COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL
ET SES APPLICATIONS

NICE du 16 au 21 JUIN 75

PROBABILITES LIEES AU DECODAGE D'UN CODE EN BLOCS. APPLICATION A DES PROBLEMES DE TRANSMISSION.

P. GODLEWSKI et G.D. COHEN

E.N.S.T. - 46 rue Barrault - 75634 PARIS Cedex 13 - Groupe Codes Correcteurs d'erreurs - C.E.T.H.E.D.E.C.

RESUME

Nous évaluons diverses expressions liées au décodage d'un code en blocs correcteur d'erreurs : probabilités de décodage, taux d'erreurs résiduelles, ..., dans le cas d'un canal binaire symétrique sans mémoire. Des expressions formelles sont obtenues, dépendant des caractéristiques du code (longueur des blocs, distance minimale d, distribution des poids), et du décodeur (capacité de correction t, $t < \frac{d}{2}$). Nous donnons ensuite, au moyen de développements limités, des formules d'approximation plus explicites.

Les systèmes de transmissions classiques, utilisant une voie de retour, ne font généralement appel qu'aux possibilités de détection. A l'aide des résultats précédents, nous étudions dans quels cas on peut combiner les propriétés de détection-correction d'un code pour accroître l'efficacité de la transmission. On considérera un modèle de transmission de blocs d'information avec accusé de réception et possibilité de répétition. On supposera le canal affecté par une combinaison d'erreurs aléatoires et par paquets. La longueur des blocs et la capacité de correction choisie pour le décodeur seront des paramètres.

SUMMARY

We evaluate several expressions dealing with the decoding of a linear block error-correcting code : decoding probability, residual error rate, ..., in the case of a binary symmetric memoryless channel. Formal expressions are obtained, depending on the code properties (block length, minimum distance d, weight distribution), and on the decoder (correction capacity t, $t < d/2$). We then give, by means of series expansions, more detailed approximation formulas.

Classical transmission systems, using feedback, most often consider only the detection possibilities. With the help of the previous results, we try to combine the detection-correction properties to increase transmission efficiency. We use a transmission model of information blocks with acknowledgement delay and possible repetition.

We suppose that the channel is corrupted by a combination of random and burst errors. Block length and correction capacity chosen for the decoder are parameters.

PROBABILITES LIEES AU DECODAGE D'UN CODE EN BLOCS.
APPLICATION A DES PROBLEMES DE TRANSMISSION.

1. INTRODUCTION

L'utilisation de codes en blocs correcteurs d'erreurs permet de se protéger des erreurs survenant dans la transmission de données binaires sur un canal bruyant. A un ensemble de k éléments binaires, on adjoint r = n-k e.b. de contrôle dépendant (généralement linéairement) des k premiers. De cette manière, on forme un bloc de n e.b.. Au décodage, la redondance ainsi introduite permet de détecter (et/ou corriger) dans un tel bloc certaines erreurs intervenant dans la transmission.

Dans un premier temps, nous allons évaluer la probabilité de décodage (P_d) et le taux d'erreurs résiduelles (T_{er}), après l'utilisation d'un (n,k) code en blocs, sur un canal binaire symétrique sans mémoire (C.B.S.).

Puis nous appliquerons ces résultats à des problèmes de transmissions. Lorsque l'on ne dispose pas de voie de retour ("feedback"), la seule solution envisageable est la correction en avant (F.E.C. : Forward Error Correction) ; mais, pour un système avec possibilité de retransmission, on a le choix entre plusieurs stratégies : la détection simple (A.R.Q. : Automatic Repeat Request) ou un compromis détection-correction (A.R.Q.-F.E.C.). A la réception, le décodeur correspondant est incomplet : il n'accepte l'information contenue dans un mot reçu que si la confiance qu'on peut lui accorder est "suffisante".

Deux modèles de canaux seront considérés : a) un C.B.S. de probabilité d'erreur p ; b) un C.B.S. sur lequel s'ajoutent indépendamment des paquets d'erreurs suivant une loi pseudo-poissonienne. Un problème sera donc de maximiser le taux effectif de transmission T_t (T.R.I.B. : Transfer Rate of Information Bits), sachant que le taux d'erreurs résiduelles doit être inférieur à β.

2. PROBABILITES LIEES AU DECODAGE D'UN CODE EN BLOCS SUR UN CANAL SANS MEMOIRE

C est un (n,k) code linéaire, c'est-à-dire un sous-espace vectoriel de dimension k de l'espace F_2^n (où $F_2 = \{0,1\}$ est le corps de Galois à 2 éléments). p(.) est le poids de Hamming, c'est-à-dire le nombre de composantes non nulles d'un n-uplet ; d(x,y) = p(x-y) est la distance de Hamming. Le code C, de distance minimale d, est donc susceptible de corriger $t_o = \lfloor \frac{(d-1)}{2} \rfloor$ erreurs, il sera caractérisé par son polynôme énumérateur de poids :

$A(z) = \sum_{i=o}^{n} A_i z^i$, où A_i est le nombre de mots de poids i.

$D_i(t)$ est le décodeur incomplet corrigeant jusqu'à t erreurs, $t \leqslant t_o$. Soit u le mot émis, e le (1,n) vecteur erreur ; $D_i(t)$ décode le mot reçu v = u+e si et seulement si il est situé à une distance inférieure à t d'un mot de code. Tous les calculs sont faits dans le cas d'un C.S.B. de probabilité d'erreur p, (q = 1-p).

Rappelons que :

- la probabilité pour que e soit égal à un mot donné de poids i est $p^i q^{n-i}$,
- la probabilité pour que le poids de e soit égal à i est $\binom{n}{i} p^i q^{n-i}$,
- la probabilité de décodage correct P_{dc} est la probabilité pour que le poids de e soit inférieur à t :

$$P_{dc} = \sum_{i=o}^{t} \binom{n}{i} p^i q^{n-i} \sim 1 - \binom{n}{t+1} p^{t+1} \left[1 - (n-t-1)\frac{(t+1)}{(t+2)} p\right] \quad (2.0)$$

2.a. Expressions formelles de P_d et de T_{er} .

Le calcul de P_d est dû à Mac Williams (cf. [1]) :

$$P_d = \sum_{i=o}^{t} f^{(i)}(0) / (i!)$$

où $f(x) = (px+q)^n A\left(\frac{qx+p}{px+q}\right)$. (2.1)

On montre de même que T_{er} peut s'écrire (voir [2] et [3]) :

$$T_{er} = \frac{1}{n} \sum_{i=o}^{t} g^{(i)}(0) / (i!) \quad P_d$$

où $g(x) = (px+q)^{n-1} (qx+p) A'\left(\frac{qx+p}{px+q}\right)$ (2.2)

2.b. Si $np << 1$ on peut développer ces expressions aux deux premiers termes significatifs. On appelle P_{de} la probabilité de décodage erroné : $P_d = P_{dc} + P_{de}$. Après des calculs explicités dans [2] et [3] on obtient :

$$P_{de} = A_d \binom{d}{t} p^{d-t} + \binom{d}{t} \left[A_{d+1} \frac{d+1}{(d-t+1)} - A_d \left\{n-d+t - \frac{t}{(d-t+1)}\right\}\right] p^{d-t+1} + \varepsilon(p^{d-t+1}) \quad (2.3)$$

Si $d=2t+1$, $t=t_o$:

$$P_d = 1 - p^{t+1}\left[\binom{n}{t+1} - A_{2t+1}\binom{2t+1}{t}\right] + p^{t+2}\left[\binom{n}{t+1}(n-t-1)\frac{(t+1)}{(t+2)} + \binom{2t+1}{t}\left\{A_{2t+2}\frac{2t+2}{t+2} - A_{2t+1}\left(n-t-1-\frac{t}{t+2}\right)\right\}\right] + \varepsilon(p^{t+2}) \quad (2.4)$$

Si d=2t+2, il suffit d'annuler A_{2t+1} dans cette dernière expression.

De même, le taux d'erreurs résiduelles peut s'écrire :

$$T_{er} = \frac{1}{n P_d}\left[d\, P_{de} + \binom{d+1}{t} A_{d+1} p^{d-t+1} + \varepsilon(p^{d-t+1}) \right] \quad (2.5)$$

Les expressions obtenues n'ont pas qu'un intérêt formel, On les utilise par exemple pour caractériser un système de transmission quand on s'impose un taux d'erreurs résiduelles très faible.

3. RECHERCHE D'UNE STRATEGIE DE TRANSMISSION

Le taux effectif de transmission T_t est le nombre d'e.b. d'information acceptés sur le nombre total d'e.b. transmis. Si N-1 est le nombre moyen de retransmissions d'un mot, alors, suivant un calcul classique :

$$N = P_d \cdot \sum_{i=1}^{\infty} i(1 - P_d)^{i-1} = 1/P_d, \qquad (3.1)$$

en désignant par a une quantité correspondant à l'accusé de réception :

$T_t = k/[n + (N-1)(n + a)]$ ou $T = k/N(n + a)$ (3.2)

suivant la procédure de transmission utilisée.

Le problème est de maximiser T_t sous la contrainte $T_{er} < \beta$. D'une manière générale, les paramètres sont n, t_o (capacité de correction du code), t (capacité de correction effectivement choisie : $t \leqslant t_o$). Nous appliquerons ce problème à une classe de codes précise : les BCH primitifs étendus : $n = 2^m$, $n-k = mt_o+1$.

3.a. Canal sans mémoire

Shannon a montré que l'emploi d'une ligne de retour ne pouvait pas améliorer la capacité d'un canal sans mémoire. Chercher l'optimum de T_t pour la classe des BCH étendus a donc peu d'intérêt même si l'on sait qu'un tel optimum existe puisque les BCH sont asymptotiquement mauvais. La solution ne peut offrir que des avantages négligeables sur le F.E.C. pur.

D'autres contraintes, de mémoire, de format (sur n) de technologie ou de coût (sur t) interviennent généralement dans la pratique. Ces cas ont fait l'objet de nombreux articles, par exemple $t_o = 1$ dans [5], ou $n = 255$, $k = 187$, $t_o = 9$ dans [4].

3.b. Erreurs par paquets

Considérons un C.B.S. sur lequel s'ajoutent des paquets d'erreurs. On suppose qu'un n-uplet est perturbé (sans possibilité de correction) par un paquet avec une probabilité μ, on pose $\lambda = 1 - \mu$. Nous appelons maintenant P'_d, T'_{er} les probabilités liées au décodage lorsque le canal est sans mémoire. On suppose qu'un mot entaché par un paquet est un vecteur aléatoire de l'espace des n-uples ; la probabilité de décodage d'un tel mot est :

$$P''_d = 2^k \left[\sum_{i=o}^{t} \binom{n}{i}\right]/2^n.$$

Si la répartition des paquets est pseudo-poissonienne, le taux d'erreurs résiduelles correspondant T''_{er} est la forme $P/n + \ell P + p + \varepsilon(t/n)$ si $\lambda \sim e^{-\ell n}$, P étant le poids moyen des paquets d'erreurs (Nous négligeons le fait qu'un paquet peut perturber deux mots consécutifs).

On a donc : $P_d = \lambda P'_d + \mu P''_d$

$$T_{er} = \frac{\lambda P'_d T'_{er} + \mu P''_d T''_{er}}{\lambda P'_d + \mu P''_d}$$

On cherche maintenant à évaluer cette expression de T_{er} en supposant que les termes dûs à la présence de paquets d'erreurs sont prédominants. Pour la classe des codes BCH étendus :

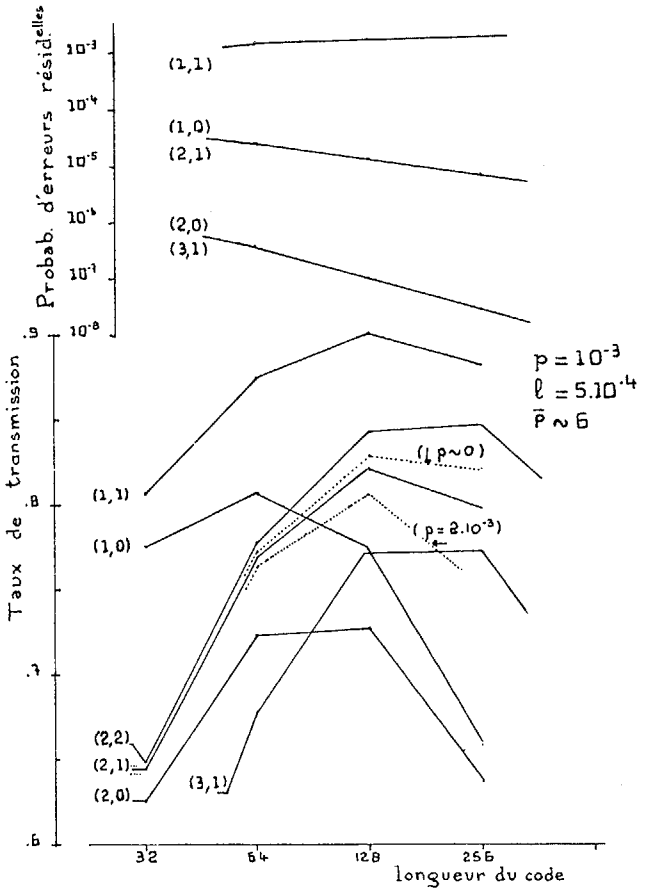
$P''_d \sim 1/t!\ 2^{m(t_o-t)}$

d'où $T_{er} \sim \mu P''_d T''_{er} \sim (1-e^{-\ell n})\left(\frac{p}{n} + \ell P + p\right)/2.t!\ 2^{n(t_o-t)}$

d'autre part,

$P_d \sim \lambda P'_d = e^{-\ell n} P'_d$ d'où $T_t = e^{-\ell n} P'_d \frac{k}{n}$ (si $a \sim o$)

et t_o fixé, un compromis est donc à faire entre le choix d'un code de longueur faible (et donc de rendement k/n peu élevé) et d'un code de grande longueur ("pénalisé" par $e^{-\ell n}$). On peut schématiser ainsi la solution choisie : le décodeur corrige les erreurs indépendantes ($P'_d \sim 1$) et détecte les paquets. Cette correction permet de limiter les retransmissions, et donc



la mémoire à l'émission : à performances égales (T_{er} et T_t égaux) la stratégie (n,t_o,t) sera supérieure à une stratégie $(n,t_o-1,t-1)$ (cf. exemple). D'autre part on se sert de "l'imperfection" des BCH pour améliorer la détection (terme 2.t! au dénominateur de T_{er}).

Exemple : on a tracé les variations de T_t et de T_{er} en fonction de n pour diverses stratégies (t_o,t), $p = 10^{-3}$, $\ell = 5.10^{-4}$, $P = 6$, $a \sim 0$.

4. CONCLUSION

Les théories classiques font généralement une dichotomie simplificatrice entre erreurs par paquets et erreurs aléatoires. Dans un cas, et si le canal a le bon goût d'être propre pendant des intervalles de garde suffisamment longs et réguliers, on peut employer un système FEC pur avec un code convolutionnel. Dans l'autre, le canal est supposé CBS, et le codage par blocs est universellement utilisé. Pour certains canaux affectés par un mélange de ces deux types de bruits, on peut parfois utiliser des méthodes simples: décorrélation des erreurs (entrelacement), ou plus sophistiquées : concaténation de codes. Si l'on n'a pas impérativement besoin de recevoir l'information à un débit constant, on a intérêt à utiliser les possibilités de retransmission. On obtient ainsi aisément une certaine adaptativité à la qualité de la voie et une plus grande indépendance vis à vis de la statistique détaillée des erreurs. Nous avons précisé dans 3., en fonction des erreurs survenant, dans quelle mesure le taux de transmission d'un système avec correction est supérieur à celui d'un système de détection simple. Si le canal est quasi-stationnaire, on peut tester les caractéristiques grâce aux propriétés de distance minimale du code, et ainsi adapter la capacité de correction à la voie.

REFERENCES

1 - E.R. BERLEKAMP, "Algebraic coding theory", New York : Mc Graw-Hill, 1968, pp.397-399.

2 - G.D. COHEN, "Développements limités des probabilités de décodage et de décodage erroné", Cahier spécial du CETHEDEC, 1974.1, pp.18-22.

3 - P. GODLEWSKI, "Probabilités liées au décodage d'un code en bloc", ibid, pp.10-16.

4 - J.R. EDWARDS, "Hybrid decoding for digital-communication-feedback systems", Proc. IEE, Vol.121, n°10, Octobre 1974, pp.1067-1075.

5 - J.H. PARK, "Information rates and errors using decision feedback", IEEE Trans. Com. Techn., Vol. COM-17, n°1, Février 1969, pp.20-24.