SHORT TEXT SUMMARIZATION EXTRACTION USING CONCISE INTERFACE

Soniya S Dadhania
Lecturer in Computer Engineering
R C Technical Institute, Ahmedabad

ABSTRACT: We focus on the problem of summarizing short text comments in the stream of a specific social networking service (SNS) message. The amount of feedback can increase at a high rate immediately after a social message is published. Motivated by the fact that users may want to get a brief understanding of a comment stream without reading the full list of comments, we try to group similar content together with the comments and generate a brief summary of the opinion for this message. Since multiple users can request synthesis at any time, existing clustering methods cannot be applied directly and cannot meet the real-time requirements of the application. Additionally, we have modeled a new incremental clustering problem for SNS stream summary comments. In addition, we propose an algorithm that can IncreSTS incremental update the group results with the latest inbound real-time feedback. In addition, we have designed a quick display interface to help users quickly and easily get an overview. From extensive experimental results and real-world demonstration, we have found that IncreSTS has the benefits of high efficiency, high scalability, and better management of abnormal values, justifying the vitality of the IncreSTS problem destination.

1. INTRODUCTION

In recent years, social network services are very widespread and have become important communication platforms in our daily life. The largest social networking site Facebook presented the statistics in 2012. According to it, an average of 3.2 billion interactions is generated each day which includes likes and comments. Besides this, Twitter also has millions of users and thus huge amount of messages are posted in a day. All such existing social platforms are very convenient to use and thus have gained high popularity among people. Due to this reason, the celebrities, corporations, and organizations also create their own social pages to interact with their fans and the public. For each message, users can express their opinions by forwarding, giving a like, and leaving comments on it. Due to popularity of these platforms, not only the quantity of comments is large, but also the generation rate is remarkably high. Mostly, celebrities and corporations have high interest to know how their fans and customers react to certain topics and content. Thus it has created the necessity to develop an advanced summarization technique for comment streams in SNS. Traditional comment streams generally express more complete information, such as the discussion on products or movies. But here the main focus is on comment streams in social network sites that are in short text style with casual language. For each social message, main objective is to cluster comments with similar content together and generate a proper opinion summary for that message. For each different groups of opinions, easy and rapid overview should be generated for users and thus an efficient and effective technique should be applied to identify the clusters of all comments of a particular social message. Grouping similar comments leads to formation of different clusters. These clusters then can be used for summarizing the comment streams from social network sites. Summarizing is defined as reducing text or any content to one-third or one-quarter its original size, clearly indicating its meaning, and retaining main thoughts expressed. The purpose of summarization is to briefly present the key points of any content in order to provide proper context for user. Summarizing is useful in many types of writing and at different points in the writing process. Summarizing is useful in many other aspects such as provide context for a paper‟s thesis, write literature reviews, and annotate a bibliography. The benefit of summarizing is it allows the reader to contextualize what people are saying, which is very vital in case of huge amount of social media contents generated every day. In addition to this, summarizing helps the user to gain a better sense of what exactly the information or content is conveying.

2. EXISTING SYSTEM

Numerous studies and systems have proposed techniques and mechanisms to generate various types of summaries on comment streams. One major category aims to extract representative and significant comments from messy discussion. Like YouTube and Face book, these popular services allow users to determine whether a comment is useful or recommendable, and the comments with the top-k most endorsements are displayed on the top of the list. This category relies on user contributions and intends to leverage the wisdom of crowds. On the other hand, some researchers model this problem as recommendation or classification tasks and employ machine learning techniques to solve it. Moreover, sentiment analysis has been applied as well to discover hidden emotions in messages. Furthermore, providing an informative presentation interface is another active research field on the summarization of social messages. Despite some effort has been spent on solving this information overload problem, a generalized approach for summarizing rapid-increasing comment streams in SNS, based on text content, is yet to be fully explored.

Disadvantages:
1. Information overloaded problem.
2. More amounts of comments are displayed as a summarized content.
3. Similar comments are not removed.
4. Unstructured texts provided by the previous approaches.

3. PROPOSED SYSTEM

We explore the problem of incremental short text summarization on comment streams from social network services. We model this problem as an incremental clustering task and propose the IncreSTS (standing for Incremental Short Text Summarization) algorithm to discover the top-k clusters including different groups of opinions towards one social message. For each comment cluster, important and common terms will be extracted to construct a keyterm cloud. This key-term cloud provides an at-a glance presentation that users can easily and rapidly understand the main points of similar comments in a cluster. Moreover, representative comments
in each group will also be identified. Our objective is to generate an informative, concise, and impressive interface that can help users get an overview understanding without reading all comments.

Advantages:
1. Provide effective short text summarization result.
2. High efficiency clustering results
3. Remove similar comments information.
4. Provide informative and impressive summarization results

SCREENSHOTS

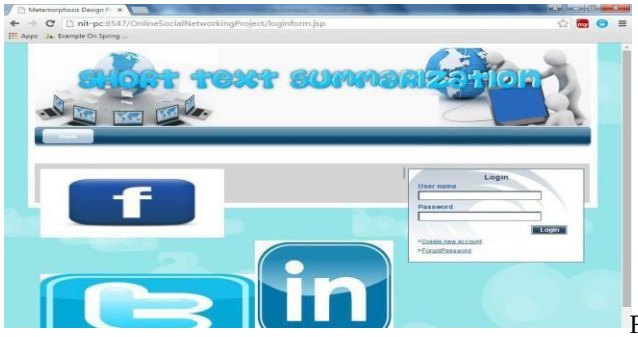


Fig : Login Page

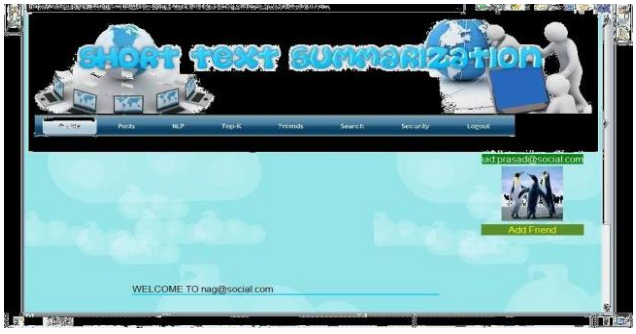


Fig: Add Friend

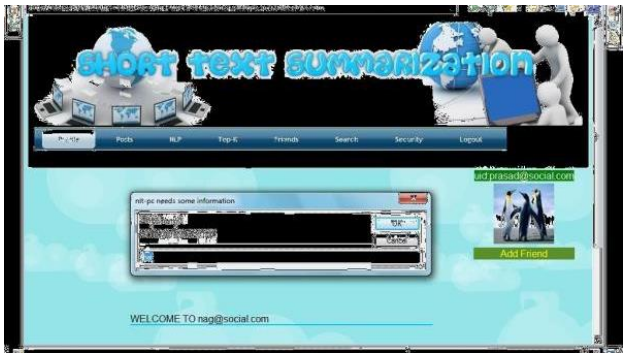


Fig: write welcome message

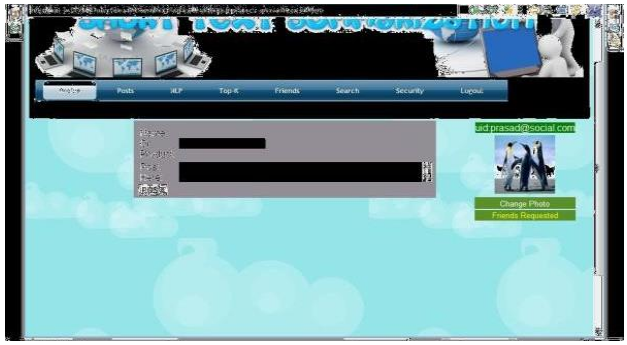


Fig: Post product

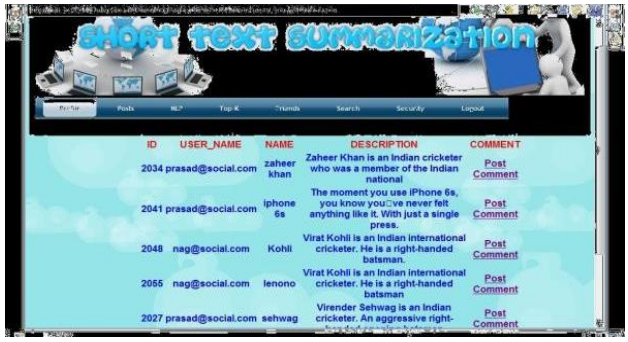


Fig: view Posted Content

Fig: Content graph

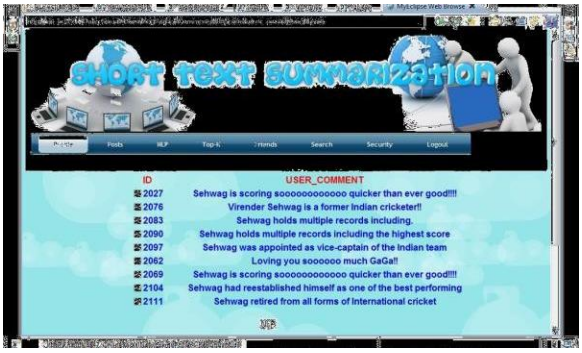
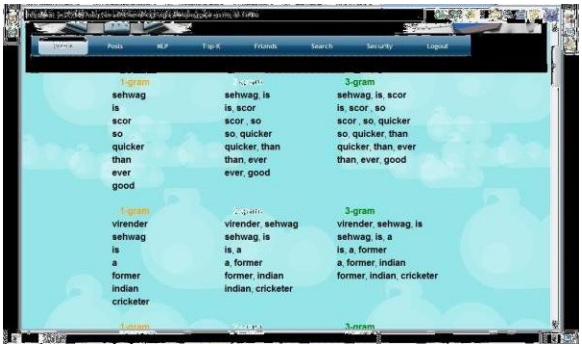


Fig: NLP

Fig: N-gram



4. CONCLUSION

The proposed system, IncreSTS (Incremental Short Text Summarization), addresses the challenges of summarizing rapidly increasing comment streams in social networking services (SNS). By modeling the problem as an incremental clustering task, the system effectively groups similar comments, extracts key terms, and generates concise summaries for users. This approach not only overcomes the limitations of existing systems, such as information overload and unstructured text, but also provides a scalable, efficient, and user-friendly interface for quick comprehension of social message discussions.

Key advantages of the system include high-efficiency clustering, removal of redundant comments, and the generation of visually appealing and informative summaries using key-term clouds. These features make IncreSTS a practical solution for handling the dynamic nature of SNS comment streams, offering users an enhanced ability to understand and manage large volumes of feedback in real time.

REFERENCES

[1] IncreSTS: Towards Real-Time Incremental Short Text Summarization on Comment Streams from Social Network Services---Cheng-Ying Liu, Ming-Syan Chen, Fellow, IEEE, and Chi-Yao Tseng.

[2] T. Preis, H. S. Moat, and E. H. Stanley, “Quantifying trading behavior in financial markets using Google trends,” Sci. Rep., vol. 3, p. 1684, 2013.

[3] H. Choi and H. Varian, “Predicting the present with Google trends,” Econ. Rec., vol. 88, no. s1, pp. 2–9, 2012.

[4] C.-T. Ho, R. Agrawal, N. Megiddo, and R. Srikant,, “Range queries in OLAP data cubes,” ACM SIGMOD Rec., vol. 26, no. 2, pp. 73–88, 1997.

[5] G. Mishne, J. Dalton, Z. Li, A. Sharma, and J. Lin, “Fast data in the era of big data: Twitter,,s real-time related query suggestion architecture,” in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2013, pp. 1147–1158.

[6] W. Liang, H. Wang, and M. E. Orlowska, “Range queries in dynamic OLAP data cubes,” Data Knowl. Eng., vol. 34, no. 1, pp. 21–38, Jul. 2000.