BrailleSUM: A News Summarization System for the Blind and Visually Impaired People

Xiaojun Wan and Yue Hu
Institute of Computer Science and Technology, The MOE Key Laboratory of Computational Linguistics, Peking University, Beijing 100871, China
{wanxiaojun, ayue.hu}@pku.edu.cn

Abstract

In this article, we discuss the challenges of document summarization for the blind and visually impaired people and then propose a new system called BrailleSUM to produce better summaries for the blind and visually impaired people. Our system considers the factor of braille length of each sentence in news articles into the ILP-based summarization method. Evaluation results on a DUC dataset show that BrailleSUM can produce shorter braille summaries than existing methods, meanwhile, it does not sacrifice the content quality of the summaries.

1 Introduction

People with normal vision can read news documents with their eyes conveniently. However, according to WHO's statistics, up to October 2013, 285 million people are estimated to be visually impaired worldwide: 39 million are blind and 246 have low vision. Unfortunately, the large number of blind and visually impaired people cannot directly or conveniently read ordinary news documents like sighted people, and they have to read braille with their fingerprints or special equipments, which brings much more burden to them. Braille is a special system with a set of symbols composed of small rectangular braille cells that contain tiny palpable bumps called raised dots used by the blind and visually impaired. It is traditionally written with embossed paper. Special equipments such as refreshable braille displays and braille embosser have been developed for the blind and visually impaired people to read or print on computers and other electronic supports.

Though some news materials have already been prepared in braille format for the blind people's reading and learning, most daily news documents are written for sighted people, and it is necessary to first translate the news documents into Braille, and then the blind people can read the news with their fingertips. Speech synthesizers are also commonly used for the task (Freitas and Kouroupetroglou, 2008), but the way of reading braille texts is still popular in the daily life of the blind people, especially for the deaf-blind people.

As we know, document summarization is a very useful means for people to quickly read and browse news articles in the big data era. Existing summarization systems focus on content quality and fluency of summaries, and they usually extract several informative and diversified sentences to form a summary with a given length. The summaries are produced for sighted people, but not for the blind and visually impaired people. A text summary can be translated into a braille summary for the blind and visually impaired people's reading, and the length of a braille summary is defined as the number of the braille cells in the summary. It is noteworthy that the shorter the braille summary is, the less burden the blind people have when reading the summary with their fingertips. The burden lies in the fact that reading a braille text by touching each braille cell with fingertips is more difficult and inconvenient than reading a normal text with eyes. So a braille summary is required to be as short as possible, while keeping the content quality and fluency.

In this study, we investigate the task of document summarization for the blind and visually impaired people for the first time. We discuss the major challenges of document summarization for the blind and visually impaired people and then propose a new system called BrailleSUM to produce better summaries for them. Our system considers the factor of braille length of each sentence in news articles into the ILP-based summarization method. Evaluation results on a DUC dataset show that BrailleSUM can produce much shorter braille summaries than existing methods, meanwhile, it does not sacrifice the content quality of the summaries.

2 Related Work

Most previous summarization methods are extraction-based, which directly rank and extract exist-

Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Short Papers), pages 578–582, Beijing, China, July 26-31, 2015.

ing sentences in a document set to form a summary. Typical methods include the centroid-based method (Radev et al., 2004), NeATS (Lin and Hovy, 2002), supervised learning based methods (Ouyang et al., 2007; Shen et al., 2007; Schilder and Kondadadi, 2008; Wong et al., 2008), graph-based ranking (Erkan and Radev, 2004; Mihalcea and Tarau, 2005), Integer Linear Programming (Gillick et al., 2008; Gillick and Favre, 2009; Li et al., 2013), and submodular function (Lin and Bilmes, 2010). Moreover, cross-language document summarization has been investigated (Wan et al., 2010), but the task focuses on how to select the translated sentences with good content quality. We can see that all existing summarization systems were proposed for sighted people, but not for the blind and visually impaired people. Document summarization for the blind and visually impaired people has its specialty and is worth exploring.

It has been a long way to help the blind and visually impaired people to browse information as conveniently as ordinary people. Special devices have been developed for achieving this long-term goal (Linvill and Bliss, 1966; Shinohara et al., 1998). After the popularity of Braille, many kinds of braille display devices have been developed for braille reading (Rantala et al., 2009). In addition, most research in this area focused on how to improve accessibility of web information for the blind people (Salampasis et al., 2005; Mahmud et al., 2007; Hadjadj and Burger, 1999).

3 Preliminaries of Braille Grades

Braille is a system of raised dots arranged in cells and it was developed by Louis Braille in the beginning of the 19th century. Braille letters, common punctuation marks, and a few symbols are displayed as raised 6 dot braille cell patterns read by using a fingertip to feel the raised dots. The number and arrangement of these raised dots within a cell distinguish one character from another. For example, the letters “a”, “b” and “c” are displayed as ⠁⠃⠉, respectively. Due to the varying needs of braille readers, there are different grades of braille. In this study we adopt grade 2 braille – EBAE (English Braille America Edition). Grade 2 braille was a space-saving alternative to grade 1 braille. In grade 2 braille, a cell can represent a shortened form of a word. Many cell combinations have been created to represent common words, making this the most popular of the grades of braille. There are part-word contractions (e.g. “stand” →⠌⠯, “without” →⠾⠳⠞), which often stand in for common suffixes or prefixes, and whole-word contractions (e.g. “every” → ⠑, “knowledge” →⠅), in which a single cell represents an entire commonly used word. Words may be abbreviated by using a single letter to represent the entire word, using a special symbol to precede either the first or last letter of the word while truncating the rest of the word, using a double-letter contraction such as "bb" or "cc", or removing most or all of the vowels in a word in order to shorten it. A complex system of styles, rules, and usage has been developed for this grade of braille.

4 System Overview

The focus of traditional summarization tasks is how to improve the content quality of a summary with a given length limit, and the content quality of a summary is measured by the overlap between the summary and reference summaries written by annotators. However, document summarization for the blind and visually impaired people is different from traditional summarization tasks. Besides the content quality, the length of a braille summary is a very important factor to be considered, because the number of braille cells in a braille summary have a direct impact on the blind and visually impaired people when they read the summary with their fingertips, and more highly contracted braille is quicker to read, as shown in previous studies such as (Veispak et al., 2012).

Given a document set, our new summarization task aims to produce a braille summary, which are translated from a traditional textual summary with a predefined length (usually measured by the count of words). The braille summary is required to keep the content quality, measured by the content quality of the textual summary. Moreover, the braille length of the summary is required to be as short as possible. The length of a braille summary is defined as the number of the rectangular braille cells in the summary. The shorter the length is, the blind and visually impaired people will spend less time reading the summary with their fingertips and thus the summary is better. For simplicity, we define the braille length of a textual summary as the length of its translated braille summary. For example, the braille length of a text “hello, world!” is 9 since the length of its translated braille text ⠓⠑⠇⠇⠕⠂ ⠸⠺⠖ is 9.

A basic solution to the new summarization task is first applying an existing summarization algorithm (e.g. the most popular ILP-based method) to produce a summary, and then translating the summary into a braille summary, which is called BasicSUM. However, the braille translation is not a simple character-to-block conversion process and there exist various contractions during the translation process, as mentioned in the previous section. Two content-similar sentences may be translated into two braille sentences with totally different lengths due to the different word lengths and conversion contractions. Therefore, our solution is to consider the new factor of braille length of each sentence during the summarization process and produce a summary with shorter braille length while keeping its content quality. In our proposed BrailleSUM system, we incorporate the factor of braille length into the ILP-based summarization framework with a new ILP formulation.

5 ILP-Based Braille Summarization

In this study, we adopt the popular ILP-based summarization framework for addressing the new task of braille summarization. The concept-based ILP method for summarization is introduced by (Gillick et al., 2008; Gillick and Favre, 2009), and its goal is to maximize the sum of the weights of the language concepts (i.e. bigrams) that appear in the summary. The ILP method is very powerful for extractive summarization because it can select important sentences and remove redundancy at the same time. Formally, the ILP method can be represented as below:

$$max \sum_{i=1}^{|B|} c_{b_i} b_i \quad (1)$$

subject to:

$$\sum_{i=1}^{N} l_i s_i \leq L_{max} \quad (2)$$

$$\sum_{i \in B_j} b_i \geq |B_j| s_j \text{, for } j = 1, \ldots, N \quad (3)$$

$$\sum_{j \in S_i} s_j \geq b_i \text{, for } i = 1, \ldots, |B| \quad (4)$$

$$b_i, s_j \in \{0,1\}, \forall i, j$$

where:

b_i, s_j are binary variables that indicate the presence of bigram *i* and sentence *j*, respectively;

c_{b_i} is the document frequency of bigram b_i;

B is the set of unique bigrams;

B_j is the set of bigrams that sentence *j* contains.

S_i is the set of sentences that contain bigram *i*.

N is the count of the sentences;

L_{max} is the maximum word count of the summary, which is set to 250 in the experiments;

l_i is the word count of sentence *i*.

Constraint (2) ensures that the total length of the selected sentences is limited by the given length limit. Inequalities (3)(4) associate the sentences and bigrams. Constraint (3) ensures that selecting a sentence leads to the selection of all the bigrams it contains, and constraint (4) ensures that selecting a bigram only happens when it is present in at least one of the selected sentences.

The new objective function for braille summarization consists of two parts: the original part reflecting the content quality and the new part reflecting the braille length factor. The function is presented as below and the constraints are the same with (2)(3)(4).

$$max\{(1-\lambda)\sum_{i=1}^{|B|} \frac{c_{b_i} b_i}{C} + \lambda \sum_{j=1}^{N} braille\_ratio_j s_j\} \quad (5)$$

where $C = \sum_{i \in B} c_{b_i}$ is a normalization constant to make the values of the two parts in the equation comparable. $\lambda \in [0, 1]$ is a combination parameter to reflect the different influences of the two parts. $braille\_ratio_j$ is a new factor to reflect the suitability level of sentence *j* to be selected, which is computed as below:

$$braille\_ratio_j = \frac{l_j}{bl_j} \quad (6)$$

where bl_j is the braille length of sentence *j*, and it is defined as the number of braille cells in the corresponding braille sentence. l_j is the word count in the original sentence. As mentioned earlier, the number of characters and signs in an English sentence is not equal to the number of the braille cells in the corresponding braille sentence, since grade 2 braille is not based on a simple one-to-one conversion from each character or sign to a braille cell. In this study, we adopt the open-source libbraille[1] tool for converting an English sentence into a braille sentence, and then get the braille length of the sentence. An example English sentence and its corresponding braille sentence are shown below:

Infected feed cannot account for four cases.

⠠⠔⠋⠑⠉⠞⠫ ⠋⠑⠫ ⠸⠉ ⠁⠒⠐⠞ ⠿ ⠋⠳⠗ ⠉⠁⠎⠑⠎⠲

We can see that the number of characters and signs in the English sentence is 38, while the number of braille cells in the braille sentence is 26, and thus the braille length bl_j is 26. We can also simply know that the word count of the sentence l_j is 7. Thus the braille ratio of the sentence is 7/26=0.269. We can see that if a sentence has a larger ratio of its word count to its braille length, then it is more suitable to be selected. Particularly, for two sentences with the same word count, the one with a shorter braille length is preferred. Note that since the sum of l_j for the sentences in a summary is fixed, the sum of bl_j for the sentences should be as small as possible in order to maximize the second part in Equation (5). For the new objective function in Equation (5), the first part ensures the content quality, and the second part tries to make the braille length of the summary as short as possible. The combination of the two

[1] http://libbraille.org/

parts can achieve the two goals of our new summarization task at the same time. If the combination parameter λ is set to 0, then the formulation in (5) is actually the same with (1).

Finally, we solve the above linear programming problem by using the IBM CPLEX optimizer and get the English summary according the value of each variable s_j. The corresponding braille summary can be produced after translation with libbraille.

6 Evaluation

In this study, we used the multi-document summarization task in DUC2006 for evaluation. DUC2006 provided 50 document sets and a summary with a length limit of 250 words was required to be created for each document set. Reference summaries have been provided by NIST annotators. For simplicity, the topic description was ignored in this study. In the experiments, our proposed BrailleSUM system with the new ILP method in Equation (5) was compared with the BasicSUM system with the traditional ILP method in Equation (1). The parameter λ in BrailleSUM is simply set to 1/4 (i.e. 0.25).

Since the aim of our system is reducing the braille length of a summary without sacrificing its content quality, we evaluate the summaries from the following two aspects: First, we evaluate the content quality of the summaries by measuring the content overlap between the summaries and the reference summaries with the ROUGE-1.5.5 toolkit (Lin and Hovy, 2003). In this study, we use three ROUGE recall scores in the experimental results: ROUGE-1 (unigram-based), ROUGE-2 (bigram-based) and ROUGE-SU4 (based on skip bigram with a maximum skip distance of 4). Second, we compute the braille length of each summary by summing the braille lengths of all the sentences in the summary, and then average the lengths across the 50 document sets.

The comparison results on summary content quality and average summary braille length are shown in Table 1. We can see that BrailleSUM and BasicSUM can achieve very similar ROUGE scores, and the score differences are non-significant because the 95% confidence intervals are highly overlapped. The scores of BrailleSUM and BasicSUM are much higher than that of the NIST baseline and the average scores of all participating systems (i.e. AverageDUC). More importantly, BrailleSUM can produce summaries with much shorter braille lengths than BasicSUM, and the braille length reduction is significant. The results demonstrate that BrailleSUM can produce much shorter braille summaries while not sacrificing the summaries' content quality. We can see that the incorporation of the braille length factor into the ILP framework is very effective for addressing the new summarization task.

In order to show the influence of parameter λ in BrailleSUM, we vary λ from 0 to 1, and show the curves of ROUGE-1 and ROUGE-2 scores, and average braille length in Figures 1-3, respectively. We can see that with the increase of λ, the average braille length of the produced summaries is decreasing steadily. The result can be easily explained by that a larger λ means more consideration of the braille length factor. We can also see from the figures that when λ is less than 0.3, the ROUGE scores usually keep steady and do not decline significantly, but when λ is becoming larger, the ROUGE scores decline obviously. The results demonstrate that the content quality factor and the braille length factor need to be balanced with a proper value of λ.

	ROUGE-1	**ROUGE-2**	**ROUGE-SU4**	**Average Braille Length**
BrailleSUM	0.39012 [0.38380-0.39590]	0.09010 [0.08617-0.09396]	0.14009 [0.13665 - 0.14332]	932* ($\triangle bl$ =103)
BasicSUM	0.38958 [0.38273-0.39586]	0.09219 [0.08791-0.09614]	0.14011 [0.13691-0.14368]	1035
AverageDUC	0.37250	0.07391	0.12928	-
NIST Baseline	0.30217	0.04947	0.09788	-

Table 1: Comparison results of summary content quality (ROUGE Recall) and average summary braille length. (The 95% confidence interval for each ROUGE score is reported in brackets; Δbl means the reduction of average braille length over BasicSUM; * means the average braille length reduction over BasicSUM is statistically significant with p-value=2.46975E-18 for t-test.)

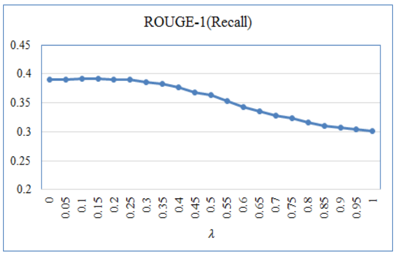

Figure 1. ROUGE-1 vs. λ

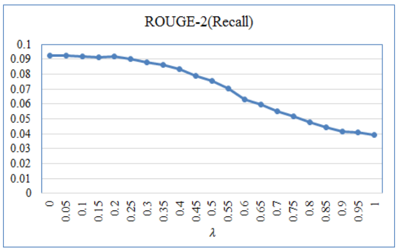

Figure 2. ROUGE-2 vs. λ

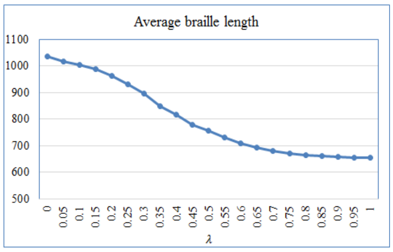

Figure 3. Average braille length vs. λ

Acknowledgments

The work was supported by National Hi-Tech Research and Development Program (863 Program) of China (2015AA015403) and National Natural Science Foundation of China (61170166, 61331011).

References

G. Erkan and D. R. Radev. 2004. LexPageRank: prestige in multi-document text summarization. In *Proceedings of EMNLP-04*.

D. Freitas and G. Kouroupetroglou. 2008. *Speech technologies for blind and low vision persons. Technology and Disability*, 20(2), 135-156.

D. Gillick, B. Favre and D. Hakkani-Tur. 2008. The ICSI summarization system at TAC 2008. In *Proceedings of the Text Understanding Conference*.

D. Gillick and B. Favre. 2009. A scalable global model for summarization. In *Proceedings of the Workshop on Integer Linear Programming for Natural Langauge Processing on NAACL*.

D. Hadjadj and D. Burger. 1999. Braillesurf: An html browser for visually handicapped people. In *Proceedings of Tech. and Persons with Disabilities Conference*, 1999.

C. Li, X. Qian and Y. Liu. 2013. Using supervised bigram-based ILP for extractive summarization. In *Proceedings of ACL* (pp. 1004-1013), 2013.

H. Lin and J. Bilmes. 2010. Multi-document summarization via budgeted maximization of submodular functions. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 912-920), 2010.

C.-Y. Lin and E.. H. Hovy. 2002. From single to multi-document summarization: a prototype system and its evaluation. In *Proceedings of ACL-02*.

C.-Y. Lin and E.H. Hovy. 2003. Automatic evaluation of summaries using n-gram co-occurrence statistics. In *Proceedings of HLT-NAACL -03*.

J. G. Linvill and J. C. Bliss. 1966. A direct translation reading aid for the blind. *Proceedings of the IEEE*, 54(1), 40-51, 1966.

J. U. Mahmud, Y. Borodin and I. V. 2007. Ramakrishnan. Csurf: a context-driven non-visual web-browser. In *Proceedings of the 16th international conference on World Wide Web* (pp. 31-40), 2007.

R. Mihalcea and P. Tarau. 2005. A language independent algorithm for single and multiple document summarization. In *Proceedings of IJCNLP-05*.

Y. Ouyang, S. Li, W. Li. 2007. Developing learning strategies for topic-focused summarization. In *Proceedings of CIKM-07*.

D. R. Radev, H. Y. Jing, M. Stys and D. Tam. 2004. Centroid-based summarization of multiple documents. *Information Processing and Management*, 40: 919-938, 2004.

J. Rantala, R. Raisamo, J. Lylykangas, V. Surakka, J. Raisamo, K. Salminen, T. Pakkanen and A. Hippula. 2009. Methods for presenting Braille characters on a mobile device with a touchscreen and tactile feedback. *Haptics, IEEE Transactions on*, 2(1), 28-39, 2009.

M. Salampasis, C. Kouroupetroglou and A. Manitsaris. 2005. Semantically enhanced browsing for blind people in the WWW. In *Proceedings of the sixteenth ACM conference on Hypertext and hypermedia* (pp. 32-34), 2005.

F. Schilder and R. Kondadadi. 2008. FastSum: fast and accurate query-based multi-document summarization. In *Proceedings of ACL-08: HLT*.

C. Shen and T. Li. 2010. Multi-document summarization via the minimum dominating set. In *Proceedings of COLING-10*.

D. Shen, J.-T. Sun, H. Li, Q. Yang, and Z. Chen. 2007. Document summarization using conditional random fields. In *Proceedings of IJCAI-07*.

M. Shinohara, Y. Shimizu and A. Mochizuki. 1998. Three-dimensional tactile display for the blind. *Rehabilitation Engineering, IEEE Transactions on*, 6(3), 249-256, 1998.

A. Veispak, B. Boets and P. Ghesquiere. 2012. Parallel versus Sequential Processing in Print and Braille Reading. *Research in Developmental Disabilities: A Multidisciplinary Journal* 33(6): 2153-2163, 2012.

X. Wan, H. Li and J. Xiao. 2010. Cross-language document summarization based on machine translation quality prediction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 917-926), 2010.

K.-F. Wong, M. Wu and W. Li. 2008. Extractive summarization using supervised and semisupervised learning. In *Proceedings of COLING-08*.