Media Mix Optimization
Applying a Quadratic Knapsack Model

Ulrich Pferschy[1], Joachim Schauer[1] and Gerhild Maier[2]
[1]*Department of Statistics and Operations Research, University of Graz, Universitaetsstrasse 15, 8010 Graz, Austria*
[2]*UPPER Network GmbH, Seering 7/2, 8141 Unterpremstätten, Austria*

Keywords: Advertising Media Optimization, Quadratic Knapsack Problem, Genetic Algorithm.

Abstract: In this contribution we present an optimization model for deciding on the best selection of advertising media to be used in a promotional campaign. The effect of each single medium and each pair of media is estimated from the evaluation data of past campaigns taking into account a similarity measure between the attributes and goals of campaigns. The resulting discrete optimization model is a Quadratic Knapsack Problem which we solve by a genetic algorithm. Then campaign budget is assigned to each selected advertising medium based on a statistical estimation from previous campaigns. Our optimization tool is integrated in the marketing management software solution MARMIND.

1 INTRODUCTION

Marketing is a crucial aspect for every company to sell its products, whatever industry or market it is concerned with. However, as a famous quotation (sometimes attributed to Henry Ford) states: "Half the money I spend on advertising is wasted, the trouble is, I don't know which half". Indeed, it is a central question of marketing management how to use the budget of a promotional campaign. In particular, the available options have increased considerably in the last decade with new possibilities such as targeted social media advertising and context sensitive web banners. Thus, the suitable selection of advertising media for a promotional campaign, i.e. deciding on the *media mix*, has become an increasingly complex task with only limited information on the actual impact of a medium on the goals of the campaign.

Contributions to finding the best media mix were given for particular industry sectors, e.g. in (Färe et al., 2004) and (Reynar et al., 2010), and from an optimization point of view in several papers going back to (Balachandran and Gensch, 1974) and more recently e.g. by (Sorato and Viscolani, 2011), (Nobibon et al., 2011) and (Sönke, 2012).

The software platform *MARMIND* produced and offered by *UPPER Network* [1] provides a wide range of tools to support the daily tasks of a marketing department from planning to realization. In collaboration with the University of Graz, Austria, an optimization tool was developed and added to the solution which computes a suggestion for the media mix of a planned promotional campaign. This tool is now an integral part of MARMIND and starts being used by marketing managers.

A central question of marketing planning concerns the effect and efficiency of advertising media (see e.g. the survey paper (Vakratsas and Ambler, 1999) and (Pergelova et al., 2010) on internet advertisements). While many statistical methods have been employed to find partial answers to this questions, these require survey data or other means of market research, which is usually not available for the full range of marketing options available to the decision maker in a typical planning scenario. Therefore, we aim to gain information from past campaigns.

The main outline of the optimization tool works as follows. MARMIND keeps a data base of all past promotional campaigns with ratings of their overall success and an evaluation of the different goals of the campaign. Based on these observations of past campaigns, we estimate the effect of every advertising medium for the currently planned campaign. To this end we take the "similarity" between planned and past campaigns into account. Moreover, we derive estimations for the pairwise effect of advertising media, since many media influence each other or

[1] `www.uppernetwork.com`

Pferschy U., Schauer J. and Maier G..
Media Mix Optimization - Applying a Quadratic Knapsack Model.
DOI: 10.5220/0004825803630370
In *Proceedings of the 3rd International Conference on Operations Research and Enterprise Systems* (ICORES-2014), pages 363-370
ISBN: 978-989-758-017-8

are dependent on each other and thus cannot be separated into unconnected decisions. Based on these effect estimations we draw up an optimization model which turns out to be a *Quadratic Knapsack Problem* (QKP). After solving this model by an improved genetic algorithm we assign the available budget to the selected advertising media by considering the proportional budget allocation of past campaigns.

In-house tests indicate that the media mix selected by the optimization tool gets highly positive appraisals from experts in the field. The various possibilities of parametrization allow a flexible adaptation for every domain.

2 FORMAL PROBLEM FORMULATION

In our setting a promotional campaign is described by a number of attributes, some of them represented by nominal values such as target groups, product classes and general strategic goals, others expressed by numerical values such as desired market share, increase in revenue, etc.

Formally, a promotional campaign t is defined by a k-dimensional vector of parameters $t(1),\ldots,t(k)$, where for some fixed k' with $0 \leq k' \leq k$ there are nominal values $t(1),\ldots,t(k')$ and positive cardinal values $t(k'+1),\ldots,t(k)$. A campaign may also consist of only a subset of these parameters and leave the remaining entries of the vector empty.

To express and measure the goals of promotional campaigns there is set of operative goals $g_1,\ldots,g_\ell$ defined such as number of new customers, awareness level, number of repeat customers, etc. Each promotional campaign t is assigned a subset G_t of these operative goals with $\ell_t := |G_t|$. For convenience we impose an upper bound $\ell_t \leq L$ on the number of selected goals, which is of moderate size in practice (think of single digit numbers), i.e. $L \ll \ell$. Furthermore, the chosen goals in G_t are ranked in a total ordering to indicate their relative importance. This preference relation between goals is represented by a rank number $r_t(g_j)$ for each goal $g_j \in G_t$, where $r_t = \ell_t$ signifies the most important, i.e. highest ranked, goal and $r_t = 1$ the least important. Clearly, each number in $1,\ldots,\ell_t$ is assigned to exactly one goal as a rank r_t.

Finally, there is a total budget B_t given for the promotional campaign t.

After completion of the promotional campaign t the responsible manager should be able to state the degree of achievement of each operative goal $g_j \in G_t$ of the campaign by assigning a numerical value representing the achieved percentage of the goal. For simplicity we will assume that this value is scaled into an achievement level $a_t(g_j) \in [0,1]$ with $a_t(g_j) = 1$ indicating perfect achievement of goal g_j. In addition, the marketing manager will be asked to evaluate the overall success of a completed promotional campaign by assigning a discrete value $s_t \in \{1,\ldots,S\}$, where S indicates the best outcome and 1 the worst. Usually, S is a single digit number.

Of course, it would be desirable to extract more information on the impact of the applied advertising media. However, one should keep in mind that an overly complicated feedback system will often be ignored or filled with data of low quality. Practical experience suggests to keep the evaluation system as simple as possible.

To reach the goals of a promotional campaign there are n different advertising media $m_1,\ldots,m_n$, available ($n \approx 200$), e.g. TV spots for different stations, newspaper ads in various publications, flyers, catalogs, social media ads, promotional events, etc., each with different characteristics.

After choosing the parameters and operative goals of a promotional campaign the central task of the marketing manager as a decision maker consists of the selection of a subset of advertising media and the allocation of a budget b_i to each selected medium m_i, such that the defined goals are met to a high degree while the available budget B_t is not exceeded. The decision on this so-called media mix is crucial for the success of any campaign.

Unfortunately, the effect of each advertising medium on the defined goals in connection with the selected parameters of the promotional campaign are mostly impossible to be quantified. Moreover, the effects of different media can not be separated but are highly interdependent, e.g., a promotional event with a celebrity will hardly have any effect without appropriate news coverage, and an evening TV spot will be better remembered if its tune is repeated by a morning radio spot. Under these circumstances, only educated guesses and general rules of thumb gained from experience can be used by the decision maker to allocate the promotional budget.

The existing software solution MARMIND can keep track of all tasks involved with the realization of a promotional campaign including accounting, managing orders with advertisement companies, etc. In this contribution we describe an optimization system developed to give the decision maker an automatically generated suggestion for the media mix.

There are two core features of our system: (1) an estimation of the direct effect and the interdependencies between advertising media based on the evaluation of past promotional campaigns by the managers, (2) the incorporation of these values into a discrete optimization model, which is basically a Quadratic Knapsack Problem (QKP), possibly with additional constraints.

3 QUADRATIC KNAPSACK MODEL

Given the parameters and operative goals of a promotional campaign t we will derive in Sections 4 and 5 an estimation of the following three values for all advertising media. For simplicity of notation we omit the reference to the current campaign t.

1. direct effect p_i on the promotional campaign caused by selecting medium m_i.
2. joint effect q_{ij} on the promotional campaign caused by selecting both media m_i and m_j.
3. estimated budget b_i allocated to medium m_i, if it is selected in the promotional campaign.

With these estimations we can set up the following mathematical optimization model with binary variables $x_i \in \{0,1\}$ representing the selection of advertising medium m_i. The objective function consists of a convex combination of a linear (direct effect) and a quadratic (joint effect) term with a parameter $\lambda \in (0,1)$ to be chosen appropriately. As a starting value we set $\lambda = 0.5$.

$$\max \quad \lambda \sum_{i=1}^{n} p_i x_i + (1-\lambda) \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_i x_j \quad (1)$$

$$\text{s.t.} \quad \sum_{i=1}^{n} b_i x_i \leq B_t \quad (2)$$

$$x_i \in \{0,1\} \quad (3)$$

The model (1)-(3) is the well-known Quadratic Knapsack Problem (QKP), see e.g. (Kellerer et al., 2004, Chapter 12) or (Pisinger, 2007).

It may seem reasonable to restrict the number of different advertising media selected for one promotional campaign by adding a cardinality constraint

$$\sum_{i=1}^{n} x_i \leq K. \quad (4)$$

However, it will turn out that the estimation of budget allocations b_i produces values of a certain proportion w.r.t. B_t which implicitly restricts the number of chosen advertising media and thus makes (4) redundant.

Practical considerations also suggest that certain advertising media (e.g. TV spots) are more costly and require a minimum budget to make sense. Thus, we will eliminate in a preprocessing step all advertising media whose minimum budget requirement would consume most of the available budget B_t.

The final suggestion of the media mix presented to the user of the system follows directly from the solution of (1)-(3). Exactly those advertising media m_i should be used whose decision variables have value $x_i = 1$ in the solution. Allocating the final budget $\bar{b}_i$ to each selected medium m_i requires a bit more care and will be treated in Section 5.2.

4 LINEAR AND QUADRATIC EFFECT ESTIMATION

It should be pointed out that all our estimations are based on the evaluation of past promotional campaigns and are not founded on some strict stochastic model. They were developed in several rounds of interaction with practitioners and validated with real-world case data. The fact that the convex combination of several terms allows the setting of a number of weighting parameters should be seen as an advantage since it permits the adaptation of the optimization system to the special customs and practices of the particular domain the system is applied in. By no means we can expect to deliver a “plug-and-play” system ready for use in any domain for every type of company.

Let $T(i)$ be the set of all past promotional campaigns containing advertising medium m_i. The linear profit value p_i will be expressed by a convex combination of the general success attributed to medium m_i in the past and the level of goal achievement reached by similar campaigns if they included m_i, i.e.

$$p_i := \lambda_p ps_i + (1-\lambda_p) pg_i \quad (5)$$

with $\lambda_p \in (0,1)$. The first term ps_i represents the average scaled success of all past promotional campaigns containing medium m_i. The underlying argument says that every medium contributed in some way to the overall success of past campaigns. Formally, we have:

$$ps_i := \frac{1}{|T(i)|} \sum_{t \in T(i)} \frac{s_t}{S} \quad (6)$$

Clearly, ps_i is in $[0,1]$.

The second term pg_i considers achievement of operative goals and similarity of parameters in more detail and will be described in the following subsection.

4.1 Considering Similarity of Campaigns

The value pg_i should reflect the principle that it is a good idea to repeat strategies that worked well in the past for campaigns with similar parameters. To formalize this principle we will express "working well" by the degree of goal achievement and "similar parameters" by introducing a similarity measure between campaigns.

Let $\tilde{T}(j)$ be the set of all past promotional campaigns containing operative goal g_j. Then the overall goal achievement a_t of a promotional campaign t will be defined as follows:

$$a_t := \frac{1}{\sum_{j \in G_t} r_t(g_j)} \left(\sum_{j \in G_t} r_t(g_j) \cdot \left(\frac{1}{2} \left(a_t(g_j) - \frac{1}{|\tilde{T}(j)|} \sum_{\tau \in \tilde{T}(j)} a_\tau(g_j) \right) + \frac{1}{2} \right) \right) \tag{7}$$

The term in the inner capital brackets computes the difference of the goal achievement for goal g_j from the average goal achievement over all promotional campaigns τ containing goal g_j. This number lies in $(-1,1)$ and is transformed to lie in $(0,1)$. Finally, the terms are weighted by their rank number and scaled by the sum of rank numbers.

Now we introduce a measure to express the similarity between two promotional campaigns t and t'. Formally, we will define a function $sim(t,t') \to [0,1]$, such that higher values of sim indicate closer similarity of two campaigns. Measures of distance and similarity are used in many fields of applied mathematics and statistics, in particular in cluster analysis (see e.g. (Everitt et al., 2011), (Guldemir and Sengur, 2006)). Our similarity function will deal separately with a linear combination of nominal and cardinal parameters of campaigns expressed by *sim_par* and with the similarity of the ordinally ranked operative goals *sim_goal*.

$$sim\_par(t,t') := \frac{1}{\sum_{i=1}^{k} c_i} \left(\sum_{i=1}^{k'} c_i \cdot sim\_nom(t(i),t'(i)) + \sum_{i=k'+1}^{k} c_i \cdot sim\_card(t(i),t'(i)) \right) \tag{8}$$

The weighting parameters $c_i \in (0,1)$ can be used to indicate the importance of different parameters.

Comparing nominal parameters is done simply by an inverted Hamming distance, i.e. assigning $sim\_nom(t(i),t'(i)) = 1$ if $t(i) = t'(i)$ and 0 otherwise, for $i = 1,\ldots,k'$. Clearly, also more complicated measures such as the Jaccard index, the Sørensen coefficient or the Tanimoto distance might be used, see e.g. (Tan et al., 2006).

For cardinal parameters $i = k'+1,\ldots,k$ the similarity is computed from the relative deviation by

$$sim\_card(t(i),t'(i)) = 1 - \frac{|t(i) - t'(i)|}{\max\{t(i),t'(i)\}}, \tag{9}$$

which is clearly in $[0,1]$. Basically, any Minkowski metric could be used and scaled into the corresponding similarity measure.

For comparing the ordered selection of goals between two campaigns in a similarity measure $sim\_goal(t,t')$, classical distance measures of orderings such as Kendall tau rank distance (similar to Kemeny distance) could be used (see (Sculley, 2007) and (Kumar and Vassilvitskii, 2010) for recent contributions). In our case, out of the available set of ℓ goals each campaign is assigned only subset of goals of small, but varying size. Hence, we use the following rather unorthodox approach.

Define a decreasing sequence of positive bonus points $\beta_1 > \beta_2 > \ldots > \beta_L$ and translate rank numbers into bonus points by assigning the goal g of a promotional campaign t with rank $r_t(g)$ exactly $\beta_{\ell_t - r_t(g)+1}$ points, i.e. the best ranked goal receives β_1 points and the lowest ranked goal with $r_t(g) = 1$ gets β_{ℓ_t} points. The remaining points $\beta_{\ell_t+1},\ldots,\beta_L$ are not assigned at all.

For any pair (t,t') of campaigns we determine the intersection of selected goals and add the bonus points accrued by every such goal in both campaigns. I.e. if some goal g' is ranked on first position in t and on third position in t', then g' contributes $\beta_1 + \beta_3$ to the total sum, while goals appearing in only one of the two campaigns do not contribute at all. This sum is scaled by the maximum possible number of points $\sum_{j=1}^{\min\{\ell_t,\ell_{t'}\}} 2\beta_j$ which guarantees a final value $sim\_goal(t,t')$ in $[0,1]$, with the desired property that identical orderings of goals yield a similarity of 1 while disjunctive sets of goals have similarity 0.

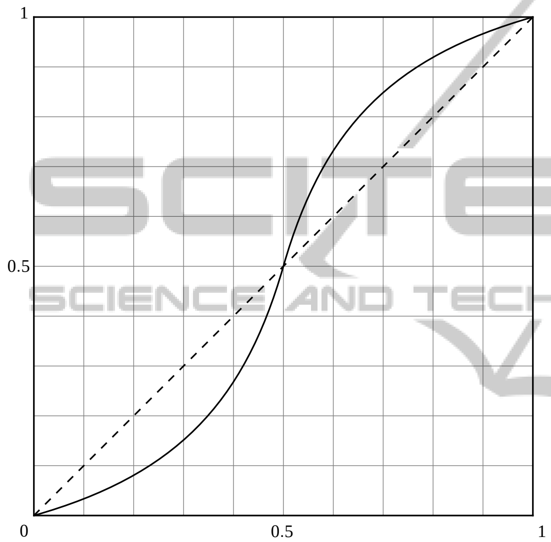
Finally, we put together the two similarity measures with a weighting parameter λ_g.

$$sim(t,t') := (1-\lambda_g)\, sim\_par(t,t') + \lambda_g \cdot sim\_goal(t,t') \tag{10}$$

A drawback of the above definitions can be found in the "averaging effect" which means that taking a linear combination over many different factors may dilute the effect of strong similarity or deviance in some components and tends to produce moderate values for almost any pair of promotional campaigns.

Thus, we aim at strengthening the influence of strong or weak similarities by increasing values closer to 1 and decreasing values closer to 0. This will be done by applying the following *sigmoid function* $F(x)$ on every partial similarity measure $sim\_nom(t(i),t'(i))$, $sim\_card(t(i),t'(i))$ and $sim\_goal(t,t')$. $F(x)$ is depicted in the following figure. It contains a tuning parameter k which we set to $k = 10$ in our implementation.

$$F(x) = \frac{1}{1+e^{(\frac{k}{2}-kx)}} + \frac{1}{1+e^{\frac{k}{2}}} \cdot (2x-1) \qquad (11)$$



It remains to put together the expressions of goal achievement and similarity. This is done by simply summing up achievement values of past campaigns weighted by their similarity to the current campaign t^c. Formally, we have

$$pg_i := \frac{1}{|T(i)|} \sum_{t \in T(i)} sim(t,t^c) \cdot a_t \qquad (12)$$

Again, pg_i is in $[0,1]$.

4.2 Estimation of Media Interaction

We proceed to estimate the effect of having two advertising media m_i and m_j together in a promotional campaign. This is done by separating from the set of all past campaigns a subset of particularly effective campaigns which stood out among the remaining campaigns. Then we will simply count the occurrence of every pair of advertising media in the effective campaigns relative to all its occurrences. Thereby, we aim to detect a systematic effect of successful pairs that happened to be chosen together in conspicuous frequency among the more effective campaigns. Note that our existing sample of campaigns is too small to allow statistical tests on this hypothesis.

Formally, we sort the set of past promotional campaigns in decreasing order of their goal achievement a_t and determine a threshold a^T such that only a prescribed percentage of campaigns exceeds this achievement value, e.g. 25%. Then we set:

$$q_{ij} := \frac{|T(i) \cap T(j) \text{ with } a_t \geq a^T|}{|T(i) \cap T(j)|} \qquad (13)$$

It turned out that there are certain pairs of media that marketing managers generally want to use together and which appear in pairs in almost all campaigns (if they appear at all), no matter whether the campaigns worked well or not. This effect is not captured by (13) which was hence extended to include the presence of pairs of media in past campaigns with strong similarities to the current campaign t^c. Let $T^c := \{t \mid sim(t,t^c) \geq \delta\}$ for some similarity threshold δ. Then we define the final quadratic effect as:

$$q'_{ij} := \lambda_q \, q_{ij} + (1-\lambda_q) \cdot \frac{|(T(i) \cap T(j)) \cap T^c|}{|T^c|} \qquad (14)$$

5 BUDGET ALLOCATION

5.1 Estimation of Budget Values

While it may seem quite reasonable that one can learn from past promotional campaigns which advertising media, resp. which combination of media, worked well to reach certain goals for campaigns with a certain set of parameters, it is less clear how to assign a budget value to an advertising medium after deciding to use it. However, one can not separate media selection from budget allocation since one may end up with a collection of advertising media that can not be realized within the given budget B_t considering the natural lower bounds on the budget for each medium.

To allow a plausible estimation of the budget values b_i in the optimization model, we consider a subset of past campaigns T^B with a budget in similar range as the current campaign t^c, i.e.

$$T^B := \{t \mid k_1 B_{t^c} \leq B_t \leq k_2 B_{t^c}\} \qquad (15)$$

with suitably chosen parameters $k_1 < 1$, $k_2 > 1$. Then we determine for each advertising medium the relative proportion of budget allocated in the past (depending on its assigned budget b_i^t) and take the mean over these values as an estimation of b_i. Formally,

$$b_i := \frac{B_{t^c}}{|T(i) \cap T^B|} \sum_{t \in T(i) \cap T^B} \frac{b_i^t}{B_t}. \qquad (16)$$

Note that different from Section 4 we do not take similarity of campaigns into account in this estimation. Discussions with marketing managers and analysis of available data exhibit that the choice of advertising media is very much tailored to the particular goals and parameters of a campaign. But once a medium is selected the invested budget is mostly dependent on technical constraints and the "size", i.e. budget, of the overall campaign. But clearly, it would be straightforward to restrict the summation in (16) to campaigns in T^c with a certain similarity to t^c.

5.2 Actual Budget Allocation

After solving the optimization model (1) - (3) we obtain a solution set $S := \{i \mid x_i = 1\}$ of all selected advertising media. Assigning the actual budget values b_i to all media $m_i \in S$ could be done by simply resorting to the estimations b_i from (16).

We suggest a more refined procedure taking into account two aspects: First, the discrete solution of optimization model will most likely leave a certain amount of budget $B_t - \sum_{i\in S} b_i$ unused and thus miss chances for a better utilization of the available budget. Secondly, and more important, it should make sense to consider the particular combination of media in S, which we already targeted specifically by the quadratic coefficients q_{ij}.

To do so, we give the relative budget proportions in a promotional campaign t, i.e. $\frac{b_i^t}{B_t}$, more weight if t shares more advertising media with the solution for the current campaign t^c. This is achieved by the following formula for every medium m_i, $i \in S$:

$$\bar{b}_i := \frac{B_{t^c}}{|S|-1}\sum_{j\in S, j\neq i}\frac{1}{|T(i)\cap T(j)\cap T^B|}\sum_{t\in T(i)\cap T(j)\cap T^B}\frac{b_i^t}{B_t} \tag{17}$$

Allocating budgets according to (17) may result in infeasible solutions or (as before) in leftover budget. We propose the following allocation process to overcome this issue.

The budget estimation b_i in (16) can be seen as an estimator in the strict statistical sense. Hence, we can also compute the associated empirical standard deviation σ_i based on the sum of squared distances from the mean and defined as follows:

$$\sigma_i := \sqrt{\frac{B_{t^c}}{|T(i)\cap T^B|-1}\sum_{t\in T(i)\cap T^B}\left(\frac{b_i^t}{B_t}-\frac{b_i}{B_{t^c}}\right)^2} \tag{18}$$

Now we start the budget allocation procedure by assigning each advertising medium $m_i \in S$ in decreasing order of profit values p_i a conservative budget value of $b_i - \sigma_i$, i.e. the estimated value reduced by one standard deviation. Then we enter into a second round and increase the budget to b_i as long as the budget B_t permits, again in decreasing order of p_i. Finally, if there is still budget left, we take a third round and increase the allocated budget to $b_i + \sigma_i$ until B_t is completely used up. Clearly, the last advertising medium considered by this procedure may obtain a budget allocation in between the three prescribed values by consuming all the remaining budget.

An analogous procedure is done for the more sophisticated budget values $\bar{b}_i$ (with the corresponding empirical standard deviation $\bar{\sigma}_i$) where it can be expected to be more relevant, since there is a larger difference from the budget values used in the optimization model. Note that in this case it may happen that we run out of budget already in the first round of allocations, since the values b_i used in the weight constraint of the optimization model may deviate considerably from $\bar{b}_i$.

6 SOLUTION OF THE QUADRATIC KNAPSACK PROBLEM

The model introduced in Section 3 is a standard Quadratic Knapsack Problem (QKP) with no additional side-constraints. This is somewhat rare, since practical applications usually require additional constraints and do not fit into the mould of standard models.

Important exact solution methods for QKP were given by (Caprara et al., 1999) and (Billionnet and Soutif, 2004). The former approach uses Lagrangian relaxation and is able to solve instances containing up to 200 variables. It is especially well suited for dense instances. (Billionnet and Soutif, 2004) uses Lagrangian decomposition and is able to solve instances of roughly the same size, however it outperforms the previous approach on instances of medium and low density.

The currently best working strategy was given by (Pisinger et al., 2007). It succeeds in reducing the size of many instances dramatically by fixing items that will or will not occur in an optimal solution. The reduced problem can then be solved by any algorithm for QKP. Combining this approach with an exact solution algorithm (Pisinger et al., 2007) were able to solve instances with up to 1500 items. Unfortunately, this code is not available, therefore we used the implementation described in (Caprara et al., 1999) for solving benchmark problems of MARMIND and managed to solve instances to optimality with up to $n = 200$ advertising media in less than 10 minutes on a simple standard PC with 2.2 GHz and 2 GB Ram.

For ensuring a good user experience *UPPER Network* however requested that the optimized marketing campaign of MARMIND has to be computed in less than 3 seconds. Moreover, we recall that all data of our QKP instances is based on estimates and does not represent assured values. Thus, we can easily settle for a good approximate solution.

For our optimization tool we implemented a genetic algorithm and imposed a time limit of 3 seconds. It turned out that this gave solutions for all instances of the required size (≥ 200 items) with an average deviation of less than 1% from optimality.

Our algorithm is a modified version of (Julstrom, 2005) which worked well for the random test instances generated according to the same method used in (Caprara et al., 1999). (Julstrom, 2005) reports test data for ten instances of 100 items and ten instances of 200 items. Every instance was solved 50 times and the algorithm was able to find the optimal solution value in about 90 percent of the runs, although the running time sometimes exceeds 1 minute. Note that our implementation was especially tuned for getting high quality results in a very short time but often succeeded to yield results similar or better than (Julstrom, 2005).

Recently (Yang et al., 2013) published a well performing metaheuristic that combined GRASP with tabu search. On 100 randomly generated benchmark instances that follow the same scheme as in (Caprara et al., 1999) the metaheuristic was able to find the optimal solution 99 times in less than 0.8 seconds. In the remaining case the gap to the optimal solution was negligibly small. Moreover, they were able to get good solutions for instances of up to 2000 variables (the solution quality was justified by comparison to known upper bounds) in less than 300 seconds.

Currently, we are working on a project to systematically test our genetic algorithm, compare it to the other existing methods listed above and to introduce harder benchmark instances for QKP. The results of this comprehensive computational study will be published as they become available.

7 CONCLUSIONS

We developed an optimization system to offer marketing managers an evidence-based suggestion for the media mix to be used for a given promotional campaign. It relies on a comparison of the current campaign to past campaigns based on their parameters and goals.

Building an optimization model with the computed direct and pairwise effect estimations gives rise to a Quadratic Knapsack Problem which can be solved almost to optimality in all real-world scenarios within a time limit of 3 seconds. The optimization tool is currently used within the industrial software solution MARMIND.

Future developments include a revision of some of the effect estimations by stochastic models as soon as a suitable set of test data derived from real world applications is available. Furthermore, the estimations will be adjusted to include a “memory” effect, i.e., giving a smaller weight to campaigns in the more distant past. It may also be interesting to take trends into accounts. Based on classical tools of statistical analysis it should be possible to detect certain trends of advertising media increasing or decreasing in importance, or in their effect for certain goals or target groups.

ACKNOWLEDGEMENTS

This research was supported by the Austrian Research Promotion Agency (FFG) under project “MARMIND media mix optimization“ and by the Austrian Science Fund (FWF): [P 23829-N13].

REFERENCES

Balachandran, V. and Gensch, D. (1974). Solving the “marketing mix” problem using geometric programming. *Management Science*, 21:160–171.

Billionnet, A. and Soutif, É. (2004). An exact method based on lagrangian decomposition for the 0–1 quadratic knapsack problem. *European Journal of operational research*, 157(3):565–575.

Caprara, A., Pisinger, D., and Toth, P. (1999). Exact solution of the quadratic knapsack problem. *INFORMS Journal on Computing*, 11:125–137.

Everitt, B., Landau, S., Leese, M., and Stahl, D. (2011). *Cluster Analysis*. Wiley, 5th edition.

Färe, R., Grosskopf, S., Seldon, B., and Tremblay, V. (2004). Advertising efficiency and the choice of media mix: a case of beer. *International Journal of Industrial Organization*, 22:503–522.

Guldemir, H. and Sengur, A. (2006). Comparison of clustering algorithms for analog modulation classification. *Expert Systems with Applications*, 30:642–649.

Julstrom, B. (2005). Greedy, genetic, and greedy genetic algorithms for the quadratic knapsack problem. In *GECCO ’05: Proceedings of the 2005 conference on*

Genetic and evolutionary computation, pages 607–614. ACM.

Kellerer, H., Pferschy, U., and Pisinger, D. (2004). *Knapsack Problems*. Springer.

Kumar, R. and Vassilvitskii, S. (2010). Generalized distances between rankings. In *Proceedings of the 19th International World Wide Web Conference*, pages 571–580. ACM.

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

Pergelova, A., Prior, D., and Rialp, J. (2010). Assessing advertising efficiency: Does the internet play a role? *Journal of Advertising*, 39:39–54.

Pisinger, D. (2007). The quadratic knapsack problem - a survey. *Discrete Applied Mathematics*, 155:623–648.

Pisinger, D., Rasmussen, A., and Sandvik, R. (2007). Solution of large quadratic knapsack problems through aggressive reduction. *INFORMS Journal on Computing*, 19:280–290.

Reynar, A., Phillips, J., and Heumann, S. (2010). New technologies drive CPG media mix optimization. *Journal of Advertising Research*, 50:416–427.

Sculley, D. (2007). Rank aggregation for similar items. In *Proceedings of the 7th SIAM International Conference on Data Mining*, pages 587–592. SIAM.

Sönke, A. (2012). Optimizable and implementable aggregate response modeling for marketing decision support. *International Journal of Research in Marketing*, 29:111–122.

Sorato, A. and Viscolani, B. (2011). Using several advertising media in a homogeneous market. *Optimization Letters*, 5:557–573.

Tan, P.-N., Steinbach, M., and Kumar, V. (2006). *Introduction to Data Mining*. Addison-Wesley.

Vakratsas, D. and Ambler, T. (1999). How advertising works: what do we really know? *Journal of Marketing*, 63:26–43.

Yang, Z., Wang, G., and Chu, F. (2013). An effective GRASP and tabu search for the 0-1 quadratic knapsack problem. *Computers and Operations Research*, 40:1176–1185.