Proceedings of the International MultiConference of Engineers and Computer Scientists 2014 Vol I,
IMECS 2014, March 12 - 14, 2014, Hong Kong

# Extracting Characteristics of Items Based on Patterns in Recommendation Graphs

Daisuke KITAYAMA, Kazutoshi SUMIYA

***Abstract*—On online shopping sites such as Amazon and Rakuten, recommended items are displayed along with the items being viewed. We consider that certain recommended items reflect the characteristics of the viewed item. For example, "DVD-R" may be recommended with "Printer," whereas "Printer" might not have the recommended item "DVD-R." In this case, we may assume that the item "Printer" can print a label on a "DVD-R." Thus, a set of items can be expressed as a directed graph structure, which comprises items as nodes and recommendation relations as edges. In this paper, we propose a method for extracting item characteristics based on the patterns in recommendation graphs. We also present an evaluation based on comparisons among items.**

***Index Terms*—Characteristics extraction, Online shopping, Recommendation graph, Recommendation system**

## I. Introduction

IN recent years, online shopping sites such as Amazon[1] and Rakuten[2] are being used increasingly. These sites include a recommendation function based on social filtering [1], [2]. These recommendation functions reflect the behavior of customers. We consider that most customers view and purchase related items based on their personal requirements. For example, "Printers" are bought for paper, photo, and label printing. However, a printer may have "Paper," "Digital camera," and "DVD-R" as recommended items. Thus, we consider that the recommended items may reflect the characteristics of the viewed item.

A set of items may be expressed as a directed graph structure, which comprises items as nodes and recommendation relations as edges. We refer to this graph as a recommendation graph. We consider that an item with specific characteristics forms a pattern with other similar (or related) recommended items in the recommendation graph. For example, a digital camera might only be recommended with a printer. Thus, we may assume that this printer can print photos from digital cameras. A pair of binoculars might only be recommended with a raincoat. Thus, we may assume that this pair of binoculars can be used in rainy weather. In contrast, a printer ink cartridge may be recommended not only with a specific printer but also with many other similar printers. We may assume that the printer has generality in its functions. Therefore, we propose a method for extracting the item characteristics based on the patterns in recommendation graphs. We use the shape of the recommendation graph and the item categories to detect patterns.

Manuscript received January 8, 2014; revised January 30, 2014.

D. Kitayama is with the Department of Computer Science, Faculty of Informatics, Kogakuin University, Japan, e-mail: kitayama@cc.kogakuin.ac.jp.

K. Sumiya is with the School of Human Science and Environment, University of Hyogo, Japan, e-mail: sumiya@shse.u-hyogo.ac.jp.

[1]http://www.amazon.com

[2]http://www.rakuten.com

We developed an approach for comparing item functions using our method based on similar items recommended when a user views an item. Our method can extract other items that reflect the characteristics of a viewed item based on the recommended items. Therefore, a user can make comparisons with similar items based on the item's characteristics.

The remainder of this paper is organized as follows. We describe our approach and other work related to this topic in Section 2. In Section 3, we explain our method for extracting characteristics using a recommendation graph and item categories. We evaluate and discuss our experimental results in Section 4. Finally, we discuss our method in Section 5.

## II. Our Approach

### A. Recommendation graphs and item categories

Recommendation functions are implemented based on a user's purchase history and the items they viewed in previous sessions. Therefore, we consider that the recommended items reflect a user's buying behavior. For example, a DVD-R may be recommended with a printer. In this case, we might assume that the printer has the function of printing DVD labels. Recommended items may reflect the characteristics of items as well. Recommendation functions show the item descriptions with item reviews and the number of item purchases by other users. A pair of binoculars is described based on its specifications and its function on online shopping sites, and user opinions of the item are provided in reviews. A pair of binoculars may also have recommended items related to the appreciation of concerts such as music CDs and Japanese fans. Thus, we may assume that the particular pair of binoculars is suitable for use at music concerts. Therefore, recommended items may reflect the implicit characteristics of items.

We define recommendation graphs and item categories as follows. A recommendation graph is a set of items with a directed graph structure, which comprises items as nodes and recommendation relations as edges. The edges connect items to recommended items. Two-way edges mean that items are recommended with each other. Figure 1 shows an example of a recommendation graph. The printer A in the center of Figure 1 has recommended items such as a DVD-R and other printers. Printers A and C have a two-way recommendation relationship.

We use the item categories to detect patterns. The item categories are defined by the online shopping site. For example, a printer may be within the category: "*Electronics > Computers & Accessories > Printers > Inkjet Printers*," whereas a DVD-R may be within the category "*Electronics > Accessories & Supplies > Blank Media > DVD-R Discs > DVD-R*." All items must belong to some category. If


Fig. 1. Example of a recommendation graph.

the target item and the recommended item are within the same category, they may be competitors or could belong to a series of items. However, if the target item and the recommended item belong to different categories, they may have a relationship that reflects their usage. Therefore, we use recommendation graphs and item categories to detect patterns that reflect the characteristics of items.

### B. *Related work*

Tsukuda et al. [3] proposed a method for the discovery of unexpected information based on subjective words. They calculated the degree of unexpectedness based on a graph analysis using Wikipedia. As an example, they analyzed the relationship between subjective words and coordinate words, and the degree of recognition of related words based on the degree of unexpectedness. This concept of graph analysis is similar to our method, although our approach includes the analysis of other aspects such as the item categories used in online shopping sites.

Hijikata et al. [4] proposed a discovery-oriented collaborative filtering method for users that was based on a recommendation process. Our method can also identify items unfamiliar to users. Therefore, their method may have a complementary relationship with our method.

Shoji [5] experimented with improving an interface to support decision making during online shopping. They targeted users who decided not to purchase items. They proposed a communication model based on sales clerks. Our method aims to present information related to items, which is based on their communication model.

## III. Extracting Relations Based on Patterns

### A. *Calculating the degree of specificity*

We use the patterns in recommendation graphs and the item categories to determine the characteristics of items. In this section, we explain the degree of specificity of recommended items, which is calculated from the recommendation graph, for a base item. We consider that an item has a high degree of specificity if the item is recommended from only the base item. The degree of specificity is calculated using the following expression:

$$Score(s,t) = 1 - \frac{|SimItem(s) \cap RecItem(t)|}{|SimItem(s)|} \quad (1)$$

where function $SimItem$ returns a set of items that are similar to item $s$. In this paper, we define similar items as items within the same category. Function $RecItem$ returns a set of items that contains the recommended item $t$. Using this expression, we can calculate the degree of specificity as $Score(s,t)$ based on the ratio of similar items with item $s$ to the number of items recommended with item $t$. The degree of specificity is between 1.0 and 0.0. If the number of items recommended with item $t$ is low, the degree of specificity is

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online) 

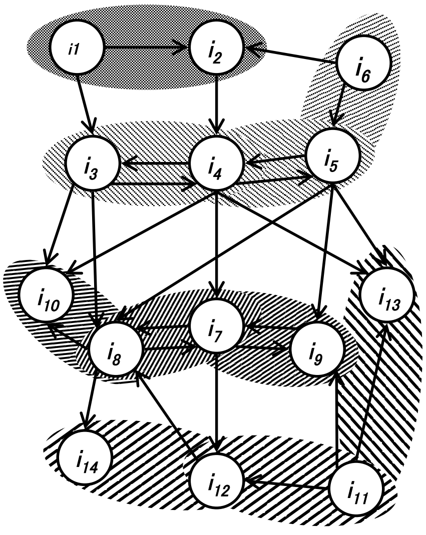


Fig. 2. Recommendation graphs and item categories.

high. Conversely, if the number of items recommended with item $t$ is high, the degree of specificity is low.

A high degree of specificity means that the recommended item reflects specific characteristics such as a DVD-label-printing function. For example, item $s$ is printer A and item $t$ is DVD-R. The DVD-R is only recommended with printer A. However, printers B, C, and D are similar to printer A, but they do not have DVD-R as a recommended item. In this case, the DVD-R has a high degree of specificity for printer A. We illustrate this in detail in Figure 2, which shows a recommendation graph where the edges represent the recommendation relationship and the edges are drawn from a base item to a recommend item. Items in the same item categories have the same color. In this case, we assume that $i_4$ is the base item $s$. The similar items $i_3$, $i_5$, and $i_6$ are within the same category. The recommended items for $i_4$ are $i_7$, $i_{10}$, and $i_{13}$. For these items, only $i_7$ is not recommended with similar items of $i_4$. In other words, $i_3$, $i_5$ and $i6$ do not have $i_7$ as recommended items. Thus, the degree of specificity for $i_7$ is 1.0.

A low degree of specificity means that a recommended item shares common characteristics such as printing paper. We determine the specificity characteristics of recommended items using the threshold $\alpha$ and common characteristics using the threshold $\beta$.

### *B. Patterns among the degrees of characteristics and the item categories*

We extract the item characteristics using the degree of specificity and the item category relation between item $a$ and $t$. We assume that item $s$ is a printer and that item $t$ is copy paper. Item $t$ is often recommended with items similar to $s$, but item $s$ and item $t$ belong to different categories. Thus, item $t$ shares a common characteristic with item $s$, i.e., printing paper.

However, we may assume that item $t$ is a popular printer. In this case, item $t$ is also recommended with items similar to $s$, but item $s$ and item $t$ belong to the same category. In this case, item $t$ may have inferior characteristics compared with item $s$. If the price of item $t$ is lower than the price of item $s$, we can estimate that item $s$ has a reasonable price. Similarly, if the printing speed of item $t$ is higher than the printing speed of item $s$, we can estimate that item $s$ has an inferior printing speed.

In this manner, we can extract the item characteristics using a combination of the degree of specificity and the item category relations between item $a$ and $t$. Table I shows different combinations of the degree of specificity and the item category relations.

TABLE I
PATTERNS AMONG THE DEGREE OF SPECIFICITY OF CHARACTERISTICS AND THE ITEM CATEGORIES

| | | Degree of specificity | |
|---|---|---|---|
| | | high | low |
| Item categories | same | This recommended item shares specificity characteristics with the flanking items. | This recommended item shares important common characteristics with the item category. |
| | different | This recommended item shares specificity characteristics with the item. | This recommended item shares common characteristics with the item. |

TABLE II
EXPERIMENTAL DATA: ITEM CATEGORIES

| Item category | Number of items |
|---|---|
| Grocery & Gourmet Food > Pasta & Noodles > Noodles | 10 |
| Grocery & Gourmet Food > Packaged Meals & Side Dishes > Rice Dishes | 10 |
| Toys & Games > Stuffed Animals & Plush | 10 |
| Electronics > Computers & Accessories > Printers | 10 |
| Electronics > Camera & Photo > Binoculars, Telescopes & Optics > Binoculars | 10 |
| Electronics > Camera & Photo > Digital Cameras | 10 |

## IV. EXPERIMENTAL

### *A. Experimental setting*

We calculated the degree of specificity using an actual recommendation graph based on Amazon.co.jp to evaluate our method. We collected items and recommended items using the Amazon Product Advertising API and constructed a recommendation graph. We selected some items as seeds and collected recommended items within two hops. Thus, we collected 9,848 seed items and 95,050 recommended items. In the experiment, we used 60 seed items from which we selected 10 items from each of six categories (see Table II). We calculated the degree of specificity using these data.

### *B. Experimental results*

We present examples of the experimental results in Tables III, IV, and V. The tables show the names of the recommended item, their degree of specificity, and whether they belong to the same categories.

Table III shows the result for a printer that can print paper up to A3 size. The results show that A3 papers have a high degree of specificity. In addition, a digital camera and papers used for printing photos have a high degree of specificity. This printer can also be used to print photos. These items belong to different categories, so we can determine that these items share specificity characteristics

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE III
EXPERIMENTAL RESULTS: CANON INK-JET PRINTER PIXUS IX6530, A3 PAPER, FIVE-COLOR, DOUBLE BLACK INK, COMPACT BUSINESS MODEL

| Recommended item | Degree of specificity | Category |
|---|---|---|
| RICOH digital camera G700, 12.1 M pixel optical zoom: X5, wide lens: 28 mm, waterproof: 5 m anti-shock: 2 m, dust-proof | 1.000 | different |
| Canon Photo paper gloss gold class A3, 20 paper GL-101A320 | 1.000 | different |
| Canon Photo paper gloss gold class A3+, 20 paper GL-101A3N20 | 1.000 | different |
| Canon high-class HR-101 A3+ HR-101A3N0BI | 1.000 | different |
| Canon copy-paper white A3 | 1.000 | different |
| Canon ink-tank BCI-326 (BK/C/M/Y) + BCI-325 multi-pack BCI-326+325/5MP | 0.959 | different |
| Canon ink-jet multifunctional machine PIXUS MX893 auto duplex printing, ADF-equipped FAX wired and wireless LAN | 0.951 | same |
| ELECOM refill ink CANON BCI-325&326-compatible set of 5-color THC-MG5230RSET | 0.894 | same |

TABLE IV
EXPERIMENTAL RESULTS: CASIO DIGITAL CAMERA EXLIM EX-ZS6 BLACK EX-ZS6BK

| Recommended item | Degree of specificity | Category |
|---|---|---|
| Sanwa Supply DG-BGM2BK compact multi-case (Black) | 0.994 | different |
| CASIO NP-80/NP-82-compatible charger | 0.975 | different |
| CASIO digital camera charger (USB-AC adapter) AD-C53U | 0.963 | different |
| CASIO EXILIM EX-Z270 EX-Z1 NP-80/NP-82 compatible battery | 0.963 | different |
| CASIO EXILIM digital camera case EX-ZS6 ZS10 ZS12 ZS20 N1 N10 | 0.963 | different |
| CASIO digital camera EXLIM EX-ZS6 silver EX-ZS6SR | 0.956 | same |
| CASIO digital camera EXLIM EX-ZS6 pink EX-ZS6PK | 0.938 | same |
| HAKUBA digital pouch CS S-size Black SDP-CSS-BK | 0.925 | different |
| ELECOM 2.7 inch covering film for liquid crystal gloss DGP-007G | 0.906 | different |
| Transcend SDHC card 16 GB Class10 TS16GSDHC10E | 0.300 | different |

with the printer. By contrast, a multifunctional machine has a high degree of specificity and shares the same category as the printer. We can determine that the common functions of this multifunctional machine and the printer share specificity characteristics. This multifunctional machine and the printer used the same type of ink cartridge, so these items are flanking products. Therefore, we consider that the calculated degree of specificity is appropriate. However, an ink cartridge that shares common characteristics with printing paper has a high degree of specificity. Ink cartridges have different model numbers for each printer, so the same ink cartridges are not recommended for similar printers. Thus, we need to consider grouping similar items based on the category hierarchy when calculating the degree of specificity.

Table IV shows the results for a digital camera. An SD memory card has a low degree of specificity in these result. SD memory cards belong to different categories for digital

TABLE V
EXPERIMENTAL RESULTS: SALT-FLAVORED RAMEN WITH A PICTURE OF A WHITE BEAR IN MARUYAMA ZOO PRODUCED BY FUJIWARA NOODLE-MAKING FACTORY 112.4 G $\times$ 10 PACKS

| Recommended item | Degree of specificity | Category |
|---|---|---|
| Mobile battery for multi-device Cheero Power Plus 2 10400mAh (Black) | 1.000 | different |
| Salt-flavored ramen with a picture of a white bear in Maruyama Zoo produced by Fujiwara noodle-making factory 112.4 g $\times$ 10 packs [special price] [limited number] | 0.970 | same |

cameras. Therefore, we can determine that this SD memory card shares common characteristics. Indeed, SD memory cards are often used to store photo data in digital cameras. Therefore, SD memory cards share the function of capturing photos with digital cameras.

Table V shows the result for instant noodles. A mobile battery has a high degree of specificity in these result. We consider that the mobile battery and the instant noodles have no relation, but our method cannot exclude incorrectly recommended items. In this case, the mobile battery was recommended with the instant noodles accidentally. We need to examine the amount of recommended items that lack relations and confirm the effects of these recommended items.

Thus, we confirmed that our method could detect item characteristics using recommendation graphs and item categories. In addition, we considered that our method may detect item compatibility. In table IV, SD memory is a compatible item for a large number of digital cameras, on the other hand, batteries, chargers and camera cases are compatible items for particular digital cameras. Naturally, the example of a digital camera and compatible SD memories is tedious. In table III, the digital camera “G700” is a compatible item for photo printers in general. However, we think that this digital camera “G700” is an incompatible item for other photo printers, because this “G700” and the printer “PIXUS IX6530” have a peculiar function of “PictBridge (Wireless LAN)” that is a direct connecting function between a digital camera and a printer via Wireless LAN without PCs. We considered that our method can discover not only explicit item compatibility but also implicit item compatibility.

### C. Application: Items’ relation viewer

Our method can be used to support item selection on online shopping sites. This function could display the item characteristics extracted using our method. The extracted characteristics can be included as item details that are not described in the item descriptions and reviews. Therefore, users can view similar items based on their extracted characteristics, as well as the item descriptions and reviews.

Figure 3 shows the interface of this application. When a user searches for printers, an online shopping site such as Amazon.com shows similar printers. In this case, our function provides the extracted characteristics with the recommended result. In this example, copy paper was extracted from printer A, B and C as common characteristics, and then a digital camera was extracted from printer C as specific characteristics. The users can then estimate that printer A can

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

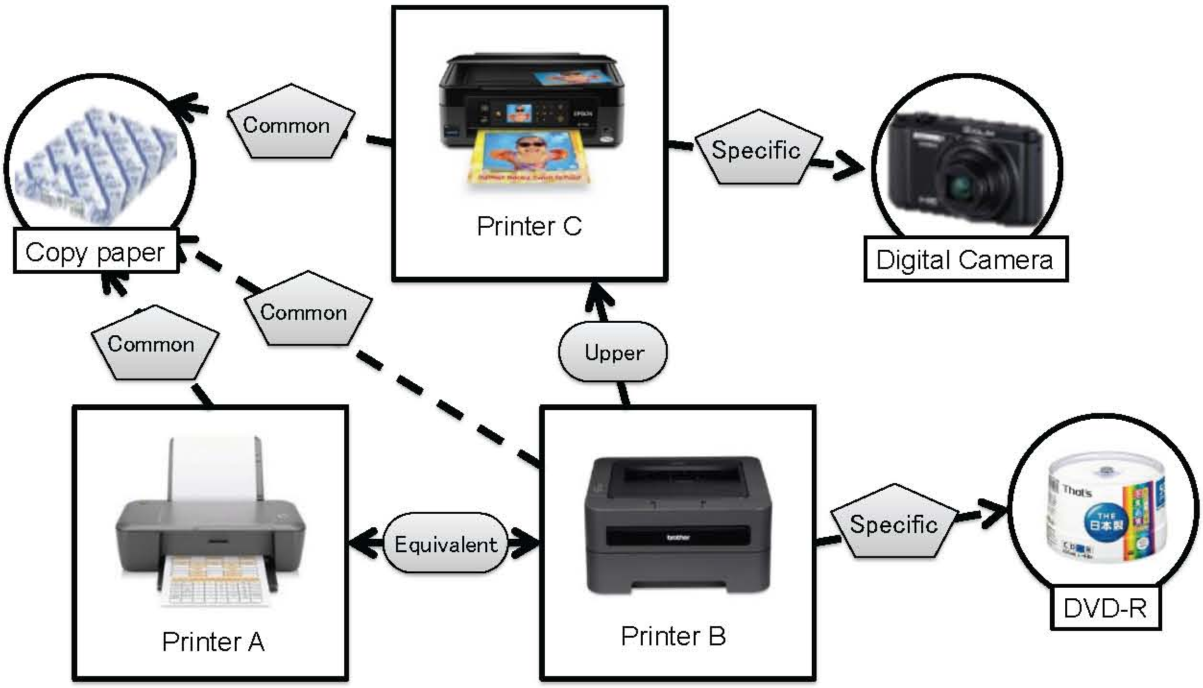


Fig. 3. Application: Items' relation viewer.

print using the paper as a common characteristic. However, printer C can print photos as a specific characteristic. If the user wants a photo printer, the user can select printer C.

## V. Conclusion

In this study, we proposed a method for extracting item characteristics based on the patterns in recommendation graphs and item categories. We identified items with common and specific characteristics as the target items. We also described the application of this method to support item selection on online shopping sites. We confirmed that our method could detect item characteristics based on recommendation graphs and item categories. We shown that a possibility of extracting implicit item compatibility such as the "PictBridge (Wireless LAN)" function between the digital camera "G700" and the printer "PIXUS IX6530". However, our method could not exclude incorrectly recommended items. We need to examine the amount of recommended items that have incorrect relations and confirm the effects of these recommended items.

In this paper, we use the Amazon Product Advertising API, however, this API only return 10 recommended items. Therefore, our recommended graph is a very sparse graph. We considered that interest characteristics of item are appeared in dense graph. Thus, we plan to develop a corroborative filtering system using dataset such as movie lens and Netflix data. Then, we will evaluate our method on developed recommender system as dense recommended graph.

We aim a knowledge discovery from recommender systems. We considered that recommendation results include customers' knowledge. Because recommendation algorithms using customers' behavior. Therefore, we consider that our method can use not only online shopping but also other recommender systems such as music, movie, recipe and so on. We will try to adapt our method to other recommender systems.

*Acknowledgment:* This research was supported in part by a Grant-in-Aid for Young Scientists (B) 24700098 from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

## References

[1] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, GroupLens: open architecture for collaborative filtering of Netnews, In Proc. Conference on Computer Supported Cooperative Work, pp. 175–186, 1994.
[2] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, Item-based collaborative filtering recommendation algorithms, In Proc. 10th International Conference on World Wide Web, pp. 285–295, 2001.
[3] K. Tsukuda, H. Ohshima, M. Yamamoto, H. Iwasaki, and K. Tanaka, Discovering unexpected information on the basis of popularity/unpopularity analysis of coordinate objects and their relationships, In Proc. 28th ACM Symposium on Applied Computing (SAC 2013), pp. 878–885, 2013.
[4] Y. Hijikata, T. Shimizu, and S. Nishida, Discovery-oriented collaborative filtering for improving user satisfaction, In Proc. International Conference on Intelligent User Interfaces 2009 (IUI2009), pp. 67–76, 2009
[5] H. Shoji, S. Conart, An interaction method that facilitates concept articulation in shopping online, AI & Society, Vol. 19, No. 1, pp. 65–83, Springer-Verlag, 2005.