Aspect Extraction Performance with POS Tag Pattern of Dependency Relation in Aspect-based Sentiment Analysis

Conference Paper · March 2018

DOI: 10.1109/INFRKM.2018.8464692

4 authors:

Ana Salwa
Universiti Teknologi MARA

Nurfadhlina Bt Mohd Sharef
Universiti Putra Malaysia

Masrah Azrifah Azmi Murad
Universiti Putra Malaysia

Azreen Azman
Universiti Putra Malaysia

Aspect Extraction Performance With POS Tag Pattern of Dependency Relation in Aspect-based Sentiment Analysis

Ana Salwa Shafie
Dept of Computer Science
Universiti Putra Malaysia
Serdang, Malaysia
anasaf01@gmail.com

Nurfadhlina Mohd Sharef
Dept of Computer Science
Universiti Putra Malaysia
Serdang, Malaysia
nurfadhlina@upm.edu.my

Masrah Azrifah Azmi Murad
Dept of Computer Science
Universiti Putra Malaysia
Serdang, Malaysia
masrah@upm.edu.my

Azreen Azman
Dept of Computer Science
Universiti Putra Malaysia
Serdang, Malaysia
azreenazman@upm.edu.my

***Abstract*—The most important task in aspect-based sentiment analysis (ABSA) is the aspect and sentiment word extraction. It is a challenge to identify and extract each aspect and it specific associated sentiment word correctly in the review sentence that consists of multiple aspects with various polarities expressed for multiple sentiments. By exploiting the dependency relation between words in a review, the multiple aspects and its corresponding sentiment can be identified. However, not all types of dependency relation patterns are able to extract candidate aspect and sentiment word pairs. In this paper, a preliminary study was performed on the performance of different type of dependency relation with different POS tag patterns in pre-extracting candidate aspect from customer review. The result contributes to the identification of the specific type dependency relation with it POS tag pattern that lead to high aspect extraction performance. The combination of these dependency relations offers a solution for single aspect single sentiment and multi aspect multi sentiment cases.**

Keywords—aspect extraction, dependency relation, POS tag patterns, extraction rule, aspect-based sentiment analysis

I. Introduction

The rapid development of Internet technologies provide users with various platforms to buy and shopping online. This medium gives the opportunities to users to freely express their opinions about a product or services through the reviews which have caused huge amount of online reviews available on the Web. These reviews are useful in customer purchase decision and help manufacturer and organizations to improve the quality of their product and services. However, it is difficult for people to read and summarize the large number of reviews accordance to their needs. Therefore, an automated approach for extracting and summarizing opinions expressed in reviews is essential. Thus, sentiment analysis has become a research interest and challenging tasks.

Sentiment analysis (SA) is the study of analysing people's opinions, sentiments, appraisals, attitudes, and emotions toward entities such as products, services, individuals and their aspects expressed in textual reviews [1]. SA can be looked from three different levels of granularity which are document-level, sentence-level and aspect-level [2], [3], [4]. In document-level SA the overall sentiment polarity expressed in reviews is identified. In sentence-level SA the sentiment polarity expressed in each sentence of the review is identified. Whilst aspect-level SA (also known as aspect-based sentiment analysis), identifies the sentiment polarities on each different aspects of one target entity [5], [6].

Aspect-based Sentiment Analysis (ABSA) has recently become one of the most interest researches in the field of sentiment analysis. In product review people usually comment on multiple aspect and give different sentiment on various aspects of that product. Specifically, in a review it might consists of four issues: (1) single aspect and single sentiment, (2) single aspect and multiple sentiments, (3) multiple aspects and single sentiment, and (4) multiple aspects and multiple sentiments. Accordingly, a new branch of sentiment analysis has emerged, called multi-aspect sentiment analysis (MASA) that aims to consider the various aspects of product or services that have been discussed within a review.

It is a challenge to deal with review sentence that consists of multiple aspects with various polarities expressed to multiple sentiments. Therefore, it is essential to identify and extract each aspect and it specific associated sentiment word correctly. Fig. 1 gives an example of review sentence that consist of multiple aspects and multiple sentiments. This review example is among the customer review that is taken from the laptop domain in SemEval 2014 dataset. The first aspect is 'display' which is associated with two sentiment words 'best' and 'long time'. Both sentiment words expressed the positive sentiment. Same as for second aspect 'battery life' also associated with two sentiment words 'long' and 'convenient' and expressed the positive sentiment.

The display on this computer is the best I've seen in a very long time, the battery life is very long and very convenient.

Fig. 1. Example of review sentence.

The most important task in ABSA is aspect and sentiment word extraction. This task aims to efficiently identify and extract aspects and sentiment word regarding that aspect from reviews. Existing works used various aspect extraction approaches which are: (1) extraction based on frequent nouns and noun phrases, (2) extraction by exploiting opinion and aspects relations, (3) extraction by supervised learning, and (4) extraction using topic modeling [7]. Researchers find that there often exists relations between aspect and sentiment words in review sentences, thus their relationships can be exploited to extract aspects [8], [9]. Previous research has shown that unsupervised methods based on dependency relations are promising for aspect extraction [10]. Such approaches firstly analyse term dependency relations in review sentences, and then apply some rules and algorithms to extract product aspects from the identified dependency relations [8].

In dependency rule-based approach, the consideration of word to be a candidate aspect or sentiment word are based on the type dependency relation, the part-of-speech (POS) tag of the word in that relation, and rule of extraction. However, the main challenges are that large numbers of aspects are not extracted by the rules and some of the extracted words are not the aspects. This issue contributes to the lower precision and recall. It is difficult to develop a generalized dependency-based rule extraction which could be implemented in any review format and domain area. Thus, it required a lot of effort and various type dependency patterns to develop the extraction rule that suit with the domain.

Thus, the main objective of this study is to perform a preliminary study in order to measure the extraction performance of different type of dependency relation in product review. This study contribution to the identification of the most potential type dependency relation with it POS tag pattern in extracting more correct aspects. The combination of these dependency relations can solve the single aspect single sentiment and multi aspect multi sentiment cases. It also will assist in developing the generalized dependency-based rule extraction.

II. RELATED WORKS

The earliest work in aspect-based sentiment analysis by [11] was employed frequent nouns and noun phrases approach to extract aspect. There are various numbers of approaches that have been studied by researchers that focus on aspect extraction problem. The most widely approach is dependency rule-based approach [12] that extracts aspect based on the grammatical relationship between word in review sentence. The existing work shows that the dependency approach has been potentially employed by different approaches to extract aspect from unstructured reviews.

Kang and Zhou [13] extracted subjective feature by extending double propagation with indirect dependency and comparative construction using dependency relation and rules. Fernandez-Gavilanes et al. [14] employed dependency parsing-based text classification to predict sentiment in online textual messages. Wang et al. [15] employed syntactic pattern and rules to extract opinion word and target in propagation process and employed dependency parser to identify the potential relations between opinion word and aspect in refinement process. Agarwal et al. [16] presented a concept extraction algorithm based on a novel concept parser scheme to extract semantic features that exploited semantic relationships between words in natural language text. Poria et al. [17] proposed a rule-based approach that exploits common-sense knowledge and dependency relations between aspects and opinion words to detect both explicit and implicit aspects.

Liu et al. [18] and Liu et al. [1] proposed an automated rule selection approach, based on dependency tree parsing, to select the most appropriate rules for every aspect. Zheng et al. [19] proposed an unsupervised dependency analysis-based approach to firstly extract Appraisal Expression Patterns (AEPs) from reviews. Raghuveer [20] used different combinations of dependencies to extract the candidate product feature opinion pairs in an effective and meaning full way. Qiu et al. [21] analyzed manually and summarized eight dependency relation-based rules to identify syntactic relations that link opinion words and target in bootstrapping process. Samha and Li [22] developed extraction rule by applying the best combination of dependency. In [23], a syntactic based approach with the help of dependency parser to extract opinion words was proposed.

All the above works employed different dependency relations in extracting aspect and sentiment word with different methodologies from one another. However, most of the previous works do not describe the dependency relation analysis in detail. The selection of typed dependency relation used in their work also did not mentioned and no specific justification and evaluation performed on each of the type dependency relations that have been used in their study. In this study, we focus on measuring the performance of different type dependency relations.

III. METHODOLOGY

The proposed methodology is carried out in four mains steps: (1) pre-processing; (2) POS tagging; (3) dependency parsing; and (4) dependency relation analysis.

A. Preprocessing

In customer review, users usually write in casual style, use various unstructured text format and using several unnecessary symbols. It is necessary to preprocess data to remove the noisy elements and normalized the language [14]. Through this, it will help to reduce the complexity of dependency relation of a review sentence. Hence, in the preprocessing task the noise element consist of useless characters and symbols such as --, *, =, /, [, :), :D have been removed. Hence, certain symbols or punctuations will be remained to preserve the authenticity dependency grammar between words.

B. POS Tagging

Part-of-Speech (POS) plays an important role in the identification of aspect and sentiment word. In this step, part-of-speech (POS) tagging is performed for each review sentence using Stanford CoreNLP. Fig. 2 shows the example of POS tag for the review sentence in Fig. 1. In this work, the POS tag is used to identify the word in the review sentence that is nouns, adjective and verb. Usually the nouns or noun phrases and adjective resulting from POS tag were represented as aspect and sentiment word respectively. In some cases, verbs and adverb also could represent as sentiment word. The POS tag used in this work is based on the Penn Treebank English POS Tag. Table I shows the list of POS tag that have been used in determining the POS tag pattern of dependency relation.

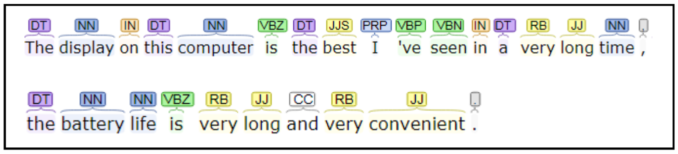

Fig. 2. Example of POS Tag.

TABLE I. LIST OF POS TAG

POS Tag	Description	Indication
NN/NNS/NNP/NNPS	Nouns	Aspect
JJ/JJR/JJS	Adjectives	Sentiment
VB/VBD/VBG/VBN/VBP/VBZ	Verbs	Sentiment
RB/RBR/RBS	Adverb	Sentiment

C. Dependency Parsing

The dependency parsing is applied to get the syntactic grammatical dependency relation between words in the review sentence using Stanford Parser (http://nlp.stanford.edu). Exploiting the appropriate dependency relation contribute to the identification of potential noun phrase aspect, sentiment word and the aspect-sentiment word pairs.

The dependency relation holds the grammatical relation between a governor (also known as head) and a dependent (modifier) [24]. In Stanford dependencies (SD), it is represented as triplets: name of the relation, governor and dependent. Fig. 3 shows the example of dependency parsing of review sentence in Fig. 1. From the dependency parsing, the type dependency relations (TDR) between governor and dependent can be identified in order to extract the most relevant aspect and sentiment word. Based on dependency parsing result in Fig. 3, the dependency relation nsubj(best-8, display-2), nsubj(long-23, life-20) and compound(life-20, battery-19) will be helpful in identifying aspects. Meanwhile the dependency relation amod(time-16 , long-15), conj(long-23, convenient-26) will be helpful in identifying sentiment words. The pre-extraction of potential product aspects has been implemented through a dependency relation analysis that has been discussed in the next topic.

root (ROOT-0 , long-23)	det (time-16 , a-13)
det (display-2 , The-1)	advmod (long-15 , very-14)
nsubj (best-8 , display-2)	**amod (time-16 , long-15)**
case (computer-5 , on-3)	nmod (seen-11 , time-16)
det (computer-5 , this-4)	det (life-20 , the-18)
nmod (display-2 , computer-5)	**compound (life-20 , battery-19)**
cop (best-8 , is-6)	**nsubj (long-23 , life-20)**
det (best-8 , the-7)	cop (long-23 , is-21)
ccomp (long-23 , best-8)	advmod (long-23 , very-22)
nsubj (seen-11 , I-9)	cc (long-23 , and-24)
aux (seen-11 , 've-10)	advmod (convenient-26 , very-25)
acl:relcl (best-8 , seen-11)	**conj (long-23 , convenient-26)**
case (time-16 , in-12)	

Fig. 3. Example of Dependency Parsing

D. Dependency Relation Analysis

The dependency relation analysis is performed to identify relevant TDR and measure the performance of each TDR in pre-extracting aspect and sentiment word. In this task, candidate aspect are identified and analysed through a common POS tag pattern of dependency relation. This task is performed in three steps: (1) select relevant TDR, (2) determine POS tag pattern, and (3) extract product aspect.

There can be many different types of dependency relations in the sentences, however not all of them are helpful and contribute in identifying aspect and sentiment word [8]. The previous study that employed dependency relation in their work have used the following TDR: ‘prep’, ‘nsubj’, ‘csubj’, ‘xsubj’, ‘dobj’, ‘iobj’, ‘conj’, ‘xcomp’, ‘acomp’, ‘amod’, ‘nmod’, ‘advmod’, ‘rcmod’, ‘acl’, ‘compound’. The definition and details explanation of these TDR can be found in [24]. Previously Raghuveer [20] only used three TDR which are “dobj”, “obj” and “xobj” for identifying the aspect terms, while Yan et al. [8] extracted four major types of dependency relations such as ‘nsubj’, ‘amod’, ‘rcmod’ and ‘dobj’ and generated a set of candidate feature-sentiment pairs from all product reviews. Meanwhile Kang and Zhou [13], used eleven TDR to extract subjective feature by extending double propagation with indirect dependency and comparative construction.

Therefore, in the first step, an experiment on these fifteen TDR has been performed on the SemEval 2014 dataset to select the relevant TDR that could identify the aspect and sentiment word. From the observation of the result, it is found that ‘nsubj’, ‘dobj’, ‘amod’ and ‘nmod’ are the most excellent TDR that have the capability to identify the potential aspect and sentiment word. The ‘acl’ relation only gives a very little result but still can be used to identify the aspect. Meanwhile, the ‘compound’ relation facilitates in the identification of the noun phrase or multi word aspect. Even though most of the ‘conj’ relation result gives the connection between two sentiment words, however there are yet exist connection between aspects as well. In other words, ‘conj’ relation could also facilitate the identification of multiple aspect of a review. There is no dependency result traced out for ‘prep’, ‘iobj’, ‘xsubj’, and ‘rcmod’ relation. A very little result obtained from ‘csubj’ relation; however, the result is not extracting the aspect. Meanwhile the ‘acomp’, ‘xcomp’ and ‘advmod’ would not be used since these relations are more relevant to identify sentiment word. Qadir [20] used ‘acomp’, ‘xcomp’ and ‘advmod’ relation in their work to identify opinion sentence.

Hence, this work only focuses on seven TDR specifically ‘nsubj’, ‘dobj’, ‘amod’, ‘nmod’, ‘acl’, ‘conj’ and ‘compound’ due to their capability to directly extract the aspect and sentiment word, and able to tackle the multi aspects and multi sentiments issue. The other TDR would be considered in the development of dependency rule extraction as they work as complement to the other TDR.

The second step is to determine the POS tag pattern for governor and dependent of each selected TDR. Each TDR may consist of several different POS tag patterns. The POS tag pattern is design based on the POS tag of governor and dependent of the relation that represent aspect and sentiment word. The list of POS tag used can be found in Table I. For example, 'nsubj' relation consists of two types of pattern. The first pattern is, governor as adjective and dependent as nouns that represent sentiment word and aspect respectively. Therefore, the POS tag pattern is nsubj (JJ, NN).

The second pattern is governor as verb and dependent as nouns that represent sentiment word and aspect respectively. The POS tag pattern is nsubj (VB, NN). TDR ID is given as nsubj1 and nsubj2 to differentiate two POS tag pattern of 'nsubj' relation. Details of TDR with it designated ID, POS tag pattern and the extraction rules that have been used in this work are tabulated in the Table II.

The third step is extracting the aspect of product using extraction rule. The extraction rule is derived based on the TDR and POS tag pattern of that TDR. For example, if dependencies parsing of a review sentence contain 'nsubj' relation, and the POS tag of governor and dependent is "JJ" and "NN" respectively, therefore the dependent is an aspect. This example describes the extraction rule of TDR ID nsubj1 as can be seen in third column of Table II. The extraction rule is applied to pre-extract the candidate aspect and sentiment word.

TABLE II. TDR POS TAG PATTERN AND EXTRACTION RULE

TDR ID	POS Tag Pattern	Extraction Rule
nsubj1	nsubj (JJ/JJR/JJS, NN/NNS/NNP)	If the relation is nsubj and match the pattern, therefore the governor is opinion and the dependent is aspect.
nsubj2	nsubj (VB/VBD/VBG/VBN/VBP/VBZ, NN/NNS/NNP)	
amod1	amod (NN/NNS/NNP, JJ/JJR/JJS)	If the relation is amod and match the pattern, therefore the governor is aspect and the dependent is opinion.
amod2	amod (NN/NNS/NNP, VB/VBD/VBG/VBN/VBP/VBZ)	
dobj	dobj (VB/VBD/VBG/VBN/VBP/VBZ, NN/NNS/NNP)	If the relation is dobj and match the pattern, therefore the governor is opinion and the dependent is aspect.
nmod1	nmod (NN, NN/NNS)	If the relation is nmod and match the pattern, therefore both words are aspects.
nmod2	nmod (JJ, NN)	If the relation is nmod and match the pattern, therefore the governor is opinion and the dependent is aspect.
acl1	acl (NN, JJ)	If the relation is acl and match the pattern, therefore the governor is aspect and the dependent is opinion.
acl2	acl (NNS, VBP)	
conjA1	conjA (NN, NN/NNS/NNP)	If the relation is conj and match the pattern, therefore both words are aspects.
conjA2	conjA (NN/NNS/NNP, JJ)	If the relation is conj and match the pattern, therefore the governor is aspect and the dependent is opinion.
conjA3	conjA (NN/NNS/NNP, VBZ)	
compound	compound (NN, NN)	If the relation is compound and match the pattern, therefore both words are aspects.

IV. PRELIMINARY RESULTS

The experiment is conducted using SemEval 2014 shared task datasets. This dataset consists of two domains: restaurant and laptop. Table III shows the details information about the datasets. For each review in both datasets, the aspect term and their polarity are already annotated, and aspect category only annotated for restaurant datasets. In this work, the experiment and evaluation are performed on training dataset only hence the evaluation not yet on the testing data.

TABLE III. INFORMATION OF SEMEVAL 2014 DATASET

Domain	Number of Review		
	Training	Testing	Total
Laptop	3045	800	3845
Restaurant	3041	800	3841

The performance is measured using evaluation metrics precision (P), recall (R) and F1-score (F1). Precision, recall and F1-score is calculated using true positive (TP), false positive (FP) and false negative (FN) as shown in Eq. (1) until Eq. (3). TP is the number of word extracted that is correct aspect, FP is the number of word extracted that is incorrect aspect and FN is the number of word that is aspect, but not extracted.

$$\text{Precision} = TP / (TP + FP) \tag{1}$$

$$\text{Recall} = TP / (TP + FN) \tag{2}$$

$$\text{F1-score} = \frac{(2 * \text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{3}$$

The experimental result is used to measure the performance of each TDR with POS tag pattern in extracting correct aspect. The correct aspect is calculated based on the number of correct word extracted compared to the number of word in actual aspect. For example, the actual aspect "battery life" consists of two words. By applying extraction rule of TDR ID nsubj1 to the dependency relation nsubj (long, life), the aspect extracted is "life". In this case, only word 'life' is correctly extracted, therefore the number of correct aspect word is considered as one out of two. Table IV shows the total number of aspect and total number of word in actual aspect for laptop and restaurant domain.

TABLE IV. ASPECT INFORMATION OF THE SEMEVAL 2014 DATASET

Aspect Information	Laptop	Restaurant
Total number of aspect	2358	3693
Total number of aspect word	3492	5120

The performance of aspect extraction in laptop and restaurant domain is summarized in Table V and Table VI. Even though some of the POS tag patterns of TDR have extracted both aspect and sentiment word, but the evaluation results only measure the performance of aspect extraction. The results have been compiled from the highest to lowest F1-score. Overall results showed restaurant domain performed better than in laptop domain. However, in both cases recall was lower than precision and consequently causes lower value to F1-score. The reason is because not all words in a multi-word aspect (aspect phrase) are extracted as well as has led to an increased number of not extracted aspects. As can be seen from Table V and Table VI, FN column presents the high number of aspects that still not extracted from the review due to this problem.

TABLE V. EXTRACTION PERFORMANCE ON LAPTOP DATASET

TDR ID	TP	FP	FN	P	R	F1
compound	1037	1489	2455	**41.05**	**29.70**	**34.46**
amod1	570	1359	2922	29.55	16.32	21.03
dobj	487	1458	3005	25.04	13.95	17.91
nsubj2	307	600	3185	33.85	8.79	13.96
conjA1	273	315	3219	**46.43**	7.82	13.38
nmod1	244	712	3248	25.52	6.99	10.97
nsubj1	179	129	3313	**58.12**	5.13	9.42
nmod2	69	178	3423	27.94	1.98	3.69
amod2	31	44	3461	41.33	0.89	1.74
conjA2	7	11	3485	38.71	0.34	0.68
acl2	5	31	2087	13.89	0.24	0.47
conjA3	7	8	3485	**46.67**	0.20	0.40
acl1	12	19	3480	38.89	0.20	0.40

TABLE VI. EXTRACTION PERFORMANCE ON RESTAURANT DATASET

TDR ID	TP	FP	FN	P	R	F1
compound	1486	1322	3634	**52.92**	**29.02**	**37.49**
amod1	1147	1105	3973	50.93	22.40	31.12
nmod2	102	175	5018	36.82	1.99	3.78
nsubj1	663	108	4457	**85.99**	12.95	22.51
conjA1	555	193	4565	**74.20**	10.84	18.92
dobj	596	698	4524	46.06	11.64	18.58
nmod1	544	846	4576	39.14	10.63	16.71
nsubj2	327	376	4793	46.51	6.39	11.23
acl1	36	25	5084	59.02	0.70	1.39
amod2	24	9	5096	72.73	0.47	0.93
conjA2	12	20	5108	37.50	0.23	0.47
acl2	7	17	5113	29.17	0.14	0.27
conjA3	1	9	5119	10.00	0.02	0.04

The aspect extraction performance between different TDR shows that 'compound' relation achieved the best performance for both domain with 41.05% of good precision, 29.70% of highest recall and 34.46% of highest F1-score value for laptop domain. Similarly, in the case of restaurant domain, 'compound' relation achieved good precision of 52.92%, highest recall of 29.02% and highest F1-score value of 37.49%. However, 'nsubj1' obtained highest precision among the others TDR with 58.12% and 85.99% of precision for laptop and restaurant domain respectively. The POS tag pattern of 'nsubj1' able to extract aspect correctly with the small number of false aspect extracted. It can be seen in TP and FP column, where the false positive value of 'nsubj1' is relatively small compared with the true positive value in both domains.

The 'conj' relation also presents the good performance. As can be seen in Table V, the TDR ID 'conjA3' and 'conjA1' achieved precision of 46.67% and 46.43% respectively. Meanwhile 'conjA1' from Table VI achieved 74.20% of precision. This indicates that POS tag pattern of conjA3 and conjA1 contribute to the high performance of multi aspects extraction. However, the result of this study cannot be compared to existing work because this result is the preliminary result of dependency relation analysis, while the existing work already implemented the dependency relation in their proposed techniques.

Based on this result, the generation of more comprehensive and generalized dependency-based rules extraction would be much easier and more reliable. However, we cannot directly rely on that relation only because the combination with other dependencies might also contribute to the finding of others potential aspect [20]. For example, 'amod', 'nsubj' and 'dobj' relation can be used to identify single aspect that associated with one or more sentiment words. However, it only returns single aspect word. Hence, the use of 'compound' relation is essential to extract aspect that consists of multi-words. Meanwhile conj:and (NN, NN) will return multiple aspect and conj:and (JJ, JJ) will return the multiple sentiment of an aspect. Therefore, the combination of these dependency relations can solve the single aspect single sentiment and multi aspect multi sentiment cases. More detail extraction rules is essential to be considered to achieve high performance and accuracy.

V. CONCLUSION

Aspect extraction is the most challenging task in multi aspect sentiment analysis. There is huge number of approaches have been studied by researches to carry out this task. Dependency based is one of promising approach for aspect extraction. This paper provides a preliminary study on the performance of dependency relation in pre-extracting the aspect using POS tag patterns. The extraction rule was generated by exploiting the type dependency relation and POS tag patterns of the governor and dependent of that relation.

From the evaluation that has been carried out, the specific type dependency relation with it POS tag pattern that could give highest extraction performance has been identified. The results presented are based on the investigation of the performance of the POS tag patterns on multi aspect multi sentiment issues. Hence it would be the basis for the generation of dependency-based extraction rule with the appropriate selection and combination of the identified TDR POS tag pattern. By means of appropriate TDR combination, the single aspect single sentiment and multi aspect multi sentiment cases can be solved. More accurate aspects extracted would be expected.

In future work, this work can be further applied to extract and evaluate the sentiment words that associated with each extracted aspect. Besides that, an appropriate pruning method will be applied to reduce the false aspects thus increase the recall. This work also will be implemented and evaluated using testing data and another domain.

ACKNOWLEDGMENT

The authors wish to thank Universiti Putra Malaysia for funding this research. The authors gratefully acknowledge the generous scholarship support of Ministry of Higher Education.

REFERENCES

[1] Q. Liu, Z. Gao, B. Liu, and Y. Zhang, "Automated rule selection for opinion target extraction," *Knowledge-Based System*, vol. 104, pp. 74–88, 2016.

[2] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no.

4, pp. 1093–1113, May 2014.

[3] K. Ravi and V. Ravi, “A survey on opinion mining and sentiment analysis: Tasks, approaches and applications,” *Knowledge-Based System*, vol. 89, pp. 14–46, 2015.

[4] T. A. Rana and Y.-N. Cheah, “Aspect extraction in sentiment analysis: comparative analysis and survey,” *Artifitial Intelligent Review*, 2016.

[5] L. Zhang and B. Liu, “Aspect and Entity Extraction for Opinion Mining,” in Data Mining and Knowledge Discovery for Big Data, Springer-Verlag Berlin Heidelberg, pp. 1–40, 2014.

[6] E. Ş. Chifu, T. Ş. Letia, and V. R. Chifu, “Unsupervised aspect level sentiment analysis using Ant Clustering and Self-organizing Maps,” in *2015 International Conference on Speech Technology and Human-Computer Dialogue, SpeD 2015*, 2015.

[7] B. Liu, “Sentiment Analysis and Opinion Mining,” *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[8] Z. Yan, M. Xing, D. Zhang, and B. Ma, “EXPRS: An extended pagerank method for product feature extraction from online consumer reviews,” *Information and Management*, vol. 52, no. 7. Elsevier B.V., pp. 850–858, 2015.

[9] S. Poria, E. Cambria, L.-W. Ku, C. Gui, and A. Gelbukh, “A Rule-Based Approach to Aspect Extraction from Product Reviews,” SocialNLP 2014, p. 28, 2014.

[10] Y. Xia, E. Cambria, and A. Hussain, “AspNet: Aspect Extraction by Bootstrapping Generalization and Propagation Using an Aspect Network,” *Cognitive Computation*, pp. 241–253, 2015.

[11] M. Hu and B. Liu, “Mining and summarizing customer reviews,” in Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 168–177, ACM, 2004.

[12] T. A. Rana and T. A. Rana, “Exploiting Sequential Patterns to Detect Objective Aspects from Online Reviews,” no. August, 2016.

[13] Y. Kang and L. Zhou, “RubE: Rule-based methods for extracting product features from online consumer reviews,” *Information and Managemant.*, vol 54, no. 2, pp. 166-176, 2016.

[14] M. Fernandez-Gavilanes, T. Alvarez-Lopez, J. Juncal-Martinez, E. Costa-Montenegro, and F. Javier Gonzalez-Castano, “Unsupervised method for sentiment analysis in online texts,” *Expert System Application.*, vol. 58, pp. 57–75, 2016.

[15] H. Wang, C. Zhang, H. Yin, W. Wang, J. Zhang, and F. Xu, “A Unified Framework for Fine-Grained Opinion Mining from Online Reviews,” *2016 49th Hawaii International Conference System Sci.*, pp. 1134–1143, 2016.

[16] B. Agarwal, S. Poria, N. Mittal, A. Gelbukh, and A. Hussain, “Concept-Level Sentiment Analysis with Dependency-Based Semantic Parsing: A Novel Approach,” *Cognitive Computation*, vol. 7, no. 4, pp. 487–499, 2015.

[17] S. Poria, E. Cambria, G. Winterstein, and G. Bin Huang, “Sentic patterns: Dependency-based rules for concept-level sentiment analysis,” *Knowledge-Based System*, vol. 69, no. 1, pp. 45–63, 2014.

[18] Q. Liu, Z. Gao, B. Liu, and Y. Zhang, “Automated rule selection for aspect extraction in opinion mining,” *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2015–Janua, no. Ijcai, pp. 1291–1297, 2015.

[19] X. Zheng, Z. Lin, X. Wang, K.-J. Lin, and M. Song, “Incorporating appraisal expression patterns into topic modeling for aspect and sentiment word identification,” *Knowledge-Based System*, vol. 61, pp. 29–47, May 2014.

[20] V. R. K. Raghuveer, “Dependency Driven Semantic Approach to Product Features Extraction and Summarization Using Customer Reviews,” in *Advances in Computing and Information Technology*, 2013, pp. 225–238.

[21] G. Qiu, B. Liu, J. Bu, and C. Chen, “Opinion Word Expansion and Target Extraction through Double Propagation,” *Computational Linguistic*, vol. 37, no. 1, pp. 9–27, 2011.

[22] A. K. Samha and Y. Li, “Aspect-Based Opinion Mining Using Dependency Relations,” *International Journal Computer Science Trends Technology*, vol. 4, no. 1, pp. 113–123, Feb. 2016.

[23] T. C. Chinsha and S. Joseph, “A syntactic approach for aspect based opinion mining,” in *Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing, IEEE ICSC 2015*, 2015, pp. 24–31.

[24] M. De Marneffe and C. D. Manning, “Stanford typed dependencies manual,” *20090110 Httpnlp Stanford*, vol. 40, no. September, pp. 1–22, 2010.