

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7644140号
(P7644140)

(45)発行日 令和7年3月11日(2025.3.11)

(24)登録日 令和7年3月3日(2025.3.3)

(51)国際特許分類

F I

G 0 6 F 18/2433(2023.01)

G 0 6 F 18/2433

G 0 6 N 20/00 (2019.01)

G 0 6 N 20/00

請求項の数 11 (全42頁)

(21)出願番号	特願2022-557900(P2022-557900)	(73)特許権者	502303739
(86)(22)出願日	令和3年1月29日(2021.1.29)		オラクル・インターナショナル・コーポ
(65)公表番号	特表2023-518866(P2023-518866		レイション
	A)		アメリカ合衆国 9 4 0 6 5 カリフォル
(43)公表日	令和5年5月8日(2023.5.8)		ニア州 レッドウッド ショアーズ, メー
(86)国際出願番号	PCT/US2021/015802		ル ストップ 5 オーピー7 オラクル パ
(87)国際公開番号	WO2021/194631		ークウェイ 5 0 0
(87)国際公開日	令和3年9月30日(2021.9.30)	(74)代理人	110001195
審査請求日	令和5年12月11日(2023.12.11)		弁理士法人深見特許事務所
(31)優先権主張番号	16/826,478	(72)発明者	ウェザービー, エドワード・アール
(32)優先日	令和2年3月23日(2020.3.23)		アメリカ合衆国、6 8 1 1 6 ネブラス
(33)優先権主張国・地域又は機関	米国(US)		カ州、オマハ、バーデット・ストリート
			、1 5 8 1 9
		(72)発明者	バクラフスキー, ケネス・ピィ
			アメリカ合衆国、0 2 4 5 3 マサチュ
			最終頁に続く

(54)【発明の名称】 機械学習モデルの結果を確実に監査できるようにするためのシステムおよび方法

(57)【特許請求の範囲】

【請求項1】

機械学習モデルの結果を監査するための、コンピュータによって実現される方法であって、

監査下のデータベースから元の時系列データ値についての状態推定値のセットを取出すステップを含み、前記状態推定値のセットは、監査下の前記データベースへの変更を記述するジャーナルエントリに関連付けられ、前記状態推定値の各々は、前記時系列データ値のうちの1つについての状態推定計算によって生成され、前記方法はさらに、

前記状態推定値の各々ごとに前記状態推定計算を反転させて、前記状態推定値の各々ごとに、再構成された時系列データ値を生成するステップと、

監査下の前記データベースから前記元の時系列データ値を取出すステップと、

前記元の時系列データ値を前記再構成された時系列データ値と対ごとに比較して、前記元の時系列と前記再構成された時系列とが一致するかどうかを判断するステップと、

監査下の前記データベースが、(i)前記元の時系列と前記再構成された時系列とが一致する場合に修正されておらず、かつ(ii)前記元の時系列と前記再構成された時系列とが一致しない場合に修正されていたという信号を生成するステップとを含む、方法。

【請求項2】

前記元の時系列値から選択されたトレーニング値のセットで、多変量状態推定モデルをトレーニングするステップと、

前記多変量状態推定モデルに関連付けられた行列を固有ベクトルのセットに分解するス

テップと、

前記固有ベクトルのセットから主固有ベクトルのサブセットを選択するステップと、

前記主固有ベクトルのサブセットから、制限された多変量状態推定モデルを作成するステップとをさらに含み、

前記状態推定計算の反転は、前記制限された多変量状態推定モデルが状態推定値を形成する計算の反転を生成するステップをさらに含む、請求項 1 に記載の方法。

【請求項 3】

監査下の前記データベースからの前記元の時系列データ値からトレーニングデータ値を除外するステップと、

残りの元の時系列データ値を前処理してセンサ障害の影響を軽減し、前記残りの元の時系列データに対する変更を記録するモデルを生成するステップと、

前記制限された多変量状態推定モデルを用いて、前記前処理された残りの元の時系列データ値について状態推定を実行して、圧縮された時系列データベースを作成するステップとをさらに含む、請求項 2 に記載の方法。

【請求項 4】

前記元の時系列値を前処理して、前記元の時系列値の品質に対するセンサ障害の影響を軽減するステップと、

前記前処理中に前記元の時系列値に対する 1 つ以上の変更を記録するデータ前処理モデルを生成するステップとをさらに含み、

前記状態推定計算の反転はさらに、

前記データ前処理モデルを取出すステップと、

前記状態推定値の各々ごとに前記データ前処理モデルによって記述される前記前処理を反転させるステップとを含む、請求項 1 ～ 3 のいずれか 1 項に記載の方法。

【請求項 5】

機械学習モデルの結果を監査するための、コンピュータによって実現される方法であって、

監査下のデータベースから元の時系列データ値についての状態推定値のセットを取出すステップを含み、前記状態推定値の各々は、前記時系列データ値のうちの 1 つについての状態推定計算によって生成され、前記方法はさらに、

前記状態推定値の各々ごとに前記状態推定計算を反転させて、前記状態推定値の各々ごとに、再構成された時系列データ値を生成するステップと、

監査下の前記データベースから前記元の時系列データ値を取出すステップと、

前記元の時系列データ値を前記再構成された時系列データ値と対ごとに比較して、前記元の時系列と前記再構成された時系列とが一致するかどうかを判断するステップと、

監査下の前記データベースが、(i) 前記元の時系列と前記再構成された時系列とが一致する場合に修正されておらず、かつ (i i) 前記元の時系列と前記再構成された時系列とが一致しない場合に修正されていたという信号を生成するステップと、

電子データ報告データ構造を生成するステップとを含み、前記電子データ報告データ構造は、

前記元の時系列値の品質に対するセンサ障害の影響を軽減するために前処理中に前記元の時系列値に対する 1 つ以上の変更を記録する前処理モデルと、

圧縮された時系列データベースとを含み、前記圧縮された時系列データベースは、制限された多変量状態推定モデルで状態推定を実行すること、および前記制限された多変量状態推定モデルのパラメータではない値を前記圧縮された時系列データベースから排除することによって生成され、前記電子データ報告データ構造はさらに、

(i) 前記元の時系列値から選択されるとともに前記制限された多変量状態推定モデルをトレーニングするために用いられるトレーニング値のセットと、(i i) 前記制限された多変量状態推定モデルとのうちの 1 つ以上を含む、方法。

【請求項 6】

機械学習モデルの結果を監査するための、コンピュータによって実現される方法であっ

て、

監査下のデータベースから元の時系列データ値についての状態推定値のセットを取出すステップを含み、前記状態推定値の各々は、前記時系列データ値のうちの1つについての状態推定計算によって生成され、前記方法はさらに、

前記状態推定値の各々ごとに前記状態推定計算を反転させて、前記状態推定値の各々ごとに、再構成された時系列データ値を生成するステップと、

監査下の前記データベースから前記元の時系列データ値を取出すステップと、

前記元の時系列データ値を前記再構成された時系列データ値と対ごとに比較して、前記元の時系列と前記再構成された時系列とが一致するかどうかを判断するステップと、

監査下の前記データベースが、(i) 前記元の時系列と前記再構成された時系列とが一致する場合に修正されておらず、かつ (i i) 前記元の時系列と前記再構成された時系列とが一致しない場合に修正されていたという信号を生成するステップとを含み、

前記状態推定計算の反転はさらに、

前記元の時系列データ値から前記状態推定値を形成するために前記状態推定計算によって実行されるステップを取消す反転状態推定計算を識別するステップと、

前記状態推定値のセットから前記元の時系列データについての反転状態推定値のセットを生成するステップとを含み、前記反転状態推定値の各々は、前記状態推定値のセットのうちの1つについて前記反転状態推定計算を実行することによって生成され、

前記状態推定値の各々についての前記再構成された時系列データ値は、前記反転状態推定値に基づいている、方法。

【請求項7】

(i) 監査下の前記データベースが修正されていなかったという信号に応答して、監査下の前記データベースが破損しておらず、かつ改ざんされていないものとして証明されていることを示す電子検証報告メッセージを生成し、

(i i) 監査下の前記データベースが修正されていたという信号に応答して、監査下の前記データベースが破損しているか、または改ざんされていることを示す電子検証報告メッセージを生成するステップと、

前記生成された電子検証報告をコンピューティングデバイスに送信して、前記電子検証報告メッセージを前記コンピューティングデバイスによって記憶させるか、または前記コンピューティングデバイスによって表示させるステップとをさらに含む、請求項1～6のいずれか1項に記載の方法。

【請求項8】

機械学習モデルの結果を監査するためのコンピュータ実行可能命令を記憶するコンピュータプログラムであって、前記コンピュータ実行可能命令は、コンピュータの少なくともプロセッサによって実行されると、前記コンピュータに、請求項1～7のいずれか1項に記載の方法を実行させる、コンピュータプログラム。

【請求項9】

機械学習モデルの結果を監査するためのコンピュータ実行可能命令を記憶するコンピュータプログラムであって、前記コンピュータ実行可能命令は、コンピュータの少なくともプロセッサによって実行されると、前記コンピュータに、

監査下のデータベースから元の時系列データ値についての状態推定値のセットを取出させ、前記状態推定値は、前記時系列データ値の各々ごとに状態推定計算によって生成されたものであり、前記コンピュータ実行可能命令は、前記コンピュータの少なくともプロセッサによって実行されると、前記コンピュータにさらに、

前記状態推定値の各々ごとに前記状態推定計算を反転させて、前記状態推定値の各々ごとに、再構成された時系列データ値を生成させ、

監査下の前記データベースから前記元の時系列データ値を取出させ、

前記元の時系列データ値を前記再構成された時系列データ値と対ごとに比較させて、前記元の時系列と前記再構成された時系列とが一致するかどうかを判断させ、

監査下の前記データベースが、(i) 前記元の時系列と前記再構成された時系列とが

10

20

30

40

50

一致する場合に修正されておらず、かつ（i i）前記元の時系列と前記再構成された時系列とが一致しない場合に修正されていたという信号を生成させ、

前記元の時系列値から選択されたトレーニング値のセットで多変量状態推定モデルをトレーニングさせ、

前記多変量状態推定モデルに関連付けられた行列を固有ベクトルのセットに分解させ、

前記固有ベクトルのセットから最大の固有値を有する固有ベクトルのサブセットを選択させ、

前記固有ベクトルのサブセットから、制限された多変量状態推定モデルを作成させ、

前記元の時系列値を前処理させて、前記元の時系列値の品質に対するセンサ障害の影響を軽減させ、

前記前処理中に前記元の時系列値に対する1つ以上の変更を記録するデータ前処理モデルを作成させ、

前記制限された多変量状態推定モデルを用いて前記元の時系列データ値について状態推定を実行させて、圧縮された時系列データベースを作成させ、

前記データ前処理モデルと、前記圧縮された時系列データベースと、（i）前記トレーニング値のセットおよび（i i）前記制限された多変量状態推定モデルのうちの1つ以上とを含む報告を生成させ、

前記状態推定値のセットは、前記報告内の前記圧縮された時系列データベースに含まれる前記状態推定値のセットであり、前記再構成された時系列データは、前記報告に基づいて生成される、コンピュータプログラム。

【請求項10】

機械学習モデルの結果を監査するためのコンピューティングシステムであって、プロセッサと、

前記プロセッサに動作可能に接続されたメモリと、

前記プロセッサおよび前記メモリに動作可能に接続されるとともにコンピュータ実行可能命令を記憶する非一時的なコンピュータ可読媒体とを含み、前記コンピュータ実行可能命令は、コンピュータの少なくともプロセッサによって実行されると、前記コンピュータに、請求項1～7のいずれか1項に記載の方法を実行させる、コンピューティングシステム。

【請求項11】

機械学習モデルの結果を監査するためのコンピューティングシステムであって、プロセッサと、

前記プロセッサに動作可能に接続されたメモリと、

前記プロセッサおよび前記メモリに動作可能に接続されたセンサインターフェイスと、

前記プロセッサおよび前記メモリに動作可能に接続されるとともにコンピュータ実行可能命令を記憶する非一時的なコンピュータ可読媒体とを含み、前記コンピュータ実行可能命令は、コンピュータの少なくともプロセッサによって実行されると、前記コンピュータに、

監査下のデータベースから前記センサインターフェイスを介して受信された元の時系列データ値についての状態推定値のセットを取出させ、前記状態推定値は、前記時系列データ値の各々ごとに状態推定計算によって生成されたものであり、さらに、

前記状態推定値の各々ごとに前記状態推定計算を反転させて、前記状態推定値の各々ごとに、再構成された時系列データ値を生成させ、

監査下の前記データベースから前記元の時系列データ値を取出させ、

前記元の時系列データ値を前記再構成された時系列データ値と対ごとに比較させて、前記元の時系列と前記再構成された時系列とが一致するかどうかを判断させ、

監査下の前記データベースが、（i）前記元の時系列と前記再構成された時系列とが一致する場合に修正されておらず、かつ（i i）前記元の時系列と前記再構成された時系列とが一致しない場合に修正されていたという信号を生成させ、

前記状態推定計算の反転のための前記命令はさらに、前記コンピューティングシステム

10

20

30

40

50

に電子データ報告データ構造を生成させ、前記電子データ報告データ構造は、

前記元の時系列値の品質に対するセンサ障害の影響を軽減するために前処理中に前記元の時系列値に対する1つ以上の変更を記録する前処理モデルと、

圧縮された時系列データベースとを含み、前記圧縮された時系列データベースは、制限された多変量状態推定モデルで状態推定を実行すること、および前記制限された多変量状態推定モデルのパラメータではない値を前記圧縮された時系列データベースから排除することによって生成され、前記電子データ報告データ構造はさらに、

(i) 前記元の時系列値から選択されるとともに前記制限された多変量状態推定モデルをトレーニングするために用いられるトレーニング値のセットと、(ii) 前記制限された多変量状態推定モデルとのうちの1つ以上を含み、

前記状態推定計算の反転のための前記命令はさらに、前記コンピュータに、

反転状態推定計算を用いて前記状態推定計算によって実行されるステップを取消させ、

前記状態推定値の各々ごとに前記前処理モデルによって記述される前記前処理を反転させる、コンピューティングシステム。

【発明の詳細な説明】

【背景技術】

【0001】

背景

センサおよびプロセッサ技術のコストの削減に伴って、センサは、センサベースの自動化を可能にするために、電氣的システム、機械的システム、配信システムおよび他のシステムの構成要素に追加されるかまたはこれら構成要素に関連付けられている。これらのセンサは、構成要素の挙動を記述する大量の情報を生成する。この大量の情報は、少なくとも一定の期間だけデータベースに記憶される。検知された情報に機械学習アルゴリズムを適用することにより、センサによって監視されるシステム構成要素の予知、異常発見／検出、および予知保全が可能になり得る。

【発明の概要】

【発明が解決しようとする課題】

【0002】

センサベースの自動化の例は、電力事業、石油およびガス、環境および水質監視、データ処理、製造、乗客および貨物輸送、さらには金融サービス部門において見出すことができる。これらの分野では、機器もしくはシステムの挙動および／または機械学習プロセスによってなされる判断は、規制機関による点検の対象になる可能性がある。このような点検は記憶されたセンサ情報に基づいてなされる可能性があり、規制に準拠しない挙動を示す記憶されたセンサ情報に基づいて、違反を犯した実体に対して多額の罰金が課される可能性もある。したがって、記憶されたセンサ情報に破損または改ざんがないことを、規制対象の実体および規制機関が確実にすることができれば有益である。

【0003】

上述の理由および他の理由から、保証すべき機械学習モデルの結果を監査できることを有効かつ効率的に確実にするための技術が必要とされている。

【課題を解決するための手段】

【0004】

概要

一実施形態では、機械学習モデルの結果を監査するための、コンピュータによって実現される方法は、監査下のデータベースから元の時系列データ値についての状態推定値のセットを取出すステップを含み、当該状態推定値の各々は、当該時系列データ値のうちの1つについての状態推定計算によって生成され、当該方法はさらに、当該状態推定値の各々ごとに当該状態推定計算を反転させて、当該状態推定値の各々ごとに、再構成された時系列データ値を生成するステップと、監査下の当該データベースから当該元の時系列データ値を取出すステップと、当該元の時系列データ値を当該再構成された時系列データ値と対ごとに比較して、当該元の時系列と当該再構成された時系列とが一致するかどうかを判断

するステップと、監査下の当該データベースが、(i) 当該元の時系列と当該再構成された時系列とが一致する場合に修正されておらず、かつ(ii) 当該元の時系列と当該再構成された時系列とが一致しない場合に修正されていたという信号を生成するステップとを含む。

【0005】

一実施形態では、機械学習モデルの結果を監査するための、コンピュータによって実現される方法はさらに、当該元の時系列値から選択されたトレーニング値のセットで、多変量状態推定モデルをトレーニングするステップと、当該多変量状態推定モジュールに関連付けられた行列を固有ベクトルのセットに分解するステップと、当該固有ベクトルのセットから主固有ベクトルのサブセットを選択するステップと、当該主固有ベクトルのサブセットから、制限された多変量状態推定モデルを作成するステップとをさらに含み、当該状態推定計算の反転は、当該制限された多変量状態推定モデルが状態推定値を形成する計算の反転を生成するステップをさらに含む。

10

【0006】

一実施形態では、機械学習モデルの結果を監査するための、コンピュータによって実現される方法はさらに、監査下の当該データベースからの当該元の時系列データ値からトレーニングデータ値を除外するステップと、残りの元の時系列データ値を前処理してセンサ障害の影響を軽減し、当該残りの元の時系列データに対する変更を記録するモデルを生成するステップと、当該制限された多変量状態推定モデルを用いて、当該前処理された残りの元の時系列データ値について状態推定を実行して、圧縮された時系列データベースを作成するステップとを含む。

20

【0007】

一実施形態では、機械学習モデルの結果を監査するための、コンピュータによって実現される方法はさらに、当該元の時系列値を前処理して、当該元の時系列値の品質に対するセンサ障害の影響を軽減するステップと、当該前処理中に当該元の時系列値に対する1つ以上の変更を記録するデータ前処理モデルを生成するステップとを含み、当該状態推定計算の反転はさらに、当該データ前処理モデルを取出すステップと、当該状態推定値の各々ごとに当該データ前処理モデルによって記述される当該前処理を反転させるステップとを含む。

【0008】

一実施形態では、機械学習モデルの結果を監査するための、コンピュータによって実現される方法はさらに、電子データ報告データ構造を生成するステップを含み、当該電子データ報告データ構造は、当該元の時系列値の品質に対するセンサ障害の影響を軽減するために前処理中に当該元の時系列値に対する1つ以上の変更を記録する前処理モデルと、圧縮された時系列データベースとを含み、当該圧縮された時系列データベースは、制限された多変量状態推定モデルで状態推定を実行すること、および当該制限された多変量状態推定モデルのパラメータではない値を当該圧縮された時系列データベースから排除することによって生成され、当該電子データ報告データ構造はさらに、(i) 当該元の時系列値から選択されるとともに当該制限された多変量状態推定モデルをトレーニングするために用いられるトレーニング値のセットと、(ii) 当該制限された多変量状態推定モデルとのうちの1つ以上を含む。

30

40

【0009】

一実施形態では、機械学習モデルの結果を監査するための、コンピュータによって実現される方法であって、当該状態推定計算の反転はさらに、当該元の時系列データ値から当該状態推定値を形成するために当該状態推定計算によって実行されるステップを取消す反転状態推定計算を識別するステップと、当該状態推定値のセットから当該元の時系列データについての反転状態推定値のセットを生成するステップとを含み、当該反転状態推定値の各々は、当該状態推定値のセットのうちの1つについて当該反転状態推定を実行することによって生成され、当該状態推定値の各々についての当該再構成された時系列データ値は、当該反転状態推定値に基づいている。

50

【0010】

一実施形態では、機械学習モデルの結果を監査するための、コンピュータによって実現される方法はさらに、(i) 監査下の当該データベースが修正されていなかったという信号に応答して、監査下の当該データベースが破損しておらず、かつ改ざんされていないものとして証明されていることを示す電子検証報告メッセージを生成し、(ii) 監査下の当該データベースが修正されていたという信号に応答して、監査下の当該データベースが破損しているか、または改ざんされていることを示す電子検証報告メッセージを生成するステップと、当該生成された電子検証報告をコンピューティングデバイスに送信して、当該検証報告メッセージを当該コンピューティングデバイスによって記憶させるか、または当該コンピューティングデバイスによって表示させるステップとを含む。

10

【0011】

一実施形態では、機械学習モデルの結果を監査するためのコンピュータ実行可能命令を記憶する非一時的なコンピュータ可読媒体であって、当該コンピュータ実行可能命令は、コンピュータの少なくともプロセッサによって実行されると、当該コンピュータに、監査下のデータベースから元の時系列データ値についての状態推定値のセットを取出させ、当該状態推定値は、当該時系列データ値の各々ごとに状態推定計算によって生成されたものであり、さらに、当該状態推定値の各々ごとに当該状態推定計算を反転させて、当該状態推定値の各々ごとに、再構成された時系列データ値を生成させ、監査下の当該データベースから当該元の時系列データ値を取出させ、当該元の時系列データ値を当該再構成された時系列データ値と対ごとに比較させて、当該元の時系列と当該再構成された時系列とが一致するかどうかを判断させ、監査下の当該データベースが、(i) 当該元の時系列と当該再構成された時系列とが一致する場合に修正されておらず、かつ(ii) 当該元の時系列と当該再構成された時系列とが一致しない場合に修正されていたという信号を生成させる。

20

【0012】

一実施形態では、非一時的なコンピュータ可読媒体であって、当該命令はさらに、当該コンピュータに、当該元の時系列値から選択されたトレーニング値のセットで多変量状態推定モデルをトレーニングさせ、当該多変量状態推定モジュールに関連付けられた行列を固有ベクトルのセットに分解させ、当該固有ベクトルのセットから最大の固有値を有する固有ベクトルのサブセットを選択させ、当該固有ベクトルのサブセットから、制限された多変量状態推定モデルを作成させ、当該元の時系列値を前処理させて、当該元の時系列値の品質に対するセンサ障害の影響を軽減させ、当該前処理中に当該元の時系列値に対する1つ以上の変更を記録するデータ前処理モデルを作成させ、当該制限された多変量状態推定モデルを用いて当該元の時系列データ値について状態推定を実行させて、圧縮された時系列データベースを作成させ、当該前処理モデルと、当該圧縮された時系列データベースと、(i) 当該トレーニング値のセットおよび(ii) 当該制限された多変量状態推定モデルのうちの1つ以上と、を含む報告を生成させ、当該状態推定値のセットは、当該報告内の当該圧縮された時系列データベースに含まれる当該状態推定値のセットであり、当該再構成された時系列データは、当該報告に基づいて生成される。

30

【0013】

一実施形態では、機械学習モデルの結果を監査するためのコンピューティングシステムであって、当該システムは、プロセッサと、当該プロセッサに動作可能に接続されたメモリと、当該プロセッサおよび当該メモリに動作可能に接続されたセンサインターフェイスと、当該プロセッサおよび当該メモリに動作可能に接続されるとともにコンピュータ実行可能命令を記憶する非一時的なコンピュータ可読媒体とを含み、当該コンピュータ実行可能命令は、コンピュータの少なくともプロセッサによって実行されると、当該コンピュータに、監査下のデータベースから当該センサインターフェイスを介して受信された元の時系列データ値についての状態推定値のセットを取出させ、当該状態推定値は、当該時系列データ値の各々ごとに状態推定計算によって生成されたものであり、さらに、当該状態推定値の各々ごとに当該状態推定計算を反転させて、当該状態推定値の各々ごとに、再構成された時系列データ値を生成させ、監査下の当該データベースから当該元の時系列データ

40

50

値を取出させ、当該元の時系列データ値を当該再構成された時系列データ値と対ごとに比較させて、当該元の時系列と当該再構成された時系列とが一致するかどうかを判断させ、監査下の当該データベースが、(i) 当該元の時系列と当該再構成された時系列とが一致する場合に修正されておらず、かつ(ii) 当該元の時系列と当該再構成された時系列とが一致しない場合に修正されていたという信号を生成させる。

【0014】

一実施形態では、機械学習モデルの結果を監査するためのコンピューティングシステムであって、当該非一時的なコンピュータ可読媒体はさらに命令を含み、当該命令は、少なくとも当該プロセッサによって実行されると、当該コンピューティングシステムに、当該元の時系列値から選択されたトレーニング値のセットで多変量状態推定モデルをトレーニングさせ、当該多変量状態推定モジュールに関連付けられた行列を固有ベクトルのセットに分解させ、当該固有ベクトルのセットから主固有ベクトルのサブセットを選択させ、当該主固有ベクトルのサブセットから、制限された多変量状態推定モデルを作成させ、当該状態推定計算の反転は、当該制限された多変量状態推定モデルが状態推定値を形成する計算の反転を生成することをさらに含む。

【0015】

一実施形態では、機械学習モデルの結果を監査するためのコンピューティングシステムであって、当該非一時的なコンピュータ可読媒体はさらに命令を含み、当該命令は、少なくとも当該プロセッサによって実行されると、当該コンピューティングシステムに、監査下の当該データベースからの当該元の時系列データ値からトレーニングデータ値を除外させ、残りの元の時系列データ値を前処理させてセンサ障害の影響を軽減させ、当該残りの元の時系列データに対する変更を記録するモデルを生成させ、当該制限された多変量状態推定モデルを用いて当該前処理された残りの元の時系列データ値について状態推定を実行させて、圧縮された時系列データベースを作成させる。

【0016】

一実施形態では、機械学習モデルの結果を監査するためのコンピューティングシステムであって、当該状態推定計算の反転のための当該命令はさらに、当該コンピューティングシステムに電子データ報告データ構造を生成させ、当該電子データ報告データ構造は、当該元の時系列値の品質に対するセンサ障害の影響を軽減するために前処理中に当該元の時系列値に対する1つ以上の変更を記録する前処理モデルと、圧縮された時系列データベースとを含み、当該圧縮された時系列データベースは、制限された多変量状態推定モデルで状態推定を実行すること、および当該制限された多変量状態推定モデルのパラメータではない値を当該圧縮された時系列データベースから排除することによって生成され、当該電子データ報告データ構造はさらに、(i) 当該元の時系列値から選択されるとともに当該制限された多変量状態推定モデルをトレーニングするために用いられるトレーニング値のセットと、(ii) 当該制限された多変量状態推定モデルとのうちの1つ以上を含み、当該状態推定計算の反転のための当該命令はさらに、当該コンピュータに、反転状態推定計算を用いて当該状態推定計算によって実行されるステップを取消させ、当該状態推定値の各々ごとに当該前処理モデルによって記述される当該前処理を反転させる。

【0017】

一実施形態では、機械学習モデルの結果を監査するためのコンピューティングシステムであって、当該非一時的なコンピュータ可読媒体はさらに命令を含み、当該命令は、少なくとも当該プロセッサによって実行されると、当該コンピューティングシステムに、当該元の時系列データ値から当該状態推定値を形成するために当該状態推定計算によって実行されるステップを取消す反転状態推定計算を識別させ、当該状態推定値のセットから当該元の時系列データについての当該反転状態推定値のセットを生成させ、当該反転状態推定値の各々は、当該状態推定値のセットのうちの1つについて反転状態推定を実行することによって生成され、当該状態推定値の各々についての当該再構成された時系列データ値は、当該反転状態推定値に基づいている。

【0018】

一実施形態では、機械学習モデルの結果を監査するためのコンピューティングシステムであって、当該非一時的なコンピュータ可読媒体はさらに命令を含み、当該命令は、少なくとも当該プロセッサによって実行されると、当該コンピューティングシステムに、(i) 監査下の当該データベースが修正されていなかったという信号に応答して、監査下の当該データベースが破損しておらず、かつ改ざんされていないものとして証明されていることを示す電子検証報告メッセージを生成させ、(ii) 監査下の当該データベースが修正されていたという信号に応答して、監査下の当該データベースが破損しているか、または改ざんされていることを示す電子検証報告メッセージを生成させ、当該生成された電子検証報告をコンピューティングデバイスに送信して、当該検証報告メッセージを当該コンピューティングデバイスによって記憶させるかまたは当該コンピューティングデバイスによって表示させる。

10

【0019】

図面の簡単な説明

明細書中に組込まれてその一部を構成する添付の図面は、開示のさまざまなシステム、方法および他の実施形態を示すものである。図に例示される要素の境界（たとえば、箱、箱の群または他の形状）がこれら境界の一実施形態を表わすことが認識されるだろう。いくつかの実施形態においては、1つの要素が複数の要素として実現されてもよく、または、複数の要素が1つの要素として実現されてもよい。いくつかの実施形態においては、別の要素の内部構成要素として示される要素が外部構成要素として実現されてもよく、この逆であってもよい。さらに、要素は縮尺通りには示されないこともある。

20

【図面の簡単な説明】

【0020】

【図1】機械学習モデルの結果を確実に監査できるようにすることに関連するシステムの一実施形態を示す図である。

【図2】機械学習モデルの結果を確実に監査できるようにすることに関連する方法の一実施形態を示す図である。

【図3】機械学習モデルの結果を確実に監査できるようにすることに関連するプロセスの一実施形態のフローチャートを示す図である。

【図4】例示的なIDPモデルにおけるインテリジェントデータ前処理からの変更記録を記憶する一実施形態を示す概略図である。

30

【図5】機械学習モデルの結果を確実に監査できるようにすることに関連するMSETモデルトレーニングプロセスの一実施形態を示す概略図である。

【図6】機械学習モデルの結果を確実に監査できるようにすることに関連するMSETモデル制限プロセスの一実施形態を示す概略図である。

【図7】機械学習モデルの結果を確実に監査できるようにすることに関連するデータ圧縮プロセスの一実施形態を示す概略図である。

【図8A】データ圧縮プロセスの一実施形態から得られる第1の例示的なデータ報告フォーマットを示す図である。

【図8B】データ圧縮プロセスの別の実施形態から得られる第2の例示的なデータ報告フォーマットを示す図である。

40

【図9】第1のデータ報告フォーマットを用いるデータ再構築プロセスの一実施形態を示す概略図である。

【図10】第2のデータ報告フォーマットを用いるデータ再構築プロセスの一実施形態を示す概略図である。

【図11】本明細書に記載の例示的なシステムおよび方法ならびに／または同等例のうちの1つ以上で構成および／またはプログラムされた例示的なコンピューティングデバイスを示す図である。

【発明を実施するための形態】

【0021】

詳細な説明

50

機械学習モデルの結果を確実に監査できるようにするシステムおよび方法を本明細書に記載する。

【0022】

モノのインターネット（Internet-of-Things：I o T）センサなどのセンサを物理デバイスに追加して、これらのデバイスの動作を監視し得る。これらのセンサは、特に、公益事業、石油およびガス、ならびに製造などの高密度センサ産業において多数存在し得る。たとえば、石油精製所は、100万を超えるセンサを備え得る。電力系統は、特に、発電所および変圧器変電所などの公益事業資産に関する監視制御およびデータ取得（supervisory control and data acquisition：SCADA）のためのセンサ、ならびに高度計量インフラストラクチャ（advanced metering infrastructure：AMI）のためのセンサを考慮に入れた場合、100万個をはるかに超えるセンサを備え得る。これらのセンサからのデータは、時系列の、時間順序でインデックス付けされた一連のデータ点、または値の対および関連する時間として記憶され得る。時系列データは、時系列データベース（時系列を記憶および提供するために最適化されたデータベースシステム）に記憶されてもよい。したがって、公益事業は、センサ読取り値をペタバイト以上のオーダーで含む非常に大規模な時系列データベースを生成することができる。

10

【0023】

なお、物理的なトランスデューサを備えたシステムから生じる大量の時系列信号を取込むシステムが、故障した物理的なトランスデューサ／センサから誤った時系列値を収集するであろうことに留意されたい。いくらかの割合のセンサが劣化している場合、および／または、入力を受けているにもかかわらず1つの値を永久的に信号で伝え続ける「縮退（stuck-at）」故障により機能停止し得る場合もあり得る。また、センサまたは上流のデータ取得および／もしくはデータ集約電子機器に断続的に問題が起こって、個々の時系列が欠落値を含む可能性もあり得る。時系列データを時系列データベースに取込む際にこれらのタイプの異常が発見されない場合、これらの異常は、時系列データベースに記憶されたデータのその後の使用に影響を及ぼす。たとえば、時系列データが製品開発または他の科学的な目的のために用いられる場合、データ分析の精度や、場合によっては、引出された結論の健全性は、上述のタイプの異常によって悪影響を受けるだろう。本明細書に記載のシステムおよび方法はまた、重要な資産を監視するセンサから生じる時系列データベースのために信号検証およびセンサ操作性検証を提供することによって、この問題を解決することを可能にする。

20

30

【0024】

機械学習（machine learning：ML）アルゴリズムは、センサによって監視されるデバイスについての予知、異常発見／検出、および予知保全を可能にするために、時系列データベースに記憶されたセンサデータに適用され得る。これは自動予知監視と称され得る。特に公益事業のための高密度センサ産業において重要性が高まっている1つの局面は、予知MLアルゴリズムが1つ以上の信号について異常なものとしてフラグを立てた理由と、これと等しく重要であるが、MLが警告を発しなかった理由とを、規制機関、保険調査員、および末端の顧客に説明する能力である。したがって、本明細書に記載のシステムおよび方法は、パターン認識ML判断のための「監査可能性保証」を提供する。

40

【0025】

自動予知監視システムを利用する公益事業などの産業は、法的要件を満たさなければならず、厳重な政府の規制を受ける。たとえば、米国における公益事業は、地域毎、群毎、州毎の規制および連邦政府による規制の対象となっており、これら規制には、当該公益事業が原子力プラントを運用する場合には、アメリカ合衆国原子力規制委員会（the United States Nuclear Regulatory Commission：USNRC）からの追加の連邦規制や、北米電力信頼度協議会（the North America Electric Reliability Corporation：NERC）および米国連邦エネルギー規制委員会（the Federal Energy Regulatory Commission：FERC）からの追加の規制が含まれる。原則として、このような要件および規制は、法的な目的または規制上の目的でデータを提出することを必要とすることもある。し

50

かしながら、これらの大規模な自動化システムによって行なわれる判断の基礎をなすペタバイトのデータの全質量を提出することは不可能である。なぜなら、データセットのサイズが膨大であるからだけではなく、データのほとんどが、このような提出を可能にし得るような態様で持続的に記憶されていないからである。

【0026】

法的な目的および規制上の目的でデータを抽出するための技術は、以下の不利点のいくつかまたは全てを被ることが多い。(i) 当該技術は、厳密な理論的根拠なしに、その場しのぎで、必要に応じて、または任意の基準を用いてデータを選択する可能性がある。(ii) 当該技術のための選択されたデータセットは、元のソースデータを復元するのに不十分である。(iii) 当該技術は、元のソースデータを復元することができないので規制要件の違反の隠蔽を可能にしてしまう恐れがある。実際のところ、第3の不利点に関して、このことが実務で実際に発生したことが報告されている。

【0027】

特に、元のソースデータの復元可能性が欠如していることにより、産業と政府との間に不必要な敵対的関係をもたらしてしまう。たとえば、カリフォルニア州で操業している電気公益事業は、各月にわたって最低10%の発電オーバーヘッドクッションを維持することが求められている。すなわち、いずれかの所与の日および時間の間に需要があったとしても、少なくとも10%のオーバーヘッドまたは「スペア」容量が存在し、このため、発電資産または送電網資産が突然故障したとしても、オーバーヘッド容量により全体が停電する可能性は低くなる。公益事業は、これら公益事業が各月にわたって最低10%の発電オーバーヘッドクッションを維持したことを実証するために、月毎に収集した時系列データをカリフォルニア公共公益事業委員会(the California Public Utilities Commission: PUC)に提出する必要がある。公益事業の発電能力が10%の発電オーバーヘッドクッションを下回るたびに、PUCは多額の罰金を課す(これは1ヶ月に複数回起こり得る)。PUCは時系列信号を非常に注意深く精査する。なぜなら、PUCは、発電能力が10%閾値を下回っていないように見せかけるために公益事業がいくつかの時系列値を調整する動機を有するものであると、習慣的に想定しているからである。同様に、公益事業は、発電能力が10%閾値をわずかに数%だけ下回ったかのように見せかけることによって当該公益事業から多額の罰金を引出すためにPUCがいくつかの時系列値を調整する動機を有するものであると、習慣的に想定している。

【0028】

これらの(両者側の不信を含む)不利点は、機械学習モデルの結果を確実に監査できるようにする、本明細書に記載のシステムおよび方法によって解消される。監査可能性を確実にするための一実施形態では、以下の情報が、たとえばデータベースに対する変更のジャーナルに保持されるべきである。

【0029】

・(たとえば、タイムスタンプ付きセンサ読取り値と、もしあれば、センサ入力に関連付けられたユーザとを含む)検知された時系列データ。

【0030】

・センサの説明(たとえば、センサに関連付けられた検知された資産、測定時のセンサの動作状態)。

【0031】

・検知された資産の説明(たとえば、測定時の資産のバージョン、当該資産と他の資産との間の関係の記録、たとえば、時間に関するインスタンスグラフまたは時間に関する部品表)。

【0032】

・測定時の「知識」状態(たとえば、1つ以上の特定の機械学習モデル、特定の条件、および分析時に用いられる特定のデータ)。

【0033】

・特定の機械学習モデルを開発するとともにセンサ読取り値を分析するために用いられ

10

20

30

40

50

る機械学習モデル作成ソフトウェアのバージョンおよびデータセットの識別。

【0034】

なお、検知された時系列データが非常に大量になる可能性があることに留意されたい。一実施形態では、大量のデータを抽出する代わりに、比較的少量の、但し、注目すべき（または最も目立つかまたは重要な）データの集合を保存することができ、このデータの集合から、元データに極めて近似するデータを回復することができる。これは法的要件および規制上の要件には十分である。回復可能性のこのような特性を備えた保存データの集合（すなわち、保存されたデータからの元データについての近似データの回復可能性）は、本明細書中では「改ざん防止」データセットと称される。一実施形態では、改ざん防止データセットの作成は、発明者D.Gawlick、K.C.Gross、Z.H.Liu、およびA.Ghoneimyにより2018年3月19日に出願され「多次元時系列データのインテリジェント前処理（Intelligent Preprocessing of Multi-Dimensional Time-Series Data）」と題され、その全体が引用により本明細書に援用されている米国特許出願連続番号第15／925,427号において記載される技術のうちの1つ以上を含み得る。悪意あるアクターが元データを変更すること、またはさらには、改ざん防止データセットを変更することが必ずしも防止されるわけではないが、本明細書に開示されるシステムおよび方法は、このような改ざんが本明細書に記載の手順に従って行なわれる監査において発見されるであろうことを積極的に保証するものである。

10

【0035】

一実施形態では、このため、本明細書に記載のシステムおよび方法は、法的要件および規制上の要件を満たすために用いることができる比較的少量の処理済みデータ（改ざん防止データセット）で、元のソースデータに極めて近似するデータの作成を可能にする。さらに、改ざん防止データセットは、これらの要件を満たすために用いることができる一方で、機械学習信号検証およびセンサ操作性検証技術にも依然として対応することができる。

20

【0036】

一実施形態では、改ざん防止データセットを抽出するための処理技術は3つの有利な特性を呈する。

【0037】

・決定論：改ざん防止データセットの処理済みデータは、入力データによって不規則性なく一義的に決定される。

30

【0038】

・圧縮：改ざん防止データセットの処理済みデータは、入力データと比較して比較的小さい。実際には、改ざん防止データセットの処理済みデータは、数オーダー分小さい可能性がある。

【0039】

・可逆性：改ざん防止データセットの処理済みデータから元のソースデータに比較的近い近似データを回復することができる。

【0040】

実際には、近似の程度は法的要件および規制上の要件によって判断されるとともに、これらの要件を満たすように調整され得る。

40

【0041】

多変量状態推定技術（Multivariate State Estimation Technique：MSE T）などの処理技術は、決定論、圧縮、および可逆性の特性を呈しているので、本発明の一実施形態において用いられる。しかしながら、上記の3つの特性を満たすデータセットを抽出するための任意の処理技術は、本明細書に記載のシステムおよび方法に従って改ざん防止データセットを作成するために用いることができる。

【0042】

一実施形態では、上記の3つの特性に加えて、改ざん防止データセットおよび履歴所在情報（これはデータベースに対する変更を追跡するが、データの完全性を保証するものではない）に基づいて、法的要件および規制上の要件が満たされたかどうかを判断するため

50

の自動化技術がある。これらの要件が満たされていない場合、システムは、適切な法的対応策または政府の罰金を決定してもよい。法的プロセスおよび規制上のプロセスをすべての当事者にとって完全に明快にすることにより、違反および敵対的関係を隠蔽しようと試みる動機付けを大幅に削減させるかまたはさらには排除することができる。

【0043】

一実施形態では、本明細書に記載のシステムおよび方法は、それらが適用される既存のML監視システムを改善することで、警告の正確性をさらに高めるとともに、ML予知異常発見のための誤警告率を低減させる。これらの改善が、本明細書に記載のシステムおよび方法の実現例によって達成され得るとともに、それらが実現されるシステム内のどこであってもハードウェアアップグレードを必要としないことに留意されたい。したがって、本明細書に記載のシステムおよび方法は、既存のいずれのIoTシステムとも直接的な後方互換性がある。このことは、従来のセンサデータ収集システムがすでに所定位置にあってアップグレードするのにかなりの労力を必要とするであろう電力公益事業、石油およびガス、製造、ならびに航空産業において特に有利である。

【0044】

本明細書に記載のシステムおよび方法は、電力公益事業セクタに関連付けて説明されているが、たとえば、石油およびガス、製造、ならびに航空セクタにおいて、IoTセンサ時系列データが収集されて用いられるような場合であればどんな用途も有することは明らかである。一実施形態では、本明細書に記載のシステムおよび方法は、発電設備内部の公益事業資産（たとえば、石炭発電所、石油発電所、原子力発電所、風力タービン、地熱発電機、ガスタービン発電所など、ならびに、配電網内の重要な資産、たとえば、変圧器、変電所およびSCADAシステムなど）についてのデジタル化されたデータをストリーミングする処理において適用され得る。

【0045】

ー公益事業予知のための時系列プロセスにおける監査可能性ー

一実施形態では、機械学習モデルの結果を確実に監査できるようにするための方法およびプロセスは、(i) 改ざん防止、(ii) スナップショット分離、(iii) ジャーナリング、(iv) アクティビティジャーナリング、および(v) データ所在記録、といった特徴を含む。

【0046】

一実施形態では、改ざん防止は、悪意あるデータ修正（改ざん）または偶発的なデータ修正（破損）が確実に検出されるようにするためのプロセスまたはシステム構成である。一実施形態では、元データに対するいずれの変更も明示するであろうコンパクトな記録を監査報告に含めることができる。

【0047】

時系列データ中の異常値は、未加工の時系列データのML処理によって識別されて、周囲データの文脈においてこれらの値が何で「あるはずである」かを示す推定値を生成し得る。ニューラルネットワーク(neural network: NN)およびサポートベクトルマシン(support vector machine: SVM)が異常検出のために採用され得る。MSET(およびオラクル(登録商標)の独自の高度なMSETパターン認識「MSET2」などの変形例)も異常検出のために採用され得る。3つのアプローチ(NN、SVM、およびMSET)はすべて、ブラックボックスレベルでは、非線形ノンパラメトリック(Nonlinear nonparametric: NLNP)回帰アルゴリズムである。NLNP回帰は、主に、時系列データフロープロセスにおける予知、異常発見/検出、および予知保全のために採用される。なぜなら、NLNP機械学習技術は、時系列「信号」の間の線形または非線形の関係についての仮説を立てるのではなく、代わりに、それらの関係を経験的に学習するからである。

【0048】

これらの3つのNLNP機械学習アプローチのうち、NNおよびSVMはともに、重みの最適化のために確率的（または明らかにランダム）なプロセスを採用する。NNの場合

、重みの確率的最適化はパーセプトロン層間で行なわれる。SVMの場合、確率的最適化は、SVM推定値におけるバイアスと分散との間のバランスを維持するために正則化パラメータの凸二次プログラミング最適化で行なわれる。たとえば、両方（NNおよびSVM）の場合、パターン認識が月曜日からのデータでトレーニングされる場合と、火曜日からのデータでトレーニングされる場合とでは、出力推定値と入力未加工信号との間の関係は極めて類似しているだろう。しかしながら、月曜日の計算用の中間重みを火曜日の計算用の中間重みと比較して、ブラックボックスを「調査する（looks into）」場合、これら中間重みは大幅に異なるものになるだろう。パターン認識のみを実験的かつ実用的に適用する場合、ブラックボックスの出力が基礎をなす時系列の正確な推定値である限り、当該ブラックボックスが実行されるたびにブラックボックス内の重みが実質的に異なるものになり得ることは問題ではない。しかしながら、監査可能性を適用する場合、NNおよびSVMによって導入される不規則性により、これらが改ざん防止データセットを抽出するのに適さないものとなる。なぜなら、それらが決定論的ではなく、このため可逆的ではないからである。

【0049】

異常検出（ならびに関連する予知および予知保全）のためにMSETを適用すると、元の未加工の時系列テレメトリ値とともに記憶され得る推定値が得られる。NNおよびSVMとは対照的に、MSETは決定論的な（ただし複雑な）数学的アルゴリズムであり、MSET推定値は上述のように可逆的であり、これは監査可能性保証における改ざん防止の鍵となる。したがって、一実施形態では、MSETは、監査可能性保証の一環として改ざん防止を可能にするために、時系列データプロセスにおける異常検出のために適用される。この異常検出プロセスによって生成される元の未加工データストリームのうちのいずれかが、悪意あるユーザによって、または、記憶媒体内のいずれかのデータ破損エラーを通じて偶発的に、修正、変更、置換または変換された場合、元の未加工データ値に対する変更は、付随するMSET推定値に基づいて検出することができる。または、元の未加工データストリームが改ざんされていないかまたは破損していない場合、元の未加工データ値は、付随する推定値に基づいて、変更されていないと検証または確認することができる。このような改ざんが無いことの証明は、MSET推定値の作成後にいつでも実行され得る。このような改ざんが無いことの証明は、決定論的で可逆的なMSETアルゴリズムを本明細書に記載の異常検出プロセスに組込むことに基づくものである。

【0050】

したがって、時系列データの元の値を推定するためのコンパクトなデータセットは、任意の時点において取込まれ、時系列データベースの寿命中になされた変更に関する他の情報とともに記憶され、監査可能性プロセスにおいて採用され得る。

【0051】

一実施形態では、スナップショット分離は、監査可能性プロセスの一部を形成する。時間に関するデータベースは、任意のバージョンの記録を記憶して取出す能力を提供する。これらのバージョンは、厳密な昇順のトランザクション時間（バージョンが問合せおよび後続処理のために利用可能であった時間または可視であった時間）によって識別される。特定のバージョンは各々、スナップショットと称されることもある。一実施形態では、特定の時間に作成された改ざん防止データセットを用いて、その作成時間に利用可能なスナップショットのデータ値を検証してもよい。

【0052】

一実施形態では、改ざん防止データセットのための特定の作成時間はジャーナルエントリの作成であってもよい。ジャーナリングプロセスは、データベースに対する全ての変更を追跡するように構成される。データベース内での各変更により、結果として、データベースに関連付けられたジャーナルにおいてジャーナルエントリが作成される。ジャーナルは、典型的には、データ損失に対して高い耐性を有する。たとえば、ジャーナリングされたデータベースにおいては、データベースに関して永久媒体上のデータに変更はなく、変更を記述するジャーナルデータが記憶されるまで、当該データベースに関する外部通知は

許可されない。これにより、障害が存在していたとしてもデータベースの監査が可能となる。ジャーナルは、時間的サポートなしでも、スナップショット問合せにおける膨大な性能負担を犠牲にして、当該システムがデータベースの任意のスナップショットを再構築することを可能にする。すべてのデータベースは最新のスナップショットのコピーを有しており、時間に関するデータベースは、時間を遡って多くのスナップショットを提供する。センサ読取り値などの不変のデータについては、スナップショットおよびジャーナルを単一のコピーとして記憶することができる。

【0053】

しかしながら、一実施形態では、監査は、ユーザのアクティビティに関する追加の知識（たとえば、「（問合せに回答して）どんな情報を誰がいつ見たか？」、または、「誰が、いつ、どんな情報を挿入、更新または削除したのか？」、および、「同じトランザクションでどんな動作が実行されたか（動作はどのように関連しているのか）？」）についての質問への回答）を用いてもよい。このような情報はアクティビティジャーナルエントリに記録され得る。アクティビティジャーナルのエントリは、標準的なジャーナルエントリと同期される。こうして、アクティビティジャーナルは、データベース変更とジャーナルエントリの付随する作成とにつながる状況を記述する追加のコンテキスト情報を提供する。

【0054】

一実施形態では、データ所在を記述する情報も標準的なジャーナルエントリに含めるかまたは当該標準的なジャーナルエントリと同期させて、データベース変更とジャーナルエントリの付随する作成とにつながる状況を記述する更に別のコンテキスト情報を提供してもよい。データ所在情報は、導き出されたデータを、対応する入力、処理ステップ、および物理的処理環境と関連付ける。たとえば、所在情報は、問合せ結果の基礎を形成するデータを識別するものである。一実施形態では、所在プロセスは、これらのデータを判断するために問合せを書換えるように構成され得る。監査の状況において、所在メタデータは、問合せ結果の基礎を形成した全てのデータの記録を保持しているので、考慮が必要なデータを実質的に低減させる。一実施形態では、所在情報は、データを提供するセンサの記述、センサの操作性状態、センサに関連付けられた資産、資産のバージョン、資産間の関係などの更に別の情報を含み得る。

【0055】

一実施形態では、改ざん防止データセットは、ジャーナルエントリの作成と同時に、データベースに対する変更に応答して作成される。一実施形態では、改ざん防止データセットは、変更に関連付けられた他のメタデータ（アクティビティジャーナリングを含む変更メタデータおよびデータ所在メタデータ）と同様にジャーナルに取込まれる。一実施形態では、改ざん防止データセットは、データベースに対する変更ごとに取込まれ、これにより、本明細書に記載のシステムおよび方法に従って行なわれる監査プロセスにおいて後続のデータ変更がいずれも検出されるであろうことを保証する（とともに、変更時に動作中の機械学習ソフトウェア構成が残りのジャーナル情報とともに取込まれて監査可能となることも保証する）。したがって、本明細書に開示されるシステムおよび方法は、時系列データベース内のデータが後で時間的に進化するにもかかわらず、データのいずれのスナップショット内の元データに対しても監査が実行されることを可能にする。このため、時系列データベースにおける機械学習の結果の監査可能性は、データベースへの変更ごとに必要なすべての情報（すなわち、ジャーナルエントリ、アクティビティジャーナルエントリ、データ所在、および改ざん防止機械学習記録）の取込みを可能にし、これにより信号のデータベースの履歴所在全体を追跡することによって、確実にされる。

【0056】

－例示的な環境－

図1は、機械学習モデルの結果を確実に監査できるようにすることに関連するシステム100の一実施形態を示す。

【0057】

一実施形態では、システム100は、インターネットなどのネットワーク115によっ

10

20

30

40

50

て接続された時系列データサービス１０５および企業ネットワーク１１０を含む。時系列データサービス１０５は、ネットワーク１２５を介してセンサ（センサ１２０など）もしくはリモートターミナルユニット（remote terminal unit：ＲＴＵ）に直接接続されるか、または、１つ以上のアップストリームデバイス１３５を介してセンサ（センサ１３０など）もしくはＲＴＵに間接的に接続される。一実施形態では、ネットワーク１１５および１２５は同じネットワークであり、別の実施形態では、ネットワーク１１５および１２５は別個のネットワークである。

【００５８】

一実施形態では、時系列データサービス１０５は、機械学習監査保証システム１４０と、センサインターフェイスサーバ１４５と、予知、異常発見および予知保全システム１５０と、ウェブインターフェイスサーバ１５５と、データストア１６０とを含み得るさまざまなシステムを含む。これらのシステム１４０～１６０の各々は、サーバ側ネットワーク１６５によって相互接続される。これらのシステム１４０～１６０の各々は、たとえば、それらが実行するものとして説明される機能を実行するためのさまざまなソフトウェアモジュールによる論理で構成される。一実施形態では、システム１４０～１６０は、専用のコンピューティングデバイスによって実現される。一実施形態では、システム１４０～１６０のうちの１つ以上は、図１では別個のユニットとして表わされているが、共通の（または共有される）コンピューティングデバイスによって実現されてもよい。

【００５９】

一実施形態では、時系列データサービス１０５は、第三者によってホストされてもよく、および／または、複数のアカウント所有者／テナントの利益のために第三者によって操作されてもよく、これら複数のアカウント所有者／テナントの各々は事業を運営しており、その各々は関連する企業ネットワーク１１０を有する。一実施形態では、時系列データサービス１０５は、電力公益事業などの公益事業エンティティに関連付けられるか、または、発電設備、変電所、もしくは他の主要な送電網構成要素などの主要な公益事業資産に関連付けられている。一実施形態では、時系列データサービス１０５は、（ｉ）時系列データベースを作成およびエクスポートするために、ならびに／または、（ｉｉ）本明細書に記載のシステムおよび方法に従って時系列データベースを監査するために、時系列データサービス１０５を動作させるためのソフトウェアモジュール等の論理で構成される。

【００６０】

一実施形態では、センサ１２０、１３０は、デバイスまたはシステムの１つ以上の構成要素の性能を検出するように取付けられ得るかまたは構成され得る。当該デバイスまたはシステムは、概して、測定可能なアクティビティを実行する構成要素を備えた任意のタイプの機械類または設備を含む。センサ１２０、１３０は、電圧センサ、電流センサ、温度センサ、圧力センサ、回転速度センサ、流量計センサ、振動センサ、マイクロフォン、電磁放射センサ、近接センサ、ジャイロスコープ、傾斜計、加速度計、全地球測位システム（global positioning system：ＧＰＳ）センサ、トルクセンサ、撓みセンサ、核放射検出器、または、検出もしくは検知された物理的挙動を記述する電気信号を生成するための多種多様な他のセンサもしくはトランスデューサのいずれかを含み得る（が、これらに限定されない）。

【００６１】

一実施形態では、センサ１２０、１３０は、ネットワーク１２５を介してセンサインターフェイスサーバ１４５に接続される。一実施形態では、センサインターフェイスサーバ１４５は、センサ１２０、１３０から読取り値を収集し、これら読取り値を時系列の観察値として、たとえばデータストア１６０に記憶するためのソフトウェアモジュールなどの論理で構成される。センサインターフェイスサーバ１４５は、たとえば、さまざまなセンサ１２０、１３０に適用可能なセンサデータフォーマットおよび通信プロトコルを用いてセンサからの読取り値を受け入れるように構成された１つ以上のアプリケーションプログラミングインターフェイス（application programming interface：ＡＰＩ）を公開することによって、センサと対話するように構成される。センサデータフォーマットは、概

して、センサデバイスによって規定される。通信プロトコルは、カスタムプロトコル（IoT実現例に先立つ従来のプロトコルなど）、または、さまざまなIoTもしくはマシンツーマシン（machine to machine：M2M）プロトコルのいずれか、たとえば、制約付きアプリケーションプロトコル（Constrained Application Protocol：CoAP）、データ配信サービス（Data Distribution Service：DDS）、ウェブサービスのためのデバイスプロファイル（Devices Profile for Web Services：DPWS）、ハイパーテキスト転送プロトコル／表現状態転送（Hypertext Transport Protocol/Representational State Transfer：HTTP／REST）、MQテレメトリ転送（MQ Telemetry Transport：MQTT）、ユニバーサル・プラグ・アンド・プレイ（Universal Plug and Play：UPnP）、拡張メッセージングおよびプレゼンスプロトコル（Extensible Messaging and Presence Protocol：XMPP）、ZeroMQ、ならびに、伝送制御プロトコルによって伝達することができる他の通信プロトコル、すなわち、インターネットプロトコルまたはユーザデータグラムプロトコル（TCP／IPまたはUDP）転送プロトコルなどであり得る。SCADAプロトコル、たとえば、プロセス制御統合アーキテクチャのためのOLE（OLE for Process Control Unified Architecture：OPCUA）、Modbus RTU、RP-570、Profibus、Conitel、IEC60870-5-101または104、IEC61850、およびDNP3なども、TCP／IPまたはUDPを介して動作するように拡張された場合に採用され得る。一実施形態では、センサインターフェイスサーバ145は、センサ120、130をポーリングしてセンサ読取り値を取出す。一実施形態では、センサインターフェイスサーバは、センサ120、130によって能動的に送信されたセンサ読取り値を受動的に受信する。

【0062】

一実施形態では、企業ネットワーク110は、電力公益事業などの公益事業エンティティに関連付けられ得る。一実施形態では、企業ネットワーク110は、政府等の規制機関エンティティと関連付けられてもよい。説明を簡潔かつ明確にするために、企業ネットワーク110は、オンサイトのローカルエリアネットワーク170によって表わされている。このローカルエリアネットワーク170には、1つ以上のパーソナルコンピュータ175またはサーバ180が、1つ以上のリモートユーザコンピュータ185とともに動作可能に接続されており、当該1つ以上のリモートユーザコンピュータ185は、ネットワーク115または他の好適な通信ネットワークもしくはネットワークの組合わせを通じて企業ネットワーク110に接続されている。パーソナルコンピュータ175およびリモートユーザコンピュータ185は、たとえば、デスクトップコンピュータ、ラップトップコンピュータ、タブレットコンピュータ、スマートフォンであり得るか、または、ローカルエリアネットワーク170もしくはネットワーク115に接続する能力を有するかもしくは他の同期機能を有する他のデバイスであり得る。企業ネットワーク110のコンピュータは、ネットワーク115または別の好適な通信ネットワークまたはネットワークの組合わせを介して時系列データサービス105とインターフェイスする。

【0063】

一実施形態では、リモートコンピューティングシステム（企業ネットワーク110のリモートコンピューティングシステム等）は、ウェブインターフェイスサーバ155を通じて、時系列データサービス105によって提供される情報またはアプリケーションにアクセスし得る。たとえば、企業ネットワーク110のコンピュータ175、180、185は、時系列データ系列データサービス105に対して時系列データベースを要求してもよい。または、たとえば、企業ネットワーク110のコンピュータ175、180、185は、本明細書に記載のシステムおよび方法に従って、時系列データベースの監査を実行してもよい。一実施形態では、リモートコンピューティングシステムは、ウェブインターフェイスサーバ155に要求を送信し得るとともにウェブインターフェイスサーバ155から応答を受信し得る。一例では、情報またはアプリケーションへのアクセスは、パーソナルコンピュータ175またはリモートユーザコンピュータ185上のウェブブラウザを用いることによって達成されてもよい。一例では、これらの通信は、ウェブインターフェイ

10

20

30

40

50

サーバ155とサーバ180との間でやり取りされてもよく、たとえば、データ交換フォーマットとしてJavaScript（登録商標）オブジェクト表記法（JavaScript object notation：JSON）を用いるリモート表現状態転送（REST）要求、または、XMLサーバへの、およびXMLサーバからのシンプルオブジェクトアクセスプロトコル（simple object access protocol：SOAP）要求の形態をとり得る。

【0064】

一実施形態では、データストア160は、センサインターフェイスサーバ145によって受信されるセンサ120、130からの時系列データを記憶および提供するように構成された1つ以上の時系列データベースを含む。一実施形態では、時系列データベースは、時系列データを記憶および提供するように構成されたOracle（登録商標）データベースである。いくつかの例示的な構成では、データストア160は、ネットワーク接続ストレージ（network-attached storage：NAS）デバイスおよび／または他の専用サーバデバイスをを用いて実現され得る。

10

【0065】

一実施形態では、アップストリームデバイス135は、IoT接続デバイスを管理するための第三者サービスであり得る。または、一実施形態では、アップストリームデバイス135は、センサ130がセンサインターフェイスサーバ145と通信することを可能にするように構成されたゲートウェイデバイスであってもよい（したがって、たとえば、センサ130がIoT対応でない場合には、センサインターフェイスサーバ145と直接通信することができない）。

20

【0066】

－MLモデル監査保証のための例示的な方法－

一実施形態では、本明細書に記載の、コンピュータによって実現される方法の各ステップは、1つ以上のコンピューティングデバイスのプロセッサ（図11を参照して図示および説明されるようなプロセッサ1110など）によって実行され得る。当該1つ以上のコンピューティングデバイスのプロセッサは、（i）メモリ（図11を参照して図示および説明されるメモリ1115および／または他のコンピューティングデバイス構成要素など）にアクセスし、（ii）システムに（図11を参照して図示および説明される機械学習監査保証論理1130などの）方法のステップを実行させるための論理で構成されている。たとえば、プロセッサは、本明細書に記載のコンピュータによって実現される方法のステップを実行するために、メモリへのアクセス、メモリからの読出し、またはメモリへの書込みを行なう。これらのステップは、（i）任意の必要な情報を取出すステップと、（ii）任意のデータを計算、判断、生成、分類または作成するステップと、（iii）計算、判断、生成、分類または作成された任意のデータを記憶するステップとを含み得る。ストレージまたは記憶について言及する場合、これは、コンピューティングデバイスのメモリまたはストレージ／ディスク（図11を参照して図示および説明されるメモリ1115、またはコンピューティングデバイス1105のストレージ／ディスク1135、またはリモートコンピュータ1165等）におけるデータ構造としてのストレージを示している。

30

【0067】

一実施形態では、方法の後続ステップは各々、少なくとも後続のステップを開始させるのに必要な程度まで前のステップが実行されたことを示す受信信号または取出し済みの記憶データを解析することに応答して開始される。概して、受信信号または取出し済みの記憶データは、前のステップの完了を示す。

40

【0068】

図2は、機械学習モデルの結果を確実に監査できるようにすることに関連する方法200の一実施形態を示す。一実施形態では、方法200のステップは、（図1を参照して図示および説明されるように）機械学習監査保証システム140または企業ネットワーク110内のコンピュータ175、180、185のいずれかによって実行される。一実施形態では、機械学習監査保証システム140またはコンピュータ175、180、185の

50

いずれかは、機械学習監査保証論理 1 1 3 0 で構成される専用のコンピューティングデバイス（コンピューティングデバイス 1 1 0 5 など）である。

【0 0 6 9】

方法 2 0 0 はさまざまなトリガに基づいて開始されてもよく、これらさまざまなトリガは、たとえば、ネットワーク上の信号の受信、または、記憶データの解析等であって、（i）時系列データサービス 1 0 5 もしくはコンピュータ 1 7 5、1 8 0、1 8 5 のユーザ（もしくは管理者）が方法 2 0 0 を開始したこと、（i i）方法 2 0 0 が定義された時間もしくは時間間隔で開始されるようにスケジュールされること、（i i i）時系列データサービス 1 0 5 またはコンピュータ 1 7 5、1 8 0、1 8 5 の（公益事業もしくは規制機関エンティティに関連付けられた）ユーザが、時系列データベースの監査を要求していたこと、または、（i v）方法 2 0 0 を開始すべきであることを示す他の何らかのトリガ、を示すものである。方法 2 0 0 は、受信信号または取出し済みの記憶データを解析し、方法 2 0 0 を開始すべきであることを当該信号または記憶データが示すと判断することに応答して、開始ブロック 2 0 5 において開始される。処理がプロセスブロック 2 1 0 に進む。

10

【0 0 7 0】

プロセスブロック 2 1 0 において、プロセッサは、監査下のデータベースから元の時系列データ値についての状態推定値のセットを取出す。状態推定値は、時系列データ値の各々ごとに状態推定計算によって生成されたものである。プロセスブロック 2 1 0 における処理が完了し、処理がプロセスブロック 2 1 5 に進む。

20

【0 0 7 1】

プロセスブロック 2 1 5 において、プロセッサは、状態推定値の各々ごとに状態推定計算を反転させて、状態推定値の各々ごとに、再構成された時系列データ値を生成する。プロセスブロック 2 1 5 における処理が完了し、処理がプロセスブロック 2 2 0 に進む。

【0 0 7 2】

プロセスブロック 2 2 0 において、プロセッサは、監査下のデータベースから元の時系列データ値を取出す。プロセスブロック 2 2 0 における処理が完了し、処理がプロセスブロック 2 2 5 に進む。

【0 0 7 3】

プロセスブロック 2 2 5 において、プロセッサは、元の時系列データ値を再構成された時系列データ値と対ごとに比較して、元の時系列と再構成された時系列とが一致するかどうかを判断する。プロセスブロック 2 2 5 における処理が完了し、処理が判断ブロック 2 3 0 に進む。

30

【0 0 7 4】

判断ブロック 2 3 0 において、プロセッサは、元の時系列が再構成された時系列と一致するかどうかを評価する。元の時系列が再構成された時系列と一致する場合（Y E S）、判断ブロック 2 3 0 における処理が完了し、処理がプロセスブロック 2 3 5 に進む。元の時系列が再構成された時系列と一致しない場合（N O）、判断ブロック 2 3 0 における処理が完了し、処理がプロセスブロック 2 4 0 に進む。

【0 0 7 5】

プロセスブロック 2 3 5 において、プロセッサは、元の時系列と再構成された時系列とが一致しているので、監査下のデータベースが修正されていなかったという信号を生成する。プロセスブロック 2 3 5 における処理が完了し、処理が終了ブロック 2 4 5 に進み、プロセス 2 0 0 が終了する。

40

【0 0 7 6】

プロセスブロック 2 4 0 において、プロセッサは、元の時系列と再構成された時系列とが一致しないので、監査下のデータベースが修正されていたという信号を生成する。プロセスブロック 2 4 0 における処理が完了し、処理が終了ブロック 2 4 5 に進み、プロセス 2 0 0 が終了する。

【0 0 7 7】

方法 2 0 0 の前述のプロセスブロックの各々について、本明細書中の他の箇所でさらに

50

詳細に説明する。

【0078】

－MSE T推定値から元の時間系列データを再構成するための準備－

図3は、機械学習モデルの結果を確実に監査できるようにすることに関連する監査可能性プロセス300の一実施形態のフローチャートを示す。MSE Tはデータフロープロセスに組み込まれており、MSE T動作は、監査可能性を求めて元の時系列データを再構成するように反転される。この図は、MSE Tのこの適用例についての監査可能性を示す。

【0079】

一実施形態では、高レベルの場合、監査可能性プロセス300は、システムが時系列データベースからMSE T推定値を取出すときに開始される。次に、当該システムは、MSE T計算を反転させて、再構成された時系列データをMSE T推定値から生成する。上述したように、MSE Tは可逆的なプロセスであり、これは、MSE T推定値に対して反転MSE T計算を実行することにより元の時系列データが再構成されるであろうことを意味する。次いで、当該システムには、監査されているデータベースから元の時系列データが与えられる。最後に、当該システムは、再構成された時系列データを元の時系列データと比較して、当該比較結果が一致を示す場合、元の時系列データに意図的な改ざん（または偶発的なデータ破損）がなかったことを証明する。

【0080】

より詳細に上述したように、監査可能性プロセス300を参照して説明した動作は、図1および図11を参照して図示および説明したメモリ、ストレージおよび／または他のコンピューティングデバイス構成要素にアクセスする1つ以上のコンピューティングデバイスのプロセッサによって実行され得る。

【0081】

一実施形態では、アーカイブされた時系列データベース305は、データベース内の元の時系列データが損なわれていないかどうか、またはデータベース内の元の時系列データが破損しているかもしくは改ざんされているかどうかを判断するために監査用に提示される。アーカイブされた時系列データベース305は、時系列のすべての観測値についての時系列値のシーケンスである1つ以上の時系列信号を含み得る。アーカイブされた時系列データベース305は、アーカイブされた時系列データベース305からの元の時系列データ値と、アーカイブされたデータベースのMSE Tモデリングから導出される再構成された時系列値との間の対差分分析310のために、かつ、これらの値のMSE T推定のために、提示されるだろう。

【0082】

一実施形態では、アーカイブされた時系列データベース305は、データベースに変更が加えられた時のデータベースのスナップショットである。加えて、データベースに変更が加えられた時点で、変更を記述するジャーナルエントリが作成され、状態推定値のセットが、アーカイブされた時系列データベース305の監査の状況下で後の対差分分析310のために作成される。一実施形態では、状態推定値のセットは、監査下のデータベース、すなわち、アーカイブされた時系列データベース305、のためにジャーナルエントリに関連付けて記憶される。たとえば、状態推定値のセットは、監査下のデータベースに対する変更を示す1つ以上のコマンドに応答して生成され、当該変更を記述するジャーナルエントリに関連付けられたデータ構造に記憶されていてもよい。（ジャーナリングされた変更をもたらす）問合せ結果の根拠となるデータを識別する所在メタデータを含む他のメタデータも、ジャーナルエントリに関連付けて記憶され得る。

【0083】

非較正バイアス、断続的な縮退故障、利得変化ドリフト、および一時的なスパイク度を含むセンサ障害は、信号品質を劣化させるとともに、IoT用途に関するML予知の際の偽アラーム（1型エラー）および不発アラーム（II型エラー）の主な原因となる。アーカイブされた時系列データベース305内の元の時系列データ値は、これらのセンサ障害のうちのいくつかを含む可能性がある。一実施形態では、アーカイブされた時系列データ

10

20

30

40

50

ベース305は、一連のインテリジェントデータ前処理315のステップを受けて、データをクレンジングし、クレンジングされたデータをMSETモデリングおよび推定のために準備する。前処理は、センサ障害の不所望な影響を軽減する役割を果たす。一実施形態では、インテリジェントデータ前処理315は、アーカイブされた時系列データベース305に含まれる1つ以上の時系列のために反復される。

【0084】

インテリジェントデータ前処理315の一実施形態では、アーカイブされた時系列データベース305内のすべての信号からのすべての観測値が最初に前処理され、次いで、たとえばOracle（登録商標）分析リサンプリングプロセス（analytical resampling process：ARP）を用いることによって、最適にリサンプリングおよび「調和」されることにより、クレンジングされて最適にリサンプリング／同期された信号についての更新されたデータベースが生成される。一実施形態では、ARPは、発明者K.C.GrossおよびG.C.Wangによって2018年10月23日に出願され「時系列信号を最適に同期させるための自動化された分析リサンプリングプロセス（Automated Analytic Resampling Process for Optimally Synchronizing Time-Series Signals）」と題され、その全体が引用により本明細書に援用されている米国特許出願連続番号第16／168,193号に記載されている技術のうちの1つ以上を含み得る。

【0085】

一実施形態では、アーカイブされた時系列データベース305は、インテリジェントデータ前処理315中に欠落値インピュテーションプロセス320を経る。欠落値インピュテーションプロセス320は、アーカイブされた時系列データベース305内の各時系列信号ごとに、欠落値チェックを用いて元の時系列信号を解析する。時系列における欠落値を解析することに対応して、欠落値インピュテーションプロセス320は、欠落値を推定された値で埋める。一実施形態では、推定された値は単純な補間である。一実施形態では、推定値は、単純な補間ではなく、MSETによって導出された系列相関関係と、既存の値との相互相関関係とに基づく非常に正確な推定値である。更新されて埋められた時系列は将来の処理のために記憶される。欠落値インピュテーションプロセス320はまた、後の反転のために推定値で満たされた時系列における欠落値の位置の記録も記憶し得る。一実施形態では、ARPは、発明者G.C.Wang、K.C.GrossおよびD.Gawlickによって2018年6月11日に出願されて「時系列センサデータの予知分析を促進するための欠落値インピュテーション（Missing Value Imputation to Facilitate Prognostic Analysis of Time-Series Sensor Data）」と題され、その全体が引用により本明細書に援用されている米国特許出願連続番号第16／005,495号に記載されている技術のうちの1つ以上を含み得る。

【0086】

一実施形態では、インテリジェントデータ前処理315は非スパイク化プロセス325を含む。一例では、非スパイク化プロセス325は、元の時系列内の欠落値が欠落値インピュテーションプロセス320によって置換えられた後に行なわれてもよい。非スパイク化プロセス325では、更新された時系列信号は、正確なセンサ読取値を表わさないデータ「スパイク」、突然の短寿命な変動を検出および除去するために外れ値チェックにより解析される。外れ値チェックは、信号についてのさまざまな統計的分布を反復的に特徴付ける（当該統計的分布の特性および挙動を記述するための記述パラメータを生成する）ことによって、信号中のスパイクを検出する。これらの特徴に基づく外れ値である時系列データ値は、データ値スパイクとしてフラグが立てられている。取込まれたスパイクは、一時的に信号平均と置換えられる。更新されて非スパイク化された時系列は、後の処理のために記憶される。非スパイク化プロセス325はまた、検出されたスパイクの時系列内の値および位置の記録を記憶し得る。一実施形態では、非スパイク化プロセス325は、発明者G.C.WangおよびK.C.Grossにより2018年12月10日に出願されて「予知監視用途のためのスパイクとの高忠実度信号の合成（Synthesizing High-Fidelity Signals with Spikes for Prognostic Surveillance Applications）」と題され、その全体が引

用により本明細書に援用されている米国特許出願連続番号第 1 6 / 2 1 5 , 3 4 5 号に記載されている技術のうちの 1 つ以上を含み得る。

【 0 0 8 7 】

一実施形態では、インテリジェントデータ前処理 3 1 5 は、非量子化プロセス 3 3 0 を含む。たとえば、非量子化プロセス 3 3 0 は、データスパイクが検出されて非スパイク化プロセス 3 2 5 によって時系列から除去された後に行なわれてもよい。非量子化プロセス 3 3 0 において、更新された信号は、「量子化」チェックを通じて解析されて、データ量子化（データの間隔が単一の代表値にグループ化されるかまたはまとめられる非可逆データ圧縮技術）により、信号値が、隣り合う代表値間で迅速に繰返し切替えられていたかどうかを判断する。言換えれば、量子化チェックにより、一定数の観察上限の間で観察点が反復的に行ったり来たりして往復している時系列の区間が識別される。データ量子化は、低解像度の「量子化された」トランスデューサまたはセンサによって起こされ得る。さらに、非量子化プロセス 3 3 0 は、時系列に検出された量子化値を高精度の連続信号に変換する。これらの連続信号は、より高い解像度のトランスデューサを用いて検出された場合に信号値が有していたであろうものに非常に近似している。更新済みの量子化されていない時系列は後の処理のために保存される。非量子化プロセス 3 3 0 はまた、検出済みの量子化された値の時系列内の値および位置の記録を記憶し得る。一実施形態では、非量子化プロセス 3 3 0 は、発明者 M.Li および K.C.Gross による 2 0 1 9 年 1 2 月 3 日に特許が付与された「高精度の予知インジケータを生成するための低解像度 I O T 信号の非量子化 (Dequantizing Low-Resolution IOT Signals to Produce High-Accuracy Prognostic Indicators)」と題され、その全体が引用により本明細書に援用されている米国特許第 1 0 , 4 9 6 , 0 8 4 号に記載されている技術のうちの 1 つ以上を含み得る。

【 0 0 8 8 】

一実施形態では、インテリジェントデータ前処理 3 1 5 は非段階化プロセス 3 3 5 を含む。たとえば、非段階化プロセス 3 3 5 は、量子化された値が検出されて非量子化プロセス 3 3 0 によって時系列で置換えられた後に行なわれてもよい。段階化は、記録システムと検出システムとの間のサンプリングレートの不一致に起因するものであり、この場合、より遅いサンプリングレート信号は、単純に、より高いサンプリングレートでそれらの最後の測定値を繰り返すので、測定された全ての信号は結果として均一なサンプリングレートをもたらすこととなる。より遅いサンプリングレートセンサの場合の時系列値は、段階ステップに似た平坦セグメントのシーケンスを有する。段階化は、低サンプリングレートデータをより高いサンプリングレート時系列に収集するための単純なアルゴリズムを用いる商用データ履歴アーカイブに関する共通の問題である。一実施形態では、非段階化プロセス 3 3 5 は、時系列を解析して、段階化された値で存在するものはいずれも識別する。非段階化プロセス 3 3 5 は、信号のうち段階化された部分をより高いサンプリングレートの信号で「埋める」。より高いサンプリングレート信号値は、一実施形態では、M S E T 推定値を用いて導出され得る。更新されて埋められた時系列は将来の処理のために記憶される。非段階化プロセス 3 3 5 はまた、後の反転のために、検出された段階化値についての時系列内の値および位置の記録を記憶し得る。一実施形態では、非段階化プロセス 3 3 5 は、発明者 K.C.Gross および G.C.Wang により 2 0 1 8 年 9 月 1 1 日に出版されて「予知監視動作を容易にするための、時間系列センサ信号内の段階化された値と推論値との置換え (Replacing Stair-Stepped Values in Time-Series Sensor Signals With Inferential Values to Facilitate Prognostic Surveillance Operations)」と題され、その全体が引用により本明細書に援用されている米国特許出願連続番号第 1 6 / 1 2 8 , 0 7 1 号に記載されている技術のうちの 1 つ以上を含み得る。

【 0 0 8 9 】

一実施形態では、インテリジェントデータ前処理 3 1 5 は均一サンプリングプロセス 3 4 0 を含む。一実施形態では、均一サンプリングプロセス 3 4 0 は、段階化された値が検出されて非段階化プロセス 3 3 5 によって時系列内で置換えられた後に行なわれてもよい。均一サンプリングプロセス 3 4 0 では、信号のサンプリングレートが異なっているかど

うかを識別するために、信号同士がサンプリングレートチェックで解析される。信号同士が、異なるサンプリングレートを示す（たとえば、同じ期間にわたって異なる数の観測値を有する）場合、より遅い信号の観測値は、信号の最高サンプリングレートに一致するようにリサンプリングされることとなる。更新されたりサンプリング済みの時系列は、将来の処理のために記憶される。均一サンプリングプロセス340はまた、元のリサンプリングされていない値と時系列内のそれら値の配置とについての記録を後の反転のために記憶し得る。

【0090】

一実施形態では、インテリジェントデータ前処理315は位相同期プロセス345を含む。たとえば、位相同期プロセス345は、均一サンプリングプロセス340の後に行なわれてもよい。または、一実施形態では、位相同期プロセス345は、均一サンプリングプロセス340と並行して行なわれてもよい。位相同期プロセス345において、更新された信号は、位相のずれた観測値を検出するために相関関係チェックにより解析される。位相のずれた観測値（または時系列における誤った時間インデックスに関連付けられた時間信号値）は、たとえば、センサなどの測定機器におけるクロック同期不一致に起因するものであり得る。位相同期プロセス345は、位相がずれた値を時間ドメイン内でシフトして、それらを正しい時間インデックスと整列させる。更新された位相同期済みの時系列は、将来の処理のために記憶される。位相同期プロセス345はまた、後の反転のために、時系列内における位相のずれた値の元の配置の記録を記憶してもよい。

【0091】

一実施形態では、均一サンプリングプロセス340および位相同期プロセスのうちの1つ以上は、Oracle（登録商標）分析リサンプリングプロセス（ARP）を用いて実行され、上記で引用によって援用されている「時系列信号を最適に同期するための自動化された分析リサンプリングプロセス（Automated Analytic Resampling process for Optimally Synchronizing Time-Series Signals）」に記載されている技術のうちの1つ以上を含み得る。信号サンプリングは、アーカイブされた時系列データベース305内の時系列信号間で変化し得ることに留意されたい。たとえば、1つの時系列は、観測値間において高いが規則的であるサンプリングレートまたは間隔を有し得る。別の時系列は、観測値間において低いが規則的であるサンプリングレートまたは間隔を有し得る。別の時系列は、観測値間において不規則または不均等な間隔を有し得る。一実施形態では、時系列信号の位相は、アーカイブされた時系列データベース305内の時系列信号のセットが観測時間に関して整列されるように調整され得る。一実施形態では、時系列信号のデータ値は、時系列信号内の新しいサンプリング間隔での観測値について推定されたデータ値を補間することによって、新しいサンプリング間隔でリサンプリングされ得る。これは、時系列信号のために共通のサンプリング間隔をもたらすために、複数の同期された時系列信号に対して実行されてもよい。

【0092】

上述の6つのインテリジェントデータ前処理315の手順は、結果として、高品質の「クレンジングされた」かまたは「拡張された」バージョンのアーカイブされた時系列データベース305をもたらす。この時系列データベース305は、記憶され得るとともに取出され、さらには、後の機械学習プロセスのために用いられ得る。なお、インテリジェントデータ前処理315において、他のデータ前処理技術も適用され得るか、またはこれら6つのすべての技術よりも少ない数の技術が実行され得ることに留意されたい。すべての信号からのすべての観測値は、ここで前処理され、最適にリサンプリングされてから、ARPを用いて「調和」される。

【0093】

インテリジェントデータ前処理315の手順は、典型的な機械学習データセットにおける共通の問題を補正する。しかし、ステップの各々は、元データのいくつかを修正する。元データを再構築することができるように、各修正ごとに記録が保持されなければならない。上記の6つのインテリジェントデータ前処理315の手順の各々の説明で述べたよう

に、各手順は、データセットに対する変更を示す記録を記憶するだろう。一実施形態では、アーカイブされた時系列データベース305に適用されるすべてのインテリジェントデータ前処理315の手順に対する変更記録は、インテリジェントデータ前処理 (intelligent data preprocessing: IDP) モデルと呼ばれる単一の電子データ構造に記憶されるだろう。

【0094】

ここで図4を参照すると、図4は、例示的なIDPモデル405におけるインテリジェントデータ前処理315からの変更記録を記憶する一実施形態の概略図400を示す。IDPモデル405は、機械学習動作に備えて当該アーカイブされた時系列データベース上で実行されるあらゆる動作についての完全な順序付き記録である。

10

【0095】

欠落値インピュテーションプロセス320は、IDPモデル405における欠落値マークデータ構造410内の推定値で満たされた時系列内の欠落値の位置の記録 (マークまたはマーカとも称される) を記憶する。一実施形態では、欠落値マークデータ構造は、アーカイブされた時系列データベース305内の時系列に関連付けられた配列のセットを含む。アーカイブされた時系列データベース305には、欠落値を含む各時系列に関連付けられた欠落値マーク配列が存在する。欠落値マーク配列は、観察インデックス (たとえば時間) 値の配列であってもよく、各々のインデックス値または時間値は、欠落値が推定された値で埋められていた関連する時系列における観察値を示す。

【0096】

20

非スパイク化プロセス325は、IDPモデル405内のスパイクデータ構造415に、検出されたスパイクの時系列内の値および位置の記録を記憶する。一実施形態では、スパイクデータ構造は、アーカイブされた時系列データベース305内の時系列に関連付けられた配列のセットを含む。スパイクが検出された各時系列に関連付けられたスパイク配列が存在する。これらの配列は、観察インデックス値 (たとえば時間) とその観察値でのスパイクの関連する振幅値 (すなわち、誤ったセンサ読取り値) とを含むタブルの配列であり得る。

【0097】

非量子化プロセス330は、IDPモデル405内の平滑化モデルデータ構造420に、検出された量子化値の時系列内の値および位置の記録を記憶する。一実施形態では、平滑化モデル420は、アーカイブされた時系列データベース305内の時系列に関連付けられた量子化された観察値配列のセットを含む。量子化された値が検出された各時系列に関連付けられた量子化された観察値配列がある。これらの配列は、観察インデックス値 (たとえば時間) とその観察値における関連する量子化値とを含むタブルの配列であり得る。

30

【0098】

非段階化プロセス335は、検出された段階化値の時系列内の値および位置の記録を平滑化モデルデータ構造420に記憶する。一実施形態では、平滑化モデル420は、アーカイブされた時系列データベース305内の時系列に関連付けられた段階化された観察値配列のセットを含む。段階化された値が検出された各時系列に関連付けられた、段階化された観察値配列が存在する。これらの配列は、観察インデックス値 (たとえば時間) およびその観察値での関連する段階化された (元のまたはまだ段階化されていない) 値を含むタブルの配列であり得る。一実施形態では、量子化された観察値配列および段階化された観察値配列は、時系列内のそれらの位置によってインデックス付けされた元の量子化された値および段階化された値のすべてを含む単一の配列である。この状況では、非段階化プロセスは、段階化された値タブルを、非量子化プロセス330によって作成された量子化済みの観察値配列にしてもよい。

40

【0099】

均一サンプリングプロセス340は、IDPモデル405におけるタイムスタンプシーケンスデータ構造425に、時系列内の元のリサンプリングされていない値およびそれらの配置の記録を記憶する。また、位相同期プロセス345は、タイムスタンプシーケンス

50

データ構造 4 2 5 に、時系列内の位相がずれた値の元の配置の記録を記憶する。一実施形態では、タイムスタンプシーケンスデータ構造 4 2 0 は、アーカイブされた時系列データベース 3 0 5 における当該時系列に関連付けられたタイムスタンプ配列のセットを含む。位相がずれた観察値が検出された各時系列に関連付けられたタイムスタンプ配列観察値配列がある。これらの配列は、時系列の各観察値ごとに観察インデックス値（たとえば時間）を含む。一実施形態では、均一サンプリングプロセス 3 4 0 および位相同期プロセス 3 4 5 の両方のために作成されるタイムスタンプ配列がある。これは、プロセスが連続して実行される場合に該当し得る。均一サンプリングプロセス 3 4 0 および位相同期プロセス 3 4 5 が一緒に実行される場合、その結果は単一のタイムスタンプ配列と一緒に記憶され得る。

10

【0100】

したがって、一実施形態では、プロセッサは、インテリジェントデータ前処理 3 1 5 に示されるように、元の時系列値を前処理して、元の時系列値の品質に対するセンサ障害の影響を軽減する。そのインテリジェントデータ前処理の過程にわたって、プロセッサは、たとえば I D P モデル 4 0 5 によって示されるように、前処理中に元の時系列値に対する 1 つ以上の変更を記録するデータ前処理モデルを生成する。一実施形態では、図 2 のプロセスブロック 2 1 5 を参照して説明した状態推定計算の反転はまた、ストレージまたはメモリから I D P モデル 4 0 5 などのデータ前処理モデルを取出すステップと、状態推定値の各々ごとに、データ前処理モデルによって記述される前処理を反転させるステップとを含む。

20

【0101】

上述した 6 つのインテリジェントデータ前処理 3 1 5 手順の各々は、I D P モデル 4 0 5 に保持される元の時系列値を取出し、拡張された時系列データベース内の対応する拡張値を元の値と置換えることによって、反転または取消しがなされてもよい。一実施形態では、インテリジェントデータ前処理 1 3 5 のさまざまなプロセスのための反転プロセスは、インテリジェントデータ前処理 1 3 5 プロセスが実行された順序とは逆の順序で実行される。

【0102】

ーモデルトレーニングー

再び図 3 を参照すると、インテリジェントデータ前処理 3 1 5 の後、M S E T モデル 3 5 0 をトレーニングするために、拡張された時系列信号が用いられる。図 5 は、機械学習モデルの結果を確実に監査できるようにすることに関連する M S E T モデルトレーニングプロセス 5 0 0 の一実施形態の概略図を示す。アーカイブされた時系列データベース 3 0 5 は、選択トレーニングデータプロセス 5 0 5 に提供される。選択トレーニングデータプロセス 5 0 5 では、時系列データからデータを選択してトレーニングデータセットを形成する。一実施形態では、トレーニングセットを形成するために時系列の観察ベクトルのサブセットが選択される。次いで、選択されたトレーニングデータセットはトレーニングデータ構造 5 1 0 に記憶される。

30

【0103】

トレーニングデータ 5 0 5 の選択が完了し、トレーニングデータ構造 5 1 0 に記憶されると、図 3 および図 4 を参照して上述したように、インテリジェントデータ前処理 3 1 5 が開始され、I D P モデル 4 0 5 が作成されて記憶される。一実施形態では、インテリジェントデータ前処理 3 1 5 の前にトレーニングデータ 5 0 5 が選択され、トレーニングデータは未拡張データから選択される。別の実施形態では、インテリジェントデータ前処理 3 1 5 の後にトレーニングデータ 5 0 5 が選択され、トレーニングデータは拡張されたデータから選択される。

40

【0104】

トレーニングデータ 5 0 5 の選択およびインテリジェントデータ前処理 3 1 5 が完了すると、トレーニングベクトル選択プロセス 5 1 5 および M S E T モデルトレーニングプロセス 5 2 0 のループが開始される。ベクトル選択の際に、トレーニングデータからトレー

50

ニングベクトルのセットが選択され、トレーニング用にMSE Tモデルに提供される。たとえば、50個または100個のベクトルが選択され得る。一実施形態では、前処理されたトレーニングデータは、モデルを改善するために2つの部分に分割される。たとえば、偶数の観測値がトレーニングセットの第1の部分形成し得るとともに、奇数の観測値がトレーニングセットの第2の部分形成し得る。選択プロセスの反復中に、トレーニングベクトルのセットは、偶数／奇数の「ホップスコッチ (hopscotch)」ベクトル選択の際に奇数の観測値から選択され、次いで、偶数の観測値から選択される。

【0105】

上述のように、MSE Tモデル350などのMSE Tモデルをトレーニングすることは、決定論的な数学的手順である。MSE Tトレーニング520は、アーカイブされた時系列データベースに記録された時系列を表わす時系列信号を用いて、時系列信号間の相関関係を学習する。MSE Tモデルトレーニングプロセス520は、拡張された信号を用いて、アーカイブされた時系列データベース305内の他の任意の信号との任意の程度の関連性を有するすべての信号を当該アーカイブされた時系列データベース305内において識別する。一実施形態では、信号間の関連性の識別は、ともに、全体として当該アーカイブされた時系列データベース305内の信号の全域に対して実行されるとともに、信号のクラスタに対しても別個に実行される。この経験的クラスタ化アプローチは、信号のアーカイブされた時系列データベース305が、公益事業エンティティの施設内の別個のシステム、または、公益事業デバイスおよびシステムの公益事業エンティティのフリート内の個別の資産に由来し得ることを認識している。MSE Tトレーニングプロセス520が最終的に出力するものは、トレーニングされたMSE Tモデル350である。一実施形態では、トレーニングプロセスはMSE T2トレーニングプロセスであり、MSE TモデルはMSE T2モデルである。

【0106】

一実施形態では、ベクトル選択515およびMSE Tトレーニングプロセス520は、判断ブロック525に示すように、MSE Tモデル350が安定するまで（1つ以上の反復の）ループ内で繰返され得る。トレーニングデータセットに対する変更ほどにはMSE Tモデル350がさほど変化しない場合、モデル350は安定しており（YES）、トレーニングは完了し、トレーニングされたMSE Tモデル350は、さらなる使用のためにデータ構造として記憶される。トレーニングデータセットを変更したときにMSE Tモデル350が依然として著しく変化する場合、モデル350はまだ安定しておらず（NO）、515で、トレーニングベクトルの追加のセットが選択され、520で、MSE Tモデル350がトレーニングベクトルの追加のセットでさらにトレーニングされる。

【0107】

再び図3を参照すると、トレーニングされたMSE Tモデル350は、アーカイブされた時系列データベース305内の各信号と他の信号との間のモデルトレーニング520中に学習された経験的な相関関係パターンに基づいて、アーカイブされた時系列データベース305内の各信号のMSE T状態推定値355を計算するために用いられる。MSE T推定値は非常に正確であるものの、その正確度はMSE Tモデルのトレーニングの程度に基づいて異なっている可能性がある。

【0108】

プロセスブロック360において、MSE T推定値、MSE Tパラメータ、およびトレーニングベクトルは報告データ構造に保存される。この報告データ構造は、データベースまたはデータ構造において、アーカイブされた時系列データベース305に対する変更を記述するジャーナルエントリと関連付けられてもよい（たとえば、アーカイブされた時系列データベースのアーカイブされた状態をもたらした変更）。一実施形態では、プロセスブロック320から355は、アーカイブされた時系列データベース305におけるデータ改ざんを確実に検出可能にするために、変更を記述するジャーナルエントリとともに含まれるようにするために、アーカイブされた時系列データベースに対する変更に応答して開始されて実行される。一実施形態では、変更に先行する1つ以上の問合せを記述するア

クティビティジャーナルメタデータと、問合せ結果の基礎をなすデータを記述する所在メタデータとが、ジャーナルエントリおよび報告データ構造とともに記憶される。

【0109】

なお、報告が監査下のデータベースに対するいずれかの変更に応答して生成され得ることで、データベースに対する各変更に関連する一続きの改ざん防止データセットが作成され得ることに留意されたい。このため、履歴所在情報（変更、当該変更に先行したアクティビティを記述するアクティビティジャーナル、および当該変更に関連付けられた問合せに応答して提示される情報の基礎を記述する所在情報を含む）と、元データを再構築するための記録とがともに、各ジャーナルエントリに記録される。したがって、監査下のデータベースの完全な履歴が記録されるとともに、監査のためにレビューすることができる。さらに、当該記録は、データベースの追加のスナップショットを維持する場合よりもはるかにコンパクトであり、結果として、性能が大幅に改善されるとともに、監査のために用いられるデータの可搬性が改善されることとなる。

10

【0110】

一実施形態では、MSET推定値355は、元の未加工信号（MSETパラメータ）とともに、アーカイブされた時系列データベース305に記憶される。また、アーカイブされた時系列データベース305には、十分に検証された場合には「1」とともに、または、元の未加工信号を計測したセンサにおいて異常が発見された場合の信号については「0」とともに、「センサ操作性フラグ」が記憶される。一実施形態では、これらのセンサ操作性フラグは、MSET推定値と元の未加工信号値との間の残差を分析し、その値を異常または異常なしとしてラベル付けするために、逐次確率比検定（sequential probability ratio test：SPRT）を用いて故障検出推定の結果に基づいて決定される。特定のセンサについての時系列データに閾値数の異常値が生じる場合、当該センサは、時系列において部分的または完全に動作不能（「0」）としてフラグが立てられ得る。動作不能なフラグは、センサ／トランスデューサの故障もしくは劣化、縮退故障、またはセンサ／トランスデューサもしくは上流のデータ収集電子機器もしくはネットワークの断続的な問題からの信号を示し得る。特定のセンサについて時系列データに生じる異常値の数が閾値数を超えない場合、センサは、時系列において動作可能（「1」）であると検証されたものとしてフラグが立てられてもよい。一実施形態では、閾値は、時系列における1つの異常な読取り値を満たすかまたは超える程度に低くてもよく、または必要に応じて高くてもよい。一実施形態では、閾値は、MSET分析などの時系列の機械学習分析によって設定され得る。一実施形態では、これらのフラグは、時系列を含むデータ構造において設定され得る。

20

30

【0111】

ーモデル制限ー

いくつかの例では、トレーニングされたMSETモデル350は、さほどコンパクトでない可能性もある。このため、可搬性を求めて、MSETモデル350のサイズを制限することが望ましい場合がある。ここで図6を参照すると、図6は、機械学習モデルの結果を確実に監査できるようにすることに関連するMSETモデル制限プロセス600の一実施形態の概略図を示す。

【0112】

一実施形態では、当該システムは、最初に、たとえば、図5を参照して図示および説明されるように、元の時系列値から選択されたトレーニング値のセットで、MSETモデル350等の多変量状態推定モデルをトレーニングする。トレーニングされた多変量状態推定モデルの複雑さは、トレーニングされた多変量状態推定モデルのための行列の主成分分析と、トレーニングされた多変量状態推定モデルを行列の主成分に限定することによって、低減され得る。たとえば、MSETモデル350は、モデル350のMSET行列の主成分分析と、モデル350を行列の主成分に限定することによって低減され得る。なお、MSETモデルは、単にMSET行列だけからなるものではなく、MSET行列はMSETアルゴリズムの中心をなすことに留意されたい。上述のように、MSETモデル350はMSET2モデルであり得る。

40

50

【0113】

一実施形態では、当該システムは、多変量状態推定モジュールに関連付けられた行列を固有ベクトルのセットに分解する。たとえば、MSETモデル350の複雑さは、MSETモデル350に含まれるMSET行列の特異値分解(singular value decomposition: SVD)605を実行することによって低減される。一実施形態では、SVD605は、固有ベクトルのシーケンスおよびそれらの関連する固有値からなる。一実施形態では、SVD605から結果として得られる固有ベクトルは、それぞれの関連する固有値によって減少順でソートされる。次いで、当該システムは、ソートされた固有ベクトルのシーケンスおよびそれらの関連する固有値610を、たとえばさらなる処理のためにSVDデータ構造として記憶する。

10

【0114】

一実施形態では、当該システムは、SVDデータ構造に記憶された固有ベクトルのセットから主固有ベクトルのサブセットを選択する。たとえば、当該システムは、ソートされた固有ベクトルおよび固有値610に基づいて、SVD605から主固有ベクトル615を選択する。最大の固有値を有する固有ベクトルは主固有ベクトルである。これらの主固有ベクトルは、時系列にわたるセンサデータの最大のばらつきの主要因であり、したがって、状態推定モデリングにとって最も有益である。一実施形態では、プロセッサは、最大の関連する固有値を有する「上位」N個の固有ベクトルを、関連する主固有値を有する主固有ベクトル620として選択する。次いで、プロセッサは、上位N個の固有ベクトルを除いて、SVDデータ構造内のすべての固有ベクトルおよび固有値を削除することで、主固有ベクトルではないすべての固有ベクトルを（それらの関連する固有値とともに）除去する。

20

【0115】

一実施形態では、当該システムは、主固有ベクトルのサブセットから「制限された」多変量状態推定モデルを作成する。たとえば、保持された上位N個の主固有ベクトルおよびそれらの関連する主固有値620が、MSETモデルリミッタ625に提供される。MSETモデル350も、MSETモデルリミッタ625に提供される。MSETモジュールリミッタ625は、MSETモデル350を上位N個の主固有ベクトルおよびそれらの固有値620に制限するように動作する。これにより、MSETアルゴリズムの結果を符号化するのに必要なデータの量が実質的に低減される。一実施形態では、MSETモデルリミッタ625は、制限されたMSET行列を主固有ベクトルおよび関連する固有値620から構築する。モデルリミッタ625は、MSETモデル350内の元のMSET行列を当該制限されたMSET行列と置換えて、制限されたMSETモデル630を作成する。制限されたMSETモデル630は、メモリまたはストレージにデータ構造として記憶される。

30

【0116】

再び図2を参照すると、一実施形態では、状態推定計算の反転はまた、制限された多変量状態推定モデル（制限されたMSETモデル630など）が状態推定を形成する計算の反転を生成することを含む。

【0117】

ーデータ圧縮ー

いくつかの例では、アーカイブされた時系列データベース305全体は非常に大きい可能性がある。したがって、可搬性を得るために、データベースのサイズを当該制限されたMSETモデル630のパラメータに制限することが望ましい場合もある。図7は、機械学習モデルの結果を確実に監査できるようにすることに関連するデータ圧縮プロセス700の一実施形態の概略図を示す。データ圧縮プロセス700において、アーカイブされた時系列データベース305は、未圧縮のまたは元のアーカイブされた時系列データベース305の監査に適した最小サイズにまで縮小される。

40

【0118】

一実施形態では、プロセスブロック705において、当該システムは、監査下のデータ

50

ベース、すなわちアーカイブされた時系列データベース305、内の元の時系列値からトレーニングデータ値を除外する。たとえば、一実施形態では、当該システムは、MSETモデルをトレーニングするために用いられるトレーニングデータを含まないアーカイブされた時系列データベース305のコピーを作成する。一実施形態では、MSETモデルのためのトレーニングデータ、たとえばMSETモデル350のためのトレーニングデータ510などは、アーカイブされた時系列データベース305のコピーから除去される。たとえば、MSETモデル350をトレーニングするために用いられるすべての観察記録は、アーカイブされた時系列データベース305のコピーから削除され得る。別の実施形態では、アーカイブされた時系列データベース305のコピーは、トレーニングデータなしで最初に作成されているため、それを削除する必要がなくなる。当該システムは、後続の処理のために当該アーカイブされた時系列データベースの縮小コピーを記憶する。

10

【0119】

一実施形態では、プロセスブロック707において、当該システムは、残りの元の時系列データ値を前処理して、センサ障害の影響を軽減するとともに、残りの元の時系列データに対する変化を記録するモデルを生成する。たとえば、(上述したような)インテリジェントデータ前処理315は、アーカイブされた時系列データベース305の縮小コピーに対して実行されてもよい。当該システムは、IDPモデル405と同様のIDP動作モデル710を形成し、この場合、IDP動作モデル710は、トレーニング観察値を除外した後に当該アーカイブされた時系列データベース305の縮小コピーに残っている残りの観察記録に対するもののみである。当該システムは、後続の処理のために、アーカイブされた時系列データベースの前処理された縮小コピーを記憶する。

20

【0120】

一実施形態では、MSET動作ブロック715において、当該システムは、圧縮された時系列データベースを作成するために、制限された多変量状態推定モデルを用いて、前処理された残りの元の時系列データ値のために状態推定(MSET状態推定など)を実行する。制限されたMSETモデル630と、トレーニングデータが除外されたアーカイブされた時系列データベース305のコピーとは、MSET動作715を実行するために用いられる。一実施形態では、上述したように、MSET動作715および制限されたMSETモデル630は、MSET2動作およびMSET2モデルである。MSET動作715は、(i)残りのパラメータに基づいて各観察値ごとに状態推定値を形成し、(ii)制限されたMSETモデル630のパラメータではない時系列変数を、アーカイブされた時系列データベースのコピーから除去する。こうして、MSET状態推定を知らせないデータ値(MSET状態推定の値に有意な影響を及ぼさないデータ値)が、アーカイブされた時系列データベースのコピーから削除されることで、データのサイズがさらに低減される。各観察値ごとに形成されたMSET状態推定値および各観察値ごとの残りのパラメータ値は、圧縮されたデータ720を形成する。圧縮されたデータ720はデータ構造としてメモリまたはストレージに記憶される。

30

【0121】

こうして、一実施形態では、アーカイブされた時系列データベースが前処理されると、制限されたMSETモデル630を用いてMSETアルゴリズムが実行される。制限されたMSETモデル630は、制限されたMSETモデル630についてのMSETパラメータとともに、元の時系列データベースの圧縮バージョン(圧縮されたデータ720)を構成する。MSET推定値がMSETモデルとともに記憶されて、元の時系列データを縮小サイズで表わす。元の未加工データストリームのいずれかが修正されたか、変更されたかまたは置換されたかを識別する必要がある場合、当該システムは、記憶されたMSET推定値およびMSETモデルを反転させて、元の時系列データを再構成することができるであろうし、さらに、これを用いて、以下に論じるように、意図される元の時系列データを有効化または無効化することができる。

40

【0122】

ーデータフォーマットー

50

図 8 A および図 8 B は、2つの例示的なデータ報告フォーマットを示す。これらの2つのフォーマットの各々は、元のアーカイブされた時系列データベース 305 などの監査下のデータベースの監査を実行するのに十分な情報を含む。一実施形態では、当該システムは、2つの例示的なデータ報告フォーマットのうちの1つに従って電子データ報告データ構造を生成する。

【0123】

一実施形態では、当該システムは、電子データ報告データ構造を生成する。当該電子データ報告データ構造は、元の時系列値の品質に対するセンサ障害の影響を軽減するために、前処理中に元の時系列値に対する1つ以上の変更を記録する前処理モデルと、圧縮された時系列データベースとを含み、当該圧縮された時系列データベースは、元の時系列値から選択されたトレーニング値のセットでトレーニングされた制限された多変量状態推定モデルを用いて状態推定を実行すること、および、制限された多変量状態推定モデルのパラメータではない値を圧縮された時系列データベースから排除すること、によって生成されるものであり、当該電子データ報告データ構造はさらに、(i) トレーニング値のセットおよび(ii) 制限された多変量状態推定モデルのうちの1つ以上を含む。

【0124】

2つのフォーマットを比較すると、第1のデータ報告フォーマット 800 は、(制限された MSET2 モデル 630 ではなく) すべてのトレーニングデータ 510 を指定する(すなわち、データ報告に含める)一方で、第2のデータ報告フォーマット 850 は、(トレーニングデータ 510 ではなく) 制限された MSET2 モデル 630 を指定する。第1のデータ報告フォーマット 800 および第2のデータ報告フォーマット 850 の両方のフォーマットは、IDP 動作モデル 710 および圧縮されたデータ 720 を指定する。

【0125】

したがって、一実施形態では、電子データ報告データ構造は、(図3、図4、および図7を参照して図示および説明されるように) 元の時系列値の品質に対するセンサ障害の影響を軽減するために、前処理中に元の時系列値に対する1つ以上の変更を記録する前処理モデル(IDP 動作モデル 710)を含む。電子データ報告データ構造はまた、圧縮された時系列データベース(圧縮されたデータ 720)を含む。当該圧縮された時系列データベースは、制限された多変量状態推定モデルで状態推定を実行すること、ならびに、(図7を参照して図示および説明されるように) 当該制限された多変量状態推定モデルのパラメータではない値を当該圧縮された時系列データベースから除外すること、によって生成される。電子データ報告データ構造はまた、(i) 元の時系列値から選択されるとともに第1のデータ報告フォーマット 800 などで当該制限された多変量状態推定モデルをトレーニングするために用いられるトレーニング値のセットと、(ii) 第2のデータ報告フォーマット 850 などでの当該制限された多変量状態推定モデルと、のうちの1つ以上を含む。

【0126】

第1のデータ報告フォーマット 800 の1つの例示的な利点は、MSET (または MSET2) 監査可能性アルゴリズムの詳細が明らかにされないことである。第2のデータ報告フォーマット (850) の1つの例示的な利点は、報告で提供されなければならないデータの量が低減されることである。なお、圧縮されたデータ 720 の量または体積は、概して、第1のデータ報告フォーマット 800 および第2のデータ報告フォーマット 850 の両方での報告のサイズに影響を及ぼし得るものであり、このため、実際には、2つのフォーマットはサイズに有意な差がない場合もあることに留意されたい。

【0127】

再び図3を参照すると、一実施形態では、当該システムは、プロセスブロック 365 に示すように、未加工データを再構成するために MSET 計算を反転させる。たとえば、データ報告フォーマット 800、850 のうちの1つに従ってデータ報告が作成され、アーカイブされた時系列データベース 305 の監査のためのデータを再構築するために用いられ得る。なお、これらの報告は、たとえば、アーカイブされた時系列データベース 305

の監査の状況下で評価を待つためにジャーナルエントリとともに記憶されることに留意されたい。いくつかの実施形態では、これらの報告は、未加工データを再構成するために用いられる前に、有意な期間にわたって記憶されてもよい。いくつかの実施形態では、アーカイブされた時系列データベース305の監査は行なわれなくてもよく、かつ、データ報告は、取出されることはなく、プロセスブロック365において未加工データを再構成するために用いられる。また、当該報告は、アーカイブされた時系列データベースの第三者のコピーの外部監査のために第三者に配布されてもよい。

【0128】

－例示的な第1のデータ報告フォーマットを用いたデータ再構築－

図9は、第1のデータ報告フォーマット800を用いるデータ再構築プロセス900の一実施形態の概略図を示しており、データ再構築プロセス900はさらに、機械学習モデルの結果を確実に監査できるようにすることに関連付けられている。

【0129】

一実施形態では、データ再構築プロセス900において、制限されたMSET（またはMSET2）モデル630は、プロセスブロック905に示すように、（たとえば、図5を参照して図示および説明されるように）IDPモデル405およびトレーニングデータ510を用いて最初にトレーニングされ、次いで（たとえば、図6を参照して図示および説明されるように）制限される。IDPモデル405は、ストレージから取出され、トレーニングおよび制限プロセス905に提供される。トレーニングデータ510は、データ報告800から読出される。第1のデータ報告フォーマット800からの制限されたMSETモデル630が無いために、制限されたMSETモデル630の作成が必要となる。

【0130】

一実施形態では、当該システムは、元の時系列データ値から状態推定値を形成するために状態推定計算によって実行されるステップを取消す反転状態推定計算を識別する。たとえば、プロセスブロック910に示すように、MSET（またはMSET2）アルゴリズムが反転され、圧縮されたデータ720に適用される。MSETアルゴリズムを反転させるために、トレーニングされたモデル630によって実行されるMSET推定値を形成するステップが解析され、離散演算のシーケンスが記録される。シーケンス内の離散演算の各々ごとに、離散演算を取消すこととなる反転演算が識別され、反転演算のシーケンスに記録される。一実施形態では、反転演算のシーケンスは、離散演算のシーケンスとは逆の順序となるべきである。たとえば、離散演算が離散演算のシーケンスにおいて最初に実行されるべきである場合、離散演算の反転演算は、反転演算のシーケンスにおいて最後に実行されるべきであり、第2の離散演算の反転演算は、反転演算のシーケンスにおいて最後から2番目に実行されるべきであり、以下同様に実行され、これにより、離散演算の反転演算は、離散演算とは逆の順序で実行されることとなる。当該システムは、後続の処理のために反転演算のシーケンスを記憶する。

【0131】

一実施形態では、当該システムは、次いで、状態推定値のセットから元の時系列データについての反転状態推定値のセットを生成する。反転状態推定値の各々は、たとえば反転演算のシーケンスによって示されるように、状態推定値のセットのうちの1つについて反転状態推定を実行することによって生成される。たとえば、当該システムは、圧縮されたデータ720に記憶された状態推定値を作成するために用いられるMSET計算の反転を、それらの状態推定値に対して実行する。各観察値ごとの元の時系列データ値の反転MSET推定値は、それらの値の状態推定値と、圧縮されたデータ720に記憶された他のパラメータの観察値とから作成される。反転演算のシーケンスは、各観察値ごとの推定されたデータ値に対して実行され、推定されたデータ値の各々ごとに反転MSET推定値を作成する。

【0132】

一実施形態では、プロセスブロック915に示すように、インテリジェントデータ前処理315も反転させて、反転MSET推定値に適用される。一実施形態では、IDP動作

モデル710がデータ報告800から読出され、IDP動作モデル710に保持される元の時系列値は反転MSET推定値内の任意の対応する値と置換され、これにより、圧縮データのためのインテリジェントデータ前処理315を反転させて、再構成データ、近似する元データ920を形成する。再構成されたデータ920は、たとえば、図3のブロック310および370～380を参照して図示および説明される検証プロセスを含む、監査プロセスにおいて後で用いるために記憶される。このため、状態推定値の各々についての再構成された時系列データ値は、反転状態推定値に基づいている。

【0133】

なお、結果として得られる再構成された時系列データは、アーカイブされた時系列データベース305内の元データと近似しており、元データとはわずかに異なってもよい。それにもかかわらず、近似は、アーカイブされた時系列データベース305が改ざんされていないことまたは破損していないことを検証するために用いられ得る。このため、この近似を用いて、規制の遵守または規制の違反を示すアーカイブされた時系列データベース305が正確にこのような遵守または違反を示すことを検証することができる。

【0134】

一例示的な第2のデータ報告フォーマットを用いたデータ再構築—

図10は、第2のデータ報告フォーマット850を用いるデータ再構築プロセス1000の一実施形態の概略図を示し、データ再構築プロセス1000はさらに、機械学習モデルの結果を確実に監査できるようにすることに関連付けられている。再構築プロセス1000は、制限されたMSETモデル630が第2のデータ報告フォーマット850にすでに記憶されているので、最初に計算される必要がないことを除いて、プロセス900と概して同じプロセスステップに従う。なお、一実施形態では、インテリジェントデータ前処理315の反転1010およびMSET反転1015を実行する順序は、プロセス1000において、プロセス900におけるインテリジェントデータ前処理315の反転915に先行してMSET反転910が行なわれるという順序から切替えられてもよいことに留意されたい。

【0135】

一実施形態では、プロセスブロック1010に示すように、インテリジェントデータ前処理315がまず反転され、圧縮されたデータ720内のMSET推定値に適用される。一実施形態では、IDP動作モデル710がデータ報告800から読出され、IDP動作モデル710内に保持される元の時系列値が、圧縮されたデータにおけるMSET推定値内の任意の対応する値と置換され、これにより、圧縮されたデータ720のためにインテリジェントデータ前処理315を反転させる。こうして、状態推定値の各々についての再構成された時系列データ値は、プロセスブロック1015を参照して説明したような反転状態推定値に基づいている。

【0136】

一実施形態では、MSET（またはMSET2）アルゴリズムは、上記のプロセスブロック910を参照して説明したのと同様の態様で、プロセスブロック1015に示すように、反転され、圧縮されたデータ720に適用される。たとえば、当該システムは、圧縮されたデータ720に記憶された状態推定値を作成するために用いられるMSET計算の反転を、それらの状態推定値に対して実行する。各観察値ごとの元の時系列データ値の反転MSET推定値は、それらの値の状態推定値と、圧縮されたデータ720に記憶された他のパラメータの観察値とから作成される。反転MSET推定値とプロセスブロック1010からの置換されたインテリジェントデータ前処理値とは、再構成されたデータ、すなわち近似する元データ1020を形成する。再構成されたデータ1020は、たとえば、図3のブロック310および370～380を参照して図示および説明される検証プロセスを含み得る監査プロセスにおいて後で使用するために記憶される。上記のプロセス900と同様に、結果として得られる時系列データは、アーカイブされた時系列データベース305内の元データと近似しており、元データとはわずかに異なる可能性があるが、アーカイブされた時系列データベース305が改ざんされていないことまたは破損していない

10

20

30

40

50

ことを検証するために用いられ得る。

【0137】

－改ざん報告－

「元の」時系列データに故意の改ざん（または偶発的なデータ破損）がなかった（比較により、統計的意味での一致が示される）場合、当該システムは、元データが損なわれていないことを証明し、それ以外の場合、元データは破損している。

【0138】

再び図3を参照すると、当該システムは、次に進んで、アーカイブされた時系列データベース305内の元の時系列の（損なわれていないか、破損している／改ざんされている）状態の監査を実行する。一実施形態では、この監査の状況下では、報告された時系列の正確さは、再構成されたデータを用いて検証される。元データは、再構成された（または再構築された）データ、たとえば、プロセス900から得られる再構成されたデータ920またはプロセス1000から得られる再構成されたデータ1020、と比較される。2つのデータストリーム、元の時系列および再構成された時系列は、プロセスブロック310に示されるように対差分分析器で対ごとに比較される。対差分分析器310は、元の時系列データ値と再構成された時系列データ値とを2つの時系列の各観察値ごとに比較して、再構成された値が元の値と一致する（かまたは閾値の範囲内で密に近似している）かどうかを確認する。

【0139】

一実施形態では、対差分分析器は、元の時系列データ値と再構成された時系列データ値とを各観察値ごとに比較して、それが事前設定された閾値量、たとえばパーセンテージ量を超えて変化するかどうかを判断する。別の実施形態では、対差分分析器は、各々の元の時系列データ値と再構成された時系列データ値とを各観察値ごとに比較して、トレーニング済みの制限されたMSE T2モデル630に含まれるトレーニング済みの故障検出モデルを用いて故障検出をトリガするかどうかを判断する。故障検出モデルは、逐次確率比検定（SPRT）を採用して、各観察値ごとに元の時系列データ値と再構成された時系列データ値との間の残差を分析することで、意図された元の時系列データが異常であるか否かを判断し得る。

【0140】

対ごとの比較の結果は判断ブロック370で評価される。2つのデータストリームが一致する場合（YES）、検証は合格とされ、処理がプロセスブロック375に進む。2つのデータストリームが一致しない場合（NO）、検証は不合格であり、処理はプロセスブロック380に進む。

【0141】

プロセスブロック375において、元の時系列データが損なわれていないことを示す「合格した」検証報告が生成される。一実施形態では、合格した検証報告は、解析時に、元の時系列データが損なわれていないことを示す信号である。一実施形態では、合格した検証報告は、元の時系列データが損なわれていないことを示す、人が読取り可能な文書である。一実施形態では、当該システムは、合格した検証報告に応答して、元の時系列データが損なわれていないという指摘をグラフィカルユーザインターフェイス（graphical user interface：GUI）上に表示する命令を生成して、実行するかまたは実行のために送信する。GUIは、公益事業エンティティまたは規制機関エンティティに関連付けられてもよい。次いで、プロセス300における処理が終了する。

【0142】

プロセスブロック380において、元の時系列データが破損しているかまたは改ざんされていることを示す「不合格」検証報告が生成される。一実施形態では、不合格検証報告は、解析時に、元の時系列データが破損しているかまたは改ざんされていることを示す信号である。一実施形態では、不合格検証報告は、元の時系列データが破損しているかまたは改ざんされていることを示す、人が読取り可能な文書である。一実施形態では、当該システムは、不合格検証報告に応答して、元の時系列データが破損しているかまたは改ざん

されているという指摘をGUI上に表示する命令を生成し、実行するかまたは実行のために送信する。GUIは、公益事業エンティティまたは規制機関エンティティに関連付けられてもよい。次いで、プロセス300における処理が終了する。

【0143】

こうして、当該システムは、監査下のデータベースが修正されていなかったという信号に応答して、検証中のデータベースが破損しておらず改ざんもされていないと証明されていることを示す電子検証報告メッセージを生成する。当該システムは、監査下のデータベースが修正されていたという信号に応答して、監査下のデータベースが破損しているかまたは改ざんされていることを示す電子検証報告メッセージを生成する。次いで、当該システムは、生成された電子検証報告をコンピューティングデバイスに送信して、検証報告メッセージをコンピューティングデバイスによって記憶させ得るかまたはコンピューティングデバイスによって表示させ得る。一実施形態では、コンピューティングデバイスは、公益事業エンティティまたは規制機関エンティティに関連付けられてもよい。

10

【0144】

ーソフトウェアモジュールの実施形態ー

一般に、ソフトウェア命令は、適切にプログラムされたプロセッサによって実行されるように設計される。これらのソフトウェア命令は、たとえば、コンピュータ実行可能コードと、当該コンピュータ実行可能コードにコンパイルされ得るソースコードとを含み得る。これらのソフトウェア命令はまた、スクリプト言語などの解釈されたプログラミング言語で書かれた命令を含み得る。

20

【0145】

複雑なシステムでは、このような命令は、典型的には、プログラムモジュールに配列され、このようなモジュールは各々、特定のタスク、プロセス、機能、または動作を実行する。モジュールのセット全体は、オペレーティングシステム（operating system：OS）または他の形態の組織化プラットフォームによって、それらの動作時に制御または調整され得る。

【0146】

一実施形態では、本明細書に記載の構成要素、機能、方法、またはプロセスのうちの1つ以上は、非一時的なコンピュータ可読媒体に記憶されたモジュールとして構成される。これらのモジュールは、メモリまたはストレージにアクセスする少なくともプロセッサによって実行されると、コンピューティングデバイスに、本明細書に記載の対応する機能を実行させる記憶されたソフトウェア命令で構成される。

30

【0147】

ークラウドシステム、マルチテナント、および企業の実施形態ー

一実施形態では、本システムは、企業組織のためのアプリケーションまたは分散アプリケーションの集合を含むコンピューティング／データ処理システムである。アプリケーションおよびコンピューティングシステムは、クラウドベースのネットワーキングシステム、サービスとしてのソフトウェア（software as a service：SaaS）アーキテクチャ、または他のタイプのネットワーク化されたコンピューティングソリューションとともに動作するように構成されてもよく、またはそれらとして実現されてもよい。一実施形態では、本システムは、少なくとも本明細書で開示される機能を提供するとともに、コンピュータネットワークを介して（サーバとして機能する）コンピューティングシステムと通信するコンピューティングデバイス／端末を介して多くのユーザによってアクセスされる、集中サーバ側アプリケーションである。

40

【0148】

ーコンピューティングデバイスの実施形態ー

図11は、本明細書に記載の例示的なシステムおよび方法、ならびに／または同等例のうちの1つ以上で構成および／またはプログラムされる例示的なコンピューティングデバイス1100を示す。例示的なコンピューティングデバイスは、バス1125によって動作可能に接続されたプロセッサ1110と、メモリ1115と、入出力ポート1120と

50

を含むコンピュータ 1105 であり得る。一例では、コンピュータ 1105 は、図 1 ～図 10 に示す論理およびシステムと同様に、機械学習モデルの結果を確実に監査できるようにすることを容易にするように構成された機械学習監査保証論理 1130 を含み得る。さまざまな例では、論理 1130 は、ハードウェア、命令が記憶された非一時的なコンピュータ可読媒体、ファームウェア、および／またはそれらの組合せで実現され得る。論理 1130 は、バス 1125 に取り付けられたハードウェア構成要素として示されているが、他の実施形態では、論理 1130 が、プロセッサ 1110 において実現され得るか、メモリ 1115 に記憶され得るか、またはディスク 1135 に記憶され得ることを認識されたい。一実施形態では、論理 1130 またはコンピュータは、説明される動作を実行するための手段（たとえば、構造：ハードウェア、非一時的なコンピュータ可読媒体、ファームウェア）である。いくつかの実施形態では、コンピューティングデバイスは、クラウドコンピューティングシステムで動作するサーバ、サービスとしてのソフトウェア（SaaS）アーキテクチャで構成されたサーバ、スマートフォン、ラップトップ、タブレットコンピューティングデバイスなどであり得る。

【0149】

当該手段は、たとえば、機械学習モデルの結果を確実に監査できるようにするためにプログラムされた ASIC として実現され得る。当該手段はまた、メモリ 1115 に一時的に記憶されてプロセッサ 1110 によってさらに実行されるデータ 1140 としてコンピュータ 1105 に提示される記憶されたコンピュータ実行可能命令として実現されてもよい。

【0150】

論理 1130 は、機械学習モデルの結果を確実に監査できるようにするための手段（たとえば、ハードウェア、実行可能命令を記憶する非一時的なコンピュータ可読媒体、ファームウェア）を提供し得る。

【0151】

コンピュータ 1105 の例示的構成を概略的に説明すると、プロセッサ 1110 は、デュアルマイクロプロセッサおよび他のマルチプロセッサのアーキテクチャを含むさまざまなプロセッサであってもよい。メモリ 1115 は、揮発性メモリおよび／または不揮発性メモリを含み得る。不揮発性メモリは、たとえば、ROM、PROM 等を含み得る。揮発性メモリは、たとえば、RAM、SRAM、DRAM などを含み得る。

【0152】

ストレージディスク 1135 は、たとえば、入出力（input/output：I/O）インターフェイス（たとえば、カード、デバイス）1145 および入出力ポート 1120 を介してコンピュータ 1100 に動作可能に接続され得る。ディスク 1135 は、たとえば、磁気ディスクドライブ、ソリッドステートディスクドライブ、フロッピー（登録商標）ディスクドライブ、テープドライブ、Zip ドライブ、フラッシュメモリカード、メモリスティックなどであり得る。さらに、ディスク 1135 は、CD-ROM ドライブ、CD-R ドライブ、CD-RW ドライブ、DVD ROM 等であってもよい。メモリ 1115 は、たとえば、プロセス 1150 および／またはデータ 1140 を記憶することができる。ディスク 1135 および／またはメモリ 1115 は、コンピュータ 1105 のリソースを制御するとともに割当てするオペレーティングシステムを格納することができる。

【0153】

コンピュータ 1105 は、I/O インターフェイス 1145 および入出力ポート 1120 を介して入出力（I/O）デバイスと対話し得る。入出力デバイスは、たとえば、キーボード 1180、マイクロフォン 1184、ポインティングおよび選択デバイス 1182、カメラ 1186、ビデオカード、ディスプレイ 1170、スキャナ 1188、プリンタ 1172、スピーカ 1174、ディスク 1135、ネットワークデバイス 1155 などであり得る。入出力ポート 1120 は、たとえば、シリアルポート、パラレルポート、および USB ポートを含み得る。

【0154】

10

20

30

40

50

コンピュータ 1105 は、ネットワーク環境において動作可能であり、このため、I/O インターフェイス 1145 および／または I/O ポート 1120 を介してネットワークデバイス 1155 に接続され得る。コンピュータ 1105 は、ネットワークデバイス 1155 を通じてネットワーク 1160 と対話し得る。コンピュータ 1105 は、ネットワーク 1160 を介してリモートコンピュータ 1165 に論理的に接続され得る。コンピュータ 1105 が対話し得るネットワークは、LAN、WAN、および他のネットワークを含み得るが、それらに限定されない。

【0155】

一定義および他の実施形態－

別の実施形態では、記載される方法および／またはそれらの同等例は、コンピュータ実行可能命令を用いて実現され得る。したがって、一実施形態では、非一時的なコンピュータ可読／記憶媒体は、マシンによって実行されると当該マシン（および／または関連する構成要素）に当該方法を実行させるアルゴリズム／実行可能アプリケーションのコンピュータ実行可能命令が格納されるように構成されている。例示的なマシンは、プロセッサ、コンピュータ、クラウドコンピューティングシステムで動作するサーバ、サービスとしてのソフトウェア（SaaS）アーキテクチャで構成されたサーバ、スマートフォンなどを含むが、これらに限定されない。一実施形態では、コンピューティングデバイスは、開示された方法のいずれかを実行するように構成された 1 つ以上の実行可能アルゴリズムを用いて実現される。

【0156】

1 つ以上の実施形態において、開示された方法またはそれらの同等例は、当該方法を実行するように構成されたコンピュータハードウェア、または、非一時的なコンピュータ可読媒体に格納されたモジュールにおいて具現化されるコンピュータ命令によって実行される。ここで、当該命令は、コンピューティングデバイスの少なくともプロセッサによって実行されると当該方法を実行するように構成された実行可能アルゴリズムとして構成されている。

【0157】

説明を簡潔にするために、図に示される例示的な方法は、アルゴリズムの一連のブロックとして図示および説明されるものであるが、当該方法がブロックの順序によって限定されないことを理解されたい。いくつかのブロックは、図示および説明されたものとは異なる順序で、および／または他のブロックと同時に実行されてもよい。さらに、例示的な方法を実現するために、例示されたすべてのブロックよりも少ないブロックが用いられることもある。ブロックは、組合わされるか、または複数の動作／構成要素に分離されてもよい。さらに、付加的および／または代替的な方法は、ブロックには示されない追加の動作を用いることができる。

【0158】

以下は、本明細書中で用いられる選択された用語の定義を含む。これらの定義は、用語の範囲内にあるとともに実装のために用いられ得る構成要素のさまざまな例および／または形態を含む。これらの例は、限定するよう意図されたものではない。用語の単数形および複数形はともに、定義の範囲内であり得る。

【0159】

「一実施形態」、「ある実施形態」、「一例」、「ある例」などと言及する場合、それらは、そのように記載される実施形態または例が、特定の特徴、構造、特性、性質、要素または限定を含み得るが、すべての実施形態または例が必ずしもその特定の特徴、構造、特性、性質、要素または限定を含むわけではないことを示している。さらに、「一実施形態においては」という表現を繰返し用いる場合、必ずしも同じ実施形態を指すわけではないが、同じ実施形態である可能性もあり、以下を含み得る。

【0160】

ASIC：特定用途向け集積回路

CD：コンパクトディスク

C D - R : 記録可能な C D
 C D - R W : 書き換え可能な C D
 D V D : デジタル多用途ディスクおよび／またはデジタルビデオディスク
 H T T P : ハイパーテキスト転送プロトコル
 L A N : ローカルエリアネットワーク
 R A M : ランダムアクセスメモリ
 D R A M : ダイナミック R A M
 S R A M : 同期 R A M
 R O M : 読取り専用メモリ
 P R O M : プログラムブル R O M
 E P R O M : 消去可能 P R O M
 E E P R O M : 電氣的に消去可能な P R O M
 U S B : ユニバーサルシリアルバス
 X M L : 拡張マークアップ言語
 W A N : 広域ネットワーク

10

本明細書で用いられる「データ構造」は、メモリ、ストレージデバイス、または他のコンピュータ化されたシステムに格納されるコンピューティングシステム内のデータの構成である。データ構造は、たとえば、データフィールド、データファイル、データ配列、データ記録、データベース、データテーブル、グラフ、ツリー、リンクされたリストなどのうちのいずれか1つであってもよい。データ構造は、他の多くのデータ構造から形成され得るとともに、それらを含み得る（たとえば、データベースは多くのデータ記録を含む）。他の実施形態に従うと、データ構造の他の例も同様に実現可能である。

20

【0161】

本明細書中で用いられる「コンピュータ可読媒体」または「コンピュータ記憶媒体」は、実行されたときに、開示される機能のうちの1つ以上を実行するように構成された命令および／またはデータを格納する非一時的な媒体を指す。データは、いくつかの実施形態においては、命令として機能し得る。コンピュータ可読媒体は、不揮発性媒体および揮発性媒体を含む形態を取り得るが、これらに限定されない。不揮発性媒体は、たとえば、光ディスク、磁気ディスクなどを含み得る。揮発性媒体は、たとえば、半導体メモリ、ダイナミックメモリなどを含み得る。コンピュータ可読媒体の一般的な形態は、フロッピー（登録商標）ディスク、フレキシブルディスク、ハードディスク、磁気テープ、他の磁気媒体、特定用途向け集積回路（application specific integrated circuit : A S I C）、プログラムブル論理デバイス、コンパクトディスク（compact disk : C D）、他の光媒体、ランダムアクセスメモリ（random access memory : R A M）、読取り専用メモリ（read only memory : R O M）、メモリチップまたはカード、メモリスティック、ソリッドステートストレージデバイス（solid state storage device : S S D）、フラッシュドライブ、および、コンピュータ、プロセッサまたは他の電子デバイスが機能することができ他の媒体を含み得るが、これらに限定されない。一実施形態において実装のために選択される場合、媒体の各タイプは、開示される機能および／または請求される機能のうちの1つ以上を実行するように構成されたアルゴリズムの格納命令を含み得る。

30

40

【0162】

本明細書で用いられる「論理」は、本明細書で開示した機能もしくは動作のいずれかを実行するために、ならびに／または、本明細書で開示した別の論理、方法および／もしくはシステムからの機能もしくは動作を実行させるために、コンピュータもしくは電気ハードウェア、実行可能なアプリケーションもしくはプログラムモジュールの命令が格納された非一時的な媒体、および／または、これらの組合わせで実現される構成要素を表わす。同等の論理は、ファームウェア、アルゴリズムでプログラムされたマイクロプロセッサ、個別論理（たとえば A S I C）、少なくとも1つの回路、アナログ回路、デジタル回路、プログラムされた論理デバイス、アルゴリズムの命令を含むメモリデバイス等を含み得るが、これらのうちのいずれも、開示された機能のうちの1つ以上を実行するように構成さ

50

れ得る。一実施形態においては、論理は、1つ以上のゲート、ゲートの組合わせ、または開示された機能のうちの1つ以上を実行するように構成された他の回路部品を含み得る。複数の論理が説明される場合、複数の論理を1つの論理に組み込むことが可能であり得る。同様に、単一の論理が説明される場合、その単一の論理を複数の論理間で分配することが可能であり得る。一実施形態においては、これらの論理のうちの1つ以上は、開示される機能および／または請求される機能を実行することに関連する対応する構造である。どのタイプの論理を実装するかを選択は、所望のシステム条件または仕様に基づき得る。たとえば、より高い速度が考慮事項である場合、機能を実現するためにハードウェアが選択されることとなるだろう。より低いコストが考慮事項である場合、機能を実現するために、格納された命令／実行可能アプリケーションが選択されることとなるだろう。

10

【0163】

「動作可能な接続」、またはエンティティが「動作可能に接続される」接続は、信号、物理的通信および／または論理通信が送信および／または受信され得る接続である。動作可能な接続は、物理的インターフェイス、電気的インターフェイスおよび／またはデータインターフェイスを含み得る。動作可能な接続は、動作可能な制御を可能にするのに十分なインターフェイスおよび／または接続のさまざまな組合わせを含み得る。たとえば、2つのエンティティは、信号を互いに直接または1つ以上の中間エンティティ（たとえば、プロセッサ、オペレーティングシステム、論理、非一時的なコンピュータ可読媒体）を介して通信するように動作可能に接続することができる。論理通信チャネルおよび／または物理通信チャネルを用いて、動作可能な接続を作成することができる。

20

【0164】

本明細書で用いられる「ユーザ」は、1以上の人、コンピュータまたは他のデバイス、またはこれらの組合わせを含むが、これらに限定されない。

【0165】

開示された実施形態をかなり詳細に図示および説明してきたが、添付の特許請求の範囲をそのような詳細に制限するかまたは何らかの形で限定することは意図していない。当然、主題のさまざまな局面を説明する目的で、構成要素または方法論の考えられ得るすべての組合わせを説明することは不可能である。したがって、本開示は、図示および説明される特定の詳細または具体的な例に限定されない。このため、本開示は、添付の特許請求の範囲内に収まる変更、修正および変形を包含するよう意図されたものである。

30

【0166】

「含む」（「includes」または「including」）という語が詳細な説明または請求項で採用される限りにおいて、この語が請求項の中で移行語として利用される際に解釈されるとおりに、「備える」（「comprising」）という語と同様に包括的であることが意図される。

【0167】

「または」という語が詳細な説明または請求項で使用される（たとえば、AまたはB）限りにおいて、「AまたはBまたはそれら両方」を意味することが意図される。出願人が「AまたはBのみであり、両方ではない」ことを示すことを意図する場合、「AまたはBのみであり、両方ではない」という表現が使用されるだろう。したがって、本明細書における「または」という語の使用は包括的な使用であり、排他的な使用ではない。

40

【図面】

【図 1】

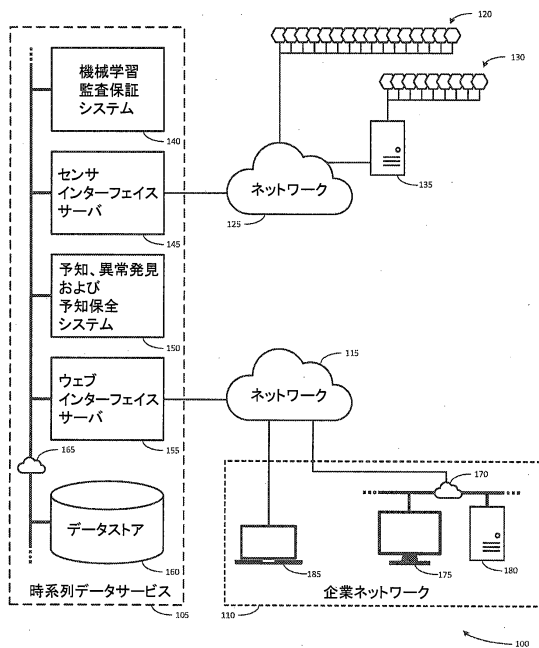


FIG. 1

【図 2】

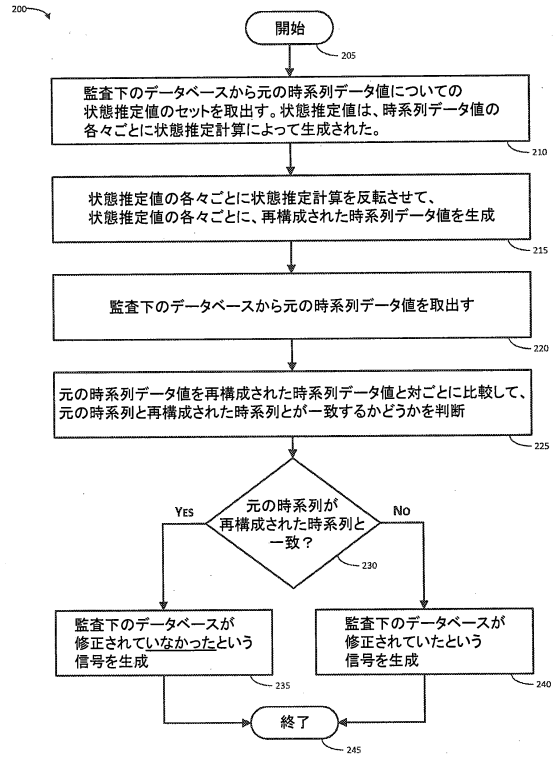


FIG. 2

【図 3】

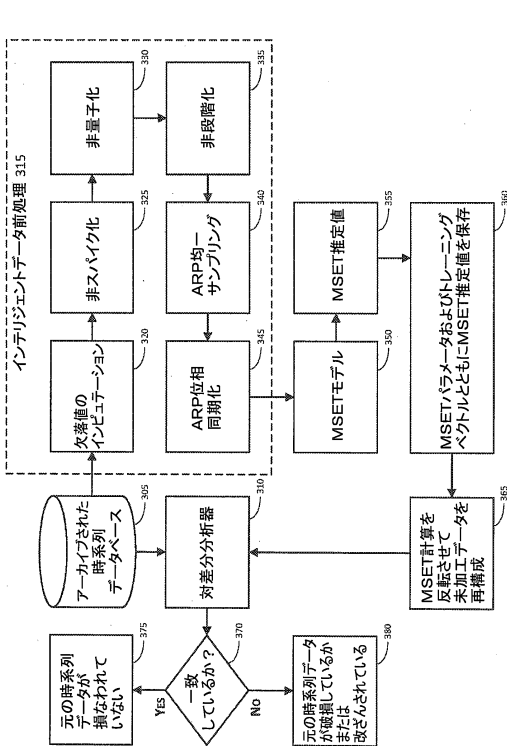


FIG. 3

【図 4】

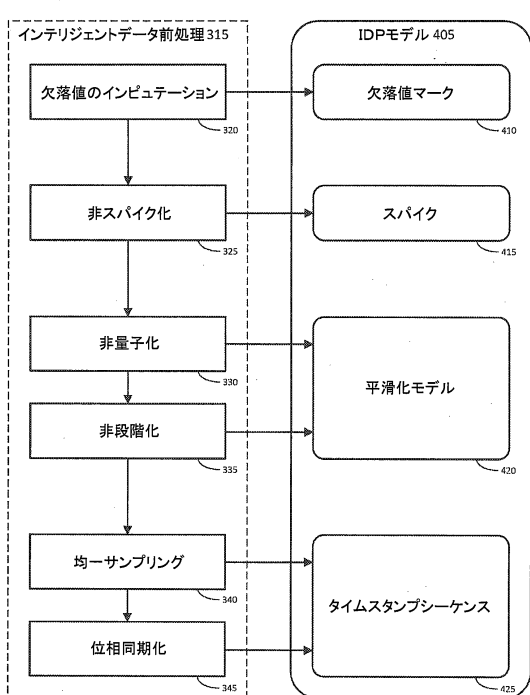


FIG. 4

10

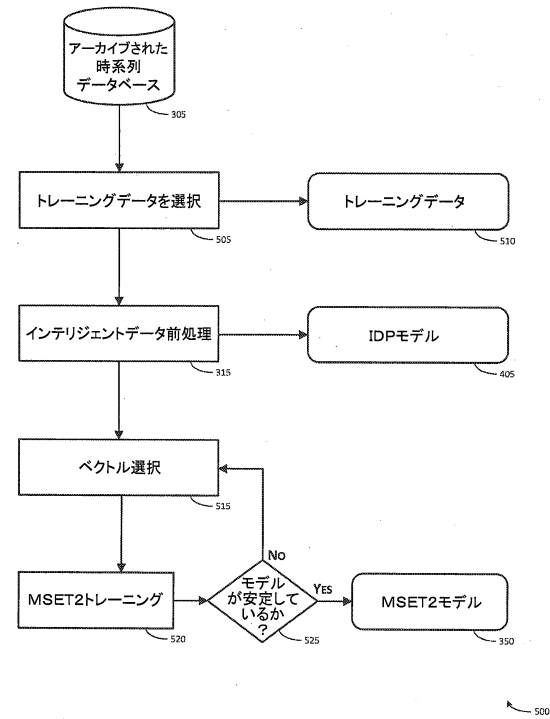
20

30

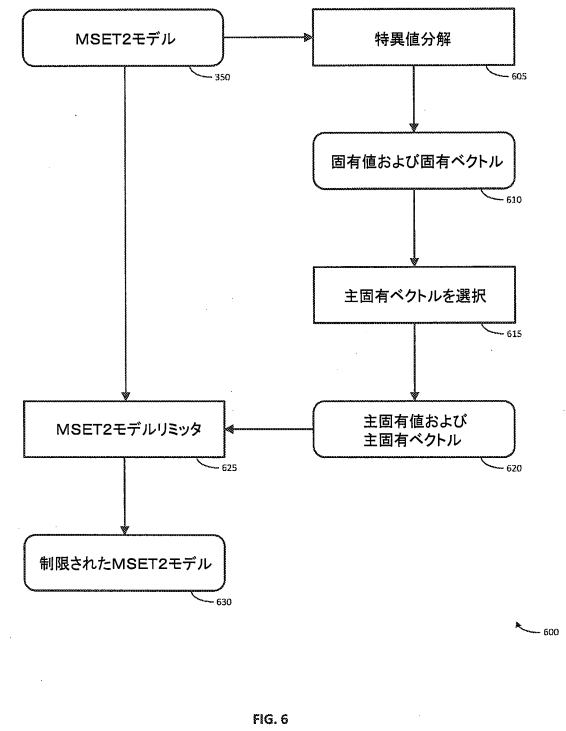
40

50

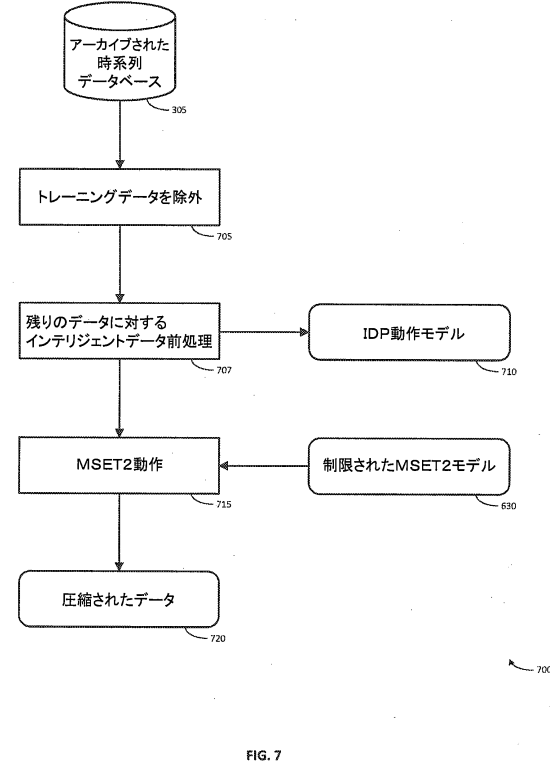
【図 5】



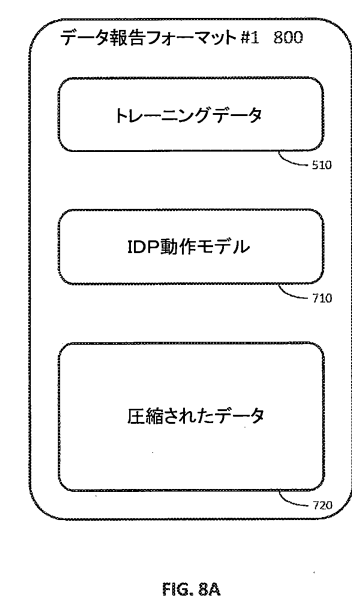
【図 6】



【図 7】



【図 8 A】



10

20

30

40

50

【図 8 B】

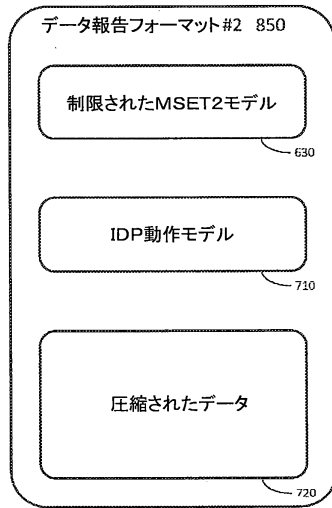


FIG. 8B

【図 9】

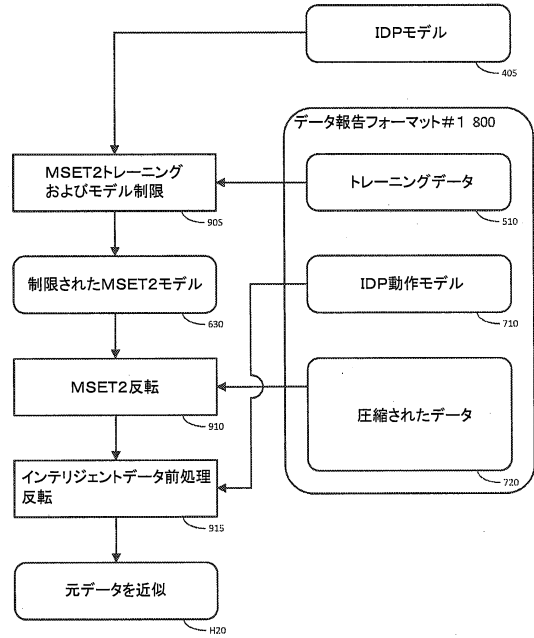


FIG. 9

【図 10】

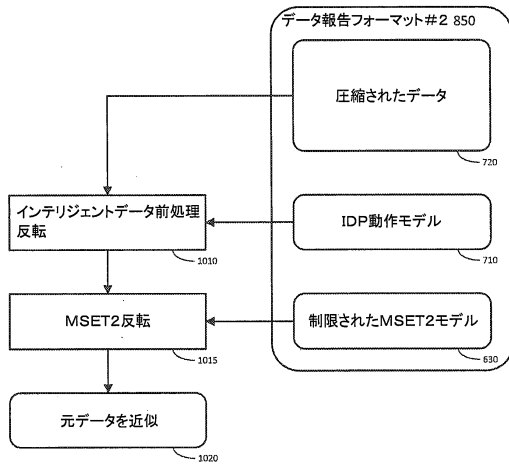


FIG. 10

【図 11】

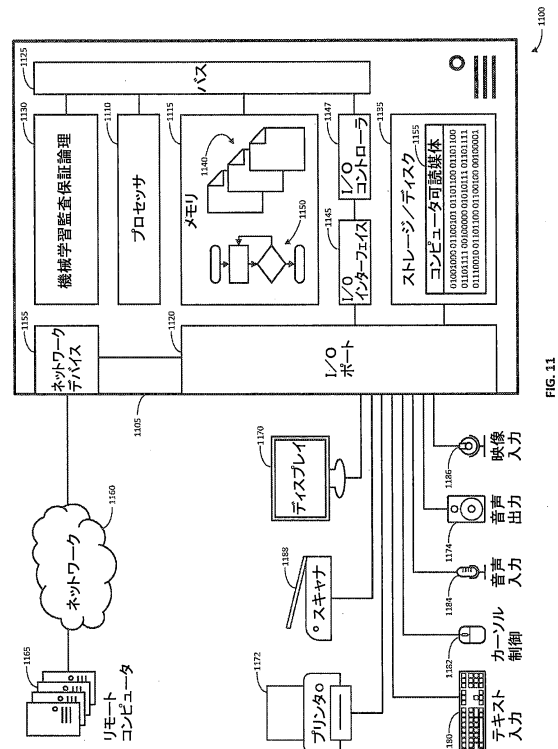


FIG. 11

フロントページの続き

- ーセッツ州、ウォルサム、フェアモント・アベニュー、35
- (72)発明者 ワン, グアン・シイ
アメリカ合衆国、92110 カリフォルニア州、サン・ディエゴ、カブリロ・ベイ・レーン、3152
- (72)発明者 グロス, ケニー・シイ
アメリカ合衆国、92025 カリフォルニア州、エスコンディード、ナポール・ストリート、2439
- (72)発明者 チスティアコバ, アンナ
アメリカ合衆国、94107 カリフォルニア州、サン・フランシスコ、ベイサイド・ビレッジ・プレイス、2、アパートメント・112
- (72)発明者 ガウリック, ディーター
アメリカ合衆国、94306 カリフォルニア州、パロ・アルト、ポール・アベニュー、757
- (72)発明者 リウ, ジェン・フア
アメリカ合衆国、94403 カリフォルニア州、サン・マテオ、ウェイン・ウェイ、1017
- (72)発明者 ソンデレグガー, リチャード・ポール
アメリカ合衆国、02124 マサチューセッツ州、ドーチェスター、アルバン・ストリート、61
- 審査官 福西 章人
- (56)参考文献 米国特許出願公開第2019/0197145 (US, A1)
米国特許出願公開第2019/0286725 (US, A1)
- (58)調査した分野 (Int.Cl., DB名)
G06N 3/00-99/00
G06F 18/00-18/40