Finding Informative Comments for Video Viewing

Seungwoo Choi
Graduate School of knowledge Service Engineering
KAIST
Daejeon, Korea
sw.choi@kaist.ac.kr

Aviv Segev
Graduate School of knowledge Service Engineering
KAIST
Daejeon, Korea
aviv@kaist.edu

Abstract—**Video is an increasingly important method of information-sharing on the Web. Services such as YouTube, Vimeo, and Liveleak are platforms that support uploading User-Generated Content. Users tend to seek related information during or after watching an informative video by finding and reading comments on Web services. However, existing services only support sorting by recentness (newest) or rating (*LIKES* score), as opposed to related information. We suggest a novel method to find informative comments by considering original content and its relevance. We conducted a qualitative study of participants watching informative videos and analyzed how users interacted with the comments, the feature preferences, and the criteria for evaluation. We developed methods to measure informativeness priority, the user-provided level of information, classify intention of information posted, and cluster duplicate themes. Analysis of 1,861 TED talk videos and 380,619 comments show suggested methods can find more informative comments compared to existing methods (*LIKES*).**

Keywords-**Information processing; Social computing; Information rates;**

I. INTRODUCTION

Recently people have become more familiar with accessing information using the Web, and information sharing using video is an important method of spreading ideas and knowledge online. For this purpose, various services, such as YouTube, Vimeo, and Liveleak, support uploading User-Generated Content (UGC) to the Web. People can access Massive Open Online Courses (MOOCs) using video services such as Coursera. Additionally, a popular conference video sharing service, TED, supports sharing and spreading ideas on the Web. Users tend to seek related information during or after watching an informative video. The best way to satisfy these information needs is to find and read the comments on Web services. However, existing services support only sorting by recentness (newest) or rating (*LIKES* score). Thus, the search for related information has limitations:

- If a video receives hundreds of comments, users find it hard to refer to all comments when looking for related information. Users prefer to consider only highly ranked ones.
- In the sorting method based on *LIKES* score, comments popular among users tend following their high ranking to be more visible than higher quality comments.

- Since users prefer highly ranked comments when there are many comments, the highly ranked comments can remain highly ranked following their exposure and accessibility.
- Unless users read all comments, there is no way for them to identify the informative comments.

We propose a method to find and select Informative Comments (ICFinder) to help users understand original contents and supply additional related information. We suggest methods based on analysis of user behavior with comments and integrate different components to find the most informative comments for situations where there are many comments. The method classified the important features identified by the users. For each of these features, an algorithm was developed or integrated to quantify them. The methods are based on analysis of 1,860 videos and 380,619 comments from TED and on a qualitative study showing how users interact with comments, which features users are more satisfied with, and which criteria users evaluate with. Prior recommendation approaches are based on recording user behavior. However, our approach is based on actual information contained in the video, contents of the comments, and meta-information. The methods are composed of measuring informativeness priority, classifying information intention, and clustering themes. The informativeness priority method calculates how much the comment refers to information coverage in the original contents. Next the method classifies the comment intention for information: whether comments invoke or evoke information. Then the method selects the most informative comments by applying clustering methods to eliminate theme duplication using rules.

To analyze the method, we randomly selected 20 videos from the TED video service and used human evaluators to judge the data. We compared the informative comments selected by human evaluators to the commonly used method of high rating (*LIKES* score) as a baseline for the ICFinder method.

The main contributions of our approach are:

- The method supplies users with related informative comments that provide useful information.
- The method of selecting informative comments is based

on matching between original contents and crowd intelligence rather than crowd numbers.

- The experiments using real-world data analyze both user preference in selecting the best comments and how to provide users with the most informative comments.
- The method shows an extended algorithm for a recommendation system by incorporating user perspective based on commenting behaviors.

II. RELATED WORK

Many Web services provide social interaction via commenting functions among users to express personal opinion or sentiment with considerable information such as URL, quotation, news articles, etc. Several studies focus on comment recommendation ranking based on criteria. The common criteria for evaluating comments' priority are usefulness and credibility.

A. Insight on Web services by comments

Hu and Liu [1] suggested an automatic model for extracting opinion features, identifying opinion sentences, and summarizing results. The work presents a system that can mine product features that have been commented on by customers automatically and summarizes the results by identifying opinion sentences and evaluating whether each opinion sentence is positive or negative.

Li et al. [2] built an adaptive recommender system by utilizing reader comments to suggest relevant news in accordance with the evolving topic. To achieve this purpose, they combined two features together. One is the method of identifying the most relevant news as further information. Another is the method of sharing information collaboratively among users by their comments.

These studies showed the possibility of containing valuable information in comments and the example of its utilization in Web services. However, these studies revealed a lack in the study of the human perspective on commenting behaviors.

B. Comments' priority by several criteria

Other studies focus on comment recommendation ranking. The common criteria for evaluating comments' priority are quality and usefulness.

Figueiredo et al. [3] investigated the quality of textual features including comments, title, description, and tags in Web service. The research indicated that collaborative features, including tags, provide more content than restrictive features, such as title. However, the research didn't show any prioritizing method by assessing comments quality itself.

Momeni et al. [4] suggested a usefulness classifier based on features such as surface-level, syntactic, semantic, and topic. They claimed that the useful comments are influenced by the entity type of the media object, its time period, and the degree of polarization among commenters.

These studies revealed that ranking the comments or contents depends on adding or applying the features or factors but did not describe how to evaluate them. Additionally, they showed the lack of study of the human perspective for understanding comments on multiple topics [5] and the need to analyze large volumes of comments [6].

III. PRELIMINARY STUDY

A. Study on TED Talks

To observe user information needs in video viewing, we chose TED, a well-known service for idea sharing. TED's slogan is "Ideas Worth Spreading." TED conferences feature a wide range of topics and are presented through a personal talk. To understand how users actively use this service, we crawled the information provided for each video, which includes the presentation script and the users' comments, for all videos from 2006 to 2014.

1) Video Statistics: The number of crawlable videos is 1,861. We analyzed TED video information to understand how much video is actively viewed. The statistics show the average view count for each TED video is over 1.2 million (Min: 45,901, Max: 29,414,125, SD: 1,651,874). The TED email count evaluates the number of emails sent by users to recommend the videos, showing how much the idea in the video is spread by users. We eliminated recently added videos with no emails. The results show the maximum number of emails is almost 100,000. The average number of emails is almost 1,000 (SD: 3,541). The actual number is higher since TED provides sharing functionality using social networks (Facebook, Twitter) embedding the service in the webpage. The comment count shows the degree to which users interact with the video. The average number of comments is 205, and the maximum number of comments is over 6,000 (SD: 288). The number of comments is high, reflecting the purpose of the web service, to spread ideas.

2) Comments Statistics: We crawled through all 380,619 comments, analyzing language and length, to evaluate how users express their opinions and feelings with TED multi-language UI. We performed language detection using the Language Detection Library for Java. 96.6% of the comments are in English. The API detected a total of 45 languages. The majority of non-English comments are in Spanish and French. In some cases the API doesn't detect any comments because of their contents: emoticon, email, and URL. We focus only on English comments due to the high proportion.

3) Script Statistics: TED also provides the video script translated into multiple languages. We only focus on English script information. We crawled 1,764 TED English video scripts in a variety of genres in TED with an average word count of 2,109 (SD: 1,010). Since some TED content is in the fields of performance and art, this content does not have written script.

Types	Contents	
Motives	- Wonder how others summarize information - Wonder about other opinions to mitigate biased opinion - Wonder how others are feeling - Wonder what is the best reply	
Evaluation — **Feature**	- **URL** - **New Information** - **Additional Opinion or Thought**	- **Reasonable Basis** - **Length** - **Explicit Information Source**
Evaluation — **Criteria**	**Positive** - Good summary - Well organized logic - Looks like an expert - Quality of question - Sympathy	**Negative** - Too much detail or too focused - Simple sentiment - Off the point - Personal experience - Narrow opinion (subjectivity)
Additional Information	- System design for tracking thread comments *(Show question and find first answer)* - Dialog form UI - Topic support UI	

Figure 1: Organized affinity diagram

B. User aspects for informative video viewing

To understand which features satisfy user information seeking, we conducted a qualitative study to identify when users find additional information in comments. This research was based on the principles of grounded theory [5]. The qualitative study consists of two steps, semi-structured interview of users after they watched a TED talk and creation of an affinity diagram.

1) Qualitative Study Settings: In the first phase, we randomly chose a video (Hans Rosling, 2006, over 9 million views) to make the semi-structured survey and used 20 randomly chosen comments. We recruited 7 men and 2 women (average age 28, range 25-36) who are graduate students or post-doc researchers. After they viewed the video, we measured using a 4-point Likert scale their satisfaction with the additional information they found in the comments. It took 20 minutes to watch the video and 25 minutes to analyze the comments. After conducting the simulated experiment, we interviewed them for 20 minutes to ask open questions about which comments were most and least satisfying and which features were most effective to measure the comments. After interviewing 9 subjects to serve as machine performance evaluators, there were no additional results that could be classified as valuable information on this topic. In the second phase, we made an open coding scheme for each transcribed interview. We used a whiteboard affinity diagramming approach in the grounded theory method [5], [7], used to organize unstructured sets

of ideas and concepts. Then we categorized them by name. We discussed these codes and their definitions in an iterative process until the diagram stabilized.

2) Findings: We created an affinity diagram from a qualitative study for extracting themes, criteria, and features related to the user satisfaction using two researchers. Figure 1 illustrates the details and organization of the affinity diagram. The classification types consist of Motives, Evaluation, and Additional Ideas. Motives illustrate why users want to seek information in comments for video viewing. Evaluation describes which features are considered to make users feel more satisfied and which criteria are considered to influence informativeness. Additional Ideas have experiment participants advise us to get better design ideas.

Their motives include finding comments in a similar situation, like reading news articles, reading comics, and watching YouTube videos for entertainment. They usually find other users' comments to obtain more information from the summary or other users' opinions to mitigate biased opinions. They want to prejudge the contents by using other users' comments before they watch the video. They wonder how other users feel after watching the TED talk. Last, they wonder why the best comment received the highest scores.

Additionally, there are evaluation features and criteria for finding informative comments. Users prefer well-organized or concise comments. Users especially give a good mark when they feel sympathy with other users after reading their comments or when they seek information from an

external source that looks qualified as an expert. Users are more satisfied with good quality questions that inspire new thoughts or call their attention to a TED video. We assumed that features are considered to be related to external information such as URL and explicit information source (based on email, parentheses, quotation). Comments mentioning a minor problem or contents irrelevant to the TED videos receive a low evaluation. Comments mentioning personal experience or overemotional comments also receive a low rating, as do comments with a claim triggering major disagreement.

Last, we explained our aim in this research to the interviewees and asked them for any new ideas for further research. One user suggested matching questions and answers in comments, and this idea was integrated into the method. Other suggestions included creating a user dialog interface and identifying comments related only to the specific video topic.

C. Characteristics of information intention

We identified several features about how the user evaluates the informativeness of comments and which criteria affect the evaluation. Next, we empirically analyzed which features from the qualitative study of TED comments are connected to which information needs. We defined two classes of information intention:

1) Invoking Information: These comments intend to purposely invoke and ask related or external information from other users. Users use the 5W1H (When, Where, What, Who, Why, How) question form and a Modal Verb (Can, Could, May, Might, Should) to ask a question explicitly. Additionally, users tend to write sentences that will trigger an answer from other users although they do not use the question form to ask questions implicitly. For instance, *I wonder, I have a question, I guess*. To find related examples, we searched the comments of Rosling's TED video to make an Affinity diagram.

2) Evoking Information: These comments intend to evoke information and aim to give information to other users. The first type of comment provides direct answer to another user's question in response to the *Invoking* class. The second type voluntarily gives information or provides an example considered to be related to the TED video although none of the users asked a question. The third type expresses users' opinion of agreement or disagreement when the speaker mentions a new idea or claim in the TED video. The fourth type expresses users' feelings after watching the TED video. We decided that the information needs should include the checking of other users' feelings, because users are motivated to check other users' opinions before or after watching the video, though some feelings are expressed simply and in short.

IV. ICFINDER METHOD

The preliminary study examined why users leave comments and which features users are more satisfied with based on TED. We chose some significant factors for designing the method based on previous results:

- Users tend to prefer comments referring to well-balanced information.
- There are two information intentions *Invoking* and *Evoking*.
- There can be various themes in the comments based on each user intention because of several reasons such as personal experience or opinion.

Figure 2 shows the ICFinder method diagram for finding informative comments. First, we considered an information coverage approach, named *Semantic Entropy*, for calculating informative priority score. Second, we developed an information classifier algorithm, which recognizes the intention of each comment. Third, for each intention, *Invoking* and *Evoking*, we clustered the comments into themes, which helped eliminate duplications. Last, we prioritized the selected comments using rules based on the informativeness score.

1) Information Coverage: Based on the qualitative study, information coverage in comments is one of the most important factors for identifying comment informativeness. If comments are too detailed compared to the original content, the user feels dissatisfaction. To avoid choosing comments with a concentration of detailed information in the script, we parse the target video script, for which we seek additional information, into sentences. Then we generate a weight matrix relating each sentence and comment with variables representing a weight score using Latent Semantic Analysis (LSA) [8]. Next we borrowed from information theory the concept of entropy [9], which represents how much information is balanced in each comment. Entropy is known as a measurement of uncertainty, the tendency of probability distribution in the event. If one event occurs more than the others, this situation looks less informative when observing the occurring event. Entropy will be close to zero when only one certain event is expected. The traditional entropy formula is

$$Entropy = \sum_i -P_i \log P_i \qquad (1)$$

where P is occurrence probability of each event. We calculate the probability

$$P_i = \frac{W_i}{\sum_{i \in S} W_i}, \; i \in S \qquad (2)$$

where S is a parsed sentence, W is each weight relating each sentence and comment, and i is specific identifier of each comment. Then, we organized entropy sets for each comment.
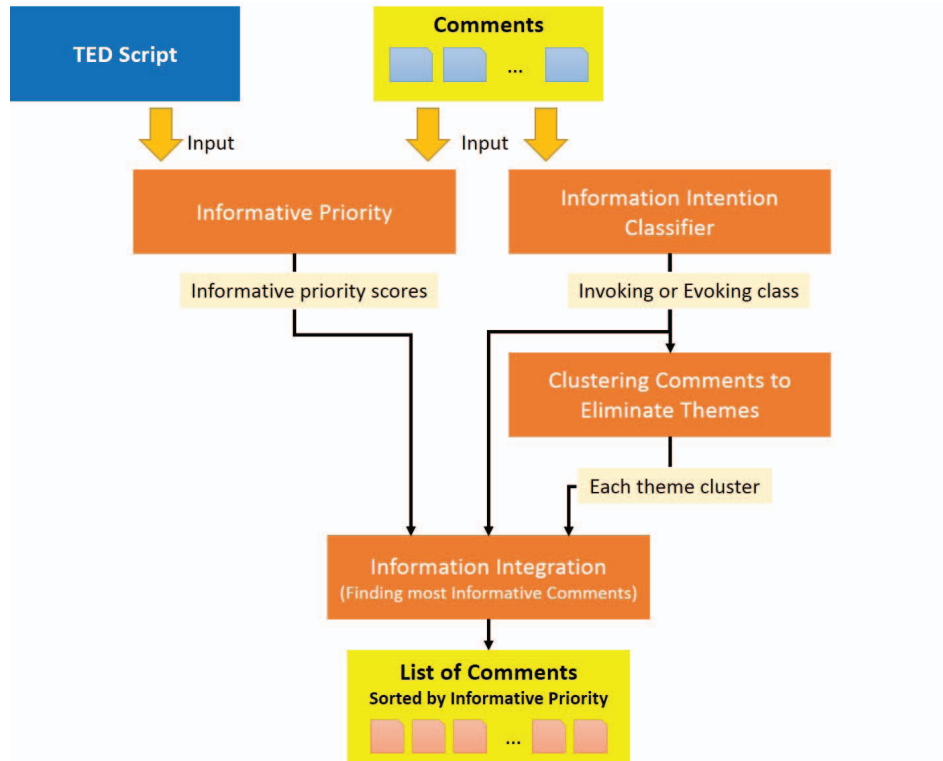
Figure 2: Overview of ICFinder method

2) Information Intention Classifier: Based on the qualitative study, it is important to classify the comment intention, using different intention approach methods, such as Q&A. To organize the information intention classifier, we picked comments randomly and made an answer set using a human evaluator. Then, we trained the classifier using the answer set. Last, we applied the classifier to judge the comments intention.

3) Clustering Comments to Eliminate Themes: Despite successful information intention classification, most comments talk about similar information, which can trigger an error because of the various themes in the comments. We used a clustering method to find duplicated groups. Before clustering, we preprocessed all comments on TED for each video. We used the Vector Space Model to extract terms from documents as a vector. We assumed a comment is defined as a document and all comments about one video are the document set. Then, we can calculate the TF-IDF score in comments. Each term corresponds to a unique weight based on term frequency (TF) and inverse document frequency (IDF).

We eliminated stopwords and stemmed all comments. The rest of the words are mentioned as *STOPSTEM* words. We treated *STOPSTEM* words in each comment as a document and calculated the TF-IDF weight set for each document. We sorted each document by descending TF-IDF weight set.

We picked up to top five *STOPSTEM* words with the highest TF-IDF score in each document. We gathered all terms from each *STOPSTEM* word and generated the dataset using these terms and TF-IDF score pair.

After we obtained the dataset for making a cluster, we used WEKA [10] API for the Expectation-Maximization clustering [11].

4) Information Integration (Finding most informative comments): We integrated all previous results in ICFinder. We ordered comments by informative priority, information intention, and number of clusters generated for each information intention. We selected the top 5 comments by integrating rules.

We consider each portion of the class *Invoking* and *Evoking* to analyze the information intention by volume, when the greater volume is more important. If a class accounts for more than 80% of the total comments, we picked 4 comments in this class. However, if the difference of class volume is lower, we picked 3 comments in the bigger class. In each class, the method picks one comment from the largest cluster volume until its entire class portion is done. In each cluster the method picks the comment with highest informative priority. We can select the best comment that covers all video information in various and duplicate themes. If the comment is classified as *Invoking* class, the method should find its answer comments, which are classified as *Evoking* class, and unite them.

We designed the method that can choose the proper answer for *Invoking* comments, defined as asking other users questions. This method also uses the LSA similarity API. In the TED service, the system supports comments and their reply as a thread, meaning comments with their answer directly input by the user interface. If an informative comment from ICFinder is in the thread, meaning a set of messages grouped, ICFinder will find the answer comment in the thread. The answer comment has to be *Evoking* class and obtain highest similarity with the original *Invoking* comment. However, ICFinder tries to search all comments if the original *Invoking* comment doesn't have a thread. ICFinder finds comments with two conditions: the comment that has a highest similarity score by LSA API and additionally the comment that is posted later than the original *Invoking* comment.

V. EXPERIMENTS

First, we evaluate the method for measuring informativeness priority by comparing one comment set with high priority and one with low priority. Then, we analyze classification accuracy for information intention. Next, we evaluate the method effectiveness in applying and organizing clusters. Last, we compare our method to other widely used information prioritizing methods. To analyze the performance, we sampled 20 videos from all videos with at least 200 comments and at least 10 minute duration.

A. User Perspective Informativeness Priority

To evaluate comment information semantic similarity with the original script, we used SEMILAR (http://www.semanticsimilarity.org) which provides LSA. We used three approaches to verify relevance between comments and script:

1) Baseline - LSA score between full script and comments
2) Semantic-Sum - We used Semantic-Sum method for each comment.
   - **Split** all scripts into single sentences, then let S denote the set of sentences from the script.
   - **Generate** LSA score as the weight between each comment and a whole set of sentences.
   - **Find** the maximum weight to normalize all weight values and calculate the sum of all weights,

$$SS = \sum_{i \in S} \frac{W_i}{MAX(W_i)} \quad (3)$$

3) *Semantic Entropy* - We used the *Semantic Entropy* method for each comment.
   - **Parse** all scripts into sentences.
   - **Generate** LSA score as the weight between each comment and the whole set of sentences.
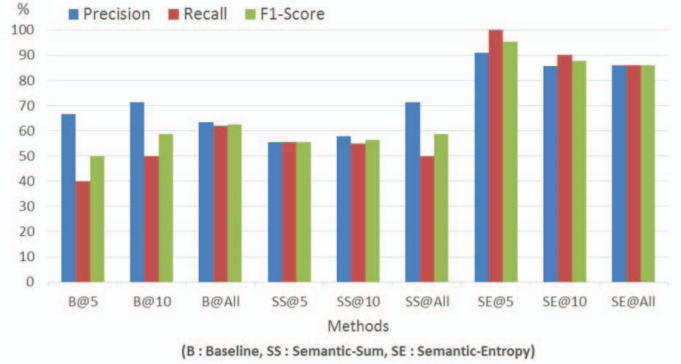

Figure 3: Informativeness priority evaluation

- **Create** probability for each sentence by using the sum of LSA weights,

$$P_i = \frac{W_i}{\sum_{i \in S} W_i}, \quad i \in S \quad (4)$$

- **Calculate** entropy based on the probability of the sentence being perceived as a balanced information covering each comment,

$$SE = \sum_{i \in S} -P_i \log P_i \quad (5)$$

Then we selected two videos to compare the efficiency of each method. For each video, we calculated each comment priority using these methods by selecting 50 comments for each video, top 25 relevant and non-relevant. Last we made 3 sets (2 topics X 50 comments) organized by methods. One researcher coded the sets. The other randomly picked one set and judged whether each comment is relevant. Cohen's Kappa measure as Inter-Annotator agreement was 0.66.

Result: Figure 3 illustrates the result of the top-K by methods selection. The X-axis represents the different selection of top-K (@K) for each method and the Y-axis the precision, recall, and F-1. The *Semantic Entropy* method outperformed the baseline and semantic sum methods for both topics analyzed. Measuring the metric with a lower value of top-K, the *Semantic Entropy* method shows even better results.

B. Information Intention Classification Accuracy Settings

1) Settings: We randomly extracted 50 comments from all previously selected 20 videos. Two researchers divided the 1,000 comments into 2 sets and coded each set separately. Overall Kappa score is 0.74 for 2 videos.

2) Features: We chose features from the user study that we can extract computationally (Figure 4).

- **URL** How many times URL appears in the comment.
- **New Information** How many Questions appear: 5W1H, modal verb, implicit question ("I guess ...", "I wonder ..."), answer form ("The answer ...", "I disagree ...."), email.
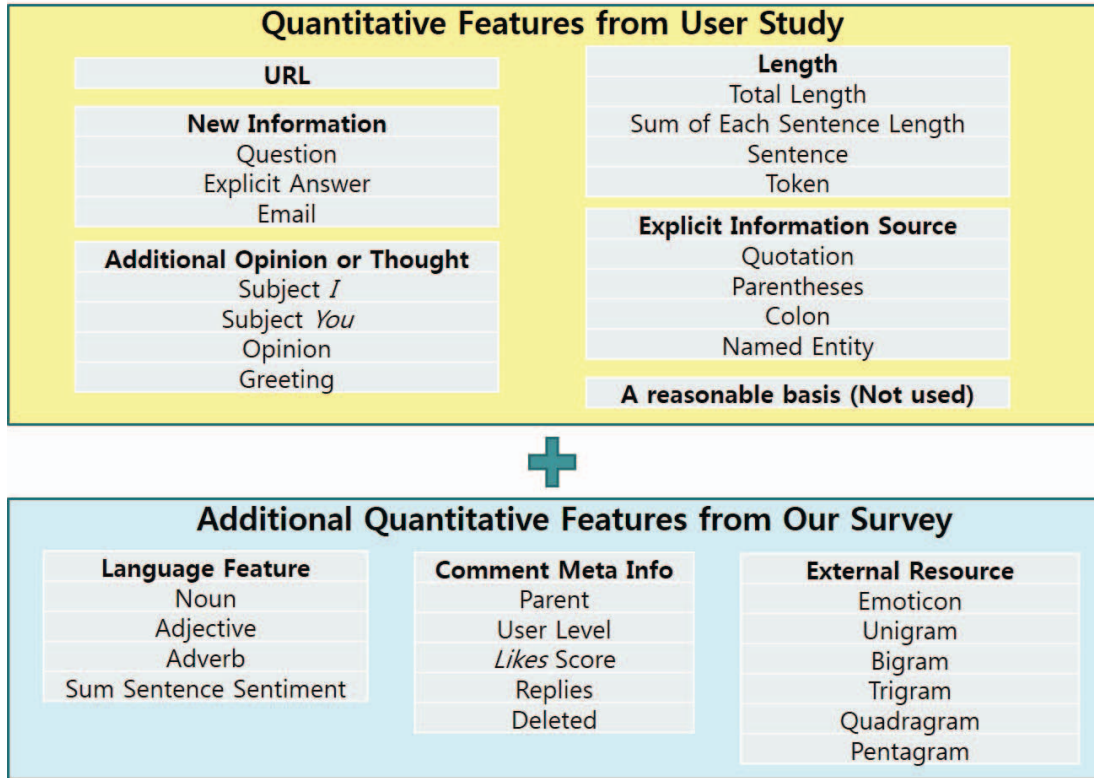
Figure 4: All features for classification

- **Opinion or Thought** How many Subjects I or You appear in comment using Stanford POS Tagger [12]. Identify opinion using Syntactic Patterns [6] or Greetings.
- **Length** Greetings, Sum of all parsed sentences length, number of sentences, and token words after POS tagger.
- **Explicit Information Source** How many Quotations, Parentheses, Colons, and Named Entities (Stanford NER [13]).

These are Basic Features from the qualitative study. We added additional quantitative features to improve accuracy.

- **Language Feature** How many Nouns, Adjectives, Adverbs, and Sums of each sentiment scores for parsed sentences.
- **Meta Feature** Is there a parent comment in thread, User level value, *LIKES* score, how many Replies, if Deleted.
- **External Feature** Emoticons and N-gram words number.

3) Classification Result: The evaluators intention classifications difference was *Invoking* (25.8%) and *Evoking* (74.2%). Table I displays results using the Basic and Additional Features with Decision Tree (J4.5) [14] and LibSVM [15] algorithms using WEKA.

The results show over 92% accuracy. Additional features are ineffective in classification improvement. We analyzed which basic feature contributes to accuracy using

$$InfoGain(Class, Attribute) =$$
$$H(Class) - H(Class|Attribute) \quad (6)$$

and

$$GainRatio(Class, Attribute) =$$
$$(H(Class) - H(Class|Attribute))/H(Attribute) \quad (7)$$

The feature contribution results show a large portion of the classification ability depends on the Count Question (InfoGain: 0.5423, GainRatio: 0.44) attribute and other features are less than 0.047 for both InfoGain and GainRatio.

C. Effectiveness of the Methods

1) Settings: We picked the top five ranked comments according to *LIKES* score and our ICFinder method, eliminating duplicates. We generated an interview form with the randomly ordered comments with a Likert-type scale ranging from 0 (Not Useful) to 3 (Very Useful). We recruited six evaluators and trained them using the qualitative study findings. Next, we had them watch allocated videos, and then they judged their interview sheet manually. The overall Kappa agreement was 0.697. We used only six evaluators to evaluate the machine performance. Nonetheless, the results indicate that our methodology can achieve results comparable to human perspective.

Table I: Classification accuracy with *Decision Tree* and *LibSVM*

| Decision Tree					LibSVM			
Precision	Recall	F-Measure	Accuracy	Features	Precision	Recall	F-Measure	Accuracy
92.6	91.6	91.8	91.6	Basic	83.9	83.1	80.7	83.1
92.5	91.6	91.8	91.6	Basic + Language (Lang)	82.6	81.5	78.2	81.5
91.1	90.7	90.8	90.7	Basic + Meta	82.3	81	77.4	81
92.2	91.4	91.6	91.4	Basic + Ext Source	81.9	80.7	76.9	80.7
91.2	91.1	91.1	91.1	Basic + Lang, Meta	82.8	80.5	76.2	80.5
91.2	91	91.1	91	Basic + Meta, Ext	83.2	80.4	75.9	80.4
92	91.2	91.4	91.2	Basic + Lang, Ext	83.3	80.8	76.6	80.8
89.4	89.5	89.5	89.5	Basic + All	82.8	79.9	75.1	79.9
93.7	**92.3**	**92.6**	**92.3**	Basic + All With Attribute Selection	**90.1**	**90.2**	**90.1**	**90.2**

Pre: Precision, Rec: Recall, F: F-Measure, Acc: Accuracy
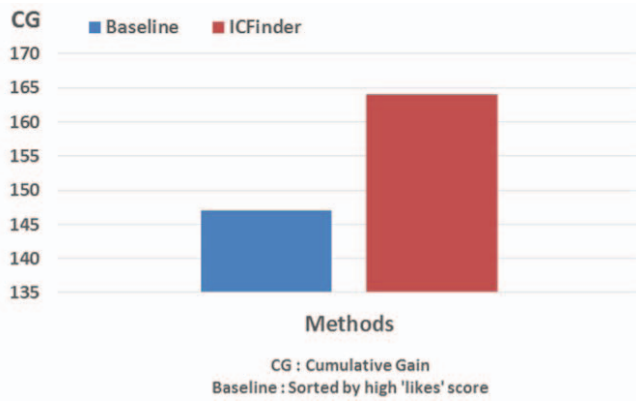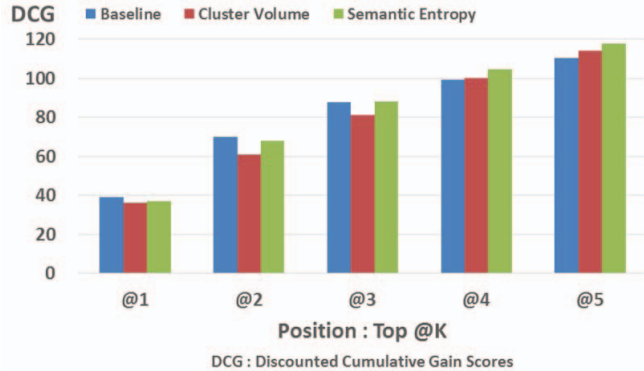

(a) Sum of Cumulative gain
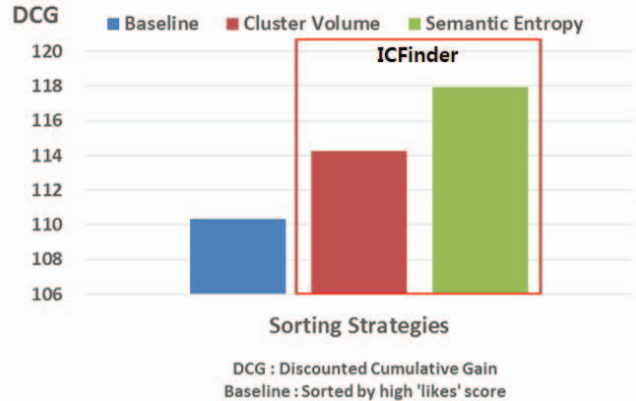
(b) Sum of Discounted Cumulative gain
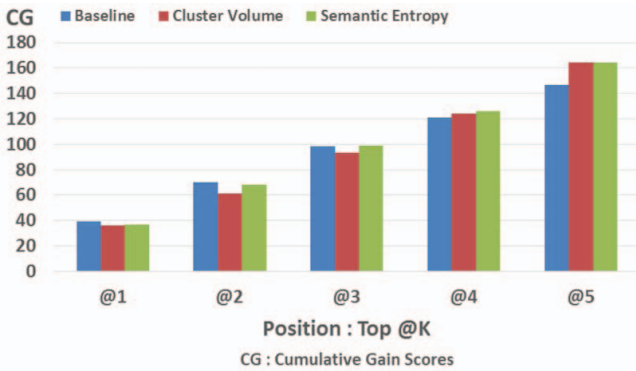
(c) Cumulative gain @K

(d) Discounted Cumulative gain @K

Figure 5: Quantitative results

2) Metric: We used Cumulative Gain (CG) and Discounted Cumulative Gain (DCG). CG shows the absolute number of scores evaluating the impact of comments. DCG measures the ranking quality based on multiple aspects and absolute informativeness score. We also used sorting strategy, consisting of sort by cluster volume and by *Semantic Entropy* score.

3) Results: Although we used a total number of 200 comments in 20 videos, the evaluation results show the method's advantages of incorporating user perspectives for recommending comments in watching video. Figure 5(a) shows the result of sum of CG. The absolute informativeness score of ICFinder is higher than the baseline by 11.56%. Figure 5(b) shows the result of the sum of DCG between baseline and two trials by different sorting strategies in our

method. As seen in the graph, the method of sorting by *Semantic Entropy* outperforms cluster volume and baseline by 3.22% and 6.84% respectively.

Figure 5(c) displays results for CG and Figure 5(d) for DCG. We analyzed the rank position @K in the graph. @ represents the Top positions and K the number of positions evaluated. Each ordinate contains the sum of CG and DCG scores for the baseline and for our method up to position (@K). The results of the top positions (@1) or (@2) show slightly better performance. However the results of low positions (@4) and (@5) outperform the baseline. The *Semantic Entropy* sorting method shows more effectiveness than cluster volume.

VI. Conclusion

The experiments show that our approach can successfully find informative comments. Our qualitative study highlights preferred features when users are watching videos. Our method shows superior results compared to the simple semantic approach. Additional work can tune the algorithm to find more fitting Question and Answer sets. It would be interesting to utilize informative comments for related information search. Our method might not be suited for cases in which users provide their information, opinions, and feelings in comments on a platform such as YouTube [16] because the purpose of leaving comments differs on the various platforms. The use of the ICFinder method for the extraction of informative comments provides the users with more relevant and useful information. We assume this data would include informative comments that are also satisfactory from the user's personal perspective.

References

[1] M. Hu and B. Liu. Mining and summarizing customer reviews. *In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168-177, 2004.

[2] Q. Li, J. Wang, Y. P. Chen and Z. Lin. User comments for news recommendation in forum-based social media. *Information Sciences*, 180(24), pages 4929-4939, 2010.

[3] F. Figueiredo, F. Belem, H. Pinto, J. Almeida, M. Goncalves, D. Fernandes, E. Moura and M. Cristo. Evidence of quality of textual features on the web 2.0. *In Proceedings of the 18th ACM conference on Information and knowledge management*, pages 909-918, 2009.

[4] E. Momeni, C. Cardie and M. Ott. Properties, Prediction, and Prevalence of Useful User-Generated Comments for Descriptive Annotation of Social Media Objects. *In Proceedings of the Seventh International Conference on Weblogs and Social Media*, 2013.

[5] H. Beyer and K. Holtzblatt. Contextual design: defining customer-centered systems. *Elsevier*, 1997.

[6] P. D. Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. *In Proceedings of the 40th annual meeting on association for computational linguistics*, pages 417-424, 2002.

[7] *Using Affinity Diagrams to make sense from Brainstorming.* http://www.leanyourcompany.com/methods/Using-Affinity-Diagrams.asp (Date last accessed on Oct. 7, 2016).

[8] T. K. Landauer, P. W. Foltz and D. Laham. An introduction to latent semantic analysis. *Discourse processes*, pages 259-284, 1998.

[9] C. Arndt. Information measures: information and its description in science and engineering. *Springer Science & Business Media*, 2012.

[10] *Weka 3: Data Mining Software in Java.* http://www.cs.waikato.ac.nz/ml/weka (Date last accessed on Oct. 7, 2016).

[11] A. P. Dempster, N. M. Laird and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1-38, 1977.

[12] K. Toutanova, D. Klein, C. D. Manning and Y. Singer. Feature-rich part-of-speech tagging with a cyclic dependency network. *In Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*, pages 173-180, 2003.

[13] J. R. Finkel, T. Grenager and C. Manning. Incorporating non-local information into information extraction systems by gibbs sampling. *In Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 363-370, 2005.

[14] I. H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. *Morgan Kaufmann*, 2005.

[15] C. C. Chang and C. J. Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3), pages 27, 2011.

[16] P. Schultes, V. Dorner and F. Lehner. Leave a Comment! An In-Depth Analysis of User Comments on YouTube. *Wirtschaftsinformatik*, pages 659-673, 2013.