

Reconnaissance du genre par analyse de dynamique de frappe au clavier sur texte libre

Romain Giot
et Christophe Rosenberger
Université de Caen, UMR 6072 GREYC
ENSICAEN, UMR 6072 GREYC
CNRS, UMR 6072 GREYC
Email : romain.giot@ensicaen.fr
Email : christophe.rosenberger@ensicaen.fr

Bernadette Dorizzi
Institut Télécom
Télécom SudParis
Email : bernadette.dorizzi@it-sudparis.eu

Résumé

Détecter, de façon automatique, le genre d'un individu lors de l'utilisation d'un ordinateur peut amener de nouveaux types d'applications (à la fois pour le confort de l'utilisateur, et pour sa sécurité). Dans cette communication, nous présentons une étude originale montrant qu'il est possible de détecter le genre d'un individu ayant saisi une phrase de texte quelconque. Les informations utilisées pour détecter le genre sont identiques aux informations utilisées en reconnaissance biométrique par dynamique de frappe au clavier. Le taux de reconnaissance d'un tel système est proche de 75%. Les résultats sont prometteurs et ouvrent la voie à une nouvelle utilisation de la dynamique de frappe au clavier.

1. Introduction

La dynamique de frappe au clavier est une modalité biométrique comportementale permettant d'authentifier des individus selon leur façon de taper au clavier [1]. Associée à la reconnaissance d'un mot de passe, la vérification statique permet d'obtenir un mécanisme d'authentification forte afin de vérifier l'identité d'un individu lors de sa connexion à un service (vérification de la correspondance du mot de passe et vérification de la façon de le saisir). Dans le cas de la vérification dynamique, le système est capable d'identifier ou de vérifier l'identité d'un individu suivant sa façon de saisir un texte quelconque : il est ainsi possible de détecter une intrusion physique d'un ordinateur (un utilisateur illégitime) en fonctionnement.

Il a également été montré que la dynamique de frappe au clavier peut fournir des informations de type « biométrie douce » [2]. De telles informations ne sont pas suffisamment discriminantes pour vérifier les individus, mais elles sont suffisantes pour pouvoir les catégoriser. Epp *et coll.* ont montré qu'il est possible de détecter l'état émotionnel d'un individu à partir de sa façon de taper au clavier [3]. Nous avons montré récemment qu'il est possible de détecter le genre d'un individu suivant sa façon de saisir un mot de passe fixé à l'avance [4]. Cette information peut également être utilisée pour améliorer les performances de vérification de systèmes existants. À notre connaissance, il n'existe pas d'autre type d'informations descriptibles étudiées dans la littérature de la dynamique de frappe au clavier.

Dans cet article, nous montrons qu'il est également possible de détecter le genre d'un individu, grâce à sa façon de taper au clavier, à partir d'un texte quelconque. Le fait de ne pas être obligé d'utiliser une phrase fixée par le système de reconnaissance augmente le nombre potentiel d'applications. Nous pouvons imaginer : d'utiliser un tel système pour surveiller des réseaux de discussion afin de détecter automatiquement des dérives potentielles, d'adapter automatiquement l'interface graphique d'un programme informatique si il est susceptible d'être utilisé différemment par les hommes et les femmes. Comme pour la reconnaissance sur texte fixe, on peut imaginer que cette information puisse être utilisée pour améliorer les performances d'un système d'authentification biométrique dynamique basée sur la frappe au clavier.

La contribution majeure de ce papier est la démonstration de la faisabilité de la reconnaissance du genre sur texte libre, chose qui n'a jamais été étudiée (à notre

connaissance) dans la littérature.

Le papier est organisé de la façon suivante. La section 2 présente l'état de l'art de la biométrie douce. La section 3 présente les données employées dans cette étude. La section 4 présente la méthode de reconnaissance développée. La section 5 présente les différents résultats obtenus. La section 6 analyse les résultats obtenus, tandis que la section 7 conclue cette communication.

2. État de l'art de la « biométrie douce »

Bien que la biométrie douce soit un type de biométrie intéressante, il y a peu de travaux sur cette technologie. Ce type de biométrie peut être qualifiée comme [2] : « un caractère fournissant de l'information sur l'individu, mais manquant d'unicité et permanence pour différencier suffisamment deux individus. »

Les informations peuvent être mesurées grâce à un capteur spécifique, déterminées de façon automatique à l'aide d'un système de reconnaissance de formes ou définies par un opérateur. Nous présentons ces trois types de définitions depuis l'état de l'art.

2.1. Utilisation de capteurs spécifiques

Jain *et coll.* présentent un framework générique permettant d'utiliser la biométrie douce dans un système de reconnaissance biométrique à l'aide des probabilités bayésiennes [2]. Ils obtiennent une amélioration de 5% pour les systèmes de reconnaissance d'empreinte digitale en utilisant des informations à propos de l'ethnie (indien ou pas), de l'âge et de la taille (artificiellement générés à l'aide d'une distribution gaussienne). Le problème principal de cette approche est la nécessité de disposer d'un matériel spécifique pour mesurer la taille de l'individu lors du processus d'authentification¹.

Ailisto *et coll.* utilisent le poids et le pourcentage de graisse dans le corps comme information supplémentaire d'un système de reconnaissance par empreinte digitale [5]. Ils réduisent l'erreur totale de 3,9% à 1,5%. Le score de poids est calculé en effectuant la différence absolue entre le poids de référence et celui mesuré (une procédure identique est employée pour la graisse). Leur expérience montre que le poids peut être une mesure discriminante et utilisé comme une biométrie douce (utilisé seul, l'erreur est de 11%), mais ça n'est pas le cas de la mesure de graisse. Une fois de plus, le choix de la modalité sélectionnée peut être problématique en raison de la nécessité d'utiliser une balance spécifique qui n'est pas ergonomique.

1. Problème contourné lors des expériences en générant les informations de façon automatique

2.2. Détermination automatique

Dans les travaux suivants, les informations sont déterminées automatiquement par des systèmes d'apprentissage au lieu d'être mesurées par des capteurs spécifiques. Jain *et coll.* utilisent des marques sur la peau du visage (tâche de rousseur, grain de beauté, cicatrice) comme biométrie douce [6]. L'intérêt de ces informations est multiple :

- ajouter des données extraites au système de reconnaissance faciale ;
- permettre une recherche plus rapide dans les grandes bases de visage ;
- faciliter la reconnaissance depuis des images partielles avec des marques.

La performance d'identification au rang 1 a été améliorée sur deux jeux de données significatifs.

Les biométries douces peuvent également être utilisées pour localiser des individus dans des bases de vidéo-surveillance [7]. Dans ce contexte, Vaquero *et coll.* ne sont pas intéressés par le système de reconnaissance faciale, mais ils veulent pouvoir trouver des individus en fonction d'une requête spécifique (par exemple : « trouve moi les individus chauves, étant entrés dans le bâtiment samedi dernier, et portant un pull rouge et des lunettes de soleil »). Les informations utilisées sont des attributs visuels tels que le type de chevelure, la présence ou non de lunettes (et leur type), le type de barbe, la couleur des habits. Certains de ces attributs peuvent être classés en tant que biométrie douce.

Le genre d'un individu peut être déterminé depuis des images de visage [8]. Alexandre présente un système fusionnant différents classificateurs utilisant des descripteurs de forme, texture et couleur d'images faciales à différentes échelles. Le système a été validé sur deux jeux de données importants et le taux de reconnaissance avoisine les 90%.

Hwang *et coll.* proposent un système de reconnaissance faciale utilisant des informations de genre afin d'améliorer les performances de reconnaissance [9]. Dans leur travail, ils partent du principe que le système de reconnaissance de genre fonctionne sans erreur. Ils utilisent trois différents modèles de visages (en utilisant LDA) : un pour les hommes, un pour les femmes, un pour les deux genres. Le système de reconnaissance utilise le modèle spécifique au genre et le modèle général. Ces travaux indiquent que la reconnaissance des femmes est plus difficile que la reconnaissance des hommes et qu'utiliser l'information du genre permet d'augmenter les performances. Les travaux sont validés sur FRGC.

Plutôt que d'utiliser une approche ensembliste pour prédire le genre d'un individu à partir de son visage, Li *et coll.* utilisent une approche basée sur des informations locales et globales en utilisant un modèle de mélange de gaussiennes spatiales (Spatial Gaussian Mixture Models) [10]. Cette approche permet d'obtenir de meilleures performances (amélioration entre 40% et 50%) qu'en utilisant des GMM et SVM sur la base de visage YGA.

2.3. Travaux en dynamique de frappe

Les travaux précédents ont montré qu'utiliser des informations de type biométrie douce permet d'augmenter le taux de reconnaissance des systèmes biométriques. Cette nouvelle information peut être mesurée à l'aide d'un capteur ou extraite depuis la donnée biométrique. La plupart des papiers de l'état de l'art ont travaillé sur la reconnaissance du genre sur des visages 2D. Une telle solution n'est pas utilisable dans tous les contextes (ordinateur sans webcam, environnement inapproprié, ...). Il peut donc être intéressant de voir si de telles informations sont également utiles dans le cas de la dynamique de frappe au clavier.

Il existe peu de travaux concernant la biométrie douce pour dynamique de frappe au clavier. Epp *et coll.* ont montré qu'il est possible de détecter l'état émotionnel d'un individu suivant sa façon de taper au clavier [3]. Il est possible de déterminer la colère et l'excitation dans 84% des cas.

Nous avons montré récemment qu'il est possible de détecter le genre d'un individu suivant sa façon de saisir un mot de passe fixé à l'avance [4]. Le taux de reconnaissance du genre est supérieur à 90% et l'utilisation de cette information en association avec un système de reconnaissance permet de diminuer la valeur du EER de 20%.

Dans cette communication, nous nous sommes intéressés à la détermination automatique du genre uniquement. Nous n'utilisons pas cette donnée additionnelle pour aider lors de la vérification de l'individu. La section suivante présente la méthode développée.

3. Données disponibles

Dans cette section, nous présentons : la base de dynamique de frappe utilisée, les données que nous avons extraites depuis cette base.

3.1. Base de données

Bello *et coll.* ont récemment diffusé une base de données de dynamique de frappe au clavier de texte

fixe [14]. Plusieurs volontaires ont saisi 15 paragraphes de texte espagnol, et 15 lignes de commandes shell UNIX courantes dans une application WEB. Pour chacun des événements claviers, l'application capture les informations suivantes :

- le code de la phrase en cours de saisie ;
- l'instant de l'évènement clavier (timestamp UNIX en millisecondes) ;
- le type d'évènement (pression ou relâchement de la touche) ;
- le code ascii de la touche actionnée.

Pour des raisons techniques liées à l'utilisation de Javascript, certaines touches ne sont pas considérées. Les erreurs de saisies sont acceptées à partir du moment où la saisie d'une phrase contient entre 90% et 130% du nombre de caractères attendus. Pour plus d'informations, se reporter à l'article original [14].

Dans nos travaux, nous utilisons les 54 volontaires ayant participé intégralement à l'étude. Chacun d'entre eux a participé à une session : chaque utilisateur n'a saisi qu'une seule fois chaque phrase. Nous ne travaillons qu'avec les données de type texte, et pas de type ligne de commande. Plusieurs méta-données sont associées à chaque utilisateur. Dans notre cas, nous nous intéressons au genre de l'utilisateur. Ainsi, la base dispose de 40 hommes et 15 femmes. Étant donnée que cette base contient différentes phrases, nous allons pouvoir nous en servir pour faire de la reconnaissance du genre sur texte libre.

3.2. Données extraites

Contrairement aux auteurs de la base de données, dans leur étude sur l'identification [15], nous n'avons pas filtré les données erronées (en se basant sur leur valeur temporelle). Nous considérons que de telles données sont toujours présentes dans des applications déployées en environnement opérationnel, et qu'il est nécessaire que les algorithmes de reconnaissance travaillent avec. Nous avons choisi d'utiliser plusieurs types de données extraites :

- l'intervalle de temps entre la pression de deux touches successives ;
- l'intervalle de temps entre le relâchement de deux touches successives ;
- le temps de pression d'une touche ;
- le délai d'attente entre le relâchement d'une touche et la pression de la touche suivante.

Pour chaque phrase saisie, nous calculons les différentes informations liées aux digraphes (deux touches successives). Parmi tous les digraphes potentiels, nous n'avons conservé que les digraphes présents au moins une fois dans chacune des phrases [14] : (« e », « a »,

« d », « de », « la », « l »). De même, nous ne conservons la durée de pression que pour les touches présentes dans les digraphes précédents (« », « a », « d », « e », « l »). Il faut noter que ces choix ont été fixés par la structure du texte de la base de données concernée. Il est possible que dans un corpus rédigé avec une langue différente, les digraphes les plus communs soient différents.

Lorsqu'un même type de temps est présent plusieurs fois dans une phrase, nous calculons sa moyenne. Ainsi, chaque phrase j de chaque utilisateur i est encodée par un vecteur $feat_j^i$ de dimension 23 :

$$feat_j^i = \{DD_d, DD_l, DD_a, DD_e, DD_{de}, DD_{la}, UU_d, UU_l, UU_a, UU_e, UU_{de}, UU_{la}, DU, DU_d, DU_l, DU_a, DU_e, UD_d, UD_l, UD_a, UD_e, UD_{de}, UD_{la}\} \quad (1)$$

avec DD_{ab} , l'intervalle de temps entre la pression de la touche a et la pression de la touche b :

$$DD_{ab} = date_pression(b) - date_pression(a) \quad (2)$$

UU_{ab} , l'intervalle de temps entre le relâchement de la touche a et la relâchement de la touche b :

$$UU_{ab} = date_relachement(b) - date_relachement(a) \quad (3)$$

DU_a , le temps de pression de la touche a :

$$DU_a = date_relachement(a) - date_pression(a) \quad (4)$$

UD_{ab} , le temps de latence entre la touche a et la touche b :

$$UD_{ab} = date_pression(b) - date_pression(a) \quad (5)$$

$date_pression(\cdot)$ et $date_relachement(\cdot)$ donnant respectivement la date (en milliseconde) de la pression et du relâchement de la touche concernée.

4. Méthode développée

4.1. Méthode de classification

Dans une étude précédente [4], nous avons montré qu'il est possible d'effectuer une reconnaissance du genre en dynamique de frappe au clavier à partir d'un texte fixe. Une des méthodes les plus efficaces consistaient à utiliser à Séparateur à Vaste Marge (SVM) [11] afin de calculer un score de reconnaissance (à comparer à un seuil pour décider du genre de l'utilisateur).

Pour cette étude, nous avons choisi d'utiliser une méthode similaire, étant donné que les descripteurs des phrases (voir la Section 3.2) sont relativement

homogènes (bien que les phrases soient hétérogènes) : ils ont le même nombre de dimension et chaque dimension représente la même chose, quel que soit le texte saisi.

Différentes variantes, présentées dans la sous-section suivante, sont testées. À chaque fois, un ensemble de descripteurs étiquetés (1 pour un texte saisi par un homme, -1 pour un texte saisi par une femme) est utilisé. Le nombre d'exemples pouvant être faible (54 exemples au minimum dans certains cas), nous avons utilisé le mécanisme de validation croisée Leave One Out. Avec un ensemble contenant n exemples, chaque exemple est utilisé successivement comme exemple de test, tandis que les $n - 1$ autres exemples servent pour l'apprentissage du modèle. De cette façon, nous pouvons disposer d'une quantité de données plus importante pour apprendre le modèle, sans avoir à la fois la même donnée dans l'ensemble d'apprentissage et l'ensemble de validation.

Nous travaillons avec des données normalisées par zscore. La moyenne μ et l'écart type σ de l'ensemble des descripteurs d'apprentissage est calculée. Chaque descripteur (celui de test et ceux d'apprentissage) est normalisé de la façon suivante :

$$feat_norm_j^i = \frac{feat_j^i - \mu}{\sigma} \quad (6)$$

Notons \mathcal{X} l'espace des descripteurs (constitué des données normalisées) et \mathcal{Y} les étiquettes (-1 ou 1) associées à cette ensemble. Apprendre le SVM va consister à trouver la fonction $y : \mathcal{X} \rightarrow \mathbb{R}$. La détermination de l'étiquette de l'exemple x est faite de la façon suivante : $l = sgn(y(x))$. La fonction $y(\cdot)$ permet de faire une séparation non linéaire basée sur un système de marges souple. Pour plus d'information, se reporter à [11].

Nous utilisons un noyau gaussien. Nous n'avons pas cherché à optimiser les paramètres nécessaires au fonctionnement du SVM : le paramètre γ du noyau et le paramètre C . Nous utilisons les valeurs par défaut proposées par la bibliothèque scikit-learn [12] (à savoir $C = 1$ et $\gamma = 1/23$).

Plutôt que d'utiliser l'étiquette prédite par le classifieur, nous utilisons la probabilité [13] que l'exemple testé soit saisi par un homme. De cette façon, nous obtenons un score plutôt qu'un booléen et nous pouvons calculer différents taux d'erreurs. Les résultats sont présentés à l'aide de la courbe ROC et de l'aire sous celle-ci. Nous utilisons également le taux d'erreur égale.

La figure 1 résume le fonctionnement du système de reconnaissance.

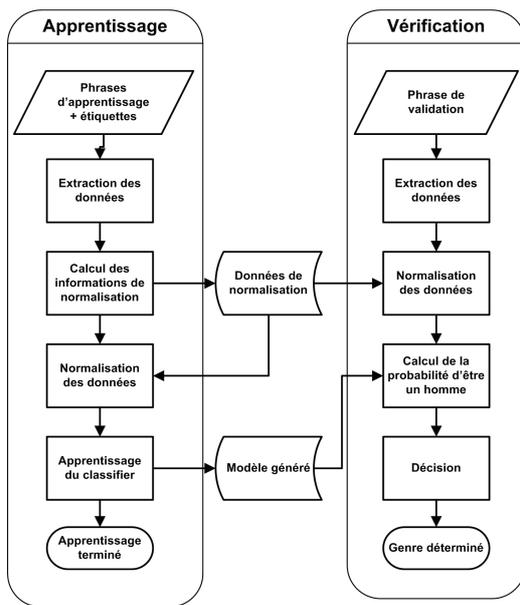


FIGURE 1. Résumé du fonctionnement du système de reconnaissance du genre par dynamique de frappe au clavier sur texte libre.

4.2. Scénarios de reconnaissance

Nous testons trois différents scénarios de configuration du mécanisme de reconnaissance du genre par dynamique de frappe :

- 1) en apprenant à reconnaître le genre à partir d'une phrase précise, puis en reconnaissant le genre uniquement sur d'autres saisies de la même phrase ;
- 2) en apprenant à reconnaître le genre à partir d'un ensemble de phrases quelconques, puis en reconnaissant le genre uniquement à partir d'une phrase inconnue ;
- 3) en utilisant la saisie de plusieurs phrases (au lieu d'une seule) pour reconnaître le genre d'un individu, le modèle ayant été appris avec d'autres phrases.

4.2.1. Reconnaissance sur phrase connue. Ce scénario est proche de la reconnaissance du genre par mot de passe [4] à la différence que dans ce cas :

- les descripteurs représentent des informations globales (la moyenne du temps de réalisation de quelques digraphes) plutôt que des informations locales (le temps entre chaque digraphe consécutif) ;
- la quantité d'information disponible est plus faible (descripteur de taille 23 au lieu de 60).

Le problème est donc plus compliqué. Dans ce scénario, nous lançons les calculs pour chaque phrase individuellement. Pour les 15 sous-scénarios possibles, nous disposons de 54 exemples (d'où la nécessité d'utiliser un mécanisme de cross-validation). Pour chaque exemple de test, le modèle est appris avec les données des autres utilisateurs ayant saisi la même phrase (il n'y a donc aucune donnée de l'individu de l'exemple de validation dans les exemples d'apprentissage).

Nous avons ajouté un 16^e sous-scénario (nommé global) qui travaille avec toutes les données possibles, soit $810 = 15 * 54$. Pour chaque exemple de test, le modèle est appris avec les données de tous les autres utilisateurs (y compris celles de la même phrase), et les données des autres phrases pour l'utilisateur ayant saisi l'exemple. Il permet de donner des performances générales en faisant abstraction de la phrase choisie.

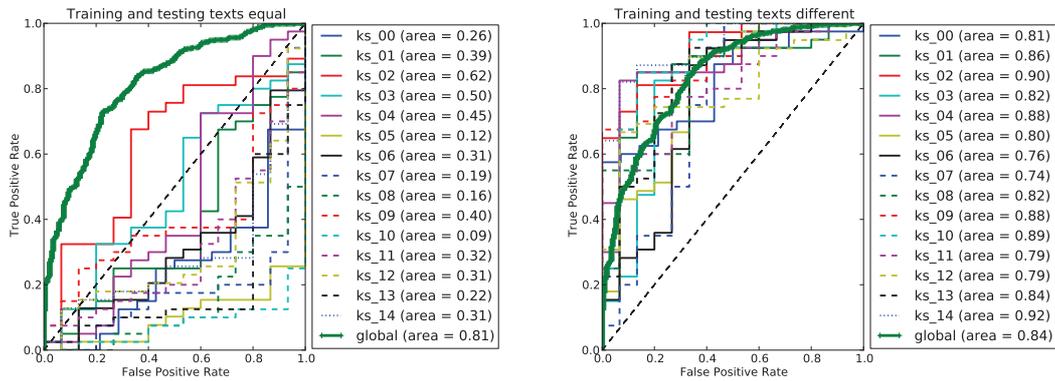
4.2.2. Reconnaissance sur phrase inconnue. Ce scénario nous semble plus proche d'un cas d'utilisation réel. Le modèle de reconnaissance de genre est appris à partir d'un ensemble de phrases, mais le test est effectué uniquement sur des phrases inconnues du modèle. Ce scénario nous semble le plus réaliste car il ne demande pas à créer un modèle exhaustif.

Pour les 15 sous-scénarios possibles, nous disposons de $756 = 14 * 54$ exemples d'apprentissage et de 54 exemples de test. Les exemples d'apprentissage étant fixés, nous n'utilisons pas de système de validation croisée.

Nous avons ajouté un 16^e scénario (nommé global) qui consiste à calculer les performances à partir de l'ensemble des scores des 15 sous-scénarios précédents. Il permet de donner des performances générales en faisant abstraction de la phrase choisie.

4.2.3. Utilisation de plusieurs phrases. Utiliser une seule phrase pour détecter le genre d'un individu peut ne pas être suffisant. Nous avons testé une méthode prenant en compte la saisie de plusieurs phrases différentes et testé l'évolution de la performance en fonction du nombre de phrases saisies. Cela permet de tester un scénario où il n'est pas nécessaire de disposer instantanément de l'information du genre de l'individu.

Pour évaluer ce scénario, nous avons utilisé le jeu de scores du scénario « Reconnaissance sur phrase inconnue ». Pour une reconnaissance à N phrases, pour chaque utilisateur, nous avons effectué la fusion de toutes les combinaisons possibles de N scores parmi les 15 de l'utilisateur. La méthode de fusion effectuée est une moyenne des scores. Le nombre de scores après fusion peut donc être différent du nombre de scores initiaux (C_N^{15} au lieu de 15 pour chaque utilisateur).



(a) Même phrase utilisée lors de l'apprentissage et de la reconnaissance. (b) Phrases différentes utilisées lors de l'apprentissage et de la reconnaissance.

FIGURE 2. Courbe ROC de la méthode de reconnaissance en fonction de la phrase utilisée et du scénario appliqué.

TABLE 1. Aire sous la courbe pour la reconnaissance du genre, en fonction de la phrase sélectionnée et du scénario (phrase d'apprentissage et de vérification différente ou égale).

Phrase	Apprentissage/Validation		Moyenne
	Identique	Différent	
ks_01	0,34	0,81	0,575
ks_02	0,31	0,86	0,585
ks_03	0,66	0,90	0,780
ks_04	0,53	0,82	0,675
ks_05	0,37	0,88	0,625
ks_06	0,12	0,80	0,460
ks_07	0,31	0,76	0,535
ks_08	0,23	0,74	0,485
ks_09	0,13	0,82	0,475
ks_10	0,35	0,88	0,615
ks_11	0,10	0,89	0,495
ks_12	0,29	0,79	0,540
ks_13	0,26	0,79	0,525
ks_14	0,15	0,84	0,495
ks_15	0,29	0,92	0,605
global	0,81	0,84	0,825
Moyenne	0,33	0,83	

TABLE 2. Taux d'erreur égal (TEE) de reconnaissance global pour les deux scénarios.

	Phrases égales	Phrases différentes
TEE (%)	25.54	26.80

5. Résultats

La figure 2 présente la courbe ROC des différentes expériences, tandis que le tableau 1 présente l'aire

sous la courbe et le tableau 2 présente le taux d'erreur égal (taux d'erreurs lorsque le ratio d'hommes considérés comme étant des femmes est égal au ratio de femmes considérées comme étant des hommes) des approches globales. Plus le point d'inflexion de la courbe est proche du coin supérieur gauche, plus les performances de reconnaissance sont bonnes. La ligne diagonale pointillée présente un comportement aléatoire. Si des courbes sont en dessous de la ligne, elles représentent des systèmes encore plus mauvais que le hasard. Plus l'aire sous la courbe est importante, plus les performances de reconnaissance sont bonnes.

Depuis ces résultats, nous pouvons faire les conclusions suivantes :

- Les performances sont grandement différentes d'une phrase à l'autre : la reconnaissance du genre fonctionne mieux sur certaines phrases (par exemple, la phrase «ks_07» donne toujours de moins bonnes performances que la phrase «ks_02»). Cela peut s'expliquer par la différence de distribution des digraphes (par exemple, la phrase «ks_07» comporte plus de doublons de digraphes que la phrase «ks_02», il y a plus de données moyennées).
- Contrairement à nos attentes, les performances obtenues avec l'utilisation des mêmes phrases en apprentissage et validation donnent de moins bons résultats que l'apprentissage et la validation sur des phrases différentes. Nous pensons que nous obtenons un tel résultat car la quantité de données est faible (à peine une cinquantaine au lieu de plusieurs centaines) lors de l'apprentissage du modèle. Les données n'étant pas suffisamment

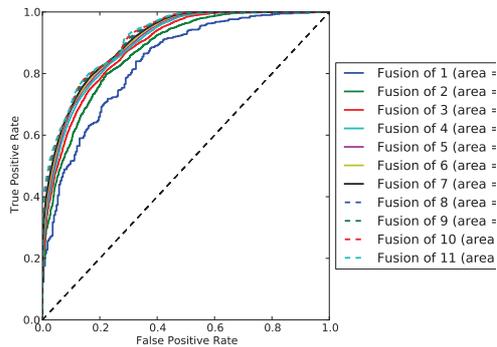


FIGURE 3. Courbe ROC en fonction du nombre de phrases utilisées pour la reconnaissance du genre.

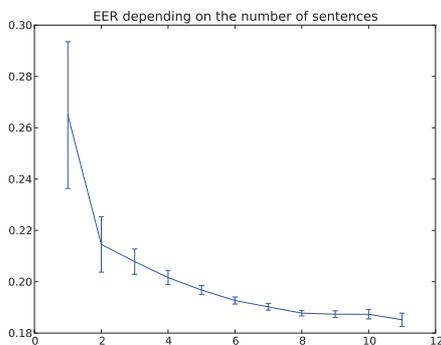


FIGURE 4. TEE (et son intervalle de confiance) en fonction du nombre de phrases utilisées pour la reconnaissance du genre.

représentatives, le modèle ne se spécialise pas suffisamment et les performances sont mauvaises. Le fait que le sous-système global donne de bonnes performances appuie ce point.

- La reconnaissance du genre à partir de texte inconnu donne des résultats acceptables : aire sous la courbe de 0,84 lorsque les phrases utilisées lors de la reconnaissance sont inconnues.

Ainsi, à partir de la saisie d'une unique phrase, nous pouvons détecter le genre d'un individu dans près de 75% des cas. Ces résultats sont inférieurs à la reconnaissance de 90% à partir de visage 2d, mais ils restent encourageants.

La figure 3 et la figure 4 représentent respectivement : la courbe ROC en fonction du nombre de phrases utilisées pour reconnaître le genre, la valeur du taux de reconnaissance égale en fonction du nombre de phrases utilisées pour la reconnaissance (ainsi que

la largeur de l'intervalle de confiance à 90% calculé avec une méthode de bootstrapping à 1000 ré-échantillonnages).

Les résultats montrent que :

- Plus le nombre de phrases utilisées est important, plus les résultats sont fiables.
- Le gain de reconnaissance décroît exponentiellement avec le nombre de phrases.

6. Discussion

Cette étude a montré la faisabilité d'un système de reconnaissance du genre basé sur la saisie de texte libre. L'étude comporte quelques limitations qu'il serait utile de prendre en compte dans des travaux futurs :

- Chaque utilisateur n'a saisi qu'une seule fois chaque phrase. Il est possible que plusieurs saisies apporteraient plus de variabilités (pouvant augmenter ou décroître les performances de reconnaissance).
- Dans certains scénarios, le même utilisateur peut être présent à la fois dans la base d'apprentissage et dans la base de validation (avec des phrases différentes, par contre). Il pourrait être utile de comparer les résultats actuels à une étude similaire prenant soin de ne pas inclure de données de l'utilisateur testé dans l'ensemble d'apprentissage. Cette nouvelle étude serait un scénario supplémentaire, où la validation est toujours faite sur des phrases inconnues, mais également sur des utilisateurs inconnus.
- Le texte est en espagnol. Les digraphes courants et présents dans tous les textes sont peut être spécifiques à cette langue. Il est possible que les conclusions varient d'une langue à l'autre.
- Nous avons vu qu'il existe une différence de performance, quelque soit le scénario, en fonction de la phrase choisie. Il pourrait être intéressant d'analyser les phrases afin de comprendre pourquoi de telles différences sont constatées.

Tous ces points pourraient être étudiés dans de futurs travaux.

7. Conclusion

La dynamique de frappe au clavier permet de vérifier l'identité d'individus selon leur façon de saisir de texte sur un clavier. Une étude précédente a également mis en évidence la possibilité de reconnaître le genre d'un individu en fonction de sa façon de taper un texte court prédéfini.

Dans cette communication, nous montrons qu'il est également possible de reconnaître le genre d'un individu selon sa façon de taper un texte non connu du système. Dans ce nouveau cas d'utilisation, le texte est plus long qu'un mot de passe, étant donné qu'il s'agit d'une phrase. Mais, en contrepartie, le genre est reconnu à partir d'un texte quelconque : les cas d'utilisations d'un tel système se démultiplient étant donné qu'il n'est pas nécessaire de créer un modèle par texte potentiellement saisissable.

Nous pouvons désormais envisager de détecter le genre d'un individu, saisissant un texte libre, à partir du moment où il a saisi suffisamment de texte (une phrase suffit pour obtenir un taux de bonne reconnaissance de près de 75%). Nous pensons qu'une telle application pourrait être utile dans des applications de détection de fraude ou de configuration automatique d'interface graphique ou de surveillance de réseaux sociaux.

De futurs travaux pourraient porter l'amélioration des performances de reconnaissance et sur la reconnaissance d'autres caractéristiques douces (tel que l'âge de l'individu) afin de caractériser de façon automatique un individu saisissant du texte sur un clavier standard.

Remerciements

Les auteurs tiennent à remercier Bello *et coll.* pour avoir rendu publique leur base de données, sans laquelle une telle étude n'aurait pas été possible.

Références

- [1] R. Giot, M. El-Abed, and C. Rosenberger, "Keystroke dynamics overview," in *Biometrics / Book 1*, D. J. Yang, Ed. InTech, Jul. 2011, vol. 1, ch. 8, pp. 157–182. [Online]. Available : <http://www.intechopen.com/articles/show/title/keystroke-dynamics-overview>
- [2] A. Jain, S. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," in *Proceedings of International Conference on Biometric Authentication*, 2004.
- [3] C. Epp, M. Lippold, and R. Mandryk, "Identifying emotional states using keystroke dynamics," in *Proceedings of the 2011 annual conference on Human factors in computing systems*, 2011, pp. 715–724.
- [4] R. Giot and C. Rosenberger, "A new soft biometric approach for keystroke dynamics based on gender recognition," *International Journal of Information Technology and Management (IJITM). Special Issue on : "Advances and Trends in Biometrics by Dr Lidong Wang*, vol. 11, no. 1/2, pp. 35–49, 2012.
- [5] H. Ailisto, E. Vildjiounaite, M. Lindholm, S.-M. Mäkelä, and J. Peltola, "Soft biometrics—combining body weight and fat measurements with fingerprint biometrics," *Pattern Recognition Letters*, vol. 27, no. 5, pp. 325 – 334, 2006.
- [6] A. Jain and U. Park, "Facial marks : Soft biometric for face recognition," in *IEEE International Conference on Image Processing (ICIP)*, 2009.
- [7] D. Vaquero, R. Feris, D. Tran, L. Brown, A. Hampapur, and M. Turk, "Attribute-based people search in surveillance environments," in *IEEE Workshop on Applications of Computer Vision (WACV)*, 2009.
- [8] L. A. Alexandre, "Gender recognition : a multiscale decision fusion approach," *Pattern Recognition Letters*, "February" 2010.
- [9] W. Hwang, H. Ren, H. Kim, S.-C. Kee, and J. Kim, "Face recognition using gender information," in *IEEE 16th International Conference on Image Processing (ICIP 2009)*. Cairo, Egypt : IEEE Signal Society, Nov. 2009.
- [10] Z. Li, X. Zhou, and T. S. Huang, "Spatial gaussian mixture model for gender recognition," in *IEEE 16th International Conference on Image Processing (ICIP 2009)*. Cairo, Egypt : IEEE Signal Society, Nov. 2009.
- [11] V. Vapnik, *Statistical learning theory*. Wiley New York, 1998.
- [12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and D. E., "Scikit-learn : Machine Learning in Python ;," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [13] C.-C. Chang and C.-J. Lin, "LIBSVM : A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27 :1–27 :27, 2011, software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [14] L. Bello, M. Bertacchini, C. Benitez, J. Carlos, Pizzoni, and M. Cipriano, "Collection and publication of a fixed text keystroke dynamics dataset," in *XVI Congreso Argentino de Ciencias de la Computacion (CACIC 2010)*, 2010.
- [15] M. Bertacchini, C. Benitez, and P. Fierens, "User clustering based on keystroke dynamics," in *XVI Congreso Argentino de Ciencias de la Computacion (CACIC 2010)*, 2010.