À LA RECHERCHE DE PASSE-PARTOUT BIOMÉTRIQUES

Master classe FIC 2023

Tanguy Gernot

Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, FRANCE

tanguy.gernot@unicaen.fr
https://gernot.fr

GREYC
Laboratoire de recherche en sciences du numérique

Normandie Université

UNICAEN

ENSI CAEN
ÉCOLE PUBLIQUE D'INGÉNIEURS
CENTRE DE RECHERCHE

cnrs

Plan

Plan

Définition de la CNIL

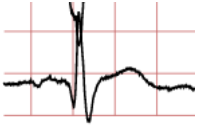
La biométrie regroupe l'ensemble des techniques informatiques permettant de reconnaître automatiquement un individu à partir de ses caractéristiques issues de différentes modalités : physiques, biologiques, ou comportementales.

Définition de la CNIL

La biométrie regroupe l'ensemble des techniques informatiques permettant de reconnaître automatiquement un individu à partir de ses caractéristiques issues de différentes modalités : **physiques**, biologiques, ou comportementales.

Définition de la CNIL

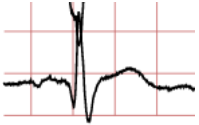
La biométrie regroupe l'ensemble des techniques informatiques permettant de reconnaître automatiquement un individu à partir de ses caractéristiques issues de différentes modalités : physiques, **biologiques**, ou comportementales.



Définition de la CNIL

La biométrie regroupe l'ensemble des techniques informatiques permettant de reconnaître automatiquement un individu à partir de ses caractéristiques issues de différentes modalités : physiques, biologiques, ou **comportementales**.

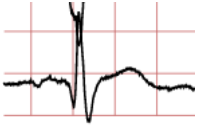





Définition de la CNIL

La biométrie regroupe l'ensemble des techniques informatiques permettant de reconnaître automatiquement un individu à partir de ses caractéristiques issues de différentes modalités : physiques, biologiques, ou comportementales.






Les données biométriques sont des données à caractère personnel, car elles permettent d'identifier une personne : ce sont des données sensibles !

Sensibilité des données biométriques

Contrairement à un mot de passe, les données biométriques ne sont pas modifiables à volonté (doigts, visages).

Sensibilité des données biométriques

Contrairement à un mot de passe, les données biométriques ne sont pas modifiables à volonté (doigts, visages).

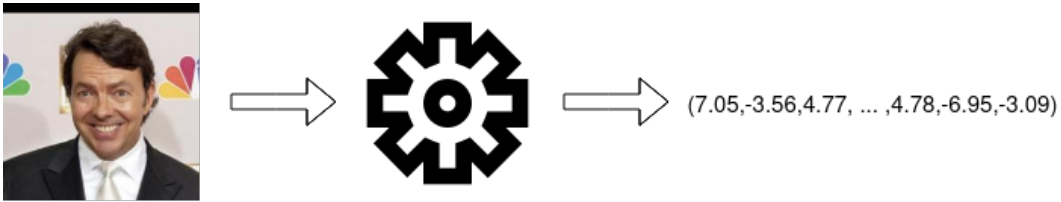
Variabilité des captures

Contrairement à la saisie d'un mot de passe qui ne tolère pas de fautes, chaque capture biométrique d'un individu est différente, mais espérée proche des autres.



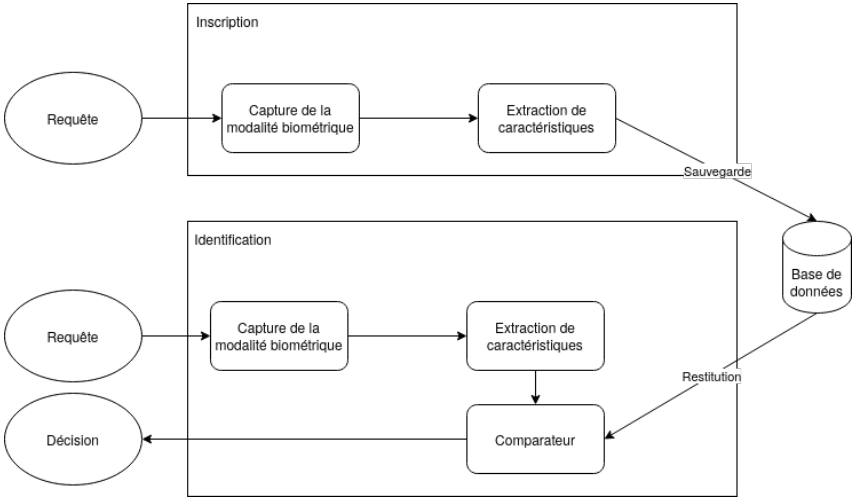
Reconnaître avec des données biométriques

Capture de la modalité $\Rightarrow$ Vecteur de caractéristiques
Réseau de neurones, filtres de Gabor, délimitation d'ondes



Introduction de la biométrie : concepts fondamentaux

Identification



Impératif de protection

Les données biométriques sont des données à caractère personnel et elles ne sont pas modifiables.

Impératif de protection

Les données biométriques sont des données à caractère personnel et elles ne sont pas modifiables.

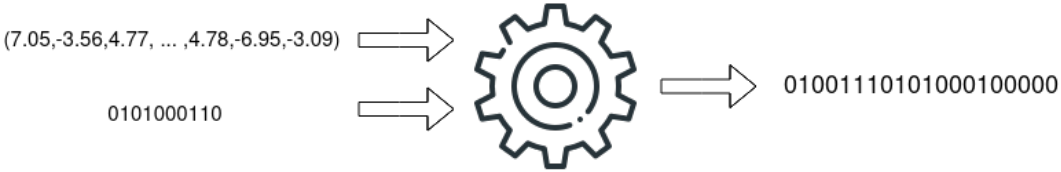
Protection $\Rightarrow$ Transformation du vecteur en un gabarit (version protégée).

Impératif de protection

Les données biométriques sont des données à caractère personnel et elles ne sont pas modifiables.

Protection $\Rightarrow$ Transformation du vecteur en un gabarit (version protégée).

- Transformation paramétrée par une graine.
- Gabarit : petit vecteur (binaire).

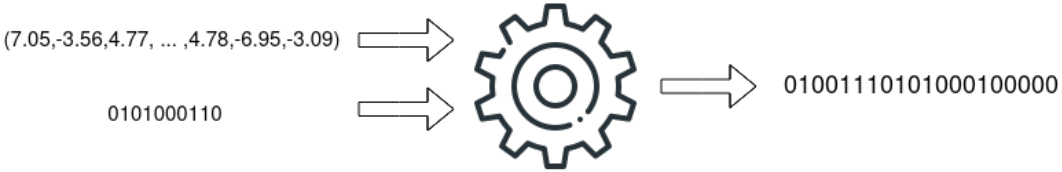


Impératif de protection

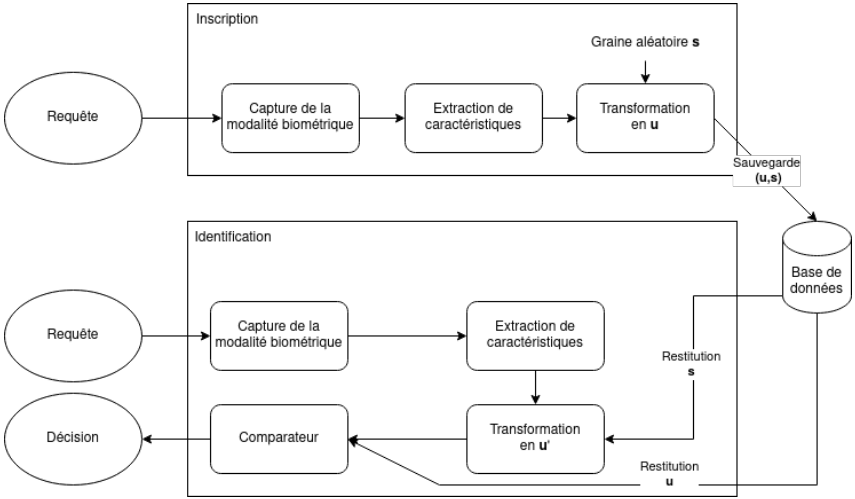
Les données biométriques sont des données à caractère personnel et elles ne sont pas modifiables.

Protection $\Rightarrow$ Transformation du vecteur en un gabarit (version protégée).

- Transformation paramétrée par une graine.
- Gabarit : petit vecteur (binaire).



- Révocable, non-inversible, performance, indistinguable.
- Comparaison dans le domaine transformé (comme pour les mots de passe).


Inscription
Graine aléatoire s
Requête
Capture de la modalité biométrique
Extraction de caractéristiques
Transformation en u
Sauvegarde (u,s)
Base de données
Identification
Requête
Capture de la modalité biométrique
Extraction de caractéristiques
Restitution s
Décision
Comparateur
Transformation en u'
Restitution u

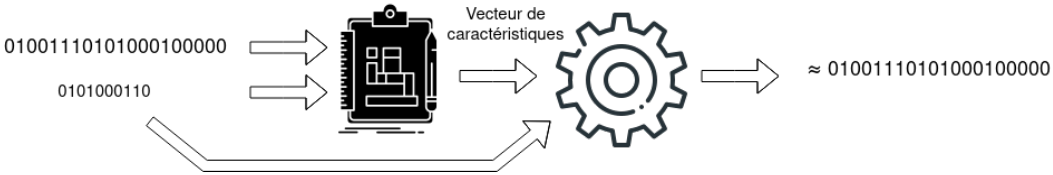
Plan

Description d'une préimage

Pour un couple (gabarit, graine), nous générons un vecteur qui, s'il est transformé avec cette graine, procure un gabarit proche.

Description d'une préimage

Pour un couple (gabarit, graine), nous générons un vecteur qui, s'il est transformé avec cette graine, procure un gabarit proche.

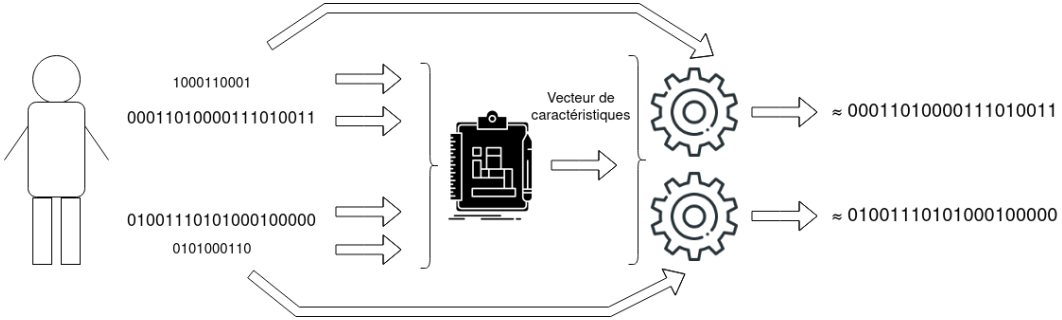


Description d'une PPR

Pour 2 couples (gabarit, graine), nous générons un vecteur qui, s'il est transformé avec les graines, procure des gabarits proches.

Description d'une PPR

Pour 2 couples (gabarit, graine), nous générons un vecteur qui, s'il est transformé avec les graines, procure des gabarits proches.

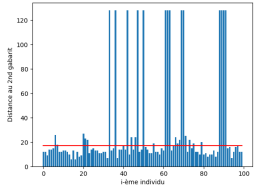


Stratégie

Pour 2 couples (gabarit, graine), nous construisons **aléatoirement** des vecteurs et nous testons s'ils sont des PPR.

Aléatoire

Base	Taux
Empreintes	68%

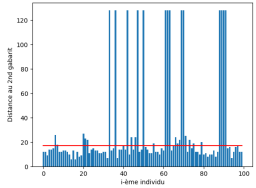


(a) Empreintes

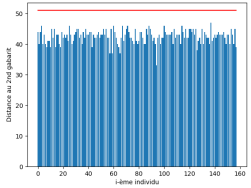
Figure: PPR aléatoires

Aléatoire

Base	Taux
Empreintes	68%
Visages	100%



(a) Empreintes

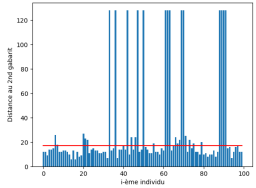


(b) Visages

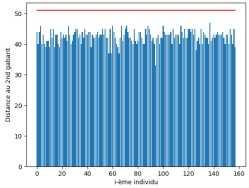
Figure: PPR aléatoires

Aléatoire

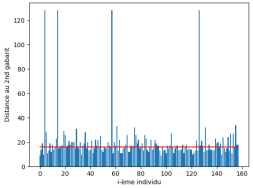
Base	Taux
Empreintes	68%
Visages	100%
ECG	42%



(a) Empreintes



(b) Visages



(c) ECG

Figure: PPR aléatoires

Stratégie

Pour 2 couples (gabarit, graine), nous construisons **avec un algorithme génétique** des vecteurs et nous testons s'ils sont des PPR.

Qu'est-ce qu'un algorithme génétique ?

- Algorithme d'optimisation : minimiser la valeur d'une fonction d'évaluation f en construisant son paramètre. Inspiré de la reproduction naturelle.

Qu'est-ce qu'un algorithme génétique ?

- Algorithme d'optimisation : minimiser la valeur d'une fonction d'évaluation f en construisant son paramètre. Inspiré de la reproduction naturelle.
- Population de taille n chromosomes = vecteurs ...

Qu'est-ce qu'un algorithme génétique ?

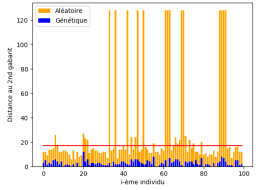
- Algorithme d'optimisation : minimiser la valeur d'une fonction d'évaluation f en construisant son paramètre. Inspiré de la reproduction naturelle.
- Population de taille n chromosomes = vecteurs ...
- ... évoluant sur t générations = itérations ...

Qu'est-ce qu'un algorithme génétique ?

- Algorithme d'optimisation : minimiser la valeur d'une fonction d'évaluation f en construisant son paramètre. Inspiré de la reproduction naturelle.
- Population de taille n chromosomes = vecteurs ...
- ... évoluant sur t générations = itérations ...
- dont $n/2$ parents se reproduisent pour donner $n/2$ enfants, avec croisement de leurs gènes et mutations aléatoires.

Génétique - Résultats des 3 bases

Base	Taux
Empreintes	100%

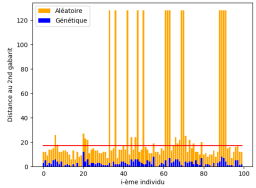


(a) Empreintes

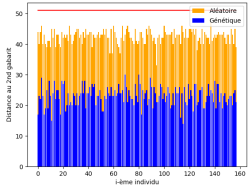
Figure: PPR génétiques

Génétique - Résultats des 3 bases

Base	Taux
Empreintes	100%
Visages	100%



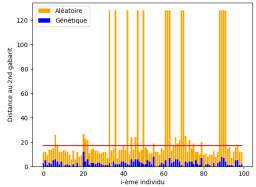
(a) Empreintes



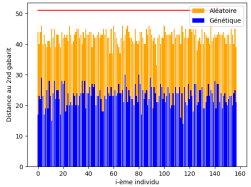
(b) Visages

Figure: PPR génétiques

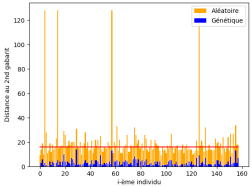
Base	Taux
Empreintes	100%
Visages	100%
ECG	100%



(a) Empreintes



(b) Visages



(c) ECG

Figure: PPR génétiques

Conclusion

- Nous obtenons des PPR dans 100% des cas pour les 3 bases.
- La proximité est bien inférieure à la méthode aléatoire, en temps similaire.

Conclusion

- Nous obtenons des PPR dans 100% des cas pour les 3 bases.
- La proximité est bien inférieure à la méthode aléatoire, en temps similaire.

Conclusion

- Nous obtenons des PPR dans 100% des cas pour les 3 bases.
- La proximité est bien inférieure à la méthode aléatoire, en temps similaire.

Pouvons-nous construire de manière similaire des vecteurs proches de plus de 2 couples ?

Plan

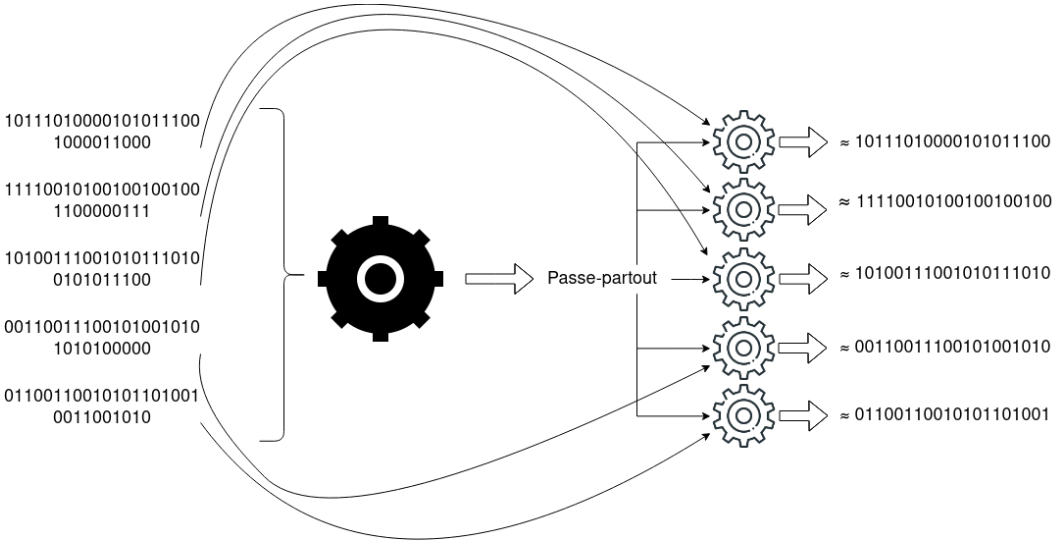
Vecteur proche de tous les couples (gabarit, graine) : un passe-partout ?

Motivation

Contourner un système de contrôle d'accès.
"Préimage universelle"

Du concept de préimage au concept de passe-partout

Description



Un vecteur passe-partout

Un vecteur x est un passe-partout pour B si pour chaque couple (gabarit, graine), le passe-partout est transformé avec la graine en un gabarit proche.

Un vecteur passe-partout

Un vecteur x est un passe-partout pour B si pour chaque couple (gabarit, graine), le passe-partout est transformé avec la graine en un gabarit proche.

Taux de couverture optimale (TCO)

TCO ϵ : B est dite ϵ-couverte par le passe-partout x s'il existe ϵn couples (gabarit, graine) pour lesquels la transformation de x est proche.

Un vecteur passe-partout

Un vecteur x est un passe-partout pour B si pour chaque couple (gabarit, graine), le passe-partout est transformé avec la graine en un gabarit proche.

Taux de couverture optimale (TCO)

TCO ϵ : B est dite ϵ-couverte par le passe-partout x s'il existe ϵn couples (gabarit, graine) pour lesquels la transformation de x est proche.

Taille optimale de dictionnaire (TOD)

TOD r : B est dite partitionnée par un ensemble de r passe-partout $\{x^1, \ldots x^r\}$ si pour chaque couple (gabarit, graine), un des r passe-partout au moins est transformé en un gabarit proche.

Du concept de préimage au concept de passe-partout

Résultats

Base	TCO (%)	TOD
Empreintes	73	5
Visages	15.2	18
ECG	61	12

Description

Quid de la couverture sur des données biométriques non utilisées pour la construction du passe-partout ?
2 ensembles : un utilisé à la construction, l'autre non.

Motivation

Il est peu probable d'avoir toutes les données d'une base pour attaquer cette même base.

Résultat

Les passe-partout conservent environ la moitié de leur couverture.

Base	TCO (%)	Report TCO (%)
Empreintes	42	73
Visages	6.3	15.2
ECG	44.3	61

À quoi ça sert ?

1. Construire un passe-partout depuis un premier lot de données biométriques transformées.
2. Obtenir un taux de couverture important sur un autre lot de données transformées.
3. Pouvoir tromper un contrôle d'accès.

Conclusion

- Jusqu'à 73% de couverture.
- Réutilisabilité : conserve environ la moitié de la couverture.

Conclusion

- Jusqu'à 73% de couverture.
- Réutilisabilité : conserve environ la moitié de la couverture.

Conclusion

- Jusqu'à 73% de couverture.
- Réutilisabilité : conserve environ la moitié de la couverture.

Pouvons-nous changer de prérequis pour augmenter le taux de couverture ?

Plan

Le vecteur passe-partout est fixé.

Le vecteur passe-partout est fixé.

Stratégie

Nous partons de la base de données biométriques et nous allons **choisir** les graines (force brute) pour construire la base de données biométriques révocables.

Le vecteur passe-partout est fixé.

Stratégie

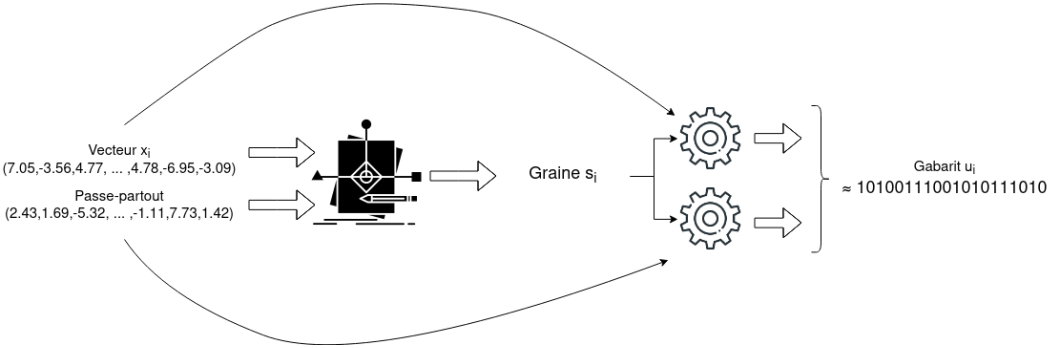
Nous partons de la base de données biométriques et nous allons **choisir** les graines (force brute) pour construire la base de données biométriques révocables.

Conséquences

- Être actif à la phase d'enrôlement.
- Cas d'usage éthique.

$\Rightarrow$ Objectif : augmenter le taux de couverture !

Pour chaque vecteur x_i, indépendamment des autres, nous choisissons une graine.
Algorithme en ligne / incrémental.



Indicateurs

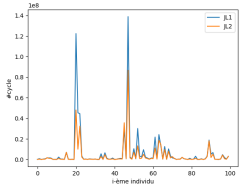
- TCO = 100%.
- TOD = 1.

Analyse

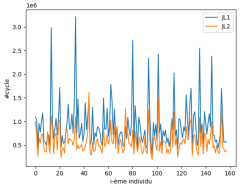
Nous avons largement augmenté le taux de couverture en un temps moindre.

Délégation de droits et passe-partout

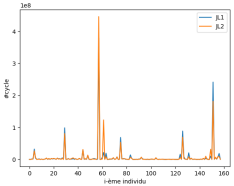
Résultats



(a) Empreintes



(b) Visages



(c) ECG

Figure: Durées de recherche

À quoi ça sert ?

- Éthique : possibilité de déléguer des droits sans modifier le circuit d'accès.

À quoi ça sert ?

- Éthique : possibilité de déléguer des droits sans modifier le circuit d'accès.
  $\rightarrow$ Au lieu d'ajouter une hiérarchie de droits sur les données biométriques, elles sont transformées en intégrant ces droits.

À quoi ça sert ?

- Éthique : possibilité de déléguer des droits sans modifier le circuit d'accès.
  $\rightarrow$ Au lieu d'ajouter une hiérarchie de droits sur les données biométriques, elles sont transformées en intégrant ces droits.
- Attaque : il faut être actif à la phase d'enrôlement et ajouter une porte dérobée sans code suspicieux et sans rendre la base suspecte.

À quoi ça sert ?

- Éthique : possibilité de déléguer des droits sans modifier le circuit d'accès.
  $\rightarrow$ Au lieu d'ajouter une hiérarchie de droits sur les données biométriques, elles sont transformées en intégrant ces droits.
- Attaque : il faut être actif à la phase d'enrôlement et ajouter une porte dérobée sans code suspicieux et sans rendre la base suspecte.
  $\rightarrow$ Elle est intégrée dans les données transformées.

À quoi ça sert ?

- Éthique : possibilité de déléguer des droits sans modifier le circuit d'accès.
  $\rightarrow$ Au lieu d'ajouter une hiérarchie de droits sur les données biométriques, elles sont transformées en intégrant ces droits.
- Attaque : il faut être actif à la phase d'enrôlement et ajouter une porte dérobée sans code suspicieux et sans rendre la base suspecte.
  $\rightarrow$ Elle est intégrée dans les données transformées.

À quoi ça sert ?

- Éthique : possibilité de déléguer des droits sans modifier le circuit d'accès.
  $\rightarrow$ Au lieu d'ajouter une hiérarchie de droits sur les données biométriques, elles sont transformées en intégrant ces droits.
- Attaque : il faut être actif à la phase d'enrôlement et ajouter une porte dérobée sans code suspicieux et sans rendre la base suspecte.
  $\rightarrow$ Elle est intégrée dans les données transformées.

L'individu dont est issu le vecteur passe-partout
souhaite couvrir la base plus tard !
$\rightarrow$ Étudier la couverture de ses futures captures (et les améliorer).

Individu passe-partout

Quid des autres captures de la personne dont est issu le vecteur passe-partout ?

Individu passe-partout

Quid des autres captures de la personne dont est issu le vecteur passe-partout ?

Ensemble de recherche

Le premier sous-ensemble de vecteurs est utilisé pour la recherche de graines.

Individu passe-partout

Quid des autres captures de la personne dont est issu le vecteur passe-partout ?

Ensemble de recherche

Le premier sous-ensemble de vecteurs est utilisé pour la recherche de graines.

Ensemble de test

Le second sous-ensemble de vecteurs, non utilisé pour la recherche de graines, permet de vérifier la performance de ”l’individu passe-partout”.

Présentation des résultats

Courbes cumulées décroissantes.
Objectif $\rightarrow$ Faire glisser la courbe rouge à droite de l'orange.

50% des vecteurs de test couvrent au moins 50% de la base ($T = 1$).

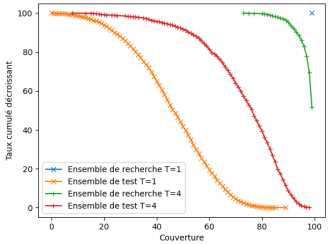


Figure: Empreintes

Présentation des résultats

Courbes cumulées décroissantes.
Objectif → Faire glisser la courbe rouge à droite de l'orange.

50% des vecteurs de test couvrent au moins 50% de la base ($T = 1$).
→ 50% des vecteurs de test couvrent au moins **80%** de la base ($T = 4$).

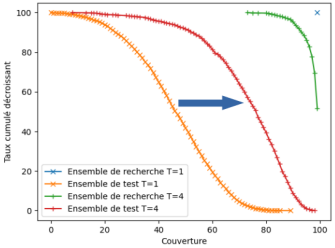
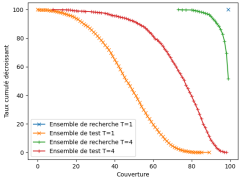


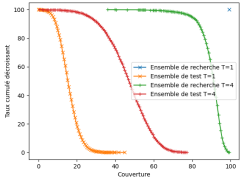
Figure: Empreintes

Délégation de droits et passe-partout

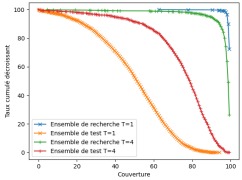
Extension - Résultats



(a) Empreintes



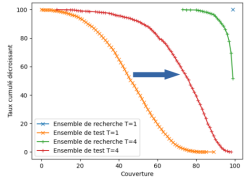
(b) Visages



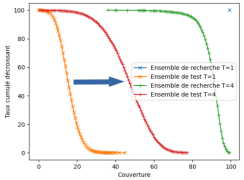
(c) ECG

Délégation de droits et passe-partout

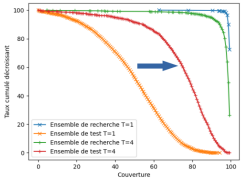
Extension - Résultats



(d) Empreintes



(e) Visages



(f) ECG

Où en est-on ?

- Nous souhaitons que de futures captures persistent à couvrir la base.
- Utiliser plusieurs vecteurs pour choisir la graine permet d'améliorer cette future couverture.

Où en est-on ?

- Nous souhaitons que de futures captures persistent à couvrir la base.
- Utiliser plusieurs vecteurs pour choisir la graine permet d'améliorer cette future couverture.

Pouvons-nous mieux choisir les vecteurs utilisés pour la recherche de graine et encore améliorer cette couverture ?

Où en est-on ?

- Nous souhaitons que de futures captures persistent à couvrir la base.
- Utiliser plusieurs vecteurs pour choisir la graine permet d'améliorer cette future couverture.

Pouvons-nous mieux choisir les vecteurs utilisés pour la recherche de graine et encore améliorer cette couverture ?

Corrélation

Corrélation bonne couverture de l'ensemble de recherche - bonne couverture de l'ensemble de test ?

Délégation de droits et passe-partout

Extension - Corrélation des performances

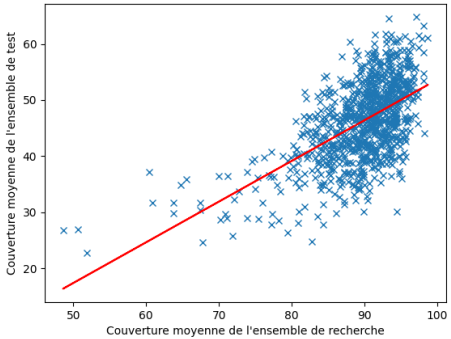


Figure: Visages

Conséquences

1. Utiliser plusieurs vecteurs pour la recherche de graines améliore la couverture d'autres vecteurs.
2. Il faut sélectionner les vecteurs utilisés pour la recherche de graines selon leur couverture pour maximiser la couverture d'autres captures.

Plan

Biométrie

- Reconnaître.
- Captures de modalités biométriques.
- Extraction de caractéristiques.
- Comparaisons.
- Données personnelles $\Rightarrow$ Protection, utilisée dans ces travaux !

Préimage ...

- ... proche (s'authentifiant avec un gabarit) ...
- ... et réutilisable (s'authentifiant pour un autre gabarit).

Outil

- Algorithme génétique.
- 100% de PPR.

Passe-partout : scénario 1

- Construction d'un passe-partout ...
- ... depuis une base de données biométriques révocables ...
- ... usurpant un maximum de gabarits !
- Partitionnement avec plusieurs passe-partout.

Résultats

- Jusqu'à 73% de couverture.
- Réutilisabilité : conserve environ la moitié de la couverture.

Passe-partout : scénario 2

- Construction d'une base de données biométriques révocables ...
- ... en choisissant les graines pour un passe-partout.

Passe-partout : scénario 2

- Construction d'une base de données biométriques révocables ...
- ... en choisissant les graines pour un passe-partout.

100% de couverture (temps variables).

Passe-partout : scénario 2

- Construction d'une base de données biométriques révocables ...
- ... en choisissant les graines pour un passe-partout.

100% de couverture (temps variables).

Individu passe-partout

- Utilisation de plusieurs vecteurs.
- Amélioration de la couverture de futurs vecteurs.
- Corrélation de couvertures : choisir de bons vecteurs.

Finalement

- Inverser 2 gabarits en une seule préimage.
- Contourner un système de contrôle d'accès en construisant un passe-partout à partir de données similaires.
- Déléguer des droits ou introduire une porte dérobée au sein des données transformées.

[Merci]

Références

- Biometric masterkeys, Computers & Security, Volume 116, 2022, 102642
- Thèse : Passe-partout biométriques. ⟨NNT : 2022NORMC240⟩. ⟨tel-03899668⟩.
- Construction et analyse de passe-partout biométriques, SSTIC, 2023

Questions ?