Objective evaluation - ZEBRA

Presenter: Andreas Nautsch

Natalia Tomashenko 1
Brij M.L. Srivastava 2
Xin Wang 3
Emmanuel Vincent 4
Andreas Nautsch 5
Junichi Yamagishi 3,6
Nicholas Evans 5
Jose Patino 5
Jean-François Bonastre 1
Paul-Gauthier Noé 1
Massimiliano Todisco 5
Mohamed Maouche 2
Benjamin O'Brien 7
Anais Chanclu 1

1 LIA – University of Avignon – France
2 Inria – France
3 NII – Tokyo – Japan
4 Inria – France
5 Audio Security and Privacy Group, EURECOM – France
6 University of Edinburgh – UK
7 Aix-Marseille University – France

4th November 2020
Recap: EER, min Cllr & ZEBRA’s “worst case”

Ideal score calibration:
1. identify bins with same ratios of errors
2. map scores to unified scale

EER: one threshold, varies on “unified scale”
min Cllr: generalised class discrimination

ZEBRA worst-case: max(abs(LLR))
Recap: min Cllr & ZEBRA’s “expectation”

If perfect privacy, then must:
- min Cllr is upper bound
- profile is symmetric

Empirical observations:
- min Cllr often close to upper bound
- lots of symmetries

Prior: chosen by adversary
⇒ inaccessible to us, but we can simulate :)

\[
\frac{\pi}{|S_A|} \sum_{a \in S_A} \log_2 \left( 1 + \frac{1 - \pi}{\alpha \pi} \right) + \frac{1 - \pi}{|S_B|} \sum_{b \in S_B} \log_2 \left( 1 + \frac{b \pi}{1 - \pi} \right)
\]

Numerical equivalent to
\[
\frac{1}{2}
\]

ECE (in bit)

Prior probability

(a) ZEBRA idea

(b) ECE plot
ZEBA: by system for all VoicePrivacy test data sets (primary)

Categorical tags of worst-case privacy disclosure

<table>
<thead>
<tr>
<th>Tag</th>
<th>Category</th>
<th>Posterior odds ratio (flat prior)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1 = 10^6</td>
<td>50% : 50% (flat posterior)</td>
</tr>
<tr>
<td>A</td>
<td>10^6 ≤ t ≤ 10^7</td>
<td>more disclosure than 50% : 50%</td>
</tr>
<tr>
<td>B</td>
<td>10^7 ≤ t ≤ 10^8</td>
<td>one wrong in 10 to 100</td>
</tr>
<tr>
<td>C</td>
<td>10^8 ≤ t ≤ 10^9</td>
<td>one wrong in 100 to 10000</td>
</tr>
<tr>
<td>D</td>
<td>10^9 ≤ t ≤ 10^10</td>
<td>one wrong in 10000 to 100000</td>
</tr>
<tr>
<td>E</td>
<td>10^10 ≤ t ≤ 10^{11}</td>
<td>one wrong in at least 1000000</td>
</tr>
</tbody>
</table>
ZEBRA: D_ECE (population) on VCTK-test
ZEBRA: worst-case privacy disclosure (individual) on VCTK-test
Metric correlation — NOT CAUSALITY

ZEBRA Population [bit] vs. ROCCH-EER [%]

ZEBRA Population [bit] vs. min Cllr

ZEBRA expected disclosure [bit] vs. worst-case disclosure

Realm of privacy preservation

Realm of the adversary

Safeguard Protected data -> Classifier Scores Decision policy Action

Error trade-offs Inference information Impact

Safety evaluation Privacy evaluation Security evaluation
ZEBRA: D_ECE (population) on VCTK-test

Sorted by oa f-test-vctk

Sorted by aa f-test-vctk
ZEBRA: worst-case privacy disclosure (individual) on VCTK-test

Sorted by oa f-test-vctk

Sorted by aa f-test-vctk
VoicePrivacy

ZEBRA: LibriSpeech (all systems)
ZEBRA: LibriSpeech (all systems) ordered by performance

VoicePrivacy
ZEBRA: VCTK-different (all systems)
ZEBrA: VCTK-different (all systems) ordered by performance
ZEGBRA: VCTK-common (all systems)
VOICE PRIVACY

ZEBRA: VCTK-common (all systems) ordered by performance

16