A Note on Mathematical Modelling of Practical Multicampaign Assignment and Its Computational Complexity

Yong-Hyuk Kim† and Yourim Yoon‡

† Department of Computer Science & Engineering, Kwangwoon University
447-1, Wolgye-dong, Nowon-gu, Seoul, 139-701, Korea
Homepage: http://soar.snu.ac.kr/~yhdfly
Email: yhdfly@kw.ac.kr

‡ School of Computer Science & Engineering, Seoul National University
Sillim-dong, Gwanak-gu, Seoul, 151-744, Korea
Email: yryoon@soar.snu.ac.kr

December 15, 2013

Abstract

Within personalized marketing, a recommendation issue known as multicampaign assignment is to overcome a critical problem, known as the multiple recommendation problem which occurs when running several personalized campaigns simultaneously. This paper mainly deals with the hardness of multicampaign assignment, which is treated as a very challenging problem in marketing. The objective in this problem is to find a customer-campaign matrix which maximizes the effectiveness of multiple campaigns under some constraints. We present a realistic response suppression function, which is designed to be more practical, and explain how this can be learned from historical data. Moreover, we provide a proof that this more realistic version of the problem is NP-hard, thus justifying the use of heuristics presented in previous work.

Keywords: Computational complexity, personalized marketing, multicampaign assignment.

1 Introduction

CRM (customer relationship management) [7] is important in obtaining and maintaining loyal customers. To maximize income and customer satisfaction, firms try to provide personalized services for customers. As personalized campaigns are frequently performed, several campaigns often happen to run at the same time or in a short period of time. It is often the case that an attractive customer for a specific campaign tends to be attractive for other campaigns. If the other campaigns are not considered when assigning customers to a specific campaign, some customers may be bombarded by a considerable number of campaigns [18] [19]. This problem is called the multiple recommendation problem. As the number of recommendations to a particular
customer increases, that customer’s interest in all campaigns decreases [3]. Eventually the rate of customer response for campaigns will drop. This clearly lowers the targeting efficiency as well as customer satisfaction, which leads to diminished customer loyalty. Unfortunately, most traditional approaches have focused on the effectiveness of individual campaigns and did not consider the problem of multiple recommendations. In the situation where several campaigns are conducted within a short period of time, we need to find the optimal assignment of campaigns to customers considering the recommendations in other campaigns.

The multicampaign assignment problem (MCAP) is a complex assignment problem in which each of \( n \) customers is assigned or not to each of \( k \) campaigns. The goal is to find a set of assignments such that the effect of the campaigns is maximized under some constraints. The main difference between this problem and with independent campaigns is that the customer response for campaigns is influenced by multiple recommendations. In previous work [17, 18, 19, 20], we defined MCAP and proposed three types of methods for it. We showed that one can solve MCAP to optimality by dynamic programming. Although the dynamic programming algorithm guarantees optimal solutions, it becomes intractable for large scale problems. We thus proposed heuristic methods that not only have practical time complexity but also showed good performance. There was also a recent study using particle swarm optimization to solve the multicampaign assignment problem [9]. However, since they are heuristics, they do not guarantee optimal solutions. Furthermore, although many approaches for MCAP have been proposed so far, nobody has yet shown whether or not the problem is computationally intractable.

In this paper, we suggest an advanced version of the multicampaign assignment problem that reflects a more realistic situation and to which all previous heuristics can be applied as they are, and then prove that it is NP-complete, using a novel technique. The remainder of this paper is organized as follows. In Section 2, we briefly review the multicampaign assignment problem and its optimal solver. We design a reasonable response suppression function and then present a practical version of the multicampaign assignment problem in Section 3. We provide the NP-completeness of multicampaign assignment in Section 4. Finally, we give our concluding remarks in Section 5.
2 Preliminaries

2.1 Multicampaign Assignment

The basic idea of the multicampaign assignment problem is from [19]. Let \( n \) be the number of customers and \( k \) be the number of campaigns. In the following, we describe input, output, constraints, and objective for MCAP.

**Input:** Each campaign has a weight. Each customer has a preference for each campaign and there is a response suppression function that gives the rate at which a customer’s preference for all campaigns decreases as the number of campaigns assigned to that customer increases.

- \( \mathbf{w} = (w_1, w_2, \ldots, w_k) \in \mathbb{Z}_+^k \): the campaign weight vector (\( w_j > 0 \) is the weight of campaign \( j \)).
- \( \mathbf{p} = (p_{ij}) \): the preference matrix. Each element \( p_{ij} \in \mathbb{Z}_+ \cup \{0\} \) is the preference value of customer \( i \) for campaign \( j \).
- \( r : \mathbb{Z}_+ \cup \{0\} \rightarrow [0,1] \): the response suppression function with respect to the number of recommendations (for convenience, we assume \( r(0) = 0 \)). The personalized extension of the response suppression function was considered in [6, 20]. We will deal with this issue again in Section 3.1.

If \( h_i \) is the number of recommendations for customer \( i \), the actual preference of customer \( i \) for campaign \( j \) becomes \( r(h_i)p_{ij} \).

**Constraints:** Let \( b^i \) be the upper bound of recommendations for campaign \( j \), and \( b_j \) be the lower bound of recommendations for campaign \( j \). Then the actual number of recommendations in campaign \( j \) is constrained to lie between \( b_j \) and \( b^j \).

- \( \mathbf{b}^* = (b_1^*, b_2^*, \ldots, b_k^*) \in \mathbb{Z}_+^k \): the upper bound capacity constraint vector
- \( \mathbf{b}_* = (b_1, b_2, \ldots, b_k) \in \mathbb{Z}_+^k \): the lower bound capacity constraint vector

These values are predefined largely depending on the budget of each campaign. The required cost of campaign \( j \) will be between \( b_j \cdot c_j \) and \( b^j \cdot c_j \), where \( c_j \) is the unit recommendation cost for campaign \( j \).

**Output:** The output is an \( n \times k \) binary campaign assignment matrix \( \mathbf{M} = (m_{ij}) \) in which \( m_{ij} \) indicates whether or not campaign \( j \) is assigned to customer \( i \).

**Objective:** The campaign preference for campaign \( j \) is defined to be the actual preference sum of recommended customers for campaign \( j \) as follows: \( \sum_{i=1}^n r(h_i)p_{ij}m_{ij} \), where \( h_i = \sum_{j=1}^k m_{ij} \).
The fitness $F(M)$ of a campaign assignment matrix $M = (m_{ij})$ is the weighted sum of campaign preferences.

$$F(M) = \sum_{j=1}^{k} \left( w_j \sum_{i=1}^{n} r(h_i) p_{ij} m_{ij} \right).$$

Note that computing $F(M)$ takes $O(nk)$ time.

The objective is to find a binary matrix $M = (m_{ij})$ that maximizes the fitness $F(M)$.

### 2.2 Getting Preferences

The preferences for a campaign can be obtained using existing methods such as collaborative filtering [24]. Collaborative filtering (CF) is widely used to predict customer preferences for campaigns [2]. In particular, CF is fast and simple [10], so it is popular for personalization in e-commerce [12, 21]. There have been a number of customer-preference estimation methods based on CF [1, 11, 13, 14, 23, 25, 26]. Nearest neighbor CF algorithms, e.g., [26], are based on computing the distance between customers based on their preference history. Predictions of how much a customer will like a campaign are computed by taking the weighted average of the opinions of a set of nearest neighbors for that campaign. Neighbors who have expressed no opinion on the campaign are ignored. This type of CF algorithm has been successfully used in previous multicampaign studies [19, 20].

### 2.3 Optimal Algorithm

We can find the optimal campaign assignment matrix of MCAP using dynamic programming [19]. The algorithm requires $O(nk \Pi_{j=1}^{k} b^j)$ space and takes $O(nk^2 \Pi_{j=1}^{k} b^j)$ time. The dynamic programming approach tries to solve all subproblems of the form “what is the best assignment for the capacity $(b'_1, b'_2, \ldots, b'_k)$, using just the first $m$ customers,” for all $0 \leq b'_1 \leq n, \ldots, 0 \leq b'_k \leq n$, and $0 \leq m \leq n$. If $k$ is a fixed number, this is a polynomial-time algorithm. However, when $k$ is not small and almost all $b^j$'s are $\Omega(n)$, it is nearly intractable. A practical version of MCAP, which is described in the next section, can also be solved by this dynamic programming algorithm with the same complexity.
3 Practical Version of Multicampaign Assignment

3.1 Realistic Response Suppression Function

In the case of multiple campaign recommendations, the customer response rate changes with recommendation patterns. We introduced the response suppression functions for the response rate with multiple recommendations [19]. But, these functions are very simple since not only are they artificially generated but the same function is also applied to all customers. Recently, this problem was weightily mentioned in [20] and furthermore a generalized Gaussian response suppression function that differs among customer classes was introduced in [6]. We design a response suppression function that is more suitable for the practical campaign situation, as mentioned in prior work [6, 20].

We also assume that each customer can have its own response suppression function since the degree of response suppression may vary from customer to customer [3, 23].

\[ r_i : \mathbb{Z}_+ \cup \{0\} \to [0, 1] \]: the response suppression function of customer \( i \) with respect to the number of recommendations (for convenience, we assume \( r_i(0) = 0 \) and \( r = (r_1, r_2, \ldots, r_n) \)). We now describe how such a response suppression function can be learned from historical data.

Assuming that the response suppression rate of a customer varies as his/her inclination, we classify the given (training) customers into a number of categories according to their profiles and preference history. This can be done by using existing data mining techniques [2, 3, 15]. We are to give the same response suppression function to all the customers belonging to the same category. Each customer has their preference for each campaign estimated by, e.g., CF. We assume that the prediction error is not large.

We perform each campaign independently to obtain multicampaign data. If two customers belong to the same category but showed different responses, we can consider that it is because of response suppression. We estimate the response suppression by using function approximation to find a function that satisfies as many of the following conditions as possible.

**Condition**: For each customer pair \((i, j)\) in \( C \) and each campaign \( k \), if customer \( i \) responded to campaign \( k \) and customer \( j \) did not,

\[ p_{ik} \cdot r_C(h_i) > p_{jk} \cdot r_C(h_j), \]
where $p_{lk}$ is the preference of customer $l$ for campaign $k$, $r_C$ is the response suppression function of category $C$, and $h_i$ is the number of recommendations for customer $l$.

Now given new customers, we classify them into the above categories according to their profiles and preference history. As the response suppression function for each new customer, we use the function of his/her corresponding category. The original multicampaign model [19] is very simple, in the sense that the number of categories is just one.

### 3.2 Formal Definition

When we use the realistic response suppression function of Section 3.1, the fitness $F(M)$ is redefined in the following formula.

$$F(M) = \sum_{j=1}^{k} \sum_{i=1}^{n} w_j r_i(h_i) p_{ij} m_{ij},$$

where $h_i = \sum_{j=1}^{k} m_{ij}$.

Now we formally define a decision version of the multicampaign assignment problem with the realistic response suppression function $r$. For convenience, we assume that $1_n$ is the $1 \times n$ matrix $(1 \ 1 \ \cdots \ 1)$.

**Definition 1** MCAP = \{⟨$w, P, r, b^*, b_*, t$⟩ | ∃ $M = (m_{ij})$ such that $b_* \leq 1_nM \leq b^*$ and $F(M) \geq t$\}

### 3.3 Heuristic Algorithms

There have been a number of heuristic algorithms for solving MCAP [6, 19, 20]. As stand-alone heuristics, the constructive assignment algorithm with $O(nk^2 \log n)$-time complexity and the iterative improvement heuristic with $O(nk \log n)$-time complexity were proposed in [19]. Recently, a novel heuristic using Lagrange multipliers with $O(nk^2)$-time complexity was devised [20]. When combined with real-coded genetic algorithms [27], this algorithm could perform near-optimally. There was also a recent heuristic algorithm of [6] based on particle swarm optimization [16]. In summary, many kinds of heuristics for MCAP have been proposed and they showed good performance. Most importantly, the practical version of the problem can also be solved by all these heuristic algorithms with the same complexity. However, up to now
nobody has shown whether or not MCAP is NP-complete. In the next section, we focus on the computational complexity of MCAP.

4 Computational Complexity of Multicampaign Assignment

Suppose there are several assignment or allocation problems, some problems are in P (e.g., the optimal assignment problem [22]). But, many of them are NP-complete [9] (e.g., the quadratic assignment problem [4], the generalized assignment problem [5], and the multiprocessor document allocation problem [8]). In this section, we prove that the practical version of MCAP is NP-complete. In the following extreme cases, MCAP is tractable, i.e., belongs to P.

- All the response suppression function are constant functions. Then the problem can be solved optimally in $O(nk \log n)$ time. If we only sort actual preferences of all the customers for each campaign $j$, i.e., the set $\{p_{1j}r_1, p_{2j}r_2, \ldots, p_{nj}r_n\}$ for each $j$, the optimal solution can be easily obtained.

- There is no bound constraint, i.e., $b_* = 0_k$ and $b^* = n1_k$. Then there exists an $O(nk \log k)$-time algorithm to solve the problem to optimality. In solving the problem, the most time-consuming task is to sort preferences of each customer $i$ for all the campaigns, i.e., the set $\{p_{i1}, p_{i2}, \ldots, p_{ik}\}$ for each $i$.

However, in the general case, the problem becomes intractable by the following theorem.

**Theorem 1** MCAP is NP-complete. Moreover, even when the campaign weight vector is ignored, i.e., $w_i = 1$, the problem is still NP-complete.

**Proof:** To show that MCAP is in NP, we let the campaign assignment matrix $M = (m_{ij})$ be the certificate. For any instance of the problem, we can check whether $b_* \leq 1_nM \leq b^*$ and $F(M) \geq t$ in polynomial time.

We show that 3-SAT $\leq_P$ MCAP. Given a 3-CNF (conjunctive normal form) formula $\phi$ over variables $x_1, x_2, \ldots, x_l$ with clauses $C_1, C_2, \ldots, C_m$, each containing exactly three distinct literals, the reduction algorithm constructs an instance $\langle w, P, r, b^*, b_*, t \rangle$ of MCAP such that the formula $\phi$ is satisfiable if and only if there is a campaign assignment matrix $M = (m_{ij})$
satisfying $b_* \leq 1_n M \leq b^*$ and $F(M) \geq t$. Without loss of generality, we make two assumptions about the formula $\phi$.

i) No clause contains both a variable and its negation. Such a clause is always satisfied by any assignment of values to the variables.

ii) Each variable appears in at least one clause. Otherwise it does not matter what value is assigned to the variable.

The reduction creates two customers, $u_i$ and $u_i'$, for each variable $x_i$ and three customers, $s_j$, $s'_j$, and $s''_j$, for each clause $C_j$, i.e., $n = 2l + 3m$. We also create $k = l + m$ campaigns and each campaign corresponds to either one variable or one clause. We label each campaign by either a clause or a variable; that is, the campaigns are labelled $C_1, C_2, \ldots, C_m, x_1, x_2, \ldots, x_l$.

We construct $w, P, r, b^*, b_*$, and $t$ as follows:

- The campaign weight vector $w$ is set to $1_n$.
- The preference matrix $P = (p_{ij})$ is constructed as follows. For each $i = 1, 2, \ldots, l$, set $p_{u_i x_i} = p_{u'_i x_i} := 10^m i - 1$. If literal $x_i$ appears in clause $C_j$, then set $p_{u_i C_j}$ to $10^{j - 1}$. If literal $\overline{x}_i$ appears in clause $C_j$, then set $p_{u'_i C_j}$ to $10^{j - 1}$. For each other variable or clause $y$, set $p_{u_i y} = p_{u'_i y} := 0$. For each $j = 1, 2, \ldots, m$, set $p_{s_j C_j} = p_{s'_j C_j} = p_{s''_j C_j} := 10^{j - 1}$. For each other variable or clause $y$, set $p_{s_j y} = p_{s'_j y} = p_{s''_j y} := 0$.
- The response suppression function $r$ is constructed as follows. For each $s_j$, if $y = 1$, set $r_{s_j}(y)$ to 1. Otherwise set $r_{s_j}(y)$ to 0. Set $r_{s'_j} = r_{s''_j} := r_{s_j}$. For each $u_i$, let $\alpha_i$ be $1 + \sum_{j: P_{u_i C_j} > 0} 1$. If $y = \alpha_i$, set $r_{u_i}(y)$ to 1. Otherwise set $r_{u_i}(y)$ to 0. For each $u_i'$, let $\alpha'_i$ be $1 + \sum_{j: P_{u'_i C_j} > 0} 1$. If $y = \alpha'_i$, set $r_{u'_i}(y)$ to 1. Otherwise set $r_{u'_i}(y)$ to 0.
- The upper bound capacity constraint vector $b^*$ is as follows. Set $b^{C_1} = b^{C_2} = \cdots = b^{C_m} := 4$ and $b^{x_1} = b^{x_2} = \cdots = b^{x_l} := 1$, i.e., $b^* = (4 \ 4 \ \cdots \ 4 \ 1 \ \cdots \ 1)$.
- The lower bound capacity constraint vector $b_*$ is equal to $b^*$.
- The value $t$ is set to the number in base 10 represented by the digit vector $b^*$, where the clause campaigns $C_1, C_2, \ldots, C_m$ stand for the $m$ least significant digits, and the variable campaigns $x_1, x_2, \ldots, x_l$ stand for the $l$ most significant digits, i.e., $t = \sum_{j=1}^k b^j 10^{j - 1} = 4 \sum_{t=1}^m 10^{j - 1} + \sum_{j=1}^l 10^m + j - 1$. 


Figure 1: An example of the reduction 3-SAT \( \leq_p \) MCAP

The reduction can be performed in polynomial time. Figure \( \Pi \) shows the reduction of an example formula with 3 variables and 4 clauses. This will help to understand the reduction algorithm.

Now we show that the 3-CNF formula \( \phi \) is satisfiable if and only if there is a campaign assignment matrix \( M = (m_{ij}) \) satisfying \( 1_n M = b^*(= b_\ast) \) and \( F(M) \geq t \). First, suppose that \( \phi \) has a satisfying assignment. For \( i = 1, 2, \ldots, l \), if \( x_i = 1 \) in this assignment, then set \( m_{u_i x_i} \) to 1. Otherwise, set \( m_{u'_i x_i} \) to 1. Having either \( m_{u_i x_i} = 1 \) or \( m_{u'_i x_i} = 1 \), but not both, for all \( i \), and setting \( m_{s_j x_i} = m_{s'_j x_i} = m_{s''_j x_i} \) to 0 for all \( i \) and \( j \), we see that for each variable-labeled campaign, the sum of recommendations must be 1, which matches the values of \( b^* \). Because each clause is satisfied, there is some literal in the clause with the value one. Then, each campaign labeled by
a clause $C_j$ has at least one $u_i$ or $u'_i$ customer that has positive preference for campaign $C_j$ and is recommended by campaign $x_i$. In fact, 1, 2, or 3 literals may be 1 in each clause, and so, among the $u_i$ and $u'_i$ customers, each clause-labeled campaign $C_j$ has 1, 2, or 3 customers that have positive preference for campaign $C_j$ and are recommended by campaign $x_i$. For each $u$ among such customers, set $m_{uC_j}$ to 1. We achieve four recommendations in each campaign labeled by clause $C_j$ by setting $m_{uC_j}$ to 1 only for each customer $s$ in the appropriate nonempty subset of customers $\{s_j, s'_j, s''_j\}$. We have matched $b^* (= b_*)$ in all campaigns. If $M$ is assigned as described above, then $r_i(h_i) = 1$ for each customer $i$ and $p_{ij} = 10^{i-1}$ whenever $m_{ij} = 1$. Hence, $F(M) = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij} r_i(h_i) p_{ij} m_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{k} p_{ij} m_{ij} = \sum_{i=1}^{k} 10^{i-1} m_{ij} = \sum_{j=1}^{k} 10^{j-1} b^j = t$. 

Now, suppose that there is a campaign assignment matrix $M = (m_{ij})$ such that $1_n M = b^* (= b_*)$ and $F(M) \geq t$. Since $r_i(h_i) \leq 1$ for each customer $i$ and $p_{ij} \leq 10^{j-1}$, $F(M) = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij} r_i(h_i) p_{ij} m_{ij} \leq \sum_{i=1}^{n} \sum_{j=1}^{k} p_{ij} m_{ij} \leq \sum_{j=1}^{k} 10^{j-1} b^j = t$. Hence, $F(M)$ is equal to $t$. The following important properties about the matrix $M$ hold.

**Property 1** For each customer $i$ and each campaign $j$, if $m_{ij} = 1$, then $p_{ij} = 10^{j-1}$.

($\because$ Suppose that $p_{ij} \neq 10^{j-1}$, i.e., $p_{ij} = 0$. Then, $F(M) = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij} r_i(h_i) p_{ij} m_{ij} \leq \sum_{i=1}^{n} \sum_{j=1}^{k} p_{ij} m_{ij} < \sum_{j=1}^{k} 10^{j-1} \sum_{i=1}^{n} m_{ij} = t$. Hence, $F(M) < t$. This is a contradiction.)

**Property 2** If $m_{ij} = 1$, then $r_i(h_i) = 1$.

($\because$ Suppose that $r_i(h_i) \neq 1$, i.e., $r_i(h_i) = 0$. Then, $F(M) = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij} r_i(h_i) p_{ij} m_{ij} < \sum_{i=1}^{n} \sum_{j=1}^{k} p_{ij} m_{ij} \leq t$. Hence, $F(M) < t$. This is a contradiction.)

**Property 3** If $m_{ij} = 1$, then $m_{iq} = 1$ for every campaign $q$ satisfying $p_{iq} > 0$.

($\because$ Since $r_i(h_i) = 1$ by Property 2, $h_i (= \sum_{j=1}^{k} m_{ij}) = \alpha_i$, i.e., $h_i$ is the same as the number of $j$’s such that $p_{ij} > 0$. Also, if $p_{ij} = 0$, $m_{ij} = 0$ by Property 1. So, if there exists campaign $q$ such that $p_{iq} > 0$ and $m_{iq} = 0$, $h_i$ cannot be $\alpha_i$.)

Consequently, $M$ chooses all elements that have positive preference in one row, or choose nothing.

For each $i = 1, 2, \ldots, l$, either $m_{u,x_i} = 1$ or $m_{u',x_i} = 1$, but not both, for otherwise the campaigns labeled by variables would not sum to 1. If $m_{u,x_i} = 1$, set $x_i$ to 1. Otherwise set $x_i$ to 0. We claim that every clause $C_j$, for $j = 1, 2, \ldots, m$, is satisfied by this statement. To prove this claim, note that to achieve four recommendations in the campaign labeled by $C_j$,
for at least one \( u_i \) or \( u'_i \), \( m_{u_iC_j} = 1 \) or \( m_{u'_iC_j} = 1 \), since the contributions of customers \( s_j \), \( s'_j \), and \( s''_j \) together sum to at most 3. If \( m_{u_iC_j} = 1 \), then \( m_{u_ix_i} = 1 \) by Property 3 and \( p_{u_iC_j} > 0 \) by Property 1, which means that the literal \( x_i \) appears in \( C_j \). Since we have set \( x_i = 1 \) when \( m_{u_ix_i} = 1 \), clause \( C_j \) is satisfied. If \( m_{u'_iC_j} = 1 \), then \( m_{u'_ix_i} = 1 \) and \( p_{u'_iC_j} > 0 \), and so the literal \( \neg x_i \) appears in \( C_j \). Since we have set \( x_i = 0 \) when \( m_{u'_ix_i} = 1 \), clause \( C_j \) is satisfied. Thus, all clauses of formula \( \phi \) are satisfied.

Q.E.D.

5 Concluding Remarks

In this paper, we suggested a practical version of the multicampaign assignment problem (MCAP). Under the practical multicampaign model, we also presented how to determine a realistic response suppression function instead of existing artificial functions of [6, 19].

In the case that the number of campaigns is fixed, the optimal campaign assignment matrix can be found in polynomial time using the dynamic programming method. However, in general, the method is intractable. The MCAP could also be solved near-optimally in a reasonable time by novel heuristics proposed in our prior work [20]. However, its intractability has been left open. In this paper, in confirmation of the intractability, we provided an elaborate proof of that the problem is NP-complete. The NP-completeness of the MCAP justifies the use of various heuristics for the problem as in [6, 19, 20].

Acknowledgments

The authors would like to thank Prof. Vic Rayward-Smith for his valuable suggestions in improving this paper.

References


