Optimal Strategies For Comparing Covariates To Solve Matching Problems

Muhammad Ahmed Shah, Raphael Olivier, Bhiksha Raj

Matching Problems

Given a gallery of N items and 1 probe, retrieve the entries from the gallery that match the probe.

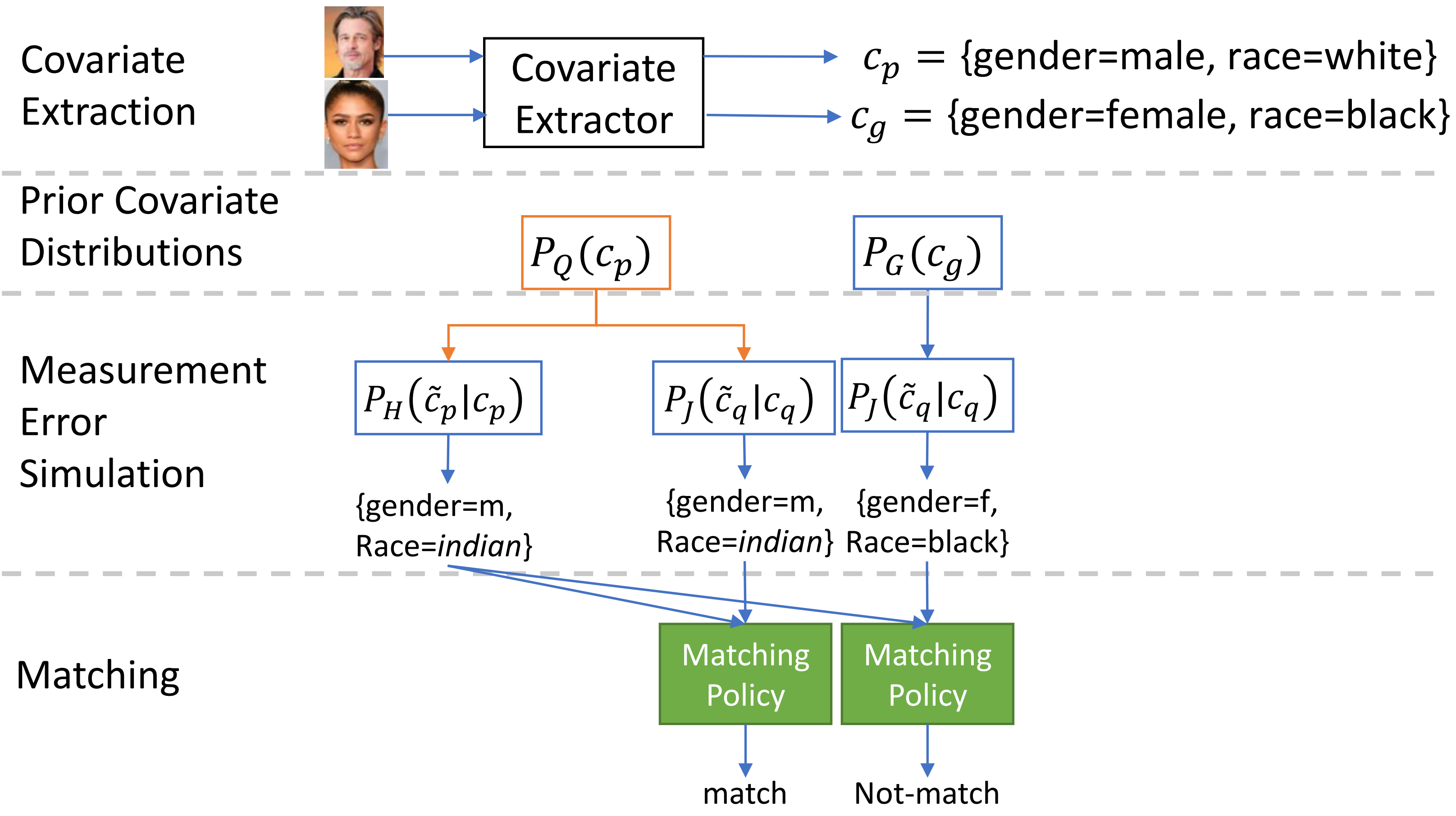
Solving Matching Problems

- Compute *covariates* that statistically relate the probe and the gallery (a.k.a features)
  - Common practice is to use deep neural networks → produces rich but extremely uninterpretable covariates.
- For many tasks common and interpretable covariates exist
  - E.g. gender, height, weight, etc. are covariates for biometric identification
  - Easy to determine the value of these covariates, e.g., by existing classifiers or public data.
  - Lead to simpler and more interpretable models.

Optimal Strategies for Matching

- **Goal**: Quantify the marginal gain of using complex models over matching interpretable covariates
- **Requisite**: Determine the maximum performance achievable by using high-level covariates
  - **Requisite**: need optimal matching strategies!

The Proposed Framework



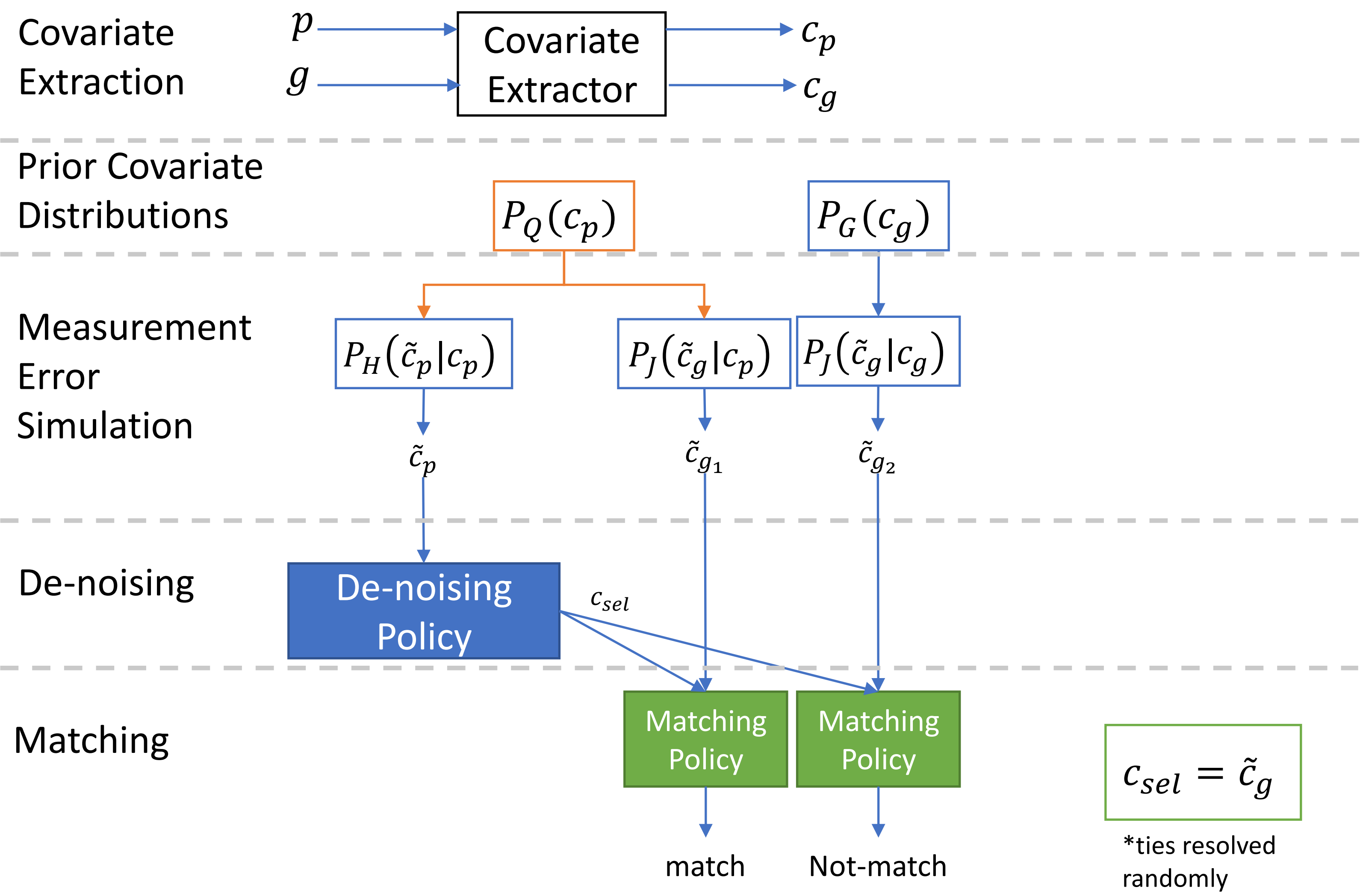
Matching Problems Considered

- Classification
  - Gallery contains prototypes for each class
  - Probes are the data to be classified
- Pairwise-Verification (e.g biometrics)
  - Gallery contains 1 prototype
  - Probes are the data claiming to be instances of the prototype
- Ranking
  - Gallery contains many data samples
  - Probes are data samples against which the gallery must be reordered

Classification

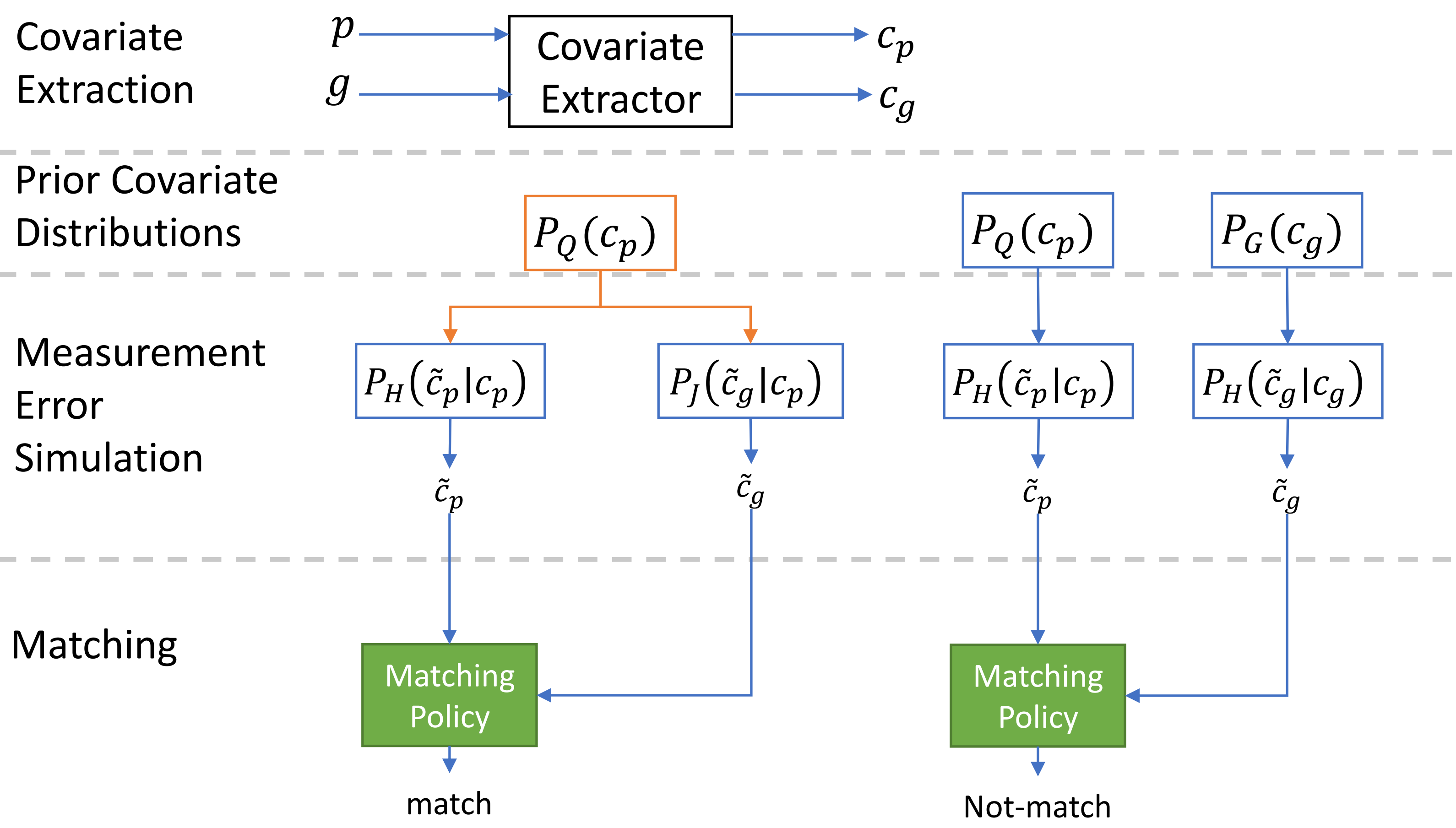
- Gallery contains prototypes for each class
- Probes are the data to be classified
- De-noising policy:

$c_{sel} = argmax_{c \in V} P_{match}(c, c_{g_1})$



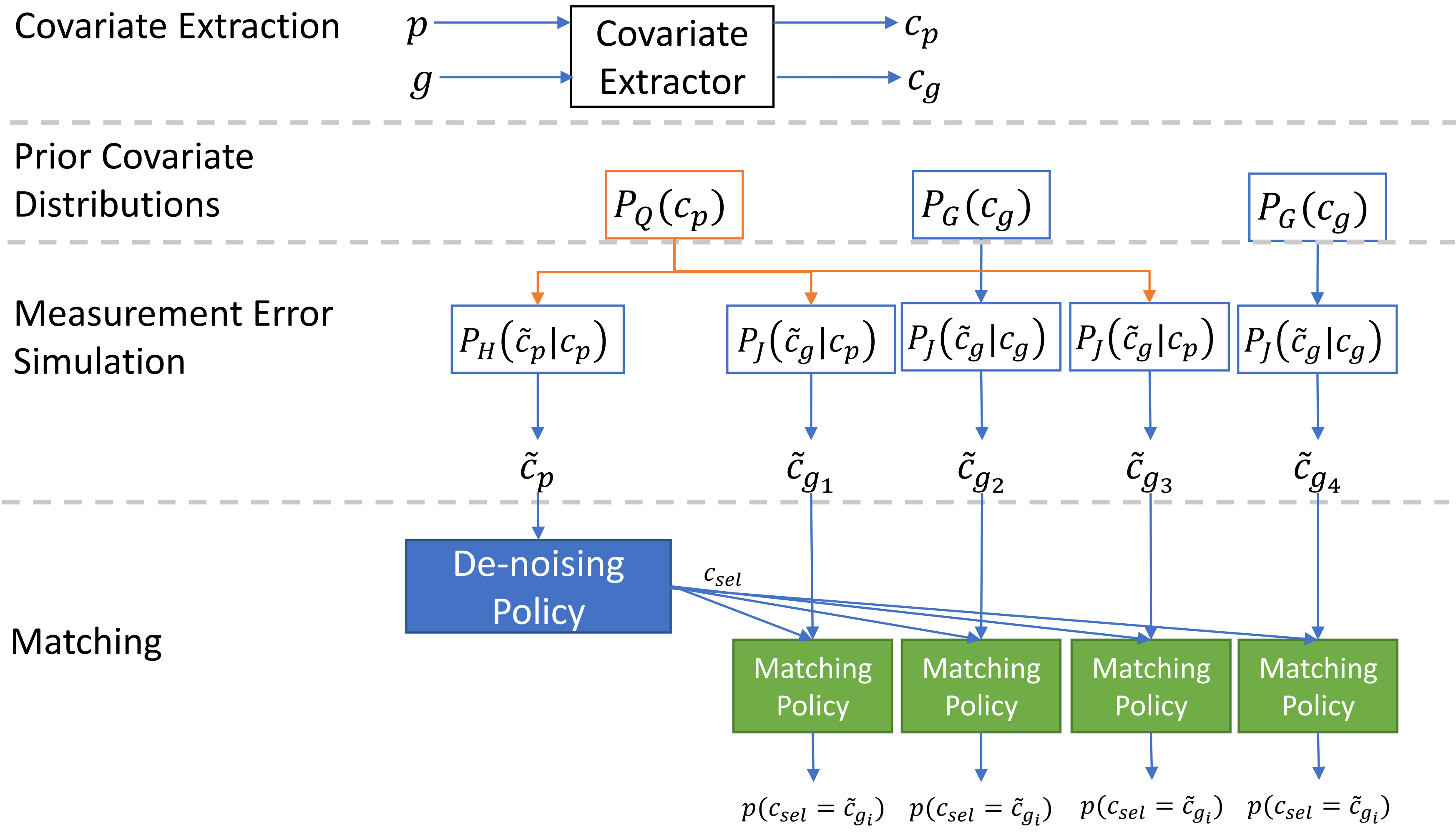
Verification

- Gallery contains 1 data sample
- Probes are the data claiming to be instances of the prototype
- Matching policy:
  $$P_{accept}(\tilde{c}_p, \tilde{c}_g)$$



Ranking

- Gallery contains N samples of which K match the probe
- Need to order gallery such that matching entries rank higher
- Policy:
  - Order by $p_{match}(c_{sel}, \tilde{c}_{g_i})$



Evaluation

- Datasets:
  - VoxCeleb:
    - 1251 celebrities
    - Covariates: Gender and Nationality
    - Avg. Covariate Collisions: 17.6
  - NIST SRE 2008
    - 15 speakers
    - Covariates: gender, age, native language, smoking habits
    - Avg. covariate collisions: 0.27

Results: Classification

- Classifying 8,250 samples of celebrity speech
- Classification Accuracy = 4.3%
- In presence of frequent collisions, the two covariates are not discriminative.
- Complex learning models are learning richer features!

Results: Pairwise-Verification

- 37,720 pairs of celebrities from VoxCeleb
- When covariates are noiseless EER is close to GMM-UBM (15.0)
- Adding some noise causes nominal deterioration
- *Suggests that complex models are not doing a lot more than predicting gender and nationality!*

$\boldsymbol{p_{corr}}$	$\boldsymbol{EER}$
0	15.9
0.1	18.8

Results: Ranking

- Ranking 248 data samples from 15 speakers
- Reference DNN model gets 100% Mean Average Precision
- *Clean covariates produce very close results*

p_{corr}	**MAP**
0	90.8
0.1	78.1
0.25	45.7

Conclusion

- We have empirically shown that in verification and ranking tasks optimally matching covariates can largely explain the performance of modern ML models.
- Based on these results we conclude that, covariate matching should be used as a standard baseline to gauge the marginal gains in performance offered by complex models.

Thank You

Questions?