Evaluation of Credibility for Reviewers and Review Scores Based on Link Analysis

Yasuyuki Shirai

***Abstract*—In recent years, there are widely used online stores on the Internet which have customer review facilities. Although these reviews help users decide their actions, there are incredible, fake or irresponsible reviews in general. It is difficult for site users and site managers to evaluate the credibility of reviews and reviewers. In this paper, we show the method to quantify the credibility of review scores and to find out credible reviewers and credible review scores based on a bipartite relation between reviewers and sale items. We also show the experimental results on the travel website. Through our experiments, we discuss these indexes for reviewers and sale items are also effective for the future prediction of review scores.**

***Index Terms*—Credibility, user review, travel website, HITS algorithm.**

I. Introduction

In recent years, there are widely used online stores on the Internet, such as bookselling sites, travel websites and daily necessaries selling sites. One of the significant characteristics of these sites is a user review facility for each sale item. User reviews are generally useful because they reflect actual impressions by the reviewers, rather than directional advertisements from sellers. For example, Rakuten Travel [1] has over 10 million reviews on the Internet which have helpful for first time users as well as regular customers.

However, some of the user reviews are incredible, because there may be deceptive or setup reviews by persons who concern those sale items in their businesses, or there may be fake (promoting or demoting) or irresponsible evaluations. Although online stores need to detect these incredible reviews, it is hard in general, to distinguish them from other truly useful reviews. In fact, it is difficult to judge the incredibility from the review text.

In this research, we try to evaluate the credibility of hotel scores and reviewers in travel website, based on the analysis of a bipartite graph that consists of sale items, i.e., hotels and reviewers. In our experiments, we use Rakuten Travel data [1], [2], provided by Rakuten Inc. and National Institute of Informatics (NII). Rakuten Travel is one of the major travel websites in Japan. It covers about 30,000 domestic hotels in Japan (2017), and the number of members is 33 million per year [3].

In this paper, we show the method to evaluate the credibility of hotel scores and reviewers based on the data until the end of 2012, and to verify the validity of our method by comparing our evaluation results with actual scores in other data from 2013 to 2015. We found that the hotels which have high scores by the credible reviewers in the data until the end of 2012 got higher scores in the future (2013-2015).

II. Related Work

There are many of researches that focus on the credibility of user reviews on Internet services, which can be broadly divided into two types: one is from the view point of marketing insights, and the other one focuses on the quantitative evaluation techniques.

For example, from the view point of marketing researches, Kusumasondjaja [4] considered the effects of positive reviews and negative reviews for site users through their experiments. They also considered the difference between the cases that the reviewer's identity is disclosed and not disclosed. Mackiewicz [5] and DeAndrea [6] discussed how users construct reviewers' credibility and how to identify credible reviewers who are likely to influence site users. Chakraborty [7] considered the impact of credible online reviews and relationship with brand images. In these researches, they used the criteria for the credibility of reviews or reviewers, and discussed how to use the credibility scores for site users.

As for the effectiveness of reviews, some researches used questionnaire survey on users. Fogg [8] carried out a large-scale investigation which reveals the site characteristics which are effective for credibility on the web sites. Helversen [9] studied the impact of customer ratings and emotional reviews, and studied the difference on online purchase intentions between in older adults and in students.

In our research, we mainly discuss the quantitative evaluation method based on the relationship between sale items and reviewers, rather than the details of the availability for site managers.

As for the quantitative evaluation methods, many of the researches adopted linguistic analysis. Mukherjee [10] used latent topic models to find fake reviews and opinion spam detection. Sharma [11] proposed the algorithm which ranks user reviews using content analysis, and credibility of the content authors. They adopted the use of grammar analysis and sentiment analysis for review sentences. In their analyses, they focused on the relevancy of review sentences and the item characteristics.

The motivation of our research is almost the same as their researches. Although our method does not use linguistic

Manuscript received September 25, 2018; revised June 22, 2019. This work was supported in part by JSPS KAKENHI Grant Numbers JP17K00441 and JP18K01891.

The author is with the Faculty of Business Administration, Daito Bunka University, Tokyo, Japan (e-mail: yasuyuki.shirai@gmail.com).

doi: 10.18178/ijmlc.2019.9.5.858

analysis but use just a relationship represented as a bipartite graph between items and reviews, the linguistics approaches for review sentences could be effective for judgment of credibility. We will consider how to combine our method in this paper with content analysis, especially sentiment analysis as a future work.

III. Overview of the Experimental Data

Rakuten Travel data set at the end of 2015 consists of 899,920 reviewers and 25,033 hotels. We split the data into a data set from 2009 to the end of 2012 (Data-1), and a data set from 2013 to 2015 (Data-2). The former data (Data-1) is used as a training data set and the latter (Data-2) for verification of our results.

In Rakuten Travel, a reviewer can provide his/her evaluation for hotels which he/she used, as ``Guest rating": 1 (very low), 2 (low), 3 (mid), 4 (high) and 5 (very high).

Fig. 1 shows the summary image of user reviews for a hotel on the Rakuten Travel website. This site displays each score such as service, place, room, amenity, bath/spa and meal as well as a total score. In the case of Fig. 1, there are 1,995 reviews in total and the total score average of the hotel is 4.26. In this paper, we will only use total scores each of which is an average value of all reviews.

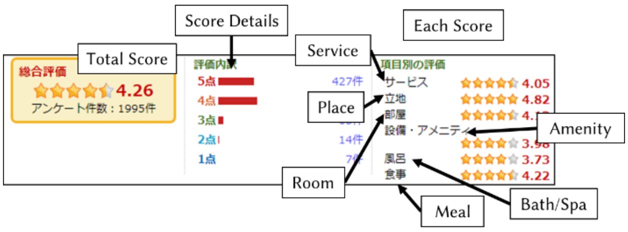


Fig. 1. User review summary.

Table I shows the distribution of total score averages of all hotels in Data-1. As shown in the table, there is strong deviation among the distribution. Especially, almost half of hotels (47.95%) have the scores from 4.0 to 4.5. Low evaluations such as 1 and 2 can be considered as reflecting unfair evaluations, because these evaluations sometimes contain abnormal or deflected claims. These evaluation scores seem to be caused by some special situations. In this sense, strong deviation among the distribution 1 to 5 can be natural and sound in our experiments.

TABLE I: Distribution of Total Score Average

	Hotels (%)
$1 \le E < 2$	0.15
$2 \le E < 3$	1.95
$3 \le E < 4$	33.68
$4 \le E \le 5$	64.22

(E: Total Score Average)

TABLE II: Number of Reviews Top 5 (Hotel)

Hotel ID	#Reviews	Average Score
19455	1412	3.91
74637	1369	4.23
129475	936	4.29
581	878	4.27
2614	853	3.91

Table II shows top 5 hotels which have large number of reviews. The average number of reviews for each hotel is 39.55. There are 1,525 hotels with reviews more than 100. Table III shows top 5 reviewers in terms of the number of posted reviews. The average number of reviews for each reviewer is 8.63. There are 3,653 reviewers with more than 10 reviews. In our experiments, we extracted reviewers and hotels that have more than or equal to 5 reviews because reviewers who have small number of reviews may be just noise in our analysis.

TABLE III: Number of Reviews Top 5 (Reviewer)

User ID	#Reviews	Average Score
user12625	101	4.36
user13848	101	3.66
user175662	101	3.66
user1765	96	4.22
user10730	90	3.96

IV. Computing Credibility

In this research, we consider that the credibility values of reviewers and hotel scores are defined via the mutual relationship between reviewers and hotel scores.

Fig. 2 shows the relationship between (in)credible hotel scores and (in)credible reviewers. Intuitively speaking, credible/incredible hotel scores are evaluated and posted by credible/incredible reviewers, and credible/incredible reviewers posts their scores to credible/incredible hotel scores, respectively. In this way, we consider these relations as a recursive structure by reviewers and hotel scores.

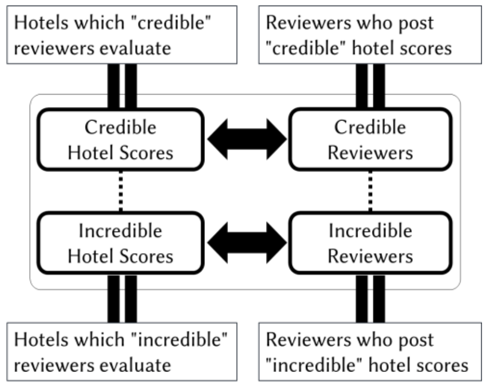


Fig. 2. Relation between (in)credible hotel scores and (in)credible reviewers.

The next subsection describes the detailed method to realize this concept quantitatively based on HITS algorithm [12].

A. Overview of the Method

In this research, we apply PageRank algorithm [13] and HITS algorithm [12] for a bipartite graph that consists of hotel scores and reviewers. HITS algorithm and PageRank algorithm are famous approach to quantify the values of web pages and widely used in search engine applications.

In HITS algorithm, it assumes all web pages are divided into authority pages and hub pages. Good authority pages are linked from good hub pages, and good hub pages link to good authority pages. Based on the recursive definitions, HITS algorithm computes the "goodness" as an eigenvector from the link structure of web pages.

In our previous researches, we extended HITS algorithm, and applied them to the analysis of the relationship between sale items and customers in retail business [14], [15], called MUJI brand, which is one of the famous and characteristic retail shops in Japan. Unlike other retail shops, MUJI has core "MUJI fan". We successfully extracted the enthusiasts for MUJI brand and the characteristic sale items of MUJI brand that are appreciated by the enthusiast users. MUJI enthusiasts tend to buy MUJI characteristic items, and MUJI characteristic items tend to be bought by MUJI enthusiasts. Some food products by MUJI brand have high values although they do not have many sales among all products.

As other applications, we also applied and extended this method to the university brand, Waseda University and Keio University which are most famous and distinctive private universities in Japan [16]. In this research, we extracted the characteristic attributes and the characteristic students of each of universities from the results of questionnaires for over 2,000 students, and analyzed the difference of two universities. We found some common features for both of two universities as well as distinctive features. Common features can be considered as the features of famous private universities in Japan.

In this paper, we extend the methods in [14]-[16], to evaluate the credibility of hotel scores and reviewers, and also try to predict future hotel scores using these indexes.

B. Detailed Methods

Let u_i be a credibility value of reviewer i and let m_j be a credibility value of a score for a hotel j. Our goal is to find a credibility vector of reviewers $\boldsymbol{u} = (u_1, u_2, \dots, u_{n_u})^T$, and credibility vectors of hotel scores $\boldsymbol{m} = (m_1, m_2, \dots, m_{n_m})^T$ where n_u represents the number of reviewers and n_m represents the number of hotels.

If a reviewer i evaluates a hotel j as a score s_{ij} $(= 1,2,\dots,5)$, we define a_{ij} as follows:

$$a_{ij} = log(L_{ij}) \times \frac{1}{(s_{ij} - \bar{s_j})/\sigma_j}$$

where L_{ij} is a comment length of the reviewer i for a hotel j (reviewers usually post their comments along with the scores), and $\bar{s_j}$, σ_j is an average and a standard deviation of scores for a hotel j, respectively. The length of the comment would be a factor of an initial values as the credibility, because a credible score seems to have a sincere and detailed comment with it in many cases. We think that longer comments reflects earnest and sincere attitude of reviewers.

Because $s_{ij} - \bar{s_j}$ is sometimes very small, the maximum value of $\frac{1}{(s_{ij} - \bar{s_j})/\sigma_j}$ is set to 5 considering the distribution of the values. If a reviewer i does not evaluate a hotel j, $a_{ij} = 0$.

We can get the vector $\boldsymbol{u}$ and $\boldsymbol{m}$ through the iterative computation of the following formula after giving the initial values $\boldsymbol{m}^{(0)}$.

In these formula, N shows the repeated number and $u^{(N)} = \left(u_1^{(N)}, u_2^{(N)}, \dots, u_{n_u}^{(N)}\right)^T$, $m^{(N)} = \left(m_1^{(N)}, m_2^{(N)}, \dots, m_{n_m}^{(N)}\right)^T$:

$$\overline{u}_i^{(N+1)} = \sum_j a_{ij}\, m_j^{(N)} \quad (i = 1, \dots, n_u),$$
$$\overline{m}_j^{(N+1)} = \sum_i a_{ij}\, u_i^{(N+1)} \quad (j = 1, \dots, n_m),$$

$$\overline{\boldsymbol{u}}^{(N+1)} = \left(\overline{u}_1^{(N+1)}, \dots, \overline{u}_{n_u}^{(N+1)}\right)^\top,$$
$$\overline{\boldsymbol{m}}^{(N+1)} = \left(\overline{m}_1^{(N+1)}, \dots, \overline{m}_{n_m}^{(N+1)}\right)^\top,$$
$$\boldsymbol{u}^{(N+1)} = \overline{\boldsymbol{u}}^{(N+1)} / \left\|\overline{\boldsymbol{u}}^{(N+1)}\right\|_2,$$
$$\boldsymbol{m}^{(N+1)} = \overline{\boldsymbol{m}}^{(N+1)} / \left\|\overline{\boldsymbol{m}}^{(N+1)}\right\|_2$$

where $\|\cdot\|_2$ shows L_2 norm, that is, each $\boldsymbol{u}^{(N+1)}$, $\boldsymbol{m}^{(N+1)}$ in the above formula is normalized so that the squared sum of all elements is 1.

Now we define a matrix $\boldsymbol{A}$ as :

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n_m} \\ \vdots & \ddots & \vdots \\ a_{n_u 1} & \cdots & a_{n_u n_m} \end{pmatrix}$$

Then, $\boldsymbol{u}^{(N+1)}$, $\boldsymbol{m}^{(N+1)}$ can be represented as

$$\boldsymbol{u}^{(N+1)} = \boldsymbol{A}\boldsymbol{m}^{(N)},\ \boldsymbol{m}^{(N+1)} = \boldsymbol{A}^T\boldsymbol{u}^{(N+1)}.$$

From the above mutual recursive equations, we can get

$$\boldsymbol{u}^{(N+1)} = \boldsymbol{A}\boldsymbol{A}^T u^{(N)}, \qquad \boldsymbol{m}^{(N+1)} = \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{m}^{(N)}.$$

It is well known that $\boldsymbol{u}^N$ and $\boldsymbol{m}^N$ converge to the principal eigenvectors of $\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}$, $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$ respectively. These normalized eigenvectors are denoted by $\boldsymbol{u}, \boldsymbol{m}$ that contain credibility sores for reviewers and hotel scores respectively.

As an example, consider the sample data shown in Table IV, which consists of 5 reviewers and 4 hotels. For example, User 1 posts his/her scores for Hotel 1 and Hotel 3. The values in the table show the initial values for credibility (a_{ij}). That is, if the value is high, it means that the posted score can be considered as a valid score as an initial value, i.e., the posted score is similar to the other scores of the hotel and the comment is long enough so that we can consider it as sincere and detailed one. Conversely, if the value is low, the posted score is apart from other scores of the hotel and does not have detailed comment.

The results for the reviewers' credibility indexes and the hotel scores' credibility indexes calculated by the method described in this section are shown in Table V. As shown in the table, Hotel 1 and Hotel 3 have high credible scores. In fact, User 1,3 and 5 posted to these hotels, and these users are considered as credible reviewers because their evaluation scores are almost similar to other scores. Remark that the score of User 3 is higher than User 1. User 1 and User 3 posts their evaluations for Hotel 1 and 3, Since the credibility score for Hotel 3 is higher than Hotel1, User 3 who posts high credible values for Hotel 3 is more credible than User 1. Additionally User 3 posts for Hotel 4 with low credible values. However, this does not contribute to reducing the credibility values of User 3 because of the credibility values of Hotel 4 is small.

Hotel 2 has two scores by User 2 and 4. If User 2 and 4 are not credible reviewers, credibility values for Hotel 2 are not so high. In fact, the posted score values by User 2 are 0.2 and 0.3, i.e., they are apart from other reviewers' scores and they do not have comments which have enough length. The credibility value of Hotel 4 is larger than Hotel 2, because a credible reviewer (User 5) posted to Hotel 4 while no credible reviewers posted to Hotel 2.

In this way, credibility values of hotel scores and reviewers are determined by the mutual relationship, and they can be calculated as eigenvectors of the matrix $\boldsymbol{AA^T}$ and $\boldsymbol{A^T A}$.

TABLE IV: SAMPLE DATA

	Hotel 1	Hotel 2	Hotel 3	Hotel 4
User 1	0.9	0.0	0.8	0.0
User 2	0.0	0.2	0.0	0.3
User 3	0.8	0.0	0.9	0.6
User 4	0.0	0.9	0.5	0.0
User 5	0.9	0.0	0.9	0.8

TABLE V: RESULTS ON SAMPLE DATA

	Credibility
Hotel 1	0.632
Hotel 2	0.072
Hotel 3	0.674
Hotel 4	0.376

	Credibility
User 1	0.478
User 2	0.055
User 3	0.577
User 4	0.173
User 5	0.637

V. EXPERIMENTS

A. Overview of the Experiments

As shown in Section III, we use Rakuten Travel Data which includes the data from 2009 to 2015. We divided the data into two set, that is, a training data set 2009-2012 (Data-1), and a test data set 2013-2015 (Data-2).

We get 1,537,803 records as a training data set (Data-1) and 1,114,929 records as a test data set (Data-2). Each data record consists of hotel ID, evaluation date, reviewer ID, an evaluation score and a comment with its length. In Data-1, we extracted reviewers and hotels that have more than or equal to 5 reviews and got 66,999 reviewers and 14,272 hotels. As for the length of the comment that is used as a weight value, we just compute the Japanese string length.

Our experiments were run on MATLAB Version 9.1.0.441655 (R2016b), Windows 10 with a Intel Core i3-6100U CPU 2.30 GHz, and 16.0 GB RAM. As for the execution time, it takes just a few seconds for each case.

The results from Data-1 are shown in Section V-B and V-C. Based on the results from Data-1, we try to predict the future scores using these indexes and compare the results with actual data in Data-2 in Section V-D.

B. Results on Reviewers

We calculated a credibility vector $\boldsymbol{u}$ for reviewers and $\boldsymbol{m}$ for hotel scores. We sort these indexes in $\boldsymbol{u}$ and $\boldsymbol{m}$ by descending order and we refer to the order as "reviewer rank" and "hotel rank" in short, respectively.

Table VI shows top 10 reviewers and bottom 10 reviewers in credibility. The values of credibility in the table are 10,000 times of the original values for easiness in reading.

Some reviewers in top 10 have many of reviews such as User2188, User23632 and User23824 but some other users such as User74703, User456451, User325008 have less than 10 reviews. These users do not post many times but their evaluation scores can be considered as credible.

On the other hand, users in bottom 10 posts their scores for incredible hotel scores and their scores are apart from average scores. As a result of our calculation, these users cannot be considered as credible reviewers.

Table VII shows the average rank in credibility of the hotels for which top 3 credible reviewers and bottom 3 (incredible) reviewers posts (The details of the hotel evaluation are shown in the next subsection). There are 14,272 hotels in total. For example, user74703 who is at second rank in reviewer rank, posts his/her scores for 7 hotels. The hotel ranks of these are 4820, 37, 693, 3113, 1, 1896, 1903 (average rank is 1780.4). While the hotel ranks by user62392 who is at second rank from the bottom in reviewer rank, are 11949, 6932, 10084, 5242, 12460, 9920, 7326 and the average rank is 9130.4.

TABLE VI: TOP/BOTTOM 10 REVIEWERS IN CREDIBILITY

Top 10 Reviewers

User ID	Credibility	#Reviews
user23632	826.94	50
user74703	630.14	7
user456451	611.33	5
user325008	597.20	6
user2580	596.56	13
user48353	586.44	22
user2188	567.41	81
user23824	551.34	44
user28387	538.21	7
user280974	534.15	8

Bottom 10 Reviewers

User ID	Credibility	#Reviews
user161845	0.00023	11
user62392	0.00036	7
user23040	0.00042	10
user466522	0.00049	6
user27316	0.00065	6
user309204	0.00083	5
user23841	0.00119	10
user246149	0.00154	8
user26525	0.00171	5
user187503	0.00255	5

We found that credible reviewers posted their scores for hotels which have high-ranking credibility scores, and that incredible reviewers posted their scores for hotels which have low-ranking credibility scores.

C. Results on Hotels

Table VIII shows top 10 hotels and bottom 10 hotels in their credibility values. The credibility values in the table are 10,000 times of the original values for easiness in reading. Unlike the results on reviewers in the last subsection, number of reviews for top 10 hotels are quite high rather than bottom 10 hotels. This is because the high credibility of hotel scores needs convergence of scores as necessary conditions.

As with the results on reviewers, we investigate that high-ranking hotels are evaluated by high-ranking reviewers. Table IX shows the average rank of reviewers who reviewed top 3 hotels and bottom 3 hotels. We found clearly that the high-ranking hotels in hotel ranks are really reviewed by credible reviewers, and that low-ranking hotels are reviewed by incredible reviewers. For example, HID 74637 which has the highest credible hotel score, has 1,369 reviews and the average rank of the reviewers is 2427.4. On the other hand, HID 129491 which has the lowest credible score, has 36 reviews and the average rank of the reviewers is 25525.1.

D. Predicting Future Scores

The evaluation results shown in the last subsection are based on the data until 2012 (Data-1). We could find out the credible reviewers and credible hotel scores from these indexes. In this subsection, we consider how to use these indexes for predicting future scores in 2013-2015 data

(Data-2).

TABLE VII: CREDIBILITY OF HOTEL REVIEWS BY TOP/BOTTOM 3 REVIEWERS

Top 3 Reviewers

UID	# Rev.	Av.(*)
user23632	50	1768.8
user74703	7	1780.4
user456451	5	2377.8

Bottom 3 Users

UID	# Rev.	Av.(*)
user161845	11	4448.6
user62392	7	9130.4
user23040	10	5146.8

"UID" : user ID,
"# Rev" : Number of reviews by the reviewer,
"Av.": Average rank of hotels reviewed by the reviewer.

TABLE VIII: TOP/BOTTOM 10 HOTELS IN CREDIBILITY

Top 10 Hotels

Hotel ID	Credibility	#Reviews
74637	7428.65	1369
2614	3360.39	853
71982	2954.54	567
1177	1500.00	778
40672	1346.06	444
51637	1272.60	577
7588	1242.97	455
29211	1224.70	553
129965	890.07	297
938	761.82	377

Bottom 10 Hotels

Hotel ID	Credibility	#Reviews
129491	0.00033	36
31902	0.00170	16
14680	0.00258	43
76886	0.00295	8
39501	0.00315	5
9469	0.00513	6
80727	0.00523	5
54665	0.00523	5
54146	0.00524	5
135932	0.00729	5

TABLE IX: CREDIBILITY OF REVIEWERS WHO REVIEW TOP/BOTTOM 3 HOTELS

Top 3 Hotels

HID	# Rev.	Av.(*)
74637	1369	2427.4
2614	853	6922.5
71982	567	3465.1

Bottom 3 Hotels

HID	# Rev.	Av.(*)
129491	36	25525.1
31902	16	43943.6
14680	43	44980.0

"HID" : hotel ID,
"# Rev" : Number of reviews for the hotel,
"Av.": Average rank of reviewers who reviewed the hotel.

Let v_{ij} be a score by a reviewer i for a hotel j in Data-1, and $\overline{s_j^1}$, $\overline{s_j^2}$ are the average scores of hotel j in Data-1 and in Data-2 respectively. If $v_{ij} < \overline{s_j^1}$, and $\overline{s_j^1} > \overline{s_j^2}$, we consider the review score v_{ij} in Data-1 has foresight of the uptrend in future review scores. Similarly, if $v_{ij} > \overline{s_j^1}$, and $\overline{s_j^1} < \overline{s_j^2}$, we consider the review score v_{ij} has foresight of the downtrend in future review scores. We think the scores in these two cases can be considered as "correct predictions" of future trends. Otherwise, we think these scores as "incorrect predictions".

We try to evaluate the correct ratio and compare them with reviewers rank. We made a data set each record of which consists of "Hotel ID, Reviewer ID, Reviewer's Credibility, Review Result in Data-1, Average Score of the Hotel in Data-1, Average Score of the Hotel in Data-2", from Data-1 and Data-2 (748,700 records in total). We divide the data equally into 5 classes according to the descending order of reviewer's credibility, Class1 to Class5, that is, Class1 includes the data of reviewers with highest credibility, while Class5 includes the data with lowest credibility.

Table X shows correct ratios by Class1 to Class5. As seen in the table, the higher credibility makes high accuracy while lower credibility makes low accuracy for future prediction. Although the difference is not so large, we can conclude that the scores of credible reviewers reflect the actual status of hotels definitely and these can be used for the future prediction, especially for the hotels that have not enough amount of reviews currently. For example, consider a hotel that the number of reviews is small and distributed, but the reviews for the hotel include those by some credible reviewers. In this case, we can guess the future score of that hotel may converge to the score by credible reviewers.

TABLE X: CORRECT RATIO OF FUTURE PREDICTION

Class	Correct Ratio
Class1	51.68%
Class2	51.31%
Class3	50.37%
Class4	49.11%
Class5	46.81%

VI. CONCLUSION

There are lots of web services which has user review facilities, such as bookselling sites, travel websites and daily necessaries selling sites. Although these reviews on the sites are helpful for users, the reviews are a mixture of credible and incredible. In recent years, it becomes the problem that there are some skeptical reviews on the shopping site. These incredible reviews are given by suspicious reviewers or incredible reviewers.

In this paper, we proposed a method to find the credibility of hotel scores and reviewers for a travel website Rakuten Travel. We think the credibility of review scores can be detected by focusing the relation between hotel scores and reviewers, and introduced mutual recursive flamework for a bipartite relation between hotel scores and reviewers. We also showed that these indexes can be used for future predictions.

As future works, we will refine the initial values in the matrix. In this research, we adopted the closeness to the average values and the length of the comment as initial values. Needless to say, the concept of the length of the comment is a little rough idea. If we can use other attributes of reviewers such as age, gender, residential area and job, and if we can use other linguistic criteria for the comments, it may be possible to use other reasonable initial values. In recent years, some researches consider how to judge the fake sentences and they can be applied to some fields.

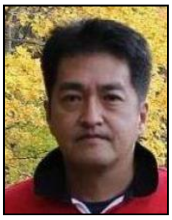
Review evaluation methods shown in this paper can be also applicable to other services such as bookselling sites and daily necessaries selling sites. We are now trying to apply and modify our method for other fields, such as cooking recipe services.

ACKNOWLEDGMENT

In this paper, we used Rakuten Travel data provided by Rakuten Inc. and the National Institute of Informatics (NII Data Repository) [2].

REFERENCES

[1] Rakuten Travel. [Online]. Available: https://travel.rakuten.co.jp
[2] Rakuten Data Set (NII Data Repository). [Online]. Available: https://www.nii.ac.jp/dsc/idr
[3] Travel Voice. [Online]. Available: https://www.travelvoice.jp/20160816-72061
[4] S. Kusumasondjaja *et al*., "Credibility of online reviews and initial trust," *Journal of Vacation Marketing*, vol. 18, no. 3, pp. 185-195, 2012.
[5] J. Mackiewicz, "The co-construction of credibility in online product reviews," *Technical Communication Quaterly*, vol. 19, no. 4, pp. 403-426, 2010.
[6] D. C. DeAndrea *et al*., "How people evaluate online reviews," *Communication Research*, vol. 45, no. 5, pp. 719-736, 2015.
[7] U. Chakraborty, "Credibility of online reviews and its impact on brand image," *Management Research Review*, vol. 41, issue 1, pp. 148-164, 2018.
[8] B. J. Fogg *et al*., "How do users evaluate the credibility of websites? A study with over 2,500 participants," in *Proc. the 2003 Conference on Designing for User Experiences (DUX '03)*, 2003, pp. 1-15.
[9] B. Helversen *et al*., "Influence of consumer reviews on online purchasing decisions in older and younger adults," *Decision Support Systems*, vol. 113, pp. 1-10, 2018.
[10] S. Mukherjee *et al*., "Credible review detection with limited information using consistency features," in *Proc. ECML PKDD 2016*, 2016, pp. 195-213.
[11] R. Sharma *et al*., "Algorithm for ranking consumer reviews on e-commerce websites," *Int'l Journal of Computer Science and Information Technologies*, vol. 6, no. 6, 2015.
[12] J. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM (JACM)*, vol. 46, no. 5, 1999.
[13] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," in *Proc. Seventh Int'l World-Wide Web Conference*, 1998.
[14] Y. Shirai, H. Morita, M. Nakamoto, and S. Oyama, "Application of popularity sensitivity and priority point for customer segmentation," *Operations Research as A Management Science Research*, vol. 59, no. 2, pp. 88-95, 2014.
[15] Y. Shirai, H. Morita, S. Cheung, M. Nakamoto, and H. Takashima, "Quantification of maniac behavior based on the coevolutional processes," *Oerations Research as A Management Science Research*, vol. 61, no. 2, pp. 80-87, 2016.
[16] H. Kozuka, H. Kawamata, T. Satomura, T. Moriguchi, and Y. Shirai, "Quantification of Waseda-Keio peculiar behaviours by the link network analysis," in *Proc. 100th Annual Conference of Japan Institute of Marketing Science*, 2016.



Yasuyuki Shirai received a BSc in mathematics in 1987, and a MEng in CS in 1989 from Tokyo Institute of Technology. He received a Dr. Eng. degree in CS from Kyushu University in 2003. His current research interests include machine learning, data mining, web mining, automated reasoning and web mining. Now He is an associate professor of the Faculty of Management, Daito Bunka University.