

# Big Data Based Analysis Framework for Product Manufacturing and Maintenance Process

Yingfeng Zhang<sup>(✉)</sup> and Shan Ren

Key Laboratory of Contemporary Design and Integrated Manufacturing  
Technology, Ministry of Education, Northwestern Polytechnical University,  
Shaanxi 710072, People's Republic of China  
zhangyf@nwpu.edu.cn

**Abstract.** With the widely use of smart sensor devices in the product lifecycle management (PLM), it creates amount of real-time and multi-source lifecycle big data. These data allow decision makers to make better-informed PLM decisions. In this article, an overview framework of big data based analysis for product lifecycle (BDA-PL) was presented to provide a new paradigm by extending the techniques of Internet of Things (IoT) and big data analysis to manufacturing field. Under this framework, the real-time lifecycle data of products can be active perception and collection. Considering the challenges of processing the lifecycle big data into useful information and exchange it among various lifecycle phase, a graphical model of big data mining was designed to achieve knowledge discovery. Finally, a case has been used to illustrate the proof-of-concept application of the proposed BDA-PL.

**Keywords:** Product lifecycle · Manufacturing · Maintenance · Big data analysis · Data mining

## 1 Introduction

Complex product manufacturers have attempted to adopt Internet of Things (IoT) and big data analysis technologies for improving the efficiency of their PLM, and increasing company's profitability and reduce the consumption of products through establishing Product Service Systems (PSS).

Recent developments in wireless sensors and communication technologies have created a new era of the IoT [1]. At present, more and more manufacturing enterprises began to widespread use IoT technology to implement and manage their business. In PLM, IoT technology can give opportunities to access and manage product data and information over the whole lifecycle [2]. In order to apply the IoT to PLM, it needs several technologies related to product identification, sensors and wireless tele-communication. Jun et al. [3] introduces an overall framework for RFID applications in PLM. Xu et al. [4] focuses on the research of closed-loop product information tracking and feedback in wireless technology-enabled from the modeling point of view. Georgiadis and Athanasiou [5] studied predictive maintenance and remanufacturing application based on closed-loop PLM.

With the widely use of IoT technology and smart sensor devices in PLM, it creates amount of real-time and multi-source lifecycle data. These data were high volume, high velocity and high variety, which belong to typical big data. It has been already caused people's extensive concern in economic, academia and manufacturing. In terms of manufacturing, big data analysis will significant impacts to product Research & Development (R&D), manufacturing, service improvement and recycling, etc. At the same time, it can also effectively promote the development of service-oriented manufacturing, green manufacturing (GM), sustainable production and consumption.

Big data research in manufacturing is still in its infancy in communication. McKinsey [6] published article "*How big data can improve manufacturing*". In-depth analysis of the issues in how big data and advanced analytics to make manufacturing more rationalization. Through the analysis of several cases, to illustrate how big data and advanced analysis applications to provide assistance for business decisions.

In modern manufacturing environments, vast amounts of data are collected in enterprise database from the whole lifecycle. Data mining has emerged as an important tool for knowledge acquisition from the manufacturing databases.

To discover unknown priority dispatching rules for the single machine scheduling problem, Bayesian algorithm are used to find hidden rules through large amounts of structured or unstructured data [7]. To address traditional performance evaluation problem of suppliers in the supply chain, Chen et al. [8] proposes an integrated model by combining K-means clustering, feature selection, and the decision tree method into a single evaluation model. Magro and Pinceti [9] presented a technique to improve the accuracy of the predictions using the rough set theory. Mavridou et al. [10] established a model applied neural network algorithm to identify bearing faults in wind turbines. Purarjomandlangrudi et al. [11] presented a data mining approach called anomaly detection to discriminate defect of rolling bearing failures.

From above literature review, main limitations can be summarized as following:

- Previous researches mainly focus on how to apply the IoT related techniques to one phase of the PLM, and the overall solution for the whole 2 lifecycle were seldom investigated.
- Little effort has been made to apply the solution of big data to manufacturing, especially to PLM or PSS. Lacking of overall system architecture of big data solution.
- Most data mining applications were only focus on one stage of whole lifecycle. Little effort has been made to data integration mining for the lifecycle.

By addressing the above limitations and questions, the research reported in this paper was concerned with big data based manufacturing applications specifically in manufacturing (including Research & Development and Manufacturing, RDM) and maintenance (including Operation and Maintenance, OM) process (hereinafter abbreviate as MMP) of product lifecycle. The focus was placed upon developing system architecture of the big data solution, and discussing the key technology related to the solution.

Several steps are taken. First, an overall framework of BDA-PL was proposed in Sect. 2. Second, key technology related to the framework was discussing in Sect. 3. An application case for axis compressor MMP was designed and demonstrated in Sect. 4. Conclusions were drawn in Sect. 5.

## 2 Overview of Big Data Based Analysis for Product Lifecycle

Figure 1 shows an overview of the proposed framework. This framework was consistent with a normal MMP. That is, for beginning of lifecycle (BOL), a product manufacturing factory has one or more production lines. Different production lines were often responsible for different production processes. For middle of lifecycle (MOL), the product was leased or sold by the manufacturer. While using the product, a customer may face several situations: the product may need maintenance after some use; the product may have some quality problems and need to be returned in compliance with the leasing contact; the product may come to the end of its life.

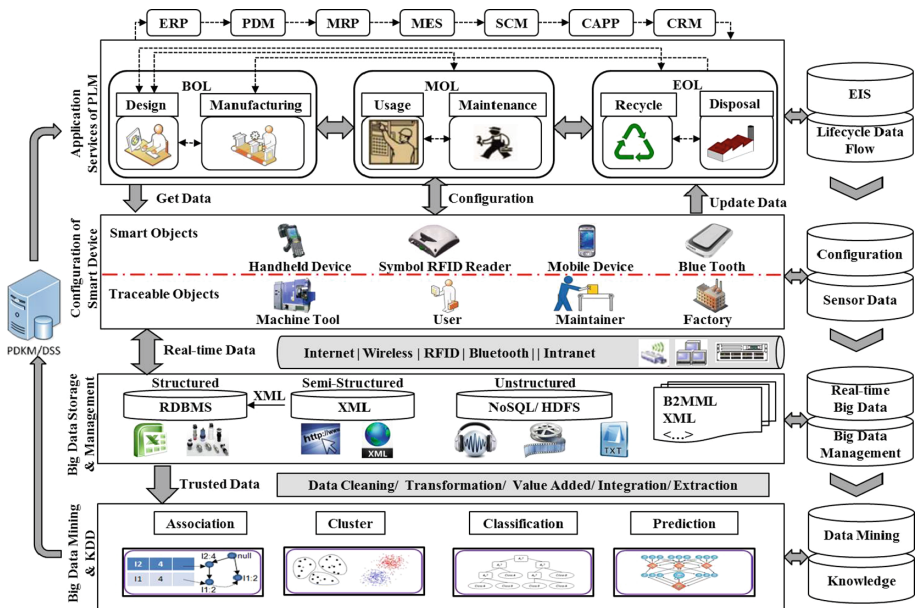


Fig. 1. Overview framework of BDA-PL

According to the big data analysis hierarchy, the proposed BDA-PL framework includes following core components:

- Application service of PLM: This layer was used to provide important real-time applications based on the integrated knowledge that mining and discovery from the lowest layer. It can be regarded as the top-level services of the PLM.

- Big data acquisition and integration: Based on the configuration of the smart device in manufacturing things and product, the real-time and complete lifecycle big data can be sensed and captured.
- Big data storage and management: Product lifecycle data consist of structured, semi-structured and un-structured data. Distributed Data Base System (DDBS), XML and Hadoop Distributed File System (HDFS) can be used to manage and store structured, semi-structured and un-structured data, respectively.
- Big data mining and Knowledge Discovery in Database (KDD): By using the theory of data mining to find valuable knowledge from lifecycle big data. Combined the knowledge with product data & knowledge management (PDKM) system and decision support system (DSS) to form a closed-loop mechanism of knowledge sharing and feedback.

Compared with existing PLM, active perception of manufacturing things, real-time monitoring of product status, value discovery of big data, dynamic optimization of PLM were the significant characteristics of the framework for BDA-PL.

### 3 Key Technologies of Big Data Based Analysis for MMP

Big data based analysis for MMP refers to the method of multiple disciplines such as information, communication, knowledge discovery and intelligent decision, etc. This paper was illustrated mainly on the realize ideas and methods of the key technologies for MMP.

#### 3.1 MMP Big Data Acquisition and Integration

Configuration of the various smart devices or product embedded information devices (PEIDs) (e.g. RFID tags, sensors) for manufacturing things and products were the foundation of real-time and multi-source big data capture.

During production phase, PEIDs can be deployed to manufacturing things, key components and products to collect real-time data of work in process (WIP) and key parts. For example, RFID readers was equipped in the fixed position such as CNC center and key equipment of assembly line, nevertheless, RFID tags was equipped in the manufacturing resources such as key parts and components. During usage and operation stage, these PEIDs and smart sensors that have been embedded (equipped) in the key components or proper locations of the product can be used to real-time monitor and capture the OM big data of product.

A huge amount of real-time data sensed at the PEIDs cannot be directly used in the up-level EISs and lifecycle management layer. So, all data produced by various departments and software have to be integrated. By information fusion technology such as definition of data relation and middleware technology, then transformed the multi-source and heterogeneous data into standard information (e.g. B2MML/XML) that can be used by PLM decision-making to realize multi-source data value-add.

### 3.2 MMP Big Data Mining Mechanism

A graphical model of MMP big data mining was designed as shown in Fig. 2. The model consists of four layers from the bottom to the top.

- Data layer was used to store the big data of MMP. The data were stored in different types of databases according to different application demand.
- Method layer mainly refer to data mining model. It was responsible for extracting suitable data from data layer and mining valuable knowledge from the data.
- Result layer was a set of data mining result. According to different application demand, suitable mining model and original data were selected to conduct data mining. Finally, achieve the knowledge set of different lifecycle stages and different application index.
- Application layer was also a demands layer, namely applied the knowledge to achieve the requirements of enterprises.

What we can see from the above analysis was that the big data mining graphical model of MMP was a close-loop structures that starting from the application demands, and ultimately meets the application of enterprises.

## 4 Case Study

This section demonstrates the usage of the proposed framework with an actual product involved, which was a kind of axial compressor manufactured by company A. The company also provides application-oriented product service for this product, in other words, the company sells its function instead of the product, e.g. through sharing and

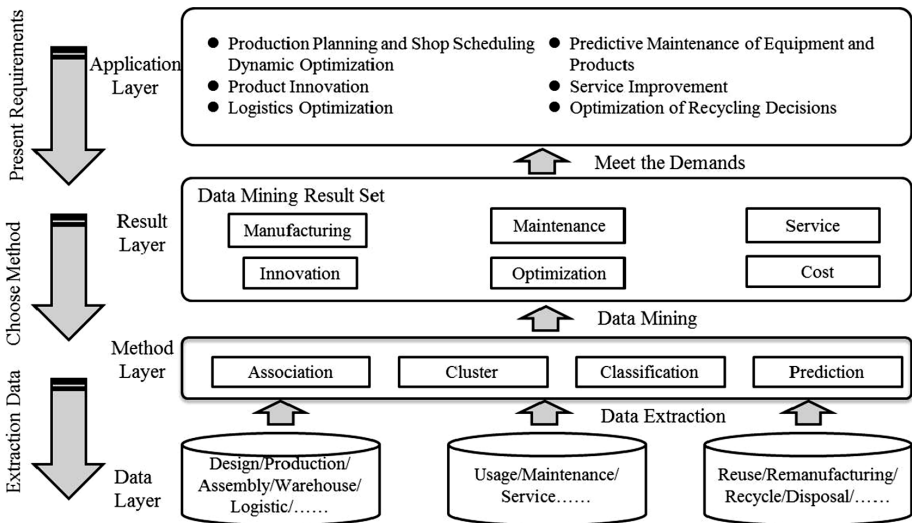


Fig. 2. Graphical model of MMP big data mining

leasing. The usage status data can be real-time collected and remote monitored by company A to improvement product design, optimization production and predictive maintenance.

Here we focus on the MMP for the critical component rotor and blades of the compressor. These two components were particularly suitable since compressor was often works under a condition of high pressure and high temperature, and easily causes the compressor to go into surging zone. When surge occurs, the signals such as flow rate, pressure and temperature, which symbolize the surge, will exhibit phenomena that were impossible under normal conventional conditions.

### 4.1 Compressor MMP Big Data Acquisition

**Create a Smart Environment for MMP with PEID-Enabled Smart Objects.** For simplicity of understanding but without lost of generality, a hypothetical MMP for the compressor as shown in Fig. 4. It consists of two plants and one warehouse, and the compressors were rented by customers. Company A provides specialized service for these customers such as specialized maintenance, remote on-line diagnoses and supply spare parts.

Figure 3 also shows a deployment of PEIDs in the MMP. UHF RFID, RFID tags and smart sensors were adopted to manufacturing resources, key components or key position of axial compressor for tracking the real-time data. The deployment information was shown in Table 1.

The warehouse consists of smart shelves, each of which is equipped with PEIDs (collectively as 1). Locations on the shelves are tagged.

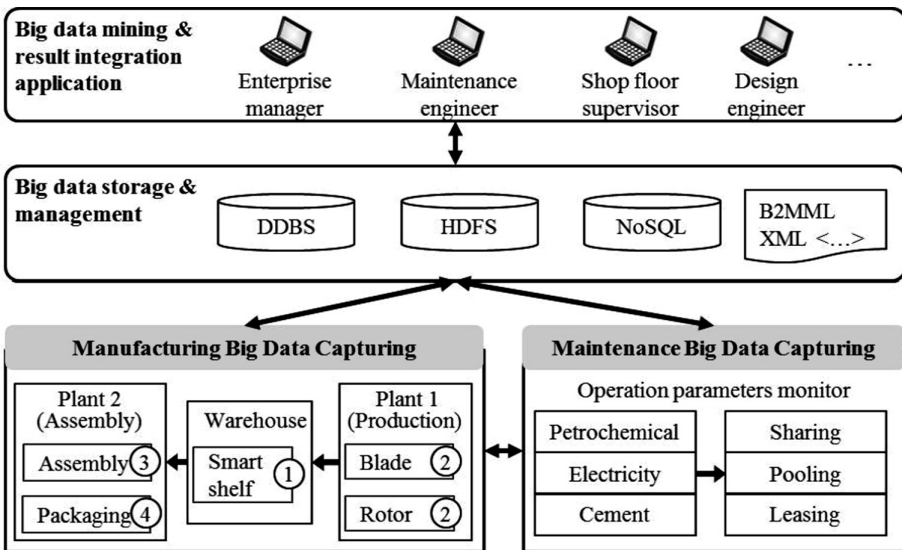


Fig. 3. Overview of the case study

**Table 1.** Deployment information of PEIDs in manufacturing shop floor and product

PEIDs	UHF/Sensors	Manufacturing recourses/Location	Objective
RFID reader	UHF	Machine	Track the pallets
RFID reader	UHF	Assembly line	Check in materials and report assembly tasks
Tag	UHF	Pallet	Track the information of pallet with materials
Tag	UHF	Critical parts	Track the real-time data from WIP to product
Sensor	Differential pressure sensors	Convergence device	Monitor flow rate of gas
Sensor	Piezoelectric velocity sensors	Both ends of rolling bearing	Monitor vibration of rotor
Sensor	Temperature sensors	Embedded in blades and rotor	Monitor temperature of gas

Plant 1 was a parts production plant. Some parts of compressors were manufactured in this plant and delivered to the warehouse. The PEIDs (collectively as 2) were configured during manufacturing process. For example, UHF RFID tags were attached on the blades and rotors during production process to track the real-time data from WIP to products, however, smart sensors were embedded in blades and rotors or configured in the proper locations of the compressor to monitor and collect the real-time field data (e.g. temperature and pressures of the gas). To procure the required real-time field data, the PEID readers were also allocated to the appropriate position near the location of the compressor and its components.

Plant 2 was assumed to be a simplified compressor assembly line. Two workstations were involved in addition to the line supervisor. Accordingly, two PEIDs (marked as 3–4) were deployed to these two workstations. Same as plant 1, sensors can be configured in proper locations of the compressor during assembling process.

**Acquisition of the MMP Big Data.** Based on the configuration of the smart environment for MMP, the real-time primitive data of PEIDs equipped to the manufacturing resources and products can be captured.

When a pallet comes to the machine, this event can be tracked by the UHF RFID reader installed at the machine. Next, UHF RFID reader check materials needed to according to this process task. If the materials were well prepared, the processing task can be executed at this machine. During the process stage, the tagged critical component's status can be tracked by the RFID UHF reader installed at the machine. Therefore, the real-time data of components and manufacturing resources can be captured. The assembly plant is not included for discussion here as the working principles are basically similar.

OM process was the phase where compressors were used by customers. Flow rate data was one of the key data to prevent the surge of compressor. In addition, the vibration of the rotor and the temperature of the compressed gas were also important symbol to surge phenomena. Therefore, through real-time monitored and captured the usage status data such as flow rate, pressure, temperature and vibration, compressor surge can be effectively prevented. As has been noted, some smart sensors have been equipped in the proper locations of the compressor can be used to monitor and collect use status data of compressor.

## 4.2 Data Mining Results Integrated Application of MMP

The big data mining results of MMP were not only useful to themselves, but also useful to other stage of the lifecycle. For example, design departments need to draw lessons from the evaluate result of supplier selection to select better supplier. Maintenance or service departments need to understand the quality factor of production to provide better sale service and maintenance service. Data mining results of marketing and customer demand data can be used by R&D departments to development new products. Easily damaged parts and supplier selection evaluation data mining results of service departments can be used by R&D departments to select better supplier and improvement design.

## 5 Conclusions

This paper has presented an innovative framework where IoT technologies and big data analysis technologies were combined to form a BDA-PL for sharing information and knowledge among actors of PLM. Two contributions are important in this research. The first contribution was the framework of BDA-PL and its key components. Under the new big data based lifecycle management strategy, lifecycle managers can use advanced analytics tool to take a deep dive into real-time and historical process data, and then optimize the factors that prove to have the greatest effect on lifecycle management. The second contribution was a novel conception of combine big data analysis with product service, which was illustrated in the case.

The proposed framework of BDA-PL just provides a new kind of useful infrastructure to improve production and maintenance efficiency by using the MMP big data. Future research works will focus mainly on how to use the advanced big data analytics tool to mining and identify hidden patterns from the sensed MMP big data for optimal production and maintenance decision.

**Acknowledgement.** Authors would like to acknowledge financial supports of National Science Foundation of China (51175435), the program for New Century Excellent Talents in University (NCET-12-0463), and the 111 Project Grant (B13044).



## References

1. Zhang, Y., Zhang, G., Wang, J., Sun, S., Si, S., Yang, T.: Real-time information capturing and integration framework of the internet of manufacturing things. *Int. J. Comput. Integr. Manuf.* (ahead-of-print), 1–12 (2014)
2. Kiritsis, D., Bufardi, A., Xirouchakis, P.: Research issues on product lifecycle management and information tracking using smart embedded systems. *Adv. Eng. Inform.* **17**(3), 189–202 (2003)
3. Jun, H.B., Shin, J.H., Kim, Y.S., Kiritsis, D., Xirouchakis, P.: A framework for RFID applications in product lifecycle management. *Int. J. Comput. Integr. Manuf.* **22**(7), 595–615 (2009)
4. Xu, D.F., Li, Q., Jun, H.B., Browne, J., Chen, Y.L., Kiritsis, D.: Modelling for product information tracking and feedback via wireless technology in closed-loop supply chains. *Int. J. Comput. Integr. Manuf.* **22**(7), 648–670 (2009)
5. Georgiadis, P., Athanasiou, E.: The impact of two-product joint lifecycles on capacity planning of remanufacturing networks. *Eur. J. Oper. Res.* **202**(2), 420–433 (2010)
6. McKinsey & Company. [http://www.mckinsey.com/insights/operations/how\\_big\\_data\\_can\\_improve\\_manufacturing](http://www.mckinsey.com/insights/operations/how_big_data_can_improve_manufacturing)
7. Premalatha, S., Baskar, N.: Implementation of supervised statistical data mining algorithm for single machine scheduling. *J. Adv. Manage. Res.* **9**(2), 170–177 (2012)
8. Chen, Y.S., Cheng, C.H., Lai, C.J.: Extracting performance rules of suppliers in the manufacturing industry: an empirical study. *J. Intell. Manuf.* **23**(5), 2037–2045 (2012)
9. Magro, M.C., Pinceti, P.: A confirmation technique for predictive maintenance using the rough set theory. *Comput. Ind. Eng.* **56**(4), 1319–1327 (2009)
10. Mavridou, E., Kehagias, D.D., Tzovaras, D., Hassapis, G.: Mining affective needs of automotive industry customers for building a mass-customization recommender system. *J. Intell. Manuf.* **24**(2), 251–265 (2013)
11. Purarjomandlangrudi, A., Ghapanchi, A.H., Esmalifalak, M.: A data mining approach for fault diagnosis: an application of anomaly detection algorithm. *Measurement* **55**, 343–352 (2014)