(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7627473号
(P7627473)

(45)発行日　令和7年2月6日(2025.2.6)

(24)登録日　令和7年1月29日(2025.1.29)

(51)国際特許分類
Ｇ０６Ｆ　21/56　(2013.01)

ＦＩ
Ｇ０６Ｆ　21/56　３２０

請求項の数　14　（全19頁）

(21)出願番号　特願2023-166560(P2023-166560)
(22)出願日　令和5年9月27日(2023.9.27)
(65)公開番号　特開2024-80605(P2024-80605A)
(43)公開日　令和6年6月13日(2024.6.13)
審査請求日　令和5年9月27日(2023.9.27)
(31)優先権主張番号　10202260284Y
(32)優先日　令和4年12月1日(2022.12.1)
(33)優先権主張国・地域又は機関
シンガポール(SG)

(73)特許権者　521354226
フレックソン　ピーティーイー　エルティーディー
ＦＬＥＸＸＯＮ　ＰＴＥ．　ＬＴＤ．
シンガポール シンガポール プラチナム２８ ゲンティングレーン ２８番 ＃０９－０３
(74)代理人　100226861
弁理士　池上 廣
(72)発明者　チャン・メイ・リン
シンガポール プラチナム２８ ゲンティングレーン　２８番 ＃０９－０３
(72)発明者　パラマシヴァム・ラジャモハン
シンガポール プラチナム２８ ゲンティングレーン　２８番 ＃０９－０３

最終頁に続く

(54)【発明の名称】　データファイルを保護するシステムおよび方法

(57)【特許請求の範囲】

【請求項１】

一連のデータファイルを保護するためのシステムであって、前記システムは以下を含む：

一連のデータファイルから変換された簡約化行列（ＲＲＥＦ）のセットを含む変換モジュール、ここで前記一連のファイルはＲＲＥＦのセットにマッピングされ、前記変換モジュールは以下より構成される：

システムより変換されたデータファイルを取得し、前記変換されたデータファイルを変換されたＲＲＥＦのセットに変換し、前記変換されたデータファイルを、変換されたＲＲＥＦのセットにマッピングし、前記変換されたＲＲＥＦのセットを訓練された人工ニューラルネットワーク（ＡＮＮ）モジュールに供給する；

ここで前記訓練されたＡＮＮモジュールは、以下より構成される：

変換されたＲＲＥＦのセットが悪質な動作を含むかどうかを識別し、変換されたＲＲＥＦのセットが悪質な動作を含むと識別された場合、変換されたＲＲＥＦのセットおよび変換されたデータファイルをシステムのメモリ内の仮想の安全な領域へ移す；

以下より構成されるクラスタリングモジュール：バックプロパゲーションモジュールを用いて、前記ＲＲＥＦのセットより、ＲＲＥＦの変換されたセットのバリエーションを取得し；前記取得されたＲＲＥＦの変換されたセットのバリエーション、および変換されたＲＲＥＦのセットをクラスタリングして、悪質な動作を含むＲＲＥＦのセットを識別する；

悪質な動作を含む識別されたＲＲＥＦのセットを訓練されたＡＮＮモジュールに供給する、ここで訓練されたＡＮＮモジュールは、識別されたＲＲＥＦのセットに関連する悪質な

動作のタイプを識別するように構成される；および
訓練されたＡＮＮモジュールによって識別された、データファイルに関連する悪質な動作のタイプに従って、悪質な動作を含む識別されたＲＲＥＦのセットにマッピングされたデータファイルを保護する。
【請求項２】
請求項１に記載のシステムにおいて、前記変換モジュールによる、前記変換されたデータファイルの前記変換されたＲＲＥＦのセットへの変換は、前記変換モジュールが以下のように構成されることを特徴とする：
前記変換されたデータファイルを中間データフレームに変換する、ここで前記中間データフレームは、マルチメディアデータフレームまたは文字ベースのデータフレームを含む；
前記中間データフレームを、１次関数を用いて行列のセットに変換する；及び行列のセットを変換されたＲＲＥＦのセットに変換する。
【請求項３】
前記クラスタリングモジュールが、悪質な動作を含まないＲＲＥＦのセットを変換モジュールに供給するように構成され、変換モジュールが、供給されたＲＲＥＦのセットをデータファイルに変換し、データファイルを供給されたＲＲＥＦのセットにマッピングするようにさらに構成される、請求項１に記載のシステム。
【請求項４】
前記変換されたＲＲＥＦのセットのバリエーションが、異なるタイムスタンプを有するＲＲＥＦのセットを含む、請求項１に記載のシステム。
【請求項５】
請求項１に記載のシステムにおいて、一連のデータファイルのＲＲＥＦのセットへのマッピング、および変換されたデータファイルの変換されたＲＲＥＦのセットへのマッピングは、各データファイルまたは各変換されたデータファイルに対して、データファイルのアドレスおよびデータファイルのメタデータ補足レイヤーにハッシュ関数を適用し、このハッシュ関数の結果を用いて各データファイルをＲＲＥＦのセットに結び付けることによって行われることを特徴とするシステム。
【請求項６】
訓練されたＡＮＮモジュールが、悪質な動作を含むＲＲＥＦのセットに基づいて、ＡＮＮモジュール内のニューラルネットワークをさらに訓練するよう構成される、請求項１に記載のシステム。
【請求項７】
請求項１に記載のシステムにおいて、変換モジュールによる一連のデータファイルのＲＲＥＦのセットへの変換において、変換モジュールが以下のように構成されることを特徴とするシステム：
前記データファイルの各々を中間データフレームに変換する、ここで前記中間データフレームは、マルチメディアデータフレームまたは文字ベースのデータフレームを含む；
前記中間データフレームを、１次関数を用いて行列のセットに変換する；及び行列のセットを変換されたＲＲＥＦのセットに変換する。
【請求項８】
一連のデータファイルから変換された簡約化行列（ＲＲＥＦ）のセットを含む変換モジュールを含むシステムを用いて一連のデータファイルを保護する方法であって、一連のデータファイルがＲＲＥＦのセットにマッピングされ、当該方法は以下のステップを含む：
変換モジュールを用いて、システムから変換されたデータファイルを取得し、変換されたデータファイルを変換されたＲＲＥＦのセットに変換するステップ；
変換モジュールを用いて、変換されたデータファイルを、変換されたＲＲＥＦのセットにマッピングし、前記変換されたＲＲＥＦのセットを訓練された人工ニューラルネットワーク（ＡＮＮ）モジュールに供給するステップ；
訓練されたＡＮＮモジュールを用いて、変換されたＲＲＥＦのセットが悪質な動作を含むかどうかを識別するステップであって、変換されたＲＲＥＦのセットが悪質な動作を含む

と識別された場合、変換されたＲＲＥＦのセットおよび変換されたデータファイルを、システムのメモリ内の仮想安全な領域に移動させるステップ；
バックプロパゲーションモジュールを用いて、ＲＲＥＦのセットから、ＲＲＥＦの変換されたセットのバリエーションを取得するステップ；
クラスタリングモジュールを用いて、取得されたＲＲＥＦの変換されたセットのバリエーションとＲＲＥＦの変換されたセットとをクラスタリングし、悪質な動作を含むＲＲＥＦのセットを識別するステップ；
クラスタリングモジュールを用いて、悪質な動作を含むと識別されたＲＲＥＦのセットを、訓練されたＡＮＮモジュールに供給するステップ、ここで訓練されたＡＮＮモジュールは、識別されたＲＲＥＦのセットに関連する悪質な動作のタイプを識別するように構成される、
クラスタリングモジュールを用いて、訓練済みのＡＮＮモジュールによって識別されたデータファイルに関連する悪質な動作のタイプに従って、悪質な動作を含む識別されたＲＲＥＦのセットにマッピングされたデータファイルを保護するステップ。
【請求項９】
請求項８に記載の方法であって、前記変換されたデータファイルを前記変換されたＲＲＥＦのセットに変換することは、以下を含む方法：
前記変換モジュールを用いて、前記変換されたデータファイルを中間データフレームに変換すること、なおここで前記中間データフレームは、マルチメディアデータフレームまたは文字ベースのデータフレームを含む；
前記変換モジュールを用いて、中間データフレームを１次関数により行列のセットに変換し、変換モジュールを用いて、行列のセットを変換されたＲＲＥＦのセットに変換すること。
【請求項１０】
クラスタリングモジュールを用いて、悪質な動作を含まないＲＲＥＦのセットを変換モジュールに供給する；ここで変換モジュールは、供給されたＲＲＥＦのセットをデータファイルに変換し、データファイルをＲＲＥＦのセットにマッピングするよう構成される、請求項８に記載の方法。
【請求項１１】
前記変換されたＲＲＥＦのセットのバリエーションが、異なるタイムスタンプを有するＲＲＥＦのセットを含む、請求項８に記載の方法。
【請求項１２】
請求項８に記載の方法において、一連のデータファイルのＲＲＥＦのセットへのマッピング、および変換されたデータファイルの変換されたＲＲＥＦのセットへのマッピングは、各データファイルまたは変換された各データファイルに対して、データファイルのアドレスおよびデータファイルのメタデータ補足レイヤーにハッシュ関数を適用し、このハッシュ関数の結果を用いて各データファイルをＲＲＥＦのセットに結び付けることによって行われることを特徴とする方法。
【請求項１３】
請求項８に記載の方法において、前記方法は、悪質な動作を含むＲＲＥＦのセットに基づいて、ＡＮＮモジュール内のニューラルネットワークを訓練するステップをさらに含むことを特徴とする方法。
【請求項１４】
請求項８に記載の方法において、前記変換モジュールによって一連のデータファイルをＲＲＥＦのセットに変換することは、以下のステップを含む：
前記変換モジュールを用いて、前記変換されたデータファイルを中間データフレームに変換するステップ、ここで前記中間データフレームは、マルチメディアデータフレームまたは文字ベースのデータフレームを含み：
前記変換モジュールを用いて、中間データフレームを、１次関数を用いて行列のセットに変換するステップ；及び変換モジュールを用いて、行列のセットを、変換されたＲＲＥＦ

のセットに変換するステップ。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、選択された一連のデータファイルから選択されたデータファイルを保護するシステムおよび方法に関する。本発明のシステムは、変換モジュール、人工ニューラルネットワーク（ＡＮＮ）、クラスタリングモジュール、およびバックプロパゲーションモジュールを含み、それらのモジュールは、マルウェアまたは異常なデータファイルを検出するように構成される。このようなデータファイルが検出されると、システムは、検出されたマルウェアに感染している可能性のある他のデータファイルを検出するための一連の検査を開始する。これらのデータファイルは、マルウェアがホストマシンおよび／またはホストマシンと接続されるストレージ／周辺機器に影響を与えないように保護される。
【背景技術】
【０００２】
マルウェアは、現代のコンピュータ環境において常に脅威となっている。より多くのデバイスが無線または有線のネットワークを通じて繋がるようになると、マルウェアはネットワークを通じて容易に拡散され、接続されるすべてのデバイスに感染する可能性がある。感染したコンピュータは、個人や組織の生産性を低下させ、さらに取り返しのつかない被害をもたらす可能性がある。例えば、感染したコンピュータは、オペレーティングシステムが損傷したり、データが破損することによって、動作が遅くなったり、完全に機能しなくなったりする可能性がある。最悪の場合、貴重なデータや情報が悪用され、乗っ取られ、他人によって使用されたり、身代金を要求されたりする可能性がある。加えて感染したシステムは、ネットワーク外の他のデバイスにマルウェアを拡散し続ける可能性がある。
【０００３】
従来のマルウェア対策ソフトウェアは、既知のマルウェアのシグネチャを検出し、また、他の既知のマルウェアに関連するプログラム内のパターンを検出して、既知のマルウェアを識別および検出を試みる場合がある。このようなマルウェアへの対策手法は、シグネチャベースの検出システムであって、ファイルやデータを既知のマルウェアのパターンと比較する。しかし、ゼロデイ攻撃は通常、既知のマルウェアとは大きく異なるため、既知のマルウェアのパターンには当てはまらず、ゼロデイ攻撃を検出できない可能性がある。
【０００４】
前記のように、貴重なデータや情報がマルウェアによって悪用される可能性がある。そこで、データのバックアップを頻繁に行い、また、暗号化されたデータを安全な環境に保管するなどの対策が取られることが多い。
【０００５】
しかし、このような対策は多くの問題を抱えている。バックアップされたデータがハッキングされたり、破損したり、ハードウェアが故障／偶発的に損傷したり、ランサムウェアやマルウェアへ感染したり、データが偶発的に削除される等が起こる可能性がある。ランサムウェアの場合、攻撃者は、ユーザーのファイルを暗号化し、データへのアクセスを可能にするために被害者へ身代金を要求する。このような攻撃で最も一般的なものは、フィッシング・スパムである。フィッシング・スパムでは、何も知らないユーザーによって不正な添付ファイルがダウンロードされ、実行される。これらのファイルが実行されると、関連するスクリプトをセルフインストールし、そこに含まれるデータを破壊することで、デバイスのオペレーティングシステムを支配する。
【０００６】
悪質なスクリプトの中には、ユーザーにより実行される必要のないものもある。このようなスクリプトは、オペレーティングシステムにおけるセキュリティの脆弱性につけ込み、そこに含まれるファイルやデータを感染させることで、デバイスに損害を与える可能性がある。さらに、リークウェアやドックスウェアと呼ばれるサイバー攻撃もある。このような攻撃では、デバイスにアクセスした悪質な第三者が、デバイスの記憶媒体に含まれる機

密データを公開すると脅迫する。
【０００７】
また、悪質な第三者は、トロイの木馬の攻撃を利用して、ユーザーのデータに損害を加えたり、データを盗んだりする。トロイの木馬は、ウイルス、ワーム、スパイウェア、ゾンビ、ボットネット、ロジック爆弾、トラップドアをも生じさせる。一般的に、これらの悪質なサイバー攻撃はすべて、ユーザーデータのセキュリティを侵すことが唯一の目的である。これらの攻撃は通常、データが格納されているアドレス・ポインタとともに、オペレーティングシステム及びオペレーティングシステムのカーネルを破壊し、結果的にデータそのものを破壊する。
【０００８】
上記の理由から、当業者は、マルウェアを含む可能性のあるデータファイルを識別および検出することによって、デバイス内のファイルを保護することができるシステムおよび方法を作り出すため鋭意努力している。
【発明の概要】
【０００９】
本発明のシステムおよび方法によって上記及びその他の課題は解決され、技術の進展がもたらされる。
【００１０】
本発明のシステムおよび方法における第一の利点は、ファイルが実行可能か否かに関わらず、すべての形式のファイルを保護するようにシステムが構成されていることにある。
【００１１】
本発明のシステムおよび方法における第二の利点は、たとえもしホストデバイスのオペレーティングシステムが侵された場合であっても、権限のないユーザーによる操作および再構築が困難なフォーマットでデータファイルを保存するように構成されていることにある。
【００１２】
本発明のシステムおよび方法の第三の利点は、本発明の実施態様に従ってファイルが一旦保存されると、システムは、ＲＲＥＦを保存されたファイルにマッピングするために用いられる関数によって、破損したファイルを簡単かつ効率的に回復できることにある。
【００１３】
本発明のシステムおよび方法の第四の利点は、ファイルが訓練されたニューラルネットワークを用いて分析される前に数値表現に変換されるため、システムがデータファイル内に含まれるマルウェアおよびゼロデイ型のマルウェアを検出するように構成されていることにある。
【００１４】
本発明のシステムおよび方法の第五の利点は、マルウェアを含むと識別されたファイルをメモリ内の仮想の安全な領域に格納するように構成され、当該格納されたファイルは許可されたユーザーによってのみ取得できることにある。
【００１５】
上述の利点は、以下に示す本発明のシステムおよび／または方法によって実現可能である。
【００１６】
本発明の第一の実施態様によれば、一連のデータファイルを保護するためのシステムであって、前記システムは以下を含む：一連のデータファイルから変換された簡約化行列（ＲＲＥＦ）のセットを含む変換モジュール、ここで前記一連のファイルはＲＲＥＦのセットにマッピングされ、前記変換モジュールは以下より構成される：
システムより変換されたデータファイルを取得し、前記変更されたデータファイルを変更されたＲＲＥＦのセットに変換し、前記変換されたデータファイルを、変換されたＲＲＥＦのセットにマッピングし、前記変換されたＲＲＥＦのセットを訓練された人工ニューラルネットワーク（ＡＮＮ）モジュールに供給する；
ここで前記訓練されたＡＮＮモジュールは、以下より構成される：
変換されたＲＲＥＦのセットが悪質な動作を含むかどうかを識別し、変換されたＲＲＥＦ

のセットが悪質な動作を含むと識別された場合、変換されたＲＲＥＦのセットおよび変換されたデータファイルをシステムのメモリ内の仮想の安全な領域へ移す；
以下より構成されるクラスタリングモジュール：バックプロパゲーションモジュールを用いて、前記ＲＲＥＦのセットより、変換された種々のＲＲＥＦのセットを取得し；前記取得された変換されたＲＲＥＦのセットおよび変換されたＲＲＥＦのセットをクラスタリングして、悪質な動作を含むＲＲＥＦのセットを識別する；悪質な動作を含む識別されたＲＲＥＦのセットを訓練されたＡＮＮモジュールに供給する、ここで訓練されたＡＮＮモジュールは、識別されたＲＲＥＦのセットに関連する悪質な動作のタイプを識別するように構成される；および
訓練されたＡＮＮモジュールによって識別された、データファイルに関連する悪質な動作のタイプに従って、悪質な動作を含む識別されたＲＲＥＦのセットにマッピングされた安全なデータファイル。

【００１７】
本発明の第一の実施態様によれば、変換モジュールによって変換されたデータファイルのＲＲＥＦへの変換は、以下の変換モジュールを含む：データファイルを中間データフレームに変換し、ここで中間データフレームは、マルチメディアデータフレームまたは文字ベースのデータフレームを含む；
中間データフレームを、１次関数を用いて行列のセットに変換する；及び行列のセットを変換されたＲＲＥＦのセットに変換する。

【００１８】
本発明の第一の実施態様によれば、クラスタリングモジュールは、悪質な動作を含まないＲＲＥＦのセットを変換モジュールに供給するように構成され、ここで変換モジュールは、供給されたＲＲＥＦのセットをデータファイルに変換し、データファイルをＲＲＥＦのセットにマッピングするように構成される。

【００１９】
本発明の第一の実施態様によれば、変換された種々のＲＲＥＦのセットは、異なるタイムスタンプを有するＲＲＥＦのセットを含む。

【００２０】
本発明の第一の実施態様によれば、データファイルのＲＲＥＦのセットへのマッピング、および変換されたデータファイルのＲＲＥＦのセットへのマッピングは、各データファイルまたは各変換されたデータファイルに対して、データファイルのアドレスおよびデータファイルのメタデータ補足レイヤーにハッシュ関数を適用し、このハッシュ関数の結果を用いて各データファイルをＲＲＥＦのセットに結び付けることによって行われる。

【００２１】
本発明の第一の実施態様によれば、訓練されたＡＮＮモジュールは、悪質な動作を含むＲＲＥＦの特定のセットに基づいて、ＡＮＮモジュール内のニューラルネットワークをさらに訓練するように構成される。

【００２２】
本発明の第一の実施態様によれば、変換モジュールによってデータファイルはＲＲＥＦのセットへ変換される。前記変換モジュールは、以下の構成を含む：前記データファイルの各々を中間データフレームに変換する、ここで前記中間データフレームは、マルチメディアデータフレームまたは文字ベースのデータフレームを含む；前記中間データフレームを、１次関数を用いて行列のセットに変換する；及び行列のセットを変換されたＲＲＥＦのセットに変換する。

【００２３】
一連のデータファイルから変換された簡約化行列（ＲＲＥＦ）のセットを含む変換モジュールを含むシステムを用いて一連のデータファイルを保護する方法であって、一連のデータファイルがＲＲＥＦのセットにマッピングされ、当該方法は以下のステップを含む：変換モジュールを用いて、システムから変換されたデータファイルを取得し、変換されたデータファイルを変換されたＲＲＥＦのセットに変換するステップ；変換モジュールを用い

て、変換されたデータファイルを、変換されたＲＲＥＦのセットにマッピングし、前記変換されたＲＲＥＦのセットを訓練された人工ニューラルネットワーク（ＡＮＮ）モジュールに供給するステップ；訓練されたＡＮＮモジュールを用いて、変換されたＲＲＥＦのセットが悪質な動作を含むかどうかを識別するステップであって、変換されたＲＲＥＦのセットが悪質な動作を含むと識別された場合、変換されたＲＲＥＦのセットおよび変換されたデータファイルを、システムのメモリ内の仮想の安全な領域に移動させるステップ；バックプロパゲーションモジュールを用いて、ＲＲＥＦのセットから、変換された種々のＲＲＥＦのセットを取得するステップ；クラスタリングモジュールを用いて、取得されたＲＲＥＦの変換されたセットのバリエーションとＲＲＥＦの変換されたセットとをクラスタリングし、悪質な動作を含むＲＲＥＦのセットを識別するステップ；クラスタリングモジュールを用いて、悪質な動作を含むと識別されたＲＲＥＦのセットを、訓練されたＡＮＮモジュールに供給するステップ、ここで訓練されたＡＮＮモジュールは、識別されたＲＲＥＦのセットに関連する悪質な動作のタイプを識別するように構成される、クラスタリングモジュールを用いて、訓練済みのＡＮＮモジュールによって識別されたデータファイルに関連する悪質な動作のタイプに従って、悪質な動作を含む識別されたＲＲＥＦのセットにマッピングされたデータファイルを保護するステップ。

【００２４】

本発明の第二の実施態様によれば、変換されたデータファイルを変換されたＲＲＥＦのセットに変換することは以下を含む：変換モジュールを用いて、データファイルを中間データフレームに変換すること、なおここで前記中間データフレームはマルチメディアデータフレームまたは文字ベースのデータフレームを含む：変換モジュールを用いて、中間データフレームを１次関数により、行列のセットに変換し、変換モジュールを用いて、行列のセットを変換されたＲＲＥＦのセットに変換すること。

【００２５】

本発明の第二の実施態様によれば、本発明の方法は以下のステップをさらに含む：クラスタリングモジュールを用いて、悪質な動作を含まないＲＲＥＦのセットを変換モジュールに供給し、ここで変換モジュールは、供給されたＲＲＥＦのセットをデータファイルに変換し、データファイルをＲＲＥＦのセットにマッピングする。

【００２６】

本発明の第二の実施態様によれば、変換された種々のＲＲＥＦのセットは、異なるタイムスタンプを有するＲＲＥＦのセットを含む。

【００２７】

本発明の第二の実施態様によれば、一連のデータファイルのＲＲＥＦのセットへのマッピング、および変換されたデータファイルの変換されたＲＲＥＦのセットへのマッピングは、各データファイル又は変換された各データファイルに対して、データファイルのアドレスおよびデータファイルのコンテキストメタデータレイヤーにハッシュ関数を適用し、ハッシュ関数の結果を用いて各データファイルをＲＲＥＦのセットに結び付けることで行われる。

【００２８】

本発明の第二の実施態様によれば、本方法は、悪質な動作を含むＲＲＥＦのセットに基づいて、ＡＮＮモジュール内のニューラルネットワークをさらに訓練するステップ含む。

【００２９】

本発明の第二の実施態様によれば、変換モジュールを用いた一連のデータファイルのＲＲＥＦのセットへの変換は以下のステップを含む：前記変換モジュールを用いて、前記変換されたデータファイルを中間データフレームに変換するステップ、ここで前記中間データフレームは、マルチメディアデータフレームまたは文字ベースのデータフレームを含み：前記変換モジュールを用いて、中間データフレームを、１次関数を用いて行列のセットに変換するステップ；及び変換モジュールを用いて、行列のセットを、変換されたＲＲＥＦのセットに変換するステップ

【図面の簡単な説明】

【００３０】
上記及びその他の解決すべき課題は、発明の詳細な説明及び以下の図面に記載される本発明のシステムおよび方法によって解決される。
【００３１】
【図１】は、本発明の実施態様に従って、ホストマシン内のデータファイルを保護するためのシステムを示したブロック図である。
【００３２】
【図２】は、本発明の実施態様を実行するためのモジュールまたはデバイスの構成要素を示したブロック図である。
【００３３】
【図３】は、本発明の実施態様に従って、一連のデータファイルを簡約化行列（ＲＲＥＦ）のセットに変換するフロー図を示す。
【００３４】
【図４】は、本発明の実施態様に従って、データファイルを保護するフロー図を示す。
【００３５】
【図５】は、本発明の実施態様に従って、データファイルから中間データフレームへの変換を示す図である。
【００３６】
【図６】は、本発明の実施態様に従って、中間ファイルから簡約化行列への変換を示す図である。
【発明を実施するための形態】
【００３７】
本発明は、一連のデータファイルから選択されたデータファイルを保護するシステムおよび方法に関する。前記システムは、変換モジュール、人工ニューラルネットワーク（ＡＮＮ）、クラスタリングモジュール、およびバックプロパゲーションモジュールを含み、これらのモジュールは、マルウェアまたは異常を含むデータファイルを識別するように構成される。このようなデータファイルが検出されると、システムは、検出されたマルウェアと同様に感染している可能性のある他のデータファイルを検出するための一連の測定を開始する。これらのデータファイルは、マルウェアがホストマシンおよび／またはホストマシンに接続されているストレージ／周辺機器に影響を与えないように保護される。
【００３８】
特に、一連のデータファイルから変換された簡約化行列（ＲＲＥＦ）を含む変換モジュールであって、一連のデータファイルがＲＲＥＦにマッピングされた変換モジュールは、以下のように構成される：システムから変換されたデータファイルを取得し、変換されたデータファイルをＲＲＥＦに変換する。変換されたデータファイルは、変換されたＲＲＥＦのセットにマッピングされ、訓練された人工ニューラルネットワーク（ＡＮＮ）モジュールに供給される。変換されたＲＲＥＦのセットを受け取ると、訓練されたＡＮＮモジュールは、変換されたＲＲＥＦのセットが悪質な動作を含むかどうかを識別し、変換されたＲＲＥＦのセットが悪質な動作を含むと識別されると、変換されたＲＲＥＦのセットと変換されたデータファイルは、システムのメモリ内の仮想の安全な領域へ移される。クラスタリングモジュールは、バックプロパゲーションモジュールを用いて、ＲＲＥＦのセットから種々の変換されたＲＲＥＦのセットを取得し、取得された種々のＲＲＥＦのセットと安全な領域内の変換されたＲＲＥＦのセットをクラスタリングして、悪質な動作を含むＲＲＥＦのセットを識別する。悪質な動作を含むＲＲＥＦのセットは、訓練されたＡＮＮモジュールに供給され、悪質な動作を含むＲＲＥＦのセットにマッピングされたデータファイルが保護される。
【００３９】
以下、本発明を、添付の図面に示された実施態様を用いて詳細に説明する。以下の説明では、本発明の実施態様の十分な理解を与えるために、多数の具体的な態様を示す。しかしながら、当業者であれば、具体的態様の一部または全部が記載されていなくとも当該具体

的な態様を実施することができることは明らかである。そのような実施態様も本発明の範囲内に含まれる。さらに、以下の特定のプロセス、ステップおよび／または構造は、本明細書に詳細に記載されていない場合があるが、この場合、対応する引用文献を参照することで実施可能である。

【００４０】

加えて当業者であれば、本明細書中の多くの機能ユニットが、モジュールとして記載されていることを理解するであろう。また当業者であれば、モジュールは、回路、ロジックチップ、または任意の構成要素であることを認識するであろう。さらに、当業者であれば、モジュールは、様々なプロセッサアーキテクチャによって実行することができるソフトウェアであることも認識するであろう。本発明の実施態様によれば、モジュールは、コンピュータ命令、ファームウェア、または受け取った命令に基づいて一連のイベントを実行するようコンピュータプロセッサに指示する実行可能なコードを含む。モジュールの選択は、設計上の選択として当業者に委ねられており、本発明の範囲を何ら限定するものではない。

【００４１】

本発明の実施態様に従ってデータファイルを保護するための例示的なプロセスまたは方法を、以下のステップで説明する。

【００４２】

ステップ１：変換モジュールを用いて、システムから変換されたデータファイルを取得し、当該データファイルをＲＲＥＦのセットに変換する；変換されたデータファイルを変換されたＲＲＥＦのセットにマッピングし、変換されたＲＲＥＦのセットを学習済み人工ニューラルネットワーク（ＡＮＮ）に供給する、ここで変換モジュールは、一連のデータファイルから変換された簡約化行列（ＲＲＥＦ）のセットを含み、加えて一連のデータファイルがＲＲＥＦのセットにマッピングされる。

【００４３】

ステップ２：訓練されたＡＮＮモジュールを用いて、変換されたＲＲＥＦのセットが悪質な動作を含むかどうかを識別し、これにより、変換されたＲＲＥＦのセットが悪質な動作を含むと識別された場合、変換されたＲＲＥＦのセットと変換されたデータファイルは、システムのメモリ内の仮想の安全な領域に移される。

【００４４】

ステップ３：バックプロパゲーションモジュールを用いて、ＲＲＥＦのセットから、変換された種々のＲＲＥＦのセットを取得する。

【００４５】

ステップ４：クラスタリングモジュールを用いて、取得された種々の変換されたＲＲＥＦのセット及び変換されたＲＲＥＦのセットをクラスタリングし、悪質な動作を含むＲＲＥＦのセットを検出する。

【００４６】

ステップ５：クラスタリングモジュールを用いて、悪質な動作を含むと検出されたＲＲＥＦのセットを訓練されたＡＮＮモジュールに供給し、悪質な動作を含むと検出されたＲＲＥＦのセットにマッピングされたデータファイルを保護する。

【００４７】

本発明の実施態様によれば、上記のステップは、ホストマシンに通信可能に接続されたシステムまたはハードウェアモジュール（サブモジュールを含む）によって実行されてもよい。また、上記のステップは、ホストマシンおよび／または接続された周辺装置のいずれかに設けられたソフトウェアモジュールによって実行されてもよい。

【００４８】

図１に、本発明の実施態様におけるホストマシン内のデータファイルを保護するシステムのブロックを示す。

【００４９】

図１は、一般的にＣＰＵ１０２、キャッシュ１０４、メモリ１０６および／またはストレ

ージ１０８を有する任意のハードウェアデバイスを含むホストマシン１１０を示す。ホストマシンには、コンピュータ、パーソナル電子デバイス、シンクライアント、および多機能デバイスが含まれるが、これらに限定されない。特に、集中管理型のメインフレーム、サーバ、またはデスクトップ型パーソナルコンピュータ（ＰＣ）を含むほぼあらゆる種類のコンピュータがホストマシンに含まれる。

【００５０】

ＣＰＵ１０２は、命令を処理することができる任意のデバイスまたは構成要素を含んでいてもよく、マイクロプロセッサ、マイクロコントローラ（ＭＣＵ）、プログラマブルロジックデバイス、または他の計算装置を含んでいてもよい。ＣＰＵ１０２は、入力信号を受信し、メモリ１０６に格納された命令に従って処理し、出力するための任意の適切な論理回路を有していてもよく、メモリアドレス空間を有するシングルコアまたはマルチコアプロセッサを含んでいていてもよい。さらに、メモリ１０６は、揮発性メモリおよび不揮発性メモリを含んでいてもよく、ストレージ１０８は、ソリッドステートデバイス（ＳＳＤ）、ハードディスクドライブ（ＨＤＤ）、光学ドライブ、または磁気ディスクドライブを含んでいてもよい。

【００５１】

ホストマシン１１０は、ＣＰＵ１０２とホストマシン１１０内の様々な電子部品との間のデータ通信を管理し、ホストマシン１１０とシステム１２０との間で交換されるデータを管理するために用いられる任意のマイクロチップを含んでいてもよいＩ／Ｏハブ（図示せず）を介して、システム１２０に通信可能に接続される。

【００５２】

システム１２０は、データファイルを簡約化行列（ＲＲＥＦ）に変換し、ＲＲＥＦをデータファイルに再構築するように構成された変換モジュール１２２と、訓練されたニューラルネットワークを構成する人工ニューラルネットワーク（ＡＮＮ）モジュール１２４と、安全および安全でない揮発性メモリおよび不揮発性メモリを含むメモリ１３０と、データファイルおよび／または種々のＲＲＥＦのセットを識別するように構成されるバックプロパゲーションモジュール１２６と、悪質な動作を含む可能性のあるＲＲＥＦのクラスタを識別するためにＲＲＥＦのクラスタを生成するように構成されるクラスタリングモジュール１２８を含む。当業者であれば、システム１２０は、本発明から逸脱することなく、追加の他のタイプのモジュールおよび／または構成要素を含むことが理解できる。

【００５３】

本発明の実施態様によれば、初期設定の段階において、システム１２０は、ホストマシン１１０からデータファイル１１１の一部または全体、もしくは選択された部分を取得し、更にデータファイル１１１は、メモリ１０６またはストレージ１０８内に格納することができる任意のファイル、またはホストマシン１１０のモジュール間で内部通信されるか、またはホストマシン１１０のモジュールによって実行可能な任意のファイルを含んでいてもよい。チェックサムは、チェックサムを有するデータファイルから抽出され、この情報は、将来の使用のためにシステム１２０のメモリ１３０に格納される。

【００５４】

一連のデータファイル１１１がホストマシン１１０より取得されるという想定の下、システム１２０は、各々のデータファイル１１１に、各種許可およびユーザー設定を関連付ける。データファイルの各々に関連付けられたメタデータが、関連付けられたファイルの真正性に関連する情報を含むことができるため、これらの設定は不可欠である。さらに、当該メタデータは、内部検証プロセスのための認証手段、例えば、ＰＥＭファイルのような証明書として使用することができる。

【００５５】

次に、変換モジュール１２２は、各々のデータファイル（許可とユーザー設定により関連付けられたデータファイルを含む場合もあれば、関連付けられていないデータファイルを含む場合もある）の変換を行う。各データファイルは、最初に中間データフレームに変換され、その後、行列のセットへ変換される。本発明の実施態様によれば、中間データフレ

ームは、画像ファイル形式を有するフレームを含むマルチメディアデータフレームを含んでいてもよく、文字ベースのデータフレームを含んでいてもよい。マルチメディアデータフレームの赤緑青（ＲＧＢ）の性質を考慮すれば、それらを数値へ変換することは容易である。本発明の他の実施態様によれば、中間データフレームは、代わりに文字ベースのデータフレームで構成されてもよい。本発明のさらなる実施態様によれば、中間データフレームは、マルチメディアデータフレーム及びキャラクタベースデータフレームの両方を含んでいてもよい。

【００５６】

中間データフレームは、それがマルチメディアであるか文字ベースのデータフレームであるかに関係なく、行列のセットに変換される。これは、データファイル５０２が変換プロセス５０６を通して行列のスナップショット５０４に変換される図５に示されている。図６は、行列のスナップショット６０４から数値の行列表現６０２への変換を示す。本発明の実施態様によれば、この変換は、所定のルックアップテーブルを用いて行われてもよく、それによって、中間データフレーム内の各英数字記号および／または文字および／または画素および／または色が、一意の番号および／または記号によって表されてもよい。次に、所定の行列生成アルゴリズムまたは関数が、ルックアップテーブルと共に使用されて、中間データフレームを行列のセットに変換することができる。本発明の他の実施態様によれば、中間データフレームを数値データフレームに変換するために、より複雑な数学関数を用いてもよく、当該数学関数は、シャッフル、回転、順序変換などの関数を含んでいてもよい。

【００５７】

本発明のさらなる実施態様では、偏微分方程式関数を数値のルックアップテーブルと共に用いて、データを導出し、従量引数を用いてデータを暗号化することができる。従量引数は、データがパラメータ化およびセグメント化されるハッシュまたは暗号化キーを含んでいてもよい。

【００５８】

各々の中間データフレームが行列のセットへ変換されると、変換モジュール１２２は、行列のセット（線形方程式で表すことができる）を簡約化行列（ＲＲＥＦ）のセットにさらに変換する。行列のセットをＲＲＥＦのセットに変換するこのステップは、変換プロセスがシステム１２０のみが知っている線形方程式の使用を含むため、データに暗号化のレイヤーを追加する。それにより、これらの線形方程式は、システム１２０のためのユニークな暗号化アルゴリズムとして機能することができる。

【００５９】

一連のデータファイル１１１が変換モジュール１２２によって対応するＲＲＥＦのセットに変換されると、各々のデータファイルはＲＲＥＦのセットにマッピングされる。これはキー／値のペアリング関数またはキー／値のハッシュ関数を用いて行われるが、これに限定されない。ペアリングを生成するために使用された関数と作成されたマッピングは、ＲＲＥＦのセットとともにメモリ１３０に安全に格納される。

【００６０】

本発明の別の実施態様では、データファイルのメタデータおよび／またはデータファイルのアドレスにハッシュ関数を適用することができる。ハッシュ関数の結果は、データファイルを対応するＲＲＥＦのセットにマッピングするために使用される。ハッシュ関数でデータファイルのメタデータ補足レイヤーを使用することにより、データファイルのアドレスが検証されるだけでなく、参照されるメタデータも検証されるため、マッピング関数にユニークな特性が追加される。

【００６１】

本発明の実施態様によれば、ハッシュ関数は、自然言語処理（ＮＬＰ）アルゴリズム及びキー／値のバインド関数の組み合わせを含んでいてもよい。ＮＬＰアルゴリズムは、文脈メタデータを処理し、情報を半構造化し、特定のキーアドレスに格納するように構成される。これにより、（マッピング関数の）検索アルゴリズムが改善され、データフレームの

位置を特定するタスクがデータファイルのチェックサムの検証とともに生成される可能性のある適切なキーをシステムが検索するため、複雑性の値が０（１）に設定される。その結果、このキーはシステム内で一意に定義される。

【００６２】

従来のオペレーティングシステムにおいては、データファイルは単にホストマシンの１６進数のメモリアドレスに格納され、データファイルの場所を特定するためにポインタが使用されるため、このステップは重要である。当該ポインタがマルウェアによって破損された場合、データファイルはホストマシンから取得できなくなるため、事実上使用できなくなる。システム１２０は、一連のデータファイル１１１の変換においてホストマシン１１０を継続的に監視する。

【００６３】

本発明の実施態様では、システム１２０がホストマシン１１０内の改ざんされたデータファイルを検出すると、システム１２０はホストマシン１１０から改ざんされたデータファイルを取得するように動作する。本発明の実施態様によれば、システム１２０は、改ざんされたデータファイルのチェックサムを検証して、改ざんされたデータファイルが不正なユーザーによって改ざんされたかどうかを識別する。改ざんされたデータファイルに、破損したチェックサムまたは不一致のチェックサムが含まれている場合、当該改ざんされたデータファイルは安全な領域に移される。しかし、悪質な第三者は、チェックサムが変換されないようにしながらデータファイルを変換する方法を発見する。したがって、変換されたデータファイルはシステム１２０によってさらに処理される必要がある。

【００６４】

変換モジュール１２２を用いて取得された変換されたデータファイルは、ＲＲＥＦの変換されたセットに変換される。本発明の実施態様においては、最初に変換されたデータファイルを中間データフレームへ変換してから、行列のセットに変換することによって行うことができる。次に、当該行列のセットは、数学的演算を用いて、対応する変換されたＲＲＥＦのセットに変換される。変換モジュール１２２は次に、変換されたデータファイルを変換されたＲＲＥＦのセットにマッピングし、このマッピングをメモリ１３０に安全に格納する。変換されたＲＲＥＦのセットは、次に、訓練された人工ニューラルネットワーク（ＡＮＮ）モジュール１２４に供給される。

【００６５】

本発明の他の実施態様では、データファイルが修正または変換された時、変換されたデータファイルを変換されたＲＲＥＦのセットに変換するために使用された多項式の平方根を用いて、変換されたＲＲＥＦのセットを変換および／または変換された数値を識別することができる。変換されたＲＲＥＦのセットの数字が多項式のルートの乗数を構成する場合、それはデータファイルが破損していないことを意味する。逆に、変換されたＲＲＥＦのセットの数字が多項式のルートの乗数を構成していない場合、それはデータファイルが破損していることを意味する。

【００６６】

本発明の実施態様において、ＡＮＮは、６６から２００の間の隠れ層を有するニューラルネットワークを含んでいてもよい。ＡＮＮは、教師なし学習方法を用いて初期訓練され、訓練のデータセットは、既知のマルウェア、侵害、および／またはサイバー攻撃に関連するＲＲＥＦのセット、および／またはマルウェアに関連しないＲＲＥＦのセットを含む。ＲＲＥＦのセットは、リード／ライト操作、メタデータ、アクセスレベルスレッド、ハードウェア固有の識別ＩＤ、消費されたメモリ、Ｉ／Ｏバースト、物理アドレスまたはソフトウェアチェックサムなどのそれぞれの特性に従って、それぞれのタイムフレームと共にラベル付けされることがあるが、これらに限定されない。ＡＮＮは教師なし学習方法を用いて訓練されるため、訓練されたＡＮＮは、脅威が検出される度に、新しい訓練を行う。本発明のさらなる実施態様において、ＡＮＮは、特定のタイプの悪意ある攻撃によってばらばらになっているＲＲＥＦのセットのパラメータを検出するように訓練されることがある。これらのパラメータが検出され、悪質な攻撃と関連付けられると、訓練されたＡＮＮ

は、将来そのような攻撃を検出できるようになる。
【００６７】
本発明のさらなる実施態様では、ＡＮＮのニューラルアーキテクチャは、既知のウイルスタイプの１００の重要なカテゴリーの訓練サンプルデータを用いて設計される。影響を受ける特徴は、入力ノードの数、隠れ層の数、および各々の隠れ層のノードの数に基づいて特定される。ＡＮＮの訓練率は、開始時に０．０１の値でパラメトリック化され、その後バッチサイズを１０００として徐々に増加させた。エポック数は５００から開始し、バッチサイズ８００、エポック数２５０で最適な結果が得られることが判明した。
【００６８】
次に、訓練済みのＡＮＮモジュール１２４は、訓練済みのＡＮＮを用いて、変換されたＲＲＥＦのセットに悪質な動作が含まれているかどうかを識別する。訓練済みのＡＮＮが、変換されたＲＲＥＦのセットにマルウェアが含まれていないと判別した場合、ＡＮＮモジュール１２４は、次に、変換モジュール１２２に、変換されたＲＲＥＦのセットを元のデータファイルに再構成させ、それを安全なファイルとしてマークさせる。これらの安全なデータファイルは、データファイル１１２としてホストマシン１１０に戻され、マッピングはメモリ１３０に格納される。
【００６９】
本発明の実施態様では、訓練済みのＡＮＮが、変換されたＲＲＥＦのセットが悪質な動作から構成されると判定した場合、変換されたＲＲＥＦのセットは、そのマッピングされた変換データファイルとともに、メモリ１３０内の仮想の安全な領域に移される。本発明の実施態様では、この仮想の安全な領域は、メモリ１３０内の仮想のサンドボックスであってもよい。
【００７０】
クラスタリングモジュール１２８は、次に、バックプロパゲーションモジュール１２６を用いて、メモリ１３０から種々の変換されたＲＲＥＦのセットを取得するように構成される。本発明の実施態様では、これらのバリエーションは、一連の過去のタイムフレームにわたる変換されたＲＲＥＦのセット、または非常に類似した値および／または数値配置を有するＲＲＥＦのセットを含んでいてもよい。
【００７１】
クラスタリングモジュール１２８は次に、変換されたＲＲＥＦの種々のセットを、メモリ１３０内の仮想の安全な領域に追加する。次に、種々の変換されたＲＲＥＦのセットが、変換されたＲＲＥＦのセットとともにクラスタリングされる。本発明の実施態様では、ｋ平均クラスタリング関数を用いて、これらのＲＲＥＦをクラスタリングする。
【００７２】
クラスタリングモジュール１２８は、次に、悪質な動作を有すると識別されたＲＲＥＦの変換されたセットに関連するＲＲＥＦのセットを特定する。以前に識別されなかったこれらＲＲＥＦのセットは、潜在的に悪質な動作を構成する可能性のあるＲＲＥＦのセットとしてフラグが立てられ、その後、ホストマシン１１０によって不用意にインストールまたは実行されないように、システム１２０によって保護される。変換されたＲＲＥＦのセットに関連しないＲＲＥＦのセットは、対応するマッピングされたデータファイルとともに、安全なファイルとして分類される。
【００７３】
本発明の実施態様によれば、悪質な動作を有すると識別されたＲＲＥＦのセットは、クラスタリングモジュール１２８によって訓練済みのＡＮＮモジュール１２４に供給される。本発明の実施態様によれば、この情報は、ＡＮＮモジュール１２４の訓練済みのＡＮＮをさらに訓練するために利用され、将来、このような類似のマルウェアをよりよく理解し、識別できるようにする。本発明の別の実施態様によれば、訓練済みのＡＮＮモジュール１２４は、ＲＲＥＦのセットに関連する悪質な動作および／またはマルウェアのタイプを識別するために使用される。この情報は、次にクラスタリングモジュール１２８に供給され、モジュール１２８は、これらのＲＲＥＦのセットに対応するデータファイルを保護する

ために、そのアプローチを調整する。例えば、訓練されたＡＮＮモジュール１２４が、バックドア攻撃の以前に発見されていないクＲＲＥＦラスタを識別した場合、訓練されたＡＮＮモジュール１２４は、この情報をクラスタリングモジュール１２８に供給する。この情報を受け取ると、クラスタリングモジュール１２８は（それ自体で、またはシステム１２０の他のモジュールと共に）、ＲＲＥＦのクラスタに関連するデータファイルに関連するすべての着信データ送信を無効にして、バックドア攻撃が行われるのを防ぐことができる。

【００７４】

本発明の別の実施態様によれば、ＡＮＮモジュール１２４は、悪質な動作に関連していることが判明したＲＲＥＦの最下位ビットを保護するロック／キー機構を誘導して、これらのＲＲＥＦが悪質な第三者によって容易に変換されないようにし、対応するデータファイルの脆弱性を無効にする。

【００７５】

本発明の実施態様を実施するためのモジュールのいずれかの中に供給され得る処理システム２００の構成要素の代表的なブロックを図２に示す。当業者であれば、これらのモジュール内に供給される各処理システムの厳密な構成は異なってもよく、図２は例示としてのみ供給されることが当然理解できる。

【００７６】

本発明の実施態様によれば、各モジュールは、コントローラ２０１およびユーザインターフェース２０２を含んでいてもよい。ユーザインターフェース２０２は、必要に応じて、ユーザーとこれらのモジュールとの間に手動の指令を可能にするように配置され、ユーザーがこれらのモジュールの各々に指示を入力するために必要な入出力の構成を含む。当業者であれば、ユーザインターフェース２０２の構成は、実施態様によって異なるが、典型的には、ディスプレイ２４０、キーボード２３５、およびトラックパッド２３６のうちの１つ以上を含むか、あるいは逆に、これらの構成のいずれも含まなくてもよいと当然理解できる。

【００７７】

コントローラ２０１は、バス２１５を介してユーザインターフェース２０２とデータ通信を行い、メモリ２２０と、本実施態様の方法を実行するための命令およびデータを処理する回路基板上に組み込まれたプロセッサ２０５と、オペレーティングシステム２０６と、ユーザインターフェース２０２と通信するための入出力（Ｉ／Ｏ）インターフェース２３０と、ネットワークカード２５０の形態の通信インターフェースとを含む。ネットワークカード２５０は、例えば、有線または無線ネットワークを介してこれらのモジュールから他の処理装置へデータを送信するため、または有線または無線ネットワークを介してデータを受信するために利用することができる。ネットワークカード２５０によって利用可能な無線ネットワークには、ワイヤレスフィデリティ（Ｗｉ－Ｆｉ）、ブルートゥース（登録商標）、近距離無線通信（ＮＦＣ）、セルラーネットワーク、衛星ネットワーク、電気通信ネットワーク、広域ネットワーク（ＷＡＮ）等が含まれるが、これらに限定されない。

【００７８】

メモリ２２０およびオペレーティングシステム２０６は、バス２１０を介してＣＰＵ２０５とデータ通信を行う。メモリは、揮発性メモリと不揮発性メモリの両方を含み、ランダムアクセスメモリ（ＲＡＭ）２２３、リードオンリメモリ（ＲＯＭ）２２５、およびストレージメモリ（図示せず）を含む各タイプのメモリを1つ以上含み、最後のものは１つ以上のソリッドステートドライブ（ＳＳＤ）で構成される。メモリ２２０は、秘密鍵（プライベート鍵）を安全に格納するための安全なストレージ２４６も含む。当業者であれば、上述したメモリは、非一過性のコンピュータ可読媒体を含み、一過性の伝搬信号を除くすべてのコンピュータ可読媒体を含むものと認識することができる。典型的に命令は、メモリにプログラムコードとして格納されるが、ハードワイヤードすることもできる。メモリ２２０は、揮発性メモリまたは不揮発性メモリのいずれかに格納可能なソフトウェアアプリケーションのようなカーネルおよび／またはプログラミングモジュールを含んでいても

よい。
【００７９】
ここで、「プロセッサ」という用語は、そのような命令を処理することができる任意のデバイスまたは構成要素を一般的に指すために使用され、マイクロプロセッサ、マイクロコントローラ、プログラマブルロジックデバイス、または他の計算装置を含んでいてもよい。すなわち、プロセッサ２０５は、入力信号を受信し、メモリに記憶された命令に従って処理され、（例えば、メモリまたはディスプレイ２４０上への）出力を生成するための任意の適切なロジック回路によって供給されてもよい。本実施態様によれば、プロセッサ２０５は、メモリアドレス可能空間を有するシングルコアまたはマルチコアプロセッサであってよい。ある実施態様によれば、プロセッサ２０５は、例えば８コアＣＰＵを含むマルチコアであってもよい。別の実施態様では、計算を加速するために並列に動作するＣＰＵコアのクラスタであってもよい。
【００８０】
図３は、本発明の実施態様に従って、ホストマシンより取得された一連のデータファイルを、対応するＲＲＥＦのセットにマッピングするためのプロセス３００を示し、プロセス３００はシステム１２０に組み込まれていてもよい。システム１２０はホストマシン１１０に通信可能に接続され、システム１２０がホストマシン１１０のバスシステム上で伝送される、またはホストマシン１１０のメモリに格納されるすべてのデータを捕捉できると仮定する。
【００８１】
プロセス３００は、ステップ３０２において、ホストマシン内に格納されている、またはホストマシンにより送信／受信されている一連のデータファイルからデータファイルを選択することから始まる。次に、プロセス３００は、ステップ３０４において、データファイルをＲＲＥＦのセットに変換する。ステップ３０６において、プロセス３００は、ＲＲＥＦのセットを記憶し、ＲＲＥＦのセットをデータファイルにマッピングする。
【００８２】
プロセス３００は次に、ステップ３０８で別のデータファイルを選択する。次に、ステップ３１０において、データファイルは、ＲＲＥＦの別のセットに変換される。ステップ３１２において、プロセス３００は、ＲＲＥＦの別のセットを記憶し、ＲＲＥＦのセットを別のデータファイルにマッピングする。その後、プロセス３００はステップ３０８に進む。別のデータファイルを選択し、データファイルを変換し、データファイルをマッピングする工程３１４は、一連のデータファイルのすべてのデータファイルが工程３００によって処理されるまで繰り返される。その後、プロセス３００は終了する。
【００８３】
図４は、プロセス４００がシステム１２０内のモジュールにより実施される、本発明の実施態様に従ってホストマシンから取得されたデータファイルを保護するためのプロセス４００を示す。プロセス４００はステップ４０２から始まり、プロセス４００によって変換されたデータファイル、すなわち最初のデータファイルがホストマシンより取得される。この変換されたデータファイルは、（図３で説明したように）プロセス３００によって以前に記憶され処理された種々のデータファイルを含んでいてもよい。次に、プロセス４００は、ステップ４０４において、第一のデータファイルを第一のＲＲＥＦのセットに変換する。ステップ４０６において、プロセス４００は、ＲＲＥＦの最初のセットがマルウェアまたは悪質な動作含むかどうかを検出する。プロセス４００が、ＲＲＥＦの最初のセットがマルウェアを含まないと検出した場合、プロセス４００はステップ４０８へ進み、第一のデータファイルとそれに対応するＲＲＥＦのセットが、対応するマッピングとともに格納される。
【００８４】
逆に、ステップ４０６において、プロセス４００は、ＲＲＥＦの最初のセットが悪質な動作に関連していると判定した場合、ステップ４１０に進む。ステップ４１０において、ＲＲＥＦの最初のセットは、そのマッピングされた最初のデータファイルとともに、仮想の

安全な領域に移される。プロセス４００は次に、ステップ４１２で、メモリから種々のＲＲＥＦの最初のセットを取得する。次に、プロセス４００によって、ＲＲＥＦの最初のセットとともに、種々のＲＲＥＦの最初のセットがクラスタリングされる。これはステップ４１４で行われる。ステップ４１４で行われたクラスタリングの結果に基づき、ステップ４１６において、プロセス４００は、悪質な動作を含む種々のＲＲＥＦの最初のセット（以前は安全であると考えられていた）を検出し、また、悪質な動作を含まない種々のＲＲＥＦの最初のセットバリエーションを検出する。次に、ステップ４１８で、プロセス４００によって、悪質な動作を含まないＲＲＥＦの最初のセットに基づいて、最初のデータファイルが復元される。ステップ４２０で、悪質な動作を含む種々のＲＲＥＦの最初のセットは、前記において説明したように、さらなる処理のために、プロセス４００によって訓練されたＡＮＮに供給される。

【００８５】

当業者が実施可能な範囲で、本明細書に具体的に開示された実施態様の一部を変換、置換、変形および修正することはすべて、添付の特許請求の範囲に含まれるものとする。

【図面】

【図１】

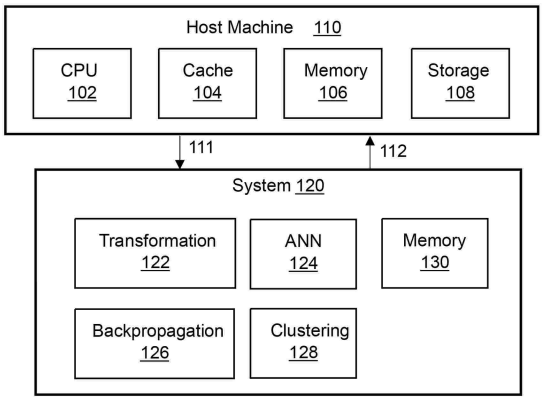

図１

【図２】

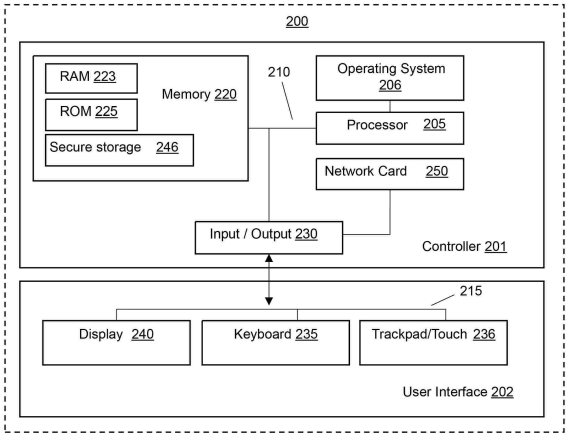

図２

【図３】

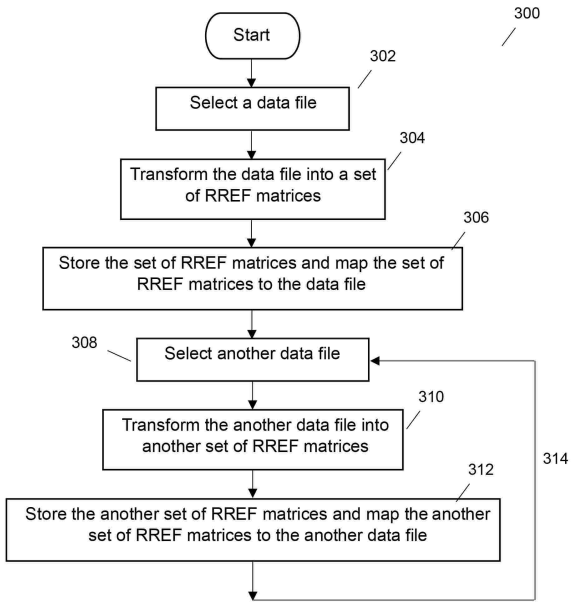

図３

【図４】

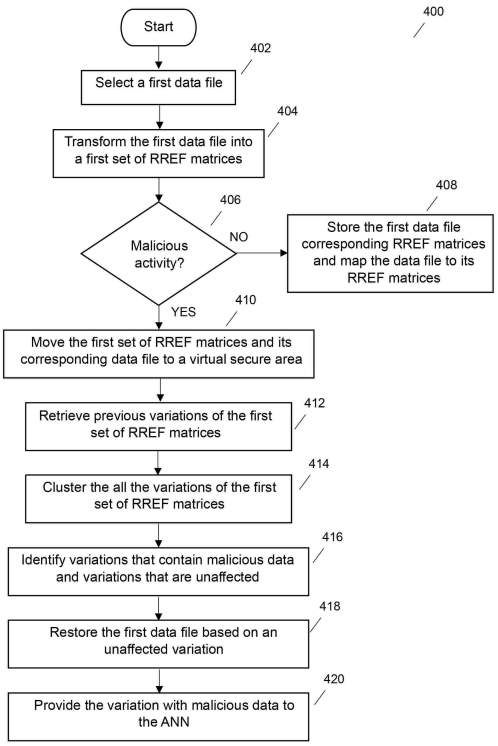

図４

【図５】

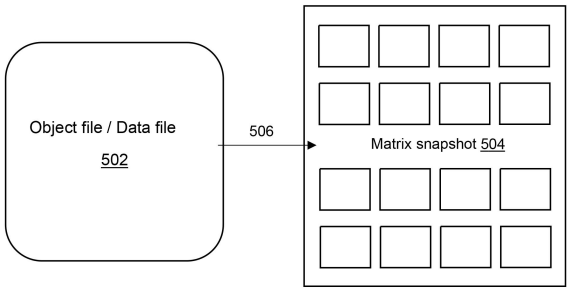
Object file / Data file
502
506
Matrix snapshot 504

図５

【図６】

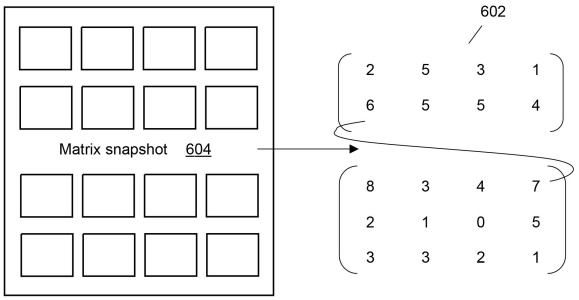
Matrix snapshot 604
602
2 5 3 1
6 5 5 4
8 3 4 7
2 1 0 5
3 3 2 1

図６

フロントページの続き

(72)発明者　タン・ホン・チュアン
シンガポール プラチナム２８ ゲンティングレーン　２８番 ＃０９－０３

審査官　上島 拓也

(56)参考文献　特開２０２０－００９４１５（ＪＰ，Ａ）
特表２０１９－５２７４４７（ＪＰ，Ａ）
米国特許出願公開第２０１９／０１６３９０４（ＵＳ，Ａ１）
米国特許出願公開第２０２０／０２５０３０９（ＵＳ，Ａ１）

(58)調査した分野　(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　２１／５６