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

特許第6152788号
(P6152788)

(45)発行日　平成29年6月28日(2017.6.28)　(24)登録日　平成29年6月9日(2017.6.9)

(51)Int.Cl.		F I		
G06F 11/07	(2006.01)	G06F	11/07	151
G06F 11/30	(2006.01)	G06F	11/07	140V
		G06F	11/30	172

請求項の数 6　(全 26 頁)

(21)出願番号　特願2013-249027(P2013-249027)
(22)出願日　平成25年12月2日(2013.12.2)
(65)公開番号　特開2015-106334(P2015-106334A)
(43)公開日　平成27年6月8日(2015.6.8)
審査請求日　平成28年8月4日(2016.8.4)

(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　100092152
弁理士　服部　毅巖
(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　多賀　実

最終頁に続く

(54)【発明の名称】障害予兆検知方法、情報処理装置およびプログラム

(57)【特許請求の範囲】

【請求項１】

監視対象のシステムから複数の種類のメッセージを収集するコンピュータが実行する障害予兆検知方法であって、

第１のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、学習に使用しない第１のメッセージの種類を判定し、

前記第１のメッセージの集合および前記システムの障害発生を示す障害情報から、障害が発生するときに現れるメッセージのパターンであって前記第１のメッセージの種類を除外した第１のメッセージのパターンを学習し、

前記第１のメッセージの集合より後に第２のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、検知に使用しない第２のメッセージの種類を判定し、

前記第２のメッセージの集合から、前記第２のメッセージの種類を除外した第２のメッセージのパターンを生成し、前記第１のメッセージのパターンと前記第２のメッセージのパターンとを比較することで前記システムの障害の予兆を検知する、

障害予兆検知方法。

【請求項２】

前記第１のメッセージの種類は、複数のメッセージの種類のうち前記第１のメッセージの集合を取得したときにおける出現頻度が閾値以上であるメッセージの種類であり、

前記第２のメッセージの種類は、前記複数のメッセージの種類のうち前記第２のメッセ

ージの集合を取得したときにおける出現頻度が閾値以上であるメッセージの種類である、
請求項１記載の障害予兆検知方法。
【請求項３】
前記システムの動作状態が変化したことを検出し、前記システムの動作状態の変化を契機として前記メッセージの種類毎の出現頻度を再計算する、
請求項１または２記載の障害予兆検知方法。
【請求項４】
前記第１のメッセージの種類が除外された前記第１のメッセージのパターンと、前記第２のメッセージの種類が除外された前記第２のメッセージのパターンとが一致するとき、前記システムの障害の予兆があると判定する、
請求項１乃至３の何れか一項に記載の障害予兆検知方法。
【請求項５】
監視対象のシステムから収集した複数の種類のメッセージと、前記システムの障害発生を示す障害情報とを記憶する記憶部と、
第１のメッセージの集合および前記障害情報から、障害が発生するときに現れる第１のメッセージのパターンを学習し、前記第１のメッセージの集合より後に取得した第２のメッセージの集合から第２のメッセージのパターンを生成し、前記第１のメッセージのパターンと前記第２のメッセージのパターンとを比較することで前記システムの障害の予兆を検知する演算部と、
を有し、前記演算部は、
前記第１のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、学習に使用しない第１のメッセージの種類を判定し、前記第１のメッセージのパターンから前記第１のメッセージの種類を除外し、
前記第２のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、検知に使用しない第２のメッセージの種類を判定し、前記第２のメッセージのパターンから前記第２のメッセージの種類を除外する、
情報処理装置。
【請求項６】
監視対象のシステムから複数の種類のメッセージを収集するコンピュータに、
第１のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、学習に使用しない第１のメッセージの種類を判定し、
前記第１のメッセージの集合および前記システムの障害発生を示す障害情報から、障害が発生するときに現れるメッセージのパターンであって前記第１のメッセージの種類を除外した第１のメッセージのパターンを学習し、
前記第１のメッセージの集合より後に第２のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、検知に使用しない第２のメッセージの種類を判定し、
前記第２のメッセージの集合から、前記第２のメッセージの種類を除外した第２のメッセージのパターンを生成し、前記第１のメッセージのパターンと前記第２のメッセージのパターンとを比較することで前記システムの障害の予兆を検知する、
処理を実行させるプログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は障害予兆検知方法、情報処理装置およびプログラムに関する。
【背景技術】
【０００２】
現在、サーバ装置やストレージや通信装置などの様々な電子機器を含む情報処理システムが利用されている。このような情報処理システムでは、ＨＤＤ（Hard Disk Drive）の故障や通信インタフェースの故障などの障害が発生することがある。そこで、監視装置が

電子機器から各種のメッセージを収集し、情報処理システムの稼働状態を監視することが行われている。例えば、監視装置は、収集したメッセージから障害を検知すると、使用するサーバ装置の切り替えや通信経路の変更を管理者に促すことが考えられる。

【０００３】

監視装置の中には、障害が発生する前に、収集したメッセージに基づいて障害の予兆を検知するものもある。例えば、監視装置は、ＨＤＤへの書き込み失敗の増加や通信遅延の急激な増大を検知すると、障害の予兆として管理者に通知することが考えられる。障害が発生する前に使用するサーバ装置の切り替えや通信経路の変更などの対策をとることができれば、情報処理の停止時間を短縮して障害の影響を軽減できる。

【０００４】

一例として、プラントなどの設備から収集するデータに基づいて障害の予兆を検知する設備状態監視方法が提案されている。この設備状態監視方法は、設備の正常状態を示す正常モデルを生成する学習フェーズと、正常モデルおよび設備から収集したデータに基づいて障害の予兆を検知する評価フェーズを含む。学習フェーズでは、正常時のデータから正常モデルとして特徴ベクトルを生成する。評価フェーズでは、現在収集したデータから特徴ベクトルを生成して正常モデルと比較する。特徴ベクトルの距離に応じた「異常測度」が閾値以上である場合、設備に障害の予兆があると判定する。

【先行技術文献】

【特許文献】

【０００５】

【特許文献１】特開２０１１－７０６３５号公報

【発明の概要】

【発明が解決しようとする課題】

【０００６】

障害の予兆を検知する方法としては、過去に障害が発生したときに現れたメッセージのパターンを学習しておき、学習したメッセージのパターンが収集したメッセージの集合の中に現れたときに、障害の予兆があると判定する方法が考えられる。学習するメッセージのパターンは、例えば、障害発生から所定時間前までに現れる確率が高いメッセージの種類の組み合わせとする。しかし、この検知方法には次のような問題がある。

【０００７】

監視対象の情報処理システムから収集されるメッセージの中には、障害との関連性が低く継続的に発生するメッセージがノイズとして含まれていることがある。例えば、使用していない通信インタフェースに対する監視機能がＯＮになっていることで発生するメッセージなど、管理者が無視できるような軽度の注意情報を含むメッセージが定期的に発生することがある。ノイズとして収集されるメッセージの種類は、情報処理システムの構成変更や情報処理システムを利用した業務プロセスの変更など、情報処理システムの動作状態が変化したときに変わる可能性がある。例えば、使用していない通信インタフェースに対する監視機能をＯＮからＯＦＦにするとノイズが削減される。

【０００８】

収集されるメッセージの中に多くのノイズが含まれている場合、障害の予兆を示すメッセージのパターンの学習結果の中にも、ノイズが混入することになる。この場合、ノイズとして継続的に発生するメッセージの種類が学習時点から変化してしまうと、学習結果と同じメッセージのパターンが収集したメッセージの中に現れなくなり、既存の学習結果を用いて障害予兆を検知することができなくなるという問題がある。これに対しては、既存の学習結果を破棄してメッセージのパターンを再学習することも考えられる。しかし、情報処理システムの動作状態が変化する毎にメッセージのパターンを再学習することは、再学習の負荷が大きく、また、障害予兆検知の精度が低下するという問題がある。

【０００９】

１つの側面では、本発明は、収集するメッセージに含まれるノイズが変化しても既存の学習結果を活用することができる障害予兆検知方法、情報処理装置およびプログラムを提

供することを目的とする。
【課題を解決するための手段】
【００１０】
１つの態様では、監視対象のシステムから複数の種類のメッセージを収集するコンピュータが実行する障害予兆検知方法が提供される。障害予兆検知方法では、第１のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、学習に使用しない第１のメッセージの種類を判定する。第１のメッセージの集合およびシステムの障害発生を示す障害情報から、障害が発生するときに現れるメッセージのパターンであって第１のメッセージの種類を除外した第１のメッセージのパターンを学習する。第１のメッセージの集合より後に第２のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、検知に使用しない第２のメッセージの種類を判定する。第２のメッセージの集合から、第２のメッセージの種類を除外した第２のメッセージのパターンを生成し、第１のメッセージのパターンと第２のメッセージのパターンとを比較することでシステムの障害の予兆を検知する。
【００１１】
また、１つの態様では、記憶部と演算部とを有する情報処理装置が提供される。記憶部は、監視対象のシステムから収集した複数の種類のメッセージと、システムの障害発生を示す障害情報とを記憶する。演算部は、第１のメッセージの集合および障害情報から、障害が発生するときに現れる第１のメッセージのパターンを学習し、第１のメッセージの集合より後に取得した第２のメッセージの集合から第２のメッセージのパターンを生成し、第１のメッセージのパターンと第２のメッセージのパターンとを比較することでシステムの障害の予兆を検知する。演算部は、第１のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、学習に使用しない第１のメッセージの種類を判定し、第１のメッセージのパターンから第１のメッセージの種類を除外する。また、演算部は、第２のメッセージの集合を取得したときにおけるメッセージの種類毎の出現頻度に基づいて、検知に使用しない第２のメッセージの種類を判定し、第２のメッセージのパターンから第２のメッセージの種類を除外する。
【００１２】
また、１つの態様では、監視対象のシステムから複数の種類のメッセージを収集するコンピュータに実行させるプログラムが提供される。
【発明の効果】
【００１３】
１つの側面では、収集するメッセージに含まれるノイズが変化しても既存の学習結果を活用することができる。
【図面の簡単な説明】
【００１４】
【図１】第１の実施の形態の情報処理装置を示す図である。
【図２】第２の実施の形態の情報処理システムを示す図である。
【図３】メッセージパターンの学習例を示す図である。
【図４】予兆検知における背景ノイズの影響例を示す図である。
【図５】予兆検知における背景ノイズの除外例を示す図である。
【図６】監視サーバのハードウェア例を示すブロック図である。
【図７】監視サーバの機能例を示すブロック図である。
【図８】メッセージテーブルの例を示す図である。
【図９】頻度テーブルの例を示す図である。
【図１０】ユーザ設定テーブルの例を示す図である。
【図１１】障害テーブルの例を示す図である。
【図１２】学習テーブルの例を示す図である。
【図１３】頻度算出の手順例を示すフローチャートである。
【図１４】パターン学習の手順例を示すフローチャートである。

【図１５】予兆検知の手順例を示すフローチャートである。
【図１６】監視サーバの他の機能例を示すブロック図である。
【図１７】期間テーブルの例を示す図である。
【図１８】頻度算出の他の手順例を示すフローチャートである。
【図１９】パターン学習の他の手順例を示すフローチャートである。
【発明を実施するための形態】
【００１５】
以下、本実施の形態を図面を参照して説明する。
［第１の実施の形態］
図１は、第１の実施の形態の情報処理装置を示す図である。
【００１６】
第１の実施の形態の情報処理装置１０は、監視対象のシステムから複数の種類のメッセージを収集し、収集したメッセージに基づいてシステムの障害の予兆を検知する。監視対象のシステムは、１または２以上の電子機器を有し、サーバ装置やストレージや通信装置などの複数の種類の電子機器を有していてもよい。情報処理装置１０が取得するメッセージの集合には、２以上または２種類以上の電子機器からのメッセージが混在していてもよい。情報処理装置１０は、コンピュータと呼ばれてもよい。情報処理装置１０は、サーバ装置（例えば、サーバコンピュータと呼ばれるもの）であってもよいし、ユーザが操作する端末装置（例えば、クライアントコンピュータと呼ばれるもの）であってもよい。
【００１７】
情報処理装置１０は、記憶部１１および演算部１２を有する。記憶部１１は、ＲＡＭ（Random Access Memory）などの揮発性の記憶装置でもよいし、ＨＤＤなどの不揮発性の記憶装置でもよい。演算部１２は、例えば、プロセッサである。プロセッサは、ＣＰＵ（Central Processing Unit）やＤＳＰ（Digital Signal Processor）であってもよく、ＡＳＩＣ（Application Specific Integrated Circuit）やＦＰＧＡ（Field Programmable Gate Array）などの特定用途の集積回路を含んでもよい。プロセッサは、ＲＡＭなどの記憶装置（例えば、記憶部１１）に記憶されたプログラムを実行する。２以上のプロセッサの集合（マルチプロセッサ）を「プロセッサ」と呼んでもよい。
【００１８】
記憶部１１は、メッセージの集合１３ａ，１３ｂおよび障害情報１４を記憶する。
メッセージの集合１３ａは、ある時点において監視対象のシステムから収集されたメッセージの集合である。メッセージの集合１３ｂは、メッセージの集合１３ａより後の時点において監視対象のシステムから収集されたメッセージの集合である。メッセージの集合１３ａは、メッセージの集合１３ｂが取得された時点で記憶部１１から削除されていてもよいし、削除されていなくてもよい。後者の場合、メッセージの集合１３ｂが取得された時点で、メッセージの集合１３ａはログ情報と見ることができる。
【００１９】
メッセージの集合１３ａ，１３ｂの中には、障害発生を示すメッセージではないが、電子機器の好ましくない動作の発生を示すメッセージが含まれる。電子機器の好ましくない動作としては、例えば、ＨＤＤへのアクセス失敗、キャッシュのオーバーフロー、通信遅延、インタフェースの初期化失敗などが挙げられる。同時期に特定の２種類以上のメッセージが発生した場合に、その後に高い確率で障害が発生することがある。ただし、メッセージの集合１３ａ，１３ｂの中には、障害との関連性が低く継続的に発生するメッセージも含まれる。このようなメッセージはノイズと言うことができる。
【００２０】
ノイズとしてのメッセージの種類は、監視対象のシステムの動作状態が変化することで変わる可能性がある。システムの動作状態の変化としては、例えば、システムの構成・設定の変更や、システムを利用した業務プロセスの変更などが挙げられる。第１の実施の形態では、メッセージの集合１３ａとメッセージの集合１３ｂとには、ノイズとして異なる種類のメッセージが含まれている。図１の例では、メッセージの集合１３ａに種類Ａ，Ｂ

，Ｘのメッセージが含まれており、メッセージの集合１３ｂに種類Ａ，Ｂ，Ｙのメッセージが含まれている。種類Ｘ，Ｙのメッセージがノイズに相当する。なお、以下では、種類Ａ，Ｂ，Ｘ，Ｙのメッセージを、メッセージＡ，Ｂ，Ｘ，Ｙと言うことがある。

【００２１】

障害情報１４は、監視対象のシステムで過去に発生した障害を示し、例えば、障害発生の時刻を示す情報を含む。システムの障害としては、例えば、ＨＤＤの故障や通信インタフェースの故障などのハードウェア障害が挙げられる。障害情報１４には、少なくとも、メッセージの集合１３ａが取得された時期に発生した障害についての情報が含まれる。障害情報１４は、ユーザが情報処理装置１０に入力してもよいし、監視対象のシステムから収集された障害発生を示すメッセージに基づいて情報処理装置１０が生成してもよい。

【００２２】

演算部１２は、メッセージの集合１３ａおよび障害情報１４から、障害が発生するときに現れるメッセージのパターン１５ａを学習する。メッセージのパターン１５ａは、例えば、過去に障害発生から所定時間前までに現れた２種類以上のメッセージの組み合わせを示す。また、演算部１２は、メッセージの集合１３ｂからメッセージのパターン１５ｂを生成する。メッセージのパターン１５ｂは、例えば、同時期に現れた２種類以上のメッセージの組み合わせを示す。そして、演算部１２は、メッセージのパターン１５ａとメッセージのパターン１５ｂとを比較することで、障害の予兆を検知する。例えば、演算部１２は、メッセージのパターン１５ｂがメッセージのパターン１５ａと一致するとき、監視対象のシステムに障害の予兆があると判断してユーザに警告する。

【００２３】

ここで、演算部１２は、メッセージのパターン１５ａを学習するにあたり、メッセージの集合１３ａが取得されたときにおけるメッセージの種類毎の出現頻度を算出する。例えば、演算部１２は、メッセージの集合１３ａに含まれるメッセージを種類毎にカウントして、メッセージの種類毎の出現確率を算出する。そして、演算部１２は、メッセージの集合１３ａに対応する出現頻度に基づいて、複数のメッセージの種類のうち学習に使用しないメッセージの種類（例えば、メッセージＸ）を判定する。学習に使用しないメッセージの種類は、例えば、出現頻度が閾値以上であるものとする。演算部１２は、メッセージのパターン１５ａから、判定した種類のメッセージ（例えば、メッセージＸ）を除外する。

【００２４】

また、演算部１２は、障害の予兆を検知するにあたり、メッセージの集合１３ｂが取得されたときにおけるメッセージの種類毎の出現頻度を算出する。例えば、演算部１２は、メッセージの集合１３ｂに含まれるメッセージを種類毎にカウントして、メッセージの種類毎の出現確率を算出する。そして、演算部１２は、メッセージの集合１３ｂに対応する出現頻度に基づいて、複数のメッセージの種類のうち検知に使用しないメッセージの種類（例えば、メッセージＹ）を判定する。検知に使用しないメッセージの種類は、例えば、出現頻度が閾値以上であるものとする。演算部１２は、メッセージのパターン１５ｂから、判定した種類のメッセージ（例えば、メッセージＹ）を除外する。

【００２５】

メッセージの集合１３ａにはノイズとしてメッセージＸが多数含まれているため、メッセージＸを除外しない場合、学習時にはメッセージＡ，Ｂ，Ｘを含むメッセージのパターンが生成される可能性が高い。また、メッセージの集合１３ｂにはノイズとしてメッセージＹが多数含まれているため、メッセージＹを除外しない場合、検知時にはメッセージＡ，Ｂ，Ｙを含むメッセージのパターンが生成される可能性が高い。この場合、２つのメッセージのパターンを単純に比較するだけでは障害の予兆を検知することが難しい。一方、上記のように生成されたメッセージのパターン１５ａ，１５ｂは、メッセージＡ，Ｂを含みメッセージＸ，Ｙを含まないため、両者の比較によって障害の予兆を検知できる。

【００２６】

第１の実施の形態の情報処理装置１０によれば、学習時のメッセージの種類毎の出現頻度に基づいて学習に使用しないメッセージの種類が判定され、判定された種類のメッセー

ジを除外したメッセージのパターン１５ａが学習される。また、検知時のメッセージの種類毎の出現頻度に基づいて検知に使用しないメッセージの種類が判定され、判定された種類のメッセージを除外したメッセージのパターン１５ｂが学習結果と比較される。これにより、システムの構成変更や業務プロセスの変更などに応じてメッセージのノイズが変化しても、既存の学習結果を利用して障害の予兆を検知することができる。その結果、再学習の負荷を抑制でき、また、障害予兆検知の精度を高めることができる。

【００２７】

［第２の実施の形態］

図２は、第２の実施の形態の情報処理システムを示す図である。

第２の実施の形態の情報処理システムは、業務で使用される各種の電子機器を集中的に管理する。この情報処理システムは、業務サーバ２１、ストレージ２２、通信装置２３、クライアント２４、管理サーバ２５および監視サーバ１００を有する。情報処理システムに含まれるこれらの装置は、ネットワーク２０に接続されている。なお、監視サーバ１００は、第１の実施の形態の情報処理装置１０の一例である。業務サーバ２１、ストレージ２２および通信装置２３の集合は、監視対象のシステムの一例である。

【００２８】

業務サーバ２１、ストレージ２２および通信装置２３は、業務で使用される電子機器の一例である。業務サーバ２１は、業務用のアプリケーションソフトウェアを実行するサーバコンピュータである。ストレージ２２は、業務に使用するデータを、磁気ディスクなどの不揮発性の記憶媒体に記憶しておく記憶装置である。ストレージ２２は、業務サーバ２１からアクセスされ得る。通信装置２３は、データを転送するルータやスイッチなどである。監視対象のシステムには、他の種類の電子機器が含まれていてもよい。

【００２９】

クライアント２４は、管理者が操作する端末装置としてのクライアントコンピュータである。監視対象のシステムの構成を変更するとき、クライアント２４は、構成変更の内容およびそのスケジュールを管理サーバ２５に登録する。構成変更の例としては、業務サーバの追加や削除、業務サーバ間での仮想マシンの移動、業務サーバ２１にインストールされたソフトウェアの更新、通信装置２３の通信ポートの設定変更などが挙げられる。

【００３０】

また、クライアント２４は、監視対象のシステムに障害が発生すると、障害を示す警告情報を監視サーバ１００から受信する。障害の例としては、業務サーバ２１やストレージ２２が備えるＨＤＤの故障、通信装置２３が備える通信ポートの故障などが挙げられる。障害を示す警告情報は、クライアント２４のディスプレイに表示される。管理者は、クライアント２４を用いて、システムを復旧する操作を行うことができる。例えば、管理者は、業務サーバ２１やストレージ２２を、予備の業務サーバやストレージに切り替える。

【００３１】

また、クライアント２４は、障害はまだ発生していないが障害の予兆があるとき、障害の予兆を示す警告情報を監視サーバ１００から受信することがある。障害の予兆の例としては、ＨＤＤへのアクセス失敗、キャッシュのオーバーフロー、通信遅延、インタフェースの初期化失敗などの好ましくない動作の組み合わせであって、所定の条件を満たす複数種類の動作の組み合わせが考えられる。障害の予兆を示す警告情報は、クライアント２４のディスプレイに表示される。管理者は、クライアント２４を用いて、障害発生前に障害の影響を小さくするための操作を行うことができる。例えば、業務サーバ２１に障害の予兆がある場合、管理者は、障害発生前に、仮想マシンを業務サーバ２１から他の業務サーバへ移動させる。また、例えば、通信装置２３に障害の予兆がある場合、管理者は、障害発生前に、通信装置２３を通過しないように通信経路の設定を変更する。

【００３２】

管理サーバ２５は、クライアント２４からの指示に応じて、監視対象のシステムの構成を変更するサーバコンピュータである。管理サーバ２５は、クライアント２４から構成変更の内容およびスケジュールが登録されると、登録された構成変更をスケジュールに従っ

て実行する。例えば、管理サーバ２５は、指定された日時に、業務サーバ２１にインストールされたソフトウェアの更新や通信装置２３の通信ポートの設定変更などを行う。

【００３３】

監視サーバ１００は、システムに障害または障害の予兆がないか監視するサーバコンピュータである。監視サーバ１００は、監視対象のシステムに属する各電子機器から継続的にメッセージを収集する。メッセージの収集には、ＳＮＭＰ（Simple Network Management Protocol）を含む任意のプロトコルを用いることができる。収集されるメッセージの集合には、ＨＤＤの故障や通信ポートの故障などの障害を示すメッセージが含まれ得る。また、収集されるメッセージの集合には、ＨＤＤへのアクセス失敗、キャッシュのオーバーフロー、通信遅延、インタフェースの初期化失敗など、障害ではないが好ましくない動作を示す注意喚起のメッセージが含まれ得る。

【００３４】

収集されたメッセージに基づいて、監視サーバ１００は、障害または障害の予兆を検知する。障害を検知すると、監視サーバ１００は、障害の種類・障害が発生した電子機器・発生時刻などを示す警告情報を生成し、クライアント２４に送信する。注意喚起のメッセージに基づいて障害の予兆を検知すると、監視サーバ１００は、予兆のある障害の種類・予兆のある電子機器・予兆の検知に用いられたメッセージ・検知時刻などを示す警告情報を生成し、クライアント２４に送信する。監視サーバ１００からクライアント２４への警告情報の送信には、電子メールを含む任意のプロトコルを用いることができる。

【００３５】

障害の予兆を検知するために、監視サーバ１００は、過去に収集されたメッセージの集合に基づいて、障害発生の所定時間前までに現れる確率の高いメッセージの種類の組み合わせを学習する。監視サーバ１００は、現在収集されたメッセージと学習結果とをリアルタイムに比較し、学習結果に合致するメッセージの系列が現れたとき障害の予兆があると判断する。以下、監視サーバ１００が行う障害予兆検知を中心に説明する。

【００３６】

図３は、メッセージパターンの学習例を示す図である。

第２の実施の形態では、同時期に現れる２種類以上のメッセージの組み合わせをメッセージのパターンとして扱う。監視サーバ１００は、収集されたメッセージの集合を用いて、障害発生と相関の高いメッセージのパターンを学習する。これにより、監視サーバ１００は、人手では発見が容易でないメッセージと障害との関係を発見することができる。

【００３７】

監視サーバ１００は、各メッセージに受信時刻の情報を付与することで、収集したメッセージを時系列に管理する。図３に示すように、監視サーバ１００は、一定の時間幅（例えば、５分間）のスライディングウィンドウを時間軸に沿ってシフトさせる。スライディングウィンドウに含まれるメッセージの種類の組み合わせが、同時期に現れたメッセージのパターンとして抽出される。このとき、メッセージのパターンにおいては、同じ種類のメッセージの数やメッセージの出現順序は考慮されない。すなわち、メッセージのパターンでは、同時期に現れたメッセージの種類が順不動で列挙されることになる。

【００３８】

例えば、ある時点でスライディングウィンドウに種類４，９，７，１のメッセージ（以下ではメッセージ４，９，７，１と表記することがある）が含まれる場合、［１，４，７，９］というパターンが抽出される。その後、スライディングウィンドウにメッセージ３が追加されると、［１，３，４，７，９］というパターンが抽出される。更にその後、スライディングウィンドウにメッセージ１０が追加されスライディングウィンドウからメッセージ４，９，７，１が追い出されると、［３，１０］というパターンが抽出される。

【００３９】

このようなメッセージのパターンは、学習時に抽出されると共に、予兆検知時に現在収集したメッセージの集合からリアルタイムに抽出される。学習時には、監視サーバ１００は、障害発生から所定時間前までに現れた回数をパターン毎にカウントすることで、パタ

ーンと障害発生の相関を学習する。あるパターンが障害発生から所定時間前までに現れるとは、例えば、スライディングウィンドウ内で末尾にある（最も新しい）メッセージの受信時刻またはスライディングウィンドウの末尾の時刻と、障害発生時刻との差が閾値以下であることである。予兆検知時には、監視サーバ１００は、障害発生との相関が高いパターンとリアルタイムに抽出するパターンとを比較して、両者の一致不一致を判定する。

【００４０】

次に、障害発生と相関の高いパターンを学習するときの問題について説明する。

図４は、予兆検知における背景ノイズの影響例を示す図である。

監視サーバ１００が収集するメッセージの中には、障害との関連性が低く継続的に発生するメッセージが含まれる。第２の実施の形態では、このような種類のメッセージを「背景ノイズ」として扱う。背景ノイズは継続的に発生するため、通常、他の種類のメッセージよりも出現頻度が高い。背景ノイズの例としては、使用していない通信ポートに対する監視機能がＯＮになっていることで発生するメッセージなど、管理者が無視できるような軽度の注意喚起のメッセージが挙げられる。システムの運用上、注意喚起のメッセージが発生するような設定を行い、管理者がこのメッセージを意図的に無視する場合がある。

【００４１】

背景ノイズとしてのメッセージの種類は、時間の経過に応じて変化することがある。例えば、システム構成を変更したときやシステムを利用した業務プロセスを変更したとき、背景ノイズが大きく変化し得る。業務プロセスの変更の例としては、ユーザがストレージ２２に格納されたファイルを直接編集するという業務手順から、業務サーバ２１で実行されるＷｅｂアプリケーションのプログラムを介して当該ファイルを編集するという業務手順に変えることが挙げられる。構成変更や業務プロセスの変更は、監視対象のシステムの規模が大きいほど高頻度で生じる。これに対し、パターン学習に用いたメッセージの集合と予兆検知時に用いる現在のメッセージの集合とには、背景ノイズとして異なる種類のメッセージが多く含まれている可能性がある。このため、次のような問題が生じる。

【００４２】

学習に用いたメッセージの集合の中に、メッセージＡ，Ｂ，Ｘが含まれているとする。メッセージＡ，Ｂの組み合わせが障害直前に高確率で現れるとする。ただし、メッセージＸが背景ノイズとして継続的に多数発生している。すると、監視サーバ１００は、スライディングウィンドウ内にメッセージＡ，Ｂ，Ｘが含まれるため、メッセージＸを除外しないと、障害と相関の高いパターンとして［Ａ，Ｂ，Ｘ］を学習してしまう。

【００４３】

一方、現在収集したメッセージの集合の中に、メッセージＡ，Ｂ，Ｙが含まれているとする。メッセージＹが背景ノイズとして継続的に多数発生しており、メッセージＸは発生していない。すなわち、構成変更や業務プロセスの変更などによって背景ノイズが変化している。すると、監視サーバ１００は、スライディングウィンドウ内にメッセージＡ，Ｂ，Ｙが含まれるため、メッセージＹを除外しないとパターンとして［Ａ，Ｂ，Ｙ］を抽出する。学習したパターン［Ａ，Ｂ，Ｘ］と現在抽出したパターン［Ａ，Ｂ，Ｙ］とは一致しないため、監視サーバ１００は、このままでは障害の予兆を検知しない。

【００４４】

図５は、予兆検知における背景ノイズの除外例を示す図である。

上記の問題に対し、第２の実施の形態では、学習時に抽出するパターンから学習時における背景ノイズを除外し、また、予兆検知時に抽出するパターンから予兆検知時における背景ノイズを除外する。予兆検知時における背景ノイズは、学習時における背景ノイズと異なる可能性がある。そして、背景ノイズが除外されたパターン同士が比較される。

【００４５】

例えば、監視サーバ１００は、学習に用いるメッセージの中から出現頻度が高いメッセージＸを検索し、学習時における背景ノイズと判定する。そして、監視サーバ１００は、メッセージＸを除外したパターン［Ａ，Ｂ］を学習する。また、監視サーバ１００は、現在収集したメッセージの集合の中から出現頻度が高いメッセージＹを検索し、予兆検知時

における背景ノイズと判定する。そして、監視サーバ１００は、メッセージＹを除外したパターン［Ａ，Ｂ］を抽出する。学習したパターン［Ａ，Ｂ］と現在抽出したパターン［Ａ，Ｂ］とは一致するため、監視サーバ１００は、障害の予兆を検知する。

【００４６】

次に、監視サーバ１００の構成について説明する。

図６は、監視サーバのハードウェア例を示すブロック図である。

監視サーバ１００は、ＣＰＵ１０１、ＲＡＭ１０２、ＨＤＤ１０３、画像信号処理部１０４、入力信号処理部１０５、媒体リーダ１０６および通信インタフェース１０７を有する。ＣＰＵ１０１は、第１の実施の形態の演算部１２の一例である。ＲＡＭ１０２またはＨＤＤ１０３は、第１の実施の形態の記憶部１１の一例である。

【００４７】

ＣＰＵ１０１は、プログラムの命令を実行する演算回路を含むプロセッサである。ＣＰＵ１０１は、ＨＤＤ１０３に記憶されているプログラムやデータの少なくとも一部をＲＡＭ１０２にロードし、プログラムを実行する。なお、ＣＰＵ１０１は複数のプロセッサコアを備えてもよく、監視サーバ１００は複数のプロセッサを備えてもよく、以下で説明する処理を複数のプロセッサまたはプロセッサコアを用いて並列実行してもよい。また、複数のプロセッサの集合（マルチプロセッサ）を「プロセッサ」と呼んでもよい。

【００４８】

ＲＡＭ１０２は、ＣＰＵ１０１が実行するプログラムやＣＰＵ１０１が演算に用いるデータを一時的に記憶する揮発性メモリである。なお、監視サーバ１００は、ＲＡＭ以外の種類のメモリを備えてもよく、複数個のメモリを備えてもよい。

【００４９】

ＨＤＤ１０３は、ＯＳやミドルウェアやアプリケーションソフトウェアなどのソフトウェアのプログラム、および、データを記憶する不揮発性の記憶装置である。なお、監視サーバ１００は、フラッシュメモリやＳＳＤ（Solid State Drive）などの他の種類の記憶装置を備えてもよく、複数の不揮発性の記憶装置を備えてもよい。

【００５０】

画像信号処理部１０４は、ＣＰＵ１０１からの命令に従って、監視サーバ１００に接続されたディスプレイ３１に画像を出力する。ディスプレイ３１としては、ＣＲＴ（Cathode Ray Tube）ディスプレイ、液晶ディスプレイ（ＬＣＤ：Liquid Crystal Display）、プラズマディスプレイ（ＰＤＰ：Plasma Display Panel）、有機ＥＬ（ＯＥＬ：Organic Electro-Luminescence）ディスプレイなどを用いることができる。

【００５１】

入力信号処理部１０５は、監視サーバ１００に接続された入力デバイス３２から入力信号を取得し、ＣＰＵ１０１に出力する。入力デバイス３２としては、マウスやタッチパネルやタッチパッドやトラックボールなどのポインティングデバイス、キーボード、リモートコントローラ、ボタンスイッチなどを用いることができる。また、監視サーバ１００に、複数の種類の入力デバイスが接続されていてもよい。

【００５２】

媒体リーダ１０６は、記録媒体３３に記録されたプログラムやデータを読み取る読み取り装置である。記録媒体３３として、例えば、フレキシブルディスク（ＦＤ：Flexible Disk）やＨＤＤなどの磁気ディスク、ＣＤ（Compact Disc）やＤＶＤ（Digital Versatile Disc）などの光ディスク、光磁気ディスク（ＭＯ：Magneto-Optical disk）、半導体メモリなどを使用できる。媒体リーダ１０６は、例えば、記録媒体３３から読み取ったプログラムやデータをＲＡＭ１０２またはＨＤＤ１０３に格納する。

【００５３】

通信インタフェース１０７は、ネットワーク２０に接続され、ネットワーク２０を介して、業務で使用される電子機器（業務サーバ２１、ストレージ２２、通信装置２３など）、クライアント２４および管理サーバ２５と通信を行うインタフェースである。通信インタフェース１０７は、ケーブルで通信装置と接続される有線通信インタフェースでもよい

し、基地局と無線リンクで接続される無線通信インタフェースでもよい。
【００５４】
なお、監視サーバ１００は、媒体リーダ１０６を備えていなくてもよく、端末装置から制御される場合には画像信号処理部１０４や入力信号処理部１０５を備えていなくてもよい。また、ディスプレイ３１や入力デバイス３２が、監視サーバ１００の筐体と一体に形成されていてもよい。業務サーバ２１、クライアント２４および管理サーバ２５も、監視サーバ１００と同様のハードウェアを用いて実現することができる。
【００５５】
図７は、監視サーバの機能例を示すブロック図である。
監視サーバ１００は、受信部１１１、障害検出部１１２、障害情報記憶部１１３および設定情報記憶部１１４を有する。また、監視サーバ１００は、頻度算出部１２１、メッセージバッファ１２２、頻度情報記憶部１２３、パターン抽出部１２４、フィルタリング部１２５、学習部１２６および学習情報記憶部１２７を有する。監視サーバ１００は、頻度算出部１３１、メッセージバッファ１３２、頻度情報記憶部１３３、パターン抽出部１３４、フィルタリング部１３５、パターン比較部１３６および警告部１３７を有する。
【００５６】
障害情報記憶部１１３、設定情報記憶部１１４、メッセージバッファ１２２，１３２、頻度情報記憶部１２３，１３３および学習情報記憶部１２７は、例えば、ＲＡＭ１０２またはＨＤＤ１０３に確保した記憶領域として実装される。上記の他のユニットは、例えば、ＣＰＵ１０１が実行するプログラムのモジュールとして実装される。
【００５７】
受信部１１１は、業務サーバ２１、ストレージ２２および通信装置２３などの電子機器からメッセージを受信する。受信部１１１が受信するメッセージには、複数個の電子機器または複数の種類の電子機器からのメッセージが混在していてよい。受信部１１１は、受信時刻を示すタイムスタンプを各メッセージに付与する。ただし、メッセージに生成時刻または送信時刻の情報が含まれている場合、別途タイムスタンプを付与しなくてもよい。
【００５８】
障害検出部１１２は、受信部１１１からメッセージを取得し、メッセージの種類を判定する。取得したメッセージが、ＨＤＤ障害やサーバソフトウェアの異常停止などの障害を示している場合、障害発生を示す障害情報を生成する。障害情報には、障害発生時刻としてメッセージに付与されている時刻や障害内容などを示す情報が含まれる。障害検出部１１２は、生成した障害情報を障害情報記憶部１１３に格納する。
【００５９】
障害情報記憶部１１３は、過去に発生した障害の内容と障害発生時刻とを対応付けた障害情報を記憶する。障害情報は、障害検出部１１２によって書き込まれることもあるし、管理者の操作に基づいてクライアント２４から書き込まれることもある。設定情報記憶部１１４は、管理者から見て障害との関連性が明らかに高いメッセージの種類および障害との関連性が明らかに低いメッセージの種類を示すユーザ設定情報を記憶する。ユーザ設定情報は、管理者の操作に基づいてクライアント２４から書き込まれる。
【００６０】
頻度算出部１２１は、受信部１１１からメッセージを取得し、直近の一定時間（例えば、２４時間）に取得されたメッセージの集合を管理し、メッセージの種類毎の出現頻度を継続的に算出する。頻度算出部１２１は、受信部１１１からメッセージを取得すると、取得したメッセージをメッセージバッファ１２２に追加し、また、一定時間より古いメッセージをメッセージバッファ１２２から削除する。そして、頻度算出部１２１は、メッセージバッファ１２２に記憶されているメッセージの集合から、メッセージの種類毎の出現頻度を示す頻度情報を生成し、生成した頻度情報を頻度情報記憶部１２３に格納する。
【００６１】
メッセージバッファ１２２は、監視サーバ１００が収集したメッセージを一定時間だけ記憶するバッファ領域である。頻度情報記憶部１２３は、メッセージの種類と出現頻度と

出現頻度に基づいて算出されるスコアとを対応付けた頻度情報を記憶する。スコアは、出現確率の逆数であり、出現頻度が高いほど小さく出現頻度が低いほど大きい値をとる。頻度情報は、頻度算出部１２１によって継続的に更新される。

【００６２】

パターン抽出部１２４は、受信部１１１からメッセージを取得し、スライディングウィンドウの時間（例えば、５分間）だけメッセージを保持し、メッセージのパターンを抽出する。パターン抽出部１２４は、受信部１１１からメッセージを取得すると、取得したメッセージが含まれるようにスライディングウィンドウをシフトし、スライディングウィンドウから外れた古いメッセージ（例えば、５分以上前のメッセージ）を削除する。そして、パターン抽出部１２４は、スライディングウィンドウに含まれるメッセージの種類を列挙したメッセージのパターンを抽出し、フィルタリング部１２５に出力する。

【００６３】

フィルタリング部１２５は、頻度情報記憶部１２３に記憶された最新の頻度情報および設定情報記憶部１１４に記憶されたユーザ設定情報を参照して、抽出されたパターンから背景ノイズを除外する。フィルタリング部１２５は、パターン抽出部１２４からパターンを取得すると、パターン内からスコアが閾値以下であるメッセージの種類（出現確率が閾値以上のメッセージの種類）を検索する。そして、フィルタリング部１２５は、検索されたメッセージの種類を背景ノイズと判定してパターン内から除外する。ただし、フィルタリング部１２５は、ユーザ設定情報によって障害との関連性が高いと指定されているメッセージの種類は除外しない。また、フィルタリング部１２５は、ユーザ設定情報によって障害との関連性が低いと指定されているメッセージの種類は除外する。フィルタリング部１２５は、フィルタリングしたパターンを学習部１２６に出力する。

【００６４】

学習部１２６は、障害情報記憶部１１３に記憶された障害情報を参照して、フィルタリングされたパターンと障害との間の相関を示す学習情報を生成し、学習情報記憶部１２７に格納する。学習部１２６は、フィルタリング部１２５からパターンを取得すると、取得したパターンが現れた時刻から一定時間以内に障害が発生したか判定する。パターンが現れた時刻としては、例えば、スライディングウィンドウの末尾の時刻やスライディングウィンドウに含まれる末尾のメッセージの受信時刻などを用いることができる。学習部１２６は、同じパターン現れた回数とそのうち一定時間以内に障害が発生した回数とをカウントし、パターンと障害との共起確率を継続的に更新していく。

【００６５】

なお、パターン抽出部１２４、フィルタリング部１２５および学習部１２６は、メッセージが受信されてすぐに当該メッセージを用いた学習を進めてもよい。ただし、学習部１２６では、パターンと障害との共起確率を算出するため、パターンが抽出されてから少なくとも一定時間待つことになる。また、パターン抽出部１２４、フィルタリング部１２５および学習部１２６は、バッチ処理のように、メッセージが受信されてからある程度時間が経った後に当該メッセージを用いた学習を進めてもよい。また、第２の実施の形態では、メッセージの集合からパターンを抽出した後に背景ノイズを除外しているが、メッセージの集合から背景ノイズを除外した後にパターンを抽出するようにしてもよい。

【００６６】

頻度算出部１２１、メッセージバッファ１２２、頻度情報記憶部１２３、パターン抽出部１２４およびフィルタリング部１２５は、学習系に属する。これに対し、頻度算出部１３１、メッセージバッファ１３２、頻度情報記憶部１３３、パターン抽出部１３４およびフィルタリング部１３５は、検知系に属しており学習系と対応している。

【００６７】

頻度算出部１３１は、受信部１１１からメッセージを取得し、直近の一定時間に取得されたメッセージの集合を管理し、メッセージの種類毎の出現頻度を継続的に算出する。メッセージバッファ１３２は、収集されたメッセージを一定時間だけ記憶するバッファ領域である。頻度情報記憶部１３３は、メッセージの種類と出現頻度と出現頻度に基づいて算

出されるスコアとを対応付けた頻度情報を記憶する。パターン抽出部１３４は、受信部１１１からメッセージを取得し、スライディングウィンドウの時間だけメッセージを保持し、メッセージのパターンを抽出する。フィルタリング部１３５は、頻度情報記憶部１３３に記憶された最新の頻度情報および設定情報記憶部１１４に記憶されたユーザ設定情報を参照して、抽出されたパターンから背景ノイズを除外する。

【００６８】

パターン比較部１３６は、学習情報記憶部１２７に記憶された学習情報を参照して、障害の予兆を検知する。パターン比較部１３６は、フィルタリング部１３５からパターンを取得すると、取得したパターンを学習情報の中から検索する。学習情報に記載されたパターンからは学習時点における背景ノイズが除外されており、現在取得したパターンからは現時点における背景ノイズが除外されている。取得したパターンと障害との間の共起確率が閾値（例えば、８０％）以上である場合、パターン比較部１３６は、障害の予兆がある、すなわち、現在から一定時間以内に障害が発生する可能性が高いと判定する。

【００６９】

警告部１３７は、パターン比較部１３６が障害の予兆を検知すると、システムの管理者に対して警告する。例えば、警告部１３７は、障害の予兆を示す警告情報を生成してクライアント２４に送信する。ただし、警告部１３７は、監視サーバ１００に接続されたディスプレイ３１に警告情報を表示するようにしてもよい。警告情報には、例えば、障害の予兆があると判定する原因となったメッセージが含まれる。

【００７０】

図８は、メッセージテーブルの例を示す図である。

メッセージテーブル１４１は、受信された複数のメッセージを格納する。メッセージテーブル１４１に相当するテーブルとして、一定時間（例えば、２４時間）分のメッセージを格納したメッセージテーブルが、メッセージバッファ１２２，１３２に記憶される。また、スライディングウィンドウの時間幅（例えば、５分間）分のメッセージを格納したメッセージテーブルが、パターン抽出部１２４，１３４によって保持されている。メッセージテーブル１４１は、時刻、種類およびメッセージの項目を含む。

【００７１】

時刻の項目には、受信部１１１がメッセージを受信した時刻が登録される。ただし、送信元の電子機器がメッセージに生成時刻または送信時刻を付与している場合、時刻の項目には、生成時刻または送信時刻が登録されてもよい。種類の項目には、メッセージの種類を示す識別情報が登録される。メッセージは、ＲＡＩＤ（Redundant Arrays of Independent Disks）のインタフェース検出失敗、カウンタのオーバーフロー、ディスク検出失敗などの発生原因に応じて、複数の種類に分類される。メッセージの種類を示す識別情報は、送信元の電子機器がメッセージに付与してもよいし、受信部１１１が付与してもよい。メッセージの項目には、メッセージに記載された不具合の具体的な内容が登録される。

【００７２】

図９は、頻度テーブルの例を示す図である。

頻度テーブル１４２は、メッセージの種類と出現頻度と出現頻度に基づいて算出されるスコアとを対応付けた頻度情報を格納する。頻度テーブル１４２に相当するテーブルとして、頻度情報記憶部１２３，１３３それぞれに頻度テーブルが記憶される。頻度テーブル１４２は、種類、出現数、総数、頻度およびスコアの項目を含む。

【００７３】

種類の項目には、メッセージの種類を示す識別情報が登録される。出現数の項目には、各種類のメッセージの受信回数が登録される。総数の項目には、全ての種類のメッセージの受信総数が登録される。頻度の項目には、出現頻度として各種類のメッセージの出現確率が登録される。ある種類のメッセージの出現確率は、当該種類のメッセージの出現数を全ての種類のメッセージの総数で割ることで算出できる。スコアの項目には、メッセージの種類毎に、出現頻度が高いほど小さく出現頻度が低いほど大きい指標値が登録される。スコアは、例えば、出現確率の逆数として算出することができる。

【００７４】
頻度情報記憶部１２３に記憶された頻度テーブルの出現数や総数は、メッセージバッファ１２２に格納された学習に使用する一定時間分（例えば、２４時間分）のメッセージの集合から算出される。新たなメッセージの受信などによってメッセージバッファ１２２に格納されたメッセージの集合が変わると、出現数・総数・頻度・スコアが更新される。頻度情報記憶部１３３に記憶された頻度テーブルに登録される出現数や総数は、メッセージバッファ１３２に格納された直近の一定時間分（例えば、直近の２４時間分）のメッセージの集合から算出される。新たなメッセージの受信によってメッセージバッファ１３２に格納されたメッセージの集合が変わると、出現数・総数・頻度・スコアが更新される。
【００７５】
図１０は、ユーザ設定テーブルの例を示す図である。
ユーザ設定テーブル１４３は、管理者によって作成されたユーザ設定情報を格納する。ユーザ設定テーブル１４３は、設定情報記憶部１１４に記憶されている。ユーザ設定テーブル１４３は、種類、除外フラグおよび非除外フラグの項目を含む。
【００７６】
種類の項目には、メッセージの種類を示す識別情報が登録される。除外フラグの項目には、当該種類のメッセージが、管理者から見て障害との関連性が低いか否かを示すフラグが設定される。障害との関連性が低いと指定されたメッセージの種類は、出現頻度が低い（スコアが大きい）場合であっても背景ノイズであると判定され、抽出されたパターンの中から除外される。非除外フラグの項目には、当該種類のメッセージが、管理者から見て障害との関連性が高いか否かを示すフラグが設定される。障害との関連性が高いと指定されたメッセージの種類は、出現頻度が高い（スコアが小さい）場合であっても背景ノイズでないと判定され、抽出されたパターンの中から除外されない。
【００７７】
図１１は、障害テーブルの例を示す図である。
障害テーブル１４４は、障害検出部１１２または管理者によって作成された障害情報を格納する。障害テーブル１４４は、障害情報記憶部１１３に記憶されている。障害テーブル１４４は、時刻および障害の項目を含む。
【００７８】
時刻の項目には、障害が発生した時刻が登録される。障害発生を示すメッセージに基づいて障害情報を生成する場合、障害発生時刻として、メッセージに記載された生成時刻や送信時刻、受信部１１１が当該メッセージを受信した時刻などを用いることができる。障害の項目には、発生した障害の内容が登録される。障害の内容としては、例えば、ＨＤＤ障害、性能低下、Ｗｅｂサーバ応答なしなどが挙げられる。
【００７９】
図１２は、学習テーブルの例を示す図である。
学習テーブル１４５は、学習部１２６が生成した学習情報を格納する。学習テーブル１４５は、学習情報記憶部１２７に記憶されている。学習テーブル１４５は、パターン、障害、出現数、予兆数および共起確率の項目を含む。
【００８０】
パターンの項目には、同時期に受信されたメッセージの種類の組み合わせを示すメッセージのパターンが登録される。パターンを抽出するにあたり、スライディングウィンドウ内でのメッセージの出現順序は考慮しなくてよい。また、パターンを抽出するにあたり、スライディングウィンドウに同じ種類のメッセージが２以上含まれていても、同じ種類のメッセージの個数は考慮しなくてよい。例えば、各メッセージの種類の識別情報を用いて、［１，３，４，７，９］、［１，４，６，１０，１２］、［３，７，１１，１４］のようにパターンが表現される。ただし、学習テーブル１４５に登録されるパターンには、学習時に背景ノイズと判定されたメッセージの種類は含まれていない。
【００８１】
障害の項目には、障害テーブル１４４に登録された障害の内容のうち、パターンが出現

してから一定時間以内に発生したことのある障害の内容が登録される。出現数の項目には、過去に各パターンが出現した回数が登録される。予兆数の項目には、パターンが出現してから一定時間以内に障害が発生した回数が登録される。共起確率の項目には、パターンと障害との間の相関を示す確率が登録される。相関が大きいほど共起確率が大きくなる。共起確率は、例えば、予兆数を出現数で割ることで算出できる。

【００８２】

次に、監視サーバ１００が実行する情報処理の手順について説明する。

図１３は、頻度算出の手順例を示すフローチャートである。

この頻度算出の手順は、学習系として、頻度算出部１２１が受信部１１１からメッセージを取得する毎に実行される。検知系として、頻度算出部１３１が受信部１１１からメッセージを取得する毎にも、頻度算出部１２１と同様の頻度算出の手順が実行される。

【００８３】

（Ｓ１０）頻度算出部１２１は、受信部１１１から取得したメッセージ（新たに受信されたメッセージ）を、メッセージバッファ１２２に格納する。

（Ｓ１１）頻度算出部１２１は、現在時刻から一定時間（例えば、２４時間）以上古いメッセージをメッセージバッファ１２２から検索し、検索したメッセージを削除する。

【００８４】

（Ｓ１２）頻度算出部１２１は、メッセージバッファ１２２に格納されているメッセージ、すなわち、直近の一定時間に収集されたメッセージの総数をカウントする。また、頻度算出部１２１は、各メッセージの種類を判定し、種類毎にメッセージバッファ１２２に格納されているメッセージの数（出現数）をカウントする。

【００８５】

（Ｓ１３）頻度算出部１２１は、ステップＳ１２でカウントした総数および種類毎の出現数を、頻度情報記憶部１２３の頻度テーブルに登録する。また、頻度算出部１２１は、総数および種類毎の出現数から種類毎の頻度および種類毎のスコアを算出し、当該頻度テーブルに登録する。例えば、頻度＝出現数÷総数とし、スコアは頻度の逆数とする。

【００８６】

このようにして、学習系である頻度算出部１２１は、新たなメッセージの受信に応じて継続的に、学習時におけるメッセージの種類毎の出現頻度およびスコアを更新する。ただし、頻度算出部１２１は、メッセージ受信からある程度の時間が経過した後に頻度算出を行ってもよいし、ある程度の量のメッセージが溜まってから頻度算出を行ってもよい。また、検知系である頻度算出部１３１は、新たなメッセージの到着に応じて継続的に、現在（検知時）におけるメッセージの種類毎の出現頻度およびスコアを更新する。

【００８７】

図１４は、パターン学習の手順例を示すフローチャートである。

このパターン学習の手順は、学習系として、パターン抽出部１２４が受信部１１１からメッセージを取得する毎に実行される。ただし、メッセージ受信からある程度の時間が経過した後や、ある程度の量のメッセージが溜まってから行うことも可能である。

【００８８】

（Ｓ２０）パターン抽出部１２４は、受信部１１１から取得したメッセージ（新たに受信されたメッセージ）をスライディングウィンドウに追加する。

（Ｓ２１）パターン抽出部１２４は、新たなメッセージの追加に応じてスライディングウィンドウを前方にシフトさせ、スライディングウィンドウから外れる古いメッセージを削除する。すなわち、パターン抽出部１２４は、保持しているメッセージの中から、新たなメッセージの受信時刻からスライディングウィンドウ時間幅（例えば、５分）以上古いメッセージを検索し、検索された古いメッセージを削除する。

【００８９】

（Ｓ２２）パターン抽出部１２４は、スライディングウィンドウに含まれるメッセージの種類を判定し、メッセージの種類を列挙したパターンを生成する。

（Ｓ２３）フィルタリング部１２５は、頻度情報記憶部１２３に記憶された頻度テーブ

ルを参照して、ステップＳ２２で生成されたパターンに含まれる複数のメッセージの種類のうち、スコアが閾値以下であるメッセージの種類を検索する。

【００９０】

（Ｓ２４）フィルタリング部１２５は、設定情報記憶部１１４に記憶されたユーザ設定テーブル１４３を参照して、ステップＳ２２で生成されたパターンに含まれる複数のメッセージの種類のうち、管理者から指定されたメッセージの種類を検索する。指定されるメッセージの種類には、前述のステップＳ１３で算出されたスコアに関係なく、背景ノイズとして除外すべきものと背景ノイズではなく除外すべきでないものとが含まれ得る。

【００９１】

（Ｓ２５）フィルタリング部１２５は、ステップＳ２２で生成されたパターンから一部のメッセージの種類をフィルタリングすることで、背景ノイズを除外する。具体的には、フィルタリング部１２５は、スコアの低いメッセージの種類を生成されたパターンから除外する。ただし、ユーザ設定テーブル１４３によって背景ノイズでないと指定されたメッセージの種類は除外されない。また、フィルタリング部１２５は、ユーザ設定テーブル１４３によって背景ノイズであると指定されたメッセージの種類を除外する。

【００９２】

（Ｓ２６）学習部１２６は、学習情報記憶部１２７に記憶された学習テーブル１４５において、フィルタリング部１２５が出力したパターンの出現数をインクリメントする。

（Ｓ２７）学習部１２６は、障害情報記憶部１１３に記憶された障害テーブル１４４を参照して、フィルタリング部１２５が出力したパターンの現れた時刻から一定時間以内に障害が発生したか判断する。パターンの現れた時刻としては、例えば、スライディングウィンドウの末尾の時刻や、スライディングウィンドウに含まれる末尾のメッセージの受信時刻などを用いることができる。一定時間以内に障害が発生した場合はステップＳ２８に処理が進み、障害が発生していない場合はステップＳ２９に処理が進む。

【００９３】

（Ｓ２８）学習部１２６は、学習テーブル１４５において、フィルタリング部１２５が出力したパターンの予兆数をインクリメントする。なお、学習テーブル１４５の障害の項目には、障害テーブル１４４に記載された障害の内容であって、パターンの現れた時刻から一定時間以内に発生した障害の内容が登録される。

【００９４】

（Ｓ２９）学習部１２６は、学習テーブル１４５において、フィルタリング部１２５が出力したパターンの共起確率を更新する。ステップＳ２７の判断がＹＥＳである場合、共起確率は、ステップＳ２８で更新した予兆数をステップＳ２６で更新した出現数で割ることで算出できる。ステップＳ２７の判断がＮＯである場合、共起確率は、更新されない現在の予兆数をステップＳ２６で更新した出現数で割ることで算出できる。

【００９５】

図１５は、予兆検知の手順例を示すフローチャートである。

この予兆検知の手順は、検知系として、パターン抽出部１３４が受信部１１１からメッセージを取得する毎に（好ましくは、リアルタイムに）実行される。

【００９６】

（Ｓ３０）パターン抽出部１３４は、受信部１１１から取得したメッセージ（新たに受信されたメッセージ）をスライディングウィンドウに追加する。

（Ｓ３１）パターン抽出部１３４は、新たなメッセージの追加に応じてスライディングウィンドウを前方にシフトさせ、スライディングウィンドウから外れる古いメッセージを削除する。すなわち、パターン抽出部１３４は、保持しているメッセージの中から、新たなメッセージの受信時刻からスライディングウィンドウ時間幅（例えば、５分）以上古いメッセージを検索し、検索された古いメッセージを削除する。

【００９７】

（Ｓ３２）パターン抽出部１３４は、スライディングウィンドウに含まれるメッセージの種類を判定し、メッセージの種類を列挙したパターンを生成する。

（Ｓ３３）フィルタリング部１３５は、頻度情報記憶部１３３に記憶された頻度テーブルを参照して、ステップＳ３２で生成されたパターンに含まれる複数のメッセージの種類のうち、スコアが閾値以下であるメッセージの種類を検索する。スコアが閾値以下である（出現頻度が閾値以上である）メッセージの種類は、過去に学習テーブル１４５が更新されたとき（学習時）と現在（検知時）とで異なる可能性がある。

【００９８】

（Ｓ３４）フィルタリング部１３５は、設定情報記憶部１１４に記憶されたユーザ設定テーブル１４３を参照して、ステップＳ３２で生成されたパターンに含まれる複数のメッセージの種類のうち、管理者から指定されたメッセージの種類を検索する。

【００９９】

（Ｓ３５）フィルタリング部１３５は、ステップＳ３２で生成されたパターンから一部のメッセージの種類をフィルタリングすることで、背景ノイズを除外する。具体的には、フィルタリング部１３５は、スコアの低いメッセージの種類を生成されたパターンから除外する。ただし、ユーザ設定テーブル１４３によって背景ノイズでないと指定されたメッセージの種類は除外されない。また、フィルタリング部１３５は、ユーザ設定テーブル１４３によって背景ノイズであると指定されたメッセージの種類を除外する。

【０１００】

（Ｓ３６）パターン比較部１３６は、フィルタリング部１３５が出力したパターンを、学習情報記憶部１２７に記憶された学習テーブル１４５から検索する。

（Ｓ３７）パターン比較部１３６は、フィルタリング部１３５が出力したパターンが学習テーブル１４５に登録されており、かつ、当該パターンの共起確率が閾値以上であるか判断する。この条件を満たす場合、パターン比較部１３６は監視対象のシステムに障害の予兆があると判断し、ステップＳ３８に処理が進む。この条件を満たさない（フィルタリング部１３５が出力したパターンが学習テーブル１４５に登録されていないか、または、当該パターンの共起確率が閾値未満である）場合、障害の予兆がないと判断する。

【０１０１】

（Ｓ３８）警告部１３７は、管理者に対して障害の予兆を警告する。例えば、警告部１３７は、障害の予兆を示す警告情報を生成してクライアント２４に送信する。

第２の実施の形態の情報処理システムによれば、学習時のメッセージの種類毎の出現頻度に基づいて学習時の背景ノイズが判定され、背景ノイズを除外したメッセージのパターンが学習される。また、検知時のメッセージの種類毎の出現頻度に基づいて検知時の背景ノイズが判定され、背景ノイズを除外したメッセージのパターンと学習結果とが比較される。これにより、監視対処のシステムの構成変更や業務プロセスの変更などに応じて背景ノイズが変化しても、既存の学習結果を利用して障害の予兆を検知することができる。その結果、再学習の負荷を抑制でき、また、障害予兆検知の精度を高めることができる。また、第２の実施の形態では、継続的に頻度情報が更新されるため、背景ノイズの変化に迅速に対応でき、学習精度および障害予兆の検知精度を向上させることができる。

【０１０２】

［第３の実施の形態］

次に、第３の実施の形態を説明する。前述の第２の実施の形態との違いを中心に説明し、第２の実施の形態と同様の事項については適宜説明を省略する。第３の実施の形態の情報処理システムは、図２と同様の構成によって実現できる。ただし、第３の実施の形態の情報処理システムは、監視サーバ１００に代えて後述する監視サーバ１００ａを含む。監視サーバ１００ａは、メッセージのパターンと障害との相関を学習するタイミングや、メッセージの種類毎の出現頻度を更新するタイミングが、監視サーバ１００と異なる。

【０１０３】

図１６は、監視サーバの他の機能例を示すブロック図である。

監視サーバ１００ａは、受信部１１１、障害検出部１１２、障害情報記憶部１１３および設定情報記憶部１１４を有する。また、監視サーバ１００ａは、パターン抽出部１２４ａ、フィルタリング部１２５ａ、学習部１２６、学習情報記憶部１２７およびログ記憶部

１２８を有する。また、監視サーバ１００ａは、頻度算出部１３１ａ、頻度情報記憶部１３３ａ、パターン抽出部１３４、フィルタリング部１３５ａ、パターン比較部１３６、警告部１３７およびログ記憶部１３８を有する。以下、監視サーバ１００ａが有するユニットのうち、第２の実施の形態の監視サーバ１００と異なるユニットについて説明する。

【０１０４】

ログ記憶部１２８は、受信部１１１で受信されたメッセージを含むログファイルを記憶する。受信部１１１で新たなメッセージが受信されると、当該メッセージがログファイルに追記される。メッセージは、パターン抽出部１２４ａから利用されると消去される。

【０１０５】

パターン抽出部１２４ａは、ログ記憶部１２８に記憶されたログファイルをバッチ方式で処理することで、ログファイルに含まれるメッセージの集合からメッセージのパターンを抽出する。例えば、パターン抽出部１２４ａは、ログファイルからメッセージを１つ読み込む毎にスライディングウィンドウをシフトし、スライディングウィンドウに含まれるメッセージの種類を列挙したパターンを生成する。全てのメッセージの読み込みを終えると、パターン抽出部１２４ａは、ログファイルを初期化（メッセージを消去）する。

【０１０６】

バッチ方式であるため、パターン抽出部１２４ａは、間欠的にログファイルを処理する。ログファイルを処理するタイミングとして、例えば、監視対象のシステムの構成が変更されたタイミングが挙げられる。システムの構成が変更されることは、管理サーバ２５に問い合わせることで知ることができる。また、ログファイルを処理するタイミングとして、所定の周期（例えば、２４時間や１ヶ月）や所定の時刻なども挙げられる。

【０１０７】

フィルタリング部１２５ａは、頻度情報記憶部１３３ａに記憶された頻度情報および設定情報記憶部１１４に記憶されたユーザ設定情報を参照して、パターン抽出部１２４ａが抽出したパターンから背景ノイズを除外する。後述するように、頻度情報記憶部１３３ａには、異なる複数の期間についての頻度情報が記憶されている。フィルタリング部１２５ａは、今回処理されたログファイルに対応する期間の頻度情報を選択して使用する。

【０１０８】

具体的には、フィルタリング部１２５ａは、パターン内からスコアが閾値以下であるメッセージの種類（出現確率が閾値以上であるメッセージの種類）を検索する。このとき、複数の期間（複数の世代）の頻度情報のうち、ログファイルに対応する期間（ログファイルと同じ世代）の頻度情報を参照する。そして、フィルタリング部１２５ａは、検索されたメッセージの種類を背景ノイズと判定してパターン内から除外する。ただし、フィルタリング部１２５ａは、ユーザ設定情報によって障害との関連性が高いと指定されているメッセージの種類は除外しない。また、フィルタリング部１２５ａは、ユーザ設定情報によって障害との関連性が低いと指定されているメッセージの種類は除外する。

【０１０９】

ログ記憶部１３８は、受信部１１１で受信されたメッセージを含むログファイルを記憶する。受信部１１１で新たなメッセージが受信されると、当該メッセージがログファイルに追記される。メッセージは、頻度算出部１３１ａから利用されると消去される。

【０１１０】

頻度算出部１３１ａは、ログ記憶部１３８に記憶されたログファイルをバッチ方式で処理することで、ログファイルに含まれるメッセージの集合からメッセージの種類毎の出現頻度を算出する。頻度算出部１３１ａは、ログファイルからメッセージを１つずつ読み込むことで、メッセージの種類毎の出現頻度を示す頻度情報を生成し、頻度算出に用いたメッセージの期間（ログファイルの世代）を示す期間情報と対応付けて頻度情報記憶部１３３ａに格納する。全てのメッセージの読み込みを終えると、頻度算出部１３１ａは、ログ記憶部１３８に記憶されたログファイルを初期化（メッセージを消去）する。

【０１１１】

なお、頻度算出部１３１ａがログファイルを処理するタイミングは、パターン抽出部１

２４ａがログファイルを処理するタイミングと同じでもよいし異なってもよい。例えば、頻度算出を行うタイミングとして、監視対象のシステムの構成が変更されたタイミング、所定の周期（例えば、２４時間や１ヶ月）、所定の時刻などが挙げられる。

【０１１２】

頻度情報記憶部１３３ａは、メッセージの種類と出現頻度と出現頻度に基づいて算出されるスコアとを対応付けた頻度情報を、１回のバッチ処理毎に記憶する。すなわち、頻度情報記憶部１３３ａは、異なる複数の期間のメッセージの集合から算出された複数の期間分（複数の世代）の頻度情報を記憶する。また、頻度情報記憶部１３３ａは、各頻度情報が何れの期間に対応するものかを管理するための期間情報を記憶する。

【０１１３】

フィルタリング部１３５ａは、頻度情報記憶部１３３ａに記憶された頻度情報および設定情報記憶部１１４に記憶されたユーザ設定情報を参照して、パターン抽出部１３４が抽出したパターンから背景ノイズを除外する。このとき、頻度情報記憶部１３３ａに記憶された複数の期間の頻度情報のうち、最新の頻度情報が使用される。

【０１１４】

図１７は、期間テーブルの例を示す図である。

期間テーブル１４６は、頻度算出部１３１ａが生成した期間情報を格納する。期間テーブル１４６は、頻度情報記憶部１３３ａに記憶されている。期間テーブル１４６は、期間、開始時刻および終了時刻の項目を含む。

【０１１５】

期間の項目には、頻度情報の生成に用いたメッセージの集合（ログファイルの世代）を識別するための識別情報が登録される。開始時刻の項目には、頻度情報の生成に用いたメッセージの集合の中で最先のメッセージの受信時刻が設定される。ただし、開始時刻の項目に、前回バッチ処理を行った時刻や、前回バッチ処理を行った後であってメッセージの受信を再開した時刻などを登録してもよい。終了時刻の項目には、頻度情報の生成に用いたメッセージの集合の中で末尾のメッセージの受信時刻が設定される。ただし、終了時刻の項目に、今回バッチ処理を行った時刻などを登録してもよい。

【０１１６】

図１８は、頻度算出の他の手順例を示すフローチャートである。

（Ｓ４０）頻度算出部１３１ａは、頻度情報を生成するタイミングが到来すると、新たな期間（新たな世代）の頻度テーブルを生成して頻度情報記憶部１３３ａに格納する。

【０１１７】

（Ｓ４１）頻度算出部１３１ａは、ログ記憶部１３８に記憶されたログファイルから、メッセージを１つ読み込む。このとき、ログファイルの先頭から順に読み込んでいく。ただし、ステップＳ４１では、任意の順序でメッセージを読み込むようにしてもよい。

【０１１８】

（Ｓ４２）頻度算出部１３１ａは、頻度テーブルに記載されたメッセージの総数をインクリメントする。また、頻度算出部１３１ａは、読み込んだメッセージの種類を判定し、頻度テーブルにおいて判定した種類に対応する出現数をインクリメントする。

【０１１９】

（Ｓ４３）頻度算出部１３１ａは、ログファイルの終端に達したか、すなわち、ログファイルに含まれる全てのメッセージが読み込まれたか判断する。終端に達した場合はステップＳ４４に処理が進み、それ以外の場合はステップＳ４１に処理が進む。

【０１２０】

（Ｓ４４）頻度算出部１３１ａは、頻度テーブルに記載されたメッセージの総数と種類毎の出現数から、種類毎の出現頻度とスコアを算出して頻度テーブルに登録する。例えば、出現頻度＝出現数÷総数とし、スコアは出現頻度の逆数とする。

【０１２１】

（Ｓ４５）頻度算出部１３１ａは、頻度情報記憶部１３３ａに記憶された期間テーブル１４６に、新たな期間の識別情報と、当該期間を示す開始時刻・終了時刻を登録する。

図１９は、パターン学習の他の手順例を示すフローチャートである。

【０１２２】

（Ｓ５０）パターン抽出部１２４ａは、学習を行うタイミングが到来すると、ログ記憶部１２８に記憶されたログファイルからメッセージを１つ読み込む。このとき、ログファイルの先頭から順に（受信時刻の早い順に）メッセージを読み込んでいく。

【０１２３】

（Ｓ５１）パターン抽出部１２４ａは、図１４のステップＳ２０～２２と同様にして、スライディングウィンドウを用いてメッセージのパターンを生成する。

（Ｓ５２）フィルタリング部１２５ａは、頻度情報記憶部１３３ａに記憶された複数の期間（複数の世代）の頻度テーブルの中から、今回処理するログファイルが属する期間（ログファイルと同じ世代）に対応する頻度テーブルを選択する。頻度テーブルの選択にあたっては、頻度情報記憶部１３３ａに記憶された期間テーブル１４６が参照される。

【０１２４】

（Ｓ５３）フィルタリング部１２５ａは、図１４のステップＳ２３～Ｓ２５と同様にして、ステップＳ５１で生成されたパターン内から背景ノイズを除外する。ただし、ステップＳ２３に相当する処理においては、ステップＳ５２で選択した頻度テーブルに記載されているメッセージの種類毎のスコア（ログファイルと同じ期間のスコア）を使用する。

【０１２５】

（Ｓ５４）学習部１２６は、学習情報記憶部１２７に記憶された学習テーブル１４５において、フィルタリング部１２５ａが出力したパターンの出現数をインクリメントする。

（Ｓ５５）学習部１２６は、障害情報記憶部１１３に記憶された障害テーブル１４４を参照して、フィルタリング部１２５ａが出力したパターンの現れた時刻から一定時間以内に障害が発生したか判断する。一定時間以内に障害が発生した場合はステップＳ５６に処理が進み、障害が発生していない場合はステップＳ５７に処理が進む。

【０１２６】

（Ｓ５６）学習部１２６は、学習テーブル１４５において、フィルタリング部１２５ａが出力したパターンの予兆数をインクリメントする。

（Ｓ５７）パターン抽出部１２４ａは、ログファイルの終端に達したか、すなわち、ログファイルに含まれる全てのメッセージが読み込まれたか判断する。終端に達した場合はステップＳ５８に処理が進み、それ以外の場合はステップＳ５０に処理が進む。

【０１２７】

（Ｓ５８）学習部１２６は、学習テーブル１４５において各パターンの共起確率を更新する。共起確率は、パターン毎に予兆数を出現数で割ることで算出できる。

第３の実施の形態の情報処理システムによれば、第２の実施の形態と同様、監視対処のシステムの構成変更や業務プロセスの変更などに応じて背景ノイズが変化しても、既存の学習結果を利用して障害の予兆を検知することができる。その結果、再学習の負荷を抑制でき、障害予兆検知の精度を高めることができる。また、第３の実施の形態では、頻度情報や学習情報がバッチ方式で更新されるため、監視サーバ１００ａの負荷を抑制できる。

【０１２８】

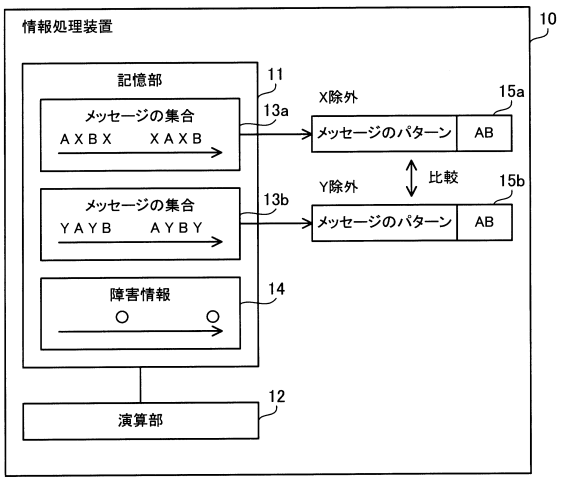
なお、前述のように、第１の実施の形態の情報処理は、情報処理装置１０にプログラムを実行させることで実現することができる。第２および第３の実施の形態の情報処理は、監視サーバ１００，１００ａにプログラムを実行させることで実現することができる。

【０１２９】

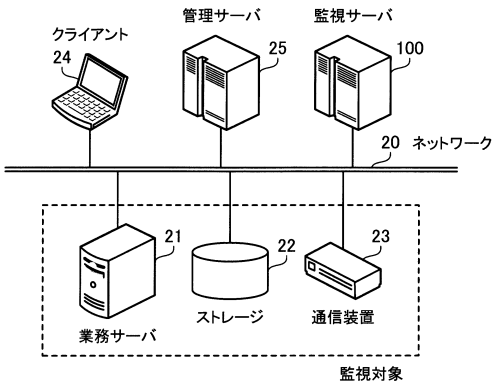
プログラムは、コンピュータ読み取り可能な記録媒体（例えば、記録媒体３３）に記録しておくことができる。記録媒体としては、例えば、磁気ディスク、光ディスク、光磁気ディスク、半導体メモリなどを使用できる。磁気ディスクには、ＦＤおよびＨＤＤが含まれる。光ディスクには、ＣＤ、ＣＤ－Ｒ（Recordable）／ＲＷ（Rewritable）、ＤＶＤおよびＤＶＤ－Ｒ／ＲＷが含まれる。プログラムは、可搬型の記録媒体に記録されて配布されることがある。その場合、可搬型の記録媒体からＨＤＤなどの他の記録媒体（例えば、ＨＤＤ１０３）にプログラムを複製して（インストールして）実行してもよい。

【符号の説明】
【０１３０】
１０　情報処理装置
１１　記憶部
１２　演算部
１３ａ，１３ｂ　メッセージの集合
１４　障害情報
１５ａ，１５ｂ　メッセージのパターン

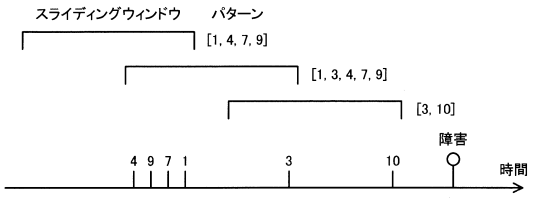
【図１】



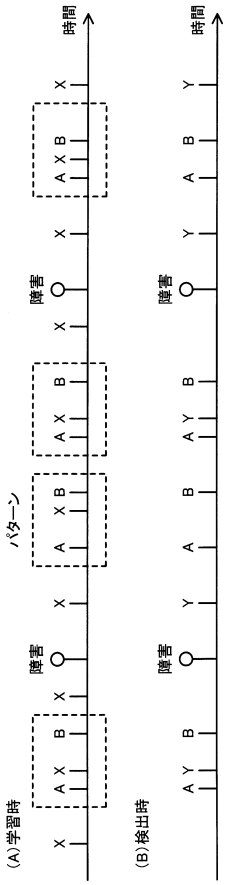
【図２】



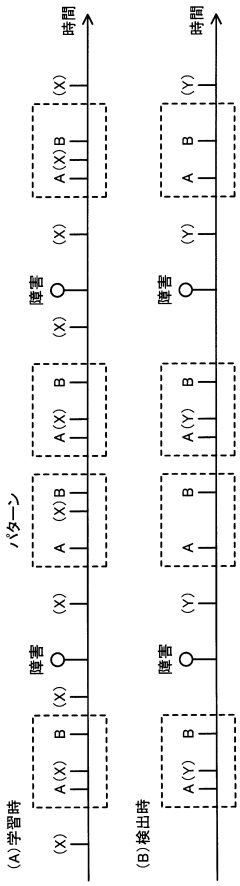
【図３】



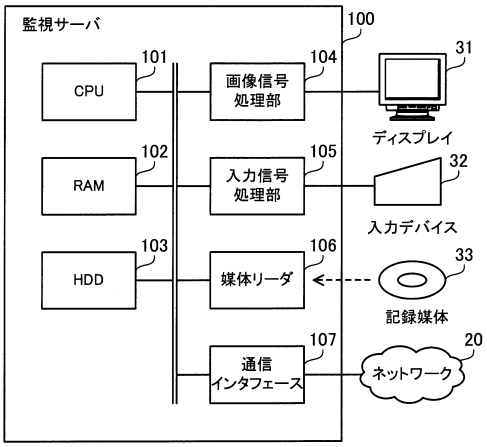
【図４】


(A)学習時
(B)検出時
パターン
障害
時間
X
Y
A
B

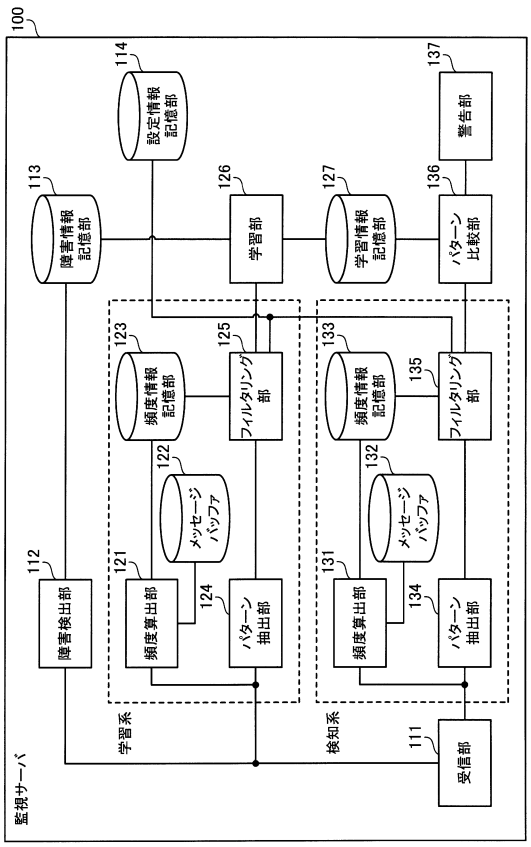
【図５】


(A)学習時
(B)検出時
パターン
障害
時間
(X)
(Y)
A
B

【図６】


監視サーバ
100
CPU
101
RAM
102
HDD
103
画像信号処理部
104
入力信号処理部
105
媒体リーダ
106
通信インタフェース
107
ディスプレイ
31
入力デバイス
32
記録媒体
33
ネットワーク
20

【図７】


100
監視サーバ
学習系
検知系
設定情報記憶部
114
警告部
137
障害情報記憶部
113
学習部
126
学習情報記憶部
127
パターン比較部
136
頻度情報記憶部
123
フィルタリング部
125
メッセージバッファ
122
頻度算出部
121
パターン抽出部
124
障害検出部
112
頻度情報記憶部
133
フィルタリング部
135
メッセージバッファ
132
頻度算出部
131
パターン抽出部
134
受信部
111

【図８】

メッセージテーブル 141

時刻	種類	メッセージ
2012-03-13 10:31:02	1	RAID: Adapter missing after reboot ...
2012-03-13 10:35:28	3	Correctable error count overflow ...
2012-03-13 10:36:18	10	Disk(4) missing after reboot ...
⋮	⋮	⋮

【図９】

頻度テーブル 142

種類	出現数	総数	頻度	スコア
1	20	102204	1.96e-04	5110.2
2	60215	102204	0.589	1.7
3	1503	102204	0.015	68.0
4	5805	102204	0.057	17.6
⋮	⋮	⋮	⋮	⋮

【図１０】

ユーザ設定テーブル 143

種類	除外フラグ	非除外フラグ
11	1 (True)	0 (False)
13	0 (False)	1 (True)
⋮	⋮	⋮

【図１１】

障害テーブル 144

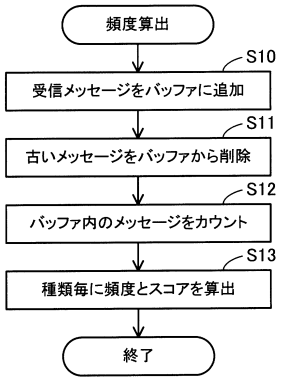
時刻	障害
2012-03-02 02:42:01	HDD障害
2012-03-05 13:35:21	性能低下
2012-03-06 07:22:50	Webサーバ応答なし
⋮	⋮

【図１２】

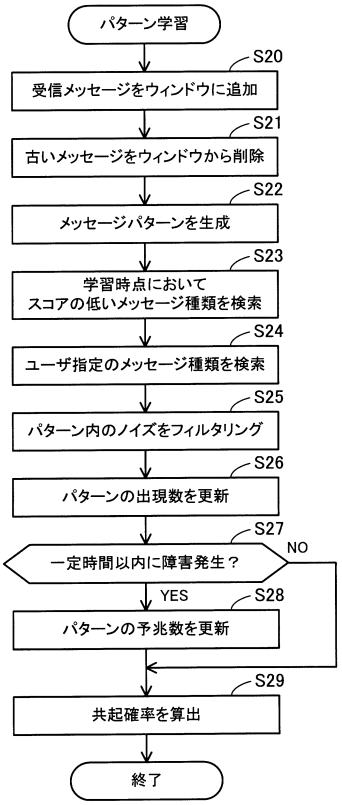
学習テーブル 145

パターン	障害	出現数	予兆数	共起確率
[1, 3, 4, 7, 9]	HDD障害	165	137	0.83
[1, 4, 6, 10, 12]	性能低下	120	108	0.90
[3, 7, 11, 14]	HDD障害	25	22	0.88

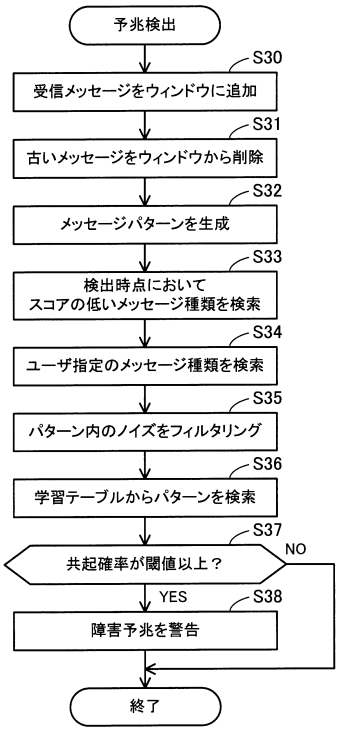
【図１３】



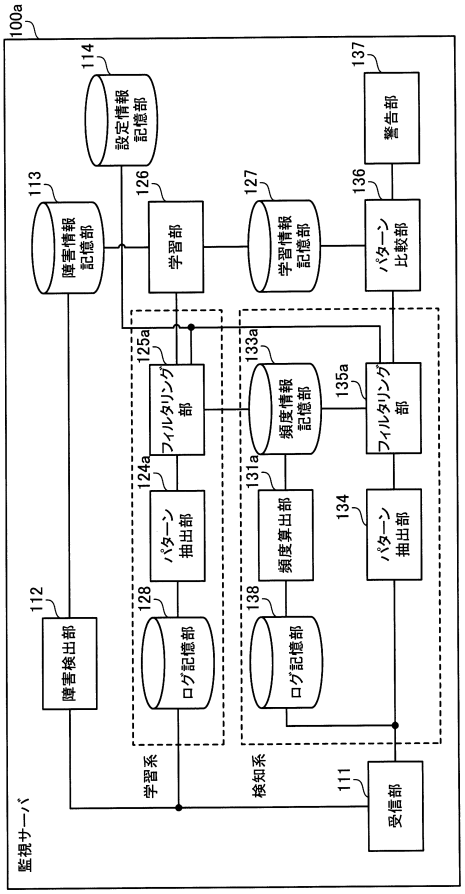
【図１４】



【図１５】



【図１６】

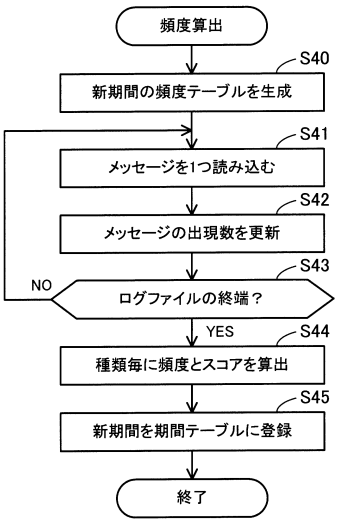


【図１７】

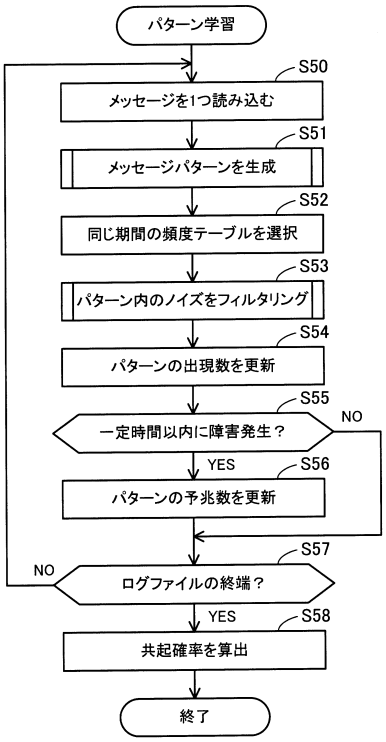
期間テーブル 146

期間	開始時刻	終了時刻
1	2012-03-13 05:00:02	2012-03-14 00:30:28
2	2012-03-14 05:00:01	2012-03-15 00:31:01
3	2012-03-15 05:00:06	2012-03-16 00:30:15
⋮	⋮	⋮

【図１８】



【図１９】


パターン学習
S50
メッセージを1つ読み込む
S51
メッセージパターンを生成
S52
同じ期間の頻度テーブルを選択
S53
パターン内のノイズをフィルタリング
S54
パターンの出現数を更新
S55
一定時間以内に障害発生？
NO
YES
S56
パターンの予兆数を更新
S57
NO
ログファイルの終端？
YES
S58
共起確率を算出
終了

フロントページの続き

(56)参考文献 特開２００６－００４３４６（ＪＰ，Ａ）
特開２００６－２６００５６（ＪＰ，Ａ）
特開２０１３－０３００９２（ＪＰ，Ａ）
国際公開第２０１３／０８８５６５（WO，Ａ1）
特開２０１２－１２３６９４（ＪＰ，Ａ）
渡辺　幸洋　外1名，「メッセージパターン学習による障害発生検知」，インターネットと運用技術シンポジウム２００９，情報処理学会，２００９年１２月，ｐｐ．２３－３０

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ１１／０７
Ｇ０６Ｆ１１／３０－１１／３４
Ｇ０５Ｂ２３／０２