MARMARA UNIVERSITY

INSTITUTE FOR GRADUATE STUDIES
IN PURE AND APPLIED SCIENCES

HEURISTIC SOLUTION TO THE PRODUCT TARGETING PROBLEM BASED ON MATHEMATICAL PROGRAMMING

FİLİZ ÇETİN

(724804984)

Ph.D. THESIS

Department of Engineering Management

Thesis Supervisor

Assoc. Prof. Dr. Çiğdem Alabaş USLU

ISTANBUL, 2016

MARMARA UNIVERSITY
INSTITUTE FOR GRADUATE STUDIES
IN PURE AND APPLIED SCIENCES

Filiz ÇETİN, a Doctor of Philosophy student of Marmara University Institute for Graduate Studies in Pure and Applied Sciences, defended her thesis entitled "Heuristic solution to the product targeting problem based on mathematical programming", on April 5, 2016 and has been found to be satisfactory by the jury members.

Jury Members

Assoc.Prof. Dr. Çiğdem Alabaş USLU (Advisor)
Marmara University

Assoc.Prof. Dr. Serol BULKAN (Jury Member)
Marmara University

Prof. Dr. Umut Rıfat TUZKAYA (Jury Member)
Yıldız Technical University

Asst. Prof. Dr. Adnan ÇORUM (Jury Member)
Bahçeşehir University

Asst. Prof. Dr. Hüseyin Selçuk KILIÇ (Jury Member)
Marmara University

APPROVAL

Marmara University Institute for Graduate Studies in Pure and Applied Sciences Executive Committee approves that Filiz ÇETİN be granted the degree of Doctor of Philosophy in Department of Engineering, Engineering Management Program on 15.04.2016 (Resolution no: 2016/10-03).

Director of the Institute

Prof. Dr. Uğur YAHŞI

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENT

I would like to express acknowledgment and appreciation to my dissertation advisory committee members, Assoc.Prof. Dr. Çiğdem Alabaş USLU, Assoc. Prof. Dr. Serol BULKAN and Prof. Dr. Umut Rıfat TUZKAYA for their assistance, direction and dedication throughout my course of study. Acknowledgment and appreciation is also due to Prof. Dr. M. Akif EYLER for his support and guidance.

It would not have been possible to write this doctoral thesis without the help of lovely people around me. I'd like to thank everyone in my beloved family, who's continuously supported me in this journey, and played with my son, Atlas, during the last few months.

I'd also like to thank to my husband, Görkem, for his continuous support. Without him, it'd have certainly been much harder for me to finish my thesis. He has made available his support in a number of ways.

April, 2016 **Filiz ÇETİN**

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

ÖZET

HEURISTIC SOLUTION TO THE PRODUCT TARGETING PROBLEM BASED ON MATHEMATICAL PROGAMMING

Müşterinin yaşam döngüsünü yönetmek firmalar için oldukça önemli bir süreç haline gelmiştir. Bu sürecin yönetilmesi noktasındaki bir strateji de farklı promosyon kampanyaları önermektir. Bu kampanyaların oluşturulmasında cevabı bulunması gereken en önemli soru ise "Karlılığı artırmak için hangi müşteriye, hangi hedefli ürünlerin pazarlaması yapılmalıdır?" olmaktadır.

Bu çalışma, yukarıda bahsi geçen soruya yanıt aramak için yapılmıştır. Bu problem özellikle sınırlı bütçe ve asgari bir satış hedefinde daha da önem arzetmektedir. Problemin NP-zorluğundan dolayı yöneylem araştırması açısından da dikkate değerdir. Bunun için ürün hedefleme problemine sezgisel yaklaşımda bulunulmuş ve matematiksel programlama önerilmiştir. Önerilen yaklaşım problemi iki aşamada çözmektedir: Birincisi, sezgisel kurallarla hangi ürünlerin kampanyaya dahil edileceğinin belirlenmesi, ikincisi ise bu ürünlerin müşterilere optimum şekilde dağıtılmasıdır. Bunlara ek olarak, tabu arama algoritması da probleme uygulanmıştır. Tabu aramanın başlangıç çözümü, bu çalışmada önerilen matematiksel programlama temelli yaklaşımlardan alınmıştır. Tabu arama algoritması kullanılmasının ana sebebi ve motivasyon kaynağı da daha iyi kar maksimizasyonu aramaktır.

Önerilen yaklaşımlarda problemin iki alt probleme bölünmesinin temel faydası, büyük boyutlu problemlerin etkin ve verimli bir şekilde çözülebilir hale getirilmesidir. Tüm önerilen sezgisel ve tabu arama algoritmaları öncelikle literatürdeki veri kümeleri üzerinde test edilmiştir. Daha sonra çok büyük boyutlu problemleri çözme kabiliyetini göstermek için yeni test problemleri oluşturulmuş ve gerçek yaşam problemlerine uygulanabilirliliği gösterilmiştir. Yapılan deneysel çalışmalarla, önerilen yaklaşımların literatürdeki mevcut yaklaşımlara göre de daha üstün sonuçlar elde ettiği gösterilmiştir. Sektörel açıdan bakıldığında bu sezgisel yöntemlerin firmaların optimum karını artırmaya yönelik güçlü bir araç olduğu görülebilir.

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

HEURISTIC SOLUTION TO THE PRODUCT TARGETING PROBLEM BASED ON MATHEMATICAL PROGAMMING

Maintaining customer lifetime longevity is a crucial issue for companies. One of the strategies for dealing with this issue is to offer different promotion campaigns. Planning these campaigns creates a problem: Which targeted products in the campaign should be offered to which customers in order to maximize profit? This problem becomes vitally important under the conditions of a limited budget and a lower bound on sales target of each product. It is also remarkable from the operational research perspective because of its NP-hardness.

This study mainly investigates solutions to these questions. For this purpose, heuristic approaches to the product targeting problem based on mathematical programming are suggested. The proposed approaches solve the problem in two parts: first, determine the products to be included in a campaign using heuristic rules and second, distribute these products to the customers optimally. Moreover, a tabu search algorithm is also applied to the problem. The initial solution of the tabu search is taken from the results of the mathematical programming based approaches proposed in this study. The main motivation of using tabu search algorithm is to find better profits.

Main advantage of the proposed approaches by dividing the problem into two sub-problems is to make very large-sized instances solvable effectively and efficiently. All the suggested heuristics and tabu search algorithm are firstly tested on the data sets from the literature. Then, new test problems are generated to show the capability of solving very large sized problems and their potential for practical applications is verified. Computational results also confirm that these approaches generate superior solutions to the problem in comparison with existing methods in the literature. From the business perspective, the heuristics proposed in this study can be viewed as a strong tool to increase optimal profit of the firms.

THIS PAGE INTENTIONALLY LEFT BLANK

CLAIM FOR ORIGINALITY

Maintaining customer lifetime longevity is a crucial issue for companies. One of the strategies for dealing with this issue is to offer different promotion campaigns. Planning these campaigns creates a problem: Which targeted products in the campaign should be offered to which customers in order to maximize profit? This problem becomes vitally important under the business conditions. It is also remarkable from the operational research perspective because of its NP-hardness. The objective of the problem is to maximize profit yielded by offering products to customers subject to a set of constraints such as budget, the upper limit of total products offered to each customer and the lower limit of the total amount of each product offered which makes the involvement of the product in the campaign possible. For this purpose, heuristic approaches to the product targeting problem based on mathematical programming are suggested. The proposed approaches solve the problem in two parts: first, determine the products to be included in a campaign using heuristic rules and second, distribute these products to the customers optimally. Moreover, a tabu search algorithm is also applied to the problem. The main motivation of using tabu search algorithm is to find better profits. All heuristics and tabu search algorithm are tested on the data sets from the literature. Computational results confirm that these approaches generate superior solutions to the problem in comparison with existing methods in the literature. The contributions of this study can be summarized as:

- Development of new heuristic approaches to deal with an approximate solution of the problem. The heuristic methods proposed in this study divide the problem into two sub-problems to make very large-sized instances solvable effectively and efficiently.
- Suggestion of a linear mathematical model which is a modification of the model existing in the literature. The modified model can be utilized to capture a strong approximation to optimal profit of the problem.

April, 2016 **Assoc.Prof.Dr. Çiğdem Alabaş Uslu** **Filiz ÇETİN**

THIS PAGE INTENTIONALLY LEFT BLANK

ABBREVIATIONS

PTP: Product targeting problem

B2B: Business to business

RFM: Recency, frequency, monetary

MQC: Minimum quantity commitment

SMS: Short message service

CLV: Customer lifetime value

H-R1: Heuristic rule 1

H-R2: Heuristic rule 2

TS: Tabu Search

H-TS: Heuristic tabu search

ANOVA: Analysis of variation

B&P: Branch-and-price

H-TS-rule1: Heuristic of tabu search with heuristic rule1 initials

H-TS-rule2: Heuristic of tabu search with heuristic rule2 initials

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

1. INTRODUCTION

The connection between a company and its customers is not a strong bond. Marketing managers focus on not only acquiring but also retaining customers. The capturing of existing customers to gain additional revenue is known as retention of customers. Reichheld and Sasser (1990) point out that retention of customers also reduces costs and can be source of high market share. Their study shows that if a company retains 5% more of its customers, this boosts the profit from 25% to 85%.

One of the strategies for retaining customers is to offer different promotion campaigns. To practically use, there are two approaches for promotions. One is mass marketing, other is direct marketing. Mass marketing focuses on large group of customer without any need of differentiation among them, while direct marketing, which is getting more importance from marketers, targets individuals and specific class of customer (Bose and Chen, 2009).

Direct marketing is an effective and measurable method of marketing (Wong et al., 2005). Bose &Chen reported that for 2005, it is expected that, on an average, $1 spent on consumer direct marketing would yield $12.66 compared to $10.10 for B2B (business-to-business) direct marketing. Offering promotion campaigns to retain the customers is the main activity of direct marketing. In other words, the goal of the promotion campaign is to keep existing customer in touch. On the other hand, the company has imbalanced data stored in customer databases and the correlation between the probability that a customer responds and the dollar amount generated by a response can be negative (Wong et al., 2005). Moreover, various business requirements must be met in order to be successful in offering promotion campaign. This makes the solution harder to implement.

Planning these campaigns exposes a problem called the product targeting problem (PTP). The product targeting problem can be described as the selection of products to be included in a campaign and the distribution of these products to customers in order to maximize profit while satisfying various business constraints. There are variations on the product targeting problem including different sets of constraints, objectives and decision problems in the literature. Customer selection, product selection or both are the main focus of the problem. Since there are differences on decision problems, solution methodologies seem to be different such as data analysis, optimization techniques or both. In the literature, there are a few studies which are closely related with the problem addressed in this study. Recently, Nobibon et al. (2011) directly address the problem and propose both optimization models and heuristic algorithms.

In this study, the optimization model by Nobibon et al. (2011) which is a basic formulation of the problem is modified in pursuit of the development of heuristic methods for targeted offers in a promotion campaign. The objective of the problem is to maximize profit yield by offering products to customers subject to a set of constraints such as budget, the upper limit of total products offered to each customer and the lower limit of the total amount of each product offered which makes the involvement of the product in the campaign possible.

The modified model allows us to derive new heuristic approaches to the problem. The heuristic methods proposed in this study divide the problem into two sub-problems to make very large-sized instances solvable effectively and efficiently. Moreover, the modified model can be utilized to capture a strong approximation to optimal profit of the problem within reasonable time.

The remainder of the study is organized as follows. Chapter 2 presents background information about direct marketing. Chapter 2 helps us to understand the environment of business marketers, and basic concepts of designing promotion campaign. Chapter 3 summarizes important studies in the literature which investigates product targeting problem from different aspects. Chapter 4 describes product targeting problem. Chapter 5 explains proposed heuristic solutions and tabu search algorithm. Chapter 5 also highlights the results of heuristics methods and gives the comparative study between the literature and proposed methods in this study. Finally, Chapter 6 discusses the implications of this research for managers and further offers suggestions to enhance the model.

2. DIRECT MARKETING

Direct marketing is now more popular than before. The report by Direct Marketing Association shows that in 2010 $154.4, in 2011 $163 billion were spent for direct marketing. From 2011 to 2016, %4.9 of increase in direct marketing sales is expected. Bose & Chen reported that for 2005, it is expected that, on an average, $1 spent on consumer direct marketing would yield $12.66 compared to $10.10 for B2B direct marketing. The interest in direct marketing in both academia and business increases because of its exponential growth.

While direct marketing was known as “mail order” or “direct mail” before, now it turns to the “direct marketing”. The term “direct marketing” was first used by Wundemlan (1974). The term was developed to cover the new understanding of targeting and long term value. (Petrison et al., 1997)

Direct marketing has been defined by the Direct Marketing Association: “as an interactive process of addressable communication that uses one or more advertising media to effect, at any location, a measurable sale, lead, retail purchase, or charitable donation, with this activity analyzed on a database for the development of ongoing mutually beneficial relationships between marketers and customers, prospects, or donors” (2011).

From a system perspective direct marketing can be represented as shown in **Figure 2.1**. Following this figure, parts of the direct marketing are explained in the successive subsections.

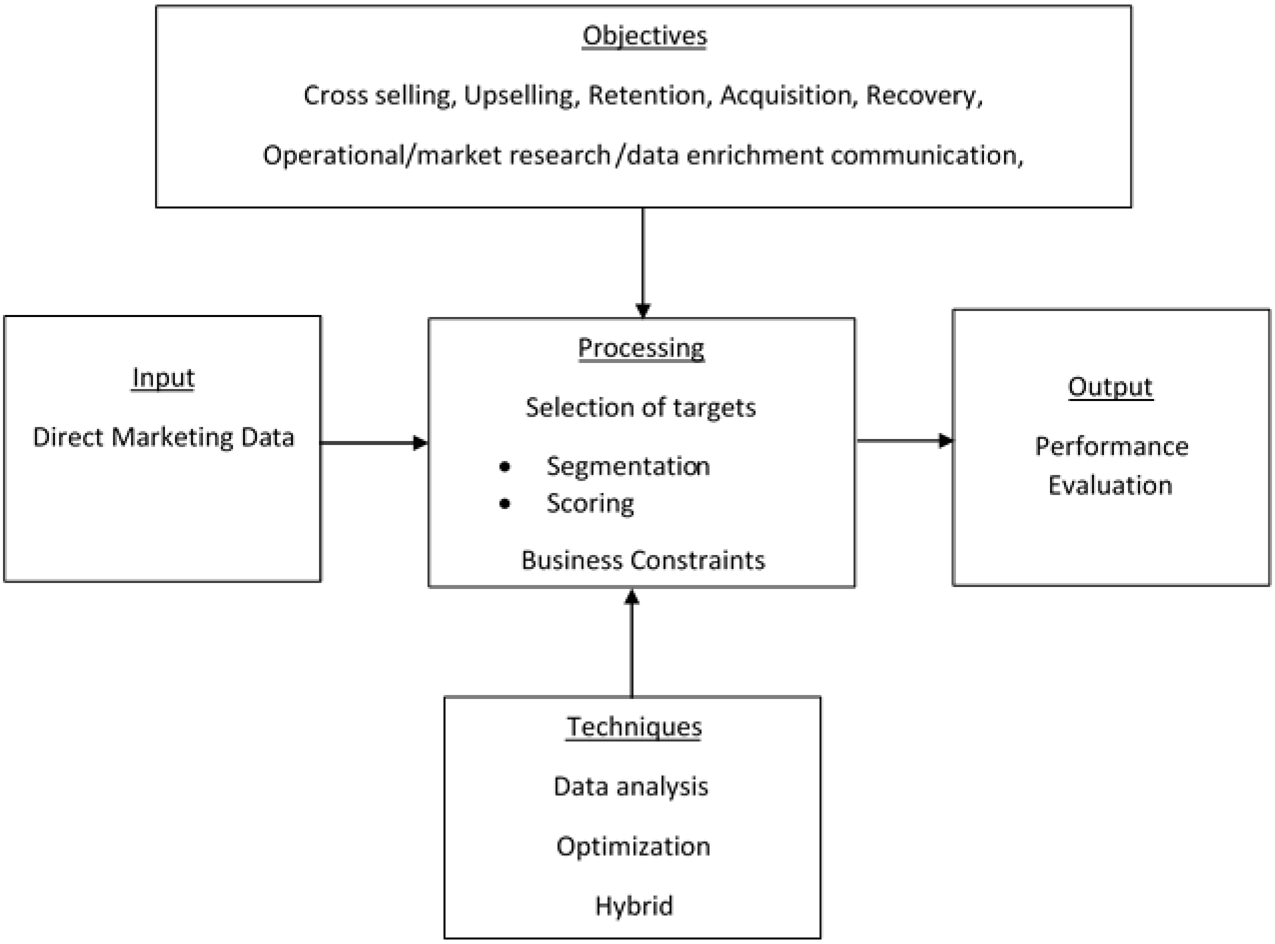

Figure 2.1 System perspective of direct marketing

Input of the direct marketing is data collection (Bose and Chen, 2009). Databases and collection of data are core process of direct marketing. The next important activity is selection of target customers since the revenue of direct marketing depends on the responses to the campaign (Bose and Chen, 2009). The customers to be targeted in a specific campaign are selected from the database, given different types of information such as demographic information and information on the customer's personal characteristics like profession, age and purchase history (Maderia and Sousa, 2002). The final activity is the evaluation of the direct marketing activities (Bose and Chen, 2009). The objectives of the campaign effect the targets, the way of selecting targets, and the methods of communication. Therefore, the marketers should focus on the aim of the campaign before studying on data in order to prepare the targeted lists.

2.1. System Perspective in Detail

From the system perspective view, the components of the direct marketing as inputs, processing, outputs, objectives and techniques are explained in detail in the following sections.

2.1.1. Objectives

As understand from the definition, in direct marketing, companies or organizations try to establish and maintain a direct relationship with their customers in order to target them individually for specific product offers (Madeira and Sousa, 2002). Offering a campaign is the way of relationship with the customer. The objectives of the campaign can be classifies as follows (Doyle, 2005):

- Cross selling: getting a customer to buy another services or products
- Upselling: getting a customer to upgrade an existing services or products
- Retention: getting a customer to retain or stay an existing services or products
- Recovery: getting a customer to re-instate an existing services or products
- Customer acquisition: communication with potential customers to sell the service or products
- Operational communication: giving important information to customer
- Market research communication: getting a customer to respond a market research
- Data enrichment communication: collecting additional information about customers

2.1.2. Inputs: Direct Marketing Data

Data in direct marketing can be layered into three levels as first, second and third type of data. The first type of data can be geographic (customer location of home, office etc), demographic (age, sex, family size, etc), lifestyle (customer habits and interest), and socio-graphic characteristics (Bult, 1993, Bult and Wittink, 1996, Van der Scheer, 1998). The second type of data includes customers' interactive behavior (transaction records, feedback from customers) (Bose and Chen, 2009). This data is suited in interactive behavior type of data. This can be called transaction data or RFM (Guido et al, 2011). RFM information helps to estimate the probability of customers will buy a certain product (Bose and Chen, 2009). Initials of RFM mean that:

- R (recency): the period since a customer's last order or the number of communication without response,
- F (frequency): the number of purchases made with during a certain period
- M (monetary or value): the amount of money that a customer spent during a certain period.

The last type of data is product and solicitation data which includes size, color, prize and design style. The marketers should know the exact characteristics of the products before planning the campaign since the customer may choose that product because of its characteristics. For example, for the upselling campaign, data on product should be known. Design style of solicitation may affect the customer choice. Before designing the campaign, type of the soliciting such as e-mail, postal mail, short message on mobile phone, etc. should be studied (Bose and Chen, 2009).

2.1.3. Processing

In direct marketing, the marketers mainly focus on the target selection under the limitations of the business constraints. These processes are explained in the following parts.

Selection of Target Customers

The two main research questions in direct marketing are "who should be selected as target for direct marketing?" and "what techniques should be used for selection of targets?" (Bose and Chen, 2009). Types of target selection are divided into two main groups: segmentation (or customer profiling) methods and scoring methods (Maderia and Sousa, 2002). Customer segmentation involves clustering similar customers together. RFM is used for customer segmentation purpose. A score is generated for an individual customer which can be a binary value or integer value indicating whether a customer would respond or not or represent a revenue, respectively (Bose and Chen, 2009). The customers are then ordered according to scores and only the customers who are likely to respond (e.g. their score is above a threshold value) are sent the product offer (Maderia and Sousa, 2002).

Business Constraints

The marketers should use available resources at the optimal level. Based on the firm strategy, the campaign designer looks for the appropriate level of using their budget, cost, inventory, etc. These limitations are defined as business constraints. There are several types of business constraints (Cohen, 2004):

- Restrictions on the minimum and maximum number of product offers that can be made in a campaign,
- Requirements on minimum expected profit from product offers,
- Limits on channel capacity,
- Limits on funding available for the campaign,
- Customer specific 'do not solicit' and credit risk limiting rules,
- Campaign return-on-investment hurdle rates

Apart from the above, there is another type of business constraints due to the technical issues or limitation of the company's technical abilities. Cohen (2004) gives minimum-quantity commitment (MQC) as an example for that types of business constraints. MQC gives the minimum requirement on the quantity of a product to include this product in the promotion campaign (Cohen, 2004). According to the banking environment, Cohen (2004) gives an example that the number of credit card offers must be greater than or equal to 20,000 or equal to zero.

These business constraints are based on the company's competitive environment and business sector. Bhaskar et al. (2009) consider two major business constraints in banking sector: camping budget and expected loan volume of each customer. In the environment of mobile advertising using SMS (Short Message Service) text messaging, De Reyck and Degraeve (2003), takes into consideration business constraints:

- Given a limited capacity of broadcast time slots (capacity constraints),
- Send ad once a day (repetition constraints)
- Two identical ads cannot be sent in the same time slot (Consecutive-type repetition constraints)

Since the aim of the business is getting revenue from the campaign, meeting the business constraint is an important criteria while planning the campaign. (Doyle, 2005; Bose and Chen, 2009)

2.1.4. Techniques

It is possible to find somewhat different classifications for target selection techniques in the related literature. The techniques used in both types of target selection are divided into three groups by Bose and Chen (2009): basic statistical techniques, advanced statistical techniques and machine learning. On the other hand, Maderia (2002) divides the techniques into four, namely; statistical regression, neural networks, decision trees, fuzzy modeling. The differences between the classifications stem from the current literature surveyed by researchers. In this study, following **Figure 2.1**, the surveyed techniques are considered in an extended outline as given in below:

- Data analysis: includes basic and advance statistical techniques.
- Optimization techniques: includes exact optimization techniques and heuristic techniques
- Hybrid techniques: includes both data analysis and optimization techniques. The constraints of the optimization problem can be derived from the result of data analysis.

In data analysis, regression models are used frequently for selection of targets as basic statistical techniques. The aim of using regression models in direct marketing is to predict customer responses. Customer responses can be represented as binary values (customer buys or don’t) or as any nonnegative value (customer spends a certain amount of money or don’t) (Levin and Zahavi, 1998). For example, when selecting target customers based on the scores given by a linear regression model, a threshold is usually set beforehand. If the score of a customer is greater than the threshold, he/she is selected (Bose and Chen, 2009). However, in linear regression models, there may be a negative return values to be estimated as a profit or return, therefore this might affect the overall sum of the observations. (Levin and Zahavi, 1998). Logit, probit, and tobit models are different from linear regression models (Maderia, 2002; Bose and Chen, 2009; Guido et al, 2011). In logit and probit models, response is binary choice (1-0 or yes/no). On the other hand, response could be revenue in tobit models. (Bose and Chen, 2009). The advanced statistical techniques are implemented for target selection by combining two simple statistical techniques (two stage models) (Bose and Chen, 2009). The first stage model the probability of responses, the second stage tries to figure out the money that a customer might spend when accept the campaign (Bose and Chen, 2009). Levin and Zahavi (1998) use logistic regression models to estimate customer response and linear regression models to estimate monetary value. If the response measure exceeds a specific value, the customer is selected for the promotion; otherwise the customer is rejected.

For the optimization techniques, mathematical programming, linear/integer programming, metaheuristics can be used to estimate the target by meeting the business constraints such as budget, cost or inventory. Beside the optimization techniques, based on the problem environment, some hybrids of these techniques exist in the literature.

2.1.5. Output: Performance Evaluation

Each direct marketing activity is evaluated by their revenue contributions. A company must run under a limited budget and a campaign necessitates a certain amount of cost. The cost of the campaign may vary according to the different business environment. Levin and Zahavi (1998) give an example about the costs: If the marketers are willing to send a mail to the customer about a promotion campaign, it should be worth doing, since expected return from an order must exceed cost invested in generating the order (means the cost of promotion). The promotion cost of this example may include brochure and postal costs. If solicitation cost is too high, this will reduce the revenue from the campaign (Bose and Chen, 2009).

Before planning the campaign, the objectives and evaluation criteria must be defined clearly. At the end, success of any target selection or product offering must be evaluated.

3. LITERATURE REVIEW

Following the system perspective of direct marketing (**Figure 2.1**), the problem addressed in this study uses a given direct marketing data. Final goal is to retain the existing customers. Offering promotion campaigns to retain the customers is the main activity within the problem environment. A promotion campaign involves a subset of existing products. Both determination of the best campaign and offering this campaign to customers constitute a kind of customer selection. Additionally, various business constraints are imposed to the problem environment. The resulting problem is called product targeting problem. Performance is measured as net profit gained by the targeted customers. The related literature with the defined problem is surveyed in detail in this section.

Variations on the product targeting problem including different sets of constraints and objectives reported in the literature are surveyed in this study. **Table 3.1** summarizes the survey according to its focus on the decision problems, the objectives, and the methods utilized. Decision problems are classified as customer selection (classification), product selection (product targeting) or both, under different prevailing conditions such as multi-period, multi-channel and uncertainty conditions. Solution methodologies suggested for these problems, on the other hand, are based on data analysis or optimization techniques. Readers are referred to the comprehensive study of Bose and Chen (2009). They investigate these methods and processes from a system perspective and provide a technical survey of the literature.

Table 3.1 Summary of the related literature with product targeting problem

Decision Problem	Objective	Method
Optimal selection of targeted customers (Bult and Wansbeek, 1995)	Maximize profit	Gains chart model
Estimation of customer response and monetary value (Levin and Zahavi, 1998)	Estimate expected return	Logistic regression models
Optimization of budget allocation btw acquisition and retention spending (Berger&Bechwati, 2001)	Maximize customer equity	Non-linear programming
Optimal design of a series of promotions (offer free gifts, discounts, special services) mailed to potential customers (Nair and Tarasewich, 2003)	Maximize the multiple purchases of customers	Genetic Algorithm
Customer selection and resource allocation (Venkatesan and Kumar, 2004)	Maximize customer lifetime value	Stochastic model to predict each customer's purchase frequency and a panel-data model that predicts contribution margin.
Customer segmentation (Jonker et al., 2004)	Maximize long term profit	RFM breakpoint segmentation, Markov decision model, and Genetic algorithm
Determination of type and number of products to be offered to each customer (Cohen, 2004)	Maximize marketing return on investment	Integer programming
Target-selection of customers (Prinzie and Van Den Poel, 2005)	Estimate binary response variable to optimize mailing depts (response or no response)	Binary logistic regression for classification
Prediction of the customer catalogs (Gönül and Hofstede, 2006)	Maximize expected profit and utility	Bayes rule
Optimization of cross-selling campaigns (Lu and Kun, 2007)	Maximize profit	TSP principle

Table 3.1 Summary of the related literature with product targeting problem – Cont.

Decision Problem	Objective	Method
Designing a marketing strategy to identify sequential optimal campaigns (Kim et al., 2009)	Maximize profit under the given defection probability	Self-Organizing Map and reinforcement Learning algorithm
Target selection process for a cross-sell campaign (Bhaskar et al., 2009)	Maximize net income	Fuzzy mathematical programming
Prediction of potential customers (Soopramanien and Juan, 2010)	Estimate return on investment	Binary logistic model
Determination of products to be offered (Schön, 2010)	Maximize profit	Non-linear programming with linear constraints
Prediction of potential customer by customer responses (Chen et al., 2011)	Predict potential customer range	Support vector machines Logistic regression
Determination of cross-selling strategies to introduce the right product to the right customer at the right time using the right comm. channel (Li et al., 2011)	Maximize long term profit	Dynamic programming
Optimization of targeted offers to each customer (Nobibon et al., 2011)	Maximize profit	Integer programming and heuristic algorithm
Optimization of product offers to group of customers using certain channels (Sundararajan et al., 2011)	Maximize profit	Markov chains, genetic algorithms, and mathematical programming
Customer classification to find new products and services with high sales potential (Ahn et al., 2011)	Expand revenue and profit	Logistic regression, artificial neural networks, and genetic algorithm
Identification of customers to whom additional products should be offered (Thuring et al., 2012)	Estimate risk profile	Multivariate credibility model
Estimation of the number of customers who will ultimately respond, the number of responses will be received by a certain period of time (Chun, 2012)	Estimate of customer responses	Monte Carlo simulation
Determination of the most valuable customers in SNS (Xu et al., 2012)	Maximize profit	Semidefinite programming

From the nineties to the present, studies which use data analysis have been popular due to advances in related technology. Among the studies given in **Table 3.1**, Venkatesan and Kumar (2004), Prinzie and Van den Poel (2005), Gönül and Hofstede (2006), Soopramanien and Juan (2010), Chen et al. (2011), Thuring et al. (2012), Chun (2012) utilize data analysis. These studies mainly examine the methodologies to estimate customer behavior. The proposed methodologies vary from basic models to more complicated response models in order to obtain more precise estimations. Bose and Chen (2009) show that RFM (recency, frequency, and monetary) data is used mostly for building target selection for existing or past customers. RFM relies on three customer variables ─how long since the last customer purchase, how often the customer purchases, how much the customer has bought─ to find valuable customers for future direct marketing campaigns (Olson et al., 2012). Verhoef et al. (2002) report that many companies use simple heuristics for the selection of targets such as ‘‘mail all customers who recently purchased a specific product” or ‘‘mail all customers who have an income above a certain limit’’. However, Bult and Wansbeek (1995) state that the disadvantage of an RFM model is its use of a limited number of variables.

Besides building simple RFM models, customer response models based on historical data are also used to predict customer behavior using advanced statistical techniques. Prinzie and Van Den Poel (2005) develop a response model using RFM data within budget constraints. Levin and Zahavi (1998), use logit models to estimate the customer response and use linear regression to estimate monetary value. Gönül and Hofstede (2006) propose a response model using Bayesian rule approaches. Chun (2012) estimates the responses by Monte Carlo simulation. Chen et al. (2011) present response models based on support vector machines and logistic regression while Soopramanien and Juan (2010) construct a decision support tool which classifies customers using a logistic regression model. Another piece of information related to customer selection is customer lifetime value (CLV) which can be calculated as the net profit/loss to the firm obtained over the lifetime of a customer (Jain and Signh, 2002). Venkatesan and Kumar (2004) take into account CLV and examine firms’ profitability within budget constraint proposing a stochastic model for prediction purposes and customer selection.

The studies mentioned above use mainly data analysis. On the other hand, companies also need to optimize various business requirements subject to available resource constraints. Berger and Bechwati (2001), Nair and Tarasewich (2003), De Reyck and Degraeve (2006), Ryals et al. (2007), Lu and Kun (2007), Bhaskar et al. (2009) Schön (2010), Sundararajan et al. (2011), Xu et al. (2012) employ optimization methodologies for this purpose. Berger and Bechwati (2001) focus on optimization techniques and offer a general approach to the budget allocation problem in order to maximize CLV. Xu et al. (2012) use semi-definite programming to determine the most valuable customers on social networking sites. Nair and Tarasewich (2003) consider the optimal design of a series of promotions mailed to potential customers. De Reyck and Degraeve (2006) investigate the problem of which ads to send out to which customers at which specific times, operating under capacity constraints while maximizing customer response and revenue. Ryals et al. (2007) develop a model for marketing portfolios in order to maximize overall return within certain risk constraints. Lu and Kun (2007) present a model based on the travelling salesman problem in order to maximize profit for optimizing cross-selling campaigns. Bhaskar et al. (2009) consider several parameters such as response propensity, expected volume, expected profit and then group the customers using similar values of these parameters. They use fuzzy mathematical programming to maximize net income from each group of customers. Though they overcome computational difficulties in large data sets, grouping schemes may affect the quality of solution. Schön (2010) suggests a non-linear integer programming model for product line design to determine the number of products to be offered, the differentiation of the products offered and their prices. Sundararajan et al. (2011) look for optimal product offers to a number of customer groups by using genetic algorithm and fuzzy linear programming techniques. Nobibon et al. (2011) develop optimization models related to the product targeting problem to maximize total profit subject to a set of limitations such as budget, the maximum number of products offered to customers, and the minimum number of products offered to customers. They also propose heuristic techniques in order to approximate optimum solutions.

Some of the studies in **Table 3.1**, namely, Jonker et al. (2004), Cohen (2004), Kim et al. (2009), Li et al. (2011), Ahn et al. (2012) emphasize both data analysis and optimization methods. Jonker et al. (2004) use RFM information to segment customers and then use the Markov decision model for maximization of profit from each segment. Additionally, they utilize genetic algorithms to obtain different segmentations. Cohen (2004) aggregates customers into groups and then optimize the number of customers who received certain products within each group. Cohen (2004), finally, assigns products to individual customers using an assignment model for each group. Li et al. (2011) propose a dynamic programming model for cross-selling campaigns in which the goal of the firm is to maximize the long-term profit from its existing customers. Ahn et al. (2011) develop a model in two steps: the first step involves the prediction of response probabilities and the second step runs a genetic algorithm in order to make final decision for target customers. Kim et al. (2009) use a self-organizing map for customer segmentation and then identify sequential optimal campaigns by using the shortest path maps.

As mentioned before, in this study, the development of heuristic methods for targeted offers in a promotion campaign is investigated. The objective of the problem is to maximize profit yield by offering products to customers subject to a set of constraints such as budget, the upper limit of total products offered to each customer and the lower limit of the total amount of each product offered which makes the involvement of the product in the campaign possible. In the literature, there are a few studies which are closely related with this problem. Recently, Nobibon et al. (2011) directly address the problem and propose both optimization models and heuristic algorithms. Since the optimization model by Nobibon et al. (2011) is a basic formulation which best fits the problem, we focus on a modification of this basic model. As explained in chapter 4 in detail, the modified model allows us to derive new heuristic approaches to the problem. The contributions of this study can be summarized as:

- Enlargement of the literature since it contains a few numbers of studies that directly focus on the problem addressed in this study.

- Development of new heuristic approaches to deal with an approximate solution of the problem. The heuristic methods proposed in this study divide the problem into two sub-problems to make very large-sized instances solvable effectively and efficiently.
- Suggestion of a linear mathematical model which is a modification of the model existing in the literature. The modified model can be utilized to capture a strong approximation to optimal profit of the problem.

THIS PAGE INTENTIONALLY LEFT BLANK

4. DESCRIPTION OF THE PRODUCT TARGETING PROBLEM

The main focus of the product targeting problem is to maximize total profit by means of offering different products to clients while satisfying various business constraints. The subset of product types offered to clients forms a promotion campaign. The business constraints are mainly the hurdle rate (the minimum acceptable rate of return), the limited budget for each product, the minimum-quantity commitment for each product to be included in the campaign and the limitation on the number of offers to each client. Following the notations of Nobibon et al. (2011), the parameters of the problem are listed as given below:

p_{ij} : expected return to the firm from offering of product j to client i,

c_{ij} : unit cost of offering product j to client i,

M_i : maximum number of offers that can be exhibited to client i,

O_j : minimum quantity commitment bound of product j,

B_j : budget for product j,

f_j : fixed cost of including product j in promotion campaign,

R : corporate hurdle rate.

Model-1 (equations 4.1-4.8) developed by Nobibon et al. (2011) provides the selection of the products out of n products which will be offered to m clients and the distribution of these products to the clients subject to the constraints to optimize the profit function defined in equation (4.1). The hurdle rate constraint in equation (4.2) is a managerial requirement while equation (4.3) represents the budget restriction, B_j, for each product j. Equation (4.4) stipulates that the total number of products offered to client i cannot excess the demand of client i, M_i. According to equations (4.5) and (4.6) the total offering number of product j must reach at least a certain level, O_j, or the product must be excluded from the promotion campaign. Binary variable y_j equals 1 for the selected (targeted) products in the campaign and zero for the excluded products. Meanwhile, binary x_{ij} is 1 if product

j is offered to client i, and otherwise x_{ij} is 0. It was proved by Nobibon et al. (2011) that even for $O_j = 1$ the problem is strongly NP-hard.

Model-1:

$$\text{Maximize} \quad \sum_{i=1}^{m}\sum_{j=1}^{n}(p_{ij} - c_{ij})x_{ij} - \sum_{j=1}^{n} f_j y_j \tag{4.1}$$

$$\text{Subject to;} \qquad \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}x_{ij} \geq (1+R)\left[\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} + \sum_{j=1}^{n} f_j y_j\right] \tag{4.2}$$

$$\sum_{i=1}^{m} c_{ij}x_{ij} \leq B_j \qquad j = 1, \cdots, n \tag{4.3}$$

$$\sum_{j=1}^{n} x_{ij} \leq M_i \qquad i = 1, \cdots, m \tag{4.4}$$

$$\sum_{i=1}^{m} x_{ij} \leq m y_j \qquad j = 1, \cdots, n \tag{4.5}$$

$$\sum_{i=1}^{m} x_{ij} \geq O_j y_j \qquad j = 1, \cdots, n \tag{4.6}$$

$$x_{ij} \in (0,1) \qquad i = 1, \cdots, m, \; j = 1, \cdots, n \tag{4.7}$$

$$y_j \in (0,1) \qquad j = 1, \cdots, n \tag{4.8}$$

5. HEURISTIC APPROACHES TO THE PRODUCT TARGETING PROBLEM BASED ON MATHEMATICAL PROGRAMMING

The following section describes the proposed solutions and model comparison in detail.

5.1. Proposed solution procedure to the problem

The main idea proposed in this study is to separate the product targeting problem into two sub-problems to deal with the solution of model-1: the selection of products which will be included in the promotion campaign and the distribution of these products to clients optimally. This strategy is carried out in two phases. In phase I, a new linear programming model is utilized to predict which products are selected for or removed from the product campaign. Once the products in the campaign are determined, the product targeting problem is reduced to a kind of generalized assignment problem. In phase II, the products selected in phase I are distributed to clients optimally by another optimization model. The two phases are connected via a heuristic rule. Two alternative heuristic rules, derived from the proposed linear programming model in phase I, are suggested to predict the products eliminated from the campaign (or equivalently the products involved in the campaign).

5.1.1. Phase I

At phase I, a new linear programming model, called model-2, is proposed as given in equations (5.1 - 5.8). Model-2 is a modification of model-1 in which a dummy client m+1 and corresponding nonnegative variables, $x_{m+1,j}$, (j = 1,...,n) are included. Unit profits and costs, $p_{m+1,j}$ and $c_{m+1,j}$, are assumed to be all zero. The utilization of the dummy variables, $x_{m+1,j}$, allows dropping of equations (4.5) and (4.8) and relaxation of equation (4.6) in model-1. Equation (4.6) is replaced with equation (5.5), therefore the total distributed quantity, including the dummy variable, of each product j must be at least O_j. Equation (5.5) guarantees that when the sum of the distributed quantity of product j to all clients (i = 1, .., m) is less than O_j, dummy variable, $x_{m+1,j}$ takes a positive level which will support equation (5.5). On the other hand, the optimization procedure tends to make x_{ij} (i = 1, .., m, j = 1,...,n) as large as possible. Since model-2 does not search for optimal assignments of binary variables, x_{ij}, these are also relaxed in equation (5.7). Binary variables y_j are dropped from model-1 by replacing with $\left(1 - x_{m+1,j} / O_j\right)$. In other words, we act $\left(1 - x_{m+1,j} / O_j\right)$ as y_j. O_j is an upper bound for each $x_{m+1,j}$ and therefore the ratio $x_{m+1,j} / O_j$ lies between 0 and 1. Hence, as the optimization procedure increases the value of $x_{m+1,j}$, the corresponding product j moves closer to exclusion from the promotion campaign. The final modification of model-1 is the replacement of B_j with $B_j\left(1 - \frac{x_{m+1,j}}{O_j}\right)$ in equation (5.3), for each product j. By the way, the offering of product j partially also necessitates reducing its budget in the same portion.

Model-2:

$$\text{Maximize } \sum_{i=1}^{m}\sum_{j=1}^{n}(p_{ij} - c_{ij})x_{ij} - \sum_{j=1}^{n} f_j\left(1 - \frac{x_{m+1,j}}{O_j}\right) \tag{5.1}$$

Subject to

$$\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}x_{ij} \geq (1+R)\left[\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} + \sum_{j=1}^{n} f_j\left(1 - \frac{x_{m+1,j}}{O_j}\right)\right] \tag{5.2}$$

$$\sum_{i=1}^{m} c_{ij}x_{ij} \leq B_j\left(1 - \frac{x_{m+1,j}}{O_j}\right) \qquad j = 1,\cdots,n \tag{5.3}$$

$$\sum_{j=1}^{n} x_{ij} \leq M_i \qquad i = 1,\cdots,m \tag{5.4}$$

$$\sum_{i=1}^{m+1} x_{ij} \geq O_j \qquad j = 1,\cdots,n \tag{5.5}$$

$$x_{m+1,j} \leq O_j \qquad j = 1,\cdots,n \tag{5.6}$$

$$0 \leq x_{ij} \leq 1 \qquad i = 1,\cdots,m,\ j = 1,\cdots,n \tag{5.7}$$

$$x_{m+1,j} \geq 0 \qquad j = 1,\cdots,n \tag{5.8}$$

5.1.2. Comparison of Model-2 with Model-1

In this subsection it is shown experimentally that model-2 gives better results than the linear programming relaxation of model-1 which consists of equations (4.2)- (4.6), and $0 \leq \mathbf{x} \leq 1, 0 \leq \mathbf{y} \leq 1$The experiments made to compare z^*_{M2} with both z^*_R and z^* are summarized in Table 5.1, where z^*_{M2}, z^*_R, and z^* are objective values of model-2, the relaxed model-1, and model-1, respectively. The table represents percentage deviations, Δ (given in equation (5.9)), of both z^*_{M2} and z^*_R from z^* in the test problems taken from Nobibon et al. (2011).

In this experiment we test client sizes (*m*), 100, 200, 300, 1000, and 2000 and product sizes (*n*), 5, 10, and 15 fixing 10% hurdle rate, an average budget capacity and considering two different cases of customer behavior (they can receive only a few offers or a higher number of offers) for each client size. The whole set of the test problems are described in the following sections in detail.

Table 5.1 shows that deviation of z^*_{M2} from z^* is less than that of z^*_R for each instance. In other words, the objective value supplied by model-2 is closer to the optimum value than that given by the relaxed model-1 for each of the instances. The average deviation over all instances is 18.3% according to model-2, compared with 49.5% according to the relaxed model-1.

From the table it is clear that $z^*_{M2} \leq z^*_R$ for each instance and $z^*_{M2} \geq z^*$ except for two instances as also illustrated in Figure 5.1. There is no guarantee that z_{M2} is an upper bound on the optimum, z^*. The experiments show that if $K^*_{M2} - K^* \geq L^*_{M2} - L^*$ then $z^*_{M2} \geq z^*$, otherwise $z^*_{M2} \leq z^*$. On the other hand, the average computer time taken to solve the relaxed model is 0.39 seconds while it takes 5.21 seconds to solve model-2.

$$\Delta = 100\left(z_i^* - z^*\right)/z^* \quad \text{,for solution approach } i \tag{5.9}$$

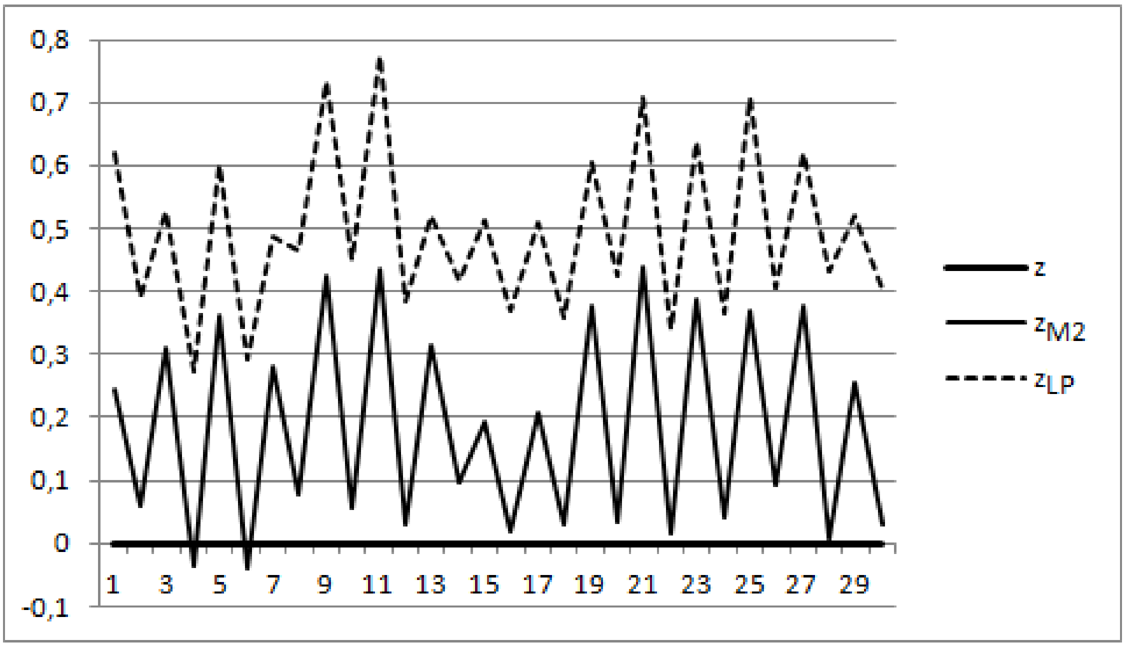

Figure 5.1 Deviations of z_{M2}^* and from the z_R^* optimal profit

Table 5.1 Comparison of z^*_{M2} and z^*_R on the selected test problems

		The Relaxed Model-1		Model-2	
Instance	z^*	Δ	Time (s)	Δ	Time (s)
100-5-l	1039	0.62	0.29	0.24	0.10
100-5-s	775	0.39	0.22	0.06	0.13
100-10-l	2322	0.53	0.31	0.31	0.22
100-10-s	1369	0.27	0.10	-0.03	0.12
100-15-l	3394	0.60	0.28	0.36	0.42
100-15-s	1776	0.29	0.26	-0.04	0.43
200-5-l	2421	0.49	0.24	0.28	0.24
200-5-s	1461	0.47	0.10	0.08	0.39
200-10-l	3710	0.73	0.34	0.43	0.64
200-10-s	2423	0.45	0.28	0.06	0.70
200-15-l	5526	0.77	0.23	0.43	1.72
200-15-s	3659	0.39	0.34	0.03	1.43
300-5-l	3379	0.52	0.22	0.31	0.35
300-5-s	2300	0.42	0.13	0.09	0.3
300-10-l	6906	0.51	0.12	0.19	1.02
300-10-s	3950	0.37	0.16	0.02	1.31
300-15-l	10699	0.51	0.23	0.21	2.10
300-15-s	5351	0.36	0.17	0.03	1.95
1000-5-l	10092	0.61	0.33	0.37	3.54
1000-5-s	7394	0.42	0.24	0.03	3.90
1000-10-l	19415	0.71	0.48	0.44	6.59
1000-10-s	13214	0.34	0.50	0.02	1.64
1000-15-l	31485	0.64	0.59	0.39	14.15
1000-15-s	17529	0.36	0.64	0.04	17.69
2000-5-l	19211	0.71	0.52	0.37	13.17
2000-5-s	15195	0.41	0.59	0.09	14.70
2000-10-l	44387	0.62	0.75	0.38	14.72
2000-10-s	24120	0.43	0.98	0.01	2.92
2000-15-l	70681	0.5	0.98	0.25	44.54
2000-15-s	34327	0.40	0.95	0.03	5.15
Average		49.5%	0.39	18.3%	5.21

5.1.3. Phase II

As explained before, in phase I, a promotion campaign is formed and in phase II, the products included in the campaign are distributed to the clients optimally. To obtain a promotion campaign two heuristic rules are proposed in this study. These rules and the incorporation of them into another optimization model in phase II are explained in the next subsections in detail.

5.2. Heuristic Rule 1

A result inferred from the basic formulation, model-1, is that unless the inequality given by equation (5.10) is satisfied, a feasible solution to the problem cannot be acquired. It follows that if $\sum_{j=1}^{n} O_j (1 - y_j) \geq \mathrm{C}$, where $\mathrm{C} = \sum_{j=1}^{n} O_j - \sum_{i=1}^{m} M_i$ is a constant, then model-1 is converted into a variety of the generalized assignment problem in which the total number of assignments of product j to clients is O_j whereas the total number of assignments of client i to products is

$$M_i \sum_{j=1}^{n} O_j y_j \leq \sum_{i=1}^{m} M_i \tag{5.10}$$

According to heuristic rule 1, $\left(1 - x^*_{m+1,j} / O_j\right)$ from the optimum solution of model-2 is treated as a measure of possibility of product j to be an offered product or similarly $x^*_{m+1,j} / O_j$ is treated as the excluding possibility of product j. Rule 1 single outs the products which have the highest exclusion possibility until the total amount of Oj of eliminated products meets the inequality in equation (5.11), where N_e is the set of excluded products.

$$\sum_{j \in N_e} O_j \geq \mathrm{C} \tag{5.11}$$

Once N_e is obtained, the remaining products represented by set N' are optimally distributed to clients by an integer programming model, model-3, below

Maximize $$\sum_{i=1}^{m}\sum_{j\in N'}(p_{ij}-c_{ij})x_{ij}-\sum_{j\in N'}f_j \tag{5.12}$$

Subject to

$$\sum_{i=1}^{m}\sum_{j\in N'}p_{ij}x_{ij} \geq (1+R)\left[\sum_{i=1}^{m}\sum_{j\in N'}c_{ij}x_{ij}+\sum_{j\in N'}f_j\right] \tag{5.13}$$

$$\sum_{i=1}^{m}c_{ij}x_{ij} \leq B_j \qquad \forall j \in N' \tag{5.14}$$

$$\sum_{j\in N'}x_{ij} \leq M_i \qquad i=1,\cdots,m \tag{5.15}$$

$$\sum_{i=1}^{m}x_{ij} \geq O_j \qquad \forall j \in N' \tag{5.16}$$

$$x_{ij} \in (0,1) \qquad i=1,\cdots,m,\ \forall j \in N' \tag{5.17}$$

Finally, the steps of the proposed heuristic approach based on rule 1, H-R1, are given below:

Step 1: Solve model-2 to obtain optimum $x^*_{m+1,j} / O_j$ for each j

Step 2: $N_e = \{\ \}$, $N = \{1, 2, ..., n\}$

Step 3: Repeat

Step 3.1: $k = \arg\max_{j \in N \setminus N_e} \left(\frac{x^*_{m+1,j}}{O_j} \right)$

Step 3.2: $N_e = N_e + \{k\}$

Until $\sum_{j \in N_e} O_j \geq C$

Step 4: Set $N' = N \setminus N_e$

Step 5: Repeat

Step 5.1: Solve model-3 to obtain a heuristic solution

Step 5.2: If the solution is infeasible then $N' = N' - \{k\}$ where

Step 5.3: $k = \arg\max_{j \in N'} \left(\frac{x^*_{m+1,j}}{O_j} \right)$

Until a feasible heuristic solution is found with profit of z_{H_R1}

Figure 5.2 Steps of the H-R1

5.3. Heuristic Rule 2

The second heuristic rule to bind phase-I and phase-II is drawn from the dual of model-2. To develop this rule, we focused on the dual constraints, which correspond to the dummy variables $x_{m+1,j}$, given in equation (5.18). Definitions of the dual variables and complete dual model, named model-4, are given in Appendix A.

$$(1+R)\frac{f_j}{O_j}u + \frac{B_j}{O_j}w_j + t_j + r_j \geq \frac{f_j}{O_j} \qquad j = 1,\cdots,n \tag{5.18}$$

Assuming dual variables u, t_j, and r_j are equal to zero at the optimum level expecting that the optimization procedure will enforce the primal constraint in equation (5.2) to exceed its right-hand-side and drive variables x_{ij} to be positive as far as possible. We also presume that the dummy variables will cause surpassing the right-hand-side of the constraint in equation (5.5). According to the complementary slackness of the dual constraint in equation (5.18), as its left-hand-side equals or is greater than the right-hand-side, the associated dummy variable, $x_{m+1,j}$, will be a non-basic variable at zero level. Hence, the second rule suggests that the higher ratio of f_j/B_j for product j is closer to being eliminated from the promotion campaign. An apparent advantage of rule 2 over rule 1 is that rule 2 does not require solving model-2. In chapter 5.5, we investigate the performances of both rules in terms of solution quality and computer time. The steps of the heuristic approach based on rule 2, H-R2, are also given as follows:

Step 1: Compute f_j / B_j for each j

Step 2: $N_e = \{ \}$, $N = \{1, 2,...,n\}$

Step 3: Repeat

Step 3.1: k $= \arg \max_{j \in N \setminus N_e} \left(\frac{f_j}{B_j} \right)$

Step 3.2: $N_e = N_e + \{k\}$

Until $\sum_{j \in N_e} O_j \geq \mathrm{C}$

Step 4: Set $N' = N \setminus N_e$

Step 5: Repeat

Step 5.1: Solve model-3 to obtain a heuristic solution

Step 5.2: If the solution is infeasible then $N' = N' - \{\mathrm{k}\}$

where $k = a\,rg \max_{j \in N'} \left(\frac{f_j}{B_j} \right)$

Until a feasible heuristic solution is found with profit of z_{H_R2}

Figure 5.3 Steps of the H-R2

Following the explanation of H-R1 and H-R2 in detail, the basic solution methodology to the product targeting problem suggested in this paper is also illustrated by Figure 5.4.

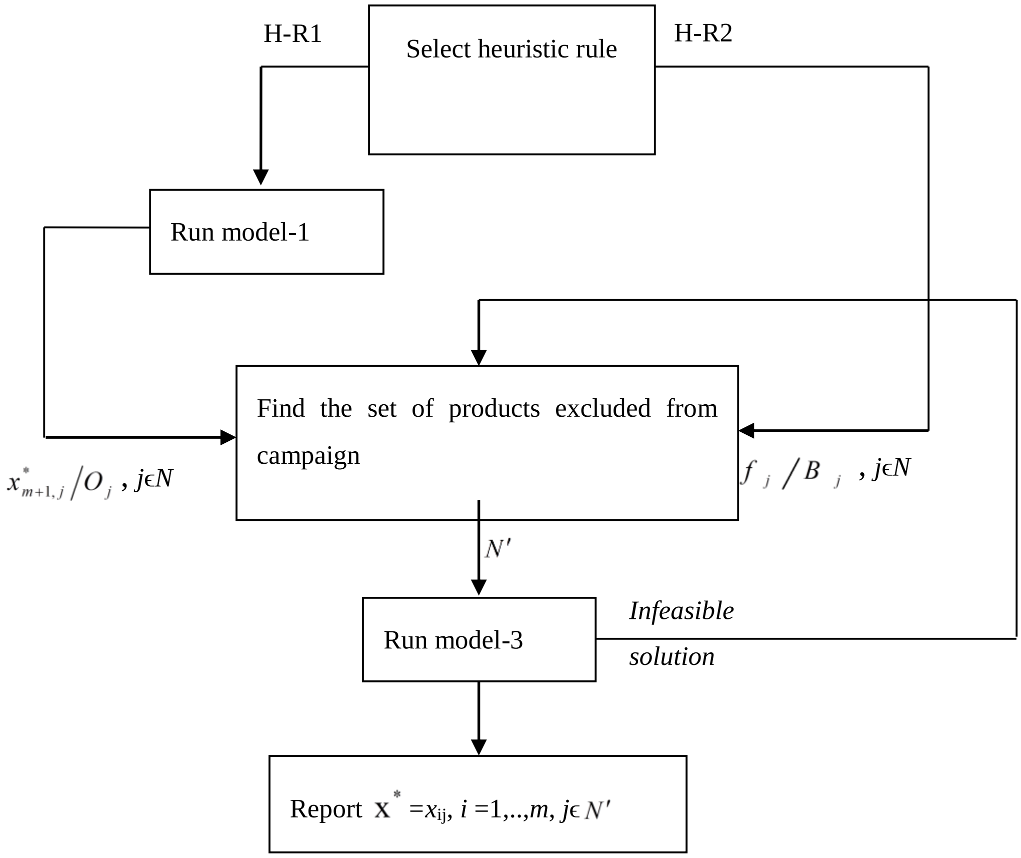

Figure 5.4 Solution methodology suggested to the product targeting problem

5.4. Tabu Search for the Product Targeting Problem

To obtain better solutions to the product targeting problem a tabu search-mathematical programming based hybrid algorithm is proposed as a third heuristic approach. Before the explanation of this heuristic in section 5.4.2 in detail, basic form of tabu search is given in section 5.4.1.

5.4.1. Basic Tabu Search

Tabu search (TS) is a higher level heuristic procedure for combinatorial optimization problems (Glover, 1990) The basic form of TS is proposed by Fred Glover (Glover and Laguna,1997), although its roots are from late 1960's and early I970's. TS is a Metaheuristic that guides a local heuristic search strategy to look for the good solutions in complex solution space beyond local optimality (Glover and Laguna,1997, Glover and Taillard, 1993).

"Metaheuristic" term was first used with the "tabu search" in the same paper in 1986 by Fred Glover. Metaheuristic is defined as a set of concepts or general algorithmic strategy that can be used to define heuristic method to be applied to many problems (Metaheuristics network project, 2015). Metaheuristics provide good solutions for solving complex optimization problems, but does not guarantee global optimum solution (Blum and Roli, 2003).

- The properties are listed below that characterize most Metaheuristics (Blum and Roli, 2003):
- Metaheuristics are strategies that guide the search process.
- The goal is to efficiently explore the search space in order to find near–optimal solutions.
- Techniques which constitute metaheuristic algorithms range from simple local search procedures to complex learning processes.
- Metaheuristic algorithms are approximate and usually non-deterministic.
- Metaheuristics are not problem-specific.

TS is a metaheuristic which guides the local search heuristic to escape local optimum traps. That's why, to define a TS algorithm, firstly the associated local search algorithm should be clearly identified. The local search heuristic uses a strategy which is based on certain moves to define the neighborhood of any given solution ((Glover and Laguna, 1997, Glover, 1986). The example of these moves can be "changing the value of variables, adding or deleting an element from a set" (Glover, 1990). Local search starts with a given solution point and then generate the neighborhood of the current solution using the pre-defined moving mechanisms. At each iteration, a local search heuristic replaced the current solution with a better solution in the neighborhood, if there exits such a solution. Otherwise, the heuristic terminates with current solution which is called a local optimum solution.

Neighborhood structure in local search is very important to cleverly search the solution space. Neighborhood generation mechanism is applied to current solution, X, to define a set of neighboring solutions in the search space, denoted N(X). Therefore, N(X) is a subset of the search space. The fundamental steps of neighborhood search method are given as below:

Step 1: Initialization

Step 1.1: Select starting solution $x^{now} \epsilon X$

Step 1.2: Record current best solution $x^* \leftarrow x^{now}$

Step 1.3: Define best_cost = c(x^*)

Step 2: Select and termination

Step 2.1: Choose solution $x^{next} \epsilon N(x^{now})$

Step 2.2: if there is no better solution to be next or another termination criteria is met then stop the search.

Step 3: Update

Step 3.1: Re-set if $c(x^{now}) > c(x^{next})$ then $x^{now} \leftarrow x^{next}$

Step 3.2: If $c(x^{now}) < c(x^*)$, go to step1-II, then return to step 2.

Figure 5.5 The fundamental steps of neighborhood search method

The word “tabu” comes from a language of Polynesia, to define things that cannot be used because they are sacred (Webster dictionary, 2015). In TS method, tabu is used to indicate the restrictions to guide the alternative candidates (Glover and Laguna, 1997).

The purpose of these restrictions on moves is to prevent cycling (Glover, 1986). Therefore with the help of those restrictions, the procedure goes beyond with points of local optimality while still in high quality solutions in each step (Glover, 1990).

Another important feature of TS is its way of memory usage. There are two types of memory in TS; long term memory and short term memory which is the core of TS. Long term memory is used primarily as a basis of strategies for intensifying and diversifying the search (Glover, 1990). Intensifying strategies are used for modifying choice rules to encourage move combinations and solution features in the past found good (Glover et al, 2007) On the other hand, diversification strategies look for the new attributes that is not included within solutions already generated (Glover et al, 2007).

There are four principle dimensions in memory structures of TS: Recency, frequency, quality and influence (Glover and Laguna, 1997).

- Recency based memory is used for keeping the track record of the solution attributes that are recently changed. Short term memory uses the recency based memory. Selected attributes which are recently visited are labeled as tabu-active, solutions are becomes tabu to prevent solutions from recent visits.
- Frequency based memory holds the frequencies to consist of ratios; transition measures (the number of iteration where an attribute changes the solution visited on a particular trajectory) and residence measure (the number of iteration where an attribute belongs to solutions visited on a particular trajectory).
- Quality dimension is defined as the ability to differentiate the merit of a solutions visited during the search.
- Influence dimension is used to impact the choices made during search.

Short term memory generally managed by a tabu list which keeps track of most recently visited solutions and forbids move toward them (Blum and Roli, 2003). Tabus are stored in a short term memory of the search (tabu lists) (Crainic and Gendrea, 2002). Management of the tabu lists strongly depend on the solution representation and moving mechanism. Tabu tenure is the number of iterations and during these iterations certain moves are recorded as tabu. A large tabu tenure makes the search process to explore larger regions, since it forbids revisiting (Crainic and Gendrea, 2002). Steps of the basic TS which utilizes the short term memory is shown below:

Step 1: Initialization

Step 1.1: Select an initial x ϵ X and let $x^* \leftarrow x$. Set k = 0, begin with an empty tabu list, T.

Step 2: If $N(x)$ - T is empty, go to Step 4.

Step 2.1: Set k: = k + 1 and select x′ ϵ $N(x)$ - T such that

x′ = BEST(xs: xs ϵ $N(x)$ – T, and s=1,.., $|N(x)|$).

Step 3: Let x ← x′. If $c(X) < c(x^*)$, let x*← x.

Step 4: If a chosen number of iterations has elapsed either in total or since x* was last improved,

or if $S(x)$ - T = 0 upon reaching this step directly from Step 2, stop.

Step 4.1: Update T by recording some elements of x′ as tabu and return to Step2.

Figure 5.6 Steps of the basic Tabu Search

5.4.2. Hybrid TS-Mathematical Modeling Algorithm for the Product Targeting Problem

In the previous sections implementation of H-R1 and H-R2 heuristics for the product targeting problem were explained. The experimental tests show that these heuristics are quite effective to find an approximate solution to the problem. In this section, application of a TS heuristic to the problem is investigated to increase qualities of the solutions obtained from H-R1 and H-R2 heuristics. Recall that the product targeting problem is divided into two sub-problems in this study: Selection of the products to be included in the promotion campaign and distribution of the selected products to the customers. Both H-R1 and H-R2 use different rules to solve the first sub-problem (called phase I) as explained previously whereas the second sub-problem (phase II) is solved optimally. Since phase II has already been optimized, obtaining different subsets of the products for the promotion campaign is crucial to get better solutions to the whole problem. Therefore, TS algorithm, named H-TS, is employed to explore the solution space of the first sub-problem. The second sub-problem is again solved by using model-3 to obtain optimum distribution of the candidate subsets of the products, which are generated by H-TS, to the customers. By the way, TS heuristic and model-3 are used together to obtain high qualified solutions to the product targeting problem. In other words, H-TS calls model-3 to obtain objective values (profits) of the solution points (candidate subsets of the products) searched throughout the progress of the TS.

Implementation of the proposed TS algorithm has three main stages. The first one is solution representation and initialization, second is neighborhood generation mechanism, third one is tabu list management.

The notations used in the developed algorithm are introduced below:

X :	array of n binary variables
$j=1,\ldots,n$:	index of the products
X_b :	best solution
X^0 :	initial solution
$x_j = \begin{cases} 1, \\ 0, \end{cases}$	if product is in the campaign otherwise
tal= $[\ldots..]_{nx1}$:	tabu add list, n sized integer array
tdl= $[\ldots..]_{nx1}$:	tabu drop list, n sized integer array
t :	iteration number
tt :	tabu tenure
r :	the number of index in which add or remove the product
T:	total number of iteration for termination

Solution Representation

A solution point is a candidate subset of products, or in other words, a candidate promotion campaign. A solution point, $X = [x_1, x_2, \ldots, x_n]$ is represented by an array by size *n* where *n* is the number of all products and x_j is binary variable. If $x_j = 1$ then j. product is included in the campaign, otherwise excluded.

For example, X= [0, 1, 0, 1, 0, 0, 1, 0, 0, 0] means that 2., 4. and 7. products are selected (offered) out of 10 products and others are eliminated.

Initial solution

Initial solution, X^0, is taken from either the phase I of H-R1 or phase I of H-R2. In other words, rule 1 or rule 2 can be utilized to start H-TS heuristic.

Neighborhood generation

To generate the neighborhood of a solution X, N(X), value of each element x_j, which corresponds product j, is shifted to 0 if $x_j = 1$ (drop move) and to 1 (add move) if $x_j = 0$. By the way, total number of neighbor solutions becomes *n*. As an example, neighbor solutions to solution X= [0, 1, 0, 1, 0, 0, 1, 0, 0, 0] are given in the following table.

Table 5.2 An example to solution representation and neighbor generation of H-TS

Solution representation	Products in the campaign
[1, 1, 0, 1, 0, 0, 1, 0, 0, 0]	1, 2, 4, 7
[0, 0, 0, 1, 0, 0, 1, 0, 0, 0]	4, 7
[0, 1, 1, 1, 0, 0, 1, 0, 0, 0]	2, 3, 4, 7
[0, 1, 0, 0, 0, 0, 1, 0, 0, 0]	2, 7
[0, 1, 0, 1, 1, 0, 1, 0, 0, 0]	2, 4, 5, 7
[0, 1, 0, 1, 0, 1, 1, 0, 0, 0]	2, 4, 6, 7
[0, 1, 0, 1, 0, 0, 0, 0, 0, 0]	2, 4
[0, 1, 0, 1, 0, 0, 1, 1, 0, 0]	2, 4, 7, 8
[0, 1, 0, 1, 0, 0, 1, 0, 1, 0]	2, 4, 7, 9
[0, 1, 0, 1, 0, 0, 1, 0, 0, 1]	2, 4, 7, 10

Computation of Objective Values

Once the neighborhood N(X) is created, each neighbor solution X'_s ($s = 1,\ldots, n$) should be evaluated by computing the corresponding objective value (profit function). To perform this computation, H-TS calls model-3 for each X'_s and acquire the optimum profit which corresponds to the promotion campaign represented by X'_s. If X'_s results in an infeasible solution, model-3 detects this situation and corresponding profit is recorded as a negative value, other words, infeasible solutions are eliminated from the N(X).

The best solution in N(X) which has the highest profit among others is recorded as the new current solution, if this solution is not obtained by a tabu move. Otherwise the next best solution, which does not contain a tabu move, is replaced by the current solution. Classification of the moves whether tabu or not is explained in the next subsection.

Short-term memory

If the current solution X is generated by shifting x_j from 1 to 0 then shifting the same x_j from 0 to 1 is classified as tabu during the immediately succeeding tt iterations.

If the current solution X is generated by shifting x_j from 0 to 1 then shifting the same x_j from 1 to 0 is classified as tabu during the immediately succeeding tt iterations.

This classification scheme is used to prevent the current solution from cycling.

There are two tabu lists in H-TS which constitute the short-term memory of the algorithm. One is tabu add list (tal), other is tabu drop list (tdl). Tabu lists include the move that is forbidden to add or drop. If a move is in the tabu add list, it is forbidden to add this move throughout the next tt iterations. If a move is in the tabu drop list, it is forbidden to drop this move throughout the next tt iterations. Lists tal and tdl hold critical information to define the neighbors which let H-TS to visit better solution spaces.

Tabu tenure, tt, is the single parameter of H-TS and should be tuned before the application of the algorithm. Pre-experiments on the test set, which includes varying product sizes such as 5, 10, 15, 20, 30, 40 products, show that when tt is set to the nearest and larger integer to $\sqrt{n}$, H-TS gives rather good results. Table 5.3 shows tabu tenure for different product size.

Table 5.3 Tabu tenure with product size

Product size, n	Tabu Tenure, tt
5	3
10, 15	4
20	5
30	6
40	7

By using tabu lists, tal and tdl, recording and classifying tabu moves are managed as explained below:

Current solution $X = [x_1, x_2,.. x_r.., x_n]$ is replaced by a neighbor solution $X'_i = [x_1, x_2,.. x'_r.., x_n]$ from $N(X)=\{X'_s \mid s = 1,\dots, n\}$, where $x_r' = \begin{cases} 1 & if\ x_r = 0 \\ 0 & if\ x_r = 1 \end{cases}$ at iteration t.

After this replacement tabu lists are updated in the following manner: If $x'_r = 1$ then tdl[r] = t, otherwise tal[r] = t.

For the next iterations, t, any trial neighbor solution X'_i is checked to see whether it contains a tabu move or not. If X'_i is obtained by an add move, $x'_r = 1$, and if tal[r] + tt ≤ t then X'_i is discarded since it contains a tabu add move. If X'_i is obtained by a drop move, $x'_r = 0$, and if tdl[r] + tt ≤ t then X'_i is discarded since it contains a tabu drop move.

Following example further explains the management of tabu lists for tabu tenure, tt = 3:

Current solution: X= [0, 1, 0, 1, 0, 0, 1, 0, 0, 0], t = 1

The best solution from N(X) is X'_3 = [0, 1, 1, 1, 0, 0, 1, 0, 0, 0]. Since tal is empty initially this neighbor is replaced by X and tdl is updated as tdl[3] = 1

Current solution: X= [0, 1, 1, 1, 0, 0, 0, 0, 1, 0], t = 5

The best solution from N(X) is X'_3 = [0, 1, 0, 1, 0, 0, 1, 0, 0, 0]. Since tdl[3] + 3 ≤ 5, converting x_3 from 1 to 0 is tabu and X'_3 is discarded. H-TS looks for the next best neighbor using the same method.

Termination Criterion

If termination criteria do not exist, TS can iterates forever, therefore one termination criterion is needed. In this study, maximum number of iteration is used. Since the initial solutions are close to the optimum solutions as seen from the results of H-R1 and H-R2, even for the small numbers of maximum number of iteration.

Step 1: Initializing

Step 1.1: Set the initial solution, $X \leftarrow X^0$, and initial profit, $f(X) = f(X^0)$

Step 1.2:Set the best solution $X^b \leftarrow X^0$, $f(X_b) = f(X^0)$

Step 1.3: Set tabu tenure = tt; tal = {}, tdl = {}, t = 1

Step 2: Searching

Step 2.1. Generate $N(X) = \{ X'_s \mid s = 1,\ldots, n\}$, $S = \{1,\ldots, n\}$

Step 2.2. Call model-3 and get $f(X'_s)$ for each X'_s

Step 2.3. Find the best neighbor X'_i, $i \in S$. If S = {} then stop

Step 3: Tabu checking

Step 3.1: If X'_i is generated by the move $x'_r = 1$ and if if tal[r] + tt $\leq$ t then re-set $S = S - \{i\}$ and go to step 2. 3

Step 3.2: If X'_i is generated by the move $x'_r = 0$ and if if tdl[r] + tt $\leq$ t then re-set $S = S - \{i\}$ and go to step 2. 3

Step 4: Updating

Step 4.1. $X \leftarrow X'i$, $f(X) = f(X'i)$

Step 4.2. If X'_i contains an add move $x'_r = 1$ then tdl[r] = t

Step 4.3. If X'_i contains a drop move $x'_r = 0$ then tal[r] = t

Step 4.4. If $f(X) > f(X^b)$ then $X^b \leftarrow X$ and $f(X^b) = f(X)$

Step 5: Terminating

Step 5.1: t = t + 1, If t $\geq$ maximum iteration then stop and X^b and $f(X^b)$

Figure 5.7 Steps of H-TS heuristic

H-TS gives high qualified solutions. Pre-experiments show that when termination criterion is set to 30 iterations, H-TS performs well. The steps of H-TS are given as above.

5.5. Results and Computational Experiments

In this section, H-R1 and H-R2 are compared with each other, statistically, using the test instances from the literature. Both approaches are also examined by comparing with the proposed methods in the literature and testing on very large sized instances of the product targeting problem. Moreover, in this section, H-TS heuristic is shown experimentally that it substantially increases quality of the solutions generated by the rule 1 and rule 2. All the proposed algorithms have been coded in Java using Netbeans and Lingo 15.0. All instances are tested on Lenovo X1 personal computer with 2.5 GHz CPU clock speed and 4 GB RAM, equipped with Windows 7 Ultimate.

5.5.1. Comparison of H-R1 with H-R2

The test problems taken from Nobibon et al. (2011) are described in Table 5.4. For example, S2-10-10-2-s means 200 customers, 0.10 hurdle rate, 10 products, average budget and small maximum offer.

Table 5.4 Description of the test set

Sample data notation	Description
m (client size)	S1-S2-S3 (100-200-300 customers) , M1-M2 (1000 - 2000 customers), L (10000 customers)
R (hurdle rate)	0.5 , 0.10 , 0.15
n (Product size)	5, 10, 15
B (budget)	1 small, 2 average, 3 large
M (maximum offer)	s small, l large

Table 5.5 represents the performance results of H-R1 and H-R2. As seen from the table, rule 1 is better than rule 2 in terms of both computational time and solution quality measured as the percentage deviation from the optimal solution, Δ. Table 5.5 represents that H-R1 and H-R2 show similar computational time performances for the problems up to 1000 clients. However, H-R2 requires more times than H-R1 for the remaining problems except the problems with 5-product. Additionally, while the average computer

time of H-R2 is approximately 74% of the average time of solving model-1 optimally, it is only 30% for H-R1. Figure 5.8 shows the solution qualities acquired by the rules according to the product sizes. As seen from the figure, H-R1 gives better solutions than H-R2 for each product size but Δ increases as the product size increases for both of the rules.

Table 5.5 Comparison of H-R1 with H-R2

		Model-1	H-R1		H-R2	
Client size, m	Product size, n	Time (s)	Time (s)	Δ	Time (s)	Δ
	5	0.59	0.22	0.15	0.22	2.02
100	10	1.54	0.35	2.38	0.37	4.84
	15	3.48	0.87	2.06	0.84	7.15
	5	1.22	0.50	0.86	0.50	2.20
200	10	2.81	0.98	1.83	0.95	5.86
	15	15.95	2.87	1.82	2.37	5.62
	5	4.14	0.83	0.31	0.85	5.73
300	10	17.96	1.99	0.71	1.56	4.31
	15	21.99	3.81	2.86	4.76	7.45
	5	27.87	7.20	0.40	7.05	1.77
1000	10	63.16	7.77	0.86	6.39	4.37
	15	133.53	26.86	2.28	19.47	3.97
	5	52.92	26.37	0.000	19.28	2.88
2000	10	165.42	35.20	0.91	91.43	3.92
	15	296.32	88.95	1.51	201.09	3.48
	5	559.91	459.64	1.24	219.25	4.06
10000	10	2109.92	1395.50	1.41	2049.74	3.47
	15	7607.88	1952.40	1.60	5437.75	4.09
Mean		615.92	222.91	1.29	447.99	4.29
Standard dev.		1814.29	547.02	0.82	1334.07	1.62
Coefficient of variation		2.95	2.45	0.64	2.98	0.38

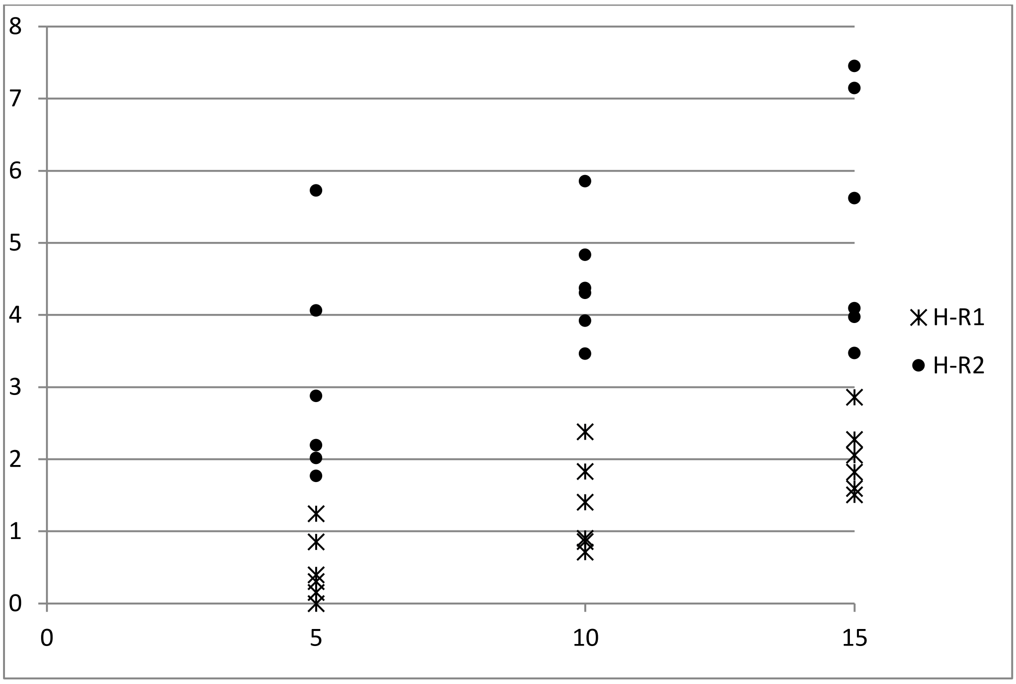

Figure 5.8 Solution quality comparison of H-R1 and H-R2 respect to the m.

In Table 5.6, the solution quality performances of H-R1 and H-R2 are classified into factors of the problem. As seen from the table, rule 1 shows a considerable improvement in comparison with rule 2 in terms of Δ.

Table 5.6 Solution qualities of H-R1 and H-R2 according to factors of the problem

Factors	Level	H-R1	H-R2	Improvement %
	5	0.49	3.11	
Product size, n	10	1.35	4.47	71.93
	15	2.02	5.31	
	100	1.53	4.67	69.52
	200	1.50	4.56	
Customer size, m	300	1.29	5.83	
	1000	1.18	3.37	
	2000	0.80	3.43	
	10000	1.41	3.88	
	s	1.61	2.51	
Maximum offer, M	l	0.96	6.08	60.13
	1	0.90	4.77	
Budget, B	2	1.17	2.96	68.92
	3	1.81	5.16	
	5	1.58	4.41	
Hurdle rate, R	10	1.20	5.09	69.53
	15	1.08	3.38	

Meanwhile, to compare rule 1 and rule 2 statistically, we performed the Wilcoxon Signed Ranks test. The test is designed to test a hypothesis about the mean of a population distribution. This test does not require the assumption that the population is normally distributed. It often involves the use of matched pairs, in this study H-R1 and H-R2, to test a mean difference of zero. The result of the statistical analysis at the significant level of .05 shows that there is no statistically significant difference between the run times of the rules where the corresponding p-value is .845. On the other hand, Δ generated by H-R1 is significantly smaller than Δ of H-R2 since the p-value is obtained as being closed to zero.

Table 5.7 Results of statistical analysis for comparing of H-R1 and H-R2

Comparison	Mean Difference	p-value
$Time_{MPH-R1} - Time_{MPH-R2} < 0$	-.196	.845
$\Delta_{MPH-R1} - \Delta_{MPH-R2} < 0$	-3.724	.000

5.5.2. Analyzing of the Factors' Effects

Since H-R1 outperforms H-R2 in terms of solution quality, in this subsection the main factors which may affect the performance of H-R1 are analyzed. As given in the **Table 5.4**, the factors of the product targeting problem are listed as product size (*n*), client size (*m*), maximum offer (*M*), budget (*B*) and hurdle rate (*R*). To understand the source of the variation in the solution quality and run time of H-R1, ANOVA (analysis of variation) is implemented, separately.

Table 5.8 shows that all factors, except *m* and *R*, have a statistically significant effect on the solution quality at a significance level of 5%. On the other hand, all the factors but B and M, affect the run time of H-R1. These results are used to generate very large scale instances of the problem, described in Subsection 5.5.3 by fixing *R*, *B* and *M*, and allowing different levels of other factors.

Table 5.8 Statistical analysis of performance of H-R1

ANOVA results for the solution quality					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	291.366[a]	12	24.280	3.363	.000
Intercept	597.396	1	597.396	82.753	.000
Customer size, *m*	29.899	5	5.980	.828	.530
Hurdle rate, *R*	8.937	2	4.468	.619	.539
Product size, *n*	164.148	2	82.074	11.369	.000
Budget, *B*	68.876	2	34.438	4.770	.009
Maximum offer, *M*	19.190	1	19.190	2.658	.104
Error	2237.899	310	7.219		
Total	3126.642	323			
Corrected Total	2529.265	322			

ANOVA results for the run time					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7.576E7	12	6313676.083	31.359	.000
Intercept	1.584E7	1	1.584E7	78.689	.000
Customer size, *m*	6.983E7	5	1.397E7	69.369	.000
Hurdle rate, *R*	2142335.538	2	1071167.769	5.320	.005
Product size, *n*	3644235.555	2	1822117.777	9.050	.000
Budget, *B*	115639.391	2	57819.695	.287	.751
Maximum offer, *M*	30293.815	1	30293.815	.150	.698
Error	6.262E7	311	201334.433		
Total	1.542E8	324			
Corrected Total	1.384E8	323			

5.5.3. Comparison of H-R1 and H-R2 with the Literature

Nobibon et al. (2011) formulate the product targeting problem as given in model -1 (equation 4.1-4.8) and also propose an alternative set-covering model in which each product is associated with a subset of clients in the optimal solution. They also develop a branch-and-price algorithm for solving it. For large size problems, Nobibon et al. (2011)

presents eight heuristics including branch-and-price (B&P) heuristics and tabu search. Among these heuristics the B&P based heuristic called H5, and the tabu search heuristic called H8 are reported to be the most effective algorithms in terms of solution quality and the feasible solution found. Thus, H-R1 and H-R2 are compared with heuristics H5 and H8 on the moderate and large sized instances given in Table 5.9. Table 5.9 shows the average results over the instances for each pair of the customer sizes and product sizes.

In the table, the best performances are represented in bold. It is apparent that H-R1 is strongly outperforms others in terms of Δ, while the next best Δ is generated by H-R2. As explained in section 3, both H-R1 and H-R2 heuristics rely on the separation of the product targeting problem into two sub-problems: selection of products in a promotion campaign (phase I) and the distribution of the selected products to customers (phase II). Therefore, there are mainly two reasons which make H-R1 and H-R2 better than others from the point of solution quality:

1) taking advantage of model-2, which gives a good approximation of the exact optimum objective value as explained in subsection 5.2.2, to select products in a campaign.

2) employing of an integer programming model, model-3, to distribute the selected products to customers optimally.

However, H5 and H8 heuristics deal with the complete problem and aim to find good feasible solutions to the problem. Table 5.9 also gives run time performances of the heuristics. Since H5 basically utilizes the B&P optimization procedure to find heuristic solutions, its run time requirement is more than other heuristics. As seen from the table, the run time results of H8 and H-R2 are close to each other. Both the heuristics give the best run time in four combinations of product-customer sizes out of a total of nine combinations (5, 10, 15-product size and 1000, 2000, 10000-customer size) while H-R1 is better than others in only the combination of 10-product and 2000-customer.

Table 5.9 Comparison of H-R1 and H–R2 with heuristics H5 and H8

	m	1000			2000			10000		
	n	5	10	15	5	10	15	5	10	15
H5	Δ	8.72	13.04	14.43	12.40	12.84	12.01	13.60	24.54	34.26
	Time	2441.7	3318.0	3274.6	1978.8	3565.3	3384.9	3730.1	3711.2	3634.3
H8	Δ	7.22	8.54	7.60	9.75	9.58	9.11	10.86	11.04	10.23
	Time	16.1	11.6	15.4	56.6	50.3	**67.2**	1268.3	**1347.2**	**1149.3**
H- R1	Δ	**0.40**	**0.86**	**2.28**	**0.00**	**0.91**	**1.51**	**1.24**	**1.41**	**1.60**
	Time	7.2	7.77	26.86	26.37	**35.20**	88.95	459.6	1385.5	1952.4
H-R2	Δ	1.77	4.37	3.97	2.88	3.92	3.48	4.06	3.47	4.09
	Time	**7.05**	**6.39**	**19.47**	**19.28**	91.43	201.1	**219.25**	2049.7	5437.7

A major difficulty in the application of heuristic approaches to a real-life case arises from the size of the problem in terms of the number of customers and products. That is why, to show the applicability of H-R1 and H-R2 to real-life cases we generated additional instances with 15000 and 20000 customers and 10, 20, 30, and 40 products using the same rules defined by Nobibon et al. (2011). It is also remarkable that Nobibon et al. (2011) derived these rules by examining the real-life data. Table 5.10 describes the new instances and also gives the best solutions and associated run times generated by H-R1 and H-R2. According to the results, H-R1 gives more qualified solutions in 35 instances out of 40 compared with H-R2, while the run time requirement of the both heuristics is close to each other. Average run time of H-R1 is less than one hour even for 20000 customers. Additionally, H-R1 and H-R2 are tested on instances with 40000-customer, and 5, 10, 15-product. The experiments demonstrate that average run time is 0.7 hour and the worst run time is 2 hours for H-R1, while these outcomes are 0.5 hour and 1.4 hours, respectively, for H-R2. These results also encourage application of both the approaches to real-life cases.

Table 5.10 Performances of H-R1 and H-R2 on the very large sized instances

Instances with m = 15000				H-R1		H-R2	
R	n	B	M	z^*_{H-R1}	Time (s)	z^*_{H-R2}	Time (s)
10	10	2	s	402793	216.8	402793	230.2
				430621	219.5	430621	244.7
				391041	222.0	391041	739.3
				376015	132.5	370921	1012.7
				423953	224.8	421198	1443.7
10	20	2	s	582660	757.8	577976	762.2
				587669	2331.7	557409	2331.4
				595134	764.3	563537	4045.7
				585756	802.3	519014	5086.6
				562717	765.0	529328	6648.5
10	30	2	s	700901	2180.8	700901	1686.0
				706801	1673.0	701653	4821.8
				672557	2025.5	666654	991.2
				679424	1751.1	679678	1655.4
				681505	1950.2	673321	3407.7
10	40	2	s	759228	3466.5	755470	2267.0
				729940	5366.1	723247	1797.3
				740488	2910.6	736306	1635.8
				718808	2892.0	712324	1682.2
				746606	2881.7	741716	1693.7
				603730.9	1676.7	592755.4	2209.2

Table 5.11 Performances of H-R1 and H-R2 on the very large sized instances –cont.

Instances with m = 20000				H-R1		H-R2	
R	n	B	M	z^*_{H-R1}	Time (s)	z^*_{H-R2}	Time (s)
10	10	2	s	495348	383.3	392940	252.4
				462338	713.6	168894	500.4
				206096	224.3	75643	244.8
				413503	612.4	413503	551.0
				185648	209.9	185648	254.6
10	20	2	s	752637	2636.1	746941	939.5
				750777	4256.8	728098	1911.5
				693131	1579.9	687474	1269.6
				751662	1643.9	752205	1929.9
				770709	4275.4	752548	3213.8
10	30	2	s	988641	3559.0	983843	3404.4
				880891	2360.3	887265	3540.5
				924886	3694.1	925056	3515.1
				903278	3685.9	901952	9086.3
				931762	9305.6	915708	14874.4
10	40	2	s	960203	5689.3	949538	6404.4
				999674	5930.1	995976	5891.5
				951009	5629.2	937709	6362.8
				970120	6168.5	968067	6670.6
				983362	5552.1	984879	5690.7
				748783.8	3405.5	717694.4	3825.4

5.5.4. Computational experiments of H-TS

In this section, H-TS are compared with H-R1, H-R2 and proposed algorithms in the literature on the test instances. All approaches are also examined on very large sized instances of product targeting problem. H-TS is employed to search the solution space of the product selection sub-problem of the product targeting problem. The initial solution set of H-TS is taken from the solutions of rule 1 and rule 2. Therefore, H-TS has been run on 166 data sets starting with the initial solutions generated by the rule 1 and 258 data sets starting with the initial solutions by the rule 2. The solution quality of H-TS shows a substantial improvement on the solutions obtained from H-R1 and H-R2. The comparative results are given in the following subsections.

5.5.5. Comparison of H-TS with H-R1 and H-R2

Performances of H-R1 and H-R2 are examined on all the test problems taken from Nobibon et al. (2011) and explained in detail in previous subsections. On the other hand, H-TS is tested on the instances which couldn't have been solved optimally by H-R1 and/or H-R2 since the main purpose of H-TS is to improve the performances of H-R1 and H-R2 further. Table 5.12 shows the performance results of H-TS including the maximum and the average number of iteration at which the best solution is found and the percentage deviation from the optimum solution, Δ.

The experimental results indicate that H-TS performs better when it utilizes the solution obtained by rule 1 as an initial solution comparing with the initial solution provided by rule 2. H-TS with initial solution by rule 1 (H-TS-rule1) gives 0.12 of percentage deviation from optimum solution notated by Δ while H-TS with rule 2 (H-TS-rule2) results in 0.27 of the deviation. It can be concluded that the initial solution from the rule 1 gives rise to closer solutions the optimum then rule2 within the termination criteria of H-TS.

Table 5.12 Comparison of H-TS with H-R1 and H-R2 with respect to solution quality.

		HR-1	HR-2	H-TS-rule1			H-TS-rule2		
Client size. *M*	Product size, *n*	Δ	Δ	Δ	Max. #of iterations until the best sol.	Avg.#of iterations until the best sol	Δ	Max. #of iterations until the best sol.	Avg.#of iterations until the best sol
	5	0.15	2.02	0.00	2	2	0.00	6	127
100	10	2.38	4.84	0.13	3	1.2	0.19	23	2.80
	15	2.06	7.15	0.00	13	3,5	0.00	13	6.33
	5	0.86	2.20	0.06	5	0.66	0.03	5	1
200	10	1.83	5.86	0.07	3	0.83	0.00	23	3.22
	15	1.82	5.62	0.00	4	1.83	0.03	6	3.22
	5	0.31	5.73	0.04	2	0.33	0.02	8	1.72
300	10	0.71	4.31	0.03	10	1.27	0.04	4	1.61
	15	2.86	7.45	0.00	4	2	0.10	6	3
	5	0.40	1.77	0.10	2	0.2	0.00	2	0.6
1000	10	0.86	4.37	0.00	2	0.6	0.08	26	2.5
	15	2.28	3.97	0.00	3	1	0.05	5	2.5
	5	0.000	2.88	0.00	0	0	0.16	5	0.9
2000	10	0.91	3.92	0.12	2	0.5	0.24	10	2.20
	15	1.51	3.48	0.21	8	1.3	1.27	25	3.7
	5	1.24	4.06	0.00	2	0.33	0.05	5	1
10000	10	1.41	3.47	0.00	10	1.66	0.06	5	1.66
	15	1.60	4.09	0.33	3	1	0.44	7	3.2
Mean		1.29	4.29	0.06			0.15		
Standard dev.		0.82	1.62	0.09			0.30		
Coefficient of variation		0.64	0.38	1.49			1.96		

As explained before, H-TS is to be ended in 30 iterations. In order to show the performance of computational time, the amount of time that elapsed until the best solution is also reported.

Table 5.13 includes both solution quality and computational time of all proposed algorithms in this study. There are 18 different test sets depicted in the table below. Tabu search heuristic, H-TS-Rule1, which uses initial data resulting from H-R1, deviates optimum solution from 0.06 percentage. Also, in 8 different test sets (out of 18), H-

TS-Rule1 finds exactly the same optimum value as the model-1. H-TS-Rule 2 approximate the optimum solution to 0.15 percentage deviation.

Table 5.13 Solution quality and run time performances of the proposed algorithms

		Model-1	H-R1		H-R2		H-TS-rule1				H-TS-rule2			
Client size. m	Product size. n	Time (s)	Time (s)	Δ	Time (s)	Δ	Init. time	Elapsed Time (s)	Total time	Δ	Init time	Elapsed Time (s)	Total time	Δ
	5	0.59	0.22	0.15	0.22	2.02	0.11	0.12	0.24	0.00	0.11	0.18	0.29	0.00
100	10	1.54	0.35	2.38	0.37	4.84	0.16	0.28	0.44	0.13	0.17	0.55	0.72	0.19
	15	3.48	0.87	2.06	0.84	7.15	0.44	1.55	1.99	0.00	0.44	2.31	2.75	0.00
	5	1.22	0.50	0.86	0.50	2.20	0.26	0.32	0.58	0.06	0.26	0.36	0.62	0.03
200	10	2.81	0.98	1.83	0.95	5.86	0.49	0.63	1.12	0.07	0.48	1.5	1.98	0.00
	15	15.95	2.87	1.82	2.37	5.62	1.42	3.02	4.44	0.00	1.33	3.5	4.83	0.03
	5	4.14	0.83	0.31	0.85	5.73	0.39	0.49	0.88	0.04	0.46	0.76	1.24	0.02
300	10	17.96	1.99	0.71	1.56	4.31	1.09	1.52	2.61	0.03	0.8	1.32	2.12	0.04
	15	21.99	3.81	2.86	4.76	7.45	2.19	2.82	5.01	0.00	2.7	5.91	8.61	0.10
	5	27.87	7.20	0.40	7.05	1.77	3.98	3.55	7.53	0.10	3.84	4.13	7.97	0.00
1000	10	63.16	7.77	0.86	6.39	4.37	2.67	5.21	7.88	0.00	3.34	12.06	15.4	0.08
	15	133.53	26.86	2.28	19.47	3.97	15.78	17.31	33.09	0.00	10.24	24.6	34.84	0.05
	5	52.92	26.37	0.000	19.28	2.88	14.82	11.56	26.38	0.00	10.7	11.38	22.08	0.16
2000	10	165.42	35.20	0.91	91.43	3.92	20.57	15.38	35.95	0.12	51.6	102.38	153.98	0.24
	15	296.32	88.95	1.51	201.09	3.48	50.33	78.73	129.06	0.21	112.58	407.38	519.96	1.27
	5	559.91	459.64	1.24	219.25	4.06	249	224.84	473.84	0.00	115	175.34	290.34	0.05
10000	10	2109.92	1395.50	1.41	2049.74	3.47	748	814.74	1562.74	0.00	947	1937	2884	0.06
	15	7607.88	1952.40	1.60	5437.75	4.09	1039	1142	2181	0.33	3260	6163	9423	0.44
Mean		615.92	222.91	1.29	447.99	4.29	119.50	129.12	248.60	0.06	251.17	491.87	743.04	0.15
Standard dev.		1814.29	547.02	0.82	1334.07	1.62	291.80	318.58	610.25	0.09	782.82	1486.96	2269.5	0.30
Coefficient of variation		2.95	2.45	0.64	2.98	0.38	2.44	2.47	2.45	1.49	3.12	3.02	3.05	1.96

Table 5.14 summarizes the solution quality according to products size, customer size and other factors of the product targeting problem. As seen from the table, the approximation of H-TS to the optimum is better than H-R1 and H-R2. Of course, it is an expected result since we designed H-TS to improve the performances of H-R1 and H-R2. For example, when testing with 10-product instance, the deviation Δ is obtained 1.35 from H-R1 and 4.47 from H-R2. As mentioned before, H-TS is tested on the instances by Nobibon et al. (2011) which couldn't have been solved optimally by H-R1 and/or H-R2. In other words, H-TS looks for better solution quality starting from initial solution set found by H-R1 and H-R2.

Table 5.14 Solution qualities of H-R1, H-R2, and H-TS according to factors of problem

Factors	Level	H-R1	H-TS-rule1	Improvement	H-R2	H-TS-rule2	Improvement %
	5	0.49	0.03	%95	3.11	0.04	%98
Product size, n	10	1.35	0.06	%96	4.47	0.10	%97
	15	2.02	0.09	%96	5.31	0.32	%93
	100	1.53	0.04	%97	4.67	0.06	%98
	200	1.50	0.04	%97	4.56	0.02	%99
Customer size, m	300	1.29	0.02	%98	5.83	0.05	%99
	1000	1.18	0.03	%97	3.37	0.04	%98
	2000	0.80	0.11	%86	3.43	0.56	%83
	10000	1.41	0.11	%92	3.88	0.19	%95
	S	1.61	0.06	%96	2.51	0.11	%95
maximum offer, M	L	0.96	0.05	%94	6.08	0.20	%96
	1	0.90	0.10	%88	4.77	0.25	%94
Budget, B	2	1.17	0.07	%94	2.96	0.20	%93
	3	1.81	0.01	%99	5.16	0.01	%99
	5	1.58	0.05	%96	4.41	0.05	%98
Hurdle rate, R	10	1.20	0.08	%93	5.09	0.30	%94
	15	1.08	0.04	596	3.38	0.11	%96

The superiority of H-TS to H-R1 and H-R2 is also figured out from Figures 5.9 – 5.13. Figure 5.9 shows the solution qualities of the heuristics accompanied by the product size, Figure 5.10 by the customer size, Figure 5.11 by the maximum offer quantity, Figure 5.12 by the budget, and finally Figure 5.13 by the hurdle rate. In each case H-TS is clearly better than other heuristics.

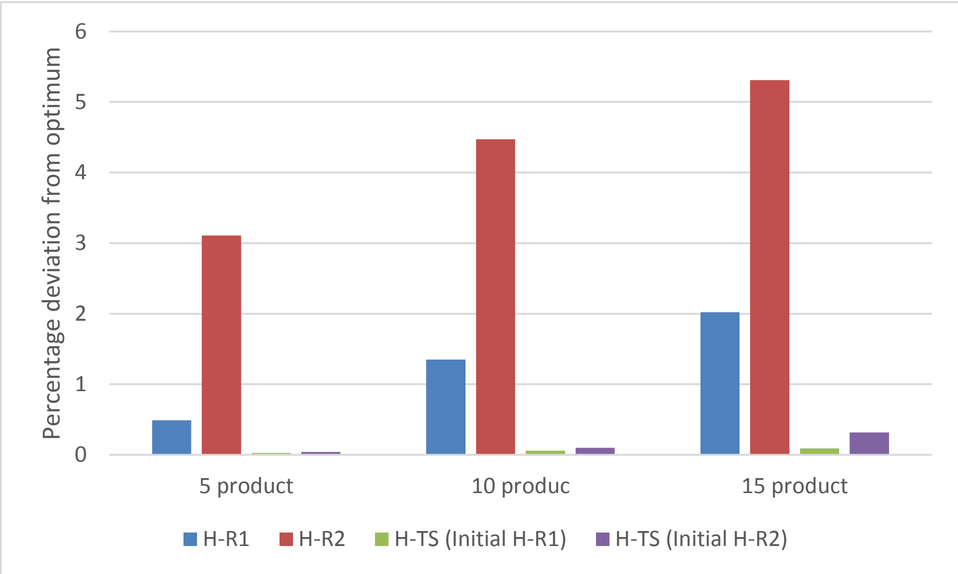

Figure 5.9 Solution qualities of the heuristics according to the product size

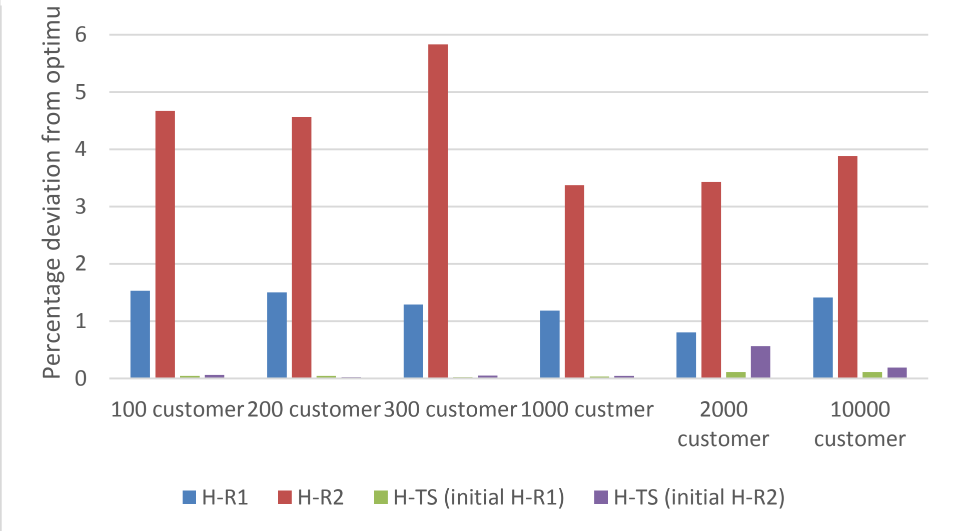

Figure 5.10 Solution qualities of the heuristics according to the size

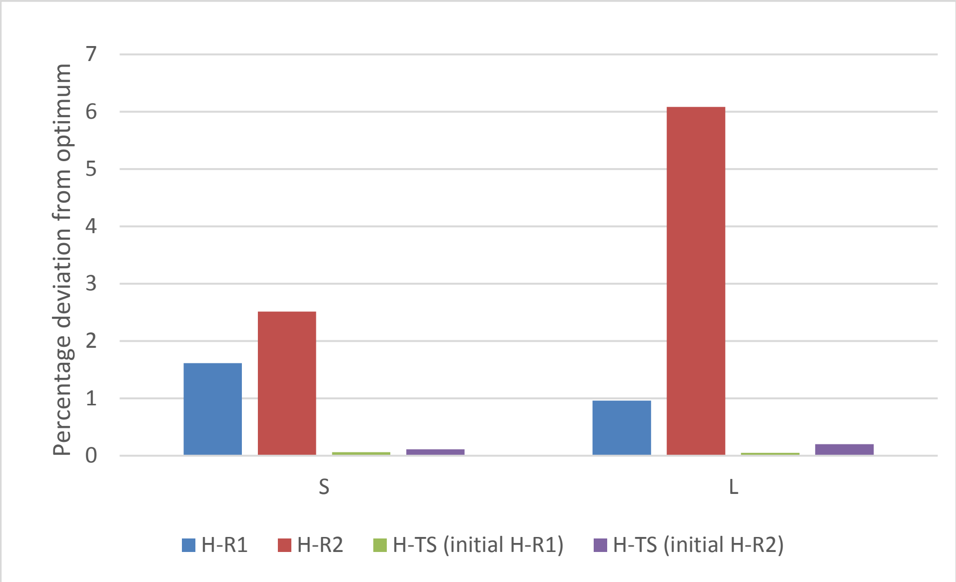

Figure 5.11 Solution qualities of the heuristics according to the maximum offer quantity

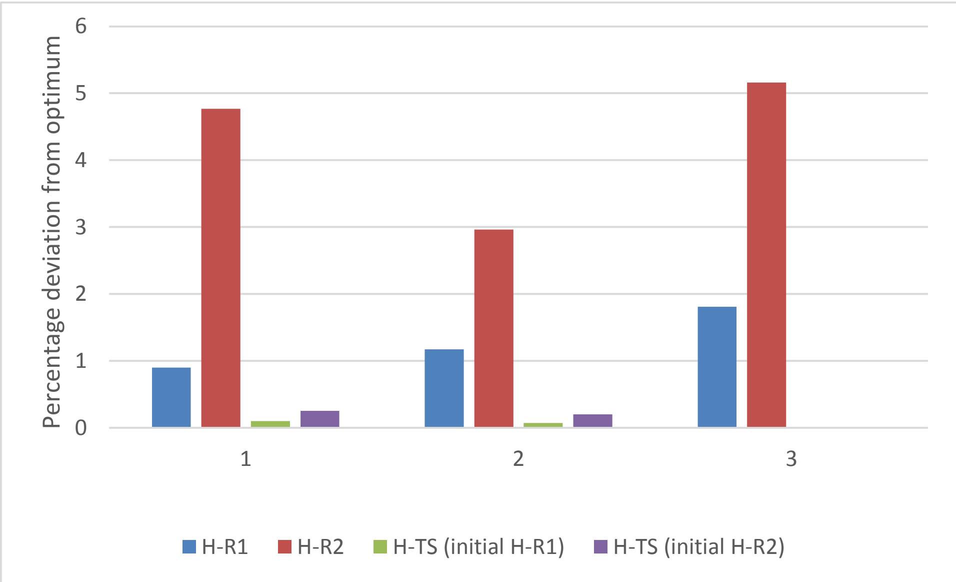

Figure 5.12 Solution qualities of the heuristics according to the budgets

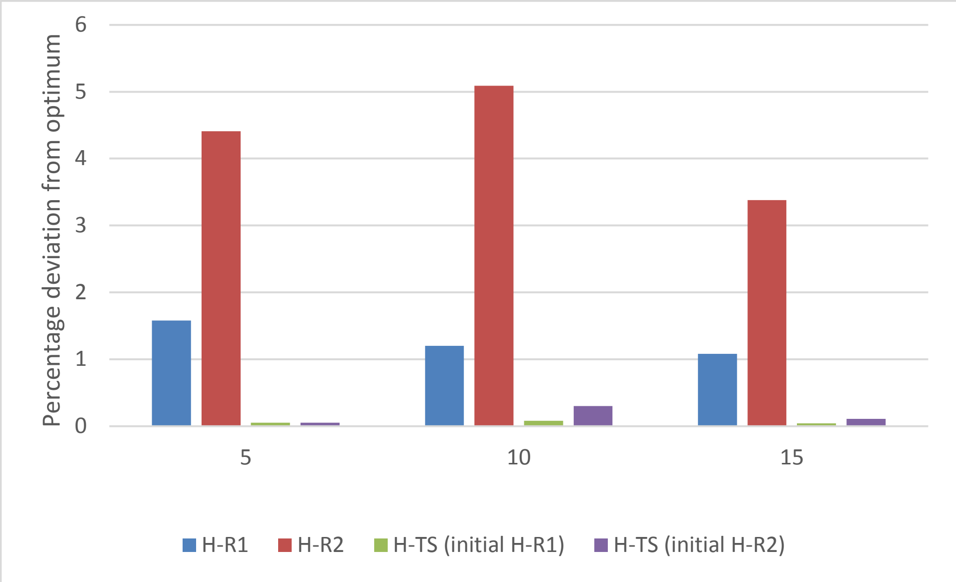

Figure 5.13 Solution qualities of the heuristics according to the hurdle rate

In Table 5.15, H-TS, H-R1 and H-R2 are also compared with the heuristics proposed in the literature. Among these heuristics B&P based heuristic called H5, and tabu search heuristic called H8 are reported to be the most effective algorithms in terms of solution quality and feasible solution found by Nobibon et al. (2011). Table 5.15 shows the average results over the instances for each pair of the customer sizes and product sizes. As seen from the table, H-TS with rule 1 is strongly outperforms others in terms of Δ. However, run time requirement of H-TS-rule 1 and H-TS-rule 2 is greater than run time of H-R1 and H-R2, respectively. This is because of the searching process of tabu search algorithm, the algorithm computes the optimum profits using model-3 for each neighbor solution. However, the average run time of H-TS-rule1 is nearly 8 hours and it is reasonable to use for planning of a promotion campaign.

Table 5.15 Comparison of H-TS, H-R1 and H–R2 with H5 and H8

	m	1000			2000			10000		
	N	5	10	15	5	10	15	5	10	15
H5	Δ	8.72	13.04	14.43	12.40	12.84	12.01	13.60	24.54	34.26
	Time	2441.7	3318.0	3274.6	1978.8	3565.3	3384.9	3730.1	3711.2	3634.3
H8	Δ	7.22	8.54	7.60	9.75	9.58	9.11	10.86	11.04	10.23
	Time	16.1	11.6	15.4	56.6	50.3	67.2	1268.3	**1347.2**	**1149.3**
H- R1	Δ	0.40	0.86	2.28	**0**	0.91	1.51	1.24	1.41	1.60
	Time	7.2	7.8	26.9	26.4	**35.2**	88.9	459.6	1385.5	1952.4
H-R2	Δ	1.77	4.37	3.97	2.88	3.92	3.48	4.06	3.47	4.09
	Time	**7.0**	**6.4**	**19.5**	**19.3**	91.4	201.1	**219.5**	2049.7	5437.8
H-TS-rule1	Δ	0.10	**0**	**0**	**0**	**0.12**	**0.21**	**0**	**0**	**0.33**
	Time	7.53	7.88	33.09	26.4	35.95	**129.06**	473.84	1562.7	2181
H-TS-rule2	Δ	**0**	0.08	0.05	0.16	0.24	1.27	0.05	0.06	0.44
	Time	7.97	15.4	34.84	22.08	153.98	519.96	290.34	2884	9423

Table 5.17 shows the solution quality (in terms of profit) and computational time (in terms of number of iterations until the best solution) performances of the algorithms proposed on the very large sized problems with 15000 and 20000 customers and 10, 20, 30, 40 products. Totally 40 instances (20 instance for 15000 customers and 20 instances for 20000 customers) are examined by each algorithm.

Apparently, H-TS increased the profit of both H-R1 and H-R2. Even though H-TS algorithm is initialized by the different solutions found by rule-1 and rule-2, the algorithm can find the same solutions in 26 instances among 40 instance in regardless of the initial solution.

The average profit for 15000 customers is 617253 according to H-TS-rule1 and 618774 to H–TS-rule2. Percentage difference between these results is only 0.2% and this shows that H-TS algorithm exhibits similar performance using the different initial solutions. On the other hand, the average iteration number of tabu search until to the best solution is computed as 2.35 for rule-1 and 4.35 for rule2. The similar results obtained on the 15000-customer instances are also valid for 20000-customer instances: The average profits are 760144 and 756918 by H-TS-rule1 and H-TS-rule2 respectively. The percentage difference between the average profits is %0.4. The average iteration number is 1.5 and 2.95 for H-TS-rule1 and H-TS-rule2 respectively. The experiments on the very large instances show that H-TS is robust against to the initial solution used. The algorithm gives the similar average profits using the different initial solutions found by H-R1 and H-R2. On the other hand, H-TS with rule 1 necessitates more iterations. Since each iteration of H-TS consumes notable computer time, from this point of view, H-TS-rule1 should be preferred.

Table 5.16 Comparison of H-TS with H-R1 and H-R2 on the very large sized instances

Instances with m = 15000				H-R1	H-R2	H-TS – rule 1		H-TS rule 2	
R	n	B	M	z^*_{H-R1}	z^*_{H-R2}	$z^*_{H-TS-rule1}$	Avg.#of iterations until the best sol	$z^*_{H-TS-rule2}$	Avg.#of iterations until the best sol
10	10	2	s	402793	402793	402793	0	402793	0
				430621	430621	430621	0	430621	0
				391041	391041	391041	0	391041	0
				376015	370921	378787	2	378787	2
				423953	421198	423953	0	423953	2
10	20	2	s	582660	577976	583718	1	583718	3
				587669	557409	599694	2	599694	6
				595134	563537	604634	2	604634	7
				585756	519014	596588	2	596588	10
				562717	529328	566590	1	566590	6
10	30	2	s	700901	700901	716962	3	716962	3
				706801	701653	719205	2	719205	4
				672557	666654	695385	4	695385	4
				679424	679678	699445	7	699445	9
				681505	673321	704578	4	704895	4
10	40	2	s	759228	755470	777767	3	780076	4
				729940	723247	757374	3	762264	4
				740488	736306	768364	3	781831	7
				718808	712324	750246	4	754813	6
				746606	741716	777315	4	782178	6
				603730.9	592755.4	617253	2.35	618774	4.35

Table 5.17 Comparison of H-TS with H-R1 and H-R2 on the very large sized instances-Cont.

Instances with m = 20000				H-R1	H-R2	H-TS – rule 1		H-TS rule 2	
R	n	B	M	z^*_{H-R1}	z^*_{H-R2}		Avg.#of iterations until the best sol		Avg.#of iterations until the best sol
10	10	2	s	495348	392940	495348	0	495348	2
				462338	168894	462341	1	462341	4
				206096	75643	206096	0	206096	2
				413503	413503	413503	0	413503	0
				185648	185648	185648	0	185648	0
10	20	2	s	752637	746941	764658	2	764568	4
				750777	728098	750777	0	750777	5
				693131	687474	693131	0	693129	2
				751662	752205	751666	1	752205	0
				770709	752548	784154	2	784154	14
10	30	2	s	988641	983843	1025609	5	1025609	6
				880891	887265	908180	4	905689	2
				924886	925056	946031	2	945819	3
				903278	901952	910575	2	910575	1
				931762	915708	956523	3	929429	3
10	40	2	s	960203	949538	986025	2	987691	3
				999674	995976	1012342	3	998012	2
				951009	937709	969096	1	969096	2
				970120	968067	984409	1	968419	1
				983362	984879	996767	1	990254	3
				748783.8	717694.4	760144	1.5	756918	2.95

THIS PAGE INTENTIONALLY LEFT BLANK

6. MANAGERIAL IMPLICATIONS AND CONCLUSIONS

In this thesis, the product targeting problem to plan a promotion campaign is addressed. Since the problem is NP-hard, optimization techniques are capable of solving only small or moderately sized problems. To deal with large sized instances, three heuristic approaches are suggested: H-R1, H-R2, and H-TS.

Among these approaches H-R1 and H-R2 solve the problem in a two-phase way. In the first phase, H-R1 optimizes a linear programming approximation model tailored to the problem to predict the targeted products (products selected for the campaign), while H-R2 utilizes dual of the linear programming model for prediction. In the first phase, both H-R1 and H-R2 generate the promotion campaign in a constructive manner. H-R1 and H-R2 offer the targeted products (the promotion campaign) to clients optimally in the second phase by solving a variation of the assignment problem. Since the proposed -R1 and H-R2 methods suggest the division of the problem into two phase, this makes large-sized problem solvable in an effective and efficient way. Moreover, modification of the existing model defined in the literature creates a new model. This new modified model can be used to catch the strong approximation to the optimal profit of the problem.

The third heuristic applied to the problem is tabu search, H-TS. The main motivation of using tabu search algorithm is to find better profits. H-TS searches the feasible region of the possible promotion campaigns relying on the principles of tabu search metaheuristic. H-TS computes the profit values for each trial promotion campaigns using the same assignment model which both H-R1 and H-R2 utilize in their second phase.

It is statistically shown that the H-R1 approach outperforms H-R2 from the point of solution quality whereas the both approaches perform similar in terms of computer times. Additionally, the results obtained from the experiments on a suite set of test problems indicate that both approaches are strongly superior to proposed methods in the literature in terms of solution quality within a shorter or reasonable computer time. The proposed approaches are also able to solve larger-sized problems up to 20000 clients and 40 products. H-TS gives the best results among the proposed algorithms in the literature and also among H-R1 and H-R2. On the other side, computer time requirement of H-TS is more than the other heuristics.

Consequently, the three methods suggested for solving of the product targeting problem are capable to find qualified solutions comparing with the existing methods in the related literature. While H-TS outperforms the others in terms of solution quality it needs more computational time. Therefore, depending on the available time to obtain a highly qualified solution the problem, H-TS heuristic is suggested. However, the performances of both H-R1 and H-R2 methods are quite remarkable to get good solutions to the product targeting problem.

7. REFERENCES

Ahn, H., Ahn, J.J., Oh, K.J., and Kim, D. H. (2011). Facilitating cross-selling in a mobile telecom market to develop customer classification model based on hybrid data mining techniques. *Expert Systems with Applications*, 38:5005–5001.

Berger, P. D., and Bechwati, N. N. (2001). The allocation of promotion budget to maximize customer equity. *OMEGA International Journal of Management Science*, 29:49-61

Bhaskar, T., Sundararajan, R. and Krishnan, P.G. (2009). A fuzzy mathematical programming approach for cross-sell optimization in retail banking. *Journal of the Operational Research Society*, 60:717-727

Blum, C., & Roli, A. (2003). Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys* (CSUR), 35(3), 268-308.

Bose, I.,and Chen, X. (2009). Quantitative models for direct marketing: A review from a systems perspective. *European Journal of Operational Research*, 195:1–16.

Bult, J. R. and Wansbeek, T. (1995). Optimal selection for direct mail, *Marketing Science*,14 (4):378–39

Chen, W.C., Hsu, C.C. and Hsu J.N., (2011). Optimal selection of potential customer range through the union sequential pattern by using a response model. *Expert Systems with Applications*, 38:7451–7461

Chun, Y. H. (2012). Monte Carlo analysis of estimation methods for the prediction of customer response patterns in direct marketing. *European Journal of Operational Research*, 217: 673–678

Cohen, M. (2004). Exploiting response models – Optimizing cross-sell and up-sell opportunities in banking. *Information Systems*, 29:327–341.

Crainic, T. G., & Gendreau, M. (2002). Cooperative parallel tabu search for capacitated network design. *Journal of Heuristics*, 8(6), 601-627.

Dengiz, B., & Alabas, C. (2000, December). Simulation optimization using tabu search. In Proceedings of the 32nd conference on Winter simulation (pp. 805-810). Society for Computer Simulation International.

De Reyck, D. and Degraeve Z. (2003). Broadcast scheduling for mobile advertising. *Operations Research*, 51:509–517

De Reyck, B., & Degraeve, Z. (2006). MABS: Spreadsheet-based decision support for precision marketing. *European journal of operational research*, 171(3), 935-950.

Direct Marketing Association, 2011, *The Power of Direct Marketing: ROI, Sales, Expenditures, and Employment in the US*. Direct Marketing Association, Incorporated.

Doyle, S. (2005). Business requirements for campaign management—A sample framework. *The Journal of Database Marketing & Customer Strategy Management*, 12(2), 177-192.

Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & operations research*, 13(5), 533-549.

Glover, F. (1990). Tabu search: A tutorial. *Interfaces*, 20(4), 74-94.

Glover, F., & Taillard, E. (1993). A user's guide to tabu search. *Annals of operations research*, 41(1), 1-28.

Glover, F., Laguna, M., & Marti, R. (2007). Principles of tabu search. *Approximation Algorithms and Metaheuristics*, 23, 1-12.

Glover, F. and Laguna, M., (1997). Tabu Search, Kluwer Academic Publishers, Boston.

Glover, F. (1989). Tabu search-part I. ORSA *Journal on computing*, 1(3), 190-206.

Gönül,F. F. and Hofstede, F. T (2006). How to Compute Optimal Catalog Mailing Decisions. *Marketing Science*, 25(1):65-74.

Guido, G. , Prete, M. I., Miraglia, S., Mare, I. D., (2011). Targeting direct marketing campaigns by neural networks, *Journal of Marketing Management*, 27:9-10, 992-1006,

Jain, D., Singh, S.S. (2002). Customer lifetime value research in marketing: A review and future directions. *Journal of interactive marketing*, 16(2): 34-46

Jonker, J. J., Piersma, N., and Van den Poel, D., (2004). Joint optimization of customer segmentation and marketing policy to maximize long-term profitability. *Expert Systems with Applications,* 27:159–168

Kim., Y. A., Song, H. S., and Kim., S. H. (2009). A new marketing strategy map for direct marketing. *Knowledge-Based Systems*, 22:327-335

Levin, N., & Zahavi, J. (1998). Continuous predictive modeling - a comparative analysis. *Journal of Interactive Marketing*, 12(2), 5-22.

Li, S., Sun, B. and Montgomery, A. L. (2011). Cross-Selling the Right Product to the Right Customer at the Right Time. *Journal of Marketing Research*, 48(4):683-700

Lu, T. Z. and Kun,W. (2007). Research on telecommunication cross-selling combinatorial optimization model, *International Conference on Management Science and Engineering 14Th*ISBN:978-0-646-47827-2

Madeira, S., & Sousa, J. M. (2002). Comparison of target selection methods in direct marketing. *In European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems*, July.

Metaheuristics network project, http://www.metaheuristics.net/ , 20 June 2015

Nair S. K., and Tarasewich, R. (2003) A model and solution method for multi-period sales promotion design. *European Journal of Operational Research*, 150 672–687

Nobibon, F. T., Leus, R. and Spieksma, F.C.R. (2011) Optimization models for targeted offers in direct marketing: Exact and heuristic algorithms. *European Journal of Operations Research*, 210:670-683

Olson, D.L. and Chae, B. (2012). Direct marketing decision support through predictive customer response modeling. *Decision Support Systems*, 54:443–451.

Petrison, L. A., Blattberg, R. C., & Wang, P. (1997). Database marketing: Past, present, and future. *Journal of Interactive Marketing*, 11(4), 109-125.

Prinzie, A. and Van Den Poel, D. (2005). Constrained optimization of data-mining problems to improve model performance: A direct-marketing application *Expert Systems with Applications*, 29:630–640

Reichheld, F. and Sasser, W. (1990). Zero defection: quality comes to services. *Harvard Business Review*, Sept-Oct:105–111

Reeves, C. R. (1996). Modern heuristic techniques. VJ Rayward-Smith, IH Osman, CR Reeves, GD Smith (eds.). Modern Heuristic Search Methods, Wiley, Chichester, 1-25.

Ryals, L, Dias, S. and Berger, M. (2007) Optimizing marketing spend: return maximization and risk minimization in the marketing portfolio. *Journal of Marketing Management*,23(9):991-1012

Schön, C. (2010).On the product line selection problem under attraction choice models of consumer behavior. *European Journal of Operational Research*, 206:260-26

Soopramanien, D.and Juan, L.H. (2010). The importance of understanding the exchange context when developing a decision support tool to target prospective customers of business insurance, *Journal of Retailing and Consumer Services*, 17:306–312

Sundararajan, S., Bhaskar, T., Sarkar, A., Dasaratha, S., Bal, D., Marasanapalle, J., Zmudza, B., and Bak, K. (2011) Marketing Optimization in Retail Banking. *Interfaces*, 41(5):485-505.

Thuring, F., Nielsen, J. P., Guillen, M. and Bolance, C. (2012). Selecting prospects for cross-selling financial products using multivariate credibility. *Expert Systems with Applications*, 39:8809-8816

Verhoef, P. C., Spring, P. N., Hoesktra, J. C., Leeflang, P. S. H. (2002), The commercial use of segmentation and predictive modeling techniques for database marketing in the Netherlands. *Decision Support Systems,* 34 : 471-481.

Venkatesan, R. and Kumar , V. (2004). A Customer Lifetime Value Framework for Customer Selections and Resource Allocation Strategy. *Journal of Marketing*, 68 (4):106-125.

Webster dictionary, http://www,webster-dictionary,org/definition/tabu, 30 June 2015

Wong, K. W., Zhou, S., Yang, Q., & Yeung, J. M. S. (2005). Mining customer value: From association rules to direct marketing. *Data Mining and Knowledge Discovery*, 11(1), 57-79.

Xu, K., Li, J. and Song, Y. (2012).Identifying valuable customers on social networking sites for profit maximization. *Expert Systems with Applications*, 39:13009-13018

THIS PAGE INTENTIONALLY LEFT BLANK

8. APPENDIX A

Model 4:

Minimize $$(1+R)\sum_{j=1}^{n} f_j u + \sum_{j=1}^{n} B_j w_j + \sum_{i=1}^{m} M_i s_i + \sum_{j=1}^{n} O_j t_j + \sum_{j=1}^{n} O_j r_j + \sum_{i=1}^{m}\sum_{j=1}^{n} l_{ij}$$

Subject to

$$\lfloor p_{ij} - (1+R)c_{ij} \rfloor u + c_{ij} w_j + s_i + t_j + l_{ij} \geq p_{ij} - c_{ij} \qquad i = 1,\cdots,m,\ j = 1,\cdots,n$$

$$(1+R)\frac{f_j}{O_j} u + \frac{B_j}{O_j} w_j + t_j + r_j \geq \frac{f_j}{O_j} \qquad j = 1,\cdots,n$$

$$u \leq 0$$
$$s_i \geq 0 \qquad i = 1,\cdots,m$$
$$w_j, r_j \geq 0, \quad t_j \leq 0 \qquad j = 1,\cdots,n$$
$$l_{ij} \geq 0 \qquad i = 1,\cdots,m,\ j = 1,\cdots,n$$

Dual variables (corresponding primal constraints in model-2):

u: hurdle rate constraint (eq. 5.2)

w_j: budget constraints for product j (eq. 5.3)

s_i: demand constraints for client i (eq. 5.4)

t_j: minimum number offering constraints for product j (eq. 5.5)

r_j: upper limit constraints on variables, $x_{m+1,j}$ (eq. 5.6)

l_{ij}: linear programming relaxation of binary variables x_{ij} (eq. 5.7)

9. CURRICULUM VITAE

Filiz Cetin is a project manager in Turkcell, Turkey's largest telecom operator in charge of CRM/BIS systems and services. Before Turkcell, Filiz worked as a business development specialist in TUBITAK, analyst in Finansbank, and project manager in Tellcom, respectively.

Filiz Çetin is studying about analytical modelling in Engineering Management Marmara University.