Agent-Based Data Analysis Towards the Dynamic Adaptation of Industrial Automation Processes

Jonas Queiroz, Paulo Leitão

Agent-Based Data Analysis towards the Dynamic Adaptation of Industrial Automation Processes

Jonas Queiroz[1,2] and Paulo Leitão[2,3]

[1] Faculty of Engineering – University of Porto, 4200-465 Porto, Portugal
[2] 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

Abstract. Industrial complex systems demand the dynamic adaptation and optimization of their operation to cope with operational and business changes. In order to address such requirements and challenges, cyber-physical systems promotes the development of intelligent production units and products. The realization of such concepts requires, amongst others, advanced data analysis approaches, capable to take advantage of increased availability of data, in order to overcome the inherent dynamics of industrial environments, by providing more modular, adaptable and responsiveness systems. In this context, this work introduces an agent-based data analysis approach to support the supervisory and control levels of industrial processes. It proposes to endow agents with data analysis capabilities and cooperation strategies, enabling them to perform distributed data analysis and dynamically improve their analysis capabilities, based on the aggregation of shared knowledge. Some experiments have been performed in the context of an electric micro grid to validate this approach.

Keywords: multi-agent system; distributed data analysis; adaptive system.

1   Introduction

Companies are subject to the market dynamics and competitiveness, demanding highly customized and quality products and services with reduced prices. Additionally, they operate in complex environments, characterized by distributed and heterogeneous systems, which require the dynamic adaptation and optimization of their processes to cope with operational changes caused by technical problems (e.g., equipment damage and resource availability) or changes in business rules (e.g., new product demands or design). In order to address such requirements, Industrie 4.0 vision [1] and Cyber-Physical Systems (CPS) principles [2] promote the use of smart machines, systems and products. To support the realization of such concepts, the use of advanced data analysis approaches should be considered, taking advantage of the increased availability of great amounts of data produced in such environments. The continuous data analysis enables to identify and predict the system operational conditions, providing valuable information to support process supervision and control.

The existing approaches handle these issues without worrying about the inherent dynamics and complexity of these environments (e.g., in face of condition changes, be capable to adapt its data analysis capabilities), which require features of modularity, adaptability and responsiveness. In industrial environments, approaches also need to support different data analysis scopes: 1) at operational level, applying distributed streaming data analysis for rapid response, and 2) at supervisory level, applying centralized and more robust data analysis for decision-making.

Multi-agent systems (MAS) [3-4] have being pointed as a suitable approach to support the design and development of distributed, flexible and dynamic systems. Thus, the goal of the ongoing work is to design an advanced distributed data analysis approach, based on MAS, to support intelligent and adaptive supervisory control applications, towards the dynamic adaptation of industrial automation processes. This approach proposes to endow agents with data analysis capabilities and cooperation strategies, enabling them to perform distributed data analysis, continuously improve and dynamically adapt their local capabilities, based on the aggregation of knowledge. Some experiments, in the context of an electrical micro grid, have being performed to consolidate and validate the proposed approach. The preliminary results shown that agents are able to perform distributed predictive data analysis of energy production.

Although this approach presents a great potential to address the recent challenges faced by industry, there are some open questions regarding to how properly endow agents with data analysis capabilities and how the extracted information could enhance agents' behaviors. The conceptual and technical answers for such questions will enable the design and development of innovative and powerful approaches.

This paper is organized as follows. Section 2 describes the contributions of the work to the realization of CPS. Section 3 presents the literature review and Section 4 presents the proposed approach. Section 5 overviews the critical analysis of this proposal and discusses the preliminary results. Finally, Section 6 wraps up the paper with the conclusions and states the research directions.

2   Contribution to Cyber-Physical Systems

CPS promotes the integration of physical and virtual worlds, the first is characterized by a large network of interacting heterogeneous hardware devices, while the second provides robust computing infrastructures, replete of software platforms, applications and information technologies. Such integration aims a more effective management of the physical environment and their processes, by embedding computational elements in physical entities and connecting such entities in a cloud-based infrastructure.

CPS have been deployed in several fields, related to smart production, grids and buildings, where large amount of devices should be efficiently sensed and controlled in a reliable, secure, real-time and distributed manner. Additionally, such devices produce large volume of data, requiring advanced data analysis approaches in order to enable the capabilities and features envisioned by CPS, namely self-adaptation, fault tolerance, automated diagnosis and proactive maintenance [5-6]. While most of existing works focus in the design of control approaches for CPS, this work intends to contribute with the issues and challenges related to supervisory aspects, considering

the Big Data features. The main objective is to provide algorithms and mechanisms to support more intelligent and adaptive monitoring and supervisory CPS.

3   Literature Review

Industrial environments are characterized by a large network of heterogeneous devices (endowed with sensors and actuators), which monitor and control related processes [7]. The industrial management systems need to integrate and coordinate such devices, automating the overall process in order to optimize and ensure the quality of outcomes, and keep the plant availability. MAS have been proposed to address the issues of industrial systems, which need to be flexible and adaptive to cope with inherent complexity required to manage dynamic, heterogeneous and distributed components [3], [4]. In MAS, several autonomous, collaborative and self-organizing decision-making entities, called agents, interact and exchange knowledge to achieve their goals [3]. 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 [4], [8], [9], and have been covered by the Industrial Agents [8] research field.

Technological advances in sensor devices have contributed to leverage their use in industrial environments and consequently the amount of collected data [10] further increases the complexity of such environments. In many cases, the produced data is underused, mainly because it is necessary a great expertise and specialized knowledge for its integration and analysis. However, the recent popularization of the Big Data concept and its potential, cached the attention of industry. 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 [11], [12], [13], [14]. However, to effectively use data analysis and extract its full potential, several challenges found in industrial scenarios need to be overcome, such as, mechanisms to integrate distributed, heterogeneous, dynamic and streaming data sources [15].

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 these technologies can provide better solutions in various domains [16], [17], [18].

4   Research Contribution and Innovation

Considering the assumptions discussed in the previous section, this work intends to design and develop an agent-based data analysis approach towards a flexible and adaptive industrial supervisory control system, capable to cope with the dynamics and high amount of distributed and heterogeneous industrial devices. This approach is more concerned with the supervisory and monitoring aspects than with the control of

processes. Therefore, the general objective of this project encompasses mechanisms and algorithms to derive information and knowledge from data of different industrial levels, and then properly provide them for decision-making and process management.

4.1.   Agent-based Data Analysis Features and Requirements

The design of the proposed approach requires the consideration of some essential requirements and features, as illustrated in Figure 1. They are directly related to ongoing and upcoming industry challenges and issues that are raised by the Industrie 4.0 vision. As already discussed in the previous section, *MAS and Data Analysis* are the basis technologies that will support this approach. The first provides the base infrastructure to achieve the required flexibility and adaptability, while the second, provides the proper tools required to take advantage of increased data availability.

   On the other hand, to cover different industry automation levels, such as, the monitoring of the operational process and the supervision of the whole plant, the proposed approach requires to support different data analysis scopes: 1) at operational level, distributed data streaming analysis for rapid response; and, 2) at supervisory level, more robust big data analysis for decision-making. *Big Data* considers the volume, variety and velocity of data, requiring dedicated and usually distributed computing infrastructures to extract valuable information from raw data, while *Data Streaming* considers the analysis of data at real or near-real time, providing simpler information that address rapid response requirements. In the literature some works already discuss approaches to address these two kinds of data analysis scopes [18].
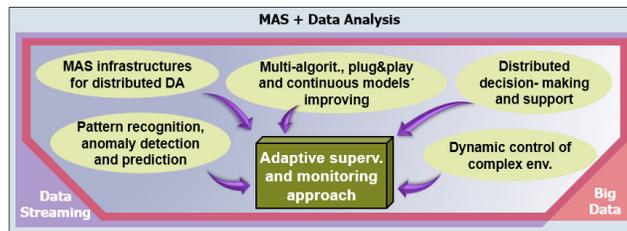

Fig. 1. Essential requirements and features of the proposed approach.

   Other requirements consider (Figure 1): 1) MAS infrastructures for distributed DA (Data Analysis), and, 2) Multi-algorithm, plug&play and continuous models' improving. The first focuses in providing a modular and scalable data analysis infrastructure, by taking advantage of the MAS approach to support and enhance the various data analysis phases. For example, agents can be employed for data retrieve, preprocess, integrate and analyze, in a distributed and cooperative way. The second comprises three related features where the focus is the utilization of MAS to provide a dynamic and adaptive infrastructure to perform data analysis. Multi-algorithms comprises the deployment of different data analysis 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. Plug&play comprises the use

of MAS to provide an open and dynamic infrastructure that enables the seamless addition of new algorithms and data sources to the system. Continuous models' improving comprises mechanisms and algorithms that enable data analysis models to be updated to fit the environments' dynamics. In this case, specialized agents could be in charge to analyze the performance of current data analysis models, updating them to enhance their current accuracy.

While the previous features are more related to infrastructural aspects, there are also others related to industrial supervisory and control aspects. In the Figure 1, the Distributed decision-making and support element comprises coordination and negotiation mechanisms for agents monitoring and diagnosing the system's conditions. The Pattern recognition, anomaly detection and prediction element represents the common application of data analysis to solve industrial problems, 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 environment or operating process conditions.

4.2. Agent-based Model

Considering the analysed features and requirements, an agent-based model is proposed (Figure 2), comprising two layers of agents. In the left side of Figure 2, at the lower layer, agents are in charge of streaming 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 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 external data, in order to provide information for high level decision-making (e.g., performance, quality or degradation indicators, event diagnosis, tends 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 (illustrated in the right side of Figure 2) that group a set of specific components, which define the agent behaviors and capabilities. The *Data Analysis module* defines the components that perform data analysis tasks, the *Decision module* defines the components that process, organize and consolidate the analysis result, 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, the *Raw/Operational data* and the *Inter agent communication*, responsible to retrieve external data from the environment and manage the agent interaction, respectively.

The components of lower layer agents *Data Analysis* module comprise:

- *Preprocess Integrate*, which prepares the raw data to be analyzed;
- *Monitoring*, which performs several types of data analysis;
- *Analysis models*, which comprises all the data analysis models used by the agent.

• *Evaluate results*, which assess the analysis model accuracy (e.g., by comparing its output with a system feedback);

The *Decision module* comprises the *Interpret*, which contextualizes and makes assumptions over the analysis result, the *Collaborative monitoring*, which realizes if the agent need any kind of information that could be provided by other agents, and, the *Context aware*, which provides a local knowledge used by the other components.
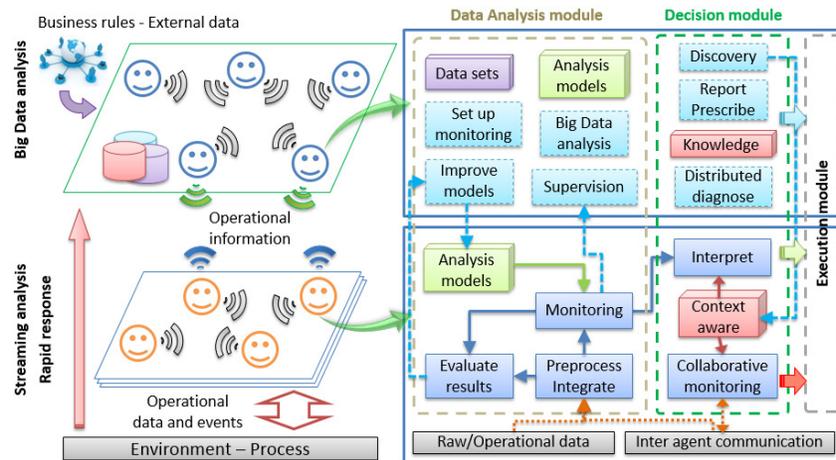

Fig. 2. Agent-based data analysis approach for adaptive industrial supervisory control systems.

The upper layer agents are also defined by several components:

• *Supervision*, which receives monitoring information from lower layer agents and uses it to obtain the status of production stations, plants or the whole process;

• *Improve models*, which retrains or rebuilds data analysis models used by lower layer agents, based on the feedback provided by these agents;

• *Set up monitoring*, which builds new data analysis models, sets up and deploys lower layer monitoring agents;

• *Big Data analysis*, which considers data from different sources, including external and historical data, in order to extract information for a broader context;

• *Analysis models*, which, like in lower level agents, comprises all the data analysis models used by the agent;

• *Data sets*, which represents the access interfaces for historical data, since external data was provided by Raw/Operational data component.

The *Decision module* of upper layer agents comprises the following components:

• *Discovery*, which monitors the system components (e.g., agents or devices) to support the dynamic adaptation of the system;

• *Report Prescribe*, 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) considering the information provided by the *Supervision, Big Data analysis*, or *Knowledge* components;

- *Knowledge*, which is related to operational and technical characteristics and constraints associated to some parts or the whole system;
- *Distributed diagnose*, which interact with other upper layer agents to collaborative identify and diagnose the whole system conditions.

5 Discussion of Results and Critical View

The described approach is being designed and verified based on a case study in the context of an electric micro grid comprising some wind turbines and photovoltaic panels. The preliminary results showed that producer agents (PA) are capable to perform distributed analysis of the energy production and weather data from sensors installed in the energy production units. PAs were able to monitor the operational conditions of production units, in order to identify abnormalities in energy production, by performing short-term prediction of energy production using different data analysis models, which were built based on historical data. 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 in a near future, which could be used by engineers, grid operators and other systems to enhance and optimize the energy distribution and balance. During energy predictions, agents were able to continuously evaluate and improve their analysis models [19].

The experiments performed so far only covered some of the lower agents' aspects of the proposed approach. The preliminary experiments showed promising results, but there are still many features and requirements that need to be addressed in order to verify the expected potentials and benefits. Moreover, these features need to be evaluated by its performance, robustness and scalability regarding the data analysis aspects. Other aspects that should be explored in this case study, comprise the development of predictive capabilities for consumer and storage agents in order to manage energy consumption and power storage of micro grids nodes.

6 Conclusions and Future Work

In industrial domain, MAS have been used as a suitable approach to design 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, the proposed approach intends to combine the features of these two technologies to contribute for the realization of CPS principles, in order to attend the requirements imposed by the Industrie 4.0 vision. This work describes an agent-based data analysis approach for intelligent and adaptive industrial supervisory control systems. Moreover, the approach covers the requirements of process monitoring and supervision automation levels. Although the promising perspectives, it is clear that to achieve the desired objectives, some aspects and issues need to be addressed, namely the dynamic, openness and rapid response requirements of industrial environments, and mechanisms for distributed, cooperative and self-improving data analysis.

Future works encompass the detailed specification and the definition of the required mechanisms and strategies to cover the more advanced aspects and features of the proposed approach. Thereafter, the current case study should be further explored, extending the preliminary experiments in order to validate and assess other aspects. Moreover, it is intended to explore another case study scenario in the manufacturing domain.

References

1. Drath, R., Horch, A.: Industrie 4.0: Hit or Hype? [Industry Forum]. IEEE Industrial Electronics Magazine, vol. 8, nº 2, pp. 56--58 (2014)
2. Lee, E.: Cyber Physical Systems: Design Challenges. In 11th IEEE International Symposium on Object Oriented Real-Time Distributed Computing, pp. 363--369 (2008)
3. Wooldridge, M.: An Introduction to Multiagent Systems. John Wiley & Sons (2009)
4. 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)
5. 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)
6. Sharma, A.B., Ivančić, F., Niculescu-Mizil, A., Chen, H., Jiang, G.: Modeling and Analytics for Cyber-Physical Systems in the Age of Big Data. ACM Sigmetrics Perform. Eval. Rev., vol. 41, nº 4, pp. 74--77 (2014)
7. Chryssolouris, G.: Manufacturing Systems: Theory and Practice. Springer Science & Business Media (2013)
8. Leitão, P., Karnouskos, S.: Industrial Agents: Emerging Applications of Software Agents in Industry. Morgan Kaufmann (2015)
9. 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)
10. 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)
11. Qin, S.J.: Survey on Data-driven Industrial Process Monitoring and Diagnosis. Annual Reviews in Control, vol. 36, nº 2, pp. 220--234, (2012)
12. Gröger, C., Niedermann, F., Mitschang, B.: Data Mining-driven Manufacturing Process Optimization. In Proc. World Congress on Engineering, vol. 3, pp. 1475--1481, (2012)
13. 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)
14. Choudhary, A. K., Harding, J. A., Tiwari, M.K.: Data Mining in Manufacturing: A Review Based on the Kind of Knowledge. Journal of Intelligent Manufacturing, vol. 20, nº 5, pp. 501--521 (2009)
15. 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)
16. 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)
17. Albashiri, K. A., Coenen, F., Leng, P.: EMADS: An Extendible Multi-agent Data Miner. Knowledge-Based Systems, vol. 22, nº 7, pp. 523--528 (2009)
18. 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)
19. Queiroz, J., Dias, A., Leitão, P.: Predictive Data Analytics for Agent-based Management of Electrical Micro Grids. In Proceedings of the IEEE IECON'15, pp 4684--4689 (2015)