

ANTOINE LOBSTEIN

GÉRARD COHEN

Sur la complexité d'un problème de codage

Informatique théorique et applications, tome 21, n° 1 (1987), p. 25-32

http://www.numdam.org/item?id=ITA_1987__21_1_25_0

© AFCET, 1987, tous droits réservés.

L'accès aux archives de la revue « Informatique théorique et applications » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

SUR LA COMPLEXITÉ D'UN PROBLÈME DE CODAGE (*)

par Antoine LOBSTEIN ⁽¹⁾ et Gérard COHEN ⁽¹⁾

Communiqué par J. E. PIN

Résumé. – Dans un article publié en 1978, Berlekamp et al. ont conjecturé l'inexistence d'un algorithme polynômial calculant la distance minimale d'un code linéaire (si $P \neq NP$). Nous renforçons ici la plausibilité de cette conjecture, en montrant la NP-complétude de quelques problèmes voisins.

Abstract. – In a paper published in 1978, Berlekamp et al. conjecture the nonexistence of a polynomial algorithm for computing the minimum weight of a linear code (if $P \neq NP$). We here provide further evidence to support this conjecture, by proving NP-completeness of a few related problems, including:

- Problem Π_7 : the minimal weight codeword searched for must begin with a "1".
- Problem Π_8 : the minimal weight codeword searched for must have a fraction $p/(p+1)$ of all its "1" on its first components.

1. INTRODUCTION

Un code binaire linéaire C , de longueur n et de dimension k , peut être caractérisé par une matrice H , dite matrice de vérification de parité, de dimensions $(n-k) \times n$, de la manière suivante : $c \in C \Leftrightarrow cH^t = \mathbf{0} \pmod{2}$, $\mathbf{0}$ désignant le vecteur nul de longueur $n-k$. Le poids d'un vecteur x , noté $|x|$, désigne le nombre de composantes non nulles de x . La distance minimale d'un code linéaire C désigne le poids minimal d'un vecteur non nul de C .

Il s'agit là d'un paramètre essentiel des codes correcteurs; en effet, un code de distance minimale d peut corriger jusqu'à $\lfloor (d-1)/2 \rfloor$ erreurs.

(*) Reçu novembre 1985, révisé juillet 1986.

⁽¹⁾ École Nationale Supérieure des Télécommunications, 46, rue Barrault, 75634 Paris Cedex 13.

Le problème de *trouver la distance minimale d'un code linéaire* peut se mettre sous la forme du problème de décision suivant :

Π Instances : H matrice binaire, $w \in \mathbb{N}$.

Question : Existe-t-il un vecteur binaire non nul y tel que

$$yH^t = \mathbf{0} \pmod{2} \quad \text{et} \quad |y| \leq w?$$

En effet, si on connaît la distance minimale, d , du code C déterminé par H , on peut répondre à la question du problème Π en comparant w et d , et si on peut résoudre le problème Π , alors, en donnant successivement à w les valeurs $1, 2, \dots$, on trouve d lorsqu'on obtient pour la première fois la réponse « oui » à la question du problème Π .

Or, la question de savoir si Π est, ou n'est pas, un problème *NP-complet* n'a pas encore reçu de réponse. Berlekamp *et al.* [1] conjecturent que Π est *NP-complet*. Dans leur article, ils montrent notamment que si, dans Π , on remplace la condition « $|y| \leq w$ » par « $|y| = w$ », on a un problème *NP-complet*. De même, il a été montré de plusieurs problèmes, liés au problème Π , qu'ils étaient *NP-complets* :

Π_1 [1] Instances : H matrice binaire, $w \in \mathbb{N}$.

Question : Existe-t-il un vecteur binaire y tel que

$$yH^t = \mathbf{0} \pmod{2} \quad \text{et} \quad |y| = w?$$

Π_2 [3] Instances : H matrice binaire, $w \in \mathbb{N}$, $K \in \mathbb{N} \setminus \{0, 1\}$.

Question : Existe-t-il un vecteur binaire y tel que

$$yH^t = \mathbf{0} \pmod{2}, \quad 0 < |y| \leq w \quad \text{et} \quad |y| \not\equiv 0 \pmod{K}?$$

Π_3 [3] Instances : H matrice binaire, $w \in \mathbb{N}^*$.

Question : Existe-t-il un vecteur binaire y tel que

$$yH^t = \mathbf{0} \pmod{2} \quad \text{et} \quad |y| \geq w?$$

Π_4 [3] Instances : H matrice binaire, $w_1 \in \mathbb{N}$, $w_2 \in \mathbb{N}$, avec $0 < w_1 < w_2$.

Question : Existe-t-il un vecteur binaire y tel que

$$yH^t = \mathbf{0} \pmod{2} \quad \text{et} \quad w_1 \leq |y| \leq w_2?$$

Π_5 [4] Instances : H matrice binaire à n colonnes.

Question : Existe-t-il un vecteur binaire y tel que

$$yH^t = \mathbf{0} \pmod{2} \quad \text{et} \quad |y| = \lfloor n/2 \rfloor?$$

Rappelons également que le problème suivant est *NP-complet* :

Π_6 [1] *Instances* : H matrice binaire, y vecteur binaire, $w \in \mathbb{N}$.

Question : Existe-t-il un vecteur binaire x tel que

$$xH^t = y \pmod{2} \quad \text{et} \quad |x| \leq w?$$

Ce problème de décision correspond au *problème du décodage linéaire* : un mot c du code C ayant été émis sur un canal bruité, on reçoit un mot r entaché d'une erreur e : $r = c + e$. On évalue alors $y = rH^t = cH^t + eH^t = eH^t$. Le vecteur y est appelé *syndrôme* de r . Il ne dépend que de l'erreur e .

Pour estimer le mot de code émis, on fait $\hat{c} = r + x$, où x est solution, de poids minimal, de l'équation $xH^t = y$. En effet, $\hat{c}H^t = rH^t + xH^t = y + y = \mathbf{0}$: \hat{c} est le mot de code le plus proche de r .

2. COMPLEXITÉ DE QUELQUES PROBLÈMES VOISINS DE Π

Soit Π_7 le problème suivant, obtenu par une spécification additionnelle dans l'énoncé de Π .

Π_7 *Instances* : H matrice binaire, $w \in \mathbb{N}$.

Question : Existe-il un vecteur binaire x tel que $xH^t = \mathbf{0} \pmod{2}$, $|x| \leq w$ et la première composante de x , soit x_1 , vaut 1 ?

PROPOSITION 1: Π_7 est *NP-complet*.

Démonstration : Posons $w' = w + 1$ et

$$H' = \begin{array}{|c|c|} \hline y & H \\ \hline \end{array}$$

Alors s'il existe une solution au problème Π_7 posé avec H' et w' , c'est-à-dire s'il existe x' tel que $x'H'^t = \mathbf{0}$, $0 < |x'| \leq w'$ et $x'_1 = 1$, en prenant le vecteur x obtenu à partir de x' en supprimant la première composante, on a $|x| \leq w$ et $xH^t = y$. Inversement, s'il existe une solution au problème Π_6 , c'est-à-dire s'il existe x tel que $xH^t = y$ et $|x| \leq w$, en ajoutant 1 devant x , on obtient un vecteur x' vérifiant $x'H'^t = \mathbf{0}$, $0 < |x'| \leq w'$ et $x'_1 = 1$.

Donc s'il existait un algorithme en temps polynômial pour Π_7 , en l'appliquant à H' et w' , on disposerait d'un algorithme en temps polynômial pour Π_6 , qui est un problème *NP-complet*.

Donc Π_7 est *NP-complet* (car il est clairement dans *NP*).

Remarque : Le problème posé avec $x_1 = 0$ est le même que le problème posé sans condition sur x_1 , c'est-à-dire que sa complexité n'est pas connue.

Nous allons montrer la *NP*-complétude d'un autre problème voisin, Π_8 ; soit $p \in \mathbb{N}$, $p \geq 3$. Considérons le problème suivant, dit du *mariage* p -dimensionnel :

p-DM Instances : T ens. fini, $U \subseteq T \times T \times \dots \times T = T^p$.

Question : U contient-il un mariage? (i.e. existe-t-il $U' \subset U$ vérifiant $|U'| = |T|$ et 2 éléments quelconques de U' n'ont pas de coordonnée commune?)

Remarque : Un mariage contenant $|T|$ éléments, chaque élément de T apparaît une et une seule fois sur chaque composante.

Exemple :

$$p=5, \quad T = \{1, 2, 3, 4\}$$

$$U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$$

(1)

$$u_1 = (1, 1, 1, 1, 1)$$

$$u_2 = (1, 2, 4, 3, 4)$$

$$u_3 = (2, 1, 1, 2, 1)$$

$$u_4 = (3, 4, 4, 4, 2)$$

$$u_5 = (3, 4, 2, 1, 3)$$

$$u_6 = (4, 3, 3, 4, 2)$$

(2)

$$u_1 = (1, 1, 1, 1, 1)$$

$$u_2 = (2, 3, 4, 2, 3)$$

$$u_3 = (3, 4, 2, 4, 2)$$

$$u_4 = (4, 4, 2, 3, 1)$$

$$u_5 = (4, 1, 3, 3, 2)$$

$$u_6 = (3, 1, 2, 1, 3).$$

Alors (2) n'admet pas de 5-DM, alors que (1) admet (u_2, u_3, u_5, u_6) .

Pour $p \geq 3$, p -DM est *NP*-complet [2]. Nous allons maintenant réduire p -DM au problème suivant :

Π_8 Instances : H matrice binaire, $w \in \mathbb{N}$, $p \in \mathbb{N}$, $p \geq 3$.

Question : Existe-t-il un vecteur binaire y tel que

$$yH^t = \mathbf{0} \pmod{2}, \quad |y| \leq w \quad \text{et} \quad y_1 = y_2 = \dots = y_{\lfloor wp/(p+1) \rfloor} = 1?$$

(i.e. existe-t-il une somme de moins de w colonnes de H qui est égale au vecteur nul, et qui contient les $\lfloor wp/(p+1) \rfloor$ premières colonnes de H ?)

PROPOSITION 2 : Π_8 est NP-complet.

Démonstration : La réduction est la suivante : à partir d'une instance quelconque de p -DM, U, T , on construit la matrice H' , contenant $p \cdot |T|$ lignes et $|U|$ colonnes, en posant pour $j \in \{1, 2, \dots, |U|\}$, $i \in \{1, 2, \dots, p\}$ et $k \in \{1, 2, \dots, |T|\}$:

H' contient « 1 » en colonne j et en ligne $(i-1)|T| + k$ si et seulement si le j -ième p -uplet de U a k pour i -ième composante.

Posons Y_p égal au vecteur de longueur $p|T|$ dont toutes les composantes valent 1. On voit que U contient un mariage ssi on peut trouver $|T|$ colonnes de H' dont la somme vaut Y_p , c'est-à-dire ssi on peut trouver un vecteur y_* vérifiant $|y_*| = |T|$ et $y_* H'^t = Y_p$. Reprenons l'exemple (1) :

$$H' = \begin{array}{cccccc|c} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 4 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 1 & 1 & 0 & 8 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 1 & 0 & 10 \\ 0 & 0 & 0 & 0 & 0 & 1 & 11 \\ 0 & 1 & 0 & 1 & 0 & 0 & 12 \\ \hline 1 & 0 & 0 & 0 & 1 & 0 & 13 \\ 0 & 0 & 1 & 0 & 0 & 0 & 14 \\ 0 & 1 & 0 & 0 & 0 & 0 & 15 \\ 0 & 0 & 0 & 1 & 0 & 1 & 16 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 & 17 \\ 0 & 0 & 0 & 1 & 0 & 1 & 18 \\ 0 & 0 & 0 & 0 & 1 & 0 & 19 \\ 0 & 1 & 0 & 0 & 0 & 0 & 20 \end{array}$$

$y_* = (011011)$.

$y_* H'^t = Y_5$ (c'est-à-dire la somme des colonnes 2, 3, 5, 6 vaut Y_5).

Posons maintenant

$$H = \left(\underbrace{Id_{p \cdot |T|}}_{p \cdot |T|} \mid \underbrace{H'}_{|U|} \right) \} p \cdot |T|$$

et $w = (p+1) \cdot |T|$ (la réduction est bien polynômiale en $|T|$). Montrons qu'il existe une solution au problème p -DM, posé avec U et T , si et seulement si il existe une solution au problème Π_8 , posé avec H et w :

(a) p -DM $\Rightarrow \Pi_8$. Supposons qu'il existe un mariage contenu dans U . Alors, il existe un vecteur y_* (de longueur $|U|$), de poids $|T|$, et vérifiant $y_* H^t = Y_p$.

Posons $y = (\underbrace{1 \dots 1}_{p \cdot |T|} y_*)$. Alors y est solution de $\Pi_8 : y H^t = \mathbf{0}$,

$$|y| = p \cdot |T| + |y_*| = (p+1) \cdot |T| = w$$

et, comme $p \cdot |T| = wp/(p+1)$, on a

$$y_1 = y_2 = \dots = y_{|wp/(p+1)|} = 1.$$

(b) $\Pi_8 \Rightarrow p$ -DM. Supposons que Π_8 admette une solution y . Les $\lfloor wp/(p+1) \rfloor$ premières composantes de y valent 1; posons donc $y = (11 \dots 1 y_*)$ où y_* est un vecteur de longueur $|U|$. Comme $|y| \leq w$, on a $|y_*| \leq w/(p+1) = |T|$. Comme $y H^t = \mathbf{0}$, on obtient $y_* H^t = Y_p$. Et en fait, comme toute colonne de H' contient exactement p fois la valeur 1, on a exactement $|y_*| = |T|$. Donc U contient un mariage.

Ceci achève la preuve de la NP-complétude de Π_8 , l'appartenance à NP étant immédiate.

Pour finir, montrons la NP-complétude du problème suivant, un peu plus éloigné du problème original Π que Π_7 et Π_8 , mais qui nous paraît intéressant.

Π_9 Instances : H matrice binaire, $w \in \mathbb{N}$.

Question : $\exists y, y H^t = \mathbf{0} \pmod{2}$, $|y| \leq w$ et $\exists z$, de support inclus dans le support de y , tel que

$$z H^t = \mathbf{1} \pmod{2}?$$

(où le support d'un vecteur est l'ensemble des coordonnées où il vaut « 1 » et $\mathbf{1}$ désigne le vecteur tout à 1).

PROPOSITION 3 : Π_9 est NP-complet.

Démonstration : On montre : 3 -DM $< \Pi_9$, c'est-à-dire on réduit 3 -DM à Π_9 . $U, T \rightarrow H'$ comme ci-dessus (avec $p=3$).

$$H = \left[\begin{array}{c|c} H' & \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \end{array} \right]$$

et $w = |T| + 1$.

Sol. à 3-DM

- somme de $|T|$ colonnes de H' égale au vecteur tout à 1;
- somme de $|T| + 1$ colonnes de H égale au vecteur nul, en ajoutant, à la somme précédente, la dernière colonne;
- solution à Π_9 .

Sol. à Π_9

→ $\exists y, yH^t = \mathbf{0}, |y| \leq |T| + 1$ et $\exists z$, de support inclus dans le support de y , tel que $zH^t = \mathbf{1}$.

(1) si y a « 1 » en dernière composante [$y = (y_* \ 1)$], alors y_*

- somme de moins de $|T|$ colonnes de H' égale à 1;
- solution à 3-DM.

(2) si y a « 0 » en dernière composante, z aussi, et donc z

- somme de moins de $|T|$ colonnes de H' égale à 1;
- sol. à 3-DM.

On conclut là encore en remarquant que $\Pi_9 \in NP$.

3. CONCLUSION

Nous avons montré l'inexistence (si $P \neq NP$) d'un algorithme polynômial de calcul de la distance minimale d'un code linéaire, soumis à une des deux restrictions suivantes :

- le mot de poids minimal cherché doit commencer par un « 1 » (problème Π_7);
- le mot de poids minimal cherché doit avoir une fraction $p/(p+1)$ de son nombre total de « 1 » sur ses premières composantes (problème Π_8).

Ce dernier résultat est presque optimal au sens suivant : si on remplace $w/(p+1)$ par $w-\lambda$, λ constante, alors le problème possède un algorithme polynômial : compléter de toutes les façons possibles avec au plus λ « 1 » le vecteur $\left(\begin{array}{c} \overbrace{11 \dots 1} \\ \longleftarrow \quad \longrightarrow \\ w-\lambda \end{array} \right)$ et tester l'appartenance au code. Le nombre de tests requis est $\sum_{i=0}^{\lambda} \binom{n-w+\lambda}{i} \simeq n^\lambda$. En faisant $\lambda \sim n/p$, au lieu de $\lambda = \text{Cte}$, on retrouve Π_8 .

BIBLIOGRAPHIE

1. E. R. BERLEKAMP, R. J. MACELIECE et H. C. A. VAN TILBORG, *On the Inherent Intractability of Certain Coding Problems*, I.E.E.E. Trans. on Information Theory, vol. IT-24, n° 3, mai 1978.
2. M. R. GAREY et D. S. JOHNSON, *Computers and Intractability: a Guide to the Theory of NP-Completeness*, San Francisco, Freeman, 1978.
3. S. C. INTAFOS et S. L. HAKIMI, *On the Complexity of Some Coding Problems*, I.E.E.E. Trans. on Information Theory, vol. IT-27, n° 6, novembre 1981.
4. R. L. GRAHAM, Communication personnelle.