

DATA VALIDATION ENGINE FOR AN ENERGY MANAGEMENT SYSTEM

Background

5 Electrical power consumption has increased over the years from the heavy usages in residential sectors, industrial sectors (specifically in manufacturing) and commercial sectors. Hence, there is a significant interest in ways to improve usage of electrical energy. In fact, United States (US) Department Of Energy (DOE) has noted that Energy Saving Performance Contracts (ESPC), contracts which provides services to monitor and optimize electrical energy usage, has resulted in 20% reduction in electrical energy use and its related cost.

15 There are multiple systems available in the market to manage energy. An example of an energy management system is monitoring software that helps to monitor the electricity usage. Such monitoring software may create awareness amongst the electricity user, whom may reduce the usage. Another example of an energy management system includes usage of energy efficient hardware. This energy efficient hardware may consume less electrical energy and therefore reduce energy wastage. However, the energy management system that uses energy efficient hardware is generally costly, and therefore not preferred.

25 These available energy management systems in the market have shown multiple drawbacks. For example, these energy management systems have shown data gaps in collection and aggregation of data. Furthermore, the available energy management system only provides method of identifying an error but not on overcoming/eliminating the errors.

Summary

Embodiments in here provide a data validation system for an energy management system and methods of performing data validation in accordance with one embodiment of the present invention.

In one embodiment, a method validating data for an energy management system include a step to receive a primary data from a plurality of apparatus within the energy management system. The method further includes a step to identify outliers within the primary data. The identification may be performed using a statistical computation of the primary data. Subsequently, the method includes a step to filter out the outlier from the primary data. Finally, the method includes a step of processing and/or profiling the filtered primary data. Furthermore, the method includes a step to validate the primary data based on a on a global level statistical computation of the primary data.

In another embodiment, a system for validating data in an energy management system includes a personal computer executing a computer code to perform data validation in the energy management system. The computer code includes a data controversy engine, a data validation engine and a data processing and profiling engine. The data controversy engine receives the primary data. The data controversy engine computes statistics for the primary data and identifies outliers within the primary data. The data processing and profiling engine receives a filtered primary data in which the identified outliers are removed. The data processing and profiling engine extracts patterns from the filtered primary data. The data validation engine receives the processed and/or profiled primary data from the data processing and profiling

engine. The data validation engine validates the primary data by comparing to predefined domain knowledge.

In an alternative embodiment, a non-transitory medium storing computer-readable, computer-executable process steps for an energy management system includes a process having a step of receiving primary data from a plurality of apparatus within the energy management system. The process further includes identifying outliers within the primary data based on a statistical computation of the primary data. Furthermore, the process includes a step to filter out the outlier from the primary data. The step includes a step of profiling the primary data for extracting patterns from the filtered primary data. The process also includes a step to validate the primary data based on a on a global level statistical computation of the primary data.

Further features of the invention, its nature, and various advantages will be more apparent from the accompanying drawings and the following detailed description.

Brief description of figures

FIG. 1 shows an illustrative energy management system in accordance with one embodiment of the present invention.

FIG. 2 shows an illustrative data validation system in accordance with one embodiment of present invention.

FIG. 3 shows a flowchart of an illustrative method of validating data in accordance with one embodiment of the present invention.

FIG. 4 shows an illustrative computer system implementing an energy management system in accordance with one embodiment of the present invention.

Detailed description

5 The making and using of the presently preferred embodiments are discussed in detail below. The specific embodiments discussed are merely illustrative of specific ways to make and use the invention, and do not limit the scope of the invention.

10 Embodiments in here provide a data validation system for an energy management system and methods of performing data validations in accordance with one embodiment of the present invention.

15 FIG. 1, meant to be illustrative and not limiting, illustrates an energy management system in accordance with one embodiment of the present invention. The energy management system 100 includes multiple components. Amongst the components are multiple sensors 110(1) - 110(N), data gateway 120 and data validation system 130.

20 Energy management system 100 may be implemented in an environment that requires the energy usage optimization and wastage management. In one exemplary embodiment, energy management system 100 may be implemented in a factory, house, commercial building, etc.

25 Energy management system 100 may be utilized to perform energy usage optimization and wastage management. In one example, energy management system 100 may collect, segregate, analyze and model received data. Furthermore, energy management system 100 may validate the received data.

Such validation may help to generate optimized solutions that are free from skewed or outlier data.

In one embodiment, the modelled data may then be supplied to next stage of the energy management system for further processing. Energy management system 100 may be modelled based on internet of things (IoT), in one embodiment. It should be appreciated that the internet of things may be referring to a network of physical objects that features an internet protocol (IP) address for internet connectivity, and the communication that occurs between these objects and other internet-enabled devices and systems.

Each sensor 110(1) - 110(N) may sense and provide operational statuses of a particular apparatus, operational durations, and a total amount of power consumed by the apparatus. In addition, each sensor 110(1) - 110(N) may also sense and provide information on temperature and electricity that is consumed by the particular apparatus.

In one embodiment, sensors 110 (1) - 110(N) may be temperature sensors that may sense and provide temperature of the apparatus. In another embodiment, sensors 110(1) - 110(N) may be electric current sensors that may sense and provide electricity amount consumed by the apparatus. In another embodiment, sensors 110(1) - 110(N) may be sound sensors that may sense and provide information on the loudness of the apparatus (in decibels). However, it should be appreciated that sensors 110(1) - 110(N) may also be other types of sensors that the user may be interested to obtain and process. Furthermore, sensors 110(1) - 110(N) may be also a mixture of multiple types of sensors mentioned above. For example, sensor 110(1) may be a temperature sensor, sensor 110(2) may

be an electric current sensor and sensor 110(3) may be a sound sensor.

5 In addition, sensors 110(1) - 110(N) may provide data in a different format or arrangement based on a location to which energy management system 100 may be implemented. In a data center, sensors 110(1) - 110(N) may sense and provide data based on a distributed network load, i.e., providing sensing data for each data load.

10 In one embodiment, sensors 110(1) - 110(N) may be placed at different locations within the apparatus. The purpose of placing the sensors at different locations within the apparatus is for a clear representation of a single parameter at different locations within the apparatus. For example, sensors 110(1) may be a temperature sensor placed
15 external to the apparatus casing/housing to sense external temperatures surrounding the apparatus. Sensor 110(2), however, may be another temperature sensor that is placed within the apparatus casing to sense an internal temperature within the apparatus. Another reason whereby sensors 110(1) -
20 110(N) may be placed at different locations is to sense different parameters. For example, an electric current sensor is placed along an electric current path.

In one embodiment, the value of N may be similar to a number of apparatus that are present in the system.
25 Alternatively, the value of N may be less than a number apparatus when more than one sensor is implemented for an apparatus. For example, when three sensors are placed in each of twenty (20) apparatuses and the value of N may be sixty (60).

The data generated by sensors 110(1) - 110(N) are primary data. These primary data are supplied to data gateway 120. In one embodiment, the data may be fed into data gateway 120 using a wireless means or using a wired means. The data
5 may be transferred from sensors 110(1) - 110(N) through an established wireless fidelity (WiFi) protocol, such as IEEE 802.11.

In one embodiment, data gateway 120 may be a router or a server. Data gateway 120 may be utilized for receiving
10 the data and collecting the data. The collected data may be in a raw form. When sensors 110(1) - 110(N) are wireless sensors, data gateway 120 may be a standard wireless router.

The collected data is then supplied to data validation system 130. Data validation system 130 may be a
15 sequence of codes that is implemented through a computer code and that may be executed on a computer. Once the data is supplied to data validation system 130, the data may be validated. In one embodiment, data validation system 130 may identify outliers after a statistical computation is performed
20 and/or the outlier is inconsistent with fundamental measurements like temperature, power, etc., using properties as should be reflected by physics. Data validation system 130 may also help to arrange the data in a manner that enables for modelling and analysis.

25 Subsequently, the data may then be used for further processing within energy management system 100.

FIG. 2, meant to be illustrative and not limiting, illustrates a data validation system in accordance with one embodiment of present invention.

Data validation system 200 may be similar to data validation system 130 of FIG. 1. In one embodiment, data validation system may be divided into portions of executable codes within the software. These executable codes perform data mining, data segregations, and data elimination.

The processes within data validation system 200 are executed using a processor within a computer. The processor could be a general purpose processor or a specific task related processor. It should be appreciated that a specific-task processor may perform the prescribed task relatively faster than a general-purpose processor. However, the specific-task processor may not be used for other tasks. The executable codes may be written in one or more programming language. In one exemplary embodiment, executable codes are written in C/C++ languages.

In one embodiment, data validation system 200 includes five steps. These five steps are performed by five different parts of codes, in one embodiment. As shown in the exemplary embodiment of FIG. 2, these five different parts of codes include system configuration 210, data controversy engine 220, data filtering 230 and data process and profiling engine 240, and data validation engine 250. These five parts of codes may occur in a sequential manner, as shown by the direction of arrows in the FIG. 2, but not necessary. Furthermore, the parts of codes may also include a loop-back so that the validity of data may be monitored with respect to consistency.

In one embodiment, primary data may include consumption information on total amount of power, temperature, flow and pressure. Preferably, the power computed can be further used for calculation of the system and/or apparatus

efficiency. The flow and temperature can be combined and computed to derive capacity of the system and/or apparatus. It should be appreciated that there may be other types of primary data that are not described in here. The apparatus may be any
5 machine and/or tool. Alternatively, the apparatus may be domestic appliances such as air conditioning unit, refrigerator, etc. The primary data may be collected using different typed of sensors mounted on the apparatus. For example, the primary data may be collected at different time
10 instances, for different statuses, and etc.

As shown in the embodiment of FIG. 2, the primary data is then supplied to system configuration 210. System configuration 210 may be configured based on where the energy management system (e.g., energy management system 100 of FIG.
15 1) may be placed. In one embodiment, system configuration 210 may adapted to perform data validation mechanism from one or more apparatuses. For example, when the energy management system is a data center, system configuration 210 may be configured to receive data from the sensors located within the
20 server. Alternatively, when the apparatus is a machine on an industrial floor, system configuration 210 may be configured to receive data from sensors located in the machines/tools.

Furthermore, system configuration 210 may also categorize the primary data into different categories. In one
25 exemplary embodiment, system configuration 210 identifies whether the primary data is a set of temperature data, a set of electrical current data or a set of energy flow data. Categorization of the data may help to ease processing of the data for validation subsequently.

30 In one embodiment, data validation system 200 may be adaptable to different types of apparatuses and systems

through inputting in configuration information into system configuration 210. In one embodiment, system configuration 210 may be configured or reconfigured by inputting fixed set of parameters. It should be appreciated that each apparatus
5 may be pre-assigned with a fixed set of parameters that requires validation for the purposes of the energy management system. In one embodiment, the fixed set of parameters may be inputted through a computer that hosts the data validating system.

10 As the data validating system may be easily adaptable, system configuration 210 may also be configured from one setting to another setting on-the-fly (i.e., while data validating system is operational). For example, at a first instance, system configuration 210 may be configured to
15 receive data from a data center. In a subsequent instance, system configuration 210 may be configured to receive data from a manufacturing facility.

Therefore, a user merely needs to configure system configuration 210 portion of the code within data validation
20 system 200. As stated above, system configuration 210 may be easily portable across different settings by inputting different settings.

The primary data are then supplied into data controversy engine 220. Data controversy engine 220 may
25 compute statistics based on the primary data. The computed statistics may be performed at a local data level, in one embodiment. It should be appreciated that the local data level in here refers to an observation on a small scale (could be in terms of a relatively short physical locality or a
30 relatively short time instances). In one exemplary embodiment, observation at the small scale level may include observations

of data sets that are just received and prior to the newly received data sets. In another exemplary embodiment, observation at the small scale level may include observation of data sets from one or more apparatus that are located
5 relatively near to each other.

In one embodiment, data controversy engine 220 may identify outliers (i.e., a garbage data) that are amongst the primary data. For instance, a particular data may be identified as a garbage data when the data is outside of a
10 confidence level range. In one embodiment, the confidence level may be computed based on a statistical computation of the local data. The data may be deemed garbage or outlier only when it is significantly outside of the confidence level range.

As stated above, data controversy engine 220 may perform multiple statistical computations on the data. The statistical computations may include computation of mean, median and standard deviation of a particular set of primary data. For example, the statistical computations may be
20 performed on the power consumption data (in kilo watts per hour (kWh)). The statistical computation may help to identify controversial data (i.e., an outlier/garbage data) from the primary data. As stated the paragraph above, the data/set of data that are outside of a particular confidence level may be
25 deemed as an outlier/garbage and therefore subsequently be eliminated. Data validation engine 220 may protect the integrity of data from getting skewed by outliers or garbage data.

In addition to that, data controversy engine 220 may
30 also compute rate of data loss. The data may get lost due to loss of connectivity, shut down, etc. Therefore, in such

instance, the primary data that are received by system configuration 210 may be random and unwanted. Such data may skew the computation of the statistics that are based on these primary data at the local level.

5 In one embodiment, the rate of data loss may be determined by comparing the consistency of incoming data. For example, in a system, a break in a data feed for a short instance before the data feed resumed would indicate that there was a state where data has been lost. Such data loss
10 needs to be identified and deemed to be as garbage data.

 Alternatively, the data rate loss may also be determined when data controversy engine 220 compares the primary data with a set of secondary data. The secondary data may be supplied into the data controversy engine 220 by a user
15 using predefined and expected performance of the apparatus. In one exemplary embodiment, the secondary data may be information derived from a specification of an apparatus.

 Data controversy engine 220 may also be utilized for performing expansion of the received data. Data expansion may
20 be reproduction of the data in an original form for a piece of information that has undergone data compression. In one embodiment, the data expansion may be expansion in accordance to a time scale. For example, by knowing the start time, the end time and the total power consumption of an apparatus, data
25 controversy engine 220 may expand the received data to obtain average power used per minute, per hour, and/or etc.

 As shown in the embodiment of FIG. 2, the data from data controversy engine 220 is then supplied to data filter 230. Data filter 230 removes the outliers identified by the
30 data controversy engine 220 from the received data. Data

filter 230 portion of data validation system 200 merely removes the data that are outlier. In one embodiment, the data may be removed by overwriting the data to a constant fixed value. For example, a power consumption value, which is
5 generally a large positive value, may be overwritten to a -1. Hence, these set of data may not be further processed.

The data is then processed by data process and profiling engine 240. Data process and profiling engine 240 may determine confidence level of the data, occurrences of the
10 data and region of interest of the data. In one embodiment, data lost distribution may be obtained upon identification of the region of interest of the data.

In one embodiment, data process and profiling engine 240 may include unsupervised machine learning. It should be
15 appreciated that unsupervised machine learning may be performed by inferring a function to describe a hidden structure from unlabeled data. The unsupervised machine learning may include segmentation and extraction of frequency patterns. Based on these two things, non-frequency patterns
20 may be identified. Segmentation may be performed by identifying the data into segments. Frequency patterns may be identified based on the segmented area.

These data profiles may then be supplied to data validation engine 250 to further verify incoming data.
25 Furthermore, these profiles may also be viewed by a monitoring device (e.g., a screen of a computer). In one embodiment, the monitoring device may show daily gap analysis, pattern symbol chart, key performance index (KPI) tracing, and abnormal pattern summary. Furthermore, the monitoring device may also
30 show parameter trend chart, component trend chart, performance evaluation chart, and correlation analysis chart. It should

be appreciated that each of these data may be shown based on validated data from data validation engine and data filter.

In one embodiment, data validation engine 250 may validate the received data by comparing these data with domain knowledge of the apparatus and by comparing at a group or global level. The domain knowledge may include knowledge of physics that is applicable to the apparatus. In one exemplary embodiment, the physics (i.e., the domain knowledge) include thermodynamics behind energy transfer within the particular apparatus. The domain knowledge may be inputted inside the data validation engine 250 through a secondary data, in one embodiment. The domain knowledge may be a set of predefined formulas or a set of hash table that refers to predicted results for a particular set external conditions. For data validation at a global level, the data may be validated by a statistical computation referring to a global scale secondary data or domain knowledge.

Data validation engine 250 may also compare the data using a set of benchmark data. It should be appreciated that the benchmark data may be different than the data obtained from the domain knowledge as the benchmark data may be obtained from the previous operational information of the apparatus.

Data validation engine 250 may also validate the data using previous profiles that are available or received. In one embodiment, the previous profiles include a set of mean, medium and or standard deviation that helps one to compare with current set of received data. Once the received data passed the validation standard or requirements of the data validation engine 250, one or more results 260 will be output. In few embodiments, the results 260 may be subjected to

further data process or mining through other module associated with the energy management system 100.

It is important to note that the data validation engine 250 may be configured to assist the disclosed system 100 to determine outliers and/or errors, which are missed or omitted by the data controversy engine 220. These omitted outliers and/or errors may only be discovered when the data is compared to or computed against the global scale secondary data or domain knowledge. These outliers and/or errors render the data failing to meet the validation standard or requirements. The data failed to meet the validation will be rejected and/or reprocessed. According to several embodiments, the data validation engine 250 is configured to deliver the validation-failed data, with newly identified outliers and/or errors, back to the data filter. Again, the data filter 230 clears the known outliers and/or errors out from the received data and deliver the filtered data to the data process and profiling engine 240 in which the filtered data is reprocessed by one or more approaches mentioned in the foregoing. The reprocessed data is again transferred to the data validation engine to examine validity of the reprocessed data. The reprocessed data is expected to pass the validation process and output as the results 260.

FIG. 3, meant to be illustrative and not limiting, illustrates a flowchart of a method of validating data for an energy management system in accordance with one embodiment of the present invention. The data validation may be performed by data validation system 200 as shown in FIG. 2. In one embodiment, steps 310 to 350 may be performed using a computer.

At step 310, a configured data may be received by the data validation engine. The configured data may be in

accordance to a system in which the data validation system may be implemented. In one embodiment, the configured data may be data for a factory setting. Alternatively, the configured data may be for a data center setting. The data may include
5 data on power consumption, temperature, status of the apparatus in question, etc.

At step 320, controversial data is identified at a local level. As stated in the embodiment of FIG. 2, identification may be performed through a statistical
10 computation. In one embodiment, the statistical computation may include, amongst other, computation of mean, median and standard deviation of a received data on the local level. Furthermore, the statistical computation may also be performed based on a reference of the secondary data (e.g., computing
15 data correlation). In one embodiment, the secondary data may include a predefined data as an input based on a specific apparatus. Any data/set of data located outside of a particular confidence level may be deemed as an outlier/garbage and therefore subsequently be eliminated.

At step 330, the data may be filtered. In one
20 embodiment, the data that has been deemed outlier or garbage in the step 320 may be filtered from the set of data. The filtering may be performed using data filter portion of codes in the disclosed system 100. In one embodiment, the garbage
25 data may be invalidated by overwriting the garbage data.

At step 340, the data may then be processed and profiled. In one embodiment, the processing and profiling may be performed by data process and profiling 240 of FIG. 2. In one embodiment, processing and profiling may determine
30 confidence level of the data, occurrences of the data and region of interest of the data.

At step 350, the data is validated on a global level preferably before being output as results. In one embodiment, the data may be validated by a statistical computation. Furthermore, the validation may also be performed by comparing
5 the data with domain knowledge (e.g., physics that relates to an apparatus). In addition, the validation may also be performed by comparing the data with a benchmark data (e.g., data obtained from profiling the data in the past). Step 330 may be performed using a data validation engine (e.g., data
10 validation engine 230 of FIG. 2), in one embodiment.

FIG. 4, meant to be illustrative and not limiting, illustrates a computer system that implements the energy management system in accordance with one embodiment of the present invention. In one embodiment, computer 410 may be a
15 device that can be instructed to carry out an arbitrary set of arithmetic or logical operations automatically. Generally, computer 410 may follow a sequence of operations, called a program. Therefore, by loading different programs into a computer, computers become flexible and useful in many
20 situations. Computer 410 may include a storage disk and a monitor. The storage disk may be used to store a software program and the monitor may be utilized to display outputs after the software is being processed. In addition to that, computer 410 may also include an input device (e.g., keyboard)
25 to input user commands.

As shown in the embodiment of FIG 4, computer 410 implements a data validation system. The data validation system may be similar to data validation engine 200 as shown in FIG. 2. The data validation system may include multiple
30 portions of codes that perform plurality of statistical process on a local level, a global level. Furthermore, the

data validation system may also perform filtering of outlier/garbage data. In one embodiment, the data validation system may perform steps 310 - 350 of FIG. 3. Subsequently, the data validation engine may also display the validated data
5 through the monitor of computer 410. In one embodiment, the computer 410 may be used with different settings (e.g., data center, a factory, a commercial building, a residential unit) by inputting in different configuration through the input device of computer 410. The input may then setup
10 configuration system (e.g., configuration system 210 of FIG.2) for that particular type of setting.

Computer 410 may then display the validated data, and may process the validated data further if needed.

Also, although the present invention and its
15 advantages have been described in detail, it should be understood that various changes, substitutions, and alterations can be made herein without departing from the spirit and scope of the invention as defined by claims on embodiments. For example, many of the processes discussed
20 above can be implemented in different methodologies and replaced by other processes, or a combination thereof.

Moreover, the scope of the present application is not intended to be limited to the particular embodiments of the process, machine, manufacture, composition of matter,
25 means, methods, and steps described in the specification. As one of ordinary skill in the art will readily appreciate from the disclosure of the present invention, processes, machines, manufacture, compositions of matter, means, methods, or steps, presently existing or later to be developed, that perform
30 substantially the same function or achieve substantially the same result as the corresponding embodiments described herein

may be utilized according to the present invention. Accordingly, claims on embodiments are intended to include within their scope such processes, machines, manufacture, compositions of matter, means, methods, or steps.

Claims

1. A method validating data for an energy management system,
5 comprising:

receiving a primary data from a plurality of
apparatus available within the energy management system;

identifying outliers within the primary data based
on a statistical computation of the primary data;

10 filtering the outliers out from the primary data to
generate a filtered primary data;

processed the filtered primary data; and

validating the processed primary data based on
15 statistical computation at a global level of the primary data.

2. The method as defined in claim 1, wherein the primary
data includes a group of data consisting of: power, temperature,
flow and pressure.

20 3. The method as defined in claim 1, wherein the
identification of outliers based on the statistical
computation of the primary data further comprises:

computing a mean value of the primary data;

25 computing a standard deviation value of the primary
data; and

identifying the outlier when a piece of the primary
data is outside a range provided by the standard deviation
value from the mean value.

30 4. The method as defined in claim 1, wherein the
identification of outliers further comprises:

comparing the primary data and secondary data,
wherein the secondary data includes a range of data; and

identifying the outliers within the primary data
that are outside of a range provided by the secondary data.

5

5. The method as defined in claim 1, wherein the
identification of the outliers at the identifying step are
performed at a local data level.

10 6. The method as defined in claim 1, wherein the validating
the processed primary data based on the statistical
computation of the primary data at the global level further
comprises:

comparing the primary data with a domain knowledge.

15

7. The method as defined in claim 1, wherein the validating
the processed primary data based on the global level
statistical computation of the primary data further comprises:

comparing the primary data with a benchmarked

20 information.

8. The method as defined in claim 1, wherein the validating
the processed primary data based on the global level
statistical computation of the primary data further comprises:

25 comparing the primary data with an expected profile
of the primary data.

9. A system for validating data in an energy management
system, wherein the system comprises:

30 a personal computer executing a computer code to
perform a method for validating data in the energy management
system, wherein the computer code comprises:

a data controversy engine that receives the primary data, wherein the data controversy engine computes statistics for the primary data and identifies outliers within the primary data;

5 a data validation engine that receives the primary data from the data controversy engine, wherein the data validation engine validates the primary data by comparing to a predefined domain knowledge; and

10 a data processing and profiling engine that receives a filtered primary data, wherein the data processing and profiling engine extracts patterns from the filtered primary data.

15 10. The system as defined in claim 9, wherein the computer code further comprises:

 a configuration unit that is adapted receive primary data from a plurality of apparatuses within the energy management system.

20 11. The system as defined in claim 9, wherein the computer code further comprises:

25 a filtering unit that receives the primary data from the data validation engine and generates the filtered primary data after obtaining a determination of the outliers data from the data controversy engine and the data validation engine.

12. The system as defined in claim 9, wherein the data controversy engine receives a secondary data and compares the primary data with the secondary data to identify the outliers.

30

13. The system as defined in claim 9, wherein the data controversy engine identifies a data loss rate by computing a data consistency rate.

14. The system as defined in claim 9, wherein the data validation engine further validates the primary data by comparing primary data with a set of benchmark data.

5

15. The system as defined in claim 9, wherein the data validation engine further validates the primary data by comparing the primary data with data profiles.

10 16. A non-transitory medium storing computer-readable, computer-executable process steps for an energy management system, wherein the process comprising:

receiving primary data from a plurality of apparatuses formed within the energy management system;

15 identifying outliers within the primary data based on a statistical computation of the primary data;

;

filtering out the outlier from the primary data;

processing the primary data to a pattern of data;

20 and

validating the processed primary data based a statistical computation of the primary data at a global level.

17. The non-transitory medium storing as defined in claim 16, wherein the identification of the outliers based on the statistical computation of the primary data further comprises:

computing a mean value of the primary data;

computing a standard deviation value of the primary data; and

30 identifying the outlier when a piece of the primary data is located outside of a range that is provided by the standard deviation value from the mean value.

18. The non-transitory medium storing as defined in claim 16,
wherein the identification of outliers further comprises:

comparing the primary data with a secondary data
that includes a range of data; and

5 identifying the outliers within the primary data
that are outside of a range provided by the secondary data.