Incremental preference elicitation for SRMP models: Application for autonomous drones

Arwa Khannoussi, Alexandru-Liviu Olteanu, Catherine Dezan, Jean-Philippe Diguet, Christophe Labreuche, Jacques Petit-Frère, Patrick Meyer

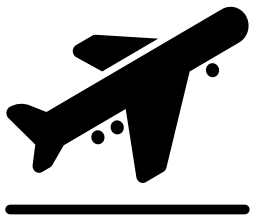
DA2PL'2018

November 23, 2018

Outline

1. Context & Research questions
2. Simple Ranking Method using Reference Profiles (SRMP)
3. Heuristics for incremental learning of SRMP model
4. Experimental Results
5. Conclusion

Drone Context



1. Mission
   - **Military**
   - **Civilian**
   - **Specified by a set of waypoints and objectives**
2. Decision Making
   - **Make autonous decisions**
   - **Take into account multiple criteria to acheive the mission's objectives**
3. Confidence
   - **Guarantee a level of trust in the drone's behavior**
   - **The drone's decisions should be consistent with the priorities of the operator**

Research questions

1. Which MCDA model to integrate into the drone ?

2. How to learn the parameters of the MCDA model ?

RQ1 : Which model?

Real-world constraints

- Decisions related to high-level actions (like land, loiter, skip a waypoint, ...)

  $\rightarrow$ choose one alternative among all possible ones

- Guarantee a high level of trust in the drone's decision

  $\rightarrow$ integrate the preference model onboard of the drone

- The preference model and its consequences are presented to the operator in order to be validated

- Heterogeneous evaluation scales of the criteria:
  quantitative (energy, ...) and qualitative (risk, ...)

RQ1: Which model?

Proposal

- Ranking or choosing in the outranking paradigm

- Simple Ranking Method using Reference Profiles (SRMP) [Rolland 2013]

  ranking can be explained through a series of rules, understandable by the drone operator

RQ2 : How to learn the parameters of the model?

Real-world constraints

- Cognitive effort of the operator should be minimized
- The operator is not an expert of the decision model, and he is probably not able to fix the parameters of the model "directly"
- However : the operator is able to make a preferential choice between pairs of actions (in a simulated environment, with a given context)
- Not all decision actions are possible (existing database of possible pairs of actions, in a given context)

RQ2 : How to learn the parameters of the model?

Proposal

- Learn the SRMP model from pairwise comparisons of alternatives (indirect elicitation)
- Elicit the SRMP model incrementally
- Confront the operator with pairs of possible actions selected from an existing database of possible pairs

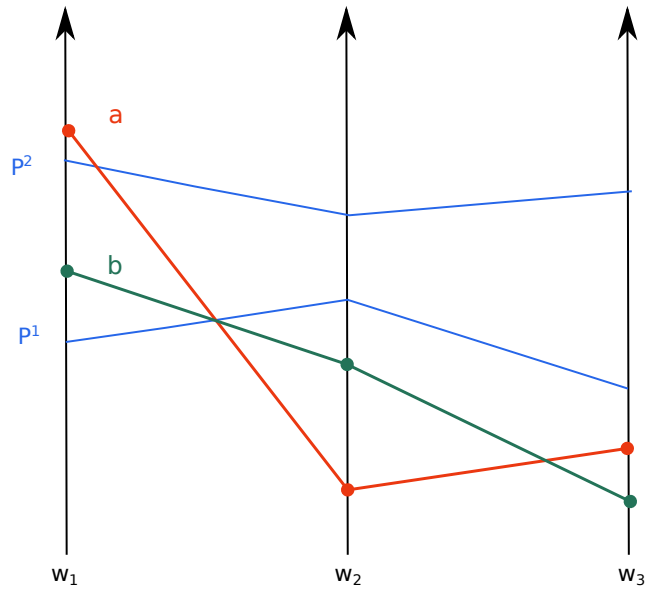
Outline

1. Context & Research questions
2. Simple Ranking Method using Reference Profiles (SRMP)
3. Heuristics for incremental learning of SRMP model
4. Experimental Results
5. Conclusion

Simple Ranking Method using Reference Profiles (SRMP)

- Preference parameters of the decision maker:
  - k reference profiles: $p^1, \ldots, p^h, \ldots, p^k$
  - a lexicographic order σ on the reference profiles
  - criteria weights $w_1, \ldots, w_m$, with $w_j \geqslant 0$ and $\sum_{j \in M} w_j = 1$

- Alternatives are compared indirectly via reference profiles (p^h)
- $a \succsim_{p^h} b$ iff a outranks p^h more strongly than b
- Create a pre-order of the alternatives



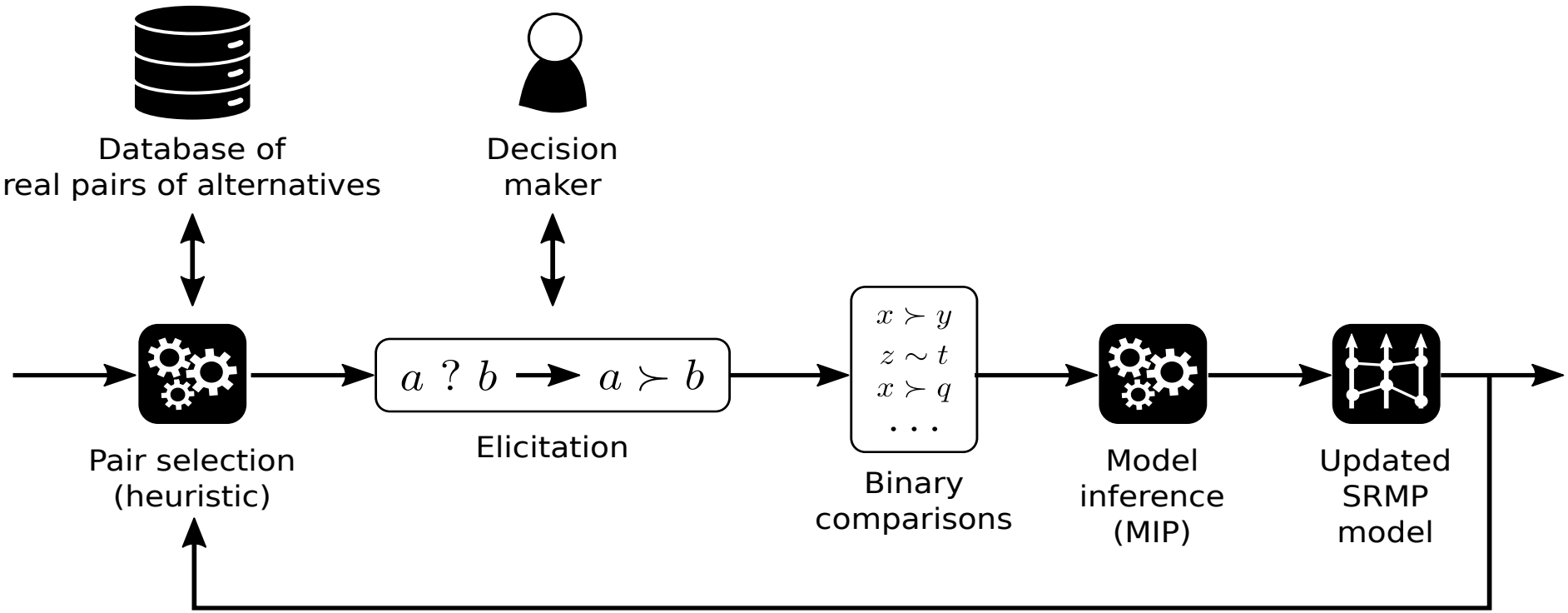
SRMP: Parameters elicitation

- Preference elicitation : tune the parameters of the SRMP model
- In our practical context : indirect elicitation
  - Determine the parameters from holistic judgements of the decision maker on pairs of alternatives (*aPb* / *aIb*)
  - Mixed Integer Program (MIP) [Olteanu et al. 2018]
    Inputs : a set of pairwise comparisons of alternatives
    Outputs : values of the parameters of the SRMP model
- Our contribution :
  - Incremental learning to limit the number of learning example
  - Heuristics to select the learning examples

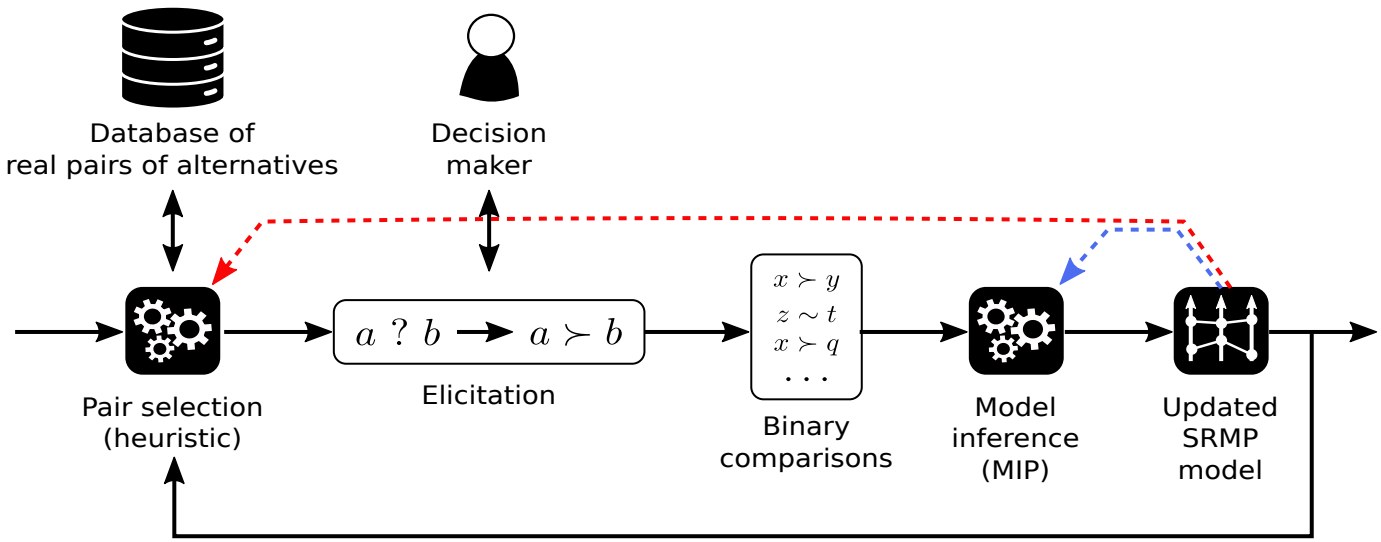
Outline

1. Context & Research questions
2. Simple Ranking Method using Reference Profiles (SRMP)
3. Heuristics for incremental learning of SRMP model
4. Experimental Results
5. Conclusion

Incremental learning of the SRMP model



Incremental learning of the SRMP model



Pair selection (heuristic)

- Random pair
- Pair of most similar alternatives
- Pair of most dissmilar alternatives
- Pair close to previous model profile
- Pair that uses the biggest number of profiles to be discriminated by the previous model (indifference, k profiles, k-1 profile, ..., 1 profile)

Model inference (MIP)

- First feasible model
- Model closest to previous one
- Model centered inside search space

Heuristics: Experimental setting

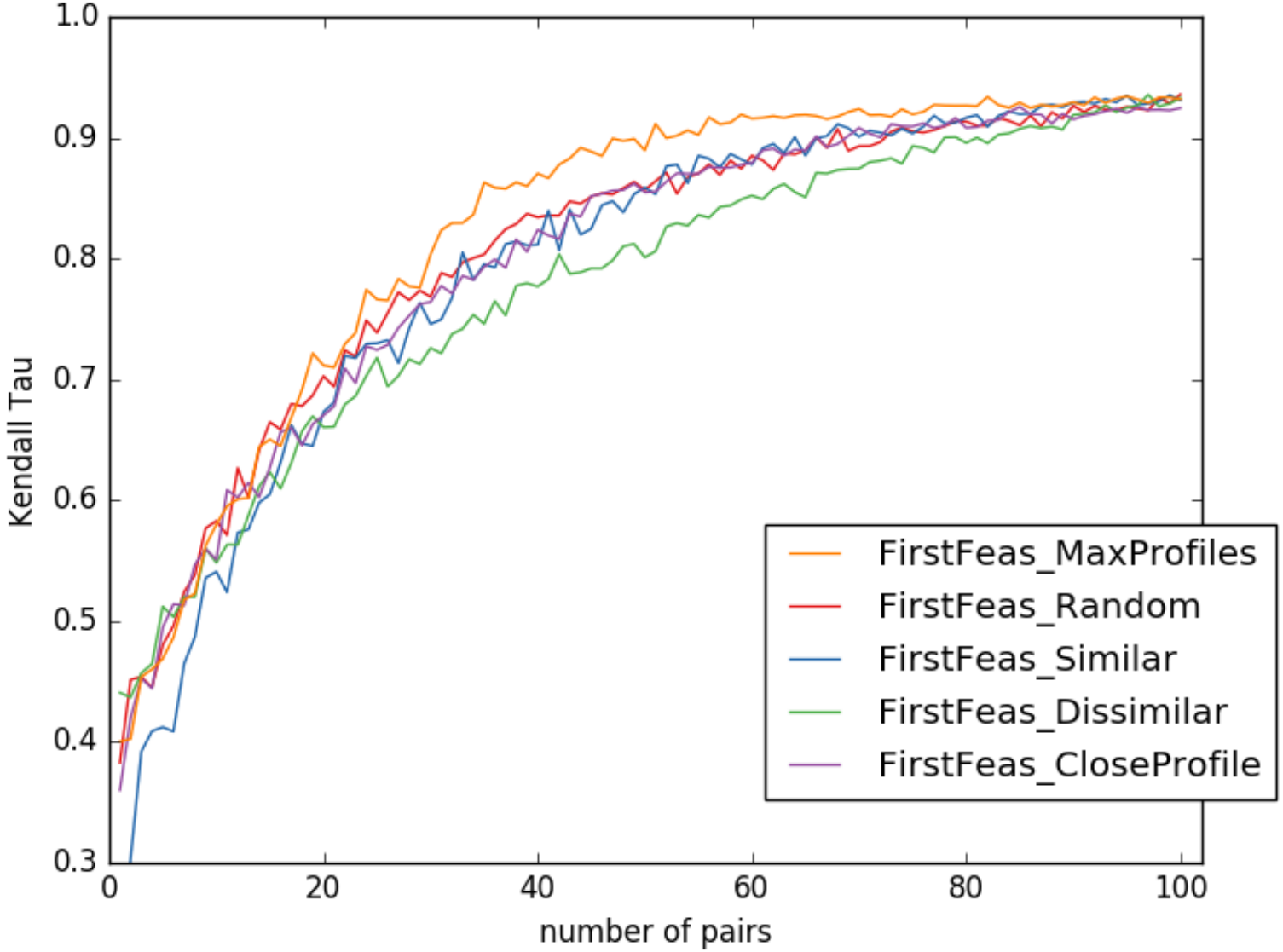
- DM simulation: random SRMP model M_{DM} (fixed k profiles, m criteria)
- Artificial datasets (different n alternatives, m criteria, k profiles)
  - elicitation phase: 100 pairs of alternatives
  - test phase: 5000 alternatives
- At each iteration:
  - a pair of alternatives is selected by the heuristic
  - use M_{DM} to add new constraints to construct a new SRMP model M_i
- Rank the test data using M_{DM} and M_i, and then compare these rankings using *Kendall*'s rank correlation measure τ.
- Each SRMP model configuration (k profiles and m criteria) is experimented 100 times

Outline

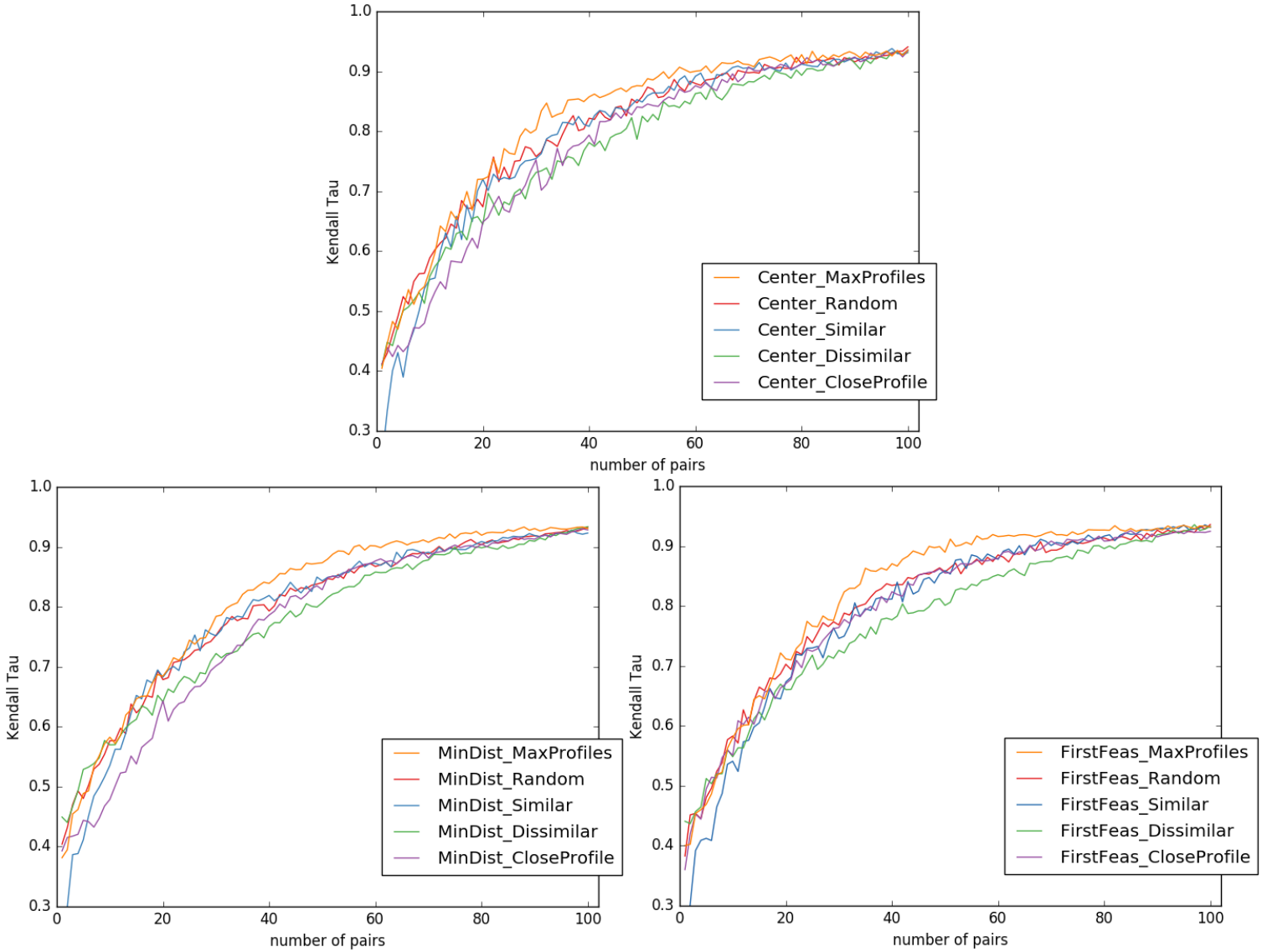
1. Context & Research questions
2. Simple Ranking Method using Reference Profiles (SRMP)
3. Heuristics for incremental learning of SRMP model
4. Experimental Results
5. Conclusion

Experimental Results

mean Kendall tau, 3 criteria, 2 profiles, 100 experiments

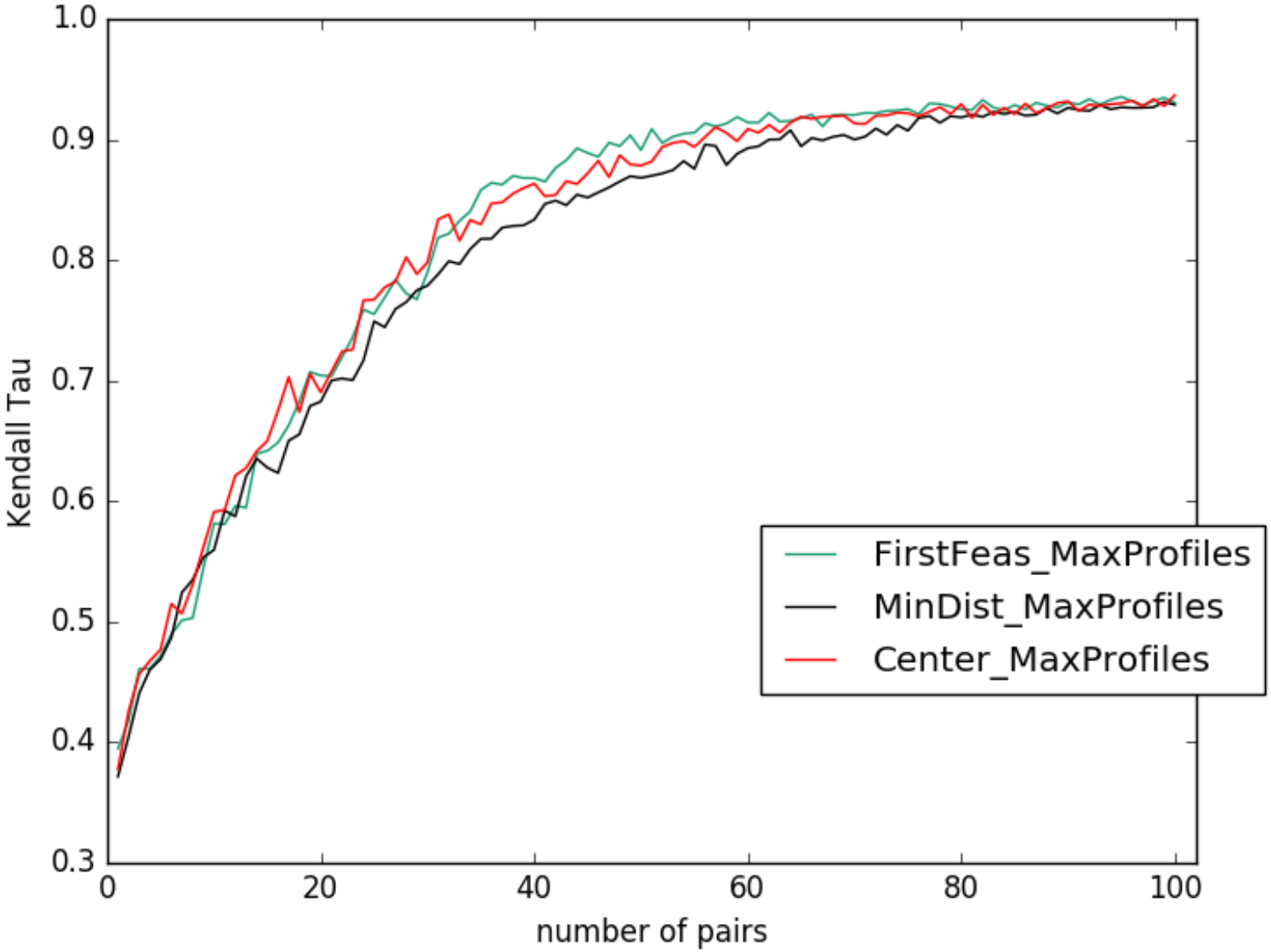


Experimental Results



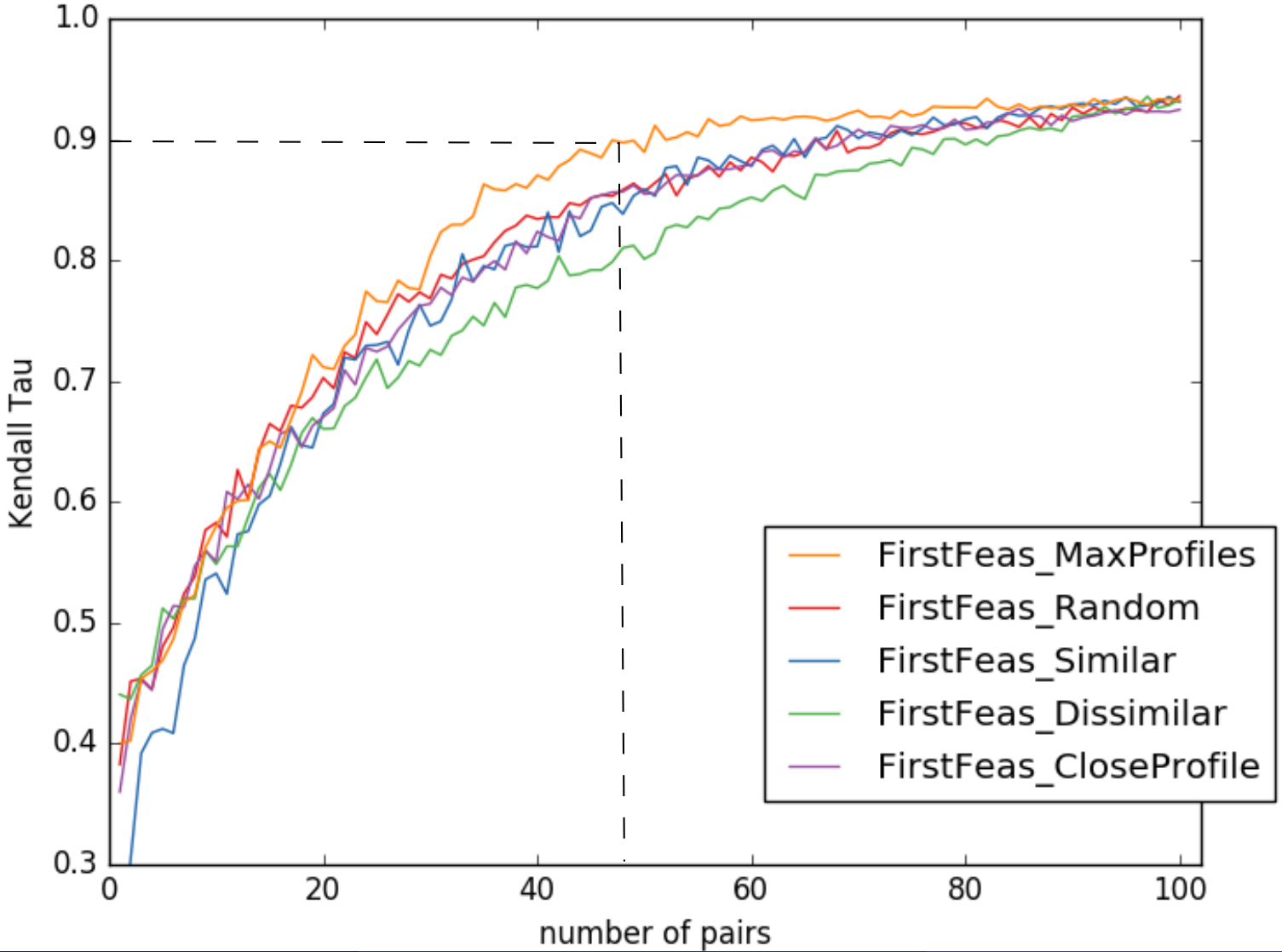
Mean Kendall Tau of 100 experiments for 2profiles 3 criteria

Experimental Results



Experimental Results

Determination of the number of pairs necessary to achieve a "good" preference model



Outline

1. Context & Research questions
2. Simple Ranking Method using Reference Profiles (SRMP)
3. Heuristics for incremental learning of SRMP model
4. Experimental Results
5. Conclusion

Conclusion

- The MaxProfiles heuristic dominates in the different MIP's configuration
- No clear difference between the MIP configurations
  Choice of the least-cost configuration in terms of computation time

- Future work :
  - Experiment new heuristics (Volume, MaxProfiles++, Alternate weights / profiles)
  - Input other preference information (typically the profiles)
  - Experiment other data configurations ((2 profiles, 7 criteria), (3 profiles, 5criteria), . . .)

    Computation time for 100 experiments: one week for 2 profiles and 3 criteria with all the different combination of heuristic and MIP's configuration. (200 cpu, 300GB RAM)
  - Integration of the learning algorithm in the drone simulator

References

Olteanu, A-L. et al. (2018). "Preference Elicitation for a Ranking Method based on Multiple Reference Profiles". working paper or preprint. URL: https://hal.archives-ouvertes.fr/hal-01862334.

Rolland, A. (2013). "Reference-based preferences aggregation procedures in multi-criteria decision making". In: *European Journal of Operational Research* 225.3, pp. 479–486. DOI: 10.1016/j.ejor.2012.10.013. URL: https://doi.org/10.1016/j.ejor.2012.10.013.