Poset – partially ordered set

Poset Description of Grid Features and Application to Off-Line Signature Verification

Paper authors:

E. N. Zois[(1)], E. Zervas[(1)], K. Barkoula[(2)] , **G. Economou**[(2)], S.Fotopoulos[(2)]

(1) Electronics Engineering Dept. Technological & Educational Institution of Athens, 12210, GREECE.

(2) Electronics Laboratory, Physics Dept., Univ. of Patras, Patras, 26500, GREECE.

14th International Conference on Frontiers in Handwriting Recognition.
Crete, Sep. 2014

Contents

Introduction to the field

- Handwriting: A behavioral way for resolving the problem of recognizing writers
- Lots of Applications: Forensics, Security, e-business, e.t.c.
- Handwriting based verification can be categorized to :
  - ✓ Context Dependent
    - \- <u>Signatures</u> or pre-defined text
  - ✓ Context Independent

- <u>Signatures</u>: The common way to declare our identity.
  - ✓ On-line and/or <u>Off-line</u>

The basic idea

- Presented a couple of years ago
- A feature extraction method with applications to:
  - ✓ Signatures
  - ✓ Coding of words and sentences
- Produced encouraging results (EER)
- Based on the probabilistic measure of predefined pixel transitions

The new proposal

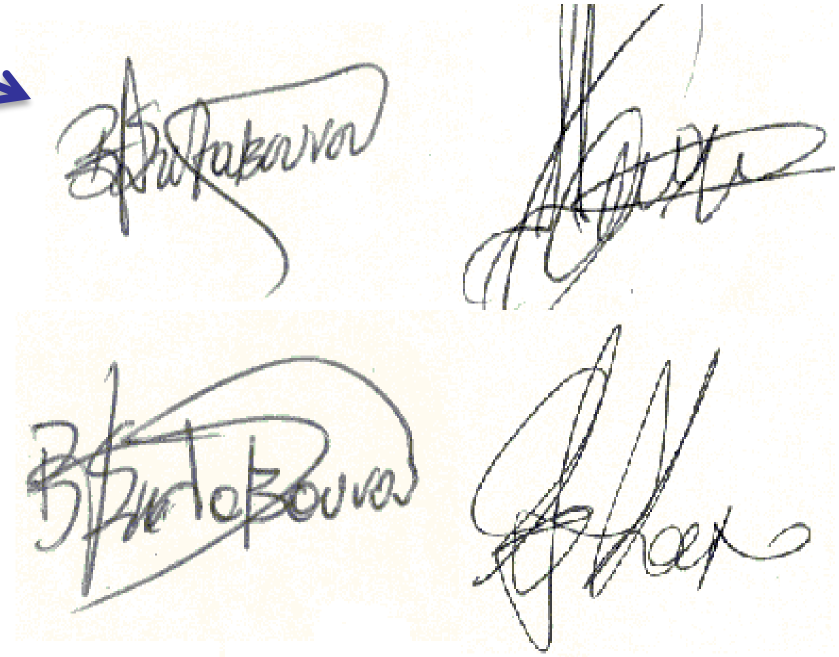
- We improved over the old feature extraction method

- Provide a new feature modeling:
  - ✓ Combine concepts from information and communication theory
  - ✓ Consider the old features as symbols
  - ✓ Use sequences of symbols to create events
  - ✓ Estimate their first order probabilities

- The outcome of this procedure is an attempt to model the handwriting process in concordance with basic elements of information and coding theory.

Databases

- **CORPUS1**: Greek database with 69 writers.
  - ✓ **Under** enrichment and restructuring
  - ✓ **Each writer**: 105 samples (genuine) and 21 skilled forgeries
  - ✓ **Development time**: One year

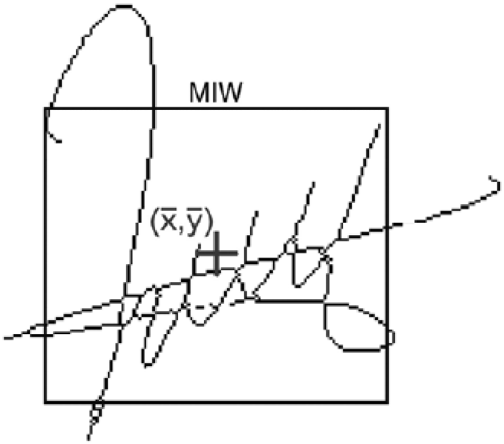


- **CORPUS2**: GPDS**300**
  - ✓ Well known
  - ✓ **Each writer**: 24 samples (genuine) and 28 skilled forgeries

No picture is displayed due to license restrictions

Signature Preprocessing

- Typical preprocessing algorithms were applied:
  - ✓ Signature Segmentation
  - ✓ Thresholding with Otsu's Algorithm
  - ✓ Thinning or skeletonization
    - ☞ *Thinning was not the best choice for GPDS300*
  - ✓ Finding '*center of mass*' of each signature
  - ✓ Most informative window (MIW)
- Feature extraction with respect to MIW section



Feature Extraction: Pixels…

- Consider a 3x3 pixel window-mask.
- Locate its starting point at the 3,1 coordinates

1,1	1,2	1,3
2,1	2,2	2,3
3,1	3,2	3,3

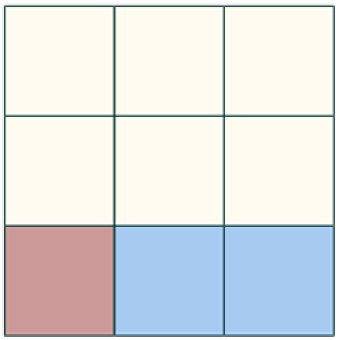
Feature Extraction: Pixels…

- Create connected binary patterns of **3** pixels
  a) **Starting** from **3,1**.
  b) **Ending** at any pixel with Chebyshev (chessboard) **distance equal to two (2).**

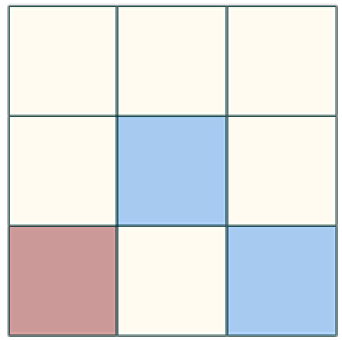
- Eight (8) primary patterns $BG_{i,}$ $i=\{0:7\}$

1,1	**1,2**	**1,3**
2,1	2,2	**2,3**
3,1	3,2	**3,3**

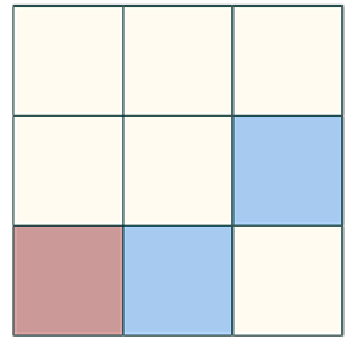
The Eight Primary Binary Grids (BG_i)



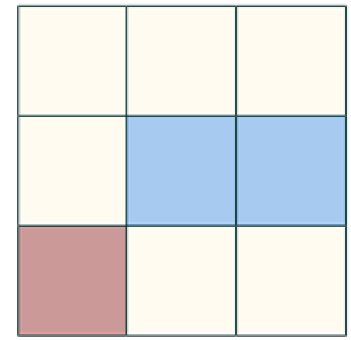
BG_0



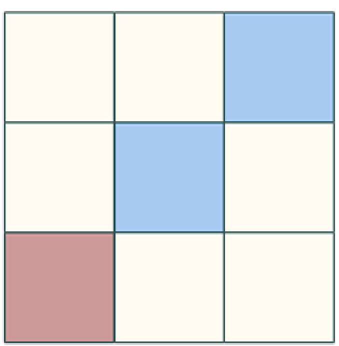
BG_1



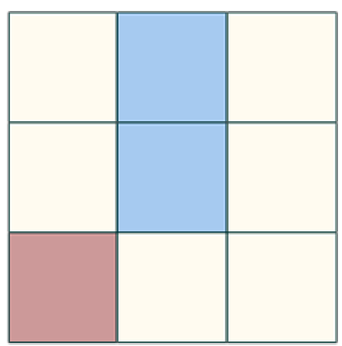
BG_2



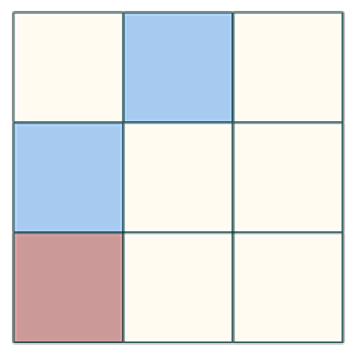
BG_3



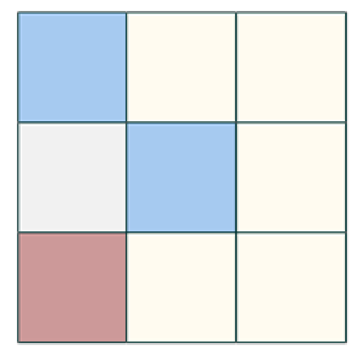
BG_4



BG_5



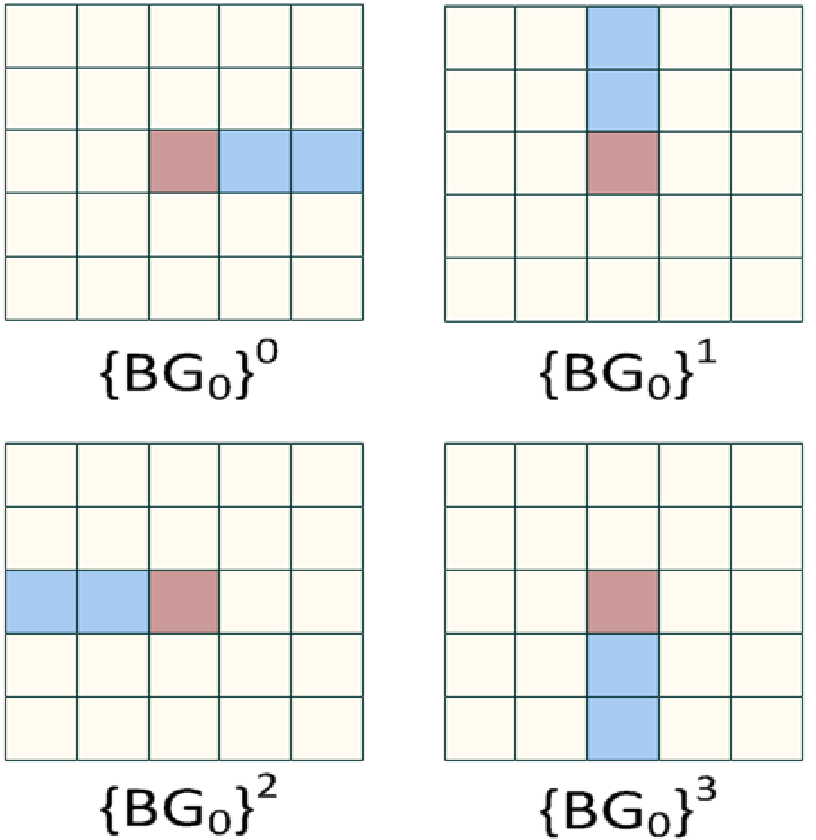
BG_6



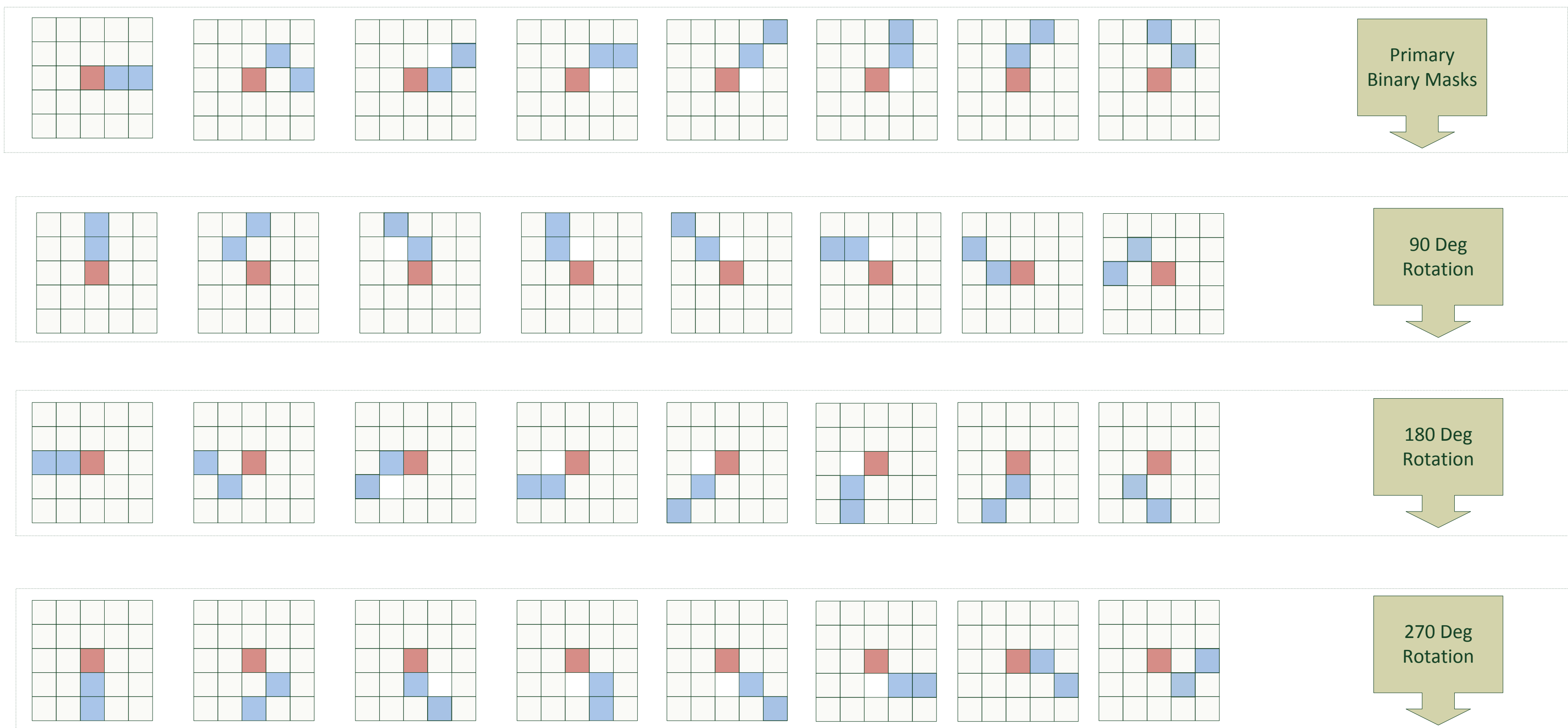
BG_7

Rotating BG_is

- Each of the BG_is is rotated by 90, 180 and 270 degrees
- The result is now positioned within a ***5x5 grid***
- An example is provided to the right for the BG_0
- Total number of alphabet symbols equals to 32

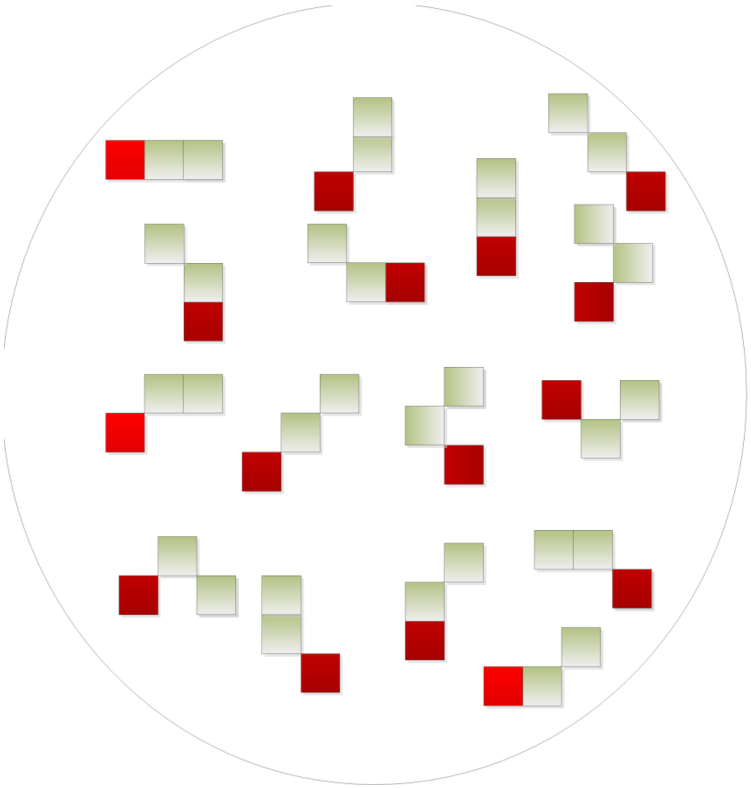


The Entire 32 Element Set



Modeling signature pixels

- Let us consider a collection of the 32-element set
  - A set of predefined **symbols**.
- The feature extraction process can be modeled as a ***discrete space – discrete alphabet source***
  - **Simple events**: *Presence of a symbol*
  - **Compound events**: *Presence of a symbol combination*



Features and Grids (a)

- The number of combinations that the 32 elements can provide is almost immeasurable ($\sim 10^{27}$).

- A reduction is applied to the number of extracted events by employing the functional and convenient concept of set partitioning.

Features and Grids (b)

- The elements of the 32 element set are grouped, (partitioned) into subsets of eight tetrads
  - *number of possible combinations still very large*

- Each one of them is called a **<u>scheme</u>**

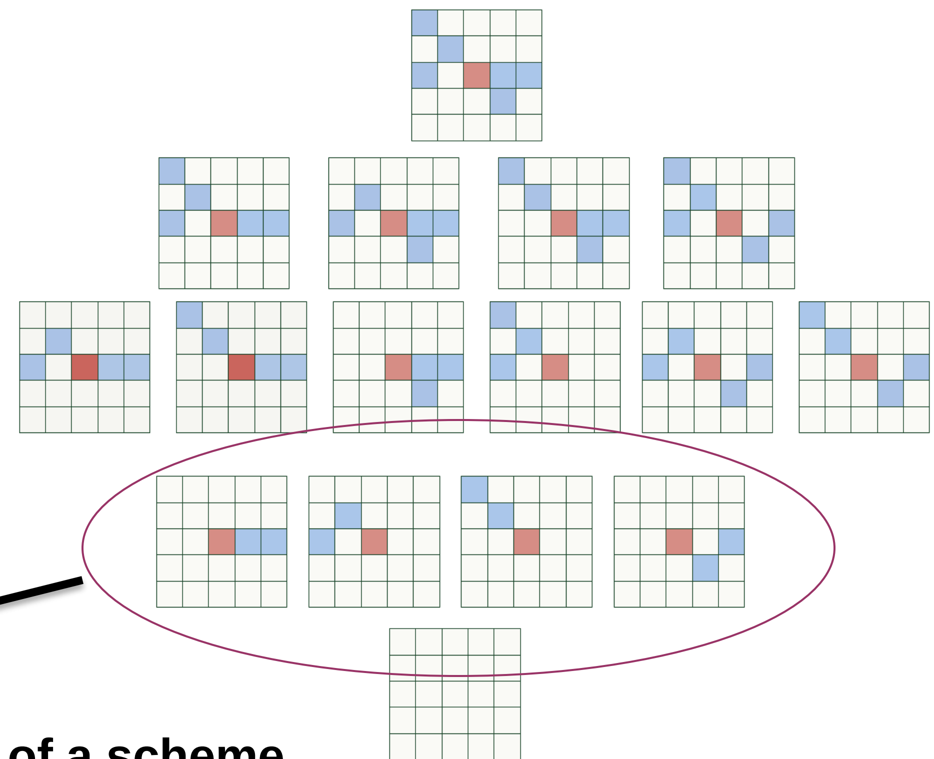
- Further reduction is achieved by selecting only **<u>orthogonal</u>** schemes
  - *Orthogonal schemes have their tetrad members arranged in such a way that no any other member can be described by the linear combination of the three remaining*

- Now, the total number of schemes is 2587

Power-Set

The set of all subsets of a set A is called the *power set* of A and denoted as $\wp(A)$ or sometimes as 2^A.
For example, if $A = \{a,b\}$, $\wp(A) = \{\emptyset, \{a\}, \{b\}, \{a,b\}\}$.

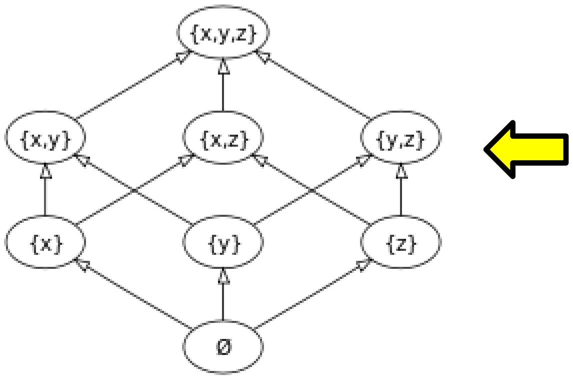
- Given a scheme:
  - For each one of its ***eight*** tetrads, create their powerset.



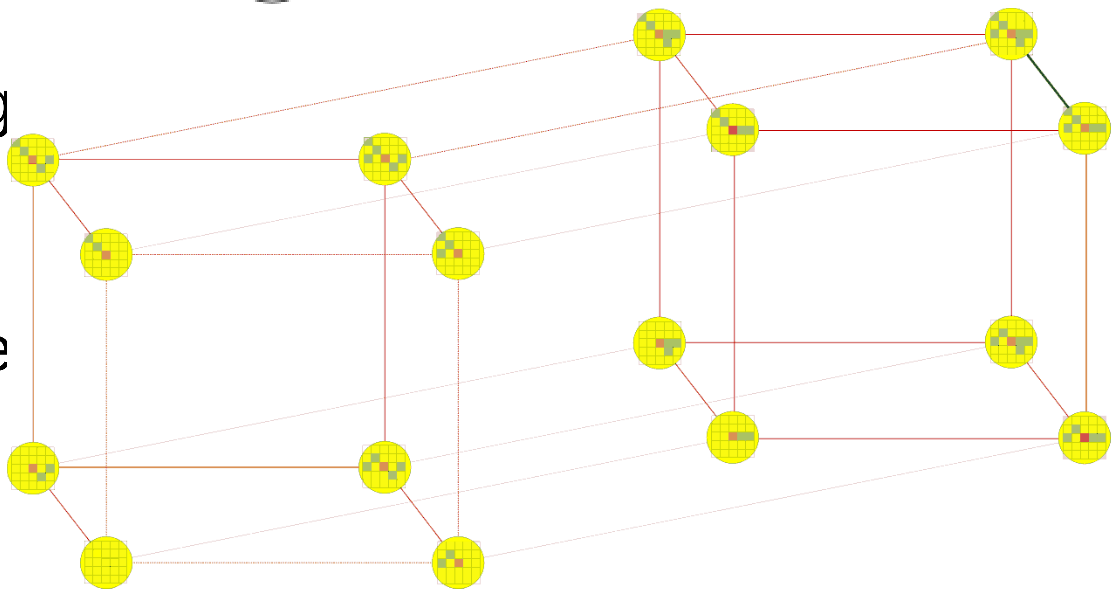
One of the eight tetrads **(layer) of a scheme and its corresponding** *(sixteen element)* **powerset.**

Partially **O**rdered **Set** – (Poset)

- The concept of ordering:
  - Each one of the eight power-sets is evaluated by ordering its elements with respect to **inclusion**.
  - Detected features are those designated as links on the poset grid.



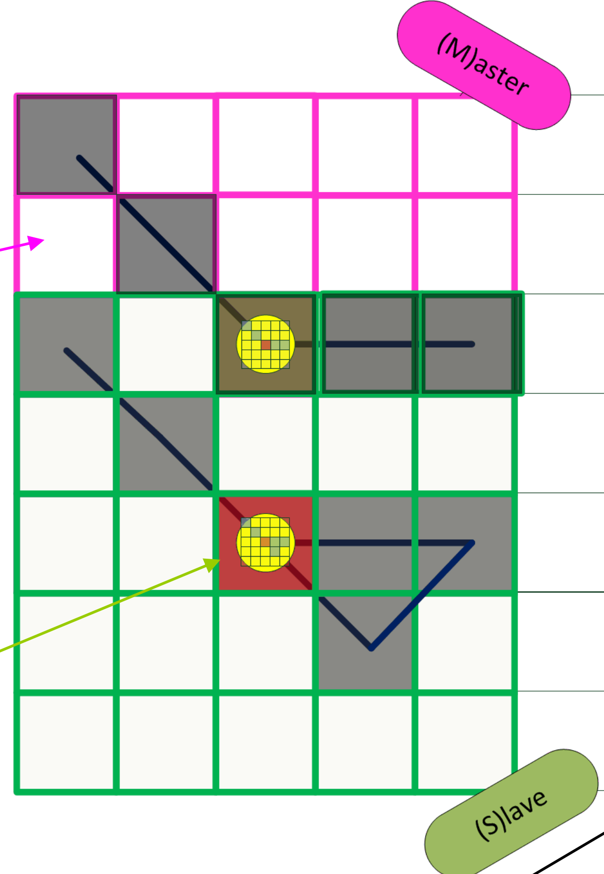
The elements of the power set of the set {x, y, z} ordered in respect to inclusion (wikipedia)



The power set of a 4-element set **ordered** by inclusion

Poset Features

Detection of a symbol

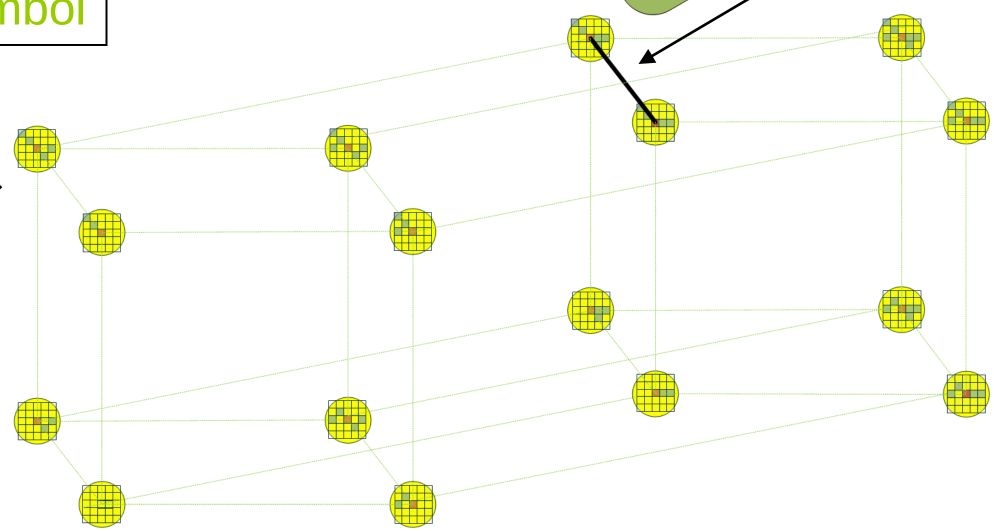


Poset features are detected using enlarged window

Feature found

Detection of a second symbol

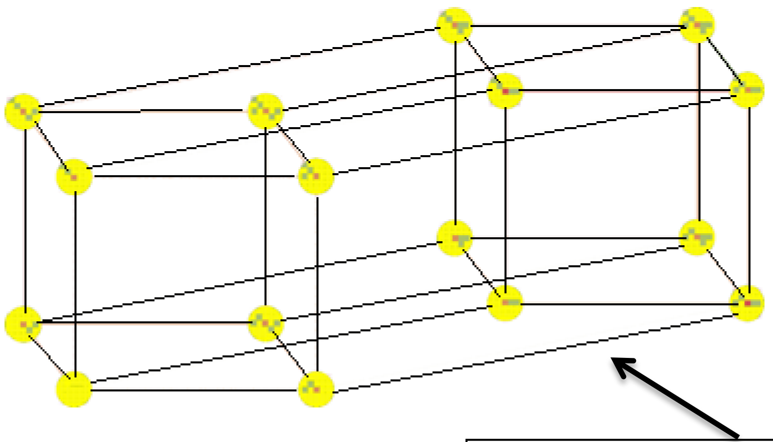
For each edge there is a corresponding probability index which is updated when the specific ordering of symbols occurs



Feature Dimensionality

Number of Features

(i) for each tetrad = **32**



12+ 12+8 =**32**

(ii) There are **8** tetrads 32x8 = **256**

(iii) In addition to the whole signature image, each signature is partitioned in **4** segments (1+4=**5**)

Thus feature dimensionality is:
5x256 = **1280**

1 + 4 =**5**

Verification Scheme

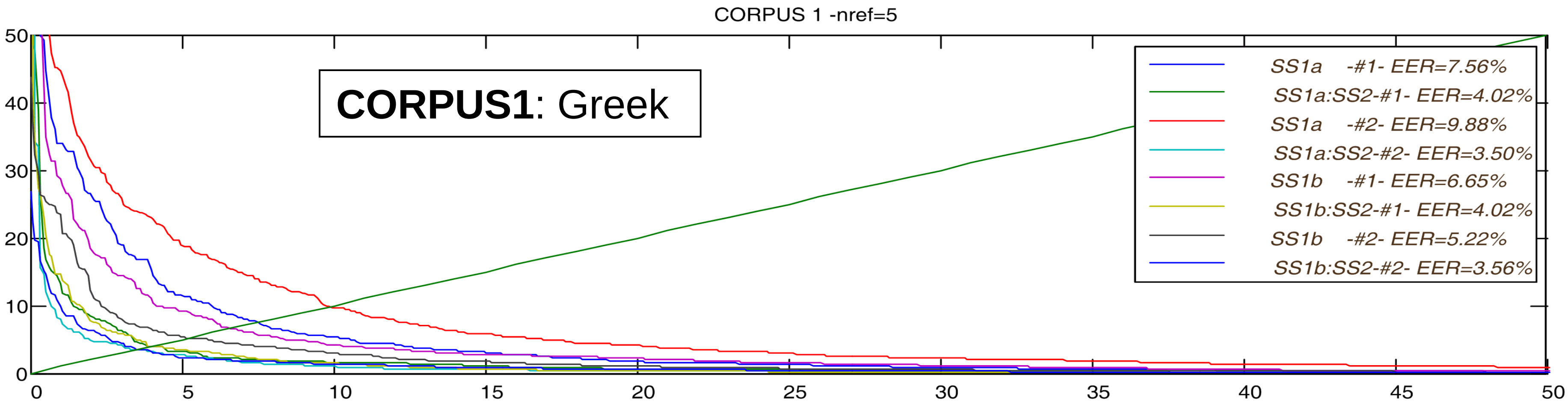
- Writer Dependent (WD) approach
- For each writer, *#nref* reference samples of genuine along with an equal number of simulated-forgery signature samples are randomly chosen in order to train the classifier
- The classifier is a hard-margin two class support vector machine (SVM) classifier using radial basis kernel
- The SVM outputs:
  - *binary class decision*
  - *a score value (equal to the distance of the tested sample from the SVM separating hyperplane)*
- There is a wide area of rbf sigma values that the system provide the reported results

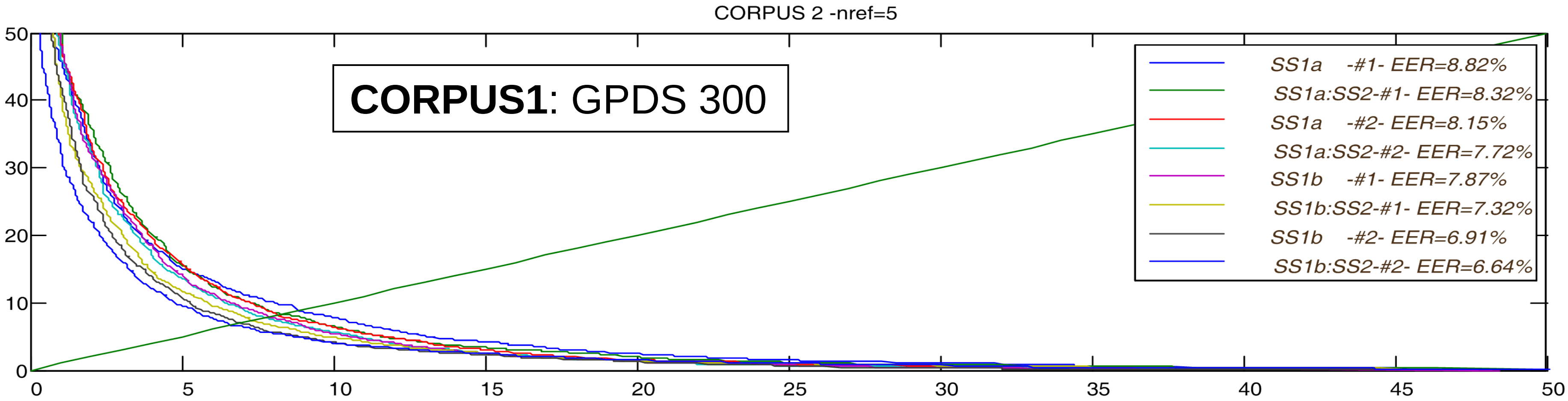
Verification Scheme

- Evaluation of the verification efficiency of the system is accomplished with the use of a global threshold applied on the overall SVM output score distribution

- Calculation of the FAR, FRR and EER

Results – ROC, EER





Results – Comparisons Corpus1

	FRR	**FAR**	**EER**
K. Tselios, [11] IET '12	-	-	9.16
K. Barkoula, [21] AFHA'13	3.29	2.18	2.79
Proposed: random scheme #1	**2.97**	**4.11**	**3.51**
Proposed: random scheme #2	**3.44**	**3.78**	**3.56**

Results – Comparisons Corpus2

Primary Author	FRR	FAR	EER
M. Ferrer, [27]	13.40	12.60	13.12
J. F. Vargas, [10]	12.06	10.53	9.02
L. Batista, [24]	16.81	16.88	-
G. Pirlo, [25]	-	-	4.6
V. Niguen, [27]	-	-	17.25
M. B. Yilmaz [28]	-	-	15.41
R. Kumar, [14]	-	-	13.76
J. R. Solar [29]	-	-	15.30
K. Tselios, [11] IET '12	-	-	12.32
K. Barkoula, [21] AFHA'13	5.23	13.03	9.04
Proposed: Random scheme #1	**4.30**	**11.56**	**7.72**
Proposed: Random scheme #2	**9.22**	**4.61**	**6.65**

Conclusions

- A new modeling of a feature extraction method
- Ordering of power set with respect to inclusion
- The method seems promising
- There are still many issues that we must address

Among others ⇨

Conclusions - issues to be addressed

- Writer Independent (WI) method – Dissimilarity framework
- Definition of first and higher order transition probs
- Application to writer verification problems
- Signature Complexity and Stability issues
  - *Preliminary results have been presented at AFHA 2014*
- Selection of the optimal scheme:
  - *sparse representation approach (preliminary results)*
- Use of multi-resolution windows
- ………….

Thank you

Questions?