

A Semi-automatic Approach for Requirement Discovery in the E-commerce Industry

Mate Kovacs and Victor V. Kryssanov

¹ **Abstract**—Analyzing and understanding customer requirements is strongly related to the long-term growth of businesses. The lack of automatic methods to elicit customer requirements makes it a time- and money-consuming task for companies to obtain useful information for product design and marketing specialists. In the case of the e-commerce industry, the vast amount of customer feedbacks makes it unfeasible to assess the needs of customers manually. In the presented study, Latent Dirichlet Allocation is applied in combination with Part of Speech selection to extract product and service requirements from a large collection of Japanese reviews. Results obtained in experiments suggest that the proposed approach can be used in practice to efficiently and effectively manage customer requirements in the e-commerce industry.

Index Terms—Requirement elicitation, customer need assessment, topic modeling, Japanese language.

I. INTRODUCTION

Companies collect information from customers to improve quality of their products and services, reduce the uncertainty associated with new product development, and learn relevant customer behavior patterns for marketing purposes. Poor understanding of customer requirements may negatively affect the whole manufacturing process, and have serious consequences for development time and cost [1]. Companies should, therefore, pay attention to the needs of the customers to reduce the market uncertainties, especially in the early stages of new product development [2].

Ever since e-commerce was born, online buying and selling services have been growing rapidly. Within this business environment, companies have several means to get customer feedback via their websites. Customers often rely on other customers' reviews, as such reviews are generally seen as relatively unbiased yet broad descriptions of the products. The reviews are also of an utmost importance for the companies, and can be a reliable source of information for product and service requirements. The immense amount of feedback from customers of e-commerce sites in these days makes it, however, hardly feasible to process and analyze the reviews manually, and to translate them into practically useful requirements. Furthermore, as customer needs and requirements usually come from a natural language source, obtaining specific product feature specifications and requirements is a time- and money-consuming effort [2].

In this paper, a semi-automatic approach is described that

allows one to discover customer requirements in the e-commerce industry. The goal of the presented study is to develop a system to help companies elicit customer requirements with minimal human intervention, and to provide useful information for domain experts in marketing and product design research.

Topic modeling was performed with a large collection of Japanese reviews, incorporating language-dependent linguistic cues to discover product and service requirements in an unsupervised way. The obtained topic definitions were then manually examined and adjusted to further improve the quality of the requirements extracted.

The rest of this paper is organized as follows. Section II provides a brief review of the related literature. Section III explains the proposed approach. Section IV presents the dataset used in experiments, the experimental procedure, and the results obtained. Finally, conclusions are formulated, and future work is outlined in Section V.

II. RELATED WORK

Assessing the needs of customers is usually done by conducting surveys and/or carefully planned interviews with the customers, often by an outside company specialized in market research. Godes *et al.* manually evaluated and analyzed customer feedback through customer forums, but the authors' approach proved to be methodologically inefficient in a long run [3].

As several studies emphasized, data scientists and marketing experts have to work together for productive requirement engineering to efficiently analyze customer feedbacks [4]. Aguwa *et al.* proposed an original approach to interpret and analyze the voice of a customer, and to convert the natural language input to a quantitative format to estimate the customer satisfaction ratio [5]. Zhang *et al.* introduced a novel methodology for ranking products based on their features, using customer reviews [6]. The authors manually defined a set of features relevant to the domain, then created a feature-specific product graph to characterize the overall quality of a product. Mining customer reviews for opinion summarization with sentiment polarity determination is a well-researched field with a number of studies conducted on Amazon customer reviews [7]. Nevertheless, using a huge amount of customer reviews collected in the e-commerce industry to automatically extract fundamental product and service requirements remains a challenge.

III. PROPOSED APPROACH

Fig. 1 depicts the processing flow of the approach

Manuscript received February 12, 2018; revised May 10, 2018.
The authors are with Graduate School of Information Science and Engineering, Ritsumeikan University, Japan (e-mail: gr0278ir@ed.ritsumei.ac.jp).

proposed in this study for requirement discovery from product reviews. After collecting customer reviews from the retailer's website, product and service aspects are extracted.

Unnecessary and unrelated requirements and requirement groups are eliminated to achieve more accurate results.

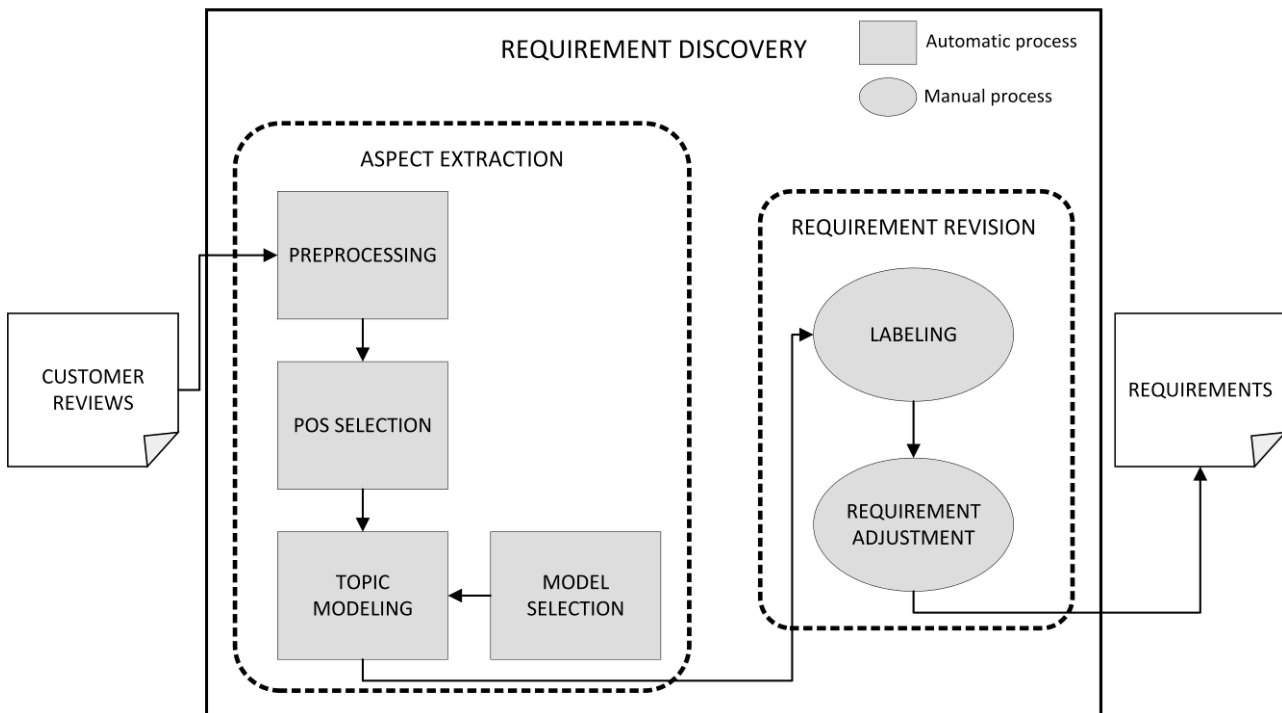


Fig. 1. The processing flow of the approach proposed for requirement discovery.

A. Aspect Extraction

Topic modeling is used to extract product and service aspects, and to formulate requirements from customer reviews. An aspect is defined as a product attribute discussed by the customers. Statistical methods of topic modeling are often used in Natural Language Processing to explore abstract semantic structures over corpora. In this interpretation of a topic model, topic words represent aspects of products and services. Latent Dirichlet Allocation (LDA) [8] is applied to build a topic model over the reviews. LDA is an unsupervised generative hierarchical clustering method that represents documents as word random mixtures over hidden topics, and it allows one to uncover these topics over a corpus. Each topic is defined by a word distribution. The iterative algorithm learns a distribution, where documents are in a latent topic space, which describes the whole corpus. In the context of customer need assessment, the algorithm discovers meaningful requirement groups over the reviews, and assigns to them product aspects with the corresponding probabilities.

To obtain more accurate results, only a fraction of words is considered. After tokenizing and stemming the reviews, a general Japanese stopword list is applied to the corpus. In preliminary experiments, the authors analyzed what Part of Speech (POS) types are important when it comes to product aspect definition. Naturally, nouns have to be considered, but not all types of nouns are relevant to product and service requirements. Important findings of the preliminary experiments are summarized as follows:

- *Common nouns* (一般名詞) are the Part of Speech most contributing when formulating product and service aspects, regardless the language.

- *Proper nouns* (固有名詞) are usually considered as brand names or names of places, and can help improve the topic interpretability.
- *Nouns that can become verbs* (サ変接続名詞) are a special category in the Japanese Part of Speech taxonomy but are just as important as common nouns.
- *Independent verbs* (自立動詞) do not necessarily represent product aspects but can help one to arrive at a better understanding of the topics mined with the LDA algorithm.

As the input for LDA is a Bag of Words (BOW) model of the documents, a BOW model is built over the corpus after preprocessing and POS selection.

B. Model Selection

Determining the correct number of clusters is an important task of unsupervised clustering. One of the most common ways to evaluate a probabilistic model is to measure the log-likelihood of the model's fit on a test set. In the case of topic models, this would mean to calculate the per-word perplexity on unseen documents. However, as Chang *et al.* [9] demonstrated that perplexity and manual evaluation of topics frequently do not correlate, a better evaluation method needs to be found for model selection.

In this study, the C_v coherence measure is used to evaluate the model quality, conditional on the number of topics. Introduced in [10], for certain probable words in a given topic, C_v is computed as the cosine similarity between every word vector pair in respect to their Normalized Pointwise Mutual Information score within a Boolean sliding window. The topic scores are averaged over the observed topics to

obtain an aggregated value. Higher coherence scores correspond to better, more consistent models.

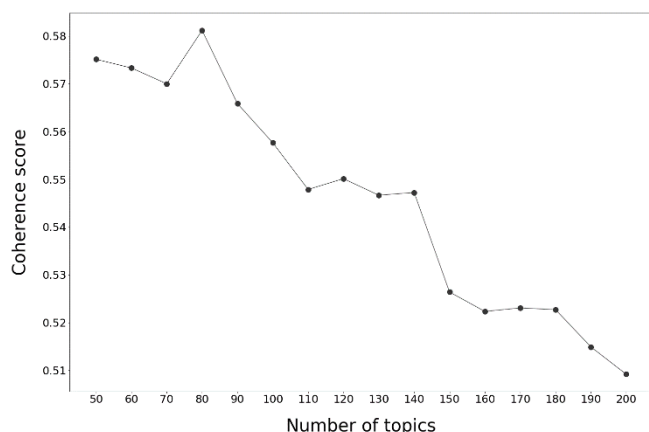


Fig. 2. Change in the coherence score with the number of topics.

C. Requirement Revision

After a topic model is built over the product reviews, the topics have to be manually labeled in accordance with their aspect words. As LDA is an unsupervised clustering algorithm, the topics obtained may include unrelated aspects, while some of them may even be meaningless. Adjustments of the requirements by eliminating unnecessary and apparently meaningless aspects, and merging similar topics need, therefore, to be done manually (preferably by domain experts) to obtain as expressive and useful requirements as possible.

IV. EXPERIMENT

A. Data

The dataset used in the study is the Rakuten Ichiba data provided by the Rakuten, Inc. (the dataset is distributed by the National Institute of Informatics). The data contains approximately 64 million product reviews from the Japanese electronic commerce site www.rakuten.co.jp. There are seven main product categories on the website: Fashion, Electronics, Foods and Drinks, Home and Garden, Beauty and Health, Cars and Sports, and Other. Inside these, there are 41 subcategories in total, with several hundred types of products.

B. Results and Discussion

LDA hyperparameters for the priors controlling the sparsity of the document-topic and topic-word distributions, the so-called α and β [8] were set to be in an inverse proportion to the topics. Passing through the corpus 2 times by chunks of 10000 reviews, the training was completed in over 12000 updates. Taking into account the number of product categories, the LDA algorithm was run for the number of topics set initially to 50, and incrementally raised to 200 with the step size of 10. Fig. 2 shows how C_v changes with the number of topics set for the given data. As the number of topics grows, the model quality decreases, and the large number of topics does not adequately represent the aspect space of the products reviewed. The coherence score peaks when the number of topics is set to 80 ($C_v=0.581$), so the model with 80 topics was selected for further analysis.

Three of the extracted topics are listed below.

- **Skincare:** 肌 (skin), 乾燥 (dryness), 化粧品 (makeup), 顔 (face), クリーム (cream), 塗る (to spread), サンプル (sample), 湿る (to be moist), 刺激 (irritation), 洗顔 (face-washing), 伸び (spreading), なじむ (get accustomed to), ツル (smooth), シミ (skin spot), 成分 (ingredient), パウダー (powder), 毛穴 (skin pores), 試す (try out), トラブル (trouble), 美容 (cosmetics), 荒れる (become rough), 潤い (moisture), コットン (cotton), アトピー (atopic dermatitis), 浸透 (penetration).
- **Customer service:** 対応 (support), 発送 (shipping), メール (e-mail), 梱包 (packaging), 配送 (delivery), 助かる (help), 連絡 (contact), 確認 (confirmation), 感謝 (gratitude), 安心 (relief), 指定 (specification), 電話 (phone call), 到着 (arrival), 希望 (expectations), 変更 (change), 問い合わせ (inquiry), 間に合う (to be in time), 遅れる (to be delayed), かかる (time it takes), 予定 (schedule), 発注 (placing an order), ミス (mistake), おまけ (extra), 丁寧 (polite), 宅配 (home delivery), 返事 (reply), キャンセル (cancellation).
- **Ochuugen and obon:** 子供 (child), タオル (towel), 大人 (adults), 喜ぶ (to be pleased), 実家 (one's parents' home), 帰る (to go home), 親 (parents), 一緒 (together), 吸水 (water absorption), 先方 (companion), フェイスタオル (face towel), パパ (father), 幼稚園 (kindergarten), 中元 (chuugen), お盆 (obon), お花 (flower), 高齢 (old age), 仏壇 (Buddhist household altar), 田舎 (countryside), 自宅 (one's home), ハンドタオル (hand towel), 祭り (festival), 贈答 (exchange of presents), 法事 (Buddhist memorial service), 差し上げる (to give, offer).

In the experiment, the requirement revision step was conducted by the authors on the top 50 most probable words (this number can be adjusted, depending on the topic contents and the domain scope). The order of aspect words corresponds to how probable is the word for that topic. The three topics above give an illustrative example of requirement groups discovered with the proposed approach. The first topic is category-specific (other similar topics discovered include *cosmetics*, *swimsuits*, *kitchen knives*, *humidifiers*, *interior design* etc.) The second is more basic, and is not restricted to product categories (similar topics include *packaging and shipping*, *quality and price*, *colors*, etc.) Although culture specific properties are observed in almost every topic, there are some requirement groups which are especially attributed to the Japanese culture. For instance, *ochuugen* midyear presents and *obon* customs (another example is *animes*, *mangas*, and *Japanese dramas*). It should be noted that some of the topics obtained consist of peculiarly abstract aspects, such as, for instance, *configuration and*

operation, beginner or novice in something, recommendation from another person, etc.

Through requirement revision, the authors found four inconsistent topics that were then eliminated, and were not included in the final requirements. Furthermore, some other topics appeared to be quite similar (e.g. those about beauty products), and could be merged but were left unchanged in this experiment. The topic labels after the requirement revision step are listed below.

Specific colors I, customer service, colors general, skirts and one pieces, clothes sizes for kids, footwear I, women underwear and shorts, bags, quality and price, out of stock, recommendation, choosing Rakuten instead of other opportunities, product similarity with illustrations, drinks, seals, clothes general, giving presents, deodorants, interior design, foods I, specific colors II, special offers, footwear II, packaging I, aromas and smells, bed and bed linen, shipping, artwork (manga, anime, drama, etc.), wallets, ochuugen and obon, cooling devices, birth stones, skincare, washing clothes, everyday clothes, furniture, chargers, hair products, foods II, beds, accessories, beauty products I, watches and accessories, sending gifts, women underwear, toys, quality and price, operation and configuration, printers, plants and seeds, shirts, humidifiers, music, kitchen knives, bathing goods, brands and origin of products, foods III, women swimsuits, bras, clothes for pregnant women, rice, recipes, cleaning, beginner or novice in something, cooking, tables and chairs, packaging II, kimono, beauty products II, curtains, computer mice, diapers, quit smoking, telescopes, defective products, Valentine's day.

The acquired requirements would be used by the related businesses selling goods on the retailer's website to improve their products and services that is often the ultimate goal of customer need elicitation and assessment. The requirements (especially, the more fundamental and abstract ones) would also be used in the retailer's website design to adjust the user interface and interaction scenarios, based on the customer priorities. Overall, the approach would help improve both business-to-customer and business-to-business relations.

V. CONCLUSIONS

In the presented study, topic modeling in combination with Part of Speech selection was applied to a large collection of Japanese customer reviews to extract customer product and service requirements. A topic coherence measure was used to assess the model quality. The results obtained demonstrated that the proposed approach can help e-commerce companies to manage product and service requirements cost -and time-efficiently.

In future work, an interactive tool will be developed for product design and marketing specialists to assist extracted requirement evaluation in a user-friendly manner. Since the described experiment was not meant to validate the proposed approach in industrial settings, future work will also include a case study, where domain experts would evaluate the usability of the proposed approach.

ACKNOWLEDGMENT

The authors acknowledge the use of the Rakuten Data provided for research purposes by the National Institute of Informatics, Japan, in cooperation with Rakuten, Inc.

REFERENCES

- [1] P. Engelbrektsson and M. Söderman, "The use and perception of methods and product representations in product development: A survey of Swedish industry," *Journal of Engineering Design*, vol. 15, pp. 141-154, 2010.
- [2] R. J. Jiao and C. H. Chen, "Customer Requirement management in product development: A review of research issues," *Concurrent Engineering*, vol. 14, pp. 173-185, 2006.
- [3] D. Godes and D. Mayzlin, "Using online conversations to study word-of-mouth communication," *Marketing Science*, vol. 23, pp. 545-560, 2004.
- [4] E. T. Bradlow, "User-generated content: The 'voice of the customer' in the 21st century," in *Marketing Intelligent Systems Using Soft Computing. Studies in Fuzziness and Soft Computing*, vol. 258, J. Casillas, F. J. Martínez-López Eds. Berlin, Heidelberg: Springer, 2010, pp. 27-29.
- [5] C. C. Aguwa, L. Monplaisir, and O. Turgut, "Voice of the customer: Customer satisfaction ratio based analysis," *Expert Systems with Applications*, vol. 39, pp. 10112-10119, Sep. 2012.
- [6] K. Zhang, R. Narayanan, and A. Choudhary, "Voice of the customers: Mining online customer reviews for product feature-based ranking," in *Proc. the 3rd Conference on Online social networks*, pp. 11-19, 2010.
- [7] A. Bhatt, A. Patel, H. Chhedha, and K. Gawande, "Amazon review classification and sentiment analysis," *International Journal of Computer Science and Information Technologies*, vol. 6, pp. 5107-5110, 2015.
- [8] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993-1022, 2003.
- [9] J. Chang, J. B. Graber, S. Gerrish, C. Wang, and D. M. Blei, "Reading tea leaves: How humans interpret topic models," in *Proc. the 22nd International Conference on Neural Information Processing Systems*, pp. 288-296, 2009.
- [10] M. Röder, A. Both, and A. Hinneburg, "Exploring the space of topic coherence measures," in *Proc. the Eighth ACM International Conference on Web Search and Data Mining*, pp. 399-408, 2015.



Mate Kovacs received his BA in Japanese Studies from Karoli Gaspar University, Hungary. Since 2016, he is at the Graduate School of Information Science and Engineering, Ritsumeikan University, Japan, pursuing a master's degree. His research interests include statistical modeling of natural languages and natural language processing for business purposes.



Victor V. Kryssanov received his Ph.D from the Russian Academy of Sciences in 1994. He currently serves as a professor at the College of Information Science and Engineering, Ritsumeikan University, Japan. Before joining Ritsumeikan University in 2004, he was a JSTA researcher at Kyoto University, a JSPS research associate at Kobe University, and a NEDO guest researcher at the JSPMI Technical Research Institute, Tokyo. His research interests include social network analysis, communication theories, and statistical modeling of complex systems.