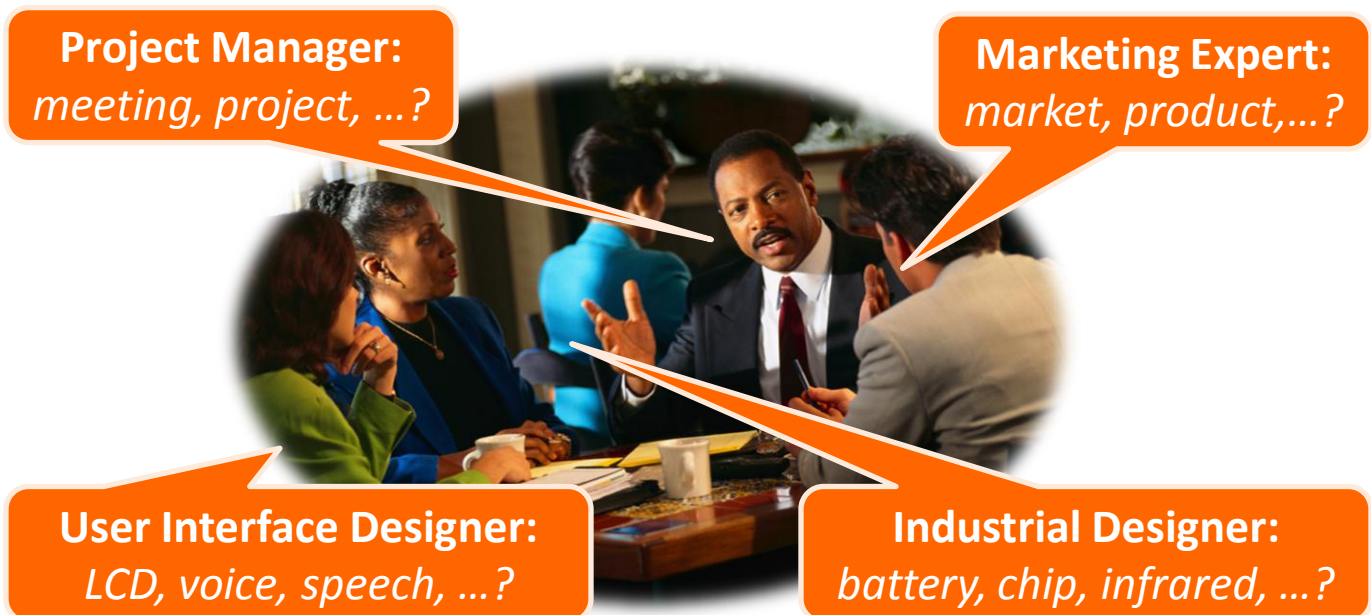


Using Participant Role in Multiparty Meetings as Prior Knowledge for Nonparametric Topic Modeling

Songfang Huang, Steve Renals, CSTR, University of Edinburgh

- A scenario meeting in which the four participants (with different roles) decide to develop a new type of television remote control:



- We are interested in:
 - The word distribution over the role $P(w|r)$
 - Will the role be helpful for topic modeling?
 - Will the role be helpful for automatic speech recognition?

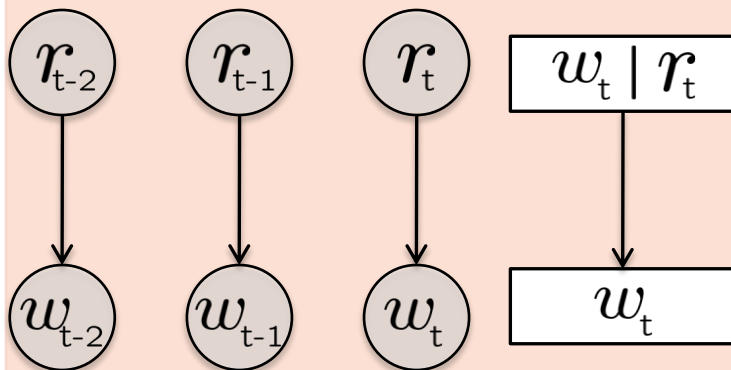
Using Participant Role in Multiparty Meetings as Prior Knowledge for Nonparametric Topic Modeling

Songfang Huang, Steve Renals, CSTR, University of Edinburgh

1. Deterministic Approach:

Factored Language Model

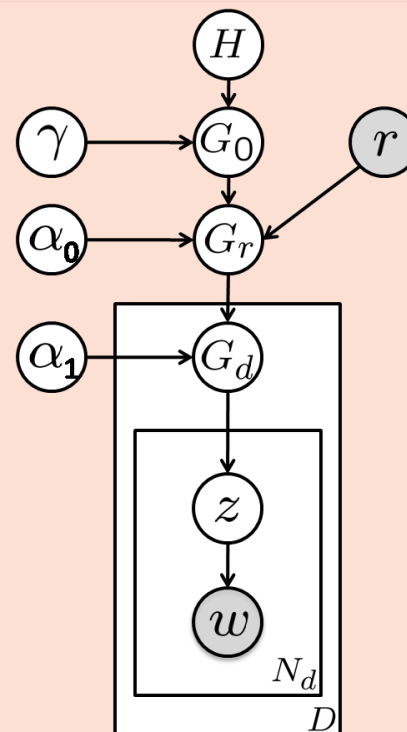
$$w_t \equiv f_t^{1:K} = \{f_t^1, f_t^2, \dots, f_t^K\}$$



$$P(w|r) = \text{count}(w, r) / \text{count}(w)$$

2. Probabilistic Approach:

Hierarchical Bayesian Model



We observed: 1) perplexity reduction for topic models comparing to the HDP, and 2) word error rate reduction for automatic speech recognition in meetings.

