A Sentiment Aggregation System based on an OWA Operator

Hossein Abbasimehr
Faculty of Information Technology and Computer Engineering
Azarbaijan Shahid Madani University
Tabriz, Iran
abbasimehr@azaruniv.ac.ir

Mostafa Shabani
IT Group, Industrial Engineering Department
KNTU University of Technology
Tehran, Iran
mshabani@mail.kntu.ac.ir

Abstract—User-Generated-Content (UGC) in the form of online reviews can be an invaluable source of information for both customers and businesses. Sentiment analysis and opinion mining tools and techniques have been proposed in the literature to extract knowledge from online reviews. Aspect-based opinion mining which has gained growing attention mainly has two tasks including aspect extraction and sentiment polarity detection. Once an aspect-based opinion mining task has been accomplished; a bag of sentiments will be achieved. In many cases, it is necessary to obtain an overall sentiment about a typical aspect. In this study, we have proposed a sentiment aggregation system based on weighted selective aggregated majority OWA (WSAM-OWA). WSAM-OWA considers both the majority and the degree of importance of information source in the process of aggregation. The proposed system exploits the helpfulness rating of reviews in determining the reliability and credibility of each sentiment. A case study was conducted to illustrate the usefulness of the proposed system. The results of this study demonstrated that the proposed sentiment aggregation system could be incorporated in opinion mining systems.

Keywords—Online Review; Opinion Mining; Sentiment Analysis; OWA; Sentiment Aggregation.

I. INTRODUCTION

Nowadays, user-generated content (UGC) [1, 2] often in the form of online reviews in social web applications and online communities [3] has become one of invaluable sources of information for both customers and business [4, 5]. From customers point of view, they can easily learn from the past experiences, ideas, and opinions of other customers to make right decisions in their purchases [6]. Also, form the businesses view point, they can exploit UGC to understand customers’ attitudes about their product or services [7, 8]. In this way, a plethora of tools and techniques in the area of opinion mining and sentiment analysis (e.g. [8-13]) have been devised and can be employed by businesses to extract actionable knowledge from online reviews. Opinion mining and especially aspect-based opinion mining approaches have been gained increasing attention during recent years [14-17].

In the aspect-based opinion mining, there are mainly two tasks including aspect extraction and sentiment polarity estimation [15, 17]. After performing the aspect-based opinion mining on a collection of reviews, for each aspect, a bag of sentiment scores results [17, 18]. For instance, suppose that for the battery of a smartphone product, a bag of sentiments \( b = \{(0.5), (1.4), (2.3), (4.5), (5.6)\} \) is computed with \( b_i = (s_i, u_i) \) where \( s_i \) is the sentiment and \( u_i \) is the cardinality of \( b_i \). To obtain an overall sentiment score about an aspect, the straightforward approach is to use arithmetic mean (AM)[18], however, it is emphasized in the literature that AM in many cases, doesn’t represent the majority opinion [18].

Ordered weighted averaging (OWA) is the popular aggregation operator that has been introduced by [19]. Furthermore, there have been several aggregation operators presented in the literature including majority additive OWA (MA-OWA) [20], selective MA-OWA (SMA-OWA) [21], selective aggregated majority-OWA (SAM-OWA)[18] and weighted selective aggregated majority OWA (WSAM-OWA)[22]. In this study, for the first time, we proposed a system based on WSAM-OWA operator for sentiment aggregation. The proposed system benefits from the advantages of the WSAM-OWA operator [22] in aggregating the sentiment scores by considering both majority and degree of credibility of sentiments. This system receives a list of aspect and a collection of reviews, then performs sentiment detection using SentiStrength [23], afterward, aggregates the results using WSAM-OWA operator. WSAM-OWA considers the degree of importance (the reliability of source) in the aggregation process [22]. In this study, we exploited the helpfulness ratings of online reviews to derive the degree of importance. The helpfulness ratings for an online review is usually given by readers which can be a sign of the helpfulness of a review. To test the effectiveness of the proposed system,
we conduct experiments using reviews gathered from Epinions.com.

The rest of this paper is organized as follows: in section 2 we describe aggregation functions and especially WSAM-OWA; section 3 describes the proposed system for sentiment aggregation. In section 4 we portray the experiments and the results. Section 5 concludes the paper.

II. BACKGROUND

In many areas of applications including group decision making, informational retrieval, opinion mining and so on, an aggregation process is essential to summarize various opinions [18, 21, 22]. Yager [19] developed the ordered weighted averaging (OWA) which is one of the widely used aggregation function. There have been various aggregation operators proposed in the literature including MA-OWA [20], MA-OWA (SMA-OWA) [21], SAM-OWA [18] and WSAM-OWA [22] In the following, we briefly describe the aggregation operator (WSAM-OWA) utilized in this study.

A. WSAM-OWA

WSAM-OWA operator is an extension of SAM-OWA [18] that considers both majority and the degrees of importance (i.e., the reliability of information source) [22]. The mathematical foundation of WSAM-OWA operator is as follows:

Let \( a = (a_1, ..., a_i, ..., a_n) \in \mathbb{R}^n \times \mathbb{N}^n \) with \( a_i = (b_i, m_i, v_i) \) representing the aggregated value, \( b_i \), and its cardinality, \( m_i \), and an associated weighting vector \( V \) of dimension \( n \) such that \( \sum_{i=1}^{n} v_i = 1 \) and \( v_i \in [0,1] \). A WSAM-OWA is a function \( F_{WSAM} : \mathbb{R}^n \times \mathbb{N}^n \rightarrow \mathbb{R} \) defined as [22]:

\[
F_{WSAM} (a_1, a_2, ..., a_n) = \sum_{l=1}^{n} w_{l,i} b_{\sigma(i)},
\]

where \( N = \max_{1 \leq i \leq n} m_i \) and \( \sigma \) denotes a permutation with respect to \( m_i \) such that \( b_{\sigma(i)} \geq b_{\sigma(i+1)} \). The weights are computed by the recurrence relations [22]:

\[
w_{l,1} = \omega_l, \quad \text{if} \quad m_l = 1,
\]

\[
w_{l,k} = \frac{\omega_l y_{l,k} y_k + \omega_{l,k-1} y_k}{z_k}
\]

\[
y_1 = 1, y_k = \begin{cases} 1, & \text{if } \sum_{j=1}^{n} \omega_j y_{j,k} = 0, \\ \sum_{j=1}^{n} \omega_j y_{j,k}, & \text{otherwise}, \end{cases}
\]

\[
z_1 = 1, z_k = \begin{cases} 1, & \text{if } \sum_{j=1}^{n} \omega_j y_{j,k} = 0, \\ 1 + \sum_{j=1}^{n} y_{j,k}, & \text{otherwise}, \end{cases}
\]

With \( 2 \leq k \leq N, 1 \leq i \leq n \), and Where

\[
y_{j,k} = \delta \frac{m_{\sigma(j)} \geq k}{1 - \delta},
\]

In which \( \delta \) is the cardinality relevance factor (CRF) with \( 0 \leq \delta \leq 1 \) [22].

III. THE PROPOSED SYSTEM

In this section, we describe the proposed system for sentiment aggregation.

A. The Proposed system

The proposed system for sentiment aggregation is illustrated in Fig. 1. To describe more, suppose that a typical product has \( n \) reviews, \( \{R_1, R_2, ..., R_n\} \) written by users on a website. Also, consider that \( p \) has \( k \) aspects, \( \{a_1, a_2, ..., a_k\} \). According to the aspect-based opinion mining, a review \( R_i \) can be viewed as a list of aspect-opinion pairs \( R_i = \{<a_j, s_k> | \text{and } a_j \in \{a_1, a_2, ..., a_k\}, s_k \in R \} \) in which \( a_j \) is an aspect, and \( s_k \) is the sentiment on \( a_j \). Following the aspect-based opinion mining, for each aspect, a list of sentiment scores is computed using a sentiment estimation approach. The proposed system uses WSAM-OWA operator [22], and aggregates the sentiment scores by considering both majority and degree of credibility of sentiments. The definition of the majority is the count of each sentiment \( s_k \). To compute the degree of credibility of sentiment, we need to have a detailed information about reviewers’ expertise. As such information may not be available, in this study, we utilize helpfulness ratings of reviews to obtain the credibility of each sentiment. In our review dataset, each review has a helpfulness rating including “Not Helpful”, “Somewhat Helpful”, “Helpful”, “Very Helpful” and “Most Helpful”. Also, there may be some reviews that have not received helpfulness rating, those were labeled as
“Not Yet”. The mapping between the helpfulness rating and the degree of credibility of sentiment is given in Table I.

As a sentiment score is calculated based on a sentiment sentence extracted from a review, the helpfulness rating of that review is considered as the degree of credibility of the sentiment. After calculating the majority and degree of credibility for each sentiment, the WSAM-OWA operator is applied to compute the aggregate sentiment on each aspect.

Suppose a bag of sentiment scores \( (b_i, m_i, w_i) \) of aspect \( a_j \) of product \( p_i \) where \( b_i \) is the sentiment score, \( m_i \) is the number of \( b_i \) and \( w_i \) is the degree of credibility of \( b_i \).

The aggregated sentiment score of aspect \( a_j \) of product \( p_i \) is calculated using WSAM-operator (7)

\[
\text{Sentiment}(p_i, a_j) = F_{\text{WSAM}}(a)
\]

Where \( F_{\text{WSAM}}(a) \) is the result of applying the WSAM-OWA operator [22].

IV. EXPERIMENTS

In this section, we describe the empirical study and the results.

A. Data Collection

In this study, we collected datasets of customer reviews for MP3 Player products from Epinions.com. To test the usefulness of the proposed aggregation system, we selected one product and conducted the experiments using it. This product has 42 reviews and 1344 review sentences. According to the proposed system in Fig. 1, the aspect list which consists of 17 aspects is given in Table II.

B. Extracting subjective review sentences

Reviews sentences are classified into subjective and objective sentences [15, 24]. Only a subjective sentence may contain an opinion. Therefore, the subjective sentences from sentences collection were extracted. In this study, sentences containing opinion words or adjectives were considered as subjective sentence.

C. Extracting sentences containing each Aspect of aspect list

In this step, all subjective review sentences containing aspects are extracted.

D. Sentiment estimation using SentiStrength software

In this study, SentiStrength\(^1\) [23, 25] polarity on each aspect in a sentence, was utilized to compute sentiment strength on each aspect in a sentence SentiStrength. It contains a dictionary of 2846 sentiment words. SentiStrength reports two sentiment strengths, for negative sentiment it uses a range

| Table I. The Mapping Between the Helpfulness Rating and the Degree of Credibility |
|---------------------------------|-------------------|
| Helpfulness rating | Corresponding degree of credibility of sentiment |
| "Not Helpful" | 0 |
| Somewhat Helpful | 1 |
| Helpful | 3 |
| Very Helpful | 4 |
| Most Helpful | 5 |
| Not Yet | 1 |

\(^1\)http://sentistrength.wlv.ac.uk/
from -1 (not negative) to -5 (extremely negative) and to compute positive sentiment it uses a range from 1 (not positive) to 5 (extremely positive). In this study, the positive and negative sentiments are added to obtain an overall sentiment strength and in the range –4 to 4. After computing the overall sentiment score of each sentences, for each aspect, the sentiment scores and their associated counts were obtained.

E. Computing the degree of credibility of each sentiment

As mentioned in the previous section, the degree of credibility of each sentiment was calculated using the helpfulness rating of the corresponding review. Based on what mentioned in the previous section, for each sentiment score, the degree of credibility (the importance weight) is achieved. At the end of this step, for each aspect, a bag of opinions was constructed.

F. Aggregation of sentiments using WSAM

As depicted in Fig. 1, we utilized the WSAM aggregator to aggregate customers’ sentiments. Table III shows the results of sentiment aggregation using different operators including AM, weighted arithmetic mean (WAM) [22], SAM, and WSAM. For SAM and WSAM the CRF $\delta=0.8$. Since, AM does not take into account the degree of important of sentiments sources and assign an equal weight to each sentiment in a given bags of sentiments. In addition, as another drawback, AM does not consider the cardinality of each sentiment in the aggregation process.

The weighted arithmetic mean (WAM) operator, which is a special type of WSAM operator. In fact, WAM is WSAM with $\delta=0.5$. Therefore, the main weakness of the WAM is that it does not consider the opinion of majority (i.e. devoting a high weight to a high majority sentiment).

The main advantage of WSAM over the other operators is considering both the majority concept and the degree of importance in the aggregation process. This differs from SAM-OWA operator that emphasizes only on the majority (cardinality) and considers the degree of importance uniform. As illustrated in Table III, for each aspect, the aggregated result of WSAM is different from the results reflects both cardinality and degree of importance of the source.

To validate the results of the proposed system, from Table III, consider the bag of sentiments for the aspect TouchScreen $\{(-1,3,0.25),(0,4,0.26),(1,4,0.23),(2,4,0.26)\}$.

For the TouchScreen aspect, the results of applying AM is 0.6, as mentioned before, AM does not consider the cardinality and the degree of importance and assign an equal weight for each item during the aggregation process.

Also, the results of using WAM is 0.5, so the resulted value of the WAM operator is below than the result of AM. The reason for this that is that for the TouchScreen aspect, the weights of sentiment -1 and 0 is greater than the weight of sentiments 1 and 2. Therefore, WAM that consider only the degree of importance of sentiment produces a value below than the value of AM. The majority’s sentiment is not taken into account using WAM operator.

Furthermore, SAM which considers only the majority vote, expected to yield a value closer to 1 according to the values of the highest cardinality sentiment. Since sentiments 0, 1, and 2 have the highest cardinality, consequently the SAM operator yields the value 0.81 which is closer to the average the sentiments 0, 1, and 2.

Finally, the WSAM operator yields 0.75. This value is closer to the sentiment having the highest cardinality and the highest degree of importance. Because the degree of importance for sentiment -1 is higher than the ones of sentiment 1, then aggregated value achieved by WSAM is lower than the output of the SAM.

The results of comparisons indicate that WSAM effectively considers the majority’s opinion and the degree of importance of sentiments in computing an aggregated value.

One of main motivations of proposing the sentiment aggregation system is to consider the reliability of source in the aggregation process. As the quality and credibility of online reviews may vary [5], the inclusion of the degree of reliability of a sentiment is a significant issue that should be considered in the development of sentiment aggregation systems. The results of this study indicated that the proposed sentiment aggregation system could be incorporated in opinion mining systems.

V. Conclusion

One of the important issues in aspect-based opinion mining is how to aggregate the resulted bag of sentiments. The straightforward approach is to use the arithmetic mean. However, it doesn’t reflect the majority opinion. In this study, a sentiment aggregation system was proposed that incorporated WSAM-OWA, a variation of OWS operator to aggregate bags of sentiments. WSAM-OWA considers both the majority and degree of importance of information in the aggregation process which is its advantages. To test the usefulness of the proposed system, we conduct a case study using reviews gathered from Epinions.com. The results of this study indicated that the proposed sentiment aggregation system could be incorporated in opinion mining systems. In this study, to compute the degree of importance of each sentiment; i.e. sentiment reliability, we proposed to use the helpfulness rating of each review. As a future work, the credibility of reviewers can be considered in the computation of the degree of importance of each sentiment.

References


### TABLE III. THE RESULTS OF DIFFERENT AGGREGATION OPERATORS FOR SENTIMENT AGGREGATION

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Bag of sentiments</th>
<th>Aggregation operators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AM</td>
<td>WAM</td>
</tr>
<tr>
<td>Application</td>
<td>0.22</td>
<td>0.11</td>
</tr>
<tr>
<td>Battery</td>
<td>0.36</td>
<td>0.37</td>
</tr>
<tr>
<td>Browser</td>
<td>0.69</td>
<td>0.47</td>
</tr>
<tr>
<td>Case</td>
<td>0.5</td>
<td>0.48</td>
</tr>
<tr>
<td>Device</td>
<td>0.85</td>
<td>0.46</td>
</tr>
<tr>
<td>Game</td>
<td>0.77</td>
<td>0.47</td>
</tr>
<tr>
<td>Memory</td>
<td>-0.38</td>
<td>-0.63</td>
</tr>
<tr>
<td>Movie</td>
<td>0.33</td>
<td>0.44</td>
</tr>
<tr>
<td>Picture</td>
<td>0.68</td>
<td>0.26</td>
</tr>
<tr>
<td>Price</td>
<td>-0.50</td>
<td>-0.48</td>
</tr>
<tr>
<td>Screen</td>
<td>0.50</td>
<td>0.09</td>
</tr>
<tr>
<td>Song</td>
<td>0.26</td>
<td>0.45</td>
</tr>
<tr>
<td>Sound</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>Speaker</td>
<td>-0.33</td>
<td>-0.2</td>
</tr>
<tr>
<td>TouchScreen</td>
<td>0.6</td>
<td>0.50</td>
</tr>
<tr>
<td>Video</td>
<td>0.3</td>
<td>1.04</td>
</tr>
<tr>
<td>Wi-Fi</td>
<td>-0.2</td>
<td>-0.38</td>
</tr>
</tbody>
</table>


H. S. Manaman, S. Jamali, and A. Al-Ahmad, "Online reputation measurement of companies based on user-generated content in online social networks," *Computers in Human Behavior*, vol. 54, pp. 94-100, 2016.


