Redesign of Gaussian Mixture Model for Efficient and Privacy-preserving Speaker Recognition

S Rahulamathavan, X Yao, R Yogachandran, K Cumanan, and M Rajarajan

CyberSA 2018
Glasgow
Motivation

Speech is unique like fingerprint

Revocable

Privacy
Privacy-preserving Solution

- Encrypted Data
- Randomised Domain
- Redesign
  - SOTA Algorithms to work On ED
- User Device (public-key, private-key)
- Data Stored at Cloud
Speaker Model

Parameters used for speaker recognition

\[ \lambda = \{ p_i, \mu_i, \Sigma_i \} \quad i = 1, \ldots, M. \]

Scalar, vector, Matrix

Verification

\[ p(\bar{x} | \lambda) = \sum_{i=1}^{M} p_i b_i(\bar{x}) \]

\[ b_i(\bar{x}) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\bar{x} - \mu_i)' \Sigma_i^{-1} (\bar{x} - \mu_i) \right\} \]
Enrollment protocol: user has enrollment data $x$ and system has the UBM $\lambda_U$. System obtains encrypted speaker model $E[\lambda_s^{(1)}]$. 

**User Enrolment**
Verification protocol: user has test data $x$ and system has the UBM $\lambda_U$ and encrypted speaker model $E[\lambda_s^{(1)}]$. The user submits encrypted data and the system outputs an accept/reject decision.
What is randomization? Example

### Table V
**RANDOMISATION EXAMPLES**

<table>
<thead>
<tr>
<th>Test speech samples</th>
<th>Corresponding random values</th>
<th>Randomised speech features</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{j1}$, $t_{j2}$, ...</td>
<td>$r_{j1}$, $r_{j2}$, ...</td>
<td>$[t_{j1}], [t_{j2}]$, ...</td>
</tr>
<tr>
<td>16.725081716161</td>
<td>420223.336542782260</td>
<td>420240.061624498421</td>
</tr>
<tr>
<td>0.0001274641</td>
<td>225758.663235026265</td>
<td>225758.663362490365</td>
</tr>
<tr>
<td>1.171659034624</td>
<td>206555.735876851157</td>
<td>206556.907535885781</td>
</tr>
<tr>
<td>46.027486334736</td>
<td>236074.104503441653</td>
<td>236120.131989776389</td>
</tr>
<tr>
<td>1910.031836695921</td>
<td>125628.018508620454</td>
<td>127538.050345316375</td>
</tr>
</tbody>
</table>
Enrolment & Authentication

1. TLS Connection
   Initial Authentication

2. Authentication Request
   + Challenge + Policy

3. Biometric Raw data

4. Feature extraction

5. Cancellable Bio Process

6. Authentication Response

7. Retrieve Encrypted biometric data

8. Crypto Matching

9. Validate and response
\( \theta = 0.4 \)

Figure 2. Combination of genuine and malicious authentication attempts. The distribution on the right hand side shows the outcome for 156 legitimate authentication attempts in numbers. The distribution on the left hand side shows the outcome for 17356 malicious attempts (in percentages).

- False Negative rate = \( \frac{6}{156} \times 100 = 3.85\% \);
- False Positive rate = \( \frac{150}{17356} \times 100 = 1.77\% \)

**Table I**

Pre-divided sets in TIMIT data corpus. The numbers in the brackets in Set 1 is randomly selected for building speaker models.

<table>
<thead>
<tr>
<th></th>
<th>Set 1 (Subset)</th>
<th>Set 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR1</td>
<td>38 (15)</td>
<td>11</td>
<td>49</td>
</tr>
<tr>
<td>DR2</td>
<td>76 (58)</td>
<td>26</td>
<td>102</td>
</tr>
<tr>
<td>DR3</td>
<td>76 (43)</td>
<td>26</td>
<td>102</td>
</tr>
<tr>
<td>DR4</td>
<td>68 (40)</td>
<td>32</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td><strong>258 (156)</strong></td>
<td><strong>95</strong></td>
<td><strong>353</strong></td>
</tr>
</tbody>
</table>
Figure 7. Android Monitor reading of smartphone memory and CPU usage during the enrolment and authentication phase of the proposed scheme.
Thank you!

www.DrRahul.uk