

Purpose-Feature Relationship Mining from Online Reviews towards Purpose-Oriented Recommendation

Sopheaktra YONG^{†a)}, *Nonmember* and Yasuhito ASANO^{†b)}, *Member*

SUMMARY To help with decision making, online shoppers tend to go through both a list of a product's features and functionality provided by the vendor, as well as a list of reviews written by other users. Unfortunately, this process is ineffective when the buyer is confronted with large amounts of information, particularly when the buyer has limited experience with and knowledge of the product. In order to avoid this problem, we propose a framework of purpose-oriented recommendation that presents a ranked list of products suitable for a designated user purpose by identifying important product features to fulfill the purpose from online reviews. As technical foundation for realizing the framework, we propose several methods to mine relation between user purposes and product features from the consumer reviews. Using digital camera reviews on Amazon.com, the experimental results show that our proposed method is both effective and stable, with an acceptable rate of precision and recall.

key words: recommendation system, review analysis, bootstrapping

1. Introduction

Nowadays, an increasing number of consumers have gradually shifted their purchasing behavior to online shopping sites. Online retailers like Amazon have created a platform where vendors can upload their products together with a description of the functions and features of the products. Consumers often browse through other users' feedback to help with their decision-making process. However, it can be difficult to go through all of the reviews since some products might have hundreds of them.

Many sites have developed recommendation systems to help users choose suitable products for their needs. The most common technique is to display recommendations for products that are similar to a buyer's prior purchases, or products that have been purchased by consumers with similar interests or purchase history. The recommendation techniques typically rely on item-centric techniques rather than modeling the preferences of each individual user by establishing item-to-item relations and using these to suggest products that are similar to the ones viewed or purchased by the users [1]. However, this system is insufficient because it does not make recommendations based on buyers' needs, which is particularly important for products designed for a variety of functions, such as a PC, smartphone, digital camera, etc. For instance, a digital camera may serve a different

purpose for different people. For example, while some may use the camera to shoot portraits or landscapes, others may use it to take photos of wildlife, sports, car racing, etc. Generally, different product features are required for different purposes. Camera makers could not describe them all in a product description or advertisement.

Various techniques with an emphasis on product features, including methods to automatically extract product features from reviews and techniques for ranking products based on features, have been proposed in [1]–[7]. By using these, for example, each function of a digital camera such as continuous shooting speed and weight can be used as the ranking of product. If the consumers are familiar with the functions and its own intended use relationship of the product, it is perhaps possible to purchase using the ranking based on the functions as described above. However, a beginner having no knowledge of the functions required for intended use, it is not easy to choose a product that matches his/her own purpose by using information provided such a ranking or shopping sites.

Remarkably, there are commodities that are essentially designed for variety of usage purposes such as PC, smartphone, digital camera etc. For example, digital single-lens reflex of Canon at the end of year 2016, EOS 1DX Mark II and EOS 7D Mark II which are equipped with high continuous shooting speed and AF performance are excellent in freezing the motion of the high-speed birds in flight. However, these cameras could not outperform camera family EOS 5Ds in shooting the importance of landscape with high resolution. The EOS 5D camera family (5D Mark III and 5D Mark IV) has an exceptional balance of continuous shooting speed, AF, resolution and high sensitivity noise which is suitable for a wide range of fields shooting.

A Beginner, even given a ranking based on the function, is still difficult to select a suitable camera, since they do not know which camera features or functions are helpful for their specific needs. For this reason, we propose a framework of purpose-oriented recommendations that can greatly improve convenience by providing product recommendations based on a consumer's intended purpose for the product they are seeking. In this paper, we describe a method for mining hidden information from online reviews in order to extract relationships between a user's intended purpose for a product and the product's features. This process forms the technical foundation of our proposed framework.

Our Contribution. The contribution of our paper can be

Manuscript received June 28, 2017.

Manuscript revised November 6, 2017.

Manuscript publicized January 18, 2018.

[†]The authors are with Graduate School of Informatics, Kyoto University, Kyoto-shi, 606–8501 Japan.

a) E-mail: sopheaktra.yong@outlook.com

b) E-mail: asano@i.kyoto-u.ac.jp

DOI: 10.1587/transinf.2017DAP0013

summarized as follows:

- We propose a purpose-oriented recommendation framework to suggest products suitable for users, especially novice users, who do not have much knowledge about the products and a methodology that is the technical foundation for realizing the purpose-oriented recommendation framework.
- We present a bootstrapping-based methodology, employing state-of-the-art NLP techniques to extract the relationship between a user's intended purpose for a product and the product's features from online reviews.
- We introduce (1) a conditioning approach for filtering new seeds by computing the semantic similarity of words to lessen the common problem of bootstrapping - semantic drift and (2) an ontology-based technique to generate more seeds for the bootstrapping process.
- We tested and confirmed the effectiveness of the proposed methods through experiments on a very large Amazon review dataset and user evaluation.

The rest of the paper is organized as follows: our framework of purpose-oriented recommendation is illustrated in Sect. 2. Section 3 presents the methodology for extracting the relationship between product features and the user's intended purpose for the product. Section 4 contains experimental results. Brief reviews of related works conducted in this area are summarized in Sect. 5, followed by the conclusion in Sect. 6. This work builds on our preliminary work [8].

2. Purpose-Oriented Recommendation Framework

As described above, buyers tend to face difficulties when deciding which products they should purchase for their specific needs if the products have a wide range of technical features and functionalities. Figure 1 illustrates the flow of the recommendation framework we wish to build in future work.

First, the system will present a list of all possible purposes for a particular product type, e.g. digital camera. In this study, we assume that the users are aware of their purposes and asked to choose from one or more purposes in a list presented by the system. Next, the system will output the features that are most relevant and associated with the chosen purposes. During this process, the system also analyses the optimal or minimum feature requirement to achieve the corresponding purposes. For example, a burst rate of at least 6 frames per seconds is required to capture a flying bird. This type of information can not only help users pick the right camera, but also provide some education on the

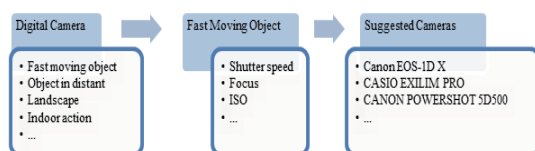


Fig. 1 Purpose-Oriented Recommendation Framework

right setting to shoot a specific target. Finally, according to the relationship between product features and purposes, the system will recommend suitable products to the users. In order to realize the framework, the most crucial part of the system is to find the association between product features and purposes expressed in the users' reviews. In the next section, we will describe the methodology in detail.

3. Methodology

The proposed methodology is composed of three main components: Component I - NLP Preprocess, Component II - Initial Seeds Generation, and Component III - Purpose-Feature Relationship Extraction Bootstrapping Model. The method extracts keywords representing purposes from the online reviews and connects them with product features. In this study, product features are manually selected as the input. They can also be extracted from several online shopping websites, such as cnet.com, dpreview.com, amazon.com, etc. There have been several studies on the automatic identification and extraction of product features, which is not the focus of this re-search. Another input component of the system is the text data of consumer reviews. In Component I, the customer reviews are preprocessed using sets of NLP tools. Next, purposes are extracted using a combined technique of Labeled LDA [9] and Word2Vec [10] by Component II. Then, starting with a few initial seeds, Component III learns patterns iteratively to extract more seeds. Figure 2 provides an overview of our method.

3.1 NLP Preprocess (Component I)

NLP is applied for linguistic preprocessing and raw text analysis such as removing stop words, sentence tokenizer, and part-of-speech tagging. The tool used in this paper is NLTK: The Natural Language Toolkit [11]. A typical review contains several sentences. Here, we employ Punkt Sentence Tokenizer for splitting reviews into sentences.

3.2 Initial Seed Generation (Component II)

Choosing seeds is arguably the most critical step in bootstrapping. However, most previous researchers either chose them manually [12], or picked the most frequently occurring words in their corpus [13] that they identified belong to the category. In this study, we propose a technique which combines topic modeling and word embedding techniques to generate initial seeds for the bootstrapping model. Our initial seed generation process is composed of three steps: candidate sentence extraction, term extraction, and term filtering.

We assume that consumers described their usage experiences or purposes inside the reviews. They often exemplified how they use the product, or identified specific functions they use to perform or achieve something. Table 1 shows some sample reviews containing intended uses or purposes of digital cameras' shutter speed functionalities.

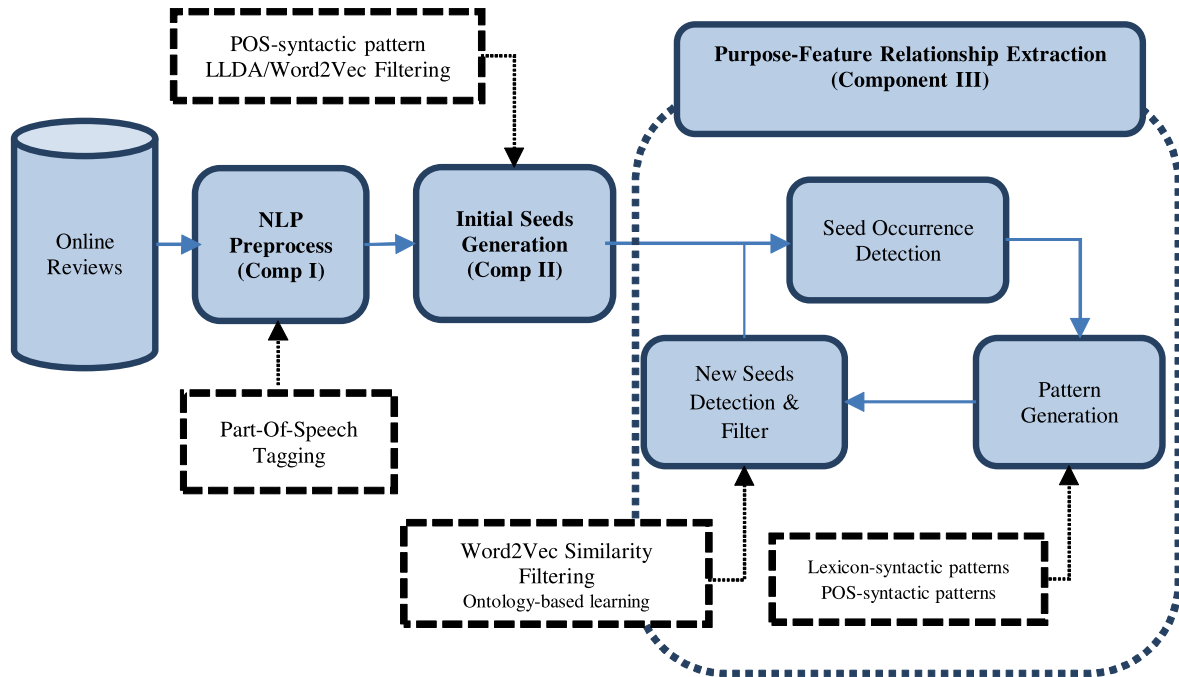


Fig. 2 Overview of our methodology.

Table 1 Customer reviews from Amazon.com. Product features and corresponding purposes are written in bold.

Review #1: The unlimited continuous shooting mode, which is perfect for catching fast-action sports shots, solves the problem of slow shutter speeds and is not found in any of the other camera's in the s3's class. Feature: shutter, Purpose: catching fast-action
Review #2: I especially liked the panoramic feature for the beautiful Alaskan landscapes and the burst feature for shooting wildlife in action. Feature: shutter, Purpose: shooting wildlife
Review #3: The shutter speed allows her to take pictures of friends in motion [dancing, skiing and other sporting events] amazing window on back for best picture view. Feature: shutter, Purpose: dancing, skiing, sporting events

Mining this type of review would enable us to detect the purposes ex-pressed by the consumers. We find “candidate sentences”, which contains product features and purpose-oriented expression, by the procedures explained below.

The process of generating initial seeds is described as follows:

1. **Candidate Sentence Extraction:** Candidate sentences are sentences containing product features as well as purpose-related expression. Sentences containing information about product features often describe whether such features are good or bad for a particular purpose. For example, consider the following statement: “The unlimited continuous shooting mode, which is perfect for catching fast-action sports shots”. Here, the phrase “Catching fast-action sports shots” is the purpose of using the “shutter” feature. This part of the review tells us that if we consider buying a camera for shooting a fast-moving object, we should choose the camera with a high “shutter” speed. To detect the

candidate sentences in the reviews, we manually define some syntactic-sentence patterns used to express purposes. An example of the pattern is <feature> + Verb + Adjective + Preposition + Noun/Noun Phrase. Next, we build the set of candidate sentences by extracting any sentences following these patterns.

2. **Term Extraction:** Once we have a list of candidate sentences, we extract the “purposes” from this set using topic modeling - Labeled LDA. Unlike the conventional LDA, Labeled LDA is a generative model which constrains the topic model to use only those topics corresponding to a document’s label set [9]. In our setting, a list of product features with their corresponding candidate sentences are prepared to train the model. The product features are trained as the topics (labels) and the candidate sentences are treated as documents. Thus, the Labeled LDA outputs the important terms per product feature. These terms are considered to be keywords or key phrases representing purposes. An example of this component output is {shutter speed: kids, wildlife, action_shots, recording, getting}, where “shutter speed” is the product feature (topic) and “kids, wildlife, action_shots, recording, getting” are the top keywords representing purposes.
3. **Term Filtering:** The list of extracted keywords may contain unrelated words since the Labeled LDA outputs the list based on word frequency alone. Therefore, we propose a filtering approach of calculating the semantic relationship of words using a recent technique, Word2Vec. It is a two-layer neural net that processes text. It can convert words into vectors by learning a corpus. Word2Vec can compute the semantic simi-

Table 2 Sample sentences and extracted patterns.

Sentences	Lexicon patterns	POS patterns
Shutter speed is good for kids.	is good for	VB JJ IN
I use burst feature for shooting wildlife.	for shooting	IN VBG

larity of words by calculating their vectors. The corpus we used to train Word2Vec is the consumer reviews. In our methodology, we utilize Word2Vec to filter out the anomaly words from the Labeled LDA's output. For example, Word2vec is able to remove the unrelated words “recording” and “getting” from the Labeled LDA output above.

3.3 Extracting Purpose-Feature Relationship (Component III)

To obtain more purpose-related sentences or purposes, we adopt a bootstrapping method as in [12] by detecting more sentence patterns for describing purposes to extract appropriate nouns and noun phrases. Because a simple bootstrapping method tends to produce noise in the results, we also proposed filtering methods to reduce this noise.

Our bootstrapping mainly iterates between the following two phases: (1) pattern generation and (2) seed extraction and filter. The algorithm begins with initial seeds and then iterates through the phases until it cannot extract any more patterns or generate new seeds with the word similarity compared to the seeds pool higher than a given threshold τ_2 .

Pattern Generation

In the pattern generation phase, we would like to detect the extraction patterns in the form of lexicon and POS patterns, which link product features and user purposes in the reviews. The process starts with a few feature-purpose seed tuples obtained from Component II. For every seed tuple $\langle x, y \rangle$, e.g. $\langle x = \text{shutter}, y = \{\text{kids}, \text{wildlife}, \dots \} \rangle$, we first retrieve all sentences containing the terms x, y . Next, all “connecting text” linking terms x and y ($\text{kids} | \text{wildlife} | \dots$) is extracted to form the candidate list. Since we also consider detecting POS syntactic patterns, the “connecting text” is parsed and tagged into the part-of-speech. As shown in the first sentence in Table 2, the phrase “is good for” (or POS-tagged “VB JJ IN”) is the pattern that connects the product feature (shutter speed) and purpose (kids). The strength of each pattern is computed by how frequently they are used to connect the seed tuples. Finally, we augment the extraction pattern if any pattern in the candidate list appears more than a specific threshold τ_1 in the reviews.

Seed Extraction and Filter

Most common bootstrapping models are pattern-based approaches in which new seeds are accepted based on the extraction patterns. This approach can lead to a well-known flaw in bootstrapping known as “semantic drift”. To avoid

Table 3 The lexicon-syntactic patterns of taxonomic relation extraction.

1. NP_0 such as $NP_1 \{, NP_2, \dots, \text{(and/or)} NP_n\}$	Hyponym(NP_i, NP_0)
2. NP_1 is a kind of NP_0	Hyponym(NP_i, NP_0)
3. $NP\{, NP\}^*\{, \}$ or other NP_0	Hyponym(NP, NP_0)
4. $NP\{, NP\}^*\{, \}$ and other NP_0	Hyponym(NP, NP_0)
5. NP_0 , including $\{NP\}^*$ or/and NP	Hyponym(NP, NP_0)

this problem, we adopt Word2Vec word similarity as a filtering method so that the current seeds do not wander away from the original semantic meaning of the initial seeds. The process is simply described as follows:

- First, our method retrieves the set of words W from the corpus that match any of the extraction patterns P . For example, in the sentence “Shutter speed is good for sport events”, “sports events” is added to W if P includes “is good for” as a pattern following the feature “shutter speed”. The words W can be expressed in multi-word terms, such as a noun or noun phrase (NP). NP is defined as $\{(\langle VBG \rangle | \langle JJ \rangle | \langle JJR \rangle | \langle JJS \rangle)^* (\langle NN \rangle | \langle NNS \rangle)^+\}$.
- The term set W is only augmented to the seed list if its similarity exceeds the predefined similarity threshold τ_2 on average.

We will introduce an ontology-based technique below, in complement to the conventional pattern-based bootstrapping approach, to generate more seeds.

Ontology-based Approach for Seed Extraction

As described in the previous section, we can only extract new seeds or purposes if the terms are expressed in the extraction patterns. Therefore, we could miss some interesting purposes because not all customers use the same sentences to express their purposes or experiences. Let us look at the 3rd review in Table 3. The terms “dancing, skiing and other sporting events” share a hyponym relation, which means that if one of them is classified as the purpose, so are the rest. For instance, the context pattern of “allows her to take pictures of friends in” is not an extraction pattern, thus the terms “dancing, skiing and other sporting events” would not have a chance to be detected. In this study, we propose an additional approach that complements pattern-based bootstrapping in detecting new seeds or purposes by mining the new purposes if they have taxonomic relations with any in the seed list.

Taxonomic relations are the most important semantic relations in a domain ontology, the extraction of which has been well studied in the field of lexicon building. We adopt the lexicon-syntactic patterns [14] for taxonomic relation extraction. A total of five lexicon-syntactic patterns are used and described in Table 3. The process of extracting taxonomic relations and generating new seeds for bootstrapping is described as follows:

- First, we (1) retrieve the set of NPs from the corpus that match any of the predefined patterns and (2) scan to determine whether at least one of the NPs is already in the seed list.

Table 4 Initial seeds used for each camera feature and outputs from the proposed model. The incorrect purposes are bold.

	Shutter	ISO	Weight	IS	Screen	Flash	Pixel	Focus
Seeds	kids sports wildlife	night light bright	vacation trip carrying	wildlife capturing	reviewing previewing low_angle	light indoor night	printing uploading	kids macros moving
Output	skiing action shots baseball sporting pets	light condition dim lit darkness dusk	bicycle_touring hiking adventures walks	kids animals fish candids sports	angle viewing inwards eye_piece composing framing	castle dim_light dark_room museums	publishing share transfer editing email	moving_targets insects portraits action shots

- Next, the rest of the NPs in the set are considered to be seed candidates and are augmented as new seeds if on average, their word semantic similarity exceeds the predefined similarity threshold τ_2 against the other terms in the seed list. Adopting this approach could help us to extract both more precise and more general purposes, which is more convenient for novice users in terms of comprehending the meaning. It could also help us to create a hierarchical classification of purposes.

4. Experimental Results

In this section, we will illustrate our experimental results and case studies. We conducted our experiments on customer reviews of digital cameras from Amazon.com. The dataset is made available via SNAP, Stanford University [15]. The dataset contains 203,773 reviews and was collected from May 1996 to July 2014.

4.1 Extracting Purpose-Feature Relationships Using a Bootstrapping Approach

In this experiment, we manually select 8 camera features and analyze the performance of each model separately by evaluating the number of purposes it could generate, precision, and (relative) recall of various models i.e. Lexicon, POS, LP, O, and LPO. The first three models are the conventional pattern-based models, whereas the LPO is the proposed model that is both pattern-based and ontology-based. The model descriptions are as follows:

- Lexicon: Generic patterns are only learnt from lexicon-syntactic patterns.
- POS: Generic patterns are only learnt from POS-syntactic patterns.
- LP: Generic patterns are learnt from new seeds generated from lexicon and POS syntactic patterns per iteration.
- O: Generic patterns are ignored in this context. New seeds are only generated from the ontology-based approach (the hyponym relation in Table 3).
- LPO: A combination of all models.

Table 4 above shows the initial seeds for each camera feature and the corresponding new seeds generated by

the proposed model - LPO. Note that some camera features' initial seeds are manually modified prior the bootstrapping process. Table 5 shows the total number of new seeds, precision, and relative recall generated by each model. LPO was able to generate the most seeds. In particular, the combination of Ontology learning and LP increased the number of seeds compared to individual models. It demonstrates that discovering new seeds via ontology relation during the bootstrapping process could increase the number of seeds accordingly.

Although it is difficult to get all the correct instances (all purposes) for each camera feature, it is possible to compare the recall of one method relative to another method. Following the work of Pantel et al. [16], the relative recall $R_{A/B}$ of method A for given method B can be calculated as:

$$R_{A/B} = \frac{R_A}{R_B} = \frac{\frac{C_A}{C}}{\frac{C_B}{C}} = \frac{C_A}{C_B} = \frac{P_A \times |A|}{P_B \times |B|}$$

where R_A (or R_B) is the recall of A (or B , respectively), C_A (or C_B) is the number of correct seeds extracted by A (or B , respectively), C is the (unknown) total number of correct seeds in the corpus, P_A (or P_B) is A 's (or B 's, respectively), precision in our experiments, and $|A|$ (or $|B|$) is the total number of seeds discovered by A (or B , respectively). We can compute the relative recall by comparing the number of seeds extracted by each model with the number of seeds retrieved by all the models together.

The total number of seeds is 272 keywords in which 199 keywords are corrected. We asked a professional photographer, with more than 10 years of experience, to annotate and evaluate the correctness of our extracted purposes. As a result, we were able to calculate the performance for all models, as shown in Table 5. The lexicon model generated a relatively small number of purposes (94 keywords) with the greatest precision, while the POS and LP models generated more purposes (235-240 keywords) with less precision. Although POS and LP generated more purposes, there is likely to be more noise in the output. However, our proposed model (LPO), which is both pattern-based and ontology-based, can not only generate more purposes, but also achieve better accuracy compared to individual components. In this experiment, the threshold τ_1 and τ_2 are set to 3 and 0.5, respectively. However, since the evaluation is performed by only one subject, the results may seem a bit biased.

Table 5 Precision, relative recall, and F-score of each method.

Methods	#New Seeds	Precision	Rel. Recall	F-score
Lexicon	94	0.84	0.40	0.54
POS	240	0.72	0.87	0.79
LP	235	0.71	0.84	0.77
O	12	0.80	0.05	0.09
LPO	263	0.75	0.99	0.85

Table 6 New purposes detected by combination of all components (LPO).

Features	Purpose keywords
Shutter	Skiing, church, sport, rides
Image Stabilizer	animals, fish, closeups, pets
Flash	castles, cathedrals, dancers, nightclubs
Focus	moving targets, insects, food, cat

Table 7 Purpose categories and individual keywords.

Purposes	Keywords
Macros	Flower, insects, foods, ...
Moving targets	Animals, kids, freezing, motion, action shots, sport events, soccer, ...
Landscape	building, nature scenes, mountains, gardens, ...
Portraits	candid, distant objects ...
Viewing and framing	previewing, histogram viewing, angle shooting, ...
Uploading	share, computer, transfer, publishing ...
Low light condition	dimmer environments, backlit daylight, dark room, ...
Bright condition	sunlight, daylight, outdoor, portraits at outdoor, ...
Traveling	trips, touring, backpack, adventures, outings, ...

4.2 Subjective Evaluation

A subjective evaluation is conducted to examine and emphasize the accuracy/effectiveness of the proposed method, LPO, with a variety of users. As noted previously, there were 263 keywords generated using the LPO model, which is inconvenient and time-consuming for users having to evaluate the results. Therefore, we categorize purposes and ask users to list important features for each category. We manually defined nine categories (shown in Table 7) because automatic categorization is out of our scope; however, the automatic categorization could be done in future by mining both taxonomic and non-taxonomic relations, as was done in [17].

Questionnaire Setup

We created 10 questions in total [18]. Question #1 asks about the respondent's photography experience and skills. Questions #2 to #10 survey the respondent's knowledge or experiences of which relevant camera features they are using for each purpose. Respondent can select answers (camera features) more than one. For example, in question #2 (as shown in Fig. 3), the respondent can select any camera features in which are useful to capture macro subjects. Thus,

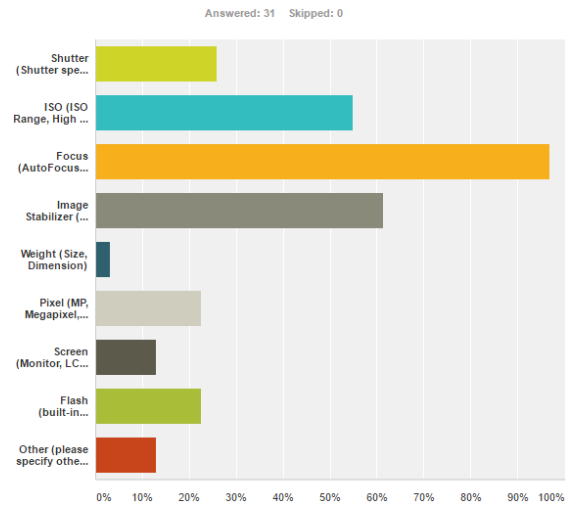
Which camera feature(s) is/are used for the following purposes: macros/closeups subjects flowers (stationary) insects foods

Fig. 3 Sample Responses - Camera features for "Macros" purposes

Table 8 Category-based precision and recall of proposed model, LPO.

Purpose Category	Ground Truth	Proposed Model	P	R
Macros/Closeups	Focus, IS, ISO	Focus, IS	1	0.67
Landscape/Scenery	ISO, Focus, IS	Focus, Shutter, IS	0.67	0.67
Moving targets	Shutter, ISO, Focus, IS	Focus, Shutter, IS	1	0.75
Portraits	Focus, IS, ISO	Focus, IS	1	0.67
Viewing and framing	Screen	Screen	1	1
Uploading, editing, and printing	N/A	Pixel	0	N/A
Low light condition	ISO, Flash	ISO, Flash	1	1
Bright condition	ISO, Shutter	ISO, Flash	0.5	0.5
Traveling	Shutter, Focus, IS	Weight	0	0
Average			0.69	0.66

we may have multiple selected camera features for a particular purpose. We consider the most relevant camera features as the ground truth to evaluate our result by selecting any answers receiving at least 50% votes from all the respondents. For instance, in Fig. 3, 3 camera features - focus, image stabilizer, and ISO receives more than 50% votes compared to other features. Thus, these features are considered as the ground truth for evaluating the experiment results.

Results

We received 31 respondents, primarily consisting of beginner users (less than 1 year of experience) and professional users (more than 5 years of experience). Their responses can be viewed as subjective evaluations from users with different opinions and levels of experience. Since we asked people to rank the camera features for different purposes, we could obtain the ground truth of relationship between user purposes and features. We were able to calculate the accu-

Table 9 Precision, recall, and F-score of each method (Category-based).

Methods	P	R	F-score
Lexicon	0.69	0.54	0.61
POS	0.69	0.61	0.65
LP	0.69	0.61	0.65
O	0.11	0.06	0.08
LPO (Proposed Model)	0.69	0.66	0.67

racy of each method, as shown in Tables 8 and 9. Though the precision of pattern-based models and LPO is the same, the number of individual keywords extracted by the LPO is higher, based on the evaluation in Sect. 4.1. In comparison to other methods, we highlight the number of new purpose keywords detected by LPO method (Table 6). Therefore, we can conclude that our combined model could extract better and precise purposes (better recall) than the pattern-based or ontology-based models with better accuracy (F-score 0.67).

5. Related Work

A number of recommendation techniques have been developed as mentioned in the introduction part. Jannach et al. [1] proposed a method to compare and quantify the combined effects of short-term and long-term profile using navigation log data. The navigation logs (such as Amazon “Customer Who Viewed This Item Also Viewed”, “Recently viewed items”) is used to estimate the user’s shopping goals as the recommendation context.

Chen et al. [19] studied the importance of “price” factors for recommendation system in unexplored categories when user has no purchase history. As a result, the authors conclude that the price factor can be a helpful factor in the unexplored categories. Price can be indicated into two perspectives either as a quantity for a product or as a distribution for user.

Peska et al. [20] tried to solve the problem of cold start and new products in recommendation system by using the Linked Open Data to enrich the content information.

Cutolo [21] used Vector Space Model technique to model a user and structural characteristic of product/services for ranking mechanism. For example, by knowing the user preference’s payment method, the products with the same payment methods will be ranked first.

Musat et al. [2] proposed a methodology to create a hybrid recommendation system where users are classified based on their opinions. The users are categorized into opinionated and non-opinionated. The system suggests products based on his interest profile for opinionated users. Otherwise, the simple average star rating can be used to recommend products for non-opinionated ones.

However, no purpose-oriented recommendation techniques have been proposed, to the best of our knowledge. Again, our goal is to detect and discover the association between user purposes and product features hidden in the user reviews. The findings from our work are to help users identifying the most important product features for purposes they want to obtain. In this section, we will take a closer

look at some of the previous researches related to the ideas and techniques used in our method, including feature-based ranking, mining semantic relations using bootstrapping, and ontology learning.

Feature-based ranking system

Mining of opinions from reviews has become a popular area of research. Users often express their usage experiences in reviews and sometimes praise some features of the product using for a particular function or purpose. Regarding the feature extraction techniques, it could be classified into supervised and unsupervised approach.

Supervised techniques require a set of pre-annotated review sentences as training sets. An example of supervised techniques, Wong and Lam [22] employ the hidden Markov model and conditional random fields as the underlying learning method for extracting product features from auction websites.

In contrast, product features can be extracted automatically using unsupervised technique presented by Hu and Liu [23]. They made assumptions that product features must be nouns or noun phrases. This utilizes the association rule mining algorithm to discover all frequent item sets within a target set of reviews. The opinion word refers to the adjectives that are adjacent to the frequent feature. Correspondingly these set of opinion words are treated as the seeds in searching for infrequent features.

Regarding the problem of feature-based summarization, Kangale et al. [5] provided a solution that could extract user opinions from reviews, and generate a rating as well as review summary of each product feature. Similarly, Kamal [6] has proposed a visualization system called OSVS - Opinion Summarization and Visualization. The system is able to extract reviews and translate into a graphical form (bar chart or pie chart) easier for users to have a quick view of product features and other users’ sentiments expressed over them. Nevertheless, beginner users who do not have much knowledge about the products, they still could decide which products are suitable for them even with the provided ranking because they do not know which product features should be taken into account.

Dong et al. [24] proposed a method to extract product features (i.e., aspect-related terms) and the accompanying opinions to build product profiles. These profiles are used to prioritize retrieved products that are similar to a user’s query product and have also been positively reviewed by users. Sinha and Roy [25] investigate the potential of preference mapping for identifying the most relevant product attributes for a marketing campaign. They test their approach on user reviews mined from the Web and provide an effective visualization of their results.

Again, our goal is to detect and discover the association between user purposes and product features hidden in the user reviews. The findings from our work are to help users identifying the most important product features for purposes they want to obtain.

Mining semantic relations with bootstrapping

Most common approach of bootstrapping to extract the semantic relations is pattern-based approach. Agichtein and Gravano [12] proposed “Snowball” algorithm to identify and extract structured relations between named entities from unstructured text. A relation is defined as a tuple that maps one entity onto another. The system is given with some initial seeds then searches the training corpus looking for documents where terms in a seed tuple occur closely to each other. It then analyzes the connecting text and surrounding context and generates an extraction pattern. Pantel et al. [16] proposed an approach based on an edit-distance technique to learn lexicon-POS patterns and could obtain both good performance and efficiency. Later Espresso [26] has been proposed by the same author to infer patterns for harvesting binary semantic relations (is-a, part-of, succession, reaction, and production).

It also describes refining techniques to deal with wide variety of relations by measuring the strength of association between patterns and seeds using pointwise mutual information. In our work, we propose an additional filtering approach for accepting new seeds by measuring the word semantic similarity to avoid semantic drift.

Ontology learning

Because we have introduced an ontology learning approach to generate more seed instances in conventional bootstrapping, we will take a look at some studies related to ontology. Hearst [14] pioneered using patterns to extract hyponym (is-a) relations. Manually building a few lexicon-syntactic patterns, Hearst sketched a bootstrapping algorithm to learn more patterns from instances, which has served as the model for most subsequent pattern-based algorithms. We have adopted this approach in our study by extracting hyponym relation of generated seeds using some predefined lexicon-syntactic patterns listed in Table 3 of Sect. 3.3.

Chen et al. [17] proposed an ontology learning framework to extract customer needs of digital cameras. To detect the ontology relation, the authors detect taxonomic and non-taxonomic relations from the customer reviews. Taxonomic relations are extracted using several methods including string matching, lexicon-syntactic patterns as in [14], and WordNet[†] taxonomic relations. Non-taxonomic relations are detected using rule-based method in the form of “Noun1, Verb, Noun2” lexical relation. In addition, the author proposed a word property-based method for extracting more non-taxonomic relations by measuring the support and confident of the co-occurrence noun or noun phrases from the corpus.

6. Conclusion and Future Work

In this paper, we proposed a method for extracting semantic relations between product features and purposes. This work is important in order to realize the proposed framework of purpose-oriented recommendation. The output of

the framework could help educate novice users to understand a product’s features, which should be taken into account for their specific needs or purposes. Our results indicate that the combination of pattern-based and ontology-based approaches could extract a large number of purposes, with good precision and recall.

For future research, we are considering quantifying the optimal setting or requirements of product features for achieving a specific purpose. For instance, if the user is interested in buying a digital camera for shooting their busy kids, then it would be helpful to not only rank the relevant features, but also the optimal setting configurations (shutter speed of 1/500 sec, 5fps burst rate, ISO 1600, continuous AFC etc. is the minimum requirement to capture a sharp image). Such quantification not only tells users which camera they should purchase, considering the important features, but also provides education on the settings required to achieve a specific goal when using the product.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 15K00423 and the Kayamori Foundation of Informational Science Advancement.

References

- [1] D. Jannach, L. Leriche, and M. Jugovac, “Adaptation and evaluation of recommendations for short-term shopping goals,” *Proceedings of the 9th ACM Conference on Recommender Systems*, pp.211–218, ACM, 2015.
- [2] C.C. Musat and B. Faltings, “Personalizing product rankings using collaborative filtering on opinion-derived topic profiles,” *IJCAI*, pp.830–836, 2015.
- [3] K. Zhang, R. Narayanan, and A. Choudhary, “Voice of the customers: Mining online customer reviews for product feature-based ranking,” *Proceedings of the 3rd Workshop on Online Social Networks, WOSN*, pp.1–9, 2010.
- [4] M. Hu and B. Liu, “Mining opinion features in customer reviews,” *Proceedings of the Nineteenth National Conference on Artificial Intelligence, Sixteenth Conference on Innovative Applications of Artificial Intelligence*, pp.755–760, 2004.
- [5] A. Kangale, S.K. Kumar, M.A. Naeem, M. Williams, and M.K. Tiwari, “Mining consumer reviews to generate ratings of different product attributes while producing feature-based review-summary,” *Int. J. Syst. Sci.*, vol.47, no.13, pp.3272–3286, 2016.
- [6] A. Kamal, “Review mining for feature based opinion summarization and visualization,” *CoRR*, vol.abs/1504.03068, 2015.
- [7] S. Uchida, T. Yamamoto, M.P. Kato, H. Ohshima, and K. Tanaka, “Object search by experience attributes,” *DBSJ Japanese Journal*, vol.14, no.7, pp.1–7, 2016.
- [8] S. Yong and Y. Asano, “Mining relationship between user purposes and product features towards purpose-oriented recommendation,” *Proceedings of the Second International Conference on Data Mining and Big Data*, vol.10387, pp.9–21, 2017.
- [9] D. Ramage, D.L.W. Hall, R. Nallapati, and C.D. Manning, “Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora,” *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, EMNLP 2009*, pp.248–256, 2009.
- [10] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” *arXiv:1301.3781*, 2013.

[†]<http://wordnet.princeton.edu/>

- [11] S. Bird, "NLTK: the natural language toolkit," Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics, pp.69–72, 2006.
- [12] E. Agichtein and L. Gravano, "Snowball: extracting relations from large plain-text collections," Proceedings of the fifth ACM conference on Digital libraries, (ACM DL2000), pp.85–94, 2000.
- [13] M. Thelen and E. Riloff, "A bootstrapping method for learning semantic lexicons using extraction pattern contexts," Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing, EMNLP 2002, pp.214–221, 2002.
- [14] M.A. Hearst, "Automatic acquisition of hyponyms from large text corpora," 14th International Conference on Computational Linguistics, COLING 1992, pp.539–545, 1992.
- [15] J.J. McAuley, R. Pandey, and J. Leskovec, "Inferring networks of substitutable and complementary products," Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.785–794, 2015.
- [16] P. Pantel, D. Ravichandran, and E.H. Hovy, "Towards terascale semantic acquisition," Proceedings of the 20th International Conference on Computational Linguistics, COLING 2004, pp.771–777, 2004.
- [17] X. Chen, C.-H. Chen, K.F. Leong, and X. Jiang, "An ontology learning system for customer needs representation in product development," The International Journal of Advanced Manufacturing Technology, vol.67, no.1-4, pp.441–453, 2013.
- [18] S. Yong and Y. Asano, "Online Survey about Digital Cameras." <https://www.surveymonkey.com/r/RMW9C7T>.
- [19] J. Chen, Q. Jin, S. Zhao, S. Bao, L. Zhang, Z. Su, and Y. Yu, "Does product recommendation meet its waterloo in unexplored categories?: No, price comes to help," Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, pp.667–676, ACM, 2014.
- [20] L. Peska and P. Vojtas, "Using linked open data to improve recommending on e-commerce," 2nd International Workshop on Semantic Technologies meet Recommender Systems & Big Data (SeRSy 2013). CEUR-WS, 2013.
- [21] A. Cutolo, G. D'Aniello, F. Orciuoli, F. Pettinati, G. Sansonetti, and C. Vitagliano, "An ontology-based recommender system in e-commerce," 2nd International Workshop on Semantic Technologies meet Recommender Systems & Big Data (SeRSy 2013). CEUR-WS, 2013.
- [22] T.-L. Wong and W. Lam, "Learning to extract and summarize hot item features from multiple auction web sites," Knowledge and Information Systems, vol.14, no.2, pp.143–160, 2008.
- [23] M. Hu and B. Liu, "Mining and summarizing customer reviews," Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.168–177, 2004.
- [24] R. Dong, M.P. O'Mahony, M. Schaal, K. McCarthy, and B. Smyth, "Sentimental product recommendation," Proceedings of the 7th ACM conference on Recommender systems, pp.411–414, ACM, 2013.
- [25] M. Sinha and R.S. Roy, "Preference mapping for automated recommendation of product attributes for designing marketing content.," CBRecSys@ RecSys, pp.65–68, 2014.
- [26] P. Pantel and M. Pennacchiotti, "Espresso: Leveraging generic patterns for automatically harvesting semantic relations," Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics, pp.113–120, Association for Computational Linguistics, 2006.



Sopheaktra Yong received the Bachelor of Science in Computer Science and Engineering and from Royal University of Phnom Penh in 2010 and Master degree in Social Informatics from Kyoto University in 2017. In 2010-2014, he had been working as a system developer in Cambodia prior to his study in Japan. He is now with Asial Corporation.



Yasuhito Asano received the BS, MS, and DS degrees in information science from the University of Tokyo in 1998, 2000, and 2003 respectively. In 2003-2005, he was a research associate in the Graduate School of Information Sciences, Tohoku University. In 2006-2007, he was an assistant professor in the Department of Information Sciences, Tokyo Denki University. He joined Kyoto University in 2008. He currently serves as an associate professor in the Graduate School of Informatics. His research interests include web mining and network algorithms. He is a member of the IEICE, IPSJ, DBSJ and OR Soc. Japan.