



A framework for shopfloor material delivery based on real-time manufacturing big data

Shan Ren¹ · Xibin Zhao² · Binbin Huang¹ · Zhe Wang¹ · Xiaoyu Song³

Received: 26 January 2018 / Accepted: 27 August 2018 / Published online: 1 September 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Although auto-identification devices such as radio frequency identification and smart sensors have been widely used in shopfloor management and control, major challenges still hinder the vision of real-time and multi-source data-driven material delivery in a manufacturing big data environment. For instance, how to collect manufacturing big data in a timely and accurate manner, and how to discover the hidden pattern from the manufacturing big data rapidly to improve the efficiency of material delivery. To address these challenges, in this paper, a framework for shopfloor material delivery based on real-time manufacturing big data is proposed. Key technologies of the proposed framework are investigated. Firstly, a solution of data sensing and acquisition is designed. Secondly, the methods of manufacturing big data preprocessing and storage are developed to integrate and share the manufacturing data in a unified data format, and to ensure the reusability of the data. Thirdly, a graphic model for the manufacturing big data mining is presented. An improved Apriori-based association analysis model is exploited to identify the frequency trajectories of material delivery. In order to demonstrate the implementation of the proposed framework, a proof-of-concept scenario is designed. The key findings and insights from the experimental results are summarized as managerial implications, which can guide manufacturers to make more informed decisions for the shopfloor management.

Keywords Manufacturing big data · RFID · Material delivery · Data mining · Shopfloor management

✉ Shan Ren
renshan@mail.nwpu.edu.cn

✉ Xibin Zhao
zxb@tsinghua.edu.cn

Binbin Huang
b.b.huang@mail.nwpu.edu.cn

Zhe Wang
894108548@qq.com

Xiaoyu Song
songx@pdx.edu

¹ Key Laboratory of Contemporary Design and Integrated Manufacturing Technology, Ministry of Education, Northwestern Polytechnical University, Shaanxi, People's Republic of China

² School of Software, Key Laboratory for Information System Security, Ministry of Education, Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing, People's Republic of China

³ Department of Electrical and Computer Engineering, Portland State University, Portland, OR 97207, USA

1 Introduction

Shopfloor logistics planning and scheduling heavily rely on the arrival of material (Ning et al. 2016), thus, decisions on logistics trajectory of material delivery are critical to improving the productivity. Generally, the trajectory of material delivery is determined by the quantity of items to be produced and the production capacity of the existing manufacturing system (Huang et al. 2007). Due to the field data of production status cannot be captured and shared within manufacturer enterprise in a timely fashion, there are some disadvantages in traditional manufacturing system such as lagged material delivery and low equipment utilization. Managers involved in the production process may make material delivery decisions based on the incomplete and inaccurate manufacturing data, which could lead to inaccurate decisions and operational inefficiencies (Zhang et al. 2017c). In addition, with the development of the customized production paradigm, manufacturers need to: (1) carry out near real-time information interaction and sharing; (2) make accurate material delivery decisions; (3) shorten the delivery

lead time and provide diverse products for its customers. Therefore, the traditional material delivery methods are no longer applicable to the new production paradigm.

Recently, with the rapid development of Internet of Things (IoT) (Tao et al. 2014; Jararweh et al. 2015), industrial internet (Posada et al. 2015), and Industrial Internet of Things (IIoT) (Tao et al. 2018a) technologies, many manufacturers have adopted these advanced information technologies to implement real-time traceability in improving the performance of shopfloor planning and control (Zhang et al. 2018a, c), and a huge amount of real-time and multi-source data has been produced during the manufacturing process. As reported by Nedelcu (2013), manufacturing sector kept more data than other sectors, estimated close to two Exabytes of new data stored in 2010. These data can be used by manufacturers to support intelligent and timely production decisions (Xue et al. 2016), to assist manufacturers to predict what they will do tomorrow (Zhang et al. 2017a).

As a core part for improving the production efficiency, material delivery in shopfloor has attracted widely attention in academia and industry. Khayat et al. (2006) proposed an integrated formulation to solve the combined production and material delivery scheduling problems. Based on the analysis of the vehicle routing problem with manual materials handling (VRPMMH), two models of VRPMMH were developed by Boonprasurt and Nanthavanij (2012) to determine the optimal fleet size and delivery routes to minimize the total cost. The self-organizing assembly system was presented by Frei et al. (2014) to spontaneously organize production machines in the shopfloor in response to the arrival of products order and materials. A future manufacturing system in the big data environment was described by Zhang et al. (2017b). The authors pointed out that the production process and material delivery will be more efficient based on the huge amount of real-time manufacturing data. To remain sustainable competitive advantage (SCA) (Liu and Liang 2015) of an automobile parts manufacturer, an automated tracking system for management of material delivery was designed by Jamaludin et al. (2018).

Despite some progress have achieved in the field of shopfloor material delivery, major challenges still exist in achieving the vision of real-time and multi-source data-driven decision-making of shopfloor material delivery in a manufacturing big data environment. They are summarized as follows:

- How to design a solution of data acquisition based on typical big data infrastructure, and then apply intelligent sensing devices and information technologies to make heterogeneous manufacturing resources have the ability of active sensing and interacting, so that the real-time material delivery planning can be achieved. By deploying the intelligent sensing devices to heterogeneous manufac-

turing resources (e.g. operator, machine, material pallet, trolley), the real-time status data of the manufacturing resources in different production stages could be tracked and captured in a timely fashion. Based on these multi-source data, the real-time data-driven frequent trajectory of material delivery can be mined to promote the delivery efficiency.

- How to develop a solution to: (1) preprocess the captured manufacturing big data; (2) store them in different data management systems, so as to ensure the availability and reutilization of the manufacturing data in future data analysis. Through the implementation of data preprocessing technologies, the multi-source and heterogeneous manufacturing data could be integrated and then shared among different production stages. Moreover, in order to promote the quick retrieval and deep analysis of data, a large amount of heterogeneous manufacturing data needs to be stored according to different data types. Based on these operations, reliable and available data could be acquired, from which reasonable and effective decisions of material delivery can be made.
- How to establish a data mining model according to the multi-sources and integrated manufacturing big data to mine the hidden pattern or association relationships among different production stages in a timely and dynamical mode, so as to facilitate shopfloor managers to make accurate and more-informed material delivery decisions. In the traditional material delivery approach, the modeling method based on multi-source manufacturing data is rarely taken into account. In fact, in order to improve material delivery decision-making, it is necessary to establish a data analysis model based on the material delivery relevant data that from different production stages and various manufacturing resources.

To address the above challenges, in this paper, a framework for shopfloor material delivery based on real-time manufacturing big data (SMD-RMBD) is proposed. The rest of the paper is organized as follows. Section 2 reviews the literature related to this research topic. Section 3 presents the framework of SMD-RMBD. Then, by using the proposed framework, several key enabling technologies are developed and discussed. An application scenario and a series of experiment are designed and conducted in Sect. 4. Conclusions and future works are given in Sect. 5.

2 Literature review

Two aspects of literature are reviewed in the following, one is real-time manufacturing data acquisition, and the other is shopfloor material delivery.

2.1 Real-time manufacturing data acquisition

In recent years, some scholars have explored the practice of IoT or IIoT technologies in shopfloor management, especially for the radio frequency identification (RFID)-based applications. An RFID-enabled system was designed to monitor the consumption of resources in a warehouse (Poon et al. 2009), where data collection and sharing were facilitated by RFID. To assist the managers' intentions to improve operational conditions under the adoption of RFID, a framework was developed to support collaboration at different levels within companies (Sari 2010). Based on market-based decision-making framework, an approach for RFID-enabled finished vehicle deployment planning was proposed by Kim et al. (2010) to coordinate real-time changes of vehicle locations. Based on RFID technology, a digital warehouse system in the tobacco industry was proposed by Wang et al. (2010). A case in the tobacco industry was studied to illustrate the feasibility of the proposed system. The authors pointed out that the system can help warehouse managers to achieve better material management and inventory control. A multi-agent based real-time production scheduling method for RFID-enabled ubiquitous shopfloor environment was proposed by Zhang et al. (2014) to acquire shopfloor data and to facilitate the realization of real-time scheduling. An RFID-based decision support system (DSS) was proposed to improve the production visibility and the decision-making performance in a distributed manufacturing environment (Guo et al. 2015). To achieve energy-saving and to prolong the lifetime of the whole manufacturing system, Wang et al. (2016) proposed a green IIoT architecture. Based on IIoT data, the authors designed a sleep scheduling and wake-up protocol to predict the sleep intervals of a system. For automobile engines remanufacturers, it is difficult to implement real-time production scheduling due to the lack of timely and accurate data of remanufacturing resources. To address this problem, an RFID-enabled remanufacturing environment was designed and tested by Zhang et al. (2018b) to monitor the real-time status of the disassembled engine parts and the remanufacturing resources. Wang et al. (2018a) explored an IoT-enabled energy efficiency optimization method for energy-intensive manufacturing enterprises. The IoT technologies were applied to sense the real-time primitive production data (e.g. energy consumption data and resources status data).

2.2 Shopfloor material delivery

Many strategies and models have been proposed by researchers in the shopfloor material delivery domain to improve the material distribution. A novel vehicle route planning model for transporting hazardous materials with multiple time-varying attributes was proposed by Meng et al. (2005). Im et al.

(2009) developed a vehicle-dispatching method to minimize the vehicle blocking and delivery times in automatic material handling systems of 300 mm semiconductor manufacturing. The discrete event simulation models were developed to evaluate the performance of the proposed method. Qu et al. (2012) illustrated a case of implementing RFID-based shopfloor material management for household electrical appliance manufacturers. Some research questions such as how technical, social and organizational issues of applying RFID for material distribution were addressed. Zhang et al. (2015) presented a dynamical model in optimizing transport tasks for shopfloor material handling. An approach to excavate logistics trajectory from RFID-enabled shopfloor data was proposed by Zhong et al. (2015). A series of experiments were designed and tested to evaluate the practicality of the proposed approach. Authors found that the proposed approach could be implemented in the supply chain management field, which is using RFID for facilitating the operations. Mohammed et al. (2017) investigated the optimization method of automated warehouse system in terms of the optimal number of storage racks and collection points. The applicability of the developed method was examined using a case study. The result showed that the proposed solution can be used to minimize the travel time of products from storage racks to collection points. An real-time information-driven optimization model of the assembly process in a synchronous line and a cyber-physical system (CPS) based smart control model for shopfloor material handling were investigated by Zhang et al. (2018d). On this basis, Zhang et al. (2018a) developed a model for production-logistics systems based on CPS and IIoT. The authors recommended that the proposed model can be used to investigate self-organizing configuration mechanisms and to improve the efficiency of shopfloor logistics. A digital twin-driven product design and manufacturing method based on cyber and physical convergence were proposed by Tao et al. (2018b). These approaches provide new capabilities to address the problems in material handling.

As stated in the foregoing sections, many studies related to material delivery mainly focused on traditional manufacturing environment. The combination of advanced information technologies and big data will bring the new opportunities to improve many manufacturing dimensions (Saucedo-Martínez et al. 2017). For instance, the large amount of data generated in the shopfloor will be used to improve the accuracy of real-time material delivery decisions. Therefore, some research gaps should be further addressed in the manufacturing big data environment.

Firstly, from the perspective of real-time manufacturing data acquisition (Sect. 2.1), many researchers only focused on the traditional manufacturing environment. In order to fulfill the real-time big data-driven material delivery, a new IoT solution based on big data infrastructure

is imperative. Within this solution, IoT devices will be deployed to heterogeneous manufacturing resources, and make them are able to sense and interact in a timely fashion. As a result, the real-time and multi-sources manufacturing big data can be captured. Meanwhile, what kind of IoT devices are needed and where these devices are installed should also be considered.

Secondly, manufacturing big data is characterized by real-time and multi-sources as well as heterogeneous (Wang et al. 2018b). Some incomplete and incorrect records are included in the raw manufacturing big data, which may affect the accuracy of decision-making. However, the issues of data processing and data integration in different production stages were seldom considered by existing material delivery relevant research (Sect. 2.2). Therefore, the data quality for modeling and analyzing is not guaranteed. In addition, various data storage solutions should be designed to store the non-real-time and heterogeneous manufacturing data to ensure the integrality and reusability of data for further deep analysis.

Thirdly, from the perspective of the data modeling, the research on integrated applications of the multi-sources manufacturing data for material delivery trajectory modeling is seldom involved (Sect. 2.2). Moreover, many research was focused on the strategy of centrally assigned according to a given task (Zhang et al. 2015). However, during the production processes, the deviation between execution and the pre-planned material delivery trajectory is often produced due to the unpredictable production exceptions. Therefore, based on the multi-sources and real-time manufacturing big data, a new data modeling method should be considered to discover the optimal logistic trajectory in a timely and dynamical mode, and reduce the deviation.

To address these research gaps, a framework for SMD-RMBD is presented and tested to achieve real-time and multi-source data-driven material delivery in the manufacturing big data environment. Generally, a framework can be used to describe the layout of the whole system, to simplify the complex relationships among all components, and to ensure the validity of the entire system (Vikhorev et al. 2013). Furthermore, according to a thesaurus dictionary by Oxford University Press, a framework is defined as “a set of beliefs, ideas or rules that forms the basis of a system” (Oxford University Press 2014). As implied by these, the framework is at the highest level when developing a complex system. In addition, the topic of the paper is in its infancy as described in paragraph 4 of Sect. 1. Therefore, it is regarded to target a framework first, which then may be built upon to develop a model or a method, to be the most effective way for advancing the knowledge for the topic. The proposed framework in Sect. 3 is a set of ideas including data acquisition, data processing and data modeling that form a basis of an analysis system for manufacturing big data and aims to

show how the processes of big data-driven material delivery could be implemented.

3 Framework for shopfloor material delivery based on real-time manufacturing big data

By applying IoT technologies in the processes of shopfloor production and management, a smart manufacturing environment is established, and the real-time and multi-source big data of heterogeneous manufacturing resources can be captured. Big data processing and analysis technologies are used to preprocess and analyze the manufacturing big data, to discover the hidden rules and association relationships from them. With the assist of the discovered association relationships, better decision-making for material delivery are provided to managers. Based on the above-mentioned procedure, a framework for SMD-RMBD is designed as shown in Fig. 1.

3.1 Manufacturing big data sensing and acquisition

The configuration of the smart manufacturing objects is significant to enhance the sensing capability of the heterogeneous manufacturing resources. The manufacturing resources are made ‘smart’ by equipping the physical manufacturing resources with intelligent sensing or auto-identification (Auto-ID) devices to achieve a certain degree of intelligence (Zhang et al. 2017c). In this framework, sensing devices such as RFID (e.g. RFID tags and readers), smart sensor, Personal Digital Assistant (PDA) and smart card are used to capture the real-time big data of manufacturing resources in the shopfloor. Networking technology such as Backbone Concentrator Node (BcN), WLAN and IPv6 are used to transmit the captured data from sensing devices to the enterprise database. The device middleware is used to collect and filter raw data from sensing devices.

The RFID is configured in different ways. In any case, trays of materials are deployed with ultra-high frequency (UHF) RFID tags and become smart objects. Therefore, the location information of the trays can be traced timely. The tags also contain information about what material is being involved and how many. Critical tools and key parts of the products are also deployed with tags due to their important roles in the following assembly process. The information of these tags will be updated during different stages of assembly operations, and are eventually carried over by finished products throughout the manufacturing processes. In addition, inventory areas, workstations and buffers for different materials, work in process (WIP) and finished products are also tagged. Each machine tool is equipped with RFID readers. Generally, UHF RFID readers are recommended

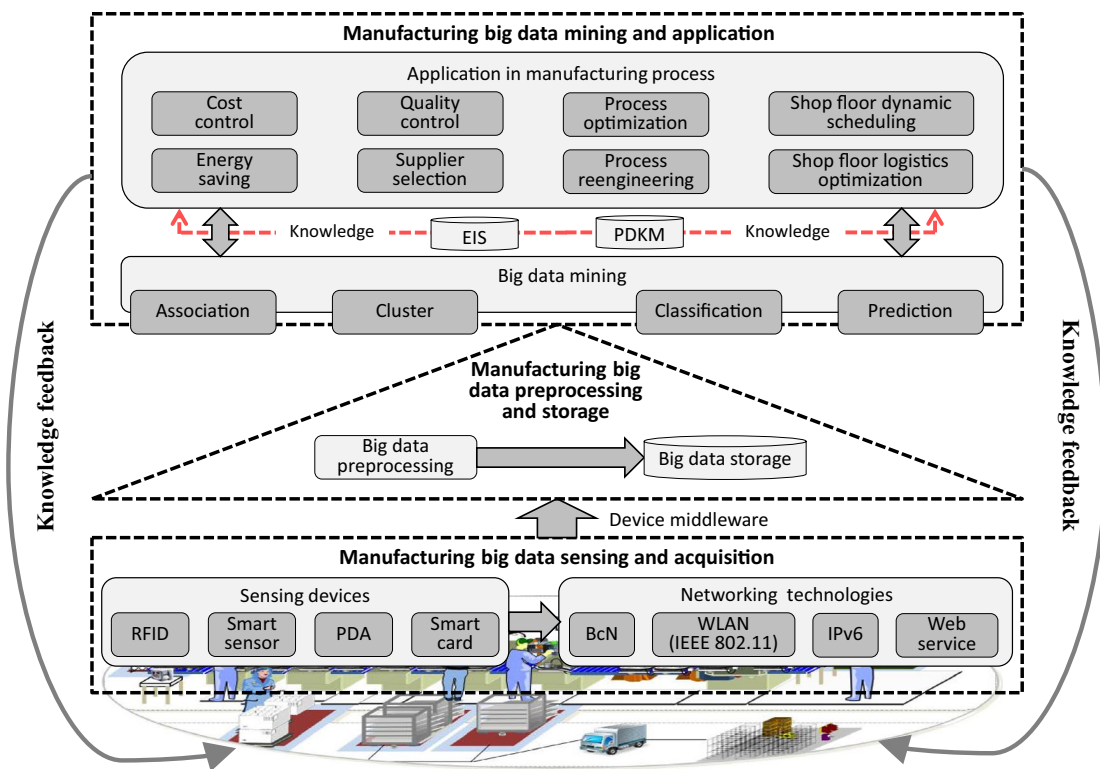


Fig. 1 Framework for shopfloor material delivery based on real-time manufacturing big data

for real-life implementation due to its affordable cost and practically acceptable reading capability. The RFID readers are also deployed to vehicles that are directly used for moving material trays, toolboxes, WIP and finished products to read the real-time status information of them. Due to the limited space, the specific methods for manufacturing big data acquisition are not repeated here. Interested readers are encouraged to read the recent publications Zhang et al. (2017b, c) of the authors.

3.2 Manufacturing big data preprocessing and storage

Based on the configuration of intelligent sensing or Auto-ID devices to the manufacturing resources, real-time status data of the manufacturing resources can be captured during the production process. However, there is some ‘noise’ (e.g. redundant, incomplete and incorrect data records) included in the raw manufacturing big data, which may not be analyzed directly and may affect the accuracy of shopfloor decisions. Therefore, the data preprocessing operations such as cleaning, integration, reduction, and transformation should be implemented so as to provide reliable and available data for further modeling and pattern discovery (Chen and Honda 2018). In addition, a huge number of manufacturing data

need to be stored to provide complete and reusable data for further deep data analysis.

3.2.1 Multi-source and heterogeneous big data preprocessing

The quality of data is foremost before running an analysis. Therefore, data preprocessing must be done to remove the ‘noise’ before the data is stored and analyzed. The data processing and storage solution is designed as shown in Fig. 2.

Firstly, the raw manufacturing big data have a large amount of redundancy, thus, a data cleansing operation should be performed to reduce the redundancy. Due to the limited space, the specific procedures for data cleaning operation are not repeated here. Interested readers are encouraged to read Zhang et al. (2017b). Secondly, the cleansed manufacturing big data is still scattered and unusable during shopfloor decisions. It is essential to carry out a data integration operation. Due to it is difficult to express the multi-source and heterogeneous data in a model, a unified data modeling method is proposed to integrate the manufacturing big data, and then to construct a logically unified framework to express the manufacturing big data. Therefore, the concept of meta-model (Molano et al. 2017) is introduced to build the unified data model. Four types of meta-model are illustrated in Table 1.

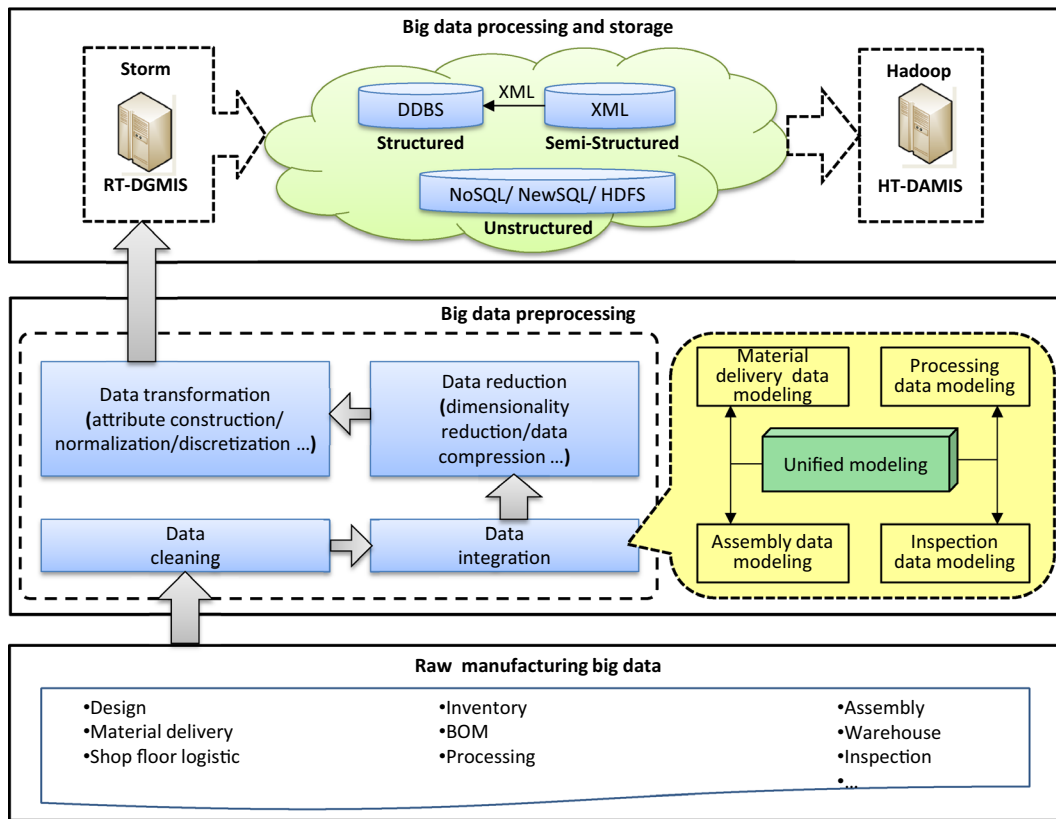


Fig. 2 Big data preprocessing and storage solution

Table 1 Four types of data meta-model and functions

Meta-model	Functions
Top-level meta-model	Describing the integrated data model
Material delivery data meta-model	Describing the multidisciplinary data and the file information during the material delivery process
Processing data meta-model	Describing the data model of technology and process, and the association relationship between them
Assemble data meta-model	Describing the data and the file information during the assembly process
Inspection data meta-model	Describing the quality data model and the association relationship among them

Take the material delivery data meta-model as an example, the modeling procedures are illustrated as follows: (1) define the data abstraction language, such as association relationship, delivery rule and standard; (2) describe the material delivery data (e.g. delivery scheme, attribute data) and their association relationship by the predefined data abstraction language; (3) establish domain ontology repository to describe the relationships between concept and attributes, and describe the constraints between attributes and relationships; (4) establish top-level meta-model to define multiple meta-model of various manufacturing data, and the overall association relationships and unified data format; (5) achieve the model sharing and data integrating through instantiation of meta-models. The other three

meta-models are not included as the modeling processes are basically similar.

Thirdly, the integrated data sets are usually still huge. Therefore, a data reduction operation should be performed to obtain a reduced representation of the data sets that are much smaller in volume, yet closely maintains the integrity of the original data.

Finally, the reduction manufacturing data must be transformed so that the patterns found may be easier to understand.

The preprocessed manufacturing big data are transmitted to the enterprise database, meanwhile, shared and applied by various manufacturing stages to optimize the production processes.

3.2.2 Multi-source and heterogeneous big data storage

According to the data classification standard in the literature (Zhang et al. 2017b), three kinds of big data management and storage solutions are elaborated as follows.

With more and more structured data is produced and collected in the shopfloor, two typical issues are appearing: (1) historical and real-time data are stored in the same database, which is impacting the performance of data processing; (2) the information value-added of historical data cannot be realized. Aiming to address these two issues, data generation management information system based on real-time data (RT-DGMIS) and data application management information system based on historical time data (HT-DAMIS) are designed. RT-DGMIS is used for managing the real-time data, while HT-DAMIS is responsible for collecting the data that are produced by RT-DGMIS. Storm and Hadoop computing framework are used to process the real-time and non-real-time data, respectively. Distributed Data Base System (DDBS) is used to store the structured data.

Semi-structured data is a type of data which is between structured and unstructured. For example, technical documents described using markup languages like XML, and real-time status data of WIP returned back to the backend system by closed-circuit television. There are not only structured data such as location and time but also unstructured data such as the picture. XML is a primary standard for expressing the structured or semi-structured data. Therefore, it can be used to describe semi-structured manufacturing big data. As a result, the semi-structured data are transformed into a standardized data format and stored in DDBS.

Unstructured data is those without format, spatial and temporal constraints. A great amount of unstructured data is producing by heterogeneous manufacturing resources. Because DDBS cannot meet the application requirements of big data in flexibility and scalability (especially the scale out), the distributed approach such as Hadoop Distributed File System (HDFS) (Haydaya and Marchildon 2012) and not only Structured Query Language (NoSQL) (Lucchese 2018) data management system are used to manage and store the unstructured manufacturing big data.

3.3 Manufacturing big data mining and application

By means of the big data analytics and mining technologies, the data mining model is established to mine the hidden pattern and knowledge from the real-time and historical lifecycle manufacturing big data. The managers of an enterprise can adjust and optimize the whole production processes according to knowledge feedback. In addition, by integrating the mined knowledge, application services such as cost and quality control, shopfloor dynamic scheduling, route optimization of material delivery, etc. can be achieved.

3.3.1 A graphic model for mining of manufacturing big data

To achieve the above-mentioned application services, a graphic model for mining of manufacturing big data is designed as shown in Fig. 3. General models and specific models are presented in this paper according to the various application requirements of shopfloor.

Firstly, the authors of this paper developed four types of general models. The functions of these general models are introduced as seen in Table 2.

Secondly, in this research, four types of specific models are also developed (as seen in Fig. 3). The special model is built for a specific application. For example, if an application service of route optimization for material delivery is proposed, the specific model of material delivery and related dataset are selected, by adjusting the control parameters of the model and optimizing the data sets, the optimal path of material delivery can be achieved. The other three specific models are not discussed here as the establishing principles are similar.

From left to right of the Fig. 3, the graphic model is divided into data source module, data mining model module, data mining result set module and application requirement module. Data source module provides completed manufacturing data for data mining. Data mining module involves data modeling operations, where general models and specific models are included. These models are responsible for extracting original data and discovering hidden knowledge from the data source module. The data mining result set module is a set of data mining results. According to different application requirements of shopfloor, suitable data and models are selected and established to perform mining operations. Application requirement module is a series of demands of shopfloor management, which applies mined knowledge of the result set module to achieve the requirements of shopfloor.

3.3.2 An improved Apriori-based association analysis model for mining material delivery trajectory

This subsection develops a model for mining material delivery trajectory based on the above-mentioned graphic model. Discovering association relationships from the dataset are vital in big data relevant applications (Hofmann 2017). Association analysis is an effective method to discover the association rules among items that occur synchronously in a given dataset. Therefore, a large amount of manufacturing data in shopfloor can be used to mine association rules and frequent patterns related to processing quality of the products, requirements of the customer, frequent trajectories of material delivery, and so on. As a typical frequent pattern mining method, the Apriori algorithm can be used to fast

Fig. 3 Data mining graphic model of manufacturing big data

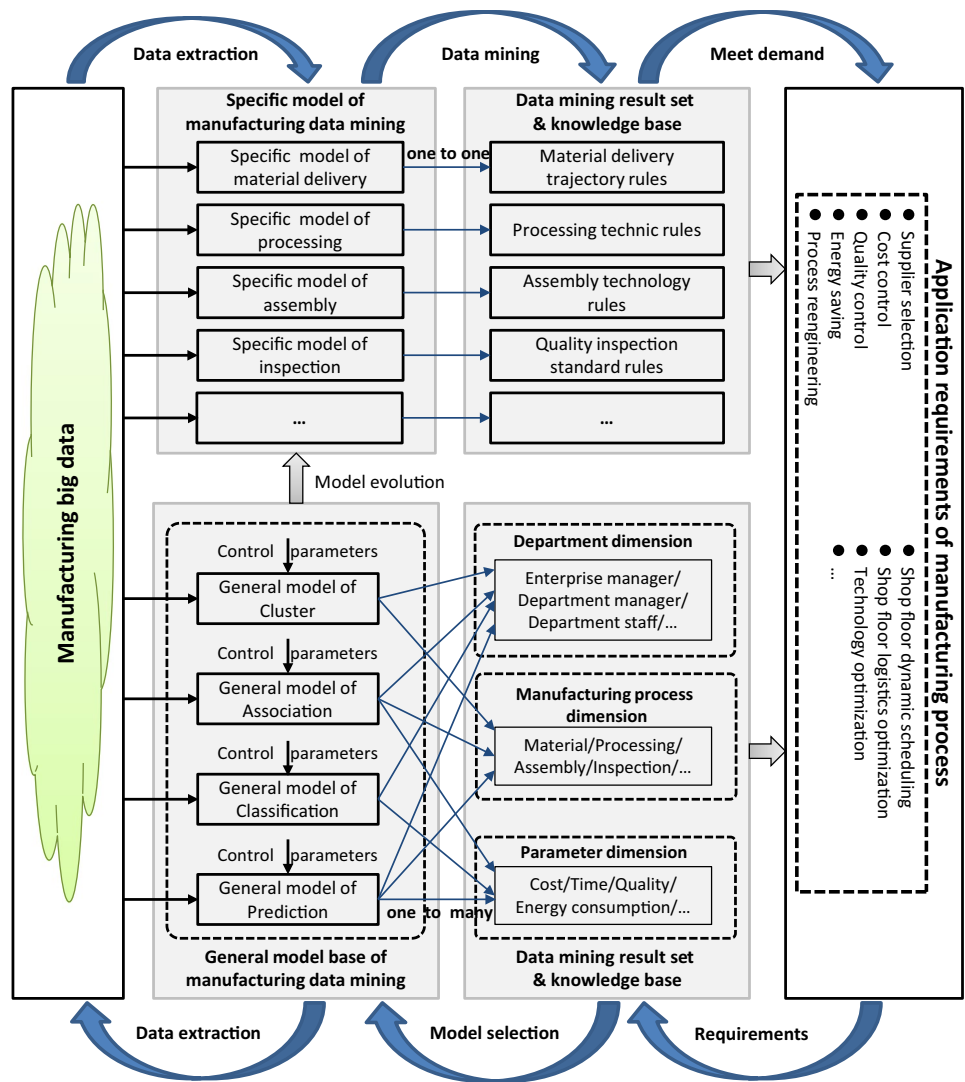


Table 2 Types and functions of the general model

Types of model	Example functions
Clustering	Formulating the production plan based on customer value analysis
Association	Analyzing the quality factor of the product and mining the logistics trajectory
Classification	Customers churn analysis and fault diagnosis
Prediction	Forecasting the order quantity and machine lifetime

extract implicit association rules from large datasets and to support decision-making (Agrawal and Srikant 1994). However, the traditional Apriori algorithms have some inherent defects when dealing with datasets that are continuously generated and dynamically updated: scanning the database frequently; generating a large number of candidate sets and hence presenting a low computational efficiency (Oswald and Sivaselvan 2018). Therefore, the traditional Apriori

algorithms may not suitable for real-time manufacturing big data environment.

With that in mind, an improved Apriori-based model for mining material delivery trajectory is proposed to enhance the efficiency of data mining and to achieve the unity of real-time and accurate planning of material delivery trajectory. It means that the improved Apriori-based model not only can accurately discover the frequent trajectories of materials delivery but also can enhance the efficiency of handling continuously updated data. The following paragraphs introduce some basic definitions in the improved Apriori-based model.

Definition 1 Given a node of the material delivery process in the form $(j_{bt}^i, p_{bt}^i, m_{bt}^i)$, j_{bt}^i represents the label of material type, p_{bt}^i represents the label of current production stage for processing the material, and m_{bt}^i represents the label of the machine that processes the material. Subscript b represents the material's batch and t represents the timestamp to start processing the material.

Definition 2 Considering two nodes $(j_{bt_1}^i, p_{bt_1}^i, m_{bt_1}^i)$ and $(j_{bt_2}^i, p_{bt_2}^i, m_{bt_2}^i)$ in the same trajectory, if $t_1 < t_2$, it indicates that the process represented by t_1 is performed before the process represented by t_2 .

Definition 3 By integrating all nodes of material delivery processes, a complete trajectory of material delivery from the first production stage to the last production stage is formed, which is denoted as a vector $T_i = [j_{bt_1}^i, m_{bt_1}^i, m_{bt_2}^i, \dots, m_{bt_n}^i]$. The location $m_{bt_k}^i$ in the trajectory vector represents the sequence number of the production stage in which it is located.

Definition 4 The *support* of the trajectory pattern P is defined by

$$support(P) = \frac{\{T_i | P \subseteq T_i, T_i \subseteq T\}}{\{T\}} \tag{1}$$

where T represents the material trajectory database and T_i represents the trajectories in T containing trajectory pattern P . If the *support* of a trajectory pattern is not less than *min_sup* (also be called minimum support threshold, which is a user-defined factor), the trajectory pattern is regarded as a frequent trajectory pattern. If the *support* of a trajectory pattern is described as $min_sup > support \geq min_cruc$, the trajectory pattern is a crucial trajectory pattern. Crucial trajectory pattern may become frequent trajectory pattern in the future as the amount of data in the database is continuous increasing (the reasons are illustrated in the following paragraph), and *min_cruc* is the minimum threshold of crucial trajectory pattern.

Definition 5 The *confidence* of the trajectory pattern $P \rightarrow N$ is defined by

$$confidence(P \rightarrow N) = \frac{\{T_k | (P, N) \subseteq T_k, T_k \subseteq T\}}{\{T_i | P \subseteq T_i, T_i \subseteq T\}} \tag{2}$$

where T_i represents the trajectories containing P , and T_k represents the trajectories containing P and N . If the *confidence* of trajectory pattern is not less than *min_conf* (also be called minimum *confidence* threshold), the trajectory pattern is considered as a frequent trajectory pattern.

For an actual production environment, the manufacturing data is continuously produced and updated. As the new data is added to the database, support and confidence of the trajectory pattern extracted from the initial database will be changed. Here, the crucial pattern is introduced to update the trajectory pattern in a database. For some logistics

trajectories, they may be not frequent trajectories in the initial database. However, these trajectories are likely to become frequent trajectories in the future as the manufacturing data is continuously increasing. The potential frequent trajectory pattern is regarded as crucial pattern, which may be significantly important for the future material delivery decision-making. Therefore, in the improved Apriori-based model, the concept of incremental learning is applied to give different weights to the support and confidence of the trajectory pattern in the historical and the real-time databases, and then to update the frequent trajectories in a timely and dynamical mode.

Based on the above-mentioned basic definitions, the key steps of the algorithm for the improved Apriori-based model are described in Table 3.

The improved Apriori-based method is enabled by some key steps equipped with suitable algorithms. They are frequent and crucial trajectories mining based on historical data (procedure 1), and frequent and crucial trajectories extracting and updating based on real-time data (procedure 2).

The purpose of procedure 1 is to mine all frequent trajectories L and crucial trajectories CS from the historical data. The input is a set of historical datasets D . The output is a set of mined frequent trajectories R and crucial trajectories CR which carry accurate information of material delivery trajectories. The purpose of procedure 2 is to: firstly mine frequent trajectories R^+ and crucial trajectories CR^+ from the real-time data; secondly extract the frequent trajectories R and crucial trajectories CR that obtained in procedure 1; thirdly combine the newly generated trajectories R^+ and CR^+ with the previous trajectories R and CR , give different weights to the support and confidence of these trajectories, and then to update the frequent and crucial trajectories of material delivery. The input is a set of real-time datasets D^+ and the historical trajectories R and CR . The output is a set of updated frequent trajectories R' and crucial trajectories CR' which with different support and confidence weights.

4 A study of an application scenario

This section describes a proof-of-concept application scenario to demonstrate how to implement the real-time manufacturing big data tracking and to discover the optimal material delivery trajectory under the presented SMD-RMBD framework. A simulation experiment based on the hypothetical motivating scenario (see Fig. 4) is carried out to validate the proposed framework.

4.1 Deployment of intelligent sensing devices

For simplicity of understanding but without losing generality of the principle, some basic manufacturing resources are

Table 3 The key steps of the algorithm for the improved Apriori-based model

<pre> // Procedure 1: Mining the frequent and crucial trajectories of material delivery based on history data // Find all frequent and crucial patterns (1) $L_1 = \text{find_frequent_1_itemsets}(D)$ (2) For($k = 2, L_{k-1} \neq \emptyset; k++$) { // Find frequent and crucial k-patterns $C_k = \text{apriori_gen}(L_{k-1});$ For each record $t \in D$ { $C_t = \text{subset}(C_k, t);$ For each itemset $c_k \in C_t$ $c_k.\text{count}++;$ } // L_k refers to candidate set of frequent k-pattern, and CS_k refers to candidate set of crucial k-pattern $L_k = \{c_k \in C_k \mid \text{support}(c_k) \geq \text{min_sup}\};$ $CS_k = \{c_k \in C_k \mid \text{min_sup} > \text{support}(c_k) \geq \text{min_cruc}\};$ } (3) Return $L = \{L_1, L_2, \dots, L_k\}$ and $CS = \{CS_1, CS_2, \dots, CS_k\}$ // Find trajectories of L and CS that satisfy min_conf (4) If $\text{confidence}(P \rightarrow N) \geq \text{min_conf}$ If $(P \rightarrow N) \in L$ $(P \rightarrow N)$ is a frequent trajectory, and stored in R Elseif $(P \rightarrow N) \in CS$ $(P \rightarrow N)$ is a crucial trajectory, and stored in CR (5) Delete the repetitive trajectories of R and CR, and R is the obtained frequent trajectory set // Procedure 2: Extracting and updating the frequent and crucial trajectories based on real-time data (1) Acquire R^+ and CR^+ of real-time data (D^+) by repeating procedure 1 (2) Extract R and CR of historical data (D) // Detect whether trajectories of R and CR are frequent or crucial trajectories (3) If $(P \rightarrow N) \in (R \cup CR)$ and $(P \rightarrow N) \in (R^+ \cup CR^+)$ // w_1 and w_2 are weights, and $w_1 + w_2 = 1$ $\text{support}'(P \cup N) = w_1 * \text{support}^D(P \cup N) + w_2 * \text{support}^{D^+}(P \cup N)$ $\text{confidence}'(P \rightarrow N) = w_1 * \text{confidence}^D(P \rightarrow N) + w_2 * \text{confidence}^{D^+}(P \rightarrow N)$ If $\text{support}'(P \cup N) \geq \text{min_sup}$ and $\text{confidence}'(P \rightarrow N) \geq \text{min_conf}$ $(P \rightarrow N)$ is a frequent trajectory, and stored in R' Elseif $\text{min_sup} > \text{support}'(P \cup N) \geq \text{min_cruc}$ and $\text{confidence}'(P \rightarrow N) \geq \text{min_conf}$ $(P \rightarrow N)$ is a crucial trajectory, and stored in CR' Elseif $(P \rightarrow N) \in (R \cup CR)$ and $(P \rightarrow N) \notin (R^+ \cup CR^+)$ // w_3 is weight, and $0 < w_3 < 1$ $\text{support}'(P \cup N) = w_3 * \text{support}^D(P \cup N)$ $\text{confidence}'(P \rightarrow N) = w_3 * \text{confidence}^D(P \rightarrow N)$ If $\text{support}'(P \cup N) \geq \text{min_sup}$ and $\text{confidence}'(P \rightarrow N) \geq \text{min_conf}$ $(P \rightarrow N)$ is a frequent trajectory, and stored in R' Elseif $\text{min_sup} > \text{support}'(P \cup N) \geq \text{min_cruc}$ and $\text{confidence}'(P \rightarrow N) \geq \text{min_conf}$ $(P \rightarrow N)$ is a crucial trajectory, and stored in CR' // Detect whether trajectories of R^+ and CR^+ are frequent or crucial trajectories (4) The trajectory detecting process of R^+ and CR^+ is the same as the procedure of step (3) (5) Delete the repetitive trajectories of R' and CR', and R' is the obtained frequent trajectory set </pre>

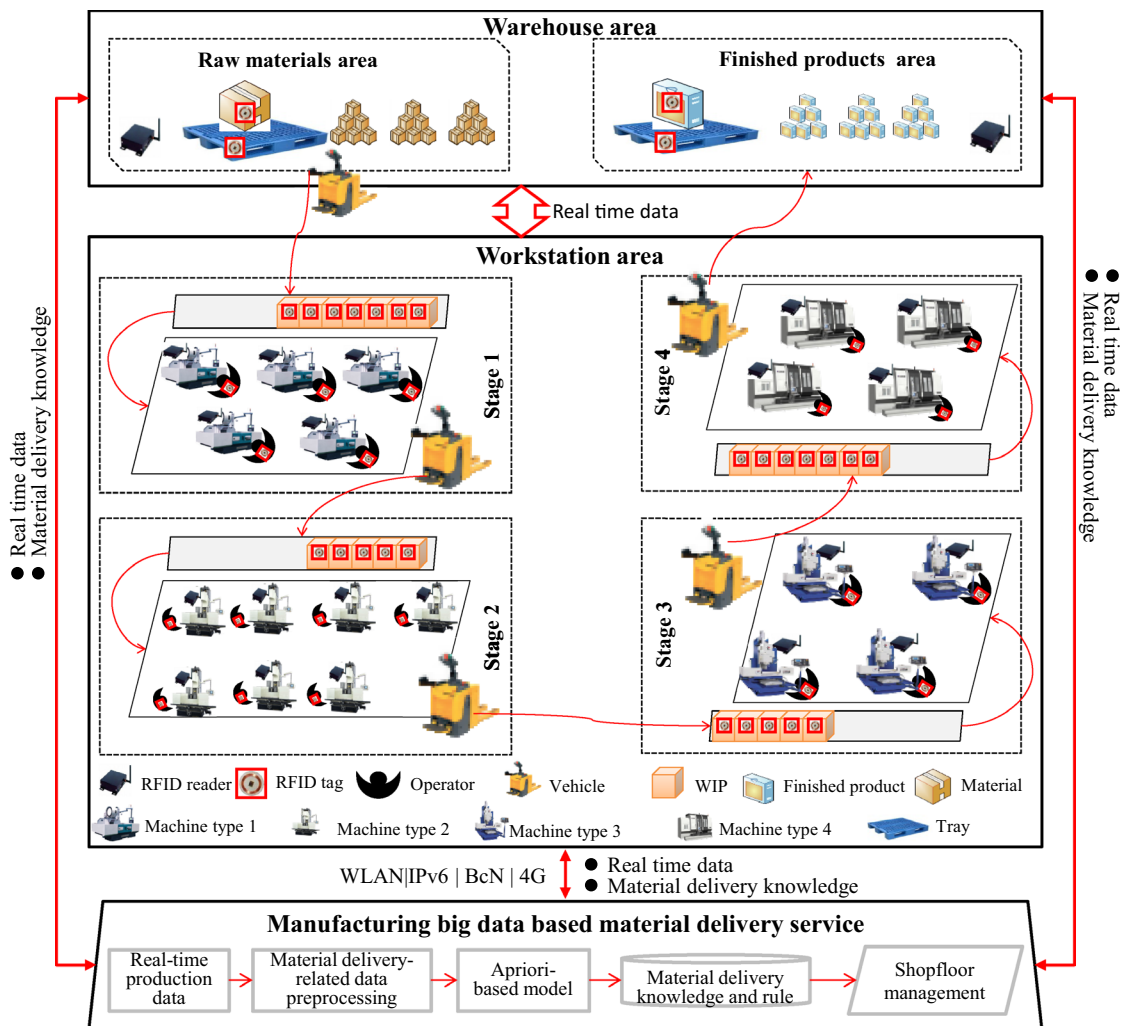


Fig. 4 Overview of the application scenario

selected to configure the intelligent sensing devices and to establish a smart manufacturing environment. In this paper, RFID devices are used to capture the real-time status data of manufacturing resources. The production process consists of one warehouse area and one workstation area (as seen in Fig. 4).

In the warehouse area, the RFID readers are deployed on raw-material loading area, finished product receiving area, and vehicles that carry material trays and finished products. The RFID readers are also deployed on warehouse gate to locate the trays to be delivered and check out the trays. On the workstation area, three types of RFID readers are deployed. For machines, they are equipped with stationary readers. For vehicles that directly used for moving material trays and WIP, they are equipped with different intelligent sensing devices. In addition, buffers for materials and WIP, critical tool, finished products, trays, etc. are deployed with various RFID tags. Operators carry handheld RFID

devices due to their frequent movement within the shopfloor. The deployment information is shown in Table 4. After the deployment of RFID devices, all the resources are converted into smart manufacturing objects, which are able to sense and interact with one another.

4.2 Multi-source RFID-enabled manufacturing data preprocessing

Based on the deployment of smart manufacturing environment, the multi-source and real-time data of manufacturing resources are captured. The original datasets of RFID-enabled material delivery within shopfloor have multi-dimensional attributes, such as Electronic Product Code (EPC), Time, Machine, etc. (see Table 5). However, the original datasets have a great number of redundancies, which will affect the trajectory mining of material delivery. Thus text mining algorithms, using the programming language R,

Table 4 Deployment information of RFID in shopfloor

RFID type	HF/UHF	Manufacturing resources	Objective
Reader	HF	Machine	Track the trays and the operators
Reader	UHF	Warehouse gate	Locate and check out the trays
Reader	UHF	Assembly station	Check in materials delivered from a warehouse, and report finished assembly tasks
Reader	UHF	Vehicle	Track the trays
Tag	HF	Operator	As an RFID staff smart card. Track the real-time data of each operator
Tag	UHF	Tray	Track the real-time data of the tray with materials and WIP
Tag	UHF	Critical tool	Track the real-time data of the critical tools
Tag	UHF	Critical component	Track the real-time data from material/WIP to the product

Table 5 Original RFID datasets

Attributes	Depth of information		
EPC	Job_ID	Operator_ID	Machine_ID
Time	Time_in	Time_out	Time_duration
AGVs	AGV_code	AGV_number	AGV_load
Machine	Falure_code	Failure_time	Repair_operator
Logistic	Process_code	Process_sequence	Location
Product	Batch_ID	Material_ID	Product_number

Table 6 Processed data used for material delivery trajectory mining

Job information	Logistic information	Time information
Job_ID	Process_code	Time_in
Batch_ID	Machine_ID	Time_out

were built to firstly clean the original RFID datasets, and secondly extract metadata for the target data: job, logistic and time information (see Table 6), which are meant to exactly describe the affected information of a material delivery action. This operation can remove the ‘noise’ and reduce the volume of the manufacturing big data, so as to improve the efficiency of data analysis. As a result, different attributes of material delivery datasets are integrated by job type to improve the efficiency of material delivery trajectory mining.

4.3 Mining the frequent trajectory of material delivery based on RFID-enabled manufacturing big data

As seen in Fig. 4, in this section, the hypothetical application scenario is simulated to mine the frequent trajectory of material delivery. The simulation experiments and data analysis are performed on a workstation (Intel(R) Core(TM) i7-7700K CPU at 4.20 GHz) with 32G of RAM. The operating system is Windows 10 Enterprise Edition with 64-bit.

Table 7 Processing time of each job in every stage

Job type	Stage 1	Stage 2	Stage 3	Stage 4
Job 1	65	11	56	31
Job 2	38	60	53	59
Job 3	37	35	58	31
Job 4	52	47	47	16
Job 5	72	21	65	34

The Matlab 2017a and Python 3 are used to simulate the production processes and to perform the association analysis.

The application scenario consists of four production stages, and each stage contains several different machines. In each production stage, a limited buffer with a volume of 1000 jobs is given. At any time, a job can only be assigned to one machine, and a machine can only process one job. The arriving time of orders in the job shop subjects to a Poisson distribution. In each order, five kinds of jobs are included. The processing time of each job in these four production stages is shown in Table 7.

The setting time of machine is included in the processing time. Random mechanical failures are considered in this application scenario. If one machine has a mechanical failure, this machine will not process job until it is repaired. The mean time between failures (MTBF) and mean time to repair (MTTR) of machines are assumed to subject to an exponential distribution.

Based on the above-mentioned assumption and setting, a virtual production process is constructed and simulated. In addition, in order to discover the material delivery trajectory, a total of $N=50$ batches of jobs are taken into account for simplicity without loss of generality and each batch contains 150 jobs. These batches of jobs traverse the whole four production stages. After the data preprocessing operation, the improved Apriori-based association analysis model is applied to mine the frequent trajectory patterns of material delivery for different jobs. The min_sup and min_conf thresholds are set as 15% and 75%, respectively. The mined

frequent trajectory patterns of material delivery for different jobs are shown in Fig. 5.

As seen in Fig. 5, the association relationships between the five jobs and M6/M7 of stage 2 are not displayed. This indicates that there is no strong logistics relationship between M6/M7 and other machines during the process of completing the five types of jobs. In addition to this, four strong association rules of material delivery trajectory for four jobs are discovered, as shown in Table 8.

For material delivery trajectory 1 (M1 → M8 → M15 → M20), the support and confidence is 63.735% and 85.565% respectively, which indicates this trajectory is most important for job 3. The job 5 have two frequent trajectories, they are (M4 → M10 → M14 → M19) and (M5 → M10 → M14 → M19). In these two trajectories, three same machines (M10, M14, and M19) are included. This means that in this shopfloor, M10, M14 and M19 play important roles for job 5. Analogously, the material delivery trajectory 4 (M2 → M10 → M13 → M18) is most important for job 4. For job 1 and job 2, no strong association rules are

found, since the support and confidence of these two jobs are less than 15% and 75%, respectively.

4.4 Comparisons and analyses

In order to evaluate the performance and applicability of the improved Apriori-based method in the manufacturing big data environment, a series of simulated experiments are executed and then compared with the traditional Apriori method that proposed by Agrawal and Srikant (1994). During this process, four groups of datasets that acquired from the above-mentioned application scenario are applied to test the efficiency of the improved method. Based on the four datasets, the execution time of the two methods for mining frequent trajectory patterns are calculated and shown in Table 9, where H-data refers to the number of historical data in a dataset and R-data refers to the number of real-time data in the same dataset. For example, the dataset (H-data: 50,000, R-data: 0) represents that there are 50,000 items of historical data and zero items of real-time data jointly contained in TA.

As seen in Table 9, for dataset TA, the time consumed in the improved method is more than the traditional one (i.e. 8.5264s vs 7.6774s). In order to discover the frequent trajectories and crucial trajectories from an initially historical dataset synchronously, a lower search threshold is given to the improved Apriori-based method (as analyzed in Sect. 3.3.2). As a result, more time is consumed to scan the whole dataset. However, the execution time of the improved method outperforms the traditional one when real-time data

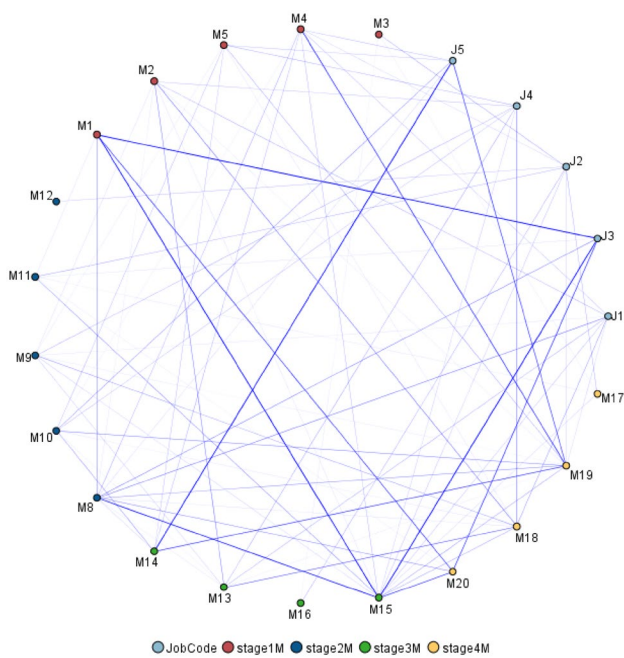


Fig. 5 Frequent trajectory patterns of material delivery material for different jobs

Table 9 A comparison of execution time for the two methods based on four datasets

Datasets	Traditional Apriori method (s)	Improved Apriori-based method (s)
TA (H-data: 50,000, R-data: 0)	7.6774	8.5264
TB (H-data: 50,000, R-data: 20,000)	10.6543	3.4482
TC (H-data: 70,000, R-data: 20,000)	13.6112	3.4913
TD (H-data: 90,000, R-data: 20,000)	16.4618	3.5023

Table 8 The strong association rules of material delivery trajectory for five types of jobs

Trajectory	Job type	Assigned machines				Support (%)	Confidence (%)
		Stage 1	Stage 2	Stage 3	Stage 4		
Trajectory 1	Job 3	M1	M8	M15	M20	63.735	85.565
Trajectory 2	Job 5	M4	M10	M14	M19	30.065	84.035
Trajectory 3	Job 5	M5	M10	M14	M19	29.400	83.673
Trajectory 4	Job 4	M2	M10	M13	M18	19.400	75.601

is involved (as seen in Table 9): TB (3.4482 vs 10.6543), TC (3.4913 vs 13.6112) and TD (3.5023 vs 16.4618). The results indicate that: execution time of the improved method for mining frequent trajectory patterns from datasets that contain real-time data is less than the traditional one, and the difference increases more and more (TB: 7.2061 vs TC: 10.1199 vs TD: 12.9595) as the continuous increase of the historical data.

As the new and real-time manufacturing data is added to the database, the traditional Apriori method scans the historical and real-time data simultaneously in the database many times to discover the frequent trajectory patterns (Guo et al. 2017). As a result, more time is consumed in this process. This indicates that the traditional method is not suitable for near real-time logistics trajectory mining in a manufacturing big data environment. However, in the same case, the improved method firstly analyzes the newly added data to obtain a series of new frequent trajectories, and secondly combines these newly generated trajectories with the previous trajectory patterns that acquired from the historical data to update the frequent trajectory patterns. In other words, the improved method firstly only analyzes and scans the newly added and real-time manufacturing data. On this base, the frequent trajectories will be actively updated according to the trajectory patterns obtained from the historical data. Therefore, the improved method achieves better performance when dealing with the datasets that contain both real-time and historical data.

By applying the improved method, the logistics trajectories relevant information and knowledge can be provided for shopfloor managers in a timely and dynamical mode. As a result, material backlog caused by the unpredictable production exceptions can be avoided to improve the production efficiency. This conclusion confirms the improved method that proposed in this paper can be used to carry out near real-time logistics trajectory planning in the manufacturing big data environment.

4.5 Managerial implications

The simulation experiment based on the hypothetical motivating scenario showed that the proposed framework was feasible to be applied to discover the frequency trajectory of material delivery. This subsection describes implications to assist managers to make more effective shopfloor decisions.

Firstly, the proposed framework can be used for monitoring and collecting the real-time materials inventory and materials consumption data during the whole production processes. The shopfloor managers can use advanced big data analytics to analyze the materials related data, to check the volume of material delivery in different time, to identify the bottleneck of shopfloor logistics, and to optimize the factors that have the greatest effects on material delivery

efficiency. Thus, the proposed framework can be used to improve the flexibility of material delivery in shopfloor.

Secondly, according to the extracted association rules, the most important and frequent logistics trajectories can be identified. The shopfloor managers should pay more attention to the machines on these frequent trajectories. That is to say, in order to ensure that these machines can operate normally and efficiently, the processing machines on these critical logistics trajectories should be regularly inspected and maintained. Meanwhile, more logistic operators should be added to share the logistics load and to avoid and reduce the deviation of material delivery in these trajectories.

Thirdly, the processing machines with a low usage rate can be identified through the mined logistics trajectories. This can assist the managers to make more informed decisions on future production planning and to analyze the potential reasons for the low utilization. For example, machines with a low usage rate maybe reflect a fact that, the machines with a high usage rate have already met the current production requirement. Therefore, to reduce the production costs, these machines with low utilization rate can be removed in the future shopfloor layout.

Fourthly, manufacturing data is continuously generated in a real-world production environment. As new manufacturing data is added in the historical database, initial support and confidence of the association rules will be changed. Through the concept of incremental learning in the improved model, different weights are given to the support and confidence of the rules in the historical and the new database. As a result, the rules with higher weights will be retained temporarily and served as a crucial reference to future production decisions. As implied by these, managers can expand this model to product design and fault diagnosis domains, due to the product design knowledge and failure modes are constantly updated with the increasing of manufacturing big data.

5 Conclusions

Recently, Auto-ID technologies such as RFID have been widely used in shopfloor management and control. Such an automatic data generation manner brings new challenges, for example, how to collect manufacturing big data in a timely and accurate manner, and how to discover the association relationships among the manufacturing big data to improve the efficiency of material delivery. To address these challenges, in this paper, a framework for SMD-RMBD is proposed to provide a new paradigm for shopfloor material delivery.

Three contributions are important in this paper. The first is the solution of data sensing and acquisition in a manufacturing big data environment. By using the solution, multi-source big data of the heterogeneous manufacturing

resources can be collected in a timely fashion, so that the real-time material delivery decisions can be achieved. The second is the manufacturing big data preprocessing and storage solution. It can be used to share and exchange the manufacturing big data among heterogeneous manufacturing resources and different production stages. Meanwhile, reliable and reusable data for further deep analysis can be acquired. The third is a data mining graphical model for manufacturing big data. And on this base, an improved model is developed to identify the frequency trajectories of material delivery in a timely and dynamical mode.

The proposed framework provides a new kind of reference infrastructure to improve the performance of shop-floor material delivery by using multi-source and real-time manufacturing big data. Future research will focus upon the following two aspects. Firstly, the data from various production stages have their own characteristics. Therefore, data preprocessing algorithms for the big data of different production stages should be investigated. Secondly, the discovered association rules can be used for improving the material delivery decisions. By using the artificial intelligent and deep learning technologies, more accurate data analysis models should be developed to identify the hidden pattern from multi-source and real-time manufacturing big data and to make more precise decisions for the optimal material delivery trajectory.

Acknowledgements The authors would like to acknowledge the financial supports of National Science Foundation of China (51675441), the 111 Project Grant (B13044).

References

- Agrawal R, Srikant R (1994) Fast algorithms for mining association rules in large databases. *J Comput Sci Technol* 15(6):487–499
- Boonprasurt P, Nanthavanij S (2012) Optimal fleet size, delivery routes, and workforce assignments for the vehicle routing problem with manual materials handling. *Int J Ind Eng Theory Appl Pract* 19(6):252–263
- Chen T, Honda K (2018) Solving data preprocessing problems in existing location-aware systems. *J Ambient Intell Humaniz Comput* 9(2):253–259
- Frei R, Șerbănuță TF, Di Marzo Serugendo G (2014) Self-organising assembly systems formally specified in Maude. *J Ambient Intell Humaniz Comput* 5(4):491–510
- Guo ZX, Ngai EWT, Yang C, Liang X (2015) An RFID-based intelligent decision support system architecture for production monitoring and scheduling in a distributed manufacturing environment. *Int J Prod Econ* 159:16–28
- Guo Y, Wang M, Li X (2017) Application of an improved Apriori algorithm in a mobile e-commerce recommendation system. *Ind Manag Data Syst* 117(2):287–303
- Haydaya P, Marchildon P (2012) Understanding product lifecycle management and supporting systems. *Ind Manag Data Syst* 112(4):559–583
- Hofmann E (2017) Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect. *Int J Prod Res* 55(17):5108–5126
- Huang GQ, Zhang YF, Jiang PY (2007) RFID-based wireless manufacturing for walking-worker assembly islands with fixed-position layouts. *Robot Comput Integr Manuf* 23(4):469–477
- Im K, Kim K, Park T, Lee S (2009) Effective vehicle dispatching method minimising the blocking and delivery times in automatic material handling systems of 300 mm semiconductor fabrication. *Int J Prod Res* 47(14):3997–4011
- Jamaludin Z, Huang CY, Abdullah L et al (2018) Automated tracking system using RFID for sustainable management of material handling in an automobile parts manufacturer. *J Telecommun Electron Comput Eng* 10(1–7):35–40
- Jararweh Y, Al-Ayyoub M, Darabseh A et al (2015) SDIoT: a software defined based internet of things framework. *J Ambient Intell Humaniz Comput* 6(4):453–461
- Khayat G, El Langevin A, Riopel D (2006) Integrated production and material handling scheduling using mathematical programming and constraint programming. *Eur J Oper Res* 175(3):1818–1832
- Kim J, Ok C-S, Kumara S, Yee S-T (2010) A market-based approach for dynamic vehicle deployment planning using radio frequency identification (RFID) information. *Int J Prod Econ* 128(1):235–247
- Liu Y, Liang L (2015) Evaluating and developing resource-based operations strategy for competitive advantage: an exploratory study of Finnish high-tech manufacturing industries. *Int J Prod Res* 53(4):1019–1037
- Lucchese F (2018) From P2P to NoSQL: a continuous metric for classifying large-scale storage systems. *J Parallel Distrib Comput* 113:227–249
- Meng Q, Lee D-H, Cheu RL (2005) Multiobjective vehicle routing and scheduling problem with time window constraints in hazardous material transportation. *J Transp Eng* 131(9):699–707
- Mohammed A, Wang Q, Alyahya S, Bennett N (2017) Design and optimization of an RFID-enabled automated warehousing system under uncertainties: a multi-criterion fuzzy programming approach. *Int J Adv Manuf Technol* 91(5–8):1661–1670
- Molano JIR, Lovelle JMC, Montenegro CE et al (2017) Metamodel for integration of internet of things, social networks, the cloud and industry 4.0. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-017-0469-5>
- Nedelcu B (2013) About big data and its challenges and benefits in manufacturing. *Database Syst J* IV(3):10–19
- Ning T, Huang M, Liang X, Jin H (2016) A novel dynamic scheduling strategy for solving flexible job-shop problems. *J Ambient Intell Humaniz Comput* 7(5):721–729
- Oswald C, Sivaselvan B (2018) An optimal text compression algorithm based on frequent pattern mining. *J Ambient Intell Humaniz Comput* 9(3):803–822
- Oxford University Press (2014) Oxford dictionary online. Oxford University Press, Oxford. <http://www.oxforddictionaries.com/>
- Poon TC, Choy KL, Chow HKH et al (2009) A RFID case-based logistics resource management system for managing order-picking operations in warehouses. *Expert Syst Appl* 36(4):8277–8301
- Posada J, Toro C, Barandiaran I et al (2015) Visual computing as a key enabling technology for Industrie 4.0 and industrial internet. *IEEE Comput Graph Appl* 35(2):26–40
- Qu T, Yang HD, Huang GQ et al (2012) A case of implementing RFID-based real-time shop-floor material management for household electrical appliance manufacturers. *J Intell Manuf* 23(6):2343–2356
- Sari K (2010) Exploring the impacts of radio frequency identification (RFID) technology on supply chain performance. *Eur J Oper Res* 207(1):174–183
- Saucedo-Martínez JA, Pérez-Lara M, Marmolejo-Saucedo JA et al (2017) Industry 4.0 framework for management and

- operations: a review. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-017-0533-1>
- Tao F, Zuo Y, Xu L, Da Zhang L (2014) IoT-based intelligent perception and access of manufacturing resource toward cloud manufacturing. *IEEE Trans Ind Inform* 10(2):1547–1557
- Tao F, Cheng J, Qi Q (2018a) IIHub: an industrial internet-of-things hub toward smart manufacturing based on cyber-physical system. *IEEE Trans Ind Inform* 14(5):2271–2280
- Tao F, Cheng J, Qi Q et al (2018b) Digital twin-driven product design, manufacturing and service with big data. *Int J Adv Manuf Technol* 94(9–12):3563–3576
- Vikhorev K, Greenough R, Brown N (2013) An advanced energy management framework to promote energy awareness. *J Clean Prod* 43:103–112
- Wang H, Chen S, Xie Y (2010) An RFID-based digital warehouse management system in the tobacco industry: a case study. *Int J Prod Res* 48(9):2513–2548
- Wang K, Wang Y, Sun Y et al (2016) Green industrial internet of things architecture: an energy-efficient perspective. *IEEE Commun Mag* 54(11):48–54
- Wang W, Yang H, Zhang Y, Xu J (2018a) IoT-enabled real-time energy efficiency optimisation method for energy-intensive manufacturing enterprises. *Int J Comput Integr Manuf* 31(4–5):362–379
- Wang JW, Yang JY, Zhang J et al (2018b) Big data driven cycle time parallel prediction for production planning in wafer manufacturing. *Enterp Inf Syst* 1–19. <https://doi.org/10.1080/17517575.2018.1450998>
- Xue S, Xiong L, Yang S, Zhao L (2016) A self-adaptive multi-view framework for multi-source information service in cloud ITS. *J Ambient Intell Humaniz Comput* 7(2):205–220
- Zhang Y, Huang GQ, Sun S, Yang T (2014) Multi-agent based real-time production scheduling method for radio frequency identification enabled ubiquitous shopfloor environment. *Comput Ind Eng* 76(1):89–97
- Zhang Y, Zhang G, Du W et al (2015) An optimization method for shopfloor material handling based on real-time and multi-source manufacturing data. *Int J Prod Econ* 165:282–292
- Zhang G, Zhang Y, Xu X, Zhong RY (2017a) An augmented Lagrangian coordination method for optimal allocation of cloud manufacturing services. *J Manuf Syst* 1–12. <https://doi.org/10.1016/j.jmsy.2017.11.008>
- Zhang Y, Ren S, Liu Y et al (2017b) A framework for Big Data driven product lifecycle management. *J Clean Prod* 159:229–240
- Zhang Y, Ren S, Liu Y, Si S (2017c) A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *J Clean Prod* 142:626–641
- Zhang Y, Guo Z, Lv J, Liu Y (2018a) A framework for smart production-logistics systems based on CPS and industrial IoT. *IEEE Trans Ind Inform*. <https://doi.org/10.1109/TII.2018.2845683>
- Zhang Y, Liu S, Liu Y et al (2018b) The ‘Internet of Things’ enabled real-time scheduling for remanufacturing of automobile engines. *J Clean Prod* 185:562–575
- Zhang Y, Ma S, Yang H, Lv J, Liu Y (2018c) A big data driven analytical framework for energy-intensive manufacturing industries. *J Clean Prod* 197:57–72
- Zhang Y, Zhu Z, Lv J (2018d) CPS-based smart control model for shop-floor material handling. *IEEE Trans Ind Inform* 14(4):1764–1775
- Zhong RY, Huang GQ, Lan S et al (2015) A big data approach for logistics trajectory discovery from RFID-enabled production data. *Int J Prod Econ* 165:260–272

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.