Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management
Bali, Indonesia, January 7 – 9, 2014

Data Mining on Return Items in a Reverse Supply Chain

Lida Mohammadi
Department of Industrial Engineering
Mazandaran University of Science and Technology
Babol, Iran

Hamed Fazlollahtabar
Faculty of Industrial Engineering
Iran University of Science and Technology
Tehran, Iran

Iraj Mahdavi1 and Ali Tajdin
Department of Industrial Engineering
Mazandaran University of Science and Technology
Babol, Iran

Abstract

In the reverse supply chain the basic concern is to categorize the returned materials in the usable or non-usable ones for performing the required operations to resend them in the supply chain. Companies can obtain more confidence from their consumers to carry out continuous revers logistics removing the defective products. To begin the reverse multi-layer multi-product supply chain and after collecting the returning commodities, we cluster them via k-mean algorithm in Matlab software environment. The results are used to perform a sampling process to deliver the commodities to the related layer for rework and repair operations. The computations show an extensive reduction in the human resource costs for separation of the returning items and increasing the efficiency.

Keywords
Reverse supply chain, data mining, clustering.

1. Introduction

Implementation of reverse logistics especially in product returns would allow not only for savings in inventory carrying cost, transportation cost, and waste disposal cost due to returned products, but also for the improvement of customer loyalty and futures sales. In a broader sense, reverse logistics refers to the distribution activities involved in product returns, source reduction, conservation, recycling, substitution, reuse, disposal, refurbishment, repair and remanufacturing (Stock, 1992). In recent years data mining has become a very popular technique for extracting information from the database in different areas due to its flexibility of working on any kind of databases and also due to the surprising results (Shahbaz et al., 2010). Data mining is the process in databases to discover and to reveal previously unknown, hidden, meaningful and useful patterns (Fayyad et al., 1996; Baker, 2010). Many approaches, methods and algorithms have been developed in the field of data mining. Data mining techniques are classified as characterization and discrimination, classification, cluster analysis, association analysis, outlier analysis and evolution analysis (Han and Kamber, 2006; Chen et al., 1996). These techniques are briefly described as below. Characterization is used for summarizing the general characteristics of any dataset. However, discrimination is utilized for determining the diversities among different datasets. The products whose sales rates are over 25% for a year in a shopping center are based on the characterization technique. Whereas, comparison of the products whose sales rates increased up to 10% and the products whose sales rates decreased up to 15% is based on the discrimination technique (Dincer, 2006). Classification is used for determining the class of a new observation utilizing available classes of the observations in training set (Larose, 2005). Grouping the customers as the ones who paid in a three-day period and the ones who paid over a three-day period is based on the classification technique. Decision trees, regression analysis, artificial neural networks, support vector machines, Naïve Bayes algorithm, k-nearest neighbor algorithm and genetic algorithm are among the classification techniques (Liao and Triantaphyllou, 2007). Cluster analysis is used for clustering similar data structures in any dataset (Tan et al., 2006). Determining the real group of the musical instruments according to their sound signals is based on the clustering technique (Essid et al., 2005). Hierarchical methods, partitioning methods, density-based methods, grid-based methods and heuristic methods are among the clustering techniques. The association analysis discovers relationships among observations and determines which observations can be realized together (Chen and Weng, 2009). A priori algorithm is one of the techniques used in association analysis. Many data mining techniques detect the exceptions as a noise but the exceptions can contain more information with respect to other observations. For this reason, outlier analysis is used in the stage of analyzing the observations that differ from the data distribution model of available dataset (Hea et al., 2004). As the last technique, the main aim of evolution analysis is to reveal time-varying tendencies of the observations within the dataset (Tan et al., 2009). This paper is organized as follows. In Section 2, a general framework and problem definition for reverse supply chain are proposed. Section 3 proposes the suggested algorithm. In Section 4, numerical experiments are presented. Finally conclusions and further researches are addressed in the last section.

2. Problem definition

The reverse supply chain under study is multi-layer, multi-product. In the designed (planned) model, the returned products after collecting and inspecting divides into two groups of disassembling and not disassembling products. The products which can be taken parted to the parts will be sent to the disassembling centers and there, they will convert to the parts. There they divide into reusable and not reusable parts. The not reusable parts will rebut safely and the reusable parts will be sent to the processing center. In the remanufacturing process, according to the production center's demand, the parts which can be used again, after processing center will be sent to the remanufacturing center and after compounding with the other parts will be changed into new products and can return to the distribution chain. In the recycling process according to the recycling center's demand the disassembled parts (which can recover again) right after disassembling centers will be sent to the recycling centers for the purpose of producing the secondary materials. In this research the aim is to clustering returned items and connecting them to the layers of the reverse supply chain by using data mining. Actually data mining is recognizing the exact , novel , useful ,perceptible samples from available inputs in an input station which cannot be reachable by using usual process. In this research most of under research returned items are home garbage. After collecting the garbage to take them back to the supply chain by regarding to their efficiency, we need so many experts and also we have time and cost categories. In this study to ease the process and the less human power, we will categorize the returned items in two phases. By using this method we can examine any kind of returned item more tenuously till its efficiency and further use of that stuff be recognized truly. In the first phase we segregate these objects by their difference in their nature and then in the second phase by attention to object's quality we can put it in the correct section and then take it back to the appropriate layer of the reverse supply chain. In this research we study 4 types of returned products which has most usage in recycle and remanufacture categories, as so 1) plastic, 2) glass, 3) paper, 4) metal.

3. Proposing an Algorithm

Clustering the returned products in the first phase can be performed with the one of data mining algorithm called k-means. This algorithm by having the number of clusters, categorizes the inputs and finally specifies the centers which according to them the clusters will be categorized. To perform this algorithm we need to an specified model or scale among the inputs which the amount of that scale should be different in any kind of returned products. By performed studies it is suggested that the suitable standard for exerting this algorithm be the attraction coefficient of any kind products against X-ray glitter. To determine the inputs of this algorithm we should take all of the under control returned products against x-ray glitter and by machines which are equipped by attracted amount calculation system in any returned product we can determine the attraction coefficient of any products. we put these attraction coefficient in k-means algorithm as inputs and by specifying k-cluster we will reach to the first clustering of performing this algorithm.

For second phase clustering we need to proficiency control and detection of the clustered products at different kinds to determine each product belongs to which categories. If the product which is returned by client be reusable or was impeccable or have repaired according to its destruction can be put in distribute system again. If product needs renovation and adding new things will be transmitted to the product assemble phase and if a part have a problem, the part will be replace and remanufactured. If there is any problem in part, chip removal will perform and if there is no way to repair, the part would be sent to the first phase of chain as a secondary material, but if it is impossible to use it as a secondary material, the company have to register it as a scrap.

In this phase of clustering to avoid wasting time and cost the suggested method of sampling can be used via an operator and from any specific pallet related to the specific kind. the proficient operator in the basis of special kind will exert sampling.

According to these samples the operator will eventually determine what kind by which percent should be returned to the noteworthy layer.

We introduce the possibilities as follow:

α : The possible percent that the returnable product will be referred to the supplier chain.

β : The possible percent that the returnable product will be referred to the manufacturer layer

λ : The possible percent that the returnable product will be referred to the distributor layer

γ : The possible percent that the returnable product will be destroyed.

In the collecting process of returned products from clients and referring them after clustering by nature to any one of supplier, manufacturer, distributor, or destroy sites it may take different operations on returned products in parts or related sub parts, so we can point out one of these operation such as duplication, or output a percent of product from supply chain as waste. We can see some of these operations in reverse supply chain process in Figure 1.

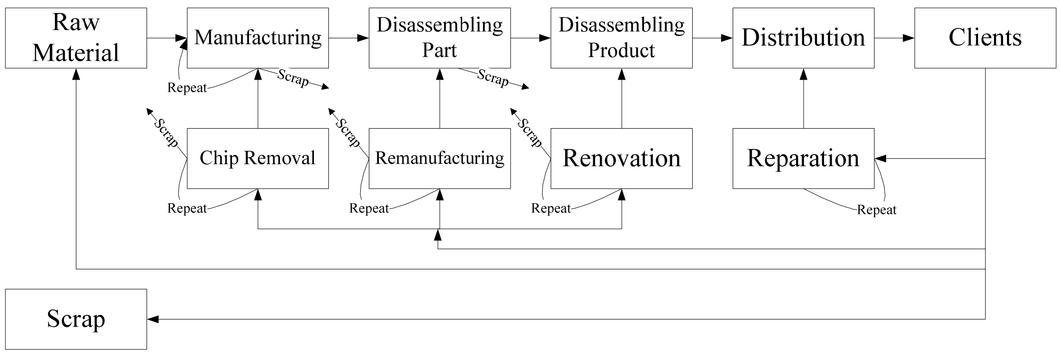

Figure 1: The recycle items illustrating duplication and output operation as waste in every part

4. Numerical experiment

For example we radiate the x-ray to the returned products collection with energy photon1 and will determine the attraction coefficient of all products with the machines which are equipped by the attraction amount calculator of this radiance by the existed products. We use these attraction coefficient as inputs and by attention to the under control kinds (plastic, glass, paper, metal) will determine the amount of clusters as follow:

K=4

i= 1 2 3 4

By performing this algorithm by MATLAB software, the inputs will be clustered and 4 cluster will be determined. This amount is shown in table 1;

Table 1: Each cluster center

C_1	0.0307	C_2	0.0320	C_3	0.0332	C_4	0.0343

By examining the given diagrams the attraction coefficient of different stuff against 1 energy photon of x-ray would be as follow:

0.03027= the attraction coefficient of plastic

0.03468= the attraction coefficient of glass

0.03074= the attraction coefficient of paper

0.03093= the attraction coefficient of metal

By assuming that the cluster centers that are obtained from performing k-means algorithm are equal to the nearest attraction obtained from remarked diagrams, we can perform clustering the first phase as follow:

C1~0.03074

C2~0.03093

C3~0.03027

C4~0.03468

So the cluster number1 demonstrate paper stuff, the cluster number 2 demonstrates metal stuff, the cluster number3 demonstrates plastic stuff and the cluster number 4 demonstrate glassy stuff.

In the second phase of clustering we use experts in any kind of stuff separately. these operators by having enough science about specific stuff pallet and by sampling them will determine that the under control returned product should be returned to which one of supplier(α), manufacturer(β) , distribution sites(γ) or with what probability should be destroyed(λ) at all. To determine these probable percent, first we should find the amount of returned products which will be sent to the anyone of these sites, so to do these we use STRATA method.

For example the operator who is the glass stuff expert will sample from the special pallet of returned glass products which has 224 stuff and determines that 22 of 45 sample should be returned to the supplier site and 17 of 45 sample to the remanufacturing site and 6 of 45 sample to the distribution site. In this sampling we don't see any destructive product. So by using STARTA method we can find the unknown amount (the number of returned glass products which entered to each site) and finally calculate the possibilities α, β, γ and λ.

In a distinct operational process , from collecting the products until sending them to the reverse supply chain and also by attention to possibilities one more duplication and waste in every part and sub part which are related to those sites, now by numerical solution we can show that how many returned products will be returned again to the supply chain.

In the Figure 2 we examine this method for returned glass products.

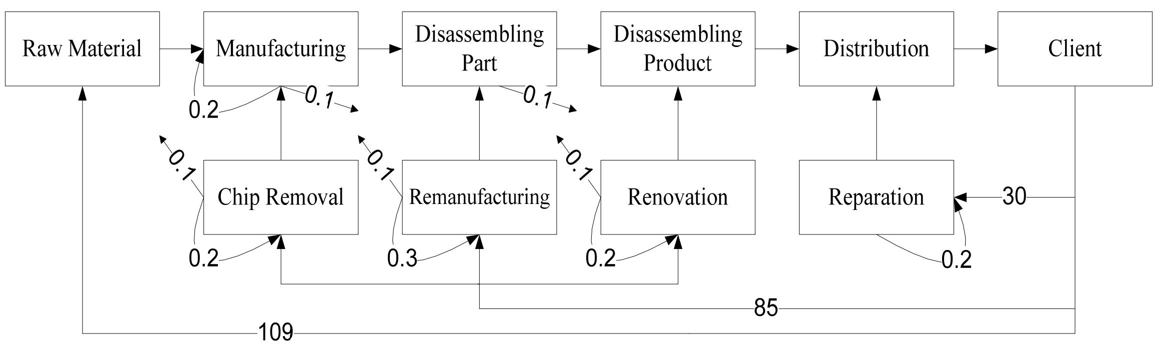

Figure 2: the recycle options with possibilities of waste and duplication.

The experts according to experiences and backgrounds first of all consider this hypothesis that from returned products which will be sent to the manufacturing site, 0.3 of them will be sent to the renovation sub part, 0.5 of them will be sent to the remanufacturing sub part and 0.2 of them will be sent to the chip removal sub part. By having these predetermined possibilities the number of products which will be sent to each one of stated sub parts will be calculated as follow:

The number of products which will be entered to the renovation sub part:

$$0.3 * 85 = 25.5 \approx 26$$

The number of products which will be entered to the remanufacturing sub part:

$$0.5 * 85 = 42.5 \approx 42$$

The number of products which will be entered to the chip removal sub part:

$$0.2 * 85 = 17$$

By assuming different phases in sequence of reverse supply chain and performed calculations: we find out that we can recycle 173 of 224 returned glass products and use it again in the supply chain.

5. Conclusions

This paper considered a reverse supply chain including multiple layers and flowing multiple products. After collecting the returned commodities a clustering of the wastes was performed to relate the materials to the corresponding layer. To optimize the clustering and reduce the inspection, material diversification is handled via X-ray. K-mean has been applied to categorize the returned commodities. Probabilities of rework activities were given and the obtained recovered products in each layer and subsequently the total reworked products were determined. Economic analysis emphasizes the effectiveness of the methodology.

References

Stock, J. K. (1992). Reverse logistics, white paper, Council of Logistics Management. IL: Oak Brook.

Muhammad Shahbaz, Syed Athar Masood, Muhammad Shaheen, Ayaz Khan. Data mining methodology in perspective of manufacturing databases. J AmSci; 2010.

Fayyad U, Shapiro GP, Smyth P. The KDD process for extracting useful knowledge from volumes of data. Communications of the ACM 1996; 39(No. 11):27e34.

Baker RSJ. Data mining. International Encyclopaedia of Education; 2010:112e8.

Han J, Kamber M. Data mining concepts and techniques. San Francisco: Elsevier Inc.; 2006. pp. 21e27.

Chen M, Han J, Yu PS. Data mining: an overview from a database perspective. IEEE Transactions on Knowledge and Data Engineering 1996; 8(No.6):866e83.

Dincer E, The k-means algorithm in data mining and an application in medicine, MSc thesis, Kocaeli Univesity, Kocaeli, 2006.

Larose DT. Discovering knowledge in data. New Jersey: Wiley & Sons Inc.; 2005. 1e25.

Liao TW, Triantaphyllou E. Recent advances in data mining of enterprise data: algorithms and applications. Singapore: World Scientific Publishing; 2007. 111e145.

Tan PN, Steinbach M, Kumar V. Introduction to data mining. Boston: Addison- Wesley; 2006. 1e37.

Essid S, Richard G, David B. Inferring efficient hierarchical taxonomies for MIR tasks: application to musical instruments. 6th International Conference on Music Information Retrieval; 2005:324e8.

Chen YL, Weng CH. Mining fuzzy association rules from questionnaire data. Knowledge-Based Systems 2009; 22(No.1):46e56.

HeaZ, Xua X, Huangb JZ, Denga S. Mining class outliers: concepts, algorithms and applications in CRM. Expert Systems with Applications 2004; 27(No.4):681e97.

Tan KC, Teoh EJ, Yua Q, Goh KC. A hybrid evolutionary algorithm for attribute selection in data mining. Expert Systems with Applications 2009; 36(No.4): 8616e30.

Biography

Lida Mohammadi has been graduated in MSc of Industrial Engineering at the Mazandaran University of Science and Technology, Babol, Iran. Her research interests are supply chain planning and decision making. She has published several research papers in international journals and conferences

Hamed Fazlollahtabar has been graduated in MSc of Industrial Engineering at Mazandaran University of Science and Technology, Babol, Iran. He received Doctorate awarded from the Gulf University of Science and Technology in Quantitative Approaches in Electronic Systems. Currently he is a PhD candidate of Industrial Engineering at Iran University of Science and Technology, Tehran, Iran. He is in the editorial board of WASET (World Academy of Science Engineering Technology) Scientific and Technical Committee on Natural and Applied Sciences, reviewer committee of International Conference on Industrial and Computer Engineering (CIE), and member of the International Institute of Informatics and Systemics (IIIS). He has become a member of Iran Elite National Foundation. His research interests are mathematical modeling, optimisation in knowledge-based systems, logistics and manufacturing systems. He has published over 150 research papers in international book chapters, journals and conferences.

Iraj Mahdavi is the Full Professor of Industrial Engineering at Mazandaran University of Science and Technology and Vice President of Research and Technology. He received his PhD from India in Production Engineering and Post-Doctorate professor from Hanyang University, Korea. He is also in the editorial board of four journals. He has published over 280 research papers. His research interests include cellular manufacturing, digital management of industrial enterprises, intelligent operation management and industrial strategy setting.

Ali Tajdin is the Assistant Professor of Industrial Engineering at Mazandaran University of Science and Technology, Babol, Iran. He has published several papers in international outlets. His research interests include network theory and optimization.