UNIVERSITÄT BERN

IWFHR 2006, LA BAULE

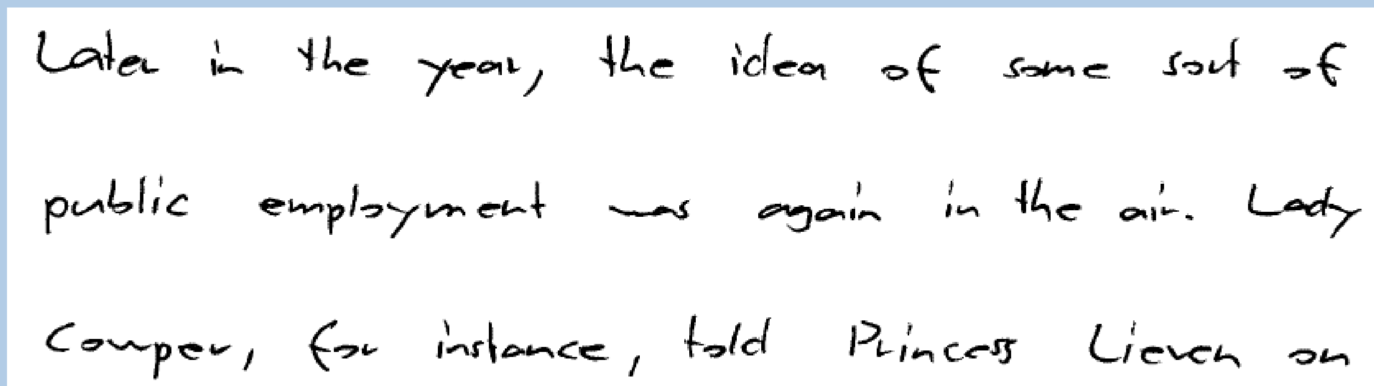
OFF-LINE WRITER VERIFICATION: A COMPARISON OF A HIDDEN MARKOV MODEL (HMM) AND A GAUSSIAN MIXTURE MODEL (GMM) BASED SYSTEM

Andreas Schlapbach and Horst Bunke
Institute of Computer Science and Applied Mathematics
University of Bern, Switzerland
{schlpbch, bunke}@iam.unibe.ch

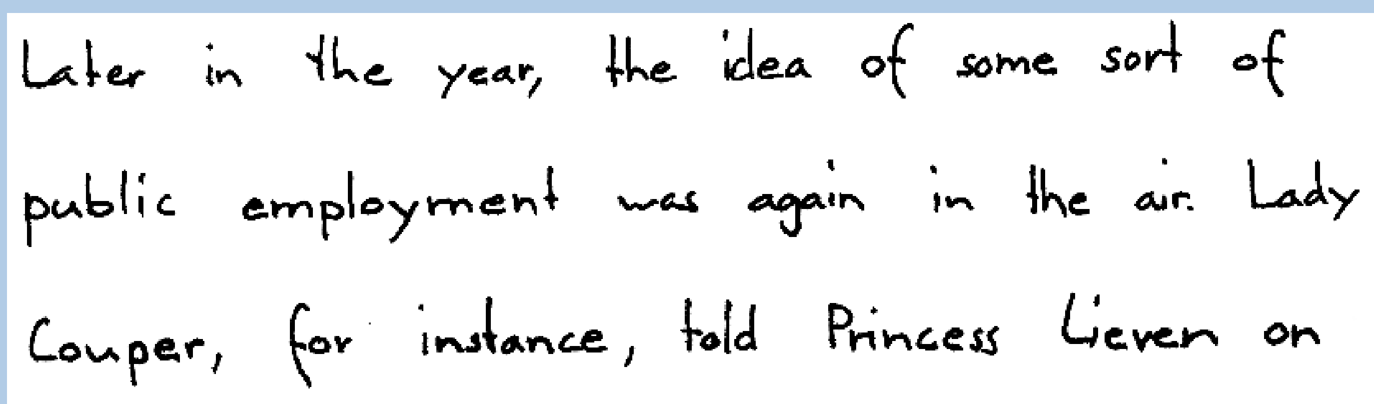
Contents

UNIVERSITÄT BERN

1. Original and Forged Text Lines



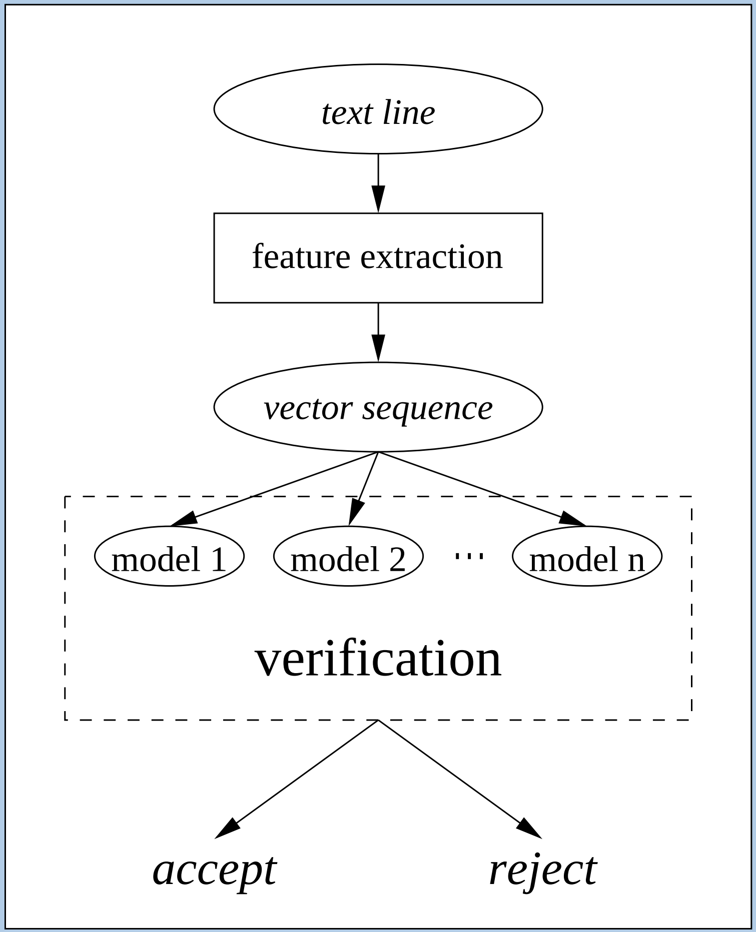
Original text lines



Skillfully forged text lines

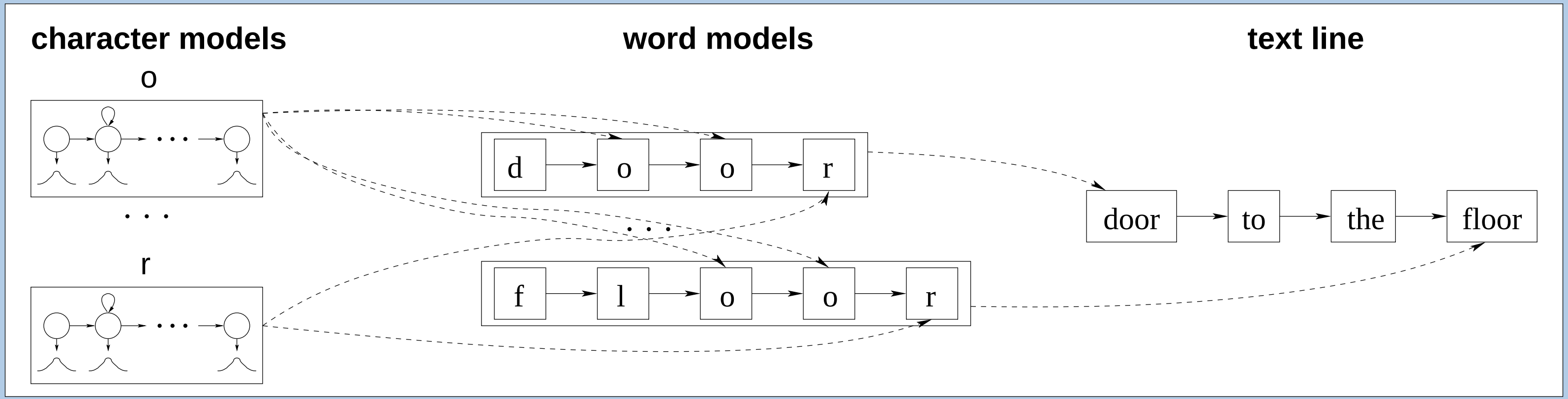
1. System Overview

- A moving sliding window extracts a sequence of feature vectors from each text line.
- For each writer, a model is build and trained with data coming from that writer only.
- A text line of an unknown writer is presented to the respective models.
- Based on the returned log-likelihood scores the decision is made to accept or reject a text line under question.



2. A HMM based System

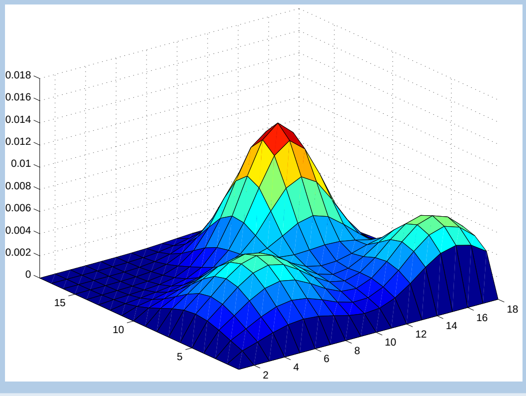
- This system is built from HMM based recognizers that are designed for the task of handwritten text line recognition.
- For every writer a verification system is build from character models using HMMs with 14 states connected in a linear topology. Gaussian output probability distributions are used.
- The character models are connected to word models which themselves are connected to form line models.



3. A GMM based System

- For each writer the model is defined by a density of a weighted sum of Gaussian densities.
- Properties of GMMs:
  - The model consists of only one state and one distribution function with one model for every writer.
  - Neither words nor characters have to be modeled, so no transcriptions of the text lines are needed.

  ▷ GMMs are conceptually much simpler than HMMs.



4. Confidence Measure 1/3

- Raw Log-Likelihood Confidence Measure
  - Uses the log-likelihood score $ll_{\mathrm{ClaimedID}}$ of the claimed identity for a text line t:

$$cm_{\mathrm{LLScore}}(t) = ll_{\mathrm{ClaimedID}}$$

  - To obtain this confidence measure the score of the claimed model has to be calculated only.

4. Confidence Measure 2/3

b UNIVERSITÄT BERN

- World Model Based Confidence Measure
  - The idea is to normalize the score of the claimed writer by a world model which is trained on a large number of samples from many writers.
  - Defined as the difference between the log-likelihood score of the claimed identity $ll_{\mathrm{ClaimedID}}$ and the world model $ll_{\mathrm{WorldModel}}$:

$$cm_{\mathrm{WorldModel}}(t) = ll_{\mathrm{ClaimedID}} - ll_{\mathrm{WorldModel}}$$

  - Two scores have to be calculated. This confidence measure is independent of the number of writers considered.

4. Confidence Measure 3/3

- Cohort Model Based Confidence Measure
  - The idea is to normalize the score of the model of the claimed writer with respect to the scores of the most competitive writers.
  - Defined as the difference between the log-likelihood scores of the claimed identity $ll_{\mathrm{ClaimedID}}$ and of the first best ranked competing writer $ll_{\mathrm{BestRankedCompeting}}$:

$$cm_{\mathrm{CohortModel}}(t) = ll_{\mathrm{ClaimedID}} - ll_{\mathrm{BestRankedCompeting}}$$

  - To obtain the score of the first best ranked competing writer the scores of all models have to be calculated. Thus the confidence measure depends on the number of writers considered.

5. Experimental Setup

UNIVERSITÄT BERN

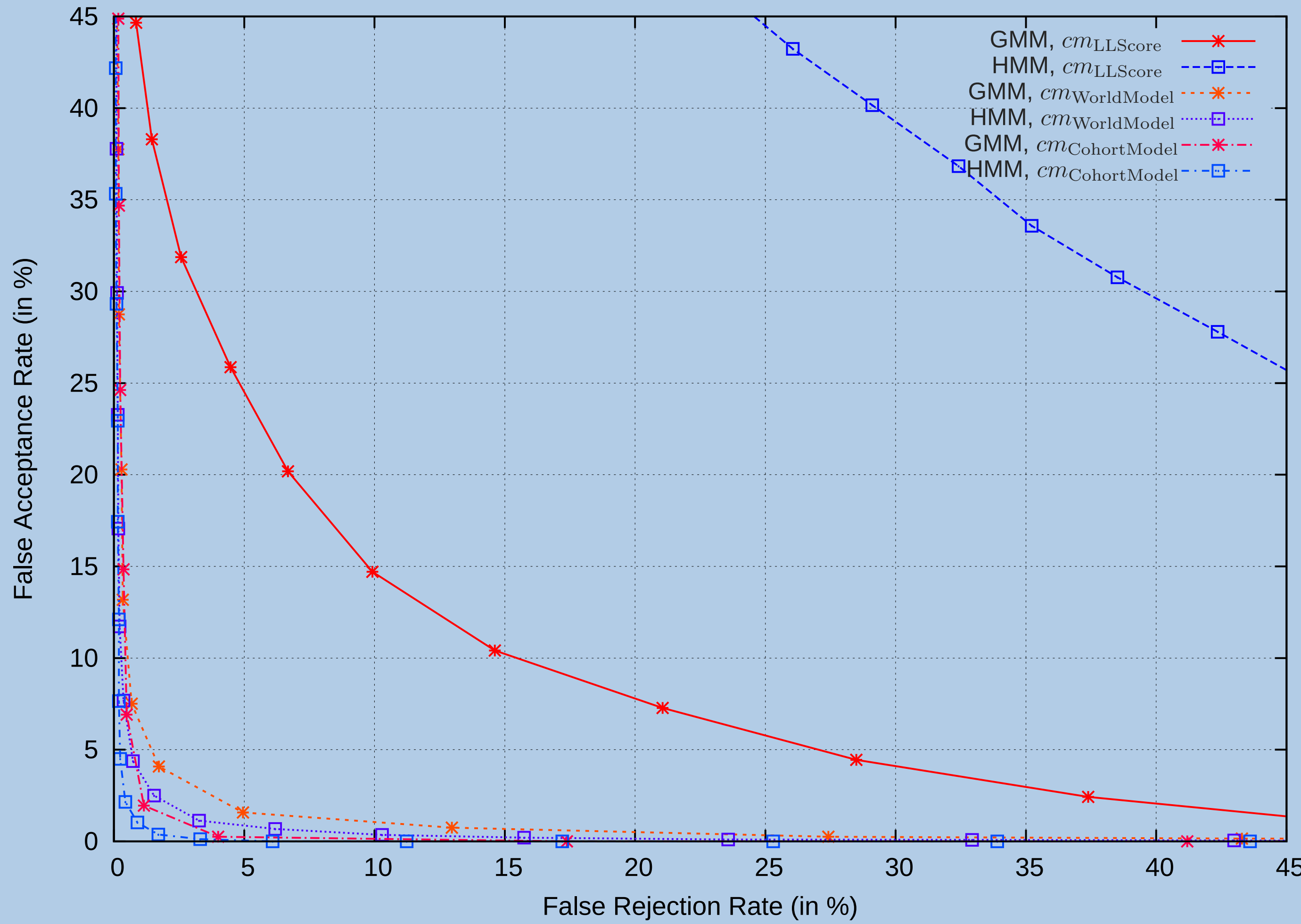
- Both writer verification systems are trained on the same training set consisting of 4,103 text lines from 100 different writers which are a subset of the IAM Handwriting Database.
- For each writer, the set of available text lines is split into four disjoint subsets to perform four-fold cross validation.
- Both systems are optimized to achieve an optimal performance on a related writer identification task.
- The IAM Handwriting Database is publicly available at: `www.iam.unibe.ch/~fki/iamdb`

5. Test Sets

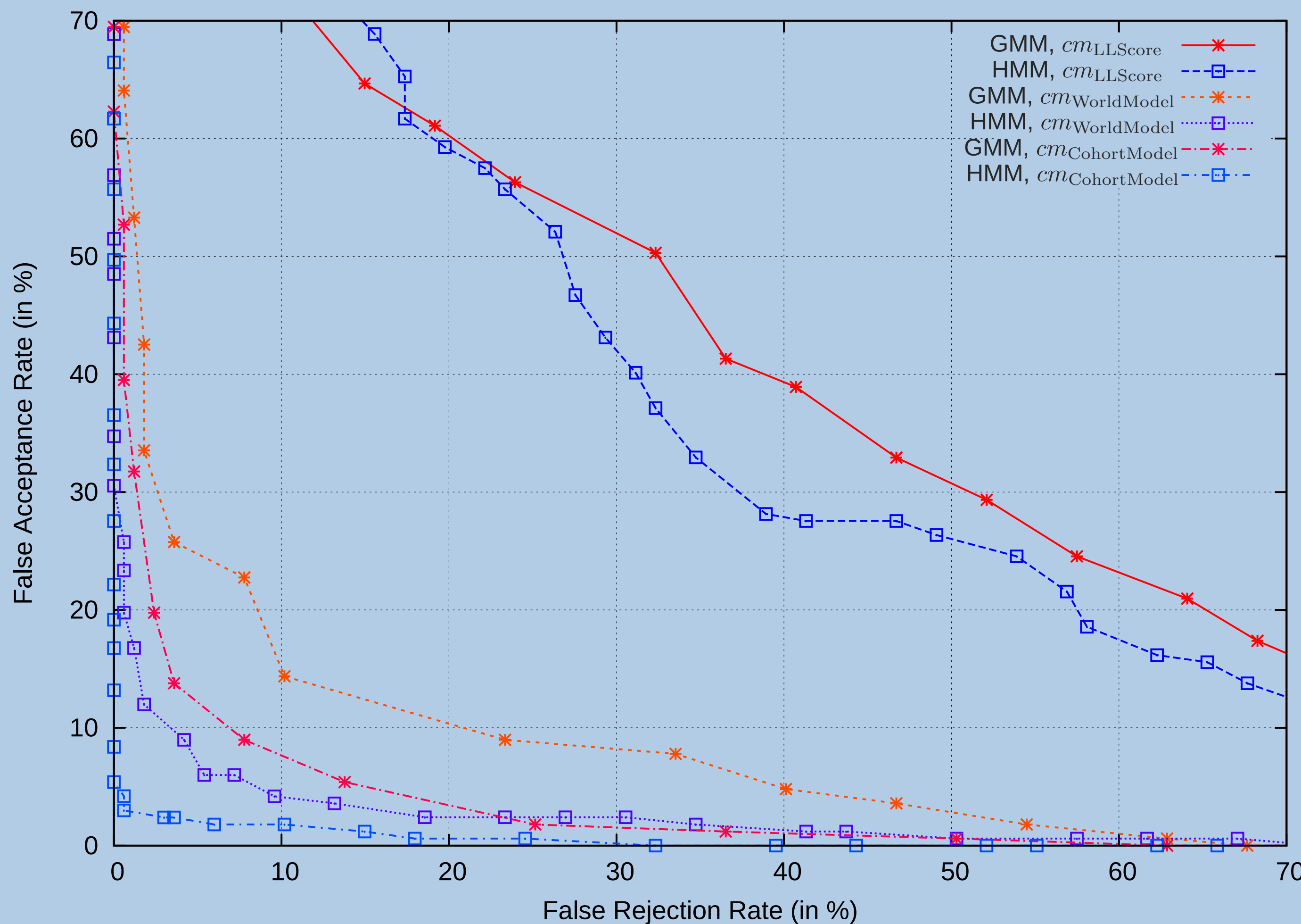
- Unskillfully Forged Test Set
  - ○ 100 writers
  - ○ 4,103 client text lines
  - ○ $7 \times 571 = 3{,}997$ impostor text lines (the identities are randomly assigned to the text lines)
  - ▷ 8,100 text lines in total
- Skillfully Forged Test Set
  - ○ 20 writers
  - ○ 169 client text lines
  - ○ the same 169 text lines are skillfully forged
  - ▷ 338 text lines in total

5. Results – Unskilled Forgeries

UNIVERSITÄT BERN



5. Results – Skilled Forgeries



5. Results – Equal Error Rates

Equal Error Rate	Unskilled Forgeries	Skilled Forgeries
GMM, cm_{LLScore}	34.0%	39.5%
HMM, cm_{LLScore}	12.5%	34.0%
GMM, $cm_{\mathrm{WorldModel}}$	3.0%	13.0%
HMM, $cm_{\mathrm{WorldModel}}$	2.0%	5.9%
GMM, $cm_{\mathrm{CohortModel}}$	1.6%	8.4%
HMM, $cm_{\mathrm{CohortModel}}$	1.0%	1.8%

6. Conclusions and Future Work

UNIVERSITÄT BERN

- Conclusions
  - The HMM and the GMM based writer verification systems perform similarly on the unskillfully forged test set.
  - The HMM based system outperforms the GMM based system on the skillfully forged test set.
  - The cohort model based confidence measure produces the best ROC curves.

6. Conclusions and Future Work

- Conclusions
  - The HMM and the GMM based writer verification systems perform similarly on the unskillfully forged test set.
  - The HMM based system outperforms the GMM based system on the skillfully forged test set.
  - The cohort model based confidence measure produces the best ROC curves.
- Future Work
  - Combine the HMM and the GMM based systems using a Multiple Classifier System (MCS) approach.