Chapter 4

Results and Discussion

In order to analyze the efficiency of the system, three types of experiments are conducted. The first preliminary experiment is set to investigate the effects of our prototype system on group meeting via electronic meeting system. Since the results from these experiments prove that our system can help the discussants see the decision factors of the problem more clearly and help reduce the distracting issues in the meeting, we then continue our investigation to study the performance of our prototype system. The second experiment is conducted to evaluate the performance of the information extraction agent and the third experiment is performed to evaluate the performance of the intelligent topic detection agent. From the experimental results, our proposed system outperforms on its duties more than 90% accuracy. The detail of each experiment and its results are clarified in the following section respectively.

4.1 The Experiment of the Effects of Our System on Group Meeting

We postulate that the proposed discussion support system can increase decision-making efficiency via electronic meeting. Thus, in order to analyze the efficiency of the system, we considered two types of discussion on electronic meeting. The first is a discussion supported by the proposed system whereas the second is a conventional discussion without our system. Both discussions are related to the same topics. The hypotheses of our experiment focus on measuring the system efficiency in two aspects: user's satisfaction and efficiency of the discussion. For the hypotheses on user's satisfaction, we have considered two types of satisfaction, according to George et al. (1990), satisfaction with the process and satisfaction with the outcomes.

The hypotheses on user's satisfaction with process include satisfaction on the discussion, the meeting time, and the information for making the decision as follows:

H1: The proposed system can encourage more discussion participation than the meeting without the proposed system.

H2: The proposed system helps the discussants reach the conclusion faster than the meeting without the proposed system.

H3: The proposed system helps the discussants obtain necessary information to support decision making more than the meeting without the proposed system.

The hypotheses on the satisfaction with the outcomes include satisfaction on the decision outcome and confidence in the decision outcome as follows:

H4: The discussant's satisfaction on the decision outcome is higher on the meeting with the proposed system than the meeting without the proposed system.

H5: The discussant's confidence in the decision outcome is higher on the meeting with the proposed system than the meeting without the proposed system.

The hypotheses on the efficiency of the discussion include the number of ideas generated by the participants and the completeness of the structure of the key decision factors. The number of ideas generated by group members has been frequently used in GSS empirical research (Davison, 1996). This measurement presents how meeting members participate in the meeting and how much they contribute to the discussion. Thus, for our experiments, we measure the number of ideas generated during the discussion from the data that we saved from the online meeting system and set the hypothesis as follows:

H6: The number of ideas generated during the discussion in the meeting with the proposed system is higher than the meeting without the proposed system.

4.1.1 Experimental Design

To investigate the concept of our system, an empirical study was conducted to study the efficiency of using computer system as facilitator in order to support group discussion via electronic meeting system. The discussants in the meeting were recruited from the combination of the university students and employees. A total number of one hundred and twenty participants were grouped with three members in each group and were randomly assigned to two sets of groups: a set of the experimental groups (with the proposed support system) and a set of control groups (without the proposed support system). We set eight topics for discussion in the meeting. Each topic concerns about the common issues in every day life such as cellular phone, a vacation plan, education, etc. Each topic was set to discuss in twenty minutes. After the discussion, the participants were requested to fill out the questionnaire (see in **Appendix D**) to measure the satisfaction with the process and the satisfaction with the outcomes according to our first five hypotheses. The questions ask the participants to specify their level of satisfaction on each criterion from the highest level of satisfaction (level 5) to the lowest level of satisfaction (level 1). The results of the experiments for the satisfaction with the process and the satisfaction with the outcomes are obtained through the data from the questionnaires, while the results of the experiments for the efficiency of the discussion are achieved from the analysis of the discussion data that we saved from the online discussion system.

4.1.2 Satisfaction With the Process

The t-test hypothesis testing has been applied to the data from the questionnaires to examine our hypotheses. The significant level (or α) of our hypothesis testing is set to 0.05. The results of the experiments on user's satisfaction with the process are illustrated in the Table 4.1.

Table 4.1 The Results of The Experiments on User's Satisfaction with the Process

Question for Testing Process Satisfaction on	Mean of Satisfaction Level with the Proposed System	Mean of Satisfaction Level without the Proposed System	T-Stat	P-value	Hypothesis Result
H1	3.22	2.61	2.5	0.01	Accept
H2	3.22	2.83	2.12	0.02	Accept
H3	3.52	3.16	1.76	0.04	Accept

From the results of the experiments, we can see that all three hypotheses are accepted. The result of the first hypothesis (H1) shows that the mean value of the meeting with our system is statistically higher than the conventional meeting with p-value 0.01, which is less than the significant level of 0.05. This shows that the participants felt that our proposed discussion support system could encourage more discussion participation than the meeting without the proposed system. The increase in the level of discussion participation is important to the new idea generation and affects the judgment on the decision alternatives. From the analysis of the result, the discussion-supported system reveals that the ideas generated by the discussants are more prominent and the system reduced irrelevant information during the meeting.

Along with the increase in the discussion participation, the third hypothesis (H3) is also accepted with the p-value of 0.04. This shows that the participants also felt that they obtained necessary information to support their decision making more than the meeting without the proposed system. Moreover, it is not only that the proposed system can increase the discussion participation and help the discussants obtain more necessary information to support decision making, the second hypothesis (H2) also shows that the proposed system can help the discussants reach the conclusion faster than the meeting without the proposed system with p-vale of 0.02. Thus, the proposed system can be beneficial for the meeting, which has time constraint.

4.1.3 Satisfaction with the Decision Outcome

The second type of the questions investigates the satisfaction on the decision outcome of the meeting. The participants were asked about their satisfaction on the decision outcome and their confidence on the decision outcome according to our fourth and fifth hypotheses (H4 and H5). The results of the satisfaction of the decision outcome are illustrated in the Table 4.2.

Table 4.2 The Results of the Experiments on the Satisfaction of the Decision Outcome

Question for Testing Outcome Satisfaction on	Mean of Satisfaction Level with the Proposed System	Mean of Satisfaction Level without the Proposed System	T-Stat	P-value	Hypothesis Result
H4	3.44	3.33	0.35	0.36	Reject
H5	3.47	3.16	1.09	0.14	Reject

From the results of the experiments on the satisfaction with the decision outcomes, the p-values of both questions are greater than the significant level of 0.05. This shows that our discussion-supported system does not increase the discussant's satisfaction with the decision outcomes. Since our discussion support system is applied to meeting system for constructing the information into the DCSML structure, this structure emphasizes on revealing the decision hierarchy of the information from the meeting, but does not make a decision for the group. The DCSML structure is applied to improve the manner in which a group makes decisions and handling the process of meeting, thus the system have clearer effects on group decision process than decision outcomes.

4.1.4 The Analysis of The Experimental Result for Pre-experiment

The results of the experiments for the sixth hypotheses (H6) were analyzed from the discussion data that we saved from the online meeting system. From the twenty minutes discussion in each meeting, the number of ideas generated for the meeting with the proposed system is very high. The results from the analysis of discussion data saved from the online meeting system are illustrated in the Table 4.3.

Table 4.3 The Results from the Analysis of Discussion Data

	% Increase in Number of Ideas Generated	% Increase in the Key Decision Factors in DCSML Structure	% Decrease in Distracting Issues
Effects of the Proposed System on the Meeting	7.41	10.02	22.93

The results show that the proposed discussion support system can increase the total number of ideas generated in the meeting by 7.41%. This information supports our sixth hypothesis (H6). The contribution of the ideas during the meeting is an important factor for the efficiency of the discussion. With our proposed system, the discussants more participate (from H1) and more idea generation than the conventional meeting. Moreover, our system can increase the amount of necessary information that can support decision making. This can be measured through the available key decision factors in the DCSML structure. From the table, we can see that the number of key decision factors in the DCSML structure increases by 10.02% with the availability of our system during the meeting. This result infers that our system can make the structure of the problem apparent and easy to understand. Furthermore, the number of the distracting issues can be reduced by 22.93% with the support of our system. This result illustrates that our system can decrease the level of distraction during the meeting and help the discussants focus more on the important issues.

The following example illustrates the comparison of the level of the completeness of the key decision factors in the DCSML structure between the meeting with our proposed system and the conventional meeting. The topic is about the choosing the type of computer. This topic has 11 criteria and 3 alternatives generated in both types of meeting. Some part of the DCSML structure of those two meetings is shown in the Figure 4.1a and 4.1b.

```
<DISCUSSANT> Tum
   <CRITERION> Cost
    <WEIGHT>
      <CRITERION> Screen </CRITERION>
    </WEIGHT>
    <PRIORITY>
      <ALTERNATIVE Value="Highest"> Desktop
      </ALTERNATIVE>

     <ALTERNATIVE Value="Lowest"> Notebook
     </ALTERNATIVE>
     <ALTERNATIVE Value="High"> Palm
     </ALTERNATIVE>
   </PRIORITY>
  </CRITERION>
  <CRITERION> Screen
    <WEIGHT>
    </WEIGHT>
    <PRIORITY>
      <ALTERNATIVE Value="Highest"> Desktop
      </ALTERNATIVE>
      <ALTERNATIVE Value="Low"> Notebook
      </ALTERNATIVE>
      <ALTERNATIVE Value="Lowest"> Palm
      </ALTERNATIVE>
  </CRITERION>
</DISCUSSANT>
```

DCSML Structure with Proposed System

Figure 4.1a DCSML of Our Proposed System

```
<DISCUSSANT> Kong
   <CRITERION> Cost
    <WEIGHT>
      <CRITERION> Screen </CRITERION>
    </WEIGHT>
    <PRIORITY>
      <ALTERNATIVE> Desktop
     </ALTERNATIVE>

     <ALTERNATIVE Value="Highest"> Notebook
     </ALTERNATIVE>
     <ALTERNATIVE Value="Lowest"> Palm
     </ALTERNATIVE>
   </PRIORITY>
  </CRITERION>
  <CRITERION> Screen
    <WEIGHT>
    </WEIGHT>
    <PRIORITY>
      <ALTERNATIVE> Desktop </ALTERNATIVE>
      <ALTERNATIVE> Notebook </ALTERNATIVE>
      <ALTERNATIVE> Palm </ALTERNATIVE>
    </PRIORITY>
  </CRITERION>
</DISCUSSANT>
```

DCSML Structure without Proposed System

Figure 4.1b DCSML of Conventional System

From the comparison of the completeness of the structure in the above two meetings, the structure of meeting with the proposed system is more complete than the conventional meeting, as seen from to the lack of value attribute in the alternative element on each criterion. From the comparison in the percentage of the completeness, we found that the structure of the meeting with the proposed system has the average completeness of 78% when compared with 42% of the conventional meeting structure. This result infers that our system can help the discussants see the decision factors of the problem more clearly. For the overall comparison, our system can increase the completeness of the DCSML structure about 2 times of the completeness in the conventional meeting.

4.2 The Experiment of the Performance of Information Extraction Agent

To investigate the performance of our proposed information extraction agent, we used a total data set of 2368 sentences from the information technology students meeting (1600 sentences) and Thai online discussion webboard (768 sentences). In this experiment, the information extraction with the model of unigram and bigram are used to compare the result of extraction. For bigram model, we compare the HMM with Viterbi algorithm and HMM with Forward algorithm. We use three measures: recall, the percentage of the key decision factors that were correctly extracted; precision, the percentage of the key decision factors, which are correct; and accuracy, which is the success rate of extraction.

Table 4.4 and 4.5 show the results of the evaluation on extracting the key decision factors from the meeting messages. Table 4.4 shows the precision-recall of the extraction with unigram and bigram model on criteria (Cri.) and decision alternatives (Alt.).

Table 4.4 The Recall and Precision

Combine	True Words	Recognize Words			Recall (%)	Precision (%)
		Cri	Alt	Other		
Unigram						
Cri	848	790	13	45	93.16	26.20
Alt	1,139	148	939	52	82.44	77.09
Other	17,110	2,077	266	14,767		
HMM Bigram Viterbi						
Cri	848	689	9	150	81.25	83.11
Alt	1,139	9	1039	91	91.22	90.19
Other	17,110	131	104	16,875		
HMM Bigram Forward						
Cri	848	670	5	173	79.00	83.64
Alt	1,139	8	1007	124	88.41	92.64
Other	17,110	123	75	16,912		

The result presented here shows that both of the bigram models can get higher precision than the unigram model. With the bigram model, the extraction with Viterbi algorithm has a higher recall for extracting criterion and alternative but lower precision for extracting the alternative when compared with the extraction with Forward algorithm. The averages of recall and precision for the HMM with the Viterbi algorithm were 86.5% and 86%, respectively.

Next evaluation of our proposed system is the accuracy. Table 4.5 shows the accuracy of the information extraction with those three models: unigram, HMM with Viterbi algorithm, and HMM with Forward algorithm. In this table, the data set is from meeting messages only, webboard only and the combination between meeting messages and webboard. The output accuracy from extracting the key decision factors is illustrated in the following table.

Table 4.5 The Accuracy of Extraction.

	Accuracy (%)		
	Messages	Webboard	Combine
Unigram	84.45	90.15	86.38
HMM Bigram Viterbi	**97.26**	**98.53**	**97.41**
HMM Bigram Forward	97.04	98.51	97.33

From Table 4.5, the HMM with Viterbi algorithm outperforms other algorithms. It was able to extract about 97.26% of the key decision factors that being discussed in the electronic meeting system and 98.53% of the key decision factors in webboard. For the combination data, it can get 97.41% of accuracy on extracting the key decision factors. On average, the HMM with Viterbi algorithm can get 97.7% accuracy of the key decision factors extraction from meeting messages. This result shows that our system was able to

trace the messages and extract the key decision factors for supporting the meeting participants.

4.3 The Experiment of the Performance of Intelligent Topic Detection Agent

During meeting, group members frequently distract from the purpose of the meeting. This can cause group-meeting performance to be inefficient and non-productivity. Therefore, intelligent topic detection agent (ITDA) is proposed to identify possible group distraction by detecting topic changing during meeting. If topic-changing occurs, a sign of possible distraction is emerged and our system will send the messages to notify the group members about such possible distraction. Then, group member can go back to discuss on the group discussion topic and complete group meeting within time limitation.

To make our system has a competence within its conditions and its aims, ITDA performance on detecting topic changing in the discussion is investigated. To evaluate the performance, 6390 sentences in Thai language from nine topics are used to be our training and testing data. Two test sets of data are prepared. The first set consists of a number of textual discussion utterances from two different topics. Second set consists of a number of textual discussion utterances from two similar topics. In each set, the discussion utterances from main topic were automatic shuffled with another topic (presume as distracting topic) under a ratio of 5:1 (5 sentences from main topic mix with 1 sentence from distracting topic), e.g., 5:1, 10:2, 15:3 etc. The total data set in each test set are a hundred set. The example of testing data in this experiment is illustrated in Figure 4.2.

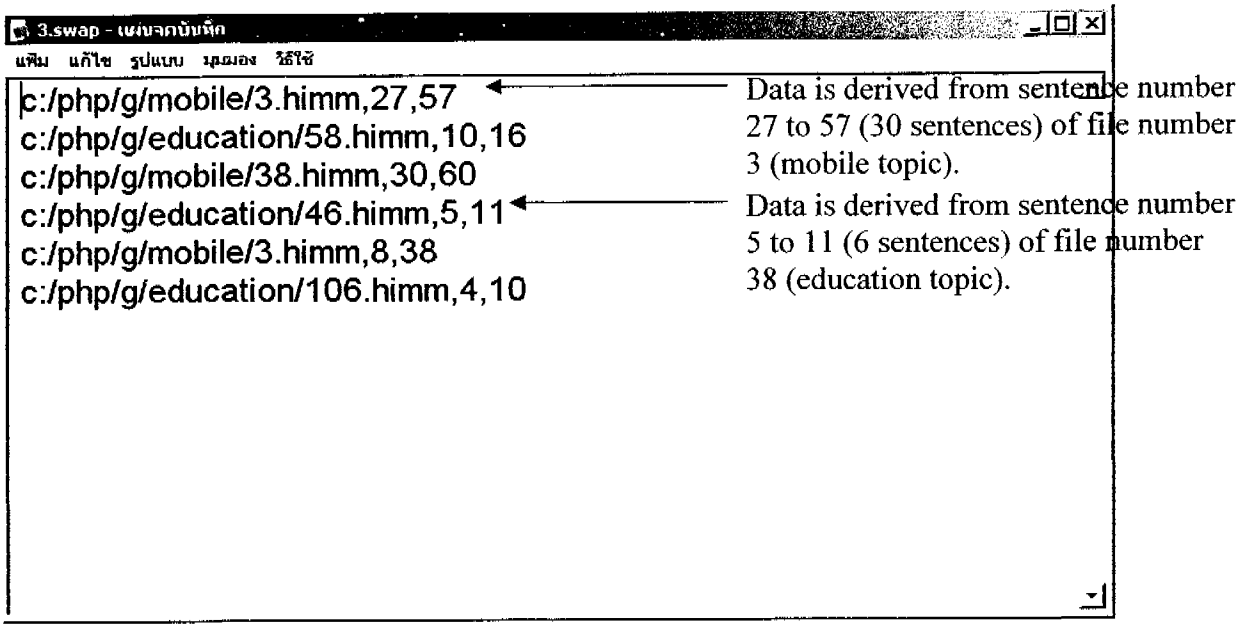

Figure 4.2 The Example of Testing Data.

In this study, we both investigate the suitable sentence blocks that are enough to detect topic accurately and investigate topic-changing performance. To determine the suitable size of sentence blocks for our proposed system, we investigate the performance of the systems for the number of blocks of sentences (VMA) from one to ten sentences per block (VMA=1 to VMA=10). The workflow of the topic-changing detection during the meeting is shown in Figure 4.3.

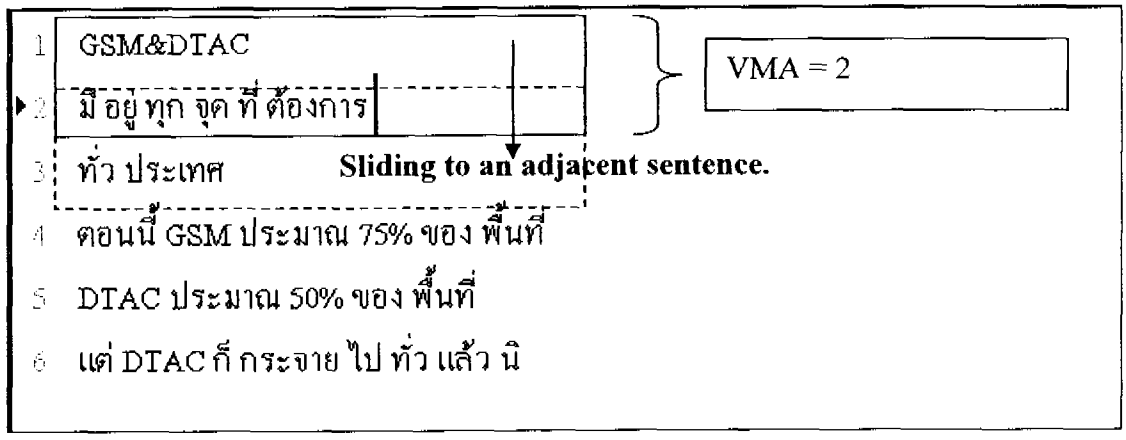

Figure 4.3 Workflow of the Sliding Sentence Block of VMA=2

From Figure 4.3, the VMA initially starts from the first sentence and waits for incoming discussion sentence through the number of VMA (from example, it is equal to 2 sentences). If the number of the discussion sentences reaches the required level, our system will start detecting the topic of an input sentences. Then, the VMA will slide to the next incoming sentence and keep detecting topic repeatedly until the end of the discussion. We use three measures to evaluate the proposed system, which are, (1) recall-the percentage of the topics that were correctly detected; (2) precision-the percentage of the topics, which are correct; and (3) accuracy-which is the success rate of detection.

Results from the experiment on topic-changing detection of different topics (the first test set) in recall and precision followed the sizes of VMA (from VMA=1 to VMA=10) against the number of the sentences from the distracting topic (1 to 6 sentences) that shuffled to the main topic are shown in Table 4.6.

Table 4.6 Recall and Precision of Experiments with Different Topics

	RECALL and PRECISION											
	The Number of Distracting Sentences											
	1sentence		2sentences		3sentences		4sentences		5sentences		6sentences	
	recall	precision	recall	precision	recall	precision	recall	precision	recall	precision	recall	precision
VMA1	98.15	97.30	98.92	98.25	99.36	98.83	99.33	98.58	99.24	99.78	99.54	99.63
VMA2	61.88	62.75	76.80	74.53	80.87	83.90	83.46	86.66	88.19	89.30	93.30	91.93
VMA3	52.05	46.60	60.21	57.80	66.54	70.78	69.63	76.11	76.32	77.33	85.71	82.89
VMA4	49.78	42.91	51.95	47.89	53.79	58.31	58.96	66.43	64.56	67.60	77.12	73.64
VMA5	49.16	40.38	49.03	42.46	46.59	49.10	50.75	57.75	54.96	57.58	70.28	65.54
VMA6	47.41	37.62	48.26	40.08	44.17	45.14	45.29	51.79	47.81	50.00	63.87	57.88
VMA7	47.90	36.81	49.84	39.44	43.24	42.18	42.25	47.37	42.70	43.97	58.13	50.00
VMA8	47.09	34.63	49.40	37.46	43.20	40.96	40.85	45.17	38.03	38.53	52.64	43.04
VMA9	46.44	32.60	48.08	34.77	43.55	40.41	39.42	42.52	35.57	35.51	47.70	37.96
VMA10	46.08	31.45	47.27	33.17	43.88	39.81	39.26	41.13	33.51	33.18	43.39	33.28
AVG	54.59	46.31	57.98	50.58	56.52	56.94	56.92	61.35	58.09	59.28	69.17	63.58

Table 4.6 illustrates that recall and precision of VMA = 1 has the highest percentage on every number of the distracting sentences. The averaged accuracy of our system on topic distraction is 99% recall and 98.73% precision compared with 58.88% recall and 56.34% precision for total average. Therefore, in case of dissimilar topics, our system can trace the utterances and detect topic changing that occurred in the online

discussion data when VMA equal to 1 (1 sentence per block). Since only one sentence is sufficient for topic-changing detection with our approach, it is not necessary to use much of information for topic-changing detection.

To check accuracy of our system in detecting group distraction, we study accuracy of the sizes of VMA (from VMA=1 to VMA=10) against the number of the sentences from the distracting topic (1 to 6 sentences) that shuffled to the main topic. The results of accuracy are shown in Table 4.7.

Table 4.7 Accuracy of Experiments with Different Topics

	ACCURACY					
	The Number of Distracting Sentences					
	1sentence	2sentences	3sentences	4sentences	5sentences	6sentences
VMA1	98.63	99.10	99.50	99.40	99.77	99.76
VMA2	77.34	84.92	90.15	91.69	94.82	95.84
VMA3	70.84	74.67	82.27	84.84	89.54	91.42
VMA4	69.69	69.31	74.59	78.75	84.48	86.70
VMA5	69.31	67.66	70.14	73.83	79.83	82.83
VMA6	68.52	67.07	68.52	70.32	76.01	79.00
VMA7	68.73	67.91	68.08	68.59	73.43	75.81
VMA8	68.44	67.63	68.05	67.67	71.27	72.95
VMA9	68.22	67.07	68.22	66.94	70.05	70.57
VMA10	68.08	66.74	68.53	67.12	69.15	68.80
AVG	72.78	73.20	75.80	76.91	80.83	82.36
TOTAL/SET	180	180	180	192	180	180

From total 180 sentences on each set (except 4 distracting sentences, which has 192 sentences), the result of experiment illustrated that VMA equal to 1 has the highest accuracy with averaged accuracy of 99.36% while the averaged accuracy on every VMA is 78.96%. Beside, from the results, the number of sentence distraction in the online discussion data by can also identify by the accuracy of dissimilar topics. In every shuffling sentences, the number of distraction is identified when the value of accuracy increasing significantly over the confidence of 79 %, for example, 2 distracting sentences has 2 values of accuracy that beyond the confidence, which are 99.1% and 84.92% respectively (the color shade). Therefore, with the result from experiment, our system can detect topic changing that occurred in the online discussion data and identify the number of sentence distraction in dissimilar topics with approximately 99% accuracy.

For the second test set, we obtained the online discussion data from two similar topics. The definition of "similar topics" in our experiment is the topics that have similar keywords such as travel and transportation. To determine the performance of our proposed system in similar topics, a set of data (similar to Figure 4.2) is used to evaluate topic-changing detection with various volume of information. Number of sentences per block (VMA) is still from one to ten sentences (VMA=1 to VMA=10). Results from the experiment on topic-changing detection of similar topics (the second test set) in recall and precision followed the sizes of VMA against the number of the sentences from the distracting topic (1 to 6 sentences) that shuffled to the main topic are shown in Table 4.8.

Table 4.8. Recall and Precision of Experiments with Similar Topics

	RECALL and PRECISION											
	The Number of Distracting Sentences											
	1 sentence		2 sentences		3sentences		4 sentences		5 sentences		6 sentences	
	recall	precision	recall	precision	recall	precision	recall	precision	recall	precision	recall	precision
VMA1	70.87	98.21	74.26	96.49	70.55	96.19	60.90	91.35	60.88	93.81	64.02	93.76
VMA2	50.88	43.50	56.59	48.75	58.99	55.82	54.40	58.01	58.74	61.95	63.98	67.83
VMA3	43.01	22.66	43.85	25.12	47.63	30.74	46.40	32.57	53.58	38.79	61.16	42.68
VMA4	42.56	17.07	37.36	16.50	38.55	16.93	39.49	22.43	43.90	23.43	52.41	29.81
VMA5	35.84	10.10	35.38	11.05	27.50	8.520	29.06	13.78	34.17	13.70	45.77	20.13
VMA6	33.78	6.596	34.78	7.637	21.95	4.650	20.26	7.200	29.6	9.319	43.97	15.73
VMA7	40.98	6.596	37.33	6.698	26.00	3.350	11.86	3.220	30.47	8.000	38.29	11.61
VMA8	39.58	5.013	35.38	5.500	33.33	2.820	10.11	2.060	26.31	4.980	30.57	7.880
VMA9	43.24	4.210	37.73	4.780	42.85	3.070	8.570	1.370	29.50	4.460	28.26	5.530
VMA10	36.36	3.150	47.72	5.010	37.03	2.560	9.520	1.370	28.84	3.722	24.00	3.810
AVG	43.71	21.71	44.04	22.75	40.44	22.46	29.06	23.34	39.60	26.22	45.24	29.88

Table 4.8 illustrates that recall and precision of VMA equal to 1 have the highest percentage on every numbers of the distracting sentences on the similar topics. However, the highest averaged recall and precision of our system in detecting topic changing decreased from 99% to 66.92% recall, and 98.73% to 94.97% precision, respectively when compared with the different topics. The averaged recall and precision also decreased from 58.88% to 40.35% recall, and 56.34% to 24.40% precision, respectively when compared with the different topics. Therefore, with the experimental results, our system still can trace the utterances and detect topic changing that occurred in the online discussion data when it encounters with the similar topic with VMA equal to 1.

To check accuracy of our system in detecting group distraction on the similar topics, we study accuracy of the sizes of VMA (from VMA=1 to VMA=10) against the number of the sentences from the distracting topic (1 to 6 sentences) that shuffled to the main topic. The results of accuracy are shown in Table 4.9.

Table 4.9 Accuracy of Experiments with Similar Topics

	ACCURACY					
	The Number of Distracting Utterances					
	1sentence	2sentences	3sentences	4sentences	5sentences	6sentences
VMA1	96.29	96.47	96.59	94.63	95.34	94.58
VMA2	90.76	91.21	93.36	92.79	94.23	93.47
VMA3	89.56	88.69	91.37	91.16	93.12	92.02
VMA4	89.40	88.07	90.58	90.31	92.06	90.65
VMA5	89.06	88.24	90.08	89.48	91.43	89.61
VMA6	89.16	88.44	90.28	89.35	91.39	89.43
VMA7	89.51	88.71	90.78	89.32	91.56	89.19
VMA8	89.55	88.72	91.06	89.70	91.64	88.81
VMA9	89.66	88.88	91.12	89.95	91.82	88.96
VMA10	89.55	89.14	91.13	90.10	91.89	88.97
AVG	90.25	89.65	91.63	90.67	92.44	90.56
TOTAL	180	180	180	192	180	180

From the total of 180 sentences on each set (except the 4 distracting sentences, which has 192 sentences), the result of the experiment illustrated that VMA equal to 1 has the highest accuracy with averaged accuracy of 95.65% while the averaged accuracy on every VMA is 90.87% on the experiments with similar topics. When we compare the results with those of different topics, the averaged accuracy on every VMA increases approximately 11.91% although the averaged accuracy on VMA equal to 1 decreases approximately 3.71%. For the identification of the number of distraction sentences, the accuracy cannot identify the number of distraction sentences in the online discussion data since the diversity of the averaged accuracy between VMA is too small.

From our experiments, the results show that our system can detect topic changing of online discussion data with approximately 97% accuracy when the number of sentence per block is 1. Moreover, the preliminary study on having an intelligent agent to facilitate group meeting shows that our agent can help group members focus on group discussion topic and reduce group distraction.