

Codes de distance trois et davantage

Malika More

(malika.more@iut.u-clermont1.fr)

Alex Esbelin

(Alex.Esbelin@math.u-bpclermont.fr)

IREM de Clermont-Ferrand

Formation ISN

30 Juin 2011

Introduction

- Avec un code de distance minimale $d \geq 3$, on est sûr que les messages erronés comportant une seule erreur (les plus courants) sont correctement corrigés
- C'est pourquoi dans la pratique, les codes qui vérifient cette propriété sont particulièrement intéressants
- Dans ce chapitre, on étudie en détails certains codes de distance minimale $d = 3$, puis on donne quelques exemples de codes de distance minimale $d > 3$

Plan du cours

- 1 Codes de Hamming
- 2 Codes de distance minimale supérieure à trois

1 Codes de Hamming

2 Codes de distance minimale supérieure à trois

Code de Hamming $\mathcal{H}_{n,m}$

Définition

La matrice de contrôle H contient TOUTES les colonnes non nulles UNE SEULE FOIS

- Redondance $r = 2$
 - 3 colonnes de 2 bits non nulles
 - Longueur des mots de code $n = 3$

$$H = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

- Attention !

Respecter la position de la matrice identité I_r

Code de Hamming $\mathcal{H}_{n,m}$

Définition

La matrice de contrôle H contient TOUTES les colonnes non nulles UNE SEULE FOIS

- Redondance $r = 3$
 - 7 colonnes de 3 bits non nulles
 - Longueur des mots de code $n = 7$

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \text{ ou } H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

ou ...

Code de Hamming $\mathcal{H}_{n,m}$

Définition

La matrice de contrôle H contient TOUTES les colonnes non nulles UNE SEULE FOIS

- Redondance $r = 4$
 - 15 colonnes de 4 bits non nulles
 - Longueur des mots de code $n = 15$

$$H = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

ou ...

Code de Hamming $\mathcal{H}_{n,m}$

Définition

La matrice de contrôle H contient TOUTES les colonnes non nulles UNE SEULE FOIS

- Redondance r fixée
 - $2^r - 1$ colonnes de r bits non nulles
 - Longueur des mots de code $n = 2^r - 1$

Il y a plusieurs codes de Hamming $\mathcal{H}_{n,m}$ différents selon l'ordre dans lequel on place les colonnes de la matrice de contrôle H

(Mais on doit respecter la position de la matrice identité I_r)

Distance minimale des codes de Hamming

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Colonne nulle? NON donc $d \geq 2$
- Deux colonnes égales? NON donc $d \geq 3$
- Trois colonnes de somme \oplus nulle? OUI
Par exemple $C_5 \oplus C_6 = C_2$ donc $d = 3$

Conclusion

Un code de Hamming a TOUJOURS une distance minimale $d = 3$
Il permet de détecter deux erreurs et d'en corriger une

Message comportant plus d'une erreur

- Nombre de messages possibles : 2^n
 - Nombre de mots de code : $2^m = 2^{n-r}$
 - Nombre de messages erronés : $2^n - 2^m = 2^n - 2^{n-r}$
- Nombre de messages (erronés) à distance 1 d'un mot de code donné : $n = 2^r - 1$
- Nombre de messages à distance 1 d'un mot de code quelconque : $n \times 2^m = (2^r - 1) \times 2^{n-r} = 2^n - 2^{n-r}$
 - Tous les messages erronés sont à distance 1 d'un mot de code

Conclusion

Avec un code de Hamming, TOUS les messages comportant plus d'une erreur sont INCORRECTEMENT corrigés

Probabilité d'exactitude après correction

Un bit reçu sur cent est faux i.e. $p = 0,01$	Code de Hamming $\mathcal{H}_{7,4,3}$ i.e. $n = 7$
--	---

- Probabilité qu'un message soit exact après détection et (éventuellement) correction

$$\begin{aligned} \mathcal{P}(X = 0) + \mathcal{P}(X = 1) &= \\ (1 - p)^n + n \times p \times (1 - p)^{n-1} &= \\ 0,99^7 + 7 \times 0,01 \times 0,99^6 &\approx 0,998 \end{aligned}$$

- 99,8% des messages sont exacts après détection et (éventuellement) correction

Rendement des codes de Hamming

- On souhaite construire un code utilisant r bits de contrôle et de rendement le plus élevé possible
- Pour corriger au moins une erreur, il faut $d \geq 3$, donc les colonnes de H doivent être non nulles et toutes différentes, donc $n \leq 2^r - 1$
- Rendement $\rho = \frac{m}{n} = \frac{n-r}{n} = 1 - \frac{r}{n}$
 - Pour r fixé, le rendement ρ est maximal lorsque $\frac{r}{n}$ est minimal, donc lorsque n est maximal
 - Comme $n \leq 2^r - 1$, la valeur maximale est atteinte pour les codes de Hamming

Conclusion

Si on veut corriger au moins une erreur, en utilisant r bits de contrôle, il n'existe pas de code avec un rendement plus élevé qu'un code de Hamming

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 1 : Liste des 2^r syndromes

Syndrome	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
Vecteur de correc- tion	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants

Syndrome	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
Vecteur de correc- tion	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteur nul : Syndrome nul

Syndrome	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
Vecteur de correc- tion	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteurs de poids 1 : Syndromes qui sont des colonnes de H

	0	0	0	0	1	1	1	1
Syndrome	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0
Vecteur de correc- tion								

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteurs de poids 1 : Syndromes qui sont des colonnes de H

	0	0	0	0	1	1	1	1
Syndrome	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
Vecteur de correc- tion	0	1	0	0	0	0	0	0

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteurs de poids 1 : Syndromes qui sont des colonnes de H

	0	0	0	0	1	1	1	1
Syndrome	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0
Vecteur de correc- tion								

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteurs de poids 1 : Syndromes qui sont des colonnes de H

	0	0	0	0	1	1	1	1
Syndrome	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0
Vecteur de correc- tion								

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteurs de poids 1 : Syndromes qui sont des colonnes de H

	0	0	0	0	1	1	1	1
Syndrome	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0
Vecteur de correc- tion								

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteurs de poids 1 : Syndromes qui sont des colonnes de H

	0	0	0	0	1	1	1	1
Syndrome	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0
Vecteur de correc- tion								

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteurs de poids 1 : Syndromes qui sont des colonnes de H

Syndrome	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
Vecteur de correc- tion	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Ligne 2 : Liste des vecteurs de correction correspondants
- Vecteurs de poids 1 : Syndromes qui sont des colonnes de H

Syndrome	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
Vecteur de correc- tion	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0

Tableau standard

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Les vecteurs de correction sont TOUS les vecteurs de poids 1

Syndrome	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1
Vecteur de correc- tion	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	1	0
	0	0	0	0	0	1	0	0
	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	0

Résumé

- Les codes de Hamming sont faciles à construire (matrice H)
- Ils ont des propriétés intéressantes (distance, probabilité, rendement, tableau standard)
- Ils ne sont pas très nombreux (n contraint par le choix de r)

r	2	3	4	5
$n = 2^r - 1$	3	7	15	31
$m = 2^r - 1 - r$	1	4	11	26
Type	$\mathcal{H}_{3,1,3}$	$\mathcal{H}_{7,4,3}$	$\mathcal{H}_{15,11,3}$	$\mathcal{H}_{31,26,3}$
$P_{exact(*)}$	0,999	0,998	0,990	0,962
ρ	0,33	0,57	0,73	0,83

(*) avec un taux d'erreurs de 1%

1 Codes de Hamming

2 Codes de distance minimale supérieure à trois

Codes de distance minimale supérieure à trois

- Un des défauts des codes de Hamming, c'est qu'ils ne permettent de corriger qu'une seule erreur par message, alors que les erreurs sont souvent groupées (rayure, perturbation atmosphérique)
- Des codes de distance minimale $d > 3$ sont donc plus intéressants que les codes de Hamming en matière de correction d'erreurs, mais ils sont plus difficiles à construire
- Ceux qu'on utilise sont basés sur des théories mathématiques sophistiquées (comme la notion de polynômes sur un corps fini).
- En pratique, les blocs sont rapidement trop longs pour être traités à la main

Codes pseudo-Hamming

Définition

La matrice de contrôle H contient les colonnes non nulles AU PLUS UNE FOIS

- Redondance r fixée
 - $2^r - 1$ colonnes de r bits non nulles POSSIBLES
 - Colonnes de la matrice identité I_r OBLIGATOIRES (et dans le bon ordre)
 - Longueur des mots de code $r \leq n \leq 2^r - 1$

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \text{ ou } H = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

ou ...

Distance minimale des codes pseudo-Hamming

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- Colonne nulle? NON donc $d \geq 2$
- Deux colonnes égales? NON donc $d \geq 3$
- Trois colonnes de somme \oplus nulle? PAS FORCÉMENT

Conclusion

- Un code pseudo-Hamming a TOUJOURS une distance minimale $d \geq 3$
- Un code pseudo-Hamming a PARFOIS une distance minimale $d > 3$

Distance minimale des codes pseudo-Hamming

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- Colonne nulle? NON donc $d \geq 2$
- Deux colonnes égales? NON donc $d \geq 3$
- Trois colonnes de somme \oplus nulle? PAS FORCÉMENT

Conclusion

- Un code pseudo-Hamming a TOUJOURS une distance minimale $d \geq 3$
- Un code pseudo-Hamming a PARFOIS une distance minimale $d > 3$

Exemple de code de distance 4

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- Les codes de distance 4 ne sont pas encore trop difficiles à fabriquer
- Malheureusement, ils ne corrigent pas mieux les erreurs que les codes de Hamming

Exemple de code de distance 5

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- Les codes de distance 5 permettent de corriger deux erreurs
- Malheureusement, ils commencent à être assez difficiles à fabriquer "à la main" (c'est-à-dire sans faire appel à des théories mathématiques élaborées)

FIN