Analysis on the Usage of Topic Model with Background Knowledge inside Discussion Activity in Industrial Engineering Context

Muhammad Luthfi*, Satoshi Goto† and Osamu Yoshie‡
Graduate School of Information, Production, and Systems
Waseda University, Kitakyushu, Japan 808-0135
**muhammad.luthfi@akane.waseda.jp, †satoshi.goto@fuji.waseda.jp, ‡yoshie@waseda.jp*

***Abstract*—Analyzing discussion activities to find any latent opinions and hidden patterns is an important problem to improve consensus building process. A lot of approaches has been proposed in forms as set of instructions and frameworks such as causal model of Consensus Building Theory (CBT) and short-term intensive workshop in strategy planning phase of Product Lifecycla Management (PLM) process. This paper will analyse a new approach to improve consensus building process by summarizing discussion activiy. The novel method is done by performing data augmentation and topic modeling with the help of background knowledge on discussion activity held within industrial engineering context. Our method produce a complete summarization of discussion activity that consists of topic distribution and the degree of similarity between topics. We also found that the usage of data augmentation will improve topic quality. We validate our findings to a professional consultant and conclude that our approach gives an adequate contribution towards summarizing discussion activity that might improve consensus building process.**

***Keywords*-topic model; background knowledge; consensus building; product lifecycle management; data augmentation**

I. Introduction

A conventional discussion activity happened when a group of people let out their own opinion with appropriate feedbacks from the other. In industries, discussions are being held in various departments to solve specific problems. We can characterize such discussions as a group of people who shares a same interest aimed to build one single consensus. Furthermore, consensus building is important because it can resolves dispute more effectively by involving people from various levels and departments in an organization [1]. Nowadays, most companies are using consensus building approach on the requirement decision part of their products, hence making such activities as a specific-themed discussion activity. The practice of consensus building often times still have frequent problems. During discussion activities, various stakeholders with different personalities and backgrounds are present might influence final conclusion [2] which will affect tendency and direction of the discussion [3].

Couple of methods can be implemented inside discussion activity to improve consensus quality such as recording, facilitation, and mediation [1]. Recording in this term stands for creating a physical record of what subject being discussed. Recording can be implemented by actually recording the whole discussion as a video file or even as simple as taking notes on participant's utterance. Facilitation in a second hand, help participants work together by providing artifact containing the discussion progress which everyone agrees on. Finally, mediation acts to help opposite parties deal with disagreement. In order to perform mediation, one independent person is needed to resolve disputes with his/her objective point of view.

II. Related Works

Researches related to consensus building has been conducted years ago. In general, researches focused on consensus building can be divided into 2 categories based on their focus point; process model and measurement model. All models are designed under the same goal: to improve consensus quality.

Process model focused on a set of rules that participants should follow under specific circumstances. A research might propose a straight forward approach e.g. using a fair, open, and freedom-focused process model [6], meaning that all perspective will be considered equal and all participants will have their freedom to disagree. A research might be focused specifically on a subprocess in consesus building e.g. Consensus Building Theory (CBT) [7] that emphasizes the cause of conflict to investigate what specific matter prevent or support a consensus to built. Other research is focused on a specific implementation of process model [3] e.g. propossing a short term and intensive workshop activity designed for Product Lifecycle Management (PLM) strategy planning phase which involves multi-party stakeholders. The workshop is then intended specifically for dicussion under digital transformation for smart, connected engineering field.

Meanwhile, measurement model focused more on the criteria to determine the consensus presence. Some popular methods are done by using standard deviation of voting results or using Kendall's coefficient on voting results [8]. A recent method showed that a digitized approach can be done by tracking every non-verbal aspects of each participant to determine consensus [5]. Another digitized approach has been done in 2 steps: the first step is inviting an external facilitator as one independent figure to direct the course of discussion, and the second step is doing text mining approach on the dialog data taken from the discussion to evaluate consensus [4].

III. Research Problem

Previous researches discussed on how consensus is being built by considering a lot of variables e.g. time consumed, participants contributions, and conflict resolution. In another hand, more variables should be valued more such as the objectivity of the final consensus and the latent opinion from each participants. This paper tries to discover latent opinion from each participants and avoid biased judgment which rarely discussed by previous researches. We proposed a text mining approach by utilizing topic model algorithms and a set of data as Background Knowledge. We are utilizing dialog data from PLM-themed discussion activity to detect hidden pattern and latent opinion from participants. Then, we will validate our findings with a professional consultant to discover the method's effectiveness. However, a preliminary study regarding this matter has been conducted [4] and this research act as the extension of it with approval from the original author.

IV. Proposed Method

In this research, we performed a digitized approach of dialog data from PLM-themed discussion activity sessions using data augmentation, topic model with background knowledge, and distribution similarity. First, the data will be prepared by a simple preprocess method and data augmentation. The clean and augmented data will then be experimented by various topic models and hyperparameters, we picked the best configuration and incorporate it into background-knowledge-backed topic model to generate topic distributions. Then, we will calculate the distribution similarity as convergence rate. Finally, a professional consultant will analyse the results to get an objective review. To summarize, we will take dialog data of discussion session and transform it into topic distributions, similarity value, and most frequent words (if necessary) from each discussion session to be validated by a professional consultant.

A. Data Augmentation

We took a real life dialog data from discussion sessions which ran for 1-2 hour long. Based on the dataset characteristics in Table I, the dataset we used is very poor. Compared to the common dataset in topic model research with specialization in short text data, our dataset size is 96.55% lower in terms of number of documents and 85.96% lower in terms of corpus size. Hence, we are using data augmentation techniques to improve dataset quality. We expand the Easy Data Augmentation [10] by adding additional processes: hypernym replacement and hyponym replacement. Hypernym and hyponym of a word is crucial as we thought the topic mixture of a sentence *s* should be the same with other sentence *s'* who has hypernym/hyponym relation with some words inside it.

Table I
Dataset Characteristics

Measures Type	PLM Workshop Dataset	Common Dataset [9]
Total Documents	383	11094
Corpus Size	686	4887
Average Length	4.83	7.84

B. Topic Model with Background Knowledge

We tried to mine latent opinion of the dataset using topic model with background knowledge. Topic model is an unsupervised learning approach where we could transform documents into document-to-topic distributions and topic-to-word distributions. In topic model point of view, document is a mixture of topic where topic itself is a mixture of word. The most popular method of topic model is Latent Dirichlet Allocation (LDA) [11], in which, most currently available topic model is proposed based on that. In LDA-based topic model, the learning process consists of generation process and sampling process. In generative process, the initial document-to-topic distributions and topic-to-word distributions are generated using hyperparameter α and β. Then, in the sampling process, distributions are evaluated by recalculating it using Gibbs Sampling for each and every word. The graphical notation of LDA topic model is shown in Fig. 1, while the generation algorithm is:

1: For each topic $k \in \{1, ..., K\}$:
2: Generate $\phi_k \sim \text{Dir}(\beta)$
3: For each document $d \in \{1, ..., D\}$:
4: Generate $\theta_d \sim \text{Dir}(\alpha)$
5: For each i, d where $d \in \{1, ..., D\}$ and $i \in \{1, ..., N_d\}$:
6: Generate $z_{id} \sim \text{Multinomial}(\theta_d)$
7: Generate $w_{id} \sim \text{Multinomial}(\phi_{z_{id}})$

where *K* is the number of topics, *D* is the number of documents, N_d is the number of words in document d, ϕ_k

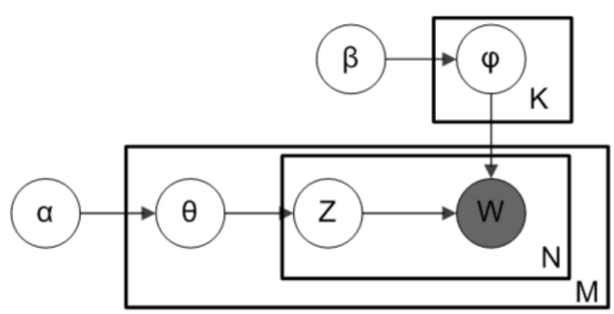


Figure 1. LDA plate notation

is topic-to-word distribution for topic k, θ_d is document-to-topic distribution for document d, z_{id} is topic for the i_{th} word in document d, and w_{id} is the i_{th} word in document d.

In our approach, we realize that our dataset has relatively smaller size compared to common topic model researches. Hence, we assembled various topic models with speciality in short text as suggested by [9]. The whole list of topic models could be seen in Table II that could be categorized into 4 types. The first type is standard, referred to the baseline topic model which is LDA. The second type, one-topic sampling based, will modify the inference process to sample only one topic for one document, meaning all words from a single document will only have one topic label. The third type, global word co-occurence based, will modify the document representation into word-network or set of biterms. Self-aggregation based as the last type will merge a number of documents into one single pseudo-document and then apply standard-type topic model to it.

After the experiment is done we will decide what is the best topic model, hyperparameters, and the number of sentence augmentation processes to use. After that, we will incorporate the result to a new background-knowledge-backed topic model called Source-LDA [12] as the most suitable topic model for our case. In Source-LDA, we are able to provide background knowledge data to influence topic labeling thus improving topic quality in the process.

C. Distribution Similarity

In this step, we aimed to picture the topic distribution into a single value that describes the rate of consensus built (agreement rate). In order to do this, we used distribution similarity calculation using Jensen-Shannon Divergence across all distributions [13]. This concludes the final step of our proposed method.

Table II
TOPIC MODEL EXPERIMENT

No.	Topic Model	Type
1	LDA [11]	Standard
2	Dirichlet Multinomial Mixture (DMM) [14]	One-topic sampling based
3	Latent-Feature LDA (LFLDA) [15]	
4	Latent-Feature DMM (LFDMM) [15]	
5	Generalized Polya Urn DMM (GPU-DMM) [16]	
6	GPU Poisson-based DMM (GPU-PDMM) [17]	
7	Biterm Topic Model (BTM) [18]	Global word co-occurence based
8	Word Network Topic Model (WNTM) [19]	
9	Self-aggregate Topic Model (SATM) [20]	Self-aggregation based
10	Pseudo-based Topic Model (PTM) [21]	

V. EXPERIMENT

We had an opportunity to utilize dialog data from requirement decisions (discussion session) of 4 Japanese companies. Data preprocessing and sentence augmentation is done to clean the data. The comparison of dataset characteristics before and after augmentation is shown in Table III. We managed to expand the dataset up to 1634.46% from the original size in terms of number of documents, and up to 145.04% in terms of corpus size.

Furthermore, the property for each sentence in dataset is presented in Table IV. There are 2 types of question which are problem, reffered to opinions retrieved at the early stage of discussion, and solution, reffered to opinions retrieved at the later stage of discussion. Response category is reffered to the participant's own divison, while organization level reffered to the participant's hiearchical level in the company.

Following the data preprocessing step, topic model experiment is conducted on all topic models in Table II. We used *topic coherence* to evaluate topic model performance because our dataset is raw and golden label for each sentence is not present [9]. The result of topic model experiment is shown in Fig. 2. The hyperparameter used in this experiment is number of iteration (1000-2000), α value (0.05-0.3), and β value (0.005-0.03). The value shown is the average of topic coherence score across various hyperparameters for each sentence augmentation processes. Based on the number of sentence augmentation process, 9 augmentation is not producing a significant result while 1 augmentation gives the best and most consistent result. Self-Augment Topic

Table III
AUGMENTED DATASET CHARACTERISTICS

Augmentation	Total Documents	Corpus Size	Average Length
No Augmentation	383	686	4.83
1 Sentence Augmentation	1017	922	5.00
9 Sentence Augmentation	5085	1519	5.10
12 Sentence Augmentation	6643	1681	5.10

Table IV
DATASET PROPERTY

Property Name	Possible value
Company ID	{1,2,3,4}
Question Type	{Problem, Solution}
Response Category	{Information Technology, Corporate Management, Business Process, Human Development}
Organization Level	{very low, low, medium, high, very high}
Opinion	{short sentence consists around 5 words}

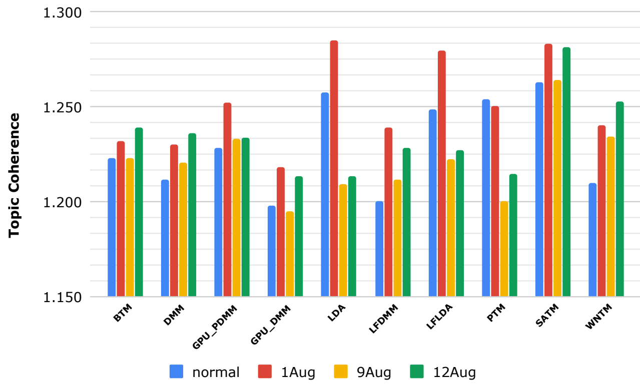


Figure 2. Topic coherence value based on topic model and sentence augmentation process performed on Dialog Data

Model (SATM) gives a good overall score regardless the number of sentence augmentation process. However, LDA held the best score of all experiment with 1 augmentation process. We picked LDA topic model with α value of 0.15, β vaue of 0.01, 2000 iteration, and 1 sentence augmentation process as the best configuration.

The next step is to incorporate this configuration into Source-LDA. Based on our research problem from Section III, PLM topics is used as the background knowledge data. We decided to use PTC Value Roadmap[1] because it contains 26 PLM Topics with complete definitions for each topics. The background knowledge dataset held a relative big size consisting of 26 topics, 1068 unique words, and 145.88 average document length.

Fig. 3 shows the topic coherence value relative to the number of sentence augmentation process applied to background knowledge dataset. The topic coherence value is increasing

[1] http://support.ptc.com/WCMS/files/28837/en/J1051_ValueRoadmap_TS.pdf

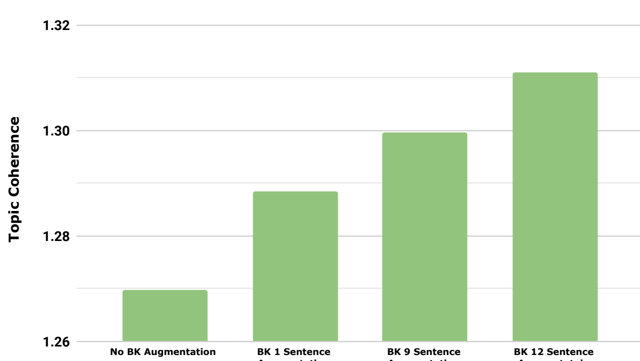


Figure 3. Topic coherence value based on background knowledge (BK) augmentation process

proportionally with the number of sentence augmentation. We only used LDA algorithm because it is the best configuriation we conclude from prior experiment. The usage of sentence augmentation on background knowledge improves topic coherence from 1.47% to 3.25%. We picked 12 sentence augmentation process on background knowledge as the best configuration. In addition, the usage of background knowledge also improves the topic coherence value from 1.307 (LDA without background knowledge) to 1.311 (LDA with background knowledge).

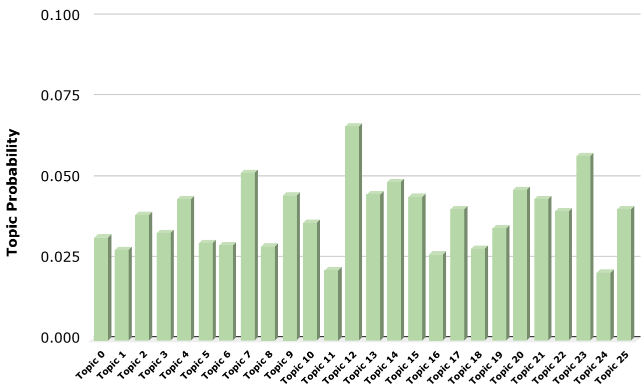
The end-result of this experiment is using LDA topic model, α value of 0.15, β value of 0.01, 2000 iteration, 1 sentence augmentation on dialog dataset, and 12 sentence augmentation on background knowledge as the go-to configuration for future use.

VI. Results and Discussion

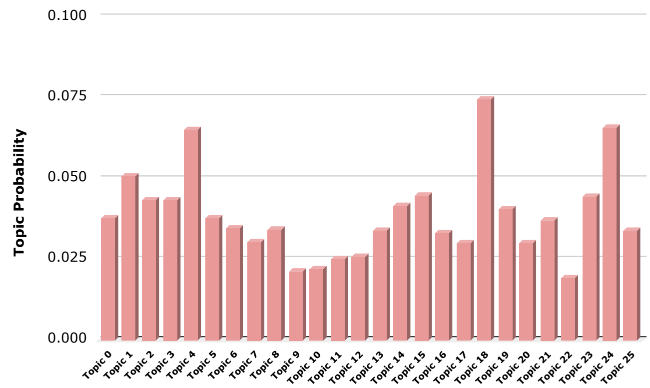
In this section, the qualitative evaluation of the result will be presented. The mapping of topic number with its actual name can be seen at Table V with a side note that the order of topic number is in alphabetical order and different with what shown in the reference (PTC Value Roadmap). There are 26 PLM Topics in total that serves as the best practice for a specific business unit.

The average topic distribution from each company calculated with the best configuration stated in Section V is shown in Fig. 4. The most probable and least probable topic is different for each company. For example, company 1 that runs in business process innovation industry discussed a lot about topic #12 (Project Management) and very little about topic #24 (Verification and Validation). Meanwhile, Company 2 from automotive manufacturer industry had a huge interest in topic #18 (Service Order Management and Field Service) but not in topic #22 (Technical and Service Parts Information Creation and Delivery). Company 3 from aqua industry has topic #19 (Service Parts Planning and Pricing) as the most probable topic and topic #2 (Component and Supplier Management) as the least probable topic. Lastly, company 4 from automotive supplier industry had topic #22 and topic #21 (System Architecture Design) as their most probable topics.

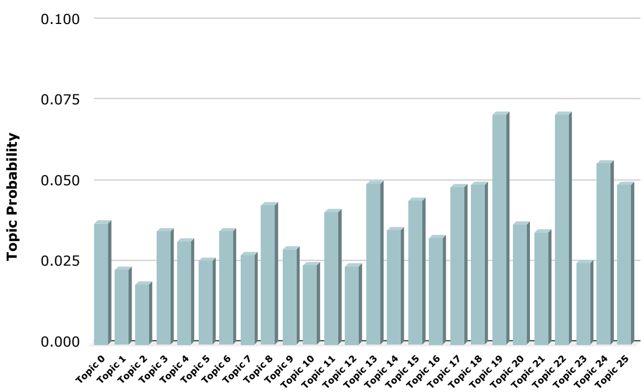
Topic distribution has finally obtained so we can proceed with similarity measurement using JS Divergence. All value approaching to 0 means that there is no variation between probability distributions, meanwhile value approaching to 1 means that there is high variation between probability distributions. The similarity of each discussion sessions can be seen at Table VI along with the top frequent words. The overall similarity achieved from each discussion is probably not too good since each one has similarity rate above 0.500. The lowest degree of similarity was achieved by Company 1 with 0.865 while the highest degree of similarity was achieved by Company 3 with 0.672. The top



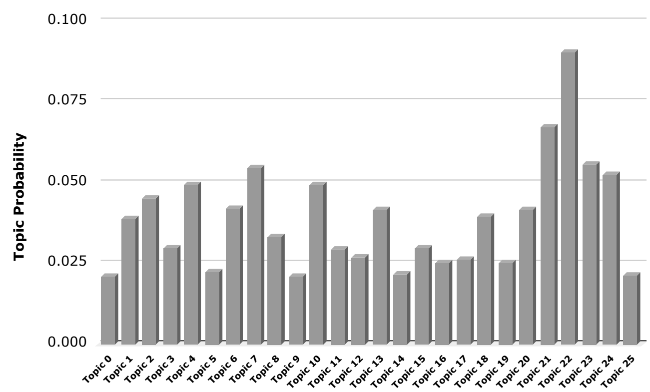
(a) Topic distributions for company 1



(b) Topic distributions for company 2



(c) Topic distributions for company 3



(d) Topic distributions for company 4

Figure 4. Average topic distributions of all company

words from each discussion acts as a support to understand the discussion more.

Given the results, here is the feedback from professional consultant for each discussion session:

1) Company 1: Company 1 had a key problem in terms of information exchange between design and manufacturing. I agree that the frequency of Design and Manufacturing topics was high. However, the topic of Project Management was rarely spoken directly by their voices. In addition, the analysis results show that there are few topics on Manufacturing Process Management. Certainly, there were few remarks on Manufacturing Process Management when the workshop was actually held. However, one of the participants was very concerned about the topic and he is one of the important people in the PLM project, so even if it is a minority opinions, I cannot ignore it as my consultant perspective. By the way, in the analysis results, the words with the highest frequency of occurrence were Information, Product, and Data. These were key words that participants often talked about during the actual workshop. As a consultant, I agree with that.

2) Company 2: The company 2 had three business unit. Thus, the participants had different opinions, as each business unit had a completely different product and each business model was different. When I looked at the results of this analysis, I thought that the reason that the topic of Verification and Validation was high was probably that they had a problem with their product quality. However, although the topic about Field Service has not been talked about in the actual workshop time, the frequency of topic 18 was high in this analysis result. In fact, this company does little field service work, so it is necessary to confirm why such analysis results were performed. In addition, the analysis results indicated that the frequency of Product Cost Management and Project Management topics was low. However, I think the discussions about costs and projects were relatively common during the actual discussions with them. Regarding the word distribution, the analysis result, it showed that the frequency of Production, Work, and Product were high. I agree with this result.

Table V
MAPPING OF PLM TOPICS

Topic No.	PLM Topics
Topic 0	Business System Support
Topic 1	Change and Configuration Management
Topic 2	Component and Supplier Management
Topic 3	Concept Development
Topic 4	Design and Manufacturing Outsourcing
Topic 5	Equipment Monitoring and Lifecycle Management
Topic 6	Manufacturing Process Management
Topic 7	Mechanical, Electrical, and Software Development
Topic 8	Performance Analysis and Feedback
Topic 9	Platform Design and Variant Generation
Topic 10	Product Cost Management
Topic 11	Product Support Analysis and Planning
Topic 12	Project Management
Topic 13	Quality and Reliability Management
Topic 14	Regulatory and Materials Compliance
Topic 15	Requirements Definition and Management
Topic 16	Service Diagnostics and Knowledge Management
Topic 17	Service Logistics and Network Management
Topic 18	Service Order Management and Field Service
Topic 19	Service Parts Planning and Pricing
Topic 20	Smart, Connected Product Enablement
Topic 21	System Architecture Design
Topic 22	Technical and Service Parts Information Creation and Delivery
Topic 23	Tooling Design and Manufacture
Topic 24	Verification and Validation
Topic 25	Warranty and Performance-based Contract Management

Table VI
SIMILARITY AND TOP WORDS

Company ID	Similarity Rate	Top words
Company 1	0.865	{Information, Product, Data}
Company 2	0.766	{Production, Work, Product}
Company 3	0.672	{Resource, Human, Product, Development}
Company 4	0.753	{Information, Data, Sharing}

3) Company 3: The motivation for Company 3 to introduce PLM was to strengthen its field service operations. Looking at the analysis results, it was found that the topics with the highest frequency were field services, such as Warranty management, Performance Based Contract Management, Technical Service Parts Information, and Service Order Management. I agree this result as a professional consultant. However, regarding the monitoring and management of equipment, it was analyzed that the topic frequency was low. This is different from the actual situation, because in the actual workshop, the story of equipment monitoring was relatively well discussed. The frequency of words of Resource, Human, Product, and Development is high. Even during the actual workshop discussion, the shortage of human resources in field service was very problematic. Thus, I agree with the analysis results.

4) Company 4: Company 4 has been practicing efforts to make its factory a smart factory. As a consultant, what I noticed in their actual workshops was their lack of information sharing between departments and insufficient training of employees. On the other hand, looking at the results of this analysis, we found that the topic # 22 was Technical and Service Part Information Creation and Delivery. At first, I wasn't interested in topic # 22. However, after reviewing the content of discussions with the workshop participants at a later date, there was an opinion that attention was paid to the management of service parts in order to contribute to sustainable sales. It seems that the results of this analysis have taught me a topic that I did not notice at first. Looking at the analysis results of the word distribution, it seems that three words, Information, Data, and Sharing, appear frequently. This was exactly the issue that was being talked about at the workshop. Additionally, The analysis results seem to indicate that there is no relationship between education and system design. Further investigation is needed as I think education topic should be highly related in the workshop.

Based on our analysis and feedback from professional consultant. We feel that our experiment on the usage of topic model with background knowledge in a industrial engineering discussion activity (in this case, PLM-themed) gives an actual contribution towards discussion overview in which, might improve consensus building process. The important takeaway of this research is that topic modeling with background knowledge will assist professional consultant to understand more towards participant's latent opinion.

VII. CONCLUSION AND FUTURE WORKS

In this paper, we analysed a new digitized approach to improve consensus building process in discussion activity held within industrial engineering context (PLM-themed). Our proposed method consists of performing data augmentation, implementing topic model with background knowledge, and calculating the distribution similarity. Finally, we validate the result on professional consultant. We received good feedback which validate our purpose of using a new approach to improve consensus building process.

However, further approach is still necessary based on two perspective: consensus building and topic modeling. From consensus building perspective, we still need to assure the emotional state of discussion participants when dialog data is recorded. Some variables might aspect the quality and consistency of participant's opinion. Meanwhile from topic modeling perspective, we are planning to expand Source-LDA so it can afford different data representation like BTM and WNTM does.

References

[1] J. Thomas-Lamar, S. Mckearnan, and L. Susskind, *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*. SAGE Publications, 1999. pp. 7–9.

[2] N. He, S. Yao, and O. Yoshie, *Emotional speech classification in consensus building*. 2014 10th International Conference on Communications (COMM), Bucharest, 2014, pp. 1-4.

[3] S. Goto, O. Yoshie, and S. Fujimura, *Preliminary Study: Text mining approach to dialog data of stakeholders on requirement decision for Enterprise Information System*. 2019 10th Annual European Decision Sciences Institute (EDSI) Conference, Nottingham, 2019.

[4] S. Goto, O. Yoshie, and S. Fujimura, *Empirical study of multi-party workshop facilitation in strategy planning phase for Product Lifecycle Management (PLM) system*. 2019 International Federation for Information Processing (IFIP) International Conference on Product Lifecycle Management, Moscow, 2019, pp. 82-93.

[5] Y. Katagiri et al., *Implicit proposal filtering in multi-party consensus-building conversations*. Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue, Columbus, 2008, pp. 100-103.

[6] C. T. L. Butler and A. Rothstein, *On conflict and consesus; A handbook on formal consesus decision making*. Food Not Bombs Publishing, Takoma Park, 2004.

[7] R. O. Briggs, G. Kolfschoten, and G. J. de Vreede, *Toward a theoretical model of consensus building*. 11th Americas Conference on Information Systems, 2005.

[8] M. M. Shepherd and W. B. Martz Jr, *Group consensus: Do we know it when we see it?*. Proceedings of the Hawaii International Conference on System Science, Los Alamitos, 2004.

[9] J. Qiang, Z. Qian, Y. Li, Y. Yuan, and X. Wu, *Short text topic modeling techniques, applications, and performance: a survey*. 2019.

[10] J. Wei and K. Zou, *EDA: Easy Data Augmentation techniques for boosting performance on text classification tasks*. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, 2019, pp. 6383-6389.

[11] D. M. Blei, A. Y. Ng, and M. I. Jordan, *Latent Dirichlet Allocation*, *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*. Vancouver, 2001, pp. 601-608.

[12] J. Wood, P. Tan, W. Wang, and C. Arnold, *Source-LDA: Enhancing probabilistic topic models using prior knowledge sources*. 33rd IEEE International Conference on Data Engineering (ICDE), San Diego, 2017, pp. 411-422.

[13] J. A. Aslam and V. Pavlu, *Query hardness estimation using Jensen-Shannon Divergence among multiple scoring functions*. 29th European Conference on Information Retrieval Research (ECIR), Rome, 2007, pp. 198-209.

[14] J. Yin and J. Wang, *A dirichlet multinomial mixture model-based approach for short text clustering*. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York,2014, pp. 233-242.

[15] D. Q. Nyugen, R. Billingsley, L. Du, and M. Johnson, *Improving topic models with latent feature word representations*. Transactions of the Association for Computational Linguistics vol. 3, 2015, pp. 299-313.

[16] C. Li et al., *Topic modeling for short texts with auxiliary word embeddings*. Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, Pisa, 2016, pp. 165-174.

[17] C. Li et al., *Enhancing topic modeling for short texts with auxiliary word embeddings*. ACM Transactions on Information Systems (TOIS), 2017, pp. 1-30.

[18] X. Cheng, X. Yan, Y. Lan, and J. Guo, *Btm: Topic modeling over short texts*, *IEEE Transactions on Knowledge and Data Engineering*. 2014, pp. 2928-2941.

[19] Y. Zuo, J. Zhao, and K. Xu, *Word network topic model: a simple but general solution for short and imbalanced texts*. Knowledge and Information Systems, 2016, pp. 379-398.

[20] X. Quan, C. Kit, Y. Ge, and S. J. Pan, *Short and sparse text topic modeling via self-aggregation*. Proceedings of the 24th International Conference on Artificial Intelligence, Buenos Aires, 2015, pp. 2270-2276.

[21] Y. Zuo et al., *Topic modeling of short texts: A pseudo-document view*. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, San Fransisco, 2016, pp. 2105-2114.

[22] K. Ondrej and J. Marlin, *Product life cycle in digital factory*. Knowledge management and innovation: a business competitive edge perspective, Cairo, 2010, pp. 1881-1886.

Appendix

Preliminary Experiment

We conduct a preliminary experiment to prove the effectiveness of our method. The dialog data we used is an original data with very limited context and small corpus size. Hence, we would like to validate our method by experimenting it on a widely-used dataset. We performed data augmentation processes on Biomedical dataset taken from [9] which has 20 topics, 4498 corpus, 19448 documents, and 7.44 average document length. Then, we perform topic modeling using LDA, BTM, and PTM algorithm. Finally, we evaluate it by calculating their *topic coherence* value. After our preliminary experiment is finished, Fig. 5 shows that data augmentation will improve topic quality on a certain degree.

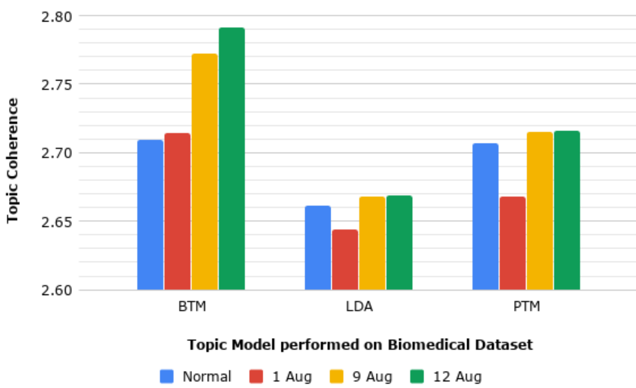


Figure 5. Topic coherence value based on topic model and sentence augmentation process performed on Biomedical Dataset