Buy Eye-mask Instead of Alarm Clock!: Graph-based Approach to Identify Functionally Equal Alternative Products

Tsukasa Hirano[1], Yoshiyuki Shoji[1,2][0000−0002−7405−9270], Takehiro Yamamoto[3][0000−0003−0601−3139], and Martin J. Dürst[1][0000−0001−7568−0766]

[1] Aoyama Gakuin University,
Sagamihara, Kanagawa 252–5258, Japan
hirano@sw.it.aoyama.ac.jp, duerst@it.aoyama.ac.jp
[2] Shizuoka University,
Hamamatsu, Shizuoka 432–8011, Japan
shojiy@inf.shizuoka.ac.jp
[3] University of Hyogo,
Kobe, Hyogo 651–2197, Japan
t.yamamoto@sis.u-hyogo.ac.jp

Abstract. This paper proposes a method to analyze product reviews to identify other products that can achieve the product's intended use. When people want to achieve the purpose of "getting up early in the morning," they generally tend to look only for an "alarm clock." In contrast, there are several products such as "smart curtains" that can directly achieve the purpose and "sleeping pills" that can indirectly achieve the purpose by replacing "getting up early in the morning" with "going to sleep early at night." The proposed method constructs a bipartite graph comprising product and uses purpose information from product review data. Then, the random walk with restart technique is employed to rank other products that can satisfy the use purpose of the input product. The proposed method was evaluated in subjective experiments, and the results suggest that the method is both accurate and useful in terms of identifying alternative products that satisfy similar use purposes.

Keywords: Information Retrieval · Review · Product Recommendation · Alternative Achievement

1 Introduction

Tunnel vision, which is also referred to as narrow-mindedness, is a significant obstacle in terms of product selection while shopping. Tunnel vision is a perceptual phenomenon whereby an individual experiences a significant reduction in peripheral vision. As a result, the individual focuses predominantly on their central vision, similar to looking through a tunnel. Most people have bought something they did not need on impulse due to tunnel vision at least once in their life. For example, after oversleeping and becoming upset, it is common for people to impulsively purchase an alarm clock in order to wake up at the necessary time. Conventional recommendation algorithms recommend various alarm clock devices when the user starts searching for an "alarm clock." In this case, the user can compare many "alarm clocks" to find the most suitable device for the given purpose, *i.e.*, waking up at the correct time.

However, solutions to avoid oversleeping are not limited to alarm clocks. For example, if you take sleeping pills the night before and go to bed early, you may reduce the risk of missing sleep the next day, and if you buy a pillow that allows you to sleep soundly, you may wake up more refreshed the next morning, even after a short sleep. As a result, when a person focuses singularly on alarm clock products, they are much more likely to purchase the product without considering any information about alternative products that are functionally equivalent." In many situations, consumers get tunnel vision and focus their attention on only a single purpose, thereby making it impossible to find an appropriate alternative product.

With the increasing ubiquity of the Internet and smartphones, various online shopping services and customer-to-customer marketplaces have become increasingly common in recent years. As a result, a wide variety of consumers can purchase desired products easily.

According to a survey conducted by the Japanese government, 73.4% of Japanese consumers have purchased items on an online shopping site. By age, 69.5% of those in their twenties and 78.0% of those aged 60 and older (regardless of gender or specific age) are using e-commerce sites [4].

A problem that can occur when casually shopping on the web in this manner is impulsive purchasing caused by tunnel vision. Online shopping services have made it considerably easier to purchase various products; however, it has also become increasingly difficult for users unfamiliar with the Internet to select the most suitable products. In fact, Zhang *et al.* [13] demonstrated that visual appeal and traditional recommendation methods create impulsive purchase intent in online shopping situations. To prevent this, it is necessary to present information that encourages logical thinking in consumers searching for a product to make them think about the purpose they need to address and the product they need once they have found it.

Generally, conventional product recommendation systems [5] have used collaborative filtering, which first recommends products of the same type, other products purchased by users who purchased that product, and products used in combination. Thus, if the user selects "alarm clock," another type of alarm clock or a related battery for such devices will be presented as a recommendation. Here, we consider the case where the user searches for an alarm clock to facilitate waking up early in the morning consistently and reliably. Many products can achieve

[4] The Ministry of Internal Affairs and Communications Japan: Usage of Internet shopping and auction flea markets (in Japanese) `https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r03/pdf/n1100000.pdf`

this purpose, such as optical clocks, smart curtains, and smartwatches; however, conventional recommendation systems will limit the displayed recommendations to various types of alarm clocks. In other words, the user cannot find a way to achieve the given purpose other than purchasing an alarm clock. Thus, the user is likely to select an alarm clock even when other suitable options are available. In terms of recommendation systems, this problem can be solved by identifying the purpose of the purchase as a query, *e.g.*, "getting up early in the morning."

Thus, in this paper, we propose a product recommendation algorithm that outputs a ranking of different functionally equivalent products by defining the purpose in a query. For example, if the target product is an alarm clock, the proposed algorithm will show candidate purposes, *e.g.*, "wake up early in the morning," "avoid being late," and "make a loud sound." The user can select a single purpose, *e.g.*, "waking up early in the morning," which is then taken as the input to the proposed algorithm. Here, the proposed algorithm determines if each product in the dataset is suitable for the identified purpose. The algorithm then identifies, ranks, and displays specific alternative products that can achieve the given purpose, *e.g.*, "morning light," "pillows," and "supplements." To achieve this, we focused on product reviews to realize a product recommendation system that takes the purchase purpose as the query input. Note that most online shopping services allow users to post reviews of products, which may include information about how the purchaser used the product.

We consider using the purpose contained in product review information to identify products with the same purpose and present them as alternatives. Specifically, assume that an algorithm recommends products that directly include the words "awake early in the morning" in their reviews. Such an algorithm would not present alternative products, which is the goal of the proposed algorithm. The first problem is that review information frequently contains spelling errors and grouping the same purposes is impossible. The second problem is that a direct search may output common or similar products as effective alternatives. For example, if the query is "alarm clock," and the purpose is to "avoid being late," it is obvious that many other types of alarm clocks will be output, *e.g.*, "light clocks" or "wrist watches with alarms." However, these alternative products do not broaden the user's perspective or change the thinking that an alarm clock is a singularly best product to avoid being late. Thus, we created a bipartite graph with the product, segmented the purpose's word groups as nodes, and performed a random walk with restart (RWR) calculation. Using bipartite graphs and RWR allows us to discover a wide range of alternative products, which cannot be realized using conventional search systems. In the context of the "alarm clock" example, by performing an edge transition for the query "prevent being late," an indirect purpose, *e.g.*, "improve the quality of sleep" may be discovered, and various products will be output, *e.g.*, "pillow" and "eye-mask."

To demonstrate the effectiveness of the proposed method, we conducted evaluation experiments using accurate data. The proposed method outputs a ranking of alternative products for as input for the purpose. In the experiments, participants were asked to label how the output products could achieve their purposes, how useful they were for their purchase decisions, and the overall accuracy of the product rankings. The usefulness of the proposed method was revealed by evaluating each of the rankings and products.

The remainder of this paper is organized as follows. In Section 2, we summarize related studies and present the position of this study. Section 3 describes the proposed method, and Section 4 describes an experiment conducted to evaluate and compare the proposed method to a baseline method. The experimental results are then discussed in Section 5. Finally, the paper is concluded in Section 6.

2 Related Work

We propose a product recommendation algorithm to identify and display alternative products for a given purpose. Thus, this study is closely related to the study of product recommendation algorithms and studies related to achievement.

2.1 Recommendation and Search of Products by Purpose

In this study, we attempt to identify products based on the purpose of the product, which is related to the study of product recommendation systems. In a previous study that considered the relationships among products, Ruining *et al.* [2] proposed Monomer, a method that treats product relationships as multiple concepts and allows recommendations that are unrestricted by categories or symmetry. This study differs because the proposed method treats the relationship between products as a single concept that utilizes the relationship between product reviews.

We attempted to identify and collect the purchase purpose from product review information. In a study that used product reviews, Lei *et al.* [14] proposed a product recommendation algorithm that divides reviews into two types, *i.e.*, one for the user's actions, and one for the product, and uses a neural network that shares a layer trained on each review. In addition, the recommendation system proposed by Sopheaktra *et al.* [12] takes product features as input and extracts and uses purposes related to the features from reviews using LDA and Word2Vec. This study differs because the proposed method does not utilize product characteristics. Instead, the proposed method seeks products based on the relationship between the purposes and products. Previous studies have also investigated product recommendation algorithms that focus on the purpose of use. For example, McAuley *et al.* [6] proposed a recommendation system that divides the products they recommend for purchase into interchangeable substitutes and complementary products purchased in addition to the identified interchangeable replacements. The current study differs in that the product is a different type of product that can realize the given purpose.

2.2 Serendipity of Recommendation

The proposed method differs from conventional recommendation systems in that it attempts to recommend products that can have a different perspective. Thus, this study is related to the study of recommending unexpected products [3]. For example, Kotkov *et al.* defined serendipitous product recommendations as those with three specific elements, *i.e.*, relevance, novelty, and unexpectedness.

As an example of serendipitous product recommendation, Kensuke *et al.* [7] proposed a recommendation method that uses a bipartite graph of users and items, as well as a bridging score that represents the degree of association between the nodes and the degree of the anomaly of the nodes. In addition, Akiyama *et al.* [1] constructed a human preference model of serendipity using data from a questionnaire about serendipity, and they proposed a recommendation method ranked by the length of item distance.

2.3 Achievement Products for the Purpose

The proposed algorithm outputs an alternative product by finding possible alternative purposes for the product. To extract potential alternative purposes, Pothirattanachaikul *et al.* [8] proposed a method to extract alternatives that can achieve the purposes from community question answering (CQA) sites. For example, assume that we are focusing on the behavior of “taking sleeping pills” to achieve the purpose of “solving sleep problems.” In that case, we present an alternative behavior of “taking a walk before bedtime.” The relationship between actions and purposes is a bipartite graph that utilizes the question-and-answer information from the CQA site. In addition, the similarity is ranked to discover alternative measures.

In addition, Yamamoto *et al.* [10] and Yang *et al.* [11] proposed a method to identify alternatives based on the relationship between the primary purposes and sub-purposes. Here, they defined “exercise” and “diet pills” as sub-goals to achieve the primary goal of “losing weight,” and they proposed a method to cluster the sub-goals using sponsored search data. Yang *et al.* defined subtasks as “selecting a hotel” and “recruiting volunteers,” *etc.*, which are required to accomplish the main task of “organizing a meeting,” and they proposed a method that connects queries with the functions described in wikiHow. Based on the definitions of these alternatives, the proposed method performs calculations to identify alternative achievement proposals.

3 Graph-based Method to Identify Alternative Products

This section describes a method that inputs the purpose for why a user wants to buy a certain product, and then outputs a ranking of alternative products that can satisfy the purpose. To realize this algorithm, the proposed method extracts sentences from product review data corresponding to the use purpose. We hypothesize that products with the same purpose are similar and substitutable. Based on this hypothesis, we assume that a graph-based computation is appropriate, and we identify potential alternative products by constructing a graph that represents the relationship between products and purposes.

3.1 Extracting Purposes from Product Review Data

In the proposed method, product review information is used as data to obtain the purpose of using a given product. The proposed method extracts phrases from strings, primarily using syntactic patterns. Note that the dataset considered in this study is in Japanese; thus, we explain the extraction method using Japanese grammar and vocabulary. This process differs depending on the target language; however, for grammatical reasons, it is generally easier in English.

In the review text, only a few parts of the text are related to the purpose; thus, the proposed method utilizes syntactic patterns to isolate the purposes from the review text data. Here, a two-step process is employed to identify the purpose. The first step involves extracting the purpose based on language patterns, and the second step involves selecting sentences using morphological analysis.

Prior to extracting the purpose, the review data are cleaned in a preprocessing step. For example, sentences that mention shop, shipping, and price are removed. We also remove sentences that contain terms that are irrelevant to the purpose, *e.g.*, "postage" and "wrapping."

An overview of the proposed purpose extraction method is illustrated in Figure 1. First, the purposes are extracted using regular expressions. Here, if the extraction utilizes typical syntactic patterns, the sentences may become redundant. Redundant sentences are those that contain many words besides the purpose word, which can result in an excessive number of nodes when using graph processing–based algorithms.

Thus, we must extract on the required words and construct subgraphs that are effective for the corresponding calculation. In Japanese, due to the grammatical order of words, it is necessary to specify the words before and after to be interposed when extracting the purpose phrase in a language pattern. The front patterns include "when" ("*-toki*" in Japanese), and the back patterns include "for this purpose" ("*-tame-ni*" in Japanese). For example, consider the sentence "*... neru toki, kuraku suru tame-ni ...*" ("When I sleep, I want to make the room dark." For this purpose, I bought an eye-mask."). Here, the phrase "*kuraku suru*" ("make the room dark") is extracted. We manually collected 15 eligible terms for the front and back patterns, and we used their combination. Sentences containing the purchase purpose are frequently followed by words like "utilize" ("*ri-you*" in Japanese); thus, the extraction process is performed when these words are found after the matching sentence. In addition, sentences exceeding a certain length are excluded because redundant sentences may be extracted even when a language pattern matching process is employed.

Next, morphological analysis is performed to filter out only those sentences extracted by syntactic patterns that contain the purpose. Note that some of these syntactic patterns are commonly used expressions. For example, if a phrase is extracted with the language pattern "For," it may express a purpose, as in "For sleeping well," or it may refer to an object, as in "For children." In such cases, it

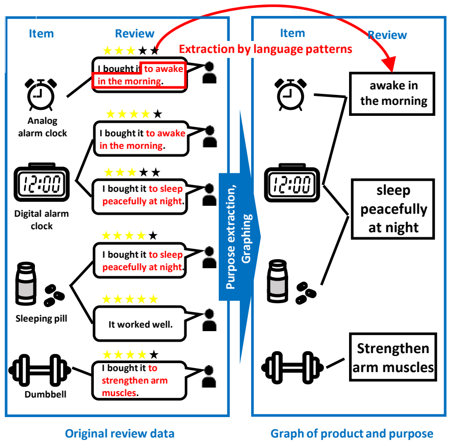

Fig. 1. Overview of generating product-purpose graphs using language pattern extraction.

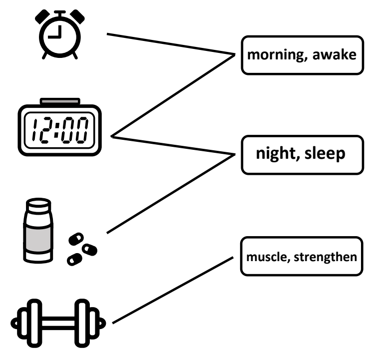

Fig. 2. Example of nodes created by splitting purpose into words.

is necessary to classify the extracted sentences by focusing on the part of speech of the words. In other words, in this case, if it were a verbal noun phrase, it would be a purpose, and if it were a noun, it would not be a purpose.

For each syntactic pattern, we created a list of the parts of speech from which the corresponding purpose could be extracted. Specific to the Japanese language, any noun can be converted into a verb using the suffix "-*suru*," and such verbs were treated as purposes by extracting the noun part.

Some other verbs related to "be" verbs and presence were listed and removed, *e.g.*, "is," "are," "do," and "exist" (in Japanese: "*aru*," "*iru*," "*suru*," and "*yaru*").

In this study, we used the MeCab morphological analysis library[5]. In addition, as the word dictionary, Mecab-ipadic-NEologd[6], which includes new words and proper technical nouns, was used for the analysis.

3.2 Creating Nodes Comprising Multiple Purpose Terms

The proposed method constructs a bipartite graph using the extracted purposes; thus, the extracted purpose phrases are converted into a set of several terms to be used as nodes.

Note that these data can contain many different representations of the same purpose. For example, the simple purpose of "getting up early in the morning"

[5] MeCab: `https://taku910.github.io/mecab/`
[6] Mecab-ipadic-Neologd : `https://github.com/neologd/mecab-ipadic-neologd`

can be expressed as “waking up before sunrise” or “early rising.” However, the size of the graph would increase significantly if each of these unique expressions are used as nodes. Thus, we extract and lemmatize keywords from the purpose phrases. In the previous example, the words “wake,” “early,” and “morning” should be extracted. The extracted set of terms is then used as a node.

When extracting important keywords, short sentences are first divided into words (note that Japanese is not separated by white space; thus, it is necessary to perform morphological analysis). Then, some words are removed according to the part of speech, *i.e.*, only noun, adjective, and verb terms are kept, and all other terms are removed. Then, as part of the lemmatization process, all verbs are set to the standard form, and nouns are set to the singular form. In addition, other keyword candidates are removed based on Japanese unique frequent terms. Here, very common words and uncommon words are removed in this step because they do not contribute to score propagation (even if treated as nodes in the graph). Finally, the graph is constructed, as shown in Figure 2.

3.3 Creating Bipartite Graph of Product and Purpose

Our algorithm assembles a bipartite graph using the created nodes. Figure 3 shows an overview of the proposed method. The algorithm is based on the hypothesis that if two products share many of the same achievable purposes, they may be substitutable for each other. Thus, the proposed method attempts to identify products that can achieve the same purpose by representing and calculating the relationship between similar products and purposes as a bipartite graph. Here, in the constructed graph, the reviewed products and word groups included in the purpose are nodes, which are connected by edges. The graphs are created as follows.

- Select a single product purpose and search for a product that has an edge to a word node contained in the purpose.
- For all obtained products, we search for products that have an edge to a word node in each purpose.
- The graph is constructed with the products obtained using the above process and the purpose word groups used in the search as nodes.

Rather than creating a graph using all product and purpose nodes, we select a particular product’s purpose and create a subgraph using only the product and purpose nodes obtained from the two searches. This process is performed first because if all nodes are used, the size of the graph matrix will become too large to compute. Second, we attempt to prevent the edges from becoming less relevant as they are followed. If all nodes are used, there is a risk that they will be calculated as being related to a particular product, even if they are dozens of edges away. This is different from the substitutability defined in this study.

In addition, if a node is explored as it is, a recommendation of a purpose related to another purpose of the first product is made. For example, assume that a pillow has the purpose of “improving the quality of sleep at night” and “curing

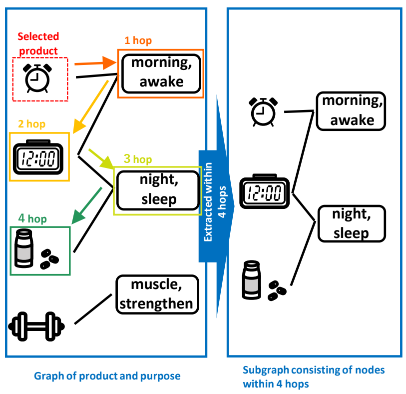

Fig. 3. Subgraph creation with graph search and extraction.

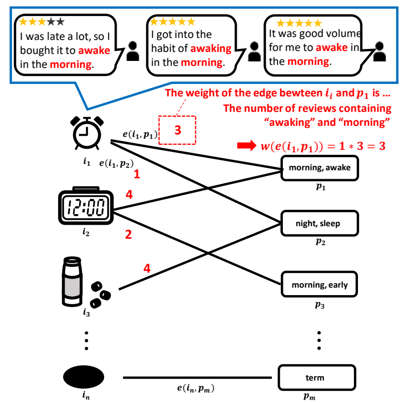

Fig. 4. Edge weighting using review texts containing purposes.

stiff shoulders," and that a search that selects the former finds the product. In this case, the following search finds the pillow again, thereby resulting in a search concerning the purpose of "curing stiff shoulders. " For this purpose, the edge to the first product is removed for the second and subsequent searches.

Here, let A be the adjacency matrix, let I be the set of commodity nodes, and let P be the set of destination nodes. In this case, matrix A comprises $(|I| + |P|) \times (|I| + |P|)$ dimensions, where the upper right and lower left of the matrix indicate the relationship from the product to the purpose and from the purpose to the product, respectively. Let $L_p(i_n)$ denote the set of P linked to i_n, and let $L_i(p_m)$ denote the set of I linked to p_n. In the upper right of the matrix, if I_n is contained in $L_i(p_m)$, the value of A_{nm} is 1; otherwise, the value is A_{nm} is 0.

3.4 Edge Weighting

The weight of each edge is set prior to performing the graph-based computation. Figure 4 shows an overview of the weighting technique used in the proposed method. Here, weighting is applied to focus on important words related to the product. Thus, weighting is performed using the occurrence frequency of words in the review text data. Some reviews were not extracted as purposes but contained the exact words as the extracted purposes. Thus, using the occurrence frequency of the purpose's words, it is possible to determine which words are essential. Here, if $|F_{nm}|$ is the frequency by which word p_m appears in the reviews of product i_n, the weight $w(e(i_n, p_m))$ of the edge $e(i_n, p_m)$ is expressed as follows:

$$w(e(i_n, p_m)) = |F_{nm}|. \tag{1}$$

3.5 Random Walk with Restart

The method outputs a ranking of products that can achieve the desired purpose by calculating the degree of relevance between the nodes in the graph. Although many algorithms have been proposed to calculate the degree of association between nodes, the proposed method employs the RWR technique. Here, RWR is used in the proposed method based on the hypothesis that products have similar purposes and that the purposes of a given product have some relationship with each other. For example, "dish detergent" is unsuitable for removing stains from baths and cars; however, it can be used on glass. In other words, we can infer that another detergent used on glass can be used on dishes. These relationships can be obtained as values using the RWR calculation.

In the following, we describe the relevance calculation method using the RWR technique. First, the adjacency matrix A, which represents the graph of products and target word groups created up to the 3.4 clause, is transformed into a transition probability matrix. This transformation is performed by normalization, where each column is divided by the sum of the weights of the edges departing the node. Here, if the transition probability matrix is A', the normalization equation is defined as follows:

$$A'_{ij} = \frac{A_{ij}}{\sum_{k=1}^{|I|+|P|} A_{kj}}. \tag{2}$$

4 Evaluation

A subject experiment using real product data acquired from an e-commerce site (Rakuten Ichiba) was conducted to evaluate the usefulness of the proposed method. Two separate evaluations were considered in this experiment, one for the top-ranked alternatives and one for the rankings.

4.1 Dataset

In this experiment, we used product review data taken from Rakuten Ichiba, which is one of Japan's most popular e-commerce sites. The review data include review sentences, product names, and category data, and such data can be used to construct appropriate graphs. However, some of the review data were unsuitable for our purpose; thus, the review data were filtered to eliminate inappropriate data.

First, we excluded categorical data. There are 39 categories in the dataset, and each product is divided into large groups, *e.g.*, "clothing," "home appliances," and "sports." The review data for products in the categories specified here were excluded. Examples of the excluded categories are only alternatives, *e.g.*, "clothing" and "food," and those that cannot be substituted, *e.g.*, "real estate" and "tickets."

Next, we excluded cases where the same review text was attached to multiple products. Products with the same review were excluded if they had the same review text at least once because this would affect the quality of the graph computation. After removing review data according to these processes, we obtained a dataset with a total of 44,834,987 reviews, and the number of products with review data was 180,910.

4.2 Implementation

The purposes were extracted from the acquired review data using regular expressions and MeCab, as described in Section 3.1. Finally, we obtained a total of 339,515 purposes.

Next, the words were segmented using MeCab to create word groups for the nodes. Here, only nouns, verbs, and adjectives were used, and the verbs were standardized. Finally, a total of 132,015 purpose nodes were obtained.

Then, the graph for the calculations was created. The preprepared query was divided into words, and the purpose node matching each word group was acquired. Then, we acquired the products related to the obtained word groups and searched for the products again using the word groups related to the products. A matrix representing the graphical relationships was created using the products and purposes obtained from these processes.

Finally, The RWR method was used to calculate the relevance of the product to the purpose. Here, the probability of a random jump to the query node was set to 0.8, and recursive calculations were performed until the results converged.

4.3 Compared Method

We prepared several variants to evaluate the usefulness of the proposed method.

1. **Proposed method**. This method calculates the relevance of the product-purpose graphs using the RWR technique.
2. **Unweighted graph**. Similar to the proposed method; however, edges are not weighted in this variant.
3. **Product description similarity**. This method calculates the similarity between the query and product descriptions using Doc2Vec.
4. **Keyword match**. This method finds products that contain all query terms in its review.
5. **Random**: This method selects products with reviews randomly.

4.4 Evaluation of Alternative Products

We prepared 25 queries and obtained product rankings in advance using the methods described in the previous section. Here, the participants labeled the top five items in all output rankings. The evaluation items included “Is it a direct solution?”, “Is it an indirect solution?”, “Did it help you in your purchasing decision?”. In this evaluation, a four-point scale was used, with four being the most applicable, and one being the most minor practical.

Table 1. Precision of each method and each evaluation factor ($** : p < 0.01, * : p < 0.05$ to **Product Descriptions** method)

Methods	Direct solution	Indirect solution	Purchase decision	# Found
Proposed Method	**0.44	0.05	**0.50	759
Unweighted Graph	****0.46**	**0.06**	****0.53**	759
Product Description	0.34	0.04	0.14	-
Keyword Match	**0.43	0.04	**0.41	24
Random	**0.02	0.00	**0.03	-

4.5 Evaluation of Ranking

Three methods that enable ranking were considered in this study, *i.e.*, the proposed method, unweighted graph calculation, and the product description similarity calculation. These three methods were divided into two cases, *i.e.*, with and without category refinement, and a total of six methods were used to label the output rankings. The evaluation items included "Have you broadened your perspective?", "Is there diversity?", "Did you get motivated to buy products?", "Comprehensive evaluation." Here, a four-point scale is used, with four being the most applicable, and one being the most minor practical.

4.6 Experimental Results

In this study, the average of the responses of two participants was used as the evaluation value. In this case, Cohen's kappa values [4] for all items in each experiment was greater than 0.2, which indicates the reliability of the results.

Table 1 shows the mean values of the reasonable rates obtained in the evaluation experiments of the alternative products. The precision for each query is one if the average of the evaluated values is more significant than three on a five-point scale. The number of found is the average number of discoveries for all queries, the number of found for the proposed method, and the unweighted method is the number of nodes used in the calculation. We found that the unweighted graphing method was rated highly for all items. The proposed method exceeded the agreement of the purpose query in all reasonable rates, except for the "direct solution" item.

Table 2 shows the average values obtained in the ranking evaluation experiments. As can be seen, the proposed method narrowing was evaluated highly in terms of "perspective," and the unweighted graph calculation with narrowing for "diversity" received a high evaluation. The unweighted graph calculation without refinement was evaluated highly for the "motivated" to buy and "overall" rating items. The proposed method exceeds the baseline results (product description's similarity) in all terms.

Table 2. Average participant rating for each method when targeted to a specific product area (1 to 4, $** : p < 0.01$, $* : p < 0.05$ to **Product Descriptions** method)

Methods	Categories	Perspective	Diversity	Motivating	Overall
Proposed method	Narrowed	**3.14**	**3.44	**3.10	**3.14
Unweighted Graph	Narrowed	**3.12	**3.50**	**3.24	**3.28
Product Descriptions	Narrowed	1.72	2.56	2.40	1.78
Proposed Method	All	**3.02	**3.30	**3.30	**3.12
Unweighted Graph	All	**3.04	**3.08	**3.60**	**3.36**
Product Descriptions	All	1.98	2.64	2.48	1.96

5 Discussion

First, an overall evaluation of the proposed method is discussed in terms of evaluating the alternative products. For all terms, we found that the proposed method exceeded the agreement on the baseline and the purpose keyword match. However, the results indicated that the unweighted graph calculation was the most effective method. Nonetheless, there were some cases where the proposed method was helpful for certain queries.

One such query was "Stretch out the wrinkles in your clothes." The higher precision can be attributed to many words related to the purpose, *e.g.*, "electric iron" and "sewing machine." For queries where the number of words is small and the types of products that can be explored are somewhat limited, words related reviews are more likely to appear.

A query for which the proposed method underperformed compared to the weighted computation was "Develop a sense of balance by training." One possible reason for the low precision for this query is that some words, *e.g.*, "house" and "home," that are unrelated to the original purpose, are assigned greater weight. Another reason may be that the weight is biased toward products with high review scores.

In addition, a query for which the results of both the proposed method and the weighted calculation were nearly the same was "Hang a picture on the wall." The reason why the precision results were similar for this query may be due to the small number of reviews for the selected products. To address this issue, the ratio of the number of reviews to the number of occurrences should be considered. The RWR algorithm enables the calculation of another indirect purpose or product that can achieve the same purpose.

The evaluation results of these experiments demonstrate that the proposed method can identify more products with higher accuracy than general search methods.

However, as a future issue, it will be necessary to reduce the number of oversight of words that express purpose. In the current search, there are cases where there is no combination of words in the query, and the search results must be examined. In addition, the extraction of purposes must be improved. In this study, the extraction process was based on language patterns; however, even if the expressions match, this does not necessarily mean that the identified product satisfies the given purpose. For example, when a product with a review that says it gets up early has a low rating, it is unlikely to solve the purpose. An effective weighting method must also be devised. The weighting technique implemented in the proposed method provides increased diversity; however, accuracy is reduced. Thus, more effective weighting methods should be investigated in future work.

6 Conclusion

In this paper, to prevent impulse buying, we have proposed a method to discover alternative products that can achieve the given purpose of using one product. For this purpose, the proposed method isolates the purpose of the product reviews according to language patterns and constructs a bipartite graph comprising products and purposes. Then, the weights of the edges are adjusted after the subgraphs are cut out for calculation. The RWR is then employed to calculate the graph, thereby making it possible to calculate another indirect purpose or product that can achieve the same purpose.

Acknowledgments

This work was supported by JSPS KAKENHI Grants Number 21H03775, 21H03774, and 22H03905. We used "Rakuten Dataset [7]" provided by Rakuten Group, Inc. via IDR Dataset Service of National Institute of Informatics [9].

References

1. Akiyama, T., Obara, K., Tanizaki, M.: Proposal and evaluation of serendipitous recommendation method using general unexpectedness. In: PRSAT@ RecSys. pp. 3–10 (2010)
2. He, R., Packer, C., McAuley, J.J.: Monomer: Non-metric mixtures-of-embeddings for learning visual compatibility across categories. arxiv **abs/1603.09473** (2016), `http://arxiv.org/abs/1603.09473`
3. Kotkov, D., Wang, S., Veijalainen, J.: A survey of serendipity in recommender systems. Knowledge-Based Systems **111**, 180–192 (2016). https://doi.org/https://doi.org/10.1016/j.knosys.2016.08.014
4. Landis, J.R., Koch, G.G.: An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics pp. 363–374 (1977)
5. Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Computing **7**(1), 76–80 (2003). https://doi.org/10.1109/MIC.2003.1167344

[7] Rakuten Dataset: `https://rit.rakuten.com/data_release/`

6. McAuley, J., Pandey, R., Leskovec, J.: Inferring networks of substitutable and complementary products. In: Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. p. 785–794 (2015). https://doi.org/10.1145/2783258.2783381
7. Onuma, K., Tong, H., Faloutsos, C.: Tangent: A novel, 'surprise me', recommendation algorithm. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. p. 657–666 (2009). https://doi.org/10.1145/1557019.1557093
8. Pothirattanachaikul, S., Yamamoto, T., Fujita, S., Tajima, A., Tanaka, K.: Mining alternative actions from community q&a corpus for task-oriented web search. In: Proceedings of the International Conference on Web Intelligence. p. 607–614 (2017). https://doi.org/10.1145/3106426.3106461
9. Rakuten Group, I.: Rakuten dataset (2020), https://doi.org/10.32130/idr.2.1, https://rit.rakuten.com/data_release/
10. Yamamoto, T., Sakai, T., Iwata, M., Yu, C., Wen, J.R., Tanaka, K.: The wisdom of advertisers: Mining subgoals via query clustering. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management. p. 505–514. CIKM '12 (2012). https://doi.org/10.1145/2396761.2396827
11. Yang, Z., Nyberg, E.: Leveraging procedural knowledge for task-oriented search. In: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. p. 513–522 (2015). https://doi.org/10.1145/2766462.2767744
12. Yong, S., Asano, Y.: Purpose-feature relationship mining from online reviews towards purpose-oriented recommendation. IEICE Transactions on Information and Systems **E101.D**(4), 1021–1029 (2018). https://doi.org/10.1587/transinf.2017DAP0013
13. Zhang, W., Leng, X., Liu, S.: Research on mobile impulse purchase intention in the perspective of system users during covid-19. Personal and Ubiquitous Computing pp. 1–9 (2020)
14. Zheng, L., Noroozi, V., Yu, P.S.: Joint deep modeling of users and items using reviews for recommendation. In: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. pp. 425–434 (2017)