Consumer-Oriented ReceiptLog Service Platform for Effective Applications

Seiki Tokunaga, Akihiro Okushi, Sachio SAIKI, Shinsuke Matsumoto, Masahide Nakamura

Graduate School of System Informatics, Kobe University

1-1, Rokkodai-cho, Nada, Kobe, Hyogo 657-8501, Japan

Email:{tokunaga,okushi}@ws.cs.kobe-u.ac.jp, sachio@carp.kobe-u.ac.jp, {shinsuke, masa-n}@cs.kobe-u.ac.jp

Abstract—Although purchase history has been used by enterprise, it also contains valuable information for consumers. This paper develops *Receipt Log Service Platform*, where every user can store the purchase history from own receipts (we call *Receipt Log*), and can use the data for various consumer services. The proposed platform consists of three components: *receipt scanner*, *ReceiptLog DB* and *ReceiptLog API*. The receipt scanner digitizes daily receipts, and the ReceiptLog DB manages the scanned data. The ReceiptLog API provides the Receipt Log as a service. The API consists of the basicAPI providing fundamental access for the Receipt Log, whereas the miningAPI performing statistical analysis over the Receipt Log. These APIs are published as Web services, which can used by multiple applications and services for various purposes. also implement two applications using the proposed platform. The one is "Sma-sho", which supports user's shopping with receipt log. The other is "ReciLog", which shares receipts within community. For each of the applications, we conduct experimental evaluation with actual subjects, in order to see the usefulness of services with Receipt Log.

Keywords—*receiptlog, Web services, database, social network service*

I. INTRODUCTION

The purchase history of consumers has been widely used by many companies to investigate and determine their sales strategy [2], [3]. By analyzing enormous quantity of the purchase history, the companies can grasp useful information such as hot selling items, preference of the consumers, and cross-selling items. Based on the information, the companies can take efficient strategies in optimizing stocks as well as making attractive goods arrangement. Thus, the purchase history has been a key for the companies not only to increase the profits, but to improve consumer satisfactions.

On the other hand, we consider that the purchase history should not be limited to the business use only. The history is also useful for the *consumers themselves*. A user can record the purchase history from his daily receipts, which we call *receipt log*. By collecting sufficient amount of the receipt log, the user can review his purchase history and may find economic information such as the bedrock price of certain goods. Integrating the receipt log with other data sources can yield more sophisticated information. For example, by using the body metrics (weight, fat, etc.) with the receipt log, the user may review the relationship between his health condition and meals.

Traditionally, the daily receipts have been recorded as the *bookkeeping*, typically written in papers/notebooks, or managed by stand-alone applications. However, the primary purpose of the bookkeeping is to grasp the daily income and outgo, which covers the financial aspect only. Beyond the conventional bookkeeping, our goal is to extend the receipt log for many other value-added services, including daily-life improvements, health checking, social networks services, etc.

The goal of this paper is to provide a consumer-oriented receipt lifelog service. In this paper, we design and implement service platform (called *ReceiptLogService*) which stores and uses receipt log for consumers. ReceiptLogService is a Web service that allows external applications to store and retrieve the receipt log easily. Our implementation consists of a receipt scanner, a MySQL database, and receipt Web API. The receipt data can be accessed via the API such as `getReceipt()` with parameters like username and date. We developped two types of ReceiptLogAPI. One is basicAPI which enables to find the receipt. The other is miningAPI which supports to analyze receipt statistics. Since the API is invoked by platform-independent and XML-based protocols (i.e., SOAP or REST), it is easy for various applications with different languages to access the API.

Furthermore, we also implement two Web applications, called *ReciLog, Smasho*, using the ReceiptLogService. Using the ReciLog the end-user can review own daily receipts just like blog, and share their receipt to comment each other. Sma-Sho is designed to cope with two typical problems in daily shopping: *duplicate purchase* and *forgotten purchase*. We also conduct a usability experiment with actual subjects, in order to evaluate the advantage and limitation of the ReciLog and Smasho. As the result of experiment, we have confirmed that ReceiptLog is helpful not only at the perspective of economic, but in decision-making of shopping [1].

II. PRELIMINARIES

A. *Purchase History*

Many companies have taken the management and sales strategies based on the purchase history of consumers. In modern systems, the purchase history is collected in database by using the Point of Sales (POS) system. The data items of the purchase history generally include date and time of purchase, age of the user, location, and so on. The RFM analysis [3] is a well-known method that finds good consumers. The analysis determines the rank of consumers with respect to **recency, frequency and monetary**. Companies use these data

---

[1]This paper, we have marshaled our research group effort [6] [7] [5]. Because of the number of page's limitation, we have skipped some technical details. Please refer above papers to grasp the technical detail.

as sale strategy. For example the company finds potentially good consumers, who are the target to send catalogue and direct mail. These analysis are methods to use purchase history for business use, so they aren't able to use for consumer themselves. Consumer needs platform to record their receipt by oneself and service which is useful for consumer by using daily receipt, so that consumer use their purchase history.

B. Using Receipt as Lifelog

The *lifelog* [4] is a social activity to record various human life in terms of events, status, and relationships. The recorded log is used to look back and improve daily life, to discover the identity, to determine actions, etc. Recently, a variety of *lifelog services* appear in the Internet. Concretely, tweet, a location, weight, beat, steps and so on. On the internet, we can access wide variety of lifelog service via API.

On the other hand, receipt has many information. For example, it has shop name, product name, total, time which a consumer bought something. So, at the perspective of consumers, the receipt provides valuable information. Concretely, the shop name in the receipt shows the consumer have bought something. And the products in the receipt also show the consumer's habit. Using the total information, consumers could review their spending. Nowadays, some applications provides which store the receipt log. However, they cope to support to keep household accounts. Thus, we still require to implement the platform which enables some potential that the receipt has.

C. Reserch Goal and Approach

The goal of this paper is to provide a benefit of recording daily receipts for consumers themselves. Recording receipts may provide gaining awareness of economic for consumers. Moreover consumer may look back his/her health status because most receipts include meal information.

To meet this goal, we design and implement a service platform, called *ReceiptLogService*, for storing receipt log. We also develop effective applications for consumers. Moreover, we also conduct an experimental evaluation that which perspective the consumer feel useful for reviewing receipts.

III. RECEIPTLOGSERVICE: SERVICE PLATFORM FOR RECORDING DAILY RECEIPTS

A. Platform Overview

We consider that recording the daily receipts by consumers themselves can be a consumer-oriented lifelog service (*single-self*) that reflects their purchase history. Moreover, integrating the receipt log with other data sources (e.g., pictures, body measurements, etc.) may implement sophisticated services (*single-share* or *multiple-share*).

To facilitate the development of such lifelog services, we design and implement a service platform, called *Receipt-LogService*, for storing and retrieving data of the daily receipts. Figure 1 shows the architecture of ReceiptLogService. First, each user digitizes daily receipts by using a receipt scanner [1]. Then, data items such as a shop name, purchase date and time, a list of items, are extracted and stored in a in a global database, called *ReceiptLogDB*. The data can be accessed from external applications via *ReceiptLogAPI*. The ReceiptLogAPI

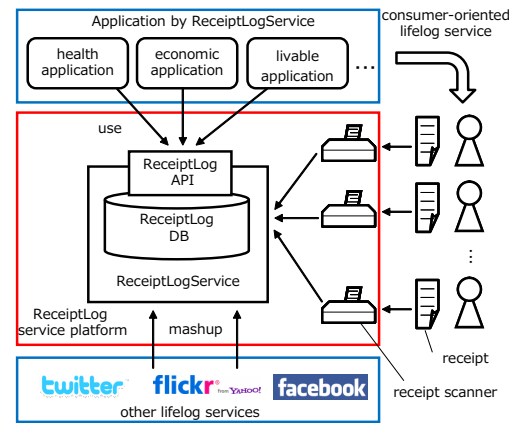Fig. 1.   ReceiptLogService architecture

```
<return type="Receipt">
    <UserID>tokunaga</userID>
    <ReceiptID>r0029</ReceiptID>
    <date>2010-05-25</date>
    <time>13:13:00</time>
    <totalMoney>225</totalMoney>
    <shopName>7-Eleven</shopName>
    <address>3-1, Rokko, Kobe, Japan</address>
</return>
```

Fig. 2.   Result of `getReceipt(tokunaga, 2010-05-25)`

has been implemented as Web service, so that the receipt log can be accessed from various platforms.

B. Receipt Scanner

As a receipt scanner, we use a commercial product *Yasasiku Kakeibo* [1] which includes a receipt scanning device and receipt data management software developed by Media Drive Corporation. First, *Yasasiku Kakeibo* reads receipt image using a scanning device. Then, receipt data items such as a shop name, purchase date, time, and a list of bought items, are extracted from the receipt image by OCR (Optical Character Reader). Each receipt data item are stored to ReceiptLogDB. Since the receipt scanning takes about 12 seconds for each receipt, user can easily records their receipt data compared with handwriting.

C. ReceiptLogDB

Receipt data items, scanned by the receipt scanner, are stored to ReceiptLogDB. By storing the receipt data in the shared ReceiptLogDB, these data are shared with multiple users. We designed and implemented the ReceiptLogDB by following two steps.

Design schema of ReceiptLogDB

ReceiptLogDB has five tables. Figure 3 depicts an ER diagram of the ReceiptLogDB. A box represents an entity (i.e., table), consisting of an entity name, a primary key and attributes. Database instances are shown at bottom of each box. We enumerate instances below each entity to support understanding. A line represents a relationship between entities, where +—∈ denotes a parent-children relationship, and

| user | UserID | name, height, weight, birthday |
| | tokunaga | Seiki Tokunaga 171.0 66.0 1987-04-16 |

receipt	UserID+ReceiptID	date, time, total money, shop name, address
	tokunaga r0029	2010-05-25 13:13:00 JPY650 7-Eleven 3-1,kobe..
	tokunaga r0030	2010-05-25 18:30:00 JPY225 McDonald 3-2,kobe..

itemDetail	UserID+ReceiptID+ItemDetailID	product name, price, CategoryID
	tokunaga r0030 i001	Filet-O-Fish JPY120 ca01
	tokunaga r0030 i002	Diet coke JPY105 ca01

comment	UserID+ReceiptID+CommentID	date, time, UserID, body
	tokunaga r0030 c001	2010-07-01 13:21 tokunaga "It's good"
	tokunaga r0030 c002	2010-07-03 14:15 okamura "feel good"

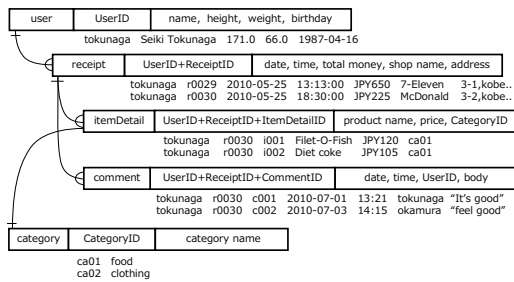| category | CategoryID | category name |
| | ca01 | food |
| | ca02 | clothing |

Fig. 3.   Data schema of ReceiptLogDB

+——·· denotes a reference relationship. We show the model which is partly normalized because of space limitations. The five tables are described as follows.

> **User table:** The user table stores personal information for each user. Concretely, the information includes user ID, user name, height, weight and more.
> **Receipt table:** The receipt table stores receipt information for each receipt. One receipt has some breakdowns which are stored in breakdown table as a children-parent relationship. Basically we assume that a query of ReceiptLogDB includes user name frequently. So we set both ReceiptID and UserID as primary keys for improving access performance.
> **Breakdown table:** Breakdown table stores detailed information of each receipt. As the receipt table, we added both ReceiptID and UserID as primary keys. Product categories are defined in category table described in bellow.
> **Category table:** Category table stores product categories such as food, clothing, daily necessities, hobby and so on.
> **Comment table:** This table stores user comments made by other users. User can make comments to each receipt with each other like a social network service. Recorded information are user name, date, comment body.

D. BasicAPI

BasicAPI allows access to the ReceiptLogDB by simple and abstract API invocation for ReceiptLogService users. Examples of ReceiptLogAPIs are follows.

- `getReceipt()` returns a list of receipts of a specified user issued on specified date.

- `getReceiptDetail()` returns details of a specified user and receipt.

- `getUserOutgo()` returns total outgo by specified user during specified period.

- `getUserComments()` returns comment list of a specified user issued on the date and the receipt.

The API can be invoked by standard Web service protocols (i.e., SOAP and REST). Figure 2 shows a REST invocation of `getReceipt()`, where we can see four receipts of user `tokunaga` issued on 2010-05-25.

E. MiningAPI

MiningAPI enables to analyze receipt log statistically. In this paper, we propose the method we analyze receipt log for consumers themselves using RFM analysisII-A. Thus, we analyze the stored receipt log, we analyze **how recency the consumer has bought a specific product, how frequency does the consumer goes a specific shop and how monetary does the consumer spend?** Using RFM analysis for consumers, they could gain much valuable lifelog service. We show the some miningAPI as the examples.

- getRecency()   How recency does the consumer go shop or buy a product?

- getFrequency() How frequency does the consumer go shop or buy the specific product?

- getMonetary()   How monetary does the consumer spends the shop or the product?

- getTimeIntervalReport() The API shows the timeline report using RFM

- getTrend   What is the trend for a user?

We introduce about the interface of `getRecency()` as the example of miningAPI.

```
getRecency(userid, target, cond, limit)
 -params
    userid
    target:{productshoparea, pricetotal}
    cond
    limi
 -return
    Receipt[]
```

We designed this API so that the API could access the latest receipt log with various perspective items. This API has targeted information which has five types of information, product name, shop name, area, price and total. So, we search the recent receipt log by above five information. Using cond parameter, we can limit of the return value of API. The API respond to the array of receipt object so that the API-user could grasp some of the recent receipt.

We have also implemented `getFrequency()` and `getMonetary()`.They are similar to the `getRecency()`. With the `getFrequency()`, we can grasp the frequency of going shop, or frequency of product the user buys. We have skipped the detail spec because of the page's number of limitation.

IV.   EFFECTIVE RECEIPTLOG APPLICATIONS

Using ReceiptLogService we could effectively develop the receipt log applications. In this section, we introduce the two effective applications "ReciLog" and "Smasho" using ReceiptLogService.

A. Social Network Application: ReciLog

First, we develop a social network web application, called *ReciLog*, as a client application of ReceiptLogService. The ReciLog is constructed by the ReceiptLogAPI and has three

main features. The main feature is the *reviewing feature*, which allows the consumers to review daily receipts just like the blog. So we can recall receipt log easily as if we look back on the diary. Moreover, the ReciLog has the *sharing feature*, with which the users can exchange comments and recommendations about goods and purchase. Also, the ReciLog provides the *reporting feature*, which summarizes the statistics of the purchase using graphs and tables. We develop ReciLog's main using basicAPI. For example, the content of receipt is composed with `getReceipt()` and Figure 4(a), (b) shows screenshots of the Recilog. We also developed the function of analyzing receipt log. Figure 4 shows the screenshots of ReciLog's analyzing screen. Concretely, screen (a) shows the reviewing feature with which the user could look back their receipt log. The reviewing feature helps a user to review the daily receipts easily. The user can add comments to each receipt, to describe the explanation of the goods, the motivation of the purchase, and the impression and so on. ReciLog has a login screen, where a user logins the service. Once the login is succeeded, the blog-like reviewing screen appears (Figure 4). In this screen, the user can review the latest receipts and add some comments. By clicking the calendar at the left side of screen, the user can move back to the older receipts issued on the specified date. RecciLog's analyzing function is developed with miningAPI. For example, the reviewing user's total payment are obtained by `getMonetary()`. And the ranking of products are obtained by `getTrend()`.

We also developed other feature, for example a favorite shop ranking feature, finding the recently product feature, analyzing the time interval cost feature and showing frequency product feature. Because of the page's number limitation we can't show all of them. Please refer [7] for detail.

B. Shopping surppot appliation "Smasho"

"Smasho" is a smart shopping support service using the receipt log. Sma-Sho is designed to cope with two typical problems in daily shopping: *duplicate purchase* and *forgotten purchase*. The duplicate purchase is that a user mistakenly buys a product although there is sufficient stock of the same product at home. The forgotten purchase is that a user forgets to buy a product although the product is out of stock. Using the miningAPIs extensively, Sma-Sho provides useful purchase histories for the user with a mobile terminal, to prevent the above two problems.

Screenshots of Sma-Sho are shown in Figure 5. A user is assumed to use Sma-Sho with a smart phone during shopping. **"When did I buy it?" screen:** shown in Figure 5(a). Using this screen, a user can prevents duplicate purchase caused by Pattern 1. The user inputs product name and presses search button. Then, Sma-sho displays the latest date with a shop name when the user bought the product. If the user bought the same product more than once, the user can search older ones by pressing next button. The user search the purchase history of "wasabi". Sma-sho responses that "okushi" bought a wasabi on 5 February 2012 at "Hankyu Oasis, Mikage" store. Thus, the user can decide whether or not he should buy wasabi now.

"Is there sufficient stock?" screen: The above algorithm is implemented in the screen shown in Figure 5(c). By using this screen, a user can prevent the forgotten purchase. When a
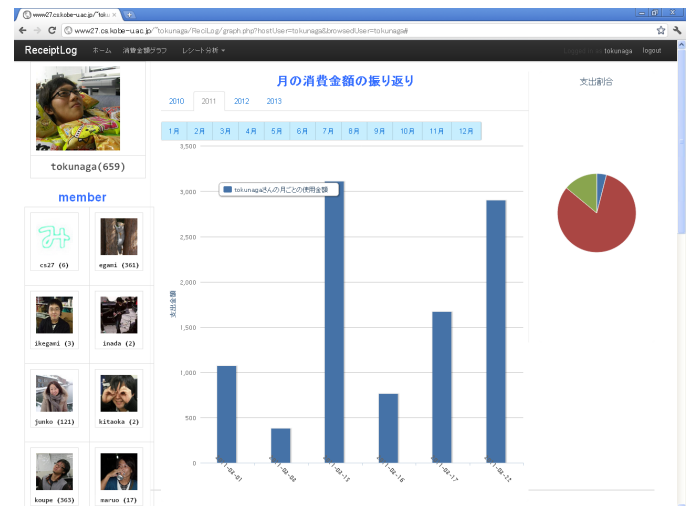
(a) reviewing user's total payment

(b) ranking of products the user bought

Fig. 4. Screenshots of Analyzing Feature of ReciLog

user presses the check button, Sma-Sho checks if each of the pre-set products is in stock or not using the algorithm. If the user presses the out-of-stock button, then Sma-Sho displays only products that are out of stock. If the user wants to know the purchase date of a product, the user selects the product from pull-down menu and presses search button.

Figure 5(c) shows a screen where Sma-Sho tells user "okushi" that tissues are likely to run out of stock.

V. EXPERIMENTAL EVALUATION

A. Experiment Overview

In order to evaluate the effective of reviewing the receipt log for users, we have conducted two experiments. Concletely, we had some questionnaires about using the applications which were introduced. The total number of receipts registered in the RecieptLogService was 1,615, which had been collected for 22 months.

B. Experimental Evaluation of ReciLog

In order to evaluate the ReciLog, we have conducted an experiment, where subjects actually use the ReciLog to review

TABLE I.   QUESTION ABOUT RECALL BY ONESELF

Pro/Cons	comments
+	I could review monthly expenditure, which I did not care before.
+	It is interesting to figure out my trend and boom.
+	It is convenient to recall when, where and for what I spent money.
+	I realized eating cost much more than I expected.
+	I figured out how I haven't been eating enough vegetables.
−	I want more detailed statistics. I cannot figure out from such simple graphs.
−	I could not figure out nutritional balance for a whole day, because I had only lunch receipts.

(a) "when did I buy it?"

(b) "what did I recently buy?"

(c) "Is there sufficient stock?"
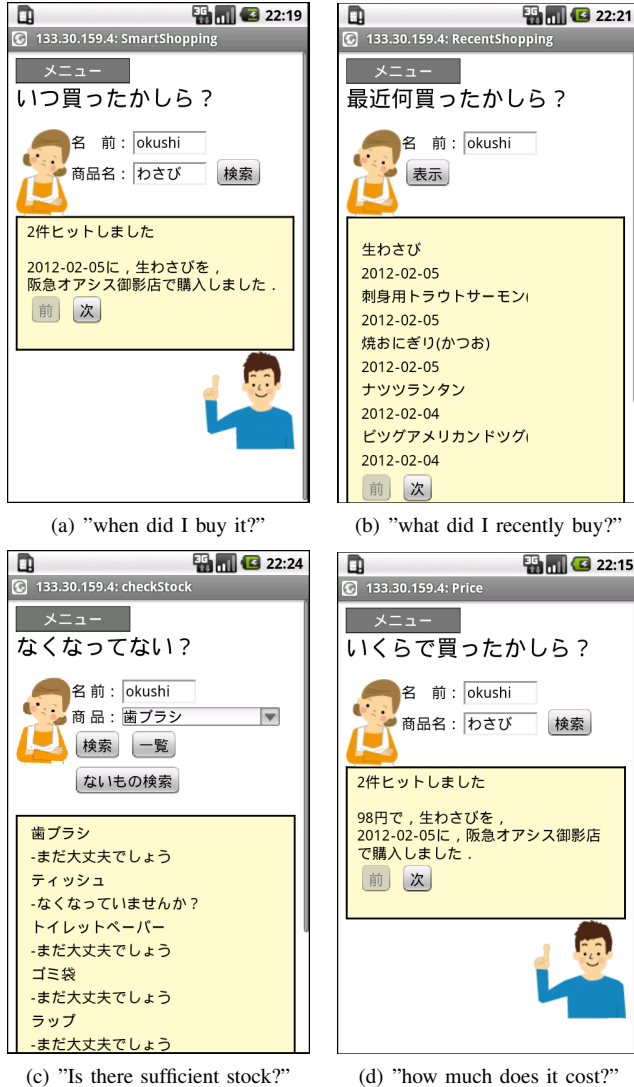
(d) "how much does it cost?"

Fig. 5.   Screenshots of Sma-Sho

their own receipt log. In this experiment, we aim to evaluate to confirm the receipt log is effective at view point of economic and healthy. The subjects are all boy students who are twenty-something. The questionnaires are asked at the viewpoints of economic and healthy.

In the experiment, we gave two types of tasks to the subjects. The first task (Task 1) is to evaluate the *self* aspect of the ReciLog, where each subject reviews own receipt log from economic and health aspects. The second task (Task 2) is to see the *share* aspect, where the subjects share the receipt log within the community.

1)   Please answer below questionnaire at the view point of your economic.

TABLE II.   EXPERIMENTAL RESULT OF RECILOG

evaluation	economic		healthy	
	task1	task2	task1	task2
5 (useful)	4	2	2	0
4	8	4	5	1
3	0	2	3	6
2	1	2	2	5
1 (unuseful)	0	3	1	1

- Is it interesting to review your receipt log.
- Is it useful to review others's receiptlog?

2)   Please answer below questionnaire at the view point of your healthy.

- Is it useful to review your receipt log.
- Is it useful to review others's receiptlog?

We also added the items of questionnaire that the subjects answer their opinion as the free comment.

The table II shows the experimental result. As the the result of economic task, twelve subjects answered that receipt log was very useful or useful. Concretely, one subject answered that "I can review the difference between my sense and real payment", and other said "I was enable to grasp that what product I had paid most". However, the evaluation of task2 was different from each subject. Remember that the task2 was the task to review other's receipt log. As the positive comments, some subjects commented that "It is useful to compare my outgo to other people's outgo in same generation. On the other hands, as the negative comments "Because of the lack of recommendation, it was too difficult to review other's receipt log".

About the reviewing the receipt log for healthy, each task1 and task2 were 3.28, 2.53 as the score which were lower than economic aspect. As the positive comments, some subjects answered "I could review own via receipt log and I noticed that I had unbalanced diet". And the subject who likes a luxury said that "I have much spent the luxury, so I have to cut it" On the other hand, the subject commented that "It is difficult for me to judge good food or bad food with reviewing other's receipt log."

C. Experimental Evaluation of Smasho

In order to evaluate Sma-Sho, we have conducted a user experiment. Eight master course students participated as subjects of the experiment. Before the experiment, they had already stored their receipts in ReceiptLog Service for one or two years. In the experiment, we asked the subjects to use Sma-Sho as they wanted. After they used Sma-Sho, we collected a questionnaire from every subject. The questionnaire consisted of the following questions. For each of the duplicate purchase and the forgotten purchase, two sets of the same questions were asked conducted.

TABLE III. Is Sma-Sho useful to prevent duplicate purchase / forgotten purchase?(1,2,3,4,5)

Score	duplicate purchase	forgotten purchase
5 (useful)	3	2
4	2	5
3	2	1
2	1	0
1 (unuseful)	0	0

■preventing duplicate purchase
    positive
        ·It is good to use Sma-Sho by mobile terminal in shopping.
        ·It is useful fot preventing duplicate purchase of product I don't frequently buy.
    point needed to improve
        ·Some products I bought didn't hit.
        ·I want the funtion preventing duplicate purcahse dedicated to foods.
■preventing forgotten purchase
    positive
        ·It is good for preventing fotgotten purchase to previously decide to display what product.
        point needed to improve.
        ·I want to receive not only judgement by Sma-Sho but also the latest date I bought the product.
        ·I want to decide list of products and consumption cycle.
        ·I want to be automatically noticed forgotten purchase from application.

Fig. 6. Comment to Smasho

D. Sma-sho's Experimental Results

- Is the Sma-sho's function is useful?
- When the scenes do you think the Sma-sho is useful?
- What is the weak point of Sma-sho?
- Please give some idea to request for Sma-sho.

The results of questionnaire are shown in Table III. The average score for the prevention of duplicate purchase was 3.8. Overall, we got positive evaluation. A subject, who marked low score (2), said "I don't have much experience of duplicate purchase". The positive comments include that; "it is convenient to use it during shopping", "it is useful for preventing duplicate purchase of what I don't buy frequently", etc. On the other hands, a subject complaint that he could not get expected search result.

The average score for the prevention of the forgotten purchase was 4.1. We received higher score than that of the duplicate purchase, which reflects that Sma-Sho is particularly useful in that context. Subjects proposed some improvement of Sma-sho, including; "I want to customize the products and their consumption cycles by myself", and "I want a feature that automatically alerts the out of stock."

E. Discussion

With the results of two experiments, we found that the receipt log was effective for reviewing their lifestyles and their decision for shopping. In other words, the user could respond next questions with receipt log. **When** did they buy some products, **where** did they buy the products and bf how much did they spend the products? Above informations are much effective methods for reviewing at economic aspect. Also, the developed miningAPI are helpful for applications.

However, the evaluation of sharing the receipt log became vary for each user. This was because the subjects couldn't understand other's receipt log just as some subjects said. "We think that some subjects responded that sharing receipt log was interesting." So, we would like to extend sharing feature as to become more interesting in reviewing other's receipt log. Concretely, we try to add new functions such as a recommendation and so on.

Finally, we discuss about the validity of evaluation. In this research, we had the limitation of resource and time, so we have requested to be subject for only college students. Thus we could not evaluate that the receipt log is efficient for other younger or older generation. Our future work is to conduct more wide variety of subjects. And we also would like to find that what kind of information in the receipt log is effect for some generations. Moreover according to various user's comment. we would like to extend ReceiptLogService platform and their applications.

VI. CONCLUSION

In this paper, we have developed a consumer-oriented ReceiptLogService platform. Concretely, we have design and implemented a service platform called ReceiptLogService which provides basic APIs for recording and reading a receipt log. Next, we also developed a social network Web application, called ReciLog and smart shopping web application Smasho. They are developed as a client Web application of the service platform, ReceiptLogService.

We then conducted an experimental evaluation of the ReciLog and Smasho with actual subjects. We confirmed the effective of receipt log for the consumers for some aspects. As the result of experiment, we have confirmed that the receipt log was useful in some aspects (e.g, economic).

REFERENCES

[1] Corp. MediaDrive. Yasasiku kakeibo. http://mediadrive.jp/.

[2] B. Davies and S. Worrall. Basket analysis: profiling british customers. *British Food Journal*, 100:102 – 109, 1998.

[3] R. Jackson and P. Wang. *Strategic Database Marketing*. McGraw-Hill, 1 edition, 7 1994.

[4] O. H. Kieron, T. Mischa, and S. Nigel. Lifelogging: Privacy and empowerment with memories for life. *Identity in the Information Society*, pages 155–172, June 2008.

[5] A. Okushi, S. Tokunaga, S. Matsumoto, and M. Nakamura. Sma-sho: Implementation and evaluation of a shopping support service using receipt log. In *Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT2012)*, pages 1–6, November 2012.

[6] S. Tokunaga, S. Matsumoto, and M. Nakamura. Implementation and evaluation of consumer-oriented lifelog service using daily receipts. In *The 13th International Conference on Information Integration and Web-based Applications & Services (iiWAS2011)*, pages 337–340, December 2011.

[7] S. Tokunaga, S. Matsumoto, and M. Nakamura. Implementation of consumer-oriented receiptlog mining service based on rfm. In *Technical Report of IEICE*, volume 111, pages 189–194, 2012.