Communication satellitaire et codes de Reed-Solomon

(Titre initial : corruption de données et codes correcteurs)

Auteur : Aurélien BLICQ

Positionnement thématique:

Informatique théorique – algèbre – informatique pratique

Mots Clés:

Français:	Anglais:
Canal binaire symétrique	Binary symmetric channel
Codes de Reed-Solomon	Reed-Solomon codes
Corps finis	Finite fields
Polynômes	Polynomials
Corruption de données	Data corruption

Bibliographie Commentée:

Explication du changement de titre : ayant d'abord pensé à un sujet sur les codes correcteurs au sens large, j'ai ensuite choisi de me concentrer sur la problématique de la communication satellitaire car ce problème, plus concret, m'a semblé mieux adapté à l'esprit de l'épreuve.

Les satellites trouvent aujourd'hui de nombreuses applications, aussi bien militaires, qu'industrielles, ou civiles. Parmi elles, on il y a la communication intercontinentale, le système GPS, ou encore la télévision numérique. Dans ces domaines de précision, la corruption de données est inacceptable. Néanmoins, lors de la transmission de données, celles-ci sont soumises à divers contraintes externes pouvant les endommager.

Parmi ces contraintes, on trouve le bruit dû à des vibrations thermiques (bruit de Johnson-Nyquist) ou des effets de rayonnement de corps comme la terre, le soleil ou d'autres objets célestes [1].

Afin de remédier à ce problème, on utilise des codes correcteurs : au lieu d'envoyer directement l'information, on lui ajoute de la redondance afin de pouvoir la retrouver en cas d'erreur lors de la transmission [2].

On représente un message par une suite de bits, ou mathématiquement, par une suite finie d'éléments de GF(2), le corps de Galois de cardinal deux [3]. On veut transmettre un message à travers un canal bruité, c'est-à-dire modifiant aléatoirement certains bits du message envoyé. L'étape de codage consiste à découper le message à transmettre en blocs de k bits puis de coder chacun de ces blocs en un mot de n bits (avec $n > k$), puis de transmettre le mot codé. A la sortie du canal bruité, on décode chacun des blocs et on reconstitue le message envoyé. Un code fonctionnant de la sorte est appelé code par bloc de paramètres (n, k)[2].

Plus généralement, un code par bloc peut gérer des données qui ne sont pas nécessairement des suites de bits mais des suites d'éléments de n'importe quel GF(q), le corps de Galois à q éléments, on ne parle alors plus de bits mais de symboles. De tels corps sont construits en considérant un polynôme irréductible à coefficients dans GF(p) = $\mathbb{Z}/p\mathbb{Z}$ $P_0(X)$ (où $p \in \mathbb{N}$ est premier) puis le corps GF(2)[X]/P_0(X) des classes d'équivalence pour la congruence modulo P_0 [2]. On obtient alors le corps GF(p^d) avec d = deg(P_0) [2].

Pour le canal bruité, une approche possible est d'utiliser le bruit blanc gaussien. En effet, celui-ci modélise bien les phénomènes physiques sous-jacents [1]. Mais il faudrait pour cela s'intéresser à la transmission des données sous forme d'ondes, aspect du problème qui mériterait un TIPE entier et auquel j'ai choisi de ne pas m'intéresser. J'utiliserais donc le modèle du canal binaire symétrique sans mémoire. Dans ce modèle, chaque bit empruntant le canal a la même probabilité p d'être inversé [2], qu'il vaille 0 ou 1 et indépendamment des autres bits du mot.

Un code de Reed-Solomon RS(n, k, t) est un code par bloc prenant une suite de k symboles et lui ajoute 2t symboles de contrôle pour donner un bloc de n symboles[3]. Généralement, les symboles sont des éléments de GF(2^8), c'est-à-dire des octets. L'intérêt des codes de Reed-Solomon est qu'ils sont performants pour corriger les bouffées d'erreurs, c'est-à-dire les erreurs affectant des symboles successifs, tout en étant également efficace pour corriger les erreurs aléatoires. Un code RS(n, k, t) peut en effet corriger 2t erreurs consécutives et t erreurs aléatoires [3]. De plus, ces codes peuvent s'implémenter facilement à l'aide de registres à décalage [3].

Pour ce code, on interprétera les messages non comme des suites finies, mais plutôt comme des polynômes à coefficients dans GF(2^8).

L'utilisation d'un code correcteur n'est cependant pas sans conséquences sur un système communiquant. Celui-ci augmente le temps de traitement en ajoutant des délais de codage et décodage. De plus, le fait d'ajouter de la redondance augmente le nombre de symboles à transmettre sans changer la quantité d'information. Il faut donc, soit accepter un débit d'information plus lent, soit augmenter le débit de symboles du matériel, quitte à augmenter la consommation d'énergie. L'utilisation d'un code correcteur présente donc des inconvénients et doit être réfléchie selon la situation [5].

Problématique :

Le consortium européen DVB (Digital Video Broadcasting) recommande l'utilisation d'un code de Reed-Solomon RS(204, 188, 8) pour la transmission de données télévisuelles [4]. En me basant sur cet exemple, je me propose d'étudier le code de Reed-Solomon, et de discuter de la pertinence de ce choix ainsi que des limitations du modèle que j'ai adopté.

Objectifs du TIPE :

- Implémentation du code de Reed-Solomon RS(204,188) à l'aide de python
- Etude des performances de ce code
- Discussion du choix du code ainsi que du modèle adopté

Références bibliographiques:

[1] Alexandru Spătaru, *Fondements de la théorie de la transmission de l'information*, presses polytechniques romandes, 1987
[2] Michel Demazure, *Cours d'algèbre,* Cassini, 2008
[3] Wiliam A. Geisel, *Tutorial on Reed-Solomon Error Correction Coding,* NASA technical memorandum 102162, 1990
[4] *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television*, ETSI EN 300 744 V1.6.2 (2015-10)
[5] Jean-Claude Belfiore, Philippe Ciblat, Michèle Wigger, *COM105 Communications Numériques et Théorie de l'Information*, cours de Telecom ParisTech, 2014