

Method and System for Determining a Condition of an Airflow Device

Field and Background

- 5 The invention relates to a method and system for determining a condition of an airflow device, in particular, but not exclusively for determining a condition of an Air Conditioning and Mechanical Ventilation (ACMV) system or a Heating, Ventilation, and Air Conditioning (HAVC) system comprising an air handling unit and/or a chiller unit.
- 10 Airflow units may be significant contributors to the energy consumption of buildings. According to building energy efficiency R&D roadmap of Singapore, commercial buildings consume around 31% of total electricity in Singapore. Cooling (60%) and ventilation (10%), together account for majority (i.e. 70%) of electricity consumption in commercial buildings in Singapore. Within Air Conditioning and Mechanical Ventilation
- 15 (ACMV) systems, the chiller unit accounts for majority of electricity consumption (55%), with the air handling unit (AHU) accounting for approximately 35% of electricity consumption (source: K. Chua, S. Chou, W. Yang and J. Yan, "*Achieving better energy-efficient air conditioning – A review of technologies and strategies*," Applied Energy, vol. 104, pp. 87-104, 2013.).
- 20 If a fault occurs in an airflow unit, power consumption of the airflow unit may increase exponentially, thereby reducing the overall efficiency of the system by as much as 50-80% in certain cases.
- 25 Moreover, equipment failures in the ACMV system may lead to additional electrical losses in the system (performance degradation) or complete system shut down, resulting in higher operational & maintenance (O&M) costs and a decrease in the stability of the system.
- 30 Off-line fault detection and diagnosis techniques have been proposed for airflow devices. However, these are applicable only when the system has stopped working due to fault, after which the off-line fault detection and diagnosis tests are performed to find the reason for the fault, for example, the faulty component. Some online fault detection and diagnosis techniques have been proposed for airflow devices based on mechanical
- 35 signatures of the system. However, typically these approaches can only detect a fault

when it reaches a particular severity level. Thus, at lower severity levels, the fault remains undetected which may result in efficiency reduction. Further, mechanical sensors may be costly and therefore unfeasible for the component level or application to a small system.

5

It is desirable to provide a method and system of determining a condition of an airflow device, for example, one comprising an air handling unit or a chiller unit, which addresses at least one of the drawbacks of the prior art and/or to provide the public with a useful choice.

10

Brief Description of Figures

Exemplary embodiments will now be described with reference to the accompanying drawings, in which:

15

Fig. 1 illustrates an airflow device according to a first embodiment, the airflow device including an AHU;

Fig. 2 illustrates a condition determination module included in the airflow device of Fig. 1;

20

Fig. 3 illustrates a method performed by a system for determining a condition of the airflow device of Fig. 1;

Fig. 4 illustrates a method of data analysis including time series analysis and frequency domain analysis performed on electrical signal data as part of the method of Fig. 3;

25

Fig. 5a schematically illustrates an artificial neural network (ANN), such as that employed as a filter during the method of Fig. 4;

Fig. 5b illustrates a method of training the ANN of Fig. 5a;

Fig. 6 schematically illustrates a decision tree, such as that used for selection of relevant Intrinsic Mode Functions as part of the method of Fig. 4;

30

Fig. 7 illustrates the method of Fig. 3 of determining the condition of the airflow device with a classification step of the method illustrated in block diagram form;

Fig. 8 schematically illustrates a mathematical model applied to determine the condition of the airflow device of Fig. 1;

Fig. 9 illustrates the data processing stages corresponding to the mathematical model of Fig. 8;

- Fig. 10 illustrates in detail the steps performed by a classification model employed in the classification step of the method of Fig. 3;
- Fig. 11a schematically illustrates a process performed by a Support Vector Machine, such as that employed during the classification step of the method of Fig. 3;
- 5 Fig. 11b-d show examples of AHU training data suitable for training the decision tree of Fig. 6 and the classification model of Fig. 10;
- Fig. 12 illustrates the Fault Impact Ratio for exemplary faults of an Air Handling Unit included in the airflow device of Fig. 1;
- Fig. 13a and 13b illustrate classification accuracy and relative time to classification, respectively, for simulated faults of an airflow device including an air handling unit with a mix of severity levels using the method according to Fig. 3 compared with results from an existing approach;
- 10 Fig. 14a and 14b illustrate classification accuracy and relative time to classification, respectively, for simulated faults of an airflow device including an air handling unit for a mix of fault conditions, using the method according to Fig. 3 compared with results from an existing approach;
- 15 Fig. 15a and 15b illustrate classification accuracy and relative time to classification, respectively, for simulated faults of an airflow device including an air handling unit with four types of fault conditions and four severity levels applied simultaneously during testing, using the method according to Fig. 3 compared with results from an existing approach;
- 20 Fig. 15c illustrates a confusion matrix for the testing data employed to obtain the results of Fig. 15a and 15b;
- Fig. 16 illustrates an airflow device according to a second embodiment, the airflow device including a chiller unit;
- 25 Fig. 17 illustrates a condition determination module included in the airflow device of Fig. 16;
- Fig. 18 illustrates a method performed by a system for determining a condition of the airflow device of Fig. 16;
- 30 Fig. 19 illustrates the method Fig. 17 of determining the condition of the airflow device with a classification step illustrated in block diagram form;
- Fig. 20 schematically illustrates a mathematical model applied to determine the condition of the airflow device of Fig. 16;
- Fig. 21 illustrates the data processing stages corresponding to the mathematical model of Fig. 20;
- 35

- Fig. 22a illustrates in detail the steps performed by a classification model employed in a classification step of the method of Fig. 17;
- Fig. 22b illustrates an example of training data suitable for training the decision tree of Fig. 6 and the classification model of Fig. 22a, specifically relating to refrigerant leakage;
- Fig. 22c and 22d illustrate the corresponding effect of refrigerant leakage on which Fig. 22b is based on the pressure of the refrigerant in the condenser and evaporator, respectively;
- Fig. 23 illustrates the Fault Impact Ratio for exemplary faults of an Air Handling Unit included in the airflow device of Fig. 16;
- Fig. 24a and 24b illustrate classification accuracy and relative time to classification, respectively, for simulated faults of an airflow device including a chiller unit with a mix of severity levels using the method according to Fig. 18 compared with results from three existing approaches;
- Fig. 25a and 25b illustrate classification accuracy and relative time to classification, respectively, for simulated faults of an airflow device including a chiller unit for a mix of fault conditions, using the method according to Fig. 18 compared with results from three existing approaches;
- Fig. 26a and 26b illustrate classification accuracy and relative time to classification, respectively, for simulated faults of an airflow device including a chiller unit with four types of fault conditions and four severity levels applied simultaneously during testing, using the method according to Fig. 18 compared with results from three existing approaches; and
- Fig. 26c illustrates a confusion matrix for the testing data employed to obtain the results of Fig. 26a and 26b.

Summary

- In a first aspect, there is provided a method of determining a condition of an airflow device. The method comprises: receiving electrical signal data representing an operating state of the airflow device; decomposing the electrical signal data into a plurality of first decomposed components of the electrical signal data representing corresponding intrinsic features of the electrical signal data; filtering each of the plurality of first decomposed components using a neural-network based filter to produce reconstructed electrical signal data of the electrical signal; decomposing the

- reconstructed electrical signal data into a plurality of second decomposed components of the reconstructed electrical signal data representing corresponding optimum intrinsic features, the optimum intrinsic features being more relevant to identifying the condition than the intrinsic features; identifying at least one most relevant intrinsic feature for
- 5 determining the condition of the airflow device from at least the optimum intrinsic features; and selecting at least one of the second decomposed components corresponding to the at least one most relevant intrinsic feature as an input to a condition classification model to determine the condition of the airflow device.
- 10 By filtering each of the plurality of first decomposed components using a neural-network based filter to produce a reconstructed electrical signal data of the electrical signal; decomposing a reconstructed electrical signal data into a plurality of second decomposed components of the reconstructed electrical signal data representing
- 15 corresponding optimum intrinsic features, the optimum intrinsic features being more relevant to identifying the condition than the intrinsic features; identifying at least one most relevant intrinsic feature for determining the condition of the airflow device from at least the optimum intrinsic features; and selecting at least one of the second decomposed components corresponding to the at least one most relevant intrinsic feature as an input to a condition classification model to determine the condition of the
- 20 airflow device, accurate and rapid determination of the condition of the airflow device may be possible, without disturbing the operation of the system.

In a specific embodiment, the at least one most relevant intrinsic feature may be identified using a decision tree.

- 25 Advantageously, the neural-network based filter may have a window size, the window size being dependent on an information entropy of the first decomposed component being filtered relative to an information entropy of the plurality of first decomposed components. It is envisaged that the window size may be further dependent on a
- 30 frequency of the first decomposed component being filtered. This may ensure optimal filtering of the decomposed components for retaining relevant information for determining the condition of the airflow device.

It is envisaged that decomposing the electrical signal data into a plurality of first

decomposed components may comprise decomposing the electrical signal data using a first empirical mode decomposition, and that decomposing the reconstructed electrical signal data into a plurality of second decomposed components may comprise decomposing the reconstructed electrical signal data using a second empirical mode
5 decomposition.

Preferably, the condition classification model may comprise a machine learning model described by at least first and second parameters, and the method may further comprise: training the machine learning model, wherein training the machine learning model
10 includes adjusting the first parameters of the machine learning model using a first optimization algorithm and adjusting the second parameters of the machine learning model using a second optimization algorithm based on a hydrological cycle (HC) algorithm. Training second parameters of the machine learning model using a second optimization algorithm based on a hydrological cycle (HC) algorithm may enable
15 improved robustness of the accuracy of condition determination under different environmental conditions.

Advantageously, the neural-network based filter may also be trained based on the hydrological cycle (HC) algorithm
20

In a specific embodiment, the airflow device may comprise an air handling unit, and the electrical signal data may be representative of an operating state of the air handling unit. As such, the method may equivalently be described as a method of determining the condition of the air handling unit. Further, the condition classification model may
25 comprise a Support Vector Machine (SVM).

In a further specific embodiment, the airflow device may comprise chiller unit, and the electrical signal data may be representative of an operating state of the chiller unit. As such, the method may equivalently be described as a method of determining the
30 condition of the chiller unit. Further, the condition classification model may comprise an Artificial Neural Network (ANN). Use of an ANN with a chiller unit may enable minimization of the computational burden for performing the method of determining the condition of the airflow device due to the relatively large number of variables available for a chiller unit.

35

In a second aspect, there is provided a method of determining a condition of an airflow device. The method comprises: receiving an electrical signal data representing an operating state of the airflow device; inputting the electrical signal data into a condition classification model to determine the condition of the airflow device, the condition
5 classification model comprising a machine learning model being described by at least first and second parameters; and training the machine learning model, wherein training the machine learning model includes adjusting the first parameters of the machine learning model using a first optimization algorithm and adjusting the second parameters of the machine learning model using a second optimization algorithm based on a
10 hydrological cycle (HC) algorithm.

In a third aspect, there is provided a system for determining a condition of an airflow device. The system comprises: an input for receiving electrical signal data representing an operating state of the airflow device; a processor configured to: decompose the
15 electrical signal data into a plurality of first decomposed components of the electrical signal data representing corresponding intrinsic features of the electrical signal data, filter each of the plurality of first decomposed components using a neural-network based filter to produce reconstructed electrical signal data of the electrical signal data, decompose the reconstructed electrical signal data into a plurality of second
20 decomposed components of the reconstructed electrical signal representing corresponding optimum intrinsic features, the optimum intrinsic features being more relevant to identifying the condition than the intrinsic features, identify at least one most relevant intrinsic feature for determining the condition of the airflow device from at least the optimum intrinsic features, and select at least one of the second decomposed
25 components corresponding to the at least one most relevant intrinsic feature as an input to a condition classification model to determine the condition of the airflow device; and an output configured to output data indicating the condition of the airflow device.

In a fourth aspect, there is provided a system for determining a condition of an airflow device, the system comprising: an input for receiving electrical signal data representing an operating state of the airflow device; and a processor configured to: input the electrical signal data into a condition classification model to determine the condition of the airflow device, the condition classification model comprising a machine learning model described by at least first parameters and second parameters, and train the machine
30 learning model, wherein training the machine learning model includes adjusting the first
35

parameters of the machine learning model using a first optimization algorithm and adjusting second parameters of the machine learning model using a second optimization algorithm based on a hydrological cycle (HC) algorithm; and an output configured to output data indicating the condition of the airflow device.

5

It is envisaged that the system may further comprise an electrical sensor configured to receive an electrical signal representing the operating state of the airflow device.

10 In a specific embodiment, the airflow device comprises an air handling unit and the electrical signal data is representative of an operating state of the air handling unit. In another specific embodiment the airflow device comprises a chiller unit and the electrical signal data is representative of an operating state of the chiller unit.

15 In a fifth aspect, there is provided an airflow device comprising a system for determining a condition of an airflow device. The system comprises: an input for receiving electrical signal data representing an operating state of the airflow device; a processor configured to: decompose the electrical signal data into a plurality of first decomposed components of the electrical signal data representing corresponding intrinsic features of the electrical signal data, filter each of the plurality of first decomposed components using a neural-
20 network based filter to produce reconstructed electrical signal data of the electrical signal data, decompose the reconstructed electrical signal data into a plurality of second decomposed components of the reconstructed electrical signal representing corresponding optimum intrinsic features, the optimum intrinsic features being more relevant to identifying the condition than the intrinsic features, identify at least one most
25 relevant intrinsic feature for determining the condition of the airflow device from at least the optimum intrinsic features, and select at least one of the second decomposed components corresponding to the at least one most relevant intrinsic feature as an input to a condition classification model to determine the condition of the airflow device; and an output configured to output data indicating the condition of the airflow device.

30

Preferably, the condition classification model may comprise a machine learning model described by at least first and second parameters, and the method may further comprise: training the machine learning model, wherein training the machine learning model includes adjusting the first parameters of the machine learning model using a first
35 optimization algorithm and adjusting the second parameters of the machine learning

model using a second optimization algorithm based on a hydrological cycle (HC) algorithm.

In a sixth aspect, there is provided a computer readable medium configured to cause a
5 computer to perform a method of determining a condition of an airflow device. The
method comprises: receiving electrical signal data representing an operating state of
the airflow device; decomposing the electrical signal data into a plurality of first
decomposed components of the electrical signal representing corresponding intrinsic
10 features of the electrical signal; filtering each of the plurality of first decomposed
components using a neural-network based filter to produce reconstructed electrical
signal data of the electrical signal; decomposing the reconstructed electrical signal data
into a plurality of second decomposed components of the reconstructed electrical
signal data representing corresponding optimum intrinsic features, the optimum
15 intrinsic features being more relevant to identifying the condition than the intrinsic
features; identifying at least one most relevant intrinsic feature for determining the
condition of the airflow device from at least the optimum intrinsic features; and
selecting at least one of the second decomposed components corresponding to the at
least one most relevant intrinsic feature as an input to a condition classification model
20 to determine the condition of the airflow device. The computer readable medium may
be transitory or non-transitory, tangible or intangible.

In a seventh aspect, there is provided a method of determining a condition of an airflow
device, comprising: receiving an electrical signal representing an operating state of the
airflow device; decomposing the electrical signal into a plurality of first decomposed
25 components of the electrical signal representing corresponding intrinsic features of the
electrical signal; filtering each of the plurality of first decomposed components using a
neural-network based filter to produce a reconstructed electrical signal of the electrical
signal; decomposing the reconstructed electrical signal into a plurality of second
decomposed components of the reconstructed electrical signal representing
30 corresponding optimum intrinsic features, the optimum intrinsic features being more
relevant to identifying the condition than the intrinsic features; identifying at least one
most relevant intrinsic feature for determining the condition of the airflow device from at
least the optimum intrinsic features; and selecting at least one of the second
decomposed components corresponding to the at least one most relevant intrinsic

feature as an input to a condition classification model to determine the condition of the airflow device.

In an eighth aspect, there is provided a method of determining a condition of an airflow device, comprising: receiving an electrical signal representing an operating state of the
5 airflow device; inputting the electrical signal into a condition classification model to determine the condition of the airflow device, the condition classification model comprising a machine learning model being described by at least first and second parameters; and training the machine learning model, wherein training the machine
10 learning model includes adjusting the first parameters of the machine learning model using a first optimization algorithm and adjusting the second parameters of the machine learning model using a second optimization algorithm based on a hydrological cycle (HC) algorithm.

15 It should be appreciated that features relevant to one aspect may also be relevant to other aspects.

Detailed Description of Embodiments

20 First Embodiment: Air Handling Unit (AHU)

Fig. 1 illustrates an airflow device 100, for example an Air Conditioning and Mechanical Ventilation (ACMV) or Heating, Ventilation, and Air Conditioning (HAVC) system,
according to a first embodiment. The airflow device 100 includes an air handling unit
25 (AHU) 101 and a system for determining a condition of the airflow device 100, specifically via the determination of a condition of the AHU 101. The airflow device 100 further includes electrical wires 107 which connect at least the AHU 101 to an external electrical power source 105, for example an electrical distribution board.

30 The AHU 101 is configured to extract air from the environment, pass the air over a heating or cooling element and return the heated or cooled air to the environment. As such the AHU 101 may comprise further devices for enabling specific aspects of the function of the AHU 101, including but not limited to a heating and/or a cooling element (for example a heating and/or cooling coil), a supply air duct, a return air duct, a fan,
35 and an economizer. The AHU 101 may also comprise a controller having the same or a

similar structure to that described in relation to Fig. 2 for controlling specific functions of the AHU 101.

The system for determining the condition of the airflow device 100 includes a condition
5 determination module 380 in the form of a computer system as well as an input to the
condition determination module 380 in the form of an electrical sensor 103, for example
one or more clamp meters hung on one or more of the electrical wires 107, for
collecting electrical signal data representing an operating state of the AHU 101, and, as
such, the airflow device 100 as a whole. Specifically, the electrical sensor 103 is
10 arranged to measure an electrical signal in the form of a current and/or power input to
the AHU 101, for example a cooling coil of the AHU.

For example, in a DC system only a single clamp meter may be employed as power may
be calculated from the current detected using a single clamp meter for a constant DC
15 load. In an AC system in which input power = Input Current × Input Voltage × Angle
(Power factor) with the angle driving the input active power, input reactive power and
input nonlinear power, according the requirements of the device. An additional clamp
meter -i.e. two clamp meters - may be employed in order to derive current and power
information.

20

It will be appreciated that the airflow device 100 may further include other systems not
shown in Fig. 1, including, but not limited to a chiller unit, as will be described in relation
to the second embodiment below.

25 Fig. 2 illustrates the condition determination module 380 in more detail.

The condition determination module 380 includes a processor 382 (which may be
referred to as a central processor unit or CPU) that is in communication with memory
devices including secondary storage 384, read only memory (ROM) 386, random
30 access memory (RAM) 388, input/output (I/O) devices 390, and network connectivity
devices 392. The processor 382 may be implemented as one or more CPU chips.

It is understood that by programming and/or loading executable instructions onto the
condition determination module 380, at least one of the CPU 382, the RAM 388, and
35 the ROM 386 are changed, transforming the condition determination module 380 in

part into a particular machine or apparatus having the novel functionality taught by the present disclosure. It is fundamental to the electrical engineering and software engineering arts that functionality that can be implemented by loading executable software into a computer can be converted to a hardware implementation by well-

5 known design rules. Decisions between implementing a concept in software versus hardware typically hinge on considerations of stability of the design and numbers of units to be produced rather than any issues involved in translating from the software domain to the hardware domain. Generally, a design that is still subject to frequent change may be preferred to be implemented in software, because re-spinning a

10 hardware implementation is more expensive than re-spinning a software design. Generally, a design that is stable that will be produced in large volume may be preferred to be implemented in hardware, for example in an application specific integrated circuit (ASIC), because for large production runs the hardware implementation may be less expensive than the software implementation. Often a

15 design may be developed and tested in a software form and later transformed, by well-known design rules, to an equivalent hardware implementation in an application specific integrated circuit that hardwires the instructions of the software. In the same manner as a machine controlled by a new ASIC is a particular machine or apparatus, likewise a computer that has been programmed and/or loaded with executable

20 instructions may be viewed as a particular machine or apparatus.

Additionally, after the condition determination module 380 is turned on or booted, the CPU 382 may execute a computer program or application. For example, the CPU 382 may execute software or firmware stored in the ROM 386 or stored in the RAM 388. In

25 some cases, on boot and/or when the application is initiated, the CPU 382 may copy the application or portions of the application from the secondary storage 384 to the RAM 388 or to memory space within the CPU 382 itself, and the CPU 382 may then execute instructions that the application is comprised of. In some cases, the CPU 382 may copy the application or portions of the application from memory accessed via the

30 network connectivity devices 392 or via the I/O devices 390 to the RAM 388 or to memory space within the CPU 382, and the CPU 382 may then execute instructions that the application is comprised of. During execution, an application may load instructions into the CPU 382, for example load some of the instructions of the application into a cache of the CPU 382. In some contexts, an application that is

35 executed may be said to configure the CPU 382 to do something, e.g., to configure the

CPU 382 to perform the function or functions promoted by the subject application. When the CPU 382 is configured in this way by the application, the CPU 382 becomes a specific purpose computer or a specific purpose machine.

- 5 The secondary storage 384 is typically comprised of one or more disk drives or tape drives and is used for non-volatile storage of data and as an over-flow data storage device if RAM 388 is not large enough to hold all working data. Secondary storage 384 may be used to store programs which are loaded into RAM 388 when such programs are selected for execution. The ROM 386 is used to store instructions and perhaps
- 10 data which are read during program execution. ROM 386 is a non-volatile memory device which typically has a small memory capacity relative to the larger memory capacity of secondary storage 384. The RAM 388 is used to store volatile data and perhaps to store instructions. Access to both ROM 386 and RAM 388 is typically faster than to secondary storage 384. The secondary storage 384, the RAM 388, and/or the
- 15 ROM 386 may be referred to in some contexts as computer readable storage media and/or non-transitory computer readable media.

- I/O devices 390 include a connection to the electrical sensor 103 and may include video monitors, liquid crystal displays (LCDs), plasma displays, touch screen displays,
- 20 keyboards, keypads, switches, dials, mice, track balls, voice recognizers, card readers, paper tape readers, or other well-known input or output devices.

- The network connectivity devices 392 may take the form of modems, modem banks, Ethernet cards, universal serial bus (USB) interface cards, serial interfaces, token ring
- 25 cards, fibre distributed data interface (FDDI) cards, wireless local area network (WLAN) cards, radio transceiver cards that promote radio communications using protocols such as code division multiple access (CDMA), global system for mobile communications (GSM), long-term evolution (LTE), worldwide interoperability for microwave access (WiMAX), near field communications (NFC), radio frequency identity (RFID), and/or
- 30 other air interface protocol radio transceiver cards, and other well-known network devices. These network connectivity devices 392 may enable the processor 382 to communicate with the Internet or one or more intranets. With such a network connection, it is contemplated that the processor 382 might receive information from the network, or might output information to the network in the course of performing the
- 35 below-described method steps. Such information, which is often represented as a

sequence of instructions to be executed using processor 382, may be received from and outputted to the network, for example, in the form of a computer data signal embodied in a carrier wave.

- 5 Such information, which may include data or instructions to be executed using processor 382 for example, may be received from and outputted to the network, for example, in the form of a computer data baseband signal or signal embodied in a carrier wave. The baseband signal or signal embedded in the carrier wave, or other types of signals currently used or hereafter developed, may be generated according to
10 several methods well-known to one skilled in the art. The baseband signal and/or signal embedded in the carrier wave may be referred to in some contexts as a transitory signal.

The processor 382 executes instructions, codes, computer programs, scripts which it
15 accesses from hard disk, floppy disk, optical disk (these various disk based systems may all be considered secondary storage 384), flash drive, ROM 386, RAM 388, or the network connectivity devices 392. While only one processor 382 is shown, multiple processors may be present. Thus, while instructions may be discussed as executed by a processor, the instructions may be executed simultaneously, serially, or otherwise
20 executed by one or multiple processors. Instructions, codes, computer programs, scripts, and/or data that may be accessed from the secondary storage 384, for example, hard drives, floppy disks, optical disks, and/or other device, the ROM 386, and/or the RAM 388 may be referred to in some contexts as non-transitory instructions and/or non-transitory information.

- 25 In an embodiment, the condition determination module 380 may comprise two or more computers in communication with each other that collaborate to perform a task. For example, but not by way of limitation, an application may be partitioned in such a way as to permit concurrent and/or parallel processing of the instructions of the application.
30 Alternatively, the data processed by the application may be partitioned in such a way as to permit concurrent and/or parallel processing of different portions of a data set by the two or more computers. In an embodiment, virtualization software may be employed by the condition determination module 380 to provide the functionality of a number of servers that is not directly bound to the number of computers in the
35 condition determination module 380. For example, virtualization software may provide

twenty virtual servers on four physical computers. In an embodiment, the functionality disclosed herein may be provided by executing the application and/or applications in a cloud computing environment. Cloud computing may comprise providing computing services via a network connection using dynamically scalable computing resources.

- 5 Cloud computing may be supported, at least in part, by virtualization software. A cloud computing environment may be established by an enterprise and/or may be hired on an as-needed basis from a third party provider. Some cloud computing environments may comprise cloud computing resources owned and operated by the enterprise as well as cloud computing resources hired and/or leased from a third party provider.

10

In an embodiment, some or all of the functionality disclosed herein may be provided as a computer program product. The computer program product may comprise one or more computer readable storage medium having computer usable program code embodied therein to implement the functionality disclosed herein. The computer

15 program product may comprise data structures, executable instructions, and other computer usable program code. The computer program product may be embodied in removable computer storage media and/or non-removable computer storage media.

- The removable computer readable storage medium may comprise, without limitation, a paper tape, a magnetic tape, magnetic disk, an optical disk, a solid state memory chip,
- 20 for example analogue magnetic tape, compact disk read only memory (CD-ROM) disks, floppy disks, jump drives, digital cards, multimedia cards, and others. The computer program product may be suitable for loading, by the condition determination module 380, at least portions of the contents of the computer program product to the secondary storage 384, to the ROM 386, to the RAM 388, and/or to other non-volatile
- 25 memory and volatile memory of the condition determination module 380. The processor 382 may process the executable instructions and/or data structures in part by directly accessing the computer program product, for example by reading from a CD-ROM disk inserted into a disk drive peripheral of the condition determination module 380. Alternatively, the processor 382 may process the executable instructions
- 30 and/or data structures by remotely accessing the computer program product, for example by downloading the executable instructions and/or data structures from a remote server through the network connectivity devices 392. The computer program product may comprise instructions that promote the loading and/or copying of data, data structures, files, and/or executable instructions to the secondary storage 384, to

the ROM 386, to the RAM 388, and/or to other non-volatile memory and volatile memory of the condition determination module 380.

5 In some contexts, the secondary storage 384, the ROM 386, and the RAM 388 may be referred to as a non-transitory computer readable medium or a computer readable storage media. A dynamic RAM embodiment of the RAM 388, likewise, may be referred to as a non-transitory computer readable medium in that while the dynamic RAM receives electrical power and is operated in accordance with its design, for example during a period of time during which the condition determination module 380
10 is turned on and operational, the dynamic RAM stores information that is written to it. Similarly, the processor 382 may comprise an internal RAM, an internal ROM, a cache memory, and/or other internal non-transitory storage blocks, sections, or components that may be referred to in some contexts as non-transitory computer readable media or computer readable storage media.

15

Fig. 3 illustrates a method performed by the system for determining a condition of the airflow device 100, which is performed at least in part by executing software on the processor 382.

20 In step 301, data collection is performed. Specifically, as noted above the electrical sensor 103 receives an electrical signal to the AHU 101 and measures one or more features of the electrical signal, for example current and/or power. The condition determination module 380 receives electrical signal data obtained in this way from the electrical sensor 103 for processing at the CPU 382.

25

In step 303, software is executed on the CPU 382 to perform data cleaning of the electrical signal data. For example, digital filters may be applied to reduce noise in the signal. It should be appreciated that other techniques of data cleaning may also be applied.

30

In step 305 software is executed on the CPU 382 to perform data analysis including time series analysis and frequency domain analysis on the electrical signal data. The individual steps of this process are illustrated in more detail in Fig. 4.

In step 201, the electrical signal data $X(t)$ is received. In the described embodiment, this is the cleaned electrical signal data obtained in step 303.

5 In step 203, the cleaned electrical signal data $X(t)$ is decomposed to obtain, in step 205, a plurality of first decomposed components of the electrical signal representing intrinsic features of the electrical signal. In this embodiment, the decomposition is performed by empirical mode decomposition (EMD) and the first decomposed components are in the form of intrinsic mode functions (IMF1, IMF2... IMF n). Open source packages for performing EMD are available.

10

After generating n -number of IMFs through EMD in steps 203 and 205, each IMF is processed through an adaptive neural network (ANN)-based filter to remove noise and to make the signal smooth.

15 Fig. 5a schematically illustrates an example ANN 19. ANNs are adaptive models trained by machine learning methods. In general, they comprise sets of algorithms configured to map inputs to outputs, which, in the described embodiment comprise the generated IMFs and filtered IMFs, respectively. The exemplary ANN 19 of Fig. 5a comprises an input layer 1901 where an IMF is input into the network, one or more hidden layers 1903 where
20 inputs are combined and an output layer 1905 at which the filtered output is received.

Note that the example ANN 19 is the simplest example of ANN and a much more complex ANN may be employed in practice; although only one hidden layer is shown in Fig. 5a, the ANN 19 may comprise a plurality of hidden layers, according to the
25 architecture employed.

The hidden layer 1903 comprises a series of biased nodes 1909. Each input to each hidden layer is weighted and combined at a node with a non-linear activation function.

30 The ANN 19 is defined by a series of parameters including those characterizing the architecture of the neural network (i.e. number of nodes and number of hidden layers), activation functions, weights and biases. The weights and biases are determined during training of the ANN 19.

An overview of the training process of the ANN 19 according to the described embodiment is shown in Fig. 5b.

In step 1109, the ANN is initialized. In practice this means that the values of the hyperparameters (the constant parameters defining the network, such as dimensions of the hidden representation, number of layers, etc.) are selected and all of the trainable parameters characterizing the model are given an initial value, for example, a randomly chosen value. Where necessary, the activation functions are also chosen. Preferably, the hyperparameters are determined heuristically, for example via ablation experiments, according to the limitations of the systems employed and the desires of the user, for example, the memory, speed constraints, desired throughput and desired accuracy.

In step 1110 an input 1001 with a known, expected output 1003 is input into the neural network with the initialized parameters. The input 1001 is processed and an output is estimated by the neural network.

In step 1111, the output produced by the neural network is compared to the target output 1003 and an error is calculated. These respective inputs 1001 and outputs 1003 comprise the training data for the ANN. For the ANN filter, therefore, a training input 1001 includes the input IMF and the target outputs 1003 include the filtered data corresponding to the input 1001.

In step 1113, the parameters of the neural network are adjusted in order to minimize the error. The process is then repeated with the adjusted neural network by returning to step 1110 and minimizing the neural network error for other items of training data.

In the described embodiment, in step 1113, the weights and biases of the ANN 19 are optimized by the HC (Hydrological Cycle) algorithm to make the filter adaptive in nature so that the error signal can be minimized.

The hydrological cycle (HC) algorithm is described in Ahmad Wedyan, Jacqueline Whalley, and Ajit Narayanan, "*Hydrological Cycle Algorithm for Continuous Optimization Problems*", Journal of Optimization, vol. 2017, Article ID 3828420, 25 pages, 2017, <https://doi.org/10.1155/2017/3828420>.

The HC algorithm, is a metaheuristic optimisation technique, whose target is to fulfil the objective of the task (here this is minimization ANN estimation error) through adjusting the control parameters (i.e. the weights and biases of the ANN technique in the described embodiment). To do this, HC initially generates a random set of weights and biases, and

5 checks the error for all weights and biases. In next step, according to the performance of previous step, all weights and biases are updated to improve the performance. Then, an error check is again performed for all updated weights and biases, and an improvement progress (performance) is determined. In next step, again the weights and biases are updated according to the performance determined in the previous step. As

10 such, HC monitors and improves the performance of the ANN.

In an example, the training dataset for training the ANN-based filter includes electrical signals with high sampling frequency, for example about 1kHz to about 2kHz.

15 Returning now to Fig. 4, implementation of the filter for each IMF is dependent on the order of the frequency (high or low frequency) of the IMF.

Specifically, in step 207, a frequency domain representation of each of the decomposed components, IMF1, IMF2... IMF_L, is obtained using a Fourier transform and the average

20 frequency of the component is determined.

If the average frequency f_{IMFi} of a component i is below a threshold frequency

$\frac{f_{IMF1}}{2} + f_{IMFL}$, where f_{IMF1} is the average frequency of the first intrinsic mode function

IMF1, i.e. the IMF with the most oscillating (high-frequency) components and f_{IMFL} is

25 the average frequency of the Lth intrinsic mode function, where the total number of

IMFs is L, a low frequency window size is calculated. Specifically, if $f_{IMFi} \leq \frac{f_{IMF1}}{2} + f_{IMFL}$

a window size of

$lf = \text{round}\left(\frac{i}{L} * (e(IMF_i))^2\right)$ is calculated, where, $e(IMF_i)$ is the Shannon entropy of

the IMF_i.

30 .

If the average frequency is above the threshold frequency, i.e. $f_{BF_i} > \frac{f_{BF1} + f_{BF2}}{2}$, then a high frequency window size for the i th IMF is calculated as

$$hf = \text{round}\left(\frac{i}{L} * e(IMF_i)\right)$$

- 5 In steps 209 and 211 the relevant decomposed component IMF is filtered using the ANN-based filter, with the determined window size lf or hf as described above, respectively. First, the Shannon energy entropy for each IMF is calculated, evaluated as:

$$E_e = -\sum_{i=1}^k \rho(i) \log(\rho(i))$$

- 10 where k is the number of nodes of ANN.

Here, $\rho(i) \left(= \frac{E_i}{E} \right)$ is the percentage value of the energy of the i^{th} IMF with respect to the

whole signal energy (E , where, $E = \sum_{i=1}^L E_i$).

- 15 Next, the ANN then filters the IMFs based on the determined window size and the Shannon energy entropy.

The window size of the ANN-based frequency filter is therefore adjusted dependent on the average frequency of the relevant decomposed component being filtered. Further, The ANN-based filtering is based on the information entropy of each IMF.

20

In step 213, once all generated IMFs are processed through the ANN based filter, the outputs are summed together to generate, or reconstruct, the filtered version of the original data signal ($X_{filter}(t)$). In other words, reconstructed electric signal data is obtained.

25

In step 215 the reconstructed signal is itself decomposed to obtain, in step 217, a plurality of second decomposed components (IMFi1, IMFi2... IMFin) of the electrical signal. In this embodiment, the decomposition is again performed by empirical mode decomposition (EMD) and the decomposed components are in the form of intrinsic mode

functions, i.e. at this point the EMD process is again applied to the filtered version of the original data and new IMFs are generated.

By filtering the original intrinsic mode functions as described in step 207-211 above and
5 reconstructing the signal in step 213, the second decomposed components are
"improved" intrinsic mode functions (IMFi1, IMFi2... IMFin), that is, decomposed
components that represent optimum intrinsic features of the original electrical signal that
are more relevant to identifying the health condition of the airflow device than the first
decomposed components, IMF1, IMF2... IMF_n obtained in step 205. The improved IMFs
10 differ from the original IMFs as each original IMF is processed through the ANN based
filter to remove noise and to make the signal smooth as possible.

In step 307 (shown in both Fig. 3 and Fig 4) at least one of the improved intrinsic mode
functions (IMFi1, IMFi2... IMFin) obtained in step 217 is selected. This is done by
15 selecting at least one improved intrinsic mode function or functions that represent one or
more intrinsic features that are most relevant for determining the condition of the AHU
101.

The most relevant feature (i.e. the IMF for use in the classification model of step 309)
20 selection is performed by inputting the IMFs and IMFis into a decision tree constructed,
for example, using the J48 algorithm.

Fig. 6 schematically illustrates a simplified example of decision tree 6001. The decision
tree 6001 consist of a series of nodes 6011 at each of which the input data is split into
25 two branches 6003. The criteria by which the data is divided at each node 6011 is
determined during training. In use, input data is subject to the splitting criteria at each
node 6011 sequentially until a terminal node 6005 is reached which gives the predicted
output value for the given input. In the described embodiment, each IMFi is input into the
decision tree and the decision tree outputs a measure of the relevancy of an input IMFi
30 for classifying the condition of the airflow device 100.

Note that the decision tree 6001 is a simplified example and that decision trees will
typically have several layers of nodes.

Thus, the decision tree predicts a relevancy of a given IMF_i for classifying the condition of the airflow device 100. In practice, the decision tree generates a relevancy score for the given IMF_i. Once all of the IMF_is have been processed through the decision tree, the IMF_is with a relevancy score higher than a particular threshold are selected as being most relevant for into the classification model.

In the described embodiment, the J48 algorithm is employed to generate the decision tree for determining the most relevant IMF_i. In general, the J48 algorithm proceeds as follows:

10

First, an input training matrix H is prepared where:

$$H = [imf_1, imf_2, imf_3, \dots, imf_n]$$

The input matrix H is represented as sample vectors $x_i \in R^n$, $i = 1, 2, \dots, I$ with labels $y \in R^l$, according to the number of features, and a split space is created with the samples with similar labels grouped together as follows:

15

Designating the data at node m in the decision tree by β , a candidate division in the data is defined as $\lambda = (j, t_m)$ having a feature j and a threshold t_m . The data is divided into $\beta_{left}(\lambda)$ and $\beta_{right}(\lambda)$ sub-groups as:

20

$$\beta_{left}(\lambda) = (x, y) | x_j \leq t_m \text{ and } \beta_{right}(\lambda) = \beta \setminus \beta_{left}(\lambda)$$

This candidate division is evaluated using an impurity function $H()$, a data set of different patterns or different signatures (related to the extracted features), the exact form of which varies according to performed task (such as classification/ regression).

25

The following equation is then obtained

$$G(\beta, \lambda) = \frac{\eta_{left}}{N_m} H(\beta_{left}(\lambda)) + \frac{\eta_{right}}{N_m} H(\beta_{right}(\lambda))$$

30

where N_m is the number of samples in the node m and η_{left} and η_{right} are the number of samples in each split at the division.

In the described embodiment, with K possible classification outcomes $0, 1, \dots, K-1$, for node m , and denoting a region of the decision tree as R_m with N_m samples as defined above then a proportion of classification observations k at node m is $p_{mk} = 1/N_m \sum_{x_i \in R_m} I(y_i = k)$ where $I(y_i = k)$ is a counting function. Example impurity functions include:

10 General evaluation of impurity (i.e., Gini): $H(X_m) = \sum_k p_{mk}(1 - p_{mk})$;

Cross-entropy: $H(X_m) = -\sum_k p_{mk} \log(p_{mk})$; and

Misclassification: $H(X_m) = 1 - \max(p_{mk})$

In particular, Misclassification is employed as the impurity function.

15

The parameters are then optimized to minimize the impurity as:

$$\lambda^* = \arg \min_{\theta} G(\beta, \lambda)$$

20 The above process is iterated again-and-again for $\beta_{left}(\lambda^*)$ and $\beta_{right}(\lambda^*)$ until $N_m < \min_{samples}$ or $N_m = 1$.

Note that the J48 algorithm is an implementation of the C4.5 algorithm which is further described in Quinlan, J.R.: 'C4.5: programs for machine learning'. Morgan Kaufmann
25 Publishers Inc., San Francisco, CA, USA., 1993. ISBN: 978-1-55860-238-0 and Quinlan, J.R., 'Improved use of continuous attributes in C4.5', J. Artif. Res., 1996, 4, pp. 77–90, Available at <https://www.jair.org/index.php/jair/article/view/10157/24078>.

Thus, the decision tree is trained based on improved IMFs obtained in step 217 (of Fig
30 4). It will generate the rank value for each IMFs and based on rank value, high ranked

IMFs will be selected for classification/diagnosis in step 309. The J48 algorithm may be advantageous because it will select only high ranked variables directly in form of a node of the decision tree.

- 5 Optionally, for an additional high level of confidence, the original IMFs obtained in step 205 are also input into the decision tree, as indicated by arrow 21, to check that the relevancy scores for the improved IMFis are higher than for the original IMFs.

10 In step 309, the one or more IMFis determined as being the most relevant for determining the condition of the AHU 101 are input into a condition classification model to determine the condition of the AHU 101. In the described embodiment, the condition classification model is in the form of a Support Vector Machine (SVM) 701.

15 Fig. 7 illustrates the method of determining the condition of the airflow device 100 of Fig. 3 with step 309 illustrated in block diagram form. Fig. 8 schematically illustrates a mathematical model applied to determine the condition of the airflow device 100 and Fig. 9 illustrates the data processing stages corresponding to the mathematical model of Fig. 8. The IMFis output selected in step 307 are input into the SVM 701 as selected features data $f(d)$ 705. Features library data, $f(l)$ 703 is also input into the SVM 701 as training data for the SVM 701. Both, $f(d)$ 705 and $f(l)$ 703 are processed through Support Vector Machine (SVM) based analysis, which outputs $f'(d)$ 707 via one of the I/O devices of the condition determination module 380 which indicates a predicted one of five health states S_1, S_2, S_3, S_4, S_5 901 of the AHU 101. Specifically, as will be appreciated from Fig. 9, the SVM 701 outputs a predicted state 903 for the AHU 101 in the form of one of the five health states S_1, S_2, S_3, S_4, S_5 901 in order to determine the output $f'(d)$ 707.

20 $f(d)$ 705 and $f(l)$ 703 are processed through Support Vector Machine (SVM) based analysis, which outputs $f'(d)$ 707 via one of the I/O devices of the condition determination module 380 which indicates a predicted one of five health states S_1, S_2, S_3, S_4, S_5 901 of the AHU 101. Specifically, as will be appreciated from Fig. 9, the SVM 701 outputs a predicted state 903 for the AHU 101 in the form of one of the five health states S_1, S_2, S_3, S_4, S_5 901 in order to determine the output $f'(d)$ 707.

25 health states S_1, S_2, S_3, S_4, S_5 901 in order to determine the output $f'(d)$ 707.

In the described embodiment, the SVM 701 is a Support Vector Machine that is guided by the Hydrological Cycle (HC) Algorithm in an approach which is referred to below as HC-SVM.

30

The steps of the condition classification model are shown in detail in Fig. 10.

In step 601, electrical signal data of the AHU 101 is received and cleaned, i.e. steps 301 and 303 as described above are performed.

35

In step 603, the input data is pre-processed i.e., steps 305 and 307 as described above are performed and the most relevant IMFis determined.

5 In steps 605 to 615 the model is trained. Support Vector Machines (SVMs) are supervised machine learning algorithms that transform data in order to find boundaries between the data which enable data to be classified into different outputs. This is shown schematically in Fig. 11a in which a boundary 2401 has been defined within a set of data points 2403 to classify data points according to whether they fall on one side 2405 of the boundary or the other side 2407 of the boundary.

10

SVMs are typically defined by both a kernel function, superparameters and hyperparameters. Kernel functions are employed to generate a pattern and classify non-linear data, through projecting onto a higher dimension space. The superparameters of the SVM are constants, which are adjusted during training. The hyperparameters of the SVM employ the kernel function to map the observations into a feature space. In an example, separate hyperplanes may be employed along with different Kernel functions.

15

Returning now to Fig. 10, in step 605, a radial basis kernel function for the SVM is selected and in step 606 and initial values of the SVM super parameters are selected. The selected kernel function and super parameter initial values are employed to generate an initial SVM in step 607.

20

It should be appreciated that the selection of the kernel function and the initial values of the super parameter will differ depending on the nature of the data step at step 603. For example, the data set distribution may be linear or nonlinear and may be handled differently by employing different selected functions and parameters according to its form. In the case of linear data set, for example, a linear splines kernel may be employed, whereas in the case of a nonlinear data set, a Gaussian kernel may be employed instead.

25

30 In step 609, library features $f(l)$ 703 are retrieved for use as input and output parameters for training the superparameters of the SVM model. Each library feature $f(l)$ 703 includes an input parameter consisting of one or more improved intrinsic mode functions obtained for the AHU 101, as described above, in other words each library feature $f(l)$ 703 is a representation of data which may be selected during the feature selection 307 for use as an input data to the SVM 701 for classification and an output parameter consisting of a

35

known health condition of the AHU 101 at the time the improved intrinsic mode function or functions was or were obtained. The improved intrinsic mode function corresponding to each library feature is input into the model and the output of the model is compared with the output parameter for the relevant library feature to calculate a fitting error in step 611. In step 613, the calculated error is compared with a threshold. If the error exceeds the threshold, then in step 615, the superparameters of the SVM are adjusted and the process returns to step 609. The error function is based on the mapping of the predicted value by the model with respect to the target value of the class. Preferably, it is at low as possible, for example, about 1% or lower may be employed as an acceptable error threshold level. If the error is not within the acceptable threshold then the super parameters will be adjusted.

Once the error is lower than the threshold value, the method proceeds to step 617 in which it is determined if stopping criteria of the model are satisfied. For example, the stopping criteria may be that the fitting error calculated in 611 is below a particular threshold, such as less than 0.01%.

If the stopping criteria are found not to be met in step 617, then the method proceeds to the hydrological cycle (HC) algorithm of steps 621 to 629 and the hyperparameters of the SVM model are adjusted. The hydrological cycle (HC) algorithm is an optimization algorithm, which adjusts the algorithm parameters according to the error, inspired by the stages of the real-life hydrological cycle undergone by water in the environment. As discussed above, the hydrological cycle (HC) algorithm is described in Ahmad Wedyan, Jacqueline Whalley, and Ajit Narayanan, "*Hydrological Cycle Algorithm for Continuous Optimization Problems*", Journal of Optimization, vol. 2017, Article ID 3828420, 25 pages, 2017, <https://doi.org/10.1155/2017/3828420>. The main concept behind the HC algorithm is that of analysing the performance of all variables of a previous iteration, and updating the variables for next iteration, after which the performance of all variables is again analysed and the variables again updated, with each of the steps of this process being inspired by a corresponding step in the real-life hydrological cycle

Steps 621 to 629 represent each stage of the hydrological algorithm described in terms of their equivalent real-life hydrological cycle counterparts.

Specifically, in step 621 the velocity is calculated according to the hyperparameter performance.

In step 623, the soil and depth are updated.

5

In step 625 the temperature, condensation and precipitation are updated.

In step 627 the water drops are updated.

10 In step 629 the hyperparameters of the SVM are updated.

Note that HC algorithm (steps 621-629) is a meta-heuristic optimization technique, where specific training data for the HC is not required. A random set of data is therefore employed for initialization of the hyperparameters and then updated in each iteration.

15 Finally, the algorithm converges on an optimal result.

Once the adjusted hyperparameters of the SVM are determined in step 629, the method returns to step 607 and an initial model with hyperparameters determined in step 629 is generated. The superparameters of the model are then trained in steps 609 to 615, based on the adjusted hyperparameters as before.

20

Thus, in this first embodiment the training of the condition classification model includes adjusting a first set of parameters of the model using a superparameter optimization algorithm and adjusting a second set of parameters, in the form of the hyperparameters of the SVM model, using a second optimization algorithm based on the hydrological cycle algorithm.

25

Once it is found in step 617 that the stopping criteria are met, then the method proceeds to step 619. Specifically, in step 619 the trained SVM is applied to the selected features data $f(d)$ 705 comprising the selected IMFs and/or IMFis and a predicted state 903 is output from the SVM. In practice, the predicted state 903 is one of the five classification states 901 and the estimated condition of the AHU 101 is then output, for example it is displayed by a display forming part of the I/O devices 390 of the condition determination module 380.

30
35

Specifically, the trained model is applied to the and the data input in step 601 is input into the SVM as trained above in steps 605 to 615 and an estimated health condition of the AHU 101, and therefore the condition of the airflow device 100 is output by the model.

- 5 In an example, the five classification states 901 include a first category (S1) indicating normal operation of the AHU 101, and the remaining four categories (S2 - S5) indicate four different types of fault conditions. Specifically, S2 indicates "Economizer opening stuck at a certain position (EO)", S3 indicates "Return air duct leakages (RAD)", S4 indicates "Supply air duct leakages (SAD)", and S5 indicates "Duct fouling (DF)".

10

In summary, therefore, in the first embodiment, an electrical signal relating to an airflow device comprising an Air Handling Unit (AHU) 101 is collected. The electrical signal is in the form of electrical current and power information and therefore represents an operating state of the airflow device. Specifically, electrical sensors measure input
15 current and power of the device. The data may consist of noise, so it is filtered by using digital filters. In the next step, data analysis is performed through time series analysis and frequency domain analysis. From the analysed data, features are extracted and the most dominating features are selected. The selected features are again processed using a condition classification model in the form of Support Vector Machine (SVM) based
20 analysis tool, which matches features and selects the most suitable combination and the related condition is the output of this monitoring system. In contrast to conventional SVM models in which hyperparameters are fixed parameters of the model, in the described embodiment, the 'hydrological cycle' optimization algorithm is used to find the optimal values of the hyperparameters. The complete diagnosis process is achieved in 4 stages, wherein stage-1 is data collection and cleaning, stage-2 is features extraction from
25 collected data, stage-3 is optimal feature selection from the extracted features, and the last stage-4 is processing of the optimal selected features. The last stage gives information about the health status of the AHU 101, and therefore the airflow device 100 as a whole.

30

Figs. 11b-11d show an example of three training data sets suitable for use in training the models described above for diagnosis of a fault in the AHU 101, specifically via the diagnosis of a fault in a damper component of an economizer forming part of the AHU 101. The datasets each comprise simulations of the cooling power 2701, i.e. electric
35 power consumed across a cooling coil side of the AHU 101 (with outdoor airflow 2703

also indicated for reference) for 500 samples taken at three different states of the damper. Specifically, in Fig. 11b, the cooling power 2701 is simulated for the damper stuck at a fully open position, in Fig. 11c, the cooling power 2701 is simulated for the damper stuck at a fully closed position, in Fig. 11d, the cooling power 2701 is simulated
5 for the damper in its normal operating condition (i.e. able to open and close).

In practice, the variation in the cooling power 2701 shown in Figs. 11b-11d may be employed for training the decision tree 6001, with the cooling power 2701 for each of the states illustrated in Figs. 11b-11d employed as target outputs (i.e. target selected IMFis).

10

Likewise, the damper conditions for which data is simulated in Figs. 11b-11d may be employed for the training process illustrated in Fig. 10, specifically the conditions of damper fault (stuck fully open Fig. 11b or fully closed, Fig. 11c) and damper operating normally (Fig. 11d) being target outputs in the training process, with the library data 703
15 including the cooling power 2701 for each of the data sets labelled with the corresponding damper condition.

It is envisaged further data sets equivalent to those in Figs. 11b-11d but corresponding to the operating condition of different components of the AHU 101 may be collected and
20 additionally employed for training the models described above.

The method and system according to the described embodiment may provide an accurate, unobtrusive and non-invasive method of determining a condition of an AHU 101. In contrast to existing approaches, the data collected may be electrical current
25 and electrical power data, which may be straightforward to collect from the operating system, without disturbing the existing operation. Moreover, the electrical data may be more accurate relative to mechanical data and enables continuous diagnosis of any problems, without having to wait for the system to be taken offline, for example due to a fault. As electrical data can be collected whenever the AHU 101 is in operation, i.e.
30 online, continuous, 24/7 health monitoring may be possible.

As described above the feature selections and signature matching are carried out by a Support Vector Machine (SVM) based analysis tools, which may significantly improve estimation accuracy. In particular, combining the techniques of the Hydrological Cycle
35 algorithm with SVM techniques in the HC-SVM approach described above may enable improved accuracy of diagnosis of the health condition of the AHU 101 due to improved

robustness to changes in environmental conditions which may affect the operation of the unit and thereby minimizing errors.

For example, during dynamic change in environmental conditions, the AHU operation
5 may also change accordingly. Therefore, under conditions of high dynamic change, the generated data may also fluctuate accordingly, for example closer to a faulty condition. In this situation, accurate condition estimation may be very important. The HC-SVM based technique may be capable of operating and estimating health accurately under a wide variety of environmental conditions.

10

Further by employing a Support Vector Machine (SVM) based discriminative classifier using separate hyperplanes, as described above, it may be possible to detect and distinguish multiple fault conditions at the same instant, potentially simplifying maintenance. The severity and exact location of the fault may also be determined.

15

Further, the method of determining the condition of the AHU 101 according to the described embodiment employs a novel improved empirical mode decomposition (I-EMD) based feature extraction and feature selection technique. Specifically, features are extracted from collected data and then further optimal feature selection from the
20 extracted features is made.

The combined effect of both this features extraction and selection followed by data processing, may enable an undesired health status, or condition of the AHU 101 to be detected even at incipient levels. This may reduce energy waste due to faults and provide
25 enough time to take remedial or schedule maintenance prior to shut down of the system. Life-span of the AHU 101 and therefore the airflow device 100 may also increase due to early intervention when faults occur.

As described above, an ANN-based filter is employed to filter the original IMFs
30 determined from the electrical signal data on the basis of library data or training data, hence potentially enabling good reconstruction and high accuracy of the improved IMF functions. In addition, a unique filter with a unique window size is generated for each IMF by varying the window size depending on frequency and information entropy of the IMF.

Systems according to the described embodiment may be able to detect a fault at an incipient level (for example 10-20% severity level) with an overall accuracy of more than 95%.

- 5 A prototype system for determining the condition of an airflow device according to the first embodiment was prepared in order to test the accuracy of the system and method described above.

10 In AHU systems, in general, four major types of faults occur. These are: Economizer opening stuck at a certain position (EO), Return air duct leakages (RAD), Supply air duct leakages (SAD), and Duct fouling (DF).

Table 1 defines four severity levels for each of these faults, which the severity level indicated in terms of the Fault Impact Ratio (FIR) for the fault, where FIR is defined as
15 the percentage difference in energy consumption between the faulted and baseline (non-faulted) scenario (e.g. FIR = 10% means 10% net site energy increase in faulted scenario compared to the baseline).

Table 1: AHU System Related Faults and Severity Level

Faults Measures	SL: Severity Level (%)			
	SL1	SL2	SL3	SL4
F1: Economizer opening stuck at certain position (EO)	0%	30%	60%	100%
F2: Return air duct leakages (RAD)	10%	15%	25%	30%
F3: Supply air duct leakages (SAD)	10%	15%	25%	30%
F4: Duct fouling (DF)	10%	20%	30%	40%

20 The FIR for each fault type at each severity level is illustrated in Fig. 12.

The prototype system was trained in accordance with the methods described above in order to diagnose these conditions.

25 An overall dataset was obtained and this overall data set was subdivided into a training dataset for training the system and a testing dataset (the datasets being mutually exclusive) for determining the performance of the trained system.

Three categories of tests were performed using the prototype system:

Case-A : Diagnosis in separate fault conditions with all mix severity levels.

- 5 In a first test, a mix of all severity levels (SL) were applied for all four fault types. This test was performed by using 'Cooling Power (CP)' information of the AHU, i.e. electric power consumed across cooling coil side of the AHU.

10 The training/testing data consisted of data samples for the healthy AHU and for each fault with four severity levels (i.e. 5 different cases in total).

Specifically, the training data consisted of 17760 simulated samples (for each case) x 5 cases = 88,800 samples in total. The testing data consisted of 17760 simulated samples (for each case) x 5 cases = 88,800 samples in total. Labelling of the data
15 samples was known from the simulation and assigned as per the known condition class in the target training file. Initialization of the SVM parameters was performed by tuning the SVM parameters using the hydrological cycle (HC) algorithm.

The obtained results of the Cooling Power (denoted "CP" below) based diagnosis tool
20 according to the first embodiment are compared with an existing mechanical-signal based approach known as 'Mixed Air Temperature (MAT)' based AHU diagnosis tool. The results were analysed based on two key parameters, specifically fault detection accuracy, and fault detection time (with respect to the maximum time taken by either technique). The results are shown in Fig. 13a and 13b, respectively.

25 Fig. 13a illustrates that for each fault type, the fault detection accuracy of the developed electrical signature (CP) based health monitoring system according to the described embodiment was accurate and comparable to the 'Mixed Air Temperature (MAT)' based AHU diagnosis tool. However, as shown in Fig. 13b, to reach this
30 accuracy, the Cooling Power based technique according to the described embodiment detected the fault in significantly less time than MAT. For each type of fault, the CP technique according to the described embodiment was 3-4 times faster with respect to MAT. Such rapid performance may enable early and accurate detection of faults of an AHU system.

35

Case-B: Diagnosis at separate severity levels with all mix fault conditions.

In a second test, a mix of all fault conditions (F1-F4) were applied for each severity level. Again, this test was performed by employing 'Cooling Power (CP)' information of the AHU to diagnose faults according to the described embodiment and results were compared with the MAT based AHU diagnosis tool.

The training/testing data consisted of data samples for the healthy AHU and for four faults each with one severity level (i.e. 5 different cases in total).

Specifically, the training data consisted of 17760 simulated samples (for each case) x 5 cases = 88,800 samples in total. The testing data consisted of 17760 simulated samples (for each case) x 5 cases = 88,800 samples in total. Labelling of the data samples was known from the simulation and assigned as per the known condition class in the target training file. Initialization of the SVM parameters was performed by tuning the SVM parameters using the hydrological cycle (HC) algorithm.

Here again, results are analysed on two key parameters, specifically fault detection accuracy, and fault detection time during different severity levels. The obtained results are shown in Fig. 14a and 14b.

Fig. 14a shows that in each case, the fault detection accuracy of the electrical signature (CP) based health monitoring system according to the described embodiment, and the MAT based AHU diagnosis tool are approximately same, which conforms the fault detection ability of both techniques at different severity levels. However, as shown in Fig. 14b, the time taken to reach this level of accuracy was much higher for the MAT technique. Again, at each severity level, the CP approach of the described embodiment was 3-4 times faster at detecting the fault with respect to the MAT. Therefore, this rapid performance demonstrates the early and accurate detection ability of methods according to the described embodiment. Rapid detection of faults may be of particular importance from a safety perspective.

Case-C: Diagnosis during all severity levels and all fault conditions.

In a third test, all four types of fault conditions, and all four severity levels were applied simultaneously during testing. Again, this test was performed by employing 'Cooling Power (CP)' information of the AHU to diagnose faults according to the described embodiment and results were compared with the MAT based AHU diagnosis tool.

5

The training/testing data consisted of data samples for the healthy AHU and for sixteen unhealthy scenarios (4 faults with 4 severity levels), i.e. 17 analysed cases.

Specifically, the training data consisted of 17760 simulated samples (for each case) x 17 cases = 301,920 samples in total. The testing data consisted of 17760 simulated
 10 samples (for each case) x 17 cases = 301,920 samples in total. Labelling of the data samples was known from the simulation and assigned as per the known condition class in the target training file. Initialization of the SVM parameters was performed by tuning the SVM parameters using the hydrological cycle (HC) algorithm.

15 Here again, results are analysed on two key parameters, specifically, fault detection accuracy, and fault detection time during different severity levels. The obtained results are shown in Fig. 15a, 15b and 15c.

Fig. 15a shows that in this complex fault condition (Case C), the fault detection
 20 accuracy of the electrical signature (CP) based health monitoring system according to the described embodiment was improved (95.8% classification accuracy) relative to the MAT based AHU diagnosis tool. (81.1% classification accuracy). Moreover, the time taken to reach even this lower level of accuracy was much higher for the MAT technique. Again, as shown in Fig. 15b at each severity level, the condition determining
 25 approach of the described embodiment was 3-4 times faster at detecting the fault with respect to the MAT technique. Therefore, this rapid performance demonstrates the early and accurate detection ability of methods according to the described embodiment. In practice, therefore, when the possibility of this type of fault is high, the proposed solution may provide quick and accurate results.

30

Fig. 15c shows a confusion matrix of the model output obtained during the Case C testing compared with the actual simulated condition (note that in Fig. 15c, S1 is labelled as NF, S2 is labelled as F1, S3 is labelled as F2, S4 is labelled as F3, and S5 is labelled as F4).

35

Thus, the test results of in each case show that methods and systems according to the described embodiment may offer improvement in terms of quick and accurate detection over existing approaches. This incipient level fault detection according to the described embodiment may improve equipment performance, reduce power wastage and
5 increase the useful operating life of an AHU.

Second Embodiment: Chiller Unit

Fig. 16 illustrates an airflow device 1100, for example an Air Conditioning and
10 Mechanical Ventilation (ACMV) or Heating, Ventilation, and Air Conditioning (HAVC) system, according to a second embodiment. The airflow device 1100 also includes a system for determining a condition of the airflow device 1100, however in this second embodiment this is achieved via determination of a condition of a chiller unit 1101 included in the airflow device 1100. Just as in the first embodiment, the airflow device
15 1100 further includes electrical wires 1107 which connect at least the chiller unit 1101 to an external electrical power source 1105, for example an electrical distribution board.

The chiller unit 1101 is configured to generate chilled liquid for cooling air. As such the
20 chiller unit 1101 may comprise further devices for enabling specific aspects of the function of the chiller unit 1101, including but not limited to a condenser, an evaporator, a refrigerant. The chiller unit 1101 may also comprise a controller having the same or a similar structure to that described in relation to Fig. 2 for controlling specific functions of the chiller unit 1101.

25 The system for determining in the condition of the chiller unit 1101 includes a second condition determination module 1380 as well as an input to the second condition determination module 1380 in the form of an electrical sensor 1103, for example a clamp meter hung on one of the electrical wires 1107, for collecting electrical signal
30 data representing an operating state of the chiller unit 1101, and, as such, the airflow device 1100 as a whole. Specifically, the electrical sensor 1103 is arranged to measure an electrical signal in the form of a current and/or power input to the chiller unit 1101.

Fig. 17 illustrates the condition determination module 1380 in more detail.

Just as in the condition determination module 380, the second condition determination module 1380 is in the form of a computer system and, as such, includes a processor 1382 (which may be referred to as a central processor unit or CPU) that is in communication with memory devices including secondary storage 1384, read only memory (ROM) 1386, random access memory (RAM) 1388, input/output (I/O) devices 1390, and network connectivity devices 1392. The processor 1382 may be implemented as one or more CPU chips. Equivalent components were described in detail above in relation to the condition determination module 380, therefore, for brevity, will not be described again here.

10

Fig. 18 illustrates a method performed by the system for determining condition of the chiller unit 1101, at least in part by executing software on the processor 1382. The method includes steps 1301 to 1309. Steps 1301 to 1307 are preformed analogously to the equivalent steps of the first embodiment, steps 301 to 307.

15

Specifically, in step 1301, data collection is performed. The electrical sensor 1103 receives an electrical signal to the chiller unit 1101 and measures one or more features of the electrical signal, for example current and/or power. The condition determination module 1380 receives the electrical signal data from the electrical sensor 1103 for processing at the CPU 1382.

20

In step 1303, data cleaning is performed on the electrical signal data. For example, digital filters may be applied to reduce noise in the signal. It should be appreciated that other techniques of data cleaning may also be applied.

25

In step 1305 data analysis including time series analysis and frequency domain analysis is performed on the electrical signal data. As in the first embodiment, this is achieved by implementing steps 201 to 217 of Fig. 4 as described above. Specifically, in step 201, in this second embodiment, the cleaned electrical signal data obtained in step 1303 is input as $X(t)$.

30

Just as in the first embodiment, in step 1307, one or more of the improved intrinsic mode functions (IMFi1, IMFi2... IMFin) obtained in step 217 is then selected. This is done by selecting the improved intrinsic mode function or functions that represent one or more intrinsic features which are most relevant for determining the condition of the

35

chiller unit 1101. Again, this selection is performed by inputting the IMFis obtained into a decision tree constructed, according to the described embodiment, using the J48 algorithm. Decision trees and the J48 algorithm were described in detail above therefore, for brevity, they will not be described again here, however, it should be appreciated that the training data for training the decision tree for IMFi selection
5 comprise data specifically corresponding to the operating states of the chiller unit 1101, an example of which is given below.

In step 1309, the improved IMFis determined as being the most relevant for determining
10 the condition of the chiller unit 1101 based on the decision tree are input into a condition classification model to determine the condition of the chiller unit 1101. In contrast to the first embodiment, however, in this second embodiment, the condition classification model is in the form of an artificial neural network (ANN) 1701. As the number of variables for the chiller unit is relatively large, using the ANN 1701 may minimize computational
15 burden. ANNs were described above in general terms in relation to Figs. 5a and 5b.

Fig. 19 illustrates the method of determining the condition of the chiller unit 1101 of Fig. 18 with step 1309 illustrated in block diagram form. Fig. 20 schematically illustrates a mathematical model applied to determine the condition of the chiller unit 1101 and Fig.
20 21 illustrates the data processing stages corresponding to the mathematical model of Fig. 20. The IMFis output from the chiller unit 1101 and selected in step 1307 are input into the ANN 1701 as selected features data D(d) 1705. Features library data, D(l) 1703 is also input into the ANN 1701 as training data for the ANN 1701. Both, D(d) 1705 and D(l) 1703 are processed through the ANN 1701, which outputs D'(d) 1707 which
25 indicates a predicted one of seven health states $F_0, F_1, F_2, F_3, F_4, F_5, F_6$ 1901 of the chiller unit 1101. Specifically, as will be appreciated from Fig. 21, the ANN 1701 outputs a predicted state 1903 for the chiller unit 1101 and the predicted state 1903 is then classified into one of the seven health states $F_0, F_1, F_2, F_3, F_4, F_5, F_6$ 1901 in order to determine the output D'(d) 1707.

30

In common with the SVM 701 of the first embodiment, in this second embodiment, the ANN 1701 is guided by the Hydrological Cycle (HC) Algorithm in an approach which is referred to below as HC-ANN. The steps of this approach are shown in detail in Fig. 22a.

In step 1501, the data is input in the model. This data is in the form of selected improved decomposed components (IMFi1, IMFi2... IMFin) obtained in step 1307, as discussed above.

- 5 In step 1503, input and output parameters are defined for the model. The input parameters consist of one or more improved intrinsic mode functions obtained for the chiller unit 1101 in step 1307 and the output parameters consist of the corresponding health conditions of the chiller unit 1101.
- 10 In step 1505, the library features D(I) 1703 are retrieved for use as input and output parameters for training and validating the ANN 1701. Each library feature D(I) 1703 for either training or validating includes an input parameter consisting of one or more improved intrinsic mode functions obtained for the chiller unit, and an output parameter consisting of a known health condition of the chiller unit at the time the improved intrinsic
- 15 mode function or functions was or were obtained.

In step 1507, an architecture of the ANN model is initialised with an initial number of hidden nodes, weights and biases.

- 20 In step 1509, the ANN 1701 is trained, that is, the input parameters for library features D(I) 1703 for training retrieved in step 1505 are input into the model and the output of the model is compared with the corresponding output parameters for each library feature D(I) 1703 to obtain an error function. It will be appreciated that a variety of error functions could be employed and may be varied according to the pattern of an error data set (for
- 25 example, a constant error data set (all error values are same), linear error data sets (error values are increasing or decreasing with a constant slope), polynomial error data sets (error values are increasing or decreasing with variable slope), or random error data sets (error values are not following any pattern)). Biases of the ANN are varied until the error is minimised.

30

- In step 1511, the ANN 1701 is validated, that is, the input parameters for library features D(I) 1703 for validating retrieved in step 1505 are input into the model and the output of the model is compared with the corresponding output parameter for each library feature D(I) 1703 to obtain an error function. Hyperparameters of the ANN 1701 are varied until
- 35 the error is minimised. The error function is based on the mapping of predicted value by

the model with respect to the target value of the class. Ideally, it is as low as possible. An example of an acceptable error level may be about 1% or below. If the error level is not within an acceptable range then the super parameters will be adjusted.

- 5 In step 1513 a model error is calculated using a test dataset.

In step 1515, the model error obtained in step 1513 is compared with stopping criteria. If the model error satisfies the stopping criteria, for example, it is below a threshold, for example less than about 0.01% then, in step 1517, the input data is input into the ANN
10 1701 and an estimated health condition of the chiller unit is output by the model.

If the stopping criteria are found not to be met in step 1515, then the method proceeds to a hydrological cycle (HC) algorithm consisting of steps 1519 to 1527 and the weights of the ANN are adjusted using this algorithm. Steps 1519 to 1527 represent each stage
15 of the hydrological algorithm described in terms of their equivalent real-life hydrological cycle counterparts.

Specifically, in step 1519 the velocity is calculated according to the ANN weight performance.
20

In step 1521, the soil and depth are updated.

In step 1523 the temperature, condensation and precipitation are updated.

- 25 In step 1525 the water drops are updated.

In step 1527 the weights of the ANN 1701 are updated.

As noted above, the HC algorithm (steps 1519-1527) is a meta-heuristic optimization
30 technique, where specific training data for the HC algorithm is not required. A random set of data is therefore employed initially and then updated in each iteration. Finally, the algorithm converges on an optimal result.

Once the updated weights of the ANN are determined in step 1527, the method returns
35 to step 1507 and an initial model with the ANN weights determined in step 1527 is

generated. The weights and hyperparameters of the ANN model are then trained and validated, respectively, in steps 1509 to 1513, as before until the stopping condition is satisfied in step 1515.

- 5 Thus, in this second embodiment, the training of the condition classification model includes adjusting a first set of parameters in the form of biases and hyperparameters of the ANN 1701 using a first optimization algorithm, and a second set of parameters, in the form of weights of the ANN 1701 using a second optimization algorithm based on the hydrological cycle algorithm.

10

Once it is found in step 1515 that the stopping criteria are met, then the method proceeds to step 1517. Specifically, in step 1517 the optimized ANN 1701 is applied to the selected features data D(d) 1705 comprising the selected IMFs and/or IMFis and a predicted state 1903 is output from the ANN 1701 which indicates one or more of the seven health states

15 $F_0, F_1, F_2, F_3, F_4, F_5, F_6$ 1901 of the chiller unit 1101 is then output, for example it is displayed by a display forming part of the I/O devices 1390 of the condition determination module 1380.

- In an example, the seven health states $F_0, F_1, F_2, F_3, F_4, F_5, F_6$ 1901 include a first
- 20 category (F_0) indicating a normal operating condition of the chiller unit 1701, and the remaining six categories ($F_1 - F_6$) indicate six different types of fault conditions. Specifically, F_1 indicates Reduced Condenser Water Flow (FWC), F_2 indicates Reduced Evaporator Water Flow (FWE), F_3 indicates Refrigerant Leak (RL), F_4 indicates Refrigerant Overcharge (RO), F_5 indicates Condenser Fouling (CF), and F_6 indicates
- 25 Non-Condensable in System (NC).

- Fig. 22b shows an example of two training data sets suitable for use in training the models described above for diagnosis of a fault in the chiller unit 1101, specifically via the diagnosis of compressor component of the chiller unit 1101. The datasets each
- 30 comprise simulations of the compressor power consumption for 3000 samples taken at two different states of the compressor. Specifically, power data during a normal operating condition of the compressor are indicated by line 2081 and power data under conditions of refrigerant leakage are indicated by line 2803.

In practice, the variation in the compressor shown in Figs. 22b may be employed for training the decision tree for selection of the improved IMFis, with the compressor power 2801, 2803 for each of the states illustrated in Figs. 22b employed as target outputs (i.e. target selected IMFis).

5

Likewise, the compressor conditions for which data is simulated in Fig. 22b may be employed for the training process illustrated in Fig. 22a, specifically the conditions of compressor fault (refrigerant leakage indicated by line 2803) and compressor operating normally (indicated by line 2801) being target outputs in the training process, with the library data 1703 including the cooling power indicated by lines 2801, 2803 for each of the data sets labelled with the corresponding compressor condition.

10

Note that due to leakage of refrigerant, pressure level of refrigerant across the condenser and evaporator is reduced as shown in Figures 22c and 22c, respectively, thereby affecting performance.

15

Thus, in the second embodiment, the condition classification model is again adapted using the hydrological cycle (HC) algorithm. In conventional approaches involving ANNs, weights are typically fixed. In contrast to this conventional approach however, the method for determining the condition of the chiller unit according to the second embodiment implements the 'hydrological cycle' (HC) optimization algorithm to determine optimal values of the weights of the ANN 1701. This may improve performance of the ANN-based classification model and may enable improved accuracy of diagnosis of the health condition of the chiller unit 1101 by increasing robustness to changes in environmental conditions which may affect the operation of the unit.

20

25

For example, during dynamic change in environmental conditions, the chiller unit operation may also change accordingly. Therefore, under conditions of high dynamic change, the generated data may also fluctuate accordingly, for example closer to a faulty condition. In this situation, accurate condition estimation may be very important. The HC-ANN based technique may be capable of operating and estimating health accurately under a wide variety of environmental conditions.

30

The combination both feature extraction and selection as embodied in step 1305 and step 1307 above followed by data processing, the method for determining the condition

35

of the chiller unit according to the second embodiment may be able to predict the health status of the chiller unit even at incipient levels. In particular, electrical signals have been shown to reflect even small changes in the functioning of chiller units. In contrast, mechanical signals may not reflect anomalies in the functioning of a chiller unit at an
5 early, or incipient level.

As the determination is based on electrical current and power data, the data may be accurate and easy to capture in a non-intrusive and non-invasive way and may leverage existing sensors already present in the Airflow device. In addition, such an “on-line”
10 diagnosis method may enable the monitoring of the condition of the chiller unit on a continuous basis in real time without having to bring the chiller unit off-line.

By selecting the most relevant feature from the electrical signal by determining improved IMFis and selecting the most relevant ones for input in to the classification model,
15 detection of faults or other “unhealthy” conditions in the chiller unit 1101 may be possible at the incipient level which will may reduce energy wastage and provide additional time to take remedial actions or schedule maintenance of the airflow. The multi-layer neuron approach employed with the ANN (i.e. the fact that one or more hidden layers is employed) may enable the detection and enable multiple fault conditions to be
20 distinguished at the same time, which may facilitate maintenance tasks. Due to potential improvement in understanding of the health-related issues of the chiller unit at the incipient level, the overall power consumption of the unit may be reduced significantly, as well as life span of the chiller unit increased.

25 A prototype system for determining the condition of a chiller unit according to the second embodiment was prepared and the accuracy of condition detection tested. The performance and accuracy of fault detection ability for different types of fault at different severity level are provided below.

30 In chiller systems, in general, four major types of faults occur. These are: Reduced Condenser Water Flow (FWC), Reduced Evaporator Water Flow (FWE), Refrigerant Leak (RL), Refrigerant Overcharge (RO), Condenser Fouling (CF), and Non-Condensable in System (NC).

Table 2 defines four severity levels for each of these faults, which the severity level indicated in terms of the Fault Impact Ratio (FIR) for the fault. As before, FIR is defined as the percentage difference in energy consumption between the faulted and baseline (non-faulted) scenario (e.g. FIR = 10% means 10% net site energy increase in faulted scenario compared to the baseline).

Table 2: Chiller Unit Related Faults and Severity Level

Chiller Faults Measures			SL: Severity Level (%)			
			SL1	SL2	SL3	SL4
F1:	Condenser Fouling	(CF)	10%	20%	30%	40%
F2:	Reduced Condenser Water Flow	(FWC)	10%	20%	30%	40%
F3:	Reduced Evaporator Water Flow	(FWE)	10%	20%	30%	40%
F4:	Non-Condensables in System	(NC)	1%	2%	3%	5%
F5:	Refrigerant Leak	(RL)	10%	20%	30%	40%
F6:	Refrigerant Overcharge	(RO)	10%	20%	30%	40%

The FIR for each fault type at each severity level is illustrated in Fig. 23.

The prototype system was trained in accordance with the methods described above in order to diagnose these conditions.

An overall dataset was obtained and this overall data set was subdivided into the training dataset and the testing dataset (the datasets being mutually exclusive).

Three categories of tests were performed using the prototype system:

Case-A : Diagnosis in separate fault conditions with combined all severity levels.

In a first test, a mix of all severity levels (SL) were applied for all six fault types. This test was performed by using electrical power (denoted P below) information of the chiller unit.

The training/testing data consisted of data samples for the healthy chiller unit and for each fault with four severity levels (i.e. 5 different cases in total). Specifically, the training data consisted of 3640 simulated samples (for each case) x 5 cases = 18200 samples in total. The testing data consisted of 1551 simulated samples (for each case) x 5 cases = 7755 samples in total. Binary labelling of the training and testing data was manually conducted for each of the 5 cases. Initialization of the ANN parameters

comprised setting the number of neurons at the input, and output layers equal to the number of input variables (N_i) and number of cases/target/class (N_o), respectively. The number of neurons at hidden layer (N_h) was evaluated as $N_h = \frac{N_i + N_o}{2} + \sqrt{N_p}$, where N_p is number of data samples. Finally, initial values of the weights and biases of the ANN were selected using the hydrological cycle (HC) algorithm.

The obtained results of the electrical based diagnosis tool according to the second embodiment were compared with existing approaches, specifically, a water flow rate (WFR) based diagnosis tool, a refrigerant pressure (RP) based diagnosis tool, and a temperature (T) based diagnosis tool. Results were analysed on two key parameters: fault detection accuracy and fault detection time, as shown in Fig. 24a and Fig. 24b, respectively.

Fig. 24a illustrates that for each fault type, the fault detection accuracy of the electrical signature (P) based health monitoring system according to the described embodiment was accurate and comparable to the other approaches. However, as shown in Fig. 24b to reach this accuracy, the electrical signature-based technique according to the described embodiment detected the fault in significantly less time. For each type of fault, the P-based technique according to the described embodiment was 3-4 times faster with respect to the other detection methods. Such rapid performance may enable early and accurate detection of faults of chiller system.

Case-B : Diagnosis at separate severity levels with combined all fault conditions.

In a second test, a mix of all fault conditions (F1-F6) were applied for each fault type. The training/testing data consisted of data samples for the healthy chiller unit and for six faults each with one severity level (i.e. 7 different cases in total).

Specifically, the training data consisted of 3640 simulated samples (for each case) x 7 cases = 25480 samples in total. The testing data consisted of 1551 simulated samples (for each case) x 7 cases = 10857 samples in total. Binary labelling of the training and testing data was manually conducted for each of the 7 cases. As in case A, Initialization of the ANN parameters comprised setting the number of neurons at the input, and output layers equal to the number of input variables (N_i) and number of cases/target/class (N_o),

respectively. The number of neurons at hidden layer (N_h) was evaluated as $N_h = \frac{N_i + N_o}{2} + \sqrt{N_p}$, where N_p is number of data samples. Finally, initial values of the weights and biases of the ANN were selected using the hydrological cycle (HC) algorithm.

5

This test was performed using the same techniques as described above and the results are illustrated in Fig. 25a and 25b.

Fig. 25a shows that in each case, the fault detection accuracy of the electrical signature (P) based health monitoring system according to the described embodiment performs well, as do the other techniques. However, as evident from Fig. 25b, the time taken to reach this level of accuracy was much faster for the P-based technique according to the described embodiment. Again, at each severity level, the condition determining approach of the described embodiment was 3-4 times faster at detecting the relevant fault with respect to the other techniques. This rapid performance demonstrates the potential early and accurate detection ability of methods according to the described embodiment.

Case-C: Diagnosis during all severity levels and all fault conditions.

20

Finally, in a third test, all the six different types of fault conditions, and all four severity levels were applied instantaneously during testing using the same techniques as for the results of Fig. 24 and 25 described above.

The training/testing data consisted of data samples for the healthy chiller unit and for 24 unhealthy scenarios (six faults each with four severity level (i.e. 25 different cases in total)).

Specifically, the training data consisted of 3640 simulated samples (for each case) x 25 cases = 91000 samples in total. The testing data consisted of 1551 simulated samples (for each case) x 25 cases = 38775 samples in total. Binary labelling of the training and testing data was manually conducted for each of the 25 cases. As in case A, Initialization of the ANN parameters comprised setting the number of neurons at the input, and output layers equal to the number of input variables (N_i) and number of cases/target/class (N_o),

30

respectively. The number of neurons at hidden layer (N_h) was evaluated as $N_h = \frac{N_i + N_o}{2} + \sqrt{N_p}$, where N_p is number of data samples. Finally, initial values of the weights and biases of the ANN were selected using the hydrological cycle (HC) algorithm.

5

The results are shown in Fig. 26a and 26b.

As is evident from Fig. 26a, the P-based technique according to the described embodiment exhibited improved accuracy relative to the other techniques employed.

10 Further, this was achieved in a much shorter time with respect to the other techniques, as evident from Fig. 26b which shows that the P-based technique according to the described embodiment was 3-4 faster than the other techniques employed.

Fig. 26c shows a confusion matrix of the model output obtained during the Case C

15 testing compared with the actual simulated condition (note that in Fig. 26c, no fault is labelled as NF).

Thus, the testing demonstrated that methods according to the described embodiment offer advantages over existing techniques in terms of quick and accurate detection. Such
20 incipient level fault detection may improve equipment performance, reduce power wastage and increase the useful operating life of the chiller unit.

The described embodiments should not be construed as limitative.

25 For example, it is envisaged that the diagnosis methods described could be employed with components of airflow devices other than AHUs and chiller units.

Further, it is envisaged that the classification model could be a machine learning model other than an SVM or ANN model. Although empirical mode decomposition is described
30 above as being employed to decompose the electrical signal data, it is envisaged that decomposing the electrical signal could be performed using mathematical techniques other than an empirical mode decomposition.

Although the parameters of the classification model are advantageously trained using
35 the hydrological cycle algorithm, it is envisaged that improved intrinsic mode functions

may be selected for use with a classifier that has not been trained based on the hydrological cycle algorithm.

Conversely, improved IMFis may not be calculated, and the whole original electrical signal data or the unfiltered IMFis (i.e. those obtained in step 205) may be input directly into a classification model with parameters trained using the hydrological cycle algorithm.

Filtering of the IMFis in order to obtain the improved IMFis may be performed using a filter other than an ANN filter which may or may not have an adjustable window size based on frequency and/or be based on information entropy. Identification of the most relevant intrinsic feature or features may be performed using a classifier other than a decision tree, for example a neural network-based classifier. A decision tree generated with an algorithm other than the J48 algorithm may be employed to determine the most relevant intrinsic feature or features.

It is envisaged that the airflow device may include both a chiller unit and an AHU and one or more systems for determining the condition of the airflow device which receive electrical signal data from one or both of the chiller unit and AHU.

Although the condition of an airflow device comprising a chiller unit is described as being determined based on electrical signal data of the chiller unit using a condition classification model based on an ANN, it is envisaged that the condition of an airflow device comprising a chiller unit may be determined based on electrical signal data of the chiller unit using a condition classification model based on an SVM.

Although the condition of an airflow device comprising AHU is described as being determined based on electrical signal data of the AHU using a condition classification model based on an SVM, it is envisaged that the condition of an airflow device comprising an AHU may be determined based on electrical signal data of the AHU using a condition classification model based on an ANN.

Although measuring the power input to the air handling unit or chiller unit as a whole may be performed, it is envisaged that the power input to a sub-component of either one of the air handling unit or chiller unit could be measured, for example a cooling coil of the air handling unit (in which case the power data may be described as the "cooling power")

or a compressor motor of a chiller unit (in which case the power data may be described as the “compressor motor power consumption”)

5 Although specific divisions of the training and testing data sets are described above in relation to the experimental data discussed, as well as specific numbers of training and testing data samples, it should be appreciated that these are intended to be only examples of suitable training data and that fewer or greater numbers of training and/or testing data samples may be employed according to embodiments.

10 Although specific categories of faults are described above in association with AHUs and chiller units, it should be appreciated that other categories of faults may exist and be monitored according to embodiments. Moreover, sub-categories of the fault categories listed above, or any other fault categories may also be classified using methods according to embodiments.

15 It will be appreciated that features of one embodiment may be employed in accordance with other embodiments.

20 Having now fully described the invention, it should be apparent to one of ordinary skill in the art that many modifications can be made hereto without departing from the scope as claimed.

CLAIMS

1. A method of determining a condition of an airflow device, comprising:
 - 5 (i) receiving electrical signal data representing an operating state of the airflow device;
 - (ii) decomposing the electrical signal data into a plurality of first decomposed components of the electrical signal data representing corresponding intrinsic features of the electrical signal data;
 - 10 (iii) filtering each of the plurality of first decomposed components using a neural-network based filter to produce reconstructed electrical signal data;
 - (iv) decomposing the reconstructed electrical signal data into a plurality of second decomposed components of the reconstructed electrical signal data representing corresponding optimum intrinsic features, the optimum
15 intrinsic features being more relevant to identifying the condition than the intrinsic features;
 - (v) identifying at least one most relevant intrinsic feature for determining the condition of the airflow device from at least the optimum intrinsic features; and
 - 20 (vi) selecting at least one of the second decomposed components corresponding to the at least one most relevant intrinsic feature as an input to a condition classification model to determine the condition of the airflow device.
- 25 2. The method of determining a condition of an airflow device according to claim 1, wherein the at least one most relevant intrinsic feature is identified using a decision tree.
3. The method of determining a condition of an airflow device according to claim 1 or 2, wherein the neural-network based filter has a window size, the window size being
30 dependent on a frequency of the first decomposed component being filtered.
4. The method of determining a condition of an airflow device according to any one of claims 1 to 3, wherein decomposing the electrical signal data into a plurality of first decomposed components comprises decomposing the electrical signal data using a first
35 empirical mode decomposition, and wherein decomposing the reconstructed electrical

signal data into a plurality of second decomposed components comprises decomposing the reconstructed electrical signal data using a second empirical mode decomposition.

5. The method of determining a condition of an airflow device according to any one of
5 claims 1 to 4, wherein the condition classification model comprises a machine learning model described by at least first and second parameters, and wherein the method further comprises:

10 training the machine learning model, wherein training the machine learning model includes adjusting the first parameters of the machine learning model using a first optimization algorithm and adjusting the second parameters of the machine learning model using a second optimization algorithm based on a hydrological cycle (HC) algorithm.

6. The method of determining a condition of an airflow device according to claim 5,
15 wherein the neural-network based filter is also trained based on the hydrological cycle (HC) algorithm.

7. A method of determining a condition of an airflow device, comprising:

- 20 (i) receiving electrical signal data representing an operating state of the airflow device;
- (ii) inputting the electrical signal data into a condition classification model to determine the condition of the airflow device, the condition classification model comprising a machine learning model being described by at least first and second parameters; and
- 25 (iii) training the machine learning model, wherein training the machine learning model includes adjusting the first parameters of the machine learning model using a first optimization algorithm and adjusting the second parameters of the machine learning model using a second optimization algorithm based on a hydrological cycle (HC) algorithm.

30

8. The method of determining a condition of an airflow device according to any one of the preceding claims, wherein the airflow device comprises an air handling unit, and wherein the electrical signal data is representative of an operating state of the air
35 handling unit.

9. The method of determining a condition of an airflow device according to claim 8, wherein the condition classification model comprises a Support Vector Machine.
- 5 10. The method of determining a condition of an airflow device according to any one of the preceding claims, wherein the airflow device comprises a chiller unit, and wherein the electrical signal data is representative of an operating state of the chiller unit.
11. The method of determining a condition of an airflow device according to claim 10,
10 wherein the condition classification model comprises an artificial neural network.
12. System for determining a condition of an airflow device, the system comprising:
an input for receiving electrical signal data representing an operating state of the
airflow device;
15 a processor configured to:
decompose the electrical signal data into a plurality of first decomposed
components of the electrical signal data representing corresponding
intrinsic features of the electrical signal data,
filter each of the plurality of first decomposed components using a neural-
20 network based filter to produce reconstructed electrical signal data,
decompose the reconstructed electrical signal data into a plurality of
second decomposed components of the reconstructed electrical signal
data representing corresponding optimum intrinsic features, the optimum
intrinsic features being more relevant to identifying the condition than the
25 intrinsic features,
identify at least one most relevant intrinsic feature for determining the
condition of the airflow device from at least the optimum intrinsic features,
and
select at least one of the second decomposed components corresponding
30 to the at least one most relevant intrinsic feature as an input to a condition
classification model to determine the condition of the airflow device; and
an output configured to output data indicating the condition of the airflow device.
13. The system for determining a condition of an airflow device according to claim 12,
35 wherein the condition classification model comprises a machine learning model

described by at least first and second parameters, and wherein training the machine learning model includes adjusting the first parameters of the machine learning model using a first optimization algorithm and adjusting the second parameters of the machine learning model using a second optimization algorithm based on a hydrological cycle (HC) algorithm.

14. System for determining a condition of an airflow device, the system comprising:
- an input for receiving electrical signal data representing an operating state of the airflow device; and
 - 10 a processor configured to:
 - input the electrical signal data into a condition classification model to determine the condition of the airflow device, the condition classification model comprising a machine learning model described by at least first parameters and second parameters, and
 - 15 train the machine learning model, wherein training the machine learning model includes adjusting the first parameters of the machine learning model using a first optimization algorithm and adjusting second parameters of the machine learning model using a second optimization algorithm based on a hydrological cycle (HC) algorithm; and
 - 20 an output configured to output data indicating the condition of the airflow device.

15. The system for determining a condition of an airflow device according to any one of claims 12 to 14, further comprising an electrical sensor configured to receive electrical signal data representing the operating state of the airflow device.

16. The system for determining a condition of an airflow device according to any one of claims 12 to 15, wherein the airflow device comprises an air handling unit and wherein the electrical signal data is representative of an operating state of the air handling unit.

17. The system for determining a condition of an airflow device according to any one of claims 12 to 15, wherein the airflow device comprises a chiller unit and wherein the electrical signal data is representative of an operating state of the chiller unit.

18. An airflow device comprising the system for determining the condition of the airflow device according to any one of claims 12 to 17.

19. A computer readable medium configured to cause a computer to perform a method of determining a condition of an airflow device according to any one of claims 1 to 11.