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

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

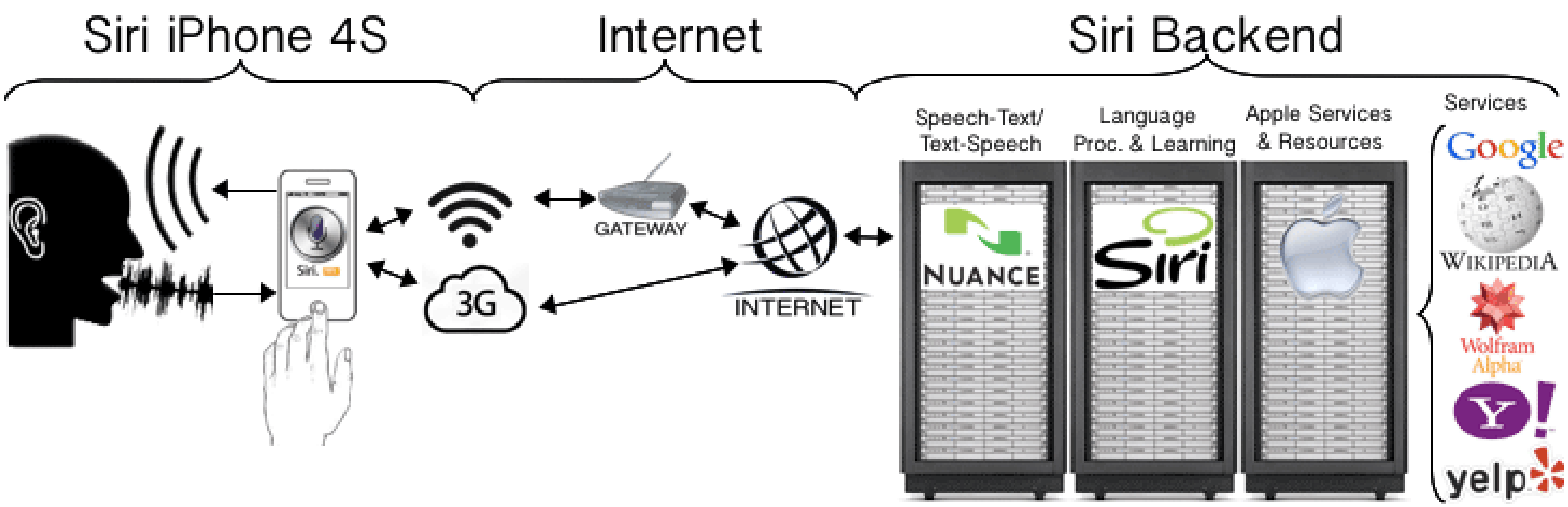
**CyberSA 2018**
**Glasgow**

# Motivation

Speech is unique like fingerprint

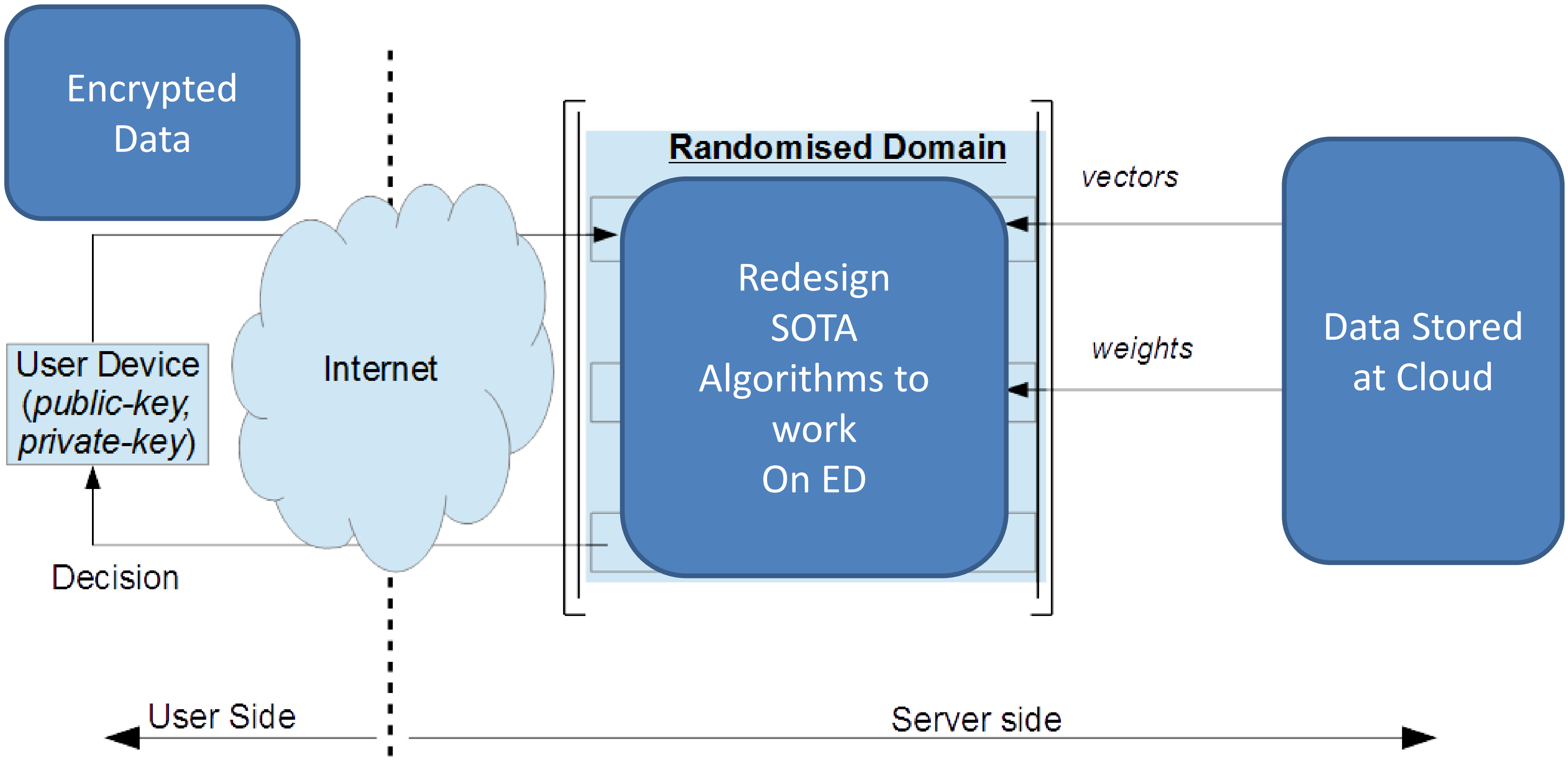
Revocable

Privacy



Siri iPhone 4S
Internet
Siri Backend
GATEWAY
3G
INTERNET
Speech-Text/ Text-Speech
Language Proc. & Learning
Apple Services & Resources
Services
NUANCE
Siri
Google
WIKIPEDIA
Wolfram Alpha
Y!
yelp

# Privacy-preserving Solution



# Speaker Model

Parameters used for speaker recognition

$$\boxed{\lambda = \{p_i, \vec{\mu}_i, \Sigma_i\} \quad i = 1, \ldots, M.}$$

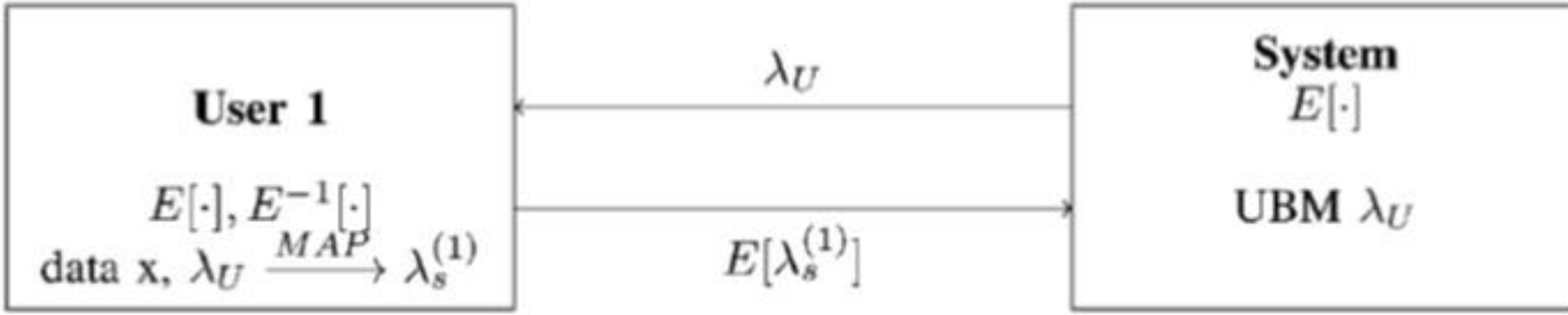
Scalar, vector, Matrix

Verification

$$\boxed{p(\vec{x} \mid \lambda) = \sum_{i=1}^{M} p_i b_i(\vec{x})}$$

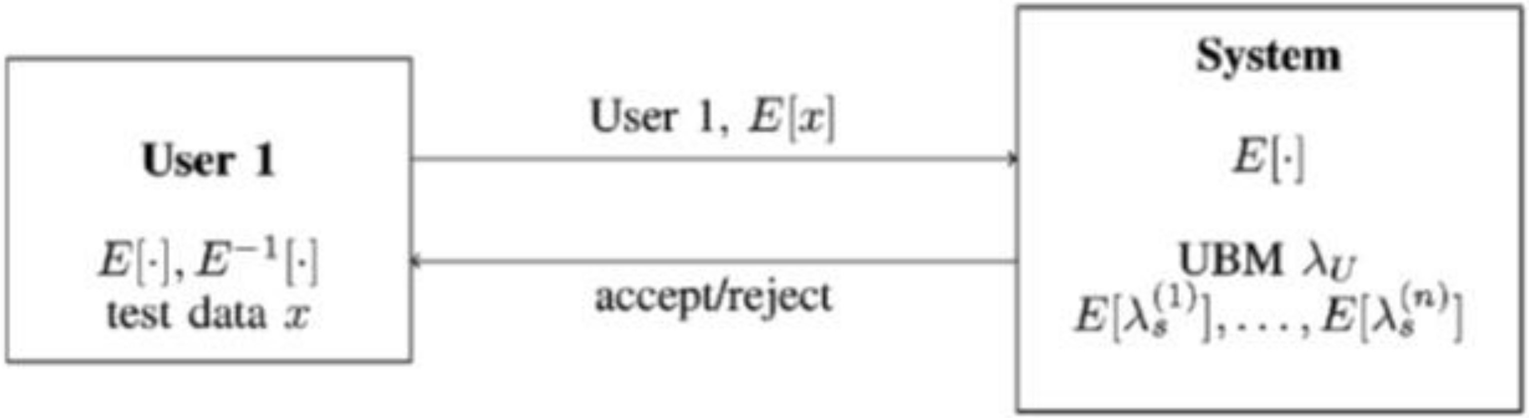
$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\vec{x} - \vec{\mu}_i)' \, \Sigma_i^{-1} \, (\vec{x} - \vec{\mu}_i)\right\}$$

(2




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.

# Verification

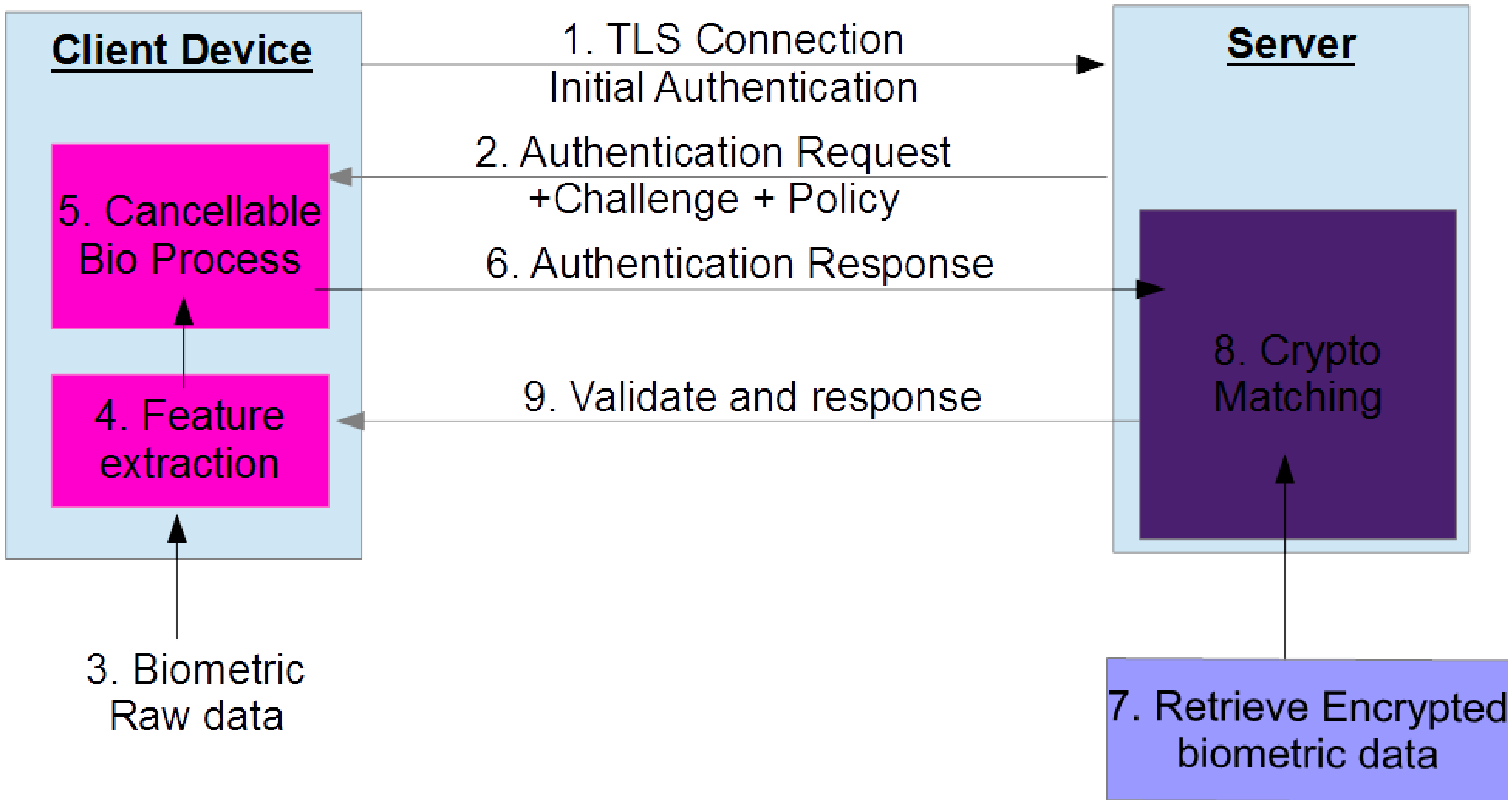
# What is randomization? Example

Table V
RANDOMISATION EXAMPLES

| Test speech samples $t_{j_1}, t_{j_2}, \ldots$ | Corresponding random values $r_{j_1}, r_{j_2}, \ldots$ | Randomised speech features $[t_{j_1}], [t_{j_2}], \ldots$ |
|---|---|---|
| 16.725081716161 | 420223.336542782260 | 420240.061624498421 |
| 0.0001274641 | 225758.663235026265 | 225758.663362490365 |
| 1.171659034624 | 206555.735876851157 | 206556.907535885781 |
| 46.027486334736 | 236074.104503441653 | 236120.131989776389 |
| 1910.031836695921 | 125628.018508620454 | 127538.050345316375 |

Client side

Client send/store this at server side




# Enrolment & Authentication

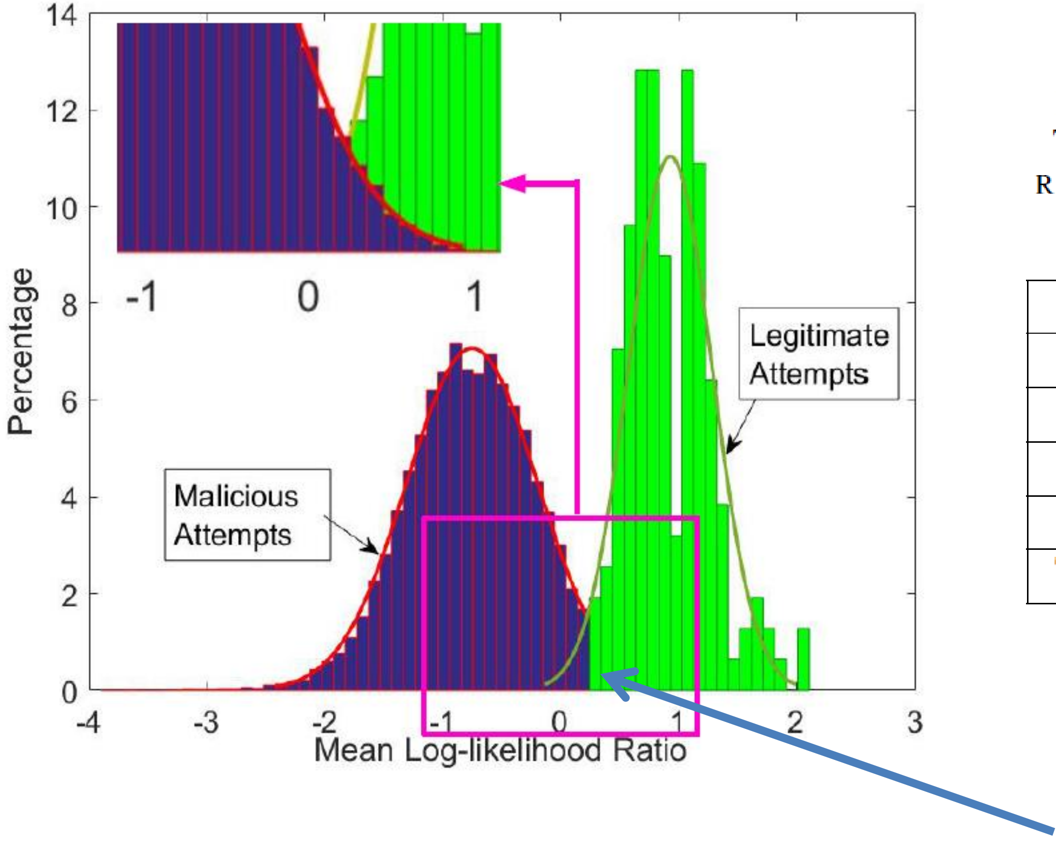



Figure 2. Combination of genuine and malicious authentication attempts. The distribution on the right hand side shows the outcome for 156 legitimate authentication attempts in numbers. The distribution on the left hand side shows the outcome for 17356 malicious attempts (in percentages).

Table I
PRE-DIVIDED SETS IN TIMIT DATA CORPUS. THE NUMBERS IN THE BRACKETS IN SET 1 IS RANDOMLY SELECTED FOR BUILDING SPEAKER MODELS.

| | Set 1 (Subset) | Set 2 | Total |
|---|---|---|---|
| DR1 | 38 (15) | 11 | **49** |
| DR2 | 76 (58) | 26 | **102** |
| DR3 | 76 (43) | 26 | **102** |
| DR4 | 68 (40) | 32 | **100** |
| Total | **258 (156)** | **95** | **353** |

- False Negative rate= $\frac{6}{156} \times 100 = 3.85\%$;
- False Positive rate = $\frac{308}{17356} \times 100 = 1.77\%$

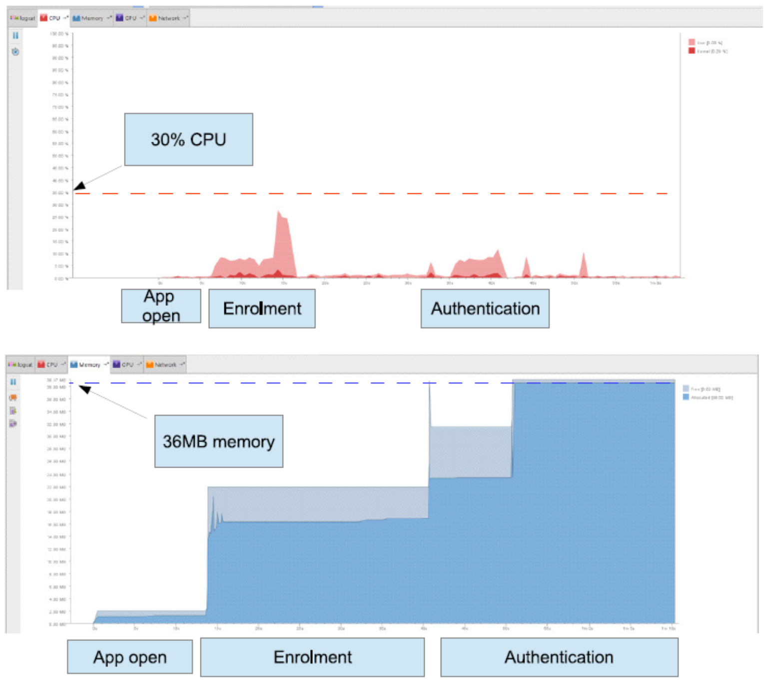



Figure 7. Android Monitor reading of smartphone memory and CPU usage during the enrolment and authentication phase of the proposed scheme.

# Experiment: Results

# Thank you!