

# Summarization Assistant for News Brief Services on Cellular Phones

Yuen-Hsien Tseng\*

## Abstract

A Chinese news summarization method is proposed in order to help humans deal with the message services of news briefs broadcast over cell phones. The problem to be solved here is unique because a strict length limit (69 or 45 characters) is imposed on the summaries for the message service. This requires some sort of automatic sentence fusion, rather than sentence selection alone. In the proposed method, important sentences were first identified based on the news content. They were matched against the news headline to determine a suitable position for concatenation with the headline to become candidates. These candidates were then ranked by their length and fitness for manual selection. In our evaluation, among 40 short news updates in the inside testing set, over 75% (80%) of the best candidates yield acceptable summaries without manual editing for the length limit of 69 (45) characters. These numbers, however, reduce to 70.7% (53.3%) for the outside testing set of 75 news stories of ordinary length. It seems that the shorter the length limit, the more difficult the problem of getting the summary from long stories. Nevertheless, the proposed method has the potential not only to reduce the cost of manual operation, but also to integrate and synchronize with other media in such services in the future.

**Keywords:** Cell Phone Service, News Brief Message, Automated Summarization, Chinese News

## 1. Introduction

The popularity of cell phones in the Taiwan area has reached the highest rate in the world during the last few years. Over 23 million cell phone numbers were used as of June 2002, which is slightly more than the population of Taiwan (Wang, 2002). To better utilize this ubiquitous communication device, a number of content providers have provided Chinese news

---

\* National Taiwan Normal University, No.162, Sec. 1, Heping East Road., Taipei City, Taiwan (R.O.C.),  
106, Tel: 886-2-77345535  
E-mail: samtseng@ntnu.edu.tw

brief services over the cell phone, such as United Daily News (United Daily News, n.d.), Central News Agency (Central News Agency, n.d.), and PC Home in Taiwan. The Asahi Shimbun in Japan (the second largest news agency in the world) has provided such news message services at an inexpensive rate since 1999, in the hope that the increase in the number of the readers of their content could lead to an increase in the subscriptions to their newspaper (China Times, n.d.). As multimedia technologies continue to improve, future news service over the cell phone may not only include text, but also speech, images, or video, integrated and synchronized. To reach this vision, however, the operation cost should be low enough to sustain such services. Therefore, automated methods of cost containment would be of great help.

The news brief shown on a cell phone is different from one on a desktop computer. Due to the limited screen size, a length limit is defined for each news message. This is usually 45 Chinese characters in PHS systems or 69 characters in other systems, including punctuation marks (United Daily News, n.d.). Summaries of this kind are longer than a news headline but shorter than a long Chinese sentence. For the benefit of the subscribers, the summaries should contain as much content as possible to reduce the frequency of retrieving the whole news story. Also the readability and coherence of the summaries are important factors that should be taken into account.

From the research perspective, the task defined above is a challenge for automatic document summarization. Previous studies have shown that the shorter the summary required, the lower the performance of machine-generated summary (Lin & Hovy, 2003), hence, the more difficult the problem is. The task of news brief summarization for cell phones falls into this difficult category. On the other hand, human summarization of news stories for cell phones is not really a difficult problem. As mentioned above, the main issue is whether one can achieve this task in a low-cost and efficient way. Strictly maintaining the length limit requires a human summarizer to pay attention to the number of characters already there while making the summarization. If a machine could suggest a number of summary candidates, each with its length shown, for human selection, not only would the human summarizer be relieved of such tedious work and improve his/her efficiency, but also the task would become less difficult for machine summarization.

This article proposes a Chinese news summarization technique to assist human summarizers in the above way, with the aim of meeting the considerations described above. Basically, our approach is a sentence fusion technique that merges the news headline with the body sentence that supplements the information carried by the headline. After a brief review of previous work in the next section, the detailed approach and its motivations are described. The performance is then evaluated and the results are shown. This is followed by a discussion of the strengths and weaknesses of the proposed method. Finally, we conclude this paper with

some other possible applications and future work for further exploring Chinese news summarization techniques.

## **2. Related Work**

Automatic news summarization techniques have been widely explored in recent years, such as the summarization tasks in DUC (DUC, n.d.) or in NTCIR (Fukushima, Okumura, & Nanba, 2002). Several practical systems (e.g. (Hovy & Lin, 1999; Evans, Klavans, & McKeown, 2004; Radev, Otterbacher, Winkel, & Blair-Goldensohn, 2005)) have been developed in the past decade. The summarization techniques used in most studies can be divided into two approaches: abstraction and extraction (Mani, 2001; Radev, Hovy & McKeown, 2002). In abstraction, advanced natural language processing (NLP) techniques are applied to analyze sentential information and then to generate concise sentences with proper semantics. Sophisticated NLP techniques, such as anaphora resolution, may be used and certain human maintained knowledge bases or corpora may be needed. In extraction, statistical techniques are applied to rank and select the text snippets for a summary. Due to its relatively low cost and high robustness across application domains and document genres, most summarization tasks adopt the extraction approach (Carbonell & Goldstein, 1998; Lin & Hovy, 2002, Tseng, *et al*, 2007). Nevertheless, abstraction-based methods move the summarization field from the use of purely extractive methods to the generation of abstracts that contain sentences not found in any of the input documents and also synthesize information across sources (Barzilay & McKeown, 2005). Thus, the need for an abstraction-based approach is sometimes inevitable.

Despite the vast literature already published, most of the studies are for English. Although some have focused on Chinese news (e.g. (Chen, Kuo, Huang, Lin & Wung, 2003)), none have been done for the problem discussed here. The problem to be solved in this paper is unique due to the facts that there is a strict length limit imposed and that the range of the length limit makes most simple sentence selection approaches invalid. Thus, an abstraction-based method or a similar one that requires sentence fusion or alteration is required.

For example, in (Takefumi, Hidetaka, & Hiroshi, 2003) the authors reported a deletion-based approach to summarize a Web news article for PC to another short article for cell phones for Japanese. There, the length limit of the short article ranges from 50 to 100 Japanese characters. The approach first computes the values of TFxIDF for each clause in advance. A few significant sentences from the original article are then extracted based on the TFxIDF values. After that, verbose descriptions corresponding to the leaves of the dependency trees, having the lowest TFxIDF, are removed from the sentences until the length of the result of summarization is within the limit.

An important issue in automatic summarization is the evaluation of machine-derived summaries. This is not an easy task. Two main approaches are commonly applied: intrinsic and extrinsic evaluation (Mani, 2001). In intrinsic evaluation, manually prepared answers or evaluation criteria are compared with those that are machine generated. In extrinsic evaluation, automated summaries are evaluated based on their performance or influence on other tasks, such as document categorization. We adopt the intrinsic approach here since it is obviously suitable for our task.

### 3. The Proposed Summarizer

To develop an automated Chinese news summarizer subject to the limitations of a cell phone, an understanding of the style of the news stories and how humans summarize them would be helpful. Table 1 lists three news examples and their English translations. As can be seen, these examples are short, with their bodies having only 1, 2, and 3 sentences, respectively. This is not uncommon for the stories to be transmitted to users' cell phones, although longer stories may be selected as well. Given such short stories, a human summarizer has very few clues as to rewrite the story thoroughly to fit the length limit. The best he or she can do may be to cut and paste the snippets from the news text with minimal editing to avoid garbling the original meaning.

The snippets to be cut and pasted can be enumerated then suggested by a computer for manual selection. Nevertheless, the possibilities of such enumeration would be huge if all substrings of the news text are blindly considered. As can be seen from the examples in Table 1, a Chinese sentence is often composed of several comma-separated clauses, which convey the meaning of the sentence in successive sequence. Chinese clauses are independent from each other in some circumstances and, thus, constitute a useful unit to be combined with others to make a new sentence. Although most of the combined sentences would be invalid, several of them would still be meaningful and sometimes more complete in content, especially for those from the beginning and ending clauses.

**Table 1. Three news examples for summarization<sup>1</sup>. The number in parenthesis is the number of characters in the preceding sentence.**

1	太空探測器在遙遠的恆星周圍發現水的痕跡 (19) 美國航空航太總署的科學家星期三稱，新近在一顆遙遠的恆星周圍發現了水存在的痕跡，這可以成爲第一個支援除我們自己存在地外生命的證據。(64) #2001/07/13#
	<b>Space Probe Sees Signs of Water Around Distant Star</b> Newly detected signs of water around a distant star are the first evidence that planetary

<sup>1</sup> The first two stories were accessed on 2005/01/04 from <http://www.1999.com.tw/english/>, while the third story was accessed on 2005/01/05 at <http://news2.ngo.org.tw/php/ens.php?id=03102302>

	systems outside our own might be able to support life, NASA scientists said on Wednesday.
2	<p><b>專家：世界人口接近頂點 90 億 (13)</b>  <u>科學家星期三預測說，在 2070 年左右，世界人口可能會達到頂峰約 90 億，然後開始下降。(38)</u>  <u>澳大利亞人口統計學家在考慮很多因素後計算出到本世紀結束時，地球上的人口會下降至 84 億人。(43)</u> #2001/08/03#</p> <p><b>Experts: World Population Set to Peak at 9 Billion</b>  The world's population will probably peak at about 9 billion around 2070 before it starts to decline, scientists predicted Wednesday.  Demographers at a think tank in Austria calculate that by the turn of the century the number of people on the planet will have dropped down to 8.4 billion people.</p>
3	<p><b>海洋生物普查行動發現數百種新生物 (16)</b>  <u>有史以來第一次的「海洋生物普查計畫」進行以來，來自 53 個國家的科學家們，平均每星期便發現 3 種新的海洋魚種。(52)</u>  <u>這項為期 10 年、耗資 10 億美元的計畫，動員了來自世界各地的科學家共同合作，目的是為了將存在海洋中所有種類的生物發掘出來並予以分類。(63)</u>  <u>計畫至今已實施了 3 年，科學家在海裡發現了 15300 多種生物，他們估計還有 5 千多種生物尚未被科學界發掘。(49)</u> #2003/10/23#</p> <p><b>Marine Life Census Finds Hundreds of New Species</b>  New marine fish species are being logged at an average rate of three per week by scientists from 53 countries engaged in the first Census of Marine Life.  The 10 year, \$1 billion global scientific collaboration aims to identify and catalog all life in the oceans.  After their first three years of work, census scientists report over 15,300 species of fish in the sea and estimate 5,000 more are still unknown to science.</p>

**Table 2. The summary candidates for the third story in Table 1. They were created by combining the last clause of each body sentence with the headline.**

1	<p>海洋生物普查行動發現數百種新生物，平均每星期便發現 3 種新的海洋魚種。(34)  Marine life census finds hundreds of new species, at an average rate of three per week.</p>
2	<p>海洋生物普查行動發現數百種新生物，目的是為了將存在海洋中所有種類的生物發掘出來並予以分類。(45)  Marine life census finds hundreds of new species, aims to identify and catalog all life in the oceans.</p>
3	<p>海洋生物普查行動發現數百種新生物，他們估計還有 5 千多種生物尚未被科學界發掘。(38)  Marine life census finds hundreds of new species; they estimate 5,000 more are still unknown to science.</p>

Take the third story from Table 1 as an example. The headline has only 16 characters, falling short of the required length of 45 or 69. The other 3 sentences have 52, 63, and 49 characters, respectively. None of them alone is an ideal summary of the required length. Nevertheless, *by concatenating the headline and the last clause of each body sentence*, as shown in Table 2, each becomes a better choice for summaries of length 45.

It is noted that the simple Select-First-N strategy which usually has been the baseline method for most news summarization tasks would not work here. As can be seen from Table 1, if the first  $n$  characters were used as the summary, the underlined text in the first sentence of the story would be chosen as the summary for the length limit 45. These summaries, however, are incomplete in their meaning. If the first clauses (ending with a comma or period and no longer than  $n$ ) were used as the summary, they may be still incomplete in meaning or too verbose to deliver the message even when their meanings are complete. For example, for the length limit 45, the first two clauses of the first story: “美國航空航太總署的科學家星期三稱，新近在一顆遙遠的恒星周圍發現了水存在的痕跡” (38 characters) would be extracted based on the Select-First-N clause strategy. This, however, is inferior to the perfect summary “太空探測器在遙遠的恒星周圍發現水的痕跡，這可以成爲第一個支援除我們自己存在地外生命的證據。” (45 characters) which is extracted based on the heuristic rule shown in Table 2. In this example, the headline perfectly replaces the first two clauses, leaving more space for including the supplemental information in the final clause: “這可以成爲第一個支援除我們自己存在地外生命的證據。”.

The above observation gives us clues to effectively enumerate the summary candidates. Nevertheless, there are other problems that need to be considered in order to further reduce the burden of human selection: (1) The number of suggested candidates should be fairly equal for each story. Long stories should not yield considerably more candidates than short ones. (2) The candidates should be ranked in some sense when they are suggested for selection.

To tackle these problems, we propose the following processing steps:

- Step 1: Sort all the sentences of a news story by their weights and select the best 5 sentences for use in the next step.
- Step 2: Generate summary candidates by matching and combining each selected sentence with the news headline. Calculate the match scores and summary lengths.
- Step 3: Sort the candidates by their lengths and scores.

In Step 1, the weight of a sentence in a story of any length is determined by the accumulated weights of the keywords occurring in that sentence, as shown below:

$$weight(S) = \sum_{w \in Keywords \in S} (0.5 + 0.5 \times tf_w / \max\_tf)$$

where  $tf_w$  is the term frequency of keyword  $w$  and  $max\_tf$  is the term frequency of the keyword which occurs most in the news story. Here, the keywords of a story are those headline words that remain from non-content-bearing word deletion and those maximally repeated patterns in the story that are extracted by Tseng's algorithm (Tseng, 2002). Tseng has shown that Chinese news stories can contain many new keywords, almost 1/3 of repeated words are unknown to a lexicon of 123,226 terms. His algorithm ensures that unknown words can be extracted as well, as long as they occur at least twice in a document.

In Step 2, since headlines are guides to a news story, they should be included in the beginning of the candidates. The ending clauses to be concatenated should supplement the content of the headline. This means that the beginning clauses of a body sentence should be as similar to the headline as possible. To spot the position for concatenation and to know the similarity, a dynamic programming (DP) technique is used.

Given two strings  $A[1..n]$  and  $B[1..m]$ , where  $n \leq m$ , the edit distance between  $A[1..i]$  and  $B[1..j]$  based on DP (Levenshtein, 1966) is:

$$d[i, j] = \min(d[i-1, j], d[i-1, j-1], d[i, j-1]) + c(A[i], B[j])$$

where  $\min$  is a function that returns the minimum of its 3 arguments, and  $c(A[i], B[j]) = 0$  if  $A[i]=B[j]$ , and 1 otherwise. The initial values for the distance are:  $d[0, 0]=0$ ,  $d[0, j]=0$  for  $j=1..m$  and  $d[i, 0] = d[i-1,0]+1$  for  $i=1..n$ .

A similarity function is defined in (Lopresti & Zhou, 1996) to convert the edit distance into the similarity:  $\exp( d[n,j] / (d[n,j] - n) )$ , where  $\exp$  is the exponent function. This similarity ranges from 0 to 1. We found, however, that its range does not distribute well for later comparison. Thus, it was changed into:

$$sim(j) = \exp\left(\frac{d[n, j]}{d[n, j] - m - n}\right)$$

where  $j$  denotes the  $j$ -th character (including the punctuation) of the body sentence. The new measure ranges from  $\exp(-n/m)$  to 1.

The starting position (the position for concatenation) of the ending clauses is first determined by the comma which most closes the character with highest similarity. Since we favor length more than similarity (here, length is a direct measure that must be met, while similarity is just an approximation of the content similarity between the headline and the body sentence), the starting position is changed to its preceding or succeeding comma whenever such changes fit the length limit better.

Figure 1 shows an example where the second row beneath the body sentence indicates an assumed similarity score for each character position. Although the last comma (defining the starting position of the proper ending clause) has a similarity score 0.5, higher than the one

with 0.25, the desired ending clauses would start from the one with 0.25 since it fits the length limit better. This changes the summary candidate from “最好的一句，其實在最後一小句。”， with 15 characters in length and 0.5 in similarity, into “最好的一句，看了以後，其實在最後一小句。”， with 20 characters in length and 0.25 in similarity.

Title:

最	好	的	一	句
---	---	---	---	---

Body Sentence:

...	,	看	了	以	後	,	其	實	在	最	後	一	小	句	。
...	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	...	...	...	...	...	

**Figure 1. An example to show the need for changing the starting position for better ending clauses.**

Sometimes, the last clause of each body sentence may be too long, causing all the concatenated candidates to exceed the required length. In this case, only the news headline will be generated as the output without any concatenation.

News stories are often written in a so-called pyramid style where the later the paragraph occurs, the more details it carries. Thus, better summaries often come from the first few sentences. Therefore, in our current implementation, we decrease the similarity of the summary candidates composed from the sentences other than the first two by a factor of 0.85, if the number of sentences in the story exceeds 3.

In Step 3, the length of the summary candidate is divided by the required length limit (45 or 69) to yield the length ratio ranging from 0 to 1. Now, we come to a problem of determining the rank of these candidates based on their length ratios and similarities. Ideally, this problem can be solved by machine learning methods, but they require manually prepared data to train a classifier to determine the best or to rank the candidates. The effectiveness of such machine classifiers depends heavily on the amount of training data. Since sufficient training data are difficult to prepare, a set of hand-crafted rules are devised instead:

- (1) From the candidate list, find the candidate with highest similarity, called X, and the candidate with largest length ratio, called Y.
  - (i) If  $X=Y$ , then output X.
  - (ii) If  $sim(X) > 1.25 * sim(Y)$  and  $ratio(X) > 0.75 * ratio(Y)$ , then output X, otherwise output Y.
  - (iii) Remove the candidate just output from the list.
- (2) Repeat (1) until there is no candidate.



#### 4. Performance Evaluation

Based on the above steps, the summary candidates of length limit 45 for the third story in Table 1 are exactly the three in Table 2. The (ratio, similarity) values for the candidates are (0.7556, 0.7351), (1.0, 0.7757), and (0.8444, 0.6718), respectively. Step 3 sorts Candidates 1, 2, and 3 into 2, 3, and 1 in decreasing order of rank. As to the quality of the candidates, Candidate 2 with rank 1 is correct and coherent in meaning and is perfect in length. Candidate 3 is fair in Chinese expressions. It would become better if the word: “他們” (“they”) in the beginning of the second clause were deleted. Candidate 1 is also correct and coherent. It carries more interesting content than Candidate 2, but it is shorter in length.

To further evaluate the above method, two sets of news stories were used. One set contained 40 Chinese news stories, which were real-time news (short stories updated every 30 minutes) from China Times<sup>2</sup> between August and September in 2003. Some of the stories, together with the above examples, were used to tune the parameters mentioned in the previous section. Therefore, this group of stories can be considered an inside testing set. The other set contains 75 normal stories, also from China Times between April 4 and 12 in 2009. The parameters and programs set for the inside testing set were used for these recent stories. Thus, they can be considered our outside testing set. Table 3 shows some statistics about these stories in the two sets.

**Table 3. Statistics of the news stories in the two testing sets.**  
(a) Various averages. (b) Number of documents in each category.

Average Statistics	Inside testing set	Outside testing set
Average number of sentences per story	2.95	8.51
Average number of clauses per sentence	4.08	3.97
Average number of characters per sentence	64.54	52.54
Average number of characters per clause	15.83	13.22
Average number of characters per headline	16.63	14.87

Category	No. of Doc. in inside testing set	No. of Doc. in outside testing set
政治/politics	2	10
社會/criminals	0	10
財經/economics	14	10
國際/international affairs	6	10
科技/technology	4	10
娛樂/entertainment	10	10
運動/sports	4	15

<sup>2</sup> “China Times” <http://www.chinatimes.com.tw/>

For each story in both sets, summary candidates were generated and ranked by the proposed method. A human summarizer chose a candidate that he/she thought to be the best among the candidates. The chosen one was then labeled in terms of its quality with one of the three tags: G (good), F (fair), or B (bad) if it was correct and coherent, correct with some readability, or unacceptable, respectively. The inside testing set was evaluated by one human summarizer, while the outside testing set was evaluated by 15 people each for 5 stories. All of the evaluators majored in library science, thus, have some sense of knowledge for manual summarization.

Table 4 shows the results for the inside testing set, where for each news story the title and the best candidates for length limit 45 and 69 are shown, respectively. The actual lengths of the best candidates are shown in the second column. In the fourth column, with a \*, the number of body sentences in that story is shown in the title row, while the rank of the chosen candidate is listed besides the candidate. The last column indicates the manual judgment of the machine-generated summary.

**Table 4. The forty news headlines, their machine-generated summaries, and manual judgment of quality for the inside testing set.**

ID	Content		*	**
1	Title	台鐵計軸器採購下周進行第 11 度招標	2	
	45	台鐵計軸器採購下周進行第 11 度招標，擁有這項產品製造技術的歐洲廠商，已摩拳擦掌準備進場搶標。	1	G
	66	台鐵計軸器採購下周進行第 11 度招標，不限定廠商使用材質，下周公告招標後，等標期約 28 天、審查作業 10 天，最快 10 月中旬可以最低價格進行決標。	1	G
2	Title	台十一線濱海公路山崩，交通中斷	5	
	45	台十一線濱海公路山崩，交通中斷，造成豐濱鄉對外交通完全中斷，民眾必須往台東縣才能找到出路。	1	G
	69	台十一線濱海公路山崩，交通中斷，形成九十度丁坡度，連日來花蓮間歇性豪雨不斷，該地段今天早上九點多終於發生小規模山崩，交通中斷阻斷來往車輛。	1	G
3	Title	台鐵與工會最後協商無交集，中秋是否停駛各說各話	4	
	45	台鐵與工會最後協商無交集，中秋是否停駛各說各話，會員現在也不敢說不上班，只是應付一下主管。	1	G
	61	台鐵與工會最後協商無交集，中秋是否停駛各說各話，工會說，這是台鐵當局的一貫技倆，會員現在也不敢說不上班，只是應付一下主管。	2	G
4	Title	兩岸航空業邁進實質合作時代	1	
	43	兩岸航空業邁進實質合作時代，這項合作也正式宣布兩岸航空貨運開始走入實質合作的經營時代。	1	G
	69	兩岸航空業邁進實質合作時代，將再度齊聚廈門，出席這項兩岸航空業首度合資的盛會，這項合作也正式宣布兩岸航空貨運開始走入實質合作的經營時代。	1	B
5	Title	高市招商，力邀重量級企業與會	2	
	30	高市招商，力邀重量級企業與會，以及多功能經貿園區的未來遠景。	1	B
	57	高市招商，力邀重量級企業與會，而行程中必定會談到世界大港高雄港和小港機場的	1	G

		海空優勢，以及多功能經貿園區的未來遠景。		
6	Title	雲縣規劃產業聚落，建立招商網路	2	
	32	雲縣規劃產業聚落，建立招商網路，發展各專區內互補特性，相互支援。	1	G
	67	雲縣規劃產業聚落，建立招商網路，並規劃以麥寮自由港區、中科雲林基地及雲林科技工業區發展為三個相互支援發展的產業聚落，爭取更多企業投資。	1	G
7	Title	中油調高桶裝瓦斯價格	4	
	34	中油調高桶裝瓦斯價格，以二十公斤裝桶裝瓦斯來看，每桶批售價調高八元。	1	G
	64	中油調高桶裝瓦斯價格，為反應進口成本上漲壓力，中油決定自四日零時起調漲各類液化石油氣產品牌價，調整幅度為二·六五%至三·九四%。	1	G
8	Title	經濟部：攤販不會就地合法	1	
	29	經濟部：攤販不會就地合法，因此不會有「就地合法」這個問題。	1	B
	54	經濟部：攤販不會就地合法，未來攤販仍須先通過地方政府審核後才能獲得營業許可，因此不會有「就地合法」這個問題。	1	G
9	Title	獅、象四連戰第二役，統一獅將派出威森掛帥	3	
	43	獅、象四連戰第二役，統一獅將派出威森掛帥，親自派遣場務人員前來台北，為威森整理投手丘。	1	G
	68	獅、象四連戰第二役，統一獅將派出威森掛帥，爭取今晚間的勝利，統一特別從台南帶著「土坯」前來新莊，賽前將由工作人員親自為威森整理投手丘。	1	G
10	Title	中華職棒大聯盟，教練護盤，「劉」住勝果	5	
	42	中華職棒大聯盟，教練護盤，「劉」住勝果，戰績繼續保持第一，領先獅隊的勝差拉開為 1.5 場。	1	F
	69	中華職棒大聯盟，教練護盤，「劉」住勝果，順利終結獅隊最後反撲，拿下 1 次救援成功，距離上次（2000 年 9 月 23 日對牛隊）贏得救援成功，已將近 3 年了。	1	F
11	Title	美國網球公開賽：阿格西驚險闖進 8 強	4	
	33	美國網球公開賽：阿格西驚險闖進 8 強。阿格西遇險，險遭丹特襲擊成功。	2	G
	65	美國網球公開賽：阿格西驚險闖進 8 強；西哥畢竟老江湖，第 2 盤穩中求勝，第 3 盤守住丹特強力攻勢，終於讓小老弟因強攻不破，右腳傷重退賽。	1	G
12	Title	娜姐送吻，小甜甜人氣下滑，克莉絲汀變旺	5	
	40	娜姐送吻，小甜甜人氣下滑，克莉絲汀變旺，克莉絲汀是「一吻成名」，一夕間躍升榜首。	3	G
	50	舌吻事件這兩天在網路上引爆熱烈討論，雖然布蘭妮、克莉絲汀都被娜姐送上香吻，但人氣指數卻呈現兩個極端。	4	G
13	Title	余詩曼睡一睡，溫碧霞脫一脫，數百萬入袋	6	
	41	余詩曼睡一睡，溫碧霞脫一脫，數百萬入袋；而溫碧霞則是小脫一下，就賺到四百多萬台幣。	1	F
	67	余詩曼睡一睡，溫碧霞脫一脫，數百萬入袋，最近港星余詩曼自稱在床上睡一睡，就有六百萬台幣入袋；而溫碧霞則是小脫一下，就賺到四百多萬台幣。	1	G
14	Title	王識賢求婚很靦腆，張鳳書當老師	3	
	42	王識賢求婚很靦腆，張鳳書當老師，反倒是張鳳書教他，求婚就該在大庭廣眾下告白才有誠意。	2	F
	68	王識賢求婚很靦腆，張鳳書當老師，導演要求他下跪求婚，王識賢靦腆的說人太多，不好意思，反倒是張鳳書教他，求婚就該在大庭廣眾下告白才有誠意。	1	G

15	Title	百慕達銀行在日本開設辦事處	2	
	13	百慕達銀行在日本開設辦事處	1	B
	64	百慕達銀行在日本開設辦事處。Bermuda Global Fund Services Limited 東京辦事處將坐落於東京，並將作為百慕達銀行旗下全球範圍的 GFS 部門與其日本客戶之間的聯繫機構。	1	G
16	Title	東芝公司同意在系統單晶片中使用 ARM 晶片	3	
	45	東芝公司同意在系統單晶片中使用 ARM 晶片，雙方已經通過新的授權協議拓展了彼此間的戰略合作關係。	1	G
	69	東芝公司同意在系統單晶片中使用 ARM 晶片，東芝公司已經同意把 ARM1026EJ-S(TM)晶片用於促成創新的系統單晶片(SOC)應用產品，從而豐富其新一代數碼產品組合。	1	G
17	Title	Inno Micro 在日本經銷並出售 nStor 產品	2	
	31	Inno Micro 在日本經銷並出售 nStor 產品，在日本出售和經銷 nStor 全系列存儲產品。	1	B
	54	Inno Micro 在日本經銷並出售 nStor 產品，日本一家私營整合商和經銷商 Inno Micro 已簽署一份協議，在日本出售和經銷 nStor 全系列存儲產品。	1	F
18	Title	登記列管繳稅營業，攤販將全面合法	2	
	34	登記列管繳稅營業，攤販將全面合法，預估有數十萬攤販可望就地「合法」。	1	G
	64	登記列管繳稅營業，攤販將全面合法，將把全台灣既存和未來可能新增的攤販，全部改以登記制統一管理，預估有數十萬攤販可望就地「合法」。	1	F
19	Title	行動攤販車可在風景區營業	4	
	41	行動攤販車可在風景區營業，甚至還成立加盟總部，鼓勵民眾只要投資數十萬元就可以創業。	1	G
	56	行動攤販車可在風景區營業，包括行動咖啡館、行動彩印店等，甚至還成立加盟總部，鼓勵民眾只要投資數十萬元就可以創業。	1	F
20	Title	輕軌工業擬改採國內標	2	
	30	輕軌工業擬改採國內標，採國內外業者共同承攬但由國內業者主導。	1	G
	57	輕軌工業擬改採國內標，放寬招商「實績」要求，提高國內業者自製率比重至五〇%，採國內外業者共同承攬但由國內業者主導。	1	G
21	Title	中共採購新規定，重擊微軟	2	
	36	中共採購新規定，重擊微軟，要購買非本國軟體系統的政府單位，一律特別呈報。	2	G
	63	中共採購新規定，重擊微軟，儘管微軟大力投資當地，並改組大中華區人事，但在大陸急力扶持國產軟件下，微軟在大陸業務可能遭致命打擊。	1	G
22	Title	扶持軟體產業，中共在融資、上市和稅收方面給予優惠措施	4	
	45	扶持軟體產業，中共在融資、上市和稅收方面給予優惠措施，成立風險投資公司，設立風險投資基金。	1	G
	62	扶持軟體產業，中共在融資、上市和稅收方面給予優惠措施，以求二〇一〇年大陸的軟體產業研究開發和生產能力達到或接近國際先進水平。	1	G
23	Title	緊縮房地產業，中共加大力道	3	
	40	緊縮房地產業，中共加大力道，要控制此類項目的建設用地供應量，或暫停審批此類項目。	1	G
	68	緊縮房地產業，中共加大力道，對高檔大戶型商品房、辦公大樓與商業性用房積壓較多的地區，要控制此類項目的建設用地供應量，或暫停審批此類項目。	1	G

24	Title	陳總統：中華民國是主權獨立國家	3	
	34	陳總統：中華民國是主權獨立國家，國軍要為捍衛中華民國主權與領土而戰。	1	G
	66	外傳前總統李登輝指「陳總統只說中華民國是國號，沒有說中華民國是國家」，而陳總統昨天則向三軍官兵強調「中華民國是一個主權獨立的國家」。	3	G
25	Title	明年總統大選，藍綠基本盤皆見鬆動	3	
	45	明年總統大選，藍綠基本盤皆見鬆動，而當年的選民，歷經政黨輪替，如今投票意向已出現明顯改變。	1	G
	64	明年總統大選，藍綠基本盤皆見鬆動，上屆大選支持泛藍的選民，陣腳略微鬆動；而之前支持陳呂配的泛綠選民，也有相當比例出現流失的現象。	1	G
26	Title	競國實業董事會決議配息配股基準日為9月12日。	1	
	39	競國實業董事會決議配息配股基準日為9月12日，9月8日起至9月12日停止股票過戶。	1	G
	39	競國實業董事會決議配息配股基準日為9月12日，9月8日起至9月12日停止股票過戶。	1	G
27	Title	國眾奪下中華電北區 FTTBL2Switch 採購案	2	
	39	國眾奪下中華電北區 FTTBL2Switch 採購案，以供中華電信協助中小企業利用寬頻網路發展商機之用。	1	G
	58	國眾奪下中華電北區 FTTBL2Switch 採購案，由國眾得標，智邦集團傳易（SMC）、和心光通、飛瑞、安捷倫及浩網等廠商負責提供相關整合產品。	1	G
28	Title	亞太電信集團跨足線上遊戲，今年營收約 2500 萬元	3	
	36	亞太電信集團跨足線上遊戲，今年營收約 2500 萬元，4C 整合的佈局儼然成形。	1	G
	69	亞太電信集團跨足線上遊戲，今年營收約 2500 萬元，推出新的娛樂事業群，亞太集團版圖橫跨了電信、網路、通訊、加值內容，4C 整合的佈局儼然成形。	1	G
29	Title	亞太電信推出「猿人在線」品牌，初期以代理為主。	3	
	31	亞太電信推出「猿人在線」品牌，初期以代理為主，朝線上遊戲邁進。	2	G
	53	亞太電信推出「猿人在線」品牌，初期以代理為主，因此結合集團內各式寬頻服務載具與平台的資源，朝線上遊戲邁進。	1	F
30	Title	友達第五代彩色濾光片廠十月起逐步量產，最大月產能 12 萬片	3	
	43	友達第五代彩色濾光片廠十月起逐步量產，最大月產能 12 萬片，使友達有效掌握上游關鍵零組件。	1	G
	64	友達第五代彩色濾光片廠十月起逐步量產，最大月產能 12 萬片，月產能 7 萬片，預估未來每月最大產能 12 萬片玻璃基板，供全球大尺寸面板需求。	1	F
31	Title	中壽投資型商品「一觸得利」狂賣，一周銷售達 13 億元	5	
	40	中壽投資型商品「一觸得利」狂賣，一周銷售達 13 億元，不僅為業界首創，引發熱賣風潮。	2	G
	57	中壽投資型商品「一觸得利」狂賣，一周銷售達 13 億元，投資標的為逆浮動+正浮動利率債券，不僅為業界首創，引發熱賣風潮。	2	G
32	Title	29 日台積電 ADR 收盤價 11.78 美元，較前交易日上漲 0.08 美元。	1	
	42	29 日台積電 ADR 收盤價 11.78 美元，較前交易日上漲 0.08 美元，漲幅為 0.68%，換算回台股每股價格約 80.54 元。	1	G
	54	29 日台積電 ADR 收盤價 11.78 美元，較前交易日上漲 0.08 美元，較前一交易日上漲 0.08 美元，漲幅為 0.68%，換算回台股每股價格約 80.54 元。	1	B

33	Title	「美夢成真」趕戲，葉全真累壞吊了點滴再上	3	
	45	「美夢成真」趕戲，葉全真累壞吊了點滴再上，不顧醫生要她吊點滴多休息的叮嚀，又回棚內拍戲去。	1	G
	59	「美夢成真」趕戲，葉全真累壞吊了點滴再上，所以她在打了兩劑粗血管針後，不顧醫生要她吊點滴多休息的叮嚀，又回棚內拍戲去。	1	B
34	Title	八點檔現拍現播，演員連連發病	4	
	43	八點檔現拍現播，演員連連發病，除了中視、華視，其餘三台都以現拍現播的方式，走本土路線。	1	G
	58	八點檔現拍現播，演員連連發病。演員日夜趕戲來趕播出，體力已受考驗，偏偏表演方式更耗費體力，病號、傷兵也因此連連爆發。	2	G
35	Title	周俊三蹲牢房，代價很值得	2	
	35	周俊三蹲牢房，代價很值得，辛苦還是有代價的，讓他獲得 3 萬元的豐厚酬勞。	1	G
	35	周俊三蹲牢房，代價很值得，辛苦還是有代價的，讓他獲得 3 萬元的豐厚酬勞。	1	G
36	Title	佼佼訪王貞治，豪華日本行。	1	
	45	佼佼訪王貞治，豪華日本行，還住在一晚高達 6 萬日幣的飯店裡，且如願吃到頂級的佐賀牛肉壽喜燒。	1	G
	67	佼佼訪王貞治，豪華日本行，除了能親眼目睹日本職棒，專訪職棒明星王貞治，還住在一晚高達 6 萬日幣的飯店裡，且如願吃到頂級的佐賀牛肉壽喜燒。	1	G
37	Title	「棋靈王圍棋入門之旅」活動開跑	3	
	34	「棋靈王圍棋入門之旅」活動開跑，使得圍棋儼然成為最新的全民益智運動。	1	G
	61	「棋靈王圍棋入門之旅」活動開跑，再加上不久前奪得今年日本本因坊頭銜的旅日棋手張栩效應，使得圍棋儼然成為最新的全民益智運動。	1	G
38	Title	周末官邸藝文沙龍，王瑁邀親子無言的交流	5	
	35	周末官邸藝文沙龍，王瑁邀親子無言的交流，激發出親子間的想像力與創造力！	1	G
	60	周末官邸藝文沙龍，王瑁邀親子無言的交流，並且藉由各式精心設計的劇場遊戲—模仿、帶領、互動，激發出親子間的想像力與創造力！	1	G
39	Title	故宮德國文物大展，開放展場設計權	3	
	39	故宮德國文物大展，開放展場設計權，舉辦公開說明會，歡迎設計師與建築師前來參與。	2	G
	65	故宮德國文物大展，開放展場設計權，故宮破天荒將公開舉辦展場競圖，預計本月 15 日下午 2 點，舉辦公開說明會，歡迎設計師與建築師前來參與。	1	G
40	Title	藝文界前輩進駐為豐樂童畫賽暖身	1	
	15	藝文界前輩進駐為豐樂童畫賽暖身	1	B
	15	藝文界前輩進駐為豐樂童畫賽暖身	1	B

Table 5 summarizes the data shown in Table 4. As can be seen, of the 40 stories, 65.0% or 62.5% of the first candidates suggested by the method for the length limit 45 and 69, respectively, were judged good. If users were able to choose from all the suggested candidates, 80% or 75% of the summaries could be obtained from a machine without manual editing. Only about 12.5% or 10% of the stories yielded summaries that were unacceptable.

**Table 5. Quality statistics for the summary candidates of the inside testing set. (a) The upper table is for length limit 45. (b) The lower table is for length limit 69.**

Quality Rank	Good	Fair	Bad
1	26 (65.0%)	2 (5.0%)	5 (12.5%)
2	5 (12.5%)	1 (2.5%)	0
3	1 (2.5%)	0	0
4	0	0	0
5	0	0	0
total	32 (80.0%)	3 (7.5%)	5 (12.5%)

Quality Rank	Good	Fair	Bad
1	25 (62.5%)	6 (15%)	4 (10%)
2	3 (7.5%)	0	0
3	1 (2.5%)	0	0
4	1 (2.5%)	0	0
5	0	0	0
total	30 (75.0%)	6 (15%)	4 (10%)

The best candidates that are unacceptable (9 cases in total in news ID 4, 5, 8, 15, 17, 32, 33, and 40) contain undesired conjunctions that break the coherence and/or readability (4 cases in 4, 5, 8, 33), clauses that duplicate the headline strings (2 cases in 17 and 32), or were nothing but the headline itself, which means that no candidates could be generated under the required length limit (3 cases in 15 and 40). The suitability of conjunctions for direct concatenation is difficult to judge, because some of them are helpful and some are not. The cases of headline duplication can be eliminated by duplication detection before concatenation. As to those candidates that contain only headlines, the clauses can be broken into smaller structures, such as phrases, for re-combination. This, however, would be a more difficult problem that would require more language analysis.

Table 6 summarizes the results for the outside testing set. As can be seen, the percentages of the first suggested candidates that were judged good decrease from 65% and 62.5% to 18.7% and 33.3%, respectively, for the length limit 45 and 69. The percentages that were judged good regardless of the rank position decreased from 80% to 53.3% for the length limit 45 and from 75% to 70.7% for the length limit 69, showing that the shorter the length limit, the less stable the method in performance. A large portion of the percentage moves to those

that were judged fair. This decrease in performance may due to the greater number of body sentences and the larger number of evaluators for the outside testing set. As more candidates (and evaluators) exist for selection, less coincidence exists for the same choice (and judgment) results. The only consistent result (compared to the inside testing set) is that those best candidates that were judged bad are still rare (less than 10%). This shows that the heuristic: “concatenating the last clauses of the body sentence with the headline” seems to work for Chinese news in this application.

**Table 6. Quality statistics for the summary candidates of the outside testing set. (a) The upper table is for length limit 45. (b) The lower table is for length limit 69.**

Quality Rank	Good	Fair	Bad
1	14 (18.7%)	11 (14.7%)	2 (2.7%)
2	14 (18.7%)	8 (10.7%)	0 (0.00%)
3	8 (10.7%)	5 ( 6.7%)	2 (2.7%)
4	4 ( 5.3%)	1 ( 1.3%)	0 (0.00%)
5	0 ( 0.00%)	4 ( 5.3%)	2 (2.7%)
total	40 (53.3%)	29 (38.7%)	6 (8.00%)

Quality Rank	Good	Fair	Bad
1	25 (33.3%)	5 ( 6.7%)	3 (4.0%)
2	13 (17.3%)	8 (10.7%)	1 (1.3%)
3	7 (9.3%)	3 ( 4.0%)	0 (0.0%)
4	4 (5.3%)	1 ( 1.3%)	0 (0.0%)
5	4 (5.3%)	1 ( 1.3%)	0 (0.0%)
total	53 (70.7%)	18 (24.0%)	4 (5.3%)

## 5. Discussion

The fact that the proposed method works for some stories is due to the characteristics of Chinese news. They tell stories in a successive sequence. Very few grammatical inversions within sentences and clauses are used. Chinese words have virtually no morphological variations. The clauses, especially at the rear part of a sentence, are sometimes quite independent of the front part. Headlines are given in a compact form to cover as many important facets as possible, such as who, what, where, when, why, and how. All of these characteristics make clause recombination a choice for summary generation. With this



heuristic strategy, the remaining work is to evaluate their fitness as summaries and rank them in a correct sense. For the news stories we tested, the proposed method applies to most of them with success. Nonetheless, for stories not of this type, such as editorials, commentaries, and lists of events, items, or prices, this method may fail. For the stories whose headlines are more eye-catching rather than informative, such that most content words do not appear in the headlines, this method may fail as well.

## **6. Conclusions**

The proposed method recombines snippets of news without modifying them. A direct advantage is that other synchronized media such as images, speech, or video of the same story can maintain synchronization with ease when they are summarized as well (like those in (ANSES, n.d.)), because the positions of where to cut and paste are known during the generation of the summary candidates. Thus, to achieve speech or video segmentation and summarization for similar services, one can use their synchronized texts based on this method.

Other practical advantages of this computer-assisted summarization include the ease of maintaining summary quality regardless of the experience of human summarizers and the reduction in the cost and time to train novices for this kind of services.

Evaluation of the quality of auto-generated summaries requires human judgment and is, thus, expensive and time-consuming for large-scale or multiple-run evaluation. To allow automatic evaluation using the methodology like those used in machine translation (Papineni, Roukos, Ward, & Zhu, 2002; Doddington, 2002), a number of test collections need to be created. Our past research projects in Chinese OCR text retrieval and Chinese document classification have results in two corresponding test collections for free use (Tseng, 2002; 2004). We hope that we can also release a Chinese collection for evaluating automatic summarization in the future.

## **Acknowledgement**

Special thanks are to the anonymous reviewers for their valuable comments and suggestions to improve the quality of this paper.

This work is partly supported by WebGenie Information Ltd. and National Science Council under the grants numbered: NSC 93-2213-E-030-007- and NSC 97-2631-S-003-003-.

## **References**

ANSES : Automatic News Summarization and Extraction System : Knowledge Media Institute : The Open University. Available: <http://technologies.kmi.open.ac.uk/anses/> [Accessed: July 22, 2009].

- Barzilay, R. & McKeown, K. R. (2005). Sentence Fusion for Multidocument News Summarization. *Computational Linguistics*, 31(3), 297-327.
- Carbonell, J. & Goldstein, J. (1998). The use of MMR, diversity-based re-ranking for reordering documents and producing summaries. In *Proceedings of SIGIR*, Melbourne, Australia, 335-336.
- Central News Agency. Price List for Stock News Brief Message from Central News Agency. (in Chinese). Available: [http://www.suio.com.tw/top/can/can\\_order\\_txt.asp](http://www.suio.com.tw/top/can/can_order_txt.asp) [Accessed: July 22, 2009]
- Chen, H.-H., Kuo, J.-J., Huang, S.-J., Lin, C.-J., & Wung, H.-C. (2003). A summarization system for Chinese news from multiple sources. *Journal of the American Society for Information Science and Technology*, 54(13), 1224-1236.
- China Times. Media Challenges: Multi-modal communications. (in Chinese). Available: [http://marketing.chinatimes.com/item\\_detail\\_page/professional\\_columnist/professional\\_columnist\\_content\\_by\\_author.asp?MMContentNoID=4369](http://marketing.chinatimes.com/item_detail_page/professional_columnist/professional_columnist_content_by_author.asp?MMContentNoID=4369) [Accessed: Dec. 3, 2003].
- Document Understanding Conferences, Available: <http://duc.nist.gov/> [Accessed: July 22, 2009]
- Doddington, G. (2002). Automatic Evaluation of Machine Translation Quality using N-gram Co-occurrence Statistics. In *Proceedings of the second international conference on Human Language Technology Research*, San Diego, California, March, 2002, 138-145.
- Evans, D. K., Klavans, J. L., & McKeown, K. R. (2004). Columbia Newsblaster: Multilingual News Summarization on the Web. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, HLT-NAACL 2004: Demonstration Papers, May 2-7, 2004, Boston USA, 1-4.
- Fukushima, T., Okumura, M., & Nanba, H. (2002). Text Summarization Challenge 2: Text Summarization Evaluation at NTCIR Workshop3. In *Proceedings of the Third NTCIR Workshop on Evaluation of Information Retrieval, Automatic Text Summarization and Question Answering*, Oct. 8-10, 2002, Tokyo, Japan, 1-6.
- Hovy, E. H., & Lin, C.-Y. (1999). *Automated Text Summarization in SUMMARIST*. In I. Mani and M. Maybury, editors, *Advances in Automated Text Summarization*, chapter 8. MIT Press.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8), 707-710.
- Lin, C.-Y. & Hovy, E. H. (2002). From single to multi-document summarization: A prototype system and its evaluation. In *Proceedings of the ACL*, Philadelphia, PA, 2002, 457-464.
- Lin, C.-Y. & Hovy, E. H. (2003). The Potential and Limitations of Sentence Extraction for Summarization. In *Proceedings of the Workshop on Automatic Summarization, post-conference workshop of HLT-NAACL-2003*, Edmonton, Canada, May 31 - June 1, 2003.

- Lopresti, D. & Zhou, J. (1996). Retrieval Strategies for Noisy Text. In *Proceedings of the Fifth Annual Symposium on Document Analysis and Information Retrieval*, April 15-17, 1996, 255-269.
- Inderjeet Mani, *Automatic Summarization*, John Benjamins, 2001.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). Bleu: A method for automatic evaluation of machine translation. In *Proceedings of the ACL*, Philadelphia, PA, 311-318.
- Radev, D. R., Hovy, E., & McKeown, K. (2002). Introduction to the Special Issue on Summarization. *Computational Linguistics*, 28(4), 399-408.
- Radev, D. R., Otterbacher, J., Winkel, A., & Blair-Goldensohn, S. (2005). NewsInEssence: summarizing online news topics. *Communications of the ACM*, 48(10), 95-98.
- Takefumi, O., Hidetaka, M., & Hiroshi, N. (2003). Web News Article Summarization and its Evaluation using Articles for Mobile Terminals. *Joho Shori Gakkai Kenkyu Hokoku*, 2003(4), 1-8. (in Japanese)
- Tseng, Y.-H. (2002). Automatic Thesaurus Generation for Chinese Documents. *Journal of the American Society for Information Science and Technology*, 53(13), 1130-1138.
- Tseng, Y.-H. (2002). FJU Test Collection for Evaluation of Chinese OCR Text Retrieval. Available: [http://www.lins.fju.edu.tw/~tseng/Collections/Chinese\\_OCR\\_IR.html](http://www.lins.fju.edu.tw/~tseng/Collections/Chinese_OCR_IR.html) [Accessed: July 22, 2009].
- Tseng, Y.-H. (2004). FJU Test Collection for Evaluation of Chinese Text Categorization. Available: [http://www.lins.fju.edu.tw/~tseng/Collections/Chinese\\_TC.html](http://www.lins.fju.edu.tw/~tseng/Collections/Chinese_TC.html) [Accessed: July 22, 2009].
- Tseng, Y.-H., Wang, Y.-M., Lin, Y.-I., Lin, C.-J., & Juang, D.-W. (2007). Patent Surrogate Extraction and Evaluation in the Context of Patent Mapping. *Journal of Information Science*, 33(6), 718-736.
- United Daily News*. egolife: Brief Message Delivery for Cell Phone. (in Chinese) Available: [http://udn.com/NASApp/LogFriend/UDNSMS/introduction\\_news.html](http://udn.com/NASApp/LogFriend/UDNSMS/introduction_news.html) [Accessed: Jan. 2, 2005]
- Wang, C. (2002). Number One in the World: Cell Phone Numbers used More Than the Populations of Taiwan. (in Chinese) ETtoday.com, 2002/08/09. Available: <http://www.ettoday.com/2002/08/09/339-1337800.htm> [Accessed: July 22, 2009].

