

SEPTIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 28 MAI au 2 JUIN 1979

SYSTEME ADAPTATIF DE COMMUNICATION NUMERIQUE UTILISANT DES CANAUX EN PARALLELE

G. Battail

Ecole Nationale Supérieure des Télécommunications, 46 rue Barrault, 75634 PARIS CEDEX 13

RESUME

Certains moyens de transmission, par exemple en ondes courtes avec sous-porteuses multiples ou à saut de fréquence, peuvent être représentés par plusieurs canaux peu fiables disponibles en parallèle. La pluralité de ces canaux constitue une redondance que l'on se propose d'exploiter pour améliorer la fiabilité de l'ensemble.

Dans cette perspective, nous considérons n canaux binaires symétriques indépendants en parallèle. Chacun d'eux est caractérisé par sa probabilité d'erreur moyenne p_i , $i = 1, 2, \dots, n$, inconnue mais supposée stationnaire ou lentement variable. La redondance nécessaire à la fiabilité globale est introduite par l'emploi d'un code redondant (n, k) , dont les symboles codés sont transmis par chacun des canaux. Nous nous proposons d'utiliser un décodage où chaque symbole est pondéré en fonction de sa probabilité d'erreur, en fonction de la fiabilité du canal qui sert à le transmettre. Cette fiabilité étant inconnue a priori, les facteurs de pondération des symboles reçus ne peuvent être déduits que du fonctionnement même du système, c'est-à-dire de manière adaptative.

A cet effet, nous utilisons un algorithme de décodage symbole par symbole dont le résultat lui-même est pondéré comme les symboles reçus. La décision finale sur chacun des symboles est comparée au même symbole avant décodage, reçu en sortie de l'un des canaux. La fréquence des différences, pondérée en fonction de la fiabilité de chaque décision, permet de déterminer le coefficient de pondération du canal correspondant.

On examine d'abord le cas le plus simple où un code $(n, 1)$, c'est-à-dire une simple répétition, est

SUMMARY

Certain transmission media e.g., HF radio with multiple sub-carriers, or frequency hopping, can be modelled as a number of poorly reliable channels in parallel. This plurality constitutes a redundancy, of which we intend to take advantage in order to improve the overall reliability.

According to this viewpoint, we consider n parallel independent binary symmetric channels. Each of them is characterized by its mean error probability p_i , $i = 1, 2, \dots, n$, which is unknown but assumed to be stationary or slowly time-varying. The needed redundancy is introduced through (n, k) coding and the resulting n symbols are transmitted in each channel. We intend to use a decoding algorithm where each symbol is weighted in terms of its error probability, depending on the reliability of the channel from which it has been received. Since this reliability is unknown a priori, the weighting coefficients can only derive from the system operation itself i.e., according to an adaptive process.

To this end, we use a symbol-by-symbol decoding algorithm, whose decision itself is weighted as the input symbols are. The decision made on each symbol is compared with the output of the corresponding channel. The frequency of the discrepancies, as weighted according to each decision likelihood, permits to determine the weighting coefficient of this channel.

The simplest case where an $(n, 1)$ code i.e., mere repetition, is used, is first considered. From the decision rule optimum in this case, we derive the algorithm to compute the weighting coefficients. Its convergence is discussed. We then generalize to the case where an (n, k) code is employed, using the concept of replication decoding.

SEPTIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS



NICE du 28 MAI au 2 JUIN 1979

RESUME

employé. On déduit de la règle de décision, optimale dans ce cas, l'algorithme de calcul des coefficients de pondération. On en discute la convergence. On généralise ensuite au cas où un code (n,k) est employé, à l'aide du concept de décodage par répliques.

SUMMARY

0 - Introduction

On considère la transmission de symboles binaires par un ensemble de n canaux peu fiables, utilisés en parallèle afin que la redondance compense leur mauvaise qualité individuelle. Chaque symbole peut être purement et simplement transmis en parallèle par l'ensemble des canaux, cas que nous désignerons par répétition, schématisé sur la figure 1. On peut aussi coder chaque bloc de k symboles avec un code de taux k/n, les symboles codés étant transmis par les n canaux (figure 2). L'ensemble de canaux peu fiables en parallèle que nous considérons est censé modéliser un certain nombre de voies physiques, notamment une voie radioélectrique en ondes courtes utilisant un système de modulation à sous-porteuses multiples, ou un système "à saut de fréquence" dans lequel on associe à chacune des fréquences employées un canal ayant ses caractéristiques propres.

Nous envisageons d'abord le cas de la répétition. Nous rappelons la règle de décision optimale dans ce cas, valable lorsque les erreurs dans les canaux en parallèle sont indépendantes et de probabilité connue. Nous examinons ensuite le cas où ces caractéristiques sont inconnues, stationnaires ou lentement variables. Nous décrirons un système de transmission dont le fonctionnement même permet d'apprécier la qualité des différents canaux, d'optimiser la décision en fonction de cette évaluation et de réaliser une adaptation dynamique lorsque la qualité des différents canaux évolue.

1 - Description du système utilisant la répétition

Nous considérons le schéma de communication de la figure 1. Un même symbole binaire est transmis en parallèle à travers n canaux symétriques, indépendants les uns des autres, dont les caractéristiques sont inconnues ou mal connues. Il s'agit de prendre une décision optimale quant au symbole émis.

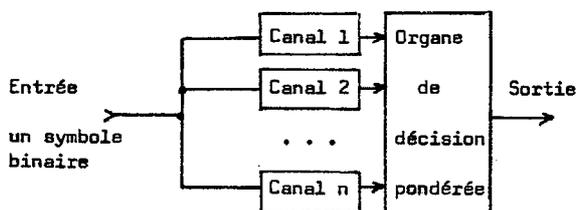


Figure 1 Système utilisant la répétition

1 - 1 - Décision optimale lorsque les caractéristiques sont connues

Supposons d'abord que les caractéristiques des canaux soient parfaitement connues, en ce sens que l'on dispose de la valeur exacte de la probabilité d'erreur sur le i-ème canal, soit p_i , $i = 1, 2, \dots, n$, à tout instant. La règle de décision optimale, au sens de la vraisemblance maximale, s'énonce ainsi [1] : nous définissons la "valeur relative a priori" du symbole binaire b_i issu du i-ème canal par :

$$(1) \quad a_i \triangleq s_i \log_e \frac{1 - p_i}{p_i}, \quad 0 < p_i \leq 1/2,$$

où s_i est le signe associé à b_i selon

$$(2) \quad s_i = (-1)^{b_i}, \quad b_i = 0, 1.$$

Alors, la décision à vraisemblance maximale est obtenue en calculant

$$(3) \quad A = \sum_{i=1}^n a_i.$$

Cette grandeur est la "valeur relative a posteriori" c'est-à-dire que le résultat de la décision à vraisemblance maximale sur l'ensemble des répliques du symbole émis, reçues à travers les n canaux, est représenté par le signe de A, selon (2). En outre, la probabilité d'erreur de cette décision est donnée par

$$(4) \quad P = 1/(1 + \exp|A|),$$

homologue de

$$(5) \quad p_i = 1/(1 + \exp|a_i|)$$

qui découle directement de (1).

On a montré [1] que cette décision comportait toujours une amélioration de la probabilité d'erreur moyenne par rapport à celle du meilleur des canaux, sauf dans quelques cas où ces probabilités d'erreur moyennes sont égales: si une seule des valeurs relatives a_i est différente de 0 (il n'y a alors en fait qu'un canal utilisable), ou si deux d'entre elles sont différentes de 0 et ont toutes deux une distribution de probabilité particulière, à deux valeurs discrètes nécessairement opposées. (Il y a cependant dans le dernier cas augmentation de l'information reçue par le destinataire, car les symboles résultant de la décision sont pondérés selon leur vraisemblance.)



G. Battail

1 - 2 - Évaluation de la qualité de chaque voie par comparaison avec le résultat de la décision

Supposons maintenant que la valeur exacte de la probabilité d'erreur p_i de chaque canal ne soit pas disponible, mais que l'on en connaisse une estimation \hat{p}_i . On peut encore appliquer la règle de décision (3) en y remplaçant a_i par \hat{a}_i , calculée d'après (1) en y remplaçant p_i par \hat{p}_i . Le résultat, soit \hat{A} , ne correspond évidemment plus à la décision optimale mais, dans la mesure où les estimations \hat{p}_i sont proches de la valeur vraie, cette décision améliore encore la probabilité d'erreur moyenne par rapport à celle du meilleur des canaux.

On peut alors envisager d'améliorer ainsi l'estimation de la probabilité d'erreur dans le i -ème canal: le résultat de la décision étant présumé meilleur que le symbole reçu en sortie du i -ème canal, la non coïncidence entre les deux signifie le plus souvent qu'une erreur s'est produite sur ce canal. Le module de la valeur relative a posteriori permet en outre de pondérer le résultat de la décision selon sa vraisemblance estimée. Le rapport du nombre (pondéré) de ces désaccords au nombre total des symboles émis, pendant un certain temps, fournira une nouvelle évaluation de la probabilité d'erreur dans le i -ème canal, soit \hat{p}_i , mieux approchée que l'estimation initiale \hat{p}_i . On substituera donc \hat{p}_i à \hat{p}_i , pour tout i , et le processus pourra être répété. On peut espérer, au moins sous certaines conditions, que cette estimation convergera vers la valeur vraie p_i de la probabilité d'erreur du canal i .

Si l'on suppose les canaux non plus stationnaires mais que leur probabilité d'erreur évolue très lentement à l'échelle du débit de symboles, on attend de ce procédé une adaptation aux conditions réelles de transmission.

1 - 3 - Relations de base

Dans le cas d'un accord entre la décision globale et le symbole reçu sur le i -ème canal, il n'y a erreur sur ce canal que si la décision est erronée, c'est-à-dire avec la probabilité P donnée par (4) si les pondérations initiales sont exactes. En cas de désaccord, l'erreur sur le canal a lieu si la décision est correcte, donc avec la probabilité $1 - P$.

La probabilité moyenne d'erreur dans le i -ème canal peut alors être exprimée par l'espérance mathématique

$$p_i = E \left[\frac{1}{2} (1 + s_i S) P + \frac{1}{2} (1 - s_i S) (1 - P) \right] \\ = \frac{1}{2} E \left[1 - s_i S (1 - 2P) \right],$$

où S est le signe de A et où P est donnée par (4). Mais on a d'après (4)

$$S(1 - 2P) = S(\exp\{A\} - 1) / (\exp\{A\} + 1) \\ = \text{th}(A/2);$$

on peut donc exprimer p_i ainsi:

$$(6) \quad p_i = \frac{1}{2} E \left[1 - s_i \text{th}(A/2) \right].$$

Si l'on ne dispose d'aucune autre information, on ne peut espérer déterminer que la fréquence moyenne de l'erreur, et donc obtenir pour la valeur relative a_i une expression de la forme

$$(7) \quad a_i = s_i v_i,$$

où v_i est une constante, égale à $\log_e \left[(1-p_i)/p_i \right]$.

On a donc alors

$$(8) \quad v_i = \log_e \frac{1 + E s_i \text{th}(A/2)}{1 - E s_i \text{th}(A/2)},$$

ce que l'on peut écrire encore

$$(9) \quad \text{th}(v_i/2) = E [s_i \text{th}(A/2)] = E [\text{th}(s_i A/2)].$$

Les relations équivalentes (8) et (9) nous serviront dans toute la suite.

Il se pourrait cependant que l'on dispose d'une information individuelle pour chaque symbole reçu, mais incomplète, par exemple la grandeur issue d'un démodulateur, soit x_i , égale à un facteur positif près à la valeur relative a priori. La valeur relative vraie serait alors

$$(10) \quad a_i = \lambda_i x_i, \quad \lambda_i > 0,$$

où λ_i est un facteur d'échelle constant mais inconnu.

Alors:

$$\lambda_i = (1/E|x_i|) E \left\{ \log_e \left[\frac{1 + \text{th}(s_i A/2)}{1 - \text{th}(s_i A/2)} \right] \right\} \\ = (1/E|x_i|) 2E \left\{ \text{argth} [\text{th}(s_i A/2)] \right\}$$

d'où finalement

$$(11) \quad \lambda_i = E(s_i A) / E|x_i|.$$

1 - 4 - Mise en oeuvre du procédé

Si les coefficients de pondération v_i ou les facteurs λ_i ont initialement leur valeur vraie, (9) ou (11) ne sont bien entendu que des identités sans intérêt pratique particulier.

Si par contre on remplace dans (9) ou (11) la valeur exacte de A par son estimation \hat{A} calculée selon (3), a_i y étant remplacée par son estimation \hat{a}_i , on attend de ces relations le moyen de réévaluer les

coefficients v_i ou λ_i . Les espérances mathématiques qu'elles comportent sont évidemment inaccessibles à l'expérience et devront y être remplacées par des moyennes temporelles qui, sous l'hypothèse de l'ergodicisme, en fourniront des évaluations.

1 - 5 - Analyse approximative de la convergence de l'algorithme tiré de (9)

Pour analyser la convergence de cet algorithme, nous allons supposer que les pondérations initiales sont voisines de la valeur exacte v_i , c'est-à-dire de la forme $\hat{v}_i = v_i + \delta v_i$, où δv_i est assez petit pour qu'une approximation du premier ordre soit valable. Nous supposons que le coefficient de pondération résultant de notre estimation est de même forme, soit $v_i + \delta'v_i$. Nous avons d'après (9):

$$\text{th} (v_i + \delta'v_i)/2 = E [\text{th}(s_i \hat{A}/2)],$$

avec, d'après (3):

$$\hat{A} \triangleq A + \delta A = \sum_{j=1}^n s_j (v_j + \delta v_j) = A + \sum_{j=1}^n s_j \delta v_j.$$

Dans une approximation du premier ordre, compte tenu de (9), nous avons donc:

$$\begin{aligned} \delta'v_i &= \frac{(\exp v_i + 1)^2}{\exp v_i} E \left[\frac{s_i \exp A}{(\exp A + 1)^2} \delta A \right] \\ &= \frac{(\exp v_i + 1)^2}{\exp v_i} E \left[\frac{s_i \exp A}{(\exp A + 1)^2} \sum_{j=1}^n s_j \delta v_j \right]. \end{aligned}$$

Ces relations s'expriment plus simplement sous la forme matricielle

$$\underline{\delta'v} = \underline{\delta v} \underline{M}$$

où $\underline{\delta'v}$ et $\underline{\delta v}$ sont les vecteurs-lignes ayant pour composantes $\delta'v_i$ et δv_i , $i = 1, 2, \dots, n$, respectivement, et où \underline{M} est la matrice carrée d'ordre n ayant pour élément général

$$m_{ij} = \frac{(\exp v_i + 1)^2}{\exp v_i} E [s_i s_j \exp A / (\exp A + 1)^2].$$

La probabilité d'une configuration de signes représentée par le vecteur ligne $\underline{s} = [s_1 \ s_2 \ \dots \ s_n]$, où chacune des composantes vaut +1 ou -1, vaut

$$2^{-n} \prod_{j=1}^n [1 + s_j \text{th}(v_j/2)]$$

ou encore

$$\prod_{j=1}^n 1/(1 + e^{-s_j v_j}),$$

car

$$\frac{1}{2} [1 + s_j \text{th}(v_j/2)] = 1/[1 + \exp(-s_j v_j)].$$

L'élément général de la matrice \underline{M} peut donc être écrit

$$m_{ij} = \frac{(e^{v_i} + 1)^2}{e^{v_i}} \sum_{\underline{s}} \left(\prod_{k=1}^n 1/(e^{-s_k v_k} + 1) \right) \frac{s_i s_j \exp \sum_k s_k v_k}{(\exp \sum_k s_k v_k + 1)^2}$$

où la première somme est étendue à tous les vecteurs \underline{s} possibles.

D'après (3) et (7), A dépend des valeurs prises par le vecteur \underline{s} . La fonction $e^x/(e^x + 1)^2$ étant paire, les valeurs de $e^A/(e^A + 1)^2$ sont égales pour des configurations de signes opposées. Ces configurations sont représentées par des vecteurs \underline{s} et $\underline{\bar{s}}$ tels que leur addition composante à composante donne le vecteur nul: $\underline{s} + \underline{\bar{s}} = \underline{0}$. Dans le calcul des espérances, on peut donc regrouper les valeurs de A qui correspondent à des vecteurs \underline{s} et $\underline{\bar{s}}$ opposés. En ne considérant plus dans les sommes que les vecteurs \underline{s} pour lesquels $s_i = +1$ et en effectuant ce regroupement, on obtient l'élément général, après calcul:

$$\begin{aligned} m_{ij} &= (e^{v_i} + 1) \exp \left(\sum_{\substack{k \neq i \\ k=1}}^n v_k \right) \left(\prod_{\substack{k \neq i \\ k=1}}^n 1/(e^{v_k} + 1) \right) \\ &\cdot \sum_{\underline{s}} s_j / \left[\exp \left(\sum_{k \in K^+} v_k \right) + \exp \left(\sum_{k \in K^-} v_k \right) \right], \end{aligned}$$

où la somme sur \underline{s} est étendue à toutes les 2^{n-1} configurations de signes telles que $s_i = +1$, où K^+ et K^- sont l'ensemble des indices des composantes de \underline{s} égales respectivement à +1 et -1.

Pour que le calcul déduit de (9) converge vers la valeur vraie v_i , pour chacun des canaux, il est nécessaire et suffisant que la plus grande des valeurs propres de la matrice \underline{M} soit de module inférieur à 1. Alors en effet $\lim_{t \rightarrow \infty} \underline{M}^t = \underline{0}$ et le vecteur des écarts $\underline{\delta'v}_i$ tend vers le vecteur nul $\underline{0}$.

La complexité de l'expression de l'élément général de la matrice \underline{M} rend difficile l'étude de ses valeurs propres. Nous n'examinerons que deux cas pour lesquels des résultats généraux peuvent être facilement obtenus: celui où $n = 2$ et celui où tous les coefficients v_i sont égaux, quel que soit n .

Dans le premier cas, l'équation aux valeurs propres de la matrice \underline{M} s'écrit:

$$(\lambda - 1) [\lambda - 4z_1 z_2 / (z_1 z_2 + 1)(z_1 + z_2)] = 0,$$

en ayant posé $\exp(v_1) = z_1$ et $\exp(v_2) = z_2$. La plus grande valeur propre de \underline{M} vaut donc 1, car

$$4z_1 z_2 / (z_1 z_2 + 1)(z_1 + z_2) < 1.$$

Il n'y a donc pas convergence dans ce cas.

Dans le second cas, on pose $\exp(v_1) = \exp(v_2) = \dots = \exp(v_n) = z$. On trouve alors pour $n = 3$ que la plus grande valeur propre est racine double de l'équation aux valeurs propres et vaut

$$4z/(z+1)^2$$



plus petite que 1 pour $z > 1$. Pour $n > 3$, la plus grande valeur propre est racine d'ordre $n-1$ et décroît avec n . La convergence a donc toujours lieu dans ce cas, pourvu que $n \geq 3$.

Les simulations paraissent montrer que:

- pour $n > 2$, la convergence a toujours lieu;
- elle a lieu également si les conditions initiales sont très différentes des valeurs vraies, c'est-à-dire lorsque l'hypothèse qui justifie l'approximation linéaire cesse d'être admissible.

Il nous est cependant impossible de prouver ces assertions en toute généralité.

2 - Système utilisant un codage (n,k)

Le schéma est alors celui de la figure 2.

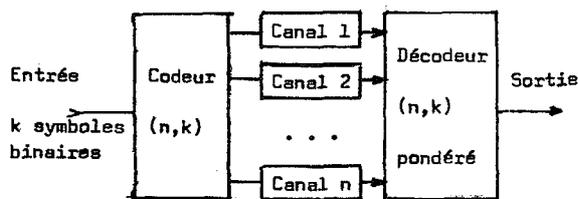


Figure 2 Système avec codage (n,k)

Ce cas n'est pas fondamentalement différent du précédent, mais la règle de décision (3) doit être remplacée par

$$(12) \quad A_i = a_i + F_i(a_1, \dots, a_n),$$

où A_i , valeur relative a posteriori, dépend maintenant du canal puisque les symboles émis diffèrent d'un canal à un autre. La fonction $F_i(\dots)$ est déduite d'une matrice de contrôle du code utilisé [1] et ne dépend pas de a_i .

Par exemple, si le code considéré est décodable à seuil, la fonction $F_i(\dots)$ s'exprime comme la somme de r grandeurs de la forme

$$\log_e \left(1 + \prod_{j \in I_m} \text{th}(a_j/2) \right) - \log_e \left(1 - \prod_{j \in I_m} \text{th}(a_j/2) \right),$$

où a_j est la valeur relative (1) du j -ème symbole et où les I_m , $m = 1, 2, \dots, r$, sont des ensembles d'indices disjoints et ne contenant pas i qui décrivent les répliques "orthogonales" du i -ème symbole.

Pour un code cyclique, la fonction F_{i+j} se déduit de la fonction F_i par un décalage circulaire par j des indices. On peut donc caractériser l'ensemble des fonctions F_i en se donnant l'une d'elles.

L'expression (12) de A doit être substituée à (3) dans (9) et (11) qui conservent leur validité. L'analyse de la convergence pourrait en principe être effectuée par les mêmes méthodes que dans le paragraphe 1 - 5, mais serait sensiblement compliquée par la substitution de (12) à (3).

Remerciements

Un problème posé par la Division Télécommunications de Thomson-CSF est à l'origine de cette étude. M. Mollat du Jourdin, alors élève de 3-ème année à l'E.N.S.T., y a effectué en 1977-1978 un stage consacré à la simulation du système décrit ci-dessus.

Référence

- [1] G. Battail et M. Decouvelaere, "Décodage par répliques", Annales des Télécom. 31 n° 11-12, Nov.-Déc. 1976, pp. 387-404.