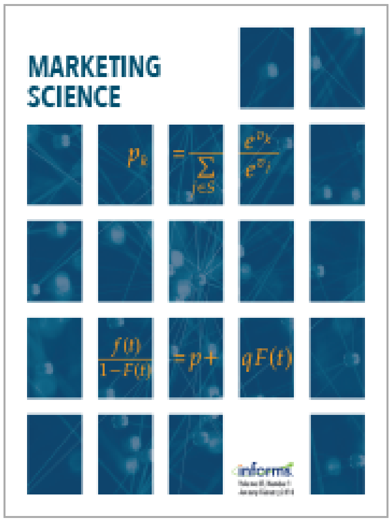


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Multicriterion Market Segmentation: A New Model, Implementation, and Evaluation

Ying Liu, Sudha Ram, Robert F. Lusch, Michael Brusco,

Marketing Science
Vol. 29, No. 5, September–October 2010, pp. 880–894
ISSN 0732-2399 | EISSN 1526-548X | 10 | 2905 | 0880
DOI 10.1287/mksc.1100.0565

Multicriterion Market Segmentation: A New Model, Implementation, and Evaluation

Ying Liu
College of Business Administration, California State University, Long Beach, Long Beach, California 90840, yliu2@csulb.edu

Sudha Ram, Robert F. Lusch
Eller College of Management, University of Arizona, Tucson, Arizona 85721
{ram@eller.arizona.edu, rlusch@email.arizona.edu}

Michael Brusco
College of Business, Florida State University, Tallahassee, Florida 32306, mbrusco@cob.fsu.edu

Market segmentation is inherently a multicriterion problem even though it has often been modeled as a single-criterion problem in the traditional marketing literature and in practice. This paper discusses the multicriterion nature of market segmentation and develops a new mathematical model that addresses this issue. A new method for market segmentation based on multiobjective evolutionary algorithms, called MMSEA, is developed. It complements existing segmentation methods by optimizing multiple objectives simultaneously, searching for globally optimal solutions, and approximating a set of Pareto-optimal solutions. We have applied and evaluated this method in two empirical studies for two firms from distinct industries: descriptive segmentation of the cell phone service market from a dual-value creation perspective and predictive segmentation of retail customers based on profit and customer sociodemographic attributes. The results provide decision makers with compelling alternatives and enhanced flexibility currently missing in existing market segmentation methods.

Key words: multicriterion market segmentation; descriptive market segmentation; predictive market segmentation; multiobjective evolutionary algorithms; multiobjective optimization; multiobjective clustering; Pareto-optimal solution set
History: Received: May 28, 2008; accepted: January 21, 2010; processed by Wayne DeSarbo. Published online in *Articles in Advance* April 8, 2010.

1. Introduction

Market segmentation is a fundamental concept of modern marketing (Wind 1978) and is one of the most pervasive activities in both the marketing academic literature and practice (DeSarbo and Grisaffe 1998). Market segmentation is intrinsically a multicriterion problem. Smith (1956, p. 6) identified this multicriterion nature in his definition of market segmentation: "Market segmentation consists of viewing a heterogeneous market as a number of smaller homogenous markets in response to differing product preferences among important market segments." This implies that market segments, which should be homogeneous within each segment, need to be related to other marketing activity variables, such as response to a particular marketing program or product position, to be useful (Myers 1996).

Identifiability, responsiveness, substantiality, accessibility, stability, actionability, differential behavior, feasibility, profitability, and projectability are a set of generally accepted criteria to evaluate whether a segmentation solution is a good one (DeSarbo and DeSarbo 2007, Frank et al. 1972, Wedel and Kamakura 2000). In practice, multiple criteria are instantiated as a set of optimization objectives or constraints. For any particular market segmentation objective, there may be many constraints originating from managerial, institutional, environmental, and resource-related restrictions (DeSarbo and Grisaffe 1998). In both descriptive and predictive segmentation applications, it is quite common for two or more of the objective criteria to be discordant. In descriptive market segmentation, different segmentation bases describe different characteristics of the customer or marketing mix and have different levels of effectiveness regarding the 10 segmentation criteria previously mentioned. For example, geodemographic data have good support for identifiability, substantiality, accessibility, and stability but often lack actionability and responsiveness. On the other hand, product or service benefit data have good support for actionability and responsiveness but only mediocre support for accessibility and stability. The key here is that two or more segmentation bases are interdependent, but the optimization objectives for the multiple bases may be antagonistic (Ramaswamy et al. 1996).

In predictive market segmentation, there is, unfortunately, often a lack of correlation between the set of descriptive variables and the set of responses. For example, clustering techniques could be used to find homogenous segments of customers that exhibit similar profiles (e.g., age, gender, etc.). However, the resulting segments are not very useful when they are used to predict differences in customer response to a marketing promotion. On the other side, segmentation based on responsiveness does not have the desired segment homogeneity with regard to customer descriptive variables. Jiang and Tuzhilin (2006) described such a kind of relationship between segment homogeneity and predictive performance.

In light of this inherent and troubling problem, a number of multicriterion approaches to segmentation have been proposed. Vriens et al. (1996) provided a review and comparison of some early procedures within the context of conjoint analysis. Krieger and Green (1996) developed a multistage segmentation method that considers one criterion at a time. In the first stage, the *K*-means method is used to cluster respondents into segments that are optimized for identifiability. In the second phase, a heuristic is used to improve the responsiveness of segments with a threshold on the increase in within-segment heterogeneity. Brusco et al. (2002) observed that this multistage approach is inherently suboptimal because the information found in one stage is not shared with the other stages in an optimal way.

A second approach to multicriterion segmentation is based on the transformation of multiple objectives into a single one. DeSarbo and Grisaffe (1998) and DeSarbo and DeSarbo (2007) defined a total utility function to integrate the different criteria and applied single-objective combinatorial optimization approaches. Most mixture models described by Wedel and Kamakura (2000) also fall into the transformation category because multiple criteria are integrated into a single maximum likelihood function. Brusco et al. (2002, 2003) used the transformation approach to solve two- and three-criteria segmentation problems. The multiple criteria were first transformed into a single optimization objective by a weighted sum function. Next, a single-objective simulated annealing heuristic was applied to find segmentation solutions.

Transformation approaches to multicriterion segmentation represent the state of the art in the marketing literature. Nevertheless, we believe that the transformation approach has several serious limitations that have not yet been resolved and that these problems create opportunities for the development of new multicriterion segmentation procedures. One obvious limitation of the transformation approach is that it is often difficult to define an appropriate total utility or set the weight of each objective. A second, and potentially more serious, disadvantage of the weighted-sum transformation approaches is that they only identify what are known as "supported" Pareto-optimal solutions. However, it is well recognized in the multiobjective programming literature that there are often a host of "unsupported" Pareto-optimal solutions that cannot be identified using any set of weights in a transformation approach (Ehrgott and Gandibleux 2000, Ulungu and Teghem 1994).

To illustrate the distinction between supported and unsupported Pareto-optimal solutions, consider a bicriterion optimization problem with two minimization objective functions and the following three Pareto-optimal solutions with objective values $(f1, f2)$: solution A (0.8, 0.6), solution B (0.6, 0.8), and solution C (0.72, 0.72). Each of the transformation approaches would use a convex combination of weights for the two objective criteria. Denoting these weights α and $(1-\alpha)$, such that $0 \leq \alpha \leq 1$, we can observe ranges for α that would enable each solution to achieve the minimum weighted sum. It is clear that solution A has the minimum weighted sum for $0 \leq \alpha < 0.5$, solutions A and B have the same weighted sum for $\alpha = 0.5$, and solution B has the minimum weighted sum for $0.5 < \alpha \leq 1$. Notice that there is no value of α for which solution C has the minimum weighted sum, indicating that this Pareto-efficient solution is unsupported and would not be identified using the state-of-the-art transformation approaches in the extant literature (Brusco et al. 2002, 2003; DeSarbo et al. 2005; DeSarbo and DeSarbo 2007; DeSarbo and Grisaffe 1998). This is unfortunate because a segmentation analyst might actually prefer solution C in light of the better trade-off between the competing criteria.

Our principal contribution in this paper is to propose a viable alternative to multistage and transformation procedures for multicriterion market segmentation. This is accomplished via the development of a metaheuristic evolutionary algorithm that *directly* estimates the entire Pareto-optimal set of solutions (both supported and unsupported). Estimation of the entire Pareto set is far more than just a mathematical advantage. It is conceivable that transformation approaches may fail to identify unsupported segmentation solutions on the Pareto frontier that are most amenable to an efficient allocation of the marketing mix. In such situations, only the direct estimation approach would uncover solutions with segments that are highly responsive to innovative promotional strategies, alternative pricing schemes, or new product initiatives. Moreover, a recurring and disturbing practical problem in market segmentation is that there is no single best solution for all optimization objectives as a result of the multicriterion nature of market segmentation. Because many existing methods do not deal with multiple criteria simultaneously, they usually produce solutions with poor objective values for one criterion or more.

Direct multicriterion clustering algorithms have a rich history in the classification literature. One of the earliest contributions was offered by Delattre and Hansen (1980), who devise an exact algorithm for generating the entire efficient set for a bicriterion problem with partition split and partition diameter as the two objective functions. Ferligoj and Batagelj (1992) also employ direct estimation procedures within the context of hierarchical clustering. Within the past 10–15 years, there has been a tremendous resurgence of interest in direct clustering methods in the machine learning and pattern recognition literature streams (Handl and Knowles 2005, 2007; Law et al. 2004; Ripon et al. 2006). Especially noteworthy is the development of evolutionary algorithms for multiobjective clustering, which is thoroughly outlined in the paper by Handl and Knowles (2007) that summarizes and extends earlier reports on the topics that were published in conference proceedings.

It is important to recognize that the applications of the evolutionary algorithm for multiobjective clustering that have been reported by Handl and Knowles (2005, 2006, 2007) have focused principally on cluster compactness (distance from centroid)- and connectedness (e.g., spanning tree)-objective criteria for a single data set. In market segmentation applications, however, there are often multiple set (or batteries) of data with competing compactness objectives or predictive objectives. The "multicriterion" in multicriterion market segmentation refers to clustering criteria as well as other segmentation criteria such as responsiveness and substantiality. These different optimization objectives and constraints of market segmentation required us to establish refined mathematical models and solution procedures for our specific applications. In short, the specific contributions of our paper are (1) a new formal model for market segmentation based on direct multiobjective optimization, (2) a new evolutionary heuristic for direct estimation of the Pareto-optimal set of segmentation solutions, and (3) two empirical studies that cover both descriptive segmentation and predictive segmentation. As a caveat, we note that the objective criteria in our applications are limited to identifiability and responsiveness (substantiality is also accommodated via segment-size constraints). Restriction to these two principal criteria is concordant with previous applications in multicriterion segmentation (Brusco et al. 2002, 2003; DeSarbo and DeSarbo 2007; DeSarbo and Grisaffe 1998; Krieger and Green 1996).

The remainder of this paper is organized as follows. We formally define multicriterion market segmentation in §2. Section 3 describes a multicriterion market segmentation method that uses multiobjective evolutionary algorithms. Two applications of the method using customer data from two firms are presented in §§4 and 5. Finally, §6 discusses conclusions and future research directions.

2. A Generalized Multicriterion Definition and Method of Market Segmentation

A general definition that directly addresses the multicriterion nature of market segmentation is as follows. To simplify discussion, without loss of generality, assume the problem is to minimize a set of objectives, where

$C =$ the set of segments in the solution, indexed by $c = 1, \ldots, K$, where K is the number of segments of a segmentation solution; one K or a range of K is specified as a constraint;

$z_i \in C =$ the segment membership of customer i, $i = 1, \ldots, N$, where N is the number of customers;

$z = [z_1, z_2, \ldots, z_N]$ is a vector of segment memberships that represents a segmentation solution for all I customers;

(1) $F(z) = [f_1(z), f_2(z), \ldots, f_M(z)]$ is the objective vector to be optimized, where M is the number of objective functions with $M \geq 1$.

(2) $G(Z) = [g_1(z), g_2(z), \ldots, g_P(z)]$ is the constraint vector to be satisfied, where P is the number of constraint functions with $P \geq 1$; P is greater than or equal to one because the range of the number of segments has to be specified.

(3) A specific segmentation solution $z^* = [z_1^*, z_2^*, \ldots, z_N^*]$ is Pareto-optimal if there does not exist another solution z^l such that

(i) z^l satisfies (2), and

(ii) $f_j(z^l) \leq f_j(z^*)$ for all $j = 1, 2, \ldots, M$, and

(iii) $f_j(z^l) < f_j(z^*)$ for at least one j.

(4) The goal is to find a set of segmentation solutions $z^* = [z_1^*, z_2^*, \ldots, z_N^*]$ that are Pareto-optimal with regard to the objective vector (1) and satisfy constraint vector (2).

Most clustering problems encountered in market segmentation are NP-hard (Krieger and Green 1996). It is usually impossible to know whether the generated solution is a real Pareto-optimal solution. Although we evaluate the solutions using alternative methods and call them Pareto-optimal, in practice they represent an approximation.

The general segmentation definition in DeSarbo and DeSarbo (2007) is close to our formulation but has a fundamental difference. The authors discussed the multicriterion nature of market segmentation and the existence of a set of Pareto-optimal solutions; however, instead of taking a direct multiobjective view of the segmentation problem, they defined a general total utility function that integrates the multiple objectives. The multicriterion market segmentation is transformed into a single objective optimization problem. Other multicriterion market studies only defined problem-specific objectives that were either addressed by a multistage approach (DeSarbo and DeSarbo 2007, Krieger and Green 1996) or were integrated into a weighted-sum function (Brusco et al. 2002, 2003). The intrinsic difficulties of multistage and transformation approaches and the practical issues discussed previously require a new perspective on the multicriterion market segmentation problems. Freitas (2004, p. 79) explained the philosophy behind the direct approach to estimating the Pareto-optimal solutions: "…instead of transforming a multi-objective problem into a single-objective problem and then solving it by using a single-objective search method, one should use a multi-objective algorithm to solve the original multi-objective problem. …One should adapt the algorithm to the problem being solved, rather than the other way around."

The range of number of segments is defined as a constraint because the number of segments is still an unresolved problem in market segmentation. The goal of finding a set of Pareto-optimal solutions reflects the reality that there are many practical, acceptable solutions for segmenting a market or a firm's customers. Moreover, a set of solutions offers good flexibility, which is missing in the single-objective optimization method. The above segmentation definition is very general in the sense that it abstracts basic elements of any segmentation problem. It has some desirable attributes compared with existing definitions. First, by definition, the single-criterion market segmentation problem is a special case where there is only one objective function and often a single optimal solution exists. Second, the goal is to find a set of Pareto-optimal solutions that show the diversity and trade-offs of practically acceptable solutions. Third, the actual value of customer membership (z_i) is problem specific. For example, if z_i is a single-segment identifier, the segmentation solution is a partition of all customers. If z_i is a set of segment identifiers, the solution is an overlapped segmentation. If z_i is a probability of membership, the segmentation is a fuzzy one. Finally, the number of segments K is a parameter of a specific segmentation. A specific K or a range of K can be specified as a constraint for a segmentation problem.

2.1. Model Selection in Multicriterion Segmentation

Model selection in cluster analysis is an inherently complex problem that can encompass decisions such as the selection and standardization of clustering variables and the determination of the appropriate number of clusters. Monte Carlo simulation studies have been conducted to compare proposed criteria and methods for variable standardization (Milligan and Cooper 1988; Steinley and Brusco 2008a, b), variable selection (Steinley and Brusco 2008b), choice of the number of clusters in nonmodel-based clustering (Dimitriadou et al. 2002, Milligan and Cooper 1985), and choice of the number of clusters in model-based clustering (Andrews and Currim 2003a, b). Although these comparative studies have provided some insight with respect to the relative performance of criteria for synthetic data structures, there are no definitive recommendations for what tends to work best in practice. Consequently, even in the case of traditional clustering methods (e.g., K-means), model selection often requires a subjective assessment of interpretability in addition to quantitative measures.

In multicriterion segmentation, the model selection process is further complicated by the fact that it is also necessary to select a solution from the estimated Pareto-optimal frontier. For direct approaches to multicriterion segmentation, the characteristics of the Pareto front make some solutions more interesting than the others. A solution from the estimated Pareto front can be selected based on a variety of criteria including (a) choosing the solution closest to the origin, (b) identifying a solution that corresponds to a "knee" in the plot of the Pareto front (in a two-dimensional solution space, a knee is the solution with the maximum marginal rate of return that is measured by the ratio of a gain in one objective to a loss in another objective, i.e., similar to a scree plot in factor analysis), and (c) using the segment profile such as segment size and statistics of segmentation bases for solution and segment screening.

3. Multicriterion Market Segmentation Using an Evolutionary Algorithm

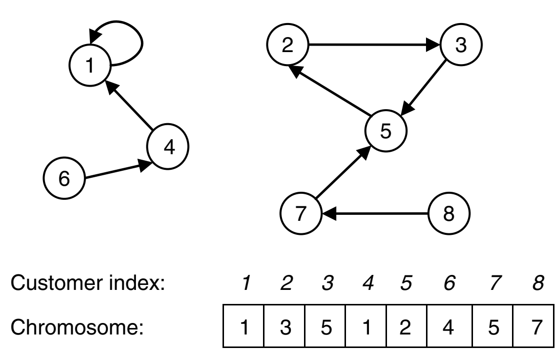
Efficient heuristic methods are required to address the computational complexity of market segmentation. Although many multiobjective optimization algorithms exist, multiobjective evolutionary algorithms are the most popular and well-studied approaches because they find a set of Pareto-optimal solutions in a single run (Coello et al. 2002). Evolutionary algorithms are heuristic search algorithms that can efficiently and effectively search for optimal solutions in a large solution space because of the built-in implicit parallelism (Goldberg 1989, Vose and Wright 2001). We propose a method called multicriterion market segmentation using an evolutionary algorithm (MMSEA).

In MMSEA, the segmentation solution is represented in a genetic structure called a chromosome. MMSEA first generates a set of initial optimal chromosomes (solutions). The initial solutions are evaluated, and only Pareto-optimal solutions are archived. A set of archived solutions evolves to the optimal solution set through a number of generations of chromosome selection, reproduction, and archiving. The selection procedure uses a fitness function to determine which solutions will be used in reproduction. MMSEA uses a crossover operation for chromosome reproduction. Crossover allows genetic information within various chromosomes to be exchanged. In the segmentation context, a crossover operation involves splicing together information from two segmentation solutions (the parents) to form new segmentation solutions (the children). After crossover, the generated child solutions are merged into the archived solution set. During the merge process, all child solutions are evaluated, and only Pareto-optimal solutions are archived at the end of each generation. The algorithm stops when the improvement of objective values of the archived solution set is very small between two generations. The algorithm details are given in the appendix, and the key algorithm design decisions are described as follows.

3.1. Chromosome Representation

Park and Song (1998) proposed a graph-based adjacency representation that has the advantage of allowing solutions with different numbers of segments to reproduce children. In the graph-based adjacency representation, a chromosome is an array as illustrated in Figure 1. Each array element represents one customer, and the element value is the array index of another customer that is in the same segment. In Figure 1, the eight customers form two connected graphs, i.e., two customer segments.

Figure 1 Chromosome Representation



3.2. Initial Solution Set Generation

The quality of the initial solution set has a big impact on the performance of evolutionary algorithms. For descriptive segmentation, a simple K-means algorithm is used to generate an initial solutions for each segmentation basis. For predictive segmentation, both K-means and clusterwise regression are used to generate an initial solution set. However, the convergence speed is slow with the initial solution set. Handl and Knowles (2006) found that aligning the initial solutions to the graphic-based adjacency representation of a minimum spanning tree helps to reduce the convergence time in multicriterion clustering algorithms. We used this method and found significant performance improvement.

3.3. Selection and Crossover

MMSEA uses a binary tournament algorithm and a density-based fitness function to select the parent solutions. The selection process first randomly selects two solutions from the current solution set. The binary tournament algorithm compares the fitness values of the two solutions and the one with a higher fitness value wins. The fitness of a solution is measured by the average Euclidean distance of the solution and its m-nearest neighbors that have the same number of segments in the objective space; m is calculated at runtime and is the square root of the number of archived solutions of the same number of segments. This simple selection process helps to select diverse solutions by allowing less crowded solutions to have more chances to be selected. MMSEA uses a uniform crossover operation that combines two parent solutions to produce two child solutions. In uniform crossover, individual elements in the two chromosome arrays are swapped with a specified probability (crossover rate).

3.4. Parameter Tuning

The performance of MMSEA is determined by two metrics: the speed to convergence and the solution set diversity. The speed to convergence measures how fast initial solutions approach to the Pareto front and is measured by the improvement of objective values between two generations. The solution diversity is the degree of representativeness of generated solutions. It is measured by the density of solutions (i.e., how close together the solutions are). Parameters include the maximum number of generations, initial solution set size, child solution set size, and crossover rate control the evolution process. The maximum number of generations was determined when the improvement of the archived solution set is small for a specified number of generations. To set other parameters, we ran a set of experiments in which we changed parameter values systematically and measured the convergence speed and solution diversity of the generated solution set. In those experiments, the initial solution set size was changed from 10 to 100 with a step length of 10. The child solution set size was changed from 1 to 100 with varied step lengths from 1 to 20. The crossover rate was changed from 0.001 to 0.5 with varied step lengths of 0.001, 0.01, and 0.1.

3.5. Selecting Solutions from the Estimated Pareto-Optimal Set

The estimated Pareto-optimal set is apt to contain hundreds if not thousands of partitions (i.e., segmentation solutions). How then should a segmentation analyst choose solutions from this set? As noted in §2.1, methods for selection include proximity to the origin, a knee in the Pareto plot, and segment profile information. One strategy that we have found effective is to visually identify several (two to five) regions of the Pareto-optimal set that have managerially favorable trade-offs among the objective criteria and then select one representative solution from each region based on distance from the origin. The fact is that, in most instances, the representative solution from each region should have basically the same cluster structure as its nearby neighbors. Next, each of the representative solutions should be profiled on the segmentation variables and possibly some exogenous variables not considered in the clustering process. The selection of a final segmentation solution can then be made on the basis of which of these segment profiles are most conducive to effective allocation of the marketing mix.

4. Evaluation One: Descriptive Market Segmentation

4.1. Segmentation Based on Customer and Firm Value

Vargo and Lusch (2004) and Boulding et al. (2005) viewed marketing as the management of the dual creation of firm (shareholder) and customer value. Accordingly, the challenge for every company is to be able to understand and differentiate heterogeneous customers based on their needs and to subsequently make a market offering that has value for both the customer and the firm. The segmentation goal of this evaluation is to model the value proposition of cell phone services by considering both the service provider's financial performance and the benefit customers receive from the provider's offerings. From the view of the service provider, the firm's value from customer (called *firm value* hereafter) may be measured by customer revenue and service cost. From the customer's perspective, the surrogates for the firm's product and service benefit to the customer (called customer benefit hereafter) could be the call minutes and number of incomplete and blocked calls.

However, it is well known that this dual value or symmetric value often does not occur because some customers may not receive much benefit from the firm but provide more value to the firm, whereas others may find high benefit in the firm's offerings but the firm may not receive high value. When the benefits or values are asymmetric, then either the customer will be motivated to switch—for example, when the value to the firm is high and the customer benefit is low—or the firm will be motivated to increase fees or lower service when the customer receives high benefit but the firm is not benefiting in terms of value received. Therefore this type of segmentation strategy not only provides homogeneous groups of customers for targeted marketing efforts but also provides strategic guidance to the firm.

4.2. The Data Set

In this study we used a data set donated by an anonymous cell phone service provider. It has a cell phone service transaction database of 31,047 customers. Because of the computer memory limit, only 10,000 cases were selected randomly from the database as the segmentation sample. The descriptive statistics for the raw sample data are shown in Table 1. Based on the value proposition segmentation model, six attributes were selected to form the customer benefit basis, and four attributes were selected to form the firm value basis. Consequently, the segmentation is a descriptive, joint market segmentation with two bases. Because the standardized data perform well in clustering algorithms (Milligan and Cooper 1988), all variables are standardized using z-scores for the segmentation process.

4.3. The Segmentation Model

The within-cluster omega squared (WCOS) was used to measure the homogeneity (a surrogate

Table 1 Descriptive Statistics for the Segmentation Bases

Attribute	Minimum	Maximum	Mean	Std. dev.
Customer benefit attributes				
Mean monthly minutes of use	0.00	7,359.25	549.44	549.75
Mean overage minutes of use	0.00	1,887.25	40.21	93.25
Mean number of outbound voice calls	0.00	409.33	26.54	35.92
Mean number of inbound voice calls	0.00	281.67	8.56	16.47
Mean number of dropped voice calls	0.00	166.33	6.17	9.36
Mean number of blocked voice calls	0.00	204.67	4.17	10.53
Firm value attributes				
Mean monthly revenue	−2.52	1,223.38	59.53	45.71
Mean total recurring charge	−8.71	341.93	47.67	24.22
Months in service	6.00	59.00	18.52	9.95
Mean number of customer care calls	0.00	327.33	2.01	5.85

of identifiability) for each segmentation basis. The WCOS is defined as follows, where

x_{ij} = the value of attribute j for customer i;
$i = 1, \ldots, I$, where I is the number of customers;
$j = 1, \ldots, J$, where J is the number of attributes in the segmentation basis;
$I(c)$ = the set of customers in the cluster c;
$c = 1, \ldots, K$, where K is the number of segments;

(1) $\bar{x}_{j_c}$ = the mean of attribute j of cluster c: $\bar{x}_{j_c} = 1/|I(c)| \sum_{i \in I(c)} x_{ij}$.
(2) $\bar{x}_j$ = the mean of attribute j for all customers: $\bar{x}_j = 1/I \sum_{i=1}^{I} x_{ij}$.
(3) Within-cluster sum of squares (WCSS) $= \sum_{c=1}^{K} \sum_{j=1}^{J} \sum_{i \in I(c)} (x_{ij} - \bar{x}_{j_c})^2$.
(4) Total sum of squares (TSS) $= \sum_{j=1}^{J} \sum_{i=1}^{I} (x_{ij} - \bar{x}_j)^2$.
(5) Within-cluster omega squared (WCOS) $= WCSS/TSS$.

As an instance of the general multicriterion market segmentation problem defined in §2, the objective functions are

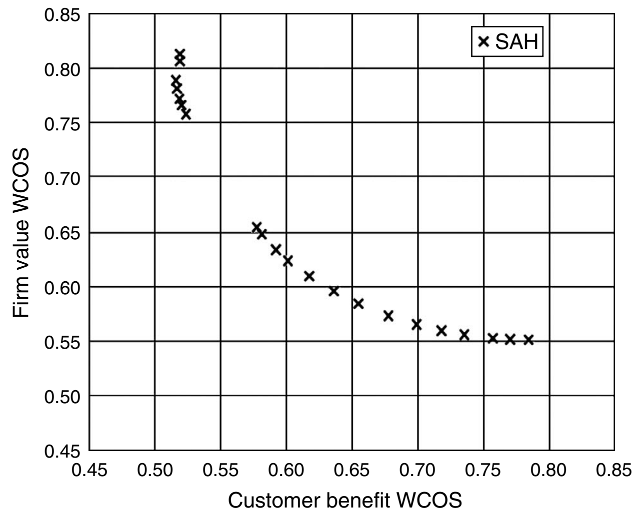
(6) $f_1(z)$ = minimize the WCOS of customer benefit basis.
(7) $f_2(z)$ = minimize the WCOS of firm value basis.
(8) $[f_1(z), f_2(z)]$ = the objective function vector.

As defined in §2, $z = [z_1, z_2, \ldots, z_I]$ is a segment solution for all I customers. This segmentation model optimizes both identifiability and responsiveness. Identifiability is satisfied because the segment homogeneity is explicitly optimized for both dimensions. Responsiveness is satisfied because the segmentation bases are customer behavioral data that represent usage of cell phone service(s).

4.4. Computational Evaluation

Given that the principal focus of our experimental analyses was to compare our MMSEA method to state-of-the-art transformation approaches in multicriterion segmentation, we wanted to adopt a model-selection process that would facilitate a fair comparison. To that end, beginning with $K = 2$ segments, we implemented 10 restarts of the simulated annealing heuristic (SAH) developed by Brusco et al. (2002) for each value of α on the interval $0 \le \alpha \le 1$ in increments of 0.05, where α is the weight placed on WCOS for customer benefit and $(1 - \alpha)$ is the weight placed on WCOS for firm value. The best weighted objective value across the 10 restarts was chosen as the solution for each value of α. To choose the best α value, we identified the solution with WCOS values closest to the origin. This criterion provides a straightforward and objective basis for selection that can also be used for MMSEA. We repeated this process for $3 \le K \le 6$ segments. The WCOS values monotonically decrease as K increases; however, the sharpness of the decline decreased substantially after $K = 4$. For this reason, we adopted a four-segment solution for interpretation.

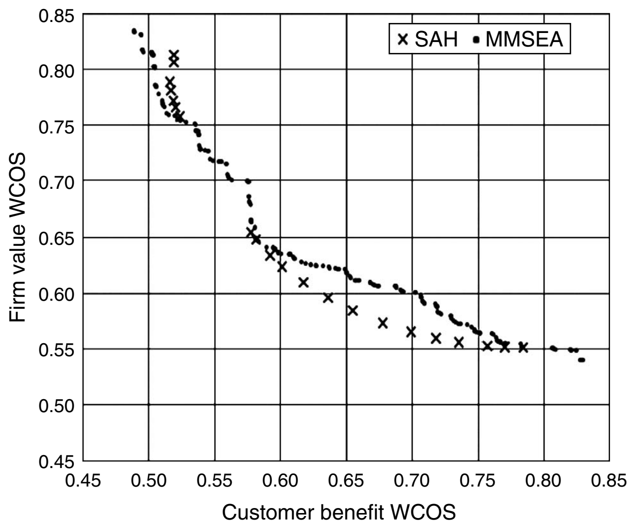
Figure 2 **SAH Four-Segment Solutions**



The estimated Pareto-optimal frontier of four-segment solutions obtained by the transformation approach is shown in Figure 2. The points in the bottom right and top left positions of the graph correspond to weighting schemes of $\alpha = 1$ and $\alpha = 0$, respectively. The remaining 19 points in Figure 2 correspond to WCOS values for the two objective criteria at the other values of α. It is interesting to observe several gaps along the frontier, particularly the large gap in the curve where WCOS (firm value) is approximately 0.76 at one point and then drops to roughly 0.65 at the next point on the frontier. Figure 2 exemplifies the potential shortcomings of the transformation approach, namely, the failure to identify a large number of unsupported points on the efficient frontier.

Figure 3 shows the comparison of SAH and MMSEA four-segment solutions. The 21 SAH solutions are the same solutions as in Figure 2. There are 200 MMSEA results that spread over the solution objective space and fill the gaps of SAH solutions. As shown in Figure 3, MMSEA is able to find many unsupported Pareto-optimal solutions that SAH or other weighted-sum transformation methods cannot uncover. To compare the membership agreement of SAH and MMSEA, we selected the solution closest to the origin from each algorithm and calculated Hubert and Arabie's (1985) adjusted Rand index (ARI) as a measure of agreement between the two partitions. The selected SAH solution has an α value of 0.5, showing that this is a balanced solution. The objective values of the selected SAH solution are (0.60, 0.62), and the objective values of the selected MMSEA solution are (0.58, 0.65). The calculated ARI value of 0.86 showed that the selected SAH and MMSEA solutions

Figure 3 Comparison of Four-Segment Solutions



are similar. The conclusion here is that MMSEA estimates supported Pareto solutions that have similar segmentation structure as those estimated by SAH. However, as we shall see, MMSEA offers much more because of its ability to identify unsupported solutions that are not uncovered by SAH.

The diversity of the MMSEA solution set allows decision makers to investigate other solutions smoothly along the Pareto front. In the gap of SAH solutions, we select another MMSEA solution whose objective values are 0.54 and 0.73. The ARI value between this solution and the selected SAH solution is 0.50. Based on ARI guidelines offered by Steinley (2004), the value of 0.50 showed that the two partitions exhibited moderate agreement.

Table 2 shows the segment-level comparison of the two solutions. The two solutions differ in both segment size and segment profiles measured by attribute means and standard deviations. The MMSEA solution has better within segment homogeneity in the customer benefit basis, whereas the SAH solution has better identifiability in the firm value basis. For example, the segment means of customer benefit attributes of the MMSEA solution are highly different from each other. Comparatively, the segment means of customer benefit attributes of the SAH solution have small differences between the two big segments (segments 1

Table 2 Comparison of MMSEA and SAH Solutions

	MMSEA solution Customer benefit WCOS = 0.54 Firm value WCOS = 0.62				SAH solution Customer benefit WCOS = 0.57 Firm value WCOS = 0.55			
	Seg 1	Seg 2	Seg 3	Seg 4	Seg 1	Seg 2	Seg 3	Seg 4
Segment size	6,294	2,893	640	173	4,915	2,334	2,321	430
Customer benefit attributes								
Mean monthly minutes of use	253.84 (191.61)	856.97 (357.24)	1,585.30 (536.89)	2,329.26 (1,021.10)	309.37 (234.46)	302.63 (259.14)	1,022.06 (404.16)	2,082.25 (781.47)
Mean overage minutes of use	11.31 (24.51)	62.69 (76.40)	93.31 (87.51)	519.42 (276.37)	14.68 (30.64)	17.98 (36.85)	70.72 (81.36)	288.19 (268.15)
Ratio of overage minutes to use minutes (%)	**4.46**	**7.32**	**5.89**	**22.30**	**4.75**	**5.94**	**6.92**	**13.84**
Mean number of outbound voice calls	10.19 (11.46)	39.91 (25.58)	105.07 (53.52)	107.20 (79.24)	12.27 (13.85)	13.77 (15.67)	52.56 (33.89)	118.51 (70.49)
Mean number of inbound voice calls	2.76 (4.99)	11.71 (11.78)	44.53 (32.90)	33.78 (40.70)	3.37 (5.87)	3.79 (6.74)	17.15 (16.99)	47.27 (42.28)
Mean number of dropped voice calls	2.54 (3.06)	9.29 (7.26)	23.18 (18.90)	23.52 (20.01)	3.02 (3.66)	3.41 (4.27)	11.84 (9.38)	26.69 (22.75)
Mean number of blocked voice calls	1.79 (3.49)	6.31 (10.58)	15.44 (26.74)	13.31 (25.05)	2.17 (4.37)	2.02 (4.23)	8.61 (15.16)	14.77 (27.13)
Firm value attributes								
Mean monthly revenue	39.51 (16.32)	80.66 (32.21)	104.44 (42.05)	268.04 (119.58)	42.01 (17.90)	44.90 (21.48)	87.55 (33.38)	187.81 (105.24)
Mean total recurring charge	37.65 (15.94)	61.04 (21.74)	73.59 (29.86)	92.25 (48.81)	40.00 (16.24)	38.91 (17.48)	64.83 (22.24)	90.19 (42.69)
Ratio of recurring charge to revenue (%)	**95.29**	**75.68**	**70.46**	**34.42**	**95.22**	**86.66**	**74.05**	**48.02**
Months in service	19.01 (9.91)	17.47 (9.64)	17.81 (10.70)	20.54 (11.51)	13.32 (4.56)	31.36 (7.62)	16.37 (8.48)	19.75 (11.33)
Mean number of customer care calls	0.75 (2.01)	3.01 (4.68)	9.07 (17.56)	5.12 (7.72)	0.95 (2.26)	0.80 (2.21)	4.24 (6.40)	8.76 (19.96)

Note. The middle and bottom panels contain the within-segment variable means and, in parentheses, the within-segment standard deviations for the two segmentation bases.

and 2). This observation is consistent with the differences of their objective values (0.54 versus 0.60 for customer benefit WCOS).

Moreover, we calculated two derived attributes in Table 2: the ratio of mean overage use of minutes to mean of monthly minutes of use and the ratio of mean total recurring charge to mean monthly revenue (in bold). The two ratios are used to determine if a customer has the right calling plan. Specifically, the segment 4 of the MMSEA solution has the highest ratio of the overage minutes of use and the lowest ratio of the total recurring charge. The segment's overage minutes of use ratio of 22.30% is three times the overall ratio of 7.32%. The overall total recurring charge ratio of 80.01% is 2.3 times the segment ratio of 34.42%. The marketing managers of this cell phone service provider may deem it crucial to identify and resolve the issues that underlie these abnormal values because this segment has the longest months in service and the highest revenue. A better understanding of the root causes for the high overage minutes of use and the low total recurring charge in this segment should enable management to devise a plan that fosters greater communication of product, service, and pricing information, which should create a winning value proposition profitably. This demonstrates the importance of insights that may be gleaned from estimating the entire Pareto-optimal set of solutions for a multicriterion segmentation problem, as opposed to using a transformation approach that only seeks supported Pareto solutions on the frontier. Nonetheless, it should be noted that the comparison will be different if a different solution is selected. The example shows an important fact of market segmentation that, for a specific set of customers, there are many diverse solutions exhibiting different properties. Predictably, part of the market segmentation problem can be quantified and automated, but other parts of the problem involve seasoned managerial judgment. This is especially important because the selection of different segment solutions can lead to different managerial conclusions and insights.

4.5. Algorithm Scalability and Cross-Sample Validation

In the above evaluation, it took MMSEA 32 minutes and 2 seconds to generate 600 solutions of different number of segments (3, 4, and 5) on a desktop PC with an Intel® Core™ i7-920 processor running a 32-bit Windows XP operating system. That was about 3.2 seconds for each solution. Program profiling showed that 90% of the run time was spent in solution initialization and objective calculation in solution evolution. Nonetheless, the memory size sets the limit of the number of customers MMSEA can process in our computer. During initialization, MMSEA needs to calculate, sort, and store the distances between each pair of customers. Let N be the number of customers; the memory size to store all the distance values is N^*N^*8 bytes (in an optimized implement, we could cut the space by 50%). MMSEA uses Java, which uses eight bytes to store a distance value. The size of 10,000 customers requires about 800 MB of memory space, which approaches the Java memory limit in the 32-bit Windows XP operating system. The computation complexity of this initialization is $O(N \log N)$. In MMSEA, all genetic operations (selection, crossover, and archiving) have a computation complexity of $O(N)$. Evolutionary algorithm can be easily customized to run in parallel because selected parents can be partitioned into multiple processes to allow parallel evolution.

MMSEA is a heuristic random search algorithm that may be sensitive to the sample data and initial seeds. If this is a severe problem, then the results may not be valid. To help provide an initial assessment of the cross-sample validity of the MMSEA results, we ran MMSEA using two nonoverlapping 10,000-case random samples. We used the four-segment MMSEA solution of sample 1 that was closest to the origin as the partition of sample 1. A four-segment partition for sample 2 (called P1) was established by assigning each case in sample 2 to the same cluster as its nearest neighbor in sample 1. Then we took the four-segment MMSEA solution of sample 2 that was closest to the origin as a second partition for sample 2 (called P2). The ARI obtained for P1 and P2 was 0.77, which reflects fairly strong agreement (Steinley 2004).

5. Evaluation Two: Predictive Market Segmentation

In this evaluation, we applied the multicriterion market segmentation to a large services and solutions retailer that has several hundred stores in the United States and Canada. We evaluated predictive market segmentation and compared the MMSEA results with those obtained using a weighted sum transforming approach (Brusco et al. 2003) and a concomitant finite mixture model (FMM) (Dayton and MacReady 1988). Both compared methods are able to segment and profile customers simultaneously (Brusco et al. 2003, Gupta and Chintagunta 1994). We used the concomitant FMM implementation by Grün and Leisch (2007) because it is a free and open source. The same MMSEA program of the first evaluation was used here, and this fact demonstrated the flexibility of the metaheuristic nature of MMSEA method.

5.1. Segmentation Model

Because of its direct managerial relevance, profit is often more effective than other segmentation bases (Bock and Uncles 2002). On the other side, sociodemographic data help to answer where the customers are and what their profiles look like, fulfilling the identifiability and accessibility criteria for market segmentation. Two optimization objectives and two constraints were defined for this model. The WCOS defined in the previous evaluation was used to measure the within-segment homogeneity of predictors (the sociodemographic variables). The second objective was measured by the total residual sum of squares (RSS) of the profit data for the segment-level linear predictive model. To meet the substantiality criterion and make the predictive model meaningful, a constraint of the minimum segment size of 30 was defined in addition to the range of the number of segments. As a result, this segmentation model explicitly optimizes identifiability (the homogeneity of predictors), responsiveness (the predictive objective), and substantiality (the minimum segment size).

5.2. The Data Set

We randomly selected 1,500 customers from 20,921 customers for this evaluation. The descriptive statistics of all variables are listed in Table 3. Profit is the dependent variable, whereas the others are independent variables. The profit is the gross profit that is the difference of sales and cost of merchandise and services sold to each customer. It is transformed using the ln( ) function to reduce its skewness and improve its normality. As in the first evaluation, all variables are standardized using z-scores for the segmentation process. We also checked that there was no significant collinearity in the linear predictive model.

The ordinary linear regression (Tables 4 and 5) on the sample shows a very weak relationship between the profit and the sociodemographic predictors. Although some predictors are significant, the goodness of model fit is not good because the seven sociodemographic predictors only explain about 3% of the profit variance. We attribute this weak relationship to the customer heterogeneity; i.e., different customer segments generate different profits. MMSEA should produce a better predictive model at the segment level.

5.3. Evaluation Results

To provide a competitor for the MMSEA procedure, we adapted the multicriterion clusterwise regression algorithm developed by Brusco et al. (2003). The method was based on the SAH solution strategy. The adaptations included a reduction from two dependent variables to one and, more importantly, the replacement of a nonnegative least-squares algorithm with a standard (no constant) ordinary least squares (OLS) regression model for fitting the within-cluster regressions. As was the case for the previous example, we began with $K = 2$ segments and implemented multiple restarts of the simulated annealing heuristic for each value of α on the interval $0 \le \alpha \le 1$ in increments of 0.05, where α is the weight placed on the proportion of variation explained in the dependent variable and $(1 - \alpha)$ is the weight placed on proportion of variation explained for the descriptor variables. Because of the increased computation time required to estimate the within-cluster regression models, we limited the number of restarts to five for each value of α. We repeated this process for $3 \le K \le 6$ segments. In light of the fact that the sharpness of the increases in explained variation moderated substantially after $K = 3$, we adopted a three-cluster solution for interpretation.

5.3.1. Comparison of Optimization Objectives. It took MMSEA 2 hours, 29 minutes, and 10 seconds to generate 600 solutions of different number of segments (three-segment, four-segment, and five-segment). That is about 14.9 seconds for each solution. Most time was spent on solution initialization (59 minutes and 22 seconds) and regression objective calculation (about 90% of solution evolution time). We ran the FMM algorithm three times and selected the result that had the best maximum likelihood value. Figure 4 shows the solution objective values of three-segment solutions for all three methods. There are 200 MMSEA solutions, 16 SAH solutions, and 1 FMM solution. The FMM solution is

Table 3 Descriptive Statistics of Model Variables

Attributes	Minimum	Maximum	Mean	Std. dev.
Age	18	86	46.80	12.67
Gender	0	1	0.72	0.45
Marital status	0	1	0.79	0.411
Working woman	0	1	0.48	0.50
Children	0	1	0.57	0.50
Adult	1	4	2.17	0.71
Income	1	9	6.28	2.05
Profit	0.84	117.22	14.49	11.98

Table 4 Predictive Model Summary

R-Squared	0.031
Adjusted R-squared	0.026
Regression sum of squares (df = 7)	45.24
Residual sum of squares (df = 1,493)	1,435.88
Total sum of squares (df = 1,500)	1,481.12

Table 5 Coefficients of Predictors

Variable	Coefficients	Significant
Age	−0.143	0.000
Gender	0.057	0.029
Marital status	−0.030	0.292
Working woman	−0.028	0.279
Children	−0.090	0.001
Adult	0.021	0.435
Income	0.086	0.001

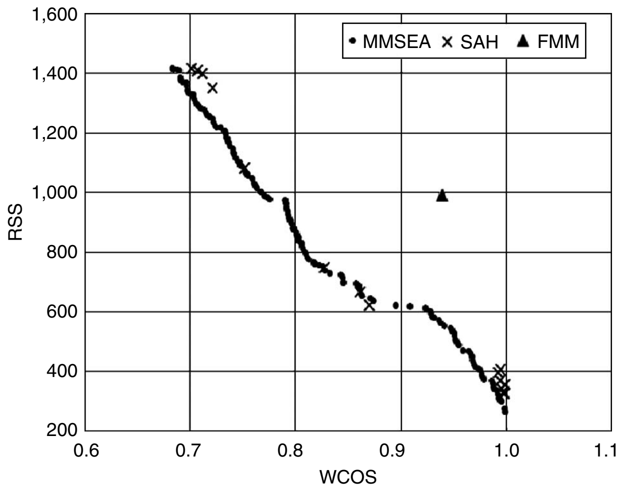


Figure 4 **Comparison of Three-Segment Solutions**

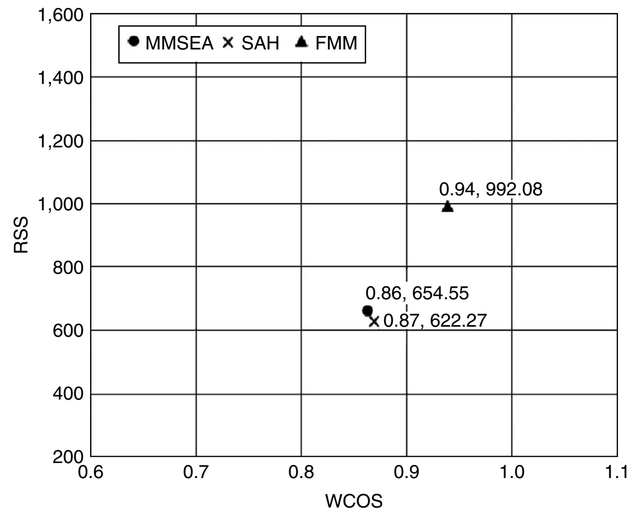


Figure 5 **Selected Three-Segment Solutions**

dominated by many solutions of MMSEA and SAH. Again, the MMSEA solutions have a better coverage of the Pareto front that includes some unsupported Pareto-optimal solutions. The Pareto front of MMSEA allows decision makers to choose solutions from two or more conflicting objectives fairly smoothly. Within the two extreme solutions of a Pareto front there are hundreds of solutions to choose from. It is clear to know how much we will gain and give up in each dimension if we select an alternative solution.

Table 6 **Comparison of Three-Segment Solutions**

	MMSEA WCOS = 0.863 RSS = 654.55			SAH WCOS = 0.870 RSS = 622.27			FMM WCOS = 0.939 RSS = 992.08		
	Seg 1	Seg 2	Seg 3	Seg 1	Seg 2	Seg 3	Seg 1	Seg 2	Seg 3
Segment size	885	405	210	630	576	294	719	519	262
R-squared	0.370	0.840	0.233	0.611	0.656	0.225	0.045	0.691	0.617
Age	**−0.112**	**−0.225**	**−0.136**	**−0.189**	**−0.074**	**−0.171**	**−0.184**	**−0.076**	**−0.202**
Gender	**0.125**	**0.129**	**−0.336**	**0.062**	**0.141**	**−0.328**	0.077	**1.310**	**−0.414**
Marital status	**−0.152**	**−0.105**	−0.020	−0.080	**−0.118**	**−0.127**	−0.019	**−0.175**	**−0.086**
Working woman	**0.481**	**−1.197**	**0.221**	**−0.806**	**0.799**	**0.115**	−0.043	−0.034	−0.050
Children	**−0.106**	**−0.214**	**−0.204**	**−0.140**	**−0.093**	−0.081	**−0.099**	**−0.058**	**−0.127**
Adult	−0.046	−0.012	**0.282**	0.037	**−0.057**	**0.155**	0.006	0.038	0.026
Income	**0.077**	**0.100**	−0.072	**0.101**	−0.000	0.076	**0.096**	**0.105**	0.047
Age	47.08	46.87	46.80	46.75	46.88	46.76	47.44	46.47	45.69
	(12.91)	(11.84)	(12.67)	(11.96)	(13.02)	(13.50)	(13.27)	(11.55)	(13.08)
Gender	0.73	0.69	0.72	0.71	0.69	0.79	0.72	1.00	0.18
	(0.44)	(0.46)	(0.45)	(0.45)	(0.46)	(0.41)	(0.45)	(0.00)	(0.38)
Marital status	0.90	0.93	0.79	0.97	0.89	0.18	0.75	0.84	0.78
	(0.30)	(0.25)	(0.41)	(0.18)	(0.31)	(0.39)	(0.43)	(0.37)	(0.42)
Working woman	0.49	0.41	0.48	0.43	0.53	0.51	0.46	0.53	0.46
	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Children	0.58	0.65	0.57	0.64	0.57	0.39	0.57	0.60	0.49
	(0.49)	(0.48)	(0.50)	(0.48)	(0.49)	(0.49)	(0.50)	(0.49)	(0.50)
Adult	2.27	2.34	2.17	2.34	2.30	1.52	2.15	2.21	2.10
	(0.63)	(0.64)	(0.71)	(0.61)	(0.63)	(0.67)	(0.72)	(0.69)	(0.70)
Income	6.51	6.66	6.28	6.71	6.52	4.88	6.44	6.31	5.77
	(1.87)	(1.96)	(2.05)	(1.85)	(1.86)	(2.19)	(1.91)	(2.09)	(2.23)
Profit	13.46	17.62	14.49	14.95	14.80	12.91	6.53	23.10	19.28
	(11.01)	(14.65)	(11.98)	(12.12)	(13.20)	(8.61)	(3.03)	(13.13)	(10.59)

Notes. The top panel contains the segment sizes, within-segment *R*-squared values, and within-segment slope coefficients (significant at 0.05 level shown in bold) for each method. The bottom panel contains the within-segment variable means and, in parentheses, the within-segment standard deviations.

5.3.2. Individual Solution Comparison. To provide more detailed information, we compared three individual solutions from the three methods. We first selected a knee from middle area in the SAH solution set. Intuitively, a knee is a solution that when compared with its neighbors in the Pareto front, its objective values decrease a lot in one dimension but increase much less in another dimension. Then in the neighborhood area of the selected SAH solution, we selected a knee from the MMSEA solution set. The objective values of the selected solutions and the FMM solution are depicted in Figure 5. Both MMSEA and SAH solutions have better (lower) objective values than the FMM solution in both dimensions. The ARI for the MMSEA and SAH solutions is 0.36, which means that the selected solutions of the two methods have a weak agreement on segment membership. For the two pairs of MMSEA-FMM and SAH-FMM solutions, the ARIs are 0.02 and 0.01, respectively, indicating that there is hardly any agreement in the solution pairs.

The details of the three selected solutions are compared side by side in Table 6. R-Squared values increased significantly for all segments except FMM segment 1. The better segment-level predictive models proved our assumption that the weak relationship between profit and predictors before segmentation was attributed to the customer heterogeneity. In all segment models of MMSEA and SAH solutions, most predictors are significant at the 0.05 level. Although MMSEA and SAH have similar objective values, there are significant differences in their segment sizes, segment members, segment profiles, and predictive models. This fact demonstrates the diversity of segmentation solutions. Decision makers may want to use different methods to find the solutions that best meet their specific requirements. For example, the MMSEA solution distinguished itself by the best predictive model (with an R-squared value of 0.84) in the high-profit segment (with the highest profit mean of 17.62). On the other hand, the SAH solution has good predictive models for the two big segments.

Table 7 shows the results of one-way analysis of variance (ANOVA) of all variables among the three segments. As expected, all three methods can generate identifiable segments for most variables because they simultaneously optimize responsiveness and identifiability. Nonetheless, MMSEA and SAH solutions are more identifiable than the FMM solution. This is consistent with the comparison of their WCOS values.

Table 7 ***p*-Value of One-Way ANOVA of Segment Means**

Attributes	MMSEA	SAH	FMM
Age	0.269	0.981	0.124
Gender	0.185	0.008	0.000
Marital status	0.000	0.000	0.001
Working woman	0.000	0.001	0.019
Children	0.000	0.000	0.013
Adult	0.000	0.000	0.107
Income	0.000	0.000	0.000
Profit	0.000	0.040	0.000

6. Conclusion

A multicriterion view of market segmentation brings a new problem definition and requires new segmentation methods. MMSEA complements existing methods by directly tackling the multicriterion nature of market segmentation using multiobjective optimization. As a result, no multicriterion aggregation or trade-offs are required before the users see the full spectrum of the solution space. Decision makers usually prefer posterior analysis to upfront trade-offs because such a priori decisions are difficult to make and impose unpredictable limitations on the search for the globally optimal solution(s). Pareto optimality compares values in multiple dimensions and eliminates the transformation issues. MMSEA estimates the whole efficient frontier (both supported and unsupported solutions), whereas the existing transformation approaches do not. This holistic view gives decision makers a lot of flexibility and insights that are missing in many existing methods.

Another distinct feature of MMSEA is that it is based on the proven metaheuristic evolutionary algorithm. The evolutionary algorithm is robust in terms of data types and objective function forms (discrete or continuous, concave or convex, single modal or multimodal) although it does not lend itself to statistical inference. As demonstrated in our evaluation, the metaheuristic nature of the algorithm allows the same MMSEA algorithm to be applied to different segmentation problems easily. Furthermore, search constraints such as segment size and solution preferences can be easily incorporated into the search process to generate solutions with desired properties.

We propose to extend our multicriterion market segmentation research in many ways. The algorithm can be parallelized to improve its performance. The current parameter tuning method can be improved with new parameter tuning algorithms (Smit and Eiben 2009). A multiobjective evolutionary algorithm is just one of many multiobjective optimization approaches. Tabu search, scatter search, ant system, and memetic algorithms may also be used to solve multiobjective optimization problems (Coello et al. 2002). More research on new multicriterion market segmentation methods and more empirical evaluation in different business settings are further topics to be investigated.

6.1. Caveats and Recommendations for Using Multicriterion Market Segmentation Methods

Multicriterion market segmentation approaches can be gainfully used for *descriptive* applications in joint segmentation settings with multiple batteries of variables. As observed by DeSarbo et al. (2005), combining multiple batteries of variables into a single set and applying *K*-means is not a viable surrogate for multicriterion segmentation methods. Merging all of the variable batteries together and running a *K*-means procedure can have at least two possible undesirable consequences: (1) the dominance of one or two batteries in the clustering solution at the expense of poor fit of the remaining batteries, and (2) a "washout" solution where a useful segmentation is not obtained for any of the batteries. Although both the transformation approach and MMSEA can be effectively used for descriptive joint segmentation applications, we propose that the latter approach has the unequivocal advantage of generating both unsupported and supported Pareto-optimal solutions, which results in a more thorough estimation of the frontier.

Following the recommendations of Brusco and Steinley (2009), we recommend that *predictive* versions of multicriterion market segmentation be used with extreme caution. In many instances, the solutions obtained in such applications do not stand up to cross validation. This occurs because the multicriterion clustering procedure capitalizes on chance variation in the model fitting process to improve the predictive fit. To illustrate, consider a situation where two customers (A and B) have exactly (or nearly) the same measurements on a set of descriptive attributes used for identifiability criterion but markedly different measurements on the dependent variable used for the responsiveness criterion. It is often the case that multicriterion procedures would place customers A and B in different segments to improve goodness of fit for the responsiveness criterion, despite the similarity of these two customers on the descriptive attributes. Now, suppose that we know the descriptive attribute measures for a new customer (customer C) and we want to establish a prediction for the dependent variable. Which within-segment regression model do we use for the prediction? Do we use the predictive model for the segment to which A was assigned, or the segment to which B was assigned? In effect, the problem that manifests itself in predictive applications is that *both* the descriptive attributes and the dependent variable are used in the model-fitting process that assigns customers to segments; however, the classification of new cases (or a holdout sample) would typically permit only the use of the descriptive information. After all, if both the descriptors and the dependent variable information were available for new cases, then there is no need to *predict* the dependent variable.

Appendix. MMSEA Algorithm

Step 1. **Initialize algorithm parameters**

Range_of_Number_of_Segments = {3, 4, 5}. *Minimum_Segment_Size* = 30.
Initial_Solution_Size = 200. *Child_Solution_Set_Size* = 40.
Archived_Solution_Set_Size = 200.
Crossover_Rate = 0.01 for evaluation one ***OR*** *Crossover_Rate* = 0.05 for evaluation two.
Number_of_Generations = 1. *Maximum_Number_of_Generations* = 600,000.
Set the initial solution set, the child solution set, and the archived solution set to an empty set.

Step 2. **Initialize solutions**

Calculate and sort the Euclidean distances between all pairs of customers using all customer attributes.
Randomly select a customer as the root of a minimum spanning tree.
Generate a minimum spanning tree in which the child node links (points) to its parent.
Link the root of the MST to its nearest neighbor to generate a MST-derived graph.

```
FOR each k in Range_of_Number_of_Segments,
  WHILE the size of the initial solution set <
    Initial_Solution_Size,
    FOR each descriptive segmentation basis,
      Use K-means to create a k-segment partition of
        customers using the segmentation basis.
    ENDFOR
    Use K-means to generate a k-segment partition of
      customers using all attributes.
    FOR each regression objective,
      Use clusterwise regression to generate a k-segment
        partition of customers.
    ENDFOR

    FOR each generated k-segment partition
      of customers,
      FOR each customer,
        Get the linked customer in the MST-derived
          graph.
        IF the linked customer is in the same segment
          as the customer,
          Set the customer's value in the chromosome
            to the index of the linked customer.
        ELSE
          Randomly select a customer that is in the
            same segment.
          Set the customer's value in the chromosome
            to the index of the selected customer.
        ENDIF
      ENDFOR
      Add the generated solution to the initial
        solution set.
    ENDFOR
  ENDWHILE
ENDFOR
```

Set the child solution set to the initial solution set.

Step 3. **Evaluate and archive child solution set**

FOR each solution in the child solution set,
Calculate the solution's objective values and compare the objective values with those of archived solutions.
Calculate the minimum Euclidean distance of objective values between the solution and the archived solutions.
IF the child solution is Pareto-optimal and its distance to the nearest archived solution is bigger than 0.001,
Add the solution to the archived solution set.
ENDIF
ENDFOR

CALL SUBROUTINE Calculate_Solution_Fitness.
WHILE the size of the archived solution set is > *Archived_Solution_Set_Size*,
Delete the solution that has the smallest fitness value from the archived solution.
CALL SUBROUTINE Calculate_Solution_Fitness.
ENDWHILE
IF *Number_of_Generations* > *Maximum_Number_of_Generations*
STOP.
ENDIF

Step 4. **Select parents by binary tournament and reproduce children by uniform crossover**
Clear the child solution set.

WHILE the size of the child solution set < *Child_Solution_Set_Size*,
Randomly select four solutions from the archived solution set and group them into two pairs.
FOR each pair of solutions,
Compare the fitness values of the two solutions.
Copy the solution with the bigger fitness value to create a child solution.
ENDFOR

FOR each element index of the child solution,
Generate a uniform random number between (0, 1), including 0 but not including 1.
IF the generated random number <= *Crossover_Rate*,
Swap the elements at the index of the two child solutions.
ENDIF
ENDFOR

FOR each child solution,
IF the number of segment of the solution is in *Range_of_Number_of_Segments*
AND the minimum segment size of the solution >= *Minimum_Segment_Size*,
Add the child solution to the child solution set.
ENDIF
ENDFOR
ENDWHILE
Increase *Number_of_Generations* by 1.
GOTO Step 3.

SUBROUTINE Calculate_Solution_Fitness
FOR each solution in the archived solution set,
Calculate m that is the rounded square root of the number of archived solutions of the same number of segments.
Calculate the average Euclidean distance between the solution and its m-nearest neighbors of the same number of segments.
Set the solution's fitness value to the calculated average Euclidean distance.
ENDFOR

References

Andrews, R. L., I. S. Currim. 2003a. A comparison of segment retention criteria for finite mixture logit models. *J. Marketing Res.* **40**(2) 235–243.

Andrews, R. L., I. S. Currim. 2003b. Retention of latent segments in regression-based marketing models. *Internat. J. Res. Marketing* **20**(4) 315–321.

Bock, T., M. Uncles. 2002. A taxonomy of differences between consumers for market segmentation. *Internat. J. Res. Marketing* **19**(3) 215–224.

Boulding, W., R. Staelin, M. Ehret, W. J. Johnston. 2005. A customer relationship management roadmap: What is known, potential pitfalls, and where to go. *J. Marketing* **69**(4) 155–166.

Brusco, M. J., D. Steinley. 2009. Cross validation issues in multiobjective clustering. *British J. Math. Stat. Psych.* **62**(2) 349–368.

Brusco, M. J., J. D. Cradit, S. Stahl. 2002. A simulated annealing heuristic for a bicriterion partitioning problem in market segmentation. *J. Marketing Res.* **39**(1) 99–109.

Brusco, M. J., J. D. Cradit, A. Tashchian. 2003. Multicriterion clusterwise regression for joint segmentation settings: An application to customer value. *J. Marketing Res.* **40**(2) 225–234.

Coello, C. A. C., D. A. V. Veldhuizen, G. B. Lamont. 2002. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Kluwer Academic Publishers, New York.

Dayton, C. M., G. B. MacReady. 1988. Concomitant variable latent class models. *J. Amer. Statist. Assoc.* **83**(401) 173–178.

Delattre, M., P. Hansen. 1980. Bicriterion cluster analysis. *IEEE Trans. Pattern Anal. Machine Intelligence* **2**(4) 277–291.

DeSarbo, W. S., C. F. DeSarbo. 2007. A generalized normative segmentation methodology employing conjoint analysis. A. Gustafsson, A. Herrmann, F. Huber, eds. *Conjoint Measurement: Methods and Applications*, 4th ed. Springer, Berlin, 321–345.

DeSarbo, W. S., D. Grisaffe. 1998. Combinatorial optimization approaches to constrained market segmentation: An application to industrial market segmentation. *Marketing Lett.* **9**(2) 115–134.

DeSarbo, W. S., C. A. Di Benedetto, M. Song, I. J. Sinha. 2005. Revisiting the miles and snow strategic framework: Uncovering interrelationships between strategic types, capabilities, environmental uncertainty, and firm performance. *Strategic Management J.* **26**(1) 47–74.

Dimitriadou, E., S. Dolničar, A. Weingessel. 2002. An examination of indexes for determining the number of clusters in binary data sets. *Psychometrika* **67**(1) 137–159.

Ehrgott, M., X. Gandibleux. 2000. A survey and annotated bibliography of multiobjective combinatorial optimization. *OR Spektrum* **22**(4) 425–460.

Ferligoj, A., V. Batagelj. 1992. Direct multicriteria clustering algorithms. *J. Classification* **9**(1) 43–61.

Frank, R. E., W. F. Massey, Y. Wind. 1972. *Market Segmentation*. Prentice Hall, Englewood Cliffs, NJ.

Freitas, A. A. 2004. A critical review of multi-objective optimization in data mining: A position paper. *ACM SIGKDD Explorations Newslett.* **6**(2) 77–86.

Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Professional, Reading, MA.

Grün, B., F. Leisch. 2007. Fitting finite mixtures of generalized linear regression in *R*. *Comput. Statist. Data Anal.* **51**(11) 5247–5252.

Gupta, S., P. K. Chintagunta. 1994. On using demographic variables to determine segment membership in logit mixture models. *J. Marketing Res.* **31**(1) 128–136.

Handl, J., J. Knowles. 2005. Improvements to the scalability of multiobjective clustering. *Proc. 2005 IEEE Congress Evolutionary Comput. (CEC 2005)*, IEEE Press, Los Alamitos, CA, 2372–2379.

Handl, J., J. Knowles. 2006. Multi-objective clustering and cluster validation. Y. Jin, ed. *Multi-Objective Machine Learning*, Vol. 16. Springer, Berlin, 21–47.

Handl, J., J. Knowles. 2007. An evolutionary approach to multiobjective clustering. *IEEE Trans. Evolutionary Comput.* **11**(1) 56–76.

Hubert, L., P. Arabie. 1985. Comparing partitions. *J. Classification* **2**(1) 193–218.

Jiang, T., A. Tuzhilin. 2006. Segmenting customers from population to individuals: Does 1-to-1 keep your customers forever? *IEEE Trans. Knowledge Data Engrg.* **18**(10) 1297–1311.

Krieger, A. M., P. E. Green. 1996. Modifying cluster-based segments to enhance agreement with an exogenous response variable. *J. Marketing Res.* **33**(3) 351–363.

Law, M. H. C., A. P. Topchy, A. K. Jain. 2004. Multiobjective data clustering. *Proc. 2004 IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognition (CVPR'04)*, IEEE Press, Los Alamitos, CA, 424–430.

Milligan, G. W., M. C. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* **50**(2) 159–179.

Milligan, G. W., M. C. Cooper. 1988. A study of standardization of variables in cluster analysis. *J. Classification* **5**(2) 181–204.

Myers, J. H. 1996. *Segmentation and Positioning for Strategic Marketing Decisions*. American Marketing Association, Chicago.

Park, Y.-J., M.-S. Song. 1998. A genetic algorithm for clustering problems. *Proc. 3rd Annual Conf. Genetic Programming*, Morgan Kaufmann, San Francisco, 568–575.

Ramaswamy, V., R. Chatterjee, S. H. Cohen. 1996. Joint segmentation on distinct interdependent bases with categorical data. *J. Marketing Res.* **33**(3) 337–350.

Ripon, K. S. N., C.-H. Tsang, S. Kwong, M.-K. Ip. 2006. Multiobjective evolutionary clustering using variable-length real jumping genes genetic algorithm. *Proc. 18th Internat. Conf. Pattern Recognition (ICPR'06)*, IEEE Computer Society, Washington, DC, 1200–1203.

Smit, S. K., A. E. Eiben. 2009. Comparing parameter tuning methods for evolutionary algorithms. *Proc. 2009 IEEE Congress Evolutionary Computation (CEC 2009)*, IEEE Press, Los Alamitos, CA, 399–406.

Smith, W. R. 1956. Product differentiation and market segmentation as alternative marketing strategies. *J. Marketing* **21**(1) 3–8.

Steinley, D. 2004. Properties of the Hubert-Arabie adjusted Rand index. *Psych. Methods* **9**(3) 386–396.

Steinley, D., M. J. Brusco. 2008a. A new variable weighting and selection procedure for *k*-means cluster analysis. *Multivariate Behav. Res.* **43**(1) 77–108.

Steinley, D., M. J. Brusco. 2008b. Selection of variables in cluster analysis: An empirical comparison of eight procedures. *Psychometrika* **73**(1) 125–144.

Ulungu, E. L., J. Teghem. 1994. Multi-objective combinatorial optimization: A survey. *J. Multicriteria Decision Anal.* **3**(2) 83–104.

Vargo, S. L., R. F. Lusch. 2004. Evolving to a new dominant logic for marketing. *J. Marketing* **68**(1) 1–17.

Vose, M. D., A. H. Wright. 2001. Form invariance and implicit parallelism. *Evolutionary Comput.* **9**(3) 355–370.

Vriens, M., M. Wedel, T. Wilms. 1996. Metric conjoint segmentation methods: A Monte Carlo comparison. *J. Marketing Res.* **33**(1) 73–85.

Wedel, M., W. Kamakura. 2000. *Market Segmentation: Conceptual and Methodological Foundations*. Kluwer Academic Publishers, Norwell, MA.

Wind, Y. 1978. Issues and advances in segmentation research. *J. Marketing Res.* **15**(3) 317–337.