

Experimentation and Self Learning in Continuous Database Marketing

James E Pearce¹
MarketEaze Solutions
jpearce@ozemail.com.au

Geoffrey I Webb
Monash University
webb@csse.monash.edu.au

Robin N Shaw
Deakin University
rshaw@deakin.edu.au

Brian Garner
Deakin University
brian@deakin.edu.au

Abstract

We present a method for continuous database marketing that identifies target customers for a number of marketing offers using predictive models. The algorithm then selects the appropriate offer for the customer. Experimental design principles are encapsulated to capture more information that will be used to monitor and refine the predictive models. The updated predictive models are then used for the next round of marketing offers.

1. Introduction

Continuous database marketing is an extension of database marketing where customers are either targeted periodically based on changes in behaviour, or made marketing offers at a touch point [1-3]. The key difference between a continuous campaign and a one-off campaign is that a continuous campaign will continue to be active for a period of time, which allows for the collection of data and ongoing modification of the campaign structure based on analysis of the results collected to date.

Our work in data mining for continuous database marketing has led to the realisation that new marketing processes are required. This paper encapsulates our thinking on the various issues that arise. The processes currently used in industry often fail to integrate data mining with other aspects of marketing campaign management. By addressing this problem and providing an experimental framework over the lifetime of the marketing campaign, our methodology allows optimised offers to be made and improves the quality of predictive models. Our particular innovations are to maintain multiple marketing models in parallel, and to institute systematic, statistically well founded processes to evaluate and refine those models.

¹ The principal author is enrolled in a research doctorate at Deakin University and acknowledges financial support from the School of Computing & Mathematics to enable him to participate in ICDM 2002.

We give a brief review of the issues surrounding continuous database marketing, and outline the usual methods employed to address them. We then examine some of the issues in more detail and propose methods of addressing them within an overall framework of a continuous database marketing campaign.

2. Background

Database marketing plays a strategic role in many organisations, particularly for business-to-consumer organisations that interact directly with their customers.

Database marketing practice is becoming increasingly complex. Companies are introducing additional channels beyond direct mail, such as outbound telemarketing, outbound email, SMS and web. Marketing campaigns are shifting from being one-off efforts to being continuous programs where customers are targeted opportunistically, based on changes in behaviour. This means that the processes such as data mining need to adapt to the continuous paradigm.

The complexity of the decision process has also increased. Rather than omitting customers unlikely to respond from a mailing list, companies now need to decide whether to make an offer, when to make an offer, which channel to use, and which offer out of many potential offers to make.

2.1. Data mining in database marketing

Data mining for database marketing has been dominated by supervised learning techniques, including logistic regression, artificial neural networks and decision-tree models [4-6]. A typical use of a model built by one of these techniques is to reduce the size of a list of customers who are to be contacted by omitting those customers unlikely to respond. This issue has been examined many times in the literature [4, 7-15].

2.2. Experimental framework

In order to test the effects on offer uptake of different factors that can be influenced by marketers, and to quantify the effects of these offers, database marketing employs an experimental framework. This framework also serves as a data collection mechanism for building new predictive models and refining existing predictive models.

Experimentation used in database marketing can vary from a very simple design to complex fractional-factorial experimental designs, where thousands of offers can be tested by varying the levels of a few factors [16, 17].

When predictive models are being used, a control cell can be formed by the inclusion of a 'random' cell. A random selection of customers are included in this cell regardless of the outcome of the predictive model usually used to determine whether an offer should be made to a customer [18]. This allows the effectiveness of predictive models to be measured.

3. Current issues in database marketing

Database marketing is shifting from a static model, where marketing campaigns were one-off exercises, to continuous marketing, where successful initiatives continue to be made over time and ongoing assessment can be used for progressive campaign refinement [2].

3.1. Continuous marketing

Continuous campaigns aim to make an appropriate offer to customers at a particular time when customers are seen to be receptive. They can take one of two forms. The first, known as 'event-triggered' marketing, is a batch process. Periodic checks of significant events in customers' attributes and behaviour are made, resulting in a relevant offer being made to all customers who match the specified criteria at regular intervals.

The second form, known as 'opportunistic marketing', occurs when an offer is made to a customer at an interactive touch point [19]. With this type of campaign the customer's visit to the touch point itself triggers the campaign; the company does not know that the customer will visit during a particular period. The decision and selection process must be performed quickly so that a dialogue can be maintained.

3.2. Decisions in continuous campaigns

To perform a continuous campaign effectively there are a number of issues that need to be addressed. They include the selection of an appropriate and significant 'event', definition of eligibility criteria that decide whether a customer is eligible to receive a particular offer, a decision on whether a particular offer should be

made to a customer, incorporation of an experimental framework, a decision on which offer should be preferred for each customer, and refinement of predictive models based on the uptake of the offer.

The criteria that govern whether a customer is eligible to receive an offer are determined by business policy; for example, a credit card customer may not be eligible for an offer of an increased credit limit if a recent statement is overdue. Such policies need to be taken into account within our framework. Different offers can be expected to have different eligibility criteria, and not all customers may be eligible for any offer.

The decision of whether a specific offer should be made to a specific customer can be aided by the use of a predictive model to select customers if the probability of uptake is greater than a threshold value [8, 9]. In this paper we assume that a model has been built for each specific offer.

4. A continuous marketing framework

Our framework covers the use of predictions and prediction intervals to determine whether a customer should receive an offer; offer selection, based on comparison of the estimated probability of response for a number of offers; the inclusion of an experimental design framework within the decision process; and a requirement to update predictive models and estimates of model accuracy with current response information.

Our innovations are to consider the outcomes of multiple predictive models when selecting an offer, to use statistical experimental design principles to augment existing models with causal data, and to evaluate and refine these models continually.

The basic procedure followed in applying the framework is outlined below. The first step is to identify those customers who are to be made an offer. The next step is to determine which offers a customer is eligible for in terms of business policy. Next we determine the customer's probability of take-up for each of those offers by using predictions from a model. These probabilities are compared with a threshold probability for each offer to determine whether a customer is to be made an offer. The next stage is to decide which offer to make to the customer. Experimental design principles are utilised to capture more information that will be used to monitor and refine the predictive models. Once the offer is made to the customer, the take-up is recorded, and the relevant predictive model updated. The updated predictive models are then used for the next round of marketing offers.

4.1. Customer eligibility

Typically decisions concerning whether a customer is eligible to receive an offer compare the estimated prob-

ability of a response π_{ij} with a threshold c , and make the offer if $\pi_{ij} > c$ [11]. The procedure used to define c may include economic considerations surrounding making the offer and subsequent returns [15, 20].

Our modification to the standard procedure chooses the threshold c_j to select customer j as being eligible to receive offer i if $\Pr(\pi_{ij} > c_j) \geq \alpha$ for some value of α . Refinements to models will change our choice of c_j as model accuracy improves. Random cell methodology is added to ensure that a randomly selected percentage of customers ξ_r is selected as being eligible regardless of the value of π_{ij} . This ensures that we continue to obtain realistic evaluations of the accuracy of our models and explore alternatives.

4.2. Offer decisions and experimentation

A naive decision process to find the best offer o_i out of m alternatives is, for customer j , select o_i such that $\pi_{ij} = \max_i \pi_{ij}$, that is, select the offer with the highest predicted probability. However, an offer of higher estimated π_i may not have a greater actual probability than others. The decision criteria need to incorporate the accuracy of the models so that a poor model will not gain preference over a good model.

Our framework includes an experimental protocol in the selection of which offer to make. The algorithm provides the offer to be made to the customer: a control offer, the best offer; or a testing offer. A customer is allocated to a control cell by ensuring that a random percentage of customers ξ_c is made no offer.

Two possible algorithms are outlined below.

Algorithm 1

```

For each customer  $j$  {
  Select offer  $o_i$  from the set of all eligible offers
   $O = \{o_1, \dots, o_m\}$  such that  $\pi_{ij} = \max_i \pi_{ij}$ 
  Add offer  $o_i$  to the set of best offers  $O' = \{o_i\}$ 
  For each offer  $k \neq i$  {
    Add offer  $o_k$  to  $O' = \{o_i, \dots\}$  if
     $\Pr(\pi_{ij} \geq \pi_{kj}) < \alpha$ 
  }
  Sample a random value  $x \in [0, 1]$ 
  If  $x > \xi_c$  then allocate an offer sampled at random from  $O'$ 
}

```

Algorithm 2

```

For each customer  $j$  {
  For each eligible offer  $o_i$  {
    Estimate  $\Pr(o_i = o_{(1)}) = p_i$ , the probability
    that offer  $o_i$  is the first ranked offer  $o_{(1)}$ 
  }
  Sample a random value  $x \in [0, 1]$ 
}

```

```

    If  $x > \xi_c$  then allocate offer  $o_i$  with probability  $p_i$ 
}

```

4.3. Model updating

In many deployments of predictive models, the model's estimation algorithm is deployed as is, to be fitted against customers as required, or alternatively as fixed estimates of take-up probability determined by fitting the models across the population. Periodical checks of model performance are undertaken, and if the model's performance has deteriorated, the model will be re-calibrated.

For an automated self-learning system, updating and monitoring of predictive models needs to be done continuously or in piecewise batch mode. The mechanics of updating models are not considered in this paper.

Once an offer has been made to a customer, that customer will either accept the offer or decline. The customer's response y_j is then captured, along with data on o , the offer made, and the estimated probability of uptake π_{ij} , to give a tuple (y_j, π_{ij}) .

Information on the current accuracy of the model can be adjusted with the new observation, allowing consistent monitoring of the model's performance. If the accuracy of the model deteriorates, model updating is required.

4.5. The overall algorithm

```

For each model  $i$  {
  Calculate  $D_i = f(y_j, \pi_{ij})$ , the measure of error,
  based on the test set
}
For each customer  $j$  who meets the trigger criteria {
  Sample a random value  $x \in [0, 1]$ 
  If  $x \leq \xi_c$ 
    Then select offer  $o$  at random from the
    set of all offers
  Else {
    For each offer  $o_i$  {
      Calculate  $\pi_{ij}$  and  $(\pi_{ij}^-, \pi_{ij}^+)$ 
      If  $\Pr(\pi_{ij} \geq c) \geq \alpha$ , add  $o_i$  to the
      set of eligible offers  $O$ 
    }
    Make offer  $o$  using Algorithm 1 or 2
  }
}
For each customer  $j$  {
  Observe the response  $y_j$  to offer  $o$ , and covariates  $x_j$ 
  Update model  $y_j = g(x_j)$  and prediction error  $D_i^*$ 
  Compare  $D_i^*$  with  $D_i$ 
  If  $D_i^* \leq D_i$ 
    Then keep new model
}

```

```

    Else
        Generate alert for model rebuilding
}

```

5. Conclusion and future work

The framework outlined encapsulates the insights we have derived from our experience in continuous database marketing. It allows continuous campaigns to be implemented before models have been fully fitted. The combination of model updating and experimental design ensures that when model accuracy is poor, more information is collected to improve accuracy, and when accuracy is high, optimised offers are made to customers.

Our specific innovations are to include both data mining and experimentation simultaneously within the continuous database marketing process, extending the single-offer paradigm to managing multiple marketing offers and models in parallel, and instituting statistically well founded processes to monitor and enhance models during a continuous marketing campaign. Our random cell methodology guards against models using biased samples.

There are several areas of work to be done. Extensions of the decision criteria can include economics factors. There are many alternative algorithms that could be used for experimentation. Response-surface methodology may be incorporated to search for an optimum.

However, the basic framework allows changes and enhancements while providing a platform from which marketing offers can be made in a continuous manner.

6. References

- [1] R. Hedgcock, "Transforming Data Into Marketing Insight," in *Automatic I.D. News*, vol. 14, 1998, pp. 50.
- [2] S. Grimes, "CRM Evolution," *Intelligent Enterprise*, vol. 2, pp. 14–17, 1999.
- [3] P. Demery, "The Decade of Marketing," *Credit Card Management*, vol. 11, pp. 74–84, 1999.
- [4] D. Haughton and S. Oulabi, "Direct Marketing Modeling with CART and CHAID," *Journal of Direct Marketing*, vol. 11, pp. 42–52, 1997.
- [5] J. Magidson, "Improved Statistical Techniques for Response Modeling: Progression Beyond Regression," *Journal of Direct Marketing*, vol. 2, pp. 6–18, 1988.
- [6] A. Berson, S. Smith, and K. Thearling, *Building Data Mining Applications for CRM*. New York: McGraw-Hill, 2000.
- [7] S. Bhattacharyya, "Direct Marketing Response Models Using Genetic Algorithms," in *The Fourth International Conference on Knowledge Discovery and Data Mining*, R. Agrawal and P. Stolorz, Eds. New York: American Association for Artificial Intelligence, 1998, pp. 144–148.
- [8] J. R. Bult and T. Wansbeek, "Optimal Selection for Direct Mail," *Marketing Science*, vol. 14, pp. 378–394, 1995.
- [9] A. E. Eiben, T. Euverman, W. Kowalczyk, E. Peelen, F. Slisser, and J. Wesseling, "Comparing Adaptive and Traditional Techniques for Direct Marketing," presented at Proceedings of the 4th European Congress on Intelligent Techniques and Soft Computing, 1996.
- [10] F. Gönül and Z. S. Meng, "Optimal Mailing of Catalogs: A New Methodology Using Estimable Structural Dynamic Programming Models," *Management Science*, vol. 44, pp. 1249–1262, 1998.
- [11] N. Levin and J. Zahavi, "Continuous Predictive Modeling: A Comparative Analysis," *Journal of Interactive Marketing*, vol. 12, pp. 5–22, 1998.
- [12] C. X. Ling and C. Li, "Data Mining for Direct Marketing: Problems and Solutions," in *The Fourth International Conference on Knowledge Discovery and Data Mining*, R. Agrawal and P. Stolorz, Eds. New York: American Association for Artificial Intelligence, 1998, pp. 73–79.
- [13] N. Piersma and J.-J. Jonker, "Determining the Direct Mailing Frequency with Dynamic Stochastic Programming," *Econometric Institute, Rotterdam EI2000-34/A*, 2000.
- [14] B. Ratner, "A Technical Illustration of Alternative Direct Marketing Response Models," *Journal of Targeting, Measurement and Analysis for Marketing*, vol. 8, pp. 379–387, 2000.
- [15] B. Masand and G. Piatetsky-Shapiro, "A Comparison of Approaches For Maximizing Business Payoff of Prediction Models," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, E. Simoudis, J. Han, and U. Fayyad, Eds. Menlo Park, California: AAAI Press, 1996.
- [16] P. D. Berger and T. L. Magliozzi, "Experimental Design in Direct Mail and the Application of Taguchi Methods," *Journal of Direct Marketing*, vol. 7, pp. 44–54, 1993.
- [17] R. C. Blattberg, "Research Opportunities in Direct Marketing," *Journal of Direct Marketing*, vol. 1, pp. 7–14, 1987.
- [18] A. Hirschowitz, "Closing the CRM Loop," *Journal of Targeting, Measurement and Analysis for Marketing*, vol. 10, pp. 168–178, 2001.
- [19] G. Herschel, "Optimizing Customer Relationship Management Through Every Interaction," Gartner, Stamford 2001.
- [20] S. Bhattacharyya, "Direct Marketing Performance Modeling Using Genetic Algorithms," *INFORMS Journal on Computing*, vol. 11, pp. 248–257, 1999.