A Software Architecture for the Analysis of Energy- and Process-Data

Björn Kroll, Sebastian Schriegel and Oliver Niggemann
Fraunhofer IOSB-INA,
Application Center for Industrial Automation, Lemgo, Germany
Langenbruch 6, 32657 Lemgo, Germany
bjoern.kroll, sebastian.schriegel, oliver.niggemann@iosb-ina.fraunhofer.de

Stefan Schramm
Institut für Mikroelektronik- und Mechatronik-Systeme gemeinnützige GmbH,
Ehrenbergstr. 27, 98693 Ilmenau, Germany
stefan.schramm@imms.de

Abstract

This paper contributes a framework that helps to fulfill the requirements of the standards DIN EN 16247 and ISO 50001 by combining (i) a synchronized data acquisition, (ii) data integration, (iii) learning of normal behavior models and (iv) a implementation of an anomaly detection as prototype. Both standards require a reliable data acquisition and energy consumption analysis for implementing a certified energy management system. It shows that this framework meets the specifications of the standards by implementing a combined data acquisition and anomaly detection approach.

1. Introduction

Energy efficient machines are much more of interest for the industry since the introduction of the ISO 50001 and the DIN EN 16247 for small and medium enterprises. Some aspects of those standards are a reliable data acquisition and analysis, which leads to anomaly and suboptimal energy detection. A reliable Data Acquisition and analysis defines multiple aspects. Those are (i) a synchronized and time critical data acquisition on the sensor level (ii) Data merging from multiple synchronized sources and (iii) learning of normal behavior to implement anomaly detection or other types of data analysis.

The contribution of this paper is a new framework, which implements a time synchronized data acquisition from various sources, a semantic data interface using OPC UA and the detection of anomalies in discrete and continuous-value as outlined in Figure 1. This type of framework is currently not available for the automation industry that covers all of the above mentioned aspects. Other systems are able to cover the whole consumption of energy and resources, but they lack the automated model generation for industry machines and do not analyze the energy data for anomalies. Also, as they are covering whole production lines, they do not have the information of small module parts. As this Framework is adapted directly into the sensor Level and as it is using OPC UA for uniform data transfer, it is very suitable for this type of systems to cover the last step, from a complete factory to a single module in a machine.

This paper is organized as follows: This section continuous with the state of the art of each component used in the framework. Section 2 briefly introduce the framework with the learning and working phase. Section 3 illustrates the results and the experimental platform, which is used to verify the software behavior. The future work is described in Section 4. Finally Section 5 concludes this paper.

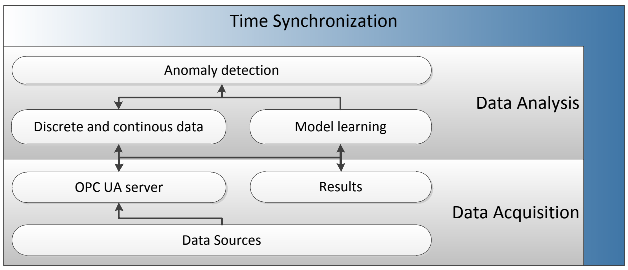

Figure 1. Framework Components.

Data Acquisition

Today's solutions for data acquisition in factory automation include different approaches. On the IO device and PLC level, IEC 61131-3 function blocks for communication are frequently used[9]. Above this level, different middleware projects have been initiated in the last few years, to provide a solution for distributed systems in automation applications; they can often be used to access PLC and MES/SCADA level information. A overview can be found in [3]. A number of projects is based on web services and its DPWS (device profile for web service). A overview about agent-based control and holonic manufacturing systems (HMS) can be found in [5]. Also Systems like OPC (Classic DA, HDA, etc.) or the OPC Unified Architecture are often found in combination with web service technologies. For energy data, either the normal (real-time) bus systems such as ProfiEnergy is used, or alternative the current IEC 61850 standard is used. The IEC 61850 Part 6, which is considered as the basic concept for an automation system in general, defines a Substation Configuration description Language (SCL).

Data Analysis

Learning timed automata is a rather new field of research, e.g. in [6]. Some of them use as well negative as positive examples. To include timing information Verwer[11] introduced a splitting operation which splits a transition if the resulting subtrees are different enough. Many industrial applications are hybrid systems, which cannot be sufficiently described with discrete event systems or strictly continuous models. In [6] a new algorithms for the learning of timed hybrid automata was introduced. In these approaches, the timed hybrid automata model the behavior of sub-systems like e.g. components of robotic systems. The division of the overall system into sub-systems is usually a-priori given by the user. The training of behavior models for the particular sub-systems can be performed in two steps [8]. In the first step, a timed automaton is trained by application of methods which are based on the above mentioned algorithms. This step requires no a-priori knowledge about the automaton structure and the number of automaton states. Similar states are merged in the second step of the automaton training.

Time Synchronization

To achieve a uniform time-base in today's distributed automation systems, several methods and protocols exist. E. g. for Ethernet based Networks the Precision Time Protocol (PTP, IEEE 1588) is able to provide synchronization of different nodes with accuracy in the range of nanoseconds while today's system-wide process data acquisition solutions only provide accuracy in the range of milliseconds [1]. Time synchronization in wireless sensor networks is quite different from synchronization in normal computer networks. Because of the limited resources of sensor nodes, computing complexity and energy consumption for those are an issue. Related works at this field are Flooding Time Synchronization Protocol (FTSP)[7] and Reference Broadcast Synchronization (RBS)[2]. There is also a NTP-like approach [4].

2 Framework Components

The framework is a combination of hard- and software. The data acquisition is able to use multiple hardware adapters depending on the hardware. The acquired data is stored in an OPC UA server and the anomaly detection itself is a software which runs on an external computer system. All modules are using OPC UA for data transmission, therefore it is possible to create a boundless vertical integration into a given industrial system.

To learn the normal behavior of machines and use it for anomaly detection, the requirements on data acquisition are the following (i) the energy and process-data must be acquired synchronous to the process and (ii) comprehensive over multiple processes. The bulk of today's industry machines is either using bus systems like PROFINET or similar. In [9] a synchronous data acquisition was introduced, this concept is able to cover the majority of machines. However the ISO50001 also claims that especially older machines near the end of their life-cycle has to arm with such data acquisition techniques. To cover those types of machines or to find a flexible approach to expand industry machines in the course of the associated project a wireless system was used. The main motivation of using a Wireless Sensor Network [10] for this purpose are the low installation costs. The sensor nodes measure the power consumption of the machine components and transmit it by radio. The physical network layer is based on the IEEE 802.15.4 standard. 6LoWPAN allows IPv6 communication with the nodes. The nodes are synchronized with the FTSP algorithm [7] and afterward the data is transmitted via a gateway-node to the OPC UA Server. The data is saved with a semantic data model on the OPC UA Server. This model features Value types, sensor location, descriptions and historic values in a most uniform way, to enable other system like MES systems for example to acquire it afterwards for other purpose. The framework outlines two operation phases. In the Learning Phase , the machine must operate at normal state and the software is learning the behavior. In the Operation Phase all data from the previous phase is used, to compare it with the actual data.

Learning Phase

The behavior models of the particular plant components are trained in this phase. For this purpose, the model (and its structure) is learned from the training data with as little a priori knowledge as possible. This approach avoids a complex manual configuration during the installation of the monitoring system. Therefore, the installation effort can be reduced so that anomaly detection approaches are even possible where rule-based approaches are infeasible. Furthermore the configuration errors and suboptimal configuration are prevented with this approach.

Operation Phase

The next step is to determine whether simulation predictions and measurements should vary significantly enough to form an anomaly. The anomaly detection is based on the model-based approach described before. Such an approach further requires little installation effort for new plants as automatic training methods can be applied. For the detection of anomalies there are several approaches that gives reliable results. On values, which represents a flow of fluids or cumulative values, the software is using a linear regression approach in consideration of each state of the learned model. When going to values that are representing raw energy consumption, there are some differences to consider. Energy-data has much sharper slopes

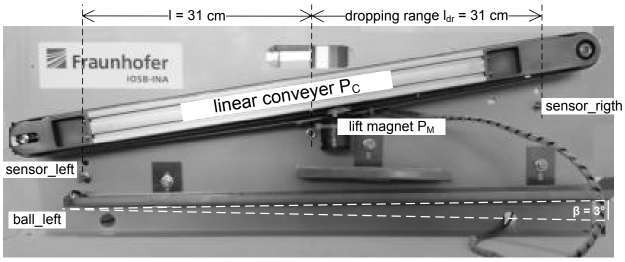

Figure 2. Experimental platform.

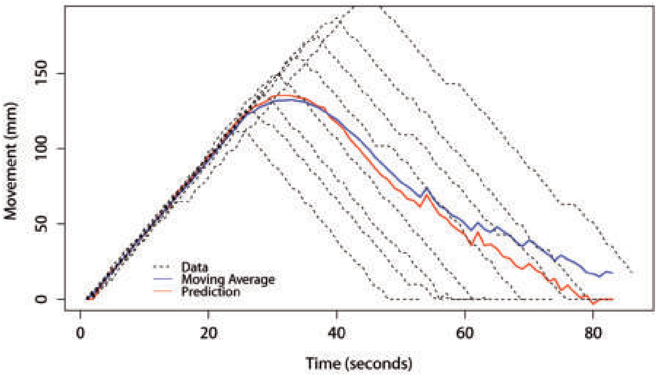

Figure 3. Learned movement data.

that make it unsuitable for a linear regression approach. In [12] the introduced Kalman-filter is used to recognize anomalies in energy data.

3 Proof of concept

Experimental platform

The construction of the experimental platform exists of an active transport process (linear conveyor) and a passive transport process (roadway). Figure 2 shows the physical process. The work piece is represented by a metal ball. The learned automaton from the experimental platform is outlined simplified in Figure 5. The process flow starts with lifting the metal ball at the bottom of the roadway by the magnet. The linear conveyer lifts up the metal ball to the dropping range and drops it. The transport process is finished when the ball and the magnet are back in the left position (bottom of the roadway) so that the next transport process can start. The energy consumption and the time duration for the whole process step depends on the handover between the transport processes. The handover between the roadway and the linear conveyor can not be influenced by the process controller. The handover between the linear conveyor and the roadway is represented by the dropping position. The transport process behavior depends on the dropping position, the angle of the roadways, the ball acceleration etc.

Results

The system is able to learn a model of the normal behavior as described in [6]. Additionally the automaton holds

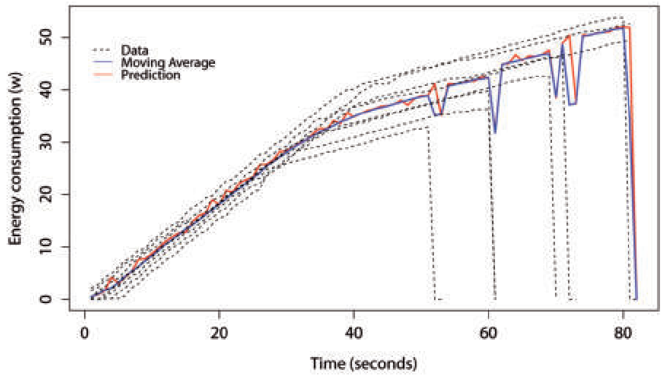

Figure 4. Learned energy consumption data.

two continuous data types: (i) the data from the movement of the magnet (Figure 3) and (ii) the accumulation of the energy used by the magnet when switched on (Figure 4). Both data types were modeled via the mechanism described in Section 2. The calculated adjusted R squared over the learned data is roughly above 0.8. This shows that the learned behavior equate with the original data. After the implementation of the outlined requirements in the next section, more detailed tests with the framework on other types of industry machines and a comparable measurement of the tests as F-measure are required.

4 Future Work

Noise handling

The learning algorithms or the data acquisition must cope with noisy training data. Furthermore, noise distributions may be trained in order to cope with disturbing noise in the operation phase. Model learning requires in most cases a compromise between different issues. A suitable trade off between model complexity, amount of training data and system noise must be considered.

Handling of simulation inaccuracies

The simulation models also come with a normal inaccuracy. This inaccuracy is mainly caused by the fact that the models are at least partially learned, i.e. they are an abstracted view onto the original observations. Since any abstraction mechanism sacrifices precision for generality, a specific degree of inaccuracy of learned models is to be expected and must be taken into consideration for the anomaly detection step.

5 Conclusion

This software architecture is able to acquire data from multiple sources and use it for anomaly detection from different data types. If industry machines do not use a bus system, it is possible to extend those machines with a wireless sensor system for data acquisition. The OPC UA integration allows a very scalable vertical configuration of

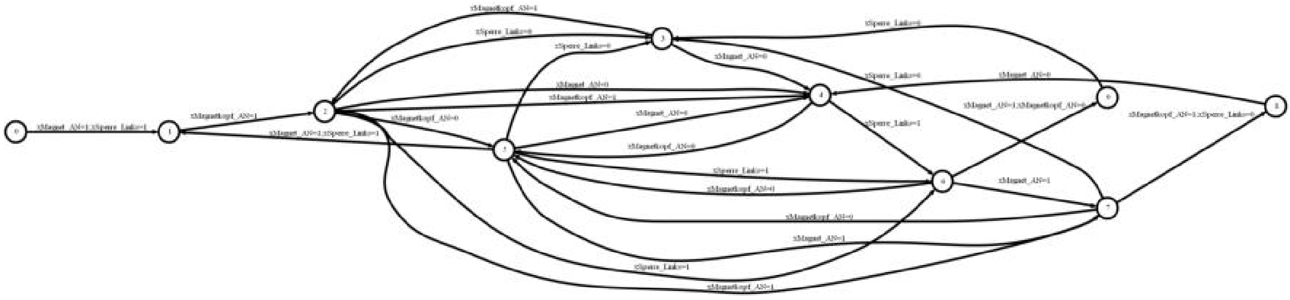

Figure 5. Learned automata

	DIN EN 16247	ISO 50001	Framework
historical data	X	X	X
energy data	X	X	X
other relevelant data	X	O	-
energy consumption	X	X	X
energy balance	X	O	X
energy consumption model	X	-	X
related energy consumption	X	X	-
predict energy consumption	-	X	X

X needed/covered, O partially, - not needed

Table 1. Specification requirements compared to the Framework

all inherited systems and if there are other components, they can be easily integrated.

Hybrid automata are a reasonable choice for anomaly detection in industrial systems, as they are able to cope with the requirements described earlier. When going to the ISO 50001 and the DIN EN 16247 this software architecture is able to provide multiple task in the process of certification and well beyond, as outlined above in Table 1. When adapted to the concept of energy performance outlined in the ISO 50001, this architecture is able to cover the aspects of energy input, energy consumption and energy efficiency.

6 Acknowledgement

This work was funded by the BMWi (German Federal Ministry of Economics and Technology), funding code IGF 16419, as a part of the project "Anubis: Analyse und Überwachung des Energieverbrauchs in der Verfahrenstechnik und der Produktionstechnik" (2012-2014)

References

[1] S. Cavalieri and G. Cutuli. Performance evaluation of opc ua. Technical report, University of Catania, Faculty of Engineering, Department of Computer Science and Telecommunications Engineering, 2010.

[2] J. Elson, L. Girod, and D. Estrin. Fine-grained network time synchronization using reference broadcasts. *SIGOPS Oper. Syst. Rev.*, 36(SI):147–163, Dec. 2002.

[3] W. Emmerich, M. Aoyama, and J. Sventek. The Impact of Research on the Development of Middleware Technology. *ACM Transactions on Software Engineering and Methodology*, 17(4):19:1–19:48, 2008.

[4] S. Ganeriwal, R. Kumar, S. Adlakha, and M. Srivastava. Network-wide time synchronization in sensor networks. Technical report, Networked and Embedded Systems Lab (NESL), Electrical Engineering Department, UCLA, 2003.

[5] P. Leitao. Agent-based distributed manufacturing control: A state-of-the-art survey. *Engineering Applications of Artificial Intelligence*, 22:979–991, 2009.

[6] A. Maier, O. Niggemann, A. Vodencarevic, R. Just, and M. Jaeger. Anomaly detection in production plants using timed automata. In *8th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*. Noordwijkerhout, The Netherlands, Jul 2011.

[7] M. Maróti, B. Kusy, G. Simon, and A. Lédeczi. The flooding time synchronization protocol. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*, SenSys '04, New York, NY, USA, 2004. ACM.

[8] O. Niggemann, B. Stein, A. Vodencarevic, A. Maier, and H. Kleine Buning. Learning behavior models for hybrid timed systems. In *Twenty-Sixth Conference on Artificial Intelligence (AAAI-12)*, Jul 2012.

[9] F. Pethig, B. Kroll, O. Niggemann, A. Maier, T. Tack, and M. Maag. A generic synchronized data acquisition solution for distributed automation systems. In *17th IEEE International Conf. on Emerging Technologies and Factory Automation ETFA'2012*, Krakow, Poland, 2012.

[10] T. Rossbach, A. Schreiber, S. Engelhardt, W. Kattanek, and E. Chervakova. Long-life sensor modules - challenges for hardware design and MAC algorithms. In *Proceedings 54. Internationales Wissenschaftliches Kolloquium*. TU Ilmenau, September 2009.

[11] S. Verwer, M. de Weerdt, and C. Witteveen. A likelihood-ratio test for identifying probabilistic deterministic real-time automata from positive data. In J. Sempere and P. Garca, editors, *Grammatical Inference: Theoretical Results and Applications*, Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010.

[12] S. Windmann, S. Jiao, O. Niggemann, and H. Borcherding. A stochastic method for the detection of anomalous energy consumption in hybrid industrial systems. In *INDIN*, 2013.