Automatic Aspect-Based Sentiment Summarization for Visual, Structured, and Textual Summaries

Tuan Anh Tran[1], Jarunee Duangsuwan[2], and Wiphada Wettayaprasit[3]

ABSTRACT: Online reviews are valuable sources of information to help companies to make good decisions for business intelligence. In this study, we propose an Automatic Aspect-based Sentiment Summarization (AAbSS) system that has two components and can generate a summary as an output. The first component is the Aspect-based Knowledge Representation and Selection (AKRS) used to represent reviews based on aspects and their polarities for selecting aspect-based knowledge. To represent and selection knowledge, a set of frequency of polarity opinion strength, a summation of frequency of aspect, and an information of aspect are initiated. The second component is the Summary Format Generation (SFG) used to automatically generate three kinds of formats. In this component, new representations for visual and structured summaries, and a new way of applied natural language generation for a textual summary are proposed. In the experiments, 15 domains from benchmark datasets of customer reviews, e.g. cell phone, digital camera, etc. are used. The proposed system not only fast generates summaries having good performance when compared to other summaries generated by other systems and easily updated when adding new reviews in the same domain but also does not spend memory capacity to save any raw data.

Keywords: Aspect-based Knowledge, Automatic Summary Generating, Structured Summary, Textual Summary, Visual Summary

DOI: 10.37936/ecti-cit.2021151.237565
Received January 29, 2020; revised May 21, 2020; accepted June 12, 2020; available online January 18, 2021

1. INTRODUCTION

Nowadays, the development of technology is so fast, especially in internet technologies. Thus, users can express their opinions, emotions, and attitudes on social networks or comment feedback on websites about products they purchased, services they used, and the experiences they had. For example, feedback about a camera from customers is "the picture quality of this camera is great". These comments usually contain users' expressed words that play the major factor in assessing users' satisfaction [1]. Hence, these comments and feedback help not only the business to enhance a quality of products or services but also the customers to have feedback from others about the interesting products. On social media, users do not usually follow grammar or language rules when they comment or feedback. Noisy texts or words out of vocabulary are also in online reviews [2]. With the big amount of these reviews, the companies/governments/customers cannot easily understand the important information by reading manually to make their decisions. Therefore, we need one system that can automatically identify, extract important knowledge, and produce a good summary. Aspect-based sentiment summarization can be applied to this system. The sentiment summarization system based on aspect has an input as customers' reviews and produces an output as a summary by having an entity and its aspects with their polarities (positive/negative). The summary which is used to enhance the businesses is a part of business intelligence [3].

Currently, the output of the existing aspect-based sentiment summarization systems is represented in a visual summary, a structured summary, or a textual summary. For the visual summary, users' opinions are

[1,2,3]Artificial Intelligence Research (AIR) Lab, Department of Computer Science, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla, Thailand, Email: tattinsphue@yahoo.com, jarunee.d@psu.ac.th and wiphada.w@psu.ac.th

This work was supported by Thailand's Education Hub for Southern Region of ASEAN Countries (TEH-AC) and PSU.GS. Financial Support for Thesis (Fiscal Year: 2019).

depicted in each aspect with its polarities by charts, symbols, or a rating meter [4]. With the structured summary, polarities are grouped together by aspects [5]. Moreover, the total positive/negative reviews of each aspect are shown. At each polarity, opinions or sentences related to aspects are selected and depicted. The textual summary means that sentences related to aspects are selected and arranged in summary [4].

In this paper, we propose an Automatic Aspect-based Sentiment Summarization System (AAbSS) for customer reviews that can produce a summary including relevant aspects. Furthermore, the system can generate the output with three kinds of format (visual/structured/textual). The proposed system can help companies/governments/customers to have various perspectives to support decision making. The input of the system can be various datasets annotated aspect terms and polarity. In addition, the proposed system can easily update the generated summary when adding new reviews. With a visual summary, new representations on a chart are proposed. With a structured summary, a new template is proposed. With a textual summary, Natural Language Generation (NLG) is applied in our system to have a readable and flexible summary. Our contributions in this work are proposing of 1) Automatic Aspect-based Sentiment Summarization system that can generate an aspect-based summary without building any tree or training any data and 2) methods to generate three kinds of output formats (visual/structured/textual).

2. RELATED WORK

In this part, we will discuss two groups: 1) previous studies with outputs of visual/structured/textual summaries and 2) commercial tools.

In the first group, previous studies related to aspectbased sentiment summarization that provided the summary output (visual/structured/textual).

For the visual summary, the summary is represented by charts, symbols, and a rating meter to help readers to have a quick view about aspects or a whole product/service that the readers are interesting. Hu and Liu [6] used a vertical bar chart to depict polarities (positive/negative) for each aspect on one chart. Carenini and Rizoli [7] also used a bar chart to indicate polarity and opinion strength. The authors can represent one aspect per chart. Abulaish, et al. [8] represented polarities of all aspects on a bar chart with a percentage of a count of reviews. Kherwa, et al. [9] used three kinds of charts to visualize aspects and their polarity. The first chart was Google-o-meter to depict polarity for one aspect. The second chart was a bar chart to describe polarity for all aspects. The last chart was a pie chart to visualize percentages for aspects. Kamal [10] used bar and pie charts to express the degree of opinion for aspects of a product. Wu, et al. [11] used a horizontal bar chart to show numbers of positive and negative opinions for all aspects. Kanbur and Aktas [12] used a vertical bar chart to represent the top 10 aspects and polarity. Most of the visual summaries represented polarities for aspects on the bar chart.

For the structured summary, polarities (positive or negative) are grouped together by aspects. At each polarity, sentences including aspects are extracted and depicted [5]. Hu and Liu [13] used statistics for opinion sentences (positive, negative) for each aspect. Zhuang, et al. [14] used statistics for opinion sentences (pro, con) for each aspect or each person's name such as actor, actress, director, etc. Blair-goldensohn, et al. [15] used quantitative and qualitative summary. The quantitative summary was in "star rating" by transferring from all sentences that were classified under a given aspect. The qualitative summary was represented by a set of sentences to represent the sentiments being expressed for each aspect. Ly, et al. [16] clustered opinion sentences into two parts (positive, negative) and chose the most representative sentence for each part. Liu, et al. [17] filtered based on LSA (LatentSemantic Analysis) and grouped reviews by polarity (positive, negative). Jmal and Faiz [18] used statistics for product and each aspect with a percentage of customer satisfaction by calculating sentence score and review score. Bafna and Toshniwal [19] used statistics opinion reviews (positive, negative) for each aspect. Yauris and Khodra [20] used statistics opinion aspects (positive, negative) for each aspect category. López Condori and Salgueiro Pardo [5] used statistics for opinion reviews (positive, negative) for each aspect. For each aspect, the best sentence containing aspect was chosen. Amplayo and Song [21] suggested a sentiment score of each aspect in all reviews. The structured summary usually grouped aspects with their polarities and listed sentences.

For the textual summary, sentences containing aspects and having high scores are extracted from the dataset. This is the main idea of the extractive summarization. Most of the previous studies used this idea to calculate sentence scores. Titov and McDonald [22] chose words having a top probability for each topic. Carenini and Rizoli [7] counted opinion strength for each aspect and built a hierarchical tree for aspects. Lu, et al. [23] chose phrases with the highest support in each aspect. Xu, et al. [24] calculated a ranking score for each sentence. Carenini, et al. [25] suggested MEAD* framework that extracted sentences from a dataset to generate a summary. These sentences contained aspects whose polarity and opinion strength scores were high. Yu, et al. [26] suggested choosing phrases by sorting aspects. If two phrases had the same aspect, these phrases were merged with only one aspect. López Condori and Salgueiro Pardo [5] used Greedy algorithm to select k sentences. Angelidis and Lapata [27] sorted polarities of comments into positive and negative, and

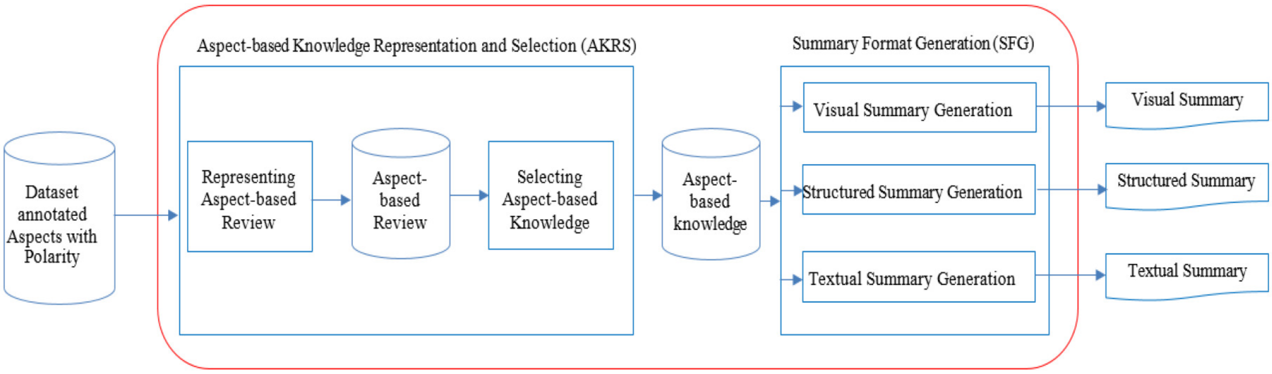


Fig.1: *An Automatic Aspect-based Sentiment Summarization (AAbSS) system.*

(S1): volume[-2]##i have excellent hearing but the volume level on this phone is especially quiet .
(S2): volume[-2]##the volume .
(S3): battery life[+3]##on the up-side , the phone has amazing battery life .
(S4): phone[+1]##overall this is a slightly better than average phone .
(S5): vibration[-1]##the vibration is not top .

(a) Cell phone review [13]

<sentence id="1004293:5">
<text> Avoid this place! </text>
<Opinions>
<Opinion target="**place**"
category="RESTAURANT#GENERAL"
polarity="**negative**" from="11" to="16"/>
</Opinions>
</sentence>

(b) Restaurant review [53]

Fig.2: *Excerpts of review benchmark datasets.*

removed neutral or redundant comments. Tran, et al. [28] proposed the method to select top interesting aspects whose total users' comments were high. After selecting, newly generated sentences and selected aspects were used to generate a text summary only. Carenini, et al. [25] used discourse relations among aspects. Gerani, et al. [29] suggested the method to select aspects from a discourse tree by using rhetorical relations. Gerani, et al. [29] built an aspect tree based on discourse structure. López Condori and Salgueiro Pardo [5] used the K-means algorithm to cluster aspects. Yang, et al. [30] proposed the text categorization task to find aspects and explore different ways to express different text categories. Gerani, et al. [31] proposed a framework to build an aspect tree based on conceptual/rhetorical/hybrid. In general, the summarization systems consumed time to build tree or listed sentences containing aspects from the dataset.

In the second group, online commercial tools analyse users' input and return results to the users. The online commercial tools are Awario [32], Hootsuite Insights [33, 34], Sentiment Viz [35, 36], Social Mention [37, 38], Social Searcher [39-41], Talk walker [42, 43], SentiStrength [44, 45], Lexalytics [46-48], Meaning Cloud [49, 50], Sentigem [51], and Sentiment Analyzer [52]. There are six tools [32-43,46-48] which are real-time search and retrieve results from the internet, e.g. Facebook, Twitter, etc. Most of these tools have their visual outputs except SentiStrength [44, 45], Meaning Cloud [49, 50], and Sentigem [51]. All these tools concern with the sentiment. For example, Lexalytics generates three kinds of outputs with the sentiment. Sentiment Viz produces two kinds of outputs concerning with sentiment. The textual outputs of these tools are generated by selecting sentences from an input.

These tools are usually free or will be paid for a fee after one month trial. Most of the tools are easy to use with a friendly interface and provide a real-time search. Keyword(s) can be an input of some tools.

3. PROPOSED METHOD

To produce automatically a summary from aspect-based customer reviews, the Automatic Aspect-based Sentiment Summarization (AAbSS) system is proposed as illustrated in Figure 1. The system consists of two components: 1) Aspect-based Knowledge Representation and Selection (AKRS), and 2) Summary Format Generation (SFG). An input of the system is customer reviews, e.g., products or services reviews, which are annotated aspect terms and their polarities in each sentence. An output of the system is a summary whose format is one of three kinds (visual, structured, and textual).

Algorithm: Aspect Review

Input: Review annotated aspects with polarity
Output: The aspect-based review A = $\{< a_i, \mathrm{F}_{a_i}, SumF_{a_i}, I_{a_i} >\}$

A = $\{< a_i, \mathrm{F}_{a_i}, SumF_{a_i}, I_{a_i} >\}$ *// the aspect-based review*
for each aspect a_i in the review do:
  if a_i is not in A then
    add new a_i to A *// new aspect*
    $\mathrm{F}_{a_i} \leftarrow \emptyset$ *// initialize frequency of aspect a_i*
    $SumF_{a_i} \leftarrow 0$ *// initialize summation of frequency for aspect a_i*
    $I_{a_i} \leftarrow 0$ *// initialize information of aspect a_i*
  if a_i has polarity opinion strength s then $f_s \leftarrow f_s + 1$ *// f_s is a frequency of polarity opinion strength s*
  else:
    if a_i polarity is negative then $f_{-1} \leftarrow f_{-1} + 1$
    if a_i polarity is neutral then $f_0 \leftarrow f_0 + 1$
    if a_i polarity is positive then $f_{+1} \leftarrow f_{+1} + 1$
for each aspect a_i in A do:
  $SumF_{a_i} \leftarrow \sum_{s=-3}^{+3} f_s$
  $I_{a_i} \leftarrow \sum_{s=-3}^{+3} (s \times f_s)$
sort A in descending order using I_a *// for ranking aspect-based review*
return the aspect-based review A

Fig.3: *Representing for aspect-based review from annotated review.*

Table 1: *The Aspect-based Review A for the Cell Phone Review in Figure 2 (a) sorted by I_a.*

i	Aspect a_i	F_{a_i} $< f_{-3}$,	f_{-2},	f_{-1},	f_0,	f_{+1},	f_{+2},	f_{+3}¿	$SumF_{a_i}$	I_{a_i}
2	battery life	< 0,	0,	0,	0,	0,	0,	1 >	1	3
3	phone	< 0,	0,	0,	0,	1,	0,	0 >	1	1
4	vibration	< 0,	0,	1,	0,	0,	0,	0 >	1	-1
1	volume	< 0,	2,	0,	0,	0,	0,	0 >	2	-4

3.1 Aspect-based Knowledge Representation and Selection Component

The first component named Aspect-based Knowledge Representation and Selection (AKRS) is used to represent an aspect-based review from datasets having aspects and their polarities. The two procedures are in the AKRS component: 1) representing aspect-based review and 2) selecting aspect-based knowledge. After representing, the aspects are selected into aspect-based knowledge. Before discussing these procedures, definitions are introduced as the following:

Let s be a polarity opinion strength, and s is an integer number in [-3, +3].

Let a be an aspect.

Definition 1: A ***frequency of polarity opinion strength*** f_s is a number of comments from users with polarity opinion strength s for the aspect.

Definition 2: A ***frequency of aspect*** $\mathrm{F_a}$ is a septuple with a frequency of each polarity opinion strength f_s for aspect a as shown in Formula (1).

$$\mathrm{F_a} =< f_{-3}, f_{-2}, f_{-1}, f_0, f_{+1}, f_{+2}, f_{+3} > \quad (1)$$

where,

f_{-3} is the frequency of polarity opinion strength -3 (or the most negative),

f_0 is the frequency of polarity opinion strength 0 (or neutral), and

f_{+3} is the frequency of polarity opinion strength +3 (or the most positive) in aspect a.

Definition 3: A ***summation of frequency*** $SumF_a$ is a value of total frequency of all polarity opinion strength for aspect a as depicted in Formula (2).

$$SumF_a = \sum_{s=-3}^{+3} f_s \quad (2)$$

Definition 4: A ***information of aspect*** I_a is a value of total information calculated from each polarity opinion strength s and its frequency for aspect a as depicted in Formula (3).

$$I_a = \sum_{s=-3}^{+3} (s \times f_s) \quad (3)$$

Algorithm: Finding Range

Input: The aspect-based review A (sorted in descending)
Output: The ranges for degrees
eliminate redundantly I_a values in A *// start finding outlier(s)*
dist_{pos} = average of $|I_{a_i} - I_{a_{i+1}}|$ values in a positive part of A
dist_{neg} = average of $|I_{a_i} - I_{a_{i+1}}|$ values in a negative part of A
$\text{index}_{\text{pos}} \leftarrow 0$ *// initialize* $\text{index}_{\text{pos}}$
for $i \leftarrow 1$ to |A| -1 do: *// find positive outlier*
  find the first $|I_{a_i} - I_{a_{i+1}}|$ value > dist_{pos} in positive part
  if the first $|I_{a_i} - I_{a_{i+1}}|$ value > dist_{pos} in positive part is found then
    $\text{index}_{\text{pos}} \leftarrow i$; $\text{outlier}^{+} \leftarrow I_{a_1}$; exit loop *// keep positive outlier*
if $\text{index}_{\text{pos}} \neq 0$ then *// outlier(s) exists in positive part*
  eliminate I_{a_1} to $I_{a_{\text{index}_{\text{pos}}}}$ from A *// remove outlier(s)*
$\text{index}_{\text{neg}} \leftarrow 0$ *// initialize* $\text{index}_{\text{neg}}$
for $i \leftarrow$ |A| -1 to 1 do: *// find negative outlier*
  find the first $|I_{a_i} - I_{a_{i+1}}|$ value > dist_{neg} in negative part
  if the first $|I_{a_i} - I_{a_{i+1}}|$ value > dist_{neg} in negative part is found then
    $\text{index}_{\text{neg}} \leftarrow i$; $\text{outlier}^{-} \leftarrow I_{a_{|A|}}$; exit loop *// keep negative outlier*
if $\text{index}_{\text{neg}} \neq 0$ then *// outlier(s) exists in negative part*
  eliminate $I_{a_{\text{index}_{\text{neg}}}}$ to $I_{a_{|A|}}$ from A *// remove outlier(s)*
find min = the minimum value of I_a in A *// start finding ranges*
find max = the maximum value of I_a in A
if max is positive and min is negative then *// a normal range*
  $\text{range}_{\text{pos}} = \lceil \max/2 \rceil$; $\text{range}_{\text{neg}} = \lceil \min/2 \rceil$
if max is positive and min is positive then *// no negative range*
  $\text{range}_{\text{pos}} = \lceil (\max - \min)/2 \rceil$; $\text{range}_{\text{neg}} = 0$
if max is negative and min is negative then *// no positive range*
  $\text{range}_{\text{pos}} = 0$; $\text{range}_{\text{neg}} = \lceil (\max - \min)/2 \rceil$
$\text{range}_{\text{hate\_very\_much}} = [\text{outlier}^{-}, \min)$; $\text{range}_{\text{hate}} = [\min, \text{range}_{\text{neg}})$; $\text{range}_{\text{dislike}} = [\text{range}_{\text{neg}}, 0)$
$\text{range}_{\text{neither\_like\_nor\_dislike}} = 0$; $\text{range}_{\text{like}} = (0, \text{range}_{\text{pos}}]$; $\text{range}_{\text{love}} = (\text{range}_{\text{pos}}, \max]$
$\text{range}_{\text{love\_very\_much}} = (\max, \text{outlier}^{+}]$
return $\text{range}_{\text{hate\_very\_much}}$; $\text{range}_{\text{hate}}$; $\text{range}_{\text{dislike}}$; $\text{range}_{\text{neither\_like\_nor\_dislike}}$; $\text{range}_{\text{like}}$; $\text{range}_{\text{love}}$; $\text{range}_{\text{love\_very\_much}}$

Fig.4: *Finding ranges for all degrees.*

Definition 5: An ***aspect-based review*** A is a set whose members have a quadruple $< a, \mathrm{F}_a, SumF_a, I_a >$ in the review as shown in Formula (4).

$$\mathrm{A} = \{< a_i, \mathrm{F}_{a_i}, SumF_{a_i}, I_{a_i} >\} \tag{4}$$

where,

i is an index of aspects, $1 \leq i \leq$ m, m is the number of aspects.

3.1.1 Representing aspect-based review procedure

The procedure aims to represent an aspect-based review from datasets that include aspects (features) and their relevant information (polarity and opinion strength). The polarity and opinion strength with frequency can help the organization staff to understand what their customers are thinking about products or services.

Excerpts of the review benchmark datasets are shown in Figure 2. Figure 2 (a) depicts a sample of sentences from cell phone review [13] that are annotated aspects, polarity, and opinion strength. For example, the aspect "*volume*" is annotated in S1, and [-2] means the polarity is negative and the opinion strength is 2. Figure 2 (b) describes a sample of a sentence containing the annotated aspect and polarity from the SemEval-2015 dataset (Task 12) [53]. For example, the sentence id 1004293:5 from the restaurant review, the annotated aspect is "*place*" and its polarity is negative.

The Aspect Review algorithm in Figure 3 is used to represent an aspect-based review from aspects and their relevant information (polarity and opinion strength). Line 1 of Figure 3 is used to initialize the aspect-based review A. Lines 2-12, annotated aspects and their relevant information are extracted. If an extracted aspect a_i is not in A, then a new aspect a_i is added into A, and its F_{a_i}, $SumF_{a_i}$, and I_{a_i} values are equal to $\varnothing$, 0, and 0 for initialization. After that, an opinion strength of the extracted aspect a_i is

Algorithm: Selecting Knowledge

Input: The aspect-based review A; the ranges; the number of interesting aspects n
Output: The aspect-based knowledge K

sort A in descending order using $SumF_a$ *// for selecting aspects having high $SumF_a$ values*
K $\leftarrow \emptyset$ *// initialize the aspect-based knowledge*
for $j \leftarrow 1$ to n do: *// n is a number of aspects in K*
  if I_{a_j} in range$_{hate_very_much}$ then $d_{a_j} \leftarrow$ -3 *// hate very much*
  if I_{a_j} in range$_{hate}$ then $d_{a_j} \leftarrow$ -2 *// hate*
  if I_{a_j} in range$_{dislike}$ then $d_{a_j} \leftarrow$ -1 *// dislike*
  if I_{a_j} in range$_{neither_like_nor_dislike}$ then $d_{a_j} \leftarrow 0$ *// neither like nor dislike*
  if I_{a_j} in range$_{like}$ then $d_{a_j} \leftarrow$ +1 *// like*
  if I_{a_j} in range$_{love}$ then $d_{a_j} \leftarrow$ +2 *// love*
  if I_{a_j} in range$_{love_very_much}$ then $d_{a_j} \leftarrow$ +3 *// love very much*
  add a_j, F_{a_j}, $SumF_{a_j}$, I_{a_j}, d_{a_j} into K
return the aspect-based knowledge K

Fig.5: *Selecting aspect(s) for each degree with the n interesting aspects..*

Table 2: *The Aspect-based Knowledge K is Selected from the Aspect-based Review A with n = 4.*

j	Aspect a_j	F_{a_j}							$SumF_{a_j}$	I_{a_j}	d_{a_j}
		$< f_{-3}$,	f_{-2},	f_{-1},	f_0,	f_{+1},	f_{+2},	f_{+3}¿			
1	volume	< 0,	2,	0,	0,	0,	0,	0 >	2	-4	-2
2	battery life	< 0,	0,	0,	0,	0,	0,	1 >	1	3	+2
3	phone	< 0,	0,	0,	0,	1,	0,	0 >	1	1	+1
4	vibration	< 0,	0,	1,	0,	0,	0,	0 >	1	-1	-1

checked. If a_i has polarity opinion strength s, its f_s value is increased by one. If a_i does not have polarity opinion strength, then its polarity is checked. If a_i polarity is negative/neutral/positive, the respective value of $f_{-1}/f_0/f_{+1}$ is increased by one. Lines 13-15, the values of $SumF_{a_i}$ and I_{a_i} are calculated for each aspect a_i in A. The value of $SumF_{a_i}$ is calculated with Formula (2). The value of I_{a_i} is calculated with Formula (3). Line 16, A is sorted in decreasing order by I_a. Line 17, the algorithm returns the aspect-based review A.

For example, the Aspect Review algorithm is applied to the cell phone review in Figure 2 (a). The result has four tuples and is depicted in Table 1. Four aspects of the result are *battery life, phone, vibration*, and *volume*. The respective values of $SumF_a$ for aspect battery life, phone, vibration, and volume are 1, 1, 1, and 2. The respective values of I_a for aspect battery life, phone, vibration, and volume are 3, 1, -1, and -4.

3.1.2 Selecting aspect-based knowledge procedure

The procedure aims to automatically select knowledge that will be used to generate a summary.

Before selecting the knowledge, the aspect is determined which degree the aspect is in. Therefore, finding ranges play a vital role in determining degree.

Let t be a number of degrees, and $t = 7$.

Let d be a degree value, and d is an integer number in [-3, +3].

Let V be a set of degree value, and V = {-3, -2, -1, 0, +1, +2, +3}.

Let L be a set of degree labels, and L = "hate very much", "hate", "dislike", "neither like nor dislike", "like", "love", "love very much".

Let VL be a relation between V and L, and VL = (-3, "hate very much"), (-2, "hate"), (-1, "dislike"), (0, "neither like nor dislike"), (+1, "like"), (+2, "love"), (+3, "love very much").

Definition 6: An ***aspect-based knowledge*** K is a set whose members have a quintuple $< a, F_a, SumF_a, I_a, d_a >$ as shown in Formula (5).

$$K = \{< a_j, F_{a_j}, SumF_{a_j}, I_{a_j}, d_{a_j} >\} \quad (5)$$

where, j is an index of aspects, $1 \leq j \leq$ n, n is the number of aspects in K, d_{a_j} is a degree value for the aspect a_j.

The Finding Range algorithm in Figure 4 is used to find the ranges for the degrees. Line 1 is used to eliminate redundantly I_a values in A because the aspects may have the same I_a value and we will consider only the distance of each adjacent I_a values. Lines 2-3 are used to calculate an average of $|I_{a_i} - I_{a_{i+1}}|$ values in the positive and negative parts. Lines 4-8 are used to find the first index i in the positive part of A if $|I_{a_i} - I_{a_{i+1}}|$ value > dist$_{pos}$. If the index i is found,

Algorithm: Visual Summary Generation

```
Input: The aspect-based knowledge K, the ranges, the number of aspect ω, options, the relation VL
Output: The visual summary
represent K on the graph with I_a as x-axis and SumF_a as y-axis    // main visual summary graph
determine the degrees among aspects with the ranges on the main visual summary graph by the line
determine ω aspects having high SumF_{a_i} value by symbols    // symbols: circle, triangle, rectangle, etc.
calculate a percentage of each degree in K
show percentage for each degree on the x-axis of the graph
calculate a percentage of each degree in the ω aspects (compare with K)
if an option is "percentage of one aspect"
      represent the percentage of degrees for one aspect on a pie chart
if an option is "percentage of each degree in the ω aspects"
      represent the percentage of each degree in the ω aspects on a pie chart
if an option is "word cloud for each part"
      represent each word cloud for each part (positive and negative parts)
if an option is "percentage of each degree in the ω aspects and K" then
      represent these percentages on a bar chart
if an option is "bubble for K"
      represent a bubble graph with different colors for K (each color for each degree)
return
```

Fig. 6: *Generating the visual summary from the aspect-based knowledge* K *with* ω *aspects.*

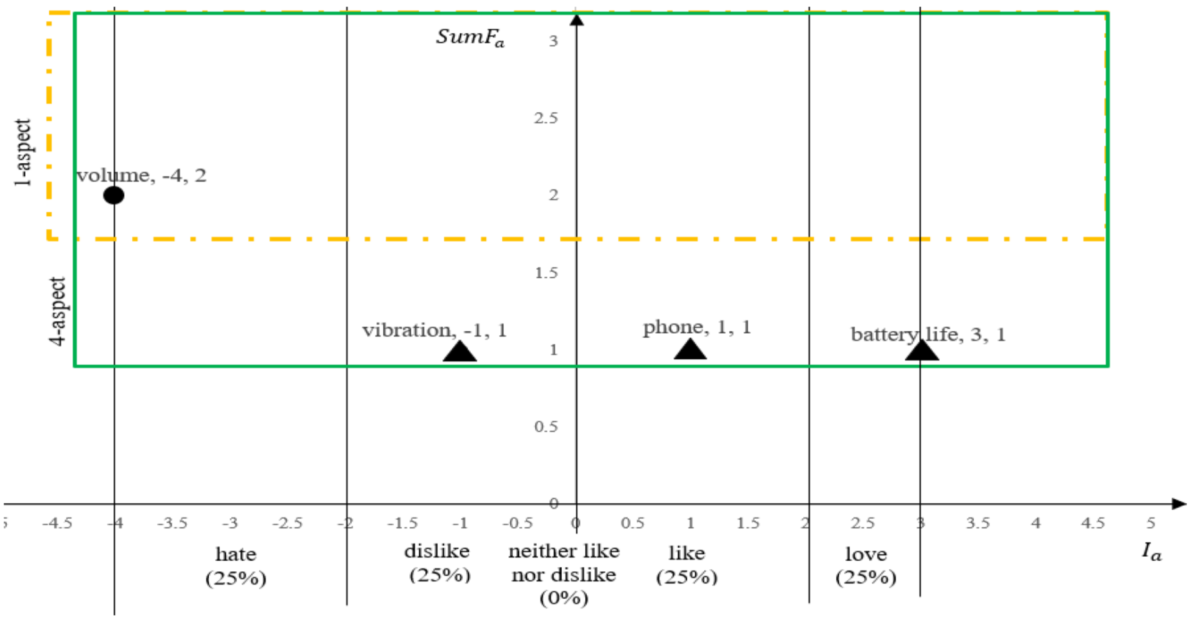


Fig. 7: *The main visual summary graph for knowledge* K *in Table 2 generated by the proposed AAbSS system.*

then the index i is saved and Outlier$^+$ $= I_{a_1}$ for the positive part. Lines 9-10, the outlier(s) in the positive part is removed if the outlier(s) is found. Lines 11-15 are used to find the first index i in the negative part of A if $|I_{a_i} - I_{a_{i+1}}|$ value $>$ dist$_{\text{neg}}$ If the index i is found, then the index i is saved and Outlier$^-$ $= I_{a_{|A|}}$ for the negative part. Lines 16-17, the outlier(s) in the negative part is removed if the outlier(s) is found. Lines 18-19 are used to find the minimum and maximum of I_a in the A. Lines 20-21 of Figure 4, the normal range (i.e. max value is positive, and min value is negative) is determined. Lines 22-23, the case of no negative range (i.e. max and min values are positive) is determined. Lines 24- 25, the case of no positive range (i.e. max and min values are negative) is determined. Lines 26-28 are used to determine ranges. Line 29, the algorithm returns the ranges of degrees. It is noted that the range$_{\text{neither_like_nor_dislike}}$ $= 0$ because it is represented the neutral value.

For example, finding the ranges for the aspect-based review A as shown in Table 1 will be as the following: No outliers; min = -4, max = 3. The result has a normal range because the min value is negative, and the max value is positive. Hence, range$_{\text{pos}}$ - $\lceil \max/2 \rceil = 2$; range$_{\text{neg}}$ - $\lceil \max/2 \rceil$ = -2; Ranges are found that are: range$_{\text{hate}}$ =

"Positive: " *np*
aspect: degree label in L
- "total comment(s): " $SumF_{a_j}$
- < polarity opinion strength " / " percentage of comments "% / " number of comment(s) " comment(s)">

"Neutral: " *ne*
aspect: degree label in L
- "total comment(s): " $SumF_{a_j}$
- < polarity opinion strength " / " percentage of comments "% / " number of comment(s) " comment(s)">

"Negative: " *nn*
aspect: degree label in L
- "total comment(s): " $SumF_{a_j}$
- < polarity opinion strength " / " percentage of comments "% / " number of comment(s) " comment(s)">

Fig.8: *Structured summary template.*

[min,range$_{\text{neg}}$)=[-4, -2); range$_{\text{dislike}}$ =[range$_{\text{neg}}$,0)=[-2,0); range$_{\text{neither_like_nor_dislike}}$ = 0; range$_{\text{like}}$=(0, range$_{\text{pos}}$]=(0, 2]; range$_{\text{love}}$ = (range$_{\text{pos}}$, max]=(2,3]. The other ranges do not exist because of their unsuitable values.

The Selecting Knowledge algorithm in Figure 5 is used to select aspects from A if their I_a values satisfy one of the ranges. The total number of the interesting aspects chosen by the algorithm equals to n and is saved in K. Line 1 is used to sort A in descending by $SumF_a$ Line 2, the aspect-based knowledge K is initialized. Lines 3-11 are used to select n aspectsfrom A whose $SumF_a$ value is high. I_{a_j} value is checked in order to assign degree value for a_j. If the I_{a_j} value belongs to one of the ranges ("hate very much", "hate", "dislike", "neither like nor dislike", "like", "love", "love very much") then d_{a_j} is assigned the respective degree value. After that, the (a_j, F$_{a_j}$, $SumF_{a_j}$, I_{a_j}) from A and d_{a_j} are added into K. Line 12, the selected knowledge saved in K is returned.

For example, with the found ranges in the previous example and n = 4, the aspect-based knowledge K selected from A (Table 1) by the Selecting Knowledge algorithm is presented in Table 2.

3.2 Summary Format Generation Component

The purpose of the Summary Format Generation (SFG) component is to automatically generate a review summary whose format is one of three kinds (visual, structured, and textual). To produce three kinds of format, three procedures are discussed in the next sections.

Let ω be the number of aspects in the summary.

3.2.1 Visual summary generation procedure

The visual summary generation procedure is used to generate a main visual summary that is based on the aspect-based knowledge K and options. The options are 1) "percentage of one aspect", 2) "percentage of the ω aspects", 3) "word cloud for each part", 4) "percentage of the ω aspects and K", and 5) "bubble for K".

The main visual summary is a graph that is based on the knowledge K. The graph has two axes that are I_a as x-axis and $SumF_a$ as y-axis. The I_a axis expresses users' love/hate with aspects. The $SumF_a$ axis expresses a number of comments for aspects. Each point on the graph has three variables $(a_i, I_{a_i}, SumF_{a_i})$ that means the aspect a_i has a level of love/hate from users and total comments from users for this aspect. Furthermore, readers can select other options to observe further information. Those options are: 1) The "percentage of one aspect" option is represented by a pie chart and shows that the percentage of users' comments loved or hated this aspect; 2) The "percentage of the ω aspects" option is represented by a pie chart and shows to the readers a number of degrees and their percentage with the ω aspects; 3) The "word cloud for each part" option is represented by two word-clouds and each shows aspects in each part (positive and negative parts); 4) The "percentage of the ω aspects and K" option is represented by a bar chart and shows a comparison between the ω aspects and K in each degree; 5) The "bubble for K" option is represented by a bubble graph and shows different colors for K (each color for each degree).

The Visual Summary Generation algorithm in Figure 6 is used to generate the visual summary with options. Lines 1-5 are used to represent the knowledge K on the graph. After representing K on the graph, aspects are separated among degrees by lines. Selected aspects are marked with different symbols with other aspects. Then the percentage of each degree in K is calculated and showed on the graph. Line 6 is used to calculate the percentage of each degree in the ω aspects. Lines 7-8, a pie chart shows that the percentage of users' comments loved or hated this aspect if the "percentage of one aspect" option is chosen. Lines 9-10 of Figure 6, the pie chart represents a number of degrees and their percentage with the ω aspects if the "*percentage of each degree in the ω as-*

```
Algorithm: Structured Summary Generation
Input: The aspect-based knowledge K, the number of aspect ω, the relation VL
Output: The structured summary
  np ← 0; ne ← 0; nn ← 0          // the number of aspects for each polarity (positive, neutral, negative)
  Sp ← ∅; Se ← ∅; Sn ← ∅          // aspects and their information for each polarity (positive, neutral, negative)
  for j ← 1 to ω do:               // ω is the number of aspects from K to generate summary
    tmp ← “”                       // tmp is used to save information of an aspect
    add aspect a_j into tmp        // tmp ← aspect “:”
    add degree label into tmp after retrieving it for a_j using d_{a_j} and VL   // tmp ← degree label \n
    add SumF_{a_j} as total comment into tmp      // tmp ← “total comment(s):” SumF_{a_j} \n
    add percentages, number of comments into tmp after calculating a percentage
        // tmp ← polarity opinion strength s “/” percentage of comments “% / ” number of comments “comment(s)”
    if a_j is positive then
        add tmp into Sp; np ← np + 1      // Sp is in a positive part
     if a_j is neutral then
        add tmp into Se; ne ← ne + 1      // Se is in a neutral part
    if a_j is negative then
        add tmp into Sn; nn ← nn + 1      // Sn is in a negative part
   if Sp ≠ ∅ then show structure for a positive part with Sp, np
   if Se ≠ ∅ then show structure for a neutral part with Se, ne
   if Sn ≠ ∅ then show structure for a negative part with Sn, nn
  return
```

Fig.9: *Generating the structured summary from the aspect-based knowledge* K *with* ω *aspects.*

Table 3: *Lexicons for Degree Labels.*

Degree Label	Lexicons
“love”	“loved”, “felt great”, “felt awesome”,...
“like”	“liked”, “felt good”, “felt satisfactory”, “felt fine”, “felt cool”,...
“neither like nor dislike”	“confused”, “neutral”, “uncertain”,...
“dislike”	“disliked”, “did not like”, “felt not good”, “felt unsatisfactory”,...
“hate”	“hated”, “detested”, “loathed”, “abhorred”,...

Table 4: *Sentences Template and Connecting Words to Generate Textual Summary.*

Category name	Sentences Template and Connecting Words
oS	“Most of the reviewers loved **(hated)** so much about **outlier_ aspect.**”; “The **outlier_ aspect** was **(were)** the best **(worst)** in the review.”; “Most of the users commented that they loved **(hated)** very much on the **outlier_aspect.**”,...
nS	“There were reviews that **degree_lexicon** with **list_aspect.**”; “There was only one **(were) list_aspect** that the users **degree_lexicon.**” ; “The **list_aspect** was **(were) degree_lexicon** by the users.” ; “The users who commented **degree_lexicon** only one **(some)** about **list_aspect.**”;...
pCW	“Moreover”, “Furthermore”, “In addition”, “Also”, “And”,...
nCW	“However”, “Otherwise”, “On the other hand”, “On the contrary”, “Nevertheless”,...

pects” option is chosen. Lines 11-12, two clouds show aspects based on their I_a (one cloud for positive and one for negative) if the “word cloud for each part” option is selected. Lines 13-14, a bar chart shows a comparison between the ω aspects and the knowledge K in each degree if the “*percentage of each degree in the* ω *aspects and* K ” option is chosen. Lines 15-16, a bubble graph shows aspects of each degree in different colors if the “*bubble for* K” option is chosen. Line 17, the algorithm returns the visual summary.

For example, the Visual Summary Generation algorithm is applied to the knowledge K (Table 2) with $\omega = 1$ and $\omega = 4$ aspect(s), and the result is depicted in Figure 7. With $\omega = 1$, a selected aspect is one circle dot and has a “hate” degree. With $\omega = 4$, selected aspects are one circle dot and three triangle dots. Note that there is no option in this example.

3.2.2 Structured summary generation procedure

The structured summary generation procedure is used to generate a structured summary that lets readers have an overview of the ω aspects. In the ω aspects having the highest comments, the summary depicts a number of aspects in positive comments or negative comments. Moreover, the readers also know the aspects are interesting or not from customers via

Algorithm: Textual Summary Generation

Input: The aspect-based knowledge K, the number of degrees t, the number of aspects ω, the relation VL
Output: The textual summary

```
tS ← Ø                                  // new sentences
wS ← “”                                 // connecting word
for i ← t to 1 do:                      // t is the number of degrees
    tmp_aspect ← “”                     // aspects
    for j ← 1 to ω do:                  // ω is the number of aspects in summary
        if d_aj = d_i then              // d_aj is degree value of aspect a_j in K; d_i is degree value in the relation VL
            add aspect a_j into tmp_aspect
    if tmp_aspect ≠ “” then
        if (d_i = +3) or (d_i = -3) then    // tmp_aspect is outlier_aspect in “love very much”/“hate very much”degree
            randomly select a sentence from oS, change tense and word in a bracket, add into tS
        else                                // tmp_aspect is list_aspect
            randomly select a sentence from nS, change tense and word in a bracket, add into tS
if exists positive part and negative part in ω aspects then     // choosing a connecting word
    if a number of sentences in a negative part are less than a number of sentences in a positive part then
        randomly select one connecting word from nCW and add into wS
    else
        randomly select one connecting word from pCW and add into wS
generate the summary from tS and wS     //generating summary by combining tS and wS
return
```

Fig. 10: *Generating the textual summary from the aspect-based knowledge* K *with* ω *aspects.*

frequencies of the aspect, a level of love/hate of the aspect from customers via the opinion strength, and percentage of like/dislike among users' comments for each aspect.

The template to generate the structured summary is proposed and depicted in Figure 8. The template has three parts of polarity that are positive, neutral, and negative. In the template, np is the total number of aspects in three degrees (“love very much”, “love”, and “like”), ne is the total number of aspects in “neither like nor dislike” degree, and nn describes the total number of aspects in three degrees (“hate very much”, “hate”, and “dislike”). At each polarity, aspects and their relevant information are in detail. The relevant information includes degree label, total comments from users, polarity opinion strength, percentage of comments, and the number of comments for polarity opinion strength.

To generate the structured summary, the Structured Summary Generation algorithm is proposed and described in Figure 9. Line 1 is used to initialize three variables that are used to save the total number of aspects for each polarity (np is the number of positive aspects, and ne is the number of neutral aspects, nn is the number of negative aspects). Line 2 is initialized three variables that are used to save aspects with the relevant information for each polarity (Sp is a positive part, Se is a neutral part, and Sn is a negative part). Lines 3-14 of Figure 9, the algorithm gets the relevant information for each aspect in each polarity and keeps it. At each aspect, five information (aspect term, degree label retrieved from VL and d_{a_j}, $SumF_{a_j}$ as a total comment, percentages by calculating the percentage of each polarity opinion strength, and number of comments for each polarity opinion strength) is used to generate a sentence for aspect a_j. After that, the sentence is saved on a temporary variable tmp. Aspect is then checked and saved in the respective polarity part. If aspect a_j is in the algorithm in Figure 9 adds tmp into Sp and increases np by one. If aspect a_j is in negative, the algorithm adds tmp into Sn and increases nn by one. Lines 15-17 are used to show aspects and the relevant information for each polarity if aspects in that polarity are not empty. Line 18, the structured summary is returned.

3.2.3 Textual summary generation procedure

The textual summary generation procedure is used to generate new sentences including aspects in a textual summary. To have a coherence among sentences, the Natural Language Generation (NLG) is applied in this procedure. Hence, lexicons and templates for sentences are defined. The lexicons for degree labels are defined and presented in Table 3 (the first column is degree label; the last column describes lexicons). The template for sentences (oS for an outstanding case, nS for a normal case) and the set of connecting words to connect between two sentences (pCW for positive, nCW for negative) are suggested and presented in Table 4. The first column of Table 4 is each category name: oS, nS, pCW, nCW; the second column depicts sets of sentences and connecting words. Note that words in a bracket and bold of the template will be flexibly chosen by the algorithm.

The **outlier_aspect** is aspects of “love very much”

Table 5: *The Result of Polarities and Selected Aspects from the Proposed AAbSS System with 10 Aspects.*

Dataset	Domain	Number of sentences	Polarity: Negative -3	-2	-1	Neutral 0	Positive +1	+2	+3	Aspects from the AAbSS system (ω = 10)	Time (second)
Customer product reviews	Cell phone	546	5	51	30	0	52	152	49	phone, size, speakerphone, radio, battery life, screen, feature, weight, reception, sound quality	0.879
	Nikon digital camera	346	1	21	9	0	20	103	49	camera, picture, picture quality, size, use, feature, scene mode, battery life, print, software	0.965
	Canon digital camera	597	7	39	15	0	43	120	62	camera, picture, viewfinder, use, picture quality, photo, control, software, feature, g3	0.584
	Mp3 player	1,716	78	172	81	0	67	301	146	player, software, price, battery, size, sound quality, case, storage, sound, navigation	0.929
	DVD player	740	90	129	19	0	30	128	37	player, dvd, play, customer service, picture, dvd player, remote, look, format, apex	1.070
SemEval-2015 (Task 12)	Restaurant	1,315	0	0	278	36	926	0	0	food, service, place, restaurant, staff, pizza, atmosphere, Sushi, decor, ambience	3.402
Additional product reviews	Canon SD500	229	1	11	15	0	24	44	53	camera, picture, LCD, image, size, flash, SD500, quality, pics, video	1.377
	Canon S100	300	4	29	22	0	40	72	51	camera, size, picture, small, software, battery, use, ease of use, s100, zoom	1.037
	Diaper Champ	375	2	36	18	0	14	71	98	diaper champ, odor, champ, use, product, diaper pail, refill, work, smell, pail	0.598
	Hitachi Router	312	1	7	71	0	89	69	28	router, price, power, adjustment, control, speed, use, collet, heavy, money	0.543
	Nokia6600	554	68	64	27	0	24	75	206	phone, bluetooth, screen, camera, feature, battery life, 6600, Nokia, interface, ringtone	0.733
	Norton	380	70	69	28	0	18	31	21	product, support, Norton, installation, install, symantec, antivirus, uninstall, firewall, update	0.681
	MicroMp3	1,011	57	57	40	0	60	227	110	player, sound, look, size, feature, software, storage, design, sound quality, touchpad	0.773
	iPod	531	12	30	20	0	20	90	19	iPod, battery, iTune, design, sound, use, capacity, Apple, interface, feature	1.104
	Linksys Router	577	10	40	13	0	10	89	56	router, setup, work, product, speed, CD, installation, security, instruction, connection	0.640

or "hate very much" degree. The **list_aspect** is aspects that are in "hate", "dislike", "neither like nor like", "like", or "love" degree. The **degree_lexicon** is lexicons for each degree that are in Table 3 randomly chosen by the algorithm.

The Textual Summary Generation algorithm in Figure 10 is to generate the textual summary. Lines 1-2 are used to initialize tS and wS to keep new sentences and connecting words, respectively. Lines 3-12, the algorithm generates sentences in summary for each degree value in the relation VL. At each degree value d_i in the relation VL, a degree value of aspect a_j is checked. If its degree value d_{a_j} is the same as d_i, then aspect a_j is added into temporary variable *tmp_aspect*. The sentence of each degree value is randomly selected from the template (outstanding case and normal case) if the *tmp_aspect* is not empty. After selecting the sentence, parameters in the sentence are suitably changed such as aspect terms, verb, etc. and added into tS. Lines 13-17 are used to select one connecting word.

If a number of sentences in a negative part are less than a number of sentences in a positive part, then the connecting word is chosen from nCW. Otherwise, the connecting word is chosen from pCW. Line 18 of Figure 10 is used to generate the summary

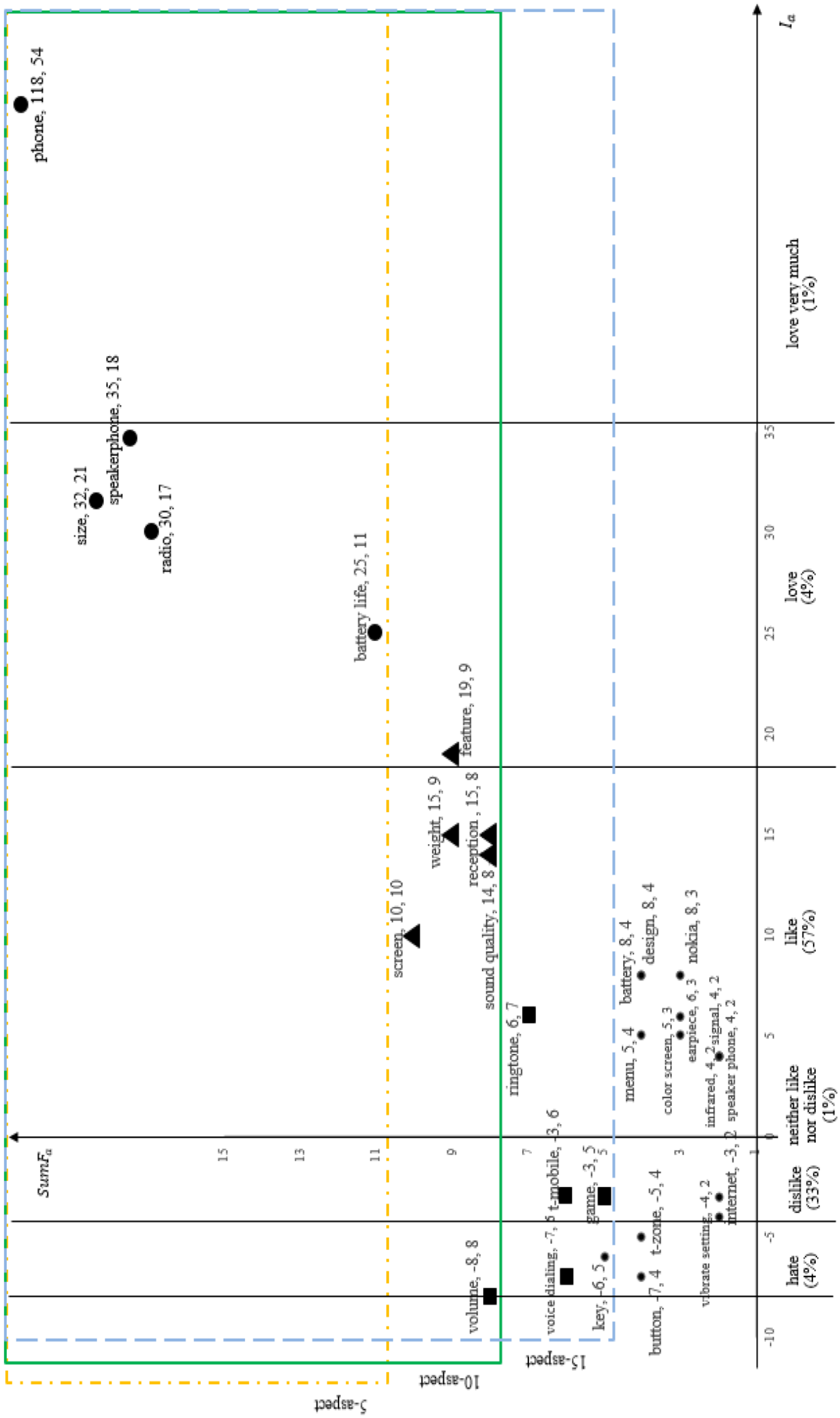


Fig.11: *The main visual summary graph of the cell phone review generated by the proposed AAbSS system.*

by combining tS and wS. Line 19, the textual summary is returned. Note that the **outlier_aspect** in the template is replaced with *tmp_aspect* if aspects are in "love very much" or "hate very much" degree. The **list_aspect** in the template is replaced with *tmp_aspect* if aspects are in "hate", "dislike", "neither like nor like", "like", or "love" degree.

A complexity time of the AAbSS system is O(m log m), where m is the number of aspects.

4. RESULTS AND DISCUSSION

4.1 Dataset

Three benchmark datasets about customer reviews are used in this study. The first dataset published by Hu and Liu [13] had five domains (Cell phone, Nikon digital camera, Canon digital camera, Mp3 player, and DVD player). The second one was released for Task 12 (Task 12 for aspect-level sentiment analysis) of the International Workshop on Semantic Evaluation 2015 (SemEval-2015) [53] and the Restaurant domain was used. The last dataset was released by Liu et al. [54] and nine domains (Canon SD500, Canon S100, Diaper Champ, Hitachi Router, Nokia, Norton, Mp3, iPod, and Linksys Router) were used. A description of the three review datasets is summarized in the first three columns of Table 5 (dataset, domains, and a number of sentences).

4.2 Data Pre-processing

There are some irregular annotations for aspects and their polarities in a dataset. To extract these aspects and polarity, the pre-processing procedure is required. In this study, two phases (before and after extracting annotated aspects) are considered.

The "before extracting annotated aspect" phase, some irregular annotations in sentences of the two datasets (Customer product reviews and Additional product reviews) are edited (e.g, removed/replaced/added). These irregular annotations are edited and described as the following:

- Remove the annotated aspects and their polarity if no opinion strength or polarity is annotated such as "*player [+]*", "look [2]".
- Add [in front of the annotated polarity if there is only one in the annotated polarity and opinion strength, e.g., the aspect *connection+3]*" becomes "*connection[+3]*".
- Add ## or # in the annotated sentence if this symbol is missed (## is used to separate the annotation and text). For example, "*volume control[-1] Weak...* becomes "*volume control[-1]## Weak...*"
- Replace { or } of the annotated polarity by [or] such as the aspect "*design[+3]*" becomes "*design[+3]*".

The "after extracting annotated aspect" phase, this phase is applied to the three datasets. The extracted aspects in the form of uppercase or lowercase letters are considered as one aspect. Furthermore, the extracted aspects in singular and plural nouns are also considered as one aspect. For example, the extracted aspect "*LCD*" and "*lcd*" are considered as one aspect. The extracted aspect "*batteries*" and "*battery*" are considered as one aspect with "*battery*".

4.3 Experimental Results

A number of polarities (negative/neutral/positive column) for each domain of the three benchmark datasets with 10 selected aspects are introduced in Table 5. In our experiments, the number of aspect-based knowledge n is equal to 100, and the number of aspects to be in a summary ω equals to 5, 10, and 15. Table 5 shows ω equals to 10.

Consuming time of the AAbSS system is by calculating the time-consuming process of constituents in the AAbSS system. The constituents include 1) pre-processing, 2) presenting for aspect-based review, 3) selecting knowledge including finding ranges, 4) generating a main visual summary, 5) generating a structured summary, and 6) generating a textual summary. The consuming time measured in second for each domain presented in the last column of Table 5.

From the three benchmark datasets, the cell phone review is discussed in details of three kinds of summary formats (visual, structured, and textual). First, the AAbSS system represents reviews based on aspects, their polarities, and opinion strength using the Aspect Review algorithm. The result is the aspect-based review A. After that, a range is determined by the I_a using the Finding Range algorithm. There is one outlier for a positive part in the cell phone review. The min and max values are -8 and 35, respectively. With the min and max values, this review has a normal range. These ranges are rangehate = [-8, -4),

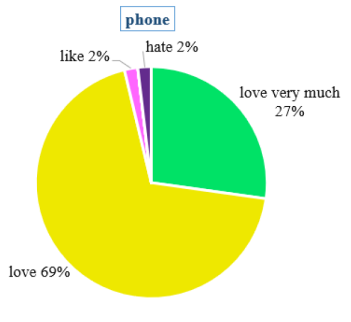


Fig.12: *A "percentage of one aspect" option for "phone" aspect in the cell phone review from the proposed AAbSS system.*

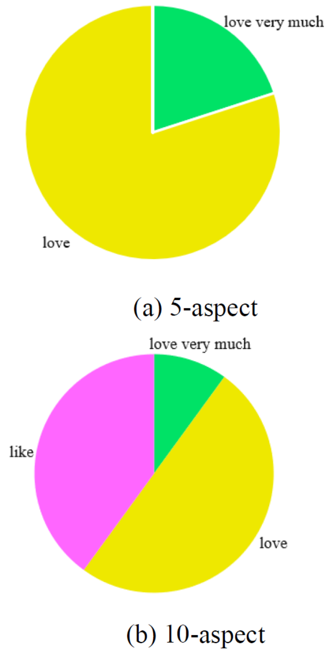


Fig. 13: *A "percentage of each degree in the 5 and 10 aspects" option for the cell phone review from the proposed AAbSS system.*

$\text{range}_{\text{dislike}}$ = [-4, 0), $\text{range}_{\text{neither\_like\_nor\_dislike}}$ = 0, $\text{range}_{\text{like}}$ = (0, 18], $\text{range}_{\text{love}}$ = (18, 35], and $\text{range}_{\text{love\_very\_much}}$ = (35, 118]. The degree of summarization or feeling is equal to six in the cell phone review because the "hate very much" degree is empty. Then the aspect-based knowledge K is selected from A with n = 100 using the Selecting Knowledge algorithm. Finally, three kinds of output are generated by the algorithms and described as the following:

The first kind of output is a visual summary. The main visual summary graph for the cell phone review in Figure 11 is generated by the Visual Summary Generation algorithm. Figure 11 is the main result of the visual summary. With $\omega = 5$, selected aspects are big circle dots. Note that 5 selected aspects in Figure 11 are in "love very much" and "love" degree. With $\omega = 10$, selected aspects are five big circle dots and five triangle dots. With $\omega = 15$, selected aspects are the previous ten aspects and five rectangle dots. Figure 12 is shown the "percentage of one aspect" option for the aspect "phone". This figure lets the readers quickly observe the aspect "phone" that has four degrees ("love very much" 27%, "love" 69%, "like" 2%, and "hate" 2%) with their percentages. Figure 13 is shown a "percentage of each degree in the ω aspects" option. Figure 13 (a) shows with 5 aspects ("love very much", "love" degree). Figure 13 (b) shows with 10 aspects ("love very much", "love", and "like" degree). Figure 14 is shown a "word cloud for each part" option. Figure 14 (a) shows aspects of the positive part. Figure 14 (b) shows aspects of the negative part. Figure 15 is shown if the readers want to compare the percentage of each degree between ω aspects and the knowledge K. In addition, the readers can observe that the "hate very much" degree does not exist in the cell phone review. In the knowledge K, the most interesting degree is "like" that has 60%. Figure 16 is shown a "bubble for K" option. This figure shows that each degree has one color, and the size of the bubble is depending on frequencies of polarity strength. The comparison of the visual summary generated by the proposed AAbSS system and existing system [13] for the Canon digital camera review is depicted in Figure 17.

The second kind of output is a structured summary. The structured summaries ω = 5/10/15 in Table 6 are generated using the Structured Summary Generation algorithm. With $\omega = 5$, the positive polarity has 5 aspects that are "phone", "size", "speakerphone", "radio", and "battery life". From the structured summary, the "phone" belongs to the "love very much" degree. The comments are 27% of comments with polarity opinion strength +3, 69% of comments with polarity opinion strength +2, 2% of comments with polarity opinion strength +1, and 2% of comments with polarity opinion strength -2. If ω = 15, the positive aspects are 11 and the negative aspects are 4. The comparison of a structured summary generated for the Canon digital camera by the proposed AAbSS system with the existing system is shown in Figure 18. The structured summary reproduced from Hu and Liu (2004) by López Condori and Salgueiro Pardo [5] showed an example from the original sentences. Note that the summary generated by the proposed AAbSS system does not include users' comments.

The third kind of output is a textual summary. The textual summary for the cell phone review is generated using the Textual Summary Generation algorithm. The first result of this algorithm is the aspects in each degree shown in Table 7. After that, the algorithm generates textual summaries based on the aspects with their degrees. The textual summaries generated by the proposed AAbSS system are shown in Table 8 from two reviews (Cell phone and Restaurant). Table 9 shows the comparison of a textual summary generated by the proposed AAbSS system with the MEAD* system [25, 29] and the rhetorical AHT system [29]. Note that some aspects are the same as the proposed AAbSS system.

The comparison of the AAbSS system with the online commercial tools in terms of format types for summaries is described in Table 10. In this table, the first four columns are the name of tools, type of input, real-time search, and retrieved sources from the internet. The rest of the columns are three types of output (visual/structured/textual). In each type, the column Sentiment shows readers whether the output has sentiment or not, and the column Description presents outputs in brief. In the visual output, the column Type of chart what kind of visualization

(a) aspects of the positive part (b) aspects of the negative part

Fig.14: *A "word cloud for each part" option for the cell phone review from the proposed AAbSS system.*

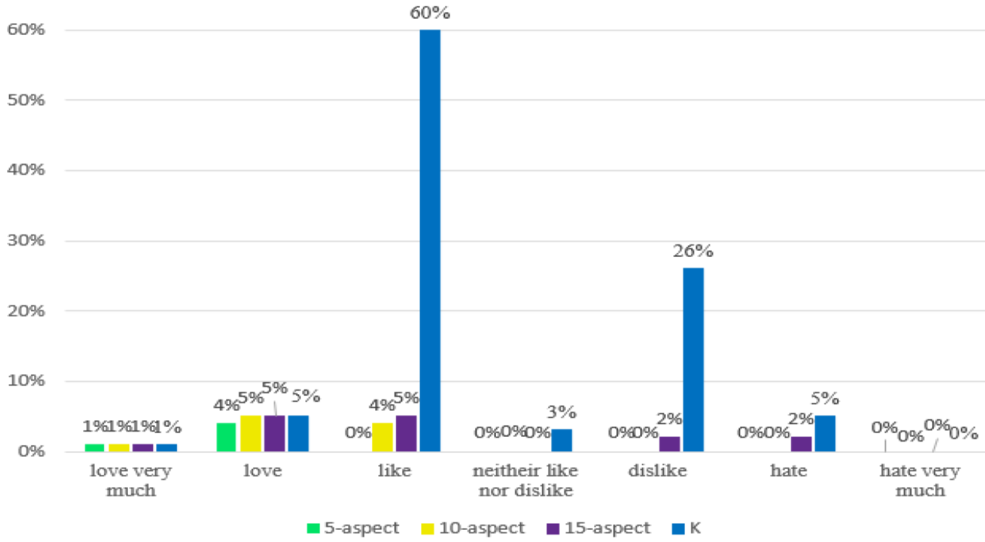


Fig.15: *A "percentage of each degree in the 5/10/15 aspects and* K*" option for the cell phone review from the proposed AAbSS system.*

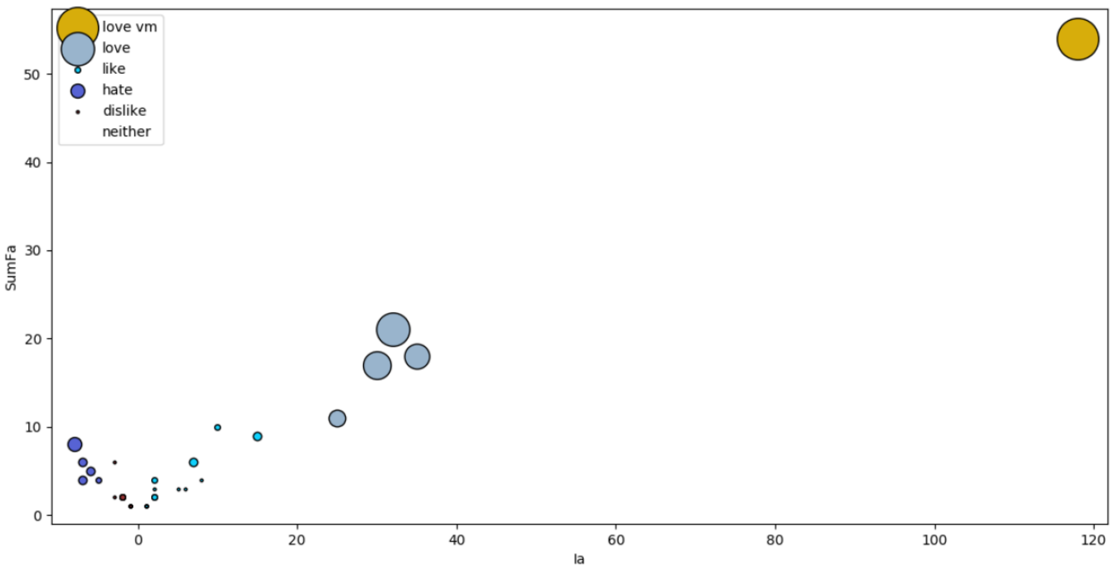
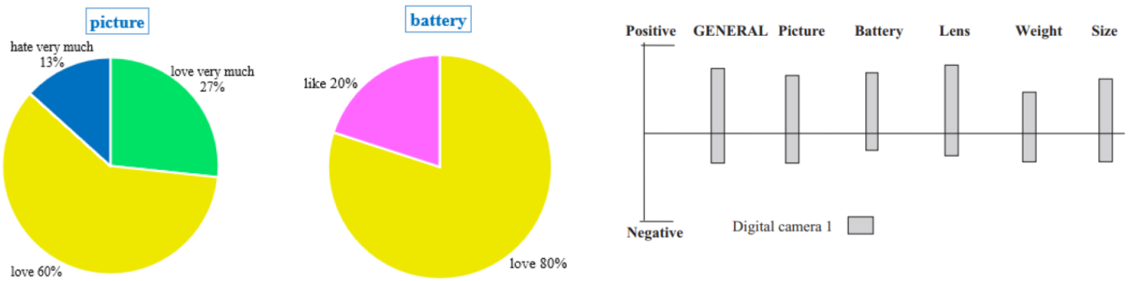


Fig.16: *A "bubble for* K*" option for the cell phone review from the proposed AAbSS system.*

each tool has. Lexalytics and the proposed AAbSS system has three kinds of outputs. The AAbSS system can provide various perspectives when compared with the visual output. The textual output of Lex-



(a) The visual summary generated by the proposed AAbSS system for each aspect, e.g., picture and battery

(b) The visual summary in [13]

Fig.17: *The example of the visual summary for the Canon digital camera review.*

Table 6: *The Structured Summaries Generated by the AAbSS System for the Cell Phone Review with ω Aspects.*

	The Structured Summary	
$\omega = 5$	$\omega = 10$	$\omega = 15$
Positive: 5 ***phone: love very much*** • total comment(s): 54 • +3 / 27% / 15 comment(s) • +2 / 69% / 37 comment(s) • +1 / 2% / 1 comment(s) • -2/ 2% / 1 comment(s) ***size: love*** • total comment(s): 21 • +3 / 5% / 1 comment(s) • +2 / 57% / 12 comment(s) • +1 / 33% / 7 comment(s) • -2/ 5% / 1 comment(s) ***speakerphone: love*** • total comment(s): 18 • +3 / 27% / 5 comment(s) • +2 / 50% / 9 comment(s) • +1 / 17% / 3 comment(s) • -1/ 6% / 1 comment(s) ***radio: love*** • total comment(s): 17 • +3 / 23% / 4 comment(s) • +2 / 59% / 10 comment(s) • +1 / 6% / 1 comment(s) • -1/ 6% / 1 comment(s) • -2/ 6% / 1 comment(s) ***battery life: love*** • total comment(s): 11 • +3 / 36% / 4 comment(s) • +2 / 55% / 6 comment(s) • +1 / 9% / 1 comment(s)	**Positive: 10** *// Same as ω = 5* ... ***screen: like*** • total comment(s): 10 • +2 / 60% / 6 comment(s) • +1 / 20% / 2 comment(s) • -2/ 20% / 2 comment(s) ***feature: love*** • total comment(s): 9 • +3 / 11% / 1 comment(s) • +2 / 89% / 8 comment(s) ***weight: like*** • total comment(s): 9 • +2 / 67% / 6 comment(s) • +1 / 33% / 3 comment(s) ***reception: like*** • total comment(s): 8 • +3 / 37% / 3 comment(s) • +2 / 50% / 4 comment(s) • -2/ 13% / 1 comment(s) ***sound quality: like*** • total comment(s): 8 • +3 / 24% / 2 comment(s) • +2 / 63% / 5 comment(s) • -2/ 13% / 1 comment(s)	**Positive: 11** *// Same as ω = 10* ... ***ringtone: like*** • total comment(s): 6 • +3 / 17% / 1 comment(s) • +2 / 17% / 1 comment(s) • +1 / 50% / 3 comment(s) • -1/ 16% / 1 comment(s) **Negative: 4** ***volume: hate*** • total comment(s): 8 • +2 / 13% / 1 comment(s) • +1 / 13% / 1 comment(s) • -1/ 13% / 1 comment(s) • -2/ 61% / 5 comment(s) ***t-mobile: hate*** • total comment(s): 6 • +2 / 33% / 2 comment(s) • -1/ 17% / 1 comment(s) • -2/ 50% / 3 comment(s) ***voice dialing: hate*** • total comment(s): 6 • +1 / 17% / 1 comment(s) • -1/ 33% / 2 comment(s) • -2/ 50% / 3 comment(s) ***game: hate*** • total comment(s): 5 • +2 / 20% / 1 comment(s) • +1/ 20% / 1 comment(s) • -2/ 60% / 3 comment(s)

alytics is generated by selecting sentences from an input. Meanwhile, the proposed AAbSS system generates the textual output by producing new sentences. The AAbSS system and Lexalytics also concern with the sentiment. The other tools have two kinds of outputs or only one kind of output. Sentiment Viz and Social Searcher have two kinds of output and concern with the sentiment. The textual output of Sentiment Viz tool is also generated by selecting sentences from retrieved results of an input. In addition, these two tools can represent only keyword(s) on their visual outputs when the visual output of the proposed AAbSS system represents all interesting aspects.

With any annotated dataset having aspects and polarity, the proposed AAbSS system for generating summaries performs well in terms of three outputs. The visual summaries provide various perspectives of aspects via the main visual, pie, bar, bubble, and word cloud as shown in Figure 11 to Figure 17. The structured summary illustrates three major groups

Positive: 73
camera: love very much
…
picture: love
- total comment(s): 15
- +3 / 27% / 4 comment(s)
- +2 / 60% / 9 comment(s)
- -3 / 13% / 2 comment(s)

…
Negative: 24
viewfinder: hate
- total comment(s): 12
- +1/ 9% / 1 comment(s)
- -1 / 33% / 4 comment(s)
- -2/ 58% / 7 comment(s)

…

(a) The structured summary generated by the proposed AAbSS system

Feature: **picture**
Positive: 12
- Overall this is a good camera with a really good picture clarity.
- The pictures are absolutely amazing - the camera captures the minutest of details.
- After nearly 800 pictures I have found that this camera takes incredible pictures.

…
Negative: 2
- The pictures come out hazy if your hands shake even for a moment during the entire process of taking a picture.
- Focusing on a display rack about 20 feet away in a brightly lit room during day time, pictures produced by this camera were blurry and in a shade of orange

(b) The structured summary reproduced from Hu and Liu 2004 by López Condori and Salgueiro Pardo [5]

Fig.18: *Comparison between the AAbSS and existing system for the Canon digital camera.*

Table 7: *Selected Aspect(s) for each Degree with Parameter ω on the Cell Phone Review.*

ω	**Aspect(s) in each degree**						
	hate very much	*hate*	*dislike*	*neither like nor dislike*	*like*	*love*	*love very much*
5	-	-	-	-	-	"size", "speakerphone", "radio", "battery life"	"phone"
10	-	-	-	-	"screen", "weight", "reception", "sound quality"	"size", "speakerphone", "radio", "battery life", "feature"	"phone"
15	-	"volume", "voice dialing"	"t-mobile", "game"	-	"screen", "weight", "reception", "sound quality", "ringtone"	"size", "speakerphone", "radio", "battery life", "feature"	"phone"

Table 8: *The Textual Summaries Generated by the Proposed AAbSS System for Two Reviews.*

Review	**No. selected Aspects**	**Textual Summary**
Cell phone	$\omega = 5$	The phone was the best in the review. The size, speakerphone, radio, and battery life of the product were felt great by the users.
	$\omega = 10$	Most of the users commented that they loved very much on the phone. The size, speakerphone, radio, battery life, and feature were felt good. The screen, weight, reception, and sound quality were liked by the users.
	$\omega = 15$	Most of the comments from the users loved so much on the phone. The size, speakerphone, radio, battery life, and feature were felt cool. The users liked the screen, weight, reception, sound quality, and ringtone. However, the users dislike the t-mobile and game. The users hated the volume and voice dialing.
Restaurant	$\omega = 5$	Most of the reviewers loved so much about food. The users felt awesome some features on service and place. The restaurant and staff were liked by the users.
	$\omega = 10$	Most of the users commented that they loved very much on the food. The users loved some features on service and place. The users felt cool some features on restaurant, staff, pizza, atmosphere, Sushi, and decor. On the contrary, the users hated the menu.
	$\omega = 15$	The most interesting from customers was food. There were service and place that the users felt great. There were reviews that liked with restaurant, staff, pizza, atmosphere, Sushi, decor, ambience, dishes, and portions. On the other hand, the waiter was felt not cool by the users. The users hated the menu, fish.

Table 9: *Comparison between Existing Systems and the Proposed AAbSS System for Nikon Digital Camera.*

System	Aspect Comparing: *The same*	Aspect Comparing: *not the same*	Textual Summary with 10 aspects
Proposed AAbSS system	*camera, picture, picture quality, feature, size, scene mode, use, software, battery life*	print	Most of the reviewers loved very much about camera. The picture, picture quality, size, and feature were loved by the users. The users who commented felt liked some features on use, scene mode, battery life, print, and software.
MEAD* [25, 29]	*camera, picture quality, feature, use, battery life*	controls, picture clarity, close-up shooting, lcd, price	It is very compact but the controls are so well designed that they 're still easy to use . It 's easy for beginners to use , but has features that more serious photographers will love , so it 's an excellent camera to grow into. But overall this is a good camera with a ' really good ' picture clarity ; an exceptional close-up shooting capability . The battery life is very good , i got about 90 minutes with the lcd turned on all the time , the first time around , and i have been using it with the lcd off every now and then , and have yet needed to recharge it . Yes , the picture quality and features which are too numerous to mention are unmatched for any camera in this price range.
Rhetorical AHT [29]	*camera, picture, picture quality, size, scene mode, use, software*	manual mode, auto mode	All reviewers (34 people), who commented on the camera, felt that it was really good mainly because of the picture. Around 26% of the reviewers expressed their opinion about the picture quality and they really liked it. Around 24% of the reviewers noted the use and they thought that it was satisfactory. Talking about the use, around 24% of the reviewers expressed their opinion about the size and they felt that it was fine. Only 6 reviewers commented about the scene mode and in overall they thought that it was satisfactory. Moreover, regarding the scene mode, 4 shoppers mentioned about the manual mode and they thought that it was satisfactory, and similarly only 4 reviewers commented about the auto mode and in overall they did not express any strong positive or negative opinion about it. Only 4 costumers mentioned the software and they felt that it was really good.

(positive, neutral, and negative) as shown in Table 6 and Figure 18. The compact textual summary is generated by using new sentences as shown in Table 8 and Table 9. Furthermore, the proposed AAbSS system allows users to update summaries by adding new reviews.

The limitation of the proposed AAbSS system is the annotated input with aspects and polarity, and it is not a real-time search.

In future work, we plan to build one component that helps to apply to the AAbSS system for the datasets without annotation. Furthermore, we also plan to add options into the AAbSS system for selecting sources that are 1) provided by users (a current function) or 2) retrieved from the internet.

5. CONCLUSION

In this work, the Aspect-based Automatic Sentiment Summarization (AAbSS) system is proposed. With any dataset annotated aspect and polarity as an input, the proposed AAbSS system can process and keep into the aspect-based knowledge. It is easy to update a generated summary by adding new reviews from the same domain into the aspect-based knowledge. From the aspect-based knowledge, the proposed AAbSS system can generate three kinds of output format without consuming time to build any tree or accessing a dataset to extract sentences. Hence, our proposed AAbSS system not only fast generates the summaries but also does not spend memory capacity to save any raw data. The proposed system only keeps the small value of the aspect-based knowledge. To automatically generate efficient, effective, and useful summaries, the new representations on a chart for a visual summary, the new template for a structured summary, and the new method of applied natural language generation for a textual summary are proposed. The visual summary not only shows the different magnitude of polarities for aspects but also specifies the magnitude of polarity for each aspect of the top interesting aspects. The structured summary depicts the most interesting aspects with the percentage of comments for each degree. The textual summary is generated by new sentences without consuming time to build any tree. In the experimental results, the user can select the number of inter-

Fig.19: *Comparison between the online commercial tools and the Proposed AAbSS System.*

No.	Tool	Input	Real time	Online source	Visual output			Structured output		Textual output	
					Senti-ment	Type of chart	Description	Senti-ment	Description	Senti-ment	Description
1	Awario [32]	Keyword	Yes	5 sources[1]	Yes	Donut, line, cloud	- Donut, line charts show the percentage of positive, negative	No	No	No	No
2	Hootsuite Insights [33, 34]	Keyword	Yes	Social media[2]	Yes	Line, bar, donut, cloud	- Line, donut, bar charts show the percentage of positive, negative, neutral	No	No	No	No
3	Sentiment Viz [35, 36]	Keyword	Yes	Tweets from the past week	Yes	Scatter, bar, cloud	- Graph with 2 axes, each point is a tweet with positive, negative - Bar chart shows a number of tweets posted at different times with their sentiment. - Cloud shows for each sentiment	No	No	Yes	- Select 20 interesting tweets
4	Social mention [37, 38]	Keyword	Yes	100+ social platform	Yes	Bar	- Percentage of positive, negative, neutral	No	No	No	No
5	Social Searcher [39 - 41]	Keyword	Yes	11 sources[3]	Yes	Donut, bar, timeline	- Donut graphs show the percentage of positive, negative, neutral for each source	Yes	- Group popular posts from sources into sentiment	No	No
6	Talk walker [42, 43]	Keyword	Yes	social channels and media from past three months	Yes	Line, o-meter, donut, cloud	- Line, o-meter, donut charts show the percentage of positive, negative	No	- Group posts by influencer	No	No
7	SentiStrength [44, 45]	Sentence	No	No	No	No	No	Yes	- Overall sentiment for an input - Sentiment for keyword(s) in text	No	No
8	Lexalytics [46 - 48]	Text (16,384 characters)	No	No	Yes	Bar, pie, cloud	- Bar charts show sentiment score for topics, locations - Cloud shows positive/negative	Yes	- Topic, entities with their sentiment score.	Yes	- Choose sentences from an input. - Sentences include sentiment lexicons

[1] social media networks, news, blogs, forums, and the web.
[2] 100 M+ Sources in 50 languages such as Facebook, Twitter, YouTube, LinkedIn, Instagram, Pinterest, etc.
[3] Facebook, Flickr, Instagram, Twitter, YouTube, Web, Reddit, Tumblr, Vimeo, VKontakte, and Dailymotion.

No.	Tool	Input	Real time	Online source	Visual output			Structured output		Textual output	
					Senti-ment	Type of chart	Description	Senti-ment	Description	Senti-ment	Description
9	Meaning Cloud [49, 50]	Text	No	No	No	No	No	Yes	- Group by levels (global, sentence, segment) with its score tag. - Each level shows sentimented entity and concept, polarity term with their score tag.	No	No
10	Sentigem [51]	Text	No	No	No	No	No	Yes	- Group sentences into sentiment (positive, negative, neutral) with an overall sentiment	No	No
11	Sentiment Analyzer [52]	Text	No	No	Yes	o-meter	- O-meter shows an overall sentiment and a level for an input	No	No	No	No
12	Proposed AAbSS	Text	No	No	Yes	Scatter, bar, pie, bubble, cloud	- Graph with 2 axes, each point is an aspect with its degree - Bar, pie charts show degrees for each aspect, each degree - Cloud shows for each positive, negative - Bubble categories based on degrees (degree includes 7 levels: love very much, love, like, neither like nor dislike, dislike, hate, hate very much)	Yes	- Group aspects into sentiment (positive, negative, neutral) - At each aspect, polarity and number of comments are mentioned	Yes	- Generate new sentences. - Sentences include sentiment lexicons

esting aspects to automatically generate three kinds of outputs using ω parameter. The summaries (visual, structured, textual) generated by the AAbSS system have good performance when the summaries are compared to other summaries generated by the other systems.

References

[1] T. Chumwatana, "Comment analysis for product and service satisfaction from Thai customers' review in social network," *J. of Inform. and Commun. Technology*, vol. 17, pp. 271-289, 2018.

[2] N. Samsudin, M. Puteh, A. R. Hamdan, and M. Z. A. Nazri, "Normalization of noisy texts in Malaysian online reviews," *J. of Inform. and Commun. Technology*, vol. 12, pp. 147–159, 2013.

[3] N. Dedić and C. Stanier, "Measuring the success of changes to existing business intelligence solutions to improve business intelligence reporting," in *Proc. 2016 Int. Conf. Research and Practical Issues of Enterprise Inform. Syst.*, Cham, pp. 225-236, 2016.

[4] W. Maharani, D. H. Widyantoro, and M. L. Khodra, "Aspect-based opinion summarization: A survey," *J. of Theoretical and Applied Inform. Technology*, vol. 95, pp. 448-456, 2017.

[5] R. E. López Condori and T. A. Salgueiro Pardo, "Opinion summarization methods: Comparing and extending extractive and abstractive approaches," *Expert Systems with Applicat.*, vol. 78, pp. 124-134, 2017.

[6] M. Hu and B. Liu, "Opinion feature extraction using class sequential rules," in *Proc. 2006 AAAI Spring Symp., Technical Report SS-06-03*, USA, pp. 61-66, 2006.

[7] G. Carenini and L. Rizoli, "A multimedia interface for facilitating comparisons of opinions," in *Proc. 14th Int. Conf. Intell. user interfaces*, USA, 2009.

[8] M. Abulaish, Jahiruddin, M. N. Doja, and T. Ahmad, "Feature and opinion mining for customer review summarization," in *Pattern Recognition and Machine Intell.*, Heidelberg, pp. 219-224, 2009.

[9] P. Kherwa, A. Sachdeva, D. Mahajan, N. Pande, and P. K. Singh, "An approach towards comprehensive sentimental data analysis and opinion mining," in *Proc. 2014 IEEE Int. Advance Computing Conf. (IACC)*, pp. 606-612, 2014.

[10] A. Kamal, "Review mining for feature based opinion summarization and visualization," *CoRR*, vol. abs/1504.03068, 2015.

[11] H. Wu, Y. Gu, S. Sun, and X. Gu, "Aspect-based opinion summarization with convolutional neural networks," in *Proc. 2016 Int. Joint Conf. on Neural Networks (IJCNN)*, pp. 3157-3163, 2016.

[12] D. Kanbur and M. S. Aktas, "Ontology based aspect oriented opinion summary methodology," in *Proc. 5th Int. Conf. on Building and Exploring Web Based Environments (WEB)*, Spain, pp. 7-16, 2017.

[13] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proc. 10th ACM SIGKDD Int. Conf. on Knowledge discovery and data mining*, USA, pp. 168-177, 2004.

[14] L. Zhuang, F. Jing, and X.-Y. Zhu, "Movie review mining and summarization," in *Proc. 15th Int. Conf. on Inform. and knowledge manage.*, USA, 2006.

[15] S. Blair-goldensohn, T. Neylon, K. Hannan, G. A. Reis, R. Mcdonald, and J. Reynar, "Building a sentiment summarizer for local service reviews," in *NLP in the Inform. Explosion Era*, 2008.

[16] D. K. Ly, K. Sugiyama, Z. Lin, and M.Y. Kan, "Product review summarization from a deeper perspective," in *Proc. 11th Annu. Int. ACM/IEEE joint Conf. on Digital libraries*, Canada, 2011.

[17] C. Liu, W. Hsaio, C. Lee, G. Lu, and E. Jou, "Movie rating and review summarization in mobile environment," *IEEE Trans.on Syst., Man, and Cybernetics, Part C (Applicat. and Reviews)*, vol. 42, pp. 397-407, 2012.

[18] J. Jmal and R. Faiz, "Customer review summarization approach using Twitter and SentiWordNet," in *Proc. 3rd Int. Conf. on Web Intell., Mining and Semantics*, Spain, 2013.

[19] K. Bafna and D. Toshniwal, "Feature based summarization of customers' reviews of online products," in *Procedia Comput. Sci.*, vol. 22, pp. 142-151, 2013.

[20] K. Yauris and M. L. Khodra, "Aspect-based summarization for game review using double propagation," in *Proc. 2017 Int. Conf. on Advanced Informatics, Concepts, Theory, and Applicat. (ICAICTA)*, pp. 1-6, 2017.

[21] R. K. Amplayo and M. Song, "An adaptable fine-grained sentiment analysis for summarization of multiple short online reviews," *Data & Knowledge Eng.*, vol. 110, pp. 54-67, 2017.

[22] I. Titov and R. McDonald, "Modeling online reviews with multi-grain topic models," in *Proc. 17th Int. Conf. on World Wide Web*, China, 2008.

[23] Y. Lu, C. Zhai, and N. Sundaresan, "Rated aspect summarization of short comments," in *Proc. 18th Int. Conf. on World wide web*, Spain, 2009.

[24] X. Xu, T. Meng, and X. Cheng, "Aspect-based extractive summarization of online reviews," in *Proc. 2011 ACM Symp. on Applied Computing*, Taiwan, 2011.

[25] G. Carenini, J. C. K. Cheung, and A. Pauls, "Multi-document summarization of evaluative text," *Computational Intell.*, vol. 29, pp. 545-576, 2013.

[26] N. Yu, M. Huang, Y. Shi, and X. Zhu, "Product review summarization by exploiting phrase prop-

erties," in *Proc. 26th Int. Conf. on Computational Linguistics: Tech. Papers (COLING)*, Japan, pp. 1113–1124, 2016.

[27] S. Angelidis and M. Lapata, "Summarizing opinions: Aspect extraction meets sentiment prediction and they are both weakly supervised," in *Proc. 2018 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, 2018.

[28] T. A. Tran, J. Duangsuwan, and W. Wettayaprasit, "A novel automatic sentiment summarization from aspect-based customer reviews," in *Proc. 15th Int. Joint Conf. on Comput. Sci. and Software Eng. (JCSSE)*, pp. 1-6, 2018.

[29] S. Gerani, Y. Mehdad, G. Carenini, R. T. Ng, and B. Nejat, "Abstractive summarization of product reviews using discourse structure," in *Proc. 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, Qatar, pp. 1602–1613, 2014

[30] M. Yang, Q. Qu, Y. Shen, Q. Liu, W. Zhao, and J. Zhu, "Aspect and sentiment aware abstractive review summarization," in *Proc. 27th Int. Conf. on Computational Linguistics (COLING)*, USA, 2018.

[31] S. Gerani, G. Carenini, and R. T. Ng, "Modeling content and structure for abstractive review summarization," *Comput. Speech & Language*, vol. 53, pp. 302-331, 2019.

[32] Awario. Awario. Available: `https://awario.com/demo/`. Last Accessed: May 13, 2020

[33] Hootsuite-Insights. Hootsuite Insights. Available: `https://hootsuite.com/`. Last Accessed: May 13, 2020.

[34] P. Grover, A. K. Kar, and P. V. Ilavarasan, "Impact of corporate social responsibility on reputation—insights from tweets on sustainable development goals by CEOs," *Int. J. of Inform. Manage.*, vol. 48, pp. 39-52, 2019.

[35] S. Viz. Tweet Sentiment Visualization. Available: `https://www.csc.ncsu.edu/faculty/healey/tweet\_viz/tweet\_app/`. Last Accessed: May 13, 2020.

[36] J. Lertsiwaporn and T. Senivongse, "Time-based visualization tool for topic modeling and sentiment analysis of Twitter messages," in *Proc. 2017 Int. MultiConf. of Eng. and Comput. Scientists (IMECS)*, Hong Kong, 2017.

[37] Social-Mention. Social mention. Available: `http://www.socialmention.com/`. Last Accessed: May 13, 2020.

[38] R. T. Khasawneh, H. A. Wahsheh, M. N. AlKabi, and I. M. Alsmadi, "Sentiment analysis of arabic social media content: A comparative study," in *Proc. 8th Int. Conf. for Internet Technology and Secured Trans. (ICITST)*, 2013, pp. 101-106.

[39] J. J. Meyers, D. L. Hansen, J. S. Giboney, and D. C. Rowe, "Training future cybersecurity professionals in spear phishing using SiEVE," in *Proc. 19th Annu. SIG Conf. on Inform. Technology Educ.*, USA, 2018.

[40] C. Borges, J. Araújo, and A. Rodrigues, "Towards an approach to elicit domain requirements from social networks: The case of emergency systems," in *Proc. 33rd Annu. ACM Symp. on Applied Computing*, France, 2018.

[41] Social-Searcher. Social media search & monitoring API. Available: `http://www.socialsearcher.com/`. Last Accessed: May 13, 2020.

[42] Talkwalker. Talkwalker. Available: `https://www.talkwalker.com`. Last Accessed: May 13, 2020.

[43] Z. Asghar, T. Ali, I. Ahmad, S. Tharanidharan, S. K. A. Nazar, and S. Kamal, "Sentiment analysis on automobile brands using Twitter data," in *Intell. Technologies and Applicat.*, Singapore, pp. 76-85, 2019.

[44] SentiStrength. SentiStrength. Available: `http://sentistrength.wlv.ac.uk/`. Last Accessed: May 13, 2020.

[45] T. Baviera, A. Sampietro, and F. J. García-Ull, "Political conversations on Twitter in a disruptive scenario: The role of "party evangelists" during the 2015 Spanish general elections," *The Commun. Review*, vol. 22, pp. 117-138, 2019.

[46] R. Feldman, B. Rozenfeld, R. Bar-Haim, and M. Fresko, "The Stock Sonar - sentiment analysis of stocks based on a hybrid approach," in *Proc. 23rd Conf. on Innovative Applicat. of Artificial Intell. (IAAI)*, USA, 2011.

[47] T. P. Lau, S. Wang, Y. Man, C. F. Yuen, and I. King, "Language technologies for enhancement of teaching and learning in writing," in *Proc. 23rd Int. Conf. on World Wide Web*, Korea, 2014.

[48] Lexalytics. Lexalytics: Data analytics with NLP & text analytics. Available: `https://www.lexalytics.com/`. Last Accessed: May 13, 2020.

[49] Meaning-Cloud. Meaning cloud. Available: `https://www.meaningcloud.com/`. Last Accessed: May 13, 2020.

[50] T. Swoboda, M. Kaufmann, and M. L. Hemmje, "Toward cloud-based classification and annotation support," in *Proc. 6th Int. Conf. on Cloud Computing and Services Sci.*, Vol. 1 and 2, Italy, 2016.

[51] F. Z. Ennaji, A. E. Fazziki, M. Sadgal, and D. Benslimane, "Social intelligence framework: Extracting and analyzing opinions for social CRM," in *Proc. 12th Int. Conf. of Comp. Syst. and Applicat. (AICCSA)*, pp. 1-7, 2015.

[52] Sentiment-Analyzer. Sentiment analyzer. Available: `https://www.danielsoper.com/sentimentanalysis/default.aspx`. Last Accessed: May 13, 2020.

[53] M. Pontiki, D. Galanis, H. Papageorgiou,

S. Manandhar, and I. Androutsopoulos, "SemEval2015 task 12: Aspect based sentiment analysis," in *Proc. 9^{th} Int. Workshop on Semantic Evaluation (SemEval 2015)*, pp. 486-495, 2015.

[54] Q. Liu, Z. Gao, B. Liu, and Y. Zhang, "Automated rule selection for aspect extraction in opinion mining," in *Proc. 24^{th} Int. Conf. on Artificial Intell.*, Argentina, 2015.

Tuan Anh Tran is a PhD. candidate at Department of Computer Science, Faculty of Science, Prince of Songkla University, Thailand. He got his BSc. in Information Technology from Hue University of Education, MSc. in Information System from Ho Chi Minh city University of Sciences, Vietnam. His research interest is Natural Language Processing.

Jarunee Duangsuwan is an assistant professor at Department of Computer Science, Faculty of Science, Prince of Songkla University, Thailand. She got BSc., MSc., PhD. in Computer Science from Chiang Mai University, Prince of Songkla University, and University of Reading, UK, respectively. Her research interests are Natural Language Processing and Machine Learning.

Wiphada Wettayaprasit is an assistant professor at Department of Computer Science, Faculty of Science, Prince of Songkla University, Thailand. She got BSc., MSc., PhD. in Computer Science from Prince of Songkla University, University of Missouri-Columbia, USA, and Chulalongkorn University, respectively. Her research interests are Artificial Intelligence, Neural Networks and Machine Learning.