

An Approach for Implementing ISA 95-Compliant Big Data Observation, Analysis and Diagnosis Features in Industry 4.0 Vision Following Manufacturing Systems

Kevin Nagorny¹(✉), Sebastian Scholze¹, José Barata², and Armando Walter Colombo³

¹ ATB - Institut für angewandte Systemtechnik Bremen, 28359 Bremen, Germany
{nagorny, scholze}@atb-bremen.de

² NOVA University of Lisbon, 1099-085 Lisbon, Portugal
jab@uninova.pt

³ University of Applied Sciences Emden/Leer, 26723 Emden, Germany
awcolombo@technik-emden.de

Abstract. Current trends are showing a technological evolution to an unified Industrial Internet of Things network where smart manufacturing devices are loosely coupled over a cloud to realize comprehensive collaboration and analysis possibilities, and to increase the dynamic and volatile of manufacturing environments. This rising complexity generates also higher ranges of error possibilities and analog a growing demand of new diagnostic approaches to handle also those highly complex systems as manufacturing systems which are following the Industry 4.0 vision. This is an ISA'95 compliant approach of a Big Data analytics methodology for analysis and observation in Industry 4.0 vision following manufacturing systems.

Keywords: Engineering · Big data · Cyber-physical systems · Context sensitivity · Diagnostics · Data mining · ISA-95

1 Introduction

By analyzing latest reports focusing on predictions of future data generation-consuming-traffic, a doubling of data growth every two years is foreseen [1]. This trend is also valid in the manufacturing domain. The Industry 4.0 vision [2] aims to establish an industrial infrastructure (the industrial internet of things (IIoT)) in which all things are able to exchange information over a network respecting the legacy ISA-95 compliant enterprise architecture. Given that the required connectivity and interoperability of the manufacturing things are guaranteed, the manufacturing environment appears as a very big data set that represents in a digital form the industrial system behind.

One future challenge is the exploitation growing complex data amounts in dynamic manufacturing environments. This work presents an approach to exploit the data through Big Data analysis for observation, analysis and diagnosis through identification, classification, filtering and analysis of data in Industry 4.0 following ISA'95 compliant manufacturing systems.

At the same time this work presents an initial idea for a potential dissertation to reach the Ph.D. degree. Following described research question and hypothesis are base for this paper.

Research Question: Is an Industry 4.0 following Big Data observation, analysis and diagnosis approach based on classical Big Data analysis technologies useful to observe, analyze and diagnose Industry 4.0 vision following manufacturing systems as well as immigrated conventional manufacturing systems?

Hypothesis: Selected existing Big Data analysis technologies are adaptable and extendable to integrate Big Data observation, analysis and diagnosis functionalities into Industry 4.0 vision following manufacturing systems as well as in conventional manufacturing systems to observe, analyze and diagnose Industry 4.0 vision following manufacturing systems as well as immigrated conventional manufacturing systems.

The following chapters are structured as following: Sect. 2 describes relationship of this work to cyber-physical systems. Section 3 is a State of the Art (SotA) and describes related works which will be used as base for this approach. Section 4 describes the overall concept of the approach, Sect. 5 describes some application scenarios and Sect. 6 concludes the paper.

2 Relationship to Cyber-Physical Systems

The work describes an approach to implement ISA-95-compliant Big Data observation, analysis and diagnosis features in Industry 4.0 vision following manufacturing systems. The whole idea of Industry 4.0 is based on cyber-physical systems and the internet of things idea. This work will present how Big Data produced by cyber-physical systems in a service-based industrial internet of things are exploitable for observation, analysis and diagnosis use cases in the manufacturing domain.

3 State of the Art and Related Works

3.1 Analysis and Diagnosis in Manufacturing

Many analysis and diagnosis approaches are available in the manufacturing area. Behind classical manual diagnosis are often individual diagnosis solution used for manufacturing systems. Researcher are working on new approaches based on e.g. (automatic (predictive)) failure detection through fault tree analysis (FTA) [3] or self-learning approaches [4] handle also systems with a higher complexity (evolvable and emergent systems with changing physical and logical conditions and behaviors).

Self-learning approaches are observing systems to learn the behavior of a system and to detect based on knowledge base systems anomalies and failures. In deterministic finite automaton (DFA) [5] algorithms observing systems to build deterministic finite automats which represents a correct behavior of the system. In case that the system gets into a state which is not covered by the model (generated in the learning phase) will be made a defined reaction as e.g. sending an alarm or to start a troubleshooting

routine [3]. Past results of this topic will influence the implementation of this approach. The relation is the idea of knowledge bases and the idea of building trees.

In the concept chapter will be shown that so called “Big Data profiles” are based on these ideas. The approach is able to identify anomalies which will be reviewed by humans and saved into a knowledge base so that time by time the knowledge will be increased. Big Data profiles describe among others a specific system state based on Big Data analyses but also possible progression which could end in this specified state. Fault tree analysis combined with the described deterministic finite automaton approach identifies errors very similar but without statistical Big Data analyses.

3.2 Data Mining and Big Data

Data mining is a broad interdisciplinary research area which covers the idea to extract for implicit, previously unknown, and potentially useful information from data [6]. Data Mining is the analysis step in the process for knowledge discovery. The process is mostly similar and can be divided into four steps [7]: (1) Focusing and selecting of the potential useful data, (2) pre-processing of the data (data cleaning and data completion), (3) transformation of data into a fitting format, (4) data mining analysis itself and (5) the evaluation/knowledge discovery of the data mining analysis results. For data mining there are a lot of algorithms and approaches available [8]. This approach aims to use latest outcomes of pattern recognition approaches as base for data mining.

One sector for Data Mining is Big Data mining. Big Data is one of the most promising technologies nowadays [9, 10]. The definition of Big Data is still in discussion and many suggestions were made. Gartner made a proposal in 2011 where was suggested to categorize Big Data through 3Vs (Volume of Data, Variety of Data and Velocity of Data) [11]. This definition is mostly accepted.

In the last years came up a range of technologies to deal with Big Data. One popular technology is the Googles “Map Reduce” Algorithm which is a programming model for parallel processing of big data sets [12]. This algorithm was implemented in the free Apache Hadoop framework which is an old solution but up to now the base/core element for many Big Data technologies. Hadoop provides basically the Hadoop Distributed File System (HDFS) and the Map Reduce Algorithm. Optionally it provides several extensions. This work will use past approaches of data mining and Big Data analysis and will extend, adapt and improve them to use these approaches in the manufacturing domain.

4 Overall Concept

This section describes a Big Data observation, analysis and diagnosis approach for Industry 4.0 vision following manufacturing systems. The approach aims to generate synergies between state-of-the-art technologies and approaches of diagnosis, pattern recognition, Big Data analysis and context extraction, to generate ISA-95 compliant functionalities for Big Data observation, analysis and diagnosis for Industry 4.0 vision following manufacturing systems.

Figure 1 shows the overall concept. The Approach is divided into a data storage part, a runtime part (where run-time modules will be deployed and executed) and an engineering part (where all engineering tools will be provided for Big Data analysis, observation and diagnosis). The approach provides functionalities for a Big Data analysis and diagnosis, and for continuous (automatic) Big Data observation. All parts and how these parts work together will be explained in the following.

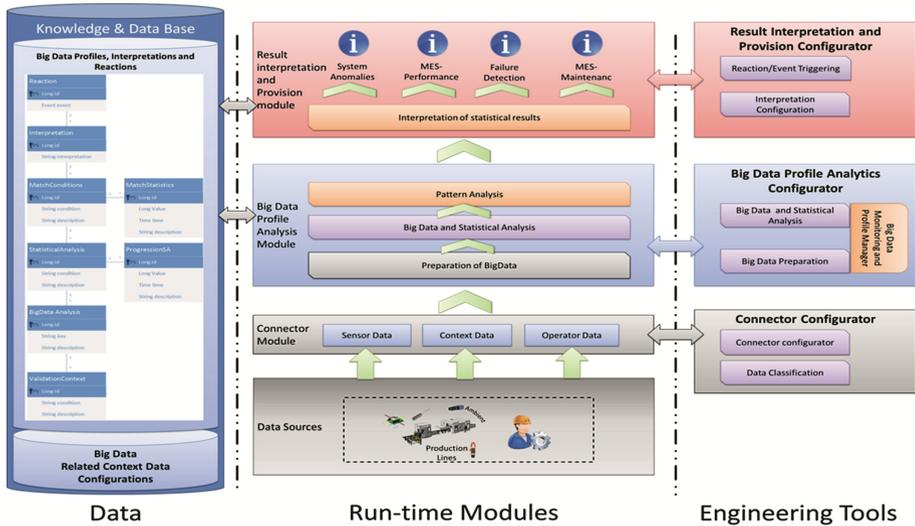


Fig. 1. Concept of a new Big Data analysis and observation approach for large-scale manufacturing systems.

4.1 Data Part

In the Data part will be stored all necessary data in a knowledge and data base. Saved will be data from data sources, related context information of data sources (location, which system, test mode y/n, etc.), Big Data Profiles for a continuous Big Data analysis and configurations (/settings) for engineering tools and run-time modules.

Input Data of data sources could be Big Data buckets (one-time/manual data downloads) for a manual analysis or could be streamed Big Data inputs (event-based of time-frame based for a continuous observation).

A pattern recognition defined in a Big Data profile can be interpreted in several ways which will be saved as interpretations in the knowledge and data base – and different interpretations are interesting for different Big Data analysis solution users (e.g. humans or systems which are connected to the Big Data analysis solution (or more precise: connected to the “Result interpretation and provision module”). Therefore “reactions” will be stored which are linked to interpretations. A reaction defines who will be how notified in case that a defined recognized pattern was found through a Big Data analysis (also called Big Data profile match).

Big Data profiles will be used to observe systems based on continuous Big Data analysis. Big Data Profiles are defining Big Data itself, statistical analyses and define conditions of a Big Data profile match to detect defined system states. Through saved pattern progressions (see Fig. 2) it is also possible to predict automatically future system behaviors and probable future Big Data profile matches.

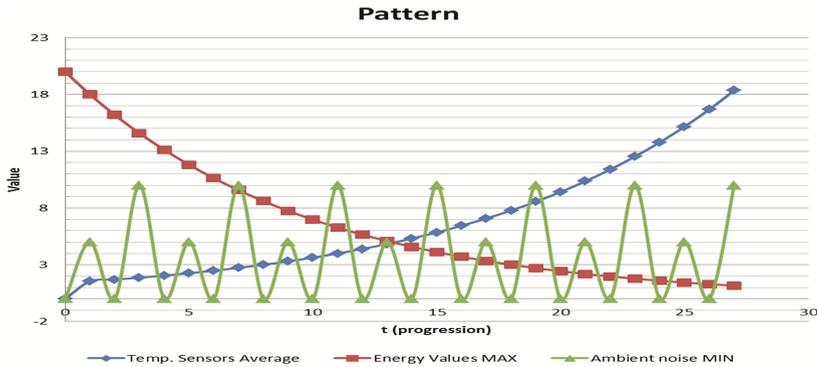


Fig. 2. Example Pattern of a Big Data Profile Pattern progression which describes an Anomaly – t27 shows (e.g.) an error state and t0–t26 shows the progression which can be used as base for future predictions. Further context data could specify under which conditions this pattern is valid or not.

4.2 Engineering Tools

This section describes the engineering tools for each module in the run-time part. The Engineering tools configure and use functionalities of run-time modules.

Connector Configurator - For the provision of data for a Big Data analysis will be used connector modules. Features/functionalities of a connector are (1.) the integration of conventional systems into a network of industry 4.0 vision following manufacturing systems (individual/application specific part to integrate proprietary interfaces) and (2.) the selection, filtering and preparation of usable system data for Big Data analysis, and (3.) the configuration of the communication with a Big Data Profile Analysis Module (data provision, event triggers, data streams, etc.).

Big Data Profile Analytics Configurator - The Big Data Profile Analytics Configurator is divided into (1.) a big data preparation, and (2.) a big data and statistical analysis configurator. These features will be used by the Big Data Monitoring and Profile Manager tool and for the Profile management. The big data preparation procedure will provide functionalities to manage input data from connectors. After selecting needed data follow the Big Data analysis. It will be provided a range of Big Data analysis algorithms which are depending on used technologies and the kind of data input (buckets or streams). The following statistical analysis step will provide functionalities to use several statistical methods (e.g. Gaussian distribution, runaways, averages, etc.) which are usable on Big Data analysis results. Provided will be also functionalities to check

the progression of statistical results. These features are useful to identify e.g. causes of a progression and will be later used for automatic predictions.

Result Interpretation and Provision Configurator - The last engineering tool is the result interpretation and provision configurator. This tool provides functionalities to configure interpretation for Big Data Profiles and to configure interpretation related reactions. As already described could the match of a Big Data Profile have several interpretations. This tool will provide functionalities to manage, generate, edit or delete interpretations and makes it possible to connect them with profiles. In case of a Big Data profile match the interpretations will be triggered. A match of a profile could have e.g. following interpretations (simple examples): “power consumption to high”, “performance is going down”, “throughput is going down”, “wear increases”, etc. Different interpretations are interesting for different users: “the energy provider wants to know that the energy consumption goes up”, “the MES system wants to know that the performance goes down”, “the Factory manager wants to know that the throughput goes down”, “the maintenance operator wants to know that the wear increases”, etc. To notify users (other systems or user) this tool will provide functionalities to configure reactions.

4.3 Runtime Part

The modules in the runtime part of Fig. 1 (expect data sources) represent the engines of the Big Data analysis approach which will be executed for a (continuous) big data observation/analysis and diagnosis in a manufacturing environment. The modules of the runtime part will be described in the following.

Data Sources - Big Data solutions are using poly-structured input data (structured data, semi structured data as XML or HTML, and unstructured data as pictures or documents) as base for their Big Data analysis. Data Sources for Big Data solutions in the manufacturing area related to maintenance tasks are mainly data from systems settled in the levels 1–3 of the ISA-95. Those information are e.g. direct sensor data from the level 1; productions process monitoring data from (SCADA) systems of the level 2; or monitoring data of the whole production process as well as scheduling, quality management, maintenance and production tracking data from (MES) systems settled in the level 3. Other data are coming from context information sources to recognize also information about the context of a system as e.g. the location, the ambient temperature, etc. Context Data will be used to get additional information of big data sources in order to make Big Data Analyses sensitive and reactive to environmental happenings. Saved Big Data profiles will be the base for continuous Big Data analyses but not under all context conditions it is necessary to observe a system so that context related conditions will be used to check the validity of a big data profile.

Connector Module - The connector module establishes the communication between the Big Data Profile Analytics Module and data sources and is responsible for the data provision.

Big Data Profile Analytics Module - This module provides Big Data analyses features and executes procedures for an (automatic) continuous Big Data analysis. The run-time module is divided into (1.) a preparation of BigData part and (2.) into a Big

Data and Statistical Analysis part. The Big Data analysis part will execute the Big Data analysis based on Big Data analysis configurations which are defined in Big Data profiles. After the Big Data analysis will be sent the results to the Statistical Analysis step. In this step will be executed the statistical analysis based on defined Statistical Analysis configurations defined in Big Data profiles. The pattern analysis step will check if a saved Big Data profile matches with the current Big Data statistics analysis results and will trigger in that case the Result interpretation and provision module in case of a match.

Result interpretation and provision module - This module interprets configured Big Data Profile matches and informs users (systems, operators, etc.) which subscribed interpretation related events.

5 Application Scenarios

This work is very application driven and should bring Big Data observation, analysis and diagnosis into ISA'95 compliant Industry 4.0 vision following manufacturing systems. There are several possible application scenarios for such an approach in case that the concept will proof the hypothesis. It follows a short list of examples.

System Anomaly/Failure detection and prediction (detected and predicted system anomalies/failures.), *Process Optimization* (system parameter values which have a positive effect of the system can be identified), *Maintenance Plan Optimization and Individualization* (optimize and individualize maintenance plans based on Big Data analyses), *Energy Efficiency Optimization*, *Decision Support* (support strategic decisions, improvements) or *Security* (through system anomaly detection can be identified system manipulations through hacking activities).

Specially focused in this paper were the needs related to manufacturing systems but there are also several other potential use cases in other domains as supply chain management, product development and improvement or product individualization.

6 Conclusions

This paper presented an approach for implementing ISA 95-compliant Big Data observation, analysis and diagnosis in Industry 4.0 vision following manufacturing systems. During this project it is aimed to proof and validate that based on the hypothesis that such an approach is suitable for application in future Industry 4.0 vision following manufacturing systems. Results of this work will bring Big Data observation, analysis and diagnosis features into future smart factories and will represent on possibility to handle the growing data amounts and complexity in such systems. There are also visible critical points: Big Data analysis can recognize changes and identify context automatically but the interpretation is an application specific issue and will need mostly human support to define goals and to build a knowledge base. It is a future challenge to develop intelligent algorithms which are able make the interpretation of anomalies identified by Big Data analyses. In this approach the interpretation (the diagnosis) will be supported by human experts.

The validation and hypothesis proof of this approach will be made in a realistic flexible manufacturing system located in the University of Applied Sciences Emden/Leer. The system consists of conventional manufacturing components and will be transferred to an Industry 4.0 vision following service-based environment as base for a validation.

Acknowledgment. This work is partly supported by the ProSEco (Collaborative Environment for Eco-Design of Product-Services and Production Processes Integrating Highly Personalized Innovative Functions) project of European Union's 7th Framework Program, under the grant agreement no. NMP2-LA-2013-609143. This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

References

1. Gantz, J., Reinsel, D.: The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East Executive Summary: A Universe of Opportunities and Challenges. International Data Corporation (IDC) (2012)
2. Drath, R., Horch, A.: Industrie 4.0: hit or hype? [Industry Forum]. IEEE Ind. Electron. Mag. **8**, 56–58 (2014)
3. Maier, A.: Online passive learning of timed automata for cyber-physical production systems. In: 12th IEEE International Conference on Industrial Informatics (INDIN 2014), pp. 60–66 (2014)
4. Stokic, D., Scholze, S., Barata, J.: Self-learning embedded services for integration of complex, flexible production systems. In: IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society, pp. 415–420 (2011)
5. Brookshear, J.G.: Theory of Computation: Formal Languages, Automata, and Complexity. Benjamin/Cummings Publishing Company, Redwood City (1989)
6. Leung, C.K.S., MacKinnon, R.K., Fan, J.: Reducing the search space for big data mining for interesting patterns from uncertain data. In: 2014 IEEE International Congress on Big Data (BigData Congress), pp. 315–322 (2014)
7. Ester, M., Sander, J.: Knowledge Discovery in Databases: Techniken und Anwendungen. Springer, Heidelberg (2013)
8. Grossman, R.L., Kamath, C., Kegelmeyer, P., Kumar, V., Namburu, R.: Data Mining for Scientific and Engineering Applications. Springer, New York (2013)
9. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., et al.: Big data: the next frontier for innovation, competition, and productivity (2011)
10. Lohr, S.: The Age of Big Data. The New York Times, 11 February 2012
11. Beyer, M.: Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data, 27 June 2011. <https://www.gartner.com/newsroom/id/1731916>
12. Dean, J., Ghemawat, S.: MapReduce: simplified data processing on large clusters. Commun. ACM **51**, 107–113 (2008)