The VoicePrivacy 2020 Challenge

Odyssey 2020

Objective evaluation- Linkability

Presenter: Mohamed Maouche [2]

Natalia Tomashenko [1]
Brij M.L. Srivastava [2]
Xin Wang [3]
Emmanuel Vincent [2]
Andreas Nautsch [4]
Junichi Yamagishi [3,5]
Nicholas Evans [4]
Jose Patino [4]
Jean-François Bonastre [1]
Paul-Gauthier Noé [1]
Massimiliano Todisco [4]
Mohamed Maouche [2]
Benjamin O'Brien [6]
Anais Chanclu [1]

[1] LIA – University of Avignon – France
[2] Inria – France
[3] NII – Tokyo – Japan
[4] Audio Security and Privacy Group, EURECOM – France
[5] University of Edinburgh – UK
[6] Aix-Marseille University – France
   Laboratoire Parole et Langage (LPL)

4th November 2020

Linkability: Definition

First introduced for biometric template protection

evaluation in [Gomez-Barrero et al. 2017].

Input: scores S of the ASV

Two types of trials: H mated and $\bar{H}$ non-mated

The goal is to measure:

$$\text{Linkability}(s) = p(H \mid s) - p(\bar{H} \mid s)$$

Linkability: Definition

$$\text{Linkability}(s) = p(H \mid s) - p(\overline{H} \mid s)$$
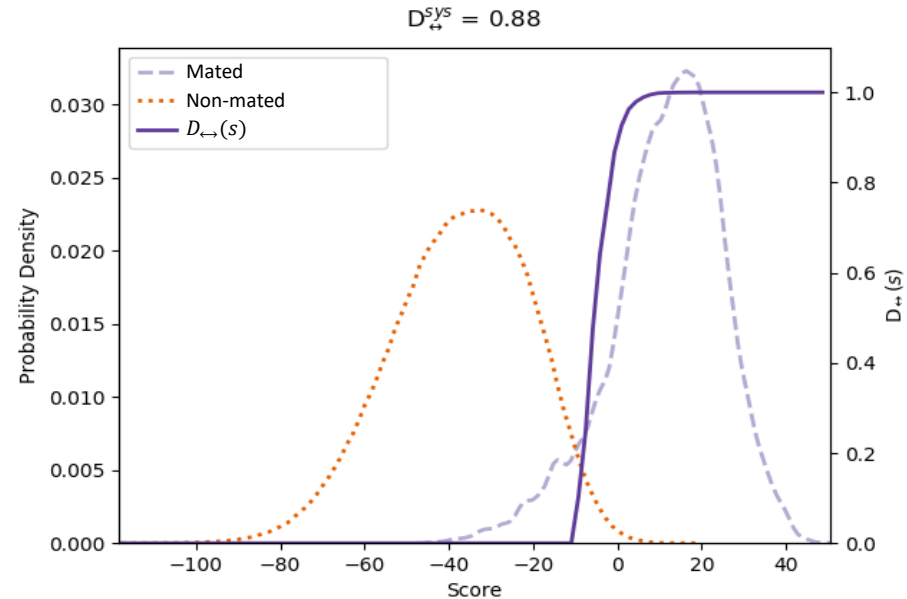
Local measure of linkability

After some transformation with the likelihood ratio

$LR(s) = \dfrac{p(s|H)}{p(s|\overline{H})}$ and the prior ratio $\omega = \dfrac{p(H)}{p(\overline{H})}$

$$D_{\leftrightarrow}(s) = \begin{cases} \dfrac{2 \cdot \omega \cdot LR(s)}{1 + \omega \cdot LR(s)} - 1 & if \ \omega \cdot LR(s) > 1 \\ 0 & otherwise \end{cases}$$

Global measure of linkability

$$D_{\leftrightarrow}^{sys} = \int p(s \mid H) D_{\leftrightarrow}(s) \, ds$$

Linkability: Advantages and Disadvantages
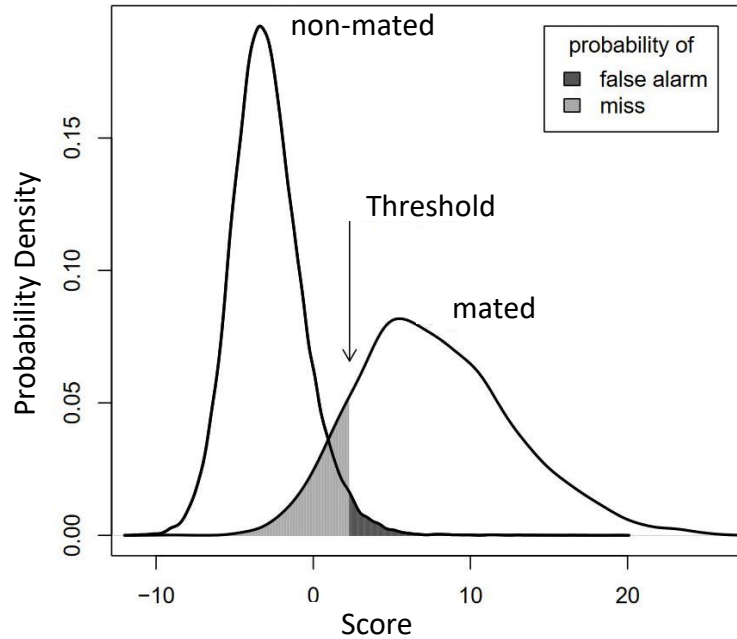
Advantages

- No notion of threshold or attacker used.

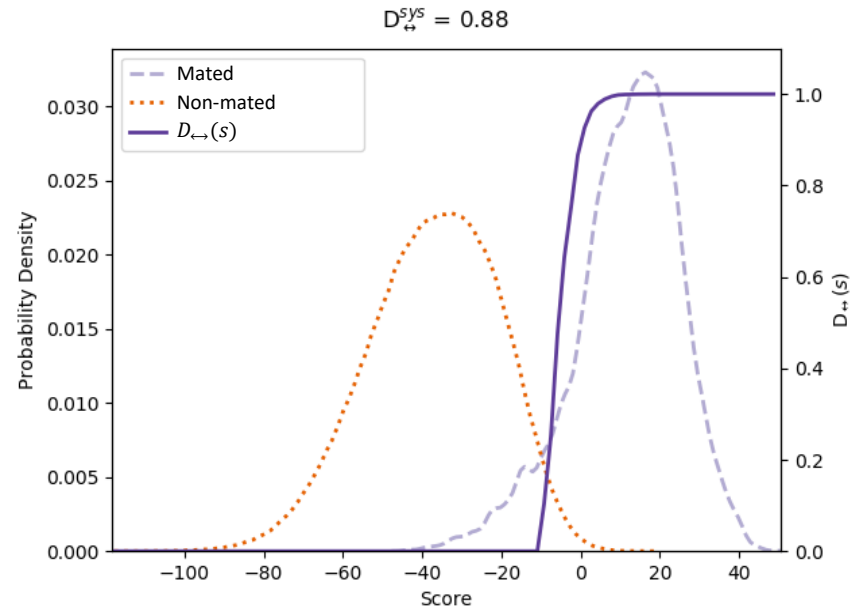- Any local separation between mated and non-mated scores can be detected.

Disadvantages

- The usage of bins to estimate the probabilities is not exact.

- Focus only on the mated cases, the strength of non-mated evidence is ignored.

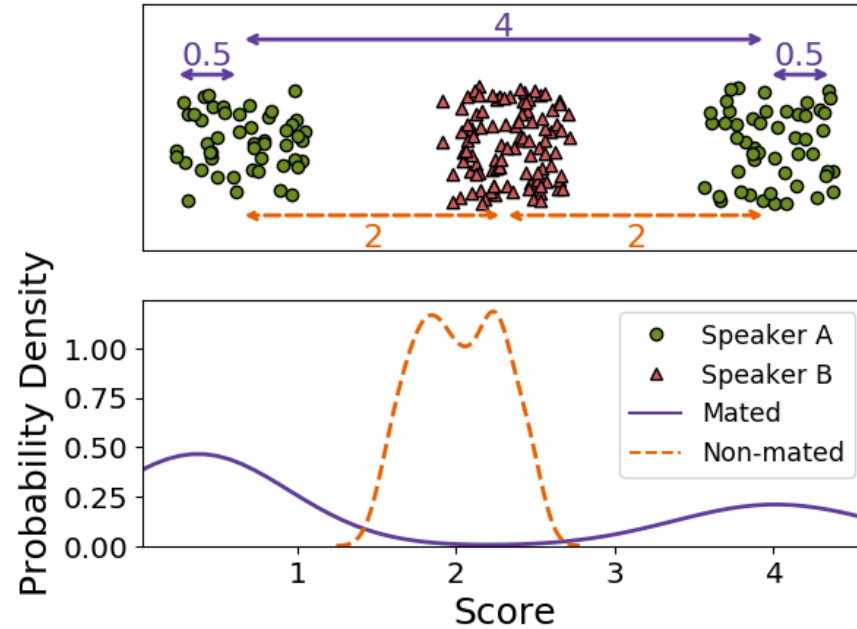# Linkability: Advantages and Disadvantages

Advantages

- No notion of threshold or attacker used.

- Any local separation between mated and non-mated scores can be detected.

Disadvantages

- The usage of bins to estimate the probabilities is not exact.

- Focus only on the mated cases, the strength of non-mated evidence is ignored.

Linkability: No threshold

[DA Van Leeuwen & N Brümmer Speaker classification 2007]

Linkability: Advantages and Disadvantages

Advantages

- No notion of threshold or attacker used.

- Any local separation between mated and non-mated scores can be detected.

Disadvantages

- The usage of bins to estimate the probabilities is not exact.

- Focus only on the mated cases, the strength of non-mated evidence is ignored.

Linkability: example of complex relation

$$D_{\leftrightarrow}^{sys} = 0.99 \text{ (low privacy)}$$
$$C_{llr}^{min} = 0.81 \text{ (high privacy)}$$

Linkability: estimation of LRs

For C_{llr}, $LLRs$ are the scores of the ASV.
For C_{llr}^{min}, $LLRs$ are the calibrated scores of the ASV (PAV algorithm [Brummer et al. CSL'06]).

For $D_{\leftrightarrow}^{sys}$: the local conditional probability are estimated using Histograms

$$LR(s) = \frac{p(s \mid H)}{p(s \mid \overline{\overline{H}})}$$ ← ——— Histogram of mated scores

← ——— Histogram of non-mated scores

⚠ Bins are used to estimate the local probabilities

Linkability: Advantages and Disadvantages

Advantages

- No notion of threshold or attacker used.

- Any local separation between mated and non-mated scores can be detected.

Disadvantages

- The usage of bins to estimate the probabilities is not exact.

- Focus only on the mated cases, the strength of non-mated evidence is ignored.

Linkability: Advantages and Disadvantages

Advantages

- No notion of threshold or attacker used.

- Any local separation between mated and non-mated scores can be detected.

Disadvantages

- The usage of bins to estimate the probabilities is not exact.

- Focus only on the mated cases, the strength of non-mated evidence is ignored.

Linkability: only mated evidence

$$D_{\leftrightarrow}(s) = \begin{cases} \dfrac{2 \cdot \omega \cdot LR(s)}{1 + \omega \cdot LR(s)} - 1 & if \boxed{\omega \cdot LR(s) > 1} \\ 0 & otherwise \end{cases}$$

$$p(H \mid s) > p(\overline{H} \mid s)$$

$$D_{\leftrightarrow}^{sys} = \int p(s \mid H) D_{\leftrightarrow}(s)\, ds$$

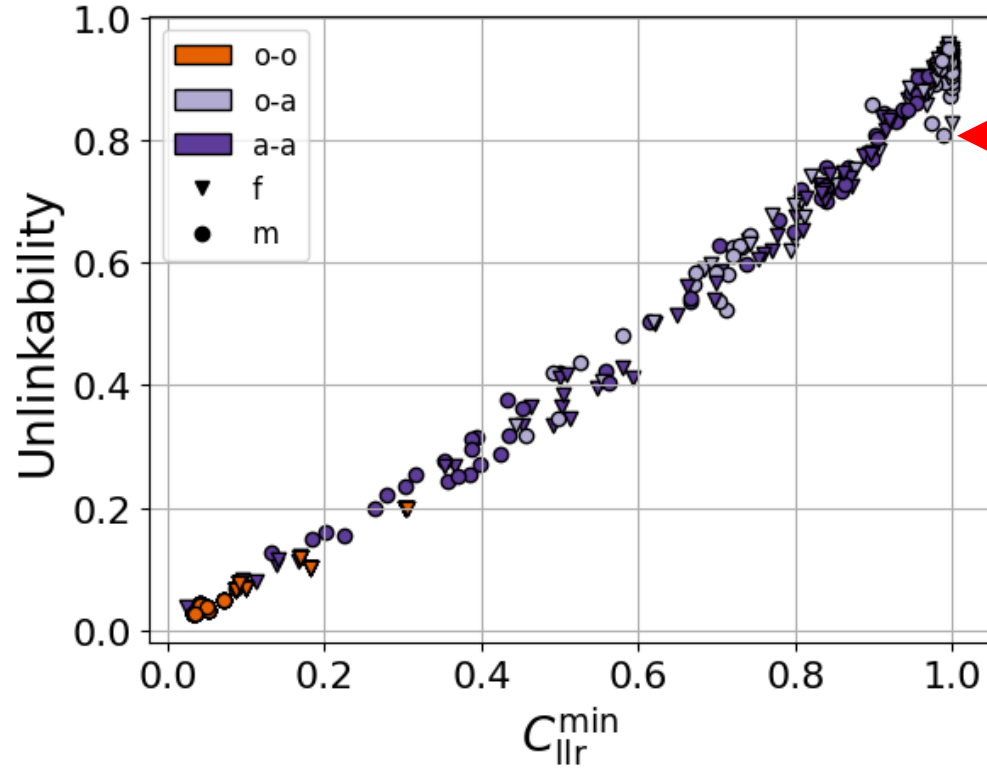# Objective evaluation
Linkability
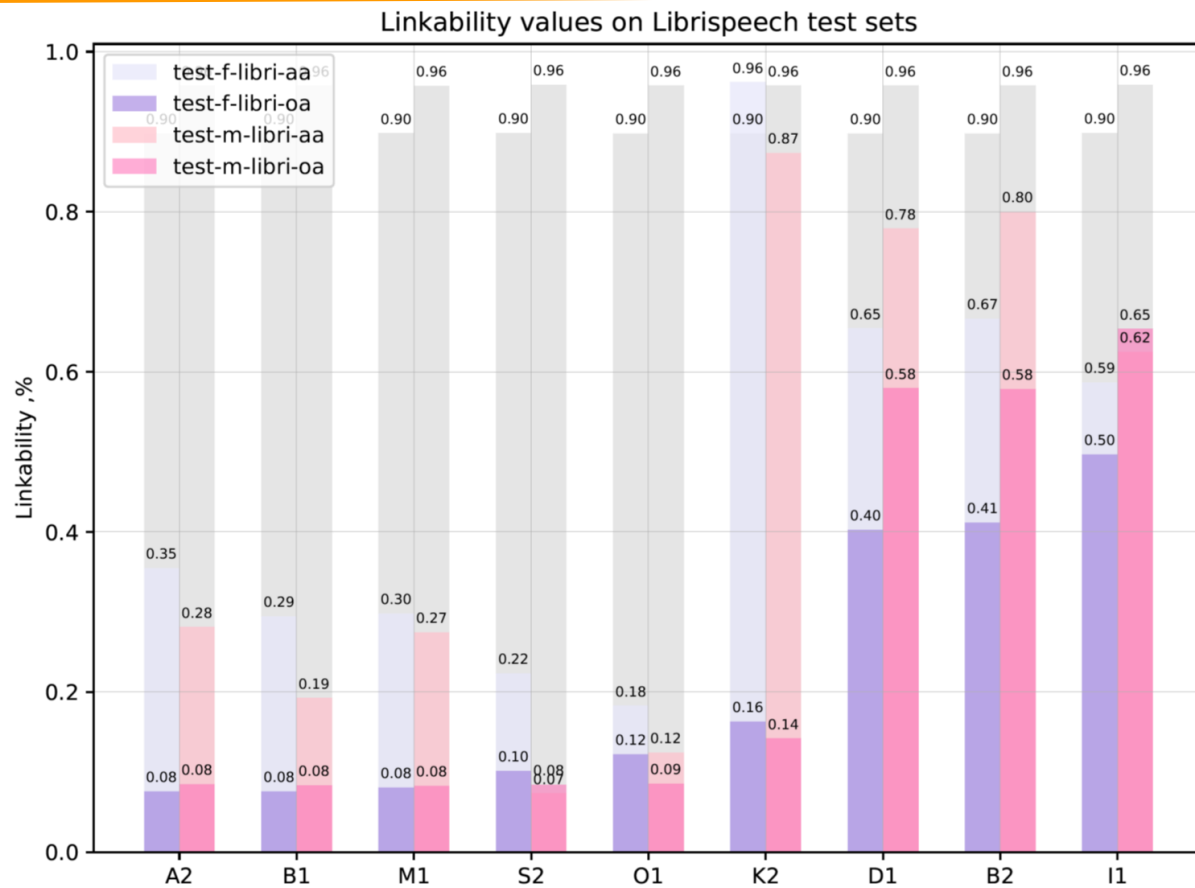
Sorted by oa test

Linkability and EER

EER vs Linkability (means over all VoicePrivacy datasets)
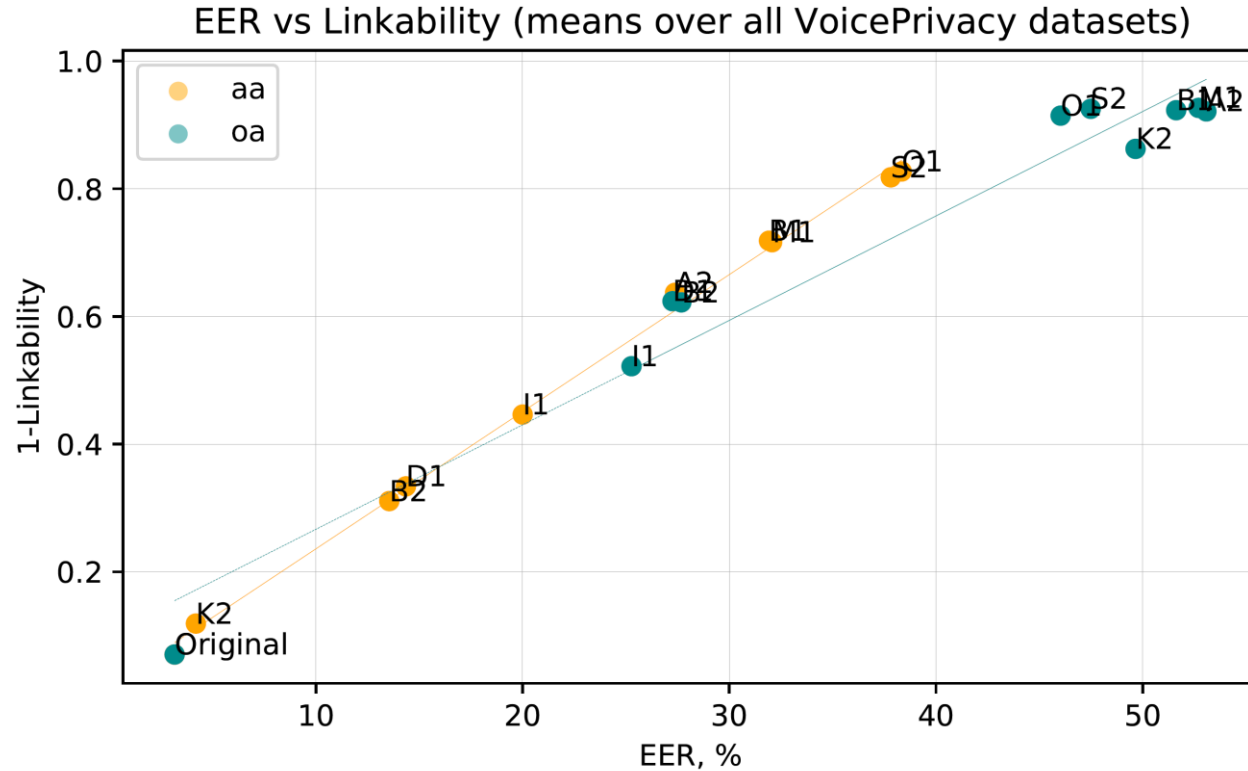
Linkability and C_{llr}^{min}

One datapoint per dataset, attacker and system
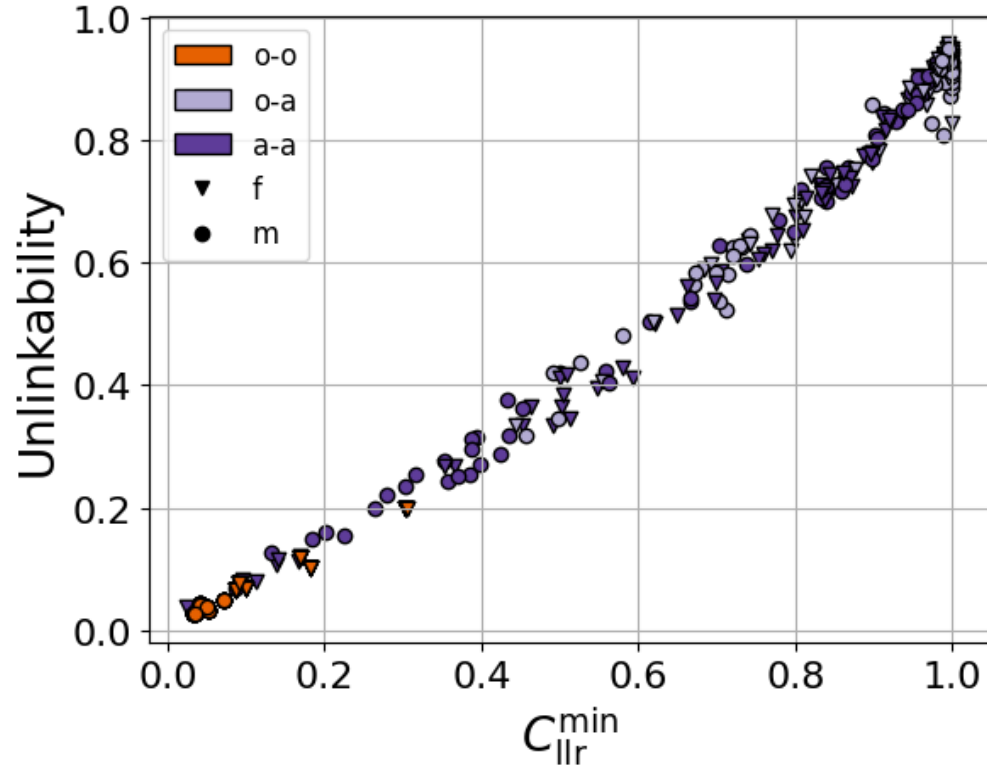
Linkability and C_{llr}^{min}

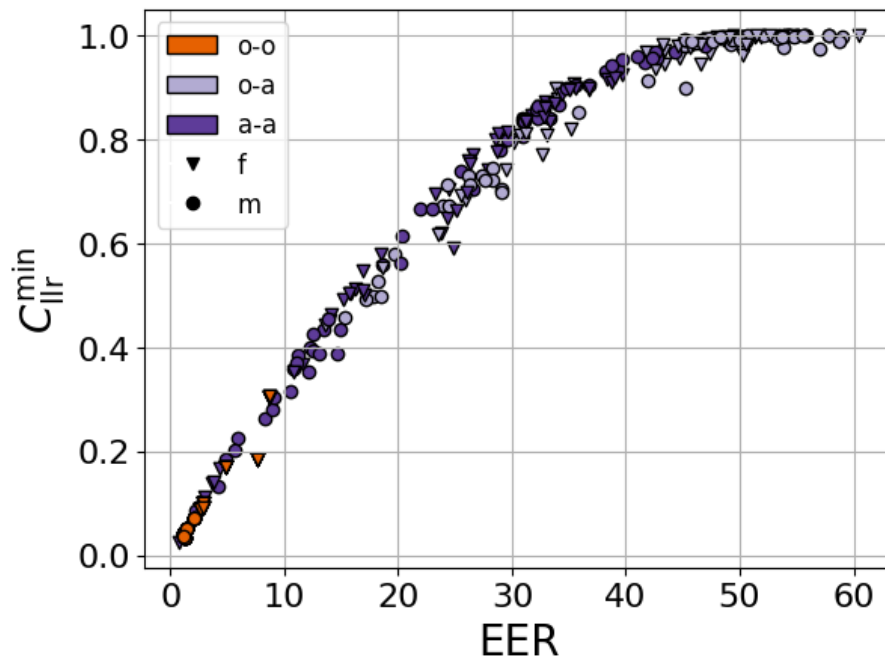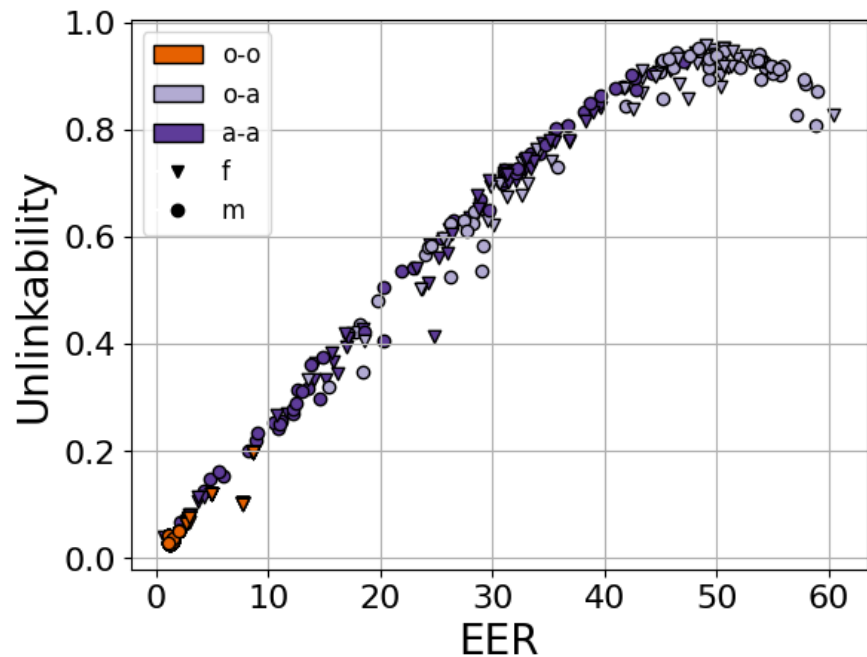One datapoint per dataset, attacker and system
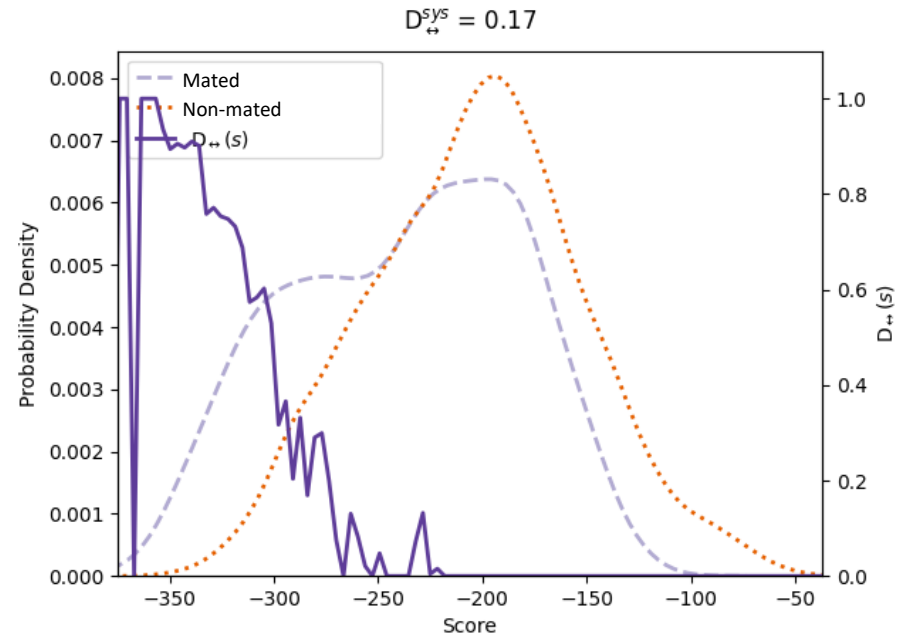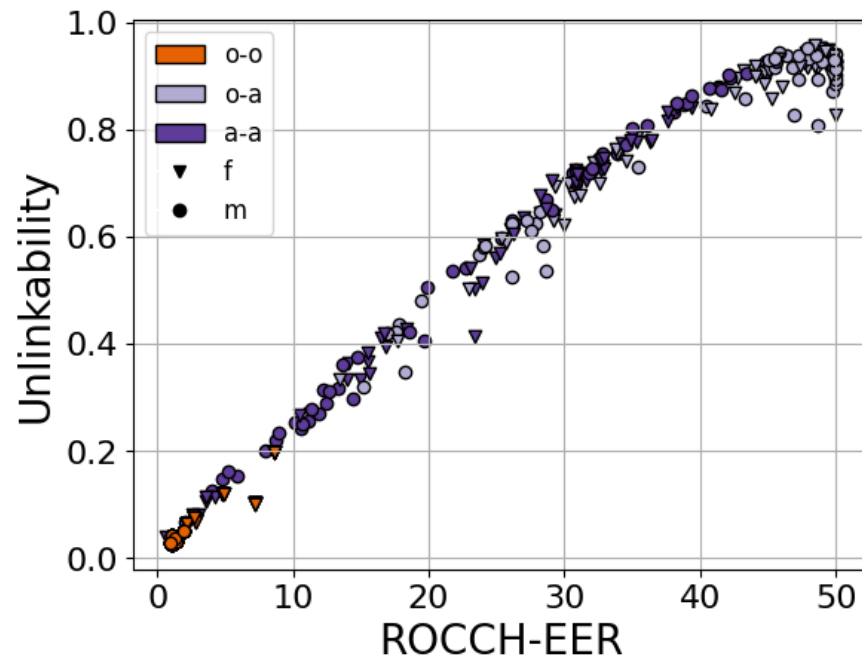
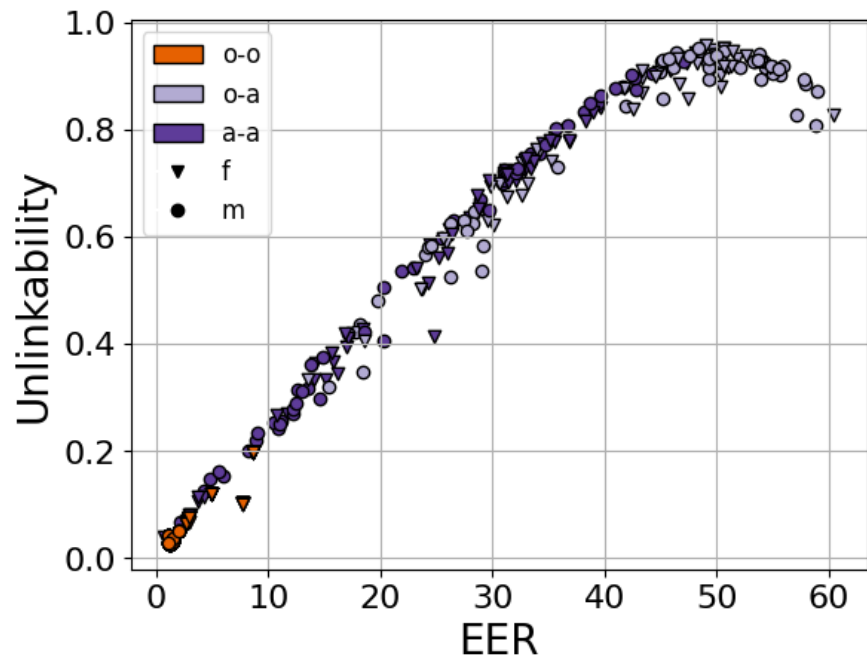# Interesting case

Vctk-test-f (o-a)

Linkability vs EER

One datapoint per dataset, attacker and system
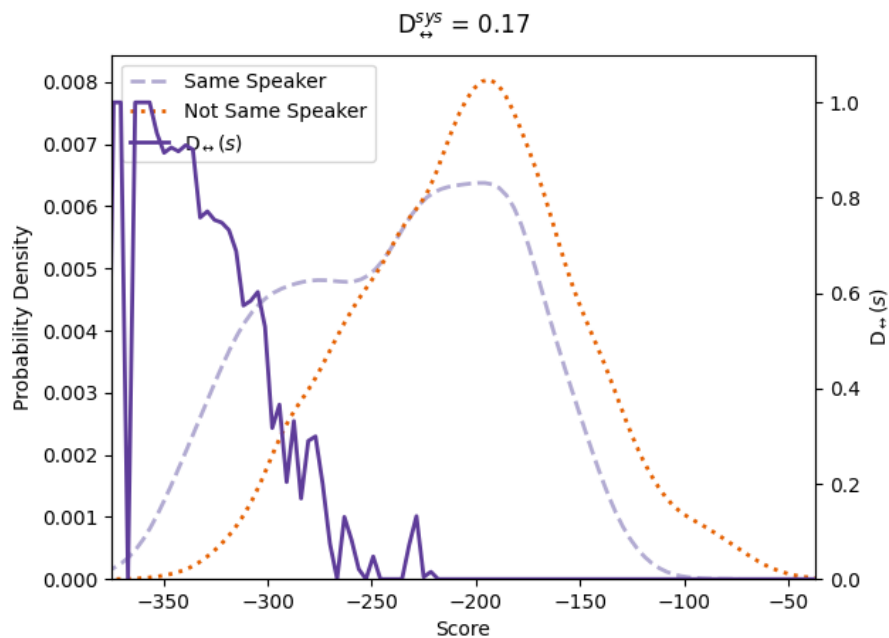
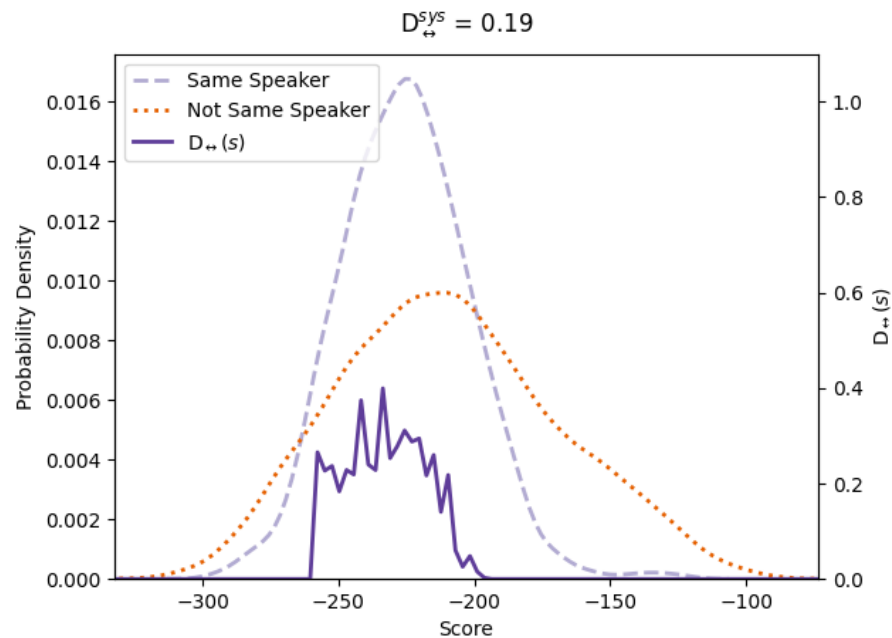# Linkability: EER vs ROCCH-EER

One datapoint per dataset, attacker and system

Interesting cases

Vctk-test-f (o-a)

Vctk-test-m (o-a)

Linkability vs Zebra

One datapoint per dataset, attacker and system