

NewsCube: Delivering Multiple Aspects of News to Mitigate Media Bias

Souneil Park, Seungwoo Kang, Sangyoung Chung, Junehwa Song

Computer Science, KAIST

Daejeon, 305-701, Republic of Korea

{spark, swkang, sychung, junesong}@nclab.kaist.ac.kr

ABSTRACT

The bias in the news media is an inherent flaw of the news production process. The resulting bias often causes a sharp increase in political polarization and in the cost of conflict on social issues such as Iraq war. It is very difficult, if not impossible, for readers to have penetrating views on realities against such bias. This paper presents NewsCube, a novel Internet news service aiming at mitigating the effect of media bias. NewsCube automatically creates and promptly provides readers with multiple classified viewpoints on a news event of interest. As such, it effectively helps readers understand a fact from a plural of viewpoints and formulate their own, more balanced viewpoints. While media bias problem has been studied extensively in communications and social sciences, our work is the first to develop a news service as a solution and study its effect. We discuss the effect of the service through various user studies.

Author Keywords

Media Bias, News, Aspect-level Browsing, News Distribution Service.

ACM Classification Keywords

K.4 Computers and Society: Social Issues, H.4 Information Systems Applications: Communication Applications

INTRODUCTION

Bias of the news media is an inherent flaw of the news production process, spanning news gathering, writing, and editing stages. At every single stage, news is probably never free from the producer's subjective valuation and external forces such as owners and advertisers. Such valuation and forces often lead to totally different *framing*¹ of realities. Producers select different aspects of reality; they cover these aspects in different tones; they present them in different styles. The resulting bias causes a sharp increase in political polarization, misunderstanding of critical issues such as the Iraq war [14], and a significant impact upon elections [4]. It is very difficult for readers to have a balanced view on realities. They mostly receive

news from restricted channels, often times through a single channel. Advanced news production and delivery models are required to break the vicious cycle of bias.

In this paper, we present *NewsCube*, a novel Internet news service that mitigates the effects of media bias. NewsCube automatically creates and promptly provides readers with multiple classified viewpoints on a news event of interest. It helps readers easily discover rich facts and compare diverse biased views of the event. As such, it effectively helps readers understand news events from a plural of viewpoints and formulate their own, more balanced viewpoints free from specific biased views. Dealing with the media bias problem from the consumers' perspective has clear advantages over doing it from the producers' perspective. News producers have made considerable efforts to reduce bias. They established and observed journalism ethics and standards². More recently, producers of radio and TV news have adopted an adversarial format such as point-counterpoint roundtable discussions³. The format attempts to give readers a balanced perspective by incorporating multiple viewpoints and opinions on a news event. However, such efforts have not succeeded in fully resolving the problem; media bias is still widespread [24]. They do not address the underlying structural factors [12] of media production. Moreover, they cannot ensure that the inherent bias of producers does not affect the framing of articles or roundtable discussions. Addressing the structural factors of media bias is a difficult, large-scale problem. It requires significant changes in the complex relationships between government, business and media over a long term. NewsCube, however, offers a practical means of mitigating media bias that can be applied immediately.

NewsCube effectively achieves its goal by classifying aspects, presenting them fairly, and recommending articles with contrasting aspects. The core of NewsCube is *aspect-level browsing*, a new method to provide readers with a classified view of a set of articles with different aspects. News producers frame news events by covering different aspects at different depths depending on their viewpoints [20]. Thus, to provide readers with diverse viewpoints on a

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.

Copyright 2009 ACM 978-1-60558-246-7/09/04...\$5.00.

¹To frame[6] is to “select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation and/or treatment recommendation”

² http://en.wikipedia.org/wiki/Journalism_ethics.

³ http://en.wikipedia.org/wiki/Media_bias#Efforts_to_correct_bias.

news event, we need to compare and contrast news articles published by multiple producers in terms of aspects. NewsCube collects different news articles reporting on an event from various sources. It identifies the salient aspects of articles, and presents readers with articles grouped together by the similarity of their aspects.

We believe that the proposed service using aspect-level browsing provides a firm basis with which to mitigate the media bias problem using computational techniques. However, effectively realizing aspect-level browsing is a technically involved problem, mainly due to the lack of semantic knowledge with which to abstract and classify the intended salient aspects of articles. We approach the problem with news structure-based extraction. News structure-based extraction exploits well-known news writing rules to automatically abstract the aspects of an article. Taking such rules into consideration, we observe that a news article is usually written to fit an inverted pyramid structure (Figure 6), and the most important contents are covered early in the article. As this rule is generally applied in news writing, news structure-based extraction can broadly be applied to many articles regardless of their topics and languages.

We conduct in-depth user studies to evaluate the effectiveness of NewsCube in terms of media bias mitigation. Through a clickstream analysis and a questionnaire, we demonstrate that the service effectively helps people read diverse aspects. We also discuss the potential impact on formulating users' own, more balanced viewpoints. Moreover, based on a focus group interview, we present how acknowledged aspects influence the viewpoints of the interviewees. In addition, we experiment and demonstrate that aspect-level classification of news articles is feasible, i.e., commonly agreeable results can be made for aspect-level classification. Finally, we study the performance of various clustering methods to realize aspect-level classification.

The contribution of this paper is in introducing an important problem to the HCI community, namely, the media bias problem, and establishing a framework for discussion. While media bias has been studied extensively in mass communication, this work is the first attempt to develop such a framework and elaborate a computational approach towards the problem. We design and develop NewsCube service as a first-round solution and study its effects. As the problem is new and complicated, the solution leaves much room for further improvement.

MEDIA BIAS PROBLEM AND RELATED WORK

Media Bias Problem

Many studies have shown the inherency of bias of news providers [11][12]. They argue that media bias is rather an inevitable structural problem of the news industry than a temporary phenomenon. During production, producers are influenced by their own political or ideological views as well as external forces such as interests of owners, advertisers, or government and corporate information

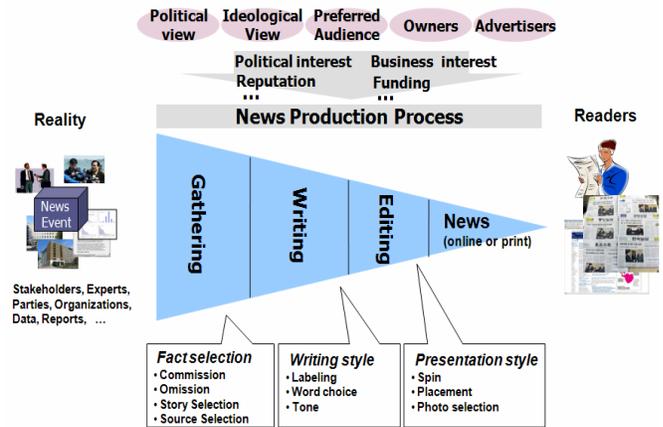


Figure 1. Causes and forms of media bias

sources [11]. Also, individual journalists are intrinsically influenced by their personal agenda and outlook.

Various causes and resulting biases have been studied. Political biases, e.g. liberal vs. conservative or left vs. right, are typically identified over many policy issues such as health care [18]. It was also argued that reliance on government or corporate information sources caused biased news written in favor of their interests [12]. Gentzkow et al. [8] have pointed out another major factor of bias, i.e., reputation of news providers on consumers, and showed that the providers slant news to meet consumers' views. McChesney's argument [16] adds class bias, which is caused by advertisers' preference on people with high income. He argues that class bias becomes more serious as media market is deregulated and advertisers have more influence on the market accordingly.

The significant impacts of media bias have also been reported. It distorts users' awareness on critical public issues and influences social decision making processes such as elections. Bernhardt et al. [4] studied the effect of media bias on political polarization and electoral outcomes. Kull et al. [14] reported misperceptions of the public specifically on the Iraq war, e.g., links between Iraq and al Qaeda, and their relation to individuals' primary news source. Gerber et al. [9] showed the influence of biased news on subscribers by conducting a field experiment during an election period. As the impact is severe, the public have been losing trust on the media. A survey by the American Society of Newspaper Editors revealed that 78 percent of the public believed that there was bias in news reporting [24]. It is also argued that media bias becomes persistent as the public often connive at bias and continue subscription despite of their distrust of the media [3].

Baker [2] has studied the different forms of embodiment of such biases in news articles. We map the forms of bias embodiments along the news production stages. (Figure 1) In the gathering stage, certain news stories are selected for publication while others are not. Detailed facts may be omitted or committed, i.e., passing along facts that support a certain viewpoint. In addition, information sources, e.g., stakeholders, experts, or think tanks may be subjectively

selected. The biases are articulated in the writing stage through labeling, word choice and tone, etc. Positive or negative words, words with a particular connotation, or extreme labels such as "far left", "far right" can be chosen. In the editing stage, biased views are further embodied indirectly by allocating the spaces for articles and their locations, and selecting photos. Certain news articles get more space and are located in more important places, e.g., front page, relative to other articles.

Solution Space for Media Bias Problem

The media bias problem is highly complicated; it seems very hard to develop a clear solution. We speculate on the potential directions of solution approaches and review the existing research efforts.

The potential approaches to the media bias problem can be viewed at two different points, i.e., production .vs. post-production stage. While the production stage approaches try to avoid the creation of bias, the post-production stage approaches deal with the created biases of produced contents. As described above, news producers have made production stage approaches by controlling their production process or changing the reporting formats. The previously mentioned journalism ethics and standards can be understood as an example self-control mechanism. However, the pursued ideas such as 'objectiveness' and 'fairness' are rather abstract and ambiguous. In addition, there is no justifiable means to enforce such control mechanisms to news producers. Different formats of reporting such as the adversarial format are being developed but they do not address the structural factors of news production and are not adopted widely. For example, the adversarial format can be adopted for only a few news events since it takes much time and efforts to report a single event in such a format; it requires careful, right and fair preparation in choosing representatives, preparing agendas and questions, and arbitrating and editing their comments.

We classify the post-production stage approaches into bias diagnosis, bias measurement, bias correction, and bias mitigation. Bias diagnosis approaches try to identify and analyze the bias of produced news articles. A number of researches in mass communication [6][23][18] and watchdog (e.g., www.fair.org) groups have tried to analyze bias in specific news events. They select a certain important public issue, e.g. health-care reform debate, and study how media framed them. These efforts heavily rely on manual analysis and can only cover few news events.

The goal of bias measurement is to quantitatively model the degree of bias. Developing an objective measure is very controversial due to the subjective nature of bias. A few works can be considered as a limited trial to approximate the degree under strong assumptions. Groseclose et al. [10] attempted to mark the position of news providers by comparing political orientation of referring think tanks. Similarly, Gamon et al. [7] tried to approximate the degree for news articles by considering the political orientation of blogs linking to the articles. Both works are based on the

assumption that the political orientation of think tanks or blogs is static and known apriori.

Correction of bias seems far more difficult than other approaches. It is hard or impossible even to clearly define what an unbiased news is. The approaches used by collective authoring communities such as Wikipedia [13] seem somewhat useful. They collect evaluations on produced contents from the public and revise biased contents. Skewz.com and NewsTrust.net also utilize mass collaboration by collecting users' evaluations on news articles. Nonetheless, such approaches are often impractical to be applied considering immediacy of news articles and their fast production cycles. Raising timely and prompt participation for evaluations and reviews and reflecting the multitudes of opinions in a short period of time are a challenging task.

Bias mitigation takes a practical approach while admitting the prevalence of biases. The approach attempts to reduce the effect of the biases on the reader's experience rather than to clearly prove or correct them. As described, it is intrinsically difficult to objectively define what bias is and so is to measure or correct it. Bias could be diagnosed to some extent based on in-depth analysis; however, such analysis takes too much time and effort to catch up the news production cycle. The approach proposed in this paper is to provide users with tools for active interactions with which they easily discover and experience diversity of existing biased views on an event. It supports users themselves to develop their own balanced views. Importantly, the delivery process is automated to quickly handle recently published news and keep up the news production cycle.

The proposed NewsCube service providing aspect-level browsing has high potential to reveal and mitigate bias. Entman [6] and Rhee [18] indicate that selection of aspects is central in the development of news frames. Also, many of the previously mentioned forms of bias embodiment such as commission, omission, story selection, and source selection are expressed by selecting and emphasizing certain aspects while deemphasizing other aspects [2]. Classifying and contrasting the aspects in news articles can successfully expose these types of bias and mitigate its effects.

News Article Processing

Much research effort has been made on news articles in various topics; news clustering [21], topic detection and tracking on news stories [1], news summarization [25], news categorization [19], etc. They have mainly targeted at helping users efficiently browse a vast quantity of news articles and easily acquire the intended information. Hence, these researches mostly saw the relation between news documents and readers, and did not look deeply into how the problems in the news production process are propagated to the readers.

Recently, many personalized news services and news-of-the-future projects are gaining much attention. The services such as Reddit.com and DailyMe.com collect and deliver news articles reflecting reader's personal interests.

The news-of-the-future projects include participatory news media such as OhmyNews [26] and social news sites such as digg.com. These media aim at supporting users to actively produce and share news stories. We believe that our work gives an important new direction for these efforts, i.e., media bias mitigation.

MODELING ASPECTS OF NEWS EVENTS

Different biased views of reporters result in different news storylines for an identical event. The reporters select different aspects of the event for their news storylines [6]. We conceptualize ‘aspect’ of an event to be Rhee’s [18] “thematic proposition”, composed of the elements of news storyline in Figure 2. The proposition has the element ‘agent’ as its subject, and other elements as its predicate. For example, an article reporting a presidential debate may mainly articulate the aspect “Obama has proposed a tax cut plan”, where Obama becomes the agent and the tax cut proposal becomes the action. An article may articulate multiple aspects. The article on the debate may further include another aspect “McCain criticized its efficacy” in the later part.

<p>Main agent: Who does the act? (need not to be a human) Action: What is the agent’s action about? Agency: How is the agent’s action done? Scene: What is the agent’s background of the action? Implications: What are the consequences?</p>
--

Figure 2. Elements of story-line

<p>1. Title: Police, Soldiers Evict Protesters from U.S. Base Site “Thousands of South Korean police and soldiers on Thursday evicted the remaining occupants and anti-U.S. civic activists from an area designated for expanded U.S. military facilities south of Seoul. ...”</p> <p>2. Title: More Violence Looms in Planned Rally at U.S. Base Site “Another violent shutdown looms as groups protesting against the relocation of U.S. military bases to Pyeongtaek, Gyeonggi Province are gearing up for a massive 10,000-strong protest this weekend...”</p>
--

Figure 3. Example articles

Selecting and articulating different aspects can lead to totally different news storylines. Figure 3 shows two news articles reporting the same event but choosing different aspects in their storylines. In the example, the first article selects the police as a main agent, and describes their stern actions evicting the protesters. However, the second article selects the protesters as a main agent, and describes their plan for a violent protest.

Extracting such semantic information, i.e., aspects, becomes very challenging when it has to be performed computationally. To abstract the aspects of articles correctly, the main agent has to be recognized and the meaning of the related descriptions have to be understood. It requires knowledge on news events which is much more complex semantic knowledge than just the meaning of words, e.g., characteristics of agents, related news events in the past, etc. News topics generally span over various broad domains and they are newly generated every day. It is impractical to encode and update such broad semantic knowledge.

Syntactic analysis of news texts can be considered as a potential approach for aspect extraction. The syntactic

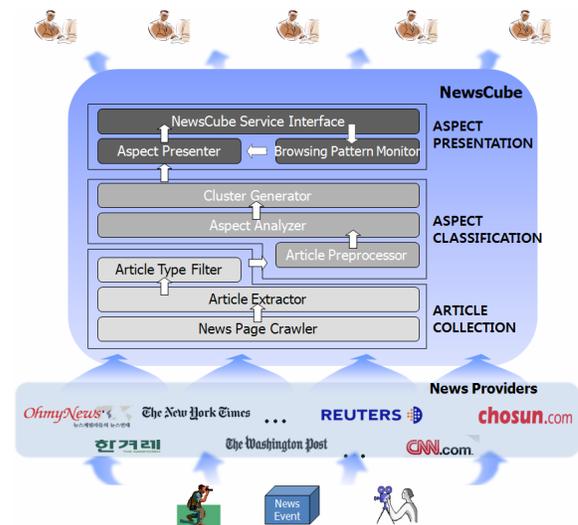


Figure 4. NewsCube Architecture

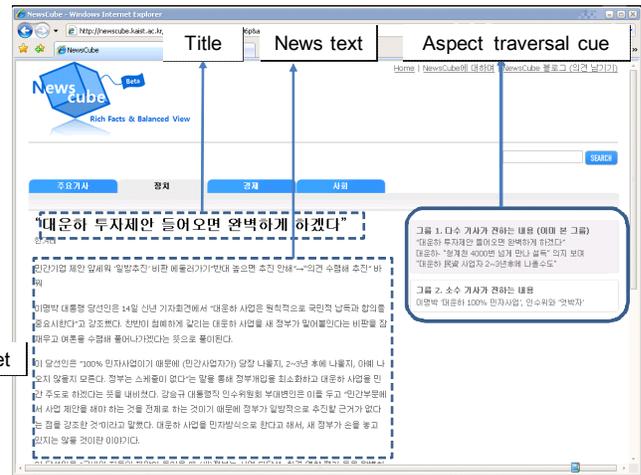
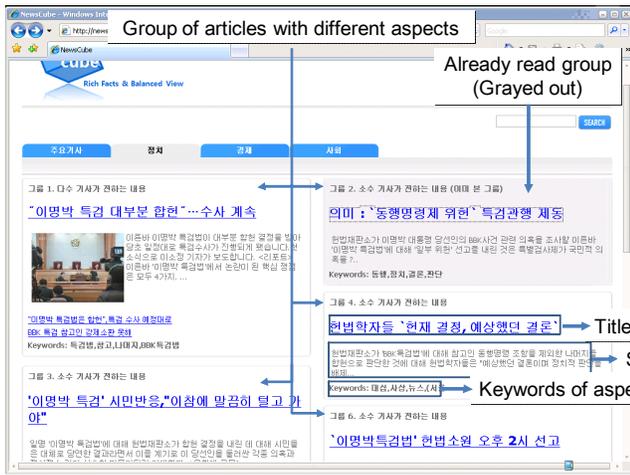
structure of news texts can give hints on the main agent and related descriptions. For example, the subjects and corresponding modifiers of sentences may be mapped to agents and descriptions, respectively. However, it is also hard to implement an accurate syntactic parser which can handle complex sentences of news. It requires sophisticated syntactic rules which can cover diverse sentence structures and disambiguation rules to resolve structural ambiguity.

To be practical, we adjusted the concept to facilitate computational approximation and devised ‘computational aspect’ based on a keyword-weight model. We note that the aspects of news articles are usually reflected in keywords in the articles. In the above example articles, core keywords such as “police, soldiers” and “violent shutdown, protesting groups” directly reveal the main aspect of each article, respectively. Therefore, we try to approximate aspects by selecting keywords and carefully calculating their importance. Although this may not fully capture the semantic meaning of texts, it is possible to distinguish difference of covered aspects by sorting out important keywords and comparing them. Our evaluation supports that the model can distinguish aspects of articles.

NEWSCUBE DESIGN CHALLENGES & APPROACHES

The primary objective in designing NewsCube is to give readers a quick understanding of diverse aspects and guide them to balanced reading. It also aims at automating the service and matching the fast news production cycles. We have developed and deployed a prototype system and it has been providing service in campus since Jan. 6, 2008. At present, the system is working on Korean news sources and news events in politics, business, and local sections are covered. Automating the NewsCube service raises a number of challenges throughout the entire process of the service: collection, classification, and presentation. We first present the overall architecture of NewsCube and discuss the challenges and approaches of each stage.

Figure 4 shows the architecture of NewsCube in high level. It consists of three major parts, i.e., collection, classification, and presentation. The collection part is



(a) Aspect view (after selecting an event)

(b) Article view

Figure 5. NewsCube interface

responsible for news page crawling and news article extraction. The Article Extractor identifies and extracts pure news text from the collected news pages, filtering out irrelevant contents such as advertisements, reader's comments, and meta-data. The Article Type Filter detects and forwards straight news articles to Article Preprocessor. The Article Preprocessor constructs annotated news text by tagging various metadata on each news article text, for example, structural information such as “<Title>” or “<SubTitle>” and morphological class of keywords. The Aspect Analyzer conducts textual as well as structural analysis on the annotated text and performs news structure-based extraction.

Aspect Presentation

For mitigation of bias, presentations should actually lead users to read diverse aspects of a news event. Users have to perceive the existence of different aspects of the event and be motivated to click different articles. Not to skew users' visual attention, articles with different aspects should be presented fairly in terms of placement and amount of allocated space. Creating such presentations becomes complex since it is impractical to allocate space equally to dozens of articles and locate them in eye-catching places in a limited space. Certain types of lay-out scheme are necessary to utilize the limited space efficiently, e.g., hiding, filtering, or ranking of articles with similar or redundant aspects.

Aspect-level browsing interface of NewsCube helps users view diverse aspects at a glance. For a selected news event, it gives a clustered overview of the aspects of the event (Figure 5(a)). Users can easily find different aspects in different article groups. To distribute users' attention evenly, each group is allocated with similar amount of space and is positioned in a 2-column grid layout. For space saving, articles in the same group covering similar aspects are hidden when necessary. Users can grasp aspects of each group through titles, snippet of news articles, and a few keywords which mainly reflect the aspects. Note that the interface of NewsCube is different from existing news

aggregation services. The existing services such as Google News, Yahoo News, and NewsInEssence [17] are limited to supporting simple event-level browsing, i.e., articles of an event are simply listed in an arbitrary order.

NewsCube further provides an aspect traversal cue to awaken and guide users for unbiased reading at every page they visit. The aspect traversal cue is a compact representation of different aspects of an event superimposed with users' traversal history. (See the right corner of Figure 5(b)) It traces and grey outs the read aspects. It helps users easily be conscious of the unread aspects and encourage them to read different aspects.

Aspect-level Classification

Classification of aspects is central to support aspect-level browsing. NewsCube approaches this problem with an unsupervised classification method, i.e., clustering. It is hard to develop and train a supervised classification method since it is impossible to predict and construct a pre-defined category of aspects for a news event.

NewsCube is the first attempt to cluster articles in aspect-level. To realize aspect-level clustering, aspects of news articles have to be extracted and articles have to be grouped regarding the similarity of the extracted aspects. We develop an aspect extraction method named *news structure-based extraction*. The method is currently customized to effectively handle straight news, the most common type of news.

Aspect Extraction

An aspect extraction method has to be able to analyze and extract aspects covered in news articles. Topics or covered events are often recognized and compared among articles by counting keywords and comparing their frequency of appearance [1]. Aspects are more detailed information than topics or covered events. Accordingly, more delicate methods are required to successfully identify and distinguish covered aspects of articles.

News structure-based extraction exploits one of the famous news writing rules, called the 'inverted pyramid style of news writing'. The rule guides journalists to

arrange information and facts in the descending order of importance within an article. From this rule, we can get conclusive hints for aspect extraction; where the semantically meaningful keywords are located and how much they are importantly addressed. Journalists are highly trained to follow this style regardless of nations or companies. Thus, most news articles, especially straight news, are frequently written in this style except informal types of articles such as columns, editorials or interviews.

The inverted pyramid structure is composed of head, sub-head, lead and main text. The head and the sub-head are decisive factors in deciding whether or not to read an article. Therefore, they usually contains core terms which reflect the aspect being focused in the article. The lead is usually the first sentence of the article, or the first two sentences in some cases. To quickly inform readers, the information ordering principle is especially applied to the lead and guides reporters to reveal the most interesting facts. In addition, the lead covers the selected aspect clearly and concisely to avoid the unreadability of long sentences. Thus, we believe the lead will contain crucial terms which reflect the mainly covered aspect. The main text is organized with collected information which supports the lead in the order of diminishing importance.

We exploit the structure of news articles in both keyword selection and weight calculation. For keyword selection, we basically extract keywords from the core parts of an article, i.e., head, sub-head and lead. The weight of each keyword is calculated based on the main text. While calculating weights, we consider not only the number of occurrences of keywords, but also the locations of occurrences. We also take into account the amount of texts used to describe the keywords at each occurrence. For example, consider the keywords “Government” and “Critics” in Figure 6. The keywords first get weight for each appearance proportional to the length of the paragraph and sentence. If a paragraph elaborates more about governments’ plan, the word "government" will get more weight. Then weight of each appearance is diminished to reflect the location of the appearance. The weight of keyword “government” in location B will be diminished more than the keyword “critics” in location A. Lastly, weight of appearances is summed for each keyword.

Beyond utilizing the news structure, the method is further tailored to capture the elements of aspects, such as agent, action, and etc. The method attempts to recognize

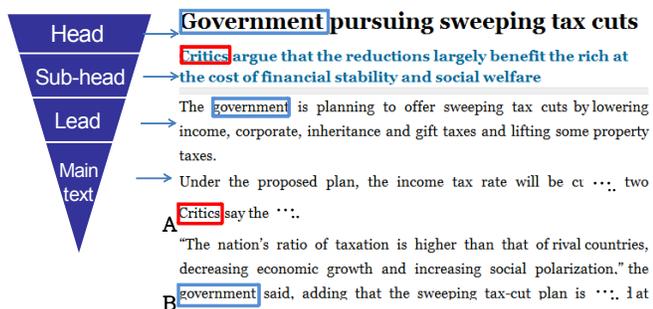


Figure 6. Inverted pyramid structure.

keywords and phrases which represent the elements. By performing a morphological analysis, the method selects only the keywords which have the potential to act as an agent or a description in an aspect. For example, proper nouns and non-descriptive nouns are selected since they have a chance to act as an agent. In order to better capture the agents of aspects, a sequence of keywords forming a noun phrase are also identified and treated as a single semantic entity. Phrases such as “President Bush” and “Justice Department” are parsed and recognized based on syntactic rules of noun phrases. Individual words in a phrase are considered as the same word and share a single weight value.

Aspect Group Creation

To create aspect groups, we adopt the hierarchical agglomerative clustering method (HAC). Starting by considering each article as an individual cluster, the method merges the most similar pair of clusters [21]. For similarity calculation, articles are mapped to vectors in a Euclidean space based on the set of the extracted keyword-weight pairs. The similarity is calculated by the cosine measure. The granularity of clustering results, i.e., the number of clusters in the results, is determined automatically by combining the method with the elbow criterion [15]. We experimentally compared the performance of HAC to that of alternative clustering methods such as k-means, spectral, etc. HAC most often performed better than other techniques.

To make the similarity measure to be robust to article length differences, variations in the number of keywords and weight scales among articles are normalized. The number of keywords can vary as the length of structural elements, e.g. title and lead, differs and so as the weight scales depending on the length of articles. The variation can distort similarity calculation and cause erroneous clustering results. For the normalization, we have sampled more than a hundred of articles and measured the distribution of the numbers of extracted keywords. Regarding the average, the number of keywords is controlled either by lengthening the lead or filtering out a few keywords whose weight is low. Weight values in each article are also normalized to follow a normal distribution with the average of 0.5 and the standard deviation of 0.25.

Aspect Collection

- Coverage of news providers: A single system cannot cover all the news articles published by every news provider due to the ever increasing scale of Internet news services. Just in Korea, more than five hundred news providers produce a huge number of articles every day. This number soars when non-traditional news sources such as blogs are included. The sources should be carefully selected considering their credibility, ideological characteristics, etc. For the careful source selection, we investigated the daily publishing rates of news providers. We observed that the amount of published articles was dominated by a few major news outlets which show high publishing rates. Currently, NewsCube covers the top twenty most publishing sources

and collect about 70% of news articles published in GoogleNews.

Another consideration is the different page formats of the news articles. As the formats differ between different providers, it is fairly complicated to recognize and extract pure news texts from the pages. For each source, by developing customized extraction rules, pure news text can be retrieved. This is important since generalized extraction techniques such as block-detection-based extraction [5] frequently lead to poor extraction results. They may not effectively filter irrelevant contents such as advertisements and seriously degrade the performance of article classification as a result.

- Heterogeneity in article types: Bias in news articles is expressed differently depending on their types such as straight news, interviews, and editorials. For example, it is more implicit in straight news than in editorials. In the latter, opinions are explicitly expressed while bias is formed by filtering or emphasizing certain facts in the former. Thus, identifying article types is another issue to enable separate handling of each type with customized processing methods.

NewsCube currently focuses on straight news, the most common type of news articles. Other types of articles such as interviews and editorials are filtered out since our clustering method is customized to straight news articles at present. We analyzed the article length distribution with more than a hundred of articles and observed that the length of articles can be a simple and effective criterion to sort out straight news articles. In the analysis, for example, articles which were longer than 2000 bytes were mostly interviews and which were shorter than 500 bytes were brief news, respectively. To additionally sort out editorials which have length similar to straight news, other features are also used, e.g. the label "editorial" in titles or distinct URL paths such as "ohmynews.com/editorial/*."

- Obsolescence of news articles: News events evolve over time and news articles may become obsolete accordingly. Users may misunderstand news events if such obsolete articles are presented together with up-to-date articles. We traced the publishing pattern of news articles over time and observed that, as expected, the publication shows a periodic behavior in a daily basis. NewsCube collects only the up-to-date articles which are published on the same date.

EVALUATION

In this section, we present the evaluation results of the NewsCube service, which consist of two parts. First, we study the performance of the developed aspect-level classification method. Second, the effectiveness of NewsCube service is examined through various user studies.

Performance of Aspect-level Classification

Feasibility of Aspect-level Classification

We first examine whether commonly agreeable results can be made for aspect-level classification. Aspect-level classification would be infeasible if people interpret aspects

of articles differently and classify them in various ways. In this experiment, we had participants create manual classification results. We then analyzed the similarity among the created results. Each participant was explained about the definition of aspects and performed classification without knowing the results that are made by others. Five participants performed classification on the article sets.

For the experiment, we sampled 30 test article sets via GoogleNews. Each set consists of articles on an event published at the same date. The article sets are sampled from politics, business, and local sections. Table 1 shows the characteristics of the sets.

Table 1. Features of article sets

Features	Values
No. article sets	30
No. news articles in total	460
Max. article set size	22
Min. article set size	7
Avg. article set size	15

To the best of our knowledge, this experiment is the first to study the feasibility of organizing news articles in aspect-level. A number of quality news text corpora such as Reuters-21578, 20NewsGroups were developed but they are inherently designed to study the feasibility of event-based classification and are not adequate to study that of aspect-level classification. In such corpora, differences in the covered aspects are not identified for the articles covering a common event [22].

Table 2. F-measure among pairs of manual clustering results

waF (avg.)	No. of article sets	waF (min.)	No. of article sets
> 0.9	9	> 0.85	3
> 0.85	18	> 0.8	10
> 0.8	23	> 0.75	15
> 0.75	26	> 0.7	22

Table 3. F-measure between ground truth and manual results

waF (avg.)	No. of article sets	waF (min.)	No. of article sets
> 0.9	15	> 0.85	7
> 0.85	23	> 0.8	18
> 0.8	28	> 0.75	22
> 0.75	29	> 0.7	25

There is a high level of similarity among the manual classification results. We demonstrate the similarity using the weighted average F-measure⁴ (waF) [25]. For each article set, waF is calculated for every pair of the results. Table 2 shows the average and the minimum waF for all article sets. For more than a half of the article sets, the average of waF over the result pairs is greater than 0.85. Such a value can be achieved, for example, when only three or four articles are grouped differently in an article set of 20 articles. The minimum waF among the result pairs is also high. For 22 article sets, every pair of results shows the waF value greater than 0.7.

Since the manual classification results are highly similar, we create ground truth classification results from them. To

⁴ F-measure is the harmonic mean of precision and recall and ranges between 0 and 1. Given two clustering results, the F-measure is calculated for each group in one result in comparison with the corresponding group in the other result. We take the weighted average F-measure as the overall measure. For details, refer to [25].

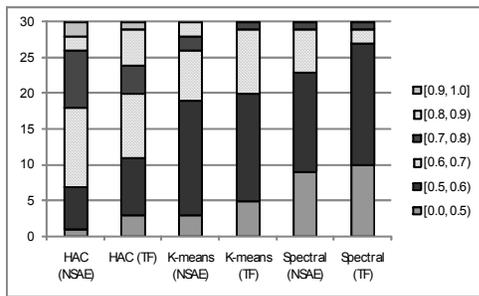


Figure 7. *waF* distribution of each method

make the ground truth, we group articles together if four or more participants agreed to group the articles together. The created ground truths are very similar to each manual clustering result. (See Table 3).

Performance of Aspect-level Clustering

We evaluate the performance of clustering methods using the developed 30 ground truth results. We compare three representative clustering methods; Hierarchical Agglomerative Clustering (HAC), K-means, and Spectral clustering. In order to demonstrate the benefits of the news structure-based extraction (NSE), each method was combined with both NSE and simple term frequency-based extraction (TF). The evaluation results show that the result of HAC-NSE is closer to the ground truth than others.

Table 4. *waF* achieved by the methods

	Average <i>waF</i> .	# of winning sets
HAC (NSE)	0.683	15
HAC(TF)	0.654	8
K-means(NSE)	0.587	0
K-means(TF)	0.567	2
Spectral(NSE)	0.542	5
Spectral(TF)	0.527	0

Table 4 shows the *waF* achieved by the methods. For each method, we calculate the average *waF* over the 30 article sets. HAC-NSE shows the highest average *waF* and it shows the best *waF* for 15 article sets. Figure 7 shows the *waF* distribution for each clustering method. HAC-NSE shows larger portions of higher values than those of other methods.

Effectiveness of NewsCube

The key to the effectiveness of NewsCube is twofold: 1) the diversity of delivered articles in terms of aspects and 2) the influence of the service on the development of balanced views. To evaluate the effectiveness, we have designed and conducted three types of user studies. First, a clickstream analysis is conducted to quantitatively measure the diversity of articles read by users. Second, a questionnaire is used to study NewsCube users' subjective evaluation on its effects in terms of aspect delivery and balanced view development. Third, a focus group interview is conducted to gain deeper understanding on the effect of NewsCube in users' opinion development process.

Participants and Procedure

1) Clickstream analysis: 75 participants (ages between 25 and 35) were recruited. They were invited via the bulletin board of our school and were chosen randomly. We inquired their

ordinary news reading habits for a fair study; they were asked how many and how often they read news articles on a news event.

The effect of the NewsCube service is compared with those of two other services. For the purpose, the participants were divided into three groups, each of which had 25 participants. The groups had comparable number of participants with similar news reading habits. Each group was provided with the same set of news articles. The article set included 19 articles about the increase of real estate tax. They were clustered into four aspect groups. None of the participants was aware of the event in detail beforehand. The first group used NewsCube and the second group used GoogleNews. The third group used RandomCube service, of which the interface is the same as NewsCube, but articles are grouped randomly. Not to influence browsing behaviors, we neither explained about the goal nor the interface of the NewsCube service.

2) Questionnaire: 33 participants were recruited from our university. They were mostly students but a few administrative staff and researchers were also included: ages ranged from 24 to 36, averaged 28; 16 males and 17 females. Participants had majors in diverse disciplines, and the news categories of their interest were also diverse including politics, business, local, technology, and science, etc. Nearly all participants used the Internet as their primary news source.

The participants first went through a short instruction session on the NewsCube service. Different from the clickstream analysis, the participants were explained about the goal and the interface of it. They were asked to evaluate NewsCube while regarding its goal after they used it for 15 minutes. Then, we inquired about a) the demand to compare and contrast multiple aspects of news events, b) the diversity of read articles, and c) the influence on their views through five closed (5 Likert scale) question items. In addition, we had a short interview with open-ended questions to gain deeper understanding on their answers and gather feedbacks.

3) Focus group interview: For the interview, three groups of participants were formed according to their position to a given news event. The selected event was 'a protest against the U.S military base expansion'⁵. In the event, many arrest warrants were requested for the protesters by the prosecution. The court rejected a number of warrants and a conflict arose between the prosecution and the court.

Each group had two participants; two who considered the protest negatively (Negative group), two who defended the position of protesters (Positive group), and two who did not have a predetermined opinion (Neutral group). One participant in a group read news through NewsCube and the other read through GoogleNews. Again, the goal of the service and interface of NewsCube was not explained not to influence their browsing behaviors. After a news reading session, we interviewed about their opinions on the event in detail. The participants were asked to explain their views on the key actors in the event, i.e., the protesters, the prosecution,

⁵ Refer to http://english.hani.co.kr/arti/english_edition/e_national/121596.html

and the court. They were also asked to point out the articles which helped them develop their views.

Results and Discussion

Table 5. Clickstream analysis results

	NewsCube	GoogleNews	Significance
Avg. # of articles read	4	2.5	$t_0=2.8449, p=.01$
Avg. # of groups read	3	1.1	$t_0=8.332, p=.01$

	NewsCube	RandomCube	Significance
Avg. # of articles read	4	2.62	$t_0=2.53, p=.02$
Avg. # of groups read	3	1.08	$t_0=8.679, p=.01$

1) Clickstream analysis: We hypothesized that NewsCube users would read more number of articles with different aspects than those of the other two services. NewsCube users can easily be aware of different articles emphasizing different aspects from the classification of aspect-level browsing. On the other hand, users of the other two services have to put extra effort to search different aspects. The GoogleNews interface simply lists similar article titles. RandomCube gives the same interface as NewsCube, but with random article classification.

We traced the articles each participant read and counted the number of the articles. To ensure the correctness of counting, we asked the participants to confirm the articles which they have actually read. We also counted and compared the diversity of read articles in terms of aspects. For this, we counted the number of article groups, i.e., the number of the groups according to the NewsCube classification from which each user read at least one article. As shown in Table 5, users of NewsCube read more articles and aspect groups. The effect of NewsCube is statistically significant according to the independent sample t-test. In case of RandomCube, the number of read articles slightly increased compared to GoogleNews. However, NewsCube users still read more articles and aspect groups in spite of the identical interface.

2) Questionnaire: The first two questions asked the demand to compare and contrast multiple aspects of news events. 79% of the participants answered that ‘they (strongly) felt that it is necessary to compare multiple articles’ in the past, (i.e., before knowing NewsCube). In addition, 70 % answered that they ‘actually compared articles once (or more) before’. They also mentioned that they did so by putting additional efforts; typically, they visit different news providers’ website or web search engines, and search related articles for comparison.

Despite the participants consider that it is important to compare and contrast multiple aspects, they did so much less often than expected. Just 30% of the participants responded that they compared multiple articles more than once a week. Moreover, only four people did so every day. To the majority of the participants, the additional efforts required for article comparison seem burdensome.

More importantly, the next two questions asked if NewsCube guides users to read more articles. 72 % said they ‘(strongly) felt like to read various articles with different

aspects’ while using the service. 70 % answered that ‘they did actually read various articles’.

Interestingly, the service motivated the users, who had not been interested in doing so, to compare and contrast multiple news articles. There were three participants who replied that ‘it is not necessary to compare multiple articles’. However, all of them answered that they ‘(strongly) felt like to read various articles’ while using the service. There were also ten participants (30%) who said that they ‘hardly compared multiple articles’. Among them, nine said that they ‘(strongly) felt like to read various articles’ as well.

In the open-ended questions to the 70% who said that they actually read various articles, many confirmed the effectiveness of the aspect traversal cue. They said that they could notice what articles are left to read through it. It made them feel that they did not complete their reading and led them to read more articles.

Among the other 30%, many participants mainly pointed out two reasons why they did not read various articles. Some said that they were simply not interested in the news events and they did not want to read articles at all. Others commented that, in some news events, they could not easily understand the difference among the classified articles and hence lost the motivation to read the articles.

As for the potential of NewsCube to help users build their own balanced view, 16 participants (48%) answered (strongly) positively that ‘they could formulate a more balanced view on the news events’ through the short test. Only two answered negatively to this question. Many of the participants who did not give positive answers further commented this single short test session was not enough to evaluate the effect on their view.

3) Focus group interview: The opinion development process of users is further examined in more detail through the interview. We observed how opinions of the participants in each group diverged depending on the service they used. As shown in Table 6, the two participants in each group read different sets of articles.

The GoogleNews users read the articles covering the aspects in favor of his/her original view points and did not read those with opposing viewpoints. For example, the one in Negatives group read the articles in article group 3, but not those in group 1. On the contrary, the NewsCube users read more diverse articles even with opposing viewpoints. After the news reading session, the two participants with the same opinion in the neutral and the negative groups began to reveal differences in their opinions.

The two participants in the Neutral group developed different views towards all three key actors. The GoogleNews user formulated a negative opinion against the protesters and emphasized the violence of the protesters. The user referred an article which discussed the violence of the protesters, which belongs to the article group 3. However, the NewsCube user defended the protesters and criticized the prosecutor for issuing the arrest warrants imprudently. Her opinion was influenced more by the articles discussing the actions of the prosecution (i.e., those in article group 1)

although she also read the same article on the protesters' violence. Interestingly, unlike the GoogleNews user, she started to criticize the articles rather than just to passively accept the information in the articles after using NewsCube. She said, "The articles in group 3 seem to intentionally exaggerate the story to formulate a negative view towards the protesters".

Table 6. Article groups whose member article is read

Group ID		1	2	3	4	5
Negatives	GoogleNews		v	v		
	NewsCube	v	v	v		v
Neutral	GoogleNews		v	v		v
	NewsCube	v	v	v	v	
Positives	GoogleNews	v	v			
	NewsCube	v	v	v	v	

Aspects articulated in each group - 1: Stern action of prosecution, 2: Conflict between court and prosecution, 3: Characteristics of key protesters, 4: Second request for arrest warrant, 5: Investigation plan for the protest leaders (Note that articles are not grouped in GoogleNews. For GoogleNews users, we checked the groups by mapping the read articles into NewsCube classification.)

In the Negative group, the opinion on the protest did not change. However, the participants developed different opinions on the court and the prosecution. The GoogleNews user criticized the court for restraining the prosecution by rejecting a number of arrest warrants. The NewsCube user, however, acknowledged the position of the court and accepted the actions of the court.

The difference of opinions was not found in Positive group although the two participants read different sets of articles. After the readings, both still defended the court and the protesters, and criticized the prosecution. The NewsCube user read the articles with the opposing view point discussing the violence of the protesters (article group 3); however, he raised questions on the credibility of the articles and considered the articles to be biased.

CONCLUSION AND FUTURE WORK

In this paper, we put forward the media bias problem and develop a computational framework for the problem. NewsCube is a novel Internet news service designed to mitigate the effect of media bias. It automatically creates and promptly provides readers with multiple classified viewpoints on a news event of interest. Users can easily discover rich facts and compare diverse biased views of the event. The evaluation results show that NewsCube leads readers to read more articles covering different aspects and helps them develop more balanced views. We are preparing for a public service in collaboration with a leading Korean Internet news media company [26].

As the problem of media bias is large and study for the problem is in its starting stage, there are lots of challenging issues for further investigation. Studies on bias diagnosis, measurement, and correction are difficult but important directions. There are also many issues for bias mitigation. Studies should be continued for a sophisticated understanding of bias in media discourse. Other types of bias embodiments, e.g., spin, placement, photo selection, space

allocation, labeling, should be explored. The bias problem should also be studied for different types of media, e.g., blogs, images, video.

REFERENCES

- Allan, J., et al. Topic Detection and Tracking Pilot Study: Final Report. In Proceedings of DARPA Broadcast News Transcription and Understanding Workshop, 1998.
- Baker, B. How to Identify, Expose and Correct Liberal Media Bias. Media Research Center, 1994.
- Baron, D. P. Persistent Media Bias. Stanford University, Graduate School of Business Research Papers: No. 1845, 2004.
- Bernhardt, D., et al. Political Polarization and the Electoral Effects of Media Bias. *CESifo Working Paper Series No. 1798*.
- Chen, Y., et al. Detecting Web Page Structure for Adaptive Viewing on Small Form Factor Devices. In *Proc. of World Wide Web*, 2003.
- Entman R.M. Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication* 43 (4): 51-8, 1993.
- Gamon, Michael., et al. BLEWS: Using Blogs to Provide Context for News Articles. In *Proc. of the ICWSM*, 2008.
- Gentzkow, M. and Shapiro, J. M. What drives media slant?: Evidence from U.S. daily newspapers (No. Working Paper No. 12707). Cambridge, MA: National Bureau of Economic Research, 2006.
- Gerber, A., et al. Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *Yale Economic Applications and Policy Discussion Paper No. 12*.
- Groseclose, T. and Milyo, J. A measure of media bias. *Quarterly Journal of Economics*, Vol. 120, Issue 4, 1191 – 1237, 2005.
- Herman, E.S. et al. Manufacturing Consent: the Political Economy of the Mass Media. New York: Pantheon Books, 1988.
- Herman, E. S. The Propaganda Model: A Retrospective. *Against All Reason*, 1: 1-14, 2003.
- Kittur, A. et al. He says, she says: Conflict and coordination in Wikipedia. In *Proc. of CHI*. 2007
- Kull, S., C. Ramsey and E. Lewis. Misperceptions, the media, and the Iraq War. *Political Science Quarterly* 118(4): 569-598, 2003.
- Mark S. Aldenderfer and Roger K. Blashfield. Cluster Analysis. SAGE Publications Inc., 1984.
- McChesney, R. W. Theses on media deregulation. *Media, Culture, and Society*, 25, 2003, 125-133.
- Radev, D. R. et al. NewsInEssence: Summarizing Online News Topics. *CACM Vol. 48, No. 10*, 2005.
- Rhee, J. W. Framing the Health Care Reform Campaign of 1993-94: News Frame, Interpretation, and Public Opinion Change. *Political Communication*, vol. 15: 274-76. 1998.
- Sebastiani, F. Machine learning in automated text categorization. *ACM Computing. Survey*: 1-47, 2002.
- Son, S.C. The Revolution of News Reading. Kaemagowon, 2006 (in Korean).
- Michael Steinbach, et al. A Comparison of Document Clustering Techniques. *Univ. of Minnesota, Technical Report #00-034*, 2000.
- Strassel, S. et al. Quality Control in Large Annotation Projects Involving Multiple Judges: The Case of the TDT Corpora. In *Proc. of LREC*, 2000.
- Tankard, James W. The Empirical Approach to the Study of Media Framing. In *Framing Public Life: Perspectives on Media and our Understanding of the Social World*. Stephen D Reese, Oscar H Gandy, and August E Grant. Mahwah, NJ: Lawrence Erlbaum Associates.
- Urban. C. Examining Our Credibility: Perspectives of the Public and the Press. American Society of Newspaper Editors, 1999.
- Xiaojun Wan, et al. CollabSum: Exploiting Multiple Document Clustering for Collaborative Single Document Summarization. In *Proc. of SIGIR*, 2007.
- <http://www.ohmynews.com>