Remise à niveau – Traitement du signal

Énoncé de la séance N°2

Dans un premier temps, il s'agira de rappeler quelques notions sur le codage de Hamming. Ces notions seront ensuite appliquées à quelques exemples concrets en utilisant le logiciel Matlab. Ce document ainsi que le matériel nécessaire pour cette séance sont librement inspirés du livre [Mac03] et sont disponibles à l'adresse suivante : `http://nicolaslerme.fr/`.

1 Éléments de compréhension

Dans le TP précédent, nous avons vu comment les codes binaires à répétition pouvaient détecter et corriger des erreurs. Nous avons également vu qu'augmenter le nombre de bits de parité permet de détecter/corriger un plus grand nombre d'erreurs mais s'accompagne malheureusement d'une diminution rapide du taux R. Peut-on faire mieux que les codes binaires à répétition ? Que se passerait-il si l'on réalisait l'encodage d'un par bloc plutôt que par bit ? Les codes de Hamming permettent de répondre à ces questions.

1.1 Cadre général

Les codes de Hamming appartiennent à la famille des codes correcteurs d'erreurs linéaires où chaque bloc de k bits est transformé en un bloc de n bits. Certaines instances de ces codes ont la particularité d'être parfaits, dans le sens où ils permettent d'atteindre le plus grand rendement $\frac{k}{n}$ possible. À l'origine, ces codes ont été inventés par R.W. Hamming en 1950 pour corriger les erreurs dans des lecteurs de cartes perforées.

1.2 Hamming$(7,4)$

Dans son article fondateur [Ham50], l'auteur focalise les développements des codes de Hamming sur cas particulier : le code Hamming$(7,4)$. Ce code ajoute 3 bits de parité à chaque bloc de 4 bits de données traité. Un code par blocs peut être vu comme une règle permettant de convertir une séquence de bits $\mathbf{s}$ de longueur k en une séquence de bits $\mathbf{t}$ de longueur n. Dans un code par blocs linéaire, les $n-k$ bits supplémentaires (appelés bits de parité) sont des fonctions linéaires des k bits source.

L'opération d'encodage peut être représentée graphiquement (voir figure 1). Les 4 premiers bits transmis $t_1t_2t_3t_4$ sont égaux aux bits source $s_1s_2s_3s_4$. Les bits de parité $t_5t_6t_7$ sont fixés de manière à ce que la parité au sein de chaque cercle soit paire (donc 0 si la somme des bits du cercle auquel il appartient est paire, 1 sinon) : le bit t_5 est la parité des trois premiers bits de

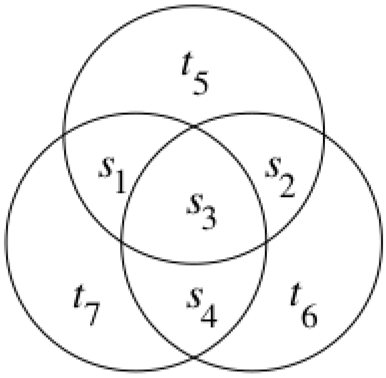

FIGURE 1 – Représentation schématique du code Hamming$(7,4)$.

$\mathbf{s}$; le bit t_6 est la parité des trois derniers bits de $\mathbf{s}$; le bit t_7 est la parité des bits s_1, s_3 et s_4. Pour un vecteur $\mathbf{s} \in \{0,1\}^4$ donné, le vecteur $\mathbf{t} \in \{0,1\}^7$ qui lui correspond s'obtient par

$$\mathbf{t} = \mathbf{G}^T \bullet \mathbf{s}, \tag{1}$$

où

$$\mathbf{G}^T = \begin{bmatrix} \mathbf{I}_4 \\ \mathbf{P} \end{bmatrix},$$

est une matrice binaire de taille 7×4 appelée matrice génératrice du code, $\mathbf{P}$ est une matrice de taille 3×4 et $\bullet$ est l'opérateur de multiplication matricielle modulo 2 (i.e. $1+1=0$ et $0+1=1$, etc.). Dans l'expression (1), $\mathbf{s}$ et $\mathbf{t}$ sont des vecteurs colonne. Si ce sont des vecteurs ligne, l'expression (1) est alors remplacée par

$$\mathbf{t} = \mathbf{s} \bullet \mathbf{G}, \tag{2}$$

où $\mathbf{G}$ est cette fois-ci une matrice de taille 4×7. Les lignes de $\mathbf{G}$ peuvent alors être vues comme une base de 4 vecteurs dans un espace binaire à 7 dimensions. Les 16 mots code peuvent alors être obtenus en effectuant toutes les combinaisons linéaires possibles de ces vecteurs.

La partie décodage est moins évidente. Supposons que certains bits du vecteur transmis $\mathbf{t}$ soient altérés, le vecteur reçu étant $\mathbf{r}$. L'idée est de vérifier si la parité des bits r_5, r_6 et r_7 est violée ou pas. On peut alors facilement créer un vecteur $\mathbf{z} \in \{0,1\}^3$ (appelé syndrome) où chaque élément vaut 1 si la parité est violée et 0 sinon. Pour le décodage, il suffit alors de répondre à la question suivante : peut-on trouver un unique bit qui se trouve à l'intérieur de tous les cercles où la parité est violée et à l'extérieur de tous les cercles où la parité n'est pas violée ? Si c'est le cas, le retournement de ce bit représenterait le syndrome observé. Si le syndrome $\mathbf{z}$ est nul, tous les bits de parité sont ok et le décodage le plus probable est donné par $r_1r_2r_3r_4$. Si le syndrome $\mathbf{z}$ est non nul, alors le vecteur transmis $\mathbf{t}$ a été altéré et $\mathbf{z}$ nous indique le motif d'erreur le plus probable. Pour un vecteur reçu $r \in \{0,1\}^7$, le syndrome $\mathbf{z}$ est obtenu ainsi :

$$\mathbf{z} = \mathbf{H}\mathbf{r},$$

où

$$\mathbf{H} = \begin{bmatrix} -\mathbf{P} & \mathbf{I}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{I}_3 \end{bmatrix},$$

est une matrice de taille 3×7 appelée matrice de contrôle de parité. Remarquons que tous les mots code $\mathbf{t} = \mathbf{G}^T\mathbf{s}$ du code satisfont

$$\mathbf{Ht} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

2 Exercices

Question 0 : Montrez que le code de Hamming(7,4) est un code linéaire.

Question 1 : Étant donné le vecteur $\mathbf{s} = (1, 0, 0, 0)^T$, calculez le vecteur transmis $\mathbf{t}$ (en supposant un bruit nul) et donnez sa représentation graphique en vous aidant de la figure 1.

Question 2 : Pour chaque vecteur $\mathbf{s} \in \{0, 1\}^4$, calculez le vecteur transmis $\mathbf{t}$ associé (en supposant toujours un bruit nul). Combien existe-t-il de mots code différents ?

Question 3 : Donnez la matrice génératrice $\mathbf{G}$ et sa transposée $\mathbf{G}^T$.

Question 4 : On considère le vecteur transmis $\mathbf{t}$ de la question 1 et on suppose qu'il est altéré par un bruit non nul. Pour les vecteurs $\mathbf{r}$ reçus suivants, donnez la représentation graphique associée, indiquez les cercles où parité n'est pas respectée ainsi que le bit à retourner selon l'algorithme de décodage :

— $\mathbf{r} = (1, 1, 0, 0, 1, 0, 1)$ (1 bit de données bruité),
— $\mathbf{r} = (1, 0, 0, 0, 0, 0, 1)$ (1 bit de parité bruité),
— $\mathbf{r} = (1, 0, 1, 0, 1, 0, 1)$ (1 bit de données bruité),
— $\mathbf{r} = (1, 0, 1, 0, 1, 0, 0)$ (2 bits bruités).

Qu'observez-vous ? Conclure.

Question 5 : Listez les syndromes $\mathbf{z}$ possibles et le bit r_k associé à retourner pour corriger l'erreur.

Question 6 : Quelle est la distance de Hamming minimale ? Comparer le rendement et les capacités de détection/correction avec ceux du code binaire à répétition.

Question 7 : Donnez la matrice $\mathbf{H}$. Vérifiez sur plusieurs exemples que les étapes de codage/décodage retournent le résultat attendu. Implémentez ensuite ces algorithmes sous Matlab et testez sur ces exemples.

Références

[Ham50] R.W. Hamming. Error detecting and error correcting codes. *Bell System Technical Journal*, 29(2) :147–160, 1950.

[Mac03] D.J.C. Mackay. *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, 2003.