Computers & Industrial Engineering 80 (2015) 1–11

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Estimating the long-term cost to serve new customers in joint distribution

Lei Sun [a,*], Mark H. Karwan [a], Banu Gemici-Ozkan [b], Jose M. Pinto [b]

[a] Department of Industrial & Systems Engineering, University at Buffalo, Buffalo, NY 14260, United States
[b] Praxair Inc., Danbury, CT 06810, United States

ARTICLE INFO

Article history:
Received 18 August 2014
Accepted 14 November 2014
Available online 26 November 2014

Keywords:
Cost estimation
Logistics
Continuous approximation
Geographical dispersion
Regression

ABSTRACT

One of the most important concerns for logistics service providers is to identify the distribution cost to serve each new customer for pricing. Compared to the analysis through cost allocation on delivery routes, cost estimation possesses the advantage of robust costing rules but is a very challenging problem due to the complex collaborative mechanisms of distribution. Based on the activities leading to a distribution cost, we analyze the relationship between multiple geographic factors and cost, and then construct appropriate attributes for estimation. Combining a data selection approach and regression or artificial neural network techniques, a prediction scheme is proposed to build models, and an explicit continuous approximation model is suggested for efficient implementation. Computational experiments demonstrate the importance of the constructed attributes and the accuracy of the proposed cost estimation method. The impacts from cost stability and delivery frequency are examined to provide further explanation and support for practical implementation.

1. Introduction

Logistics costs comprise a significant proportion of business costs, often exceeding 10% of company turnover (Engblom, Solakivi, Tyli, & Ojala, 2012). According to a recent report from the Council of Supply Chain Management Professionals (2012), total logistics costs in the United States account for 8.5% of the nations gross domestic product, among which approximately 63% comes from distribution activities. Typically, distribution services are provided by either a third-party logistics service provider or the supplier in a supply chain, and customers pay the shipping fee. Estimating distribution cost too high or too low can damage a business. If charging too much, a company may fail to gain new business or even lose existing customers to competitors. If charging too little, profit margins can quickly disappear. This motivates logistics providers to analyze and understand the appropriate distribution costs of their services.

If a customer is always served on a stand-alone route from a supplier, the distribution cost to serve can be easily observed or estimated. However, in practice, it is common that a single supplier serves several customers at dispersed locations on the same delivery route. An apparent benefit of the joint distribution strategy is that combining multiple customers on the same route can increase truck utilization and reduce total operational cost. Customers are charged lower fees than stand-alone transportation service would allow. Even though joint distribution provides substantial discounts for the group as a whole, each customer should be charged separately. The route-sharing mechanism makes it a challenge to determine the cost to serve each single customer in joint distribution systems.

Distribution costs are highly dependent on routes employed to deliver products. Due to different accessibilities to delivery routes, there are two kinds of cost analyses which can be conducted. First, logistics collaborators negotiate distribution costs when the delivery route and associated cost are known. This analysis focuses on guaranteeing the fairness of cost allocation to individuals on the route (Frisk, Gthe-Lundgren, Jrnsten, & Rnnqvist, 2010). Second, in most industries, the supplier provides a shipping quote to customers before customers make any order decisions. In this analysis, very little determinate information exists for a new customer. For example, the mixture of customers with different ordering rules leads to unpredictable routes to serve a single customer in the future. Thus, estimating distribution cost for new customers is more challenging and involves more variability. However, for the purpose of pricing, cost estimation takes on a vital role to help suppliers make correct decisions, especially for long-term contracts.

* Corresponding author at: Praxair Inc., Tonawanda, NY 14150, United States. Tel.: +1 7167252776.

E-mail address: leisun@buffalo.edu (L. Sun).

http://dx.doi.org/10.1016/j.cie.2014.11.012
0360-8352/

The two cost analysis processes vary on methodology, available inputs and expected outcomes. Since there is no fixed route for a given customer, the costs allocated to it through a mathematical programming model of cost allocation are different at each delivery, and then the costs of nearby customers can be much different. However, cost estimation relying on continuous approximations can lead to a stable cost value for each customer and smooth cost changes based on customer locations. A comparison between cost allocation and estimation is summarized in Table 1, and more detailed data support will be provided in the following sections.

Even though multiple variable factors are involved in the actual one-time delivery costs, a company in practice will not frequently change the charge for a customer or to offer significantly different prices to customers in the same neighborhood. Thus, the purpose of cost estimation analysis is not simply to approximate the cost allocation results, but to develop a practical pricing mechanism. For general purposes, the distribution cost considered in this paper is only transportation related, for example, not including loading service. This consideration is common in the literature, e.g. Turkensteen and Klose (2012) and Frisk et al. (2010). Since the distribution cost and the travel distance are approximately linearly related (Turkensteen & Klose, 2012), distance is used as a surrogate measure for cost. In this paper, we will analyze the effect of the dispersion of geographical customer locations on the expected distribution costs, determine what these effects are and how they can be quantified, and derive accurate cost estimation.

The remainder of this paper is organized as follows. In Section 2, we review the related literature. Section 3 describes the problem scenario with a preliminary analysis of observed data. In Section 4, we analyze the factors related to cost and construct attributes for predictive modeling. Section 5 describes the prediction scheme and modeling methods for the distribution cost estimation problem. We demonstrate the advantages of the constructed attributes and the approximation model, and explore the impacts of cost instability and delivery frequency through computational experiments in Section 6. We end with some concluding remarks in Section 7.

2. Literature review

Cost estimation is a fundamental activity of many engineering and business processes, e.g. production cost in mechanical manufacturing (Qian & Ben-Arieh, 2008) and maintenance cost of product management (Seo & Ahn, 2006). Since cost estimation is often performed for new products or processes, for which good quality historical data does not exist, the cost model must make the most of sparse, noisy and approximate information (Smith & Mason, 1997).

A typical cost estimation process includes two major steps. First, activity-based costing provides an approach to recognize cost components. For example, Qian and Ben-Arieh (2008) analyzed the design and development phases of machined rotational parts, and then obtained parametric cost representations for an estimation model. In the case of distribution cost, due to the complexity of routing and allocation models, an explicit cost representation cannot be observed directly and thus some intermediate factors are explored for further cost modeling. Second, as procedures to generate numeric results, regression and machine learning techniques have been used widely for predictive modeling. For example, researchers have applied them for plastic injection cost estimation (Wang, 2007) and oil price estimation (Azadeh, Moghaddam, Khakzad, & Ebrahimipour, 2012). For transportation and logistics research, a detailed comparison of regression and neural networks on their differences and similarities is analyzed by Karlaftis and Vlahogianni (2011).

Table 1
Comparison between cost allocation and estimation.

	Cost allocation	Cost estimation
Model type	Discrete	Continuous
Availability for new customers	No	Yes
Stable cost values	No	Yes
Smooth cost changes	No	Yes
Accurately reflect the real cost	Yes	Approximate

Previous research on logistics networks mainly focuses on optimization problems, i.e. finding optimal routes with the minimum total distribution cost. There is a large body of literature on the vehicle routing problem (VRP) from either theoretic or application aspects (Toth & Vigo, 2001). Researchers built various models according to different constraints, e.g. capacity restrictions and time windows, and sought more efficient algorithms due to the NP-hardness of VRP. A more recent interest in distribution is cost analysis on given delivery routes, the purpose of which is mainly for service pricing. Frisk et al. (2010) investigated a number of sharing mechanisms based on economic models to allocate cost to each collaborator on a route. Among various cost allocation methods, a basic and widely accepted fairness criterion is the core of the corresponding cooperative game, which can be approached by a model proposed by Faigle, Fekete, Hochstttler, and Kern (1998). Because of the known routes, cost allocation is treated as post hoc analysis, while the cost estimation problem in this paper provides decision makers a prior knowledge of cost before delivery.

The distribution cost approximation problem without detailed routing information has been studied by researchers in the field of continuous approximation. A continuous model aggregates discrete parameters regarding communities or aggregated groups of customers. The aggregation of parameters allows the model to avoid operational details and focus on higher level decisions (Huang, Smilowitz, & Balcik, 2013). Daganzo (1984) proposed a simple formula for the route length of a capacity constrained VRP based on the number of customers, the area of delivery, the average distance and the vehicle capacities. The results have been improved and extended to scenarios with more constraints by Robust, Estrada, and Lpez-Pita (2004) and Figliozzi (2008, 2009). For application, Turkensteen and Klose (2012) adopted continuous approximation methods to analyze the expected costs of customer groups in the necessary physical distribution system for marketing segmentation research. To apply the continuous approximation models, the locations and demands of potential cooperating customers should be known exactly as parameters. These however are not available when estimating the future long-term cost of single customers. All of the papers reviewed above are applicable only to groups of customers for strategic planning, while this paper is the first to involve impacts of both routing and allocation models to approximate costs to serve single customers for pricing purposes.

The problem of estimating distribution cost to serve a single customer has rarely been addressed in the literature. Kone and Karwan (2011) attempted to predict the cost to serve new customers in an industrial gas delivery network. They developed a new method for clustering customers and then adopted segmented regression. Since only simple observed factors were utilized as variables for modeling, e.g. distance from plant, product consumption rate and some special product related features, their model performed with modest predictive accuracy. We however thoroughly analyze distribution related factors based on routing and allocation activities, construct attributes to reveal potential correlation to cost, and develop a generic method for estimating distribution cost.