Un nouveau générateur pseudo-aléatoire

René Blacher

Un nouveau générateur pseudo-aléatoire

Rene BLACHER

Résumé : Nous introduisons un nouveau générateur pseudo-aléatoire. Il transforme une suite aléatoire de taille N en un très grand nombre de suites aléatoires. La réunion de ces suites forme une suite de taille extrêmement longue. Cette suite a des lois marginales uniformes. De plus, tous les couples et tous les triplets sont indépendants. Les dépendances n'existent que pour les p-uplets tels que $p \geq 4$. Mais ils n'y a qu'une probabilité infime de trouver une telle dépendance tant que $p \leq N$. De plus, pour $p \geq N$, les dépendances peuvent exister mais elles sont impossibles à déterminer.

Summary : We study a new random generator. It transforms a random sequence with size N in a very great number of random sequences. The union of these sequences forms a sequence with a extremely long size. This sequence has marginal distributions which are uniform. Moreover, dependences exist only for p-tuples such that $p \geq 4$. But they is only a negligible probability to find a such dependence as long as $p \leq N$. Moreover, for $p \geq N$, dependences can exist but they are impossible to determine.

Mots-clefs : Générateur pseudo-aléatoire, Nombres aléatoires, dépendance.

Key words : Pseudo random generator, random numbers, dependence.

1 Introduction

Nous allons définir dans ce rapport un nouveau générateur pseudo-aléatoire. Il sera particulièrement efficace car il utilise des suites déjà réellement aléatoires.

On les transforme alors par un algorithme $\mathcal{B}$ qui est simple et rapide. Celui-ci va transformer ces suites réellement aléatoires en suites presque aléatoires. Pour démontrer cela, on utilisera essentiellement les propriétés des nombres aléatoires.

1.1 Description de l'algorithme $\mathcal{B}$

Nous allons transformer une suite aléatoire de taille NI en une suite de taille quasi-infinie. Pour cela, on la réécrit sous forme de matrice $x_{i,j} \in \{0, 1, ..., m-1\}$ à I lignes et N colonnes (par exemple I=20, $N = 1.000.000$) :

$$\begin{pmatrix} x_{1,1}x_{1,2}\cdots\cdots x_{1,N} \\ x_{2,1}x_{2,2}\cdots\cdots x_{2,N} \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ x_{I,1}x_{I,2}\cdots\cdots x_{I,N} \end{pmatrix} .$$

Ensuite on décale chaque ligne i, sauf la première, de $d_i \in \{0, 1, ..., N-1\}$. Par exemple, le décalage de 1 de la ligne 2 est défini par

$$x(2,:) = \big(x_{2,1}, x_{2,2}, ..., x_{2,N}\big) \ \mapsto \ \big(x_{2,2}, ..., x_{2,N}, x_{2,1}\big) \ .$$

Le décalage de 2 est défini par

$$x(2,:) = \big(x_{2,1}, x_{2,2}, x_{2,3}, ..., x_{2,N}\big) \ \mapsto \ \big(x_{2,3}, ..., x_{2,N}, x_{2,1}, x_{2,2}\big) \ .$$

On obtient donc une matrice décalée $\{x_{i,j}^d\}$ ou $d = (d_2, d_3, ..., d_I)$:

$$\begin{pmatrix} x_{1,1} & x_{1,2}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots x_{1,N} \\ x_{2,1+d_2} & x_{2,2+d_2}\cdots\cdots\cdots x_{2,N} \quad x_{2,1} \ \ x_{2,2}\cdots\cdots x_{2,d_2} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_{I,1+d_I} & x_{I,2+d_I}\cdots\cdots x_{I,N} \quad x_{I,1} \ x_{I,2}\cdots\cdots\cdots x_{I,d_I} \end{pmatrix} .$$

Ensuite on somme les I lignes modulo m. On obtient donc une suite de N termes (dans l'exemple 10^6 termes) :

$$x_1^d = \overline{x_{1,1} + x_{2,1+d_2} + ... + x_{I,1+d_I}}$$
$$x_2^d = \overline{x_{1,2} + x_{2,2+d_2} + ... + x_{I,2+d_I}},$$
$$..$$

avec la notation suivante

Notations 1.1 *Soit $h \in \mathbb{Z}$ et $m \in \mathbb{N}^*$. On définit $\overline{h}^m$ par*
1) $\overline{h}^m \equiv h$ modulo m.
2) $0 \leq \overline{h}^m < m$.

Quand le choix de m est clair, on simplifie $\overline{h}^m$ en $\overline{h}$.

On obtient donc une suite de N nombres aléatoires x_n^d, n=1,2,...,N, pour chaque décalage $d = (d_2, d_3, ..., d_I)$. On verra qu'il est mieux de ne garder que les $N - q$ premiers. On obtient donc une suite x_n^d, n=1,2,3,...,N-q. En général, on choisit $q = \lfloor N/2 \rfloor$, la partie entière de N/2.

Il y a N décalages d_i possibles par ligne i. On décale I-1 lignes. Donc au total il y a N^{I-1} décalages $d = (d_2, d_3, ..., d_I)$ possibles : par exemple, si $N = 10^6$ et I=21, il y a $(10^6)^{20} = 10^{120}$ décalages possibles.

C'est cela, l'algorithme $\mathcal{B} : \{x_{i,j}\} \mapsto \mathcal{B}(\{x_{i,j}\})$. Il permet de produire des suites x_n a peu près aléatoire de taille quasi infinie en prenant toutes les suites x_j^d possibles que l'on appelera blocs.

Définition 1.2 *On appelle bloc les différentes matrices $\{x_{i,j}^d\}$associées à des décalages vectoriels $d = (d_2, d_3, ..., d_I)$ ou bien la matrice ligne résultant des sommes modulo m que l'on note par $\{x_j^d\}$.*

Bien sûr, il serait mieux de noter l'algorithme $\mathcal{B}$ par $\mathcal{B}^{N,I}$ plutôt que $\mathcal{B}$. Mais pour simplifier on se contente de $\mathcal{B}$.

On obtient donc la suite x_n en réunissant dans un certain ordre toutes les suites x_n^d, n=1,2,....,N-q : $d = 1, 2, ..., N^{I-1}$ lorsque on indicie linéairement les décalages, i.e. si on définit les décalages $d(t) = (d_2(t), d_3(t), ..., d_I(t))$ en fonction de $t = 1, 2, ..., N^{I-1}$.

Donc x_n est de taille $(N-q)N^{I-1}$, ce que l'on peut considérer comme quasi infinie dès que les paramêtres sont assez grands.

Pour obtenir la suite x_n comme suite pseudo-aléatoire, le mieux, c'est d'opérer successivement : quand on a obtenu la première matrice décalée $\{x_{i,j}^1\}$, on opère les nouveaux décalages, sur la matrice $\{x_{i,j}^1\}$, et non sur la matrice originale $\{x_{i,j}\}$, et ainsi de suite.

1.2 Etude des suites x_n^d

Comme $x_{i,j}$ est un échantillon IID (Independent Identically Distributed), on peut admettre l'hypothèse suivante.

Hypothèse 1.1 *On suppose qu'il existe une matrice aléatoire IID $\{X_{i,j}\}$, $X_{i,j} \in \{0, 1, ..., m-1\}$, définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ qui soit un modèle correct de $\{x_{i,j}\}$ (cf [13]) : il existe $\omega \in \Omega$ vérifiant $x_{i,j} = X_{i,j}(\omega)$ for all (i,j).*

On pourra donc écrire aussi $x_j^d = X_j^d(\omega)$ et $x_j = X_j(\omega)$.

On utilisera alors le théorème suivant (th 5 page 74 of [11]).

Théorème 1 *Soient X et Y deux vecteurs aléatoires indépendants, $X,\ Y \in \{0, 1, ..., m-1\}^p$. On suppose que X est de loi uniforme sur $\{0, 1, ..., m-1\}^p$. Alors, $\overline{X+Y} \in \{0, 1, ..., m-1\}^p$ suit aussi la loi uniforme.*

On en déduira le théorème suivant.

Théorème 2 *La suite X_j^d, j=1,2,....,N est IID.*

Théorème 3 *Soient 3 décalages d^1, d^2, d^3. Soient $j_s \in \{1, 2, ..., N\}$, s=1,2,3, tels que $X_{j_1}^{d^1} \neq X_{j_2}^{d^2}$, $X_{j_1}^{d^1} \neq X_{j_3}^{d^3}$, $X_{j_2}^{d^2} \neq X_{j_3}^{d^3}$.*

Alors, $(X_{j_1}^{d^1}, X_{j_2}^{d^2}, X_{j_3}^{d^3})$ suit la loi uniforme sur $\{0, 1, ..., m-1\}^3$.

Ce dernier théorème est prouvé en section 2.3. Donc les suites X_j^d vérifient les propriétés de dépendance et d'uniformité les plus importantes. En effet, chaque X_j^d suit la loi uniforme sur $\{0, 1, ..., m-1\}$. De plus les couples $(X_{j_1}^{d^1}, X_{j_2}^{d^2})$ et les triplets $(X_{j_1}^{d^1}, X_{j_2}^{d^2}, X_{j_3}^{d^3})$ sont indépendants. Enfin, chaque suite X_j^d, j=1,2,....,N, est IID

Bien sûr, pour que toute la suite X_j soit IID, il faudrait qu'elle vérifie l'indépendance de chaque $(X_{j_1}^{d^1}, ..., X_{j_p}^{d^p})$ pour tout $p \in \mathbb{N}^*$. On se doute bien que ce ne sera pas le cas.

D'ailleurs on verra qu'il existe des dépendances de quadruplets (cf section 8) . Mais dès que les paramêtres sont bien choisis, il y a très peu de chances d'en trouver. En effet, on a le théorème suivant (cf Corollaire 3.3).

Proposition 1.1 *Si on choisit une quadruplet $(X_{j_1}^{d^1}, X_{j_2}^{d^2}, X_{j_3}^{d^3}, X_{j_4}^{d^4})$ au hasard, il y a au plus*

$$\frac{3^I N^{2I}}{N^I(N^I-1)(N^I-2)(N^I-3)}$$

chances de tomber au hasard sur un quadruplet dépendant.

Par exemple si $I = 21$, $N = 10^{10}$, on a moins d'une chance sur 10^{390} de trouver un quadruplet dépendant au hasard : l'immense majorité des quadruplets seront indépendants.

Maintenant, un cryptanalyste pourrait quand même vouloir trouver une dépendance : cela lui apporterait quelques renseignements sur une petite partie de la suite x_n. Mais, à cause de cette propriété 1.1, il n'a aucune chance de trouver une telle dépendance en procédant au hasard.

Maintenant, en étudiant les dépendances, on verra en sections 5, 6 et 7 que on peut supprimer certaines dépendances lorsque on choisit m premier et $p \leq m$.

C'est encore vrai pour des p-uplets où p est plus grand. En effet, la probabilité est de plus en plus faible de trouver des p-uplets dépendants lorsque p augmente, tout au moins au début. On peut donc par exemple imposer que ce soit vrai pour tout $p \leq N$.

Hypothèse 1.2 *On choisira m premier et vérifiant $N \leq m$.*

De la sorte on élimine des dépendances pour tout $p \leq N$.

Cependant lorsque p est assez grand, cela change et la probabilité diminue. Mais ce n'est pas grave parce que trouver un N-uplet au hasard a encore une probabilité infime.

Or, cela ne sert à rien d'étudier des p-dépendances lorsque $p \geq log(n_0)/log(2)$ si on a un échantillon de taille n_0 (cf Remark 2.1.1 page 23 de [11]). Ici $n_0 \leq (N-q)N^{I-1}$. Par exemple si I=20, $N = 10^6$, $n_0 \leq (N-q)N^{I-1} \leq 10^{120}$. Donc $log(n_0)/log(2) \leq 120Log(10)/log(2) \approx 398.6$. Il n'y a donc à etudier les dépendances que jusqu'à p=398.

Maintenant, pour des p-uplets où p est plus grand que N, on verra en section 4 que cela n'a guère d'importance si $m \geq p$. En effet, les dépendances se traduiront par des relations linéaires modulo m du type $\sum_{s=1}^{p} \beta_s x_{j_s}^{d^s} \equiv \alpha$ où $\beta_s \in \mathbb{N}$ avec $\beta_s \leq p$.

Or ces dépendances restent identiques tant que chaque $x_{j_s}^{d^s}$ reste dans un bloc de décalage d^s. Après les décalages changent et les dépendances disparaissent ou bien ce ne sont plus les mêmes. Dans ce cas, les relations changent. Donc on devra seulement trouver des relations $\sum_{s=1}^{p} \beta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour $r = 0, 1, 2, ..., N_1$ où $N_1 \leq N$.

Or il est facile de voir que, *même s'il y a indépendance*, on trouvera toujours des équations vérifiant $\sum_{s=1}^{p} \beta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour $r = 0, 1, 2, ..., N_1$ si $m \leq p$ et $p \geq N$: cf section 4.2.3.

Si on veut trouver une relation linéaire, on pourra donc toujours le faire même s'il y a indépendance. En fait, on risque même d'en trouver un très grand nombre. Cela veut dire que l'on ne peut pas détecter de dépendance si $m \leq p$ lorsque $p \geq N$.

Pour que ce soit vrai pour tout $p \geq N$, on imposera donc l'hypothese suivante

Hypothèse 1.3 *On choisira m et N tel que $m \leq N$.*

Donc finalement on prendra m=N.

De plus, si un cryptanalyste essaie de retrouver une partie de la suite x_n par les dépendances, le plus simple pour lui sera de chercher les dépendances les plus faciles à trouver : celles des quadruplets.

Dans ce cas, si il y a dépendance, pendant N_1 termes successifs, $N_1 \leq N$, on aura une relation linéaire, par exemple, $x_{j_1+r}^{d^1} - x_{j_2+r}^{d^2} + x_{j_3+r}^{d^3} - x_{j_4+r}^{d^4} \equiv 0$ modulo m pour $r = 0, 1, ..., N_1$.

Cela voudra dire que s'il connait $x_{j_1+r}^{d^1}$, $x_{j_2+r}^{d^2}$, $x_{j_3+r}^{d^3}$, il connaitra $x_{j_4+r}^{d^4}$.

Mais ce sera pendant un certain temps seulement, un temps plus petit que N_1, et il ne pourra rien déduire d'autre sur le comportement de la suite x_j^d. Ce n'est donc pas une découverte trop intéressante pour lui.

1.2.1 Recherche de dépendances particulières

Maintenant, on a calculé la probabilité de trouver des dépendances au hasard. Il se pose alors la question de savoir si il n'y a pas certaines dépendances qu'on a plus de chances de trouver si on les prend de façon réfléchie, et non au hasard.

On verra que c'est le cas. Pour cela on prendra d'abord le cas des quadruplets. Prendre des quadruplets au hasard, cela revient a priori à choisir une dépendance de 4 termes, chacun appartenant à un bloc différent : donc 4 blocs. Mais on verra que on peut trouver des 4-dépendances entre deux blocs : cf section 8.1.

Dans ce cas, la première ligne devra être composée des élements $X_{1,n}$ et $X_{1,n+D}$ pour chacun des deux blocs. En plus il faut que D=N/2, et donc que N soit pair. Donc si ce n'est pas le cas, une telle dépendance n'existera pas. De même si on prend que les N/2 premiers termes de chaque blocs ($q = \lfloor N/2 \rfloor$), une telle dépendance ne pourra pas apparaitre.

Mais s'il y a une telle dépendance, la probabilité de la trouver est égale à $(2/N)^{I-1}$. On voit que c'est une probabilité bien plus grande que celle obtenue quand on cherche les quadruplets au hasard.

Mais c'est encore une probabilité infime : la probabilité de trouver un tel quadruplet, si I=21, $N = 10^{10}$, est égale à $2^{20}/10^{200}$.

On peut alors se demander si il n'y a pas des dépendances qu'il soit plus facile de trouver. La réponse est oui : par exemple, il peut y avoir une dépendance entre deux blocs lorsque $N = 2N'$ et lorsque la première ligne de chacun de ces bloc est $(X_{1,2}, X_{1,4}, X_{1,6}, ..., X_{1,2N'})$ et $(X_{1,2}, X_{1,4}, X_{1,6}, ..., X_{1,2N'})$.

Dans ce cas, si il y a dépendance, on aura environ une chance sur 2^{I-1} d'avoir une telle dépendance. Par exemple si I=21, on a une chance sur 2^{20} de trouver une dépendance. C'est beaucoup plus que pour les dépendances de quadruplets que nous venons d'étudier.

Mais pour trouver une telle dépendance, il faut que $N = 2N'$ et que l'on ne prenne pas seulement les N/2 premiers termes de chaque bloc. Donc il ne faut pas que N soit premier pour pouvoir trouver une telle dépendance. Aussi on imposera les hypothèses suivante

Hypothèse 1.4 *On imposera que N soit premier et que $q = \lfloor N/2 \rfloor$.*

Or, ce résultat restera vrai dans beaucoup de cas : il y a beaucoup de dépendances potentielles qui seront éliminées avec cette hypothèse. On en étudiera quelques unes en section 9.2.

En fait il semble que, de la sorte, on aura éliminé les dépendances qu'il y a une chance non-infime d'obtenir.

Donc, récapitulons : on imposera les hypothèses suivantes.

Hypothèse 1.5 *On imposera que N=m soit premier et que $q = \lfloor N/2 \rfloor$.*

Maintenant, on peut aussi prendre des hypothèses moins restrictives.

Hypothèse 1.6 *On imposera que N soit premier, que $m \leq N$ et que $q = \lfloor N/2 \rfloor$.*

Sous cette hypothèse, on peut par exemple choisir m=2, c'est à dire supposer que les $x_{i,j}$ soient des bits aléatoires. Cela peut simplifier les calculs électroniques. Mais de la sorte certaines dépendances de probabilité infime n'ont pas été éliminées (cf section 5, par exemple proposition 5.1).

1.3 Système de cryptage

Ayant un générateur pseudo-aléatoire, on a un système de cryptage à clef secrète. On va voir qu'il est extrêmement performant car indécryptable, simple et rapide.

On vient de voir sa définition. Il est donc clair qu'il est simple et rapide. Il reste à voir qu'il est bien indécryptable.

On choisira donc une matrice aléatoire $\{x_{i,j}\} \in \{0, 1, ..., m\}$ telle que m=N est premier. A chaque décalage d, on prendra seulement les $\lfloor N/2 \rfloor$ premiers termes de la suite x_j^d.

1.3.1 Définition linéaire des décalages

Pour définir la suite x_n, réunion des x_j^d, on définit les décalages en fonction d'un seul paramêtre t : d=d(t). On peut utiliser des congruences pour cela. Mais, on peut aussi définir ces décalages de façon à ce qu'ils soient choisis au hasard (cf section 10.2).

1.3.2 Système de cryptage

On a déjà vu qu'il y a indécryptabilité des dépendances. Maintenant, on a aussi l'indécryptabilité complète du système de cryptage. Il y a en effet 2 cas possibles.

Cas où $\{x_{i,j}\}$ est connue Dans ce cas, l'indécryptabilité signifie que si on a une suite x_j^d on ne peut pas retrouver le décalage "d" en un temps correct. Dans ce sens, normalement, l'algorithme $\mathcal{B}$ est incassable *même quand* $\{x_{i,j}\}$ *est connue*. En effet, cette question est une variante du "Subset Sum Problem" (cf p 117-122 [9]).

Cas où $\{x_{i,j}\}$ est inconnue Il est clair que, si le système est indécryptable *même lorsque* $\{x_{i,j}\}$ *est connue*, il le sera de façon absolue lorsque $\{x_{i,j}\}$ est inconnue. En effet, on devra d'abord retrouver la matrice $\{x_{i,j}\}$, ce qui est impossible.

Par exemple, si $\{x_{i,j}\}$ est une matrice de 10^8 chiffres, on devra essayer chaque matrice possible $\{x_{i,j}\}$. Parce qu'il y a $10^{100.000.000}$ telles matrices possibles $\{x_{i,j}\}$, il est tout à fait impossible de casser le système : cf aussi section 12.3.

Applications D'abord, on peut facilement avoir une sécurité équivalente à celle du cas où $\{x_{i,j}\}$ est inconnue en faisant une première modification de la matrice $\{x_{i,j}\}$: on transforme la matrice $\{x_{i,j}\}$ par une transformation cryptographiquement forte, mais moins rapide. Ensuite, celle-ci pourra être considérée comme inconnue. On pourra donc y appliquer l'algorithme $\mathcal{B}$ qui, lui est rapide. On peut appliquer cette méthode pour crypter les conversations téléphoniques.

D'autre part, contrairement au VOTP, le système peut aussi servir à l'authentification.

Enfin, ce système sera particulièrement efficace dans un réseau avec ordinateur central où la matrice aléatoire $\{x_{i,j}\}$ sera inconnue. Dans ce cas, on a bien une méthode pour utiliser autrement le Vernam One Time Pad.

Tout ceci est détaillé en section 12.

1.4 Généralisation

On peut généraliser l'algorithme $\mathcal{B}$ de plusieurs façons.

D'abord on peut supprimer les dépendances en ajoutant un système cryptographique quelconque : au lieu de transformer un message M par $\overline{M(j) + x_j^d}$, on peut le transformer par $\overline{C(M(j) + x_j^d)}$ où C est un algorithme cryptographique quelconque, par exemple le DES. Cela peut permettre de donner une nouvelle vie à ces systèmes.

D'autre part, les décalages sont des permutations très particulières. On peut donc remplacer les décalages de chaque ligne par des permutations assez rapides Pe^s. On obtient alors en sommant modulo m les colonnes des suites de nombres $x_j^{Pe^s}$ construites par blocs.

Troisièmement, plutôt qu'utiliser les sommes sur toutes les lignes de la matrice, on peut sommer seulement certaines lignes différentes pour chaque bloc.

Enfin chaque suite $x_j^{d^s}$ peut être réécrite sous forme de matrice et transformée par l'algorithme $\mathcal{B}$ en une suite infiniment plus longue.

2 Propriétés élémentaires

2.1 Généralisation des notations

Certaines des propriétés obtenues avec des décalages restent vraies avec des permutations. Il est donc plus simple de démontrer tout de suite ces propriétés sous ces hypothèses.

En effet, l'algorithme $\mathcal{B}$ peut facilement être généralisé à d'autre cas que les décalages : les décalages sont donc des permutations définies par

$$\big(x_{i,1}, x_{i,2}, ..., x_{i,N}\big) \ \mapsto \ \big(x_{i,d_i+1}, x_{i,d_i+2}, ..., x_{i,N}, x_{i,1}, x_{i,2}, , ..., x_{i,d_i}\big) \ .$$

On étend donc l'utilisation de l'algorithme $\mathcal{B}$ aux permutations Pe.

Notations 2.1 *Soient Pe_s, s=1,2,....,I, I permutations de $\{1,2,...,N\}$. On suppose $Pe_1 = Id$. On note par $\{X_{i,j}^{Pe}\}$ la matrice aléatoire telle que, pour tout $i \in \{1,2,3...,I\}$, $(X_{i,1}^{Pe}, X_{i,2}^{Pe}, ..., X_{i,N}^{Pe}) = (X_{i,Pe_i(1)}, X_{i,Pe_i(2)}, ..., X_{i,Pe_i(N)})$. On note par $(X_1^{Pe}, ..., X_N^{Pe})$ le vecteur aléatoire tel que $X_j^{Pe} = \overline{X_{1,j}^{Pe} + X_{2,j}^{Pe} + ... + X_{I,j}^{Pe}}$.*

2.2 Etude des suites x_j^{Pe}

L'étude de la dépendance repose sur les théorèmes suivants.

Théorème 4 *Pour toute permutation Pe, pour tout $j \in \{1,2,...,N\}$, X_j^{Pe} suit la loi uniforme sur $\{0,1,....,m-1\}$.*

Ce théorème est une conséquence du théorème 1.

Théorème 5 *Soit $p \in \mathbb{N}^*$. Soit $(X_{j_1}^{Pe^1}, ..., X_{j_p}^{Pe^p})$ un vecteur aléatoire tel que les $X_{j_s}^{Pe^s}$ soient tous différents. On suppose qu'il existe i_0 tel que $X_{i_0,j_1}^{Pe^1}$ est indépendant des autres $X_{i_0,j_s}^{Pe^s}$, $s \neq 1$. Alors, pour tout $p \in \{2,3,....,N\}$, pour tout $(j_1,....,j_p) \in \{1,2,....,N\}^p$,*

$$P\Big\{X_{j_1}^{Pe^1} = b_1\} \cap ..... \cap \{X_{j_p}^{Pe^p} = b_p\}\Big\} = (1/m) P\Big\{\{X_{j_2}^{Pe^2} = b_2\} \cap ..... \cap \{X_{j_p}^{Pe^p} = b_p\}\Big\} \ .$$

Démonstration Pour prouver ce résultat, nous utiliserons le lemme suivant. Sa démonstration est évidente.

Lemme 2.1 *On pose $(X'_1, ..., X'_p) = (X_{j_1}^{Pe^1}, ..., X_{j_p}^{Pe^p})$. Alors il existe des variables aléatoires U, Y_s, s=1,2,...,r, W_j, $j \in \mathcal{E} = \{s_1, s_2, ..., s_r\} \subset \{2,3,...,p\}$ et T_j, $j \notin \mathcal{E}$, telles que*

a) $X'_1 = \overline{U + Y_1 + ... + Y_r}$,

b) $X'_j = \overline{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h}$ si $j \in \mathcal{E}$ où $\delta_{j,h} = 0$ ou bien $\delta_{j,h} = 1$ quand h=1,2,...,r,

c) $X'_j = T_j$, si $j \notin \mathcal{E}$,

où

1) U est de loi uniforme,

2) U est indépendant de $(Y_1, ..., Y_r, W_{s_1}, ..., W_{s_r}, T_{j_1}, ..., T_{j_{p'}})$, où $p' = p - r - 1$.

3) Les T_j sont indépendants des Y_r.

Grâce a ce lemme, on a les égalités suivantes

$$P\Big\{X'_1 = b_1\} \cap ..... \cap \{X'_p = b_p\}\Big\}$$

$$= P\Big\{\{\overline{U + Y_1 + ... + Y_r} = b_1\} \cap \big\{ \cap_j \{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h = b_j\}\big\} \cap \big\{ \cap_t \{T_t = b_t\}\big\} \ \Big\}$$

$$= P\Big\{\Big[\cup_{d_1,..,d_r} \{Y_1 = d_1\} \cap .... \cap \{Y_r = d_r\}\Big] \cap \Big[\{\overline{U + Y_1 + .. + Y_r} = a_1\} \cap \Big\{ \cap_j \{\overline{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h} = b_j\}\Big\} \cap \Big\{ \cap_t \{T_t = b_t\}\Big\} \ \Big]\Big\}$$

$$= P\Big\{ \cup_{d_1,\ldots,d_r} \Big[\{Y_1 = d_1\} \cap \ldots\ldots\ldots \cap \{Y_r = d_r\} \cap \{\overline{U + Y_1 + .. + Y_r} = a_1\} \cap \Big\{ \cap_j \{\overline{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h} = b_j\} \Big\} \cap \Big\{ \cap_t \{T_t = b_t\} \Big\} \Big] \Big\}$$

$$= \sum_{d_1,\ldots,d_r} P\Big\{ \{Y_1 = d_1\} \cap \ldots\ldots\ldots \cap \{Y_r = d_r\} \cap \{\overline{U + Y_1 + .. + Y_r} = a_1\} \cap \Big\{ \cap_j \{\overline{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h} = b_j\} \Big\} \cap \Big\{ \cap_t \{T_t = b_t\} \Big\} \Big\}$$

$$= \sum_{d_1,\ldots,d_r} P\Big\{ \{Y_1 = d_1\} \cap \ldots\ldots\ldots \cap \{Y_r = d_r\} \cap \{U = \overline{a_1 - d_1 - \ldots. - d_r}\} \cap \Big\{ \cap_j \{W_j = \overline{b_j - \sum_{h=1}^{r} \delta_{j,h} d_h}\} \Big\} \cap \Big\{ \cap_t \{T_t = b_t\} \Big\} \Big\}$$

$$= \sum_{d_1,\ldots,d_r} P\{U = \overline{a_1 - d_1 - \ldots. - d_r}\} P\Big\{ \{Y_1 = d_1\} \cap \ldots\ldots\ldots \cap \{Y_r = d_r\} \cap \Big\{ \cap_j \{W_j = \overline{b_j - \sum_{h=1}^{r} \delta_{j,h} d_h}\} \Big\} \cap \Big\{ \cap_t \{T_t = b_t\} \Big\} \Big\}$$

$$= (1/m) \sum_{d_1,\ldots,d_r} P\Big\{ \{Y_1 = d_1\} \cap \ldots\ldots\ldots \cap \{Y_r = d_r\} \cap \big\{ \cap_j \{W_j = \overline{b_j - \sum_{h=1}^{r} \delta_{j,h} d_h}\} \big\} \cap \big\{ \cap_t \{T_t = b_t\} \big\} \Big\}$$

$$= (1/m) \sum_{d_1,\ldots,d_r} P\Big\{ \{Y_1 = d_1\} \cap \ldots\ldots\ldots \cap \{Y_r = d_r\} \cap \big\{ \cap_j \{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h = b_j\} \big\} \cap \big\{ \cap_t \{T_t = b_t\} \big\} \Big\}$$

$$= (1/m) P\Big\{ \cup_{d_1,\ldots,d_r} \Big[\{Y_1 = d_1\} \cap \ldots\ldots\ldots \cap \{Y_r = d_r\} \cap \big\{ \cap_j \{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h = b_j\} \big\} \cap \big\{ \cap_t \{T_t = b_t\} \big\} \Big] \Big\}$$

$$= (1/m) P\Big\{ \Big[\cup_{d_1,\ldots,d_r} \{Y_1 = d_1\} \cap \ldots\ldots\ldots \cap \{Y_r = d_r\} \Big] \cap \Big[\big\{ \cap_j \{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h = b_j\} \big\} \cap \big\{ \cap_t \{T_t = b_t\} \big\} \Big] \Big\}$$

$$= (1/m) P\Big\{ \big\{ \cap_j \{W_j + \sum_{h=1}^{r} \delta_{j,h} Y_h = b_j\} \big\} \cap \big\{ \cap_t \{T_t = b_t\} \big\} \Big\}$$

$$= (1/m) P\Big\{ \{X'_2 = b_2\} \cap \ldots\ldots\ldots \cap \{X'_p = b_p\} \Big\} \ . \ \blacksquare$$

Ce résultat signifie donc que $X_{j_1}^{Pe^1}$ est indépendant de $(X_{j_2}^{Pe^2}, ..., X_{j_p}^{Pe^p})$. Donc pour découvrir les dépendances sous ces hypothèses, il ne sert à rien de considérer $X_{j_1}^{Pe^1}$. Il faut donc se concentrer sur l'étude de $(X_{j_2}^{Pe^2}, ..., X_{j_p}^{Pe^p})$ bien que le vecteur $(X_{j_1}^{Pe^1}, ..., X_{j_p}^{Pe^p})$ puisse être considéré comme dépendant.

Il convient donc de définir les vecteurs qu'il est bon d'étudier

Définition 2.2 *On dit qu'un vecteur aléatoire $(Z_1, Z_2, ..., Z_p)$ est complètement dépendant ou a une dépendance complète s'il n'existe pas deux vecteurs indépendants $(Z_{\phi(1)}, Z_{\phi(2)}, ..., Z_{\phi(p')})$ et $(Z_{\phi(p'+1)}, Z_{\phi(p'+2)}, ..., Z_{\phi(p'+p")})$ où p'+p"=p, p', p" $\in \mathbb{N}^*$ et où ϕ est une permutation de $\{1, 2, ..., p\}$.*

En particulier si $(X_{j_1}^{Pe^1}, ..., X_{j_p}^{Pe^p})$ est complètement dépendant, $X_{j_1}^{Pe^1}$ n'est pas indépendant de $(X_{j_2}^{Pe^2}, ..., X_{j_p}^{Pe^p})$.

Donc, on déduit du théorème précédent une condition pour qu'il y aie dépendance complète.

Théorème 6 *Soit $(X_{j_1}^{Pe^1}, ..., X_{j_p}^{Pe^p})$ un vecteur aléatoire où les $X_{j_s}^{Pe^s}$ sont tous distincts.*

Alors pour qu'il y aie dépendance complète, il faut que pour tout i, pour tout s, il existe $t \neq s$ tel que $X_{i,j_s}^{Pe^s} = X_{i,j_t}^{Pe^t}$.

2.3 Indépendance des couples et des triplets

On étudie maintenant la dépendance entre deux vecteurs aléatoires. On considère donc un vecteur aléatoire $(X_{j_1}^{d^1}, X_{j_2}^{d^2})$ où

$$X_{j_1}^{d^1} = U * \begin{pmatrix} X_{1,j_1} \\ X_{2,\overline{j_1+d_2^1}} \\ \dots\dots \\ \dots\dots \\ X_{I,\overline{j_1+d_I^1}} \end{pmatrix}, \; X_{j_2}^{d^2} = U * \begin{pmatrix} X_{1,j_2} \\ X_{2,\overline{j_2+d_2^2}} \\ \dots\dots \\ \dots\dots \\ X_{I,\overline{j_2+d_I^2}} \end{pmatrix},$$

quand $U = (1, 1, ..., 1)$ et où $\overline{j_1 + d_2^1} \equiv j_1 + d_2^1$ modulo N. Ici on utilise des décalages plutôt que des permutations car cela ne change rien.

D'apres le théorème 6, pour que $X_{j_1}^{d^1}$ et $X_{j_2}^{d^2}$ soient dépendants, il faut que, pour tout i, $X_{i,j_1+d_i^1} = X_{i,j_1+d_i^2}$. Comme $d_1^1 = d_1^2 = 0$ pour la première ligne, $j_1 = j_2 = n$. Donc, X_{1,j_1} et X_{1,j_2} appartiennent à deux blocs X^{d^1} et X^{d^2} différents : $d^1 \neq d^2$.

Pour les autre lignes, il faut que $n + d_i^1 = n + d_i^2$, c'est à dire $d_i^1 = d_i^2$ pour i=2,3,...,I.

Comme tous les décalages sont égaux, $d^1 = d^2$, ce qui est une contradiction.

Donc, les couples sont toujours indépendants.

De la même façon, à 3 dimensions, on prend un vecteur aléatoire $(X_{j_1}^{d^1}, X_{j_2}^{d^2}, X_{j_3}^{d^3})$ où

$$X_{j_1}^{d^1} = U * \begin{pmatrix} X_{1,j_1} \\ X_{2,\overline{j_1+d_2^1}} \\ \dots\dots \\ \dots\dots \\ X_{I,\overline{j_1+d_I^1}} \end{pmatrix}, \; X_{j_2}^{d^2} = U * \begin{pmatrix} X_{1,j_2} \\ X_{2,\overline{j_2+d_2^2}} \\ \dots\dots \\ \dots\dots \\ X_{I,\overline{j_2+d_I^2}} \end{pmatrix}, \; X_{j_3}^{d^3} = U * \begin{pmatrix} X_{1,j_3} \\ X_{2,\overline{j_3+d_2^3}} \\ \dots\dots \\ \dots\dots \\ X_{I,\overline{j_3+d_I^3}} \end{pmatrix}.$$

D'apres le théorème 6 et le fait que les couples sont indépendants, pour que $X_{j_1}^{d^1}$, $X_{j_2}^{d^2}$ et $X_{j_3}^{d^3}$ soient dépendants, il faut que, pour tout i, $X_{i,\overline{j_1+d_i^1}} = X_{i,\overline{j_1+d_i^2}} = X_{i,\overline{j_3+d_i^3}}$. Comme $d_1^1 = d_1^2 = d_1^3 = 0$, pour la première ligne, il faut donc $j_1 = j_2 = j_3 = n$. Donc, X_{1,j_1}, X_{1,j_2} et X_{1,j_3} proviennent de trois blocs différents : $d^1 \neq d^2$, $d^1 \neq d^3$, $d^2 \neq d^3$.

Pour les autre lignes, il faut que $n + d_i^1 \equiv n + d_i^2 \equiv n + d_i^3$, c'est à dire $d_i^1 = d_i^2 = d_i^3$ pour i=2,3,...,I.

Comme tous les décalages sont égaux, les trois vecteurs $\{X_{i,n}^{d^1}\}$, $\{X_{i,n}^{d^2}\}$ et $\{X_{i,n}^{d^3}\}$ doivent être égaux : $d^1 = d^2 = d^3$, ce qui est donc impossible.

Donc, il n'y a pas de triplets dépendants.

On trouve donc que les variables aléatoires $X_j^{d_s}$ sont de loi uniforme et ont des couples et des triplets indépendants. Ces conditions doivent, bien sûr, être vérifiées par des suites IID.

Malheureusement, cela ne continue pas et nous verrons en section 8 qu'il existe des quadruplets dépendants. Mais ce n'est pas génant car il y en a peu comme nous allons le voir maintenant.

3 Probabilité d'avoir une dépendance

On va maintenant étudier quelle est la probabilité d'avoir une p-dépendance complète lorsqu'on choisit un p-uplet au hasard. D'abord le nombre de p-uplet possibles est donné par la proposition suivante.

Proposition 3.1 *Soient* $\left\{X_{i,j_s}^{Pe^s}\right\}_{i=1,2,..,I}$ *s=1,2,....,p, p vecteurs colonnes tous différents. Alors, il y a au total* $N^I(N^I-1)....(N^I-(p-1))$ *tels vecteurs colonnes possibles.*

Démonstration Il y a N^I colonnes possibles. Il y a donc N^I façons de choisir la première colonne. Comme la deuxième colonne ne peut être égale a la première, il y a N^I-1 façons de choisir la deuxième colonne. Et ainsi de suite. ■

.

Remarque 3.1 *On considère que le quadruplet* (Z_1,Z_2,Z_3,Z_4) *est différent de* (Z_2,Z_1,Z_3,Z_4) *par exemple. On considère donc des arrangements.*

C'est normal : on étudie ce qui se passe quand on prend des p-uplets au hasard. Or, si on prend un p-uplet au hasard, on peut tomber sur l'un ou l'autre arrangement.

3.1 Cas p=4

Pour mieux comprendre la démonstration du cas général, on va d'abord étudier le cas des quadruplets. Le nombre maximum de quadruplets dépendants est donné par la proposition suivante.

Proposition 3.2 *Il y a au plus* $(3N^2)^I$ *quadruplets dépendants.*

Démonstration : Il ne peut s'agir que de quadruplets ayant une dépendance complète car les couples et les triplets sont indépendants. Soit $\left\{X_{i,j_s}^{Pe^s}\right\}_{i=1,2,..,I}$ s=1,2,3,4, un tel quadruplet dépendant.

D'apres le théorème 6 , pour chaque $X_{i,j_s}^{Pe^s}$, il existe $X_{i,j_r}^{Pe^r}$ tel que $s\neq r$ et que $X_{i,j_s}^{Pe^s}=X_{i,j_r}^{Pe^r}$. Donc, pour chaque ligne i, on aura des termes égaux deux par deux : si quatre sont égaux, c'est un cas particulier de deux couples égaux.

Il faut donc que le i-ème terme de la première colonne et de la deuxième soient égaux, ou bien le i-ème terme de la première colonne et de la troisième, ou bien le i-ème terme de la première colonne et de la quatrième.

Or, il y a N façons de choisir le i-ème terme de la première colonne et de la deuxième. A ce moment, il y a encore au plus N façons de choisir le i-ème terme de la troisième et quatrième colonnes. Donc au total il y a au plus N^2 façons de choisir les termes des 4 colonnes dans ce cas.

Il y a encore N façons de choisir le i-ème terme de la première colonne et de la troisième. Donc au total il y a encore au plus N^2 façons de choisir les termes des 4 colonnes dans ce cas.

Il y a N façons de choisir le i-ème terme de la première colonne et de la quatrième. Donc au total il y a toujours au plus N^2 façons de choisir les termes des 4 colonnes dans ce cas.

Donc au total il y a au plus $3N^2$ possibilités de choix pour chaque ligne i.

Donc au total il y a au plus $(3N^2)^I$ quadruplets dépendants. ■

On en déduit une majoration de la probabilité de trouver un quadruplet dependant au hasard.

Corollaire 3.3 *Si on choisit une quadruplet au hasard, il y a une probabilité de*

$$\frac{3^I N^{2I}}{N^I(N^I-1)(N^I-2)(N^I-3)}$$

au plus de trouver un quadruplet dépendant.

Remarque 3.2 *Dans cette démonstration, c'est toujours les arrangements des $X_{j_s}^{Pe^s}$ que l'on considère. En effet, simplifions en considérant des lignes de 4 termes chacunes. Par exemple, écrivons les deux premières lignes $X_{1,j_1}, X_{1,j_2}, X_{1,j_1}, X_{1,j_2}$ et $X_{2,j_1}, X_{2,j_1}, X_{2,j_2}, X_{2,j_2}$.*

Supposons que l'on choisisse un arrangement $X_{1,j_1}, X_{1,j_2}, X_{1,j_1}, X_{1,j_2}$ pour la première ligne. Comme chaque X_{i,j_s}, s=1,2, peut représenter chacune des N variables aléatoires $X_{i,1}, X_{i,2}, ..., X_{i,N}$, on a bien aussi comme suite possible $X_{1,j_2}, X_{1,j_1}, X_{1,j_2}, X_{1,j_1}$. On a des résultats semblables pour les autres lignes. Donc on considère bien tous les arrangements possibles dans le nombre de quadruplets dépendants que nous obtenons.

3.2 Cas général

On étudie les p-uplets qui sont complètement dépendants. D'apres le théorème 6, pour qu'un vecteur colonne soit dépendant des autres, il faut que pour chaque $X_{i,j_s}^{Pe^s}$ de ce vecteur colonne, il existe $X_{i,j_r}^{Pe^r}$ tel que $s \neq r$ vérifiant $X_{i,j_s}^{Pe^s} = X_{i,j_r}^{Pe^r}$.

Donc, pour chaque ligne i, les termes sont égaux par groupe de 2, ou de 3, ou de 4, ou de 5, etc.

De façon évidente, les groupes de 2, 4, 6,... peuvent se ramener à l'étude des groupes de 2 : 1 groupe de 4 = 2 groupes de 2. Les groupes de 3, 5, 7,... peuvent se ramener à l'étude des groupes de 3 auquels on associe des groupes de 2.

Si p=3p", $p" \in \mathbb{N}^*$, il y a au plus p" groupes de 3 possibles.

Si p=3p"+1, il y a au plus (p"-1) groupes de 3 possibles et, dans ce cas, deux groupes de 2 : ce n'est pas possible qu'il y aie p" groupes de 3, il resterait un élément solitaire et donc l'indépendance d'un élément.

Si p=3p"+2, il y a au plus p" groupes de 3.

En plus il faut tenir compte des cas où p" est pair ou impair (cf ci-apres).

Ces différents cas donnent à peu près le même type de résultats. On va donc se contenter d'étudier ici le cas où p=6p', $p' \in \mathbb{N}^*$.

Proposition 3.4 *Soient $\left\{X_{i,j_s}^{Pe^s}\right\}_{i=1,2,..,I}$ s=1,2,..,p, p vecteurs colonnes distincts. On suppose p=6p' pair. Alors, il y a au plus*

$$\Phi(p) = \left[\frac{(6p')!N^{3p'}}{2^{3p'}} \sum_{q'=0}^{p'} \frac{2^{q'}}{3^{2q'}} \frac{1}{[(2q')!](3p'-3q')!N^{q'}}\right]^I$$

p-uplets complètement dépendants.

Démonstration On peut choisir q groupes de 3 pour q=0,2,4,...,p" ou p"=2p' : q=2q' est pair. En effet, si p-3q = 3(p"-q) est impair - i.e. si q est impair - il restera un nombre impair de $X_{i,j_s}^{Pe^s}$ pour chaque ligne i. On ne peut pas les grouper par paquet de deux. C'est donc contraire à la définition.

Donc p-3q=3(p"-q) doit être pair, c'est à dire q=2q' pair. On considère donc p-3q, pour q=0,2,4,...,2p', i.e. p-3q = 3(2(p'-q')). Il reste alors 3(p'-q') couples.

Maintenant, pour chaque q, il y a C_p^{3q} façons de prendre 3q nombres parmis les p. On choisit donc un tel groupe de 3q termes.

Groupe de 3 Il faut alors savoir combien il y a de façons de choisir q groupes de 3 termes dans l'ensemble de 3q termes $X_{i,j_{s_t}}^{Pe^{s_t}}$, $t = 1, .., 3q$, que l'on vient de choisir.

On choisit donc un premier groupe de 3 termes dépendants. On peut choisir dans ce premier groupe le terme $X_{i,j_{s_1}}^{Pe^{s_1}}$ car il appartient forcément à un des groupes.

Il faut donc choisir des triplets dépendants contenant ce premier terme $X_{i,j_{s_1}}^{Pe^{s_1}}$ dans le groupe $(X_{i,j_{s_2}}^{Pe^{s_2}}, X_{i,j_{s_3}}^{Pe^{s_3}}, ..., X_{i,j_{s_{3q}}}^{Pe^{s_{3q}}})$. Il faut donc choisir deux éléments parmis les 3q-1 restants. Il y a au plus C_{3q-1}^2 tels deux éléments. Il y a donc C_{3q-1}^2 tels triplets contenant $X_{i,j_{s_1}}^{Pe^{s_1}}$.

On choisit ensuite un deuxième groupe de 3 termes dépendants parmis les 3q-3 termes restants (c'est à dire auquel on a retiré les 3 termes du premier goupe). On a donc fait le premier pas d'une récurence. On choisit donc d'abord comme premier terme, le premier terme de la suite restante $X_{i,j_{s_t}}^{Pe^{s_t}}$, i.e., celui où t est minimal dans cette suite. Ce terme appartient donc au groupe de 3q-3 termes restants.

Ce groupe a 3q–3 termes auquel on retire le premier : il reste 3q-4 termes. Il faut donc choisir deux éléments parmis au plus 3q-4.

Il y a donc au plus C_{3q-4}^2 tel triplets.

etc jusqu'au q-ème groupe ;

On a donc au plus,

$$C_{3q-1}^2 C_{3q-4}^2 .....C_{3q-3q+5}^2 C_{3q-3q+2}^2$$

$$= \frac{(3q-1)!}{2!(3q-1-2)!}\ \frac{(3q-4)!}{2!(3q-4-2)!} \cdots\cdots\cdots\cdots \frac{(3q-3q+5)!}{2!(3q-3q+5-2)!}\ \frac{(3q-3q+2)!}{2!(3q-3q+2-2)!}$$

$$= \frac{\big[(3q-1)(3q-2)\big]\big[(3q-4)(3q-5)\big].....\big[(3q-3q+5)(3q-3q+4)\big]\big[(3q-3q+2)(3q-3q+1)\big]}{2^q}$$

$$= \frac{\big[(3q-1)(3q-2)\big]\big[(3q-4)(3q-5)\big].....\big[5.4\big]\big[(2.1\big]}{2^q}$$

tels groupes de 3 possibles.

Maintenant, pour chacune de ces combinaisons possibles, il y a N^q choix possibles parmis les variables aléatoires égales à $X_{i,j}$, j=1,..,N. Donc au total, il y a au plus

$$\frac{\big[(3q-1)(3q-2)\big]\big[(3q-4)(3q-5)\big].....\big[5.4\big]\big[(2.1\big]N^q}{2^q}$$

$$= \frac{\big[(3q-1)(3q-2)(3q-3)\big]\big[(3q-4)(3q-5)(3q-6)\big].....3.2.1.N^q}{(3q-3)(3q-6).....6.3*2^q}$$

$$= \frac{(3q-1)!.N^q}{3^{q-1}(q-1)(q-2).....2.1*2^q} = \frac{(3q-1)!N^q}{3^{q-1}2^q.(q-1)!}$$

tels ensembles de triplets dépendants

Maintenant, il faut choisir q pair : q= 2q'. Donc au total, il y a au plus

$$\frac{(6q'-1)!N^{2q'}}{3^{2q'-1}2^{2q'}.(2q'-1)!}$$

tels ensembles de triplets dépendants.

En particulier, si $p = 3q = 6q'$, il y a au plus

$$\frac{[(p-1)!]N^{p/3}}{2^{2p'}3^{2p'-1}(2p'-1)!}$$

tel ensembles de triplets dépendants.

Couples Maintenant, il reste $q" = p - 3q = 6p' - 6q'$ termes dépendants. Il reste donc $3(p' - q')$ couples dépendants.

Il faut savoir combien il y a de façons de choisir des groupes de 2 termes dans l'ensemble de q" termes restants $(X_{i,j_{t_1}}^{Pe^{t_1}}, X_{i,j_{t_2}}^{Pe^{t_2}}, ..., X_{i,j_{t_{q"}}}^{Pe^{t_{q"}}})$.

Si p'-q'=0, il n'y a rien à chercher. Supposons donc $p' > q'$.

On choisit donc d'abord deux termes dont le premier est $X_{i,j_{t_1}}^{Pe^{t_1}}$. Il y a au plus p-3q-1 tels couples.

On prend ensuite le premier terme restant comme celui ayant le plus petit r dans l'ensemble $X_{i,j_{t_r}}^{Pe^{t_r}}$ restant. On choisit un terme associé à ce terme pour composer un second couple. Il y a p-3q-3 possibilités, et ainsi de suite.

Au total il y a au plus, $(p - 6q' - 1)(p - 6q' - 3)...3.1$ tels ensembles de couples possibles.

Maintenant, il faut prendre de tels ensembles avec tous les $X_{i,j}$ possibles. Donc, au total il y a au plus, pour $q' < p'$,

$$(p - 6q' - 1)(p - 6q' - 3)...3.1 N^{3p'-3q'}$$

$$= \frac{(p - 6q' - 1)(p - 6q' - 2)(p - 6q' - 3)...3.2.1.N^{3p'-3q'}}{(p - 6q' - 2)(p - 6q' - 4)...4.2}$$

$$= \frac{(6p' - 6q')(p - 6q' - 1)(p - 6q' - 2)(p - 6q' - 3)...3.2.1.N^{3p'-3q'}}{(p - 6q')(p - 6q' - 2)(p - 6q' - 4)...4.2}$$

$$= \frac{(6p' - 6q')!N^{3p'-3q'}}{2^{3p'-3q'}(3p' - 3q')!}$$

tels ensembles de couples de $X_{i,j}$ dépendants possibles.

Sommes Comme

$$C_p^{3q} = \frac{p!}{[(3q)!][(p - 3q)!]} = \frac{(6p')!}{[(6q')!][(6p' - 6q')!]} ,$$

au total, il y a au plus,

$$\sum_{q'=0}^{p'} \frac{(6p')!}{[(6q')!][(6p' - 6q')!]} \frac{(6q' - 1)!N^{2q'}}{3^{2q'-1}2^{2q'}.(2q' - 1)!} \frac{(6p' - 6q')!N^{3p'-3q'}}{2^{3p'-3q'}(3p' - 3q')!}$$

$$= \sum_{q'=0}^{p'} \frac{(6p')!}{(6q')} \frac{N^{2q'}}{3^{2q'-1}2^{2q'}.(2q' - 1)!} \frac{N^{3p'-3q'}}{2^{3p'-3q'}(3p' - 3q')!}$$

$$= \frac{(6p')!N^{3p'}}{2^{3p'}} \sum_{q'=0}^{p'} \frac{2^{q'}}{3^{2q'-1}(6q')} \frac{1}{[(2q' - 1)!](3p' - 3q')!N^{q'}}$$

$$= \frac{(6p')!N^{3p'}}{2^{3p'}} \sum_{q'=0}^{p'} \frac{2^{q'}}{3^{2q'}} \frac{1}{[(2q')!](3p' - 3q')!N^{q'}}$$

tels ensembles de variables dépendantes possibles pour chaque ligne i. Donc, pour toutes les lignes, il y a au plus

$$\left[\frac{(6p')!N^{3p'}}{2^{3p'}} \sum_{q'=0}^{p'} \frac{2^{q'}}{3^{2q'}} \frac{1}{[(2q')!](3p' - 3q')!N^{q'}}\right]^I$$

p-uplets dépendants possibles. ■

Corollaire 3.5 *Si p=6p' et si on choisit un p-uplet au hasard, il y a au plus une probabilité de* $\frac{\Phi(p)}{N^I(N^I-1)(N^I-2).....(N^I-(p-1))}$ *de trouver ainsi un p-uplet dépendant.*

Par exemple si p=6, $N = 10^9$, I=20;

$$\Phi(p) = \left[\frac{(6p')!N^{3p'}}{2^{3p'}} \sum_{q'=0}^{p'} \frac{2^{q'}}{3^{2q'}} \frac{1}{[(2q')!](3p'-3q')!N^{q'}}\right]^I$$

$$= \left[\frac{(6)!N^3}{2^3} \sum_{q'=0}^{1} \frac{2^{q'}}{3^{2q'}} \frac{1}{[(2q')!](3-3q')!N^{q'}}\right]^I$$

$$= \frac{(720)^I N^{3I}}{2^{3I}} \left[\frac{1}{3!} + \frac{2}{9} \frac{1}{[2!]N}\right]^I$$

$$= \frac{(720)^{20} 10^{9*60}}{2^{60}} \left[\frac{1}{6} + \frac{2}{9} \frac{1}{2*10^9}\right]^{20} .$$

Donc il y a une probabilité de tomber au hasard sur un 6-uplet dépendant qui est majorée par environ

$$\frac{(720)^{20}}{6^{20}2^{60} * 10^{9*60}} < \frac{1}{10^{516}} .$$

3.3 Cas p=N

Dans l'exemple précédent, on a vu qu'il y a très peu de chances de trouver au hasard une telle dépendance lorsque p=6. On va voir que c'est vrai aussi lorsque p=N.

Proposition 3.6 *On suppose p=N=6p' grand. Il y a au plus environ*

$$\left[\frac{(6p')!N^{3p'}}{2^{3p'}}\right]^I \left[\frac{(p'+1)2^{0.2956*p'}}{3^{0.5912.p'}} \frac{1}{[(0.5912.p')!](3p'-0.8868p')!N^{0.2956*p'}})\right]^I$$

p-uplets complètement dépendants.

Démonstration Posons q'=h. Alors,

$$\frac{2^{q'}}{3^{2q'}} \frac{1}{[(2q')!](3p'-3q')!N^{q'}} = \frac{2^h}{3^{2h}} \frac{1}{[(2h)!](3p'-3h)!N^h} = \psi(h) .$$

Donc,

$$\psi(h+1)/\psi(h) = \left[\frac{2^{h+1}}{3^{2(h+1)}} \frac{1}{[(2h+2)!](3p'-3h-3)!N^{h+1}}\right] \Big/ \left[\frac{2^h}{3^{2h}} \frac{1}{[(2h)!](3p'-3h)!N^h}\right]$$

$$= \left[\frac{2^{h+1}}{3^{2(h+1)}} \frac{1}{[(2h+2)!](3p'-3h-3)!N^{h+1}}\right] \left[\frac{3^{2h}}{2^h} \frac{[(2h)!](3p'-3h)!N^h}{1}\right]$$

$$= \left[\frac{2}{3^2}\ \frac{1}{[(2h+2)!](3p'-3h-3)!N}\right]\left[\frac{[(2h)!](3p'-3h)!}{1}\right]$$
$$= \frac{2}{3^2N}\ \frac{(3p'-3h)(3p'-3h-1)(3p'-3h-2)}{(2h+2)(2h+1)}\ .$$

On suppose donc N=p= 6p'. On cherche maintenant

$$\frac{2}{3^2N}\ \frac{(3p'-3h)(3p'-3h-1)(3p'-3h-2)}{(2h+2)(2h+1)} = 1\ .$$

On pose c=p', h=ac On cherche donc

$$\frac{2}{3^26c}\ \frac{(3c-3h)(3c-3h-1)(3c-3h-2)}{(2h+2)(2h+1)} = 1\ .$$

Donc

$$(3c-3h)(3c-3h-1)(3c-3h-2) = \frac{3^26c}{2}\big[(2h+2)(2h+1)\big]\ .$$

Donc,

$$(3c-3ac)(3c-3ac-1)(3c-3ac-2) = 3^3c\big[(2ac+2)(2ac+1)\big]\ .$$

Donc,

$$3^3c^3(1-a)(1-a-1/[3c])(1-a-2/[3c]) = 3^3c*4(ac)^2(1+1/[ac])(1+1/[2ac])\ .$$

Posons A=1-a. Alors,

$$(A)(A-1/[3c])(A-2/[3c]) = 4a^2(1+1/[ac])(1+1/[2ac])\ .$$

Donc,

$$A^3-(1/[3c]+2/[3c])A^2+2A/[9c^2] = 4a^2[1+(1/[ac]+1/[2ac])+1/(2a^2c^2)]\ .$$

Donc,

$$A^3-A^2/c+2A/[9c^2] = 4a^2[1+3/[2ac]+1/(2a^2c^2)]\ .$$

Donc,

$$(1-a)^3-(1-a)^2/c+2(1-a)/[9c^2] = 4a^2+6a/c+2/c^2\ .$$

Donc,

$$1-3a+3a^2-a^3-1/c-a^2/c+2a/c+2/[9c^2]-2a/[9c^2] = 4a^2+6a/c+2/c^2\ .$$

Donc,

$$(1-1/c+2/[9c^2])+(-3+2/c-2/[9c^2])a+[3-1/c]a^2-a^3 = 4a^2+6a/c+2/c^2\ .$$

Donc,

$$(1-1/c-16/[9c^2])+(-3-4/c-2/[9c^2])a+[-1-1/c]a^2-a^3 = 0\ .$$

Donc $1-3a-a^2-a^3 \approx 0$ si N est grand. Donc, on peut admettre

$$1-3a-a^2-a^3 = 0\ .$$

La seule racine réelle est $a \approx 0.295597742522085$.
La racine exacte si $c = 10^7$ est 0.295597683617988 et si $c = 10^6$, 0.295597153481165

Finalement à peu près

$$\frac{\psi(h+1)}{\psi(h)} = \frac{2}{3^2 6c} \frac{(3c-3h)(3c-3h-1)(3c-3h-2)}{(2h+2)(2h+1)} > 1$$

si $0 \leq h \leq 0.2956 * c$ et

$$\frac{\psi(h+1)}{\psi(h)} = \frac{2}{3^2 6c} \frac{(3c-3h)(3c-3h-1)(3c-3h-2)}{(2h+2)(2h+1)} < 1$$

si $n' \geq h \geq 0.2956 * c$.

Donc le maximum de $\psi(h)$ est atteint à peu près pour $h = h_0 = \lfloor 0.2956 * p' \rfloor$.

Il vaut donc

$$\psi(h_0) = \frac{2^{h_0}}{3^{2h_0}} \frac{1}{[(2h_0)!](3p'-3h_0)!N^{h_0}}$$

$$= \frac{2^{0.2956*p'}}{3^{2*0.2956*p'}} \frac{1}{[(2*0.2956*p')!](3p'-3*0.2956*p')!N^{0.2956*p'}})$$

$$= \frac{2^{0.2956*p'}}{3^{0.5912.p'}} \frac{1}{[(0.5912.p')!](3p'-0.8868p')!N^{0.2956*p'}}) \text{ . } \blacksquare$$

Proposition 3.7 *On suppose p=N=6p' grand et $p^2/N^I << 1$. Si on choisit un N-uplet au hasard il y a au plus une probabilité de*

$$\frac{N^{3I/2}[1+p^2/N^I][0.1060]^{p'I}[1+6/N]^I}{6^I.\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}}$$

de tomber au hasard sur un N-uplet dépendant.

Démonstration En supposant qu'aucun $X_{j_s}^{Pe_s} = X_{j_t}^{Pe_t}$, il y a $N^I(N^I-1)(N^I-2)....(N^I-(p-1))$ p-uplet dépendants $(X_{j_1}^{Pe_1}, ..., X_{j_p}^{Pe_p})$ possibles.

D'apres la formule de Stirling, $p! \approx \sqrt{2\pi p}\big(\frac{p}{e}\big)^p$ et $\Gamma(z) = \sqrt{\frac{2\pi}{z}}\Big(\frac{z}{e}\Big)^z\Big(1+O(1/z)\Big)$. Donc,

$$N^I(N^I-1)(N^I-2)....(N^I-(p-1)) = \frac{N^I!}{(N^I-p)!}$$

$$= \frac{\sqrt{2\pi N^I}\big(\frac{N^I}{e}\big)^{N^I}}{\sqrt{2\pi(N^I-p)}\big(\frac{N^I-p}{e}\big)^{N^I-p}} = \frac{\sqrt{2\pi N^I}(N^I)^p}{e^p\sqrt{2\pi(N^I-p)}}\Big(\frac{N^I}{N^I-p}\Big)^{N^I-p}$$

$$= \frac{N^{Ip}}{e^p\sqrt{(1-p/N^I)}}\Big(\frac{1}{1-p/N^I}\Big)^{N^I-p} \geq \frac{N^{Ip}}{e^p}\Big(\frac{1}{1-p/N^I}\Big)^{N^I-p} .$$

Or, si $q/h << 1$,

$$Log\Big(\Big(\frac{1}{1-q/h}\Big)^{h-q}\Big) = (q-h)Log(1-q/h) = (h-q)\Big[q/h + q^2/[2h^2] + q^3/[3h^3] + ...\Big]$$

$$= q + q^2/[2h] + q^3/[3h^2] + ... - q^2/h - q^3/[2h^2] - q^4/[3h^3] - .....$$

$$= q - q^2/[2h] - q^3/[6h^2] - ..... \ .$$

Donc,

$$\Big(\frac{1}{1-p/N^I}\Big)^{p-N^I}$$

$$= exp\big(-p + p^2/[2N^I] + p^3/[6N^{2I}] + ...\big)$$

$$= e^{-p}exp\big(p^2/[2N^I] + p^3/[6N^{2I}] + ...\big) \ .$$

$$= e^{-p}\Big[1 + \big(p^2/[2N^I] + p^3/[6N^{2I}] + ...\big) + \frac{\big(p^2/[2N^I] + p^3/[6N^{2I}] + ...\big)^2}{2!} + ...\Big] \ .$$

Donc, parce que $p^2/N^I << 1$, on peut supposer

$$\big(p^2/[2N^I] + p^3/[6N^{2I}] + ...\big) + \frac{\big(p^2/[2N^I] + p^3/[6N^{2I}] + ...\big)^2}{2!} + ... < p^2/N^I \ .$$

Dans ce cas,

$$\frac{1}{N^I(N^I-1)(N^I-2)....(N^I-(p-1))} \leq \frac{e^p}{N^{Ip}}\Big(\frac{1}{1-p/N^I}\Big)^{p-N^I}$$

$$\leq \frac{e^p}{N^{Ip}}\ e^{-p}\Big[1 + p^2/N^I\Big] = N^{-Ip}\big[1 + p^2/N^I\big] \ .$$

D'autre part, comme $p! \approx \sqrt{2\pi p}\big(\frac{p}{e}\big)^p$, il y au plus

$$\Big[\frac{(6p')!N^{3p'}}{2^{3p'}}\Big]^I\Big[\frac{(p'+1)2^{0.2956*p'}}{3^{0.5912.p'}}\ \frac{1}{[(0.5912.p')!](3p'-0.8868p')!N^{0.2956*p'}})\Big]^I$$

$$= \Big[\frac{\sqrt{2\pi N}\big(\frac{N}{e}\big)^{6p'}N^{3p'}}{2^{3p'}}\Big]^I\Big[\frac{(p'+1)2^{0.2956*p'}}{3^{0.5912.p'}}\ \frac{1}{[(0.5912.p')!](3p'-0.8868p')!N^{0.2956*p'}}\Big]^I$$

$$= \frac{\sqrt{(2\pi)^I N^I}\big(\frac{N}{e}\big)^{6p'I}N^{3p'I}}{2^{3p'I}}\ \frac{(p'+1)^I 2^{0.2956*p'I}}{3^{0.5912.p'I}N^{0.2956*p'I}}.....$$

$$.....\Big[\frac{1}{\sqrt{2\pi 0.5912.p'}\big(\frac{0.5912.p'}{e}\big)^{0.5912.p'}\sqrt{2\pi(3p'-0.8868p')}\big(\frac{3p'-0.8868p'}{e}\big)^{3p'-0.8868p'}}\Big]^I$$

$$= \frac{\sqrt{(2\pi)^I N^I} N^{9p'I}}{2^{3p'I-0.2956*p'I} e^{6p'I}} \frac{N^I/6^I[1+6/N]^I}{3^{0.5912.p'I} N^{0.2956*p'I}} \ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots \left[\frac{e^{3p'-0.8868p'+0.5912.p'}}{\sqrt{(2\pi)^2 0.5912.p'(3p'-0.8868p')}(0.5912.p')^{0.5912.p'}(3p'-0.8868p')^{3p'-0.8868p'}} \right]^I$$

$$= \frac{\sqrt{(2\pi)^I} N^{9p'I-0.2956*p'I+3I/2}}{2^{3p'I-0.2956*p'I+I} e^{6p'I}} \frac{[1+6/N]^I}{3^{0.5912.p'I+I}} \ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots \frac{e^{(3p'-0.8868p'+0.5912.p')I}}{\sqrt{(2\pi)^{2I}[0.5912.p'(3p'-0.8868p')]^I}(0.5912.p')^{0.5912.p'I}(3p'-0.8868p')^{(3p'-0.8868p')I}}$$

$$= \frac{N^{9p'I-0.2956*p'I+3I/2}[1+6/N]^I}{2^{3p'I-0.2956*p'I+I}.3^{0.5912.p'I+I}} \ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots \frac{e^{(-3p'-0.2956.p')I}.(1/0.5912)^{0.5912.p'I} 6^{0.5912.p'I}}{\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}(6p')^{0.5912.p'I}(2.1132*p')^{(3p'-0.8868p')I}}$$

$$= \frac{N^{9p'I-0.2956*p'I+3I/2-0.5912.p'I}[1+6/N]^I}{2^{3p'I-0.2956*p'I+I}.3^{0.5912.p'I+I}} \ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots \frac{e^{(-3p'-0.2956.p')I}.(1.6915)^{0.5912.p'I} 6^{0.5912.p'I}(6/2.1132)^{(3p'-0.8868p')I}}{\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}(6p')^{(3p'-0.8868p')I}}$$

$$= \frac{N^{9p'I-0.2956*p'I+3I/2-0.5912.p'I-(3p'-0.8868p')I}[1+6/N]^I}{2^{3p'I-0.2956*p'I+I}.3^{0.5912.p'I+I} e^{3p'I+0.2956.p')I}} \ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \frac{(1.6915)^{0.5912.p'I} 2^{0.5912.p'I} 3^{0.5912.p'I}(2.8393)^{(3p'-0.8868p')I}}{\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}}$$

$$= \frac{N^{6p'I+3I/2}[1+6/N]^I}{2^{3p'I-0.8868*p'I+I}.3^I.e^{3p'I+0.2956.p')I}} \frac{(1.6915)^{0.5912.p'I}(2.8393)^{(3p'-0.8868p')I}}{\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}}$$

$$= \frac{N^{6p'I+3I/2}[1+6/N]^I}{3^I.\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}} \frac{(1.6915)^{0.5912.p'I}(2.8393/e)^{3p'I}}{2^{3p'I-0.8868*p'I+I}.e^{0.2956.p'I}(2.8393)^{0.8868p'I}}$$

$$= \frac{N^{6p'I+3I/2}[1+6/N]^I}{3^I.\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}} \frac{(1.6915)^{0.5912.p'I}(1.1396)^{p'I}}{2^{2.1132*p'I+I}.e^{0.2956.p'I}(2.8393)^{0.8868p'I}}$$

$$= \frac{N^{3I/2}N^{6p'I}[1+6/N]^I}{6^I.\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}} \left[\frac{(1.6915)^{0.5912}(1.1396)}{2^{2.1132}.e^{0.2956}(2.8393)^{0.8868}}\right]^{p'I}$$

$$= \frac{N^{3I/2}N^{6p'I}[1+6/N]^I}{6^I.\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}} \left[0.1060\right]^{p'I}$$

N-uplets dépendants. Donc, la probabilité de trouver au hasard un N-uplet dépendant est

$$\frac{N^{3I/2}N^{6p'I} * N^{-Ip}[1+p^2/N^I][1+6/N]^I}{6^I.\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}} \left[0.1060\right]^{p'I}$$

$$= \frac{N^{3I/2}[1+p^2/N^I][0.1060]^{p'I}[1+6/N]^I}{6^I.\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}} \; . \blacksquare$$

Par exemple si $N = 6*10^6$, I=20, la probabilité de tomber au hasard sur un N-uplet dépendant est de l'ordre de

$$\frac{N^{3I/2}[1+p^2/N^I][0.1060]^{p'I}[1+6/N]^I}{6^I.\sqrt{(2\pi)^I[0.5912.p'(3p'-0.8868p')]^I}}$$
$$= \frac{6^{30}10^{30}[1+1/[6^{18}10^{6*18}]][0.1060]^{20*1.000.000}[1+1/10^6]^{20}}{6^{20}.(2\pi)^{10}[0.5912.*1000000(2.1132*1000000)]^{10}} = O(\frac{1}{10^{19493887}}) \; .$$

3.4 Dépendances non complètes

On voit donc que si p n'est pas trop grand, les dépendances qu'on a le plus de chances de trouver sont celles des quadruplets. Cependant, il se pose un problème : supposons que nous cherchions au hasard toutes les dépendances d'ordre 6 dans un 6-uplet. Il faudra trouver toutes les dépendances complètes d'ordre 4, toutes celles d'ordre 5 et celles d'ordre 6. Il y en aura $C_6^4 = 15$ quadruplets inclus dans le sextuplet et $C_6^5 = 6$ quintuplets.

Il faudra donc ajouter les probabilités des 15 dépendances d'ordre 4 et des 6 d'ordre 5 à celle d'ordre 6.

Il y a donc a priori plus de chances de trouver un p-uplet dépendant en choisissant un 6-uplet plutôt qu'un quadruplet. Mais du point de vue du cryptanalyste, cela ne sert à rien. En effet, il devra passer plus de temps pour tester d'abord tous les quadruplets possibles. Il est en effet bien évident que si on prend plusieurs quadruplets, on a plus de chance de tomber au hasard sur une dépendance que si on en prend qu'un. Mais cela ne change donc rien pour un cryptanalyste car il doit passer plus de temps dessus.

3.5 Conclusion

On voit donc que si on prend au hasard des p-uplets, on a une chance très faible de trouver une dépendance lorsque $p \leq N$ dès que N et I sont suffisament grands.

Cependant la question qui se pose est : pourra on prédire de manière plus sûre certaines parties de la suite x_n en choisissant mieux les dépendances? On verra en sections 8 et 9 que, dans certains cas, c'est possible si on choisit mieux les blocs. Mais, pour remédier à ce problème, il suffira encore une fois de bien choisir les paramêtres.

4 Etude des dépendances de grande taille

On va voir maintenant qu'on ne peut pas détecter les dépendances entre p-uplets si $p \geq N$ et $N \geq m$. Pour mieux se rendre compte de cela, on va d'abord supposer que q=0 (définition de q en section 1.1), i.e. on ne supprime aucun terme des blocs x_j^d : j=1,2,....,N. On verra qu'il sera difficile de détecter des dépendances dans ces cas-là. On verra aussi que ce sera vrai a fortiori si on ne prend que $\lfloor N/2 \rfloor$ termes à chaque bloc.

4.1 Etude d'introduction

4.1.1 Equation linéaire

On verra en sections 5, 6 et 7 que, si il y a dépendance, celle-ci se traduira par l'écriture d'équations du type $X_t = \overline{\sum_{s=1}^{Q} \delta_{t,s} U_s}$ où les U_s sont IID et où $\delta_{t,s} = 0$ ou 1. Cela signifie que toute dépendance se traduira sous forme d'équations linéaires modulo m.

Ce résultat est normal : supposons qu'on aie plus de X_t que de U_s, dans la plupart des cas, on a les équation réciproques : $U_s \equiv A_s(X_1, ..., X_p)$ pour s=1,...,Q. Donc on va trouver des relations linéaires $X_t \equiv \sum_{s=1}^{Q} \delta_{t,s} A_s(X_1, ..., X_p)$ pour t=Q+1,Q+2,...,p.

Maintenant, s'il n'y a pas plus de X_t que de U_s il est très possible qu'il n'y aie pas de dépendance linéaire.

4.1.2 Dépendance élémentaire

En fait, il y aura plus de chances qu'il y aie dépendance si il y a plus de 2N "X_n" (deux fois le nombre de $X_{1,j}$) : comme cela on a plus de chances de trouver dans les vecteurs colonnes au moins deux $X_{i,j_s}^{d^s}$ et $X_{i,j_t}^{d^t}$ égaux (cf théorème 6).

On va étudier d'abord la dépendance la plus simple qui puisse apparaitre, celle qui existe lorsque q=0 et qu'il y a 2N "$X_{j_s}^{d^{t_s}}$" qui se suivent.

Pour la voir, on prend les suites $X_n^{d^{t_1}}$, n=1,2,....,N, et $X_n^{d^{t_2}}$, n=1,2,....,N, correspondant à deux blocs. Il est donc clair que, pour tout i, pour chaque $X_{i,j}^{d^{t_1}}$, il existe $X_{i,j'}^{d^{t_2}}$: $X_{i,j}^{d^{t_1}} = X_{i,j'}^{d^{t_2}}$. Donc il peut y avoir dépendance.

En fait, il y a effectivement dépendance : $X_1^{d^{t_1}} + X_2^{d^{t_1}} + ... + X_N^{d^{t_1}} \equiv X_1^{d^{t_2}} + X_2^{d^{t_2}} + ... + X_N^{d^{t_2}}$ modulo m car $X_1^{d^{t_1}} + X_2^{d^{t_1}} + ... + X_N^{d^{t_1}} \equiv \sum_i \sum_{j=1}^{N} X_{i,j}^{d^{t_1}} \equiv \sum_i \sum_{j=1}^{N} X_{i,j}^{d^{t_2}} \equiv X_1^{d^{t_2}} + X_2^{d^{t_2}} + ... + X_N^{d^{t_2}}$ modulo m.

Pour éliminer cette dépendance trop élémentaire, on supprimera au moins un terme, le dernier $X_N^{d^1}$: i.e. on ne considère que les suites $X_1^{d^{t_1}}, X_2^{d^{t_1}}, ..., X_{N-1}^{d^{t_1}}, X_1^{d^{t_2}}, X_2^{d^{t_2}}, ..., X_{N-1}^{d^{t_2}}$. Donc $q \geq 1$.

Plus généralement on verra en section 9 qu'il est mieux de supposer que l'on en élimine $q = \lfloor N/2 \rfloor + 1$.

4.1.3 Exemple de dépendance

La dépendance précédente fournit un bon exemple des dépendances qui peuvent se produire. Par exemple, pour 3 blocs qui se suivent, les dépendances s'écrivent sous la forme de **deux** équations : $X_1^{d^{t_1}} + X_2^{d^{t_1}} + ... + X_N^{d^{t_1}} \equiv X_1^{d^{t_2}} + X_2^{d^{t_2}} + ... + X_N^{d^{t_2}} \equiv X_1^{d^{t_3}} + X_2^{d^{t_3}} + ... + X_N^{d^{t_3}}$ modulo m

Plus généralement, si il y a une dépendance entre $X_{j_1}^{d^1}$, ..., $X_{j_p}^{d^p}$, on sait qu'elle se traduira par une ou **plusieurs** équations linéaires. Par exemple soient 4 blocs tels qu'il y aie une dépendance entre quatre termes $X_1^{d^1}$,$X_1^{d^2}$, $X_{55}^{d^3}$, $X_{55}^{d^4}$: $X_1^{d^1} - X_1^{d^2} - X_{55}^{d^3} + X_{55}^{d^4} \equiv 0$.

Dans ce cas, si q=0, on a les équations linéaires. $X_{1+n}^{d^1} - X_{1+n}^{d^2} - X_{55+n}^{d^3} + X_{55+n}^{d^4} \equiv 0$ pour $n = 0, 1, ..., n_1$ ou $n_1 < N$ et $X_1^{d^1} + X_2^{d^1} + ... + X_N^{d^1} = X_1^{d^2} + X_2^{d^2} + ... + X_N^{d^2} = X_1^{d^3} + X_2^{d^3} + ... + X_N^{d^3} =$

$X_1^{d^4} + X_2^{d^4} + ... + X_N^{d^4}$.

Donc les dépendances peuvent se traduire par plusieurs équations linéaires modulo m.

4.1.4 Relations entre les dépendances

Certaines relations montrent que la dépendance complète entre p variables n'implique pas forcement la dépendance linéaire entre p+p' variables les contenant.

Supposons par exemple m=2 et que $X_1 + ... + X_{p-1} \equiv 0$. Alors, $X_1 + ... + X_{p-1} + X_p \equiv X_p$. Il n'y a donc pas moyen de prévoir X_p connaissant $(X_1, .., X_{p-1})$. En effet, il ne peut y avoir d'autre relation linéaire : par exemple, si $X_2 + ... + X_p \equiv 0$, alors $X_1 + ... + X_p \equiv X_p \equiv X_1$. Comme X_1 et X_p sont généralement indépendants, c'est généralement impossible.

Remarquons que les dépendances de ce type ne sont sûrement obtenues que lorsque m=2. Sinon, a priori, on peut avoir encore des relations linéaires avec X_p.

Par exemple supposons $X_1 - 2X_2 + 2X_3 - X_4 + X_5 - 2X_6 \equiv 0$. Cela ne veut dire pas dire qu'il n'est pas possible que $X_2 + X_3 + X_4 + X_5 - 2X_6 - X_7 \equiv 0$. Il suffit de prendre X_2, X_3, X_4, X_5, X_6 IID et X_1 et X_7 par ces relations.

4.1.5 Probabilité d'avoir une dépendance complète

Supposons que l'on choisisse un p-uplet au hasard $X_{j_s}^{d^s}$, s=1,2,...,p. Cela revient à peu près à avoir choisi au hasard p termes parmis N pour chaque ligne.

Or, s'il y a dépendance complète, il n'y a pas un $X_{i,j_s}^{d^s}$ tel que $X_{i,j_s}^{d^s} \neq X_{i,j_t}^{d^t}$ pour tout $t \neq s$. Donc, si on prend p assez grand, on a de grandes chances que cette condition soit vérifiée. On choisit donc $p = hN$, $h \in \mathbb{N}$, pour savoir à partir de quel h on a une chance raisonable d'avoir un dépendance complete.

Pour la ligne i, il s'agit donc de prendre au hasard un échantillon de p termes $X_{i,t_s}^{d^s}$, s=1,2,...,p, parmis N, ce qui revient à choisir au hasard p "t_s", $t_s \in \{1, 2, ..., N\}$, i.e. un p-échantillon de loi uniforme avec p=hN.

On considère alors les valeurs de $V_r = \sum_{s=1}^{p} 1_r(T_s)$ pour tout $r \in \{1, 2, ..., N\}$, où les T_s sont IID et de loi uniforme. Cela revient à savoir combien de $X_{i,r}^{d^s}$ (ou de $T_s = r$) on trouve quand on prend un p-uplet au hasard pour r fixe.

Maintenant on sait que la probabilité que $\sum_{s=1}^{p} 1_r(T_s) = k$ est $\frac{p!}{k!(p-k)!}(1/N)^k(1-1/N)^{p-k}$ (cf page 50 [1]).

Par exemple, si p=hN, elle vaut en k=0,

$$(1-1/N)^{hN} = e^{Log(1-1/N)(hN)} = e^{-(1/N+1/[2N^2]+..)(hN)} \approx e^{-h} .$$

Elle vaut en k=1,

$$hN(1/N)(1-1/N)^{hN-1} \approx h.e^{-h} .$$

Donc,

$$P\Big\{ \sum_{s=1}^{p} 1_r(T_s) = 1 \Big\} \approx h.e^{-h} .$$

Or, pour qu'il n'y a pas un $X_{i,j_s}^{d^s}$ tel que $X_{i,j_s}^{d^s} \neq X_{i,j_t}^{d^t}$ lorsque $t \neq s$, il faut donc que pour aucun r, $\sum_{s=1}^{p} 1_r(T_s) = 1$. Quelle est donc la probabilité qu'un tel évènement arrive lorsque l'on a choisi au hasard les $X_{i,j_s}^{d^s}$, s=1,..,hN ?

Le problème qui se pose pour calculer cette probabilité est que les $\sum_{s=1}^{p} 1_r(T_s)$ ne sont pas indépendants lorsque r=1,2,...,N. Mais le nombre de r tels que $\sum_{s=1}^{p} 1_r(T_s) = 1$ se comporte à peu près comme le résultat d'un échantillon indépendant. On peut le voir par simulation.

Maintenant, pour qu'il y aie dépendance complète, il ne faut donc pas qu'il y aie un r tel que $\sum_{s=1}^{p} 1_r(T_s) = 1$. Il faut donc que he^{-h} soit assez petit pour que $\sum_{s=1}^{p} 1_r(T_s) = 1$ pour r=1,2,....,N, aie peu de chances d'arriver. Maintenant, sur N tirages indépendants la probabilité qu'un évènement de probabilité p_r n'arrive pas est, d'après la loi binomiale, $(1-p_r)^N$. Si p_r est petit, $(1-p_r)^N \approx 1 - p_r N$. Donc, pour donner un ordre d'idée, si $p_r N \approx 1$, cet évènement (qu'il y aie un r tel que $\sum_{s=1}^{p} 1_r(T_s) = 1$) a une chance raisonable d'arriver. Ici $p_r = he^{-h}$. Donc si $N.he^{-h} \approx 1$, cet évènement a une chance raisonable d'arriver.

Maintenant, il faut que pour toutes les lignes i, $\sum_{s=1}^{p} 1_r(T_{i,s}) \neq 1$.

Or, $P\Big\{ \cap_i \{\sum_{s=1}^{p} 1_r(T_{i,s}) \neq 1\} \Big\} = [(1-p_r)^N]^I \approx 1$, i.e. $(1-p_r)^{NI} \approx 1$.

Par exemple si $NI.he^{-h} \approx 1$, cet évènement (qu'il y aie un r et un i tel que $\sum_{s=1}^{p} 1_r(T_{i,s}) = 1$) a une chance raisonable d'arriver.

Or $NI.he^{-h} \approx 1$ si $Log(NIh) = Log(N) + log(I) + Log(h) = h$. Par exemple si $N = 10^7$, I=20, $h > 7log(10) + Log(20) = 19.11$.

Ce qui veut dire que, pour avoir une chance raisonable de ne pas avoir de dépendance complète, il faut prendre $h \leq Log(N) + log(I)$.

Si il n'y a pas dépendance complète, il peut toujours y avoir des sous-dépendances. Mais il faudra les trouver : cela se ramenera à trouver les $\sum_s \delta_s X_s \equiv 0$ modulo 2 quand m=2 et $\delta_s = 0$ ou 1, ce qui sera impossible, comme nous allons le voir maintenant.

4.2 Equation linéaire de dépendance

Le problème quand on a une dépendance est de savoir si celle-ci va permettre à un cryptanalyste de prévoir un $x_{j_t}^{d^t}$ connaissant des $x_{j_s}^{d^s}$. Or les décalages changent à chaque bloc. Donc comme un bloc a au plus N termes, il faut pouvoir trouver les dépendances en utilisant moins de N p-uplets qui se suivent. On va donc voir que si $p \geq N$ et $m \leq N$, on ne peut pas détecter ces dépendances.

On verra en section 5, 6, et 7 que, si il y a dépendance et si m=2, celle-ci se traduira par l'écriture d'équations du type $X_t \equiv \sum_{s=1}^{Q} \delta_{t,s} U_s$ où les U_s sont IID pour s=1,2,...,p. Si m est quelconque, toute dépendance se traduira sous forme de M équations linéaires $\beta_1^r X_{j_1}^{d^1} + ... + \beta_p^r X_{j_p}^{d^p} \equiv \alpha_r$ où $\beta_s^r \in \{0, 1, 2, ..., \}$ pour r=1,2,...,M.

4.2.1 Détection de dépendance dans le cas de décalages

Il est bien clair lorsque l'on a des décalages que la dépendance de $(X_{j_1}^{d_1}, .., X_{j_p}^{d_p})$ est la même que celle de $(X_{j_1+1}^{d_1}, .., X_{j_p+1}^{d_p})$, $(X_{j_1+2}^{d_1}, .., X_{j_p+2}^{d_p})$, si tous les termes $X_{j_s+r}^{d_s}$, r=0,1,2, restent dans le même bloc leur correspondant. Il y aura donc au plus N dépendances successives identiques : après il n'y a plus dépendance ou on change de dépendance. Par exemple, si p=4, on peut trouver une q_0 tel que $X_{1+n}^{d_1} + X_{1+n}^{d_2} + X_{q_0+n}^{d_3} + X_{q_0+n}^{d_4} \equiv 0$ pour $n = 1, 2, ..., N_1$ ou $N_1 < N$: cf ci-après.

Par exemple, en simpliant $X_{1,j}$ en U_j , $X_{2,j}$ en Y_j, $X_{3,j}$ en Z_j et $X_{4,j}$ en T_j , on peut trouver les blocs suivants
$(U_1, U_2, \mathbf{U_3}, \mathrm{U_4}, \mathrm{U_5}, \mathrm{U_6}), (\mathrm{U_1}, \mathrm{U_2}, \mathbf{U_3}, \mathrm{U_4}, \mathrm{U_5}, \mathrm{U_6}), (\mathrm{U_1}, \mathrm{U_2}, \mathrm{U_3}, \mathrm{U_4}, \mathbf{U_5}, \mathrm{U_6}),\ (\mathrm{U_1}, \mathrm{U_2}, \mathrm{U_3}, \mathrm{U_4}, \mathbf{U_5}, \mathrm{U_6})$
$(Y_5, Y_6, \mathbf{Y_1}, \mathrm{Y_2}, \mathrm{Y_3}, \mathrm{Y_4}), .(\mathrm{Y_5}, \mathrm{Y_6}, \mathbf{Y_1}, \mathrm{Y_2}, \mathrm{Y_3}, \mathrm{Y_4}), .(\mathrm{Y_4}, \mathrm{Y_5}, \mathrm{Y_6}, \mathrm{Y_1}, \mathbf{Y_2}, \mathrm{Y_3}), .(\mathrm{Y_4}, \mathrm{Y_5}, \mathrm{Y_6}, \mathrm{Y_1}, \mathbf{Y_2}, \mathrm{Y_3})$
$(Z_6, Z_1, \mathbf{Z_2}, \mathrm{Z_3}, \mathrm{Z_4}, \mathrm{Z_5}).., ...(\mathrm{Z_5}, \mathrm{Z_6}, \mathbf{Z_1}, \mathrm{Z_2}, \mathrm{Z_3}, \mathrm{Z_4}).., ..(\mathrm{Z_4}, \mathrm{Z_5}, \mathrm{Z_6},\ \mathrm{Z_1}, \mathbf{Z_2},\ \mathrm{Z_3}).., .(\mathrm{Z_3}, \mathrm{Z_4}, \mathrm{Z_5}, \mathrm{Z_6}, \mathbf{Z_1}, \mathrm{Z_2},)$
$(T_4, T_5, \mathbf{T_6}, \mathrm{T_1}, \mathrm{T_2}, \mathrm{T_3},), .(\mathrm{T_1}, \mathrm{T_2}, \mathbf{T_3}, \mathrm{T_4}, \mathrm{T_5}, \mathrm{T_6}), .(\mathrm{T_5}, \mathrm{T_6}, \mathrm{T_1}, \mathrm{T_2}, \mathbf{T_3}, \mathrm{T_4},).., (\mathrm{T_2}, \mathrm{T_3}, \mathrm{T_4}, \mathrm{T_5}, \mathbf{T_6}, \mathrm{T_1})$

En sommant modulo m les lignes pour chaque colonne en gras, on obtient un quadruplet dependant.

Mais comme colonnes en gras, on peut aussi prendre les suivantes
$(U_1, U_2, U_3, U_4, U_5, \mathbf{U_6}), (\mathrm{U_1}, \mathrm{U_2}, \mathrm{U_3}, \mathrm{U_4}, \mathrm{U_5}, \mathbf{U_6}), (\mathrm{U_1}, \mathbf{U_2}, \mathrm{U_3}, \mathrm{U_4}, \mathrm{U_5}, \mathrm{U_6}), (\mathrm{U_1}, \mathbf{U_2}, \mathrm{U_3}, \mathrm{U_4}, \mathrm{U_5}, \mathrm{U_6})$
$(Y_5, Y_6, Y_1, Y_2, Y_3, \mathbf{Y_4}), .(\mathrm{Y_5}, \mathrm{Y_6}, \mathrm{Y_1}, \mathrm{Y_2}, \mathrm{Y_3}, \mathbf{Y_4}), .(\mathrm{Y_4}, \mathbf{Y_5}, \mathrm{Y_6}, \mathrm{Y_1}, \mathrm{Y_2}, \mathrm{Y_3}), .(\mathrm{Y_4}, \mathbf{Y_5}, \mathrm{Y_6}, \mathrm{Y_1}, \mathrm{Y_2}, \mathrm{Y_3})$
$(Z_6, Z_1 Z_2, Z_3, Z_4, \mathbf{Z_5}).., ...(\mathrm{Z_5}, \mathrm{Z_6}, \mathrm{Z_1}, \mathrm{Z_2}, \mathrm{Z_3}, \mathbf{Z_4}).., .(\mathrm{Z_4}, \mathbf{Z_5}, \mathrm{Z_6}, \mathrm{Z_1}, \mathrm{Z_2}, \mathrm{Z_3}).., ...(\mathrm{Z_3}, \mathbf{Z_4}, \mathrm{Z_5}, \mathrm{Z_6}, \mathrm{Z_1}, \mathrm{Z_2},)$
$(T_4, T_5, T_6, T_1, T_2, \mathbf{T_3},), .(\mathrm{T_1}, \mathrm{T_2}, \mathrm{T_3}, \mathrm{T_4}, \mathrm{T_5}, \mathbf{T_6}), .(\mathrm{T_5}, \mathbf{T_6}, \mathrm{T_1}, \mathrm{T_2}, \mathrm{T_3}, \mathrm{T_4},).., (\mathrm{T_2}, \mathbf{T_3}, \mathrm{T_4}, \mathrm{T_5}, \mathrm{T_6}, \mathrm{T_1})$

On voit donc que chaque fois que l'on décale un quadruplet de 1, on conserve pendant un certain temps le même type de dépendance et donc la même relation linéaire si elle existe. Par exemple, modulo 2, $X_1^{d^1} + X_1^{d^2} + X_3^{d^3} + X_3^{d^4} \equiv 0$, $X_2^{d^1} + X_2^{d^2} + X_4^{d^3} + X_4^{d^4} \equiv 0$, $X_3^{d^1} + X_3^{d^2} + X_5^{d^3} + X_5^{d^4} \equiv 0$, $X_4^{d^1} + X_4^{d^2} + X_6^{d^3} + X_6^{d^4} \equiv 0$.

Maintenant, on aura aussi $X_5^{d^1} + X_5^{d^2} + X_1^{d^3} + X_1^{d^4} \equiv 0$, $X_6^{d^1} + X_6^{d^2} + X_2^{d^3} + X_2^{d^4} \equiv 0$. Mais, ce ne sera plus le même type de relation linéaire.

En effet, supposons que les sommes modulo m des 4 matrices ci-dessus représentent la suite X_j, j=1,2,...,24. Alors, ces équations s'écrivent

$X_1 + X_7 + X_{15} + X_{21} \equiv 0$, $X_2 + X_8 + X_{16} + X_{22} \equiv 0$, $X_3 + X_9 + X_{17} + X_{23} \equiv 0$, $X_4 + X_{10} + X_{18} + X_{24} \equiv 0$, et $X_5 + X_{11} + X_{13} + X_{19} \equiv 0$, $X_6 + X_{12} + X_{14} + X_{20} \equiv 0$,

Ce n'est donc plus le même type de relation linéaire : il y en a donc de deux types différents : $X_{1+n}^{d^1} + X_{1+n}^{d^2} + X_{q_0+n}^{d^3} + X_{q_0+n}^{d^4} \equiv 0$ et $X_{1+n}^{d^1} + X_{1+n}^{d^2} + X_{q_1+n}^{d^3} + X_{q_1+n}^{d^4} \equiv 0$.

Donc, si on a trouvé la dépendance et si on connait les $x_{j_1+r}^{d^1}$, $x_{j_2+r}^{d^2}$ et $x_{j_3+r}^{d^3}$, on peut connaitre les $x_{j_4+r}^{d^4}$ pour $r = 0, 1, 2, ...N_1$ ou $N_1 \leq N$: on aura par exemple des relations du type $x_{1+n}^{d^1} + x_{1+n}^{d^2} + x_{q_0+n}^{d^3} + x_{q_0+n}^{d^4} \equiv 0$ pour tous les $n = 1, 2, ..., N_1$.

C'est général : si on connait la dépendance et si on connait $x_{j_t}^{d_t}$, t=1,2,...,p-1, on connait $x_{j_p}^{d_p}$. Donc on peut prévoir $x_{j_p}^{d_t}$ pendant un certain temps! En cryptographie cela pourrait poser un problème. Mais on a vu qu'il y a une chance infime de trouver une telle dépendance.

4.2.2 Recherche de la dépendance. Cas m=2

On se pose donc la question de savoir combien il y a de combinaisons possibles $\sum_{s=1}^{p} \delta_s x_{j_s}^{d^s} \equiv 0$ où $\delta_s = 0$ ou 1 quand m=2? De façon évidente, il y a 2^p combinaisons possibles $\sum_{s=1}^{p} \delta_s x_{j_s}^{d^s}$.

Alors combien y aura-t-il d'équations du type $\sum_{s=1}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$ modulo m=2, où $\alpha = 0$ ou 1 pour $r = 0, 1, ...n_1$, quand $n_1 < N$?

Si on choisit $p > N$, il y a $2^p > 2^N$ combinaisons possibles $\sum_{s=1}^{p} \delta_s x_{j_s}^{d^s}$. Donc, si au moins un $x_{j_s}^{d^s} \neq 0$, il y aura 2^{p-1} combinaisons possibles $\sum_{s=1}^{p} \delta_s x_{j_s}^{d^s} \equiv 0$ et autant pour $\sum_{s=1}^{p} \delta_s x_{j_s}^{d^s} \equiv 1$.

Remarquons que, de toutes façons, pour qu'il y aie une relation linéaire modulo m entre des variables aléatoires $X_{j_s}^{d^s}$, il faut que celle-ci soit de la forme $\sum_{s=1}^{p} \delta_s X_{j_s}^{d^s} \equiv 0$ étant donné que ce sera vrai lorsque $X_{j_s}^{d^s} = 0$ pour tout s. Le cas où tous les $x_{j_s}^{d^s} = 0$ n'est donc pas génant.

De même, il y aura donc $2^{p-1}/2 = 2^{p-2}$ combinaisons possibles $\sum_{s=1}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour r=0,1.

Pour prouver ce résultat, on doit pouvoir utiliser des théorèmes mathématiques. Mais la démonstration semble un peu longue. Aussi, le plus simple est de le vérifier par simulation.

Nous avons donc employé les programmes Matlab suivants

```
function y = NombComb(p,d,X)
X1 = X(1:p);
X2= X(2:p+1);
fork = 1 : d^p
.....H= DecompNomb2(k-1,p,d);
.....Y1(k) = X1*H';
.....Y1(k) = Y1(k) -d*fix(Y1(k)/d);
.....Y2(k) = X2*H';
.....Y2(k) = Y2(k) -d*fix(Y2(k)/d);
end
U = ones(d^p, 1);
for i=1:d
.....for j=1:d
.....I = Indicat(Y1,i-3/2,i-1/2).*Indicat(Y2,j-3/2,j-1/2);
.....N(i,j) = I*U;
.....end end
y=N;
```

où

function y = DecompNomb2(A,p,d)
for h=1:p
..... j=p-h;
.....$D(h) = fix(A/d^j)$;
.....$A = A - d^j * D(h)$;
end
y=D;

Ce sera la même chose pour trois suites : il y aura donc $2^{p-1}/2^2 = 2^{p-3}$ combinaisons possibles $\sum_{s=1}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour r=0,1,2.

Finalement, il y aura $2^p/2^R$ combinaisons possibles $\sum_{s=1}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour r=1,...,R si au moins un $x_s^{d_s} \neq 0$.

Maintenant, pour chaque décalage d^s, on doit se limiter au N-q premiers $X_s^{d^s}$ au maximum : après les décalages auront changé et la dépendance cessera. Donc, il y aura 2^{p-N+q} combinaisons possibles $\sum_{s=t}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$, r=1,...,N-q, si $p \geq N - q$ alors qu'il y aura peut être indépendance.

Donc, si $p > N$, et s'il y a une seule dépendance linéaire $\sum_{s=t}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$, on ne peut pas retrouver cette dépendance à partir de l'échantillon étant donné que l'on peut toujours trouver des suites δ_s telles que $\sum_{s=t}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour tout $r = 1, 2, ..., n_1$.

De toutes façons, si on veut être relativement sûr qu'il y a dépendance, on a vu qu'on doit prendre au moins 2N termes, c'est à dire $p \geq 2N$, et donc, il y aura il y aura toujours statistiquement $2^{p-N} \geq 2^N$ combinaisons possibles vérifiant $\sum_{s=1}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour $r = 1, 2, ..., n_1$. On ne peut donc pas trouver la dépendance s'il y a une seule dépendance.

Ce sera la même chose si $2N \geq p \geq N$, et donc, il y aura toujours 2^{p-N} combinaisons possibles vérifiant $\sum_{s=1}^{p} \delta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour $r = 1, 2, ..., n_1$.

En particulier, si on ne prend que les $\lfloor N/2 \rfloor$ premiers termes de chaque bloc, il sera tout à fait impossible de trouver la dépendance même si elle existe dès que $p \geq N$.

4.2.3 Cas m quelconque

Remarquons d'abord que une dépendance serait plus facile à trouver si on avait par exemple $\delta_1 x_{j_1}^{d^1} + ... + \delta_p x_{j_p}^{d^p} \equiv 0$ lorque m= 1125497 et p=4. Mais le nombre de combinaisons linéaires possibles augmente très rapidement si p grandit (i.e. les coefficients des $x_{j_s}^{d^s}$).

Le problème est donc de savoir combien d'équations linéaires exprimant une dépendance sont possibles. Théoriquement, d'après les resultats obtenus pour p=4,5,6,7,8, en sections 5, 6 et 7, on pourrait avoir $\sum_{s=1}^{p} \beta_s x_{j_s+r}^{d^s} \equiv \alpha$ ou $\beta_s \in \{0, 1, ..., \lfloor p/2 \rfloor\} \cup \{-1, -2, ..., -\lfloor p/2 \rfloor\}$. On voit ainsi que lorsque p=6, on trouve des équations du type $\sum_{s=1}^{p} \beta_s x_{j_s+r}^{d^s} \equiv 0$ ou $\beta_s = \pm 1$ ou ± 2.

Maintenant, combien y a t il de combinaisons possibles $\sum_{s=1}^{p} \beta_s x_{j_s}^{d^s}$ où $\beta_s \in \mathcal{E}$, quand $card(\mathcal{E}) = P_q$? Clairement, il y en a $(P_q)^p$.

Donc, si $\beta_s \in \mathcal{E}$, pour que $\sum_{s=1}^{p} \beta_s x_{j_s}^{d^s} \equiv \alpha$, il y a statistiquement $(P_q)^p/m$ combinaisons possibles vu qu'il y a m valeurs possibles pour α.

Pour que $\sum_{s=1}^{p} \beta_s x_{j_s+r}^{d^s} \equiv \alpha$, $r = 0, 1, ..., M$, il y a statistiquement $(P_q)^p/m^M$ combinaisons possibles.

D'autre part, d'après nos résultats des sections 5, 6 et 7, on déduit que $P_q \leq Min(p, m)$.

Supposons, $p \geq N$ et $P_q = m$. On a au plus M=N (après la dépendance change, et donc n'est plus détectable pour les $x_{j_s+r}^{d_s}$).

On a donc au plus $(P_q)^p/m^N \geq m^N/m^N \geq 1$ combinaisons possibles. Donc même s'il y a indépendance, on peut toujours trouver une relation linéaire verifiant $\sum_{s=1}^{p} \beta_s x_{j_s+r}^{d^s} \equiv \alpha$ pour $r = 0, 1, ..., M$.

Supposons, $p \geq N$ et $P_q = p$. On a toujours au plus M=N.

On a donc au plus $(P_q)^p/m^N \geq p^N/m^N$ combinaisons possibles. Donc si on veut que, même s'il y a indépendance, on puisse toujours trouver une relation linéaire, il faut imposer $N \geq m$.

4.2.4 Plusieurs équations linéaires

L'exemple donné en section 4.1.3 montre qu'il peut y avoir plusieurs équations linéaires.

Donc on peut avoir des relations du type $\sum_{s=1}^{p} \beta_s^t x_{j_s+r}^{d^s} \equiv \alpha^t$ pour $r = 1, 2, ..., n_1^t$, t=1,2,....,T. En effet, on peut toujours supposer que p reste fixe quite à supposer $\beta_s^t = 0$.

D'après ce qui précède, pour que $\sum_{s=1}^{p} \beta_s x_{j_s+r}^{d^s} \equiv \alpha$, $r = 0, 1, ..., M$, il y a statistiquement $(P_q)^p/m^M$ combinaisons possibles.

Donc, avec $p \geq N$, N=m, et q = $\lfloor N/2 \rfloor$, il y a $(P_q)^p/N^M \geq N^p/N^M \geq N^N/N^M$ combinaisons possibles. Comme $M \leq N/2$, il y a au moins $N^{N/2}$ combinaisons possibles.

Cela veut dire que N/2 relations linéaires pourront toujours être vérfiées. C'est bien suffisant

Il s'agit ici de relations linéaires de grande taille : $p \geq N$. On n'étudie pas dans cette section les relations de petites tailles qui sont en nombre infime comme nous l'avons déjà vu.

4.3 Conséquences

Donc, pour trouver à coup sûr des dépendances, il faut d'abord prendre des p uplets où $p \geq 2N$. Si $N \geq m$, il y a beaucoup de combinaisons linéaires possibles entre N termes qui se suivent même s'il y a indépendance. C'est à dire qu'il n'y a donc aucun moyen de trouver une dépendance de taille plus grande que N.

Donc, si on veut détecter des dépendances, il faut donc chercher des dépendances où p est petit en particulier celle qui a la plus grande probabilité d'arriver, celle des quadruplets.

De toutes façons, le résultat général est clair : il sera extrêmement difficile à un cryptanalyste de trouver des dépendances, de petite ou de grande taille.

5 Etude des dépendances de quadruplets

On s'intéresse maintenant au cas p=4. On va voir qu'il y a des cas différents où il y a dépendance. On considère donc un vecteur $(X_{j_1}^{d_1}, X_{j_2}^{d_2}, X_{j_3}^{d_3}, X_{j_4}^{d_4})$. D'après les théorèmes 6 et 1, on peut l'écrire sous la forme ($\overline{X+Y+U+W}$, $\overline{X+Y+V+T}$, $\overline{X+Z+U+T}$, $\overline{X+Z+V+W}$) où (X,Y,Z,U,V,W,T) est un vecteur aléatoire de loi uniforme, et donc indépendant : en effet, supposons par exemple que dans les colonnes 1 et 2, on aie $X_{2,j_1}^{d^1} = X_{2,j_2}^{d^2}$ et $X_{3,j_1}^{d^1} = X_{3,j_2}^{d^2}$. Alors, $\overline{X_{2,j_1}^{d^1} + X_{3,j_1}^{d^1}} = \overline{X_{2,j_2}^{d^2} + X_{3,j_2}^{d^2}}$ est de loi uniforme d'apres le théorème 1. Avec les notations que nous voulons introduire, on peut par exemple noter cette variable aléatoire $\overline{X_{2,j_1}^{d^1} + X_{3,j_1}^{d^1}}$ par Y.

Pour étudier la 4-dépendance, on peut utiliser le théorème suivant.

Proposition 5.1 *On suppose (X,Y,Z,U,V,W,T) est un vecteur aléatoire indépendant et de loi uniforme. Alors,*

$$P\Big\{\{\overline{X+Y+U+W}=a_1\}\cap\{\overline{X+Y+V+T}=a_2\}\cap\{\overline{X+Z+U+T}=a_3\}\cap\{\overline{X+Z+V+W}=a_4\}\Big\}$$

$= \frac{1}{m^4}$ *si m est impair,*

$= 0$ *si m est pair et* $\overline{a_2-a_1-a_4+a_3}$ *est impair,*
$= \frac{2}{m^4}$ *si m est pair et* $\overline{a_2-a_1-a_4+a_3}$ *est pair.*

Démonstration On a les égalités suivantes

$$P\Big\{\{\overline{X+Y+U+W}=a_1\}\cap\{\overline{X+Y+V+T}=a_2\}\cap\{\overline{X+Z+U+T}=a_3\}\cap\{\overline{X+Z+V+W}=a_4\}\Big\}$$

$$= \sum_{x,u,v,w,t} P\Big\{\{X=x\}\cap\{U=u\}\cap\{V=v\}\cap\{W=w\}\cap\{T=t\}.....$$

$$.....\cap\{\overline{X+Y+U+W}=a_1\}\cap\{\overline{X+Y+V+T}=a_2\}\cap\{\overline{X+Z+U+T}=a_3\}\cap\{\overline{X+Z+V+W}=a_4\}\Big\}$$

$$= \sum_{x,u,v,w,t} P\Big\{\{X=x\}\cap\{U=u\}\cap\{V=v\}\cap\{W=w\}\cap\{T=t\}.....$$

$$.....\cap\{Y=\overline{a_1-x-u-w}\}\cap\{Y=\overline{a_2-x-v-t}\}\cap\{Z=\overline{a_3-x-u-t}\}\cap\{Z=\overline{a_4-x-v-w}\}$$

$$= (1/m^5)\sum_{x,u,v,w,t} P\Big\{\{Y=\overline{a_1-x-u-w}\}\cap\{Y=\overline{a_2-x-v-t}\}\cap\{Z=\overline{a_3-x-u-t}\}\cap\{Z=\overline{a_4-x-v-w}\}$$

$$= (1/m^5)\sum_{x,u,v,w,t,\ \overline{a_1-x-u-w}=\overline{a_2-x-v-t}\ ,\ \overline{a_3-x-u-t}=\overline{a_4-x-v-w}} (1/m^2)$$

$$= (1/m^7)\sum_{x,u,v,w,t,\ \overline{a_1-u-w}=\overline{a_2-v-t}\ ,\ \overline{a_3-u-t}=\overline{a_4-v-w}} 1$$

$$= (1/m^7)\sum_{x,u,v,w,t,\ v=\overline{a_2+u+w-a_1-t}\ ,\ v=\overline{a_4+u+t-a_3-w}} 1$$

$$= (1/m^7)\sum_{x,u,w,t,\ \overline{a_2+u+w-a_1-t}=\overline{a_4+u+t-a_3-w}} 1$$

$$= (1/m^7)\sum_{x,u,w,t,\ \overline{a_2+w-a_1-t}=\overline{a_4+t-a_3-w}} 1$$

$$= (1/m^5) \sum_{w,t,\ \overline{a_2+w-a_1-t}=\overline{a_4+t-a_3-w}} 1$$

$$= (1/m^5) \sum_{w,t,\ \overline{a_2-a_1-a_4+a_3}=\overline{2t-2w}} 1 \ .$$

On utilise alors la propriété suivante : soit $\overline{c} = \overline{d}$ et $h \in \mathbb{N}$. Alors $\overline{hc} = \overline{hd}$.
Posons $b = \overline{a_2 - a_1 - a_4 + a_3}$.

Supposons m=2m'+1. Alors 2(m'+1) =m+1 et 2 est inversible.
Supposons $b = \overline{2t - 2w}$.
Alors, $b = \overline{2t-2w} \iff \overline{2^{-1}b} = \overline{2^{-1}2(t-w)} \iff \overline{2^{-1}b} = \overline{t-w} \iff t = \overline{2^{-1}b + w}$.

Alors,

$$\frac{1}{m^5} \sum_{w,t,\ \overline{a_2-a_1-a_4+a_3}=\overline{2t-2w}} 1 = \frac{1}{m^5} \sum_{w,t,\ t=\overline{2^{-1}b+w}} 1 = \frac{1}{m^5} \sum_{w} 1 = \frac{1}{m^4} \ .$$

Donc,

$$P\Big\{\{\overline{X+Y+U+W} = a_1\} \cap \{\overline{X+Y+V+T} = a_2\} \cap \{\overline{X+Z+U+T} = a_3\} \cap \{\overline{X+Z+V+W} = a_4\}\Big\} = \frac{1}{m^4} \ .$$

Supposons m =2m'. Alors $\{2(t-w) \mid t, w\} \equiv \{2r | r = 0, 1, ..., m-1\} \equiv \{2r | r = 0, 1, ..., m'-1\}$.
Donc, si b impair $b \neq \overline{2(t-w)}$. Alors,

$$\frac{1}{m^5} \sum_{w,t,\ \overline{a_2-a_1-a_4+a_3}=\overline{2t-2w}} 1 = 0 \ .$$

Si b= 2c, $\overline{2(t-w)} = 2c$ ou $0 \leq c < m'$.
Quel sont les nombres x, $0 \leq x < m$, tels que $2x \equiv 2c$? Cette relation est équivalente à 2x=2c+2km', qui est équivalente à x=c+km', c'est à dire, x égal à c ou c+m'.
Donc, $\overline{2(t-w)} = 2c \iff \overline{t-w} = c$ ou bien $\overline{t-w} = c+m' \iff t = \overline{w+c}$ ou $t = \overline{w+c+m'}$.
Donc,

$$\frac{1}{m^5} \sum_{w,t,\ \overline{a_2-a_1-a_4+a_3}=\overline{2t-2w}} 1 \ = \ \frac{1}{m^5} \sum_{w,t,\ 2c=\overline{2t-2w}} 1$$

$$= \frac{1}{m^5} \sum_{w,t,\ t=\overline{w+c}} 1 \ + \ \frac{1}{m^5} \sum_{w,t,\ t=\overline{w+c+m'}} 1$$

$$= \frac{1}{m^4} + \frac{1}{m^4} = \frac{2}{m^4} \ . \blacksquare$$

Corollaire 5.2 *On suppose m pair. On suppose que (X,Y,Z,U,V,W,T) est un vecteur aléatoire indépendant et de loi uniforme. On suppose* $X_1 = \overline{X+Y+U+W}$, $X_2 = \overline{X+Y+V+T}$, $X_3 = \overline{X+Z+U+T}$ *et* $X_4 = \overline{X+Z+V+W}$.
Alors, $\overline{X_2 - X_1 + X_3 - X_4}$ *est pair.*

Corollaire 5.3 *On suppose $m = 2$. On suppose que (X,Y,Z,U,V,W,T) est un vecteur aléatoire indépendant et de loi uniforme. On suppose $X_1 = \overline{X+Y+U+W}$, $X_2 = \overline{X+Y+V+T}$, $X_3 = \overline{X+Z+U+T}$ et $X_4 = \overline{X+Z+V+W}$.*
Alors, $\overline{X_2 - X_1 + X_3 - X_4} = 0$.

On étudie maintenant un deuxième cas.

Proposition 5.4 *On suppose (Y,Z,U,V,W,T) est un vecteur aléatoire indépendant et de loi uniforme. Alors,*

$$P\Big\{\{\overline{Y+U+W} = a_1\} \cap \{\overline{Y+V+T} = a_2\} \cap \{\overline{Z+U+T} = a_3\} \cap \{\overline{Z+V+W} = a_4\}\Big\}$$

$= \frac{1}{m^4}$ *si m est impair,*

$= 0$ *si m est pair et $\overline{a_2 - a_1 - a_4 + a_3}$ est impair,*
$= \frac{2}{m^4}$ *si m est pair et $\overline{a_2 - a_1 - a_4 + a_3}$ est pair.*

Démonstration On a les relations suivantes :

$$P\Big\{\{\overline{Y+U+W} = a_1\} \cap \{\overline{Y+V+T} = a_2\} \cap \{\overline{Z+U+T} = a_3\} \cap \{\overline{Z+V+W} = a_4\}\Big\}$$

$$= \sum_{u,v,w,t} P\Big\{\{U=u\} \cap \{V=v\} \cap \{W=w\} \cap \{T=t\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots \cap \{\overline{Y+U+W} = a_1\} \cap \{\overline{Y+V+T} = a_2\} \cap \{\overline{Z+U+T} = a_3\} \cap \{\overline{Z+V+W} = a_4\}\Big\}$$

$$= (1/m^4) \sum_{u,v,w,t} P\Big\{\{Y = \overline{a_1 - u - w}\} \cap \{Y = \overline{a_2 - v - t}\} \cap \{Z = \overline{a_3 - u - t}\} \cap \{Z = \overline{a_4 - v - w}\}\Big\}$$

$$= (1/m^4) \sum_{u,v,w,t,\ \overline{a_1-u-w}=\overline{a_2-v-t}\ ,\ \overline{a_3-u-t}=\overline{a_4-v-w}} (1/m^2) \ .$$

C'est la même équation que dans la démonstration de la proposition 5.1. ■

On étudie enfin un dernier cas.

Proposition 5.5 *On suppose (Y,Z,U,V) est un vecteur aléatoire indépendant et de loi uniforme. Alors,*

$$P\Big\{\{\overline{Y+U} = a_1\} \cap \{\overline{Y+V} = a_2\} \cap \{\overline{Z+U} = a_3\} \cap \{\overline{Z+V} = a_4\}\Big\}$$

$= 0$ *si $\overline{a_2 - a_1} \neq \overline{a_4 - a_3}$,*

$= 1/m^3$ *si $\overline{a_2 - a_1} = \overline{a_4 - a_3}$.*

Démonstration On a les relations suivantes

$$P\Big\{\{\overline{Y+U}=a_1\}\cap\{\overline{Y+V}=a_2\}\cap\{\overline{Z+U}=a_3\}\cap\{\overline{Z+V}=a_4\}\Big\}$$

$$=\sum_{u,v}P\Big\{\{U=u\}\cap\{V=v\}\cap\{\overline{Y+U}=a_1\}\cap\{\overline{Y+V}=a_2\}\cap\{\overline{Z+U}=a_3\}\cap\{\overline{Z+V}=a_4\}\Big\}$$

$$=(1/m^2)\sum_{u,v}P\Big\{\{Y=\overline{a_1-u}\}\cap\{Y=\overline{a_2-v}\}\cap\{Z=\overline{a_3-u}\}\cap\{Z=\overline{a_4-v}\}\Big\}$$

$$=(1/m^2)\sum_{u,v\ \overline{a_1-u}=\overline{a_2-v}\ ,\ \overline{a_3-u}=\overline{a_4-v}}(1/m^2)$$

$$=(1/m^4)\sum_{u,v,\ v=\overline{a_2+u-a_1}\ ,\ v=\overline{a_4+u-a_3}}1$$

$$=(1/m^4)\sum_{u,\ \overline{a_2+u-a_1}=\overline{a_4+u-a_3}}1$$

$$=(1/m^4)\sum_{u,\ \overline{a_2-a_1}=\overline{a_4-a_3}}1$$

$$=(1/m^3)\sum_{\overline{a_2-a_1}=\overline{a_4-a_3}}1$$

$= 0$ si $\overline{a_2-a_1}\neq\overline{a_4-a_3}$,

$= 1/m^3$ si $\overline{a_2-a_1}=\overline{a_4-a_3}$. ■

Corollaire 5.6 *On suppose* $X_1=\overline{Y+U}$, $X_2=\overline{Y+V}$, $X_3=\overline{Z+U}$, $X_4=\overline{Z+V}$.

Alors, $\overline{X_2-X_1}=\overline{X_4-X_3}$.

Dans ce dernier cas, il y a toujours dépendance alors que dans les cas précédents cela dépendait de m.

On peut aussi prouver la proposition 5.5 par la méthode des matrices.

Proposition 5.7 *On suppose (Y,Z,U,V) est un vecteur aléatoire indépendant et de loi uniforme. On suppose* $X_1=\overline{Y+U}$, $X_2=\overline{Y+V}$, $X_3=\overline{Z+U}$, $X_4=\overline{Z+V}$.

Alors, $\overline{X_2-X_1}=\overline{X_4-X_3}$.

Démonstration La matrice associée est

$$C=\begin{pmatrix}1&0&1&0\\1&0&0&1\\0&1&1&0\\0&1&0&1\end{pmatrix}.$$

Or det(C)= 0. Donc il y a une relation linéaire modulo m entre les X_s.
On voit facilement que $X_1 - X_2 = X_3 - X_4$ modulo m. ■

Si on veut chercher les dépendances par la méthode des matrices, dans certains cas ce sera peut être possible comme le cas précédent.

Par exemple pour trouver la dépendance linéaire entre u=(3,1,-1) , v=(-1,1,2) , w=(1,-1,1) et t=(5-2,3), on sait que l'on doit résoudre le système au+bv+cw+dt=0, i.e. on résoud le système au+bv+cw=dt et on trouve a,b,c, le résultat dépendant de d.

Dans le cas qui nous occupe ici, il faudra aussi résoudre les equations linéaires en fonction de a_1, a_2, a_3, a_4 i.e. en fonction de X_1, X_2, X_3, X_4 . En effet, il faut resoudre le système $\overline{Y+U} = a_1$, $\overline{Y+V} = a_2$, $\overline{Z+U} = a_3$, $\overline{Z+V} = a_4$
i.e. $Y = \overline{a_1 - u}$, $Y = \overline{a_2 - v}$, $Z = \overline{a_3 - u}$, $Z = \overline{a_4 - v}$,
i.e. $\overline{a_1 - u} = \overline{a_2 - v}$, $\overline{a_3 - u} = \overline{a_4 - v}$.

5.1 Conclusion

On voit donc que lorsque p=4 des dépendances sont possibles. On remarque aussi que ce n'est pas parce que, pour tout i, pour tout s, il existe $t \neq s$ tel que $X_{i,j_s}^{d^s} = X_{i,j_t}^{d^t}$ qu'il y a dépendance.

6 Etude des dépendances de 6-uplets

On ne peut pas étudier ici tous les cas possibles lorsque p=6 ou 8 . Mais les dépendances que nous allons étudier éclairent la nature des dépendances entre 6 termes.

Proposition 6.1 *On suppose que* $(Y, Z, U, V, W, A, B, C, D, R, S, T, E, F, H)$ *est un vecteur aléatoire de loi uniforme.*

Alors,

$$P\Big\{\{\overline{Y+Z+U+V+W}=a_1\}\cap\{\overline{Y+A+B+C+D}=a_2\}\cap\{\overline{Z+A+R+S+T}=a_3\}.....$$

$$....\cap\{\overline{U+B+R+E+F}=a_4\}\cap\{\overline{V+C+S+E+H}=a_5\}\cap\{\overline{W+D+T+F+H}=a_6\}\Big\}$$

$= (1/m^6)$ *si m=2m'+1.*

$= (2/m^6)$ *si m=2m' et si* $\overline{a_3-a_4-a_1+a_2-a_6+a_5}$ *est pair.*
$= 0$ *si m=2m' et si* $\overline{a_3-a_4-a_1+a_2-a_6+a_5}$ *est impair.*

Démonstration On a les égalités suivantes

$$P\Big\{\{\overline{Y+Z+U+V+W}=a_1\}\cap\{\overline{Y+A+B+C+D}=a_2\}\cap\{\overline{Z+A+R+S+T}=a_3\}.....$$

$$....\cap\{\overline{U+B+R+E+F}=a_4\}\cap\{\overline{V+C+S+E+H}=a_5\}\cap\{\overline{W+D+T+F+H}=a_6\}\Big\}$$

$$=\sum_{z,u,v,w,a,b,c,d,s,t,e,f} P\Big\{\{Z=z\}\cap\{U=u\}\cap\{V=v\}\cap\{W=w\}\cap\{A=a\}\cap\{B=b\}$$

$$.....\cap\{C=c\}\cap\{D=d\}\cap\{S=s\}\cap\{T=t\}\cap\{E=e\}\cap\{F=f\}.....$$
$$\cap\{\overline{Y+Z+U+V+W}=a_1\}\cap\{\overline{Y+A+B+C+D}=a_2\}\cap\{\overline{Z+A+R+S+T}=a_3\}.....$$
$$....\cap\{\overline{U+B+R+E+F}=a_4\}\cap\{\overline{V+C+S+E+H}=a_5\}\cap\{\overline{W+D+T+F+H}=a_6\}\Big\}$$

$$=\sum_{z,u,v,w,a,b,c,d,s,t,e,f} P\Big\{\{Z=z\}\cap\{U=u\}\cap\{V=v\}\cap\{W=w\}\cap\{A=a\}\cap\{B=b\}$$

$$.....\cap\{C=c\}\cap\{D=d\}\cap\{S=s\}\cap\{T=t\}\cap\{E=e\}\cap\{F=f\}.....$$
$$\cap\{Y=\overline{a_1-z-u-v-w}\}\cap\{Y=\overline{a_2-a-b-c-d}\}\cap\{R=\overline{a_3-z-a-s-t}\}.....$$
$$....\cap\{R=\overline{a_4-u-b-e-f}\}\cap\{H=\overline{a_5-v-c-s-e}\}\cap\{H=\overline{a_6-w-d-t-f}\}\Big\}$$

$$= (1/m^{12}) \sum_{z,u,v,w,a,b,c,d,s,t,e,f} P\Big\{\{Y = \overline{a_1 - z - u - v - w}\} \cap \{Y = \overline{a_2 - a - b - c - d}\} \cap \{R = \overline{a_3 - z - a - s - t}\} ...$$

$$.... \cap \{R = \overline{a_4 - u - b - e - f}\} \cap \{H = \overline{a_5 - v - c - s - e}\} \cap \{H = \overline{a_6 - w - d - t - f}\}\Big\}$$

$$= (1/m^{12}) \sum_{z,u,v,w,a,b,c,d,s,t,e,f;\ \overline{a_1-z-u-v-w}=\overline{a_2-a-b-c-d},\ \overline{a_3-z-a-s-t}=\overline{a_4-u-b-e-f},\ \overline{a_5-v-c-s-e}=\overline{a_6-w-d-t-f}}$$

$$P\Big\{\{Y = \overline{a_1 - z - u - v - w}\} \cap \{Y = \overline{a_2 - a - b - c - d}\} \cap \{R = \overline{a_3 - z - a - s - t}\} .....$$

$$.... \cap \{R = \overline{a_4 - u - b - e - f}\} \cap \{H = \overline{a_5 - v - c - s - e}\} \cap \{H = \overline{a_6 - w - d - t - f}\}\Big\}$$

$$= m^{-12} \sum_{z,u,v,w,a,b,c,d,s,t,e,f;\ z=\overline{a_1-a_2+a+b+c+d-u-v-w},\ z=\overline{a_3-a_4+u+b+e+f-a-s-t},\ \overline{a_5-v-c-s-e}=\overline{a_6-w-d-t-f}} m^{-3}$$

$$= (1/m^{12}) \sum_{u,v,w,a,b,c,d,s,t,e,f;\ \overline{a_1-a_2+a+b+c+d-u-v-w},=\overline{a_3-a_4+u+b+e+f-a-s-t},\ \overline{a_5-v-c-s-e}=\overline{a_6-w-d-t-f}} (1/m^3)$$

$$= (1/m^{12}) \sum_{u,v,w,a,b,c,d,s,t,e,f;\ \overline{a_1-a_2+c+d-v-w}=\overline{a_3-a_4+2u+e+f-2a-s-t},\ \overline{a_5-v-c-s-e}=\overline{a_6-w-d-t-f}} (1/m^3)$$

$$= (1/m^{15}) \sum_{u,v,w,a,b,c,d,s,t,e,f;\ t=\overline{a_3-a_4-a_1+a_2-c-d+v+w+2u+e+f-2a-s},\ t=\overline{a_6-a_5+v+c+s+e-w-d-f}} 1$$

$$= (1/m^{15}) \sum_{u,v,w,a,b,c,d,s,e,f;\ \overline{a_3-a_4-a_1+a_2-c-d+v+w+2u+e+f-2a-s}=\overline{a_6-a_5+v+c+s+e-w-d-f}} 1$$

$$= (1/m^{15}) \sum_{u,v,w,a,b,c,d,s,e,f;\ \overline{a_3-a_4-a_1+a_2+2u-2a}=\overline{a_6-a_5+2c+2s-2w-2f}} 1$$

$$= (1/m^{15}) \sum_{u,v,w,a,b,c,d,s,e,f;\ \overline{a_3-a_4-a_1+a_2-a_6+a_5}=\overline{2a-2u+2c+2s-2w-2f}} 1$$

$= (1/m^6)$ si m=2m'+1,

et, par le même raisonnement qu'en section 5,

$= (2/m^6)$ si m=2m' et si $\overline{a_3 - a_4 - a_1 + a_2 - a_6 + a_5}$ est pair.
$= 0$ si m=2m' et si $\overline{a_3 - a_4 - a_1 + a_2 - a_6 + a_5}$ est impair.

Corollaire 6.2 *On suppose m=2. Alors,*

$$P\Big\{\{\overline{Y+Z+U+V+W} = a_1\} \cap \{\overline{Y+A+B+C+D} = a_2\} \cap \{\overline{Z+A+R+S+T} = a_3\}.....$$

$$.... \cap \{\overline{U+B+R+E+F} = a_4\} \cap \{\overline{V+C+S+E+H} = a_5\} \cap \{\overline{W+D+T+F+H} = a_6\}\Big\}$$

$= (1/m^5)$ *si* $\overline{a_3 - a_4 - a_1 + a_2 - a_6 + a_5} = 0.$
$= 0$ *si* $\overline{a_3 - a_4 - a_1 + a_2 - a_6 + a_5} = 1.$

Proposition 6.3 *On suppose que* (Y, Z, U, V, W, T) *est un vecteur aléatoire de loi uniforme. Alors,*

$$P\Big\{\{\overline{Y+U} = a_1\} \cap \{\overline{Y+V} = a_2\} \cap \{\overline{Z+U} = a_3\} \cap \{\overline{Z+W} = a_4\} \cap \{\overline{T+V} = a_5\} \cap \{\overline{T+W} = a_6\}\Big\}$$

$= (1/m^5)$ *si* $\overline{a_2 - a_5 + a_3 - a_1} = \overline{a_4 - a_6}$,
$= 0$ *si* $\overline{a_2 - a_5 + a_3 - a_1} \neq \overline{a_4 - a_6}$.

Démonstration On a les égalités suivantes

$$P\Big\{\{\overline{Y+U} = a_1\} \cap \{\overline{Y+V} = a_2\} \cap \{\overline{Z+U} = a_3\} \cap \{\overline{Z+W} = a_4\} \cap \{\overline{T+V} = a_5\} \cap \{\overline{T+W} = a_6\}\Big\}$$

$$= P\Big\{ \cup_{y,z,t} \{Y = y\} \cap \{Z = z\} \cap \{T = t\}.....$$

$$.....\cap \{\overline{Y+U} = a_1\} \cap \{\overline{Y+V} = a_2\} \cap \{\overline{Z+U} = a_3\} \cap \{\overline{Z+W} = a_4\} \cap \{\overline{T+V} = a_5\} \cap \{\overline{T+W} = a_6\}\Big\}$$

$$= \sum_{y,z,t} P\Big\{ \{Y = y\} \cap \{Z = z\} \cap \{T = t\}.....$$

$$\{U = \overline{a_1 - y}\} \cap \{V = \overline{a_2 - y}\} \cap \{U = \overline{a_3 - z}\} \cap \{W = \overline{a_4 - z}\} \cap \{V = \overline{a_5 - t}\} \cap \{W = \overline{a_6 - t}\}\Big\}$$

$$= (1/m^3) \sum_{y,z,t} P\Big\{\{U = \overline{a_1 - y}\} \cap \{V = \overline{a_2 - y}\} \cap \{U = \overline{a_3 - z}\} \cap \{W = \overline{a_4 - z}\} \cap \{V = \overline{a_5 - t}\} \cap \{W = \overline{a_6 - t}\}\Big\}$$

$$= (1/m^3) \sum_{y,z,t,\ \overline{a_1 - y} = \overline{a_3 - z}\ ,\ \overline{a_2 - y} = \overline{a_5 - t},\ \overline{a_4 - z} = \overline{a_6 - t}} (1/m^3)$$

$$= (1/m^6) \sum_{y,z,t,\ y = \overline{a_1 - a_3 + z}\ ,\ y = \overline{a_2 - a_5 + t},\ \overline{a_4 - z} = \overline{a_6 - t}} 1$$

$$= (1/m^6) \sum_{z,t,\ \overline{a_1 - a_3 + z} = \overline{a_2 - a_5 + t},\ \overline{a_4 - z} = \overline{a_6 - t}} 1$$

$$= (1/m^6) \sum_{z,t,\ z=\overline{a_2-a_5+a_3-a_1+t},\ z=\overline{a_4-a_6+t}} 1$$

$$= (1/m^6) \sum_{t,\ \overline{a_2-a_5+a_3-a_1+t}=\overline{a_4-a_6+t}} 1$$

$$= (1/m^6) \sum_{t,\ \overline{a_2-a_5+a_3-a_1}=\overline{a_4-a_6}} 1$$

$= (1/m^5)$ si $\overline{a_2 - a_5 + a_3 - a_1} = \overline{a_4 - a_6}$,
$= 0$ si $\overline{a_2 - a_5 + a_3 - a_1} \neq \overline{a_4 - a_6}$. ■

Proposition 6.4 *On suppose que (X, Y, Z, U, V, W, T, R) est un vecteur aléatoire de loi uniforme. Alors,*

$$P\Big\{\{\overline{X+Y+R} = a_1\} \cap \{\overline{X+U+T} = a_2\} \cap \{\overline{V+Y+T} = a_3\}.....$$

$$..... \cap \{\overline{Z+U+R} = a_4\} \cap \{\overline{W+Z+T} = a_5\} \cap \{\overline{W+V+R} = a_6\}\Big\}$$

$= 1/m^6$ *si m=3m'+1 ou bien m=3m'+2.*

$= 0$ *si m=3m' et* $\overline{a_2 - 2a_3 - a_1 - a_4 + a_5 + 2a_6} \neq 3c$.

$= 3/m^6$ *si m=3m' et* $\overline{a_2 - 2a_3 - a_1 - a_4 + a_5 + 2a_6} = 3c$.

Démonstration On a les égalités suivantes

$$P\Big\{\{\overline{X+Y+R} = a_1\} \cap \{\overline{X+U+T} = a_2\} \cap \{\overline{V+Y+T} = a_3\}.....$$

$$..... \cap \{\overline{Z+U+R} = a_4\} \cap \{\overline{W+Z+T} = a_5\} \cap \{\overline{W+V+R} = a_6\}\Big\}$$

$$= \sum_{x,y,z,u,v,w} P\Big\{\{X = x\} \cap \{Y = y\} \cap \{Z = z\} \cap \{U = u\} \cap \{V = v\} \cap \{W = w\}$$

$$..... \cap \{\overline{X+Y+R} = a_1\} \cap \{\overline{X+U+T} = a_2\} \cap \{\overline{V+Y+T} = a_3\}.....$$

$$..... \cap \{\overline{Z+U+R} = a_4\} \cap \{\overline{W+Z+T} = a_5\} \cap \{\overline{W+V+R} = a_6\}\Big\}$$

$$= \sum_{x,y,z,u,v,w} P\Big\{\{X = x\} \cap \{Y = y\} \cap \{Z = z\} \cap \{U = u\} \cap \{V = v\} \cap \{W = w\}.....$$

$$..... \cap \{R = \overline{a_1 - x - y}\} \cap \{T = \overline{a_2 - x - u}\} \cap \{T = \overline{a_3 - y - v}\}.....$$

$$..... \cap \{R = \overline{a_4 - u - z}\} \cap \{T = \overline{a_5 - w - z}\} \cap \{R = \overline{a_6 - w - v}\}\Big\}$$

$$= (1/m^6) \sum_{x,y,z,u,v,w} P\Big\{\{R = \overline{a_1 - x - y}\} \cap \{T = \overline{a_2 - x - u}\} \cap \{T = \overline{a_3 - y - v}\} \ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \cap \{R = \overline{a_4 - u - z}\} \cap \{T = \overline{a_5 - w - z}\} \cap \{R = \overline{a_6 - w - v}\}\Big\}$$

$$= (1/m^6) \sum_{x,y,z,u,v,w,\ \overline{a_1-x-y}=\overline{a_4-u-z}=\overline{a_6-w-v},\ \overline{a_2-x-u}=\overline{a_3-y-v}=\overline{a_5-w-z}} (1/m^2)$$

$$= (1/m^6) \sum_{x,y,z,u,v,w,\ x=\overline{a_1-a_4+u+z-y},\ \overline{a_4-u-z}=\overline{a_6-w-v},\ x=\overline{a_2-a_3-u+y+v},\ \overline{a_3-y-v}=\overline{a_5-w-z}} (1/m^2)$$

$$= (1/m^8) \sum_{x,y,z,u,v,w,\ x=\overline{a_1-a_4+u+z-y},\ \overline{a_1-a_4+u+z-y}=\overline{a_2-a_3-u+y+v},\ \overline{a_4-u-z}=\overline{a_6-w-v},\ \overline{a_3-y-v}=\overline{a_5-w-z}} 1$$

$$= (1/m^8) \sum_{y,z,u,v,w,\ \overline{a_1-a_4+u+z-y}=\overline{a_2-a_3-u+y+v},\ \overline{a_4-u-z}=\overline{a_6-w-v},\ \overline{a_3-y-v}=\overline{a_5-w-z}} 1$$

$$= (1/m^8) \sum_{y,z,u,v,w,\ z=\overline{a_2-a_3-a_1+a_4-2u+2y+v},\ z=\overline{a_4-a_6+w+v-u},\ z=\overline{a_5-a_3+y+v-w}} 1$$

$$= (1/m^8) \sum_{y,u,v,w,\ \overline{a_2-a_3-a_1+a_4-2u+2y+v}=\overline{a_4-a_6+w+v-u},\ \overline{a_4-a_6+w+v-u}=\overline{a_5-a_3+y+v-w}} 1$$

$$= (1/m^8) \sum_{y,u,v,w,\ \overline{a_2-a_3-a_1-u+2y}=\overline{-a_6+w},\ \overline{a_4-a_6-u}=\overline{a_5-a_3+y-2w}} 1$$

$$= (1/m^8) \sum_{y,u,v,w,\ u=\overline{a_2-a_3-a_1+a_6-w+2y},\ u=\overline{a_4-a_6-a_5+a_3-y+2w}} 1$$

$$= (1/m^8) \sum_{y,v,w,\ \overline{a_2-a_3-a_1+a_6-w+2y}=\overline{a_4-a_6-a_5+a_3-y+2w}} 1$$

$$= (1/m^8) \sum_{y,v,w,\ \overline{a_2-2a_3-a_1-a_4+a_5+2a_6}=\overline{3w-3y}} 1 \ .$$

Posons $b = \overline{a_2 - 2a_3 - a_1 - a_4 + a_5 + 2a_6}$.

Supposons m=3m'+1 ou m=3m'+2. Si m=3m'+1, Alors $3(2m' + 1) = 6m' + 3 = 2(3m' + 1) + 1 = 2m + 1 \equiv 1$ modulo m et 3 est inversible.

Si m=3m'+2. Alors $3(m' + 1) = 3m' + 2 + 1 = m + 1 \equiv 1$ et 3 est inversible.

Supposons $b = \overline{3w - 3y}$.

Alors, $b = \overline{3w - 3y} \iff \overline{3^{-1}b} = \overline{3^{-1}3(w - y)} \iff \overline{3^{-1}b} = \overline{w - y} \iff w = \overline{3^{-1}b + y}$.

Alors,

$$(1/m^8) \sum_{y,v,w,\ \overline{a_2-2a_3-a_1-a_4+a_5+2a_6}=\overline{3w-3y}} 1 = (1/m^8) \sum_{y,v,w,\ w=\overline{3^{-1}(a_2-2a_3-a_1-a_4+a_5+2a_6)+y}} 1$$

$$= (1/m^8) \sum_{y,v} 1 = (1/m^6) \ .$$

Donc il y a indépendance.

Supposons m =3m' Alors 3 n'est pas inversible. Alors $\{3(t-w) \mid t, w\} \equiv \{3r | r = 0, 1, ..., m-1\} \equiv \{3r | r = 0, 1, ..., m' - 1\}$. Donc, $\{\overline{3(t - w)} \mid t, w\} = \{3r | r = 0, 1, 2, ..., m' - 1\}$.

Donc, si $b \neq \overline{3r}$. Alors,

$$(1/m^8) \sum_{y,v,w,\ \overline{a_2-2a_3-a_1-a_4+a_5+2a_6}=\overline{3w-3y}} 1 = 0 \ .$$

Si b= 3c, $\overline{3(w - y)} = 3c$ ou $0 \leq c < m'$.

Quel sont les nombres x modulo 3m', $0 \leq x < m$, tels que $3x \equiv 3c$? Cette relation est équivalente à 3x=3c+3km', qui est équivalente à x=c+km', c'est à dire, x égal à c, c+m' ou c+2m'.

Donc, $\overline{3(w - y)} = 3c \iff \overline{w - y} = c$ ou $\overline{w - y} = c + m'$ ou $\overline{w - y} = c + 2m' \iff y = \overline{w - c}$ ou $y = \overline{w - c - m'}$ ou $y = \overline{w - c - 2m'}$. Donc,

$$(1/m^8) \sum_{y,v,w,\ \overline{a_2-2a_3-a_1-a_4+a_5+2a_6}=\overline{3w-3y}} 1 = (1/m^8) \sum_{y,v,w,\ 3c=\overline{3w-3y}} 1$$

$$= (1/m^8)\Big[\sum_{y,v,w,\ y=\overline{w-c}} 1 + \sum_{y,v,w,\ y=\overline{w-c-m'}} 1 + \sum_{y,v,w,\ y=\overline{w-c-2m'}} 1\Big]$$

$$= (1/m^8)\Big[\sum_{v,w} 1 + \sum_{v,w} 1 + \sum_{v,w} 1\Big] = (3/m^6) \ . \ \blacksquare$$

Proposition 6.5 *On suppose que* $(X, Y, Z, U, V, W, A, B, C, D, R, S, T, E, F, G, H, K, L, M)$ *est un vecteur aléatoire de loi uniforme. Alors,*

$$P\Big\{\{\overline{X+Z+T+U+V+W+R+S+A+B}=a_1\}\cap\{\overline{X+Z+T+U+C+D+E+F+G+H}=a_2\}$$
$$\cap\{\overline{X+V+W+R+C+D+E+K+L+M}=a_3\}\cap\{\overline{Y+Z+V+S+A+C+F+G+K+L}=a_4\}$$
$$\cap\{\overline{Y+T+W+S+B+D+F+H+K+M}=a_5\}\cap\{\overline{Y+U+R+A+B+E+G+H+L+M}=a_6\}\Big\}$$

$= (1/m^6)$ *si m=3m'+1 ou m=3m'+2,*

$= 0$ *si m=3m' et si* $\overline{a_2+a_6+a_1-2a_5-2a_3+a_4}\neq 3c$,

$= (3/m^6)$ *si m=3m' et si* $\overline{a_2+a_6+a_1-2a_5-2a_3+a_4}=3c$.

Démonstration On a les égalités suivantes

$$P\Big\{\{\overline{X+Z+T+U+V+W+R+S+A+B}=a_1\}\cap\{\overline{X+Z+T+U+C+D+E+F+G+H}=a_2\}$$
$$\cap\{\overline{X+V+W+R+C+D+E+K+L+M}=a_3\}\cap\{\overline{Y+Z+V+S+A+C+F+G+K+L}=a_4\}$$
$$\cap\{\overline{Y+T+W+S+B+D+F+H+K+M}=a_5\}\cap\{\overline{Y+U+R+A+B+E+G+H+L+M}=a_6\}\Big\}$$

$$=\sum_{x,y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s} P\Big\{\{X=x\}\cap\{Y=y\}\cap\{Z=z\}\cap\{M=m\}\cap\{T=t\}\cap\{L=l\}\cap\{K=k\}$$
$$\cap\{U=u\}\cap\{V=v\}\cap\{H=h\}\cap\{W=w\}\cap\{G=g\}\cap\{R=r\}\cap\{F=f\}\cap\{E=e\}\cap\{A=a\}\cap\{D=d\}\cap\{S=s\}$$
$$\cap\{\{\overline{X+Z+T+U+V+W+R+S+A+B}=a_1\}\cap\{\overline{X+Z+T+U+C+D+E+F+G+H}=a_2\}$$
$$\cap\{\overline{X+V+W+R+C+D+E+K+L+M}=a_3\}\cap\{\overline{Y+Z+V+S+A+C+F+G+K+L}=a_4\}$$
$$\cap\{\overline{Y+T+W+S+B+D+F+H+K+M}=a_5\}\cap\{\overline{Y+U+R+A+B+E+G+H+L+M}=a_6\}\Big\}$$

$$=\sum_{x,y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s} P\Big\{\{X=x\}\cap\{Y=y\}\cap\{Z=z\}\cap\{M=m\}\cap\{T=t\}\cap\{L=l\}\cap\{K=k\}$$
$$\cap\{U=u\}\cap\{V=v\}\cap\{H=h\}\cap\{W=w\}\cap\{G=g\}\cap\{R=r\}\cap\{F=f\}\cap\{E=e\}\cap\{A=a\}\cap\{D=d\}\cap\{S=s\}$$
$$\cap\{\overline{x+z+t+u+v+w+r+s+a+B}=a_1\}\cap\{\overline{x+z+t+u+C+d+e+f+g+h}=a_2\}$$
$$\cap\{\overline{x+v+w+r+C+d+e+k+l+m}=a_3\}\cap\{\overline{y+z+v+s+a+C+f+g+k+l}=a_4\}$$
$$\cap\{\overline{y+t+w+s+B+d+f+h+k+m}=a_5\}\cap\{\overline{y+u+r+a+B+e+g+h+l+m}=a_6\}\Big\}$$

$$=(1/m^{18})\sum_{x,y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$
$$P\Big\{\{B=\overline{a_1-x-z-t-u-v-w-r-s-a}\}\cap\{C=\overline{a_2-x-z-t-u-d-e-f-g-h}\}$$
$$\cap\{C=\overline{a_3-x-v-w-r-d-e-k-l-m}\}\cap\{C=\overline{a_4-y-z-v-s-a-f-g-k-l}\}$$
$$\cap\{B=\overline{a_5-y-t-w-s-d-f-h-k-m}\}\cap\{B=\overline{a_6-y-u-r-a-e-g-h-l-m}\}\Big\}$$

$$= (1/m^{20}) \sum_{x,y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{\overline{a_1-x-z-t-u-v-w-r-s-a}=\overline{a_5-y-t-w-s-d-f-h-k-m}=\overline{a_6-y-u-r-a-e-g-h-l-m}}$$

$$\sum_{\overline{a_2-x-z-t-u-d-e-f-g-h}=\overline{a_3-x-v-w-r-d-e-k-l-m}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

$$= (1/m^{20}) \sum_{x,y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{x=\overline{a_1-a_5-z-t-u-v-w-r-s-a+y+t+w+s+d+f+h+k+m}}$$

$$\sum_{\overline{a_5-y-t-w-s-d-f-h-k-m}=\overline{a_6-y-u-r-a-e-g-h-l-m}}$$

$$\sum_{\overline{a_2-x-z-t-u-d-e-f-g-h}=\overline{a_3-x-v-w-r-d-e-k-l-m}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

$$= (1/m^{20}) \sum_{x,y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{x=\overline{a_1-a_5-z-u-v-r-a+y+d+f+h+k+m}}$$

$$\sum_{\overline{a_5-y-t-w-s-d-f-h-k-m}=\overline{a_6-y-u-r-a-e-g-h-l-m}}$$

$$\sum_{\overline{a_2-x-z-t-u-d-e-f-g-h}=\overline{a_3-x-v-w-r-d-e-k-l-m}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

(alors vu que $x = a_1 - a_5 - z - u - v - r - a + y + d + f + h + k + m$,
$-x = a_5 - a_1 + z + u + v + r + a - y - d - f - h - k - m$.

Donc, $a_2 - x - z - t - u - d - e - f - g - h$
$= a_2 + a_5 - a_1 + z + u + v + r + a - y - d - f - h - k - m - z - t - u - d - e - f - g - h$
$= a_2 + a_5 - a_1 + v + r + a - y - 2d - 2f - 2h - k - m - t - e - g$

$a_3 - x - v - w - r - d - e - k - l - m$
$= a_3 + a_5 - a_1 + z + u + v + r + a - y - d - f - h - k - m - v - w - r - d - e - k - l - m$
$= a_3 + a_5 - a_1 + z + u + a - y - 2d - f - h - 2k - 2m - w - e - l)$

$$= (1/m^{20}) \sum_{x,y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{x=\overline{a_1-a_5-z-u-v-r-a+y+d+f+h+k+m}}$$

$$\sum_{\overline{a_5-y-t-w-s-d-f-h-k-m}=\overline{a_6-y-u-r-a-e-g-h-l-m}\}}$$

$$\sum_{\overline{a_2+a_5-a_1+v+r+a-y-2d-2f-2h-k-m-t-e-g}=\overline{a_3+a_5-a_1+z+u+a-y-2d-f-h-2k-2m-w-e-l}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{\overline{a_5-y-t-w-s-d-f-h-k-m}=\overline{a_6-y-u-r-a-e-g-h-l-m}\}}$$

$$\sum_{\overline{a_2+a_5-a_1+v+r+a-y-2d-2f-2h-k-m-t-e-g}=\overline{a_3+a_5-a_1+z+u+a-y-2d-f-h-2k-2m-w-e-l}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{\overline{a_5-t-w-s-d-f-k}=\overline{a_6-u-r-a-e-g-l}\}}$$

$$\sum_{\overline{a_2+a_5-a_1+v+r+a-y-2d-2f-2h-k-m-t-e-g}=\overline{a_3+a_5-a_1+z+u+a-y-2d-f-h-2k-2m-w-e-l}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{t=\overline{a_5-a_6+u+r+a+e+g+l-w-s-d-f-k}}$$

$$\sum_{\overline{a_2+a_5-a_1+v+r+a-y-2d-2f-2h-k-m-t-e-g}=\overline{a_3+a_5-a_1+z+u+a-y-2d-f-h-2k-2m-w-e-l}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

(donc comme $t = a_5 - a_6 + u + r + a + e + g + l - w - s - d - f - k$,
$-t = a_6 - a_5 - u - r - a - e - g - l + w + s + d + f + k$

Alors, $a_2 + a_5 - a_1 + v + r + a - y - 2d - 2f - 2h - k - m - t - e - g$
$= a_2+a_5-a_1+v+r+a-y-2d-2f-2h-k-m-e-g+a_6-a_5-u-r-a-e-g-l+w+s+d+f+k$
$= a_2 + a_6 - a_1 + v - y - d - f - 2h - m - 2e - 2g - u - l + w + s)$

$$= (1/m^{20}) \sum_{y,z,m,t,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{t=\overline{a_5-a_6+u+r+a+e+g+l-w-s-d-f-k}}$$

$$\sum_{\overline{a_2+a_6-a_1+v-y-d-f-2h-m-2e-2g-u-l+w+s}=\overline{a_3+a_5-a_1+z+u+a-y-2d-f-h-2k-2m-w-e-l}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{\overline{a_2+a_6-a_1+v-y-d-f-2h-m-2e-2g-u-l+w+s}=\overline{a_3+a_5-a_1+z+u+a-y-2d-f-h-2k-2m-w-e-l}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{\overline{a_2+a_6-a_1+v-y-d-f-2h-m-2e-2g-u-l+w+s}=\overline{a_3+a_5-a_1+z+u+a-y-2d-f-h-2k-2m-w-e-l}}$$

$$\sum_{\overline{a_3+a_5-a_1+z+u+a-y-2d-f-h-2k-2m-w-e-l}=\overline{a_4-y-z-v-s-a-f-g-k-l}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{\overline{a_2+a_6+v-h-e-2g-2u+2w+s}=\overline{a_3+a_5+z+a-d-2k-m}}$$

$$\sum_{\overline{a_3+a_5-a_1+2z+u+2a-2d-h-k-2m-w-e}=\overline{a_4-v-s-g}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,e,a,d,s}$$

$$\sum_{\overline{a_2+a_6+v-h-e-2g-2u+2w+s}=\overline{a_3+a_5+z+a-d-m-2k}}$$

$$\sum_{e=\overline{a_3+a_5-a_1-a_4+2z+u+2a-2d-h-k-2m-w+v+s+g}} 1$$

(comme $e = a_3 + a_5 - a_1 - a_4 + 2z + u + 2a - 2d - h - k - 2m - w + v + s + g$
$-e = a_1 - a_5 - a_3 + a_4 - 2z - u - 2a + 2d + h + k + 2m + w - v - s - g$

Donc, $a_2 + a_6 + v - h - e - 2g - 2u + 2w + s$
$= a_2 + a_6 + v - h - 2g - 2u + 2w + s + a_1 - a_5 - a_3 + a_4 - 2z - u - 2a + 2d + h + k + 2m + w - v - s - g$
$= a_2 + a_6 + a_1 - a_5 - a_3 + a_4 - 3g - 3u + 3w - 2z - 2a + 2d + k + 2m$

$= a_3 + a_5 + z + a - d - m - 2k.$

Donc,
$a_2 + a_6 + a_1 - 2a_5 - 2a_3 + a_4 - 3g - 3u + 3w - 3z - 3a + 3d + 3m + 3k = 0.$

Donc,
$a_2 + a_6 + a_1 - 2a_5 - 2a_3 + a_4 = 3g + 3u - 3w + 3z + 3a - 3d - 3m - 3k)$

$$= (1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,e,a,d,s} \sum_{\overline{a_2+a_6+a_1-2a_5-2a_3+a_4}=\overline{3(g+u-w+z+a-d-m-k)}} \sum_{e=\overline{a_3+a_5-a_1-a_4+2z+u+2a-2d-h-k-2m-w+v+s+g}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,a,d,s,\ \overline{a_2+a_6+a_1-2a_5-2a_3+a_4}=\overline{3(g+u-w+z+a-d-m-k)}} 1 \ .$$

Posons $b = \overline{a_2 + a_6 + a_1 - 2a_5 - 2a_3 + a_4}$.

Supposons m=3m'+ ou m=3m'+2. Alors il existe 3^{-1} (cf plus haut).
Supposons $b = \overline{3(g + u - w + z + a - d - m - k)}$. Alors,
$b = \overline{3(g + u - w + z + a - d - m - k)} \iff \overline{3^{-1}b} = \overline{3^{-1}3(g + u - w + z + a - d - m - k)}$
$\iff \overline{3^{-1}b} = \overline{g + u - w + z + a - d - m - k} \iff g = \overline{3^{-1}b - u + w - z - a + d + m + k}.$

Alors,

$$(1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,a,d,s,\ \overline{a_2+a_6+a_1-2a_5-2a_3+a_4}=\overline{3(g+u-w+z+a-d-m-k)}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,a,d,s,\ g=\overline{3^{-1}b-u+w-z-a+d+m+k}} 1$$

$$= (1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,r,f,a,d,s,} 1 = (1/m^6) \ .$$

Donc il y a indépendance

Supposons m =3m' Alors 3 n'est pas inversible.
Alors $\{3(g + u - w + z + a - d - m - k) \mid g, u, w, z, a, d, m, k\} \equiv \{3r | r = 0, 1, ..., m - 1\} \equiv \{3r | r = 0, 1, ..., m' - 1\}$. Donc, $\{\overline{3(g + u - w + z + a - d - m - k)} \mid g, u, w, z, a, d, m, k\} = \{3r | r = 0, 1, 2, ..., m' - 1\}$.
Donc, si $b \neq \overline{3r}$. Alors,

$$(1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,a,d,s,\ \overline{a_2+a_6+a_1-2a_5-2a_3+a_4}=\overline{3(g+u-w+z+a-d-m-k)}} 1 = 0 \ .$$

Si b= 3c, $\overline{3(g+u-w+z+a-d-m-k)} = 3c$ ou $0 \leq c < m'$.

Quel sont les nombres x modulo 3m', $0 \leq x < m$, tels que $3x \equiv 3c$? Ce sont c, c+m' et c+2m'.

Donc, $\overline{3(g+u-w+z+a-d-m-k)} = 3c$

$\iff \overline{g+u-w+z+a-d-m-k} = c$ OU $\overline{g+u-w+z+a-d-m-k} = c+m'$ OU $\overline{g+u-w+z+a-d-m-k} = c+2m'$

$\iff w = \overline{g+u+z+a-d-m-k-c}$ OU $w = \overline{g+u+z+a-d-m-k-c-m'}$ OU $w = \overline{g+u+z+a-d-m-k-c-2m'}$. Donc,

$$(1/m^{20}) \sum_{y,z,m,l,u,k,v,h,w,g,r,f,a,d,s,\ \overline{a_2+a_6+a_1-2a_5-2a_3+a_4}=\overline{3(g+u-w+z+a-d-m-k)}} 1$$

$$= (1/m^{20})\Big[\sum_{y,z,m,l,u,k,v,h,w,g,r,f,a,d,s,\ w=\overline{g+u+z+a-d-m-k-c}} 1 + \sum_{y,z,m,l,u,k,v,h,w,g,r,f,a,d,s,\ w=\overline{g+u+z+a-d-m-k-c-m'}} 1 + \sum_{y,z,m,l,u,k,v,h,w,g,r,f,a,d,s,\ w=\overline{g+u+z+a-d-m-k-c-2m'}} 1\Big]$$

$$= (1/m^{20})\Big[\sum_{y,z,m,l,u,k,v,h,g,r,f,a,d,s} 1 + \sum_{y,z,m,l,u,k,v,h,g,r,f,a,d,s} 1 + \sum_{y,z,m,l,u,k,v,h,g,r,f,a,d,s} 1\Big] = (3/m^6)\,. \blacksquare$$

7 Etude des dépendances de 8-uplets

Les dépendances entre 8 termes montrent que l'on obtient des dépendances du même type que pour les cas à 4 ou 6 dimensions. En fait, c'est vrai aussi pour les cas à 5 et 7 dimensions.

Proposition 7.1 *On suppose que* $(X, Y, U, V, W, A, B, R, S, T)$ *est un vecteur aléatoire de loi uniforme. Alors,*

$$P\Big\{\{\overline{X+U+S}=a_1\}\cap\{\overline{X+U+V}=a_2\}\cap\{\overline{X+W+A}=a_3\}\cap\{\overline{X+T+B}=a_4\}....$$

$$....\cap\{\overline{X+R+B}=a_5\}\cap\{\overline{Y+R+S}=a_6\}\cap\{\overline{Y+T+V}=a_7\}\cap\{\overline{Y+W+A}=a_8\}\Big\}$$

$$= 0 \ si\ \overline{a_5-a_4-a_6+a_7-a_2+a_1}\neq 0,$$

$$= 1/m^7 \ si\ \overline{a_5-a_4-a_6+a_7-a_2+a_1}=0.$$

Démonstration On a les égalités suivantes

$$P\Big\{\{\overline{X+U+S}=a_1\}\cap\{\overline{X+U+V}=a_2\}\cap\{\overline{X+W+A}=a_3\}\cap\{\overline{X+T+B}=a_4\}....$$

$$....\cap\{\overline{X+R+B}=a_5\}\cap\{\overline{Y+R+S}=a_6\}\cap\{\overline{Y+T+V}=a_7\}\cap\{\overline{Y+W+A}=a_8\}\Big\}$$

$$= \sum_{u,v,w,a,b,r,s,t} P\Big\{\{U=u\}\cap\{S=s\}\cap\{W=w\}\cap\{A=a\}\cap\{T=t\}\cap\{B=b\}\cap\{R=r\}\cap\{V=v\}$$
$$\cap\{\overline{X+U+S}=a_1\}\cap\{\overline{X+U+V}=a_2\}\cap\{\overline{X+W+A}=a_3\}\cap\{\overline{X+T+B}=a_4\}....$$
$$....\cap\{\overline{X+R+B}=a_5\}\cap\{\overline{Y+R+S}=a_6\}\Big\}\cap\{\overline{Y+T+V}=a_7\}\cap\{\overline{Y+W+A}=a_8\}\Big\}$$

$$= (1/m^8)\sum_{u,v,w,a,b,r,s,t} P\Big\{\{X=\overline{a_1-u-s}\}\cap\{X=\overline{a_2-u-v}\}\cap\{X=\overline{a_3-w-a}\}\cap\{X=\overline{a_4-t-b}\}....$$
$$....\cap\{X=\overline{a_5-r-b}\}\cap\{Y=\overline{a_6-r-s}\}\Big\}\cap\{Y=\overline{a_7-t-v}\}\cap\{Y=\overline{a_8-w-a}\}\Big\}$$

$$= (1/m^8)\sum_{u,v,w,a,b,r,s,t,\ \overline{a_1-u-s}=\overline{a_2-u-v}=\overline{a_3-w-a}=\overline{a_4-t-b}=\overline{a_5-r-b},\ \overline{a_6-r-s}=\overline{a_7-t-v}=\overline{a_8-w-a}} (1/m^2)$$

$$= (1/m^{10}) \sum_{u,v,w,a,b,r,s,t,\ \overline{a_1-s}=\overline{a_2-v},\ \overline{a_2-u-v}=\overline{a_3-w-a}=\overline{a_4-t-b}=\overline{a_5-r-b},\ \overline{a_6-r-s}=\overline{a_7-t-v}=\overline{a_8-w-a}} 1$$

$$= (1/m^{10}) \sum_{u,v,w,a,b,r,s,t,\ v=\overline{a_2-a_1+s},\ \overline{a_2-u-v}=\overline{a_3-w-a}=\overline{a_4-t-b}=\overline{a_5-r-b},\ \overline{a_6-r-s}=\overline{a_7-t-v}=\overline{a_8-w-a}} 1$$

$$= (1/m^{10}) \sum_{u,v,w,a,b,r,s,t,\ v=\overline{a_2-a_1+s},\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-w-a}=\overline{a_4-t-b}=\overline{a_5-r-b},\ \overline{a_6-r-s}=\overline{a_7-t-a_2+a_1-s}=\overline{a_8-w-a}} 1$$

$$= (1/m^{10}) \sum_{u,w,a,b,r,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-w-a}=\overline{a_4-t-b}=\overline{a_5-r-b},\ \overline{a_6-r-s}=\overline{a_7-t-a_2+a_1-s}=\overline{a_8-w-a}} 1$$

$$= (1/m^{10}) \sum_{u,w,a,b,r,s,t,\ a=\overline{a_8-a_6+r+s-w},\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-w-a}=\overline{a_4-t-b}=\overline{a_5-r-b},\ \overline{a_6-r-s}=\overline{a_7-t-a_2+a_1-s}} 1$$

$$= m^{-10} \sum_{u,w,a,b,r,s,t,\ a=\overline{a_8-a_6+r+s-w},\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-w-a_8+a_6-r-s+w}=\overline{a_4-t-b}=\overline{a_5-r-b},\ \overline{a_6-r-s}=\overline{a_7-t-a_2+a_1-s}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,r,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-w-a_8+a_6-r-s+w}=\overline{a_4-t-b}=\overline{a_5-r-b},\ \overline{a_6-r-s}=\overline{a_7-t-a_2+a_1-s}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,r,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-w-a_8+a_6-r-s+w}=\overline{a_4-t-b}=\overline{a_5-r-b},\ r=\overline{a_6-a_7+a_2-a_1+t}} 1$$

$$= m^{-10} \sum_{u,w,b,r,s,t} 1$$

$$\sum_{\overline{a_2-u-a_2+a_1-s}=\overline{a_3-w-a_8+a_6-s+w-a_6+a_7-a_2+a_1-t}=\overline{a_4-t-b}=\overline{a_5-b-a_6+a_7-a_2+a_1-t},\ r=\overline{a_6-a_7+a_2-a_1+t}} 1$$

$$= m^{-10} \sum_{u,w,b,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-w-a_8+a_6-s+w-a_6+a_7-a_2+a_1-t}=\overline{a_4-t-b}=\overline{a_5-b-a_6+a_7-a_2+a_1-t}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-a_8-s+a_7-a_2+a_1-t}=\overline{a_4-t-b}=\overline{a_5-b-a_6+a_7-a_2+a_1-t}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-a_8-s+a_7-a_2+a_1-t}=\overline{a_4-t-b},\ \overline{a_4-t-b}=\overline{a_5-b-a_6+a_7-a_2+a_1-t}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-a_8-s+a_7-a_2+a_1-t}=\overline{a_4-t-b},\ \overline{a_5-a_4-a_6+a_7-a_2+a_1}=0} 1$$

$= 0$ si $\overline{a_5 - a_4 - a_6 + a_7 - a_2 + a_1} \neq 0$,

et si $\overline{a_5 - a_4 - a_6 + a_7 - a_2 + a_1} = 0$,

$$= (1/m^{10}) \sum_{u,w,b,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-a_8-s+a_7-a_2+a_1-t}=\overline{a_4-t-b}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,s,t,\ \overline{a_2-u-a_2+a_1-s}=\overline{a_3-a_8-s+a_7-a_2+a_1-t},\ \overline{a_3-a_8-s+a_7-a_2+a_1}=\overline{a_4-b}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,s,t,\ \overline{a_1-u}=\overline{a_3-a_8+a_7-a_2+a_1-t},\ s=\overline{a_3-a_8+a_7-a_2+a_1-a_4+b}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,t,\ \overline{a_1-u}=\overline{a_3-a_8+a_7-a_2+a_1-t}} 1$$

$$= (1/m^{10}) \sum_{u,w,b,t,\ t=\overline{a_3-a_8+a_7-a_2+u}} 1$$

$$= (1/m^{10}) \sum_{u,w,b} 1 = (1/m^7) \ . \blacksquare$$

Corollaire 7.2 *Si on peut écrire* $(X_1, ..., X_8)$ *sous la forme*

$$(X_1, ..., X_8)$$

$$= \left(\overline{X+U+S}, \overline{X+U+V}, \overline{X+W+A}, \overline{X+T+B}, \overline{X+R+B}, \overline{Y+R+S}, \overline{Y+T+V}, \overline{Y+W+A} \right)$$

Alors $\overline{X_1 - X_2 - X_4 + X_5 - X_6 + X_7} = 0$,

Proposition 7.3 *On suppose que*
$(X, Y, Z, Q, R, S, T, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, RR, SS, XX, YY, ZZ)$ *est un vecteur aléatoire de loi uniforme. Alors*

$$P\Big\{ \{\overline{X+Y+Z+Q+R+S+T} = a_1\} \cap \{\overline{X+A+B+C+D+E+F} = a_2\}$$

$$\cap \{\overline{Y+A+G+H+I+J+K} = a_3\} \cap \{\overline{Z+B+G+L+M+N+O} = a_4\} ....$$

$$.... \cap \{\overline{Q+C+H+L+P+XX+YY} = a_5\} \cap \{\overline{R+D+I+M+P+ZZ+RR} = a_6\}$$

$$\cap \{\overline{S+E+J+N+XX+ZZ+SS} = a_7\} \cap \{\overline{T+F+K+O+YY+RR+SS} = a_8\} \Big\}$$

$= (1/m^8)$ *si m=2m'+1,*

$= 0$ *si m=2m', et* $\overline{a_7 - a_8 - a_6 + a_5 - a_1 + a_2 + a_3 - a_4} \neq 2c$,

$= (2/m^8)$ *si m=2m', et* $\overline{a_7 - a_8 - a_6 + a_5 - a_1 + a_2 + a_3 - a_4} = 2c$.

Démonstration On a les égalités suivantes

$$P\Big\{ \{\overline{X+Y+Z+Q+R+S+T} = a_1\} \cap \{\overline{X+A+B+C+D+E+F} = a_2\}$$

$$\cap \{\overline{Y+A+G+H+I+J+K} = a_3\} \cap \{\overline{Z+B+G+L+M+N+O} = a_4\} ....$$

$$.... \cap \{\overline{Q+C+H+L+P+XX+YY} = a_5\} \cap \{\overline{R+D+I+M+P+ZZ+RR} = a_6\}$$

$$\cap \{\overline{S+E+J+N+XX+ZZ+SS} = a_7\} \cap \{\overline{T+F+K+O+YY+RR+SS} = a_8\} \Big\}$$

$$= \sum_{y,z,q,r,s,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,zz,rr}$$

$$P\Big\{ \{Y=y\} \cap \{Z=z\} \cap \{Q=q\} \cap \{R=r\} \cap \{S=s\} \cap \{T=t\} \cap \{A=a\} \cap \{B=b\}$$

$$\cap \{C=c\} \cap \{D=d\} \cap \{E=e\} \cap \{F=f\} \cap \{H=h\} \cap \{I=i\} \cap \{J=j\} \cap \{K=k\} \cap \{L=l\} \cap \{M=m\}$$

$$\cap\{N=n\}\cap\{O=o\}\cap\{XX=xx\}\cap\{YY=yy\}\cap\{ZZ=zz\}\cap\{RR=rr\}$$
$$\cap\{\overline{X+Y+Z+Q+R+S+T}=a_1\}\cap\{\overline{X+A+B+C+D+E+F}=a_2\}$$
$$\cap\{\overline{Y+A+G+H+I+J+K}=a_3\}\cap\{\overline{Z+B+G+L+M+N+O}=a_4\}....$$
$$....\cap\{\overline{Q+C+H+L+P+XX+YY}=a_5\}\cap\{\overline{R+D+I+M+P+ZZ+RR}=a_6\}$$
$$\cap\{\overline{S+E+J+N+XX+ZZ+SS}=a_7\}\cap\{\overline{T+F+K+O+YY+RR+SS}=a_8\}\Big\}$$

$$=(1/m^{24})\sum_{y,z,q,r,s,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,zz,rr}$$
$$P\Big\{\{X=\overline{a_1-y-z-q-r-s-t}\}\cap\{X=\overline{a_2-a-b-c-d-e-f}\}$$
$$\cap\{G=\overline{a_3-y-a-h-i-j-k}\}\cap\{G=\overline{a_4-z-b-l-m-n-o}\}....$$
$$....\cap\{P=\overline{a_5-q-c-h-l-xx-yy}\}\cap\{P=\overline{a_6-r-d-i-m-zz-rr}\}$$
$$\cap\{SS=\overline{a_7-s-e-j-n-xx-zz}\}\cap\{SS=\overline{a_8-t-f-k-o-yy-rr}\}\Big\}$$

$$=(1/m^{24})\sum_{y,z,q,r,s,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,zz,rr,\ \overline{a_1-y-z-q-r-s-t}=\overline{a_2-a-b-c-d-e-f}}$$
$$\sum_{\overline{a_3-y-a-h-i-j-k}=\overline{a_4-z-b-l-m-n-o}}$$
$$\sum_{\overline{a_5-q-c-h-l-xx-yy}=\overline{a_6-r-d-i-m-zz-rr}}$$
$$\sum_{\overline{a_7-s-e-j-n-xx-zz}=\overline{a_8-t-f-k-o-yy-rr}}(1/m^4)$$

$$=(1/m^{28})\sum_{y,z,q,r,s,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,zz,rr}$$
$$\sum_{y=\overline{a_1-a_2+a+b+c+d+e+f-z-q-r-s-t}}$$
$$\sum_{y=\overline{a_3-a_4+z+b+l+m+n+o-a-h-i-j-k}}$$
$$\sum_{zz=\overline{a_6-a_5+q+c+h+l+xx+yy-r-d-i-m-rr}}$$
$$\sum_{zz=\overline{a_7-a_8+t+f+k+o+yy+rr-s-e-j-n-xx}}1$$

$$= (1/m^{28}) \sum_{z,q,r,s,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,rr} \frac{\sum}{a_1-a_2+a+b+c+d+e+f-z-q-r-s-t=a_3-a_4+z+b+l+m+n+o-a-h-i-j-k} \frac{\sum}{a_6-a_5+q+c+h+l+xx+yy-r-d-i-m-rr=a_7-a_8+t+f+k+o+yy+rr-s-e-j-n-xx} 1$$

$$= (1/m^{28}) \sum_{z,q,r,s,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,rr} \frac{\sum}{a_1-a_2+c+d+e+f-q-r-s-t=a_3-a_4+2z+l+m+n+o-2a-h-i-j-k} \frac{\sum}{a_6-a_5+q+c+h+l-r-d-i-m=a_7-a_8+t+f+k+o+2rr-s-e-j-n-2xx} 1$$

$$= (1/m^{28}) \sum_{z,q,r,s,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,rr} \frac{\sum}{s=a_1-a_2-a_3+a_4-2z-l-m-n-o+2a+h+i+j+k+c+d+e+f-q-r-t} \frac{\sum}{s=a_7-a_8-a_6+a_5-q-c-h-l+r+d+i+m+t+f+k+o+2rr-e-j-n-2xx} 1$$

$$= (1/m^{28}) \sum_{z,q,r,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,rr} \frac{\sum}{a_1-a_2-a_3+a_4-2z-l-m-n-o+2a+h+i+j+k+c+d+e+f-q-r-t=a_7-a_8-a_6+a_5-q-c-h-l+r+d+i+m+t+f+k+o+2rr-e-j-n-2xx} 1$$

$$= (1/m^{28}) \sum_{z,q,r,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,rr} \frac{\sum}{a_1-a_2-a_3+a_4-2z-2m-2o+2a+2h+2j+2c+2e-2r-2t=a_7-a_8-a_6+a_5+2rr-2xx} 1$$

$$= (1/m^{28}) \sum_{z,q,r,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,rr} \frac{\sum}{2xx-2rr-2z-2m-2o+2a+2h+2j+2c+2e-2r-2t=a_7-a_8-a_6+a_5-a_1+a_2+a_3-a_4} 1$$

et par la même démonstration que précédemment,

si 2 est inversible (m=2m'+1)

$$= (1/m^{28}) \sum_{z,q,r,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy,rr} \sum_{rr=\overline{-2^{-1}(a_7-a_8-a_6+a_5-a_1+a_2+a_3-a_4)+xx-z-m-o+a+h+j+c+e-r-t}} 1$$

$$= (1/m^{28}) \sum_{z,q,r,t,a,b,c,d,e,f,h,i,j,k,l,m,n,o,xx,yy} 1 = (1/m^{28-20}) = (1/m^8) \ ,$$

$= 0$ si m=2m', et $\overline{a_7 - a_8 - a_6 + a_5 - a_1 + a_2 + a_3 - a_4} \neq 2c$.

$= (2/m^8)$ si m=2m', et $\overline{a_7 - a_8 - a_6 + a_5 - a_1 + a_2 + a_3 - a_4} = 2c$. ■

Proposition 7.4 *On suppose que* $(X, Y, Z, U, V, W, R, S, T, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O)$ *est un vecteur aléatoire de loi uniforme. Alors,*

$$P\Big\{\{\overline{X+Z+S+V+C+G+I+L} = a_1\} \cap \{\overline{X+Z+S+W+G+C+J+N} = a_2\}$$
$$\cap\{\overline{X+Z+T+A+D+H+I+O} = a_3\} \cap \{\overline{X+R+T+B+D+H+J+L} = a_4\}....$$
$$.... \cap \{\overline{X+R+U+B+E+G+I+M} = a_5\} \cap \{\overline{Y+R+T+W+E+G+K+N} = a_6\}$$
$$\cap\{\overline{Y+R+U+A+F+H+I+M} = a_7\} \cap \{\overline{Y+R+S+V+F+H+K+O} = a_8\}\Big\}$$
$$= 1/m^8.$$

Démonstration On a les égalités suivantes

$$P\Big\{\{\overline{X+Z+S+V+C+G+I+L} = a_1\} \cap \{\overline{X+Z+S+W+G+C+J+N} = a_2\}$$
$$\cap\{\overline{X+Z+T+A+D+H+I+O} = a_3\} \cap \{\overline{X+R+T+B+D+H+J+L} = a_4\}....$$
$$.... \cap \{\overline{X+R+U+B+E+G+I+M} = a_5\} \cap \{\overline{Y+R+T+W+E+G+K+N} = a_6\}$$
$$\cap\{\overline{Y+R+U+A+F+H+I+M} = a_7\} \cap \{\overline{Y+R+S+V+F+H+K+O} = a_8\}\Big\}$$

$$= \sum_{x,y,z,r,s,t,u,v,w,a,b,g,h,i,j,k,l,m,n,o}$$
$$P\Big\{\{X=x\} \cap \{Y=y\} \cap \{Z=z\} \cap \{R=r\} \cap \{S=s\} \cap \{T=t\} \cap \{U=u\}$$
$$\cap\{V=v\} \cap \{W=w\} \cap \{A=a\} \cap \{B=b\} \cap \{G=g\} \cap \{H=h\} \cap \{I=i\} \cap \{J=j\}$$
$$\cap\{K=k\} \cap \{L=l\} \cap \{M=m\} \cap \{N=n\} \cap \{O=o\}$$
$$\cap\{\overline{X+Z+S+V+C+G+I+L} = a_1\} \cap \{\overline{X+Z+S+W+G+C+J+N} = a_2\}$$

$$\cap\{\overline{X+Z+T+A+D+H+I+O}=a_3\}\cap\{\overline{X+R+T+B+D+H+J+L}=a_4\}....$$
$$....\cap\{\overline{X+R+U+B+E+G+I+M}=a_5\}\cap\{\overline{Y+R+T+W+E+G+K+N}=a_6\}$$
$$\cap\{\overline{Y+R+U+A+F+H+I+M}=a_7\}\cap\{\overline{Y+R+S+V+F+H+K+O}=a_8\}\Big\}$$

$$=(1/m^{20})\sum_{x,y,z,r,s,t,u,v,w,a,b,g,h,i,j,k,l,m,n,o}$$
$$P\Big\{\{C=\overline{a_1-x-z-s-v-g-i-l}\}\cap\{C=\overline{a_2-x-z-s-w-g-j-n}\}$$
$$\cap\{D=\overline{a_3-x-z-t-a-h-i-o}\}\cap\{D=\overline{a_4-x-r-t-b-h-j-l}\}....$$
$$....\cap\{E=\overline{a_5-x-r-u-b-g-i-m}\}\cap\{E=\overline{a_6-y-r-t-w-g-k-n}\}$$
$$\cap\{F=\overline{a_7-y-r-u-a-h-i-m}\}\cap\{F=\overline{a_8-y-r-s-v-h-k-o}\}\Big\}$$

$$=(1/m^{20})\sum_{x,y,z,r,s,t,u,v,w,a,b,g,h,i,j,k,l,m,n,o}$$
$$\sum_{\overline{a_1-x-z-s-v-g-i-l}=\overline{a_2-x-z-s-w-g-j-n}}$$
$$\sum_{\overline{a_3-x-z-t-a-h-i-o}=\overline{a_4-x-r-t-b-h-j-l}}$$
$$\sum_{\overline{a_5-x-r-u-b-g-i-m}=\overline{a_6-y-r-t-w-g-k-n}}$$
$$\sum_{\overline{a_7-y-r-u-a-h-i-m}=\overline{a_8-y-r-s-v-h-k-o}}(1/m^4)$$

$$=(1/m^{24})\sum_{x,y,z,r,s,t,u,v,w,a,b,g,h,i,j,k,l,m,n,o}$$
$$\sum_{l=\overline{a_1-a_2+x+z+s+w+g+j+n-x-z-s-v-g-i}}$$
$$\sum_{l=\overline{a_4-a_3+x+z+t+a+h+i+o-x-r-t-b-h-j}}$$
$$\sum_{m=\overline{a_5-a_6+y+r+t+w+g+k+n-x-r-u-b-g-i}}$$
$$\sum_{m=\overline{a_7-a_8+y+r+s+v+h+k+o-y-r-u-a-h-i}}1$$

$$= (1/m^{24}) \sum_{x,y,z,r,s,t,u,v,w,a,b,g,h,i,j,k,n,o} \sum_{\overline{a_1-a_2+x+z+s+w+g+j+n-x-z-s-v-g-i}=\overline{a_4-a_3+x+z+t+a+h+i+o-x-r-t-b-h-j}} \sum_{\overline{a_5-a_6+y+r+t+w+g+k+n-x-r-u-b-g-i}=\overline{a_7-a_8+y+r+s+v+h+k+o-y-r-u-a-h-i}} 1$$

$$= (1/m^{24}) \sum_{x,y,z,r,s,t,u,v,w,a,b,g,h,i,j,k,n,o} \sum_{\overline{a_1-a_2+w+j+n-v-i}=\overline{a_4-a_3+z+a+i+o-r-b-j}} \sum_{\overline{a_5-a_6+y+t+w+k+n-x-u-b-i}=\overline{a_7-a_8+s+v+k+o-u-a-i}} 1$$

$$= (1/m^{24}) \sum_{x,y,z,r,s,t,u,v,w,a,b,g,h,i,j,k,n,o} \sum_{\overline{a_1-a_2+w+2j+n-v-2i}=\overline{a_4-a_3+z+a+o-r-b}} \sum_{\overline{a_5-a_6+y+t+w+n-x-b}=\overline{a_7-a_8+s+v+o-a}} 1$$

$$= (1/m^{24}) \sum_{x,y,z,r,s,t,u,v,w,a,b,g,h,i,j,k,n,o} \sum_{w=\overline{a_4-a_3-a_1+a_2-2j-n+v+2i+z+a+o-r-b}} \sum_{w=\overline{a_7-a_8-a_5+a_6-y-t-n+x+b+s+v+o-a}} 1$$

$$= (1/m^{24}) \sum_{x,y,z,r,s,t,u,v,a,b,g,h,i,j,k,n,o} \sum_{\overline{a_4-a_3-a_1+a_2-2j-n+v+2i+z+a+o-r-b}=\overline{a_7-a_8-a_5+a_6-y-t-n+x+b+s+v+o-a}} 1$$

$$= (1/m^{24}) \sum_{x,y,z,r,s,t,u,v,a,b,g,h,i,j,k,n,o} \sum_{\overline{a_4-a_3-a_1+a_2-2j+2i+z+2a-r-2b}=\overline{a_7-a_8-a_5+a_6-y-t+x+s}} 1$$

$$= (1/m^{24}) \sum_{x,y,z,r,s,t,u,v,a,b,g,h,i,j,k,n,o} \sum_{t=a_7-a_8-a_5+a_6-a_4+a_3+a_1-a_2+2j-2i-z-2a+r+2b-y+x+s} 1$$

$$= (1/m^{24}) \sum_{x,y,z,r,s,u,v,a,b,g,h,i,j,k,n,o} 1$$

$$= (1/m^{24-16}) = (1/m^8) \ . \blacksquare$$

7.1 Conclusion

On voit que le type de dépendance ne change pas vraiment lorsque p augmente (tout en restant petit). Finalement, il y a des dépendances linéaires modulo m du type $\sum_s a_s X_s \equiv \alpha$ où $a_s \in \mathbb{N}$ et $|a_s| \leq p/2$.

On voit que si m est premier, certaines dépendances seront supprimées. Si m est de l'ordre de N, beaucoup de petites dépendances seront supprimées, ce qui rend a priori la recherche de dépendances plus difficile.

8 Recherches des dépendances particulières

On va étudier toutes les dépendances possibles de quadruplets et on verra que si on choisit certains paramêtres de manière précise, on a plus de chances de trouver des dépendances que dans le cas général où on choisit les dépendances au hasard.

On prend donc un vecteur aléatoire $(X_{j_1}^{d^1}, X_{j_2}^{d^2}, X_{j_3}^{d^3}, X_{j_4}^{d^4})$ où

$$X_{j_1}^{d^1} \equiv U * \begin{pmatrix} X_{1,j_1} \\ X_{2,\overline{j_1+d_2^1}} \\ \cdots\cdots \\ \cdots\cdots \\ X_{I,\overline{j_1+d_I^1}} \end{pmatrix}, \; X_{j_2}^{d^2} \equiv U * \begin{pmatrix} X_{1,j_2} \\ X_{2,\overline{j_2+d_2^2}} \\ \cdots\cdots \\ \cdots\cdots \\ X_{I,\overline{j_2+d_I^2}} \end{pmatrix},$$

$$X_{j_3}^{d^3} \equiv U * \begin{pmatrix} X_{1,j_3} \\ X_{2,\overline{j_3+d_2^3}} \\ \cdots\cdots \\ \cdots\cdots \\ X_{I,\overline{j_3+d_I^3}} \end{pmatrix}, \; X_{j_4}^{d^4} \equiv U * \begin{pmatrix} X_{1,j_4} \\ X_{2,\overline{j_4+d_2^4}} \\ \cdots\cdots \\ \cdots\cdots \\ X_{I,\overline{j_4+d_I^4}} \end{pmatrix},$$

et où $U = (1, 1....., 1)$ et où $\overline{j_s + d_i^s} \equiv j_s + d_i^s$ modulo N.

D'après le théorème 5, pour que $X_{j_1}^{d^1}$, $X_{j_2}^{d^2}$, $X_{j_3}^{d^3}$, $X_{j_4}^{d^4}$ soient dépendants, il faut que, pour tout i,

ou $X_{i,\overline{j_1+d_i^1}} = X_{i,\overline{j_2+d_i^2}}$ et $X_{i,\overline{j_3+d_i^3}} = X_{i,\overline{j_4+d_i^4}}$,
ou $X_{i,\overline{j_1+d_i^1}} = X_{i,\overline{j_3+d_i^3}}$ et $X_{i,\overline{j_2+d_i^2}} = X_{i,\overline{j_4+d_i^4}}$,
ou $X_{i,\overline{j_1+d_i^1}} = X_{i,\overline{j_4+d_i^4}}$ et $X_{i,\overline{j_3+d_i^3}} = X_{i,\overline{j_2+d_i^2}}$.

Donc, pour la première ligne,

ou $j_1 = j_2 = n$ et $j_3 = j_4 = n + D$,
ou $j_1 = j_3 = n$ et $j_2 = j_4 = n + D$,
ou $j_1 = j_4 = n$ et $j_3 = j_2 = n + D$.

Donc, il n'y a que 3 cas possibles.

Si les $X_{1,n}$ et les $X_{1,n+D}$ appartiennent à deux blocs différents, $d^1 \neq d^2$, où $D \neq 0$:
$(....X_{1,n}.....X_{1,n+D}.....)....(....X_{1,n}.....X_{1,n+D}....)$.

Si les $X_{1,n}$ et les $X_{1,n+D}$ appartiennent à trois blocs différents $d^1 \neq d^2$, $d^1 \neq d^3$ et $d^3 \neq d^2$, où $D \neq 0$, à l'ordre des blocs près :
$(....X_{1,n}....).....(....X_{1,n+D}.....)....(....X_{1,n}.....X_{1,n+D}....)$.

Si les $X_{1,n}$ et les $X_{1,n+D}$ appartiennent à quatre blocs différents $d^1 \neq d^2$, $d^1 \neq d^3$, $d^1 \neq d^4$, $d^2 \neq d^3$, $d^2 \neq d^4$, $d^3 \neq d^4$, à l'ordre des blocs près :
$(....X_{1,n}.....).....(.....X_{1,n}.....)....(....X_{1,n+D}.....).....(.....X_{1,n+D}....)$.

8.1 Cas de 2 blocs différents

On est donc dans le cas où la première ligne est
$(....X_{1,n}.....X_{1,n+D}.....)....(....X_{1,n}.....X_{1,n+D}....)$ avec $D \neq 0$.

8.1.1 Etudes des décalages d^1 et d^2

Il n'y a que deux décalages possibles par ligne : d_i^1 et d_i^2. Donc, la ligne i s'écrira sous la forme

$(....X_{i,\overline{n+d_i^1}}.....X_{i,\overline{n+D+d_i^1}}.....)....(....X_{i,\overline{n+d_i^2}}.....X_{i,\overline{n+D+d_i^2}}....)$.

D'apres le théorème 6, on doit avoir
ou $n+d_i^1 \equiv n+d_i^2$ et $n+D+d_i^1 \equiv n+D+d_i^2$,
ou $n+d_i^1 \equiv n+D+d_i^1$ et $n+d_i^2 \equiv n+D+d_i^2$,
ou $n+d_i^1 \equiv n+D+d_i^2$ et $n+D+d_i^1 \equiv n+d_i^2$.

Donc,
Ou $n+d_i^1 \equiv n+d_i^2$ et $n+D+d_i^1 \equiv n+D+d_i^2$,
Dans ce cas, $d_i^1 = d_i^2$.

Ou $n+d_i^1 \equiv n+D+d_i^1$ et $n+d_i^2 \equiv n+D+d_i^2$,
Dans ce cas, $D=0$, ce qui est contraire à l'hypothèse.

Ou $n+d_i^1 \equiv n+D+d_i^2$ et $n+D+d_i^1 \equiv n+d_i^2$,
Dans ce cas, $d_i^1 \equiv D+d_i^2$ et $D+d_i^1 \equiv d_i^2$.
Donc, $D+D+d_i^2 \equiv d_i^2$, Donc, $2D \equiv 0$.
Comme, il s'agit d'égalités modulo N, cela signifie que,
Ou D=0, ce qui est contraire à l'hypothèse.
Ou $D = N/2$. Il faut donc N pair.

Par exemple, avec I=2, en simplifiant $X_{1,i}$ en U_i et $X_{2,i}$ en Y_i, on peut avoir un quadruplet dépendant dans le cas suivant.
$(U_1, \mathbf{U_2}, \mathrm{U}_3, \mathrm{U}_4, \mathbf{U_5}, \mathrm{U}_6).,.(\mathrm{U}_1, \mathbf{U_2}, \mathrm{U}_3, \mathrm{U}_4, \mathbf{U_5}, \mathrm{U}_6)$
$(Y_4, .\mathbf{Y_5}, \mathrm{Y}_6, \mathrm{Y}_1, \mathbf{Y_2}, \mathrm{Y}_3).,.(\mathrm{Y}_1, \mathbf{Y_2}, \mathrm{Y}_3, \mathrm{Y}_4, \mathbf{Y_5}, \mathrm{Y}_6,)$

On se place donc dans le cas où N est pair. On ne peut choisir d^1 et d^2 que de la façon suivante.
$d^1 = (0, \underline{d_2, d_3, ..., d_I)}$ pour le premier bloc
$d^2 = (0, \overline{d_2 + \delta_2 N/2}\ ,\ \overline{d_3 + \delta_3 N/2, ..., d_I + \delta_I N/2})$ où $\delta_s = 0$ ou 1 .

Donc, pour chaque n, pour chaque ligne i, il y a $2N$ couples de décalages d_i^1 et d_i^2 possibles pour avoir des quadruplets dépendants. Donc, pour chaque n il y a $(2N)^{I-1}$ couples de décalages d^1 et d^2 possibles.
Il y a donc, $(2N)^{I-1}$ quadruplets dépendants possibles associés à un couple de blocs pour un n donné.

Maintenant, pour les deux blocs distincts, il y a au total $(N^{I-1})(N^{I-1}-1)$ décalages d^1 où d^2 possibles quand n est donné.

Donc, pour n donné, en choisissant au hasard deux blocs, la probabilité de trouver un quadruplet dépendant est au plus de

$$\frac{2^{I-1}N^{I-1}}{(N)^{I-1}(N^{I-1}-1)} = \frac{2^{I-1}}{(N^{I-1}-1)} \approx (2/N)^{I-1} \ .$$

Par exemple si $N = 2 * 10^6$ et $I = 21$, cette probabilité est de l'ordre de $\frac{1}{10^{120}}$.

Remarquons que l'on n'a pas une probabililité de l'ordre de $(1/N)^{2I}$ comme en section 3.1 car on suppose n connu et D=N/2, et on ne choisit donc pas j_1, j_2, j_3 et j_4 au hasard.

8.2 3 blocs différents

On est dans le cas où la première ligne est
$(....X_{1,n}....).....(....X_{1,n+D}.....)....(....X_{1,n}.....X_{1,n+D}....)$ avec $D \neq 0$.

Comme il n'y a que trois blocs pour chaque ligne, il n'y a que trois décalages possibles d_i^1, d_i^2 et d_i^3. Donc, on a pour la ligne i,
$(....X_{i,n+d_i^1}...).....(...X_{i,n+D+d_i^2}.....)....(....X_{i,n+d_i^3}.....X_{i,n+D+d_i^3}....)$

D'apres le théorème 6, les termes de chaque ligne doivent être égaux deux à deux. Donc,
ou $n+d_i^1 \equiv n+d_i^3$ et $n+D+d_i^2 \equiv n+D+d_i^3$,
ou $n+d_1^i \equiv n+D+d_i^2$ et $n+d_i^3 \equiv n+D+d_i^3$,
ou $n+d_i^1 \equiv n+D+d_i^3$ et $n+D+d_2^i \equiv n+d_i^3$.

Donc,
Ou $n+d_i^1 \equiv n+d_i^3$ et $n+D+d_i^2 \equiv n+D+d_i^3$. Dans ce cas, $d_i^1 = d_i^2 = d_i^3$.
Ou $n+d_i^1 \equiv n+D+d_i^2$ et $n+d_i^3 \equiv n+D+d_i^3$. Dans ce cas, $D=0$, ce qui est contraire aux hypothèses.
Ou $n+d_i^1 \equiv n+D+d_i^3$ et $n+D+d_i^2 \equiv n+d_i^3$. Dans ce cas, $d_i^1 \equiv D+d_i^3$ et $D+d_i^2 \equiv d_i^3$. Donc, $d_i^1 \equiv 2D+d_i^2$ et $d_i^3 \equiv D+d_i^2$.

Par exemple, avec D=-1, on a un quadruplet dépendant :
$(U_1, U_2, \mathbf{U_3}, \mathrm{U}_4, \mathrm{U}_5, \mathrm{U}_6).,.(\mathrm{U}_1, \mathbf{U_2}, \mathrm{U}_3, \mathrm{U}_4, \mathrm{U}_5, \mathrm{U}_6).,.(\mathrm{U}_1, \mathbf{U_2}, \mathbf{U_3}, \mathrm{U}_4, \mathrm{U}_5, \mathrm{U}_6)$
$(Y_3, .Y_4, .\mathbf{Y_5}, \mathrm{Y}_6, \mathrm{Y}_1, \mathrm{Y}_2).,.(\mathrm{Y}_5, \mathbf{Y_6}, \mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Y}_3, \mathrm{Y}_4).,.(\mathrm{Y}_4, \mathbf{Y_5}, \mathbf{Y_6}, \mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Y}_3)$

Probabilité de trouver une dépendance On veut connaitre la probabilité de trouver au hasard un quadruplet dépendant sachant que l'on connait n et D et les 3 arrangements de la première ligne (l'emplacement du bloc ayant deux éléments).

Donc, avec $\delta_s = 0$ ou 1, on ne peut choisir d^1 , d^2 et d^3 que de la façon suivante :
$d^1 = (0, d_2, d_3, ..., d_I)$ pour le premier bloc,
$d^2 = (0, \overline{d_2 - 2\delta_2 D}, \overline{d_3 - 2\delta_3 D}, ..., \overline{d_I - 2\delta_I D})$,
$d^3 = (0, \overline{d_2 - \delta_2 D}, \overline{d_3 - \delta_3 D}, ..., \overline{d_I - \delta_I D})$.

Quel est le nombre de quadruplets dépendants lorsque l'on choisit trois blocs au hasard?

On suppose n et D donnés.

Pour chaque ligne i, $d_i^1 = d_i$ peut prendre N valeurs. De plus, $d_i^2 = d_i - 2\delta_i D$ peut prendre 2 valeurs. Enfin, $d_i^3 = d_i - \delta_i D$ peut prendre 1 valeur, fonction de δ_i. Pour chaque ligne, il y a donc, 2N décalages d_i^1, d_i^2, et d_i^3 possibles.

Donc, il y a, au plus, $(2N)^{I-1}$ décalages d^1, d^2, et d^3 possibles qui fournissent des quadruplets dépendants.

Maintenant, pour n donné il y a au total $(N^{I-1})(N^{I-1}-1)(N^{I-1}-2)$ triplets de décalages possibles distincts.

Donc, la probabilité de trouver un quadruplet dépendant est au plus de

$$\frac{2^{I-1}N^{I-1}}{N^{I-1}(N^{I-1}-1)(N^{I-1}-2)}$$

$$= \frac{2^{I-1}}{(N^{I-1}-1)(N^{I-1}-2)} \approx 2^{I-1}/N^{2(I-1)} \ .$$

Par exemple si $N = 2^*10^6$ et $I = 21$, cette probabilité est de l'ordre de $\frac{1}{2^{20}10^{240}}$.

8.3 4 blocs différents

On peut toujours se ramener au cas où la première ligne, est
$(....X_{1,n}.....).....(.....X_{1,n}.....)....(....X_{1,n+D}.....).....(.....X_{1,N+D}....)$.

Donc, on a pour la ligne i,
$(....X_{i,\overline{n+d_i^1}}.....).....(.....X_{i,\overline{n+d_i^2}}.....)....(....X_{i,\overline{n+D+d_i^3}}.....).....(.....X_{i,\overline{N+D+d_i^4}}....)$.

Comme ces termes sont toujours égaux deux à deux, on a pour la ligne i,
ou $n+d_i^1 \equiv n+d_i^2$ et $n+D+d_i^3 \equiv n+D+d_i^4$,
ou $n+d_i^1 \equiv n+D+d_i^3$ et $n+d_i^2 \equiv n+D+d_i^4$,
ou $n+d_i^1 \equiv n+D+d_i^4$ et $n+d_i^2 \equiv n+D+d_i^3$.

Donc,
Si $n+d_i^1 = n+d_i^2$ et $n+D+d_i^3 = n+D+d_i^4$, alors, $d_i^1 = d_i^2$ et $d_i^3 = d_i^4$
Si $n+d_i^1 = n+D+d_i^3$ et $n+d_i^2 = n+D+d_i^4$, alors, $d_i^1 = D+d_i^3$ et $d_i^2 = D+d_i^4$,
Si $n+d_i^1 = n+D+d_i^4$ et $n+d_i^2 = n+D+d_i^3$, alors, $d_i^1 = D+d_i^4$ et $d_i^2 = D+d_i^3$.

Par exemple, avec D=2, et en simplifiant $X_{1,j}$ en $U_j, X_{2,j}$ en $Y_j, X_{3,j}$ en Z_j et $X_{4,j}$ en T_j ,

$(U_1, U_2, \mathbf{U_3}, \mathrm{U}_4, \mathrm{U}_5, \mathrm{U}_6), (\mathrm{U}_1, \mathrm{U}_2, \mathbf{U_3}, \mathrm{U}_4, \mathrm{U}_5, \mathrm{U}_6), (\mathrm{U}_1, \mathrm{U}_2, \mathrm{U}_3, \mathrm{U}_4, \mathbf{U_5}, \mathrm{U}_6), (\mathrm{U}_1, \mathrm{U}_2, \mathrm{U}_3, \mathrm{U}_4, \mathbf{U_5}, \mathrm{U}_6)$
$(Y_5, Y_6, \mathbf{Y_1}, \mathrm{Y}_2, \mathrm{Y}_3, \mathrm{Y}_4), .(\mathrm{Y}_5, \mathrm{Y}_6, \mathbf{Y_1}, \mathrm{Y}_2, \mathrm{Y}_3, \mathrm{Y}_4), .(\mathrm{Y}_4, \mathrm{Y}_5, \mathrm{Y}_6, \mathrm{Y}_1, \mathbf{Y_2}, \mathrm{Y}_3), .(\mathrm{Y}_4, \mathrm{Y}_5, \mathrm{Y}_6, \mathrm{Y}_1, \mathbf{Y_2}, \mathrm{Y}_3)$
$(Z_6, Z_1, \mathbf{Z_2}, \mathrm{Z}_3, \mathrm{Z}_4, \mathrm{Z}_5)., ...(\mathrm{Z}_5, \mathrm{Z}_6, \mathbf{Z_1}, \mathrm{Z}_2, \mathrm{Z}_3, \mathrm{Z}_4)., ...(\mathrm{Z}_4, \mathrm{Z}_5, \mathrm{Z}_6, \mathrm{Z}_1, \mathbf{Z_2}, \mathrm{Z}_3)., ...(\mathrm{Z}_3, \mathrm{Z}_4, \mathrm{Z}_5, \mathrm{Z}_6, \mathbf{Z_1}, \mathrm{Z}_2,)$
$(T_4, T_5, \mathbf{T_6}, \mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3,), .(\mathrm{T}_1, \mathrm{T}_2, \mathbf{T_3}, \mathrm{T}_4, \mathrm{T}_5, \mathrm{T}_6), .(\mathrm{T}_5, \mathrm{T}_6, \mathrm{T}_1, \mathrm{T}_2, \mathbf{T_3}, \mathrm{T}_4,)., (\mathrm{T}_2, \mathrm{T}_3, \mathrm{T}_4, \mathrm{T}_5, \mathbf{T_6}, \mathrm{T}_1)$

Probabilité de trouver une dépendance On veut connaitre la probabilité de trouver au hasard un quadruplet dépendant sachant que l'on connait n et D donné ainsi que l'arrangement de la première ligne (parmis les 3 possibles). Alors, pour pour chaque ligne $i > 1$, il y a

Ou N façons de choisir d_i^1 et N façons au plus de choisir d_i^3 : $d_i^1 = d_i^2$ et $d_i^3 = d_i^4$,

Ou N façons de choisir $d_1^i = D + d_3^i$ et N façons au plus de choisir $d_2^i = D + d_4^i$: $d_1^i = D + d_3^i$ et $d_2^i = D + d_4^i$,

Ou N façons de choisir $d_1^i = D + d_4^i$ et N façons au plus de choisir $d_2^i = D + d_3^i$: $d_1^i = D + d_4^i$ et $d_2^i = D + d_3^i$.

C'est a dire $3N^2$ possibilités de choix au plus pour chaque ligne.

Donc au total, il y a $[3N^2]^{I-1}$ possibilités pour les quadruplets dépendants.

Quel est le nombre de quadruplets lorsque l'on choisi 4 blocs au hasard? Comme précédemment, il y a au total $N^{I-1}(N^{I-1}-1)(N^{I-1}-2)(N^{I-1}-3)$ possibilités. Donc, la probabilité de trouver un quadruplet dépendant est au plus de

$$\frac{3^{I-1}N^{I-1}}{(N^{I-1}-1)(N^{I-1}-2)(N^{I-1}-3)} \approx (3/N)^{2(I-1)} \ .$$

8.4 Conclusion

On voit donc que si on veut avoir le plus de chances possible de trouver une dépendance, il faut choisir la dépendance entre deux blocs dans le cas où N est pair. On peut supprimer ce cas en choisissant N premier ou en ne prenant que les N/2 premiers termes de chaque bloc. On va s'apercevoir que c'est général.

9 Autres Dépendances

On a donc vu dans la section précédente que dans le cas où p est petit, le cas où il y a le plus de chances de trouver une dépendance, c'est celui de deux blocs différents lorsque N est pair. Mais il facile de voir qu'on peut généraliser ce résultat : c'est ce que l'on va voir maintenant.

9.1 Cas $N = 2N'$

On suppose donc $N = 2N'$ et on va étudier la p-dépendance sur deux blocs qui se suivent. Pour trouver une dépendance, pour chaque ligne i, on peut imposer qu'elle soit définie pour la première ligne par $(X_{1,2}, X_{1,4}, X_{1,6}, ..., X_{1,2N'})$ et $(X_{1,2}, X_{1,4}, X_{1,6}, ..., X_{1,2N'})$.

Pour la ligne i on a
$(X_{i,\overline{2+d_i^1}}, X_{i,\overline{4+d_i^1}}, X_{i,\overline{6+d_i^1}}, ..., X_{i,\overline{2N'+d_i^1}})$ et $(X_{i,\overline{2+d_i^2}}, X_{i,\overline{4+d_i^2}}, X_{i,\overline{6+d_i^2}}, ..., X_{i,\overline{2N'+d_i^2}})$.

Donc si $d_i^2 \equiv d_i^1 + 2k_i$, $k_i \in \{0, 1, ..., N'\}$,
$\{X_{1,\overline{2+d_i^1}}, X_{1,\overline{4+d_i^1}}, X_{1,\overline{6+d_i^1}}, ..., X_{1,\overline{2N'+d_i^1}}\} = \{X_{1,\overline{2+d_i^2}}, X_{1,\overline{4+d_i^2}}, X_{1,\overline{6+d_i^2}}, ..., X_{1,\overline{2N'+d_i^2}}\}$.
Donc il y peut y avoir dépendance. Donc, pour chaque ligne il y a $N * (2N')/2 = N^2/2$ décalages d_i^1 et d_i^2 qui conservent cette éventuelle dépendance.
Comme il y a $N^{I-1}(N^{I-1} - 1)$ couples de décalages possibles, on aura donc environ une chance sur 2^{I-1} d'obtenir au hasard une telle dépendance.
Par exemple si I=20, on a une chance sur 2^{20} de trouver une dépendance, ce qui est une chance beaucoup plus grande que pour les dépendances de quadruplets que nous venons d'étudier.

Maintenant, est ce qu'une telle dépendance serait utile pour un cryptanalyste? En fait, elle ne servirait pas à grand chose : au mieux, elle permettrait de calculer les 2 derniers termes $X_{2^p}^{d^2}$ connaissant les $x_n^{d^1}$ et $x_{n'}^{d^2}$, $n' \leq N - 1$.
De même si on avait une dépendance pour des élément répartis de 4 en 4, cela permettrait de connaitre les 4 derniers éléments du deuxième bloc et ainsi de suite.
La dépendance linéaire associée sera $X_2^{d^1} + X_4^{d^1} + ... + X_{2N'}^{d^1} \equiv X_2^{d^2} + X_4^{d^2} + ... + X_{2N'}^{d^2}$. Ce sera la généralisation de la dépendance élémentaire étudiée en section 4.1.2.
Cela ne servirait sans doute donc pas à grand chose de connaitre ces dépendances pour un cryptanalyste.
Mais il est tout de même mieux de supprimer ces dépendances : c'est pourquoi on peut choisir N premier ou ne prendre à chaque bloc que les $\lfloor N/2 \rfloor$ premiers termes.

Maintenant, on a un résultat de ce type si au lieu de prendre un X_j^d sur 2, on prend un X_j^d sur 3 ou sur 4, etc. Pour avoir une idée de ce qui se passe, on va d'abord étudier l'exemple suivant.

On se place dans le cas où la premiere ligne est
$(....X_{1,n}.....X_{1,n+D}.....X_{1,n+D'}.....)....(....X_{1,n}.....X_{1,n+D}....X_{1,n+D'}.....)$.

Donc,
ou $n + d_i^1 = n + d_i^2$ et $n + D + d_i^1 = n + D + d_i^2$ et $n + D' + d_i^1 = n + D' + d_i^2$,
ou $n + d_i^1 = n + d_i^2$ et $n + D + d_i^1 = n + D' + d_i^2$ et $n + D' + d_i^1 = n + D + d_i^2$,
ou $n + d_i^1 = n + D + d_i^2$ et $n + D + d_i^1 = n + d_i^2$ et $n + D' + d_i^1 = n + D' + d_i^2$,
ou $n + d_i^1 = n + D + d_i^2$ et $n + D + d_i^1 = n + D' + d_i^2$ et $n + D' + d_i^1 = n + d_i^2$,
ou $n + d_i^1 = n + D' + d_i^2$ et $n + D + d_i^1 = n + d_i^2$ et $n + D' + d_i^1 = n + D + d_i^2$,
ou $n + d_i^1 = n + D' + d_i^2$ et $n + D + d_i^1 = n + D + d_i^2$ et $n + D' + d_i^1 = n + d_i^2$.

Si $n + d_i^1 = n + d_i^2$ et $n + D + d_i^1 = n + D + d_i^2$ et $n + D' + d_i^1 = n + D' + d_i^2$,

$d_i^1 = d_i^2$.

Si $n + d_i^1 = n + d_i^2$ et $n + D + d_i^1 = n + D' + d_i^2$ et $n + D' + d_i^1 = n + D + d_i^2$,
$d_1^i = d_i^2$ et D=D' : c'est impossible.

Si $n + d_i^1 = n + D + d_i^2$ et $n + D + d_i^1 = n + d_i^2$ et $n + D' + d_i^1 = n + D' + d_i^2$,
$d_i^1 = d_i^2$ et D=0 : c'est impossible.

Si $n + d_i^1 = n + D + d_i^2$ et $n + D + d_i^1 = n + D' + d_i^2$ et $n + D' + d_i^1 = n + d_i^2$,
$d_i^1 = D + d_i^2$, $D + d_i^1 = D' + d_i^2$ et $D' + d_i^1 = d_i^2$,
Donc, $2D + d_i^2 = D' + d_i^2$. Donc, $2D = D'$.
De plus, $D' + D + d_i^2 = d_i^2$. Donc, $D' + D = 0$. Donc $3D = 0$. Donc D=N/3 ou 2N/3. Donc D'=2N/3 ou N/3, ce qui est impossible si N est premier.

Si $n + d_i^1 = n + D' + d_i^2$ et $n + D + d_i^1 = n + d_i^2$ et $n + D' + d_i^1 = n + D + d_i^2$. C'est le même que le cas précédent en intervertissant D et D'.

Si $n + d_i^1 = n + D' + d_i^2$ et $n + D + d_i^1 = n + D + d_i^2$ et $n + D' + d_i^1 = n + d_i^2$.
Donc, $d_1^i = d_2^i$ et D'=0 : c'est impossible.

Donc si N est premier ou si on ne prend que les N/2 premiers termes de chaque bloc, on ne peut trouver de telles dépendances.

9.2 Autres dépendances

Mais pour en être plus sûr, il faut étudier les dépendances entre p termes lorsque p=6,8,10,12, pour voir de quelle façon peuvent se créer les dépendances.

Maintenant comme le problème est d'avoir une chance raisonable de trouver une dépendance, le mieux est que les termes dépendants soient groupés dans le moins de blocs possibles.

9.2.1 Dépendance sur 3 blocs : cas 1

On va d'abord étudier les dépendances qui existent lorsque on prend trois blocs. On va voir qu'il y a peu de chances de tomber au hasard sur une dépendance surtout si N est premier ou si on supprime les N/2 derniers termes.

On se place dans le cas ou, avec $D, D' \neq 0$, la première ligne est
$(....X_{1,n}....).....(....X_{1,n+D}.....X_{1,n+D'}....)....(....X_{1,n}.....X_{1,n+D}.....X_{1,n+D'}.....)$.

Donc, on a pour la ligne i,
$(....X_{i,n+d_i^1}...)....(...X_{i,n+D+d_i^2}....X_{i,n+D'+d_i^2}....)....(...X_{i,n+d_i^3}...X_{i,n+D+d_i^3}...X_{i,n+D'+d_i^3}....)$.

D'apres le théorème 6, pour qu'il y aie dépendance, les termes de chaques lignes doivent être egaux deux à deux. Donc,
ou $n + d_i^1 \equiv n + d_i^3$, $n + D + d_i^2 \equiv n + D + d_i^3$ et $n + D' + d_i^2 \equiv n + D' + d_i^3$,
ou $n + d_i^1 \equiv n + d_i^3$, $n + D + d_i^2 \equiv n + D' + d_i^3$ et $n + D' + d_i^2 \equiv n + D + d_i^3$,
ou $n + d_i^1 \equiv n + D + d_i^3$ et $n + D + d_i^2 \equiv n + d_i^3$ et $n + D' + d_i^2 \equiv n + D' + d_i^3$.
ou $n + d_i^1 \equiv n + D + d_i^3$ et $n + D + d_i^2 \equiv n + D' + d_i^3$ et $n + D' + d_i^2 \equiv n + d_i^3$.
ou $n + d_i^1 \equiv n + D' + d_i^3$ et $n + D' + d_i^2 \equiv n + d_i^3$ et $n + D + d_i^2 \equiv n + D + d_i^3$.
ou $n + d_i^1 \equiv n + D' + d_i^3$ et $n + D' + d_i^2 \equiv n + D + d_i^3$ et $n + D + d_i^2 \equiv n + d_i^3$.

Donc,

Cas 1) $d_i^1 = d_i^3 = d_i^2$.
Cas 2) $d_i^1 \equiv d_i^3$ et $D + d_i^2 \equiv D' + d_i^3$ et $D' + d_i^2 \equiv D + d_i^3$. Donc D-D'=D'-D. Donc, D-D'=N/2.
Cas 3) $d_i^1 \equiv D + d_i^3$ et $D + d_2^i \equiv d_i^3$ et $D' + d_i^2 \equiv D' + d_i^3$. Donc D=0 : impossible.
Cas 4) $d_i^1 \equiv D + d_i^3$ et $D + d_2^i \equiv D' + d_i^3$ et $D' + d_i^2 \equiv d_i^3$. Donc D-D'=D'. Donc D=2D'.
Cas 5) $d_i^1 \equiv D' + d_i^3$ et $D' + d_2^i \equiv d_i^3$ et $D + d_i^2 \equiv D + d_i^3$. Donc D'=0 : impossible.
Cas 6) $d_i^1 \equiv D' + d_i^3$ et $D + d_2^i \equiv d_i^3$ et $D' + d_i^2 \equiv D + d_i^3$. Donc D=D'-D. Donc D'=2D.

Donc suivant le cas, il n'y aura que une possibilité au plus parmis celles-ci : D-D'=N/2, D=2D', D'=2D.

Par exemple, si D=N/2+D' et D=2D', alors, D'=N/2, donc D=0.

Ou bien si 2D= D' et D=2D', alors, D=-D' et 4D=D. Donc par exemple D=N/3 et D'=2N/3, ce qui est impossible si N est premier.

Quel est le nombre de quadruplets dépendants? D'après ce qui précède, si N est premier, suivant le choix de D et D', il n'y a que 2 valeurs possibles à chaque ligne pour d_i^1, d_i^2 et d_i^3. Donc, pour chaque ligne, il y a donc, 2N décalages d^1, d^2, et d^3 possibles.

Donc, il y a, au plus, $(2N)^{I-1}$ décalages d^1, d^2, et d^3 possibles qui fournissent des 6-uplets dépendants.

Donc, en choisissant bien D et D', la probabilité de trouver un 6-uplet dépendant est

$$\frac{2^{I-1}N^{I-1}}{N^{I-1}(N^{I-1}-1)(N^{I-1}-2)} \approx 2^{I-1}/N^{2(I-1)} \ .$$

C'est la même que celle de la section 8.2. Il n'y a donc pas plus de chances de trouver une dépendance dans ce cas.

9.2.2 Dépendance sur 3 blocs : cas 2

On se place dans le cas où, avec $D, D' \neq 0$, la première ligne est du type
$(....X_{1,n}.....X_{1,n+D}...).....(....X_{1,n}....X_{1,n+D'}.....)....(....X_{1,n+D}.....X_{1,n+D'}.....)$.

Donc, on a pour la ligne i,
$(...X_{i,n+d_i^1}....X_{i,n+D+d_i^1}...)....(...X_{i,n+d_i^2}....X_{i,n+D'+d_i^2}...)....(...X_{i,n+D+d_i^3}....X_{i,n+D'+d_i^3}....)$.

Donc, pour avoir une dépendance il faut que
ou $n + d_i^1 \equiv n + d_i^2$, $n + D + d_i^1 \equiv n + D + d_i^3$ et $n + D' + d_i^2 \equiv n + D' + d_i^3$,
ou $n + d_i^1 \equiv n + d_i^2$, $n + D + d_i^1 \equiv n + D' + d_i^3$ et $n + D' + d_i^2 \equiv n + D + d_i^3$,
ou $n + d_i^1 \equiv n + D' + d_i^2$, $n + D + d_i^1 \equiv n + D + d_i^3$ et $n + d_i^2 \equiv n + D' + d_i^3$,
ou $n + d_i^1 \equiv n + D' + d_i^2$, $n + D + d_i^1 \equiv n + D' + d_i^3$ et $n + d_i^2 \equiv n + D + d_i^3$,

ou $n + d_i^1 \equiv n + D + d_i^3$, $n + D + d_i^1 \equiv n + D' + d_i^2$ et $n + d_i^2 \equiv n + D' + d_i^3$,
ou $n + d_i^1 \equiv n + D + d_i^3$, $n + D + d_i^1 \equiv n + d_i^2$ et $n + D' + d_i^2 \equiv n + D' + d_i^3$,
ou $n + d_i^1 \equiv n + D' + d_i^3$, $n + D + d_i^1 \equiv n + D' + d_i^2$ et $n + d_i^2 \equiv n + D + d_i^3$,
ou $n + d_i^1 \equiv n + D' + d_i^3$, $n + D + d_i^1 \equiv n + d_i^2$ et $n + D' + d_i^2 \equiv n + D + d_i^3$,

Donc,

Cas 1) $d_i^1 = d_i^3 = d_i^2$.

Cas 2) $d_i^1 \equiv d_i^2$, $D + d_i^1 \equiv D' + d_i^3$ et $D' + d_i^2 \equiv D + d_i^3$. Donc D-D'=D'-D. Donc D-D'=N/2.

Cas 3) $d_i^1 \equiv D' + d_i^2$, $D + d_i^1 \equiv D + d_i^3$ et $d_i^2 \equiv D' + d_i^3$. Donc, $d_i^1 = d_i^3$. Donc, $d_i^3 \equiv D' + d_i^2$ et $d_i^2 \equiv D' + d_i^3$. Donc, $d_i^3 - D' \equiv D' + d_i^3$. Donc 2D'=0. Donc D'=N/2.

Cas 4) $d_i^1 \equiv D' + d_i^2$, $D + d_i^1 \equiv D' + d_i^3$ et $d_i^2 \equiv D + d_i^3$. Donc, $d_i^1 \equiv D' + D + d_i^3$ et $d_i^1 \equiv D' - D + d_i^3$. Donc $2D \equiv 0$. Donc D=N/2.

Cas 5) $d_i^1 \equiv D + d_i^3$, $D + d_i^1 \equiv D' + d_i^2$ et $d_i^2 \equiv D' + d_i^3$. Donc, $D + D + d_i^3 \equiv D' + D' + d_i^3$. Donc $2D \equiv 2D'$. Donc, D=D'+N/2.

Cas 6) $d_i^1 \equiv D + d_i^3$, $D + d_i^1 \equiv d_i^2$ et $D' + d_i^2 \equiv D' + d_i^3$. Donc, $d_i^2 \equiv d_i^3$. Donc, $D + D + d_i^3 \equiv d_i^3$. Donc 2D=0. Donc D=N/2.

Cas 7) $d_i^1 \equiv D' + d_i^3$, $D + d_i^1 \equiv D' + d_i^2$ et $d_i^2 \equiv D + d_i^3$. Donc, $D + (D' + d_i^3) \equiv D' + (D + d_i^3)$. Donc D'+D=D+D'. C'est toujours possible.

Cas 8) $d_i^1 \equiv D' + d_i^3$, $D + d_i^1 \equiv d_i^2$ et $D' + d_i^2 \equiv D + d_i^3$. Donc, $D + (D' + d_i^3) \equiv d_i^2$. Donc, $D' + (D + D' + d_i^3) \equiv D + d_i^3$. Donc 2D'=0. Donc D'=N/2.

Donc si N est premier ou si on ne prend que les N/2 premiers termes de chaque bloc, la seule dépendance aura lieu quand $d_i^1 = d_i^3 = d_i^2$ ou quand $d_i^1 \equiv D' + d_i^3$, $D + d_i^1 \equiv D' + d_i^2$ et $d_i^2 \equiv D + d_i^3$.

Il y a donc au plus 2N décalages possibles par ligne et donc $(2N)^{I-1}$ possibilités pour toutes les lignes si n, D, et D' sont donnés. Maintenant, il y a $N^{I-1}(N^{I-1}-1)(N^{I-1}-2)$ possibilités. Donc on a une probabilité de l'ordre de $2/N^{2(I-1)}$ de trouver une telle dépendance qui est donc beaucoup plus petite que celle du cas de deux blocs.

9.2.3 Dépendance sur 3 blocs : cas 3

On étudie le cas d'une dépendance par triplets pour la première ligne.

$(....X_{1,n}.....X_{1,n+D}...)....(....X_{1,n}.....X_{1,n+D}.....)....(....X_{1,n}.....X_{1,n+D}.....)$.

Donc pour la ligne i avec $D \neq 0$, étudions d'abord les dépendances par couple
$(....X_{i,n+d_i^1}.....X_{i,n+D+d_i^1}...)....(....X_{i,n+d_i^2}...X_{i,n+D+d_i^2}....)....(...X_{i,n+d_i^3}...X_{i,n+D+d_i^3}...)$.

Donc,
ou $n + d_i^1 = n + d_i^2$ et $n + D + d_i^1 = n + D + d_i^3$ et $n + D + d_i^2 = n + d_i^3$,
ou $n + d_i^1 = n + d_i^2$ et $n + D + d_i^1 = n + d_i^3$ et $n + D + d_i^2 = n + D + d_i^3$,
ou $n + d_i^1 = n + D + d_i^2$ et $n + D + d_i^1 = n + d_i^3$ et $n + d_i^2 = n + D + d_i^3$,
ou $n + d_i^1 = n + D + d_i^2$ et $n + D + d_i^1 = n + D + d_i^3$ et $n + d_i^2 = n + d_i^3$,
OU
ou $n + d_i^1 = n + d_i^3$ et $n + D + d_i^1 = n + d_i^2$ et $n + D + d_i^2 = n + D + d_i^3$,
ou $n + d_i^1 = n + d_i^3$ et $n + D + d_i^1 = n + D + d_i^2$ et $n + d_i^2 = n + D + d_i^3$,
ou $n + d_i^1 = n + D + d_i^3$ et $n + D + d_i^1 = n + d_i^2$ et $n + D + d_i^2 = n + d_i^3$,
ou $n + d_i^1 = n + D + d_i^3$ et $n + D + d_i^1 = n + D + d_i^2$ et $n + d_i^2 = n + d_i^3$,

Donc,

Si $n + d_i^1 = n + d_i^2$ et $n + D + d_i^1 = n + D + d_i^3$ et $n + D + d_i^2 = n + d_i^3$.
Alors, $d_i^1 = d_i^2$ et $d_i^1 = d_i^3$ et $D + d_i^2 = d_i^3$. Donc, D=0 : c'est impossible.

Si $n + d_i^1 = n + d_i^2$ et $n + D + d_i^1 = n + d_i^3$ et $n + D + d_i^2 = n + D + d_i^3$.
Alors, $d_i^1 = d_i^2$ et $D + d_i^1 = d_i^3$ et $d_i^2 = d_i^3$. Donc, D=0 : c'est impossible.

Si $n + d_i^1 = n + D + d_i^2$ et $n + D + d_i^1 = n + d_i^3$ et $n + d_i^2 = n + D + d_i^3$.
Alors, $d_i^1 = D + d_i^2$ et $D + d_i^1 = d_i^3$ et $d_i^2 = D + d_i^3$.
Donc, $D + D + d_i^2 = d_i^2 - D$. Donc 3D=0. Donc D=aN/3 ou a=1 ou 2.
Donc $d_i^1 = aN/3 + d_i^2$ et $d_i^2 = aN/3 + d_i^3$. Donc $d_i^1 = 2aN/3 + d_i^3$.
Si a=1, $d_i^1 = N/3 + d_i^2$ et $d_i^1 = 2N/3 + d_i^3$.
Si a=2, $d_i^1 = 2N/3 + d_i^2$ et $d_i^1 = N/3 + d_i^3$.

Si $n + d_i^1 = n + D + d_i^2$ et $n + D + d_i^1 = n + D + d_i^3$ et $n + d_i^2 = n + d_i^3$.
Alors, $d_i^1 = D + d_i^2$ et $d_i^1 = d_i^3$ et $d_i^2 = d_i^3$. Donc D=0 : c'est impossible.

Quant au 4 derniers cas, il suffit de permuter les rôles des blocs pour avoir les mêmes conclusions.
Donc si N est premier, il n'y a aucune dépendance par couples.

Voyons maintenant les dépendances par triplets

$(....X_{i,n+d_i^1}.....X_{i,n+D+d_i^1}...)....(....X_{i,n+d_i^2}...X_{i,n+D+d_i^2}....)....(...X_{i,n+d_i^3}...X_{i,n+D+d_i^3}...)$.

Donc,
ou $n + d_i^1 = n + d_i^2 = n + d_i^3$ et $n + D + d_i^1 = n + D + d_i^2 = n + D + d_i^3$,
ou $n + d_i^1 = n + d_i^2 = n + D + d_i^3$ et $n + D + d_i^1 = n + D + d_i^2 = n + d_i^3$,
ou $n + d_i^1 = n + D + d_i^2 = n + d_i^3$ et $n + D + d_i^1 = n + d_i^2 = n + D + d_i^3$,
ou $n + d_i^1 = n + D + d_i^2 = n + D + d_i^3$ et $n + D + d_i^1 = n + d_i^2 = n + d_i^3$.

Donc,

Si $n + d_i^1 = n + d_i^2 = n + d_i^3$ et $n + D + d_i^1 = n + D + d_i^2 = n + D + d_i^3$.
Alors, $d_i^1 = d_i^2 = d_i^3$.

Si $n + d_i^1 = n + d_i^2 = n + D + d_i^3$ et $n + D + d_i^1 = n + D + d_i^2 = n + d_i^3$.
Alors, $d_i^1 = d_i^2 = D + d_i^3$ et $D + d_i^1 = D + d_i^2 = d_i^3$.
Donc, $d_i^1 = D + d_i^3$ et $d_i^1 = d_i^3 - D$. Donc 2D=0, D=N/2.

Si $n + d_i^1 = n + D + d_i^2 = n + d_i^3$ et $n + D + d_i^1 = n + d_i^2 = n + D + d_i^3$.
Alors, $d_i^1 = D + d_i^2 = d_i^3$ et $D + d_i^1 = d_i^2 = D + d_i^3$.
Donc, $d_i^1 = D + d_i^2$ et $D + d_i^1 = d_i^2$. Donc 2D=0, D=N/2.

Si $n + d_i^1 = n + D + d_i^2 = n + D + d_i^3$ et $n + D + d_i^1 = n + d_i^2 = n + d_i^3$.
Alors, $d_i^1 = D + d_i^2 = D + d_i^3$ et $D + d_i^1 = d_i^2 = d_i^3$.
Donc, $d_i^1 = D + d_i^2$ et $D + d_i^1 = d_i^2$. Donc 2D=0, D=N/2.

Donc si N est premier, ce cas de dépendance n'existe pas. Donc, il n'y a pas de dépendance par triplet.

9.2.4 Dépendance sur 3 blocs pour 8-uplets : cas 1

On se place d'abord dans le cas où, avec $D,\ D' \neq 0$, la première ligne est
$(..X_{1,n}..)....(..X_{1,n+D}..X_{1,n+D'}...X_{1,n+D"}...)..(..X_{1,n}...X_{1,n+D}...X_{1,n+D'}...X_{1,n+D"}..)$.

Donc, on a pour la ligne i,
$(..X_{1,n+d_i^1}...)....(...X_{1,n+D+d_i^2}....X_{1,n+D'+d_i^2}....X_{1,n+D"+d_i^2}...).....$
$.....(....X_{1,n+d_i^3}.....X_{1,n+D+d_i^3}...X_{1,n+D'+d_i^3}...X_{1,n+D"+d_i^3}....).$

Donc, pour avoir une dépendance, il faut que
ou $n+d_i^1 \equiv n+d_i^3$, $n+D+d_i^2 \equiv n+D+d_i^3$, $n+D'+d_i^2 \equiv n+D'+d_i^3$, et $n+D"+d_i^2 \equiv n+D"+d_i^3$,
ou $n+d_i^1 \equiv n+d_i^3$, $n+D+d_i^2 \equiv n+D+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+D'+d_i^3$,
ETC
ou $n+d_i^1 \equiv n+d_i^3$, $n+D+d_i^2 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+D+d_i^3$,
ETC
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^2 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+d_i^3$, et $n+D"+d_i^2 \equiv n+D"+d_i^3$,
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^2 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+d_i^3$,
ETC

Les ETC désignent des cas qui donneront le même type de résultats lorsque N est premier ou si on ne prend que les N/2 premiers termes de chaque bloc. Il n'est pas utile d'étudier tous les cas des ETC : c'est trop long à écrire.

Donc,
Cas 1) $d_i^1 = d_i^3 = d_i^2$.
Cas 2) $d_i^1 \equiv d_i^3$, $D+d_i^2 \equiv D+d_i^3$, $D'+d_i^2 \equiv D"+d_i^3$, et $D"+d_i^2 \equiv D'+d_i^3$: D'=D": impossible.
Cas 3) $d_i^1 \equiv d_i^3$, $D+d_i^2 \equiv D'+d_i^3$, $D'+d_i^2 \equiv D"+d_i^3$, et $d_i^2+D" \equiv D+d_i^3$. Donc, $d_i^2 \equiv D'-D+d_i^3$ et $D'+D'-D \equiv D"$, et $D"+D'-D \equiv D$. Donc 2D'=D+D" et 2D=D'+D". Donc, 2D'-D=2D-D'. Donc, 3D=3D'. Donc $D-D' = \alpha N/3$ ou $\alpha = 1$ ou 2. Donc, $D" = 2D'-D = 2D'-D'-\alpha N/3 = D'-\alpha N/3$. Donc, D,D', D" représentent des termes distants de N/3 et il y en a au moins deux de ces termes distants de 2N/3.
Cas 4) $d_i^1 \equiv D+d_i^3$, $D+d_i^2 \equiv D'+d_i^3$, $D'+d_i^2 \equiv d_i^3$, et $D"+d_i^2 \equiv D"+d_i^3$. Donc, $d_i^2 \equiv d_i^3$ et donc D'=0: impossible.
Cas 5) $d_i^1 \equiv D+d_i^3$, $D+d_i^2 \equiv D'+d_i^3$, $D'+d_i^2 \equiv D"+d_i^3$, et $D"+d_i^2 \equiv d_i^3$. Donc, $D+d_i^2 \equiv D'+D"+d_i^2$, $D'+d_i^2 \equiv D"+D"+d_i^2$. Donc D=D'+D" et D'=2D". Donc D=3D".

Il n'y a donc que dans le cas 5 (celui où on ne retrouve aucun des 0,D,D',D", présent dans les deux termes d'une égalité) que l'on peut trouver des dépendances autres que l'egalité. Comme cas semblables, il y a
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^2 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+d_i^3$,
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^2 \equiv n+D"+d_i^3$, $n+D'+d_i^2 \equiv n+d_i^3$, et $n+D"+d_i^2 \equiv n+D'+d_i^3$,
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^2 \equiv n+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+D'+d_i^3$,

ou $n+d_i^1 \equiv n+D'+d_i^3$, $n+D+d_i^2 \equiv n+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+D+d_i^3$,
ou $n+d_i^1 \equiv n+D'+d_i^3$, $n+D+d_i^2 \equiv n+D"+d_i^3$, $n+D'+d_i^2 \equiv n+d_i^3$, et $n+D"+d_i^2 \equiv n+D+d_i^3$,
ou $n+d_i^1 \equiv n+D'+d_i^3$, $n+D+d_i^2 \equiv n+D"+d_i^3$, $n+D'+d_i^2 \equiv n+D+d_i^3$, et $n+D"+d_i^2 \equiv n+d_i^3$,

ou $n+d_i^1 \equiv n+D"+d_i^3$, $n+D+d_i^2 \equiv n+d_i^3$, $n+D'+d_i^2 \equiv n+D+d_i^3$, et $n+D"+d_i^2 \equiv n+D'+d_i^3$,
ou $n+d_i^1 \equiv n+D"+d_i^3$, $n+D+d_i^2 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+d_i^3$, et $n+D"+d_i^2 \equiv n+D+d_i^3$,
ou $n+d_i^1 \equiv n+D"+d_i^3$, $n+D+d_i^2 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+D+d_i^3$, et $n+D"+d_i^2 \equiv n+d_i^3$,

Donc,
Cas 5) $d_i^1 \equiv D+d_i^3$, $D+d_i^2 \equiv D'+d_i^3$, $D'+d_i^2 \equiv D"+d_i^3$, et $D"+d_i^2 \equiv d_i^3$. Donc, $D+d_i^2 \equiv D'+D"+d_i^2$, $D'+d_i^2 \equiv D"+D"+d_i^2$. Donc D=D'+D" et D'=2D". Donc D=3D".
Cas 6) $d_i^1 \equiv D+d_i^3$, $D+d_i^2 \equiv D"+d_i^3$, $D'+d_i^2 \equiv d_i^3$, et $D"+d_i^2 \equiv D'+d_i^3$. Donc, $D+d_i^2 \equiv D"+D'+d_i^2$, et $D"+d_i^2 \equiv D'+D'+d_i^2$. Donc, D=D"+D' et D"=2D'. Donc D=3D'.
Cas 7) $d_i^1 \equiv D+d_i^3$, $D+d_i^2 \equiv d_i^3$, $D'+d_i^2 \equiv D"+d_i^3$, et $D"+d_i^2 \equiv D'+d_i^3$. Donc, $D'+d_i^2 \equiv D"+D+d_i^2$, et $D"+d_i^2 \equiv D'+D+d_i^2$. Donc, D'=D"+D et D"=D'+D. Donc, D'=D'+D+D. Donc 2D=0. Donc D=N/2.

Cas 8) $d_i^1 \equiv D' + d_i^3$, $D + d_i^2 \equiv d_i^3$, $D' + d_i^2 \equiv D" + d_i^3$, et $D" + d_i^2 \equiv D + d_i^3$. Donc, $D' + d_i^2 \equiv D" + D + d_i^2$, et $D" + d_i^2 \equiv D + D + d_i^2$. Donc D"=2D et D'=3D.
Cas 9) $d_i^1 \equiv D' + d_i^3$, $D + d_i^2 \equiv D" + d_i^3$, $D' + d_i^2 \equiv d_i^3$, et $D" + d_i^2 \equiv D + d_i^3$. Donc, $D + d_i^2 \equiv D" + D' + d_i^2$, et $D" + d_i^2 \equiv D + D' + d_i^2$. Donc, D=D"+D' et D"=D+D'. Donc, D=D+D'+D'. Donc 2D'=0. Donc D'=N/2.
Cas 10) $d_i^1 \equiv D' + d_i^3$, $D + d_i^2 \equiv D" + d_i^3$, $D' + d_i^2 \equiv D + d_i^3$, et $D" + d_i^2 \equiv d_i^3$. Donc, $D + d_i^2 \equiv D" + D" + d_i^2$, $D' + d_i^2 \equiv D + D" + d_i^2$. Donc D=2D" et D'=3D".

Cas 11) $d_i^1 \equiv D" + d_i^3$, $D + d_i^2 \equiv d_i^3$, $D' + d_i^2 \equiv D + d_i^3$, et $D" + d_i^2 \equiv D' + d_i^3$. Donc, $D' + d_i^2 \equiv D + D + d_i^2$, et $D" + d_i^2 \equiv D' + D + d_i^2$. Donc, D'=2D et D"=3D.
Cas 12) $d_i^1 \equiv D" + d_i^3$, $D + d_i^2 \equiv D' + d_i^3$, $D' + d_i^2 \equiv d_i^3$, et $D" + d_i^2 \equiv D + d_i^3$. Donc, $D + d_i^2 \equiv D' + D' + d_i^2$, et $D" + d_i^2 \equiv D + D' + d_i^2$. Donc, D=2D' et D"=3D'.
Cas 13) $d_i^1 \equiv D" + d_i^3$, $D + d_i^2 \equiv D' + d_i^3$, $D' + d_i^2 \equiv D + d_i^3$, et $D" + d_i^2 \equiv d_i^3$. Donc, $D + d_i^2 \equiv D' + D" + d_i^2$, $D' + d_i^2 \equiv D + D" + d_i^2$. Donc D=D'+D" et D'=D+D". Donc, D'=D'+D" +D". Donc 2D"=0. Donc, D"=N/2.

Finalement si N est premier ou si on prend seulement N/2 termes de chaque bloc, il ne peut y avoir que un cas en dehors de $d_i^1 \equiv d_i^2 \equiv d_i^3$, le cas où par exemple D=3D" lorsque D" est petit, (par exemple D"=1 : si $D \geq N/2$, cela n'a plus d'importance). Cela fait que pour chaque ligne on a 2N décalages possibles.

Donc, en choisissant bien D, D' et D" (par exemple D=3D"), la probabilité de trouver un 8-uplet dépendant est

$$\frac{2^{I-1}N^{I-1}}{N^{I-1}(N^{I-1}-1)(N^{I-1}-2)} \approx 2^{I-1}/N^{2(I-1)} \ .$$

C'est la même que celle de la section 8.2. Il n'y a donc pas plus de chances de trouver une dépendance dans ce cas.

9.2.5 Dépendance sur 3 blocs pour 8-uplets : cas 2

On se place dans le cas où, avec $D, D' \neq 0$, la première ligne est
$(..X_{1,n}..X_{1,n+D}...).....(...X_{1,n+D'}...X_{1,n+D"}...)....(..X_{1,n}...X_{1,n+D}...X_{1,n+D'}...X_{1,n+D"}....)$.

Donc, on a pour la ligne i,
$(...X_{i,n+d_i^1}...X_{i,n+D+d_i^1}...).....(...X_{i,n+D'+d_i^2}...X_{i,n+D"+d_i^2}...).....$
$.....(..X_{i,n+d_i^3}...X_{i,n+D+d_i^3}..X_{i,n+D'+d_i^3}...X_{i,n+D"+d_i^3}...).$

Donc, pour avoir une dépendance
ou $n+d_i^1 \equiv n+d_i^3$, $n+D+d_i^1 \equiv n+D+d_i^3$, $n+D'+d_i^2 \equiv n+D'+d_i^3$, et $n+D"+d_i^2 \equiv n+D"+d_i^3$,
ou $n+d_i^1 \equiv n+d_i^3$, $n+D+d_i^1 \equiv n+D+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+D'+d_i^3$,
ou $n+d_i^1 \equiv n+d_i^3$, $n+D+d_i^1 \equiv n+D"+d_i^3$, $n+D'+d_i^2 \equiv n+D'+d_i^3$, et $n+D"+d_i^2 \equiv n+D+d_i^3$,
ETC
ou $n+d_i^1 \equiv n+d_i^3$, $n+D+d_i^1 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+D+d_i^3$,
ETC
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^1 \equiv n+d_i^3$, $n+D'+d_i^2 \equiv n+D'+d_i^3$, et $n+D"+d_i^2 \equiv n+D"+d_i^3$,
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^1 \equiv n+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+D'+d_i^3$,
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^1 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+d_i^3$, et $n+D"+d_i^2 \equiv n+D"+d_i^3$,
ou $n+d_i^1 \equiv n+D+d_i^3$, $n+D+d_i^1 \equiv n+D'+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+d_i^3$,
ETC
ou $n+d_i^1 \equiv n+D'+d_i^3$, $n+D+d_i^1 \equiv n+D+d_i^3$, $n+D'+d_i^2 \equiv n+d_i^3$, et $n+D"+d_i^2 \equiv n+D"+d_i^3$,
ou $n+d_i^1 \equiv n+D'+d_i^3$, $n+D+d_i^1 \equiv n+D+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+d_i^3$,

ETC
ou $n+d_i^1 \equiv n+D'+d_i^3$, $n+D+d_i^1 \equiv n+d_i^3$, $n+D'+d_i^2 \equiv n+D+d_i^3$, et $n+D"+d_i^2 \equiv n+D"+d_i^3$,
ou $n+d_i^1 \equiv n+D'+d_i^3$, $n+D+d_i^1 \equiv n+d_i^3$, $n+D'+d_i^2 \equiv n+D"+d_i^3$, et $n+D"+d_i^2 \equiv n+D+d_i^3$,
ETC
ou $n+d_i^1 \equiv n+D'+d_i^3$, $n+D+d_i^1 \equiv n+D"+d_i^3$, $n+D'+d_i^2 \equiv n+D+d_i^3$, et $n+D"+d_i^2 \equiv n+d_i^3$,
ETC

Donc,
Cas 1) $d_i^1 = d_i^3 = d_i^2$.

Cas 2) $d_i^1 \equiv d_i^3$, $D' + d_i^2 \equiv D" + d_i^3$, et $D" + d_i^2 \equiv D' + d_i^3$. Donc D'-D"=D"-D'. Donc D'=N/2+D".

Cas 3) $d_i^1 \equiv d_i^3$, $D + d_i^1 \equiv D" + d_i^3$, $D' + d_i^2 \equiv D' + d_i^3$, et $D" + d_i^2 \equiv D + d_i^3$. Donc D=D" : impossible.

Cas 4) $d_i^1 \equiv d_i^3$, $D + d_i^1 \equiv D' + d_i^3$, $D' + d_i^2 \equiv D" + d_i^3$, et $D" + d_i^2 \equiv D + d_i^3$. Donc D=D': impossible.

Cas 5) $d_i^1 \equiv D+d_i^3$, $D+d_i^1 \equiv d_i^3$, $D'+d_i^2 \equiv D'+d_i^3$, et $D"+d_i^2 \equiv D"+d_i^3$. Donc 2D=0 : D=N/2.

Cas 6) $d_i^1 \equiv D+d_i^3$, $D+d_i^1 \equiv d_i^3$, $D'+d_i^2 \equiv D"+d_i^3$, et $D"+d_i^2 \equiv D'+d_i^3$. Donc 2D=0 : D=N/2.

Cas 7) $d_i^1 \equiv D + d_i^3$, $D + d_i^1 \equiv D' + d_i^3$, $D' + d_i^2 \equiv d_i^3$, et $D" + d_i^2 \equiv D" + d_i^3$. Donc, $d_i^2 = d_i^3$. Donc D'=0 : impossible.

Cas 8) $d_i^1 \equiv D + d_i^3$, $D + d_i^1 \equiv D' + d_i^3$, $D' + d_i^2 \equiv D" + d_i^3$, et $D" + d_i^2 \equiv d_i^3$. Donc D=D'-D et D'-D"=D". Donc D'=2D=2D". Donc D=D"+N/2.

Cas 9) $d_i^1 \equiv D' + d_i^3$, $D + d_i^1 \equiv D + d_i^3$, $D' + d_i^2 \equiv d_i^3$, et $D" + d_i^2 \equiv D" + d_i^3$. Donc D'=0 : impossible.

Cas 10) $d_i^1 \equiv D' + d_i^3$, $D + d_i^1 \equiv D + d_i^3$, $D' + d_i^2 \equiv D" + d_i^3$, et $D" + d_i^2 \equiv d_i^3$. Donc D'=0 : impossible.

Cas 11) $d_i^1 \equiv D' + d_i^3$, $D + d_i^1 \equiv d_i^3$, $D' + d_i^2 \equiv D + d_i^3$, et $D" + d_i^2 \equiv D" + d_i^3$. Donc D=D' : impossible.

Cas 12) $d_i^1 \equiv D' + d_i^3$, $D + d_i^1 \equiv d_i^3$, $D' + d_i^2 \equiv D" + d_i^3$, et $D" + d_i^2 \equiv D + d_i^3$. Donc, D=-D' et D'-D"=D"-D. Donc, 2D"=0. Donc D"=N/2.

Cas 13) $d_i^1 \equiv D' + d_i^3$, $D + d_i^1 \equiv D" + d_i^3$, $D' + d_i^2 \equiv D + d_i^3$, et $D" + d_i^2 \equiv d_i^3$. Donc D=D"-D' et D'-D"=D. Donc 2D=0 : D=N/2.

Donc si N est premier ou si on ne prend que les N/2 premiers termes de chaque blocs, on ne peut trouver d'autres dépendances que l'identité, ce qui est impossible.

9.2.6 3 blocs différents pour 12-uplets

On s'intéresse aux dépendances où les termes sont groupés par groupe de 2. Pour ce type de dépendance, il faut qu'il y aie deux éléments du premier groupe égaux à 2 éléments du deuxième etc. On est donc dans le cas
$(X_{1,n}..X_{1,n+D}..X_{1,n+E}..X_{1,n+F})(X_{1,n}..X_{1,n+D}..X_{1,n+G}..X_{1,n+H})(X_{1,n+E}..X_{1,n+F}..X_{1,n+G}..X_{1,n+H})$.

Comme précédemment, il est trop long d'étudier tous les cas. Mais beaucoup de cas sont du même type. On étudie seulement qu'un cas par type. Donc

OU

$n + d_1^i = n + d_2^i$, $n + D + d_1^i = n + D + d_2^i$,
$n + E + d_1^i = n + E + d_3^i$, $n + F + d_1^i = n + F + d_3^i$,
$n + G + d_2^i = n + G + d_3^i$, $n + H + d_2^i = n + H + d_3^i$.
Alors, $d_1^i = d_2^i = d_3^i$.

OU

$n + d_1^i = n + d_2^i$, $n + D + d_1^i = n + D + d_2^i$,
$n + E + d_1^i = n + F + d_3^i$, $n + F + d_1^i = n + E + d_3^i$,
$n + G + d_2^i = n + G + d_3^i$, $n + H + d_2^i = n + H + d_3^i$.
alors, $d_1^i = d_2^i = d_3^i$,

OU

$n + d_1^i = n + d_2^i$, $n + D + d_1^i = n + D + d_2^i$,
$n + E + d_1^i = n + F + d_3^i$, $n + F + d_1^i = n + G + d_3^i$,
$n + G + d_2^i = n + H + d_3^i$, $n + H + d_2^i = n + E + d_3^i$.
Alors, $d_1^i = d_2^i$,
$E + d_1^i = F + d_3^i$, $F + d_1^i = G + d_3^i$,
$G + d_2^i = H + d_3^i$, $H + d_2^i = E + d_3^i$.
Donc, $d_1^i = F + d_3^i - E$,
$F + F + d_3^i - E = G + d_3^i$, $G + F + d_3^i - E = H + d_3^i$, $H + F + d_3^i - E = E + d_3^i$.
Donc, 2F=E+G, G+F=E+H, H+F=2E.
Donc 2F-E = G =E+H-F. Donc 2F-E+F=E+H.
Donc, 2F-E+F=E+2E-F. Donc, 4F=4E.
Donc, si N est premier E=F, ce qui est impossible.

OU

$n + d_1^i = n + d_2^i$, $n + D + d_1^i = n + D + d_2^i$,
$n + E + d_1^i = n + G + d_3^i$, $n + F + d_1^i = n + H + d_3^i$,
$n + G + d_2^i = n + E + d_3^i$, $n + H + d_2^i = n + F + d_3^i$
Alors, $d_1^i = d_2^i$, $E + d_1^i = G + d_3^i$, $F + d_1^i = H + d_3^i$, $G + d_2^i = E + d_3^i$, $H + d_2^i = F + d_3^i$
Donc,
$E + d_1^i - G = d_3^i$,
$F + d_1^i = H + E + d_1^i - G$, $G + d_1^i = E + E + d_1^i - G$, $H + d_1^i = F + E + d_1^i - G$;
Donc, 2G=2E. Donc si N premier, E=G impossible

OU

$n + d_1^i = n + D + d_2^i$, $n + D + d_1^i = n + d_2^i$,
$n + E + d_1^i = n + F + d_3^i$, $n + F + d_1^i = n + E + d_3^i$,
$n + G + d_2^i = n + H + d_3^i$, $n + H + d_2^i = n + G + d_3^i$.
Alors, $d_1^i = D + d_2^i$, $D + d_1^i = d_2^i$.
Donc, 2D=0, ce qui est impossible si N est premier.

OU

$n + d_1^i = n + D + d_2^i$, $n + D + d_1^i = n + G + d_2^i$,
$n + E + d_1^i = n + F + d_3^i$, $n + F + d_1^i = n + E + d_3^i$,
$n + d_2^i = n + H + d_3^i$, $n + H + d_2^i = n + G + d_3^i$
Alors, $E + d_1^i = F + d_3^i$, $F + d_1^i = E + d_3^i$,
Donc, $d_1^i = F + d_3^i - E$.

Donc, $F + F + d_3^i - E = E + d_3^i$. Donc 2F=2E et F= E+N/2 , ce qui est impossible si N est premier.

OU

$n + d_1^i = n + D + d_2^i$, $n + D + d_1^i = n + G + d_2^i$,
$n + E + d_1^i = n + F + d_3^i$, $n + F + d_1^i = n + G + d_3^i$,
$n + d_2^i = n + H + d_3^i$, $n + H + d_2^i = n + E + d_3^i$.
Alors, $d_1^i = D + d_2^i$, $D + d_1^i = G + d_2^i$, $E + d_1^i = F + d_3^i$, $F + d_1^i = G + d_3^i$, $d_2^i = H + d_3^i$, $H + d_2^i = E + d_3^i$.

Donc,

$2D + d_2^i = G + d_2^i$, $2H + d_3^i = E + d_3^i$, $E + D + d_2^i = F + d_3^i$, $F + D + d_2^i = G + d_3^i$.

Donc,

$2D = G$, $2H = E$,

$E + D + H + d_3^i = F + d_3^i$, $F + D + H + d_3^i = G + d_3^i$.

Donc,

$E + D + H = F$, $F + D + H = G$, Donc, E+D+H+D+H=G.

Donc, E+2D+2H=G Donc E+G+E=G. Donc 2E=0. Donc E=N/2, ce qui est impossible si N est premier.

OU

$n + d_1^i = n + D + d_2^i$, $n + D + d_1^i = n + H + d_2^i$,
$n + E + d_1^i = n + F + d_3^i$, $n + F + d_1^i = n + G + d_3^i$,
$n + d_2^i = n + E + d_3^i$, $n + G + d_2^i = n + H + d_3^i$.
Alors,
$d_1^i = D + d_2^i$, $D + d_1^i = H + d_2^i$,
$E + d_1^i = F + d_3^i$, $F + d_1^i = G + d_3^i$,
$d_2^i = E + d_3^i$, $G + d_2^i = H + d_3^i$.
Donc,
$D + D + d_2^i = H + d_2^i$,
$E + D + d_2^i = F + d_3^i$, $F + D + d_2^i = G + d_3^i$,
$G + E + d_3^i = H + d_3^i$.

Donc, 2D=H, G+E=H,

$E + D + E + d_3^i = F + d_3^i$, $F + D + E + d_3^i = G + d_3^i$.

Donc,

$2E + D = F$, $F + D + E = G$.

Donc, $F + F - 2E + E = G$, Donc, 2F=G+E.

Donc 2F=H=2D. Donc F=D+N/2, ce qui est impossible si N est premier.

OU

$n + d_1^i = n + D + d_2^i$, $n + D + d_1^i = n + H + d_2^i$,
$n + E + d_1^i = n + G + d_3^i$, $n + F + d_1^i = n + H + d_3^i$,
$n + d_2^i = n + E + d_3^i$, $n + G + d_2^i = n + F + d_3^i$.
Alors,
$d_1^i = D + d_2^i$, $D + d_1^i = H + d_2^i$,
$E + d_1^i = G + d_3^i$, $F + d_1^i = H + d_3^i$,
$d_2^i = E + d_3^i$, $G + d_2^i = F + d_3^i$.

Donc,

$D + D + d_2^i = H + d_2^i$,
$E + D + d_2^i = G + d_3^i$, $F + D + d_2^i = H + d_3^i$,
$G + E + d_3^i = F + d_3^i$.

Donc,
2D=H, G+E=F,
$E + D + E + d_3^i = G + d_3^i$, $F + D + E + d_3^i = H + d_3^i$.
Donc,
2E+D=G, F+D+E=H=2D, Donc, F+E=D.
Donc,
2E+F+E=3E+F=G, et G+E=F.
Donc, 3E+G+E=G. Donc 4E=0. Donc E=aN/4, a=1,2,3.
Donc 2aN/4+D=2E+D=G. Donc fatalement il y aura des éléments distant d'au moins N/2. De toutes façon, si N est premier c'est impossible.

OU
$n + d_1^i = n + G + d_2^i$, $n + D + d_1^i = n + H + d_2^i$,
$n + E + d_1^i = n + G + d_3^i$, $n + F + d_1^i = n + H + d_3^i$,
$n + d_2^i = n + E + d_3^i$, $n + D + d_2^i = n + F + d_3^i$.
Alors,
$d_1^i = G + d_2^i$, $D + d_1^i = H + d_2^i$,
$E + d_1^i = G + d_3^i$, $F + d_1^i = H + d_3^i$,
$d_2^i = E + d_3^i$, $D + d_2^i = F + d_3^i$.
Donc,
$D + G + d_2^i = H + d_2^i$,
$E + G + d_2^i = G + d_3^i$, $F + G + d_2^i = H + d_3^i$,
$D + E + d_3^i = F + d_3^i$.
Donc,
D+G=H, D+E=F,
$E + G + E + d_3^i = G + d_3^i$, $F + G + E + d_3^i = H + d_3^i$.
Donc, 2E=0, E=N/2, ce qui est impossible si N est premier.

Rappelons qu'il reste à étudier le cas des triplets si on veut être parfaitement certain qu'il n'y a pas de dépendances possibles.

9.2.7 Conclusion

Donc, si on prend à chaque bloc seulement N/2 nombres ou si N est premier, on n'a pas moyen d'obtenir une dépendance avec une probabilité plus forte que celle qu'on a déjà obtenue. Donc, on ne trouvera sans doute pas de dépendances qui puissent servir à casser le système.

10 Système de cryptage

Le Vernam One time Pad (VOTP) est le seul système de cryptage à clef secrète prouvé indécryptable (cf p 110 de [6]). Malheureusement, il est difficilement utilisable. Il n'est donc utilisé que dans les cas où une sécurité extrême est nécessaire comme le téléphone rouge Washington-Moscou.

Ce que nous avons fait dans ce rapport, c'est présenter un nouveau moyen d'utiliser le VOTP. Il devient utilisable, mais reste indécryptable. De plus, il est simple et rapide. C'est donc un systeme extrêmement performant.

Certes il existent déjà plusieurs moyens d'utiliser autrement le VOTP. Le plus simple est d'utiliser des générateurs pseudo-aléatoires cryptographiquement sûrs, par exemple le BBS [2], le Yarrow [3], Isaac [4] ou Fortuna [5].

En fait, notre méthode consiste bien à créer un générateur pseudo-aléatoire. Mais celui-ci n'utilise pas des fonctions de hachage ou les propriétés des nombres premiers, mais, par l'algorithme $\mathcal{B}$, il va transformer des suites réellement aléatoires en suites presque aléatoires. Pour démontrer cela, on a utilisé essentiellement les propriétés des nombres aléatoires.

Dans le cas où on utilise un ordinateur central, chaque ordinateur périphérique sera muni d'une suite réellement aléatoire qui servira de moyen de cryptage selon la méthode du VOTP. Mais la suite de l'ordinateur central, i.e. la matrice $\{x_{i,j}\}$ sera inconnue. On retrouve bien ainsi le Vernam One Time Pad. Il ne s'agit que d'une légère variante et on peut donc dire qu'on l'utilise autrement.

10.1 Choix des paramêtres

On va donc choisir les paramêtres pour définir un système de cryptage. On utilise les résultats que nous avons vu dans les chapitres précédents.

Aussi, on choisira donc une matrice aléatoire $\{x_{i,j}\} \in \{0, 1, ..., m\}$ telle que m est premier. Le mieux est de choisir m=N. Mais on peut aussi choisir m=2 par exemple. Dans chaque bloc, on prendra seulement les $\lfloor N/2 \rfloor$ premiers termes de la suite x_j^d.

10.2 Définition linéaire des décalages

On va maintenant donner un exemple pour definir les décalages en fonction d'un seul paramêtre. On imposera aussi qu'ils soient choisis au hasard.

Pour etre sûr de ne pas retrouver le même décalage, on va définir les premiers décalages grâce à une congruence linéaire, et les derniers au hasard. Si la congruence linéaire est bien choisie, on utilisera au plus une seule fois chaque décalage. Avec cette méthode, on n'obtiendra pas une suite x_n aussi longue que la suite réunion de toutes les suites x_j^d possibles. Ce n'est pas génant : celle-ci est beaucoup trop longue pour pouvoir être utilisée en entier.

Par exemple, on utilise I_1 lignes pour définir les premiers décalages : on prend une congruence linéaire T_1 modulo N^{I_1} de période N^{I_1} et on redécompose chaque $T_1^n(y)$ en I_1 chiffres appartenant à $\{1, 2, ..., N\}$: il suffit d'écrire $T_1^n(y) \in \{0, 1, 2, ..., N^{I_1} - 1\}$ en base N. Avec ces I_1 chiffres, on définit les I_1 premiers décalages associés au n-ème bloc, i.e. les décalages des lignes $i = 2, 3, ..., I_1 + 1$.

Ensuite, on définit les derniers décalages au hasard : pour cela, pour les I_2 lignes restantes, on emploie une suite aléatoire (non IID) $da_n \in \{0, 1, ..., N - 1\}$, $n = 1, 2, ..., I_2N_3$, où N_3 est le nombre de fois que l'on veut utiliser des décalages : $N_3 \leq N^{I_1}$. Si I_2N_3 est important, on peut obtenir cette suite en appliquant l'algorithme $\mathcal{B}$ sur une deuxième matrice aléatoire $\{x'_{i,j}\}$: pour cette deuxième application de l'algorithme $\mathcal{B}$, les décalages sont tous définis par les congruences.

Par exemple, on applique T_1 sur une autre matrice aléatoire $x'_{i,j}$, $i = 1, 2, ..., I_1 + 1$, j=1,2,....,N, où $x'_{i,j} \in \{0, 1, ..., m' - 1\}$ avec $m' = N$. On utilise alors l'algorithme $\mathcal{B}$ pour obtenir des suites $x'^{d'}_n$ et on prend les I_2 premiers termes de cette suite à chaque décalage d'(t).

Enfin, plutôt que d'utiliser une nouvelle matrice $\{x'_{i,j}\}$, on peut aussi utiliser la matrice $\{x_{i,j}\}$ d'une autre façon.

10.3 Système de cryptage

On veut se servir de l'algorithme $\mathcal{B}$ pour transmettre des messages. On suppose donc que l'on a un message M(n) composé de nombres à valeurs dans $\{0, 1, ..., m-1\}$.

Transmission entre deux personnes On utilise donc la suite x_n réunion des suites $x_j^{d(t)}$ definie par exemple comme précédemment.

Si on utilise rigoureusement la même méthode que pour le Vernam One time Pad, alors, à chaque message, il faut donc associer n, le nombre de termes de la suite définitivement utilisés.

On suppose donc que le k-ème message est de longueur n'_k. La longueur de la suite utilisée après ce k-ème message est donc $n_k = n'_0 + n'_1 + n'_2 + ... + n'_k$ où $n'_0 = 0$.

Donc pour ce (k+1)-ème message on utilise la suite $x_n \in \{0, 1, ..., m-1\}$, $n = n_k + 1, n_k + 2, ...n_k + n'_{k+1}$.

On crypte le message $M_n \in \{0, 1, ..., m-1\}$ en posant $C_n = \overline{M_n + x_{n_k+n}}$ modulo m.

Maintenant, on peut utiliser une méthode legerement différente de celle du Vernam One time Pad : à chaque message, il faut donc associer n, le nombre de blocs déjà utilisés. On gaspille ainsi des nombres aléatoires. Mais ce n'est pas grave étant donné la taille de la suite x_n.

On suppose donc que le k-ème message utilise n'_k blocs, pas forcément entièrement : i.e. la fin du message utilise seulement le début du n'_k-ème bloc. Le nombre de blocs utilisés après ce k-ème message est donc $n"_k = n'_0 + n'_1 + n'_2 + ... + n'_k$ où $n'_0 = 0$.

Donc pour ce k+1-ème message on utilise la partie de la suite $x_n \in \{0, 1, ..., m-1\}$, qui commence avec le bloc $n_k + 1$.

L'avantage de cette variante, c'est qu'il sera impossible de casser un message codé par la méthode des dépendances. En effet, il faudrait trouver la dépendance dès qu'on commence un nouveau message. Or celle-ci est inconnue. Il faut d'abord la trouver si tant est qu'elle existe. Dans l'autre technique, on pourrait éventuellement avoir trouvé la dépendance grâce à l'envoi des messages précédents.

De plus le message codé implique que l'on a ajouté le message M(n) qui est inconnu. Donc, l'ajout de ce message inconnu fait disparaitre cette dépendance.

Transmission entre plusieurs personnes Dans ce cas, on peut supposer que $\{x_{i,j}\}$ est connue.

Pour correspondre entre S personnes, on choisit des clefs $Cl_{s,s'}$ toutes différentes.

Chaque clef $Cl_{s,s'}$ détermine une transformation de la matrice $\{x_{i,j}\}$ en une matrice $\{x_{i,j}^{s,s'}\}$. Par exemple, on peut la transformer en une ligne x_j^d, j=1,...,N, par l'algorithme $\mathcal{B}$ où d est determiné par s, s'. Ensuite, on réécrit x_j^d sous forme d'une matrice $\{x_{i,j}^{s,s'}\}$, i=1,2,...,I' , j=1,2,....,N' ou $N' \leq N/I'$.

Ensuite, par l'algorithme $\mathcal{B}$, cette matrice est transformée en une suite $x_n^{s,s'}$.

Pour cela on utilise une congruence T'_1 pour les I'_1 décalages par congruence.

Quant au I'_2 décalages au hasard, on peut utiliser une suite aléatoire non IID de longueur qui détermine plus de S^2 suites de longueur N'^{I_2}.

Chaque clef $Cl_{s,s'}$ doit pouvoir indiquer quelle partie de cette suite on utilise si on la décompose en S^2 sous-suites.

Alors, pour retrouver le message, si on procède par bloc, il faut retrouver les décalages. Comme ils sont aléatoires, il faut tous les essayer. Il est plus simple d'essayer les clefs.

Cas d'un ordinateur central On peut supposer que $\{x_{i,j}\}$ est inconnue car implantée seulement sur l'ordinateur central.

La matrice $\{x_j^{d^s}\}$ de chaque ordinateur périphérique est inconnue. Mais même si elle était connue cela ne serait pas génant. il suffirait d'appliquer la technique ci-dessus pour transmettre en retransformant $\{x_j^{d^s}\}$ en matrice $\{x_{i,j}^{d^s}\}$.

Dans le cas d'un ordinateur central, on peut sans doute se passer de décalages au hasard. On peut, le cas écheant, prendre plusieurs congruences pour définir les différents décalages.

L'essentiel c'est que, à un numero "s" d'un périphérique, correspondent les décalages $T_1(s)$ et donc la ligne $x_j^{d^s}$ que l'on peut réécrire sous forme de matrice.

Conservation du n Pour que le système soit indecryptable, il faut toujours garder de manière sûre le nombre de termes de la suite x_n deja utilisées ou de décalages déja utilisés. Si on la garde en mémoire dans l'ordinateur, il faut vérifier qu'elle n'a pas été changée, de manière à éviter que l'on emploie de nouveau la même suite.

De toutes façons, il faut que les deux personnes qui correspondent vérifient qu'ils ont le même n.

On peut aussi imposer que le décalage, soit fourni par la date, jour, heure, minute, seconde : comme cela, on n'utilisera jamais la même partie de la suite, ce qui revient à supprimer les n premiers termes déjà utilisés.

11 Généralisation

11.1 Utilisation d'un second cryptosystème

Le générateur pseudo-aléatoire que nous avons défini n'est pas parfait puisqu'il reste des dépendances possibles.

Aussi, il vient à l'idée de supprimer les dépendances restantes : pour cela on peut ajouter un système cryptographique quelconque.

Pour crypter un messsage M, au lieu de le transformer par $\overline{M(j)+x_j^d}$, on peut le transformer par $\overline{C(M(j)+x_j^d)}$ ou C est un algorithme cryptographique quelconque, par exemple le DES.

Cela doit normalement supprimer les dépendances ou les rendre très difficiles à découvrir.

Dans ce cas, on peut éventuellement utiliser des matrices $\{x_{i,j}\}$ plus petites.

Cette technique a aussi l'avantage de rendre une nouvelle vie au systeme C utilisé, et peut être même de le rendre définitivement indécryptable. Par exemple, on a un nouveau moyen d'utiliser le DES

En fait, "on crypte des suites aléatoires au lieu de crypter directement des messages". Donc des suites aléatoires assez longues sont nécessaires. Les suites x_n obtenues en utilisant l'algorithme $\mathcal{B}$ joueront ce rôle.

Cette seule technique (chiffrer des suites aléatoires au lieu de chiffrer directement des messages) est une méthode simple et efficace. Par exemple,

1) L'attaque à texte clair choisi est impossible : ainsi la cryptanalyse différentielle contre le DES est impossible s'il crypte des suites aléatoires.

2) Le problème de la redondance ne se pose pas.

Quant à savoir si, après transformation, une dépendance lineaire du type $X_{j_1}^{d^1} \equiv X_{j_2}^{d^2} - X_{j_3}^{d^3} + X_{j_4}^{d^4}$ modulo m sera empiriquement transformée en suite indépendante comme c'est le cas pour des dépendances pas trop grandes pour certaines congruences, cela dépend des propriétés de C, bien sûr.

11.2 Généralisation des décalages

Les décalages sont des permutations très particulières. On peut donc généraliser les décalages de chaque ligne en les remplaçant par des permutations Pe_i. On obtient alors en sommant les colonnes des suites de nombres construites par blocs $x_j^{Pe^s}$.

Le problème est de choisir des permutations assez rapides à effectuer. Par exemple, on peut choisir les permutations du type $P_{P^a}(e_1,e_2,e_3)$, $1 \le e_1 < e_2 < e_3 < N$, qui permutent entre eux les ensembles $\{x_{i,1},....,x_{i,e_1}\}$, $\{x_{i,e_1+1},....,x_{i,e_2}\}$, $\{x_{i,e_2+1},....,x_{i,e_3}\}$, $\{x_{i,e_3+1},....,x_{i,N}\}$ avec une permutation P^a définie sur 4 éléments.

C'est surtout dans le cas où on utilise $\mathcal{B}$ comme générateur aléatoire que cette technique est efficace parce que les raisonnements sur les dépendances des $x_j^{d^s}$ restent valables. Mais, plus ces permutations portent sur de petites sous-suites, moins les mêmes dépendances risquent de se reproduire longtemps : par exemple $x_{j_1+r}^{d^1} - x_{j_1+r}^{d^2} + x_{j_3+r}^{d^3} - x_{j_3+r}^{d^4} \equiv 0$ pour r=0,1,2,3,4,5, seulement. Dans ce cas, il n'est pas possible de pouvoir détecter ces dépendances comme on le faisait pour les décalages avec des relations du type $\delta_1 x_{j_1+r}^{d^1} + \delta_2 x_{j_1+r}^{d^2} + \delta_3 x_{j_3+r}^{d^3} + \delta_4 x_{j_3+r}^{d^4} \equiv 0$ qui duraient pour $r = 0,1,2,...,N_2$ quand N_2 est de l'ordre de N. Si $N_2 = 1$, on ne pourra pas détecter une telle dependance!

11.3 Généralisation des sommes

Plutot qu'utiliser les sommes sur toutes les lignes de la matrice, on peut sommer seulement certaines lignes différentes pour chaque bloc. Cela diminue la probabilité d'obtenir une dépendance

parce que pour qu'il y aie dépendance entre blocs il faut qu'il y aie les mêmes lignes pour chaque bloc.

De plus on peut prendre beaucoup de lignes : par exemple I=100 et en choisir 10 ou 20. Dans ce cas, c'est exactement le Subset Sum problem. Or celui-ci est admis incassable.

11.4 Utilisation double de l'algorithme $\mathcal{B}$

Plaçons nous dans le cas d'un ordinateur central. Alors chaque ordinateur périphérique "s" disposera pour correspondre d'une suite $x_j^{d^s}$. Si elle n'est pas assez grande ou si on veut plus de securité, on peut la transformer par l'algorithme $\mathcal{B}$ en une suite infiniment plus longue. Pour cela, on transforme la suite $x_j^{d^s}$ en une suite $x_j^{d^s,d'^t}$ en appliquant l'algorithme $\mathcal{B}$ avec des décalages d'^t.

On va voir que dans certains cas, il y a moins de dépendances de petite taille. Pour les dépendances de grande taille, on a vu que ce n'est pas important si m n'est pas trop grand : cf section 4.

Supposons qu'il y aie une seule dépendance entre 4 blocs. Voyons maintenant comment les dépendances se répartissent dans chaque ligne.

Par exemple, avec D=2, et en écrivant $X_j^{d^1} = X_j$, $X_j^{d^2} = Y_j$, $X_j^{d^3} = Z_j$, $X_j^{d^4} = T_j$, on a vu dans l'exemple de la section 8.3 que l'on a les quadruplets de dépendance suivants.

$(X_1, X_2, \mathbf{X_3}, \mathrm{X}_4, \mathrm{X}_5, \mathrm{X}_6, ..), (\mathrm{Y}_1, \mathrm{Y}_2, \mathbf{Y_3}, \mathrm{Y}_4, \mathrm{Y}_5, \mathrm{Y}_6, ..), (\mathrm{Z}_1, \mathrm{Z}_2, \mathrm{Z}_3, \mathrm{Z}_4, \mathbf{Z_5}, \mathrm{Z}_6, ..), (\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3, \mathrm{T}_4, \mathbf{T_5}, \mathrm{T}_6, ..)$

Donc on a aussi la dépendence

$(.., X_7, X_8, \mathbf{X_9}, \mathrm{X}_{10}, \mathrm{X}_{11}, \mathrm{X}_{12}), (.., \mathrm{Y}_7, \mathrm{Y}_8, \mathbf{Y_9}, \mathrm{Y}_{10}, \mathrm{Y}_{11}, \mathrm{Y}_{12})(.., \mathrm{Z}_7, \mathrm{Z}_8, \mathrm{Z}_9, \mathrm{Z}_{10}, \mathbf{Z_{11}}, \mathrm{Z}_{12})(.., \mathrm{T}_7, \mathrm{T}_8, \mathrm{T}_9, \mathrm{T}_{10}, \mathbf{T_{11}}, \mathrm{T}_{12})$

Supposons que ces dernieres suites représentent la deuxième ligne lorsque l'on réécrit les $x_j^{d^t}$ sous forme de matrice. Décalons ces lignes par l'algorithme $\mathcal{B}$. On a par exemple

$(X_1, X_2, \mathbf{X_3}, \mathrm{X}_4, \mathrm{X}_5, \mathrm{X}_6), (\mathrm{Y}_1, \mathrm{Y}_2, \mathbf{Y_3}, \mathrm{Y}_4, \mathrm{Y}_5, \mathrm{Y}_6), (\mathrm{Z}_1, \mathrm{Z}_2, \mathrm{Z}_3, \mathrm{Z}_4, \mathbf{Z_5}, \mathrm{Z}_6), (\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3, \mathrm{T}_4, \mathbf{T_5}, \mathrm{T}_6)$

$(X_7, X_8, \mathbf{X_9}, \mathrm{X}_{10}, \mathrm{X}_{11}, \mathrm{X}_{12}), (\mathrm{Y}_8, \mathbf{Y_9}, \mathrm{Y}_{10}, \mathrm{Y}_{11}, \mathrm{Y}_{12}, \mathrm{Y}_7)(\mathrm{Z}_{10}, \mathbf{Z_{11}}, \mathrm{Z}_{12}, \mathrm{Z}_7, \mathrm{Z}_8, \mathrm{Z}_9), (\mathrm{T}_9, \mathrm{T}_{10}, \mathbf{T_{11}}, \mathrm{T}_{12}, \mathrm{T}_7, \mathrm{T}_8)$

On voit que dans ce cas, il n'y aura pas dépendance (cf ci-après).

Notons maintenant par $d_i'^s$ les nouveaux décalages associés au bloc "s" pour la ligne i. D'après ce qui précède on voit que pour avoir le même type de dépendance, il faudrait d'abord que $d_2'^1 = d_2'^2$ et $d_2'^3 = d_2'^4$.

Il faudra aussi $d_2'^3 = d_2'^1$. En effet, si on choisit la dépendance dont le premier terme est X_{10}, celle-ci sera définie par

$(X_8, X_9, \mathbf{X_{10}}, \mathrm{X}_{11}, \mathrm{X}_{12}, \mathrm{X}_7), (\mathrm{Y}_8, \mathrm{Y}_9, \mathbf{Y_{10}}, \mathrm{Y}_{11}, \mathrm{Y}_{12}, \mathrm{Y}_7)(\mathrm{Z}_8, \mathrm{Z}_9, \mathrm{Z}_{10}, \mathrm{Z}_{11}, \mathbf{Z_{12}}, \mathrm{Z}_7), (\mathrm{T}_8, \mathrm{T}_9, \mathrm{T}_{10}, \mathrm{T}_{11}, \mathbf{T_{12}}, \mathrm{T}_7)$

Donc on aura dépendance si les lignes sont de la forme suivante :

$(X_1, X_2, \mathbf{X_3}, \mathrm{X}_4, \mathrm{X}_5, \mathrm{X}_6), (\mathrm{Y}_1, \mathrm{Y}_2, \mathbf{Y_3}, \mathrm{Y}_4, \mathrm{Y}_5, \mathrm{Y}_6), (\mathrm{Z}_1, \mathrm{Z}_2, \mathrm{Z}_3, \mathrm{Z}_4, \mathbf{Z_5}, \mathrm{Z}_6), (\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3, \mathrm{T}_4, \mathbf{T_5}, \mathrm{T}_6)$

$(X_8, X_9, \mathbf{X_{10}}, \mathrm{X}_{11}, \mathrm{X}_{12}, \mathrm{X}_7), (\mathrm{Y}_8, \mathrm{Y}_9, \mathbf{Y_{10}}, \mathrm{Y}_{11}, \mathrm{Y}_{12}, \mathrm{Y}_7)(\mathrm{Z}_8, \mathrm{Z}_9, \mathrm{Z}_{10}, \mathrm{Z}_{11}, \mathbf{Z_{12}}, \mathrm{Z}_7), (\mathrm{T}_8, \mathrm{T}_9, \mathrm{T}_{10}, \mathrm{T}_{11}, \mathbf{T_{12}}, \mathrm{T}_7)$

Ce type de condition doit être vrai pour chaque ligne : $d_i'^1 = d_i'^2 = d_i'^3 = d_i'^4$.

Maintenant prouvons que si ces conditions sur les décalages ne sont pas verifiées, la dépendance disparait. Reprenons l'exemple ci-dessus où les dépendances ne correspondent pas.

$(X_1, X_2, \mathbf{X_3}, \mathrm{X}_4, \mathrm{X}_5, \mathrm{X}_6), (\mathrm{Y}_1, \mathrm{Y}_2, \mathbf{Y_3}, \mathrm{Y}_4, \mathrm{Y}_5, \mathrm{Y}_6), (\mathrm{Z}_1, \mathrm{Z}_2, \mathrm{Z}_3, \mathrm{Z}_4, \mathbf{Z_5}, \mathrm{Z}_6), (\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3, \mathrm{T}_4, \mathbf{T_5}, \mathrm{T}_6)$

$(X_7, X_8, \mathbf{X_9}, \mathrm{X}_{10}, \mathrm{X}_{11}, \mathrm{X}_{12}), (\mathrm{Y}_8, \mathbf{Y_9}, \mathrm{Y}_{10}, \mathrm{Y}_{11}, \mathrm{Y}_{12}, \mathrm{Y}_7)(\mathrm{Z}_{10}, \mathbf{Z_{11}}, \mathrm{Z}_{12}, \mathrm{Z}_7, \mathrm{Z}_8, \mathrm{Z}_9), (\mathrm{T}_9, \mathrm{T}_{10}, \mathbf{T_{11}}, \mathrm{T}_{12}, \mathrm{T}_7, \mathrm{T}_8)$

Donc, le vecteur de la deuxième ligne correspondant à (X_3, Y_3, Z_5, T_5) qui est dépendant est (X_9, Y_{10}, Z_8, T_7) qui est indépendant. La somme de ces deux vecteurs modulo m est donc indépendante d'après le théorème 1 .

D'une façon générale, pour avoir 4-dépendance pour la somme des lignes, il faut avoir la 4-dépendance pour chaque ligne.

On voit donc dans cet exemple qu'il semble y avoir une amélioration : le nombre de 4-dépendances n'augmente pas alors que le nombre de possibilités augmente, i.e. la probabilité

diminue. Mais il faudrait voir si c'est bien le cas sous toutes les hypothèses possibles.

Maintenant, l'exemple précédent est un cas de p-dépendance lorsque p est petit.

Supposons maintenant que nous soyons dans le cas de grandes dépendance pour $p \geq 2N/I_3$ où I_3 est le nombre de lignes des nouveaux décalages. Dans ce cas, il semble possible qu'il y aie des dépendances. Mais on peut y appliquer les résultats déja vus en section 4 : il ne sera pas possible de les trouver.

12 Applications

12.1 Indécryptabilité des dépendances

On a vu que dès que les paramêtres sont bien choisis, il y a très peu de chances de trouver une dépendance au hasard.

Maintenant quelle serait l'inconvénient qu'un cryptanalyste trouve une p-dépendance? Eh bien, il pourrait casser une partie de la suite x_n. Il trouverait par exemple une relation du type $x_{j_4}^{d^4} = \overline{x_{j_1}^{d^1} + x_{j_2}^{d^2} + x_{j_3}^{d^3}}$ modulo m entre 4 blocs. Il pourrait donc prévoir pendant un certain temps $x_{j_4}^{d^4}$ si il connait $x_{j_1}^{d^1}$, $x_{j_2}^{d^2}$, $x_{j_3}^{d^3}$. Mais ce sera pendant un certain temps seulement : quand on change de blocs, les décalages changent. Donc, la dépendance change.

Remarquons que, si on a des permutations au lieu de décalages, on ne pourra pas prévoir aussi longtemps $x_{j_4}^{d^4}$, voire on ne pourra pas le prévoir du tout. Cela dépend du type de permutation.

De toutes façons, même avec les décalages, il sera extrêmement difficile de trouver au hasard de telles prévisions.

A l'heure actuelle, le meilleur ordinateur a une puissance de 10^{15} opérations par seconde (RoadRunner)

Si $N = 10^7$, I=20+1, il y a au maximum une probabilité de l'ordre de $1/N^I = 1/10^{140}$ de trouver une dépendance au hasard (4 termes sur 2 blocs). On peut donc la considérer comme introuvable.

On ne peut donc espérer casser le système de cryptage (ou plutôt une de ses parties) par les dépendances.

12.2 Cas où $\{x_{i,j}\}$ est connue

Normalement l'algorithme $\mathcal{B}$ est incassable *même quand* $\{x_{i,j}\}$ *est connue.* En effet, la question est : si on a une suite x_j^d et la matrice $\{x_{i,j}\}$, peut on retrouver le décalage d? Eh, bien, normalement, non! En effet ce problème est une variante du "Subset Sum Problem" (cf p 117-122 [9]). Il faut en effet retrouver les termes dont les sommes modulo m sont égaux a x_j^d.

12.3 Cas où $\{x_{i,j}\}$ est inconnue

Il est clair que, si le système est indécryptable *même lorsque* $\{x_{i,j}\}$ *est connue*, il le sera de façon absolue lorsque $\{x_{i,j}\}$ est inconnue. En fait, dans ce cas, on a une sécurité presque équivalente à celle du Vernam One Time Pad.

En effet, supposons que l'on veuille casser le système. A cause des propriétés d'aléarité de $\{x_{i,j}\}$, la seule technique d'attaque possible est probablement l'attaque par force brute.

D'abord on doit trouver $\{x_{i,j}\}$. Par exemple, si $\{x_{i,j}\}$ est une matrice de 10^8 chiffres, on doit essayer chaque matrice possible $\{x_{i,j}\}$. Or, il y a $10^{100.000.000}$ matrices possibles $\{x_{i,j}\}$. C'est un nombre qui peut être considéré comme infini. Il est donc impossible de casser le système. Même un ordinateur quantique faisant $10^{1.000}$ opérations à la seconde ne le casserait pas, et de loin!

En fait, on devra essayer toutes les matrices $\{x_{i,j}\}$ possibles avec toutes les clefs possibles.

12.4 Cas où $\{x_{i,j}\}$ peut etre considérée comme inconnue

Supposons que $\{x_{i,j}\}$ soit connue. Alors, une première transformation peut être faite : on crypte $\{x_{i,j}\}$ en une seconde matrice aléatoire $\{x'_{i,j}\}$. Ce sera $\{x'_{i,j}\}$ qui sera utilisée avec l'algorithme $\mathcal{B}$: $\{x'_n\} = \mathcal{B}(\{x'_{i,j}\})$. De très longues clefs peuvent être utilisées pour que tout décryptage soit impossible.

Quand $x'_{i,j}$ a été obtenue, on peut la considerer comme inconnue. De cette façon, on a une securité plus proche de celle du VOTP.

12.5 Un moyen pour utiliser le Vernam One Time Pad

C'est le cas où la matrice $\{x_{i,j}\}$ est inconnue ou peut être considérée comme inconnue. On a ainsi une méthode pour utiliser autrement le Vernam One Time Pad.

12.5.1 Avantage par rapport au Vernam One Time Pad

Si l'algorithme $\mathcal{B}$ est utilisé au lieu du VOTP, on a plusieurs avantages.

1) On n'a pas à échanger plusieurs fois la matrice secrète $\{x_{i,j}\}$: la suite x_n a une taille suffisante pour n'importe quel utilisation.

2) Il n'y a pas besoin de beaucoup de mémoire pour le stockage de $\{x_{i,j}\}$.

3) Le calcul de x_n est rapide.

12.5.2 Exemple : communication téléphonique

Supposons que l'on veuille téléphoner avec une sécurité absolue.

Alors, le téléphone crypte $\{x_{i,j}\}$. Il la transforme en une autre matrice aléatoire $\{x'_{i,j}\}$: c'est le cas 12.4. Pour cela, 10 ou 20 secondes seront nécessaires (plus suivant la sécurité désirée).

Alors, la conversation est chiffrée par $x'_n = \mathcal{B}(\{x'_{i,j}\})$. La vitesse de $\mathcal{B}$ permet de crypter une conversation téléphonique sans problème.

Puisque la matrice $\{x_{i,j}\}$ peut être considérée comme inconnue, on peut téléphoner avec une sécurité absolue.

12.5.3 Exemple : Réseau avec un ordinateur central

Dans ce cas, la matrice $\{x_{i,j}\}$ est seulement dans la mémoire de l'ordinateur central. Aussi $\{x_{i,j}\}$ est inconnue : on a la sécurité absolue de 12.3

Maintenant, chaque ordinateur périphérique "q" contient une suite aléatoire $\{x_n^{d^q}\} = \mathcal{B}_{d^q}(\{x_{i,j}\})$. Donc le système est sûr même si un grand nombre de suites $\{x_n^{d^q}\}$ sont connues par un cryptanalyste.

Natuellement, pour communiquer, l'ordinateur central et l'ordinateur "q" utilisent $\{x_n^{d^q}\}$: $C = M \oplus \{x_n^{d^q}\}$ [1] .

12.6 Authentification

Avec cette algorithme $\mathcal{B}$, on peut aussi construire une méthode d'authentification des données bien que l'on sache que le VOTP ne peut pas servir directement à l'authentification.

12.6.1 Rappels : authentification et masque jetable

On admettait jusque là que le Vernam One Time Pad ne pouvait pas servir à l'authentification des données pour la raison suivante.

Supposons qu'un message M soit authentifié par un message crypté C utilisant un masque jetable x(n) : C=x+M. Un tripatouilleur habile pourrait avoir déposé 10 euros sur son compte. La banque envoie le message M disant que 10 euros ont été déposés sur le compte N. Le tripatouilleur intercèpte le message M et connait donc le masque jetable x égal à C-M. Il ne lui reste qu'à changer la somme qu'il s'attribue (100 000 euros par exemple) et envoyer le message correspondant.

12.6.2 Méthode d'authentification

Pour avoir une méthode d'authentification sûre, il faudra donc utiliser le VOTP d'une manière un peu plus compliquée. On va maintenant donner un exemple pour montrer comment on peut authentifier un message M.

[1] Si la taille de M est trop grande on remplace $\{x_n^{d^q}\}$ par $\{\varsigma_n^q\} = \mathcal{B}_{d'}(\{x_n^{d^q}\})$.

On veut authentifier un n-ème message $M \in \{0, 1, ..., 10^8 - 1\}$. On dispose d'une matrice aléatoire $\{x(i,j)\}$, i=1,2,3, $j = 1, 2, ..., 10^{10}$, où $x(i,j) \in \{0, 1, ..., 10^8 - 1\}$.

Pour définir le décalage de la ligne 2, on utilise $\overline{aM + x(3,n)}$ modulo $m = 10^8$ où a est choisi de façon à ce que la congruence $T(x) \equiv ax$ mod m soit de période maximum.

On pose $d_2(n, M) = \lfloor \frac{10^{10}(\overline{aM+x(3,n)}\ )}{m} \rfloor$.

On a donc une ligne décalée $x^{d_2}(2, j)$, $j = 1, 2, ..., 10^{10}$. Soit $x^{d_2}(2, n)$ le n-ème terme de cette ligne.

On definit alors le code d'authentification numero n par $Ca = \overline{M + x(1,n) + x^{d_2}(2,n)}$ modulo 10^8 [2].

Pour authentifier le message on envoie $[Ca, M, n]$.

On pourrait avoir des systèmes d'authentification plus simples, par exemple a=1 ou supprimer la troisième ligne. Cela ferait gagner du temps. Mais cela diminuerait la securité *théorique* du système (mais pas forcément pratique). De toutes façons, ces calculs sont très rapides.

[2] On peut aussi poser r $Ca = \overline{x(1,n) + x^{d_2}(2,n)}$.

References

[1] JOHNSON L.N. KOTZ S. (1969) Discrete distributions, Wiley, New York.

[2] BLUM L, BLUM M. and SHUB M. (1986) A Simple Unpredictable Pseudo-Random Number Generator, SIAM Journal on Computing, volume 15, pages 364383, May 1986.

[3] KELSEY J. SCHNEIER B. and FERGUSON N. Ferguson (1999) Yarrow-160: Notes on the Design and Analysis of the Yarrow Cryptographic Pseudorandom Number Generator Sixth Annual Workshop on Selected Areas in Cryptography, Springer Verlag.

[4] JENKINS R. I. (1996) Isaac. Lecture Notes in Computer Science, volume 1039 , pages 4149. Springer,1996.

[5] FERGUSON N. SCHNEIER B. (2003) Practical Cryptography, published by Wiley in 2003.

[6] BUCHMANN J. A. (2000) Introduction to cryptography. Springer, New York.

[7] KNUTH D.E. (1998) the Art of Computer Programming; Vol 2. Third Edition Addison-Wesley, Reading, Massachusetts.

[8] GENTLE J. (1984) Random Number Generation and Monte Carlo Method, Springer 13, 61-81.

[9] MENEZES A., VAN OORSCHOT P. , VANSTONE S. (1996) Handbook of Applied Cryptography, CRC Press, 1996.

[10] SCHNEIER B (1996) Applied Cryptography 2nd Edition, John Wiley and sons, Inc

[11] BLACHER R. (2009) A Perfect Random Number Generator. Rapport de Recherche LJK Universite de Grenoble. http://hal.archives-ouvertes.fr/hal-00426555/fr/

[12] BLACHER R. (2010) A Perfect Random Number Generator II. Rapport de Recherche LJK Universite de Grenoble. http://hal.archives-ouvertes.fr/hal-00443576/fr/.

[13] BLACHER R. (2010) Correct models. Rapport de Recherche LJK Universite de Grenoble. http://hal.archives-ouvertes.fr/.