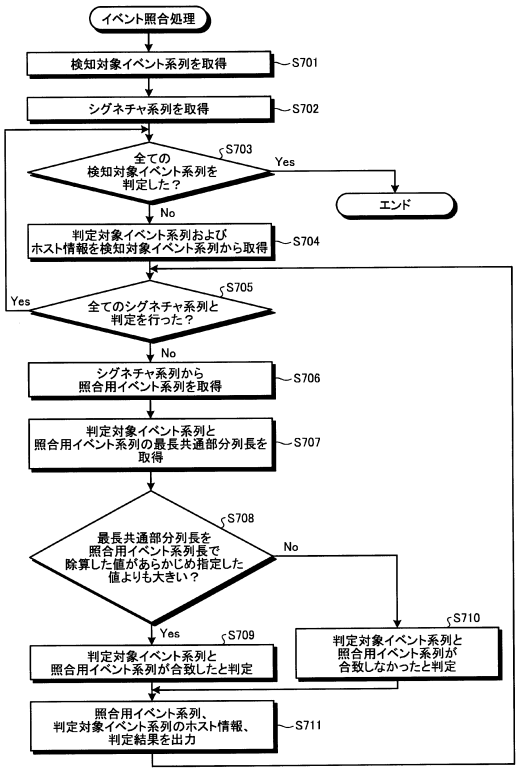
【図１４】



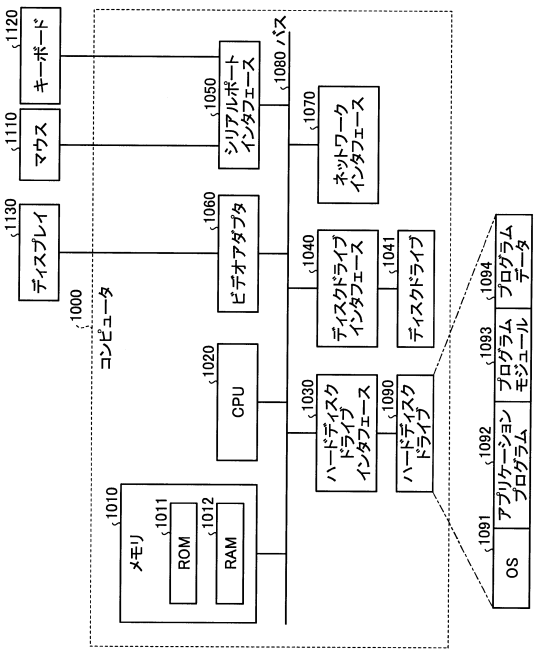
【図１５】



【図１６】



【図１７】



フロントページの続き

(56)参考文献 国際公開第２０１６／０７６３３４（WO，Ａ１）
米国特許出願公開第２０１４／０１２３２８０（ＵＳ，Ａ１）
米国特許出願公開第２０１３／０２７６１１７（ＵＳ，Ａ１）
米国特許出願公開第２００７／０１３６４５５（ＵＳ，Ａ１）
神谷 和憲 ほか，Ｆｉｒｅｗａｌｌログを用いたマルウェア感染端末の検知手法，情報処理学会第７７回（平成２７年）全国大会講演論文集，日本，一般社団法人情報処理学会，２０１５年3月１７日，４Ｅ－０３，ｐｐ．３－４３３～３－４３４

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ ２１／５６