https://dx.doi.org/10.1016/j.jclepro.2020.123144

This is the Pre-Published Version.

An integrated framework for active discovery and optimal allocation of smart manufacturing services

Geng Zhang[a,b], Chun-Hsien Chen[a**], Pai Zheng[c*], Ray Y. Zhong[d]

a School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore

b Delta-NTU Corporate Laboratory for Cyber-Physical System, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

c Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong

d Department of Industrial and manufacturing systems Engineering, The University of Hong Kong, Hong Kong

[*] Corresponding author: Pai Zheng pai.zheng@polyu.edu.hk, Chun-Hsien Chen: MCHchen@ntu.edu.sg

© 2020. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/.

An integrated framework for active discovery and optimal allocation of smart manufacturing services

Abstract

Smart manufacturing is gradually recognized and widely adopted due to the promising features of sustainability, flexibility, and collaboration. Service discovery and allocation in smart manufacturing aim to provide on-demand manufacturing capabilities for meeting customized production requirements. They are tightly coupled in practice, whereas they are usually considered as two independent processes and investigated separately in most research. Meanwhile, the collaboration relationship and decision autonomy of service providers are rarely taken into account to perform sustainable and flexible production. To deal with these challenges, this paper proposes an integrated framework to holistically describe the active discovery and optimal allocation of smart manufacturing services. A mechanism is designed to consider the collaborative relationship of manufacturing resource and promote collaborative production. The distributed optimization model based on analytical target cascading method is introduced to maintain the decision autonomy of service providers and achieve the optimal allocation of smart manufacturing services. A case study is further provided to demonstrate the effectiveness of the proposed framework.

Keywords: smart manufacturing service; active discovery; optimal allocation; analytical target cascading

1 Introduction

Currently, due to the significant revolutions in economic globalization and market competition, industrial enterprises are changing their business modes to meet the raised personalized product demands and sustainable manufacturing requirements. With the rapid development of information and communication technologies (ICT), such as Internet of Things (IoT) (Asghari et al., 2019), cloud computing (Botta et al., 2016), cyber physical system (CPS) (Serpanos, 2018), and digital twin (Schleich et al., 2017), many service-oriented smart manufacturing paradigms (e.g. IoT-based manufacturing (Yang et al., 2019), cloud manufacturing (Ren et al., 2017; Tao et al., 2014; Xu, 2012), social manufacturing (Ding et al., 2018; Leng and Jiang, 2016)) have been proposed to facilitate industrial enterprises to implement sustainable production.

Smart manufacturing service (SMS) is a basic element in the newly emerged service-oriented manufacturing paradigms. The efficient management of SMSs plays an important role in promoting smart manufacturing. Many research work has been conducted in this area, such as service management platforms (Alexopoulos et al., 2018), service models (Quintanilla et al., 2016), and service management methods (Lartigau et al., 2015). Although great progress has been made in improving enterprises' service management level, some challenges still exist in realizing smart manufacturing, especially in the active discovery and optimal allocation of SMSs. Firstly, the discovery and allocation of SMSs are tightly coupled in practice. Nevertheless, they are usually considered as two independent processes and investigated separately in most existing studies. A holistic description of the discovery and allocation of SMSs is thus essential for improving the services management level. Secondly, collaborative production is one of the most important characteristics of smart manufacturing. Existing investigations rarely consider the collaborative relationship between manufacturing resources, which will affect the efficiency of SMS management. Thirdly, centralized methods have been widely used to tackle service allocation problems. However, they can hardly protect the decision autonomy of service providers because only one decision model is adopted in the allocation process. Since decision autonomy is the core factor for service providers to keep their flexibility and sustainable advantages during the production processes, distributed optimization strategies are hence needed to facilitate the allocation process of SMSs.

Given the aforementioned challenges, this research aims to establish an integrated framework for active discovery and optimal allocation of SMSs to promote sustainable, flexible, and collaborative production. Under this framework, manufacturing resources can be timely perceived through the application of ICT. SMSs can thus be obtained based on the dynamic collaboration mechanism of manufacturing resources. Then, the optimal allocation of SMSs can be implemented by a distributed optimization model.

The rest of this paper is organized as follows. Section 2 reviews related literature. Section 3 briefly introduces the overall framework for active discovery and optimal allocation of SMSs. The key technologies for implementing the proposed architecture are illustrated in section 4. A case study is given in section 5. Section 6 draws conclusions and highlights the future work.

2 Literature review

Two streams of literature are relevant to this research, namely the ICT-driven smart manufacturing paradigms and service management in smart manufacturing.

2.1 ICT-driven smart manufacturing paradigms

Due to the promising feature of linking physical world and cyber world, ICT have been increasingly recognized by academia and industry, and widely used in the manufacturing field (Tao and Qi, 2019). Many smart manufacturing paradigms are proposed and the key technologies of which are further investigated as summarized in Table 1.

Table 1 Related studies on ICT driven smart manufacturing paradigms

Paradigms	ICT category	Remarks/Contributions
Internet of manufacturing things	IoT	Capturing real-time data of manufacturing resources and making better-informed enterprises decisions (Bi et al., 2014; Zhang et al., 2015); improving energy-aware production management (Shrouf and Miragliotta, 2015), smart production-logistics (Qu et al., 2014), supply chain management (Ben-Daya et al., 2019), production planning and scheduling (Tian et al., 2019; Wang et al., 2019)
Cloud manufacturing	IoT, cloud computing	Providing on-demand capabilities of distributed manufacturing resources for meeting personalized manufacturing requirements (Tao et al., 2014; Xu, 2012; Zhang et al., 2017b)
Edge and Fog-based	IoT, edge and	Enabling the application of business logic between

manufacturing	fog computing	the downstream data of services and the upstream data of devices in the smart factory (Chen et al., 2018; Wu et al., 2017)
Social manufacturing	IoT, CPS	Extending the crowdsourcing idea to the manufacturing area and establishing a cyber-physical-social connection (Jiang et al., 2016)
Smart product service system	IoT, CPS, big data analytics, Digital twin	Combining smart, connected products and e-services into on-demand solutions to satisfying the needs of individual consumers (Zheng et al., 2018; Liu et al., 2019; Ren et al., 2019; Zheng et al., 2019)
Cyber physical production system	CPS, IoT	Leading to the 4th industrial revolution (Monostori et al., 2016); applying in efficient management for energy-intensive industries (Ma et al., 2019), proactive material handling (Wang et al., 2020), self-organizing and self-adaptive intelligent shop-floor (Zhang et al., 2017a)
Digital twin-based manufacturing	Digital twin, IoT	Facilitating product design (Tao et al., 2019), prognostics of complex equipment (Tao et al., 2018), rapid individualized design (Liu et al., 2019), additive manufacturing (Knapp et al., 2017)

2.2 Service management in smart manufacturing

Service management can achieve the circulation, transaction, and sharing of distributed manufacturing resource capabilities, it thus become a crucial issue for realizing the purpose of smart manufacturing and sustainable production (Tao et al., 2015). The literature related to service management in smart manufacturing can be categorized into three aspects.

The first aspect is about the service management platform. A collaborative cloud platform was designed by integrating additive and subtractive manufacturing resources efficiently and seamlessly to optimize the production plan, improve resource utilization, and avoiding energy waste (Qian et al., 2019). Through the seamless integration of digital design and rapid prototyping, a service platform was constructed to promote the research and development of personalized parts and meet customized requirements. (Xie et al. 2019). A service-oriented simulation integration platform was developed for making hierarchical manufacturing planning, reducing part production variance, and improving production performance (Xu et al. 2016).

The second aspect is about the service model. By integrating time delay and reward/punishment during service transaction, a trust evaluation model was constructed to quantize the service satisfaction and meet customers' manufacturing requirements (Yang et al., 2019). Based on the theory of complex network, a function availability model of high-end manufacturing equipment was developed by integrating product-service to achieve effective service configuration (Chang et al., 2018). In order to improve the agile supply chain, a framework was designed by canonicalizing manufacturing service capabilities and sharing service information models precisely (Kulvatunyou et al., 2015). Through considering the user data, a self-organizing evaluation model was proposed to distinguish effective services and users, then to facilitate the ordering and utilization of manufacturing services (Huang et al., 2018).

The third aspect is the optimization methods for the service allocation or refers to service composition and optimal selection, which is the core part of the literature of service management including two issues. One issue focused on the different service allocation objectives and constraints for specific requirements, such as energy aware for cleaner production (Yang et al., 2019), sustainability consideration for sustainable manufacturing (Wu et al., 2019), quality of service (QoS)-aware (Bouzary and Chen, 2019; Zhang et al., 2019) for optimizing manufacturing cost, time, and quality, urgent task-aware for improving customers' satisfaction (Wang et al., 2018), and long/short-term utility aware for improving the efficiency of resource utilization (Zhang et al., 2019). The other issue emphasized the novel algorithms and approaches for improving the service allocation, such as ensemble optimization approach to select optimal cloud-based service composition (Fazeli et al., 2019), multi-population differential artificial bee colony algorithm for simultaneously optimizing conflicting evaluation criteria (Zhou et al., 2018), extended flower pollination algorithm for solving the multiple-objectives optimization problems (Zhang et al., 2018), clustering network-based approach (Li et al., 2017) to get the candidate service sets effectively and efficiently, social network analysis approach for maximizing the synergy effect of services (Ren et al., 2018).

2.3 Research gaps

Despite the significant progress has been made in aforementioned literature, some research gaps still need to be fulfilled.

1) In terms of the ICT driven smart manufacturing paradigms, most research focused on the macro aspects. Limited effort was made on the discovery and allocation of SMSs. Hence, an integrated framework is proposed to holistically illustrate the processes of services discovery and allocation in the context of smart manufacturing.

2) Regarding to the service model, most studies provided services based on the intrinsic attributes of manufacturing resources and rarely considered the collaboration relationship between different resources. To fill this gap, a collaboration mechanism of manufacturing resources is introduced to provide value-added services for meeting the requirements of smart manufacturing.

3) In respect of the service allocation, most studies adopted centralized methods to obtain the optimization results. Those methods with a single decision model may not be suitable to maintain the decision autonomy and sustainable competitiveness advantages of service providers. Therefore, a distributed hierarchical model is depicted to perform the optimal allocation of SMSs.

3 Overview of active discovery and optimal allocation of SMSs

The discovery and allocation of SMSs can be defined as the integration of encapsulating manufacturing resources into services and providing on-demand services for meeting different manufacturing requirements.

As shown in Fig. 1, the overall framework for active discovery and optimal allocation of SMSs consists of three modules, i.e. manufacturing resource perception, manufacturing service discovery, and manufacturing service allocation.

1) Manufacturing resource perception module is used to perceive the real-time information of manufacturing resources. The sensing capability of the shop-floor can be enhanced by configuring ICT devices (e.g. sensors). Hence, the real-time data of manufacturing resources can be sensed by the inserted sensors. Then the real-time information can be actively acquired by the data processing, and further provided for different manufacturing processes.

2) Manufacturing service discovery module is used to encapsulate the capabilities of manufacturing

resources into smart manufacturing services. Two components are included in this module. The first component is responsible for developing a resource collaboration mechanism. Value-added services can be designed based on the collaboration mechanism. The second component is used to supply value-added services and form a service pool in the cloud-based manufacturing platform.

3) Manufacturing service allocation module is used to allocate optimal services to complete personalized manufacturing tasks. Two components are included in this module. The first component is to identify and illustrate the workflow of the optimal allocation of SMSs. In the second component, a distribution solution is designed and used to achieve optimal allocation of SMSs in a distributed manner.

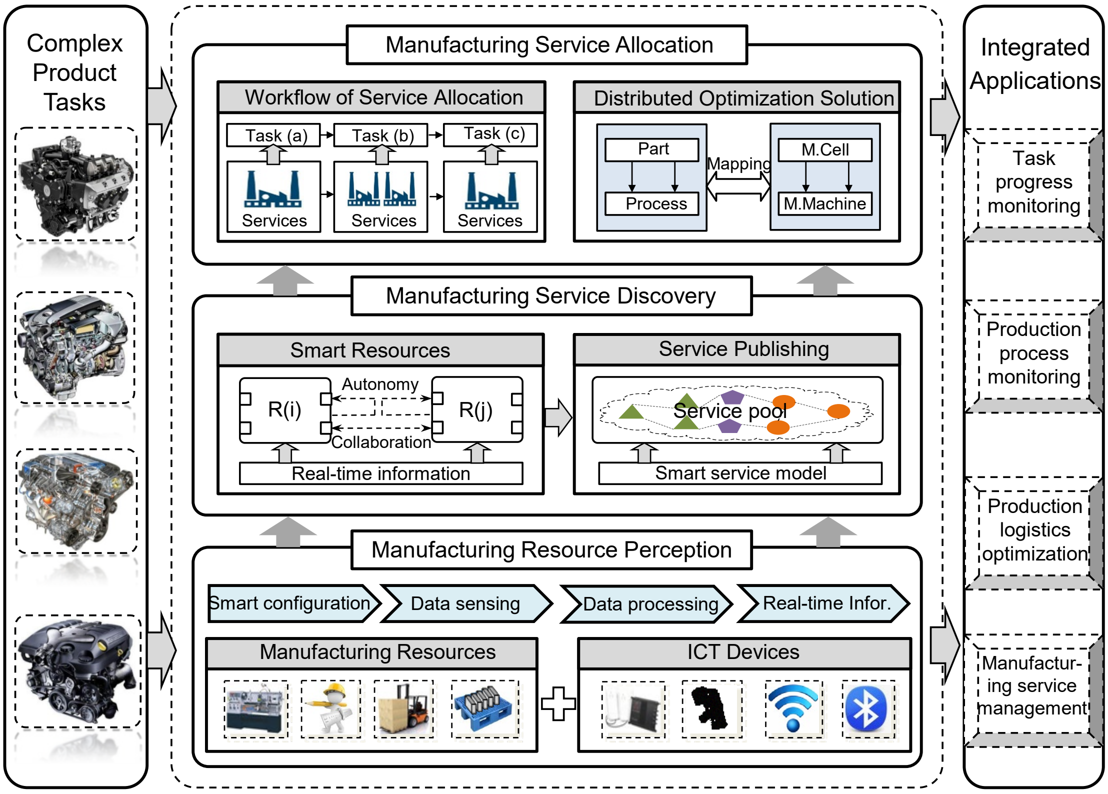

Fig. 1. Framework for active discovery and optimal allocation of SMSs

Fig. 2 describes the working logic of the methodology and coupled relationship between modules in the proposed framework. Manufacturing resource perception module is the basis for the other two modules. Through constructing sensing environment, it can provide accurate and real-time information for the manufacturing service discovery module to encapsulate manufacturing resources into services. Manufacturing resource perception module also supplies real-time information for the manufacturing service allocation module to acquire more reliable candidate manufacturing services to meet personalized requirements, and then obtain more satisfied services allocation results. Manufacturing service discovery module aims to acquire SMSs which are core elements for implementing the optimal allocation of SMSs. Thus, it provides necessary input (i.e. SMSs) for implementing manufacturing service allocation module. Meanwhile, manufacturing service discovery module can give manufacturing resource perception module positive feedback on constructing better sensing environment to acquire more accurate and reliable manufacturing information. Manufacturing service allocation module can offer positive feedback for the other two modules. It can give manufacturing service discovery module suggestions on how to effectively design the services to efficiently participate in the allocation process of SMSs. In addition, it can also provide advice for manufacturing resource module on developing efficient solution of acquiring real-time information to achieve the seamless interaction between different manufacturing stages.

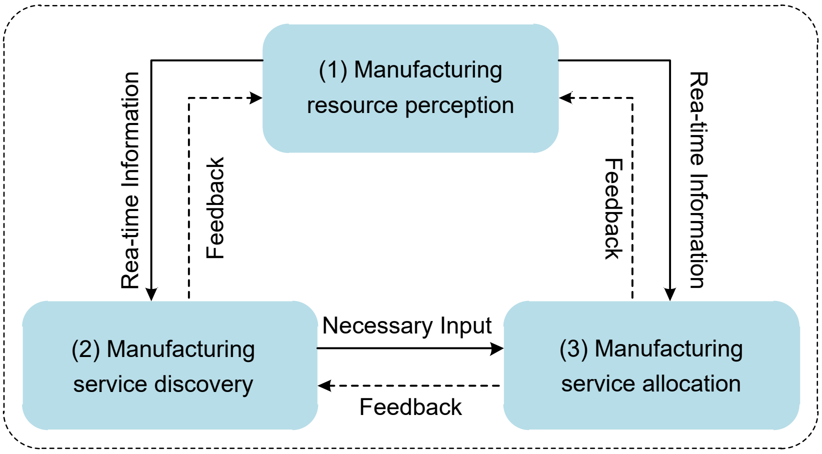

Fig. 2. Relationship between modules in the proposed framework

4 Implementation of active discovery and optimal allocation of SMSs

4.1 Active perception of manufacturing resources

Active perception of manufacturing resources aims to acquire real-time information of manufacturing resources in the shop-floor which can facilitate up-level managers to make better-informed decisions. The procedure for implementing this technology is described in Fig. 3. Three components are included, namely smart configuration, real-time data sensing, real-time data processing.

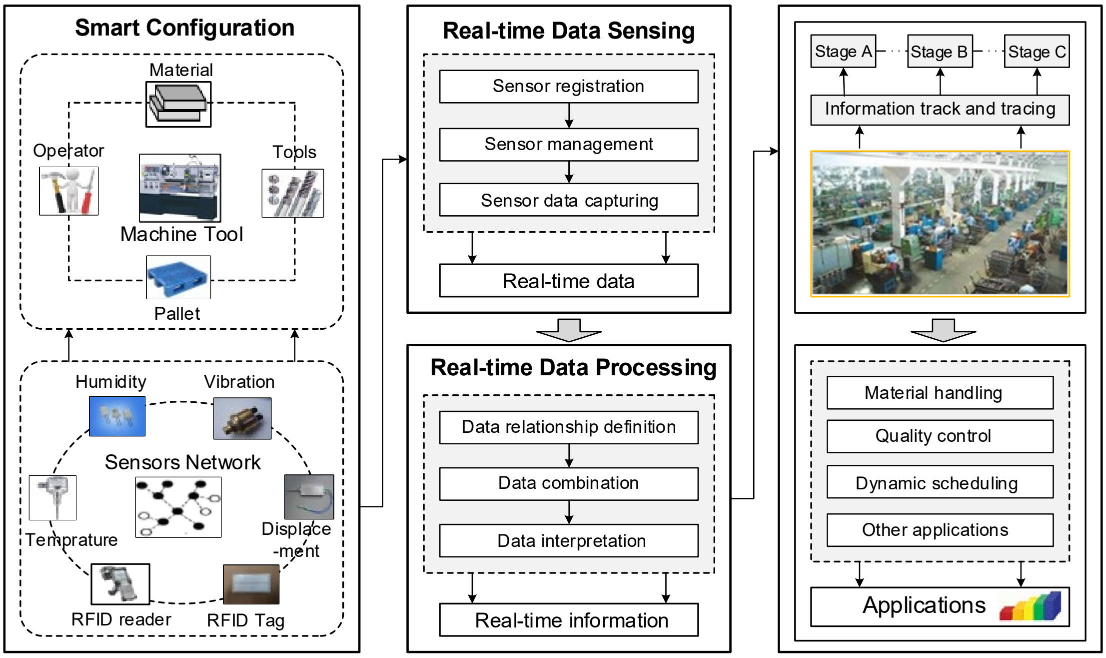

Fig. 3. Active perception of manufacturing resources

4.1.1 *Smart configuration*

This component is used to configure a sensor network to improve the sensing capability of traditional manufacturing resources in the shop-floor by the application of ICT devices. The following steps can be used to execute this component. The first step is to identify the data that is needed to capture, such as the location, temperature, pressure, vibration. The second step is to select appropriate

ICT devices to capture the required data. For example, RFID devices are available for capturing the location data of specific manufacturing resources (e.g. machine tools, materials). Then, a sensor network among related manufacturing resources can be formed based on the selected devices. The third step is to optimize the sensor network. The objective of the optimization model is to minimize the cost of all the selected sensors subject to the constraints for the accuracy and reliability of the captured data.

4.1.2 Real-time data sensing

After the smart configuration, the real-time data of and manufacturing resources and production processes can be actively sensed and captured. Three steps are included in the implementation of real-time data sensing. The first step is sensor registration and aims to input the basic information of the sensors. When a sensor is selected to construct the sensor network, its basic information is registered on the management system, including the sensor type, sensor location, key function parameters of the sensor, etc. The second step is sensor management and consists of two sub-steps. The first sub-step is to use driver methods to activate the selected sensor. Two kinds of methods can be involved in this step, i.e. standard interface (e.g. USB, COM) and specific driver. If the sensor can be activated by the standard interface, it can be used in a plug and play manner. Other, the sensor can only be used after it being activated by the specific driver provided the third-party. Then, the sensorThe second sub-step is to manage the sensors to build a mapping relationship with associated manufacturing resources and meet diverse requirements for data capturing. For example, use an RFID tag to identify the trolley in the shop-floor (i.e. build a mapping relationship between the tag the trolley) and track the location of the trolly (i.e. data capturing requirement). The third step is sensor data capturing. This step aims to capture the data from sensors and bind the sensed data with corresponding manufacturing resources. The data in this step is categorized into metadata and status data. The metadata is used to identify the sensor and manufacturing resources, such as sensor ID, resource ID. The status data can describe the real-time status of corresponding manufacturing resources, such as the real-time location data. The bonded data then will be transferred to the up-level system. For example, a machine tool is equipped with RFID devices, and an operator is equipped with an RFID tag which stores related data of the operator. When the operator enters the sensing space of the machine tool, the real-time location data of the operator can be sensed and captured, then transferred to the following component for processing.

4.1.3 Real-time data processing

This component aims to translate the captured real-time data into meaningful information for decision-making. Three steps are followed to implement this component. The first step is to define the relationship between the captured data and principles for data processing. The second step is data combination to judge whether the captured data could match with the principles. The third step is data interpretation which acts as an engine to processing the data into meaningful information according to related principles. For better understanding, take the following process as an example to explain each step. Firstly, define the principles for machine tool A: 1) When A senses operator data in its processing zone means that the real-time location of the operator is the processing zone of A; 2) When A senses material data in its in-buffer area means that the material is available in A; 3) An operator in the processing zone of A uses the material in the in-buffer area to complete assigned task. Secondly, combine the captured data to judge whether the defined principles should be triggered. Assume that the data of operator B is captured in the processing zone of A and the data of material C is captured in the in-buffer area of A. Thirdly, perform the data interpretation, and the following meaningful information can be acquired: 1) Operator B is currently located in the processing zone of machine tool A; 2) Material C is available in machine tool A; 3) Operator B will complete the assigned task by machine tool A and material C. The acquired information then can be used for up-level decision making.

The real-time information of manufacturing resources can be acquired by executing the aforementioned components. The real-time track and tracing of manufacturing resources can thus be achieved, which can promote applications such as material handling, quality control, and dynamic scheduling.

4.2 Active discovery of SMSs

In order to implement the active discovery of SMSs, two components are designed in this section: 1) collaboration mechanism of manufacturing resources, and 2) services encapsulation and publishing.

4.2.1 Collaboration mechanism of manufacturing resources

The collaborative production aims to achieve real-time information sharing and autonomous decision-making of manufacturing resources. As shown in Fig. 4, a mechanism is designed to perform the collaboration production.

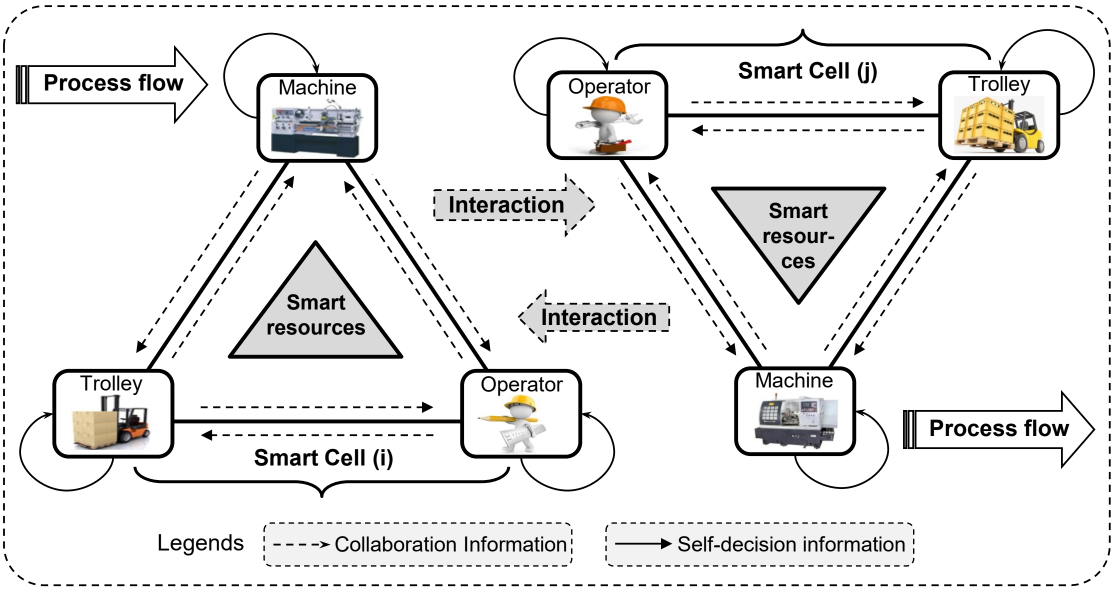

Fig. 4. Collaboration mechanism of manufacturing resources

The proposed collaboration mechanism can be described as two processes. The first process is real-time information sharing which aims to share information with smart manufacturing resources (SMRs) and make better-informed decisions. SMRs mainly refer to the resources which have the capabilities of active sensing and autonomous decision-making. During the real-time information sharing, the SMR sends real-time information to its upstream and downstream SMRs, and receives real-time information from its upstream and downstream SMRs. For example, smart machine tool A in smart cell (*i*) can send the real-time information (e.g. the required material information and task operation information) to the trolley and operator; it can also receive the real-time information from the trolley and operator; moreover, the machine tool can also send/receive the real-time information to/from its downstream machine tool in the same cell (*i*) or different cell (*j*). The interaction between different cells can then be achieved based on the real-time information sharing. The second process is the autonomous decision-making. It can be performed as the following steps. The first step is to identify the real-time status of SMRs according to the shared information. The second step is to optimize the manufacturing processes according to the identified status of SMRs. The third step is to share the local optimization information with related SMRs. For example, machine tool A in cell (*i*) will firstly identify its and the collaborated SMRs' real-time status to judge whether exceptions (e.g. material lack, resources breakdown) occur during production processes. Secondly, machine tool A executes local manufacturing process optimization according to the identified exceptions. For example, machine tool A will optimize its task queue when exceptions occur to its upstream/downstream SMRs to avoid extra work-in-progress (WIP) cost. Then, the obtained local optimization information of machine tool A will be shared with its upstream/downstream SMRs.

4.2.2 Services encapsulation and publishing

This component aims to construct the service model of SMRs and publish it on the cloud-based platform. A shown in Fig. 5, two modules are included to implement this component, i.e. 1) service encapsulation and 2) publishing services.

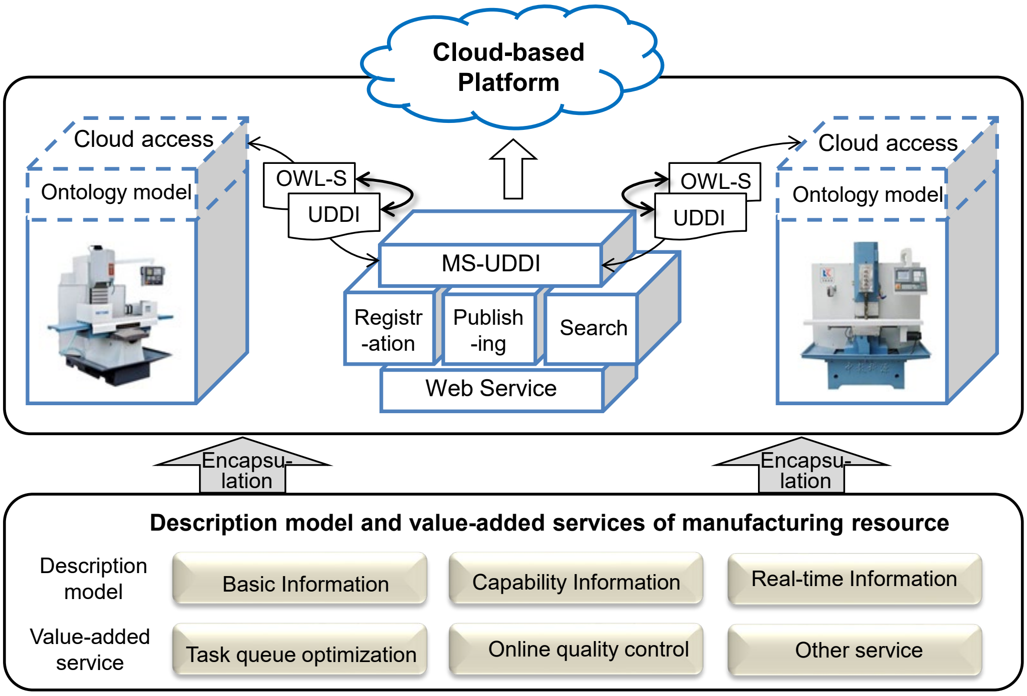

Fig. 5. Service encapsulation and publishing

■ Services encapsulation

Services encapsulation aims to encapsulate SMRs into SMSs by constructing the description model and designing value-added services of SMRs.

The description model is to describe the capabilities of an SMR . The description model of an SMR consists of four parts, i.e. basic information, processing capability information, real-time status information, and quality of service information.

● Basic information of SMR

Basic information of an SMR is the basic part of the description model, which could facilitate up-level managers to find the service provider efficiently. The basic information mainly includes the information of name, specification, type, identity document (ID), etc.

● Processing capability information of SMR

Processing capability information is the core part of the description model. It is the most important evaluation criterion to judge whether the provided services can meet manufacturing requirements and participate in the optimal allocation process of SMSs. The processing capability information includes processing methods (e.g. machining and heat treatment), machining features (e.g. hole, groove, and plane), machinable materials (e.g. steel, cast iron, and alloy), etc.

- Real-time status information of SMR

The real-time status information is an important part of the description model, which provides accurate and timely information for the optimal allocation of SMSs. The real-time information of the SMR includes service status (e.g. working, idle, maintenance), material information (e.g. material kinds, material quantity), task queue (e.g. task ID, due time, task progress), workload, etc.

- Quality of service information of SMR

The quality of service information reflects the service level of the SMR. It is the key parameter to evaluate the candidate services and form the optimal allocation results of SMSs. The quality of service information includes manufacturing cost, manufacturing time, product quality, service times, customer satisfaction, on-time delivery, etc.

Based on the collaboration mechanism of SMRs, some value-added services can be designed, such as task queue optimization, online quality control, and active material handling.

- Task queue optimization

This service aims to timely optimize the local task queue of SMRs according to the shared real-time information. The general optimization model of this service can be stated as formula (1)-(3). $\boldsymbol{OptQue}_{SMR}$ represents the optimal task queue of the SMR, $\boldsymbol{TSet}_{SMR}$ represents the assigned task set of the SMR, $\boldsymbol{I}_{SMR}$ represents the shared information acquired the SMR, $\boldsymbol{G}_{SMR}$ and $\boldsymbol{H}_{SMR}$ represent the inequality and equality constraints for completing $\boldsymbol{TSet}_{SMR}$ by SMR.

$$\text{Objective} \quad \boldsymbol{OptQue}_{SMR} = \mathrm{F}_{OptQue}(\boldsymbol{TSet}_{SMR}, \boldsymbol{I}_{SMR}) \tag{1}$$

$$\text{Subject to} \qquad \boldsymbol{G}_{SMR}(\boldsymbol{TSet}_{SMR}, \boldsymbol{I}_{SMR}) \le 0 \tag{2}$$

$$\boldsymbol{H}_{SMR}(\boldsymbol{TSet}_{SMR}, \boldsymbol{I}_{SMR}) = 0 \tag{3}$$

- Online quality control

This service aims to monitor and diagnose the processing quality of SMRs online. The workflow of online quality control can be described as follows. Firstly, set up an allowable deviation $\boldsymbol{\delta}$ for the processing quality. Secondly, get the standard quality requirement $\boldsymbol{\tau}_S$ and real-time quality information $\boldsymbol{\tau}_R$. Thirdly, judge whether exceptions of processing quality occur according to formula (4). When the exceptions are monitored or diagnosed, the up-level system will be informed, and measures will be taken to tackle the exceptions.

$$\text{If} \begin{cases} \|\boldsymbol{\tau}_R - \boldsymbol{\tau}_S\| \le \boldsymbol{\delta}, \text{exceptions do not occur.} \\ \|\boldsymbol{\tau}_R - \boldsymbol{\tau}_S\| > \boldsymbol{\delta}, \text{exceptions occur.} \end{cases} \tag{4}$$

- Active material handling

This service aims to actively deliver the required material to SMRs to complete the assigned manufacturing tasks. The workflow of active material handling can be described as follows. The first is to check the real-time status of the required material according to the shared information. If the required material is not enough to complete the assigned manufacturing tasks, the second is to send the information (e.g. material type, material quantity) of the required material to the up-level system. Then, trolleys will get material handling tasks and deliver the required material to associated SMRs.

- Publishing services

After obtaining the description model and value-added services of SMRs, the next is to publish them on the cloud-based manufacturing platform. In this section, MS-UDDI (UDDI for manufacturing service) which integrates ontology web language for services (OWL-S) and UDDI technologies is introduced to complete the SMSs publishing. Three modules are included in the MS-UDDI, i.e. registration module, publishing module, and search module. In the registration module, the ontology model of SMSs is constructed by specific development tools (e.g. Protégé) and described by OWL-S. Then, the OWL-S description of SMSs is transformed into UDDI data according to the mapping relationship between the OWL-S profile and UDDI (Srinivasan et al., 2005). In the publishing module, all the related UDDI data of the SMSs is published on the cloud-based manufacturing platform through the standard interface. Then, a service pool is formed by the published SMSs. The information of SMSs can be acquired and invoked by the search module, which can make the SMRs quickly respond to different manufacturing requirements and participate in the allocation process of SMSs. More details about MS-UDDI can be referred to the research work (Zhang et al., 2017b).

4.3 Optimal allocation of SMSs

4.3.1 Working logic of optimal allocation of SMSs

The optimal allocation of SMSs can be considered as the process of SMSs composition and optimal selection for completing specific manufacturing tasks. Three different roles participate in the SMSs allocation process, i.e. customer, resource owner, and cloud-based manufacturing platform. The customer is the demander of SMSs, and submits manufacturing tasks to the cloud-based manufacturing platform. The resource owner acts as the SMSs provider, and provides capabilities for manufacturing tasks. The cloud-based manufacturing platform is responsible for providing transaction mechanisms and tools for the customer and resource owner.

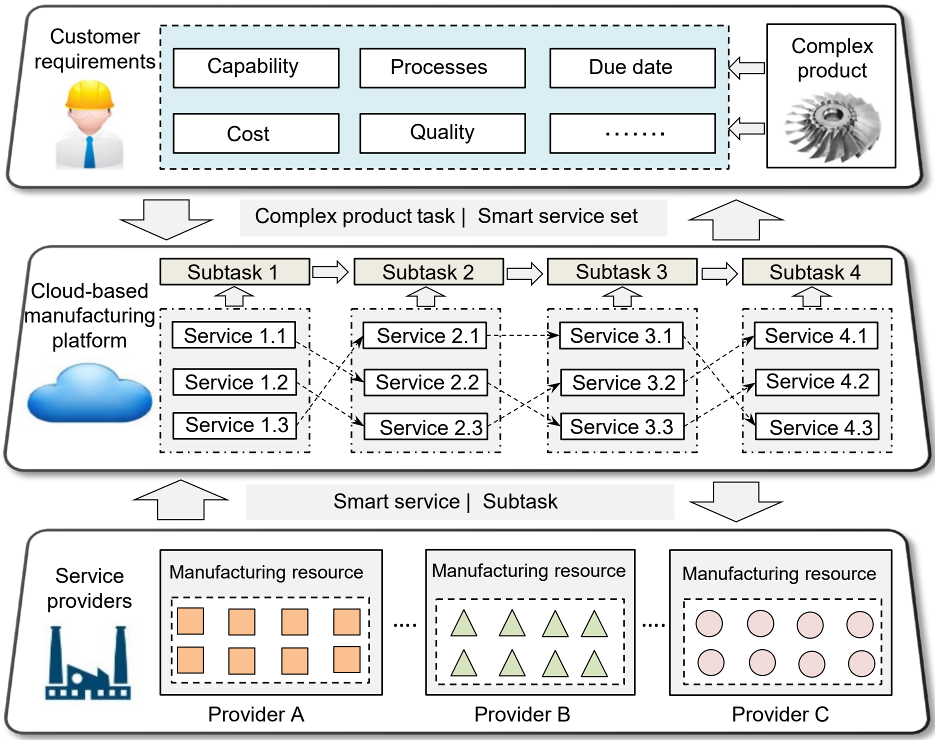

Fig. 6. Working logic of allocation of SMSs

As shown in Fig. 6, the working logic of optimal allocation of SMSs can be described as follows. Firstly, customers submit the manufacturing task information and requirements to the cloud-based platform. Resource owners published the SMSs which are acquired from the SMRs on the cloud-based manufacturing platform, and then an SMSs pool is formed. The submitted task will be decomposed into several subtasks according to its production processes. Based on the manufacturing requirements for the capability and QoS of SMSs, candidate service sets are selected from the service pool. Then, the optimal allocation of SMSs will be implemented by invoking optimization methods. After the allocation, the service provider can get a piece of subtasks to complete, and the customer can obtain a set of SMSs to perform its manufacturing task. All the production information can be timely acquired and shared with the resource owner and customer by the cloud-based manufacturing platform.

4.3.2 Distributed optimization model of SMS allocation

To maintain the decision autonomy of service providers, a distributed optimization model is proposed for the optimal allocation of SMSs.

As shown in Fig. 7, the distributed optimization model is constructed based on the hierarchical structure of complex manufacturing task and manufacturing resources. Three levels are included in the proposed model, i.e. system level, cell level, and machine level. Each level has different elements. Only one element which represents the manufacturing system is at the system level. It is used to get the overall requirements of the submitted manufacturing task, and cascade manufacturing targets to the elements in the cell level. The elements in the cell level represent the manufacturing requirements for the decomposed parts. They aim to select the right manufacturing cells to fulfill the targets cascaded from the system level and cascade targets to the elements at the machine level. The elements in the machine level are employed to select the right machine tools to accomplish specific manufacturing processes to fulfill the cascaded manufacturing targets from the cell level. Each element in the model can perform local optimization within its decision autonomy, and send the feedback to its upper level element. When elements in the lower level cannot fulfill the cascaded targets, the elements in the upper level will adjust the targets according to the feedback. The interaction hence can be achieved through the loop of cascade and feedback. The optimal allocation results can be obtained by the iterative interaction.

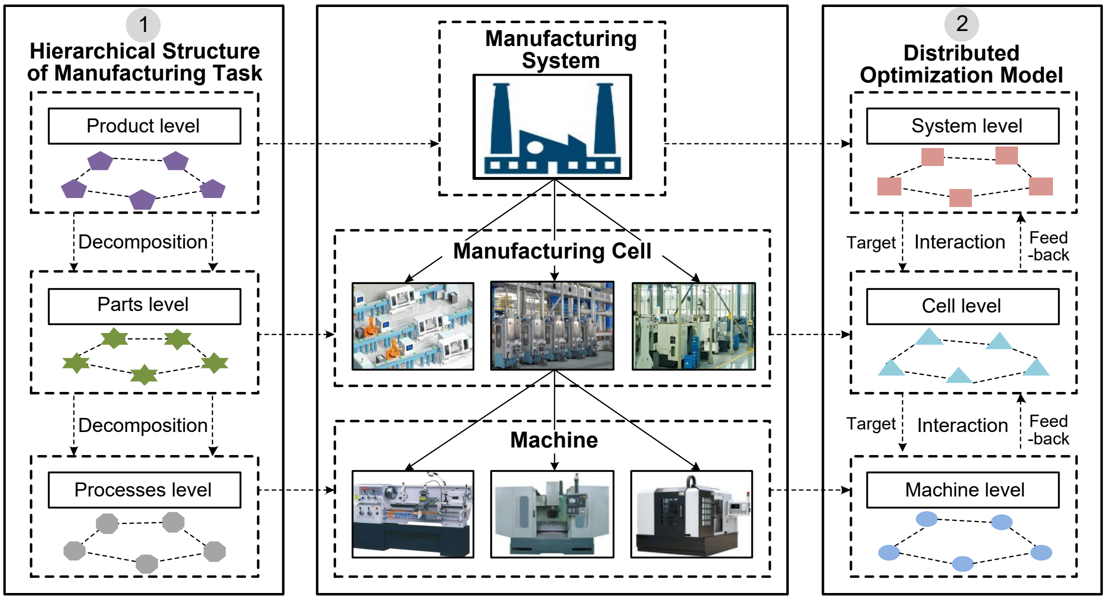

Fig. 7. Distributed optimization model of SMS allocation

4.3.3 Solutions for the distributed optimization model of SMS allocation

Since the constructed optimization model is distributed, a distributed optimization solution/method is needed to solve the proposed model. Analytical target cascading (ATC) is a decomposition-based, multilevel, and hierarchical optimization method. During the implementation of the ATC method, the whole problem is decomposed into a few individual elements in a hierarchical structure. Each element is iteratively solved within its own decision autonomy, and then the final optimization results can be achieved. More detailed information about the implementation of ATC can refer to Kim et al. (2003), Tosserams et al. (2006), and Zhang et al. (2017c).

Given the promising features of ATC in solving engineering problems, it is applicable to solve the distributed optimization model of SMSs allocation. According to the conventions of ATC, the formulation of elements in the model can be presented below and related notations are listed in Table 2.

Table 2 Notations

Notations	Description
$a.b$	bth element in the ath level, a=1,2,3: 1 represents the system level, 2 represents the cell level, and 3 represents the machine level
$\boldsymbol{x}_{a.b}$	Vector of local parameters in the element $a.b$
$f(\overline{\boldsymbol{x}_{a.b}})$	Local optimization objective of bth element in the ath level
$\boldsymbol{g}_{a.b}$	Inequality constraints for the element $a.b$
$\boldsymbol{h}_{a.b}$	Equality constraints for the element $a.b$
$\boldsymbol{T}_{a.b}$	Vector of target variables for element $a.b$
$\boldsymbol{R}_{a.b}$	Vector of response variables for element $a.b$

$T^c_{a.b}$	Cascaded target for the total manufacturing cost of element $a.b$
$T^t_{a.b}$	Cascaded target for the total manufacturing time of element $a.b$
$R^c_{a.b}$	Backtracked response for the total manufacturing cost of element $a.b$
$R^t_{a.b}$	Backtracked response for the total manufacturing time of element $a.b$
$\boldsymbol{v}_{a.b}$	Vector of Lagrangian multiplier parameters of element $a.b$
$v^c_{a.b}$	Lagrangian multiplier parameter for the total manufacturing cost of element $a.b$
$v^t_{a.b}$	Lagrangian multiplier parameter for the total manufacturing time of element $a.b$
$\boldsymbol{w}_{a.b}$	Vector of weight coefficients of element $a.b$
$w^c_{a.b}$	Weight coefficient for the total manufacturing cost of element $a.b$
$w^t_{a.b}$	Weight coefficient for the total manufacturing time of element $a.b$
$O^k_{a.b}$	kth candidate service for completing the element $a.b$
$s^k_{a.b}$	Selection coefficient for service $O^k_{a.b}$ (If service $O^k_{a.b}$ is selected, $s^k_{a.b}=1$; otherwise, $s^k_{a.b}=0$)
$c_{a.b}$	Local manufacturing cost for completing the element $a.b$
$t_{a.b}$	Local manufacturing time for completing the element $a.b$
$c^k_{a.b}$	Manufacturing cost per unit time (for element in the cell level)/Manufacturing cost (for element in other levels) for $O^k_{a.b}$ completing the element $a.b$
$t^k_{a.b}$	Manufacturing time for $O^k_{a.b}$ completing the element $a.b$

- Element in the system level`

Objective function $\min f(\overline{\boldsymbol{x}_{1.1}}) + \| \boldsymbol{w}_{1.1} \circ (\boldsymbol{T}_{1.1} - \boldsymbol{R}_{1.1}) \|_2^2 + \sum_{2.i \in \nu(1.1)} \phi_{AL}^{2.i}(\boldsymbol{T}_{2.i} - \boldsymbol{R}_{2.i})$ (5)

Subject to $\overline{\boldsymbol{x}_{1.1}} = [\boldsymbol{x}_{1.1}, \boldsymbol{R}_{1.1}, (\boldsymbol{T}_{2.i} \mid 2.i \in \nu(1.1))]$ (6)

$$\boldsymbol{g}_{1.1}(\overline{\boldsymbol{x}_{1.1}}) \le \boldsymbol{0} \quad (7)$$

$$\boldsymbol{h}_{1.1}(\overline{\boldsymbol{x}_{1.1}}) = \boldsymbol{0} \quad (8)$$

$$\phi_{AL}^{2.i}(\boldsymbol{T}_{2.i} - \boldsymbol{R}_{2.i}) = \boldsymbol{v}_{2.i}^T(\boldsymbol{T}_{2.i} - \boldsymbol{R}_{2.i}) + \| \boldsymbol{w}_{2.i} \circ (\boldsymbol{T}_{2.i} - \boldsymbol{R}_{2.i}) \|_2^2 \quad (9)$$

Three parts are included in the objective function of the element. The first part is used to minimize the local objective of the system level. The second part is used to minimize the deviation between the system responses and overall manufacturing targets. The third part is used to cascade targets for the elements at the cell level. Eq. (6) presents the variables in this element. Eqs. (7)-(8) present the constraints in the system level. Eq. (9) presents the augmented Lagrangian relaxation technique for the target cascading process, the detailed information of which can refer to Tosserams et al. (2006).

- Element in the cell level

Objective function $\min f(\overline{\boldsymbol{x}_{2.i}}) + \phi_{AL}^{2.i}(\boldsymbol{T}_{2.i} - \boldsymbol{R}_{2.i}) + \sum_{3.j \in \nu(2.i)} \phi_{AL}^{3.j}(\boldsymbol{T}_{3.j} - \boldsymbol{R}_{3.j})$ (10)

$$\text{Subject to} \quad \overline{\boldsymbol{x}_{2.i}} = [\boldsymbol{x}_{2.i}, \boldsymbol{R}_{2.i}, (\boldsymbol{T}_{3.j} \mid 3.j \in \nu(2.i))] \tag{11}$$

$$\boldsymbol{g}_{2.i}(\overline{\boldsymbol{x}_{2.i}}) \leq \boldsymbol{0} \tag{12}$$

$$\boldsymbol{h}_{2.i}(\overline{\boldsymbol{x}_{2.i}}) = \boldsymbol{0} \tag{13}$$

$$\phi_{AL}^{3.j}(\boldsymbol{T}_{3.j} - \boldsymbol{R}_{3.j}) = \boldsymbol{v}_{3.j}^{T}(\boldsymbol{T}_{3.j} - \boldsymbol{R}_{3.j}) + \| \boldsymbol{w}_{3.j} \circ (\boldsymbol{T}_{3.j} - \boldsymbol{R}_{3.j}) \|_2^2 \tag{14}$$

Similarly, three parts are included in the objective function of the element. The first part is to minimize the local objective of element *2.i*. The second part is used to minimize the deviation between the responses of the current element and the cascaded targets from the system level. The third part is used to cascade targets for the elements at the machine level. Eq. (11) presents the variables in element *2.i*. Eqs. (12)-(13) represent the manufacturing constraints for completing element *2.i*. Eq. (14) presents the augmented Lagrangian relaxation technique for the target cascading process.

- Element in the machine level

$$\text{Objective function} \quad \min f(\overline{\boldsymbol{x}_{3.j}}) + \phi_{AL}^{3.j}(\boldsymbol{T}_{2.i} - \boldsymbol{R}_{2.i}) \tag{15}$$

$$\text{Subject to} \quad \overline{\boldsymbol{x}_{2.i}} = [\boldsymbol{x}_{3.j}, \boldsymbol{R}_{3.j}] \tag{16}$$

$$\boldsymbol{g}_{3.j}(\overline{\boldsymbol{x}_{3.j}}) \leq \boldsymbol{0} \tag{17}$$

$$\boldsymbol{h}_{3.j}(\overline{\boldsymbol{x}_{3.j}}) = \boldsymbol{0} \tag{18}$$

Because the machine level is the bottom level of the distributed optimization model, the elements in this level have no targets to cascade. Thus, the objective function of the element in the machine level is composed of two parts. The first part is the local objective of element *3.j*. The second part aims to minimize the deviation between the responses of element *3.j* and the cascaded targets from element *2.i*. Eq. (16) presents the variables in element *3.j*. Eqs. (17)-(18) represent the manufacturing constraints for completing element *3.j*.

After the formulation of each element, the constructed model can be solved according to the steps of implementing the ATC method, and the optimal results of the allocation of SMSs can be achieved.

5 Case study

The case referred in this paper is a task of producing some key parts of an automotive engine (Zhang et al., 2018). Firstly, the effectiveness of ATC in solving the distributed optimization model of SMS allocation is to verify. Then, a prototype system is used to achieve a proof-of-concept validation of the proposed framework.

5.1 Effectiveness of the distributed optimization model

As shown in the upper part of Fig. 8, six subtasks (i.e. valve, crankcase, connecting rod, oil pan, gear housing, and EGR passage) are included for completing the assembly process of the engine. The engine has its specific process flow and logistics flow. For simplicity of understanding, the connecting rod, one of the subtasks, is considered as an example to explain the effectiveness of ATC in solving the distributed optimization model of SMS allocation.

As shown in the lower left part of Fig. 8, when the manufacturing task is submitted to the cloud-based platform, the manufacturing system can receive the manufacturing requirements for producing the connecting rod. Then, the connecting rod is decomposed into three processes, i.e. milling, drilling, and boring. A distributed optimization model in a hierarchical structure can be constructed based on the decomposition of the subtask. As shown in the lower right part of Fig. 8, the distributed model consists of five elements which are in a three-level structure. Element 1.1 is in the system level, element 2.1 is in the cell level, and the other three elements (i.e. 3.1, 3.2, and 3.3) are at the machine level. Based on analyzing the manufacturing capabilities of the registered SMSs in the cloud-based platform, the qualified candidate services are acquired from the service pool and listed near the related elements. Table 3 lists the information of each candidate SMS which comes from an automotive parts association. Note that the data in Table 3 is treated to keep the confidentiality of their key businesses.

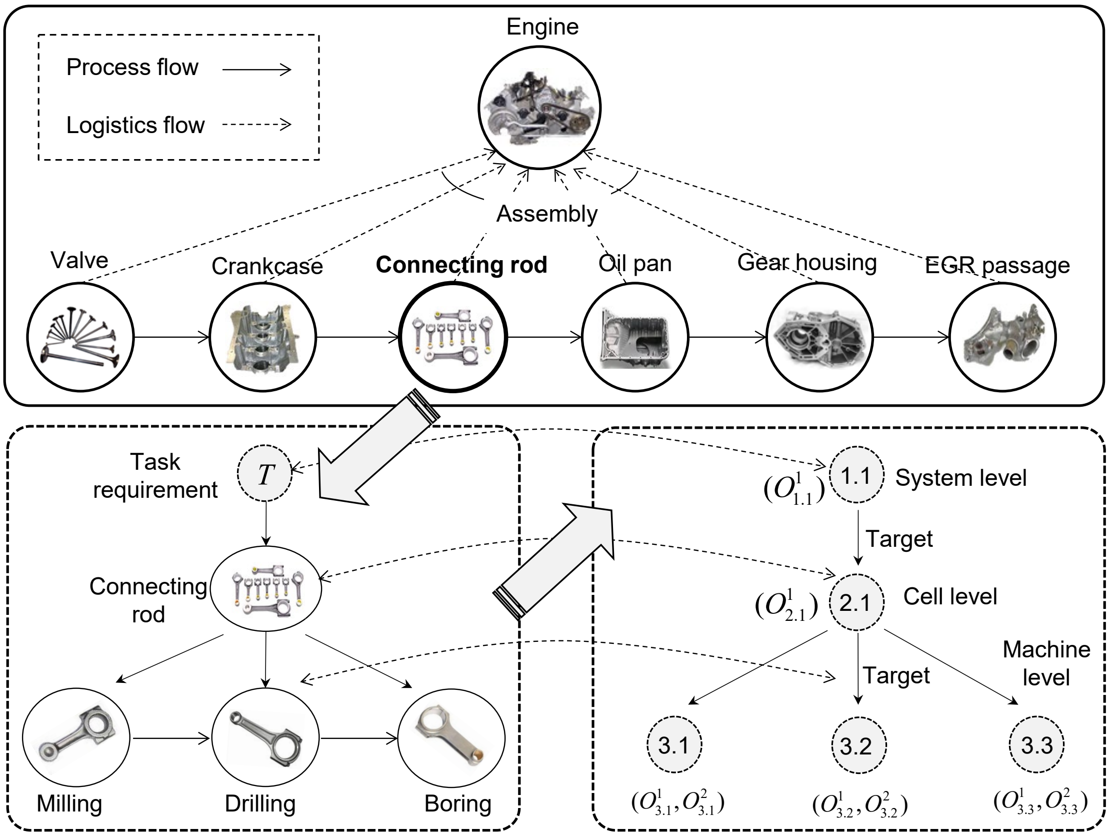

Fig. 8. Distributed model of the case

Table 3 Information of Candidate Services

Element		Candidate manufacturing services			
Number	Element Level	Service Option	Service Style	$c_{a.b}^{k}$	$t_{a.b}^{k}$
1.1	System Level	$O_{1.1}^{1}$	Manufacturing System	-	-
2.1	Cell Level	$O_{2.1}^{1}$	Manufacturing Cell	0.2	-
3.1	Machine Level	$O_{3.1}^{1}$	Machine Tool	14	4
		$O_{3.1}^{2}$	Machine Tool	11	5
3.2	Machine Level	$O_{3.2}^{1}$	Machine Tool	5	4
		$O_{3.2}^{2}$	Machine Tool	6	3

3.3	Machine Level	$O_{3.3}^{1}$	Machine Tool	12	5
		$O_{3.3}^{2}$	Machine Tool	11	8

After obtaining the distributed model of the case, the elements in each level can be formulated according to the formulations described in section 4.3. The formulations of the elements in system level, cell level, and machine level are presented as Table 4, 5, and 6, respectively. In this case, the overall manufacturing targets for total manufacturing cost and time are 31 and 15, and the weight coefficients are 0.3 and 0.7. As shown in Table 4, the overall manufacturing requirements of the connecting rod are included in the objective function of element 1.1. It can be considered as the local objective of element 1.1 at the system level. The total manufacturing cost/time of each element is generally composed of its local manufacturing cost/time and the total manufacturing cost/time of its upstream elements. Because no local manufacturing cost and time are produced in element 1.1, the total manufacturing cost and time of element 1.1 are equal to its upstream element 2.1. As shown in Table 5, no local manufacturing time is produced, thus the total manufacturing time of element 2.1 is equal to the total manufacturing time of its upstream elements at the machine level. The local manufacturing cost of element 2.1 denotes the manufacturing cost which is consumed for running the cell service. Since the machine level is the bottom level in the distributed model, no upstream elements are related to the elements at the machine level. Therefore, as shown in Table 6, the total manufacturing cost/time of the element at the machine level is equal to its local manufacturing cost/time.

The distributed optimization model can be solved according to the procedure of implementing the ATC method. Table 7 lists the obtained optimization results. The total manufacturing cost is 30.8 and the total manufacturing time is 14. Though the results are not exactly the same as the given targets (31 and 15), the deviations are allowable in practice. All the service options are the same as the results obtained by the centralized optimization method (e.g. particle swarm optimization). It means that the proposed distributed model is effective in dealing with the optimal allocation of SMSs.

Table 4 Formulation of the element in the system level

Element in the system level (1.1)	
Objective function	$\min \| \boldsymbol{w}_{1.1} \circ (\boldsymbol{T}_{1.1} - \boldsymbol{R}_{1.1}) \|_2^2 + \phi_{AL}^{2.1}(\boldsymbol{T}_{2.1} - \boldsymbol{R}_{2.1})$
Subject to	$\overline{\boldsymbol{x}_{1.1}} = [\boldsymbol{R}_{1.1}, \boldsymbol{T}_{2.1}]$, $\boldsymbol{T}_{1.1} = [T_{1.1}^{c}, T_{1.1}^{t}]$, $\boldsymbol{R}_{1.1} = [R_{1.1}^{c}, R_{1.1}^{t}]$, $\boldsymbol{T}_{2.1} = [T_{2.1}^{c}, T_{2.1}^{t}]$, $\boldsymbol{v}_{2.1} = [v_{2.1}^{c}, v_{2.1}^{t}]$, $\boldsymbol{w}_{2.1} = [w_{2.1}^{c}, w_{2.1}^{t}]$, $\phi_{AL}^{2.1}(\boldsymbol{T}_{2.1} - \boldsymbol{R}_{2.1}) = \boldsymbol{v}_{2.1}^{T}(\boldsymbol{T}_{2.1} - \boldsymbol{R}_{2.1}) + \| \boldsymbol{w}_{2.1} \circ (\boldsymbol{T}_{2.1} - \boldsymbol{R}_{2.1}) \|_2^2$, $\boldsymbol{R}_{1.1} - \boldsymbol{T}_{2.1} = \boldsymbol{0}$, $\overline{\boldsymbol{x}_{1.1}} \geq \boldsymbol{0}$

Table 5 Formulation of the element in the cell level

Element in the cell level (2.1)	
Objective function	$\min \phi_{AL}^{2.1}(\boldsymbol{T}_{2.1}-\boldsymbol{R}_{2.1})+\sum_{3.j\in\nu(2.1)}\phi_{AL}^{3.j}(\boldsymbol{T}_{3.j}-\boldsymbol{R}_{3.j})$
Subject to	$\overline{\boldsymbol{x}_{2.1}}=[\boldsymbol{x}_{2.1},\boldsymbol{R}_{2.1},(\boldsymbol{T}_{3.j}\mid 3.j\in\nu(2.i))]$, $\boldsymbol{x}_{2.1}=[c_{2.1},t_{2.1}]$, $\boldsymbol{R}_{2.1}=[R_{2.1}^{c},R_{2.1}^{t}]$, $\boldsymbol{T}_{3.j}=[T_{3.j}^{c},T_{3.j}^{t}]$, $\boldsymbol{v}_{3.j}=[v_{3.j}^{c},v_{3.j}^{t}]$, $\boldsymbol{w}_{3.j}=[w_{3.j}^{c},w_{3.j}^{t}]$, $\phi_{AL}^{3.j}(\boldsymbol{T}_{3.j}-\boldsymbol{R}_{3.j})=\boldsymbol{v}_{3.j}^{T}(\boldsymbol{T}_{3.j}-\boldsymbol{R}_{3.j})+\Vert\boldsymbol{w}_{3.j}\circ(\boldsymbol{T}_{3.j}-\boldsymbol{R}_{3.j})\Vert_2^2$, $R_{2.1}^{t}=t_{2.1}^{1}+\sum_{3.j\in\nu(2.1)}T_{3.j}^{t}$, $R_{2.1}^{c}=c_{2.1}^{1}\cdot R_{2.1}^{t}+\sum_{3.j\in\nu(2.1)}T_{3.j}^{c}$, $\overline{\boldsymbol{x}_{2.1}}\geq\boldsymbol{0}$

Table 6 Formulation of the element in the machine level

Element in the machine level (3.*j*)	
Objective function	$\min \phi_{AL}^{3.j}(\boldsymbol{T}_{3.j}-\boldsymbol{R}_{3.j})$
Subject to	$\overline{\boldsymbol{x}_{3.j}}=[\boldsymbol{x}_{3.j},\boldsymbol{R}_{3.j}]$, $\boldsymbol{x}_{3.j}=[c_{3.j},t_{3.j},s_{3.j}^{k},c_{3.j}^{k},t_{3.j}^{k}]$, $\boldsymbol{R}_{3.j}=[R_{3.j}^{c},R_{3.j}^{t}]$, $\boldsymbol{v}_{3.j}=[v_{3.j}^{c},v_{3.j}^{t}]$, $\boldsymbol{w}_{3.j}=[w_{3.j}^{c},w_{3.j}^{t}]$, $R_{3.j}^{t}=t_{3.j}$, $R_{3.j}^{c}=c_{3.j}$, $t_{3.j}=\sum s_{3.j}^{k}\cdot t_{3.j}^{k}$, $c_{3.j}=\sum s_{3.j}^{k}\cdot c_{3.j}^{k}$, $\sum s_{3.j}^{k}=1$, $\overline{\boldsymbol{x}_{3.j}}\geq\boldsymbol{0}$

Table 7 Optimal service allocation results

Element		Service Allocation Results			
Number	Element Level	Service Option	Service Provider	$c_{a.b}^{k}$	$t_{a.b}^{k}$
1.1	System Level	$O_{1.1}^{1}$	Manufacturing System	-	-

2.1	Cell Level	$O_{2.1}^{1}$	Manufacturing Cell	0.2	-
3.1	Machine Level	$O_{3.1}^{2}$	Machine Tool	11	5
3.2	Machine Level	$O_{3.2}^{1}$	Machine Tool	5	4
3.3	Machine Level	$O_{3.3}^{1}$	Machine Tool	12	5
Total Manufacturing Cost			30.8		
Total Manufacturing Time			14		

The most important reason to adopt the distributed optimization model is the feature of maintaining the decision autonomy of service providers. An assumption is made as follows to verify this feature. The SMS $O_{2.1}^{1}$ is assigned many tasks, and the service provider prefers to complete the connecting rod task as soon as possible to release the workload of $O_{2.1}^{1}$. Then, a local objective $\min R_{2.1}^{t}$ will be added to element 2.1 to minimize its total manufacturing time. The objective function of element 2.1 can be revised as Eq. (19).

$$\min R_{2.1}^{t} + \phi_{AL}^{2.1}(\boldsymbol{T}_{2.1} - \boldsymbol{R}_{2.1}) + \sum_{3.j \in v(2.1)} \phi_{AL}^{3.j}(\boldsymbol{T}_{3.j} - \boldsymbol{R}_{3.j}) \tag{19}$$

Then, the optimization results can be obtained by implementing the ATC method. Table 8 compares the SMSs allocation results with and without the decision autonomy of the service provider of $O_{2.1}^{1}$. When the decision autonomy is not considered in the SMSs allocation process, the total manufacturing time $R_{2.1}^{t}$ for $O_{2.1}^{1}$ completing element 2.1 is 14. When the decision autonomy is considered, the $R_{2.1}^{t}$ is 12. It can be seen that the decision autonomy of the service provider of $O_{2.1}^{1}$ can be maintained by the distributed optimization model. However, the total manufacturing cost is increased from 30.8 to 34.4. That because reducing the total manufacturing time of $O_{2.1}^{1}$ brings extra manufacturing cost to complete element 3.1 and 3.2, i.e. $R_{3.1}^{c}$ is increased from 11 to 14, $R_{3.2}^{c}$ is increased from 5 to 6. Though their cost is increased, the manufacturing time is saved.

Table 8 Optimization results of with and without decision autonomy

Element	Without $\min R_{2.1}^{t}$			With $\min R_{2.1}^{t}$		
	Service option	$c_{a.b}^{k}$	$t_{a.b}^{k}$	Service option	$c_{a.b}^{k}$	$t_{a.b}^{k}$
1.1	$O_{1.1}^{1}$	-	-	$O_{1.1}^{1}$	-	-
2.1	$O_{2.1}^{1}$	0.2	0	$O_{2.1}^{1}$	0.2	-
3.1	$O_{3.1}^{2}$	11	5	$O_{3.1}^{1}$	14	4
3.2	$O_{3.2}^{1}$	5	4	$O_{3.2}^{2}$	6	3
3.3	$O_{3.3}^{1}$	12	5	$O_{3.3}^{1}$	12	5

$R_{2.1}^{t}$	14	12
Total manufacturing cost	30.8	34.4
Total manufacturing time	14	12

5.2 Prototype System

In order to better explain the key technologies proposed in this research, a proof-of-concept prototype system is developed to acts as a platform to achieve SMS management. The process of active discovery and optimal allocation of SMSs can be demonstrated by the designed system.

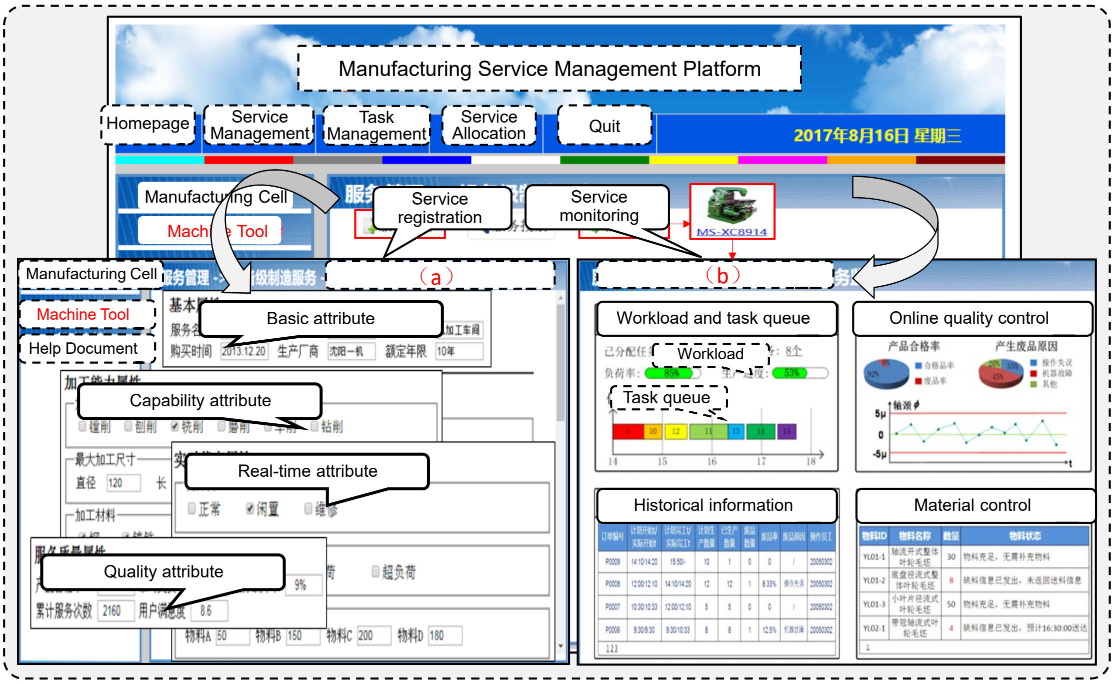

Fig. 9. SMSs management in the prototype system

Fig. 9 shows the SMS management module of the prototype system. A machine tool for milling is taken as an example to explain this module. The first function of this module is to achieve service registration. Four aspects of attributes are required to describe in this module. The basic attribute is used to describe the basic information of the machine tool. The capability attribute is used to describe the processing capability information of the machine tool. The real-time attribute is used to describe the real-time information of the machine tool. The quality attribute is used to describe the quality of service information the machine tool. After completing all the related attributes, the machine tool service can be registered at the prototype system. The second function of this module is to achieve service monitoring. As can be seen in Fig. 9, the real-time production information of the machine tool can be timely monitored, such as the real-time workload and task queue of the machine tool, online quality information, the historical production information, and the required material information.

Fig. 10 shows the workflow of SMS allocation in the prototype system. A task of producing key parts of an automotive engine is taken as the example. Firstly, the customer describes the information (e.g. due time, maximum cost) of the task and submits it to the system. The prototype system partitions the complex task into many subtasks, and assigns each subtask to the qualified service provider. The subtask, such as the connecting rod, will be decomposed into several elements in a hierarchical structure according to its detailed process information, and the qualified candidate SMSs will be acquired from the services pool. After implementing the ATC-based service allocation, the optimization results will be obtained. When the service (e.g. machine tool for milling) gets the manufacturing task, the service provider will timely optimize the task queue based on its real-time status, and perform the obtained task according to the updated queue. During the whole production process, all the real-time manufacturing information of the task can be monitored and shared between the participants (i.e. customer, service provider, and manufacturing system). Once an exception occurs, it can be identified quickly, and the re-allocation of SMSs will be implemented to ensure that the task can be accomplished on time.

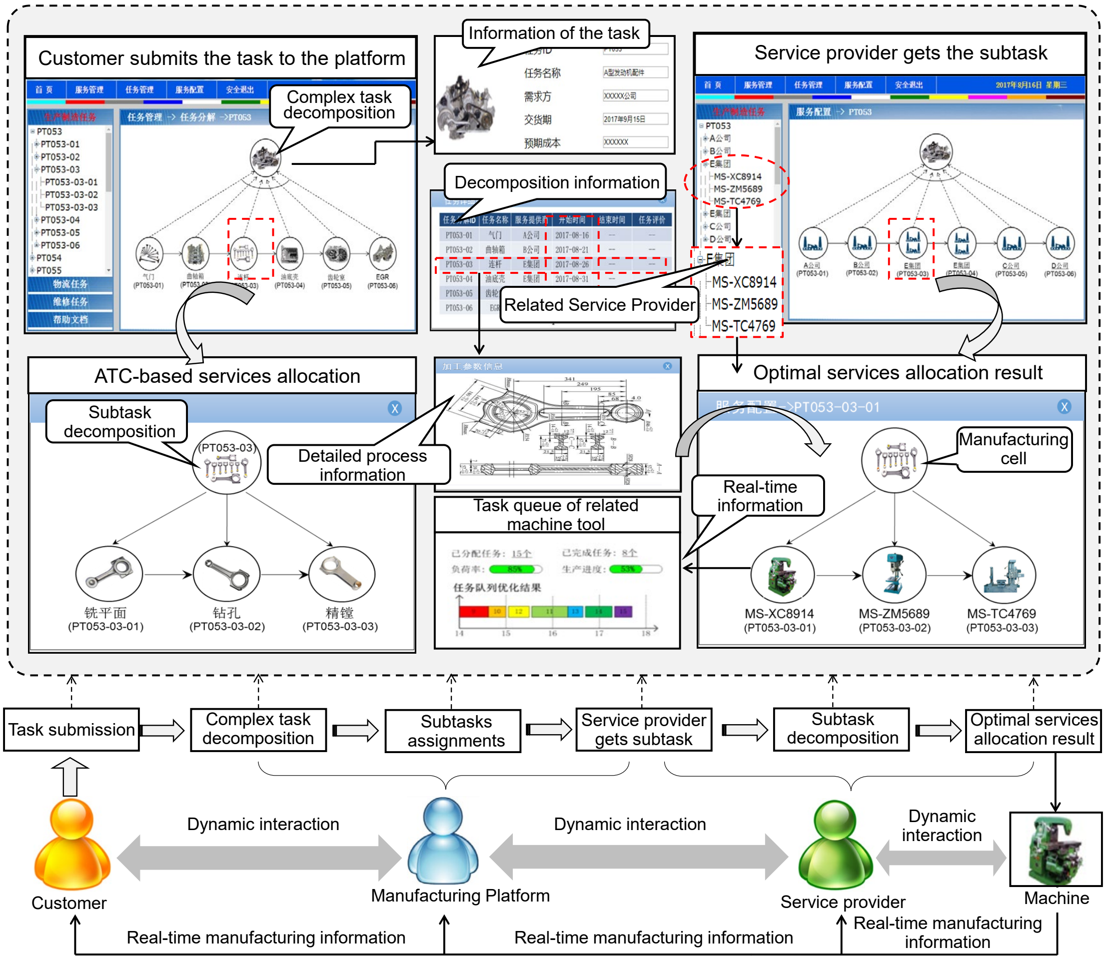

Fig. 10. Workflow of SMS allocation in the prototype system

5.3 Discussions and Managerial implications

Based on the research results, some advantages of the proposed framework for participants in smart manufacturing can be found and discussed as follows. For the service providers, they can have autonomous decision rights during the SMS allocation process. Hence, the collaboration relationship between different service providers can be protected, and the competitiveness of service alliances or communities can be further enhanced. For operators of the cloud-based manufacturing platform, they can get more accurate and transparent information of SMSs and tasks. The utilization rate of services can be increased. Then, idle manufacturing resources and unnecessary energy consumption can be reduced. In addition, the operators can have an alternative optimization tool (i.e. the proposed distributed optimization strategy) to perform global or local SMS allocation regarding different circumstances. For the customers, they can obtain the production information of their tasks through the cloud-based manufacturing platform. Also, they can timely provide their internal feedback for the service providers to improve the production processes. Finally, they can get satisfactory products efficiently.

Moreover, the case study demonstrates that the proposed framework has potential to be applied in industry. Specifically, it can promote the shift of traditional manufacturing paradigm to service-oriented smart manufacturing, and contribute to enhance the sustainability, flexibility, and collaboration of production processes. The implications to improve production management of enterprises can be depicted as follows. Firstly, the proposed framework can be applicable for monitoring and acquiring real-time information from manufacturing processes. The up-level managers can utilize the information to analyze service status of manufacturing resources, identify the task progress, and optimize manufacturing processes. Thus, the proposed framework can be used to support dynamic optimization management of manufacturing execution systems. Secondly, the application of the proposed framework presents the potential of bringing more benefits to enterprises. The decision autonomy of service providers can be maintained under this framework. Hence, the service management can be more flexible, which will benefit the enterprises to keep their sustainable competitive advantages. Thirdly, the proposed framework provides the possibility for facilitating enterprises to efficiently deal with production exceptions. Based on the acquired real-time information, production exceptions can be quickly identified. The distributed optimization model in the framework allows enterprises to solve production exceptions with only considering local optimization parameters, which is more efficient than traditional centralized optimization method. Ultimately, the production of enterprises can proceed normally.

6 Conclusions

Nowadays, smart manufacturing has attracted wide attention from both academia and industry. This research is considered an attempt in the active discovery and optimal allocation of SMSs to facilitate sustainable production.

The major contributions of this paper are summarized as follows. Firstly, an integrated framework is proposed to accomplish the holistic description of the active discovery and optimal allocation of SMSs and improve the SMS management. Secondly, three key technologies are identified to implement the proposed framework, i.e. active perception of manufacturing resources, active discovery of SMSs, and optimal allocation of SMSs. Thirdly, a collaboration mechanism is designed to achieve the real-time information sharing between manufacturing resources, promote collaborative production, and provide a basis for the active discovery of SMSs. Fourthly, the working logic of optimal allocation of SMSs is identified, and a distributed optimization model based on the ATC method is constructed to maintain the decision autonomy of service providers and promote the flexibility and sustainability of smart manufacturing. A case study is further implemented to verify the effectiveness of the proposed methodology in this research.

The future work may follow several aspects. Firstly, how to design a more comprehensive and standard model to describe the services in the context of smart manufacturing? This research is limited to describe the service model of manufacturing resources in the shop-floor. Actually, all the resources in smart manufacturing can provide services for meeting different manufacturing requirements. Secondly, how to accurately describe the submitted manufacturing tasks and accomplish the matching between tasks and candidate services? Effective and accurate description of manufacturing tasks plays an important role in promoting efficient service allocation. Thirdly, how to develop a platform to implement the active discovery and optimal allocation of SMSs in real-life? A platform in practice can not only improve the proposed method but also popularize the applications of smart manufacturing.

Acknowledgements

The authors wish to acknowledge the financial support from the National Research Foundation (NRF) Singapore and Delta Electronics International (Singapore) Pte Ltd., under the Corporate Laboratory@ University Scheme (Ref. SCO-RP1; RCA-16/434) at Nanyang Technological University, Singapore, and the Departmental General Research Funds (G-UAHH), at The Hong Kong Polytechnic University, China. Special thanks go to Guangdong Zhaoqing Automotive Parts Industry Association.

References

Alexopoulos, K., Sipsas, K., Xanthakis, E., Makris, S., Mourtzis, D., 2018. An industrial Internet of things based platform for context-aware information services in manufacturing. Int. J. Comput. Integr. Manuf. 31, 1111–1123. https://doi.org/10.1080/0951192X.2018.1500716

Asghari, P., Rahmani, A.M., Javadi, H.H.S., 2019. Internet of Things applications: A systematic review. Comput. Networks 148, 241–261. https://doi.org/10.1016/j.comnet.2018.12.008

Ben-Daya, M., Hassini, E., Bahroun, Z., 2019. Internet of things and supply chain management: a literature review. Int. J. Prod. Res. https://doi.org/10.1080/00207543.2017.1402140

Bi, Z., Xu, L. Da, Wang, C., 2014. Internet of things for enterprise systems of modern manufacturing. IEEE Trans. Ind. Informatics 10, 1537–1546. https://doi.org/10.1109/TII.2014.2300338

Botta, A., De Donato, W., Persico, V., Pescapé, A., 2016. Integration of Cloud computing and Internet of Things: A survey. Futur. Gener. Comput. Syst. 56, 684–700. https://doi.org/10.1016/j.future.2015.09.021

Bouzary, H., Frank Chen, F., 2019. A hybrid grey wolf optimizer algorithm with evolutionary operators for optimal QoS-aware service composition and optimal selection in cloud manufacturing. Int. J. Adv. Manuf. Technol. 101, 2771–2784. https://doi.org/10.1007/s00170-018-3028-0

Chang, F., Zhou, G., Xiao, X., Tian, C., Zhang, C., 2018. A function availability-based integrated product-service network model for high-end manufacturing equipment. Comput. Ind. Eng. 126, 302–316. https://doi.org/10.1016/j.cie.2018.09.043

Chen, B., Wan, J., Celesti, A., Li, D., Abbas, H., Zhang, Q., 2018. Edge Computing in IoT-Based Manufacturing. IEEE Commun. Mag. 56, 103–109. https://doi.org/10.1109/MCOM.2018.1701231

Ding, K., Jiang, P., Su, S., 2018. RFID-enabled social manufacturing system for inter-enterprise monitoring and dispatching of integrated production and transportation tasks. Robot. Comput. Integr. Manuf. 49, 120–133. https://doi.org/10.1016/j.rcim.2017.06.009

Fazeli, M.M., Farjami, Y., Nickray, M., 2019. An ensemble optimisation approach to service composition in cloud manufacturing. Int. J. Comput. Integr. Manuf. 32, 83–91. https://doi.org/10.1080/0951192X.2018.1550679

Gamboa Quintanilla, F., Cardin, O., L'Anton, A., Castagna, P., 2016. A modeling framework for manufacturing services in Service-oriented Holonic Manufacturing Systems. Eng. Appl. Artif. Intell. 55, 26–36. https://doi.org/10.1016/j.engappai.2016.06.004

Huang, S., Chen, Y., Zhou, H., Gu, X., 2018. Self-organizing evaluation model and algorithm for manufacturing cloud services driven by user behavior. Int. J. Adv. Manuf. Technol. 95, 1549–1565. https://doi.org/10.1007/s00170-018-1651-4

Jiang, P., Leng, J., Ding, K., Gu, P., Koren, Y., 2016. Social manufacturing as a sustainable paradigm for mass individualization. Proc. Inst. Mech. Eng. Part B J. Eng. Manuf. 230, 1961–1968. https://doi.org/10.1177/0954405416666903

Kim, H.M., Michelena, N.F., Papalambros, P.Y., Jiang, T., 2003. Target Cascading in Optimal System Design. J. Mech. Des. 125, 474–480. https://doi.org/10.1115/1.1582501

Knapp, G.L., Mukherjee, T., Zuback, J.S., Wei, H.L., Palmer, T.A., De, A., DebRoy, T., 2017. Building blocks for a digital twin of additive manufacturing. Acta Mater. 135, 390–399. https://doi.org/10.1016/j.actamat.2017.06.039

Kulvatunyou, B., Lee, Y., Ivezic, N., Peng, Y., 2015. A framework to canonicalize manufacturing service capability models. Comput. Ind. Eng. 83, 39–60. https://doi.org/10.1016/j.cie.2015.01.027

Lartigau, J., Xu, X., Nie, L., Zhan, D., 2015. Cloud manufacturing service composition based on QoS with geo-perspective transportation using an improved Artificial Bee Colony optimisation algorithm. Int. J. Prod. Res. 53, 4380–4404. https://doi.org/10.1080/00207543.2015.1005765

Leng, J., Jiang, P., 2016. A deep learning approach for relationship extraction from interaction context in social manufacturing paradigm. Knowledge-Based Syst. 100, 188–199. https://doi.org/10.1016/j.knosys.2016.03.008

Li, F., Zhang, L., Liu, Y., Laili, Y., Tao, F., 2017. A clustering network-based approach to service composition in cloud manufacturing. Int. J. Comput. Integr. Manuf. 30, 1331–1342. https://doi.org/10.1080/0951192X.2017.1314015

Liu, B., Zhang, Y., Zhang, G., Zheng, P., 2019. Edge-cloud orchestration driven industrial smart product-service systems solution design based on CPS and IIoT. Adv. Eng. Informatics 42. https://doi.org/10.1016/j.aei.2019.100984

Liu, Q., Zhang, H., Leng, J., Chen, X., 2019. Digital twin-driven rapid individualised designing of automated flow-shop manufacturing system. Int. J. Prod. Res. 57, 3903–3919. https://doi.org/10.1080/00207543.2018.1471243

Ma, S., Zhang, Y., Lv, J., Yang, H., Wu, J., 2019. Energy-cyber-physical system enabled management for energy-intensive manufacturing industries. J. Clean. Prod. 226, 892–903. https://doi.org/10.1016/j.jclepro.2019.04.134

Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., Ueda, K., 2016. Cyber-physical systems in manufacturing. CIRP Ann. 65, 621–641. https://doi.org/10.1016/j.cirp.2016.06.005

Qian, C., Zhang, Y., Liu, Y., Wang, Z., 2019. A cloud service platform integrating additive and subtractive manufacturing with high resource efficiency. J. Clean. Prod. 241, 118379(1)–118379(16). https://doi.org/10.1016/j.jclepro.2019.118379

Qu, T., Lei, S., Chen, Y., Wang, Z., Luo, H., Huang, G.Q., 2014. Internet-of-Things-enabled smart production logistics execution system based on cloud manufacturing, in: ASME 2014 International Manufacturing Science and Engineering Conference, MSEC 2014 Collocated with the JSME 2014 International Conference on Materials and Processing and the 42nd North American Manufacturing Research Conference. https://doi.org/10.1115/MSEC2014-4194

Ren, L., Zhang, L., Wang, L., Tao, F., Chai, X., 2017. Cloud manufacturing: key characteristics and applications. Int. J. Comput. Integr. Manuf. 30, 501–515. https://doi.org/10.1080/0951192X.2014.902105

Ren, M., Ren, L., Jain, H., 2018. Manufacturing service composition model based on synergy effect: A social network analysis approach. Appl. Soft Comput. J. 70, 288–300. https://doi.org/10.1016/j.asoc.2018.05.039

Ren, S., Zhang, Y., Liu, Y., Sakao, T., Huisingh, D., Almeida, C.M.V.B., 2019. A comprehensive review of big data analytics throughout product lifecycle to support sustainable smart manufacturing: A framework, challenges and future research directions. J. Clean. Prod. 210, 1343–1365. https://doi.org/10.1016/j.jclepro.2018.11.025

Schleich, B., Anwer, N., Mathieu, L., Wartzack, S., 2017. Shaping the digital twin for design and production engineering. CIRP Ann. - Manuf. Technol. 66, 141–144. https://doi.org/10.1016/j.cirp.2017.04.040

Serpanos, D., 2018. The Cyber-Physical Systems Revolution. Computer (Long. Beach. Calif). 51, 70–73. https://doi.org/10.1109/MC.2018.1731058

Shrouf, F., Miragliotta, G., 2015. Energy management based on Internet of Things: Practices and framework for adoption in production management. J. Clean. Prod. 100, 235–246. https://doi.org/10.1016/j.jclepro.2015.03.055

Srinivasan, N., Paolucci, M., Sycara, K., 2005. An efficient algorithm for OWL-S based semantic search in UDDI. Semant. Web Serv. Web Process Compos. 3387, 96–110.

Tao, F., Cheng, Y., Xu, L. Da, Zhang, L., Li, B.H., 2014. CCIoT-CMfg: Cloud computing and internet of things-based cloud manufacturing service system. IEEE Trans. Ind. Informatics 10, 1435–1442. https://doi.org/10.1109/TII.2014.2306383

Tao, F., Qi, Q., 2019. New IT driven service-oriented smart manufacturing: Framework and characteristics. IEEE Trans. Syst. Man, Cybern. Syst. 49, 81–91. https://doi.org/10.1109/TSMC.2017.2723764

Tao, F., Sui, F., Liu, A., Qi, Q., Zhang, M., Song, B., Guo, Z., Lu, S.C.Y., Nee, A.Y.C., 2019. Digital twin-driven product design framework. Int. J. Prod. Res. 57, 3935–3953. https://doi.org/10.1080/00207543.2018.1443229

Tao, F., Zhang, L., Liu, Y., Cheng, Y., Wang, L., Xu, X., 2015. Manufacturing Service Management in Cloud Manufacturing: Overview and Future Research Directions. J. Manuf. Sci. Eng. Trans. ASME 137, 040912(1)–040912(11). https://doi.org/10.1115/1.4030510

Tao, F., Zhang, M., Liu, Y., Nee, A.Y.C., 2018. Digital twin driven prognostics and health management for complex equipment. CIRP Ann. 67, 169–172. https://doi.org/10.1016/j.cirp.2018.04.055

Tian, S., Wang, T., Zhang, L., Wu, X., 2019. The Internet of Things enabled manufacturing enterprise information system design and shop floor dynamic scheduling optimisation. Enterp. Inf. Syst. https://doi.org/10.1080/17517575.2019.1609703

Tosserams, S., Etman, L.F.P., Papalambros, P.Y., Rooda, J.E., 2006. An augmented Lagrangian relaxation for analytical target cascading using the alternating direction method of multipliers. Struct. Multidiscip. Optim. 31, 176–189. https://doi.org/10.1007/s00158-005-0579-0

Wang, J., Zhang, Y., Liu, Y., Wu, N., 2019. Multiagent and bargaining-game-based real-time scheduling for internet of things-enabled flexible job shop. IEEE Internet Things J. 6, 2518–2531. https://doi.org/10.1109/JIOT.2018.2871346

Wang, W., Zhang, Y., Zhong, R.Y., 2020. A proactive material handling method for CPS enabled shop-floor. Robot. Comput. Integr. Manuf. 61, 101849. https://doi.org/10.1016/j.rcim.2019.101849

Wang, Y., Dai, Z., Zhang, W., Zhang, S., Xu, Y., Chen, Q., 2018. Urgent task-aware cloud manufacturing service composition using two-stage biogeography-based optimisation. Int. J. Comput. Integr. Manuf. 31, 1034–1047. https://doi.org/10.1080/0951192X.2018.1493230

Wu, D., Liu, S., Zhang, L., Terpenny, J., Gao, R.X., Kurfess, T., Guzzo, J.A., 2017. A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing. J. Manuf. Syst. 43, 25–34. https://doi.org/10.1016/j.jmsy.2017.02.011

Wu, Y., Jia, G., Cheng, Y., 2019. Cloud manufacturing service composition and optimal selection with sustainability considerations: a multi-objective integer bi-level multi-follower programming approach. Int. J. Prod. Res. https://doi.org/10.1080/00207543.2019.1665203

Xie, J., Yang, Z., Wang, X., Lai, X., 2019. A cloud service platform for the seamless integration of digital design and rapid prototyping manufacturing. Int. J. Adv. Manuf. Technol. 100, 1475–1490. https://doi.org/10.1007/s00170-018-2809-9

Xu, D., Nageshwaraniyer, S.S., Son, Y.J., 2016. A service-oriented simulation integration platform for hierarchical manufacturing planning and control. Int. J. Prod. Res. 54, 7212–7230. https://doi.org/10.1080/00207543.2016.1221535

Xu, X., 2012. From cloud computing to cloud manufacturing. Robot. Comput. Integr. Manuf. 28, 75–86. https://doi.org/10.1016/j.rcim.2011.07.002

Yang, H., Kumara, S., Bukkapatnam, S.T.S., Tsung, F., 2019. The internet of things for smart manufacturing: A review. IISE Trans. 51, 1190–1216. https://doi.org/10.1080/24725854.2018.1555383

Yang, X., Wang, S., Yang, B., Ma, C., Kang, L., 2019. A service satisfaction-based trust evaluation model for cloud manufacturing. Int. J. Comput. Integr. Manuf. 32, 533–545. https://doi.org/10.1080/0951192X.2019.1575982

Yang, Y., Yang, B., Wang, S., Liu, W., Jin, T., 2019. An Improved Grey Wolf Optimizer Algorithm for Energy-Aware Service Composition in Cloud Manufacturing. Int. J. Adv. Manuf. Technol. 105, 3079–3091. https://doi.org/10.1007/s00170-019-04449-9

Zhang, G., Zhang, Y., Xu, X., Zhong, R.Y., 2018. An augmented Lagrangian coordination method for optimal allocation of cloud manufacturing services. J. Manuf. Syst. 48, 122–133. https://doi.org/10.1016/j.jmsy.2017.11.008

Zhang, S., Xu, Y., Zhang, W., Yu, D., 2019. A new fuzzy QoS-aware manufacture service composition method using extended flower pollination algorithm. J. Intell. Manuf. 30, 2069–2083. https://doi.org/10.1007/s10845-017-1372-9

Zhang, W., Yang, Y., Zhang, S., Yu, D., Li, Y., 2018. Correlation-aware manufacturing service composition model using an extended flower pollination algorithm. Int. J. Prod. Res. 56, 4676–4691. https://doi.org/10.1080/00207543.2017.1402137

Zhang, Y., Qian, C., Lv, J., Liu, Y., 2017a. Agent and cyber-physical system based self-organizing and self-adaptive intelligent shopfloor. IEEE Trans. Ind. Informatics 13, 737–747. https://doi.org/10.1109/TII.2016.2618892

Zhang, Y., Tao, F., Liu, Y., Zhang, P., Cheng, Y., Zuo, Y., 2019. Long/Short-Term Utility Aware Optimal Selection of Manufacturing Service Composition Toward Industrial Internet Platforms. IEEE Trans. Ind. Informatics 15, 3712–3722. https://doi.org/10.1109/TII.2019.2892777

Zhang, Y., Zhang, G., Liu, Y., Hu, D., 2017b. Research on services encapsulation and virtualization access model of machine for cloud manufacturing. J. Intell. Manuf. 28, 1109–1123. https://doi.org/10.1007/s10845-015-1064-2

Zhang, Y., Zhang, G., Qu, T., Liu, Y., Zhong, R.Y., 2017c. Analytical target cascading for optimal configuration of cloud manufacturing services. J. Clean. Prod. 151, 330–343. https://doi.org/10.1016/j.jclepro.2017.03.027

Zhang, Y., Zhang, G., Wang, J., Sun, S., Si, S., Yang, T., 2015. Real-time information capturing and integration framework of the internet of manufacturing things. Int. J. Comput. Integr. Manuf. 28, 811–822. https://doi.org/10.1080/0951192X.2014.900874

Zheng, P., Lin, T.-J., Chen, C.-H., Xu, X., 2018. 1 A systematic design approach for service innovation of smart product-service systems. Journal of Cleaner Production 201, 657–667. https://doi.org/10.1016/j.jclepro.2018.08.101

Zheng, P., Lin, Y., Chen, C.H., Xu, X., 2019. Smart, connected open architecture product: an IT-driven co-creation paradigm with lifecycle personalization concerns. Int. J. Prod. Res. 57, 2571–2584. https://doi.org/10.1080/00207543.2018.1530475

Zhou, J., Yao, X., Lin, Y., Chan, F.T.S., Li, Y., 2018. An adaptive multi-population differential artificial bee colony algorithm for many-objective service composition in cloud manufacturing. Inf. Sci. (Ny). 456, 50–82. https://doi.org/10.1016/j.ins.2018.05.009