European Journal of Operational Research 226 (2013) 115–121

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Optimal resource allocation in survey designs

Melania Calinescu [a,*], Sandjai Bhulai [a], Barry Schouten [b]

[a] VU University Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
[b] Statistics Netherlands, Henri Faasdreef 312, 2492 JP The Hague, The Netherlands

ARTICLE INFO

Article history:
Received 3 April 2012
Accepted 29 October 2012
Available online 14 November 2012

Keywords:
Resource allocation
Nonlinear optimization
Dynamic programming
Markov decision theory
Survey design
Scheduling

ABSTRACT

Resource allocation is a relatively new research area in survey designs and has not been fully addressed in the literature. Recently, the declining participation rates and increasing survey costs have steered research interests towards resource planning. Survey organizations across the world are considering the development of new mathematical models in order to improve the quality of survey results while taking into account optimal resource planning. In this paper, we address the problem of resource allocation in survey designs and we discuss its impact on the quality of the survey results. We propose a novel method in which the optimal allocation of survey resources is determined such that the quality of survey results, i.e., the survey response rate, is maximized. We demonstrate the effectiveness of our method by extensive numerical experiments.

1. Introduction

Surveys are used all around the world to measure socio-economic status and well-being of people, to test theories, or to make investment decisions, driven by the impossibility of observing the entire population of interest (see [6]). No matter what the framework of a survey is, its success relies on the active participation of the sampled households and businesses. Nonresponse occurs when members of a sample cannot or will not participate in the survey. The impact of nonresponse appears in the inability of computing a full-sample estimator of the population mean. Thus, a bundle of practical issues is created, including bias in point estimates, bias in estimators of precision, and inflation of the variance of point estimators. The error caused by nonresponse is one of the several sources of error in surveys and it has attracted a great deal of interest among researchers across the world (see [6]). An apparent solution to the problem is to increase the frequency of attempts to gather information from reluctant sample members. Under these circumstances, the costs of conducting surveys increase significantly, which leads to new problems, such as budget overruns. Therefore, a constant scientific challenge to the survey community concerns developing new survey designs to accommodate the presence of both nonresponse and high costs.

Modeling the bundle of processes behind a survey and understanding the numerous interactions between these processes have been a constant obstacle for researchers in their attempts to design quality but cost-effective surveys. As a consequence, only few processes have been investigated from a cost perspective, e.g., call scheduling in [9]. More advanced studies investigate the relationship between costs, quality and few survey features (e.g., the interview mode, the schedule of calls). For example, in [7,11], the main idea is to identify a set of design features that potentially influence the survey costs and errors in the estimates and to monitor them throughout the survey run. This information helps in subsequent phases to alter the design features such that a desired balance between costs and errors is achieved.

When person or household characteristics (e.g., social and financial) are employed to adjust the design features to a given set of characteristics (i.e., different design features can be applied to sample units with different characteristics) the resulting survey design is termed *adaptive* (introduced in [16,14]). Adaptive designs render realistic survey models and can be used to capture the interactions between survey features, sample unit characteristics, survey costs and quality.

In the present paper, adaptive survey designs are analyzed from the perspective of resource allocation problems. To our knowledge, this is the first paper that addresses designing surveys from a resource allocation perspective. Given a budget, a set of household characteristics, and a list of survey features that influence costs and quality, we model the allocation of survey resources such that quality is maximized while costs meet the budget constraint. Our interest in the problem is motivated by the increased difficulty (e.g., higher costs, and higher nonresponse) survey organizations are faced with in order to obtain high-quality survey estimates. Statistics Netherlands is among the first organizations to consider

* Corresponding author. Address: VU University Amsterdam, Faculty of Sciences, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands. Tel.: +31 205987687.
E-mail address: melania.calinescu@vu.nl (M. Calinescu).

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2012.10.046

redesigning their surveys such that planning of resources is taken into account.

Resource allocation problems can be found in a wide range of applications. In [12], the author investigates applications where the resource allocation can be modeled as a continuous convex nonlinear problem. Algorithms to solve such problems are also surveyed and they most often involve finding the optimal value of the Lagrange multiplier for the explicit constraint (mainly through some type of line search). There are also numerous applications that require a relaxation of the condition on strict convexity and differentiability of the cost constraints, which increases significantly the complexity of the problem. The auction algorithm presented in [1] finds a near optimal solution of this problem in finite time.

In its integer or mixed-integer formulation, the resource allocation problem has an NP-complete worst case complexity (see [8]). Therefore, only few such applications have been addressed in the literature, e.g., optimal sample allocation in stratified sampling (see [10,3]), manufacturing capacity planning (see [4]). The proposed algorithms take advantage of the convexity in the objective function and/or constraints. Applications where the objective function and/or the constraints include separable nonconvex functions are often encountered (e.g., due to economies of scale). In this case, additional difficulties in solving the problem are posed by the presence of several local optima. In [5], an approach is suggested to solve such problems, namely solve a convex lower bounding problem (e.g., the convex envelope) at every node of the branch-and-bound search tree. Using the branch-and-bound framework developed by [2], the optimal solution is reached in a finite number of iterations. However, no implementation results or optimality gap assessments are reported.

The resource allocation problem for survey designs has specific features that lead to a formulation as a nonconvex integer nonlinear problem, which prohibits the application of many algorithms that are found in the literature. A possible approach could be to implement solutions of convex approximations of the problem, however, this may result in major errors in survey estimates. We present an algorithm that solves the problem to optimality using Markov decision theory. The algorithm reaches the optimal solution in a finite number of iterations. The numerical experiments discussed here displayed short computational times on an Intel Xeon L5520 processor.

The remainder of the paper is structured as follows. Section 2 discusses the mathematical model and Section 3 discusses the algorithm to derive optimal adaptive survey design policies. Section 4 presents a range of practical problems that can be handled through this model and solution method. Numerical examples of these situations are given in Section 5. Section 6 concludes the results of the paper and gives directions for future research.

2. Problem formulation

Consider a survey sample consisting of N units that can be clustered into homogeneous groups based on characteristics, such as age, gender, and ethnicity (information that can be extracted from external sources of data). Let $\mathcal{G} = \{1, \ldots, G\}$ be the set of homogeneous groups with size N_g for group $g \in \mathcal{G}$ in the survey sample. The survey fieldwork is divided into time slots, denoted by the set $\mathcal{T} = \{1, \ldots, T\}$, at which units in a group can be approached for a survey. The survey itself can be conducted using certain interview modes, such as a face-to-face, phone, web/paper survey; the set of different modes is denoted by $\mathcal{M} = \{1, \ldots, M\}$. At each time slot $t \in \mathcal{T}$ one can decide to approach units in group $g \in \mathcal{G}$ for a survey using mode $m \in \mathcal{M}$. In doing so, successful participation in the survey depends on first establishing contact, and then be responsive by answering the questionnaire. From historical data group-dependent contact probabilities $p_g(t,m)$ and participation probabilities $r_g(t,m)$ can be estimated, which we consider as given quantities in our problem. Note that from historical data it can also be observed that certain time slots (e.g., morning, evening) have an influence on the availability of the unit and the willingness to respond. Therefore, to employ most of the available information, the contact and response probabilities are modeled at the level of time slots for each group as well rather than the mode only.

Denote by $x_g(t,m)$ a binary 0–1 decision variable that denotes if units in group g are approached for a survey at time t using mode m. Note that at time t only one mode can be employed to approach a group, yielding the constraint $\sum_{m\in\mathcal{M}} x_g(t,m) \leqslant 1$. When a successful contact is established and the unit agrees to participate, the survey ends with success; this happens with probability $p_g(t,m)\, r_g(t,m)$. Note that we assume independence between participation and contact. However, if the unit refuses participation after successful contact, the unit is not considered for a future survey approach; this happens with probability $p_g(t,m)(1 - r_g(t,m))$. Only in the case that the unit is not contacted successfully, the unit can be considered for a future survey approach (see Fig. 1); this happens with probability $1 - p_g(t,m)$. Thus, if the unit is approached again at time t' using mode m', then the probability of a successful approach is $(1 - p_g(t,m))p_g(t',m')r_g(t',m')$, and the probability of a contact failure is $(1 - p_g(t,m))(1 - p_g(t',m'))$. In general, the probability that a contact fails up to time t' is denoted by $f_g(t')$ given by

$$f_g(t') = \prod_{t=1}^{t'} \prod_{m\in\mathcal{M}} [x_g(t,m)(1 - p_g(t,m)) + (1 - x_g(t,m))]$$
$$= \prod_{t=1}^{t'} \prod_{m\in\mathcal{M}} [1 - x_g(t,m)p_g(t,m)].$$

Note that this is a highly non-linear expression in the decision variables, which can be recursively computed by

$$f_g(t') = \prod_{m\in\mathcal{M}} [x_g(t',m)(1 - p_g(t',m)) + (1 - x_g(t',m))] f_g(t' - 1)$$
$$= \prod_{m\in\mathcal{M}} [1 - x_g(t',m)p_g(t',m)] f_g(t' - 1), \quad (1)$$

using the fact that $f_g(0) = 1$. Using this definition, the response rate for group g can then be computed by

$$\sum_{t\in\mathcal{T}} \sum_{m\in\mathcal{M}} f_g(t-1)\; x_g(t,m)p_g(t,m)r_g(t,m).$$

The clustering of the N units usually results in groups that are not of the same size or importance. Therefore, the response rates for the groups are usually weighted by a factor w_g (e.g., $w_g = N_g/N$ is taken in practice). Hence, the objective of the decision maker becomes to maximize

$$\sum_{g\in\mathcal{G}} \sum_{t\in\mathcal{T}} \sum_{m\in\mathcal{M}} w_g f_g(t-1) x_g(t,m) p_g(t,m) r_g(t,m), \quad (2)$$

by setting the decision variables $x_g(t,m)$ optimally. The decision variables are subject to constraints, though, due to scarcity in resources. In practice, due to resource management constraints, the number of times that a group can be approached by mode m is limited to $\bar{k}_g(m)$ times, leading to the constraint $\sum_{t\in\mathcal{T}} x_g(t,m) \leqslant \bar{k}_g(m)$. By combining the objectives with all the constraints, we can draft our optimization problem as a binary programming problem in the following manner.