

# Big Data and Machine Learning for the Smart Factory—Solutions for Condition Monitoring, Diagnosis and Optimization

Alexander Maier, Sebastian Schriegel and Oliver Niggemann

## 1 Introduction

In this chapter, Machine Learning solutions for Cyber-Physical Systems (CPSs) in a Smart Factory are outlined using production plants as an example. The increasing complexity of production plants is still a present issue within the industry. Due to an increasing number of product variances, an increasing product complexity and increasing pressure for efficiency in a distributed and globalized production chain, production systems are evolving rapidly: They are becoming modular, can be parameterized and comprise a growing set of sensors. These new features are needed from a production point of view, but they overstrain the capabilities of most users: Too many parameters to optimize and too complex inter-dependencies need to be understood.

According to Lee [8], CPSs “are complex engineering systems that integrate physical, computation and networking, and communication processes. CPSs can be illustrated as physical devices, objects, equipment that is translated into cyberspace as a virtual model. With networking capabilities, the virtual model can monitor and control its physical aspect, while the physical aspect sends data to update its virtual model”. The physical world and the cyber world are decoupled by a data-to-feature level, where the data is processed to deliver information in cyberspace at different

---

A. Maier · S. Schriegel · O. Niggemann (✉)  
Fraunhofer IOSB-INA, Langenbruch 6, Lemgo, Germany  
e-mail: oliver.niggemann@hs-owl.de

A. Maier  
e-mail: alexander.maier@iosb-ina.fraunhofer.de

S. Schriegel  
e-mail: sebastian.schriegel@iosb-ina.fraunhofer.de

O. Niggemann  
Institute Industrial IT–inIT, Langenbruch 6, Lemgo, Germany

levels of abstraction, beginning with information about the condition of a machine component leading to the overall throughput and quality risk of a manufacturing line. Therefore, CPSs aim to develop user functions for removing this burden from user capabilities, e.g. for self-diagnosis and self-optimization.

CPSs often tackle the challenge of self-diagnosis and self-optimization in a new and innovative way: Instead of relying on human expertise and additional engineering steps formalizing the necessary knowledge, self-diagnosis and self-optimization is performed in a data-driven way, i.e. by applying ideas from the field of Big Data. Data such as sensor and actuator signals will be stored and analyzed both locally in sensor and control devices, and also globally in Big Data platforms. Based on these data, algorithms will learn models of the normal system behavior. Hence, these models will be used to detect deviations from normal workflow as anomaly (e.g. condition monitoring), find root causes for such anomalies (e.g. diagnosis), calculate optimal parameters and predict the future system behavior (e.g. predictive maintenance, energy optimization).

In this chapter, the following topics are covered:

1. Application of Big Data platforms in CPSs (Sect. 2).
2. Capture relevant data in sufficient quality (Sect. 3).
3. Machine learning algorithms to abstract system observations into such models (Sects. 4 and 5).
4. Usage of models for condition monitoring, predictive maintenance and diagnosis (Sect. 4).
5. Usage of models for the automatic system optimization, e.g. energy consumption (Sect. 5).
6. Application scenarios from the SmartFactoryOWL and from industrial applications (Sect. 6).

## 2 Big Data in CPSs

The term “Big Data” is not clearly defined in the literature. In common parlance, it describes any data analysis dealing with a huge amount of data. This is not a formal definition, but hit the core of the problem. According to the International Data Corporation (IDC), the volume of data is doubling every two years (<http://idcdocserv.com/1678>). In year 2014 the world-wide amount of data was about 4.4 Zettabyte. In 2020, the amount will be ten times bigger, around 44 Zettabytes. Werner Vogels, Amazon’s chief technology officer, told the BBC (<http://www.bbc.com/news/business-26471415>): “You can never have too much data—bigger is definitely better. The more data you can collect the finer-grained the results can be.” For companies, this means diverse challenges. One of the challenges is the interconnection of the data: A huge amount of the data volumes exists without any connection to other data. Therefore, a challenge to Big Data concepts is to connect data to gain competitive advantages and savings, and to form new business.

Although the best-known applications of Big Data refer to customer data in internet (Google, Facebook, Amazon and others), Big data in manufacturing holds a similar potential. Every used kilowatt hour, from each produced screw to each car, even each switching of a proximity sensor and each change of a temperature sensor generates raw data that holds an enormous potential if it is stored and provided for intelligent analysis. The acquisition, handling and analysis of these data present several challenges.

Due to the heterogeneous data sources providing a different time base, the data has to be synchronized. Several solutions have been proposed for the time synchronization (e.g. IEEE 1588) as well as for the transmission (e.g. Profinet) and semantic annotation of the data (e.g. OPC UA).

The analysis of historical process data throughout the product lifecycle requires new architectures and platforms for dealing with the enormous volume of data, and the variation and speed of data. The data analysis drives the conventional data acquisition and storage to their limits, that is why Big Data platforms are needed.

One of the most used Big Data platforms is the Hadoop ecosystem. A typical Big Data platform is structured as follows: The CPS is connected via a standardized interface (e.g. by OPC UA) with a Hadoop ecosystem. Hadoop itself is a software framework for scalable, distributed computing. The process data is stored in a non-relational database (HBase), which is based on a Hadoop Distributed File System (HDFS). In addition to HBase, a time series database OpenTSDB serves as an interface for data analysis. This database provides simple statistical functions such as mean values, sums or differences that are usually not available in a non-relational data storage. The interfaces of OpenTSDB or Hadoop thus enable data analysis directly on the storage system. Because the algorithms can process the data locally, the volume of a historical dataset does not need to be loaded into a single computer system. Via a web interface, both the data and the calculated results can be visualized (e.g. using Grafana).

Figure 1 illustrates the architecture of a Big Data platform using a Hadoop ecosystem.

While classical condition monitoring is mainly about monitoring of specified signals and checking the crossing of some predefined thresholds to find anomalies, Big data is about discovering the data and finding relationships in the data without specifically looking for something special. This requires the use of novel

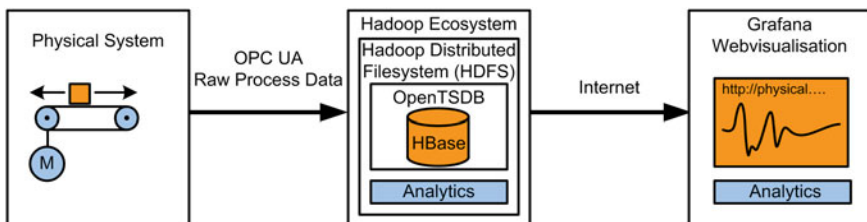


Fig. 1 Example big data architecture

approaches and algorithms. A main issue is the high amount of data which possibly is stored in the cloud and has to be transferred via internet. The limited bandwidth would lead to a delayed analysis. Therefore, the in-memory algorithms are required that perform (a part of) the analysis directly on the data hosting server. Further approaches for the speed-up of the analysis point towards the direction of parallelization. An example mentioned here is the MapReduce technology, which divides the problem into several sub-problems that can be parallelized to gain a computation speed increase, more details will follow in Sect. 4.2.

### 3 Requirements and Challenges to Data Quality

Smart Factories are built from many assets, consisting of a multitude of different components. The Reference Architecture Model Industry 4.0 (RAMI 4.0, for details see: [www.plattform-i40.de](http://www.plattform-i40.de)) includes the hierarchically ordered assets Product, Field Device, Control Device, Station, Work Centers, Enterprise and the Connected World. All of these physical assets generate or apply data over the application's life cycle. The assets are connected by an Integration and Communication Layer to the functional information processing. The Integration and Communication Layer technologies influence the quality of the information about the physical assets. Low sampling rates, unsynchronized and patchy data do not represent the exact physical state/behavior of the asset, and valuable information is lost. Often, small changes in the time series of signals or energy consumption of specific actors caused by wear of the asset point to a future failure and maintenance requirement. Modern industrial communication systems can fulfill these high performance requirements regarding the data acquisition:

- PROFINET IRT offers sampling rates down to 32,15  $\mu$ s.
- OPC UA includes semantic data annotations, historical data access and the vertical communication including the Internet and Cloud solutions.
- Time synchronization with PTP—Precision Time Protocol (IEEE 1588) enables factory wide data synchronization with a precision of 100 ns and less [15].

But due to different reasons like device availability, costs, all-over system integration and interoperability, these technologies today are often not integrated in one application. Many Machines are communicating with PROFINET, but they are not synchronized to other machines and they do not have historical data access. Often a specific gateway copies the data from the PROFINET system to the OPC UA address space—this works, but due to the performance of the Gateway and due to the performance of the (non-real time capable) OPC UA interface, the data quality is lost: The sampling rate decreases; due to jitter and latencies the synchronization accuracy of the data decreases and often data is lost completely due to bandwidth problems. For this reason, the handling of poor process data is a basic feature of machine data analytic solutions.

On top of these challenges in data quality and data acquisition, it depends on the asset (Product, Field Device, Control Device, Station, Work Centers) what kind of model formalism is useful for further functional data analytics: Discrete manufacturing work centers like straight-line machine sequences can be represented using state machines; continuous processes like rotating machines (generators) need modeling formalisms like clustering.

## 4 Condition Monitoring and Diagnosis

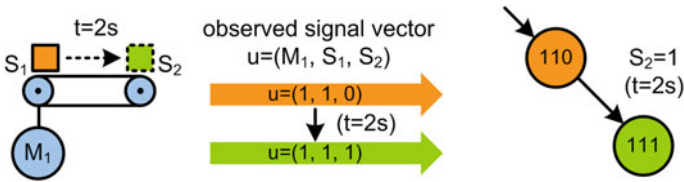
### 4.1 *Anomaly Detection Using Identified Hybrid Timed Automata*

Modern production plants in manufacturing industry are mostly programmed using state machines. The advantage is that the meaning of state machines is easy understandable to humans since they think in a similar way. Therefore, the usage of state machines for anomaly detection in CPSs is obvious. In addition, state machines can be identified automatically from observations. Since production plants are mostly dependent on time, this factor has to be considered as well resulting in a timed automaton as used formalism.

While the identification of untimed stochastic automata is an adequately studied research area [1, 3, 11, 16], the identification of discrete-time behavior in CPSs is a new research area, which arose in the last few years. Various algorithms have been developed to study such timed automata, including RTI + [17] and HyBUTLA [10], which also identifies a hybrid timed automaton. As extension, Markov machines can be used if the probability of the next event depends on a number of previous events, an example is given in Ray [13]. An example for the identification of Bayesian networks is given in de Campos [5].

Most of the aforementioned automaton learning algorithms use the state merging approach for the identification of the resulting automaton. Based on the initially constructed prefix tree acceptor, which comprises all prefixes of recorded production cycles, all pairs of states are checked for compatibility in an iterative manner. If two states are compatible according to a specified criterion, they are merged. After the compatibility of all pairs of states has been computed and the compatible states have been merged, the resulting automaton represents the normal behavior of the considered CPS.

The automaton identification algorithms, which are based on the state merging approach mentioned above, work in an offline manner, i.e. all recordings have to be stored and be available for the state merging procedure. The algorithm OTALA, introduced in Maier [9] is the first algorithm for the identification of timed automata in an online manner. It is based on the assumption that each state in the observed CPS can be represented by a signal vector and each signal vector corresponds to one state in the final automaton. Using OTALA for identification, the observations



**Fig. 2** Online timed automaton learning algorithm (OTALA)

do not have to be stored since each observation is included directly into the final automaton. The basic functionality of the algorithm is shown in Fig. 2.

Basically, the identification of timed automata does not require expert knowledge (except of some parameters which are mostly equal for the different use cases). The states and transitions with the corresponding events are identified by the algorithm. However, structural information about the system (e.g. asynchronous subsystems) is needed to identify one separated model for each asynchronously running subsystem.

The identified automata are finally used for anomaly detection. Much work has been done in this area (c.f. Sarkar et al. [14] for a comprehensive overview). In Vodencarevic et al. [18], the algorithm ANODA (Anomaly Detection Algorithm) has been presented which takes the identified timed automaton as input and detects anomalies in the CPS. The behavior of the CPS can be considered as a path through the identified automaton. Whenever an event is observed in the CPS that cannot be depicted by the automaton in a certain state or if the observed timing does not fit into the identified time range, an anomaly is signaled.

## 4.2 Identification of Behavior Models Using Map/Reduce Technology

Due to the enormous amount of data, conventional data analysis methods are overcharged. In the context of Big Data, extended analysis methods are required to cope with the data flood. As described in the previous sections, the behavior models can be identified from data collected for the system and its components in normal, fault-free operation, using algorithms such as (Hy-) BUTLA or OTALA. These algorithms can be used in Big Data context as well, however, the data handling has to be adapted. Since CPSs create a huge amount of data, the MapReduce technology can be applied to parallelize the identification of behavior models. One possible solution is presented in [19]: OTALA is applied for model learning of the discrete states, and quadratic regression models (QRM) are generated for continuous behavior. Both model learning algorithms have been parallelized applying the MapReduce technology. The MapReduce version of OTALA allows to distribute the workload on  $|I|$  nodes, and therefore a speedup is achieved as each transition

$T$  can be processed in parallel to the REDUCE function. For the MapReduce version of QRM, distribution of workload on  $|S|$  nodes is possible by processing the states  $S$  of the automaton in parallel to the REDUCE function. Furthermore, online algorithms have been proposed, which efficiently handle novel observations to update the models that have been created from large historical data sets.

### 4.3 Condition Monitoring in Continuous Processes

Condition Monitoring and Anomaly Detection in continuous processes require different approaches than in discrete manufacturing processes. As stated before, Big Data is more than just detecting the crossing of a predefined threshold of a specified signal. It requires unsupervised machine learning techniques, which autonomously find relationships in the data and use these for condition monitoring, for instance.

Clustering approaches are commonly used approaches for these tasks. However, as the data becomes high dimensional, clustering methods (such as DBSCAN, e.g. see Zhu [20]) reach their limit since density is more difficult to define in high dimensional space. Therefore, dimension reduction methods are needed to pre-process the data. After reducing the dimensionality of the input data, conventional clustering algorithms can be used to create a behavior model. In [6] self-organizing maps are used to reduce the input dimension and to generate a two dimensional map which visualizes the observed process. A further possibility is the Multidimensional Scaling (MDS, e.g. in Borg and Groenen [2]) which is a set of techniques from the mathematical statistics. The goal is the arrangement of objects and their relation to each other. The farther the objects are from each other, the more dissimilar they are and the closer they are, the more similar they are. There are thus collected information about pairs of objects to identify them to metric information about objects.

The Principal Component Analysis (PCA) is a commonly used method to reduce the dimensionality of the input data. The method basically assumes that features with a low variance provide a small contribution to the final model and therefore can be neglected. To minimize the information loss, the PCA computes the principal components, which are new features that are mostly uncorrelated. The dimensionality of the dataset is then reduced by using the most relevant (see step six in the subsequent procedure for calculating the PCA) of the principal components to describe the dataset. When the PCA is used for visualization purpose, the three first principal components are selected creating a three-dimensional figure. The negligence of the remaining principal components is possible, because most of the variance of the original dataset, i.e. the information, is represented by the first few principal components [7].

The principal components are determined by performing the following six steps:

1. Center and scale the data matrix  $\bar{X}$ .
2. Compute covariance matrix of data.

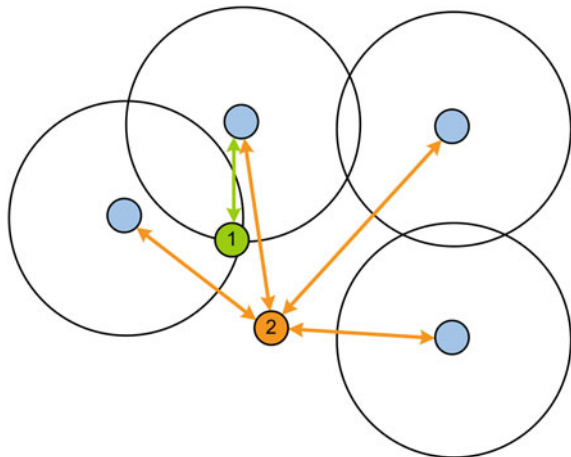
3. Calculate the eigenvectors covariance matrix.
4. Sort eigenvectors decreasingly by corresponding eigenvalues.
5. Resulting eigenvector matrix is used as transformation/rotation matrix.
6. Choose  $x$  highest dimensions as reduction for PCA.

The PCA is an effective method for dimension reduction. For example, reducing a dataset from 30 signals to 3 principal components (reduction of 90 %), which allows keeping 80 % of the information (corresponds to a loss of 20 %), is a quite effective way of dimensionality reduction.

After the model has been identified using PCA and DBSCAN, it is used for anomaly detection. The anomaly detection is based on the assumption denoted by Chandola et al. [4]: “Normal data instances lie close to their closest cluster centroid, while anomalies are far away from their closest cluster centroid.” Therefore, the following two steps are performed for each observation:

1. Transform data from input space to feature space: The transformation matrix (see step 6 from above) is used to transform each incoming observation from the original input space to the reduced feature space.
2. Compute distance to model (see Fig. 3): To determine whether the new observation belongs to the normal operation phase, distances to each point of the model is checked. Several distance functions can be used (e.g. Euclidean or Mahalanobis Distance). The core distance is determined using the Marr Wavelet. In contrast to the Euclidean Distance, the Marr Wavelet crosses the abscissa, the corresponding values are used as core distance.
3. Check thresholds: If the calculated distance exceeds a predefined threshold (identified core distances from the previous step), an anomaly is signaled. As it can be seen in Fig. 3, the first observation was classified as normal, whereas the second observation is identified as an anomaly.

**Fig. 3** Distance of observations to normal behavior model





The PCA can also be used for predictive analytics. For this, two core distances have to be used. Crossing the first threshold, a warning is given, crossing the second threshold, an anomaly is signaled.

## 5 System Optimization

Another application of smart services in manufacturing is the self-optimization of industrial processes. Optimization can be carried out regarding different influencing variables (e.g. time or speed), but in this section we focus on the optimization of the energy consumption. The goal is the analysis and improvement of the performance and efficiency of a manufacturing plant, leading to an optimized operation. Due to increasing energy prices, a special focus for this smart service is the optimization of the energy efficiency in industrial automation systems.

Typically, the optimization of energy efficiency is a manual process, performed by experts of the plant by exchanging old and inefficient drives against new and efficient drives. This is a useful and necessary step, however, it still requires man power and finance investments. Further methods require a manual time planning of the production steps in the Manufacturing Execution System (MES) to obtain an energy-efficient process, or special energy controllers that are typically located at the energy meter and monitor the trend of the energy consumption. If the trend points to unwanted levels, the controller switches off equipment, based on certain priorities and other rules. Typical time periods are in the range of 15–30 min [12].

When it comes to active methods to implement real-time optimization at second or millisecond intervals, only little research has been carried out, because many applications require process parameters to be rapidly adapted to changing operating conditions due to their process dynamics. An active optimization of the process requires models that are able to predict the future process behavior for various parameter combinations in order to determine the optimum process parameters.

Especially in transportation and logistics applications, electrical drive systems are the most voracious energy consumers. Therefore in this application use case, optimization is performed regarding to the energy consumption and load management of electrical drive systems. Figure 4 illustrates an exemplary plant setup of two drives.

The main optimization problem in this application is the minimization of the overall energy consumption of the two converters ( $E_{L,WR}$ ) and motors ( $E_{L,mot}$ ) for time instances  $k = 0 \dots n-1$ . Additionally, the energy  $E_g(k)$  which is fed back into the internal DC link (difference between regenerative and motive power) has to be minimized. To achieve this, the energy consumption and the energy feedback are weighted with the factor  $\lambda$  according to the primary usage of power feedback (consumption at braking resistances, regenerative feedback to the mains, and temporary storage in capacitors). Altogether, the optimal driving speeds  $v$  of the conveying system can be obtained as solution of the following optimization problem:

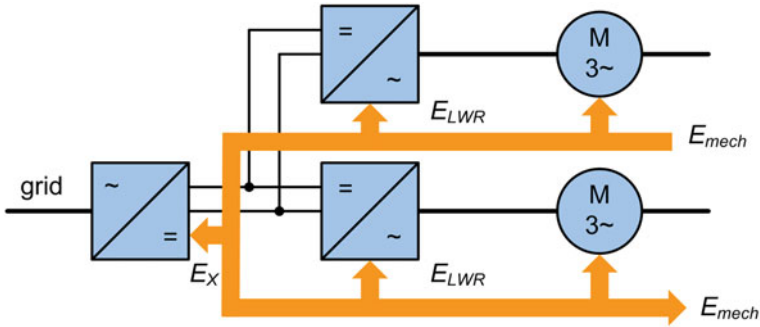


Fig. 4 Two motors connected to a DC circuit

$$v = \operatorname{argmin} \left\{ \sum_{k=0}^{n-1} \sum_{c=0}^{n_c-1} y_c(k) + \lambda \sum_{k=0}^{n-1} E_g(k) \right\}$$

Further information regarding the optimization constraints of the energy flows and the sequences of movements (starting position, end position, speed limits, and positioning time) can be found in [19]. The obtained optimization problem is a mixed integer quadratically constrained problem (MIQP). It can be efficiently solved with standard methods. Evaluation results for energy-efficiency optimization of movements in conveying systems verify energy savings of around 7 %. The common optimization of the two drives leads to a motion profile in which the first drive is slowed down just during the acceleration phase of the second drive. In this particular application case, the complete regenerative energy of the first drive has been recovered as input for the second drive. Thus, the automatic coordination of motor speeds prevents undesired energy feedback.

## 6 Smart Services and Applications

The described technologies, data-driven condition monitoring and optimization of machines or production lines are the basic feature for new technical services (also known as Smart Services), like remote machine operation and new business models. Today, the most usual life cycle of a machine is: The machine builder constructs and sells the machine to the producing company. The producing company operates the machines and produces their products. But the machine-operator is not the overall machine expert (that is the machine builder). Because of this reason, machines are not often driven with the best operation parameters: best product quality, high output, save material and energy, ecology. Summarized: the total production costs are high.

A data-driven condition monitoring and optimization of machines enables two basic features: The machines operate more independent from human operators than before, and the remote operation with machine experts and process experts becomes possible. Due to this, the producing (product) company does not have to be the operator any more. The machine builder could fulfill the role of the machine operator. With this the business model can be transformed: The machine itself is no longer sold to the producing company, it remains the machine builder. Production capacities are leased to the product company. With this, the producing company can concentrate to their main competencies (products) and save investment cost. With this, the machine builder implements a complete new business model, and can use his know-how in their own machine for (1) the construction of new machines and (2, new) for efficient operation of machines.

Fraunhofer evaluated the technical potential of data-driven Smart Services and Applications like condition monitoring and optimization in production lines at various companies like Miele, Wesergold, Audi and Deutsche Windtechnik. These examples are described below. Problem statements for each application are given and the data-driven Smart Service solution is explained.

Miele is a manufacturer of domestic appliances, commercial equipment and fitted kitchens, based in Gütersloh, Germany. Miele is owner-managed company, founded in 1899. The production of dish washers for the very cost-sensitive consumer market has to be as much efficient as production time (machine output), material and energy. Today, process, production and product experts at Miele are running process and machine optimization based on their specific product and machine knowledge. This guarantees maximum product quality, which is the most important objective at Miele. But for an efficient production, the machine optimization has to involve aims like a high machine output, material and energy savings. Often, these are conflicting objectives that cannot be solved by a specific expert. Because of this reason, Fraunhofer evaluated a data-driven Smart Services for Miele. The machine signals of specific production processes were acquired and learned using the machine learning algorithm OTALA. The result was a timed discrete automaton. The automaton showed the time variance for specific process states. Based on this knowledge, better parameters for a higher output were identified.

Tönsmeier is a German company with more than 70 locations, 3000 employees, 1100 vehicles and 30 processing facilities in the waste management industry. The headquarter is located in Porta Westfalica, Germany. Tönsmeier was founded in 1927 by Karl Tönsmeier. In the processing facilities, Tönsmeier is sorting waste with the objective of recycling. The availability of sorting machines is critical because the input of the machine (waste) is often very different and causes failures like plugging and stoppage of the whole sorting machine. With the aim of the prediction and preventing of the plugging problem, Fraunhofer evaluated the performance of data-driven real-time condition monitoring for Tönsmeier.

At the automobile manufacturer Audi, Fraunhofer tested machine learning algorithms with the objective of functionality checks in new production systems. The production systems at Audi are very complex. In the start-up phase of new systems, a lot of time is needed for failure detection and system tests. By optimizing this start-up phase, Fraunhofer evaluated a data-driven assistance system that collects process data from the PROFINET communication of machines, and learns the machine behavior as a discrete state machine. The state machine is a reduced visualization of the complex machine behavior. This simple visualization is used by Audi process engineers searching problems and helps to program the PLCs, and it is also used for acceptance and certification tests of the machines at the end of the start-up phase. This Smart Service is a machine start-up assistance system that helps decreasing start-up phases and increasing the PLC programming quality.

In a research project, Fraunhofer and Resolto implemented a condition monitoring systems that is able to predict wear in wind power plants. The objective was to detect even small changes in processes before they mature into a system error. The system was tested in a wind power plant of the project partner Deutsche Windtechnik. All data of the wind turbines (as sensor values, energy data, weather data, etc.) were collected and analyzed by the self-learning system to model normal behavior of the process. By the help of an adapter and the industrial communication protocol OPC UA, the wind turbines were connected to the Internet. The process data were transmitted to a central server making data analysis. The analysis algorithms were applied to detect discrepancies between the real-time data and the identified model. The models were generated by a principal component analysis.

## 7 Summary and Outlook

One of the biggest challenges in Big Data is the generation of added value from data obtained in CPS.

In this chapter, we presented some algorithms that address the challenges of Big Data and automatically learn models of normal behavior. The identified models are used for condition monitoring, anomaly detection and predictive analytics. The main benefit of the proposed approaches is that only little expert knowledge is required. The behavior models are largely identified automatically, as well as the parameters needed for anomaly detection and predictive analytics.

A further benefit is the possibility of self-optimization. Based on the identified behavior models and given restrictions, new moving profiles in a storage can be calculated.

Various case studies proved the applicability of the presented machine learning approaches for condition monitoring, anomaly detection and optimization.

## References

1. Adenis P, Mukherjee K, Ray A (2011) State splitting and state merging in probabilistic finite state automata. In: American control conference, San Francisco, CA, USA, pp 5145–5150
2. Borg I, Groenen P (2003) Modern multidimensional scaling: theory and applications. *J Educ Meas* 40:277–280. doi:10.1111/j.1745-3984.2003.tb01108.x
3. Carrasco R, Oncina J (1994) Learning stochastic regular grammars by means of a state merging method. In: Carrasco RC, Oncina J (eds) Grammatical inference and application. Lecture notes in computer science. Springer, Heidelberg, pp 139–152
4. Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. *ACM Comput Surv* 41(3), Article 15
5. de Campos LM (2006) A scoring function for learning bayesian networks based on mutual information and conditional independence tests. *J Mach Learn Res* 7:2149–2187
6. Frey CW (2008) Diagnosis and monitoring of complex industrial processes based on self-organizing maps and watershed transformations. In: IEEE international conference on computational intelligence for measurement systems and applications
7. Jolliffe IT (2002) Principal component analysis. Springer, New York
8. Lee J (2015) Smart factory systems. *Informatik Spektrum* 28(3):230–235
9. Maier A (2014) Online passive learning of timed automata for cyber-physical production systems. In: The 12th IEEE international conference on industrial informatics (INDIN 2014)
10. Niggemann O, Stein B, Vodencarevic A et al (2012) Learning behavior models for hybrid timed systems. In: Twenty-sixth conference on artificial intelligence (AAAI-12), pp 1083–1090
11. Oncina J, Garcia P (1992) Inferring regular languages in polynomial update time. In: Bunke H (ed) Advances in structural and syntactic pattern recognition. Machine perception and artificial intelligence, vol 5. World Scientific, Singapore
12. Palensky P, Dietrich D (2011) Demand side management: demand response, intelligent energy systems, and smart loads. *IEEE Trans Ind Inf* 7:381–388
13. Ray (2004) Symbolic dynamic analysis of complex systems for anomaly detection. *Sig Process* 84(7):1115–1130
14. Sarkar S, Sarkar S, Ray A (2013) Data-enabled health management of complex industrial systems, fault detection: classification. Techniques and role in industrial systems. Nova Science Publishers, New York
15. Schumacher M, Wisniewski L, Jasperneite J et al (2013) Node to node synchronization accuracy requirements of dynamic frame packing. In: 2013 international IEEE symposium on precision clock synchronization for measurement, control and communication (ISPCS 2013)
16. Thollard F, Dupont P, de la Higuera C (2000) Probabilistic DFA inference using Kullback-Leibler divergence and minimality. In: Proceedings of the 17th international conference on machine learning. Morgan Kaufmann, San Francisco, pp 975–982
17. Verwer S (2010) Efficient identification of timed automata: theory and practice. Ph.D. thesis, Delft University of Technology
18. Vodencarevic A, Kleine Büning H, Niggemann O et al (2011) Identifying behavior models for process plants. In: Proceedings of the 16th IEEE international conference on emerging technologies and factory automation (ETFA 2011), pp 937–944
19. Windmann S, Niggemann O, Stichweh H (2015) Energy efficiency optimization by automatic coordination of motor speeds in conveying systems. In: IEEE international conference on industrial technology (ICIT 2015)
20. Zhu L, HeF, Tong Y et al (2015) Fault detection and diagnosis of belt weigher using improved DBSCAN and bayesian regularized neural network. *Mechanika* 21(1):70–77