Tatouage Robuste et Réversible pour la Traçabilité de Bases de Données Relationnelles en Santé

Javier Franco Contreras, Gouenou Coatrieux, Frédéric Cuppens, Nora Cuppens-Bouhlahia, Emmanuel Chazard, Christian Roux

Tatouage Robuste et Réversible pour la Traçabilité de Bases de Données Relationnelles en Santé

J. Franco-Contreras[1], G. Coatrieux[1], F. Cuppens[2], N. Cuppens-Boulahia[2], E. Chazard[3], C. Roux[1]

1. Institut Mines-TELECOM, TELECOM Bretagne, U INSERM 1101 LaTIM, Brest 29238, France
2. Institut Mines-TELECOM, TELECOM Bretagne, UMR CNRS 3192 Lab-STICC, Rennes 35576, France
3. Dép. Santé Publique, CHU Lille; UDSL EA 2694; Univ Lille Nord de France; F-59000 Lille, France.

Résumé— **Dans cet article, nous proposons un schéma de tatouage avec pour objectif la traçabilité des bases de données de santé. Celui-ci s'appuie sur une modulation de tatouage réversible et robuste proposée à l'origine par De Vleeschouwer *et al.* pour les images et est fondée sur une transformation bijective d'histogrammes circulaires. Ici, nous adaptons cette modulation aux attributs numériques des bases à protéger. Afin d'identifier ou vérifier l'origine de bases regroupées au sein d'un entrepôt de données ou plus simplement mélangées en une seule base, nous proposons de dissimuler une marque qui identifie le fournisseur de la base. Nous montrons expérimentalement l'adéquation de notre schéma à ce problème de traçabilité sur la base d'un extrait de la base nationale PMSI-MCO 2011.**

Mots-clés
Tatouage, Bases de Données de Santé, Sécurité de l'information Médicale

I. INTRODUCTION

Le partage de bases de données de santé sert différents objectifs de la gestion des soins aux évaluations médico-économiques en passant par la surveillance sanitaire. De plus en plus de professionnels de santé sont concernés et mettent à disposition l'information par le biais d'entrepôts de données ou encore sur le *cloud* à des fins de stockage et d'analyse. Bien évidemment, les besoins en termes de sécurité sont dans le même temps accrus, notamment en matière de traçabilité. Dans un tel contexte, pouvoir identifier des jeux de données scindés en une seule et même base ou entrepôt de données offre des perspectives intéressantes tant pour tracer les données d'un fournisseur que localiser une fuite d'information.
Pour ce type de problème, le tatouage est une solution pertinente [1]. Dans son principe, il s'appuie sur une distorsion contrôlée des données à protéger pour y dissimuler un message qui peut ensuite aider à vérifier si elles ont été distribuées illégalement. Pour une image, le message est inséré en modulant de manière imperceptible ses niveaux de gris. Dans le cas d'une base de données, ce sont le plus souvent les valeurs de quelques attributs qui sont modifiées.
A l'instar des documents multimédia, le contrôle de la distorsion se pose également pour les bases de données. Pour y parvenir, une alternative est de s'appuyer sur le tatouage réversible. La propriété de réversibilité garantie la possibilité de récupérer la base de données originale une fois la marque retirée [2] [3].
Dans la suite de cet article, nous décrivons en section II l'architecture du système proposé ainsi que la modulation de tatouage réversible utilisée. En section III, nous montrons expérimentalement les performances de ce schéma en termes de traçabilité, dans le cas d'un extrait de la base nationale PMSI-MCO 2011 qui décrit la totalité des séjours d'hospitalisation terminés en 2011 en France. La section IV conclue cet article et donne quelques perspectives.

II. SYSTÈME PROPOSÉE

II.1. Architecture du système

Une base de données relationnelle est composée d'un ensemble fini de relations R_i. Pour des questions de simplicité, nous considérons ici une base constituée d'une seule relation de M tuples $\{t_u\}_{u=1,...,M}$ non ordonnés, où chaque tuple t_u est une suite de N attributs distincts : $\{t_u.A_1, t_u.A_2, ..., t_u.A_N\}$. t_u est identifié de manière unique par sa clé primaire $t_u.PK$ qui peut être un attribut ou un ensemble d'attributs.
La majorité des schémas de tatouage de bases de données suivent la démarche donnée en figure 1. Pour s'assurer que la marque (i.e. le message tatoué) ne dépend pas de la structure de stockage de la base, les tuples sont artificiellement réorganisés en les répartissant en N_g groupes de tuples $\{G_i\}_{i=1,...,N_g}$. Une stratégie classique pour la création des groupes consiste à déterminer le groupe d'appartenance n_u du tuple t_u de la manière suivante :

$$n_u = H(K_s|H(K_s|t_u.PK))modN_g \qquad (1)$$

où H est une fonction de hachage cryptographique (ex. SHA - *Secure Hash Algorithm*), '|' l'opérateur de

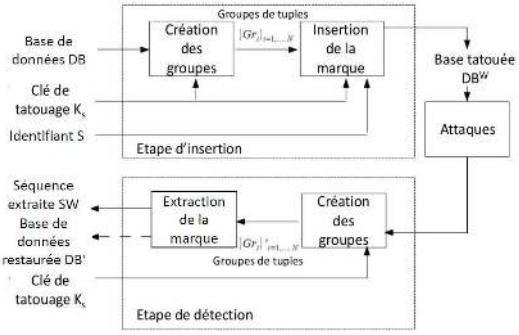

FIGURE 1 – Principe général d'une chaine de tatouage de base de données.

concaténation et K_s la clé secrète de tatouage. Le choix d'une fonction de hachage de type cryptographique garantit une répartition uniforme des tuples dans les groupes. Une fois la base ainsi restructurée, un bit ou un symbole du message est généralement inséré par groupe à l'aide d'une modulation de tatouage. Dans le contexte de notre application, i.e. l'identification de l'origine d'une base de données, le message que nous dissimulons est une séquence pseudo-aléatoire $S = \{s_i\}_{i=1...M}$, $s_i \in \{+1, -1\}$, générée secrètement à l'aide d'une clé K_W. Une fois les tuples tatoués, ils sont substitués à leurs versions originales dans la base. Le processus de détection suit une démarche similaire. Les groupes sont reconstitués et, de chacun, une valeur s_i^W est extraite. La prise de décision sur l'origine de la base de données s'appuie sur une mesure de corrélation entre S et la séquence extraite S^W. S sera détectée dans une base si $\langle S, S^W \rangle \geq T_r$, où T_r est un seuil de décision fixé *a priori*.

II.2. Modulation

Sur la base du schéma précédent, un élément s_i de la séquence S est inséré dans un groupe de tuples G^i à l'aide de la modulation réversible à l'origine développée pour les images par De Vleeschouwer *et al.* dans [4]. Pour ce faire, les tuples de G^i sont également répartis en deux sous-groupes $G^{A,i}$ et $G^{B,i}$ sur le même principe que précédemment :

$$n_s g = H(K_s | H(K_s | t_u.PK)) mod 2 \qquad (2)$$

Supposons maintenant que A_n soit l'attribut considéré pour l'insertion. A_n est un attribut de type numérique, i.e. il prend des valeurs dans une plage donnée. Les histogrammes de $G^{A,i}$ et $G^{B,i}$ sont calculés et projetés sur un cercle comme illustré en figure 2a. Le sous-groupe $G^{A,i}$ (resp. $G^{B,i}$) est ensuite caractérisé par son centre de masses $C^{A,i}$ (resp. $C^{B,i}$) auquel est associé un vecteur $V^{A,i}$ (resp. $V^{B,i}$) (cf. figure 2b). Le processus d'insertion consiste alors à moduler l'angle $\beta_i = (\widehat{V^{A,i}, V^{B,i}})$, pour insérer la valeur $s_i = \pm 1$ de la manière suivante :

$$\beta_i^W = \beta_i - 2\alpha \, si \, s_i = -1, \beta_i^W = \beta_i + 2\alpha \, si \, s_i = +1 \quad (3)$$

Où β_i^W est l'angle tatoué et α l'amplitude de marquage. C'est donc le signe β_i qui porte la valeur du symbole enfoui. L'amplitude de marquage résulte d'une opération de décalage des histogrammes des groupes $G^{A,i}$ et $G^{B,i}$. Prenons l'exemple d'un attribut qui peut prendre L valeurs différentes, cf. figure 2b. Dans ce cas β_i sera modifié de $\mp\alpha = \frac{2\Delta\pi}{L}$ où Δ est l'amplitude de décalage des histogrammes.
Observant β_i^W, le lecteur peut retrouver aisément sa valeur originalle, i.e. β_i. À partir du signe de β_i^W, le sens de rotation est identifié ; il suffit alors d'appliquer le décalage d'histogramme inverse pour récupérer les valeurs originales des attributs. Cependant, tous les groupes de tuples ne peuvent pas forcément porter un symbole. Ces groupes dits "non-porteurs" sont ceux pour lesquels la valeur α est trop faible par rapport à l'angle β_i, ne permettant pas de modifier son signe. Pour éviter une erreur de décodage à la lecture, ces groupes doivent être modifiés. Dans ce cas, β_i est augmenté de $\mp 2\alpha$ en fonction de son signe (voir [3] pour plus de précision). Du fait de la présence de ces "non-porteurs", la séquence insérée S^W sera légèrement différente de S, avec une mesure de corrélation $\langle S, S^W \rangle \leq 1$.

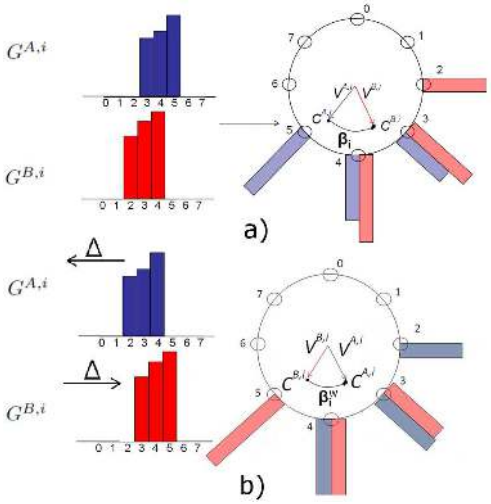

FIGURE 2 – a) Projection des histogrammes sur le cercle. b) Modulation de l'angle β_i et modification correspondante sur l'histogramme.

III. OBJECTIF DE TRAÇABILITÉ ET EXPÉRIMENTATION

Notre objectif est de pouvoir identifier des bases regroupées au sein d'une seule et unique base. Dans ce scénario, une base de données est tatouée avant d'être mêlée avec d'autres. Le message tatoué est un identifiant unique et secret indiquant l'origine de

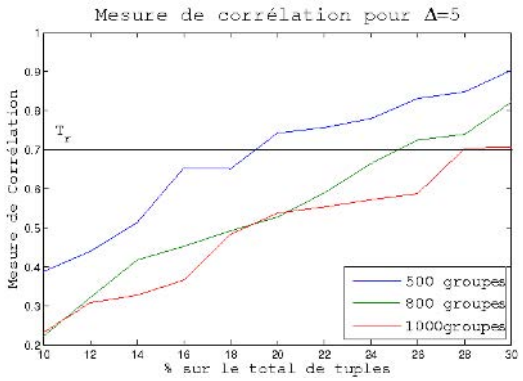

FIGURE 3 – Mesure de corrélation pour $\Delta = 5$.

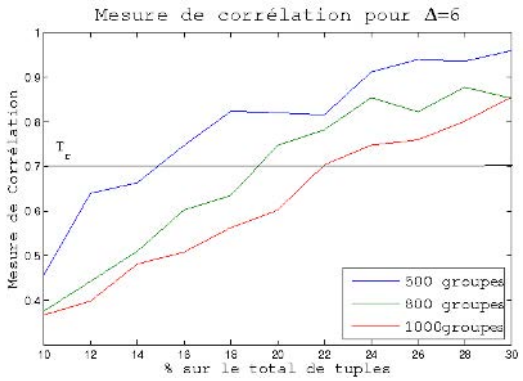

FIGURE 4 – Mesure de corrélation pour $\Delta = 6$.

la base, un hôpital par exemple. Afin de vérifier expérimentalement l'adéquation de notre schéma, nous avons considéré un extrait de la base nationale PMSI-MCO 2011, de plus d'un million de tuples avec des attributs relatifs aux séjours hospitaliers de patients. L'attribut numérique " âge du patient " est utilisé pour dissimuler le message. Les attributs " identifiant de l'hôpital " et " identifiant du séjour " servent quant à eux de clés primaires (cf. section II).
Pour évaluer les performances de détection (ou d'identification) de notre schéma, le plan d'expérience suivant a été exécuté une dizaine de fois :

1. Création de dix bases de données à partir de notre extrait. Nous ne chercherons à vérifier que la présence d'une seule de ces bases dans le mélange, e.g la base n°1.
2. Insertion d'un identifiant unique dans chacune de ces bases considérant une amplitude de décalage d'histogramme (i.e. Δ) et un nombre de groupes de tuples (i.e. N_g) donnés.
3. Mélange de ces bases considérant que la base n°1 représente D% de ce mélange.
4. Détection de la base n°1 dans le mélange.

Au cours de ces tests, nous avons pris différentes valeurs de taille de groupe (N_g= 500, 800 et 1000), d'amplitude de décalage (Δ= 5, 6) et de taux d'occupation de la base n°1 dans le mélange (entre 10% et 30%). Le seuil de décision T_r a quant à lui été fixé arbitrairement à 0.7. Les résultats en termes de mesure de corrélation entre l'identifiant de la base n°1 et celui extrait du mélange sont donnés en moyenne par les figures 3 et 4.
Nous pouvons constater que les capacités de détection de notre schéma dépendent fortement du nombre de groupes et de l'amplitude de décalage utilisés. Pour un même taux de détection, il est possible de jouer sur l'un ou l'autre de ces paramètres. Dans [3], le lecteur pourra vérifier que ce compromis dépend également des propriétés statistiques de l'attribut considéré pour l'insertion. Il apparaît également que si la base d'intérêt représente un pourcentage très faible du mélange (moins de 10%), il faut jouer fortement sur l'amplitude de décalage Δ pour assurer la détection de la marque au risque de nuire à"l'imperceptibilité" de la marque.

IV. CONCLUSION

Dans cet article, nous avons proposé un schéma de tatouage réversible de bases de données relationnelles dans l'objectif de pouvoir les tracer. Les résultats obtenus sur des données réelles montrent que les performances de détection de ce schéma dépendent de la taille de la base à tracer et de sa représentativité après l'agrégation des données. Une base de trop faible représentativité implique un marquage de très forte distorsion pour détecter sa présence.

V. REMERCIEMENTS

Ces travaux ont été réalisés avec le Soutien de l'ANR - Projet ANR ARPEGE PAIRSE.

RÉFÉRENCES

[1] G. Coatrieux, H. Maitre, B. Sankur, Y. Rolland, and R. Collorec, "Relevance of watermarking in medical imaging," in *Proc.Int. Conf. IEEE EMBS ITAB*, 2000, pp. 250 –255.

[2] G. Coatrieux, E. Chazard, R. Beuscart, and C. Roux, "Lossless watermarking of categorical attributes for verifying medical data base integrity," in *Proc. 33th Int. Conf. IEEE-EMBS.* IEEE, Sep. 2011, pp. 8195–8198.

[3] J. Franco-Contreras, G. Coatrieux, E. Chazard, N. Cuppens-Boulahia, C. Roux, and F. Cuppens, "Robust Lossless Watermarking based on Circular Interpretation of Bijective Transformations for the Protection of Medical Databases," in *Proc. 34th Int. Conf. IEEE-EMBS*, 2012, pp. 5875 – 5879.

[4] C. De Vleeschouwer, J.-F. Delaigle, and B. Macq, "Circular interpretation of bijective transformations in lossless watermarking for media asset management," *Multimedia, IEEE Trans. on*, vol. 5, no. 1, pp. 97 – 105, march 2003.