

Scoring and Direct Methods for the Interpretation of Evidence in Forensic Speaker Recognition

Context

- Forensic Speaker Recognition
- Bayesian Interpretation of Evidence

Key Points

- Speech as Evidence
- The Strength of Evidence
- Likelihood Ratio (LR)
- Evaluation of the Strength of Evidence
- Distribution of Likelihood Ratios
- Tippett Plots

Speech as Evidence

Evidence (E) : In the case of questioned recording (trace), the evidence does not consist of the speech itself, but in the quantified degree of similarity (likelihood) between speaker dependent features extracted from the recorded speech of the suspect, represented by his/her model.

Strength of Evidence

Likelihood Ratio (LR) for two alternative hypotheses:

H_0 : the suspect is the source of the trace
 H_1 : the speaker at the origin of the trace is not the suspect

$$\frac{p(H_0|E)}{p(H_1|E)} = \frac{p(E|H_0)}{p(E|H_1)} \cdot \frac{p(H_0)}{p(H_1)}$$

a posteriori
probability ratio
Likelihood Ratio
(LR)
a priori
probability ratio

Calculation of Likelihood Ratio

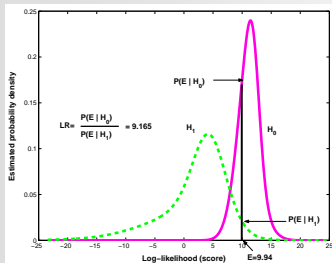
Two Methods

- Scoring Method
 - Direct Method
- Scoring Method** : Likelihood calculated from distribution of scores which model within-source and between-sources variability
- H_0 : distribution of scores of within-source variability
 - H_1 : distribution of scores of between-sources variability
- 3 databases :
- Suspect Reference Database (R)
 - Potential Population Database (P)
 - Suspect Control Database (C)
- Direct Method** : Likelihood directly calculated from GMM of suspect and GMMs of the potential population
- H_0 : GMMs of the suspect
 - H_1 : GMMs of the potential population
- 2 databases : Suspect Reference Database (R)
 Potential Population Database (P)

Databases Used:

- R= 5 utterances per speaker (2-3 min each)
- P = 100 speakers (2-3 min each)
- C = 30-40 utterances per speaker (10-20 sec each)

Scoring Method

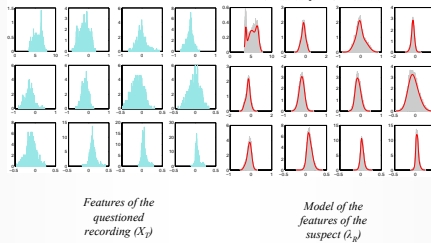


$$LR_{scoring} = \frac{p(E|H_0)}{p(E|H_1)}$$

$$= \frac{p(L(X_T|\lambda_{R_i})|H_0)}{p(L(X_T|\lambda_{P_j})|H_1)}$$

for $i = 1, \dots, N_R$

Direct Method



$$LR_{direct} = \frac{N_R \sqrt{\prod_{i=1}^{N_R} p(X_T|\lambda_{R_i})}}{N_P \sqrt{\prod_{j=1}^{N_P} p(X_T|\lambda_{P_j})}}$$

Experimental Framework

Significance of the Strength of Evidence

If the speaker of the trace is known, likelihood ratios can be calculated for mock cases where:

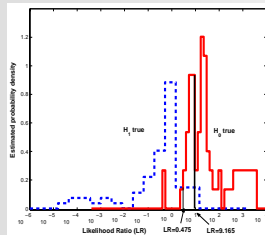
- H_0 is true
- H_1 is true

Experimental Framework:

- 60 cases in which H_0 is true
- 60 cases in which H_1 is true
- 60 traces from 15 male speakers (12-15 sec each)
- Potential population of 100 male speakers

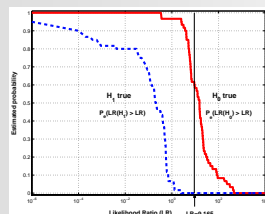
Scoring Method: Distributions of Likelihood Ratio

Distribution of likelihood ratios when the speaker of the trace is known.



Probability density Plot (LRs calculated by Scoring Method)

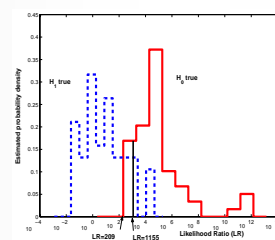
Probabilities of estimated LR when the hypothesis H_0 is true and when H_1 is true.



Tippett Plot (LRs calculated by Scoring Method)

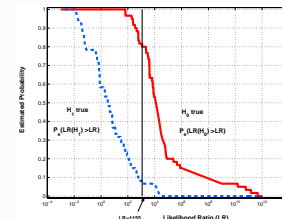
Direct Method: Distributions of Likelihood Ratio

Distribution of likelihood ratios when the speaker of the trace is known.



Probability density Plot (LRs calculated by Direct Method)

Probabilities of estimated LR when the hypothesis H_0 is true and when H_1 is true.



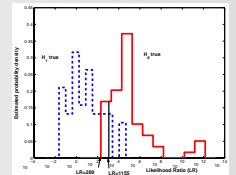
Tippett Plot (LRs calculated by Direct Method)

Significance Probabilities of Strength of Evidence

Significance Probability: A measure of compatibility of the data with the hypothesis.

$$Z_{H_0} = \frac{LR_E - \mu_{LR_{H_0}}}{\sigma_{LR_{H_0}}}$$

$$Z_{H_1} = \frac{LR_E - \mu_{LR_{H_1}}}{\sigma_{LR_{H_1}}}$$



Under the assumption of normality, the values of Z_{H_0} and Z_{H_1} correspond to the probabilities in the Tippett plots for the values LR(E)

Scoring Method:

LR(E) corresponds to 75 % and 7.08 % significance probability for the hypotheses H_0 and H_1

Direct Method:

LR(E) corresponds to 62.2 % and 5.4 % significance probability for the hypotheses H_0 and H_1

Conclusions

1. The Bayes' model, current interpretation framework used in forensic science, is adapted for forensic automatic speaker recognition.
2. Two methods, i.e., the *Scoring Method* and *Direct Method*, can be applied for calculating likelihood ratio as the value of the strength of evidence when using Gaussian Mixture Models of the speakers.
3. Distributions of likelihoods when H_0 and H_1 are true and Tippett Plots can be used for evaluation of the strength of evidence in forensic speaker recognition.