Industrial Cyber Physical Systems Supported by Distributed Advanced Data Analytics

Jonas Queiroz[1,2], Paulo Leitão[2,3], Eugénio Oliveira[1,2]

[1] Faculty of Engineering - University of Porto, 4200-465 Porto, Portugal
[2] LIACC - Artificial Intelligence and Computer Science Laboratory, 4169-007 Porto, Portugal
[3] Polytechnic Institute of Bragança, Campus Sta Apolónia, 5300-253 Bragança, Portugal,
`jonas.queiroz@fe.up.pt, pleitao@ipb.pt, eco@fe.up.pt`

Abstract. The industry digitization is transforming its business models, organizational structures and operations, mainly promoted by the advances and the mass utilization of smart methods, devices and products, being leveraged by initiatives like Industrie 4.0. In this context, the data is a valuable asset that can support the smart factory features through the use of Big Data and advanced analytics approaches. In order to address such requirements and related challenges, Cyber Physical Systems (CPS) promote the development of more intelligent, adaptable and responsiveness supervisory and control systems capable to overcome the inherent complexity and dynamics of industrial environments. In this context, this work presents an agent-based industrial CPS, where agents are endowed with data analysis capabilities for distributed, collaborative and adaptive process supervision and control. Additionally, to address the different industrial levels' requirements, this work combines two main data analysis scopes, at operational level, applying distributed data stream analysis for rapid response monitoring and control, and at supervisory level, applying big data analysis for decision-making, planning and optimization. Some experiments have been performed in the context of an electric micro grid where agents were able to perform distributed data analysis to predict the renewable energy production.

1   Introduction

The technological advances and reduced costs for collecting, storing, sharing and analyzing data, as well as the mass adoption of digital and internet based services by consumers, enterprises and governments have led to a digital transformation of the economy [1]. This process, called digitization, is drastically affecting companies and industries, requiring the transformation and reshaping of their business models, organizational structures, as well as their operations, where the traditional relationship they have with their products and customers has been changed to become continuous and unbounded, throughout the products' life-cycle [2]. In this domain, the digitization has been powered and promoted by the advances and the mass utilization of embedded and networked sensors and cloud computing technologies. It has also contributed to a considerable increase in the access and availability of diverse types of data and information,

as well as with the emergence of the smart devices and products, which are mainly researched inside the context of the Internet of Things (IoT) and Services [3].

Furthermore, these concepts go beyond the industrial domain (smart factories, machines and products), covering buildings and homes (smart buildings) and even complete cities and their infrastructures (smart cities, transportation and grids). In the digitized industry, data represents the main factor responsible for the changes and transformation in both its operations and business, and because of this it should be understood and actively managed as a valuable asset that can leverage the generated value [4]. In this sense, the data from smart machines and products has a great value, which can be further increased when integrated and combined with other data, e.g., from historical and others external, contextual and enterprise information [2]. Additionally, the high volume of data, which some years ago was underused, mainly because of the lack of tools and expertise to process and analyze it, nowadays can have its fully value extracted through the use of Big Data and advanced analytics approaches [5, 6].

In this context, data analysis has been widely applied in industrial domain, e.g., at operational level for process monitoring, diagnosis, optimization and control, and at business level for customer relationship management, supply chain, sales and others [7, 8]. For example, the continuous analysis of the available data in industrial environment can enhance the operation management, as well as online products reviews can be analyzed to improve the way companies treat their products and customers, optimizing and enhancing their profits. According to [4] the use of Big Data and advanced analytics can result in 20 to 25 percent increase in production volume and up to a 45 percent reduction in downtimes.

These new concepts of smart machines and products require industries to build and support several new technologies and infrastructures. Among them, the data management (e.g., retrieval, integration and analysis) comprises an essential requirement to support the smart capabilities and features, which pose several challenges, regarding mechanisms to integrate distributed, heterogeneous, dynamic and stream data sources [6]. Also, a cloud-based infrastructure can support all the remote and local software features, such as the monitoring of health and performance, diagnostic, remote control, optimization and reconfiguration strategies and algorithms, as well as to enable the autonomy of smart products and machines, enabling them to learn and adapt to their environment, user preferences and operate on their own, which are essential requirements to achieve Industrie 4.0 goals [9, 10]. All these aspects should be combined to develop supervisory and control systems capable to support the management of the large amount of heterogeneous, distributed and dynamic components, as well as the processes and operations, which are constantly subject to disturbances or evolving to attend new requirements, optimize and ensure the quality of outcomes and reduce the downtimes. In this context, Cyber Physical Systems (CPS) [10] proposes the integration of physical and virtual worlds to support all these requirements and capabilities, by embedding computational elements in physical entities and connecting such entities in a cloud-based infrastructure, aiming to provide a more effective management of the physical environment and their processes.

In order to realize the CPS features, namely self-adaptation, reconfiguration, responsiveness, fault tolerance, automated diagnosis and proactive maintenance [10, 11],

Multi-agent systems (MAS) [12] have being pointed as a suitable approach. In MAS, several autonomous, collaborative and self-organizing decision-making entities, called agents, interact and exchange knowledge to achieve their goals [12]. Additionally, MAS employ standardized communication interfaces and protocols which ensures more flexibility, modularity and openness for such systems. The application of agent-based technology in the industrial domain, to solve problems related to production automation and control, supervision and diagnosis, production planning, and supply chain and logistics, is surveyed in [13, 14], and have been covered by the field of Industrial Agents [15].

While most of existing works focus in the design of CPS control approaches, this work intends to contribute with the issues and challenges related to the supervisory aspects, considering the advanced data analysis as a key enabler for these systems smart features. The main objective is to provide a conceptual model to support more intelligent and adaptive supervisory and control industrial CPS. For this purpose, this work addresses two industrial levels, combining two main data analysis scopes: 1) at operational level, applying distributed data stream analysis for rapid response monitoring and control, and 2) at supervisory level, applying more robust and big data analysis for decision-making, planning and optimization. This advanced distributed and collaborative data analysis approach proposes to endow agents with data analysis capabilities and cooperation strategies, enabling them to perform local and collaborative data analysis, continuously improve and dynamically adapt their capabilities, based on the aggregation of knowledge. Some experiments, in the context of an electrical micro grid, have being performed where the preliminary results shown that agents are able to perform distributed predictive data analysis of renewable energy production.

The rest of the paper is organized as follows. Section 2 presents the main concepts of the proposed approach for the distributed advanced data analysis. Section 3 overviews and discusses the preliminary experiments and results by considering an electrical micro grid case study. Finally, Section 4 rounds up the paper with the conclusions and points out the future work.

2 Combining Data Analysis and MAS to Realize Industrial CPS

Considering the issues previously discussed, this work intends to design and develop an agent-based data analysis approach towards a flexible and adaptive industrial supervisory and control CPS, capable to cope with the requirements of Industrie 4.0. Additionally, this approach is more concerned with the supervisory and monitoring aspects than with the system and process control, covering the models and mechanisms to obtain and provide the required information for process management and decision-making, based on the data of different industrial automation levels.

2.1 Agent-based Data Analysis Powering Industrial CPS

The design of the proposed approach requires the consideration of some essential requirements and features related to ongoing and upcoming challenges and issues raised

by the Industrie 4.0 vision, as illustrated in the left side of Fig. 1. Such features cover different industry automation levels, that can be generalized in the operational level, characterized by the monitoring and control of processes, and the supervisory and business level, characterized by the supervision of the whole plant and its operation, integrated with the business management. These two levels present specific characteristics and requirements which are covered by CPS principles (Fig. 1 - center), where the operational level is mainly related with the physical world, characterized by the IoT and their smart devices, also demanding the processing and analysis of real time data streams [16] in order to attend the rapid response monitoring and control requirements. On the other hand, the supervisory level is hosted in a virtual world defined by a cloud-based infrastructure where robust software applications are used to attend the requirements of complex system management and high level information for decision-making, supported by big data analytics.
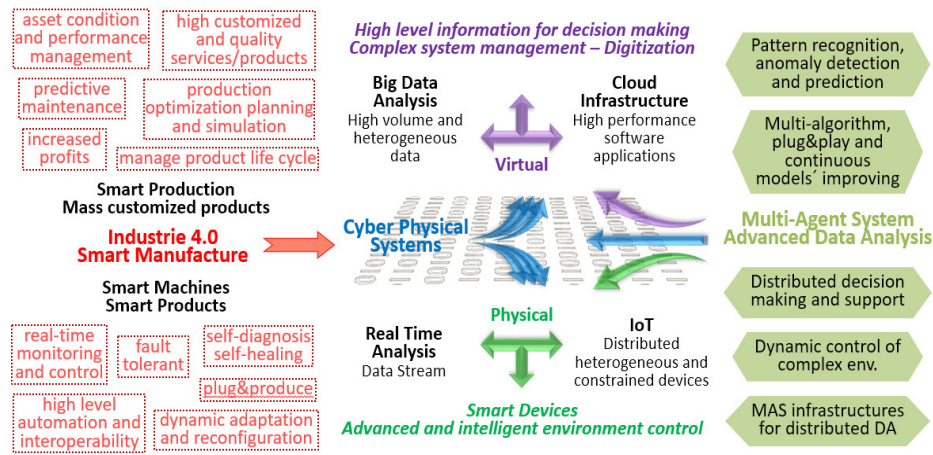

Fig. 1. Essential requirements and features of the proposed approach

In this approach, *MAS* and *Data Analysis(DA)* (Fig. 1 - right) are the basis technologies proposed to address the requirements and features envisioned by Industrie 4.0, and also taking in consideration the CPS principals. The first provides the conceptual framework to realize the underline system infrastructure that is required to achieve the desired flexibility and adaptability levels, while the second, provides the proper tools capable to analyze and obtain the required information to fulfill the desired system functionalities, also taking advantage of the increased data availability. In general, MAS and data analysis have been used successfully, but separately, to address several issues in industrial domain. In particular, MAS is used to develop adaptive and intelligent control systems, while data analysis to provide effectively data-driven decision-making algorithms. In this sense, several works leverage and discuss the potentials and how the integration of them can provide better solutions [17].

Besides combining MAS with DA, this approach also intends to address two data analysis scopes: Big Data and Data Stream analysis. The first is related with the analysis of great volumes of heterogeneous data to extract valuable information for supporting

the decision making, optimization and planning, while second is related with the analysis of the continuous incoming operational data, at real or near real-time, providing simpler information, but addressing the process monitoring and control rapid response requirements. In this context, there are already some works that cover the combination of these data analysis levels. For example, the *lambda architecture* that is proposed in [18] to address Big Data batch processing and real time data stream processing was adapted by [19] for an architecture of recommender systems, and also by [20] for an architecture to a data analysis system for a vehicular network. Other works discuss similar approaches, e.g., in [21] is presented a distributed highway traffic stream mining system that consider a central server to perform more robust DA tasks and support the monitoring components at lower level.

In this context, other desired features that can be achieved by the combination of MAS and data analysis include (see Fig. 1 - right): *MAS infrastructures for distributed DA*, and, *multi-algorithm, plug&play and continuous models' improving*. The first focuses on the use of MAS architectures and organizations to support and enhance the various data analysis phases, providing a flexible and scalable data analysis infrastructure. For example, agents can be employed to perform, in a distributed and cooperative way, the data retrieval, preprocessing, integration and analysis. The second comprises three related aspects which focus in the use of MAS to provide a dynamic and adaptive infrastructure to perform DA. For instance, the use of multiple DA algorithms and models, e.g., one per agent, which can perform the same task over the data and at the end the results could be combined to obtain more accurate information. Also, the use of MAS to provide an open and dynamic infrastructure that enables the seamless addition (plug&play) of new algorithms and data sources to the system, as well as mechanisms and algorithms that enable DA models to be continuously updated to fit the environment's dynamics.

While the previous features are more related to infrastructural aspects, there are also others related to industrial supervisory and control aspects. In the Fig. 1, the *distributed decision-making and support* element comprises coordination and negotiation mechanisms and algorithms for agents, based on distributed and local DA, take collaborative decisions for the monitoring and diagnosis system's conditions. In this sense, agents could use different DA algorithms and model to analyze environment variables, interacting with other agents to enhance or rectify the outputs. The *pattern recognition, anomaly detection and prediction* element represents the common application of DA for monitoring, prediction, diagnosis and to obtain other information about the environment and their elements conditions, usually aiming the identification of problems and possible system improvements, while the *dynamic control of complex environments* element comprises the support of dynamic adaptation and optimization of operations and processes in face of changes in the operating conditions.

2.2 Agent-based Model

Considering the features and requirements analyzed in the previous section, an agent-based model is proposed (Fig. 2 - left), comprising two layers of agents and a set of components that define the agents' capabilities (Fig. 2 - right). In the left side of Fig. 2,

at the lower layer, agents are in charge of stream data analysis, providing simple information about the processes (e.g., operation status, triggers and events), but attending rapid response constraints. In this layer, each agent is responsible to retrieve and analyze the data from process devices, in order to support monitoring and control actions. These agents could be embedded into devices to perform distributed data analysis and intelligent monitoring, cooperating to identify problems or provide information about the system. At the upper layer, agents are responsible to process and analyze great amounts of historical and incoming data from plant operations, business and also contextual or external data, in order to provide information for high level decision-making, systems optimization or activities planning (e.g., performance, quality or degradation indicators, event diagnosis, trends and forecasts). These agents could be deployed in a cloud-based computing environment, taking advantage of such kind of infrastructure and other tools to perform their tasks and also to manage the lower level agents.

In this approach, the agents of each layer comprise three modules (Fig. 2 - right) that group a set of specific components, which define the agent behaviors and capabilities. The *Data Analysis module* defines the components that perform DA tasks, the *Decision module* defines the components that process, organize and consolidate the analysis outputs, and the *Execution module* defines the components that use the consolidated information to act in the environment. Agents from both layers have two common components, *Raw/Operational data* and *Inter agent communication*, responsible to retrieve external data from the environment and manage the agent interaction, respectively.
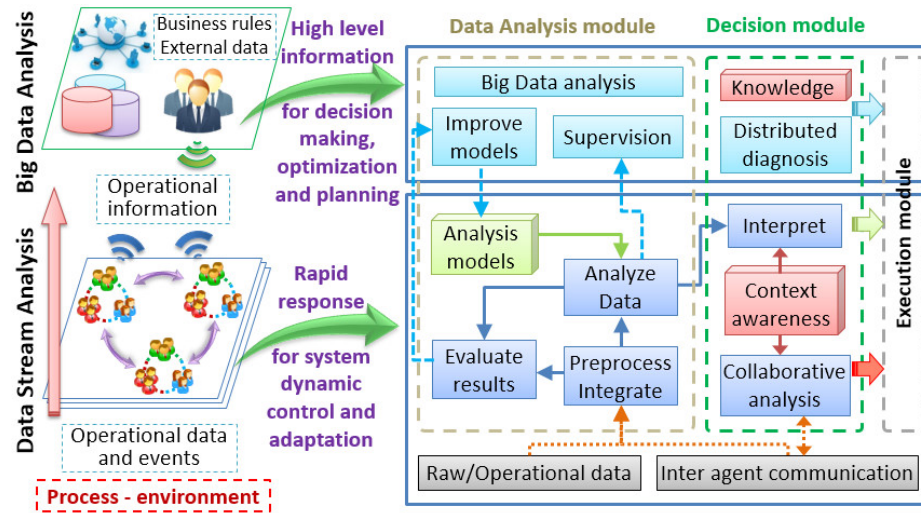

Fig. 2. Agent-based data analysis approach for adaptive industrial supervisory control systems

The proposed model comprehends only two types of agents, the lower and upper layer agents, with their monitoring and supervisory capabilities, respectively. In this sense, this model doesn't suggest any specialized role for the agents, which is application depend, i.e., given an application other capabilities can be added to each agent in order to address its role in the application. For example, in a manufacturing system, the

agents responsible to monitor the machines should present additional components that are different from the agents that monitor the products. In the model presented in Fig. 2, the *Execution module* generalizes where these specific components should be defined. This module is not detailed here, since the components of this module is application dependent and also the focus of this project is in the data analysis aspects.

The components of lower layer agents were designed to provide desired features like collaborative and scalable data analysis for a real-time monitoring and control, as well as a local autonomy and intelligence, required for the system dynamically adaptation. In this sense, the components of lower layer agents *Data Analysis* module comprise:

- *Preprocess Integrate*, which prepares the raw data to be analyzed, through the application of algorithms and methods for filtering, features extraction, data transformation and integration, as well as others preprocessing tasks that should be implemented by this component;
- *Analyze Data*, which performs several types of data analysis, such as classification, prediction, clustering. This component implements the methods required to continually apply the analysis models using the available data (provided by previous component), in order to obtain information that will be used by the agent to take actions and to monitor the related components. This component can be simple, using a single DA model, or more complex using and combining several data analysis models;
- *Analysis models*, which comprises all the data analysis models used by the agent *Analyze Data* component. One agent can have several models that can be used simultaneously (e.g., agent could apply multiple models and combine the information for better results) or for specific context or environment condition, which should be defined by the *Context aware* component;
- *Evaluate results*, which assess the analysis model accuracy to improve the agent data analysis capabilities (e.g., by comparing the model output with a system feedback). The results of this process could be used by the own agent or send to other, in order to improve its analysis models, if their accuracy were not good enough as before;

The *decision module* defines the components that process, organize and consolidate the analysis result, comprising:

- *Interpret*, which contextualizes and makes assumptions over the analysis outputs. It should implement the required decision making mechanism to enable the agent to handle with multiple, complementary, contradictory or incomplete information;
- *Collaborative analysis*, which realizes if the agent need any kind of information that can be provided by other agents, taking the required actions by interacting with other agents, to support its analysis tasks, decisions and actions;
- *Context aware*, which provides a local knowledge to support the tasks of the other components. It provides all kinds of contextual information, such as the current state or condition of the environment, its components and operations, including system and user requirements and business rules, in a passive or active way.

Like in the lower layer, the agents' components of upper layer were designed to attend some desired features regarding the analysis of aggregated data from different

sources in order to obtain high level information for assisting the decision-making, system optimization and planning tasks. In this sense, the *data analysis* module comprehends the following components:

- *Supervision*, which receives data analysis outputs and monitoring information from lower layer agents and uses it to obtain the status of a component, an operation or the whole system. This higher level information is used by these agents to support their tasks;
- *Improve models*, which (re)builds or (re)trains data analysis models used by lower layer agents, based on their feedback. This component can provide information about the confidence of the analysis outputs provided by lower level agents;
- *Big Data analysis*, which considers data from different sources, including contextual, external and historical data, as well as current data and events coming from lower layer agents (that can be firstly aggregated by the Supervision component), in order to extract information for a broader context;

The *decision module* of upper layer agents is responsible to monitor the system components (e.g., agents or devices) to support the system dynamic adaptation and it comprises the following components:

- *Knowledge*, which keeps the knowledge related to operational and technical characteristics and constraints associated to some parts or the whole system;
- *Distributed diagnose*, which compiles and provides information about the conditions of some parts or whole system, and suggests actions and their possible consequences in the system (what-if information). It also considers the information provided by the other components to interact with other upper layer agents to collaboratively identify and diagnose the whole system conditions.

3　Preliminary Experiments and Results

The described approach is being validated on a case study in the context of an electric micro grid comprising some wind turbines and photovoltaic panels. In micro grids, it is required distributed and flexible approaches capable to manage the local requirements and autonomy of each micro grid node, and at the same time assure the global micro grid energy efficiency and self-sustainability. In this context, one challenge is the integration and analysis of multiple distributed data streams produced by energy resources (micro grid nodes and their energy elements). Another challenge regards the prediction of renewable energy production and consumption that are directly dependent on the intermittent weather conditions. It means that the accuracy of the predictions is lower for future time periods, requiring the dynamic adaptation of the system for the current conditions.

In this sense, the proposed approach can address the global and local requirements and features. The use of agent-based approach can provide the required distribution and flexibility to manage the autonomy of each grid node and their energy elements. Additionally, the use of advanced data analysis approaches can provide the appropriate tools

that can take advantage of the large amount of data streams produced by the energy elements to obtain the necessary information to support the management and control of the energy production, distribution, storage and consumption.

The performed experiments only covered the lower layer aspects of the proposed approach. In this layer, three types of agents were developed (see Fig. 3), Storage (SA), Consumer (CA) and Producer (PA) agents, responsible to manage each kind of electric elements. In this sense, each kind of agent can present different data analysis capabilities according with their roles. For instance, CAs can perform the prediction of the energy demand based on the user profile, while PAs can predict the renewable energy production based on the weather conditions [22]. This anticipated information could be used to support the energy distribution along elements, as well as optimize the schedule/shift of loads, minimize peak demand and even negotiate in an energy market or provide information to the main grid operator manage the high scale energy generation. In the current experiments the upper layer agents were not developed, however such information can be send to these agents that will be in charge to perform such tasks.
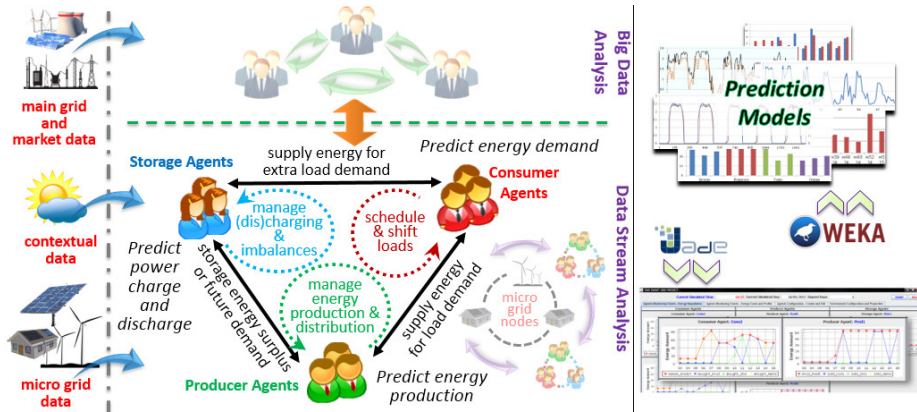

Fig. 3. Micro grid agent-based data analysis developed experiment

In order to test the proposed approach, a simple simulation environment was developed, where agents were implemented using JADE framework and data analysis models using WEKA Java API. In this environment SAs and CAs implement simple and static battery and consumption models, while PAs use real data from micro grid photovoltaic (PV) panels and wind turbines to simulate the energy production [22]. Also for the PAs different data analysis models were developed for the prediction of renewable energy production, using historical data, as well as external data from weather forecasts. The historical data used in the experiments comprehends three years of monitoring the renewable energy production operations, which record one sample at each five minutes. The development of prediction models was performed using four different algorithms provided by WEKA (M5P, M5Rules, Linear Regression, Multilayer Perceptron) [23].

The experimental results showed that PAs are capable to perform distributed data analysis for different time scopes and goals. For instance, PAs performed short-term prediction of the energy production for the current 5 minutes period, 1 step ahead (next

5 minutes period) and 3 steps ahead. The first aims the monitoring of the operational conditions of the production units to identify abnormality in energy production, while the second and third aim to provide information to correct energy imbalances. Also, through a mid-term prediction of energy production, performed by integrating external weather forecasting data, PAs were able to provide information about the amount of energy expected to be produced for the next hour interval, which could be used by engineers, grid operators and other systems to enhance and optimize the energy distribution [22]. During the energy predictions, agents were able to continuously evaluate and improve their analysis models.

The Table 1 summarizes the results of the developed experiments, illustrating the accuracy of the renewable energy production prediction, based on the models built by M5P algorithm, which was the one that in general presented better results. The high correlation coefficients (*Corr.Coef.*) illustrate the applicability of DA for renewable energy production, while the relative absolute error (*RAE*) illustrates that the prediction models presented better results for the short-term and the PV system, which can be justified by the high instability of weather conditions which is higher for the wind system. Even presenting a higher error rate for the long-term prediction, using the short-term prediction the system can dynamically adapt its operation to correct the imbalances. Most of the prediction deviation that was responsible for the higher RAE in the PV system was observed in cloudy days. It means that the prediction models should be tuned for these kinds of days, or even specific models can be built for these kind of days. In general, the wind system, presented bad results, where the days with high variable wind speed were the main responsible. Because of such bad results the experiments were not performed for the next hour period.

Table 1. The prediction performance for renewable energy production

	PV system			
	Current step	*1 step ahead*	*3 steps ahead*	*Next hour*
Corr.Coef.	0.990	0.991	0.977	0.951
RAE (%)	4.635	10.899	15.723	25.997
	Wind system			
Corr.Coef.	0.892	0.839	0.779	-
RAE (%)	38.157	40.252	46.128	-

4 Conclusions and Future Work

The increased industrial digitization and market demands for high customized products and services have been posed new requirements and challenges for industries and companies. The Industrie 4.0 initiative is promoting the related concepts and their benefits in order to improve the management of manufacturing processes and operations, optimizing and ensuring the quality of outcomes and reducing the downtimes. Among these technologies, the advanced DA and MAS can be used to attend the Industrie 4.0 vision,

contributing for the realization of industrial CPS features, namely self-adaptation, re-configuration, responsiveness, fault tolerance, automated diagnosis and proactive maintenance. In industrial domain, MAS have been used as a suitable approach to de-sign and develop flexible and adaptable industrial control systems, while data analysis is being used to provide effective algorithms to support data-driven decision-making. In this context, this work proposes the combination of these two technologies to design an agent-based data analysis approach for intelligent and adaptive industrial supervi-sory control systems, where agents are endowed with data analysis capabilities to ena-ble the dynamic distributed and collaborative process supervision and control. Addi-tionally, to address the different industrial automation levels' requirements, the pro-posed approach covers two data analysis scopes, applying data stream analysis at oper-ational level, attending the monitoring and control rapid response requirements, com-bined with a more robust Big Data analysis at supervisory and business levels to support decision making, planning and optimization tasks.

The preliminary experiments showed promising results, where several data analysis capabilities were successfully embedded in agents, including the self-improving of data analysis model. The agents were able to distributed and collaboratively perform the prediction of renewable energy production for different time scopes, and use such in-formation for the management of micro grid nodes.

Future work encompasses the detailed specification of the mechanisms and strate-gies to cover the more advanced aspects and features of the proposed approach. Re-garding the case study, it can be further explored, e.g., by developing predictive capa-bilities for CAs and SAs in order to manage energy consumption and power storage of micro grids nodes, also extending the preliminary experiments in order to validate and assess other aspects. Moreover, it is intended to explore another case studies scenario in the manufacturing domain.

5 References

1. Sabbagh, K., Friedrich, R., El-Darwiche, B., Singh, M., Koster, A.: Digitization for eco-nomic growth and job creation: Regional and industry perspective. The global information technology report 2013 - Growth and Jobs in a Hyperconnected World, Geneva, World Eco-nomic Forum, pp. 35--42 (2013)
2. Porter, M.E., Heppelmann, J.E.: How Smart, Connected Products are Transforming Com-panies. Harvard Business Review, nº October 2015, pp. 96--112 (2015)
3. Aggarwal, C.C., Ashish, N., Sheth, A.P.: The Internet of Things: A Survey from the Data-Centric Perspective. In Managing and Mining Sensor Data, pp. 383--428, Springer (2013)
4. Wee, D., Kelly, R., Cattel, J., Breunig, M.: Industry 4.0 - How to Navigate Digitization of the Manufacturing sector. McKinsey & Company (2015)
5. Sagiroglu, S., Sinanc, D.: Big data - A review. In International Conference on Collaboration Technologies and Systems (CTS) (2013)
6. Obitko, M., Jirkovský, V., Bezdíček, J.: Big Data Challenges in Industrial Automation. In Industrial Applications of Holonic and Multi-Agent Systems, pp. 305--316. Springer (2013)
7. Qin, S.J.: Survey on Data-driven Industrial Process Monitoring and Diagnosis. Annual Re-views in Control, vol. 36, nº 2, pp. 220--234, (2012)

8. Harding, J.A., Shahbaz, M., Kusiak, A.: Data Mining in Manufacturing: A Review. Journal of Manufacturing Science and Engineering, vol. 128, nº 4, pp. 969--976 (2006)
9. Drath, R., Horch, A.: Industrie 4.0: Hit or Hype? [Industry Forum]. IEEE Industrial Electronics Magazine, vol. 8, nº 2, pp. 56--58 (2014)
10. Lee, E.: Cyber Physical Systems: Design Challenges. In 11th IEEE International Symposium on Object Oriented Real-Time Distributed Computing, pp. 363--369 (2008)
11. Lee, J., Lapira, E., Bagheri, B., Kao, H.A.: Recent Advances and Trends in Predictive Manufacturing Systems in Big Data Environments. Manufacturing Letters, vol. 1, nº 1, pp 38--41 (2013)
12. Wooldridge, M.: An Introduction to Multiagent Systems. John Wiley & Sons (2009)
13. Leitão, P.: Agent-based Distributed Manufacturing Control: A State-of-the-art Survey. Engineering Applications of Artificial Intelligence, vol. 22, nº 7, pp. 979--991 (2009)
14. Metzger, M., Polakow, G.: A Survey on Applications of Agent Technology in Industrial Process Control. IEEE Trans. on Industrial Informatics, vol. 7, nº 4, pp. 570--581 (2011)
15. Leitão, P., Karnouskos, S.: Industrial Agents: Emerging Applications of Software Agents in Industry. Morgan Kaufmann (2015)
16. Gaber, M.M., Zaslavsky, A., Krishnaswamy, S.: Mining data streams: a review. ACM SIGMOD Record, vol. 34, nº 2, pp. 18--26 (2005)
17. Cao, L., Gorodetsky, V., Mitkas, P.: Agent Mining: The Synergy of Agents and Data Mining. IEEE Intelligent Systems, vol. 24, nº 3, pp. 64--72 (2009)
18. Marz, N., Warren, J.: Big Data: Principles and Best Practices of Scalable Realtime Data Systems. Chapter 1, ed. 1st, Manning Publications, p. 328 (2014)
19. Twardowski, B., Ryzko, D.: Multi-agent Architecture for Real-time Big Data Processing. In Int'l Joint Conf. on Web Intelligence and Intelligent Agent Technol., pp. 333--337 (2014)
20. Daniel, A., Paul, A., Ahmad, A.: Near real-time big data analysis on vehicular networks. In International Conference on Soft-Computing and Networks Security (ICSNS-2015), (2015)
21. Liu, Y., Choudhary, A., Zhou, J., Khokhar, A.: A scalable distributed stream mining system for highway traffic data. In Proceedings of the 10th European conference on Principle and Practice of Knowledge Discovery in Databases (PKDD'06) (2006)
22. Queiroz, J., Leitão, P., Dias, A.: Predictive Data Analysis Driven Multi-Agent System Approach for Electrical Micro Grids Management. In Proceedings of the IEEE ISIE'16, pp. 738--743 (2016)
23. Witten, I.H., Frank, E., Hall, M.A.: Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 3rd ed. (2011)