Indoor lighting fault detection and diagnosis using a data fusion approach

F. Marino[1], A. Capozzoli[2], M. Grossoni[1], F. Lauro[1,2], F. Leccese[1], F. Moretti[1,3], S. Panzierei[1] & S. Pizzuti[1,3]
[1]*Department of Electronic Engineering, Roma Tre University, Italy*
[2]*Department of Energy, Turin Polytechnic, Italy*
[3]*Energy New technologies and sustainable Economic Development Agency (ENEA), Italy*

Abstract

In this paper, an innovative and automated fault detection and diagnosis (FDD) approach based on high-level correlation rules in order to improve reliability, safety and efficiency of a supervised building is presented. The proposed method is based on the data fusion of different measurements, using their fuzzification and aggregation through suitable operators, in order to get dimensionless severity indicators able to diagnose faults and to identify the possible causes (ranked according their severity) generating them. Thus, a set of possible anomalies that can occur in a building and the correlation with measured physical quantities were identified. Experimentation of this FDD technique was applied to indoor lighting of a real office building. The proposed method was validated over a one-month period with the aim of detecting anomalous consumption events, considering when and in which circumstances they occurred. After this stage, the FDD system was performed in real time operation.
Keywords: fault detection and diagnosis, smart building, BEMS, data fusion approach, fuzzy logic, sustainability.

1 Introduction

Recent studies have shown that the building sector is responsible for more than 40% of European energy consumption (IEA [1]). It has also been estimated that two-thirds of all the energy used in a typical office is related to lighting, electronic devices and HVAC (*heating, ventilating, air-conditioning*) systems.

WIT Transactions on Ecology and The Environment, Vol 190, © 2014 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/EQ140101

Poorly maintained and improperly controlled equipment wastes an estimated 15% to 30% of energy (Katipamula and Brambley [2]): the prevention of this waste and the potential energy savings can be achieved with an appropriate energy management trough widespread automated diagnostics and control activities (Schneider Electric [3]).

There is plenty of literature on process fault diagnosis (Katipamula and Brambley [2], Ding [4]), ranging from analytical methods (Khan *et al.* [5]) to statistical and artificial intelligence approaches (Lauro *et al.* [6]). Some methods require accurate process models, quantitative or semi-quantitative models (Kavuri *et al.* [7]), or qualitative models (Simani [8], Kavuri *et al.* [9]), and, fundamentally, a priori knowledge of the physics principles of the process itself; on the other side, there are methods that don't assume any form of model information and only rely on historical process data in an empiric way (e.g. black box, PCA) (Marzat *et al.* [10], Kavuri *et al.* [11], Ma and Jiang [12]). In fig. 1, a scheme of the most known and used diagnostics methods is illustrated.

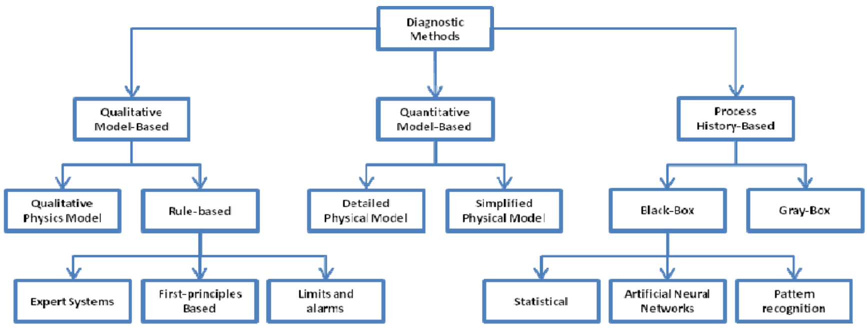

Figure 1: Scheme of diagnostic methods.

Nowadays *Fault Detection and Diagnosis* (FDD) concerning buildings' energy behavior is a critical issue. Its fundamental aspects are the improvement of the reliability, safety and efficiency of the supervised system and also the operating stability to ensure high performances and to prevent overall damages (Hyvfirinen and Kirki [13], Dexter and Pakanen [14]).

Since *Building Energy Management Systems* (BEMS) are increasingly taking place, much information is getting available, thus human and automated monitoring are getting harder and harder and, at present, automated supervisory and control activities are limited to low-level logics (e.g. thresholds, basic correlation rules).

In this paper, an innovative and automatic system of fault detecting that can be applied to commercial building networks not only to a single structure, in order to obtain a scale economy with the monitoring of the main working parameters is proposed. Another important advantage is about energy savings, and thus economic benefits, only through the management of its electrical consumptions minimizing deterioration of building and its facilities. The last innovative aspect is the thought and tested method itself: from the data acquired by sensor and reported to the database, dimensionless indicators referring to several measurements are defined, combined with data fusion operations and organized from a low level to a high one, that fully describe the physical state and energy behavior characterizing the building itself. From this information one can find out easily and quickly the reason of the abnormal event and provide to take back all the parameters in the correct operating range.

The paper is further organized as follows: section 2 introduces the used algorithm to detect peaks of electric consumption. Section 3 describes the FDD method that is going to be tested and evaluated for a particular diagnostics rule, explained in details on section 4. Section 5 reports results of this experimentation. Finally section 6 draws conclusions and states possibilities for future work.

2 Peak detection algorithm

Identifying and analyzing peaks in a given time-series play a very important role because they show sudden and unexpected events in many applications. It is necessary to formalize the notion of a peak through the definition of appropriate algorithms that are able to automatically detect them, in order to avoid misleading subjective considerations.

In this experimentation, the method used to detect peaks of electrical consumption is here reported, based on the idea of Palshikar [15].

Let $T = x_1, x_2, \ldots, x_N$ be a given uniformly sampled time-series containing N values. Let x_i be a given i^{th} point in T ($i=1, 2,\ldots,N$). Let S be a given *peak function*, which associates a score (a non-negative real number) $S(i, x_i, T)$ to each element of T. This function computes the average of the maximum among the signed distances of x_i from its k left neighbors and the maximum among the signed distances of xi from its k right neighbors, where k is a positive integer parameter (generally 3 to 5) that determines the size of the local window, $2k$, around the x_i element. Then for all the items of the given time series that return a positive value of $S(k, i, x_i, T)$ is calculated the mean and standard deviation, m' and s'.

Finally, if the eqn (1)

$$S(k,i,x_i,T) - m' \geq h * s' \tag{1}$$

is satisfied that the considered element is identified as a peak (h is a constant value specified by the user on the degree of abnormality).

An important consideration was about the choice of the size of the time series: as shown in eqn (1), a peak is detected depending on the standard deviation and the mean of the series itself. A more or less wide time-window determines a different system’s sensitivity in discriminating outliers. After a careful study and several experimental trials, a time span of 24 hours that resulted the optimum between speed and reliability of anomalies’ detection, was chosen.

Thus, the analysis takes in consideration the first 144 samples (each sample of power consumption is acquired every 10 minutes). With the occurrence of a new sample, the chronologically older will be eliminated and so on to the n^{th} iteration. The *S* value of the peak detected will be inserted in the database and will be object of further proceedings.

3 The FDD method

The proposed FDD method is based on the data fusion of different kind of measurements about electrical and thermal consumptions, weather conditions, and working hours, in order to define, through their aggregation, dimensionless and normalized severity indicators able to isolate faults and identify the possible causes (ranked according their severity) generating them.

The diagnostic methodology is aimed at supporting the complex decision-making process related to the identification of possible causes of anomalies (values or abnormal trends) by a pre-processing procedure of monitored data.

The adopted approach requires the definition of three different lists of:

1) pre-processing indices of energy and environment monitored data, for each final energy use representing the *symptoms*;
2) events related to the anomalies (*situations assessment*);
3) actual *causes*.

and the operations among them (fig. 2).

A set of potential anomalies that can occur in a building and the correlation with measured physical quantities was first defined, associated with decisional trees build on expert systems.

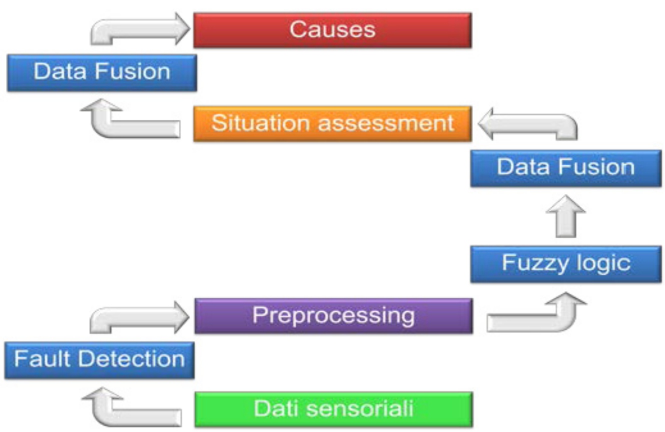

Figure 2: Process to detect faults.

The whole diagnostics process starts with the fault detection analysis performed through the peak detection method above described on the monitored data set: a detected peak is considered as a potential fault for the considered cause. Then, through data fusion operations (based on fuzzy logic and suitable aggregation) of different kind of measurements, a ranking about the severity and reliability of the evaluated anomalous values is carried out. When a fault is detected then is stored with all the other associated and relevant information on a database: in this way a real-time diagnostic and also off-line statistical analysis is performed. Moreover an email notification is sent to the energy manager to make him up to date about the energy performances of the building under exam.

4 The diagnostics rule

In order to verify the correctness and the suitability of the proposed method a diagnostics rule, representing the logic form of the cause under exam, was defined. To this purpose the experimentation was performed for the evaluation of anomalies related to the "indoor lights switched on in relation to occupied rooms and working hours" (for an occupied room we intend the presence of the employee in the research centre through the badge pass).

A decision trees approach was adopted and its construction was based on the identification of the relationship existing between each cause and the connected situations assessment and between these and the output of the pre-processing procedure. Through this methodology was possible to evaluate all the potential associations of each cause related to the events and then to the monitored variables. In fig. 3 a scheme with the considered diagnostics rule and the preprocessing (P), situations (S) and causes (C) indicators is shown.

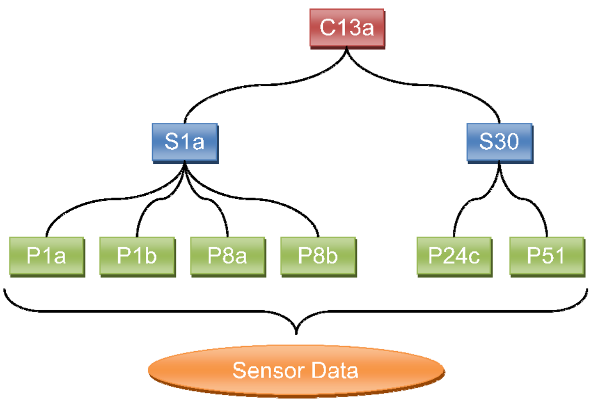

Figure 3: Scheme of the diagnostic rule.

Below, the meanings of each preprocessing indicators considered are reported:

P1a: Outlier electrical power;
P1b: Outlier electrical energy;

P8a: Abnormal electricity trend than the historical one of the same group of weekdays (Mon/Tue–Thur/Fri/weekend);
P8b: Abnormal electricity trend than the previous week;
P24c: Percentage of occupied rooms;
P51: Working hours.

The situations indicators are obtained through the aggregation of particular preprocessing:

S1a: Simultaneous lights switched on of an abnormal number of electrical utilities;
S30: Large number of occupied rooms compared to working hours.

Finally, the indicator of the fault itself, *C13a*, represents also an important indication of the users' behavior and is obtained by the combination of the situations previously specified.

From the mathematical point of view the relations among these indicators are (eqns (2), (3), (4)):

$$S1a = P1a\ OR\ P1b\ OR\ (0.7 * P8a + 0.3 * P8b) \quad (2)$$

$$S30 = P51\ AND\ P24c \quad (3)$$

$$C13a = 0.7 * S1a + (1 - 0.7) * S30 \quad (4)$$

where the *AND* and *OR* operations correspond in the fuzzy logic, respectively, to the minimum and maximum calculation, Novàk *et al.* [16] . Fuzzy logic was applied with the aim to make sensor data, coming from heterogeneous sources, homogeneous and integrated to each other. In this experimentation only the aspect about the outlier of electrical power (*P1a*) was considered but in future testing operations we'll be extended to other preprocessing indictors (e.g. the above defined *P1b*, *P8a*, and *P8b*). *P24c* indicate the presence of 3 employees of 4 on each floor and, for its definition, is already a blurred value between 0 and 1. *P1a* has been fuzzyfied with a sigmoid expression to normalize the severity of the electrical peak detected; finally *P51* has been fuzzyfied with an inverse Gaussian function to empathize the condition representing out of working hours. The equations used to represent these functions are defined below, eqns (5), (6).

$$sigm(x) = \frac{1}{1 + e^{-\frac{(x-c)}{t}}} \quad (5)$$

$$gauss(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

In table 1 the membership functions and parameters for the particular kind of fuzzy set are reported in detail.

Table 1: Membership functions and parameters for different fuzzy set.

Fuzzy set	Membership function	Parameters
P1a	Sigmoid	c=0.08 t=0.1
P51	Gaussian	μ=12 σ=6

These parameters of the membership functions have been tuned with offline analysis in order to catch the manually generated faults in a reliable way.

5 Experimentation

The proposed FDD method has been implemented in a multithreading environment application written in JAVA for a one-month period. Electrical consumption's peaks detected become the basis of the following evaluation of their severity on the overall F40 building performances.

This building can be considered, for many aspects, as an example of many old structures belonging to the overall service industry and, in particular, to the public administration. It is especially constituted by offices and labs, is organized in 3 floors and is highly equipped with sensors and actuators. This structure is located inside the ENEA Research Center 'Casaccia' and about 50 employees work there daily, starting from 8 to 18. Regarding the energy aspect, this building has 4 electric lines related to lighting, electromotive force, conditioning and the emergency line.

Results of one month working of the FDD technique were collected and then examined and considerations concerning when and in which circumstances abnormal events occurred were made.

As an example, in fig. 4 is reported the screenshot of the ground floor's total active power referred to the 29 March 2013, the day in which a fault was manually generated with the aim to tune the parameters of the membership functions and to reveal effectively and in a reliable way the anomalous event. This trend shows the detected peaks with the peak detection method, highlighted with red squares, in the particular context and for the sensibility of the analysis; for each of them the instant value (in kW) and the date and time when they happened are reported.

In fig. 5 the final result of the complete diagnostics system is reported and only two indicators, comparing to the several detected peaks, are shown. These indices don't correspond to the fault severity but represent the plausibility that a definite peak is associated to the cause under experimentation. So, with this FDD method, even in presence of high values of the S function potentially dangerous, peaks couldn't be considered as anomalies for certain situations while could be extremely significant in other circumstances. In this case, the two values, 0.768 and 0.773, reflect the fuzzification of the examined cause (C13a) and, since they exceed the threshold set by the user (0.5), will be reported in the database.

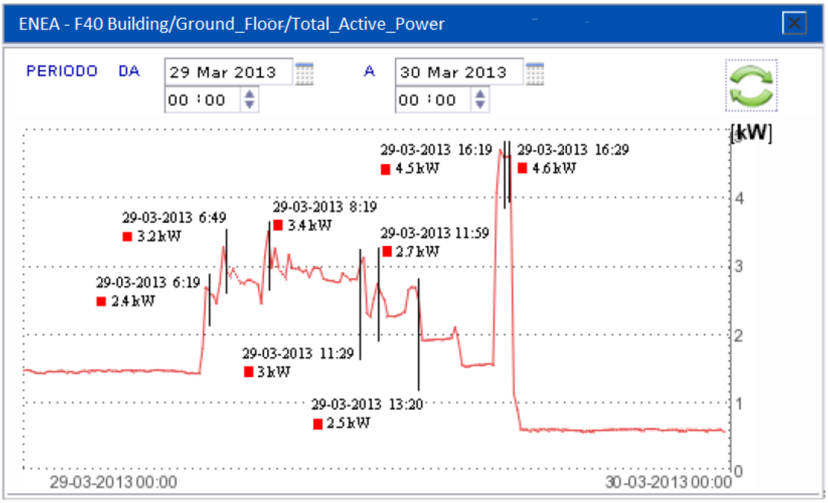

Figure 4: Trend of total active power – Ground Floor – 29/03/2013. Detected peaks are shown with red squares.

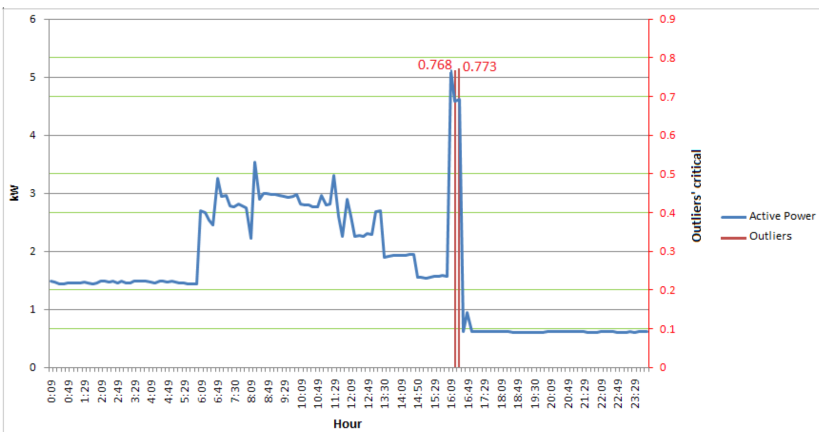

Figure 5: Total active power and outliers' indicators – Ground Floor – 29/3/2013.

Promptly, an e-mail will be sent to the energy manager with information about the date and time when the outliers occurred, the floor of the building, the ID code and the meaning of the cause that generated them; their critically determine the priority to intervene and to restore the correct operative parameters.

At the end of the one-month lasting experimentation of the diagnostic method, a lot of information was collected and some important considerations about building's performances and strictly correlated user's energy behavior were made.

Firstly, only the number of the outliers occurred on each building's floor, during the specified period (March 2013), were counted reporting also the average and the maximum values of the relative indicators found (as shown in table 2), but in that way we have only a partial idea of the building's energy behavior.

Table 2: Number of faults, average and maximum value of the relative indicators on each floor.

Floor	Number of faults	Average value	Max value
0	102	0.716593173	0.857360006
1	57	0.764186715	0.852502757
2	72	0.799734897	0.902425976

For a more detailed analysis, considerations about which outliers occurred hour by hour unexpectedly or in a systematic way due to a known and justifiable reason were made.

As shown in table 3, referring to the ground floor, most of these events took place in the early morning (from 6 am to 8 am for cleaning activities) and at the end of the working day (from 16 to 18) corresponding to the rapid changes in power consumption due, respectively, to the massive switching on and off of the rooms. The number of these faults and the relative average and maximum value were also reported.

Electrical trend's irregularities, practically negligible, found during the night refer to slight fluctuations in the values of active power among samples closest each other.

Moreover, it can be recognized that with the indication of an outliers of 0.902 degree during the night, a temporary malfunction in the Wattmeter that wasn't able to work properly, acquiring and reporting only null values except one correct active power's sample of 0.4 kW.

After this period of validation, the FDD system has been put in real time operation and actually is working properly, collecting and analyzing data.

6 Conclusions

In this paper an innovative and automatic faults detecting system, which can be applied to commercial building networks is proposed. From the data acquired by sensors and through their aggregation with data fusion operations, dimensionless and normalized indicators referring to different kind of several measurements are

Table 3: Number of faults, average and maximum value of the relative indicators referring to the ground floor for different hours.

Ground Floor			
Hour	**Number of faults**	**Average value**	**Maximum value**
0	0	0	0
1	0	0	0
2	1	0.902425976	0.902425976
3	0	0	0
4	0	0	0
5	2	0.77934432	0.780269321
6	30	0.77454554	0.788900629
7	1	0.760844281	0.760844281
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0	0
14	0	0	0
15	0	0	0
16	0	0	0
17	1	0.787360283	0.787360283
18	33	0.816434209	0.842447717
19	2	0.831920717	0.839945571
20	2	0.864528445	0.866887408
21	0	0	0
22	0	0	0
23	0	0	0

defined in order to isolate faults and identify the possible reasons (ranked according their severity) generating them. Moreover this technique is also able to discriminate anomalous events that happen in an unexpected or in a systematic way and to calculate the level of plausibility that they can occur in relation to a specific cause of malfunction.

On the basis of this experimentation that has shown important results about the method's correctness and its reliability during real time operations with the building performance always under control, the activity will keep going on collecting data in order to have a more accurate framework of building's energy situation and eventually to provide operations of maintenance.

Future developments will include the study of new diagnostic rules applied to conditioning related to both electrical and thermal measurements. As is known, these aspects with the lighting one are the most important factors that affect energy consumption and continually require to be monitored.

Finally, this diagnostics method will be applied to a buildings network, no more only to a single one, with the precise aim at monitoring more structures thinking in a future that it would be adopted in a large scale, in small districts or cities.

References

[1] IEA (International Energy Agency), *Electricity/Heat Data for IEA Europe*, 2008.

[2] Katipamula, S. & Brambley, M.R., *Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems - A Review, Part I,* HVAC&R RESEARCH, Vol.11, Issue 1, January 2005.

[3] TAC by Schneider Electric, *Leading Techniques for Energy Savings in Commercial Office Buildings*, 2006.

[4] Ding, S.X., *Model-based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*, Springer, 2008.

[5] Khan, I., Capozzoli, A., Corgnati, S. & Cerquitelli, T., *Fault Detection Analysis of Building Energy Consumption Using Data Mining Techniques*, Energy Procedia 42, pp. 557–566, 2013.

[6] Lauro, F., Capozzoli, A. & Pizzuti, S., *Building Energy Consumption Modeling with Neural Ensembling Approaches for Fault Detection Analysis*, Sustainability in Energy and Buildings: Research Advances, KES Open Access LibrAry, Vol. 2, pp. 7–12, 2013.

[7] Kavuri, S.N., Venkatasubramanian, V., Rengaswamy, R. & Yin, K., *A review of process fault detection and diagnosis Part I: Quantitative model-based methods*, 2002.

[8] Simani, S., *Fault diagnosis of dynamic systems using model-based and filtering approaches*, 2008.

[9] Kavuri, S.N., Venkatasubramanian, V. & Rengaswamy, R., *A review of process fault detection and diagnosis Part II: Qualitative modes and search strategies*, 2002.

[10] Marzat, J., Piet-Lahanier, H., Damongeot, F. & Walter, E., *Autonomous fault diagnosis: state of the art and aeronautical benchmark*, 3rd European Conference for Aero-Space Sciences, EUCASS 2009.

[11] Kavuri, S.N., Venkatasubramanian, V., Rengaswamy, R. & Yin, K., *A review of process fault detection and diagnosis Part III: Process history based methods*, 2002.

[12] Ma, J. & Jiang, J., *Applications of fault detection and diagnosis methods in nuclear plants: A review*, 2010.

[13] Hyvfirinen, J. & Kirki, S., *Building Optimization and Fault Diagnosis Source Book*, International Energy Agency, Real Time Simulation of HVAC Systems for Building Optimisation, Fault Detection and Diagnosis, ANNEX 25, 1996.

[14] Dexter, A. & Pakanen, J., *Demonstrating Automated FDD Methods in Real Buildings*, International Energy Agency, Energy Conservation in Buildings and Community Systems, ANNEX 34, 2001.
[15] Palshikar, G.K., *Simple Algorithms for Peak Detection in Time-Series,* Proc. of 1st Int. Conf. Advanced Data Analysis, Business Analytics and Intelligence, 2009.
[16] Novàk, V., Perfilieva, I. & Mockor, J., *Mathematical Principles of Fuzzy Logic,* The Springer International Series in Engineering and Computer Science, Vol. 517, 1999.