2023 HAI

Leveraging Interactional Sociology for Trust Analysis in Multiparty Human-Robot Interaction

Marc Hulcelle*, Léo Hemamou°, Giovanna Varni¤, Nicolas Rollet*, Chloé Clavel*

* Télécom-Paris, Institut Polytechnique de Paris
Palaiseau, France

° iCIMS
Paris, France

¤ Department of Information Engineering and Computer Science (DISI),
University of Trento, Povo, Italy

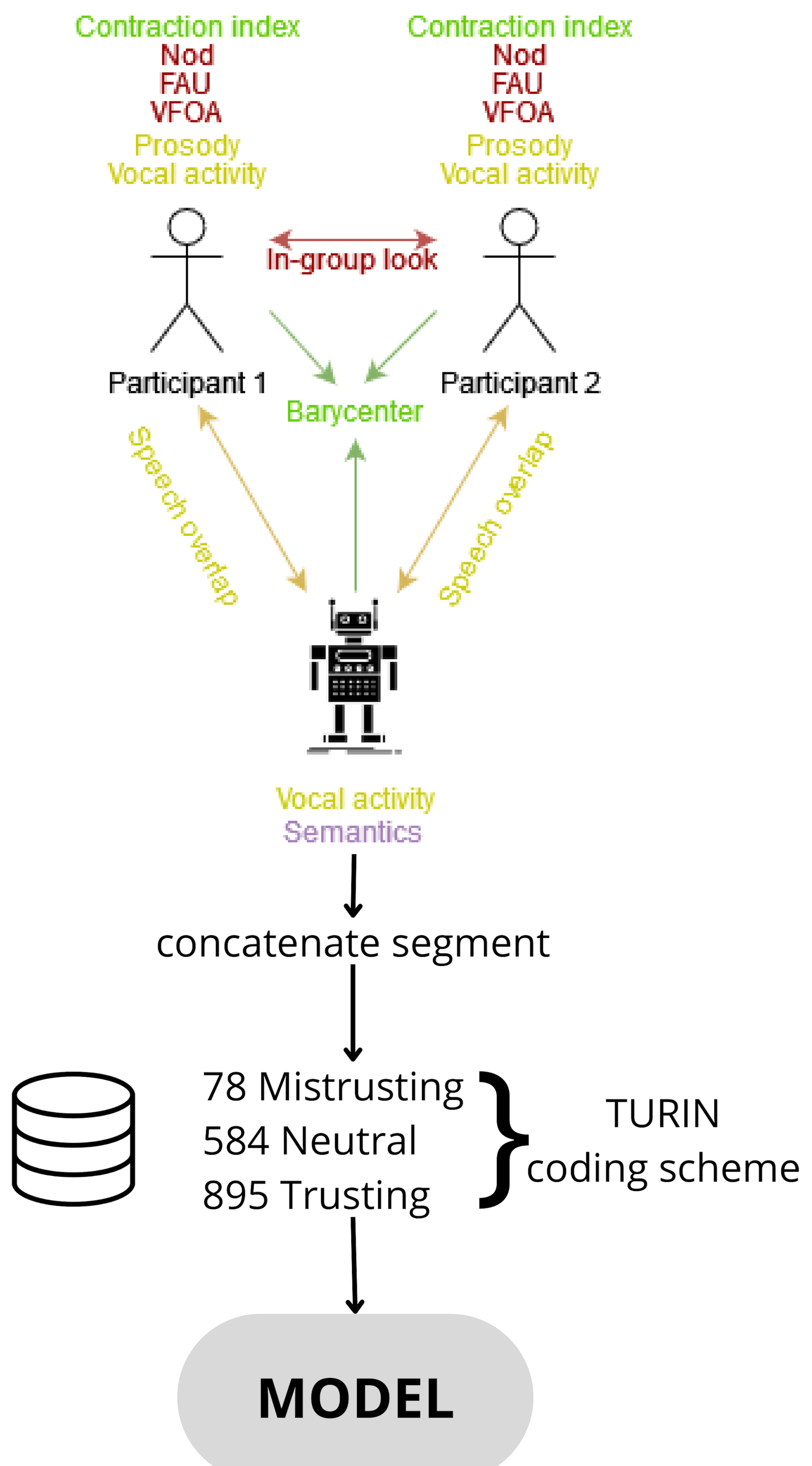
TELECOM Paris
IP PARIS

ABSTRACT

By leveraging Interactional Sociology theories, multimodal behavioral features and recurrent neural architectures, we incrementally build computational models for trust analysis in multiparty human-robot interactions (HRI). We show that the model's performance improves when i) modeling group dynamics with different granularities (i.e. group member, dyadic, and group as a whole), and ii) modeling users-robot interactions as a question-answer sequence.

METHODOLOGY

10 interactions from Vernissage dataset



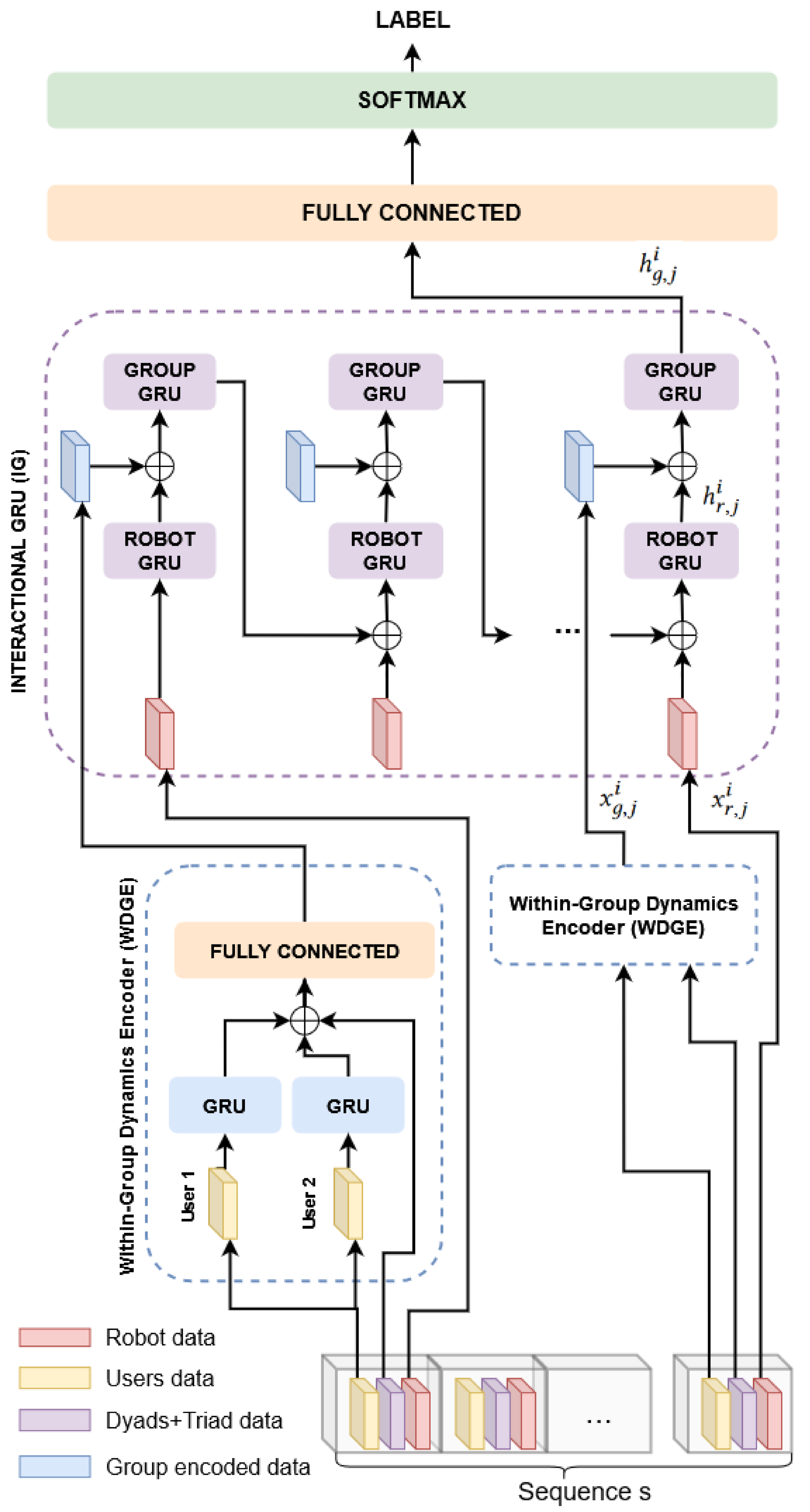
Train as Leave-One-Group-Out

ASSUMPTIONS

A1 [RNN] : users' actions are relevant within the sequence of previous behaviors of all users, and produced in response to another's speaking turn

A2 [WGDE] : participants can either be speakers - addressing the whole group or a part of it - or be listeners - by actively or passively being engaged. It is necessary to analyze the interaction between all users to fully understand the group dynamics

A3 [IG] : participants continuously exchange social signals shaping the interactional context which other participants use to build their answer, and hence renewing the context at each speaking turn



RESULTS

No optimal sequence length ⟶ Sometimes few context is enough

The WGDE module leads to increased performance ⟶ Trust should be analyzed at different levels of the group

The IG module alone does not improve performance ⟶ Segmentation does not properly capture speaking turns ? Hierarchical turn taking modeling is not optimal ?

τ	1	2	3	4	5	6	7	8
SG	0.565 ±.164	0.571 ±.158	0.575 ±.144	0.578 ±.150	0.578 ±.152	0.585 ±.147	0.577 ±.140	0.566 ±.147
WGDE-SG	0.591 ±.138	0.607 ±.134	0.596 ±.138	0.598 ±.133	0.590 ±.138	0.597 ±.137	0.596 ±.130	0.605 ±.144
IG	0.541 ±.149	0.536 ±.136	0.556 ±.123	0.547 ±.132	0.538 ±.128	0.552 ±.150	0.525 ±.136	0.537 ±.161
WGDE-IG	0.572 ±.141	0.580 ±.126	0.584 ±.137	0.585 ±.156	0.596 ±.144	0.592 ±.141	0.580 ±.170	0.560 ±.163

Table 1: Mean and std balanced accuracy on the test sets of the models in the multi-class classification task for $\tau \in [1, 8]$.

τ = Length of history (length of sequence-1)

ERROR ANALYSIS

- Some interactions have far more errors than others
- Segments that were the hardest to classify :
  - Trusting : contained annotations of "Gaze", "Facial Expression", and "F-formation"
  - Mistrusting : contained annotations of "Gaze", "Facial Expression", and "Intonation"
- Most frequent annotations for segments with highest error rate : "Alignment", "Compliance"

PERSPECTIVES

- Method only for offline detection
  - Change segmentation method
  - Automatically extracted features only
- Additional features
  - Semantics for users
  - F-formation features ?
  - Alignment features ?
- More data collected with a trust-specific scenario