

Visualizing Group Dynamics based on Multiparty Meeting Understanding

Ni Zhang¹, Tongtao Zhang¹, Indrani Bhattacharya², Heng Ji¹ and Richard J. Radke²

¹Department of Computer Science

²Department of Electrical, Computer, and Systems Engineering

Rensselaer Polytechnic Institute

{zhangn5, zhangt13, bhatti, jih}@rpi.edu rjradke@ecse.rpi.edu

Abstract

Group discussions are usually aimed at sharing opinions, reaching consensus and making good decisions based on group knowledge. During a discussion, participants might adjust their own opinions as well as tune their attitudes towards others' opinions, based on the unfolding interactions. In this paper, we demonstrate a framework to visualize such dynamics; at each instant of a conversation, the participants' opinions and potential influence on their counterparts is easily visualized. We use multi-party meeting opinion mining based on bipartite graphs to extract opinions and calculate mutual influential factors, using the Lunar Survival Task as a study case.

1 Introduction

Group meetings are pervasive in modern workplaces, consuming workers' time and energy. Reaching consensus and making decisions more efficiently are major challenges. For example, during a meeting, some of the participants might insist on their own opinions towards the discussed items or topics, while others might rapidly change opinions and attitudes. As a meeting unfolds, we can observe developing leadership characteristics among the participants; for example, some participants may speak more assertively to drive decisive conclusions and steer the meeting, while others may follow the crowd and merely deliver tiny ideas.

In order to track the dynamics that reflect the change of opinions and the procedure of decision making, we require a meeting assistant that works in real time. That is, the assistant should keep track of the agenda and discussion process as a minute or note-taker, as well as record and assess the influence and contribution of each participant.

In this paper, using the NASA Lunar Survival

Task ¹ as a study case, we present an automatic meeting assistant with the following functionalities:

- The assistant detects and extracts participants' opinions from their speech and visualizes the groups' instantaneous state (ranking of items) based on current and previous utterances.
- The assistant visualizes an influence factor for each participant using current and previous utterances in real time. Using this information, emerging leadership in the group can be visualized.

The proposed assistant begins with speech recognition output, and detects the opinions from the speakers with Natural Language Processing (NLP) tools. We propose a bipartite graph formalism to assess participants' influence.

2 System Overview

2.1 Study Case Introduction

The NASA Lunar Survival Task is a widely used group consensus exercise that helps encourage the development of communication, cooperation, and decision making skills (Hall and Watson, 1970). In small groups of 3–4, participants discuss a hypothetical survival scenario and rank the value of supplies that may aid in their survival and safe rendezvous with their mothership. Before the discussion, each participant is asked to independently rank the items. Next, the participants are asked to reach consensus on the ranking with active verbal interaction. Each member of the group must agree upon the final ranking, which acts as the group decision.

¹<https://t.co/5e56cHayji>

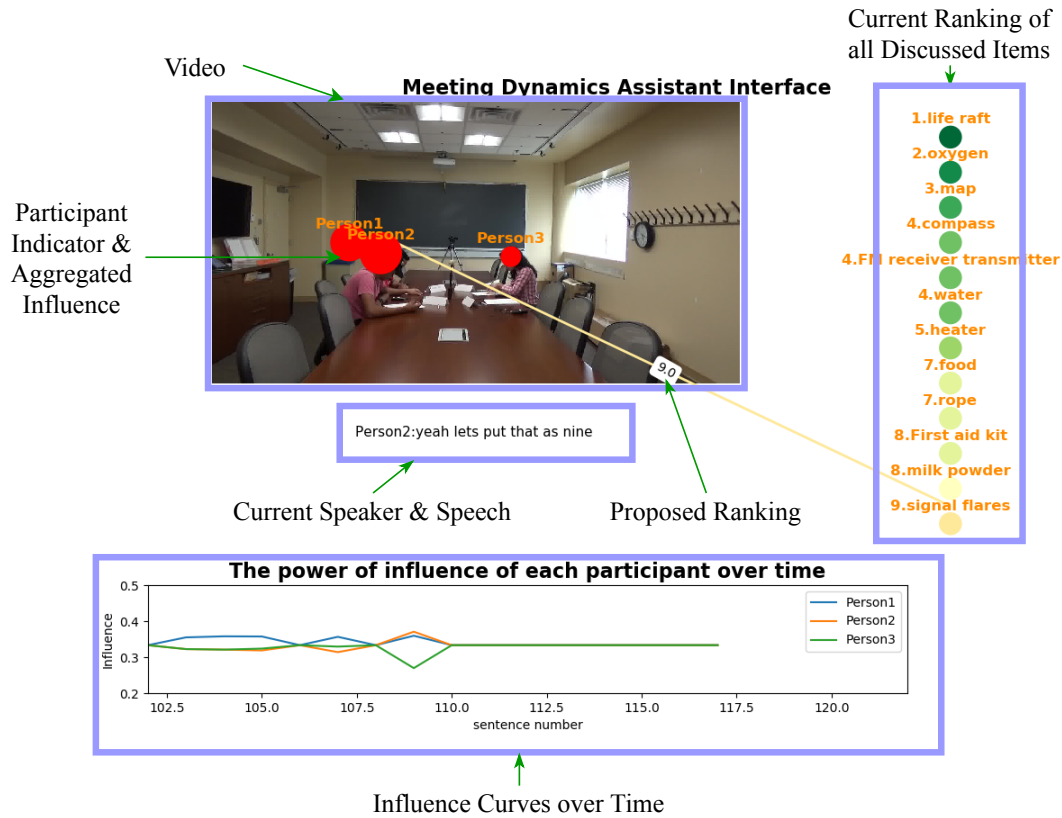


Figure 1: Main interface of our proposed meeting assistant.

2.2 Interface Details

Figure 1 illustrates the interface of the meeting assistant².

A video window shows the meeting scene. In this window, we use red circles to denote the participants. The sizes of the circles denote the aggregated influence factors of participants; the larger the circle is, the more influence (i.e., contribution to the conversation) the participant possesses.

Beneath the video window we provide the current speaker and speech. The raw speech is processed by IBM Watson’s Speech to Text System³, and we use the text output to detect discussion/focus items and extract speakers’ opinions as detailed further below.

On the right side, we place a real-time ranking list for items that have been discussed. As the discussion proceeds, the ranking list expands with newly involved items. We also illustrate the current focus item of each participant and her/his proposed rank of the focused item with a colored edge. The colors of item circles and opin-

ion edges denote different rankings; greener items have higher rankings and redder items have lower rankings.

At the bottom, we illustrate curves indicating instantaneous influence factors from each participant in real time based on the current speech.

Figure 2 illustrates a series of screenshots from our proposed meeting assistant. Three participants attend the meeting and start the discussion about several items. In Figure 2a and 2b, the meeting was in an early stage, and there is no difference among the participants in terms of influence. As the discussion continues in Figure 2c and 2d, where more items have been discussed, the curves of influence fluctuate and imply differences in activity among participants. Moreover, the rankings of discussed items are adjusted according to the speech and extracted opinions as mentioned above.

Finally, as shown in Figure 2f, from the size of the red circles representing aggregated influences and the historical record of the curves, we can conclude that Person 1 and Person 2 are contributing more in the discussion, while Person 3 is less active.

²A short video clip illustrating the meeting assistant can be viewed at https://youtu.be/3_YS0ZGQNQo.

³<https://www.ibm.com/watson/services/speech-to-text/>

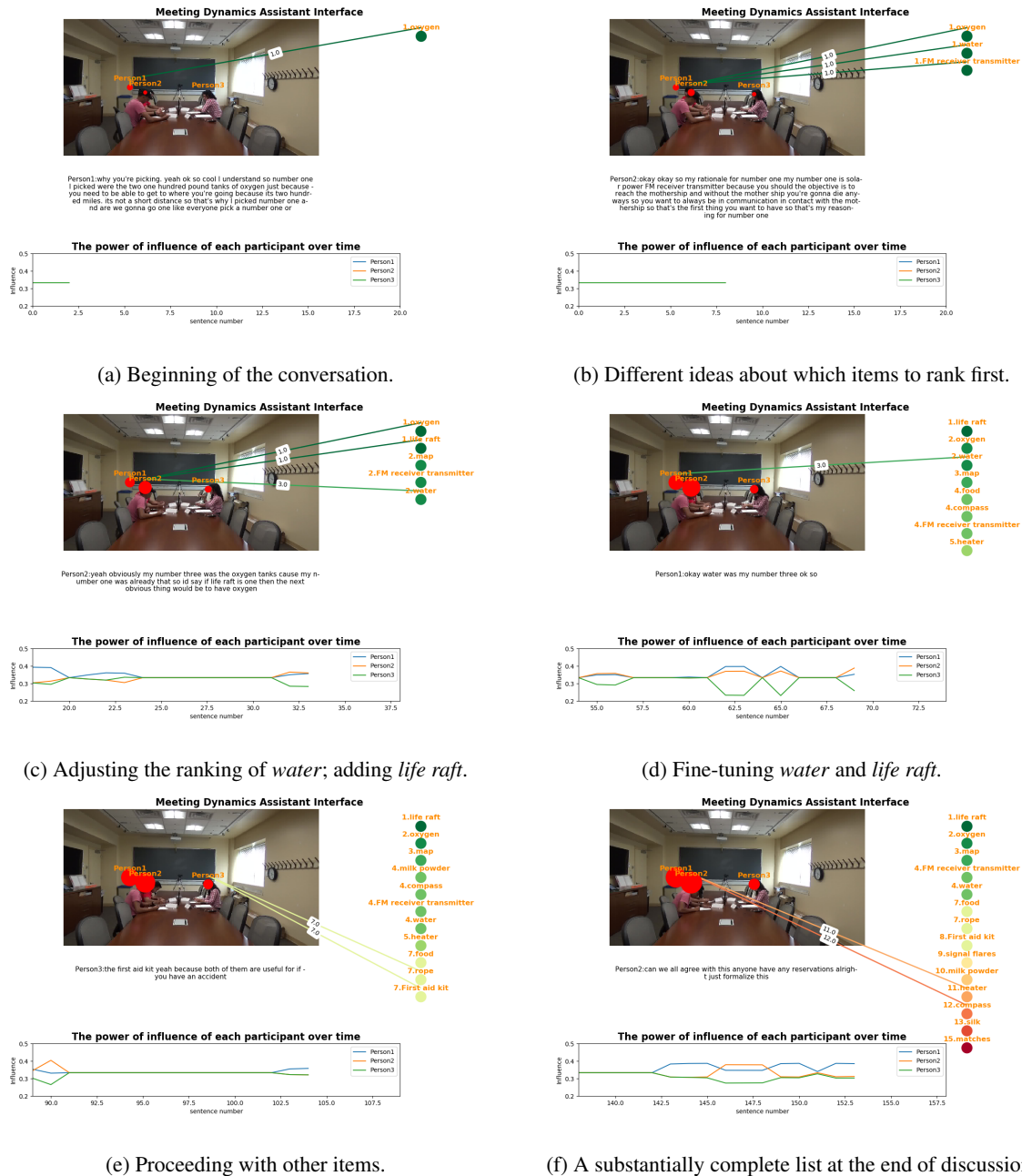


Figure 2: Screenshots of the meeting assistant at different points in time.

3 Opinion Detection and Extraction

Our system takes as input transcribed meeting speech, sentence by sentence, and outputs real-time rankings (opinion words) of the items after each participant expresses her/his thoughts.

3.1 Opinion Word Identification

In the context of the Lunar Survival Task discussion, we observed that participants express their opinions of item rankings in multiple ways, including

1. Explicitly mentioning an item with its rank-

ing(e.g., “*In my opinion, we should put water as the second most important.*”)

2. Agreement or disagreement (e.g., “*Yeah, I agree.*”)
3. Comparison of the items by relative ranking (e.g., “*Matches are less important than signal flares because they don’t work on the moon.*”)

In the first scenario of a participant proposing an item ranking, we use the Stanford CoreNLP (Manning et al., 2014) name tagger to extract the NUMBERS and ORDINALS mentioned in the discussion (Finkel et al., 2005). We also eliminated

numbers beyond 15 and numbers that are parts of pronouns such “this one”. Additionally, in this specific discussion, people use “last” or “least” to imply they are ranking the item at 15 and we also implemented this rule.

As for the second scenario, people typically express agreement/disagreement with the person who talked immediately before them (Abu-Jbara et al., 2012). For agreement, we assume the current speaker accepts the previous speaker’s stated opinion, which means we pass the weights captured for the previous person to the current speaker if we find the expression of agreement in the current sentence. We found that expressions of disagreement are not useful since people typically express their own opinion following their disagreement.

We currently do not deal with the third scenario of relative rankings, because no definitive ranking can be extracted from such statements.

3.2 Target Identification and Ranking

In this step, we identify discussed items in the discussion. As participants must have a very condensed discussion of these 15 items in a relatively short time, they usually mention the items with the exact words from the list they are given. Thus, we take the nouns and noun phrases as chunks and if any word matches with the nouns in the given list, it is recognized as the item in the given list.

So far, we have the opinion words and potential targets annotated in the conversation, and we want to pair them up and find the target of the ranking. It has been shown in previous work on relation extraction that the shortest dependency path between any two entities captures the information required to assert a relationship between them (Bunescu and Mooney, 2005). Based on the observation that people tend to mention items and their related ranks close to each other, we pair the item with the rank found in its shortest dependency path.

4 Bipartite Graph Construction

We propose an assessment method of the influence factors among participants based on bipartite graphs.

4.1 Dynamic Update of Weights and Vertices

We construct a directed bipartite graph $G = (U \cup V, E)$, where the vertices U represent the participants in the discussion, the vertices V represent

the items, and E denotes the edges between these vertices. u_i denotes the i th vertex or participants’ cumulative informative score in U . v_j is the j th vertex or item ranking in V .

In the Lunar Survival scenario, we observed that the information given in the conversation is very helpful in getting the right result and reaching consensus. To reflect this observation, we have total number of sentences so far for each speaker in the conversation as an informativeness indicator. The edges of the bipartite graph carry weights w_{ij} , representing the relationship between vertices u_i and v_j , i.e., u_i ’s current ranking of v_j . Thus, we can represent all the edge weights of the graph as a $|U| \times |V|$ matrix $W = [w_{ij}]$. With the item-rank pair extracted, we dynamically update W , and calculate v_j as $\frac{\sum_i w_{ij}}{i}$.

4.2 Influence Model

We implemented an influence model (IM) (Basu et al., 2001) to track and understand the participants’ opinion behaviors. We model the participants’ opinion shifts as a Markov chain with each state representing a user’s opinion on the item. We use the coupled HMM to correlate the influence of the opinions among multiple participants. Each participant i has a chain of rankings on the items at time t denoted S_t^i . We assume that

$$P(S_t^i | S_{t-1}^1, \dots, S_{t-1}^N) = \sum_j \alpha_{ij} P(S_t^i | S_{t-1}^j), \quad (1)$$

where α_{ij} (calculated from the model) can tell us how much the state transition of person i is influenced by the given neighbor j .

This observed IM is characterized by (Φ, A) , where Φ is the state transition probability matrix, and A is the influence strength vector. At any time t , we calculate the pairwise transition probability matrix $P(S_t^i | S_{t-1}^j)$ by counting, and determine α_{ij} using the constrained gradient ascent method to maximize per-chain likelihood.

5 Experiments

5.1 Dataset Construction

We curated 5 meetings and transferred the recorded voice to the texts using IBM Watson’s Speech-to-Text API (Saon et al., 2017). The conversations are 10–15 minutes long and have an average of 412 sentences. we collected the initial and final rankings of the items from each person using pre- and post-discussion questionnaires. We

Team ID	Precision	Recall	F1 score
Team 1	0.53	0.21	0.31
Team 2	0.65	0.55	0.6
Team 3	0.71	0.62	0.66
Team 4	0.65	0.33	0.44
Team 5	0.59	0.70	0.64

Table 1: Information extraction measures

performed the opinion extraction and target pairing as described in Section 3 for the 5 meetings. The extraction precision and accuracy compared to human annotated ground truth is summarized in Table 1. The precision is defined as the fraction of correct ranks among all ranks retrieved from the conversation, and recall is the fraction of correct ranks that have been retrieved over all the ranks supposed to be retrieved as in ground truth.

5.2 Meeting Dynamics Analysis

We see that groups have very distinct opinions on the 15 items before each meeting, and that they all achieved consensus at the final stage of the meeting. From the playback of the meeting assistant videos, we have a very clear view of the unfolding speech content that influences the participants' state of mind.

We observed the following patterns that correlate with a participant's influence:

1. Approval of other people first, followed by stating clearly the opinion on an item (e.g., Figure 3)
2. Detailed explanation of the reason to choose a specific rank (e.g., "It's a two hundred mile trek so you need some sort of sustenance for the human body.")
3. Drawing attention before a statement (e.g., "OK so proposition hear me out. We're at nine now right? If we're going forward so what if we put milk powder as ten?")

6 Related Work

Opinion target extraction and pairing: In our context, the targets are constrained to 15 items given beforehand, but they appear in different forms in the conversation. In the context of product review mining, (Hu and Liu, 2004) extracted frequent nouns and noun phrases as product feature candidates. Following that method, (Abu-Jbara et al., 2012) extracted frequent noun phrases

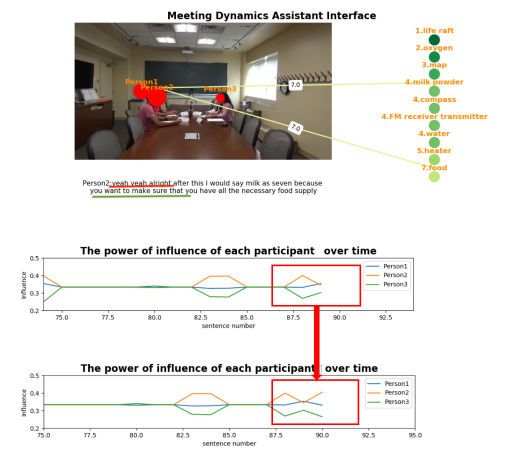


Figure 3: Example of arguments with stronger influence.

and named entities mentioned by different discussants.

As for opinion extraction, various methods were used in different contexts. (Kim and Hovy, 2006) collected opinion-bearing words and classified them into 3 classes. (Ortigosa et al., 2014) also studied opinion from 3 classes. Since in our case, the opinion on an item is restricted to a ranking of 1–15, we used name tagging results to identify the ordinals and numbers mentioned in the conversation.

Group dynamics studies: Most group dynamics studies to date of role recognition or influence studies are based on non-linguistic features. (Rienks and Heylen, 2005) used audio-only features including a collection of nonverbal and verbal cues to perform three-way classification of the participants dominance level. (Beyan et al., 2018) acquired audio and visual features and predicted emergent leadership with multiple kernel learning. Our group has extended the system proposed here to include non-verbal and visual cues to accurately predict emergent leadership and contribution (Bhattacharya et al., 2018).

When modeling opinion shifts, we referred to (Chen et al., 2017) but noticed that these are less complicated in face-to-face conversation than in a social network. (Asavathiratham, 2001) first proposed a simplified coupled-HMM influence model to understand the behaviors of a large number of interacting components. (Basu et al., 2001) expanded the theory and proposed a gradient ascent method to learn the influence model.

7 Conclusion and future work

In this paper we demonstrate a system for meeting assistance, visualizing real-time opinion extraction and group dynamics. We use the Lunar Survival Task to observe how people gradually change their opinions and make decisions. With the current meeting assistant tool, we have a closed set of 15 items given in advance and a fixed set of ranks. In future work, we plan to develop information extraction systems that handle open sets and detect multiple topics in a meeting. The opinions extracted could be used to study group dynamics and recognize roles in meetings, extending the scope of the meeting assistant to more general scenarios.

8 Acknowledgements

Thanks to Mike Foley, Christoph Riedl and Brooke Foucault Welles at Northeastern University for the experimental design. This work was supported by the U.S. National Science Foundation No. IIP-1631674. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

- Amjad Abu-Jbara, Mona Diab, Pradeep Dasigi, and Dragomir Radev. 2012. Subgroup detection in ideological discussions. In *50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*, pages 399–409.
- Chalee Asavathiratham. 2001. *The Influence Model: A Tractable Representation for the Dynamics of Networked Markov Chains*. Ph.D. thesis, Massachusetts Institute of Technology.
- Sumit Basu, Tanzeem Choudhury, Brian Clarkson, Alex Pentland, et al. 2001. Learning human interactions with the influence model. NIPS.
- Cigdem Beyan, Francesca Capozzi, Cristina Becchio, and Vittorio Murino. 2018. Prediction of the leadership style of an emergent leader using audio and visual nonverbal features. *IEEE Transactions on Multimedia*, 20(2):441–456.
- Indrani Bhattacharya, Michael Foley, Ni Zhang, Tongtao Zhang, Christine Ku, Cameron Mine, Heng Ji, Christoph Riedl, Brooke Foucault Welles, and Richard J. Radke. 2018. A multimodal-sensor-enabled room for unobtrusive group meeting analysis. In *ACM International Conference on Multimodal Interaction*.
- Razvan C Bunescu and Raymond J Mooney. 2005. A shortest path dependency kernel for relation extraction. In *Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 724–731. Association for Computational Linguistics.
- Chengyao Chen, Wenjie Li, Dehong Gao, and Yuexian Hou. 2017. Exploring interpersonal influence by tracking user dynamic interactions. *IEEE Intelligent Systems*, 32(3):28–35.
- Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating non-local information into information extraction systems by Gibbs sampling. In *43rd Annual Meeting on Association for Computational Linguistics*, pages 363–370. Association for Computational Linguistics.
- Jay Hall and Wilfred Harvey Watson. 1970. The effects of a normative intervention on group decision-making performance. *Human Relations*, 23(4):299–317.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 168–177. ACM.
- Soo-Min Kim and Eduard Hovy. 2006. Extracting opinions, opinion holders, and topics expressed in online news media text. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text*.
- Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. In *52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60.
- Alvaro Ortigosa, José M Martín, and Rosa M Carro. 2014. Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, 31:527–541.
- Rutger Rienks and Dirk Heylen. 2005. Dominance detection in meetings using easily obtainable features. In *International Workshop on Machine Learning for Multimodal Interaction*, pages 76–86. Springer.
- George Saon et al. 2017. English conversational telephone speech recognition by humans and machines. In *Proceedings of INTERSPEECH 2017*.