Online Shopping Recommender System Using Hybrid Method

Ade Romadhony
Fakultas Informatika
Institut Teknologi Telkom
Bandung, Indonesia
ade@ittelkom.ac.id

Said Al Faraby
Fakultas Informatika
Institut Teknologi Telkom
Bandung, Indonesia
saf@ittelkom.ac.id

Bambang Pudjoatmodjo
Fakultas Informatika
Institut Teknologi Telkom
Bandung, Indonesia
bbp@ittelkom.ac.id

***Abstract*— As Internet usage becomes very common nowadays, including in Indonesia, people do their activities by relying on information gathered from world wide web. One of the examples is gathering information before buying product. Customers will search the product reviews before deciding whether to buy the products or not. Product reviews can be found by searching through search engine, reading on personal blogs, accessing from certain websites that provide reviews, or searching on electronic forum. Because searching on the net can be a daunting task, we propose a system that can provide product reviews, together with recommendations, focusing on Indonesia market. The recommendations will become a valuable resource for customers to narrow the search space on searching the specific products suitable for their needs.**

We took a sample on two groups with different characteristics. The focus of this research is to design good personal recommendations. Our system generates two types of recommendation: personal and item-based. User-based collaborative filtering is implemented to produce personal recommendation. And we implement item-based collaborative filtering on item's recommendation. The evaluation has been taken by conducting a survey on the users. The users should state whether they like the recommendation or not. The result shows that most of the users like the item-based recommendations. But the personal recommendation is only preferred by the user who already did some activities: rate or view an item.

Keywords—recommender system; personal; user-based; item-based;collaborative filtering *(key words)*

I. INTRODUCTION

Higher online transaction is one of the biggest effects of easier Internet usage in Indonesia. Nowadays, it is very common that people place the web as the source of information to buy or sell something. Therefore, many online shops exist, in different forms, from website with private domain to thread in online forum. This condition leads to advantages and also disadvantages for customer. The main advantage is customer has more options to perform buying. But it can also be disadvantageous because with so many options, customer will find difficulties to choose which product that is more suitable, and which shop that can be trusted, has good customer service, and also offers good price. Thus, the main problem is there is no one-stop place to search comprehensive information about online selling and buying. The needed information related to online selling and buying is list of products, list of online shops, and also recommendations about choosing product and shop.

In order to perform online transaction, customer has to trust online shop/seller. Trust is seen as long term relationship between buyer and seller [1]. Trust to seller/shop is often built by references from other customers. Trust building highly depends on what so called "*word of mouth communication*", which will influence customers to choose product or shop [2]. According to a research about Indonesian Online Shopping Behaviour [3], there are four main factors influencing customer to do online shopping: safety, comfort, quality sites, and similar attitudes and behavior among the users. But searching information about product or shop with regard to those factors is not an easy task, and it can make customers spend more time on browsing. Even more than 80% of online customers leave the web page before they find what they need. [4].

Another proof that trust is very important for customer is the highest amount of online transaction recorded in Indonesia that happens on Facebook and Kaskus [5]. In Indonesia, there is no a website like www.amazon.com which contains review and recommendation about a product or shop. Customers still have strong belief in social media and online community (forum). Kaskus is claimed to be the largest online community in Indonesia. But discovering recommended product or shop in social media or online forum is not easy because the data is very large and not well-structured. One possible solution to solve the problem about difficulty to search information related to choosing suitable product and online shop is a recommender system. With a recommender system, customers can choose suitable products, and recommended online shop easier. The sellers will get advantages also because the more customers recommend them, the more new customers they will get.

Generally, there are two popular methods on generating recommendation: content-based filtering, and collaborative

978-1-4673-4992-5/13/$31.00 ©2013 IEEE

filtering. Content-based filtering has some disadvantages, such as there is no subjectivity covered [6]. Subjectivity factor is covered in collaborative-filtering, but the performance will not be good enough for new item addition that has no rating yet [7]. Collaborative filtering has been used widely in e-commerce area, including product recommendation application. There are two approaches on collaborative filtering: user-based and item based. Amazon uses item-based collaborative filtering to produce similar/related item recommendation.

A web-based online shopping recommender system prototype has been developed to simulate customer activity on searching product/shop information, rate product/shop, and get recommendations [8]. Recommendations in that prototype are generated based on web usage mining using apriori algorithm [9], and item-based collaborative filtering. But the quality of recommendation is not good enough (below 50% of users who like the recommendations they get), especially in personal recommendations. So the focus of this study is to improve the system so it can produce better personal recommendation quality.

II. SYSTEM OVERVIEW

The application has been uploaded on the Internet and can be accessed at www.panduanbelanja.com

A. Functionalities

The functionalities of application developed are as follows:

1. Rating and review input from customers
   User can give rating and review to a product or shop. The rating ranges from 1 to 5. For shop's rating, users rate it based on some criteria, such as: price competitiveness, shipping, and customer service.
2. Product recommendation
   User gets personal recommendations on the main page after login, and collaborative recommendations in product- detail page.
3. Product and shop search
   User can easily look for detail information on specific product or shop.
4. Favorite products and shop list
   This list is a recommendation for non member user, and new user who has no view and rating activity yet. This list is generated based on the highest and most frequent rated products and shops.
5. Most viewed product list
   This list is a recommendation for non member and new users also. It is generated based on the most frequent viewed products by members.

B. User characteristics

Users in this study were chosen by snowball sampling [10], generally classified into two groups:

1. Woman, age ranging from 20 to 35 years old, who frequently viewed and rated products in baby/kids stuff and woman stuff category.
2. Man and woman, age ranging form 18 to 23 years old, who frequently viewed and rated products in electronic stuff category.

Both groups could be considered as suitable representation of Indonesia majority as online shopping buyer..

III. RECOMMENDATION PROCESS

There are two recommendations generated on the system developed: personal and related item recommendation. Recommendations are generated based on user activity (viewing and rating on a product). Based on viewing and rating data, item similarity and user similarity matrix will be built. The algorithm to generate item similarity can be seen in **Algorithm 1** and **Algorithm 2**.

```
For each product on list, P1
   For each user U who rate P1
      For each product P2 rated by U
         Save that a user rated P1 and P2
   For each product P2
      Compute similarity between P1 and P2
(based on rating)
```

Algorithm 1 Item-Based Collaborative Fiktering Based on Rating

```
For each product on list, P1
   For each user U who view P1
      For each product P2 viewed by U
         Save that a user viewed P1 and
P2
   For each product P2
      Compute similarity between P1 and P2
(based on viewing)
```

Algorithm 2 Item-Based Collaborative Filtering Based on Viewing

And the algorithm to generate user similarity matrix can be seen in **Algorithm 3** and **Algorithm 4**.

```
For each user on list, U1
   For each product P rated by U1
      For each user U2 rate product P
         Save that a product rated by U1
and U2
   For each user U2
      Compute similarity between U1 and U2
(based on rating)
```

Algorithm 3 User-Based Collaborative Filtering Based on Rating

```
For each user on list, U1
   For each product P which viewed by U1
      For each user U2 view product P
         Save that a product viewed by U1
and U2
   For each user U2
      Compute similarity between U1 and U2
(based on rating)
```

Algorithm 4 User-Based Collaborative Filtering Based on Viewing

This research is funded by Dikti Indonesia, under AUPT scheme

Item and user similarity computed based on cosine similarity formula [11] is defined on (1).

$$\text{Similarity}\left(\vec{A},\vec{B}\right)=\cos\left(\vec{A},\vec{B}\right)=\frac{\vec{A}\bullet\vec{B}}{\|\vec{A}\|\times\|\vec{B}\|} \quad (1)$$

Here is the explanation of cosine similarity usage. For item similarity based on rating, vector A represents all user rating on item A (0 if a user has not rated the item yet) and vector B represents all user rating on item B. The explanation is similar for other similarity computation.

A. *Personal Recommendation*

For personal recommendation generation, the input is user similarity, based on rating and viewing activity. And the procedures to generate the recommendation are as follows:

1. Check if a user has similarity with other user (similarity > 0). Pick 3 other users with highest similarity, referred as similar users. Products that will be recommended are the highest rated products which have been rated by similar user, but have not been rated by the user.
2. If a user has no similar user based on rating, check if he/she has similar user based on view. If user has similar user based on view, recommended products are most frequent viewed products, but has not been rated by user.
3. If a user has no similar user based on view also, so the recommendations are favorite products.

A sample snapshot of personal recommendation for a user can be seen in **Figure 1**.

Figure 1 Snapshot of Personal Recommendation

B. *Related Item Recommendation*

Related item recommendations are generated based on similarity score between items. The first step is checking on item similarity table based on rating. Recommended products are six products with highest similarity score (> 0) and in the same category. If there is no similar product based on rating, the system will find on item similarity matrix table based on viewing. The recommended products are also restricted to product from similar category. This procedure is to ensure that recommended products have close relationship with current viewed product. If there is no similar product based on viewing also, so the recommendations are products with highest rating on the same subcategory/category. A sample snapshot of related item recommendation can be seen in **Figure 2**.

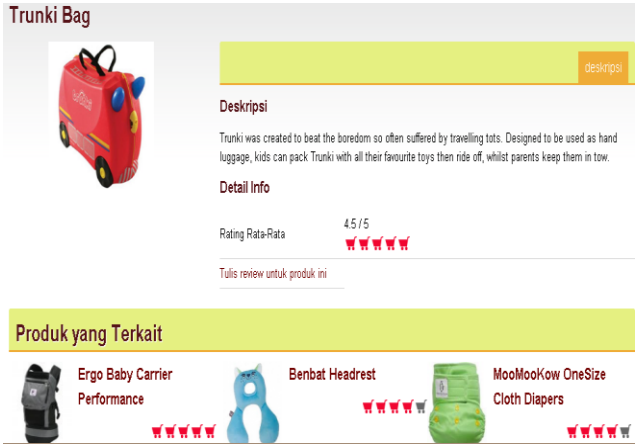

Figure 2 Snapshot of Related Item Recommendation

IV. EVALUATION

Recommendation of both personal and related items has been evaluated through user study method [12]. Users are asked to answer some questions regarding recommendations they get. Here is the list of questions:

1. Are personal recommendations you get suitable for yourself? (Answer options: yes, no, yes for some products, no for the rest)
2. Do related item recommendations you get have related property with currently-view product? (Answer options: yes, no, yes for some products, no for the rest)
3. Are you interested in clicking the personal recommendation item? (Answer options: yes, no, yes for some products, no for the rest)
4. Are you interested in clicking item on related product section? (Answer options: yes, no, yes for some products, no for the rest)
5. Do rating score influence you to be interested in a product? (Answer options: yes, no)

Totally 20 users are participated in this user studies. The characteristics summarized in **Table 1**.

Table 1 Characteristics of User Studies Participants

Sex	56% female, 44% male
Category preferences	50% electronic, 5% apparel, 45% baby and kids
Total rated products	25% rated >=5 products, 75% rated < 5 products

And the answers from participants have been summarized in **Table 2**.

Table 2 Summary of User Studies Answers

Question	Answer		
	Yes	Yes for some products, No for the rest	No
1	30%	35%	35%
2	65%	10%	25%
3	65%	10%	25%
4	50%	50%	-
5	100%	-	-

Based on user studies result, it can be concluded that personal recommendations still need improvement. Most of users who answer no or some yes some no for personal recommendations are new users or users with similar users but the similarity score is low. New users with no or few activity tend to have no similar users. And if the user similarity score is low, then the probability of recommended product suitable for user will be low.

As for related items recommendation, some comment from users is related to category or subcategory relationship. The majority of users prefer that related items should have very close relationship, for example, they are the same products with the same brand, but different type. It happens probably because of misunderstanding about the term used, "related items", and there is no explanation why those items are related to each other.

V. DISCUSSION

Based on evaluation result, it can be concluded that there is improvement on recommendation quality, better than prior prototype developed. A special case that still needs improvement is for new item addition, or never been rated, and for new user. Pseudo rating could be used to overcome new item or item with no rating. New user and low user similarity score cause poor personal recommendation performance as related/collaborative recommendation. Therefore, other strategies to produce better recommendation are still needed, such as by clustering users based on user profile. The user profile can be acquired explicitly or implicitly by user's activity.

As for case domain, this system is prepared to be widely used by online shopper, so we hope to be able to gather large enough dataset. But the number of users who participate actively on giving rating is still low, so we think that we might consider recommendation method concentrated on small dataset for the next system enhancement.

REFERENCES

[1] Implications of Buyer Decision Theory forDesign of Ecommerce Situs. International Journal of Human-Computer Studies, 55, (2001), 815-844.

[2] Chrysanthos Dellarocas and Ritu Narayan. What motivates consumers to review a product online? A study of the product-specific antecedents of online movie reviews

[3] Nanang Husin. Internet User Behavior Analysis in Online Shopping on Indonesia. ICACSIS 2011, 137-142

[4] Young Ae Kim and Jaideep Srivastava. Impact of Social Influence in E-Commerce Decision Making. ICEC'07, August 19–22, 2007, Minneapolis, Minnesota, USA. Copyright 2007 ACM 978-1-59593-700-1/07/0008

[5] eCommerce in Indonesia. Presentation by Veritrans and Daily Social. August 2012

[6] M. Balabanovi and Y. Shoham. Fab: content-based,collaborative recommendation. Commun. ACM, 40(3):66{72,1997.

[7] G. Linden, B. Smith, and J. York. Amazon Recommendations: Item-to-Item Collaborative Filtering. Internet Computing, 7(1):76{80, 2003

[8] Ade Romadhony. Pemberian Rekomendasi Layanan Belanja Online. Thesis. Program Magister Sekolah Teknik Elektro dan Informatika Institut Teknologi Bandung. 2010

[9] Ziming Zeng. An Intelligent E-Commerce Recommender System Based-on Web Mining :International Journal of Business and Management, Vol 4. No.7 2009

[10] Snowball Sampling, Experiment-resources.com, http://www.experiment-resources.com/snowball-sampling.html

[11] G. Linden, B. Smith, and J. York. *Amazon Recommendations: Item-to-Item Collaborative Filtering.* Internet Computing, 7(1):76{80, 2003}.

[12] D. Jannach, M.Zanker, A.Felfernig, G.Friedrich. Recommender Systems : An Introduction. Cambridge University Press. 2010