(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6535809号
(P6535809)
(45)発行日　令和1年6月26日(2019.6.26)　(24)登録日　令和1年6月7日(2019.6.7)

(51)Int.Cl.　F I
H04L 12/70 (2013.01)　H04L 12/70 100Z

請求項の数 10 (全 21 頁)

(21)出願番号　特願2018-506692(P2018-506692)
(86)(22)出願日　平成28年3月24日(2016.3.24)
(86)国際出願番号　PCT/JP2016/059330
(87)国際公開番号　WO2017/163352
(87)国際公開日　平成29年9月28日(2017.9.28)
審査請求日　平成30年7月2日(2018.7.2)

(73)特許権者　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)代理人　110000279
特許業務法人ウィルフォート国際特許事務所
(72)発明者　橋本　恭佑
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
(72)発明者　藪崎　仁史
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
(72)発明者　木下　順史
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
最終頁に続く

(54)【発明の名称】異常検出装置、異常検出システム、及び、異常検出方法

(57)【特許請求の範囲】
【請求項１】
データフローの異常を検出する異常検出装置であって、プロセッサ及びメモリを有し、
前記プロセッサは、
複数のデータフローを、データフローのデータ量の時系列変化の類似性に基づいて分類し、
同じ分類に属する少なくとも２つのデータフローの間について、通常時における相関係数と、或るタイミングにおける相関係数とを算出し、
前記通常時における相関係数と前記或るタイミングにおける相関係数との差分が所定の閾値よりも大きい場合、前記少なくとも２つのデータフローの内の少なくとも何れかが異常であると判定する
異常検出装置。
【請求項２】
前記データフローとは、発信元から着信先へ通信ネットワークを介して流れるデータの流れである
請求項１に記載の異常検出装置。
【請求項３】
前記プロセッサは、データ量の時系列変化の周波数成分の特性が類似するデータフローを、同じ分類に属させる
請求項２に記載の異常検出装置。

【請求項４】
前記周波数成分の特性が類似するとは、所定の閾値以上の周波数成分を含む周波数帯域の少なくとも一部が重複することである
請求項３に記載の異常検出装置。
【請求項５】
データフローのデータ量の時系列変化に対して相関係数の算出対象の範囲として設定される対比時間は、同じ分類に属するデータフローにおいて共通である
請求項２に記載の異常検出装置。
【請求項６】
前記対比時間は、前記同じ分類に属するデータフローのデータ量の時系列変化に対して共通に設定される離散化幅の倍数として算出される
請求項５に記載の異常検出装置。
【請求項７】
前記共通に設定される離散化幅は、当該同じ分類に属するデータフロー毎にデータ量の時系列変化に基づいて算出した離散化幅のうち、最長の離散化幅である
請求項６に記載の異常検出装置。
【請求項８】
前記プロセッサは、データフローが異常であると判定した場合、当該異常を検出したタイミングと、当該データフローの発信元及び着信先の情報とを通知し、当該タイミングにおいて発生した障害内容の入力を受け付ける
請求項１に記載の異常検出装置。
【請求項９】
データフローの異常を検出する異常検出システムであって、分析装置及びネットワーク装置を有し、
前記分析装置は、
ネットワーク装置から複数のデータフローのデータ量の時系列変化の情報を収集し、
それら収集した複数のデータフローを、データフローのデータ量の時系列変化の類似性に基づいて分類し、
同じ分類に属する少なくとも２つのデータフローの間について、通常時における相関係数と、或るタイミングにおける相関係数とを算出し、
前記通常時における相関係数と前記或るタイミングにおける相関係数との差分が所定の閾値よりも大きい場合、前記少なくとも２つのデータフローの内の少なくとも何れかが異常であると判定する
異常検出システム。
【請求項１０】
データフローの異常を検出する計算機装置による異常検出方法であって、
複数のデータフローを、データフローのデータ量の時系列変化の類似性に基づいて分類し、
同じ分類に属する少なくとも２つのデータフローの間について、通常時における相関係数と、或るタイミングにおける相関係数とを算出し、
前記通常時における相関係数と前記或るタイミングにおける相関係数との差分が所定の閾値よりも大きい場合、前記少なくとも２つのデータフローの内の少なくとも何れかが異常であると判定する
異常検出方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、データの異常検知に関する。
【背景技術】
【０００２】

近年、クラウドコンピューティングシステム（以下「クラウドシステム」という）や仮想計算機の進展に伴い、アプリケーションの性能劣化による障害、及び、アプリケーションのバージョンアップデートに含まれるソースコードのバグによる障害など、いわゆるサイレント障害の検出が求められている。

【０００３】

特許文献１には、性能種目又は被管理装置を要素とし、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し、この相関関数に基づいて相関モデルを生成し、この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モデル生成部と、各要素間の各相関モデルを順次探索して最適な相関モデルを決定し、この決定された相関モデルに基づいて第１の要素の性能情報から第２の要素の性能情報を予測するモデル探索部を含む、運用管理装置が開示されている。

【先行技術文献】

【特許文献】

【０００４】

【特許文献１】米国特許出願公開第２００９／０２１６６２４号明細書

【発明の概要】

【発明が解決しようとする課題】

【０００５】

しかし特許文献１の場合、クラウドのようにデータフローの通信量が大きくなると、必要な計算量及び計算資源量も大きくなり、また計算時間も長くなる。故に、データフローの組に対する相関係数の計算量はさらに大きくなる。そこで本発明の目的は、データの異常検出における相関分析の処理負荷を低減することにある。

【課題を解決するための手段】

【０００６】

一実施例に係る、データフローの異常を検出する異常検出装置は、プロセッサ及びメモリを有する。

当該プロセッサは、

複数のデータフローを、データフローのデータ量の時系列変化の類似性に基づいて分類し、

同じ分類に属する少なくとも２つのデータフローの間について、通常時における相関係数と、或るタイミングにおける相関係数とを算出し、

通常時における相関係数と前記或るタイミングにおける相関係数との差分が所定の閾値よりも大きい場合、前記少なくとも２つのデータフローの内の少なくとも何れかが異常であると判定する。

【発明の効果】

【０００７】

本発明によれば、データの異常検出における相関分析の処理負荷を低減することができる。

【図面の簡単な説明】

【０００８】

【図１】本実施例に係るデータセンタの構成例を示す図。

【図２】ネットワーク装置の構成例を示す図。

【図３】分析システムの構成例を示す図。

【図４】フロー情報テーブルの構成例を示す図。

【図５】フロー特性テーブルの構成例を示す図。

【図６】フロー群情報テーブルの構成例を示す図。

【図７】相関情報テーブルの構成例を示す図。

【図８】異常情報テーブルの構成例を示す図。

【図９】通信量テーブルの構成例を示す図。

【図１０】フロー群生成処理の一例を示すシーケンスチャート。
【図１１】フロー群生成処理の一例を示すフローチャート。
【図１２】異常フロー検出処理の一例を示すシーケンスチャート。
【図１３】異常フロー検出処理の一例を示すフローチャート。
【発明を実施するための形態】
【０００９】
以下、図面を参照しながら実施例を説明する。なお、要素の数等（個数、数値、量、範囲等を含む）に言及する場合、特に明示した場合及び原理的に明らかに特定の数に限定される場合などを除き、その特定の数に限定されるものではなく、特定の数以上でも以下でも良い。また、各情報の内容を説明する際に、「識別情報」、「識別子」、「名」、「名前」、「ＩＤ」という表現を用いることがあるが、これらについてはお互いに置換が可能である。また、構成要素（要素ステップなどを含む）は、特に明示した場合及び原理的に明らかに必須であると考えられる場合などを除き、必ずしも必須のものではない。また、「ｘｘｘテーブル」又は「ｘｘｘリスト」の表現にて情報を説明することがあるが、情報は、どのようなデータ構造で表現されていてもよい。すなわち、情報がデータ構造に依存しないことを示すために、「ｘｘｘテーブル」又は「ｘｘｘリスト」を「ｘｘｘ情報」と呼ぶことができる。また、「プログラム」を主語として処理を説明する場合があるが、プログラムは、プロセッサ（例えばＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ））によって実行されることで、定められた処理を、適宜に記憶資源（例えばメモリ）及び通信インターフェイスデバイスのうちの少なくとも１つを用いながら行うため、処理の主語が、プロセッサ、そのプロセッサを有する装置とされてもよい。プロセッサが行う処理の一部又は全部が、ハードウェア回路で行われてもよい。コンピュータプログラムは、プログラムソースからインストールされてよい。プログラムソースは、プログラム配布サーバ又は記憶メディア（例えば可搬型の記憶メディア）であってもよい。また、以下の説明では、同種の要素を区別して説明する場合には、「計算機５０－１」、「計算機５０－２」のように、参照符号を使用し、同種の要素を区別しないで説明する場合には、「計算機５０」のように参照符号のうちの共通番号のみを使用することがある。
【００１０】
本実施例に係るシステムは、データフロー（以下単に「フロー」という場合がある）の通信量の時系列変化を相関分析し、その分析結果の相関係数が通常時（正常時）の相関係数と比べて所定よりも低い場合、当該フローを通常と異なる挙動を示した異常フローとして検出する。当該システムは、例えば、通常と異なる挙動を示したアプリケーションシステムなどを検出できる。当該システムは、異常フローを検出するにあたり、フロー通信量の時系列変化の特性（周期特性又は周波数成分特性など）が類似するフローを同じフロー群に分類する第１の処理と、同じフロー群に属するフロー同士で相関分析を行う第２の処理とを実行する。これにより、異常フローを検出するための相関分析において、フローの組合せ数を削減することができる。すなわち、相関分析の計算量を削減し、相関分析の処理に要する時間を短縮することができる。
【００１１】
また、本実施例に係るシステムは、フロー通信量に基づいて、相関分析の対象とする２つのデータフローの適切なウィンドウサイズ（対比時間）を算出する。クラウドシステムに流れるデータ通信量は膨大である為、サンプリングされて計測されることが多い。データ通信量がサンプリングされたパケット数から算出される場合、データ通信量が他のフローと比べて相対的に少ないフローはほとんどサンプリングされない。この場合、データ通信量の計測時間（離散化幅）を長くすることが考えられる。しかし、フロー離散化幅を長くすると、瞬間的な異常を検出しづらくなる。したがって、本実施例では、データ通信量に基づいて、フロー毎の適切な離散化幅（フロー離散化幅）を算出する。例えば、データ通信量が小さい場合はフロー離散化幅を長く、データ通信量が大きい場合はフロー離散化幅を短くする。これにより、データ通信量が比較的大きいフローの瞬間的な異常と、データ通信量が比較的小さいフローの長時間に渡る異常との何れも検出することができる。

【００１２】
　また、本実施例では、各フローのフロー離散化幅に基づいて、フロー群に属するフローに共通の離散化幅（フロー群離散化幅）を算出する。各フローのフロー離散化幅がばらばらであると、相関分析の対象とする少なくとも２つのフローの離散化幅を一致させる処理が必要となる。すなわち、相関分析の対象とするフローの組み合せ毎に、フロー離散化幅を一致させるための計算処理が必要となる。本実施例は、フロー群に属する各フローに対して共通のフロー群離散化幅を設定する。これにより、フローの組み合わせ毎にフロー離散化幅を一致させるための計算処理を省略することができ、相関分析に要する処理時間を短縮することができる。
【００１３】
　また、本実施例では、異常を検出したフローに関する情報を管理者に通知する。当該フローに関する情報は、例えば、当該フローの５タプル及び／又は仮想ネットワークＩＤ（ＶＬＡＮタグなど。以下同じ）などの情報である。これにより、管理者は、通知されたフローの情報から、通常と異なる挙動を示した機能及び機器などを特定することができる。
【００１４】
　なお、本実施例において、フローは、データ通信のパケットヘッダに含まれる、着信先ＭＡＣアドレス、発信元ＭＡＣアドレス、着信先ＩＰアドレス、発信元ＩＰアドレス、Ｌ４ポート番号、及び、仮想ネットワークＩＤによって一意に決まるデータ通信であってよい。又は、フローは、着信先ＩＰアドレス、発信元ＩＰアドレス、Ｌ４ポート番号、及び、仮想ネットワークＩＤによって一意に決まるデータ通信であってよい。又は、フローは、着信先ＩＰアドレス、発信元ＩＰアドレス、及び、仮想ネットワークＩＤによって一意に決まるデータ通信であってよい。
【００１５】
　図１は、本実施例に係るデータセンタの構成例を示す。データセンタは、管理システム１０、分析システム１００、制御ネットワーク２１、複数のネットワーク装置３０、及び、複数の計算機５０を含む。複数のネットワーク装置３０及び複数の計算機５０は、通信ネットワークで接続されたデータネットワーク３を構成してよい。データネットワーク３は、制御ネットワーク２１に接続されてよい。管理システム１０及び／又はネットワーク装置３０は、仮想的に実装されてもよい。
【００１６】
　ネットワーク装置３０は、計算機５０に仮想的に実装されてもよい。ネットワーク装置３０及び分析システム１００の詳細については、それぞれ図２、図３を用いて後述する。
【００１７】
　管理システム１０は、管理者が、顧客システムを構成するデータネットワーク３を管理するために使用するシステムである。管理システム１０は、所定のネットワーク２０を介して、分析システム１００と接続されている。管理システム１０は、分析システム１００から送信された各種情報を管理者へ提示してよい。例えば、管理システム１０は、分析システム１０から送信された異常フローの情報を管理者に通知する。管理者は、その通知された異常フローの情報に基づいて、顧客システムにおいて発生した異常を分析してよい。また、管理者は、異常フローが検出されたときに顧客システムにおいて発生した異常の内容を、管理システム１０のＧＵＩを介して、分析システム１００に登録してもよい。また、管理者は、管理システム１０を介して、過去に発生した顧客システムの異常と、そのときに通知された異常フローの情報との対応関係を参照できてよい。
【００１８】
　データネットワーク３は、顧客システム毎に論理的に分離されていてよい。例えば、１つのデータネットワーク３が、１つの顧客システムであってよい。顧客システムとは、少なくとも１つのアプリケーションによって構成される顧客毎のアプリケーションシステムであってよい。例えば、データセンタを利用する企業毎に１つの顧客システムが構成されてよい。データネットワーク３のプロトコルの例は、ネイティブなＩＰ通信である。
【００１９】

制御ネットワーク２１は、ネットワーク装置３０と分析システム１００とを接続するネットワークである。各データネットワーク３のデータは、当該制御ネットワーク２１を介して、分析システム１００に収集されてよい。

【００２０】

計算機５０は、ＣＰＵ、メモリ及びストレージなどの計算資源を有し、顧客システムにおけるアプリケーションを実行する。アプリケーションは、例えば、ＷＥＢサーバ、アプリケーションサーバ，ＤＢ（Ｄａｔａｂａｓｅ）サーバ等のプログラムである。アプリケーションは、ＶＭ（Ｖｉｒｔｕａｌ　Ｍａｃｈｉｎｅ）内に実装されてもよい。

【００２１】

図２は、ネットワーク装置３０の構成例を示す。ネットワーク装置３０は、例えば、ルータやスイッチ等によって実現される通信装置である。ネットワーク装置３０は、機能として、スイッチ３１、スイッチ管理部３２、フロー統計管理部３３、転送部３４、ポート３５、及び、管理ポート３６を有してよい。

【００２２】

スイッチ３１は、ポート３５から受信した通信パケットを、当該通信パケットのヘッダ情報に適合する出力先ポートに転送する、イーサネット（登録商標）ファブリックのスイッチであってよい。

【００２３】

スイッチ管理部３２は、スイッチ３１を管理する。スイッチ管理部３２は、例えば管理用端末から送信されるデータ参照要求や設定要求などを処理してよい。管理用端末とやり取りするプロトコルは、例えば、ＳＮＭＰ（Ｓｉｍｐｌｅ　Ｎｅｔｗｏｒｋ　Ｍａｎａｇｅｍｅｎｔ　Ｐｒｏｔｏｃｏｌ）、ｓＦｌｏｗなどである。

【００２４】

フロー統計管理部３３は、ネットワーク装置３０が受信した通信パケットのフロー毎の通信量又は通信パケット数をカウントする。フロー統計管理部３３は、ｓＦｌｏｗのプロトコルに対応してもよい。

【００２５】

転送部３４は、フロー統計管理部３３がカウントした値（計測値）を、分析システム１００へ送信する。

【００２６】

ポート３５は、計算機５０との間で通信パケットを送受信する為の物理ポートである。

【００２７】

管理ポート３６は、例えば管理用端末との間でデータを送受信する為の物理ポートである。また、管理ポート３６は、フロー統計管理部３３の計測値を分析システム１００へ送信するための物理ポートである。

【００２８】

図３は、分析システム１００の構成例を示す。分析システム１００は、データネットワーク３におけるデータフロー（データ通信量）を分析するためのシステムである。分析システム１００は、ＣＰＵ１５０、通信Ｉ／Ｆ１３０、入力Ｉ／Ｆ１４０、メモリ１１０、及び、ストレージ１２０などを備える計算機によって構成されてよい。メモリ１１０は、例えば、ＤＲＡＭ（Ｄｙｎａｍｉｃ　Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）、ＦｅＲＡＭ（Ｆｅｒｒｏｅｌｅｃｔｒｉｃ　Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）、ＭＲＡＭ（Ｍａｇｎｅｔｏｒｅｓｉｓｔｉｖｅ　Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）などである。ストレージ１２０は、例えば、ＳＳＤ（ｓｏｌｉｄ　ｓｔａｔｅ　ｄｒｉｖｅ，）、ＨＤＤ（Ｈａｒｄ　Ｄｉｓｋ　Ｄｒｉｖｅ）などである。

【００２９】

入力Ｉ／Ｆ１４０は、分析システム１００に接続される管理システム１０の操作画面等を介して、検出した異常なフローを管理者に通知したり、管理者から障害内容の入力を受け付けたりする為の（ノースバンド）インタフェースである。

【００３０】

通信Ｉ／Ｆ１０３は、ネットワーク装置３０から計測結果を受信するための（サウスバンド）インタフェースである。

【００３１】

メモリ１１０には、機能として、フロー群生成部１１１、相関算出部１１２、及び、異常検出部１１３が格納されてよい。これらの機能は、ストレージ１２０に保持されているプログラムがメモリ１１０に読み出されてＣＰＵ１５０に実行されることにより、実現されてよい。プログラムは、予めストレージ１２０に格納されてもよいし、所定のネットワークを介して又は可搬型記憶媒体を介して外部からインストールされても良い。なお、これらの機能１１１、１１２、１１３をまとめてフロー分析部と呼んでもよい。

【００３２】

ストレージ１２０には、データとして、フロー情報テーブル１２１、フロー特性テーブル１２２、フロー群情報テーブル１２３、相関情報テーブル１２４、異常情報テーブル１２５、通信量テーブル１２６、及び、離散化後通信量テーブル１２７が格納されてよい。

【００３３】

以下、各テーブルについて説明する。なお、以下のテーブルは一例であり、各テーブルは、複数のテーブルとして正規化されものであってもよいし、他のテーブルと結合されたものであってもよい。

【００３４】

図４は、フロー情報テーブル１２１の構成例を示す。フロー情報テーブル１２１は、フローに関する情報（「フロー情報」という）を管理する。

【００３５】

フロー情報テーブル１２１は、データ項目として、フローＩＤ２００、着信先ＩＰアドレス２０１、発信元ＩＰアドレス２０２、着信先ＭＡＣアドレス２０３、発信元ＭＡＣアドレス２０４、着信先ポート番号２０５、発信元ポート番号２０６、トランスポート層２０７、ネットワーク層２０８、及び、仮想ネットワークＩＤ２０９を有してよい。

【００３６】

フローＩＤ２００は、データネットワーク３を流れるフローを一意に識別するための値である。フローＩＤ２００は、ネットワーク装置３０のフロー統計管理部３３によって付与されてよい。

【００３７】

着信先ＩＰアドレス２０１は、フローＩＤ２００のフローの着信先のＩＰアドレスを示す。発信元ＩＰアドレス２０２は、フローＩＤ２００のフローの発信元のＩＰアドレスを示す。

【００３８】

着信先ＭＡＣアドレス２０３は、フローＩＤ２００のフローの着信先のＭＡＣアドレスを示す。発信元ＭＡＣアドレス２０４は、フローＩＤ２００のフローの発信元のＭＡＣアドレスを示す。

【００３９】

着信先ポート番号２０５は、フローＩＤ２００のフローの着信先のポート番号を示す。発信元ポート番号２０６は、フローＩＤ２００のフローの発信元のポート番号を示す。

【００４０】

トランスポート層２０７は、フローＩＤ２００のフローのトランスポート層の種類（ＴＣＰ、ＵＤＰなど）を示す。

【００４１】

ネットワーク層２０８は、フローＩＤ２００のフローのネットワーク層の種類（ＩＰｖ４、ＩＰｖ６、ＩＣＭＰ（Ｉｎｔｅｒｎｅｔ Ｃｏｎｔｒｏｌ Ｍｅｓｓａｇｅ Ｐｒｏｔｏｃｏｌ）など）を示す。

【００４２】

仮想ネットワークＩＤ２０９は、フローＩＤ２００のフローが属する仮想ネットワークのＩＤを示す。

【００４３】
これらの情報は、フローを構成するＩＰパケットのヘッダ情報から判明する。なお、フロー情報テーブル１２１の１つのレコードは、１つのＩＰパケットから判明した情報であってよい。すなわち、フロー情報テーブル１２１には、同じフローＩＤ２００を有する複数のエントリが存在してもよい。
【００４４】
図５は、フロー特性テーブル１２２の構成例を示す。フロー特性テーブル１２２は、フローの通信量の時系列変化の特性に関する情報（「フロー特性」という）を管理する。フロー特性テーブル１２２は、データ項目として、フローＩＤ３００、計測時間３１０、通信量平均３２０、通信量標準偏差３３０、フロー群ＩＤ３４０、フロー離散化幅３５０、及び、周波数成分３６０を有してよい。
【００４５】
フローＩＤ３００は、図４のフローＩＤ２００と同じである。
【００４６】
計測時間３１０は、フローＩＤ３００のフロー通信量の計測時間を示す。
【００４７】
通信量平均３２０は、フローＩＤ３００のフロー通信量の単位時間当たりの平均を示す。通信量平均３２０は、計測時間３１０内に計測されたフロー通信量から算出されてよい。
【００４８】
通信量標準偏差３３０は、フローＩＤ３００のフロー通信量の単位時間当たりの標準偏差を示す。フロー通信量の標準偏差３３０は、計測時間３１０内に計測されたフロー通信量から算出されてよい。
【００４９】
フロー群ＩＤ３４０は、フロー群を一意に識別するための番号である。同じフロー群ＩＤ３４０を有するフローＩＤ３００のフローは、同じフロー群に属する。フローＩＤ３００が分類されるフロー群は、計測時間３１０、通信量平均３２０及び通信量標準偏差３３０に基づいて決定されてよい。分類方法の詳細については後述する。
【００５０】
フロー離散化幅３５０は、フローＩＤ３００のフローの離散化幅（時間）を示す。フロー離散化幅３５０は、フロー間の相関係数を算出する際に用いられる。フロー離散化幅３５０の初期値は、管理者によって設定されてもよい。フロー離散化幅３５０の算出方法の詳細については後述する。
【００５１】
周波数成分３６０は、フローＩＤ３００のフロー通信量の時系列変化の周波数成分を示す。周波数成分３６０には、所定の閾値以上の周波数成分を含む周波数帯域が格納されてもよい。周波数成分３６０の算出方法については後述する。
【００５２】
図６は、フロー群情報テーブル１２３の構成例を示す。フロー群情報テーブル１２３は、フロー群に関する情報を管理する。フロー群情報テーブル１２３は、データ項目として、フロー群ＩＤ４００、フロー群離散化幅４１０、及び、ウィンドウサイズ４２０を有してよい。
【００５３】
フロー群ＩＤ４００は、図５のフロー群ＩＤ３４０と同じである。フロー群離散化幅４１０は、フロー群ＩＤ４００のフロー群に対する離散化幅を示す。ウィンドウサイズ４２０は、フロー群ＩＤ４００のフロー群に対するウィンドウサイズを示す。
【００５４】
フロー群ＩＤ４００に属する全てのフローＩＤ３００のフローには、共通のフロー群離散化幅４１０及びウィンドウサイズ４２０が適用される。相関係数の算出対象とされるウィンドウサイズ（対比時間）は、フロー群離散化幅４１０の所定の倍数として算出されて

よい。
【００５５】
したがって、分析システム１００の相関算出部１１２は、同じフロー群ＩＤに属するフローＩＤ間の相関係数を算出する際、フロー群情報テーブル１２３において当該フロー群ＩＤに対応付けられているウィンドウサイズ４２０（対比時間）を用いればよい。つまり、本実施例によれば、相関係数を算出する毎に離散化幅を一致させる必要がなくなる。
【００５６】
図７は、相関情報テーブル１２４の構成例を示す。相関情報テーブル１２４は、相関分析の結果に関する情報を管理する。相関情報テーブル１２４は、データ項目として、フローＩＤ５００、対フローＩＤ５０１、相関係数５０２、相関係数算出回数５０３、相関係数平均５０４、相関係数標準偏差５０５、及び、相関係数変化時刻５０６を有してよい。
【００５７】
フローＩＤ５００、及び、対フローＩＤ５０１は、図４のフローＩＤ２００と同じである。
【００５８】
相関係数５０２は、フローＩＤ５００のフローと、対フローＩＤ５０１のフローとの間の相関係数を示す。フローＩＤ５００及び対フローＩＤ５０１は同じフロー群に属する。したがって、当該相関係数５０２は、フロー群情報テーブル１２３において、当該フローＩＤ５００及び対フローＩＤ５０１が属するフロー群ＩＤ４００に対応付けられているウィンドウサイズ４２０を用いて算出された値である。
【００５９】
相関係数算出回数５０３は、相関係数５０２を算出した回数を示す。
【００６０】
相関係数平均５０４は、相関係数５０２の平均を示す。すなわち、相関係数平均５０４は、元の相関係数平均５０４に、今回算出した相関係数５０２を含めたときの平均である。つまり、相関係数平均５０４は、相関係数５０２を算出する毎に更新されてよい。
【００６１】
相関係数標準偏差５０５は、相関係数５０２の標準偏差を示す。すなわち、相関係数標準偏差５０５は、元の相関係数標準偏差５０５に、今回算出した相関係数を含めたときの標準偏差である。つまり、相関係数標準偏差５０５は、相関係数５０２を算出するごとに更新されて良い。
【００６２】
相関係数変化時刻５０６は、相関係数５０２に顕著な変化が発生した時刻（タイミング）である。例えば、相関係数５０２と相関係数平均５０４との差分が所定の閾値よりも大きい場合の、当該相関係数５０２に係るフローＩＤ５００又は対フローＩＤ５０１が検出された時刻である。相関係数変化時刻５０６は、相関係数５０２に顕著な変化が発生していない場合は空白（ＮＵＬＬ）であってよい。
【００６３】
図８は、異常情報テーブル１２５の構成例を示す。異常情報テーブル１２５は、異常と検出されたフロー（異常フロー）に関する情報を管理する。異常情報テーブル１２５は、データ項目として、フローＩＤ６００、対フローＩＤ６０１、異常内容６０２、異常継続時間６０３、及び、異常改善方法６０４を有してよい。
【００６４】
フローＩＤ６００、及び、対フローＩＤ６０１は、異常と検出されたフローＩＤである。フローＩＤ６００及び対フローＩＤ６０１は、相関情報テーブル１２４の相関係数変化時刻５０６に時刻が格納されている、フローＩＤ５００及び対フローＩＤ５０１であってもよい。
【００６５】
異常内容６０２は、フローＩＤ６００及び対フローＩＤ６０１と関連付けられる、顧客システムにおいて発生した異常の内容を示す。

【００６６】
異常継続時間６０３は、顧客システムにおいて、異常内容６０２の異常が継続した時間を示す。
【００６７】
異常改善方法６０４は、顧客システムにおける、異常内容６０２の異常に対する改善方法の情報を示す。
【００６８】
異常内容６０２には、相関情報テーブル１２４において、フローＩＤ６００及び対フローＩＤ６０１と対応する相関係数変化時刻５０６において、顧客システムで発生した異常の内容が格納されてよい。
【００６９】
異常内容６０２、異常継続時間６０３、及び／又は、異常改善方法６０４は、管理者によって入力されてよい。例えば、分析システム１００が相関係数変化時刻５０６を、管理システム１０を介して管理者へ提示し、管理者に、その相関係数変化時刻において顧客システムで発生した異常の内容、その異常が継続した時間、及び／又は、その異常に対する改善方法を入力してもらってもよい。
【００７０】
図９は通信量テーブル１２６の構成例を示す。通信量テーブル１２６は、各フローの各時刻におけるデータ通信量を管理する。通信量テーブル１２６は、データ項目として、フローＩＤ７００、時刻７０１、及び、通信量７０２を有してよい。
【００７１】
フローＩＤ７００は、図４のフローＩＤ２００と同じである。
【００７２】
時刻７０１は、フローＩＤ７００のフローの通信量７０２が計測された時刻である。時刻７０１は、分析システム１００がネットワーク装置３０から通信量の情報を受領した時刻であっても良いし、ネットワーク装置３０が当該通信量を計測した時刻であってもよい。
【００７３】
通信量７０２は、フローＩＤ７００のフローの、時刻７０１における通信量である。通信量７０２は、ネットワーク装置３０が、実際に計測した値であっても良いし、サンプリングしたデータ（パケット）から算出した値であってもよい。
【００７４】
なお、離散化後通信量テーブル１２７の有するデータ項目は、図９の通信量テーブル１２６と同じであってよい。よって、離散化後通信量テーブル１２７の図面については省略する。
【００７５】
図１０は、フロー群の生成処理の一例を示すシーケンスチャートである。フロー群の生成処理は、分析システム１００の導入時、定期的、アプリケーションの新規デプロイや構成時、又は、所定のイベント発生時などに実行されてよい。図１０は、計算機５０－１が計算機５０－２へ送信したデータの通信量をネットワーク装置３０が計測し、分析システム１００がその計測結果に基づいてフロー群を生成する処理の例である。
【００７６】
（ステップ１０００）計算機５０－１は、着信先を計算機５０－２とするデータを、ネットワーク装置３０へ送信する。当該データは、ＩＰパケットであってよい。
【００７７】
（ステップ１０１０）ネットワーク装置３０は、発信元の計算機５０－１から送信されたデータを、着信先の計算機５０－２へ転送する。
【００７８】
（ステップ１０２０）ネットワーク装置３０は、転送データのフロー通信量を計測し、当該フローの情報及び計測結果を分析システム１００へ送信する。フロー情報は、転送デ

ータ（ＩＰパケット）のヘッダに含まれる情報（すなわちフロー情報テーブル１２０のデータ項目に対応する値）であってよい。フローの計測結果は、サンプリングに基づく統計情報（例えば計測時間３１０、通信量平均３２０、通信量標準偏差３３０）であってよい。ネットワーク装置３０は、当該ステップ１０２０の処理を、データ転送毎に実行しても良いし、定期的に実行しても良いし、データ転送回数が所定回数に達する毎に実行してもよい。なお、フローＩＤは、ネットワーク装置３０によって付与されても良いし、分析システム１００によって付与されてもよい。ネットワーク装置３０は、ｓＦｌｏｗプロトコルに従って、フローの計測結果を分析システム１００へ送信してよい。

【００７９】

（ステップ２０１０）分析システム１００は、フロー群生成処理を実行する。次に当該処理を説明する。

【００８０】

図１１は、フロー群生成処理の例を示すフローチャートである。本処理は、図１０のステップ２０１０の処理に相当する。

【００８１】

（ステップ５０１０）フロー群生成部１１１は、各フローの通信量を算出する。フロー群生成部１１１は、各フローＩＤについて、次の（Ａ１）乃至（Ａ４）の処理を実行してよい。

【００８２】

（Ａ１）フロー群生成部１１１は、フロー情報テーブル１２１から、フローＩＤ２００が一致するエントリを数える。

【００８３】

（Ａ２）フロー群生成部１１１は、そのエントリ数に基づいて、フローＩＤのフローのパケット数を算出する。パケット数は、「ネットワーク装置３０におけるサンプリングレート×エントリ数」として算出されてよい。サンプリングレートは、ネットワーク装置３０及び分析システム１００に初期設定されてよい。

【００８４】

（Ａ３）フロー群生成部１１１は、パケット数と、平均パケット長と、計測時間とに基づいて、フローＩＤの通信量を算出する。通信量は、「パケット数×平均パケット長／計測時間」として算出されてよい。平均パケット長及び計測時間は、ネットワーク装置３０及び分析システム１００に初期設定されても良いし、ネットワーク装置３０によって計測されても良い。

【００８５】

（Ａ４）フロー群生成部１１１は、フローＩＤと、ステップ１０２０で計測結果を受領した時刻と、その算出した通信量と、を対応付けて通信量テーブル１２６へ格納する。なお、計測結果を受領した時刻は、ネットワーク装置３０がデータを受信した時刻であってもよい。

【００８６】

（ステップ５０１５）フロー群生成部１１１は、各フローの通信量平均３２０及び通信量標準偏差３３０を算出する。フロー群生成部１１１は、各フローＩＤについて、次の（Ｂ１）乃至（Ｂ２）の処理を実行してよい。

【００８７】

（Ｂ１）フロー群生成部１１１は、通信量テーブル１２６から、フローＩＤ７００が同じエントリを抽出する。そして、フロー群抽出部１１１は、その抽出したエントリの時刻７０１から、最古の時刻と最新の時刻を特定する。

【００８８】

（Ｂ２）フロー群生成部１１１は、フロー特性テーブル１２２の上記（Ｂ１）で特定したフローＩＤに対応する計測時間３１０に、最古の時刻から最新の時刻までの時間を格納する。フロー群抽出部１１１は、フロー特性テーブル１２２の上記（Ｂ１）で特定したフローＩＤに対応する通信量平均３２０及び通信量標準偏差３３０に、上記（Ｂ１）で抽出

した通信量７０２から算出した平均及び標準偏差を格納（上書き）する。

【００８９】

（ステップ５０２０）フロー群生成部１１１は、各フローのフロー離散化幅３５０及び周波数成分３６０を算出する。以下、フロー離散化幅３５０及び周波数成分３０６の算出方法を説明する。

【００９０】

（フロー離散化幅３５０の算出方法）

フロー毎にフロー離散化幅を算出する理由は次の通りである。フローの通信量が非常に小さい場合、そのフローに対してサンプリングされるパケット数も少ない。したがって、その少数のサンプリングされたパケット数に基づいて上記（Ａ３）のように通信量を算出するにあたり、サンプリングされるパケット数が少し増減するだけで、算出される通信量が大きく変動してしまう。この場合、検出される通信量の変動が、実際に通信量の増減によるものか（つまり有意な変動であるのか）、それとも、サンプリングされたパケット数がたまたま増減しただけなのか（つまり無意な変動であるのか）を判断することができない。

【００９１】

そこで、本実施例では、各フローの通信量の大きさに基づいて、各フローの適切な（統計的な信頼度が所定以上となる）サンプリング時間を算出する。このサンプリング時間を、「フロー離散化幅」と呼ぶ。フロー離散化幅３５０は、「分析可能通信量／通信量平均」として算出されてよい。この分析可能通信量は、所定値であってよい。この通信量平均は、フロー特性テーブル１２２においてフローＩＤと対応付けられている通信量平均３２０であってよい。

【００９２】

（周波数成分３６０の算出方法）

フローが異常か否かは、例えば次のように判定する方法が考えられる。すなわち、計測されたフローの全ての組み合わせについて、それぞれ、通常時（正常時）における通信量の時系列変化に係る相関係数を算出しておく。そして、全ての組み合わせについて相関係数を算出し、その算出した相関係数と通常時の相関係数との差分が所定よりも大きい場合、当該組み合わせに係るフローを異常と判定する。

【００９３】

しかし、フロー数が多くなるとフローの組み合わせ数が膨大となり、全ての組み合わせについて相関係数を算出することが困難となる。そこで、本実施例では、通常時（正常時）における通信量の時系列変化の特性が類似するフローを同じフロー群に分類しておく。そして、フロー群に属するフローの組み合わせについて相関係数を算出し、その算出した相関係数と通常時の相関係数とを比較することにより、フローの異常を判定する。これにより、フローの組み合わせ数が少なくなるので、相関係数の算出に要する処理負荷を低減することができる。周波数成分は、各フローをフロー群に分類する際に用いられる指標である。以下、各フローをフロー群に分類する方法を説明する。

【００９４】

例えば、フローの通信量の時系列変化の特性（「フロー特性」）を、（Ｃ１）非定常かつ規則性のあるフロー特性（以下「周期性の高いフロー特性」という）、（Ｃ２）定常的なフロー特性、（Ｃ３）非定常かつ不規則なフロー特性（以下「周期性の低いフロー特性」という）に分類することができる。フロー特性が類似する場合、相関関係も高くなる可能性が高い。反対に、フロー特性が類似しない場合、相関関係も低くなる可能性が高い。周期の特性は周波数の特性として表現が可能であるので、「周期性の高い」は「特定の周波数成分が強い」と表現することができる。以下、（Ｃ１）乃至（Ｃ３）について説明する。

【００９５】

（Ｃ１）周期性の高いフロー特性同士では、周波数成分（周期）と位相とが類似するほど、相関係数が高くなる可能性が高い。

【００９６】
（Ｃ２）定常的なフロー特性は、周期が非常に大きく且つ振幅が非常に小さい周期性の高いフロー特性と表現することもできる。振幅が非常に小さい為、位相のずれは、相関係数にあまり影響を与えない可能性が高い。したがって、定常的なフロー特性同士では、周波数成分が類似するほど、相関係数が高くなる可能性が高い。
【００９７】
（Ｃ３）周期性の低いフロー特性には、特徴的な周波数成分（周期）や位相は存在しない可能性が高い。例えば、アプリケーションシステムがユーザからのアクセス等のイベントを契機に送受信するデータは、周期性の低いフロー特性を有する可能性が高い。しかしながら、例えば、ＷＥＢの３階層モデルにおいて、ＷＥＢサーバからアプリケーションサーバへ送信されるデータと、アプリケーションサーバからＤＢサーバへ送信されるデータとは、連動している（同じようなタイミングで送信される）可能性が高い。このように、同一のイベントを契機に同じようなタイミングで送信されるデータは、パルス波の挙動に近いため、周期性は存在しないものの、ほぼ同一の周波数帯域に高い周波数成分を有する可能性が高い。
【００９８】
各フローを、上述の（Ｃ１）乃至（Ｃ３）のように分類しても良いが、もう少し緩い条件で分類してもよい。例えば、フロー特性の周波数成分のみを用いて分類してもよい。この分類方法は、上述の分類方法と比較して、同じ分類に相関関係の低いフローの組み合わせが存在する（フォールスポジティブの）可能性が高くなり、相関係数の算出処理の負荷が高くなるが、反対に、同じ分類に相関関係の高いフローの組み合わせが存在しない（フォールスネガティブの）可能性が低くなる。
【００９９】
次に、周波数成分の算出処理の一例を示す。
【０１００】
（Ｄ１）フロー群生成部１１１は、通信量テーブル１２６から、フローＩＤ７００が同じエントリを抽出する。
【０１０１】
（Ｄ２）フロー群生成部１１１は、その抽出した複数のエントリの時刻７０１を、当該フローＩＤに対応するフロー離散化幅３５０の間隔で分割する。そして、その分割した各エントリの通信量７０２の合計（又は平均）を算出する。例えば、フロー離散化幅３５０が「１分」の場合、その抽出した複数のエントリの時刻７０１を、１分間隔で分割する。そして、その分割した各１分間の通信量の合計（又は平均）を算出する。これにより、フロー離散化幅３５０で再計算された通信量の時系列データ（以下「離散化後フロー通信量」という）が生成される。
【０１０２】
（Ｄ３）フロー群生成部１１１は、上記（Ｄ２）で算出した離散化後フロー通信量に対して周波数解析を行い、周波数成分を算出する。
【０１０３】
（Ｄ４）フロー群生成部１１１は、フローＩＤ、フロー離散化幅に対応する時刻、離散化後フロー通信量を、離散化後通信量テーブル１２７（不図示）に格納する。
【０１０４】
（Ｄ５）フロー群生成部１１１は、フロー特性テーブル１２２において、フローＩＤ３００に対応する、フロー離散化幅３５０及び周波数成分３６０に、上述で算出したフロー離散化幅及び周波数成分を格納（上書き）する。フロー群生成部１１１は、全てのフローＩＤについて、上記（Ｄ１）乃至（Ｄ５）の処理を行う。
【０１０５】
（ステップ５０２５）フロー群生成部１１１は、フロー特性テーブル１２２から各フローの周波数成分３６０の大きい周波数帯域を特定する。例えば、フロー群抽出部１１１は、上位Ｎ（Ｎは正の整数）個の周波数成分が属する周波数帯域を特定してもよい。又は、

フロー群抽出部１１１は、所定の閾値以上の周波数成分が属する周波数帯域を特定してもよい。

【０１０６】

そして、フロー群抽出部１１１は、その特定した周波数帯域に基づいて各フローを、各フロー群に分類する。例えば、フロー群生成部１１１は、その特定した周波数帯域が、所定の閾値よりも大きい方に属しているか、それとも、小さい方に属しているか基づいて、各フローを２つのフロー群に分類してもよい。例えば、フロー群生成部１１１は、その特定した周波数帯域が、複数の異なる区間の何れに属するかに基づいて、各フローを複数のフロー群に分類しても良い。例えば、フロー群抽出部１１１は、特定した周波数帯域を属性として、Ｋ－ＭＥＡＮＳ法等の公知のクラスタリング手法によって、各フローを複数のフロー群に分類しても良い。

【０１０７】

そして、フロー群生成部１１１は、フロー群抽出部１２２において、同じフロー群に分類したフローＩＤ３００に対応するフロー群ＩＤ３４０に、共通のフロー群ＩＤを付与する。

【０１０８】

（ステップ５０３０）フロー群生成部１１１は、各フロー群について、フロー群離散化幅とウィンドウサイズとを算出する。フローの組み合わせについて相関係数を算出するためには、それらのフローの離散化幅が一致している必要があるからである。そこで本実施例では、以下のように、各フロー群に対してフロー群離散化幅を設定する。

【０１０９】

（Ｅ１）フロー群生成部１１１は、フロー特性テーブル１２２から、フロー群ＩＤが同じエントリを抽出する。

【０１１０】

（Ｅ２）フロー群生成部１１１は、その抽出したエントリのうち、最大のフロー離散化幅を特定する。

【０１１１】

（Ｅ３）フロー群生成部１１１は、その特定した最大のフロー離散化幅（フロー群離散化幅）に所定値を掛けて、ウィンドウサイズを算出する。この所定値は、予め設定された１以上の値であってよい。

【０１１２】

（Ｅ４）フロー群生成部１１１は、フロー群情報テーブル１２３において、フロー群ＩＤ３００に対応するフロー群離散化幅４１０及びウィンドウサイズ４２０に、それぞれ、（Ｅ３）で算出した最大のフロー離散化幅及びウィンドウサイズを格納（上書き）する。なお、フロー群生成部１１１は、上記（Ｅ１）のフロー群ＩＤがフロー群情報テーブル１２３に存在しない場合は、新規エントリを作成してよい。

【０１１３】

（ステップ５０３５）フロー群生成部１１１は、通信量テーブル１２６を用いて、ステップ５０２０の（Ｄ１）乃至（Ｄ５）と同様の手順で、各フローＩＤについて、フローＩＤの属するフロー群ＩＤ３４０のフロー群離散化幅４１０に対応する時刻、離散化後フロー通信量を算出し、離散化後通信量テーブル１２７（不図示）に格納（上書き）する。

【０１１４】

以上の処理により、データ量の時系列変化が類似するフローを、同じフロー群に分類することができる。また、フロー群に対して共通の、フロー群離散化幅及びウィンドウサイズを算出することができる。

【０１１５】

図１２は、異常フローの検出処理の一例を示すシーケンスチャートである。異常フローの検出処理は、随時実行されてよい。図１２は、計算機５０－１が計算機５０－２へ送信したデータの通信量をネットワーク装置３０が計測し、分析システム１００がその計測結果に基づいて、異常フローを検出する処理の例である。

【０１１６】
ステップ２０００からステップ２０２０までの各処理は、図１０のステップ１０００からステップ１０２０までの各処理と同じである。よって、ここでは説明を省略する。
【０１１７】
（ステップ２０３０）分析システム１００は、異常フロー検出処理を実行する。当該処理の詳細については後述する（図１３参照）。
【０１１８】
（ステップ２０４０）分析システム１００は、異常フローを検出した場合、当該異常フローに関する情報（フロー情報テーブル１２１のデータ項目、相関情報テーブル１２４の相関係数変化時刻５０６など）を管理システム１０へ送信する。
【０１１９】
（ステップ２０５０）管理者は、管理システム１０を介して、その通知された異常フローの発生時に顧客システムにおいて発生した異常内容などを入力する。管理システム１０は、この入力された異常内容などを、分析システム１００へ送信する。分析システム１００は、この送信された異常内容などを、異常情報テーブル１２５の異常フローＩＤに対応するエントリに格納する。これにより、異常フローと、顧客システムにおいて発生した異常内容などが対応付けられる。
【０１２０】
図１３は、異常フロー検出処理の一例を示すフローチャートである。本処理は、図１２のステップ２０３０の処理に相当する。
【０１２１】
（ステップ６０１０）相関算出部１１２は、処理対象のフロー群ＩＤを選択する。
【０１２２】
（ステップ６０２０）相関算出部１１２は、ステップ６０１０で選択したフロー群ＩＤを有する２つのフローＩＤ（フローＩＤ５００及び対フローＩＤ５０１）の間の相関係数を算出し、相関情報テーブル１２４の相関係数５０２に格納する。例えば、以下の（Ｆ１）乃至（Ｆ４）の処理により、相関係数を算出する。
【０１２３】
（Ｆ１）相関算出部１１２は、離散化後通信量テーブル１２７から、フローＩＤ５００及び対フローＩＤ５０１に対応するレコードを抽出する。例えば、フローＩＤ５００「Ｘ」の時刻「ｉ」における通信量を「Ｘｉ」、対フローＩＤ「Ｙ」の時刻「ｉ」における通信量を「Ｙｉ」とすると、フローＩＤ「Ｘ」と対フローＩＤ「Ｙ」との間の相関係数「ｒ」は下記の式（１）で算出される。
【０１２４】
【数１】

$$r=\frac{\frac{1}{N}\sum_{i=1}^{N}(Xi-\overline{X})(Yi-\overline{Y})}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Xi-\overline{X})^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Yi-\overline{Y})^2}}$$

・・・（１）
【０１２５】
ここで、「Ｎ（Ｎは正の整数）」は、離散化後通信量テーブル１２７におけるフローＩＤ「Ｘ」（又はフローＩＤ「Ｙ」）のエントリ数である。同じフロー群に属するフローは同じフロー群離散化幅で離散化されているので、フローＩＤ「Ｘ」及び「Ｙ」の当該エントリ数は同じ「Ｎ」となる。
【０１２６】
（Ｆ２）相関算出部１１２は、相関情報テーブル１２４の、フローＩＤ「Ｘ」及び対フローＩＤ「Ｙ」に対応する相関係数５０２に、その算出した相関係数「ｒ」を格納する。
【０１２７】
（Ｆ３）相関算出部１１２は、今回算出した相関係数「ｒ」を用いて、相関情報テーブ

ル１２４の過去に算出された相関係数平均値５０４、及び、相関係数標準偏差５０５を更新する。また、相関算出部１１２は、相関係数算出回数５０３をインクリメントする。

【０１２８】

（Ｆ４）相関算出部１１２は、上記（Ｆ１）乃至（Ｆ３）の処理を、ステップ６０１０で選択したフロー群ＩＤに属するフローＩＤの全ての組み合せについて実行する。

【０１２９】

相関算出部１１２は、上記（Ｆ１）乃至（Ｆ４）の処理を、全てのフロー群ＩＤについて実行する。これにより、相関係数の計算回数が、同じフロー群に属するフローの組み合せ数（フロー群に属するフロー数の２乗とフロー群数との積）となる。この計算回数は、全てのフローの組み合せ数（フロー数の２乗）よりも少ない。よって、本実施例によれば、相関係数の算出に要する計算リソース及び／又は計算時間を削減することができる。

【０１３０】

（ステップ６０３０）相関算出部１１２は、相関情報テーブル１２４における相関係数５０２と相関係数平均５０４との差分を算出し、当該差分が所定の閾値よりも大きいエントリを特定する。そして、相関算出部１１２は、それら特定したエントリのフローＩＤ５００及び対フローＩＤ５０１を、異常情報テーブル１２５のフローＩＤ６００及び対フローＩＤ６０１に格納する。なぜなら、相関係数が平均的な相関係数よりも大きく外れている場合（相関係数に顕著な変化がある場合）、その相関係数に係るフロー及び／又は対フローが異常である可能性が高いからである。なお、上記差分に対する所定の閾値は、相関係数の標準偏差に基づく閾値として定義されても良い。

【０１３１】

上述した実施例は、本発明の説明のための例示であり、本発明の範囲を実施例にのみ限定する趣旨ではない。当業者は、本発明の要旨を逸脱することなしに、他の様々な態様で本発明を実施することができる。

【符号の説明】

【０１３２】

３：データネットワーク　１０：管理システム　２１：制御ネットワーク　３０：ネットワーク装置　５０：計算機　１００：分析システム

【図１】

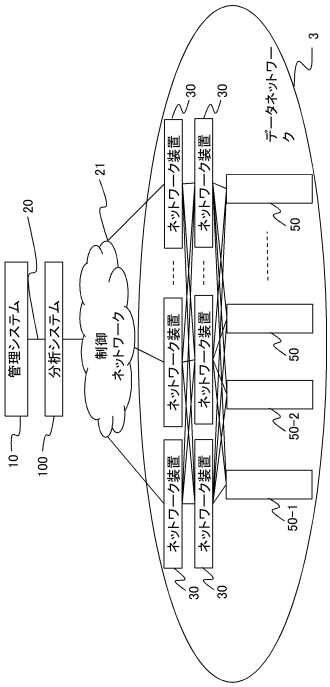

図1

【図２】

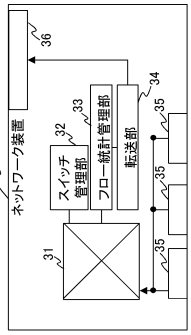

図2

【図３】

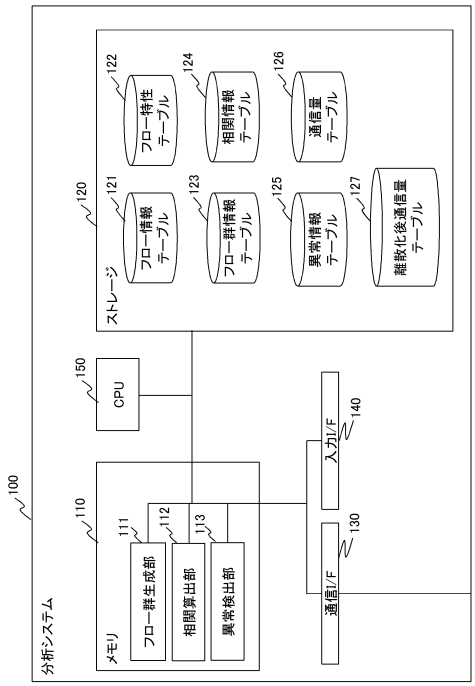

図3

【図４】

フロー情報テーブル 121

フローID 200	着信先IPアドレス 201	発信元IPアドレス 202	着信先MACアドレス 203	発信元MACアドレス 204	着信先ポート番号 205	発信元ポート番号 206	トランスポート層 207	ネットワーク層 208	仮想ネットワークID 209
1									
1									
2									
3									
：									

図4

【図５】

フロー特性テーブル 122

フローID (300)	計測時間 (310)	通信量平均 (320)	通信量標準偏差 (330)	フロー群ID (340)	フロー離散化幅 (350)	周波数成分 (360)
1				A		
2				B		
3				A		
：				：		

図5

【図６】

フロー群情報テーブル 123

フロー群ID (400)	フロー群離散化幅 (410)	ウィンドウサイズ (420)
A		
B		

図6

【図７】

相関情報テーブル 124

フローID (500)	対フローID (501)	相関係数 (502)	相関係数算出回数 (503)	相関係数平均 (504)	相関係数標準偏差 (505)	相関係数変化時刻 (506)
1	2					
：	3					
：	：					
：	N					
2	3					
：	4					
：	：					
：	N					
：	：					
N-1	N					

図7

【図８】

異常情報テーブル 125

フローID (600)	対フローID (601)	異常内容 (602)	異常継続時間 (603)	異常改善方法 (604)

図8

【図９】

通信量テーブル 126

700 フローID	701 時刻	702 通信量[bps]
1	20165031 1-140001	10
1	20165031 1-140002	11
1	20165031 1-140003	12
1	20165031 1-140004	11
2	20165031 1-140001	10

図9

【図１０】

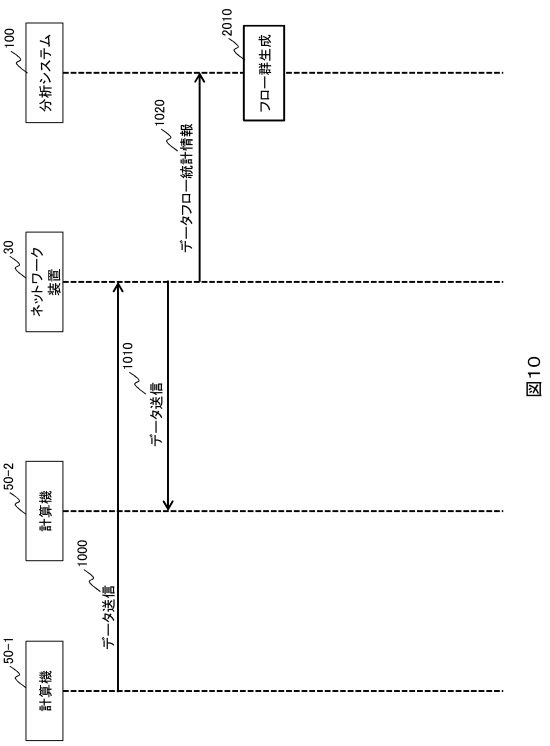

図10

【図１１】

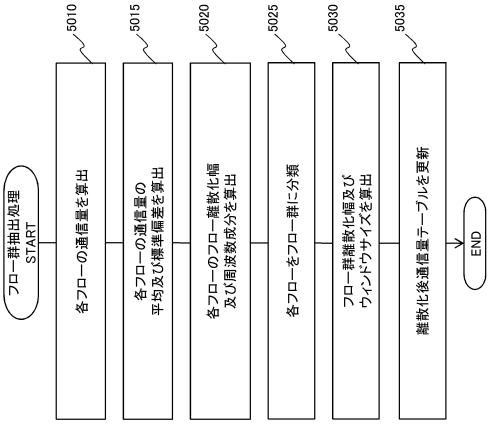

図11

【図１２】

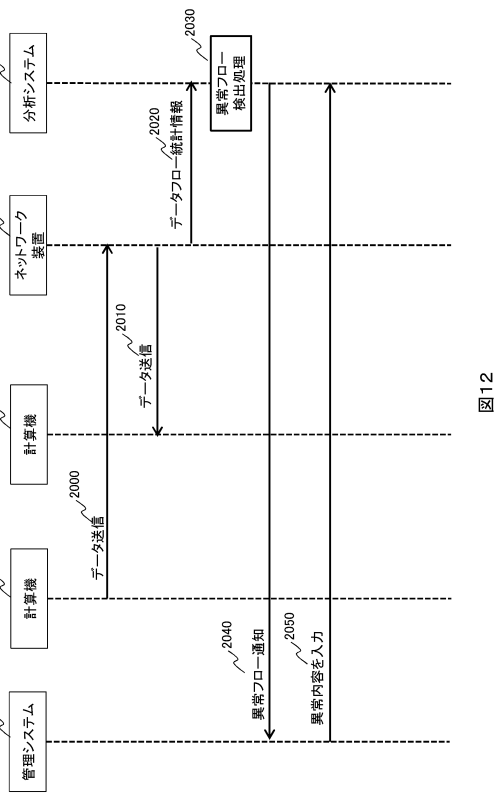

図12

【図１３】

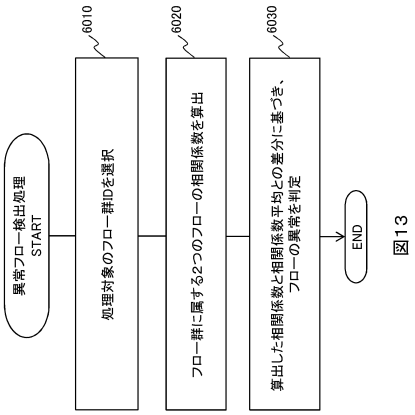
異常フロー検出処理
START
6010
処理対象のフロー群IDを選択
6020
フロー群に属する2つのフローの相関係数を算出
6030
算出した相関係数と相関係数平均との差分に基づき、
フローの異常を判定
END

図13

フロントページの続き

審査官　西村　純

(56)参考文献　特許第４１１２５８４（ＪＰ，Ｂ２）
特開平１１－１７７５４９（ＪＰ，Ａ）
特開２００６－１１５１２９（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｈ０４Ｌ　１２／００－１２／９５５