Implicit Aspect Detection in Restaurant Reviews

R. Panchendrarajan, M. N. Nazick Ahamed, B. Murugaiah, S. Prakhash,
S. Ranathunga, A. Pemasiri
Department of Computer Science and Engineering
University of Moratuwa
Katubedda 10400, Sri Lanka

Abstract

For aspect-level sentiment analysis, the important first step is to identify the aspects and their associated entities present in customer reviews. Aspects can be either explicit or implicit, where the identification of the latter is more difficult. For restaurant reviews, this difficulty is escalated due to the vast number of entities and aspects present in reviews. The problem of implicit aspect identification has been studied for customer reviews in different domains, including restaurant reviews. However, the existing work for implicit aspect identification in customer reviews has the limitation of choosing at most one implicit aspect for each sentence. Furthermore, they deal only with a limited set of aspects related to a particular domain, thus have not faced the problem of ambiguity that arises when an opinion word is used to describe different aspects. This paper presents a novel approach for implicit aspect detection, which overcomes these two limitations. Our approach yields an F1-measure of 0.842 when applied for a set of restaurant reviews collected from Yelp.

1 Introduction

Entities in a restaurant refer to products (e.g. food), services, individuals, events etc., and the aspects are the attributes or components of these entities (Zhang and Liu, 2014). For example, *smell* is an aspect of the food entity. As the customers go to restaurants for various purposes, rating values for individual aspects related to a restaurant are important as well as the overall rating value of the restaurant. For example, if someone is trying to select a restaurant to have a party, she will be interested in the rating value of parking facilities.

To determine the opinion on an aspect expressed in reviews, aspect-level sentiment analysis (or opinion mining) should be carried out. This consists of two core parts: detecting aspects and classifying sentiment score for each aspect (Schouten et al., 2015). Sentiment score is calculated using the positive or negative sentiments indicated by opinion words such as good, excellent, poor, and bad. Opinions are associated with opinion targets, which are entities on which opinions are expressed (Qiu et al., 2011). For example, in the sentence “Staff was very kind”, the *staff* entity is the target of the opinion *kind*.

Rating value for an aspect is obtained by aggregating all the sentiment scores assigned for that aspect across the reviews. However, restaurant domain deals with a vast number of aspects such as food, individual food item, drink, appetizer, furniture, staffs, different places or areas, and offers, which are interrelated. When the relationships are modeled, one entity becomes the aspect of another. For example, entities food items, drinks, desserts and appetizers come under the category of food entity as its sub-aspects and each of those aspect has sub-aspects such as taste, quality, and price. Therefore the rating value of an aspect has a significant

Proceedings of NAACL-HLT 2016, pages 128–136,
San Diego, California, June 12-17, 2016.

impact on the rating value of its parent aspect(s). This behavior of different aspects can be modeled as hierarchical relationships so that the rating value of an aspect can be calculated as a composite score of its sub-aspects.

Aspects can be explicit or implicit. When aspects are mentioned literally in a text, they are called explicit aspects, whereas implicit aspects are only implied by the sentence but not literally mentioned (Schouten et al., 2015). For example consider the sentences, “Taste of food in that restaurant is great” and “Food is delicious in that restaurant”. In the first sentence, aspect *taste* of food entity is explicitly mentioned and in the second one we can infer that it refers to the aspect *taste* of food entity, even though it is not explicitly mentioned. In the data set we considered, an average of 15.6% of the sentences contains one or more implicit aspects.

When compared with explicit aspect identification, identifying implicit aspects is much difficult. The problems escalate with the possibility to associate an aspect with multiple entities. For example, in the sentences “Pizza is very small” and “Pizza size is very small”, aspect *size* is associated with the entity food item. In the sentences “Restaurant is small” and “Restaurant size is small”, same aspect *size* is associated with the entity *Restaurant*. Similarly, the opinion word *small* that gets attached to the aspect *size* refers to size of pizza in the first two sentences, and to size of restaurant in the second two sentences, thus leading to ambiguity. This problem escalates when we deal with a large number of aspects

There can even be aspects that do not have a direct attachment to any entity. For example, in the sentences “I would recommend this restaurant” and “I will definitely be back”, the overall experience of the customer is the aspect, however this aspect does not have any directly associated entity.

Customers also have a tendency to mention multiple aspects in a single sentence. For example, consider the sentences, “Even though food is expensive it was delicious” and “Food was delicious in that small restaurant”. In the first sentence, two different aspects - *price* and *taste* of food are mentioned implicitly. The second one mentions two different aspects belonging to two different entities.

Implicit aspect identification is not a problem specific to restaurant reviews. Previous research has explored solutions for the same, for domains such as mobile phone reviews (Hai et al, 2011; Wang et al., 2013; Zhang and Zhu, 2013; Schouten et al., 2014, Schouten et al. 2015), restaurant reviews (Schouten et al., 2014, Schouten et al. 2015) and clothing reviews (Zhang and Zhu, 2013). However, none of this research is capable of identifying multiple implicit aspects appearing in a sentence. Moreover, they have dealt with only a limited number of high-level aspects, disregarding the hierarchical relationships they may have with other aspects.

This paper presents a method to detect implicit aspects mentioned in restaurant reviews. It is capable of identifying multiple implicit aspects appearing in a sentence. In our approach, each opinion word is considered as implying an implicit aspect. Using a model trained using manually tagged data, a list of aspects that can be implied by each of these opinions are identified. These aspects are given a score using the co-occurrence between opinion word and other words in the sentence. This is achieved by extending the work of Schouten et al. (2014) that identifies at most one implicit aspect in a given sentence. Aspect with the highest score is chosen as the potential candidate aspect. Opinion targets and opinions are extracted and checked whether they have any relationship with (parent or sibling) the predicted aspect.

In order to identify the relationships, different entities with different aspects are modeled as a hierarchy. Such a comprehensive model cannot be found in related literature for the restaurant domain. This hierarchy enables to verify whether the predicted implicit aspect is correct or not by utilizing two different relationships between aspects: aspect-parent and aspect-sibling. Such a verification technique cannot be found in the existing literature.

The rest of the paper is organized as follows. Section 2 discusses previous research related to implicit aspect identification, and Section 3 discusses our approach for the same. Section 4 presents the model we developed to capture the hierarchical relationships between aspects in the restaurant domain. Section 5 evaluates our system, and section 6 concludes the paper.

2 Related Work

As mentioned earlier, implicit aspect identification has been explored in the context of customer reviews for different domains. To identify implicit aspects, most of the previous research uses association between aspects and opinion words occurring in a sentence, where the aspect is implied by the opinion word.

Hu and Liu (2005) have used association rules and generate patterns to identify both explicit and implicit aspects. Popescu and Etzioni (2005) suggest an approach using Point wise Mutual Information (PMI) based semantic association analysis. No quantitative experimental results have been reported in this work. Su et al. (2008) propose a similar approach that clusters aspects as well as the opinions to generate association rules that map a set of aspects to opinions. However, this approach considers only adjectives as opinion words and no quantitative results are given.

Hai et al. (2011) propose a two-phase approach based on co-occurrence association rule mining (coAR) and the experiment was carried out for a set of Chinese mobile phone reviews. Main deficiency in their work is, they use only the co-occurrence of opinion words to identify an implicit aspect. A hybrid association rule mining technique proposed by Wang et al. (2013) overcomes this issue, by extracting indicators for aspects. They experimented on a Chinese data set of mobile phone reviews. These indicators are both opinion words and other words. Zhang and Zhu (2013) overcome the same issue by making use of the associations between an aspect and the rest of the notional words in the clause. A corpus of mobile phone reviews in Chinese and a collection of clothes reviews in Chinese were used in the research.

All these approaches have a drawback of identifying an implicit aspect only if it is available explicitly in the training data set. Schouten et al. (2014) overcome this issue by utilizing the co-occurrence between notional words and either the explicit or implicit aspect so that an implicit aspect can be identified even if it does not appear explicitly. The authors later extended this work by employing word sense disambiguation and utilizing the semantic relations between words (Schouten et al., 2015). Restaurant reviews and product reviews were used for the experiment purposes in both studies. However, both these works are only capable of choosing at most one implicit aspect for each sentence. Furthermore they identify implicit aspects only of five categories, food, service, ambience, price, and anecdotes/miscellaneous. They do not consider different types of entities with different aspects, and their relationships.

In summary, none of this previous work is capable of identifying multiple implicit aspects occurring in a sentence. Moreover, they deal with only a limited number of aspects of the respective domain, and ignore the relationships between aspects at different levels. Thus they have not faced the problem of ambiguity when predicting an implicit aspect, in cases where the same opinion can be associated with different aspects.

3 Implicit Aspect Identification

This section presents our implicit aspect identification method that overcomes the following limitations discussed in section 2.

- Finding implicit aspects only of a set of limited categories in restaurant domain, thus ignoring the ambiguity in attaching opinions to aspects
- Finding only one implicit aspect in a sentence

Two models are created to identify implicit and explicit aspects separately. Both use a training data set with explicit and implicit aspects manually labeled. First model uses maximum entropy classification technique to identify explicit aspects. Second model identifies opinion words in a sentence and predicts the implicit aspect implied by that opinion word using the co-occurrence of words. Accuracy of the identified implicit aspect prediction is checked using a set of rules and the hierarchical relationships among aspects.

3.1 Training Data Set

In the data set, aspects (both explicit and implicit) and entities are manually labeled. For example, in the sentence "Pizza was small in that big restaurant", *pizza* and *restaurant* are identified as entities or explicit aspects and are labeled as *Food item* and *Restaurant*, respectively. *small* and *big* are opinion words that identify implicit aspects. Therefore the opinion words are labeled with the implicit aspect they indicate. For example, in the above sentence *small* and *big* are labeled as *Food_item_size* and *Environment_size*, respectively.

3.2 Training Phase

In the training phase, two models are separately trained to identify explicit aspects and implicit aspects. For explicit aspect identification, a standard maximum entropy classifier (Opennlp.apache.org, 2016) is used to create the model M1 using our annotated corpus. N-grams are used as features where n varies from 2 to 5.

In order to train the next model M2, training data set is scanned and a list of opinion words O is created by identifying the opinion words labeled as implicit aspects. In the second iteration of scanning, only the sentences with implicit aspects are extracted. Words labeled as explicit aspects in those extracted sentences are replaced with their explicit aspect label or entity label. For example, the sentence "Pizza was small in that big restaurant" is modified as "Food_item was small in that big restaurant". Modified sentences are stored under each identified opinion word along with their label. For example, the modified sentence "Food item was small in that big Restaurant" is stored under both opinion words *small* and *big* with the candidate aspect labels *Food_item_size* and *Environment_size*, respectively. All the possible aspects that can be implied by an opinion word are now available in the model as aspect-sentence pairs. For example, consider another sentence "Restaurant is not suitable for parties as it is very small". Here "Restaurant" and "it" are replaced by the explicit aspect tag *Restaurant*. This sentence is stored in the model under the opinion word *small* along with the candidate aspect label *Environment_size*. Finally model appears as follows for these two sentences:

small

Food item_size – Food_item was small in that big Restaurant

Environment_size- Restaurant is not suitable for parties as Restaurant is very small

3.3 Testing Phase

When a new restaurant review is given, explicit aspects and entities are identified using the trained model M1. Same as the training phase, words identified as entities or explicit aspects in the test data are replaced with their predicted explicit aspect or entity label. Modified test data are processed word by word within a sentence for opinion words available in the list O. For each identified opinion word in a sentence, the list of candidate aspects A is extracted using the model M2. With the list of candidate aspects, identifying the winning implicit aspect is a two-step process as described below.

3.3.1 Step 1

As the first step, one implicit aspect from the list of candidate aspects is chosen as the potential candidate aspect using the co-occurrence between the opinion word and other words in the sentence. If there is only one candidate aspect, it is chosen as the potential candidate. Otherwise, for each candidate aspect Ai, a score is computed using equation (1). This equation is a modified version of the equation used by Schouten et al. (2014). The limitation of Schouten et al.'s equation is, it does not consider the distance between an opinion word and other words in the sentence while calculating the co-occurrence of words to obtain the score. In our modified equation, we add a weight when calculating the sum of co-occurrence frequency of words. Distance between the opinion word and other words in the sentence are used as the weight, thus removing the impact of faraway words on the sum of co-occurrence. Co-occurrence frequency between opinion word and other words in the sentence is calculated using the sentences attached to the opinion words in model M2 for a particular candidate aspect.

$$\text{Score } A_i = 1/n \sum C_{ij}/f_j * 1/d_j \quad (1)$$

In equation (1), n is the number of words in the given sentence, A_i is the i^{th} candidate aspect in A for which the score is computed, j represents j^{th} word in the sentence, C_{ij} is the co-occurrence frequency of aspect A_i and i^{th} word, f_i is the frequency of the i^{th} word and d_j is the distance between the j^{th} word and the opinion word. $1/d_j$ operates as weight.

Co-occurrence of stop words is not considered to get the sum. Highest scoring aspect that exceeds the threshold becomes the potential aspect for the next step. If the highest score is lower than the threshold, identified opinion word is discarded. Optimal threshold is identified based on the training data using a simple linear search. Threshold is

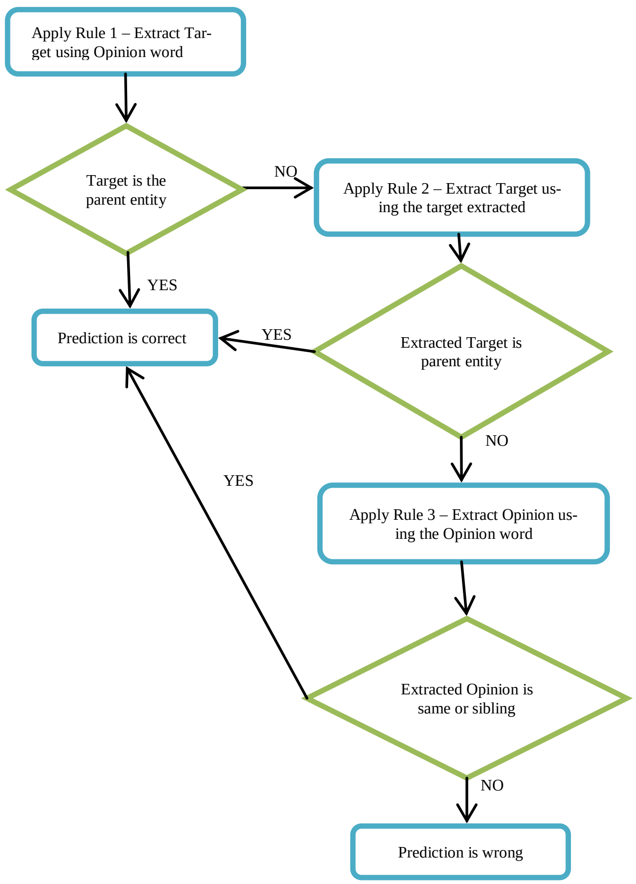

Figure 1:Flow of Step 2

increased from 0 by a step size of 0.01 until the optimum value for F1-measure is obtained.

However, evaluations (Table 1 – row 3) showed that step 1 is not sufficient to identify and discard wrong predictions for the aspect implied by an opinion word. For example, consider the sentence, “We came as a small group for the dinner”. Here, opinion word *small* describes the size of the group. However while processing this sentence, *small* is identified as an opinion word available in O. Suppose either *Food_item_size* or *Environment_size* is chosen as the winning candidate entity. If the process stops at that point, *small* will be identified as either *Food_item_size* or *Environment_size*.

3.3.2 Step 2

In this step we validate the predicted implicit aspect implied by an opinion word. Step 2 works as the flow shown in Figure 1. Once the potential candidate aspect is chosen, next step is to extract its opinion target to check whether the prediction is correct or not. If the opinion target is the parent of the potential candidate, it is chosen as the winning candidate. Otherwise the potential candidate aspect is discarded. Opinion targets are extracted using the double propagation approach proposed by Qui et al (2011), which propagates information back and forth between opinion words and targets using grammar rules.

These grammar rules are based on the dependency relations between words. The dependency relations describe relations between opinion words and targets. We used the dependency relations mod, pnmod, subj, s, obj, obj2 and desc as defined by Qiu et al. (2011). Following are the rules we used:

Rule 1 - Using the given opinion word, target is extracted using grammar rules. Example: In the sentence “staff were very kind”, staff is identified as the target using the rule *kind* -> *mod* -> *Staff.*

Rule 2–Target extracted using Rule 1 is used to extract further targets in the sentence using grammar rules. Example: In the sentence “food and desserts are tasty in that restaurant”, when the opinion word *tasty* is processed, dessert is identified as its target in the previous step. However, it is not the parent of *Food_item_Taste*. Therefore the flow moves to Rule 2 and *Food* is identified as target using the rule *Dessert* ->*conj* ->*Food*. Since *Food* is the grand parent of *Food_item_Taste* it is chosen as the winning candidate aspect.

Rule 3 - Using the identified sibling or the same type of implicit aspects, opinion words are identified using grammar rules. Example: In the sentence “Staff was kind and available”, *kind* is identified as *Staff_behaviour*. When *available* is processed, *kind* is identified as an opinion word using the rule *kind* ->*conj* -> *available*. Since both *Staff_behaviour* and *Staff_availability* are siblings in the hierarchy of aspects, *Staff_availability* is chosen as the winning aspect for the opinion word available.

Rules are applied one after other and checked whether the prediction is correct or not. If the prediction fails in all three rules aspect is discarded. In order to record these hierarchical relationships between aspects, we developed a comprehensive model for the restaurant domain, as described in Section 4.

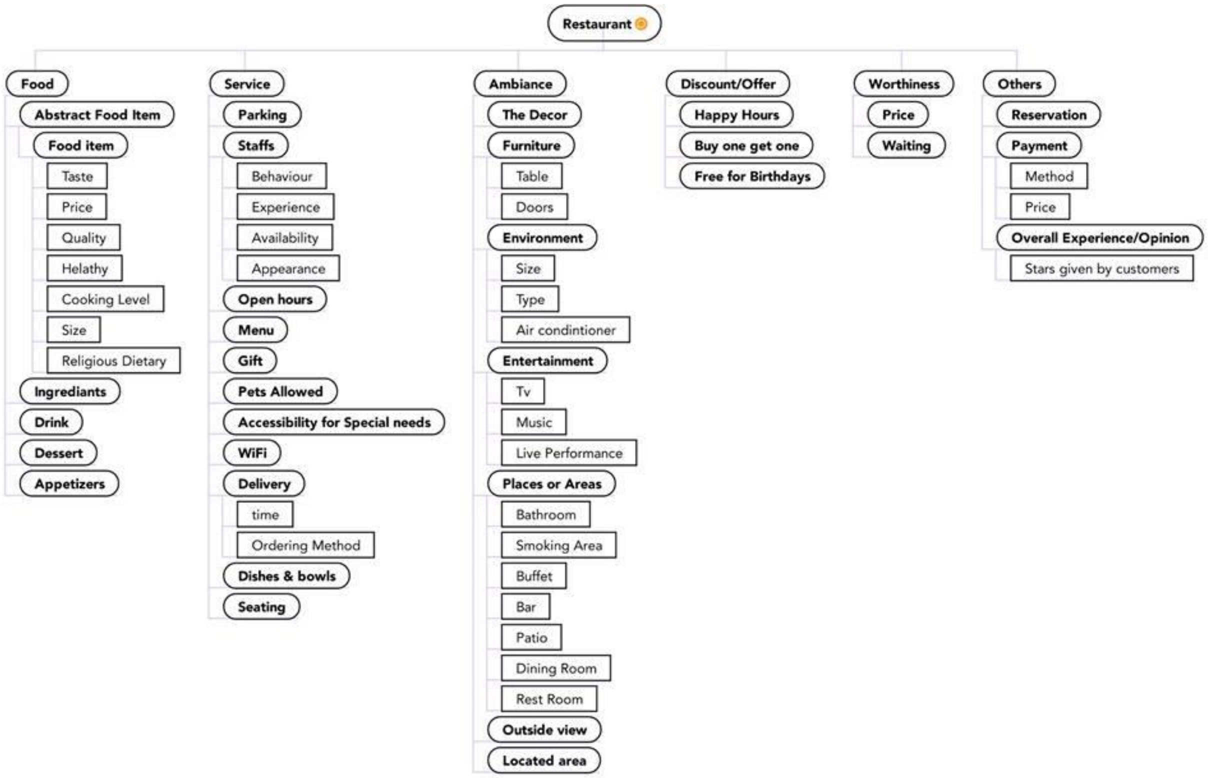

Figure 2: Hierarchy of Aspects

We also tried out two modifications for the above approach that we implemented: (1) implement Step 1 only where the chosen potential candidate aspect is considered as the winning aspect (i.e. do not execute step 2). (2) Consider the occurrence of stop words while calculating the weighted sum of co-occurrence.

4 Modeling the Hierarchy of Aspects

Restaurant industry deals with a vast number of entities such as food items, drinks, furniture, staff, offers, etc. which are interrelated to each other.

Figure 2 shows the model we developed to represent the hierarchical relationships between different entities and aspects. This model was manually developed using a random sample of 400 reviews and was validated and refined using another set of 400 reviews collected from Yelp (2016).

Model consists of aspects up to four levels. Level 1 one is restaurant and it has six sub main sub aspects as food, service, ambience, offers, worthiness and other aspects. Each sub aspect is further categorized. For example, aspect *Service* has *Staff* as its one of the sub aspects which has four sub aspects, *Behavior, Experience, Appearance and Availability*.

As discussed in section 3, identifying the relationship between various aspects enable to check whether the predicted implicit aspect is correct or not. For example consider the two sentences, “Food item was very expensive” and “Food was really delicious”. In both sentences, aspects *Food_item_Price* and *Food_item_Taste* will be identified as implicit aspects, respectively. In order to check whether the prediction is correct or not, the opinion targets are extracted. In the above example, the opinion targets of *expensive* and *delicious* are *Food_item* and *Food* respectively. Since those are the parent and grandparent of the aspects *Food_item_Price* and *Food_item_Taste* respectively, *Food_item_Price* is chosen as the winning aspectfor the opinion word *expensive,* and *Food_item_Taste*is chosen as the winning aspect for the opinion word*delicious*.

Now consider the earlier discussed example, “I am a big fan of that restaurant”. Here, “I” is identified as the opinion target of the opinion word *big* with the prediction of either *Food item_Size* or *Environment_Size*as higher scoring one will be cho-

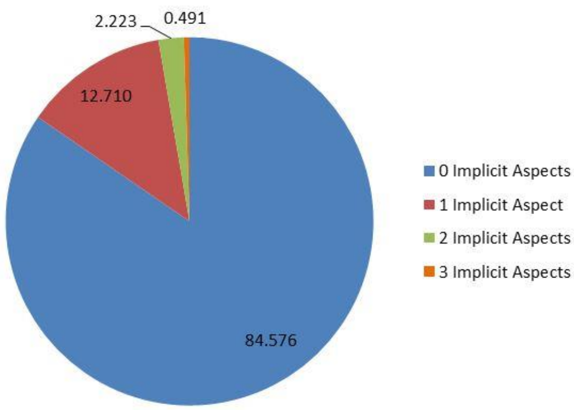

Figure 3: Average distribution of sentences in the restaurant review data set, according to the number of implicit features they contain in 1000 reviews.

sen as the potential candidate aspect. However, this prediction will be discarded as the opinion target is not even an entity in the model.

5 Data Set and Initial Analysis

1000 restaurant reviews collected from Yelp (2016) are used as the training data set. Both explicit and implicit aspects were labeled manually in this data set. Even though the restaurant domain deals with a vast amount of entities with various aspects, not all the sentences in restaurant reviews contain implicit aspects. As shown in Figure 3, 15.6% of the sentences contain one or more implicit aspects in 1000 restaurant reviews.

However it is essential to identify that small fraction of sentences and all the aspects mentioned implicitly in those sentences since important aspects are most likely to be used in the sentence implicitly. For example, more than 92% of aspects of staff entity appear implicitly in restaurant reviews, as it can be seen in Figure 4.

Figure 5 shows the distribution of Level 2 aspects (food, service, ambience, offers, worthiness and others) by considering all the aspects as sub aspects of level 2 aspects. Distribution of an aspect was obtained by calculating the frequency of occurrence of an aspect and its sub aspects in the training data set.

6 Evaluation

Evaluations are performed using 10-fold-cross validation with a training data set of 1000 reviews.

Figure 4: Percentage of aspects of Staff entity appearing explicitly or implicitly in 1000 reviews

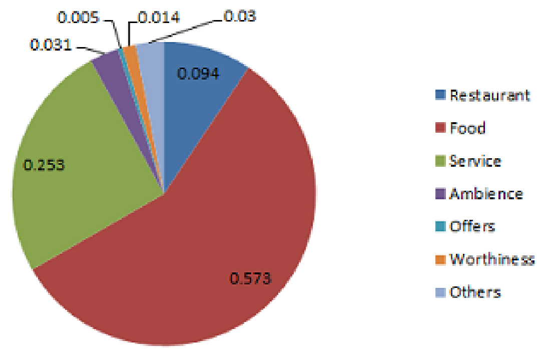

Figure 5: Distribution of level 2 aspects and restaurant in the training data set

For each instance of the algorithm, 900 reviews are used as the training set and the remaining 100 reviews are used for testing.

Table 1 shows the evaluation results of 10-fold-cross validation for the several methods. Methods 1 to 4 assume that the model M1 can identify explicit aspects and entities with an accuracy of 100%. Method 5 shows the results for our approach by using the trained model M1.

The accuracy of the M1 model was tested using an additional tagged set of 400 reviews. Model M1 identified explicit aspects with an F1-Measure of 0.88 (Precision – 0.931, Recall, 0.835).

Method	**Precision**	**Recall**	**F1-Measure**
1. Initially given solution	**0.947**	**0.758**	**0.842**
2.Using the approach suggested by Schouten et al. [6]	0.495	0.929	0.645
3. Modification 1	0.916	0.752	0.826
4. Modification 2	0.931	0.754	0.834

5. Initially given solution with the trained model M1	0.886	0.694	0.779

Table 1: Evaluation Results

It can be seen in Table 1 that our approach gives the best result. Moreover it is worth noting that the precision drops drastically from 0.947 to 0.529 in Modification 2 as it does not execute Step 2.

Moreover the approach suggested by Schouten et al. (2014) fails in the case of identifying large number of inter-related implicit aspects. Therefore adding step 2 to Schouten's work (Modification 1) improves precision from 0.49 to 0.91. The result for our approach is slightly higher than this, as it considers the distance between opinion words and other words in the sentence. Moreover, the occurrence of stop words does not have any impact as the F1-Measure obtained using Modification 2 is very close to the same obtained using proposed solution.

Table 2 shows the evaluation results for 10-fold-cross validation of our solution for sentences with more than one aspect and it can be observed that the F1-Measure is above 0.82.

sentence type	Precision	Recall	F1-Measure
1. Sentences with two implicit aspects	0.978	0.709	0.822
2. Sentences with more than two implicit aspects	0.975	0.725	0.832

Table 2: Evaluation results for sentences with multiple implicit aspects

In order to measure the inter-rater-reliability (IRR) of aspect annotation, three data sets, each with 100 reviews were picked. Each set was tagged by two different annotators. Two types of measure of consistency were computed; absolute agreement, and Kappa coefficient. The absolute agreement was calculated by dividing the total number of times all annotators agreed on a tag over the total number of tags. Kappa coefficient (Carletta, 1996) is calculated as follows,

$$\text{Kappa Coefficient} = P(A) - P(E)/(1 - P(E)) \qquad (2)$$

where P(A) is the proportion of times the annotators actually agree and P(E) is the proportion of times the annotators are expected to agree due to chance.

File	Absolute Agreement	Cohen's Kappa
1. Test data set 1	0.93	0.861
2. Test data set 2	0.928	0.855
3. Test data set 3	0.893	0.785
Average	**0.917**	**0.834**

Table 3: Annotator Agreement Test Results

An acceptable agreement for Cohen's Kappa value for most NLP classification tasks lies between 0.7 and 0.8 (Carletta, 1996). Table 3 shows the results for IRR test and it can be seen that average Kappa coefficient value for the test data sets is 0.83. Therefore the training data set with aspects labeled is acceptable.

7 Conclusion

This paper presented an approach to identify multiple implicit aspects in a sentence. Co-occurrence between opinion word and other words in the sentence is used to identify an aspect that is implied in an opinion word. Double propagation technique is used to extract opinion target to check whether the identified aspect is correct or not. Relationships between different entities with different aspects are modeled as a hierarchy, which helps in improving the accuracy of implicit aspect identification in the presence of a large number of inter-related aspects.

As future work, it would be interesting to extend this work dynamically to improve the model, as new entities and aspects are found. Furthermore, this work can be extended to other domains as well by identifying relationships between aspects specific to a domain and modeling them as a hierarchy.

References

Carletta, J. Assessing Agreement on Classification Tasks: The Kappa Statistic. Computational Linguistics Volume 22 Issue 2, pages 248-254, 1996.

Hai, Z., Chang, K. and Kim, J. Implicit Feature Identification via Co-occurrence Association Rule Mining. In 12th International Conference on Computational Linguistics and Intelligent Text processing, volume 6608, pages 393–404. Springer, 2011.

Hu, M., Liu, B. and Cheng, J. Opinion Observer: Analyzing and Comparing Opinions on the Web. In Proceedings of the 14th International Conference on World Wide Web, pages 342–351. ACM, 2005.

Opennlp.apache.org,. "Apache Opennlp - Welcome To Apache Opennlp". N.p., 2016. Web. 7 Jan. 2016.

Popescu, A. and Etzioni, O. OPINE: Extracting Product Features and Opinions from Reviews. In Proceedings of HLT/EMNLP (Demonstration Abstracts), pages 32–33. Vancouver, 2005.

Qiu, G., Liu, B., Bu, J. and Chen, C. Opinion Word Expansion and Target Extraction through Double Propagation. Computational Linguistics, volume 37, pages 9–27, 2011.

Schouten, K. and Frasincar, F. Finding Implicit Features in Consumer Reviews for Sentiment Analysis. In 14th International Conference on Web Engineering, volume 8541, pages 130–144. Springer, 2014.

Schouten et al., Semantics-Driven Implicit Aspect Detection in Consumer Reviews. In Proceedings of the 24th International Conference on World Wide Web, pages 109-110. ACM, 2015.

Su, Q., Xu, X., Guo, H., Guo, Z., Wu, X., Zhang, X., Swen, B. and Su, Z. Hidden Sentiment Association in Chinese Web Opinion Mining. In Proceedings of the 17th International Conference on World Wide Web, pages 959–968. ACM, 2008.

Wang, W., Xu., H. and Wan, H. Implicit Feature Identification via Hybrid Association Rule Mining. Expert Systems with Applications: An International Journal Volume 40 Issue 9, pages 3518-3531, 2013.

Yelp,. "Yelp". N.p., 2016. Web. 7 Jan. 2016.

Zhang,L.and Liu, B.Aspect and Entity Extraction for Opinion Mining, Data Mining and Knowledge Discovery for Big Data, Volume 1, pages 1-49. Springer, 2014.

Zhang, Y. and Zhu, W. Extracting Implicit Features in Online Customer Reviews for Opinion Mining. In Proceedings of the 22nd International Conference on World Wide Web, pages 103–104. ACM, 2013.