Accepted Manuscript

A data-driven Scheduling Approach to Smart Manufacturing

Daniel Alejandro Rossit , Fernando Tohmé , Mariano Frutos

PII: S2452-414X(18)30047-5
DOI: https://doi.org/10.1016/j.jii.2019.04.003
Reference: JII 93

To appear in: *Journal of Industrial Information Integration*

Please cite this article as: Daniel Alejandro Rossit , Fernando Tohmé , Mariano Frutos , A data-driven Scheduling Approach to Smart Manufacturing, *Journal of Industrial Information Integration* (2019), doi: https://doi.org/10.1016/j.jii.2019.04.003

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

A data-driven Scheduling Approach to Smart Manufacturing

Daniel Alejandro Rossit [a,b]* , Fernando Tohmé [b,c] and Mariano Frutos[a,d]
[a] *Engineering Department, Universidad Nacional del Sur, Bahía Blanca, Argentina;*
[b] *INMABB UNS CONICET, Departamento de Matemática, Av. Alem 1253, Bahía Blanca, Buenos Aires CP 8000, Argentina;*
[c] *Economics Department, Universidad Nacional del Sur, Bahía Blanca, Argentina;*
[d] *INMABB UNS CONICET, Departamento de Matemática, Av. Alem 1253, Bahía Blanca, Buenos Aires CP 8000, Argentina;*
***corresponding author**: Departamento de Ingeniería, Universidad Nacional del Sur, Av. Alem 1253 1° Piso, Bahía Blanca CP 8000, Argentina, daniel.rossit@uns.edu.ar

---

Abstract. Traditional methods of scheduling are mostly based on the use of pieces of information directly related to the performance of schedules, as for instance processing times, delivery dates, etc., assuming that the production system is operating normally. In the case of malfunctions, the literature concentrates on the ensuing corrective operations, like scheduling with machine breakdowns or under remanufacturing considerations. These event-driven approaches are mainly used in dynamic scheduling or rescheduling systems. Unlike those, Smart Manufacturing and Industry 4.0 production environments integrate the physical and decision-making aspects of manufacturing processes in order to achieve their decentralization and autonomy. On these grounds we propose a data-driven architecture for scheduling, in which the system has real time access to data. Then, scheduling decisions can be made ahead of time, on the basis of more information. This promising approach is based on the architecture of cyber-physical systems, with a data-driven engine that uses, in particular, Big Data techniques to extract vital information for Industry 4.0 systems.

Keywords: scheduling; Industry 4.0; Smart Manufacturing; Cyber-Physical systems; Big Data; Data driven; decision-making;

1. Introduction

The last few years have witnessed the emergence of a fourth industrial revolution, referred to as Industry 4.0 (Lu [2017]; Xu et al. [2018]). It amounts to a paradigm change in manufacturing processes, based on the heavy use of automated tools (Tao et al. 2017). Hermann et al. (2016) compiled all the practitioner and academic information on this issue and stated that Industry 4.0 is embodied in several technologies of which Cyber-Physical Systems (CPS) and the Internet of Things (IoT) provide the

ACCEPTED MANUSCRIPT

foundations. CPS ï which integrate computational systems with physical processes (Lee 2008) ï are, in particular, fundamental components of this new paradigm. Wang et al. (2015) present a thorough analysis of the future impact of CPS on manufacturing environments, based on their ten more salient characteristics. Many of them involve the decision-making procedures in production settings, endowing CPS with the ability to respond, autonomously and flexibly, to unforeseen situations. CPS have the additional advantage of generating virtual replicas of the production process (Lee et al. [2015]; Chen [2017a]), facilitating its remote administration through Cloud Computing (Wang et al. [2014]; Gao et al. [2015]). This, in turn, allows viewing the production environment as the embodiment of Cloud Manufacturing (Xu [2012]; Wang et al. [2015]). These and other applications of CPS have led to the wider concept of Cyber-Physical Production Systems (CPPS) (Monostori [2014]; Monostori et al. [2016]), representing the ensemble of sub-systems connected to the environment and among them in these enhanced Industry 4.0 settings (Lu 2017). One of the benefits of CPS is the possibility of linking directly the shop floor with a high-level Decision Support System (DSS) (Rossit & Tohmé 2018). This allows providing real time data to the DSS as well as giving the shop floor the ability to rapidly adapt to the output of the DSS.

On the other hand, a CPPS is able to collect data about its environment and the processes in which it intervenes, using sensors and wireless technologies. The amount of data generated by contemporary manufacturing industries has grown explosively, generating more than 1000 exabytes per year (Yin & Kaynak 2015). This information covers a wide range, from the properties of materials and the parameters of the processes to data about clients and suppliers (Kusiak [2017a]; Xu & Duan [2018]). So, for instance, Singh et al. (2017) explore Twitter to find out the opinion of clients about the quality of foods with the goal of improving the supply chain of this industry. Similar searches through non-traditional sources of information are extremely useful for managing manufacturing processes. A recent article in *Nature* (Kusiak 2017b) emphasizes the vital role that data analysis will play in the development of Smart Manufacturing processes. But, as pointed out by Choudhary et al. (2009), the analysis of attributes and patterns in very large databases demands intelligent and automatized data analysis methodologies. CPPS may be able to run such methodologies through Cloud Computing connections to high-performance computing facilities. This will allow translating data into knowledge, by means of intelligent tools, methods and frameworks (Choudhary et al. [2009]; Shukla & Tiwari [2017]).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

These new manufacturing structures will induce changes in the way production planning is carried out (Almada-Lobo [2016]; Wang et al. [2016]). We propose here a *data-driven* approach to the solution of production scheduling problems, which are known for being NP-hard (Pinedo 2012), and thus computationally demanding even for small instances, as for example Flow Shop cases with at least four machines (Garey et al. 1976). Since schedules are usually chosen for short time frames, the planning process has to be repeated frequently (sometimes several times in a single week). The quality of the solutions has a direct economic impact on the benefits of companies, and thus on their long-term ability to thrive in competitive markets (Framinan et al. 2014). This intrinsic criticality of scheduling processes becomes even more salient in their incorporation to the Smart Manufacturing processes of Industry 4.0. As pointed out by Monostori (2014), scheduling processes constitute one of the main challenges in the design of CPPS. Data mining procedures can be, in principle, applied during the normal course of a fabrication process, in general without having to devote extra resources to data collection (Choudhary et al. 2009).

Scheduling problems are traditionally solved taking as data parameters as processing times, delivery dates, preparation times, sequence dependent preparation times, etc. Negative impacts on the schedule, like machine breakdowns or quality losses, are handled by taking into consideration also all other aspects of the system as extra constraints on potential solutions. In this way we get alternative approaches, like flow-shop scheduling under machine breakdowns or job shop scheduling with remanufacturing, when a product does not satisfy the specifications and has to be processed again (Kis & Pesch [2005]; Ruiz & Vázquez-Rodríguez [2010]; Rossit et al [2018a]). These two approaches are naturally event-driven, since they address the issues once they have happened and afterwards proceed to restate the scheduling problem (Vieira et al [2003]; Ouelhadj & Petrovic [2009]; Rossit et al [2018b]). The importance of those constraints is that inside organizations there exist system specifically destined to provide support in those areas, as quality control or maintenance systems, usually associated to level 3 of the control standard ANSI/ISA-95.

Here we intend to take advantage of the potentialities of CPPS to use data from previous experiences, collecting it through the access to the physical processes of production and processing it to yield useful information (Chen 2017b). Besides, CPPS will absorb a good deal of the business functions of manufacturing companies (Monostori 2014), gaining access to other data, as for instance on inventory

ACCEPTED MANUSCRIPT

management, quality control and predictive maintenance. All this digitalized information can be shared and used to solve problems that might arise in carrying out a production plan. These possibilities lead us to propose a data-driven DSS for scheduling activities in Smart Manufacturing environments, using both the classical data used in the search for optimal schedules as well as that on other functionalities that may have an impact on the performance or execution of production plans. Our contribution lays the groundwork for addressing rescheduling from a new perspective. Instead of waiting for an event to happen, the data already available becomes accessible in real time. This allows the system to anticipate events as machine breakdowns. In order to be able to do that, the system must be based on a data-driven architecture in which all level 3 and below data in ISA-95 have to accessible to all systems, in particular to the scheduler, which executes the orders of clients. Our scheme of data-driven DSS oriented to scheduling will, thus, run algorithms of Big Data Analytics (BDA), profiting from the data-intensive capabilities of CPPS. We will discuss the features of this scheme and the shortcomings that can be overcome by means of the technologies of CPPS.

The plan of the paper is as follows. In section 2 covers the impact of new technologies, fundamentally CPS and CPPS on production planning. Also, we will show the impact of Big Data on Smart Manufacturing and scheduling. Furthermore, we will present a methodology of design, based on which to develop a data-driven DSS. In section 3 we will analyze the large volume of data that may influence the scheduling process. Section 4, describes the decision-making process following the ANSI/ISA 95 protocol, showing that decisions are made at different levels in several autonomous components. In section 5, we present our design of a data-driven scheduling DSS for Smart Manufacturing. We will also show how the standard architecture of CPS makes it possible to link all the components required to develop such DSS. We will validate this alternative design, comparing it to possible implementations in more traditional configurations.

2. Smart Manufacturing and Big Data Scenario

In this section we introduce the main features of Smart Manufacturing production environments as well as a brief review on the use of Big Data in Manufacturing. For the former part we will focus on the technologies collectively denominated Industry 4.0, with emphasis on CPS. For the uses of Big Data in Manufacturing we will review the

ACCEPTED MANUSCRIPT

literature on data-driven applications in production activities. Finally, we will present the approach of Helu et al. (2016) to the development of data-driven systems.

2.1. CPS and CPPS

As briefly discussed in the Introduction, CPS are some of the main components of Industry 4.0 systems, facilitating the confluence of physical and virtual spaces, integrating computational and communication processes in interaction with physical processes, adding new capabilities to physical systems (Wang et al. [2015]; Lu & Da Xu [2018]). Unlike traditional embedded systems, in which components tend to be independent, CPS feature a network of interactive I/O physical elements. Later years have witnessed great advances in this area. New intelligent CPS spur innovation and competition in different industries (aerospace, automobile, chemical, energy, infrastructure, transportation, etc.). A relevant instance of CPS is constituted by intelligent manufacturing lines, in which a single machine can carry out a variety of procedures communicating with the other components (Wang & Wang 2018b).

A more precise description of CPS can be given in terms of the five-level architecture introduced in Lee et al. (2015). They define a 5C architecture outlining the main design levels of CPS: 1) *Connection* level, 2) *Conversion* level, 3) *Cyber* level, 4) *Cognition* level and 5) *Configuration* level. Figure 1 illustrates this architecture.

The Connection level is the one at which information from the environment is collected, coming from sensors, controllers or enterprise manufacturing systems (ERP, SCM, etc). At this level it is necessary to have well designed protocols (managing different types of data) and select the proper sensors. Then, the Conversion level is the one at which data is transformed into useful information, bringing some sort of self-awareness to the machines. The Cyber level is the third one, playing a central role in the architecture, since it gathers information from all the components of the system. The fourth level is the Cognition level, at which a thorough knowledge of the system is generated. This knowledge can be used by expert users and supports the decision-making process. The final level is the Configuration level, where the information at the cyber space is fed back to the physical space. This fifth level acts as a resilience control system (RCS).

ACCEPTED MANUSCRIPT

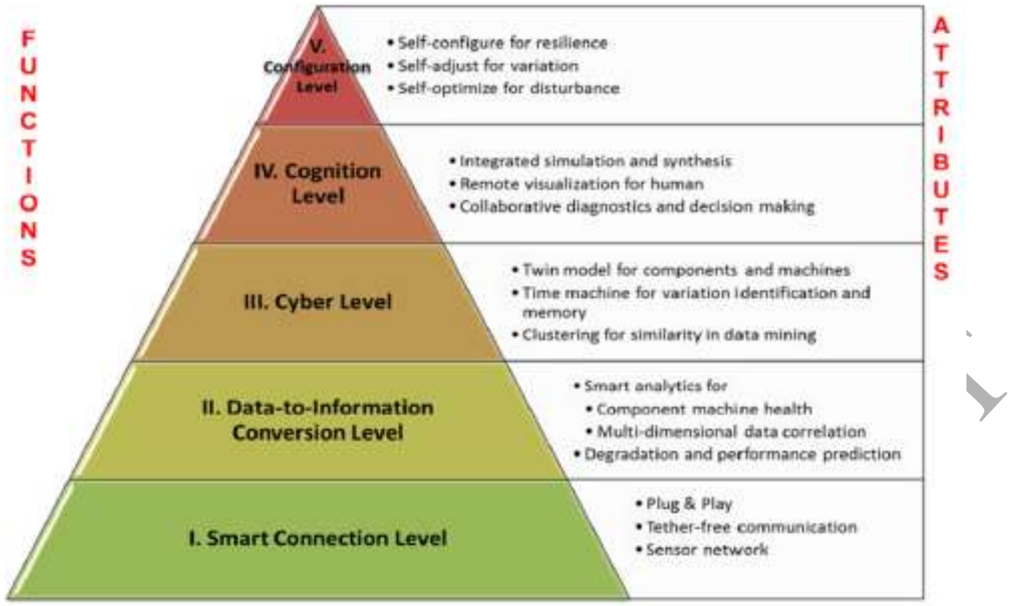

Figure 1. Architecture of Cyber-Physical Systems. Lee et al. (2015)

Lee et al. (2015) apply this five-level architecture in a Prognostics and Health Management (PHM) application, aiming to ensure the correct maintenance of the physical assets. Beyond this, we consider this architecture as roadmap for the characterization of CPS and the study of new aspects of them, in our case, incorporating scheduling into CPS.

CPS with manufacturing-specific implementations have given raise to Cyber-Physical Production Systems (CPPS). According to Monostori (2014) CPPS consist of autonomous and cooperating elements and subsystems interconnected in such way that, depending on the setting, cover all the stages of the production process, from the shop floor to the logistic networks. One of the main challenges posed by these systems is the need to develop robust approaches to scheduling, in order to face adequately to the different and unforeseen stresses on distributed production processes.

2.2. Big Data in Smart Manufacturing and scheduling

Big Data in manufacturing involves the analysis of large volumes of heterogeneous and multi-source data generated along the life cycle of industrial production (Li et al. [2015]; Hassani et al. [2018]). The ensuing databases are characterized by 5 Vôs (Chen et al. 2014): *volume* (large amounts of data), *variety* (data comes in different formats and is generated by different sources), *velocity* (data is generated and renewed by fast processes), *veracity* (data is used to reduce error levels, inconsistencies, incompleteness, ambiguities, noise and other kinds of inaccuracies) and *value* (the marginal worth conveyed by data). Smart Manufacturing environments are ready for the implementation of BDA, thanks to being fabrication systems in which sensors provide information of events and states, while managing levels provide market data (Babiceanu & Seker [2016]; Tao et al. [2018]). The systematic computational analysis of data outputs from these environments will help to make ñinformedò decisions, improving the quality of the Smart Manufacturing processes. For these reasons, data-driven systems are necessary requirements for the implementation of Smart Manufacturing environments (Tao et al. 2018).

Given the considerations advanced by Tao et al. (2018) it is relevant to revisit some contributions of BD scheduling. While there have been some BDA applications for the solution of scheduling problems in Smart Manufacturing environments, these developments are still in an initial phase. Zhong & Xu (2015) postulate a Physical Internet procedure for Job-shop scheduling, with data fed in real time by RFID devices. Metan et al. (2010) study a novel scheduling procedure of real-time selection of dispatch rules, combining simulation, data mining and statistical control process charts. Premalatha et al. (2012) analyze the results of running a (Bayesian) supervised learning algorithm on the data of production plus information on the attributes of the system. Wang & Jiang (2016) study the effects of real-time variations of work parameters and how to use the ensuing data to reschedule the production process. More recently, Ji & Wang (2017) consider a less error-prone scheduling procedure in which input data is contrasted to failure patterns detected by a BDA procedure in databases of previous runs of the system. Similarly, Vallhagen et al. (2017) propose to incorporate a BDA procedure on the data inputs to an already running scheduling process.

While these contributions improve on the typical management of scheduling data, they do not incorporate information from other functions. Our goal is to fill out this gap, by outlining how to integrate all that information in a single CPPS and its dynamics.

2.3. A data-driven framework

Helu et al. (2016), present a seven steps framework, depicted in Figure 2, based on the ideas previously presented in Helu et al. (2015). This framework provides the blueprint for the design and implementation of a data-driven DSS. The first step involves defining the *scope* of the DSS, i.e. the problem domain to which the decision-making procedure will be applied. The second step involves *identifying* the requirements of the processes implementing the decisions made, including the resources involved and the data flows in them.

The third step defines what data and from which source will be *collected*, defining the key parameters to be captured. The fourth step involves the *transmission* of data to the analysis center. The fifth step is the actual *analysis* of data, taking into account the following procedures:

- Identification of analysis techniques appropriate to the problem.
- Preparation of data for the analysis.
- Visualization methods.
- Validation and verification.

Figure 2. Data-driven decision-making framework (Helu et al. 2016)

The next step consists in *sharing* information, while the last one involves *retrieving* the information, delivering it to the right user in the corresponding format. Some activities at the last step are:

- Ensuring the accessibility of data resources.
- Presenting the data.
- Searching and querying the data through formal methods
- Managing the traceability of data.

Once established a data-driven procedure along these lines, it is possible to identify the tools and technologies that can contribute to its implementation.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

3. Scheduling planning and levels of ISA-95

The choice of a schedule of operations is part of a production planning process. More precisely, it is the last stage of planning before the actual execution of the plan (Pinedo 2012). Scheduling involves the allocation of the available production resources in a work flow generated in a previous planning stage. The choice of a schedule demands a detailed description of the production process and amounts to handle a large volume of information (Framinan et al. [2014]; Rossit et al. [2018b]). As it is intuitively evident, these decision problems have a strong combinatorial nature and consequently a high complexity.

Formally, a scheduling problem is the allocation of a family N of jobs, N = {1, 2, é , n} on a set M of machines, M = {1, 2, é , m}. Each job j consists of a class O_j of operations, where operation O_{ij} of job j must be carried out on machine i. Each operation O_{ij} has an associated processing time $p_{ij} \in$ on machine i. Each job j will have associated an ordering R_j of the operations of O_j, reflecting the precedence ordering among operations. The whole point of scheduling is to find a schedule ´ of jobs over machines yielding an optimal value F(´), where F denotes some objective function.

Scheduling problems are highly dependent on the actual details of the production setting (Job Shop, Flow Shop, etc.). This implies that different parameters (delivery dates, preparation times, waiting times, etc.) and different objective functions (makespan, total tardiness, maximal tardiness, etc.) require alternative statements of the general problem.

3.1. Scheduling data

A scheduling procedure requires a large amount of data to yield a reliable schedule of operations. Framinan et al. (2014) present a classification distinguishing between a class of model-based data and another of instance-based data. Pinedo (2012), instead, classifies scheduling data in static or dynamic, being the former independent of the actual schedule, while the latter depend on it.

Static data include data about the jobs and machines that are only inputs for the scheduling procedure. These data are relatively stable, involving the amounts to be produced, the processing times, the due dates, the release dates and the precedence constraints. The priority ordering of jobs yield also static data since it does not depend

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

on the schedule nor it is defined either by the scheduler or by means of a formula based on information about the clients. The same is true for the *contention factor*, indicating the number of alternative machines able to run operation *O*.

With respect to production resources, the static data include the velocities of machines and the programmed maintenance times, as well as more involved pieces of information, like the size of the buffers between machines or the *machine homogeneity* η, which is the ratio between the standard deviation of the number of operations that can be processed in each machine and the average number of operations that can be run on each machine (Park et al. 1997).

On the other hand, the dynamic data, depending on the schedule, involve the starting and finishing times on machines, the inactivity times, the transition between jobs times, the sequence of jobs over machines, the number of delayed jobs, etc. Other pieces of information are also dynamic, even if they do not seem directly related to the schedule. Examples of this are, for instance, the *rate of use of the system* at a given point of time (the time average of the number of machines in use over the total number of machines); or the *relative machine workload* (the rate between the maximum workload in a machine of the system over the average workload); or the *flow allowance*, representing the lead time to the due date (Pinedo 2012).

3.2. Scheduling System

All these data are fed into the scheduling procedure, which uses libraries of algorithms to generate schedules and modify them when needed. To illustrate how this works, consider the representation of a scheduling system, drawn from Pinedo (2012) and depicted in Figure 3. We can see that this system uses information from the work order and data collected in the shop floor. This is processed by a Database Management facility, a module with a series of basic functions, including multiple routines of edition, classification and search. The outputs of this process are submitted to the Schedule Generator, which uses algorithmic tools to solve the scheduling problem. Once obtained a schedule it can be modified through a Schedule Editor. The performance of either the schedule coming out from the Generator or from the Editor is evaluated. A graphic output presents all this information to the user.

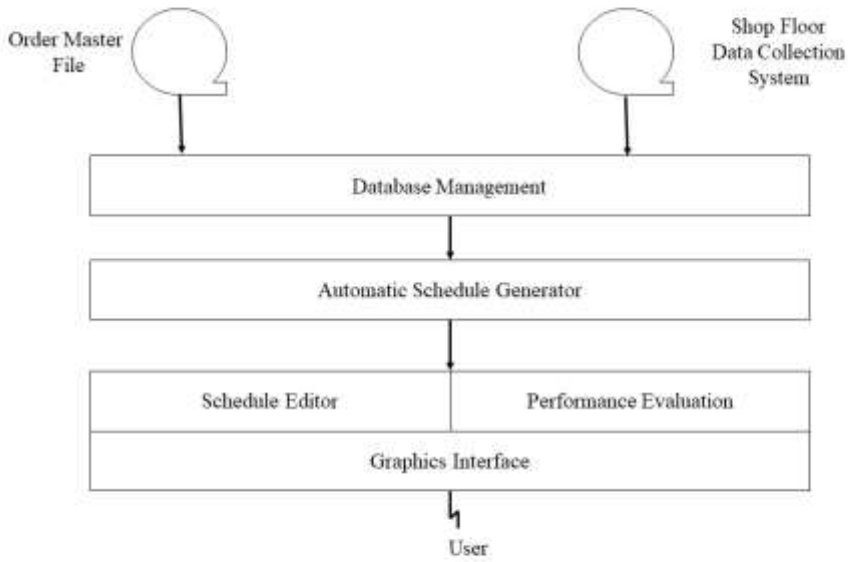

Figure 3. Scheduling System. Pinedo (2012)

The aforementioned scheduling process is part of a wider Production Planning and Control system. Again, drawing from Pinedo (2012), we show this system in Figure 4. We can see that this larger process starts by receiving orders from the clients and forecasts of demand. These are translated into material and resource use requirements. This is fed into the Scheduling module. Its output, in turn, is transmitted to the shop floor, which feedbacks information about the actual production process.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

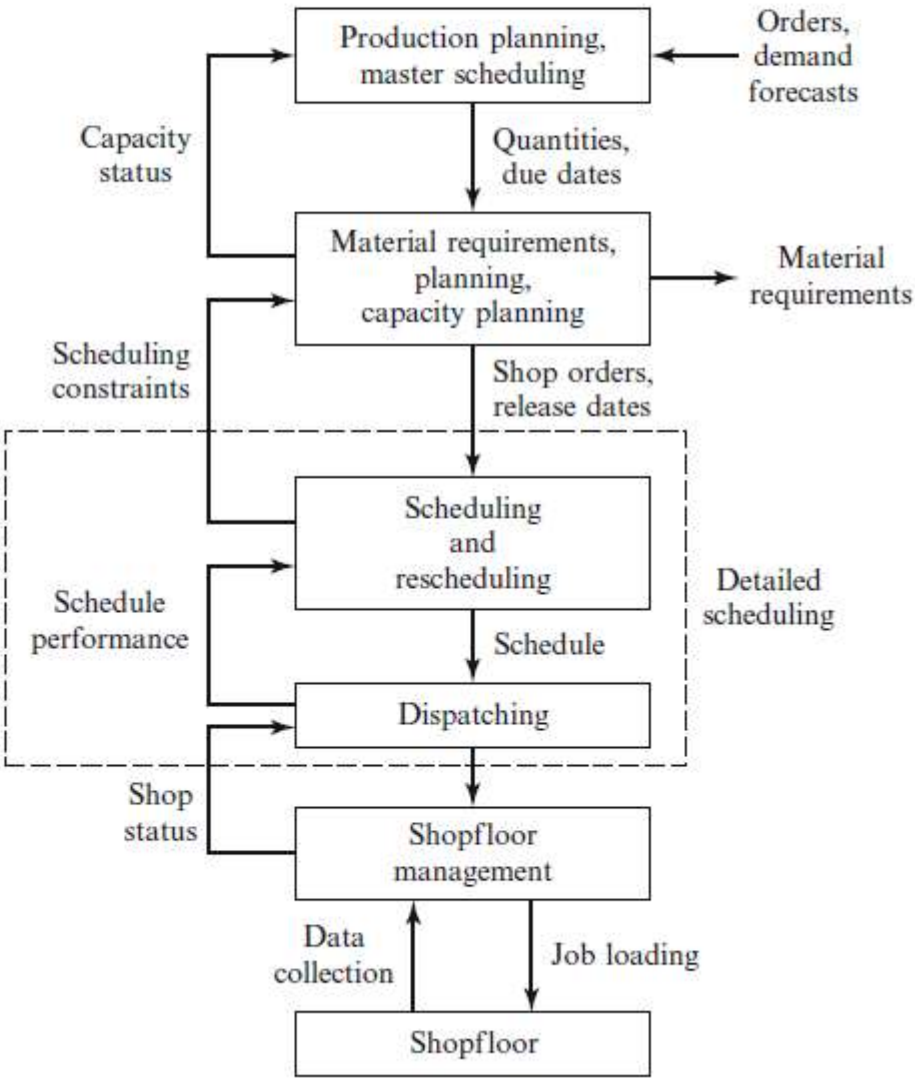

Figure 4. Production Planning and Control scheme. Pinedo (2012)

3.3. Standard ANSI/ISA 95

In order to show how we think all the aforementioned functionalities should be integrated into the scheduling system we have to briefly present the ANSI/ISA 95 standard. This standard provides a widely accepted representation of the architecture of a firm and its different levels of decision-making.

The ANSI/ISA 95 is a standard that can provide a framework for an automated interface between production facilities and control systems. It can yield a common ground for the communication of all the participants in a production process and gives a representation of how information can be modelled and used. It organizes the different levels of decision-making hierarchically. It is based on the ñPurdue Enterprise Reference

Architectureò (PERA) which distinguishes five levels, as shown in Figure 5. Level 0 is associated to the physical process of manufacturing. Level 1 involves the intelligent devices that measure and manipulate the physical process are located. Typical instruments at this level are sensors, analyzers, effectors and related instruments. Level 2 represents the control and supervision of the underlying activities. Systems acting on ISA-95 Level 2 are Supervisory Control and Data Acquisition (SCADA), Programmable Logic Controllers (PLC), Distributed Control Systems (DCS) and Batch Automation Systems. Level 3 involves the management of the operations and the production work flow in the production of the desired products. Some of the systems comprised at this level are Batch Management, manufacturing execution/operations management systems (MES/MOMS), the laboratory, maintenance and plant performance management systems, data historians and related middleware. This level has special importance for our work, since it is here where the scheduling process takes place. Finally, level 4 is associated to the business activities of the entire firm. This architecture represents, in a synthetic way, the different activities and functions of a production system. Besides, it establishes the way in which the different levels communicate; in traditional productions settings, in particular, each level interacts only with its adjacent levels (Rossit & Tohmé 2018).

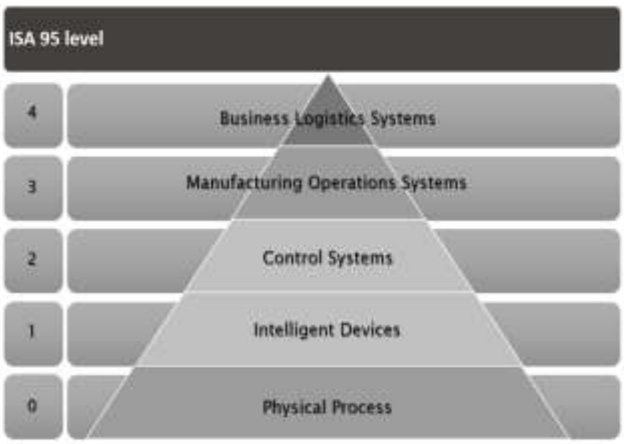

Figure 5. Control structure ANSI/ISA 95 (Rossit & Tohmé 2018)

3.4. Decision making in CPPS

CPPS change the way in which decisions are made in the realm of industrial planning and control. To introduce our view on this topic, we show in Figure 6 the levels of ISA 95 that should be incorporated into CPPS. This integration ensues from the capacities of

CPPS, which can enact physical processes (level 0), measure and handle the instruments reading the physical processes (level 1) and implement control actions over its operations (level 2). Furthermore, given the computing power of CPPS, they will also be able to plan, evaluate and manage the entire production process (level 3).

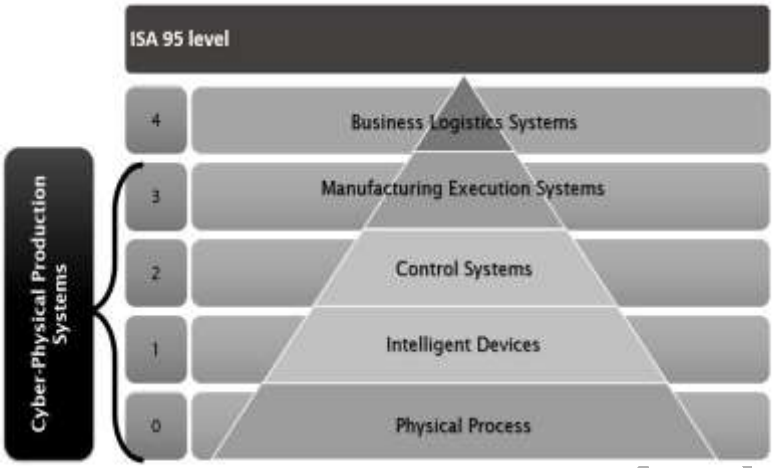

Figure 6. Control structure of a CPPS.

This integration of functionalities will yield direct benefits, as for instance increasing the flexibility of the production system in response to unexpected events; or the higher integration and transmission of information, given that a CPPS by itself can translate the data obtained at level 1 into the higher-level language used at level 3, bypassing the adjacency constraints inherent in PERA.

On the other hand, decision-making, focused on production planning, will be also impacted by the development of Industry 4.0. This will give rise to a new structure, which, while keeping PERAôs levels, will be managed by two large systems: ERP (Enterprise Resource Planning) and the CPPS. Figure 7 shows this.

ACCEPTED MANUSCRIPT

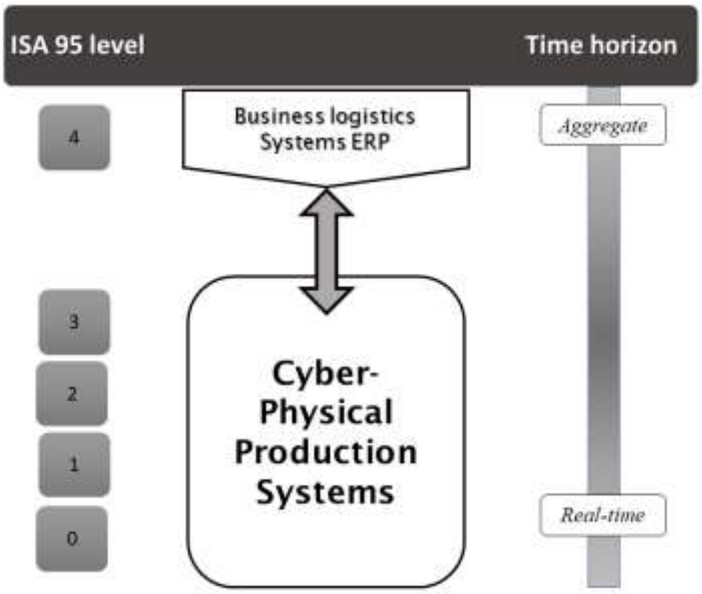

Figure 7. Distribution of ISA 95 levels between ERP and CPPS. The representation of time is drawn from the model of the Manufacturing Enterprise Solutions Association (MESA) International.

Figure 7 shows that the decisions about both the aggregate level and the goals to be pursued will be handled by the Enterprise Resource Planning (ERP) systems (tuned to smart manufacturing environments). All other decisions will be automatically and systematically run by CPPS, including the execution of the production plan in real time. In this structure, the CPPS can be seen as a set of autonomous elements collaborating to reach the goals set by the ERP system. This means, in particular, that current Manufacturing Execution Systems (MES), which take care of dispatching work orders and their scheduling in the shop floor, will be absorbed by CPPS. This will yield information of better quality, useful for both making the decisions at this level and minimizing response times, increasing the flexibility of the entire system.

A good deal of the decisions made by ERP systems (as inventory control, management of databases, handling information about suppliers, etc.) will be managed by CPPS. But we leave them separated as to indicate at what point the system becomes autonomous and up to which human interventions may be needed, particularly in the area of production planning. The linkage with human decision-makers will be at the aggregate or strategic level. They will define the goals and guidelines for the firm and the system. An ERP system will get them and will translate these guidelines for the rest of the system, in particular to the CPPS that handle the production system. The latter are

ACCEPTED MANUSCRIPT

thus not completely autonomous since they keep an open loop with the ERP system, at least on production planning (see section 2.4.1.1 in Framinan et al. 2014). Basically, this procedure plans and controls the fabrication tasks in a given manufacturing infrastructure, but not planning the infrastructure itself. In this line, Almada-Lobo (2016) postulates the absorption of Manufacturing Execution Systems in Industry 4.0 systems. On the contrary, higher levels of management, defining more general aspects of the fabrication, should not be included.

Despite this, since CPPS integrate horizontally the 3 level of ISA-95, they will gain access to quality control systems, inventory management, maintenance planning and the historical records of these systems. This endows CPPS with the ability to handle data that can influence both the schedule as the inventory issues that may generate delays or scarcity of resources, affecting its performance (Kuo & Kusiak 2018). The same is true for real time monitoring systems linked to maintenance, in which the stress on machines may force an earlier maintenance stop, affecting their availability (Lee et al. 2017). Even if the impact of these aspects on the schedule may be indirect, it is clear that having access to data on them helps to find better schedules. Our design is based on the incorporation of these new functionalities into CPPS and on their redirection towards obtaining improved scheduling processes.

Besides these new capabilities, CPPS handle the classical data required to solve the scheduling problems defined in Figures 2 and 3, but with serious advantages over traditional approaches. Since CPPS control physical processes and collect data on them, their computing power allows also to process this information. This yields an integrated view in which data is collected by the same entity that will execute the actions that ensue from processing them. Furthermore, the decision-making process can be also carried out by CPPS. Thus, these systems will at the same time collect the data, process them, make decisions and execute them. This vertical integration improves the quality of the results and the entire performance of the production system. Our DSS based on these potentialities of CPPS is intended as a step towards implementing the aforementioned vertical integration in Smart Manufacturing environments.

4. Data-driven scheduling DSS for CPPS

Given the production structure defined by the CPPS (Figure 7), we can present our own data-driven decision support system. We use the architecture defined by Lee et al.

ACCEPTED MANUSCRIPT

(2015), associating to each layer of the architecture the DSS functions running on the data-driven logic. After that we will discuss the ensuing architecture along the lines of Helu et al. (2016), evaluating which large bodies of data should be processed. Finally, we present examples, exhibiting the running behavior of this architecture.

We are now ready to present a data-driven Scheduling DSS in the aforementioned seven steps framework of Helu et al. (2016). We will discuss each step in the light of traditional scheduling methods, evaluating possible roadblocks for their implementation. We also relate the data-driven Scheduling DSS with the CPS architecture in Lee et al. (2015). Then, we analyze the CPS and Smart Manufacturing technologies that may address those limitations. We also analyze the 5V of Big Data affecting our design, postulating solutions for them. Finally, we present and overarching architecture for our design.

4.1. Data-driven Scheduling DSS

Consider the seven steps in Helu et al. (2016) applied to the development of a data driven Scheduling DSS:

Scope. The scope of a Scheduling DSS is the design of production schedules, assigning jobs to the different shop floor resources as to satisfy the goals of the organization. The very nature of scheduling procedures is data-intensive and thus, even if conceiving it as a DSS may present some complications, incorporating explicitly data sources seems a reasonable development.

Identification. The identification of the manufacturing processes, with their associated resources and flows of data are fundamental for a sound scheduling procedure. In order to yield a schedule, it has to be closely linked to the shop floor, receiving timely information about its state and evolution. This amounts to use as data and interaction sources the systems at level 3 of ISA-95 (e.g. the laboratory information management system, the warehouse management system and the maintenance management system). Possible limitations to fully implement this step arise in the area of asset management, since the coordination of all the components of the system can be highly complex, given the large amount of information required to ensure the soundness of solutions. Furthermore, since this coordination must be stable in time, the complexity is even higher. A comprehensive study on the integration of industrial information from

ACCEPTED MANUSCRIPT

different sources, providing inputs to the decision-making process can be found in Chen (2016).

Collection. Data is fed into a scheduling procedure through three channels. One comes down from the higher levels of planning, a second one comes from the shop floor where the actual production process is carried out and the third channel stems from other business functions at level 3. The data flow may face obstacles because of the traditional design of machines and other production assets, more ready to receive orders than to provide information. The same is true for control units that, in general, keep the information confined to the controlled devices. Another source of possible complications to the reception of useful information is that the connection between machines and applications requires the compatibility between different data formats. The limitations of connections at level 3 can be due to differences in information formats. In Yan et al (2017) a controller system based on IoT accomplishes this function, overcoming some of these complications.

Transmission. This step involves the physical transmission of data. In a scheduling process data have to be transmitted with high reliability, ensuring a necessary level of security. This, in turn, requires the traceability of the data. But this may be hampered by the presence of different permission levels, incompatibilities among transmission devices and, particularly, by possible asymmetries induced by wireless broadcasts. Worse yet, the latter type of transmission can suffer interferences from other, non-transmission-related, devices generating electromagnetic waves in the shop floor (Wang & Wang 2018a). Cheng et al. (2018) present a 5G-based IoT architecture industrial information transmission through mobile networks.

Analysis. This step refers to the core of any data-driven system, in which the goals of the DSS get entangled with the data received. Important aspects that arise in a scheduling context are the costs of the different assignments, the delays in deliverance, product quality, etc. If data are lacking about some of these features or the applications are not programmed to take all of them into account, the analysis will be hampered. Other limitations may arise if some analyses do not run in real time. Furthermore, since some of the analytical tools may be heuristic or trained by machine learning procedures, they may fail in the presence of unconventional scenarios. In Cheng et al (2017) several Big Data tools and manufacturing implementations are discussed, including even some scheduling applications.

ACCEPTED MANUSCRIPT

Sharing. In a scheduling procedure, this step involves generating and submitting real-time reports of the state of the shop floor associated to the orders received. This report should only reach the decision-makers with the adequate authorizations. This, in turn, requires increasing the security of this information sharing procedure and ensuring the interpretability of the report (Wang & Wang 2018c).

Retrieving. This final step amounts to the generation of a database of the analyses performed and of the patterns of behavior detected and addressed by the DSS. A difficulty here is that those pieces of information may also be accompanied by data about the context, increasing the volume of data to be stored. The data recorded has to be also traceable as to ensure the detection of the source of errors and their future correction (Lu 2017).

4.2. Architecture of the data-driven Scheduling DSS based on CPPS

The requirements and limitations to a data-driven Scheduling DSS have been established. Now we can see how to implement it in CPPS. We will see how the CPPS technologies address the seven steps in our schema and allow solving the aforementioned problems that may hamper the efficiency of scheduling processes.

The implementation of a Scheduling DSS in CPPS does not present complications with the definition of the scope, which is standard in manufacturing settings. But the identification step requires addressing the potential coordination problems among different and inhomogeneous components. But CPPS consist of autonomous systems strongly interconnected and collaborating to achieve common objectives (Monostori 2014). Their sensors and control units provide real-time access to the processes being executed and ensure the connection to different levels of decision-making.

The data collection step is also facilitated by the capabilities of CPPS, which integrate the different levels in the control structure and provide full connection among machines and applications. This ensures the compatibility among all the data formats in the system and their use in real time by any decision-making facility. The data transmission step, in turn, is supported by the design of CPPS, with intrinsic permissions and authorizations. The possible inconveniences with wireless transmissions are also diminished by the design of the entire CPPS, which makes all its components compatible.

ACCEPTED MANUSCRIPT

The data analysis step can be improved by means of the use of CPPS, thanks of their connection to high performance computing centres through cloud computing. This allows accessing more powerful analytical tools, overcoming the limitations of using only local computing resources. With respect to sharing information, the use of CPPS facilitates the compatibility among levels of decision as well as the regulation of the interactions, enabling two-way exchanges between users and the DSS.

Finally, the retrieval step will be greatly benefitted from the use of Knowledge Engineering methods that can be incorporated to the toolbox of CPPS, allowing in many cases to dispense of contexts, which can be inferred. AI techniques can be applied to improve searches in databases as well as to use those results to correct errors detected in the analysis of patterns of behaviour of the system.

Figure 8 shows how the different levels in CPS can be associated to the data-driven scheme described above. It can be seen that this association makes the CPS a ñdata-driven engineò of the Scheduling DSS, facilitating the interaction with the upper echelons in the architecture of the entire system.

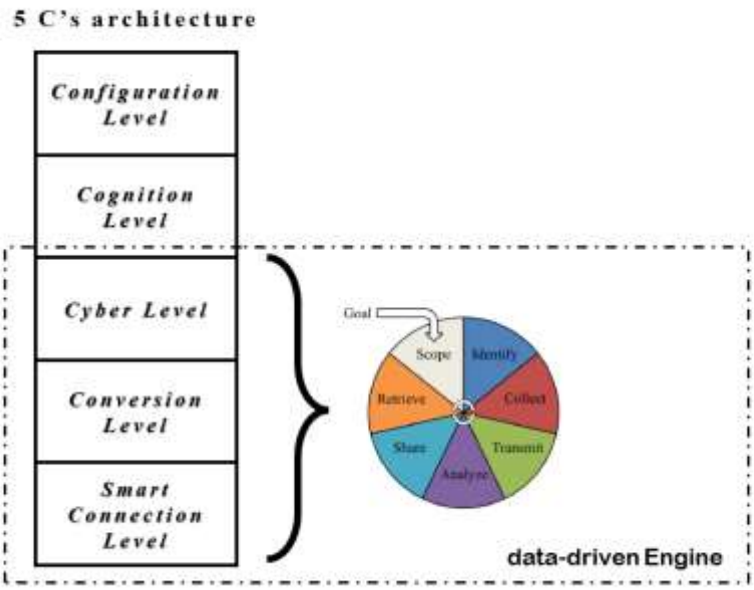

Figure 8. The data-driven engine in a CPS.

The data-driven engine of Figure 8 covers only some levels of the CPPS, corresponding to its underlying CPS, but the other modules of a scheduling procedure can also be assigned in the architecture. As shown in Figure 2, a scheduling procedure has to be able to generate schedules and the ensuing work orders.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Given the architecture of CPS (Figure 4), the Cognition level is where, according to Lee et al. (2015), the decision-making procedures become associated, in particular the high level processes that solve the scheduling problem. In our proposal this level is right above the data-driven engine. One of the outputs of the data-driven system is the information about both strategic and operational aspects of scheduling. This is fed into the Cognition level, where the choice of an optimal schedule is made.

In turn, the output of the Cognition level becomes an input for the Configuration level, where the final decision is made.

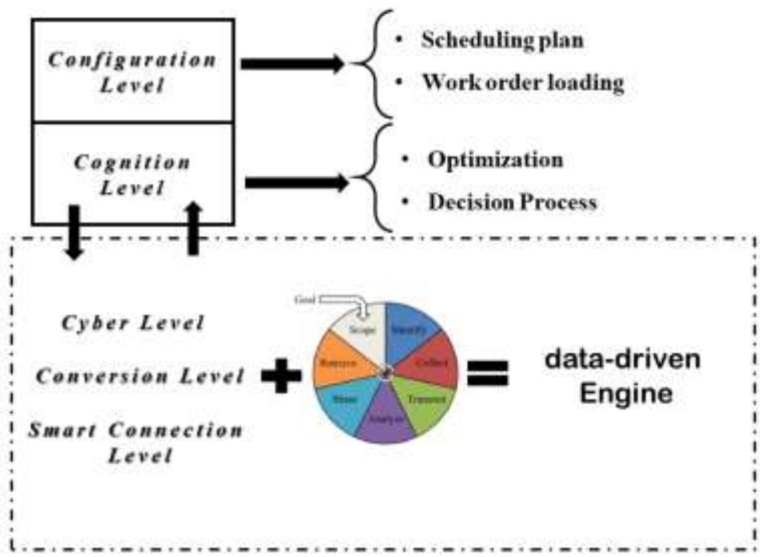

Figure 9. Full data-driven Scheduling DSS System.

4.3. Big Data 5Vs approach

The integration of business functions into the scheduling process amounts to a large flow of data. The characterization of this flow can be done in terms of the 5V approach to Big Data.

The first V corresponds to *volume*, which is indeed one of the main features in our system. Scheduling processes are already data-intensive (Pinedo 2012), because of the requirement of information on the operations and their parameters and their conditions (waiting times, their dependence on the sequence, etc.), as well as on parameters on resources, due times, timing of the arrival of orders, the way in which materials are handled and many others (Toncovich et al. 2019). Classical data for Flow Shop systems is usually well handled by means of the existing technologies and the design strategies for scheduling systems (Framinan & Ruiz 2010). But the incorporation of new functionalities complicates the whole picture. To address the increased volume of data the idea is to notice that, since the new functions are handled by CPPS the data are already in digital form, allowing via IoT to connect to Cloud Computing services, which can take on the burden of processing. This will in turn yield fast high quality results.

The second aspect is the *variety* of data. This aspect of classical scheduling systems is already well handled by existing approaches (Framinan & Ruiz 2010). But the vertical integration in CPPS yields extra layers of data in different format. In our approach the shop floor data used by classical scheduling systems (Figure 2), will be obtained in a direct real time way. While the classical systems interact with level 2 of ISA-95 (usually a SCADA-like system) in our case these levels are integrated, gaining access to data from sensors or measuring systems with different formats. The horizontal integration of business functions in our system also increases the diversity of sources (as for instance, inventory and storing facility management or quality control, just to mention just a few). Therefore, addressing this variety is a central issue for our design, in order to facilitate the flow of data in the DSS. This calls for a unification of communication protocols, on one hand the protocol based on the ISA-95 architecture and on the other protocols for the operational and tactical aspects of production. The latter should run under an open loop control procedure to allow humans in the company to understand it. B2MML (Business to Manufacturing Markup Language) seems to be a good choice for achieving this goal. It is a XML (Extensible Markup Language) allowing both the communication between digital systems as well as human interaction. The classical structure for scheduling (Harjunkoski & Bauer 2014) can be easily extended to our more integral design.

The other Vs correspond to the *velocity*, *veracity* and *value* of the data in our system. The speed of data generation is tightly linked to the dynamics of the production system. With a variety of products and short production cycles, the velocity of generation of data will be fast, while if the products are few or modular it might be slower. In both cases the data are all under the control of the CPPS which should be adequate for the traffic of data that can be generated. With respect to the veracity of data, in our case it should not be an issue, being all generated internally by the same system and its sensors. Finally, the marginal value of data (which indicate what data can be safely discarded) required for the direct solution of the scheduling problem (Figures 2 and 3) is high, while the rest of the data that arise in the vertical and horizontal integration of the system has a high aggregate value, due to their influence in increasing the flexibility and efficiency of the system. So, for instance, the real time information on the evolution of the system (vertical integration) or on the quality of the production (horizontal integration) is highly valuable for the improvement of the system.

5. Comparing the data-driven scheduling DSS system to an event-driven scheduling system: an example

To show the differences between our proposal and event-driven systems we present the following case. Consider a production system with parallel machines with several stages and assume that it runs at a stationary state. Assume that at a certain point of time a machine starts to increase its vibration. This anomalous behavior can lead to the breakdown of the machine (or require shutting it down for maintenance) or at least to generate defective products. Let us see the different responses to this situation by event-driven and our own data-driven systems.

The event driven system has a hierarchical centralized structure, in which the information relevant to the scheduling process is handled by one component, the information about maintenance by another and finally that about the quality of production by another one. Thus, the sensors detecting the vibrations do not transmit this information to the scheduling system but to the maintenance one. The latter will compare these signals with the established limits, sending a maintenance alert when the permitted threshold becomes surpassed. This alarm will come with a proposal of maintenance tasks that can, potentially, solve the problem. The maintenance supervisor will verify that analysis and define the concrete tasks to be implemented. Only once these activities are started will the scheduling system be notified to proceed to reschedule as response to this event. The quality assessment system may, in parallel detect the anomaly in the measurement of the outcomes, and this just if they indicate a persistent bias. Then, this system may also release an alert signal that can reach the scheduling system, mediated by the event of starting corrective actions. The anomalous vibrations may also trigger reprocessing operations on the defective products.

Our architecture, on the contrary, allows the sensor data become available to all level 3 systems in the ISA-95 structure. Then, they will not only sent to the maintenance system but will make it available also for the scheduling one, which will become aware before it would in an event-driven architecture. The same would happen with the control

ACCEPTED MANUSCRIPT

graphs of the quality management system, which through the horizontal integration and flexibility provided by Industry 4.0 designs, can also reach in real time the scheduling system. This allows the latter to make decisions before it would its event-driven counterpart.

In the case of maintenance, the data-driven design can anticipate the shutdown of the defective machine, by modifying the schedule, assigning a lower load of work to that the machine and deriving the more demanding operations to other parallel machines. In the same way, the scheduling system may also reassign tasks in such way that if there exists some planned idle time it can be concentrated on the defective machine. With the data from the quality control graphs the scheduling system can anticipate the bias towards increasing vibrations can assign the machine to other tasks and assess the persistence or not of the bias. The integration of systems allows the information of the anomaly to reach the scheduling system from two sources, maintenance and quality control. The scheduling system can thus assess the simultaneous reception of data, improving the efficiency of its response.

It is clear that a data-driven architecture will improve over an event-driven one by accelerating the flow of information and facilitating the anticipation of problematic situations. Furthermore, the integration of systems allows to benefit more from data, since they all can access the same sources and make use of the information..

6. Conclusions

We have presented the architecture of a scheduling system based on data-driven procedures in Smart Manufacturing environments. The implementation of this architecture profits from technologies implemented in CPPS, which allow overcoming some of the roadblocks to the implementation in more traditional settings. The advantages of this design can be found in its vertical and horizontal integration. The former allows to establish direct contact between the physical and decision-making levels, while the horizontal integration will incorporate into the CPPS business functions that traditionally are carried out independently of the scheduling activities. The resulting embedded DSS runs on a data-driven engine, which among other methodologies, applies Big Data techniques to extract vital information for a smooth functioning of Smart Manufacturing organizations.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

Almada-Lobo, F. (2016). The Industry 4.0 revolution and the future of manufacturing execution systems (MES). *Journal of Innovation Management*, *3*(4), 16-21.

Babiceanu, R. F., & Seker, R. (2016). Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *Computers in Industry*, *81*, 128-137.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171-209.

Chen, Y. (2016). Industrial information integration ð A literature review 2006ï2015. *Journal of Industrial Information Integration*, *2*, 30-64.

Chen, H. (2017a). Theoretical Foundations for Cyber-Physical Systems: A Literature Review. *Journal of Industrial Integration and Management*, *2*(03), 1750013.

Chen, H. (2017b). Applications of cyber-physical system: a literature review. *Journal of Industrial Integration and Management*, *2*(03), 1750012.

Cheng, J., Da Xu, L., Chen, W., Tao, F., & Lin, C. L. (2018). Industrial IoT in 5G Environment towards Smart Manufacturing. *Journal of Industrial Information Integration*, 10, 10-19

Cheng, Y., Chen, K., Sun, H., Zhang, Y., & Tao, F. (2018). Data and knowledge mining with Big data towards smart production. *Journal of Industrial Information Integration*, 9, 1-13.

Choudhary, A. K., Harding, J. A., & Tiwari, M. K. (2009). Data mining in manufacturing: a review based on the kind of knowledge. *Journal of Intelligent Manufacturing*, 20(5), 501.

Framinan, J. M., Leisten, R., & García, R. R. (2014). *Manufacturing Scheduling Systems*. Springer London.

Framinan, J. M., & Ruiz, R. (2010). Architecture of manufacturing scheduling systems: Literature review and an integrated proposal. *European Journal of Operational Research*, *205*(2), 237-246.

Garey, M. R., Johnson, D. S., & Sethi, R. (1976). The complexity of flowshop and jobshop scheduling. *Mathematics of operations research*, *1*(2), 117-129.

Gao, R., Wang, L., Teti, R., Dornfeld, D., Kumara, S., Mori, M., & Helu, M. (2015). Cloud-enabled prognosis for manufacturing. *CIRP Annals-Manufacturing Technology*, 64(2), 749-772.

Harjunkoski, I., & Bauer, R. (2014). Sharing data for production scheduling using the ISA-95 standard. *Frontiers in Energy Research*, *2*, 44.

Hassani, H., Huang, X., & Silva, E. (2018). Banking with blockchain-ed big data. *Journal of Management Analytics*, *5*(4), 256-275.

Helu, M., Morris, K., Jung, K., Lyons, K., & Leong, S. (2015). Identifying performance assurance challenges for smart manufacturing. *Manufacturing letters*, *6*, 1-4.

Helu, M., Libes, D., Lubell, J., Lyons, K., & Morris, K. C. (2016, August). Enabling smart manufacturing technologies for decision-making support. In *ASME 2016 international design engineering technical conferences and computers and information in engineering conference*.

Hermann, M., Pentek, T., & Otto, B. (2016, January). Design principles for industrie 4.0 scenarios. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on* (pp. 3928-3937). IEEE.

ACCEPTED MANUSCRIPT

Ji, W., & Wang, L. (2017). Big data analytics based fault prediction for shop floor scheduling. *Journal of Manufacturing Systems*, *43*, 187-194.

Kis, T., & Pesch, E. (2005). A review of exact solution methods for the non-preemptive multiprocessor flowshop problem. *European Journal of Operational Research*, *164*(3), 592-608.

Kuo, Y. H., & Kusiak, A. (2018). From data to big data in production research: the past and future trends. *International Journal of Production Research*, 1-26.

Kusiak, A. (2017a). Smart manufacturing. *International Journal of Production Research*, 1-10.

Kusiak, A. (2017b). Smart manufacturing must embrace big data. *Nature*, *544*(7648), 23.

Lee, E. A. (2008, May). Cyber physical systems: Design challenges. In Object oriented real-time distributed computing (isorc), *2008 11th ieee international symposium on* (pp. 363-369). IEEE.

Lee, J., Kao, H. A., & Yang, S. (2014a). Service innovation and smart analytics for industry 4.0 and big data environment. *Procedia Cirp*, *16*, 3-8.

Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, *3*, 18-23.

Lee, J., Jin, C., & Bagheri, B. (2017). Cyber physical systems for predictive production systems. *Production Engineering*, *11*(2), 155-165.

Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big data in product lifecycle management. *The International Journal of Advanced Manufacturing Technology*, *81*(1-4), 667-684.

Lu, Y. (2017). Industry 4.0: a survey on technologies, applications and open research issues. *Journal of Industrial Information Integration*, *6*, 1-10.

Lu, Y., & Xu, L. (2018). Internet of Things (IoT) Cybersecurity Research: A Review of Current Research Topics. IEEE Internet of Things Journal, online published. DOI: 10.1109/JIOT.2018.2869847.

Metan, G., Sabuncuoglu, I., & Pierreval, H. (2010). Real time selection of scheduling rules and knowledge extraction via dynamically controlled data mining. *International Journal of Production Research*, *48*(23), 6909-6938.

Monostori, L. (2014). Cyber-physical production systems: Roots, expectations and R&D challenges. *Procedia CIRP*, *17*, 9-13.

Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., ... & Ueda, K. (2016). Cyber-physical systems in manufacturing. *CIRP Annals-Manufacturing Technology*, *65*(2), 621-641.

Ouelhadj, D., & Petrovic, S. (2009). A survey of dynamic scheduling in manufacturing systems. *Journal of scheduling*, *12*(4), 417.

Park, S. C., Raman, N., & Shaw, M. J. (1997). Adaptive scheduling in dynamic flexible manufacturing systems: a dynamic rule selection approach. *IEEE Transactions on Robotics and Automation*, *13*(4), 486-502.

Pinedo, M. (2012). *Scheduling*. New York: Springer.

Premalatha, S., & Baskar, N. (2012). Implementation of supervised statistical data mining algorithm for single machine scheduling. *Journal of Advances in Management Research*, *9*(2), 170-177.

Rossit, D. & Tohmé, F. (2018). Scheduling research contributions to Smart manufacturing. *Manufacturing Letters*. 15 (B), 111-114.

ACCEPTED MANUSCRIPT

Rossit, D. A., Tohmé, F., & Frutos, M. (2018a). The non-permutation flow-shop scheduling problem: a literature review. *Omega*, 77, 143-153

Rossit, D. A., Tohmé, F. & Frutos, M. (2018b). Industry 4.0: Smart Scheduling. *International Journal of Production Research*. **In press**

Ruiz, R., & Vázquez-Rodríguez, J. A. (2010). The hybrid flow shop scheduling problem. *European journal of operational research*, *205*(1), 1-18.

Shukla, M., & Tiwari, M. K. (2017). Big-data analytics framework for incorporating smallholders in sustainable palm oil production. *Production Planning & Control*, *28*(16), 1365-1377.

Singh, A., Shukla, N., & Mishra, N. (2017). Social media data analytics to improve supply chain management in food industries. *Transportation Research Part E: Logistics and Transportation Review*.

Tao, F., Cheng, Y., Zhang, L., & Nee, A. Y. (2017). Advanced manufacturing systems: socialization characteristics and trends. *Journal of Intelligent Manufacturing*, *28*(5), 1079-1094.

Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems*.

Toncovich, A., Rossit, D.A., Frutos, M. & Rossit, D.G. (2019). Solving a multi-objective manufacturing cell scheduling problem with the consideration of warehouses using a simulated annealing based procedure. *International Journal of Industrial Engineering Computations*, *10*(1), 1-16.

Vallhagen, J., Almgren, T., & Thörnblad, K. (2017). Advanced use of Data as an Enabler for Adaptive Production Control using Mathematical Optimization ïAn Application of Industry 4.0 Principles. *Procedia Manufacturing*, *11*, 663-670.

Vieira, G. E., Herrmann, J. W., & Lin, E. (2003). Rescheduling manufacturing systems: a framework of strategies, policies, and methods. *Journal of scheduling*, *6*(1), 39-62.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84.

Wang, L., Wang, X. V., Gao, L., & Váncza, J. (2014). A cloud-based approach for WEEE remanufacturing. *CIRP Annals-Manufacturing Technology*, *63*(1), 409-412.

Wang, P., Gao, R. X., & Fan, Z. (2015). Cloud computing for cloud manufacturing: benefits and limitations. *Journal of Manufacturing Science and Engineering*, *137*(4), 040901.

Wang, L., Törngren, M., & Onori, M. (2015). Current status and advancement of cyber-physical systems in manufacturing. *Journal of Manufacturing Systems*, 37(Part 2), 517-527

Wang, S., Wan, J., Zhang, D., Li, D., & Zhang, C. (2016). Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. *Computer Networks*, *101*, 158-168.

Wang, C., & Jiang, P. (2016). Manifold learning based rescheduling decision mechanism for recessive disturbances in RFID-driven job shops. *Journal of Intelligent Manufacturing*, 1-16.

Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. *Journal of Manufacturing Systems*.

ACCEPTED MANUSCRIPT

Wang, L., & Wang, X. V. (2018a). Latest Advancement in CPS and IoT Applications. In *Cloud-Based Cyber-Physical Systems in Manufacturing* (pp. 33-61). Springer, Cham.

Wang, L., & Wang, X. V. (2018b). Machine Availability Monitoring and Process Planning. In *Cloud-Based Cyber-Physical Systems in Manufacturing* (pp. 83-103). Springer, Cham.

Wang, L., & Wang, X. V. (2018c). Challenges in Cybersecurity. In Cloud-Based Cyber-Physical Systems in Manufacturing (pp. 63-79). Springer, Cham.

Xu, X. (2012). From cloud computing to cloud manufacturing. *Robotics and computer-integrated manufacturing*, *28*(1), 75-86.

Xu, L. D., Xu, E. L., & Li, L. (2018). Industry 4.0: state of the art and future trends. *International Journal of Production Research*, *56*(8), 2941-2962.

Xu, L. D., & Duan, L. (2019). Big data for cyber physical systems in industry 4.0: a survey. *Enterprise Information Systems*, 13(2), 148-169.

Yan, H., Jun, L., Bo, P. Z., Yue, X., & Su, H. (2017). Mixed time-triggered and event-triggered industrial controller in IoT environment. *Journal of Industrial Information Integration*, 11, 11-18.

Yin, S., & Kaynak, O. (2015). Big data for modern industry: challenges and trends [point of view]. Proceedings *of the IEEE*, *103*(2), 143-146.

Zhong, R. Y., & Xu, C. (2015, August). A Job-Shop Scheduling Model with Real-Time Feedback for Physical Internet-Based Manufacturing Shopfloor. In *Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), 2015 IEEE 12th Intl Conf on* (pp. 638-641). IEEE.

ACCEPTED MANUSCRIPT