Identifying Emergent Leadership in Small Groups using Nonverbal Communicative Cues

Dairazalia Sanchez-Cortes[1,2], Oya Aran[1], Marianne Schmid-Mast[3] and Daniel Gatica-Perez[1,2]

[1] Idiap Research Institute, Martigny, Switzerland
[2] Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[3] University of Neuchâtel (UNINE), Switzerland

(dscortes,oaran,gatica)@idiap.ch, marianne.schmid@unine.ch

ABSTRACT

This paper addresses firstly an analysis on how an emergent leader is perceived in newly formed small-groups, and secondly, explore correlations between perception of leadership and automatically extracted nonverbal communicative cues. We hypothesize that the difference in individual nonverbal features between emergent leaders and non-emergent leaders is significant and measurable using speech activity. Our results on a new interaction corpus show that such an approach is promising, identifying the emergent leader with an accuracy of up to 80%.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing

General Terms

Human Factors

Keywords

Emergent leadership, Nonverbal behavior, Speech activity

1. INTRODUCTION

Leadership and the interaction between co-workers in organizations are critical variables for the success of many of the faced tasks. Nowadays visionary organizations are hiring team leaders based on different interviews using task-solving problems, and through the observation of emergent leaders in assessment centers [6].

An emergent leader is defined as the one who arises from an interacting group and has his base of power from followers rather than from a higher authority [13]. Emergent leadership is a key research area in social psychology and there are a number of works that analyze verbal and nonverbal behaviors displayed by emergent leaders, how it is perceived by observers and how it can be measured.

In [13] observers were able to identify emergent leadership in small groups from both verbal and nonverbal information. Verbal communication was transcribed from videotapes. Nonverbal communication was tested with a visual only setup and an audio-visual setup, where the audio is filtered such that it provides only some vocal nonverbal information. For emergent leadership, highest correlated values were obtained between filtered speech and participation which is defined as the relative amount of time each group member spent talking).

The relationship between leadership and several personality traits is also of interest to social psychologists. It is showed that cognitive ability and the personality traits of extroversion and openness to experience were predictive of emergent leadership behaviors [9]. Another study [8] investigated the relationship between the leadership style and sociable and aggressive dominance. Although both types of dominance have characteristics that lead to leadership, there is a high correlation between leadership and sociable dominance. Sociably dominant people receive more frequent and longer-lasting glances from the group, they look at others more while speaking and use more gestures. Their scores correlate with total speaking time and average turn duration [10]. On the other hand, aggressively dominant people attempt to interrupt more, and they look at others less while listening.

The relationship between dominance and influence in face-to-face groups was analyzed in [2]. They conclude that dominant people influence their group more than individuals lower in dominance by acting competent. In order to attain this influence, they speak the most, and gain more control over the group and the group decisions.

To our knowledge, there has been no work focused on computational models to predict emergence of the leadership in small group conversations. However, several recent works are focused on automatically analyzing group interaction, human behavior and personality from nonverbal cues [5]. Automatically extracted audio and visual nonverbal features are used to predict personality traits such as extroversion and locus of control [11] and individual behaviors like dominance, role or status [12, 7], just as group behaviors like influence, cooperation, and competition [4, 7]. Our work shows some similarities with these recent works in terms of sensors and nonverbal features, however we are focusing on a different aspect of social interaction, emergent leadership.

Ours is perhaps the first study on automatic prediction of emergent leadership. We study two novel research questions

in the context of predicting emergent leadership in small groups based on automatic sensing. Firstly, is there a correlation between how the emergent leader is perceived and his nonverbal behavior? And secondly, can we predict emergent leadership using automatically extracted acoustic nonverbal cues?, using a new interaction corpus explicitly collected to study this task. Our findings reveal a positive correlation between nonverbal behavior and the emergent leader, as well as a strong relation between perceived dominance and emergent leadership. Our method can correctly identify the emergent leader with an accuracy of up to 80%.

We describe our approach in Section 2. In Section 3, we present the data corpus used in this study. Section 4 explains the nonverbal feature extraction and leadership estimation methods. In Section 5, we present the analysis of our results and conclude in Section 6.

2. OUR APPROACH

To analyze emergence of leadership two sets of data per group interaction were collected. The first set includes audio-visual recordings of a group performing a task. The second set includes questionnaires filled by each group member.

From the questionnaires, we coded and averaged several variables, as described in Section 3. From the audio recordings, we automatically extract a number of nonverbal cues to characterize individual participants. We then analyze the correlation between variables coded from questionnaires and acoustic features. We also present a method to automatically infer the emergent leader using acoustic nonverbal cues. Figure 1 shows our approach.
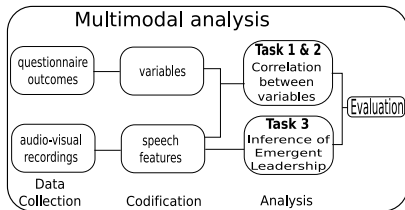

Figure 1: Our approach.

In our work we address three goals

1. Investigate the existence of correlations between questionnaire outcomes and how participants are perceived.

2. Verify if there is correlation between how participants are perceived and their nonverbal acoustic cues.

3. Predict the emergent leadership in small groups using simple acoustic nonverbal features.

3. DATA

The Emergent LEAder data corpus (ELEA) consists of approximately 300 minutes of recorded meetings. There are 12 four-person meetings and eight three-person meetings in newly formed groups, i.e. composed of previously unacquainted people. Participants in ELEA meetings are asked to participate in the Winter Survival Task with no roles assigned.

Sensing infrastructure: Audio recordings were gathered using the Microcone, a commercial microphone array [1]. For video recordings we use two setups, one static setup

with six cameras (four close-ups, two side-views, and one center-view), and one portable setup with two web-cameras (see Figure 2).

Scenario: To recruit participants we posted advertisements in a university asking for volunteers to participate in a study on casual social interactions. Volunteers were asked to participate in the study for approximately one hour. The recruitment process, questionnaires and test were available both in English and French.

Participants first chose a letter to keep their names anonymous, then they filled questionnaires about themselves, performed the winter survival task, and finally were asked to fill in questionnaires based on the perceived interaction.

Survival task: For our study we chose the Winter Survival Test given that it is the most cited test in studies related to small group performance and leadership [9]. It is focused on ranking a list of items in order to survive an airplane crash in winter. The ranking is performed first individually and then as a team, to favour interaction between the participants and allow the emergence of the leader. At the end, the individual scores are calculated based on the absolute difference with the group ranking, and the group score is calculated based on the absolute difference with the survival experts ranking.

Questionnaires: Participants were asked to answer 17 statements that capture how they perceive each participant (including themselves). 16 of the statements were evaluated on a five-point scale. Variables included in these statements are: perceived leadership (PLead), perceived dominance (PDom), and perceived competence (PComp). One other statement asks for the ranking of group dominance (RDom) for all participants in the group, assigning 1 to the person considered the most dominant during the interaction, and 3 or 4 for the least dominant. As a result, we obtain four questionnaire outputs for each participant, which reflect the participants perception.
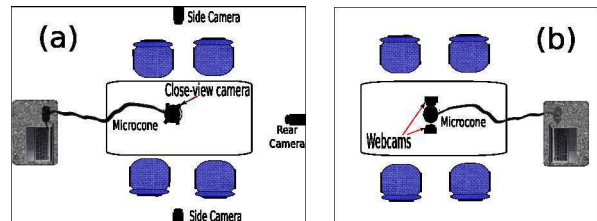

Figure 2: The recording setups of the ELEA corpus: (a) Static setup and (b) Portable setup.

4. AUTOMATIC LEADERSHIP ANALYSIS

Our aim in this work is to build computational models to analyze and predict emergent leadership in small group converations. Among the three goals that we address, the first two require automatically extracted nonverbal cues. The third goal also requires the design of a model which estimates the emergent leader given the automatically extracted nonverbal features. This section details the feature extraction and estimation techniques that we use.

4.1 Extracting Nonverbal Cues

In this work we study in detail the nonverbal cues related to speaking activity. The analysis of visual features is clearly relevant but has been deferred as part of future work.

	PLead	PDom	RDom	PComp
PLead	1	0.84	0.78	0.44
PDom	0.84	1	0.73	0.38
RDom	0.79	0.72	1	0.29
PComp	0.44	0.38	0.29	1

Table 1: Correlation values between variables from questionnaires outcomes.

Acoustic nonverbal cues. The microcone automatically generates the speaker segmentations [1]. This results in a binary segmentation for each participant where status 1 represents speaking, and status 0 represents non-speaking. From the binary segmentation we compute the following features for each participant.

Speaking Length (TSL_i): The total time that participant i speaks according to the binary speaking status.

Speaking Turns (TST_i): Number of turns accumulated over the entire meeting for each participant i, where each turn is a continuous speaking status. We added a variant $TSTf_i$ which accumulates turns bigger than two seconds.

Speaking Turn Duration ($TSTD_i$): The average turn duration per participant i over the entire meeting.

Successful Interruptions (TSI_i): We use two ways to calculate the successful interruptions:

TSI_i^1:Participant i successfully interrupts participant j if i starts talking when j is already speaking, and j finishes his/her turn before i does.

TSI_i^2:Participant i successfully interrupts participant j if i starts talking when j is already speaking, and when i finishes his/her turn, j is not speaking anymore.

For each of the two cases, we added a variant ($TSIf_i^1$ and $TSIf_i^2$) which accumulates interruptions for turns bigger than two seconds.

4.2 Predicting the Emergent Leader

For the task of predicting the emergent leader, our hypothesis is that the emergent leader in a group is the one who has the highest value of a nonverbal feature (i.e. the participant with the highest total speaking time). We define a rule-based estimator that selects the participant with the maximum feature value in the group as the emergent leader. Thus, we estimate the leader by

$$EL_m^f = \arg\max_p(f_p^m), p \in \{1, 2 \ldots P\}, \qquad (1)$$

where p is the participant number, f_p^m is the value of feature f for participant p in group m, and P is the number of participants ($P = 3$ or 4 in our case).

We also applied score level fusion to investigate whether the combination of features has an advantage on using single features. We fused rule based estimators defined on different individual features and used the normalized feature values as the scores of each estimator as proposed in [3]. The normalization is done via z-normalization. For group m, using feature combination $\mathcal{C}$, we sum up the scores for each participant and select the participant with the highest total score:

$$EL_m^{\mathcal{C}} = \arg\max_p(\sum_{f \in \mathcal{C}} \hat{f}_p^m), \ \ \mathcal{C} \subseteq \mathcal{F}, \qquad (2)$$

where $\hat{f}_p^m$ is the score of participant p using feature f in group m, and $\mathcal{F}$ is the set of all features.

5. RESULTS

	PLead	PDom	RDom	PComp
TSL	0.51	0.46	0.49	0.28
TSTD	0.44	0.39	0.40	0.19
TST	0.34	0.35	0.36	0.36
TSTf	0.60	0.60	0.53	0.27
TSI^1	0.60	0.56	0.52	0.34
$TSIf^1$	0.62	0.60	0.54	0.26
TSI^2	0.40	0.41	0.35	0.17
$TSIf^2$	0.57	0.57	0.55	0.27

Table 2: Correlation values between variables from questionnaires and nonverbal acoustic features.

Questionnaire output analysis. For the first task, we analyzed the correlation of the questionnaire outputs. Each perceived variable was averaged over all participants per group, and group ranking was normalized according to the number of participants per group. Table 1 shows the Pearson correlation values. PLead shows correlation with PDom and RDom (0.84 and 0.78 respectively). These results suggest that the emergent leader is perceived as a dominant person by the other participants. Interestingly, the correlation between perceived leadership and competence is less strong and lower between perceived or ranked dominance and competence, which suggests that participants might not have used the latter construct as part of their judgements.

Nonverbal cues-questionnaire analysis. For the second task, Table 2 shows Pearson correlation values between questionnaire outputs and individual nonverbal features. As we can see there is a correlation between several nonverbal features and PLead, suggesting that emergent leadership perception has a connection to the person who talks the most, has more turns, and interrupts the most. Furthermore, several nonverbal cues have also correlation (although with lower values) with perceived or ranked dominance. This confirms previous work that shows that these features are reasonably correlated with dominance in groups [5, 10]. Finally, the nonverbal features have a medium correlation with judgment of competence [2], which suggests that they might still have some discriminating power.

Automatic inference. For the evaluation of our approach, we use the variables from questionnaires as a ground truth. We calculated the accuracy of the rule-based estimator by comparing the selected emergent leader with the participant who has the highest value for each of the nonverbal cues. Figure 3 shows the accuracy using single features, where the best accuracy for variable PLead is achieved using $TSIf^2$ and TSTD with 70%, followed by TSTf and $TSIf^1$ with 65%. Random behavior in this case is 28.3%.

Table 3 shows the results of score level fusion. For each of the variables, we observe an increase in the accuracy, with 10% for PLead, 5% for PDom, 10% for RDom and 5% for PComp. This shows that cue integration is beneficial.

Additionally for this task, we performed a thin slice analysis to explore the temporal support needed by our approach. We compute the same features, originally computed for the whole interaction, for smaller slices, and then estimate the emergent leader with the rule-based estimator using the features computed for each slice. We define the duration of the slices as multiples of 1/8 of the original duration, where each slice starts from the beginning of the meeting. Figure 4 shows the accuracy of the thin slice based estimation with respect to the variables PLead and PDom. We see that after
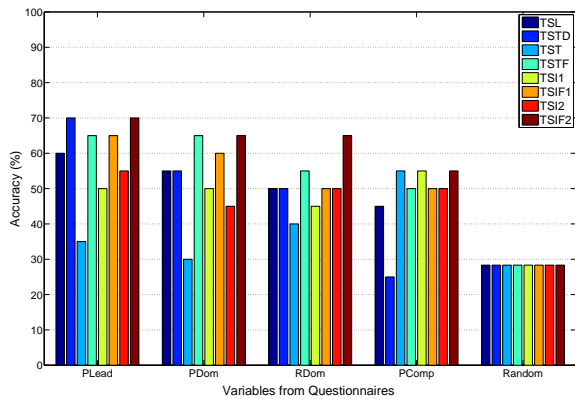
Figure 3: Accuracy of individual features on predicting emergent leadership and related concepts

Variable	Accuracy (%)	Features combined
PLead	80	TSTD, TSI^1
PDom	70	TSTD, TSI^1
	70	TSTD, TSIf^2
	70	TSTf, TSIf^2
RDom	70	TSTD, TSI^2
	70	TSTD, TSIf^2
PComp	60	TSL, TST
	60	TST, TSI^1
	60	TSTf, TSIf^2

Table 3: Results of score level fusion for emergent leadership and related concepts

the first half of the recording (7.49 minutes on average), the estimations follow a trend and change only slightly.
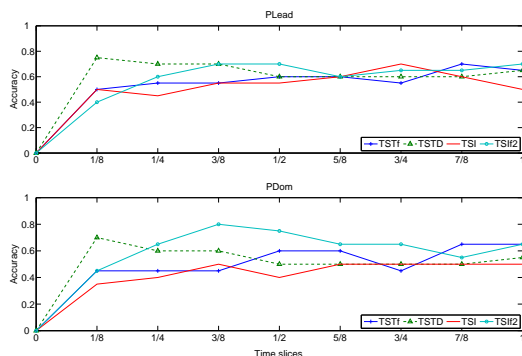

Figure 4: Analysis of time slices from audio. Variables Plead (top) and PDom (bottom) for best performing features in score level fusion

The above results show that the rule based estimator, which selects the participant with the highest value as the emergent leader, performs reasonably well in comparison to the ground truth. To see the potential of this approach to be used as a baseline for more complex estimators, we analyzed the case where we consider also the participant with the second highest feature value. In this case we assume that the estimation is correct if the participant with either the highest or the second highest feature value matches the ground truth. We observed that the accuracy to predict RDom in small groups increases up to 90% for TST and for PLead accuracy increases up to 85% using TSIf. The random performance in this case is 56%. These results suggest that

quite often the second ranked person is indeed the leader (or dominant) person when the first one is not.

6. CONCLUSIONS

In this work, we identified emergent leadership in newly formed groups on a new group interaction corpus, by computing speech activity based nonverbal cues. We evaluated the effectiveness of both individual cues and score level fusion in detecting the emergent leader. Based on the results we noticed that the emergent leader was perceived by his/her peers as a dominant person, who talks the most, and has more turns and interruptions (i.e. makes more suggestions). In addition to perceived dominance, the speaking behavior make emergent leaders sometimes appear to be competent. Clearly, such observations are limited given the moderate size of our dataset. In the future, we plan to expand our dataset to verify the generality of our findings, analyze another prosodic speech features and to extract visual features in addition to audio.

7. REFERENCES

[1] http://www.dev-audio.com/.
[2] C. Anderson and G. J. Kilduff. Why do dominant personalities attain influence in face-to-face groups? the competence- signaling effects of trait dominance. *Journal of Personality and Social Psychology*, 96(2):491–503, 2009.
[3] O. Aran and D. Gatica-Perez. Fusing audio visual nonverbal cues to detect dominance in small group conversations. In *Proc. ICPR*, Aug 2010.
[4] S. Basu and et al. Towards measuring human interactions in conversational settings. In *IEEE CVPR-CUES*, Dec 2001.
[5] D. Gatica-Perez. Automatic nonverbal analysis of social interaction in small groups: a review. *Image and Vision Computing*, 1(12), Dec 2009.
[6] L. D. Goodstein and R. I. Lanyon. Applications of personality assessment to the workplace: a review. *Journal of Business and Psychology*, 13(3), 1999.
[7] D. Jayagopi, B. Raducanu, and D. Gatica-Perez. Characterizing conversational group dynamics using nonverbal behavior. In *ICME*, New York, June 2009.
[8] A. K. Kalma, L. Visser, and A. Peeters. Sociable and aggressive dominance: Personality differences in leadership style? *Leadership Quarterly*, 4(1):45–64, 1993.
[9] J. Kickul and G. Neuman. Emergent leadership behaviours: The function of personality and cognitive ability in determining teamwork performance and ksas. *Journal of Business and Psychology*, 15(1), 2000.
[10] M. S. Mast. Dominance as expressed and inferred through speaking time. a meta-analysis. *Human Communication research*, 28(3):420–450, 2002.
[11] F. Pianesi, N. Mana, and A. Cappelletti. Multimodal recognition of personality traits in social interactions. In *ICMI*, Chania, Oct 2008.
[12] B. Raducanu and D. Gatica-Perez. You are fired! nonverbal role analysis in competitive meetings. In *IEEE ICASSP*, Apr 2009.
[13] R. T. Stein. Identifying emergent leaders from verbal and nonverbal communications. *Personality and Social Psychology*, 32(1):125–135, 1975.