

Big Data Technology for Resilient Failure Management in Production Systems

Volker Stich, Felix Jordan^(✉), Martin Birkmeier, Kerem Oflazgil, Jan Reschke, and Anna Diewes

Institute for Industrial Management, RWTH Aachen University (FIR), Aachen, Germany
{volker.stich, felix.jordan, martin.birkmeier, kerem.oflazgil, jan.reschke, anna.diewes}@fir.rwth-aachen.de

Abstract. Due to a growing complexity within value chains the susceptibility to failures in production processes increases. The research project BigPro explores the applicability of Big Data to realize a pro-active failure management in production systems. The BigPro-platform complements structured production data and unstructured human data to improve failure management. In a novel approach, the aggregated data is analyzed for reoccurring patterns that indicate possible failures of the production system, known from historic failure events. These patterns are linked to failures and respective countermeasures and documented in a catalog. The project results are validated in three industrial use cases.

Keywords: Failure management · Big data · Complex event processing · Production control · Pattern management

1 Introduction

The amount of data generated in production companies is continually growing. One reason for this development is the advancing integration of system control and measurement utilities within the production, due to new cost-efficient, high-performance information technologies. These allow for an intelligent connection of different production systems units and in general an increased interconnectedness of the production systems in total. The idea of interconnected machines and the overall production integration is labelled as “*Industry 4.0*” in Germany. Industry 4.0 aims at the systematic network integration of machines to make efficient use of the company’s available information resources [1]. As part of this development, the value of production data and the hereby generated information has obtained increasing value for a company. The following approach illustrates a new strategy on how big amounts of data can be systematically used for a failure management system in production.

In a world with complex production procedures and globally operating corporate groups, an efficient failure management system can be a significant advantage in competition. With downtime costs on average as high as 22.000\$ per minute, failures should be avoided or at least detected as soon as possible [2].

The research project BigPro addresses this issue by creating a Big Data driven, proactive failure management system, capable of processing various data from the production environment. Within the platform, the generated production data will be analyzed for data patterns that indicate possible failures in the production system.

2 Literature Review

2.1 Big Data

In 2001 the Meta Group (Gartner) proposed a report about future data management, proposing the three dimensions: Variety, Velocity and Volume [3]. The term Big Data was not yet invented, but the data classification into the three V's prevailed and has been supplemented in 2013 by the dimension Veracity in an IBM study [4].

The dimension *Volume* is still the most common perception of Big Data and describes the amount of data that is generated and processed, at times comprising of petabytes of data. *Velocity* describes the speed at which the data is generated and processed, with special emphasis on the increasing significance of real-time data transmission. The fact that the majority of data is unstructured or semi-structured is regarded by the dimension *Variety*. The newly introduced dimension *Veracity* covers the aspect of the uncertain quality of data and the outcome of data analyses, taking into account that data is partially imprecise, nuanced, and may be redundant or incomplete [5].

Big Data introduces new capabilities of data storage, processing, and analysis. With increasing data sources in companies the available data in companies exceeds their processing capabilities. This not only holds true for the data volume, but also for its variety. With roughly 80 % of the data being unstructured or semi-structured, the ability to consider all kinds of data for analytical tasks, infused by Big Data technology, is of great importance for a company's success. The processing data in real-time is another important aspect that makes Big Data technology capable for failure management systems, since a short reaction time to identify failures is of crucial economic importance for production companies [6].

2.2 Complex Online Optimization and Complex Event Processing

The term *complex online optimization* summarizes hard to solve optimization problems with high response time requirements while including different decision makers and project phases [7]. These challenges exist especially when a failure occurs or is suspected and the production system needs to be stabilized. In most classic failure management approaches, production managers try to cushion failures by including buffers within the production plan. However, there are new approaches which introduce a dynamic component to adapt production plans to occurring failures, e.g. simulation-based rescheduling. Most of these approaches concentrate on a particular machine, ignoring the succeeding production steps and the changes that come with the adjustment of the production plan for the following machines. The BigPro approach includes different kinds of data from several decision-making levels to create a comprehensive failure management in production. This approach includes not only the already mentioned rescheduling

concept, but also approaches of event-based failure identification and prevention activities, which are part of an automated data analysis.

Complex event processing (CEP) describes the direct tracking, processing and analyzing of data streams in near real-time. The aim of complex event processing is to gain insight in data patterns and identify meaningful business events within a complex data context [8, 9]. The advantage of complex event processing is that these event streams can be processed directly on the data stream. This technology shows great potential for the use in an intelligent and agile production, where great amounts of data from different sources such as sensor-data streams, service data and external data need to be analyzed on-the-fly. In BigPro this technology will be used to analyze failure patterns to initialize preventive actions. Here, not only current but also past event patterns are considered to create a larger information basis and make the forecasting system more reliable and resilient.

2.3 FMEA Incident Management

The failure mode and effects analysis is an established systematic technique, used to identify and analyze failures and failure types. The FMEA analysis enables the detection of failure possibilities and weak points within a process and identifies proactive measures to prevent these failures [10]. Furthermore, FMEA optimizes existing processes and can even be used to bundle all information regarding past detected failures and their connection for further use. The FMEA method therefore is a suitable tool to define failure groups as part of the reactive failure management in BigPro.

2.4 Mood Tracking and Sentiment Analysis

Monitoring human related data such as emotions and physical activities have gained increasing awareness in many different research areas [11]. However, stress management in production context is a rather new research area. Due to newly developed biosensors it is possible to measure different parameters such as heart rate variability, heart beat or skin conductance which are reliable indicators for stress. This information can be merged in a production environment to identify stressful situations and prevent failure or production downtime by taking measures accordingly.

Sentiment Analysis refers to the analysis of written human interaction to identify the emotional state of the author, at the time the message was written. A message can contain not only an informative, but also an emotional message [12]. The analysis of human data will be included in BigPro as another potential failure indicator to gain better insight into the production system, and improve failure management.

2.5 Identified Research Gap

The integration of Big Data technology into a failure management system has not yet been put to the test. This enables the merge of structured and unstructured data in a production context and to create a more precise virtual image of the production system. It also requires more sophisticated CEP algorithms to better process and merge

structured and unstructured data in a failure management context. To ensure portability of the solution, another challenge is to cover three different use cases with very distinct information systems and business cases.

After the data is processed and a potential failure is recognized, a user-oriented visualization is necessary to suggest or initiate countermeasures. Depending on the failure's seriousness and impact on production, countermeasures need to be taken by persons from different hierarchies with different authorization levels in the company. Hence, a user-oriented visualization (management decides on aggregated information, while production workers require actual information on machine status) of failures needs to be developed. Furthermore, the integration of human data as an indicator for failures requires new data privacy concepts.

3 Big Data for Production Failure Management

3.1 Failure Recognition with Big Data to Increase Production Resilience

To detect possible failures all production data (e.g. sensor data, order data from the ERP system and other information systems, production environment data, ...) will be gathered and analyzed. As part of the BigPro approach the influence of the persons within the production - the heart of a production- will be considered as well. In fact, the worker's input and his working experience is of great importance to gain better insights in the production system's condition. Unusual observations such as growing noise emission or oil leakage stay mostly unnoticed, but can be detected by experienced workers. As part of this project, different sources of human input are tested regarding their failure management suitability: text analyses of intranet department news, maintenance comments as well as voice recognition within the production itself are potential data sources.

Human data, as well as data from production assets will be automatically analyzed and handled by complex event processing methods. In addition, not only current but also past information from failure situations are processed to detect reoccurring patterns and improve the platform's failure forecasting capabilities. After patterns are detected, the probability of an occurring failure is determined to define the data's quality and to decide whether correcting actions will be taken.

Applying Big Data – technology to the production data allows for the consideration of all data (structured or unstructured) relevant for the production process, making the digital representation of the production more comprehensive. Thus, the more data and information is available in real-time, the better the planning, controlling and managing of production systems can be performed, while responsiveness to unforeseeable events increases. All these aspects pave the way to a more resilient production system suffering of fewer unplanned production downtimes.

3.2 Big Data for Failure Prevention and Reaction Management

After patterns have been detected, adequate countermeasures need to be defined for the pro-active character of the failure management system in BigPro. As a supporting tool

for the creation and evaluation of specific reactive actions the FMEA analysis will be used. For known patterns a reactive action will be defined in the failure management platform and documented in a countermeasure catalog. For an identified pattern with a high probability rate the previously defined countermeasure might be initiated automatically by the system. Patterns with a lower probability of occurrence can be forwarded to the person in charge as a failure warning with a reaction proposal. Thus, the risk of the production from going into downtime is reduced. The catalog will be extended in an ongoing validation process. To eventually use this technology for different production branches cross-sector solutions need to be generated.

3.3 Failure Visualization

As a subordinate theme this research aims to visualize information about possible failures, their urgency, and possible reasons with proposed countermeasures.

Information should be visualized differently for different groups of employees. While the production manager needs failure notifications about urgent failures, the machine operator needs all types of information about the machines that are in his area of responsibility. He also needs a different degree of information and is used to more technically detailed information. His information may include information about resource shortage or signs of increasing wear as well as a drop in oil pressure. This personalized way of failure visualization creates a more transparent and user-oriented workflow while increasing efficiency of the failure management system.

4 BigPro for a Resilient Failure Management in Production

The project BigPro unites new data processing approaches with an emphasis on failure management strategies. The aim of the project is the creation of new usable concepts and tools in regards to the failure detection, failure handling and failure visualization. The project takes place in close collaboration with three project partners to test created solutions in action within their production systems.

The project partners are of varying size with a range of different production systems to study and ensure the manifold application possibilities of the BigPro platform.

4.1 The Overall Approach

Information plays an important role in this project. To realize an effective and efficient failure management system, it is important to consider the right pieces of information in the right context. The project's Big Data approach allows for the consideration of all kinds of data and information, without the need to specify relevant information beforehand. Thus, all available data can be gathered, analyzed, and used in the BigPro platform for a data-based failure management.

BigPro will extend the data processed for analysis from the production environment (production machine data, environmental data, and order data) with unstructured, human data. Thus, a more complete digital image of the production is gained. Impressions, such

as unusual machine noises or flawed machine operations are difficult to track with ordinary sensors. BigPro will be capable to capture and understand human input, and will use this additional information for the failure management.

The overall goal of BigPro is to enable a pro-active failure management for producing companies. This goal is carried out by developing algorithms for data pattern analysis. These algorithms examine existing data pools for patterns during production failures. Detected data patterns will be correlated with the related failure and included in the catalog of countermeasures. BigPro platform will use this data base to compare current data stream from the production environment with the known patterns. In case of a match, the system will warn, that a specific failure might occur. If a known and established countermeasure is recognized, it will be suggested to the responsible user.

Next, to initialize and conduct pro-active or re-active countermeasures, it is important to identify the appropriate management/decision level to address the failure. Here, it is important to provide the required information visualized user-oriented and in the right aggregation level.

The project comprises of the following tasks to implement a Big Data platform for failure management in production systems:

- Creating an information landscape for each use case, and developing a concept to determine data and information reliability for the failure management system,
- Evolving algorithms for CEP data pattern management as basis for a pro-active failure management,
- Creating an expandable catalog of countermeasures, correlated with identified data patterns, and
- Developing new, user-oriented visualization concepts for different decision levels.

4.2 Use Cases Descriptions

The first use case is part of a research environment to test the interaction between practice and research. Based on a real production environment, electrically powered pedal carts are being assembled in a small-batch production. The factory is equipped with modern machinery and assisted by voice-based systems such as Pick-by Voice commissioning. Due to research activities, the data environment is extended on a regular basis. This leads to a dynamic data generation environment and a high variety and veracity of data. As part of research it is possible to study employees as indicators of disturbance in more detail than in actual companies.

The second partner has started to digitalize its hand moulding shop by installing RFID technology linked to the ERP system to increase process transparency. These data are extended by data pulled from the involved production machines. This use case represents the data availability of a typical SME. The company does not yet have a total failure management system but with up to six weeks of throughput time for each product, it is of the utmost importance, that failures and resulting production disturbances can be avoided.

The production process of the third partner requires the interaction of a high number of production machines, each creating a significant volume of data points that need to

be merged to extend the already existing failure management system. The integration of human created content promises further insights into the production process and its stability.

4.3 Challenges in BigPro

The three use cases and their diverse production and business backgrounds mean a significant challenge to BigPro. Each partner demands for a specific problem solution in a specific context. To ensure transferability of the solution, three measures need to be taken: First, the partner specific problems need to be generalized to examine transferability options. Second, a set of standard BigPro elements to address the generic problems will be defined. These sets comprise of involved information objects, as well as required information sources (e.g. sensors). Third, the catalog's logic to gather countermeasures needs to be receptive for all three partner's requirements.

Further challenges arise from the integration of structured and unstructured data. Especially the aspired inclusion of human-generated content poses a challenge for the BigPro platform. On the one hand, it is necessary to generate data without interfering with the workers' working routines. Thus, analyses were run to identify already existing human interfaces within the treated use cases. On the other hand, there is still the complexity of digitalizing input and processing the retrieved data into context-related content. Therefore, the system will be taught in terms and context by reading in documents and manuals of the respective process.

5 Conclusion and Outlook

An efficient failure management plays an important role for production companies. Scrap and downtime are cost drivers that need to be avoided. Since data and information play an increasingly important role in companies and for decision makers, it seems natural to use data for a failure management system. BigPro introduces comprehensive approach by using Big Data methods for more precise failure detection. A Big Data platform will be developed capable of processing structured and unstructured data, generated in the production environment.

Unlike other approaches, BigPro not only uses data from production machines and environment sensors, but stresses the worker's capabilities to indicate disturbances and failures. By digitalizing the human input, and merging it with machine data on the BigPro platform, the digital image of the production is more complete, and serves a better decision basis. On this basis, data pattern analyses are run to detect looming failures in production. This goal drives another challenge: the combination of historic and real-time data, as well as the correlation with data patterns and related failures.

Finally, a concept for a user-oriented visualization to better support decision makers is required. This concept ensures that only information relevant for a person is shown (management decides on aggregated information, while production workers require actual information on machine status).

In the first project phase the technical and business-driven use case requirements have been gathered, discussed and documented. Next, the BigPro platform will be initiated based on the documented requirements. In parallel, the information landscape is drawn, to identify relevant information objects. Based on the information objects, the data pattern analysis will start.

Acknowledgements. The BigPro project is promoted by the German Federal Ministry of Education and Research (Project number 01IS14011A-I), in the context of the operational program *ICT 2020 – Research for Innovation*. The project will end in August 2017.

Special thanks go to: Asseco AG, C. Grossmann Stahlguss GmbH, Cognesys GmbH, i2solutions GmbH, Robert Bosch GmbH, Software AG, FZI Forschungszentrum Informatik, Werkzeugmaschinenlabor der RWTH Aachen.

References

1. McKinsey Digital. Industry 4.0- How to navigate digitization of the manufacturing sector. Industry report, McKinsey & Company (2015)
2. Industry News. Downtime Costs Auto Industry \$22 k/Minute. <http://news.thomasnet.com/companystory/downtime-costs-auto-industry-22k-minute-survey-481017> (2006)
3. Laney, D.: 3-D Data Management: Controlling Data Volume, Velocity and Variety. META Group Original Research Note (2001)
4. Schroeck, M., et al.: Analytics: the real-world use of big data. IBM Institute for Business Value - executive report, IBM Institute for Business Value (2012)
5. Bloehdorn, S., Fromm, H.: Big Data—Technologies and Potential. Enterprise-Integration (2014)
6. Zikopoulos, P., Eaton, C.: Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data. McGraw-Hill Osborne Media, New York (2011)
7. Grötschel, M., Krumke, S.O., Rambau, J.: Online Optimization of Complex Transportation Systems. Springer, Berlin (2001)
8. Cugola, G., Margara, A.: Processing flows of information: from data stream to complex event processing. *ACM Comput. Surv. (CSUR)* **44**(3), 15 (2012)
9. Grauer, M., et al.: An approach for real-time control of enterprise processes in manufacturing using a rule-based system. In: Multikonferenz Wirtschaftsinformatik (2010)
10. Goebbels, S., Jakob, R.: Geschäftsprozess-FMEA: Fehlermöglichkeits-und Einfluss-Analyse für IT-gestützte Geschäftsprozesse. Symposion Publishing GmbH (2004)
11. Mishne, G., De Rijke, M.: MoodViews: tools for blog mood analysis. In: AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs (2006)
12. Pang, B., Lee, L.: Opinion mining and sentiment analysis. *Found. Trends Inf. Retrieval* **2**(1–2), 1–135 (2008)