Predicting Future Ranking of Online Novels based on Collective Intelligence

Shimizu, Kazunori
Research Institute for IT | Grad. School of ISEE

Ito, Eisuke
Grad. School of ISEE | Research Institute for IT

Hirokawa, Sachio
Research Institute for IT | Grad. School of ISEE

https://hdl.handle.net/2324/25940

出版情報：Proceedings of ICDIPC2013 (The Third Int'l Conf on Digital Info. Processing and Communications), pp.261-272, 2013-02-14. SWDIC
バージョン：

Predicting Future Ranking of Online Novels based on Collective Intelligence

Kazunori Shimizu
Grad. School of ISEE

Eisuke Ito
Research Institute for IT

Sachio Hirokawa
Research Institute for IT

Kyushu University
Hakozaki 6-10-1 Higashiku, Fukuoka, 812-8581, Japan
{2IE11061G@s, ito.eisuke.523@m, hirokawa@cc}.kyushu-u.ac.jp

ABSTRACT

A large number of novels are being uploaded as online novels. The present paper proposes a ranking algorithm based on the users' favorite lists (bookmarks). Empirical evaluation has been conducted with respect to each genre of novels. In several genres, it is confirmed that the top ranked novels in July are predicted from the bookmarks of May.

KEYWORDS

Collective Intelligence, Ranking, User generated type contents services, Online novels, Bookmark, HITS

1 INTRODUCTION

The search and recommendation mechanism are crucial to help users to discover their favorite contents among a huge number of contents that were posted by users. Measurement to evaluate the quality of contents and method to categorize the contents are the most important mechanism in such systems. Collective intelligence has been paid much attention as a method to realize the two mechanisms. In [2, 9, 10], tags and comments by other readers are used to search and recommend related contents for a reader.

User generated type contents services, such as Youtube[1)], Nicovideo[2)] and YouKu[3)], are popular. The popularity of online novels are has been g [1] oing up gradually, in recent years. The number of contents and the number of readers are increasing in the site of "Qidian" in China [1] and the site of "Syosetu" in Japan [12]. Those online novels might be on the sales targets of e-book terminals such as Amazon Kindle.

Recommendation and categorization have been gaining much interest as research target. Not only recommendation of commercial products but also the recommendation of movies [9, 10] and scientific papers [2] have been the targets of those research. Faceted analysis of documents [11] is very close to those researches. Ranking of the online novels, which is the subject of the present paper, is a new research target. There are a few works which considers ranking of the online novels in the site of "syosetu.com" [5, 6], where ranking method and categorization methods are proposed.

The site of "syosetu.com" contains more than 13 million online novels as of September 2012 and is still drastically in-

1) http://www.youtube.com/
2) http://www.nicovideo.jp/
3) http:www.youku.com/

creasing the number of contents and readers. The quality evaluation is one of the most important and difficult task not only for online novels but also for movies and products. Most of those online novels are written by amature writers and might not be of good quality. However, there are some high quality novels found as well. It is necessary to find appropriate data source for estimating the quality.

In ”syosetu.com”, the readers can write their feedback to the author of the novels. Readers can register a favorite novel in the site, if they want to read the next episodes in the novel. The system automatically notifies the readers concerning the new addition of section or episode to their favorite novels. A user can registrate his/her favorite authors as well. The system notifies the latest section of the novel or the latest novel by the authors. Readers can write their comments on a novel if the author of the novel admits. Those comments become as strong encouragement for writers. Some comments point out the typographic errors and improvement of the contents.

The feedback from readers expresses the desirability to the novel. The readers' favorite registration can be considered as collective intelligence. The total registration number by readers in May is used as ranking measurement. The present paper proposes a ranking method of online novels based on the bipartite graph structure of novels and readers. Moreover, we focus on the difference of weights for each genre according to each reader. There are some expert readers for particular genres. But it is not always the case that they can appropriately evaluate any kinds of novels.

The rest of the paper is organized as follows. Section 2 reviews the related work. Section 3 describes the structure of “syosetu.com”, the basic statistics such as the number of novels and the number of readers, as well as the distribution of keyword frequencies in the novels. In section 4, we propose a ranking of novels based on the bipartite graph structure. Section 4 shows examples of ranked novels and the evaluation of the method in predicting the popularity. Section 6 concludes the paper and shows further work.

2 RELATED WORK

There are many researches and services to recommend commercial products. The basic information for recommendation are easily collected as the reputation of the product by users. For example, [3] propose a recommendation system which retrieves products of recommendation by analyzing the relationship of products, users and tags from social media such as SNS.

On the other hand, there is no research or service to recommend online novels, as far as the authors know. One of the reasons of this difference, that the uthors guess, is in the ratio between the number of targets and the number of readers. Concerning to product reputation, the number of users is very large compared to the number of products to be evaluated. One product may require a large amount of money to create it as a commercial product. On the other hand, CGM (Consumer Generated Media) requires almost no money. At least a user is not required to pay any money to post his/her comment in such a site. The online novels uploaded in “syosetu.com” are created free by amature writers. As a result, the number of contents and the number of readers are very close (Table 1). This fact indicates a difficulty in collecting many reputations against each content and then a difficulty in applying well known recommendation method such as collaborative filtering.

Tags are recognized as core information inWeb2.0 to infer the contents of the target items. However, some of tags may be nothing but noises, since they are assigned too easily by users. [8] tried to get rid of those noises by assigning appropriate weight to each tag depending on user's preference.

In "syosetu.com", a writer can assign at most 15 keywords as tags to his/her novel. However, those keywords can be determined by the writer's choice and cannot be used as controlled vocabulary for classification of novels.

Extraction of meaning and evaluation from the relation of novels, readers and keywords that are assigned by users to novels is the crucial point in ranking and recommendation system. It would be natural to think of applying graph theory whose nodes are novels, readers and keywords and whose edges are their co-occurrence relations. Thus, the graph will be formalized as a tripartite graph. In the present paper, however, the authors consider bipartite graph whose nodes are novels and readers. This choice of formulation has the following two merits. First merit is in the simplicity and the visibility of the graph where we can focus on novels and readers. The second merit is in the performance as computation time. If we chose all of three objects, i.e., novels, readers and keywords, the total number of edges would be quite large compared to our choice. The reasonable performance speed is very important factor in real services.

Conventional recommendation systems use collaborative filtering to optimize the precision of recommendation. However, the recommended results tend to contain items that a user already knows before. [4] proposed the notion of novelty and three algorithms to discover new items as recommendation. They evaluated that their algorithm work well in recommending new interesting items to users. However, their evaluation has nothing to do with time-axis. In the present paper, we are interested in recommending items that will be interesting to a user in near future even though the user does not know now. We make an empirical evaluation of the proposed recommendation method in time-axis. In this sense, our method can be considered as a prediction method as well as recommendation method. In other words, our proposed method predicts the novelty of the novels and the future ranking of novels.

3 STATISTICS ON "syosetu.com"

In this section, we explain the organization of the site "syosetu.com", the number of novels and the number writers. We also describe a basic statistics concerning to the distribution of frequencies of words which were assigned to novels by users.

3.1 Organization of "syosetu.com"

The site "syosetu.com" is served by HINA PROJECT inc. where anyone can subscribe and read their novels freely.

Table 1 The number of Novels, Readers and Authors

	May 2012	July 2012	September 2012
Novels	159,090	168,396	137,730
Readers	240,730	258,478	272,512
Authors	53,396	56,214	44,585
My Pages	92,418	---	---

Table 1 shows the number of writers, readers and novels as of May, July and September 2012. There is no distinction between writers and readers except that a writer registrates at least one novel of his/her own. The number of "my pages" in the table represents the number users who bookmarked their favorite novels in the site. The list of favorite novels are kept by the system for each user. We will explain in detail on "my pages" later. The decline of the numbers of novels and writers in September 2012 are due to the removal by the administrator of non-original novels, i.e., the secondary writings.

Figure 1 illustrates the interaction between an author and his/her readers. An author can upload his/her novel to the system. A reader can not only read a novel but also can write his/her comments to the novel. A novel may contain several sections. If a novel contains only one section, it is a short novel. Other novels are serial novels where the author can upload the latest section as he/she writes. In the case of a serial novel, the author can specify a conclusion. When an author contributes a novel, he/she can set the title, the author's name, the genre, keywords and the outline as metadata. Genre has to be chosen from 15 keywords provided by the site management side. The author can choose arbitrary keywords freely with a limited length. The outline can be Figure 1 illustrates the interaction between an author and his/her readers. An author can upload his/her novel to the system. A reader can not only read a novel but also can write his/her comments to the novel. A novel may contain several sections. If a novel contains only one section, it is a short novel. Other novels are serial novels where the author can upload the latest section as he/she writes. In the case of a serial novel, the author can specify a conclusion. When an author contributes a novel, he/she can set the title, the author's name, the genre, keywords and the outline as metadata. Genre has to be chosen from 15 keywords provided by the site management side. The author can choose arbitrary keywords freely with a limited length. The outline can be described as he/her wishes with a limited length.

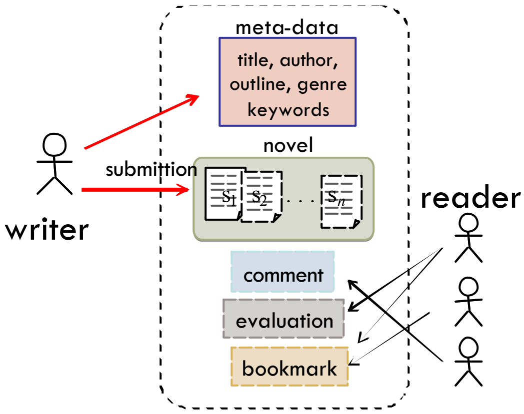

Figure 1 Data Model

Anyone can read the contributed novels without user registration. The person who carried out user registration to the site can send his/her feedback by scoring and commenting to a novel. The registered users can use their bookmarks of favorite novels and favorite writers. Users have their own "my page" to see the registered information as well as related novels. Figure 2 illustrates the structure of the my page of a user. In my page, a user can check the latest series of his/her favorite novel and the latest novels of his/her favorite writer.

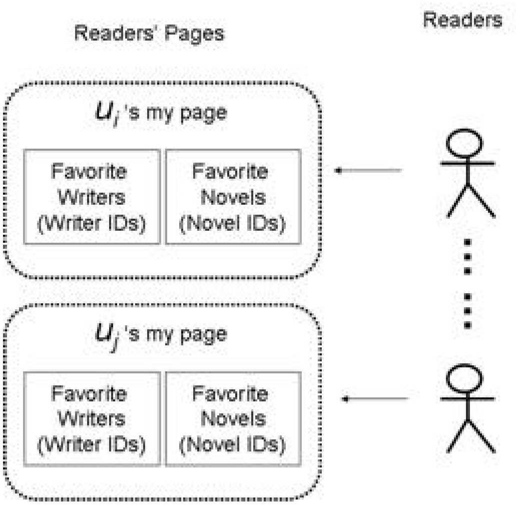

Figure 2 My Page of a User

3.2 Distribution of Keywords Frequencies

We collected the HTML pages that ontain meta-data from syosetu.com in April 2012. At the same time, we collected the list of score of novels and the list of favorite novels with respect to each user. The followings are a basic analysis on the frequency distribution of genres and keywords in meta-data of novels. Note that our analysis is based on the meta-data and is not on the whole content of the novels. There are 128,115 unique words written as keywords of novels by the writers. Table 2 shows the top 20 words together with the frequency of each word. Some the top words are marked as warning words by the management side. The top word "cruel depiction" and the third "R15" in Table 2 are examples of those warning words. The warning words are used in filtering of novels. The reader who does not like cruel contents can limit to the novels which do not contain the keyword "cruel depiction".

Table 2 Top 20 Frequent Words

Rank	Word	Freq.
1	cruel depiction	27,696
2	love	26,669
3	R15	21,718
4	modern	21,547
5	fantasy	20,247
6	high school student	15,633
7	serious	12,900
8	tender	11,651
9	different world	11,433
10	youth	9,303
11	magic	8,673
12	girl	8,169
13	comedy	7,893
14	school	7,277
15	friendship	6,960
16	Boy	6,696
17	campus	6,689
18	happy end	6,685
19	literature	4,859
20	dark	4,742

Figure 3 plots the frequencies of the words in decreasing order. We can observe that the distribution follows the Zipf's law.

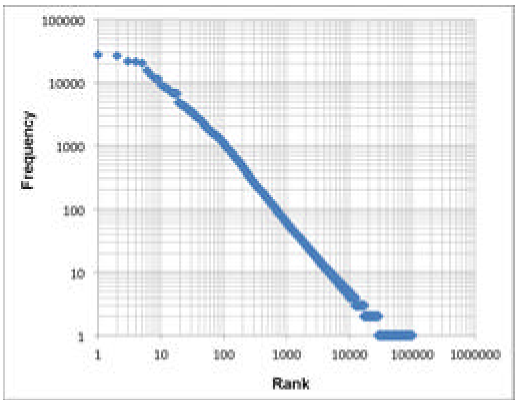

Figure 3 Distribution of Keyword Frequencies

Most of the keywords have very low frequencies. Table 3 displays the number of words whose frequencies are smaller than 11. There are 99,481 words that appear only once. There are 11,440 words that appear twice. Over 86% of all words are covered by those words that appear at most twice.

Table 3 Rates of Infrequent Words

frequency	unique word	Rate (%)
1	99,483	77.7
2	11,440	8.9
3	4,760	3.7
4	2,490	1.9
5	1,667	1.3

3.3 AUTHORS' TREND ANALYSIS

Next, the tendency for every author was investigated. We collected author IDs indicated in novel metadata and analyzed the tendency of the number of contributed novels and the number of keywords for every author. Table 4 shows the IDs and the number of contributed novels for the top 10 writers.

Table 4 Top 10 Writers

rank	ser-ID	the number of novels
1	73	1,218
2	26055	491
3	107085	425
4	26407	300
5	153402	290
6	34896	265
7	9272	254
8	47590	233
9	126858	230
10	200	200

Figure 4 shows the distribution of the number of novels in descending order. Note that both axes are drawn in log-scale. We can observe that the distribution follows the Zipf's law.

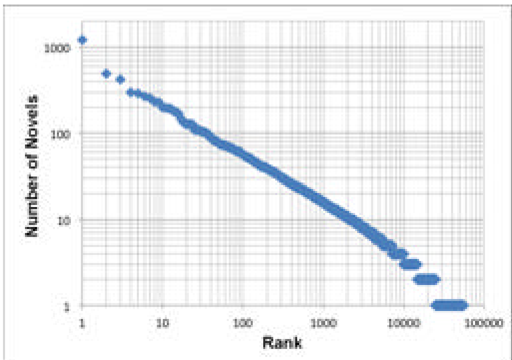

Figure 4 Distribution of the Number of Novels by a Writer

At the next step, we analyzed each writer if he/she focuses on a particular genre or writes a variety of novels in many genres. We used the correlation between the number of contributed novels and the number of keywords for each writer. The scatter diagram of the number of novels and the number of keywords is shown in Figure 5.

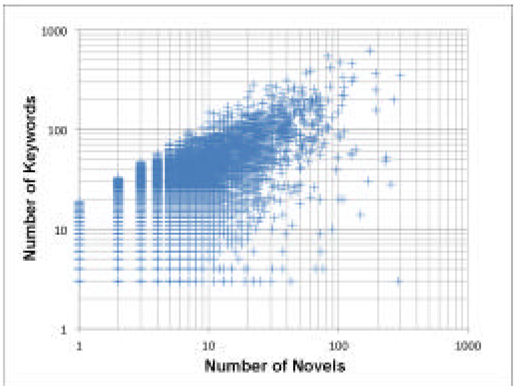

Figure 5 The number of contributed novels and the number of keywords by a writer

We can observe that the points appear below a line. This is because of the limit of the number of keywords that can be assigned to a novel. The maximal number of keywords for a novel is 15 in syosetu.com. As the consequence of this restriction, the number of keywords is smaller than 15 times of the number of contributed novels for each writer. This is why we see the straight line showing a maximum number of keywords. We can observe that the number of keywords and the number of novels are correlated. It means that most of the writers contribute their novels with a variety of keywords. However, we can observe that there are several authors with very few keywords. The right top point represents the writer who has the largest contribution number of novels (175) and with the largest number of keywords (618).

4 RANKING OF NOVELS BASED ON BIPARTITE GRAPH

4.1 Bipartite Graph of Novels and Readers

We introduce a graph structure to formalize the ranking of novels. We use the symbols U, C, A, C_u and U_c to represent the following data sets.

- U : The set of user u (u $\in$ U) represents a user.
- C : The set of contributed novels c (c $\in$ C) represents a novel.
- A : The set of authors a (a $\in$ A) represents a writer.
- C_u : The set of favorite novels of a user.
- U_c : The set of users who registered the novel c as a favorite novel

The relation between the writers, the novels and the readers can be drawn as a tripartite graph. The set of nodes of the tripartite graph consists of U, C and A. An edge (u, c) shows the fact that the user u registered the novel c in his/her favorite. An edge (u, a) represents the fact that the user u registered the author a as his/her favorite. An edge (a, c) represents the fact that the novel c is written by the author a. Figure 6 illustrates the linkages between users, novels and authors.

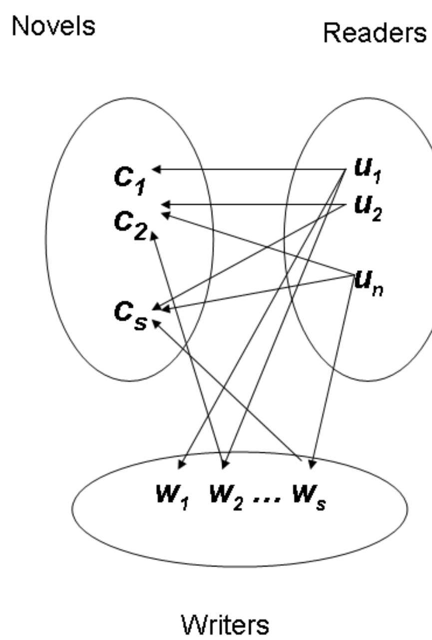

Figure 6 Tripartite Graph of Users, Novels and Authors

Figure 7 illustrates a bipartite graph of readers and novels, where edges represent bookmarks. This is obtained from Figure 6 by deleting writers.

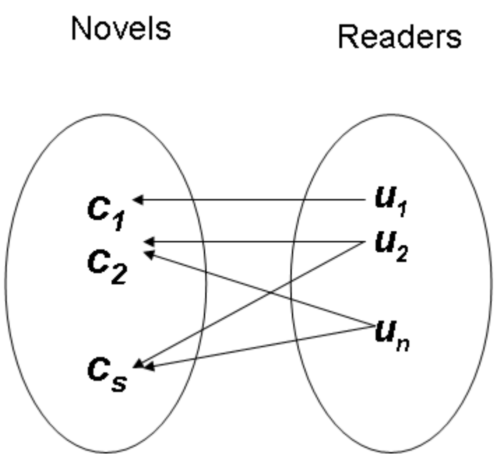

Figure 7. A Bipartite Graph of eaders and Novels

4.2 Problems of Conventional Ranking Methods

Two ranking methods are widely used in many user contribution type contents sites. The first method is the ranking by popularity and the other is the ranking by freshness. Popularity ranking is measured by the number of readers, the evaluation scores by users and the number of bookmarks.

In the site of "syosetu.com", the ranking of a novel c is determined by the following formula (1), where $|U_c|$ denotes the number of bookmarks that contain the novel c, $t_{u;c}$ denotes the sentence evaluation score (1..5 points) and $s_{u;c}$ denotes the story evaluation score (1..5 point) of the novel c by a user u. $p(c)$ denotes the overall score of the novel.

$$p(c) = (2|U_c|) + \sum_{u \in U} (t_{u,c} + s_{u,c}) \qquad (1)$$

The top ranked novels by this formula are read by many readers, have high evaluation scores and are stored in the bookmarks of many readers. Since this score is accumulated, the old novels tend to gain high ranking. On the other hand, it is very rare for a new novel to be ranked in the upper position, even if the quality of the novel is good enough.

This is not a problem specific to syosetu.com but is a common problem of the accumulation type scores. To overcome this problem, many sites utilize the popularity ranking in recent period. In fact, "syosetu.com" and "nicovideo movie" provide the daily, weekly and monthly popularity rankings.

Popularity ranking is useful for beginners and light users. Those users have almost no experience of reading the novels in the site. They will be satisfied in reading top ranked popular novels most of which are old. On the other hand, the ranking in the latest popularity is useful for heavy users. They spend much time at the site and read many contents. Chances would be very hight that those user have read most of the popular novels. A heavy user reads and hunts for the new novel contributed. Therefore, the ranking by freshness will be fine for them.

There is no appropriate popularity ranking for the middle-class users. They like reading novels but do not have much time to read and hunt for many novels. This paper proposes a ranking method for those middle-class users.

4.3 Intuition of Ranking Algorithm

We apply the HITS algorithm [7] assuming that the bookmarks represent a kind of link structure. HITS stands for Hyper Induced Topic Selection. The HITS algorithm calculates the hub-degrees and the authority-degrees of Web pages by propagating the score along the linkage. As a result, highly scored Web pages can be extracted.

The reason why we apply HITS to bookmarks of novels is that we have the following hypothesis. There are a small number of readers, or judges, who can make excellent choices of novels among many heavy users who haunt novels. They take the initiative in other readers, and have the capability to select the novel of high quality out of the novel contributed recently. Middle-class readers would keep a good novel in their bookmark if they find a good novel by chance.

If we consider the novels as authorities and the readers as hubs and the bookmarks as links, we arrive the following mutual definitions of good readers and good novels. Good readers are those who bookmark good novels. Good novels are those novels which are bookmarked by good readers. However, a naive application of HITS will be identical to the popularity ranking.

We are not interested in the final scores of the HITS algorithm. Instead, we consider the intermediate values in the iteration process of HITS algorithm. In this formulation, we think that we can simulate how the influence of the good readers are propagated.

4.4 SHIP Algorithm

Now we describe the proposed algorithm SHIP (SHimizu Ito Process). The algorithm propagates the scores of novel assigned by a small number of judges. We use the following notation in the definition of SHIP.

$w(u)$ represents the weight of a reader u. $w(c)$ represents the weight of a novel c. We denote the weight of a reader and a novel after n times iteration by $w_n(u)$ and $w_n(c)$ respectively. There are many possibilities as the initial weights of $w_0(u)$ and $w_0(c)$. We used the same constant value as the initial weights in the experiment of the next section. The ranking of novels are given by the weight in descending order.

The SHIP algorithm calculates $wn(u)$ and $wn(c)$ by the following procedure.

```
// Initialization
for each user u in U do
  w(u) = 1
for each content c in C do
  w(c) = 1
// Calculation of Weights
function SHIP(U,C)
  for step from 1 to n do
    norm = 0
    for each user u in U do
      w(u) = 0
      for each content c in C_u do
          w(u) += w(c)/|C_u|
      norm += square(w(u))
    norm = sqrt(norm)
    for each novel c in C do
      w(u) = w(u)/norm
    norm = 0
    for each content c in C do
      w(c) = 0
      for each u in U_c do
        w(c) += w(u)/|U_c|
      norm += square(w(c))
     norm = sqrt(norm)
     for each page c in C do
       w(c) = w(c)/norm
```

The time complexity of the algorithm is $O(n ? r)$, where n is the number of edges of the bipartite graph, i.e., n = S(U_c) = S(C_u) and r is the iteration count. We confirmed that that the iteration count r can be lower than 20 to get the best estimation.

5 EXPERIMENT AND EVALUATION

This chapter describes the experiment and the evaluation of the prediction performance of the proposed ranking method. The hypothesis is that we can predict the future ranking of novels based on bookmarks that we already have. In actual experiment, the ranking obtained by SHIP using the bookmarks as of May, 2012 is compared with the real popularity ranking as of July, 2012.

5.1 Dataset

We collected the meta-data of online novels and the bookmarks as of May, 2012 and as of July, 2012 from syosetu.com. Table 5 lists the details of the dataset that we used for evaluation.

Table 5 Dataset

	May 2012	July 2012
|C|	159,090	168,396
|C'|	64,519	-
|U|	240,730	258,478
|U'|	92,418	-
|E|	5,435,508	-

In Table 5, *U* represents the set of all users, *U'* represents the set of all users who register at least one novel in his/her bookmark. *C* represents the set of all novels, *C'* represents the set of novels which are registered by at least one user. $|E|$ is the total number of registered novels in bookmarks, i.e., the number of links from users to novels.

It is worthwhile to note that the ranking algorithm yields the same result when it is applied to either *U* and *C* or *U'* and *C'*. Moreover, the size of total sets *U* and *C* are quite large compare to those of the restricted sets *U'* and *C'*. So, we apply the SHIP algorithm to *U'* and *C'*. It is very common in the case of user contribution type sites that there are enormous trial contribution. Those trial contents are not bookmarked or assigned scores. A large number of such contents locates at the bottom in the ranking. We remove those trial contents in our evaluation experiment and use meaningful *C'* for which at least one reader recognized as good quality.

5.2 Evaluation

Now we explain how we evaluate the ranking. We think that the proposed method can predict and can find the novels which future popularity take off a little bit earlier. To confirm this hypothesis, we compare the ranking obtained by SHIP from data at the time of the past and the ranking at the future time. Specifically, the prediction ranking from the data as of May, 2012 is compared with the ranking as of July, 2012.

The site of syosetu.com uses the formula (1) as the popularity ranking and lists the novels in descending order with respect to the score. We list the novels *c* in descending order of the weight $w_n(c)$ obtained by SHIP algorithm. We compare the top *k* novels in the original site and the top *k* novels in SHIP ranking. We conducted the comparison with respect to each genre. The following symbols denote the experiment datasets.

*R*5 : Top *k* novels ranked as of May, 2012

*R*7 : Top *k* novels ranked as of July, 2012

*P*5 : Top *k* novels predicted by SHIP based on the dataset of May, 2012

Table 6 shows the prediction performance of SHIP with respect to the top k = 100 novels and the iteration count $r = 20$. Each column of Table 6 represents the following value.

(R-UP) The number of novels in top 100 as of July but not in top 100 as of May, identical to |*R*5 *n* *R*7|.

(P-UP) The number of novels in top 100 by SHIP based on dataset as of May but not in top 100 as May, identical to |*P*5 *n* *R*5 *n* *R*7|

(Performance) Prediction performance calculated by PUP/ R-UP

Table 6 Prediction Performance

Genre	R-UP	P-UP	Performance (%)
all genres	14	1	7.14
literature	6	1	16.67
love	3	3	100.00

history	6	1	16.67
mystery	3	1	33.33
fantasy	15	6	40.00
SF	2	1	50.00
horror	6	5	83.33
comedy	8	5	62.50
adventure	6	4	66.67
school	7	1	14.29
war	4	1	25.00
fairy tale	6	4	66.67
poetry	11	4	36.36
essay	7	6	85.71
other	5	3	60.00

The second column (P-UP) of Table 6 contains at least one novel. It means that SHIP succeeds to predict at least one novel that was not in top 100 as of May but was in top 100 as of July. The performance in all genres, the first line, is very low. In fact, only one new novel is succeeded to predict in top k = 100. However, SHIP gains over 50% the prediction performance in more than half genres. Remember that "fantasy" and "love" are the genres which contains the largest number of novels in Table 2. SHIP succeeds to predict 100% in "love" and 40% in "fantasy".

We do not have a clear theory to explain why the performance varies with respect to genre. However, such an hypothesis would be reasonable that the influence of judges are very weak in the genres whose prediction performances are low. The novels in such low prediction genres might be tagged with the keywords of other genre. For example, it is reported, in [6], that most novels in the history genre tend to be tagged with "love" and "fantasy". The readers who have good eyes on historical novels would not read the love historical novels or the fantasy historical novels. At least, they would not make a bookmark on those mixture novels. As a result, the number of bookmarks on the history genre becomes small and then the prediction performance of SHIP becomes low. The number of readers who like historical novels might be very small at all.

6 CONCLUSION

The contents services of the user participatory type are becoming in fashion and popular in recent years. In search of interesting contents, search, a classification, and the recommendation technique are important, since contents are huge.

There are no editor or no trained librarian concerning to online novels. Neither quality evaluation nor a classification technique is clear. On the other hand, there are many readers who perform comment and favorite registration. If we can use the collective intelligence gained by a large number of readers, search, the ranking, and a classification of good quality may be possible.
The present paper proposed a ranking method to predict the future popularity based on the linkage information between novels and readers. The method is based on the HITS algorithm. Performance evaluation is conducted using real dataset in a Japanese online novel site syosetu.com. As a result, it turned out that the proposed method is sufficiently effective with respect to main genres. More precise evaluation will be necessary. We used only two datasets of May 2012 and July 2012. Evaluation in the long range is needed to confirm if the method is stable or not. In the present paper, we used the ranking obtained after r = 20 times. It is worth while to evaluate the effect of the iteration number.

This work was supported by JSPS KAKENHI Grand Number 2350099 and

24500176.

REFERENCES

1. QiDian http://www.qidian.com/.
2. Kensuke Baba, Eisuke Ito, and Sachio Hirokawa, Co-occurrence analysis of access log of institutional repository, In Proc. of JCAICT2011, pp. 25-29, 2011
3. Ido Guy, Naama Zwerdling, Inbal Ronen, David Carmel, and Erel Uziel, Social media recommendation based on people and tags, In Proc. of SIGIR 2010, pp. 194-201, 2010
4. Yoshinori Hijikata, Takuya Shimizu, and Shogo Nishida, Discovery-oriented collaborative filtering for improving user satisfaction, In Proc. of IUI2009, pp.67-76, 2009
5. Eisuke Ito, Sachio Hirokawa, and Kazunori Shimizu, Introducing faceted views in diversity of online novels, In Proc. of ICDIM2012 (Seventh International Conference on Digital Information Management), pp. 145-148, 2012
6. Eisuke Ito and Kazunori Shimizu, Frequency and link analysis of online novels toward social contents ranking, In Proc. of SCA2012 (The 2nd International Conference on Social Computing and its Applications), pp. 531-536, 2012
7. Jon M. Kleinberg, Authoritative sources in a hyperlinked environment, Journal of ACM, 46, pp. 604-632, 1999
8. Huizhi Liang, Yue Xu, Yuefeng Li, Richi Nayak, and Xiao- hui Tao, Connecting users and items with weighted tags for personalized item recommendations, In Proc. of HT2010, pp. 51-60, 2010
9. Naomichi Murakami and Eisuke Ito, Tag hierarchy analysis in Nicovideo movie service (in Japanese), IPSJ Tech Report 2010-MPS-81(17), pp. 1-6, 2010
10. Naomichi Murakami and Eisuke Ito, Emotional video rank- ing based on user comments, In Proc. of ACM iiWAS2011, pp. 499-502, 2011
11. Chengjiu Yin, Sachio Hirokawa, Jane Yin-Kim Yau, Tetsuya Nakatoh, Kiyota Hashimoto, and Yoshiyuki Tabata, Analyzing research trends with cross tabulation search engine, International Journal of Distance Education Technologies, 11(1), (in print) 2012
12. HINA Project inc., “Syosetu ka ni narou”, http://www.syosetu.com