(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
2019年12月19日(19.12.2019)

WIPO | PCT

(10) 国際公開番号
WO 2019/240038 A1

(51) 国際特許分類:
G06F 21/55 (2013.01) *H04L 12/70* (2013.01)

(21) 国際出願番号： PCT/JP2019/022738

(22) 国際出願日： 2019年6月7日(07.06.2019)

(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 2018-113154 2018年6月13日(13.06.2018) JP

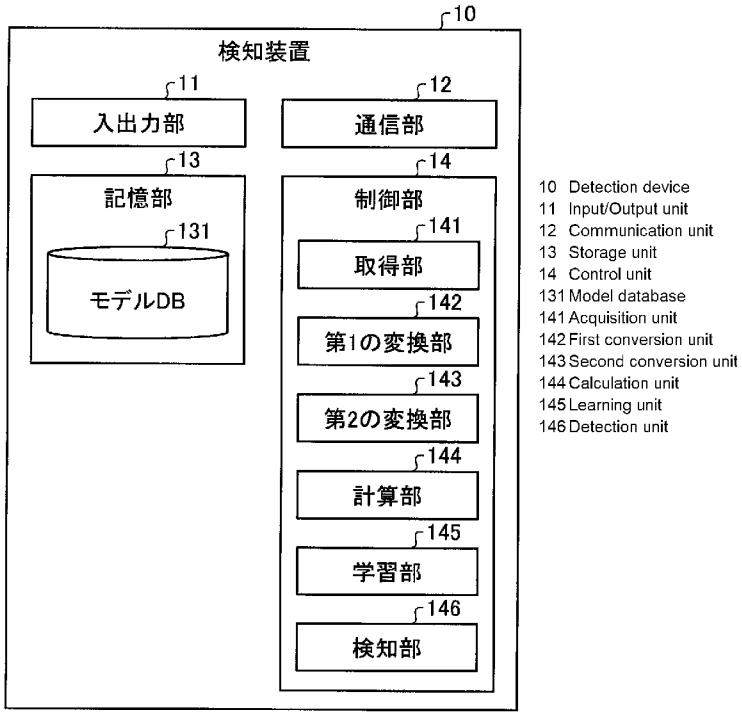
(71) 出願人: 日本電信電話株式会社 (**NIPPON TELEGRAPH AND TELEPHONE CORPORATION**) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目５番１号 Tokyo (JP).

(72) 発明者: 山田 真徳 (**YAMADA, Masanori**); 〒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). 五十嵐 弓将 (**IGARASHI, Yuminobu**); 〒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). 山中 友貴 (**YAMANAKA, Yuki**); 〒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).

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

(54) 発明の名称： 検知装置及び検知方法



(57) **Abstract:** This detection device (10) acquires a network log and a host log of an apparatus. The detection device (10) converts the network log into network feature values that are in a format enabling the values to be input into a multimodal generative model, which generates output data on the basis of a plurality of latent variables that are expressed as random variables. The detection device (10) converts the host log into host feature values that are in a format such that the values can be input into the generative model. The detection device (10) inputs at least one from among the network feature

WO 2019/240038 A1

[続葉有]

(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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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)）

values and the host feature values into the generative model, and calculates output data. The detection device (10) uses anomaly scores calculated on the basis of the output data to detect anomalies in the apparatus.

(57) 要約：検知装置（１０）は、機器のネットワークログ及びホストログを取得する。また、検知装置（１０）は、ネットワークログを、確率変数で表される複数の潜在変数を基に出力データを生成する生成モデルであって、マルチモーダルな生成モデルに入力可能な形式のネットワーク特徴量に変換する。また、検知装置（１０）は、ホストログを、生成モデルに入力可能な形式のホスト特徴量に変換する。また、検知装置（１０）は、ネットワーク特徴量及びホスト特徴量のうちの少なくとも一方を生成モデルに入力し、出力データを計算する。また、検知装置（１０）は、出力データを基に計算したアノマリスコアを用いて、機器の異常の検知を行う。

明　細　書

発明の名称 ： 検知装置及び検知方法

技術分野

[0001] 本発明は、検知装置及び検知方法に関する。

背景技術

[0002] 近年、ＩｏＴが普及し始め、これまであらゆる機器がネットワークを介して協調して動作し様々な価値を生み出そうとしている。一方で、様々な機器がネットワークを介してつながると思いもしない脆弱性を生む。また、現在はＩｏＴ普及の黎明期ということもあり、セキュリティ対策が十分でない機器が大量にネットワークにつながっている。

[0003] このような事情によりＩｏＴの異常検知技術は重要になってくる。異常検知器は、リスト型の検知器と学習型の検知器にわけることができる。リスト型の検知器は、各ＩｏＴ機器に合わせて検知条件を人が設計するタイプのものである。また、学習型の検知器は、データから検知条件を学んでいくタイプのものである。

[0004] ＩｏＴ機器は種類が多いので、学習型が主流になると考えられる。さらに学習型の検知手法には、正常状態を学び、正常状態からのズレをもって異常を検知するタイプのものと、異常状態を学び、異常状態への近さを利用して異常を検知するタイプのものとがある。

[0005] 例えば、正常状態からのズレをもって異常を検知するタイプの検知手法として、ＩｏＴ機器が接続された正常状態のネットワークのログを学習した検知モデルを用いた、アノマリ検知型の検知手法が知られている。

先行技術文献

非特許文献

[0006] 非特許文献1：Jinwon An, Sungzoon Cho, ”Variational Autoencoder based Anomaly Detection using Reconstruction Probability” [online]、[平成３０年６月４日検索]、インターネット（http://dm.snu.ac.kr/st

atic/docs/TR/SNUDM-TR-2015-03.pdf)

非特許文献2：Diederik P Kingma, Max Welling, "Auto-Encoding Variational Bayes" [online]、[平成３０年６月４日検索]、インターネット（https://arxiv.org/pdf/1312.6114.pdf)

非特許文献3：Masahiro Suzuki, Kotaro Nakayama, Yutaka Matsuo, "JOINT MULTIMODAL LEARNING WITH DEEP GENERATIVE MODELS" [online]、[平成３０年６月４日検索]、インターネット（https://arxiv.org/pdf/1611.01891.pdf)

非特許文献4：CERT NetSA Security Suite, "YAF" [online]、[平成３０年６月４日検索]、インターネット（https://tools.netsa.cert.org/yaf/index.html)

発明の概要

発明が解決しようとする課題

[0007] しかしながら、従来の技術には、ＩｏＴ機器の異常を高い精度で検知することが困難な場合があるという問題がある。例えば、ネットワークログを学習するアノマリ検知型の検知手法は、ＤｏＳ攻撃やArp spoofingによる異常を検知するのに有効である一方で、ランサムウェアによる異常を検知することができない場合がある。これは、ランサムウェアによる異常が、ネットワーク側の異常としては現れにくく、ホスト側の異常として現れやすいためである。逆に、ホストログログを学習する検知手法は、ランサムウェアによる異常の検知には有効であるが、ＤｏＳ攻撃やArp spoofingによる異常の検知には有効でない場合がある。

発明の効果

[0008] 本発明によれば、ＩｏＴ機器の異常を高い精度で検知することができる。

図面の簡単な説明

[0009] [図1]図１は、第１の実施形態に係る検知システムの構成の一例を示す図である。

[図2]図２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3]図３は、ＶＡＥについて説明するための図である。

[図4]図４は、第１の実施形態に係る生成モデルの一例を示す図である。

[図5]図５は、第１の実施形態に係る生成モデルの一例を示す図である。

[図6]図６は、第１の実施形態に係る粒度をそろえる方法について説明するための図である。

[図7]図７は、第１の実施形態に係る検知装置の処理の流れを示すフローチャートである。

[図8]図８は、第１の実施形態の効果を説明するための図である。

[図9]図９は、第１の実施形態の効果を説明するための図である。

[図10]図１０は、第１の実施形態の効果を説明するための図である。

[図11]図１１は、第１の実施形態の効果を説明するための図である。

[図12]図１２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

[0010] 以下に、本願に係る検知装置及び検知方法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。

[0011] ［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システム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システムの構成の一例を示す図である。図１に示すように、検知システム１は、検知装置１０、ゲートウェイ２０、機器３０を有し、ゲートウェイ２０は外部ネットワーク４０と接続されている。

[0012] 例えば、検知装置１０は、機器３０と外部ネットワーク４０との通信であって、ゲートウェイ２０を通過する通信のログを取得する。また、例えば、検知装置１０は、機器３０のホストログを取得する。また、検知装置１０は、取得したログを用いて学習した生成モデルを用いて、機器３０の異常の検

知を行う。

[0013] また、機器３０は、監視カメラやウェアラブルデバイスといったＩｏＴ機器である。例えば、機器３０が監視カメラである場合、検知装置１０は、監視カメラの解像度を変化させたときのネットワークログ及びホストログを取得する。

[0014] 次に、図２を用いて、検知装置１０の構成について説明する。図２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に示すように、検知装置１０は、入出力部１１、通信部１２、記憶部１３及び制御部１４を有する。

[0015] ここで、検知装置１０は、ＶＡＥ（Variational Autoencoder）を用いて検知及び学習を行う。図３を用いて、ＶＡＥについて説明する。図３は、ＶＡＥについて説明するための図である。

[0016] 図３に示すように、ＶＡＥの生成モデルは、オートエンコーダである。また、ＶＡＥは、ｉｎ層に入力された学習データを基にエンコーダq_ϕ（ｚ｜ｘ）により確率変数で表される潜在変数ｐ（ｚ）を生成し、ｐ（ｚ）から確率的に決定されたｚを基にデコーダp_θ（ｘ｜ｚ）により出力データを生成し、ｏｕｔ層に出力する。

[0017] ここで、エンコーダq_ϕ（ｚ｜ｘ）、潜在変数ｐ（ｚ）、及びデコーダp_θ（ｘ｜ｚ）は、いずれも分布を仮定するものである。また、エンコーダq_ϕ（ｚ｜ｘ）と潜在変数ｐ（ｚ）との間は確率的であるため、逆誤差伝搬が不可能である。一方、潜在変数ｐ（ｚ）とデコーダp_θ（ｘ｜ｚ）との間は決定的であるため、逆誤差伝搬が可能である。

[0018] また、ＶＡＥは、ｌｏｇｐ（ｘ）の再構成誤差項からＫＬ ｄｉｖｅｒｇｅｎｃｅによるｐ（ｚ）の束縛を表す正則化項を引いた変分下限の部分を目的関数として、当該目的関数が最大化されるように学習を行う。

[0019] また、本実施形態の検知装置１０は、生成モデルの学習を行う際には、マルチモーダルな学習を行う。マルチモーダルな学習とは、ネットワークログとホストログのような異なるドメインのデータを利用した学習のことである

。なお、１つのドメインのデータを使う学習は、シングルモーダルな学習と呼ばれる。

[0020] ここで、図４を用いて、マルチモーダルな生成モデルの学習について説明する。図４は、第１の実施形態に係る生成モデルの一例を示す図である。図４に示すように、検知装置１０は、生成モデルに、ネットワークログに基づく学習データ及びホストログに基づく学習データの両方を入力することができる。

[0021] また、検知装置１０は、ネットワークログに基づく特徴量を層２０１ａに入力し、抽象的な意味を持つデータを得る。一方、検知装置１０は、ホストログに基づく特徴量を層２０１ｂに入力し、抽象的な意味を持つデータを得る。

[0022] さらに、検知装置１０は、層２０１ａに特徴量が入力された場合は、層２０５ａを経由して出力データを出力させる。一方、検知装置１０は、層２０１ｂに特徴量が入力された場合は、層２０５ｂを経由して出力データを出力させる。

[0023] つまり、生成モデルは、異なるドメインのデータから得られた特徴量が各ドメインに対応する層に入力され、当該層から出力されるデータが中間層で合流し、さらに当該中間層から出力されるデータが出力層の手前の層で各ドメインに対応する層に分岐するようなニューラルネットワークを備えている。このような生成モデルにより、本実施形態では、異なるドメインの学習データを利用可能なマルチモーダルな学習を実現している。

[0024] なお、層２０１ａ、層２０１ｂ、層２０５ａ及び層２０５ｂは、それぞれ複数の層であってもよい。また、以降の説明では、層２０１ａ及び層２０１ｂを抽象化層と呼ぶ場合がある。また、層２０５ａ及び層２０５ｂを具体化層と呼ぶ場合がある。

[0025] また、ＶＡＥにおいては、入力される特徴量をｘとするとｘ→ｚ→ｘのようにして潜在変数ｚが計算される。本実施形態では、入力される特徴量ｘの元になったデータのドメインによらず、抽象化されたデータを用いて、潜在

変数ｚを計算することができる。

[0026] 一例として、ネットワークログからは、パケットの数やサイズに関する特徴量が得られる一方で、ＣＰＵ（Central Processing Unit）やメモリ等の使用量に関する特徴量は得られない場合がある。逆に、ホストログからは、ＣＰＵやメモリ等の使用量に関する特徴量が得られるが、パケットの数やサイズに関する特徴量は得られない場合がある。このように、異なるドメインのデータからは異なる特徴量が得られる。

[0027] 例えば、抽象化層によって、各特徴量が、「珍しさの度合い」、「分散の度合い」といった抽象的な意味を持つデータに変換されれば、異なるドメインのデータから得られた特徴量を同じ基準で評価することが可能になる。なお、抽象的な意味は、生成モデル内で解釈可能なものであればよく、上記の例のように言語で端的に表現できるようなものでなくてもよい。

[0028] ただし、各ドメイン間で、抽象化層及び具体化層の次元数が大きく異なっている場合、学習において、次元数が大きい方のドメインが重視されてしまうことがある。そのため、本実施形態の生成モデルでは、各ドメイン間の抽象化層及び具体化層の次元数がなるべく同じオーダーになるように設計されている。さらに、検知装置１０は、入力される特徴量の次元数を生成モデルに合わせて調整する。

[0029] また、図５に示すように、生成モデルには、一方のドメインのデータに基づく特徴量のみが入力されてもよい。図５は、第１の実施形態に係る生成モデルの一例を示す図である。図５の例では、ネットワークログに基づく特徴量が入力されているのに対し、ホストログに基づく特徴量は入力されていない。

[0030] 図２に戻り、入出力部１１は、ユーザからのデータの入力を受け付ける。入出力部１１は、例えば、マウスやキーボード等の入力装置、及びディスプレイやタッチパネル等の表示装置を含む。通信部１２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。例えば、通信部１２はＮＩＣ（Network Interface Card）である。通信部１２は、例えばゲートウェイ２０

との間でデータ通信を行う。

[0031] 記憶部１３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１３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１３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１３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１３は、モデルＤＢ１３１を有する。モデルＤＢ１３１は、学習済みの生成モデルのパラメータ等を記憶する。

[0032] 制御部１４は、検知装置１０全体を制御する。制御部１４は、例えば、ＣＰＵ、ＧＰＵ（Graphics Processing Unit）、ＴＰＵ（Tensor 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１４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１４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１４は、取得部１４１、第１の変換部１４２、第２の変換部１４３、計算部１４４、学習部１４５及び検知部１４６を有する。

[0033] 取得部１４１は、機器３０のネットワークログ及びホストログを取得する。例えば、取得部１４１は、ＹＡＦ（Yet Another Flowmeter）（例えば、非特許文献４を参照）を用いてipfix形式のネットワークログを取得することができる。また、取得部１４１は、機器３０のＯＳがＬｉｎｕｘ（登録商標）であれば、「/proc/diskstats」、「/proc/loadavg」、「/proc/meminfo」、「/proc/stat」といったディレクトリに存在するＣＰＵ、メモリ、ディスクＩ／Ｏに関する情報をホストログとして取得することができる。

[0034] 第１の変換部１４２は、ネットワークログを、確率変数で表される複数の

潜在変数を基に出力データを生成する生成モデルであって、マルチモーダルな生成モデルに入力可能な形式のネットワーク特徴量に変換する。なお、ネットワーク特徴量は、第１の特徴量の一例である。

[0035] 第１の変換部１４２は、ネットワークログに含まれる量的データを量的データの所定の統計量に変換することができる。例えば、ネットワークログに含まれる量的データは、通信のバイト数、パケット数、フロー数、パケットサイズ等である。また、統計量は、平均、最大、最小、変動係数、レート等である。

[0036] また、第１の変換部１４２は、ネットワークログに含まれる質的データをｋ－ｈｏｔ（ただし、ｋは１以上の整数）ベクトルに変換する。例えば、第１の変換部１４２は、ネットワークログに含まれるｉｐアドレス、ｍａｃアドレス、ｐｏｒｔのｓｒｃ及びｄｓｔを１－ｈｏｔベクトルに変換することができる。また、第１の変換部１４２は、ネットワークログに含まれるプロトコルをｋ－ｈｏｔベクトルに変換することができる。例えば、ｋを２とし、０番から３番までの４つのプロトコルのうち０番と３番のものがネットワークログに含まれている場合、第１の変換部１４２は、プロトコルを［１，０，０，１］と表すことができる。

[0037] 以下に、ネットワーク特徴量の例を示す。なお、ｕｐは、機器３０から外部ネットワーク４０へ向かう方向を示している。また、ｄｏｗｎは、外部ネットワーク４０から機器３０へ向かう方向を示している。

・ｕｐとｄｏｗｎそれぞれのバイト数
・ｕｐとｄｏｗｎそれぞれのパケット数
・ｕｐとｄｏｗｎそれぞれのフロー数
・ｕｐとｄｏｗｎそれぞれの平均パケットサイズ
・ｕｐとｄｏｗｎそれぞれの最大パケットサイズ
・ｕｐとｄｏｗｎそれぞれの最小パケットサイズ
・ｕｐとｄｏｗｎそれぞれの平均パケットサイズの平均
・ｕｐとｄｏｗｎそれぞれの変動係数（平均パケットサイズの標準偏差を平

均パケットサイズの平均で割ったもの）

・ｕｐとｄｏｗｎそれぞれの平均フローレート（フロー数を時間で割る）

・ｕｐとｄｏｗｎそれぞれの平均パケットレート（パケット数を時間で割る）

・ｉｐアドレス、ｍａｃアドレス、プロトコルのｋ－ｈｏｔベクトル

[0038] このように、第１の変換部１４２によれば、ネットワークログの各データの加工及び組み合わせにより複数の特徴量を得ることができる。このため、第１の変換部１４２は、ネットワーク特徴量を所定の次元数に調整することができる。

[0039] 第２の変換部１４３は、ホストログを、生成モデルに入力可能な形式のホスト特徴量に変換する。なお、ホスト特徴量は、第２の特徴量の一例である。例えば、第２の変換部１４３は、時間経過で累積するようなデータは、単位時間ごとの増加量に変換することができる。また、第２の変換部１４３は、対数等によるスケール調整を行うことができる。例えば、第２の変換部１４３は、Ｌｉｎｕｘで取得可能な以下の項目について、対数によるスケール調整を行う。

・SectorsRead

・TimeSpentReading

・SectorsWritten

・TimeSpentWriting

・TimeSpentDoing_I_Os

・WeightedTimeSpentDoing_I_Os

[0040] 第２の変換部１４３は、ホストログに含まれる時系列の累積データを単位時間ごとのデータに変換し、さらに、リソースの使用量に関するデータを全リソース量で割ることで正規化する。例えば、第２の変換部１４３は、メモリに関する値をTotal Mmoryで割り、１以下になるように変換する。また、第２の変換部１４３は、実行中のプロセス数を全プロセス数で割り、１以下になるように変換する。

[0041] また、第２の変換部１４３は、ＣＰＵ使用状況に関する以下の項目については、全項目の和を各項目の値で割り、１以下になるように変換する。

・Cpu_user

・Cpu_Nine

・Cpu_system

・Cpu_Idle

・Cpu_Iowait

・Cpu_Irq

・Cpu_Softirq

[0042] ここで、図６に示すように、ネットワークログとホストログは、出力間隔が異なり、１対１に対応しない場合がある。この場合、第２の変換部１４３は、ネットワーク特徴量とホスト特徴量とが１対１で対応するように、ネットワークログとホストログの粒度をそろえる処理を行う。図６は、第１の実施形態に係る粒度をそろえる方法について説明するための図である。

[0043] 第２の変換部１４３は、１つのネットワークログに複数のホストログが対応している場合、複数のホストログの各要素の最大、最小、平均及び分散のうちの少なくともいずれかを計算することで、複数のホストログを１つのホスト特徴量に変換する。ネットワークログは、出力するインタフェースによって出力間隔が異なるものとする。

[0044] 図６の例では、interface１のネットワークログには２つのホストログが含まれるので、第２の変換部１４３は、当該２つのホストログの各要素の最大、最小、平均及び分散のうちの少なくともいずれかを計算し、１つのホスト特徴量に変換する。

[0045] 一方で、interface３のネットワークログには、対応するホストログが存在しない（１つ未満）ため、第２の変換部１４３は、ホスト特徴量の変換を行わない。この場合、ネットワーク特徴量のみが生成モデルに入力される。

[0046] 計算部１４４は、ネットワーク特徴量及びホスト特徴量のうちの少なくとも一方を生成モデルに入力し、出力データを計算する。計算部１４４の処理

により、入力データに対応する出力データが得られる。ここで、検知装置１０は、入力データと出力データとの類似度合いに基づいて以降の処理を行う。

[0047] 学習部１４５は、出力データと生成モデルに入力した各特徴量との差分が小さくなるように生成モデルの学習を行う。具体的には、学習部１４５は、図３のｌｏｇｐ（ｘ）が最適化されるようにｐ（ｚ）のパラメータを更新する。

[0048] 検知部１４６は、出力データを基に計算したアノマリスコアを用いて、機器３０の異常の検知を行う。例えば、検知部１４６は、図３のｌｏｇｐ（ｘ）の値をアノマリスコアとし、アノマリスコアが閾値を超えている場合、機器３０に異常が発生していると判定することができる。

[0049] ［第１の実施形態の処理］

図７を用いて検知装置１０の処理について説明する。図７は、第１の実施形態に係る検知装置の処理の流れを示すフローチャートである。図７に示すように、まず、検知装置１０は、機器３０のネットワークログ及びホストログを取得する（ステップＳ１０１）。

[0050] 次に、検知装置１０は、ネットワークログをネットワーク特徴量に変換する（ステップＳ１０２）。また、検知装置１０は、ホストログをホスト特徴量に変換する（ステップＳ１０３）。そして、検知装置１０は、ネットワーク特徴量及びホスト特徴量をモデルに入力し、出力データを計算する（ステップＳ１０４）。

[0051] ここで、学習を行う場合（ステップＳ１０５、学習）、検知装置１０は、出力データを基にモデルを更新する（ステップＳ１０６）。一方、検知を行う場合（ステップＳ１０５、検知）、検知装置１０は、出力データから計算したアノマリスコアを用いて、異常を検知する（ステップＳ１０７）。

[0052] ［第１の実施形態の効果］

第１の実施形態において、検知装置１０は、機器３０のネットワークログ及びホストログを取得する。また、検知装置１０は、ネットワークログを、

確率変数で表される複数の潜在変数を基に出力データを生成する生成モデルであって、マルチモーダルな生成モデルに入力可能な形式のネットワーク特徴量に変換する。また、検知装置１０は、ホストログを、生成モデルに入力可能な形式のホスト特徴量に変換する。また、検知装置１０は、ネットワーク特徴量及びホスト特徴量のうちの少なくとも一方を生成モデルに入力し、出力データを計算する。また、検知装置１０は、出力データを基に計算したアノマリスコアを用いて、機器３０の異常の検知を行う。このように、検知装置１０は、ネットワークログ及びホストログの両方から変換した特徴量を使って異常の検知を行うため、ＩｏＴ機器の異常を高い精度で検知することができる。例えば、検知装置１０は、ＤｏＳ攻撃やArp spoofingによる異常、及びランサムウェアによる異常の両方を検知することができる。

[0053] また、検知装置１０は、出力データと生成モデルに入力した各特徴量との差分が小さくなるように生成モデルの学習を行うことができる。このように、検知装置１０は、検知に用いるモデルの学習をさらに行うことができる。

[0054] また、検知装置１０は、ネットワークログに含まれる量的データを量的データの所定の統計量に変換し、ネットワークログに含まれる質的データをｋ－ｈｏｔ（ただし、ｋは１以上の整数）ベクトルに変換することができる。これにより、検知装置１０は、特徴量の次元数を調整することができる。

[0055] また、検知装置１０は、ホストログに含まれる時系列の累積データを単位時間ごとのデータに変換し、さらに、リソースの使用量に関するデータを全リソース量で割ることで正規化する。これにより、検知装置１０は、データを正規化し、特徴量の次元数を調整することができる。

[0056] また、検知装置１０は、１つのネットワークログに複数のホストログが対応している場合、複数のホストログの各要素の最大、最小、平均及び分散のうちの少なくともいずれかを計算することで、複数のホストログを１つのホスト特徴量に変換することができる。これにより、検知装置１０は、ネットワーク特徴量とホスト特徴量の粒度をそろえることができる。

[0057] ここで、図８から図１１を用いて、第１の実施形態の検知装置１０を用い

て行った実験の結果を示し、実施形態の効果を説明する。図８から図１１は、第１の実施形態の効果を説明するための図である。

[0058] 実験では、機器３０は、小型コンピュータのrasberry piに動画撮影用のカメラを備えたものであるとする。また、実験では、機器３０を用いて動画をストリーミングした際のネットワークログ及びホストログを正常状態のデータとした。

[0059] まず、機器３０のストリーミング中の動画の画質を変更することで、ネットワークの異常を模擬した際の結果を図８及び図９に示す。図８及び図９に示すように、動画の画質を高画質から低画質に変更し、検知装置１０を用いてその際に取得したネットワークログ及びホストログからアノマリスコアを計算した結果、変更に応じてアノマリスコアが増加した。

[0060] 次に、機器３０によるストリーミング中に、ファイルの暗号化を実行させることで、ランサムウェアを模擬した際の結果を図１０及び図１１に示す。図１０及び図１１に示すように、暗号化を行っていない状態から、暗号化させるファイルのサイズを増加させていき、検知装置１０を用いてその際に取得したネットワークログ及びホストログからアノマリスコアを計算した結果、サイズの増加に応じてアノマリスコアが増加した。

[0061] このように、第１の実施形態の検知装置１０によって計算されるアノマリスコアは、機器３０に発生した異常に応じて増加する。このとき、適当な閾値を設定すれば、検知装置１０による異常の検知が可能となる。

[0062] ［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・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・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。

[0063] 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。

[0064] ［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。

[0065] 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、ネットワークログ及びホストログを入力とし、検知結果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。

[0066] 図１２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。

[0067] 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。

[0068] 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。

[0069] 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が、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に読み出して実行する。

[0070] 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

[0071] １０ 検知装置
１１ 入出力部
１２ 通信部
１３ 記憶部
１４ 制御部
２０ ゲートウェイ
３０ 機器
４０ 外部ネットワーク
１４１ 取得部
１４２ 第１の変換部
１４３ 第２の変換部
１４４ 計算部
１４５ 学習部
１４６ 検知部

請求の範囲

[請求項1] 機器のネットワークログ及びホストログを取得する取得部と、

前記ネットワークログを、確率変数で表される複数の潜在変数を基に出力データを生成する生成モデルであって、マルチモーダルな生成モデルに入力可能な形式の第１の特徴量に変換する第１の変換部と、

前記ホストログを、前記生成モデルに入力可能な形式の第２の特徴量に変換する第２の変換部と、

前記第１の特徴量及び前記第２の特徴量のうちの少なくとも一方を前記生成モデルに入力し、前記出力データを計算する計算部と、

前記出力データを基に計算したアノマリスコアを用いて、前記機器の異常の検知を行う検知部と、

を有することを特徴とする検知装置。

[請求項2] 前記出力データと前記生成モデルに入力した各特徴量との差分が小さくなるように前記生成モデルの学習を行う学習部をさらに有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] 前記第１の変換部は、前記ネットワークログに含まれる量的データを前記量的データの所定の統計量に変換し、前記ネットワークログに含まれる質的データをｋ－ｈｏｔ（ただし、ｋは１以上の整数）ベクトル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] 前記第２の変換部は、前記ホストログに含まれる時系列の累積データを単位時間ごとのデータに変換し、さらに、リソースの使用量に関するデータを全リソース量で割ることで正規化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5] 前記第２の変換部は、

１つのネットワークログに複数のホストログが対応している場合、前記複数のホストログの各要素の最大、最小、平均及び分散のうちの少なくともいずれかを計算することで、前記複数のホストログを１つの前記第２の特徴量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6] コンピュータによって実行される検知方法であって、

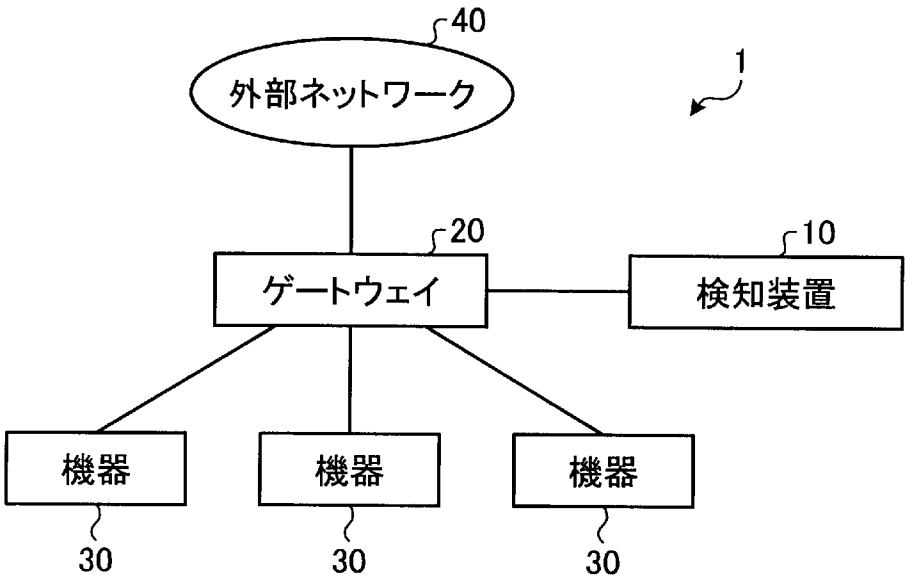
機器からネットワークログ及びホストログを取得する取得工程と、

前記ネットワークログを、確率変数で表される複数の潜在変数を基に出力データを生成するマルチモーダルな生成モデルに入力可能な形式の第１の特徴量に変換する第１の変換工程と、

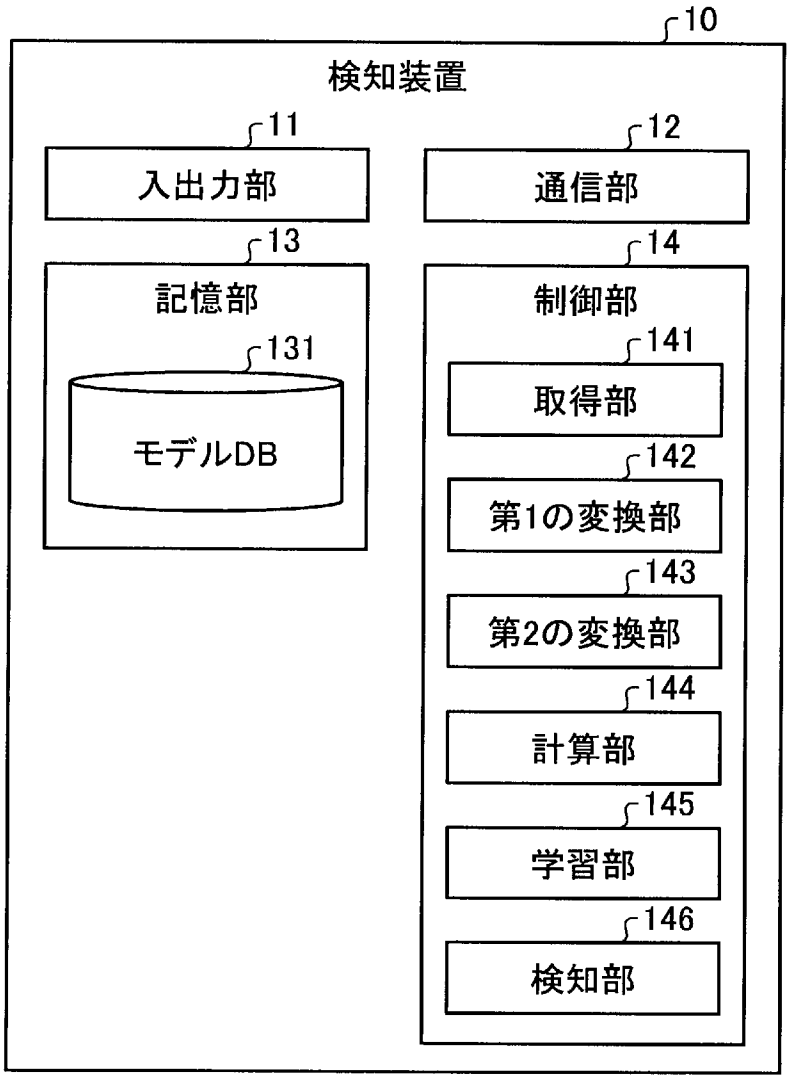
前記ホストログを、前記生成モデルに入力可能な形式の第２の特徴量に変換する第２の変換工程と、

を含むことを特徴とする検知方法。

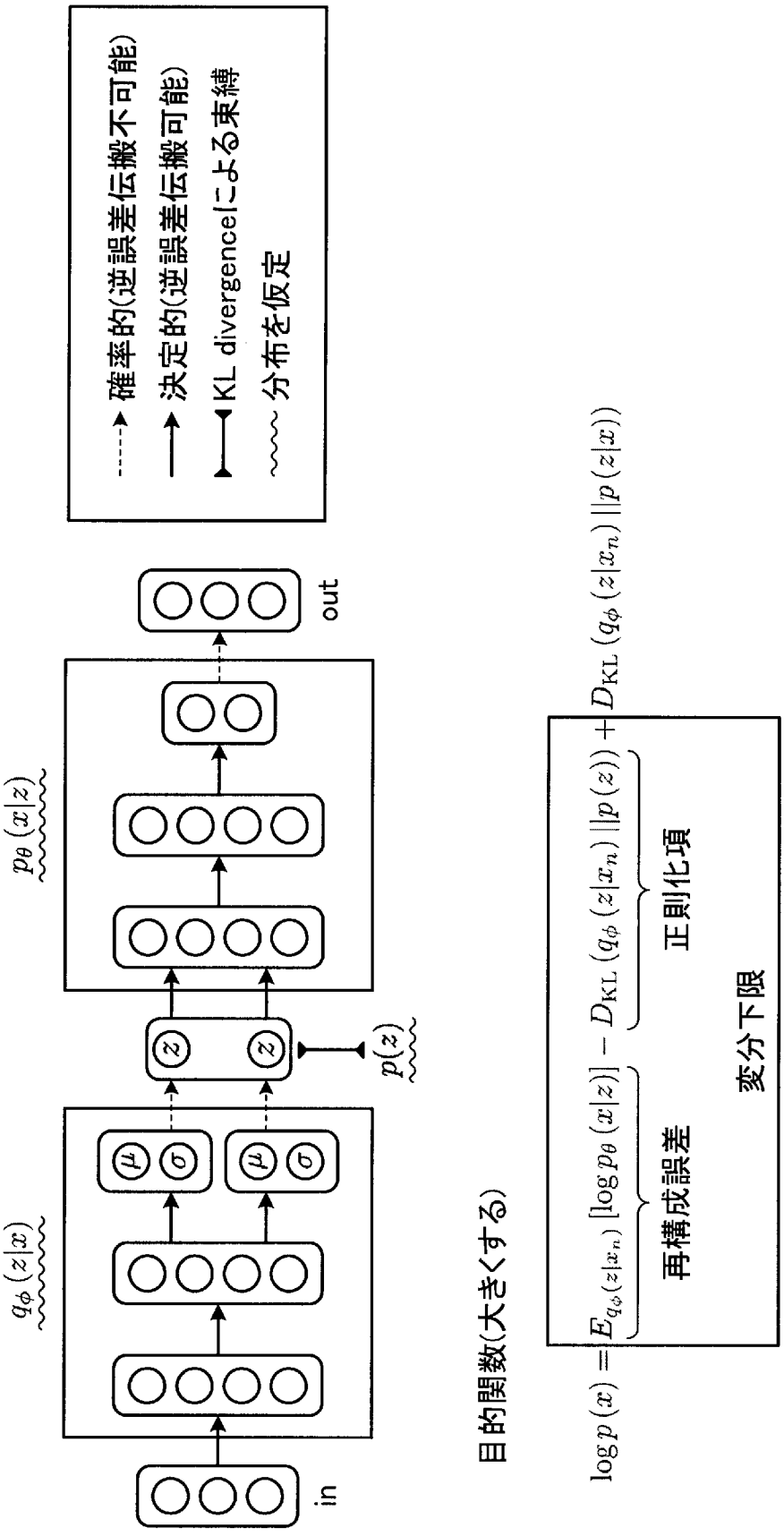
[図1]


40
外部ネットワーク
1
20
ゲートウェイ
10
検知装置
機器
機器
機器
30
30
30

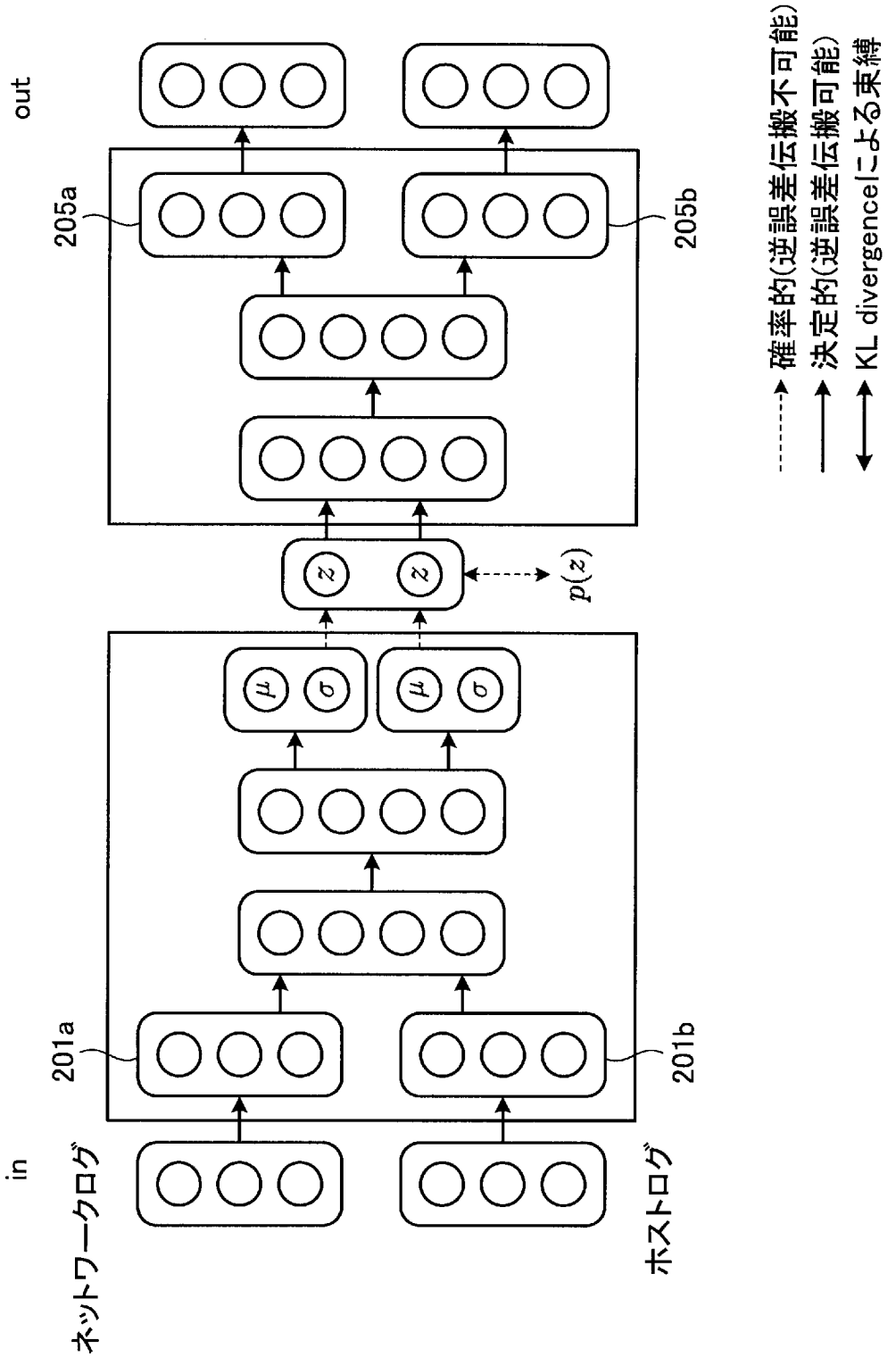
[図2]


10
検知装置
11
入出力部
12
通信部
13
記憶部
131
モデルDB
14
制御部
141
取得部
142
第1の変換部
143
第2の変換部
144
計算部
145
学習部
146
検知部

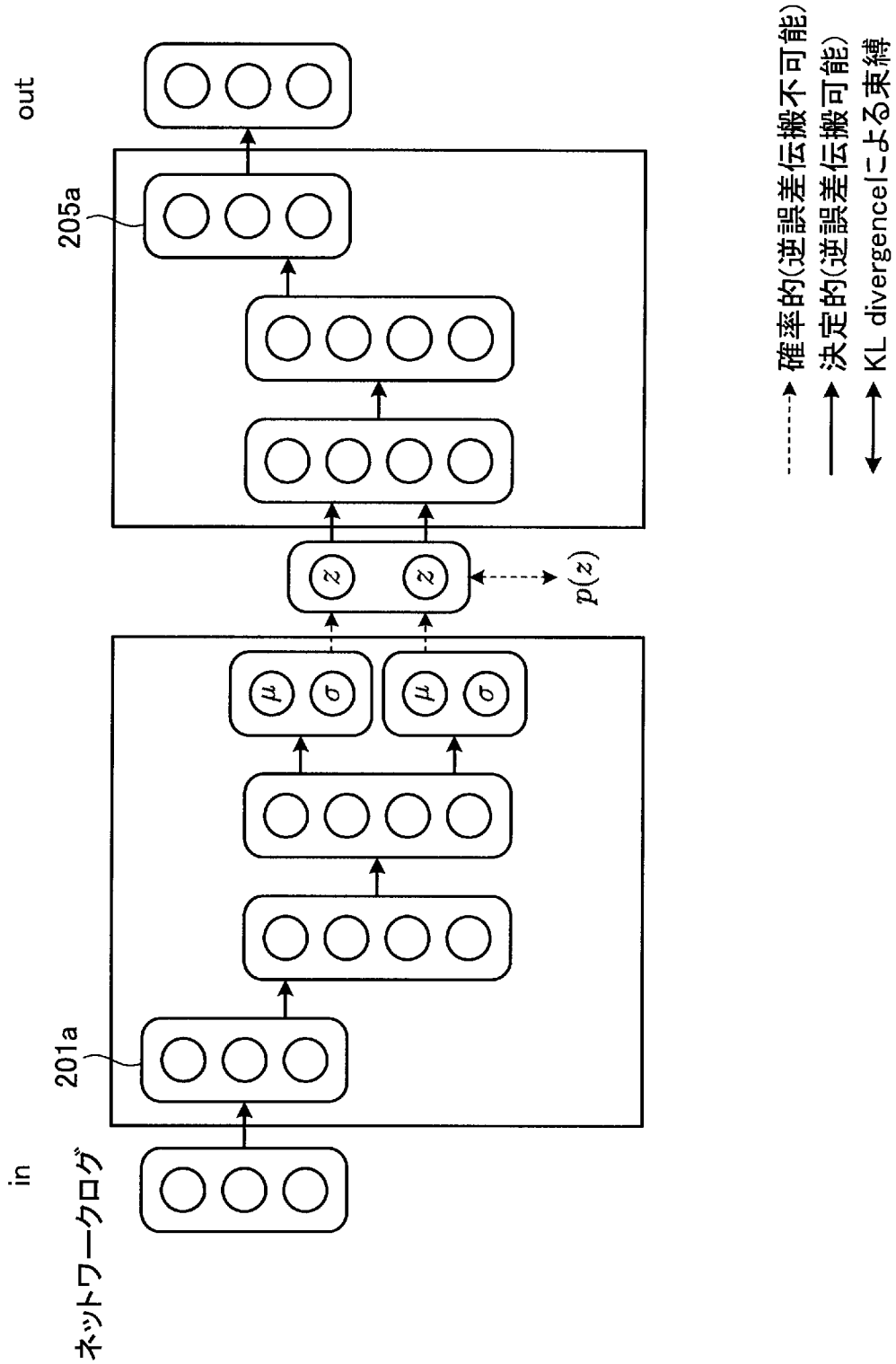
[図3]


確率的(逆誤差伝搬不可能)
決定的(逆誤差伝搬可能)
KL divergenceによる束縛
分布を仮定
in
out
$q_\phi(z|x)$
$p_\theta(x|z)$
$p(z)$
目的関数(大きくする)
$\log p(x) = \underbrace{E_{q_\phi(z|x_n)}[\log p_\theta(x|z)]}_{\text{再構成誤差}} - \underbrace{D_{\mathrm{KL}}(q_\phi(z|x_n)||p(z))}_{\text{正則化項}} + D_{\mathrm{KL}}(q_\phi(z|x_n)||p(z|x))$
変分下限

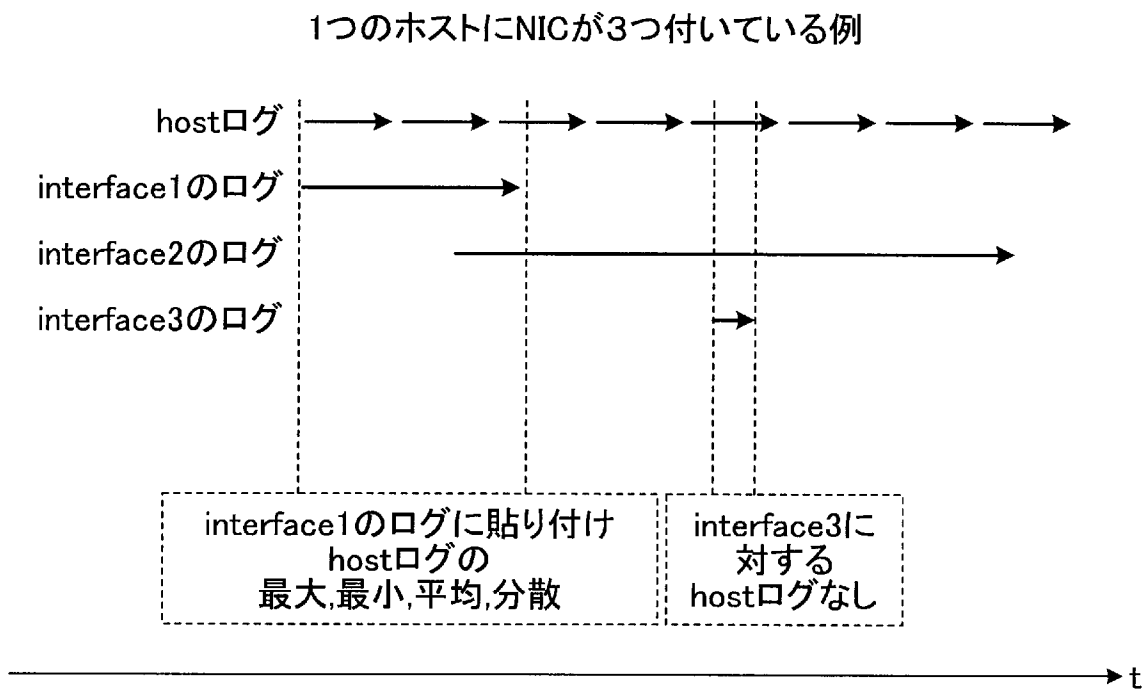
[図4]


out
205a
205b
in
201a
201b
ネットワークログ
ホストログ
μ
σ
z
$p(z)$
確率的(逆誤差伝搬不可能)
決定的(逆誤差伝搬可能)
KL divergenceによる束縛

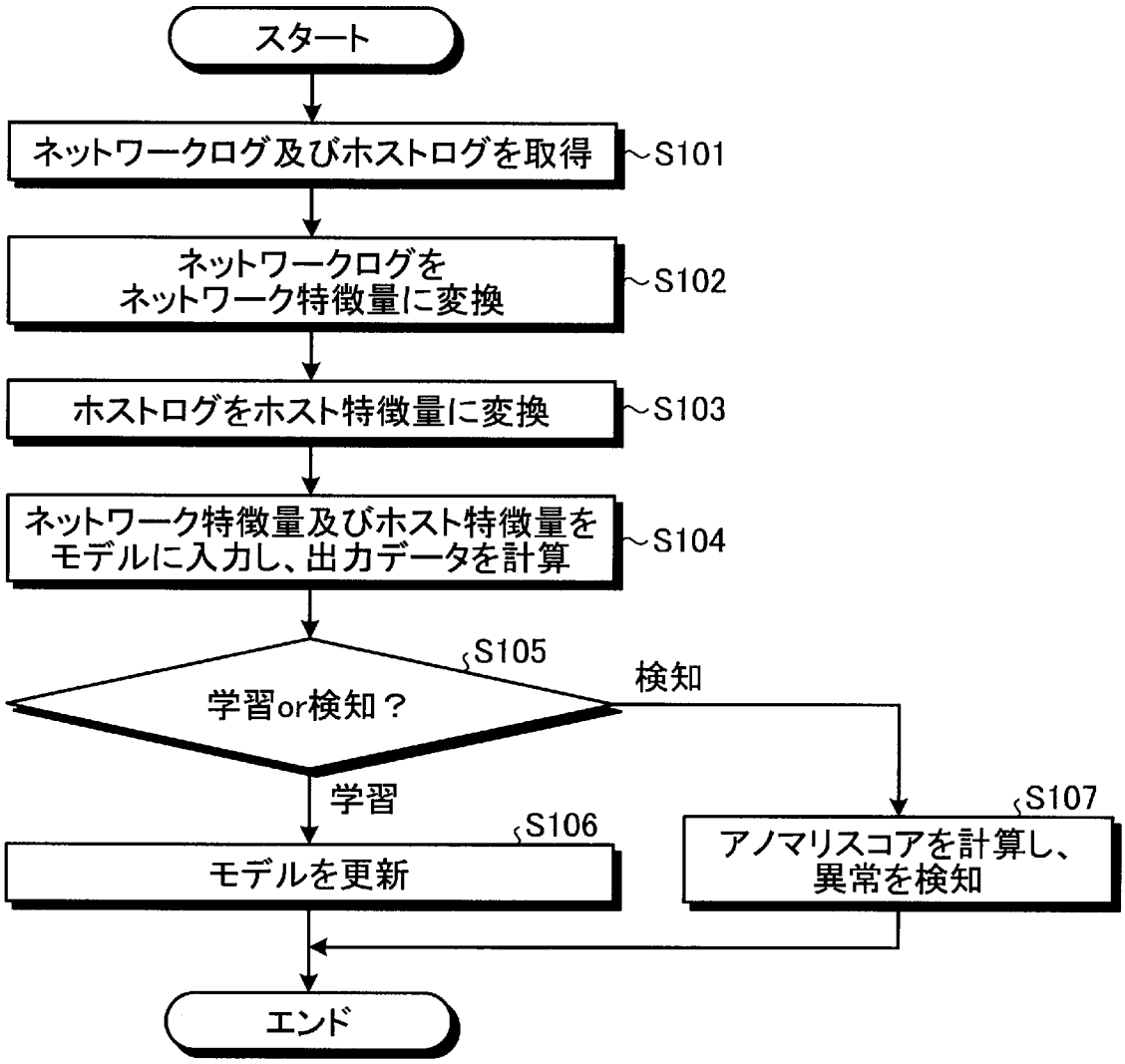
[図5]


out
205a
$p(z)$
201a
in
ネットワークログ
確率的(逆誤差伝搬不可能)
決定的(逆誤差伝搬可能)
KL divergenceによる束縛

[図6]


1つのホストにNICが3つ付いている例
hostログ
interface1のログ
interface2のログ
interface3のログ
interface1のログに貼り付け
hostログの
最大,最小,平均,分散
interface3に
対する
hostログなし
t

[図7]


スタート
ネットワークログ及びホストログを取得
S101
ネットワークログを
ネットワーク特徴量に変換
S102
ホストログをホスト特徴量に変換
S103
ネットワーク特徴量及びホスト特徴量を
モデルに入力し、出力データを計算
S104
S105
学習or検知？
検知
学習
S106
モデルを更新
S107
アノマリスコアを計算し、
異常を検知
エンド

[図8]

データ取得の状況(ネットワークの異常)

データ系列	カメラの画質	状態
正常1	高画質(1280x720 720p)	カメラ配信のみ
異常1	低画質(640x480 VGA)	カメラ配信のみ
異常2	低画質(640x480 VGA)	カメラ配信のみ

[図9]

アノマリスコア(ネットワークの異常)

データ系列	アノマリスコア
正常1	497
異常1	602
異常2	623

[図10]

データ取得の状況(ホストの異常)

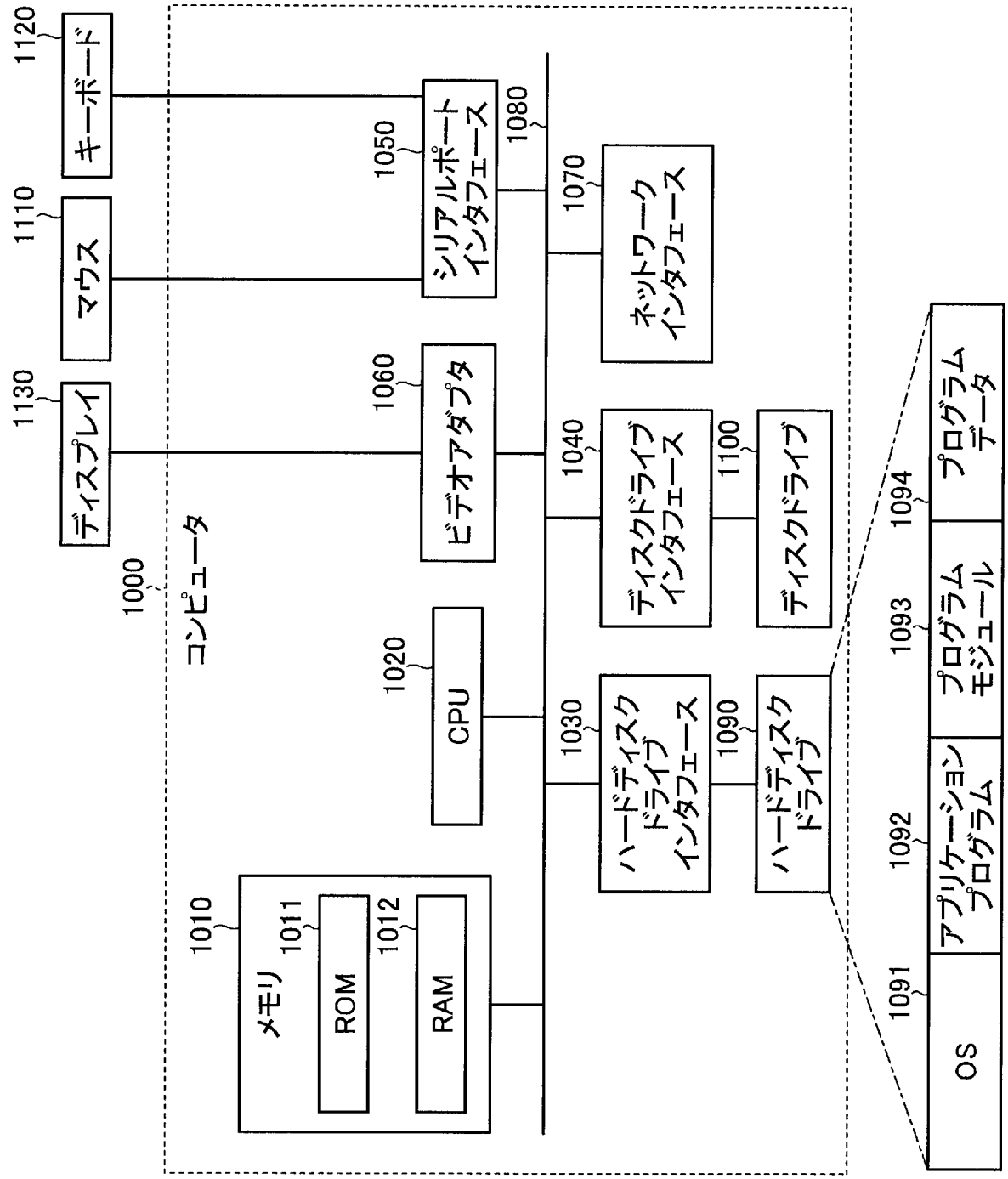
データ系列	カメラの画質	状態
正常1	高画質(1280x720 720p)	カメラ配信のみ
異常1	高画質(1280x720 720p)	カメラ配信, 1MBのファイル暗号化
異常2	高画質(1280x720 720p)	カメラ配信, 10MBのファイル暗号化
異常3	高画質(1280x720 720p)	カメラ配信, 100MBのファイル暗号化

[図11]

アノマリスコア(ホストの異常)

データ系列	アノマリスコア
正常1	611
異常1	746
異常2	776
異常3	927

[図12]


1000
コンピュータ
1010
メモリ
1011
ROM
1012
RAM
1020
CPU
1030
ハードディスクドライブインタフェース
1040
ディスクドライブインタフェース
1050
シリアルポートインタフェース
1060
ビデオアダプタ
1070
ネットワークインタフェース
1080
1090
ハードディスクドライブ
1100
ディスクドライブ
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
1110
マウス
1120
キーボード
1130
ディスプレイ

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2019/022738

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl. G06F21/55(2013.01)i, H04L12/70(2013.01)i

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)
Int.Cl. G06F21/55, H04L12/70

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	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

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 2018-73258 A (NIPPON TELEGRAPH AND TELEPHONE CORP.) 10 May 2018, abstract, paragraphs [0011], [0025]-[0034] (Family: none)	6 1-2,4 3,5
Y A	JP 2018-25936 A (OKUMA CORPORATION) 12 February 2018, paragraph [0002] (Family: none)	1-2,4, 3,5
Y A	菊地 啓太, ほか2名, IoT家電の故障検知を目的とした外れ値検知モデルの構築, 電気学会研究会資料, 22 March 2018, pp. 61-65, in particular, p. 63, right column, line 22 to p. 64, left column, line 10, (KIKUCHI, Keita et al., "Outlier Detection Model on the Failure Detection for IoT Home Appliance", Documents of research group of the Institute of Electrical Engineering of Japan)	4 3,5

☐ 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 14 August 2019 (14.08.2019)	Date of mailing of the international search report 27 August 2019 (27.08.2019)
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１９／０２２７３８

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F21/55(2013.01)i, H04L12/70(2013.01)i

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））

Int.Cl.　G06F21/55, H04L12/70

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	１９７１－２０１９年
日本国実用新案登録公報	１９９６－２０１９年
日本国登録実用新案公報	１９９４－２０１９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 Y A	JP 2018-73258 A（日本電信電話株式会社）2018.05.10, 要約、段落0011,0025-0034（ファミリーなし）	6 1-2,4 3,5
Y A	JP 2018-25936 A（オークマ株式会社）2018.02.15, 段落0002（ファミリーなし）	1-2,4 3,5

☑　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矛盾するものではなく、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
１４．０８．２０１９	２７．０８．２０１９

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｓ　１２０６
吉田　歩
電話番号　０３－３５８１－１１０１　内線　３５４６

Ｃ（続き）．　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
Y	菊地　啓太，ほか２名，ＩｏＴ家電の故障検知を目的とした外れ値検知モデルの構築，電気学会研究会資料，2018.03.22，p.61-65，特にp.63 右欄第22行－p.64 左欄第10行	4
A		3,5