# An E-Commerce Recommender System using Complaint Data and Review Data


**Toshinori Hayashi**
Kwansei Gakuin University
den82687@kwansei.ac.jp

**Yuanyuan Wang**
Yamaguchi University
y.wang@yamaguchi-u.ac.jp

**Yukiko Kawai**
Kyoto Sangyo University
kawai@cc.kyoto-su.ac.jp

**Kazutoshi Sumiya**
Kwansei Gakuin University
sumiya@kwansei.ac.jp



## ABSTRACT

In recent years, the use of e-commerce recommender systems has become more widespread, with key applications including tracking user purchase histories, considering value estimates and product review comments, and recommending higher-rated related items. However, traditional recommender systems are not consistent when recommending alternative items based on user input. Although users choose options on an existing product or service, (e.g., screen size and quality), it is still difficult to satisfy users' requirements. To solve this problem, we propose a novel item recommender system that analyzes two kinds of data: complaint data from the Fuman Kaitori Center and reviewer comments on e-commerce. First, the system generates the negative vectors of user-checked items from complaint data and positive vectors of related item data by subtracting lower-rated reviews from higher-rated reviews. Next, the system calculates the similarities between these two vectors and determines which reviews can resolve complaints related to user-checked items. Thus, the proposed system can provide suitable substitutes for user checked items. In this paper, we describe our proposed recommendation method based on complaint data and review data, and verify its efficacy using qualitative evaluation.



## ACM Classification Keywords

H.1.2 User/Machine Systems: Human information processing

## Author Keywords

E-Commerce; Recommender System; Complaint; Review;


## INTRODUCTION

As the prevalence of e-commerce has increased, many recommender systems have been proposed by researchers [1, 2, 5, 6, 9]. These systems help users to find items and promote their purchase. In addition, a user can view other consumers' product reviews to acquire information about a given item before purchasing. However, some reviews contain both positive and negative information. In these cases, users often cannot recognize genuine complaints about a given item. Moreover, even consumers find reliable information, they still face difficulty finding items that meet their needs. Thus, we propose a novel recommender system that extracts the drawbacks of products [4] and reviews. Consumer's needs tend to come from complaints; therefore, in our proposed system facilitates users finding suitable items based on their complaints. Finally, this paper is focused on complaints related to products (and not services). Hence, users can find suitable alternative items (item $Y_1$, item $Y_2$, and item $Y_3$) that will address complaints regarding their original purchase (item X) (see Figure1).


**Figure 1. Example of Recommendation**

## RELATED WORK

E-commerce recommenders have been extensively studied. Yandi et al. [7] proposed a recommender system that used coupons. In addition, different methods of recommending substitutes have been studied and are well-documented. McAuley et al. [3] considered the relationship between substitutes and complements based on items with reviews and their cost. They proposed methods for clarifying the relationship between substitutes and complements for a given topic. Zheng et al. [10] also analyzed the relationship between substitutes and complements by applying economic principles. In a similar vein, our proposed method recommends items based on complaint data and review data.

Finally, review bias has also been extensively studied. Zhang et al. [8] calculated review bias using 2 factors: user preferences and prejudices caused by reading other reviews. Since we propose a recommender system based on review data, we will analyze the reliability of said data.



*WII'18*, March 11, 2018, Tokyo, Japan

Table 1. Categories of Complaints

| Main | # of Sub | Main | # of Sub |
|---|---|---|---|
| Life | 14 | Vehicle | 14 |
| Fashion | 10 | Hobby | 13 |
| Food | 9 | Restaurant | 9 |
| Medical | 5 | Outdoor | 6 |
| Consumer Electronics | 6 | Industry | 14 |
| Health | 5 | Sightseeing | 3 |
| Public environment | 6 | Education | 10 |
| International Culture | 5 | Politics | 4 |
| Human relationships | 8 | Jobs | 5 |
| Pets | 6 | Other | 1 |

## DATASET OF COMPLAINTS

One of the datasets used with our proposed system is complaint data provided by Insight Tech Inc from the Fuman Kaitori Center website, which provides a platform for consumer complaints. Here, users post complaints with their own account on the website, receiving points each time they post. Furthermore, they can exchange these points for coupons, which can be used on purchases. Consumers post complaints regarding a wide range of products, services, and subjects.

Each complaint contains the following metadata: Posted user ID, Category, and Complaints. Our proposed system uses Complaints, Product_name, Category, and Sub_category. Since Product_name is not required, the system uses complaint data with Product_name.

Table 1 shows complaint categories, and each category has several subcategories. In total, these amount to 20 main categories and 153 subcategories. Users choose a category and subcategory when they post complaints.

## PROPOSED SYSTEM

### System Overview

The purpose of this system is to recommend alternatives to items with which users are dissatisfied. Figure 2 shows the overview of our proposed system. For item X, which is an item that the user peruses, this system finds item Y which solves one of the problems with item X.

To accomplish this, the system first extracts the item's corresponding complaint data and generates vectors from negative feature words. At the same time, the system extracts the review data of related items (item Ys) from e-commerce sites to generate positive feature vectors (we use related items from the same categories). However, many reviews contain both positive and negative words. To remove the negative feature words from the generated vectors, the system generates 2 vectors from high-evaluation and low-evaluation reviews. Next, by subtracting low-evaluation vectors from high-evaluation vectors, the vectors with positive feature words are calculated. Finally, the system finds a suitable alternative (item Y) by calculating similarity between negative feature vectors and positive feature vectors. It calculates the vectors using cosine similarity as follows:

$$Sim(v_F, v_R) = \frac{\sum w_n \cdot (wp_i - wn_i)}{\sqrt{\sum_{i=1}^{|V|}(w_n)^2} \cdot \sqrt{\sum (wp_i - wn_i)^2}} \quad (1)$$

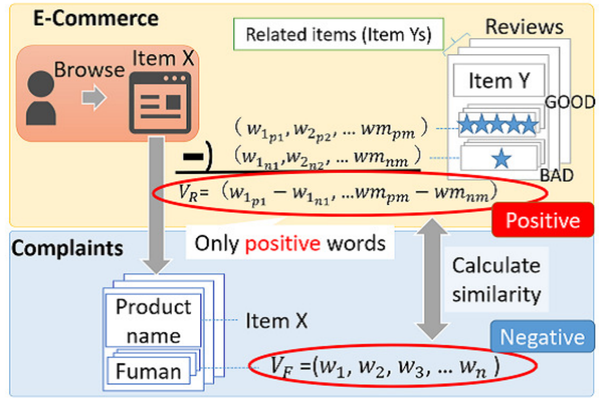


Figure 2. System Overview

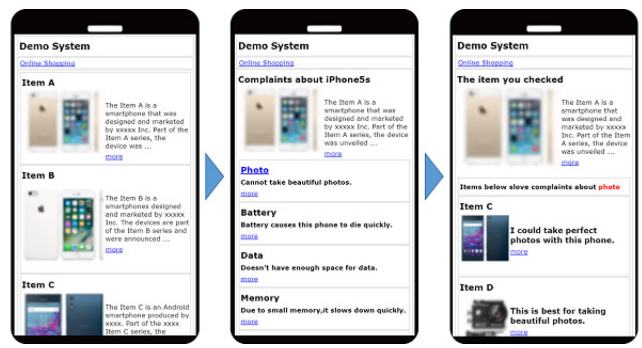


Figure 3. System Interface

$V_F$ denotes the vectors comprising negative feature words from complaints. $V_R$ denotes the vectors with positive feature words from reviews. In the next section we further explain this vector generation.

In basic terms, this system recommends other items which complaints for an item X by finding an item Y which has same feature words and has been positively evaluated. This system uses similarity values to rank item Ys for each complaint.

### Extraction of Negative Feature Words of Complaints

Our proposed system uses a complaint dataset. 50% of complaint data ia labeled "product_name" which denotes the object of a complaint. This system only uses complaint data which has "product_name" in its metadata.

First, the system extracts nouns from complaints by each item. Following this, it calculates the weight of each feature word using the term frequency–inverse document frequency (TF-IDF) method as follows:

$$tf_{i,j} \cdot idf_i = \frac{n_{i,j}}{\sum_k n_{k,j}} \cdot \log \frac{|D|}{df_i} \quad (2)$$

$d$ denotes the document that is integrated by all complaints for one item

### Extraction of Positive Feature Words of Item Review

This system analyzes items checked by consumers on e-commerce sites, and extracts reviews of related items (item Y) to identify positive and negative words related to both. For related items, this system generates feature vectors by extracting positive words from review data.

Table 2. Top 15 negative feature Words of item A with evaluation.

| Complaint | Evaluation Score | Review | Evaluation Score |
|---|---|---|---|
| Stereo | No-N | initial | No-N |
| music | No-N | friend | No-N |
| XPERIA | Negative | download | No-N |
| battery | Negative | Car navigation | No-N |
| car | No-N | self | No-N |
| sound | No-N | Sim | No-N |
| memory | Negative | Beauty | No-N |
| movie | Negative | Initially | No-N |
| USB | Negative | Packaging | Negative |
| sebum | No-N | necessary | No-N |
| LCD | Negative | Terminal | Negative |
| package | Negative | ROM | No-N |
| brightness | Negative | A moment | No-N |
| Skype | Negative | Upper-part | No-N |
| group | Negative | level | No-N |

Table 3. Average precision of item A to item D.

| Average precision | Complaint | Review |
|---|---|---|
| item A | **0.49** | 0.16 |
| item B | **0.46** | 0.41 |
| item C | 0.53 | **0.65** |
| item D | **0.36** | 0.27 |

First, this system extracts two types of item reviews (high-evaluation reviews and low-evaluation) for each item, and generates two types of vectors based on these reviews. Evaluation values are distinguished by the number of stars on five grading criteria. Reviews of 4 or 5 stars are evaluated as high, and reviews of 1 or 2 stars are evaluated as low.

To generate vectors, this system uses the same formula as (2). D denotes twice the number of same categorized items on an e-commerce site, as the system generates two vectors per item. Thus, the feature vectors $V_rp$ and $V_rn$ are generated.

$$v_{rp} = (wp_1, wp_2, \cdots, wp_i, \cdots, wp_m) \quad (3)$$
$$v_{rn} = (wn_1, wn_2, \cdots, wn_i, \cdots, wn_m) \quad (4)$$

Finally, this system subtracts the weight of low-evaluation vectors $V_rn$ from the weight of high-evaluation vectors $V_rp$ on each item. This calculation generates vectors containing only positive words for each item. Feature words with negative values imply negative subjects. This system does not require negative words for generated vectors, so it removes feature words with negative values. Thus, the generated vectors have only positive feature words, and each value implies the level of positivity for these feature word. For example, the feature words with values at or near 1 are considered very positive:

$$v_R = (v_{rp} - v_{rn}) = (wp_1 - wn_1, \cdots, wp_i - wn_i, \cdots, wp_m - wn_m) \quad (5)$$

### Interface

This system finds items with positive reviews relating to negative points of a user-checked item. Based on common feature words of the negative vectors of item Xs and positive vectors of item Ys, this system recommends alternative items for each consumer. These are ranked based on the calculated similarities between X and Y items. Figure 3 shows the interface of our proposed system.

The system displays items as general e-commerce websites (left, Figure 3). When users click on an item, its details are shown along with the number and nature of complaints posted by other users.(center, Figure 3). Finally, when the user clicks a type of complaint, the system shows an item that solves the problem expressed in that complaint (right, Figure 3).

Table 4. Top 10 negative feature words from complaints.

| Items | Top 10 feature words |
|---|---|
| item B | scratch, cover, Sheet, processing, book, Built in, Type, Grip sensor, player, flash. |
| item C | Exchange, Photo, Recently, input, character, capacity, battery, Radio wave, Repair, SoftBank. |
| item D | Call out, Going, number, phone fingertip, purpose, Partner, limit worry, History, browser, Incoming. |

Table 5. Top 10 positive feature words from reviews.

| Items | Top 10 feature words |
|---|---|
| item 1 | compact, heat, volume, position, button, sound, Conpact, Replacement, Authentication, Book |
| item 2 | Radio wave, going out, reaction, Real racing, Preparation, frequency, Pursuit, defect, heat, China |
| item 3 | charging, Body, Caution, scam, sim, Purchase, A case, Trouble, Upper-left, regret |
| item 4 | Professional, Notice, Completion, Reconstruction, charging, Excitement, Contact, Spoofing, Malignant, Free of charge |
| item 5 | Activation lock, OCN, SIM, Cheap, Safety, Lock, Profile, Relief, Sim, Connection. |
| item 6 | action, attachment, Abundance, image quality, camera, Attached, Travel, Success, Snorkeling, Record |
| item 7 | Windows, keyboard, board, PC, Key, computer, 64bit, cost, USB, Allowance |

## EVALUATION

Our evaluations were performed using two data types: complaint data (provided from Insight Tech Inc.) and review data from Amazon. Each item had approximately 50 complaint posts. In the Amazon data, each item had approximately 60 high-evaluation and low-evaluation reviews. We evaluated 4 items with 200 complaint posts, and 7 items with 420 reviews.

### Comparison of Complaints and Negative Review

We assumed that complaint data shows more negative features of items than review data. Users referred to opened reviews on e-commerce sites during their shopping; however, the contents of complaint data were more straightforward than those of review data (review data is closed and users tend to only post when complaining). To verify those differences, we evaluated via a five-subject questionnaire.

We extracted feature words from both complaint and review data for the same items from the same category (phones). Following this, we subtracted high-evaluation reviews from low-evaluation reviews to remove positive words.

**Table 6. Similarity between Negative Complaints and Positive Reviews.**

| Similarity | item A | item B | item C | item D |
|---|---|---|---|---|
| item 1 | **0.60** | 0.16 | 0.46 | 0.05 |
| item 2 | 0.40 | 0.17 | 0.46 | 0.20 |
| item 3 | 0.18 | 0.20 | 0.28 | 0.10 |
| item 4 | 0.43 | 0.27 | **0.60** | 0.06 |
| item 5 | 0.33 | 0.25 | **0.63** | 0.12 |
| item 6 | 0.03 | 0.06 | 0.23 | 0.02 |
| item 7 | 0.21 | 0.17 | 0.53 | 0.04 |

**Table 7. Recommendation on Proposed System of Each Complaint.**

| Values | item 1 | item 2 | item 3 | item 4 | item 5 |
|---|---|---|---|---|---|
| Complaint1 | 0.00 | 0.00 | 0.50 | 0.25 | **1.00** |
| Complaint2 | **1.00** | **0.60** | **0.8** | 0.40 | 0.20 |
| Complaint3 | 0.50 | 0.00 | 0.25 | 0.00 | **1.00** |
| Complaint4 | 0.00 | 0.50 | 0.25 | **0.75** | **1.00** |
| Complaint5 | 0.00 | **1.00** | **0.75** | 0.25 | 0.50 |

Table 2 shows the negative feature words of item A from both complaints and reviews. We also calculated for items B, C, and D, and verified the difference between the feature words from complaints and reviews via five-subject questionnaire.

Table 2 also shows the results of this evaluation for item A. "Negative" means they recognized as negative objects of the item. "No-N" means they did recognized as non-negative objects. We adopted majority rule to calculate precision for their different answers. Subjects evaluated each feature words after checking complaints and reviews of each item.

Table 3 shows the average precision from items A–D. The value of precision for extracting negative words from complaints is higher than that from reviews on all items except item C. Some words extracted from reviews on item C were too general. We could see more negative feature words extracted from complaints than from reviews. In future work, we must verify this discrepancy for many items to remove general terms on TF-IDF methods.

### Comparison of System and Results Manually

We evaluated this system compared to the results of manually selecting alternative items. We first calculated the number of recommended items using our proposed method, the complaint data of 4 items, and the review data of 7 items. The categories for these items are described below.

**item A to item D** :items from complaints categorized phone

**item 1 to item 5** :items from reviews categorized phone

**item 6** :items from reviews categorized Camera

**item 7** :items from reviews categorized PC

Table 4 shows the top 10 feature words we extracted for items B–D using complaint data. The feature words extracted for item A are shown in table 2. Most of these feature words imply the objects of complaints.

Table 5 shows the top 10 feature words extracted for items 1-7 using review data. It may prove beneficial to only extract positive words using our method; however, we will also examine extracting additional parts of speech in future work.

**Table 8. Average of Recommendation Evaluated Manually.**

| Values | item 1 | item 2 | item 3 | item 4 | item 5 |
|---|---|---|---|---|---|
| Complaint1 | 0.00 | 0.00 | 0.33 | 0.33 | 0.33 |
| Complaint2 | 0.50 | **0.83** | 0.33 | **0.67** | **1.00** |
| Complaint3 | 0.17 | 0.17 | 0.00 | 0.00 | 0.00 |
| Complaint4 | 0.00 | **0.83** | **0.67** | **0.83** | **1.00** |
| Complaint5 | 0.33 | **0.83** | **0.83** | 0.33 | **1.00** |

Table 6 shows the values of the similarities between negative vectors from complaints and positive vectors from reviews used to rank items. High values of similarity could be seen after normalizing for items A and C. However, the values for items B and D are lower because their vectors have many dimensions. We can also see that our proposed methods is effective for recommending other categorized items (such as cameras and PCs).

Table 7 shows the results of system recommendations by complaint for item A. For example, the system recommended 3 items for complaint 1 based on feature words. The values were calculated using the order of ranking score and similarity values between items.

- Complaint1: Complaint about battery
- Complaint2: Complaint about Internet connection
- Complaint3: Complaint about charger
- Complaint4: Complaint about photo and memory
- Complaint5: Complaint about reaction of screen

To evaluate our proposed system, we produced answer data for recommendations via questionnaire. 6 subjects rated whether items 1–5 could solve the problem of certain complaints. Subjects scored 1.0 if the recommended item could solve the problem of each complaint, and 0.0 if it could not. Table 8 shows the averages of the rating results.

The correlation coefficient value between the results of proposed system and the manual answer data was 0.45. We can see that our proposed system performs with similar accuracy to manual recommendation. In future work, we plan to add more factors to raise the correlation coefficient value.

## CONCLUSION

In this paper, we proposed a recommender system that uses complaint and review data to recommend alternative purchase items on e-commerce websites. We extracted feature words from complaints with low-evaluation reviews, and verified the efficacy of our proposed system with 200 complaints and 420 reviews. Our results showed that complaint data was an effective information source for our purposes, and that our proposed system performed well when compared with manual recommendation methods.

In future work, we plan to further validate the accuracy of complaint data by analyzing consumer complaints for many more items. Furthermore, we will consider new extraction methods to enhance the precision of the proposed system.


## ACKNOWLEDGEMENTS

In this paper, we used an FKC Data Set provided for research purposes by the National Institute of Informatics in cooperation with Insight Tech Inc.

## REFERENCES


1. L. N. Anh-Thu, H-H Nguyen, and N. Thai-Nghe. 2016. A Context-aware implicit feedback approach for online shopping recommender systems. In *Asian Conference on Intelligent Information and Database Systems*. Springer, 584–593.
2. U. Leimstoll and H. Stormer. 2007. Collaborative recommender systems for online shops. *AMCIS 2007 Proceedings* (2007), 156.
3. J. McAuley, R. Pandey, and J. Leskovec. 2015. Inferring Networks of Substitutable and Complementary Products. In *Proc. of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15)*. 785–794.
4. K. Mitsuzawa, M. Tauchi, M. Domoulin, M. Nakashima, and T. Mizumoto. 2016. FKC Corpus: a Japanese Corpus from New Opinion Survey Service. In *Proc. of the Novel Incentives for Collecting Data and Annotation from People: types, implementation, tasking requirements, workflow and results*. 11–18.
5. A. O Omondi and A. W Mbugua. 2017. An Application of association rule learning in recommender systems for e-Commerce and its effect on marketing. (2017).
6. B. Smith and G. Linden. 2017. Two Decades of Recommender Systems at Amazon. com. *IEEE Internet Computing* 21, 3 (2017), 12–18.
7. Y. Xia, G. Di Fabbrizio, S. Vaibhav, and A. Datta. 2017. A Content-based Recommender System for E-commerce Offers and Coupons. (2017).
8. Xiaoying Zhang, Junzhou Zhao, and John C.S. Lui. 2017. Modeling the Assimilation-Contrast Effects in Online Product Rating Systems: Debiasing and Recommendations. In *Proceedings of the Eleventh ACM Conference on Recommender Systems (RecSys '17)*. ACM, New York, NY, USA, 98–106. DOI: **http://dx.doi.org/10.1145/3109859.3109885**
9. Qi Zhao. 2016. E-commerce Product Recommendation by Personalized Promotion and Total Surplus Maximization. In *Proc. of the 9th ACM International Conference on Web Search and Data Mining*. 709–709.
10. J. Zheng, X. Wu, J. Niu, and A. Bolivar. 2009. Substitutes or Complements: Another Step Forward in Recommendations. In *Proc. of the 10th ACM Conference on Electronic Commerce (EC '09)*. 139–146.