



Opinion summarization on spontaneous conversations[☆]

Dong Wang, Yang Liu

Department of Computer Science, The University of Texas at Dallas, 800 West Campbell Road, Richardson, TX 75080, United States

Received 31 January 2014; received in revised form 2 February 2015; accepted 30 April 2015

Available online 11 May 2015

Abstract

In this study we explore opinion summarization on spontaneous conversations using unsupervised and supervised approaches. We annotate a phone conversation corpus with reference extractive and abstractive summaries for a speaker's opinion on a given topic. We investigate two methods: the first is an unsupervised graph-based method, which incorporates topic and sentiment information, as well as sentence-to-sentence relations extracted based on dialogue structure; the second is a supervised method that casts the summarization problem as a classification problem. Furthermore, we investigate the use of pronoun resolution in this summarization task. We develop various features based on pronoun coreference and incorporate them in the supervised opinion summarization system. Our experimental results show that both the graph-based method and the supervised method outperform the baseline approach, and the pronoun related features can help to generate better summaries.

© 2015 Elsevier Ltd. All rights reserved.

Keywords: Opinion summarization; Switchboard corpus; Sentiment analysis; Graph-based summarization; Pronoun resolution

1. Introduction

There is a growing interest in sentiment analysis in the Natural Language Processing (NLP) community. Most of the previous research focuses on identifying the polarity or subjectivity of a text document or sentence. However, in many cases people's opinion is mixed or vague, and it is hard to measure using simple scales such as positive, negative or neutral. In the following examples, the first one shows two sentences extracted from product reviews. These short sentences express mixed opinions of pros and cons. The second example is a person's opinion on "gun control", extracted from a conversation transcript. This person is against gun control but agrees that there should be some legislation.

[Example 1] (Ganesan et al., 2010)

iPhone's battery is bulky but it is cheap...

iPhone's battery is bulky but it lasts long

[Example 2]

well, on a scale of one to ten, uh, being ten, no, kind of legislation and zero being, uh, total ban, I probably would lean more towards six or seven.

[☆] This paper has been recommended for acceptance by Edward J Briscoe.

E-mail addresses: dongwang@hlt.utdallas.edu (D. Wang), yangl@hlt.utdallas.edu (Y. Liu).

*um, I feel like a total ban on guns is just going to put the guns in the hands the criminals.
I shouldn't say I don't know if it was any worse,
but it certainly didn't get any better.
I think that the law are way too lax.
I, I agree the thing that scares me, uh, though about where I would definitely want some sort of legislation
I don't believe that people should be allowed to carry guns in their vehicles.*

For cases shown above, it is hard to determine whether the sentiment is positive or negative. Moreover, users may be more interested in the reasons behind the opinions, rather than a simple answer such as “support or against” or “like it or not”. Under these circumstances, an opinion-based summary is able to better represent people’s opinion.

Different from generic summarization, an opinion-based summary summarizes people’s opinion rather than objective facts. It can be used in summarizing reviews or comments where we care more about what people think about a person or a product. In this study, we investigate opinion summarization on spontaneous conversations. The task is defined as, given a conversation and a topic, a summarization system needs to generate a summary of the speaker’s opinion towards the topic. This is useful for many applications, especially for processing the increasing amount of conversation recordings (e.g., telephone conversations, customer service, round-table discussions, interviews in broadcast programs) where we often need to find a person’s opinion or attitude, for example, “what does the speaker think about capital punishment and why?”. This kind of questions can be treated as a topic-oriented opinion summarization task.

In this work, we annotate a subset of the Switchboard corpus with human summaries. This is one of the first corpora for this study. We compare two methods that have been widely used in extractive summarization: graph-based and supervised methods. Our system attempts to incorporate more information about topic relevance and sentiment scores. In addition, in the graph-based method we propose to incorporate the dialogue structure information in order to select salient summary sentences. Next we exploit the use of pronoun resolution in spontaneous conversation summarization. We analyze the referent types in spontaneous conversations and study the potential improvement by using pronoun information. Then we propose a variety of pronoun related features and incorporate them in the supervised summarization approach. Our experimental results show that both the unsupervised and supervised methods achieve better results compared to the baseline, and the pronoun features are helpful to extract better summary sentences. In addition, we measure the performance of three automatic coreference resolution systems by comparing with human annotation, and find that all of the three tools cannot perform well on the spontaneous conversation domain for our defined pronoun set. It shows the limit of using automatically generated pronoun resolution results in conversation summarization.

This paper is an extension of our previous work in [Wang and Liu \(2011\)](#). Our new main contributions in this paper are: (i) a new study of sentiment analysis for conversations, including data annotation, automatic classification, and an analysis of relationship between subjective sentences and opinion summaries, (ii) use of a supervised approach and comparison with the graph-based method for opinion summarization, (iii) using both human transcript and speech recognition output for summarization, and (iv) exploration of pronoun information in summarization of conversations.

2. Related work

Research in document summarization has been well established over the past decades. Previous studies have used various domains, including news articles, scientific articles, web documents and reviews. Recently there is an increasing research interest in speech summarization, such as conversational telephone speech ([Zhu and Penn, 2006](#); [Zechner, 2002](#)), broadcast news ([Maskey and Hirschberg, 2005](#); [Lin et al., 2009](#)), lectures ([Zhang et al., 2007](#); [Furui et al., 2004](#)), meetings ([Murray et al., 2005](#); [Xie and Liu, 2010](#)), and voice mails ([Koumpis and Renals, 2005](#)). In general speech domains seem to be more difficult than well-written text for summarization.

In previous work, both unsupervised and supervised methods have been studied for text or speech summarization.

- Unsupervised methods: The Maximal Marginal Relevance (MMR) approach selects the most salient sentences and at the same time avoids redundancy. The salience is measured by the similarity of a sentence and the whole document, and redundancy is measured using the similarity of a sentence and the sentences already selected in the summary. Latent Semantic Analysis (LSA) applies singular value decomposition (SVD) on a term by sentence matrix and then chooses the most important sentences. Graph-based methods ([Erkan and Radev, 2004](#); [Garg et al., 2009](#); [Chen](#)

Download English Version:

<https://daneshyari.com/en/article/559013>

Download Persian Version:

<https://daneshyari.com/article/559013>

[Daneshyari.com](https://daneshyari.com)