(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7555494号
(P7555494)

(45)発行日　令和6年9月24日(2024.9.24)

(24)登録日　令和6年9月12日(2024.9.12)

(51)国際特許分類
Ｈ０４Ｌ　43/022 (2022.01)

ＦＩ
Ｈ０４Ｌ　43/022

請求項の数　24　（全22頁）

(21)出願番号　特願2023-535466(P2023-535466)
(86)(22)出願日　令和3年12月9日(2021.12.9)
(65)公表番号　特表2023-553140(P2023-553140A)
(43)公表日　令和5年12月20日(2023.12.20)
(86)国際出願番号　PCT/US2021/062713
(87)国際公開番号　WO2022/125842
(87)国際公開日　令和4年6月16日(2022.6.16)
審査請求日　令和5年8月2日(2023.8.2)
(31)優先権主張番号　17/120,050
(32)優先日　令和2年12月11日(2020.12.11)
(33)優先権主張国・地域又は機関
米国(US)

(73)特許権者　502208397
グーグル エルエルシー
Ｇｏｏｇｌｅ ＬＬＣ
アメリカ合衆国 カリフォルニア州 ９４０４３ マウンテン ビュー アンフィシアター パークウェイ １６００
１６００ Ａｍｐｈｉｔｈｅａｔｒｅ Ｐａｒｋｗａｙ ９４０４３ Ｍｏｕｎｔａｉｎ Ｖｉｅｗ， ＣＡ Ｕ．Ｓ．Ａ．
(74)代理人　110001195
弁理士法人深見特許事務所
(72)発明者　ペパン，フランソワ
アメリカ合衆国、９４０４３ カリフォルニア州、マウンテン・ビュー、アンフィシアター・パークウェイ、１６００

最終頁に続く

(54)【発明の名称】　サービス拒否攻撃の自動検出および軽減

(57)【特許請求の範囲】

【請求項１】

コンピュータにより実現される方法（６００）であって、データ処理ハードウェア（１４４）によって実行されると、前記データ処理ハードウェア（１４４）に、動作を実行させ、前記動作は、

ネットワークサービス（３０）によって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージ（２２）の第１のセットを取得することと、

第１のモデル（３１０）を介して、前記ネットワークトラフィックメッセージ（２２）の第１のセットに基づいてネットワーク不正使用が発生しているかどうかを判断することと、

前記ネットワーク不正使用が発生しているとき、

前記ネットワークサービス（３０）によって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージ（２２）の第２のセットを取得することと、

前記ネットワークトラフィックメッセージの第２のセット（２２）内の各ネットワークトラフィックメッセージ（２２）について、第２のモデル（１７２）を介して、前記ネットワークトラフィックメッセージ（２２）を、前記ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージ（２２）または前記ネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージ（２２）としてラベル付けすることと、

第３のモデル（４１０）を介して、少なくとも１つのネットワークトラフィックルール（４０２）を生成し、各ネットワークトラフィックルール（４０２）は、ファイアウォールによって実現されるよう構成され、実現されると、前記不正ネットワークトラフィックメッセージ（２２）の影響を低減する、方法（６００）。

【請求項２】

前記ネットワークトラフィックメッセージ（２２）の第１のセットは、複数のネットワークトラフィックウィンドウ（２１０）を含み、前記複数のネットワークトラフィックウィンドウ（２１０）の各々は、異なる離散的な時間の部分に関連付けられる、前記ネットワークトラフィックメッセージ（２２）の第１のセットのネットワークトラフィックメッセージ（２２）のサブセットを含む、請求項１に記載の方法（６００）。

【請求項３】

前記動作は、前記複数のネットワークトラフィックウィンドウ（２１０）の各々について、前記ネットワークトラフィックメッセージ（２２）のサブセットから、サンプリングされたネットワークトラフィックメッセージ（２２Ｓ）のセットをサンプリングすることをさらに含み、前記サンプリングされたネットワークトラフィックメッセージ（２２Ｓ）のセットは、前記ネットワークトラフィックメッセージ（２２）のサブセットの全体を表す、請求項２に記載の方法（６００）。

【請求項４】

前記動作は、前記複数のネットワークトラフィックウィンドウ（２１０）の各々について、前記サンプリングされたネットワークトラフィックメッセージ（２２Ｓ）のセットの特性を、データ構造に記憶することをさらに含む、請求項３に記載の方法（６００）。

【請求項５】

前記ネットワーク不正使用が発生しているかどうかを判断することは、

前記第１のモデル（３１０）が、不正確率スコア（３１２）を生成することと、

前記不正確率スコア（３１２）は不正確率閾値を満たす、と判断することとを含む、請求項１～４のいずれか１項に記載の方法（６００）。

【請求項６】

前記第１のモデル（３１０）は、ラベル付けされたネットワークトラフィックメッセージ（２２Ｌ）のセットでトレーニングされたニューラルネットワークを含む、請求項１～５のいずれか１項に記載の方法（６００）。

【請求項７】

前記動作は、前記ネットワーク不正使用が発生していると判断した後に、

前記発生しているネットワーク不正使用が偽陽性または実際のネットワーク不正使用のいずれかであったことを示すフィードバック（３２０）を受信することと、

前記フィードバック（３２０）に基づいて前記第１のモデル（３１０）を更新することとをさらに含む、請求項１～６のいずれか１項に記載の方法（６００）。

【請求項８】

前記動作はさらに、

前記ネットワークサービス（３０）によって以前に受信されたネットワークトラフィックを表す履歴ネットワークトラフィックメッセージ（２２Ｈ）のセットを取得することを含み、

前記履歴ネットワークトラフィックメッセージ（２２Ｈ）のセットは、前記ネットワーク不正使用前に前記ネットワークサービスによって以前に受信されたネットワークトラフィック（２２Ｈ）を表し、

前記ネットワークトラフィックメッセージ（２２）を、前記ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージ（２２）または前記ネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージ（２２）としてラベル付けすることは、前記履歴ネットワークトラフィックメッセージ（２２Ｈ）のセットに基づく、請求項１～７のいずれか１項に記載の方法（６００）。

【請求項９】

前記動作はさらに、
前記生成された少なくとも１つのネットワークトラフィックルール（４０２）を、前記ネットワークサービス（３０）に関連付けられるユーザ（１２）に提供することと、
前記ユーザ（１２）から、前記生成された少なくとも１つのネットワークトラフィックルール（４０２）のうちの１つを受け入れる指示を受信することと、
前記ユーザ（１２）によって示される前記受け入れられたネットワークトラフィックルール（４０２）を実現することとを含む、請求項１～８のいずれか１項に記載の方法（６００）。
【請求項１０】
前記動作はさらに、
所望のネットワークトラフィックルール（４０２）に関連付けられるユーザプリファレンス（５１０）を受信することと、
前記ユーザプリファレンス（５１０）に基づいて、前記生成された少なくとも１つのネットワークトラフィックルール（４０２）のうちの１つを選択することと、
前記選択されたネットワークトラフィックルール（４０２）を実現することとを含む、請求項１～９のいずれか１項に記載の方法（６００）。
【請求項１１】
前記ユーザプリファレンス（５１０）は、前記選択されたネットワークトラフィックルール（４０２）によって影響されてもよい非不正ネットワークトラフィックメッセージ（２２）の量を含む、請求項１０に記載の方法（６００）。
【請求項１２】
前記ネットワーク不正使用はサービス拒否攻撃を含む、請求項１～１１のいずれか１項に記載の方法（６００）。
【請求項１３】
システム（１００）であって、
データ処理ハードウェア（１４４）と、
前記データ処理ハードウェア（１４４）と通信するメモリハードウェア（１４２）とを備え、前記メモリハードウェア（１４２）は、前記データ処理ハードウェア（１４４）上で実行されると前記データ処理ハードウェア（１４４）に動作を実行させる命令を記憶し、前記動作は、
ネットワークサービス（３０）によって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージ（２２）の第１のセットを取得することと、
第１のモデル（３１０）を介して、前記ネットワークトラフィックメッセージ（２２）の第１のセットに基づいてネットワーク不正使用が発生しているかどうかを判断することと、
前記ネットワーク不正使用が発生しているとき、
前記ネットワークサービス（３０）によって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージ（２２）の第２のセットを取得することと、
前記ネットワークトラフィックメッセージの第２のセット（２２）内の各ネットワークトラフィックメッセージ（２２）について、第２のモデル（１７２）を介して、前記ネットワークトラフィックメッセージ（２２）を、前記ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージ（２２）または前記ネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージ（２２）としてラベル付けすることと、
第３のモデル（４１０）を介して、少なくとも１つのネットワークトラフィックルール（４０２）を生成し、各ネットワークトラフィックルール（４０２）は、ファイアウォールによって実現されるよう構成され、実現されると、前記不正ネットワークトラフィックメッセージ（２２）の影響を低減する、システム（１００）。
【請求項１４】

前記ネットワークトラフィックメッセージ（２２）の第１のセットは、複数のネットワークトラフィックウィンドウ（２１０）を含み、前記複数のネットワークトラフィックウィンドウ（２１０）の各々は、異なる離散的な時間の部分に関連付けられる、前記ネットワークトラフィックメッセージ（２２）の第１のセットのネットワークトラフィックメッセージ（２２）のサブセットを含む、請求項１３に記載のシステム（１００）。

【請求項１５】

前記動作は、前記複数のネットワークトラフィックウィンドウ（２１０）の各々について、前記ネットワークトラフィックメッセージ（２２）のサブセットから、サンプリングされたネットワークトラフィックメッセージ（２２Ｓ）のセットをサンプリングすることをさらに含み、前記サンプリングされたネットワークトラフィックメッセージ（２２Ｓ）のセットは、前記ネットワークトラフィックメッセージ（２２）のサブセットの全体を表す、請求項１４に記載のシステム（１００）。

【請求項１６】

前記動作は、前記複数のネットワークトラフィックウィンドウ（２１０）の各々について、前記サンプリングされたネットワークトラフィックメッセージ（２２Ｓ）のセットの特性を、データ構造に記憶することをさらに含む、請求項１５に記載のシステム（１００）。

【請求項１７】

前記ネットワーク不正使用が発生しているかどうかを判断することは、

前記第１のモデル（３１０）が、不正確率スコア（３１２）を生成することと、

前記不正確率スコア（３１２）は不正確率閾値を満たす、と判断することとを含む、請求項１３～１６のいずれか１項に記載のシステム（１００）。

【請求項１８】

前記第１のモデル（３１０）は、ラベル付けされたネットワークトラフィックメッセージ（２２Ｌ）のセットでトレーニングされたニューラルネットワークを含む、請求項１３～１７のいずれか１項に記載のシステム（１００）。

【請求項１９】

前記動作は、前記ネットワーク不正使用が発生していると判断した後に、

前記発生しているネットワーク不正使用が偽陽性または実際のネットワーク不正使用のいずれかであったことを示すフィードバック（３２０）を受信することと、

前記フィードバック（３２０）に基づいて前記第１のモデル（３１０）を更新することとをさらに含む、請求項１３～１８のいずれか１項に記載のシステム（１００）。

【請求項２０】

前記動作はさらに、

前記ネットワークサービス（３０）によって以前に受信されたネットワークトラフィックを表す履歴ネットワークトラフィックメッセージ（２２Ｈ）のセットを取得することを含み、

前記履歴ネットワークトラフィックメッセージ（２２Ｈ）のセットは、前記ネットワーク不正使用前に前記ネットワークサービス（３０）によって以前に受信されたネットワークトラフィックを表し、

前記ネットワークトラフィックメッセージ（２２）を、前記ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージ（２２）または前記ネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージ（２２）としてラベル付けすることは、前記履歴ネットワークトラフィックメッセージ（２２Ｈ）のセットに基づく、請求項１３～１９のいずれか１項に記載のシステム（１００）。

【請求項２１】

前記動作はさらに、

前記生成された少なくとも１つのネットワークトラフィックルール（４０２）を、前記ネットワークサービス（３０）に関連付けられるユーザ（１２）に提供することと、

前記ユーザ（１２）から、前記生成された少なくとも１つのネットワークトラフィック

ルール（４０２）のうちの１つを受け入れる指示を受信することと、
前記ユーザ（１２）によって示される前記受け入れられたネットワークトラフィックルール（４０２）を実現することとを含む、請求項１３～２０のいずれか１項に記載のシステム（１００）。
【請求項２２】
前記動作はさらに、
所望のネットワークトラフィックルール（４０２）に関連付けられるユーザプリファレンス（５１０）を受信することと、
前記ユーザプリファレンス（５１０）に基づいて、前記生成された少なくとも１つのネットワークトラフィックルール（４０２）のうちの１つを選択することと、
前記選択されたネットワークトラフィックルール（４０２）を実現することとを含む、請求項１３～２１のいずれか１項に記載のシステム（１００）。
【請求項２３】
前記ユーザプリファレンス（５１０）は、前記選択されたネットワークトラフィックルール（４０２）によって影響されてもよい非不正ネットワークトラフィックメッセージ（２２）の量を含む、請求項２２に記載のシステム（１００）。
【請求項２４】
前記ネットワーク不正使用はサービス拒否攻撃を含む、請求項１３～２３のいずれか１項に記載のシステム（１００）。
【発明の詳細な説明】
【技術分野】
【０００１】
本開示は、サービス拒否（ＤｏＳ）攻撃などのネットワーク不正使用の自動検出および軽減に関する。
【背景技術】
【０００２】
背景
ネットワーク不正使用は、典型的には、ネットワークの許容可能な使用ポリシーによって禁止される目的のためのコンピュータネットワークの使用として定義される。ネットワーク不正使用の一般的な例は、サービス拒否（ＤｏＳ）攻撃である。ＤｏＳ攻撃は、正当なユーザが、１人以上の悪意のある行為者の行為に起因して、情報システム、デバイス、または他のネットワークリソースにアクセスできないときに発生する。影響を受けるサービスは、電子メール、ウェブサイト、オンラインアカウント（例えば、銀行業務）、または影響を受けるコンピュータもしくはネットワークに依存する他のサービスを含み得る。ＤｏＳ攻撃は、典型的には、ターゲットが応答することができなくなる（例えば、クラッシュする）まで、ターゲットホストまたはネットワークをネットワークトラフィックでフラッディングすることによって達成され、正当なユーザのためのアクセスを妨げる。多くのタイプのネットワーク不正使用は、検出および軽減を回避するために正当なトラフィックを模倣する。これらの攻撃および他のネットワーク不正使用は、リソースおよびサービスをアクセス不可能にすることによって、組織に時間および費用の両方を費やさせ得る。
【発明の概要】
【０００３】
概要
本開示の一局面は、データ処理ハードウェアによって実行されると、データ処理ハードウェアに動作を実行させる、コンピュータにより実現される方法を提供する。動作は、ネットワークサービスによって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージの第１のセットを取得することを含む。動作はまた、第１のモデルを介して、ネットワークトラフィックメッセージの第１のセットに基づいてネットワーク不正使用が発生しているかどうかを判断することを含む。ネットワーク不正使用が発生しているとき、動作は、ネットワークサービスによって現在受信されているネット

ワークトラフィックを表すネットワークトラフィックメッセージの第２のセットを取得することを含む。動作はまた、ネットワークトラフィックメッセージの第２のセット中の各ネットワークトラフィックメッセージについて、第２のモデルを介して、ネットワークトラフィックメッセージを、ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージまたはネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージとしてラベル付けすることを含む。本方法は、第３のモデルを介して、少なくとも１つのネットワークトラフィックルールを生成することを含む。各ネットワークトラフィックルールは、ファイアウォールによって実現されるよう構成され、実現されると、不正ネットワークトラフィックメッセージの影響を低減するよう構成される。

【０００４】

本開示の実現例は、以下の任意選択の特徴のうちの１つ以上を含んでもよい。いくつかの実現例では、ネットワークトラフィックメッセージの第１のセットは、複数のネットワークトラフィックウィンドウを含む。複数のネットワークトラフィックウィンドウの各々は、ネットワークトラフィックメッセージの第１のセットのうち、異なる離散的な時間の部分に関連付けられるネットワークトラフィックメッセージのサブセットを含む。いくつかの例では、動作は、複数のネットワークトラフィックウィンドウの各々について、ネットワークトラフィックメッセージのサブセットから、サンプリングされたネットワークトラフィックメッセージのセットをサンプリングすることをさらに含む。サンプリングされたネットワークトラフィックメッセージのセットは、ネットワークトラフィックメッセージのサブセットの全体を表す。任意選択で、動作は、複数のネットワークトラフィックウィンドウの各々について、サンプリングされたネットワークトラフィックメッセージのセットの特性を、データ構造に記憶することをさらに含む。

【０００５】

いくつかの実現例では、ネットワーク不正使用が発生しているかどうかを判断することは、第１のモデルが、不正確率スコアを生成することと、不正確率スコアが不正確率閾値を満たす、と判断することとを含む。いくつかの例では、第１のモデルは、ラベル付けされたネットワークトラフィックメッセージのセットでトレーニングされたニューラルネットワークを含む。動作は、ネットワーク不正使用が発生していると判断した後、発生しているネットワーク不正使用が偽陽性または実際のネットワーク不正使用のいずれかであったことを示すフィードバックを受信することと、フィードバックに基づいて第１のモデルを更新することとをさらに含んでもよい。

【０００６】

いくつかの例では、動作は、過去にネットワークサービスによって受信されたネットワークトラフィックを表す履歴ネットワークトラフィックメッセージのセットを取得することをさらに含む。履歴ネットワークトラフィックメッセージのセットは、ネットワーク不正使用前にネットワークサービスによって受信されたネットワークトラフィックを表してもよい。ネットワークトラフィックメッセージを、ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージまたはネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージとしてラベル付けすることは、履歴ネットワークトラフィックメッセージのセットに基づいてもよい。動作は、生成された少なくとも１つのネットワークトラフィックルールを、ネットワークサービスに関連付けられるユーザに提供することと、ユーザから、生成された少なくとも１つのネットワークトラフィックルールのうちの１つを受け入れる指示を受信することとをさらに含んでもよい。動作はまた、ユーザによって示される受け入れられたネットワークトラフィックルールを実現することをさらに含んでもよい。

【０００７】

任意選択で、動作は、所望のネットワークトラフィックルールに関連付けられるユーザプリファレンスを受信することと、ユーザプリファレンスに基づいて、生成された少なくとも１つのネットワークトラフィックルールのうちの１つを選択することとをさらに含む。動作はまた、任意選択で、選択されたネットワークトラフィックルールを実現すること

をさらに含む。いくつかの例では、ユーザプリファレンスは、選択されたネットワークトラフィックルールによって影響されてもよい非不正ネットワークトラフィックメッセージの量を含む。いくつかの実現例では、ネットワーク不正使用はサービス拒否攻撃を含む。

【０００８】

本開示の別の局面は、ネットワーク不正使用を検出および軽減するためのシステムを提供する。本システムは、データ処理ハードウェアと、データ処理ハードウェアと通信するメモリハードウェアとを含む。メモリハードウェアは、データ処理ハードウェア上で実行されるとデータ処理ハードウェアに動作を実行させる命令を記憶する。動作は、ネットワークサービスによって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージの第１のセットを取得することを含む。動作はまた、第１のモデルを介して、ネットワークトラフィックメッセージの第１のセットに基づいてネットワーク不正使用が発生しているかどうかを判断することを含む。ネットワーク不正使用が発生しているとき、動作は、ネットワークサービスによって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージの第２のセットを取得することを含む。動作はまた、ネットワークトラフィックメッセージの第２のセット中の各ネットワークトラフィックメッセージについて、第２のモデルを介して、ネットワークトラフィックメッセージを、ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージまたはネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージとしてラベル付けすることを含む。本方法は、第３のモデルを介して、少なくとも１つのネットワークトラフィックルールを生成することを含む。各ネットワークトラフィックルールは、ファイアウォールによって実現されるよう構成され、実現されると、不正ネットワークトラフィックメッセージの影響を低減するよう構成される。

【０００９】

この態様は、以下の任意選択の特徴のうちの１つ以上を含んでもよい。いくつかの実現例では、ネットワークトラフィックメッセージの第１のセットは、複数のネットワークトラフィックウィンドウを含む。複数のネットワークトラフィックウィンドウの各々は、ネットワークトラフィックメッセージの第１のセットのうち、異なる離散的な時間の部分に関連付けられるネットワークトラフィックメッセージのサブセットを含む。いくつかの例では、動作は、複数のネットワークトラフィックウィンドウの各々について、ネットワークトラフィックメッセージのサブセットから、サンプリングされたネットワークトラフィックメッセージのセットをサンプリングすることをさらに含む。サンプリングされたネットワークトラフィックメッセージのセットは、ネットワークトラフィックメッセージのサブセットの全体を表す。任意選択で、動作は、複数のネットワークトラフィックウィンドウの各々について、サンプリングされたネットワークトラフィックメッセージのセットの特性を、データ構造に記憶することをさらに含む。

【００１０】

いくつかの実現例では、ネットワーク不正使用が発生しているかどうかを判断することは、第１のモデルが、不正確率スコアを生成することと、不正確率スコアが不正確率閾値を満たす、と判断することとを含む。いくつかの例では、第１のモデルは、ラベル付けされたネットワークトラフィックメッセージのセットでトレーニングされたニューラルネットワークを含む。動作は、ネットワーク不正使用が発生していると判断した後、発生しているネットワーク不正使用が偽陽性または実際のネットワーク不正使用のいずれかであったことを示すフィードバックを受信することと、フィードバックに基づいて第１のモデルを更新することとをさらに含んでもよい。

【００１１】

いくつかの例では、動作は、過去にネットワークサービスによって受信されたネットワークトラフィックを表す履歴ネットワークトラフィックメッセージのセットを取得することをさらに含む。履歴ネットワークトラフィックメッセージのセットは、ネットワーク不正使用前にネットワークサービスによって受信されたネットワークトラフィックを表してもよい。ネットワークトラフィックメッセージを、ネットワーク不正使用に関与している

不正ネットワークトラフィックメッセージまたはネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージとしてラベル付けすることは、履歴ネットワークトラフィックメッセージのセットに基づいてもよい。動作は、生成された少なくとも１つのネットワークトラフィックルールを、ネットワークサービスに関連付けられるユーザに提供することと、ユーザから、生成された少なくとも１つのネットワークトラフィックルールのうちの１つを受け入れる指示を受信することとをさらに含んでもよい。動作はまた、ユーザによって示される受け入れられたネットワークトラフィックルールを実現することをさらに含んでもよい。

【００１２】

任意選択で、動作は、所望のネットワークトラフィックルールに関連付けられるユーザプリファレンスを受信することと、ユーザプリファレンスに基づいて、生成された少なくとも１つのネットワークトラフィックルールのうちの１つを選択することとをさらに含む。動作はまた、任意選択で、選択されたネットワークトラフィックルールを実現することをさらに含む。いくつかの例では、ユーザプリファレンスは、選択されたネットワークトラフィックルールによって影響されてもよい非不正ネットワークトラフィックメッセージの量を含む。いくつかの実現例では、ネットワーク不正使用はサービス拒否攻撃を含む。

【００１３】

本開示の１つ以上の実現例の詳細を、添付の図面および以下の説明に記載する。他の局面、特徴、および利点は、説明および図面、ならびに特許請求の範囲から明らかになるであろう。

【図面の簡単な説明】

【００１４】

【図１】ネットワーク不正使用を検出および軽減するための例示的なシステムの概略図である。

【図２】ネットワークトラフィックメッセージのための例示的なネットワークトラフィックウィンドウの概略図である。

【図３】図１のシステムの例示的な構成要素の概略図である。

【図４】図１のシステムの追加の例示的な構成要素の概略図である。

【図５】ネットワークトラフィックルールに対するユーザプリファレンスの概略図である。

【図６】ネットワーク不正使用を検出および軽減する方法のための動作の例示的な構成のフローチャートである。

【図７】本明細書で説明されるシステムおよび方法を実現するために使用されてもよい例示的なコンピューティングデバイスの概略図である。

【発明を実施するための形態】

【００１５】

様々な図面における同様の参照符号は、同様の要素を示す。

詳細な説明

ネットワークサービスがますますユビキタスになるにつれ、ネットワーク不正使用も同様に日常的になっている。サービス拒否（ＤｏＳ）攻撃などの攻撃は、正当なユーザに対するリソースおよびサービスに対するアクセスを拒否するために頻繁に使用される。これらの攻撃および他のネットワーク不正使用は、組織に時間および費用の両方を費やさせる著しい混乱を引き起こし得る。組織およびクラウドサービスプロバイダは、ネットワークトラフィックルールを積極的に使用してネットワーク不正使用を軽減しようとするが、存在する攻撃は多種多様であり、多くの攻撃は正当なトラフィックのように見える。正当なトラフィックを遮断することは、ネットワーク不正使用を軽減する目的をくじくので、ネットワーク不正使用の検出および軽減は、概して、複雑で労働集約的な努力である。

【００１６】

従来、ネットワーク管理者は、（例えば、ファイアウォールを介して適所に置かれる）手動ネットワークトラフィックルールの混合を通して、および特定のアプリケーションのためのソフトウェア定義ルールを介して、ネットワーク不正使用防御を実施する。しかし

ながら、オフプレミスの、クラウドベースのサービスの出現に伴い、ネットワーク管理者は、今や、代わりに、ネットワークトラフィックルールおよび／またはポリシーを管理ならびに実施するために、クラウドプロバイダのインフラストラクチャに依拠することが多い。従来の技術は、細かい粒度で問題に対処できない。すなわち、いくつかの従来の技術は、いつ攻撃が発生しているかを識別するが、ネットワークトラフィックのどの部分が攻撃の一部であり、ネットワークトラフィックのどの部分が正当であるかを識別することはできない。これらの技術は、典型的には、トラフィックの増加がネットワーク不正使用であるか正当なトラフィックであるかを検出することが苦手であり、偽陽性につながる。加えて、従来の技術は、いかなる適切なスケーリングも欠いており、代わりに、ライブ（すなわち、「リアルタイム」）で実行するには計算費用がかかりすぎ、および／または反応する熟練者を必要とするかもしくはそのような熟練者に依存する。そのような熟練者は、概して、反応時間が相対的に遅く、費用がかかる。

【００１７】

本明細書での実現例は、軽減を、ネットワーク不正使用がいつ発生するかをリアルタイムまたは近リアルタイムで判断する高速検出フェーズと、相対的により低速なトラフィック分類フェーズとに分割する、ネットワーク不正使用軽減部を対象とする。検出アルゴリズムは、ネットワーク不正使用が進行中であるときを検出するために連続的に動作する。検出アルゴリズムが、ネットワーク不正使用が進行中であると判断すると、署名生成アルゴリズムが、ネットワークトラフィックのいくつかまたはすべてを、正常な（すなわち、正当な）トラフィックまたは不正トラフィックのいずれかとして分類する。署名生成アルゴリズムは、不正トラフィックの鍵となる特性を抽出し、不正トラフィックの影響を低減する（たとえば、不正トラフィックを遮断する）ように設計された、１つ以上のネットワークトラフィックルールを生成する。したがって、いくつかの実現例では、ネットワーク不正使用軽減部は、人的介入を必要とせずにリアルタイムでネットワーク不正使用を検出および軽減する。

【００１８】

図１を参照すると、いくつかの実現例では、例示的なシステム１００は、ネットワーク１１２を介して複数の遠隔コンピューティングデバイス２０、２０ａ～ｎと通信する遠隔システム１４０を含む。遠隔システム１４０はスケーラブル／弾性コンピューティングリソース１４４（例えば、データ処理ハードウェア）および／もしくはストレージリソース１４２（例えば、メモリハードウェア）を有する、単一のコンピュータ、複数のコンピュータ、または分散型システム（例えば、クラウド環境）であってもよい。遠隔コンピューティングデバイス２０は、ネットワーク１１２を介して（例えば、インターネットを介して）アクセス可能な任意のネットワーク接続されたコンピューティングデバイスであってもよい。

【００１９】

遠隔システム１４０はまた、１つ以上のネットワークサービス３０も含む。ネットワークサービス３０は、ユーザデバイス１０を介して、同じであるかまたは異なるネットワーク１１２を通じて遠隔システム１４０と通信するユーザ１２（例えば、遠隔システム１４０の顧客、クライアント、管理者など）のために、ネットワーク関連サービスを提供する。ユーザデバイス１０は、デスクトップワークステーション、ラップトップワークステーション、またはモバイルデバイス（すなわち、スマートフォン）などの任意のコンピューティングデバイスに対応してもよい。ユーザデバイス１０は、コンピューティングリソース１８（例えば、データ処理ハードウェア）および／またはストレージリソース１６（例えば、メモリハードウェア）を含む。

【００２０】

ネットワークサービス３０は、ホスティング、負荷分散、ルーティングなどのサービスのためにクラウドネットワーク（例えば、遠隔システム１４０）のリソースをユーザ１２に提供する。例えば、ユーザ１２は、リソース（例えば、ウェブサイト）をホストし、ネットワークサービス３０は、遠隔システム１４０のクラウド環境を通して、インターネッ

ト（すなわち、ネットワーク１１２）を介して、リソースへのアクセスを複数の遠隔コンピューティングデバイス２０に提供する。ネットワークサービス３０をネットワーク不正使用から保護するのを助けるために、遠隔システム１４０は、ネットワーク不正使用軽減部１６０を実行する。

【００２１】

ネットワーク不正使用軽減部１６０は、ネットワーク不正使用のために遠隔コンピューティングデバイス２０とネットワークサービス３０との間で通信されるネットワークトラフィックメッセージ２２を監視する。ネットワーク不正使用は、ユーザポリシーによって禁止された態様でのコンピュータネットワークの使用として定義される。典型的なネットワーク不正使用は、サービス拒否（ＤｏＳ）攻撃、分散型サービス拒否（ＤＤｏＳ）攻撃、ページススクレイピング、およびクレデンシャルスタッフィング攻撃などの行為を含む。ネットワーク不正使用は、典型的には、リソース（例えば、時間、金銭など）の著しい混乱および損失を引き起こし得る、サービスの劣化または損失をもたらす。たとえば、ＤＤｏＳ攻撃は、大量のネットワークトラフィックでネットワークサービス（またはそのサポートインフラストラクチャ）を圧倒することによってネットワークサービスへのネットワークトラフィックを混乱させる悪意のある試みである。しかしながら、正当なトラフィックを遮断することは、それがＤＤｏＳ攻撃の目標の達成になり、望ましくないので、そのようなネットワーク不正使用を軽減することは困難である。

【００２２】

ネットワーク不正使用軽減部１６０は、不正検出部３００を含む。不正検出部３００は、ネットワークサービス３０によって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージ２２の第１のセット２２ａを取得する。すなわち、不正検出部３００は、遠隔コンピューティングデバイス２０のうちの１つ以上からネットワークサービス３０宛てのリアルタイムまたはほぼリアルタイムのネットワークトラフィックメッセージ２２を取得する。いくつかの例では、ネットワークトラフィックメッセージは、Ｌ７（すなわち、ＨＴＴＰ）ヘッダフィールドのうちのいくつかまたはすべてである。

【００２３】

ここで図２を参照すると、いくつかの例では、ネットワークトラフィックメッセージ２２の第１のセットは、複数のネットワークトラフィックウィンドウ２１０、２１０ａ～ｎを含む。概略図２００に示されるように、各ネットワークトラフィックウィンドウ２１０は、ネットワークトラフィックメッセージ２２の第１のセットのうち、異なる離散的な時間の部分に関連付けられるネットワークトラフィックメッセージ２２のサブセットを含む。概略図２００では、ｘ軸は時間の経過を示し、各ネットワークトラフィックウィンドウ２１０はその時間の一部にわたる。すなわち、ネットワークトラフィックメッセージ２２の第１のセットが、不正検出部３００によって経時的に取得されたすべてのネットワークトラフィックメッセージ２２を含むとき、各ネットワークトラフィックウィンドウ２１０は、それらのネットワークトラフィックメッセージ２２のうち、離散的な時間の部分（例えば、１５秒、３０秒、１分など）にわたって受信されるサブセットを表す。各ネットワークトラフィックウィンドウ２１０の「サイズ」（すなわち、離散的な時間の部分の長さ）は、均一であってもよく、または変化するトラフィック傾向に基づいて動的であってもよい。たとえば、ネットワークトラフィックメッセージ２２の密度が増加すると、１つ以上のネットワークトラフィックウィンドウ２１０のサイズは減少してもよい。

【００２４】

いくつかの実現例では、複数のネットワークトラフィックウィンドウ２１０の各々について、不正検出部３００は、ネットワークトラフィックウィンドウ２１０のネットワークトラフィックメッセージ２２のサブセットから、サンプリングされたネットワークトラフィックメッセージ２２のセット２２Ｓａ～ｎをサンプリングする。サンプリングされたネットワークトラフィックメッセージのセット２２Ｓは、ネットワークトラフィックウィンドウ２１０のネットワークトラフィックメッセージ２２のサブセットの全体を表す。すな

わち、いくつかのシナリオでは、ネットワークトラフィックメッセージ２２の量は、あらゆるネットワークトラフィックメッセージ２２を分析または処理することを実行不可能または望ましくなくし、代わりに、不正検出部３００は、サンプリングされたネットワークトラフィックメッセージの代表的なセット２２Ｓをサンプリングして、ネットワークトラフィックウィンドウ２１０のネットワークトラフィックメッセージ２２のサブセット全体の任意の特性または傾向を同時に保持しながら、分析すべきネットワークトラフィックメッセージ２２の総数を減少させる。不正検出部３００は、サンプリングされたネットワークトラフィックメッセージのセット２２Ｓを取得するために、任意の従来のサンプリング技術を使用してもよい。

【００２５】

任意選択で、不正検出部３００は、複数のネットワークトラフィックウィンドウ２１０の各々について、サンプリングされたネットワークトラフィックメッセージのセット２２Ｓの特性をデータ構造に記憶する。すなわち、各ネットワークトラフィックウィンドウ２１０についてサンプリングされたネットワークトラフィックメッセージ２２Ｓの各々に関連付けられる情報の全体を記憶する代わりに、不正検出部３００は、サンプリングされたネットワークトラフィックメッセージのセット２２Ｓの記憶要件を低減するために、サンプリングされたネットワークトラフィックメッセージのセット２２Ｓから、特定の特性（例えば、帯域幅、サイズ、ソースアドレス、宛先アドレス、領域など）を抽出してもよい。たとえば、不正検出部３００は、サンプリングされたネットワークトラフィックメッセージ２２Ｓのヘッダフィールドのいくつかもしくはすべて、またはサンプリングされたネットワークトラフィックメッセージ２２Ｓの１つ以上のヘッダフィールドの値の傾向もしくは集約された統計を記憶してもよい。任意選択で、集約された統計は、確率的データ構造（例えば、カウント－最小スケッチ）を使用する。

【００２６】

再び図１を参照すると、不正検出部３００は、検出モデル３１０を介して、ネットワークトラフィックメッセージ２２の第１のセットに基づいてネットワーク不正使用が発生しているかどうかを判断する。すなわち、以下でより詳細に論じるように、不正検出部３００は、検出モデル３１０を用いてネットワークトラフィックメッセージ２２を処理することに基づいて、ネットワーク不正使用が現在発生しているかどうかを判断する。検出モデル３１０は、ネットワークトラフィックメッセージ２２の統計またはニューラルネットワークに基づくアルゴリズムを含む任意のタイプのモデルを組み込むことができる。いくつかの例では、不正検出部３００による不正判定３０２は、単一のネットワークトラフィックウィンドウ２１０に依存する。他の例では、不正判定３０２は、複数のネットワークトラフィックウィンドウ２１０に依存する。すなわち、これらの例では、不正検出部は、現在のネットワークトラフィックウィンドウ２１０においてネットワーク不正使用が発生しているかどうかを判断しながらいくつかの数の以前のネットワークトラフィックウィンドウ２１０の局面を保持するのに充分なメモリを含む。たとえば、不正検出部３００は、現在のネットワークトラフィックウィンドウ２１０および以前の３つのネットワークトラフィックウィンドウ２１０に基づいて、ネットワーク不正使用が現在発生しているかどうかを判断してもよい。

【００２７】

いくつかの実現例では、不正検出部３００は、検出モデル３１０にデータを提供する前に、ネットワークトラフィックメッセージ２２の第１のセットを時系列のセットに変換する。例えば、不正検出部３００は、観測値の相対頻度を求めるために、ネットワークトラフィックメッセージ２２の１つ以上の特性の分布（例えば、各Ｌ７ヘッダフィールドについての分布）を求める。不正検出部３００は、時系列のセットを生成するために、例えば、ジェンセン・シャノン情報量を使用して、異なるネットワークトラフィックウィンドウ２１０におけるこれらの分布に基づく差を求めてもよい。検出モデル３１０は、これらの時系列において、（例えば、現在のメッセージ／秒またはネットワークトラフィックメッセージのヘッダ内の様々なフィールドのような）ネットワークトラフィックメッセージの

様々な特性および傾向を表すトラフィック異常を検出するようにトレーニングされる。

【００２８】

いくつかの例では、検出モデル３１０は、不正確率スコア３１２を生成する。不正検出部３００が、不正確率スコアが不正確率閾値を満たす、と判断すると、不正検出部３００は、検出モデル３１０が現在のネットワーク不正使用を検出したと判断し、不正判定３０２をメッセージラベル付け部１７０に送信する。

【００２９】

ここで図３を参照すると、いくつかの例では、検出モデル３１０は、トレーニングネットワークトラフィックサンプル２２Ｔのセットでトレーニングされたニューラルネットワーク（例えば、ロングショートタームメモリ（ＬＴＳＭ）ニューラルネットワーク）を含む。いくつかの実現例では、各トレーニングネットワークトラフィックサンプル２２Ｔはラベル付けされる。すなわち、各トレーニングネットワークトラフィックサンプル２２Ｔは、トレーニングネットワークトラフィックサンプル２２Ｔが、検出モデル３１０が不正ネットワークトラフィックサンプルまたは正当なネットワークトラフィックサンプルとして予測すべきサンプルであるかどうかを示すラベルを含む。いくつかの例では、トレーニングネットワークトラフィックサンプル２２Ｔのサブセットは、正当なネットワークトラフィックサンプルを示す陽性ラベルを含み、トレーニングネットワークトラフィックサンプル２２Ｔの残りのサブセットは、不正なネットワークトラフィックサンプルを示す陰性ラベルを含む。検出モデル３１０は、ラベル付けされたトレーニングネットワークトラフィックサンプル２２Ｔの予測を行った後、検出モデル３１０は、その予測を（例えば損失公式を介して）ラベルと比較し、それに応じてモデル３１０のパラメータを（例えば、勾配降下を介して）調整して、適切なラベルを予測するよう学習する。トレーニングネットワークトラフィックサンプル２２Ｔは、シミュレートされたワークトラフィックサンプルまたは履歴ネットワークトラフィックサンプルであってもよい。履歴ネットワークトラフィックサンプルは、検出モデルが監視するようにトレーニングされたネットワークサービス３０によって以前に受信されたネットワークトラフィックメッセージ、および／または他のネットワークサービスによって以前に受信されたネットワークトラフィックメッセージであってもよい。

【００３０】

いくつかの実現例では、ネットワーク不正使用軽減部１６０は、不正検出部３００が不正判定３０２を生成した（すなわち、不正検出部３００が、ネットワーク不正使用が現在発生していると判断した）後、不正検出部３００がその判定を行ったときにネットワーク不正使用が実際に発生していたかどうかを示すフィードバック３２０を受信する。すなわち、フィードバック３２０は、不正検出部３００が正しかったかまたは誤っていた（すなわち、偽陽性）かを示す。ネットワーク不正使用軽減部１６０は、ユーザ１２からユーザデバイス１０を介してフィードバック３２０を受信してもよい。例えば、ユーザ１２は、ネットワーク不正使用の通知４０４（図４）を受信した後、ユーザ１２は、ネットワーク不正使用が起こっていたかまたは起こっていなかったことを調査および判断し、それに応じてフィードバック３２０を提供する。遠隔システム１４０も、フィードバック３２０を提供してもよい。たとえば、ネットワーク不正使用が発生していると不正検出部が判断した後、より洗練されたおよび／またはより時間がかかるネットワークトラフィック分析部が、ネットワーク不正使用が発生したかどうかを判断してもよい。

【００３１】

いくつかの例では、検出モデル３１０は、ユーザ１２および検出モデル３１０に関連付けられるネットワークサービス３０（すなわち、検出モデル３１０が監視しているネットワークサービス）によって受信されたネットワークトラフィックメッセージ２２に基づいてのみ更新される。他の例では、検出モデル３１０は、他のネットワークサービス３０を監視する他の検出モデル３１０によって受信されたネットワークトラフィックメッセージ２２に基づいて更新される。例えば、クラウド環境は、多くのネットワークサービス３０を実行し、多くの異なる検出モデル３１０を使用して、これらのネットワークサービス３

０を保護してもよい。各検出モデル３１０は、他の検出モデル３１０を更新するために使用されてもよい異なるネットワークトラフィックメッセージ２２を受信してもよい。いくつかの例では、ユーザ１２は、グローバル検出モデル３１０更新をオプトアウトまたはオプトインしてもよい。

【００３２】

ネットワーク不正使用軽減部１６０は、フィードバック３２０を、判断されたネットワーク不正使用の期間中に取得されたネットワークトラフィックメッセージ２２に適用するラベルとして扱ってもよい。すなわち、ネットワーク不正使用軽減部１６０は、検出モデル３１０の精度を高めるために、フィードバック３２０から生成されたフィードバックサンプル２２Ｆに基づいて、検出モデル３１０を更新または微調整または再トレーニングしてもよい。

【００３３】

再び図１を参照すると、ネットワーク不正使用が発生している間（すなわち、不正検出部３００がメッセージラベル付け部１７０に不正判定３０２を送信した）、メッセージラベル付け部１７０は、ネットワークサービス３０によって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージ２２の第２のセット２２ｂを取得する。いくつかの例では、ネットワークトラフィックメッセージ２２の第２のセットは、ネットワークトラフィックメッセージ２２の第１のセット（すなわち、不正検出部３００が取得したネットワークトラフィックメッセージ２２のセット）と同じである。他の例では、メッセージラベル付け部は、ネットワークトラフィックメッセージ２２の、少なくとも部分的に異なるセットを取得する。たとえば、メッセージラベル付け部１７０は、ネットワークトラフィックメッセージ２２の、より最近の、またはより大きい、またはより小さいセットを取得する。

【００３４】

メッセージラベル付け部１７０はまた、いくつかの例では、履歴ネットワークトラフィックメッセージ２２Ｈのセットを取得する。これらの履歴ネットワークトラフィックメッセージ２２Ｈ（または履歴ネットワークトラフィックメッセージ２２Ｈの特性）は、ネットワーク不正使用に先行する可能性がある１つまたは以前のネットワークトラフィックウィンドウ２１０の間にキャプチャされたネットワークトラフィックメッセージ２２を含む。たとえば、不正検出部３００は、前の期間（例えば、２４時間、４８時間など）のネットワークトラフィックメッセージ２２の第１のセット（またはサンプリングされたネットワークトラフィックメッセージ２２Ｓのセットもしくはサンプリングされたネットワークトラフィックメッセージ２２Ｓのセットの特性）を記憶してもよい。メッセージラベル付け部１７０は、履歴ネットワークトラフィックメッセージ２２Ｈの取得されたセットが主に正当なトラフィックに関連付けられる尤度が高いという前提で、履歴ネットワークトラフィックメッセージ２２Ｈのセットを取得する。すなわち、履歴ネットワークトラフィックメッセージ２２Ｈのセットによって表される期間中にネットワーク不正使用が起こっていなかった可能性が高く、したがって履歴ネットワークトラフィックメッセージ２２Ｈは主に正当なトラフィックを表す。必要ではないが、履歴ネットワークトラフィックメッセージ２２Ｈのセットは、ネットワークサービス３０のためのベースライントラフィックの潜在的なコンテキストをメッセージラベル付け部１７０に提供することによって、メッセージラベル付け部１７０の品質および／または精度を改善してもよい。

【００３５】

ネットワークトラフィックメッセージの第２のセット内の各ネットワークトラフィックメッセージ２２について、メッセージラベル付け部１７０は、ラベル付けモデル１７２を介し、履歴ネットワークトラフィックメッセージ２２Ｈのセットを使用して、各ネットワークトラフィックメッセージを、ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージ２２またはネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージ２２のいずれかとしてラベル付けする。すなわち、メッセージラベル付け部１７０は、正当なトラフィックに主に関連付けられる履歴ネットワーク

トラフィックメッセージ２２Ｈのセットに基づいて、ネットワークトラフィックメッセージ２２の第２のセット中の各ネットワークトラフィックメッセージ２２を、正当なトラフィックまたは非正当なトラフィック（すなわち、不正トラフィック）のいずれかとしてラベル付けする。したがって、メッセージラベル付け部１７０は、ネットワークトラフィックメッセージ２２の第２のセットのネットワークトラフィックメッセージ２２の各々についてのラベル（すなわち、不正または正当）を含む、ラベル付けされたネットワークトラフィックメッセージ２２Ｌのセットを生成する。

【００３６】

いくつかの実現例では、ラベル付けモデル１７２は、履歴ネットワークトラフィックメッセージのセット（またはヘッダフィールドの分布などのメッセージの特性）を使用して、ネットワークサービス３０のためのベースライントラフィックまたは正常トラフィックを決定する。ラベル付けモデル１７２は、履歴ネットワークトラフィックメッセージ２２Ｈのセット内において特定の値を見る確率と、ネットワークトラフィックメッセージ２２の第２のセット（すなわち、現在のメッセージ）内に同じ特定の値を見る確率とに基づいて、点を多次元空間（例えば、ネットワーク不正使用軽減部１６０が監視する各ヘッダフィールドにつき１次元）にマッピングすることによって、正常トラフィックを不正トラフィックから分離してもよい。ラベル付けモデル１７２は、多次元空間内の点をクラスタ化して、正常トラフィックと不正トラフィックとの分離を可能にしてもよい。いくつかの実現例では、ラベル付けモデル１７２は、正常トラフィックのクラスタおよび不正トラフィックのクラスタについてのベースライン分布から、各ヘッダフィールドについてのジェンセン・シャノン情報量の合計を測定する。ラベル付けモデル１７２は、より高いジェンセン・シャノン情報量を有するクラスタは不正トラフィックのクラスタである、と判断してもよい。

【００３７】

いくつかの例では、ネットワーク不正使用軽減部１６０は、検出モデル３１０または不正検出部３００を含まず、メッセージラベル付け部１７０は、不正判定３０２を受信しない。このシナリオでは、メッセージラベル付け部１７０は、充分な数のネットワークトラフィックメッセージ２２が不正としてラベル付けされるときにネットワーク不正使用が発生していると判断することによって、検出モデル３１０として機能してもよい。たとえば、ある総数またはパーセンテージのネットワークトラフィックメッセージ２２がしきい値を満たすとき、メッセージラベル付け部１７０は、ネットワーク不正使用が発生していると判断してもよい。

【００３８】

メッセージラベル付け部１７０は、ラベル付けされたネットワークトラフィックメッセージ２２Ｌのセットをルール生成部４００に送信する。ルール生成部４００は、ルールモデル４１０を介して、少なくとも１つのネットワークトラフィックルール４０２を生成し、これは、実現されると、不正ネットワークトラフィックメッセージ２２の影響を低減する。すなわち、いくつかの例では、ルール生成部４００は、ラベル付けされたネットワークトラフィックメッセージ２２Ｌを受信し、ラベル付けされたネットワークトラフィックメッセージ２２Ｌに基づいてネットワーク不正使用を軽減するネットワークトラフィックルール４０２を導出する。たとえば、ネットワークトラフィックルールは、不正としてラベル付けされた、ラベル付けされたネットワークトラフィックメッセージ２２Ｌの一部を遮断または破棄する。ネットワークトラフィックルールは、任意に、ファイアウォールにおいて実現される。特定の例において、ルール生成部４００は、不正なラベル付けされたネットワークトラフィックメッセージ２２Ｌの大部分が特定の地域から発生していると判断してもよい。このシナリオでは、ルール生成部４００は、その特定の地域からのネットワークトラフィックメッセージを遮断するネットワークトラフィックルール４０２を生成してもよい。ネットワークトラフィックルール４０２は、不正トラフィックが最も共通する特性に基づいて、ネットワークトラフィックメッセージ２２の任意の数の特性（たとえばヘッダフィールド）に基づいて、ネットワークトラフィックメッセージ２２を遮断また

は迂回させてもよい。

【００３９】

ここで図４を参照すると、ルール生成部４００は、適用可能なネットワークトラフィックルール４０２であるルールを生成するルール生成アルゴリズムであるルールモデル４１０を含む。概して、ルール生成部４００は、最大量の不正なネットワークトラフィックを軽減し、最小量の正当なトラフィックに影響を及ぼすルールの生成を試みる。いくつかの例では、これらの目標は競合している。すなわち、ネットワーク不正使用のいくつかの例では、遮断される不正トラフィックがより多いほど、同様に遮断される正当なトラフィックはより多くなる。この目的のため、ルール生成部４００は、いくつかの実現例では、複数のネットワークトラフィックルール４０２、４０２ａ～ｎを生成する。ネットワーク不正使用軽減部１６０は、進行中のネットワーク不正使用を示す通知４０４とともに、ユーザデバイス１０を介してネットワークサービス３０に関連付けられるユーザ１２に複数のネットワークトラフィックルール４０２を送信してもよい。いくつかの例では、ネットワーク不正使用軽減部１６０は、通知４０４だけをユーザ１２に送信し、何らかのアクション（たとえば、１つ以上のネットワークトラフィックルール４０２を生成すること）を行う前にユーザ１２からの応答を待つ。

【００４０】

ユーザ１２は、各提案されたネットワークトラフィックルール４０２が現在のネットワークトラフィックに及ぼす影響を評価し、ユーザの目標または希望に最も受け入れ可能なネットワークトラフィックルール４０２を選択してもよい。ネットワーク不正使用軽減部１６０は、提案されたネットワークトラフィックルール４０２の各々の効果をユーザ１２に明確に示すモデル、報告、および／または統計を含んでもよい。たとえば、第１のネットワークトラフィックルール４０２は、不正なトラフィックの９５パーセントを遮断するが、正当なトラフィックの１０パーセントも遮断してもよく、一方、第２のネットワークトラフィックルール４０２は、不正なトラフィックの８０パーセントのみを遮断するが、正当なトラフィックの１パーセントのみを遮断してもよい。

【００４１】

ユーザ１２は、提案される生成されたネットワークトラフィックルール４０２のうちの１つの受け入れを示す指示４０６を提供して、ネットワーク不正使用軽減部１６０に戻してもよい。ルール実現部（例えば、ファイアウォール）は、選択されたネットワークトラフィックルール４０２を実現してもよい。いくつかの例では、ネットワーク不正使用軽減部１６０は、ユーザ介入なしに、ネットワークトラフィックルール４０２のうちの１つを自動的に実現する。たとえば、ルール生成部４００は、ネットワークトラフィックルール４０２を、様々な要素（例えば、遮断された不正トラフィックの量、遮断された正当なトラフィックの量、サービスの劣化など）から求められる最高スコアとともに与えてもよく、ルール実現部４２０は、生成されたネットワークトラフィックルール４０２を直ちに実現してもよい。いくつかの例では、ユーザ１２は、後の時点で、ネットワークトラフィックルール４０２を無効にするか、または異なるネットワークトラフィックルール４０２を選択することによって、ネットワークトラフィックルール４０２をオーバーライドする。

【００４２】

ここで図５を参照すると、任意選択で、ネットワーク不正使用軽減部１６０は、ユーザ１２から（例えば、ユーザデバイス１０を介して）、所望のネットワークトラフィックルール４０２または所望のネットワークトラフィックルールパラメータに関連付けられるユーザプリファレンス５１０を受信する。すなわち、ネットワーク不正使用軽減部１６０は、ユーザプリファレンス５１０に基づいて、ネットワークトラフィックルール４０２を生成し、および／またはどのネットワークトラフィックルール４０２を実現するかを選択してもよい。ユーザプリファレンスは、付帯的損害感受性等のパラメータを含んでもよい。付帯的損害とは、ネットワークトラフィックルールによって影響を受ける非不正トラフィックまたは正当なトラフィックの量を指す。したがって、付帯的損害に対して高い感受性を有するユーザ１２は、正当なトラフィックを遮断することに対する強い嫌悪を有するが

、付帯的損害に対して低い感受性を有するユーザ１２は、正当なトラフィックを遮断することに対する強い嫌悪を有さない。この感度は、ユーザ１２の選好に加えて、ネットワークサービス３０のタイプおよび性質に大きく依存してもよい。

【００４３】

他のユーザプリファレンス５１０は、リソースプロビジョニング（すなわち、いくつのリソースがネットワークサービス３０および／またはユーザ１２に割り振られるか、およびいくつのリソースがネットワークサービス３０に割り振られてもよいか）、コスト感度（すなわち、ユーザ１２がネットワークサービス３０および／またはそのインフラストラクチャによって消費されるリソースに対して支払いを行うとき）、ならびにインパクト感度（すなわち、ユーザ１２がネットワークサービス３０へのアクセスの低下または喪失にどの程度敏感であるか）を含む。ユーザプリファレンス５１０に基づいて、ネットワーク不正使用軽減部１６０は、ルール生成部４００によって生成された複数のネットワークトラフィックルール４０２のうちの１つを選択してもよく、ルール実現部４２０は、人的介入なしに、選択されたネットワークトラフィックルール４０２を自動的に実現してもよい。

【００４４】

図６は、サービス拒否攻撃などのネットワーク不正使用を自動的に検出および軽減する方法６００のための動作の例示的な構成のフローチャートである。方法６００は、動作６０２において、ネットワークサービス３０によって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージ２２の第１のセットを取得することを含む。動作６０４において、方法６００は、第１のモデル３１０を介して、ネットワークトラフィックメッセージ２２の第１のセットに基づいてネットワーク不正使用が発生しているかどうかを判断することを含む。ネットワーク不正使用が発生しているとき、方法６００は、動作６０６において、ネットワークサービス３０によって現在受信されているネットワークトラフィックを表すネットワークトラフィックメッセージ２２の第２のセットを取得することを含む。

【００４５】

動作６０８において、方法６００は、ネットワークトラフィックメッセージ２２の第２のセット中の各ネットワークトラフィックメッセージ２２について、第２のモデル１７０を介して、ネットワークトラフィックメッセージ２２を、ネットワーク不正使用に関与している不正ネットワークトラフィックメッセージ２２またはネットワーク不正使用に関与していない非不正ネットワークトラフィックメッセージ２２としてラベル付けすることを含む。動作６１０において、方法６００は、第３のモデル４１０を介して、少なくとも１つのネットワークトラフィックルール４０２を生成することを含む。各ネットワークトラフィックルール４０２は、ファイアウォールによって実現されるよう構成され、実現されると、不正ネットワークトラフィックメッセージ２２の影響を低減するよう構成される。

【００４６】

図７は、本文書で説明されるシステムおよび方法を実現するために用いられ得る例示的なコンピューティングデバイス７００の概略図である。コンピューティングデバイス７００は、ラップトップ、デスクトップ、ワークステーション、携帯情報端末、サーバ、ブレードサーバ、メインフレーム、および他の適切なコンピュータなど、様々な形態のデジタルコンピュータを表すことが意図されている。本明細書に示された構成要素、それらの接続および関係、ならびにそれらの機能は、例示的なものにすぎず、本文書に記載および／または特許請求される本発明の実現例を限定するものではない。

【００４７】

コンピューティングデバイス７００は、プロセッサ７１０と、メモリ７２０と、ストレージデバイス７３０と、メモリ７２０および高速拡張ポート７５０に接続する高速インターフェイス／コントローラ７４０と、低速バス７７０およびストレージデバイス７３０に接続する低速インターフェイス／コントローラ７６０とを含む。構成要素７１０、７２０、７３０、７４０、７５０および７６０の各々は、様々なバスを用いて相互接続され、共通のマザーボード上に、または必要に応じて他の方法で実装されてもよい。プロセッサ７

１０は、高速インターフェイス７４０に結合されたディスプレイ７８０などの外部入力／出力装置上にグラフィカルユーザインターフェイス（ＧＵＩ）のためのグラフィカル情報を表示するために、メモリ７２０またはストレージデバイス７３０に記憶された命令を含む、コンピューティングデバイス７００内で実行するための命令を処理することができる。他の実現例では、複数のプロセッサおよび／または複数のバスが、必要に応じて、複数のメモリおよびメモリのタイプとともに用いられてもよい。また、複数のコンピューティングデバイス７００が接続されてもよく、各デバイスは、（たとえば、サーババンクとして、ブレードサーバのグループとして、またはマルチプロセッサシステムとして）必要な動作の部分を提供する。

【００４８】

メモリ７２０は、コンピューティングデバイス７００内で情報を非一時的に記憶する。メモリ７２０は、コンピュータ可読媒体、揮発性メモリユニット、または不揮発性メモリユニットであってもよい。非一時的メモリ７２０は、コンピューティングデバイス７００による使用のためにプログラム（例えば、命令のシーケンス）またはデータ（例えば、プログラム状態情報）を一時的または永続的に記憶するために用いられる物理デバイスであってもよい。不揮発性メモリの例は、フラッシュメモリおよび読み出し専用メモリ（ＲＯＭ）／プログラマブル読み出し専用メモリ（ＰＲＯＭ）／消去可能プログラマブル読み出し専用メモリ（ＥＰＲＯＭ）／電子的消去可能プログラマブル読み出し専用メモリ（ＥＥＰＲＯＭ）（たとえば、ブートプログラムなどのファームウェアに典型的に用いられる）を含むが、これらに限定されない。揮発性メモリの例には、ランダムアクセスメモリ（ＲＡＭ）、ダイナミックランダムアクセスメモリ（ＤＲＡＭ）、スタティックランダムアクセスメモリ（ＳＲＡＭ）、相変化メモリ（ＰＣＭ）、およびディスクまたはテープが含まれるが、これらに限定されない。

【００４９】

ストレージデバイス７３０は、コンピューティングデバイス７００のための大容量ストレージを提供することができる。いくつかの実現例では、ストレージデバイス７３０はコンピュータ可読媒体である。様々な異なる実現例では、ストレージデバイス７３０は、フロッピー（登録商標）ディスクデバイス、ハードディスクデバイス、光ディスクデバイス、もしくはテープデバイス、フラッシュメモリもしくは他の同様のソリッドステートメモリデバイス、またはストレージエリアネットワークもしくは他の構成におけるデバイスを含むデバイスのアレイであってもよい。さらなる実現例では、コンピュータプログラム製品は、情報担体において有形に具現化される。コンピュータプログラム製品は、実行されると上述の方法などの１つ以上の方法を実行する命令を含む。情報担体は、メモリ７２０、ストレージデバイス７３０、またはプロセッサ７１０上のメモリなどのコンピュータ可読媒体または機械可読媒体である。

【００５０】

高速コントローラ７４０は、コンピューティングデバイス７００のための帯域幅集約型動作を管理し、低速コントローラ７６０は、より低い帯域幅集約型動作を管理する。そのような役割の割り当ては、例示的なものにすぎない。いくつかの実現例では、高速コントローラ７４０は、メモリ７２０、ディスプレイ７８０（たとえば、グラフィックスプロセッサまたはアクセラレータを通して）、および様々な拡張カード（図示せず）を受け入れ得る高速拡張ポート７５０に結合される。いくつかの実現例では、低速コントローラ７６０は、ストレージデバイス７３０および低速拡張ポート７９０に結合される。低速拡張ポート７９０は、様々な通信ポート（たとえば、ＵＳＢ、Bluetooth（登録商標）、イーサネット（登録商標）、無線イーサネット（登録商標））を含んでもよく、キーボード、ポインティングデバイス、スキャナ、またはスイッチもしくはルータなどのネットワーキングデバイスなどの１つ以上の入力／出力装置に、たとえばネットワークアダプタを介して結合されてもよい。

【００５１】

コンピューティングデバイス７００は、図に示されるように、いくつかの異なる形態で

実現されてもよい。例えば、標準サーバ７００ａとして、もしくはそのようなサーバ７００ａのグループ内で複数回、ラップトップコンピュータ７００ｂとして、またはラックサーバシステム７００ｃの一部として実現されてもよい。

【００５２】

本明細書に記載のシステムおよび技術のさまざまな実現例は、デジタル電子および／もしくは光学回路系、集積回路系、特別に設計されたＡＳＩＣ（特定用途向け集積回路）、コンピュータハードウェア、ファームウェア、ソフトウェア、ならびに／またはそれらの組合せで実現されてもよい。これらのさまざまな実現例は、少なくとも１つのプログラマブルプロセッサを含むプログラマブルシステム上で実行可能および／または解釈可能な１つ以上のコンピュータプログラムにおける実現例を含んでいてもよく、当該プロセッサは専用であっても汎用であってもよく、ストレージシステム、少なくとも１つの入力装置、および少なくとも１つの出力装置からデータおよび命令を受信するように、かつこれらにデータおよび命令を送信するように結合されている。

【００５３】

ソフトウェアアプリケーション（すなわち、ソフトウェアリソース）は、コンピューティングデバイスにタスクを実行させるコンピュータソフトウェアを指してもよい。いくつかの例では、ソフトウェアアプリケーションは、「アプリケーション」、「アプリ」、または「プログラム」と呼ばれることがある。アプリケーションの例には、システム診断アプリケーション、システム管理アプリケーション、システム保守アプリケーション、ワード処理アプリケーション、スプレッドシートアプリケーション、メッセージングアプリケーション、メディアストリーミングアプリケーション、ソーシャルネットワーキングアプリケーション、およびゲームアプリケーションが含まれるが、これらに限定はされない。

【００５４】

これらのコンピュータプログラム（プログラム、ソフトウェア、ソフトウェアアプリケーションまたはコードとしても知られる）は、プログラム可能なプロセッサのための機械命令を含み、高水準手続き型および／もしくはオブジェクト指向型プログラミング言語で、ならびに／またはアセンブリ／機械言語で実現することができる。本明細書で使用されるとき、用語「機械可読媒体」および「コンピュータ可読媒体」は、機械命令を機械可読信号として受信する機械可読媒体を含む、機械命令および／またはデータをプログラマブルプロセッサに提供するよう使用される任意のコンピュータプログラム製品、非一時的コンピュータ可読媒体、装置および／またはデバイス（例えば、磁気ディスク、光ディスク、メモリ、プログラマブルロジックデバイス（ＰＬＤ））を指す。「機械可読信号」という用語は、機械命令および／またはデータをプログラマブルプロセッサに提供するために使用される任意の信号を指す。

【００５５】

本明細書で説明するプロセスおよび論理フローは、データ処理ハードウェアとも呼ばれ、入力データに対して演算し出力を生成することによって１つ以上のコンピュータプログラムを実行して機能を実行する１つ以上のプログラマブルプロセッサによって実行されることができる。プロセスおよび論理フローはまた、専用論理回路、たとえば、ＦＰＧＡ（フィールドプログラマブルゲートアレイ）またはＡＳＩＣ（特定用途向け集積回路）によって実行され得る。コンピュータプログラムの実行に好適であるプロセッサは、例として、汎用マイクロプロセッサおよび特殊目的マイクロプロセッサの両方、ならびに任意の種類のデジタルコンピュータの任意の１つ以上のプロセッサを含む。概して、プロセッサは、読み取り専用メモリもしくはランダムアクセスメモリまたは両方から命令およびデータを受信することになる。コンピュータの必須要素は、命令を実行するためのプロセッサ、ならびに命令およびデータを記憶するための１つ以上のメモリデバイスである。一般に、コンピュータはさらに、たとえば磁気ディスク、光磁気ディスクまたは光ディスクといった、データを格納するための１つ以上の大容量記憶装置を含むか、当該１つ以上の大容量記憶装置からデータを受取るかもしくは当該１つ以上の大容量記憶装置にデータを転送するよう作動的に結合されるか、またはその両方を行うことにもなる。しかしながら、コンピ

ュータは、そのようなデバイスを有する必要はない。コンピュータプログラム命令およびデータを記憶するのに好適なコンピュータ可読媒体は、例として、半導体メモリデバイス、たとえば、ＥＰＲＯＭ、ＥＥＰＲＯＭ、およびフラッシュメモリデバイス；磁気ディスク、たとえば内蔵ハードディスクまたはリムーバブルディスク；光磁気ディスク；およびＣＤ　ＲＯＭおよびＤＶＤ－ＲＯＭディスクを含む、あらゆる形態の不揮発性メモリ、媒体、ならびにメモリデバイスを含む。プロセッサおよびメモリは、特殊目的論理回路によって補足され得るか、または特殊目的論理回路に組み込まれ得る。

【００５６】

ユーザとの対話を提供するために、本開示の１つ以上の局面は、たとえばＣＲＴ（陰極線管）、ＬＣＤ（液晶ディスプレイ）モニタまたはタッチスクリーンといったユーザに対して情報を表示するための表示装置と、選択肢的にキーボードおよびたとえばマウス、トラックボールといったユーザがコンピュータに入力を提供可能であるポインティングデバイスとを有するコンピュータ上で実現され得る。他の種類のデバイスを用いて、ユーザとの対話を提供することもでき、たとえば、ユーザに提供されるフィードバックは、任意の形態の感覚フィードバック、たとえば、視覚フィードバック、聴覚フィードバック、または触覚フィードバックであり得、ユーザからの入力は、音響入力、音声入力、または触覚入力を含む、任意の形態で受信することができる。加えて、コンピュータは、ユーザが用いるデバイスにドキュメントを送信し、ユーザが用いるデバイスからドキュメントを受信することによって、たとえば、ユーザのクライアントデバイス上のウェブブラウザから受信された要求に応答してそのウェブブラウザにウェブページを送信することによって、ユーザと対話し得る。

【００５７】

いくつかの実現例について説明した。それにもかかわらず、本開示の精神および範囲から逸脱することなく、様々な修正がなされ得ることが理解されるであろう。したがって、他の実現例は特許請求の範囲内にある。

【図面】

【図１】

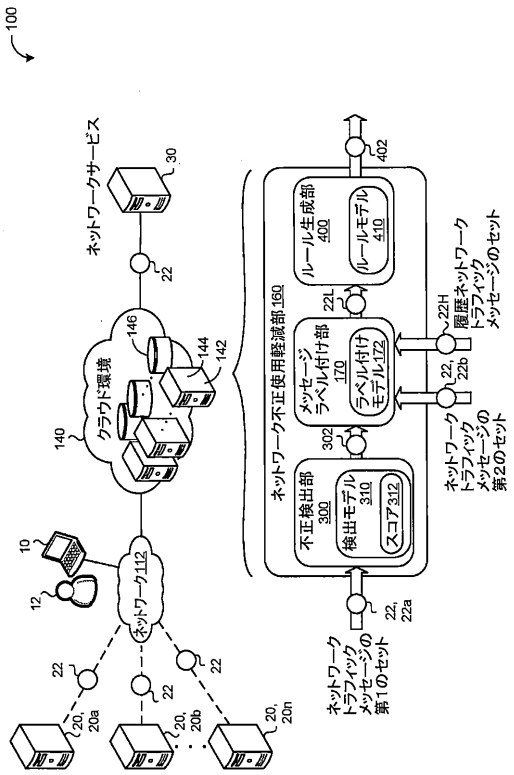

FIG. 1

【図２】

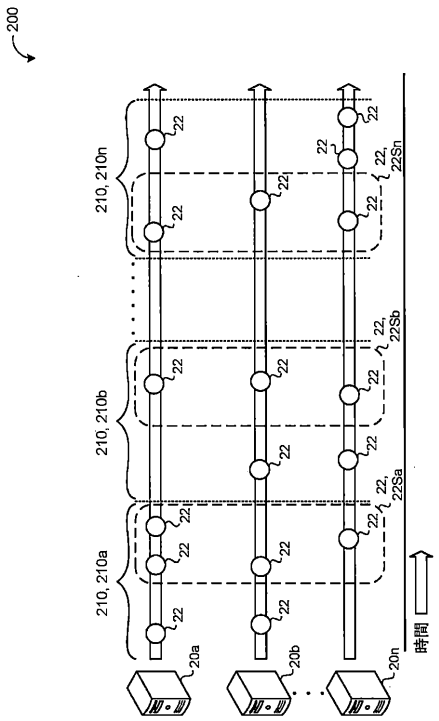

FIG. 2

【図３】

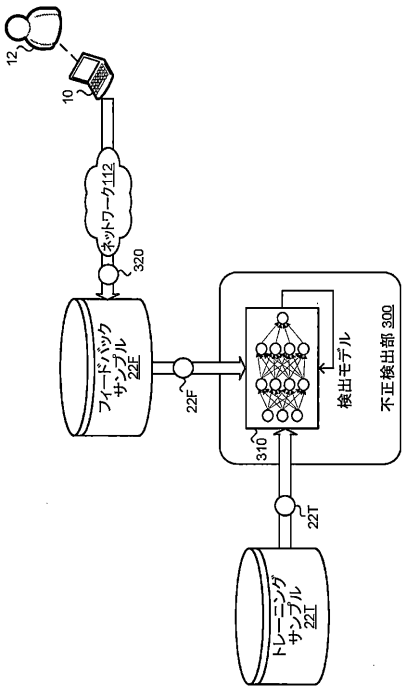

FIG. 3

【図４】

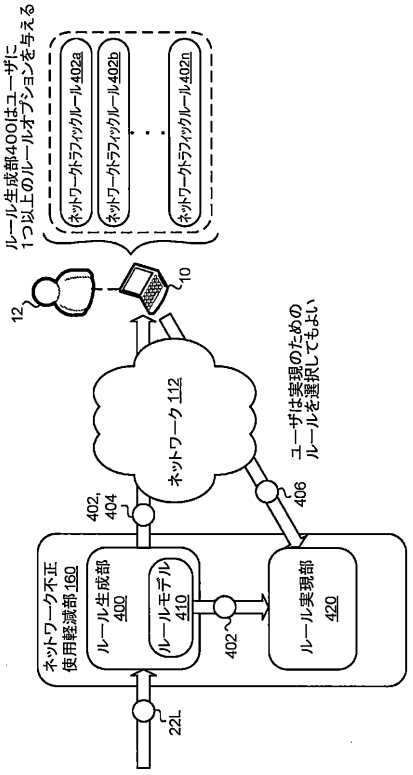

FIG. 4

【図５】

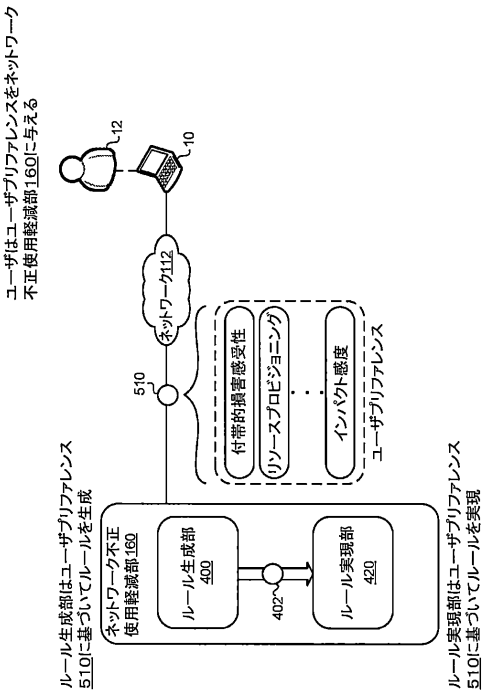

FIG. 5

【図６】

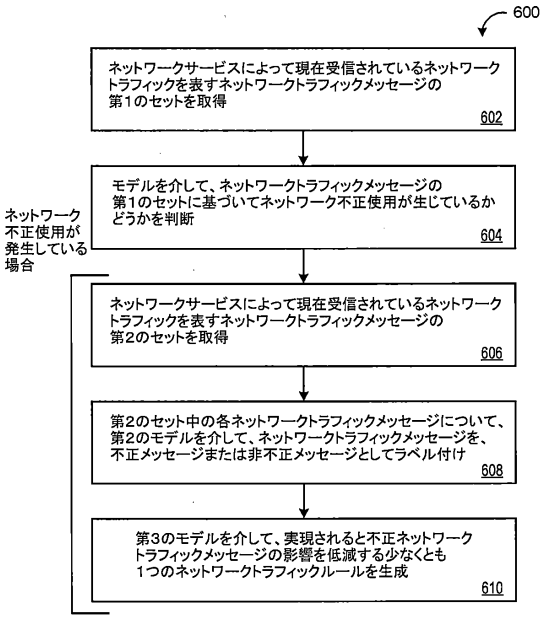

FIG. 6

【図７】

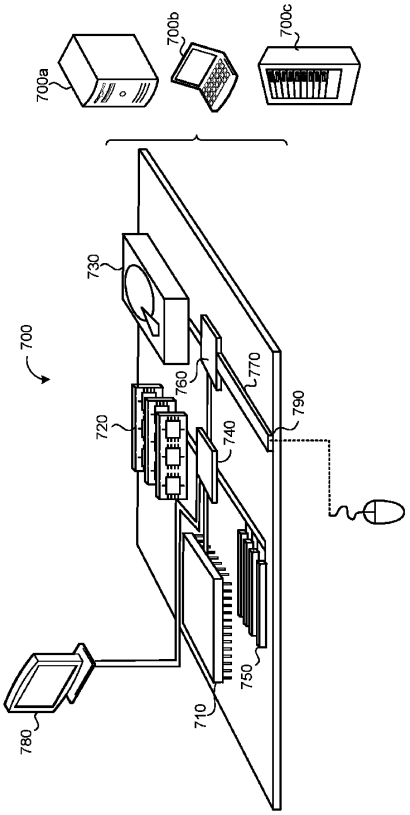

FIG. 7

フロントページの続き

(72)発明者　ハルダー，アンドレ・ロイド・ペルレ
アメリカ合衆国、９４０４３　カリフォルニア州、マウンテン・ビュー、アンフィシアター・パークウェイ、１６００
(72)発明者　ジョシ，プラジャクタ
アメリカ合衆国、９４０４３　カリフォルニア州、マウンテン・ビュー、アンフィシアター・パークウェイ、１６００
(72)発明者　ロイ，アミタバ
アメリカ合衆国、９４０４３　カリフォルニア州、マウンテン・ビュー、アンフィシアター・パークウェイ、１６００
(72)発明者　タラガデーディービ，サイラ
アメリカ合衆国、９４０４３　カリフォルニア州、マウンテン・ビュー、アンフィシアター・パークウェイ、１６００
(72)発明者　カイナー，エミール
アメリカ合衆国、９４０４３　カリフォルニア州、マウンテン・ビュー、アンフィシアター・パークウェイ、１６００
(72)発明者　クオ，チアータン
アメリカ合衆国、９４０４３　カリフォルニア州、マウンテン・ビュー、アンフィシアター・パークウェイ、１６００
(72)発明者　イェ，ジアユ
アメリカ合衆国、９４０４３　カリフォルニア州、マウンテン・ビュー、アンフィシアター・パークウェイ、１６００

審査官　宮島 郁美
(56)参考文献　国際公開第２０２０／２２２８７２（WO，Ａ１）
特開２００５－０３９７２１（ＪＰ，Ａ）
米国特許出願公開第２０１６／０３５２７６５（ＵＳ，Ａ１）
(58)調査した分野　(Int.Cl.，ＤＢ名)
Ｈ０４Ｌ１２／００－１２／６６，１３／００，４１／００－６９／４０