Performances d'un Cryptosystème Numérique à base de Chaos

V. GUGLIELMI[1], M. BONNEFONT[1], D. FOURNIER-PRUNARET[1], P. PINEL[1], A.K. TAHA[1]

[1]LESIA, Institut National des Sciences Appliquées, 135 avenue de Rangueil, 31077 Toulouse Cedex

veronique.guglielmi@insa-toulouse.fr

Résumé – Nous présentons les résultats de la cryptanalyse d'un système numérique de cryptographie, système réalisé sur DSP et basé sur l'utilisation de suites chaotiques générées par une récurrence discrète de dimension 2. Nous nous plaçons dans le pire scénario possible quant à la sécurité du chiffrement, ce qui ramène notre attaque à un problème d'estimation des paramètres d'une trajectoire chaotique. La sécurité du cryptosystème est grandement améliorée dès lors que l'on s'interdit la transmission de 2 points successifs de la trajectoire.

Abstract *– We present the results of the cryptanalysis of a digital cryptographic system, implemented on DSP and based on the use of chaotic suites issued from a two-dimensional discrete recurrence. We consider the worst case as regards the cipher security, which reduces our attack to estimating the parameters of a chaotic trajectory. Then, the security of the system is greatly improved if we prohibit the transmission of 2 successive points of the trajectory.*

1. Contexte

L'utilisation de signaux chaotiques pour sécuriser les transmissions est un sujet d'études depuis plusieurs années [1][2][3]. En effet, la grande sensibilité du chaos fait que la même récurrence discrète non linéaire permet de générer, en modifiant légèrement les valeurs de ses conditions initiales ou de ses paramètres, des suites non périodiques de points aux distributions analogues mais qui pourtant ne prendront en fait jamais les mêmes valeurs. Les attracteurs chaotiques peuvent ainsi être assimilés à des générateurs de variables pseudo-aléatoires, chaque suite étant considérée comme la réalisation d'un processus stochastique discret.

Nous avions implémenté sur DSP (Digital Signal Processors) un algorithme de cryptage reposant sur la difficulté à distinguer des trajectoires différentes d'un même attracteur chaotique (construit à partir de récurrences de dimension deux) [4]. Cette réalisation nous avait permis de vérifier le bon fonctionnement de notre cryptosystème et de montrer que toute attaque « à force brute » nécessiterait des temps de calculs prohibitifs [5].

Ici, nous présentons l'analyse cryptographique plus complète que nous avons depuis menée. Nous nous plaçons dans le contexte le pire possible, en ayant au préalable inhibé les niveaux supplémentaires de sécurité que possède le cryptosystème, et en simulant des attaques « à clair choisi ». Le principe de nos attaques dépasse alors le cadre spécifique de notre algorithme de cryptage : il s'agit désormais de retrouver les valeurs des paramètres d'une récurrence chaotique à partir de la seule connaissance de points issus de cette récurrence. Le problème posé est donc maintenant un problème d'estimation non linéaire, qui risque de se retrouver dans tout système de transmission « sécurisée » utilisant des signaux chaotiques en tant que suites pseudo-aléatoires. En effet, l'aspect pseudo-aléatoire des signaux chaotiques repose sur la difficulté à prédire exactement les valeurs prises par ces signaux. Cette difficulté repose entièrement sur la sensibilité du chaos par rapport aux paramètres ; elle n'existe donc plus dès lors qu'il est possible d'estimer les valeurs des paramètres de manière suffisamment précise pour pouvoir reconstruire la trajectoire correspondante de l'attracteur chaotique.

2. Description du cryptosystème

Supposons que le message en clair $\{p_n\}$ soit sous forme binaire, et considérons une récurrence non inversible amenant à des séquences chaotiques $\{s_n\}$. Pour l'implémentation de notre système, les récurrences paraissant pour l'instant les plus appropriées sont des récurrences de dimension égale au minimum à deux [6] :

$$\begin{cases} x_{n+1} = f(x_n, y_n, a, b) \\ y_{n+1} = g(x_n, y_n, a, b) \end{cases}$$

où (x,y) sont les variables réelles, f et g des fonctions non linéaires et a et b des paramètres réels.

Nous prenons pour la séquence chaotique $\{s_n\}$ une seule des deux coordonnées (x_n , y_n), et nous notons pour le point à l'instant n : $s_n = F^n(s_0)$, où s_0 représente la condition initiale de la récurrence. Considérons maintenant deux séquences différentes $\{s_n\} = F^n(s_0)$ et $\{t_n\} = F^n(t_0)$. La n-ième valeur cryptée c_n sera soit s_n soit t_n, selon la valeur du n-ième bit correspondant p_n du message en clair (message supposé sous forme binaire donc) : $c_n = s_n$ si $p_n = 0$; $c_n = t_n$ si $p_n = 1$.

Nous utilisons deux trajectoires du même attracteur chaotique pour les séquences de chiffrement $\{s_n\}$ et $\{t_n\}$, où s_0

et t_0 sont deux conditions initiales différentes choisies dans le bassin d'attraction de l'attracteur. De ce fait, par définition du chaos, les valeurs réelles prises par les suites $\{s_n\}$ et $\{t_n\}$ peuvent être très proches, et distribuées de la même manière. Mais, en même temps, s_n n'est jamais égal à t_n !

Le récepteur déchiffre le message par comparaison des valeurs du texte chiffré c_n avec celles des deux séquences s_n et t_n : $p_n = 0$ si $c_n = s_n$; $p_n = 1$ si $c_n = t_n$. Pour que ce déchiffrement soit possible, il est clair qu'il faut pouvoir reconstruire exactement les deux suites $\{s_n\}$ et $\{t_n\}$ au niveau du récepteur. Nous supposons, afin d'éviter le problème de la synchronisation du chaos entre émetteur et récepteur [7], que les caractéristiques du signal chaotique (paramètres et conditions initiales, en supposant donc l'équation de récurrence connue de tous) ont été échangées au préalable entre les interlocuteurs à titre de clés de communication.

La figure 1 présente le schéma complet du système de cryptage. Notre cryptosystème possède ainsi deux niveaux supplémentaires de sécurité, reposant l'un sur l'utilisation d'une troisième suite pseudo-aléatoire comme bruit additif, l'autre sur l'implémentation d'un masque binaire au niveau du codage en virgule flottante sur le DSP des points issus des suites chaotiques, mais ces deux niveaux sont inhibés pour les attaques présentées ici puisque nous nous plaçons dans le pire contexte possible quant à la sécurité.

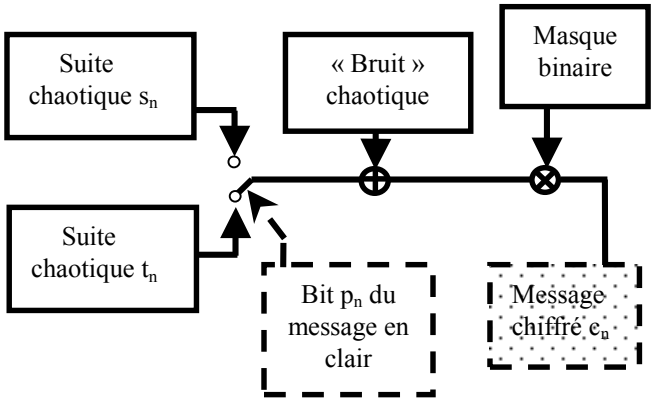

FIG. 1 : Schéma-bloc du codeur

3. Attaques numériques

Dans le cadre de cette communication, nous mènerons notre cryptanalyse en prenant la récurrence cubique suivante :

$$\begin{cases} x_{n+1} = y_n \\ y_{n+1} = a\left(-{x_n}^3 + x_n\right) + b\left(-{y_n}^3 + y_n\right) \end{cases}$$

où (x,y) sont les variables réelles, et a et b des paramètres réels.

Nous avons déjà pu voir par ailleurs [5] qu'il ne s'agit sans doute pas là de la récurrence la meilleure possible quant à la sécurité du système, parce qu'elle est linéaire par rapport aux paramètres a et b, et qu'il s'agit en fait d'une récurrence à une dimension mais de rang 2. Néanmoins, il nous a semblé intéressant de mener nos attaques par rapport à cette récurrence, relativement simple et bien connue, pour en dégager certains résultats sur la sécurité du cryptosystème valables quelle que soit la récurrence utilisée.

Les hypothèses faites pour l'attaque du système [8] sont celles correspondant à une attaque « à clair choisi » : l'attaquant dispose d'une machine de chiffrement, qu'il peut utiliser à sa guise. De ce fait, l'attaquant peut choisir de ne crypter par exemple que des "0", ce qui a pour effet de supprimer le basculement entre les deux suites chaotiques $\{s_n\}$ et $\{t_n\}$: la n-ième valeur cryptée c_n sera toujours égale à s_n pour tout instant n. On suppose de plus que l'attaquant connaît le modèle de récurrence utilisé, mais n'a par contre pas accès à la clé privée que constituent les paramètres a et b et la condition initiale s_0. L'attaque cherche alors à estimer les valeurs de a et de b à partir des points de la suite $\{s_n\}$.

3.1 Des solutions analytiques

La connaissance de 4 points consécutifs d'une même suite, ou même seulement la connaissance de 2 séries de 3 points consécutifs d'une même suite, amène à un système linéaire de 2 équations à 2 inconnues. En effet, supposons par exemple que nous connaissions 4 points successifs s_0, s_1, s_2, s_3 d'une même suite. Cela amène à :

$$\begin{cases} s_2 = b(s_1 - {s_1}^3) + a(s_0 - {s_0}^3) \\ s_3 = b(s_2 - {s_2}^3) + a(s_1 - {s_1}^3) \end{cases}$$

Le problème admet alors une solution analytique :

$$\begin{cases} a = \dfrac{s_3(s_1 - {s_1}^3) - s_2(s_2 - {s_2}^3)}{(s_1 - {s_1}^3)^2 - (s_2 - {s_2}^3)(s_0 - {s_0}^3)} \\ b = \dfrac{s_2(s_1 - {s_1}^3) - s_3(s_0 - {s_0}^3)}{(s_1 - {s_1}^3)^2 - (s_2 - {s_2}^3)(s_0 - {s_0}^3)} \end{cases}$$

Pour remédier à ceci, nous introduisons des décalages dans la récurrence (ce qui correspond à augmenter la dimension de la récurrence), afin d'interdire au pirate la connaissance de points successifs. A partir de la connaissance de 4 points d'indices non consécutifs d'une même suite : $s_0, s_{d1}, s_{d2}, s_{d3}$, où les décalages vérifient: $d3 > d2 > d1 \geq 1$, nous construisons désormais un système polynomial de 3 équations à 3 inconnues :

$$\begin{cases} s_{d1} = Q_1(a,b,s_0) \\ s_{d2} = Q_2(a,b,s_0) \\ s_{d3} = Q_3(a,b,s_0) \end{cases}$$

Les degrés des polynômes Q_1, Q_2, Q_3 varient de manière exponentielle par rapport aux décalages $d1, d2, d3$. Par exemple, pour le polynôme Q_1, pour tout décalage $d \geq 1$, on a :

- Degrés en *a* et *b* : $(3^{d-1}-1)/2$;
- Degrés en s_0 : 3^{d-1}

On peut donc obtenir simplement des polynômes de degrés élevés :

- $d=3$ amène aux degrés 4 en *a* et *b* et 9 en s_0
- $d=4$ amène aux degrés 13 en *a* et *b* et 27 en s_0
- $d=10$ amène aux degrés 9841 en *a* et *b* et 19683 en s_0.

Ceci d'une part empêche toute résolution analytique, et d'autre part amène à un système admettant non pas une solution unique mais bien au contraire un grand nombre de solutions parasites. L'existence de ces solutions parasites va compliquer considérablement la recherche des valeurs correctes des paramètres.

3.2 Par la méthode de Newton-Raphson

Avec les décalages introduits au paragraphe précédent, la résolution analytique du problème posé devient impossible, et il faut passer à une méthode numérique d'optimisation. Celle que nous mettons en œuvre est la méthode itérative de Newton-Raphson.

Dans un premier temps, on s'intéresse au cas où les 2 premiers points des 4 points $s_0, s_{d1}, s_{d2}, s_{d3}$ supposés connus sont en fait 2 points d'indices consécutifs (i.e. $d1=0$).

Le système à résoudre n'a alors plus que 2 inconnues au lieu de 3. La méthode de Newton-Raphson en permet la résolution, en amenant aux valeurs correctes de *a* et de *b*, et ce même pour des décalages importants (de l'ordre de 15 par exemple) et pour une initialisation quelconque (sans information a priori) de l'algorithme d'optimisation.

Par contre, dès lors que les 2 premiers points des 4 points connus $s_0, s_{d1}, s_{d2}, s_{d3}$ ne sont pas des points successifs (i.e. $d1 \geq 1$), l'attaque n'est que marginalement efficace car l'optimisation, même pour des valeurs initiales relativement proches des valeurs cherchées et pour des décalages modestes, n'aboutit pas souvent. Elle converge vers une solution parasite du système, ou bien elle fait intervenir une matrice singulière, ou encore elle diverge.

4. Attaques dans l'espace d'état

4.1 Estimation en une dimension

Le comportement statistique des suites chaotiques est lié aux paramètres. En estimant la densité de probabilité des points issus d'une suite, on montre que cette densité dépend essentiellement des valeurs de *a* et de *b*, et qu'elle ne dépend que faiblement des valeurs des conditions initiales de la récurrence ou des valeurs des décalages entre les points de la suite. Par exemple, pour *a*=2.2 et *b*=-0.91, un attracteur chaotique existe et peut être observé dans le plan de phase de la figure 2. Sa densité de probabilité est donnée par la figure 3.

La comparaison de la densité obtenue pour le message chiffré avec les densités de la base de donnée que le pirate aura pu constituer (pour différentes valeurs des paramètres) amène à une première estimation de *a* et *b*, relativement imprécise mais utile pour initialiser la méthode de Newton-Raphson.

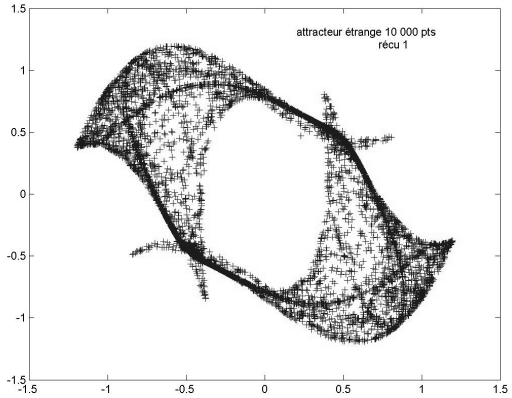

FIG. 2 : Attracteur étrange

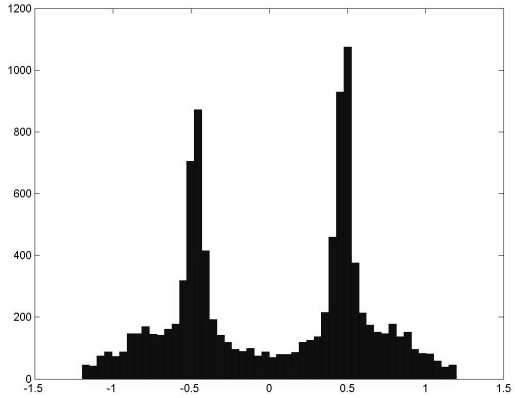
FIG. 3 : Densité de probabilité

4.2 Estimation en dimension supérieure

De manière similaire, le tracé de l'attracteur chaotique dans l'espace d'état (i.e. le tracé du nuage des points (s_n, s_{n+d}) si on note *d* le décalage utilisé lors du cryptage) peut lui aussi permettre une première estimation des paramètres *a* et *b*, et ce afin d'initialiser l'algorithme de Newton-Raphson. En effet, la forme de l'attracteur dans l'espace d'état ne dépend pas des conditions initiales de la suite $\{s_n\}$; elle est caractéristique de zones dans l'espace des paramètres. De plus, ces tracés donnent des informations sur la valeur du décalage *d*, ainsi que sur les positions où surviennent les changements de valeur binaire (de "0" à "1" ou de "1" à "0") dans la suite des bits du message en clair.

Dans le cas où le message en clair n'est pas choisi par le pirate mais correspond à un message réellement transmis, il est intéressant d'utiliser cette possibilité de détection des changements de bits pour essayer de décrypter directement le message. Par exemple, la figure 4 donne le tracé des points (c_n, c_{n+1}) pour un message chiffré avec un décalage constant égal à 2.

Toutefois, cette détection est trop imprécise pour permettre à elle seule un déchiffrement correct. Elle doit donc plutôt être utilisée en support à un algorithme d'estimation des paramètres *a* et *b*, quand l'attaque n'est pas « à clair choisi ».

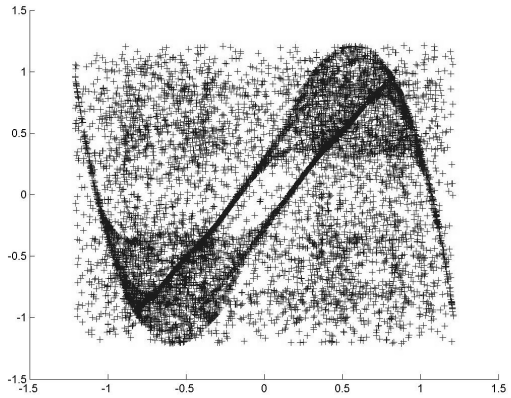

FIG. 4 : Points (c_n, c_{n+1}) d'un message crypté

5. Conclusions

Pour tester la sécurité d'un cryptosystème numérique réalisé dans notre laboratoire, nous effectuons des attaques « à clair choisi ». L'attaque s'interprète alors comme un problème d'estimation non linéaire des paramètres d'une récurrence chaotique, et dépasse ainsi le cadre de la cryptanalyse pour s'attacher à l'aspect pseudo-aléatoire du chaos. L'estimation des paramètres de la récurrence pouvant se ramener à la résolution de systèmes polynomiaux, nous développons un algorithme d'attaque basé sur une méthode numérique d'optimisation (la méthode de Newton-Raphson) et initialisé via l'étude des attracteurs chaotiques dans l'espace d'état. Un résultat principal apparaît : l'attaque réussit dès lors que l'on connaît 2 points d'indices consécutifs de la récurrence ; elle échoue la plupart du temps sinon. Ceci prouve la nécessité d'introduire des décalages dans la suite des points, et d'augmenter ainsi fortement la dimension de la récurrence utilisée. Ces décalages peuvent être constants ou mieux, varier dynamiquement en fonction du temps et des variables du cryptosystème, ce qui améliore la sécurité du chiffrement.

Remerciements

Ce travail a été supporté en partie par le Fonds National de la Science de la République Française (projet TRANSCHAOS de l'ACI 'Sécurité Informatique' 2003).

Références

[1] L. Kocarev, *Chaos-based cryptography: a brief overview*, IEEE Circuits and Systems Magazine, 2001, 1, (3), pp. 6-21.

[2] T. Stojanovski et L. Kocarev, *Chaos-based random number generators-part I: analysis [cryptography]*, IEEE Trans. on Circuits and Systems I, 2001, 48, (3), pp. 281-288.

[3] T. Stojanovski et L. Kocarev, *Chaos-based random number generators-part II: practical realization*, IEEE Trans. on Circuits and Systems I, 2001, 48, (3), pp. 382-385.

[4] L. Bénéteau, D. Fournier-Prunaret, V. Guglielmi, P. Pinel, S. Rouabhi et A.K. Taha, *Two encryption schemes using the chaotic dynamics of two-dimensional noninvertible maps*, Complements to Proceedings of International Conference on Nonlinear Dynamics of Electronic Systems (NDES'02), Juin 2002, Izmir, Turquie.

[5] V. Guglielmi, H. Poonith, D. Fournier-Prunaret, A.K. Taha, *Security performances of a chaotic cryptosystem*, IEEE International Symposium on Industrial Electronics (ISIE '04), Mai 2004, Ajaccio, France.

[6] C. Mira, D. Fournier-Prunaret, L. Gardini, H. Kawakami et J.C. Cathala, *Basin bifurcations of two-dimensional non invertible maps – Fractalization of basins*, International Journal of Bifurcation and Chaos, 1994, 4, (2), pp. 343-382.

[7] Y.-H. Chu et S. Chang, *Dynamical cryptography based on synchronised chaotic systems*, Electronics Letters, 1999, 35, (12), pp. 974-975.

[8] M.I. Sobhy et A.R. Shehata, *Methods of attacking chaotic encryption and countermeasures*, Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP'01), Mai 2001, Salt Lake City, USA.