(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6126891号
(P6126891)

(45)発行日　平成29年5月10日(2017.5.10)　(24)登録日　平成29年4月14日(2017.4.14)

(51)Int.Cl.　F I
G06F 11/07 (2006.01)　G06F 11/07 190
G06F 11/34 (2006.01)　G06F 11/07 140A
G06F 11/34 166

請求項の数 5　(全 70 頁)

(21)出願番号　特願2013-74784(P2013-74784)
(22)出願日　平成25年3月29日(2013.3.29)
(65)公開番号　特開2014-199579(P2014-199579A)
(43)公開日　平成26年10月23日(2014.10.23)
審査請求日　平成27年12月4日(2015.12.4)

(73)特許権者　000005223
富士通株式会社
神奈川県川崎市中原区上小田中４丁目１番１号
(74)代理人　100074099
弁理士　大菅　義之
(74)代理人　100133570
弁理士　▲徳▼永　民雄
(74)復代理人　100157967
弁理士　菅田　洋明
(72)発明者　大塚　浩
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

最終頁に続く

(54)【発明の名称】検出方法、検出プログラム、および検出装置

(57)【特許請求の範囲】
【請求項１】
コンピュータシステムに含まれるハードウェア、及びソフトウェアの少なくとも一方を含む複数の構成アイテムから所定期間に出力された複数のメッセージに基づいて、ある種別の障害の発生が予測された場合、前記複数の構成アイテムの各々の構成アイテムについて、該構成アイテムが出力した出力メッセージと同じ種別のメッセージが、前記ある種別の過去に発生した障害の発生時点より前に出力された頻度を表す第１の頻度と、前記発生時点より前にいずれかのメッセージが出力された出力時点から所定時間だけ遡るウィンドウ期間中に前記出力メッセージと同じ種別のメッセージが出力され、かつ、前記ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせである第２のパターンに基づいて前記ある種別の障害の発生が予測された頻度を表す第２の頻度とを取得し、
前記複数の構成アイテムの各々の構成アイテムについて取得した前記第１の頻度と前記第２の頻度とから算出した統計値に基づいて、前記複数の構成アイテムの中で相対的に高い蓋然性で前記ある種別の障害が発生すると予測される１つ以上の構成アイテムを示す結果情報を生成する、
ことをコンピュータが実行する検出方法。
【請求項２】
統計値は、前記第１の頻度に対して単調減少するとともに前記第２の頻度に対して単調増加することを特徴とする請求項１に記載の検出方法。
【請求項３】

前記結果情報を生成する処理が、
前記複数のメッセージの各々のメッセージについて、該メッセージと同じ種別のメッセージを、前記ある種別の障害の発生が予測された前記ウィンドウ期間に出力した第１の構成アイテムと、前記ある種別の過去に発生した障害が実際に起きた第２の構成アイテムとの間の関係である第１の関係が、前記複数のメッセージのうちの前記メッセージと同じ種別のメッセージを出力した第３の構成アイテムとの間に成り立つ関連構成アイテムを、前記複数の構成アイテムの中から検索し、
前記第３の構成アイテムについて前記関連構成アイテムが見つかった場合は、前記ある種別の障害が前記関連構成アイテムにおいて将来発生する蓋然性に関する評価値を、前記第３の構成アイテムについて算出した前記統計値に基づいて決定し、
前記関連構成アイテムについて決定した前記評価値に基づいて、前記結果情報を生成する
ことを含むことを特徴とする請求項１または２に記載の検出方法。
【請求項４】
コンピュータシステムに含まれるハードウェア、及びソフトウェアの少なくとも一方を含む複数の構成アイテムから所定期間に出力された複数のメッセージに基づいて、ある種別の障害の発生が予測された場合、前記複数の構成アイテムの各々の構成アイテムについて、該構成アイテムが出力した出力メッセージと同じ種別のメッセージが、前記ある種別の過去に発生した障害の発生時点より前に出力された頻度を表す第１の頻度と、前記発生時点より前にいずれかのメッセージが出力された出力時点から所定時間だけ遡るウィンドウ期間中に前記出力メッセージと同じ種別のメッセージが出力され、かつ、前記ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせである第２のパターンに基づいて前記ある種別の障害の発生が予測された頻度を表す第２の頻度とを取得し、
前記複数の構成アイテムの各々の構成アイテムについて取得した前記第１の頻度と前記第２の頻度とから算出した統計値に基づいて、前記複数の構成アイテムの中で相対的に高い蓋然性で前記ある種別の障害が発生すると予測される１つ以上の構成アイテムを示す結果情報を生成する、
ことを含む処理をコンピュータに実行させる検出プログラム。
【請求項５】
コンピュータシステムに含まれるハードウェア、及びソフトウェアの少なくとも一方を含む複数の構成アイテムから所定期間に出力された複数のメッセージに基づいて、ある種別の障害の発生が予測された場合、前記複数の構成アイテムの各々の構成アイテムについて、該構成アイテムが出力した出力メッセージと同じ種別のメッセージが、前記ある種別の過去に発生した障害の発生時点より前に出力された頻度を表す第１の頻度と、前記発生時点より前にいずれかのメッセージが出力された出力時点から所定時間だけ遡るウィンドウ期間中に前記出力メッセージと同じ種別のメッセージが出力され、かつ、前記ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせである第２のパターンに基づいて前記ある種別の障害の発生が予測された頻度を表す第２の頻度とを取得する取得手段と、
前記複数の構成アイテムの各々の構成アイテムについて取得した前記第１の頻度と前記第２の頻度とから算出した統計値に基づいて、前記複数の構成アイテムの中で相対的に高い蓋然性で前記ある種別の障害が発生すると予測される１つ以上の構成アイテムを示す結果情報を生成する生成手段と、
を備える検出装置。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、コンピュータシステムに発生する障害（failure）を管理する技術に関する。
【背景技術】
【０００２】

コンピュータシステムに発生する障害に関しては、例えば以下のような様々な観点から、様々な研究が行われている。

【０００３】

・実際に障害が発生した場合に、いかにして障害箇所（point of failure）または障害原因を特定するかという観点。

・障害の発生をいかにして予測するかという観点。

・障害に対処するシステム管理者（system administrator）等の人間の負担をいかにして軽減するかという観点。

【０００４】

例えば、あるネットワークシステム性能診断方法によれば、ネットワークシステム設計情報とネットワーク機器の稼動統計情報とがリンクされる。また、ＩＰ（Internet Protocol）層やＡＴＭ（Asynchronous Transfer Mode）層といった異なるプロトコル層の設計情報と稼動統計情報とがリンクされ、統合管理される。そして、サーバからクライアントへの経路に沿った稼動統計情報を一覧表示することにより、障害予兆発生範囲や原因個所が特定される。

【０００５】

また、情報システムに発生したトラブルの原因を突き止めて解決するための、ある種のトラブルシュート支援技術では、性能情報データベースが参照されることがある。さらに、先行する挙動の系列が後続の挙動に影響を与え得る動作対象に対して、異常動作の検出および原因の特定を可能にすることを目的とした異常挙動検出装置も提案されている。

【０００６】

また、ある運用管理装置は、障害の予兆を検出し、障害の発生場所を特定することを可能とすることを目的とし、相関モデル生成部と相関変化分析部を含む。相関モデル生成部は、少なくとも第１の要素に関する性能情報の時系列変化を示す第１の性能系列情報と、第２の要素に関する性能情報の時系列変化を示す第２の性能系列情報との相関関数を導出する。各要素は、性能種目または被管理装置である。相関モデル生成部は、相関関数に基づいて相関モデルを生成する。具体的には、相関モデル生成部は、相関モデルを各要素間の組み合わせについて求める。相関変化分析部は、被管理装置から新たに検出され取得される性能情報に基づいて、相関モデルの変化を分析する。

【０００７】

また、ある故障解析方法によれば、重障害の故障箇所と、当該重障害の予兆となる軽障害の故障箇所とが、１つの故障グループとして関連づけられて、故障関連づけテーブルに格納される。そして、障害発生時には、障害情報から障害種別が判別され、障害情報が障害種別とともに障害ログデータとして格納される。また、障害発生時には、故障関連づけテーブルが参照されて、対応する故障グループ番号が特定され、特定された故障グループ番号が障害ログデータに関連づけられて格納される。重障害発生時には、当該重障害と同じ故障グループに属する軽障害の障害ログデータが参照され、故障検出箇所が特定される。

【０００８】

さらに、機器の構成や設定が変更された場合にも、メッセージパターンに基づく障害検知を適切に行うことを目的とする管理装置も提案されている。当該管理装置は、判別手段と更新手段を備える。

【０００９】

ここで、情報処理システムに障害が発生したときに、情報処理システムから一定期間に受信されたメッセージを含むメッセージ群を示す第１のメッセージパターンが検出された検出回数が、障害共起情報に記憶されているものとする。判別手段は、障害共起情報から検出回数を読み出し、検出回数に基づいて障害と第１のメッセージパターンとの共起確率を算出する。そして、共起確率が閾値以上の場合に、判別手段は、障害が発生したと判別する。

【００１０】

また、更新手段は、構成要素が変更されると、変更された構成要素が出力するメッセージを第１のメッセージパターンから除いたメッセージ群を示す第２のメッセージパターンを作成する。そして、更新手段は、障害共起情報に記憶された第１のメッセージパターンを第２のメッセージパターンに更新する。

【００１１】

そのほかにも、コンピュータシステムの障害検知のための作業負担を軽減させることを目的としたプログラムが提案されている。ここで、構成情報記憶部には、情報処理システムの構成要素の識別情報に対応づけて、当該構成要素の種別情報が記憶されているものとする。上記プログラムがコンピュータに実行させる処理は、情報処理システムより出力され識別情報を含むメッセージに対応する種別情報を、構成情報記憶部を用いて判定することを含む。また、上記プログラムがコンピュータに実行させる処理は、複数のメッセージが含まれる第１のメッセージ群と第２のメッセージ群とを照合することを含む。ここで、第２のメッセージ群は、具体的には、メッセージ群記憶部に記憶されており、第２のメッセージ群に含まれる各メッセージには、他の情報処理システムの構成要素の種別情報が関連づけられているものとする。上記プログラムがコンピュータに実行させる処理は、さらに、上記の照合で一致しないメッセージ同士については、それぞれに係る種別情報に関して照合を行うことを含む。

【先行技術文献】

【特許文献】

【００１２】

【特許文献１】特開２００２－９９４６９号公報
【特許文献２】国際公開ＷＯ２０１０／０１０６２１号公報
【特許文献３】特開２００５－１４１４５９号公報
【特許文献４】特開２００９－１９９５３３号公報
【特許文献５】特開２００９－２３０５３３号公報
【特許文献６】特開２０１２－１２３６９４号公報
【特許文献７】特開２０１２－１４１８０２号公報

【発明の概要】

【発明が解決しようとする課題】

【００１３】

コンピュータシステムにおける障害の発生を未然に防ぐことは、コンピュータシステムの可用性を高めるうえで有益である。しかしながら、障害の発生を未然に防ぐための技術は、まだ発展途上であり、改善の余地がある。

【００１４】

例えば、単に「コンピュータシステムに障害が発生しそうかどうか」を予測するだけでは、「障害の発生を防止する」という目的が十分に達せられないことがあり得る。具体的には、「コンピュータシステム内のどの構成アイテム（configuration item）に対して対策をとれば障害の発生を防止するうえで有益なのか」ということが不明だと、「障害の発生を防止する」という目的が十分に達せられないことがあり得る。

【００１５】

そこで、本発明は、１つの側面では、障害の発生を防止するうえで有益な情報を検出することを目的とする。

【課題を解決するための手段】

【００１６】

一態様によれば、コンピュータシステムを管理するコンピュータが実行する検出方法が提供される。

前記検出方法は、第１のパターンに基づき、ある種別の障害の発生が予測される場合に、前記コンピュータが、Ｑ個の構成アイテム（１≦Ｑ）の各々について、第１の頻度と第２の頻度とに基づいて、統計値を算出することを含む。ここで、前記第１のパターンは、複数の構成アイテムのうちの前記Ｑ個から所定時間以下の長さの期間に出力されるＰ個（

１≦Ｑ≦Ｐ）のメッセージの組み合わせである。そして、前記複数の構成アイテムの各々は、前記コンピュータシステムに含まれるハードウェア、ソフトウェア、または両者の組み合わせである。また、前記統計値は、前記ある種別の障害が当該構成アイテムで将来発生する蓋然性に関するものである。

【００１７】

ここで、前記第１の頻度は、前記ある種別の障害が過去に発生した発生時点より前に、前記Ｐ個のメッセージのうち当該構成アイテムが出力した出力メッセージと同じ種別のメッセージが出力された頻度である。また、前記発生時点より前にいずれかのメッセージが出力された出力時点から前記所定時間だけ遡る期間を、ウィンドウ期間ということにする。そして、前記ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせを、第２のパターンということにする。前記第２の頻度は、前記ウィンドウ期間中に前記出力メッセージと同じ種別のメッセージが出力され、かつ、前記第２のパターンに基づいて前記ある種別の障害の発生が予測された頻度である。

【００１８】

さらに、前記検出方法は、前記コンピュータが、前記複数の構成アイテムの中で相対的に高い蓋然性で前記ある種別の障害が発生すると予測される１つ以上の構成アイテムを示す結果情報を、前記統計値に基づいて生成することを含む。

【発明の効果】

【００１９】

上記の検出方法によれば、障害の発生を防止するうえで有益な情報が検出される。

【図面の簡単な説明】

【００２０】

【図１】第１実施形態のコンピュータが実行する処理のフローチャートである。

【図２】コンピュータのハードウェア構成図である。

【図３】コンピュータシステムの例を示す図である。

【図４】第２実施形態の検出サーバの動作を例示する図である。

【図５】第２実施形態の検出サーバのブロック構成図である。

【図６】第２実施形態で利用される各種テーブルの例を示す図である。

【図７】第２実施形態の検出サーバが行う処理のフローチャートである。

【図８】第３実施形態における関係情報の学習を説明する図である。

【図９】第３実施形態におけるランキングの改良について説明する図である。

【図１０】第３実施形態の検出サーバのブロック構成図である。

【図１１】第３実施形態で利用される各種テーブルの例を示す図である。

【図１２】第３実施形態において検出サーバが関係情報を学習する処理のフローチャートである。

【図１３】第３実施形態の検出サーバが、学習した関係情報を使って改良ランキング情報を生成する処理のフローチャート（その１）である。

【図１４】第３実施形態の検出サーバが、学習した関係情報を使って改良ランキング情報を生成する処理のフローチャート（その２）である。

【発明を実施するための形態】

【００２１】

以下、実施形態について、図面を参照しながら詳細に説明する。具体的には、図１を参照して第１実施形態についてまず説明し、その後、第１～第３実施形態に共通する点について図２～３の例を参照しながら説明する。そして、図４～８を参照して第２実施形態について説明し、図９～１３を参照して第３実施形態について説明する。最後にその他の変形例についても説明する。

【００２２】

図１は、第１実施形態のコンピュータが実行する処理のフローチャートである。第１実施形態のコンピュータは、コンピュータシステムを管理する。

コンピュータシステムには、複数の構成アイテム（configuration items）が含まれる

。構成アイテムの数は任意である。例えば、クラウド環境では、構成アイテムの数が数千から数万のオーダである場合もある。

【００２３】

各構成アイテムは、コンピュータシステムに含まれるハードウェア、ソフトウェア、または両者の組み合わせである。例えば、物理サーバ、Ｌ２（layer 2）スイッチ、Ｌ３（layer 3）スイッチ、ルータ、ディスクアレイ装置などのハードウェア装置は、いずれも、構成アイテムの例である。また、ＯＳ（Operating System）、ミドルウェア、アプリケーションソフトウェアなどの種々のソフトウェアは、いずれも、構成アイテムの例である。構成アイテムの粒度（granularity）によっては、例えば、あるハードウェア装置と、当該ハードウェア装置上で動作するソフトウェアの組み合わせが、１つの構成アイテムとして扱われてもよい。例えば、ある１つの構成アイテムは、ルータと、当該ルータ上で動作するファームウェアの組み合わせであってもよい。

【００２４】

コンピュータシステムの構成によっては、ある構成アイテムが、ある物理マシン上で直接動作する（running）ＯＳであってもよい。また、別のある構成アイテムは、ハイパーバイザにより仮想化された物理マシン上で動作する仮想マシンのＯＳであってもよい。もちろん、ハイパーバイザ以外の仮想化技術が使われていてもよい。

【００２５】

ハイパーバイザ上で実行される仮想マシンは、実装に応じて、「仮想マシン」、「ドメイン」、「論理ドメイン」、「パーティション」などの名前で呼ばれることがある。また、ハイパーバイザ上では２つ以上の仮想マシンが実行され得るが、ある種の実装によれば、ある特定の仮想マシンが特別な役割を果たす。この特定の仮想マシンは「ドメイン０」や「制御ドメイン」などと呼ばれることがあり、その他の仮想マシンは「ドメインＵ」や「ゲストドメイン」などと呼ばれることがある。

【００２６】

また、特定の仮想マシン上のＯＳは「管理ＯＳ」または「ホストＯＳ」などと呼ばれることがあり、他の仮想マシン上のＯＳは「ゲストＯＳ」などと呼ばれることがある。例えば、ある種の実装によれば、ゲストＯＳは、ハイパーバイザを介してホストＯＳのデバイスドライバの機能を利用することにより、ハードディスク装置などのデバイスにアクセスする場合がある。

【００２７】

ところで、コンピュータシステムにおける障害（failure）の予兆（predictor）（すなわち障害の兆候（sign））を検出するための技術はいくつか提案されているが、単に予兆を検出するだけでは、実際の障害の発生を未然に防ぐには不十分な場合がある。具体的には、「コンピュータシステム内のどの構成アイテムに対して対策をとれば障害の発生を防止するうえで有益なのか」ということが不明だと、「障害の発生を防止する」という目的が十分に達せられないことがあり得る。例えば、「コンピュータシステム内のどの構成アイテムで障害が発生しそうなのか」ということが不明だと、「どの構成アイテムに対して対策をとることが有益なのか」も不明である。

【００２８】

そこで、第１実施形態のコンピュータは、図１のフローチャートにしたがって、「コンピュータシステム内のどの構成アイテムに対して対策をとれば、障害の発生を防止するうえで有益なのか」を示唆する情報を生成し、出力する。つまり、第１実施形態によれば、障害の発生を防止するうえで有益な情報が検出される。

【００２９】

まず、ステップＳ１において、コンピュータは、複数の種別のうちのある種別の障害の発生を予測する。または、ステップＳ１において、コンピュータは、当該ある種別の障害の発生が予測されることを示す予測通知を受け取る。

【００３０】

具体的には、コンピュータ自体が予測を行う場合、コンピュータは、Ｐ個のメッセージ

の組み合わせパターンである第１のメッセージパターンに基づいて、上記ある種別の障害の発生を予測する。第１のメッセージパターンは、換言すれば、Ｐ個のメッセージの組み合せであるような、第１のパターンである。ここで、Ｐ個のメッセージの各々は、コンピュータシステムの上記複数の構成アイテムのうちのＱ個の構成アイテムのいずれかから出力されたメッセージである（１≦Ｑ≦Ｐ）。また、Ｐ個のメッセージは、ある所定時間（以下では「第１の所定時間」という）以下の長さの期間中に出力されたものとする。Ｐ個のメッセージの各々は、具体的には、イベントの発生を知らせるメッセージである。

【００３１】

第１の所定時間の長さは実施形態に応じて任意である。例えば、第１の所定時間の長さは、１～５分間程度の長さであってもよいし、それより短くても、それより長くてもよい。

【００３２】

例えば、第１の所定時間の長さが５分間であり、コンピュータシステムに１０００個の構成アイテムが含まれ、ある５分間の間に１０００個の構成アイテムのうち３０個から、合計で５０個のメッセージが出力されるとする。この場合、Ｑ＝３０かつＰ＝５０である。このようにＱ＜Ｐの場合、少なくとも１つの構成アイテムは、上記期間中に２個以上のメッセージを出力している。もちろん、上記３０個の構成アイテムの中には、上記期間中に１個しかメッセージを出力しないものがあってもよい。

【００３３】

また、各メッセージにより通知されるイベントの種別は任意である。例えば、「あるデバイスがオープンされた」、「ウェブページへのアクセスが拒否された（denied）」、「物理サーバが再起動された」など、様々な種別のイベントがあり得る。イベントを通知するメッセージは、「イベントログ」や「メッセージログ」などの名前で呼ばれることもあり、単に「ログ」と呼ばれることもある。

【００３４】

コンピュータは、予め、「ある特定の１つ以上の種別のイベントが第１の所定時間以下の長さの期間中に生じる場合は、ある特定の種別の障害が発生しやすい」といった共起情報を学習してもよい。コンピュータは、学習した共起情報に基づいて、ステップＳ１で上記第１のメッセージパターン（すなわちＰ個のメッセージの組み合わせパターン）に基づいて、上記ある種別の障害の発生を予測してもよい。

【００３５】

あるいは、上記のとおり、コンピュータは、ステップＳ１において、自ら予測を行う代わりに、予測通知を受け取ってもよい。予測通知は、例えば、予測を行う他のコンピュータから、ネットワークを介して送信されてもよい。予測通知は、具体的には、「第１のメッセージパターンから上記ある種別の障害の発生が予測される」ということを示す。

【００３６】

いずれにしろ、コンピュータは、「第１のメッセージパターンが、上記ある種別の障害の予兆である」ということを認識することができる。しかし、上記のとおり、単に障害の予兆が検出されるだけでは、不十分である。

【００３７】

つまり、「コンピュータシステム中のどの構成アイテムに対して対策をとるのが有効なのか」ということが不明だと、障害の回避に失敗することがあり得る。一方、コンピュータシステムの可用性の向上という効果を得るには、障害を未然に防ぐことが有益である。そして、障害を未然に防ぐには、適宜の対策をとることが有益である。対策の例として、例えば、ハードウェアの交換、ハードウェアの増設、ハードウェアまたはソフトウェアの再起動、ソフトウェアのアップグレード、ソフトウェアの再インストールなどが挙げられる。

【００３８】

第１実施形態のコンピュータは、どの構成アイテムに対して対策をとることが有益なのかを示す情報を、システム管理者等の人間に提示するために、さらにステップＳ２～Ｓ４

の処理を行う。つまり、第１のパターンに基づき上記ある種別の障害の発生が予測される場合に、コンピュータは、ステップＳ２～Ｓ４の処理を行う。

【００３９】

ステップＳ２でコンピュータは、Ｑ個の構成アイテムの各々について統計値（statistic）を算出する。ある構成アイテムについて算出される統計値は、具体的には、第１のメッセージパターンから予測される上記ある種別の障害が、当該ある構成アイテムにおいて将来発生する蓋然性に関する値である。

【００４０】

なお、当該統計値は、蓋然性そのものの値である必要はない。例えば、当該統計値は、蓋然性が高いほど大きくなるような適宜の値であってよい。

【００４１】

コンピュータは、具体的には、以下に説明する第１の頻度と第２の頻度に基づいて、統計値を算出する。

【００４２】

ここで、予測された上記ある種別の障害が、過去に実際に発生した時点を、「発生時点」ということにする。また、Ｐ個のメッセージのうち、統計値の算出対象たる上記ある構成アイテムが出力したメッセージを、「出力メッセージ」ということにする。そして、発生時点より前に出力メッセージと同じ種別のメッセージが出力された頻度を、「第１の頻度」ということにする。ここでの「頻度」は、何らかの広い意味における頻度であってよく、したがって、第１の頻度の具体的な数学的定義は様々であってよい。つまり、「発生時点より前に、出力メッセージと同じ種別のメッセージが、コンピュータシステムに含まれる複数の構成アイテムからどれほど多く出力されたか」を示すような種々の頻度が、第１の頻度として利用可能である。

【００４３】

例えば、第１の頻度は、出力メッセージと同じ種別のメッセージが、上記複数の構成アイテムのいずれかから発生時点より前に出力された頻度の生の値そのものであってもよい。あるいは、何らかのメッセージ（出力メッセージと同じ種別のメッセージでもよいし、出力メッセージと別の種別のメッセージでもよい）が出力された時点を含み、当該時点から第１の所定時間だけ遡る期間が「ウィンドウ期間」として定義されてもよい。第１の頻度は、出力メッセージと同じ種別のメッセージが、発生時点よりも前の全ウィンドウ期間に合計で何回出現するかを示す値であってもよい。または、第１の頻度は、発生時点よりも前の全ウィンドウ期間のうち、出力メッセージと同じ種別のメッセージを含むウィンドウ期間の数であってもよい。

【００４４】

例えば、メッセージが出力されるタイミングと、第１の所定時間の長さによっては、出力メッセージと同じ種別の１つのメッセージが、３つのウィンドウ期間に含まれる場合があり得る。この場合、第１の頻度の具体的定義に応じて、当該１つのメッセージに対応して、第１の頻度は、１だけインクリメントされてもよいし、３だけインクリメントされてもよい。いずれにせよ、第１の頻度は、「出力メッセージと同じ種別のメッセージが発生時点より前にどれほど多く出力されたか」ということを示す。また、第１の頻度は、絶対頻度であってもよいし、相対頻度であってもよい。

【００４５】

なお、１つのコンピュータシステムの中に同じ種別の２つ以上の構成アイテムが含まれる場合などには、２つ以上の構成アイテムが同じ種別のメッセージを出力することもあり得る。しかし、コンピュータが第１の頻度を数える際には、「どの構成アイテムからメッセージが出力されたか」は問わない。第１の頻度は、障害の発生とは関係なく、「出力メッセージはどの程度一般的な種別のメッセージなのか」ということを示す尺度である。第１の頻度が高ければ、出力メッセージは一般的な種別のメッセージであり、第１の頻度が低ければ、出力メッセージは珍しい種別のメッセージである。

【００４６】

また、上記の発生時点より前に（具体的には、上記の発生時点から第２の所定時間以内の過去において）、いずれかのメッセージが出力された時点を、「出力時点」ということにする。そして、出力時点を含み、かつ、出力時点から第１の所定時間だけ遡る期間を、「ウィンドウ期間」ということにする。なお、発生時点から第２の所定時間以内の過去には、２つ以上のメッセージが出力された可能性もある。その場合は、各メッセージについて、出力時点とウィンドウ期間が定義される。

【００４７】

第１の所定時間と第２の所定時間は、どちらが長くてもよいし、互いに等しくてもよい。例えば、第１の所定時間が５分間で、第２の所定時間が１時間の場合、ウィンドウ期間は、実際に上記ある種別の障害が発生した発生時点から１時間以内の過去において何らかのメッセージが出力された時点を終了時点とする、長さ５分間の期間である。この５分間のウィンドウ期間中に出力されたメッセージの数は、１つの場合もあり得るし、２つ以上の場合もあり得る。以下では、ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせパターンを「第２のメッセージパターン」ということにする。第２のメッセージパターンは、換言すれば、ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせであるような、第２のパターンである。

【００４８】

さらに、ウィンドウ期間中に出力メッセージと同じ種別のメッセージが上記複数の構成アイテムのいずれかから出力され、かつ、第２のメッセージパターンに基づいて上記ある種別の障害の発生が予測された頻度を、「第２の頻度」ということにする。第２の頻度の具体的な数学的定義も、様々であってよい。例えば、第２の頻度は、絶対頻度であってもよいし、相対頻度であってもよい。

【００４９】

ここで、「第２のメッセージパターンに基づいて上記ある種別の障害の発生が予測された」とは、換言すれば、「第２のメッセージパターンに基づく過去の予測が正しかった」ということを意味する。なぜなら、発生時点とは、上記ある種別の障害が過去に実際に発生した時点であり、上記の定義より、第２のメッセージパターン中の各メッセージが出力された時点は、発生時点よりも前のウィンドウ期間内に属するからである。

【００５０】

よって、「第２のメッセージパターンに基づいて上記ある種別の障害の発生が予測された」という条件下では、「ウィンドウ期間中に出力メッセージと同じ種別のメッセージが上記複数の構成アイテムのいずれかから出力される」ことは、以下のことを意味する。すなわち、これは、過去の正しい予測において予測の根拠に使われた第２のメッセージパターンの中に、出力メッセージと同じ種別のメッセージが含まれていたことを示す。

【００５１】

よって、第２の頻度は、上記ある種別の障害に関して、出力メッセージと同じ種別のメッセージを含むメッセージパターンを根拠として過去に行われた予測が、正解だった頻度を示す。ある観点によれば、第２の頻度は、上記ある種別の障害に関する正しい予兆検出に、出力メッセージと同じ種別のメッセージが、どの程度深く関連しているかを示す尺度である。

【００５２】

なお、第１のメッセージパターンと第２のメッセージパターンは、偶然同じパターンである場合もあり得るし、互いに異なる場合もあり得る。換言すれば、２つ以上の異なるメッセージパターンから、同じ１つの種別の障害が予測される可能性がある。つまり、ある１つの種別の障害の予兆は、２通り以上存在する可能性がある。

【００５３】

一方で、同じ１つの種別の障害の予兆となる（be predictive of）２つ以上のメッセージパターンの中には、ある共通の１つの種別のメッセージが含まれる可能性もある。よって、ある観点によれば、第２の頻度は、正解した過去の１回または複数回の予測においてそれぞれ根拠として使われたメッセージパターン中に、どの程度頻繁に、出力メッセージ

と同じ種別のメッセージが含まれていたかを示す尺度である。

【００５４】

ステップＳ２における統計値の算出は、以上のような第１の頻度と第２の頻度に基づいて行われる。第１の頻度と第２の頻度から統計値を導出するための計算式は、実施形態に応じて適宜定義されてよいが、統計値は、第１の頻度に対して単調減少するとともに第２の頻度に対して単調増加する値であることが好ましい。

【００５５】

なぜなら、このように統計値が定義されると、予測された上記ある種別の障害とはとりわけよく共起する（しかし、他の種別の障害とはあまり共起しない）ような種別のメッセージを出力した構成アイテムに対して、大きな値が統計値として算出されるからである。つまり、予測された上記ある種別の障害を特徴づけるような、特定の種別のメッセージを出力した構成アイテムに対して、大きな値が統計値として算出されるからである。

【００５６】

後述の第２～第３実施形態で使われる統計値ＷＦ－ＩＤＦ（ｆ，ｎ）は、第１の頻度に対して単調減少するとともに第２の頻度に対して単調増加する統計値の一例である。

なお、第１の頻度は、図１の処理を実行するコンピュータ自体が数えてもよいし、他のコンピュータが数えてもよい。例えば、図１の処理を実行するコンピュータは、コンピュータシステムに含まれる複数の構成アイテムのいずれかからメッセージが出力されるたびに、当該メッセージの種別に対応づけられて記憶装置に記憶された第１のカウント値を更新してもよい。この場合、コンピュータは、第１のカウント値から第１の頻度を算出してもよい。

【００５７】

同様に、第２の頻度も、図１の処理を実行するコンピュータ自体が数えてもよいし、他のコンピュータが数えてもよい。例えば、図１の処理を実行するコンピュータは、複数の種別のうちのいずれかの種別の障害が実際に発生するたびに、下記の２つの種別の組み合わせに対応づけられて記憶装置に記憶された第２のカウント値を更新してもよい。

【００５８】

・当該発生した障害を正しく予測する根拠となった第２のメッセージパターンに含まれる各メッセージの種別。

・当該発生した障害の種別。

【００５９】

例えば、第２のメッセージパターンに４個のメッセージが含まれ、それぞれの種別が互いに異なる場合、コンピュータは、４個のメッセージそれぞれに対応する４個の第２のカウント値をそれぞれ更新する。このように第２のカウント値が使われる場合、コンピュータは、第２のカウント値から第２の頻度を算出してもよい。

【００６０】

さて、以上説明したようにしてステップＳ２でＱ個の構成アイテムの各々について統計値を算出した後、コンピュータは、ステップＳ３の処理を実行する。具体的には、コンピュータは、Ｑ個の構成アイテムの各々について算出した統計値に基づいて、結果情報を生成する。結果情報は、コンピュータシステムに含まれる複数の構成アイテムの中で、相対的に高い蓋然性で、「第１のメッセージパターンから予測される上記ある種別の障害が、将来発生する」と予測される、１つ以上の構成アイテムを示す。具体的には、結果情報は、当該１つ以上の構成アイテムをそれぞれ識別する識別情報を含む。

【００６１】

識別情報は、例えばＩＰアドレスであってもよいし、その他の情報であってもよい。例えば、以下に挙げる情報のいずれか１つ、または、以下に挙げる情報の２つ以上の組み合わせが、識別情報として使われてもよい。

【００６２】

・ＩＰアドレス。

・ＴＣＰ（Transmission Control Protocol）ポート番号。

・ホスト名。
・ホスト名を含むＦＱＤＮ（Fully Qualified Domain Name）。
・ＭＡＣ（Media Access Control）アドレス。
・アプリケーション名。
・ＣＭＤＢ（Configuration Management Database）において各構成アイテムに割り当てられている識別子。
・ハードウェア装置の製造シリアル番号。

【００６３】
そして、ステップＳ４でコンピュータは、結果情報を出力する。具体的には、コンピュータは、例えば、結果情報をディスプレイに表示してもよいし、マイクから音声的に結果情報を出力してもよいし、プリンタに結果情報を出力してもよい。また、コンピュータは、結果情報を含む電子メールまたはインスタントメッセージを生成し、生成した電子メールまたはインスタントメッセージを、システム管理者宛に送信してもよい。もちろん、コンピュータは、結果情報を不揮発性記憶装置に出力してもよい。このように、ステップＳ４における出力の具体的方法は、実施形態に応じて様々である。ステップＳ４の出力後、図１の処理は終了する。

【００６４】
なお、結果情報は、例えば、Ｑ個の構成アイテムのうちで統計値が最大の構成アイテムを識別する識別情報を含むことが好ましい。なぜなら、統計値が最大の構成アイテムは、ある観点によれば、障害が発生する蓋然性が最も高いと推定され、障害の予測において最も重要と推定されるからである。場合によっては、重要と推定された構成アイテムそのものに対して何らかの対策をとることが、障害の発生を未然に防ぐうえで有益なこともある。管理者等は、障害の予測において重要と推定された各構成アイテムに関して、何らかの対策をとるかどうかを判断し、判断に応じて適宜の対策をとってもよい。

【００６５】
実施形態によっては、コンピュータは、ステップＳ３において、Ｑ個の構成アイテムを統計値にしたがってソートしてもよく、ソート結果に応じてＱ個の構成アイテムに順位をつけてもよい。そして、コンピュータは、Ｑ個の構成アイテムすべて（あるいはＱ個のうち、相対的に順位が上のいくつかの構成アイテム）それぞれの識別情報を、順位および／または統計値と対応づけてもよい。結果情報は、以上のようにしてそれぞれ順位および／または統計値と対応づけられた、Ｑ個（またはそれ以下）の識別情報を含む情報であってもよい。

【００６６】
また、コンピュータは、ステップＳ３において、Ｑ個の構成アイテムそれぞれの統計値に基づいて、それらＱ個の構成アイテム以外の構成アイテムも含めて、いくつかの構成アイテムについて、上記ある種別の障害が将来発生する蓋然性を評価してもよい。そして、コンピュータは、その評価の結果に基づく結果情報を、ステップＳ３で生成してもよい。

【００６７】
例えば、コンピュータは、Ｐ個のメッセージの各々について、以下に説明する「関連構成アイテム」を検索してもよい。具体的には、コンピュータは、コンピュータシステムに含まれる複数の構成アイテム間の関係を示す構成情報を用いて、関連構成アイテムを検索してもよい。

【００６８】
ここで、以下の２つの条件に当てはまるメッセージを出力した構成アイテムのことを、「第１の構成アイテム」ということにする。

【００６９】
・Ｐ個のメッセージのうちで関連構成アイテムを検索する対象として現在コンピュータが着目している当該メッセージと、同じ種別のメッセージである。
・上記ある種別の障害の発生が過去に正しく予測された際の予測に使われた、上記第２のメッセージパターンに含まれるメッセージである。

【００７０】
また、過去に正しく予測された上記ある種別の障害が、実際に発生した構成アイテムのことを、「第２の構成アイテム」ということにする。そして、第１の構成アイテムと第２の構成アイテムとの間の関係を「第１の関係」ということにする。
【００７１】
コンピュータは、Ｐ個のメッセージの各々について、当該メッセージを出力した構成アイテムとの間に、第１の関係と等価な第２の関係が成り立つ（hold true）ような構成アイテムを、「関連構成アイテム」として検索してもよい。より具体的には、コンピュータは、コンピュータシステムに含まれる上記複数の構成アイテムの中から、構成情報を用いて、上記のごとき関連構成アイテムを検索してもよい。
【００７２】
なお、構成情報により示される関係は、例えば、以下のようないずれの関係であってもよい。
【００７３】
・２つの構成アイテム間の論理的依存関係（logical dependency）。例えば、物理サーバと、当該物理サーバ上で動作するホストＯＳとの間の関係や、ホストＯＳとゲストＯＳとの間の関係など。
・２つの構成アイテム間の物理的接続関係。例えば、物理サーバと、当該物理サーバに接続されるＬ２スイッチとの間の関係など。
・２つ以上の論理的依存関係の合成（composition）。例えば、物理サーバとホストＯＳとの間の論理的依存関係と、ホストＯＳとゲストＯＳとの間の論理的依存関係との合成（すなわち、物理サーバとゲストＯＳとの間の、間接的な論理的依存関係）など。
・２つ以上の物理的接続関係の合成。例えば、物理サーバとＬ２スイッチとの間の物理的接続関係と、Ｌ２スイッチとルータとの間の物理的接続関係との合成（すなわち、物理サーバとルータとの間の、間接的な物理的接続関係）など。
・１つ以上の論理的依存関係と１つ以上の物理的接続関係の合成。例えば、ホストＯＳと、当該ホストＯＳが動作する物理サーバに接続されたストレージ装置との間の関係や、１台のＬ２スイッチに接続された２台の物理サーバ上でそれぞれ動作する２つのホストＯＳ同士の関係など。
【００７４】
さて、上記のような構成情報を用いた検索の結果、Ｑ個の構成アイテムのうちのある構成アイテムについて、関連構成アイテムが見つかった場合は、コンピュータは、次のような処理を行ってもよい。すなわち、コンピュータは、第１のメッセージパターンから予測される上記ある種別の障害が、当該関連構成アイテムにおいて将来発生する蓋然性に関する評価値を決定してもよい。当該関連構成アイテムについての評価値の決定は、具体的には、Ｑ個の構成アイテムのうち、当該関連構成アイテムが見つかった当該ある構成アイテムについてステップＳ２で算出済みの統計値に基づく。
【００７５】
なお、Ｑ個の構成アイテムのうちの１つの構成アイテムについて、２つ以上の関連構成アイテムが見つかる場合もあり得る。また、Ｑ個の構成アイテムのうちの２つ以上の構成アイテムについて、たまたま同じ構成アイテムが、それぞれの関連構成アイテムとして見つかる場合もあり得る。いずれにしろ、コンピュータは、統計値を算出した対象の構成アイテムに関して見つかった関連構成アイテムの評価値に、当該統計値を反映させる。
【００７６】
以上のような処理により、検索の結果見つかった各関連構成アイテムについて、評価値が決定されてもよい。この場合、コンピュータは、検索の結果見つかった各関連構成アイテムについて決定した評価値に基づいて、結果情報を生成してもよい。
【００７７】
例えば、Ｑ個の構成アイテムの中の少なくとも１つに関して、複数の構成アイテムの中から、関連構成アイテムとして、検索の結果見つかった構成アイテムが、１つ以上あると

する。この場合、結果情報は、これらの１つ以上の関連構成アイテムのうちで、評価値が最大の構成アイテムを識別する識別情報を含んでもよい。なぜなら、ある観点によれば、評価値が最大の構成アイテムは、障害が発生する蓋然性が最も高いと推定され、障害の予測において最も重要と推定されるからである。障害の予測において最も重要と推定される構成アイテムに対して対策をとることが、障害の発生を未然に防ぐうえで有益な場合もあり得る。

【００７８】

また、コンピュータは、評価値の決定された全構成アイテム（すなわち、検索の結果見つかった全関連構成アイテム）を、評価値にしたがってソートしてもよく、ソート結果に応じてこれらの構成アイテムに順位をつけてもよい。そして、コンピュータは、順位づけした全構成アイテム（または、そのうち順位が上のいくつかの構成アイテム）それぞれの識別情報を、順位および／または評価値と対応づけてもよい。結果情報は、以上のようにしてそれぞれ順位および／または評価値と対応づけられた、いくつかの識別情報を含む情報であってもよい。

【００７９】

以上のような構成情報を用いた検索と評価値の決定が行われるにしろ、行われないにしろ、ステップＳ３では、Ｑ個の統計値に基づいて結果情報が生成される。そして、ステップＳ４では結果情報が出力される。よって、システム管理者等の人間は、結果情報を参照することにより、「予測された障害はどの構成アイテムと関連性が高いか」を適切に判断することができる。管理者等は、結果情報に基づいて、「障害の発生を防ぐうえで、どの構成アイテムについて対策を講じることが有益か」ということを適切に判断することもできる。結果情報は当該判断を助ける情報である。なお、構成情報を用いた検索と評価値の決定に関する更に詳しい例は、第３実施形態とともに後述する。

【００８０】

さて、図２は、コンピュータのハードウェア構成図である。図１の処理を実行するコンピュータは、具体的には、図２のコンピュータ１００であってもよい。

コンピュータ１００は、ＣＰＵ（Central Processing Unit）１０１と、ＲＡＭ（Random Access Memory）１０２と、通信インタフェイス１０３を有する。コンピュータ１００はさらに、入力装置１０４と、出力装置１０５と、記憶装置１０６と、コンピュータ読み取り可能な記憶媒体１１０の駆動装置１０７を有する。コンピュータ１００のこれらの構成要素は、互いにバス１０８で接続されている。

【００８１】

ＣＰＵ１０１は、シングルコアまたはマルチコアのプロセッサの一例である。コンピュータ１００は複数のプロセッサを有していてもよい。ＣＰＵ１０１はプログラムをＲＡＭ１０２にロードし、ＲＡＭ１０２をワーキングエリアとしても利用しながら、プログラムを実行する。例えば、ＣＰＵ１０１は、図１の処理のためのプログラムを実行してもよい。

【００８２】

通信インタフェイス１０３は、例えば、有線ＬＡＮ（Local Area Network）インタフェイス、無線ＬＡＮインタフェイス、またはその組み合わせである。コンピュータ１００は、通信インタフェイス１０３を介してネットワーク１２０に接続される。

【００８３】

通信インタフェイス１０３は、具体的には、外付けのＮＩＣ（Network Interface Card）でもよいし、オンボード型のネットワークインタフェイスコントローラでもよい。例えば、通信インタフェイス１０３は、物理層の処理を行う「ＰＨＹチップ」と呼ばれる回路と、ＭＡＣ副層の処理を行う「ＭＡＣチップ」と呼ばれる回路を含んでいてもよい。

【００８４】

入力装置１０４は、例えば、キーボード、ポインティングデバイス、またはその組み合わせである。ポインティングデバイスは、例えば、マウスでもよいしタッチパッドでもよいしタッチスクリーンでもよい。

【００８５】
出力装置１０５は、ディスプレイ、スピーカ、またはその組み合わせである。ディスプレイはタッチスクリーンであってもよい。
【００８６】
記憶装置１０６は、具体的には、１つ以上の不揮発性の記憶装置である。記憶装置１０６は、例えば、ＨＤＤ（Hard Disk Drive）でもよいし、ＳＳＤ（Solid-State Drive）でもよいし、両者の組み合わせでもよい。さらにＲＯＭ（Read Only Memory）が記憶装置１０６として含まれていてもよい。
【００８７】
記憶媒体１１０の例は、ＣＤ（Compact Disc）やＤＶＤ（Digital Versatile Disk）などの光ディスク、光磁気ディスク、磁気ディスク、フラッシュメモリなどの半導体メモリカードなどである。
【００８８】
ＣＰＵ１０１が実行するプログラムは、予め記憶装置１０６にインストールされていてもよい。あるいは、プログラムは、記憶媒体１１０に格納されて提供され、記憶媒体１１０から駆動装置１０７により読み取られて記憶装置１０６にコピーされ、その後、ＲＡＭ１０２にロードされてもよい。または、ネットワーク１２０上のプログラム提供者１３０から、ネットワーク１２０と通信インタフェイス１０３を介して、プログラムがコンピュータ１００にダウンロードされ、インストールされてもよい。プログラム提供者１３０は、具体的には、他のコンピュータである。
【００８９】
なお、ＲＡＭ１０２、記憶装置１０６、および記憶媒体１１０は、いずれも、コンピュータ読み取り可能な有形の（tangible）媒体であり、信号搬送波のような一時的な（transitory）媒体ではない。
【００９０】
図２のコンピュータ１００は、図１に関して説明したコンピュータシステムと、ネットワーク１２０を介して接続されていてもよい。
コンピュータ１００は、コンピュータシステムに含まれる任意の構成アイテムから、ネットワーク１２０と通信インタフェイス１０３を介してメッセージを受信してもよく、受信したメッセージを記憶装置１０６に記憶してもよい。あるいは、構成アイテムから出力された各メッセージは、当該メッセージを出力した構成アイテムの識別情報（例えばＩＰアドレス）とともに、不図示の他のコンピュータの記憶装置に記憶されてもよい。コンピュータ１００は、ネットワーク１２０と通信インタフェイス１０３を介して記憶装置にアクセスし、記憶されたメッセージを読み出してもよい。
【００９１】
いずれにしろ、コンピュータ１００は、図１のステップＳ１に関して説明したＰ個のメッセージを取得することができる。よって、コンピュータ１００（より具体的にはＣＰＵ１０１）は、Ｐ個のメッセージから、上記ある種別の障害の発生を予測することができる。
【００９２】
あるいは、コンピュータ１００がＰ個のメッセージ自体を取得しない実施形態も可能である。つまり、コンピュータ１００は、上記ある種別の障害の発生が予測されることを示す予測通知を、ネットワーク１２０と通信インタフェイス１０３を介して、ステップＳ１で受信してもよい。この場合、予測通知には、Ｐ個のメッセージの各々がどの構成アイテムから出力されたのかを示す情報（例えばＰ個のＩＰアドレス）が含まれる。
【００９３】
よって、ステップＳ１でコンピュータ１００自体が予測を行うにしろ、コンピュータ１００が予測通知を受け取るにしろ、コンピュータ１００は、各メッセージを出力した構成アイテムを認識することもできる。
【００９４】

また、図１のステップＳ２に関して説明したように、第１の頻度は、コンピュータ１００（より具体的にはＣＰＵ１０１）自体によって数えられてもよい。この場合、第１の頻度（またはその算出に利用される第１のカウント値）は、記憶装置１０６またはＲＡＭ１０２に記憶される。あるいは、第１の頻度は、他のコンピュータによって数えられてもよい。この場合、コンピュータ１００は、ネットワーク１２０と通信インタフェイス１０３を介して、第１の頻度を取得してもよい。

【００９５】

第２の頻度も同様に、ＣＰＵ１０１によって数えられてもよいし、ネットワーク１２０と通信インタフェイス１０３を介して取得されてもよい。つまり、第２の頻度（またはその算出に利用される第２のカウント値）も、記憶装置１０６またはＲＡＭ１０２に記憶されてもよい。

【００９６】

いずれにしろ、コンピュータ１００（より具体的にはＣＰＵ１０１）は、上記Ｐ個のメッセージの組み合わせパターンである上記第１のメッセージパターンと、第１の頻度と、第２の頻度を認識することができる。また、コンピュータ１００は、Ｐ個のメッセージの各々がどの構成アイテムから出力されたのかということも、認識することができる。したがって、コンピュータ１００は、ステップＳ２で、上記Ｑ個の構成アイテムのそれぞれについて統計値を算出することができる。

【００９７】

さらに、コンピュータ１００は、算出したＱ個の統計値を用いて、ステップＳ３で結果情報を生成することもできる。なお、コンピュータ１００が結果情報の生成に構成情報を利用する場合、構成情報は、コンピュータ１００自体の記憶装置１０６に記憶されていてもよい。あるいは、ネットワーク１２０を介してコンピュータ１００に接続された記憶装置に構成情報が記憶されていてもよい。

【００９８】

また、ステップＳ４では、コンピュータ１００は、結果情報を出力装置１０５に出力してもよく、記憶装置１０６に出力してもよく、駆動装置１０７を介して記憶媒体１１０に出力してもよい。コンピュータ１００は、ネットワーク１２０を介して接続された他の装置（例えば、他のコンピュータ、ネットワークストレージ装置、プリンタなど）に結果情報を出力してもよい。また、コンピュータ１００は、結果情報を含む電子メールまたはインスタントメッセージを生成し、生成した電子メールまたはインスタントメッセージを、通信インタフェイス１０３とネットワーク１２０を介して送信してもよい。

【００９９】

以上説明したように、図１の処理は、図２のコンピュータ１００により実行されてもよい。

【０１００】

さて、図３は、コンピュータシステムの例を示す図である。図３には、コンピュータ２００と、コンピュータ２００が接続されたネットワーク２１０と、ネットワーク２１０に接続されたコンピュータシステム２３０が例示されている。コンピュータ２００は、具体的には、図１の処理を実行するコンピュータである。コンピュータ２００が図２のコンピュータ１００であってもよく、その場合、ネットワーク２１０は図２のネットワーク１２０である。

【０１０１】

コンピュータシステム２３０は、４台の物理サーバと２台のＬ２スイッチと１台のＬ３スイッチを含む。具体的には、図３の例では、物理サーバ２４０と２５０がＬ２スイッチ２８０に接続されており、物理サーバ２６０と２７０がＬ２スイッチ２８１に接続されており、Ｌ２スイッチ２８０と２８１がＬ３スイッチ２９０に接続されている。そして、Ｌ３スイッチ２９０はネットワーク２１０に接続されている。

【０１０２】

物理サーバ２４０はハイパーバイザ２４１により仮想化されている。具体的には、ハイ

パーバイザ２４１上で、ホストＯＳ２４２とゲストＯＳ２４３とゲストＯＳ２４４が動作する。

【０１０３】

同様に、物理サーバ２５０はハイパーバイザ２５１により仮想化されている。具体的には、ハイパーバイザ２５１上で、ホストＯＳ２５２とゲストＯＳ２５３とゲストＯＳ２５４が動作する。

【０１０４】

同様に、物理サーバ２６０はハイパーバイザ２６１により仮想化されている。具体的には、ハイパーバイザ２６１上で、ホストＯＳ２６２とゲストＯＳ２６３が動作する。

【０１０５】

同様に、物理サーバ２７０はハイパーバイザ２７１により仮想化されている。具体的には、ハイパーバイザ２７１上で、ホストＯＳ２７２とゲストＯＳ２７３が動作する。

【０１０６】

例えば、以下に挙げるハードウェアとソフトウェアは、コンピュータシステム２３０に含まれる構成アイテムの例である。

【０１０７】

・物理サーバ２４０、２５０、２６０、および２７０の各々。
・Ｌ２スイッチ２８０と２８１の各々。
・Ｌ３スイッチ２９０。
・ハイパーバイザ２４１、２５１、２６１、および２７１の各々。
・ホストＯＳ２４２、２５２、２６２、および２７２の各々。
・ゲストＯＳ２４３、２４４、２５３、２５４、２６３、および２７３の各々。
・ゲストＯＳ上で動作する不図示の各アプリケーション。

【０１０８】

なお、構成アイテムの粒度は実施形態に応じて様々であってよい。各構成アイテムを識別する識別情報は、構成アイテムの粒度に応じて、個々の構成アイテムを識別可能な情報であれば、どのような情報であってもよい。識別情報の例は上述したとおりである。

【０１０９】

識別情報の粒度によっては、いくつかのハードウェアの集合、いくつかのソフトウェアの集合、または１つ以上のハードウェアと１つ以上のソフトウェアの集合が、１つの構成アイテムとして扱われてもよい。例えば、識別情報としてＩＰアドレスが使われる場合には、ゲストＯＳと複数のアプリケーションを含む集合全体が、１つの構成アイテムとして扱われてもよい。なぜなら、ゲストＯＳと、ゲストＯＳ上の複数のアプリケーションは、同じＩＰアドレスからメッセージを送信するからである。

【０１１０】

また、各構成アイテムがメッセージを送信するのに用いるプロトコルは、実施形態に応じて任意であってよい。構成アイテムの種別に応じて、異なるプロトコルが使われてもよい。メッセージの送信に使われるプロトコルの例は、ＩＣＭＰ（Internet Control Message Protocol）やＳＮＭＰ（Simple Network Management Protocol）などである。もちろん、他のプロトコルが使われてもよい。

【０１１１】

以上説明した第１実施形態によれば、ある種別の障害の発生が予測されたときに、結果情報が生成され、出力される。出力される結果情報は、予測された障害の発生する蓋然性が高そうな構成アイテムを示す。したがって、結果情報は、「どの構成アイテムに対して対策をとることが有益なのか」を示唆する。つまり、第１実施形態によれば、障害の発生を防ぐための対策をとることが望ましい構成アイテムが、１つ以上検出される。よって、第１実施形態は、障害の発生を未然に防ぐうえで効果的である。

【０１１２】

続いて、図４～７を参照して第２実施形態について説明する。第２実施形態では、ＩＰアドレスが構成アイテムの識別情報として使われる。また、第２実施形態では、障害の発

生もメッセージにより通知される。

【０１１３】

図４は、第２実施形態の検出サーバの動作を例示する図である。図４には「学習フェーズ」と「検出フェーズ」という２つのフェーズの動作が示されている。検出フェーズの動作が第１実施形態の図１の動作に対応する。

【０１１４】

第２実施形態の検出サーバは、学習フェーズにおいて、第１実施形態に関して説明した「第２の頻度」に相当する情報を学習する。その後、検出フェーズでは、ある種別の障害の予兆が検出される。障害の予兆が検出されると、検出サーバは、第１実施形態に関して説明した統計値に相当する値を算出し、算出した統計値に基づいて、第１実施形態に関して説明した結果情報に相当する情報を生成および出力する。

【０１１５】

以下、図４の学習フェーズの詳細について説明する。なお、図４では便宜上、「１７２．１６．１．２」、「１０．０．７．６」、および「１０．０．０．１０」というＩＰアドレスを、それぞれ「Ａ」、「Ｂ」、および「Ｃ」という文字で表してある。

【０１１６】

学習フェーズは、実際の障害の発生を契機として、検出サーバが、障害の発生に先立つ期間においてなされた１回以上の予兆検出の結果に基づく学習を行うフェーズである。例えば、図４には以下の動作シーケンスが例示されている。

【０１１７】

・時刻ｔ１に、ＩＰアドレスＡの構成アイテムから、「１」という種別のメッセージＭ１が出力された。

・時刻ｔ２に、ＩＰアドレスＢの構成アイテムから、「２」という種別のメッセージＭ２が出力された。

・時刻ｔ３に、ＩＰアドレスＣの構成アイテムから、「３」という種別のメッセージＭ３が出力された。

・時刻ｔ４に、ＩＰアドレスＡの構成アイテムから、「４」という種別のメッセージＭ４が出力された。

・時刻ｔ５に、ＩＰアドレスＢの構成アイテムから、「２」という種別のメッセージＭ５が出力された。

・時刻ｔ６に、ＩＰアドレスＡの構成アイテムから、「３」という種別のメッセージＭ６が出力された。

・時刻ｔ７に、ＩＰアドレスＡの構成アイテムから、「１」という種別のメッセージＭ７が出力された。

・時刻ｔ８に、ＩＰアドレスＢの構成アイテムから、「２」という種別のメッセージＭ８が出力された。

・時刻ｔ９に、ＩＰアドレスＢの構成アイテムから、「７」という種別のメッセージＭ９が出力された。

【０１１８】

なお、図４の例では、「７」という種別のメッセージは、「ある特定の種別の障害が発生した」というイベントを通知するためのメッセージである。他方、「１」、「２」、「３」、および「４」という種別のメッセージは、障害の発生以外のイベントを通知するためのメッセージである。以下では説明の簡単化のため、「７」という種別のメッセージにより発生が通知される特定の種別の障害のことを、単に「障害＃７」と表記することがある。また、「障害＃ｆ」などの同様の表記を用いることもある。「７」という種別は、メッセージの種別でもあり、障害の種別でもある。

【０１１９】

さて、第２実施形態では、ウィンドウ３０１を用いて障害予兆が検出される。以下ではウィンドウ３０１の長さを「Ｔ１」と表記することもある。ウィンドウ３０１の長さＴ１は、第１実施形態に関して説明した「第１の所定時間」に対応する。図４に矢印で示すよ

うに、ウィンドウ３０１は時間軸に沿ってスライドしてゆく。

【０１２０】

また、第２実施形態では、個々のメッセージパターンが検出される時点から始まる、ある所定の長さの期間内における障害の発生が予測される。当該期間を以下では「予測対象期間」という。予測対象期間の長さは、第１実施形態に関して説明した「第２の所定時間」に対応し、以下では予測対象期間の長さを「Ｔ２」と表記することもある。

【０１２１】

実際に時刻ｔ９に障害＃７が起きると、検出サーバは、メッセージＭ９を受信する。検出サーバは、メッセージＭ９の受信により、障害＃７の発生を認識し、学習フェーズの処理を開始する。

【０１２２】

具体的には、検出サーバは、まず、時刻ｔ９における障害＃７の予兆として正しく検出された障害予兆（つまり、時刻ｔ９における障害＃７の発生についての正しい予測）を検索する。詳しくは後述するとおり、第２実施形態では、障害予兆が検出されるたびに、検出結果が記憶される。よって、検出サーバは、記憶装置を検索することで、時刻ｔ９の障害の発生に先立つ期間においてなされた１回以上の予兆検出の結果を認識することができる。

【０１２３】

ここで、第２実施形態における障害の発生の予測は、上記のとおり、予測対象期間内の未来に関して行われる。よって、時刻ｔ９における障害＃７の発生についての正しい予測は、もし存在するとすれば、時刻ｔ９を終了時点とする長さＴ２の期間内に存在する。図４では、時刻ｔ９を終了時点とする予測対象期間３０２が、両向き矢印により示されている。

【０１２４】

そこで、検出サーバは、具体的には、時刻ｔ９を終了時点とする予測対象期間３０２内に行われた予測の結果を検索する。図４は、時刻ｔ１、ｔ２、ｔ３、ｔ５、ｔ６、およびｔ８に行われた６回の予測が正しかったことを示す。具体的には、図４は、以下のことを示している。なお、図４においては、正解した予測に関して検出された障害予兆（つまりメッセージパターン）は実線で囲われており、不正解の予測に関して検出された障害予兆は破線で囲われている。

【０１２５】

・時刻ｔ１にメッセージＭ１が出力される。時刻ｔ１を終了時点とするウィンドウ３０１内には、メッセージＭ１のみが含まれる。そのため、検出サーバは、メッセージＭ１のみを含むメッセージパターンから、障害の発生を予測する。こうして時刻ｔ１に行われた予測では、検出サーバは、「長さＴ２の予測対象期間中に障害＃７が発生するだろう」と予測した。この予測が正解であることは、時刻ｔ９に判明する。

・時刻ｔ２にメッセージＭ２が出力される。時刻ｔ２を終了時点とするウィンドウ３０１内には、メッセージＭ１とＭ２が含まれる。そのため、検出サーバは、メッセージＭ１とＭ２を含むメッセージパターンから、障害の発生を予測する。こうして時刻ｔ２に行われた予測では、検出サーバは、「長さＴ２の予測対象期間中に障害＃７が発生するだろう」と予測した。この予測が正解であることは、時刻ｔ９に判明する。

・時刻ｔ３にメッセージＭ３が出力される。時刻ｔ３を終了時点とするウィンドウ３０１内には、メッセージＭ１とＭ２とＭ３が含まれる。そのため、検出サーバは、メッセージＭ１とＭ２とＭ３を含むメッセージパターンから、障害の発生を予測する。こうして時刻ｔ３に行われた予測では、検出サーバは、「長さＴ２の予測対象期間中に障害＃７が発生するだろう」と予測した。この予測が正解であることは、時刻ｔ９に判明する。

・時刻ｔ４にメッセージＭ４が出力される。時刻ｔ４を終了時点とするウィンドウ３０１内には、メッセージＭ３とＭ４が含まれる。そのため、検出サーバは、メッセージＭ３とＭ４を含むメッセージパターンから、障害の発生を予測する。こうして時刻ｔ４に行われた予測では、検出サーバは、「長さＴ２の予測対象期間中には障害が発生しないだろ

う」と予測したか、または、「長さＴ２の予測対象期間中に障害＃ｆ（ただしｆ≠７）が発生するだろう」と予測した。この予測が不正解であることは、時刻ｔ９に判明する。

・時刻ｔ５にメッセージＭ５が出力される。時刻ｔ５を終了時点とするウィンドウ３０１内には、メッセージＭ４とＭ５が含まれる。そのため、検出サーバは、メッセージＭ４とＭ５を含むメッセージパターンから、障害の発生を予測する。こうして時刻ｔ５に行われた予測では、検出サーバは、「長さＴ２の予測対象期間中に障害＃７が発生するだろう」と予測した。この予測が正解であることは、時刻ｔ９に判明する。

・時刻ｔ６にメッセージＭ６が出力される。時刻ｔ６を終了時点とするウィンドウ３０１内には、メッセージＭ４とＭ５とＭ６が含まれる。そのため、検出サーバは、メッセージＭ４とＭ５とＭ６を含むメッセージパターンから、障害の発生を予測する。こうして時刻ｔ６に行われた予測では、検出サーバは、「長さＴ２の予測対象期間中に障害＃７が発生するだろう」と予測した。この予測が正解であることは、時刻ｔ９に判明する。

・時刻ｔ７にメッセージＭ７が出力される。時刻ｔ７を終了時点とするウィンドウ３０１内には、メッセージＭ６とＭ７が含まれる。そのため、検出サーバは、メッセージＭ６とＭ７を含むメッセージパターンから、障害の発生を予測する。こうして時刻ｔ７に行われた予測では、検出サーバは、「長さＴ２の予測対象期間中には障害が発生しないだろう」と予測したか、または、「長さＴ２の予測対象期間中に障害＃ｆ（ただしｆ≠７）が発生するだろう」と予測した。この予測が不正解であることは、時刻ｔ９に判明する。

・時刻ｔ８にメッセージＭ８が出力される。時刻ｔ８を終了時点とするウィンドウ３０１内には、メッセージＭ７とＭ８が含まれる。そのため、検出サーバは、メッセージＭ７とＭ８を含むメッセージパターンから、障害の発生を予測する。こうして時刻ｔ８に行われた予測では、検出サーバは、「長さＴ２の予測対象期間中に障害＃７が発生するだろう」と予測した。この予測が正解であることは、時刻ｔ９に判明する。

【０１２６】

以上のごとき図４の例では、検出サーバは、時刻ｔ９における上記の検索（つまり予測対象期間３０２の範囲内での正解した予測の検索）の結果、以下のことを認識する。

【０１２７】

・予測対象期間３０２中に行われた予測のうち、時刻ｔ９の障害＃７の発生を正しく当てたのは、時刻ｔ１、ｔ２、ｔ３、ｔ５、ｔ６、およびｔ８に行われた６回の予測である。

・これら正解した６回の予測のうち、障害予兆を示すメッセージパターン（つまり予測に使われたウィンドウ３０１に含まれるメッセージのパターン）中に、種別「１」のメッセージが含まれるのは、４回である。

・これら正解した６回の予測のうち、障害予兆を示すメッセージパターン中に、種別「２」のメッセージが含まれるのは、５回である。

・これら正解した６回の予測のうち、障害予兆を示すメッセージパターン中に、種別「３」のメッセージが含まれるのは、２回である。

・これら正解した６回の予測のうち、障害予兆を示すメッセージパターン中に、種別「４」のメッセージが含まれるのは、２回である。

【０１２８】

以下では、障害＃ｆ（つまり「ｆ」という種別のメッセージにより通知される障害）の発生についての正しい予測のうち、「ｎ」という種別のメッセージが「予兆パターン」（predictive pattern）に含まれる相対頻度を「ＷＦ（ｆ，ｎ）」と表記する。なおここで、「予兆パターン」とは、障害の発生の予測に使われたメッセージパターンのことであり、換言すれば、障害の予兆として検出されるメッセージパターンのことである。

【０１２９】

また、第２実施形態では、メッセージパターンは、メッセージの出力される時間的順序とは無関係な組み合わせパターンである。また、第２実施形態では、ウィンドウ３０１内に同じ種類のメッセージが２つ以上含まれる場合、メッセージの重複は無視される。例えば、以下に挙げる４つの場合は、同じ１つのメッセージパターン（以下、便宜上「［１，

２］」と表記することがある）に該当する。
【０１３０】
・「１」という種別のメッセージが先に出力され、その後、「２」という種別のメッセージが出力され、ウィンドウ３０１内にはこれら２つのメッセージのみが含まれる場合。
・「２」という種別のメッセージが先に出力され、その後、「１」という種別のメッセージが出力され、ウィンドウ３０１内にはこれら２つのメッセージのみが含まれる場合。
・「１」という種別のメッセージが先に出力され、次に「２」という種別のメッセージが出力され、その後、「１」という種別のメッセージが出力され、ウィンドウ３０１内にはこれら３つのメッセージのみが含まれる場合。
・「１」という種別のメッセージが先に出力され、次に「２」という種別のメッセージが出力され、その後、「２」という種別のメッセージが出力され、ウィンドウ３０１内にはこれら３つのメッセージのみが含まれる場合。
【０１３１】
上記の４つの場合以外にも、メッセージパターン［１，２］に該当する場合が存在し得ることは明らかである。実施形態によっては、ウィンドウ３０１内に同じ種類のメッセージが含まれる回数に応じた違いが考慮に入れられてもよい。例えば、メッセージパターン［１，２］と［１，１，２］と［１，２，２］が区別される実施形態も可能である。
【０１３２】
なお、図４の例では、時刻ｔ９の学習フェーズにおけるＷＦ（ｆ，ｎ）の値は、以下に示すとおりである。
【０１３３】
ＷＦ（７，１）＝４／６
ＷＦ（７，２）＝５／６
ＷＦ（７，３）＝２／６
ＷＦ（７，４）＝２／６
【０１３４】
なお、ＷＦ（ｆ，ｎ）は、図１に関して説明した「第２の頻度」の具体例である。図１と図４の対応関係をより詳しく説明すれば、以下のとおりである。
図１に関して説明した「発生時点」は、図４では、時刻ｔ９に対応する。よって、図１に関して説明した「発生時点から第２の所定時間以内の過去」は、図４では、時刻ｔ９を終了時点とする予測対象期間３０２に対応する。したがって、図４において予測対象期間３０２に含まれる時刻ｔ１～ｔ８のそれぞれは、図１に関して説明した「出力時点」に相当する。よって、図４において、各時刻ｔｊ（１≦ｊ≦８）を終了時点とするウィンドウ３０１の範囲が、図１に関して説明した各「ウィンドウ期間」に相当する。
【０１３５】
ここで、図１に関して説明した「第２のメッセージパターン」は、「ウィンドウ期間」に含まれる１つ以上のメッセージの組み合わせパターンである。よって、図４においては、各時刻ｔｊ（１≦ｊ≦８）に行われた予測に使われた各メッセージパターンが、「第２のメッセージパターン」に対応する。
【０１３６】
時刻ｔ９よりも後のある時刻（例えば、後述する検出フェーズにおける時刻ｔ１１）において、障害＃７の発生が予測される場合があり得る。具体的には、Ｑ個の構成アイテムから出力されたＰ個のメッセージの組み合わせパターンである「第１のメッセージパターン」に基づいて、障害＃７の発生が予測される場合があり得る（１≦Ｑ≦Ｐ）。この場合に、Ｑ個の構成アイテムのうち、「第１のメッセージパターン」に含まれる「ｎ」という種別のメッセージを出力した構成アイテムについての「統計値」の算出において使われる「第２の頻度」が、ＷＦ（７，ｎ）に対応する。
【０１３７】

図４では、予測対象期間３０２内の最後の「出力時点」である時刻ｔ８の下に、ＷＦ（７，１）とＷＦ（７，２）の上記の値（すなわち４／６と５／６）が例示されている。ＷＦ（７，３）とＷＦ（７，４）の値は、図４では紙幅の都合上、省略されている。

【０１３８】

ところで、第２実施形態におけるＷＦ（ｆ，ｎ）は、上記のように相対頻度である。具体的には、ＷＦ（ｆ，ｎ）は、障害＃ｆの発生についての正しい予測のうちで「ｎ」という種別のメッセージが予兆パターンに含まれる予測の回数を、障害＃ｆの発生についての正しい予測の回数で割った値である。より正確には、ＷＦ（ｆ，ｎ）の分子と分母それぞれの値を数える対象は、実際に障害＃ｆが発生した「発生時点」を終了時点とする予測対象期間３０２の範囲に限られる。

【０１３９】

図４では、理解の助けとするために、予測対象期間３０２の範囲内で時刻ｔ１から順にＷＦ（７，１）の分子と分母を数える場合の、分子と分母それぞれの値も、「ＷＦ（７，１）」の行に例示されている。例えば、時刻ｔ５の下には「３／４」と書いてあるが、これは以下のことを示す。

【０１４０】

・時刻ｔ５における予測は、予測対象期間３０２の中で障害＃７の発生を正しく予測した４番目の予測である（時刻ｔ４の予測が不正解であることに注意）。

・上記４回の正しい予測のうち、予兆パターンが「１」という種別のメッセージを含むのは、３回である（「１」という種別のメッセージは、時刻ｔ１、ｔ２、およびｔ３の予兆パターンには含まれるが、時刻ｔ５の予兆パターンには含まれないことに注意）。

【０１４１】

同様に、図４では、理解の助けとするために、予測対象期間３０２の範囲内で時刻ｔ１から順にＷＦ（７，２）の分子と分母を数える場合の、分子と分母それぞれの値も、「ＷＦ（７，２）」の行に例示されている。

【０１４２】

以上のようにして、第２実施形態の学習フェーズでは、実際の障害の発生を契機として、検出サーバが、障害の発生に先立つ期間においてなされた１回以上の予兆検出の結果に基づく学習を行う。

【０１４３】

なお、時刻ｔ９での障害＃７の発生に先立つ時刻ｔ１、ｔ２、ｔ３、ｔ５、ｔ６、およびｔ８において正しい予測が可能な理由は、時刻ｔ１よりも前の時点で少なくとも１回は既に障害＃７が発生したことがあるからである。つまり、時刻ｔ１より前に障害＃７が発生したときに、障害＃７が発生する直前の予測対象期間中の各ウィンドウのメッセージパターンが、障害＃７と共起するメッセージパターンとして学習される。何回か実際に障害＃７が発生すると、各メッセージパターンと障害＃７の共起頻度が算出可能である。検出サーバは、例えば共起頻度に基づいて、学習した各メッセージパターンを重みづけしてもよい。もちろん、検出サーバは、他の種類の障害についても同様に学習を行う。

【０１４４】

以上のようにして、検出サーバは、学習済みのメッセージパターンに基づいて、時刻ｔ１～ｔ８のそれぞれにおける予測を行う。その結果、図４の例では、たまたま、時刻ｔ１、ｔ２、ｔ３、ｔ５、ｔ６、およびｔ８における６回の予測が当たったわけである。

【０１４５】

以上の説明から分かるように、最初に障害＃７が生じたときには、障害＃７の予兆となるメッセージパターンは、まだ１つも学習されていない。したがって、障害＃７の最初の発生の前には、障害＃７の発生は予測されない。よって、障害＃７の最初の発生の直前の予測対象期間の中で正解した予測の回数も０回である。この場合、ＷＦ（７，ｎ）は、例えば０と定義されてもよい。

【０１４６】

さて、続いて、上記の学習フェーズの学習結果を利用する検出フェーズについて説明す

る。図４の例では、時刻ｔ９の後の時刻ｔ１０において、ＩＰアドレスＢの構成アイテムから、「２」という種別のメッセージＭ１０が出力される。また、時刻ｔ１１において、ＩＰアドレスＡの構成アイテムから、「１」という種別のメッセージＭ１１が出力される。

【０１４７】

なお、時刻ｔ９とｔ１０の間に、さらに１つ以上のメッセージが出力されていてもよい。検出サーバは、メッセージが出力されるたびに、当該メッセージの出力された時点を終了時点とするウィンドウ内のメッセージパターンに基づいて、障害の発生に関する予測を行う。

【０１４８】

例えば、検出サーバは、時刻ｔ１１にメッセージＭ１１を受信すると、時刻ｔ１１を終了時点とするウィンドウ３０３に含まれるメッセージパターン［１，２］（つまりメッセージＭ１０とＭ１１の２つを含むパターン）に基づく予測を行う。図４の例では、時刻ｔ１１における予測において、検出サーバが、「障害＃７が長さＴ２の予測対象期間以内に発生するだろう」と予測したものとする。

【０１４９】

なお、図４の例では、時刻ｔ９以降に、障害＃７の発生が予測されたのは、時刻ｔ１１が初めてであったものとする。つまり、時刻ｔ１０における予測（および、時刻ｔ９とｔ１０の間に１つ以上のメッセージが出力される場合は、各メッセージの出力時点を終了時点とするウィンドウに基づく予測）では、障害＃７の発生は予測されなかったものとする。

【０１５０】

こうして時刻ｔ１１に障害＃７の発生が予測されると、検出サーバは、「予測される障害＃７の発生を未然に防ぐには、コンピュータシステム中のどの構成アイテムに対して対策をとることが有効か」を示唆する情報を、生成および出力する。以下では当該情報のことを「ランキング情報」という。ランキング情報は図１の「結果情報」に対応する。つまり、第２実施形態の検出フェーズの処理は、図１の処理に対応する。

【０１５１】

例えば、図４の例では、時刻ｔ１１における予測が図１のステップＳ１に対応する。この場合、予測にはウィンドウ３０３に含まれる２個のメッセージＭ１０とＭ１１が使われるので、図１における「Ｐ」の値は２である。また、図４の例では、メッセージＭ１０の発信元（sender）たる構成アイテムと、メッセージＭ１１の発信元たる構成アイテムは異なるので、図１における「Ｑ」の値は２である。

【０１５２】

図１のステップＳ２と同様に、第２実施形態でも、Ｑ個の構成アイテムの各々について、予測された障害＃７が当該構成アイテムにおいて将来発生する蓋然性に関する統計値が算出される。第２実施形態では、統計値の具体例として、式（１）により定義されるＷＦ－ＩＤＦ（ｆ，ｎ）が使われる。ＷＦ－ＩＤＦ（ｆ，ｎ）は、障害＃ｆの発生が予測されたときに予測の根拠として使われたメッセージパターン（すなわち予兆パターン）中の、「ｎ」という種別のメッセージを出力した構成アイテムについて算出される統計値である。

$$WF\text{-}IDF(f,n)=WF(f,n)\times\log_{10}(1/DF(n)) \qquad (1)$$

【０１５３】

式（１）中のＷＦ（ｆ，ｎ）は、学習フェーズに関して上述したものである。上述のとおり、ＷＦ（ｆ，ｎ）は、図１に関して説明した「第２の頻度」に対応する。一方、式（１）中のＤＦ（ｎ）は、図１に関して説明した「第１の頻度」の具体例である。つまり、ＤＦ（ｎ）は、「ｎ」という種別のメッセージがどれほど多く出力されるかを示す。

【０１５４】

具体的には、ＤＦ（ｎ）も相対頻度である。ある時刻ｔにおけるＤＦ（ｎ）は、当該ある時刻ｔまでに検出サーバが分析したウィンドウの総数のうち、「ｎ」という種別のメッ

セージを含むウィンドウの数を示す、相対頻度である。

【０１５５】

換言すれば、ある時刻ｔにおけるＤＦ（ｎ）の分母は、当該ある時刻ｔまでに検出サーバが障害予兆の検出のためにメッセージパターンを分析した回数である。そして、当該ある時刻ｔにおけるＤＦ（ｎ）の分子は、分析されたすべてのメッセージパターンのうち、「ｎ」という種別のメッセージを含むものの数である。

【０１５６】

なお、上記のとおり第２実施形態では、メッセージパターンの定義において、ウィンドウ内での同じ種類のメッセージの重複は無視される。よって、上記ある時刻ｔにおけるＤＦ（ｎ）の分子は、分析されたすべてのメッセージパターン中の「ｎ」という種別のメッセージについて重複を無視して数えた数でもある。

【０１５７】

上記のとおり、ウィンドウ内での同じ種類のメッセージの重複が考慮に入れられる実施形態も可能である。その場合、ＤＦ（ｎ）の分子は、ウィンドウ内での同じ種類のメッセージの重複を無視して数えられる値（つまり「ｎ」という種別のメッセージを含むウィンドウの数）であってもよい。あるいは、ＤＦ（ｎ）の分子は、ウィンドウ内での同じ種類のメッセージの重複を考慮して数えられる値（つまり「ｎ」という種別のメッセージの総数）であってもよい。

【０１５８】

図４では、紙幅の都合上、時刻ｔ１１におけるＤＦ（１）の値（すなわち１２００／１２０００）とＤＦ（２）の値（すなわち（６／１２０００）のみが例示されている。図４ではＤＦ（３）やＤＦ（４）などは省略されているが、ＤＦ（ｎ）は、各種別についてカウントされる。

【０１５９】

なお、ＤＦ（１）とＤＦ（２）を比べると、「１」という種別のメッセージよりも「２」という種別のメッセージの方が遥かに珍しいことが分かる。それにもかかわらず、ＷＦ（７，１）とＷＦ（７，２）には大きな差がなく、むしろ、ＷＦ（７，２）の方がＷＦ（７，１）よりも大きいくらいである。つまり、「２」という種別のメッセージは、「他の種別の障害と比べて障害＃７ととりわけよく共起し、障害＃７を特徴づける予兆である」と推定される。式（１）のＷＦ－ＩＤＦ（ｆ，ｎ）は、このような推定を反映した統計値の例である。

【０１６０】

式（１）から明らかなように、式（１）のＷＦ－ＩＤＦ（ｆ，ｎ）は、「第１の頻度」としてのＤＦ（ｎ）に対して単調減少するとともに、「第２の頻度」としてのＷＦ（ｆ，ｎ）に対して単調増加する統計値の一例である。ＷＦ－ＩＤＦ（ｆ，ｎ）は、ＤＦ（ｎ）に対して単調減少するとともにＷＦ（ｆ，ｎ）に対して単調増加するように定義されていれば、式（１）以外の式により定義されていてもよい。

【０１６１】

例えば、式（１）における対数の底は、実施形態に応じて変更されてもよい。また、対数を使わない式により、ＷＦ－ＩＤＦ（ｆ，ｎ）が定義されてもよい。もちろん、適宜の係数の加算または乗算などを含む式が、ＷＦ－ＩＤＦ（ｆ，ｎ）を定義するのに使われてもよい。

【０１６２】

例えば、図４の例では、障害＃７の発生が時刻ｔ１１で予測されたときの予兆パターンは、メッセージＭ１０とＭ１１を含む。そして、メッセージＭ１１の種別は「１」である。よって、検出サーバは、メッセージＭ１１の発信元（つまりＩＰアドレスＡの構成アイテム）についての統計値として、ＷＦ－ＩＤＦ（７，１）を算出する。同様に、検出サーバは、「２」という種別のメッセージＭ１０の発信元（つまりＩＰアドレスＢの構成アイテム）についての統計値として、ＷＦ－ＩＤＦ（７，２）を算出する。

【０１６３】

ところで、情報検索（information retrieval）の分野で使われるＴＦ－ＩＤＦ（term frequency-inverse document frequency）は、ＴＦとＩＤＦの積である。ＴＦだけを用いる場合、特定の文書にのみ頻出する用語と、多くの文書に頻出する一般的な用語との区別が困難であるが、ＩＤＦを利用することで、一般的な用語の影響を少なくすることができる。つまり、ＩＤＦは、一種のノイズフィルタの役割を果たす。よって、ある特定の文書と、当該特定の文書を特徴づける用語（つまり特定の文書にのみ頻出する用語）とのペアに対して算出されるＴＦ－ＩＤＦは、上記特定の文書と、様々な文書に頻出する一般的な用語とのペアに対して算出されるＴＦ－ＩＤＦよりも大きい。

【０１６４】

式（１）における「$\times \log_{10}(1/DF(n))$」という乗算も、一種のノイズフィルタの役割を果たす。例えば、ある構成アイテムが、「ｎ」という種別のメッセージを、恒常的に比較的高い頻度で繰り返し出力する場合があり得る。この場合、どの時刻で予測が行われるにせよ、ウィンドウ内には「ｎ」という種別のメッセージが含まれる蓋然性が高い。そして、恒常的に繰り返し出力されるメッセージは、特定の種別の障害との間でのみ高頻度で共起するわけではないから、特定の種別の障害との関連性は低い。「ｎ」という種別のメッセージが、恒常的に比較的高い頻度で繰り返し出力される場合、特定の種別の障害を予測する上では、この「ｎ」という種別のメッセージを出力する構成アイテムの重要度は、低いと推定される。

【０１６５】

式（１）における「$\times \log_{10}(1/DF(n))$」という乗算は、以上のように恒常的に比較的高い頻度で繰り返し出力されるメッセージの影響を軽減するための、ノイズフィルタの役割を果たす。つまり、式（１）における「$\times \log_{10}(1/DF(n))$」という乗算は、特定の種別の障害の予測において、より重要性の高い構成アイテムを、より適切に見出すために、行われる。換言すれば、上記「第１の頻度」に対して単調減少するように上記「統計値」を定義することで、ノイズの影響が軽減され、そのため、提示される結果情報の精度も高まる。

【０１６６】

「ｎ」という種別のメッセージを含むメッセージパターンから障害＃ｆの発生が予測されたとすると、ＷＦ－ＩＤＦ（ｆ，ｎ）は、以下のことを示す。つまり、ＷＦ－ＩＤＦ（ｆ，ｎ）は、「ｎ」という種別のメッセージを出力した構成アイテムの重要性を示す。より詳しくは、ＷＦ－ＩＤＦ（ｆ，ｎ）は、「『ｎ』という種別のメッセージを出力した構成アイテムからメッセージが出力されることが、障害＃ｆの発生を予測するうえで、どれほど重要なのか」ということを示す。別の観点から述べれば、ＷＦ－ＩＤＦ（ｆ，ｎ）は、「『ｎ』という種別のメッセージを出力した構成アイテムにおいて、当該メッセージの出力の原因となった事象に対して対策を講じることが、障害＃ｆの発生とどれほど強く関連するのか」を示す。

【０１６７】

図４の例では、時刻ｔ１１に、ウィンドウ３０３内の２つのメッセージＭ１０とＭ１１を含むメッセージパターンに基づいて、障害＃７の発生が予測される。こうして障害＃７に関して時刻ｔ１１で検出した予兆パターンに関する情報が、図４には詳細予兆情報３０４として例示されている。詳細予兆情報３０４は、予兆パターン内の各メッセージについて、当該メッセージを出力した発信元の構成アイテムのＩＰアドレスと、当該メッセージの種別とを対応づける情報である。

【０１６８】

図４の例では、「１」という種別のメッセージＭ１１は、ＩＰアドレスＡ（１７２．１６．１．２）の構成アイテムから出力されたので、ＩＰアドレスＡと「１」という種別が対応づけられている。また、「２」という種別のメッセージＭ１０は、ＩＰアドレスＢ（１０．０．７．６）の構成アイテムから出力されたので、ＩＰアドレスＢと「２」という種別が対応づけられている。

【０１６９】

検出サーバは、予兆パターンに含まれる各メッセージの発信元の構成アイテムについて、上記のとおりＷＦ－ＩＤＦ（ｆ，ｎ）を算出する。図４の例では、検出サーバは、メッセージＭ１１の発信元（つまりＩＰアドレスＡの構成アイテム）について、式（２）のようにＷＦ－ＩＤＦ（７，１）を算出する。また、検出サーバは、メッセージＭ１０の発信元（つまりＩＰアドレスＢの構成アイテム）について、式（３）のようにＷＦ－ＩＤＦ（７，２）を算出する。

$$\begin{aligned} WF\text{-}IDF(7,1) &= WF(7,1)\times\log_{10}(1/DF(1)) \\ &= 4/6\times\log_{10}(12000/1200) \\ &= 0.67 \end{aligned} \quad (2)$$

$$\begin{aligned} WF\text{-}IDF(7,2) &= WF(7,2)\times\log_{10}(1/DF(2)) \\ &= 5/6\times\log_{10}(12000/6) \\ &= 2.75 \end{aligned} \quad (3)$$

【０１７０】

第２実施形態では、検出サーバは、算出した各ＷＦ－ＩＤＦ（ｆ，ｎ）の値に基づいて、予兆パターンに含まれるメッセージの発信元の構成アイテムに順序をつける。そして、検出サーバは、順序づけの結果を示すランキング情報３０５を生成する。ランキング情報３０５は、図１のステップＳ３に関して説明した「結果情報」の一例である。

【０１７１】

図４に示すように、ランキング情報３０５は、予兆パターンに含まれるＰ個のメッセージの発信元のＱ個の構成アイテムの各々について、以下の４種類の情報を対応づける情報である（１≦Ｑ≦Ｐ）。

【０１７２】

・当該構成アイテムの順位（つまり、ＷＦ－ＩＤＦ（ｆ，ｎ）によるソートの結果として与えられた順位）。

・当該構成アイテムのＩＰアドレス（つまり、当該構成アイテムを識別する識別情報）。

・予兆パターンに含まれるメッセージのうち、当該構成アイテムが出力したメッセージの種別。

・当該構成アイテムに関して算出されたＷＦ－ＩＤＦ（ｆ，ｎ）。

【０１７３】

なお、１つの構成アイテムから、予兆パターンに含まれる２つ以上のメッセージが出力される場合もあり得る。つまり、図１に関して説明したように、Ｑ＜Ｐの場合があり得る。

【０１７４】

例えば、障害＃ｆの予兆パターンの中には「ｎ１」という種別のメッセージと「ｎ２」という種別のメッセージがともに含まれ、かつ、両メッセージは同じ１つの構成アイテムから出力されたとする。この場合、これら２つのメッセージを出力した当該構成アイテムに関して、ＷＦ－ＩＤＦ（ｆ，ｎ１）とＷＦ－ＩＤＦ（ｆ，ｎ２）の両方を検出サーバが算出する。そして、検出サーバは、ＷＦ－ＩＤＦ（ｆ，ｎ１）とＷＦ－ＩＤＦ（ｆ，ｎ２）のうちの大きい方の値を採用する。こうして採用された値が、Ｑ個の構成アイテムをソートする際のソートキーとして使われる。

【０１７５】

ランキング情報３０５の生成後、検出サーバは、ランキング情報３０５を出力する。ランキング情報３０５の出力は図１のステップＳ４に対応する。ランキング情報３０５は、予兆パターンに含まれるＰ（＝２）個のメッセージ出力したＱ（＝２）個の構成アイテムのうちで、統計値としてのＷＦ－ＩＤＦ（ｆ，ｎ）が最大の構成アイテムを識別する識別情報（すなわち、ＩＰアドレスＢ）を含む。つまり、ランキング情報３０５は、時刻ｔ１１より後の将来に発生しそうだと予測される障害＃７に関して、「障害＃７の予測にとって最も重要度が高い」と推定される構成アイテムを識別する情報として、ＩＰアドレスＢを含む。したがって、例えばシステム管理者などの人間は、出力されたランキング情報３

０５を見ることにより、障害＃７との関連性が高い構成アイテムを認識することができる。システム管理者などは、障害＃７の発生を防ぐための適切な対策を立案することもできる。

【０１７６】

また、ランキング情報３０５は、順位とＩＰアドレスだけでなく、算出されたＷＦ－ＩＤＦ（ｆ，ｎ）も含む。例えば、１位と２位の構成アイテムのＷＦ－ＩＤＦ（ｆ，ｎ）の値に大きな差がない場合などは、システム管理者は、１位と２位の構成アイテムの双方に対して対策を講じることに決めてもよい。

【０１７７】

このように、ランキング情報３０５は、障害＃ｆの発生を防ぐうえで有益な情報である。別の観点から見れば、第２実施形態の検出サーバは、予測された障害の発生を防ぐための作業を行うシステム管理者などを強力に支援するものである。

【０１７８】

なお、ランキング情報３０５情報の出力（およびシステム管理者による対策の実行）にも関わらず、不幸にして、時刻ｔ１１よりも後に、実際に障害＃７が発生してしまう可能性もある。その場合は、障害＃７の発生を契機として、検出サーバは、再び学習フェーズの処理を行う。仮に、障害＃７が、時刻ｔ１１から予測対象期間の長さＴ２以内の未来に実際に発生した場合には、時刻ｔ１１での予測は、再度の学習フェーズにおいて、「正解した予測」として扱われ、新たなＷＦ（７，１）とＷＦ（７，２）の算出において考慮に入れられる。

【０１７９】

続いて、図５～７を参照して、図４を参照して説明した第２実施形態のさらなる詳細について説明する。

図５は、第２実施形態の検出サーバのブロック構成図である。図４の学習フェーズと検出フェーズの処理を行う検出サーバは、具体的には、図５の検出サーバ４００であってもよい。

【０１８０】

検出サーバ４００は、コンピュータシステム内の種々の構成アイテムからメッセージ４２０を入力として受け取り、推定結果情報４３０を出力する。推定結果情報４３０は、具体的には、例えば図４のランキング情報３０５であってもよい。

【０１８１】

検出サーバ４００は、ログ情報記憶部４０１、障害予兆検知部４０２、辞書情報記憶部４０３、および障害予兆情報記憶部４０４を有する。検出サーバ４００はさらに、ログ統計算出部４０５、ログ統計情報記憶部４０６、予兆統計算出部４０７、予兆統計情報記憶部４０８、ランキング生成部４０９、およびランキング情報記憶部４１０を有する。

【０１８２】

ログ情報記憶部４０１にはメッセージ４２０が蓄積される。例えば、図４のメッセージＭ１～Ｍ１１はいずれもログ情報記憶部４０１に蓄積される。ログ情報記憶部４０１の詳細は、図６とともに後述する。

【０１８３】

障害予兆検知部４０２は、検出サーバ４００が１つのメッセージ４２０を受信すると、メッセージ４２０の受信時点を終了時点とするウィンドウ内のメッセージパターンに基づいて、障害が発生しそうかどうかを予測する。障害の発生が障害予兆検知部４０２により予測される場合とは、換言すれば、障害の予兆（具体的には予兆パターン）が障害予兆検知部４０２により検知される場合である。例えば、図４には、時刻ｔ１～ｔ８とｔ１１における予測の実行が例示されている。

【０１８４】

なお、障害予兆検知部４０２は、具体的には、辞書情報記憶部４０３に記憶される辞書情報を利用して予兆を検知する。詳しくは図６とともに後述するとおり、第２実施形態では２種類の辞書情報が使われる。

【０１８５】
また、障害予兆検知部４０２は、障害の予兆を検知すると、検知した結果を障害予兆情報記憶部４０４に記憶する。障害予兆情報記憶部４０４の詳細は図６とともに後述する。
【０１８６】
ところで、図４に関する上記の説明から明らかなように、どのｎに関しても、ＤＦ（ｎ）の値は、検出サーバ４００が１つのメッセージ４２０を受信するたびに変化する。ログ統計算出部４０５は、各ｎについてのＤＦ（ｎ）の値の算出に使うための１種の統計値（具体的には、ＤＦ（ｎ）の分子の値と分母の値）を算出する。
【０１８７】
そして、ログ統計算出部４０５は、算出した値をログ統計情報記憶部４０６に記憶する。ログ統計情報記憶部４０６の詳細については、図６とともに後述する。
【０１８８】
また、検出サーバ４００が受信したメッセージ４２０が、障害が実際に発生したことを知らせる種別のものであった場合、検出サーバ４００は、図４の学習フェーズの処理を行う。
【０１８９】
例えば、図４のメッセージＭ９は、障害＃７の発生を知らせるメッセージ４２０の例である。検出サーバ４００が時刻ｔ９にメッセージＭ９を受信すると、予兆統計算出部４０７は、障害予兆情報記憶部４０４に記憶された情報を参照して、予測対象期間３０２に行われた予測の結果を読み出す。そして、予兆統計算出部４０７は、読み出した情報に基づいて、ＷＦ（ｆ，ｎ）の算出に使うための１種の統計値（すなわちＷＦ（ｆ，ｎ）の分子と分母の値）を算出する。図４の例ではｆ＝７であり、ｎ＝１，２，３，４である。
【０１９０】
予兆統計算出部４０７は、算出結果を予兆統計情報記憶部４０８に記憶する。予兆統計情報記憶部４０８の詳細は、図６とともに後述する。
【０１９１】
さて、例えば図４の時刻ｔ１１に例示されるように、障害予兆検知部４０２が障害の発生を予測すると、ランキング生成部４０９は推定結果情報４３０を生成する。上記のとおり、推定結果情報４３０は、例えばランキング情報３０５のような情報である。具体的には、ランキング生成部４０９は、ログ統計情報記憶部４０６と予兆統計情報記憶部４０８を参照してＷＦ－ＩＤＦ（ｆ，ｎ）を算出し、算出したＷＦ－ＩＤＦ（ｆ，ｎ）に基づいて推定結果情報４３０を生成する。
【０１９２】
そして、ランキング生成部４０９は、生成した推定結果情報４３０を出力する。例えば、ランキング生成部４０９は、推定結果情報４３０をランキング情報記憶部４１０に蓄積してもよい。実施形態によっては、ランキング情報記憶部４１０が省略されてもよい。また、ランキング生成部４０９は、推定結果情報４３０をディスプレイに出力してもよい。ランキング生成部４０９は、推定結果情報４３０を含む電子メールまたはインスタントメッセージを、システム管理者に宛てて送信（すなわち出力）してもよい。
【０１９３】
ところで、図５の検出サーバ４００は、具体的には図２のコンピュータ１００であってもよい。検出サーバ４００がコンピュータ１００により実現される場合、図２と図５は以下のように対応する。
【０１９４】
検出サーバ４００は、通信インタフェイス１０３を介してメッセージ４２０を受信する。また、検出サーバ４００は、推定結果情報４３０を出力装置１０５に出力してもよく、記憶装置１０６に出力してもよく、駆動装置１０７を介して記憶媒体１１０に出力してもよい。もちろん、検出サーバ４００は、通信インタフェイス１０３とネットワーク１２０を介して推定結果情報４３０を送信してもよい。
【０１９５】

ログ情報記憶部４０１、辞書情報記憶部４０３、障害予兆情報記憶部４０４、ログ統計情報記憶部４０６、予兆統計情報記憶部４０８、およびランキング情報記憶部４１０は、記憶装置１０６により実現されてもよい。障害予兆検知部４０２、ログ統計算出部４０５、予兆統計算出部４０７、およびランキング生成部４０９は、プログラムを実行するＣＰＵ１０１により実現されてもよい。

【０１９６】

また、図５の検出サーバ４００は、図３のコンピュータ２００であってもよい。この場合、メッセージ４２０は、コンピュータシステム２３０内の種々の構成アイテムから出力されて、ネットワーク２１０を介して、検出サーバ４００としてのコンピュータ２００に受信される。また、コンピュータシステム２３０のシステム管理者は、検出サーバ４００から出力される推定結果情報４３０を参照して、コンピュータシステム２３０内のどの構成アイテムに対して対策をとるかを決め、適宜の対策を実行する。

【０１９７】

続いて、図５中の種々の記憶部に記憶される情報の具体例について、図６を参照して説明する。図６は、第２実施形態で利用される各種テーブルの例を示す図である。

ログテーブル５０１は、ログ情報記憶部４０１に記憶される情報の一例である。ログテーブル５０１の各エントリは、検出サーバ４００が受信した各メッセージ４２０に対応する。ログテーブル５０１の各エントリは、例えば以下の４つのフィールドを含んでもよい。

【０１９８】

・検出サーバ４００がメッセージ４２０を受信した時刻。
・メッセージ４２０を出力した構成アイテムを識別するＩＰアドレス。
・メッセージ４２０に含まれる文字列。
・メッセージ４２０の種別。

【０１９９】

例えば、ログテーブル５０１の１番目のエントリは、２０１２年７月３１日２３時４２分ちょうどに、ＩＰアドレスＢ（１０．０．７．６）により識別される構成アイテムから検出サーバ４００が受信したメッセージ４２０に対応する。当該メッセージは、「Permission Denied」という文字列を含み、この文字列に対応する種別は「２」という種別である。検出サーバ４００は、メッセージ４２０を受信するたびに、受信したメッセージ４２０に対応する新しいエントリをログテーブル５０１に追加する。

【０２００】

詳しくは図７のステップＳ１０４に関して後述するが、ログテーブル５０１のメッセージ種別は省略されてもよい。逆に、ログテーブル５０１がメッセージ種別を含む場合、メッセージ種別は以下のようにして記録されてもよい。

【０２０１】

検出サーバ４００は、メッセージ４２０を受信すると、以下に説明するメッセージ辞書テーブル５０２を参照する。そして、検出サーバ４００は、メッセージ辞書テーブル５０２とメッセージ４２０に含まれる文字列とに基づいて、メッセージ４２０の種別を判断し、判断結果をログテーブル５０１にメッセージ種別として記録する。

【０２０２】

メッセージ辞書テーブル５０２は、辞書情報記憶部４０３に記憶される情報の一例である。メッセージ辞書テーブル５０２の各エントリは、メッセージの１つの種別に対応する。上記のとおり、いくつかの種別のメッセージは、それぞれ障害の発生を示し、他の種別のメッセージは、それぞれ障害の発生以外のイベントを示す。メッセージ辞書テーブル５０２の各エントリは、例えば以下の２つのフィールドを含んでもよい。

【０２０３】

・メッセージ種別。
・当該メッセージ種別に分類されるメッセージに含まれる文字列。

【０２０４】

例えば、メッセージ辞書テーブル５０２の２番目のエントリは、「Permission denied」という文字列を含むメッセージ４２０が、「２」という種別に分類されることを示す。そのため、ログテーブル５０１の１番目のエントリのメッセージ種別は、上記のとおり「２」と記録されている。

【０２０５】

なお、個々のメッセージ４２０に含まれる実際の文字列は、種別によって予め決められた固定の文字列と、環境等に応じて可変の文字列とを含む文字列であってもよい。この場合、メッセージ辞書テーブル５０２のメッセージ文字列と、受信されたメッセージ４２０に含まれる文字列との完全一致ではなく、部分一致に基づいて、メッセージ辞書テーブル５０２を用いたメッセージの種別の判断が行われてもよい。

【０２０６】

なお、メッセージ辞書テーブル５０２は、予め用意された静的なテーブルであってもよいし、動的に学習されてもよい。メッセージ辞書テーブル５０２の学習は、例えば、公知の方法にしたがって行われてもよい。

【０２０７】

さて、パターン辞書テーブル５０３も、辞書情報記憶部４０３に記憶される情報の一例である。パターン辞書テーブル５０３の各エントリは、例えば以下の３つのフィールドを含んでもよい。

【０２０８】

・障害の種別（図６の例では、具体的には、当該種別の障害の発生を通知するメッセージの種別により表される）。

・当該種別の障害の予兆パターン（つまり、当該種別の障害の予兆となるメッセージパターンであり、図６の例では、具体的には、当該メッセージパターンに含まれるメッセージの種別のリストにより表される）。

・当該予兆パターンから、どの程度の蓋然性で、当該種別の障害の発生が予測されるのかを示すスコア。

【０２０９】

なお、実施形態によってはスコアは省略されてもよい。検出サーバ４００は、例えば公知の方法にしたがって、パターン辞書テーブル５０３を動的に学習してもよい。スコアは、例えば、学習の過程で観察された、実際の障害とメッセージパターンとの共起頻度に基づく値であってもよい。

【０２１０】

例えば、図４の時刻ｔ１１で障害予兆検知部４０２は、ウィンドウ３０３内には２つのメッセージＭ１０とＭ１１が含まれることを認識する。また、ログテーブル５０１がメッセージ種別を含む場合は、障害予兆検知部４０２は、ログテーブル５０１から、メッセージＭ１０とＭ１１それぞれの種別を認識してもよい。あるいは、障害予兆検知部４０２は、ログテーブル５０１のメッセージ文字列とメッセージ辞書テーブル５０２に基づいて、メッセージＭ１０とＭ１１それぞれの種別を認識してもよい。

【０２１１】

いずれにしろ、障害予兆検知部４０２は、メッセージＭ１０とＭ１１それぞれの種別が「２」と「１」であることを認識する。つまり、障害予兆検知部４０２は、ウィンドウ３０３に対応するメッセージパターン［１，２］を認識する。

【０２１２】

よって、障害予兆検知部４０２は、パターン辞書テーブル５０３内でメッセージパターン［１，２］を検索する。その結果、図６の例では、パターン辞書テーブル５０３の１番目のエントリが見つかる。

【０２１３】

したがって、障害予兆検知部４０２は、メッセージパターン［１，２］から予測される障害の種別が「７」である、ということを認識する。以上のようにして、障害予兆検知部４０２は、時刻ｔ１１において、障害＃７の予兆として、メッセージパターン［１，２］

を検出する。なお、障害予兆検知部４０２は、ウィンドウに対応するメッセージパターンを障害の予兆として検出するか否かを、スコアの値と閾値に基づいて、決めてもよい。

【０２１４】

また、障害予兆検知部４０２は、１つのメッセージパターンから２つ以上の種別の障害の発生を予測してもよい。つまり、パターン辞書テーブル５０３において、異なる障害種別に対応する２つ以上のエントリの予兆パターンが、たまたま同じメッセージパターンである場合もあり得る。

【０２１５】

さて、障害予兆テーブル５０４は、障害予兆情報記憶部４０４に記憶される情報の一例である。障害予兆検知部４０２は、１つの予兆パターンを検出するたびに、新規エントリを障害予兆テーブル５０４に追加する。障害予兆テーブル５０４の各エントリは、例えば以下の５つのフィールドを含んでもよい。

【０２１６】

・障害予兆テーブル５０４内で個々のエントリを識別するＩＤ（identification）。

・障害予兆検知部４０２が発生を予測した障害の種別。

・当該種別の障害について障害予兆検知部４０２が検知した予兆パターン（つまり、当該種別の障害の予測の根拠として障害予兆検知部４０２が使ったメッセージパターン）。

・障害予兆検知部４０２が予測を実行した時刻。

・当該種別の障害がいつから始まりそうか（つまり当該種別の障害がいつ発生しそうか）ということを障害予兆検知部４０２が予測する場合は、その予測された開始時刻。

【０２１７】

なお、実施形態によっては開始時刻が省略されてもよい。逆に、予測された当該種別の障害がいつまでに発生しそうかということを障害予兆検知部４０２が予測する場合は、その予測された時刻を示す終了時刻フィールドがさらにあってもよい。いつからいつまでの期間に障害が発生しそうかを障害予兆検知部４０２が予測する場合は、開始時刻と終了時刻の両方のフィールドがあってもよい。

【０２１８】

ログ統計テーブル５０５は、ログ統計情報記憶部４０６に記憶される情報の一例である。ログ統計テーブル５０５には、図４に関して説明したＤＦ（ｎ）を算出するための情報が記憶される。具体的には、ログ統計テーブル５０５の各エントリは、以下の３つのフィールドを含む。

【０２１９】

・当該エントリを識別するＩＤ。

・メッセージ種別。

・カウント。

【０２２０】

任意のメッセージ種別「ｎ」について、メッセージ種別が「ｎ」のエントリのカウントは、ＤＦ（ｎ）の分子を示す。また、第２実施形態では、どのｎについても、ＤＦ（ｎ）の分母は共通の値（すなわち、障害予兆検知部４０２によって今までに分析されたウィンドウの総数）である。この共通の値が、メッセージ種別として便宜上「＊」と書かれたエントリにおいて、カウントとして記録される。

【０２２１】

図６には、図４の時刻ｔ１１におけるログ統計テーブル５０５の５つのエントリが例示されている。なお、ログ統計テーブル５０５は、「１」～「４」以外のメッセージ種別に対応する他のエントリをさらに含み得るが、図６ではそれらのエントリは省略されている。

【０２２２】

予兆統計テーブル５０６は、予兆統計情報記憶部４０８に記憶される情報の一例である。予兆統計テーブル５０６には、図４に関して説明したＷＦ（ｆ，ｎ）を算出するための

情報が記憶される。具体的には、予兆統計テーブル５０６の各エントリは、以下の４つのフィールドを含む。

【０２２３】

・当該エントリを識別するＩＤ。
・障害種別。
・メッセージ種別。
・カウント。

【０２２４】

任意のｆとｎの組み合わせについて、障害種別が「ｆ」でメッセージ種別が「ｎ」のエントリのカウントは、ＷＦ（ｆ，ｎ）の分子を示す。また、第２実施形態では、ある１つの「ｆ」という障害種別に関しては、どのｎについても、ＷＦ（ｆ，ｎ）の分母は共通の値（すなわち、障害が発生した時点を終了時点とする予測対象期間内で行われた予測のうち、正解だった予測の回数）である。この共通の値が、メッセージ種別として便宜上「＊」と書かれたエントリにおいて、カウントとして記録される。

【０２２５】

図６には、図４の時刻ｔ１１における予兆統計テーブル５０６の５つのエントリが例示されている。換言すれば、図６には、図４の時刻ｔ９での障害＃７の発生を契機に学習された内容が例示されている。なお、予兆統計テーブル５０６は、「７」以外の障害種別に対応する他のエントリをさらに含み得るが、図６ではそれらのエントリは省略されている。

【０２２６】

ランキングテーブル５０７は、図４の検出フェーズで生成される。ランキングテーブル５０７は、下記の「予兆ＩＤ」以外は、図４のランキング情報３０５と同様である。つまり、ランキングテーブル５０７の各エントリは、障害予兆検知部４０２により検知された予兆パターン中の、いずれか１つ以上のメッセージの発信元たる構成アイテムに対応する。また、ランキングテーブル５０７の各エントリは、以下の５つのフィールドを含む。

【０２２７】

・ランキングテーブル５０７の生成の契機となった予測を識別するＩＤ（以下「予兆ＩＤ」ともいう）。つまり、ランキングテーブル５０７の生成の契機となった予測の結果を障害予兆検知部４０２が障害予兆テーブル５０４に記録するときに使われたＩＤ。
・順位。
・ＩＰアドレス。
・メッセージ種別。
・スコア（具体的にはＷＦ－ＩＤＦ（ｆ，ｎ））。

【０２２８】

なお、予兆ＩＤは、複数回の予測にそれぞれ対応するランキング情報同士を、ランキング情報記憶部４１０内で区別するための識別情報である。よって、ランキングテーブル５０７が推定結果情報４３０として出力される際には、予兆ＩＤは省略されてもよい。

【０２２９】

また、予兆パターン中の２つ以上のメッセージを出力した構成アイテムに対応するエントリでは、メッセージ種別のフィールドには、それら２つ以上のメッセージの種別のリストが記憶される。

【０２３０】

ランキングテーブル５０７は、推定結果情報４３０として、例えば、出力装置１０５に出力されてもよいし、検出サーバ４００の外部の他の装置に出力されてもよい。また、ランキングテーブル５０７の各エントリは、ランキング情報記憶部４１０に記憶されてもよい。

【０２３１】

続いて、図７のフローチャートを参照して、検出サーバ４００が行う処理について説明する。なお、検出サーバ４００が行う種々の処理のうち、ログ情報記憶部４０１へのメッ

セージ４２０の蓄積と、パターン辞書テーブル５０３の学習と、障害予兆検知部４０２による障害予兆の検知は、公知の処理と同様であってよい。よって、図７ではこれらの処理は省略されている。図７には、具体的には、ログ統計算出部４０５と予兆統計算出部４０７とランキング生成部４０９により行われる処理が示されている。

【０２３２】

ステップＳ１０１で検出サーバ４００は、何らかのイベントの発生を待つ。そして、「障害発生通知以外のメッセージ４２０が受信された」というイベントが発生すると、ログ統計算出部４０５がステップＳ１０２の処理を行う。他方、「障害発生通知であるメッセージ４２０が受信された」というイベントが発生すると、予兆統計算出部４０７がステップＳ１０３の処理を行う。また、「障害予兆検知部４０２により障害予兆が検知された」というイベントが発生すると、ランキング生成部４０９がステップＳ１０４～Ｓ１１３の処理を行う。

【０２３３】

例えば、図４の時刻ｔ１～ｔ８、ｔ１０、およびｔ１１のいずれにおいても、ステップＳ１０２の処理が実行される。また、図４の時刻ｔ９では、ステップＳ１０３の処理が実行される。そして、図４の時刻ｔ１～ｔ８やｔ１１のように、何らかの種別の障害の発生が障害予兆検知部４０２により予測された場合には、ステップＳ１０４～Ｓ１１３の処理が実行される。

【０２３４】

さて、ステップＳ１０２でログ統計算出部４０５は、ログ統計情報を更新する。具体的には、ログ統計算出部４０５は、ログ統計情報記憶部４０６内のログ統計テーブル５０５中の２つ以上のエントリを更新する。

【０２３５】

まず、ログ統計算出部４０５は、ステップＳ１０１でメッセージ４２０が受信された時点を終了時点とする長さＴ１のウィンドウに含まれるメッセージを、ログテーブル５０１から検索する。検索の結果、ステップＳ１０１で受信されたメッセージ４２０を少なくとも含む、１つ以上のメッセージが見つかる。例えば、図４の時刻ｔ３でのメッセージＭ３の受信を契機にステップＳ１０２の処理が実行される場合、メッセージＭ１～Ｍ３が見つかる。

【０２３６】

ログ統計算出部４０５は、見つかった各メッセージについて、ログ統計テーブル５０５において当該メッセージの種別に対応するエントリのカウントを１だけインクリメントする。さらに、ログ統計算出部４０５は、ログ統計テーブル５０５において「＊」というメッセージ種別のエントリのカウントも、１だけインクリメントする。ステップＳ１０２の処理が完了すると、検出サーバ４００は、再びステップＳ１０１でイベントの発生を待つ。

【０２３７】

例えば、図４の時刻ｔ１１でメッセージＭ１１が受信された場合のステップＳ１０２の動作は以下のとおりである。時刻ｔ１１を終了時点とするウィンドウ３０３には、２つのメッセージＭ１０とＭ１１が含まれ、それぞれの種別は「２」と「１」である。よって、この場合、ステップＳ１０２でログ統計算出部４０５は、ログ統計テーブル５０５においてメッセージ種別が「２」と「１」と「＊」の３つのエントリそれぞれのカウントを１だけインクリメントする。

【０２３８】

さて、ステップＳ１０３で予兆統計算出部４０７は、予兆統計情報を更新する。具体的には、予兆統計算出部４０７は、予兆統計情報記憶部４０８内の予兆統計テーブル５０６中の特定のいくつかのエントリを次のようにして更新する。

【０２３９】

予兆統計算出部４０７は、ステップＳ１０１で受信されたメッセージ４２０によって通知された障害の種別を検索キーとして用いて、予兆統計テーブル５０６を検索する。検索

の結果見つかった全エントリが、ステップＳ１０３での更新対象のエントリである。
【０２４０】
例えば、図４の時刻ｔ９にステップＳ１０３が実行される場合、障害種別が「７」の全エントリが見つかる。予兆統計算出部４０７は、予兆統計テーブル５０６の中から見つけた各エントリのカウントを、０に初期化する。
【０２４１】
また、予兆統計算出部４０７は、ステップＳ１０１で受信されたメッセージ４２０によって通知された障害の発生に先立つ長さＴ２の予測対象期間に行われた予測結果を、障害予兆情報記憶部４０４から検索する。
【０２４２】
例えば、図４の時刻ｔ９にステップＳ１０３が実行される場合、予兆統計算出部４０７が障害予兆情報記憶部４０４を検索すると、時刻ｔ１～ｔ８の各々で行われた８回の予測の結果が見つかる。つまり、検索の結果、障害予兆テーブル５０４の８つのエントリが見つかる。
【０２４３】
予兆統計算出部４０７は、障害予兆テーブル５０４の中から見つけた各エントリについて、当該エントリの障害種別が、ステップＳ１０１で受信されたメッセージ４２０によって通知された障害の種別と同じか否かを判断する。
【０２４４】
これら２つの種別が互いに異なる場合、予兆統計算出部４０７は、障害予兆テーブル５０４中の当該エントリを無視する。なぜなら、障害予兆テーブル５０４中の当該エントリは外れた予測を表しているからである。
【０２４５】
逆に、２つの種別が等しい場合、予兆統計算出部４０７は、障害予兆テーブル５０４中の当該エントリに記録されている予兆パターン（すなわち、正解と判明した予兆パターン）を参照する。そして、予兆統計算出部４０７は、当該予兆パターンに含まれる各メッセージ種別について、以下の処理を行う。
【０２４６】
・予兆統計テーブル５０６において、ステップＳ１０１で受信されたメッセージ４２０によって通知された障害の種別と、上記予兆パターンに含まれる当該メッセージ種別のペアに対応づけられたカウントを、１だけインクリメントする処理。
・予兆統計テーブル５０６において、ステップＳ１０１で受信されたメッセージ４２０によって通知された障害の種別と、「＊」という種別のペアに対応づけられたカウントを、１だけインクリメントする処理。
【０２４７】
例えば、図４の時刻ｔ９にステップＳ１０３が実行される場合、予兆統計算出部４０７は、障害予兆テーブル５０４から見つかった８個のエントリのうち、時刻ｔ４とｔ７の予測に対応する２個のエントリを無視する。他方、予兆統計算出部４０７は、残りの６個のエントリの各々の予兆パターンに含まれる各メッセージ種別に関して、上記の処理を行う。その結果、予兆統計テーブル５０６におけるＩＤが「１」～「５」の５つのエントリそれぞれのカウント値は、図６に示す値に更新される。
【０２４８】
以上のようにして、ステップＳ１０３では、図４の学習フェーズの処理が行われ、予兆統計テーブル５０６に学習結果が反映される。ステップＳ１０３の処理が完了すると、検出サーバ４００は、再びステップＳ１０１でイベントの発生を待つ。
【０２４９】
さて、ステップＳ１０４～Ｓ１１３の処理は、障害予兆検知部４０２によって障害の発生が予測されたとき（すなわち障害予兆が検知されたとき）に、ランキング生成部４０９により実行される。ステップＳ１０４～Ｓ１１３の処理は、図１のステップＳ２～Ｓ４に対応し、図４の検出フェーズに対応する。

【０２５０】
ステップＳ１０４でランキング生成部４０９は、障害予兆検知部４０２が障害の予測に用いたウィンドウに含まれる全メッセージの情報を取得し、ランキング情報（具体的には、ランキングテーブル５０７）を空に初期化する。
【０２５１】
例えば、障害予兆検知部４０２は、長さＴ２の予測対象期間の範囲内の未来において障害が発生しそうだ、と予測すると、予測に用いたウィンドウの開始時点と終了時点を、予測結果とともにランキング生成部４０９に通知してもよい。すると、ランキング生成部４０９は、ログテーブル５０１から、上記ウィンドウに含まれる全メッセージのエントリを取得することができる。なお、ランキング生成部４０９は、ログテーブル５０１中のフィールドのうち、少なくともＩＰアドレスとメッセージ種別さえ取得すれば十分である。
【０２５２】
実施形態によっては、障害予兆検知部４０２がランキング生成部４０９に、上記ウィンドウに含まれる各メッセージの発信元のＩＰアドレスと、各メッセージの種別を、予測結果とともに通知してもよい。この場合、ランキング生成部４０９は、ログテーブル５０１を参照しなくても、ウィンドウに含まれる全メッセージについてのＩＰアドレスとメッセージ種別を取得することができる。また、この場合、ログテーブル５０１のメッセージ種別は省略可能である。
【０２５３】
例えば、図４の時刻ｔ１１で障害予兆検知部４０２が障害＃７の発生を予測したとする。この場合、ランキング生成部４０９は、ステップＳ１０４で、ログテーブル５０１または障害予兆検知部４０２から、ウィンドウ３０３に含まれる全メッセージに関して、少なくともメッセージ種別と発信元のＩＰアドレスを取得する。つまり、ステップＳ１０４でランキング生成部４０９は、少なくとも、図４に詳細予兆情報３０４として例示されている情報を取得する。
【０２５４】
また、上記のとおりステップＳ１０４でランキング生成部４０９は、ランキングテーブル５０７を初期化する。
【０２５５】
次に、ステップＳ１０５でランキング生成部４０９は、ステップＳ１０４で情報を取得したメッセージの中に未処理のメッセージがあるか否かを判断する。未処理のメッセージが残っていれば、ランキング生成部４０９は、次にステップＳ１０６の処理を実行する。逆に、ステップＳ１０４で情報を取得した全メッセージについての処理が完了していれば、ランキング生成部４０９は、次にステップＳ１１３の処理を実行する。
【０２５６】
ステップＳ１０６でランキング生成部４０９は、未処理のメッセージを１つ選択する。例えば、ランキング生成部４０９は、ステップＳ１０４で図４のメッセージＭ１０とＭ１１についての情報を取得した場合、メッセージＭ１０とＭ１１のうちの１つを選択する。以下、ステップＳ１０６で選択されたメッセージを「選択メッセージ」という。
【０２５７】
次に、ステップＳ１０７でランキング生成部４０９は、選択メッセージの種別に関するログ統計情報と予兆統計情報を取得する。説明の便宜上、選択メッセージの種別が「ｎ」であるものとし、障害予兆検知部４０２により障害＃ｆが予測されたものとする。この場合、ステップＳ１０７でランキング生成部４０９は、具体的には以下の４つの値を取得する。
【０２５８】
ランキング生成部４０９は、ログ統計テーブル５０５においてメッセージ種別の値が「ｎ」のエントリを参照し、カウントの値を読み取る。こうして読み取られた値は、ＤＦ（ｎ）の分子に相当する。
【０２５９】

さらに、ランキング生成部４０９は、ログ統計テーブル５０５においてメッセージ種別の値が「＊」のエントリを参照し、カウントの値を読み取る。こうして読み取られた値は、ＤＦ（ｎ）の分母に相当する。

【０２６０】

また、ランキング生成部４０９は、予兆統計テーブル５０６において障害種別の値が「ｆ」かつメッセージ種別の値が「ｎ」のエントリを参照し、カウントの値を読み取る。こうして読み取られた値は、ＷＦ（ｆ，ｎ）の分子に相当する。

【０２６１】

そして、ランキング生成部４０９は、予兆統計テーブル５０６において障害種別の値が「ｆ」かつメッセージ種別の値が「＊」のエントリを参照し、カウントの値を読み取る。こうして読み取られた値は、ＷＦ（ｆ，ｎ）の分母に相当する。

【０２６２】

例えば、選択メッセージが図４のメッセージＭ１０である場合、ステップＳ１０７では、図４に例示されたＤＦ（２）の分子と分母（すなわち６と１２０００）と、図４に例示されたＷＦ（７，２）の分子と分母（すなわち５と６）が取得される。

【０２６３】

続いて、ステップＳ１０８でランキング生成部４０９は、ステップＳ１０７で取得した４つの値を用いて、式（１）にしたがって、ＷＦ－ＩＤＦ（ｆ，ｎ）の値を算出する。例えば、選択メッセージが図４のメッセージＭ１０である場合は、式（３）に示すように、約２．７５という値が算出される。他方、選択メッセージが図４のメッセージＭ１１である場合は、式（２）に示すように、約０．６７という値が算出される。

【０２６４】

次に、ステップＳ１０９でランキング生成部４０９は、選択メッセージの発信元のＩＰアドレスが既にランキングテーブル５０７に含まれているか否かを判断する。

【０２６５】

例えば、選択メッセージが図４のメッセージＭ１０である場合、ランキング生成部４０９は、メッセージＭ１０の発信元の構成アイテムを識別するＩＰアドレスＢ（１０．０．７．６）を検索キーとして用いて、ランキングテーブル５０７を検索する。検索の結果、エントリが見つかれば、ランキング生成部４０９は、「選択メッセージの発信元のＩＰアドレスが既にランキングテーブル５０７に含まれている」と判断する。逆に、エントリが見つからなければ、ランキング生成部４０９は、「選択メッセージの発信元のＩＰアドレスはランキングテーブル５０７に含まれていない」と判断する。

【０２６６】

選択メッセージの発信元のＩＰアドレスがランキングテーブル５０７に含まれていない場合、ランキング生成部４０９は、次にステップＳ１１０の処理を行う。逆に、選択メッセージの発信元のＩＰアドレスが既にランキングテーブル５０７に含まれている場合、ランキング生成部４０９は、次にステップＳ１１１の処理を行う。

【０２６７】

ステップＳ１１０でランキング生成部４０９は、ランキングテーブル５０７に、以下の４つの値を含む新規エントリを追加する。

【０２６８】

・ステップＳ１０１で障害予兆検知部４０２から通知された予測結果に関するＩＤ（つまり予兆ＩＤ）。

・選択メッセージの発信元のＩＰアドレス。

・選択メッセージの種別。

・選択メッセージのスコアとしてステップＳ１０８で算出されたＷＦ－ＩＤＦ値。

【０２６９】

例えば、障害予兆検知部４０２があるメッセージパターンからある障害の発生を予測し、その予測結果を障害予兆テーブル５０４に「ｐ」というＩＤとともに記憶したとする。この場合、ステップＳ１０１では、予測結果とともに障害予兆検知部４０２から「ｐ」と

いうＩＤがランキング生成部４０９に通知される。以上のように通知された「ｐ」というＩＤが、ステップＳ１１０における予兆ＩＤである。

【０２７０】

なお、ステップＳ１１０で追加される新規エントリにおいて、順位のフィールドは空でよい。エントリの追加後、ランキング生成部４０９は、再びステップＳ１０５の判断を行う。

【０２７１】

他方、ステップＳ１１１は、１つの構成アイテムから出力された２つ以上のメッセージがウィンドウ内に含まれる場合に、それら２つ以上のメッセージのうち、２番目以降にステップＳ１０６で選択されたメッセージに関して、実行される。

【０２７２】

具体的には、ステップＳ１１１でランキング生成部４０９は、ステップＳ１０９でのランキングテーブル５０７の検索の結果見つかったエントリにおけるメッセージ種別フィールドのリストに、選択メッセージの種別を追加する。また、ステップＳ１１１でランキング生成部４０９は、ランキングテーブル５０７中のスコアが、ステップＳ１０８で算出したＷＦ－ＩＤＦ（ｆ，ｎ）以上か否かを判断する。なおここで、「ランキングテーブル５０７中のスコア」とは、具体的には、ステップＳ１０９でのランキングテーブル５０７の検索の結果見つかったエントリ内のスコアのことである。

【０２７３】

ランキングテーブル５０７中のスコアが、算出したＷＦ－ＩＤＦ（ｆ，ｎ）以上である場合、上記エントリのスコアを更新する必要はない。よって、この場合、ランキング生成部４０９は、次にステップＳ１０５の判断を行う。

【０２７４】

逆に、ランキングテーブル５０７中のスコアが、算出したＷＦ－ＩＤＦ（ｆ，ｎ）未満の場合、ランキング生成部４０９は、次に、ステップＳ１１２でランキングテーブル５０７のスコアを更新する。具体的には、ランキング生成部４０９は、ステップＳ１０９でのランキングテーブル５０７の検索の結果見つかったエントリ中のスコアを、ステップＳ１０８で算出したＷＦ－ＩＤＦ（ｆ，ｎ）の値に置き換える。

【０２７５】

以上のようなステップＳ１１２でのスコアの更新の後、ランキング生成部４０９は、ステップＳ１０５の判断を再び行う。

【０２７６】

例えば、障害＃ｆの予兆パターンの中には「ｎ１」という種別のメッセージと「ｎ２」という種別のメッセージがともに含まれ、かつ、両メッセージが同じ１つの構成アイテムから出力された、という場合があり得る。以上のステップＳ１０９～Ｓ１１２によれば、このような場合に、ＷＦ－ＩＤＦ（ｆ，ｎ１）とＷＦ－ＩＤＦ（ｆ，ｎ２）のうちの大きい方の値がスコアとして採用される。

【０２７７】

例えば、「ｎ１」という種別のメッセージは、障害＃ｆとの共起頻度が他の種別の障害との共起頻度と比べて低いか、または、どの種類の障害との共起頻度も比較的高いものとする。つまり、ＷＦ（ｆ，ｎ１）が小さいか、または、ＤＦ（ｎ１）が大きいものとする。他方、「ｎ２」という種別のメッセージは、障害＃ｆとの共起頻度が比較的高く、かつ、他の種類の障害との共起頻度は比較的低いものとする。つまり、ＷＦ（ｆ，ｎ２）が大きく、かつ、ｆ≠ｇなるｇについてＷＦ（ｇ，ｎ２）が小さい（別の観点から換言すればＤＦ（ｎ２）が比較的小さい）ものとする。

【０２７８】

この場合、ＷＦ－ＩＤＦ（ｆ，ｎ１）よりもＷＦ－ＩＤＦ（ｆ，ｎ２）の方が大きい。また、この場合、「ｎ１」という種別のメッセージと障害＃ｆとの関連性よりも、「ｎ２」という種別のメッセージと障害＃ｆとの関連性の方が高い。つまり、「ｎ２」という種別のメッセージは、「ｎ１」という種別のメッセージよりも、一層よく障害＃ｆを特徴づ

けている。よって、障害＃ｆの予測にとっての重要性がより高い構成アイテムは、「ｎ２」という種別のメッセージの発信元の構成アイテムの方である。

【０２７９】

よって、ランキング生成部４０９は、ステップＳ１０９～Ｓ１１２にしたがって、１つの構成アイテムに関して算出した２つ以上のＷＦ－ＩＤＦ（ｆ，ｎ）値のうちの最大のものを採用する。

【０２８０】

さて、ステップＳ１０４で情報が取得された全メッセージについてのステップＳ１０６～Ｓ１１２の処理が完了すると、ランキング生成部４０９は、ステップＳ１１３で、スコア（つまりＷＦ－ＩＤＦ値）の降順に、ランキングテーブル５０７のエントリをソートする。そして、ランキング生成部４０９は、ソート結果に応じた順位を各エントリに記録する。図６には、以上のようにして順位づけされたランキングテーブル５０７が例示されている。

【０２８１】

さらに、ランキング生成部４０９は、ステップＳ１１３で、ランキングテーブル５０７を推定結果情報４３０として出力する。例えば、ランキング生成部４０９は、ランキングテーブル５０７の全エントリをランキング情報記憶部４１０に追加してもよい。ランキング生成部４０９は、ディスプレイ等の出力装置１０５に、ランキングテーブル５０７を出力してもよいし、通信インタフェイス１０３を介して他の装置にランキングテーブル５０７を出力してもよい。ランキング生成部４０９は、例えば、ランキングテーブル５０７を含む電子メールやインスタントメッセージなどを送信してもよい。

【０２８２】

ステップＳ１１３の出力後、検出サーバ４００は、再びステップＳ１０１でイベントの発生を待つ。

【０２８３】

以上の第２実施形態によれば、障害の発生を未然に防ぐうえで有益な示唆を与える推定結果情報４３０が、検出サーバ４００から出力される。よって、システム管理者は、推定結果情報４３０を参照することで、「障害の発生を未然に防ぐうえではどの構成アイテムに対して対策を講じるのが有効なのか」ということを、簡単に判断することができる。例えば、図６のランキングテーブル５０７を見たシステム管理者は、「障害＃７の予測と関連性が高い構成アイテムは、ＩＰアドレスＢ（１０．０．７．６）で識別される構成アイテムである」と判断することができる。場合によっては、システム管理者は、ランキングテーブル５０７に基づいて、「ＩＰアドレスＢ（１０．０．７．６）で識別される構成アイテムに対して対策をとることが、予測された障害＃７の発生を予防するうえで重要である」と判断してもよい。

【０２８４】

したがって、第２実施形態は、コンピュータシステムにおける障害の発生を予防することでコンピュータシステムの可用性を向上させる効果を奏する。

【０２８５】

続いて、図８～１４を参照して第３実施形態について説明する。第３実施形態では、第２実施形態の検出フェーズで生成されるランキング情報から、より信頼度の高い情報（以下、「改良（refined）ランキング情報」という）が生成される。具体的には、改良ランキング情報の生成においては、コンピュータシステムに含まれる構成アイテム間の関係（例えば論理的依存関係や物理的接続関係など）を示す情報が学習され、利用される。そして、第３実施形態の検出サーバは、生成した改良ランキング情報を出力する。

【０２８６】

第３実施形態は、コンピュータシステム内に、互いに同じかまたは互いに類似する複数の部分を含むような環境に特に好適である。なぜなら、第３実施形態によれば、コンピュータシステム内のある部分に生じる可能性のある障害を防ぐのに有益な改良ランキング情報を、当該ある部分と同じかまたは類似する他の部分に過去に生じた障害に応じて学習さ

れた情報から得ることも可能だからである。

【０２８７】

例えば、第３実施形態は、クラウド環境のインフラストラクチャを提供するためにデータセンタ内に設けられる大規模なコンピュータシステムに適用されてもよい。上記のような大規模なコンピュータシステムは、多数の物理サーバを含む。場合によっては、コンピュータシステムは、ディスクアレイ装置などのストレージ装置をさらに多数含むこともある。この種の環境では、例えば、いくつかの物理サーバが１つのネットワークデバイス（例えばＬ２スイッチなど）に接続される。また、各物理サーバが仮想化されることも多く、各物理サーバ上でそれぞれ複数の論理サーバが動作することも多い。

【０２８８】

したがって、コンピュータシステム内のある一部分（例えばある１つのブロードキャストドメイン）のネットワークトポロジが、他の一部分のネットワークトポロジと同じか、または類似している場合も多い。同様に、ある物理サーバ上のソフトウェア構成が、他の物理サーバ上のソフトウェア構成と同じか、または類似している場合も多い。つまり、上記のような大規模なコンピュータシステムは、互いに同じかまたは互いに類似する複数の部分を含むことが多い。よって、この種の大規模なコンピュータシステムには、第３実施形態が適用されることが好ましい。

【０２８９】

さて、図８は、第３実施形態における関係情報の学習を説明する図である。図８の例では、時刻ｔ２１にメッセージＭ２１が出力され、時刻ｔ２２にメッセージＭ２２が出力され、時刻ｔ２３にメッセージＭ２３が出力されたものとする。また、時刻ｔ２３を終了時点とするウィンドウには、メッセージＭ２１、Ｍ２２、およびＭ２３のみが含まれていたものとする。

【０２９０】

そして、メッセージＭ２１、Ｍ２２、およびＭ２３を含むメッセージパターン６０１に基づいて、障害＃３９の発生が予測されたものとする。つまり、メッセージパターン６０１が、障害＃３９の予兆パターンとして検知されたものとする。さらに、その後の時刻ｔ２４において、実際に障害＃３９が発生したことを通知するメッセージＭ２４が出力されたものとする。なお、図８では、メッセージＭ２１、Ｍ２２、Ｍ２３、およびＭ２４それぞれの発信元の構成アイテムのＩＰアドレスが、「Ｘ」、「Ｚ」、「Ｗ」、および「Ｙ」と示されている。

【０２９１】

時刻ｔ２４における実際の障害＃３９の発生により、時刻ｔ２３に行われた予測が正しかったことが判明する。つまり、時刻ｔ２３に検知されたメッセージパターン６０１が正しい予兆パターンであったことが、時刻ｔ２４に判明する。そこで、第３実施形態では、正しいことが判明した予兆パターン内の各メッセージの発信元の構成アイテムと、障害が発信した構成アイテムとの間の関係が、時刻ｔ２４（またはそれ以降）に学習される。

【０２９２】

図８には、例として、コンピュータシステムに含まれる複数の構成アイテムのうち、１７個の構成アイテムの間の関係が、グラフ６０２の形式で示されている。なお、構成アイテム間の関係を示す構成情報は、図８～９では理解の助けとするためにグラフの形式で示されている。しかし、構成情報の具体的なデータ形式は、実施形態に応じて任意である。

【０２９３】

グラフ６０２は、１７個の構成アイテムを示す１７個のノードＮ１～Ｎ１７を含む。なお、以下では説明の簡単化のため、あるノードＮｉにより表される構成アイテムのことも、単に「ノードＮｉ」ということがある（１≦ｉ）。

【０２９４】

ノードＮ１～Ｎ６は、ゲストＯＳのレイヤに属する。ノードＮ１、Ｎ２、Ｎ３およびＮ４が表す構成アイテムのＩＰアドレスは、それぞれ、「Ｘ」、「Ｙ」、「Ｚ」、および「Ｗ」である。なお、ゲストＯＳのレイヤは、論理サーバのレイヤのうちの１つである。

【０２９５】
また、図８～９の例では、ゲストＯＳと、当該ゲストＯＳ上で動作する全アプリケーションを含む集合が、ゲストＯＳのレイヤの１つの構成アイテムとして扱われる。ただし、以下では説明の簡単化のため、例えばノードＮ１により表される構成アイテム（すなわちアプリケーションを含む構成アイテム）のことを、単に「ゲストＯＳ」という場合もある。

【０２９６】
なお、図８～９の例では、メッセージの発信元がいずれもゲストＯＳのレイヤの構成アイテムであるが、これは偶然である。他のレイヤの構成アイテムがメッセージを出力することも、もちろんある。

【０２９７】
ノードＮ７～Ｎ１０は、ホストＯＳのレイヤに属する。なお、ホストＯＳのレイヤも、論理サーバのレイヤのうちの１つである。

【０２９８】
また、図８～９の例では、ハイパーバイザと、当該ハイパーバイザ上で動作するホストＯＳとを含む集合が、ホストＯＳのレイヤの１つの構成アイテムとして扱われる。ただし、以下では説明の簡単化のため、例えばノードＮ７により表される構成アイテムのことを単に「ホストＯＳ」という場合もある。

【０２９９】
ノードＮ１１～Ｎ１４は、物理サーバのレイヤに属する。また、ノードＮ１５～Ｎ１６はＬ２スイッチのレイヤに属し、ノードＮ１７はＬ３スイッチのレイヤに属する。

【０３００】
グラフ６０２によれば、ノードＮ１７により表されるＬ３スイッチ（例えば図３のＬ３スイッチ２９０）には、ノードＮ１５とＮ１６により表される２台のＬ２スイッチ（例えば図３のＬ２スイッチ２８０と２８１）が接続されている。グラフ６０２では、このようなネットワークデバイス間の直接的かつ物理的な接続関係は、２つのノード間のエッジにより表される。

【０３０１】
また、グラフ６０２によれば、ノードＮ１５により表されるＬ２スイッチには、ノードＮ１１とＮ１２により表される２台の物理サーバ（例えば図３の物理サーバ２４０と２５０）が接続されている。また、ノードＮ１６により表されるＬ２スイッチには、ノードＮ１３とＮ１４により表される２台の物理サーバ（例えば図３の物理サーバ２６０と２７０）が接続されている。

【０３０２】
グラフ６０２では、このようなネットワークデバイスと物理サーバの間の直接的かつ物理的な接続関係も、２つのノード間のエッジにより表される。また、例えばノードＮ１１からノードＮ１５を通ってノードＮ１７に至るパスは、物理サーバとＬ３スイッチの間の間接的な接続関係を示す。

【０３０３】
さらに、グラフ６０２によれば、ノードＮ１１により表される物理サーバ（例えば図３の物理サーバ２４０）上で、ノードＮ７により表されるホストＯＳ（例えば図３のホストＯＳ２４２）が動作する。また、ノードＮ１とＮ２により表されるゲストＯＳ（例えば図３のゲストＯＳ２４３と２４４）は、ノードＮ７により表されるホストＯＳの機能を利用する。グラフ６０２では、このようなハードウェアとソフトウェアの間の論理的依存関係や、２つのソフトウェア間の論理的依存関係も、２つのノード間のエッジにより表される。

【０３０４】
また、グラフ６０２によれば、ノードＮ１２により表される物理サーバ（例えば図３の物理サーバ２５０）上で、ノードＮ８により表されるホストＯＳ（例えば図３のホストＯＳ２５２）が動作する。また、ノードＮ３とＮ４により表されるゲストＯＳ（例えば図３

のゲストＯＳ２５３と２５４）は、ノードＮ８により表されるホストＯＳの機能を利用する。

【０３０５】

そして、グラフ６０２によれば、ノードＮ１３により表される物理サーバ（例えば図３の物理サーバ２６０）上で、ノードＮ９により表されるホストＯＳ（例えば図３のホストＯＳ２６２）が動作する。また、ノードＮ５により表されるゲストＯＳ（例えば図３のゲストＯＳ２６３）は、ノードＮ９により表されるホストＯＳの機能を利用する。

【０３０６】

さらに、グラフ６０２によれば、ノードＮ１４により表される物理サーバ（例えば図３の物理サーバ２７０）上で、ノードＮ１０により表されるホストＯＳ（例えば図３のホストＯＳ２７２）が動作する。また、ノードＮ６により表されるゲストＯＳ（例えば図３のゲストＯＳ２７３）は、ノードＮ１０により表されるホストＯＳの機能を利用する。

【０３０７】

例えば以上のようなグラフ６０２により表される構成情報を用いて、第３実施形態の検出サーバは、関係情報を学習する。具体的には、検出サーバは、検知した予兆パターンが正しかったことを認識すると、予兆パターン内の各メッセージと、障害を通知するメッセージを、グラフ６０２にマッピングする。

【０３０８】

例えば、図８の例では、メッセージＭ２１の発信元の構成アイテムは、「Ｘ」というＩＰアドレスで識別され、かつ、ノードＮ１により示される。また、メッセージパターン６０１が正しい予兆パターンであることが、時刻ｔ２４に判明する。よって、検出サーバは、メッセージＭ２１を、ノードＮ１にマッピングする。同様に、検出サーバは、メッセージＭ２２をノードＮ３にマッピングし、メッセージＭ２３をノードＮ４にマッピングする。

【０３０９】

また、時刻ｔ２４に障害＃３９が発生した構成アイテム（すなわち、障害＃３９の発生を通知するメッセージＭ２４の発信元）は、「Ｙ」というＩＰアドレスで識別され、かつ、ノードＮ２により示される。よって、検出サーバは、メッセージＭ２４をノードＮ２にマッピングする。

【０３１０】

そして、検出サーバは、予兆パターン内のメッセージがマッピングされたノードと、障害の発生を通知するメッセージがマッピングされたノードとの関係を学習する。２つのノード間の関係は、２つのノード間の最短パスにより一意に表される。よって、第３実施形態では、２つのノード間の最短パスが、２つのノードによりそれぞれ表される構成アイテム同士の関係を示す関係情報として学習される。具体的には、図８の例では、検出サーバはパスＰ１～Ｐ３を学習する。

【０３１１】

パスＰ１は、メッセージＭ２１の発信元の構成アイテムと、障害＃３９の発生した構成アイテムとの間の関係を示す。具体的には、パスＰ１は、ノードＮ１から始まり、ノードＮ７を通って、ノードＮ２に至るパスである。つまり、パスＰ１は、「正しい予測に使われた『１』という種別のメッセージの発信元は、予測された障害＃３９が実際に発生したゲストＯＳによって機能が利用されるホストＯＳの機能を利用する、他のゲストＯＳである」ということを示す。

【０３１２】

パスＰ２は、メッセージＭ２２の発信元の構成アイテムと、障害＃３９の発生した構成アイテムとの間の関係を示す。具体的には、パスＰ２は、ノードＮ３から始まり、ノードＮ８、Ｎ１２、Ｎ１５、Ｎ１１、およびＮ７を通って、ノードＮ２に至るパスである。つまり、パスＰ２は、「正しい予測に使われた『２』という種別のメッセージの発信元は、予測された障害＃３９が実際に発生したゲストＯＳが動作している物理サーバとＬ２スイッチを介して接続された他の物理サーバ上の、ゲストＯＳである」ということを示す。

【０３１３】
パスＰ３は、メッセージＭ２３の発信元の構成アイテムと、障害＃３９の発生した構成アイテムとの間の関係を示す。具体的には、パスＰ３は、ノードＮ４から始まり、ノードＮ８、Ｎ１２、Ｎ１５、Ｎ１１、およびＮ７を通って、ノードＮ２に至るパスである。つまり、パスＰ３は、「正しい予測に使われた『３』という種別のメッセージの発信元は、予測された障害＃３９が実際に発生したゲストＯＳが動作している物理サーバとＬ２スイッチを介して接続された他の物理サーバ上の、ゲストＯＳである」ということを示す。
【０３１４】
なお、２つのノードを結ぶパスは、複数あり得る。例えば、ノードＮ１からＮ２までの可能なパスの中には、例えば、ノードＮ１から始まって、ノードＮ７とＮ１１を通り、再度ノードＮ７に戻ってから、ノードＮ２に至るようなパスも、存在する。しかし、このパスは、ループを含み、したがって最短ではない。このように最短ではないパスは、ノードＮ１とＮ２の間の関係を示す関係情報としては使われない。
【０３１５】
検出サーバは、例えばワーシャル・フロイド法（Warshall-Floyd algorithm）などの公知のアルゴリズムを利用することで、最短パスを認識することができる。
【０３１６】
さて、第３実施形態の検出サーバは、以上のようにして障害の実際の発生に応じて学習した関係情報を、後に同じ種別の障害の発生が予測された際のランキング情報の改良に用いる。具体的には、第３実施形態の検出サーバは、何らかの種別の障害の発生を予測すると、まず、第２実施形態の検出サーバ４００と同様にしてランキング情報を生成する。そして、第３実施形態の検出サーバは、生成したランキング情報と、学習した関係情報に基づいて、改良ランキング情報を生成する。
【０３１７】
図９は、第３実施形態におけるランキングの改良について説明する図である。図９は、図８のパスＰ１～Ｐ３が学習された後に、メッセージＭ３１～Ｍ３３が出力され、メッセージＭ３１～Ｍ３３を含むメッセージパターンから、障害＃３９の発生が予測された場合を例示している。
【０３１８】
なお、メッセージＭ３１の種別は「３」であり、メッセージＭ３２の種別は「２」であり、メッセージＭ３３の種別は「１」であるものとする。また、障害＃３９の予測に使われるウィンドウ内には、メッセージＭ３１～Ｍ３３のみが含まれていたとする。
【０３１９】
ここで、コンピュータシステムには、図８に例示した１７個の構成アイテムだけでなく、さらに、図９に例示する１０個の構成アイテムが少なくとも含まれているものとする。図９では、これら１０個の構成アイテムの間の関係が、グラフ６０３の形式で示されている。
【０３２０】
具体的には、グラフ６０３は、１０個の構成アイテムを示す１０個のノードＮ２１～Ｎ３０を含む。ノードＮ２１～Ｎ２５はゲストＯＳのレイヤに属する。ノードＮ２１～Ｎ２５がそれぞれ表す構成アイテムのＩＰアドレスは、図９では、便宜上、「Ａ」、「Ｂ」、「Ｃ」、「Ｄ」、および「Ｅ」という文字により表されている。以下、説明の便宜上、例えば、ＩＰアドレスＡは１７２．１６．１．２であり、ＩＰアドレスＢは１０．０．７．６であり、ＩＰアドレスＣは１０．０．０．１であり、ＩＰアドレスＤは１０．０．０．１０であり、ＩＰアドレスＥは１０．０．０．３であるものとする。
【０３２１】
ノードＮ２６～Ｎ２７は、ホストＯＳのレイヤに属する。ノードＮ２８～Ｎ２９は、物理サーバのレイヤに属する。そして、ノードＮ３０は、Ｌ２スイッチのレイヤに属する。Ｌ３スイッチのレイヤはグラフ６０３では省略されている。
【０３２２】

さて、グラフ６０３によれば、ノードＮ３０により表されるＬ２スイッチには、ノードＮ２８とＮ２９により表される２台の物理サーバが接続されている。

【０３２３】

そして、グラフ６０３によれば、ノードＮ２８により表される物理サーバ上で、ノードＮ２６により表されるホストＯＳが動作する。また、ノードＮ２１、Ｎ２２、およびＮ２３により表される３つのゲストＯＳは、いずれも、ノードＮ２６により表されるホストＯＳの機能を利用する。

【０３２４】

さらに、グラフ６０３によれば、ノードＮ２９により表される物理サーバ上で、ノードＮ２７により表されるホストＯＳが動作する。また、ノードＮ２４およびＮ２５により表される２つのゲストＯＳは、いずれも、ノードＮ２７により表されるホストＯＳの機能を利用する。

【０３２５】

ここで、メッセージＭ３１の発信元が、ノードＮ２１により表されるゲストＯＳ（すなわち、ＩＰアドレスＡ（１７２．１６．１．２）で識別される構成アイテム）であるものとする。また、メッセージＭ３２の発信元が、ノードＮ２３により表されるゲストＯＳ（すなわち、ＩＰアドレスＣ（１０．０．０．１）で識別される構成アイテム）であるものとする。そして、メッセージＭ３３の発信元が、ノードＮ２５により表されるゲストＯＳ（すなわち、ＩＰアドレスＥ（１０．０．０．３）で識別される構成アイテム）であるものとする。

【０３２６】

また、上記のとおり、メッセージＭ３１～Ｍ３３を含むメッセージパターンから、障害＃３９の発生が予測されたものとする。したがって、この場合、第３実施形態の検出サーバは、第２実施形態の検出サーバ４００と同様にして、メッセージＭ３１～Ｍ３３の発信元たる３つの構成アイテムのそれぞれについてＷＦ－ＩＤＦ（ｆ，ｎ）を算出する。そして、検出サーバは、算出した３つの値を使って、ランキング情報６０４を生成する。ランキング情報６０４の形式は、図４のランキング情報３０５と同様である。

【０３２７】

ランキング情報６０４によれば、メッセージＭ３３を出力した構成アイテムについて算出されたＷＦ－ＩＤＦ（３９，１）は、２．０００００であり、３つの値の中で最大である。また、メッセージＭ３２を出力した構成アイテムについて算出されたＷＦ－ＩＤＦ（３９，２）は０．００４３である。同様に、メッセージＭ３１を出力した構成アイテムについて算出されたＷＦ－ＩＤＦ（３９，３）も０．００４３である。よって、ＩＰアドレスＥで識別される構成アイテムの順位は１位であり、ＩＰアドレスＣとＡでそれぞれ識別される２つの構成アイテムの順位はいずれも２位である。

【０３２８】

第３実施形態の検出サーバは、学習済みの関係情報（具体的には図８のパスＰ１～Ｐ３）を用いて、ランキング情報６０４から改良ランキング情報６０５を生成する。ここで、図９のランキング情報６０４と改良ランキング情報６０５の例から分かるように、ランキング情報と改良ランキング情報には以下のような違いがある。

【０３２９】

・ランキング情報では、障害の予測に用いられたメッセージパターンに含まれるメッセージを少なくとも１つ出力したすべての構成アイテムに、スコアが与えられている。

・ランキング情報では、障害の予測に用いられたメッセージパターンに含まれるメッセージを１つも出力していない構成アイテムに対しては、スコアは与えられない。

・改良ランキング情報では、障害の予測に用いられたメッセージパターンに含まれるメッセージを１つも出力していない構成アイテムに対しても、スコアが与えられる場合があり得る。

・改良ランキング情報では、障害の予測に用いられたメッセージパターンに含まれるメッセージを少なくとも１つ出力した構成アイテムについて、スコアが与えられない場合

があり得る。

【０３３０】

以下、検出サーバが改良ランキング情報６０５を生成する方法について、具体的に説明する。

メッセージＭ３１の種別は「３」であり、「３」というメッセージ種別に関して学習された関係情報は、図８のパスＰ３である。そこで、検出サーバは、パスＰ３で示される関係と等価な関係が、メッセージＭ３１の発信元との間で成り立つような構成アイテム（以下、「関連構成アイテム」ともいう）を検索する。具体的には、グラフ６０３において、メッセージＭ３１の発信元を表すノードＮ２１から始まり、かつ、パスＰ３とトポロジ的に相似なパスを、検出サーバが、たどってゆく（traverse）。そして、検出サーバは、パスＰ３と相似な当該パスの終点のノードにより表される構成アイテムを、メッセージＭ３１にとっての関連構成アイテムとして、認識する。

【０３３１】

なお、図９の例では、パスＰ３と相似なパスは複数ある。しかし、「パスＰ３と相似なパス自体が、始点たるノードＮ２１と、パスＰ３と相似な当該パスの終点との間の最短パスである」という条件（以下「最短パス条件」という）を満たすパスは２つだけである。メッセージＭ３１にとっての関連構成アイテムは、より正確には、パスＰ３と相似なパスのうち、最短パス条件を満たすパスの終点のノードにより表される構成アイテムである。

【０３３２】

図８に示すように、パスＰ３は、ゲストＯＳのレイヤのノードから始まる。そして、パスＰ３は、ホストＯＳのレイヤのノード、物理サーバのレイヤのノード、Ｌ２スイッチのレイヤのノード、物理サーバのレイヤのノード、および、ホストＯＳのレイヤのノードを通って、ゲストＯＳのレイヤのノードに至る。グラフ６０３において、ノードＮ２１から始まって上記のパスＰ３と同じ順に種々のレイヤのノードを通るパスは複数ある。しかし、最短パス条件を満たすパスは２つのみである。

【０３３３】

例えば、ノードＮ２１から始まって、ノードＮ２６、Ｎ２８、Ｎ３０、Ｎ２８、およびＮ２６を通って、ノードＮ２２に至るパスは、パスＰ３と相似ではあるが、最短パス条件を満たさない。それに対して、以下の２つのパスは、いずれも、パスＰ３と相似であり、かつ、最短パス条件を満たす。

【０３３４】

・ノードＮ２１から始まって、ノードＮ２６、Ｎ２８、Ｎ３０、Ｎ２９、およびＮ２７を通って、ノードＮ２４に至るパス（このパスは、図９にパスＰ１３として示されている）。

・ノードＮ２１から始まって、ノードＮ２６、Ｎ２８、Ｎ３０、Ｎ２９、およびＮ２７を通って、ノードＮ２５に至るパス。

【０３３５】

よって、検出サーバは、「３」という種別のメッセージＭ３１にとっての関連構成アイテムとして、ノードＮ２４とＮ２５で表される２つの構成アイテムを認識する。つまり、メッセージＭ３１にとっての関連構成アイテムは、ＩＰアドレスＤとＥによりそれぞれ識別される２つの構成アイテムである。

【０３３６】

さて、メッセージＭ３２の種別は「２」であり、「２」というメッセージ種別に関して学習された関係情報は、図８のパスＰ２である。そこで、グラフ６０３において、メッセージＭ３２の発信元を表すノードＮ２３から始まり、かつ、パスＰ２とトポロジ的に相似であり、かつ、最短パス条件を満たすパスを、検出サーバがたどってゆく。検出サーバは、こうしてたどったパスの終点のノードにより表される構成アイテムを、メッセージＭ３２にとっての関連構成アイテムとして認識する。具体的には、ノードＮ２３から始まり、かつ、パスＰ２と相似であり、かつ、最短パス条件を満たすようなパスは、以下の２つである。

【０３３７】
・ノードＮ２３から始まって、ノードＮ２６、Ｎ２８、Ｎ３０、Ｎ２９、およびＮ２７を通って、ノードＮ２４に至るパス（このパスは、図９にパスＰ１２として示されている）。
・ノードＮ２３から始まって、ノードＮ２６、Ｎ２８、Ｎ３０、Ｎ２９、およびＮ２７を通って、ノードＮ２５に至るパス。
【０３３８】
よって、検出サーバは、「２」という種別のメッセージＭ３２にとっての関連構成アイテムとして、ノードＮ２４とＮ２５で表される２つの構成アイテムを認識する。つまり、メッセージＭ３２にとっての関連構成アイテムも、ＩＰアドレスＤとＥによりそれぞれ識別される２つの構成アイテムである。
【０３３９】
さて、メッセージＭ３３の種別は「１」であり、「１」というメッセージ種別に関して学習された関係情報は、図８のパスＰ１である。そこで、グラフ６０３において、メッセージＭ３３の発信元を表すノードＮ２５から始まり、かつ、パスＰ１とトポロジ的に相似であり、かつ、最短パス条件を満たすパスを、検出サーバがたどってゆく。
【０３４０】
ここで、ノードＮ２５から始まり、かつ、パスＰ１と相似なパスは、２つある。１つは、ノードＮ２５から始まり、ノードＮ２７を通って、ノードＮ２５に戻るパスである。しかし、このパスは最短パス条件を満たさない。もう１つは、ノードＮ２５から始まり、ノードＮ２７を通って、ノードＮ２４に至るパスＰ１１である。パスＰ１１は最短パス条件を満たす。
【０３４１】
よって、検出サーバは、「１」という種別のメッセージＭ３３にとっての関連構成アイテムとして、パスＰ１１の終点のノードＮ２４により表される構成アイテムを認識する。
【０３４２】
以上より、ＩＰアドレスＤで識別される構成アイテムは、メッセージＭ３１にとっての関連構成アイテムでもあり、メッセージＭ３２にとっての関連構成アイテムでもあり、メッセージＭ３３にとっての関連構成アイテムでもある。よって、検出サーバは、メッセージＭ３１とＭ３２とＭ３３それぞれの発信元について算出したＷＦ－ＩＤＦ（３９，３）とＷＦ－ＩＤＦ（３９，２）とＷＦ－ＩＤＦ（３９，１）のうちの最大値を、ＩＰアドレスＤで識別される構成アイテムのスコアに決定する。
【０３４３】
ここで、図９のランキング情報６０４によれば、ＷＦ－ＩＤＦ（３９，３）＝０．００４３であり、ＷＦ－ＩＤＦ（３９，２）＝０．００４３であり、ＷＦ－ＩＤＦ（３９，１）＝２．０００００である。よって、ＩＰアドレスＤで識別される構成アイテムのスコアは、２．０００００である。
【０３４４】
また、ＩＰアドレスＥで識別される構成アイテムは、メッセージＭ３１にとっての関連構成アイテムでもあり、メッセージＭ３２にとっての関連構成アイテムでもある。よって、検出サーバは、メッセージＭ３１とＭ３２それぞれの発信元について算出したＷＦ－ＩＤＦ（３９，３）とＷＦ－ＩＤＦ（３９，２）のうちの最大値を、ＩＰアドレスＥで識別される構成アイテムのスコアに決定する。つまり、ＩＰアドレスＥで識別される構成アイテムのスコアは、０．００４３である。
【０３４５】
ＩＰアドレスＤとＥで識別される２つの構成アイテム以外の構成アイテムは、メッセージＭ３１とＭ３２とＭ３３のいずれにとっても、関連構成アイテムではない。よって、検出サーバは、上記２つの構成アイテムについて決定したスコアに基づいて、上記２つの構成アイテムの順位を決定する。すなわち、２．００００というスコアが与えられた構成アイテム（つまり、ＩＰアドレスＤで識別される構成アイテム）の順位が１位であり、０．

００４３というスコアが与えられた構成アイテム（つまり、ＩＰアドレスＥで識別される構成アイテム）の順位が２位である。

【０３４６】

改良ランキング情報６０５では、以上のようにして決定された順位とスコアが、スコアの付与の根拠となったメッセージの種別とともに、ＩＰアドレスに対応づけられている。

【０３４７】

以上の例では、障害＃３９の予測に使われたウィンドウの中では、たまたま、ＩＰアドレスＤで識別される構成アイテムからは何もメッセージが出力されていないが、それにもかかわらず、ＩＰアドレスＤで識別される構成アイテムが１位と判定される。このように、改良ランキング情報６０５の生成においては、正解した予兆パターンたるメッセージパターン６０１中のメッセージの発信元と、時刻ｔ２４に実際に障害が発生した構成アイテムとの間の関係と等価な関係が利用される。

【０３４８】

こうして生成された改良ランキング情報６０５は、ＷＦ－ＩＤＦ（ｆ，ｎ）のような統計値に基づくだけでなく、関係情報にも基づいているため、ランキング情報６０４と比べて信頼性がより高い。よって、第３実施形態によれば、検出サーバは、障害の発生を防ぐための対策を講じることが望ましい構成アイテムを、より高い信頼性をもって示唆する情報を提供することが可能である。

【０３４９】

また、以上のように関係情報を利用する第３実施形態は、互いに同じかまたは互いに類似する複数の部分（例えば、グラフ６０２で示される部分とグラフ６０３で示される部分）を含む大規模なコンピュータシステムに特に好適である。なぜなら、関係情報の利用により、予兆パターンの学習に関するデータスパースネス問題が軽減され、検出サーバが提示する情報の信頼度が高まるからである。

【０３５０】

続いて、図１０～１４を参照して、図８～９を参照して説明した第３実施形態のさらなる詳細について説明する。

図１０は、第３実施形態の検出サーバ７００のブロック構成図である。検出サーバ７００は、コンピュータシステム内の種々の構成アイテムからメッセージ７２０を入力として受け取り、推定結果情報７３０を出力する。推定結果情報７３０は、具体的には、例えば図９の改良ランキング情報６０５であってもよい。

【０３５１】

検出サーバ７００は、第２実施形態の検出サーバ４００内のコンポーネントと類似のいくつかのコンポーネントを含む。具体的には、検出サーバ７００は、ログ情報記憶部７０１と、障害予兆検知部７０２と、辞書情報記憶部７０３と、障害予兆情報記憶部７０４を含む。また、検出サーバ７００は、ログ統計算出部７０５と、ログ統計情報記憶部７０６と、予兆統計算出部７０７と、予兆統計情報記憶部７０８と、ランキング生成部７０９と、ランキング情報記憶部７１０も含む。

【０３５２】

さらに、検出サーバ７００は、検出サーバ４００には存在しないいくつかのコンポーネントも含む。具体的には、検出サーバ７００は、トポロジ関係学習部７１１と構成情報記憶部７１２と関係情報記憶部７１３と推定部７１４をさらに含む。

【０３５３】

ログ情報記憶部７０１にはメッセージ７２０が蓄積される。ログ情報記憶部７０１、障害予兆検知部７０２、辞書情報記憶部７０３、障害予兆情報記憶部７０４、ログ統計算出部７０５、ログ統計情報記憶部７０６、予兆統計算出部７０７、および予兆統計情報記憶部７０８は、第２実施形態の各コンポーネントと同様である。

【０３５４】

ランキング生成部７０９は、第２実施形態のランキング生成部４０９と同様にランキング情報（例えば図９のランキング情報６０４）を生成し、生成したランキング情報をラン

キング情報記憶部７１０に記憶する。しかし、第３実施形態では、ランキング生成部７０９の生成したランキング情報自体ではなく、ランキング情報から得られる改良ランキング情報（例えば図９の改良ランキング情報６０５）が、推定結果情報７３０として出力される。

【０３５５】

ランキング情報記憶部７１０は、第２実施形態のランキング情報記憶部４１０と同様にランキング情報を記憶する。さらに、ランキング情報記憶部７１０は、改良ランキング情報も記憶する。

【０３５６】

トポロジ関係学習部７１１は、図８に例示したように、障害予兆検知部７０２により検知された予兆パターンが正解と判明した場合に、その正しい予兆パターンに含まれる各メッセージの発信元と、障害が実際に発生した構成アイテムとの間の関係情報を学習する。そして、トポロジ関係学習部７１１は、学習した関係情報を関係情報記憶部７１３に記憶する。具体的には、第３実施形態のトポロジ関係学習部７１１は、ログ情報記憶部７０１、障害予兆情報記憶部７０４、ランキング情報記憶部７１０、および構成情報記憶部７１２を参照して、関係情報を学習する。

【０３５７】

なお、実施形態によっては、トポロジ関係学習部７１１は、必ずしも、ログ情報記憶部７０１とランキング情報記憶部７１０を参照する必要はない。例えば、障害予兆情報記憶部７０４内に、検知された予兆パターンに含まれる各メッセージの発信元のＩＰアドレスが記憶される場合、トポロジ関係学習部７１１は、障害予兆情報記憶部７０４と構成情報記憶部７１２を参照して関係情報を学習してもよい。トポロジ関係学習部７１１による学習の詳しい手順の例は、図１２とともに後述する。

【０３５８】

構成情報記憶部７１２には、コンピュータシステムの複数の構成アイテム間の関係を表す構成情報が記憶される。構成情報は、コンピュータシステムの構成（configuration）が変更されると、それに応じて変更される。例えば、新たな構成アイテムの追加、既存の構成アイテムの削除、またはマイグレーションなどが行われると、構成情報は変更される。構成情報記憶部７１２は、公知の構成管理データベース（Configuration Management Database；ＣＭＤＢ）であってもよい。

【０３５９】

なお、図８のグラフ６０２と図９のグラフ６０３は、いずれも、構成情報の一部を、便宜上、グラフ形式で視覚的に表現したものである。構成情報記憶部７１２内の構成情報の実際のデータ形式は、実施形態に応じて任意である。例えば、テーブル形式が利用されてもよいし、ＸＭＬ（Extensible Markup Language）などの所定言語を用いた形式が利用されてもよい。

【０３６０】

また、第３実施形態の構成情報においては、各構成アイテムは、識別情報としてのＩＰアドレスにより識別されるものとする。よって、推定部７１４は、例えば図９のようにパスの終点を探すことにより、パスの終点の構成アイテムのＩＰアドレスを認識することが可能である。

【０３６１】

関係情報記憶部７１３には、トポロジ関係学習部７１１により学習された関係情報が記憶される。関係情報記憶部７１３の詳細は図１１とともに後述する。

【０３６２】

推定部７１４は、ランキング生成部７０９の生成したランキング情報と、関係情報記憶部７１３に記憶されている学習済みの関係情報と、構成情報記憶部７１２に記憶されている構成情報を用いて、改良ランキング情報を生成する。換言すれば、推定部７１４は、障害予兆検知部７０２により予測された障害との関連性が高い構成アイテム（つまり障害が発生する蓋然性の高い構成アイテム）を、コンピュータシステム内の構成アイテム間の関

係に基づいて推定する。推定の結果が改良ランキング情報である。また、障害との関連性が高いと推定される構成アイテムは、場合によっては、対策を講じることで障害の発生を予防する効果が得られる見込みの高い構成アイテムそのものである。

【０３６３】

なお、ある障害が、他の障害により、直接的または間接的に引き起こされることもあり得る。よって、「ある障害が発生する蓋然性が高い」と推定された構成アイテムそのものではなく、原因となる他の障害が生じそうな他の構成アイテムに対して、対策をとることが有益な場合もあり得る。しかし、その場合でも、システム管理者等は、改良ランキング情報から、「どの構成アイテムに対して対策をとることが障害の発生を防ぐうえで有益なのか」に関する示唆を得ることができる。なぜなら、改良ランキング情報は、「上記ある障害がどの構成アイテムにおいて発生する蓋然性が高いのか」を示すので、対策をとる対象の構成アイテムの候補を絞り込むのに役立つからである。

【０３６４】

推定部７１４は、生成した改良ランキング情報（例えば図９の改良ランキング情報６０５）を推定結果情報７３０として出力する。例えば、推定部７１４は、推定結果情報７３０としての改良ランキング情報を、ディスプレイに出力してもよいし、ランキング情報記憶部７１０に出力してもよい。推定部７１４は、改良ランキング情報を含む電子メールまたはインスタントメッセージを、システム管理者に宛てて送信してもよい。実施形態によっては、推定部７１４がログ情報を参照してもよい。

【０３６５】

ところで、図１０の検出サーバ７００は、具体的には図２のコンピュータ１００であってもよい。検出サーバ７００がコンピュータ１００により実現される場合、図２と図１０は以下のように対応する。

【０３６６】

検出サーバ７００は、通信インタフェイス１０３を介してメッセージ７２０を受信する。また、検出サーバ７００は、推定結果情報７３０を出力装置１０５に出力してもよく、記憶装置１０６に出力してもよく、駆動装置１０７を介して記憶媒体１１０に出力してもよい。もちろん、検出サーバ７００は、通信インタフェイス１０３とネットワーク１２０を介して推定結果情報７３０を送信（つまり出力）してもよい。

【０３６７】

ログ情報記憶部７０１、辞書情報記憶部７０３、障害予兆情報記憶部７０４、ログ統計情報記憶部７０６、予兆統計情報記憶部７０８、ランキング情報記憶部７１０、構成情報記憶部７１２、および関係情報記憶部７１３は、記憶装置１０６により実現されてもよい。障害予兆検知部７０２、ログ統計算出部７０５、予兆統計算出部７０７、ランキング生成部７０９、トポロジ関係学習部７１１、および推定部７１４は、プログラムを実行するＣＰＵ１０１により実現されてもよい。

【０３６８】

また、図１０の検出サーバ７００は、図３のコンピュータ２００であってもよい。この場合、メッセージ７２０は、コンピュータシステム２３０内の種々の構成アイテムから出力されて、ネットワーク２１０を介して、検出サーバ７００としてのコンピュータ２００に受信される。また、コンピュータシステム２３０のシステム管理者は、検出サーバ７００から出力される推定結果情報７３０を参照して、コンピュータシステム２３０内のどの構成アイテムに対して対策をとるかを決め、適宜の対策を実行する。

【０３６９】

続いて、図１０中の種々の記憶部に記憶される情報の具体例について、図１１を参照して説明する。図１１は、第３実施形態で利用される各種テーブルの例を示す図である。

なお、ログ情報記憶部７０１と辞書情報記憶部７０３中のテーブルについては、図１１では図示を省略した。例えば図６のログテーブル５０１と同様のテーブルがログ情報記憶部７０１に記憶されてもよい。また、図６のメッセージ辞書テーブル５０２およびパターン辞書テーブル５０３と同様のテーブルが辞書情報記憶部７０３に記憶されてもよい。

【０３７０】
さて、図１１の障害予兆テーブル８０１は、障害予兆情報記憶部７０４に記憶される情報の一例である。障害予兆テーブル８０１中に例示された種々の値は、図６の障害予兆テーブル５０４中に例示された種々の値とは異なるが、障害予兆テーブル８０１の形式は障害予兆テーブル５０４と同様である。
【０３７１】
なお、障害予兆テーブル５０４と同様に、障害予兆テーブル８０１も、予測された障害の終了時刻を示すフィールドをさらに含んでいてもよい。また、実施形態によっては、障害予兆テーブル８０１には、障害予兆検知部７０２により検知された予兆パターンに含まれる各メッセージの種別だけでなく、各メッセージの発信元のＩＰアドレスがさらに記憶されていてもよい。
【０３７２】
図１１の障害予兆テーブル８０１には、図８の時刻ｔ２３にメッセージパターン６０１に基づいて行われた予測の結果が、「１」というＩＤのエントリに記憶されている。また、図９に示した予測の結果が、「２」というＩＤのエントリに記憶されている。
【０３７３】
ログ統計テーブル８０２は、ログ統計情報記憶部７０６に記憶される情報の一例である。ログ統計テーブル８０２に例示された種々の値は、図６のログ統計テーブル５０５中に例示された種々の値とは異なるが、ログ統計テーブル８０２の形式はログ統計テーブル５０５と同様である。
【０３７４】
なお、図１１には、図９でランキング情報６０４が生成される時点におけるログ統計テーブル８０２の４つのエントリが例示されている。また、ログ統計テーブル８０２は、「１」～「３」以外のメッセージ種別に対応する他のエントリをさらに含み得るが、図１１ではそれらのエントリは省略されている。
【０３７５】
予兆統計テーブル８０３は、予兆統計情報記憶部７０８に記憶される情報の一例である。予兆統計テーブル８０３に例示された種々の値は、図６の予兆統計テーブル５０６中に例示された種々の値とは異なるが、予兆統計テーブル８０３の形式は予兆統計テーブル５０６と同様である。
【０３７６】
図１１には、図９でランキング情報６０４が生成される時点における予兆統計テーブル８０３の４つのエントリが例示されている。換言すれば、図１１には、図８の時刻ｔ２４での障害＃３９の発生を契機に学習された内容が例示されている。予兆統計テーブル８０３は、「時刻ｔ２４を終了時点とする予測対象期間内で障害＃３９の予測に成功していたのは、１回だけ（つまり時刻ｔ２３での予測だけ）であった」ということを示している。なお、予兆統計テーブル８０３は、「３９」以外の障害種別に対応する他のエントリをさらに含み得るが、図１１ではそれらのエントリは省略されている。
【０３７７】
さて、トポロジ関係テーブル８０４は、関係情報記憶部７１３に記憶される関係情報の一例である。障害の発生が正しく予測され、その正しい予測において検知された予兆パターンがＰ（１≦Ｐ）個のメッセージを含む場合、トポロジ関係学習部７１１により、トポロジ関係テーブル８０４にＰ個のエントリが追加される。トポロジ関係テーブル８０４の各エントリは、例えば以下の５つのフィールドを含んでもよい。
【０３７８】
・上記の正しい予測を表すエントリを障害予兆テーブル８０１の中で識別するＩＤ（以下「予兆ＩＤ」という）。
・トポロジ関係テーブル８０４内で個々のエントリを識別するＩＤ。
・上記の正しく予測された障害の種別。
・上記の正しい予測で使われたメッセージパターン（つまり、検知された予兆パター

ン）中の個々のメッセージの種別。
・上記予兆パターンに含まれるメッセージのうちで、当該エントリのメッセージ種別で表されるメッセージを出力した、発信元の構成アイテムと、上記の正しく予測された障害が生じた構成アイテムとの間の関係を示すパス。
【０３７９】
なお、トポロジ関係テーブル８０４における上記パスは、第３実施形態では、具体的には、図８のグラフ６０２のようなグラフにおける、発信元の構成アイテムのノードから、障害が生じた構成アイテムのノードに至るパスである。また、第３実施形態では、このように２つの構成アイテム間の関係を示すパスは、具体的には、ＸＰａｔｈ形式で表される。ＸＰａｔｈ形式でのパスの表現は、ある種のＦＣＭＤＢ（federated CMDB）でのクエリに利用されているので、ここでは詳しい説明を省略する。第３実施形態との関連という観点から、ＸＰａｔｈ形式でのパスの表現について概略を説明すれば、以下のとおりである。
【０３８０】
トポロジ関係テーブル８０４の３つのエントリのパスは、それぞれ、図８のパスＰ１、Ｐ２、およびＰ３を表す。例えば、２番目のエントリ中のＸＰａｔｈ式は、パスＰ２を表す。図８に示すように、パスＰ２は、以下に示すノードとエッジの系列（sequence）である。
【０３８１】
・論理サーバのレイヤ（具体的にはゲストＯＳのレイヤ）のノードＮ３（すなわち、「２」という種別のメッセージの発信元を示すノード）。
・ノードＮ３から、論理サーバのレイヤ（具体的にはホストＯＳのレイヤ）のノードＮ８に至るエッジ。
・ノードＮ８。
・ノードＮ８から、物理サーバのレイヤのノードＮ１２に至るエッジ。
・ノードＮ１２。
・ノードＮ１２から、ネットワークデバイスのレイヤ（具体的にはＬ２スイッチのレイヤ）のノードＮ１５に至るエッジ。
・ノードＮ１５。
・ノードＮ１５から、物理サーバのレイヤのノードＮ１１に至るエッジ。
・ノードＮ１１。
・ノードＮ１１から、論理サーバのレイヤ（具体的にはホストＯＳのレイヤ）のノードＮ７に至るエッジ。
・ノードＮ７。
・ノードＮ７から、論理サーバのレイヤ（具体的にはゲストＯＳのレイヤ）のノードＮ２（すなわち、障害＃３９が実際に発生した構成アイテムを示すノード）に至るエッジ。
・ノードＮ２。
【０３８２】
ところで、トポロジ関係テーブル８０４におけるＸＰａｔｈ式は、図９に関して説明したように、具体的には、トポロジ的に相似なパスの検索のために使われる。よって、第３実施形態では、パスＰ２そのものを具体的に示す情報ではなく、パスＰ２がどのレイヤのノードをどういう順で通るのかを示すＸＰａｔｈ式が使われる。
【０３８３】
例えば、トポロジ関係テーブル８０４の２番目のエントリ中のＸＰａｔｈ式は、以下のことを示している。パスＰ２と相似なパスを検索するには、このようなＸＰａｔｈ式により表される、多少一般化された形式の関係情報だけで十分である。
【０３８４】
・パス上の１番目のノード（つまりパスの始点）は、論理サーバのレイヤのノードである。

・パス上の２番目のノードは、論理サーバのレイヤのノードである。
・パス上の３番目のノードは、物理サーバのレイヤのノードである。
・パス上の４番目のノードは、ネットワークデバイスのレイヤのノードである。
・パス上の５番目のノードは、物理サーバのレイヤのノードである。
・パス上の６番目のノードは、論理サーバのレイヤのノードである。
・パス上の７番目のノードは、論理サーバのレイヤのノードであり、この７番目のノードがパスの終点である。

【０３８５】
なお、実施形態に応じて、ＸＰａｔｈ以外の形式によってパスが表現されてもよいことは無論である。ＸＰａｔｈ式は、２つの構成アイテム間の関係を示すための所定フォーマットのデータの一例に過ぎない。

【０３８６】
さて、ランキングテーブル８０５は、ランキング生成部７０９が第２実施形態のランキング生成部４０９と同様にして生成するテーブルである。よって、ランキングテーブル８０５の形式は、図６のランキングテーブル５０７の形式と同じである。

【０３８７】
図１１のランキングテーブル８０５には、図９のランキング情報６０４に対応する３つのエントリが例示されている。また、ランキングテーブル８０５の各エントリにおける予兆ＩＤは、当該エントリのスコア（すなわちＷＦ－ＩＤＦ（ｆ，ｎ））を算出する契機となった予測を識別するためのＩＤであり、具体的には、障害予兆テーブル８０１内のエントリを識別するＩＤである。

【０３８８】
例えば、ランキングテーブル８０５に例示した３つのエントリの予兆ＩＤは、いずれも「２」である。つまり、これら３つのエントリは、障害予兆テーブル８０１において「２」というＩＤを有する２番目のエントリの予測（すなわち図９の予測）の際に生成されたランキング情報に対応する。

【０３８９】
改良ランキングテーブル８０６は、ランキングテーブル８０５に基づいて推定部７１４が生成するテーブルである。改良ランキングテーブル８０６の形式はランキングテーブル８０５と同じである。例えば、改良ランキングテーブル８０６に例示されている２つのエントリは、図９の改良ランキング情報６０５に対応する。なお、改良ランキング情報６０５は、障害予兆テーブル８０１において「２」というＩＤで識別される予測が行われると生成される。よって、図１１の改良ランキングテーブル８０６中の２つのエントリの予兆ＩＤは、いずれも「２」である。

【０３９０】
第３実施形態では、ランキングテーブル８０５と改良ランキングテーブル８０６の双方が、ランキング情報記憶部７１０に記憶される。また、図１１のランキングテーブル８０５には、予兆ＩＤが「２」の３つのエントリのみが例示されているが、ランキング情報記憶部７１０中のランキングテーブル８０５は、予兆ＩＤが「１」の３つのエントリも含む。つまり、ランキング情報記憶部７１０中のランキングテーブル８０５には、図９の予測に応じて得られたランキング情報だけでなく、図８の時刻ｔ２３における予測に応じて得られたランキング情報も、記憶されている。

【０３９１】
続いて、検出サーバ７００が行う処理について、さらに詳しく説明する。なお、第２実施形態と同様に、検出サーバ７００が行う種々の処理のうち、ログ情報記憶部７０１へのメッセージ７２０の蓄積と、パターン辞書テーブル５０３の学習と、障害予兆検知部７０２による障害予兆の検知は、公知の処理と同様であってよい。また、検出サーバ７００も図７と類似の処理を実行するが、図７のステップＳ１０３とＳ１１３が、第３実施形態では変形される。

【０３９２】

具体的には、第３実施形態では、図７のステップＳ１０３が以下のように変形される。

・予兆統計算出部７０７が、第２実施形態のステップＳ１０３と同様の方法により、予兆統計情報記憶部７０８を更新する。

・トポロジ関係学習部７１１が、図１２のフローチャートにしたがって、図８に例示したように関係情報を学習する。

【０３９３】

また、第３実施形態では、図７のステップＳ１１３が以下のように変形される。

・ランキング生成部７０９が、第２実施形態のステップＳ１１３と同様にして、ランキングテーブル８０５のエントリをソートし、各エントリに順位をつける。また、ランキング生成部７０９は、ランキングテーブル８０５の各エントリをランキング情報記憶部７１０に追加する。

・さらに、ランキング生成部７０９は、ランキングテーブル８０５を推定部７１４にも出力する。その際、ランキング生成部７０９は、障害予兆検知部７０２により予測された障害の種別も、推定部７１４に通知する。なお、障害予兆検知部７０２により予測された障害の種別は、既にステップＳ１０１で、障害予兆検知部７０２からランキング生成部７０９へと通知されている。

・推定部７１４は、ランキング生成部７０９から受け取ったランキングテーブル８０５のメッセージ種別フィールドに基づいて、予測に使われたメッセージパターンを認識する。例えば、図１１のランキングテーブル８０５からは、メッセージパターン［１，２，３］が認識される。

・そして、推定部７１４は、認識したメッセージパターンと、ランキング生成部７０９から通知された障害の種類の組み合わせに対応して既に学習された関係情報を、関係情報記憶部７１３において検索する。

・検索の結果、学習済みの関係情報が見つかった場合、推定部７１４は、図１３～１４のフローチャートにしたがって、図９に例示したように改良ランキング情報（例えば図１１の改良ランキングテーブル８０６）を生成および出力する。

・検索の結果、学習済みの関係情報が見つからなかった場合は、推定部７１４は、受け取ったランキングテーブル８０５自体をメッセージ７２０として出力してもよい。

【０３９４】

なお、実施形態によっては、検索の結果、学習済みの関係情報が見つからなかった場合、推定部７１４は以下のような処理を行ってもよい。

【０３９５】

推定部７１４は、受け取ったランキングテーブル８０５から認識したメッセージパターンを包含するメッセージパターンと、ランキング生成部７０９から通知された障害の種類との組み合わせに対応して既に学習された関係情報を、検索してもよい。なおここで、第１のメッセージパターンに含まれる全メッセージが第２のメッセージパターンにも含まれる場合、「第２のメッセージパターンは第１のメッセージパターンを包含する」ということにする。例えば、メッセージパターン［１，２］は、メッセージパターン［１，２，３，４］に包含される。

【０３９６】

例えば、メッセージパターン［１，２］から障害＃５が予測されたが、メッセージパターン［１，２］と障害＃５の組み合わせに対応して学習済みの関係情報がまだ存在しない場合があり得る。この場合、仮にメッセージパターン［１，２，３，４］と障害＃５の組み合わせに対応して学習済みの関係情報があれば、推定部７１４は、当該関係情報を利用してもよい。つまり、メッセージパターン［１，２］を包含する他のメッセージパターンと障害＃５との組み合わせに関する再検索の結果、関係情報が見つかれば、推定部７１４は、再検索の結果に基づいて、ランキングテーブルから改良ランキングテーブルを生成してもよい。そして、推定部７１４は、そのようにして生成した改良ランキングテーブルを、推定結果情報７３０として出力してもよい。

【０３９７】

あるいは、推定部７１４は、受け取ったランキングテーブル８０５から認識したメッセージパターンと類似するメッセージパターンと、ランキング生成部７０９から通知された障害の種類との組み合わせに対応して既に学習された関係情報を検索してもよい。例えば、メッセージパターン［１，２］から障害＃５が予測されたが、メッセージパターン［１，２］と障害＃５の組み合わせに対応して学習済みの関係情報がまだ存在しない場合があり得る。この場合、推定部７１４は、例えば、メッセージパターン［１，１０］と障害＃５の組み合わせや、メッセージパターン［２，１８］と障害＃５の組み合わせに対応して学習された関係情報を検索してもよい。２つのメッセージパターンが類似するかどうかの基準は実施形態に応じて任意だが、互いに類似するメッセージパターン同士は、少なくとも１つの同じ種別のメッセージを含む。

【０３９８】

さて、図１２は、第３実施形態において検出サーバ７００（具体的にはトポロジ関係学習部７１１）が関係情報を学習する処理のフローチャートである。第３実施形態では、障害が発生すると、トポロジ関係学習部７１１が図１２の処理を実行する。

【０３９９】

なお、トポロジ関係学習部７１１は、検出サーバ７００の受信するメッセージ７２０から、障害の発生を認識してもよいし、ログ情報記憶部７０１へのエントリの追加を監視することで障害の発生を認識してもよい。あるいは、障害の発生に応じて図７のステップＳ１０３の処理を実行する予兆統計算出部７０７が、障害の発生をトポロジ関係学習部７１１に通知してもよい。いずれにせよ、何らかの障害が発生すると、トポロジ関係学習部７１１は図１２の処理を開始する。

【０４００】

ステップＳ２０１でトポロジ関係学習部７１１は、既に検知された予兆パターンのうち、今回発生した障害を正しく予測していた各予兆パターンについての障害予兆情報を取得する。換言すれば、トポロジ関係学習部７１１は、既に行われた予測のうち、今回発生した障害を正しく予測していた各予測についての障害予兆情報を取得する。具体的には、トポロジ関係学習部７１１は、今回の障害の発生に先立つ長さＴ２の予測対象期間に行われた予測結果を、障害予兆情報記憶部７０４から検索する。この検索は、図７のステップＳ１０３で予兆統計算出部４０７が行う検索と類似である。

【０４０１】

例えば、図８の時刻ｔ２４で障害＃３９が発生すると、トポロジ関係学習部７１１は図１２の処理を実行し始める。図８の例では、時刻ｔ２４と時刻ｔ２３の差が長さＴ２以下であるものとする。よって、トポロジ関係学習部７１１は、障害予兆テーブル８０１の障害種別と予測実行時刻のフィールドを参照して検索を行うと、障害予兆テーブル８０１の１番目のエントリ（つまり時刻ｔ２３での予測結果を示すエントリ）を取得する。なお、こうして１番目のエントリが取得されることは、「時刻ｔ２４に実際に発生した障害＃３９について、時刻ｔ２３（図１１の例では２０１２年８月３１日２３時）に検知された予兆パターン［１，２，３］は、正しいと判明した」ということを意味する。

【０４０２】

なお、発生した障害に対して、長さＴ２の予測対象期間内の過去においては１回も正しい予測に成功していなかった場合もあり得る。また、発生した障害に対して、長さＴ２の予測対象期間内の過去において、１回だけ正しい予測に成功していた場合もあり得るし、２回以上正しい予測に成功していた場合もあり得る。よって、ステップＳ２０１で障害予兆情報記憶部７０４から取得されるエントリの数は、０個の場合もあり得るし、１個の場合もあり得るし、２個以上の場合もあり得る。

【０４０３】

次に、ステップＳ２０２でトポロジ関係学習部７１１は、ステップＳ２０１で取得した正しい予兆パターンのうち、未処理の予兆パターンがあるか否かを判断する。つまり、トポロジ関係学習部７１１は、ステップＳ２０１で取得したエントリのうち、ステップＳ２０３以降の処理の対象としてまだ選択していないエントリがあるか否かを判断する。

【０４０４】
ステップＳ２０１で１個もエントリが取得されなかった場合か、または、ステップＳ２０１で取得された全エントリが既にステップＳ２０３以降の処理の対象として選択済みの場合、未処理の予兆パターンは存在しない。よって、図１２の関係情報の学習は終了する。

【０４０５】
逆に、ステップＳ２０１で１個以上のエントリが取得され、その中に、ステップＳ２０３以降の処理の対象としてまだ選択されていないエントリがある場合、未処理の予兆パターンが存在する。よって、この場合、トポロジ関係学習部７１１は次に、ステップＳ２０３で、未処理の予兆パターンを１つ選択する。つまり、ステップＳ２０３でトポロジ関係学習部７１１は、ステップＳ２０１で取得した、ある１つのエントリを選択する。以下では説明の便宜上、ステップＳ２０３で選択されたエントリの予兆パターンを「選択予兆パターン」ともいう。

【０４０６】
さらに、ステップＳ２０３でトポロジ関係学習部７１１は、選択予兆パターンが検知されたときにＷＦ－ＩＤＦ値が算出された１つまたは複数の構成アイテムそれぞれについてのエントリを、ランキング情報記憶部７１０内のランキングテーブル８０５から取得する。

【０４０７】
例えば、図８の時刻ｔ２４での障害＃３９の発生を契機として、トポロジ関係学習部７１１が図１２の処理を実行する場合、ステップＳ２０１では、時刻ｔ２３での予測に対応するエントリが取得される。つまり、この場合、障害予兆テーブル８０１の１番目のエントリがステップＳ２０１で取得され、ステップＳ２０３で選択される。

【０４０８】
すると、トポロジ関係学習部７１１は、ステップＳ２０３でさらに、障害予兆テーブル８０１の１番目のエントリのＩＤを読み取る。そして、トポロジ関係学習部７１１は、読み取ったＩＤの値を検索キーとして用いて、ランキング情報記憶部７１０内のランキングテーブル８０５を検索する。図１１では省略されているが、ランキングテーブル８０５には、図８の時刻ｔ２３での予測に応じて、メッセージＭ２１、Ｍ２２、およびＭ２３それぞれの発信元の構成アイテムについて追加された３つのエントリがある。

【０４０９】
よって、トポロジ関係学習部７１１は、検索の結果、３つのエントリを取得することができる。つまり、トポロジ関係学習部７１１は、「Ｘ」、「Ｚ」、および「Ｗ」というＩＰアドレスでそれぞれ識別される３つの構成アイテムについて、時刻ｔ２３の予測の際ににランキングテーブル８０５に追加された３つのエントリを取得する。

【０４１０】
次に、ステップＳ２０４でトポロジ関係学習部７１１は、ステップＳ２０３で取得したエントリのうち、未処理の構成アイテムについてのエントリがまだ残っているか否かを判断する。つまり、トポロジ関係学習部７１１は、正しいと判明したある１つの予兆パターンに含まれるメッセージを少なくとも１つ出力した構成アイテムのうち、まだ関係情報の学習が済んでいないものが残っているか否かを判断する。

【０４１１】
具体的には、ステップＳ２０３でランキングテーブル８０５から取得したエントリのうち、ステップＳ２０５～Ｓ２０８の処理対象としてまだ選択されていないものが残っていれば、図１２の学習処理は、次にステップＳ２０５に進む。逆に、ステップＳ２０３でランキングテーブル８０５から取得された全エントリについて、ステップＳ２０５～Ｓ２０８が実行済みであれば、図１２の学習処理は、ステップＳ２０２に戻る。

【０４１２】
そして、ステップＳ２０５でトポロジ関係学習部７１１は、未処理の構成アイテムを１つ選択する。つまり、トポロジ関係学習部７１１は、ステップＳ２０３でランキングテー

ブル８０５から取得したエントリのうちの、未処理の１つを選択する（ランキングテーブル８０５の１つのエントリは１つの構成アイテムに対応することに注意されたい）。以下では説明の便宜上、ステップＳ２０５で選択された構成アイテムを「選択構成アイテム」ともいう。

【０４１３】

次に、ステップＳ２０６でトポロジ関係学習部７１１は、構成情報記憶部７１２に記憶されている構成情報を参照して、選択構成アイテムから今回障害が発生した構成アイテムまでの最短パスを認識する。

【０４１４】

例えば、上記のようにステップＳ２０４で、図８の「Ｘ」、「Ｚ」、および「Ｗ」というＩＰアドレスでそれぞれ識別される３つの構成アイテムについての３つのエントリが、ランキング情報記憶部７１０中のランキングテーブル８０５から取得されたとする。そして、ステップＳ２０５では、「Ｘ」というＩＰアドレスで識別される構成アイテムに対応するエントリが選択されたとする。また、図８によれば、時刻ｔ２４に実際に障害＃３９が発生した構成アイテムは、「Ｙ」というＩＰアドレスで識別される。よって、この場合、ステップＳ２０６でトポロジ関係学習部７１１は、構成情報を参照して、図８のパスＰ１を認識する。パスＰ１が最短パスであることは、図８から明らかである。

【０４１５】

なお、構成情報は、図８にグラフ６０２の形式で示されるような構成アイテム間の関係を定義するだけでなく、さらに、任意の２つの構成アイテム間の最短パスに関する情報を含んでいてもよい。例えば、検出サーバ７００は、予めワーシャル・フロイド法などの公知のアルゴリズムを利用して、任意の２つの構成アイテム間の最短パスを求めてもよい。こうして事前に判明した最短パスが構成情報記憶部７１２に記憶されていてもよい。この場合、トポロジ関係学習部７１１は、記憶されている最短パスの情報を読み出すだけで最短パスを認識することができる。もちろん、トポロジ関係学習部７１１は、ステップＳ２０６で、例えばダイクストラ法（Dijkstra's algorithm）などの公知のアルゴリズムを利用して、動的に最短パスを探索してもよい。

【０４１６】

いずれにせよ、トポロジ関係学習部７１１は、最短パスを認識した後、ステップＳ２０７において、認識した最短パスを表すＸＰａｔｈ式を生成する。例えば、ステップＳ２０６でトポロジ関係学習部７１１が図８のパスＰ１を最短パスとして認識した場合、トポロジ関係学習部７１１は、図１１のトポロジ関係テーブル８０４の１番目のエントリに例示されているようなＸＰａｔｈ式を、ステップＳ２０７で生成する。

【０４１７】

そして、次のステップＳ２０８でトポロジ関係学習部７１１は、生成したＸＰａｔｈ式をトポロジ関係テーブル８０４に記録する。具体的には、トポロジ関係学習部７１１は、ステップＳ２０５でランキングテーブル８０５から選択したエントリのメッセージ種別フィールドに記憶されている種別の数と同数の新規エントリを、トポロジ関係テーブル８０４に追加する。

【０４１８】

例えば、ある正しい予兆パターンに含まれるメッセージのうちの３つが１つの構成アイテムから出力されており、当該構成アイテムについてのランキングテーブル８０５のエントリがステップＳ２０５で選択されたとする。この場合、ステップＳ２０８では、３つのエントリがトポロジ関係テーブル８０４に追加される。

【０４１９】

トポロジ関係テーブル８０４に追加される各新規エントリのメッセージ種別の値は、ステップＳ２０５で選択したエントリのメッセージ種別フィールドに記憶されている各種別の値に等しい。また、トポロジ関係学習部７１１は、各新規エントリに対して、当該新規エントリを識別するためのＩＤを新たに発行する。

【０４２０】

なお、ステップＳ２０８でトポロジ関係テーブル８０４に追加される各新規エントリにおいて、予兆ＩＤの値は、ステップＳ２０１で障害予兆テーブル８０１から取得されたエントリのうち、ステップＳ２０３で選択されたエントリのＩＤである。また、各新規エントリにおける障害種別は、図１２の処理をトポロジ関係学習部７１１が開始する契機となった障害の種別である。そして、各新規エントリにおけるパスは、ステップＳ２０７で生成されたＸＰａｔｈ式である。

【０４２１】

以上のようにしてステップＳ２０８でトポロジ関係テーブル８０４に１つ以上のエントリが追加されると、図１２の学習処理は、再びステップＳ２０４に戻る。

【０４２２】

さて、図１３～１４は、第３実施形態の検出サーバ７００（具体的には推定部７１４）が、学習した関係情報を使って改良ランキング情報を生成する処理のフローチャートである。上記のとおり、図１３～１４の処理は、あるメッセージパターンに基づいてある種類の障害の発生が予測され、かつ、当該あるメッセージパターンと当該ある種類の障害との組み合わせについて関係情報が学習済みの場合に、実行される。

【０４２３】

さて、ステップＳ３０１で推定部７１４は、改良ランキングテーブル８０６を空に初期化する。

【０４２４】

なお、図１１に関してはあまり詳しく説明しなかったが、第３実施形態に関しては、「改良ランキングテーブル」という名称を以下の２つのテーブルに共通に用いて説明をしている。

【０４２５】

・ある１回の予測に対応して、推定部７１４がローカルに生成するテーブル。

・推定部７１４により生成されたテーブルの各エントリが蓄積される、ランキング情報記憶部７１０内のテーブル。

【０４２６】

つまり、ある観点から見れば、図１１の改良ランキングテーブル８０６は、図９に例示された１回の予測に対応して、推定部７１４がローカルに生成した、２つのエントリを有するテーブルを示したものである。一方、別の観点から見れば、図１１の改良ランキングテーブル８０６は、ランキング情報記憶部７１０内で改良ランキング情報を記憶するテーブルについて、２つのエントリのみを抜粋して例示的に示したものである。

【０４２７】

しかし、説明の簡単化のため、本明細書では、両者とも単に「改良ランキングテーブル８０６」と呼んでいる。同様に、ランキング生成部７０９がローカルに生成するテーブルと、ランキング情報記憶部７１０内に蓄積されるテーブルの双方も、本明細書では共通の「ランキングテーブル８０５」という名前で参照している。

【０４２８】

図１３～１４の説明における改良ランキングテーブル８０６は、より詳しくは、推定部７１４がローカルに生成するテーブルの方である。よって、ステップＳ３０１では、ローカルなテーブルが初期化される。

【０４２９】

次に、ステップＳ３０２で推定部７１４は、ランキング生成部７０９から出力されたランキングテーブル８０５に未処理のエントリがあるか否かを判断する。ランキングテーブル８０５の全エントリについて、ステップＳ３０３～Ｓ３１２の処理が完了していれば、推定部７１４は次にステップＳ３１３の処理を実行する。逆に、ランキングテーブル８０５の中に未処理のエントリが残っていれば、推定部７１４は次にステップＳ３０３の処理を実行する。

【０４３０】

ステップＳ３０３で推定部７１４は、ランキング生成部７０９から出力されたランキン

グテーブル８０５中の未処理のエントリを１つ選択する。ステップＳ３０３で選択されたエントリを、以下では便宜上「選択エントリ」ともいう。

【０４３１】

次に、ステップＳ３０４で推定部７１４は、選択エントリからスコア（すなわち、選択エントリの構成アイテムについて算出されたＷＦ－ＩＤＦ（ｆ，ｎ））を読み取る。

【０４３２】

そして、ステップＳ３０５で推定部７１４は、選択エントリ中の各メッセージ種別と、障害予兆検知部７０２によって今回予測された障害の種別の組み合わせに対応するパスを、トポロジ関係テーブル８０４から読み取る。より具体的には、選択エントリのメッセージ種別フィールドには、１つ以上の種別のリストが記憶されている。よって、推定部７１４は、リスト中の各種別について、以下の３つの条件をすべて満たすエントリをトポロジ関係テーブル８０４の中から検索し、見つかったエントリからパスを読み取る。

【０４３３】

・予兆ＩＤフィールドの値により識別される、障害予兆テーブル８０１中のエントリにおける予兆パターンが、障害予兆検知部７０２が今回検知した予兆パターンと等しい（なお、後者の予兆パターンは、換言すれば、推定部７１４がランキング生成部７０９から受け取ったランキングテーブル８０５の予兆ＩＤフィールドの値により識別される、障害予兆テーブル８０１中のエントリに記憶されている予兆パターンである）。

・障害種別フィールドの値は、障害予兆検知部７０２が今回予測した障害の種別（つまり、推定部７１４がランキング生成部７０９から通知された種別）と等しい。

・メッセージ種別フィールドの値は、選択エントリ中のメッセージ種別フィールドのリスト中のいずれかの値に等しい。

【０４３４】

なお、ステップＳ３０５で読み取られるパスの数は、１つの場合もあり得るし、複数の場合もあり得る。例えば、選択エントリが図１１のランキングテーブル８０５の２番目のエントリである場合、ステップＳ３０５では、図１１のトポロジ関係テーブル８０４の２番目のエントリのパス（すなわち、図８のパスＰ２を示すＸＰａｔｈ式）が得られる。また、例えば、ある特定のメッセージパターンに基づくある特定の種別の障害の予測が、過去に２回以上当たっていた場合は、ステップＳ３０５で２つ以上のパスが得られる場合がある。選択エントリのメッセージ種別フィールドに２つ以上の種別が記録されている場合にも、ステップＳ３０５で２つ以上のパスが得られる場合がある。

【０４３５】

さて、次に、ステップＳ３０６で推定部７１４は、構成情報記憶部７１２に記憶された構成情報を参照して、選択エントリのＩＰアドレスを持つ構成アイテムを始点として、ステップＳ３０５で読み取ったパスと相似なパスをたどるとたどりつく終点の構成アイテムを検索する。以下では説明の便宜上、検索の結果見つかった構成アイテムを「終点構成アイテム」という。なお、図９に関して説明したように、ステップＳ３０６では、最短パス条件を満たすパスの終点の構成アイテムのみが検索される。

【０４３６】

なお、上記のように、構成情報において各構成アイテムは、ＩＰアドレスにより識別されている。よって、推定部７１４は、終点構成アイテムのＩＰアドレスも、検索の結果として取得することができる。

【０４３７】

例えば、選択エントリが図１１のランキングテーブル８０５の１番目のエントリである場合、ステップＳ３０５ではトポロジ関係テーブル８０４の１番目のエントリのパス（すなわち図８のパスＰ１を示すＸＰａｔｈ式）が得られる。また、選択エントリのＩＰアドレスは、ＩＰアドレスＥである。よって、推定部７１４は、ＩＰアドレスＥを持つ構成アイテム（つまり図９のノードＮ２５で表される構成アイテム）を始点として、パスＰ１と相似なパスＰ１１をたどる。すると、終点構成アイテムとして、ノードＮ２４で表される構成アイテム（すなわち、ＩＰアドレスＤで識別される構成アイテム）が見つかる。

【０４３８】
また、選択エントリが図１１のランキングテーブル８０５の２番目のエントリである場合、図９に関する説明から分かるように、２つの終点構成アイテムが見つかる。つまり、ノードＮ２４とＮ２５により表される２つの構成アイテムが見つかる。同様に、選択エントリが図１１のランキングテーブル８０５の３番目のエントリである場合も、ノードＮ２４とＮ２５により表される２つの構成アイテムが、終点構成アイテムとして見つかる。
【０４３９】
以上のように、ステップＳ３０６では、１つだけ終点構成アイテムが見つかる場合もあるし、複数の終点構成アイテムが見つかる場合もある。しかし、場合によっては、ステップＳ３０６で１つも終点構成アイテムが見つからない場合もあり得る。
【０４４０】
なお、ステップＳ３０５で２つ以上のパスが読み取られた場合、ステップＳ３０６では、各パスについて、終点構成アイテムの検索が行われる。その結果、複数の終点構成アイテムが得られる場合もあり得るし、たまたま、２つ以上のパスについて得られた終点構成アイテムが同じ場合もあり得る。
【０４４１】
そこで、ステップＳ３０７で推定部７１４は、未処理の終点構成アイテムがあるか否かを判断する。ステップＳ３０６で１つも終点構成アイテムが見つからなかったか、または、ステップＳ３０６で見つかったすべての終点構成アイテムについてステップＳ３０８～Ｓ３１２の処理が完了している場合、推定部７１４は、再度ステップＳ３０２の判断を行う。
【０４４２】
逆に、ステップＳ３０６で１つ以上の終点構成アイテムが見つかり、そのうちステップＳ３０８～Ｓ３１２の処理の対象として未選択のものが残っている場合は、推定部７１４は、次に、ステップＳ３０８で、未選択の終点構成アイテムを１つ選択する。以下では説明の便宜上、ステップＳ３０８で選択された終点構成アイテムを「選択終点構成アイテム」という。
【０４４３】
続いて、ステップＳ３０９で推定部７１４は、選択終点構成アイテムのＩＰアドレスが既に改良ランキングテーブル８０６に含まれているか否かを判断する。
【０４４４】
例えば、選択構成アイテムが、図９のノードＮ２４で表される構成アイテム（つまりＩＰアドレスＤにより識別される構成アイテム）である場合、推定部７１４は、ＩＰアドレスＤを検索キーとして用いて改良ランキングテーブル８０６を検索する。検索の結果、エントリが見つかれば、推定部７１４は、「選択終点構成アイテムのＩＰアドレスが既に改良ランキングテーブル８０６に含まれている」と判断する。逆に、エントリが見つからなければ、推定部７１４は、「選択終点構成アイテムのＩＰアドレスは改良ランキングテーブル８０６に含まれていない」と判断する。
【０４４５】
選択終点構成アイテムのＩＰアドレスが改良ランキングテーブル８０６に含まれていない場合、推定部７１４は、次にステップＳ３１０の処理を行う。逆に、選択終点構成アイテムのＩＰアドレスが既に改良ランキングテーブル８０６に含まれている場合、推定部７１４は、次にステップＳ３１１の処理を行う。
【０４４６】
ステップＳ３１０で推定部７１４は、改良ランキングテーブル８０６に、以下の４つの値を含む新規エントリを追加する。
【０４４７】
・推定部７１４がランキング生成部７０９から受け取ったランキングテーブル８０５の全エントリに共通の予兆ＩＤの値。この予兆ＩＤの値は、図１３～１４の処理を推定部７１４が開始する契機となった予測の結果を障害予兆検知部７０２が障害予兆情報記憶部

７０４に記憶する際に用いたＩＤに等しい。
・選択終点構成アイテムを識別するＩＰアドレス。
・選択エントリのＩＰアドレスを持つ１つの構成アイテムに関して、現在の選択終点構成アイテムをステップＳ３０６で検索した際に用いたパスが１つだけの場合は、当該１つのパスがステップＳ３０５で読み取られた際に検索キーとして使われたメッセージ種別。現在の選択終点構成アイテムをステップＳ３０６で検索した際に用いたパスが２つ以上ある場合は、当該２つ以上のパスがステップＳ３０５で読み取られた際に検索キーとしてそれぞれ使われたメッセージ種別のリスト。
・ランキングテーブル８０５中の選択エントリからステップＳ３０４で読み取られたスコア。

【０４４８】
なお、ステップＳ３１０で追加される新規エントリにおいて、順位のフィールドは空でよい。エントリの追加後、推定部７１４は、再びステップＳ３０７の判断を行う。

【０４４９】
他方、ステップＳ３１１は、例えば、ランキングテーブル８０５の２つ以上のエントリに対応する２つ以上の構成アイテムをそれぞれ始点とするパスのそれぞれの終点として、たまたま同じ１つの構成アイテムが見つかった場合に、実行され得る。例えば、図９の例では、パスＰ１１の終点も、パスＰ１２の終点も、パスＰ１３の終点も、ノードＮ２４である。よって、ノードＮ２４で表される構成アイテム（つまりＩＰアドレスＤで識別される構成アイテム）についてのエントリが、ステップＳ３０９における検索の結果として見つかる場合が２回ある。

【０４５０】
具体的には、ステップＳ３１１で推定部７１４は、改良ランキングテーブル８０６内のスコアが、ステップＳ３０４でランキングテーブル８０５の選択エントリから読み取ったスコアより大きいか否かを判断する。なおここで、「改良ランキングテーブル８０６内のスコア」とは、具体的には、ステップＳ３０９での改良ランキングテーブル８０６の検索の結果見つかったエントリ内のスコアのことである。

【０４５１】
改良ランキングテーブル８０６内のスコアが、ステップＳ３０４で選択エントリから読み取ったスコアより大きい場合、ステップＳ３０９の検索で見つかったエントリを更新する必要はない。よって、この場合、推定部７１４は、次にステップＳ３０７の判断を行う。

【０４５２】
逆に、改良ランキングテーブル８０６内のスコアが、ステップＳ３０４で選択エントリから読み取ったスコア以下の場合、推定部７１４は、次に、ステップＳ３１２で改良ランキングテーブル８０６のエントリを更新する。すなわち、推定部７１４は、ステップＳ３０９での改良ランキングテーブル８０６の検索の結果見つかったエントリを更新する。具体的には以下のとおりである。

【０４５３】
改良ランキングテーブル８０６内のスコアが、ステップＳ３０４で読み取ったスコアより小さい場合、推定部７１４は、スコアフィールドの値を、ステップＳ３０４で読み取ったスコアに置き換える。また、この場合、推定部７１４は、メッセージ種別フィールドを次の内容に置き換える。

【０４５４】
・選択エントリのＩＰアドレスを持つ１つの構成アイテムに関して、現在の選択終点構成アイテムをステップＳ３０６で検索した際に用いたパスが１つだけの場合は、当該１つのパスがステップＳ３０５で読み取られた際に検索キーとして使われたメッセージ種別。
・現在の選択終点構成アイテムをステップＳ３０６で検索した際に用いたパスが２つ以上ある場合は、当該２つ以上のパスがステップＳ３０５で読み取られた際に検索キーと

してそれぞれ使われたメッセージ種別のリスト。
【０４５５】
一方、改良ランキングテーブル８０６内のスコアと、ステップＳ３０４で読み取ったスコアが互いに等しい場合、推定部７１４は、スコアフィールドは更新しないが、メッセージ種別フィールドのリストに次の内容を追加する。
【０４５６】
・選択エントリのＩＰアドレスを持つ１つの構成アイテムに関して、現在の選択終点構成アイテムをステップＳ３０６で検索した際に用いたパスが１つだけの場合は、当該１つのパスがステップＳ３０５で読み取られた際に検索キーとして使われたメッセージ種別。
・現在の選択終点構成アイテムをステップＳ３０６で検索した際に用いたパスが２つ以上ある場合は、当該２つ以上のパスがステップＳ３０５で読み取られた際に検索キーとしてそれぞれ使われたメッセージ種別。
【０４５７】
以上のような更新の後、推定部７１４は、ステップＳ３０７の判断を行う。なお、ステップＳ３０９～Ｓ３１２によれば、改良ランキングテーブル８０６のメッセージ種別フィールドには、「どの種別のメッセージの発信元との関係に基づいて、終点構成アイテムにスコアが与えられたのか」という情報が示されることになる。
【０４５８】
ところで、推定部７１４がランキング生成部７０９から受け取ったランキングテーブル８０５の全エントリが選択済みの場合、図１３～１４の処理は、ステップＳ３０２からステップＳ３１３に移行する。
【０４５９】
ステップＳ３１３では、推定部７１４は、スコアの降順に、改良ランキングテーブル８０６のエントリをソートする。そして、推定部７１４は、ソート結果に応じた順位を各エントリに記録する。図１１には、以上のようにして順位づけされた改良ランキングテーブル８０６が例示されている。
【０４６０】
ステップＳ３１３ではさらに、推定部７１４は、改良ランキングテーブル８０６を推定結果情報７３０として出力する。例えば、推定部７１４は、以上のようにしてローカルに生成した改良ランキングテーブル８０６の各エントリを、ランキング情報記憶部７１０内のテーブルに追加してもよい。推定部７１４は、ディスプレイ等の出力装置１０５に、改良ランキングテーブル８０６を出力してもよいし、通信インタフェイス１０３を介して他の装置に改良ランキングテーブル８０６を出力してもよい。推定部７１４は、例えば、改良ランキングテーブル８０６を含む電子メールやインスタントメッセージなどを送信してもよい。
【０４６１】
ステップＳ３１３での出力後、図１３～１４の処理は終了する。よって、検出サーバ７００は、再度、図７のステップＳ１０１でイベントの発生を待つ。
【０４６２】
以上、図８～１４を参照して説明した第３実施形態によれば、関係情報が考慮に入れられた、より信頼度の高い改良ランキング情報が提示される。また、第３実施形態では、「大規模コンピュータシステムには、互いに類似する構成を持つ複数の部分が含まれる場合が多い」という特徴が利用される。この特徴を利用することで、大規模コンピュータシステムに関する学習のデータスパースネス問題も軽減される。
【０４６３】
なお、関係情報を用いない第２実施形態において推定結果情報４３０として出力されるランキング情報も、実用上、十分に信頼度が高い情報である。
【０４６４】
なぜなら、一般的な傾向として、障害＃ｆに対して大きなＷＦ－ＩＤＦ（ｆ，ｎ）が算

出されるような「ｎ」という種別のメッセージは、障害＃ｆと偶然よく共起するというよりも、障害＃ｆと直接または間接の因果関係がある場合が多いからである。そして、このように障害＃ｆと密接に関連する「ｎ」という種別のメッセージの発信元は、経験的には、障害＃ｆが生じる構成アイテム自体である場合が比較的多い。

【０４６５】

したがって、「大きなＷＦ－ＩＤＦ（ｆ，ｎ）が算出されるような『ｎ』という種別のメッセージの発信元の構成アイテムに対して、何らかの対策を講じることが、障害＃ｆの発生を予防するうえで有益である」という場合も、多いのである。よって、たとえ第２実施形態のように関係情報が使われなくても、実用上、十分に信頼度が高くて十分に有益なランキング情報が得られる。

【０４６６】

なお、ある種別の障害の予兆として検知されたメッセージパターンに含まれるいずれかのメッセージの発信元において、たまたま、当該メッセージパターンから予測された種別の障害が生じることもあり得る。

【０４６７】

例えば、図８の例において、メッセージＭ２２が、「Ｚ」というＩＰアドレスで識別される構成アイテムではなく、「Ｙ」というＩＰアドレスで識別される構成アイテムから出力されたとする。この場合、障害＃３９の予兆として検知されたメッセージパターン６０１に含まれるメッセージＭ２２の発信元は、予測された障害＃３９が生じる構成アイテムと、たまたま等しい。よって、この場合にメッセージＭ２２に関して学習されるパスは、「Ｙ」というＩＰアドレスで識別される構成アイテムから、「Ｙ」というＩＰアドレスで識別される構成アイテム自身へ至る最短パスである。つまり、この場合、メッセージＭ２２に関しては、空パスが学習される。なお、ある構成アイテムから当該構成アイテム自身へ至る空パスは、空パスを表すための特定の文字列（空文字列ではない文字列）により表現されてもよい。

【０４６８】

関係情報として空パスが学習され、空パスが図１３のステップＳ３０５で読み取られる場合、ステップＳ３０６で見つかる終点構成アイテムは、パスの始点の構成アイテム自体（つまり選択エントリのＩＰアドレスで識別される構成アイテム）である。

【０４６９】

なお、本発明は第１～第３実施形態に限られるものではなく、第１～第３実施形態は様々に変形可能である。以下に、第１～第３実施形態を変形するいくつかの観点を例示する。以下に述べる変形は、相互に矛盾しない限り、任意に組み合わせることが可能である

図６と図１１には種々のテーブルを例示したが、種々の情報の形式は、実施形態に応じて任意である。テーブル以外のデータ形式が利用されてもよいし、例示した以外のフィールドをさらに含むテーブルが使われてもよい。

【０４７０】

また、式（１）のＷＦ－ＩＤＦ（ｆ，ｎ）以外の統計値が使われてもよい。ＷＦ－ＩＤＦ（ｆ，ｎ）の各種変形については上述したとおりである。

【０４７１】

ところで、推定結果情報４３０の例としてランキングテーブル５０７を示し、推定結果情報７３０の例として改良ランキングテーブル８０６を示したが、推定結果情報の形式は実施形態に応じて任意である。

【０４７２】

例えば、順位が上位Ｕ位までの構成アイテムの識別情報のみが、推定結果情報として出力されてもよい（１≦Ｕ）。また、順位とスコア（すなわちＷＦ－ＩＤＦ（ｆ，ｎ））のうち少なくとも一方が、構成アイテムの識別情報と対応づけられて推定結果情報に含まれていれば、それで十分である。つまり、必ずしも順位とスコアの双方が必要なわけではない。また、推定結果情報においては、メッセージ種別は省略可能である。もちろん、ランキングテーブル８０５と改良ランキングテーブル８０６の双方を含む情報が、推定結果情

報７３０として出力されてもよい。

【０４７３】

そして、第１実施形態に関しても説明したとおり、ＷＦ－ＩＤＦ（ｆ，ｎ）等の値による評価対象の構成アイテムの粒度は、実施形態に応じて様々であってよい。例えば、ゲストＯＳとアプリケーションが別々の構成アイテムとして扱われる実施形態も可能であるし、ゲストＯＳと、ゲストＯＳ上で動作するアプリケーションの集合が１つの構成アイテムとして扱われる実施形態も可能である。各構成アイテムを識別する識別情報は、構成アイテムの粒度に応じた適宜の情報であってよい。

【０４７４】

ところで、第２～第３実施形態についての説明では、障害の発生を知らせるメッセージとそれ以外のイベントを知らせるメッセージを区別した。しかし、実施形態によっては、障害予兆検知部４０２または７０２が、ある種の障害（例えば軽微な障害）の発生を知らせるメッセージを含むメッセージパターンから、別種の障害（例えば重大な障害）の発生を予測することがあってもよい。

【０４７５】

例えば第２実施形態がこのように変形される場合、ログ統計算出部４０５は、「受信されたメッセージ４２０が、障害の発生についての通知であるのか、それとも、その他のイベントについての通知であるのか」ということによらず、ステップＳ１０２と同様にログ統計テーブル５０５を更新してもよい。受信されたメッセージ４２０が障害の発生についての通知である場合には、さらに、予兆統計算出部４０７がステップＳ１０３の処理を実行する。なお、この場合、ステップＳ１０３がステップＳ１０２より先に実行されてもよい。第３実施形態も同様に変形されてもよい。

【０４７６】

第２～第３実施形態におけるランキング情報の生成においては、図７のステップＳ１０９～Ｓ１１２に示すように、いくつかの値のうちの最大値を採用する処理が行われる場合がある。同様に、第３実施形態における改良ランキング情報の生成においても、図１４のステップＳ３０９～Ｓ３１２に示すように、いくつかの値のうちの最大値を採用する処理が行われる場合がある。

【０４７７】

しかし、実施形態によっては、いくつかの値のうちの最大値を採用する処理の代わりに、いくつかの値の算術和または重みづけ和を採用する処理が行われてもよい。例えば、図９の例において、推定部７１４は、ノードＮ２４で表される構成アイテムに対して、ＷＦ－ＩＤＦ（３９，１）とＷＦ－ＩＤＦ（３９，２）とＷＦ－ＩＤＦ（３９，３）の最大値の代わりに、これら３つの値の算術和または重みづけ和を与えてもよい。

【０４７８】

ところで、上記の説明においては、ある構成アイテムに障害が発生したとき、当該構成アイテム自体が、障害の発生を通知するメッセージを送信するものと仮定している。

【０４７９】

しかし、実施形態によっては、ある構成アイテムに障害が発生したとき、他の構成アイテムが、前者の構成アイテムにおける障害の発生を通知するメッセージを出力してもよい。例えば、後者の構成アイテムは、前者の構成アイテムに障害が発生しているか否かを監視し、前者の構成アイテムにおける障害の発生に応じて、メッセージを出力してもよい。

【０４８０】

例えば、図８の例において、「Ｙ」というＩＰアドレスで識別される構成アイテムに時刻ｔ２４において障害が発生したとき、他のＩＰアドレス（便宜上「Ｙ２」とする）で識別される構成アイテムが、メッセージＭ２４と類似のメッセージを出力してもよい。出力されるメッセージには、障害が発生した構成アイテムを識別するための「Ｙ」というＩＰアドレスが含まれるものとする。なお、「Ｙ２」というＩＰアドレスで識別される構成アイテムから以上のようにして出力される当該メッセージの種別も、「３９」と分類される。

【０４８１】
この場合、トポロジ関係学習部７１１は、予兆パターンに含まれる各メッセージの発信元と、「Ｙ２」というＩＰアドレスで識別される構成アイテムとの間の関係を学習するのではないことに注意されたい。すなわち、この場合も、トポロジ関係学習部７１１は、予兆パターンに含まれる各メッセージの発信元と、「Ｙ」というＩＰアドレスで識別される構成アイテムとの間の関係を学習する。
【０４８２】
もちろん、第１実施形態に関して説明したように、ＩＰアドレスは識別情報の一例に過ぎない。実施形態によっては、ＩＰアドレス以外の識別情報が利用されてもよい。
【０４８３】
なお、検出サーバ４００は、図５のコンポーネントのうち、少なくともランキング生成部４０９を含んでいればよい。他のコンポーネントは、検出サーバ４００と通信可能な他のコンピュータ上に実装されていてもよい。例えば、障害予兆検知部４０２が他のコンピュータ上に実装されている場合、検出サーバ４００は、図１のステップＳ１に関して説明したような予測通知を受信することにより、障害が予測されたことを認識してもよい。
【０４８４】
同様に、検出サーバ７００は、図１０のコンポーネントのうち、少なくともランキング生成部７０９と推定部７１４を含んでいればよい。例えば、トポロジ関係学習部７１１が他のコンピュータ上に実装されている場合、検出サーバ７００の推定部７１４は、他のコンピュータのトポロジ関係学習部７１１により学習された関係情報を参照すればよい。
【０４８５】
ところで、検出サーバ４００と７００は、いずれも、以下のような構成要素を有する検出装置の具体例である。
【０４８６】
・図１のステップＳ１と同様に、障害の発生を予測するか、または、予測通知を受け取る予兆検知手段。
・図１のステップＳ２と同様に、統計値を算出する算出手段。
・図１のステップＳ３と同様に、結果情報を生成する生成手段。
・図１のステップＳ４と同様に、結果情報を出力する出力手段。
【０４８７】
例えば、障害予兆検知部４０２と７０２はいずれも、障害の発生を予測するタイプの予兆検知手段の例であり、ＣＰＵ１０１により実現され得る。予測通知を受け取るタイプの予兆検知手段の例は、例えば、通信インタフェイス１０３とＣＰＵ１０１の組み合わせである。
【０４８８】
また、検出サーバ４００におけるランキング生成部４０９は、算出手段の例でもあり、生成手段の例でもある。検出サーバ７００におけるランキング生成部７０９は、算出手段の例であり、検出サーバ７００における推定部７１４は、生成手段の例である。ある観点によれば、ログ統計算出部４０５および７０５、ならびに、予兆統計算出部４０７および７０７は、ＷＦ－ＩＤＦ（ｆ，ｎ）の算出に使うための情報を生成しているので、算出手段の一部を実現しているとも見なせる。いずれにせよ、算出手段は、例えばＣＰＵ１０１により実現されてもよい。
【０４８９】
また、出力手段の例としては、出力装置１０５や、通信インタフェイス１０３などが挙げられる。
【０４９０】
ところで、上記のとおり、第３実施形態では、図１２の処理は、何らかの障害が実際に発生したときに実行される。しかし、実施形態によっては、検出サーバ７００は、図１２と類似のバッチ処理により、関係情報を学習してもよい。
【０４９１】

例えば、ログ情報記憶部７０１には、今までに実際に発生したα個の障害についてのエントリが含まれており、障害予兆情報記憶部７０４には、それらα個の障害に関して障害予兆検知部７０２が正解したβ回の予兆検知についてのエントリが含まれているとする。なお、α個の障害の中には、正しく予測されなかった障害もあり得るし、１回の予測だけが正解した障害もあり得るし、２回以上の予測が正解した予測もあり得る。よって、α＜β、α＞β、およびα＝βのいずれの場合もあり得る。

【０４９２】

いずれにしろ、トポロジ関係学習部７１１は、１つの障害が発生するたびに図１２の処理を実行する代わりに、図１２と類似のバッチ処理を実行してもよい。つまり、１回のバッチ処理の実行により、トポロジ関係学習部７１１は、α個の障害（すなわち、発生したことがログ情報記憶部７０１に記録済みの、過去の複数の障害）のそれぞれについての関係情報を学習してもよい。

【０４９３】

最後に、上記の種々の実施形態に関して、さらに下記の付記を開示する。

（付記１）

コンピュータシステムを管理するコンピュータが、

各々が前記コンピュータシステムに含まれるハードウェア、ソフトウェア、または両者の組み合わせである複数の構成アイテムのうちのＱ個（１≦Ｑ）から所定時間以下の長さの期間に出力されるＰ個（１≦Ｑ≦Ｐ）のメッセージの組み合わせである第１のパターンに基づき、ある種別の障害の発生が予測される場合、前記Ｑ個の構成アイテムの各々について、前記ある種別の障害が過去に発生した発生時点より前に、前記Ｐ個のメッセージのうち当該構成アイテムが出力した出力メッセージと同じ種別のメッセージが出力された第１の頻度と、前記発生時点より前にいずれかのメッセージが出力された出力時点から前記所定時間だけ遡るウィンドウ期間中に前記出力メッセージと同じ種別のメッセージが出力され、かつ、前記ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせである第２のパターンに基づいて前記ある種別の障害の発生が予測された第２の頻度とに基づいて、前記ある種別の障害が当該構成アイテムで将来発生する蓋然性に関する統計値を算出し、

前記複数の構成アイテムの中で相対的に高い蓋然性で前記ある種別の障害が発生すると予測される１つ以上の構成アイテムを示す結果情報を、前記統計値に基づいて生成する

ことを特徴とする検出方法。

（付記２）

前記統計値は、前記第１の頻度に対して単調減少するとともに前記第２の頻度に対して単調増加する

ことを特徴とする付記１に記載の検出方法。

（付記３）

前記結果情報は、前記Ｑ個の構成アイテムのうちで前記統計値が最大の構成アイテムを識別する識別情報を含む

ことを特徴とする付記１または２に記載の検出方法。

（付記４）

前記結果情報を生成する処理が、

前記Ｐ個のメッセージの各々について、前記Ｐ個のメッセージのうちの当該メッセージと同じ種別のメッセージであって、前記ある種別の障害の発生が過去に正しく予測された際の予測に使われた前記第２のパターンに含まれるメッセージを出力した第１の構成アイテムと、過去に正しく予測された前記ある種別の障害が実際に発生した第２の構成アイテムとの間の第１の関係と等価な第２の関係が、前記Ｐ個のメッセージのうちの当該メッセージを出力した構成アイテムとの間に成り立つ関連構成アイテムを、前記複数の構成アイテム間の関係を示す構成情報を用いて、前記複数の構成アイテムの中から検索し、

前記Ｑ個の構成アイテムのうちのある構成アイテムについて前記関連構成アイテムが見つかった場合は、前記ある種別の障害が前記関連構成アイテムにおいて将来発生する蓋

然性に関する評価値を、前記Ｑ個の構成アイテムのうちの当該ある構成アイテムについて算出した前記統計値に基づいて決定し、
　検索の結果見つかった各関連構成アイテムについて決定した前記評価値に基づいて、前記結果情報を生成する
ことを含むことを特徴とする付記１から３のいずれか１項に記載の検出方法。
（付記５）
前記結果情報は、前記Ｑ個の構成アイテムの中の少なくとも１つに関して前記関連構成アイテムとして見つかった１つ以上の構成アイテムのうちで、前記評価値が最大の構成アイテムを識別する識別情報を含む
ことを特徴とする付記４に記載の検出方法。
（付記６）
前記構成情報により示される前記関係は、２つの構成アイテム間の論理的依存関係であるか、２つの構成アイテム間の物理的接続関係であるか、２つ以上の前記論理的依存関係の合成であるか、２つ以上の前記物理的接続関係の合成であるか、または、１つ以上の前記論理的依存関係と１つ以上の前記物理的接続関係の合成である
ことを特徴とする付記４または５に記載の検出方法。
（付記７）
前記コンピュータはさらに、
　前記複数の構成アイテムのいずれかからメッセージが出力されるたびに、当該メッセージの種別に対応づけられて記憶装置に記憶されたカウント値を更新し、
　前記第１の頻度を、前記カウント値から算出する
ことを特徴とする付記１から６のいずれか１項に記載の検出方法。
（付記８）
前記コンピュータはさらに、
　複数の種別のうちのいずれかの種別の障害が実際に発生するたびに、当該発生した障害を正しく予測する根拠となった前記第２のパターンに含まれる各メッセージの種別と、当該発生した障害の前記種別との組み合わせに対応づけられて記憶装置に記憶されたカウント値を更新し、
　前記第２の頻度を、前記カウント値から算出する
ことを特徴とする付記１から６のいずれか１項に記載の検出方法。
（付記９）
コンピュータシステムを管理するコンピュータに、
各々が前記コンピュータシステムに含まれるハードウェア、ソフトウェア、または両者の組み合わせである複数の構成アイテムのうちのＱ個（１≦Ｑ）から所定時間以下の長さの期間に出力されるＰ個（１≦Ｑ≦Ｐ）のメッセージの組み合わせである第１のパターンに基づき、ある種別の障害の発生が予測される場合、前記Ｑ個の構成アイテムの各々について、前記ある種別の障害が過去に発生した発生時点より前に、前記Ｐ個のメッセージのうち当該構成アイテムが出力した出力メッセージと同じ種別のメッセージが出力された第１の頻度と、前記発生時点より前にいずれかのメッセージが出力された出力時点から前記所定時間だけ遡るウィンドウ期間中に前記出力メッセージと同じ種別のメッセージが出力され、かつ、前記ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせである第２のパターンに基づいて前記ある種別の障害の発生が予測された第２の頻度とに基づいて、前記ある種別の障害が当該構成アイテムで将来発生する蓋然性に関する統計値を算出し、
前記複数の構成アイテムの中で相対的に高い蓋然性で前記ある種別の障害が発生すると予測される１つ以上の構成アイテムを示す結果情報を、前記統計値に基づいて生成する
ことを含む処理を実行させる検出プログラム。
（付記１０）
各々がコンピュータシステムに含まれるハードウェア、ソフトウェア、または両者の組み合わせである複数の構成アイテムのうちのＱ個（１≦Ｑ）から所定時間以下の長さの期

間に出力されるＰ個（１≦Ｑ≦Ｐ）のメッセージの組み合わせである第１のパターンに基づき、ある種別の障害の発生が予測される場合、前記Ｑ個の構成アイテムの各々について、前記ある種別の障害が過去に発生した発生時点より前に、前記Ｐ個のメッセージのうち当該構成アイテムが出力した出力メッセージと同じ種別のメッセージが出力された第１の頻度と、前記発生時点より前にいずれかのメッセージが出力された出力時点から前記所定時間だけ遡るウィンドウ期間中に前記出力メッセージと同じ種別のメッセージが出力され、かつ、前記ウィンドウ期間に含まれる１つ以上のメッセージの組み合わせである第２のパターンに基づいて前記ある種別の障害の発生が予測された第２の頻度とに基づいて、前記ある種別の障害が当該構成アイテムで将来発生する蓋然性に関する統計値を算出する算出手段と、

前記複数の構成アイテムの中で相対的に高い蓋然性で前記ある種別の障害が発生すると予測される１つ以上の構成アイテムを示す結果情報を、前記統計値に基づいて生成する生成手段と

を備える検出装置。

【符号の説明】

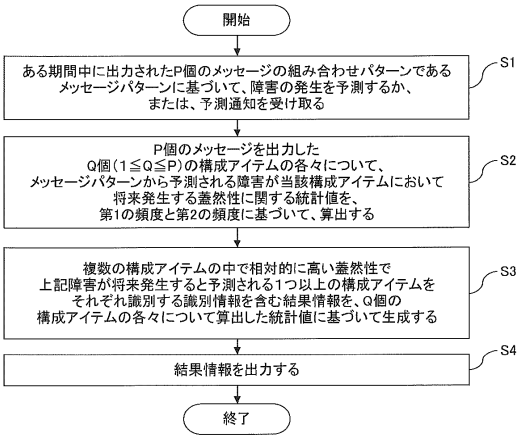
【０４９４】

１００、２００　コンピュータ
１０１　ＣＰＵ
１０２　ＲＡＭ
１０３　通信インタフェイス
１０４　入力装置
１０５　出力装置
１０６　記憶装置
１０７　駆動装置
１０８　バス
１１０　記憶媒体
１２０、２１０　ネットワーク
１３０　プログラム提供者
２３０　コンピュータシステム
２４０、２５０、２６０、２７０　物理サーバ
２４１、２５１、２６１、２７１　ハイパーバイザ
２４２、２５２、２６２、２７２　ホストＯＳ
２４３、２４４、２５３、２５４、２６３、２７３　ゲストＯＳ
２８０、２８１　Ｌ２スイッチ
２９０　Ｌ３スイッチ
３０１、３０３　ウィンドウ
３０２　予測対象期間
３０４　詳細予兆情報
３０５、６０４　ランキング情報
４００、７００　検出サーバ
４０１、７０１　ログ情報記憶部
４０２、７０２　障害予兆検知部
４０３、７０３　辞書情報記憶部
４０４、７０４　障害予兆情報記憶部
４０５、７０５　ログ統計算出部
４０６、７０６　ログ統計情報記憶部
４０７、７０７　予兆統計算出部
４０８、７０８　予兆統計情報記憶部
４０９、７０９　ランキング生成部
４１０、７１０　ランキング情報記憶部

４２０、７２０　メッセージ
４３０、７３０　推定結果情報
５０１　ログテーブル
５０２　メッセージ辞書テーブル
５０３　パターン辞書テーブル
５０４、８０１　障害予兆テーブル
５０５、８０２　ログ統計テーブル
５０６、８０３　予兆統計テーブル
５０７、８０５　ランキングテーブル
６０１　メッセージパターン
６０２、６０３　グラフ
６０５　改良ランキング情報
７１１　トポロジ関係学習部
７１２　構成情報記憶部
７１３　関係情報記憶部
７１４　推定部
８０４　トポロジ関係テーブル
８０６　改良ランキングテーブル
Ｍ１～Ｍ１１、Ｍ２１～Ｍ２４、Ｍ３１～Ｍ３３　メッセージ
ｔ１～ｔ１１、ｔ２１～ｔ２４　時刻
Ｎ１～Ｎ１７、Ｎ２１～Ｎ３０　ノード
Ｐ１～Ｐ３、Ｐ１１～Ｐ１３　パス
Ａ～Ｅ、Ｗ～Ｚ　ＩＰアドレス

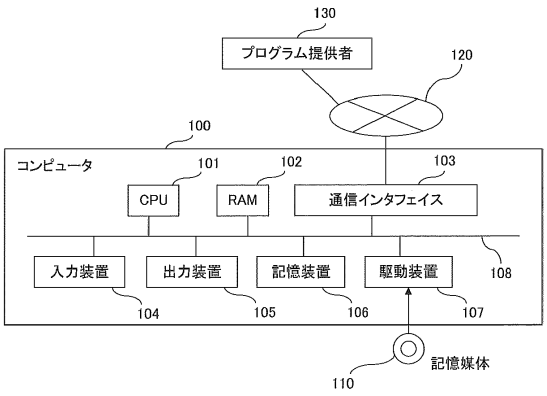
【図１】

第1実施形態のコンピュータが実行する処理のフローチャート



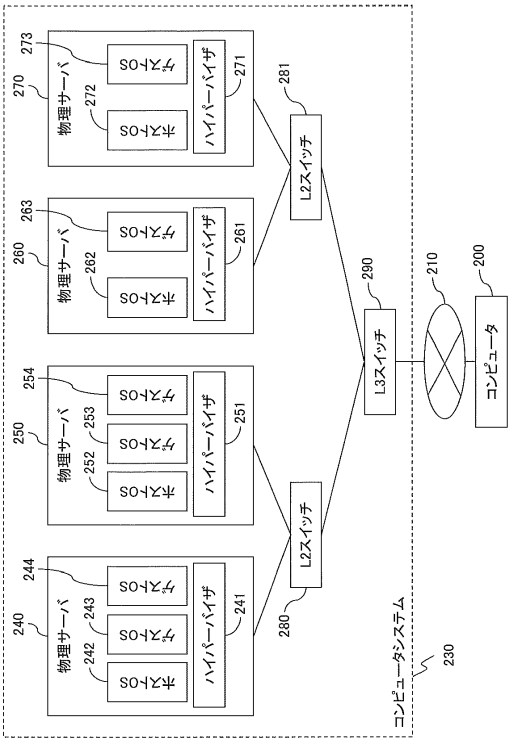
【図２】

コンピュータのハードウェア構成図



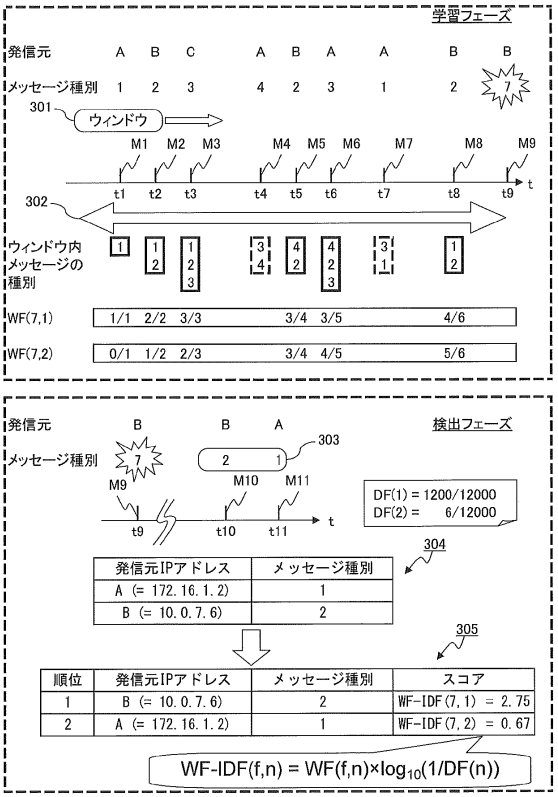
【図３】

コンピュータシステムの例を示す図



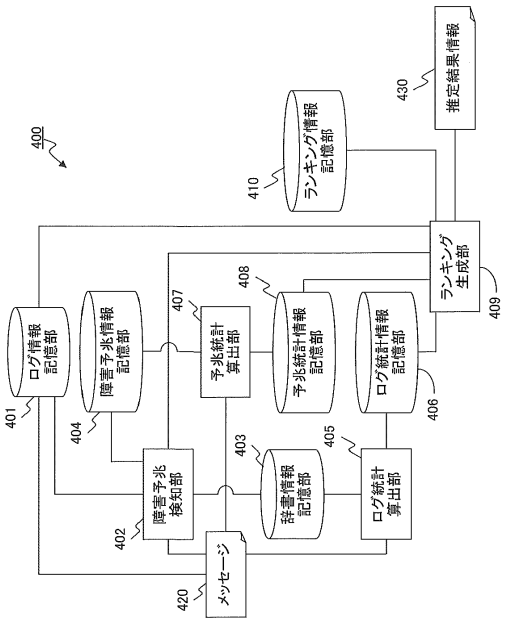
【図４】

第2実施形態の検出サーバの動作を例示する図



【図５】

第2実施形態の検出サーバのブロック構成図



【図６】

第2実施形態で利用される各種テーブルの例を示す図

501

時刻	IPアドレス	メッセージ文字列	メッセージ種別
2012/7/31 23:42:00	10.0.7.6	Permission denied	2
2012/7/31 23:42:15	10.0.0.1	Device: /dev/sda, opened	5

502

メッセージ種別	メッセージ文字列
1	Restoring network configuration
2	Permission denied

503

障害種別	予兆パターン	スコア
7	1,2	0.95
7	2,3,4	0.83

504

ID	障害種別	予兆パターン	予測実行時刻	開始時刻
1	7	1,2	2012/7/31 23:00:00	2012/8/1 00:00:00

505

ID	メッセージ種別	カウント
1	1	1200
2	2	6
3	3	2
4	4	500
5	*	12000

506

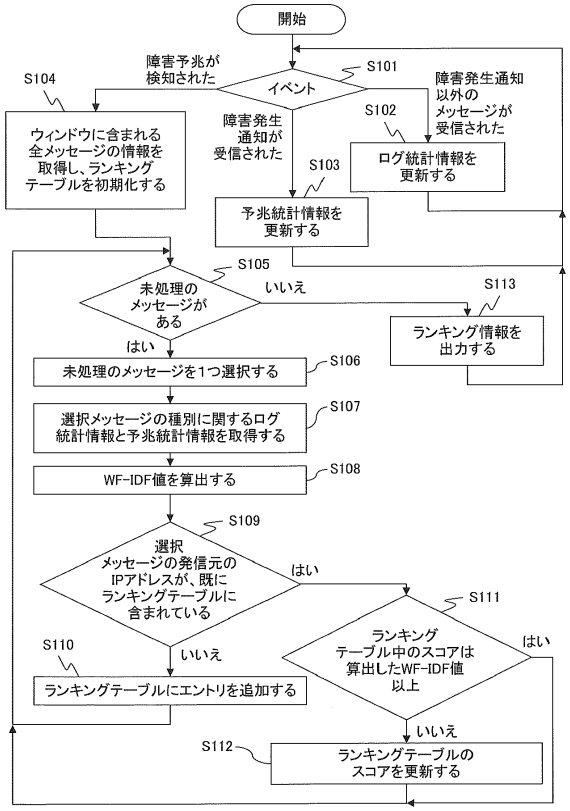
ID	障害種別	メッセージ種別	カウント
1	7	1	4
2	7	2	5
3	7	3	2
4	7	4	2
5	7	*	6

507

予兆ID	順位	IPアドレス	メッセージ種別	スコア
1	1	10.0.7.6	2	2.7509
1	2	172.16.1.2	1	0.6667

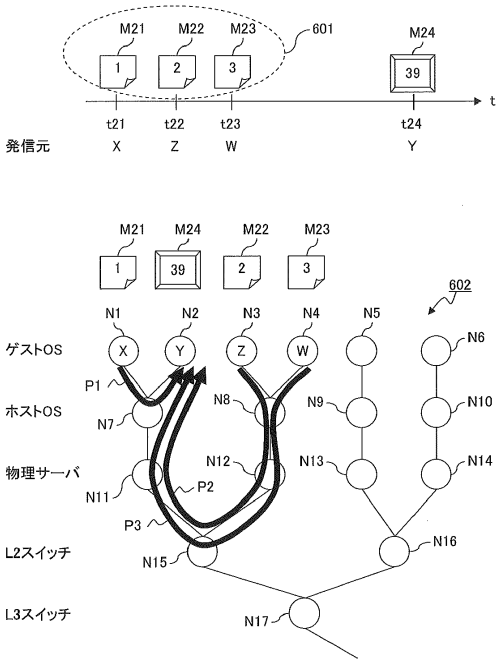
【図７】

第2実施形態の検出サーバが行う処理のフローチャート



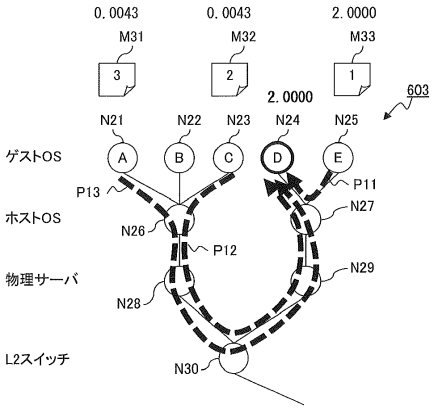
【図８】

第3実施形態における関係情報の学習を説明する図



【図９】

第3実施形態におけるランキングの改良について説明する図



順位	発信元IPアドレス	メッセージ種別	スコア
1	E (= 10.0.0.3)	1	2.0000
2	C (= 10.0.0.1)	2	0.0043
2	A (= 172.16.1.2)	3	0.0043

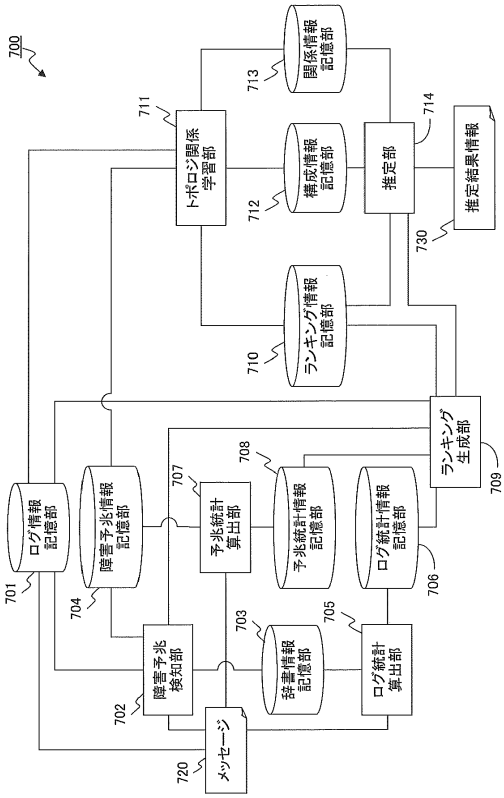
604

605

順位	発信元IPアドレス	メッセージ種別	スコア
1	D (= 10.0.0.10)	1	2.0000
2	E (= 10.0.0.3)	2, 3	0.0043

【図１０】

第3実施形態の検出サーバのブロック構成図



【図１１】

第3実施形態で利用される各種テーブルの例を示す図

801

ID	障害種別	予兆パターン	予測実行時刻	開始時刻
1	39	1, 2, 3	2012/8/31 23:00:00	2012/9/1 00:00:00
2	39	1, 2, 3	2012/9/3 9:15:00	2012/9/3 10:15:00

802

ID	メッセージ種別	カウント
1	1	100
2	2	9901
3	3	9901
4	*	10000

803

ID	障害種別	メッセージ種別	カウント
1	39	1	1
2	39	2	1
3	39	3	1
4	39	*	1

804

予兆ID	ID	障害種別	メッセージ種別	パス
1	1	39	1	%LogicalServer/&*/%LogicalServer/&*/%LogicalServer
1	2	39	2	%LogicalServer/&*/%LogicalServer/&*/%Server/&*/%NetworkDevice/%Server/&*/%LogicalServer/&*/%LogicalServer
1	3	39	3	%LogicalServer/&*/%LogicalServer/&*/%Server/&*/%NetworkDevice/%Server/&*/%LogicalServer/&*/%LogicalServer

805

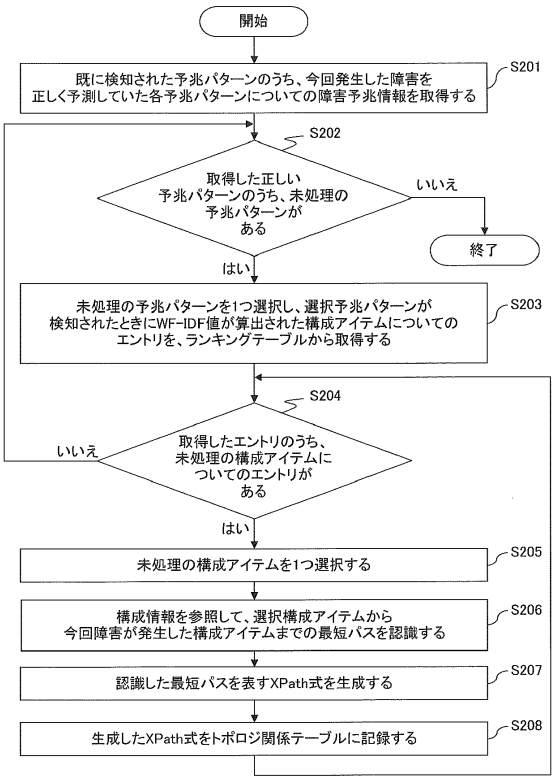
予兆ID	順位	IPアドレス	メッセージ種別	スコア
2	1	E (= 10.0.0.3)	1	2.0000
2	2	C (= 10.0.0.1)	2	0.0043
2	2	A (= 172.16.1.2)	3	0.0043

806

予兆ID	順位	IPアドレス	メッセージ種別	スコア
2	1	D (= 10.0.0.10)	1	2.0000
2	2	E (= 10.0.0.3)	2, 3	0.0043

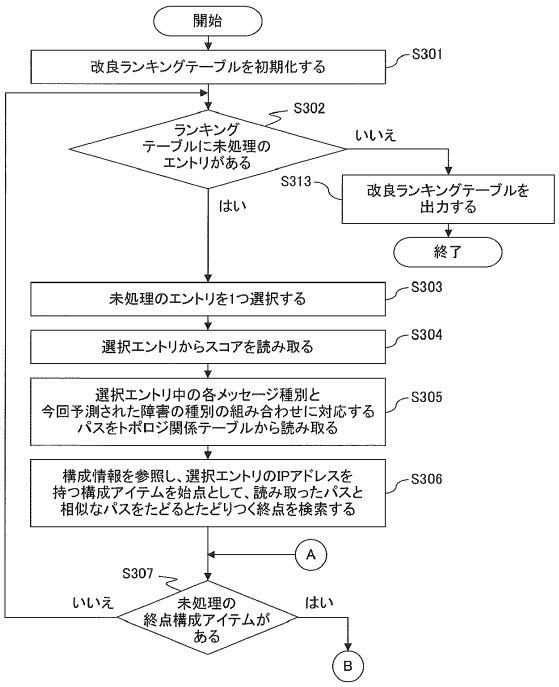
【図１２】

第3実施形態において検出サーバが
関係情報を学習する処理のフローチャート



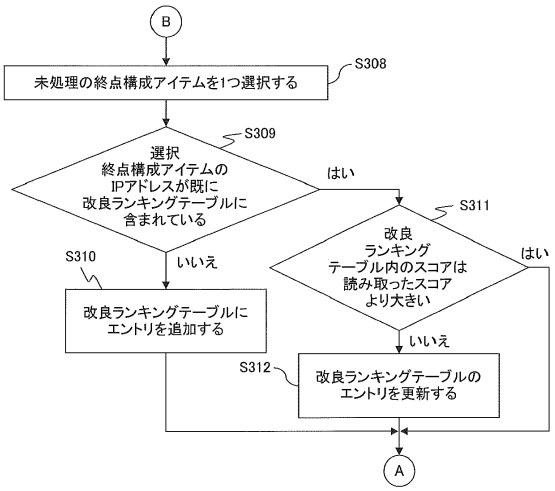
【図１３】

第3実施形態の検出サーバが、学習した関係情報を使って
改良ランキング情報を生成する処理のフローチャート（その1）



【図１４】

第3実施形態の検出サーバが、学習した関係情報を使って
改良ランキング情報を生成する処理のフローチャート（その2）



フロントページの続き

(72)発明者　渡辺　幸洋
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内
(72)発明者　松本　安英
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

審査官　大塚　俊範

(56)参考文献　特開２０１１－００２９０６（ＪＰ，Ａ）
特開２００７－１７２１３１（ＪＰ，Ａ）
国際公開第２０１２／１２７５８８（WO，Ａ１）
国際公開第２０１０／１３１７４６（WO，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　１１／０７
Ｇ０６Ｆ　１１／３０－１１／３４