ZESZYT NAUKOWY
Wyższa Szkoła Zarządzania i Bankowości w Krakowie

mgr inż. Jakub Pizoń
Lublin University of Technology, Phd Student at Mechanical Facutly
jakub.pizon@pollub.edu.pl
dr Tomasz Cieplak
Lublin University of Technology, Department of Organisation of Enterprise
t.cieplak@pollub.pl

REMOTE MONITORING OF PRODUCTION RESOURCES WITH THE USE OF MODERN STREAM AND MACHINING LEARNING TOOLS

Introduction

Today's world of production is on the edge of another new technological revolution. This revolution takes different names but one of the most popular is so-called Industry 4.0[1]. That name was attached to it in Germany and become one of main aims for German economy. Therefore, it is intensively implement in German industry but also worldwide. Adequately as previous production revolutions, revolution of Industry 4.0 is caused by new discoveries and transformations in technology.

What is good to mention is that the key change that Industry 4.0 brings is not strictly connected to some production based innovation or theories but rather to intensive development of information technology and their expansion into different fields. Therefore actual wheel of revolution are informatics technology solutions which are increasingly adapted to needs of technological environment and technological processes. This is particularly visible in perspective of manufacturing machines that deliver more and more advanced products and semi-finish products controlled by highly specialized software that can automatically work parts, work offsets, depending on the program code.

Significant fact is that the trend of high specialization of key production machines is also accompanied by increasing possibilities to identify individual components of the production process. Therefore, traceability is not only provided by bar code stickers (which is still one of the dominant ways), but also by RFID tags or more sophisticated solutions such as beacons (electronic tag that use Bluetooth Low Energy technology). This entire process is fascinating in its meaning but also it leads to the provision of production data that is essential for carrying out various databased analyses like Exploratory Data Analysis (EDA) or Machine Learning

[1] *Recommendations for implementing the strategic initiative INDUSTRIE 4.0*, April 2013,http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Material_fuer_Sonderseiten/Industrie_4.0/Final_report__Industrie_4.0_accessible.pdf, access date: 25.06.2017.

(ML).

Because of that its important what manufacturing tools we have but also what datasets those tools can deliver. Then what kind of software we have and if it has ability to processed gathered data in efficient way. Answering to these questions is extremely important, basically from the economic perspective. Doing data analyses, it is possible to find significant patterns, models, solutions that can save noteworthy sums of money. Thus, the data obtained is not only strictly technical, but is also the economic dimension of production enterprise. Even though analytical model based on data is some sort of forecast and it contains some errors. Nevertheless, they can indicate critical issues and be basis for business decisions. Moreover, they can also indicate which elements needs to be strictly controlled, monitored and repaired to get economic added value.

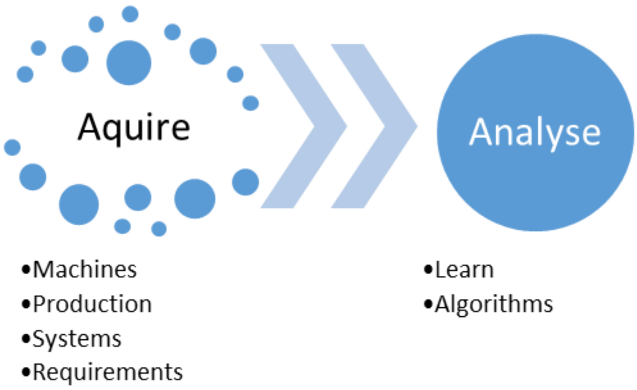
It is also significant that, by means of cost-optimized, increased specialization and centralization of production resources, modern systems must remain flexible. To achieve this great aim efficient system of control should be applied. To apply such system required is to have data about the current state of production. Information acquired from data will is necessary to identify state of the machinery to precisely monitor production and be able to plan it. Thus, obtaining information is crucial. To provide information, it is necessary to obtain data on the components of the production system in a brief time, in real time or close to real. The way to get this result is to process production data streams and machine learning. By processing streaming data, it is possible to obtain insight on selected aspects of production using appropriate models and visualization techniques. Consequently, the data, along with the relevant production context, allow you to obtain information in close-to-real time. What can be the basis for an almost immediate decision to allow direct reaction. A reaction that foresees failure or detection of system bottlenecks. On the other hand, machine learning provides a way to bring intelligence to production solutions. Even though the description seems simple – this is not a trivial task and therefore is subject to a lot of research, machine learning tools projects conducted by scientific centers but also IT companies and automotive industry.

The result of these aspirations is that in the era of flexible production systems and centralization of production capabilities increase role of machine learning systems that perform the task assigned and produce models that can be used for analysis and management optimization.

1. Remote monitoring of production resources

Modern production management and monitoring systems operate within specific operational positions and areas. Different machines, different production methods, different systems occur in production. All these forms lead to the delivery of the product, which is to have the right quality, meet the right standards, economical to the end customer. Due to the substantial number of these requirements, the organization of the manufacturing process is one of the key factors in meeting the requirements. To meet these requirements, it is necessary to perform management functions such as monitoring, control and verification that all requirements are fulfilled. To deliver those functions two steps needs to be executed. First is to gather data about the process and second is to learn from data gathered as shown on Figure 1.

Figure 1. Acquire and analyses production data



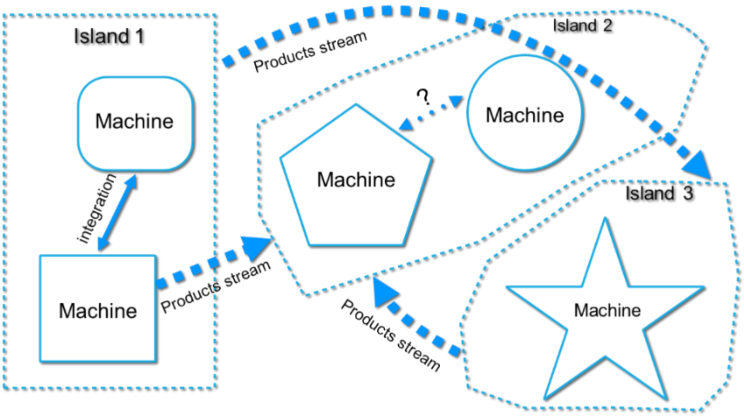
Source: Own work.

That is necessary to set goals, identify solutions and increase quality. Because of that special production systems are developed. Nevertheless, there are organized in so called technological islands.

1.1.Technological islands

Technological islands (Figure 2.) can be defined as areas in which production systems (production sockets, individual machines, production lines, branches or factories) generate data but are not exchanged with external systems. Monitoring and control systems create dedicated and hermetically sealed architectures of systems - implemented without the possibility of exchanging data in an open format. Because of that, many non-integrated data sources are produced. Data is not exchanged electronically but there is flow of intermediates and products between islands which are the subject of production. Thus, as far as it is possible to implement monitoring functions in individual islands, then production monitoring in the flow perspective is a new type of task. That makes it difficult or even impossible to analyze and identify key factors of performance – to notice significant changes – in real time. That is a cause of significant risk that companies need to cover, even though they use advance IT software to support manufacture processes.

Figure 2. Production islands



Source: Own work.

Despite so many differences, it is important to remember that what connects all the production is data. Data, regardless of the form in which they are, reproduce the technological process. Therefore, to minimize the risks associated with lack of monitoring and control, it is necessary to use the appropriate technologies to obtain data. What's more, they will automatically provide their analysis and learning.

1.2.Data acquisition and analytics

To conduct monitoring appropriate control models, we need the data of the machine or process under investigation. Acquisition of these data is a prerequisite for further action. In the case of a production environment, the acquisition of data largely depends on the type of machine park. For CNC machines or PLCs, it is possible to obtain data from the controller's internal memory. Nevertheless, in the case of information on the amount of work in progress, collecting the relevant data is a challenge.

However, with the advent of the Industrial Internet of Things (IIoT), the predictions of maintenance parameters gain more and more attention in the industry. First and foremost, IoT data collection and processing technologies that are considered mature enough to generate, transmit, store, and analyze all kinds of data in batches or in real time[2].

In case of monitoring and control data acquisition is needed to address problems that involve both high operational risk due to unexpected failures and limited insight into the causes of problems in complex business environments. Most of these problems can be addressed in the form of questions such as:

- What is the probability that a machine will crash soon?
- What is the remaining life of the equipment?
- What are the causes of the failure and what maintenance should be done to solve these problems?

By collecting data and analytics in this dimension, it is possible to use the predictive machine learning methods to:

- Reduce operational risk and increase return on assets,
- Reduce unnecessary maintenance and control maintenance costs,
- Reduce inventory costs by reducing inventory levels related to repair and maintenance activities,
- Identify the relationship between the various maintenance issues.

Moreover, by implementing an analytical solution it is possible to identify key performance indicators for the enterprise. That will indicate the state of health of the machine park or what is the estimated value of remaining life of assets. Depending on these indicators, it is possible to propose maintenance activities as well as an indication of estimated order dates for interchangeable parts[3].

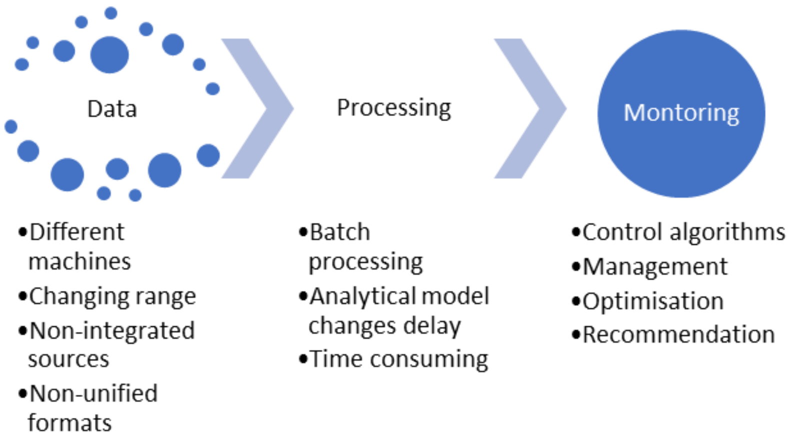
Therefore, with the development of the machine park, it is also necessary to develop monitoring and data processing systems. So far, the dominant concept in this dimension is Computer Integrated Manufacturing (CIM), but CIM itself assumes the vertical integration of production control systems[4]. Hierarchical structure causes that data before it is processed and

[2] Pizoń J., Lipski J., *Manufacturing Process Support Using Artificial Intelligence*, Applied Mechanics and Materials, Vol. 791, pp. 89-95, Sept. 2015.
[3] Pizoń J., *Koncepcja wdrożenia technologii „Internetu rzeczy" w systemie logistycznym przedsiębiorstwa*, Systemy Logistyczne Wojsk, nr 43, 2015.
[4] Poyser T.D., *CIM - COMPUTER INTEGRATED MANUFACTURING, A STRATEGIC MODEL AND APPLICATION OF CURRENT TECHNOLOGIES*, AEG Westinghouse Industrial Automation Corporation, Factory Automation Systems Division, P.O. Box 490, Pittsburgh 1990.

reaches the management layer on one hand passes time, and on the other hand not all data are relevant from the perspective of monitoring processes (Figure 3.). Moreover, in this dimension, monitoring systems perform supervisory functions through total control over the manufacturing process. What's more, in terms of the large dimension of production as well as the lack of integration between individual dimensions of production processes, is a problem as well as a hindrance to data processing. What is more, it requires the provision of appropriate IT architectures that will provide adequate computing power.

Moreover, CIM systems are based on batch processing. This means that even data obtained from a production machine is analyzed only when the database model of the monitoring program is found. The analysis of this data is already carried out by queries sent to the data model. It is only after these processes that the relevant system processes the data and allows to indicate the monitored aspects. In addition, even if the data is cumulative to the form of the warehouse, significant delays in data analysis will continue to be significant.

Figure 3. Hierarchical processing



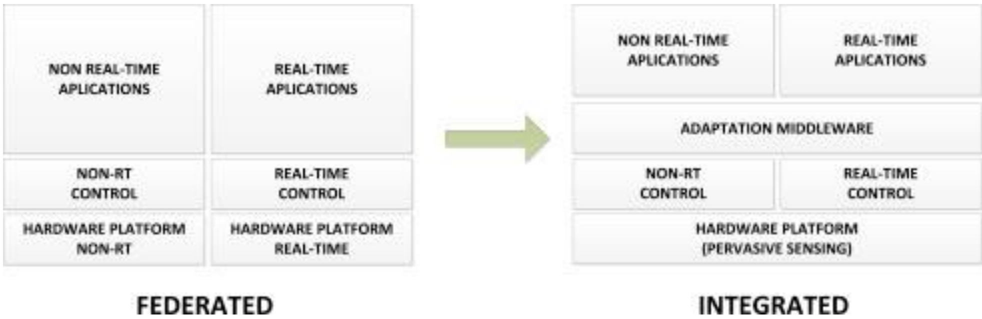
Source: Own work.

Because of profound change in manufacturing CIM systems evolve towards cyber-physical systems (CPS called. Cyber-Physical Systems), and cloud production (Cloud Manufacturing, CM). Cyber-physical systems implement strict integration layer calculations and physical processes[5]. Most often in the form of embedded systems and network monitoring and controlling physical processes operating in a feedback loop, where the physical processes are the source of data for calculating a control signal objects.

[5] Lee E., Cyber Physical Systems: Design Challenges, Electrical Engineering and Computer Sciences University of California at Berkele, January 2008.

Hence, there is a second dimension of IIoT integration that CPS implementations provides. Such solutions, in addition to hierarchical integration, allow for horizontal integration (Figure 5.) and the implementation of monitoring and management functions. That on the other hand leads to evolution from traditional federated model of production to integrated CPS systems (Figure 4.)[6]. That process is highly supported by stream processing that provides basics for fast and effective monitoring of manufacturing process.

Figure 4. Evolution from traditional federated systems to integrated CPS



Source: Bansal K., Sanisha, Comparative Study of Open Source Processing Frameworks for Analysis of Big Data, International Journal of Advanced Research in Computer Science, Volume 8, No. 5, May – June 2017.

Stream processing systems compute data as it enters into the system. Its processing model is different from the batch processing. Stream processor define operations to be applied to data item as it enters the system instead of defining operations to apply to an entire dataset.

The dataset is unbounded in stream processing systems. It means that total dataset is defined as the amount of data that has entered the system so far. The working dataset is limited to a single item at a time. Processing is event-based and results are immediately available and will be updated continuously as new data arrives.

Stream processing systems have the capability of handling large amount of data, but they can process only one record (true stream processing) or very few (micro-batch processing) items at a time. Analytics, server or application error logging, and other time-based metrics are some fields where stream processing best fits[7].

Therefore, there is a strong need to analyze ongoing data to pick the most important attributes in accordance with the manufacture logic to provide complete management information.

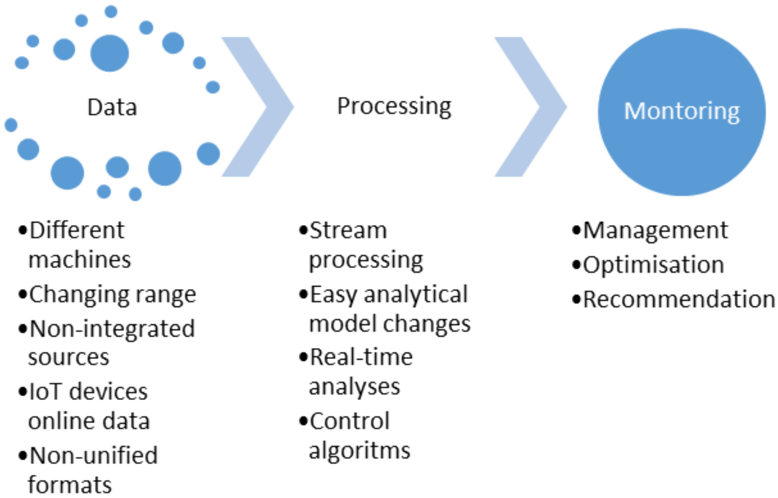
---

[6] Lee E., Cyber Physical Systems: Design Challenges, Electrical Engineering and Computer Sciences University of California at Berkele, January 2008.

[7] Lee J., Bagheri B., An Kao H., *A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems*, Manufacturing Letters, Volume 3, 2015, Pages 18-23.

That can be base for effective analyses performed with use of machine learning and increase quality and lower the costs of quality. This opens a very large perspective for deployments of continuous monitoring and analysis of streams of production data with use of stream computing.

Figure 4. Horizontal processing



Source: Own work.

The key is that remote monitoring allows one place to collect all production data and there, through appropriate algorithms, central control of the production solution and centrally controllable. For this reason, it is important that monitoring at the level of individual products - individually - allows for a precise indication at what time the product will be processed and transported. Conducting such analysis at such a low level poses significant challenges, but in the long run may provide the basis for large scale savings. Moreover, it is worth noting that due to the electronic form of data and the broadband Internet it is not required that these machines be in one location but in several various locations and there can be some analytics using machine learning.

2. Machine learning tools

In view of the changes in the production environment associated with the industrial revolution. Tools are needed to convert the received data into useful management information. Therefore, meaningful information must be inferred from the data. Currently, there are several tools and methodologies available for the data to information conversion level. In recent years, extensive focus has been applied to develop these algorithms specifically for prognostics and health management applications[8].

That is why, tool that can provide decision support system in face of cognition and use of CPS is machine learning. Machine learning: is an artificial intelligence technique that uses automated algorithms to churn through data sets more quickly than data scientists can do via conventional analytical modeling.

Machine learning is closely related to big data analytics that uses advanced analytic techniques to discover hidden patterns from big data. It applies the data mining, predictive analytics and machine learning tools to sets of big data that can be structured, unstructured and semi-structured data. It involves examining large data sets, to discover hidden patterns, unknown correlations, market trends, customer choice and other useful information that can help organizations make efficient business decisions.

Analyzing production data using a machine learning cloud computing environment allows to solve business issues, and more specifically to answer the questions like "What is the probability that a device will crash due to a component failure?" The indicated question can be obtained by means of multicriteria classification and learning algorithm to create a predictive model that learns from historical data collected from machines. Currently machine learning is more and more provided in machine learning as a service and provide possibility to build analytical model online. Moreover, machine learning as a service offers the distinct advantage of scalable machine resources as and when they are needed. Among available solutions it is always good to use, tools selected below:

- **Microsoft Machine Learning Studio** features a library of time-saving sample experiments, R and Python packages and best-in-class algorithms from Microsoft businesses like Xbox and Bing. Azure ML also supports R and Python custom code.
- **Amazon Machine Learning** is a service that makes it easy for developers of all skill levels to use machine learning technology. Amazon Machine Learning provides visualization tools and wizards that guide you through the process of creating machine learning models without having to learn complex ML algorithms and technology. Amazon Machine Learning makes it easy to get predictions for application using simple APIs, without having to implement custom prediction generation code, or manage any infrastructure. One-year free access is available.

[8] László Monostori, *Cyber-physical Production Systems: Roots, Expectations and R&D Challenges*, Procedia CIRP, Volume 17, 2014, Pages 9-13.

- **BigML** is a cloud based machine learning platform with an easy to use graphical interface. It also provides simple mechanisms to incorporate predictive models into production applications through its REST API. The platform combines supervised learning (to build predictive models), unsupervised learning (to understand behavior), anomaly detection (used in fraud detection), data visualization tools (scatter-plots and Sunburst diagrams) and many mechanisms for exploring data. BigML is a pragmatic, low cost, easy to use platform for building powerful predictive models. A free version is also available.
- **DataRobot** is a cloud based machine learning service that takes much of the grunt out of predictive model building. It automatically searches for the best features, selects the most appropriate algorithms, test the model and provides an API for model deployment. It takes best-in-class algorithms from R, Python, H20, Spark, and other sources, and employs text mining, variable type detection, encoding, imputation, scaling, transformation, and automated feature engineering. The infrastructure is massively powerful for rapid processing, and big data is well supported with certification on Cloudera Enterprise 5. A free community version is available.
- **IBM's Watson Analytics** offers predictive analytics and data visualization, and a conversational type interface. It automatically does the hard math to show the most relevant facts, patterns and relationships. A free version is offered with limitations on data volumes[9].

Choosing the right tool depends on the specificity of the task being analyzed. Therefore, before starting to formulate a solution, it is recommended to test several solutions and to critically evaluate the solutions available.

Summary

This article elaborates that integration of advanced production sites can be accomplished only with a use of data processing and analytics. Therefore, questions arise how such data is transmitted and processed is still a large area for systematization, implementation and proposal of innovative solutions. Therefore, this article refers to that how to effectively acquire and

[9] *Machine Learning as a Service Platforms*, http://www.butleranalytics.com/free-machine-learning-service-platforms/, access date: 25.06.2017.

process data.

The article characterizes and identifies aspects of the changes in contemporary production environments. It presents successively changes related to the process of organization of production processes related to the fourth industrial revolution as well as to the progressive degree of computerization of production systems. The article outlines the solution and context in which they can be applied to composing IT solutions for modern production environments. The article is just one of first steps to verify theses of use of additional information technology paradigms in the field of manufacturing and in consequence the development of modern manufacturing technologies.

Literature

[1] Bansal K., *Sanisha, Comparative Study of Open Source Processing Frameworks for Analysis of Big Data*, International Journal of Advanced Research in Computer Science, Volume 8, No. 5, May – June 2017
[2] Bordel B., Alcarria R., Robles T., Martín D., *Cyber–physical systems: Extending pervasive sensing from control theory to the Internet of Things*, Pervasive and Mobile Computing, Volume 40, September 2017
[3] László Monostori, *Cyber-physical Production Systems: Roots, Expectations and R&;D Challenges*, Procedia CIRP, Volume 17, 2014, Pages 9-13,
[4] Lee E., *Cyber Physical Systems: Design Challenges*, Electrical Engineering and Computer Sciences University of California at Berkele, January 2008.
[5] Lee J., Bagheri B.,An Kao H., *A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems*, Manufacturing Letters, Volume 3, 2015, Pages 18-23
[6] *Machine Learning as a Service Platforms*, http://www.butleranalytics.com/free-machine-learning-service-platforms/, access date: 25.06.2017
[7] Pizoń J., *Koncepcja wdrożenia technologii „Internetu rzeczy" w systemie logistycznym przedsiębiorstwa*, Systemy Logistyczne Wojsk, nr 43, 2015
[8] Pizoń J., Lipski J., *Manufacturing Process Support Using Artificial Intelligence*, Applied Mechanics and Materials, Vol. 791, pp. 89-95, Sept. 2015
[9] Poyser T.D., *CIM - COMPUTER INTEGRATED MANUFACTURING, A STRATEGIC MODEL AND APPLICATION OF CURRENT TECHNOLOGIES*, AEG Westinghouse Industrial Automation Corporation, Factory Automation Systems Division, P.O. Box 490, Pittsburgh 1990.
[10] *Recommendations for implementing the strategic initiative INDUSTRIE 4.0*, April 2013, http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Material_fuer_Sonderseiten/Industrie_4.0/Final_report__Industrie_4.0_accessible.pdf, access date: 25.06.2017

Abstract

The article characterizes and identifies aspects of the changes in contemporary production environments. It presents successively changes related to the process of organization of production processes related to the fourth industrial revolution as well as to the progressive degree of computerization of production systems. The article outlines the solution and context

in which they can be applied to composing IT solutions for modern production environments.