A Consideration to Estimate Spoiling Pages in Comics

Yoshiki MAKI, Yuji SHIRATORI, Kenta SATO, and Satoshi NAKAMURA

Graduate School of Advanced Mathematical Sciences, Meiji University, Tokyo 164-8525, Japan tekkanomaki01@gmail.com

Abstract: Previous studies suggested that spoilers might increase the enjoyment of novels. However, the problem of spoilers has not been sufficiently clarified. The objective of our work is to clarify the effect of comic spoilers and to apply clarified knowledge for applications. In our past work, we constructed a spoiler dataset and investigated the spoilers' effect by changing the spoiler timing for readers. However, in the past work, we could not clarify the characteristics of spoilers. In this work, we clarified that spoilers reduced the interest in continuing reading the comics and analyzed the characteristics of spoilers by using the dataset. Then, we considered how to construct a comic spoiler dataset and investigated how to determine the spoiling pages automatically using image processing, and character detection and so on.

Keywords: Comics, Spoiler, Automatic Detection, Characteristics

1. INTRODUCTION

People enjoy reading with the variety of literary forms such as novels, comics, and so on. Comics, as one of those forms, excite readers and evoke laughter, surprise, and other emotions through pictures and stories. The popularity of comics in Japan is very high; the size of the Japanese comic market reached nearly 3.77 billion U.S. dollars in 2013. Japanese comics became popular around the world, partly due to the promotion by the Japanese government. Comics are also of interest to the field of information engineering, in which various studies on comics are being conducted. Among them, spoilers stand out as one of the significant problems that attract the interest of the researchers.

Spoilers ruin contents by revealing the endings, e.g., the winning or losing of matches, or the life or death of characters. Readers often want to avoid encountering spoilers before enjoying the content. Nakamura et al. [4, 5] proposed several methods for reducing the risk of encountering spoilers.

There are several studies on the characteristics of story spoilers. Leavitt et al. [1] studied the effect of novels' spoilers, and Levine et al. [2] studied the effects of novels' spoilers based on their timing. However, they did not look into comics' spoilers and considered spoilers as plots in and of these stories. Hassoun et al. [3] studied the effect of comics' spoilers depending on the comic panel layout. However, they only focused on spoilers on one page and did not focus on the entire of comics.

In our previous study [6], we investigated the degree of effect of the spoiler depending on the reading progress. We revealed that there were two types of comics: one is the spoiler-acceptable type that did not reduce the reader's enjoyment after providing spoiler information, and the other is the spoiler-unacceptable type that reduced the reader's enjoyment after providing spoiler information. However, our previous study could not show the essential spoiler effect because we only compared the emotions at the end of the reading. In addition, it is necessary to clarify what kind of page people felt spoilers because in estimating spoiler pages, there is no study to define what kind of elements we should use.

In this paper, we firstly analyze the comic spoiler's effect more deeply such as the degree of reducing enjoyment after encountering a spoiler depending on the spoiler timing and any other characteristics in chapter 2. In addition, in chapter 3, we analyze the aspects of spoilers which were constructed in the previous work. Then, in chapter 4, we discuss the guideline to construct the useful spoiler dataset and a method to identify which pages recognize as spoilers.

2. ANALYSIS OF SPOILER EFFECT

2.1 Experiment on spoiler effect

In our previous study [6], we prepared eight comics and constructed the dataset of spoiler information which made by four participants who were hired to label spoiling pages and conducted an experimental test about the spoiler effect depending on the spoiling timing. To do that, we divided each comic into three parts according to the number of volumes. In addition, we prepared four patterns of providing spoiler information and asking to answer questionnaires (see Figure 1). Note that subjects who were assigned to Pattern 0 did not encounter spoiler information while reading the target comic.

42 undergraduate students at Meiji University participated in the experimental study. We assigned five subjects to each spoiling pattern in each comic randomly. Each subject had never read their target comic. As a result, 20 subjects enjoyed and tried the experimental test in each comic.



Figure 1: The timing of asking to answer questionnaires and providing a spoiler page in each pattern.

2.2 Results

Figure 2 shows the average of enjoyment value at the end of reading the comics in each pattern. The difference between Pattern 0 and others was not significant. In fact, there was no difference in the final enjoyment between subjects who were exposed to the spoilers and subjects who were not. In this experimental test, we forced subjects to read the comics to the last. In addition, the comics that we prepared were very popular and successful in their sales. In other words, this result only clarified that people could enjoy popular comics, regardless of whether or not to encounter spoiler information, when they read such comics to the last.



Figure 2: The average last enjoyment value in each pattern.

Here, in this paper, we analyzed the spoiler effect by comparing the level of enjoyment between before and after encountering the spoiler to reveal the impact of the spoiler.

There were two independent variables in this

experiment: the difference in just before and after the spoiler (spoiler factor) and the difference in the spoiler timing (spoiler timing factor). There were also dependent variables, i.e., two questions in the questionnaire about the subjects' subjective impressions.

The dependent variables in the questionnaire on enjoyment value and interest in continuation value were analyzed by two (spoiler factor) x two (spoiler timing factor) mixed-design analysis ANOVA (see Figures 3 and 4).

In the questionnaire on enjoyment value (see Figure 3), we observed the significant differences between Pattern 1 and 2 (p < 0.05), and between Pattern 1 and 3 (p < 0.05). However, there was no interaction effect.

In the questionnaire on interest in continuation (see Figure 4), we observed that there were significant differences between Pattern 1 and 2 (p < 0.05), and between Pattern 1 and 3 (p < 0.05). There was an interaction effect between the spoiler factor and the spoiler timing factor (F2, 117 = 8.04, p < 0.01). In addition, there were significant differences in Pattern 1 and 2 (p < 0.05) in the spoiler factor.

Figure 5 and 6 show a transition graph evaluating enjoyment and interest in continuation about reading progress and spoiler timing. In Figure 5, enjoyment tends to increase from the start to the end. Also, in Figure 5 and 6, in the pattern in which the subjects encountered spoilers, enjoyment and interest in continuation decreased after viewing the spoiler. But, in the pattern 1, interest in continuation increases after viewing the spoiler.



Figure 3: The difference of enjoyment before and after spoiler in each pattern.



Figure 4: The difference in interest to continue before and after spoiler in each pattern.



Figure 5: Transition of enjoyment value in relation to reading progress in each pattern.



Figure 6: Transition of interest value to continue in relation to reading progress in each pattern.

Table 1 shows the data on the interest in continuation depending on the conditions ("With spoiler" condition vs. "Without spoiler" condition). In this table, Pattern 0 is classified to "Without spoiler," and Pattern 1, 2 and 3 are classified to "With spoiler." This result clarified that 10% of the subjects who did not encounter a spoiler lost interest in continuation, and 37% of the subjects who saw a spoiler lost interest in continuation. In other words, spoilers reduce the readers' interest in continuation.

Table 1: The ratio of lost in	nterest with a spoiler or
-------------------------------	---------------------------

without a spoiler.				
	Number of	Lost		
	subjects	interest	Ratio	
Without spoiler	40	4	10%	
With spoiler	120	44	37%	

2.3 Discussion

As shown in Figure 2, there was no significant difference in the degree of enjoyment value in each pattern. So, we could say that spoilers do not affect enjoyment if readers complete reading whole of the major comics. It means that even with a spoiler, readers can still enjoy other excellent parts and read to the last. In addition, Figure 3 shows that there was no significant difference in the degree of enjoyment between before and after encountering spoiler information.

Here, we could classify these eight comics into two groups. One is a spoiler-acceptable comic, and the other is a spoiler-unacceptable comic from the previous study [6]. Furthermore, we found a difference in ideas about spoiler information among the experimental subjects. For that reason, we should redesign a method to construct spoiler dataset in comic. We discuss the method to construct spoiler dataset in section 3. Then, we should increase the number of each group's comic for experimental test and analyze them depending on the group in the future work.

In Figure 4, we could see the interaction effect between the spoiler factor and the spoiler timing factor. At first, the result of Pattern 1 showed that spoilers increased the interest in the continuation if the readers who only read the outline encountered the spoiler information. On the other hand, the result of Pattern 2 showed that the interest in continuation value dropped if the readers encountered the spoiler information before reading the final part of the comic. In addition, Table 1 shows that the readers might give up reading the comics if they encountered the spoiler information.

These results showed that it might be possible to motivate people who have never read the target comic to buy it by providing its outline and spoiler information. On the other hand, we also might be able to demotivate people who have already read the target comic except the final part to buy a next volume of the target comic by providing its spoiler information.

3. PRE-ANALYSIS OF SPOILERS IN DATASET

In the future work, we will realize a system which estimates the degree of spoiling of each facing pages in a comic based on the spoiler dataset which we will construct as soon as possible.

In this paper, we preliminarily find out the characteristics of spoiler pages by analyzing the current spoiler dataset which is constructed in the previous work [6]. In the dataset, four the participants selected three spoiler pages and ranked them in eight comics in four genres.

3.1 Characteristics of spoiling pages in a comic

Figure 7 shows the degree of spoiling in each page in each comic. The horizontal axis shows the page number of the final volume in each comic, and the vertical axis indicates the total score of spoiling evaluated by four the experimental participants. First-ranked spoiling page got five points, second-ranked spoiling page got three points, and third-ranked spoiling page got one point. YuYu Hakusho(by Yoshihiro Togashi) and All You Need Is Kill(by Takeshi Obata and Hiroshi Sakurazaka)are classified into the battle genre, Ping Pong(by Taiyo Matsumoto) and GIANT KILLING(by Tsujimoto) are classified into sports genre, Prophecy(by Tetsuya Tsutsui) and The Shadow Spirit(by Natsuhiko Kyogoku and Aki Shimizu) are classified into the mystery genre, and *ReRe* Hello(by Toko Minami) and Snow White With The Red Hair(by Sorata Akizuki) are classified into the romance genre.

In the battle, sports and romance genres, the experimental participants selected one or two scenes as a spoiler. On the other hand, in the mystery genre, the participants selected some pages as a spoiler. The reason for this result on the mystery genre would be that there are many highlighting scenes such as scenes about a mystery, mysterious killer, trick, solution of a mystery and so on. Thus, in the mystery genre, each comic has many mysteries and these mysteries drawn more than three pages. Then, the participants could not select all the spoiler pages.

As seen in Figure 7, the participants frequently chose the continuous pages, so it seems that several pages illustrated one spoiler content. Therefore, it is necessary to reconsider how to instruct the participants to select a spoiler in comics.

We found from Figure 7 that the comics can be classified into two types. One type tends to increase the score of spoiling (YuYu Hakusho, All You Need Is Kill, and GIANT KILLING, Ping Pong) and the other type tends to decrease the score of spoiling (ReRe Hello, and Snow White With The Red Hair). In our previous study, we also classified these comics into two types. One type is comics in which spoiler is acceptable (spoiler-acceptable comics), and the other type is comics in which spoiler is unacceptable (spoiler-unacceptable comics). Note that the spoiler-acceptable comics and the comics that tend to increase the score of spoiling are similar. In addition, the spoiler-unacceptable comics and the comics that tend to decrease the score of spoiling are similar. It seems that there is a relationship between the classification of comics into spoiler acceptable and spoiler unacceptable and the way of gingering up the comic stories.

3.2 Characteristics of content on a spoiler page



Figure 7: The spoiling score of each page in each comic.

In the battle genre, there were specific terms on the pages chosen as a spoiler that are related to the result of the battle, such as "*I won*" or "*Finished*." In addition, on the pages close to a spoiler page, bigger frames were used than on other pages.

In the sports genre, in the selected pages, there was character information which clarifies Win-Lose. The difference between the sports genre and the battle genre can be seen in the types of characters selected. In the sports genre, the numeric character is selected as the spoiler. For example, in the *GIANT KILLING* and *Ping-Pong*, there are little numeric characters in near the spoiler page. In addition, there are big frames near the spoiler page.

On the other hand, in the comics of the mystery genre, there are no specific terms selected as a spoiler. However, there are many big frames near the spoiler page. In particular, in *The Shadow Spirit*, the size of terms is small on the first, fourth and fifth-ranked spoiler page. In the *Prophecy*, there are small terms in the second, third, and fourth-ranked spoiler page.

In the romance genre, there are terms "I love you" and "I know" near the spoiler page with a big frame. All of the above suggests that there are specific terms, a specific frame size, and a big frame near a spoiler page.

4. HOW TO CONSTRUCT SPOILER DATASET

From the experimental results and pre-analysis of spoilers, we conclude that the spoiler effect depends on how much the book is read. For example, when he/she is reading volume 1 of a long story, he/she cannot identify spoiler information in volume 20 and may ignore the spoiler. Therefore, to construct spoiler dataset, we must take hold readers' condition to select a spoiler page.

In our previous spoiler dataset, we asked the participants to select three spoiling pages and to rank these pages subjectively. Since there was a possibility that the participants' criteria for judgment were different, we could not choose the real spoiler page in each comic. In addition, it might be very difficult for the participants to reduce the number of the spoiler pages up to three (see Figure 7). In the first place, they could not find any spoiler pages in a comic. In addition, the participants had a conflict to judge what page should be a spoiler.

In order to solve these problems, we must set a rule whether the page is the spoiler page or not and take hold a condition to think about a spoiler.

At first, we define facing pages as a spoiler in a comic. In addition, we will enable the participants (readers) to set three levels of spoiler in each facing pages in a comic. Next, to take hold readers' condition, if facing pages in the next story of a comic reduces the enjoyment of readers who already read up to the most current story in the comic, we define that the facing pages is the spoiler. For example, if a reader already read until the story number N in a comic, the reader evaluates whether each page is a spoiler or not in the story number N+1 in the comic.

Here, it is difficult for the participants to judge whether the page is the spoiler or not by thinking about the degree of reducing their enjoyment. Then, we will change the direction as follows; "Please evaluate the degree of spoiling in each page based on the following basis. Imagine the situation that you show facing pages in the story number N+1 to your friends who have read the comic until the story number N and have not read the story number N+1. If you assume that your friends will have an awful time by reading the page, the page is a spoiler. If you assume that your friends will have no thought by reading the page, the page is not a spoiler." By changing the direction like this, we expect that the evaluation becomes easier for the participants than the previous one because it requires thinking about the relationship with others.

Furthermore, in the previous dataset, we asked the participants who have read the whole of the comic to decide spoiling pages. Here, it was difficult for these participants to judge spoiler pages because they knew everything about the story, such as who will die, what will happen, how things go and so on. In the future, we will recruit new participants who have never read the target comic.

5. HOW TO ESTIMATE SPOILER PAGES

The pre-analysis of spoilers in the previous dataset showed several characteristics. However, it is not clear what kind of dimensions there are on the spoiler page.

It is considered that there are many features of spoiler pages, one of them being the frame size. There are small frames in the general scenes and big frames in the exciting scenes. Exciting scenes include scenes crucial to the story, such as battle results or truth about the incident, so these scenes look like spoilers. In other words, it might be possible to identify spoiler pages by their frame size. In the estimate frame method, Tanaka [7] proposed a frame estimating technique utilizing the direction of the concentration gradient. We are considering using this method to estimate the frame size.

Another feature is the terms in the speech bubbles or

cutline. Comics show the results using not only the picture but also the cutline and the characters' remarks. In the previous study [6], the spoiler page feature was the frame with the terms expressing the results of games or battles. So, we will extract the terms from each page and use these terms to estimate spoiler pages. As for the character recognition method, we are considering using Google's Vision API [8]. In addition, we are going not only to use the frame size and the terms but also to extract and use face information and bubble speech.

6. SUMMARY

In this paper, we attempted to clarify the effects of spoilers depending on the reading progress. We revealed that the degree of enjoyment was not changed even if the reader encountered a spoiler. Based on the results of the test on the interest in continuation, the number of the subjects who lost interest after seeing a spoiler was about four times higher than in the case where no spoilers were encountered. So, we conclude that the willingness to buy comics decreases if the reader sees spoilers.

We found many problems in the spoiler dataset. In this paper, we explained about how to construct the spoiler dataset. In the future work, we will construct the new spoiler dataset based on the method. In addition, from the analysis of spoiler pages, we discovered that the features of the spoiling pages were different depending on the genres. However, there are some common features around the spoiling pages, e.g., the size of the frame and the number of terms, and so on. In the future work, we plan to estimate the spoiler pages by using machine learning techniques.

ACKNOWLEDGMENTS

This work was supported in part by JST ACCEL Grant Number JPMJAC1602, Japan.

REFERENCES

- J. D. Leavitt and N. J. S Christenfeld; Story Spoilers Don't Spoil Stories. Psychological Science; 22 (9), pp.1152–1154, 2011.
- [2] W. H. Levine, M. Betzner and K. S. Autry; The Effect of Spoilers on the Enjoyment of Short Stories. Journal of Discourse Processes; 53 (7), pp.513–531, 2016.
- [3] D. Hassoun; Sequential outliers. The role of spoilers? Need for cognition, need for affect, and narrative selection and enjoyment. Psychology of Popular Media Culture; 4 (2), pp.346–358, 2013.
- [4] S. Nakamura and K. Tanaka; Temporal Filtering

System for Reducing the Risk of Spoiling a User's Enjoyment. 2007 International Conference on Intelligent User Interfaces; (IUI2007), pp.345–348, 2007.

- [5] S. Nakamura and T. Komatsu; Study of Information Clouding Methods to Prevent Spoilers of Sports Match. Advanced Visual Interfaces; (AVI 2012), pp.661–664, 2012.
- [6] Y. Maki and N. Satoshi; Do Manga Spoilers Spoil Manga? The Sixth Asian Conference on Information Systems; (ACIS 2017), 2017.
- [7] T. Tanaka, K. Shoji, F. Toyama, and J. Miyamichi; Layout Analysis of Tree-Structured Scene Frames in Comic Images, Proc. 20th International Joint Conference on Artificial Intelligence; pp.2885–2890, 2007.
- [8] Google Cloud Platform; Cloud Vision API; 2018.