(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

(43) Date de la publication internationale
24 novembre 2005 (24.11.2005)

PCT

(10) Numéro de publication internationale
WO 2005/111915 A2

(51) Classification internationale des brevets[7] : G06K 9/00

(21) Numéro de la demande internationale :
PCT/EP2005/052151

(22) Date de dépôt international : 11 mai 2005 (11.05.2005)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0405236 14 mai 2004 (14.05.2004) FR

(71) Déposant *(pour tous les États désignés sauf US)* : **GEMPLUS** [FR/FR]; Avenue du Pic de Bertagne Parc, d'activité de Gémenos, F-13420 Gémenos (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants *(pour US seulement)* : **NACCACHE, David** [FR/FR]; 52, rue Letort, F-75018 Paris (FR). **BRIER, Eric** [FR/FR]; 486, avenue du 21 Août 1944, F-13400 Aubagne (FR). **CORON, Jean-Sébastien** [FR/FR]; 17, avenue d'Argenteuil, F-92600 Asnières sur Seine (FR). **CARDONNEL, Cédric** [FR/FR]; Résidence "Les Hauts de, Cassis" Bât.D, F-13470 Carnoux en Provence (FR).

(81) États désignés *(sauf indication contraire, pour tout titre de protection nationale disponible)* : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) États désignés *(sauf indication contraire, pour tout titre de protection régionale disponible)* : ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— *sans rapport de recherche internationale, sera republiée dès réception de ce rapport*

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(54) **Title:** METHOD OF MASKING A DIGITAL DATUM, SUCH AS A BIOMETRIC PRINT, AND USE THEREOF FOR SECURING A SECURITY DOCUMENT

(54) **Titre :** PROCEDE DE MASQUAGE D'UNE DONNEE NUMERIQUE, TELLE QUE PAR EXEMPLE UNE EMPREINTE BIOMETRIQUE, ET APPLICATION A LA SECURISATION D'UN DOCUMENT DE SECURITE

$$m = \prod_{i=1}^{n} q_i^{b_i} \bmod p \qquad (I)$$

(57) **Abstract:** The invention relates to a method of masking a plain datum b having n bits. The inventive method is characterised in that a masked datum m is produced using the following masking function: (I), wherein p is a prime number, b_i is the bit at position i of plain datum b, and q_i is the prime number at position i in a set of prime numbers (q_1, ..., q_n). The invention also relates to a method of masking a biometric print, consisting in: determining a set of s real minutiae which are characteristic of the print; mixing and arranging the real minutiae with t false minutiae; and forming a mixed biometric datum b having n = s + t bits, such that, for any i: $b_i = 1$ if position i corresponds to a real minutia, and $b_i = 0$ if position i corresponds to a false minutia. The invention can be used to secure a security document such as a bank cheque.

(57) **Abrégé :** L'invention concerne un procédé de masquage d'une donnée claire b de n bits, caractérisé en ce qu'on produit une donnée masquée m en utilisant la fonction de masquage n suivante : (I), mod p , où p est un nombre premier, b_i est le bit de rang i de la donnée claire b, qi est le nombre premier de rang i d'un ensemble de nombres premiers (q_1, ..., q_n). L'invention concerne également un procédé de masquage d'une empreinte biométrique au cours duquel on détermine un ensemble de s minuties réelles caractéristiques de la dite empreinte, on mélange et on range les minuties réelles avec t fausses minuties, on forme une donnée biométrique mélangée b de n = s + t bits telle que, pour tout i : $b_i = 1$ si le rang i correspond à une minutie réelle et $b_i = 0$ si le rang i correspond à une fausse minutie. Application à la sécurisation d'un document de sécurité tel qu'un chèque bancaire.

WO 2005/111915 A2

PROCEDE DE MASQUAGE D'UNE DONNEE NUMERIQUE, TELLE QUE PAR EXEMPLE UNE EMPREINTE BIOMETRIQUE, ET APPLICATION A LA SECURISATION D'UN DOCUMENT DE SECURITE

L'invention a pour objet les systèmes d'identification et / ou d'authentification biométrique. Ces systèmes manipulent des données biométriques de tous types telles que par exemple des empreintes digitales, des empreintes numériques de l'œil, de la peau, du visage, ou même de la voix.

L'utilisation d'empreintes biométriques est de plus en plus envisagée pour compléter des mots de passe utilisateur ou une signature manuelle, notamment pour des applications nécessitant un haut niveau de sécurité. En effet, l'utilisation d'une empreinte biométrique est un bon complément à un mot de passe ou une signature manuelle, dans la mesure où une empreinte biométrique peut difficilement être dérobée à son propriétaire réel et ne peut pas non plus être imitée, copiée. En contrepartie de cette sécurité et dans la mesure où une empreinte biométrique ne peut pas être remplacée, il est indispensable d'empêcher l'accès direct à cette empreinte afin de garantir la sécurité des personnes et la fiabilité de l'empreinte.

Pour cela, on peut par exemple envisager des procédés de masquage connus pour masquer l'empreinte biométrique à sécuriser. L'empreinte masquée peut être ensuite utilisée en lieu et place de l'empreinte claire, pour signer un message, authentifier l'identité d'une personne, etc. L'intérêt de tels procédés est l'utilisation de fonctions de hachage, qui sont des fonctions à sens unique, c'est-à-dire qu'elles ne peuvent pas être inversées. En

d'autres termes, connaissant une empreinte masquée par une fonction de hachage à partir d'une empreinte claire, il n'est pas possible de retrouver l'empreinte claire, même en connaissant tous les paramètres de la fonction de hachage.

Il est bien connu par ailleurs qu'il n'est pas possible de prendre deux empreintes biométriques strictement identiques d'un même individu à des instants différents. D'abord parce qu'il est très difficile de positionner, de manière strictement identique mais à des instants différents, un même instrument de mesure adapté pour relever la dite empreinte biométrique. Ensuite par ce que l'environnement (température, humidité, etc.) et l'état de santé général (stress, maladie de peau, etc.) de l'individu au moment où l'empreinte est relevée peut perturber le résultat du relevé.

Or, avec les fonctions de hachage connues, partant de deux données initiales peu différentes, les données masquées correspondantes sont très différentes et décorrélées, de sorte qu'il n'est pas possible, en les comparant, d'en déduire si les données initiales sont identiques à quelques erreurs près ou pas. Il n'est donc pas possible d'utiliser les fonctions de hachage connues pour masquer une empreinte biométrique.

Un premier objet de l'invention est de proposer un procédé de masquage utilisant une nouvelle fonction de hachage, mieux adaptée pour masquer des empreintes biométriques que les fonctions de hachage connues. Dans un mode préféré de mise en œuvre, le procédé de masquage selon l'invention est utilisé pour sécuriser une empreinte biométrique.

Enfin, un deuxième objet de l'invention est une utilisation du procédé de masquage de l'invention pour sécuriser un document de sécurité tel que par exemple un chèque bancaire.

Le premier objet de l'invention est atteint par un procédé de masquage d'une donnée claire b de n bits, caractérisé en ce qu'on produit une donnée masquée m en utilisant la fonction de hachage suivante : $m = \prod_{i=1}^{n} q_i^{b_i} \bmod p$, où p est un nombre premier, b_i est le bit de rang i de la donnée claire b, q_i est le nombre premier de rang i d'un ensemble de nombres premiers (q_1, ..., q_n). De préférence, p est un nombre premier de grande taille et les éléments de l'ensemble de nombres premiers sont de petite taille.

Par rapport aux fonctions de hachage connues, la fonction $m = \prod_{i=1}^{n} q_i^{b_i} \bmod p$ a pour principal intérêt d'être tolérante aux erreurs, comme on le verra mieux par la suite, elle est de ce fait particulièrement bien adaptée pour masquer des données biométriques.

Dans un mode préféré de mise en œuvre, le procédé de masquage ci-dessus est appliqué à une empreinte biométrique. Pour cela, au cours du procédé, on détermine un ensemble de s minuties réelles caractéristiques de la dite empreinte, on mélange et on range les minuties réelles avec t fausses minuties, on forme une donnée mélangée b de n = s + t bits telle que, pour tout i :

$b_i = 1$ si le rang i correspond à une minutie réelle et

$b_i = 0$ si le rang i correspond à une fausse minutie.

et on applique la fonction de hachage selon l'invention à cette donnée mélangée pour produire une donnée masquée.

De préférence, les minuties réelles et les fausses minuties sont mélangées de façon aléatoire.

Le deuxième objet de l'invention concerne quant à lui un procédé de sécurisation d'un document de sécurité, par exemple d'un chèque bancaire, au cours duquel, après avoir obtenu une donnée de référence par masquage d'une empreinte biométrique selon un procédé tel que décrit ci-dessus,

- on mémorise la dite donnée de référence sur ou dans le document de sécurité, ou
- on associe à la dite donnée de référence un code - barre que l'on mémorise sur ou dans le document de sécurité, la donnée de référence et le code - barre étant également mémorisés dans une table.

Des exemples préférés de mise en œuvre de l'invention sont décrits ci-dessous.

On va tout d'abord détailler le procédé de masquage selon l'invention. Pour masquer une donnée claire $b = (b_n, \ldots, b_1)$ de n bits, on utilise la fonction de hachage suivante:

$$m = \prod_{i=1}^{n} q_i^{b_i} \bmod p$$

La fonction utilise comme paramètres un ensemble $(q_n, \ldots, q_1)$ de petits nombres premiers, par exemple des nombres entiers d'environ 60 bits. La fonction utilise également un paramètre p, qui est un entier de grande taille, par exemple d'environ 1024 bits. p est choisi de préférence tel que $2*q_n\text{^}2t < p < 4*q_n\text{^}2t$, où t est un nombre d'erreurs acceptées.

Contrairement aux fonctions de hachage connues, la fonction selon l'invention est peu sensible aux erreurs, c'est-à-dire que, connaissant deux données m, μ masquées

par cette fonction, il est possible de dire si les données claires d'origine b, β correspondantes sont identiques, à au maximum t erreurs près.

En effet, m, μ sont obtenues par les relations :

$$m = \prod_{i=1}^{n} q_i^{b_i} \bmod p \text{ et } \mu = \prod_{i=1}^{n} q_i^{\beta_i} \bmod p,$$

On définit de plus :

$$\lambda = \frac{m}{\mu} \bmod (p) = \frac{a}{\alpha} \bmod p$$

avec $a = \prod_{i \in \Delta_i} q_i \bmod p$ et $\alpha = \prod_{i \in \Gamma_i} q_i \bmod p$

où Δ_i est l'ensemble des indices i compris entre 1 et n pour lesquels $b_i = 1$ et $\beta_i = 0$, et où Γ_i est l'ensemble des indices i compris entre 1 et n pour lesquels $b_i = 0$ et $\beta_i = 1$. La somme des tailles des ensembles Δ_i et Γ_i est au plus égale à t, t étant le nombre de bits de β différents des bits de b de même rang, correspondant au nombre maximum d'erreurs acceptées.

a et α, qui sont des produits de petits nombres premiers q_i, sont également des petits nombres, qui vérifient de plus la relation : $a*\lambda = \alpha \bmod p$. A partir de cette dernière égalité et du nombre λ, il est alors possible de retrouver les nombres a et α. Une décomposition de a et α en nombres premiers permet finalement factoriser a et α. La décomposition est facilitée en tirant partie du fait que a et α se décomposent en principe en de petits nombres premiers. Si a et α se décomposent sur l'ensemble $(q_n, \ldots, q_1)$, alors on en déduit que les données d'origine b et β sont identiques, à au plus t erreurs près.

On va maintenant décrire un mode préféré de mise en œuvre du procédé de masquage utilisant la fonction de masquage décrite ci-dessus, pour le masquage d'une empreinte biométrique.

Dans l'exemple ci-dessous, l'empreinte biométrique physique que l'on cherche à masquer est une empreinte digitale caractérisée par un nombre s prédéfini de minutiae réelles. Une minutie réelle est un détail d'une empreinte en un point donné de l'empreinte physique, comme par exemple une rupture d'une ligne, une fourche sur une ligne, etc. Numériquement, une minutie peut être traduite par une chaîne de caractères incluant des informations sur la position et la forme de la minutie.

Selon l'invention, pour masquer l'empreinte physique, on ajoute d'abord à l'ensemble des minuties réelles un ensemble de t fausses minuties, également définies par une chaîne de caractères mais qui ne correspondent pas à une minutie réelle de l'empreinte physique. Dans un exemple, une fausse minutie est définie de manière totalement aléatoire, et on ajoute un ensemble de t = 80 fausses minuties à un ensemble de s = 20 minuties réelles.

L'ordre des minuties réelles et des fausses minutiae est mélangé, par exemple de façon aléatoire, puis on forme une donnée mélangée $b = (b_n, \ldots, b_1)$ de $n = s + t$ bits telle que, pour tout i :

$b_i = 1$ si le rang i correspond à une minutie réelle et

$b_i = 0$ si le rang i correspond à une fausse minutie.

La donnée mélangée b est ensuite masquée par le procédé de masquage selon l'invention pour produire une donnée masquée m telle que :

$$m = \prod_{i=1}^{n} q_i^{b_i} \bmod p$$

La donnée masquée m peut ensuite être mémorisée dans une base de données, sur une carde d'identité, dans une mémoire d'une carte à puce, etc. La donnée masquée m peut être utilisée comme donnée de référence, par exemple pour vérifier l'identité d'une personne, de la manière suivante.

Pour vérifier l'identité d'une personne, il suffit :

- de relever une nouvelle empreinte biométrique physique sur la personne puis de calculer l'ensemble de s minuties réelles correspondant,
- d'ajouter t fausses minuties, de mélanger les fausses minuties et les vraies minuties, de déterminer la donnée mélangée β associée à la nouvelle empreinte biométrique, puis de masquer β par la fonction $\mu = \prod_{i=1}^{n} q_i^{\beta_i} \bmod p$ pour obtenir une nouvelle donnée masquée μ,
- de déterminer s'il y a concordance entre la donnée référence m précédemment mémorisée et la donnée μ masquée obtenue à partir de la nouvelle empreinte réelle qui vient d'être relevée.

Pour déterminer s'il y a concordance entre m et μ :

- on calcule $\lambda = \frac{m}{\mu} \bmod p = \frac{a}{\alpha} \bmod p$, puis a et α à partir de la relation $a*\lambda = \alpha \bmod p$, avec a et α petits devant l'entier p, par l'algorithme des fractions continues par exemple.

- on décompose ensuite a et α en facteurs premiers, puis
  - il y a concordance si a et α se décomposent sur au plus t éléments de l'ensemble des nombres premiers $(q_n, \ldots, q_1)$,
  - il n'y a pas concordance sinon

Une application envisagée du procédé de masquage selon l'invention vise à sécuriser un document de sécurité tel que par exemple un chèque bancaire. Pour cela, selon l'invention, une empreinte biométrique du propriétaire du document de sécurité est masquée par un procédé de masquage tel que décrit ci-dessus, pour produire une donnée de référence.

Selon un premier mode de réalisation, la donnée de référence est mémorisée sur ou dans le document de sécurité, par exemple par impression.

Selon un deuxième mode de réalisation, la donnée de référence est associée à un code - barre, le couple donnée de référence / code - barre associé est mémorisé dans une base de donnée, et le code - barre est mémorisé, par exemple par impression, sur le document de sécurité.

Il suffit ensuite, lorsque le document de sécurité est transmis par exemple, de prendre simultanément avec le document de sécurité une empreinte biométrique de la personne qui transmet le dit document puis de vérifier que l'empreinte biométrique de la personne qui transmet le document correspond bien à l'empreinte incluse dans la donnée de référence mémorisée sur le document ou associée au code - barre mémorisé sur le document.

Dans le premier mode de réalisation, la vérification pourra être faite par toute personne , la donnée de

référence étant mémorisée directement sur le document. Dans le deuxième mode de réalisation, la vérification pourra être faite par toute personne ayant accès à la base de données, et qui n'est pas nécessairement la personne qui reçoit le document.

Le code - barre est réalisé selon des techniques connues, on pourra utiliser par exemple un code - barre à une dimension, constitué d'une série de barres verticales d'épaisseur et d'écartement variables. Le choix de la forme du code - barre est en pratique fonction du nombre de données de référence à mémoriser, chaque donnée de référence correspondant à des personnes différentes.

De préférence, la base de données dans laquelle les couples donnée de référence / code - barre associé sont mémorisés est accessible pour vérification uniquement à un nombre restreint de personnes, selon le niveau de sécurité souhaité : l'accès peut par exemple être autorisé pour toute personne amenée à recevoir des documents de sécurité ou, de manière plus restreinte, uniquement à une autorité certificatrice.

Dans un exemple pratique, le document de sécurité est un chèque bancaire et l'empreinte digitale de son propriétaire est mémorisée sur le chèque sous la forme d'un code - barre. Un commerçant dispose d'un dispositif de lecture et de masquage d'une empreinte doté de moyens pour lire une empreinte, la masquer puis imprimer la donnée masquée associée. La banque émettrice du chèque a seule le droit d'accès à la base de données dans laquelle sont mémorisés la donnée masquée de référence (correspondant à l'empreinte initiale masquée) et le code - barre associé ; cet accès lui permet de vérifier que l'empreinte laissée par la personne qui a présenté le chèque au commerçant, et que ce dernier a masqué et

imprimé sur le chèque, correspond bien à celle du propriétaire du chèque.

REVENDICATIONS

1. Procédé de masquage d'une donnée claire b de n bits, caractérisé en ce qu'on produit une donnée masquée m en utilisant la fonction de hachage suivante : $m = \prod_{i=1}^{n} q_i^{b_i} \bmod p$, où p est un nombre premier, b_i est le bit de rang i de la donnée claire b, q_i est le nombre premier de rang i d'un ensemble de nombres premiers (q_1, ..., q_n).

2. Procédé de masquage selon la revendication 1, dans lequel p est un nombre premier de grande taille et les éléments de l'ensemble de nombres premiers sont de petite taille.

3. Procédé de masquage selon l'une quelconque des revendications 1 et 2, appliqué à une empreinte biométrique, caractérisé en ce qu'on détermine un ensemble de s minuties réelles caractéristiques de la dite empreinte, en ce qu'on mélange et on range les minuties réelles avec t fausses minuties, en ce qu'on forme une donnée mélangée b de n = s + t bits telle que, pour tout i :

b_i = 1 si le rang i correspond à une minutie réelle et

b_i = 0 si le rang i correspond à une fausse minutie,

et en ce qu'on applique la fonction de hachage à cette donnée mélangée pour produire une donnée masquée.

4. Procédé selon la revendication 3, dans lequel les minuties réelles et les fausses minuties sont mélangées de façon aléatoire.

5. Procédé de sécurisation d'un document de sécurité, par exemple d'un chèque bancaire, au cours duquel, après avoir obtenu une donnée de référence par masquage d'une empreinte biométrique selon un procédé selon la revendication 3 et 4,

- on mémorise la dite donnée de référence sur ou dans le document de sécurité, ou
- on associe à la dite donnée de référence un code - barre que l'on mémorise sur ou dans le document de sécurité, la donnée de référence et le code -barre étant par ailleurs mémorisés dans une table.

6. Procédé de vérification d'un document de sécurité sécurisé par un procédé selon la revendication 5, procédé de vérification au cours duquel :

- on numérise une empreinte biométrique physique d'une personne présentant le document de sécurité,
- on masque l'empreinte numérisée par un procédé selon l'une des revendications 3 à 4, pour produire une donnée masquée,
- on compare la donnée masquée avec une donnée de référence, puis
- on accepte le document de sécurité si la donnée masquée et la donnée de référence sont identiques à un taux d'erreur près prédéfini, ou on refuse le titre sinon.

7. Procédé selon la revendication 6, dans lequel, lors de l'étape de comparaison, si un code - barre associé à la donnée de référence est mémorisé sur le document de sécurité, alors :

- on lit le code - barre et on recherche dans une table la donnée de référence associée au code - barre, puis
- on compare la donnée de référence avec la donnée masquée.