Knowledge-Based Systems 22 (2009) 327–335

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A new marketing strategy map for direct marketing

Young Ae Kim [a,*], Hee Seok Song [b], Soung Hie Kim [a]

[a] Business School, Korea Advanced Institute of Science and Technology (KAIST), 207-43 Cheongyangri 2-dong, Dongdaemoon-gu, Seoul 130-722, South Korea
[b] Department of Management Information Systems, Hannam University,133 Ojung-dong, Daeduk-gu, Daejon 306-791, South Korea

ARTICLE INFO

Article history:
Received 2 May 2007
Received in revised form 25 September 2008
Accepted 17 February 2009
Available online 6 March 2009

Keywords:
Sequential decision making
Reinforcement Learning
Direct marketing strategy
Customer relationship management
Marketing strategy map

ABSTRACT

Direct marketing is one of the most effective marketing methods with an aim to maximize the customer's lifetime value. Many cost-sensitive learning methods which identify valuable customers to maximize expected profit have been proposed. However, current cost-sensitive methods for profit maximization do not identify how to control the defection probability while maximizing total profits over the customer's lifetime. Unfortunately, optimal marketing actions to maximize profits often perform poorly in minimizing the defection probability due to a conflict between these two objectives. In this paper, we propose the *sequential* decision making method for profit maximization under the given defection probability in direct marketing. We adopt a Reinforcement Learning algorithm to determine the sequential optimal marketing actions. With this finding, we design a marketing strategy map which helps a marketing manager identify sequential optimal campaigns and the shortest paths toward desirable states. Ultimately, this strategy leads to the ideal design for more effective campaigns.

1. Introduction

Direct marketing is one of the most effective marketing methods with an aim to maximize the expected profits [13]. A number of cost-sensitive learning methods which focus on predicting profitable customers have been proposed for direct marketing [2,3,13,15]. However, a common objective of these methods is to only maximize the short-term profit associated with each marketing campaign. They ignore the interactions among decision outcomes when sequences of marketing decisions are made over time. These independent decision-making strategies cannot guarantee the maximization of total profits generated over a customer's lifetime because they often inundate profitable customers with frequent marketing campaigns or encourage radical changes in customer behavior [10]. This approach can decrease customer profitability because of the annoyance factor or their budgetary limits per unit time.

Some researchers have recognized the importance of sequential decision making to overcome the limitations of isolated decision making. For example, Pednault et al. [10] and Abe et al. [1] proposed sequential cost-sensitive learning methods for direct marketing. These sequential cost-sensitive methods, however, fail to consider the cost generated from customer defections. Although a primary objective of direct marketing is to maximize total profit, it is also important to control the probability of customer defection, keeping it under a desirable or acceptable level because the occurrence of a customer defection brings about tangible and intangible loss, (i.e., an increase of acquisition cost of a new customer, loss of word-of-mouth effects, and loss of future cash flows and profits). Since customer switching costs are much lower in e-commerce marketplaces, a company always needs to pay more attention to customer defection. However, current sequential cost-sensitive methods for maximizing profit do not indicate how to control the probability of customer defection while maximizing total profits over the customer's lifetime. Unfortunately, optimal marketing actions designed to maximize profits often perform poorly in minimizing the probability of customer defection due to a conflict between a profit maximization and defection probability minimization. For example, an optimal marketing action for profit maximization is liable to give up unprofitable customers who are most likely to defect but are profitable from a *long-term* perspective. In contrast, an optimal marketing action for the minimization of defection probability is apt to unnecessarily sacrifice loyal customers' profit with excessive marketing cost.

To overcome this conflict, we regard the customer defection probability as a constraint and try to control it under the given threshold because, in general, controlling defection probability under the threshold is more cost effective than completely avoiding customer defection with 0%. We also think that most companies have more interest in a strategy which guarantees the maximization of total profits while the defection probability is bounded by a desirable or acceptable level.

* Corresponding author. Tel.: +82 2 958 3684; fax: +82 2 958 3604.
E-mail addresses: kya1030@business.kaist.ac.kr, kya1030@business.kaist.edu (Y.A. Kim), hssong@hannam.ac.kr (H.S. Song), seekim@business.kaist.ac.kr (S.H. Kim).

0950-7051/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2009.02.013

In this paper, we have developed a sequential decision-making methodology for profit maximization under the given defection probability constraint. For effective sequential learning, we have adopted the Reinforcement Learning algorithm. We have also suggested the concept of a marketing strategy map which visualizes the results of learning such as an optimal marketing action in each state and customer's behavior dynamics according to suggested marketing actions. This marketing strategy map can help a company identify sequential optimal campaigns and the shortest paths toward desirable states. Ultimately, this strategy leads to the ideal design for more effective campaigns.

The rest of this paper is organized in the following manner: In Section 2, a Self-Organizing Map and Reinforcement Learning that are prerequisites for our study are briefly introduced. Section 3 details our method for direct marketing and Section 4 reports experimental results with real-world data sets. Section 5 describes a marketing strategy map and its applications. Finally, Section 6 summarizes our works and contributions.

2. Background

The proposed method adopts a Self-Organizing Map (SOM) and Reinforcement Learning for effective sequential learning and its visualization.

2.1. Self-Organizing Map (SOM)

The SOM [8,11] is a sophisticated clustering algorithm in terms of the visualization of its clustering results. It clusters high-dimensional data points into groups and represents the relationships between the clusters onto a map that consists of a regular grid of processing units called "neurons." Each neuron is represented by an n-dimensional weight vector, $\mathbf{m} = [m_1, m_2, \ldots, m_n]$ where n is equal to the dimension of the input features. The weight vector of each neuron is updated during iterative training with input data points. The SOM tends to preserve the topological relationship of the input data points so the similar input data points are mapped onto nearby output map units. This topology-preserving property of SOM facilitates the ability to design the marketing strategy map in our proposed method. In our method below, we define the possible customer states using SOM, and with the output map of SOM, we design the marketing strategy map.

2.2. Reinforcement learning

Reinforcement Learning [9,12] is characterized by goal-directed learning from interaction with its environment. At each discrete time t, the learning agent observes the current *state* $s_t \in S$, where S is the set of possible states in a system and selects an action $a_t \in A(s_t)$, where $A(s_t)$ is the set of actions available in state s_t. As a consequence of its action a_t in state s_t, the agent receives an immediate positive or negative*reward* r_{t+1}, and next state s_{t+1}. Based on these interactions, the agent attempts to learn a policy $\pi: S \to A$ which is a function of mapping states to actions to maximize the expected sum of its immediate rewards, $R = \sum_{t=0}^{\infty} \gamma^t r_t$ [where γ (i.e.,$0 \leqslant \gamma < 1$) is *a discount rate*]. Thus, Reinforcement Learning is particularly well suited to multi-step decision problems where the decision criteria can be represented in a recursive way as a function of the immediate numerical value [4].

3. The proposed method

We suggest the following method for profit maximization under the control of defection probability in direct marketing. As shown in Fig. 1, we prepared customer episodes with campaigns and the response history data and adopted the Reinforcement Learning algorithm to determine an optimal policy. We then design a marketing strategy map. To provide more simple and practical business intelligence, we designed a method for segmentation marketing instead of for individualized marketing.

3.1. Definition of states and actions

States are representations of the environment that the agent observes and are the basis on which agent's decisions are made. In this method, states would be represented as customer segments which have similar purchase patterns and response behaviors against promotion (e.g., recency, frequency, and monetary value) at the time of each campaign. In the rest of the paper, the following terms are used interchangeably: "state" and "customer segment." Thus,

$$S = \{s1, s2, \ldots, sN\}$$

where S is the set of states, N is the total number of states.

The actions are defined as all of the marketing campaigns conducted in a company. As the number of campaigns increases, companies feel the need to analyze the effects of diverse competing campaigns in each state (e.g., customer segments) in a systematic way. Thus,

$$A = \{a1, a2, \ldots, aM\}$$

where A is the set of actions, M is the total number of actions (i.e. campaigns)

3.2. Definition of profit and defection probability

The agent achieves both profit and defection probability as immediate rewards at each transition. An immediate profit P is the net profit which is computed as the purchase amount minus the cost of action. An immediate defection probability D is computed as the probability of falling into a fatal state (i.e., defection state). The concept of fatal state was first introduced by Geibel [5,6] who noted that processes, in general, have a dangerous state which the agent wants to avoid by the optimal policy. For example, a chemical plant where temperature or pressure exceeds some threshold may explode. Thus, the optimal strategy of operating a plant is not to completely avoid the fatal state when considering the related control costs, but to control the probability of entering a fatal state (i.e., an exploration) under a threshold.

In this method, a fatal state means the status of customer defection. Like an exploration in a chemical plant, customer defection is fatal to a company and brings about tangible and intangible loss. However, it is difficult to reflect both the tangible and intangible loss from defection to the reward of profit. It is also impossible and cost-ineffective to completely avoid customer defection, but customer defection could be controlled under the threshold – an acceptable or desirable level for a company. The defection probability means the customer defection rate of each state as well as the defection probability of a customer in each state. The immediate defection probability D on transition from s to s' under action a is defined by:

$$D(s, a, s') = \begin{cases} 1 & \text{if} \quad s \text{ is a non-fatal state, } s' \text{ is a fatal state} \\ 0 & \text{else} \end{cases} \tag{1}$$

If the agent enters a fatal state from a non-fatal state, the immediate defection probability is 1 and the immediate profit is 0. It is natural to consider a fatal state as a final state (i.e., an absorbing state) in which the agent ends its learning with the current sequence (i.e., a sequence of (s, a, r) sets).

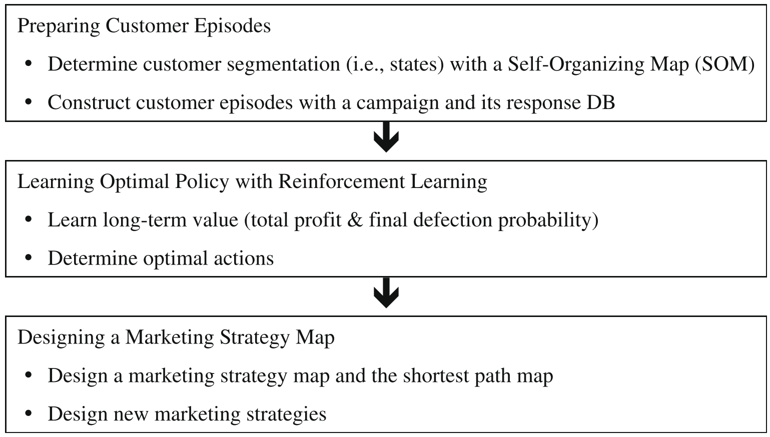

Fig. 1. The proposed framework for sequential marketing campaigns.

3.3. Learning strategy

The objective of the proposed method is to maximize the total profit while the defection probability is controlled under the given threshold for all states, as follows:

$$V_P^\pi(s) = E\left(\sum_{t=0}^{\infty} \gamma_P^t P_t\right) \rightarrow \max$$
$$V_D^\pi(s) = E\left(\sum_{t=0}^{\infty} \gamma_D^t D_t\right) \leqslant \theta \qquad (2)$$

where $V_P^\pi(s)$ is the cumulative profits and $V_D^\pi(s)$ is the probability that an agent ends in a defection state, when it starts in state s. An immediate profit P and an immediate defection probability D are discounted by discount rates γ_P $(0 < \gamma_P < 1)$ and $\gamma_D = 1$, respectively. Since γ_P is lower than 1, the agent will try to reach a more profitable state as quickly as possible, controlling the defection probability under the threshold. In addition, since γ_D is 1 and D is defined by (1), a value of $\sum_{t=o}^{k} \gamma_D^t D_t$ is 1, if and only if a customer in state s enters a defection state and ends his relationship with the company at time k. If not, a value of $\sum_{t=o}^{k} \gamma_D^t D_t$ is zero and a customer in state s continuously has a relationship with a company.

In order to construct an optimal policy π^*, the *state–action value function* $Q^\pi(s,a)$, which is a value of taking action a in state s under policy π, is computed by Watkin's Q-learning algorithm [14]. The state–action value function $Q_P(s,a)$ and $Q_D(s,a)$ can be defined by:

$$Q_P(s,a) = \mathrm{E}[P(s,a) + \gamma_P V_P^*(s')]$$
$$Q_D(s,a) = \mathrm{E}[D(s,a) + \gamma_D V_D^*(s')] \qquad (3)$$

where $P(s,a)$ and $D(s,a)$ are an immediate profit and defection probability of taking action a in state s, respectively, and $V_P^*(s')$ and $V_D^*(s')$ are optimal values of the next state s' under the optimal policy π^*. To optimize the total profit under the given defection probability, the optimal policy is selected by *a reverse-1st lexicographic ordering* (i.e., action a is preferred to a' if $Q_1 < Q_1'$, or if $Q_1 = Q_1'$ and $Q_2 \geqslant Q_2'$):

$$\pi^*(s) = \arg\max\nolimits_a \succeq \left[\max(Q_D^\pi(s,a),\theta), Q_P^\pi(s,a)\right] \qquad (4)$$

where $\max\left(Q_D^\pi(s,a),\theta\right)$ is higher value between $Q_D^\pi(s,a)$ and θ.

The agent prefers an action a to a' if $\max\left(Q_D^\pi(s,a),\theta\right) < \max\left(Q_D^\pi(s,a'),\theta\right)$ or if $\max\left(Q_D^\pi(s,a),\theta\right) = \max\left(Q_D^\pi(s,a'),\theta\right)$ and $Q_P^\pi(s,a) \geqslant Q_P^\pi(s,a')$. If several marketing actions would have $Q_D^\pi(s,a)$ less than threshold θ, they have the same value θ compared to the first component, $\max\left(Q_D^\pi(s,a),\theta\right)$. Then, the agent compares the second component $Q_P^\pi(s,a)$and selects the action with the highest profit value as an optimal action.

Fig. 2 shows the algorithm for learning $Q_P(s,a)$ and $Q_D(s,a)$ and achieving the optimal policy $\pi^*(s)$. Input training data E is a set of episodes where each episode is a sequence of events, and each event consists of a state, an action, profit, and defection probability. Each episode represents the campaign interactions between a customer and a company as time goes on. Note that we introduce a dummy state s_{def} and a dummy action a_{def} for a defection state and its action for technical reasons. If a customer falls into a defection state, we construct his last event with $\langle s_{def}, a_{def}\rangle$. We also compute both $Q_P(s_{def}, a_{def})$ and $Q_D(s_{def}, a_{def})$ at zero (i.e., no values), because customers entering a defection state have no permanent rewards. $Q_P(s,a)$ and $Q_D(s,a)$, which are N- by- M matrix where N is the number of states and M is the number of actions, are updated with each episode from line 2 to 8. At line 4, α is a step-size-parameter which affects the rate of convergence to $Q^*(s,a)$. Since α is set up to be a decreasing function of t (at line 5), we can assure convergence to $Q^*(s,a)$ as $t \rightarrow \infty$, for all (s,a). After the learning for all episodes, the agent can achieve the optimal policy in each state at line 9 and 10. However, the agent changes the optimal action a^* into "*no action*" if the cumulative profit (i.e.,$Q_P(s,a^*)$) is negative regardless of the Q_D (s,a^*) value. If companies determine that "no action," is needed for the state, they have to give up customers in the state or develop new campaigns which are especially effective for the state.

4. Experiments

To the best of our knowledge, this is the first study suggesting sequential optimal marketing actions for maximizing long-term total profit while keeping the defection probability below a given threshold, and the first study to design a marketing strategy map for these purposes. To evaluate our sequential decision making method's feasibility in direct marketing, we experimented with a part of KDD-CUP-98 datasets [7] which concerns direct-mail promotions soliciting donations.

4.1. Data sets and pre-processing

The dataset for experiments consists of 95,412 records. Each record contains each donor's direct-mail promotion pattern (e.g., which direct-mail was sent or not, when it was sent, etc.) for 22 campaigns conducted monthly for about two years. Other information such as response behavior against each promotion (e.g., whether a donor responded or not, how much was donated) was also collected.

Input: a set of episodes $E = \{E_1, E_2, \ldots, E_i\}$

where $E_i = \{< s_{i,1}, a_{i,1}, P_{i,1}, D_{i,1} >, < s_{i,2}, a_{i,2}, P_{i,2}, D_{i,2} >, \ldots, < s_{i,j}, a_{i,j}, P_{i,j}, D_{i,j} >\}$,

Threshold θ, $\gamma_P (0 < \gamma_P < 1)$, $\gamma_D = 1$.

Output: state-action value function $Q_P(s,a), Q_D(s,a)$,

optimal policy(action) $\pi^*(s)$ in each state.

1. $Q_P(s,a) = 0; Q_D(s,a) = 0;$
2. For each episode E_i
3. For $l = 1$ to j -1
4. $\alpha_t = \dfrac{1}{1+t};$
5. where $t = visit(s,a)$; //total number of times that (s,a) has been visited up to
6. $Q_P^t(s_{i,l}, a_{i,l}) = (1-\alpha_t) Q_P^{t-1}(s_{i,l}, a_{i,l}) + \alpha_t \left(P_{i,l} + \gamma_P Q_P^{t-1}(s_{i,l+1}, a^*)\right);$
7. $Q_D^t(s_{i,l}, a_{i,l}) = (1-\alpha_t) Q_D^{t-1}(s_{i,l}, a_{i,l}) + \alpha_t \left(D_{i,l} + \gamma_D Q_D^{t-1}(s_{i,l+1}, a^*)\right);$
8. where $a^* = \arg\max_a \succeq \left[\max\left(Q_D^{t-1}(s_{i,l+1}, a), \theta\right), Q_P^{t-1}(s_{i,l+1}, a)\right]$
9. For each state s
10. Find $\pi^*(s) = \arg\max_a \succeq \left[\max\left(Q_D(s,a), \theta\right), Q_P(s,a)\right];$

Fig. 2. The learning strategy with profit maximization under the given defection probability.

For effective experiments, we classified the original dataset into two donor groups. The first group included donors who often responded to campaigns except for the last (22nd) campaign. We collected data from the "active donors" group by excluding the data from the last two campaigns from each donor in this group. The second group included donors who had previously actively donated, but stopped donations long before the last (22nd) campaign. By definition of Paralyzed Veterans of America (PVA, a donor of KDD-CUP-98 datasets), the second group included "the lapsed donors" who had not made a donation within the last 12 months. Out of this group, we prepared data from the defector's group by collecting campaigns and response history until the donors became lapsed donors. We defined the lapsed donors as a fatal state (i.e., a defection state) and had an agent learn the optimal policy controlling the probability of being a lapsed donor. The original dataset had some fields showing whether a donor would become a lapsed donor or not at each promotion. We sampled 10,000 records which consisted of 50% active donors and 50% defectors to equally ascertain information from both groups.

In the original dataset, the set of actions, A had 11 types of actions (e.g., direct-mail campaigns), and we gave a number to each action from $a1$ to $a11$. In order to determine a set of possible states S, we conducted customer segmentation using a Self-Organizing Map (SOM). The input features of the SOM include 14 features regarding the promotion pattern and response behavior which were collected at the time of each campaign (See Table 1).

Since the SOM was given no information about the optimal number of states, we had to experiment with the number of states of the SOM. For experiments to test our method, we chose two SOM models, 6×8 SOM (48 states) and 6×7 SOM (42 states). The 6×8 SOM outperformed the others by achieving the highest average total profit over all the states and was optimal for profit maximization. The 6×7 SOM outperformed the others by achiev-

Table 1
Input features of SOM.

Category		Features	Descriptions
Promotion pattern		tot_num_pro	Total number of promotions to date
		tot_num_pro_6m	Total number of promotions in the last 6 months
Response behavior	History	tot_amt_don	Total amount of donations to date
		tot_num_don	Total number of donations to date
		amt_per_don	Average amount per donation to date (*tot_amt_don/tot_num_don*)
		frequency	Response rate to date (*tot_num_don/tot_num_pro*)
		amt_per_pro	Average amount per promotion to date (*tot_amt_don/tot_num_pro*)
	Recent (6 months)	tot_amt_don_6m	Total amount of donations in the last 6 months
		tot_num_don_6m	Total number of donations in the last 6 months
		amt_per_don_6m	Average amount per donation in the last 6 months (*tot_amt_don_6m/tot_num_don_6m*)
		frequency_6m	Response rate in the last 6 months (*tot_num_don_6m/tot_num_pro_6m*)
		amt_per_pro_6m	Average amount per promotion in the last 6 months (*tot_amt_don_6m/tot_num_pro_6m*)
	Last	recency	Number of months since the last donation
		last_amt	Amount of the last donation

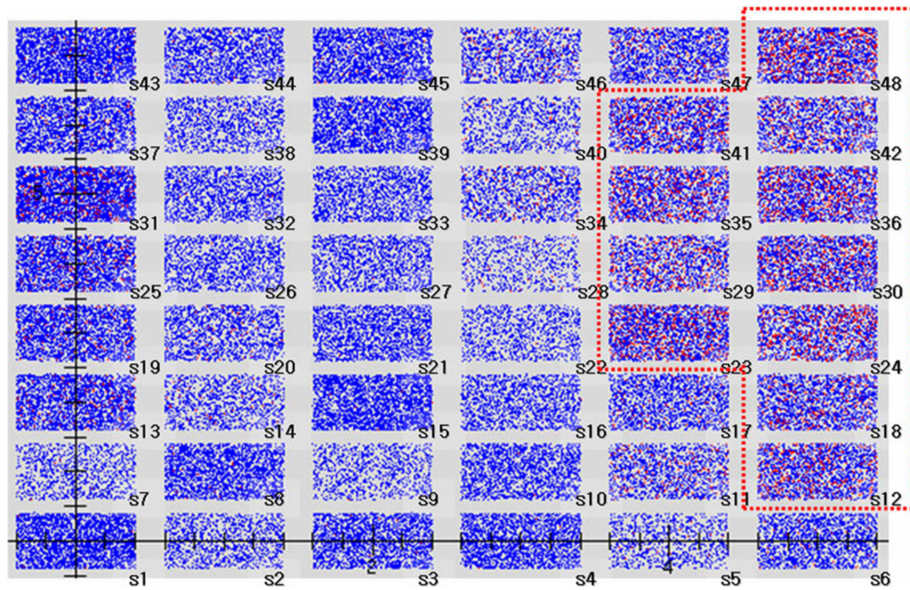

Fig. 3. A set of states on the output map of SOM (6 × 8 SOM).

ing the lowest average defection probability over all the states and was optimal for defection probability minimization.

Fig. 3 illustrates the result of the 6 × 8 SOM which we conducted with Clementine, a SPSS data mining tool. Each state's number was assigned to distinguish the states, (e.g., $s1, s2, \ldots, s48$). In Fig. 3, each point represents a donor at the time of each campaign and each red point is a donor who became a lapsed donor in the next period according to the current campaign. We can see that the states in the red box ($s12, s18, s23, s24, s29, s30, s35, s36, s41, s42, s48$) have more red points than the other states. The probability of being a lapsed donor in the next period in these risky states was over 10%. To be more specific, the average probability of being a lapsed donor in the next period over these risky states was 16.1%, while the average over all states was 5.2%.

After determining the customer segmentation, we constructed a customer episode for each donor. Whenever an action a was conducted, we determined a state s through the trained SOM model. We then calculated the defection probability and a profit value of the state–action set (s,a). The immediate defection probability was 0 or 1 according to whether the donor was a lapsed donor the next time or not. The profit value was computed by the donation amount – the cost of the campaign ($0.68) including the mail cost. In the same way, we prepared 10,000 episodes as input training data.

4.2. Results

In our method, the threshold reflects a desirable or acceptable level of customer defections for each company. The decision of the threshold is invariably dependent on several factors including the conditions of the market and the characteristics and goals of each company. Therefore, we experimented with our method using various levels of thresholds and the selected SOM models (i.e., 6 × 8 SOM and 6 × 7 SOM). We then observed the change of the average Q_P and Q_D values. For finite experiments on the threshold, we increased the value of the threshold by 0.05 (5%) within a meaningful range of thresholds.

Table 2 shows the results of the basic models which were learned for profit maximization or defection probability minimization, respectively. The highest defection probability of each basic model provided the lower and upper bound on experiment thresholds. In the case of the 6 × 8 SOM model, if the agent learns to minimize the defection probability, the agent is able to control the defection probability under 0.0307 over all states. This value of 0.0307 is the lowest threshold which the agent is able to achieve with the 6 × 8 SOM. In contrast, if the agent learns to maximize the total profit, the agent does not consider the defection probability and, therefore, retains the defection probability under 0.3025 over all states. The value of 0.3025 is the upper bound of the threshold. Values over this threshold are meaningless in the 6 × 8 SOM because the agent is not able to achieve more than the average total profit in the total profit maximization model (i.e., $28.91) even though the threshold is increased over 0.3025. Based on the results in Table 2, we changed the thresholds by 0.05 between 0.0307 and 0.3025 in the 6 × 8 SOM model and between 0 and 0.2846 in the 6 × 7 SOM model.

Fig. 4 shows the performance comparison of our methods with different thresholds. We compared the average Q_P and Q_D values over all starting states assuming equal distribution of donors into all states. Note that the first bar and triangular point correspond to the basic model for minimization of the defection probability, and the last bar and triangular point are the basic model for maximization of the total profits. As mentioned earlier, we were able to observe the conflict between the two marketing objectives: maximization of the total profit and minimization of the defection probability. As we achieve more total profit by alleviating the

Table 2
Results of basic models for the decision of a meaningful range of thresholds.

	Min. of defection probability			Max. of total profit		
	Average total profit	Average defection pro.	The highest defection pro.	Average total profit	Average defection pro.	The highest defection pro.
6 × 8 SOM	7.29	0.0013	**0.0307**	28.91	0.1333	**0.3025**
6 × 7 SOM	7.41	0	**0**	27.28	0.1400	**0.2846**

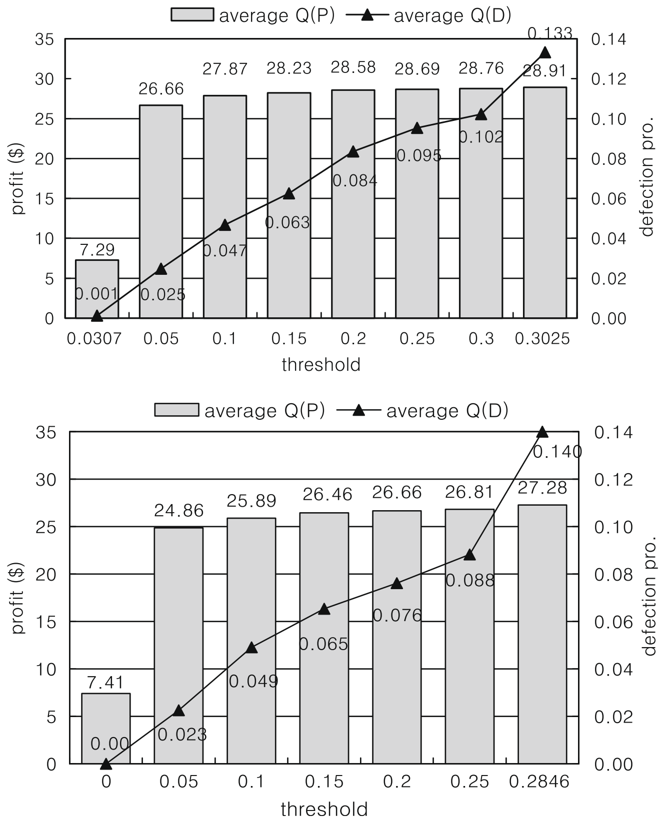

Fig. 4. Performance comparison of different thresholds. (a) Average Q_P and Q_D value in the 6 × 8 SOM model; (b) Average Q_P and Q_D value in the 6 × 7 SOM model.

constraints of the defection probability, the number of donors who are apt to defect increases more. Fig. 4 also shows that the basic model for minimization of the defection probability achieves very poor performance of the total profits, $7.29. In addition, the basic model for maximization of the total profits achieves poor performance of the defection probability, 0.133 (13.3%), and the threshold, 0.3025 (30.25%), because it disregards the other objective. However, when learning with each threshold in our model, we could achieve far more satisfactory results of both the total profit and the defection probability.

At this point, marketing experts in each company could decide the threshold based on observed tradeoffs and acquired knowledge. For further analysis, we selected 0.05 (5%) based on Duncan's test (significance level = 0.05) of the average Q_P and Q_D values. Duncan's test in the 6 × 8 SOM showed that as the threshold increased from 0.05 to 0.3, the average Q_D value significantly increased at each step, but the average Q_P value was not significantly different. Duncan's test in the 6 × 7 SOM also showed similar results, so as the threshold increased from 0.05 to 0.15, the average Q_D value significantly increased at each step, but the average Q_P value was not significantly different. Therefore, when selecting 0.05 as the threshold in both the 6 × 8 SOM and 6 × 7 SOM, we could achieve a significantly lower average Q_D value than the models with thresholds of 0.1 or 0.15. We could also achieve the same average Q_P value as models with thresholds of 0.1 or 0.15 achieved.

To choose a better model, we took a *T*-test (significance level = 0.05) with two SOM models (the 6 × 7 SOM and 6 × 8 SOM with a threshold of 0.05). The *T*-test demonstrated that the 6 × 8 SOM model significantly outperformed the 6 × 7 SOM model. The 6 × 8 SOM achieved a significantly higher average Q_P value ($26.66) than the 6 × 7 SOM ($24.86), but a significant difference in the average Q_D value was not observed between the 6 × 8 SOM (0.025) and the 6 × 7 SOM (0.023).

Table 3 shows the performance comparison of our model with a threshold of 0.05 and the basic models in the 6 × 8 SOM. The last row in Table 3 gives the results when campaigning without any optimization model. The $Q_P(s,a)$and $Q_D(s,a)$ values in no optimization model are calculated as follows:

$$Q(s,a) = E[r(s,a) + \gamma V(s')] \\ = E[r(s,a)] + \gamma \sum_{s'} p(s'|s,a) Q(s',a') \tag{5}$$

Unlike other optimization models, there is no strategy to select optimal action a^* in transition state s'. It uses the Q_P (s',a') and $Q_D(s',a')$ values observed from a training dataset instead of the optimal values $Q_P^*(s',a^*)$ and $Q_D^*(s',a^*)$.

As shown in Table 3, our method significantly outperformed the basic models in terms of expected revenue (significance level = 0.1). The expected revenue is the average expected revenue generated from the surviving customers who do not defect. It is computed by multiplying the average total profit by the rate of surviving customers as follows:

$$\frac{\sum_S Q_P(s,a^*)(1 - Q_D(s,a^*))}{N} \tag{6}$$

where *N* is the total number of states, Q_P is the average total profits and $(1 - Q_D)$ is the rate of surviving customers, when following optimal action a^* in each state.

The last three columns of Table 3 show the improvement over no optimization model. Our model increased the average expected revenue by 7.48 times over no optimization model, while the basic models increased the average by 7.46 times and 2.07 times, respectively. Our method was able to improve the total profit and control the customer defection probability under the given threshold over all states and ultimately, achieve higher expected revenues. In contrast, the basic model was effective in optimizing each objective, but ineffective in considering its conflicting objective.

5. A marketing strategy map

We suggest the concept of a marketing strategy map and describe how to utilize the map for designing new marketing strategies.

5.1. Design of a marketing strategy map

To clearly show an optimal action and customer behavior dynamics in each state, we need to design a marketing strategy map. Fig. 5a illustrates the marketing strategy map of our experi-

Table 3
Performance comparison with basic models.

	Average Q_P	Average Q_D	Expected revenue	Threshold	Q_P Improvement	$(1 - Q_D)$ Improvement	Exp. revenue improvement
Max. of total profit	28.91	0.1333	25.95	0.3025	4.5 (28.91/6.38)	1.55 (0.8667/0.56)	7.46 (25.95/3.48)
Our model (θ = 0.05)	26.66	0.0248	**26.02**	0.05	4.2 (26.66/6.38)	1.74 (0.9752/0.56)	**7.48**(26.02/3.48)
Min. of defection Pro.	7.29	0.0013	7.20	0.0307	1.1 (7.29/6.38)	1.78 (0.9987/0.56)	2.07 (7.20/3.48)
No optimization model	6.38	0.44	3.48	0.67	–	–	–

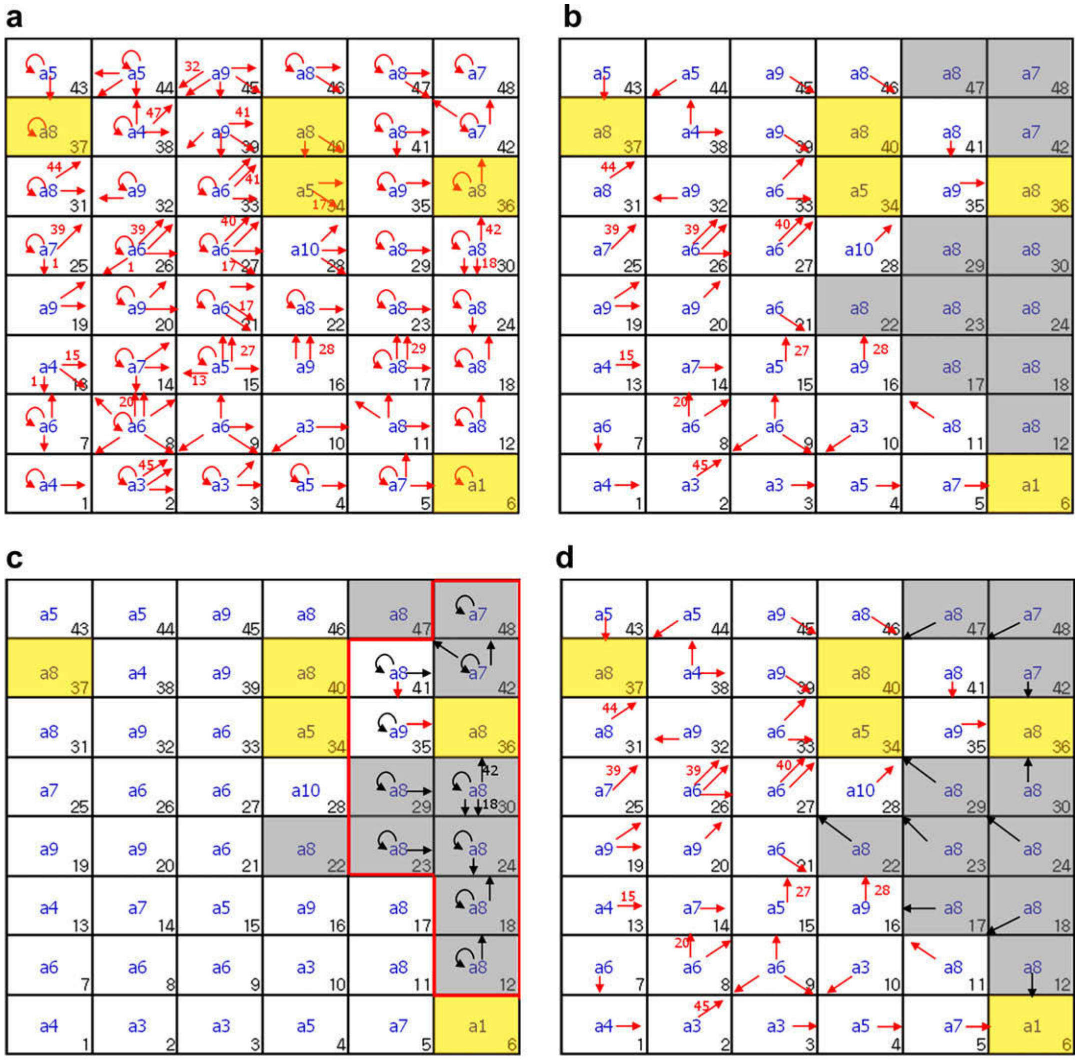

Fig. 5. The marketing strategy map and its applications.(a) Optimal actions and major customer paths; (b) Customers' shortest paths to the desirable states ($s6$, $s34$, $s36$, $s37$, and $s40$); (c) Customer paths in risky states (high defection probability states) and (d) A marketing campaign strategy for the ineffective states. * $s35$ (D_1), $s36$ (D_0) and $s41$ (D_2); * Red arrows in $s35$ and $s36$ are the shortest paths to the desirable state $s36$.

ments. To find customers' paths in each state, we exploited the association rules of the form ($state = s$ & $action = a^*$) → ($next$ $state = s'$), where action a^* is the optimal action of state s. We selected the association rules in order of high confidence until the sum of confidences from the selected rules was over 70%. Therefore, the strategy map can explain at least 70% of customer transitions by the optimal campaign. On average, our strategy map shows 77.3% of customer transitions over all states.

As expected, most customers shift from a current state to nearby states on the strategy map by the targeted campaign because input behavior patterns between two nearby states are mostly similar according to the topology-preserving property of the SOM. However, some customers significantly change their behavior states. We describe these transitions with a direction arrow and a state number on the map. For example, customers in state $s2$ move into state $s3$ (24.3%), $s9$ (21.6%) or $s45$ (16.2%) or remain in state $s2$ (21.6%) by action $a3$ in Fig. 5a.

The marketing strategy map also shows the desirable states to which a company attempts to drive customers. We selected the top 10% of states in terms of total profits as the desirable states (i.e. $s6$, $s34$, $s36$, $s37$ and $s40$). However, we did not consider defection probability, because defection probability is controlled under the given threshold over all states. Among these desirable states, state $s36$ was one of the risky states from a short-term perspective (See Fig. 3). However, it can be finally transformed into a desirable state through sequential optimal campaigns in our method.

5.2. Marketing strategy map applications

A marketing manager can utilize the marketing strategy map to design the shortest path map. The shortest path map can be used to design more effective campaign strategies as well as to identify sequential optimal campaigns and the shortest paths towards desirable states.

To design the shortest path map, we found a set of states with the shortest paths which lead to the desirable states after n period (i.e.D_n) (See Table 4). We first selected all rules of the form ($state = s \in (S - D_0)$ & $action = a^*$) → ($next$ $state = s' \in D_0$), where S is a set of all possible states and D_0 is a set of desirable states. With these rules, we found all states (i.e., D_1) and their paths which led to the desirable states after 1 period. We then selected all rules of the form ($state = s \in (S - \cup_{i=0,1} D_i)$ & $action = a^*$) → ($next$ $state = s' \in D_1$). With these rules, we found all states (i.e.D_2) and their paths which led to the desirable states after 2 periods via one of the states in D_1. By repeating this process, we finally found all the states from D_1 to D_4 and their paths. There was no further

Table 4
A set of states leading to the desirable states after n periods (D_n).

D_n	States
D_0	$s6$, $s34$, $s36$, $s37$, $s40$
D_1	$s5$, $s27$, $s33$, $s35$, $s39$, $s43$, $s44$, $s45$
D_2	$s2$, $s4$, $s15$, $s20$, $s25$, $s26$, $s28$, $s31$, $s38$, $s41$
D_3	$s1$, $s3$, $s8$, $s9$, $s13$, $s14$, $s16$, $s19$, $s32$, $s46$
D_4	$s7$, $s19$, $s11$, $s21$

D_n = {s| state s has the paths which lead to the desirable states after n periods}.

$D_n(n \geqslant 5)$. A state in D_n has direct paths to states in D_{n-1} and leads to states in D_0 via states in $D_{n-i}(i = 1,2,\ldots n-1)$, sequentially.

Fig. 5b illustrates the shortest path map. A marketing manager can identify sequential optimal campaigns and the shortest paths toward desirable states in each state. For example, customers in state $s38$ can go to the desirable states after 2 periods through 2 different paths: $s38$(by $a4$) → $s44$(by $a5$) → $s37$ or $s38$(by $a4$) → $s39$(by $a9$) → $s34$. The shortest path map can then be used to determine if a current campaign is effective or not by identifying states which have no path to the desirable states. In Fig. 5b, a total of 10 states, gray color states (e.g., $s12$,$s17$,$s18$), cannot lead to desirable states even though the suggested campaign is optimal. Among these 10 ineffective states, 8 states are risky states in which the probability of being a defector in next period is over 10%.

Fig. 5c shows customers' transition in risky states. One of the marketing manager's major concerns about risky states is to detect customer defections and drive potential defectors to safer and more profitable states (i.e., desirable states). However, customers in risky states except state $s35$ (D_1), $s36$ (D_0) and $s41$ (D_2) fail to escape these risky states through the current optimal campaigns. They are kept in risky states and cannot go to desirable states. Therefore, a marketing manager needs to develop new campaign strategies for ineffective states.

The objective of designing new strategies is to provide the shortest path to get to the desirable states. We adopt a gradual approach which suggests next state with the fastest path to desirable states among immediate neighbors of an ineffective state. This approach is based on the fact that it is very difficult to significantly change customer behavior in such a short period. Fig. 5 (d) illustrates new campaign strategies for the ineffective states. For example, we designed this strategy to drive customers in state $s22$ to state $s27$ (D_1) among its immediate neighbors, [i.e. $s15$ (D_2), $s16$ (D_3), $s21$ (D_4), $s27$ (D_1), and $s28$ (D_2)] (See Table 4) because state $s27$ has the fastest path to state $s34$, as one of the states in D_1. In the case of state $s29$, we could select two strategies, (i.e., $s34$ (D_0) and $s36$ (D_0)), but we selected state $s34$ with a higher total profit as the next state. In the case of state $s24$, it has no immediate neighbor states which have the shortest path to the desirable states. Therefore, after designing new strategies of its neighbors (i.e., $s17$,$s18$,$s23$,$s29$, and $s30$), we designed its strategy based on new strategies.

As shown in Fig. 5d, most of the shortest paths to desirable states including the shortest paths designed by new strategies have a trend towards state $s34$ and $s40$ because these two states are the top 2 total profit states. Therefore, we can say that our proposed method gives a good performance in terms of suggesting optimal marketing campaigns and designing new campaigns.

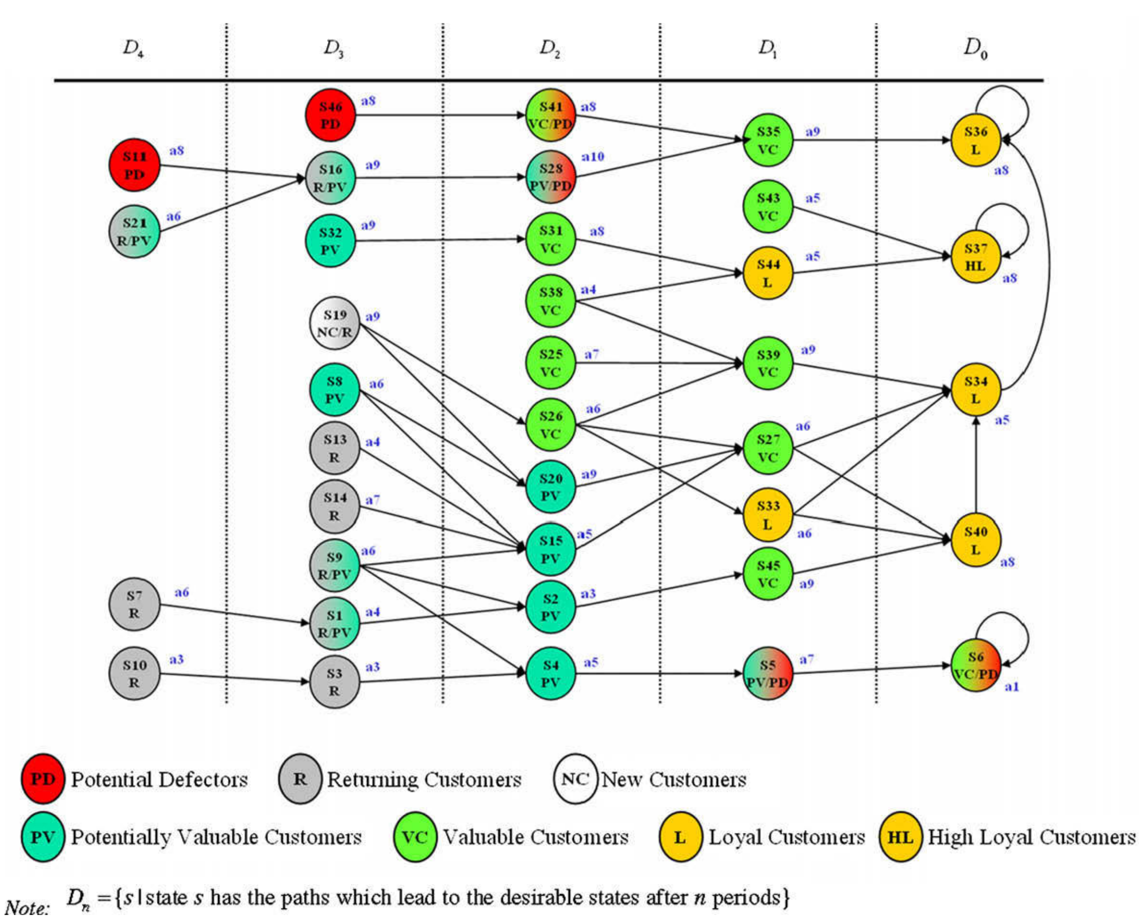

Fig. 6. Understanding of customer dynamics in a marketing strategy map. Note: {D_n = {s | state s has the paths which lead to the desirable states after n periods}.

5.3. Understanding of customer dynamics

In order to meaningfully interpret customer dynamics by sequential marketing campaigns, we need to understand each state in a marketing strategy map in Fig. 5a or b. We analyzed each state's input features of the SOM in Table 1 and then classified all states into 7 customer groups based on the knowledge of domain experts: (1) potential defectors whose total number of donations and total amount of donations are very low and did not give donations in the last 6 months; (2) returning customers who started to give donations again and increased donations in the last 6 months; (3) new customers who started giving donations; (4) potentially valuable customers whose total amount of donations is low, but are increasing in the last 6 months and who continuously donate a little amount of money; (5) valuable customers who continuously donate an average amount of money and also have responded to recent direct mailings; (6) loyal customers who donate more than an average amount of money and whose total number of donations and total amount of donations are higher than the average, and (7) high loyal customers who are the most highly profitable customers in the company.

Based on the analysis results of each state and the shortest path map in Fig. 5b, we could design a new form of a shortest path map in Fig. 6. Basically, this map shows the same shortest paths to the desirable states and optimal marketing campaigns as the shortest path map in Fig. 5b but, it is more meaningful and practical for a marketing manager to understand customer dynamics by a sequential marketing campaign.

With the shortest path map in Fig. 6, we can identify how a current customer status is changed by sequential optimal campaigns before becoming a loyal customer. For example, customers in state $s46$ (potential pefector) can become a loyal customer after 3 periods through the path of 'PD $s46$ (by $a8$) → VC/PD $s41$ (by $a8$) → VC $s35$ (by $a9$) → L $s36$' where VC/PD of state $s41$ means that a state consists of both valuable customers (VC) and potential defectors (PD). We can also find that most customers change their states from returning customers or potential defectors to potential valuable customers or valuable customers and then finally become loyal customers with these sequential marketing strategy.

As for customer dynamics by marketing campaigns, marketing experts mentioned that as shown in the marketing strategy map in Fig. 5a and the shortest path map in Fig. 5b and Fig. 6, marketing campaign rules and customer behavior dynamics are not simple in competitive marketplaces. Therefore, they should be different depending on where a customer is before, what kind of marketing campaign is conducted and where a customer is now by the marketing action.

6. Conclusion

While direct marketing has garnered a great deal of attention, few studies have addressed the tradeoff between two conflicting objectives such as the profit and defection probability even though these tradeoffs are of great interest to companies. To solve this tradeoff conflict, we have developed a sequential decision-making methodology for profit maximization under the given defection probability constraint. Our method suggests sequential optimal marketing actions for maximizing long-term total profit while controlling the defection probability under the threshold over a customer's lifetime. In addition, the suggested marketing strategy map clearly shows an optimal action and customers' behavior dynamics in each state. It also helps a marketing manager identify sequential optimal campaigns and the shortest paths toward desirable states and, ultimately, a design for more effective campaigns. Our experiments demonstrate the feasibility of our proposed method in direct marketing. The proposed method is a practical implementation procedure for direct marketing in telecommunications, online shopping malls, and other highly competitive marketplaces suffering from profit loss and customer defections.

References

[1] N. Abe, N. Verma, C. Apte, R. Schroko, Cross channel optimization marketing by reinforcement learning, in: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'04), 2004, pp. 767–772.
[2] P. Domingos, MetaCost: a general method for making classifiers cost sensitive, in: Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'99), ACM Press, 1999, pp.155–164.
[3] W. Fan, S.J. Stolfo, J. Zhang, P.K. Chan, AdaCost: misclassification cost-sensitive boosting, in: Proceedings of the. Sixteenth International Conference on Machine Learning (ICML'99), 1999, pp. 97–105.
[4] Z. Gabor, Z. Kalmar, C. Szepesvari, Multi-criteria reinforcement learning, in: Proceedings of the Fifteenth International Conference on Machine Learning (ICML'98), 1998, pp.197–205.
[5] P. Geibel, Reinforcement learning with bounded risk, in: Proceedings of the Eighteenth International Conference on Machine Learning (ICML '01), 2001, pp. 162–169.
[6] P. Geible, F. Wysotzki, Risk-sensitive reinforcement learning applied to control under constraints, Journal of Artificial Intelligence Research 24 (2005) 81–108.
[7] KDD Cup 1998 Data Web site, http://kdd.ics.uci.edu/databases/kddcup98/kddcup98.html, 1998.
[8] T. Kohonen, Self-Organizing and Associative Memory, Springer-Verlag, Berlin, 1995.
[9] T.M. Mitchell, Machine Learning, McGraw-Hill, 1997.
[10] E. Pednault, N. Abe, B. Zadrozny, Sequential cost-sensitive decision making with reinforcement learning, in: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD '02), ACM Press, 2002, pp. 259–268.
[11] O. Simula, P. Vasara, J. Vesanto, R. Helminen, The Self-Organizing Map in Industry Analysis, CRC Press, 1999.
[12] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, MIT Press, Cambridge, 1998.
[13] K. Wang, S. Zhou, J.M.S. Yeung, Mining customer value: From association rules to direct marketing, Data Mining and Knowledge Discovery 11 (2005) 57–79.
[14] C.J.C.H. Watkins, P. Dayan, Q-learning, Machine Learning 8 (1992) 279–292.
[15] B. Zadrozny, C. Elkan, Learning and making decision when costs and probabilities are both unknown, in: Proceedings of the Seventh ACM SIDKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'01), ACM Press, 2001, pp. 204–213.