Computers and Electrical Engineering 000 (2016) 1–11

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Data quality management for service-oriented manufacturing cyber-physical systems

Zhiting Song[a], Yanming Sun[a], Jiafu Wan[b,*], Peipei Liang[a]

[a] School of Business Administration, South China University of Technology, China
[b] School of Mechanical & Automotive Engineering, South China University of Technology, China

ARTICLE INFO

Article history:
Received 7 April 2016
Revised 17 August 2016
Accepted 17 August 2016
Available online xxx

Keywords:
Data quality
Cyber-physical systems
Service-oriented manufacturing
Workflow nets

ABSTRACT

Service-oriented manufacturing (SOM) is a new worldwide manufacturing paradigm, and a cyber-physical system (CPS) is accepted as a strategic choice of SOM enterprises looking to provide bundles of satisfying products and services to customers. The issue of data quality is common in any CPS and poses great challenges to its efficient operation. This paper focuses on defective data generated by the improper operation of physical and cyber components of a service-oriented manufacturing CPS (SMCPS), and develops effective managerial policies to deal with such data. First, formal semantics of workflow nets (WF-nets) are employed to construct process-oriented ontology for the SMCPS. Second, a two-stage optimization model together with algorithms is designed to find optimal policies that balance local and global management objectives. Finally, our model is illustrated through a case. Results show that the proposed control strategy outperforms one-stage control and random control in guaranteeing data quality and saving control costs.

1. Introduction

Servitization, when integrated into traditional manufacturing, produces a new manufacturing mode, which is known as service-oriented manufacturing (SOM) [1]. Advancements in vertical, horizontal and end-to-end integration covered by Industry 4.0 have led to the widespread use of SOM enterprises in the manufacturing industry [2,3]. Emerging technologies which have ushered in the SOM phenomenon include cyber-physical system (CPS) [4], wireless sensor networks [5,6], cloud computing [7,8], Internet of Things [9,10] and big data [11]. Of these technologies, an efficient CPS with a well-designed structure and intelligent management level is a remarkable feature of competitive SOM enterprises. Establishing such a CPS has become widely adopted as a significant strategy of SOM enterprises attempting to create new values hidden in servitization and to undergird their competitive advantages.

A service-oriented manufacturing CPS (SMCPS) is composed of data-intensive cyber and physical processes [12,13], and designed to supply customers with high-quality products and excellent services. The efficient operation of a SMCPS relies heavily on the quality of data contained in those processes. One of the leading causes of defective data is the failure of data-related SMCPS components, such as computers, communication devices, and machines with embedded sensors, as shown in Fig. 1. Flawed data is and will continue to be the most common issue in any CPS, thus having a hugely negative impact on CPS performance [14]. Under the SOM paradigm, a CPS is a network of diverse and heterogeneous components. In addition, every enterprise is closely connected to other enterprises in the SOM network. Once data is flawed at one network node,

* Corresponding author.
E-mail address: mejwan@scut.edu.cn (J. Wan).

http://dx.doi.org/10.1016/j.compeleceng.2016.08.010
0045-7906/

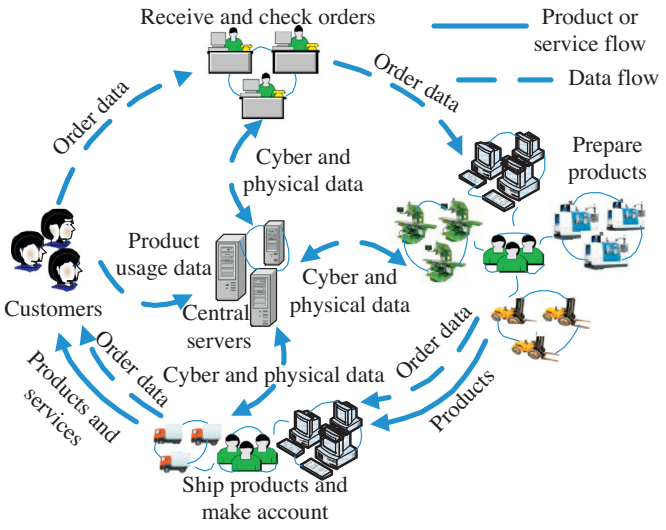

Fig. 1. Framework of a SMCPS.

consequent negative effects (defined as data quality risks [15]) increase exponentially within the network. Improving the data quality of a SMCPS has become an important priority for all SOM enterprises.

A considerable number of studies focus on improving the data quality of a CPS from a technical perspective. Fewer studies focus on the quantitative management of data quality from a managerial perspective. This article presents a methodology to manage flawed data caused by various SMCPS components from a managerial perspective. The proposed methodology models a SMCPS by applying workflow nets (WF-nets). This process in turn enables the formulation of a two-stage optimization problem generating optimal control resource configurations. The two-stage model analyzes data error control, along with the risks caused by data errors. The model also seeks out the optimal control policies which will satisfy the task-level and system-level management objectives by using proposed algorithms.

The novel aspects of this paper are presented as follows: (1) This paper employs WF-nets to formally visualize the structure and model behavior of a SMCPS, which in turn facilitates the capture of the static and dynamic attributes of data errors. We further provide a quantitative analysis of control costs and the risks associated with data errors. (2) Our article establishes both a task-level optimization model (which focuses on local data quality and control costs), and a system-level optimization model (which focuses on global risks and control costs), to balance local and global data quality management objectives. This study contributes to the existing theories and practical aspects of the ex-ante data quality management-based design of a SMCPS and the ex-post data quality control of an existing SMCPS. With effective strategies in place to manage data quality in a SMCPS, manufacturers can make full use of the opportunities arising from combined CPS and servitization, and create steady revenue streams.

The rest of the paper is organized as follows. In Section 2, a review of relevant literature is presented. Section 3 presents process-oriented ontology for a SMCPS. A two-stage optimization model and corresponding algorithms are proposed in Section 4. Section 5 displays a simulation example, and analyzes the results. Finally, Section 6 concludes the research, and suggests directions for future work.

2. Related literature

Two streams of existing literature are relevant to our study. The first one is related to data quality and data errors, and the second one addresses recent efforts to improve the data quality of a CPS.

Data quality, originally defined as data accuracy, is a multidimensional concept beyond accuracy. Although there is no standard definition of data quality dimensions, the frequently mentioned and widely accepted dimensions are accuracy, consistency, timeliness, and completeness [16]. In fact, there exist one-to-one relationships between data quality dimensions and data error types [17]. For example, when data accuracy does not meet requirements, accuracy errors occur. For any SMCPS, data accuracy is the primary operational basis, because ultimately, accuracy represents the objectivity and authenticity of data. As a SMCPS contains numerous heterogeneous informational components generating heterogeneous data [18], data consistency is vital for the exchange and fusion of data. SOM enterprises are characterized by their rapid response to random demands. As such, data timeliness matters significantly in a SMCPS. Therefore, data quality dimensions in this study refer to accuracy, consistency, and timeliness. Accordingly, error types are accuracy errors, consistency errors, and timeliness errors.

Efforts were previously made to improve data quality using various methods. To disseminate sensor data from the physical space in a timely and reliable manner, Jatzkowski et al. [19] incorporated Ethernet and CAN-based communication protocols and their synchronization mechanisms into software architecture. Luo et al. [20] created a fully-fledged data forwarding

Please cite this article as: Z. Song et al., Data quality management for service-oriented manufacturing cyber-physical systems, Computers and Electrical Engineering (2016), http://dx.doi.org/10.1016/j.compeleceng.2016.08.010