Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 72 (2018) 1362–1367

www.elsevier.com/locate/procedia

51st CIRP Conference on Manufacturing Systems

A Big Data based Cost Prediction Method for Remanufacturing End-of-Life Products

Zhouyang Ding[a], Zhigang Jiang[b]*, Ying Liu[c], Yan Wang[d], Congbo Li[e]

[a]Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science & Technology, Wuhan 430081, China
[b]Key Laboratory of Metallurgical Equipment and Control Technology, Wuhan University of Science & Technology, Wuhan 430081, China
[c] Institute of mechanical and manufacturing engineering,School of Engineering, Cardiff University, Cardiff CF24 3AA
[d]Department of Computing, Engineering and Mathematics, University of Brighton, Brighton, BN2 4GJ, United Kingdom
[e] State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing, 400030, China
* Corresponding author. Tel.: +86 27 68862500; fax: +86 27 68862500. *E-mail address:* jiangzhigang@wust.edu.cn (Z.G. Jiang)

Abstract

Remanufacturing is considered as an important industrial process to restore the performance and function of End-of-Life (EOL) products to a like-new state. In order to help enterprises effectively and precisely predict the cost of remanufacturing processes, a remanufacturing cost prediction model based on big data is developed. In this paper, a cost analysis framework is established by applying big data technologies to interpret the obtained data, identify the intricate relationship of obtained sensor data and its corresponding remanufacturing processes and associated costs. Then big data mining and particle swarm optimization Back Propagation (BP) neural network algorithm are utilized to implement the cost prediction. The application of presented model is verified by a case study, and the results demonstrates that the developed model can predict the cost of the remanufacturing accurately allowing early decision making for remanufacturability of the EOL products.

Peer-review under responsibility of the scientific committee of the 51st CIRP Conference on Manufacturing Systems.

Keywords: Remanufacturing; Cost Prediction; End-of-Life Products; Big Data; BP Neural Network

1. Introduction

As a major national strategy, remanufacturing plays a significant role in current manufacturing industry, it is considered as a significant value recovery approach and manufacturing method that can generate End-of-Life (EOL) products as good as new ones [1,2] and potentially achieve considerable economic, environmental and social benefits in many applications [3]. In order to help enterprises implement remanufacturing better, advance decision-making has to be executed. It can be seen In previous studies that remanufacturing cost is an important reference indicator for evaluating remanufacturability of EOL products [4]. Due to its objectivity and accuracy, the cost prediction will be the key part and prerequisite conditions for remanufacturing, and more beneficial to the remanufacturing decision [5].

Many scholars have studied the cost prediction of remanufacturing processes, the methodologies are mainly focused on cost driver, linear regression, artificial neural network and so on. Qin et al. [6] proposed a neural network based cost estimation model to estimate the remanufacturing cost of engineering machinery hydraulic cylinder. Furthermore, Liu et al. [7] and Xiang et al. [8] developed cost prediction models for remanufacturing of mechanical and electrical products based on linear regressive and support vector machine respectively. Wang et al. [9] presented an integrated approach that considers various key cost factors in the remanufacturing operations to assess the economic effects of component reuse. Regrettably, these studies only have established prediction models or methods of remanufacturing cost for specific objects and merely consider a certain characteristic or relationship that affects cost.

To be brief, current difficulties of cost prediction can be summed up in three major issues: 1) difference in the failure types and failure degree of the EOL components, and the diversification of repair schemes; 2) remanufacturing cost is a

2212-8271 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 51st CIRP Conference on Manufacturing Systems.
10.1016/j.procir.2018.03.129

dynamic random variable and contains many known or unknown influencing factors, and its research involves the whole life cycle of EOL products; 3) in the era of big data, short timeliness, widely dispersion, massive data, diversified structural are important features of remanufacturing data. These features make it impossible to obtain accurate results using traditional prediction models or methods.

In the area of remanufacturing, big data refers to a large amount of data (eg, historical work maintenance data of the recycling parts etc.) that is produced during the manufacturing, maintenance, and remanufacturing phases [10]. Big data technology is the process of data integration, processing, analysis and interpretation. Due to its high value and great potentiality, the increasing big data has been widely used in medical treatment, e-commerce and other industries. Predictions based on big data have also been successfully applied in many fields. Based on big data processing and data mining technology, Liu et al. established a distribution cost prediction model for improving benefit of power grid enterprises [11]. Hammer et al. successfully implemented a profit per hour an agile control approach by using big data and advanced algorithms, and brought huge profits [12]. Ye proposed a big data-based economic assessment system to help enterprises conduct economic assessments [3]. All of these successful cases show that it is feasible to develop a cost prediction model by using big data.

With the development of detection technology and the extensive use of sensors, the remanufacturing industry also has accumulated a large amount of data, which can be used for further decision-making by enterprises. On the basis of the previous researches, big data technologies has been employed to establish a cost prediction framework in this paper.

In short, the main novelty of this paper is that, presenting a big data analysis framework of remanufacturing process cost, revealing the change rules of remanufacturing cost, developing a cost prediction model of remanufacturing process based on big data, and proposing a estimation method remanufacturing cost based on big data and optimized BP neural network.

2. Analysis framework of remanufacturing cost based on big data

2.1. Cost structure of EOL products for remanufacturing

The remanufacturing process of EOL products mainly includes disassembly, cleaning, inspection, recondition, reassembly and testing [13]. In the complete remanufacturing system, the corresponding costs are generated in each stage.

According to the above analysis, the total remanufacturing cost for EOL products consists of the expenses of its parts remanufacturing and the expenses incurred by the whole machine in other processes such as recovery, disassembly, cleaning, testing and transportation, shown in Fig.1. The cost structure can be broadly classified as follows.

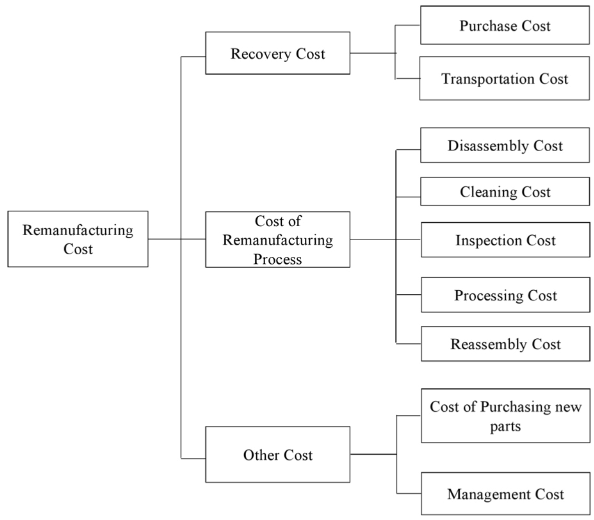

Fig. 1. The structure of remanufacturing cost for EOL products

(1) *Recovery Cost:* recovery costs includes cost of purchasing EOL products from customers and the transportation cost incurred in the course of products are shipped back to remanufacturing plant.

$$C_{\mathrm{r}} = C_{\mathrm{p}} + (l \cdot T_w)/n \tag{1}$$

where C_r is the unit recovery cost (£), l is the shipment distance (km) , T_{w} is transportation costs per unit distance (£/km) and is the number of products.

(2) *Cost of Remanufacturing Process:* a series of costs incurred in the process of remanufacturing, such as disassembly cost, cleaning cost, inspection cost, processing cost and so on.

$$C_p = C_l + C_m + C_t + C_e \tag{2}$$

where C_p is the total cost of processing (£), C_l is the total labor cost of each process (£), C_m , C_t and C_e respectively represent the materials cost (£), tool cost (£) and energy consumption cost (£).

(3) *Other Costs:* this part is mainly composed of the costs of purchasing new parts and other management expenses such as inventory costs and equipment maintenance costs.

$$C_O = C_a + C_n = C_a + m \cdot f \tag{3}$$

where C_o is other indirect costs (£), C_a is the management expenses (£), C_n represents the cost of purchasing new parts (£), m and f respectively represent quantity and unit cost of new parts.

The management cost is charged at $x\%$ of the other total cost. In summary, the total cost of remanufacturing is the sum of the above costs:

$$C_{total} = (C_r + C_p + C_n) \cdot (1 + x\%) \tag{4}$$

2.2. Influential factors of remanufacturing cost for EOL products

There are many influential factors that affect the cost of remanufacturing, including the characteristic parameters of EOL products themselves and market environment factors, they constitute the main cost drivers in the remanufacturing process. In this study, the key influential factors have been mainly identified by applying big data technology to analyze the remanufacturing cost data obtained, combining with reviewed literature and consulting field experts.

The key influential factors of remanufacturing process are listed in Table 1. L represents this factor as a qualitative indicator, while N is a quantitative indicator. Rating of the impact for remanufacturing cost are employed to illustrate the impact extent of this factor on total cost. The rating of the impact for remanufacturing cost is the quantization value obtained according to results of Spearman Rank Correlation Coefficient (equation (5)) calculation combined with the expert evaluation, the greater the numeric value, the greater the influence of this factor on the total cost of remanufacturing. There are also other methods to evaluate influencing factors, such as Pearson R, Refief F-algorithms and Stepwise Regression Approaches, but it chooses Spearman Rank Correlation Coefficient because it can ignore the overall distribution of variables, subgroup size, and solve the "multi-collinearity" problem.

$$R_s = \frac{\sum_{t=1}^{n}(R_t - \overline{R})(Q_t - \overline{Q})}{\sqrt{\sum_{t=1}^{n}(R_t - \overline{R})^2}\sqrt{\sum_{t=1}^{n}(Q_t - \overline{Q})^2}} \tag{5}$$

where pairs observations of the two variables x, y were ranked from small to large, ranks are denoted by R_t and Q_t respectively.

Table 1 The influential factors of the remanufacturing cost

Key Influential factors	Qualitative or Quantitative	Rating of the impact for remanufacturing cost
Quality of products procured	L	6
Price paid to the customer for the old products	N	7
Transportation cost	N	4
Ratio of parts remanufactured	N	9
Failure types	L	8
Complexity of repairing technology	L	9
Processing time for remanufacturing	N	7
Direct labor cost	N	6
Replacement rate of parts	N	5
Cost of purchasing new parts	N	9

2.3. Change rules of remanufacturing cost for EOL products

This paper mainly focuses on the data provided by the remanufacturing workshop of a machine tool factory. The analysis of change rules is the basis of choosing prediction method. By using big data technology to analyze the obtained data, it can find that the change rules of remanufacturing cost of EOL products are as follows:

(1) Remanufacturing process cost has greater randomness and uncertainty. Since the EOL products themselves have different failure types, quality status and failure levels, it is impossible for any EOL product to be remanufactured through the same process of inspection, processing and reassembly. This has led to the diversification of resources, energy, etc. used in the remanufacturing process, which ultimately leads to uncertainty in the total cost.

(2) There is a nonlinear relationship between the total cost and its influencing factors. This determines that it is impossible to use traditional prediction methods such as linear regression and data statistics to predict the cost of remanufacturing process.

(3) Some influence factors play a decisive role in the change of total cost. Some factors, such as cost of purchasing new parts and ratio of parts remanufactured, have a great impact on total cost, and their fluctuation will cause great changes in cost prediction.

All rules show that the traditional prediction model based on small range and small amount of data is ineffective and not conducive to business decision-making. Neural network is a computational method that simulates human brain big data analysis mechanism and is the most successful method in big data analysis, especially for non-linear and irregular data. Therefore, considering the characteristics of big data and the change rules of remanufacturing cost, the combination of the existing big data technology and the neural network-based prediction algorithm will provide a true and effective result for the cost prediction of enterprises remanufacturing.

3. Cost Prediction Model of Remanufacturing Process Based on Big Data

3.1. Cost prediction model

3.1.1 Technology architecture of prediction model based on Hadoop

Hadoop big data platform can efficiently process large scale data by connecting multiple network distributed servers, and it can make full use of existing common resources. Two key technologies of Hadoop are MapRduce and HDFS, HDFS provides storage of large amounts of data, and MapRduce provides the calculation of these data [14, 15]. The proposed architecture of big data-based remanufacturing cost prediction system is shown in Fig. 2, it consists of four levels.

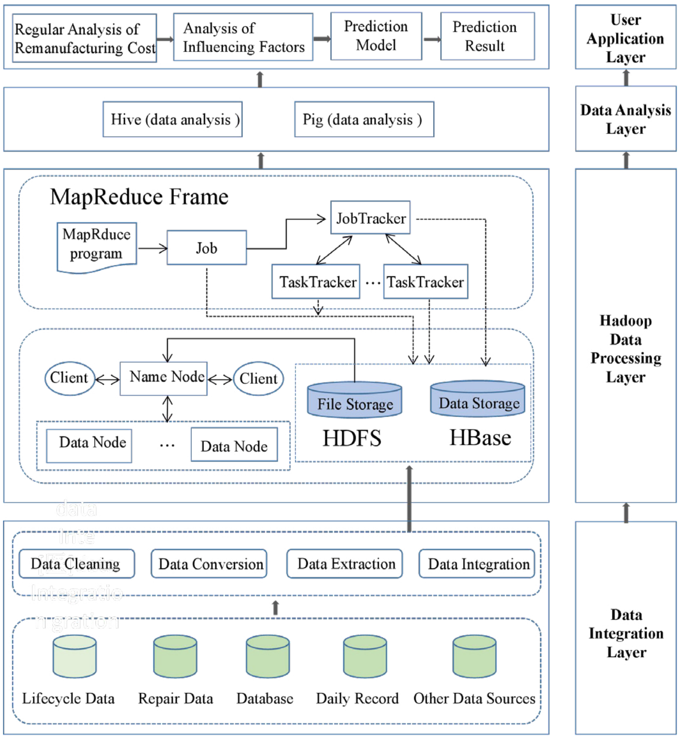

Fig. 2. Big data processing architecture of remanufacturing cost prediction model based on Hadoop

(1) Data integration layer

The data integration layer is at the bottom of the whole architecture, and it is the data source that the system needs to handle. In this system, the first step is a lot of data is imported into the entire data processing platform. These data have a variety of structures, types of changeable features, some can be directly stored in HDFS, and some can be directly processed by the MapRduce program. All these data are imported into Hadoop big data processing module, and distributed storage and computation of data are implemented according to different applications [16].

(2) Hadoop data processing layer

For this layer, it mainly includes file storage, data storage and programming model. In the overall architecture, the file storage layer links down with the data source and data integration layer to access specific storage resources, and provides the file access service upward for the programming model and data storage layer. The function of data storage section is to provide the storage and management ability of a large number of distributed and extensible data tables. Currently, in the field of big data processing based on cloud computing, the MapRduce model occupies a dominant position, and it is also the core of Hadoop technology [17].

(3) Data analysis layer

Data analysis plays an important role in the overall architecture. It provides data analysts with advanced analytical tools to improve their productivity. For example, Pig provides a higher level of data processing power and Hive converts structured data into a table for inquiries.

(4) User application layer

After being integrated, processed, analyzed and excavated, the collected data will enter the business application layer. At this level, users can realize the direct application of data. The cost prediction model can be established by analyzing change rule and influencing factors of remanufactring cost, and using the specific algorithm to predict the cost.

3.1.2 MapReduce programming model

MapReduce framework is the core of Hadoop technology, which was first proposed by Google and is widely used computing model for distributed processing of massive data [].

In brief, MapReduce is divided into two stages. The mapper function is responsible for filtering, transforming and breaking up the raw data in Map stage. Then obtained data is input to the Reduce stage and processed by the reduceer function to get the final result. Its process shown in Fig.3. Mapreduce takes the form of key-value pairs as input and output throughout the running process. MapReduce can take full advantage of idle devices which are widely distributed and have computing-function in the Internet of Things (IoT).

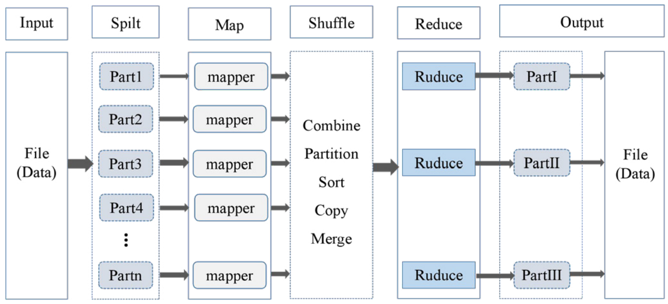

Fig.3. Process flow chart of MapReduce

3.1.3 Realized schemes of prediction model

A prediction scheme of remanufacturing cost is established based on big data technology. By applying a series of data in the whole life cycle of the EOL products to the big data analysis framework, the result can be accurately predicted. The specific process is shown in Fig. 4.

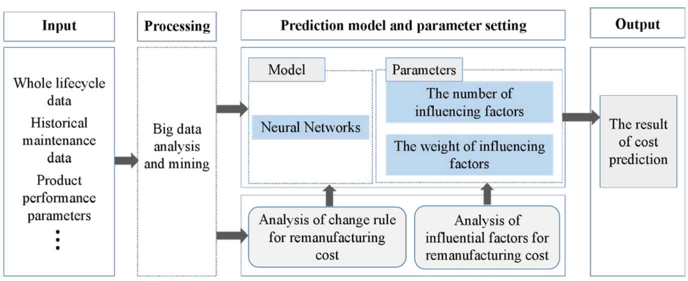

Fig.4. Flow chart of remanufacturing cost prediction

BP neural network is employed to predict remanufacturing cost and particle swarm is utilized to optimize it. The influencing factor analysis is used to select and optimize the parameters of prediction model. Several parameters mentioned in the second part have an impact on the result of the cost prediction, and many factors should be considered in the selection of parameters.

3.2. BP neural network

3.2.1 Modeling process of BP neural network and PSO algorithm

Neural network is composed of a large number of neurons nonlinear system, the function and structure of each neuron are relatively simple, but the neural network system composed of a great number of neural networks is very complex. BP neural network is a one-way propagation of multi-layer network, there are three layers, input layer, hidden layer and output layer [18], it is divided into forward and backward propagation. The weight of each layer can be adjusted through the forward propagation and back propagation error, the weight adjustment process is the BP neural network learning and training process. The process of reducing the output error is a cycle of reciprocation until it reaches the termination condition.

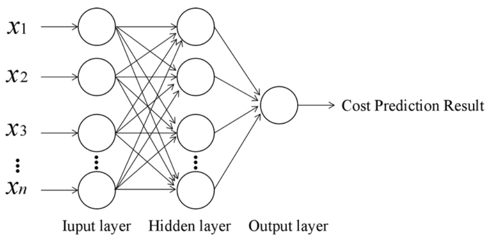

Fig.5. Neural network prediction model

There are several nodes in the input layer, which are the normalized and quantized data of the some influencing factors. The output layer has 1 node, which represents the cost prediction result. The number of hidden layer nodes is determined by the empirical formula.

Despite BP neural network has good ability of approximating non-linear mapping, simple BP neural network has slow convergence speed and local extremum. It is difficult to determine the number of hidden layers and hidden nodes. To this end, we need to improve BP algorithm. Many algorithms such as genetic algorithm (GA) and simulated annealing algorithm (SAA) can optimize neural network, but these algorithms are prone to problems such as large fitting error, poor stability and slow speed. So in this paper, particle swarm optimization (PSO) algorithm is introduced to optimize the initial weights and thresholds of BP to speed up the convergence of BP neural network.

PSO algorithm is a stochastic global optimization technology based on swarm intelligence theory, it searches the solution space intelligently and thus finds the optimal solution through the interaction between particles. The idea of PSO algorithm is that individuals are regarded as particles in multi-dimensional space. Each particle has an initial velocity, and modifies its own state based on its own flight experience and group experience.

3.2.2 PSO neural network algorithm

Particle swarm optimization and neural network are different optimization strategies, there are many differences in the application of problems and the way of information processing. In this paper, the global search ability of particle swarm optimization is adapted to optimize the weights and thresholds of neural networks to improve the learning ability and forecast efficiency of neural networks. Algorithm flow shown in Fig. 6.

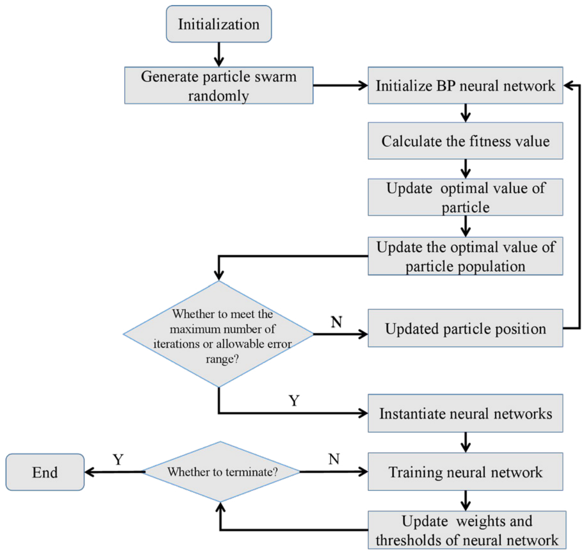

Fig. 6. Flow chart of particle swarm optimization neural network algorithm

4. Case Study

This section will show the practical application of the proposed cost prediction model through a practical case. The MapRduce program is written and distributed to four nodes, and the big data analysis framework shown in Fig.2 is applied to the remanufacturing workshop of a machine tool factory for data collection, analysis and mining. The results are utilized to the cost prediction model.

Taking remanufacturing implementation for C6132A1 EOL machine tool of a machine tool factory as an example, based on big data analysis and considering the change rules of cost, according to machine tool remanufacturing process, 6 factors affecting the remanufacturing cost have been identified and quantified. They are ratio of parts remanufactured, failure types, complexity of repairing technology, processing time for remanufacturing, direct labor cost and cost of purchasing new parts.

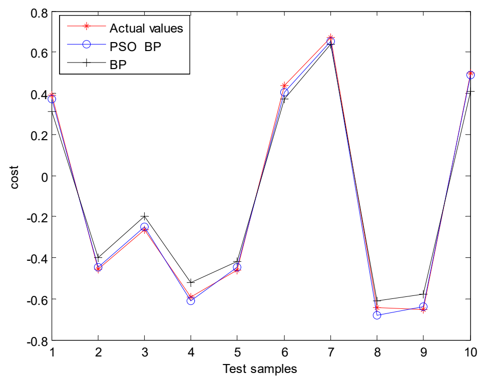

Fig. 7. PSO neural network compared with general BP neural network and its actual value

These six factors together with the total cost are the inputs and outputs of the neural network model. Considering the accuracy of the results, the number of hidden layer nodes is calculated as 4, 5, 6, and 7 respectively, and the number of nodes finally selected is 4.

By comparing the actual value, the prediction result of neural network and the prediction result of PSO neural network (as shown in Fig . 7), it can be seen that the accuracy of the PSO neural network for predicting the remanufacturing cost is higher than that of the unoptimized neural network.

A BP neural network model based on particle swarm optimization is used to predict the remanufacturing cost of a EOL machine tool. The prediction results and errors are shown in Table 2.

Table 2 Predict results and errors

Serial Number	Predictive Value	Actual Value	Error	Relative Error
NO.1	0.37066	0.39207	0.02141	5.78%
NO.2	-0.44815	-0.45710	-0.00895	2.00%
NO.3	-0.24920	-0.26452	-0.01532	6.15%
NO.4	-0.61186	-0.59080	0.02106	-3.44%
NO.5	-0.44682	-0.45930	-0.01248	2.79%
NO.6	0.40339	0.43860	0.03521	8.73%
NO.7	0.65032	0.67352	0.02320	3.57%
NO.8	-0.68244	-0.64502	0.03742	-5.48%
NO.9	-0.63645	-0.65487	-0.01842	2.89%
NO.10	0.48820	0.49910	0.01090	2.23%

It can be seen from Table 2 that the relative errors of the predictions are no more than 10%, and the prediction results are shown in Table 3.

Table 3 Comparison of prediction results

Model	Average relative error	Training steps	Generalization ability
Basic BP neural network	4.02%	187	weak
PSO neural network	2.52%	6	strong

In light of the analysis above, PSO neural network algorithm for remanufacturing cost prediction model is superior to other general neural network model, mainly because of its higher prediction accuracy, stronger generalization ability, faster convergence speed and shorter training time. This will provide more reliable support for the implementation of remanufacturing.

5. Conclusion

In sum, the research discusses the accuracy and legitimacy of the proposed framework of big data analysis and the practicability of the presented prediction model through a case study. The method proposed in this paper is suitable for the cost prediction of remanufactured enterprises with a large amount of data accumulation. It will greatly improve the prediction accuracy and provide strong support for enterprises in making remanufacturing decisions. Future research works should focus on the construction of integrated prediction software system combined with intelligent technology and the development of dynamic data mining algorithm in order to realize the rapid, accurate and convenient prediction of cost.

Acknowledgements

The work described in this paper was supported by the National Natural Science Foundation of China (51675388). These financial contributions are gratefully acknowledged.

References

[1] Zhang Y, Ren S, Liu Y, et al. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. Journal of Cleaner Production, 2016.
[2] Sabharwal S, Garg S. Determining cost effectiveness index of remanufacturing: A graph theoretic approach. International Journal of Production Economics, 2013, 144(2):521-532.
[3] Ye Bowen. The research of application big Data analytics to the parts remanufacturing economic evaluation. Huazhong University of Science and Technology, 2014.
[4] Xu Y, Feng W. Develop a cost model to evaluate the economic benefit of remanufacturing based on specific technique. Journal of Remanufacturing, 2014, 4(1):4.
[5] Ghazalli Z, Murata A. Development of an AHP CBR evaluation system for remanufacturing: end-of-life selection strategy. International Journal of Sustainable Engineering, 2011, 4(1):2-15.
[6] Qin P P, Yun S, Shu J W, et al. Estimation of the remanufacturing cost of engineering machinery hydraulic cylinder// 3d International Conference on Advanced Information and Communication Technology for Education. 2015.
[7] LIU Zhi-feng BI, Dao-kun LIU Tao et al. Estimating for remanufacturing cost of mechanicaI and electricaI products based on linear regression. Machinery Design＆Manufacture, 2011(9):258-260.
[8] XIANG Hong, XIA Xu-hui, WANG Lei, JIANG Zhi-gang et al. A Prediction Method for Remanufacturing Cost of Mechanical and Electrical Products Based on Support Vector Machine. Machinery Design & Manufacture, 2016(4):268-272.
[9] Wang W, Mo D Y, Wang Y, et al. Assessing the cost structure of component reuse in a product family for remanufacturing. Journal of Intelligent Manufacturing, 2016:1-13.
[10] Zhang Y, Ren S, Liu Y, et al. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products[J]. Journal of Cleaner Production, 2016.
[11] Liu J, Cai Y, Xing L. Research and Application of Grid Cost Structure and Synthetical Benefit Analysis Based on Big Data. Electrical Automation, 2017.
[12] Hammer M, Somers K, Karre H, et al. Profit Per Hour as a Target Process Control Parameter for Manufacturing Systems Enabled by Big Data Analytics and Industry 4.0 Infrastructure ☆[J]. Procedia Cirp, 2017, 63:715-720.
[13] Parker S D. Remanufacturing: The ultimate form of recycling. 1997.
[14] Jun Liu. Hadoop big data processing. Beijing: People Post Press, 2013.
[15] Yan Lou. Big Data Technology Introduction. Beijing: Tsinghua University Press, 2017.
[16] ZHANG Suxiang, LIU Jianming, ZHAO Bingzhen, CAO Jinpin et al. Cloud Computing-Based Analysis on Residential Electricity Consumption Behavior. Power System Technology, 2013, 37(6):1542-1546.
[17] Makun Long. Short Term Distributed Load Forecasting Method Based on Big Data. Hunan University, 2014.
[18] Changguo Lv. Research on Grid Resource Scheduling Based On BP Algorithm. Harbin University of Science and Technology, 2007.