Service Switching in Case-based Decisions following Bad Experiences: Online reviews data of Japanese hairdressing salons *

Hiroki Takahashi[1], Nariaki Nishino[1], and Ryuichiro Ishikawa[2][0000−0002−0290−1107]

[1] Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
{takahashi@css, nishino@tmi}.t.u-tokyo.ac.jp

[2] SILS, Waseda University, Shinjuku, Tokyo 169-8050, Japan
r.ishikawa@waseda.jp
http://www.f.waseda.jp/r.ishikawa/

Abstract. This paper examines consumer service switching from the perspective of case-based decision theory (CBDT), developed by Gilboa and Schmeidler. In contrast with the consumption of physical goods, it is difficult for consumers to evaluate their utility from service provision in advance because of intangibility. CBDT is a decision criterion that reflects consumers' past experiences, and enables us to examine their reasoning for service switching. Our paper empirically examines consumer choice behavior based on past experiences, using data from Japanese hairdressing salons, which consist of salon introductions and individual reviews of those salons. We focus on bad service experiences because CBDT suggests that after experiencing bad service, consumers will choose services that are less similar for their next salon appointment. Our paper examines whether CBDT accurately predicts the switching process for the service consumption. The results indicate that prior experiences have no significant effect on service choices.

Keywords: Case-based decision theory · Similarity functions · Online review data · Service switching · Service science .

1 Introduction

The paper examines a behavior hypothesis whereby people switch their service choice after a bad experience, which is derived from case-based decision theory (CBDT). CBDT is developed by Gilboa and Schmeidler [3, 4] within the

* This research is conducted based on the "Recruit dataset" provided by Recruit Technologies Co., Ltd. via the IDR Dataset Service of the National Institute of Informatics. Ishikawa is supported by JSPS KAKENHI Grant number JP17K18569. We thank Dr Amanda Fitzgibbons of Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

decision theory literature on and it employs case-based reasoning, as shown by Schank [12]. However, the authors main purpose is to axiomatize a new theory in a decision theoretical framework and to provide an alternative to the expected utility theory of von Neumann & Morgenstern [15] and Savage [11]. In contrast, this paper examines how CBDT can explain consumers' choices based on behavioral data on past experiences.

Our research is related to the emerging literature of service science, which considers the characteristics and improvement of the service industry. In contrast to physical goods, consumers cannot correctly estimate their utility before consuming a service. The qualities of services depend on both the ability of service providers and the chemistry between purveyors and consumers. In this sense, several different uncertainties affect service provision.

Therefore, researchers considering services focus on consumers' experiences and investigate how/what service should be provided given their various experiences. In this paper, we make use of CBDT to represent the consumers' decision criteria regarding their experiences, and examine how consistent their behavior is with the theory. In particular, our focus is the 'bad' service experiences because consumers may switch their choices after bad experiences. CBDT potentially has a role in explaining the process of consumers' choices.

Indeed, failure in service provision can often occur because of the different qualities involved and a 'lack of chemistry' between service providers and consumers. Once a consumer becomes disgruntled or dissatisfied with a service, he or she will not visit the service provider again. Naturally, service providers try to prevent such situations and improve on poor service provision. One task of service science researchers is to examine how services can be improved after such failures [5, 13, 14].

In this paper, we investigate how the bad services experiences influence service switching using a unique data set on Japanese hairdressing salon services available online (Recruit [8]). The data set provides individuals' reviews and details of why they chose the hairdressing salons that they visited. Therefore, if an individual assesses a hairdressing salon as offering poor service and then chooses a different salon for their next appointment, we can say they altered their choice due to the previous bad experience.

The remainder of this paper is organized as follows. The next section explains the composition and characteristics of the Japanese hairdressing salon data that we use in our analysis. Section 3 provides our empirical results from analyzing the decision-making of consumers after they have had bad service experiences. The final section summarizes the research and provides some comments regarding future research.

2 Hot Pepper Beauty: Japanese Beauty Salon Data

We use data from the largest booking website for Japanese hairdressing salons, *Hot Pepper Beauty*, provided by Recruit Technologies Co., Ltd [8]. There are many hairdressing salons in Japan, competing intensely to acquire customers.

Many of the salons join the Hot Pepper Beauty website and provide marketing information on their salons as well as discounts for customers booking salon appointments through the website. As customers can post reviews of the salon services, the website is useful because it allows customers to compare the services of many salons online..

As stated in the Introduction, the services offered by hairdressing salons are a typical service good in that the quality and satisfaction level of customers cannot be evaluated in advance. As an example, in the case of hairdressers, each customer will have a specific preference regarding hair styles and he or she may fail to communicate these preferences effectively to the hairdresser. Even when a customer explains his or preferences in detail, different hairdressers may understand the instructions in different ways. Moreover, the satisfaction level of customers may be decided not only by the actual hair styles achieved, but also by the atmosphere of the salon and the customers' communication with the hairdressers. Therefore, the Hot Pepper Beauty data set can assist us to understand the process of consumers' salon choices/switches because we can observe their repeated uses of the service every few months.

2.1 Descriptive Statistics

The data set is provided through a Japanese research organization, the National Institute of Informatics, and it is public data that are recorded from January 11, 2012 to January 9th, 2014. We use information provided by the salon on the hairdressers and information contained in reviews in the data. We excluded some incomplete data on hairdressers. Table 1 reports the descriptive statistics of the data that we actually used in our study.

Table 1. Descriptive statistics of the Japanese hairdressing salon data

Descriptive Statistics		Mean	S.D.
	#Reviews written by an individual	1.228	(0.749)
	#Reviews written to each hairdresser	8.838	(13.397)
	Evaluations (out of 5)	4.682	(0.660)
Count for Data		Count	
Review information	#Hairdressers reviewed	11,929	
Hairdresser	#Hairdresser	53,029	

Each component of the table is explained as follows:

Review Information: Each date-stamped review consists of the review text and an evaluation of the hairdresser who provided the service to the customer who has wrote the review. Each of the hairdressers is assigned an ID number. Then, we use this to refer to the hairdressers that each customer reviews.

The review information contains an evaluation of the service, indicated by a number from 1 to 5, where a higher number indicates a better service.

As each reviewer has an ID that is provided in the data set information with each review, we can develop a time series of review information by each reviewer.

Hairdresser Information: Hairdresser information in the data set includes the hairdressers who belong to a hairdressing salon. The hairdressers also provide
three types of information for consumers: their selling points, strengths, and a self introduction. All three types of information provided by the hairdressers are collected together in one long *document*, which we then use to estimate topics, as explained below.

Next, we examine consumers' switching choices from the perspective of CBDT by constructing a similarity measure from the data.

2.2 Choice Switching in CBDT

The CBDT developed by Gilboa and Schmeidler [3, 4] considers a decision-maker who has to make a choice. Unfortunately, he or she does not have any information on the utilities/payoffs from the choices. Therefore, he or she refers to similar experiences that he or she has had previously.

The situation is formulated as a triple (P, A, R) where P is a set of decision problems, A is a set of acts, and R is a set of possible outcomes. Then, a case (p, a, r) is an element of $P \times A \times R \equiv C$. Because C includes all the cases, these cases do not necessarily happen. The decision-maker has only partial experiences of the cases, which are remembered as memory $M \subset C$.

He or she also has a similarity function s from $P \times P$ to the closed interval between 0 and 1; $s : P \times P \to [0, 1]$. This function represents the degree of similarity between the two cases. We assume that $s(p, p) = 1$ for any identical problem p, and that $0 \leq s(p, p') < 1$ for any two different cases, p and p'.

As the decision-maker enjoys utilities from past outcomes, we represent them by a utility function $u : R \to \mathbb{R}$. Then, he or she chooses an action $a \in A$ to maximize the following objective function:

$$U(a) \equiv \sum_{(q,a,r) \in M} s(p, q)u(r).$$

The above equation represents the fact that the decision-maker would like to choose an identical action to that which yielded higher utilities on average in similar cases in the past. Note that the function is given as a weighted average of the utilities from his or her choice, which takes similarity into account.

From this perspective, we hypothesize in our data analysis that a customer switches from a hairdresser to a dissimilar hairdresser once he or she has had a bad experience. In the following, we examine how much her choice differs between the past hairdresser and the next hairdresser chosen over two successive periods.

2.3 Customers' Evaluations and Choice Switching

To analyze choice switching, we introduce some notations. Let the set of individuals be I and the set of review information be $\mathfrak{R}$. The review information consists of the set of hairdressers denoted by S and the evaluations $E = \{1, 2, 3, 4, 5\}$, that is $\mathfrak{R} = S \times E$. Note that each customer can write an online review anytime and that each is given a time stamp in the review information.

To examine customer switching between hairdressers, we compare the reviews that each of the individuals posted in two successive periods. A set of review information for individual $i \in I$ is denoted by $\mathfrak{R}_i$. We write the t-th review of individual i as $\mathfrak{r}_t^i \in \mathfrak{R}_i$. We call a pair of individual i's reviews $(\mathfrak{r}_t^i, \mathfrak{r}_{t+1}^i)$ *successive-period information.* We denote the set of all individual i's successive-periods information by $\mathscr{P}^i = \{(\mathfrak{r}_t^i, \mathfrak{r}_{t+1}^i)|\ \mathfrak{r}_t^i, \mathfrak{r}_{t+1}^i \in E \times S\}$. When we restrict the evaluation of the first period in the successive periods to $E' \subset E$, we write $\mathscr{P}^i(E')$, that is $\mathscr{P}^i(E') = \{(\mathfrak{r}_t^i, \mathfrak{r}_{t+1}^i)|\ \mathfrak{r}_t^i \in E' \times S, \mathfrak{r}_{t+1}^i \in E \times S\}$. For example, in focusing on the reviews where the evaluations consist of 1 and 2, we write $\mathscr{P}^i(\{1, 2\})$. We denote all the successive information restricted on E by $\mathscr{P}(E) = \cup_{i \in I} \mathscr{P}^i(E)$.

First, we examine how each customer switches his or her choice, taking into account the evaluations. We denote the set of successive-period information for customers who switch hairdressers by $\mathscr{P}^h(E)$. Table 2 reports the numbers for each element of the data. We regard a good evaluation as one where the ranking is higher than 3 and a bad evaluation as one where it is lower than 3. An evaluation of 3 is regarded as neutral. As shown in the table, although less than half of all the consumers chose the different hairdressers in successive periods, around 90 percent of those consumers who provided a bad evaluation did then switch hairdressers in the next period.

Table 2. Descriptive statistics of the consecutive data

		$\lvert\mathscr{P}(E')\rvert$	$\lvert\mathscr{P}^h(E')\rvert$
Evaluation	$E' = \{5\}$	15366	4605 (30.0%)
	$E' = \{4\}$	3697	1670 (45.2%)
	$E' = \{3\}$	335	266 (79.4%)
	$E' = \{2\}$	141	134 (95.0%)
	$E' = \{1\}$	55	54 (98.2%)
Total	$E = \{1, 2, 3, 4, 5\}$	19594	6729(34.3%)

Note: The percentages show the proportion of hairdressers who were affected by consumer switching.

3 Relationships between Evaluations and Switching

As our hypothesis is that the customers choose a different type of hairdresser after having a bad experience with the hairdresser previously chosen, we focus on the data for which a different hairdresser is chosen in each of the two periods analyzed. Following the CBDT, we estimate a similarity function s. If our hypothesis is valid, the estimated similarity for the information involving bad evaluations might be lower than that for the information involving better evaluations.

To estimate similarity, we use topic analysis for the *documents* with the hairdresser information and characterize them based on topics. We evaluate the difference/similarity of the hairdressers that the customers choose based on the estimated similarity function. The construction of similarity is illustrated in Figure 1.

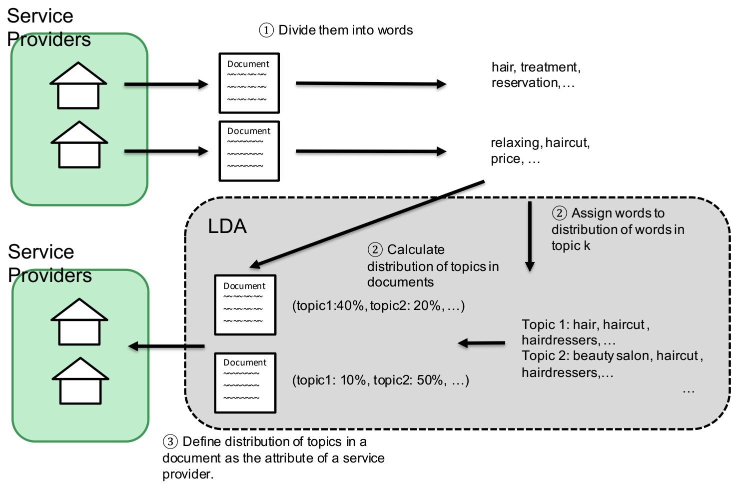

Fig. 1. The process of assigning attributes of documents using latent Dirichlet allocation.

3.1 Construction of Similarity Values by Topic Analysis

In our data set, there are no numerical attribute data. Thus, we use *documents* for the hairdressers, and derive their attributes using topic analysis. We use *latent Dirichlet allocation* (LDA), which is a probabilistic model of topic analysis [1]. Variables are estimated by Bayes estimation. Each topic is defined by words that appear in the text.

We used mecab-ipadic-NEologd as a dictionary to extract Japanese texts with a space between words [10]. For text preprocessing, we normalized texts and excluded stop words. In addition, we excluded alphabets, marks, and numbers to extract Japanese text only and we excluded geographical nouns, which are not proper attributes. Then, we used only nouns and adjectives[3] to derive the characteristics of hairdressers.

Recall that there is a set of hairdressers S. Each hairdresser $a \in S$ has a *document* d_a. Each d_a has a finite set of words $W_{d_a} = \{w_1, w_2, \dots w_n\}$. In the LDA, using W_{d_a}, the topic distribution $\theta_{d_a} = (\theta_{d_a 1}, ..., \theta_{d_a K})$ of d_a is an estimated probability distribution over K topics. Each topic has particular words that characterize it. The word distribution indicates which words belong to each topic.

In this research, we define similarity based on the topic distribution θ_{d_a} of hairdresser a's document d_a. In the prior empirical research based on CBDT, there are two definitions of similarity. Gayer et al. [2] define similarity as the reciprocal of distance and Ossadnik et al. [7] define it by the number of the same features objectives. In our paper, we adopt the former definition because we use continuous variables. That is, for $a, a' \in S$, we define similarity $s(a, a')$ by $\theta_{d_a}, \theta_{d'_a}$ as follows:

$$s(a, a') = \frac{1}{1 + \sqrt{\sum_{k=1}^{K} (\theta_{d_a k} - \theta_{d'_a k})^2}}. \tag{1}$$

We used the Gensim package in Python to estimate posterior topic distributions [9]. Considering preliminary experiments, we set the number[4] of topics to 15. We conducted 10 repetitive estimations of topics and checked each result, given that the result of the estimation changes every time.

3.2 Estimating Topic Distributions and Similarities

First, we extracted words in topics regarding hairdressers from the data. Table 3 provides an example of the results from the retrieval. It shows the top five relevant words in each topic regarding hairdressers. We translated each Japanese word to English[5].

We observed that some topics are difficult to understand and distinguish from other topics. This is because we use the same genre of documents on hairdressers,

[3] In our topic analysis, we did not consider the relationship between two successive words (e.g., "very good"). To avoid mistaking the effect of these relationships, we only checked nouns and adjectives.

[4] We checked the relationship between the number of topics and the coherence value [6], which enabled us to find appropriate number of topics, given our understanding of the topics. However, the relationship showed that it was difficult to determine the optimal number of topics. Therefore, we checked the estimation results for the number of topics between 5 and 50 and decided to use 15 topics because the result appeared to be acceptable.

[5] We have excluded unnecessary words in preprocessing to the extent possible, but some remain owing to our limitations (e.g. 彡).

Table 3. Top five relevant words for each topic regarding hairdressers

topic1	topic2	topic3	topic4	topic5
color	style	once-in-a-lifetime meeting	nail	lasting
extension	proposal	lifetime	eyelash	short bob cut
color	hair	short bob cut	extension	limited
highlight	cut	information	natural feminine	optimization
strength	customer	price	sweet	debut
topic6	**topic7**	**topic8**	**topic9**	**topic10**
hair	style	kindness	waiting	costumer
customer	customer	彡	come to a store	Mr,Ms
proposal	together	transfer	massage	movement
style	enjoyment	reception	shampoo	cute
cut	trouble	maternity leave	technique	casual
topic11	**topic12**	**topic13**	**topic14**	**topic15**
high tone	strength	beauty	shrink	qualification
produce	price	dressing	correction	O.K.
simple	a lot	master	hair	do one's best
special	love	contest	mother	personal
nuanced curly	after	hair set	please	please

whereas topic analysis is usually used for documents, such as news articles, that cover various topics. However, we made use of the results without any adjustment.

Next, we estimate average similarity values for successive-period information that relates to hairdressers who experienced customer switching. We estimate similarity values for those with $\mathscr{P}^h(\{1,2\})$ and $\mathscr{P}^h(\{4,5\})$, and examine how experiences with bad evaluations affected service switching behavior using a Mann-Whitney U test.

Table 4 describes the results of 10 repetitive estimations. It shows that the estimated word distributions of estimations 4, 6, and 10 have significant differences at $\alpha < .01$, but most of the cases have no significance. That is, our comparisons did not show clear evidence of service switching by consumers to a low similarity service following a bad experience. We discuss these results further in the next section.

3.3 Results and Discussion

Our estimations show little clear evidence of service switching from the perspective of CBDT. That is, after a bad experience, consumers do not choose a less similar service next time. In other words, bad service experiences do not affect the service choice as far as similarity of services is concerned. Based on the analysis, we provide the following concluding remarks. We focus on the following four aspects of our analysis: the data problem, the similarity function problem, selection bias, and the fact that we excluded those customers who chose the same service in two successive periods from our data set.

Table 4. Comparing similarity of $\mathscr{P}^h(\{4,5\})$ and $\mathscr{P}^h(\{1,2\})$.

Estimation No.	Average Similarity of $\mathscr{P}^h(\{4,5\})$	Average Similarity of $\mathscr{P}^h(\{1,2\})$	U statistic	p value
1	0.656	0.659	581135	0.365
2	0.630	0.622	562666	0.140
3	0.655	0.652	580945	0.366
4	0.615	0.586	494869	0.000***
5	0.617	0.605	550457	0.059
6	0.612	0.595	542034	0.028*
7	0.677	0.664	559216	0.111
8	0.658	0.646	558010	0.102
9	0.654	0.645	563516	0.148
10	0.604	0.585	529149	0.008**

* Significant at $\alpha < .05$** Significant at $\alpha < .01$
*** Significant at $\alpha < .001$

First, regarding the data problem, we used review data for our consumer behavior analysis that did not comprise data on the whole usage of the service or satisfaction data which happened in the real world. In CBDT, the summation of the past experiences of an individual is considered. If there is a very large effect arising from these prior experiences, which we did not consider in our analysis, then the posterior experience that we could observe from the review data will have little effect on decision-making. In future research, we could add three or more successive sets of review data to examine whether the effect of the evaluations on service switching behavior changes.

Regarding the problem of estimating similarity, it may not be appropriate to use only topic distributions to form similarity. The estimated topics may differ from the impression that consumers obtain when they read those messages. It is possible that consumers not only see text messages, but also images, prices, menus, and other features. For future research, we may use other variables in combination with topic distribution to form our estimates of similarity.

In addition, there may be a problem of selection bias in that those customers who write bad reviews may differ from those who only write good reviews. For example, customers who write bad reviews may actually be satisfied with a service but have a policy of writing bad reviews to prompt service providers to improve their service further. If this is the case, CBDT decision criteria will suggest a similar service because the customers are satisfied.

Finally, a problem may arise because we excluded those who chose the same service for two consecutive periods and used only data for switching consumers. However, CBDT does not exclude the possibility that consumers will choose the same service. Thus, to conduct a proper analysis, we need to obtain data for choice situations where consumers cannot choose the same service providers in consecutive periods.

4 Conclusion

In this research, we focused on service switching behavior after consumers experienced bad services, and analyzed the similarities of the hairdresser services chosen by consumers. We conducted a topic analysis on marketing messages written by the hairdressers to determine similarities. Our results indicated that there were few significant relationships between the evaluations and similarities.

In future research, we will examine service switching behavior using estimations of similarities on a more complete data set and analyze the construction of similarity using additional variables.

References

1. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation David. J. Mach. Learn. Res. 3, 993-1022 (2003). https://doi.org/10.1162/jmlr.2003.3.4-5.993.
2. Gayer, G., Gilboa, I., Lieberman, O.: Rule-Based and Case-Based Reasoning in Housing Prices. B.E. J. Theor. Econ. 7, (2007). https://doi.org/10.2202/1935-1704.1284
3. Gilboa, I., Schmeidler, D.: Case-Based Decision Theory. Q. J. Econ. 110, 605-639 (1995).
4. Gilboa, I., Schmeidler, D.: A Theory of Case-Based Decisions. Cambridge University Press. (2001).
5. Keaveney, S.M.: Customer Behavior in Services Industries An Exploratory Study. J. Mark. 59, 71-82 (1995).
6. Mimno, D., Wallach, H.M., Talley, E., Leenders, M., McCallum, A.: Optimizing Semantic Coherence in Topic Models. Proc. 2011 Conf. Empir. Methods Nat. Lang. Process. 262-272 (2011).
7. Ossadnik, W., Wilmsmann, D., Niemann, B.: Experimental evidence on case-based decision theory. Theory Decis. 75, 211-232 (2013). https://doi.org/10.1007/s11238-012-9333-4
8. Recruit Technologies Co., Ltd.: Hot Pepper Beauty data. *Informatics Research Data Repository*, National Institute of informatics (2014) (dataset). https://doi.org/10.32130/idr.4.1
9. Řehůřek, R., Sojka, P.: Software Framework for Topic Modelling with Large Corpora. In: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks. pp. 45-50. ELRA, Valletta, Malta (2010).
10. Sato, T., Hashimoto, T., Okumura, M.: Implementation of a word segmentation dictionary called mecab-ipadic-NEologd and study on how to use it effectively for information retrieval. In: Proceedings of the Twenty-three Annual Meeting of the Association for Natural Language Processing. The Association for Natural Language Processing (2017).
11. Savage, L. J.: The Foundations of Statistics. Wiley Publications in Statistics (1954).
12. Schank, R. P.: Explanation Patterns: Understanding Mechanically and Creatively. L. Erlbaum Associates (1986).
13. Smith, A. K., Bolton, R. N., Wagner, J.: A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery. J. Mark. Res. 36, 356 (1999). https://doi.org/10.2307/3152082
14. Thomas, J. S., Blattberg, R. C., Fox, E. J.: Recapturing Lost Customers. J. Mark. Res. 41, 31-45 (2004). https://doi.org/10.1509/jmkr.41.1.31.25086

15. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press (1944).