An Improved Call-scheduling Method Using Call History and Frame Information

L. Reedman and M. Robinson, Statistics Canada
Laurie Reedman, Statistics Canada, RHC 11-K, Tunney's Pasture, Ottawa, ON, Canada, K1A 0T6

Abstract
An important feature of Computer Assisted Telephone Interviewing is that it facilitates efficient call scheduling which can help reduce costs and non-contacts. Using call history data from a recent RDD survey at Statistics Canada, improved methods for scheduling calls to unreached telephone numbers in the sample are developed. In addition the logistic regression model used takes account of the likely business/residence status of each sampled telephone number. In this paper the model is explained, the fit of the model is evaluated, and implementation issues are described.

Keywords: call scheduling; call history; non-contact; logistic regression; random digit dialling

1. Introduction

At Statistics Canada an increasing number of surveys are conducted using Computer Assisted Telephone Interviewing (CATI). The benefits of CATI in data processing are well known; CATI also has the potential to manage interviewer case loads under a complex set of conditions. For instance, case management software can be developed with a call-scheduling module to form queues of cases sharing certain characteristics (such as language, tracing needs, refusal conversion, or new cases), present cases to the appropriate interviewer when an appointment has been made, and record critical information about each call attempt, such as the time, interviewer identification, telephone number dialled, length of call, and outcome, to name a few. In this paper, a method of assigning priorities to cases in a queue is explored. The scope of this paper is further restricted to household surveys, with emphasis on random digit dialling (RDD) surveys. Section 2 of this paper describes the methodology. In Section 3 implementation options and some considerations in adapting this method to a non-RDD setting are discussed. Conclusions follow in Section 4.

2. Methodology

The goal of the scheduling module in a CATI application is to make efficient and effective use of interviewer resources, so as to reduce costs, reduce non-productive telephone calls, and to reduce non-response bias. Scheduling software exists which arranges cases into queues depending on the kind of attention needed next, such as tracing, refusal conversion, or an appointment. Currently at Statistics Canada the scheduling software used has the ability to form various queues. Within the queue for new cases and cases requiring a "cold" call-back, the scheduler gives higher priority to cases which have already been attempted than to cases which have never been called. When data collection is done from the head office, interviewers attempt to make telephone calls to the same number at different times of the day and on different days of the week. The scheduler in centralized collection also has the ability to shift the priority of cases from east to west, depending on the time zone. In de-centralized data collection, the interviewer caseload is generally small enough that each interviewer can schedule his or her cases efficiently. This research focuses on a single queue for cases which have not yet been attempted (first calls) and cases which have been attempted but no contact has yet been made. A logistic regression model is used to predict the best time to attempt a call, based on call history and frame information.

2.1 Sample Design

RDD surveys at Statistics Canada follow a list assisted design, similar to that described by Casady and Lepkowski (1993). The "list" or frame at Statistics Canada is primarily a composite of telephone billing files from all the major telephone companies in Canada. As described in Ciok (1994), the frame is updated on average four times each year, from updated telephone billing files. It is organized into banks of 100 numbers having the first 8 digits in common. This part of the frame contains a flag derived from information on the billing files, indicating the status of every telephone number. The status indicator is represented on the frame as a "B" if the number is billed as a business or an "R" if the number is billed as a residence. If the number is not in use, or is unlisted and was consequently removed from the billing file before being sent to Statistics Canada, or the number was newly listed since the time the previous billing file was sent to Statistics Canada, then the billing status is essentially unknown at Statistics Canada. In these cases, the status indicator is left blank. Banks with valid first 8 digits of telephone numbers but no working numbers are termed zero-listed banks, and are omitted from the frame. Similarly, entire banks which are reserved for business use only (such as for large companies or government agencies) are omitted from the frame. There are remote areas of Canada for which billing files are not available; here, all possible 100-banks with valid area code and prefix combinations are included in the frame. No status indicator is available for these telephone numbers; however, as will be demonstrated in Sections 2.2 and 2.3, this indicator is a strong predictor of the best time to make a telephone call, and is especially useful in scheduling the first call.

The non-telephone population of households in Canada has been estimated at 1.0%. Undercoverage due to the omission of zero-listed banks has not been formally measured, but is believed to be small (Casady and Lepkowski 1993). It is therefore reasonable to use a truncated frame in Canada. For more description of list assisted methods in general see Casady and Lepkowski (1993), and for more detail of RDD sampling at Statistics Canada see Dolson (1996).

2.2 The first call

Until recently, first calls were made essentially at random between 9:00 am and 9:00 pm on weekdays. A contact rate of approximately 32% was attained. It has been found that the contact rate improves as time approaches evening. Statistics Canada conducted an RDD survey in 1996 which gathered information about Sun Exposure. In this survey, first calls to sampled cases with a "B" indicator on the frame were made between 9:00 am and 4:00 pm; first calls to sampled cases with an "R" were made between 4:00 pm and 9:00 pm; and first calls to sampled cases with a blank indicator were made at random. This resulted in an increase in contact rate on that first call of approximately five percentage points. Of the sampled cases with a "B" indicator 68% were indeed businesses, while 8% were private households, 12% were non-working, and 12% were unresolved by the end of the collection period. Of the sampled cases with an "R" indicator 86% were private residences, while 3% were businesses, 9% were non-working, and 2% were unresolved by the end of the collection period. Of the sampled cases with a blank status indicator, 61% were ultimately classified as non-working; however, 24% of the sampled cases were private residences, 11% were businesses, and 4% were unresolved.

These findings suggest that it is more efficient to schedule first calls based on the status indicator on the frame than to make them at random.

It should be noted that some telephone call attempts at Statistics Canada are made on Sunday; however, first calls on Sunday in the Sun Exposure survey were not included in this analysis because they were too few to support reliable conclusions.

2.3 Subsequent calls

A logistic regression model has been suggested by Stokes and Greenberg (1990) for modelling the probability of making contact on call history information. This was done using data from another recent Statistics Canada RDD survey, in which 7,995 telephone numbers were sampled and 25,419 call attempts were recorded using CATI. As in the Sun Exposure survey, some calls in this survey were made on Sunday; however, due to their small number, all cases which included a call attempt made on Sunday were discarded from the analysis. The dependent variable in the regression was the probability of making contact; therefore, all calls after the initial contact were discarded from the analysis. A different model would result if the outcome to be predicted was the probability of contact with a residence, or a completed interview, or a refusal; these alternatives will be discussed in Section 3.3. All calls to numbers which were eventually classified as non-working were also discarded from the analysis. Remaining were 10,411 call attempts to be modelled.

Explanatory variables as suggested by Stokes and Greenberg (1990) were available on the call history file. Also included was the status indicator from the frame. Each of the variables tested is explained below.

i. Billing file indicator

This variable is the status indicator from the frame. For the analysis, the variable BILLIND has four levels: business, residential, unknown status but the 100-bank is known to exist, and unknown status with no knowledge of the bank's existence (for those with no billing file information).

ii. Number of previous attempts

This variable counts the number of unsuccessful call attempts to a telephone number. It was found that the likelihood of making contact decreases as the number of attempts increases. For the analysis, the variable COUNT has 5 levels: 1 previous attempt, 2 previous attempts, and so on up to 5 or more previous attempts.

iii. Amount of time since last call

This variable represents the amount of time since the last call attempt to a telephone number. It was found that redialling (calling again within two hours) is recommended following a busy signal, but otherwise it is most advantageous to wait two or more days before calling again. The variable LAGDAY has four levels: redial (within two hours), same day call back (after more than two hours), next day call back, and 2 or more days wait to call back.

iv. Timing of previous calls

Stokes and Greenberg (1990) suggest that if a call has previously been made with no success during the daytime, then it may be favourable to try next in the evening, and vice versa. The variable representing this characteristic is called PAST, and has three levels: number previously attempted during the day, number previously attempted during the evening, and number previously attempted during both time periods.

v. Previous outcome

As demonstrated by Stokes and Greenberg (1990), the data here also showed that the outcome of an unsuccessful call to a number is significant in predicting the probability of contact on the next call attempt. Initially, a variable which distinguished between a busy signal, an unanswered ring and any other type of non-contact (fax machine, answering machine, or "dead air") was tested. However, it was found that the model performed equally well when the unanswered rings were combined with other types of non-contact. In the final model, the variable PREVBUSY has two levels: previously busy, and previously another non-contact.

vi. Local time of current call

The variable for local time of the number being called is intuitively significant; the data confirmed this, with the late evening time period having the highest likelihood of contact. For the analysis this variable is called LTIME, and has three levels: 9:00 am to 4:00 pm weekdays, 4:00 pm to 7:00 pm weekdays, and 7:00 pm to 9:00 pm weekdays.

Each of the variables individually proved to be significant at the 1% level. Two-way interactions were also investigated. Table 1 below shows the variables and two-way interactions in the final model. For a full description of model creation and variable refinement, see Robinson (1996).

Table 1: Variables in the model.

Variable	**Wald Chi squared**	**degrees of freedom**
BILLIND	165.2	3
COUNT	265.8	4
BILLIND x LTIME	98.52	6
BILLIND x PREVBUSY	23.36	3
BILLIND x PAST	51.89	6
LAGDAY x PREVBUSY	54.93	3
LTIME x PAST	25.10	4

For the data in this analysis, the logistic regression model provided an excellent fit. The histogram of standardized residuals from the model was found to have a mean of 0.00095 and a standard deviation of 0.99653.

The data for building this model came from one Statistics Canada RDD survey, rather than an experiment designed for the investigation of factors affecting the success of call-backs. Before proceeding with implementation steps, it would be wise to validate and possibly modify the model by applying it to different sets of data.

When comparing the model found here to that of Stokes and Greenberg (1990), we observe remarkable similarities, given that both models were derived from single sources of data. The obvious difference is the variable BILLIND which was only available on Statistics Canada's frame. Each interaction variable involving BILLIND in the model here corresponds to a variable in the Stokes and Greenberg model, with some differences in degrees of freedom.

3. Implementation options

The probability of contact can be calculated using this model for every possible combination of the explanatory variables. A look-up table can be incorporated into the scheduling software such that for any vector of characteristics in the call history of a case, the probability of contact is found.

Table 2 below shows some calculated probabilities of contact. For each of the three levels of the variable LTIME (local time of current call) only the top ten probabilities are shown.

Table 2: Probabilities of contact on subsequent calls.

Number of previous attempts	Time since last call	Previous outcome	Previously called in:	Billing file indicator	Predicted probability of contact
			Daytime		
1	redial	busy	daytime	residential	0.540
1	next day	busy	evening	business	0.525
1	2 or more days	other	evening	business	0.415
2	redial	busy	daytime	residential	0.389
2	redial	busy	both	residential	0.383
1	next day	other	evening	business	0.378
2	next day	busy	evening	business	0.374
1	next day	busy	evening	residential	0.358
1	2 or more days	other	daytime	residential	0.357
1	redial	busy	daytime	business	0.343
			Early Evening		
1	redial	busy	daytime	residential	0.724
2	redial	busy	daytime	residential	0.586
1	redial	busy	evening	residential	0.566
1	same day	other	daytime	residential	0.553
2	redial	busy	both	residential	0.525
1	2 or more days	other	daytime	residential	0.511
3	redial	busy	daytime	residential	0.498
1	next day	other	daytime	residential	0.473
1	redial	busy	evening	business	0.452
3	redial	busy	both	residential	0.436
			Late Evening		
1	same day	other	daytime	residential	0.694
1	2 or more days	other	daytime	residential	0.657
1	redial	busy	evening	residential	0.635
1	next day	other	daytime	residential	0.621
2	same day	other	daytime	residential	0.550
3	redial	busy	both	residential	0.539
2	2 or more days	other	daytime	residential	0.508
2	same day	busy	daytime	residential	0.494
2	next day	other	daytime	residential	0.470

3.1 How to integrate first and subsequent calls

In Section 2.2 it was noted that an improvement in contact rate can be gained by making first calls to business numbers during the day and making first calls to residential numbers during the evening. The Sun Exposure survey data were used to calculate probabilities of making contact on the first call for all combinations of local time of current call and the status indicator on the frame. The probabilities are shown in Table 3.

Table 3: Probabilities of contact on first call.

Local time of current call	Billing file indicator	Probability of contact
daytime	residential	0.42
daytime	business	0.43
daytime	blank	0.17
early evening	residential	0.58
early evening	business	0.36
early evening	blank	0.18
late evening	residential	0.58
late evening	business	0.23
late evening	blank	0.18

The probabilities in Tables 2 and 3 can be included in a single look-up table in a scheduling module. For the variables which have no value prior to the first call (number of previous attempts, time since last call, previous outcome and timing of previous call) a null or "not stated" value can be assigned.

3.2 Prioritizing within the queue

Greenberg and Stokes (1990) describe a Markov decision process to model an optimal calling strategy for working telephone numbers. The state of the system is described by call history information similar to the variables in our logistic regression model, with the exception of the billing file status indicator. The objective of their study was to minimize the number of telephone call attempts required to contact a household. The optimal strategy they found was that in general, there should not be two calls made to the same number in the same day, except toward the end of the collection period. Similarly, following unsuccessful calls, subsequent attempts should be made closer together as the collection period runs out.

In light of these findings, the following suggestions can be made. A batch job within the scheduling software can be automatically triggered at user defined time intervals. Early in the collection period it may not be necessary to update the probabilities more than once every hour (or even once every shift), because it is more efficient to spend the early part of collection on new cases than on call-backs. Towards the end of the collection period, however, when the optimal strategy may be to shift the emphasis from non-contact cases to ones where a busy signal has been recorded in the call history, one might want to trigger the batch job to update probabilities more frequently, allowing call-backs with a better chance of making contact to be brought to the front of the queue more frequently. A Markov decision process using data from a Statistics Canada survey where the billing file status indicator is available can be developed to further test these hypotheses.

Intuitively, cases with a higher probability should be given a higher priority in the queue. However, it may not be advantageous to begin calling the cases for a given time period in the order of their probabilities. If the time period is actually broader than the true best time window, then all the most promising cases may be "wasted" by attempting them too early. A better approach might be to randomize cases in the same queue in the same time period. Once the probabilities of contact have been updated (both for first calls and for subsequent calls) the batch job can order the cases by probability of contact, and group them into priority classes, such as deciles or quartiles. Within a priority class, cases can be randomized, then ranked and placed in the queue. This updated queue replaces the former one, and the scheduler continues to supply cases to interviewers as before.

3.3 Re-defining success

The desired outcome of the process described above was contact with a respondent. However, the real goal of survey takers is a completed interview, not just contact. Some thought was given to modelling completed interviews on the variables described in Section 2.3, but it was soon decided that these variables would not be adequate, and further that some of the most influential characteristics cannot be accurately measured at all. For instance, when predicting contact, all calls after the first contact are discarded. When predicting a completed interview, prior contact calls would need to be included, but how would an appointment be represented in the model? Or an appointment followed by a soft refusal followed by another appointment and then a completed interview? It would also be necessary to include a human factor in the equation. A respondent's willingness to complete a survey depends not only on the likelihood of answering the telephone, but also on the subject matter of the survey, and the inclination to answer questions at the time of the call.

Conversely, one could design a model whose outcome is a refusal, and avoid making calls to telephone numbers with a high probability of refusal in a particular time period. As suggested by Greenberg and Stokes (1990), certain times of the day may be better than others for making contact, but they may also have high refusal rates.

One must also consider calls to telephone numbers which ultimately reach businesses. It is desirable to resolve these cases in as few calls as possible. A polytomous model could perhaps be used, in which one outcome to be predicted is the probability of confirming a business case, and another outcome is the probability of a completed interview, or a refusal, or contact with a household.

3.4 A non-RDD survey

The sample sizes in non-RDD surveys are typically smaller than those in RDD surveys. Survey frames, in many cases, are more up-to-date than the billing files used for RDD, and contact information can be of relatively good quality. Given these features, it is generally possible to resolve every case in a non-RDD sample. Even if the final outcome of a case is "unable to contact", survey takers at Statistics Canada can expect that at least five and usually more attempts will have been made at different times of the day and different days of the week. An improvement in contact rate might be possible if a priority system similar to the one described for RDD surveys was employed. The explanatory variables would not include the status indicator from the RDD frame, but there could be other auxiliary information available on the frame, such as sex, date of birth, or household size, which may prove to be significant in predicting the best time to make contact. In a repeated or cohort survey, for example, a logistic regression model can be fit to the first cohort data, and the derived probabilities can be used to schedule calls in subsequent cohorts. A truly flexible scheduler would allow the user to specify the model, or the look-up table of call history charateristics to be used.

4. Conclusions

This paper describes a method of improving the efficiency of scheduling calls to unreached telephone numbers in a CATI sample. Contact on the first call is predicted by the likely business/residential status from the frame and the local time of the current call. Contact on subsequent calls is modelled on six characteristics: the status indicator from the frame, the number of previous attempts, the amount of time since the last call, the timing of the last call, the outcome of the previous attempt, and the local time of the current call. An implementation method is suggested, whereby the predicted probabilities are sorted by size, divided into priority classes, randomized within priority class, assigned a rank based on this random order, and placed back on the queue. This paper also suggests a number of research directions, such as using logistic regression models with different definitions of success, or polytomous models, or using a Markov model to determine a calling strategy. This research might lead to further improvements in efficiency in call scheduling.

Acknowledgement

The authors wish to thank Dave Dolson for his collaboration in the work that led to this paper, and for his helpful comments on an earlier version.

References

Casady, R. and Lepkowski, J. (1993) "Stratified Telephone Survey Designs." *Survey Methodology, 19, 103-113.*

Ciok, R. (1994) "Random Digit Dialing Telephone Frame Documentation." *Statistics Canada internal report.*

Dolson, D. (1996) "RDD Sampling Methods." *Statistics Canada internal report.*

Greenberg, B.S. and Stokes, S.L. (1990) "Developing an Optimal Call Scheduling Strategy for a Telephone Survey." *Journal of Official Statistics, 6:421-435.*

Robinson, M. (1996) "A Logistic Regression to Model Telephone Survey Response Data." *Statistics Canada internal report.*

Stokes, S.L. and Greenberg, B.S. (1990) "A priority System to Improve Callback Success in Telephone Surveys." *Proceedings of the American Statistical Association, Survey Research Methods Section, 742-747.*