Service innovation and smart analytics for Industry 4.0 and big data environment

Jay Lee and Hung An Kao

NSF I/UCRC Center for Intelligent Maintenance Systems (IMS), University of Cincinnati, Cincinnati, OH 45221-0072, USA
* Corresponding author. Tel.: +1-513-556-3412; fax: +1-513-556-4647; E-mail address: jay.lee@uc.edu

Abstract

Today, in an Industry 4.0 factory, machines are connected as a collaborative community. Such evolution requires the utilization of advance prediction tools so that data can be systematically processed into information that can explain the uncertainties and thereby make more “informed” decisions. Cyber-Physical System based manufacturing and service innovations are two inevitable trends and challenges for manufacturing industries. This paper addresses the trends of manufacturing service transformation in big data environment as well as the readiness of smart predictive informatics tools to manage big data to achieve transparency and productivity.

Keywords: Manufacturing servitization; predictive maintenance; industrial big data

1. Introduction

In today’s competitive business environment, companies are facing challenges in dealing with big data issues for rapid decision making for improved productivity. Many manufacturing systems are not ready to manage big data due to the lack of smart analytics tools. Germany is leading a transformation toward 4th Generation Industrial Revolution (Industry 4.0) based on Cyber-Physical System based manufacturing and service innovation. As more software and embedded intelligence are integrated in industrial products and systems, predictive technologies can further intertwine intelligent algorithms with electronics and tether-free intelligence to predict product performance degradation and autonomously manage and optimize product service needs.

Nowadays, smart factories focus mostly on control-centric optimization and intelligence. However, more intelligence can be achieved by interacting with different surrounding systems that have direct impact to machine performance. Achieving such seamless interaction with surrounding systems turns regular machines into self-aware and self-learning machines and consequently improves overall performance and maintenance management. Although the autonomous computing methodology has been implemented successfully in computer science, self-learning machines are still far from implementation in current industries. Transformation from today’s status into more intelligent machines requires further advancement in the science by tackling several fundamental issues. These issues can be divided into five distinct categories as follows:

- Manager and Operator Interaction: Currently, operators control machines, managers design logistic schedules and machines are only performing the assigned tasks. Although these tasks are usually optimized by expert operators and managers but a significantly important factor is missing in these decisions that is the health condition of the machine components.

- Machine Fleet: It is very common that similar or identical machines (machine fleet) are being exposed to completely different working conditions for different tasks. In contrast, most predictive and prognostic methods are designed to support single or limited number of machines and working
conditions. Currently, available prognostic and health management methods are not taking advantage of considering these identical machines as a fleet by gathering worthwhile knowledge from different instances.

• Product and Process Quality: As the final outcome of the manufacturing process, product quality can provide much insight of machine condition via backward reasoning algorithms. Product quality can provide feedback for system management, which can be used to improve production scheduling. Currently such feedback loop does not exist and needs further research.

• Big Data and Cloud: Data management and distribution in Big Data environment is critical on achieving self-aware and self-learning machines. The importance of leveraging additional flexibility and capabilities offered by cloud computing is inevitable but adapting prognostics and health management algorithms to efficiently implement current data management technologies requires further research and development.

• Sensor and Controller Network: Sensors are machine’s gateway to sense its surrounding physical environment. However, sensor failure and degradation may pass wrong and inaccurate readings to decision-making algorithms, which will result in incorrect outcome.

With these issues in mind, the objective of the paper is to review how current manufacturing industries evolve for the upcoming industrial big data environment, and propose the key technology for sustainable innovative service. The paper is organized as follows. Section 2 focuses on trends of service innovation in manufacturing industries and unmet needs of an Industry 4.0 factory. Section 3 describes the proposed self-aware and self-maintenance machine systems based on industrial big data analysis. Section 4 presents two case studies that have been conducted to demonstrate the feasibility of the proposed framework. Section 5 concludes the paper with some perspectives.

2. Trends and unmet needs for Industry 4.0 era

The discovery of new technologies has escorted the industry development from the early adoption of mechanical systems to support production processes, to today’s highly automated assembly lines in order to be responsive and adaptive to current dynamic market requirement and demands. Under Industry 4.0 concept, the astounding growth in the advancement and adoption of information technology and social media networks has increasingly influenced consumers’ perception on product innovation, quality, variety and speed of delivery. This requires establishing the factory with capability of self-awareness, self-prediction, self-comparison, self-reconfiguration and self-maintenance. Accompany with these new technology, two types of innovative development are receiving more attentions by academia and industries, which are service innovation and industrial big data. In this section, previous research on these two topics will be reviewed and discussed.

2.1. Manufacturing servitization and innovation

Many advanced countries, whose economic base is manufacturing industry, made efforts to transform their economy and reinvigorate manufacturing industry because they suffer threats from emerging markets and global manufacturing supply chain. Therefore, manufacturing firms not only seek manufacturing technique innovation but also began to focus on induction and impetus of service. This way, the fuzzy boundary of manufacturing industry and service industry drive and stimulate the development of manufacturing servitization.

Servitization was proposed by Vandermerwe and Rada in 1988 [1]. They emphasized the concept of customer-focus, combining product, service, support, and knowledge is the most important element. Furthermore, the authors also asserted that not only service industries but also manufacturing industries should focus on innovative value-added service development in order to quickly enhance their core competencies. Baines defined manufacturing servitization as innovation of organizational capabilities and process, from product sales to integrated product-service [2].

Servitization is defined as the strategic innovation of an organization’s capabilities and processes to shift from selling products to selling an integrated product and service offering that delivers value in use, i.e. a Product-Service System [3]. The concept of a Product Service-System (PSS) is a special case of servitization. Mont defines PSS as a system of products, services, supporting networks, and infrastructure that is designed to be competitive, satisfy customers’ needs, and have a lower environmental impact than traditional business models [4]. In PSS business model, industries develop product with value-added service instead of single product itself, and provide their customers services that are needed. In this relationship, the market goal of manufacturers is not one-time product selling, but continuous profit from customers by total service solution, which can satisfy unmet customers’ needs.

2.2. Industrial big data environment

Recently, big data becomes a buzzword on everyone’s mouth. It has been used widely in data mining after human-generated content has been a boost to social network. It is also called web 2.0 eras since late 2004 [5]. Lots of research organizations and companies have devoted themselves to this new research topic and most of them focus on social or commercial mining, including sales prediction, user relationship mining and clustering, recommendation systems, opinion mining, etc. [6–10]. However, these research focus on ‘human-generated or human-related data’ instead of ‘machine-generated data or industrial data’, which may include machine controllers, sensors, manufacturing systems, etc.

Under the above-mentioned Industry 4.0 era, intelligent analytics and cyber physical systems are teaming together to realize a new thinking of production management and factory
transformation. Using appropriate sensor installations, various signals such as vibration, pressure, etc. can be extracted. In addition, historical data can be harvested for further data mining. Communication protocols, such as MTConnect [11] and OPC, can help users to record controller signals. When all the data are aggregated, this amalgamation is called “Big Data”. The transforming agent consists of several components: an integrated platform, predictive analytics and visualization tools. The deployment platform is chosen based on: speed of computation, investment cost, ease of deployment and update, etc. [12]. The actual processing of big data into useful information is then the key of sustainable innovation within an Industry 4.0 factory.

3. Self-aware and self-maintenance machines for industrial big data environment

The recent developments of an Internet of Things (IOT) framework and the emergence of sensing technology have created a unified information grid that tightly connects systems and humans together, which further populates a big data environment in industry. With more advanced analytics, the advent of cloud computing, and a Cyber Physical Systems (CPS) framework, future industry will be able to achieve a fleet wide information system that helps machines to be self-aware and actively prevent potential performance issues. A self-aware and self-maintenance machine system is defined as a system that can self-assess its own health and degradation and further use similar information from other peers for smart maintenance decisions to avoid potential issues. Smart analytics for achieving such intelligence will be used at the individual machine and also at the fleet level.

For a mechanical system, self-awareness means being able to assess current or past condition of a machine and react to the assessment output. Such health assessment can be performed by using a data-driven algorithm to analyze data/information collected from the given machine and its ambient environment. Real time machine condition can be feedback to machine controller for adaptive control and machine managers for in-time maintenance. However, for most industrial applications, especially for a fleet of machines, self-awareness of machines is still far from being realized. Current diagnosis or prognosis algorithms are usually for a specific machine or application and are not adaptive or flexible enough to handle more complicated information. The reasons for why a self-aware machine has not been fully realized are summarized as follows:

**Lack of a closely coupled human-machine interaction**: a major influential factor for machine condition and performance is the human operation and management. Productivity and production quality can be greatly affected by task design and scheduling. Current machines can only passively listen to operators command and react, even when the assigned task is not optimal for its current condition. A smarter machine system on the other hand, should be able to actively suggest task arrangement and adjust operational parameters to maximize productivity and product quality.

**Lack of adaptive learning and full utilization of available information**: PHM systems cannot be widely implemented in industry because of its low level of adaptability, which eventually leads to a lack of robustness in the health monitoring algorithms. The problem behind such issue is that for a PHM system, development and implementation are usually separated. PHM algorithm is developed by data collected from experiments and does not change during implementation unless being re-trained by experts. In most cases the algorithm only handles condition monitoring data from real machines using a pre-defined procedure without attempting to learn from it. Such situation is far from optimal because real-time data collected from machines in the field is usually from more machine units and of a much longer time duration, which means it contains much more information than the lab generated data. Algorithms that are capable of learning from such data will be able to achieve optimal flexibility and robustness on handling different situations.

In order to solve the aforementioned research gaps, a unified Cyber Physical System framework for self-aware and self-maintenance machines has been developed that can extract meaningful information from big data more efficiently and further perform more intelligent decision-making. The proposed system framework is shown in Figure 1.

Within the scope of this research, physical space is considered to include:

- A fleet of machines, including
  - Condition Monitoring (CM) data collected now and previously
  - Controller parameters
  - Digitized machine performance (e.g. product quality measurement)
  - Machine and component configuration, model information
  - Utilization history, tasks being performed
- Human actions, including
  - Maintenance activities
  - Human controlled operating parameters and usage pattern

While in the cyber (computational) space, firstly, the data and information format needs to be properly defined so that information collected from the physical space can be recorded and managed. Secondly, the cyber space is designed to be
able to summarize and accumulate knowledge on machine degradation, so that such knowledge can be used for health assessment of new machines. Lastly health assessment results should be feedback in time to the physical space so that proper reaction can be taken.

3.1. Machine health awareness analytics with self-learning knowledge base

Unlike most of the existing CPS, which are control or simulation oriented, the proposed CPS uses a knowledge base and related algorithms to represent machine degradation and performance behaviour in the physical world. Machine health awareness analytics is designed to fulfill such task. Using adaptive learning and data mining algorithms, a knowledge base representing machine performance and degradation mechanism can be automatically populated. The knowledge base will be able to grow with new data to eventually enhance its fidelity and capability of representing complex working conditions that happen to real-world machine. With data samples and associated information collected from machines, both horizontal (machine to machine) and vertical (time to time) comparison will be performed using specifically designed algorithms for knowledge extraction. Whenever health information of a particular machine is required, the knowledge base will provide necessary information for health assessment and prediction algorithms. Because of the comprehensiveness of the knowledge base, PHM algorithms can be more flexible on handling unprecedented events and more accurate on PHM result generation.

![Figure 2: Adaptive learning for machine clustering](image)

The adaptive learning and knowledge extraction is further explained in Figure 2. Considering a machine fleet, similarity always exists among machines – machines that are performing similar tasks or at similar service time may have similar performance and health condition. Based on such similarity, machine clusters can be built, as a knowledge base representing different machine performance and working condition.

Algorithm wise, unsupervised learning algorithms such as Self-Organizing Map (SOM) and Gaussian Mixture Model (GMM) can be used for autonomously creating clusters for different working regime and machine conditions. The adaptive clustering methodology Figure 2 utilizes an on-line update mechanism: the algorithm compares the latest input to the existing cluster and tries to identify one cluster that is most similar to the input sample using multidimensional distance measurement. Search of similar cluster can end with two results: 1) similar cluster found. If it is this case, then the machine from which the sample has been collected will be labelled as having the health condition defined by the identified cluster. Meanwhile depending on deviation between existing cluster and the latest sample, the algorithm will update the existing cluster using new information from the latest sample. 2) No similar cluster found. In this case, the algorithm will hold its operation with the current sample until it sees enough count of out-of-cluster samples. When number of out-of-cluster samples exceeds a certain amount, it means that there exists a new behavior of the machine that has not been modeled so that the algorithm will automatically create a new cluster to represent such new behavior. In such case the clustering algorithm can be very adaptive to new conditions. Moreover the self-grow cluster will be used as the knowledge base for health assessment in the proposed cyber space. With such mechanism, different machine performance behavior can be accumulated in the knowledge base and utilized for future health assessment.

3.2. Decision support analytics for self-maintenance

The main objective of design, control and decision making of machine operations is to meet the production goal with effective and efficient production planning and maintenance scheduling. The actual system performance often deviates from the designed productivity target because of low operational efficiency, mainly due to significant downtime and frequent machine failures. In order to improve the system performance, two key factors need to be considered: (1) the mitigation of production uncertainties to reduce unscheduled downtime and increase operational efficiency and (2) the efficient utilization of the finite resources on the throughput-critical sections of the system by detecting its bottleneck components. With the advent of PHM development in a CPS framework, rich PHM knowledge is utilized to assist and enhance the capability of decision making in production control and maintenance scheduling to achieve high reliability and availability.

3.3. Advantages of the CPS framework

The key innovation of such CPS framework is that it realizes a self-aware and self-maintenance system by integrating both sensor data as well as fleet wide information so the data volume can be reduced and similar pattern can be identified. Such strategy further ensures that information
hidden under the industry Big Data can be properly utilized. The key advantages of the designed framework can be summarized into the following perspectives:

1. Unified Cyber Physical System frameworks for machine-to-machine health modeling: the proposed CPS is not a CPS for one machine but for a fleet of machines and human operators. The system enables machines to gather information from its peer, human operator and other surrounding environment so that machines can achieve self-awareness of its health condition via comparing with and learning from the past history of other peers.

2. Enable self-aware and self-maintenance intelligence using self-learning PHM algorithms: rigidity and inability of handling unprecedented events are major hurdles that prevent current PHM algorithms from being widely implemented in industry. This paper proposes a solution of developing adaptive capability for anomaly detection, health assessment and degradation prediction. Adaptive algorithms also enable the system to learn from in-field data and accumulate in-field knowledge that can hardly be gained in a lab test environment.

3. Smart decision support system for proactive maintenance scheduling: with connected machines and awareness of machine condition across the fleet, tasks and maintenance plans will be scheduled and optimized from the fleet level. By balancing and compensating working load and stress for each machine according to their individual health condition, production and machine performance can be maximized.

4. Case study: Mechanical system components with self-aware intelligence

The increasing feeding rate and high acceleration value have improved the machine tools efficiency and accuracy and also reduce the machining times. The each component of the machine tools can directly affect the overall precision of the machine tools system. The critical components can be defined based on their criticality level, low failure frequency but high impact. The condition-based monitoring can enable the manufacturing factories to observe the degradation of each critical component and calculate their the health value, therefore the overall health condition of the machine will be evaluated based on the health condition of each component. This health indicator like health value, confidence value can provide and predict the information of the machine’s health condition. So the predictive analytical algorithms provide significant improvements of adding PHM to traditional maintenance schemes. Machine data is effectively handled method, the physical model and experiment knowledge are also applied to fully understand the performance of these components, like the FEA model, mode analysis, etc. In the cyber space, all of these models are integrated to convert the data to information, with which the customers can schedule the timely maintenance before the failure.

In order to model the components’ behavior quickly and effectively, the accelerated degradation test in the lab is utilized in the machine reliability analysis, the different stress factors are screened and the important stress factors are picked according to the quantization matrix. After that, the characteristics tests are conducted to identify the initial condition. Thus the data collected from multiple degradation process under different stress levels can generate the degradation pattern for the component, thus the time conversion between these different degradation pattern will be validated, which can assist to identify the degradation trend in the actual industrial application. The case is shown in Figure 3. After establishing the degradation model, this PHM model will be integrated in the cyber space. The data collected from the machine components will be processed through this PHM model and thus a virtual component is modeled in the cyber space, so it can simulate and predict the health pattern of the machine component in the reality.
5. Conclusion

Industry 4.0 proposes the predictive manufacturing in the future industry. The machines are connected as a collaborative community. Such evolution requires the utilization of advance prediction tools so that data can be systematically processed into information that can explain the uncertainties and thereby make more “informed” decisions.

In this paper, trends and unmet needs accompany with the upcoming Industry 4.0 era has been presented in this article. It includes manufacturing servitization, which changes manufacturers’ value proposition, and industrial big data, which makes manufacturing analytics more important than past decades. To sustain under these trends, a systematic framework is proposed for self-aware and self-maintain machines. The framework includes the concepts of cyber-physical system and decision support system. Lastly, a case study is presented in order to demonstrate the feasibility of the proposed work.

In sum, the prognostics monitoring system is trend of the smart manufacturing and industrial big data environment. There are many areas that are foreseen to have an impact with the advent of the fourth industrial revolution. Of which four key impact areas emerge:

- Machine health prediction reduces the machine downtime, and the prognostics information will support the ERP system to optimize manufacturing management, maintenance scheduling, and guarantee machine safety.
- The information flow among the production line, business management level and supply chain management make the industrial management more transparent, organized.
- The new trend of industry will reduce the labor cost, and provide better working environment.
- Eventually it will reduce the cost by energy saving, optimized maintenance scheduling and supply chain management.

References