A New Approach for A Domain-Independent Turkish Sentiment Seed Lexicon Compilation

Ekin Ekinci and Sevinç Omurca
Department of Computer Engineering, Kocaeli University, Turkey

Abstract: *Sentiment analysis deals with opinions in documents and relies on sentiment lexicons; however, Turkish is one of the poorest languages in regard to having such ready-to-use sentiment lexicons. In this article, we propose a domain-independent Turkish sentiment seed lexicon, which is extended from an initial seed lexicon, consisting of 62 positive/negative seeds. The lexicon is completed by using the beam search method to propagate the sentiment values of initial seeds by exploiting synonym and antonym relations in the Turkish Semantic Relations Dataset. Consequently, the proposed method assigned 94 words as positive sentiments and 95 words as negative sentiments. To test the correctness of the sentiment seeds and their values the first sense, the total sum and weighted sum algorithms, which are based on SentiWordNet and SenticNet 3, are used. According to the weighted sum, experimental results indicate that the beam search algorithm is a good alternative to automatic construction of a domain-independent sentiment seed lexicon.*

Keywords: *Sentiment lexicon, beam search, pattern generation, turkish language, unsupervised framework.*

Received March 29, 2016; accepted September 22, 2016

1. Introduction

Sentiment analysis is the process of identifying the opinion (e.g., negative or positive) of a given document. While people are making decisions about their choices they are mostly affected by other people's opinions. The Internet, which has an important place in our daily lives, has become an indispensable medium for data analysis as a result of the constant and unavoidable increase in stored data through new resources that have millions of users such as blogs, dictionaries, news portals, ecommerce and social media websites. Unlimited streaming of information and review on the Internet enables the analysis of user reviews from many different domains. Due to the increase of review pages, blogs, dictionaries, etc. sentiment analysis has become an important topic in recent years and sentiment lexicons are essential for a research in this field.

A sentiment lexicon is crucial for sentiment analysis as it provides initial sentiment information about the documents and consists of words and phrases which are assigned a positive or negative score reflecting its sentiment polarity [28]. A sentiment learning algorithm initially starts with sentiment seeds such as good or bad which have domain-independent positive or negative scores. The algorithm then propagates these sentiment seeds to estimate the sentiment score of each new sentiment word. These kinds of algorithms use a thesaurus such as WordNet [22] to find new sentiment words and calculate their sentiment scores due to the similarity between words. However parsing all synonyms from a thesaurus to give them sentiment scores is not a suitable way to construct an efficient sentiment lexicon.

Constructing a general lexicon, which can be used efficiently in every domain, is a difficult and time consuming task because the sentiment expressions of the words changes from one domain to another. A prime example of this is the word short, which in an electronics review would take a negative connotation, as in the short battery life of the product, however, when it appears in restaurant reviews it takes on a positive meaning, as in the short service duration. Due to this problem of different domains some words have both positive and negative scores, so if the approach for generating a sentiment seed lexicon can distinguish the ambiguous words; then these kinds of words can be filtered and not included in the lexicon. In result, a more accurate sentiment seed lexicon can be generated for sentiment classification.

In sentiment classification applications, either a general purpose sentiment lexicon is adapted to a specific domain using some domain-specific data or a sentiment lexicon is constructed in a given domain starting from a seed word set. General-purpose sentiment lexicons such as SentiWordNet [3] cannot capture sentiment variations across different domains; they only provide fast and scalable approach to sentiment analysis [16]. On the other hand, constructing a lexicon in a specific domain starting from a very small seed word set is time-consuming. Pang *et al*. [37] showed that it is difficult to get good coverage of a target domain from manually selected words.

Apart from these approaches, in this study, the seed set which is the largest as possible and contains the words has unambiguous positive or negative orientation is assumed as the best initial lexicon for sentiment analysis. In Turkish there is a gap in this field which needs to be focused on.

With regard to this consideration, we propose a Turkish sentiment seed lexicon compilation method which is initially made up of manually created seeds and their associated sentiments, and then we expand it with a beam search algorithm. A beam search is a heuristic search algorithm in artificial intelligence which explores the state-space graph by expanding a specific number (denoted by beam width and filter width) of the most promising nodes at each level of the graph. The beam width and filter width are user defined parameters of the algorithm [46].

We evaluated the proposed seed lexicon using four different evaluation methods, the first sense, total sum and weighted sum based on SentiWordNet and SenticNet 3 algorithms as explained in "experimental results section". Experimental results show that the proposed method results reliable sentiment seed words.

The remainder of the article is organized as follows: Section 2 presents some related work; section 3 describes the proposed approach for generating the Turkish sentiment seed lexicon; section 4 presents and discusses the empirical study; finally we outline conclusions and discuss the future work in section 5.

2. Related Work

So far, numerous sentiment lexicons of varying sizes have been constructed to enable sentiment classification in documents. SentiWordNet [3], SentiSense [15], WordNet-Affect [42] are some of these based on the English lexical databases such as WordNet [22], SenticNet [10], Opinion Lexicon [23], The Semantic Orientation CALculator (SO-CAL) [43], AFINN [34], Subjectivity Lexicon [39]. Generally, the similarity of the words is used to construct the lexicon. Due to obtaining the similarity measures, lexicon construction methods are classified into two. The first one is thesaurus based and the second one is a corpus based approach [44]. In this study, the thesaurus based approach is used.

For the thesaurus based approach, some interesting studies can be summarized. General Inquirer [41], which is a hand-made lexicon constituted by lemmas, can be thought of as the first sentiment lexicon [27]. Started with a manually constructed small set of seed words, and then sorted them by polarity into positive and negative lists. Finally these lists were grown by adding words obtained from WordNet.

Kamps *et al*. [25] developed a distance measure on WordNet, and showed how it could be used to determine the semantic orientation of adjectives. They constructed a synonyms graph by using words provided from WordNet and determined the word polarity by distance from words in the graph. Esuli and Sebastiani [20] presented a method for determining the orientation of subjective terms based on semi-supervised learning and applied it to term representations which were obtained by using term glosses from an online dictionary. Vossen *et al.* [47] presented a semantic database cornetto that combines Duch WordNet, which is similar to the Princeton WordNet for English, with Referentie Bestand Nederlands (RBN) which includes frame-like information. Ding *et al.* [18] obtained an opinion lexicon or the set of opinion words through a bootstrapping process using WordNet. Baccianella *et al*. [3] constructed SentiWordNet by automatically annotating all WordNet synsets according to their degrees of positivity, negativity, and neutrality. Mahyoub *et al*. [31] presented an Arabic sentiment lexicon that assigns sentiment scores to the words found in the Arabic WordNet. They started from a small seed list of positive and negative words then used semi-supervised learning to propagate the scores in the Arabic WordNet by exploiting the synset relations.

As it is mentioned in Cruz *et al*. [13], according to the number of cites, the two most used lexicons nowadays are Bing Liu's Opinion Lexicon [23] and SentiWordNet [3, 21]. When the literature is examined it can be observed that there are not many sentiment lexicons for languages other than English. Some of these are: Hindu and French [38], Arabian [1], German [12], Japanese [24], Chinese [30, 49], Romanian [4], Indian [14], Spanish [8] and Punjabi [26] sentiment lexicons.

While there are many studies in semantic analysis for English, there are limited studies for Turkish. There has been one previous effort for developing a general purpose sentiment lexicon [16], and there have been a few studies for sentiment analysis. However; we haven't seen any study about establishing a Turkish sentiment seed lexicon that contains domain independent seeds. Dehkharghani *et al*. [16] presented the first comprehensive Turkish polarity lexicon, SentiTurkNet and they assigned three polarity scores to each synset in the Turkish WordNet, indicating its positivity, negativity, and objectivity levels. Furthermore, some related studies are summarized as in the following. Cakmak *et al*. [9] tried to build fuzzy models for Turkish emotion words and investigated 197 emotion words by using fuzzy logic techniques. Vural *et al*. [48] presented a framework for unsupervised sentiment analysis in Turkish documents. They translated the lexicon of the SentiStrength sentiment analysis library to Turkish and implemented their framework to the classification problem of movie reviews. Özsert and Özcan [36] indicated that WordNets are easily connected to each other by linking the words in one WordNet to their similar meanings in the other WordNets. They used General Inquirer as a source for English seed words and used Turkish as a foreign language which does not have a resource such as General Inquirer. They proposed a semi-automated method to produce foreign seed words and generated 1,398 positive and 1,414 negative seed words for Turkish.

3. Design of the Lexicon Model

In this section, the brief overviews of relevant concepts such as beam search and the proposed system are provided.

3.1. Beam Search

Because of its simplicity and effectiveness, the beam search method which is a heuristic tree search method is applied to many problems in several domains. This method is preferred especially for very large state space graphs [46] and has been applied to scheduling, sequencing and combinatorial optimization problems [2, 5, 7, 29, 33, 40, 50]. Like Breadth-First search, beam search progresses level by level with no backtracking but unlike Breadth-First search it moves downward from only the selected nodes which are the most promising nodes according to the evaluation functions at each level. Using only the most promising nodes for branching the next level, beam search is admitted as an adaptation of Branch and Bound algorithm that uses the incomplete derivates so it is an approximate method [45]. In this tree search method the number of nodes to be selected is determined by the user, thus the user can control time and space complexity of the problem [6]. Beam search uses a combination of two evaluation functions for progressing. The first one is called local evaluation function which is fast but can discard good solutions and the second one is called global evaluation function which is accurate but computationally more expensive.

The beam search tree shows how the partial solutions are constructed and at each level child nodes, which have the same parent, compete with each other based on the local evaluation function firstly. In local evaluation function child nodes are filtered according to the filter width (α), thus partial solutions, which are called as beams, are obtained. Then these partial solutions are subjected to the global evaluation function. According to the beam width (β), child nodes, which are the same level and form the partial solution, are filtered in the global evaluation. The success of the beam search depends on the local and global evaluation functions, α and β [35, 40]. With filtering, the aim is to reduce computational burden of the searching algorithm. During the algorithm, searching with filtering can provide performance improvement so this increases effectiveness. With the larger value of α, β good solutions can be found but can cause high computational time [6]. The Figure 1 illustrates the each step of the beam search.

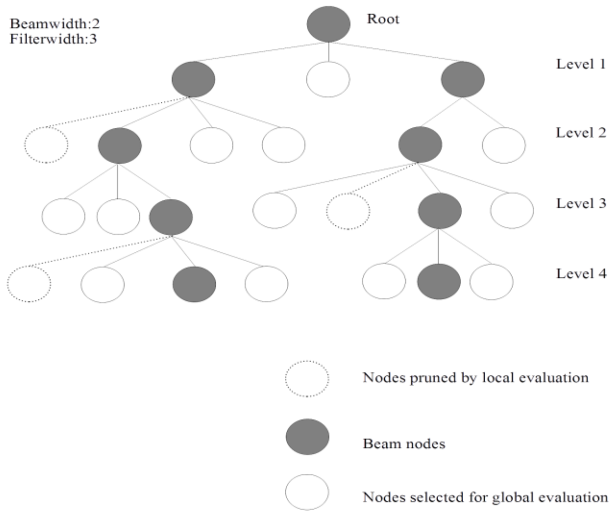

Figure 1. Beam search tree representation [40].

3.2. The Proposed Model

Most sentiment lexicon compilation approaches have two main steps, these being sentiment seed collection and sentiment value propagation. In the first step, seeds with accurate sentiment values are collected manually from existing dictionaries. In this study, 62 sentiment seeds, which have precise sentiment values because of domain independence, have been determined and collected from a Turkish lexicon. In the second step, an existing word/phrase/concept graph is used as the foundation. Sentiment values are propagated from seeds to the remaining parts of the foundation graph. To this end, ConceptNet-based Turkish sentiment lexicon which is open access for academic use called the Turkish Semantic Relations Dataset, which is obtained from a natural language research group of Yıldız Technical University[1], is used to improve value propagation with relation selection. In this lexicon there are 85,674 words and the lexicon contains at least one relation for every word with another, so some words are presented more than one time in it. When this lexicon is converted to the state space graph based on relations, the graph would be too large.

Propagation with SYNONYM and ANTONYM relations by using the graph without any optimization criteria increases significantly the computational burden and reduces effectiveness. Therefore, in this study it is convenient to use beam search, which is preferred for very large state space graphs and can provide optimization [46].

As mentioned in the previous section, the success of the beam search depends on these four factors: the local and global evaluation function, filter width (α) and beam width (β) [40]. For the purposed beam search, firstly, local and global evaluation functions are determined based on a heuristic rule which has

[1]http://www.kemik.yildiz.edu.tr/

been created taking into account Turkish language structures. The heuristic rule says that if a word has more SYNONYM and ANTONYM relations than other words this word is more important. This rule is used as both a local and a global evaluation function for the purpose of propagation. The proposed algorithm is reflected in Algorithm 1.

4. Empirical Study

The number of new seeds obtained from every depth of the solution tree due to local and global evaluation function has been determined with α and β. After determining these functions, selection of the optimal α and β values is very important for the performance of the functions. While compiling the sentiment lexicon, as a result of the experiments, it has been observed that the optimal values for both α and β are 4. If the values of α and β are smaller, then the important seed words are discarded, if these values are larger, the new seed is not added but only increase the repeat number of already existing ones. Furthermore determining the depth of the tree is very important, so to do this, experiments are conducted then a decision is given by the obtained seeds which are examined based on the selected depths.

Algorithm 1:Proposed beam search algorithm.

```
procedure BEAMSERACH(SeedDictionary)
  for all SentimentSeed s_i ϵ SeedDictionary do
    CandidateSeedList←FindCandidate(s_i)
    SeedList←GloballyEvaluate(CandidateSeedList)
    depth←2
    while depth ≠4 do
      for all SentimentSeed ss_j ϵ SeedList do
        CandidateSeedList←FindCandidate(ss_j)
        LocalList←LocallyEvaluate(CandidateSeedList)
      end for
      SeedList←GloballyEvaluate(LocalList)
      depth←depth+1
    end while
  end for
end procedure
procedure FINDCANDIDATE(SentimentSeed s_i)
  for all CandidateSeed c_j ϵ TurkishDictionary (ConceptNetbased) do
    if c_j.equals(s_i.SYNONYM) then
      CandidateSeedList.add(c_j)
      Orientation(c_j) ← Orientation(s_i)
    else if c_j.equals(s_i.ANTONYM) then
      CandidateSeedList.add(c_j)
      Orientation(c_j) ← Orientation(s_i)'
    end if
    return CandidateSeedList
  end for
end procedure

procedure GLOBALLYEVALUATE(LocalList, beamWidth)
  if LocalList.size<beamWidth then
    SeedList←LocalList
  else
    for all CandidateSeed s_i ϵLocalList do
      count(s_i) ← count(s_i) + 1
      for all Word w_j ϵTurkishDictionary do
        if w_j.equals(s_i.SYNONYM) then
          count(s_i) ← count(s_i) + 1
        else if w_j.equals(s_i.ANTONYM) then
          count(s_i) ← count(s_i) + 1
        end if
      end for
    end for
  end if
  TemporaryList←sort all s_i from maximum to minimum based on the count(s_i)
  SeedList←first four element of TemporaryList
  return SeedList
end procedure
procedure LOCALLYEVALUATE(CandidateSeedList, filterWidth)
  if LocalList.size<filterWidth then
    SeedList←LocalList
  else
    for all CandidateSeed s_i ϵLocalList do
      count(s_i) ←0
      for all Word w_j ϵTurkishDictionary do
        if w_j.equals(s_i.SYNONYM) then
          count(s_i) ← count(s_i) + 1
        else if w_j.equals(s_i.ANTONYM) then
          count(s_i) ← count(s_i) + 1
        end if
      end for
    end for
  end if
  TemporaryList←sort all s_i from maximum to minimum based on the count(s_i)
  SeedList←first four element of TemporaryList
  return SeedList
end procedure
```

When the depth is increased it can be clearly seen that there is no change in the variety of adjectives; only the same adjectives repeat multiple times and causes unnecessary time complexity. So in this context, tree depth is determined to be 4.

For each of the 62 seeds, in every depth, based on the heuristic rule first locally then globally maximum four seeds are determined. These steps are performed recursively until the fourth level of the tree is reached and at the end 494 new seeds are obtained. The proposed method assigns 277 words as positive and 217 as negative sentiments. These new 494 seeds are examined and some operations are performed on them. These operations are described as in below:

- If there are same pairs of sentiment seeds and sentiment values in the sentiment lexicon, only one of these pairs will be kept, and others will be deleted.
- If there is a sentiment seed with different sentiment values (<seed1, positive>, <seed1, negative>) in the sentiment lexicon, all pairs like these will be deleted.

As a result of these steps, the number of seeds has fallen from 494 to 189 (94 positive and 95 negative). For these 189 seeds, there is not a baseline method for Turkish natural language processing to check whether their sentiment value is true or not. Therefore, a global English sentiment lexicon is needed in this study. For this purpose, SentiWordNet and SenticNet 3, which have been found very helpful in sentiment analysis and opinion mining problems, are preferred. SentiWordNet is a public lexical resource and devised for using sentiment classification and opinion mining applications [3]. In SentiWordNet, each synset is associated to three numerical scores P (positive), O (objective) and N (negative), describing how positive, objective, negative the terms contained in the synset are [21]. These P, O and N scores are in range [0.0-1.0] and summation of these scores is equal to 1.0. The method used in SentiWordNet is based on quantitative analysis of the glosses associated to synsets [19].

SenticNet 3 is a freely available semantic and affective resource for sentic computing and exploits energy flows to provide the semantics and sentics associated with 30.000 multi-word expressions [11]. SenticNet[2] is presented as RDF XML format and this resource consists of a sentic vector, which contains pleasantness, attention, sensitivity, and aptitude, and a polarity value, which takes the interval from -0.999 to +0.999, for each concept.

To use the SentiWordNet and SenticNet 3 in this study, first Turkish sentiment seeds are translated to English with a fast translator. Thus, in our Java code gtranslateapi-1.0.jar is used. Gtranslateapi[3] provides a simple, unofficial, Java client API for using Google Translate. 189 Turkish sentiment seeds are translated into English using this API. After translation, the results are examined with Seslisozluk[4], which is top ranked dictionary website in Turkey and 6th in the world. It is observed that there are some problems about translation into English. Some Turkish words cannot be translated to English and these words are manually translated based on their meaning and lexical category and some of them are translated but it has been seen that their meanings are considerably different. One of the main reasons of this situation is semantic differences between the two languages. To solve this problem WordNet, which is a lexical dictionary [32], is used to verify correctness of the translation result. This verification process is carried out and adjustments are made manually. As a result of these steps, the solution tree is obtained. The English version of the tree is presented in Figure 2.

To evaluate the acceptability of the proposed method, SentiWordNet based algorithms, which are the first sense, the total sum and the weighted sum, and SenticNet 3 are implemented and described below.

For a sentiment seed s_i its synset defined as:

$$Synset_i = \{S_{\#1}, S_{\#2}, ..., S_{\#n}\}. \tag{1}$$

[2]http://sentic.net/senticnet-3.0.zip

[3]http://code.google.com/p/google-translate-api-v2-java/

[4]http://www.seslisozluk.net/

In the first sense algorithm, only the first sense score of the sentiment seed is used. The basic idea of this algorithm is that the first sense of the sentiment seed offers the most common usage [17]. The algorithm for this approach is displayed in Algorithm 2.

Algorithm 2: First sense algorithm.

```
procedure FIRSTSENSE(SeedDictionary)
  for all SentimentSeed s_i ϵ SeedDictionary do
     if Positive(S_#i) > Negative(S_#i) then
        SeedValue(s_i) ←Positive
     else if Positive(S_#i) < Negative(S_#i) then
        SeedValue(s_i) ←Negative
     else
        SeedValue(s_i) ←Objective
     end if
  end for
end procedure
```

The results of the algorithm are displayed in Table 1. In the first sense approach the sentiment seed can take three different values, as stated here: positive, objective and negative. As it is seen in the table below, the sentiment value of 41 sentiment seeds are estimated incorrectly, this corresponds to approximately 21% of all seeds.

Table 1. Confusion matrix of first sense algorithm.

Actual	**Predicted**		
	Positive	**Objective**	**Negative**
Positive	74	15	5
Objective	0	0	0
Negative	9	12	74

When the first sense is examined in terms of incorrectly classified sentiment seeds, it is observed that 24 seeds which are positive or negative are labeled as objective. To give an example, we can consider the Turkish word “bayağı”. The translation API translates the adjective “bayağı” to the English word “plebeian”. First sense scores of “plebeian” in SentiWordNet are P: 0, O: 1, N: 0. Positive and negative sentiment scores of this seed are equal so this seed is labeled as objective. On the other hand, “bayağı” definitely has a negative meaning in Turkish.

The English translation of the Turkish adjective “kart” is “old”. The first sense scores of “old” are P: 0.375, O: 0.625, N: 0.0. Even the first sense algorithm predicts the sentiment value of *“old”* as positive; the Turkish adjective *“kart”* is used for describing negative situations. There are 8 more seeds like this situation.

“Körpe”, which is the antonym of “kart”, is found as negative based on its first sense score. Its English translation is “fresh” and sentiment scores of this adjective are P: 0.375, O: 0.0, N: 0.625. First sense algorithm predicts 5 positive sentiment seeds as negative. The performance measure of the algorithm is given in Table 2 (TP: TP Rate, FP: FP Rate, P: Precision, R: Recall, ACC: Accuracy).

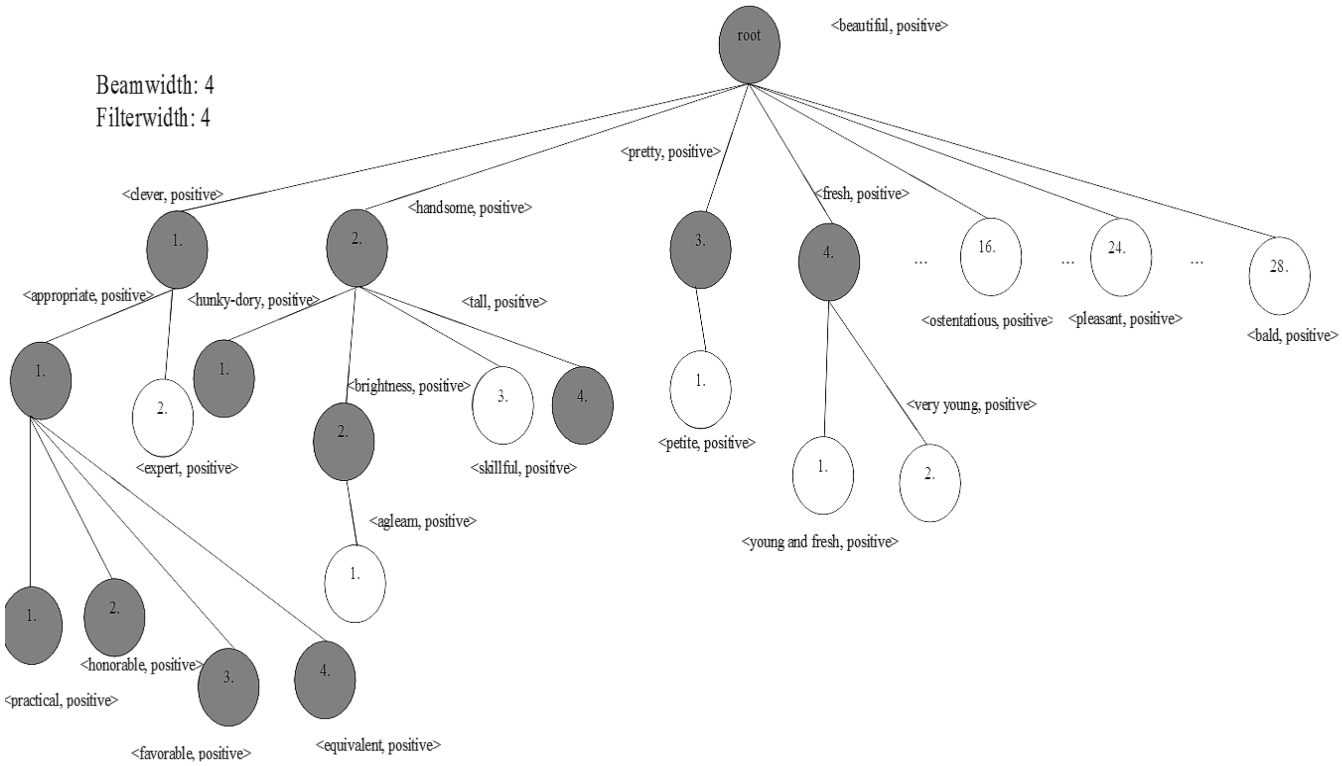

Figure 2. Beam search tree for English.

Table 2. Performance measures for first sense algorithm.

Class	TP	FP	P	R	ACC	F-Score
Positive	0,79	0,09	0,89	0,79	0,85	0,84
Objective	0	0,17	0	0	0,86	0
Negative	0,77	0,06	0,94	0,77	0,86	0,85
Weighted	0,52	0,1	0,61	0,52	0,86	0,56

To evaluate the performance of the first sense algorithm precision, recall, accuracy and F-measure values are calculated. When the results evaluated due to F-measure a low success rate is observed. This is because of the instances which are labeled as objective. In this study objectivity cannot be ignored. So for the purpose of performance evaluation, the usage of accuracy measure would be more appropriate. The first sense algorithm has been observed with an accuracy rate of 86% in average.

In the total sum algorithm, for a sentiment seed the positive and negative sentiment scores in all synsets are summed up separately. The sentiment value of a sentiment seed is obtained by comparing positive and negative scores. If the positive score is greater than the negative score the sentiment value is positive; if negative score is greater than the positive score the sentiment value is negative and it is objective when there is no difference between positive and negative scores. The purposed algorithm for the evaluation is given in the Algorithm 3.

Algorithm 3: Total Sum Algorithm.

:procedure TOTALSUM(SeedDictionary)
for all SentimentSeed $s_i \in$ SeedDictionary do
PositiveSum←0
NegativeSum←0
for all Synset $S_{\#j} \in$ Synset$_i$ do
PositiveSum←PositiveSum+Positive($S_{\#j}$)
NegativeSum←NegativeSum+Negative($S_{\#j}$)
end for
if PositiveSum> NegativeSum then
SeedValue(s_i) ←Positive
else if PositiveSum < NegativeSum then
SeedValue(s_i) ←Negative
else
SeedValue(s_i) ←Objective
end if
end for
end procedure

The results of total sum algorithm are given in Table 3. Like the first sense approach, in this algorithm the sentiment seed can take positive, objective and negative values. Compared with the first sense there has been a huge decline in the number of incorrectly predicted sentiment values. 29 sentiment values, approximately 15% of all the seeds, are predicted incorrectly in this approach.

Table 3. Confusion matrix of total sum algorithm.

Actual	Predicted		
	Positive	**Objective**	**Negative**
Positive	79	10	5
Objective	0	0	0
Negative	6	8	81

When the total sum algorithm is applied to propagating the sentiment seeds only the 29 sentiment seeds are predicted incorrectly. It can be seen from

Table 3 that, 18 of the positive or negative seeds are classified as objective, 5 positive seeds are classified as negative and 6 negative seeds are classified as positive. In Turkish, the adjective “genç” is used in a positive sense. Its English translation is “young” and according to the total sum this adjective is labeled as objective. The total positive and negative scores for young are 0.75. The “hamarat” sentiment seed is used as positive in Turkish sentences but its English counterpart “industrious” is predicted as negative with the total sum algorithm. While the total positive value for “industrious” is 0.375, the total negative value is 0.5. The translation API translates the Turkish adjective “gelişigüzel”, which is used in a negative sense, as “haphazard”. When the adjective “haphazard” is evaluated based on the total sum algorithm it is observed that the total positive value for this adjective is 0.875 and total negative value is 0.5. So this seed is labeled as positive by SentiWordNet. The performance measure of the total sum algorithm is displayed in Table 4.

Table 4. Performance measures for total sum algorithm.

Class	TP	FP	P	R	ACC	F-Score
Positive	0,84	0,06	0,93	0,84	0,89	0,88
Objective	0	0,1	0	0	0,9	0
Negative	0,85	0,05	0,94	0,85	0,9	0,89
Weighted	0,56	0,07	0,62	0,56	0,9	0,59

When the algorithm is evaluated in terms of performance, it can be clearly seen that the F-measure rate is low because of the objective class. There has been a 7% increase in accuracy rate and this algorithm has been observed with an average accuracy rate of 9%.

In the weighted sum algorithm, positive and negative values of the first synset are multiplied separately with the maximum weight which is equal to synset length of the seed. The positive and negative values of other synsets are multiplied with weights in decreasing order respectively [40]. At last the obtained positive and negative scores are summed up separately. As in the total sum, these positive and negative scores are compared to determine the sentiment value of the seed. The algorithm of weighted sum is displayed Algorithm 4.

Algorithm 4:Weighted sum algorithm.

```
procedure WEIGHTEDSUM(SeedDictionary)
 for all SentimentSeed s_i ϵ SeedDictionary do
    PositiveSum←0
    NegativeSum←0
    for all Synset S_#j ϵ Synset_i do
       PositiveSum ← PositiveSum + Positive(S_#j) * (Synset_i.length-j+1)
       NegativeSum←NegativeSum+Negative(S_#j) * (Synset_i.length-j+1)
    end for
    if PositiveSum> NegativeSum then
       SeedValue(s_i) ←Positive
    else if PositiveSum < NegativeSum then
       SeedValue(s_i) ←Negative
    else
       SeedValue(s_i) ←Objective
    end if
 end for
end procedure
```

When Table 5 is examined, it can be seen that only the sentiment values of 26 sentiment seeds are found incorrectly. This corresponds to approximately 13% of all of the process. Although the results have almost the same with the total sum, the weighted sum algorithm is more successful. The performance measure of the algorithm is given in Table 6.

Table 5. Confusion matrix of weighted sum algorithm.

Actual	**Predicted**		
	Positive	**Objective**	**Negative**
Positive	83	7	4
Objective	0	0	0
Negative	7	8	80

In this Algorithm, 15 sentiment seeds are predicted as objective, 4 positive sentiment seeds are predicted as negative and 7 negative sentiment seeds are predicted as positive. The weighted sum scores of some adjectives, which are predicted incorrectly, are exemplified here. Despite this, the Turkish adjective “öldürücü” is used in a negative sense; the positive and negative sentiment score of its English translation “lethal” is 0.0 so this seed is labeled as objective based on the weighted sum. The “bonkör” sentiment seed is used as positive in Turkish sentences but its English translation “generous” is predicted as negative with the weighted sum algorithm. While the positive weighted sum for “generous” is 0.5, the negative weighted sum value is 1.25. Weighted sum scores of “gelişigüzel”, which are calculated as based on its English translation “random”, are 0.125 and 0.0, respectively. Thus it is labeled as positive.

Table 6. Performance measures for first sense algorithm.

Class	TP	FP	P	R	ACC	F-Score
Positive	0,88	0,07	0,92	0,88	0,9	0,9
Objective	0	0,08	0	0	0,92	0
Negative	0,84	0,04	0,95	0,84	0,9	0,9
Weighted	0,57	0,06	0,62	0,57	0,91	0,6

Among these three algorithms, we can remark that the best accuracy rate is obtained with Weighted Sum algorithm. As with other algorithms, the low F-measure rate is obtained because of the objective class. In this study, intended sentiment seed lexicon is expected to include seeds which are labeled as positive and negative only. Because while the initial seed are determining for a sentiment analysis job only the negative and positive ones are included. However; in SentiWordNet, each seed is associated with one of three labels positive, objective and negative. For this reason, precision, recall and F-measure have been observed as zero for objective class. So for the F-measure rate, SentiWordNet based algorithms get the weighted rate of between 56-60%. The same situation is observed for SenticNet 3.

As mentioned above, each concept in SenticNet 3 takes polarity value range from -0.999 to+0.999. The sentiment value of a seed is determined according to this value. If the polarity value is greater than zero the sentiment value is positive; if the polarity value is less than zero the sentiment value is negative and the seed is not available in resource the polarity value is objective. The results of the SenticNet 3 are displayed in Table 7.

If the Table 7 inspected carefully, then it can be noticed that, there has been an increase in the number of incorrectly predicted sentiment values. 42 sentiment values, approximately equal to 22%, are predicted incorrectly.

Table 7. Confusion matrix of SenticNet 3.

Actual	Predicted		
	Positive	Objective	Negative
Positive	77	14	2
Objective	0	0	0
Negative	7	19	70

With the usage of SenticNet 3, 33 positive or negative sentiment seeds are classified as objective, 2 positive sentiment seeds are classified as negative and 7 negative seeds are classified as positive. SenticNet 3 classifies 33 adjectives as objective this is because the resource does not contain these seeds. If we explain this situation with an example; in Turkish, adjective “edalı” is used as positive sense, as “coquettish”. This adjective is classified as positive by SentiWordNet based algorithms. However, “coquettish” is classified as objective by using SenticNet 3. Once the positive seeds, which are predicted as negative, are examined, one of the seed we are faced with “sağlam”. The word “sağlam” is used as positive in Turkish sentences, but its English translation “sturdy” is predicted as negative with the polarity value of -0.053 in SenticNet 3. And it is also possible to give example of sentiment seed which has negative meaning in Turkish but has positive meaning in English. While “geveze” is used as negative sense in Turkish, its English translation “garrulous” is predicted as positive with polarity value 0.878. After all, the performance measure based on SenticNet 3 is summarized in Table 8.

Table 8. Performance measures for SenticNet 3.

Class	TP	FP	P	R	ACC	F-Score
Positive	0,83	0,07	0,92	0,83	0,88	0,87
Objective	0	0,17	0	0	0,83	0
Negative	0,73	0,06	0,94	0,73	0,85	0,82
Weighted	0,52	0,1	0,62	0,52	0,85	0,56

By looking at the results, it can be seen that SenticNet 3 provides similar performance with First Sense algorithm. As compared with the other two algorithms, the performance of SenticNet 3 is slightly worse on the seed lexicon this is because the resource which doesn’t contain all seeds in the lexicon.

When the incorrectly predicted sentiment seeds and their values are examined, it can be clearly seen that the problem arises from the semantic differences between the two languages. As a result of experiments it has been observed that the purposed seed lexicon, which is expanded by using Beam Search algorithm, can be a base lexicon for Turkish sentiment analysis. Especially for the extraction of domain specific sentiment seeds and their values there is a need for such a lexicon.

5. Conclusions

In this article, we present a beam search tree approach for sentiment lexicon compilation in Turkish. A beam search approach is applied for the first time to this kind of a problem. Initially, 62 sentiment seeds and their associated sentiment values are collected, and then this set of seeds is expanded by exploiting word relationships by using a beam search tree approach. It should be noted that, for the optimality of the algorithm, the parameters such as local evaluation function, global evaluation function, filter width and beam width are very important. In our study, while global and local evaluation functions are constituted by considering language specific features; the filter width and beam width are determined by the results of some experiments. As a result of the study 251 accurate seeds are obtained. The correctness of sentiment values of these seeds is determined with SentiWordNet and SenticNet 3; however, at first all of them are translated to English. To test the correctness of sentiment values via SentiWordNet first sense, the total sum and the weighted sum have been conducted. Among these four algorithms, the weighted sum has been observed to be most successful with an average accuracy rate of 91% for lexicon in hand. Our experimental results show that the beam search tree approach is an effective method for construction of the sentiment seed lexicon when the results are evaluated due to an English sentiment lexicon.

For a specific domain, the sentiment analysis studies make use of both a domain-specific and general-purpose sentiment lexicons. In the first approach, constructing such a lexicon from a very small seed word set is time consuming and costly because of manual effort involved. And in this approach it is difficult to get good coverage of a target domain from manually selected words and it requires reconstruction of seed set for every new domain. In the second approach, the lexicon is not capable of capturing the sentiment variations across different domains. Considering all these cases, the main contribution of the study is, constructing a seed lexicon, which is largest as possible, contains only the words has unambiguous positive or negative orientation for sentiment analysis studies in Turkish.

The second one is, for the resource poor languages like Turkish, there haven't been a baseline method which can be used to test the effectiveness of a new sentiment lexicon. Thereby, it is concluded that, the weighted sum is an accurate method to measure the performance of such a lexicon.

Acknowledgement

This study is supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under project number 114E422.

References

[1] Abdul-Mageed M., Diab M., and Korayem M., "Subjectivity and Sentiment Analysis of Modern Standard Arabic," *in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Shortpapers*, Portland, pp. 587-591, 2011.

[2] Araya I. and Riff M., "A Beam Search Approach to the Container Loading Problem," *Computers and Operations Research*, vol. 43, pp. 100-107, 2014.

[3] Baccianella S., Esuli A., and Sebastiani F., "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining," *in Proceedings of the 7th International Conference on Language Resources and Evaluation*, Valletta, pp. 2200-2204, 2010.

[4] Banea C., Mihalcea R., and Wiebe J., "A Bootstrapping Method for Building Subjectivity Lexicons for Languages with Scarce Resources," *in Proceedings of the 6th International Conference on Language Resources and Evaluation*, Marrakech, pp. 2764-2767, 2008.

[5] Bautista J., Pereira J., and Adenso-Diaz B., "A Beam Search Approach for the Optimization Version of the Car Sequencing Problem," *Annals of Operations Research*, vol. 159, no. 1, pp. 233-244, 2008.

[6] Bennell J. and Song X., "A Beam Search Implementation for the Irregular Shape Packing Problem," *Journal of Heuristics*, vol. 16, no. 2, pp. 167-188, 2010.

[7] Blum C., "Beam-ACO-Hybridizing Ant Colony Optimization with Beam Search: An Application to Open Shop Scheduling," *Computers and Operations Research*, vol. 32, no. 6, pp. 1565-1591, 2005.

[8] Brooke J., Tofiloski M., and Taboada M., "Cross-Linguistic Sentiment Analysis: from English to Spanish," *in Proceedings of International Conference Recent Advances in Natural Language Processing*, Borovets, pp. 50-54, 2009.

[9] Cakmak O., Kazemzadeh A., Yildirim S., and Narayanan S., "Using Interval Type-2 Fuzzy Logic to Analyze Turkish Emotion Words," *in Proceedings of Signal and Information Processing Association Annual Summit and Conference*, Hollywood, pp. 1-4, 2012.

[10] Cambria E., Havasi C., and Hussain A., "SenticNet 2: A Semantic and Affective Resource for Opinion Mining and Sentiment Analysis," *in Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference*, Palo Alto, pp. 202-207, 2012.

[11] Cambria E., Olsher D., and Rajagopal D., "SenticNet 3: A Common and Common-Sense Knowledge Base for Cognition-Driven Sentiment Analysis," *in Proceedings of the 28th AAAI Conference on Artificial Intelligence*, Québec City, pp. 1515-1521, 2014.

[12] Clematide S. and Klenner M., "Evaluation and Extension of a Polarity Lexicon for German," *in Proceedings of the 1st Workshop on Computational Approaches to Subjectivity and Sentiment Analysis*, Portugal, pp. 7-13, 2010.

[13] Cruz F., Troyano J., Pontes B., and Ortega F., "Building Layered, Multilingual Sentiment Lexicons At Synset and Lemma Levels," *Expert Systems with Applications*, vol. 41, no. 13, pp. 5984-5994, 2014.

[14] Das A. and Bandyopadhyay S., "SentiWordNet for Indian Languages," *in Proceedings of the 8th Workshop on Asian Language Resources*, Beijing, pp. 56-63, 2010.

[15] De Albornoz J., Plaza L., and Gervás P., "Sentisense: An Easily Scalable Concept-Based Affective Lexicon for Sentiment Analysis," *in Proceedings of the 8th International Conference on Language Resources and Evaluation*, Istanbul, pp. 3562-3567, 2012.

[16] Dehkharghani R., Saygın Y., Yanıkoğlu B., and Oflazer K., *Language Resources and Evaluation*, Springer-Netherlands, pp. 1-19, 2015.

[17] Demirtaş E., Cross-Lingual Sentiment Analysis with Machine Translation, Master's Thesis, Eindhoven University of Technology, 2013.

[18] Ding X., Liu B., and Yu P., "A Holistic Lexicon-Based Approach to Opinion Mining," *in Proceedings of the International Conference on Web Search and Data Mining*, Palo Alto, pp. 231-240, 2008.

[19] Duyu T., Bing Q., LanJun Z., KamFai W., YanYan Z., and Ting L., "Domain-Specific Sentiment Word Extraction by Seed Expansion and Pattern Generation," Retrieved September 1,

2015 from http://arxiv.org/pdf/1309.6722v1.pdf, Last Visited, 2013.

[20] Esuli A. and Sebastiani F., "Determining the Semantic Orientation of Terms Through Gloss Classification," *in Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, Bremen, pp. 617-624, 2005.

[21] Esuli A. and Sebastiani F., "SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining," *in Proceedings of the 5th International Conference on Language Resources and Evaluation*, Genoa, pp. 417-422, 2006.

[22] Fellbaum C., *WordNet: An Electronic Lexical Database*, MA: MIT Press, 1998.

[23] Hu M. and Liu B., "Mining and Summarizing Customer Reviews," *in Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Seattle, pp. 168-177, 2004.

[24] Kaji N. and Kitsuregawa M., "Building Lexicon for Sentiment Analysis from Massive Collection of HTML Documents," *in Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Prague, pp. 1075-1083, 2007.

[25] Kamps J., Marx M., Mokken R., and De Rijke M., "Using Wordnet to Measure Semantic Orientation of Adjectives," *in Proceedings of the 4th International Conference on Language Resources and Evaluation*, Lisbon, pp. 1115-1118, 2004.

[26] Kaur A. and Gupta V., "A Novel Approach for Sentiment Analysis of Punjabi Text using SVM," *The International Arab Journal of Information Technology*, vol. 14, no. 5, pp. 707-712, 2014.

[27] Kim S. and Hovy E., "Determining the Sentiment of Opinions," *in Proceedings of the 20th International Conference on Computational Linguistics*, Geneva, pp. 1367-1373, 2004.

[28] Liu B., *Sentiment Analysis and Opinion Mining*, Morgan and Claypool Publishers, 2012.

[29] López-Ibáñez M. and Blum C., "Beam-ACO for the Travelling Salesman Problem with Time Windows," *Computers and Operations Research*, vol. 37, no. 9, pp. 1570-1583, 2010.

[30] Lu B., Song Y., Zhang X., and Tsou B., "Learning Chinese Polarity Lexicons by Integration of Graph Models and Morphological Features," *in Proceedings 6th Asia Information Retrieval Societies Conference*, Taipei, pp. 466-477, 2010.

[31] Mahyoub F., Siddiqui M., and Dahab M., "Building an Arabic Sentiment Lexicon Using Semi-Supervised Learning," *Journal of King Saud University-Computer and Information Sciences*, vol. 26, no. 4, pp. 417-424, 2014.

[32] Miller G., "WordNet: A Lexical Database for English," *Communications of the ACM*, vol. 38, no. 11, pp. 39-41, 1995.

[33] Mousavi S., Bahri F., and Tabataba F., "An Enhanced Beam Search Algorithm for the Shortest Common Supersequence Problem," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 3, pp. 457-467, 2011.

[34] Nielsen F., "A new ANEW: Evaluation of a Word List for Sentiment Analysis in Microblogs," *in Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big Things Come in Small Packages (MSM)*, Heraklion, pp. 93-98, 2011.

[35] Ow P. and Morton T., "Filtered Beam Search in Scheduling," *International Journal of Production Research*, vol. 26, no. 1, pp. 35-62, 1988.

[36] Özsert C. and Özcan A., "Word Polarity Detection Using a Multilingual Approach," *in Proceedings of International Conference on Computational Linguistics and Intelligent Text Processing*, Samos, pp. 75-82, 2013.

[37] Pang B., Lee L., and Vaithyanathan S., "Thumbs up Sentiment Classification using Machine Learning Techniques," *in Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Philadelphia, pp. 79-86, 2002.

[38] Rao D. and Ravichandran D., "Semi-Supervised Polarity Lexicon Induction," *in Proceedings of the 12th Conference of the European Chapter of the ACL*, Athens, pp. 675-682, 2009.

[39] Riloff E. and Wiebe J., "Learning Extraction Patterns for Subjective Expressions," *in Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Sapporo, pp. 105-112, 2003.

[40] Sabuncuoğlu I. and Bayiz M., "Job Shop Scheduling with Beam Search," *European Journal of Operational Research*, vol. 118, no. 2, pp. 390-412, 1999.

[41] Stone P., Dunphy D., Smith M., and Ogilvie D., *The General Inquirer: A computer Approach to Content Analysis*, MA: MIT Press, 1966.

[42] Strapparava C. and Valitutti A., "WordNet-Affect: an Affective Extension of WordNet," *in Proceedings of the 4th International Conference on Language Resources and Evaluation*, Lisbon, pp. 1083-1086, 2004.

[43] Taboada M., Brooke J., Tofiloski M., Voll K., and Stede M., "Lexicon-Based Methods for Sentiment Analysis," *Computational Linguistics* vol. 37, no. 2, pp. 267-307, 2011.

[44] Tan S. and Wu Q., "A Random Walk Algorithm for Automatic Construction of Domain-Oriented Sentiment Lexicon," *Expert Systems with*

Applications, vol. 38, no. 10, pp. 12094-12100, 2011.

[45] Tang J., Guan J., Yu Y., and Chen J., "Beam Search Combined With MAX-MIN Ant Systems and Benchmarking Data Tests for Weighted Vehicle Routing Problem," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 4, pp. 1097-1109, 2014.

[46] Vadlamudi S., Aine S., and Chakrabarti P., "Incremental Beam Search," *Information Processing Letters*, vol. 113, no. 22-24, pp. 888-893, 2013.

[47] Vossen P., Maks I., Segers R., and Van Der Vliet H., "Integrating Lexical Units, Synsets, and Ontology in the Cornetto Database," *in Proceedings of the 6th International Conference on Language Resources and Evaluation*, Marrakech, pp. 1006-1013, 2008.

[48] Vural A., Cambazoglu B., Senkul P., and Tokgoz Z., A Framework for Sentiment Analysis in Turkish: Application to Polarity Detection of Movie Reviews in Turkish, *in Computer and Information Sciences*, Springer London, pp. 437-445, 2013.

[49] Xu G., Meng X., and Wang H., "Build Chinese Emotion Lexicons Using A Graph-based Algorithm and Multiple Resources," *in Proceedings of the 23rd International Conference on Computational Linguistics*, Beijing, pp. 1209-1217, 2010.

[50] Yavuz M., "Iterated Beam Search for the Combined Car Sequencing and Level Scheduling Problem," *International Journal of Production Research*, vol. 51, no. 12, pp. 3698-3718, 2013.

Ekin Ekinci has received her BEng. in Computer Engineering from Çanakkale Onsekiz Mart University in 2009 and ME. in Computer Engineering from Gebze Teknik University in 2012. She is currently working towards Ph.D. degree in Computer Engineering from Kocaeli University, Turkey. Also, she is currently a Research Assistant of Computer Engineering Department at Kocaeli University in Turkey. Her main research interests include text mining, sentiment analysis, natural language processing and machine learning.

Sevinç Omurca was born in Burdur in Turkey at 23th December 1979. She is an associate professor at the Kocaeli University Computer Engineering Department in Turkey. She has a Ph.D. at the Kocaeli University Electronics and Communications Engineering**.** Her main research interests include text mining, sentiment analysis, natural language processing, machine learning and data mining.