(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5656136号
(P5656136)
(45)発行日　平成27年1月21日(2015.1.21)　(24)登録日　平成26年12月5日(2014.12.5)

(51)Int.Cl.　F I
G06F 21/56 (2013.01)　G06F 21/00 156G

請求項の数 20 (全 19 頁)

(21)出願番号	特願2013-508131 (P2013-508131)	(73)特許権者	501113353
(86)(22)出願日	平成23年4月25日 (2011.4.25)		シマンテック コーポレーション
(65)公表番号	特表2013-529335 (P2013-529335A)		Ｓｙｍａｎｔｅｃ Ｃｏｒｐｏｒａｔｉｏｎ
(43)公表日	平成25年7月18日 (2013.7.18)		アメリカ合衆国, カリフォルニア州 ９４０４３, マウンテン ビュー, エリス ストリート ３５０
(86)国際出願番号	PCT/US2011/033829	(74)代理人	100107456
(87)国際公開番号	WO2011/137083		弁理士 池田 成人
(87)国際公開日	平成23年11月3日 (2011.11.3)	(74)代理人	100148596
審査請求日	平成26年4月16日 (2014.4.16)		弁理士 山口 和弘
(31)優先権主張番号	12/769,262	(74)代理人	100123995
(32)優先日	平成22年4月28日 (2010.4.28)		弁理士 野田 雅一
(33)優先権主張国	米国(US)		
早期審査対象出願			

最終頁に続く

(54)【発明の名称】クラスタリングを使用した行動シグネチャの生成

(57)【特許請求の範囲】
【請求項１】
悪意のあるソフトウェア（マルウェア）を検出するために行動シグネチャを生成するコンピュータ実装方法であって、
コンピュータを使用して、マルウェアデータセットにマルウェアのマルウェア挙動トレースを収集する工程であり、前記マルウェア挙動トレースは、前記マルウェアによって実行された連続挙動について説明する、工程と、
コンピュータを使用して、グッドウェアデータセットにグッドウェアのグッドウェア挙動トレースを収集する工程であり、前記グッドウェア挙動トレースは、前記グッドウェアによって実行された連続挙動について説明する、工程と、
前記マルウェアに対する前記マルウェア挙動トレースを正規化してマルウェア挙動シーケンスを生成する工程と、
前記グッドウェアに対する前記グッドウェア挙動トレースを正規化してグッドウェア挙動シーケンスを生成する工程と、
同様のマルウェア挙動シーケンス及びグッドウェア挙動シーケンスをクラスタにまとめてクラスタリングする工程であり、前記クラスタ内の前記マルウェア挙動シーケンスは、マルウェアファミリの挙動について説明し、前記マルウェアファミリは、関連付けられる一連のマルウェアを含む、工程と、
前記クラスタを分析して前記マルウェアファミリのみに共通の挙動サブシーケンスを特定する工程と、

前記マルウェアファミリのみに共通の前記挙動サブシーケンスを使用して前記マルウェアファミリに対する行動シグネチャを作成する工程と
を含む、方法。
【請求項２】
前記行動シグネチャを作成する工程の後に、
以前は前記マルウェアデータセットのメンバーではない、新しいマルウェアを特定する工程と、
前記新しいマルウェアを前記マルウェアデータセットに追加する工程と、
前記新しいマルウェアの挙動トレースを収集する工程と、
前記新しいマルウェアに対する前記挙動トレースを正規化して前記新しいマルウェアに対する挙動シーケンスを生成する工程と、
前記新しいマルウェアに対する前記挙動シーケンスがマルウェア挙動シーケンス及びグッドウェア挙動シーケンスのクラスタと整合するかどうか判断する工程と、
前記クラスタと整合する前記新しいマルウェアに対する前記挙動シーケンスに応じて、前記クラスタを分析して、前記クラスタ内の前記マルウェア挙動シーケンスおよび前記新しいマルウェアに対する前記挙動シーケンスのみに共通の新しい挙動サブシーケンスを特定する工程と、
前記新しい挙動サブシーケンスを使用して前記マルウェアファミリに対する新しい行動シグネチャを作成する工程と
をさらに含む、請求項１に記載の方法。
【請求項３】
前記マルウェア挙動トレース及びグッドウェア挙動トレースは、実行されたアプリケーションプログラミングインターフェース（ＡＰＩ）呼び出しについて説明する、請求項１に記載の方法。
【請求項４】
前記マルウェアに対する前記挙動トレースを正規化してマルウェア挙動シーケンスを生成する工程は、
マルウェア挙動トレース内の関連連続挙動をまとめて分類してオペレーションを形成する工程を含み、マルウェア挙動シーケンスは、１つまたは複数の連続挙動および１つまたは複数のオペレーションを含む、請求項１に記載の方法。
【請求項５】
同様のマルウェア挙動シーケンス及びグッドウェア挙動シーケンスをクラスタにまとめてクラスタリングする工程は、
前記マルウェア挙動シーケンス及びグッドウェア挙動シーケンスの間で編集距離を決定する工程と、
前記決定された編集距離に応じて前記マルウェア挙動シーケンス及びグッドウェア挙動シーケンスをクラスタリングする工程と
を含む、請求項１に記載の方法。
【請求項６】
前記クラスタを分析して前記マルウェアファミリのみに共通の挙動サブシーケンスを特定する工程は、
前記クラスタ内の前記マルウェア挙動シーケンスのみに共通の複数の候補サブシーケンスを特定する工程と、
前記マルウェア挙動シーケンス内のどこで前記候補サブシーケンスが現れるかを特定する工程と、
前記マルウェア挙動シーケンス内のどこで前記候補サブシーケンスが現れるかに応じて、前記候補サブシーケンスの中から前記挙動サブシーケンスを選択する工程と
を含む、請求項１に記載の方法。
【請求項７】
前記マルウェア挙動シーケンス内のどこで前記候補サブシーケンスが現れるかに応じて

、前記候補サブシーケンスの中から前記挙動サブシーケンスを選択する工程は、
他の候補サブシーケンスより早期に前記マルウェア挙動シーケンスに現れた前記挙動サブシーケンスに応じて前記挙動サブシーケンスを選択する工程を含む、請求項６に記載の方法。
【請求項８】
クライアントのセキュリティモジュールに前記行動シグネチャを分配する工程をさらに含み、前記セキュリティモジュールは、前記行動シグネチャを使用して前記クライアント側に存在するマルウェアを検出するよう適合される、請求項１に記載の方法。
【請求項９】
悪意のあるソフトウェア（マルウェア）を検出するために行動シグネチャを生成するコンピュータシステムであって、
マルウェアデータセットにマルウェアのマルウェア挙動トレースを収集する工程であり、前記マルウェア挙動トレースは、前記マルウェアによって実行された連続挙動について説明する、工程と、
グッドウェアデータセットにグッドウェアのグッドウェア挙動トレースを収集する工程であり、前記グッドウェア挙動トレースは、前記グッドウェアによって実行された連続挙動について説明する、工程と、
前記マルウェアに対する前記マルウェア挙動トレースを正規化してマルウェア挙動シーケンスを生成する工程と、
前記グッドウェアに対する前記グッドウェア挙動トレースを正規化してグッドウェア挙動シーケンスを生成する工程と、
同様のマルウェア挙動シーケンス及びグッドウェア挙動シーケンスをクラスタにまとめてクラスタリングする工程であり、前記クラスタ内の前記マルウェア挙動シーケンスは、マルウェアファミリの挙動について説明し、前記マルウェアファミリは、関連付けられる一連のマルウェアを含む、工程と、
前記クラスタを分析して前記マルウェアファミリのみに共通の挙動サブシーケンスを特定する工程と、
前記マルウェアファミリのみに共通の前記挙動サブシーケンスを使用して前記マルウェアファミリに対する行動シグネチャを作成する工程と
を含む工程を実行するために実行可能なコンピュータプログラムモジュールを格納する非一時的なコンピュータ可読記憶媒体と、
前記コンピュータプログラムモジュールを実行するためのコンピュータプロセッサと
を備える、コンピュータシステム。
【請求項１０】
前記工程は、
前記行動シグネチャを作成する工程の後に、
以前は前記マルウェアデータセットのメンバーではない、新しいマルウェアを特定する工程と、
前記新しいマルウェアを前記マルウェアデータセットに追加する工程と、
前記新しいマルウェアの挙動トレースを収集する工程と、
前記新しいマルウェアに対する前記挙動トレースを正規化して前記新しいマルウェアに対する挙動シーケンスを生成する工程と、
前記新しいマルウェアに対する前記挙動シーケンスがマルウェア挙動シーケンス及びグッドウェア挙動シーケンスのクラスタと整合するかどうか判断する工程と、
前記クラスタと整合する前記新しいマルウェアに対する前記挙動シーケンスに応じて、前記クラスタを分析して、前記クラスタ内の前記マルウェア挙動シーケンスおよび前記新しいマルウェアに対する前記挙動シーケンスのみに共通の新しい挙動サブシーケンスを特定する工程と、
前記新しい挙動サブシーケンスを使用して前記マルウェアファミリに対する新しい行動シグネチャを作成する工程と

をさらに含む、請求項９に記載のコンピュータシステム。
【請求項１１】
前記マルウェアに対する前記挙動トレースを正規化してマルウェア挙動シーケンスを生成する工程は、
マルウェア挙動トレース内の関連連続挙動をまとめて分類してオペレーションを形成する工程を含み、マルウェア挙動シーケンスは、１つまたは複数の連続挙動および１つまたは複数のオペレーションを含む、請求項９に記載のコンピュータシステム。
【請求項１２】
前記クラスタを分析して前記マルウェアファミリのみに共通の挙動サブシーケンスを特定する工程は、
前記クラスタ内の前記マルウェア挙動シーケンスのみに共通の複数の候補サブシーケンスを特定する工程と、
前記マルウェア挙動シーケンス内のどこで前記候補サブシーケンスが現れるかを特定する工程と、
前記マルウェア挙動シーケンス内のどこで前記候補サブシーケンスが現れるかに応じて、前記候補サブシーケンスの中から前記挙動サブシーケンスを選択する工程と
を含む、請求項９に記載のコンピュータシステム。
【請求項１３】
前記マルウェア挙動シーケンス内のどこで前記候補サブシーケンスが現れるかに応じて、前記候補サブシーケンスの中から前記挙動サブシーケンスを選択する工程は、
他の候補サブシーケンスより早期に前記マルウェア挙動シーケンスに現れた前記挙動サブシーケンスに応じて前記挙動サブシーケンスを選択する工程を含む、請求項１２に記載のコンピュータシステム。
【請求項１４】
悪意のあるソフトウェア（マルウェア）を検出するために行動シグネチャを生成するための実行可能なコンピュータプログラムであって、
コンピュータに、
マルウェアデータセットにマルウェアのマルウェア挙動トレースを収集するための機能であり、前記マルウェア挙動トレースは、前記マルウェアによって実行された連続挙動について説明する、機能と、
グッドウェアデータセットにグッドウェアのグッドウェア挙動トレースを収集するための機能であり、前記グッドウェア挙動トレースは、前記グッドウェアによって実行された連続挙動について説明する、機能と、
前記マルウェアに対する前記マルウェア挙動トレースを正規化してマルウェア挙動シーケンスを生成するための機能と、
前記グッドウェアに対する前記グッドウェア挙動トレースを正規化してグッドウェア挙動シーケンスを生成するための機能と、
同様のマルウェア挙動シーケンス及びグッドウェア挙動シーケンスをクラスタにまとめてクラスタリングするための機能であり、前記クラスタ内の前記マルウェア挙動シーケンスは、マルウェアファミリの挙動について説明し、前記マルウェアファミリは、関連付けられる一連のマルウェアを含む、機能と、
前記クラスタを分析して前記マルウェアファミリのみに共通の挙動サブシーケンスを特定するための機能と、
前記マルウェアファミリのみに共通の前記挙動サブシーケンスを使用して前記マルウェアファミリに対する行動シグネチャを作成するための機能と
を実現させる、コンピュータプログラム。
【請求項１５】
前記行動シグネチャの作成の後に、
以前は前記マルウェアデータセットのメンバーではない、新しいマルウェアを特定する機能と、

前記新しいマルウェアを前記マルウェアデータセットに追加する機能と、
前記新しいマルウェアの挙動トレースを収集する機能と、
前記新しいマルウェアに対する前記挙動トレースを正規化して前記新しいマルウェアに対する挙動シーケンスを生成するための機能と、
前記新しいマルウェアに対する前記挙動シーケンスがマルウェア挙動シーケンス及びグッドウェア挙動シーケンスのクラスタと整合するかどうか判断するための機能と、
前記クラスタと整合する前記新しいマルウェアに対する前記挙動シーケンスに応じて、前記クラスタを分析して、前記クラスタ内の前記マルウェア挙動シーケンスおよび前記新しいマルウェアに対する前記挙動シーケンスのみに共通の新しい挙動サブシーケンスを特定するための機能と、
前記新しい挙動サブシーケンスを使用して前記マルウェアファミリに対する新しい行動シグネチャを作成するための機能と
をさらに実現させる、請求項１４に記載のコンピュータプログラム。
【請求項１６】
前記マルウェアに対する前記挙動トレースを正規化してマルウェア挙動シーケンスを生成する機能は、
マルウェア挙動トレース内の関連連続挙動をまとめて分類してオペレーションを形成する機能を含み、マルウェア挙動シーケンスは、１つまたは複数の連続挙動および１つまたは複数のオペレーションを含む、請求項１４に記載のコンピュータプログラム。
【請求項１７】
前記クラスタを分析して前記マルウェアファミリのみに共通の挙動サブシーケンスを特定する機能は、
前記クラスタ内の前記マルウェア挙動シーケンスのみに共通の複数の候補サブシーケンスを特定する機能と、
前記マルウェア挙動シーケンス内のどこで前記候補サブシーケンスが現れるかを特定する機能と、
前記マルウェア挙動シーケンス内のどこで前記候補サブシーケンスが現れるかに応じて、前記候補サブシーケンスの中から前記挙動サブシーケンスを選択する機能と
を含む、請求項１４に記載のコンピュータプログラム。
【請求項１８】
コンピュータを使用して、前記マルウェア挙動トレースを収集する工程は、
仮想コンピューティング環境で前記マルウェアの実行をエミュレートする工程と、
前記マルウェアによって行われる、オペレーティングシステムへのアプリケーションプログラミングインターフェース（ＡＰＩ）呼び出しのシーケンスを記録するために、エミュレートされた前記マルウェアの実行をモニタする工程と、
を含む、請求項１に記載の方法。
【請求項１９】
前記マルウェア挙動トレースを収集する工程は、
仮想コンピューティング環境で前記マルウェアの実行をエミュレートする工程と、
前記マルウェアによって行われる、オペレーティングシステムへのアプリケーションプログラミングインターフェース（ＡＰＩ）呼び出しのシーケンスを記録するために、エミュレートされた前記マルウェアの実行をモニタする工程と、
を含む、請求項９に記載のコンピュータシステム。
【請求項２０】
前記マルウェア挙動トレースを収集する機能は、
仮想コンピューティング環境で前記マルウェアの実行をエミュレートする機能と、
前記マルウェアによって行われる、オペレーティングシステムへのアプリケーションプログラミングインターフェース（ＡＰＩ）呼び出しのシーケンスを記録するために、エミュレートされた前記マルウェアの実行をモニタする機能と、
を含む、請求項１４に記載のコンピュータプログラム。

【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、概して、コンピュータセキュリティに関し、具体的には、悪意のあるソフトウェアを検出するための行動シグネチャの生成に関する。
【背景技術】
【０００２】
最新のコンピュータを攻撃することができる多種多様の悪意のあるソフトウェア（マルウェア）が存在する。マルウェアの脅威は、コンピュータウイルス、ワーム、トロイの木馬プログラム、スパイウェア、アドウェア、クライムウェアおよびフィッシング詐欺ウェブサイトを含む。マルウェアは、例えば、ログイン、パスワード、銀行口座名およびクレジットカード番号などの重要な情報をひそかに取得することができる。同様に、マルウェアは、危殆化したコンピュータに対する攻撃者のアクセスおよび制御を可能にする隠しインターフェースを提供することができる。
【０００３】
最近のマルウェアは、相対的に少数のコンピュータのみを標的として配信される場合が多い。例えば、トロイの木馬プログラムは、特定企業の特定部門のコンピュータを標的とするよう設計され得る。そのようなマルウェアは、同じマルウェアの例があまり存在しないため、セキュリティソフトウェアが検出することは難しく、セキュリティソフトウェアは、それを認識するように構成されていない恐れがある。さらに、マルウェアは検出を回避するよう設計されたポリモーフィズムを含むことができるため、大量に分配されたマルウェアの検出でさえ難しくなりつつある。
【発明の概要】
【発明が解決しようとする課題】
【０００４】
マルウェアの検出が困難となるのに応じて、セキュリティソフトウェアは、ヒューリスティックベースの検出へ向けて発展している。このタイプの検出は、悪意のある挙動を示す挙動について説明する行動シグネチャなどの一連のヒューリスティックを使用してマルウェアを特定する。しかし、膨大な量の異なるマルウェアは、マルウェアに対する行動シグネチャの維持を困難にする。多くの行動シグネチャを有することにより、シグネチャの維持および使用が複雑化する。例えば、１つのソフトウェアが悪意のあるものであるかどうかを判断するために使用される分析は、より複雑性を増し、この判断を下すために消費される計算リソースは、行動シグネチャの数の増加に伴い増加する。
【課題を解決するための手段】
【０００５】
上記のおよび他の問題は、マルウェアを検出するための行動シグネチャを生成するための方法、コンピュータシステムおよびコンピュータ可読記憶媒体によって対処される。本方法の実施形態は、コンピュータを使用して、マルウェアデータセットにマルウェアの挙動トレースを収集する工程を含む。挙動トレースは、マルウェアによって実行された連続挙動について説明する。本方法は、マルウェアに対する挙動トレースを正規化してマルウェア挙動シーケンスを生成する工程と、同様のマルウェア挙動シーケンスをクラスタにまとめてクラスタリングする工程とをさらに含む。クラスタ内のマルウェア挙動シーケンスは、マルウェアファミリの挙動について説明する。それに加えて、本方法は、クラスタを分析してマルウェアファミリに共通の挙動サブシーケンスを特定する工程と、挙動サブシーケンスを使用してマルウェアファミリに対する行動シグネチャを作成する工程とを含む。
【０００６】
コンピュータシステムの実施形態は、工程を実行するための実行可能なコンピュータプログラムモジュールを格納する非一時的なコンピュータ可読記憶媒体を備える。工程は、マルウェアデータセットにマルウェアの挙動トレースを収集する工程を含む。挙動トレー

スは、マルウェアによって実行された連続挙動について説明する。工程は、マルウェアに対する挙動トレースを正規化してマルウェア挙動シーケンスを生成する工程と、同様のマルウェア挙動シーケンスをクラスタにまとめてクラスタリングする工程とをさらに含む。クラスタ内のマルウェア挙動シーケンスは、マルウェアファミリの挙動について説明する。それに加えて、工程は、クラスタを分析してマルウェアファミリに共通の挙動サブシーケンスを特定する工程と、挙動サブシーケンスを使用してマルウェアファミリに対する行動シグネチャを作成する工程とを含む。また、コンピュータシステムは、コンピュータプログラムモジュールを実行するためのコンピュータプロセッサも備える。

【０００７】

媒体の実施形態は、マルウェアデータセットにマルウェアの挙動トレースを収集するための実行可能なコンピュータプログラムモジュールを格納する非一時的なコンピュータ可読記憶媒体を備える。挙動トレースは、マルウェアによって実行された連続挙動について説明する。また、モジュールは、マルウェアに対する挙動トレースを正規化してマルウェア挙動シーケンスを生成し、同様のマルウェア挙動シーケンスをクラスタにまとめてクラスタリングする。クラスタ内のマルウェア挙動シーケンスは、マルウェアファミリの挙動について説明する。それに加えて、モジュールは、クラスタを分析してマルウェアファミリに共通の挙動サブシーケンスを特定し、挙動サブシーケンスを使用してマルウェアファミリに対する行動シグネチャを作成する。

【図面の簡単な説明】

【０００８】

【図１】一実施形態によるコンピューティング環境の高レベルのブロック図である。

【図２】セキュリティサーバまたはクライアントとして使用するための典型的なコンピュータを示す高レベルのブロック図である。

【図３】一実施形態によるセキュリティサーバのシグネチャ生成モジュールの詳細な概観を示す高レベルのブロック図である。

【図４】行動シグネチャを生成および分配するために、セキュリティサーバの一実施形態によって実行される工程を示すフローチャートである。

【図５】新しいマルウェアサンプルを検出するための更新された行動シグネチャを生成および分配するために、セキュリティサーバの一実施形態によって実行される工程を示すフローチャートである。

【０００９】

図面は、単なる例示を目的として一実施形態を描写する。当業者であれば、本明細書で示される構造および方法の代替の実施形態を、本明細書に記載される原理から逸脱することなく使用できることが以下の説明から容易に理解されよう。

【発明を実施するための形態】

【００１０】

図１は、一実施形態によるコンピューティング環境１００の高レベルのブロック図である。図１は、ネットワーク１１４によって接続されたセキュリティサーバ１１０および３つのクライアント１１２を示す。説明を簡素かつ明確にするため、図１では３つのクライアント１１２のみが示される。コンピューティング環境１００の実施形態は、数千または数百万ものクライアント１１２を有することができる。また、いくつかの実施形態は、複数のセキュリティサーバ１１０も有する。

【００１１】

クライアント１１２は、悪意のあるソフトウェアをホストすることができる電子デバイスである。一実施形態では、クライアント１１２は、例えば、Ｍｉｃｒｏｓｏｆｔ Ｗｉｎｄｏｗｓ互換オペレーティングシステム（ＯＳ）、Ａｐｐｌｅ ＯＳ Ｘおよび／またはＬｉｎｕｘの分配を実行する従来のコンピュータシステムである。また、クライアント１１２は、携帯情報端末（ＰＤＡ）、携帯電話、テレビゲームシステムなどのコンピュータ機能性を有する別のデバイスでもあり得る。クライアント１１２は、通常、悪意のあるソフトウェアをホストすることができる多数のコンピュータファイルを格納する。

【００１２】
「マルウェア」と呼ばれる場合がある悪意のあるソフトウェアは、一般に、クライアント１１２上でひそかに実行するソフトウェアまたは何らかの秘密の機能性を有するソフトウェアとして定義される。マルウェアは、多くの形態を取ることができ、正当なファイルに添付される寄生ウイルス、コンピュータを感染して他のコンピュータに広げるためにコンピュータのセキュリティの弱点を突くワーム、正当に見えるが実際は悪意のある隠しコードを含むトロイの木馬プログラム、ならびに、機密情報の取得もしくは広告の表示のためにキーストロークおよび／またはコンピュータ上の他の動作をモニタするスパイウェアなどが挙げられる。
【００１３】
クライアント１１２は、クライアント上のマルウェアの存在を検出するためのセキュリティモジュール１１６を実行する。セキュリティモジュール１１６は、例えば、クライアント１１２のＯＳに組み込んでも、別々の包括的なセキュリティパッケージの一部であってもよい。一実施形態では、セキュリティモジュール１１６は、セキュリティサーバ１１０を動作する同じ実体によって提供される。セキュリティモジュール１１６は、ネットワーク１１４を介してセキュリティサーバ１１０と通信し、クライアント１１２側でマルウェアを検出するための検出データを入手する。
【００１４】
セキュリティモジュール１１６によって入手された検出データは、行動シグネチャを含む。行動シグネチャは、マルウェアの特徴を示す挙動のシーケンスについて説明する。セキュリティモジュール１１６は、クライアント側で実行するソフトウェアの挙動を観察してその挙動が行動シグネチャのいずれかと一致するかどうか判断することによって、クライアント１１２側でマルウェアを検出する。
【００１５】
また、検出データは、クライアント１１２側でマルウェアを検出する他の方法について説明するデータも含むことができる。例えば、検出データは、マルウェアの特徴を示す、コンピュータファイル内で見出されたデータのシーケンスを特定するシグネチャ文字列、ある所定のソフトウェアが悪意のあるものかどうか評価するためにセキュリティモジュール１１６が使用できるレピュテーション、および／または、マルウェア攻撃を示す可能性があるクライアント１１２の状態を特定するヒューリスティックを含むことができる。
【００１６】
一実施形態では、セキュリティモジュール１１６は、検出データを使用してクライアント１１２をモニタし、マルウェアを検出した場合にレポートを生成する。レポートは、検出されたマルウェアについて、クライアント１１２のユーザおよび／またはクライアント１１２の管理者などの別の実体に通知する。また、セキュリティモジュール１１６は、悪意のある挙動を阻止する、マルウェアを隔離する、マルウェアを除去するなど、１つまたは複数の動作を実行してマルウェアを修復することもできる。
【００１７】
セキュリティサーバ１１０は、検出データを生成してクライアント１１２に分配するよう構成されたハードウェアデバイスおよび／またはソフトウェアモジュールである。セキュリティサーバ１１０の一例としては、セキュリティソフトウェアおよびサービスをクライアント１１２のセキュリティモジュール１１６に提供するウェブベースのシステムが挙げられる。実施形態に応じて、セキュリティサーバ１１０の機能の１つまたは複数を、クラウドコンピューティング環境によって提供することができる。本明細書で使用される場合、「クラウドコンピューティング」は、ネットワーク１１４上のサービスとして、動的に拡張可能であり仮想化される場合が多いリソースが提供されるコンピューティングのスタイルを指す。クライアント１１２およびセキュリティモジュール１１６に属する機能もまた、クラウドコンピューティング環境によって提供することができる。
【００１８】
セキュリティサーバ１１０の一実施形態は、セキュリティモジュール１１６がマルウェ

アの検出に使用する行動シグネチャを生成するためのシグネチャ生成モジュール１１８を含む。シグネチャ生成モジュール１１８は、同様の挙動を実行するソフトウェアのクラスタを特定する。所定のクラスタに対して、シグネチャ生成モジュール１１８は、クラスタ内でマルウェアを検出するために使用できる行動シグネチャを特定する。さらに、新しいマルウェアが特定されると、シグネチャ生成モジュール１１８は、可能であれば、新しいマルウェアを既存のクラスタに割り当てる。新しいマルウェアが割り当てられたクラスタに対する既存のシグネチャを使用して新しいマルウェアを検出できない場合は、シグネチャ生成モジュール１１８は、新しいマルウェアを含む、クラスタ内でマルウェアを検出するために使用できる新しい行動シグネチャを生成する。セキュリティサーバ１１０の動作は、一実施形態では自動化され、これにより、手動操作なしで行動シグネチャを生成することが可能になる。

【００１９】

分配モジュール１２０は、クライアント１１２のセキュリティモジュール１１６に行動シグネチャおよび他の検出データを分配する。一実施形態では、分配モジュール１２０は、新しいシグネチャが作成されるとローリングベースで行動シグネチャを分配する。別の実施形態では、行動シグネチャは、既定のスケジュールでおよび／またはセキュリティモジュール１１６による要求と同時に提供される。

【００２０】

セキュリティサーバ１１０によって実行されたクラスタリングベースのシグネチャ生成技法は、こうして個別の行動シグネチャを効果的に使用して、複数のマルウェアサンプルを検出する。その結果、セキュリティサーバ１１０は、それぞれのシグネチャが単一のマルウェアサンプルに特有である場合に必要とされるよりも少ない行動シグネチャをセキュリティモジュール１１６に分配する。さらに、マルウェアを検出するためにセキュリティモジュール１１６によって使用される行動シグネチャのセットのサイズは、従来の技法と比べて低減される。この効率は、新しく発見されたマルウェアに対して行動シグネチャが生成される場合でさえ、維持される。したがって、シグネチャ生成技法は、効果的で高性能のマルウェア検出を実現する。

【００２１】

ネットワーク１１４は、ネットワーク上のセキュリティサーバ１１０と、クライアント１１２と、他の任意の実体との間の通信経路を表す。一実施形態では、ネットワーク１１４は、インターネットであり、標準の通信技術および／またはプロトコルを使用する。したがって、ネットワーク１１４は、イーサネット８０２．１１、マイクロ波アクセスのための世界規模の相互運用（ＷｉＭＡＸ）、３Ｇ、デジタル加入者線（ＤＳＬ）、非同期転送モード（ＡＴＭ）、ＩｎｆｉｎｉＢａｎｄ、ＰＣＩ Ｅｘｐｒｅｓｓ Ａｄｖａｎｃｅｄ Ｓｗｉｔｃｈｉｎｇなどの技術を使用してリンクを含むことができる。同様に、ネットワーク１１４上で使用されるネットワークプロトコルは、マルチプロトコルラベルスイッチング（ＭＰＬＳ）、伝送制御プロトコル／インターネットプロトコル（ＴＣＰ／ＩＰ）、ユーザデータグラムプロトコル（ＵＤＰ）、ハイパーテキスト転送プロトコル（ＨＴＴＰ）、簡易メール転送プロトコル（ＳＭＴＰ）、ファイル転送プロトコル（ＦＴＰ）などを含むことができる。ネットワーク１１４上で交換されるデータは、ハイパーテキストマークアップ言語（ＨＴＭＬ）、拡張可能マークアップ言語（ＸＭＬ）などを含む技術および／または形式を使用して表すことができる。それに加えて、セキュアソケット層（ＳＳＬ）、トランスポート層セキュリティ（ＴＬＳ）、仮想プライベートネットワーク（ＶＰＮ）、インターネットプロトコルセキュリティ（ＩＰｓｅｃ）などの従来の暗号化技術を使用して、リンクのすべてまたは一部を暗号化することができる。他の実施形態では、実体は、上記で説明されるものの代わりにまたはそれに加えて、カスタムおよび／または専用のデータ通信技術を使用する。

【００２２】

図２は、セキュリティサーバ１１０またはクライアント１１２として使用するための典型的なコンピュータ２００を示す高レベルのブロック図である。チップセット２０４と結

合されたプロセッサ２０２が示される。また、チップセット２０４には、メモリ２０６と、記憶装置２０８と、キーボード２１０と、グラフィックスアダプタ２１２と、ポインティングデバイス２１４と、ネットワークアダプタ２１６も結合される。ディスプレイ２１８は、グラフィックスアダプタ２１２と結合される。一実施形態では、チップセット２０４の機能性は、メモリコントローラハブ２２０およびＩ／Ｏコントローラハブ２２２によって提供される。別の実施形態では、メモリ２０６は、チップセット２０４の代わりにプロセッサ２０２と直接結合される。

【００２３】

記憶装置２０８は、ハードドライブ、コンパクトディスク読み取り専用メモリ（ＣＤ－ＲＯＭ）、ＤＶＤまたはソリッドステートメモリデバイスなどの非一時的なコンピュータ可読記憶媒体である。メモリ２０６は、プロセッサ２０２によって使用される命令およびデータを保持する。ポインティングデバイス２１４は、マウス、トラックボールまたは他のタイプのポインティングデバイスであり、キーボード２１０と組み合わせて使用して、コンピュータシステム２００にデータを入力する。グラフィックスアダプタ２１２は、ディスプレイ２１８上にイメージおよび他の情報を表示する。ネットワークアダプタ２１６は、ネットワーク１１６とコンピュータシステム２００を結合する。

【００２４】

当技術分野で公知の通り、コンピュータ２００は、図２に示されるものとは異なるコンポーネントおよび／またはそれ以外のコンポーネントを有することができる。それに加えて、コンピュータ２００は、示されるある特定のコンポーネントを欠くことがある。一実施形態では、セキュリティサーバとして機能するコンピュータ２００は、複数のブレードコンピュータから形成され、キーボード２１０、ポインティングデバイス２１４、グラフィックスアダプタ２１２および／またはディスプレイ２１８を欠く。さらに、記憶装置２０８は、コンピュータ２００からローカルおよび／またはリモート操作することができる（ストレージエリアネットワーク（ＳＡＮ）内で具体化されるなど）。

【００２５】

この説明では、指定の機能性を提供するためのコンピュータプログラム論理について言及する際に用語「モジュール」を使用する。モジュールは、ハードウェア、ファームウェアおよび／またはソフトウェアで実装することができる。モジュールは、通常、記憶装置２０８などのコンピュータ可読記憶媒体上に格納され、メモリ２０６にロードされ、プロセッサ２０２によって実行される。

【００２６】

図３は、一実施形態によるセキュリティサーバ１１０のシグネチャ生成モジュール１１８の詳細な概観を示す高レベルのブロック図である。図３に示されるように、シグネチャ生成モジュール１１８自体が複数のモジュールを含む。いくつかの実施形態では、機能は、本明細書に記載されるものとは異なる様式でこれらのモジュールに分配される。

【００２７】

格納モジュール３１０は、シグネチャ生成モジュール１１８によって使用されるデータを格納する。そのようなデータの例には、評価中のソフトウェア、シグネチャ生成プロセス中に作成され使用される中間データおよび作成された行動シグネチャが含まれる。データ格納モジュール３１０は、リレーショナルデータベースまたは別のタイプのデータベースを含み得る。

【００２８】

図３に示されるように、格納モジュール３１０の実施形態は、異なるクラスのソフトウェアを保持するデータセットを格納する。マルウェアデータセット３１２は、既知のマルウェアのサンプルを格納する。データセット３１２内のマルウェアは、クライアント１１２からおよび／または他のソースから入手することができる。それに加えて、格納モジュール３１０は、既知のグッドウェアのサンプルを格納するグッドウェアデータセット３１４を格納する。グッドウェアは、悪意のない（すなわち、正当な）ものとして知られているソフトウェアである。グッドウェアは、クライアント１１２上に存在する場合が多い、

ありふれたおよび／または一般のソフトウェアプログラムを含むことができる。データセット内のマルウェア３１２およびグッドウェア３１４は、まとめて「ソフトウェア」と呼ばれる。

【００２９】

挙動収集モジュール３１６は、データセット３１２、３１４にソフトウェアに挙動トレースを収集する。ソフトウェアのサンプルに対する「挙動トレース」は、ソフトウェアが実行される際に行う挙動のシーケンスである。一実施形態では、挙動収集モジュール３１６は、データセット３１２、３１４内の各ソフトウェアサンプルを適切な環境で実行する。環境は、実際のものでもシミュレートされたものでもよい。例えば、特定のソフトウェアサンプルが、Ｍｉｃｒｏｓｏｆｔ Ｗｉｎｄｏｗｓ ＯＳを改変したものを有するコンピュータ上で実行することを目的としたものであれば、挙動収集モジュール３１６は、Ｗｉｎｄｏｗｓベースのコンピュータをシミュレートする仮想コンピューティング環境でソフトウェアをエミュレートすることができる。

【００３０】

挙動収集モジュール３１６は、ソフトウェアによって実行される挙動のシーケンスを観察するのに十分長い時間、ソフトウェアを実行できるようにする。実行の長さは、例えば、実行される命令の数または実行時間の長さによって指定することができる。挙動収集モジュール３１６は、環境内で指定の実行量を実行するためにソフトウェアが必要とするいかなるパラメータおよび／または環境リソースもソフトウェアに提供する。

【００３１】

挙動収集モジュール３１６は、実行する挙動のシーケンスを収集する（記録する）ため、実行しているソフトウェアをモニタする。一実施形態では、挙動収集モジュール３１６は、実行しているソフトウェアによって行われる、ＯＳへのアプリケーションプログラミングインターフェース（ＡＰＩ）呼び出しを特にモニタし、したがって、ソフトウェアに対する挙動トレースは、そのＡＰＩ呼び出しシーケンスである。挙動収集モジュール３１６の他の実施形態は、モニタして、挙動トレースに異なるおよび／または追加のデータを含める。

【００３２】

トレース正規化モジュール３１８は、実行されたソフトウェアの挙動トレースを正規化する。正規化により、異なるトレースを比較できるように挙動トレースを標準的な表現にする。一実施形態では、トレース正規化モジュール３１８は、ＡＰＩ呼び出しで参照されるファイル経路、ドライブ名およびフォルダ位置などの異なる実行インスタンスにおいて異なり得るパラメータの標準的な表現によって挙動トレースを正規化する。例えば、パラメータは、挙動トレースから取り除く、および／または、ワイルドカードシンボルに置き換えることができる。正規化された挙動トレースは、「挙動シーケンス」と呼ばれる。

【００３３】

また、トレース正規化モジュール３１８は、挙動トレース内の挙動を正規化の一環としてオペレーションに分類することもできる。「オペレーション」は、一緒に分類される一連の関連連続挙動である。例えば、ファイルからデータを入手するソフトウェアは、「ファイルを開く」というＡＰＩ呼び出しに続いて「ファイルを読む」という呼び出しを行う場合が多い。トレース正規化モジュール３１８は、これらの２つの呼び出しを「ファイルへアクセスする」という単一のオペレーションに組み合わせることができる。この様式でオペレーションを使用することにより、その中に含まれる情報内容を実質的に変更することなく、シーケンスのサイズを低減し、したがって、後段の分析の効率を高める。正規化モジュール３１８によって実行される正規化のタイプは実施形態ごとに異なり得、および／または、異なる実施形態では省略され得る。

【００３４】

クラスタリングモジュール３２０は、同様の挙動シーケンスがまとめてクラスタリングされるように、ソフトウェアの挙動シーケンスをクラスタリングする。クラスタリングモジュール３２０は、各挙動シーケンスを他の挙動シーケンスと比較して、一連のクラスタ

を識別し、各クラスタは一連の同様の挙動シーケンスを含む。行動シーケンスはマルウェア３１２とグッドウェア３１４の両方から得られるため、結果として得られるクラスタは、単一のクラス（例えば、すべてマルウェア）のシーケンスからなるものでも、両方のクラスの組合せを含んでもよい。

【００３５】

一実施形態では、クラスタリングモジュール３２０は、類似性の尺度として編集距離を使用して（すなわち、距離測定）、編集の閾値数値内の同一の挙動シーケンスをまとめてクラスタリングする。実施形態に応じて編集距離閾値は異なり得る。それに加えて、クラスタリングモジュール３２０の異なる実施形態は、異なるおよび／または追加のクラスタリング技法および類似性の尺度を使用して、挙動シーケンスをクラスタリングする。

【００３６】

クラスタ分析モジュール３２２は、挙動シーケンスのクラスタを分析し、マルウェアに対する行動シグネチャとして使用することができるサブシーケンスを特定する。一実施形態では、クラスタ分析モジュール３２２は、各クラスタを分析し、クラスタ内の挙動シーケンスに代表されるソフトウェアのクラスを決定する。上記のように、クラスタは、マルウェア、グッドウェアまたは２つのクラスの組合せからのシーケンスを含むことができる。

【００３７】

一実施形態では、クラスタ分析モジュール３２２は、大部分がソフトウェアの一クラスから得られたシーケンスを含むクラスタを特定する。例えば、モジュール３２２は、独占的にソフトウェアの一クラスから得られたシーケンスを含むクラスタおよび／またはソフトウェアの一クラスから得られた閾値量（例えば、９５％）を超えるシーケンスを含むクラスタを特定することができる。

【００３８】

通常、関連ソフトウェアから得られた挙動シーケンスは、まとめてクラスタリングされる傾向にある。同じグッドウェアの異なるバージョン（例えば、異なるパッチレベル）は、実質的に同じ挙動を実行する場合が多く、したがって、異なるバージョンから得られたシーケンスは、まとめてクラスタリングされる傾向にあることになる。同様に、ポリモーフィックマルウェアファミリの異なるインスタンスから得られたシーケンスは、ポリモーフィズムにもかかわらずマルウェアの挙動が一貫した状態で維持されるため、まとめてクラスタリングされる傾向にある。したがって、大部分がマルウェアから得られたシーケンスを含むクラスタ（「マルウェアクラスタ」と呼ばれる）は、単一のマルウェアファミリ、例えば、ポリモーフィズム、共通のコードベースまたは別の関係性を通じて関連付けられる一連のマルウェアの挙動を説明するものと推定される。

【００３９】

クラスタ分析モジュール３２２は、マルウェアクラスタを分析し、クラスタに代理されるファミリ内のマルウェアの検出に有用な挙動サブシーケンスを特定する。クラスタに代理されるマルウェアファミリに対するそのようなサブシーケンスを見出すため、モジュール３２２は、そのクラスタ内の挙動シーケンスのすべてに共通の１つまたは複数のサブシーケンスを特定する。言い換えれば、モジュール３２２は、クラスタ内の挙動シーケンスのすべてにおいて見出される挙動シーケンスの１つまたは複数の部分を特定する。あるいは、クラスタがマルウェアとグッドウェアの両方を含む実施形態では、クラスタ分析モジュール３２２は、クラスタ内のマルウェア挙動シーケンスのみに共通の１つまたは複数の挙動サブシーケンスを特定する。クラスタ内の挙動シーケンスに共通の挙動サブシーケンスは、行動シグネチャを作成できる候補対象を代表するため、本明細書では「候補シーケンス」と呼ばれる。

【００４０】

一実施形態では、クラスタ析モジュール３２２は、サブシーケンスが閾値長さより長い場合にのみ、サブシーケンスを候補対象として特定する。例えば、モジュール３２２は、１０を超える挙動を含む候補シーケンスを特定することができ、ここでの「挙動」はＡＰ

Ｉ呼び出しまたはオペレーションである。この方法では、クラスタ析モジュール３２２は、候補対象がマルウェアファミリに限ったことではないため、短過ぎて誤判定のマルウェア検出をもたらす可能性がある候補シーケンスを除外する。

【００４１】

複数の候補シーケンスが存在する場合は、クラスタ分析モジュール３２２の実施形態は、候補対象を評価し、クラスタ内の挙動シーケンスの中で最も早期に起こる候補対象を特定する。異なる候補対象は、挙動シーケンスの異なる場所で起こり得る。ある候補対象がクラスタ内の挙動シーケンスの終了間近で起こる傾向にあり得る一方で、別の候補対象は挙動シーケンスの開始間近で起こる傾向にあり得る。この状況では、クラスタ分析モジュール３２２は、より早期に現れた候補シーケンスを、ファミリ内のマルウェアの検出に使用するシーケンスとして選択する。より早期のシーケンスを使用することは、クライアント１１６側でのマルウェアの早期検出を可能にするため、有利である。他の実施形態は、追加のおよび／または異なる基準を使用して、複数の候補シーケンスの中からシーケンスを選択する。

【００４２】

シグネチャ作成モジュール３２４は、クラスタ分析モジュール３２２によって選択された候補シーケンスに基づいて行動シグネチャを作成する。所定のクラスタに対して選択された候補シーケンスの場合、シグネチャ作成モジュール３２４は、候補シーケンスをその挙動トレース形式に変換し返す。この変換は、場合によりパラメータの正規化を維持する一方で、候補シーケンス内の任意のオペレーションを拡大して、元の挙動シーケンスに戻す工程を伴う。シグネチャ作成モジュール３２４は、変換された候補シーケンスから行動シグネチャを生成する。行動シグネチャは、対応するクラスタに代表されるマルウェアファミリによって実行された挙動のシーケンスについて説明する。したがって、クライアント１１２のセキュリティモジュール１１６でシグネチャを使用して、クライアント側でマルウェアファミリのインスタンスを検出することができる。この様式では、シグネチャ作成モジュール３２４の実施形態は、マルウェアクラスタのすべてまたは選択されたサブセットに対応するシグネチャを作成する。

【００４３】

シグネチャ更新モジュール３２６は、マルウェアデータセット３１２に追加された新しいマルウェアサンプルの観点から行動シグネチャを更新する。一実施形態では、シグネチャ更新モジュール３２６は、シグネチャ生成モジュール１１８の他のモジュールを使用して、新しいマルウェアをカバーするシグネチャを効果的に生成する。新しいマルウェアサンプルがマルウェアデータセット３１２に追加されると、シグネチャ更新モジュール３２６は、挙動収集モジュール３１６およびトレース正規化モジュール３１８を使用して、新しいマルウェアサンプルに対する挙動シーケンスを生成する。また、シグネチャ更新モジュール３２６は、クラスタリングモジュール３２０を使用して、新しい挙動シーケンスが既存のクラスタの１つと整合する（クラスタリングする）かどうか判断する。

【００４４】

新しい挙動シーケンスが既存のクラスタと整合する場合、シグネチャ更新モジュール３２６は、クラスタ分析モジュール３２２モジュールを使用して、新しく追加された挙動シーケンスの観点からクラスタを分析し、必要ならば、新しい候補シーケンスを生成する。一実施形態では、この分析は、現行の選択された候補シーケンス（すなわち、クラスタに対する行動シグネチャが生成された候補シーケンス）が、新しいマルウェアサンプルに対する挙動シーケンスについても説明するかどうか判断する工程を伴う。説明する場合は、クラスタに対する既存の行動シグネチャを使用して新しいマルウェアを検出することができ、クラスタに対するシグネチャを更新する必要はない。

【００４５】

現行の選択された候補シーケンスが新しい挙動シーケンスについて説明しない場合は、シグネチャ更新モジュール３２６は、クラスタ分析モジュール３２２を使用して、クラスタに対する新しい候補シーケンスを生成する。新しい候補シーケンスは、新しいマルウェ

アサンプルの挙動シーケンスおよび既にクラスタ内に存在していた挙動シーケンスに共通である。次いで、シグネチャ更新モジュール３２６は、クラスタ分析モジュール３２２を使用して、新しい候補シーケンスの中から選択し、シグネチャ作成モジュール３２４を使用して、選択された候補シーケンスに基づいてクラスタに対する新しい行動シグネチャを生成する。シグネチャ生成モジュール１１８は、クライアント１１２のセキュリティモジュール１１６にこの新しい行動シグネチャを分配する。

【００４６】

新しい挙動シーケンスが既存のクラスタと整合しない場合、シグネチャ更新モジュール３２６の実施形態は、クラスタリングモジュール３２０を使用して挙動シーケンスに対する新しいクラスタを作成する。シグネチャ更新モジュール３２６は、クラスタ分析３２２およびシグネチャ作成３２４モジュールを使用して、新しいクラスタに対する候補シーケンスを特定し、候補対象の中から選択されたシーケンスに対する新しい行動シグネチャを作成する。シグネチャ生成モジュール１１８は、クライアント１１２のセキュリティモジュール１１６に新しいクラスタに対する行動シグネチャを分配する。

【００４７】

一実施形態では、新しいマルウェアサンプルの挙動シーケンスを既存のクラスタとクラスタリングすることを試みるよりむしろ、シグネチャ更新モジュール３２６は、データセット３１２、３１４内のソフトウェアのすべての挙動シーケンス（新しいマルウェアに対する挙動シーケンスを含む）を再クラスタリングする。再クラスタリングは、カレンダーベースのスケジュールで、指定量の新しいマルウェアが特定されるとき（例えば、最後のクラスタリング以降、５０個の新しいマルウェアサンプルが特定されるとき）および／またはそれ以外のとき、新しいマルウェアが特定されるごとに実行することができる。それに加えて、シグネチャ更新技法を使用して、新しいグッドウェアの観点からおよび／またはマルウェア３１２もしくはグッドウェア３１４のデータセットからソフトウェアが取り除かれる際、更新されたシグネチャを生成することもできる。

【００４８】

図４は、行動シグネチャを生成および分配するために、セキュリティサーバ１１０の一実施形態によって実行される工程を示すフローチャートである。他の実施形態は、異なるおよび／または追加の工程を実行することができる。さらに、他の実施形態は、異なる順番で工程を実行することができる。その上、セキュリティサーバ１１０以外の実体によって工程の一部またはすべてを実行することができる。

【００４９】

初めに、ソフトウェアデータセットを確立する（４１０）。ソフトウェアデータセットは、マルウェアデータセット３１２およびグッドウェアデータセット３１４を含む。セキュリティサーバ１１０は、例えば、エミュレーション環境でソフトウェアを実行することによって、データセットにソフトウェアに対する挙動トレースを収集する（４１２）。セキュリティサーバ１１０は、トレースを標準的な表現にすることによって、ソフトウェアの挙動トレースを正規化する（４１４）。また、正規化の一環として、セキュリティサーバ１１０は、トレース内の関連連続挙動をオペレーションに分類して挙動シーケンスを形成する（４１４）。

【００５０】

セキュリティサーバ１１０は、例えば、類似性の尺度として編集距離を使用して、同様の挙動シーケンスをまとめてクラスタリングする（４１６）。セキュリティモジュール１１０は、マルウェアの特定に使用できる挙動の候補シーケンスを特定するため、大部分がマルウェアから得られたシーケンスを含むクラスタを分析する（４１８）。セキュリティモジュール１１０は、マルウェアクラスタに対する候補シーケンスの中から選択し、選択された候補シーケンスを使用してクラスタに代表されるマルウェアファミリに対する行動シグネチャを生成する（４１８）。セキュリティサーバ１１０は、クライアント１１２のセキュリティモジュール１１６にマルウェアクラスタに対して生成されたシグネチャを分配する（４２０）。

【００５１】
図５は、新しいマルウェアサンプルを検出するための更新された行動シグネチャを生成および分配するために、セキュリティサーバ１１０の一実施形態によって実行される工程を示すフローチャートである。図５に示される工程と同様に、他の実施形態は、異なるおよび／または追加の工程を実行することができ、工程は、異なる順番でまたは異なる実体によって実行することができる。
【００５２】
初めに、新しいマルウェアサンプルを特定し、マルウェアデータセット３１２に追加する（５１０）。セキュリティサーバ１１０は、新しいマルウェアに対する挙動トレースを収集し、挙動トレースを正規化して挙動シーケンスを生成する（５１２）。セキュリティサーバ１１０は、可能であれば、この挙動シーケンスを既存のクラスタと整合させる（５１４）。挙動トレースはクラスタと整合することを想定すると、セキュリティサーバ１１０は、クラスタを分析し、新しいマルウェアサンプルおよび必要であればクラスタ内に既に存在していた他のマルウェアを包含するクラスタに対するシグネチャを再生成する（５１６）。セキュリティサーバ１１０は、クライアント１１２のセキュリティモジュール１１６に生成されたシグネチャを分配する（５１８）。新しいシグネチャは、クラスタに対する以前のシグネチャの代用として分配することができる。
【００５３】
したがって、本明細書に記載される技法により、マルウェアのファミリを検出できる行動シグネチャの小さく有効なセットの自動生成が可能になる。さらに、本技法は新しいマルウェアおよびマルウェア亜種が発見されると、新しいシグネチャを効果的に生成する。
【００５４】
上記の説明は、ある特定の実施形態の動作を示すために含まれるものであり、本発明の範囲を限定するためのものではない。本発明の範囲は、以下の特許請求の範囲によってのみ限定されるものとする。上記の論考から、当業者には、本発明の精神および範囲によってさらに包含されるであろう多くの変形形態が明らかになるであろう。

【図１】

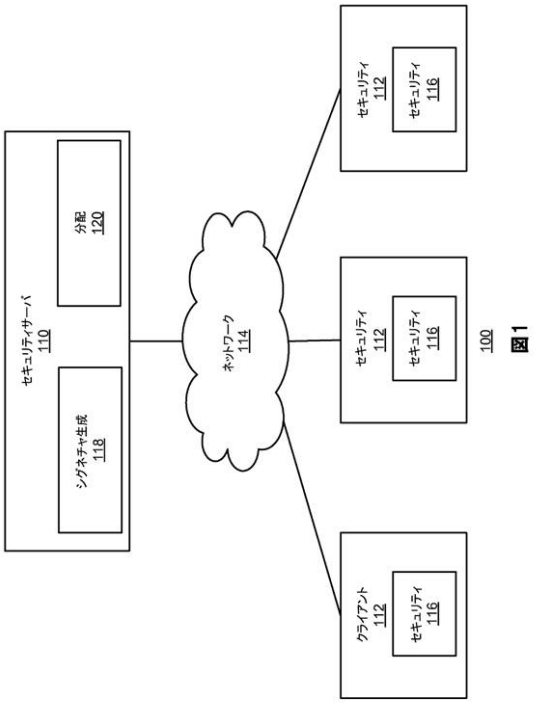
セキュリティサーバ
110
分配
120
シグネチャ生成
118
ネットワーク
114
セキュリティ
112
セキュリティ
116
セキュリティ
112
セキュリティ
116
クライアント
112
セキュリティ
116
100

図１

【図２】

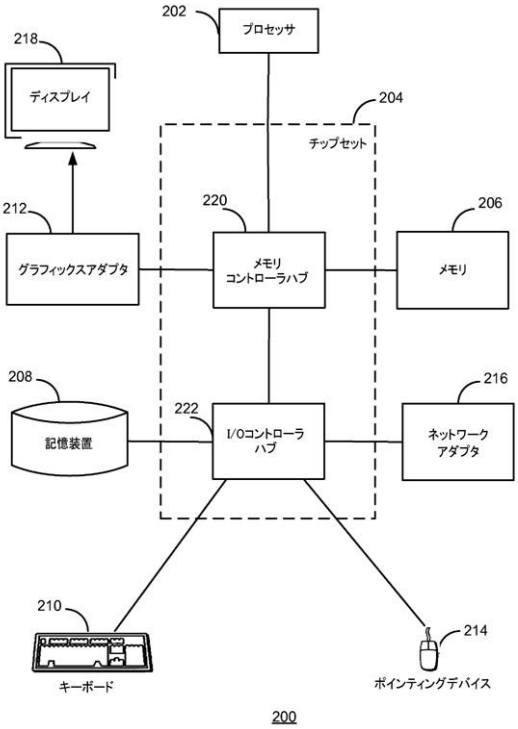
202
プロセッサ
218
ディスプレイ
204
チップセット
212
220
206
グラフィックスアダプタ
メモリ
コントローラハブ
メモリ
208
216
222
記憶装置
I/Oコントローラ
ハブ
ネットワーク
アダプタ
210
214
キーボード
ポインティングデバイス
200

図２

【図３】

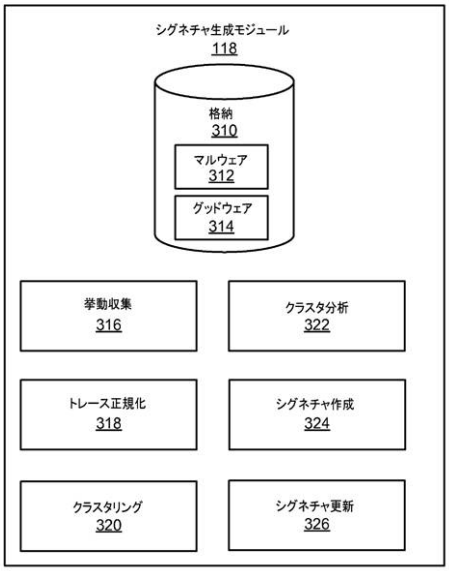
シグネチャ生成モジュール
118
格納
310
マルウェア
312
グッドウェア
314
挙動収集
316
クラスタ分析
322
トレース正規化
318
シグネチャ作成
324
クラスタリング
320
シグネチャ更新
326

図３

【図４】

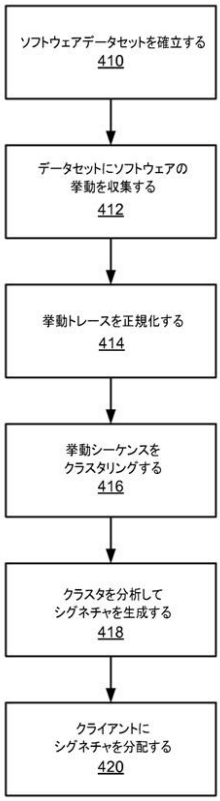
ソフトウェアデータセットを確立する
410
データセットにソフトウェアの
挙動を収集する
412
挙動トレースを正規化する
414
挙動シーケンスを
クラスタリングする
416
クラスタを分析して
シグネチャを生成する
418
クライアントに
シグネチャを分配する
420

図４

【図５】

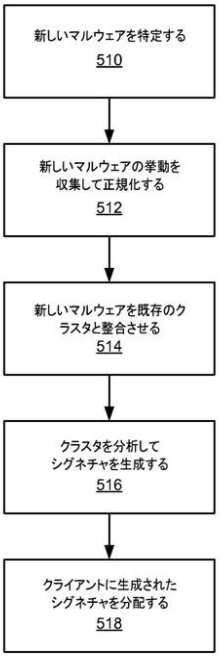
新しいマルウェアを特定する
510
新しいマルウェアの挙動を
収集して正規化する
512
新しいマルウェアを既存のク
ラスタと整合させる
514
クラスタを分析して
シグネチャを生成する
516
クライアントに生成された
シグネチャを分配する
518

図5

フロントページの続き

(72)発明者　サティッシュ・ソーラブ
アメリカ合衆国　カリフォルニア州　９４０４３　マウンテンビュー　エリスストリート　３５０
シマンテックコーポレーション内
(72)発明者　ペレイラ・シェーン
アメリカ合衆国　カリフォルニア州　９４０４３　マウンテンビュー　エリスストリート　３５０
シマンテックコーポレーション内

審査官　木村　励

(56)参考文献　米国特許出願公開第２００７／０１３６４５５（ＵＳ，Ａ１）
欧州特許出願公開第０２１２８７９８（ＥＰ，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　２１／００　－　２１／８８