Characterizing Small Group Dynamics: classification and mining

Dinesh Babu Jayagopi, Idiap Research Institute
Daniel Gatica-Perez, Idiap Research Institute
Related Work – small group, face-to-face

- Individual behaviour
  - Dominance and Status (Jayagopi et al. 2009)
  - Roles (Vinciarelli 2007)
  - Personality (Pianesi et al. 2008)

- Group behaviour
  - Small set of meetings – interactivity and centrality (Otsuka et al. 2006)

- Conversational analysis
  - Group meeting activities – monologue, discussion, presentation (Zhang et al 2005, Dielmann et al 2007)
Modeling - Individual or Group behaviour

gaining ahead

gaining along
Applications: characterizing groups

- Monitoring team cohesiveness
- Identify leadership skills
- Identify irresponsible behaviour
Two different approaches

• Classification task
  – discriminative training
  – requires ground-truth

• Mining task
  – Clustering
  – does not require ground-truth
Two different approaches

Classification task
  discriminative training
  requires ground-truth

Mining task
  Clustering
  does not require ground-truth
Samples – two very different groups

AMI meeting dataset:
To design a remote control together
Each participant had a role

The Apprentice dataset:
To fire a participant
6th season of a TV show

COOPERATIVE

COMPETITIVE
Motivation and hypotheses

Motivation

Objectives of a group => Group Dynamics (Mc Grath 1984)

Hypothesis 1

Hypothesis 2

Group Dynamics

Nonverbal Cues
Our Approach
Meeting Dataset

34 five-min meetings
4 participants

15 meetings
Average dur - 6 min
Participants median - 7
Nonverbal cues: speaking activity based

- from head set microphones
  - from speaking-turn segmentation
    - speaking length (TSL)
    - number of turns (TST)
    - number of successful interruptions (TSI)
Nonverbal cues: speaking activity based

Fraction of Overlapped Speech \( FO = \frac{O}{T} \)

Fraction of Silence \( FS = \frac{S}{T} \)

Fraction of Non-overlapped Speech \( FN = \frac{N}{T} \)
Nonverbal cues: speaking activity based

**Features characterizing individuals**

- **Speaker Segmentation**
- **TSL, TST, TSI**

**Features characterizing the group**

- **FO, FS, FN**
- **GSL, GST, GSI, GIT**
- **GLEM, GTEM, GIEM**

**Group Speaking Length (GSL)**

\[ \sum_i TSL(i) / \text{MeetingDuration} \]

**Group Speaking Turns (GST)**

\[ \sum_i TST(i) / \text{MeetingDuration} \]

**Group Speaking Interruptions (GSI)**

\[ \sum_i TSI(i) / \text{MeetingDuration} \]

**Group Speaking Interruption-to-Turns Ratio (GIT)**

\[ \sum_i TSI(i) / \sum_j TST(j) \]
Nonverbal cues: speaking activity based

Features characterizing individuals
- Speaker Segmentation

Features characterizing the group
- FO
- FS
- FN
- GSL
- GST
- GSI
- GIT
- GLEM
- GTEM
- GIEM

Egalitarian measure - Bhatttacharya Distance
- BD(input, egal. Vector)
- 0 if egalitarian

TST - [20 15 10 5]
- Input - [0.4 0.3 0.2 0.1]
- Egal - [0.25 0.25 0.25 0.25]
- GTEM = 0.17

Group Speaking Length Egalitarian Measure
Group Speaking Turns Egalitarian Measure
Group Speaking Interruptions Egalitarian Measure
Meeting Prediction using 2 classifiers

Using naive bayes classifier (likelihood ratio)

\[ \log\left( \frac{P(A/(f_1, f_2, \ldots, f_N))}{P(B/(f_1, f_2, \ldots, f_N))} \right) \]

Using SVM with quadratic kernel

\[ \log\left( \frac{P(f_1/A).P(f_2/A)\ldots P(f_N/A).P(A)}{P(f_1/B).P(f_2/B)\ldots P(f_N/B).P(B)} \right) \]
Results

• Group Interruption To Turn Ratio (GIT), Group Turn Egalitarian Measure (GTEM)
  Best single cues

  - Combination of GIT and GTEM using an SVM classifies 100 % Correct

  - Fraction of overlap, Group Speaking Ratio did not predict well
Our two hypotheses – ‘competitive and cooperative meetings have different group dynamics’ and ‘nonverbal features capture this difference’ and was verified

Characterization of entire group by the aggregation (both temporal and person-wise) of their nonverbal behaviour is promising

GIT, GTEM are best single cues

Combination of GIT, GTEM using an SVM – 100% accuracy

The current limitation is the moderate size of dataset

Future work would expand dataset, features and classification classes.
Two different approaches

Classification task
    discriminative training
    requires ground-truth

Mining task
    Clustering
    does not require ground-truth
Mining task

10 different groups
Features

Speaker Segmentation

Nonverbal cues related to social verticality

TSL, TST, TSI

Bag-of-NVPs

Generic group patterns

Leadership patterns

Quantize TSL, TST, TSI into one of five Classes (15 words)

Describe the position of the leader (6 words)

Features characterizing individuals

Features characterizing the group

Bag-of-NVPs
- allows fusion of cues
- removes noise in the features
- allows comparison of groups of different size
Generic group patterns - construction

Silence (0)

One (1)

Two (2)

Rest (3)

Equal (4)

$TSL, TST, TSI$

$\Rightarrow 2 * \frac{1}{P}$

$\Rightarrow 3 * \frac{1}{P}$

$\Rightarrow \delta$

$\Rightarrow \text{participants}$
Generic group patterns – at diff. scales

The diagrams illustrate the distribution of speaking length, speaker turns, and successful interruptions across different time intervals (0.5, 1, 2, 3, 4, and 5 minutes). Each panel shows the frequency of different speaking lengths, turns, and interruptions, with bars indicating the proportion of occurrences for each category.
Leadership patterns – at diff. scales

2 min

5 min
Latent Dirichlet Allocation (LDA) model

\[ P(w_i) = \sum_{t=1}^{T} P(w_i | z_i = t)P(z_i = t) \]

Input (example):
SL-One
ST-Two
SI-Equal
SL-L
ST-L
SI-NL

Say T = 3
5 min documents = 873 (overlapping)
2 min documents = 501 (non-overlapping)
### LDA based discovery at 2 min & 5 min

<table>
<thead>
<tr>
<th>Topic 1 - LDA</th>
<th>Topic 2 - LDA</th>
<th>Topic 3 - LDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(z) = 0.32$</td>
<td>$P(z) = 0.33$</td>
<td>$P(z) = 0.35$</td>
</tr>
<tr>
<td>‘Autocratic’</td>
<td>‘Participative’</td>
<td>‘Free-rein’</td>
</tr>
<tr>
<td>Word</td>
<td>$P(w</td>
<td>z)$</td>
</tr>
<tr>
<td>SL-L</td>
<td>0.2</td>
<td>ST-Equal</td>
</tr>
<tr>
<td>ST-L</td>
<td>0.2</td>
<td>SL-Equal</td>
</tr>
<tr>
<td>SI-L</td>
<td>0.16</td>
<td>ST-NL</td>
</tr>
<tr>
<td>SI-Two</td>
<td>0.11</td>
<td>SI-NL</td>
</tr>
<tr>
<td>SI-Rest</td>
<td>0.08</td>
<td>SI-NL</td>
</tr>
<tr>
<td>ST-Rest</td>
<td>0.07</td>
<td>SI-Equal</td>
</tr>
</tbody>
</table>

Table 1: LDA based discovery at 5-minute scale

<table>
<thead>
<tr>
<th>Topic 1 - LDA</th>
<th>Topic 2 - LDA</th>
<th>Topic 3 - LDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(z) = 0.31$</td>
<td>$P(z) = 0.35$</td>
<td>$P(z) = 0.34$</td>
</tr>
<tr>
<td>‘Autocratic’</td>
<td>‘Participative’</td>
<td>‘Free-rein’</td>
</tr>
<tr>
<td>Word</td>
<td>$P(w</td>
<td>z)$</td>
</tr>
<tr>
<td>ST-L</td>
<td>0.22</td>
<td>ST-Equal</td>
</tr>
<tr>
<td>SL-One</td>
<td>0.20</td>
<td>SI-L</td>
</tr>
<tr>
<td>SI-L</td>
<td>0.19</td>
<td>SL-Equal</td>
</tr>
<tr>
<td>ST-One</td>
<td>0.15</td>
<td>ST-NL</td>
</tr>
<tr>
<td>SI-Silence</td>
<td>0.13</td>
<td>SI-Two</td>
</tr>
<tr>
<td>ST-Two</td>
<td>0.06</td>
<td>ST-Rest</td>
</tr>
</tbody>
</table>

Table 2: LDA based discovery at 2-minute scale
Average topic distribution over groups

5 min slices

2 min slices
We investigated the problem of discovering nonverbal group patterns using topic models.

We proposed a novel bag-of-NVPs approach to characterize groups.

Generic group patterns – described the group and leadership patterns coded the position of the leader.

Using an LDA model, the topics discovered fairly mimic well the three classic leadership styles of Lewin et al – autocratic, participative and free-rein.

Future work – annotation to further validate our claims, expand the bag-of-NVPs, try author-topic model etc.
ACKNOWLEDGEMENTS

• Bogdan Raducanu

• Daniel Gatica-Perez