

SPECIFICATION

TITLE OF INVENTION

PROGRAM, ABNORMALITY DETECTION METHOD, AND SYSTEM

TECHNICAL FIELD

[0001] This invention relates to techniques for detecting persistent abnormalities in measured values.

BACKGROUND ART

[0002] When there are plural datasets each of which contains plural values of items related to a performance of a device to be analyzed, such as a ship, which were measured at a substantially same timing, a relationship between the items can be identified by known methods. Before identifying the relationship between the items, it is desirable that datasets that are likely to contain an abnormal measured value are excluded. For this purpose, one or more items are selected as items for filtering, datasets containing measured values of the items for filtering that are likely to be normal are extracted from the population datasets, and then the relationship between two or more items other than the items for filtering (hereinafter referred to as "items for analysis") in the extracted datasets is identified. The inventor of this invention invented a technique to exclude datasets that are likely to contain abnormal values from the population datasets with a minimal reduction in the number of datasets, without being influenced by the intuition, experience, etc. of the operator. The technology is described in Patent Document 1 (WO 2020/262038).

PRIOR ART DOCUMENT

PATENT DOCUMENT

[0003] Patent Document 1: WO 2020/262038

SUMMARY OF THE INVENTION

PROBLEM TO BE SOLVED BY THE INVENTION

[0004] In the method described in Patent Document 1, for each of the items of filtering, datasets containing measured values of the item of filtering within a predetermined range are extracted from the population datasets as a candidate group of datasets that are likely to contain normal values. Among the candidate groups of datasets each of which is extracted from the population datasets for one of the items of filtering groups, the group of datasets whose "filtering efficiency" is the largest is adopted as datasets after extraction. "Filtering efficiency" is a ratio of a decrease in an indicator of statistical variation of measured values of the items for analysis in the datasets before and after the extraction to a decrease in a number of datasets before and after the extraction.

[0005] The inventor of this invention conceived that it is possible to detect when persistent abnormalities are occurring in measurements of values of an item related to a performance of the device to be analyzed by use of the filtering efficiency.

[0006] This invention provides a means of detecting persistent abnormalities in measurements of values of an item related to a performance of a device to be analyzed by use of datasets each of which contains plural values of items measured at a substantially same timing.

MEANS FOR SOLVING THE PROBLEM

[0007] To solve the problem described above, a first aspect of this invention includes a program causing a computer to execute: when a group of datasets each of which contains a measured value of a first item related to a performance of a device to be analyzed and a measured value of a

second item affecting the performance of the device, the value of the first item and the value of the second item contained in each dataset being measured at a substantially same timing, a step for acquiring, for each of two or more different time periods, a group of datasets each of which contains values measured in the time period; a step for extracting, for each of the two or more different time periods, datasets that contain measured values of the second item within a range determined in accordance with a predetermined rule; a step for calculating, for each of the two or more different time periods, a ratio of a decrease in an indicator of statistical variation of measured values of the first item in datasets before and after the extraction to a decrease in a number of datasets before and after the extraction; and a step for detecting persistent abnormalities in measurements of values of the first item or the second item based on a comparison among the ratios each of which is calculated for one of the two or more different time periods.

[0008] A second aspect of this invention includes a program according to the first aspect, when the first item includes a plurality of sub items, causing the computer to execute: a step for specifying, for each of the two or more different time periods, a relational expression among the sub items of the first item based on the extracted datasets; and a step for calculating the indicator, for each of the datasets before and after the extraction for each of the two or more different time periods, that represents a statistical variation in difference between a value of the first item according to the relational expression and a measured value of the first item contained in each dataset.

[0009] A third aspect of this invention includes a method executed by a data processing device comprising: when a group of datasets each of which contains a measured value of a first item related to a performance of a device to be analyzed and a measured value of a second item affecting

the performance of the device, the value of the first item and the value of the second item contained in each dataset being measured at a substantially same timing, a step for acquiring, for each of two or more different time periods, a group of datasets each of which contains values measured in the time period, a step for extracting, for each of the two or more different time periods, datasets that contain measured values of the second item within a range determined in accordance with a predetermined rule; a step for calculating, for each of the two or more different time periods, a ratio of a decrease in an indicator of statistical variation of measured values of the first item in datasets before and after the extraction to a decrease in a number of datasets before and after the extraction; and a step for detecting persistent abnormalities in measurements of values of the first item or the second item based on a comparison among the ratios each of which is calculated for one of the two or more different time periods.

[0010] A fourth aspect of this invention includes a system comprising: when a group of datasets each of which contains a measured value of a first item related to a performance of a device to be analyzed and a measured value of a second item affecting the performance of the device, the value of the first item and the value of the second item contained in each dataset being measured at a substantially same timing, a means for acquiring, for each of two or more different time periods, a group of datasets each of which contains values measured in the time period; a means for extracting, for each of the two or more different time periods, datasets that contain measured values of the second item within a range determined in accordance with a predetermined rule; a means for calculating, for each of the two or more different time periods, a ratio of a decrease in an indicator of statistical variation of measured values of the first item in datasets before and after the extraction to a decrease in a

number of datasets before and after the extraction; and a means for detecting persistent abnormalities in measurements of values of the first item or the second item based on a comparison among the ratios each of which is calculated for one of the two or more different time periods.

EFFECTS OF THE INVENTION

[0011] According to this invention, persistent abnormalities in measurements of values of an item related to a performance of a device to be analyzed can be detected by use of datasets each of which contains plural values of items measured at a substantially same timing.

[0012] [FIG. 1] Figure 1 shows an example of a hardware configuration of a performance analyzing device according to prior art.

[FIG. 2] Figure 2 shows a functional configuration realized by the performance analyzing device according to prior art.

[FIG. 3] Figure 3 shows examples of measured values stored by the performance analyzing device according to prior art.

[FIG. 4] Figure 4 shows an example of an item selection screen displayed by the performance analyzing device according to prior art.

[FIG. 5] Figure 5 shows examples of stored datasets temporarily stored by the performance analyzing device according to prior art.

[FIG. 6] Figure 6 is a graph showing an example of a relational expression calculated by the performance analyzing device according to prior art.

[FIG. 7] Figure 7 shows an example of an operating flow in an extraction process executed by the performance analyzing device according to prior art.

[FIG. 8] Figure 8 is an example of graph showing a transition of a statistical variation and a number of datasets obtained as a result of extraction processes repeatedly performed by the performance

analyzing device according to prior art.

[FIG. 9] Figure 9 is a graph showing an example of datasets output by the performance analyzing device according to prior art.

[FIG. 10] Figure 10 is a graph showing an example of datasets output by the performance analyzing device according to prior art.

[FIG. 11] Figure 11 is a graph showing an example of datasets extracted by a more classical method than that of the performance analyzing device according to prior art.

[FIG. 12] Figure 12 is an example of graph showing a transition of a statistical variation and a number of datasets obtained as a result of extraction processes repeatedly performed by a modified example of the performance analyzing device according to prior art.

[FIG. 13] Figure 13 is an example of graph showing a transition of a statistical variation and a number of datasets obtained as a result of extraction processes repeatedly performed by a modified example of the performance analyzing device according to prior art.

[FIG. 14] Figure 14 shows a functional configuration realized by the abnormality detecting device according to an embodiment of this invention.

[FIG. 15] Figure 15 shows an example of a screen displayed by the abnormality detecting device according to the embodiment of this invention.

[FIG. 16] Figure 16 is a graph showing an example of changes in the number of datasets and changes in the indicator of statistical variation in the value of the item for analysis in datasets before and after an extraction of datasets performed by the abnormality detecting device according to an embodiment of this invention for each of the items for filtering, for each of different time periods.

[FIG. 17] Figure 17 shows examples of the filtering efficiency

calculated by the abnormality detecting device according to an embodiment of this invention.

[FIG. 18] Figure 18 is a graph showing an example of changes in the filtering efficiency for each of the items for filtering.

[FIG. 19] Figure 19 is a graph showing an example of changes in the filtering efficiency for each of the items for filtering.

MODES FOR CARRYING OUT THE INVENTION

[0013] A configuration and operation of an abnormality detecting device according to an embodiment of this invention are partially in common with those of the performance analyzing device described in Patent Document 1. Therefore, prior to a description of the abnormality detecting device according to the embodiment of this invention, the performance analyzing device described in Patent Document 1 is described below.

[0014] [1] Exemplary embodiment of prior art described in Patent Document 1 (WO 2020/262038)

Figure 1 shows an example of a hardware configuration of Performance Analyzing Device 10 according to an exemplary embodiment of the prior art described in Patent Document 1, i.e. WO 2020/262038. Performance Analyzing Device 10 is a device that executes various types of data processing to analyze a performance of a vehicle, machinery, etc. In this exemplary embodiment, Performance Analyzing Device 10 executes various types of data processing to analyze a performance of a ship. Performance Analyzing Device 10 is a computer comprising Processor 11, Memory 12, Storage 13, Communication Device 14, Data Input Device 15, Data Output Device 16 and Data Bus 17.

[0015] Processor 11 controls the entire computer by executing data

processing in accordance with programs such as an operating system. Processor 11 comprises a central processing unit (CPU) equipped with an interface for peripheral devices, a control unit, an arithmetic unit, a register, etc. Processor 11 reads out programs from at least one of Storage 13 and Communication Device 14 to Memory 12, and executes various types of data processing in accordance with the programs.

[0016] Memory 12 is a computer-readable recording medium, such as a read only memory (ROM) and a random access memory (RAM). Storage 13 is a computer-readable storage medium such as a hard disk drive. Communication Device 14 is a device that performs data communications with other computers via at least one of wired and wireless networks.

[0017] Data Input Device 15 is a device that receives data input from external devices such as a switch, a button, a sensor, etc. The external devices include an imaging device such as a digital camera that takes pictures. Data Output Device 16 is a device that outputs data to external devices such as a display, a speaker, an LED lamp, etc.

[0018] Data Input Device 15 and Data Output Device 16 may be configured as a single device such as a touch screen equipped with a stacked display and touch panel. Each device, such as Processor 11 and Memory 12, is connected with Data Bus 17 and communicates with other units of Performance Analyzing Device 10 via Data Bus 17. Data Bus 17 may be constituted of a single bus or of plural busses with one for each unit.

[0019] Functions of Performance Analyzing Device 10 are realized when Processor 11 executes data processing according to a program read from Storage 13 to Memory 12, and controls operations of other units of Performance Analyzing Device 10 such as data communications of Communication Device 14.

[0020] Figure 2 illustrates a functional configuration of Performance

Analyzing Device 10. Performance Analyzing Device 10 comprises, as functional components, Measured Value Storing Unit 101, Item Selection Receiving Unit 102, Dataset Extracting Unit 103, Condition Sufficiency Judging Unit 104, Relational Expression Calculating Unit 105 and Processing End Judging Unit 106. Measured Value Storing Unit 101 stores measured values of plural items related to an object of performance analysis. In this exemplary embodiment, Measured Value Storing Unit 101 stores values measured during one voyage of the ship, a performance of which is analyzed. In this specification, one voyage means a voyage performed by the ship from departure from a port to arrival at a next port.

[0021] In detail, Measured Value Storing Unit 101 stores measured values of plural items including, for example, a sailing speed (log speed and/or over-the-grounds speed), fuel consumption, engine speed (RPM), ship displacement, trim, change in sailing speed, wind speed, wind direction, wave height, current speed, current direction, pitch, roll, rudder angle, etc. Values are measured by a sensor on the ship, by a crew of the ship or by an external system such as a weather forecasting system, and the measured values are inputted to Performance Analyzing Device 10. Input of the measured values to Performance Analyzing Device 10 may be performed automatically via Communication Device 14 or manually via Data Input Device 15.

[0022] Measured Value Storing Unit 101 stores measured values of each item for plural measurement periods. Figure 3 illustrates examples of measured values stored in Measured Value Storing Unit 101. As shown in Figure 3, Measured Value Storing Unit 101 stores, for example, measured values of each item for measurement periods t1, t2, t3, t4, t5, etc., in correspondence with a name of ship where the values were measured. In the examples of measured values, the sign at the end of

each measured value such as "-t1" indicates a measurement period in which the value was measured.

[0023] A time interval of measured values stored in Measured Value Storing Unit 101 is, for example, a shortest time interval (hereinafter, "minimum measurement interval") among measurement time intervals for each item. For example, if the minimum measurement interval is once every minute, Measured Value Storing Unit 101 stores measured values of each item every minute. Measured Value Storing Unit 101 stores values complemented by interpolation or by extrapolation from available measured values before and after a measurement period for which no measured value is available.

[0024] Measured Value Storing Unit 101 may, for example, interpolate an average of available measured values before and after measurement periods as a value for the measurement period for which no measured value is available. Measured Value Storing Unit 101 may also calculate an approximate formula indicating a time-series change of measured values, and interpolate or extrapolate a value that the approximate formula indicates for each measurement period as a value for the measurement period for which no measured value is available. With regard to some items, such as ship displacement, a value is measured only once a day because a value of the item rarely fluctuates during navigation. For those items, Measured Value Storing Unit 101 may store the value measured on the day for all of measurement periods in the day.

[0025] For example, if a minimum measurement interval is every 10 minutes, and measurement intervals of two items are the same every 10 minutes but measurement timings of the items are not the same, Measured Value Storing Unit 101 stores the measured values of the items measured at different timings as values measured in the same

measurement period. Namely, values of different items of the same measurement period may be measured at different timings in the measurement period. Alternatively, Measured Value Storing Unit 101 may calculate interpolated values or extrapolated values of different items for the same timings in each measurement period, and store the interpolated values or the extrapolated values of different items for the measurement period.

[0026] Item Selection Receiving Unit 102 receives selections of items for analysis and selections of items for filtering from among plural items of measured values related to performance analysis. Items for analysis are two or more items relations of which are determined and used for the performance analysis. Items for filtering are one or more items that are used for filtering datasets containing measured values of items for analysis that adequately represent the aforementioned relations.

[0027] Item Selection Receiving Unit 102 displays an "item selection screen" that receives operations made by a user for selecting items for analysis and items for filtering. Figure 4 shows an example of item selection screen. Item selection screen A1 shown in Figure 4 contains Input Field C1 for receiving a name of ship to be analyzed, Input Field C2 for receiving one or more names of items to be included in the items for analysis, Input Field C3 for receiving one or more names of items to be included in the items for filtering, Input Field C4 for receiving a period in which values to be used for performance analysis are measured, and Button B1 for receiving an instruction to start the performance analysis.

[0028] Input of data to each input field in the item selection screen is made by a user of Performance Analyzing Device 10, such as a user who is in charge of performance analysis of ships. The input of data to each input field may be made by typing text, checking selected items from a

pull-down list showing candidate items, dragging selected items from a list of candidate items, etc. Item Selection Receiving Unit 102 receives data in each input field in the item selection screen as data indicating a ship selected by the user, items selected by the user as items for analysis, items selected by the user as items for filtering, or a period determined by the user as a period for selecting datasets to be used for the performance analysis, when Button B1 is pressed by the user to start performance analysis.

[0029] When Item Selection Receiving Unit 102 receives the data input by the user to the item selection screen, it sends the received data to Dataset Extracting Unit 103. Dataset Extracting Unit 103 extracts, from among pre-extracted datasets such as datasets stored in Measured Value Storing Unit 101, post-extracted datasets containing values of the items for filtering within a specified range for each item. Each dataset stored in Measured Value Storing Unit 101 contains values of plural items measured in the same measurement period.

[0030] Dataset Extracting Unit 103 reads from Measured Value Storing Unit 101 datasets that are associated with the name of ship indicated by the data sent from Item Selection Receiving Unit 102, contain values of the items for analysis and the items for filtering indicated by the data sent from Item Selection Receiving Unit 102, and contain values measured in a measurement period within the period indicated by the data sent from Item Selection Receiving Unit 102. Dataset Extracting Unit 103 causes Memory 12 to temporarily store the datasets read from Measured Value Storing Unit 101 as pre-extracted datasets.

[0031] Figure 5 shows examples of pre-extracted datasets stored in Memory 12. Figure 5 shows, for example, values of the items for analysis, i.e. log speed and fuel consumption, and values of the items for filtering, i.e. wind speed and wind direction, measured in measurement period t1,

i.e. LOG-t1, FOC-t1, Wind-t1 and WD-t1, are stored as a dataset named DS-t1 in Memory 12.

[0032] Similarly, datasets named DS-t2, DS-t3, DS-t4, DS-t5, etc., corresponding to measurement periods t2, t3, t4, t5, etc., are stored in Memory 12. The area in Memory 12 where pre-extracted datasets as shown in Figure 5 are temporality stored is hereafter referred to as a "pre-extracted dataset-storing area." The pre-extracted datasets that are read out from Measured Value Storing Unit 101 by Dataset Extracting Unit 103 and temporarily stored in the pre-extracted dataset-storing area are hereafter referred to as "initial pre-extracted datasets."

[0033] Dataset Extracting Unit 103 extracts post-extracted datasets containing values within the specified range for each of the items for filtering from the initial pre-extracted datasets stored in the pre-extracted dataset-storing area, and causes Memory 12 to store the post-extracted datasets in an area of Memory 12 other than the pre-extracted dataset-storing area named as a "post-extracted dataset-storing area." In this embodiment, Dataset Extracting Unit 103 executes the extraction of datasets by adjusting the specified ranges for each of the items for filtering such that a number of datasets excluded in the extraction is approximately a predetermined amount.

[0034] In the following description, the predetermined amount of datasets excluded in the extraction is defined as 90% of the number of pre-extracted datasets, and wind speed is used as one of the items for filtering. Dataset Extracting Unit 103 determines a range less than a predetermined percentage, for example 90% of the maximum wind speed of the pre-extracted datasets as a provisional range of wind speed for filtering datasets. When a range that is less than 90% of the maximum wind speed of the pre-extracted datasets is determined as the provisional range for filtering, Dataset Extracting Unit 103 excludes

datasets that contain measured values of wind speeds greater than 90% of the maximum wind speed of the pre-extracted datasets, and tentatively extracts datasets that contain measured values of wind speeds less than 90% of the maximum wind speed.

[0035] When a number of extracted datasets does not reach a predetermined amount, i.e. 10% of the number of pre-extracted datasets, Dataset Extracting Unit 103 tentatively extracts datasets that contain measured values of wind speeds that are less than 80% of the maximum wind speed of the pre-extracted datasets. Dataset Extracting Unit 103 repeatedly limits the range of wind speed for extracting datasets and tentatively extracts datasets containing values of wind speeds within the limited range until a number of datasets excluded in the tentative extraction reaches the predetermined amount, i.e. 10% of the number of pre-extracted datasets. When a number of datasets excluded in the tentative extraction reaches the predetermined amount, Dataset Extracting Unit 103 determines the datasets extracted in the final tentative extraction as post-extracted datasets for use in the subsequent processing, and stops the repeated tentative extractions.

[0036] A method of limiting a range of a value of an item for filtering is not limited to the above-mentioned method. For example, Dataset Extracting Unit 103 may repeatedly limit the range of wind speed by any percentage of increase, such as 1%, rather than 10%. Further, Dataset Extracting Unit 103 may re-extract datasets using a range (for example, a range less than 85% of the maximum value) that is greater than a range used in a previous extraction (for example, a range less than 80% of the maximum value) when a number of datasets excluded in the previous extraction is much larger than the predetermined amount.

[0037] In connection with the predetermined amount of excluded datasets in

the extraction, Dataset Extracting Unit 103 may use a fixed number instead of a fixed percentage of a number of pre-extracted datasets. After Dataset Extracting Unit 103 completes the extraction of datasets for wind speed as described above, it performs extractions of datasets for each of items for filtering other than wind speed in the same way as for wind speed. Dataset Extracting Unit 103 causes Memory 12 to temporarily store each group of post-extracted datasets extracted for each of the items for filtering in different post-extracted dataset-storing areas prepared for each of the items for filtering.

[0038] Condition Sufficiency Judging Unit 104 judges whether a number of post-extracted datasets and a statistical variation in measured values of the items for analysis contained in the post-extracted datasets satisfy a predetermined extraction condition. The extraction condition is a condition used for filtering datasets. A statistical variation is an indicator of variation in measured values in the present embodiment. The smaller the statistical variation in the measured values is, the clearer the trend shown by the measured values is. Accordingly, the smaller the statistical variation in the measured values is, the better the measured values are for performance analysis. Details of statistical variation in measured values will be explained later.

[0039] In this embodiment, a reduction ratio is defined as a ratio of a reduction in statistical variations in measured values to a reduction in numbers of datasets in the extraction. Condition Sufficiency Judging Unit 104 judges that the extraction condition is satisfied when a reduction ratio is larger than a predetermined threshold. Condition Sufficiency Judging Unit 104 calculates a reduction in numbers of datasets by subtracting the number of post-extracted datasets stored in the post-extracted dataset-storing area from the number of pre-extracted datasets stored in the pre-extracted dataset-storing area, for each of the

items for filtering.

[0040] Condition Sufficiency Judging Unit 104 instructs Relational Expression Calculating Unit 105 to calculate a relational expression among the items for analysis based on measured values of the pre-extracted datasets or the post-extracted datasets. The relational expression is used for calculating a reduction in statistical variations in values of each of the items for analysis. When Relational Expression Calculating Unit 105 receives the instruction, it calculates a relational expression among the items for analysis by use of the pre-extracted datasets stored in the pre-extracted datasets. Relational Expression Calculating Unit 105 also calculates a relational expression among the items for analysis by use of the post-extracted datasets stored in the post-extracted dataset-storing area for each of the items for filtering.

[0041] Relational Expression Calculating Unit 105 calculates the relational expression according to, for example, a curve fitting method. The curve fitting method is a method for determining a curve expressing a relation among parameters using a regression analysis such as a nonlinear least squares method, a nonlinear least absolute value method, etc. The type of regression analysis may be any one of a normal regression analysis, a principal component regression analysis, a geometric mean regression analysis, etc. The type of curve expressing a relation among the items for analysis may be any one of a curve following the square-cubic law expressed as $y=ax^3$, a more generic curve expressed as $y=ax^b$, etc.

[0042] The curve calculated by Relational Expression Calculating Unit 105 according to the curve fitting method may be a curve expressed by a polynomial, a curve with an intercept, etc. A type of curve calculated by Relational Expression Calculating Unit 105 may be selected by the user from among plural types of curve. In this case, the item selection screen shown in Figure 4 may contain an input field for receiving a type of

curve selected by the user. Relational Expression Calculating Unit 105 may calculate a relational expression that represents a relationship between, for example, values of log speed and values of fuel consumption using the datasets illustrated in Figure 5.

[0043] Figure 6 shows an example of a curve representing a relationship between values of log speed and values of fuel consumption. In Figure 6, Coordinate Space G1 is a coordinate space with a horizontal axis indicating log speed and a vertical axis indicating fuel consumption. Each of the plots in Coordinate Space G1 indicates any of the combinations of values of log speed and fuel consumption contained in the pre-extracted datasets shown in Figure 5. Curve D1 in Coordinate Space G1 indicates a relational expression calculated by Relational Expression Calculating Unit 105 based on the values contained in the pre-extracted datasets. Relational Expression Calculating Unit 105 calculates a relational expression for the pre-extracted datasets. Relational Expression Calculating Unit 105 also calculates a relational expression for the post-extracted datasets for each of the items for filtering. Relational Expression Calculating Unit 105 sends the calculated relational expressions to Condition Sufficiency Judging Unit 104.

[0044] Condition Sufficiency Judging Unit 104 calculates a reduction in statistical variations in values of the items for analysis in the extraction using the notified relational expressions for each of the items for filtering. More specifically, Condition Sufficiency Judging Unit 104 calculates, for each of the relational expressions handed over from Relational Expression Calculating Unit 105, a statistical variation, such as a mean, a standard deviation, etc., of differences in the direction of a Y-axis (or differences in the direction of an X-axis, shortest distances, etc.) between the values of items for analysis and the curve indicating

the relationship among the items for analysis. Namely, Condition Sufficiency Judging Unit 104 calculates a statistical variation in values of the items for analysis contained in the pre-extracted datasets, which is referred to hereafter as a pre-extraction statistical variation. Condition Sufficiency Judging Unit 104 also calculates, for each of the items for filtering, a statistical variation in values of the items for analysis contained in the post-extracted datasets that are extracted based on values of the item for filtering, which is referred to hereafter as a post-extraction statistical variation. Then, Condition Sufficiency Judging Unit 104 calculates, for each of the items for filtering, a reduction in statistical variations by subtracting the post-extraction statistical variation from the pre-extraction statistical variation.

[0045] Then, Condition Sufficiency Judging Unit 104 calculates, for each of the items for filtering, a reduction ratio by dividing the reduction in statistical variations by the reduction in numbers of datasets. Condition Sufficiency Judging Unit 104 judges, for each of the items for filtering, whether the extraction condition is satisfied based on the reduction ratio.

[0046] In this embodiment, Condition Sufficiency Judging Unit 104 selects the item for filtering whose reduction ratio is the largest among the reduction ratios of the items for filtering, and judges that only the extraction condition for the selected item for filtering is satisfied.

[0047] As described above, in this embodiment, Condition Sufficiency Judging Unit 104 judges whether the extraction condition is satisfied based on the statistical variation represented by differences between the values of the items for analysis and the relational expression calculated by Relational Expression Calculating Unit 105. Condition Sufficiency Judging Unit 104 sends the post-extracted datasets corresponding to the selected item for filtering to Dataset Extracting Unit 103.

- [0048] Dataset Extracting Unit 103 repeats the extraction process described above using the post-extracted datasets sent from Condition Sufficiency Judging Unit 104 as new pre-extracted datasets.
- [0049] For example, Dataset Extracting Unit 103 instructs Memory 12 to delete the initial pre-extracted datasets from the pre-extracted dataset-storing area and store the new pre-extracted datasets in the pre-extracted dataset-storing area. Dataset Extracting Unit 103 also instructs Memory 12 to delete the post-extracted datasets from each of the post-extracted dataset-storing areas corresponding to any of the items for filtering. Alternatively, Dataset Extracting Unit 103 may instruct Memory 12 to prepare a new pre-extracted dataset-storing area and new post-extracted dataset-storing areas, and store the new pre-extracted datasets in the new pre-extracted dataset-storing area. When Dataset Extracting Unit 103 completes the extraction process using the new pre-extracted datasets, Condition Sufficiency Judging Unit 104 judges which extraction condition is satisfied based on the reduction ratio calculated for new post-extracted datasets extracted from the new pre-extracted datasets by use of the extraction condition for each of the items for filtering.
- [0050] As the extraction process and the judgment process are repeated as described above, a number of the last selected post-extracted datasets, i.e. a number of the latest pre-extracted datasets, is gradually reduced. Each time a set of the extraction process and the judgment process is executed, Condition Sufficiency Judging Unit 104 notifies a number of the last selected post-extracted datasets to Processing End Judging Unit 106.
- [0051] Processing End Judging Unit 106 judges whether the extraction process and the judgment process, which are executed repeatedly, should be terminated. In this embodiment, Processing End Judging Unit

106 judges that a predetermined end condition is satisfied when the number of the last selected post-extracted datasets notified from Condition Sufficiency Judging Unit 104 is less than a predetermined threshold. When Processing End Judging Unit 106 judges that the end condition is satisfied, it instructs Condition Sufficiency Judging Unit 104 to terminate the judgment process. Following the instruction from Processing End Judging Unit 106, Condition Sufficiency Judging Unit 104 does not send the last selected post-extracted datasets to Dataset Extracting Unit 103. As described above, Dataset Extracting Unit 103 and Condition Sufficiency Judging Unit 104 repeatedly execute the extraction process and the judgment processes until the end condition is satisfied.

[0052] As described above, Performance Analyzing Device 10 executes the extraction process to extract post-extracted datasets from pre-extracted datasets. Figure 7 illustrates an example of an operation flow in the extraction process executed by Performance Analyzing Device 10. The extraction process is initiated when a user in charge of performance analysis instructs Performance Analyzing Device 10 to display the item selection screen shown in Figure 4. First, at Step S11, Performance Analyzing Device 10 (Item Selection Receiving Unit 102) receives data input to the input fields in the item selection screen, such as items for analysis input to Input Field C2 and items for filtering input to Input Field C3.

[0053] Then, at Step S12, Performance Analyzing Device 10 (Dataset Extracting Unit 103) selects one of the items for filtering as an item of focus for filtering, and extracts post-extracted datasets from the initial pre-extracted datasets by use of the extraction condition corresponding to the item of focus for filtering. Then, at Step S13, Performance Analyzing Device 10 judges whether there are one or more items for

filtering that have not yet been selected in already executed Step S12.

[0054] When Performance Analyzing Device 10 judges that there are one or more items for filtering that have not yet been selected in already executed Step S12, it repeats the processes at Step S12 and subsequent steps. When Performance Analyzing Device 10 judges that there are no items for filtering that have not yet been selected in already executed Step S12, the process at Step S14 is executed. At Step S14, Performance Analyzing Device 10 (Condition Sufficiency Judging Unit 104) judges which extraction condition is satisfied among the extraction conditions each for any of the items for filtering, based on a number of post-extracted datasets and a statistical variation of values of the items for analysis in the post-extracted datasets.

[0055] Then, at Step S15, The Performance Analyzing Device 10 (Processing End Judging Unit 106) judges whether the end condition is satisfied. When Performance Analyzing Device 10 judges that the end condition is not satisfied, it repeats the processes at Step S12 and subsequent steps. When Performance Analyzing Device 10 judges that the end condition is satisfied, it terminates the series of processes following the flow shown in Figure 7. The series of processes executed in the flow shown in Figure 7 is referred to hereafter as a set of repeated extractions.

[0056] Figure 8 is an example of a graph showing a transition of a statistical variation and a number of datasets while the set of repeated extractions is executed. The coordinate space shown in Figure 8 has a vertical axis indicating a statistical variation and a horizontal axis indicating a number of datasets. Black Circle E0 in the graph indicates a combination of the statistical variation of values of the items for analysis and the number of datasets in connection with the initial pre-extracted datasets. White Inverted Triangle E1 in the graph indicates

a combination of the statistical variation and the number of datasets in connection with the first post-extracted datasets of the first extraction process in the set of repeated extractions corresponding to the first item for filtering. Similarly, White Square E2, White Triangle E3 and Black Diamond E4 respectively indicate combinations of the statistical variation and the number of datasets regarding the first post-extracted datasets corresponding to the second item for filtering, the third item for filtering and the fourth item for filtering.

[0057] Hereinafter, the first item for filtering shall be trim, the second item for filtering shall be wave height, the third item for filtering shall be current speed, and the fourth item for filtering shall be wind speed. In this embodiment, the extraction condition is determined such that a reduction in numbers of datasets in each extraction process is approximately a predetermined number N. The graph shows that a size relation of reductions in statistical variations corresponding to the items of filtering in the first extraction process is trim, wave height, current speed and wind speed in ascending order. Namely, when the extraction condition regarding wind speed is used, the maximum reduction in statistical variations is achieved. Therefore, the first post-extracted datasets extracted by use of the extraction condition regarding wind speed is selected as a new pre-extracted datasets, i.e. the second pre-extracted datasets used for the second extraction process.

[0058] As shown in the graph, a size relation of reductions in statistical variations corresponding to the items of filtering in the second extraction process is wave height, trim, wind speed and current speed in ascending order. Therefore, the second post-extracted datasets extracted by use of the extraction condition regarding current speed is selected as the third pre-extracted datasets used for the third extraction process. The graph also shows that a size relation of reductions in statistical

variations corresponding to the items of filtering in the third extraction process is current speed, wind speed, trim and wave height in ascending order. Therefore, the third post-extracted datasets extracted by use of the extraction condition regarding wave height is selected.

[0059] Threshold Th1 shown on the horizontal axis in Figure 8 indicates the threshold used in the end condition. In the example shown in the graph of Figure 8, after the third extraction process, a number of each of the third post-extracted datasets corresponding to the items for filtering is less than Threshold Th1. Therefore, the set of repeated extractions is terminated. Dataset Extracting Unit 103 outputs to a predetermined destination the last selected datasets, i.e. the third post-extracted datasets extracted by use of the extraction condition regarding wave height.

[0060] The predetermined destination may be, for example, Storage 13, an external device connected to Data Output Device 16 such as a display, an external storage device, etc. In this embodiment, the datasets output to the predetermined destination are used for performance analysis of the ship. Figures 9 and 10 are each graphs showing examples of datasets used in the performance analysis of the ship.

[0061] The graphs of Figures 9 and 10 show log speeds of the ship measured each month. A faster log speed indicates a higher performance in terms of fuel efficiency.

[0062] Datasets used to draw Graph F1 of Figure 9 are the initial pre-extracted datasets fuel consumptions of which are within a predetermined range. Datasets used to draw Graph F2 of Figure 10 are the selected last post-extracted datasets in the set of repeated extractions.

[0063] In this example, maintenance was performed on the ship in the fourth month, and the log speed of the ship increased after the maintenance.

This indicates that maintenance reduced propulsion resistance during sailing of the ship, and fuel efficiency improved. It is not readily apparent Graph F1 shown in Figure 9 that fuel consumption had decreased at the timing when the maintenance was performed because the statistical variation in log speeds is large.

[0064] On the other hand, it is readily apparent from Graph F2 in Figure 10 that fuel consumption had decreased at the timing when the maintenance was performed because statistical variation in log speeds is small. Thus, as the statistical variation of measured values decreases, trends in measured values become more apparent. According to this embodiment, datasets are filtered based on measured values of the items for filtering contained in the datasets such that statistical variation of measured values of the items for analysis contained in the filtered datasets decrease. As a result, datasets are obtained that show a trend in values of the items for analysis more clearly than the original datasets.

[0065] In general, when extraction of datasets is properly performed a statistical variation in measured values decreases as a number of values is reduced. However, a problem arises when a number of datasets after extraction becomes too small. Graph F4 shown in Figure 11 comparative to Graph F2 shown in Figure 10 is illustrative of the problem. Datasets used to draw Graph F4 of Figure 11 are datasets extracted from the initial pre-extracted datasets used to draw Graph F1 using an extraction condition that filters out a large number of datasets in one extraction process. The number of plots in Graph F4 is smaller than the number of plots in Graph F2.

[0066] Graph F4 shows an unreliable trend such that the longer the elapsed time after the maintenances is, the higher the fuel efficiency is. This problem may be caused by the existence of abnormal values in the initial pre-extraction data, and by the existence of correlations between

the item for filtering used in the extraction, and the log speed.

[0067] When a number of datasets after the extraction process is excessively small, reliability of values contained in remaining datasets cannot be ensured, since the values may be subject to the influence of abnormal values contained in the initial pre-extracted datasets. According to Performance Analyzing Device 10, datasets are extracted in such a way that a number of datasets filtered out in the set of repeated extractions is minimized to reduce a statistical variation in values of the items for analysis, by selecting the item for filtering used for each of the extraction processes in the set of repeated extractions from among the plural items for filtering based on reduction ratios corresponding to the items for filtering. As a result, a sufficient number of datasets remain after the set of repeated extractions performed to obtain the target statistical variation.

[0068] For example, a relationship between log speed and fuel consumption should be maintained constant under constant conditions. However, values of log speed and fuel consumption that are measured during sailing of the ship include values that do not accurately reflect a relationship between log speed and fuel consumption due to various factors each of which change in a variety of ways during sailing. Datasets for performance analysis of a ship should include as few datasets as possible that show a relationship among values of items for analysis under circumstances that rarely occur, and as many datasets as possible that show a relationship among values of items for analysis under circumstances that occur frequently. Circumstances that occur frequently are referred to hereinafter as "usual circumstances," and the circumstances that rarely occur are referred to hereinafter as "unusual circumstances."

[0069] A relationship between values of log speed and values of fuel

consumption measured under usual circumstances is more reproducible than other relationships. When datasets showing a relationship between values of log speed and values of fuel consumption measured under unusual circumstances are excluded, a reduction in statistical variations is larger than when datasets showing a relationship between values of log speed and values of fuel consumption measured under usual circumstances are excluded, and a statistical variation in values of log speed and values of fuel consumption is reduced.

[0070] Performance Analyzing Device 10 uses a reduction ratio to specify the extraction condition that achieves a most efficient reduction of datasets, and uses the specified extraction condition for the extraction of datasets. As a result, datasets containing values measured under unusual circumstances are preferentially excluded, compared to a case where the extraction condition used for the extraction is not determined based on the reduction ratio. Thus, datasets extracted by Performance Analyzing Device 10 contain values of the items for analysis that show a relationship that is more reproducible than that for other relationships.

[0071] Items for filtering, which are used for filtering out datasets containing values measured under unusual circumstances, are not limited to items that have an obvious influence on a relationship between log speed and fuel consumption, such as wind speed, current speed, etc. Even if changes in circumstances affecting the relationship between log speed and fuel consumption and changes in measured values are coincidentally linked to each other, if the reduction ratio satisfies the extraction conditions, datasets extracted by Performance Analyzing Device 10 include measured values showing a reproducible relationship between log speed and fuel consumption.

[0072] In this embodiment, a number of datasets is not reduced by one reduction process, but rather a number of datasets is gradually reduced

by plural reduction processes using different extraction conditions as shown in the flow in Figure 8. By reducing a small number of datasets in each extraction and repeating reductions using different extraction conditions, datasets containing values measured under different unusual circumstances can be removed with high probability.

[0073] In this embodiment, Performance Analyzing Device 10 judges whether an extraction condition is satisfied and whether an end condition is satisfied. Therefore, reliable datasets can be obtained consistently without being affected by an ability or experience of a user in charge of performance analysis of the ship. In this embodiment, a number of datasets to be reduced in a single extraction is limited, as a result of which a number of datasets is prevented from becoming too small.

[0074] In this embodiment, datasets are extracted such that statistical variation, which is represented by differences between a relational expression among the items for analysis and measured values of the items for analysis, is reduced. Namely, the relational expression among the items for analysis is used as the reference for calculating the statistical variation. As a result, values of the items for analysis contained in the datasets after extraction more accurately show a relationship among the items for analysis.

[0075] [2] Modifications of prior art described in Patent Document 1 (WO 2020/262038)

The above-described embodiment is an exemplary embodiment of the prior art, and may be modified in various ways. Following are examples of modifications of the above-described embodiment. Two or more of the above-described embodiment and the following modifications may be combined.

[0076] [2-1] End condition

In the above-described embodiment, Processing End Judging Unit 106 judges that the set of repeated extractions should be terminated when a number of datasets is less than a predetermined threshold. Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a different condition is satisfied. For example, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a statistical variation in measured values of the items for analysis contained in the last post-extracted datasets satisfies a predetermined condition. For example, when a statistical variation in measured values of the items for analysis contained in the last post-extracted datasets is less than a predetermined threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated.

[0077] Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a reduction in statistical variations in values of the items for analysis in the last extraction satisfies a predetermined condition. For example, when a reduction in statistical variations in values of the items for analysis in the last extraction is more than a predetermined threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated. Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a reduction in numbers of datasets in the last extraction satisfies a predetermined condition. For example, when a reduction in numbers of datasets in the last extraction is more than a predetermined threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated.

[0078] When a number of datasets or a reduction in numbers of datasets is used in the end condition, the set of repeated extractions can be

terminated before a number of remaining datasets becomes too small. In such cases, termination of the set of repeated extractions may occur when a statistical variation is not small enough. On the other hand, when a statistical variation or a reduction in statistical variations is used in the end condition, a statistical variation can be terminated after a statistical variation becomes sufficiently small.

[0079] Processing End Judging Unit 106 may judge whether the set of repeated extractions should be terminated based on two or more of a statistical variation in values of the items for analysis contained in the post-extracted datasets extracted in the last extraction, a number of the post-extracted datasets extracted in the last extraction, a reduction in statistical variations in values of the items for analysis in the last extraction, and a reduction in numbers of datasets in the last extraction. For example, when the statistical variation is less than a threshold and the number of datasets is less than another threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated.

[0080] Processing End Judging Unit 106 may calculate an indicator using two or more of the above four parameters, and judge whether the set of repeated extractions should be terminated based on the indicator. For example, when an indicator obtained by adding a reduction in statistical variations multiplied by a predetermined factor and a number of datasets multiplied by another predetermined factor is more than a predetermined threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated.

[0081] Depending on how the thresholds are set, it is possible, for example, that a number of datasets may be reduced to an extent that extraction of datasets may result in a situation where the reduction ratio or statistical variation increases. To avoid such a situation, Processing End Judging

Unit 106 may judge that the set of repeated extractions should be terminated when a reduction ratio or a statistical variation increases in the last extraction process for more than a predetermined number of items for filtering, in addition to the end condition being satisfied. In this case, Dataset Extracting Unit 103 may select the post-extracted datasets of the last extraction process with the highest reduction ratio as the final datasets to be output, or Dataset Extracting Unit 103 may select the pre-extracted datasets of the last extraction process as the final datasets to be output.

[0082] According to this modification, reliable datasets can be stably extracted without dependence on knowledge or experience of a user in charge of performance analysis. Further, according to this modification, it is possible to extract datasets with a good balance between prevention of an excessive reduction of numbers of datasets and improvement of reduction in statistical variations, by combining two or more parameters.

[0083] [2-2] Reduction of numbers of datasets in each extraction process

In the above-described embodiment, extraction conditions are adjusted such that a reduction in numbers of datasets in each extraction process matches or approximates a predetermined number. Alternatively, extraction conditions may be adjusted such that a reduction in statistical variations in each extraction process matches or approximates a predetermined value. In this modification, Dataset Extracting Unit 103 adjusts a range of values of the item for filtering in the extraction condition such that a reduction in statistical variations in each extraction process matches or approximates a predetermined value.

[0084] In this modification, Dataset Extracting Unit 103 tentatively extracts, for each of the items for filtering, post-extracted datasets from pre-extracted datasets. Relational Expression Calculating Unit 105

calculates, for each of the group of post-extracted datasets, a relational expression among the items for analysis based on values of the items for analysis contained in the post-extracted datasets. Dataset Extracting Unit 103 calculates, for each of the group of post-extracted datasets, a reduction in statistical variations in the tentative reduction process by use of the relational expression in the same way as in the above-described embodiment. When the reduction in statistical variations in the tentative reduction process does not approximate the predetermined value, Dataset Extracting Unit 103 changes, for each of the items for filtering, the range of values of the item for filtering in the extraction condition such that a reduction in statistical variations approaches the predetermined value. More specifically, when the reduction in statistical variations in the tentative reduction process is substantially smaller than the predetermined value, Dataset Extracting Unit 103 widens the range in the extraction condition, and when the reduction in statistical variations in the tentative reduction process is substantially larger than the predetermined value, Dataset Extracting Unit 103 narrows down the range in the extraction condition. Dataset Extracting Unit 103 repeats the tentative extraction and the adjustment of extraction condition until the reduction in statistical variations substantially matches the predetermined value for all of the items for filtering.

[0085] Figure 12 is an example of a graph showing a transition of a statistical variation and a number of datasets while the set of repeated extractions is executed according to this modification. In Figure 12, as in the example in Figure 8, Black Circle E0 indicating a combination of the statistical variation and the number of the initial pre-extracted datasets, White Inverted Triangle E1 indicating a combination of the statistical variation and the number of the post-extracted datasets of

each extraction process using the extraction condition in connection with trim, White Square E2 indicating a combination of the statistical variation and the number of the post-extracted datasets of each extraction process using the extraction condition in connection with wave height, White Triangle E3 indicating a combination of the statistical variation and the number of the post-extracted datasets of each extraction process using the extraction condition in connection with current speed, and Black Diamond E4 indicating a combination of the statistical variation and the number of the post-extracted datasets of each extraction process using the extraction condition in connection with wind speed, are shown. In this modification, the extraction condition is determined such that a reduction in statistical variations in values of the items for analysis contained in datasets in each extraction process becomes approximately a predetermined value σ .

[0086] The graph shows that a size relation of reductions in numbers of datasets corresponding to the items of filtering in the first extraction process is trim, wave height, current speed and wind speed in ascending order. In this case, a reduction ratio of datasets extracted using the extraction condition in connection with wind speed becomes the largest. Therefore, the post-extracted datasets of the first extraction process extracted using the extraction condition in connection with wind speed is selected as the second pre-extracted datasets used in the second extraction process. For the second and third extraction processes, similarly, datasets with the smallest reduction in numbers of datasets is selected.

[0087] In this modification, Processing End Judging Unit 106 judges that the set of repeated extractions should be terminated when a statistical variation is less than a predetermined threshold. In the graph of Figure 12, Threshold Th2 used in the end condition is shown on the vertical

axis, and a statistical variation becomes less than Threshold Th2 after the third extraction process. Therefore, Performance Analyzing Device 10 terminates the set of repeated extractions after the third extraction process, and the last post-extracted datasets a number of which is the largest is selected as the final datasets to be output.

[0088] In this modification, when the threshold used in the end condition is fixed, a number of extraction processes executed in the set of repeated extractions is also fixed. In the above-described embodiment, when a fixed threshold such as Threshold Th1 shown in Figure 8 is used in the end condition, a number of extraction processes executed in the set of repeated extractions is also fixed.

[0089] In this modification, a threshold of a number of datasets such as Threshold Th1 shown in Figure 8 may be used in the end condition instead of Threshold Th2. Similarly, in the above-described embodiment, a threshold of statistical variation such as Threshold Th2 shown in Figure 12 may be used in the end condition instead of Threshold Th1. In either case, a number of datasets, a statistical variation in values of the items for analysis, or their reduction can be reduced to the threshold used in the end condition.

[0090] [2-3] Change in reduction in each extraction process

In the above-described embodiment, a reduction in numbers of datasets in each extraction process is substantially fixed. The reduction in numbers of datasets may be changed in each extraction process. In the above modification, a reduction in statistical variations in each extraction process is substantially fixed. The reduction in statistical variations may be changed in each extraction process. For example, Dataset Extracting Unit 103 may change the extraction condition based on a number of extraction processes that have already been executed in the set of repeated extractions.

[0091] Figure 13 is an example of a graph showing a transition of a statistical variation and a number of datasets while the set of repeated extractions is executed in this modification. The information represented by Black Circle E0, White Inverted Triangle E1, White Square E2, White Triangle E3 and Black Diamond E4 in the graph in Figure 13 is the same as those in Figures 8 and 12. In the example of Figure 13, when at least one of a statistical variation for any of the items for filtering reaches Threshold Th3 the set of repeated extractions is terminated.

[0092] In this modification, Dataset Extracting Unit 103 adjusts the extraction condition used for the first extraction process such that a reduction in numbers of datasets becomes approximately N1, adjusts the extraction condition used for the second extraction process such that a reduction in numbers of datasets becomes approximately N2, which is smaller than N1, and adjusts the extraction condition used for the third extraction process such that a reduction in numbers of datasets becomes approximately N3, which is smaller than N2. Namely, until the statistical variation is reduced to a certain degree, a large number of datasets are excluded in one extraction process, and when the statistical variation approaches the threshold, a small number of datasets are excluded in one extraction process.

[0093] Similarly, Dataset Extracting Unit 103 may adjust the extraction condition such that a reduction in statistical variations reaches a threshold that becomes smaller with the number of extraction processes already executed. In this case, the set of repeated extractions may be terminated when a number of the last post-extracted datasets reaches a threshold. According to this modification, without increasing a number of extraction processes in a set of repeated extractions, datasets showing a reliable relationship among the items for analysis are obtained

[0094] [2-4] Change of items for filtering

In the above-described embodiment, the items for filtering are not changed when the set of repeated extractions is executed. However, the items for filtering may be changed when the set of repeated extractions is executed. In this modification, Dataset Extracting Unit 103 may change the items for filtering for each of extraction processes executed in the set of repeated extractions. For example, Dataset Extracting Unit 103 may remove from the current items for filtering an item corresponding to the post-extracted datasets selected as the pre-extracted datasets for a next extraction process, and add a new item to the items for filtering, each time an extraction process is completed.

[0095] For example, in the example in Figure 8, Dataset Extracting Unit 103 removes wind speed from the items for filtering after the first extraction process is completed, since the post-extracted datasets extracted using the extraction condition in connection with wind speed are selected as the second pre-extracted datasets. Then, Dataset Extracting Unit 103 adds a new item, such as engine revolution speed, to the items for filtering for the second extraction process. It may occur that, once a range of values of an item has been limited, further limiting the range of values of the same item may result in a significant reduction in numbers of datasets.

[0096] According to this modification, an item already used for filtering out unreliable datasets is not repeatedly used in a set of repeated extractions and thus the above problem is avoided. Instead of replacing an item for filtering already used for the extraction process with another item, the item may simply be removed from the items for filtering, or one or more new items may be added to the items for filtering without removing the item already used for the extraction process. In either case, as long as the items for filtering are appropriately changed,

reliable datasets can be efficiently extracted.

[0097] [2-5] How to limit measured values of the items for filtering

A method for limiting measured values of the items for filtering is not limited to that described in the above-described embodiment. For example, Dataset Extracting Unit 103 may extract datasets in which measured values of two or more items for filtering satisfy a predetermined condition.

[0098] For example, Dataset Extracting Unit 103 may extract, from the pre-extracted datasets, datasets containing measured values of wind direction indicating a head wind and measured values of wind speed indicating values that are less than the first threshold, and containing measured values of wind direction indicating a tail wind and measured values of a wind speed indicating values that are less than the second threshold, or containing measured values of a wind direction indicating a side wind and measured values of a wind speed indicating values that are less than the third threshold. For example, Dataset Extracting Unit 103 may calculate, for each of the pre-extracted datasets, an indicator by adding a measured value of pitch and a measured value of roll, and extract, from the pre-extracted datasets, datasets in which indicators are less than a threshold.

[0099] A relational expression among the items for filtering may be used for extracting datasets. For example, Relational Expression Calculating Unit 105 calculates a relational expression among two or more items for filtering based on measured values of the items for filtering contained in the pre-extracted datasets. Then, Dataset Extracting Unit 103 extracts datasets by excluding, from the pre-extracted datasets, datasets containing values of the items for filtering with a large difference from the relational expression such that a statistical variation in differences between values of the items for filtering and the relational expression

becomes less than a threshold.

[0100] For example, Dataset Extracting Unit 103 calculates a relational expression expressing a relationship between engine revolution speed and change in sailing speed based on values of these items for filtering contained in the pre-extracted datasets. Then, Dataset Extracting Unit 103 calculates a statistical variation, such as a mean, a standard deviation, etc., of differences between values of the items for filtering and a curve indicating the relational expression (differences in the direction of a Y-axis, differences in the direction of an X-axis, or shortest distances, etc.). When the statistical variation is more than a threshold, Dataset Extracting Unit 103 removes datasets from the pre-extracted datasets in the order of their difference from the curve indicating the relational expression until a statistical variation in values of the items for filtering contained in the remaining datasets reaches the threshold. The datasets that remain when the statistical variation reaches the threshold are used as post-extracted datasets.

[0101] When there is a correlation among values of two or more items for filtering, the larger the statistical variation in differences between measured values and the relational expression, the more likely it is that a large number of abnormal values due to failure of a measuring device, misjudgment of a measurer, etc. are contained in the pre-extracted datasets. In many cases, the abnormal values are not detected by monitoring measured values of only one of the items for filtering. According to this modification, such abnormal values can be eliminated by using the relational expression as a reference.

[0102] [2-6] Extraction condition

Extraction conditions different from those used in the above-described embodiment may be used. In the above-described embodiment, Condition Sufficiency Judging Unit 104 selects

post-extracted datasets with the largest reduction ratio as the datasets satisfying the extraction condition. Alternatively, for example, Condition Sufficiency Judging Unit 104 may select a predetermined number of groups of post-extracted datasets in decreasing order of reduction ratios.

[0103] In such a case, Dataset Extracting Unit 103 executes the next extraction process for each of the selected groups of post-extracted datasets. For example, if two groups of post-extracted datasets are selected by Condition Sufficiency Judging Unit 104 in each extraction process and extraction processes are executed three times for the set of repeated extractions, eight groups of post-extracted datasets are obtained as a result of the set of repeated extractions. In such a case, all of the post-extracted datasets may be used for performance analysis, or one only may be selected for use for performance analysis.

[0104] For example, Condition Sufficiency Judging Unit 104 may select post-extracted datasets with a smallest statistical variation or a largest number of datasets as the final datasets to be output. In the above-described embodiment, Condition Sufficiency Judging Unit 104 compares reduction ratios calculated for each of the items for filtering to judge whether the extraction condition is satisfied. Instead of using a relative condition, an absolute condition may be used as the extraction condition.

[0105] For example, Condition Sufficiency Judging Unit 104 judges that the extraction condition is satisfied when the reduction ratio is larger than a predetermined threshold. In such a case, when plural groups of datasets are extracted, Condition Sufficiency Judging Unit 104 may select one of them as the final datasets to be output. Condition Sufficiency Judging Unit 104 may use an extraction condition including a combination of an absolute condition and a relative condition.

[0106] In such a case, when there are plural groups of datasets that satisfy the absolute condition, Condition Sufficiency Judging Unit 104 may select, for example, a group of datasets that satisfies the relative condition as the datasets to be used in the next extraction process. Whichever extraction condition is used, reliable datasets showing a correct relationship among the items for analysis are obtained without need to depend on a knowledge and experience of a user.

[0107] [2-7] Number of extraction processes

In the above-described embodiment, the extraction process is executed more than once for the set of repeated extractions. The extraction process may be executed only once for the set of repeated extractions. In such a case, when one group of post-extracted datasets of the first execution process is selected to be output, the set of repeated extractions is terminated. Even in such a case, since one group of datasets is selected from among plural groups of datasets based on efficiency of extraction expressed by a reduction ratio, a large number of datasets with only a small statistical variation can be output. Thus, according to this modification, datasets that show a reliable relationship among the items for analysis are obtained.

[0108] [2-8] Datasets to be output

In the above-described embodiment, Dataset Extracting Unit 103 outputs datasets that are judged to satisfy the extraction condition. However, Dataset Extracting Unit 103 may output datasets that are judged not to satisfy the extraction condition, in addition to datasets that are judged to satisfy the extraction condition. In such a case, the user may select one group of datasets from among the plural groups of datasets output from Performance Analyzing Device 10 for performance analysis. In such a case, as long as the user appropriately selects a group of datasets, the user can obtain datasets containing measured values that

show a reproducible relationship among the items for analysis. Namely, according to this modification, the user can readily obtain datasets that are useful for performance analysis with the assistance of Performance Analyzing Device 10.

[0109] When Condition Sufficiency Judging Unit 104 determines datasets to be output, datasets that are useful for performance analysis are obtained regardless of a knowledge and experience of the user. However, a skilled user who has extensive knowledge and experience may select a better group of datasets for performance analysis than a group of datasets automatically selected by Performance Analyzing Device 10. According to this modification, the knowledge and experience of a skilled user can be utilized to select datasets.

[0110] [2-9] Datasets containing values measured during several different voyages

In the above-described embodiment, datasets are extracted from ones that contain values measured during the same voyage of a ship whose performance is analyzed. Datasets may be extracted from ones that contain values measured during any one of plural different voyages of the ship.

[0111] In this modification, Measured Value Storing Unit 101 stores datasets classified into plural groups corresponding to different voyages of the same ship whose performance is analyzed. These groups are referred to hereinafter as voyage groups. Each of the voyage groups contains plural datasets that contain values measured during a voyage corresponding to the group. Dataset Extracting Unit 103 executes the first extraction process for each of the voyage group. As a result, plural groups of post-extracted datasets each corresponding to any of the items for filtering are extracted for each of plural voyage groups. Condition Sufficiency Judging Unit 104 calculates, for each of the groups of

post-extracted datasets, a statistical variation in values of the items for analysis contained in the datasets. Then, Condition Sufficiency Judging Unit 104 calculates an indicator for each of the items for filtering using the statistical variations calculated for the item for filtering.

[0112] For example, Condition Sufficiency Judging Unit 104 calculates the reduction ratio, i.e. a ratio of reduction in statistical variations in values of the items for analysis contained in datasets to reduce a number of datasets in the extraction process for each of the groups of post-extracted datasets. Then, Condition Sufficiency Judging Unit 104 calculates an average of the reduction ratios for each of the items for filtering as an indicator of reduction efficiency.

[0113] Condition Sufficiency Judging Unit 104 selects, for each of the voyage group, the post-extracted datasets corresponding to the item for filtering an indicator of which is the largest as the pre-extracted datasets for the second extraction process.

[0114] Similar to the above-described embodiment, in this modification, Dataset Extracting Unit 103 and Condition Sufficiency Judging Unit 104 repeatedly execute the extraction process and the selection process until the end condition is satisfied. A method for calculating the indicator of reduction efficiency is not limited to the above method. For example, Condition Sufficiency Judging Unit 104 may calculate a sum of the reduction ratios, a value obtained by multiplying the reduction ratios, etc., as the indicator of reduction efficiency.

[0115] The voyage groups may correspond to voyages of different ships, and each voyage group may contain datasets containing values measured during a voyage of a corresponding ship. In such a case, datasets containing values measured under unusual circumstances are excluded, and datasets containing values showing a reproducible relationship among the items for analysis are obtained.

[0116] The voyage groups may correspond to different time periods, and each voyage group may contain datasets containing values measured during the voyage of the ship in the corresponding time period. In this modification, Condition Sufficiency Judging Unit 104 calculates the indicator of reduction efficiency for each of the different time periods. The indicators may be used to analyze changes over time in parameters that affect a performance of the ship.

[0117] [2-10] Type of statistical variation

In the above-described embodiment, a statistical variation in differences between values of the items for analysis and a relational expression showing a relationship among the items for analysis is used for executing the set of repeated extractions. Types of statistical variation used for executing the set of repeated extractions are not limited to that in the above-described embodiment. For example, a statistical variation in values of one item for analysis may be used for executing the set of repeated extractions.

[0118] [2-11] Number of items for analysis

In the above-described embodiment, log speed and fuel consumption are used as examples of the items for analysis. A number of items for analysis is not limited to two, and one or more than two items for items for analysis may be used.

[0119] [2-12] How to determine and how to obtain items for analysis and items for filtering

In the above-described embodiment, the items for analysis and the items for filtering are selected by the user from candidate items measured values of which are stored in Measured Value Storing Unit 101. A method to determine the items for analysis and the items for filtering is not limited to that in the above-described embodiment. The items for analysis and the items for filtering may be predetermined.

[0120] In the above-described embodiment, measured values of the items for analysis and measured values of the items for filtering are stored in Measured Value Storing Unit 101. A place where these values are stored is not limited to Measured Value Storing Unit 101. These values may be stored in an external system with which Performance Analyzing Device 10 can communicate, and Dataset Extracting Unit 103 may obtain the values from the external system.

[0121] [2-13] Functional configuration

A functional configuration of a device according to the prior art is not limited to that shown in Figure 2. For example, in the above-described embodiment, Dataset Extracting Unit 103 outputs the final datasets. A unit shown in Figure 2 other than Dataset Extracting Unit 103 or a unit that is not shown in Figure 2 may output the final datasets. For example, the extraction process executed by Dataset Extracting Unit 103 and the selection process executed by Condition Sufficiency Judging Unit 104 in the above-described embodiment may be executed by a single unit.

[0122] [2-14] Hardware configuration

A hardware configuration of a device according to the present invention is not limited to that shown in Figure 1. For example, Performance Analyzing Device 10 may consist of two or more separate devices. In such a case, the devices that constitute Performance Analyzing Device 10 may contain a resource comprising a cloud service.

[0123] [3] Exemplary embodiment of present invention

Abnormality Detecting Device 20 according to an embodiment of this invention is described below. Abnormality Detecting Device 20 is a device that detects persistent abnormalities occurring in measurements of values of an item related to a performance of a device to be analyzed or in measurements of values of an item affecting the performance of

the device to be analyzed, based on datasets each of which contains measured values of the items.

[0124] The device to be analyzed is a device that performs any kind of work or processing whose type is not limited. For example, the device to be analyzed may be a moving device such as a ship, a data processing device such as a computer, etc.

[0125] The performance of the device, i.e., the efficiency of the work, processing, etc. performed by the device, is affected by the environment in which the device operates, etc. For example, when the device to be analyzed is a ship, fuel consumption (ton/mile) is one of the items related to the performance of the device. Fuel consumption is affected by wind speed, wind direction, etc.

[0126] Therefore, when the performance of the ship is analyzed by use of values of fuel consumption of the ship that are continuously measured, i.e. repeatedly measured at different timings, values of fuel consumption that were measured at substantially same timings when values of wind speed, wind direction, etc., within normal ranges of each of the items were measured, are extracted from the obtained measured values of fuel consumption, and the extracted measured values of fuel consumption are used for the analysis of the performance of the ship.

[0127] A normal range in this application means a range that does not include an outlier, i.e., a value that is extremely large or extremely small in comparison with other values. For example, the normal range may be determined by any one of the following methods.

(1) A method in which the x-th largest value among the measured values is determined as the upper limit of the normal range of the measured values, and the x-th smallest value among the measured values is determined as the lower limit of the normal range of the measured values, when x is a predetermined number.

(2) A method in which a value obtained by adding a predetermined value to the average of the measured values is determined as the upper limit of the normal range of the measured values, and a value obtained by subtracting a predetermined value from the average of the measured values is determined as the lower limit of the normal range of the measured values.

(3) A method in which the upper and lower limits of the normal range of the measured values are calculated in accordance with a statistical theory such as test of rejection of Smirnoff-Grubbs, etc.

[0128] A substantially same timing in this application means a concept that is not limited to strictly the same timing, but includes different timings with sufficiently small differences, which can be considered the same timing. The differences among timings that can be considered as a substantially same timing may differ depending on a rate of change of measured value.

[0129] In this application, a measured value means a concept that is not limited to a value actually measured, but includes a value calculated by use of the actually measured value. For example, values obtained by interpolating actually measured values, a moving average of actually measured values, etc., are also treated as measured values in this application.

[0130] There may be plural items related to the performance of the device to be analyzed. For example, a performance of a ship to transport cargo is higher when it consumes less fuel if the time required for the transportation is the same, and it is higher when the time required for the transportation is shorter if the fuel consumption required for the transportation is the same. Therefore, when considering the performance of a ship, it is necessary to analyze two items related to the performance, i.e. fuel consumption and log speed (knot), at the same time.

Accordingly, a relational expression between fuel consumption and log speed of a ship is usually used as information showing the performance of the ship to transport cargo.

[0131] Fuel consumption and log speed of a ship are items related to the performance of the ship, in other words, items indicating attributes of the ship. On the other hand, wind speed, wind direction, etc., affect the performance of the ship, but they are not attributes of the ship. Accordingly, wind speed, wind direction, etc., are not items related to the performance of the ship, but items affecting the performance of the ship. There may be plural items affecting the performance of the device to be analyzed, like wind speed and wind direction for the ship.

[0132] Hereafter, in the same way as described in Patent Document 1, one or more items related to the performance of the device to be analyzed (an example of a first item or examples of plural sub items of the first item) are referred to as items for analysis, and one or more items that affect the performance of the device to be analyzed (examples of second items) are referred to as items for filtering.

[0133] The performance analyzing device described in Patent Document 1 extracts, from the obtained datasets, for each of plural items for filtering, datasets containing values of the item for filtering within a normal range. The performance analyzing device calculates, for each of the plural items for filtering, a filtering efficiency. A filtering efficiency is $\Delta\sigma/\Delta N$, when ΔN is a decrease in a number of datasets before and after the extraction, and $\Delta\sigma$ is a decrease in an indicator of statistical variation, such as variance, of values of the item for analysis contained in datasets before and after the extraction. The performance analyzing device selects a group of extracted datasets whose filtering efficiency is the maximum for analyzing the performance of the device, so that statistical variation of values of the item for analysis is efficiently reduced as

much as possible without reducing number of values of the item for analysis.

[0134] A type of the indicator of statistical variation is not limited to variance, but may be standard deviation, range size, interquartile range, mean difference, mean absolute deviation, etc.

[0135] Abnormality Detecting Device 20 of this embodiment is in common with the performance analyzing device of Patent Document 1 in that it uses filtering efficiencies. However, Abnormality Detecting Device 20 differs from the performance analyzing device of Patent Document 1 in that filtering efficiencies are used for detecting persistent abnormalities occurring in measurements of values of an item for analysis or an item for filtering, not for analyzing the performance of the device to be analyzed.

[0136] Persistent abnormalities in measurements of values mean a state in which abnormal values are included in the measured values of a measurement device at a higher frequency than normal as a result of a persistent malfunction of the measurement device or peripheral devices of the measurement device such as a communication cable. On the other hand, temporary abnormalities in measurements of values mean a state in which the measurement device and the peripheral devices of the measurement device are functioning normally, but abnormal values (noise) are accidentally output by the measurement device due to the measurement environment or other reasons.

[0137] A hardware configuration of Abnormality Detecting Device 20 is the same as the hardware configuration of Performance Analyzing Device 10 of Patent Document 1; namely, a computer shown in Figure 1. Therefore, a description of the hardware configuration of Abnormality Detecting Device 20 is not repeated here.

[0138] Figure 14 shows a functional configuration realized by Abnormality

Detecting Device 20. In other words, when a computer used as hardware of Abnormality Detecting Device 20 executes data processing in accordance with a program of this embodiment, Abnormality Detecting Device 20 comprising the functional components shown in Figure 14 is realized.

[0139] The functional configuration of Abnormality Detecting Device 20 is described below. In the following description, the device to be analyzed is assumed to be a ship.

[0140] Dataset Obtaining Unit 201 obtains datasets containing measured values. Dataset Obtaining Unit 201 receives datasets from external devices, and stores the datasets.

[0141] A composition of datasets obtained by Dataset Obtaining Unit 201 is the same as that of datasets stored by Measured Value Storing Unit 101 described in Patent Document 1, as shown in Figure 3. Namely, Dataset Obtaining Unit 201 obtains datasets in connection with each of different ships, such as Ship A, Ship B, etc. Dataset Obtaining Unit 201 obtains, for each of the ships, datasets each of which contains values of items such as log speed, O. G. speed (speed over ground), etc., measured at a substantially same timing. Different datasets contain values measured at substantially different timings from each other, such as measurement period t1, measurement period t2, etc. A group of data stored in each data row of the table shown in Figure 3 constitutes one dataset.

[0142] Items having values contained in each dataset are classified into either the items for analysis or the items for filtering. For example, log speed and fuel consumption indicate the performance of the ship, and are targets of analysis, and thus are classified into the items for analysis. On the other hand, wind speed and wind direction do not indicate the performance of the ship but affect the performance of the ship. Therefore, wind speed and wind direction are classified into the items

for filtering.

[0143] Dataset Obtaining Unit 201 reads from the stored datasets, for each of two or more different time periods, a group of datasets containing values measured in the time period, in accordance with conditions received by Condition Receiving Unit 202 from a user, and hands them over to Dataset Extracting Unit 203.

[0144] Condition Receiving Unit 202 receives from the user conditions for specifying datasets that are read by Dataset Obtaining Unit 201 and handed over to Dataset Extracting Unit 203.

[0145] Figure 15 shows an example of a screen (hereinafter referred to as "condition reception screen") that Condition Receiving Unit 202 displays for receiving conditions from the user. On the condition reception screen, the user enters data indicating a target ship, items for analysis, items for filtering, time period, and number of divisions. In the "target ship" field, the user enters an identification of the ship, such as a name of the ship, for which the user wants to detect persistent abnormalities in measurements. In the "items for analysis" field, the user enters one or more identifications of items of analysis, such as names of the items, for which the user wants to detect persistent abnormalities in measurements. In the "items for filtering" field, the user enters one or more identifications of items of filtering, such as names of the items, for which the user wants to detect persistent abnormalities in measurements. In the "time period" field, the user enters information on a time period, such as a starting date and an ending date of the time period, for which the user wants to detect persistent abnormalities in measurements. In the "number of divisions" field, the user enters a number of divisions for the time period entered in the "time period" field.

[0146] Abnormality Detecting Unit 206 detects persistent abnormalities in

measurements of values based on results of comparisons between two or more filtering efficiencies calculated by Filtering Efficiency Calculating Unit 205 for each of plural time periods. The number entered in the "number of divisions" field on the condition reception screen (hereafter referred to as "number of divisions") is used to specify each of the plural periods for which a filtering efficiency is calculated by Filtering Efficiency Calculating Unit 205. Namely, Filtering Efficiency Calculating Unit 205 calculates, for each of the divided time periods, a filtering efficiency using a group of datasets containing values measured in the time period.

[0147] When the user clicks the "start" button on the condition reception screen, Condition Receiving Unit 202 hands over the information entered by the user on the condition reception screen to Dataset Obtaining Unit 201.

[0148] Dataset Obtaining Unit 201 reads a group of datasets to be handed over to Dataset Extracting Unit 203 from the stored datasets according to the information handed over from Condition Receiving Unit 202. For example, when Dataset Obtaining Unit 201 receives the information illustrated in Figure 15 from Condition Receiving Unit 202, Dataset Obtaining Unit 201 reads from the stored datasets (see Figure 3) for Ship A, which contains values measured during each of the following three periods.

[0149] (First period) from January 1, 2020 to April 30, 2020

(Second period) from May 1, 2020 to August 1, 2020

(Third period) from September 1, 2020 to December 31, 2020

[0150] Dataset Obtaining Unit 201 generates, for each of the above three time periods, a group of datasets containing values of the items for filtering specified by the user, such as log speed, fuel consumption, etc., and values of the items for filtering specified by the user, such as trim,

wave height, current speed, wind speed, etc., and hands the generated groups of datasets over to Dataset Extracting Unit 203.

[0151] Dataset Extracting Unit 203 extracts, for each of the time periods and for each of the items for filtering, from the group of datasets provided by Dataset Obtaining Unit 201, datasets each of which contains a value of the filtering item within a normal range for the filtering item determined according to a predetermined rule.

[0152] Various "predetermined rules" can be employed by Dataset Extracting Unit 203 to extract datasets. Some examples of the predetermined rule are shown below.

[0153] (1) Identify the normal range for which outliers are no longer produced by tests following statistical theories such as Smirnov-Grubbs test.

(2) When a standard deviation of the values is σ , an arbitrarily determined positive number is α , and a mean of the values is μ , $(\mu - \alpha \times \sigma)$ is the lower limit of the normal range, and $(\mu + \alpha \times \sigma)$ is the upper limit of the normal range.

(3) When a number of the values is N , and an arbitrarily determined positive number less than 1 is β , $(\beta \times N + 1)$ th value from the lower end of the values is the lower limit of the normal range, and $(\beta \times N + 1)$ th value from the higher end of the values is the upper limit of the normal range

(4) Using the method described in (2) above, while gradually changing α , extraction of datasets is repeatedly performed until an amount of decrease in the indicator of statistical variation of values of the item for analysis before and after the extraction reaches a predetermined threshold value. Then, the normal range when the amount of decrease in the indicator reaches the threshold value is adopted.

(5) Using the method described in (2) above, while gradually changing α , extraction of datasets is repeatedly performed until a decrease ratio of the indicator of statistical variation of values of the item for analysis

before and after the extraction reaches a predetermined threshold value. Then, the normal range when the decrease ratio of the indicator reaches the threshold value is adopted.

(6) Using the method described in (3) above, while gradually changing β , extraction of datasets is repeatedly performed until an amount of decrease in the indicator of statistical variation of values of the item for analysis before and after the extraction reaches a predetermined threshold value. Then, the normal range when the amount of decrease in the indicator reaches the threshold value is adopted.

(7) Using the method described in (3) above, while gradually changing β , extraction of datasets is repeatedly performed until a decrease ratio of the indicator of statistical variation of values of the item for analysis before and after the extraction reaches a predetermined threshold value. Then, the normal range when the decrease ratio of the indicator reaches the threshold value is adopted.

[0154] When the indicator of statistical variation of values of the item for analysis is used for extracting datasets as in (4) to (7) above, Dataset Extracting Unit 203 calculates the indicator.

[0155] When the number of items for analysis specified by the user is one, Dataset Extracting Unit 203 calculates the indicator of statistical variation of values of the item for filtering. As previously mentioned, any type of indicator of statistical variation, such as variance, standard deviation, etc., may be employed.

[0156] When the number of items for analysis specified by the user is two or more, Dataset Extracting Unit 203 hands over the datasets to Relation Specifying Unit 204. Relation Specifying Unit 204 specifies a relational expression between the items for analysis based on the values of the items contained in the datasets handed over from Dataset Extracting Unit 203, and hands over the specified relational expression to Dataset

Extracting Unit 203. Dataset Extracting Unit 203 calculates an indicator of statistical variation of difference between a value calculated by use of the relational expression specified by Relation Specifying Unit 204 and a measured value contained in the dataset, as the indicator of statistical variation of values of the items for filtering.

[0157] The difference between a value calculated by use of the relational expression and a measured value contained in a dataset may be, for example, calculated by one of the following methods.

[0158] (1) Select one item from the items for filtering, and calculate a value of the selected item by substituting measured values of unselected items contained in a dataset into the relational expression. Then, a difference between the calculated value of the selected item and a measured value of the selected item contained in the dataset is calculated.

[0159] For example, when log speed (hereinafter referred to as "S") and fuel consumption (hereinafter referred to as "F") are specified as the items for filtering by the user, the relational expression specified by Relation Specifying Unit 204 is $F=f(S)$, i.e. fuel consumption F is a function of log speed S, and measured values of the items for filtering contained in the datasets are (S_1, F_1) , (S_2, F_2) , ..., (S_n, F_n) , where n is the number of the datasets, Dataset Extracting Unit 203 selects, for example, fuel consumption F, and calculates $\Delta F_1=f(S_1)-F_1$, $\Delta F_2=f(S_2)-F_2$, ..., $\Delta F_n=f(S_n)-F_n$. Then, Dataset Extracting Unit 203 calculates an indicator of statistical variation, such as variant, of $\Delta F_1, \Delta F_2, \dots, \Delta F_n$.

[0160] When the relational expression specified by Relation Specifying Unit 204 is $S=f(F)$, i.e. log speed S is a function of fuel consumption F, Dataset Extracting Unit 203 selects, for example, log speed S, and calculates $\Delta S_1=f(F_1)-S_1$, $\Delta S_2=f(F_2)-S_2$, ..., $\Delta S_n=f(F_n)-S_n$. Then, Dataset Extracting Unit 203 calculates an indicator of statistical variation, such as variant, of $\Delta S_1, \Delta S_2, \dots, \Delta S_n$.

[0161] (2) Calculate a shortest distance between a line in a coordinate space showing the relational expression specified by Relation Specifying Unit 204 and a coordinate point in the coordinate space showing measured values of the items for filtering contained in a dataset.

[0162] For example, when log speed (hereinafter referred to as "S") and fuel consumption (hereinafter referred to as "F") are specified as the items for filtering by the user, the relational expression specified by Relation Specifying Unit 204 is $F=f(S)$, i.e. fuel consumption F is a function of log speed S, and measured values of the items for filtering contained in the datasets are (S1, F1), (S2, F2), ..., (Sn, Fn), where n is the number of the datasets, Dataset Extracting Unit 203 calculates, for example, the following values.

$$\Delta S1F1 = \min \{[(S-S1)^2 + (F-F1)^2]^{(1/2)}\},$$

$$\Delta S2F2 = \min \{[(S-S2)^2 + (F-F2)^2]^{(1/2)}\},$$

...

$$\Delta S_n F_n = \min \{[(S-S_n)^2 + (F-F_n)^2]^{(1/2)}\},$$

where S and F are arbitrary values satisfying $F=f(S)$.

Then, Dataset Extracting Unit 203 calculates an indicator of statistical variation, such as variant, of $\Delta S1F1, \Delta S2F2, \dots, \Delta S_n F_n$.

[0163] As previously mentioned, Relation Specifying Unit 204 specifies a relational expression between the items for analysis based on the datasets handed over from Dataset Extracting Unit 203. Various methods may be employed by Relation Specifying Unit 204 to specify a relational expression, similar to Relational Expression Calculating Unit 105 of Patent Document 1. For example, Relation Specifying Unit 204 may use a regression analysis technique such as the nonlinear least squares method or the nonlinear least absolute value method to specifying the relational expression. A type of relational expression specified by Relation Specifying Unit 204 may be selected from among,

for example, exponential approximation equation, linear approximation equation, logarithmic approximation equation, polynomial approximation equation, power approximation equation, etc., depending on a combination of the items for analysis.

[0164] For example, when log speed and fuel consumption are the items for analysis, Relation Specifying Unit 204 specifies a relational expression shown by Graph D1 in Figure 6.

[0165] Relation Specifying Unit 204 hands the specified relational expression over to Dataset Extracting Unit 203.

[0166] Dataset Extracting Unit 203 performs the extraction process described above focusing on each of the items for filtering, i.e. trim, wave height, current speed, and wind speed, for each of the time periods, i.e. the first period, the second period and the third period, by use of the datasets handed over from Dataset Obtaining Unit 201.

[0167] When Dataset Extracting Unit 203 completes the extraction process, Dataset Extracting Unit 203 hands over the following information to Filtering Efficiency Calculating Unit 205.

[0168] (1-1) Datasets before and after the extraction process focusing on trim for the first period.

(1-2) Datasets before and after the extraction process focusing on trim for the second period.

(1-3) Datasets before and after the extraction process focusing on trim for the third period.

[0169] (2-1) Datasets before and after the extraction process focusing on wave height for the first period.

(2-2) Datasets before and after the extraction process focusing on wave height for the second period.

(2-3) Datasets before and after the extraction process focusing on wave height for the third period.

- [0170] (3-1) Datasets before and after the extraction process focusing on current speed for the first period.
- (3-2) Datasets before and after the extraction process focusing on current speed for the second period.
- (3-3) Datasets before and after the extraction process focusing on current speed for the third period.
- [0171] (4-1) Datasets before and after the extraction process focusing on wind speed for the first period.
- (4-2) Datasets before and after the extraction process focusing on wind speed for the second period.
- (4-3) Datasets before and after the extraction process focusing on wind speed for the third period.
- [0172] Filtering Efficiency Calculating Unit 205 receives from Dataset Extracting Unit 203 datasets before and after the extraction focusing on each of the items for filtering for each of the time periods, and calculates a filtering efficiency for each of the combinations of time period and item for filtering by use of the pairs of datasets before and after extraction. Then, Filtering Efficiency Calculating Unit 205 hands over the calculated filtering efficiencies to Abnormality Detecting Unit 206.
- [0173] Namely, Filtering Efficiency Calculating Unit 205 calculates a filtering efficiency for each of the above (1-1) to (4-3).
- [0174] When a decrease in the number of datasets before and after extraction is ΔN and a decrease in the indicator of statistical variation of measured values of the items for analysis before and after extraction is ΔV , a filtering efficiency P is expressed as $P=(\Delta V/\Delta N)$.
- [0175] Filtering Efficiency Calculating Unit 205 calculates an indicator of statistical variation of measured values of the items for analysis for each of the pairs of datasets received from Dataset Extracting Unit 203 in the

same manner as Dataset Extracting Unit 203.

- [0176] When Dataset Extracting Unit 203 has already calculated the indicators of statistical variation of the measured values of the items for analysis for the extraction of datasets, Filtering Efficiency Calculating Unit 205 may use the indicators to calculate filtering efficiencies.
- [0177] Filtering Efficiency Calculating Unit 205 hands over the filtering efficiencies calculated for each of the pairs of datasets of the above (1-1) to (4-3) (hereinafter referred to as filtering efficiency P1-1, filtering efficiency P1-2, ..., filtering efficiency P4-3) to Abnormality Detecting Unit 206.
- [0178] Abnormality Detecting Unit 206 detects persistent abnormalities in measurements of values of either an item for analysis or an item for filtering based on results of comparisons between the filtering efficiencies handed over from Filtering Efficiency Calculating Unit 205.
- [0179] Figure 16 is a graph showing changes in the number of datasets and changes in the indicator of statistical variation in the values of the items for analysis in each of the pairs of datasets before and after the extraction of datasets.
- [0180] Figure 16(a) is a graph relating to the first period. Point E_0 in Figure 16(a) shows the number of the datasets and the indicator of statistical variation of the measured values of the items for analysis contained in the datasets for the first period before the extraction is performed. Point E_1 in Figure 16(a) shows the number of the datasets and the indicator of statistical variation of the measured values of the items for analysis contained in the datasets for the first period after the extraction focusing on trim is performed. Similarly, each of points E_2 to E_4 in Figure 16(a) shows the number of datasets and the indicator of statistical variation of the measured values of the items for analysis contained in the datasets for the first period after the extraction focusing on wave height, current

speed, and wind speed respectively. Figure 16(b) is a graph relating to the second period, and Figure 16(c) is a graph relating to the third period.

[0181] Figure 17 shows examples of the filtering efficiency calculated by Filtering Efficiency Calculating Unit 205. Figure 18 is an example of graph showing changes in the filtering efficiency for each of the items for filtering.

[0182] Abnormality Detecting Unit 206 compares filtering efficiencies for two or more different periods with respect to each of the items for filtering shown in Figure 17, and detects persistent abnormalities in measurements of values of an item for filtering based on the results of the comparisons.

[0183] In the examples shown in Figures 16 to 18, the filtering efficiency with respect to wind speed in the third period (P_{4-3}) is significantly smaller than the filtering efficiencies with respect to wind speed in the first period (P_{4-1}) and the second period (P_{4-2}). In this case, Abnormality Detecting Unit 206 detects the occurrence of persistent abnormalities in measurements of values of wind speed during the third period.

[0184] For example, Abnormality Detecting Unit 206 may judge whether or not persistent abnormalities in measurements of values are occurring by any one of the following methods.

(1) Judge that persistent abnormalities in measurements of values are occurring when an amount of change in filtering efficiency, i.e. ($P_{4-2}-P_{4-1}$) or ($P_{4-3}-P_{4-2}$), exceeds a predetermined threshold value.

(2) Judge that persistent abnormalities in measurements of values are occurring when a rate of change in filtering efficiency, i.e. (P_{4-2}/P_{4-1}) or (P_{4-3}/P_{4-2}), exceeds a predetermined threshold value.

(3) Judge that persistent abnormalities in measurements of values are occurring when one or more statistical outliers are contained in the

filtering efficiencies.

[0185] Figure 19 is another graph showing an example of changes in the filtering efficiency for each of the items for filtering. The graph of Figure 19 shows that filtering efficiencies in the third period are significantly lower than the filtering efficiencies in the first and second periods for all of the items for filtering. In this case, Abnormality Detecting Unit 206 determines that persistent abnormalities in measurements of values of an item for analysis, i.e., log speed or fuel consumption, occur during the third period.

[0186] When Abnormality Detecting Unit 206 detects an occurrence of persistent abnormalities in measurements of values of any item, Abnormality Notifying Unit 207 notifies the user of the occurrence of persistent abnormalities. Abnormality Notifying Unit 207 may notify the user of the occurrence of persistent abnormalities by, for example, pronouncing an audio message, displaying a graphical message, sending a text message to a terminal device used by the user, etc.

[0187] According to Abnormality Detecting Device 20, when persistent abnormalities in measurements of values of an item related to or affecting the performance of the device to be analyzed occur, the user can know the occurrence of persistent abnormalities.

[0188] For example, if an anemometer that was working properly during the first and second periods fails during the third period and starts to output wrong values of wind speed, it is not easy for the user to find the failure of the anemometer based on the values output by the anemometer, since the values output by the anemometer are constantly changing. However, according to Abnormality Detecting Device 20, the user can easily know the failure of the anemometer.

[0189] [4] Modifications of the embodiment of present invention

The above-described embodiment is an exemplary embodiment of

the present invention, and may be modified in various ways. Following are examples of modifications of the above-described embodiment. Two or more of the above-described embodiment and the following modifications may be combined.

[0190] [4-1] Configuration of device

In the above-described embodiment, Abnormality Detecting Device 20 is realized by a computer executing data processing in accordance with a program. Alternatively, Abnormality Detecting Device 20 may be a dedicated device that comprises the components shown in Figure 14.

[0191] [4-2] System

In the above-described embodiment, this invention is realized by one device, i.e. Abnormality Detecting Device 20, but the invention may be realized by a system including plural devices that operate cooperatively with each other. In such a case, each of the components of Abnormality Detecting Device 20 (see Figure 14) may be located in any of the plural devices included in the system.

[0192] [4-3] Datasets before extraction

In the above-described embodiment, datasets before an extraction by Dataset Extracting Unit 203 are read from datasets stored by Dataset Obtaining Unit 201, but other datasets may be used as datasets before the extraction.

[0193] For example, among datasets read from the datasets stored by Dataset Obtaining Unit 201, only datasets containing normal values of the items for analysis may be used in the extraction process performed by Dataset Extracting Unit 203.

[0194] Datasets after an extraction process performed by Dataset Extracting Unit 203 may be used in a subsequent extraction process performed by Dataset Extracting Unit 203. In this case, Dataset Obtaining Unit 201 obtains datasets after an extraction process as datasets used in a

subsequent extraction process.

[4-4] Categories of Invention

This invention can be understood as a device exemplified by Abnormality Detecting Device 20 in the above-described embodiment, a system exemplified by the system of the above described modification [4-2], a program for causing a computer to execute the data processing performed by the device exemplified by Abnormality Detecting Device 20, and an abnormality detection method performed by the device exemplified by Abnormality Detecting Device 20. The program according to this invention may be provided in a form of a recording medium such as an optical disk on which the program is stored, or may be provided in a downloadable form to a computer via a network such as the Internet and installed in the computer.

DESCRIPTION OF REFERENCE NUMERALS

- [0195] 20: Abnormality Detecting Device
- 201: Dataset Obtaining Unit
- 202: Condition Receiving Unit
- 203: Dataset Extracting Unit
- 204: Relation Specifying Unit
- 205: Filtering Efficiency Calculating Unit
- 206: Abnormality Detecting Unit
- 207: Abnormality Notifying Unit

CLAIMS

- [Claim 1] A program causing a computer to execute:
- when a group of datasets each of which contains a measured value of a first item related to a performance of a device to be analyzed and a measured value of a second item affecting the performance of the device, the value of the first item and the value of the second item contained in each dataset being measured at a substantially same timing,
- a step for acquiring, for each of two or more different time periods, a group of datasets each of which contains values measured in the time period,
- a step for extracting, for each of the two or more different time periods, datasets that contain measured values of the second item within a range determined in accordance with a predetermined rule,
- a step for calculating, for each of the two or more different time periods, a ratio of a decrease in an indicator of statistical variation of measured values of the first item in datasets before and after the extraction to a decrease in a number of datasets before and after the extraction, and
- a step for detecting persistent abnormalities in measurements of values of the first item or the second item based on a comparison among the ratios each of which is calculated for one of the two or more different time periods.
- [Claim 2] A program according to Claim 1, when the first item includes a plurality of sub items, causing the computer to execute:
- a step for specifying, for each of the two or more different time periods, a relational expression among the sub items of the

first item based on the extracted datasets, and

a step for calculating the indicator, for each of the datasets before and after the extraction for each of the two or more different time periods, that represents a statistical variation in difference between a value of the first item according to the relational expression and a measured value of the first item contained in each dataset.

[Claim 3]

A method executed by a data processing device comprising:

when a group of datasets each of which contains a measured value of a first item related to a performance of a device to be analyzed and a measured value of a second item affecting the performance of the device, the value of the first item and the value of the second item contained in each dataset being measured at a substantially same timing,

a step for acquiring, for each of two or more different time periods, a group of datasets each of which contains values measured in the time period,

a step for extracting, for each of the two or more different time periods, datasets that contain measured values of the second item within a range determined in accordance with a predetermined rule,

a step for calculating, for each of the two or more different time periods, a ratio of a decrease in an indicator of statistical variation of measured values of the first item in datasets before and after the extraction to a decrease in a number of datasets before and after the extraction, and

a step for detecting persistent abnormalities in measurements of values of the first item or the second item based on a comparison among the ratios each of which is calculated for one of

the two or more different time periods.

[Claim 4]

A system comprising:

when a group of datasets each of which contains a measured value of a first item related to a performance of a device to be analyzed and a measured value of a second item affecting the performance of the device, the value of the first item and the value of the second item contained in each dataset being measured at a substantially same timing,

a means for acquiring, for each of two or more different time periods, a group of datasets each of which contains values measured in the time period,

a means for extracting, for each of the two or more different time periods, datasets that contain measured values of the second item within a range determined in accordance with a predetermined rule,

a means for calculating, for each of the two or more different time periods, a ratio of a decrease in an indicator of statistical variation of measured values of the first item in datasets before and after the extraction to a decrease in a number of datasets before and after the extraction, and

a means for detecting persistent abnormalities in measurements of values of the first item or the second item based on a comparison among the ratios each of which is calculated for one of the two or more different time periods.