Scientific Talk

Indexing Protected Deep Face Templates by Frequent Binary Patterns

Dailé Osorio-Roig

da/sec – Biometrics and Internet Research Group

Hochschule Darmstadt

Contents

Introduction

Motivation

- EU GDPR 2016/679 defines biometric information as ***sensitive data***.
- ***Reconstruction*** of facial images from their corresponding ***embeddings***.
- ***Limited*** application of the ***biometric template protection*** methods.

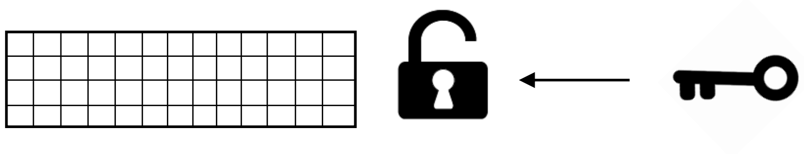
Contribution

- Introduction of a new approach based on the search of ***frequent binary patterns***.
  1. High application: for indexing and retrieval of protected binary templates across different cancelable schemes, i.e. ***Computational workload reduction***.
  2. ***Agnostic*** approach: in terms of biometric modality and cancelable scheme.

Introduction

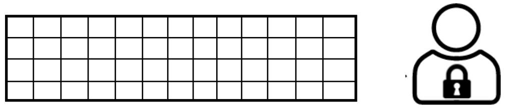
Biometric template protection schemes

- Biometric Cryptosystems.
  - ✓ Key-binding or Key-generation schemes.
  - ✓ Homomorphic encryption.



- ○ Enable complex process of biometric comparison by verifying the correctness of a retrieved key.
- ○ Optimizations are depending on the encoding schemes.
- ○ Limited operations in the encrypted domain.

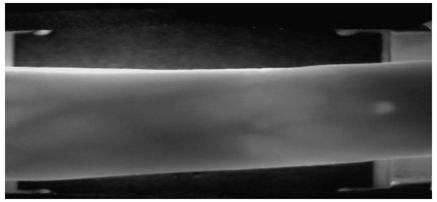
- Cancelable schemes
  - ✓ Feature transformation step.

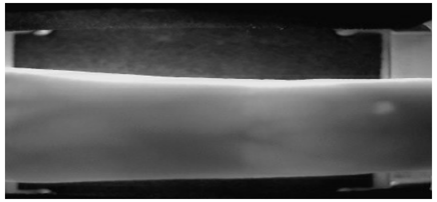
$f(.) \longrightarrow$ 

  - ✓ Retaining efficient biometric comparators.

- ○ High application in identification systems.
- ○ Even, for exhaustive searches.
- ○ Workload reduction accelerates one-to-one comparison.

Properties of biometric data





$0\mathbf{0}\underbrace{011}_{\mathbf{111}}10010$

$0\mathbf{1}\underbrace{011}_{\mathbf{111}}10010$

Fuzzy data

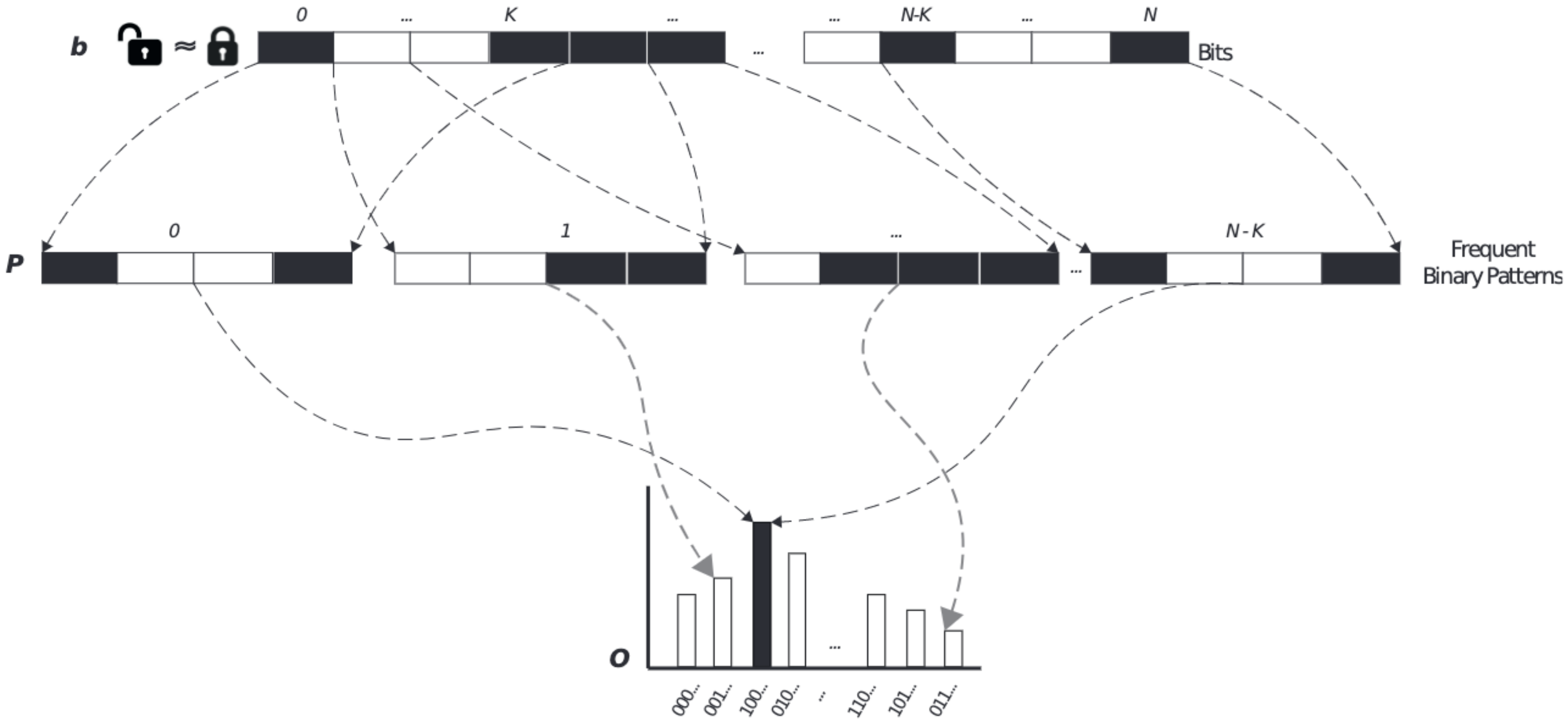
$01\mathbf{11}\underbrace{1100}_{\mathbf{1100}}10$

$01\mathbf{00}\underbrace{1100}_{\mathbf{1100}}10$

- ➢ Common ***binary patterns*** can be suitable for biometric indexing, even, for cancelable templates ?.

Proposed system

- Let $b \in \{0,1\}^{N}$ be a bit-string of size $\boldsymbol{N}$ and $\boldsymbol{K} < \boldsymbol{N}$ a given frequent pattern length.
- A set $\boldsymbol{P}$ of binary patterns are extracted from $\boldsymbol{N}$ bits.
- Frequent patterns are defined according to their corresponding number of occurrences in $\boldsymbol{N}$.



Proposed system

- Enrolment step: the frequent binary pattern with ***maximum number of occurrence*** is selected as a bin.
- Retrieval step: the search is through ***all*** their ***most frequent binary patterns*** (sorted) until a match is found.
- Definition of the workload reduction.

$$W = \sum_{i=1}^{z} |l_i|$$

$|l_i|$ Denotes the number of references stored in bin l_i

$z \leq R$ Denotes a threshold for the maximum number of bins visited.

Workload reduction can be easily controlled by the number of bins visited !

Experimental setup

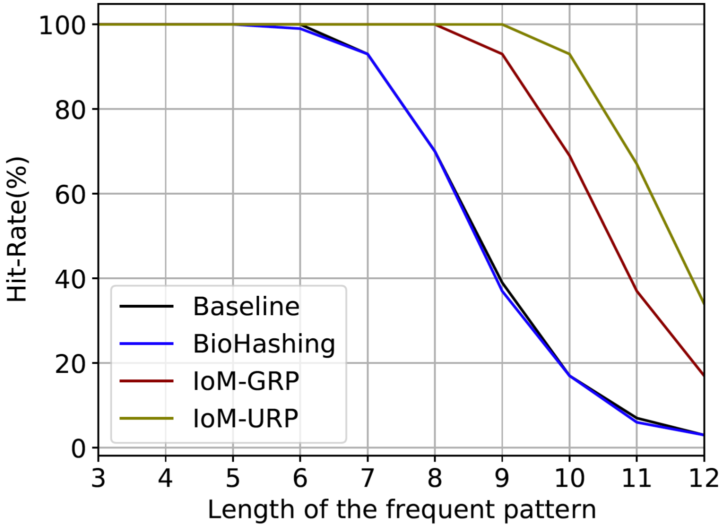
- LFW and SDUMLA image databases for Face and Finger-vein.
- Embedding features from ArcFace (Face) and ImageNet (Finger-vein (set of testing)) pre-trained models.
- Sub-sampling over 10 rounds.
- Number of enrolled subjects in Face and Finger-vein: 1680 and 318, respectively.
- For impostor comparisons:
  - sub-set of subjects containing a single sample (LFW).
  - sub-set of samples selected for the training set (SDUMLA).
- Evaluation in closed-set scenario in terms of hit-rate (HR) and penetration rate (PR).
- Evaluation in open-set scenario in terms of DET curves.

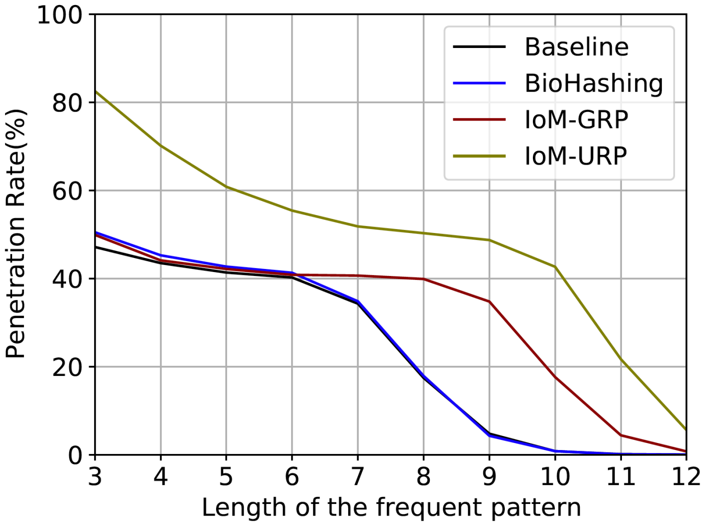
Experimental setup

- Generic cancelable schemes, representing the current state-of-the-art for biometric template protection:
  - BioHashing – a single binary representation.
  - Variants of Index-of-Maximum Hashing: Uniformly Random Permutation (IoM-URP) and Gaussian Random Projection (IoM-GRP).
  - Encoding step over IoM-URP and IoM-GRP.
  - Baseline: original embeddings are binarized by using the function sign with threshold 0, representing the unprotected system.
- Normal- (where each user’s key is assumed to be secret) and stolen-token- (where the impostor has access to the genuine user’s secret key) scenarios for cancelable schemes.

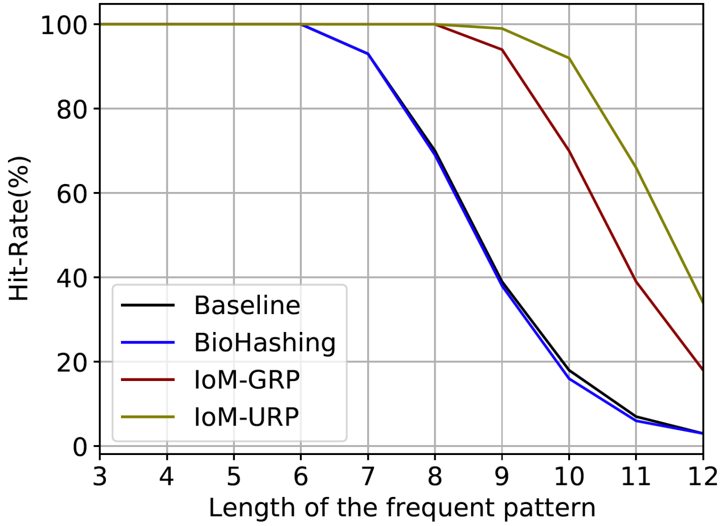
Results over Face (closed-set scenario)

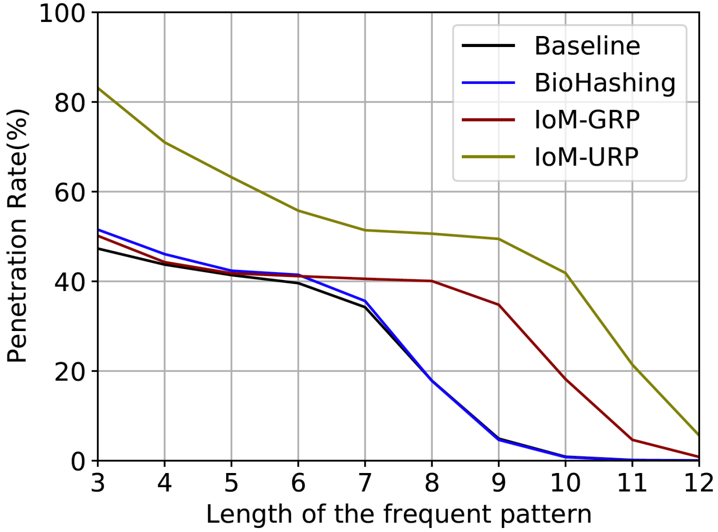
Normal-scenario





Stolen-token-scenario



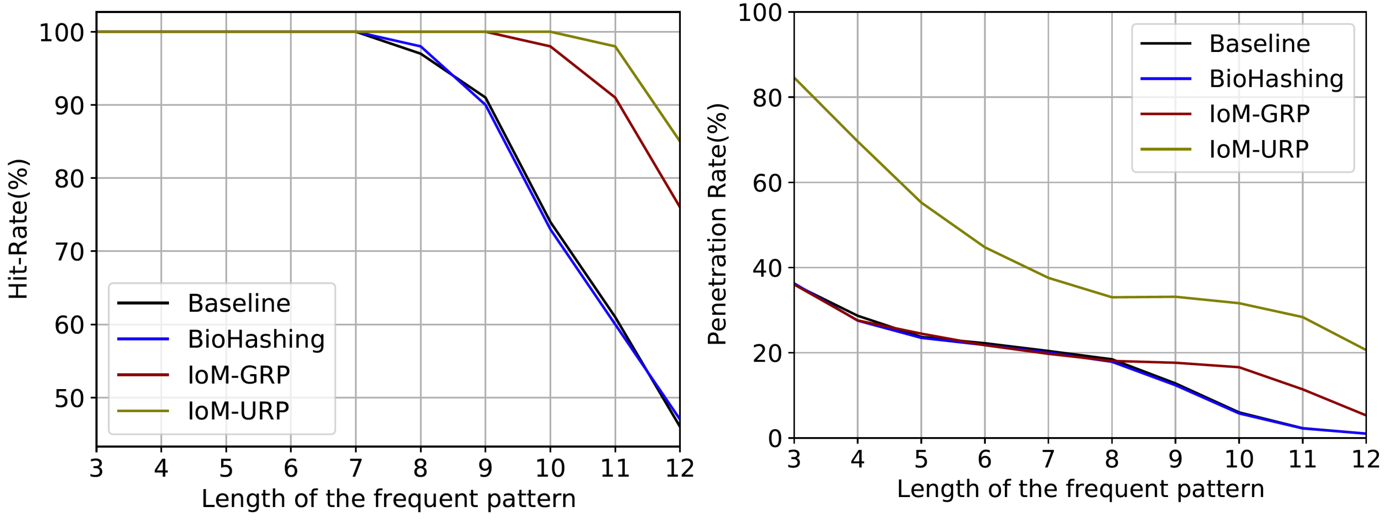


fbp = frequent binary pattern

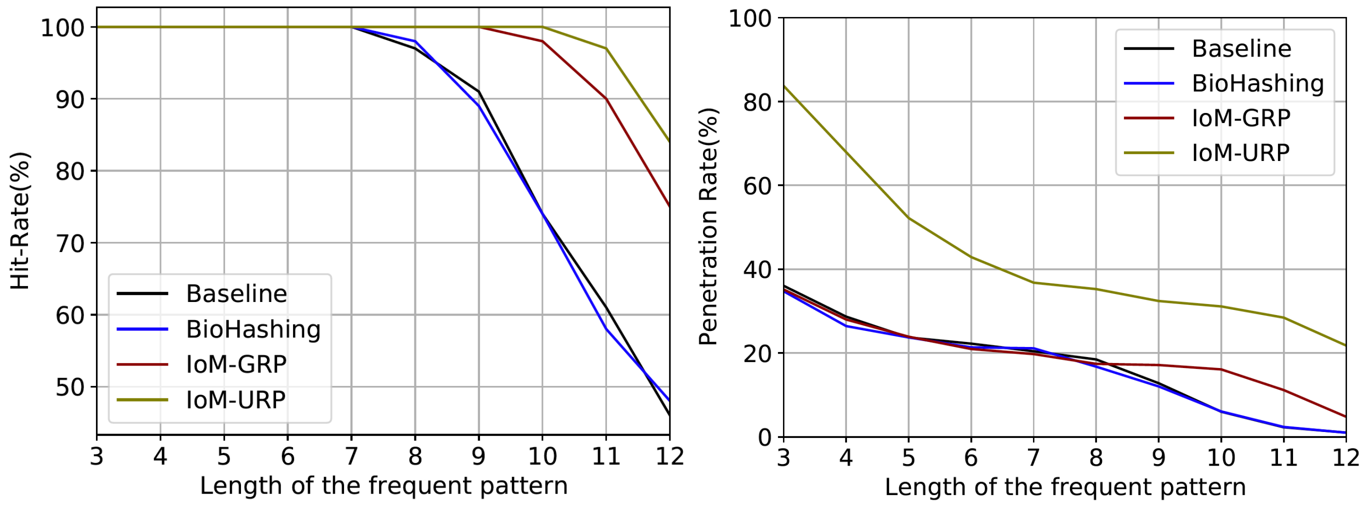
- Non-significant differences between Normal- and Stolen-token- scenario.
- H-R 100% is maintained to a certain length of the fbp depending on the BTP scheme.
- P-R < 52% reducing the baseline workload (i.e 100%).
- Lowest WR (P-R < 36%) by a drop in the H-R (down to 93%) while fixing a length fbp across BTP schemes.

Results over Finger-vein (closed-set scenario)

Normal-scenario



Stolen-token-scenario



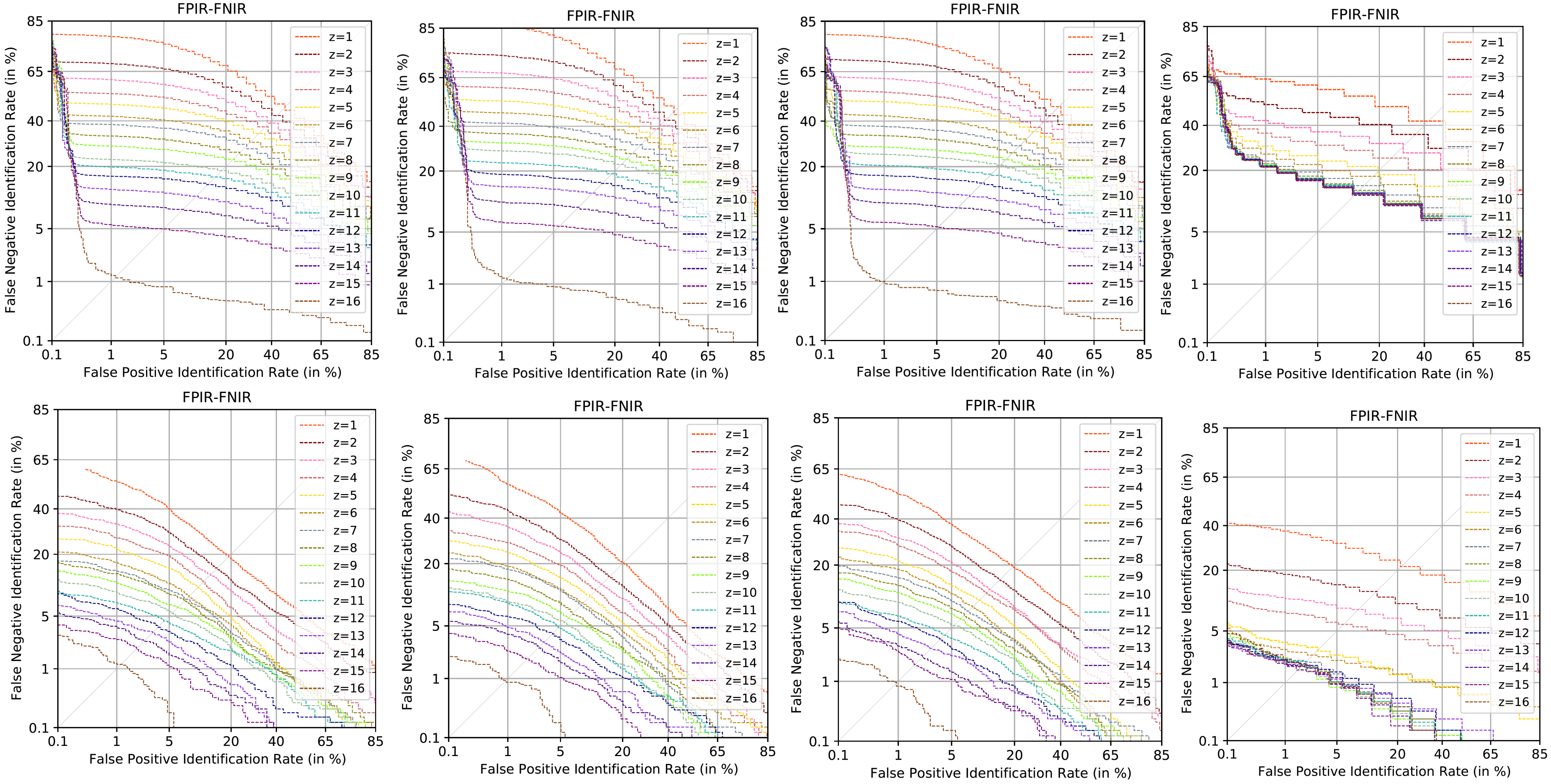
fbp = frequent binary pattern

- Non-significant differences between Normal- and Stolen-token- scenario.
- H-R 100% is maintained to a certain length of the fbp depending on the BTP scheme.
- For larger lengths than face, while maintaining H-R 100%.
- P-R < 22% reducing the baseline workload (i.e 100%) – IoM-URP (PR < 45%) over H-R 100%.

Results over Open-set scenario

Stolen-token-scenario – challenging scenario

baseline | BioHashing | IoM-GRP | IoM-URP



Face

$$W = \sum_{i=1}^{z} |l_i|$$

For $z = [1,16]$

For $k = 4$

Finger-Vein

Best results over Open-set scenario

Face

BTP approach	Normal-scenario			Stolen-token-scenario		
	FNIR@FPIR=1.0%	z	P-R(%)	FNIR@FPIR=1.0%	z	P-R(%)
Baseline	19.76	11	66.08	19.76	11	66.08
BioHashing	23.30	11	66.27	23.14	11	66.44
IoM-GRP	19.57	11	66.28	20.37	11	66.61
IoM-URP	22.33	5	87.90	29.99	5	88.59

- A rejection rate for genuine identification transactions of less than 24%.
- 66% of workload-reduction.

Finger-Vein

BTP approach	Normal-scenario			Stolen-token-scenario		
	FNIR@FPIR=1.0%	z	P-R(%)	FNIR@FPIR=1.0%	z	P-R(%)
Baseline	14.40	7	45.60	14.40	7	45.60
BioHashing	17.86	8	50.17	14.81	8	49.15
IoM-GRP	16.54	8	50.31	13.49	8	50.89
IoM-URP	11.45	3	84.95	11.16	3	84.04

- A rejection rate for genuine identification transactions of less than 18%.
- 51% of workload-reduction.

Future Work

- Extension of the proposed system to multi-biometrics :
  - Frequent binary patterns are combined from multiple biometric characteristics.
- Generalise the idea to other biometric characteristics (e.g. deep iris features).
- Selection of the best configuration in terms of the number of bins visited from the probe is a challenge:
  - Include evaluation of the accuracy in terms of rank-1.
- Larger datasets.

Conclusions

- Search of frequent binary patterns over binary representations seems to be suitable for biometric indexing.
- High application for indexing templates protected by Cancelable schemes.
- Different statistical data in terms of workload-reduction could be perceived from different biometric characteristics.
- Many future work avenues.

Thank you for your attention!
Questions ?