JOURNAL OF INDUSTRIAL AND
MANAGEMENT OPTIMIZATION

doi:10.3934/jimo.2021202

AN APPLICATION OF APPROXIMATE DYNAMIC PROGRAMMING IN MULTI-PERIOD MULTI-PRODUCT ADVERTISING BUDGETING

Majid Khalilzadeh and Hossein Neghabi*

Department of Industrial Engineering, Ferdowsi University of Mashhad, Iran
Azadi Square, Mashhad, Iran

Ramin Ahadi

Faculty of Management, Economics and Social Sciences, University of Cologne
Cologne, Germany

(Communicated by Ada Che)

Abstract. Advertising has always been considered a key part of marketing strategy and played a prominent role in the success or failure of products. This paper investigates a multi-product and multi-period advertising budget allocation, determining the amount of advertising budget for each product through the time horizon. Imperative factors including life cycle stage, BCG matrix class, competitors' reactions, and budget constraints affect the joint chain of decisions for all products to maximize the total profits. To do so, we define a stochastic sequential resource allocation problem and use an approximate dynamic programming (ADP) algorithm to alleviate the huge size of the problem and multi-dimensional uncertainties of the environment. These uncertainties are the reactions of competitors based on the current status of the market and our decisions, as well as the stochastic effectiveness (rewards) of the taken action. We apply an approximate value iteration (AVI) algorithm on a numerical example and compare the results with four different policies to highlight our managerial contributions. In the end, the validity of our proposed approach is assessed against a genetic algorithm. To do so, we simplify the environment by fixing the competitor's reaction and considering a deterministic environment.

1. Introduction.

A large number of organizations and companies have been expanded greatly over time. A large company could have several business divisions and a number of product lines within each division. At the same time, there are a host of other companies producing the same products, which creates a competitive environment where they struggle to win the largest share of the market. This issue is crucial that even the retails shelf-space allocation is taken into account [15]. Accordingly, companies must find optimal strategies considering all aspects of the business. Among them, marketing strategies play a specific role in the success or failure of the companies due to the wide range of customers' options in such a competitive environment. Advertising is a fundamental component of marketing

2020 *Mathematics Subject Classification.* Primary: 90C39, 90C40.

Key words and phrases. Dynamic programming, Marketing, Advertising budgeting, Multi-period multi-product.

The second author is supported by Iran's National Elites Foundation.

*Corresponding author: Hossein Neghabi.

strategy [4]. It goals to identify and distinguish products, giving information, encouraging customers, and building value, preference, and loyalty. The media, as a well-known type of advertising, encompass any method used to convey a message to the public such as newspapers, magazines, radio, television, the Internet, and billboard [36].

Due to the importance of advertising, companies allocate a significant share of their promotional budget to advertising [9]. To efficiently use the budget, appropriate media planning is essential to determine when and through which media the budget should be spent. Planning depends on a number of factors including advertising goals, customers' profile of different market segments, market and competitive state, competitors' profile, brand positioning, product portfolio analysis, and the life cycle stage [18], that must be taken into account to avoid sub-optimal planning decisions. For instance, in the introductory stage, a significant portion of the promotional budget is often allocated to induce trial. In the growth stage, planners seek to capture maximum market potential, and moving to the maturity stage, advertising acts as a reminder to keep customers aware of the product. Finally, companies usually withdraw their support for advertisement in the decline stage although it depends on specific circumstances.

Almost all marketers make their decisions based on prior-period advertising responses while circumstances may change over time and should be considered in media planning due to their significant effects. Therefore, companies need to devise appropriate plans based on the current market scenarios. Indeed, selecting an optimal media mix depends on the current state of the system (e.g., on the stage of the product life cycle in the market), which is affected by the previous system state and action. Accordingly, this is a dynamic media planning where a chain of optimal actions needs to be taken. This dynamic nature is captured by dividing the planning horizon into multiple time periods. Important factors such as customers' behavior and the cost and popularity of different media must be analyzed at the beginning of each time period. Then, marketers can start planning based on this information to maximize the total reach of customers. In addition, successful media advertising should consider to which kind of media their customers are exposed. In other words, the effect of different media varies for each product since they have specific types of customers, meaning that advertising in media is effective only if it covers a significant number of the targeted customers. Hence, media planners need to manage their advertising budget for different products simultaneously. This ends in a multi-period multi-product advertising where at each period a set of actions for all products needs to be taken while taking its long-term effect on the following time steps into account.

Considering competitors, this problem is a complicated game where companies seek to gain higher market share. Competitors make their reactions with a probability conditioned to our decisions. So, while making decisions, we should consider the competitor's reactions and the impact on our objective function. For the sake of simplicity, we do not contemplate each individual competitor, but their mean reaction is considered. This is not the only uncertain dimension of the problem, the behavior of customers and efficiency of the decisions in the market are also stochastic and should be considered in the decision-making process. The last and probably the most serious challenge is the curse of dimensionality in such a problem, caused by the substantial number of products, actions, and reactions of rivals.

In this regard, various methods are able to address these challenges such as Monte Carlo simulation, dynamic programming, stochastic dynamic programming, mathematical programming, and fuzzy and robust programming. However, none of these methods considers all three challenges regarding media planning at the same time. As an example, using exact solutions of dynamic programming is not applicable due to the extremely high computational costs, the number of scenarios in stochastic mathematical programming and Monte Carlo simulation easily gets intractable, and robust/fuzzy programming cannot cover all complexities of the problem like probabilities of competitors reactions and the associate impact on our decisions. To the best of our knowledge, the most suitable method is approximate solutions for the formulated Markov Decision Process (MDP). Since the transition probabilities (i.e., dynamics of the environment) are known in this case, we employ a model-based approach called approximate dynamic programming, which basically is a planning approach considering the sequential nature of the problem and solves the curse of dimensionality in the light of approximation. Hence, in this paper, we develop an ADP algorithm to address the advertising management problem for companies running their business in a competitive and uncertain environment. The goal of the planning is to maximize profits. Two general algorithmic strategies including approximate value iteration (AVI) and approximate policy iteration (API) exist for obtaining approximate solutions to our computational stochastic optimization problem. First, the AVI algorithm is used to gain answers based on the decisions maximizing the objective function. Next, the API algorithm is proposed to obtain policies that map the system states to the actions providing insights to senior managers.

We make a couple of contributions to this paper. First, to the best of our knowledge, this is the first study addressing the multi-period multi-product advertising budgeting problem considering complex assumptions of a real market. In addition to the stochastic effects of the actions, we consider a competitive environment where other companies' reactions affect the results. Methodologically, we apply an ADP algorithm on the advertising management problem addressing the multi-period planning in an uncertain and competitive environment. We evaluate the proposed approach in a real-size numerical application and benchmark it against a heuristic approach (genetic algorithm). There are also some managerial implementations in this paper. By employing the proposed approach, managers can make more profit compared to using pre-defined policies. They also need to note that each product needs different share of the advertising budget and a different media to maximize the joint sales of all products.

The rest of the paper is organized as follows. Section 2 reviews the literature on media planning and advertising management as well as some applications of ADP on similar works. The problem definition is presented in section 3 discussing the main assumptions and details of the problem. Then, in section 4, we apply and adopt an ADP algorithm to solve the problem. Section 5 discusses the results and benchmark them with similar methods, and finally, we conclude our paper in section 6.

2. **Literature review.** The literature of our work is divided into two streams. In the first one, papers related to media planning have been discussed, and the second one involves the number of research that used ADP as their solution approach.

2.1. **Media planning.** There has been various research about media planning in which advertising decisions are indicated according to different factors including single/multi-product(s), single/multi-media, mass or segmented specific media, and time planning horizon. In early studies, simple models are considered to discuss interactions between advertising and sales in the present and future ([30, 41]). Introductory papers are mentioned in Danaher and Rust [11] where authors proposed simple models that determine optimal expenditure spending in media for the case of a single product and a single period.

The problem of media planning has been extended in many terms, for example, Doyle and Saunders [12] analyzed a multi-product advertising budgeting problem for a single-stage while the model considered in Sriram and Kalwani [37] solves the problem for a single product over multi periods. In the following, authors in [9] in addition to determining an optimal advertising budget proposed a media allocation model including both multiperiod and multiproduct simultaneously. A multi-period dynamic advertising program has been proposed in [16] which demand at each period depends on current and past advertising efforts.

In related literature, customer lifetime value has been considered as an important factor in allocating the promotion budgets using various approaches ([26, 19]). The experience Fischer et al. [14] of marketing budget allocation showing that profit increase from improving allocation across products or regions is more effective than expanding the total budget. They have studied the marketing budget allocation problem for multi-product, multi-country firms. In the same direction, Yang and Xiong [43] discussed the optimization problem of the advertising budget allocation for revenue management subject to an inventory constraint. Their method was a nonparametric learning-while-doing, which means performing a series of advertising experiments to predict the market sales reaction by observing realized sales (learn and exploration) then allocating budget planning based on the learned sales function (exploitation). This policy balances the advertising budget and the inventory budget simultaneously. With some improvement, allocating advertising resources under constraints for multiple markets with diverse goals and objectives was the subject of Wang et al. [42]. They have presented a model based on the Vidale–Wolfe [41] advertising response model, which considers the relationship between market conditions and specific objectives (strategic concerns), saturation level, market responses, and advertising spending. Yang et al. [44] extended that work by employing a generalized Vidale–Wolfe model as an advertising dynamic for budget allocation decisions under a finite time horizon, which is very similar to our work regarding consideration of a dynamic market.

Media planning has been addressed in literature by different approaches including linear and non-linear programming, goal programming, game theory, and dynamic programming. A linear programming model is presented in [6] allocating budget to different media considering various constraints in order to maximize segmented exposure audience. Another one is integer-linear programming with the objective function of maximizing people exposed to their advertisements; this work is single-media planning that addressed the question of which channels, what time, and how often they need to promote [28]. Abedi [1] applied a nonlinear and nonseparable knapsack problem to allocate a limited marketing budget between multiple channels at multiple markets.

In [21], mixed-integer goal programming is developed to determine the types of media and advertising frequency in the dual consumer/industrial market. Branch and bound integer programming and goal programming techniques have been used to maximize the advertising revenue in news websites [25]. Moreover, Jea et al. [17] proposed an advertisement scheduling plan and used branch and bound techniques to maximize the profit in online advertising agencies. From a game theory perspective, there are continuous interactions among actors through analysis, evaluate and learn. The learning cycle is drastically quickened and developed with machine learning, online learning, and different implementations of artificial intelligence [3]. In the face of constant environmental changes, any actor can use methods such as reinforcement learning to continuously adapt their behaviors [45].

In the case of competitive markets, game theory can be applied as a powerful approach to recognize interactions among marketing activities and rivals. Naik et al. [29] have investigated a planning problem in oligopoly dynamic markets using the Lanchester model; furthermore, they develop an estimation method to calibrate dynamic models of oligopoly using market data. An analysis of optimal advertising allocation has been carried out in [34] where market share for each firm depends on rivals' reactions as well as their decisions. Lu et al. [24] have dealt with differential games in the dynamic advertising models. They considered a supply chain consisting of a dominant retailer and a manufacturer, in which the two-channel members take the effect of dynamic advertising into account as a Stackelberg differential game (leader and follower).

Due to changeable marketing conditions over the course of time, the time horizon of media selection problems should be divided into multiperiod. This means that the problem is dynamic with sequential factors and decisions in real cases. For example, in [27], a dynamic programming model is proposed to determine optimal advertising decisions in order to maximize profit at the end of the time horizon. Li and Sun [22], used the dynamic pricing policy and dynamic programming technique to manage products demands.

The most challenging factor in our problem is the uncertainty that has been addressed in the literature from different perspectives. Stochastic game theory has been developed by Prasad and Sethi [34]. In Bass et al.[5] the effect of different themes in the advertising budgeting problem has been discussed using a Bayesian linear model. Regarding uncertain conditions, effects of advertising would not be wholly determined and accordingly, a new objective function of maximizing expected market utility has been introduced in [13] as a special Markov decision process.In [8], a stochastic advertising budgeting model for multiproduct and multiperiod has been developed and its results are compared with a deterministic model for a real case study. Besides, uncertainty has been addressed with other approaches such as fuzzy optimization [31] and robust optimization [2].

2.2. **ADP.** In media planning, decisions determine which percentage of budget ought to be assigned to which product over time. Decisions in different periods affect each other such that a chain of decisions should be made instead of choosing separately. Moreover, at the beginning of each period, new information such as the price of products, market share, and competitors' performance will be received, which can affect our actions. These features lead to a sequential problem with a great deal of uncertainty. As such, we formulate it as a Markov decision process (MDP), but unfortunately, due to the high dimensionality of such a problem, exact dynamic programming cannot be utilized to give us the optimal solution. Instead, an approximate dynamic programming (ADP) algorithm is used, which is suitable for huge dimensions and a high amount of uncertainty. ADP is a methodology estimating the value of future states throughout an iterative process. It first had been studied in Bellman and Dreyfus [7] using statistics to estimate the value of further states, then developed in the context of reinforcement learning by Sutton and Barto in [38]. In the following, various works ([32, 39]) make a significant contribution to the development of the field and nowadays, we have books of Puterman [35], Powell [33], and Bertsekas [10] as references for ADP approach.

3. **Problem description.** This paper addresses the problem of selecting advertising packages of different media for a multi-product company in a multi-period time horizon. Various packages are offered by an advertising agency including different types such as television, magazines, billboards, and the Internet. The cost and effectiveness of each package vary for each product, and significantly affect budget allocation to products and packages. The purpose of the model is to find the best packages of advertisement, within the allowable budget assigned for different media, in order to maximize the profits or sales.

In a multi-product case (i.e., companies manufacturing more than one product), it is essential to promote all the products simultaneously with a joint advertising strategy. Assumed that these products have a shared promoting budget while having an independent market and no cross effect (i.e., are entirely distinct from each other). For example, a dishwasher and an air conditioner are usually manufactured in the same company and are the typical needs of people. Accordingly, if a person purchases an air conditioner, it will have no effect on purchasing a dishwasher. The company should promote such distinct products in separate advertisements while jointly allocating the budget to have the most significant impact on total sales. In order to devise a detailed media plan, companies should be completely familiar with the market nature. In the current market, people have different attitudes toward their needs and have various options, presenting a number of challenges to the media planners. To provide a better service to different customers, companies usually segment the market into their target customers. Then, they place advertisements in a way that they have the highest impact on their target customers [40]. Accordingly, media can be divided into mass media, targeting all the segments in a similar way, and segment-specific media with different plans for different segments. The present study considered mass media for advertising, indicating that the media have general audiences.

Another important aspect of advertising that should be taken into account is that advertising influences differently in each product life cycle stage. In the initial phase, advertising raises awareness about products among the public whereas it tries to retain customers' loyalty in later phases. In this study, the product life cycle is considered a classic cycle and is made up of four primary stages: introduction, growth, maturity, and decline (i.e., each stage is determined using the number of sales). The effectiveness of each media varies among different stages (e.g., television might influence more in the introduction stage rather than maturity stage). This challenge could be addressed by dividing the whole planning horizon into multiple time periods to distinguish different life cycle stages and consider their effects on the planning. Then the decisions could adopt in each period, depending on all the factors which might change over the course of time, such as product life cycle and product price. We also consider the Boston Consulting Group's product portfolio matrix (BCG matrix), which is designed to help with long-term strategic planning, assist a business in considering growth opportunities by reviewing its portfolio of products, and decide where to invest, discontinue, or develop products. The matrix is divided into four quadrants based on an analysis of the market growth and relative market share, where for each quadrant advertising packages can have different effects.

Furthermore, since competitors' actions can strongly affect the results of our decisions, we require to analyze them to approach more to the global optimal answers in a real media planning problem. After identifying and analyzing competitors, we can design our own advertising program to draw customers' attention to our product in the presence of many other options. Noted that competitors are aware of decisions and might modify their strategies based on the state of the market. Thus, to improve our strategy, we should include the probability of competitors' reactions conditioned to our decisions.

The most challenging part of this problem is that all influencing factors of advertising management need to be modeled under uncertain circumstances due to the lack of information in the following periods. For instance, it does not guarantee that an advertising package certainly affects the sales as we expected. There are so many other factors that might be neglected such as an introduction of a new product by another company. Moreover, competitors' reaction is also uncertain as we do not fully aware about their strategy in the future. However, the best we can do is that we must consider these uncertainties in the planning part. Accordingly, we address a decision-making problem optimally allocating the advertising budget to different products by selecting an advertisement package for each within a multi-period time horizon. Our approach must deal with the stochastic and dynamic nature of the problem and makes decisions regarding uncertainties that might occur and adjust decisions in further periods. In the next section, we apply and adopt a suitable method for our problem covering the existing challenges.

4. **Methodology.** This section describes our approach addressing a decision-making problem of advertising planning using approximate dynamic programming (ADP). The proposed method has been emerging as a powerful discrete-time technique for solving multistage stochastic control processes that would otherwise be computationally intractable. The approach involves "cost-to-go" values calculation for all costs states that can be expected under the optimal policy [33]. In the following, we define necessary terms to allows us to map our problem to the scales of dynamic programming, and develop an approximate solution dealing with challenges raised by huge dimensions and uncertainty.

Suppose that the planning horizon is finite and divided into different periods and let $\mathcal{T} = \{1, 2, ..., T\}$, $(T < \infty)$ be the set of decision periods. The number of periods (T) is fixed and decisions are taken at the beginning of each. Let $S_{t,t\in\mathcal{T}}$ represent the system state at period t (i.e., the initial state (S_0) has been specified beforehand). To consider the longer-term impact and temporal dependency of our decision, we formulate our problem as a Markov decision process (MDP), which models a sequential decision process aiming to optimize a long-term objective function. Here, each state depends on the previous pair of state-action and random conditions that occurred during the time. In each period, a decision must be made according to the value index, the sum of the probability of going to the next, and the constraints. In addition to states, an MDP contains also a set of actions and a reward function that is represented in the following.

4.1. **States, actions, and rewards.** The state of the system is represented by the life cycle stage, class in the BCG matrix, competitors' strategy, and the available budget. Noted, it is assumed that there is a finite number of products defined as $\mathcal{M} = \{1, 2, \ldots, M\}$. Thus, three first terms of the system state are specified for each product while the budget is shared between them (i.e., each state is a vector with $3M + 1$ elements.).

The product life cycle state is defined as $LC_t = (lc_{t1},\, lc_{t2}, ..., lc_{tM})$, where, $lc_{tm} \in \{Introduction,\, Growth, Maturity, Decline\}$ is the life cycle stage of product m, at time t. This state depends on the sales volume of the product at the beginning of period t, which is denoted by $Y_t\,(m)$.

The product class in the BCG matrix is defined as $BCG_t = \,(bcg_{t1}, bcg_{t2},\, \ldots, bcg_{tM})$ where $bcg_{tm} \in \{\ Question\ marks,\ Stars,\ Cash\ cows,\ Dogs\ \}$ is the status of the product m in BCG matrix, at time t, which depends on industry attractiveness (growth rate of that industry) and competitive position (relative market share).

In addition, the competitors' strategy is defined as, $CS_t = (cs_{t1},\, cs_{t2}, ..., cs_{tm})$, where, $cs_{tm} \in \{$ High defensive marketing, High offensive marketing, Low defensive marketing, Low offensive marketing $\}$is the strategy that competitors follow over a period t. Let $p\,(cs_{tm}) = \hat{P}\,(cs_{tm}|X_{t-1,m})$ denotes a conditional probability distribution of the competitors' strategy. Our action conditions this distribution in a previous period (ie.$X_{t-1,m}$) that will determine the likelihood of competitors' reaction to the market. Moreover, each reaction has an uncertain effect on the market, which is represented as a bounded interval.

The remaining budget status is assumed to be discrete and shown by B_t that is an integer number between zero and total budget. Regarding these components, the state of the system at decision period t is defined as, $S_t = (LC_t, BCG_t, CS_t, B_t)$, $S_t \in S$, where S is the set of all possible states. Also associated state with product m in period t is defined as, $S_{tm} = (lc_{tm}, bcg_{tm}, cs_{tm})$.

Regarding the set of actions, at each period t, we must decide to choose an advertising package among allowable budget choices. The selected action vector for all products is defined as $X_t = \,(x_{t1}, x_{t2}, x_{t3},\, \ldots, x_{tM})$, where, $x_{tm} \in D\ \forall m \in \mathcal{M},\, t \in \mathcal{T}$, is our action for the product m and D is set of all advertising packages plus inaction; $D = \{$Inaction, Advertising package 1, Advertising package 2, $\ldots$, Advertising package $k\}$.

The reward of any action is equivalent to the related effects on the product's sales which is determined by the corresponding state and action. This effectiveness is a random range varying in each component of vector S_{tm} indicated by s_{tm}.

$$R\,(s_{tm}, x_{tm}) = rand\,(r_{1,s_{tm},x_{tm}}, r_{2,s_{tm},x_{tm}}) \qquad \forall t \in \mathcal{T} \tag{1}$$

where, $r_{1,s_{tm},x_{tm}}$ and $r_{2,s_{tm},x_{tm}}$ respectively represent the lower bound and upper bound of change in sales volume of product m. This amount is dependent on the state of product m, s_{tm}, and the action, x_{tm}. Indeed, this randomness indicates uncertainty in the problem and is determined by experts and decision makers. Therefore, assuming the independence among aforesaid uncertainties, the resultant effect on product sales is equal to $\prod\limits_{s_{tm}\,\in\, S_{tm}} R\,(s_{tm}, x_{tm})$. To evaluate the results of our possible decisions at each period we need to estimate sales volume at the end of this period according to our action. The estimated sales volume for product m

at the end of period t (i.e., $\hat{A}\left(S_{tm}, x_{tm}\right)$), is obtained by:

$$\hat{A}\left(\mathrm{S_{tm}}, \mathrm{x_{tm}}\right) = \mathrm{Y_t}\left(\mathrm{m}\right) \prod_{s_{tm} \in S_{tm}} \mathrm{R}\left(\mathrm{s_{tm}}, \mathrm{x_{tm}}\right) \ \forall m \in \mathcal{M}, \qquad \forall t \in \mathcal{T} \tag{2}$$

The reward of each advertising package depends on the product price and the corresponding sales volume. The price of products over a specific period is not certain and can be obtained through statistical forecasting tools such as regression, time series technique, neural network, etc. We use a simple regression in this problem to estimate the price at each time since the details of the price prediction are beyond the scope of this paper. Moreover, $Y\left(S_t, X_t\right)$, the total value of estimated sales of action X_t, is obtained by the summation of multiplying the estimated sales volume of product m, and associated market price ($\xi_m\left(t\right)$) that is defined as:

$$Y\left(S_t,\ X_t\right) = \sum_{m=1}^{M} \xi_m\left(t\right) \hat{A}\left(S_{tm}, x_{tm}\right), \qquad \forall t \in \mathcal{T},\ x_{tm} \in X_t \tag{3}$$

where, $\xi_m\left(t\right)$ is the price of product m, at decision period t.

In addition, assumed to have the budget constraint:

$$c\left(X_t\right) \leq B_t \tag{4}$$

where, $c\left(X_t\right)$ is the cost of action X_t that have made at decision period t. In the following, a new variable is defined to remove infeasible actions:

$$u\left(X_t\right) = \begin{cases} 1, & c\left(X_t\right) \leq B_t \\ 0, & otherwise \end{cases}, \tag{5}$$

where, $u\left(X_t\right)$ represents the possibility of action $\left(X_t\right)$ at current period t.

4.2. **Transition and value functions.** In MDP, transition functions and transition probability are defined to identify how the system changes as a result of decisions and information from one state to another [33]. It is described in a stochastic system as $S_{t+1} = S^M\left(S_t, X_t, W_{t+1}\right)$ where S^M is the state transition function and W_{t+1} represents the information uncertainty variable that will arrive between periods t and $t+1$. Here, uncertainties include the effectiveness of advertising packages on sales volume, products price, total market volume, and the competitors' reaction. The competitors' strategy expresses the reactions to the market, which is utterly dependent on our action in the previous period.

In this study, in order to determine the next states, including life cycle state and the BCG class of product m (i.e., the competitors' strategy is independent of the sales volume) we need to obtain sales volume at the next period, which is defined as follows:

$$Y_{t+1}\left(m\right) = \hat{A}\left(S_{tm}, x_{tm}\right) + \tilde{\omega} \tag{6}$$

where $\tilde{\omega}$ is a random variable representing the exogenous information in sales estimation in addition to uncertainties that have been considered to the estimated value; because it is necessary to update the actual sales volume at the next period.

BCG matrix class relies on the growth rate of the market, and the relative market share of the company. These factors are completely dependent on the total market volume indicated by $\Gamma_(tm)$ and predicted by regression or time-series techniques, at each period. This amount is uncertain by nature and determined in Eq 7 where $\hat{\omega}_{tm}$ is a random number added to the predicted amount of total market volume.

$$\Gamma_{tm} = \hat{\Gamma}_{tm} + \hat{\omega}_{tm} \tag{7}$$

Position in the BCG matrix is determined as follows [20]:

$$BCGmatrix = \begin{cases} Question\ marks, & \psi_t > \psi_t^*,\ \Omega_t \le \Omega_t^* \\ Stars, & \psi_t > \psi_t^*,\ \Omega_t > \Omega_t^* \\ Cash\ cows, & \psi_t \le \psi_t^*,\ \Omega_t > \Omega_t^* \\ Dogs, & \psi_t \le \psi_t^*, \Omega_t \le \Omega_t^* \end{cases} \tag{8}$$

$$low\ growth\ market \ \le \psi_t^* \le high\ growth\ market$$

$$low\ market\ share \ \le \Omega_t^* \le high\ market\ share$$

where, ψ_t is the market growth rate, and Ω_t is the relative market share at the beginning of decision period t; ψ_t^* and Ω_t^* represent breaking boundaries of market growth rate and market share, respectively that determine the BCG matrix classes.

In this study, we use the classic version of the life cycle for all our products. In the classic life cycle, stages are obtained only by corresponding sales volume. Boundaries between stages that specify the life cycle states are given by experts. Regarding budgeting, there is only an initial budget at the beginning of our planning horizon, and the available budget at each period is determined according to the previous budget and taken action:

$$B_{t+1} = B_t - c\left(X_t\right) \tag{9}$$

Finally, the best set of advertising packages throughout the whole-time horizon is chosen by this objective function:

$$Max \quad E \sum_{t=1}^{T} Y\left(S_t, X_t\right), \tag{10}$$

where E denotes the expected value of sales over the total period.

To find a solution, the Bellman equation of our objective function is introduced as:

$$V\left(S_t\right) = \max_{\forall X_t} u\left(X_t\right)\left(y\left(S_t,\ X_t\right)\ + \gamma E\left\{V\left(S_{t+1}\right)|S_t, X_t\right\}\right) \tag{11}$$

Eq. (2) indicates our value function that ranks decisions based on present sales estimation over period t and expected future sales, for all products. Thus, $V(S_t)$ includes expected sale value y and estimated value of next states $V\left(S_{t+1}\right)$ that determined according to the best corresponding decisions discounted by γ. As mentioned before, $u(X_t)$ limits our decisions to the allowable budget.

This equation is reformulated to demonstrate expected sale value as:

$$V\left(S_t\right) = max\ u\left(X_t\right)\left(y\left(S_t,\ X_t\right) + \gamma \sum_{s^{,} \in S}\ p(S_{t+1} = s^{,}|S_t, X_t)V\left(S_{t+1}\right)\}\right) \tag{12}$$

$$p\left(S_{t+1} = s^{,}|\mathrm{S_t}, \mathrm{X_t}\right) = \begin{cases} \prod_{m=1}^{M} p\left(\mathrm{cs_{tm}}\right) & \forall \mathrm{m} \ \text{ if } \mathrm{p_{bcgtm}} = 1 \text{ and } \mathrm{p_{lctm}} = 1 \\ & \text{and } Rb_{t+1} = Rb_t - c\left(X_t\right) \\ 0 & \text{Otherwise} \end{cases} \tag{13}$$

p is a part of transition probability function from state S_t to S_{t+1}, where $\mathrm{p_{bcgtm}} = 1$ and $\mathrm{p_{lctm}} = 1$ represent the possibilities regarding life cycle and BCG matrix. That is to say that this probability for states that are not achievable in future periods are not considered. Note that here, we only consider the probability of competitor's reaction to our decisions in the probability transition function. The effectiveness of each decision (advertising package) that changes the sale values (i.e., product life cycle and BCG matrix as well) is considered as an uncertainty of the environment and is addressed in the simulation by generating a random number between the lower bound and upper bound of the effectiveness value. Hence, definitely, the random effectiveness of the packages affects the next state of the system and is taken into account in our simulation.

4.3. **ADP formulation.** Due to the sequential nature of this problem, it can be modeled as dynamic programming; where, the existence of probabilities and uncertainties in the environment makes it stochastic. Primarily, exact dynamic programming algorithms deal only with small state and action spaces. However, in this study, we face a huge sized state space caused by the multiple numbers of products as well as a variety of essential factors impacting our decisions that must be considered as a part of the system state. For example, with four products and 12 periods, the dimension of our problem will be 2.01×10^8. This challenge is known as a curse of dimensionality in the literature and has been addressed by approximate approaches. Among them, ADP is appropriate for highly complex and extremely large problems, particularly in the presence of high uncertainty in the environment [33]. Thus, in the following, we employ an ADP algorithm that fits our problem.

4.4. **Algorithmic strategy.** Algorithm 1 shows a tabular approximate dynamic programming formulation using a lookup-table representation adapted to our problem including budget constraint, probability of competitors' reaction, qualitative state for the BCG matrix and the product life cycle stages, and estimation of products' price. Since our planning horizon is finite, a finite MDP is employed to model it. This algorithm makes decisions using an approximate value function that steps forward in time and considers the value of the states based on their probabilities. As we have access to the model of the environment, we use a model-based algorithm that bootstraps the next steps but does not sample them. In other words, it makes decisions based on the expected value of immediate reward plus value function of the estimated following state that contains the long-term return of our current decision. Note that this is a modified version of the classic AVI [33] while containing a budget constraint for all horizon time (i.e., without this modification the results are not feasible as the algorithm always takes an action with the highest estimated additional value, even if it violates the budget limitation).

In the first step, the value of all states at all periods must be initialized. Also, we need to determine the initial state of our system including the life cycle stage, BCG matrix quadrant, competitors' reaction for each product as well as the total budget. In step 1, at first the effectiveness of different decisions and price of each product are generated, and then the optimization equation is solved within allowable decisions by which sample estimate $\bar{\nu}_t^n$ is calculated. Moreover, as greedy strategies could lead to a local answer, we use a mixed exploitation and exploration strategy. In this strategy, the next step will be determined randomly with probability $\bar{p}$ and using the best decision with probability $1-\bar{p}$. The value of $\bar{p}$ in each iteration is obtained by Eq (14), which demonstrates that in initial iterations, $\bar{p}$ is a significant number and decisions mostly are made randomly to update states as much as possible, it declines further. Therefore in final iterations, the majority of decisions are made based on the highest sales volume.

$$\bar{p}_{n-1} = \frac{b}{b+n-1} \tag{14}$$

In the following, the amount of $\bar{\nu}_t^n$ is applied to update the value function approximation $\bar{V}_t^n(S_t)$ using lookup table representation as

$$\bar{V}_t^n(S_t) = (1-\alpha_{n-1})\,\bar{V}_{t+1}^{n-1}(S_t^n) + \alpha_{n-1}\bar{\nu}_t^n \tag{15}$$

where, α_{n-1} is the corresponding stepsize, which is a significant factor affecting the convergence of the ADP algorithm. This smoothing factor is considered to be harmonic represented by $\alpha_{n-1} = \frac{a}{a+n-1}$, $a \in (0, +\infty)$ where higher amounts of a slow the rate at which α_n drops to zero.

Algorithm 1: An approximate dynamic programming algorithm to solve the media allocation problem

---

Step 0. Initialization:

step 0a : Initialize $\bar{V}_t^0$ for all states S_t ;

step 0b : Initialize S_0 ;

step 0b : set $n = 1.$;

Step 1. Do for $t = 1, 2, \ldots, T$:;

step 1a : For all products generate a sample of $R(s_{tm}, x_{tm})$, ξ_{tm} ;

step 1b : If $c(X_t) > B_t$ then $u(X_t) = 0$, otherwise $u(X_t) = 1$;

step 1c : With probability $(1 - \bar{p}_{n-1})$, choose the best action and compute $\bar{v}_t^n$ (exploitation):

$$\bar{v}_t^n = \max_{X_t} \; u(X_t)\left(y(S_t^n,\ X_t) + \gamma \sum_{s' \in S} p(S_{t+1}^n = s' | S_t^n, X_t)\bar{V}_{t+1}^{n-1}(s')\right)$$

With probability p_{n-1} choose a random decision and compute $\bar{v}_t^n$ (exploration).

step 1d : Update the value function $\bar{V}_t^{n-1}(S_t)$ using

$$\bar{V}_t^n(S_t) = \begin{cases} (1-\alpha_{n-1})\bar{V}_{t+1}^{n-1}(S_t^n) + \alpha_{n-1}\bar{v}_t^n & \mathrm{S}_t = S_t^n \\ \bar{V}_{t+1}^{n-1}(S_t^n) & \text{otherwise} \end{cases}$$

step 1e : Choose a sample path $\omega^n = (\tilde{\omega}_{tm}, \widehat{\omega}_{tm})$

step 1f : Compute $S_{t+1} = S^M(S_t, X_t, W_{t+1})$, where $W_{t+1}(\omega^n)$ is new information.

Step 2. Increment n to $n+1$. if $n \leq N$, go to Step 1.

Step 3. Return the value functions, $\{\bar{v}_t^n = 1, 2, ..., T\}$

The presented algorithm works based on value iteration where decisions are made based on value maximization. However, we can also use different policies which map system states into defined decisions. A policy is defined as a rule or function that determines decisions in different states based on given information [33]. Thus, let X_t^π be a decision function that determines our action at period t based on policy $\pi \in \Pi$, where Π is the set of policies. We are going to find the best policy π^* maximizing the total expected sales value over all periods. Regarding the number of periods T, our objective function is:

$$\max_{\pi\in\Pi} E \sum_{t=1}^{T} y\left(S_t, X_t^{\pi}\right) \tag{16}$$

To determine the following states, X_t^n is obtained based on a policy of the model and a random number added to the state which is considered as exogenous information. We solve this algorithm for a fixed policy, and finally a comparison between different policies is carried out.

5. Numerical example.

In this section, we implement the proposed model on a numerical example to solve the advertising management problem of a selling company. This company manufactures durable and independent products and aims to improve its sales by investing money in promotion. For each product, there are five separate advertising packages of various combinations of media, and managers must decide either among these packages for each product or do nothing in each decision period. The time horizon of the problem is one year which is divided into 12 months so that decisions are made at the beginning of each period.

5.1. Parameters and data.

Advertising packages considered in this study cost differently. These packages are shown in Figure 1 that from left to right, arranged from aggressive to defensive. Moreover, it shows that aggressive packages have higher costs than defensive ones, and have more significant impacts on sales. For example, in package 1, which is the most aggressive, 60 percent of the budget is assigned to TV and radio, 30 percent to outdoor, 5 percent to magazines, and 5 percent to the Internet. In contrast, package 5 targets specified segments and tries to remain the current market state the same as before. In addition, the company does not always have to promote its products and can take no action, called inaction from now on.

The impact of each package on products' sales volume is shown in Table 1 and Table 2, this effectiveness varies among products and depends on life cycle stage, class of BCG matrix, and competitors' reactions (i.e., these are deterministic numbers but intervals between a lower and upper bound). Moreover, competitors' reaction relies on our decisions with different probabilities shown in Table 3. It means that even the current state of the system depends on the action we want to take.

Since the products are durable, the life cycle of products is assumed to be classic and their curves are included in the Appendix 1, which determines the bounds between life cycle stages. Value of growth rate (ψ_t) and relative share market (Ω_t) are obtained by Eq (17) and Eq (18), where total market volume ($\hat{\Gamma}_{tm}$) is estimated (see Appendix 1) using a regression. In the following, as can be seen in Figure 2, our position in the BCG matrix is determined according to them. Moreover, to estimate the price of products, two different regressions are applied to real-world sales data in previous years which is shown in the Appendix 1.

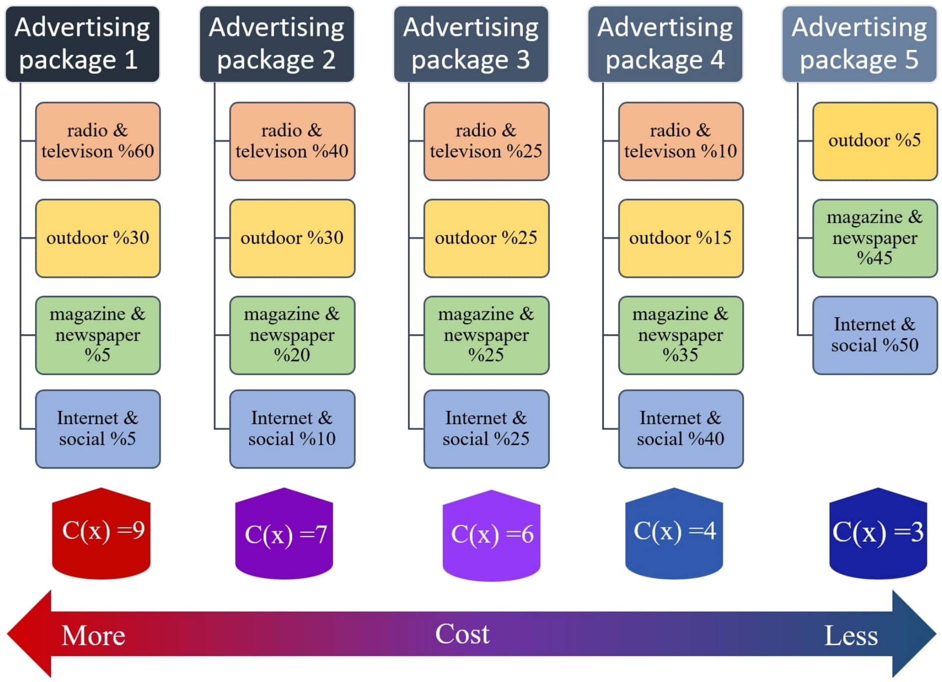


FIGURE 1. Costs and budget percentages of each media in the agency's advertising packages.

TABLE 1. Effectiveness of advertising packages according to the system state for product 1

		Product 1											
		AP 1		AP 2		AP 3		AP 4		AP 5		Inaction	
		r_1	r_2	r_1	r_2	r_1	r_2	r_1	r_2	r_1	r_2	r_1	r_2
Product life cycle	Int	1.12	1.26	1.12	1.21	1.20	1.08	1.02	1.20	1.01	1.15	0.98	1.08
	Gr	1.24	1.51	1.30	1.35	1.25	1.38	1.10	1.43	1.09	1.42	0.98	1.18
	Ma	1.00	1.20	0.98	1.16	0.96	1.15	0.94	1.13	0.93	1.10	0.80	0.93
	Dec	0.90	1.12	0.93	1.05	0.90	1.05	0.88	1.04	0.87	1.01	0.65	0.90
Competitive strategy	H-Def	0.83	1.02	0.88	0.93	0.84	0.92	0.79	0.92	0.78	0.87	0.71	0.84
	H-Off	0.80	0.86	0.70	0.90	0.68	0.81	0.54	0.89	0.53	0.88	0.55	0.70
	L-Def	0.82	0.95	0.80	0.92	0.77	0.88	0.68	0.92	0.67	0.89	0.64	0.78
	L-Off	0.80	0.93	0.78	0.90	0.75	0.85	0.65	0.89	0.64	0.86	0.62	0.76
BCG Matrix	Qus	1.12	1.26	1.11	1.22	1.06	1.16	1.04	1.10	1.01	1.09	0.89	1.09
	Str	1.25	1.52	1.30	1.36	1.25	1.36	1.11	1.44	1.09	1.37	0.95	1.15
	C-Co	1.00	1.29	1.00	1.19	0.98	1.21	0.95	1.18	0.93	1.12	0.89	0.98
	Dg	1.00	1.04	0.96	1.00	0.94	0.98	0.90	0.95	0.88	0.95	0.72	0.80

Int: Introduction, Gr: Growth, Ma: Maturity, Dec: Decline,
H-Def: High defensive H-Off: High offensive L-def: Low defensive L-Off: Low offensive
Qus: Question marks Str: Stars C-Co: Cash cows Dg: Dogs

$$\psi_t = \frac{\hat{\Gamma}_{tm} - Market\ volume\ last\ year\ (t)}{Market\ volume\ last\ year\ (t)} \tag{17}$$

$$\Omega_t = \frac{\mathrm{Y_t\,(m)}}{\hat{\Gamma}_{tm}} \tag{18}$$

TABLE 2. Effectiveness of advertising packages according to the system state for product 2

		Product 2											
		AP 1		AP 2		AP 3		AP 4		AP 5		Inaction	
		r_1	r_2	r_1	r_2	r_1	r_2	r_1	r_2	r_1	r_2	r_1	r_2
Product life cycle	Int	1.17	1.25	1.14	1.18	1.09	1.20	1.04	1.18	1.03	1.12	0.96	1.03
	Gr	1.32	1.40	1.24	1.42	1.25	1.38	1.12	1.40	1.11	1.39	0.99	1.10
	Ma	1.10	1.28	1.14	1.16	1.09	1.17	1.00	1.18	1.00	1.15	0.87	1.00
	Dec	0.98	1.03	0.94	1.03	0.90	1.00	0.85	1.00	0.80	1.00	0.77	0.92
Competitive strategy	H-Def	0.80	1.04	0.90	0.90	0.86	0.89	0.78	0.92	0.80	0.88	0.73	0.84
	H-Off	0.75	0.88	0.73	0.84	0.70	0.82	0.60	0.84	0.55	0.85	0.60	0.69
	L-Def	0.77	0.99	0.80	0.90	0.81	0.86	0.76	0.86	0.75	0.83	0.68	0.78
	L-Off	0.74	0.94	0.70	0.90	0.75	0.82	0.74	0.80	0.70	0.77	0.62	0.76
BCG Matrix	Qus	1.18	1.30	1.15	1.20	1.12	1.19	1.06	1.16	1.03	1.12	0.92	1.08
	Str	1.30	1.45	1.25	1.43	1.26	1.39	1.13	1.31	1.12	1.21	0.98	1.05
	C-Co	1.14	1.24	1.14	1.18	1.12	1.17	1.06	1.15	1.04	1.12	0.90	0.99
	Dg	1.00	1.10	0.98	1.05	0.96	1.04	0.90	0.99	0.91	0.95	0.68	0.88

TABLE 3. Competitors' reactions probabilities

	Product 1						Product 2					
	AP 1	AP 2	AP 3	AP 4	AP 5	Inaction	AP 1	AP 2	AP 3	AP 4	AP 5	Inaction
H-Def	0.15	0.2	0.22	0.21	0.25	0.3	0.18	0.2	0.21	0.23	0.26	0.32
H-Off	0.35	0.3	0.22	0.14	0.12	0.3	0.3	0.25	0.19	0.12	0.1	0.16
L-Def	0.3	0.3	0.3	0.34	0.36	0.2	0.2	0.35	0.33	0.3	0.24	0.23
L-Off	0.2	0.2	0.26	0.31	0.27	0.2	0.32	0.2	0.27	0.35	0.4	0.29

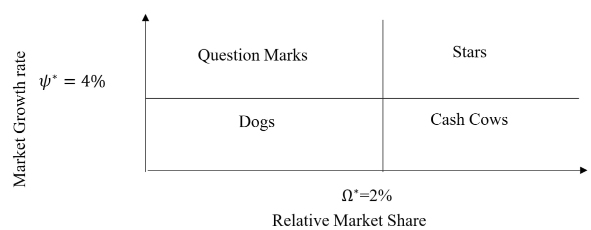


FIGURE 2. The *BCG* matrix of the company studied and its parameters

To evaluate the results, since our approach is model-based (i.e., we have access to dynamics of the environment) we only need to simulate the algorithm but not the environment. After having formulated our algorithm and adjust the parameters for our numerical example, we solve the problem using a *dual Intel Core i3-7100 CPU 3.90 GHz* workstation having *4* gigabytes of *RAM* and *Python 3.6*. In the following, we present the numerical results in addition to insights that our interpretation of the learned optimal policies and value functions.

5.2. **Computational results.** To find a converged answer, we let the algorithm run for 30 million iterations, where in each state, decisions are made basically according to the expectation maximization of the value function, and also a small percentage of decisions are chosen randomly to update the possible states. As a result, the amount of stepsize is considered fairly significant to explore more in initial iterations and exploit more in the last iterations. In other words, the algorithm is converged when the most probable states get updated several times. The results are shown in Figure 3 where the objective value levels off just after 1×10^7 iterations at around 4.05×10^4, which shows the total revenue over time.

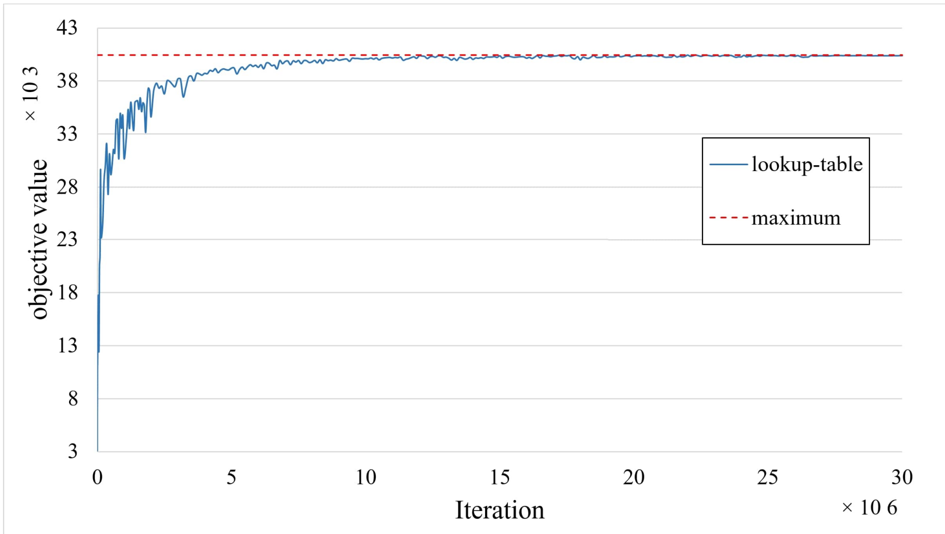


FIGURE 3. Appropriate number of iterations to get a converged answer

5.2.1. *Step size and exploration.* As said in section 4.4, we utilize a harmonic stepsize rule in the look-up table ADP algorithm. Due to the high impact of stepsize on the convergence of the model, we plot the approximate value function up to 10 million iterations, under the harmonic stepsize rule with different amounts its parameter. Moreover, different probabilities of random decisions have been considered to find an appropriate amount of p and avoid local answers.

In Figure 4, compression between the different amount of parameters a and b is carried out, and as it can be seen when a is small (e.g., a = 1.5 Million and b = 750 Thousand), the algorithm generally does not perform well enough and can converge to a local optimum, as the stepsize decreases to a small number too fast. A larger value of the parameter a prevents the stepsize from dropping too fast and helps to create a better convergence. However, the too large amount of a generates a long stream of large stepsizes which causes fluctuation in the answers as it is highly sensitive to new observations. In general, an appropriate value of the parameter a depends on the number of iterations and the expected rate of convergence. Moreover, this figure shows that lower amounts of b, which means to go further based on the best decision, leads to local answers. The reason is that the algorithm chooses the states that are obtained by the best actions and updates the same states in different iterations, meaning that many states are not updated. On the other hand, a substantial value of b causes a significant fluctuation at beginning iterations due to a high probability of making random decisions.

In each stage, decisions are made according to the state of the system. Although there are a massive number of states in our model, we just consider those that have been updated more than a certain number. In the following, the percentage of selected decisions in different periods is shown in Figure 5 for product 1 and product 2. As can be seen, in the first period, decisions are certain because the system's state is precisely known. However, in the following stages, decisions can be varied according to uncertainties considered in our problem. Aggressive packages have been roughly chosen in the middle stages which are matched with the high price and life cycle stage (see Appendix 1).

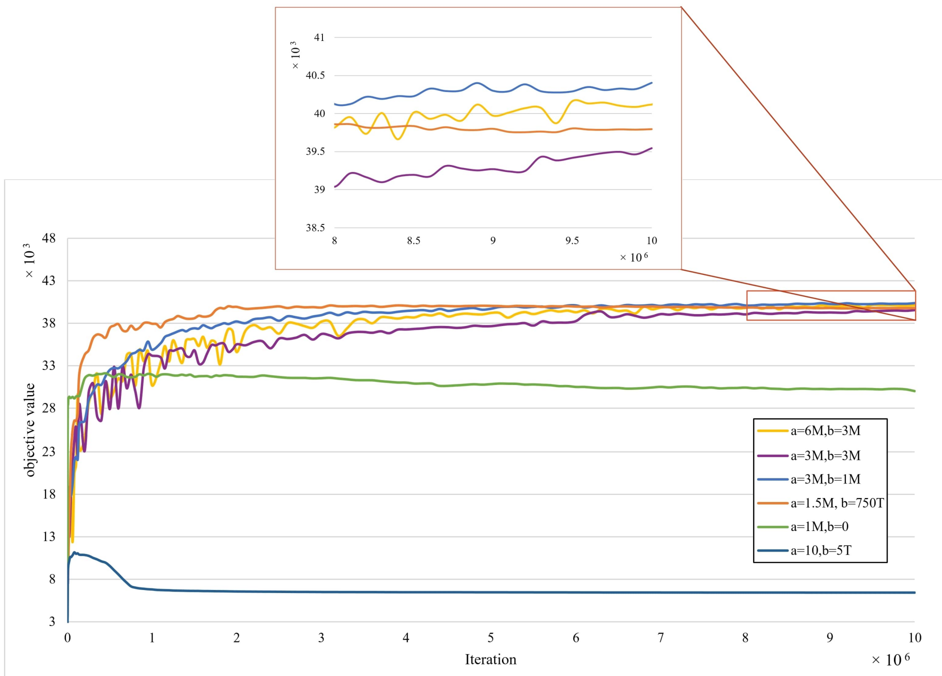
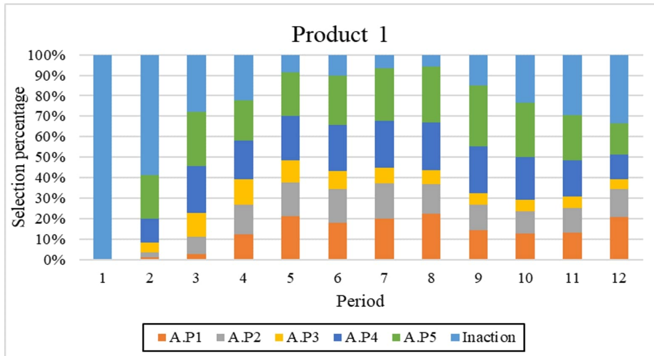


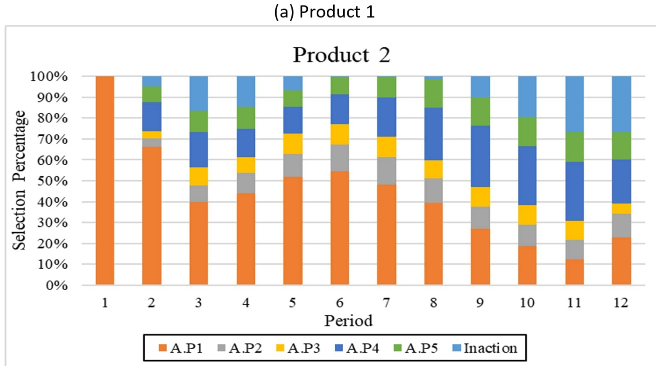
FIGURE 4. Sales volume with different number of parameters a and b

Concerning different costs and effectiveness of the packages, we plot the average cost and added value function during the time. The additional value of each stage is the difference between the current value and the previous one, which shows the value is made by decisions in this stage. What stands out from Figure 6 is that the majority of the budget is assigned in the middle stages and therefore a more prominent profit is made in these stages which can be justified according to the price and life cycle diagrams, as the price and the sale volume are both significantly large in middle stages which convinces the model to assign more budget then. Also, it is specified that the average cost for product 2 is more significant than that of product 1.

One of the most critical factors in our problem is the budget and accordingly, we plot the percentage of decisions for both products, as shown in Figure 7. As can be seen, when the budget is high enough, the model chose more aggressive packages for product 2. However, aggressive packages for product 1 are mostly chosen in middle budget levels.

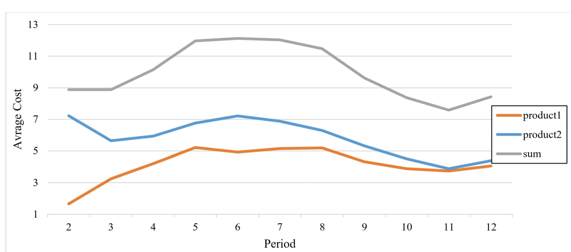


(a) Product 1

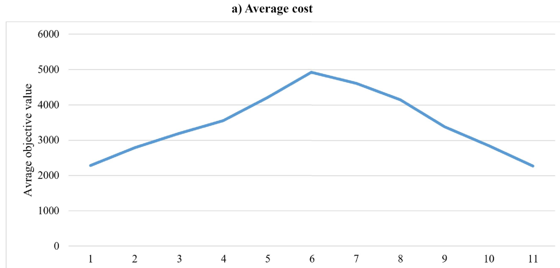


(b) Product 2

FIGURE 5. Percentage of selected advertising packages in each period

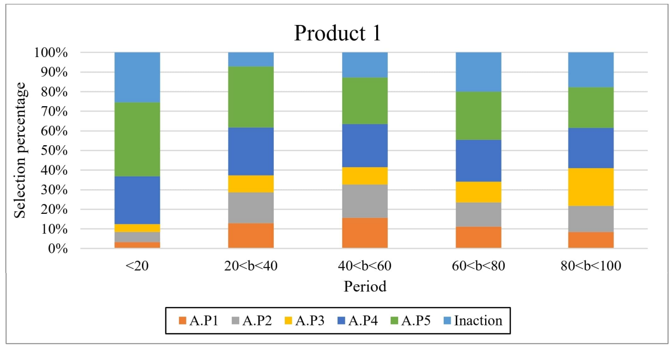


a) Average cost

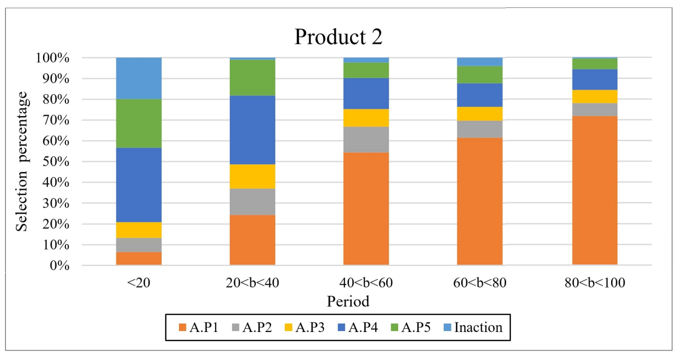


b) Average Objective value

FIGURE 6. Costs and additional value in each period



a) product 1



b) product 2

FIGURE 7. Percentage of selected advertising packages in different budget levels

In addition to applying the AVI approach to our problem, we utilize an API algorithm to discuss the results for different policies and obtain the most appropriate one. In such a way, decisions are made based on a certain policy; for example, price policy is a rule based on which decisions are made, and other factors are not considered. In this paper, we solve the problem with four different policies that are as follows: a. life cycle policy, b. BCG policy, c. competitors' reaction policy, and d. price policy. Moreover, in all these policies, budget constraint is also considered. Finally, the algorithm is solved for policies, and its results are shown in Figure 8, where competitors' reaction policy has the best result, and life cycle and price policies are closed to it but BCG policy yields poor results.

5.2.2. *Evaluation.* To evaluate the proposed method, we compare the results with a heuristic approach under simplified circumstances. We employ a genetic algorithm to make advertising decisions (i.e., selecting an advertising package among available options at the beginning of each decision time step), assuming that our actions have no effect on competitors' actions, and for the sake of simplicity, instead of stochastic pay off of each action we consider a deterministic – an expected – value. To do so, we restrict the decisions to a limited budget shared among the whole period, and also, we put an upper bound for the market share to avoid unrealistic results caused by independent reactions of competitors (See Appendix 2 for more details). Using the genetic algorithm, we assume that, the possibilities are fixed, competitors have been removed and for each product, a maximum of 10% of the total market is allocated. Figure 9 shows the comparison between the results by of proposed approach – ADP – and a genetic algorithm. The genetic algorithm converges extremely fast while ADP needs a higher number of iterations (i.e., due to the exploration and exploitation strategies). Finally, both methods converge to the same objective value indicating the fact that after training, policies of ADP are approximately optimal. However, in practical cases, where the uncertainty of payoffs is significantly high and players are interconnected, we need a dynamic approach like ADP which is reliable even in a highly uncertain environment.

6. **Conclusion.** This study develops a multi-period advertising budget allocation for different products with the aim of maximizing the total profits. To make the best decision, we consider four primary marketing factors: life cycle, BCG matrix, competitors' reaction, and available budget. We assume that there are multiple independent products with different price profiles sharing the same advertising budget and to avoid semi-optimal answers, decisions for all products must be taken jointly. The most challenging parts of this problem are its dynamic nature and considerable uncertainties in real-world cases in addition to the huge size of the problem. To deal with them, we envelope an ADP algorithm to optimally plan marketing decisions in such a stochastic environment.

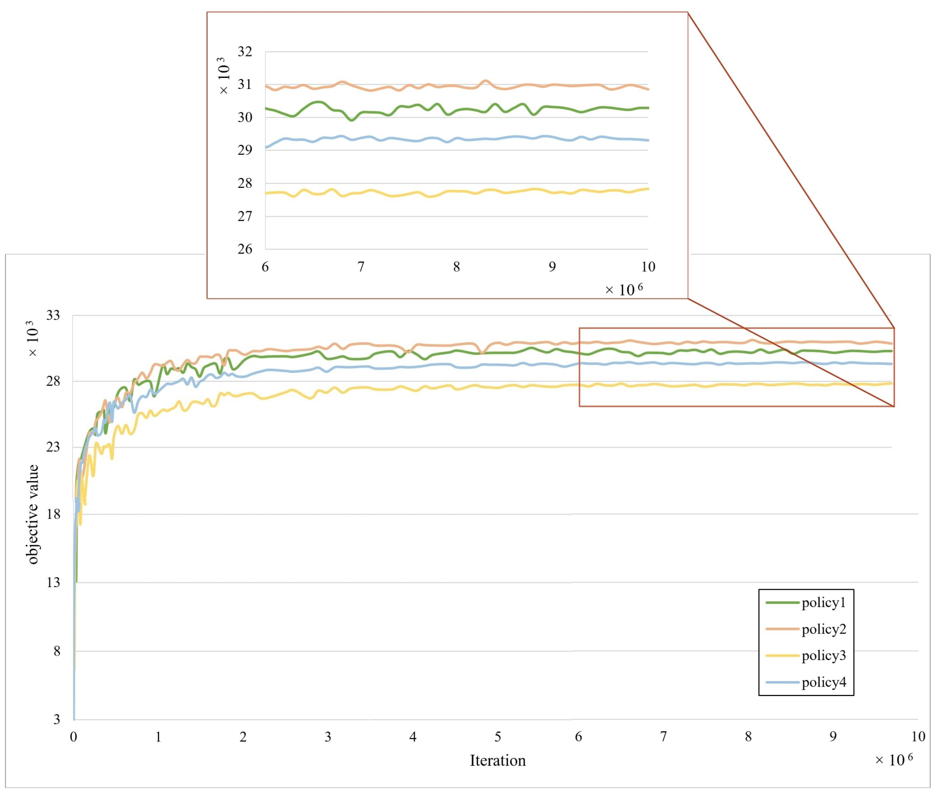


FIGURE 8. A comparison between different policies

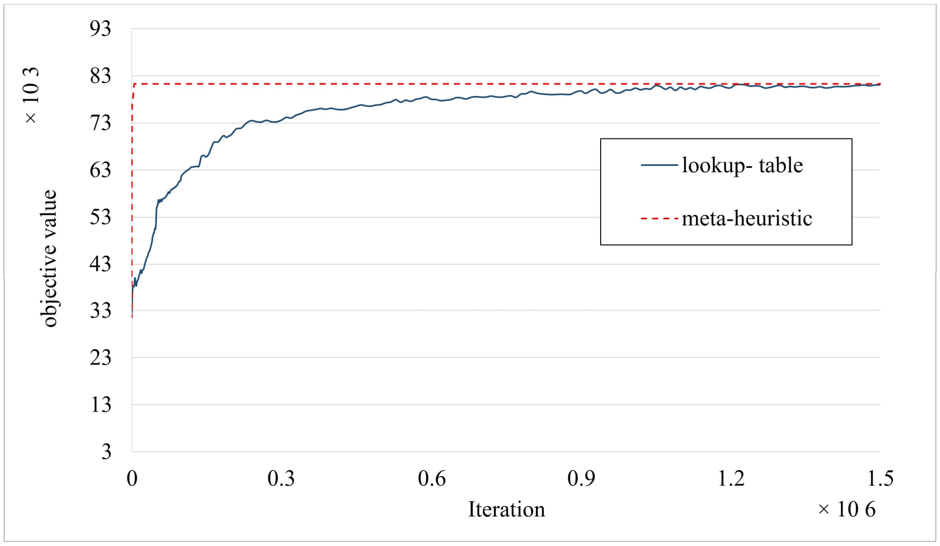


FIGURE 9. Comparison between the results by our proposed approach – ADP – and a genetic algorithm

Next, a case study is presented for two products and twelve-time steps (i.e., monthly) when decisions are chosen among five different advertising packages. This problem is first solved using AVI algorithm and the results indicate that it converges successfully to the optimal solution. The model majorly allocates more budget to product 2 rather than product 1 due to its special price profile, life cycle, and the effectiveness of packages for the product. With respect to the time, it can be seen that more aggressive packages were selected in the middle stages to impact as much as possible.

To give more satisfactory answers to the managers, we introduce several rule-based policies (e.g., a price-based policy selects packages just based on the prices of products) and compare them together and with the proposed model. In our problem, the life cycle and competitor policies seem to gain more desired results than others. Finally, to validate the model, the results are assessed against a heuristic algorithm considering simplified assumptions, showing that our approach yields the optimal solutions.

Future studies are recommended to consider other factors such as the time of introducing the new product, pricing policies, inflation, and unexpected events. Regarding the methodology, this work could be improved by developing multi-agent learners competing with each other in a shared market instead of modeling competitor's reactions as exogenous information. Also using model-free algorithms such as reinforcement learning could add more values in the case of complex and unknown reward and transition functions.

Appendixes.

Appendix 1. Classical life cycle is included four stages based on sales volume of each products that is shown in Figure 10. Note that for each of these product initial sales has been considered at the beginning of planning time horizon.

TABLE 4. The market volume over the past year and estimation for future periods

Periods	1	2	3	4	5	6	7	8	9	10	11	12
MV	3.18*	2.36	2.15	2.24	2.13	1.90	1.64	1.49	1.36	1.37	1.53	2.27
MV(LY)	2.88*	2.07	1.85	1.98	1.88	1.70	1.50	1.45	1.43	1.49	1.69	2.47

$*:\times 10^4$, MV: Market value, MV(LY): Market volume last year.

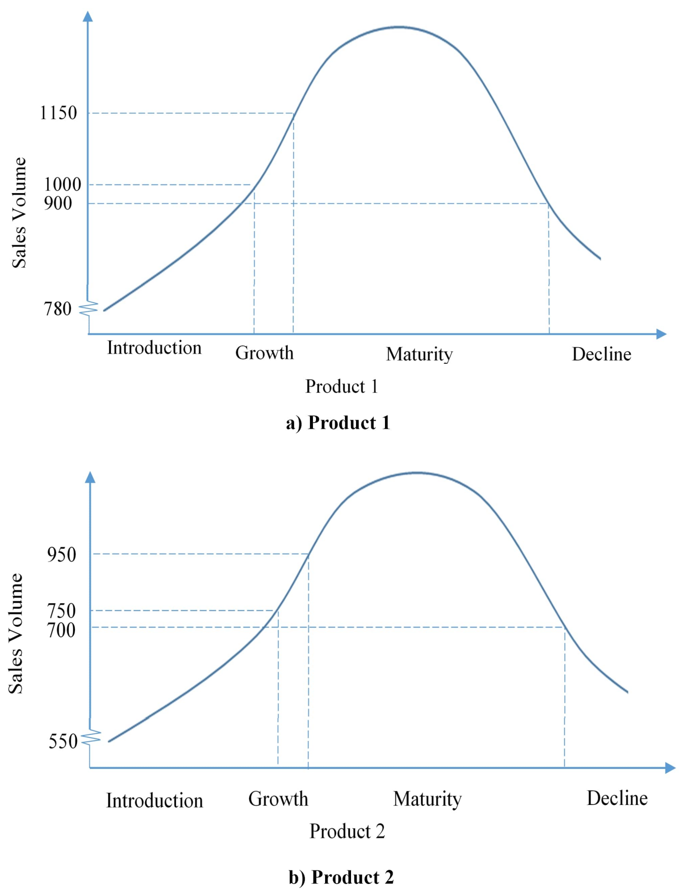


FIGURE 10. Classical life cycle curves

Statistics in Table 4 show the estimation of monthly total market volume for the planning horizon and the actual corresponding amount for previous year, which is used to determine BCG matrix class in the proposed model.

Figure 11 (a) and (b) indicate the price changes over the course of time. It can be seen that for both products, having increased at first the price is followed with a decline after almost 6^{th} month.

TABLE 5. Policy 1 based on budget and life cycle stages

Budget bound	Product 1				Product 2			
	Int	Gr	Ma	Dec	Int	Gr	Ma	Dec
0-20	AP 5	AP 4	AP 3	AP 2	AP 5	AP 4	AP 3	AP 3
20-40	AP 4	AP 3	AP 2	AP 2	AP 4	AP 3	AP 3	AP 2
40-60	AP 3	AP 2	AP 2	AP 1	AP 3	AP 3	AP 2	AP 1
60-80	AP 2	AP 1	AP 1	AP 1	AP 2	AP 2	AP 1	AP 1
80-100	AP 2	AP 1	AP 1	AP 1	AP 1	AP 1	AP 1	AP 1

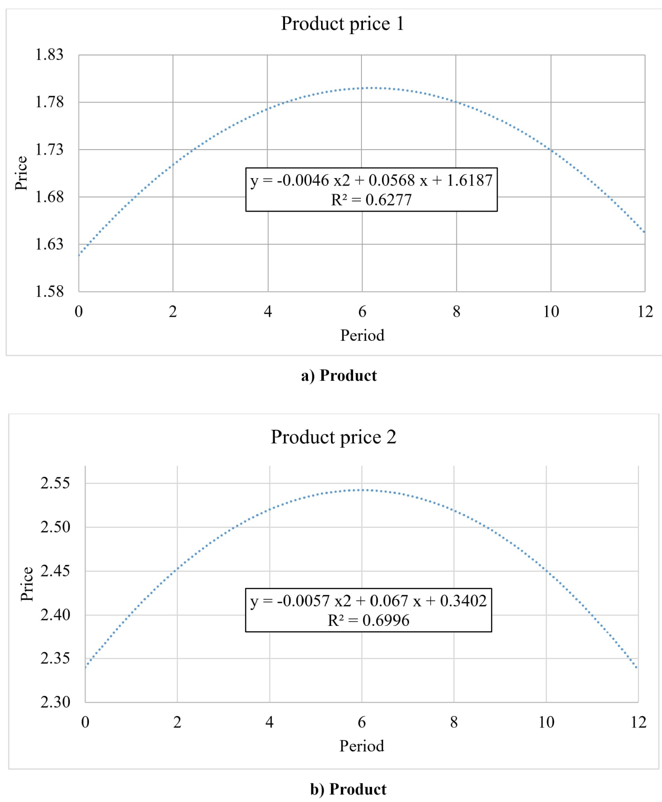


FIGURE 11. Price forecast regression in 12 future decision periods

Through Table 5 to Table 8, different policies used in our model is given determining that in various states which package should be chosen. In addition to the system budget all these policies are based on one of the contributing factors such as life cycle, competitor's reaction, BCG matrix and price level.

Appendix 2. We use a classical genetic algorithm in this paper to benchmark the proposed approach against a heuristic decision-making strategy. We first initialize the first generation with a random uniform selection of product advertising package 1 and 2 for 12 periods sets generated between 0 and 5. An example is showed in Figure 12.

Table 6. Policy 2 based on budget and competitors' reaction

Budget bound	Product 1				Product 2			
	H-Def	H-Off	L-Def	L-Off	H-Def	H-Off	L-Def	L-Off
0-20	AP 5	AP 4	AP 5	AP 4	AP 5	AP 3	AP 4	AP 4
20-40	AP 5	AP 3	AP 4	AP 3	AP 4	AP 3	AP 3	AP 3
40-60	AP 4	AP 2	AP 3	AP 3	AP 3	AP 2	AP 2	AP 3
60-80	AP 3	AP 2	AP 2	AP 2	AP 2	AP 1	AP 2	AP 1
80-100	AP 2	AP 1	AP 2	AP 2	AP 1	AP 1	AP 1	AP 1

Table 7. Policy 3 based on budget and BCG matrix class

Budget bound	Product 1				Product 2			
	Qus	Str	C-Co	Dg	Qus	Str	C-Co	Dg
0-20	AP 5	AP 5	AP 5	AP 5	AP 2	AP 2	AP 5	AP 5
20-40	AP 5	AP 4	AP 5	AP 4	AP 1	AP 1	AP 4	AP 4
40-60	AP 4	AP 3	AP 4	AP 2	AP 1	AP 1	AP 4	AP 3
60-80	AP 3	AP 2	AP 3	AP 1	AP 1	AP 1	AP 3	AP 2
80-100	AP 2	AP 1	AP 2	AP 1	AP 1	AP 1	AP 2	AP 1

Table 8. Policy 4 based on budget and price product

Budget bound	Product 1 (price)			Product 2 (price)		
	[0, 1.67]	[1.67, 1.73]	[1.73, $+\infty$)	[0, 2.4]	[2.4, 2.47]	[2.47, $+\infty$)
0-20	AP 5	AP 4	AP 3	AP 5	AP 5	AP 3
20-40	AP 5	AP 4	AP 2	AP 5	AP 4	AP 3
40-60	AP 4	AP 3	AP 2	AP 4	AP 3	AP 2
60-80	AP 4	AP 3	AP 2	AP 3	AP 2	AP 1
80-100	AP 3	AP 2	AP 1	AP 3	AP 1	AP 1

Advertising package

Product 1	1	2	3	4	5	6	7	8	9	10	11	12
	0	0	1	1	3	2	1	5	5	2	4	0
Product 2	1	2	3	4	5	6	7	8	9	10	11	12
	1	1	3	2	1	5	2	5	4	0	3	1

Figure 12. An example of initial chromosomes for two products

To evaluate each chromosome, we use a fitness function that firstly checks the feasibility of randomly generated answers regarding the budget constraint (we regenerate the infeasible chromosomes to find the possible ones). After that, sales value (fitness value), product life cycle and BCG matrix are determined for each product. To consider the effect of each package on the sales values we use a deterministic efficiency (i.e., the average of interval efficiency). Another modification is that we consider the competitors' reaction impact as a static factor. The genetic algorithm selects the parents among high value chromosomes to create the next generation (i.e., this selection is based on a probability function, the higher fitness function, the higher chance for selecting as a parent). To generate new chromosomes, we consider four crossover function. The first one is a four-point crossover, where each gene is three cuts, and their average is considered as new values (see

Figure 13). The second one is a point averaging, where parents are combined together and individual rounded average for each gene is considered (see Figure 14). The third one is chromosome symmetry, where genes are mirrored based on the one in the middle. Finally, the fourth one is a strand crossover; we exchange the string of products (see Figure 15).

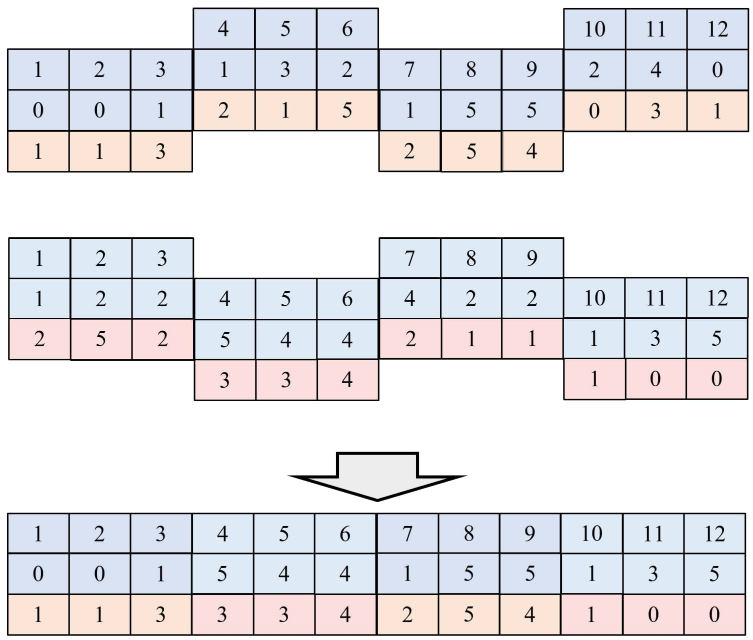


FIGURE 13. A four-point crossover example

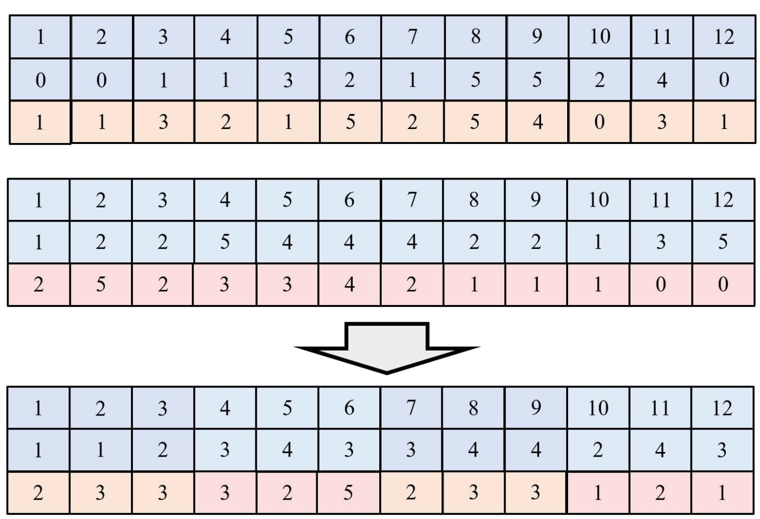


FIGURE 14. A point average crossover example

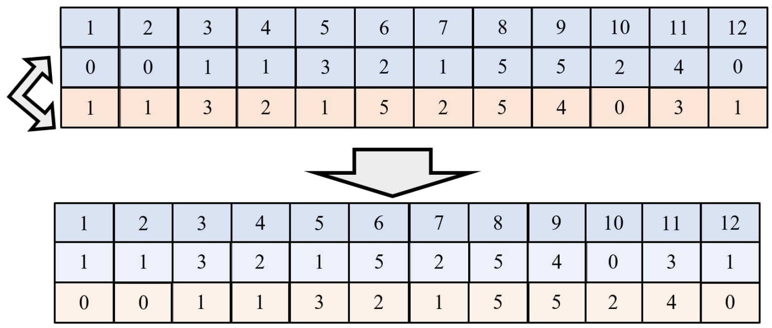


FIGURE 15. A strand crossover example

REFERENCES

[1] V. S. Abedi, Allocation of advertising budget between multiple channels to support sales in multiple markets, *J. Operational Research Society*, **68** (2017), 134–146.

[2] A. Albadvi and H. Koosha, A robust optimization approach to allocation of marketing budgets, *Management Decision*, **49** (2011), 601–621.

[3] T. Araujo, J. R. Copulsky, J. L. Hayes, S. J. Kim and J. Srivastava, From purchasing exposure to fostering engagement: Brand–consumer experiences in the emerging computational advertising landscape, *J. Advertising*, **49** (2020), 428–445.

[4] W. Arens and L. B. Courtland, *Contemporary Advertising*, 5^{nd} edition, McGraw-Hill Education, 1994.

[5] F. M. Bass, N. Bruce, S. Majumdar and B. P. S. Murthi, Wearout effects of different advertising themes: A dynamic bayesian model of the advertising-sales relationship, *Marketing Science*, **26** (2007), 179–195.

[6] F. M. Bass and R. T. Lonsdale, An exploration of linear programming in media selection, *Mathematical Models in Marketing*, **132** (1976), 137–139.

[7] R. Bellman and S. Dreyfus, Functional approximations and dynamic programming, *Math. Tables Aids Comput.*, **(**13) (1959), 247–251.

[8] C. Beltran-Royo, L. F. Escudero and H. Zhang, Multiperiod multiproduct advertising budgeting: Stochastic optimization modeling, *Omega*, **59** (2016), 26–39.

[9] C. Beltran-Royo, H. Zhang, L. A. Blanco and J. Almagro, Multistage multiproduct advertising budgeting, *European J. Oper. Res.*, **225** (2013), 179–188.

[10] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 4^{th} edition, Athena Scientific, 2012.

[11] P. J. Danaher and R. Rust, Determining the optimal return on investment for an advertising campaign, *European J. Oper. Res.*, **95** (1996), 511–521.

[12] P. Doyle and J. Saunders, Multiproduct advertising budgeting, *Marketing Science*, **9** (1990), 97–113.

[13] R. Du, Q. Hu and S. Ai, Stochastic optimal budget decision for advertising considering uncertain sales responses, *European J. Oper. Res.*, **183** (2007), 1042–1054.

[14] M. Fischer, S. Albers, N. Wagner and M. Frie, Dynamic marketing budget allocation across countries, products, and marketing activities, *J. Marketing Research*, 2009.

[15] H. K. Gajjar and K. G. Adil, A dynamic programming heuristic for retail shelf space allocation problem, *Asia-Pac. J. Oper. Res.*, **28** (2011), 183–199.

[16] T. P. Hsieh, C. Y. Dye and K. K. Lai, A dynamic advertising problem when demand is sensitive to the credit period and stock of advertising goodwill, *J. Operational Research Society*, **71** (2020), 948–966.

[17] K. F. Jea, J. Y. Wang and C. W. Hsu, Two-agent advertisement scheduling on physical books to maximize the total profit, *Asia-Pac. J. Oper. Res.*, **36** (2019), 1950014.

[18] A. Kaul, S. Aggarwal, M. Krishnamoorthy and P. C. Jha, Multi-period media planning for multi-products incorporating segment specific and mass media, *Ann. Oper. Res.*, **269** (2018), 317–359.

[19] H. Koosha and A. Albadvi, Allocation of marketing budgets to maximize customer equity, *Operational Research*, **20** (2020), 561–583.

[20] P. Kotler and K. L. Keller *Marketing Management*, 14^{th} edition, Prentice Hall, 2012.

[21] N. K. Kwak, C. W. Lee and J. H. Kim, An MCDM model for media selection in the dual consumer/industrial market, *European J. Operational Research*, **166** (2005), 255–256.

[22] G. Li and B. Sun, Optimal dynamic pricing for used products in remanufacturing over an infinite horizon, *Asia-Pac. J. Oper. Res.*, **31** (2014), 1450012.

[23] X. Li, Y. Li and W. Cao, Cooperative advertising models in O2O supply chains, *International J. Production Economics*, **215** (2019), 144–152.

[24] F. Lu, J. Zhang and W. Tang, Wholesale price contract versus consignment contract in a supply chain considering dynamic advertising, *Int. Trans. Oper. Res.*, **26** (2019), 1977–2003.

[25] P. Manik, A. Gupta, P. C. Jha and K. Govindan, A goal programming model for selection and scheduling of advertisements on online news media, *Asia-Pac. J. Oper. Res.*, **33** (2016), 1650012.

[26] M. Memarpour, E. Hassannayebi, N. F. Miab and A. Farjad, Dynamic allocation of promotional budgets based on maximizing customer equity, *Operational Research*, **21** (2021), 2365–2389.

[27] H. I. Mesak and H. Zhang, Optimal advertising pulsation policies: A dynamic programming approach, *J. Operational Research Society*, **52** (2001), 1244–1255.

[28] A. Mihiotis and I. Tsakiris, A mathematical programming study of advertising allocation problem, *Appl. Math. Comput.*, **148** (2004), 373–379.

[29] P. A. Naik, K. Raman and R. S. Winer, Planning marketing-mix strategies in the presence of interaction effects, *Marketing Science*, **24** (2005), 25–34.

[30] M. Nerlove and K. J. Arrow, Optimal advertising policy under dynamic conditions, *Mathematical Models in Marketing*, **132** (1962), 167–168.

[31] B. Pérez-Gladish, I. González, A. Bilbao-Terol and M. Arenas-Parra, Planning a TV advertising campaign: A crisp multiobjective programming model from fuzzy basic data, *Omega*, **38** (2010), 84–94.

[32] W. B. Powell, An operational planning model for the dynamic vehicle allocation problem with uncertain demands, *Transportation Research Part B: Methodological*, **21** (1987), 217–232.

[33] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, John Wiley & Sons, 2007.

[34] A. Prasad and S. P. Sethi, Competitive advertising under uncertainty: A stochastic differential game approach, *J. Optim. Theory Appl.*, **123** (2004), 163–185.

[35] M. L. Puterman, Markov decision processes, *Handbooks Oper. Res. Management Sci.*, **2** (1990), 331–434.

[36] J. Z. Sissors and R. B. Baron, *Advertising Media Planning*, 7^{th} edition, Mc Graw hill Publishin, 2010.

[37] S. Sriram and M. U. Kalwani, Optimal advertising and promotion budgets in dynamic markets with brand equity as a mediating variable, *Management Science*, **53** (2007), 46–60.

[38] R. S. Sutton and A. G. Barto, Toward a modern theory of adaptive networks: Expectation and prediction, *Psychological Review*, **88** (1981), 135–170.

[39] J. N. Tsitsiklis, Asynchronous stochastic approximation and Q-learning, *32nd IEEE Conference on Decision and Control*, **16** (1994), 185–202.

[40] R. Van der Wurff, P. Bakker and R. G. Picards, Economic growth and advertising expenditures in different media in different countries, *J. Media Economics*, **21** (2008), 28–52.

[41] M. L. Vidale and H. B. Wolfe, An operations-research study of sales response to advertising, *Operations Research*, **5** (1957), 370–381.

[42] X. Wang, F. Li and F. Jia, Optimal advertising budget allocation across markets with different goals and various constraints, *Complexity, 2020*, 2020.

[43] C. Yang and Y. Xiong, Nonparametric advertising budget allocation with inventory constraint, *European J. Oper. Res.*, **285** (2020), 631–641.

[44] Y. Yang, B. Feng, J. Salminen and B. J. Jansen, Optimal advertising for a generalized Vidale–Wolfe response model, *Electronic Commerce Research*, **285** (2021), 1–31.

[45] T. Zhao, W. Zhang, H. Zhao and Z. Jin, A reinforcement learning-based framework for the generation and evolution of adaptation rules, *In 2017 IEEE International Conference on Autonomic Computing (ICAC)*, (2017), 103–112.

Received March 2021; revised August 2021; early access November 2021.

E-mail address: makhz2014@gmail.com
E-mail address: hosseinneghabi@um.ac.ir
E-mail address: ramin.ahadi.z@gmail.com