Automatic Detection of Irrelevant Comments in an Electronic Meeting

Milam Aiken, Bart Garner

School of Business Administration, University of Mississippi, University, MS 38677

ABSTRACT

Groups exchanging ideas in electronic meetings often generate irrelevant or off-topic comments that can detract from the conversation. Here, we describe a system that seeks to identify this immaterial text using previously identified keywords. Results of an experiment with the system show that group members believe meetings do have irrelevant comments that waste time, but participants often enjoy them. The system achieved an F measure of 42.3% for recall and precision, and further research is necessary to determine if this is sufficient or what can be done to improve this score.

Indexing terms/Keywords

Electronic meetings, Group Support Systems, Relevancy

Academic Discipline And Sub-Disciplines

Computer Science (text analysis, natural language processing)

TYPE (METHOD/APPROACH)

Lab Experiment

INTRODUCTION

Many discussions include off-topic comments as group members attempt to joke or kill time. However amusing they may be to some, these comments can be a distraction to others, especially in a serious discussion [8]. In addition, irrelevant comments could contain flames or personally abusive text [9] [15].

The purpose of this paper is to investigate a new electronic meeting system that seeks to automatically detect and exclude comments considered to be non-topical and not adding value. First, we discuss prior studies of group comments and then we describe the software. The paper then describes two experiments testing the system and presents a discussion of the results.

LITERATURE REVIEW

Several experiments using electronic meeting systems have noted that when anonymity is provided, groups often become disinhibited [10] [19] and sometimes generate irrelevant text, particularly towards the end of the session [20]. For example, in one study, groups discussing the parking problem on campus generated 12.5% irrelevant comments, while groups discussing who should be president generated 21.9%. [7]. In another study of 14 meetings with different topics [3], the percentage of relevant comments ranged from 54% to 89%, and thus, irrelevant comments ranged from 11% to 46%, as shown in Table 1:

Table 1. Relevant comments in meetings

	Meeting Topic	Relevant Comments	Unique Comments	% Relevant	% Unique
1	What makes for success in our culture?	34	14	67	27
2	How can we improve the parking problem on campus?	46	11	84	20
3	How can the spread of AIDS be reduced?	36	16	63	28
4	How can we encourage more tourists to visit the city?	31	17	74	40
5	How can we improve the parking problem on campus?	54	19	77	27
6	How can the spread of AIDS be reduced?	47	18	81	31
7	What makes for success in our culture?	49	10	89	18
8	Do you have some ideas about how to make a class more interesting?	35	15	73	31
9	Do you fear that Iraq is slowly becoming another Vietnam?	49	9	66	12
10	Do you fear that Iraq is slowly becoming another Vietnam?	59	23	87	34
11	Do you have some ideas about how to make a class more interesting?	26	15	68	39

I S S N 2 2 7 8 - 5 6 1 2
V o l u m e 1 2 N u m b e r 1
I n t e r n a t i o n a l J o u r n a l o f M a n a g e m e n t a n d I n f o r m a t i o n T e c h n o l o g y

12	What type of soft drink should be in the vending machines on campus?	79	19	78	19
13	What type of soft drink should be in the vending machines on campus?	85	14	73	12
14	How can we encourage more tourists to visit the city?	28	13	54	25

Group members might add irrelevant comments (e.g., "When will this meeting be over?") when they become bored, and they sometimes type short, nonsensical text (e.g., "abc") to test the software [5]. The topic and meeting techniques can be factors in the number of irrelevant comments generated [2] [6] [12], and meeting participant interest and culture are also important [4] [18].

IDENTIFYING IRRELEVANT TEXT

It is difficult to define exactly what is relevant and what is irrelevant. What might be valued by some group members might not be valued by others. Yet, text that has no clear relationship to the current topic is probably not pertinent. For example, if the topic of the meeting is "How can we improve sales next quarter?" a comment such as "What are you doing this weekend?" clearly has no place in the discussion.

One simple way of possibly identifying irrelevant text is keyword matching. That is, if a new comment has at least one word from a list of previously identified words relevant to the topic, it can be admitted. For example, in a meeting about the parking problem on campus, key words could include several nouns such as "bus," "campus," "car," and "bicycle" as well as several verbs and adjectives such as "abolish," "allowed," "commute," "difficult," and "easier." In addition, the size of the keyword list can be further reduced by "stemming" the content words [14], i.e., removing the prefix, suffix, or both so that only the root of the word remains.

From prior meetings of groups discussing the parking problem on campus, we identified nouns and verbs that could be important in future discussions of the same topic. However, if no meetings have been held on the topic before, text from Web sites related to the subject could be analyzed for relevant keywords.

When the participant submits a comment, the system parses the comment for the presence of any of the keywords. If none is present, it reports that the comment is not acceptable. For example, as shown in Figure 1, if the user types: "What do you plan to do this weekend?" the system shows the message "This comment was not considered relevant and was not recorded." The user then can revise the comment and submit it again, or think of a different idea to contribute. The participant also has the opportunity to read what others are writing.

Figure 1. Sample screens of the system in use

3124 | P a g e
M a y 2 0 1 7
w w w . c i r w o r l d . c o m
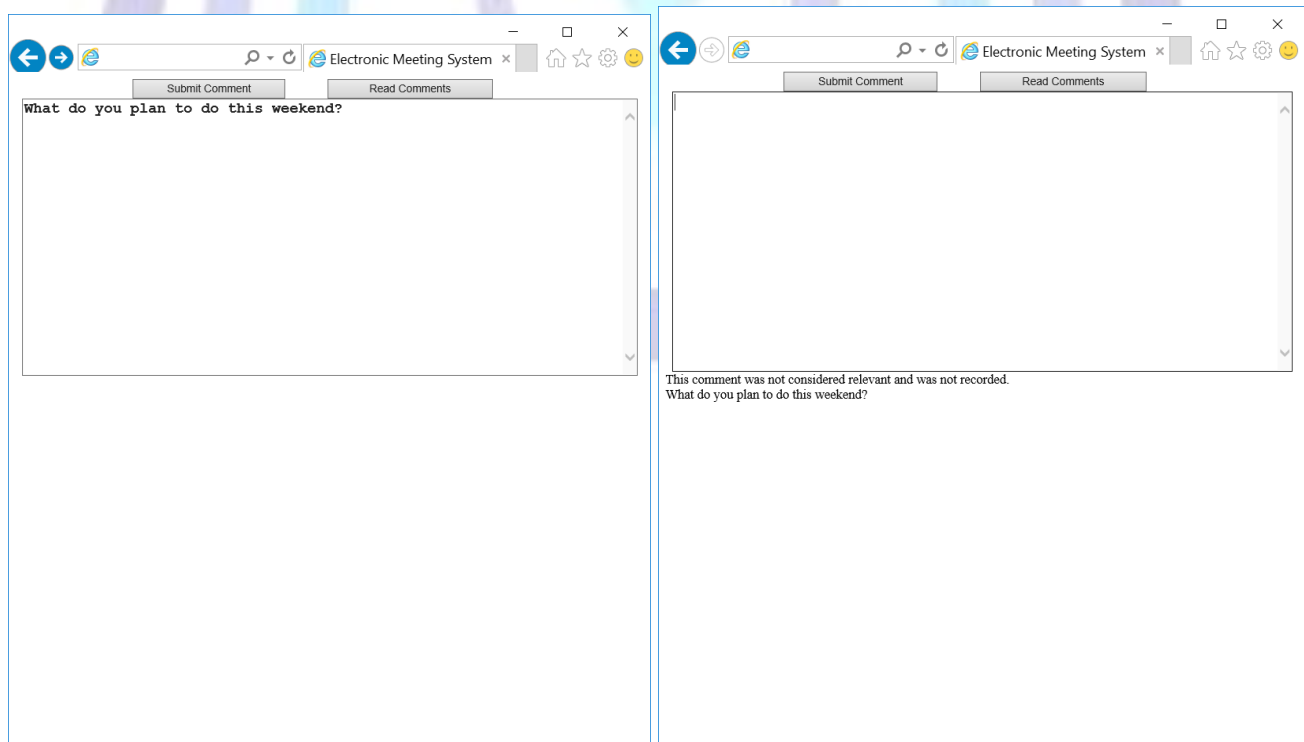
EXPERIMENTAL STUDIES

Purpose

The purpose of these two studies was to evaluate how well an electronic meeting system can detect possible irrelevant comments in a group discussion.

Study 1

In the first trial, we used 50 undergraduate business students to discuss the parking problem on campus for about 10 minutes, a topic and time limit that have been used for many experiments in the past (e.g., [16] [17] [21]). As it was a test of the system's text exclusion capability, students were encouraged to write some irrelevant comments. Afterwards, students completed a survey asking the following questions using a scale of 1 = disagree to 7 = totally agree:

Q1: I enjoy reading irrelevant comments in a serious discussion.

Q2: Reading irrelevant comments does not waste my time.

Q3: Irrelevant comments add value to the discussion.

Q4: Most discussions I have had included irrelevant comments.

A total of 94 comments were generated that were deemed topical by the system, and 10 comments were excluded from the discussion including   "hhh," "testing," "Segways," and "pizza." However, 11 comments not excluded were deemed irrelevant by a human reviewer, including "This comment is not about parking," "I heard parking lots like to get dressed up and go clubbing on the weekend," "cars and tacos," "parking is life," "I will try at least one more silly comment after this one, but I need to go to my car," "Is this an irrelevant comment?", and "you can type whatever you want, then put cars." One completely blank comment was also allowed to pass, but should have been blocked. As noted in one of the comments, several students realized that if they included certain keywords such as "car" or "parking" then their text would not be blocked by the system.

Study 2

In the second test, we used 30 graduate business students to discuss the same topic in the same amount of time. In this session, the group facilitator encouraged participants to type more irrelevant comments as the number from the first trial was not deemed sufficient for an adequate evaluation of the software.

A total of 95 comments were accepted into the discussion by the system, and 39 comments were kept out. Examples of the latter include "Go Knicks," "warriors blew a 3-1 lead," "hey," "frank has terrible suggestions," and "there are too many squirrels." However, three excluded comments were deemed relevant by the external reviewer, including ""Tear down Tad Pad but do not touch the cemetery" and "cut back on the construction during peak hours." Of the 95 included comments, eight were determined to be irrelevant in addition to four blank comments.

Table 2. Overall summary of variables

Variable	Min	Max	Mean	Std. Dev
Overall				
1: Enjoy irrelevant comments	1	7	4.15	1.84
2: Irrelevant comments not a waste of time	1	7	3.44*	1.91
3: Irrelevant comments add value	1	7	3.11**	1.53
4: Most discussions have irrelevant comments	1	7	5.20**	1.43
Group 1				
1: Enjoy irrelevant comments	1	7	4.22	1.81
2: Irrelevant comments not a waste of time	1	7	3.78	1.92
3: Irrelevant comments add value	1	7	3.26**	1.55
4: Most discussions have irrelevant comments	1	7	5.06**	1.57
Group 2				
1: Enjoy irrelevant comments	1	7	4.03	1.92
2: Irrelevant comments not a waste of time	1	7	2.87**	1.77
3: Irrelevant comments add value	1	7	2.87**	1.48
4: Most discussions have irrelevant comments	3	7	5.43**	1.14

* Significantly different from the neutral measure of 4 at α = 0.05

** Significantly different from the neutral measure of 4 at α = 0.01

Analysis

As shown in Table 2, students tended to think irrelevant comments are a waste of time and do not add value. Moreover, they believed that most discussions have irrelevant comments. However, they were neutral about how much they enjoy the non-topical comments. Although they can be a waste of time, they are often amusing.

Those who enjoyed irrelevant comments (Q1) thought they were not a waste of time (Q2) (R = 0.53, p < 0.01) and added value (Q3) (R = 0.52, p < 0.01). In addition, Q2 and Q3 were significantly correlated (R = 0.51, p < 0.01). None of the variables was significantly correlated with Q4 (Most discussions have irrelevant comments).

As described above, 9.6% of the total comments written in study 1 were excluded by the system and none of these was considered relevant. However, 11.7% of the comments accepted were determined to be irrelevant by the reviewer. That is, the precision (how useful the search results are) was 100% and the recall (how complete the results are) was 47.6% with an overall F measure (the harmonic mean of precision and recall) of 64.4% [13].

In study 2, a larger percentage of the total comments were excluded (29.1%) because the facilitator encouraged more off-topic ideas. Of these excluded comments, 10.3% were determined to be actually relevant. Further, 12.6% of the accepted comments were determined by the reviewer to be irrelevant. However, one-third of these non-topical comments were simply blank. Thus, the software needs to be revised to exclude comments with no text. For study 2, the precision was 92.3% and the recall was 27.5% with an overall F measure of 42.3%

CONCLUSION

Summary

Group members used an electronic meeting system with automatic off-topic comment detection in an attempt to increase the relevancy of the final transcript. Students agreed that many meetings included irrelevant comments that are a waste of time, but many enjoy them anyway because of the humorous content. The system was able to exclude much immaterial text, but did not detect all.

Limitations

The first limitation in the studies is the crude keyword-matching algorithm used by the software. Many irrelevant comments could have one of the keywords and get through, but a relevant comment without a previously identified keyword could be barred. In fact, several students noticed this and added keywords to off-topic comments just to get them accepted by the system.

Second, the studies focused on two ad-hoc, informal meetings in which irrelevant comments probably do not matter much. In a more serious conversation, irrelevant text would be less acceptable.

Future Research

It is not clear what the acceptable rate of exclusion should be, in addition to precision and recall. If more comments are not accepted by the software, there is a risk of more relevant comments also not being allowed. Future studies should investigate this trade-off. In addition, the analysis of textual content possibly could be improved with statistical learning techniques [1] [11] or context-aware semantics [22] rather than through a simplistic use of keywords.

REFERENCES

[1] Adevaa, J., Atxaa, J., Carrillob, M., and Zengotitabengoab, E. (2014). Automatic text classification to support systematic reviews in medicine. Expert Systems with Applications, 41(4), 1498-1508.

[2] Aiken, M. (2002). Topic effects on electronic meeting comments. Academy of Information and Management Sciences, 5(1/2), 115-126.

[3] Aiken, M., Gu, L., and Wang, J. (2009). Electronic meeting topic effects. Best Practices and Conceptual Innovations in Information Resources Management: Utilizing Technologies to Enable Global Progressions. Chapter 19, Advances in Information Resources Management, Volume 20, IGI Global Books, Hershey, PA

[4] Aiken, M., Gu, L., Wang, J., and Vanjani, M. (2008). Topic influences on electronic meeting relevant comments. Issues in Information Systems, 9(2), 300-304.

[5] Aiken, M., Park, M., and Garner, B. (2012). Translation of relevant and irrelevant multilingual group support system comments. International Journal of Intercultural Information Management, 3(1), 45-58.

[6] Aiken, M., Rebman, C., and Vanjani, M. (2007). Comment generation with three electronic brainwriting techniques. Journal of Management Information and Decision Sciences, 10(1), 11-30.

[7] Aiken, M. and Waller, B. (2000). Flaming among first-time group support system users. Information and Management, 37(2), 95-100.

3126 | P a g e
M a y 2 0 1 7
w w w . c i r w o r l d . c o m

[8] Aiken, M., Wang, J., Gu, L., and Paolillo, J. (2011). An exploratory study of how technology supports communication in multilingual groups. International Journal of e-Collaboration, 7(1), 17-29.

[9] Alonzo, M. and Aiken, M. (2004). Flaming in electronic communication. Decision Support Systems, 36(3), 205-213.

[10] Dennis, A. and Williams, M. (2003). Electronic brainstorming. Group creativity: Innovation through collaboration, 160-178.

[11] Fitzgerald, J., Elmore, J., Koons, H., Hiebert, E., Bowen, K., Sanford-Moore, E., and Stenner, A. (2015). Important text characteristics for early-grades text complexity. Journal of Educational Psychology, 107(1), 4-29.

[12] Gu, L., Aiken, M., and Wang, J. (2007). Topic effects on process gains and losses in electronic meetings. Information Resources Management Journal, 20(4), 1-11.

[13] Hripcsak, G. and Rothschild, A. (2005). Agreement, the f-measure, and reliability in information retrieval. Journal of the American Medical Informatics Association, 12(3), 296-298.

[14] Ismailov, A., Jalil, M., Abdullah, Z., and Rahim, N. (2016). A comparative study of stemming algorithms for use with the Uzbek language. 3rd International Conference on Computer and Information Sciences (ICCOINS), 7-12.

[15] Johnson, N., Cooper, R., and Chin, C. (2008). The effect of flaming on computer-mediated negotiations. European Journal of Information Systems, 17, 417–434.

[16] Lindblom, T., Aiken, M., and Vanjani, M. (2009). Electronic facilitation of large meetings. Communications of the IIMA, 9(3), 23-38.

[17] Park, M., Aiken, M., and Ghosh, K. (2010). A study of factors affecting electronic meeting participation. International Journal of Business and Systems Research, 4(3), 264-277.

[18] Reinig, B. and Mejias, R. (2004). The effects of national culture and anonymity on flaming and criticalness in GSS-supported discussions. Small Group Research, 35(6), 698-723.

[19] Suler, J. (2004). The online disinhibition effect. Cyberpsychology & Behavior, 7(3), 321-326.

[20] Underhill, C. and Olmsted, M. (2003). An experimental comparison of computer-mediated and face-to-face focus groups. Social Science Computer Review, 21(4), 506-512.

[21] Wong, Z. and Aiken, M. (2006). The effects of time on computer-mediated communication group meetings: An exploratory study using an evaluation task. International Journal of Information Systems and Change Management, 1(2), 138-158.

[22] Xie, S., Wang, J., Amin, M., Yan, B., Bhasin, A., Yu, C., and Philip, S. (2015). A context-aware approach to detection of short irrelevant texts. In Data Science and Advanced Analytics (DSAA), 36678 2015. IEEE International Conference on (pp. 1-10). IEEE

Authors' biographies

Dr. Milam Aiken is a Professor and Chair of Management Information Systems in the School of Business Administration at the University of Mississippi. His research interests include neural networks and electronic meeting systems.

Dr. Bart Garner is a Clinical Assistant Professor of Management Information Systems in the School of Business Administration at the University of Mississippi. His research interests include multilingual meeting systems, artificial intelligence, and business ethics.

3127 | P a g e
M a y 2 0 1 7
w w w . c i r w o r l d . c o m