La carte codée optiquement: Aspects système et sécurité

CDU: 351.755.6:535.2:535.411

H. Lienhard

Dans cette publication, nous présenterons les premiers produits se basant sur l'optique cohérente et sur la technologie de carte de la Landis & Gyr: Le système PHONOCARD® et le système de contrôle d'accès, ID 2000. Nous exposerons le concept du système ainsi que quelques réflexions du point de vue de la théorie d'information sur la sécurité de tels appareils et installations.

1. Introduction

Les produits mentionnées sont des exemples de deux domaines distincs d'application utilisant des cartes de type différent: la *carte prépayée* et la *carte d'identification* ou en abrégé, *carte ID*.

Chaque carte contient un certain nombre d'unités codées spécialement, qui dans un cas représentent des unités de *valeur* pouvant être consommées, dans l'autre cas, des unités d'*information*.

En fait, il s'agit de deux vieilles connaissances: le ticket (dévalorisé par le poinçonnement de trous) et la carte d'identité (carte de crédit....). Puisque ces documents ne peuvent être fabriqués resp. émis que par des intances spécialement autorisées, leur authenticité doit être contrôlable et difficilement imitable.

Contrairement aux variantes habituelles, nous nous intéressons ici seulement aux documents qui doivent être acceptés et examinés exclusivement par des machines.

Aussi complexes qu'elles puissent être, l'aptitude de ces machines (appelées par la suite accepteurs) a reconnaître des structures (patterns) complexes est négligeable vis-a-vis de celle de hommes.

Pour réaliser quand même des systèmes de haute sécurité, il faut poser des exigences spéciales aux caractéristiques d'authenticité des documents, *en particulier le document et l'accepteur doivent être exactement adaptés l'un a l'autre.*

De l'accepteur, on exigera d'accepter si possible sans faute les bons (vrais) documents et de rejeter les contrefaçons avec une grande probabilité. En plus, pour beaucoup d'applications, il devra être bon marché.

2. Les systèmes W et E

Nous distinguons deux systèmes fondamentalement différents utilisant des documents pouvant être lus par machine.

Le système W (rewriting system) – l'information peut être modifiée par l'accepteur.

Le système E (erasing system) – l'information peut au plus être effacée par l'accepteur.

Donc, pour les cartes prépayées, nous avons dans le cas W des cartes rechargeables avec de nouvelles unités, dans le cas E en revanche, ces unités de valeur sont physiquement détruites.

Un accepteur relativement sûr peut alors être réalisé avec une dépense raisonnable, si le choix et la qualité des paramètres importants pour l'accepteur,

- rendent improbable une imitation des unités *vraies* vu les dépenses et la difficulté technologiques.
- permettent une bonne discrimination des contrefaçons les plus usuelles (une mesure de cette discrimination sera introduite en appendice ainsi que des procédures possibles de décision).

Si l'accepteur doit être normalement accessible, cette exigeance ne pourra guère être satisfaite pour le système W.

Tout d'abord, la difficulté technologique de création de ces unités dans l'accepteur est hors de question, ensuite, il suffit de voler un tel appareil pour réaliser de telles unités.

Il s'ensuite qu'*intrinsèquement* les systèmes W ne sont *pas sûrs* (exemple classique: la carte magnétique).

Pour cette raison nos systèmes de cartes, *PHONOCARD* et ID 2000, sont tous deux conçus comme systèmes E.

3. Niveaux du système

Bien que les systèmes de prépaiement et de contrôle d'acces (ID) servent des buts différents, ils peuvent être traités ensemble du point de vue purement technique.

Les cartes prépayées contiennent en plus des unités de valeur, certaines

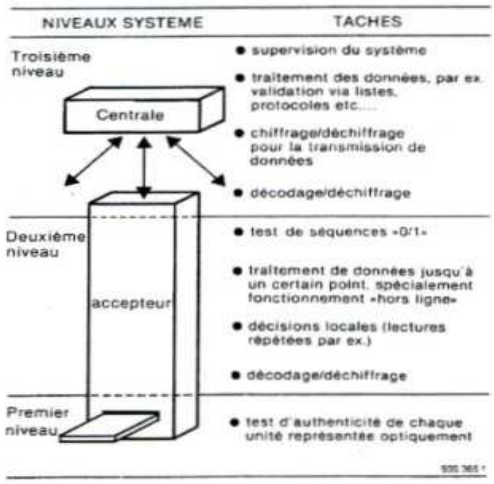
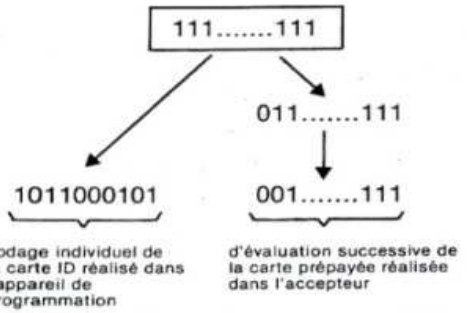


Fig. 1 Niveaux logiques et leurs fonctions dans un système de cartes.

informations d'identification (par ex: la date d'émission); les cartes ID seront individualisées par effacement, cependant cette opération au contraire de l'effacement des cartes prépayées ne se fait pas dans l'accepteur mais à l'aide d'un appareil de programmation placé dans un endroit sûr.

Par la suite nous représenterons une unité de valeur encore valable comme le «un»binaire par «1» et une unité de valeur effacée comme le «zéro» binaire par «0».

Les cartes prépayées et les cartes en blanc ID contiennent toutes deux une suite de «1», qui après sera modifiée par effacement.



Dans un système de cartes, on distingue trois niveaux logiques plus ou moins liés (voir *fig. 1).*

- Au *premier niveau*, l'unité d'information (ou de valeur) sera chiffrée. L'unité est *représentée physiquement* à ce niveau.
- Déjà au *deuxième niveau* – le niveau code – on opéra avec les symboles abstraits («0» et «1»).
  Les séquences de tels symboles seront combinés en mots ("bit-strings"). A ce niveau l'accepteur contrôlera le codage, éventuellement aussi le chiffrage (de la carte.
- Le *troisième niveau* comprend le *traîtement* proprement dit *des données* comme le contrôle de listes noires etc.
  Dans le cas de lecteurs autonomes ces fonctions sont réalisées à l'aide d'un microordinateur se trouvant dans l'accepteur.
  Dans le cas de système de contrôle d'accès, l'ordinateur de la centrale pourra executer ces fonctions.

4. Effacement illegal

Au premier niveau, on examine l'authenticité des «1» et «0». La création de «1» authentiques (structures codées optiquement) est très difficile et par conséquent la probabilité de création illégale de «1» est très faible; en revanche l'effacement c'est-à-dire la création de vrais «0» est bien plus simple.

Dans le cas de la carte prépayée, un effacement illégal dévalue seulement la carte, dans le cas de la carte ID en revanche, on pourrait essayer de modifier l'information (par ex. un droit d'accès) par effacement. Il existe une assymétrie dans la sécurité des «0» et «1».

Cette assymétrie peut être éliminée aux deuxième niveau par un codage adéquat de la séquence binaire. Par exemple, on utilisera un codage qui aura pour séquence binaire légale un nombre fixe de «0» et «1».

Lors de la lecture d'une carte, on contrôlera au deuxième niveau le nombre de «0» et de «1».

Exemple: la séquence binaire comprend 96 bits, elle ne sera acceptée que si elle contient 48 «0» et 48 «1». De telles séquences,
Il en existe $\binom{96}{48} (\cong 6{,}4 \cdot 10^{27})$ Sans cette restriction, il existe $2^{96} (\cong 8 \cdot 10^{28})$ séquences. Avec un code «48 parmi 96», on perd donc environ une decade par rapport au codage purement binaire.

5. Chiffrage

Comme indiqué ci-dessus *(fig. 1)*, nous distinguons deux types de chiffrage. Le chiffrage de la carte et le chiffrage de la transmission des données depuis l'accepteur jusqu'à la centrale. Dans le premier cas, il s'agit surtout de la protection de la banque des données, dans le second cas, de la protection de la transmission des données.

Si la ligne de transmission est rendue inaccessible, le deuxième chiffrage perd son intérêt.

Par le chiffrage de la carte, on veut empêcher que des personnes ayant affaire avec la fabrication ou la mise en service d'un tel système puissent abuser de leur connaissances sur la banque des données (par ex. liste des personnes autorisées etc.) et sur le traitement des données. On veut donc empêcher que par ex. un programmeur connaissant le décodage et le déchiffrage ainsi que les codes légaux comme ils sont représentés dans la banque des données, puisse retrouver les codes des cartes elles-mêmes. Ceci peut être évité en utilisant par ex. ce qui est appelé «trap door one way functions» [1].

La *figure 2* montre comment une information valable (CI) de carte peut être stockée dans une banque de donnée. La carte ne contient pas cette information mais une forme chiffrée de celle-ci.

Cette dernière est générée par un appareil de programmation (se trouvant en lieu sûr) à l'aide d'une fonction maintenue secrète $f(\cdot)$:

Code de carte CC' = f(CI)

La fonction inverse $f^{-1}(\cdot)$ est appliquée a CC' dans l'accepteur ou eventuellement seulement dans l'unité centrale.

Dans un système de contrôle d'accès par ex. on contrôlera si la fonction $f^{-1}(CC')$ est présente dans la liste autorisant l'accès.

La fonction f sera choisie de telle façon que sans la connaissance d'une information «trap-door» il soit pratiquement impossible de retrouver f à partir de f^{-1}.

De cette manière, la connaissance de la «liste ayant accès» ne sert pratiquement plus.

Si la liaison accepteur – centrale d'un système de contrôle d'accès peut-être écoutée en dehors du domaine de sécurité, on pourra avoir accès illegalement sans pour cela utiliser l'accepteur. Il suffit d'enregistrer une séquence de signaux (qui mène à l'ouverture de la porte) et d'injecter sur la liaison une telle séquence à un moment propice. Un dialogue compliqué entre l'accepteur et la centrale rend une telle opération plus difficile, mais en fait la situation n'est pas changée.
Nous imposerons à ce chiffrage la condition suivante: à savoir-rendre pratiquement impossible, à partir de l'observation de séquences. S_i de n signaux, de retrouver avec succès une séquence propre de signaux E_{n+1}

$S_1, S_2, \ldots\ldots S_n \longmapsto E_{n+1}$

CODE DE LA FIRME	123	930 366 f
IDENTIFICATION	321654987	
ZONE DE LIEU	05	
ZONE DE TEMPS	0	
NUMERO D'EMISSION	0	
GENERATION DU P.I.N. (Numero d'Identification Personnel)	NOMBRE DE CARACTERES SECRETS	0
	ALGORITHME	6

Fig. 2 Information en texte clair de la carte (CI) pour une carte de contrôle d'accès.

En principe, un chiffrage/déchiffrage dépendant du temps et de la date peut réaliser cette condition [1].

Le chiffrage de carte mentionné d'abord, qui s'occupait principalement de la protection de l'opérateur du système vis-à-vis du producteur des cartes et de l'accepteur, peut être réalisée de différentes façons.

Les «trap-door one-way functions» déjà mentionnées offrent une possibilité intéressante. Ici l'algorithme de déchiffrage peut être rendu complètement accessible sans pour autant mettre le système en danger.

Si on utilise un chiffrage de bloc classique [2], la clef K de l'opérateur du système doit rester secrète.

Pour transmettre la clef avec sécurité, on pourrait, par ex. utiliser une «carte clé» qui sera lue dans l'accepteur.

La *figure 3* illustre une séquence possible de codage/décodage pour un système ID.

6. Structure de l'accepteur

Dans sa structure, l'accepteur constitue un système de contrôle et de décision. La *figure 4* illustre le cas général. La carte insérée est illuminée par les sources, la lumière réfléchie résultante est reçue par les détecteurs et convertie en signaux électriques analogiques qui seront ensuite traîtés par la partie électronique.

Ainsi les signaux des détecteurs S_i (t) sont transformés en un point d'un espace de décision multidimensionnel, c'est-à-dire un vecteur $\underline{Y}$ de valeurs digitales.

L'algorithme de décision proprement dit est programmé dans le microordinateur de l'accepteur (voir en appendice).

Le microordinateur contient également des algorithmes de commande et de contrôle des sources lumineuses, de l'effacement et du transport des cartes. De cette manière, on contrôlera non seulement l'authenticité de la structure optique mais aussi par ex. le comportement à l'effacement du matériel de la carte.

Le choix de ce concept pour l'accepteur apporte en plus d'une sécurité optimale, une grande flexibilité.

7. Le problème de décision

Les procédures de décision doivent assurer d'une part une grande sécurité contre toute fraude, d'autre part être réalisés de manière simple et efficiente par le logiciel du microordinateur.

De telles procédures de décision peuvent être déduites à l'aide de méthodes statistiques semblables à celles utilisées dans la théorie des estimations et des tests d'hypothèses.

Comme mentionné ci-dessus, les signaux des détecteurs sont représentés après un traîtement électronique

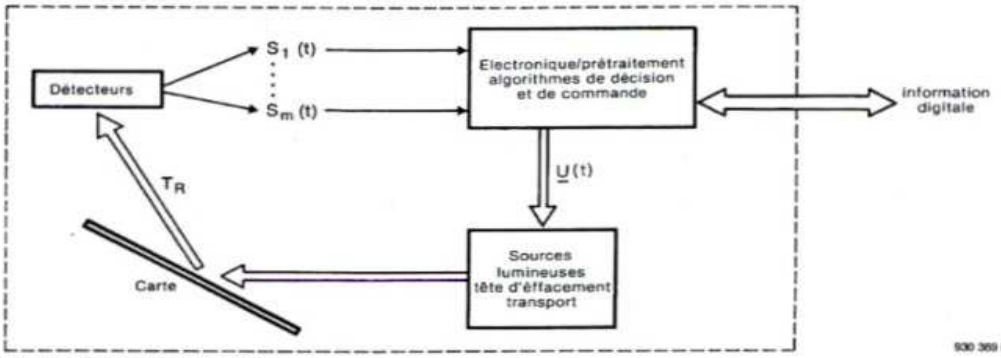


Fig. 4 Structure de l'accepteur

préalable (par une opération T_V) en un vecteur $\underline{Y}$ d'un espace de décision $\mathfrak{G}$ à N dimensions.

Si N > 2, il est pratiquement impossible de trouver intuitivement une bonne procédure de décision.

En appendice, nous montrerons comment arriver à un test optimal en rendant minimum le risque (appelé risque de Bayes) dû à une décision erronnée (équation 7).

Ce procédé n'est pas très commode à réaliser dans l'accepteur; c'est pourquoi nous déduirons de ce test optimal, un test *suboptimal* plus sévère mais conduisant à une procédure de décision plus simple. Supposons une répartition statistique de Gauss. Par ce procédé l'espace de décision $\mathfrak{G}$ sera divisé par trois hyperplans en trois domaines Γ_0, Γ_1 et le reste.

Si le vecteur $\underline{Y}$ se trouve dans Γ_0, on se décidera pour un «0», s'il se trouve dans Γ_1, on se décidera pour un «1». S'il ne se trouve ni dans Γ_0, ni dans Γ_1, l'unité ne sera pas acceptée.

Numériquement parlant, cela signifie le test d'un nombre d'inégalités du type

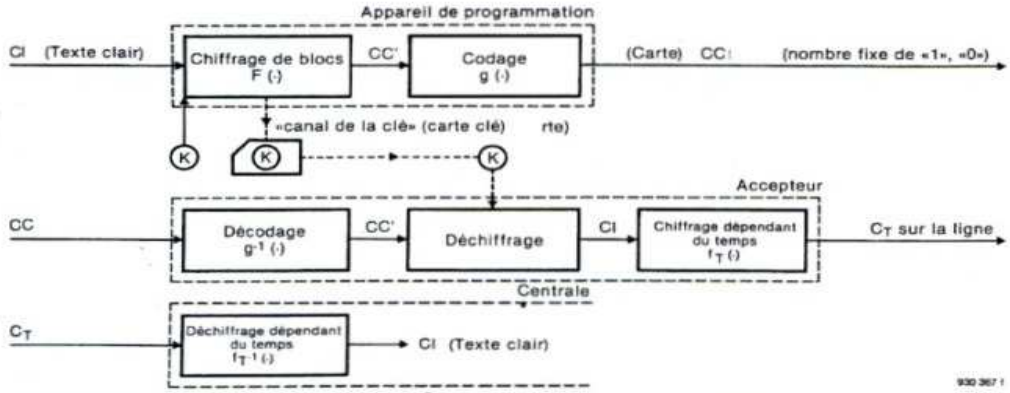


Fig. 3a Chiffrage de la carte à l'aide de chiffrage de blocs. La clé est connue par l'opérateur de système seul.

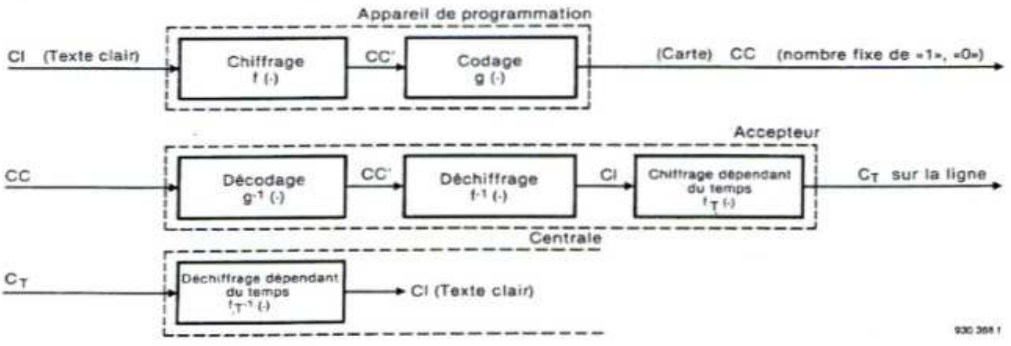


Fig. 3b Chiffrage de la carte réalisé avec «trap-door one-way function f», f est connue de l'opérateur du système seulement, f est donnée au fabricant de l'accepteur par l'opérateur du système.

$\underline{\alpha}_{ik}^{T} \underline{Y} < D_{ki}$

ou $\underline{\alpha}_{ik}^{T}$ représente un vecteur ligne précalculé et D_{ki} une constante précalculée (équation (17) à (19)).

Nous introduirons également en appendice avec le procédé de décision, une mesure de discrimination, l'information de discrimination. Cette mesure permet un jugement quantitatif de différentiation entre diverses alternatives, par ex. la différentiation des «1» par rapport à une certaine contrefaçon. Cette mesure peut être utilisée comme aide pour la spécification des paramètres du document ainsi que pour le «design» de l'accepteur.

8. Liaison au «problème inverse»

(voir page 7)

Avant d'utiliser une procédure de décision, nous représenterons comme mentionné ci-dessus les mesures après transformation T_v en un vecteur $\underline{Y}$ dans l'espace de décision.

Si le «problème inverse», en abrégé IP, est résoluble (cela signifie que par ex. on puisse retrouver, a partir de mesures d'intensité du champ éloigné, le réflecteur optique), alors on pourra utiliser l'*espace de représentation* du réflecteur lui-même comme notre domaine de décision. On se limitera ici au cas paramétrique: Supposons que la structure du réflecteur soit décrit par un vecteur paramétrique $\underline{Y}$ de dimension N. Nous déduisons la transformation suivante à partir de la solution du «problème inverse»

$$\begin{pmatrix} T_v^*\text{:espace} \\ \text{de mesure} \end{pmatrix} \longrightarrow \begin{pmatrix} \text{espace de décision} \\ (= \text{espace de repré-} \\ \text{sentation}) \end{pmatrix}$$

Le problème inverse ne doit pas être necessairement résoluble pour le procédé de décision mentionné dans le dernier paragraphe. Pour cela nous supposerons connues à priori les contrefaçons dangereuses. Dans ce cas, il suffit de se protéger contre ces contrefaçons facilement réalisables; i. e. dans l'espace de décision, nous avons besoin d'une «distance» suffisante entre les structures authentiques et ces contrefaçons. Pour pouvoir quantifier cette «distance» numériquement, nous utilisons une mesure de discrimination appropriée.

Si nous ne connaissons rien à priori de ces contrefaçons, nous pouvons procéder différemment. Nous devons garantir que les valeurs mesurées qui conduisent a une décision «authentique» proviennent avec grande probabilité seulement d'une certaine structure (authentique). Dans ce cas, le IP doit avoir une solution stable et si possible unique. En principe, une ambiguïté ne peut être acceptée que si chaque solution c'est-à-dire chacune des structures possibles soit difficile à réaliser. Si nous transformons les valeurs mesurées dans l'espace de décision, aucune information de discrimination doit être perdue.

A l'aide de la mesure de discrimination, nous pouvons définir dans l'espace de décision, les zones Γ_0 et Γ_1 acceptant les structures permises «0» et «1».

Pour ce faire, ces domaines seront choisis aussi étroits que le degré minimal d'acceptation, pour des unités authentiques, soit encore atteint [3].L'exigence de trouver une solution au IP requiert de nombreuses mesures; ce procédé n'est praticable que si l'effort requis par ces mesures est acceptable.

Dans l'accepteur, le réflecteur passif est illuminé produisant des signaux dans un espace de mesure $\mathfrak{M}$; appelons ceci la transformation à la lecture T_R. De plus appelons $\mathfrak{D}$ l'espace de représentation du réflecteur (ou structure) et Γ la fonction de décision, nous pouvons dès lors formuler schématiquement le procédé d'acceptation comme dans la *figure 5*.

9. Séparation des systèmes

Nous avons mentionné deux methodes de décision. Dans l'une nous supposons la connaissance de contrefaçons dangereuses, dans l'autre non. En réalité, certaines contrefaçons sont généralement bien connues. Il ne s'agit pas seulement dans un système tel que PHONOCARD (voir page 40), de se protéger contre des contrefaçons, mais également de se différentier par rapport à d'autres systèmes. Les cartes d'un pays ne doivent naturellement pas être acceptées par un accepteur d'un autre pays. Pour obtenir une séparation entre les différents systèmes, les unités (des «1» pour les cartes prépayées) seront, pour chaque système, représentées différemment.

Des cartes étrangères au système doivent être rejetées avec grande probabilité comme les contrefaçons.

10. Appendice: Procédés de décision et mesure de discrimination

Nous allons ici essayer d'obtenir des procédures rationnelles de décision à l'aide de méthodes statistiques.

Des considérations intuitives n'aboutissent plus à un résultat pour des espaces de décision de dimensions N plus élevées. Tout d'abord nous déduirons un test optimal dans un certain sens. A partir de celui-ci nous développerons un test suboptimal plus sévère qui se laisse réaliser plus facilement dans l'accepteur. Comme le montre la *figure 6*, les signaux bruts des détecteurs (vecteur $\underline{S}(t)$) sont prétraités et transformés en un vecteur (ou point) de l'espace de décision $\mathfrak{G}$ de dimension N:

La transformation

$$T_v: [\underline{S}(t), 0 \leq t \leq T] \longrightarrow \underline{Y} \qquad (1)$$
$$\underline{Y} \in \mathfrak{G}$$

ou [0, T] represente l'intervalle d'observation

Bien que par la suite nous traiterons exclusivement le problème de décision, notons quelques remarques à propos du prétraitement, c'est-à-dire de la transformation $T_v(\cdot)$:

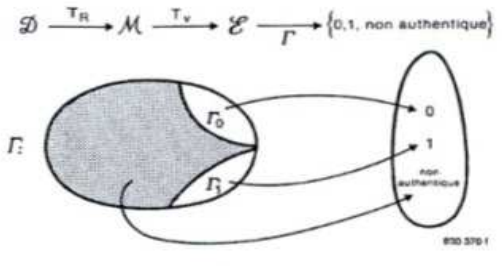


Fig. 5 Procédé de lecture et de test dans l'accepteur

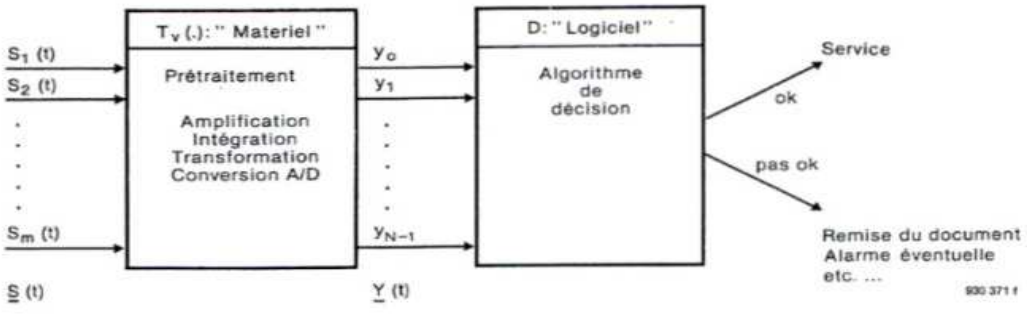


Fig. 6 Structure du traitement du signal dans l'accepteur

- comme on le verra l'*information de discrimination* I(i:j) entre les diverses hypothèses H_i et H_j joue dans l'algorithme de décision un rôle central. Idealement cette information ne devrait pas être amoindrie par la transformation $T_y(\cdot)$; d'après Kullback [3] $\underline{y}$ serait une «sufficient statistic» pour la discrimination.
- Normalement cette procédure de décision D (voir *figure 6)* sera réalisée dans le logiciel du microordinateur; d'autre part le prétraitement demande dans la plupart des cas un matériel analogique encore plus exigeant.

10.1 Le test optimal

Définissons dans l'espace $\mathfrak{G}$ des sous-ensembles Γ_j – dans un certain sens optimaux – de telle sorte que l'hypothese H_j découle avec grande sécurité de $\underline{Y} \in \Gamma_j$. Pour cela nous introduirons le «risque de Bayes» (voir par ex. [4]) que nous essayerons de rendre minimum.

Avec E{} = l'espérance mathématique et Prob ( ),P( ) = la probabilité, nous décrirons de la manière suivante, le risque (c'est-à-dire les coûts auxquels il faut s'attendre).

$$E\{\text{coûts}\} = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} c_{ij}\ \text{Prob (décision pour } H_j, \text{ en réalité } H_i)$$

$$= \sum_{i,j} c_{ij} \cdot P(\underline{Y} \in \Gamma_j \text{ et } H_i) ; \qquad (2)$$

c_{ij} = facteurs de couts

Ce faisant nous considérons M hypothèses différentes.

Pour clarifier les ésprits, nous considérons un problème plus concret. Supposons que les hypothèses H_0 et H_1 représentent des *alternatives légitimes* (c'est-à-dire «0» ou «1» pour une carte d'accès ou de crédit) et $H_2 \ldots H_{M-1}$ des contrefaçons.

Avec une expérience croissante, M pourra plus tard augmenter, ce qui conduira seulement à une modification du logiciel de l'accepteur.

Théoriquement, il y a naturellement d'innombrables fausses alternatives immaginables.

L'expérience montre cependant que seules les contrefaçons reproductibles avec relativement peu d'efforts sont vraiment dangereuses pour le système. Dans cet esprit, les $H_2 \ldots H_{M-1}$ représenteront les alternatives les plus dangereuses.

Maintenant, nous choisirons les *facteurs de coûts* c_{ij} correspondants.

$c_{ii} = 0$:

pas de perte si la décision est correcte.

$c_{ij} = 0$ pour i, j ≥ 2:

pas de perte s'il y a confusion de différentes structures fausses.

$c_{ij} \ll c_{ji}$ pour $i = 0,1;\ j \geq 2$

Nous posons $c_{ij} = \underline{c}$; $c_{ji} = \overline{c}$

c'est-à-dire facteur de coût plus bas si de bonnes structures sont refusées. En revanche coûts élevés

$C_{01} = c_{10} = \underline{c}$

facteur de coût plus faible si 0 et 1 sont confondus.

ou en forme matricielle:

$$\begin{bmatrix} 0 & \underline{c} & \underline{c} & \underline{c} & \cdots & \underline{c} & \underline{c} \\ \underline{c} & 0 & \underline{c} & \underline{c} & \cdots & \underline{c} & \underline{c} \\ \overline{c} & \overline{c} & & & & & \\ \overline{c} & \overline{c} & & & & & \\ \vdots & \vdots & & & & & \\ \vdots & \vdots & & & 0 & & \\ \vdots & \vdots & & & & & \\ \vdots & \vdots & & & & & \\ \overline{c} & \overline{c} & & & & & \end{bmatrix} = [c_{ij}]$$

En introduisant la probabilité conditionnelle $p(\underline{y}|H_i)$, nous pouvons écrire

$$E\{\text{coûts}\} = \sum_{i,j} \int_{\Gamma_j} c_{ij}\, P(H_i)\, p(\underline{y} \mid H_i)\, d\underline{y}$$

$$= \sum_{j} \int_{\Gamma_j} \sum_{i} c_{ij}\, P(H_i)\, p(\underline{y} \mid H_i)\, d\underline{y} \qquad (3)$$

Pour rendre minimum le risque, nous choisirons les sous-ensembles Γ_j de telle manière que les intégrands dans (3) deviennent aussi petits que possible.

$\underline{Y} \in \Gamma_j \longleftrightarrow$ tous $k \neq j$:

$$\sum_i c_{ij} P(H_i)\, p(\underline{y}|H_i) < \sum_i c_{ik} P(H_i)\, p(\underline{y}|H_i) \qquad (4)$$

On pose pour les probabilités à priori $P(H_j)$:

$$P(H_0) = P(H_1) = p \qquad (5)$$

Comme mentionné auparavant, les hypothèses avec $i \geq M$ auront une probabilité négligeable. Donc on aura:

$$\sum_{i=2}^{M-1} P(H_i) = 1 - 2p \qquad (6)$$

Des relations (4), (5) et des hypothèses sur les coefficients c_{ij}, il s'ensuit

$$\underline{Y} \in \Gamma_0 \longleftrightarrow \left\{ \begin{array}{l} p(\underline{y} \mid H_1) < p(\underline{y} \mid H_0) \\ \displaystyle\sum_{i=2}^{M-1} P(H_i) \frac{p(\underline{y} \mid H_i)}{p(\underline{y} \mid H_0)} < (\underline{c}/\overline{c})p \end{array} \right\}$$

$$\underline{Y} \in \Gamma_1 \longleftrightarrow \left\{ \begin{array}{l} p(\underline{y} \mid H_0) < p(\underline{y} \mid H_1) \\ \displaystyle\sum_{i=2}^{M-1} P(H_i) \frac{p(\underline{y} \mid H_i)}{p(\underline{y} \mid H_1)} < (\underline{c}/\overline{c})p \end{array} \right\}$$

autrement – pas accepté. (7)

10.2 Le test suboptimal

On obtient un test plus simple, plus sévère avec (8)

$$\underline{y} \in \Gamma_0 \longleftrightarrow \left\{ \begin{array}{l} p(\underline{y} \mid H_1) < p(\underline{y} \mid H_0) \\ i \geq 2 : \dfrac{p(\underline{y} \mid H_i)}{p(\underline{y} \mid H_0)} < \gamma_i \end{array} \right\}$$

$$\underline{y} \in \Gamma_1 \longleftrightarrow \left\{ \begin{array}{l} p(\underline{y} \mid H_0) < p(\underline{y} \mid H_1) \\ i \geq 2 : \dfrac{p(\underline{y} \mid H_i)}{p(\underline{y} \mid H_1)} < \gamma_i \end{array} \right\}$$

si non, pas accepté, (8)

avec $\gamma_i = (\underline{c}/\overline{c}) \cdot \dfrac{p}{(M-2)\,P(H_i)}$; $(i \geq 2)$

ou $\gamma i = (\underline{c}/\overline{c})\ \dfrac{p}{1-2p}$;

si $P(H_i) = \dfrac{1-2p}{M-2}$; $(i \geq 2)$ (9)

(c'est-à-dire les contrefaçons ont la même probabilité).

Posons $\gamma_0 = \gamma_1 = 1$
et définissons

$$L_{Ki} = \log \frac{p(\underline{y} \mid H_k)}{p(\underline{y} \mid H_i)} = \log \frac{p_k(\underline{y})}{p_i(\underline{y})} ;$$

$$p_i(\underline{y}) = p(\underline{y}|H_i) \qquad (10)$$

alors au lieu de (8) nous pouvons écrire

$\underline{y} \in \Gamma_0 \longleftrightarrow L_{oi} > \log \gamma_i^{-1}$ für $i \neq 0$.

$y \in \Gamma_1 \longleftrightarrow L_{1i} > \log \gamma_i^{-1}$ für $i \neq 1$. (11)

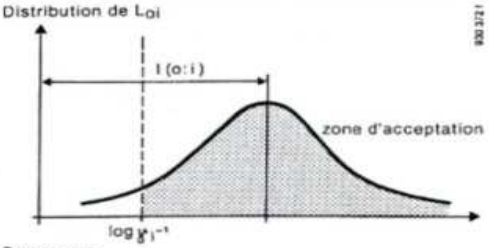
Nous définissons maintenant les grandeurs suivantes

$$I(0:i) = E\{L_{oi} \mid H_o\} = \int p_o(\underline{y}) \log \frac{p_o(\underline{y})}{p_i(\underline{y})}\, d\underline{y}$$

$$\left(= I(p_o : p_i) \right) \qquad (12)$$

Nous appelons I(0:i) *l'information de discrimination* de l'hypothèse H_o par rapport à l'hypothèse H_i.

La *figure 7* montre comment le degré d'acceptation (c'est-à-dire décision pour H_o, avec H_o donné) est influencé par l'information de discrimination.



Supposons

$p(\underline{y} \mid H_o) = n(r_o, \sigma^2)$; c'est-à-dire distribution normale avec valeur moyenne r_o et variante σ^2

$p(\underline{y} \mid H_i) = n(r_i, \sigma^2)$

et posons

$$\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{1}{2}t^2}\, dt, \text{ alors}$$

$$P\{L_{oi} > \log \gamma_i^{-1} \mid H_o\} = \Phi\left\{ \sqrt{\tfrac{1}{2}} \cdot \frac{\log \gamma_i^{-1}}{\sqrt{2I}} \right\}$$

$I = I(o:i) = E\{L_{oi} \mid H_o\}$

Fig. 7 Influence de I (o:i)

En particulier, pour des grandes valeurs de $\log \gamma_i^{-1}$ (par ex. pour une probabilité relativement grande de contrefaçon) nous avons besoin d'un I(0:i) relativement grand pour un degré d'acceptation raisonnable.

10.3 L'information de discrimination

C'est une généralisation de la «mutual information» de Shannon.

Il découle immédiatement de l'inéquation de Jensen [5] que

$$I(p:q) \geq 0 \qquad (13)$$

avec $I = 0 \longleftrightarrow p = q$ (avec une probabilité 1 pour des densités générales p, q). De plus, ce qui suit est valable pour des caractéristiques statistiquement independantes $(y_1, y_2, \ldots y_N)$:

$$I(k:i;y_1, y_2 \ldots y_N) = \sum_{l=1}^{N} I(k:i;y_l) \qquad (14)$$

(Factorisation des densités $p_k(y_1, \ldots y_N)$; $p_i()$).

L'information de discrimination est *invariante* pour des transformations non singulières [3].

$$\underline{X} = T(\underline{Y}) \longrightarrow I(k:i;\underline{X}) = I(k:i;\underline{Y}) \qquad (15)$$

Il découle de (13) à (15) que (I(k:i)) peut être interpretée comme une *distance géométrique* entre les hypothèses H_i et H_k après une transformation correspondante (rotation) des vecteurs $\underline{Y}$. Cette situation devient particulièrement transparente dans le cas de distributions normales avec covariances identiques.

Supposons $p_i(\underline{y}) = n(\underline{r_i}, \underline{\underline{I}})$ (16)

c'est-à-dire les distributions normales avec valeur ri et matrice unité comme matrice de covariance.

De (8) et (11) il découle:

$$\underline{\alpha}_{ik}^{T}(\underline{y} - \underline{r_k}) < d_{ki}$$

$$\text{avec } \underline{\alpha}_{ik}^{T} = (\underline{r_i} - \underline{r_k})^T / (2\, I(k:i))^{1/2} \qquad (17)$$

$$\sqrt{2}\, d_{ki} = (I(k:i))^{1/2} - \log \gamma_i^{-1} / (I(k:i))^{1/2}$$

où maintenant

$$2\, I(k:i) = (\underline{r_k} - \underline{r_i})^T (\underline{r_k} - \underline{r_i}) = \| \underline{r_k} - \underline{r_i} \|^2$$

avec $\| \underline{r_k} - \underline{r_i} \|$ distance vectorielle euclidienne.

L'inégalité (17) signifie simplement que la *projection* du vecteur $(\underline{Y} - \underline{r_k})$ sur le vecteur $(\underline{r_i} \underline{r_k})$ doit être plus petite que d_{ki};

dans un exemple à deux dimensions:

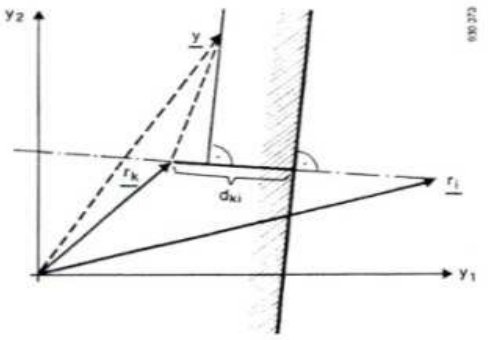


Fig. 8 Projection d'un vecteur dans un exemple à deux dimensions.

c'est-à-dire que $\underline{Y}$ doit être dans le demi-plan hachuré. Dans l'espace à N dimensions, l'ensemble d'acceptation sera limité par des hyperplans. Ceci est vrai en général pour des distributions de probabilités appartenant à la famille «exponentielle» c'est-à-dire pour les distributions les plus importantes en pratique comme la distribution normale, de Poisson etc...
Si nous supposons des distributions normales avec une matrice de covariance Σ, nous aurons au lieu de (17)

$$\underline{\beta}_{ik}^{T}(\underline{y} - \underline{r_k}) < d_{ki} \qquad (18)$$

$$\text{avec } \underline{\beta}_{ik}^{T} = (\underline{r_i} - \underline{r_k})^T \Sigma^{-1} / (2\, I(k:i))^{1/2}$$

$$\text{et } 2\, I(k:i) = (\underline{r_i} - \underline{r_k})^T \Sigma^{-1} (\underline{r_i} - \underline{r_k}) \qquad (19)$$

Dans ce cas, l'information de discrimination est donnée par la position relative des vecteurs r_i, r_k et par la matrice de covariance. La *figure 9* représente le cas simplifié (réduit à 2 dimensions) d'un accepteur.

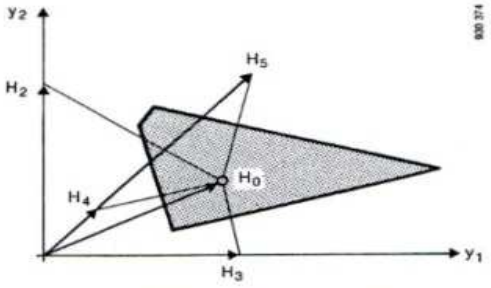


Fig. 9 Exemples avec une seule hypothese permise H_0 espace de décision pour H_0 limites vis-à-vis de $H_2 H_3$. H_4. H_5

Les cas les plus dangereux sont ici:

- considération d'une seule caractéristique (hypothèses H_2, H_3)
- valeurs égales, petites de Y_1 et Y_2 (H_4)
- valeurs égales, maximales de Y_1 et Y_2 (H_5).

Tous les γ_i sont égaux à 1 dans ce cas; c'est-à-dire le rapport des coûts (9) est égal au rapport probabilité de contrefaçon / $P(H_c)$.

Ceci n'est raisonnable que si les cas de contrefaçons ont une petite probabilité.

La *figure 10* montre encore le cas d'une détection 0/1.

10.4 Critères de réalisation

Si tous les autres paramètres sont fixes, une variation de (c/ c̄) changera la probabilité d'acceptation par rapport à la «power of rejection». Une diminution de (c/ c) diminue le degré d'acceptation. Le test en sera plus sévère contre les contrefaçons. En pratique, il faut calculer un nombre de cas pour arriver à un compromis raisonnable.

Les rapports de probabilité $\dfrac{p}{(M\text{-}2)P(H_i)}$

(voir (9)) sont en réalité totalement inconnus.

Nous devrons prendre un «intelligent guess» en fonction de la technologie utilisee.

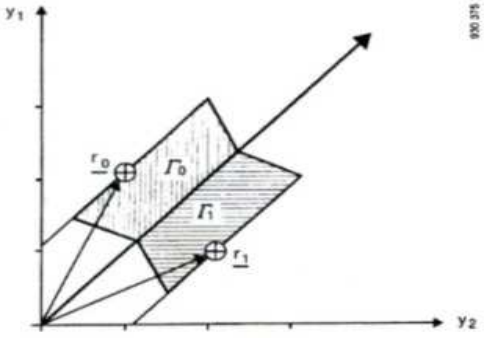


Fig. 10 Exemple de detection 0/1
$y \in \Gamma_0 \rightarrow$ "0" $y \in \Gamma_1 \rightarrow$ "1"
sinon: contrefaçon

L'information de discrimination ne dépend pas seulement de la réalisation de l'accepteur mais aussi de la nature du document.
Si les paramètres de test du document ont une grande dispersion, cette information sera petite conformément à (19).

La réalisation de l'accepteur (à des coûts limités) dépend fortement de la qualité de production du document.

En résumé nous pouvons retenir:
Le choix et la qualité des paramètres du document devraient

1) conduire à une probabilité d'imitation minimale (technologie difficile)

2) apporter une grande information de discrimination contre les contrefacons les plus usuelles.

11. Bibliographie

[1] Diffie, W., Hellman, M.E.: New directions in cryptography; IEEE Transactions on Information Theory, Vol IT-22, No. 6, Nov. 1976.

[2] Feistel, H.; Notz, W.A., Smith, D.L.: Some cryptographic techniques for machine-to-machine data communications. Proceedings of the IEEE, Vol. 63, No. 11, Nov. 1975.

[3] Kullback, S. Information Theory and Statistic (1959), Dover, 1968.

[4] Fukunaga, K. Introduction to Statistical Pattern Recognition, Academic Press NY, 1972.

[5] Feller, W. An Introduction to Probability Theory and its Applications; Vol. II. John Wiley, 1966.

[6] Rivert, R.L.: Shavric, A.: Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems.

Auteur: Heinz Lienhard
LGZ Landis & Gyr Zoug SA
CH-6301 Zoug (Suisse)

Traducteur: J. Clarinval
LGZ Landis & Gyr Zoug SA