Toward Digital Twin for Cyber Physical Production Systems Maintenance: Observation Framework Based on Artificial Intelligence Techniques

Farah Abdoune1, Maroua Nouiri2, Pierre Castagna2, and Olivier Cardin2

1 Manufacturing Engineering Laboratory of Tlemcen (MELT), B.P N° 119, Tlemcen, Algeria
2 LS2N, UMR CNRS 6004, Université de Nantes, IUT de Nantes, 44 470 Carquefou, France
Maroua.nouiri@univ-nantes.fr

Abstract. Manufacturing Systems are considered complex engineering systems given the large number of integrated entities and their interactions. Unplanned events and disruptions that can happen at any time in real-word industrial environments increase the complexity of manufacturing production systems. In the fourth industrial revolution (so called Industry 4.0), the industrial sector is rapidly changing with emerging technologies like Cyber-Physical Production System (CPPS), Internet of Thing (IoT), Artificial Intelligence (AI), etc. However, the efficiency and reliability of these systems are still questionable in many circumstances. To address this challenge, an observation framework based on AI techniques aimed at elaborating predictive and reactive planning of the maintenance operations of CPPS is proposed in this paper. The proposed tool aims to improve the system’s reliability and helps the maintenance supervisors to adjust maintenance decisions. In order to assess the performance of the proposed tool, a case study on an industry-type learning factory is considered. A proof of concept shows the efficiency of the framework.

Keywords: Predictive maintenance · Reactive maintenance · Cyber-physical production system · Artificial intelligence

1 Introduction

Industrial companies today face two important problems. Customer demand is increasingly diverse, while customers are more and more demanding. At the same time, globalization implies significant competition. Faced with these challenges, industries search to improve the efficiency, reliability and availability of their services to be more competitive. Many studies indicate that the maintenance service and related activities have a direct impact on the efficiency of the production [1].

In fact, a good maintenance strategy reduces significantly the operating costs of the concerned systems and increases their reliability and overall availability to undertake operations. Maintenance activities aim to restore an item for correction or to achieve
better status \cite{2}. The lack of knowledge about the production systems, its equipment and the associated processes complicate more the management of these systems. However with the developments of Information and Communications Technologies (ICT) and the rise of the fourth industrial revolution, the focus on the maintenance of Cyber-Physical Production Systems (CPPS) is increasing. Several definitions have been proposed for CPPS, mostly related to varying contexts. According to \cite{3}, Cyber-Physical Production Systems are defined as “systems of systems of autonomous and cooperative elements connecting with each other in situation dependent ways, on and across all levels of production, from processes through machines up to production and logistics networks, enhancing decision-making processes in real-time, response to unforeseen conditions and evolution along time”.

Our research target in this work is the maintenance of CPPS. Recently, many techniques have been proposed in order to help managers, supervisors and operators to optimize maintenance decisions. An observation-based predictive maintenance framework for CPPS is proposed. Our work exploits the concept of CPPS and uses Industrial Internet of Things (IIoT) technologies to deploy an intelligent tool for predictive and reactive maintenance. The proposed observation-based predictive maintenance framework is based on real-time data acquisition and analysis through Artificial Intelligence (AI) techniques to detect and treat dysfunctions.

The rest of this article is organized as follows: a review of recent works showing the evolution of CPPS maintenance is given in the next section. Section 3 describes the details of the proposed framework. Section 4 discusses the results of our experimentation on a case study based on a learning factory. A conclusion and some future directions are given in Sect. 5.

2 Literature Review on Maintenance

The maintenance strategies are classified in two main groups: reactive and preventive \cite{4}. In the reactive category, the maintenance activity is triggered by an occurrence of a failure. In opposition, the preventive category aims at avoiding failure occurrence.

In order to avoid the significant downtime and repair costs due to a classical corrective maintenance, the manufacturers are more and more interested in the predictive maintenance strategy \cite{5}, which is based on continuous measurements to detect faults and anticipate problems. In \cite{6}, methods and tools related to predictive maintenance in manufacturing systems are reviewed and an integrated predictive maintenance platform is proposed. In \cite{7}, a review on simulation-based approaches is made, which have been widely used in the maintenance context. In these study, the behaviour of the system is reproduced and simulated.

Recently, AI techniques and Big Data applications provide technical support for the efficient development of manufacturing systems by accurately and timely data collection, data analysis, data processing, root-cause identification, and deriving valuable insights for maintenance improvement \cite{8}. In literature, there have been several reviews on the role of AI for the maintenance of manufacturing systems \cite{9}.

Machine learning (ML) is widely used in condition monitoring, fault prediction and predictive analytics. ML techniques are a branch of AI methods based on the use of
huge amounts of data to learn and to identify patterns [8]. In [10], they proposed a concept model for proactive decision support system based on real-time predictive analytics, designed for maintenance of a cyber-physical systems in order to minimize its downtime. A Hierarchical Modified Fuzzy Support Vector Machine (HMFSVM) is proposed in [11] to understand the trends of vehicle faults. This method is compared with commonly used approaches like logistic regression, random forests and support vector machines. A reference architecture based on deep learning for CPS is proposed in [12]. The concept for a CNC machine utilized in shop floor is explored.

The Centre for Intelligence Maintenance System (IMS) created a Watchdog Agent Technology - an approach for product performance degradation assessment and prediction, for modeling and decision making with human interaction [13]. This technology includes time domain analysis, Principal Component Analysis (PCA), Fuzzy Logics System (FLS), Logistic Regression (LR), Artificial Neural Network (ANN), Bayesian belief networks and Support Vector Machines (SVM) [14].

An Intelligent System for predictive maintenance (SIMAP) [15] has been developed for real time diagnosis of industrial processes based on neural networks that detect anomalies. The fuzzy logic method is used to provide behaviour modelling of a maintainer experience integrated into an intelligent maintenance system [16]. To estimate failure degradation of bearings and to predict failure probability, LR has also been used in combination with relevance vector machine (RVM) [17].

The various applications of ANN applications in fault risk assessment and early fault detection analysis have been reviewed with examples of their usage in predictive maintenance cases [18]. SVM has been used for fault diagnosis of automobile hydraulic brake system [19]. Recently, new methods based on hybridization between supervised and unsupervised learning techniques have been developed; an example is the root cause analysis and faults prediction for intelligent transportation systems (ITSs) based on the coupling of K-means Algorithm and ANN [20]. The method was tested on the Train Door System at Bombardier Transport (BT) as a case study.

We conclude from the literature review that several methods have been proposed for maintenance (mathematical modelling, simulation-based techniques, AI tools, etc.). However, the previous cited methods consider the mass of data accumulated over the years from integrated embedded sensors of CPPS (historical data) to make effective maintenance decisions. Few works use real time data to detect the deviation of the system and treat in real time the system malfunctions.

3 Observation-Based Maintenance Framework

3.1 The Proposed Framework

The faults in CPPS may be due to internal causes (for example: machine breakdown) or external causes. The proposed framework aims to get early discovery of system faults that may compromise the reliability of the production system.

In this work, an observation-based predictive and reactive maintenance model framework is proposed. The main objective is to identify and localize the disruptions, assess their criticalities and then notify maintenance managers or operators via IoT tools.
Figure 1 presents the flowchart of the proposed framework. The framework is structured in four main parts detailed hereinafter.

**The Cyber-Physical Production System (CPPS)**

The first part of the proposed framework is the CPPS. It is decomposed into two parts: a physical part, including workstations, storage and transfer means and a cyber part responsible to control the physical components of CPPS. This logical part of the CPPS includes infrastructures such as programmable logic controllers (PLC), a manufacturing execution system (MES) and other elements.

**Data Acquisition (DA)**

A fundamental part of our framework is the acquisition of data from the equipment. This function is important because it allows knowing the state and the behaviour of the CPPS. There are two sources of information: i) CPPS components such as MES that provide equipment data and production-specific data; ii) other components such as PLCs refer to measurable data concerning a product being processed, and also information from the sensors which is necessary for control. The Data Communication module allows this exchange with the CPPS (denoted (a) in Fig. 1).

Some information might also be provided by some other sources than the CPPS itself. For example, with the IIoT technologies, it is possible to sense the surrounding physical environment. Various signals such as vibration, pressure, or temperature can
be extracted through sensors. In Fig. 1, this flow of data was simplified and drawn as collected from IIoT devices (b).

**Set of Detectors**
The third part of the framework consists of a set of *n* Detector modules (D₁ to Dₙ) used to detect CPPS malfunctions. The core of each detector is a real-time observation model. Based on information from the CPPS and IoT sensors, given by the Data Acquisition (DA), this model predicts what the “ideal” (or nominal) behaviour of the system should be. The real time observation model Mᵢ determines the difference (Δ) between the nominal behaviour (e) and the actual behaviour of the CPPS (c). The difference provides a valuable information about the dysfunction occurring in the CPPS. Each detector is assigned to a specific aspect of the CPPS (real-time events, thermal behaviour, energy management, economic, etc.).

**Intelligent Maintenance Decision Support Center (IMaDeSC)**
The fourth part, denoted Intelligent Maintenance Decision Support Center (IMaDeSC), makes it possible to deal with the discrepancies detected between the nominal virtual behaviour and the actual behaviour. The observed malfunctions (d) detected by the observation model Mᵢ are stored in a database, hence creating a history of how the equipment has been used over time.

These historical data are used as input of a ML algorithm. The analysis of these data aims to detect abnormal patterns and identify recurring failure scenarios. The purpose of using machine learning methods is to improve the system’s overall efficiency.

In fact, the historical data analysis allows defining preventive maintenance strategies. Thus the prediction of the future behaviour of the system indicates the potential occurrence of a failure within a time window and the notification of the operator of the dysfunction as well as of its criticality. Therefore, a preventive maintenance task schedule can be established before the malfunction occurs.

Analysis of immediate data enables decisions to be made regarding reactive maintenance, based for example on an estimated threshold set beforehand by the operator. The results of this analysis can be recommendations transmitted to the operator to help him undertake reactive responses that will be applied to the CPPS. These recommendations for adjusting maintenance decisions can be transmitted via some IIoT devices (connected watches, iPhone, virtual reality tools, etc.).

### 3.2 Artificial Intelligence Techniques

As presented in Sect. 2, machine learning methods have been used to deal with maintenance issues and to improve the efficiency of the system. ML techniques are a branch of AI methods based on the use of huge amounts of data to learn, identify patterns and make decisions with minimal human intervention [20]. Classification and regression are the major supervised ML learning techniques used for prediction target.

Supervised learning is based on using labelled data with a correct output to learn and be able to classify. Classification is the process of finding mathematical and statistical models to separate data into multiple classes where data can be divided into binary or multiple discrete labels. It is used to classify normal and abnormal patterns, i.e. “faulty”
and “healthy” within certain time periods in predictive maintenance. In opposition, regression is the process of finding a model to distinguish the data into continuous real values. It is generally used to predict the remaining useful time of the industrial equipment.

Here, the challenge is to choose the appropriate and most efficient technique in order to present a solid decision-making tool. There are several methods able to detect failures. According to literature, the most relevant ones and adequate in maintenance predictive model include logistic regression (LR), artificial neural network (ANN), and support vector machine (SVM).

**Logistic regression** is a classification technique used to assign observations to a discrete set of classes for analysing problems, where there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). The logistic regression uses the logistic function, which is an S-discriminant one, to squeeze the output of a linear equation between 0 and 1. It is used in predictive maintenance cases to model the probability of a certain event existing such as healthy or faulty event.

An **Artificial Neural Network** (ANN) is a method inspired by the brain biological structure, built to process information that are linked together. The basic building block is similar to a real neural network. Thus, it has inputs, outputs and processing elements. It is very powerful when there are massive datasets where they can learn tasks by considering input data [18]. It has three basics types of layers: an input layer, hidden layers and an output one. Due to their ability to learn from examples, ANN has received an important attention, and it shows promising results for evaluating data in order to support predictive maintenance activities.

A **Support Vector Machine** (SVM) method is a hyperplane that separates two sets with a great margin. The margin is the sum of distances from the plane to the closest point of each set [14]. The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space with the maximum margin that distinctly classifies the data points. This method has often been used in condition monitoring and fault diagnosis.

### 4 Case Study on Assembly Line: Implementation Architecture

In order to assess the performance of the proposed framework, a case study on an assembly line is considered.

#### 4.1 Learning Factory Description

Figure 2 presents a global view of a virtual representation of an assembly line in the University of Nantes. This line includes six workstations and a pallet storage. Pallets transport the products and are each equipped with RFID tags enabling the storage of the list of services needed to be executed on the transported products.

Pallets move through a complex network of conveyors. At each switching point, a RFID Read/Write Unit allows to decide the orientation of the pallets. Seventeen Read/Write Units are used to identify the position of the pallets. Four PLCs and an
Industrial MES control the system. The aim of the tool is to improve the pallet movement system by detecting minor blockages causing the system to slow down and major blockages leading to the total immobilization of pallets.

![Fig. 2. The learning factory](image)

4.2 Implementation Architecture

A single detector is implemented for this proof of concept. Figure 3 presents the implemented flowchart. This detector is constructed using a discrete event simulation model running in real time; the FlexSim software was chosen. The simulated digital model communicates with CPPS components using an emulation module exchanging data such as: ID of the incoming pallet, signals from sensors and actuators through an OPC UA server.

The model simulates the nominal behaviour of the real system, detects the real-time malfunctions and saves them directly in the MySQL database. The database in turn feeds the AI tool using the chosen machine-learning algorithm, a logistic regression, analysing the data and alerting the operator about the event of a potential failure through a portable IIoT device.

4.3 Observation Model Behaviour and Data Logging

A very important step is to synchronize the observation model with actual data coming from the CPPS. Figure 4 presents the behaviour of the implemented detector. Two different phenomena can occur: (i) the actual pallet is late compared to the virtual pallet; this happens especially when the actual pallet is blocked, or (ii) the actual pallet is ahead of the virtual pallet.

For each of the seventeen RFID read/write units, two events must therefore be detected: the arrival of the virtual pallet at the location of the virtual RFID unit and the arrival of the actual pallet at the actual RFID unit location.
Fig. 3. The flowchart of the implemented framework for the case study

Fig. 4. Detector’s behaviour model
When the virtual pallet arrives at the virtual RFID unit point, it is blocked to wait for the arrival of the actual pallet. But it is important to generate an alert before the arrival of the actual pallet, especially in case of blockage of this actual pallet. This is the function of the loop denoted (d) in Fig. 4. Of course, if the actual pallet arrives during the $D_{\text{lim}}$ period, the alert is not logged in the database.

When the actual pallet arrives at the RFID unit location, the virtual pallet is resynchronized at the location of the virtual unit. This is important, so that the behaviour of the virtual model is consistent with reality.

The data analysis and decision centre is of course a central element of the proposed framework. It is based on a learning algorithm. Through literature review, it has been shown that many artificial intelligence techniques are used. In this work, the logistic regression seems to be suitable for our case.

### 4.4 Proof of Concept

Due to the COVID 19 pandemic, we did not have access to the actual learning factory. A proof of concept based on an emulation of the actual system was developed in order to validate our framework proposal (Fig. 5).

A virtual PLC programmed with Schneider Unity was used. This PLC communicates with the FlexSim observation model via an OPC UA server. The Flexsim simulator was connected to a MySQL database. A program written in Python enables the analysis of the recorded data.

This proof of concept is only made on one conveyor, containing an entry point A and an RFID read/write module $L_1$ located at a distance $D$ from point A. $V$ is the speed of the conveyor. At time $t_0$, the virtual PLC sends the information leading to the creation of a pallet in the simulator. On the date $t_1$, the PLC sends the information indicating the arrival of the actual pallet at point $L_1$. $t_1$ was programmed such that:

$$t_1 = t_0 + \frac{D}{V} + \varepsilon$$
with $\varepsilon$ being a random variable such as $\varepsilon = \text{Uniform}(-\frac{D}{v}, \frac{D}{v})$ allowing us to introduce a perturbation.

Each time a pallet is created, a new value of $\varepsilon$ is fired. As shown in Table 1, the virtual PLC generates a population of values of $\varepsilon$ that can be split into five classes. If $\varepsilon$ is negative, this corresponds to an early arrival of the pallet at L1. Otherwise, this corresponds to a late arrival.

| Class-2 | $-3700\text{ ms} \leq \varepsilon < -2000\text{ ms}$ | Arrival very early |
|---------|-----------------------------------------------|--------------------|
| Class-1 | $-2000\text{ ms} \leq \varepsilon < -500\text{ ms}$ | Arrival early |
| Class0  | $-500\text{ ms} \leq \varepsilon < 500\text{ ms}$ | Arrival at time |
| Class1  | $500\text{ ms} \leq \varepsilon < 2000\text{ ms}$ | Late arrival |
| Class2  | $2000\text{ ms} \leq \varepsilon < 3700\text{ ms}$ | Very late arrival |

The detector analyses the differences between the arrival dates of the virtual pallets in the observation model and the actual arrival dates provided by the virtual PLC. This difference $\Delta$ is logged in the SQL database together with the timestamp, the identification of the tag reader unit and the identification of the pallet. We created a toy application in Python to demonstrate the possibility to analyze these data. To do so, it is meant to reconstruct the classes of Table 1. Table 2 shows the number of data logs in the SQL database for each class of $\varepsilon$ together with the number of logs the Python application considered in each class. Globally, the number of logs is coherent, which demonstrates the accuracy of the observations made.

| Classes | Number of delays $\varepsilon$ generated by the virtual PLC | Number of delays $\Delta$ observed by the Python application |
|---------|----------------------------------------------------------|----------------------------------------------------------|
| Class-2 | 15                                                       | 16                                                       |
| Class-1 | 11                                                       | 10                                                       |
| Class0  | 6                                                        | 6                                                        |
| Class1  | 7                                                        | 7                                                        |
| Class2  | 11                                                       | 11                                                       |

The next step will be to implement the framework on the real learning factory. In parallel, machine learning methods will be implemented to interpret the results and propose probable causes of dysfunctions. Several packages are needed for logistic regression in Python. The most popular data science and machine learning libraries such as Scikit Learn, Numpy, Panda and Matplotlib allow writing elegant and compact code and also implementing models and solving problems. We already used these packages in the toy
application in order to train our LR algorithm and improve its accuracy using simulated observations to predict faulty patterns, while waiting to experiment it on the real learning factory and adjust its parameters. For now, the model and application were only about detecting the gaps between the real system and its nominal model. However, we shall introduce in a next version an extra analysis aiming at detecting pallet defectiveness or RFID read/write units’ latency and propose solutions based on the problem criticality.

5 Conclusions

In this work, an observation framework based on artificial intelligence techniques is proposed to deal with CPPS maintenance. The proposed framework aims to help managers and supervisors to detect and predict dysfunctions of the CPPS. The objective of the integrated tool is to improve the system’s reliability and to adjust maintenance decisions. To assess its performance, the implementation of the framework is detailed and tested on a real case study. Due to the COVID 19 pandemic, the proof of concept is detailed and only based on a virtual implementation. Tests on the full learning factory are the first objectives of future works. Different AI techniques will be implemented on the framework and tested in order to select the best one which is able to optimize maintenance strategies.

Another future direction is to generalize the framework to get a first design and implementation in the context of integration inside the Digital Twin of CPPS. The idea would be to add other detectors with different objectives: minimize energy consumption, improve productivity, etc. The analysis of historical data in the IMaDeSC will then be based on multiple objectives. Thus the Digital Twin will be able to help supervisors to make compromise-based decisions by providing advanced data for decision support.

Additional research will be carried out on the optimization of the implementation of intelligent sensors (location, frequency of sending data, cycle, etc.) to get more pertinent data via IIoT technology.

References

1. Jardine, A.K.S., Tsang, A.H.C.: Maintenance, Replacement, and Reliability: Theory and Applications, 2nd edn. CRC Press, Taylor & Francis, Boca Raton (2013)
2. Shafiee, M., Chukova, S.: Maintenance models in warranty: a literature review. Eur. J. Oper. Res. 229(3), 561–572 (2013)
3. Cardin, O.: Classification of cyber-physical production systems applications: proposition of an analysis framework. Comput. Ind. 104, 11–21 (2019)
4. Khazraei, K., Deuse, J.: A strategic standpoint on maintenance taxonomy. J. Facil. Manag. 9, 96–113 (2011)
5. Verhagen, W.J.C., De Boer, L.W.M.: Predictive maintenance for aircraft components using proportional hazard models. J. Ind. Inf. Integr. 12, 23–30 (2018)
6. Efthymiou, K., Papakostas, N., Mourtzis, D., Chryssoulis, G.: On a predictive maintenance platform for production systems. Procedia CIRP 3, 221–226 (2012)
7. Nguyen, A.-T., Reiter, S., Rigo, P.: A review on simulation-based optimization methods applied to building performance analysis. Appl. Energy 113, 1043–1058 (2014)
8. Zhu, L., Yu, F.R., Wang, Y., Ning, B., Tang, T.: Big data analytics in intelligent transportation systems: a survey. IEEE Trans. Intell. Transp. Syst. 20(1), 383–398 (2018)
9. Rault, R., Trentesaux, D.: Artificial intelligence, autonomous systems and robotics: legal innovations. In: Service Orientation in Holonic and Multi-Agent Manufacturing. Studies in Computational Intelligence, pp. 1–9. Springer, Cham (2018)
10. Shcherbakov, M.V., Glotov, A.V., Cheremisinov, S.V.: Proactive and predictive maintenance of cyber-physical systems. In: Kravets, A., Bolshakov, A., Shcherbakov, M. (eds.) Cyber-Physical Systems: Advances in Design & Modelling. Studies in Systems, Decision and Control, vol. 259. Springer, Cham (2019)
11. Chaudhuri, A.: Predictive Maintenance for Industrial IoT of Vehicle Fleets using Hierarchical Modified Fuzzy Support Vector Machine, arXiv180609612 Cs, June 2018 (2018)
12. Lee, J., Azamfar, M., Singh, J., Siahpour, S.: Integration of digital twin and deep learning in cyber-physical systems: towards smart manufacturing. IET Collab. Intell. Manuf. 2(1), 34–36 (2020)
13. Djurdjanovic, D., Lee, J., Ni, J.: watchdog agent - an infotronics-based prognostics approach for product performance degradation assessment and prediction. Adv. Eng. Inf. 17, 109–125 (2003)
14. Raza, J., Liyanage, J.P., Al Atat, H., Lee, J.: A comparative study of maintenance data classification based on neural networks, logistic regression and support vector machines. J. Qual. Maint. Eng. 16, 303–318 (2010)
15. Garcia, M.C., Sanz-Bobi, M.A., del Pico, J.: SIMAP: intelligent system for predictive maintenance: application to the health condition monitoring of a wind turbine gearbox. Comput. Ind. 57, 552–568 (2006)
16. Niu, G., Li, H.: IETM centred intelligent maintenance system integrating fuzzy semantic inference and data fusion. Microelectron. Reliabil. 75, 197–204 (2017)
17. Caesarendra, W., Widodo, A., Yang, B.-S.: Application of relevance vector machine and logistic regression for machine degradation assessment. Mech. Syst. Signal Process. 24, 1161–1171 (2010)
18. Krenek, J., Kuca, K., Blazek, P., Krejcar, O., Jun, D.: Application of artificial neural networks in condition based predictive maintenance. In: Krol, D., Madeyski, L., Nguyen, N.T. (eds.) Recent Developments in Intelligent Information Database Systems, pp. 75–86. Springer, Cham (2016)
19. Jegadeeshwaran, R., Sugumaran, V.: Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines. Mech. Syst. Signal Process. 52–53, 436–446 (2015)
20. Mbili, J., Nouiri, M., Trentesaux, D., Baert, D.: Root causes analysis and fault prediction in intelligent transportation systems: coupling unsupervised and supervised learning techniques. In: IEEE International Conference on Control, Automation and Diagnosis (ICCAD), pp. 1–6 (2019)