

METHOD AND APPARATUS FOR PROCESSING DATA, DEVICE, AND STORAGE MEDIUM

TECHNICAL FIELD

[0001] The present disclosure relates to the field of computer technologies, and in particular, relates to a method and apparatus for processing data, a device, and a storage medium.

BACKGROUND

[0002] With the increasing development of computer technologies, there is an increasing demand for data storage.

[0003] In the related art, due to the limitation of regional conditions and safety factors, for example, the wind turbine generator set is deployed in a wild area, data transmission over networks is challenging. The operation files of the wind turbine generator set cannot be transmitted to the operation and maintenance personnel at a remote distance in real time, and instead, the operation and maintenance personnel periodically copy the operation files from a computer device storing operation files near the wind turbine generator set.

[0004] However, the data storage capacity of the computer device is limited, and the operation files cannot be fully stored. How to store operation files in the computer device is an urgent problem to be solved.

SUMMARY

[0005] The present disclosure provides a method and apparatus for processing data, a device, and a storage medium. The technical solutions are as follows.

[0006] According to one aspect of the present disclosure, a method for processing data is provided. The method is performed by a computer device. The method includes:

[0007] acquiring occupation information of storage space, the occupation information being configured to indicate occupation of the storage space by a plurality of operation files of a production device;

[0008] reserving and storing operation files in a first file subset and deleting operation files in a second file subset in the case that the occupation information exceeds a first storage threshold and

does not exceed a second storage threshold, wherein in a file set including the plurality of operation files, the first file subset and the second file subset are complementary to each other; and

[0009] reserving and storing operation files in a third file subset and deleting operation files in a fourth file subset in the case that the occupation information exceeds the second storage threshold, wherein in the file set, the third file subset and the fourth file subset are complementary to each other.

[0010] The first file subset includes a first number of operation files whose generation times belong to a first time period, and the third file subset includes a second number of operation files whose generation times belong to the first time period, the first number being greater than the second number.

[0011] In some embodiments of the present disclosure, each of the plurality of operation files features a corresponding time type, and generation times of at least two operation files of a same time type both belong to a same time period;

[0012] reserving and storing the operation files in the first file subset and deleting the operation files in the second file subset in the case that the occupation information exceeds the first storage threshold and does not exceed the second storage threshold includes:

[0013] in the case that the occupation information exceeds the first storage threshold and does not exceed the second storage threshold, determining a files by sampling from operation files of a same time type, and constructing the first file subset, a being a positive integer;

[0014] determining the second file subset from the plurality of operation files based on the first file subset; and

[0015] reserving and storing the operation files in the first file subset and deleting the operation files in the second file subset.

[0016] In some embodiments of the present disclosure, determining the a files by sampling from the operation files of the same time type further includes at least one of:

[0017] determining the a files from the operation files of the same time type in a random manner;

[0018] determining the a files from the operation files of the same time type in a manner that a generation time interval of adjacent files is a predefined time interval;

[0019] determining the a files from the operation files of the same time type under a constraint that file contents of the a files are different from each other; or

[0020] determining the a files from the operation files of the same time type under a constraint

that a file content difference between adjacent files exceeds a difference threshold.

[0021] In some embodiments of the present disclosure, each of the plurality of operation files features a corresponding time type, and generation times of at least two operation files of a same time type both belong to a same time period;

[0022] said reserving and storing the operation files in the third file subset and deleting the operation files in the fourth file subset in the case that the occupation information exceeds the second storage threshold includes:

[0023] in the case that the occupation information exceeds the second storage threshold, determining b files by sampling from the operation files of the same time type, and constructing the third file subset, b being a positive integer;

[0024] determining the fourth file subset from the plurality of operation files based on the third file subset; and

[0025] reserving and storing the operation files in the third file subset and deleting the operation files in the fourth file subset.

[0026] In some embodiments of the present disclosure, the method further includes:

[0027] acquiring generation times of the plurality of operation files; and

[0028] determining at least two time types for the plurality of operation files based on the generation times, durations of time periods corresponding to different time types being the same.

[0029] In some embodiments of the present disclosure, the method further includes:

[0030] acquiring a file cleanup directory, the file cleanup directory including storage paths of a part of the plurality of operation files, wherein a storage path of an operation file in the second file subset and/or the fourth file subset belongs to the file cleanup directory.

[0031] In some embodiments of the present disclosure, the production device includes a wind turbine generator set. The method further includes:

[0032] acquiring a data cleanup label, wherein a storage path of an operation file in the second file subset and/or the fourth file subset belongs to the file cleanup directory, and the operation file is provided with the data cleanup label.

[0033] The data cleanup label is configured to indicate operation data of the wind turbine generator set in the case that a wind turbine rotation speed of the wind turbine generator set is less than a rotation speed threshold or operation data of the wind turbine generator set in the case that a pitch angle of the wind turbine generator set belongs to an angle range.

[0034] In some embodiments of the present disclosure, the method further includes:

[0035] deleting a part of operation files at a head position in a file sequence in the case that the occupation information exceeds a third storage threshold, wherein the file sequence is obtained by arranging the plurality of operation files by the generation times from the earliest to the latest, wherein the occupation information does not exceed the third storage threshold after the part of operation files is deleted.

[0036] In some embodiments of the present disclosure, the method further includes:

[0037] acquiring a file cleanup type, the file cleanup type is a file data type of a part of the plurality of operation files, wherein a file data type of an operation file in the second file subset and/or the fourth file subset belongs to the file cleanup type.

[0038] In some embodiments of the present disclosure, the method further includes:

[0039] acquiring current time information and a historical reading time of the computer device, the historical reading time being a time at which the computer device transmits a historical operation file to a reading device; and

[0040] determining the first storage threshold and/or the second storage threshold based on a storage duration between the current time information and the historical reading time, wherein the storage duration is in a positive correlation with the first storage threshold, and/or the storage duration is in a positive correlation with the second storage threshold.

[0041] In some embodiments of the present disclosure, the method further includes:

[0042] acquiring a historical reading interval, the historical reading interval being configured to indicate an adjacent time interval in which the computer device transmits the historical operation file to the reading device;

[0043] determining the first storage threshold and/or the second storage threshold based on the storage duration between the current time information and the historical reading time includes:

[0044] determining the storage duration between the current time information and the historical reading time; and

[0045] determining the first storage threshold and/or the second storage threshold based on a duration ratio of the storage duration in the historical reading interval, wherein the duration ratio is in a positive correlation with the first storage threshold, and/or the duration ratio is in a positive correlation with the second storage threshold.

[0046] In some embodiments of the present disclosure, the production device includes a wind

turbine generator set;

[0047] acquiring the occupation information of the storage space includes:

[0048] acquiring wind turbine file occupation information of the storage space by a plurality of wind turbine operation files of the wind turbine generator set;

[0049] reserving and storing the operation files in the first file subset and deleting the operation files in the second file subset in the case that the occupation information exceeds the first storage threshold and does not exceed the second storage threshold includes:

[0050] reserving and storing wind turbine operation files in the first file subset and deleting wind turbine operation files in the second file subset in the case that the wind turbine file occupation information exceeds the first storage threshold and does not exceed the second storage threshold;

[0051] reserving and storing the operation files in the third file subset and deleting the operation files in the fourth file subset in the case that the occupation information exceeds the second storage threshold includes:

[0052] reserving and storing wind turbine operation files in the third file subset and deleting wind turbine operation files in the fourth file subset in the case that the wind turbine file occupation information exceeds the second storage threshold.

[0053] According to another aspect of the present disclosure, an apparatus for processing data is provided. The apparatus includes an acquiring module and a processing module.

[0054] The acquiring module is configured to acquire occupation information of storage space, the occupation information being configured to indicate occupation of the storage space by a plurality of operation files of a production device.

[0055] The processing module is configured to reserve and store operation files in a first file subset and delete operation files in a second file subset in the case that the occupation information exceeds a first storage threshold and does not exceed a second storage threshold, wherein in a file set including the plurality of operation files, the first file subset and the second file subset are complementary to each other.

[0056] The processing module is further configured to reserve and store operation files in a third file subset and delete operation files in a fourth file subset in the case that the occupation information exceeds the second storage threshold, wherein in the file set, the third file subset and the fourth file subset are complementary to each other.

[0057] The first file subset includes a first number of operation files whose generation times

belong to a first time period, and the third file subset includes a second number of operation files whose generation times belong to the first time period, the first number being greater than the second number.

[0058] According to another aspect of the present disclosure, a computer device is provided. The computer device includes a processor and a memory storing at least one instruction, at least one program, at least one code set, or at least one instruction set, wherein the at least one instruction, the at least one program, the at least one code set, or the at least one instruction set, when loaded and executed by the processor, cause the processor to perform the method for processing data according to the above aspects.

[0059] According to another aspect of the present disclosure, a computer-readable storage medium is provided. The computer-readable storage medium stores at least one instruction, at least one program, at least one code set, or at least one instruction set, wherein the at least one instruction, the at least one program, the at least one code set, or the at least one instruction set, when loaded and executed by a processor, cause the processor to perform the method for processing data according to the above aspects.

[0060] According to another aspect of the present disclosure, a computer program product is provided. The computer program product includes one or more computer instructions stored in a computer-readable storage medium, wherein the one or more computer instructions, when read from the computer-readable storage medium and executed by a processor, cause the processor to perform the method for processing data according to the above aspects.

[0061] The technical solutions according to the present disclosure achieve at least the following beneficial effects.

[0062] The file subset to be reserved and stored and the file subset to be deleted are determined based on the relationship between the occupation information and the first storage threshold and the relationship between the occupation information and the second storage threshold, so that the operation files are thinned in the case of insufficient storage space, thereby reducing the load on the storage space. Moreover, different numbers of files for different reservation and storage are determined based on different load pressures on the storage space, thereby dynamically reducing the load pressure on the storage space and providing a new storage method for operation files.

BRIEF DESCRIPTION OF DRAWINGS

[0063] To describe the technical solutions in embodiments of the present disclosure more clearly, the following briefly describes the accompanying drawings required for describing the embodiments. Apparently, the accompanying drawings in the following description show merely some embodiments of the present disclosure, and those of ordinary skill in the art may still derive other accompanying drawings from these accompanying drawings without creative efforts.

[0064] FIG. 1 is a schematic diagram of a computer system according to some exemplary embodiments of the present disclosure;

[0065] FIG. 2 is a schematic diagram of a method for processing data according to some exemplary embodiments of the present disclosure;

[0066] FIG. 3 is a first flowchart of a method for processing data according to some exemplary embodiments of the present disclosure;

[0067] FIG. 4 is a second flowchart of a method for processing data according to some exemplary embodiments of the present disclosure;

[0068] FIG. 5 is a third flowchart of a method for processing data according to some exemplary embodiments of the present disclosure;

[0069] FIG. 6 is a fourth flowchart of a method for processing data according to some exemplary embodiments of the present disclosure;

[0070] FIG. 7 is a fifth flowchart of a method for processing data according to some exemplary embodiments of the present disclosure;

[0071] FIG. 8 is another schematic diagram of a method for processing data according to some exemplary embodiments of the present disclosure;

[0072] FIG. 9 is a sixth flowchart of a method for processing data according to some exemplary embodiments of the present disclosure;

[0073] FIG. 10 is a structural block diagram of an apparatus for processing data according to some exemplary embodiments of the present disclosure; and

[0074] FIG. 11 is a structural block diagram of a server according to some exemplary embodiments of the present disclosure.

[0075] The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate embodiments consistent with the present disclosure and, together with the specification, serve to explain the principles of the present disclosure.

DETAILED DESCRIPTION

[0076] For clearer descriptions of the objectives, technical solutions, and advantages of the present disclosure, embodiments of the present disclosure are further described in detail hereinafter with reference to the accompanying drawings.

[0077] Reference is made in detail to the exemplary embodiments, examples of which are illustrated in the accompanying drawings. When the accompanying drawings are involved in the following description, the same numbers in different accompanying drawings represent the same or similar elements unless otherwise indicated. The embodiments described in the following exemplary embodiments do not represent all embodiments consistent with the present disclosure. Rather, they are merely examples of apparatus and methods consistent with some aspects of the present disclosure, as detailed in the appended claims.

[0078] The terms used in the present disclosure are for the purpose of describing particular embodiments only and are not intended to be limiting to the present disclosure. As used in the present disclosure and the appended claims, the singular forms "a", "an", and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It should also be understood that the term "and/or" as used herein refers to and encompasses any or all possible combinations of one or more associated listed items.

[0079] It should be noted that the user information (including but not limited to user device information, user personal information, and the like) and data (including but not limited to data for analysis, stored data, displayed data, and the like), which are referred to in the present disclosure, are information and data authorized by the user or fully authorized by various parties, and the collection, use, and processing of the relevant data are required to comply with relevant laws and regulations and standards in relevant countries and regions. For example, information such as the operation file involved in the present disclosure is acquired under sufficient authorization.

[0080] It should be understood that although the terms "first", "second", and the like may be used in the present disclosure to describe various information, such information should not be limited to these terms. These terms are only used to distinguish one type of information from another. For example, a first parameter may also be referred to as a second parameter, and similarly, a second parameter may also be referred to as a first parameter, without departing from the scope of the present disclosure. The word "if", as used herein, may be interpreted as "in the case that", "in the case of", or "in response to determining that", depending on the context.

[0081] FIG. 1 shows a schematic diagram of a computer system according to some embodiments of the present disclosure. The computer system may be implemented as a system architecture of a method for processing data. The computer system may include a terminal 100 and a server 200.

[0082] The terminal 100 may be an electronic device such as a mobile phone, a tablet computer, or a personal computer (PC). A client of a target application may be installed and run on the terminal 100. The target application may be a management application of a production device or the like, or may be another application that provides data processing, which is not limited in the present disclosure. In addition, the form of the target application is not limited in the present disclosure. The target application includes, but is not limited to, an application (App) installed on the terminal 100, a mini program, and the like, or may be in the form of a web page.

[0083] The server 200 may be an independent physical server, or may be a server cluster or a distributed system formed by a plurality of physical servers, or may be a cloud server that provides a cloud computing service. The server 200 may be a backend server of the target application, configured to provide a backend service for the client of the target application.

[0084] In the method for processing data according to the embodiments of the present disclosure, each step may be performed by a computer device. The computer device refers to an electronic device with data calculation, processing, and storage capacities. The solution implementation environment shown in FIG. 1 is used as an example. The terminal 100 may perform the method for processing data (for example, the client of the target application installed and running on the terminal 100 performs the method for processing data), the server 200 may perform the method for processing data, or the terminal 100 and the server 200 perform the method for processing data through interaction and cooperation, which is not limited in the present disclosure.

[0085] In addition, the technical solutions of the present disclosure may be combined with blockchain technology. For example, some data (such as an operation file) involved in the method for processing data disclosed in the present disclosure may be stored on a blockchain. The terminal 100 and the server 200 may communicate with each other over a network, such as a wired or wireless network.

[0086] FIG. 2 shows a schematic diagram of a method for processing data according to some exemplary embodiments of the present disclosure.

[0087] The occupation information 322 of storage space 310 is acquired. The storage space 310 stores the operation files of a production device, and the occupation information 322 is configured

to indicate the occupation of the storage space by a plurality of operation files of the production device. Exemplarily, the occupation information 322 is data in the form of a percentage, configured to indicate the occupation ratio of the storage space by the plurality of operation files.

[0088] With respect to the operation file, the production device being a wind turbine generator set is used as an example. The operation file records at least one of information such as the wind turbine rotation speed, generation power, or blade pitch angle of the wind turbine generator set in a process where the wind turbine generator set generates electricity by means of air blowing.

[0089] Exemplarily, n operation files in the storage space 310 are used as an example for description. The generation times of the n operation files are later than a moment T_0 and earlier than a moment T_1 .

[0090] In the case that the occupation information exceeds a first storage threshold 332 and does not exceed a second storage threshold 334, the occupation ratio of the storage space by the plurality of operation files exceeds 60% and does not exceed 80%. x operation files are reserved and stored, and $(n-x)$ operation files excluding the x operation files between the moment T_0 and the moment T_1 are deleted.

[0091] In the case that the occupation information exceeds the second storage threshold 334, the occupation ratio of the storage space by the plurality of operation files exceeds 80%. y operation files are reserved and stored, and $(n-y)$ operation files excluding the y operation files between the moment T_0 and the moment T_1 are deleted. As the load pressure on the storage space increases, a smaller number of operation files whose generation times are within a time period between the moment T_0 and the moment T_1 are reserved and stored; that is, y is less than x . Different numbers of files for different reservation and storage are determined based on different load pressures on the storage space to dynamically reduce the load pressure on the storage space.

[0092] Next, the method for processing data is described with reference to the following embodiments.

[0093] FIG. 3 shows a first flowchart of a method for processing data according to some exemplary embodiments of the present disclosure. The method may be performed by a computer device. The method includes:

[0094] In step 510, the occupation information of storage space is acquired.

[0095] Exemplarily, the storage space is configured to store the operation files of a production device, and the storage space is provided by the computer device or is storage space to which the

computer device is granted read permission. Exemplarily, the occupation information is configured to indicate the occupation of the storage space by a plurality of operation files of the production device. The occupation information may indicate the occupation ratio of the storage space by the plurality of operation files in the form of a percentage, or may indicate the occupation volume of the plurality of operation files and the total volume of the storage space numerically. The data format of the occupation information is not limited in the present disclosure.

[0096] Exemplarily, the production device is a device that alters at least one of the properties, performance, or form of raw materials input into the production device, or enhances appearance value in the social production process. Exemplarily, the production device may be a wind turbine generator set, a solar power generator set, a power transmission device, or the like used in the power production process; may be a mining device, a mineral transmission device, a mineral sorting device, or the like used in the mining production process; or may be a logistics sorting device, a logistics packaging device, or the like used in the logistics sorting production process. It may be understood that production devices in more production processes may be further included.

[0097] In step 520, in the case that the occupation information exceeds a first storage threshold and does not exceed a second storage threshold, operation files in a first file subset are reserved and stored, and operation files in a second file subset are deleted.

[0098] Exemplarily, in the case that the occupation information exceeds the first storage threshold and does not exceed the second storage threshold, the storage space reaches a first-level load pressure for storing operation files. The load in the storage space for storing operation files is reduced by deleting a part of operation files in the storage space.

[0099] Exemplarily, in a file set including a plurality of operation files, the first file subset and the second file subset are complementary to each other. The files, excluding the operation files in the second file subset, in the plurality of operation files are the operation files in the first file subset.

[0100] In step 530, in the case that the occupation information exceeds the second storage threshold, operation files in a third file subset are reserved and stored, and operation files in a fourth file subset are deleted.

[0101] In the case that the occupation information exceeds the second storage threshold, the storage space reaches a second-level load pressure for storing operation files. The load in the storage space for storing operation files is reduced by deleting a part of operation files in the storage space.

[0102] Exemplarily, in the file set, the third file subset and the fourth file subset are complementary to each other. The files, excluding the operation files in the fourth file subset, in the plurality of operation files are the operation files in the third file subset.

[0103] It should be noted that the first file subset includes a first number of operation files whose generation times belong to a first time period, and the third file subset includes a second number of operation files whose generation times belong to the first time period, the first number being greater than the second number.

[0104] It can be seen that in the case that the occupation information indicates that there are different load pressures, there are correspondingly different reservation and storage methods for operation files. Exemplarily, in the case that the storage space reaches the second-level load pressure for storing operation files, compared with the first-level load pressure, a smaller number of operation files whose generation times belong to the first time period are reserved and stored in the storage space (that is, the first number is greater than the second number), and the load on the storage space is reduced by reserving and storing the smaller number of operation files.

[0105] The first time period may be a predefined time period, for example, one day, one week, or 12 hours. In some embodiments, in the case that the occupation information exceeds the first storage threshold and does not exceed the second storage threshold, the operation files in the first file subset are reserved and stored, and the number of operation files whose generation times belong to the last day in the first file subset is 10; in the case that the occupation information exceeds the second storage threshold, the operation files in the third file subset are reserved and stored, and the number of operation files whose generation times belong to the last day in the third file subset is 5. As the load pressure on the storage space increases, the number of files reserved and stored over a time period decreases.

[0106] In summary, in the method according to the embodiments, the file subset to be reserved and stored and the file subset to be deleted are determined based on the relationship between the occupation information and the first storage threshold and the relationship between the occupation information and the second storage threshold, so that the operation files are thinned in the case of insufficient storage space, thereby reducing the load on the storage space. Moreover, different numbers of files for different reservation and storage are determined based on different load pressures on the storage space, thereby dynamically reducing the load pressure on the storage space and providing a new storage method for operation files.

[0107] FIG. 4 shows a second flowchart of a method for processing data according to some exemplary embodiments of the present disclosure. The method may be performed by a computer device. That is, in the embodiments shown in FIG. 3, step 520 is implemented as steps 522, 524, and 526.

[0108] In step 522, in the case that the occupation information exceeds the first storage threshold and does not exceed the second storage threshold, a files are determined by sampling from operation files of the same time type, and the first file subset is constructed.

[0109] Exemplarily, each of the plurality of operation files features a corresponding time type, and the generation times of at least two operation files of the same time type both belong to the same time period. The time period may be obtained through division based on a predefined time interval. For example, based on a time interval such as one day, one week, or 12 hours, a plurality of adjacent time periods are obtained through division by using the current moment as the end moment of a time period, or a plurality of adjacent time periods are obtained through division by using the earliest historical moment as the start moment of a time period. The earliest historical moment corresponds to the generation time of a first operation file in the storage space, and the first operation file is a file with the earliest generation time.

[0110] Exemplarily, a is a positive integer. Exemplarily, the value of a may be predefined, or may be determined based on the number of operation files of the same time type. The method for determining the value of a is not limited in the present disclosure. a files are determined by sampling from operation files of each same time type. The number of files in the first file subset is a product of a and the number of time types.

[0111] In some embodiments of the present disclosure, the method for determining the a files by sampling is described. The step is implemented as at least one of:

[0112] ·The a files are determined from the operation files of the same time type in a random manner.

[0113] Exemplarily, the a files reserved and stored in the first file subset are determined in a random manner, and a technical solution for determining the a files is provided.

[0114] ·The a files are determined from the operation files of the same time type in a manner that the generation time interval of adjacent files is a predefined time interval.

[0115] Exemplarily, the predefined time interval is used as the generation time interval between adjacent files, so that the a files are determined by uniform sampling based on the generation time,

thereby ensuring uniformity in generation time among the determined a files.

[0116] In operation files of one time type, files excluding the a files are files deleted from the second file subset. The operation files in the first file subset have uniformity in generation time, and operation files with close times can be used to replace files deleted from the second file subset.

[0117] Exemplarily, the value of a being 8 and the total number of the operation files of the same time type denoted as $N = 50$ is used as an example for description. Specifically:

[0118] $\text{step} = \text{float64}(N - 1) / \text{float64}(a - 1)$

[0119] $\text{index} = i * \text{step} + 1;$

[0120] $\text{lowIndex} = \text{math.Floor}(\text{index});$

[0121] $\text{roundIndex} = \text{math.Round}(\text{index});$

[0122] $\text{fraction} = \text{index} - \text{float64}(\text{lowIndex});$

[0123] $\text{indices}[i] = \{$

[0124] $\text{lowIndex}, \text{fraction equal to } 0$

[0125] $\text{roundIndex}, \text{fraction not equal to } 0\}$

[0126] step represents the step size of determining the a files; for example, $a = 8$ and $N = 50$, and the step size of determining 8 files from 50 files is 7. Index represents the reference index serial number of the a files among a total of N operation files of the same time type, and the value of i ranges from 0 to $a-1$, corresponding to the reference index serial numbers of the a files in total.

[0127] lowIndex represents the rounding-down operation result of index, roundIndex represents the rounding operation result of index, and fraction represents a difference value obtained by subtracting lowIndex from index.

[0128] In the case that fraction is equal to 0, the actual index serial numbers of the a files are lowIndex; in the case that fraction is not equal to 0, the actual index serial numbers of the a files are roundIndex.

[0129] The a files are determined from the operation files of the same time type under a constraint that the file contents of the a files are different from each other.

[0130] Exemplarily, the file contents of the a files are different from each other, so that more abundant file content of the operation files of the same time type can be reserved and stored, facilitating statistical analysis of the operation files of the production device in different states.

[0131] Exemplarily, in the case that the production device is in a test use state, the a files are determined under the constraint that the file contents of the a files are different from each other, so

that a tester of the production device performs statistical analysis on the operation files of the production device in different states, for example, draws a power consumption curve of the production device.

[0132] The a files are determined from the operation files of the same time type under a constraint that a file content difference between adjacent files exceeds a difference threshold.

[0133] Exemplarily, in the case that the file content difference between adjacent files exceeds the difference threshold, the reserved and stored operation file represents the information of a moment at which the operation state of the production device fluctuates greatly, which can avoid a problem that the operation file is deleted in the case the operation of the production device is abnormal.

[0134] Exemplarily, in the case that the production device is in an actual production use state, the a files are determined under the constraint that the file content difference between adjacent files exceeds the difference threshold, so that the operation files in the case that the operation of the production device is abnormal can be reserved and stored for the maintenance personnel of the production device to perform subsequent cause analysis on the abnormal operation of the production device.

[0135] In step 524, the second file subset is determined from the plurality of operation files based on the first file subset.

[0136] Exemplarily, in the file set, the first file subset and the second file subset are complementary to each other. The operation files in the first file subset and the operation files in the second file subset are different from each other, and the second file subset is constructed by files, excluding the operation files in the first file subset, in the plurality of operation files.

[0137] In step 526, the operation files in the first file subset are reserved and stored, and the operation files in the second file subset are deleted.

[0138] Exemplarily, reserving and storing the operation files in the first file subset may be implemented as not performing an editing operation on the operation files in the first file subset. The operation files in the second file subset are deleted to reserve and store the operation files in the first file subset.

[0139] In summary, in the method according to the embodiments, the file subset to be reserved and stored and the file subset to be deleted are determined based on the relationship between the occupation information and the first storage threshold and the relationship between the occupation information and the second storage threshold. Different file determination methods are provided

for the file subset to be deleted, which fully consider the operation features of the production device in different use phases, so that the operation files are thinned in the case of insufficient storage space, thereby reducing the load on the storage space. Moreover, different numbers of files for different reservation and storage are determined based on different load pressures on the storage space, thereby dynamically reducing the load pressure on the storage space and providing a new storage method for operation files.

[0140] FIG. 5 shows a third flowchart of a method for processing data according to some exemplary embodiments of the present disclosure. The method may be performed by a computer device. That is, on the basis of the embodiments shown in FIG. 4, the method further includes steps 505 and 506. Step 530 may be implemented as steps 532, 534, and 536.

[0141] In step 505, the generation times of the plurality of operation files are acquired.

[0142] Exemplarily, the generation time is a time at which the operation file is written into the storage space or a time at which the operation file is edited. The generation time is acquired at a moment at which the time type of the operation file needs to be determined.

[0143] In step 506, at least two time types are determined for the plurality of operation files based on the generation times.

[0144] Exemplarily, the time type of the operation file is determined based on a time period to which the generation time belongs. Exemplarily, a plurality of operation files belonging to the same time period feature the same time type, and time periods corresponding to different time types have the same duration.

[0145] Exemplarily, the time period may be obtained through division based on a predefined time interval. For example, based on a time interval such as one day, one week, or 12 hours, a plurality of adjacent time periods are obtained through division by using the current moment as the end moment of a time period, or a plurality of adjacent time periods are obtained through division by using the earliest historical moment as the start moment of a time period. The earliest historical moment corresponds to the generation time of a first operation file in the storage space, wherein the first operation file is a file with the earliest generation time.

[0146] It should be noted that steps 505 and 506 in the embodiments may be combined with the steps in FIG. 4 to form new embodiments for separate implementation, which is not limited in the present disclosure. Exemplarily, steps 505 and 506 in the embodiments are performed before step 510, but a case in which any one of steps 505 and 506 is performed after or simultaneously with

step 510 in other embodiments is not excluded.

[0147] In step 532, in the case that the occupation information exceeds the second storage threshold, b files are determined by sampling from the operation files of the same time type, and the third file subset is constructed.

[0148] Exemplarily, each of the plurality of operation files features a corresponding time type, and the generation times of at least two operation files of the same time type both belong to the same time period.

[0149] Exemplarily, b is a positive integer. The number of files in the third file subset is a product of b and the number of time types. In some embodiments, the value of a is greater than that of b.

[0150] It should be noted that, for an implementation method for determining the b files by sampling, reference can be made to the above description of the sampling method for the a files, which is not repeated herein.

[0151] In step 534, the fourth file subset is determined from the plurality of operation files based on the third file subset.

[0152] Exemplarily, in the file set, the third file subset and the fourth file subset are complementary to each other. The operation files in the third file subset and the operation files in the fourth file subset are different from each other, and the fourth file subset is constructed by files, excluding the operation files in the third file subset, in the plurality of operation files.

[0153] In step 536, the operation files in the third file subset are reserved and stored, and the operation files in the fourth file subset are deleted.

[0154] Exemplarily, reserving and storing the operation files in the third file subset may be implemented as not performing an editing operation on the operation files in the third file subset. The operation files in the fourth file subset are deleted to reserve and store the operation files in the third file subset.

[0155] It should be noted that steps 532, 534, and 536 in the embodiments may be combined with steps 510 and 520 in FIG. 3 to form new embodiments for separate implementation, which is not limited in the present disclosure. In the embodiments, steps 505, 506, 532, 534, and 536 may be combined with steps 510 and 520 in FIG. 3 to form new embodiments for separate implementation, which is not limited in the present disclosure.

[0156] In summary, in the method according to the embodiments, the file subset to be reserved and stored and the file subset to be deleted are determined based on the relationship between the

occupation information and the first storage threshold and the relationship between the occupation information and the second storage threshold, so that the operation files are thinned in the case of insufficient storage space, thereby reducing the load on the storage space. Moreover, different numbers of files for different reservation and storage are determined based on different load pressures on the storage space, thereby dynamically reducing the load pressure on the storage space and providing a new storage method for operation files.

[0157] FIG. 6 shows a fourth flowchart of a method for processing data according to some exemplary embodiments of the present disclosure. The method may be performed by a computer device. That is, on the basis of the embodiments shown in FIG. 4, the method further includes steps 515 and 516.

[0158] In step 515, a file cleanup directory is acquired.

[0159] Exemplarily, the file cleanup directory includes the storage paths of a part of the plurality of operation files. A part of operation files in the file cleanup directory is configured to construct the second file subset and/or the fourth file subset. Correspondingly, the storage path of an operation file in the second file subset and/or the fourth file subset belongs to the file cleanup directory.

[0160] Exemplarily, the file cleanup directory is configured to indicate a storage path on which file thinning needs to be performed. In the case that step 515 is separately performed, performing the file thinning refers to deleting a part of operation files in the file cleanup directory (that is, the second file subset and/or the fourth file subset are constructed based on a part of operation files in the file cleanup directory).

[0161] The first file subset includes files in the file cleanup directory excluding the second file subset, and operation files not belonging to the file cleanup directory. The third file subset includes files in the file cleanup directory excluding the fourth file subset, and operation files not belonging to the file cleanup directory.

[0162] It should be noted that step 515 in the embodiments may be combined with the steps in FIG. 3 to form new embodiments for separate implementation, which is not limited in the present disclosure.

[0163] In step 516, a data cleanup label is acquired.

[0164] Exemplarily, the data cleanup label is a file label for which file thinning needs to be performed. Exemplarily, the operation file is provided with a corresponding file label, and in the

case that the file label of the operation file is the same as the data cleanup label, the file is the operation file on which file thinning needs to be performed.

[0165] In some embodiments, in the file cleanup directory, each operation file whose file label is the data cleanup label belongs to a file cleanup set. Exemplarily, a part of operation files whose file labels are the data cleanup label in the file cleanup directory is deleted (that is, the second file subset and/or the fourth file subset are constructed based on a part of operation files whose file labels are the data cleanup label in the file cleanup directory). The second file subset and/or the fourth file subset are subsets of the file cleanup set. Exemplarily, the storage path of an operation file in the second file subset and/or the fourth file subset belongs to the file cleanup directory, and the operation file is provided with the data cleanup label.

[0166] The first file subset includes files in the file cleanup directory excluding the second file subset, and operation files not belonging to the file cleanup directory. The third file subset includes files in the file cleanup directory excluding the fourth file subset, and operation files not belonging to the file cleanup directory.

[0167] In the embodiments, the production device includes a wind turbine generator set. In some embodiments, the data cleanup label is configured to indicate the operation data of the wind turbine generator set in the case that the wind turbine rotation speed of the wind turbine generator set is less than a rotation speed threshold. In the case that the wind turbine rotation speed of the wind turbine generator set is less than the rotation speed threshold, the wind turbine generator set is blown by air with a low wind speed and is in a low power generation state. Exemplarily, the load in the storage space for storing operation files is reduced by deleting a part of the data of the wind turbine generator set in the low power generation state. The deleted operation file in the low power generation state is data of low importance in the wind turbine generator set and does not affect the maintenance or analysis of the operation of the wind turbine generator set.

[0168] In some other embodiments, the data cleanup label is configured to indicate the operation data of the wind turbine generator set in the case that the pitch angle of the wind turbine generator set belongs to an angle range. In the case that the pitch angle of the wind turbine generator set belongs to the angle range, the angle at which the wind turbine generator set is subjected to air blowing is within a preset operation angle, and the wind turbine generator set is in a normal-wind direction operating state. Exemplarily, the load in the storage space for storing operation files is reduced by deleting a part of the data of the wind turbine generator set in the normal-wind direction

operating state. The normal-wind direction operating state will not cause damage to the wind turbine generator set due to air blowing. The deleted operation file in the normal-wind direction operating state is data of low importance in the wind turbine generator set and does not affect the maintenance or analysis of the operation of the wind turbine generator set.

[0169] In summary, in the method according to the embodiments, the file subset to be reserved and stored and the file subset to be deleted are determined based on the relationship between the occupation information and the first storage threshold and the relationship between the occupation information and the second storage threshold. The file range of the operation file to be deleted is constrained by using the acquired file cleanup directory and data cleanup label, so as to preferentially delete the operation file that has less impact on the maintenance or analysis of the production device. Meanwhile, the operation files are thinned in the case of insufficient storage space, thereby reducing the load on the storage space. Moreover, different numbers of files for different reservation and storage are determined based on different load pressures on the storage space, thereby dynamically reducing the load pressure on the storage space and providing a new storage method for operation files.

[0170] FIG. 7 shows a fifth flowchart of a method for processing data according to some exemplary embodiments of the present disclosure. The method may be performed by a computer device. That is, on the basis of the embodiments shown in FIG. 4, the method further includes steps 517 and 540.

[0171] In step 517, a file cleanup type is acquired.

[0172] Exemplarily, the file cleanup type refers to the file data type of a part of the plurality of operation files. Exemplarily, the file data type indicated by the file cleanup type may be at least one of: a distributed file system, for example, a Seaweed File System (SeaweedFS); an object storage file system, for example, an Amazon Simple Storage Service (Amazon S3); or an operating system file system, for example, a Linux file system. Exemplarily, the file data type of an operation file that needs to be deleted is indicated by recording the file cleanup type in storage media parameters.

[0173] Exemplarily, a part of operation files of the file cleanup type is used to construct the second file subset and/or the fourth file subset. Correspondingly, the file data type of an operation file in the second file subset and/or the fourth file subset is of the file cleanup type. Operation files in different data storage methods that need to be deleted are indicated by acquiring the file cleanup

type.

[0174] It should be noted that step 517 in the embodiments may be combined with the steps in FIG. 3 to form new embodiments for separate implementation, which is not limited in the present disclosure.

[0175] In step 540, in the case that the occupation information exceeds a third storage threshold, a part of operation files at a head position in a file sequence is deleted.

[0176] Exemplarily, in the case that the occupation information exceeds the third storage threshold, the storage space reaches a third-level load pressure for storing operation files. The third storage threshold is greater than the second storage threshold. The load in the storage space for storing operation files is reduced by deleting a part of operation files at the head position in the file sequence.

[0177] In the case that the occupation information exceeds the third storage threshold, the storage space is at risk of being fully occupied. The occupation information of the storage space is reduced to less than the third storage threshold by deleting a part of operation files at the head position in the file sequence. Exemplarily, after the part of operation files is deleted, the occupation information does not exceed the third storage threshold.

[0178] Exemplarily, the file sequence is obtained by arranging the plurality of operation files by the generation times from the earliest to the latest. For example, an operation file that has been generated for a long time is preferentially deleted.

[0179] It should be noted that step 540 in the embodiments may be combined with the steps in FIG. 3 to form new embodiments for separate implementation, which is not limited in the present disclosure.

[0180] FIG. 8 shows another schematic diagram of a method for processing data according to some exemplary embodiments of the present disclosure.

[0181] In the case that the occupation information does not exceed the first storage threshold 332, for example, the occupation information does not exceed 60%, for operation files whose generation times belong to the first time period, the storage space reserves and stores a plurality of operation files, for example, fully stores all operation files.

[0182] In the case that the occupation information exceeds the first storage threshold 332 and does not exceed the second storage threshold 334, for example, the occupation information exceeds 60% and does not exceed 80%, for operation files whose generation times belong to the first time

period, a medium number of operation files are reserved and stored, for example, a first number of operation files in the first time period are reserved and stored.

[0183] In the case that the occupation information exceeds the second storage threshold 334, for example, the occupation information exceeds 80%, for operation files whose generation times belong to the first time period, a small number of operation files are reserved and stored, for example, a second number of operation files in the first time period are reserved and stored. The second number is less than the first number.

[0184] In the case that the occupation information exceeds the third storage threshold 336, for example, the occupation information exceeds 90%, the storage space is at risk of being fully occupied. The operation files are deleted by the generation times from the earliest to the latest, and thus the occupation information of the storage control is reduced to less than the third storage threshold.

[0185] Exemplarily, the file cleanup type is recorded in storage media parameters to indicate the file data type of an operation file that needs to be deleted, so that the method for deleting and reserving and storing operation files is applicable to a plurality of file data types.

[0186] In summary, in the method according to the embodiments, the file subset to be reserved and stored and the file subset to be deleted are determined based on the relationship between the occupation information and the first storage threshold and the relationship between the occupation information and the second storage threshold, so that the operation files are thinned in the case of insufficient storage space, thereby reducing the load on the storage space. Moreover, different numbers of files for different reservation and storage are determined based on different load pressures on the storage space, and in the case that the storage space is at risk of being fully occupied, the occupation information of the storage control is reduced to less than the third storage threshold by deleting a part of operation files at the head position in the file sequence, thereby dynamically reducing the load pressure on the storage space and providing a new storage method for operation files.

[0187] FIG. 9 shows a sixth flowchart of a method for processing data according to some exemplary embodiments of the present disclosure. The method may be performed by a computer device. That is, on the basis of the embodiments shown in FIG. 4, the method further includes steps 502 and 504.

[0188] In step 502, the current time information and a historical reading time of the computer

device are acquired.

[0189] Exemplarily, the historical reading time is a time at which the computer device transmits a historical operation file to a reading device. In some embodiments, since the production device is deployed in a wild area, data transmission over networks is challenging, and therefore, the operation and maintenance personnel copy the operation files from the storage space using the carried reading device. The reading device is a device for copying the operation files from the storage space. Exemplarily, the historical reading time is a time, recorded by the computer device, at which the historical operation file is transmitted to the reading device.

[0190] In step 504, a first storage threshold and/or a second storage threshold are determined based on the storage duration between the current time information and the historical reading time.

[0191] Exemplarily, the storage duration is in a positive correlation with the first storage threshold, and/or the storage duration is in a positive correlation with the second storage threshold. As the storage duration increases, a probability that the computer device transmits the operation files to the reading device again increases. For example, as the duration between the current time and the historical reading time increases, a probability that the operation and maintenance personnel copy the operation files from the storage space using the carried reading device increases, and the determined first storage threshold and/or the determined second storage threshold increase, so that more space in the storage space is allowed to store the operation files.

[0192] In some embodiments, the method further includes:

[0193] ·A historical reading interval is acquired.

[0194] Exemplarily, the historical reading interval is configured to indicate an adjacent time interval in which the computer device transmits the historical operation file to the reading device. In some embodiments, since the production device is deployed in a wild area, data transmission over networks is challenging, and therefore, the operation and maintenance personnel periodically copy the operation files from the storage space using the carried reading device. The historical reading interval is a time interval between two consecutive times that the operation and maintenance personnel copy the operation files from the storage space using the carried reading device.

[0195] Correspondingly, step 504 may be implemented as:

[0196] ·The storage duration between the current time information and the historical reading time is determined.

[0197] Exemplarily, the storage duration is a time interval between the current time information and the historical reading time. Exemplarily, the historical reading time is a time at which the operation and maintenance personnel last copied the operation files from the storage space using the carried reading device.

[0198] The first storage threshold and/or the second storage threshold are determined based on a duration ratio of the storage duration in the historical reading interval.

[0199] Exemplarily, as the duration ratio increases, the current time approaches the next scheduled time at which the operation and maintenance personnel copy the operation files from the storage space using the carried reading device, and the determined first storage threshold and/or the determined second storage threshold increase, so that more space in the storage space is allowed to store the operation files. Exemplarily, the duration ratio is in a positive correlation with the first storage threshold, and/or the duration ratio is in a positive correlation with the second storage threshold.

[0200] In summary, in the method according to the embodiments, the file subset to be reserved and stored and the file subset to be deleted are determined based on the relationship between the occupation information and the first storage threshold and the relationship between the occupation information and the second storage threshold, so that the operation files are thinned in the case of insufficient storage space, thereby reducing the load on the storage space. The storage thresholds are determined based on the historical reading time and the historical reading interval, which fully considers the actual operating state of copying the operation files from the storage space by the operation and maintenance personnel using the carried reading device, thereby dynamically determining the load pressure borne by the storage space based on time and making full use of the storage space to store operation files.

[0201] In some embodiments of the present disclosure, the production device includes a wind turbine generator set. Correspondingly:

[0202] The wind turbine file occupation information of storage space by a plurality of wind turbine operation files of the wind turbine generator set is acquired.

[0203] Exemplarily, the storage space is configured to store the wind turbine operation files of the wind turbine generator set, and the wind turbine operation file records at least one of information such as the wind turbine rotation speed, generation power, or blade pitch angle of the wind turbine generator set in a process where the wind turbine generator set generates electricity

by means of air blowing.

[0204] In the case that the wind turbine file occupation information exceeds a first storage threshold and does not exceed a second storage threshold, wind turbine operation files in a first file subset are reserved and stored, and wind turbine operation files in a second file subset are deleted.

[0205] Exemplarily, in the case that the wind turbine file occupation information exceeds the first storage threshold and does not exceed the second storage threshold, the storage space reaches a first-level load pressure for storing wind turbine operation files. The load in the storage space for storing wind turbine operation files is reduced by deleting a part of wind turbine operation files in the storage space.

[0206] Exemplarily, in a file set including a plurality of wind turbine operation files, the first file subset and the second file subset are complementary to each other. The files, excluding the wind turbine operation files in the second file subset, in the plurality of wind turbine operation files are the wind turbine operation files in the first file subset.

[0207] In the case that the wind turbine file occupation information exceeds the second storage threshold, wind turbine operation files in a third file subset are reserved and stored, and wind turbine operation files in a fourth file subset are deleted.

[0208] In the case that the wind turbine file occupation information exceeds the second storage threshold, the storage space reaches a second-level load pressure for storing wind turbine operation files. The load in the storage space for storing wind turbine operation files is reduced by deleting a part of wind turbine operation files in the storage space.

[0209] Exemplarily, in the file set, the third file subset and the fourth file subset are complementary to each other. The files, excluding the wind turbine operation files in the fourth file subset, in the plurality of wind turbine operation files are the wind turbine operation files in the third file subset.

[0210] It should be noted that the first file subset includes a first number of wind turbine operation files whose generation times belong to a first time period, and the third file subset includes a second number of wind turbine operation files whose generation times belong to the first time period, the first number being greater than the second number.

[0211] Those of ordinary skill in the art can understand that the above embodiments may be implemented independently, or the above embodiments may be freely combined to form new

embodiments to implement the method for processing data according to the present disclosure.

[0212] FIG. 10 shows a structural block diagram of an apparatus for processing data according to some exemplary embodiments of the present disclosure. The apparatus includes an acquiring module 810 and a processing module 820.

[0213] The acquiring module 810 is configured to acquire the occupation information of storage space, the occupation information being configured to indicate the occupation of the storage space by a plurality of operation files of a production device.

[0214] The processing module 820 is configured to reserve and store operation files in a first file subset and delete operation files in a second file subset in the case that the occupation information exceeds a first storage threshold and does not exceed a second storage threshold, where in a file set including the plurality of operation files, the first file subset and the second file subset are complementary to each other.

[0215] The processing module 820 is further configured to reserve and store operation files in a third file subset and delete operation files in a fourth file subset in the case that the occupation information exceeds the second storage threshold, where in the file set, the third file subset and the fourth file subset are complementary to each other.

[0216] The first file subset includes a first number of operation files whose generation times belong to a first time period, and the third file subset includes a second number of operation files whose generation times belong to the first time period, the first number being greater than the second number.

[0217] In some embodiments, each of the plurality of operation files features a corresponding time type, and the generation times of at least two operation files of the same time type both belong to the same time period. The processing module 820 is further configured to:

[0218] in the case that the occupation information exceeds the first storage threshold and does not exceed the second storage threshold, determine a files by sampling from operation files of the same time type, and construct the first file subset, a being a positive integer;

[0219] determine the second file subset from the plurality of operation files based on the first file subset; and

[0220] reserve and store the operation files in the first file subset and delete the operation files in the second file subset.

[0221] In some embodiments, the processing module 820 is further configured to perform at least

one of:

[0222] determining the a files from the operation files of the same time type in a random manner;

[0223] determining the a files from the operation files of the same time type in a manner that the generation time interval of adjacent files is a predefined time interval;

[0224] determining the a files from the operation files of the same time type under a constraint that the file contents of the a files are different from each other; or

[0225] determining the a files from the operation files of the same time type under a constraint that a file content difference between adjacent files exceeds a difference threshold.

[0226] In some embodiments, each of the plurality of operation files features a corresponding time type, and the generation times of at least two operation files of the same time type both belong to the same time period. The processing module 820 is further configured to:

[0227] in the case that the occupation information exceeds the second storage threshold, determine b files by sampling from the operation files of the same time type, and construct the third file subset, b being a positive integer;

[0228] determine the fourth file subset from the plurality of operation files based on the third file subset; and

[0229] reserve and store the operation files in the third file subset and delete the operation files in the fourth file subset.

[0230] In some embodiments, the acquiring module 810 is further configured to:

[0231] acquire the generation times of the plurality of operation files;

[0232] the processing module 820 is further configured to:

[0233] determine at least two time types for the plurality of operation files based on the generation times, the durations of time periods corresponding to different time types being the same.

[0234] In some embodiments, the acquiring module 810 is further configured to:

[0235] acquire a file cleanup directory, the file cleanup directory including the storage paths of a part of the plurality of operation files, where the storage path of an operation file in the second file subset and/or the fourth file subset belongs to the file cleanup directory.

[0236] In some embodiments, the production device includes a wind turbine generator set; the acquiring module 810 is further configured to:

[0237] acquire a data cleanup label, where the storage path of an operation file in the second file

subset and/or the fourth file subset belongs to the file cleanup directory, and the operation file is provided with the data cleanup label.

[0238] The data cleanup label is configured to indicate the operation data of the wind turbine generator set in the case that the wind turbine rotation speed of the wind turbine generator set is less than a rotation speed threshold or the operation data of the wind turbine generator set in the case that the pitch angle of the wind turbine generator set belongs to an angle range.

[0239] In some embodiments, the processing module 820 is further configured to:

[0240] delete a part of operation files at a head position in a file sequence in the case that the occupation information exceeds a third storage threshold, where the file sequence is obtained by arranging the plurality of operation files by the generation times from the earliest to the latest, where the occupation information does not exceed the third storage threshold after the part of operation files is deleted.

[0241] In some embodiments, the acquiring module 810 is further configured to:

[0242] acquire a file cleanup type, the file cleanup type is the file data type of a part of files in the plurality of operation files, where the file data type of an operation file in the second file subset and/or the fourth file subset belongs to the file cleanup type.

[0243] In some embodiments, the acquiring module 810 is further configured to:

[0244] acquire the current time information and a historical reading time of the computer device, the historical reading time being a time at which the computer device transmits a historical operation file to a reading device;

[0245] the processing module 820 is further configured to:

[0246] determine the first storage threshold and/or the second storage threshold based on the storage duration between the current time information and the historical reading time, where the storage duration is in a positive correlation with the first storage threshold, and/or the storage duration is in a positive correlation with the second storage threshold.

[0247] In some embodiments, the processing module 820 is further configured to:

[0248] acquire a historical reading interval, the historical reading interval being configured to indicate an adjacent time interval in which the computer device transmits the historical operation file to the reading device;

[0249] the processing module 820 is further configured to:

[0250] determine the storage duration between the current time information and the historical

reading time; and

[0251] determine the first storage threshold and/or the second storage threshold based on the duration ratio of the storage duration in the historical reading interval, where the duration ratio is in a positive correlation with the first storage threshold, and/or the duration ratio is in a positive correlation with the second storage threshold.

[0252] In some embodiments, the production device includes a wind turbine generator set;

[0253] the acquiring module 810 is further configured to:

[0254] acquire the wind turbine file occupation information of the storage space by a plurality of wind turbine operation files of the wind turbine generator set;

[0255] the processing module 820 is further configured to:

[0256] reserve and store wind turbine operation files in the first file subset and delete wind turbine operation files in the second file subset in the case that the wind turbine file occupation information exceeds the first storage threshold and does not exceed the second storage threshold; and

[0257] reserve and store wind turbine operation files in the third file subset and delete wind turbine operation files in the fourth file subset in the case that the wind turbine file occupation information exceeds the second storage threshold.

[0258] It should be noted that in the case that the apparatus according to the above embodiments implements the functions thereof, the division of the functional modules is merely exemplary. In practical applications, the above functions may be assigned to different functional modules according to actual needs, that is, the internal structure of the device may be divided into different functional modules, so as to implement all or a part of the above functions.

[0259] With regard to the apparatus in the above embodiments, the specific manner in which each module performs the operation has been described in detail in the embodiments related to the method. The technical effects achieved by each module performing the operation are the same as the technical effects in the embodiments related to the method, which will not be described in detail herein.

[0260] The embodiments of the present disclosure further provide a computer device, including a processor and a memory storing one or more computer programs. The one or more computer programs, when executed by the processor, cause the computer device to perform the method for processing data according to the above method embodiments.

[0261] In some embodiments, the computer device is a server. Exemplarily, FIG. 11 is a structural block diagram of a server according to some exemplary embodiments of the present disclosure.

[0262] Typically, the server 200 includes a processor 2301 and a memory 2302.

[0263] The processor 2301 includes one or more processing cores, such as a 4-core processor and an 8-core processor. In some embodiments, the processor 2301 is implemented in at least one hardware form of a digital signal processor (DSP), a field-programmable gate array (FPGA), or a programmable logic array (PLA). The processor 2301 further includes a main processor and a coprocessor. The main processor is a processor configured to process data in an awake state, and is also referred to as a central processing unit (CPU). The coprocessor is a low-power processor configured to process data in a standby state. In some embodiments, the processor 2301 is integrated with a graphics processing unit (GPU) that is responsible for rendering and drawing contents that need to be displayed on a display screen. In some embodiments, the processor 2301 further includes an artificial intelligence (AI) processor for processing computing operations related to machine learning.

[0264] The memory 2302 includes one or more computer-readable storage media, which are non-transitory. The memory 2302 further includes a high-speed random access memory, and a non-volatile memory, such as one or more magnetic disk storage devices and flash storage devices. In some embodiments, the non-transitory computer-readable storage medium in the memory 2302 is configured to store at least one instruction. The at least one instruction, when executed by the processor 2301, causes the processor to perform the method for processing data according to the method embodiments of the present disclosure.

[0265] In some embodiments, the server 200 further includes an input interface 2303 and an output interface 2304. The processor 2301, the memory 2302, the input interface 2303, and the output interface 2304 are connected via buses or signal lines. The peripheral devices are connected to the input interface 2303 and the output interface 2304 via a bus, signal line, or circuit board. The input interface 2303 and the output interface 2304 are configured to connect at least one peripheral device related to input/output (I/O) to the processor 2301 and the memory 2302. In some embodiments, the processor 2301, the memory 2302, the input interface 2303, and the output interface 2304 are integrated on the same chip or circuit board. In some other embodiments, any one or two of the processor 2301, the memory 2302, the input interface 2303, and the output interface 2304 are implemented on a separate chip or circuit board, which is not limited in the

embodiments of the present disclosure.

[0266] Those skilled in the art understand that the structures illustrated above do not constitute a limitation on the server 200, and that more or fewer components than those illustrated are included, or some of the components are combined, or a different arrangement of components is employed.

[0267] In some exemplary embodiments, a chip is further provided. The chip includes one or more programmable logic circuits and/or one or more program instructions. The chip, when running on a computer device, is caused to perform the method for processing data according to the above embodiments.

[0268] In some exemplary embodiments, a computer program product is further provided. The computer program product includes one or more computer instructions stored in a computer-readable storage medium. A processor of a computer device reads the one or more computer instructions from the computer-readable storage medium. The one or more computer instructions, when read from the computer-readable storage medium and executed by the processor, cause the computer device to perform the method for processing data according to the above method embodiments.

[0269] In some exemplary embodiments, a computer-readable storage medium is further provided. The computer-readable storage medium stores one or more computer programs. The one or more computer programs, when loaded and executed by a processor, cause the processor to perform the method for processing data according to the above method embodiments.

[0270] It should be understood by those of ordinary skill in the art that all or part of the steps for implementing the above embodiments are completed by hardware, or are completed by instructing relevant hardware by a program stored in a computer-readable storage medium. The storage medium mentioned above is a read-only memory, a magnetic disk, a compact disk, or the like.

[0271] Those skilled in the art should recognize that in the one or more embodiments described above, the functions described in the embodiments of the present disclosure are implemented using hardware, software, firmware, or any combination thereof. The functions, when implemented using software, are stored in a computer-readable medium or transmitted as one or more instructions or codes on the computer-readable medium. The computer-readable medium includes a computer storage medium and a communication medium, where the communication medium includes any medium that facilitates the transfer of a computer program from one place to another. The storage medium is any available medium that is accessible by a general-purpose or special-

purpose computer.

[0272] Described above are merely optional embodiments of the present disclosure and are not intended to limit the present disclosure. Any modifications, equivalents, improvements, and the like, made within the spirit and principle of the present disclosure should fall within the protection scope of the present disclosure.

CLAIMS

1. A method for processing data, performed by a computer device, comprising:

acquiring current time information and a historical reading time of the computer device, the historical reading time being a time at which the computer device transmits a historical operation file to a reading device;

acquiring a historical reading interval, the historical reading interval being configured to indicate an adjacent time interval in which the computer device transmits the historical operation file to the reading device;

determining a storage duration between the current time information and the historical reading time;

determining a first storage threshold and a second storage threshold based on a duration ratio of the storage duration in the historical reading interval, wherein the duration ratio is in a positive correlation with the first storage threshold, and the duration ratio is in a positive correlation with the second storage threshold;

acquiring wind turbine file occupation information of storage space, the wind turbine file occupation information being configured to indicate occupation of the storage space by a plurality of wind turbine operation files of a wind turbine generator set;

in a case that the wind turbine file occupation information exceeds the first storage threshold and does not exceed the second storage threshold, determining a files from wind turbine operation files of a same time type for the wind turbine generator set in a test use state under a constraint that file contents of the a files are different from each other and/or for the wind turbine generator set in an actual production use state under a constraint that a file content difference between adjacent files exceeds a difference threshold, and constructing a first file subset, wherein generation times of at least two wind turbine operation files of a same time type both belong to a same time period, and a is a positive integer;

acquiring a file cleanup directory, the file cleanup directory comprising storage paths of a part of the plurality of wind turbine operation files; and determining a second file subset from the plurality of wind turbine operation files based on the first file subset, wherein a storage path of a wind turbine operation file in the second file subset and/or a fourth file subset belongs to the file cleanup directory, and the wind turbine operation file is provided with a data cleanup label, the data cleanup label being configured to indicate operation data of the wind turbine generator set in

a case that a wind turbine rotation speed of the wind turbine generator set is less than a rotation speed threshold or operation data of the wind turbine generator set in a case that a pitch angle of the wind turbine generator set belongs to an angle range;

reserving and storing wind turbine operation files in the first file subset and deleting wind turbine operation files in the second file subset, wherein the first file subset comprises files in the file cleanup directory excluding the second file subset, and operation files not belonging to the file cleanup directory; and

reserving and storing wind turbine operation files in a third file subset and deleting wind turbine operation files in the fourth file subset in a case that the wind turbine file occupation information exceeds the second storage threshold, wherein the third file subset comprises files in the file cleanup directory excluding the fourth file subset, and operation files not belonging to the file cleanup directory,

wherein the first file subset comprises a first number of operation files whose generation times belong to a first time period, and the third file subset comprises a second number of operation files whose generation times belong to the first time period, the first number being greater than the second number.

2. The method according to claim 1, wherein determining the a files by sampling from the wind turbine operation files of the same time type further comprises at least one of:

determining the a files from the wind turbine operation files of the same time type in a random manner; or

determining the a files from the wind turbine operation files of the same time type in a manner that a generation time interval of adjacent files is a predefined time interval.

3. The method according to claim 1, wherein each of the plurality of wind turbine operation files features a corresponding time type, and generation times of at least two wind turbine operation files of a same time type both belong to a same time period;

reserving and storing the wind turbine operation files in the third file subset and deleting the wind turbine operation files in the fourth file subset in the case that the wind turbine file occupation information exceeds the second storage threshold comprises:

in the case that the wind turbine file occupation information exceeds the second storage

threshold, determining b files by sampling from the wind turbine operation files of the same time type, and constructing the third file subset, b being a positive integer;

determining the fourth file subset from the plurality of wind turbine operation files based on the third file subset; and

reserving and storing the wind turbine operation files in the third file subset and deleting the wind turbine operation files in the fourth file subset.

4. The method according to claim 1 or 3, further comprising:

acquiring generation times of the plurality of wind turbine operation files; and

determining at least two time types for the plurality of wind turbine operation files based on the generation times, durations of time periods corresponding to different time types being the same.

5. The method according to any one of claims 1 to 3, further comprising:

deleting a part of wind turbine operation files at a head position in a file sequence in a case that the wind turbine file occupation information exceeds a third storage threshold, wherein the file sequence is obtained by arranging the plurality of wind turbine operation files by the generation times from the earliest to the latest,

wherein the wind turbine file occupation information does not exceed the third storage threshold after the part of wind turbine operation files is deleted.

6. The method according to any one of claims 1 to 3, further comprising:

acquiring a file cleanup type, the file cleanup type is a file data type of a part of the plurality of wind turbine operation files,

wherein a file data type of a wind turbine operation file in the second file subset and/or the fourth file subset belongs to the file cleanup type.

7. An apparatus for processing data, comprising an acquiring module and a processing module, wherein:

the acquiring module is configured to acquire current time information and a historical reading time of a computer device, the historical reading time being a time at which the computer

device transmits a historical operation file to a reading device;

the acquiring module is further configured to acquire a historical reading interval, the historical reading interval being configured to indicate an adjacent time interval in which the computer device transmits the historical operation file to the reading device;

the processing module is configured to determine a storage duration between the current time information and the historical reading time;

the processing module is further configured to determine a first storage threshold and a second storage threshold based on a duration ratio of the storage duration in the historical reading interval, wherein the duration ratio is in a positive correlation with the first storage threshold, and the duration ratio is in a positive correlation with the second storage threshold, wherein

the acquiring module is further configured to acquire wind turbine file occupation information of storage space, the wind turbine file occupation information being configured to indicate occupation of the storage space by a plurality of wind turbine operation files of a wind turbine generator set;

the processing module is further configured to:

in a case that the wind turbine file occupation information exceeds the first storage threshold and does not exceed the second storage threshold, determine a files from wind turbine operation files of a same time type for the wind turbine generator set in a test use state under a constraint that file contents of the a files are different from each other and/or for the wind turbine generator set in an actual production use state under a constraint that a file content difference between adjacent files exceeds a difference threshold, and construct a first file subset, wherein generation times of at least two wind turbine operation files of a same time type both belong to a same time period, and a is a positive integer;

acquire a file cleanup directory, the file cleanup directory comprising storage paths of a part of the plurality of wind turbine operation files; and determine a second file subset from the plurality of wind turbine operation files based on the first file subset, wherein a storage path of a wind turbine operation file in the second file subset and/or a fourth file subset belongs to the file cleanup directory, and the wind turbine operation file is provided with a data cleanup label, the data cleanup label being configured to indicate operation data of the wind turbine generator set in a case that a wind turbine rotation speed of the wind turbine generator set is less than a rotation speed threshold or operation data of the wind turbine generator set in a case that a pitch angle of the wind turbine

generator set belongs to an angle range; and

reserve and store wind turbine operation files in the first file subset and delete wind turbine operation files in the second file subset, wherein the first file subset comprises files in the file cleanup directory excluding the second file subset, and operation files not belonging to the file cleanup directory; and

the processing module is further configured to reserve and store wind turbine operation files in a third file subset and delete wind turbine operation files in the fourth file subset in a case that the wind turbine file occupation information exceeds the second storage threshold, wherein the third file subset comprises files in the file cleanup directory excluding the fourth file subset, and operation files not belonging to the file cleanup directory,

wherein the first file subset comprises a first number of operation files whose generation times belong to a first time period, and the third file subset comprises a second number of operation files whose generation times belong to the first time period, the first number being greater than the second number.

8. A computer device, comprising: a processor and a memory storing at least one program, wherein the at least one program, when executed by the processor, causes the computer device to perform the method for processing data as defined in any one of claims 1 to 6.

9. A computer-readable storage medium, storing one or more executable instructions, wherein the one or more executable instructions, when loaded and executed by a processor, cause the processor to perform the method for processing data as defined in any one of claims 1 to 6.

10. A computer program product, comprising one or more computer instructions stored in a computer-readable storage medium, wherein the one or more computer instructions, when read from the computer-readable storage medium and executed by a processor, cause the processor to perform the method for processing data as defined in any one of claims 1 to 6.