(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年9月3日(03.09.2020)

WIPO | PCT

(10) 国際公開番号
WO 2020/175147 A1

(51) 国際特許分類:
G06F 11/07 (2006.01) *G06N 3/08* (2006.01)

(21) 国際出願番号： PCT/JP2020/005474

(22) 国際出願日： 2020年2月13日(13.02.2020)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2019-037021 2019年2月28日(28.02.2019) JP

(71) 出願人: 日本電信電話株式会社 (**NIPPON TELEGRAPH AND TELEPHONE CORPORATION**) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目５番１号 Tokyo (JP).

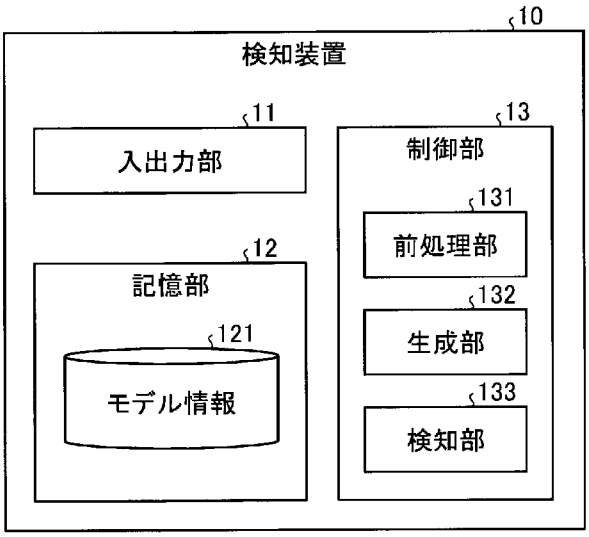
(72) 発明者: 東羅 翔太郎 (**TORA, Shotaro**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ 知的財産センタ内 Tokyo (JP). 外山 将司 (**TOYAMA, Masashi**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ 知的財産センタ内 Tokyo (JP). 豊田 真智子(**TOYODA, Machiko**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ 知的財産センタ内 Tokyo (JP).

(74) 代理人: 特許業務法人酒井国際特許事務所 (**SAKAI INTERNATIONAL PATENT OFFICE**); 〒1000013 東京都千代田区霞が関３丁目８番１号 虎の門三井ビルディング Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,

(54) **Title:** DETECTION DEVICE AND DETECTION PROGRAM

(54) 発明の名称： 検知装置及び検知プログラム



10 Detection device
11 Input/output unit
12 Storage unit
13 Control unit
121 Model information
131 Preprocessing unit
132 Generation unit
133 Detection unit

(57) **Abstract:** A preprocessing unit (131) processes training data and detection target data. In addition, a generation unit (132) generates a normal state model by means of deep learning on the basis of the training data processed by the preprocessing unit (131). Additionally, a detection unit (133) calculates an anomaly level on the basis of output data obtained by inputting, into the model, the detection target data processed by the preprocessing unit (131), and detects an anomaly in the detection target data on the basis of the anomaly level.

(57) 要約：前処理部（１３１）は、学習用のデータ及び検知対象のデータを加工する。また、生成部（１３２）は、前処理部（１３１）によって加工された学習用のデータを基に、深層学習により正常状態のモデルを生成する。また、検知部（１３３）は、前処理部（１３１）によって加工された検知対象のデータをモデルに入力して得られた出力データを基に異常度を計算し、異常度を基に検知対象のデータの異常を検知する。

WO 2020/175147 A1

［続葉有］

HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

明　細　書

発明の名称 ： 検知装置及び検知プログラム

技術分野

[0001]　本発明は、検知装置及び検知プログラムに関する。

背景技術

[0002]　従来、深層学習（Deep　Learning）を用いたモデルであるオートエンコーダ（ＡＥ：autoencoder）やリカレントニューラルネットワーク（ＲＮＮ：Recurrent　neural　network、ＬＳＴＭ：Long　short-term　memory、ＧＲＵ：Gated　recurrent　unit）を使った異常検知技術が知られている。例えば、オートエンコーダを用いた従来技術では、まず正常なデータの学習によりモデルが生成される。そして、検知対象のデータと、当該データをモデルに入力して得られる出力データとの間の再構成誤差が大きいほど異常の度合いが大きいと判断される。

先行技術文献

非特許文献

[0003] 非特許文献1：池田　泰弘、石橋　圭介、中野　雄介、渡辺敬志郎、川原　亮一、「オートエンコーダを用いた異常検知におけるモデル再学習手法」、信学技報　IN2017-84

発明の概要

発明が解決しようとする課題

[0004]　しかしながら、従来の技術には、深層学習を使って異常検知を行う場合に、検知精度が低下する場合があるという問題がある。例えば、従来技術では、異常検知のための学習用のデータ又は検知対象のデータに対する適切な前処理が行われない場合がある。また、従来技術では、モデルの生成が乱数依存であるため、学習データに対して一意のモデルかどうかを確認するのが難しい。また、従来技術では、学習データに異常が含まれている可能性を考慮していない場合がある。いずれの場合も、異常検知における検知精度が低下

することが考えられる。なお、ここでいう検知精度の低下は、異常なデータを異常と検知する検知率の低下、及び正常なデータを異常と検知する誤検知率の上昇を指すものとする。

課題を解決するための手段

[0005] 上述した課題を解決し、目的を達成するために、検知装置は、学習用のデータ及び検知対象のデータを加工する前処理部と、前記前処理部によって加工された学習用のデータを基に、深層学習によりモデルを生成する生成部と、前記前処理部によって加工された検知対象のデータを前記モデルに入力して得られた出力データを基に異常度を計算し、前記異常度を基に前記検知対象のデータの異常を検知する検知部と、を有することを特徴とする。

発明の効果

[0006] 本発明によれば、深層学習を使って異常検知を行う場合の学習データの前処理や選定、モデルの選択を適切に行えるようになり、検知精度を向上させることができる。

図面の簡単な説明

[0007] [図1]図１は、第１の実施形態に係る検知装置の構成の一例を示す図である。

[図2]図２は、オートエンコーダについて説明するための図である。

[図3]図３は、学習について説明するための図である。

[図4]図４は、異常検知について説明するための図である。

[図5]図５は、特徴量ごとの異常度について説明するための図である。

[図6]図６は、変動が小さい特徴量の特定について説明するための図である。

[図7]図７は、逓増するデータ及び逓減するデータの一例を示す図である。

[図8]図８は、逓増するデータ及び逓減するデータを変換したデータの一例を示す図である。

[図9]図９は、モデルが安定する場合の例を示す図である。

[図10]図１０は、モデル安定しない場合の例を示す図である。

[図11]図１１は、固定期間学習の結果の一例を示す図である。

[図12]図１２は、スライディング学習の結果の一例を示す図である。

[図13]図１３は、正規化手法ごとの異常度の一例を示す図である。

[図14]図１４は、第１の実施形態に係る検知装置の学習処理の流れを示すフローチャートである。

[図15]図１５は、第１の実施形態に係る検知装置の検知処理の流れを示すフローチャートである。

[図16]図１６は、テキストログの異常度の一例を示す図である。

[図17]図１７は、テキストログの異常度と障害情報との関係の一例を示す図である。

[図18]図１８は、障害発生時のテキストログ及び特徴量ごとの異常度の一例を示す図である。

[図19]図１９は、異常度が上昇したときのテキストログ及び特徴量ごとの異常度の一例を示す図である。

[図20]図２０は、異常度が上昇した前後の時刻のテキストログＩＤのデータ分布の一例を示す図である。

[図21]図２１は、数値ログの異常度の一例を示す図である。

[図22]図２２は、数値ログの異常度と障害情報との関係の一例を示す図である。

[図23]図２３は、障害発生時の数値ログ及び特徴量ごとの異常度の一例を示す図である。

[図24]図２４は、数値ログの特徴量ごとの入力データと出力データの一例を示す図である。

[図25]図２５は、検知プログラムを実行するコンピュータの一例を示す図である。

発明を実施するための形態

[0008] 以下に、本願に係る検知装置及び検知プログラムの実施形態を図面に基づいて詳細に説明する。なお、本発明は、以下に説明する実施形態により限定されるものではない。

[0009] ［第１の実施形態の構成］

まず、図１を用いて、第１の実施形態に係る検知装置の構成について説明する。図１は、第１の実施形態に係る検知装置の構成の一例を示す図である。図１に示すように、検知装置１０は、入出力部１１、記憶部１２及び制御部１３を有する。

[0010] 入出力部１１は、データの入出力を行うためのインタフェースである。例えば、入出力部１１は、ネットワークを介して他の装置との間でデータ通信を行うためのＮＩＣ（Network Interface Card）であってもよい。

[0011] 記憶部１２は、ＨＤＤ（Hard Disk Drive）、ＳＳＤ（Solid State Drive）、光ディスク等の記憶装置である。なお、記憶部１２は、ＲＡＭ（Random Access Memory）、フラッシュメモリ、ＮＶＳＲＡＭ（Non Volatile Static Random Access Memory）等のデータを書き換え可能な半導体メモリであってもよい。記憶部１２は、検知装置１０で実行されるＯＳ（Operating System）や各種プログラムを記憶する。さらに、記憶部１２は、プログラムの実行で用いられる各種情報を記憶する。また、記憶部１２は、モデル情報１２１を記憶する。

[0012] モデル情報１２１は、生成モデルを構築するための情報である。実施形態では、生成モデルはオートエンコーダであるものとする。また、オートエンコーダは、エンコーダ及びデコーダにより構成される。エンコーダ及びデコーダはいずれもニューラルネットワークである。このため、例えば、モデル情報１２１は、エンコーダ及びデコーダの層の数、各層の次元の数、ノード間の重み、層ごとのバイアス等を含む。また、以降の説明では、モデル情報１２１に含まれる情報のうち、重み及びバイアス等の学習により更新されるパラメータをモデルパラメータと呼ぶ場合がある。また、生成モデルを単にモデルと呼ぶ場合がある。

[0013] 制御部１３は、検知装置１０全体を制御する。制御部１３は、例えば、ＣＰＵ（Central Processing Unit）、ＭＰＵ（Micro Processing Unit）等の電子回路や、ＡＳＩＣ（Application Specific Integrated Circuit）、ＦＰＧＡ（Field Programmable Gate Array）等の集積回路である。

また、制御部１３は、各種の処理手順を規定したプログラムや制御データを格納するための内部メモリを有し、内部メモリを用いて各処理を実行する。また、制御部１３は、各種のプログラムが動作することにより各種の処理部として機能する。例えば、制御部１３は、前処理部１３１、生成部１３２、検知部１３３及び更新部１３４を有する。

[0014] 前処理部１３１は、学習用のデータ及び検知対象のデータを加工する。また、生成部１３２は、前処理部１３１によって加工された学習用のデータを基に、深層学習によりモデルを生成する。また、検知部１３３は、前処理部１３１によって加工された検知対象のデータをモデルに入力して得られた出力データを基に異常度を計算し、異常度を基に検知対象のデータの異常を検知する。なお、実施形態では、生成部１３２は、深層学習にオートエンコーダを用いる。また、以降の説明では、学習用のデータ及び検知対象のデータを、それぞれ学習データ及びテストデータと呼ぶ。

[0015] このように、検知装置１０は、制御部１３の各部の処理により、学習処理及び検知処理を行うことができる。また、生成部１３２は、生成したモデルの情報をモデル情報１２１として記憶部１２に格納する。また、生成部１３２は、モデル情報１２１を更新する。検知部１３３は、記憶部１２に記憶されたモデル情報１２１を基にオートエンコーダを構築し、異常検知を行う。

[0016] ここで、図２を用いて、実施形態におけるオートエンコーダについて説明する。図２は、オートエンコーダについて説明するための図である。図２に示すように、オートエンコーダを構成するＡＥネットワーク２は、エンコーダとデコーダを有する。ＡＥネットワーク２には、例えば、データに含まれる１つ以上の特徴量の値が入力される。そして、エンコーダは、入力された特徴量群を圧縮表現に変換する。さらにデコーダは、圧縮表現から特徴量群を生成する。このとき、デコーダは、入力されたデータと同様の構造を持つデータを生成する。

[0017] このように、オートエンコーダがデータを生成すること再構成と呼ぶ。また、再構成されたデータを再構成データと呼ぶ。また、入力されたデータと

再構成データとの誤差を再構成誤差と呼ぶ。

[0018] 図３を用いて、オートエンコーダの学習について説明する。図３は、学習について説明するための図である。図３に示すように、学習の際には、検知装置１０は、各時刻の正常なデータをＡＥネットワーク２に入力する。そして、検知装置１０は、再構成誤差が小さくなるようにオートエンコーダの各パラメータを最適化する。このため、十分に学習が行われると、入力データと再構成データが同値になる。

[0019] 図４を用いて、オートエンコーダによる異常検知について説明する。図４は、異常検知について説明するための図である。図４に示すように、検知装置１０は、正常であるか異常であるかが未知のデータをＡＥネットワーク２に入力する。

[0020] ここで、時刻t_1のデータは正常であるものとする。このとき、時刻t_1のデータに対する再構成誤差は十分に小さくなり、検知装置１０は、時刻t_1のデータが再構成可能、すなわち正常であると判断する。

[0021] 一方、時刻t_2のデータは異常であるものとする。このとき、時刻t_2のデータに対する再構成誤差は大きくなり、検知装置１０は、時刻t_1のデータが再構成不可能、すなわち異常であると判断する。なお、検知装置１０は、再構成誤差の大小を閾値により判定してもよい。

[0022] 例えば、検知装置１０は、複数の特徴量を持つデータを使って学習及び異常検知を行うことができる。このとき、検知装置１０は、データごとの異常度だけでなく、特徴量ごとの異常度を計算することができる。

[0023] 図５を用いて、特徴量ごとの異常度について説明する。図５は、特徴量ごとの異常度について説明するための図である。図５の例では、特徴量は、コンピュータにおける各時刻のＣＰＵ使用率、メモリ使用率、ディスクＩＯ速度等である。例えば、入力データの特徴量と再構成されたデータの特徴量とを比較すると、ＣＰＵ１とｍｅｍｏｒｙ１の値に大きく差がついている。この場合、ＣＰＵ１とｍｅｍｏｒｙ１を原因とする異常が発生している可能性があると推定することができる。

[0024] また、オートエンコーダのモデルは、学習データのサイズに依存せずにコンパクトにすることが可能である。また、モデルが生成済みであれば、検知は行列演算によって行われるため、高速に処理することが可能になる。

[0025] 実施形態の検知装置１０は、検知対象の装置から出力されるログを基に、当該装置の異常を検知することができる。例えば、ログは、センサによって収集されるセンサデータであってもよい。例えば、検知対象の装置は、サーバ等の情報処理装置であってもよいし、ＩｏＴ機器であってもよい。例えば、検知対象の装置は、自動車に搭載された車載器、医療用のウェアラブル測定機器、生産ラインで使用される検査装置、ネットワークの末端のルータ等である。また、ログの種類は、数値及びテキストが含まれる。例えば、情報処理装置であれば、数値ログはＣＰＵやメモリなどの装置から収集される測定値、テキストログはsyslogやＭＩＢといったメッセージログである。

[0026] ここで、単にオートエンコーダの学習を行い、学習済みのモデルを使って異常検知を行うだけでは、十分な検知精度が得られない場合がある。例えば、各データに対し適切な前処理が行われない場合や、複数回学習した場合のモデル選択を誤った場合、学習データに異常が含まれている可能性を考慮していない場合に検知精度が低下することが考えられる。そこで、検知装置１０は、以下に説明する各処理の少なくともいずれかを実行することで、検知精度を向上させることができる。

[0027] （１．変動が小さい特徴量の特定）

学習データにおいては変動が小さかった特徴量が、検知対象データにおいてわずかでも変動した場合、検知結果に大きな影響を与える場合がある。この場合、本来異常でないデータに対する異常度が過剰に大きくなり、誤検知が起きやすくなる。

[0028] そこで、前処理部１３１は、特徴量の時系列データである学習用のデータから、時間に対する変動の大きさの度合いが所定値以下である特徴量を特定する。また、検知部１３３は、検知対象のデータの特徴量のうち、前処理部１３１によって特定された特徴量、又は、前処理部１３１によって特定され

た特徴量以外の特徴量の少なくともいずれかを基に異常を検知する。

[0029] つまり、検知部１３３は、テストデータの特徴量のうち、学習データにおいて変動が大きかった特徴量のみを用いて検知を行うことができる。これにより、検知装置１０は、学習データにおいては変動が小さかった特徴量が、検知対象データにおいてわずかでも変動した場合の異常度の影響を抑えることができ、異常でないデータの誤検知を抑制することができる。

[0030] 一方、検知部１３３は、テストデータの特徴量のうち、学習データにおいて変動が小さかった特徴量のみを用いて検知を行うことができる。この場合、検知装置１０は、検知における異常度のスケールを大きくする。これにより、検知装置１０は、検知対象データにおいて変動が大きくなった場合のみを異常として検知することができる。

[0031] 図６は、変動が小さい特徴量の特定について説明するための図である。図６の上部の表は、特徴量の学習データの標準偏差（ＳＴＤ）を計算し、閾値を設定したときの、各閾値に該当する特徴量の数である。例えば、閾値を０．１とした場合、ＳＴＤ≧０．１となる特徴量数（Ｇｒｏｕｐ１性能値数）は１３２個である。また、そのとき、ＳＴＤ＜０．１となる特徴量数（Ｇｒｏｕｐ２性能値数）は４８個である。

[0032] 図６の例では、特徴量の標準偏差の閾値を０．１とする。このとき、前処理部１３１は、学習データから、標準偏差が０．１未満である特徴量を特定する。そして、検知部１３３が、テストデータから特定された特徴量を用いて検知を行った場合（ＳＴＤ＜０．１）、異常度は６．９×10^{12}～３．７×10^{16}程度であった。一方、検知部１３３が、テストデータから特定された特徴量を除いて検知を行った場合（ＳＴＤ≧０．１）、異常度はＳＴＤ＜０．１の場合と比べて非常に小さくなり、最大でも２０，０００程度であった。

[0033] （２．逓増又は逓減するデータの変換）

サーバシステムから出力されるデータのように、逓増又は逓減するデータが存在する。このようなデータの特徴量は、学習データとテストデータにお

いて取り得る値の範囲が異なる場合があり、誤検知の原因になる。例えば、累積値の場合、累積値そのものよりも値の変化度合い等に意味がある場合がある。

[0034] そこで、前処理部１３１は、学習用のデータ及び検知対象のデータの一部又は全てを、当該データの所定の時刻間の差又は比に変換する。例えば、前処理部１３１は、時刻間のデータの値の差分を取ってもよいし、ある時刻のデータの値を１つ前の時刻のデータの値で割ってもよい。これにより、検知装置１０は、学習データとテストデータのデータ取り得る範囲の違いによる影響を抑えることができ、誤検知の発生を抑え、さらに、テストデータにおいて、学習時と異なる変化をする特徴量の異常を検知しやすくなる。図７は、遁増するデータ及び遁減するデータの一例を示す図である。また、図８は、遁増するデータ及び遁減するデータを変換したデータの一例を示す図である。

[0035] （３．最適なモデルの選択）

モデルの学習を行う際には、モデルパラメータの初期値等をランダムに決定する場合がある。例えば、オートエンコーダを含むニューラルネットワークを用いたモデルの学習を行う際には、ノード間の重み等の初期値をランダムに決定する場合がある。また、誤差逆伝播の際にドロップアウトの対象になるノードがランダムに決定される場合がある。

[0036] このような場合、層（レイヤ）の数、層ごとの次元数（ノード数）が一定であっても、最終的に生成されるモデルが毎回同じものになるとは限らない。そのため、ランダム性のパターンを変えて学習を複数回試行すると、複数のモデルが生成されることになる。ランダム性のパターンを変えることは、例えば初期値として使用する乱数を発生させ直すことである。このような場合、各モデルから計算される異常度が異なることもあり、異常検知に使用するモデルによっては、誤検知が起きる原因になる。

[0037] そこで、生成部１３２は、複数のパターンごとに学習を行う。つまり、生成部１３２は、学習用のデータに対して複数回学習を行う。そして、検知部

１３３は、生成部１３２によって生成されたモデルのうち、互いの関係の強さに応じて選択されたモデルを用いて異常を検知する。生成部１３２は、関係の強さとして、同一のデータを入力したときの再構成データから計算される異常度間の相関係数を計算する。

[0038] 図９は、モデルが安定する場合の異常度の例を示す図である。各矩形内の数値は、各モデルの異常度間の相関係数である。例えば、trial3とtrial8との相関係数は０．８である。図９の例では、モデル間の相関係数が最低でも０．７７と高いため、生成部１３２がどのモデルを選択しても大きな差は生じないと考えられる。

[0039] 一方、図１０は、モデルが安定しない場合の異常度の例を示す図である。図９と同様に、各矩形内の数値は、モデル間の相関係数である。例えば、trial3という試行で生成されたモデルと、trial8という試行で生成されたモデルとの相関係数は０．９２である。図１０の場合、これらのモデルからは選択せず、層の数や層ごとの次元数などを変更して、再度モデルの生成をやり直すのがよい。

[0040] （４．データの分布の時間変化への対応）

サーバシステムから出力されるデータのように、時間の経過に応じて分布が変化するデータがある。このため、分布の変化前に収集された学習データを使って生成されたモデルを用いて、分布の変化後に収集されたテストデータの検知を行った場合、テストデータの正常分布を学習していないために、正常なデータの異常度が大きくなることが考えられる。

[0041] そこで、前処理部１３１は、時系列データである学習用のデータを所定の期間ごとのスライディングウィンドウで分割する。そして、生成部１３２は、前処理部１３１によって分割されたスライディングウィンドウごとのデータのそれぞれを基に、モデルを生成する。また、生成部１３２は、固定期間の学習データに基づくモデルの生成（固定期間学習）と、当該固定期間をスライディングウィンドウで分割した期間それぞれの学習データに基づくモデルの生成（スライティング学習）の両方を行ってもよい。また、スライディ

ング学習は、分割されたスライディングウィンドウごとのデータを基に生成されるモデルをすべて使用するのではなく、その中からいずれかを選択して使用してもよい。例えば、前日から一定期間遡ったデータを使って作成したモデルを、翌日１日の異常検知に適用することを繰り返してもよい。

[0042] 図１１は、固定期間学習の結果の一例を示す図である。図１２は、スライディング学習の結果の一例を示す図である。図１１及び図１２は、各モデルから計算された異常度を表している。スライディング学習は、前日までの2週間のデータで作成したモデルで、翌日1日の異常度を算出している。スライディング学習の方が、固定期間学習と比べて異常度が上昇する期間が多い。これは、短期的に見ると、データ分布が細かく変化しているためであると見ることができる。

[0043] （５．学習データからの異常データの除去）

検知装置１０はいわゆるアノマリ検知を行うものであるため、学習データはなるべく正常なデータであることが望ましい。一方で、収集した学習データの中には、人が認識することが難しい異常データや外れ度が高いデータが含まれていることがある。

[0044] 前処理部１３１は、学習用のデータに対する複数の異なる正規化手法ごとに生成されたモデル群、又は、それぞれに異なるモデルパラメータが設定されたモデル群のうちの少なくとも一方のモデル群に含まれる少なくとも１つのモデルを使って計算された異常度が所定の値より高いデータを学習用のデータから除外する。なお、この場合のモデルの生成及び異常度の計算は、それぞれ生成部１３２及び検知部１３３によって行われてもよい。

[0045] 図１３は、正規化手法ごとの異常度の一例を示す図である。図１３に示すように、各正規化手法に共通して、０２／０１以降の異常度が高い。この場合、前処理部１３１は、学習データから０２／０１以降のデータを除外する。また、前処理部１３１は、少なくとも１つの正規化手法で異常度が高くなるデータを除外することができる。

[0046] また、異なるモデルパラメータが設定されたモデル群を用いて異常度を計

算する場合も、図１３に示すような複数の異常度の時系列データが得られる。その場合も同様に、前処理部１３１は、いずれかの時系列データで異常度が高いデータを除外する。

[0047] ［第１の実施形態の処理］

図１４を用いて、検知装置１０の学習処理の流れについて説明する。図１４は、第１の実施形態に係る検知装置の学習処理の流れを示すフローチャートである。図１４に示すように、まず、検知装置１０は、学習データの入力を受け付ける（ステップＳ１０１）。次に、検知装置１０は、逓増又は逓減する特徴量のデータを変換する（ステップＳ１０２）。例えば、検知装置１０は、各データを所定の時刻間の差又は比に変換する。

[0048] ここで、検知装置１０は、バリエーションごとの学習データに対する正規化を実行する（ステップＳ１０３）。バリエーションとは、正規化の手法であり、図１３に示すmin-max正規化、標準化（Z-score）、ロバスト正規化等が含まれる。

[0049] 検知装置１０は、生成モデルを用いて、学習データからデータを再構成する（ステップＳ１０４）。そして、検知装置１０は、再構成誤差から異常度を計算する（ステップＳ１０５）。そして、検知装置１０は、異常度が高い期間のデータを除外する（ステップＳ１０６）。

[0050] ここで、未試行のバリエーションがある場合（ステップＳ１０７、Ｙｅｓ）、検知装置１０は、ステップＳ１０３に戻り、未試行のバリエーションを選択して処理を繰り返す。一方、未試行のバリエーションがない場合（ステップＳ１０７、Ｎｏ）、検知装置１０は、次の処理へ進む。

[0051] 検知装置１０は、ランダム性のパターンを設定した上で（ステップＳ１０８）、生成モデルを用いて学習データからデータを再構成する（ステップＳ１０９）。そして、検知装置１０は、再構成誤差から異常度を計算する（ステップＳ１１０）。

[0052] ここで、未試行のパターンがある場合（ステップＳ１１１、Ｙｅｓ）、検知装置１０は、ステップＳ１０８に戻り、未試行のパターンを設定して処理

を繰り返す。一方、未試行のパターンがない場合（ステップＳ１１１、Ｎｏ）、検知装置１０は、次の処理へ進む。

[0053] 検知装置１０は、各パターンの生成モデルの相関の大きさを計算し、相関が大きい生成モデル群の中から生成モデルを選択する（ステップＳ１１２）。

[0054] 図１５を用いて、検知装置１０の検知処理の流れについて説明する。図１５は、第１の実施形態に係る検知装置の検知処理の流れを示すフローチャートである。図１５に示すように、まず、検知装置１０は、テストデータの入力を受け付ける（ステップＳ２０１）。次に、検知装置１０は、逓増又は逓減する特徴量のデータを変換する（ステップＳ２０２）。例えば、検知装置１０は、各データを所定の時刻間の差又は比に変換する。

[0055] 検知装置１０は、学習時と同じ手法でテストデータを正規化する（ステップＳ２０３）。そして、検知装置１０は、生成モデルを用いてテストデータからデータを再構成する（ステップＳ２０４）。ここで、検知装置１０は、学習データにおいて変動が小さい特徴量を特定する（ステップＳ２０５）。このとき、検知装置１０は、特定した特徴量を異常度の計算対象から除外してもよい。そして、検知装置１０は、再構成誤差から異常度を計算する（ステップＳ２０６）。さらに、検知装置１０は、異常度を基に異常を検知する（ステップＳ２０７）。

[0056] ［第１の実施形態の効果］

前処理部１３１は、学習用のデータ及び検知対象のデータを加工する。また、生成部１３２は、前処理部１３１によって加工された学習用のデータを基に、深層学習によりモデルを生成する。また、検知部１３３は、前処理部１３１によって加工された検知対象のデータをモデルに入力して得られた出力データを基に異常度を計算し、異常度を基に検知対象のデータの異常を検知する。このように、実施形態によれば、深層学習を使って異常検知を行う場合の学習データの前処理や選定、モデルの選択を適切に行えるようになり、検知精度を向上させることができる。

[0057] 前処理部１３１は、特徴量の時系列データである学習用のデータから、時間に対する変動の大きさの度合いが所定値以下である特徴量を特定する。また、検知部１３３は、検知対象のデータの特徴量のうち、前処理部１３１によって特定された特徴量、又は、前処理部１３１によって特定された特徴量以外の特徴量の少なくともいずれかを基に異常を検知する。これにより、検知装置１０は、検知精度を低下させるデータを除外することができる。

[0058] 前処理部１３１は、学習用のデータ及び検知対象のデータの一部又は全てを、当該データの所定の時刻間の差又は比に変換する。これにより、検知装置１０は、学習データが特徴量の取りうる範囲を網羅していなくても誤検知を抑制できる。また、上昇または下降のトレンド成分の影響を取り除くことで、時間による値の範囲の変化の影響を抑えることができる。

[0059] 生成部１３２は、深層学習にオートエンコーダを用いる。これにより、検知装置１０は、再構成誤差による異常度の計算及び異常検知を行うことができるようになる。

[0060] 生成部１３２は、学習用のデータに対して複数回学習を行う。また、検知部１３３は、生成部１３２によって生成された各モデルのうち、互いの関係の強さに応じて選択されたモデルを用いて異常を検知する。これにより、検知装置１０は、最適なモデルを選択することができる。

[0061] 前処理部１３１は、時系列データである学習用のデータを所定の期間ごとのスライディングウィンドウで分割する。また、生成部１３２は、前処理部１３１によって分割されたスライディングウィンドウごとのデータのそれぞれを基に、モデルを生成する。これにより、検知装置１０は、データ分布の変化に早期に追従したモデルを生成することができ、データ分布が変化することの影響による誤検知を抑えることができる。

[0062] 前処理部１３１は、学習用のデータに対する複数の異なる正規化手法ごとに生成されたモデル群、又は、それぞれに異なるモデルパラメータが設定されたモデル群のうちの少なくとも一方のモデル群を使って計算された異常度が所定の値より高いデータを学習用のデータから除外する。これにより、検

知装置１０は、検知精度を低下させるデータを除外することができる。

[0063] ［検知装置の出力例］

ここで、検知装置１０による検知結果の出力例について説明する。検知装置１０は、例えばテキストログ及び数値ログの学習及び検知を行うことができる。例えば、数値ログの特徴量は、各種センサが計測した数値及び数値に統計的な処理を施した値である。また、例えば、テキストログの特徴量は、各メッセージを分類してＩＤを付与し、一定時刻ごとの各ＩＤの出現頻度を表す値である。

[0064] 以降の出力結果を得るための設定等について説明する。まず、使用したデータは、OpenStack系のシステムの３つのコントローラノードから取得した数値ログ（約３５０メトリクス）及びテキストログ（約３０００～４５００ＩＤ）である。また、データの収集期間は５／１～６／３０であり、収集間隔は５分である。また、期間中に、メンテナンス日を含めて８回の異常イベントが発生した。

[0065] 検知装置１０は、コントローラノードごとにモデルを生成した。また、検知装置１０は、各モデルを使って検知を行った。学習期間は５／１～６／５である。また、検知の対象となる評価期間は、５／１～６／３０である。

[0066] 図１６は、テキストログの異常度の一例を示す図である。図１６に示すように、検知装置１０は、メンテナンスがあった５／１２や障害が発生した６／１９に高い異常度を出力している。また、図１７は、テキストログの異常度と障害情報との関係の一例を示す図である。図１７に示すように、検知装置１０は、異常が発生した５／７や６／１９に高い異常度を出力している。

[0067] 図１８は、障害発生時のテキストログ及び特徴量ごとの異常度の一例を示す図である。なお、outlierが特徴量ごとの異常度を表しており、値が大きな上位１０のログメッセージを示している。図１８に示すように、rabbit関連の障害が発生した６／１９の該当時刻のログメッセージを見ると、rabbitに関する内容が多く、一部ではERRORが記されている。ここから、rabbitで何かあったことが原因ではないかと推測することが可能となる。

[0068] 図１９は、異常度が上昇したときのテキストログ及び特徴ごとの異常度の一例を示す図である。また、図２０は、異常度が上昇した前後の時刻のテキストログＩＤのデータ分布の一例を示す図である。図１９に示すように、上位１０のログの特徴ごとの異常度は一致しており、これは４００以上のログで同じ値であった。また、図２０に示すように、各コントローラで発生したテキストログのＩＤが類似しており、１０：３１以前の時刻にこれらのＩＤは全く出現していなかったことがわかる。これより、１０：３１に普段出ないログが大量に出力されるような異常が発生したことが示唆されている。

[0069] 図２１は、数値ログの異常度の一例を示す図である。図２１に示すように、検知装置１０は、メンテナンスがあった５／１２や障害が発生した５／２０に高い異常度を出力している。また、図２２は、数値ログの異常度と障害情報との関係の一例を示す図である。図２２に示すように、検知装置１０は、異常が発生した６／１４や６／１９に高い異常度を出力している。

[0070] 図２３は、障害発生時の数値ログ及び特徴量ごとの異常度の一例を示す図である。図２３に示すように、rabbit関連の障害が発生した６／１９の同時刻に検知装置１０はrabbitのメモリ関連の特徴量が高い異常度を出力している。図２４は、６／１９の同時刻前後の数値ログの特徴量ごとの入力データと再構成データの一例を示す図である。originalが入力データを、reconstructが再構成データを表す。図２４に示すように、特徴量ごとの異常度が一番大きなtop1のrabbitのメモリ関連の特徴量は、該当時刻に入力データが大幅に小さな値となったが、その変化がうまく再構成できていないことがわかる。一方、残りの３つの特徴量は、該当時刻の再構成データが大きく上昇している様子が確認される。これは、rabbitのメモリ関連の特徴量が下降したときは、残りの３つの特徴量は上昇すると学習された結果であるが、下降度合いが学習期間の想定以上に大きかったため、異常度が大きくなったと推測される。

[0071] ［その他の実施形態］

これまで、生成部１３２が深層学習にオートエンコーダを用いる場合の実

施形態を説明した。一方で、生成部１３２は、深層学習にリカレントニューラルネットワーク（以降、ＲＮＮ）を用いてもよい。つまり、生成部１３２は、深層学習にオートエンコーダ又はＲＮＮを用いる。

[0072] ＲＮＮは、時系列データを入力とするニューラルネットワークである。例えば、ＲＮＮを使った異常検知方法には、予測モデルを構築する方法と、sequence-to-sequenceのオートエンコーダモデルを構築する方法がある。また、ここでいうＲＮＮには、単純なＲＮＮだけでなく、ＲＮＮのバリエーションであるＬＳＴＭやＧＲＵも含まれる。

[0073] 予測モデルを構築する方法では、検知装置１０は、再構成誤差の代わりに、元のデータの値と予測値の誤差を基に異常を検知する。例えば、予測値は、所定の期間の時系列データを入力した場合のＲＮＮの出力値であり、ある時刻の時系列データの推定値である。検知装置１０は、実際に収集されたある時刻のデータと当該時刻の予測値との誤差の大きさを基に異常を検知する。例えば、検知装置１０は、誤差の大きさが閾値を超えている場合に、当該時刻に異常が発生したことを検知する。

[0074] sequence-to-sequenceのオートエンコーダモデルを構築する方法は、オートエンコーダを構築する点は第１の実施形態と共通しているが、ニューラルネットワークがＲＮＮである点、及び入力データ及び出力データ（再構成データ）が時系列データである点が異なる。この場合、検知装置１０は、時系列データの再構成誤差を異常度とみなし、異常を検知することができる。

[0075] ［システム構成等］

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

アとして実現され得る。

[0076] また、本実施形態において説明した各処理のうち、自動的に行われるものとして説明した処理の全部又は一部を手動的に行うこともでき、あるいは、手動的に行われるものとして説明した処理の全部又は一部を公知の方法で自動的に行うこともできる。この他、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、特記する場合を除いて任意に変更することができる。

[0077] ［プログラム］

一実施形態として、検知装置１０は、パッケージソフトウェアやオンラインソフトウェアとして上記の検知を実行する検知プログラムを所望のコンピュータにインストールさせることによって実装できる。例えば、上記の検知プログラムを情報処理装置に実行させることにより、情報処理装置を検知装置１０として機能させることができる。ここで言う情報処理装置には、デスクトップ型又はノート型のパーソナルコンピュータが含まれる。また、その他にも、情報処理装置にはスマートフォン、携帯電話機やＰＨＳ（Personal Handyphone System）等の移動体通信端末、さらには、ＰＤＡ（Personal Digital Assistant）等のスレート端末等がその範疇に含まれる。

[0078] また、検知装置１０は、ユーザが使用する端末装置をクライアントとし、当該クライアントに上記の検知に関するサービスを提供する検知サーバ装置として実装することもできる。例えば、検知サーバ装置は、学習データを入力とし、生成モデルを出力とする検知サービスを提供するサーバ装置として実装される。この場合、検知サーバ装置は、Ｗｅｂサーバとして実装することとしてもよいし、アウトソーシングによって上記の検知に関するサービスを提供するクラウドとして実装することとしてもかまわない。

[0079] 図２５は、検知プログラムを実行するコンピュータの一例を示す図である。コンピュータ１０００は、例えば、メモリ１０１０、ＣＰＵ１０２０を有する。また、コンピュータ１０００は、ハードディスクドライブインタフェース１０３０、ディスクドライブインタフェース１０４０、シリアルポート

インタフェース１０５０、ビデオアダプタ１０６０、ネットワークインタフェース１０７０を有する。これらの各部は、バス１０８０によって接続される。

[0080] メモリ１０１０は、ＲＯＭ（Read Only Memory）１０１１及びＲＡＭ１０１２を含む。ＲＯＭ１０１１は、例えば、ＢＩＯＳ（Basic Input Output System）等のブートプログラムを記憶する。ハードディスクドライブインタフェース１０３０は、ハードディスクドライブ１０９０に接続される。ディスクドライブインタフェース１０４０は、ディスクドライブ１１００に接続される。例えば磁気ディスクや光ディスク等の着脱可能な記憶媒体が、ディスクドライブ１１００に挿入される。シリアルポートインタフェース１０５０は、例えばマウス１１１０、キーボード１１２０に接続される。ビデオアダプタ１０６０は、例えばディスプレイ１１３０に接続される。

[0081] ハードディスクドライブ１０９０は、例えば、ＯＳ１０９１、アプリケーションプログラム１０９２、プログラムモジュール１０９３、プログラムデータ１０９４を記憶する。すなわち、検知装置１０の各処理を規定するプログラムは、コンピュータにより実行可能なコードが記述されたプログラムモジュール１０９３として実装される。プログラムモジュール１０９３は、例えばハードディスクドライブ１０９０に記憶される。例えば、検知装置１０における機能構成と同様の処理を実行するためのプログラムモジュール１０９３が、ハードディスクドライブ１０９０に記憶される。なお、ハードディスクドライブ１０９０は、ＳＳＤにより代替されてもよい。

[0082] また、上述した実施形態の処理で用いられる設定データは、プログラムデータ１０９４として、例えばメモリ１０１０やハードディスクドライブ１０９０に記憶される。そして、ＣＰＵ１０２０は、メモリ１０１０やハードディスクドライブ１０９０に記憶されたプログラムモジュール１０９３やプログラムデータ１０９４を必要に応じてＲＡＭ１０１２に読み出して、上述した実施形態の処理を実行する。

[0083] なお、プログラムモジュール１０９３やプログラムデータ１０９４は、ハ

ードディスクドライブ１０９０に記憶される場合に限らず、例えば着脱可能な記憶媒体に記憶され、ディスクドライブ１１００等を介してＣＰＵ１０２０によって読み出されてもよい。あるいは、プログラムモジュール１０９３及びプログラムデータ１０９４は、ネットワーク（ＬＡＮ（Local Area Network）、ＷＡＮ（Wide Area Network）等）を介して接続された他のコンピュータに記憶されてもよい。そして、プログラムモジュール１０９３及びプログラムデータ１０９４は、他のコンピュータから、ネットワークインタフェース１０７０を介してＣＰＵ１０２０によって読み出されてもよい。

符号の説明

[0084] １０　検知装置

１１　入出力部

１２　記憶部

１３　制御部

１２１　モデル情報

１３１　前処理部

１３２　生成部

１３３　検知部

請求の範囲

[請求項1] 学習用のデータ及び検知対象のデータを加工する前処理部と、

前記前処理部によって加工された学習用のデータを基に、深層学習によりモデルを生成する生成部と、

前記前処理部によって加工された検知対象のデータを前記モデルに入力して得られた出力データを基に異常度を計算し、前記異常度を基に前記検知対象のデータの異常を検知する検知部と、

を有することを特徴とする検知装置。

[請求項2] 前記前処理部は、特徴量の時系列データである前記学習用のデータから、時間に対する変動の大きさの度合いが所定値以下である特徴量を特定し、

前記検知部は、前記検知対象のデータの特徴量うち、前記前処理部によって特定された特徴量、又は、前記前処理部によって特定された特徴量以外の特徴量の少なくともいずれかを基に異常を検知することを特徴とする請求項１に記載の検知装置。

[請求項3] 前記前処理部は、前記学習用のデータ及び検知対象のデータの一部又は全てを、当該データの所定の時刻間の差又は比に変換することを特徴とする請求項１に記載の検知装置。

[請求項4] 前記生成部は、深層学習にオートエンコーダ又はリカレントニューラルネットワークを用いることを特徴とする請求項１から３のいずれか１項に記載の検知装置。

[請求項5] 前記生成部は、前記学習用のデータに対して複数回学習を行い、

前記検知部は、前記生成部によって生成された各モデルのうち、互いの関係の強さに応じて選択されたモデルを用いて異常を検知することを特徴とする請求項４に記載の検知装置。

[請求項6] 前記前処理部は、時系列データである前記学習用のデータを所定の期間ごとのスライディングウィンドウで分割し、

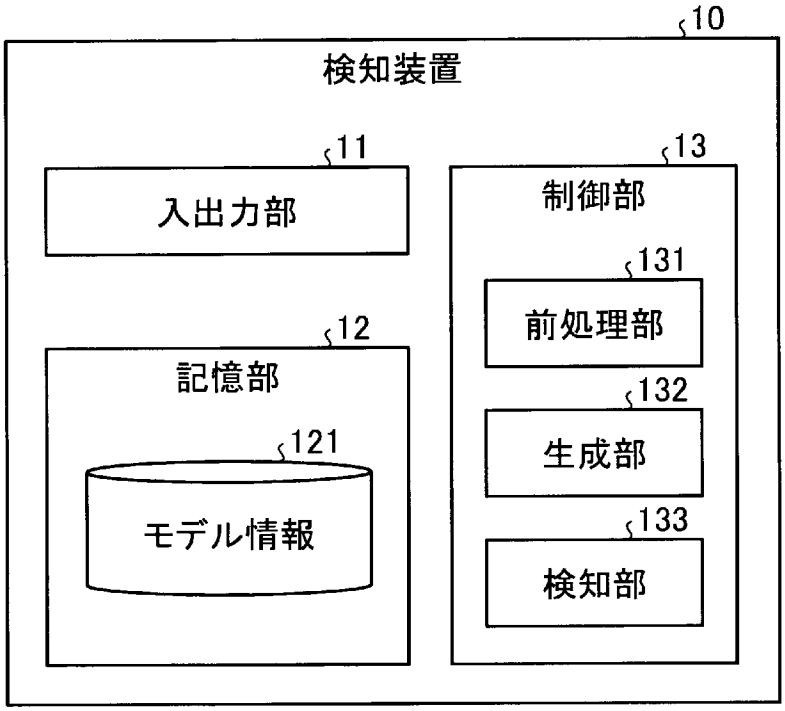
前記生成部は、前記前処理部によって分割されたスライディングウ

ィンドウごとのデータのそれぞれを基に、モデルを生成することを特徴とする請求項４又は５に記載の検知装置。

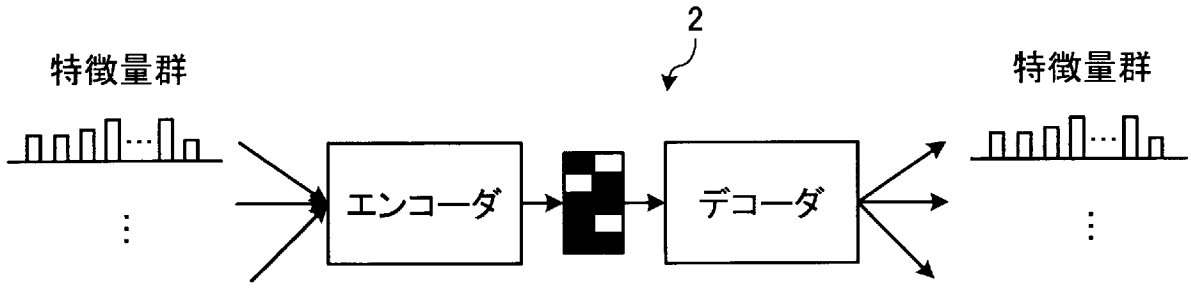
[請求項7] 前記前処理部は、前記学習用のデータに対する複数の異なる正規化手法ごとに生成されたモデル群、又は、それぞれに異なるモデルパラメータが設定されたモデル群のうちの少なくとも一方のモデル群に含まれる少なくとも１つのモデルを使って計算された異常度が所定の値より高いデータを前記学習用のデータから除外することを特徴とする請求項４から６のいずれか１項に記載の検知装置。

[請求項8] コンピュータを、請求項１から７のいずれか１項に記載の検知装置として機能させるための検知プログラム。

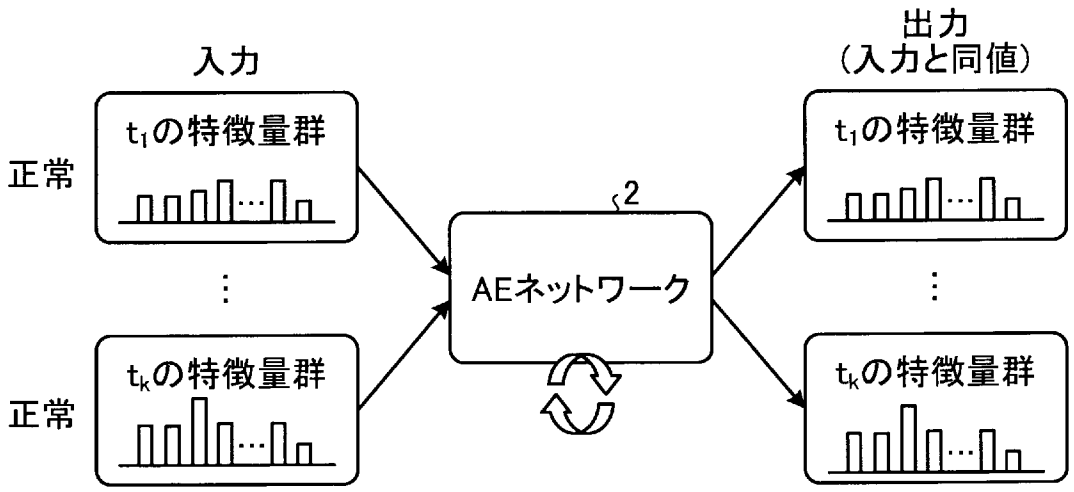
[図1]


10
検知装置
11
入出力部
12
記憶部
121
モデル情報
13
制御部
131
前処理部
132
生成部
133
検知部

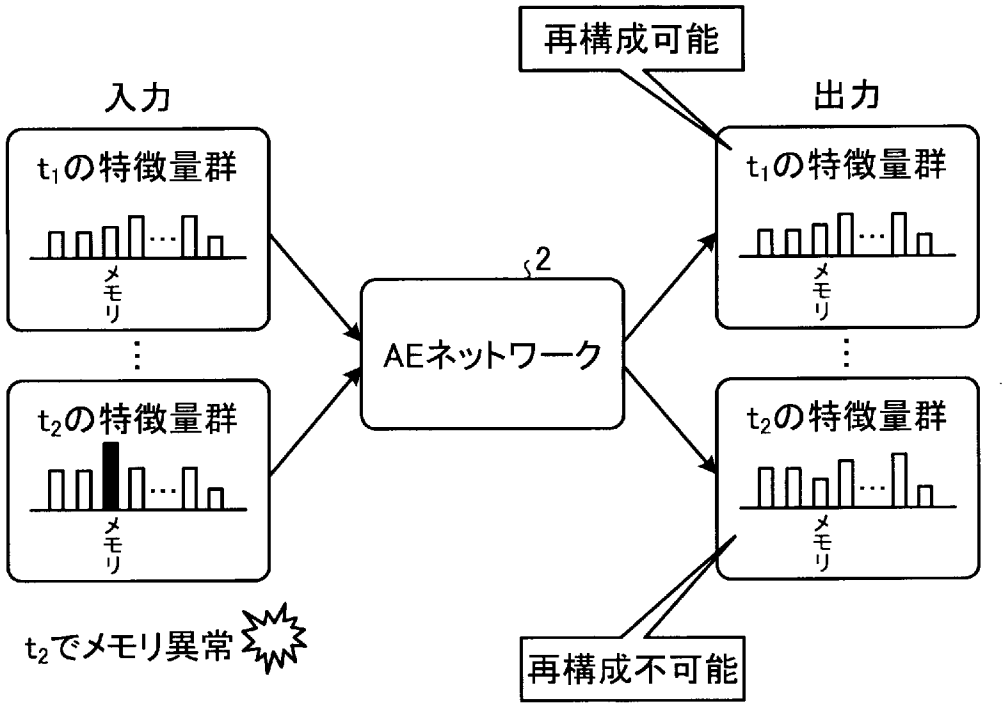
[図2]


2
特徴量群
エンコーダ
デコーダ
特徴量群

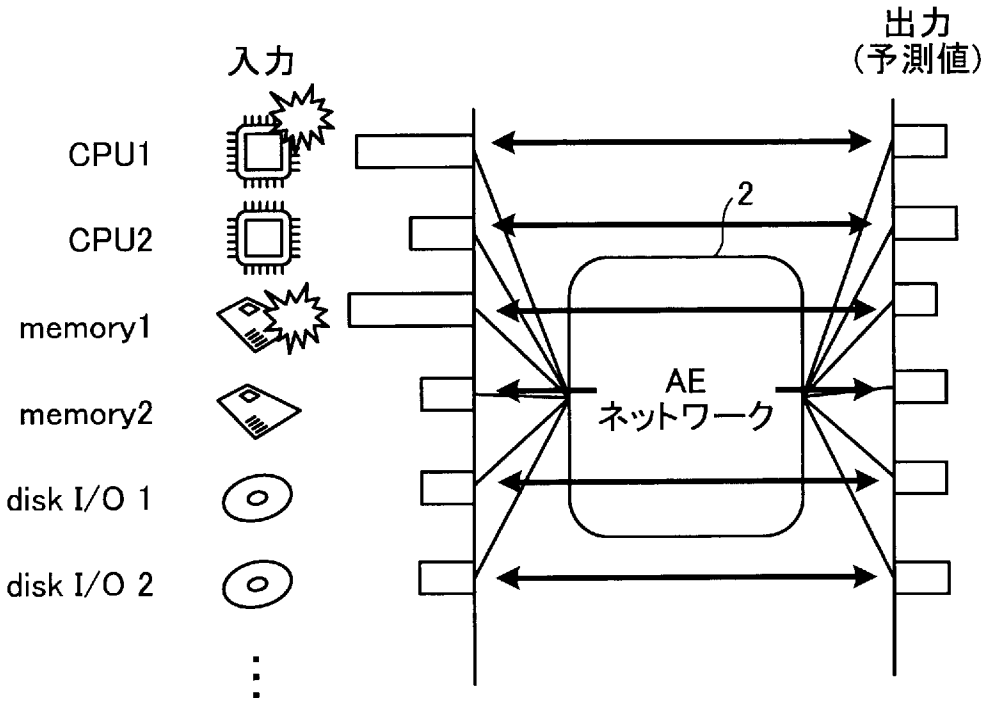
[図3]


入力
出力
(入力と同値)
正常
t_1の特徴量群
正常
t_kの特徴量群
2
AEネットワーク
t_1の特徴量群
t_kの特徴量群

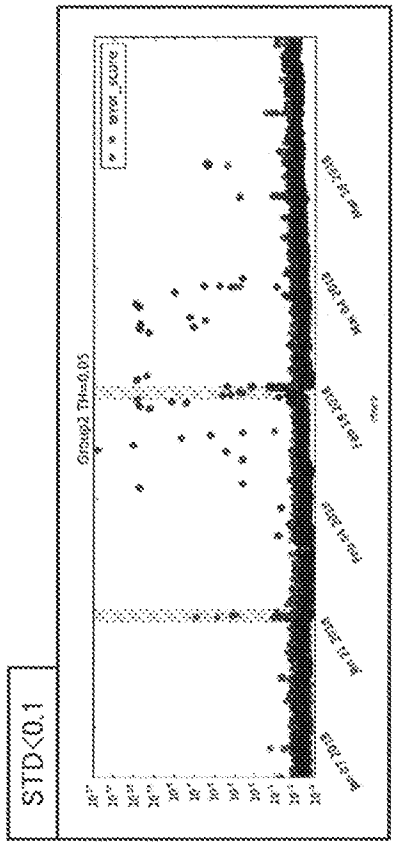
[図4]


再構成可能
入力
出力
t_1の特徴量群
メモリ
t_2の特徴量群
メモリ
2
AEネットワーク
t_1の特徴量群
メモリ
t_2の特徴量群
メモリ
t_2でメモリ異常
再構成不可能

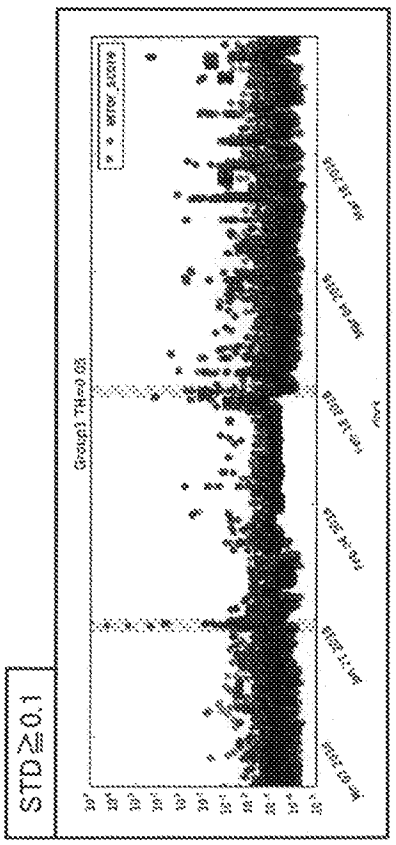
[図5]


入力
出力
(予測値)
CPU1
CPU2
memory1
memory2
disk I/O 1
disk I/O 2
2
AE
ネットワーク

[図6]

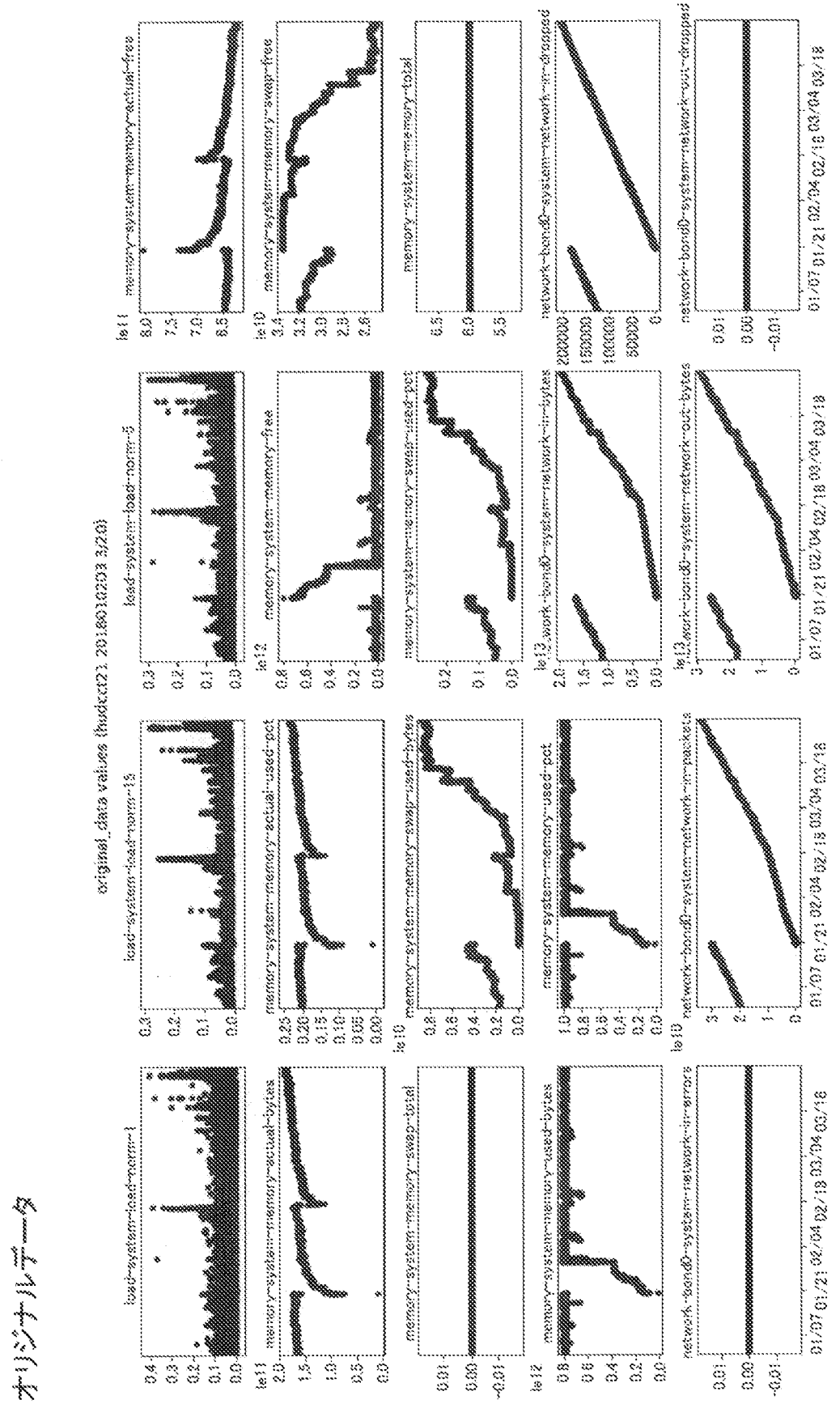


No.	日時	異常度
1	2018-02-09 21:54:23	3.728099e+16
2	2018-02-09 21:49:23	3.728099e+16
3	2018-02-10 11:54:23	1.259676e+13
4	2018-02-10 11:59:23	1.259676e+13
5	2018-02-27 16:44:23	6.676465e+12

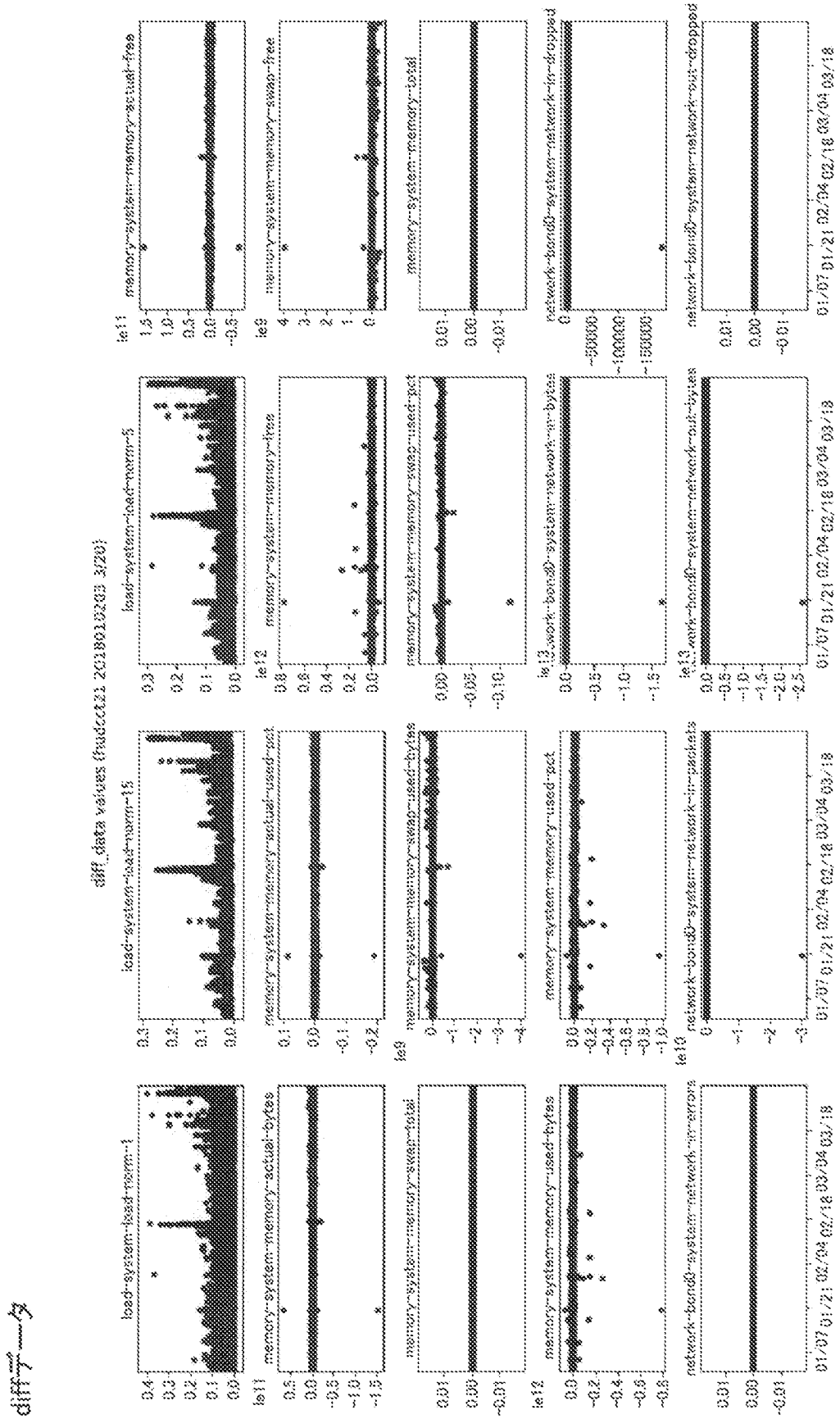


No.	日時	異常度
1	2018-01-20 11:39:23	20696.283079
2	2018-01-20 11:44:23	2232.494126
3	2018-03-29 16:44:23	280.962202
4	2018-03-29 16:39:23	280.959160
5	2018-01-20 11:59:23	239.251599

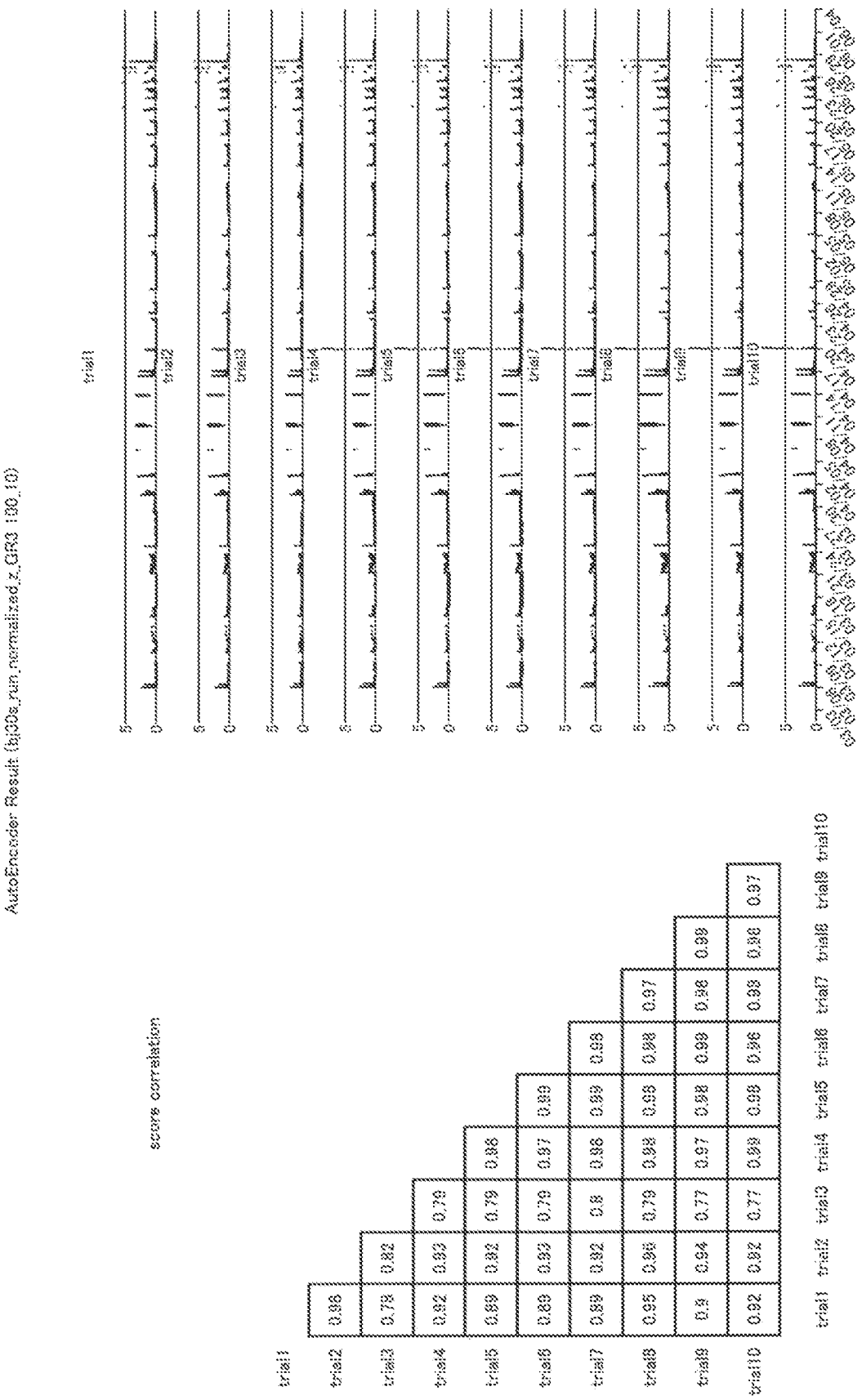
[図7]


オリジナルデータ
original_data values (index121 20180102 03:20)
load-system-load-norm-1
load-system-load-norm-15
load-system-load-norm-5
memory-system-memory-actual-free
memory-system-memory-actual-bytes
memory-system-memory-actual-used-pct
memory-system-memory-swap-used-bytes
memory-system-memory-swap-free
memory-system-memory-swap-total
memory-system-memory-used-bytes
memory-system-memory-used-pct
network-bond0-system-network-in-bytes
network-bond0-system-network-out-bytes
network-bond0-system-network-in-dropped
network-bond0-system-network-out-dropped
network-bond0-system-network-in-errors
network-bond0-system-network-in-packets
01/07 01/21 02/04 02/18 03/04 03/18

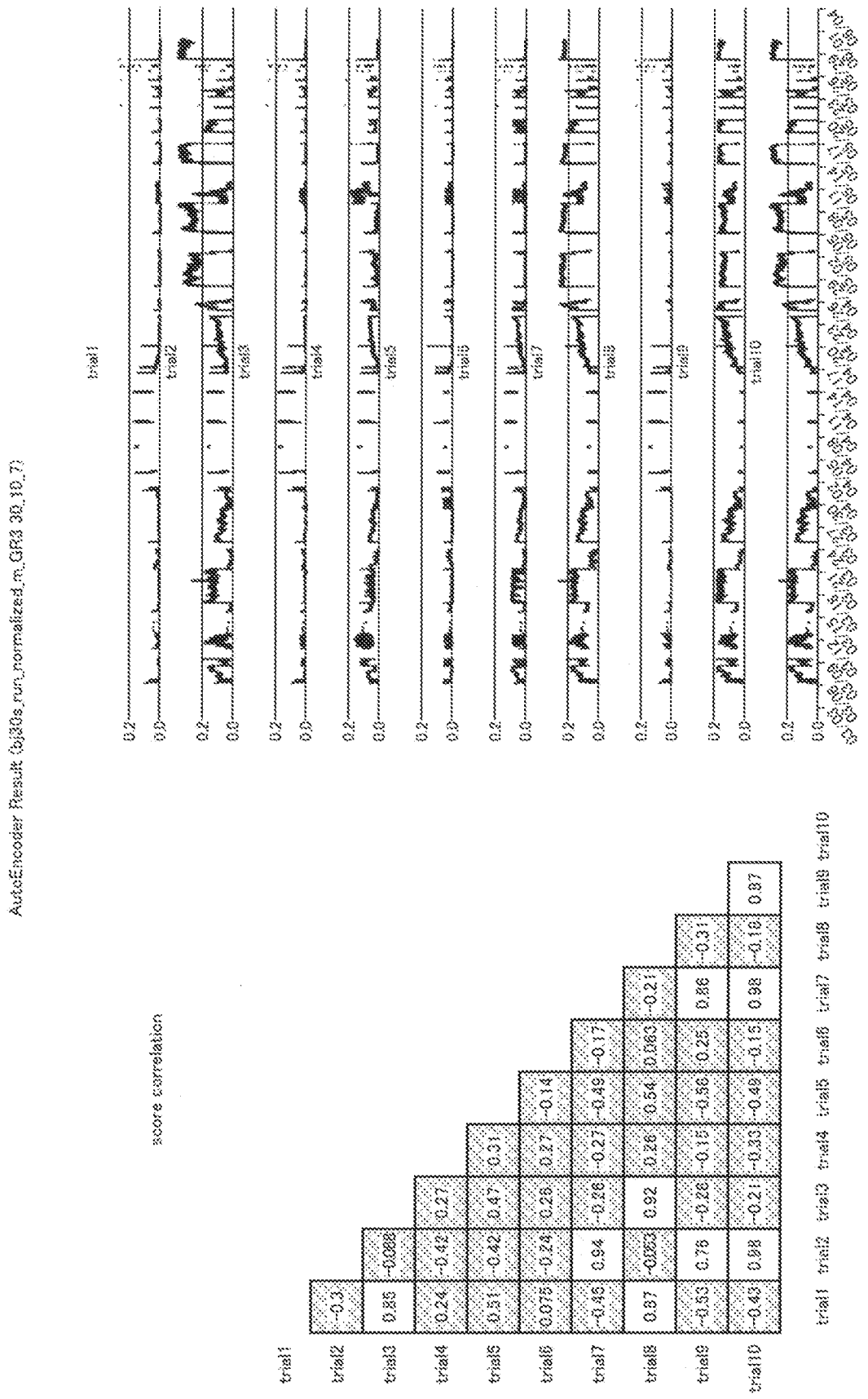
[図8]


diff_data values (fmdcc121 20180103203 3720)
load-system-load-norm-1
load-system-load-norm-15
load-system-load-norm-5
memory-system-memory-actual-free
memory-system-memory-actual-bytes
memory-system-memory-actual-used-pct
memory-system-memory-swap-free
memory-system-memory-swap-total
memory-system-memory-swap-used-bytes
memory-system-memory-swap-used-pct
memory-system-memory-total
memory-system-memory-used-bytes
memory-system-memory-used-pct
network-bond0-system-network-in-bytes
network-bond0-system-network-in-dropped
network-bond0-system-network-in-errors
network-bond0-system-network-in-packets
network-bond0-system-network-out-bytes
network-bond0-system-network-out-dropped
01/07 01/21 02/04 02/18 03/04 03/18

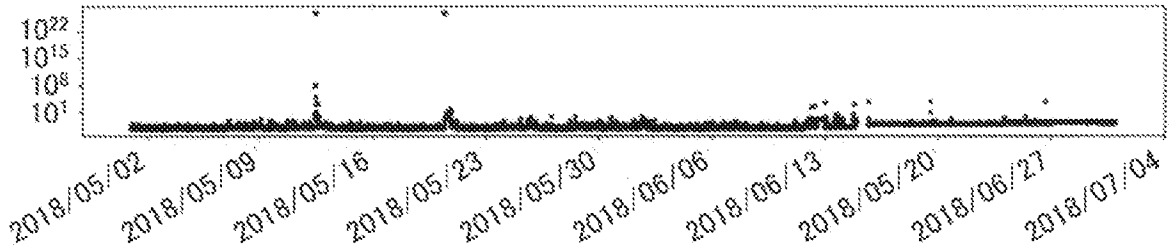
[図9]


AutoEncoder Result (bj30s_run_normalized_z_GR3 i30_10)
scors correlation
trial1
trial2
trial3
trial4
trial5
trial6
trial7
trial8
trial9
trial10
0.96
0.78 0.82
0.92 0.93 0.79
0.89 0.82 0.79 0.86
0.89 0.92 0.79 0.87 0.89
0.89 0.92 0.8 0.86 0.89 0.93
0.95 0.96 0.79 0.88 0.88 0.98 0.97
0.9 0.94 0.77 0.87 0.88 0.98 0.98 0.98
0.92 0.92 0.77 0.89 0.88 0.96 0.98 0.96 0.97

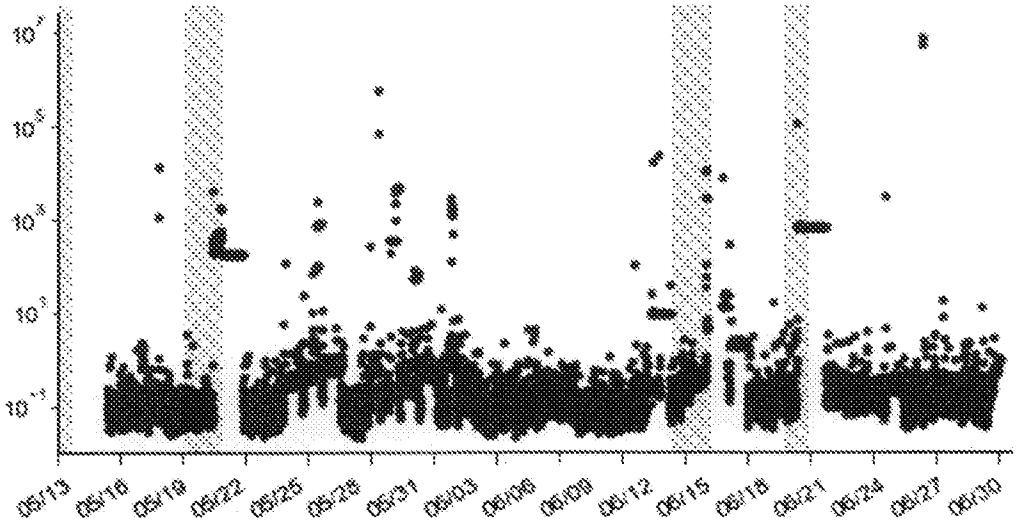
[図10]


AutoEncoder Result (bj30s_run_normalized_m_GR3 30_10_7)
score correlation
trial1
trial2
trial3
trial4
trial5
trial6
trial7
trial8
trial9
trial10

[図11]

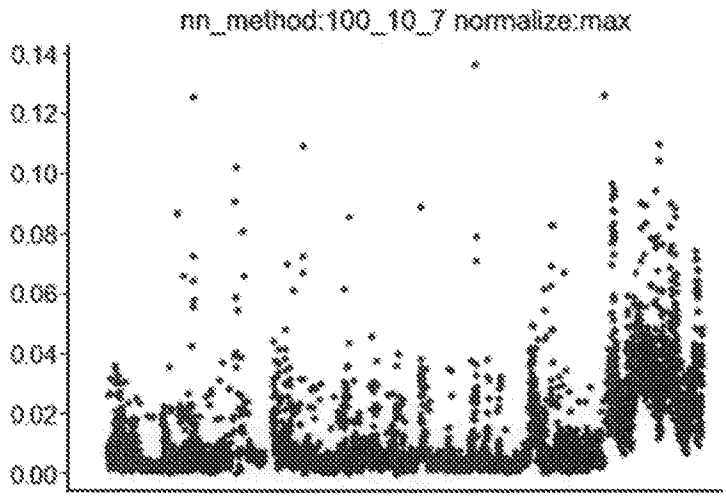

10^{22}
10^{15}
10^{8}
10^{1}
2018/05/02
2018/05/09
2018/05/16
2018/05/23
2018/05/30
2018/06/06
2018/06/13
2018/05/20
2018/06/27
2018/07/04

[図12]

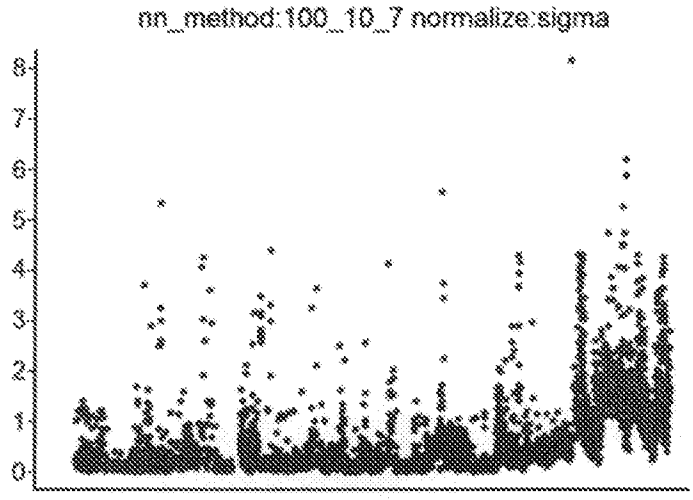

05/13
05/16
05/19
05/22
05/25
05/28
05/31
06/03
06/06
06/09
06/12
06/15
06/18
06/21
06/24
06/27
06/30

[図13]

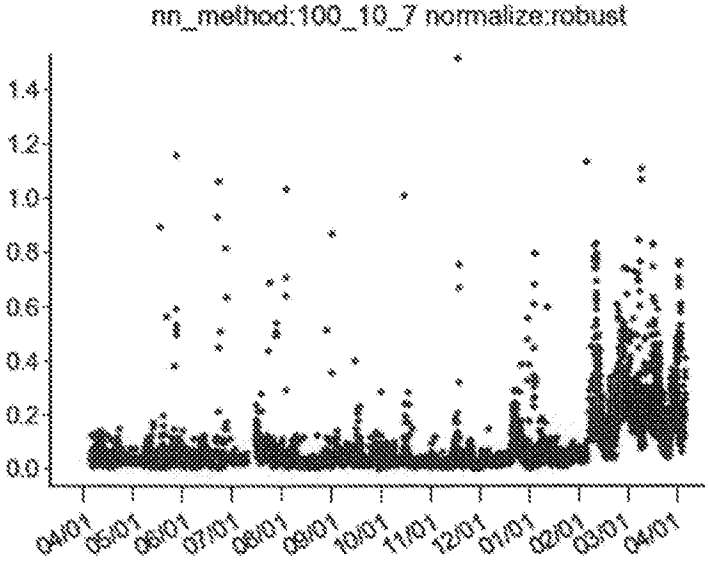
min-max正規化



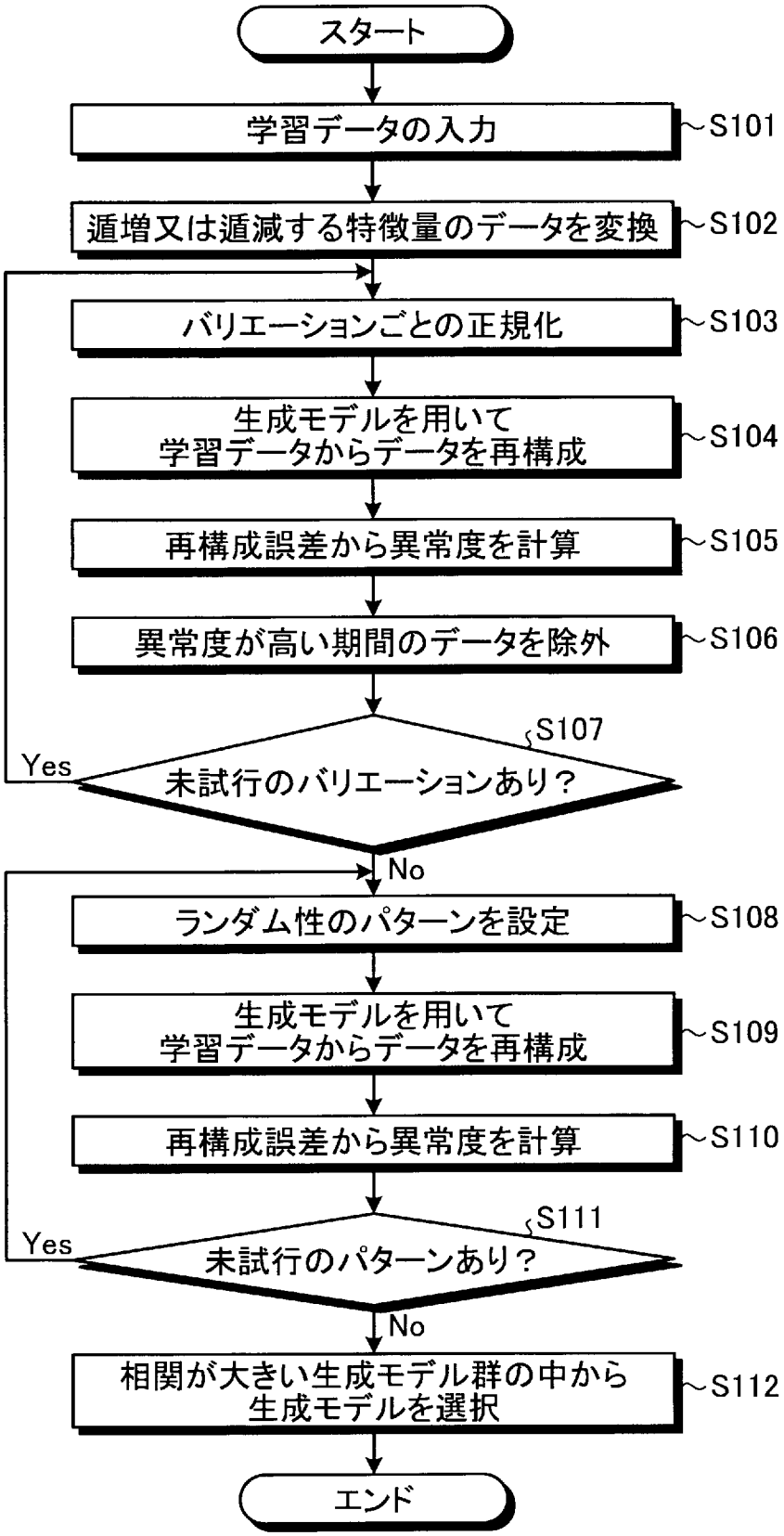
標準化
(Z-score正規化)



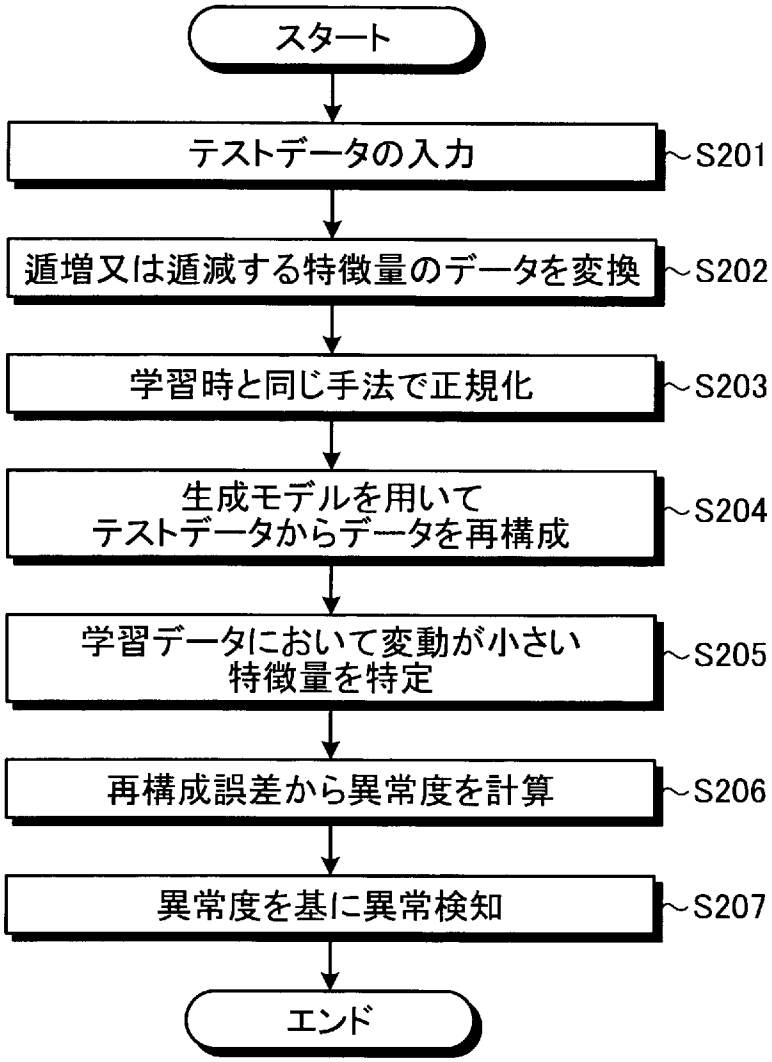
ロバスト正規化



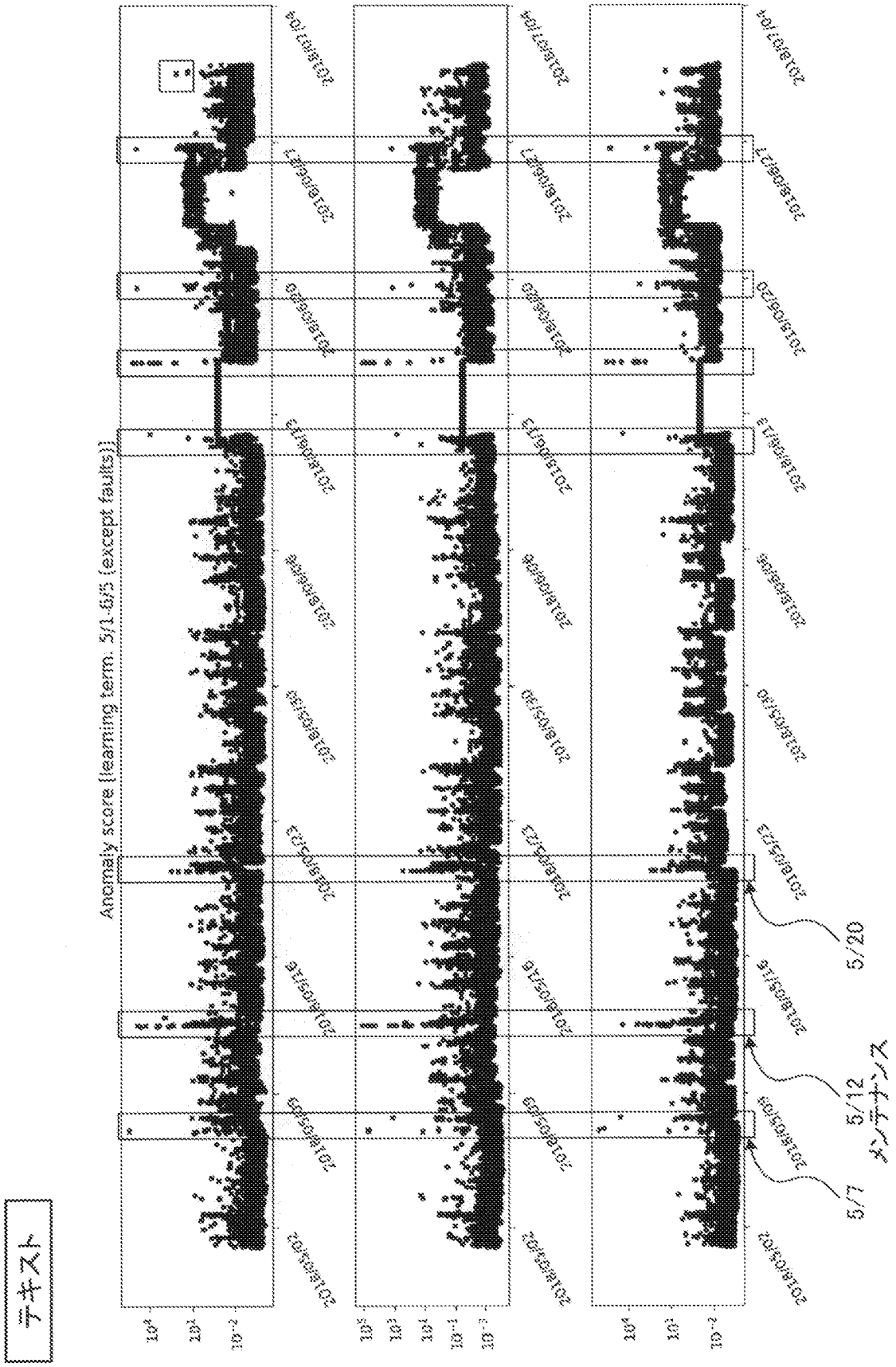
[図14]


スタート
学習データの入力
S101
遁増又は遁減する特徴量のデータを変換
S102
バリエーションごとの正規化
S103
生成モデルを用いて
学習データからデータを再構成
S104
再構成誤差から異常度を計算
S105
異常度が高い期間のデータを除外
S106
S107
未試行のバリエーションあり？
Yes
No
ランダム性のパターンを設定
S108
生成モデルを用いて
学習データからデータを再構成
S109
再構成誤差から異常度を計算
S110
S111
未試行のパターンあり？
Yes
No
相関が大きい生成モデル群の中から
生成モデルを選択
S112
エンド

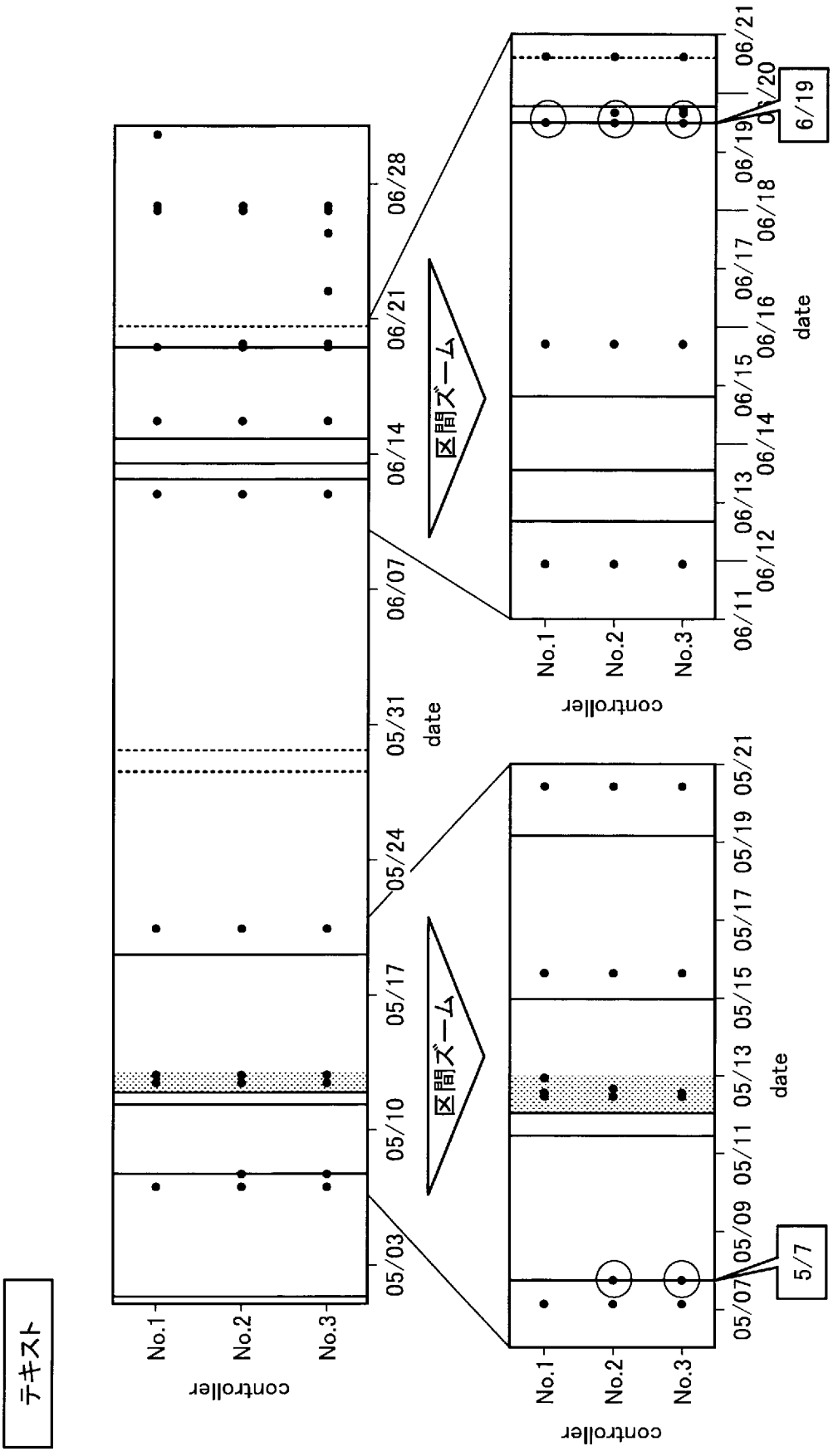
[図15]


スタート
テストデータの入力
S201
遁増又は遁減する特徴量のデータを変換
S202
学習時と同じ手法で正規化
S203
生成モデルを用いて
テストデータからデータを再構成
S204
学習データにおいて変動が小さい
特徴量を特定
S205
再構成誤差から異常度を計算
S206
異常度を基に異常検知
S207
エンド

[図16]


テキスト
Anomaly score (learning term: 5/1-6/5 (except faults))
5/20
5/12
5/7
タイムスタンプ

[図17]


テキスト
No.1
No.2
No.3
controller
05/03
05/10
05/17
05/24
05/31
06/07
06/14
06/21
06/28
date
区間ズーム
05/07
05/09
05/11
05/13
05/15
05/17
05/19
05/21
5/7
06/11
06/12
06/13
06/14
06/15
06/16
06/17
06/18
06/19
06/20
06/21
6/19

[図18]

テキスト

controller3

2018/6/19 12:47

	id	log message	outlier
1	1447	ERROR oslo.messaging._drivers. impl_rabbit AMQP server is unreachable CONNECTION_FORCED – broker forced connection closure with reason 'shutdown'. Trying again	1069.0
2	1430	ERROR oslo.messaging._drivers. impl_rabbit AMQP server is unreachable Broken pipe. Trying again	415.7
3	1424	INFO oslo.messaging._drivers.impl_rabbit A recoverable connection/channel error occurred, trying to reconnect	162.9
4	1425	INFO oslo.messaging._drivers.impl_rabbit A recoverable connection/channel error occurred, trying to reconnect Broken pipe	116.3
5	1462	ERROR oslo.messaging._drivers.impl_rabbit AMQP server is unreachable : ECONNREFUSED. Trying again	102.5
6	1443	Mirrored queue in vhost '/' : Slave saw deaths of mirrors	90.8
7	1422	ERROR oslo.messaging._drivers.impl_rabbit AMQP server is unreachable : CONNECTION_FORCED – broker forced connection closure with reason 'shutdown'. Trying again	74.0
8	1433	ERROR oslo.messaging._drivers.impl_rabbit ECONNREFUSED. Trying again	73.7
9	959	INFO oslo.messaging._drivers.impl_rabbit Reconnected to AMQP server via [amqp] clientwith port	71.4
10	175	accepting AMQP connection	69.9

[図19]

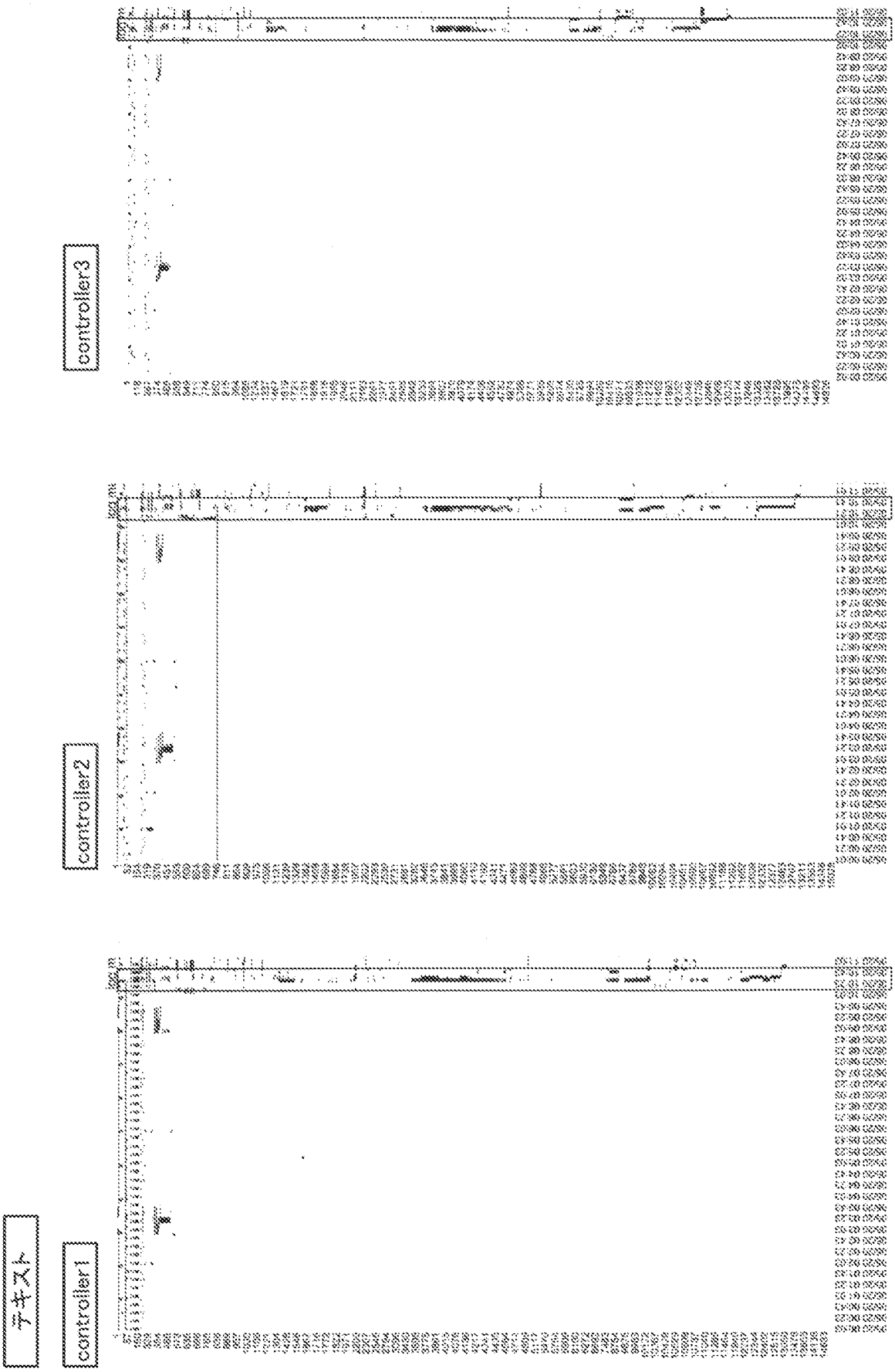
テキスト

controller2

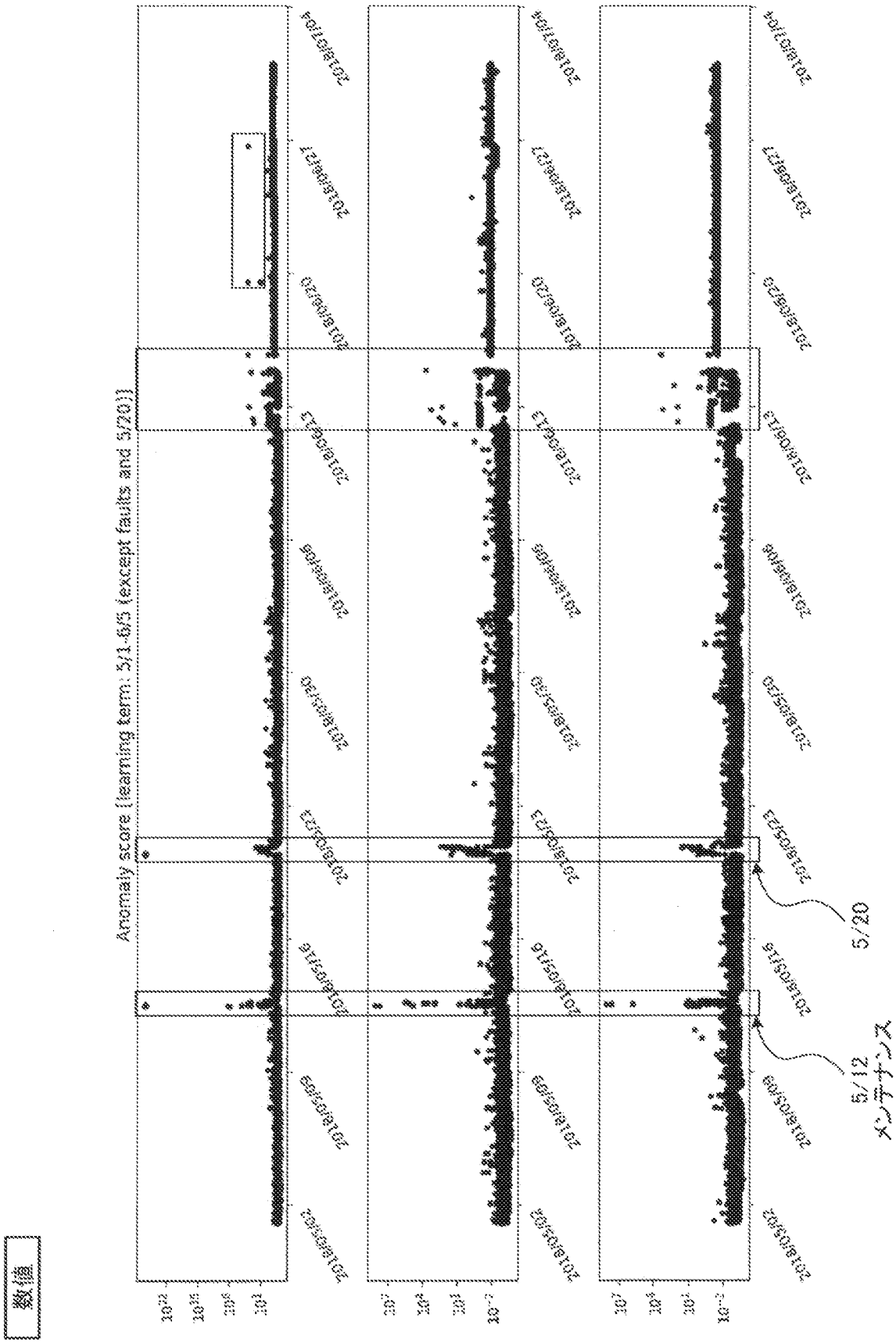
2018/5/20 10:31

	id	log message	outlier
1	12716	ERROR router_ids = self.plugin_rpc.get_router_ids (context)	94.3
2	12372	INFO heat.engine.resource deleting Port	94.3
3	4069	ERROR cinder.service model server went away・・・	94.3
4	12441	ERROR neutron.db.agentschedulers_db DBConnectionError : (pymysql.err.OperationalError) 'Lost connection to MySQL server during query')	94.3
5	10280	[Note] WSREP : State transfer to complete.	94.3
6	12668	ERROR neutron.agent.dhcp.agent raise result	94.3
7	11027	ERROR oslo_service.periodic_task QUOTAS.expire (context)	94.3
8	4075	ERROR cinder.service return self.dbapi.connect (*cargs, **cparams)	94.3
9	11043	ERROR oslo_service.periodic_task return self._execute_and_instances (context)	94.3
10	12487	ERROR oslo.service.loopingcall ectxt.value = e.inner_exc	94.3

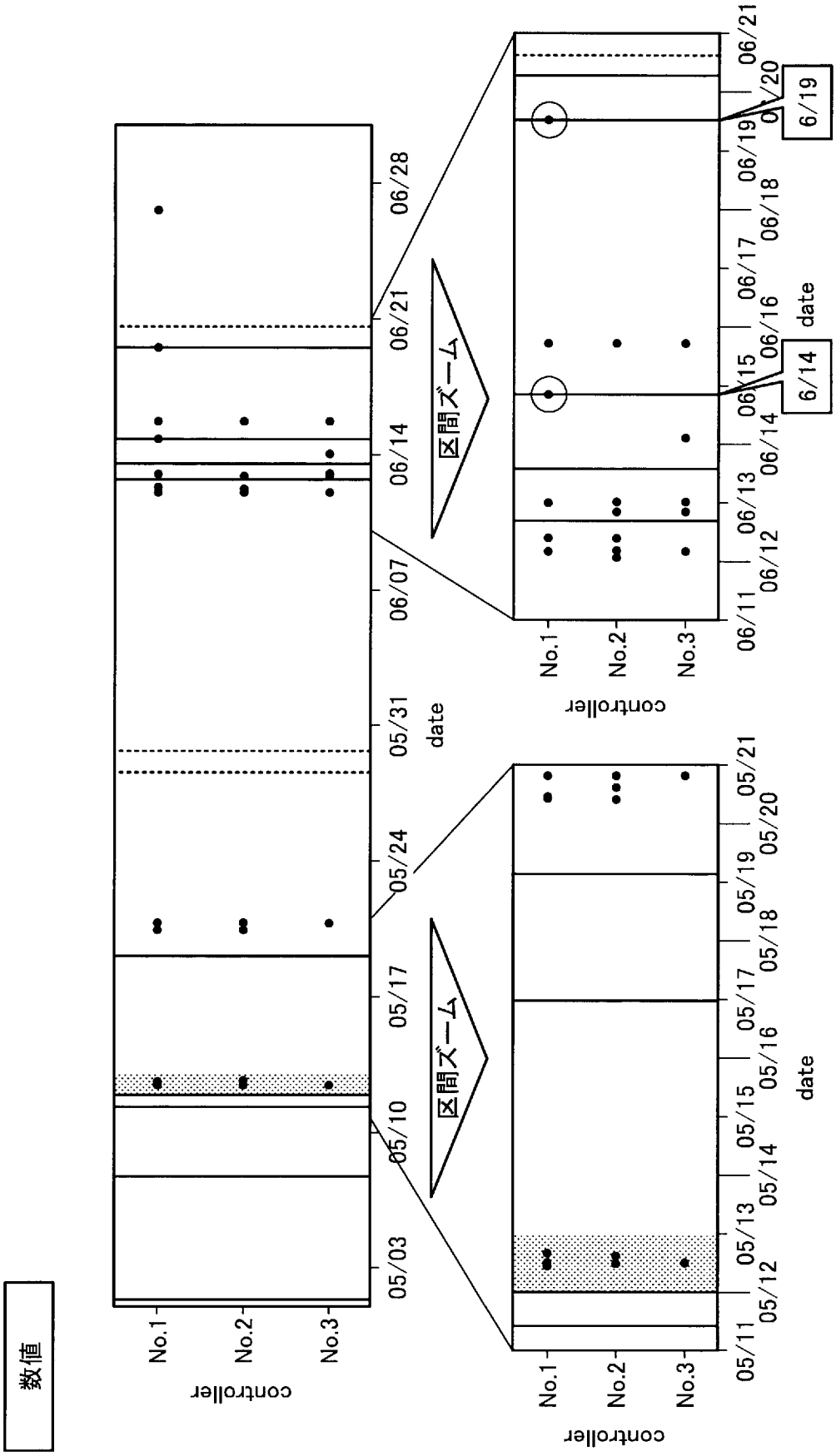
[図20]



[図21]


Anomaly score (learning term: 5/1-6/5 (except faults and 5/20))
5/20
5/12
メンテナンス
数値

[図22]


数値
controller
No.1
No.2
No.3
05/03
05/10
05/17
05/24
05/31
06/07
06/14
06/21
06/28
date
区間ズーム
06/11
06/12
06/13
06/14
06/15
06/16
06/17
06/18
06/19
06/20
06/21
6/14
6/19
05/11
05/12
05/13
05/14
05/15
05/16
05/17
05/18
05/19
05/20
05/21

[図23]

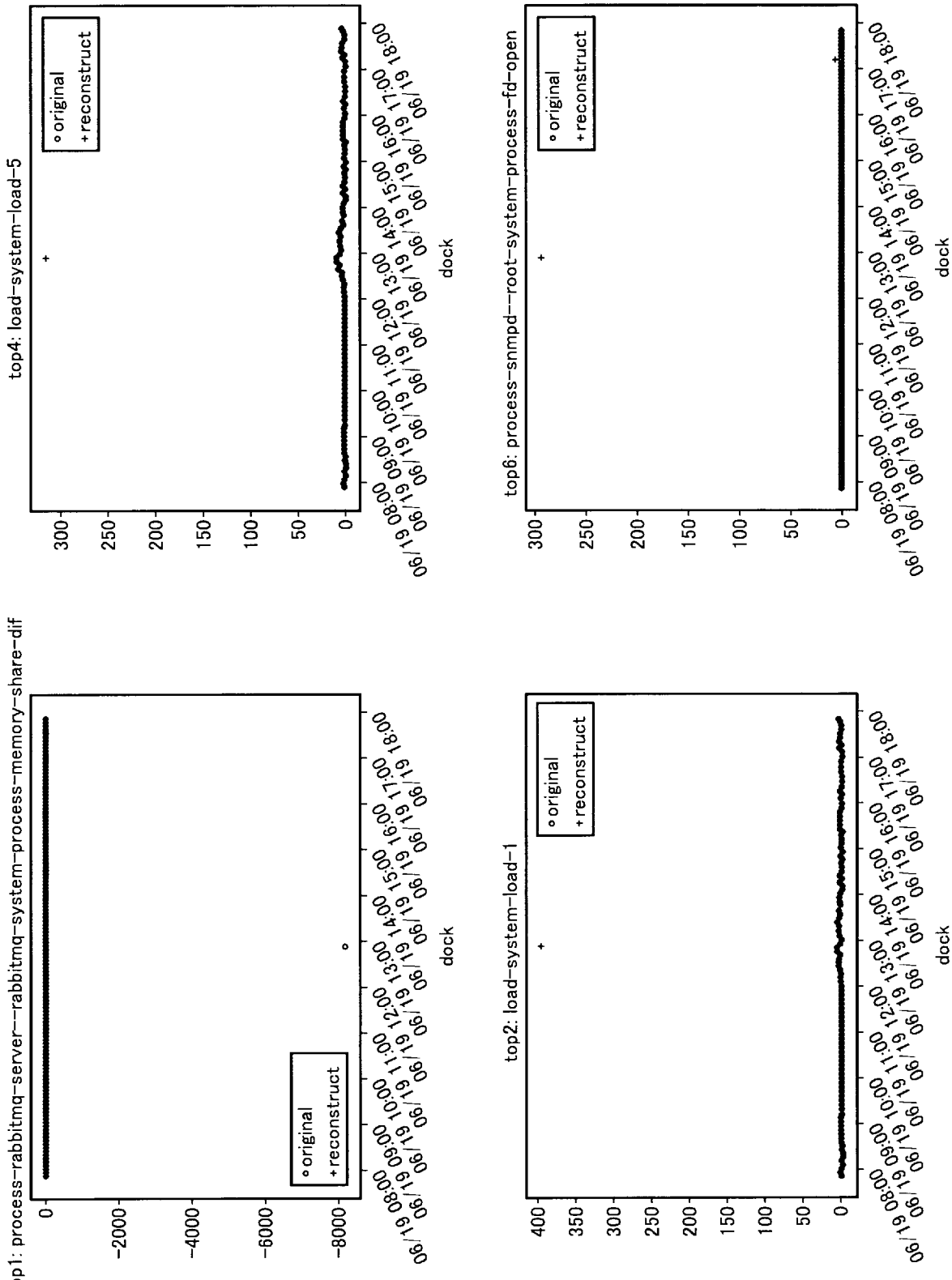
数値

controller1

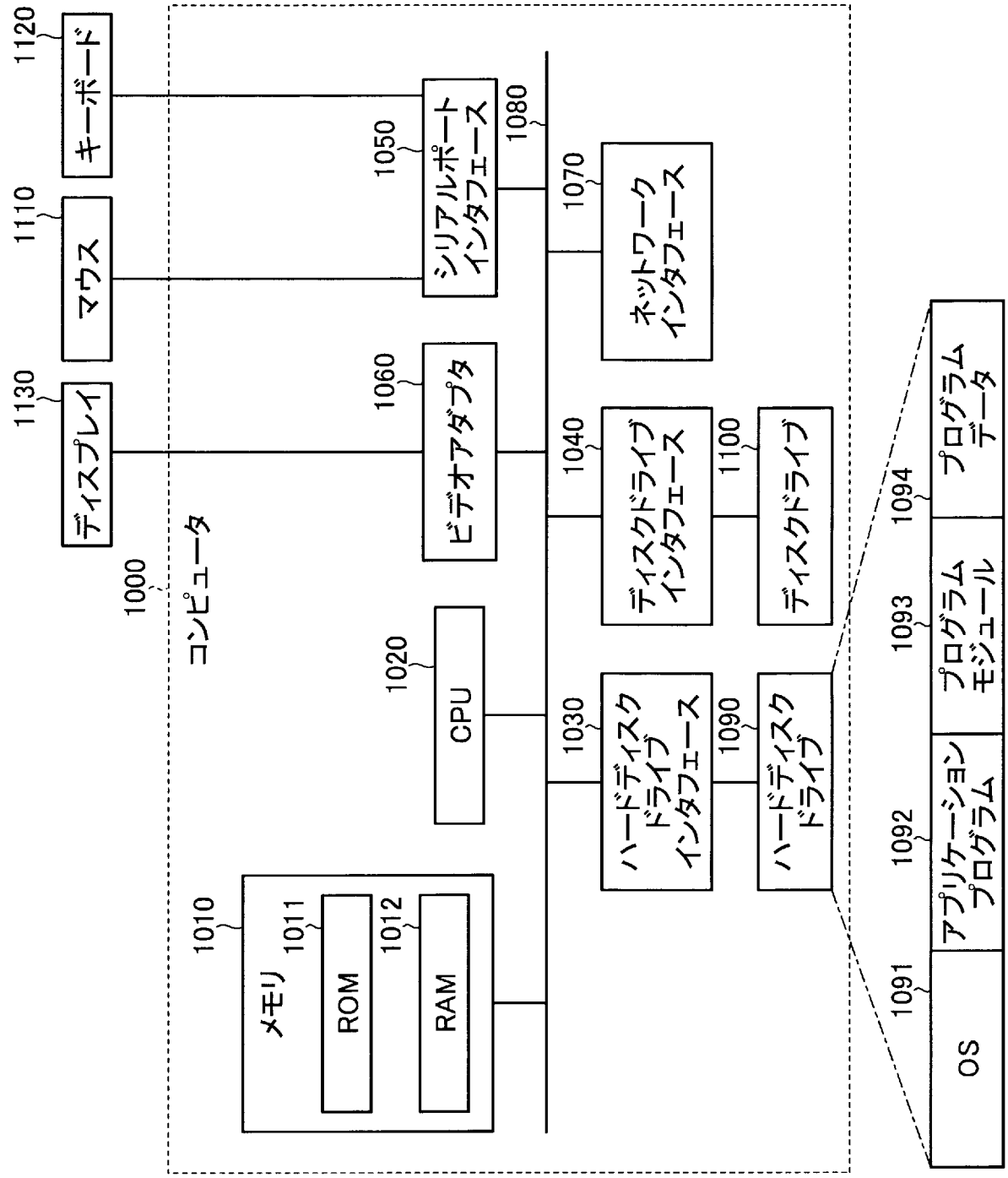
2018/6/19 12:53

	metrics	outlier
1	process-rabbitmq-server--rabbitmq-system-process-memory-share-diff	8190.9
2	load-system-load-1	392.1
3	load-system-load-norm-1	392.0
4	load-system-load-5	307.2
5	load-system-load-norm-5	307.2
6	process-snmpd--root-system-process-fd-open	293.9
7	process-httpd--apache-system-process-cpu-total-pct	292.1
8	load-system-load-15	237.5
9	load-system-load-norm-15	237.5
10	process-httpd--apache-system-process-memory-rss-pct	199.5
11	process-httpd--apache-system-process-memory-rss-bytes	196.8
12	process-httpd--apache-system-process-fd-open	176.1
13	process-httpd--apache-system-process-memory-size	150.5
14	cpu-system-cpu-system-pct	125.7
15	process-libvirtd--root-system-process-cpu-total-pct	78.4

[図24]


top1: process-rabbitmq-server--rabbitmq-system-process-memory-share-dif
top2: load-system-load-1
top4: load-system-load-5
top6: process-snmpd--root-system-process-fd-open
original
reconstruct
dock

[図25]


1120
キーボード
1110
マウス
1130
ディスプレイ
1000
コンピュータ
1050
シリアルポート
インタフェース
1080
1070
ネットワーク
インタフェース
1060
ビデオアダプタ
1040
ディスクドライブ
インタフェース
1100
ディスクドライブ
1020
CPU
1030
ハードディスク
ドライブ
インタフェース
1090
ハードディスク
ドライブ
1010
メモリ
1011
ROM
1012
RAM
1091
OS
1092
アプリケーション
プログラム
1093
プログラム
モジュール
1094
プログラム
データ

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2020/005474

A. CLASSIFICATION OF SUBJECT MATTER
G06F 11/07(2006.01)i; G06N 3/08(2006.01)i
FI: G06N3/08; G06F11/07 151

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F11/07; G06N3/08

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2020
Registered utility model specifications of Japan	1996-2020
Published registered utility model applications of Japan	1994-2020

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A	JP 2018-148350 A (NIPPON TELEGRAPH AND TELEPHONE CORP.) 20.09.2018 (2018-09-20) paragraphs [0031]-[0069]	1,4,8 2-3,6-7 5
Y	柏木 陽佑 他, 識別的推定法に基づく音声の構造的表象を制約として用いたニューラルネットワーク音響モデルの話者適応, 情報処理学会 研究報告音声言語情報処理(SLP), 21 July 2016, vol. 2016-SLP-112, no. 1, pp. 1-6, ISSN:2188-8663, chapter 3.3, (KASHIWAGI, Yosuke et al., "Speaker adaptation for deep neural network acoustic models based on discriminative estimation of structural constrainst", IPSJ SIG Technical report:Spoken Language Processing (SLP))	2
Y	JP 2008-117381 A (OMRON CORP.) 22.05.2008 (2008-05-22) paragraphs [0046]-[0063], fig. 1	3

☒ Further documents are listed in the continuation of Box C.
☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search 27 April 2020 (27.04.2020)	Date of mailing of the international search report 12 May 2020 (12.05.2020)
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2020/005474

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	納谷 太 他, センサネットワークを用いた業務の計測と分析, 電子情報通信学会技術研究報告, 09 July 2009, vol. 109, no. 131, pp. 127-134, ISSN:0913-5685, page 130, right column, lines 2-5, (NAYA, Futoshi et al., "Workflow Measurement and Analysis with Wireless Sensor Network Systems", IEICE technical report)	6,7
Y	JP 2018-112863 A (TOSHIBA CORP.) 19.07.2018 (2018-07-19) paragraphs [0010]-[0025], fig. 1	7

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2020/005474

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
JP 2018-148350 A	20 Sep. 2018	(Family: none)	
JP 2008-117381 A	22 May. 2008	(Family: none)	
JP 2018-112863 A	19 Jul. 2018	CN 110121724 A paragraphs [0026]-[0042], fig. 1	

国際調査報告

国際出願番号
PCT/JP2020/005474

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06F 11/07(2006.01)i; G06N 3/08(2006.01)i
FI: G06N3/08; G06F11/07 151

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G06F11/07; G06N3/08

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２０年
日本国実用新案登録公報	１９９６－２０２０年
日本国登録実用新案公報	１９９４－２０２０年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	JP 2018-148350 A (日本電信電話株式会社) 20.09.2018 (2018-09-20) 段落0031-0069	1, 4, 8
Y		2-3, 6-7
A		5
Y	柏木 陽佑 他, 識別的推定法に基づく音声の構造的表象を制約として用いたニューラルネットワーク音響モデルの話者適応, 情報処理学会 研究報告 音声言語情報処理（ＳＬＰ）, 2016.07.21, 第２０１６－ＳＬＰ－１１２巻 第１号, 第1頁-第6頁, ISSN:2188-8663 第3.3節	2
Y	JP 2008-117381 A (オムロン株式会社) 22.05.2008 (2008-05-22) 段落0046-0063, 図1	3
Y	納谷 太 他, センサネットワークを用いた業務の計測と分析, 電子情報通信学会技術研究報告, 2009.07.09, 第109巻 第131号, 第127頁-第134頁, ISSN:0913-5685 第130頁右欄第2行目-第5行目	6, 7
Y	JP 2018-112863 A (株式会社東芝) 19.07.2018 (2018-07-19) 段落0010-0025, 図1	7

☐ Ｃ欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

＊ 引用文献のカテゴリー
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
“O” 口頭による開示、使用、展示等に言及する文献
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
“Y” 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
“&” 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
27.04.2020	12.05.2020

名称及びあて先

日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）

三坂　敏夫 5B 4178

電話番号 03-3581-1101 内線 3545

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2020/005474

引用文献	公表日	パテントファミリー文献	公表日
JP 2018-148350 A	20.09.2018	（ファミリーなし）	
JP 2008-117381 A	22.05.2008	（ファミリーなし）	
JP 2018-112863 A	19.07.2018	CN 110121724 A 段落0026-0042，図1	