Pattern Mining in Meeting using FP-Growth Algorithm

Bincy K Sunny
PG Student
Department of Computer Science & Engineering
Mount Zion College of Engineering Kadamanitta, Kerala, India

Rincy Roy Oommen
Assistant Professor
Department of Computer Science & Engineering
Mount Zion College of Engineering Kadamanitta, Kerala, India

Abstract

Human interaction in meeting schedule provides a way for analyzing the outcome of the meeting. The proposed system develops a meeting application for a large firm company. It allow users scattered among various places to post comments for the topic given by the admin. The main thing is that each user must login at the exact time of the meeting section. Otherwise they will not be permitted to comment on the topic. Each of them can express their own ideas, views directly on the meeting. At the end of the meeting admin will check whether positive or negative comments are supported by most of the users. Based on this admin can conclude the topic by listing reasons and distributing it to all the participants in that meeting. Interaction types used in this application are proposal, positive comment and negative comment. FP-Growth algorithm is used for finding frequent patterns.

Keywords: Human interaction, Sentiment analysis, N-gram algorithm, FP-Growth

I. INTRODUCTION

Human interaction is an important characteristics of any group activities. It includes group discussion, meeting, panel discussion, debates etc. In meetings status is shared, new decisions are made, alternatives are discussed. There by generate new ideas and made strategies. Mining human interactions is important for accessing and understanding meeting content. Human interactions in a meeting discussion are defined as public actions or communicative actions taken by meeting participants corresponding to the current topic. Various interactions imply different user roles, attitudes, and intentions about a topic during a discussion. The definition of interaction types naturally varies according to usage. In this paper, mainly focus on the task-oriented interactions that address task-related aspect. Here create a set of interaction types based on a standard utterance-unit tagging scheme: propose, acknowledgement, positive comment, negative comment. Frequent pattern can be mined by using FP-Growth algorithm.

A. Scope of the Work

The human interaction in meeting schedule is used to ensure that equal opportunity for the members to share ideas, knowledge, and views. Evaluate the performance of members in a meeting and thus helps in finding conclusion based on the interactions at the end of meeting.

II. RELATED WORKS

Zhiwen Yu, Zhiyong Yu, Xingshe Zhou, Christian Becker and Yuichi Nakamura et.al. (2012) [2]. Human interactions in a meeting discussion are defined as social behaviours or communicative actions taken by meeting participants corresponding to the current topic. Various interactions imply different user roles, attitudes, and intentions about a topic during a discussion. The definition of interaction types naturally varies according to usage. In this paper, mainly focus on the task-oriented interactions that address task-related aspect. The other communicative actions that concern the meeting and the group itself are not included. For generalizability, we create a set of interaction types based on a standard utterance-unit tagging scheme: propose, comment, acknowledgement, requestInfo, askOpinion, posOpinion, and negOpinion.

Theresa Wilson, Janyce Wiebe, Paul Hoffmann et.al (2006)[3]. Sentiment analysis is the task of identifying positive and negative opinions, emotions, and evaluations. Most work on sentiment analysis has been done at the document level, for example distinguishing positive from negative reviews. However, tasks such as multi-perspective question answering and summarization, opinion-oriented information extraction, and mining product reviews require sentence-level or even phrase-level sentiment analysis. This paper presents a new approach to phrase-level sentiment analysis that first determines whether an expression is neutral or polar and then disambiguates the polarity of the polar expressions. With this approach, the system is able to automatically identify the contextual polarity for a large subset of sentiment expressions, achieving results that are significantly better than baseline.
Samiksha Kankane, Vikram Garg et.al (2015) [4]. Web data mining is an emerging research area where mining data is an important task and various algorithms has been proposed in order to solve the various issues related to the web mining in existing dataset. This paper focuses the concept of data mining and FP-Growth algorithm. As for FP-Growth algorithm, the effectiveness is limited by internal memory size because mining process is on the base of large tree-form data structure. This Research work concentrates on web usage mining and in particular focuses on discovering the web usage patterns of web sites from the server log files. This paper finds the procedure to work with the proposed technique which can be possible to remove the drawback of limitation of the existed technique in the web mining area. The various web usages mining technique can further work on various scientific area, medical area and social media application to approach for the research and security related area. A detail and pattern growth technique can help in getting more data and further on using line up algorithm we can illustrate the data states presentation effectively.

III. EXISTING SYSTEM

A smarter pattern mining system for human interactions in group discussion is designed which classifies the input and extracts the frequent patterns from human interactions[1]. That can be used to ensure the performance analysis of human interactions and to make winning strategies. This system will take text file containing conversation of group discussion as an input and will generate frequent patterns of interaction as output. It will then parse the input interaction, identify the roles (persons) involved in the discussion and assign the class label to persons’ statements. System will generate interaction tree and subtrees of it.

A. Limitations of Existing System

1) Consisting of less members
2) Consumes a lot of time
3) It doesn’t make final output

IV. PROPOSED SYSTEM

Developing a meeting application for a large firm company. It consisting of member in that firm. Each of them can express their own ideas, views directly in meeting. Admin can get output at the end of discussion. Interaction types used in this application are proposal, positive comment and negative comment. Proposal means admin can proposes a topic, user can comment positively or negatively based on that proposal. FP Growth algorithm is used for creating tree based mining and finding frequent pattern. Generally meeting means an occasion when people come together to discuss or decide something. Meeting member to get together and discuss about a problem or issue or a special matter. It is also defined as a situation in which two or more people meet together in order to take decision. It is an effective and important tool in the communication process.

A. Advantages of Proposed System

– Save time: Individual can meet a number of persons at a period interactively, a meeting can save time.
– Cope with information explosion: New knowledge and new principles are coming thick and fast. Meetings enable us to cope with this situation.
– Feeling of being consulted: Members get the feeling that they have been consulted and this is useful in getting their intelligent and willing cooperation.
– Democratic functioning: Democracy aims at achieving all people’s welfare by all people’s involvement. This is possible through meetings.
– Idea development: Ideas are systematically exchanged, analyzed and improved by a group.
– Bolder decisions: Collectively we can take more bold decisions because of united strength.
– Preventing mistakes: A meeting helps to avoid mistakes by a collective and many-angled focus on issues.

V. SYSTEM ARCHITECTURE

The user login to the system and comment on the specified topic given by the admin. Data pre-processing can be applied on the comment. Data pre-processing consists of three steps tokenization, stopword elimination and stemming. Tokenization is used find the margins. The output of this is applied to the next step called stopword elimination. It is used to remove the repeated words such as (the, a, an, are, as, etc…). In computing, stop words are words which are filtered out before or after processing of natural language data (text). Though stop words usually refer to the most common words in a language, there is no single universal list of stop words used by all natural language processing tools, and indeed not all tools even use such a list. Some tools specifically avoid removing these stop words to support phrase search. Stemming is used to reduce the words into their base roots. It is used as an approximate method for grouping words with a similar basic meaning together. Then categorise the statement. Here statement can be splits and check the sentiments of the words. Here sentiment of word can be analysed by checking the collection of positive and negative words from the database. If the word is positive, then the count value is increment by one. If the word is negative, then the count value is decrement by -1. If the word is neither positive nor negative
then it is a neutral and its count is zero. After that calculate the support count for the comments which is more supported by the user. For this N-gram algorithm is used. After that collect the stored comments from the database for getting the remarks of the meeting. To know who were actively participate in the meeting and who were proposes a lot of ideas, who were speak a lot or a little etc. For this FP-Growth algorithm is used to mine the patterns. Then analyse the patterns for getting the outcome. The output from the analysis is helps to make decision of the meeting at the end of discussion.

### Fig. 5.1: System Architecture

#### B. FP-Growth Algorithm

Procedure:

FPGrowth(DB, ξ)
Define and clear F-list : F[];
foreach Transaction Tᵢ in DB do
    foreach Item aⱼ in Tᵢ do
        F[aⱼ] ++;
    end
end
Sort F[];
Define and clear root of FP-tree : r;
foreach Transaction Tᵢ in DB do
    Make Tᵢ ordered according to F;
    Call ConstructTree(Tᵢ, r);
end
foreach Item aⱼ in I do
    Call Growth(r, aⱼ, ξ);
end
VI. CONCLUSION AND FUTURE SCOPE

The meeting application will help the large firm to take decision based on proposal. In proposed system ensure that equal opportunity for the members to share ideas, knowledge, and views. It also consumes less time. Here proposed a tree-based mining method for discovering frequent patterns of human interaction in meetings. The mining results would be useful for summarization of meetings. It can be also used for interpretation of human interaction in meetings. In this paper, mainly focus on the task-oriented interactions that address task-related aspect. Here create a set of interaction types based on a standard utterance-unit tagging scheme: propose, acknowledgement, positive comment, negative comment. Frequent pattern can be mined by using FP-Growth algorithm. In future, expecting more exploration in the field of data mining and finding more interesting patterns.

REFERENCES