Decentralized fault detection in building services by means of tensor decomposed qualitative models

Thorsten Müller-Eping, Nicolas Réhault, Tim Rist
Fraunhofer Institute for Solar Energy Systems ISE, Heidenhofstr. 2, 79110 Freiburg, Germany
E-mail: thorsten.mueller-eping@ise.fraunhofer.de

Abstract. Fault detection and diagnosis (FDD) methods are most of the time deployed in buildings as supervisory solutions on a management level or as cloud computing solution. The deployment of Internet of Things devices will enable to embed FDD methods as edge-computing solutions directly on subsystems like heat pumps or air handling units. This paper shows how qualitative models can be used for fault detection in building services and how tensor decomposition methods can enable their integration on decentralized subsystems as edge-computing solution.

1. Introduction

Fault detection and diagnosis (FDD) methods for buildings aim at detecting and identifying faults occurring in building services timely in order to trigger a corrective action and therefore to save energy or enhance comfort. Currently, these methods are hosted on the management tier of Building Automation Systems (BAS) or as Software-as-a-Service (SaaS) in cloud-computing based solutions. In this framework, data from sensors and actuators is transmitted over communication protocols like BACnet or KNX to the supervisory layer where it is processed, analyzed and where FDD results can be treated. This method is well adapted to legacy systems where devices at the edge of the system are not designed to embed FDD methods or do not provide enough computational resources. It offers also a high flexibility in data processing by using powerful and cloud-based computational resources and data storage capabilities. However, the transfer of time-series data with high time resolution lead often to latency and congestion issues in BAS networks and thus to data and time losses that impair the application of FDD methods. The integration of FDD methods on edge devices has the potential to enable data processing and analysis directly at the source where data is produced. FDD methods can thus be embedded for example on a heat pump or an Air Handling Unit (AHU) controller that can then communicate only the state (faulty/non faulty) and meta data of the detected fault of the subsystem to the management layer. Compared to the transmission of raw time-series data, the functional information can be drastically reduced as only condensed results of the FDD analysis are transmitted. Furthermore, the FDD analysis can be executed on different devices or nodes, which are independent from each other, and do not depend anymore on the reliability of the time-series data transmission over the network or on the availability of the SaaS as cloud solution. In a near future, Internet-of-Things (IoT) devices like sensors, actuators and machines will become ubiquitous in buildings and provide means to integrate analytical methods like FDD at the edge of networks, where the raw data is produced. One challenge is to adapt FDD methods, which can be computationally expensive, to small and low-cost edge-computing devices with low computational resources.
In this paper, we recall the basics of qualitative models in section 2. For the supervision of complex processes, qualitative models show limitations as both model complexity and storage amount increase rapidly with a rising number of system signals: the model suffers from “the course of dimensionality”. Reducing the complexity of qualitative models of large discrete-time systems like building services is the main challenge for making them applicable on edge-computational units. In section 2.4, we show how tensor decomposition methods can be applied to qualitative models and how they enable to reduce the dimensionality and computational resources. Then, in section 3 we provide two application examples of distinct complexity. First, we show the results of the implementation of a qualitative model for fault detection on a heating loop with a low system complexity. Second, a more complex model of an AHU heating coil is presented that demonstrates the added-value of model reduction by CP tensor decomposition for the integration of FDD methods on decentralized devices.

2. Qualitative models

2.1. Dynamical systems

A dynamical system is a process, where the output changes over time as a reaction of the input [15]. Accordingly to this definition, building services like air handling units (AHU), water loops or heat pumps are dynamical systems. Usually, dynamical systems have multiple inputs and outputs (MIMO systems) which are summarized in an input vector $u(k) \in \mathbb{R}^m$ and an output vector $y(k) \in \mathbb{R}^q$ (s. Figure 1).

![Figure 1: Dynamical system](image)

The inputs and outputs of building systems usually appear in form of discrete-time continuous-variable signals, which are also called time series whereby the integer variable $k \in \mathbb{N}$ denotes the discrete time. How the output of a dynamical system reacts on a given input depends on the so-called states, which are stock variables describing the amount of energy stored in a system and represented by the state vector $x(k) \in \mathbb{R}^n$. Usually, the goal of modeling dynamical systems is to describe their input-output behavior (I/O behavior), which can be realized by white, gray or black box models. For fault detection (FD) the relation between system input and output plays an important role, because a system that operates in a nominal condition shows another I/O behavior than a system that is subject to some fault. As shown in [4, 5], a wide range of model-based fault detection approaches are available and have been successfully applied to building systems.

In this paper, a nonlinear probabilistic black box modeling approach, a so-called qualitative model, is used for representing the dynamics of building services and to detect faults during their operation. Qualitative models are stochastic models which describe the behavior of dynamical systems only approximately. They have been proposed by [7, 9, 10, 14] who have shown, that they can be efficiently implemented for fault detection in systems with a small number of input, state and output signals. The base for qualitative modeling is a quantized system.

2.2. Quantized systems

A quantized system as shown in Figure 2 consists of a process as depicted in Figure 1 and the quantizers. The quantizers transform the discrete-time continuous-variable input, state and output signals of the process into discrete-time and discrete-variable signals, which are called quantized or qualitative signals. The quantized signal values are denoted in brackets $[\cdot]$. The right-hand side of figure 2 depicts the quantization of a continuous variable signal (black dots) into a quantized signal where the signal values at time $k$ are given by the symbols low, medium, high or very high or, alternatively by the integer numbers 1, 2, 3 or 4.
Despite the fact that the input \( u(k) \in \mathbb{R}^m \), state \( x(k) \in \mathbb{R}^n \) and output \( y(k) \in \mathbb{R}^q \) vectors are representing a number of \( m \) input signals, \( n \) state signals and \( q \) output signals, due to the quantization, their respective qualitative value is given by a single integer number. Thereby \( M \) denotes the number of qualitative inputs, \( N \) is the number of qualitative states and \( Q \) is the number of qualitative outputs. Further information can be found e.g. in [7, 14].

2.3. Stochastic automata as qualitative models
As qualitative model a stochastic automaton

\[
A = (U, X, Y, L, p_0)
\]

is used which consists of the set of qualitative inputs \( U := \{1, \ldots, M\} \), the set of qualitative states \( X := \{1, \ldots, N\} \), the set of qualitative outputs \( Y := \{1, \ldots, Q\} \), the vectorial initial probability distribution of the quantized states \( p_0 \in [0, 1]^N \) and the behavior relation \( L \)

\[
L(\bar{z}, w | z, v) = \Pr\left( [x(k+1)] = \bar{z} \left| x(k) = z, u(k) = v \right. \right). \tag{2}
\]

The behavior relation \( L \) describes the conditional probability that the system moves from a quantized state with qualitative value \( z \in X \) to a successor state \( \bar{z} \in X \) while receiving a quantized input with qualitative value \( v \in U \) and giving a quantized output with qualitative value \( w \in Y \) \( [11] \). The conditional probabilities \( L \) are identified by a black box method called stochastic qualitative identification which is given in [7]. For the model identification historical measurement data reflecting the nominal behavior of the underlying system is used. The conditional probability \( L(\bar{z}, w | z, v) \) is calculated for all combinations of qualitative values \( z, \bar{z}, v \) and \( w \) and stored as an element \( l(\bar{z}, w, z, v) = L(\bar{z}, w | z, v) \) of a four-dimensional, so-called behavior tensor \( [13] \)

\[
L \in [0, 1]^{N \times Q \times N \times M}. \tag{3}
\]

2.4. CP-decomposition of the behavior tensor
Qualitative models suffer from their large storage requirement. The number of values to be stored is dependent of the size of the behavior tensor in equation (3) and is given by \( \alpha = QN^2M \).

For the decomposition of the qualitative models and the reduction of the storage amount, we used a factorization method referred to as non-negative CP tensor decomposition \([3, 6, 12, 13] \). The simplified expression of the CP decomposition of the non-negative four-dimensional
behavior tensor $L \in [0, 1]^{N \times Q \times N \times M}$ is defined as a factorization into a weighting vector $\lambda \in \mathbb{R}_{\geq 0}^R$ and a number of four so-called factor matrices $X^{(1)} \in \mathbb{R}_{\geq 0}^{N \times R}$, $X^{(2)} \in \mathbb{R}_{\geq 0}^{Q \times R}$, $X^{(3)} \in \mathbb{R}_{\geq 0}^{N \times R}$, $X^{(4)} \in \mathbb{R}_{\geq 0}^{M \times R}$ that give the CP representation of the behavior tensor of the qualitative model

$$L \approx \left[ \lambda; X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)} \right].$$

The CP decomposition (4) is usually an approximation of the original tensor whereby the accuracy and the data storage requirements of the approximation depends on the so-called rank of the decomposition $R \in \mathbb{N}_{\geq 0}$. In summary, the concept of storage reduction is based on storing only the factor matrices $X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)}$ and the vector $\lambda$ instead of the full tensor $L$. For the CP representation of qualitative models, the number of values to be stored is given by

$$\alpha_{CP} = R \cdot (Q + 2N + M) + R.$$

2.5. Qualitative fault detection

The qualitative model $A$ in equation (1) is used as a qualitative observer. Based on the measured quantized input $[u(k)]$ and output $[y(k)]$ (I/O pair) at time $k$, the model gives a prediction of the conditional probability distribution of the quantized states which is expressed by a vector $p_z(k+1) \in [0, 1]^N$ (s. Figure 2). Observation algorithms for the full model (3) and the reduced model (4) can be found e.g. in [12, 13]. The detection of faults is based on a proof of consistence [14]. That means, as long as the qualitative model generates a state probability vector with non-vanishing probabilities for the measured I/O pair, the system is said to be in nominal condition and the I/O pair is called consistent with the qualitative model. On the opposite, if an I/O pair is inconsistent with the qualitative model, all components of the state probability vector are zero and the system is assumed to be in a faulty condition.

3. Application examples

In this section, we describe the implementation of qualitative models for fault detection on two application examples consisting of a heating circuit and a heating coil of an AHU. The objective was to demonstrate the possibility to embed complex FDD methods on small and low-cost computational units that can be directly integrated to the devices. The models have been implemented on a RASPBERRY Pi (RPi) microcomputer with a quad core 1.2GHz 64bit CPU and 1GB of RAM. In the first example, the microcomputer has been connected directly to the DDC station that controls the heating circuit over BACon/IP. The second example has been realized in an offline context, without connection with the real plant.

3.1. Fault detection in a heating circuit

This example shows the supervision of a simple heating circuit by a qualitative model which was implemented on a RPi environment. This use-case was generated during the project OBSERVE and is also published in the final report [1]. In this example, the qualitative model was trained by the use of measurement data (260 days) representing the nominal behavior of the heating circuit (s. Figure 3 right). Figure 4 shows the result of the qualitative model for a selected time.
range of 300 minutes. The upper part of the figures shows the probabilities of the quantized states for the state variable $T_{Water,Out}$. The different shades of gray represent the probabilities for each of the qualitative values of the state variable $T_{Water,Out}$, where a dark color depicts a high probability. The green dots are the measurement values. The signal below the figures indicates whether the system is in nominal ($b(k) = 1$) or faulty condition ($b(k) = 0$). As can be seen, in the interval $[44 \leq k \leq 111]$ the qualitative model does not generate any positive probabilities due to the inconsistence of the measured I/O pair and for the signal $b(k) = 0$ holds. This fault is related to a manual switch off of the heat generator due to an oil delivery [1].

Figure 4: State probabilities of the water outlet temperature and behavior signal [1]. One time step $k$ equals 1 minute.

As Table 1 shows, the qualitative model used for this example is of low complexity: the memory consumption is only 0.37 MB and model reduction techniques are not necessary.

Table 1: Model properties

| Qualitative inputs, states, outputs | Heating circuit | Heating coil |
|-----------------------------------|-----------------|-------------|
| $M = 216, N = 6, Q = 6$           | $M = 36, N = 64, Q = 64$ |
| Storage amount of (3)             | $\alpha = 46
d656$ values, 0.37 MB | $\alpha = 9,437,184$ values, 72 MB |
| Rank of decomposition             | $\alpha_{CP} = 15,700$ values, 0.4 MB | $R = 100$ |

3.2. Fault detection in an AHU heating coil

In this application example, the operation of a heating coil of a large AHU is supervised by a qualitative model. The measurement data of the AHU was determined within the European project CASCADE [2] and the qualitative fault detection results presented here are given in a similar form in [12, 13]. The qualitative model was trained with two month of measurement data representing the nominal behavior of the heating coil. After the training phase the qualitative model was applied to a set of new measurement data (3 weeks). For a selected time range, the results of the full model (3) are shown on the left-hand side of Figure 5. The right-hand side of the figure depicts the results given by the tensor decomposed qualitative model (4). As the figure shows, a fault based on a too low water inlet temperature in heating mode occurs during the time interval $\approx [2800 \leq k \leq 2930]$. As can be seen, both models show similar results, although in case of the decomposed model the number of values to be stored is reduced by a factor of more than 600 and the amount of stored data could be reduced from 72 MB to 0.4 MB (Table 1).

Of course, in this case, even the full model with 72 MB memory consumption could have been stored on the RPi environment. But keeping in mind that an AHU consists of many components (cooling coil, humidifier etc.) to be supervised, the storage requirements for a qualitative model representing the whole plant would exceed the capacity of the microcomputer. Furthermore, the higher the data storage requirements, i.e. the size of the behavior tensor (3), the higher the computing times. That is, the decomposed qualitative model (4) provides two main advantages: less storage capacity is required and the calculation times are significantly reduced [12].
Figure 5: State probabilities of the water outlet temperature and behavior signals. Results of the full model (left) and the decomposed model (right). One time step $k$ equals 30 minutes.

4. Conclusion
This paper shows that modern mathematical techniques like CP tensor decomposition can contribute to the integration of complex FDD methods like qualitative models on small computing devices. The model reduction significantly lowers the computational effort and thus enables the deployment of these method as edge-computing solution in future IoT-systems. With this, qualitative models can be directly integrated on buildings subsystems like air handling units.

References
[1] Boeck, M., J. Buchholz, H. Przybilla, and A. Wolfram (2019). “AP.A5”. In: OBSERVE - Optimierung und Betriebsführung komplexer Gebäudeenergieversorgungsanlagen. URL: [ob-serve.de](http://ob-serve.de).
[2] CASCADE - ICT for Energy Efficient Airports (2015). URL: [www.cascade-eu.org](http://www.cascade-eu.org).
[3] Cichocki, A., R. Zdunek, H. A. Phan, and S.-I. Amari (2009). Nonnegative Matrix and Tensor Factorizations. John Wiley & Sons, Ltd.
[4] Katipamula, S. and M. Brambley (2005). “Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems – A Review, Part I”. In: HVAC&R Research 11.1, pp. 3–25.
[5] Kim, W. and S. Katipamula (2018). “A Review of Fault Detection and Diagnostics Methods for Building Systems”. In: Science and Technology for the Built Environment 24.1, pp. 3–21.
[6] Kolda, T. G. and B. W. Bader (2009). “Tensor Decompositions and Applications”. In: SIAM Review 51.3, pp. 455–500.
[7] Lichtenberg, G. (1998). Theorie und Anwendung der qualitativen Modellierung zeitdiskreter dynamischer Systeme durch nichtdeterministische Automaten. Dissertation. Vortragsberichte VDI, Reihe 8: Mess-, Steuerungs- und Regelungstechnik, Nr. 686. Düsseldorf: VDI Verlag GmbH.
[8] Lichtenberg, G. and J. Lunze (1997). “Observation of Qualitative States by Means of a Qualitative Model”. In: International Journal of Control 66.6, pp. 885–903.
[9] Lunze, J. (1992). “Qualitative Modelling of Continuous-Variable Systems by Means of Nondeterministic Automata”. In: Intelligent Systems Engineering 1.1, pp. 22–30.
[10] Lunze, J. and J. Schröder (2001). “State Observation and Diagnosis of Discrete-Event Systems Described by Stochastic Automata”. In: Discrete Event Dynamic Systems 11.4, pp. 319–369.
[11] Lunze, J. (1998), “Qualitative Modelling of Dynamical Systems Motivation, Methods, and Prospective Applications”. In: Mathematics and Computers in Simulation 46.5-6, pp. 465–483.
[12] Müller-Eping, T. (2019). “Tensordekomposition qualitativer Modelle zur Fehlererkennung – Anwendung in der Gebäudeautomatik”. Unpublished PhD thesis, submitted 4th of Feb., 2019.
[13] Müller-Eping, T., G. Lichtenberg, and V. Vogelmann (2017). “Fault Detection Algorithms Based on Decomposed Tensor Representations for Qualitative Models”. In: IFAC-PapersOnLine 50.1: 20th IFAC World Congress, Toulouse, pp. 5622–5629.
[14] Schröder, J. (2003). Modelling, State Observation and Diagnosis of Quantised Systems. Lecture Notes in Control and Information Sciences 282. Berlin: Heidelberg: Springer-Verlag.
[15] Unbehauen, H. (2008). Regelungstechnik I - Klassische Verfahren zur Analyse und Synthese linearer kontinuierlicher Regelsysteme, Fuzzy-Regelsysteme. Wiesbaden: Vieweg + Teubner.