Multiobjective Evolutionary Computation for Market Segmentation

Ying Liu[a]
College of Business, California State University Long Beach, Long Beach, CA, U.S.A.

Keywords: Market Segmentation, Multiobjective Evolutionary Computation, Pareto Optimal.

Abstract: The market segmentation, by its computational essence, is a NP-hard multicriteria problem. Multiobjective evolutionary algorithms are developed to optimize multiple objectives simultaneously and can generate a set of Pareto optimal solutions. As a proven meta-heuristic technique, multiobjective evolutionary computation is robust in handling different data types, various business constraints and different objective function forms. The generated Pareto optimal solution set gives a holistic view of possible solutions that bring business insights and allow big flexibility in solution selection. These features make the multiobjective evolution computation a good fit for market segmentation problems. There are challenges in every phase in implementation of multiobjective evolutionary computation for market segmentation.

1 INTRODUCTION

Market segmentation provides business decision makers a useful perspective to understand and differentiate customers by their needs and behaviors (Dolnicar et al., 2018). Though market segmentation has had an easy-to-understand conceptual definition (Smith, 1956) for more than a half century, Wedel and Kamakura (2000) observed that "the development of market segmentation method has been partly contingent on the availability of marketing data, the advances of analytical techniques and the progress of segmentation methodology." As the size and richness of customer data increase, academicians and practitioners demand more efficient, robust, and scalable techniques to meet the new challenges in customer segmentation. The recent advances in machines learning and distributed data processing bring new capabilities and new methods to segment customers.

Wedel and Kamakura (2000) classified market segmentation methods into predictive and descriptive methods. Descriptive methods use two or more sets of variables to describe the customer segments while predictive methods analyze the relationship between a set of independent variables and one or more dependent variables. This classification is helpful but gives few clues of the complexity and the great variance of segmentation methods in problem definition and solution implementation. There is an abundance of different market segmentation methods. However, there is relatively little work done on the computational issues of the market segmentation, especially its multicriteria nature, to help people understand existing methods and develop new segmentation methods. This research investigates the computational properties of the market segmentation problem in section 2. Section 3 shows the complexity and the multicreteria nature of the market segmentation problem. Section 4 shows that the multiobjective evolutionary computation is a good candidate to solve the market segmentation problems. The conclusions and future research directions are discussed in Section 5.

2 THE COMPUTATIONAL VIEW OF MARKET SEGMENTATION

2.1 Clustering Is a Subproblem

The challenges of market segmentation roots in its computational properties. A fundamental task of market segmentation is grouping customers based on similarities in their needs and preferences. Clustering is a common tool for this purpose (Punj and Stewart, 1983). Clustering could be generally defined as a set

[a] https://orcid.org/0000-0002-3419-0733

Liu, Y.
Multiobjective Evolutionary Computation for Market Segmentation.
DOI: 10.5220/0010684400003063
In *Proceedings of the 13th International Joint Conference on Computational Intelligence (IJCCI 2021)*, pages 149-154
ISBN: 978-989-758-534-0; ISSN: 2184-3236

of techniques that group entities that are similar in measured characteristics. It aims to maximize the homogeneity within the segment while maximizing the heterogeneity among segments. Each segment is a class of customers that marketer can identify, target, and communicate with. In the early market segmentation research, due to the limits of customer data and the lack of computational power, clustering was used widely market segmentation. However, cluster analysis is a solution only good for single-basis descriptive segmentation. A segmentation basis is a set of variables used to describe a certain aspect of customers. Different segmentation bases describe different features of the customer or marketing mix and have different levels of effectiveness regarding the segmentation criteria. In a general model of descriptive market, more than one segmentation bases are used to take advantage of the benefits of each segmentation basis. The model is called joint descriptive market segmentation (Morwitz and Schmittlein, 1992). From the view of segmentation objectives, descriptive methods are optimized for segment identifiability while predictive methods are optimized for segment responsiveness. In predictive segmentation, decision makers do not segment customers for “clustering” purposes only. They want actionable segments that will let them formulate effective marketing campaign in an objective way. As a result, in both descriptive and predictive market segmentation, clustering is a subproblem of a general model of market segmentation.

2.2 Computation Definition

Though the conceptual definition of market segmentation is simple, the computational definitions of market segmentation have been in many forms. A common practice is to define the segmentation problem according to different segmentation solution techniques. Market segmentation was framed as a clustering problem (Punj and Stewart, 1983) in early research when the clustering techniques were used. When the focus was shifted from descriptive variables to response variables, the market segmentation was framed as a segmentation problem solved by solution procedures such as chi-squared automatic interaction detector (CHAID) (Kass, 1980), classification and regression trees (CART) (Breiman et al., 1984), and clusterwise regress (Spath, 1982). Unlike the clustering definition that aims to maximize within-segment homogeneity, the segment problem (Kleinberg et al., 2004) aims to maximize a general utility function (usually not the segment homogeneity to distinguish it from the clustering problem) of all segments.

The identifiability, responsiveness and other criteria of market segmentation demonstrate the characteristics of both clustering and segmentation. But market segmentation problems did not have a computational multiobjective definition until DeSarbo and Grisaffe (1998) used combinatorial optimization approaches as the solution techniques. Since then, many multiobjective optimization approaches (Krieger and Green, 1996, Brusco et al., 2003) were developed to solve the multicriteria market segmentation problems. Though DeSarbo and Grisaffe (1998) pointed out that there exists a set of Pareto optimal solutions for a multiobjective problem definition of market segmentation, those methods do not generate the Pareto optimal solution set because they are essentially single objective solution techniques. Giving that evolutionary algorithm generate good results for many multiobjective optimization problems (Coello et al., 2002), Liu et al., (2010) applied a multiobjective evolutionary algorithm to market segmentation. The algorithm directly tackles the multiobjective segmentation problem and generates a Pareto optimal solution set.

2.3 Discriminative vs. Generative Methods

Assumptions about the segmentation data play an important role in segmentation methods. Like the clustering method classification scheme proposed by Zhong and Ghosh (2003), market segmentation methods can be classified into discriminative (or distance/similarity-based) approaches and generative (or model-based) approaches from their computational assumption about the nature of data. Discriminative methods calculate distances or similarity between customers and segment customers based on these measures. K-means, hierarchical clustering, and Self-Organizing Map are typical discriminative clustering methods. Generative methods assume customers are from different statistical models and try to find the parameters of the corresponding models. Each type has its advantages and disadvantages. Because of the direct optimization of within-segment customer similarity, discriminative-based segmentation methods are usually efficient and intuitive. However, the results usually are used as-is and no statistical inference could be drawn from the results. There are several advantages of generative methods. If the distribution assumption of data is correct, they usually generate better results than discriminative methods. The results are more interpretable and enable statistic inference. But generative methods such as finite mixture model are computationally expensive when the number-of-segments is big or there are many segmentation variables (Wedel and Kamakura, 2000). Both the discriminative and generative approaches may be formulated as a data mining and/or optimization problem whose solutions need to have confidence and support, information content and unexpectedness (Padmanabhan and Tuzhilin, 2003).

3 THE COMPUTATIONAL ISSUES OF MARKET SEGMENTATION

3.1 The Similarity Measures and Clustering Process

Clustering is a subproblem of market segmentation. Clustering by itself is vaguely defined and the clustering process is hard and fuzzy (Jain et al., 1999). The vagueness lies in the measurement of so-called “homogeneity” or “similarity”. Punj and Stewart (1983) discussed problems of determining the appropriate similarity measure in market segmentation. From several empirical experiments, they found that each similarity measure has different characteristics and different distance measures lead to different clustering results. Skinner (1978) identified three aspects of similarity measures: elevation, scatter, and shape. In a rough sense, elevation could be thought of as the mean of all attributes of a given subject. Scatter is about deviation, while shape is about the direction (up/down) of the data. The most important finding of was that a specific distance or similarity measure may not cover all aspects.

The hardness and fuzziness of clustering process is explained by the impossibility theory of clustering proved by Kleinberg (2002). It is intuitive to think of three desired properties of any clustering process. Scale-invariance property means that changing the unit of distance measure should not change the clustering result. Richness requires that a clustering process should be able to generate all possible partitions of clustering entities. Finally, consistency is satisfied when the clustering result stays the same when we increase the distance among clusters and decrease distances within clusters. The impossibility of clustering showed that there is no clustering process that can satisfy all three properties simultaneously. To avoid the limitation of clustering process, Penaloza et al., (2017) developed a multi-objective clustering algorithm that uses multiple criteria to measure the quality of cluster cohesion.

Zhong and Ghosh (2003) proposed to classify the clustering method into discriminative methods and generative methods. In discriminative methods, also called distance/similarity-based methods, the similarity function is defined between pairs of objects. In generative methods, also called model-based methods, the similarity is defined indirectly through the assumption of data distribution. Those methods assume that the overall distribution of the data is a mixture of probability distributions, each being a different cluster (Fraley and Raftery 1998). Even the probabilistic clustering methods assume similarity measure, though in an indirect way. The distinction of discriminative and generative methods helps to understand the similarity measures among clustering algorithms.

3.2 The Computational Complexity

The trend of big data and quick market response time raise attentions to computational complexity of market segmentation. Aloise et al. (2009) showed that clustering is NP-hard even for the 2-cluster problem using the very simple Euclidean distance to measure similarity. Kleinberg et al., (1998) proved that most optimization problems become NP-complete if they are defined in a segmentation form. Krieger and Green (1996) defined the market segmentation problem as a 0-1 programming problem whose computational complexity is NP-hard. Consequently, the market segmentation problem, even framed as a clustering problem, cannot be solved in polynomial time. Existing methods either transform the problem into an easy to solve version or apply heuristic techniques to solve the problem.

3.3 The Multicriteria Nature

Marketing researchers realized that market segmentation is a multicriteria problem from the very beginning because customers in a segment should have similar profiles (identifiability) and respond similarly to a marketing mix (responsiveness) (Smith, 1956). For example, customers in a segment should have similar demographic attributes such age, educational level, location, etc. Identifiability makes it easy to target a specific customer segment. Responsiveness can be measured by customer behaviors such as response rate or transactional values of a marketing promotion. Simultaneously clustering customers and predicting their responses to marketing mix is a long-standing problem facing marketing researchers. During the evolution of market segmentation theories, more and more criteria are added. In addition to the identifiability and responsiveness, Wedel and Kamakura (2000) added substantiality, accessibility, stability and actionability criteria to evaluate whether a segmentation solution is good or not. DeSarbo and DeSarbo (2009) added four more criteria of differential behavior, feasibility, profitability and projectability. At the general conceptual level, clustering only addresses the identifiability criterion (Brusco et al., 2003) while other criteria such as responsiveness, profitability and actionability must be addressed by augmented methods.

3.4 Determining the Number-of-Segments

The issue of determining the number-of-segments appears in both predictive segmentation and joint segmentation. In predictive segmentation, the criterion of the predictive power is as important as the criterion of the segment homogeneity. As the number-of-segments increases, the within segment homogeneity usually increases but the predictive power may increase or decrease independent of the within-segment homogeneity. Joint segmentation consists of clustering on multiple segmentation bases, in that, each can be thought of as an independent clustering problem and an overall trade-off must be made in selecting the "right" number-of-segments. The multicriteria nature of market segmentation means that determining the "right" number-of-segments is a multicriteria decision and often involves marketers' domain knowledge. Consequently, decision makers would like to see a set of segmentation solutions that have different numbers-of-segments. Those solutions give them flexibility in investigating solutions and select the most appropriate ones for a specific business scenario.

4 MULTIOBJECTIVE EVOLUTIONARY COMPUTATION

4.1 Why Multiobjective Evolutionary Computation?

The multicriteria nature of market segmentation raises many issues that cannot be addressed appropriately by traditional market segmentation methods such as K-means and cluster-wise regression because they only optimize one objective. Many heuristic multiobjective methods have been developed to address the multicriteria requirement of market segmentation. These methods can be classified in three categories: multi-stage method, transformation method, and multiobjective method.

The multi-stage method solves one criterion at one stage. For example, Kriger and Green (1996) used K-means method to optimize group identifiability in stage one and a heuristic algorithm to optimize responsiveness of segments in stage two. The disadvantage of the multi-stage approach is that information found in one stage is not used by the other stages because of the separated processing phases. It is not efficient in the sense of information sharing. The order of objective optimization often matters. Furthermore, because each stage optimizes a single objective, the result is often suboptimal regarding all objectives.

The transformation method transforms multiple criteria into one (Green and Krieger, 1991, Brusco et al., 2003), therefore the problem can be solved by many established single objective optimization methods. However, it is often difficult, if not impossible, to define an appropriate total utility or weighted sum function to represent the multiple criteria. Multiple criteria may be incommensurate. For example, one criterion is the within-segment homogeneity measure by within-segment sum of variance and another criterion is the predictive power in logistic regression measure by maximum likelihood. Additionally, the transformation procedure may put unnecessary limitations on the search space. Global optimal solution could be lost in transformation (Freitas, 2004).

Multiobjective evolutionary algorithms such as NGSA II, SPEA2, and FAME (Santiago et al., 2019) optimize multiple objectives simultaneously and generate a set of Pareto optimal solutions. These methods have some much-desired features. First, multiple criteria can be independently defined in terms of multiple optimization objects, computational constraints, and decision variables. This avoids the difficulties of combining multiple criteria. The multiple objectives optimization can incorporate both generative and discriminative measures. Second, a multiobjective optimization method generates a set of Pareto optimal solutions representing trade-offs among multiple possibly conflicting objectives. Third, a single run can generate solutions with different number-of-segments. There is no upfront need to determine the number-of-segments.

4.2 A Good Fit for Market Segmentation

The evolutionary computation is the most widely used meta-heuristic approach to solve multiobjective optimization problems (Coello et al., 2002). It is a good fit for market segmentation problems for several reasons. First, it searches for optimal solution(s) using a set of objectives simultaneously. This property makes the algorithm efficient because all objectives are directly used in optimization. Second, the evolutionary algorithm can be used to find a set of solutions that has the desired diversity. The results are usually representative of the possible trade-offs that are important for decision making.

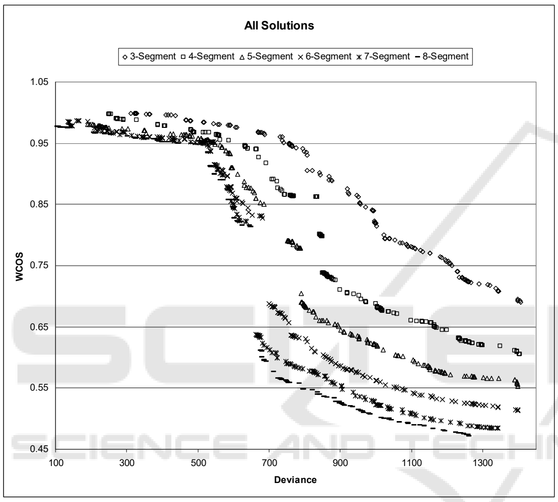

Figure 1: A sample solution set that has different number-of-segments.

The solutions in Figure 1 are generated from a multiobjective genetic algorithm (MOEA) applied in consumer data that include both demographic and transactional attributes. The two minimum optimization objectives are deviance of linear regression and WCOS (with cluster omega squared) for within cluster homogeneity. with the aid of data visualization and analysis tools, the shape and the parameters of the solution set give insights and improve the decision-making process.

Additionally, evolution computation is less susceptible to the shape of continuity of the Pareto solution set because it does not make any assumptions about data and problem properties. Finally, the evolutionary algorithm is independent of objective functions and decision variables. This is a very attractive feature because the algorithm could be used in a broad range of market segmentation problems in different objective function forms (discrete or continuous, concave or convex, single modal or multimodal).

4.3 Computational Issues

Nonetheless, the application of multiobjective evolutionary computation in market segmentation is relatively new (O'Brien et al., 2020) and it brings some challenges in all computational phases

The quality and diversity of the initial solution set affects the effectiveness of evolutionary computation. For market segmentation, existing clustering and segmentation algorithms optimized for different segmentation objectives can be used to generate the initial solution set. There are not many theories to guide the implementation and selection of initialization algorithms.

The parameter setting in evolutionary algorithms is another challenging task since the parameters interact in highly non-linear ways (Lobo et al., 2007). Grid search is often used but Bayesian optimization can reach or surpass human expert-level optimization on many algorithms such as structured SVM and coevolutionary neural network (Snoek et al., 2012). It is interesting to check its effectiveness in evolutionary algorithm in market segmentation.

Selecting the most appropriate solution from a set of Pareto optimal solutions consisting of solutions with different number-of-segments is another interesting research topic. The shape of the Pareto front, the parameters of each solution, and the practical constraints are factors to be considered. Visualization tools and multiobjective data analysis techniques are helpful in determining the number-of-segments and solution selection.

5 CONCLUSIONS

There is a good match between the computational essence of market segmentation problem and the multiobjective evolutionary computation. Multiobjective evolutionary computation brings a new perspective to multicriteria market segmentation in its computational model definition, optimization process and the solution set analysis. It comes with challenges in all phases of evolutionary computation. Given the meta-heuristic nature of the multiobjective evolutionary computation, it is a research topic to use new algorithms in solution initialization, parameter setting and solution selection. There is a need for more empirical evaluation in different business settings.

REFERENCES

Aloise, D., Deshpande, A., Hansen, P. & Popat, P. (2009). NP-hardness of Euclidean sum-of-squares clustering. *Machine Learning*, 75(2), 245-248.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. I. (1984). *Classification and Regression Trees.* Wadsworth International Group.

Brusco, M. J., Cradit, J. D., & Tashchian, A. (2003). Multicriterion Clusterwise Regression for Joint Segmentation Settings: An Application to Customer Value. *Journal of Marketing Research*, 40(2), 225-234.

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

DeSarbo, W. S., & Desarbo, C. F. (2007). A Generalized Normative Segmentation Methodology Emply Conjoint Analysis. *Conjoint Measurement: Methods and Applications,* 4th ed, Springer Berlin.

DeSarbo, W. S., & Grisaffe, D. (1998). Combinatorial Optimization Approaches to Constrained Market Segmentation: An Application to Industrial Market Segmentation. *Marketing Letters,* 9(2), 115-134.

Dolnicar, S., Grun, B., & Leisch, F. (2018). *Market Segmentation Analysis*. Springer.

Fraley, C., & Raftery, A. E. (1998). How Many Clusters? Which Clustering Methods? Answers Via Model-Based Cluster Analysis. *The Computer Journal,* 41(8): 578-588.

Freitas, A. (2004). A Critical Review of Multi-Objective Optimization in Data Mining: a Position Paper. *SIGKDD Explorations,* 6(2), 77-86.

Green, P. E., & Krieger, A. M. (1991). Segmenting Markets with Conjoint Analysis. *Journal of Marketing,* 55(4), 20-31.

Jain, A. K., Murthy, M. N., & Flynn, P. J. (1999). Data Clustering: A review. *ACM Computing Surveys,* 31(3), 264-323.

Kass, G. V. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics,* 29(2), 119-127.

Kleinberg, J. (2002). An Impossibility Theorem for Clustering. *Proceedings of the 15th Conference on Neural Information Processing Systems (NIPS). Vancouver, Canada,* 463-470.

Kleinberg, J., Papadimitriou, C., & Raghavan, P. (2004). Segmentation Problems. *Journal of the ACM,* 51(2), 263-280.

Krieger, A. M., & Green, P. E. (1996). Modifying Cluster-Based Segments to Enhance Agreement with an Exogenous Response Variable. *Journal of Marketing Research,* 33, 351-363.

Liu, Y, Ram, S., Lusch, R. F., & Brusco, M. (2010). Multicriterion Market Segmentation: A New Model, Implementation, and Evaluation. *Marketing Science*, 29(5), 880-894.

Lobo, F. G., Lobo, F. J., & Lima, C. F., Michalewicz, Z. (2007). *Parameter Settings in Evolutionary Algorithms*. Springer.

Morwitz, V. G., & Schmittlein, D. W. (1992). Using Segmentation to Improve Sales Forecasts Based on Purchase Intent: Which Intenders Actually Busy?. *Journal of Marketing Research*, 29, 391-405.

O'Brien, M., Liu, Y., Chen, H., & Lusch. R. (2020). Gaining insight to B2B Relationships Through New Segmentation Approaches: Not All Relationships Are Equal. *Expert Systems with Applications,* 161.

Padmanabhan, B., & Tuzhilin, A. (2003). On the Use of Optimization for Data Mining: Theoretical Interactions and eCRM Opportunities. *Management Science*, 49(10), 1327-1343.

Penaloza, M. G. M., Montes, E. M., Ramirez, N. C., Mesa, H. G. A. M., & Figueroa, H. V. R. (2017). Improve Multi-objective clustering with automatic determination of the number of clusters. *Neural Computing and Application,* 28(8), 2255-2275.

Punj, G., & Stewart, D. W. (1983). Cluster Analysis in Marketing Research: Review and Suggestion for Application. *Journal of Marketing Research,* 20, 134-148.

Santiago, A., Dorronsoro, B., Nebro, A. J., Durillo, J. J., Castillo, O., Fraire, H. J. (2019). A Novel Multi-objective Evolutionary Algorithm with Fuzzy Logic Based Adaptive Selection of Operators: FAME. *Information Science,* 471, 233-251.

Skinner, H. A. (1978). Differentiating the Contribution of Elevation, Scatter and Shape in Profile Similarity. *Educational and Psychological Measurement,* 38, 297-308.

Smith, W. R. (1956). Product Differentiation and Market Segmentation as Alternative Marketing Strategies. *Journal of Marketing,* 21, 3-8.

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems. 25,* 2960-2968.

Spath, H. (1982). A Fast Algorithm for Clusterwise Linear Regression. *Computing, 29,* 175-181.

Wedel, M., & Kamakura, W. (2000). *Market Segmentation: Conceptual and Methodological Foundations.* Kluwer Academic Publisher.

Zhong, S., & Ghosh, J. (2003). A Unified Framework for Model-based Clustering. *Journal of Machine Learning Research,* 4, 1001-1037.