Generating Comparative Summaries of Contradictory Opinions in Text

Yahoo-DAIS Seminar  Jan 26, 2010

Hyun Duk Kim (hkim277@illinois.edu) and ChengXiang Zhai (czhai@cs.uiuc.edu)

Department of Computer Science
University of Illinois, Urbana-Champaign
Web 2.0

• Huge volume of opinionated data
  – Blog, forum, reviews…

Google Chrome after a year: Sporting a new stable release

9/15/2008 10:08:00 AM
At Google Chrome’s first birthday, just two weeks ago, we looked forward to an action-packed year for the browser and the web. Thanks to a full year of great feedback from our users, we’re kicking off our second year of Google Chrome with a brand new stable release. This release comes hot on the heels of 51 developer, 21 beta and 15 stable updates and 3,605 bug fixes in the past year.

For you, that means significant speed improvements for the browser as well as a fresh redesign of some of its most loved features. To walk through the top highlights in the spirit of a September of epic sporting tournaments, here’s a play-by-play comparison of our brand new release against our previous releases:

A wicked serve, volley, and return: Fast, fast and fast

This new release of Google Chrome is faster than ever, as we continue to provide a modern browser that starts up quickly from your desktop, and is fast to load web pages and web applications.
Web 2.0

• Huge volume of opinionated data
  – Blog, forum, reviews...

Need of automatic summarization
Existing Opinion Summarization

- Focused on separating pos/neg opinions

[Hu and Liu ‘04, Liu et al. ‘05]

Still need to read texts!
Need of Contrastive Mining

• Biased opinion
  → General summarization

user reviews
  
appearance 93%

view: positive comments (78) | negative comments (6)

• Controversial opinion

user reviews
  
battery life 50%

view: positive comments (31) | positive comments (31)

Need Contrastive Mining!
Research Questions

1. Can we further summarize pos/neg separated opinions?
2. Can we show contradictory opinions more effectively?

Contrastive Opinion Summarization (COS)
Rest of the Talk

1. Problem definition

2. Method
   - Optimization framework
   - Similarity function
   - Approximation algorithm

3. Experiments

4. Demo Systems

5. Related work & Conclusion
Problem Definition
Contrastive Opinion Summarization

Input:
Pos/Neg sentence set

\[ X = \{x_1, \ldots, x_n\} \]

\[ Y = \{y_1, \ldots, y_m\} \]

Aligning Sentences
Contrastive Opinion Summarization

Output:

\[ X = \{x_1, \ldots, x_n\} \]

\[ Y = \{y_1, \ldots, y_m\} \]

\[ U = \{u_1\}_{i=1}^{k} \subset X \]

\[ V = \{v_1\}_{i=1}^{k} \subset Y \]

Contrastive Opinion Summary
Optimization Framework
Similarity Functions

\[\phi(s_1, s_2) \in [0,1]\]

\(s_1, s_2\):
Same polarity

\[\psi(u, v) \in [0,1]\]

\(u, v\):
Different polarity

Content Similarity Function

Contrastive Similarity Function
Contrastive Opinion Summarization

X (Pos)  U  V  Y (Neg)
X1  u1  v1  y1
X2  u2  v2  y2
X3  ...
X4  uk  vk  yk
X5  ...
Xn

Representativeness
Contrastiveness
Contrastive Opinion Summarization

\[ r(U, V) = \frac{1}{|X|} \sum_{x \in X} \max_{i \in [1,k]} \phi(x, u_i) + \frac{1}{|Y|} \sum_{y \in Y} \max_{i \in [1,k]} \phi(y, v_i) \]

Representativeness

Contrastiveness

\[ c(U, V) = \frac{1}{k} \sum_{i=1}^{k} \psi(u_i, v_i) \]
Objective Function

\[(U, V)^* = \arg \max_{(U, V)} (\lambda r(U, V) + (1 - \lambda)c(U, V))\]

\[= \arg \max_{(U, V)} \left( \frac{\lambda}{|X|} \sum_{x \in X} \max_{i \in [1, k]} \phi(x, u_i) + \frac{\lambda}{|Y|} \sum_{y \in Y} \max_{i \in [1, k]} \phi(y, v_i) \right)\]

\[+ \frac{1 - \lambda}{k} \sum_{i=1}^{k} \psi(u_i, v_i)\]

1. Define an appropriate content similarity function \(\Phi\)
2. Define an appropriate contrastive similarity function \(\psi\)
3. Solve the optimization problem efficiently.
1. Similarity Functions
Content Similarity Function

\[ \sum_{u \in s_1} \max_{v' \in s_2} \omega(u, v') \]

\[ \omega(u, v) \in [0,1]: \text{term similarity function} \]

Sentence 1: It has short battery time
Sentence 2: The battery life Isn’t great

Find max sim among words
Content Similarity Function

$$\sum_{u \in s_1} \max_{v' \in s_2} \omega(u, v') + \sum_{v \in s_2} \max_{u' \in s_1} \omega(u', v)$$

$$\omega(u, v) \in [0,1]: \text{term similarity function}$$

Sentence 1: It has short battery time

Sentence 2: The battery life Isn’t great

Same for the opposite side
Content Similarity Function

\[
\phi(s_1, s_2) = \frac{\sum_{u \in s_1} \max_{v' \in s_2} \omega(u, v') + \sum_{v \in s_2} \max_{u' \in s_1} \omega(u', v)}{|s_1| + |s_2|}
\]

\(\omega(u, v) \in [0, 1] : \text{term similarity function}\)

Sentence 1: It has short battery time

Sentence 2: The battery life Isn’t great

Normalize by the total number of words
Word Similarity Function

1. Word Overlap (WO):
   \[ \omega_{WO}(u, v) = \begin{cases} 1 & \text{iff } u = v, \\ 0 & \text{otherwise} \end{cases} \]
   – Same as Jaccard similarity function

2. Semantic Word Matching (SEM):
   \[ \omega_{SEM}(u, v) = \begin{cases} 1 & \text{if } u = v, \\ \gamma \text{sim}(u, v) & \text{otherwise} \end{cases} \]
   – \( \gamma \): strength of semantic matching
   \[ \gamma = 0 \quad \Rightarrow \quad \omega_{SEM}(u, v) = \omega_{WO}(u, v) \]
   – \( \text{sim}(u, v) \): semantic term similarity by WordNet:Similarity tool
Contrastive Similarity Function

• How well two sentences with opposite opinions match up with each other

Sentence 1: It has long battery time
Sentence 2: The battery life isn’t great

• Remove sentiment related words negation words and adjectives
→ Define $\psi$ in the same way as $\Phi$
2. Approximation Algorithms
Optimization Algorithm

Too many Combinations

→ Two greedy ways to optimize the objective function
Two approximation Algorithms

\[(U, V)^* = \arg \max_{(U,V)} (\lambda r(U,V) + (1 - \lambda)c(U,V))\]

- **Representativeness First**
- **Contrastiveness First**
Representativeness-First Approximation

X (Pos)

X1
X2
X3
X4
X5
...
Xn

Y (Neg)

y1
y2
y3
y4
...
ym
Representativeness-First Approximation

Find k cluster,
Representativeness-First Approximation

Find k cluster,
→ Find contrastive pairs
Representativeness-First Approximation

• Generate k-clusters, $X \rightarrow \{U_1, \ldots, U_k\}$, $Y \rightarrow \{V_1, \ldots, V_k\}$

• Reformulate objective function

$$g(S) = \sum_{i=1}^{k} g_i(u_i, v_i)$$

$$g_i(u_i, v_i) = \lambda \left( \frac{1}{|X|} \sum_{x \in U_i} \phi(x, u_i) + \frac{1}{|Y|} \sum_{y \in V_i} \phi(y, v_i) \right) + \frac{1 - \lambda}{k} \psi(u_i, v_i)$$

$$(u_i^*, v_i^*) = \arg \max_{u_i \in U_i, v_i \in V_i} g_i(u_i, v_i)$$

**Representativeness**

**Within cluster**

**Contrastiveness**

**of the pair**
Contrastiveness-First Approximation

X (Pos)          Y (Neg)

X_1
X_2
X_3
X_4
X_5
...
X_n

y_1
y_2
y_3
y_4
...
y_m
Contrastiveness-First Approximation

Find contrastive pair,
Contrastiveness-First Approximation

**X (Pos)**

<table>
<thead>
<tr>
<th>X1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X2</td>
<td></td>
</tr>
<tr>
<td>X3</td>
<td></td>
</tr>
<tr>
<td>X4</td>
<td></td>
</tr>
<tr>
<td>X5</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>Xn</td>
<td></td>
</tr>
</tbody>
</table>

**Y (Neg)**

<table>
<thead>
<tr>
<th>y1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>y2</td>
<td></td>
</tr>
<tr>
<td>y3</td>
<td></td>
</tr>
<tr>
<td>y4</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>ym</td>
<td></td>
</tr>
</tbody>
</table>

Find contrastive pair, → Select representative pairs
Contrastiveness-First Approximation

- Reformulate objective function

\[ (u_i^*, v_i^*) = \arg \max_{u_i, v_i} \lambda r(S_i) + (1 - \lambda)c(S_i) = \arg \max_{u_i, v_i} g(u_i, v_i) \]

\[ = \arg \max_{u_i, v_i} \lambda \left( \frac{1}{|X|} \sum_{x \in X_{u_i}} \phi(x, u_i) + \frac{1}{|Y|} \sum_{y \in Y_{v_i}} \phi(y, v_i) \right) + \frac{1 - \lambda}{k} \psi(u_i, v_i) \]

\[ S_{i-1} = \{(u_j, v_j)\}_{j=1}^{i-1} \text{ : already chosen } i-1 \text{ pairs} \]

\[ X_{u_i} = \{x \in X \mid \phi(x, u_i) > \phi(x, u_j) \forall j = 1, \ldots, i-1\} \]

\[ Y_{v_i} = \{y \in Y \mid \phi(y, v_i) > \phi(y, v_j) \forall j = 1, \ldots, i-1\} \]

- After choosing the first pair \((u_1, v_1)\), we would iteratively choose \((u_i, v_i)\) to maximize the “gain function”, \(g(u_i, v_i)\)
How to set $k$?

• $k$: the number of target summary

• Set heuristically by the following formula

\[ k = 1 + \log_2(|X| + |Y|) \]

• In tuition
  – Larger $k$ if we have more sentences to summarize
  – The growth saturate as the number of sentences becomes very large
Experiments
Design

• Opinion separated data set
  – 12 product + 1 non-product data

• Two human assessors for gold standards

• Measures
  – Precision of pairing
  – Aspect Coverage of clusters
Human Gold-standard

X (Pos)

X1
X4
X3
X2
X7
...
Xn

Find clusters

Y (Neg)

y1
y2
y4
y6
...
ym
Human Gold-standard

Find clusters and align them
Human Gold-standard

Measures:
Precision, Aspect Coverage
<table>
<thead>
<tr>
<th>No</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>oh ... and file transfers are fast &amp; easy.</td>
<td>you need the software to actually transfer files</td>
</tr>
<tr>
<td>2</td>
<td>i noticed that the micro adjustment knob and collet are well made and work well too.</td>
<td>the adjustment knob seemed ok, but when lowering the router, i have to practically pull it down while turning the knob.</td>
</tr>
<tr>
<td>3</td>
<td>the navigation is nice enough, but difficult navigation - i wo n’t necessarily say &quot; difficult ,” but i do n’t enjoy the scrollwheel to navigate.</td>
<td>scrolling and searching through thousands of tracks , hundreds of albums or artists , or even dozens of genres is not conducive to save driving</td>
</tr>
<tr>
<td>4</td>
<td>i imagine if i left my player untouched (no backlight) it could play for considerably more than 12 hours at a low volume level.</td>
<td>there are 2 things that need fixing  first is the battery life.  it will run for 6 hrs without problems with medium usage of the buttons.</td>
</tr>
</tbody>
</table>
**Sample Result**

<table>
<thead>
<tr>
<th>No</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>oh ... and file transfers are fast &amp; easy.</td>
<td>you need the software to actually transfer files</td>
</tr>
<tr>
<td>2</td>
<td>i noticed that the micro adjustment knob and collet are well made and work well too.</td>
<td>the adjustment knob seemed ok, but when lowering the router, i have to practically pull it down while turning the knob.</td>
</tr>
</tbody>
</table>

**Different polarities of opinions made from different perspectives.**
<table>
<thead>
<tr>
<th>No</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>The navigation is nice enough, but scrolling and searching through thousands of tracks, hundreds of albums or artists, or even dozens of genres is not conducive to save driving</td>
<td>Difficult navigation - I won’t necessarily say &quot;difficult,&quot; but I don’t enjoy the scrollwheel to navigate.</td>
</tr>
</tbody>
</table>

**Positive vs. negative**

Not much disagreement
<table>
<thead>
<tr>
<th>No</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>i imagine if i left my player untouched (no backlight) it could play for considerably more than 12 hours at a low volume level.</td>
<td>there are 2 things that need fixing first is the battery life. it will run for 6 hrs without problems with medium usage of the buttons.</td>
</tr>
</tbody>
</table>
Questions to answer

1. Representative-first approximation (RF) vs. contrastive-first approximation (CF), which optimization algorithm performs better?

2. Can semantic matching help to improve performance?

3. Is excluding sentimental words when computing the contrastive similarity beneficial?
Rep-fist vs. Contrast-first

AC/P

RF-AC

CF-AC

RF-P

CF-P

RF-Precision

RF-Aspect Coverage

CF-Precision

CF-Aspect Coverage
Rep-fist vs. Contrast-first

Aspect coverage is higher than precision → Easier to achieve representativeness than contrastiveness.
Rep-fist vs. Contrast-first

CF-AC
RF-AC
CF-P
RF-P

AC/P

RF outperforms RF
→ More important to optimize contrastiveness
Semantic Term Matching

AC/P

No Semantic Matching  Full Semantic Matching

RF-Precision  RF-Aspect Coverage  CF-Precision  CF-Aspect Coverage
Semantic term matching does not help. 
→ as increase the value of $\gamma$, the performance even drops.
Contrastive Similarity Heuristic

• For contrastive similarity function.

→ The heuristic of removing sentimental words is effective.

<table>
<thead>
<tr>
<th>Opt. Method</th>
<th>Precision</th>
<th></th>
<th>Aspect Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RF</td>
<td>CF</td>
<td>RF</td>
</tr>
<tr>
<td>WO+all words</td>
<td>0.484</td>
<td>0.531</td>
<td>0.737</td>
</tr>
<tr>
<td>WO</td>
<td>0.503</td>
<td>0.537</td>
<td>0.737</td>
</tr>
<tr>
<td>SEM+all words</td>
<td>0.470</td>
<td>0.507</td>
<td>0.718</td>
</tr>
<tr>
<td>SEM</td>
<td>0.500</td>
<td>0.540</td>
<td>0.763</td>
</tr>
</tbody>
</table>

_table. Effectiveness of removing sentimental words in computing contrastive similarity_
Demo System
Data Set & Demo System

• Data set
  – Data set used for experiment with human evaluation

• Demo Systems
  1. Simple COS Generation
  2. Open Opinion Search Engine – Ooops!

• http://sifaka.cs.uiuc.edu/ir/data/cos/ [link]
Contrastive Opinion Summarization Research Data

- This page has links to the data set and demo page for contrastive opinion summary.
  *Hyun Duk Kim and ChengXiang Zhai, Generating Comparative Summaries of Contradictory Opinions in Text, CIKM 2009*
  
- Page created by Hyun Duk Kim and ChengXiang Zhai of the TIMan group at UIUC

**[Demo system]**

- Demo2: Coops - Open Opinion Search engine using Contrastive opinion summary
- Demo1: Contrastive opinion summary generation with user input.

**[Data set]**

- Data set used for the experiment of the 'Contrastive Opinion Summarization' [link]
  
- When use this data set, please use the following publication as a reference.
  *Hyun Duk Kim and ChengXiang Zhai, Generating Comparative Summaries of Contradictory Opinions in Text, CIKM 2009*

- Details about how this data set was made and used are on the publication.

- Contacts:
  *Hyun Duk Kim (dkim277 at illinois.edu), [HomePage]*
  *ChengXiang Zhai (czhai at cs.uic.edu), [HomePage]*

- File name
  *ProductName_[Feature]*

- Format
  *[(polarity)] (sentence number) - (sentence)*

- [1]: positive
- [0]: negative

Ex:

[1] 1 - the height adjustment knob is a bit flimsy, but it works well and provides good accuracy.

[1] 2 - i noticed that the micro adjustment knob and collet are well made and work well too.

[1] 3 - i don't have problems with the depth adjustment, the speed control, the collet (i haven't tried the 1/4" sleeve yet).

[1] 4 - it is a great router for a table and with the depth adjustment knob it is easy to adjust depth precisely.

[1] 5 - the depth adjustment and b
**How to use:**

- Input one sentence set (ex: positive) on the first textarea, the opposite sentiment sentence set (ex: negative) on the next textarea, and click the 'Generate' button.
- One line of input is assumed as one sentence sentence.
- To get meaningful results, input at least 2 sentences to each textarea.
- Following are sample inputs and outputs:
  - **ApexAD2600 Progressive-scan DVD player player**
  - **MicroMP3 battery-life**
  - **MicroMP3 design**
  - **MicroMP3 headphones**
  - **MicroMP3 software**
  - **Nokia 6600 battery-life**
  - **Creative Labs Nomad Jukebox Zen Xtra 40GB navigation**
  - **Creative Labs Nomad Jukebox Zen Xtra 40GB software**
  - **Creative Labs Nomad Jukebox Zen Xtra 40GB size**
  - **Creative Labs Nomad Jukebox Zen Xtra 40GB weight**
  - **Creative Labs Nomad Jukebox Zen Xtra 40GB transfer**
  - **Hitachi router adjustment**
  - **aspartic acid**

[Generate Comparative Summary!](#)
Contrastive Opinion Summary

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  it works great.</td>
<td>it has never worked.</td>
</tr>
<tr>
<td>2  this is the best dvd player i've purchased.</td>
<td>i have had this player for 3 months and have been able to see a total of 6 dvd's, after the frustration and aggravation of getting the player to do its job.</td>
</tr>
<tr>
<td>i took it to my father's house to play a tom jones concert dvd and he loved the player so much i gave it to him.</td>
<td>i try navigating using the player's built in menu, and it still wouldn't work!</td>
</tr>
<tr>
<td>4  so far, we've only used this player a few times to watch movies and it has worked fine.</td>
<td>when the player was working it was great, but...</td>
</tr>
<tr>
<td>5  i just hooked it up and everything seemed to be going fine.</td>
<td>this player is not worth any price and i recommend that you don't purchase it.</td>
</tr>
<tr>
<td>6  i bought mine in december of 2003, and have had no real problems with it.</td>
<td>2600 about a month ago and have had numerous problems with it.</td>
</tr>
<tr>
<td>i've had the player for about 2 years now and it still performs nicely with the exception of an occasional wwhhhrr sound from the motor.</td>
<td>this item broke just two weeks after i gave it to my kids for christmas.</td>
</tr>
</tbody>
</table>
Simple COS Generation

• COS generation
  – CF approximation with no semantic matching
  – $\Lambda = 0.5$

• HTML, Php, Python
Contrastive Opinion Summarization Research Data

- This page has links to the data set and demo page for contrastive opinion summary.
  
  Hyun Duk Kim and ChengXiang Zhai, Generating Comparative Summaries of Contradictory Opinions in Text, CIKM 2009
  
- Page created by Hyun Duk Kim and ChengXiang Zhai of the TIM group at UIUC.

[Demo system]

- Demo2: Coopst - Open Opinion Search engine using Contrastive opinion summary
- Demo1: Contrastive opinion summary generation with user input.

[Data set]

- Data set used for the experiment of the 'Contrastive Opinion Summarization' link.
- When use this data set, please use the following publication as a reference.
  
  Hyun Duk Kim and ChengXiang Zhai, Generating Comparative Summaries of Contradictory Opinions in Text, CIKM 2009
- Details about how this data set was made and used are on the publication.
- Contacts:
  
  Hyun Duk Kim (dkim277 at illinois.edu), [HomePage]
  ChengXiang Zhai (czhai at cs.uiuc.edu), [HomePage]
- Filename
  
  [ProductName] [Feature]
- Format
  
  [(polarity)] (sentence number) - (sentence)

| [1]: positive |
| [0]: negative |

Ex:

[1] 1 - the height adjustment knob is a bit flimsy, but it works well and provides good accuracy.
[1] 2 - i noticed that the micro adjustment knob and collet are well made and work well too.
[1] 3 - i don't have problems with the depth adjustment, the speed control, the collet (i haven't tried the 1/4" sleeve yet).
[1] 4 - it is a great router for a table and with the depth adjustment knob it is easy to adjust depth precisely.
[1] 5 - the depth adjustment and b
Searched for: the lord of the rings
Results 1 - 50 of 7597 matches

1. [14411]
   ...CNO>14411</DOCNO> -Title- One of the best series ever! -Rating- 5.0 -Pros- It's THE LORD OF THE RINGS! -Cons- The...
   (Score: 0.495531)

2. [14459]
   .../DOCNO> -Title- What I Think: Lord of the Ring:Updated 4/24/02 -Rating-...
   (Score: 0.495434)

3. [14399]
   ... <DOCNO>14399</DOCNO> -Title- Lord of the Rings Changed Publishing ...
   (Score: 0.490453)
1. [14411]
   - Even if the characters were just talking, I was really interested in what was said.
   - I'm not sure how something can be so good, it can't belong in a group.
   - The Lord of The Rings is a series that you will want to read over and over, just to see all of the characters again.
   - -TitleOne of the best series ever! -Rating 5.0 -ProsIt's THE LORD OF THE RINGS! -ConsThere are only 4 books in the series -The Bottom LineIf you like fantasy, or are a complete newbie, you will need to read Lord of the Rings. -Full Review-
   - I can't imagine a fantasy fan who hasn't read this yet, but I have heard of them.
   - If you are worried about not liking it because you're not a die hard fantasy fan, have no worries.
1. [14411]

- Even if the characters were just talking, I was really interested in what was said.
- I'm not sure how something can be so good, it can't belong in a group.
- The Lord of The Rings is a series that you will want to read over and over, just to see all of the characters again.
- TitleOne of the best series ever! - Rating 5.0 - Pros It's THE LORD OF THE RINGS! - Cons There are only 4 books in the series.
Even if the characters were just talking, I was really interested in what was said.

I'm not sure how something can be so good, it can't belong in a group.

The Lord of The Rings is a series that you will want to read over and over, just to see all of the characters again.

-TitleOne of the best series ever! -Rating5.0 -ProsIt's THE LORD OF THE RINGS! -ConsThere are only 4 books in the series

-The Bottom LineIf you like fantasy, or are a complete newbie, you will need to read Lord of the Rings. -Full Review-
Contrastive Opinion Summary

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Lord of The Rings is a series that you will want to read over and</td>
<td>All of the characters are a breath of fresh air after reading many books where big guys were always stupid, girls were always willing to</td>
</tr>
<tr>
<td>over, just to see all of the characters again.</td>
<td>go out with you, and the heroes never die.</td>
</tr>
<tr>
<td>I would read the book called The Hobbit before I read the lord of the</td>
<td>If you are worried about not liking it because you're not a die hard fantasy fan, have no worries.</td>
</tr>
<tr>
<td>rings so that I could understand the lord of the rings a lot better.</td>
<td></td>
</tr>
<tr>
<td>TitleWhat I Think: Lord of the Rings.Updated 4/24/02.Rating:5.0</td>
<td>The story comes to an end - The Bottom Line would recommend that you by it.</td>
</tr>
<tr>
<td>ConsThe story comes to an end - The Bottom Line would recommend that you</td>
<td>It is really hard for me to explain what it is that is so great about the series.</td>
</tr>
<tr>
<td>by it</td>
<td></td>
</tr>
<tr>
<td>I'm not sure how something can be so good, it can't belong in a group.</td>
<td>I can't imagine a fantasy fan who hasn't read this yet, but I have heard of them.</td>
</tr>
</tbody>
</table>

1. [14411]
   🍂 Even if the characters were just talking, I was really interested in what was said.
   🍂 I'm not sure how something can be so good, it can't belong in a group.
   🍂 The Lord of The Rings is a series that you will want to read over and over, just to see all of the characters again.
   🍂 TitleOne of the best series ever! -Rating:5.0 -ProsIt's THE LORD OF THE RINGS! -ConsThere are only 4 books in the series -The Bottom LineIf you like fantasy, or are a complete newbie, you will need to read Lord of the Rings. -Full Review-
   🍂 I can't imagine a fantasy fan who hasn't read this yet, but I have heard of them.
   🍂 If you are worried about not liking it because you're not a die hard fantasy fan, have no worries.
   🍂 It's really hard for me to explain what it is that is so great about the series.
Ooops! – Open Opinion Search

• Architecture
  – Lemur CGI interface, Opinion finder
  – Php, Python, Google visualization API
Ooops! – Open Opinion Search

• Reviews from Epinions.com
  – 13,000 reviews of 6 different products groups

• COS generation
  – CF approximation with no semantic matching
  – $\Lambda = 0.5$
Related work & Conclusion
Related Work

• Opinion summarization
  – Probabilistic methods [Mei et al. ’07, Titov and McDonald ‘08]
  – Heuristic rule based methods [Liu et al. ‘05, Hu and Liu ’04, Hu and Liu ’06, Zhuang ’06, Pang et al. ‘02, Ku et al. ‘06]
  – Others: opinion integration [Lu and Zhai ‘08]

• Contradictory mining
  – Comparative sentence mining
    [Zindal and Liu ‘06, Marneffe et al. ‘08, Thomas et al. ‘06]
  – Visualization [Chen et al. ‘06]

• Sentence similarity measures
  – [Achananuparp et al. ‘08]
Conclusion

• Novel summarization problem
  – Contrastive opinion summary (COS)
  – Optimization framework / Approximation methods
  – Explored different similarity measures
  – Demo systems
Future Work

• Improve algorithm
  – More experiments on additional larger data sets
  – Use advanced semantic term matching methods
  – Study how to develop better approximate solutions

• Improve system
  – Better opinion analysis module
  – Feed back from demo systems

• Contradictory mining
  – Finding contradiction in general
Thank you

Hyun Duk Kim (hkim277@illinois.edu)
ChengXiang Zhai (czhai@cs.uiuc.edu)
Reference


Reference


Reference


Reference


