ACCQ204

Enseignant: Aslan Tchamkerten

1 Récapitulatif

- Hamming: $[n, n - \log_2(n+1), 3]_2$ d'où $R = 1 - O((\log n)/n)$ et $\delta = O(1/n)$. Taux élevé, distance faible, faible complexité.
- RS: faible complexité, atteignent la borne de Singleton $[n = 2^t, k, d = n - k + 1]_q$ mais $q \geq n$. Or on voudrait pouvoir choisir q indépendamment de n car c'est le message a envoyer qui dicte l'alphabet à utiliser, et non le code qui dicte l'alphabet utilié pour écrire le message...
- codes BCH ("sous-ensemble-sous-corps" d'un code RS): faible complexité mais dsi $R > 0$ on a $\delta = 0$

Comment atteindre $\delta, R > 0$ sur un alphabet de taille donnée et avec une complexité polynomiale en codage/décodage?

Idée! Supposons qu'on veuille un code binaire de taux 1/2. On pourrait partir d'un code RS $[n, k = n/2, n/2+1]_q$ sur un alphabet à $q = n = 2^t$ éléments et écrire chaque symbol sur t bits (voir TD3). On obtient donc un code $[n \log_2 n, (n \log_2 n)/2, n/2 + 1]_2$. D'où $R \to 1/2$ mais $\delta \to 0$ car la distance minimale reste inchangée (chaque symbole du code de RS est "codé" avec un code de distance 1. Et si on augmentait cette distance par un véritable code correcteur d'erreurs au niveau des symbols du code RS? Ceci aboutit aux codes dit "concaténés".

2 Codes concaténés

Soit $q \geq 2, k \geq 1$ entiers et $Q = q^k$.

Soit

$$C_{out} : [Q]^K \to [Q]^N$$

code dit extérieur et

$$C_{in} : [q]^k \to [q]^n$$

code dit intérieur.

Pour un message

$$m = (m_1, \ldots, m_K)$$

on a

$$C_{out}(m) = [C_{out}(m)_1, \ldots, C_{out}(m)_N]$$

et ensuite on utilise C_{in} et on obtient

$$[C_{in}(C_{out}(m)_1), \ldots, C_{in}(C_{out}(m)_N)] \equiv C_{in} \circ C_{out}(m)$$

le concaténation de codes C_{in} et C_{out}.

Theorem 1 *$C_{in} \circ C_{out}$ est un code $[n \cdot N, k \cdot K, d \cdot D]_q$. De plus, si C_{in} et C_{out} sont linéaires, alors $C_{in} \circ C_{out}$ est linéaire. (voir exercices).*

Corollary 1 *Si C_{out} et C_{in} ont les taux R et r et distances δ_{out} et δ_{in}, respectivement, alors $C_{in} \circ C_{out}$ a taux $R \cdot r$ et distance minimale $\delta_{out} \cdot \delta_{in}$.*

2.1 Décodage simple de code concaténés

On suppose que C_{in} a distance minimale d et C_{out} distance minimale D. On considère décoder C_{in} puis ensuite C_{out}.

- Si le $i$ème bloc contient $< d/2$ erreurs il sera décodé juste.
- Si il y a moins de $D/2$ blocs erronés, alors $C_{in} \circ C_{out}$ sera décodé juste.

Si le nombre total d'erreurs est $< d \cdot D/4$ alors le nombre de blocs avec au moins $d/2$ erreurs est au plus $D/2$.

Donc on peut corriger jusqu'a $d \cdot D/4$ et on sait qu'on peut corriger jusqu'a $d \cdot D/2$. Le décodage simple, bien qu'a faible complexité, n'est donc pas optimal. Peut on faire mieux avec un décodeur de faible complexité? "Oui", mais avec le modèle de Shannon (erreurs aléatoires). Rappel: dans le modèle de Hamming, pour une distance minimale d donnée on peut espérer au mieux corriger $d/2$ erreurs. Malheureusement, le meilleur compromis taux/distance minimale n'est pas connu (voir bornes fondamentales). De plus, pour les codes concaténés on ne sait pas non plus comment atteindre $d/2$ (i.e., $d \cdot D/2$) avec une faible complexité. A l'inverse, dans le modèle de Shannon la capacité correspond au meilleur compromis taux/correction d'erreur.

Dans la prochaine section on va voir comment atteindre la capacité d'un BSC(p) avec un décodeur de faible complexité, à l'aide du codage en petit blocs. Cette construction permet d'atteindre une probabilité d'erreur qui décroît polynomialement avec la longueur du bloc. Ensuite on va voir comment, à l'aide d'une construction en concatétation, on peut atteindre les mêmes peformances mais en plus avec une probabilité d'erreur qui décroit exponentiellement avec la taille du bloc.

3 Concaténation et erreurs aléatoires (modèle de Shannon)

3.1 Codage en petits blocs et modèle de Shannon: Partie A

Rappel: Shannon nous dit que il existe un code $(N, R = 1 - H_b(p) - \varepsilon)$ t.q. $Pr(erreur) \leq 2^{-\tilde{\varepsilon}N}$ (où $\tilde{\varepsilon} > 0$ pour tout $\varepsilon > 0$) avec complexité au décodage $e^{O(N)}$.

Idée: décomposer la transmission sur N bits en petites transmission de longueur $N' = c \log N$. Chaque petite transmission opère au même taux $R = 1 - H_b(p) - \varepsilon$ et donc le taux du code global est R.

En appliquant le théorème de Shannon sur chaque bloc on a que

$$Pr(\text{bloc i décodé incorrectement}) \leq N^{-\tilde{\varepsilon}c}$$

et par la borne de l'union

$$Pr(\text{il existe un bloc i décodé incorrectement}) \leq \frac{N}{c \log N} N^{-\tilde{\varepsilon}c}.$$

Puisque c peut être choisi arbitrairement, en prenant c suffisamment grand on peut guarantir que la borne ci-dessus tend vers zéro lorsque $N \to \infty$.

Complexité au décodage (recherche exhaustive pour tous les blocs): $O(N \exp(c' \log N)) = poly(N)$.

La construction en "petits blocs" ci-dessus permet d'atteindre la capacité du BSC avec une faible complexité au décodage et une probabilité dérreur qui decroît polynomialement.

Avec une construction en concaténation on peut faire mieux, rendre la proba. d'erreur exponentiellement décroissante (tout en gardant une compéxité au décodage sous exponentielle).

3.2 Code concaténés et modèle de Shannon: Partie B (Forney 1965, PhD thesis)

C_{out} : $(N, R = 1 - \varepsilon) - RS$ ($\Rightarrow \delta = \varepsilon$ et on peut corriger jusqu'à une fraction $\varepsilon/2$ d'erreurs.
C_{in} : $(n, r = 1 - H_b(p) - \varepsilon)$
$C_{in} \circ C_{out}$: $\mathcal{R} = 1 - H_b(p) - \varepsilon'$ avec $\varepsilon' \to 0$ si $\varepsilon \to 0$.

Vu:

$$Pr(\text{bloc i décodé incorrectement}) \leq e^{-\Omega(n)} \quad \text{(Shannon)}$$

Probabilité d'erreur:

$$\begin{aligned} Pr(C_{in} \circ C_{out} \text{ décodé incorrectement}) &= Pr(\text{au moins } \varepsilon N/2 \text{ blocs décodés incorrectement}) \\ &= \sum_{i=\varepsilon N/2}^{N} (e^{-\Omega(n)})^i \binom{N}{i} \\ &\leq (e^{-\Omega(n)})^{\varepsilon N/2} 2^N \\ &= e^{-\Omega(nN)} \\ &= e^{-\Omega(\mathcal{N})}. \end{aligned}$$

où $\mathcal{N} = nN$ est la longueur du code concaténé. De plus, la complexité au decodeur est $poly(N)\exp(n) = poly(\mathcal{N})$. On déduit donc:
Théorème (Forney) On peut atteindre la capacité d'un canal BSC avec complexité enc./déc en $poly(\mathcal{N})$ et probabilité d'erreur $e^{-\Omega(\mathcal{N})}$.

4 Décodage en liste

Definition:

Soit $0 \leqslant \rho \leqslant 1$ et $L \geqslant 1$.

Un code $C \subseteq \Sigma^n$ est (ρ, L)-liste décodable, si $\forall y \in \Sigma^n$

$$|\{c \in C : \Delta(c, y) \leqslant \rho n\}| \leqslant L$$

Remarque:

Le décodage est dit "efficace" si $L = e^{o(n)}$.

Ce décodage peut être utilisé de plusieurs façon. Par exemple, si $|\,\text{liste}\,| = 1$, déclarer l'unique élément et si $|\,\text{liste}\,| \geq 2$, déclarer "erreur." Un autre exemple est de déterminer le message envoyé dans la liste à l'aide d'information extérieure, si telle est disponible.

Théorème:

Soit $q \geqslant 2$ entier et $0 < \rho < 1 - \frac{1}{q}$

i. Pour tout entier $L \geqslant 1$ et $R \leqslant 1 - H_q(\rho) - \frac{1}{L}$ il existe un code (ρ, L)-décodable.

ii. Si un (ρ, L) code a taux $R \geqslant 1 - H_q(\rho) + \varepsilon$ alors $L \geqslant 2^{\Omega(\varepsilon n)}$.

Remarques: avec $L = cste$ il est possible de corriger jusqu'a ρn erreurs avec un taux $1 - H_q(\rho) - 1/L$ (L n'a pas besoin de grandir comme $e^{o(n)}$!!!). Si le taux est $> 1 - H_q(\rho)$ alors corriger une fraction $\rho > 0$ d'erreurs implique une liste immense $L = e^{\Omega n}$.

Preuve:

i. $|C| = q^k$ et $\forall m$ et on génère chaque mot code indépendamment aléatoirement $C(m) \sim$ uniform$[q]^n$. On défini l'évènement erreur

$$\mathcal{E} = \{\exists m_{\alpha_1}, \dots m_{\alpha_{L+1}} \text{ et } y \in [q]^n \text{ tel que } C(m_{\alpha_i}) \in \mathcal{B}(y, \rho n) \forall i = 1 \dots L+1\}$$

On va montrer que si $R \leq 1 - H_q(\rho) - 1/L$ alors $P(\mathcal{E}) < 1$, et il s'ensuit que $\exists$ un code C t.q. $\forall y$, $\mathcal{B}(y, \rho n)$ contient au plus L mots codes. On a

$$\mathcal{E} = \bigcup_{\alpha_1 \dots \alpha_{L+1}, y} \mathcal{E}(\alpha_1, \dots, \alpha_{L+1}, y)$$

où on définit

$$\mathcal{E}(\alpha_1, \dots, \alpha_{L+1}, y) = \{C(m_{\alpha_i}) \in \mathcal{B}(y, \rho n), \forall i = 1 \dots L+1\}.$$

On a

$$P(C(m_{\alpha_i}) \in \mathcal{B}(y, \rho n)) = \frac{\text{Vol}_q(n, \rho n)}{q^n} \leqslant \frac{q^{nH_q(\rho)}}{q^n} = q^{-n(1-H_q(\rho))}$$

et donc

$$P(\mathcal{E}(\alpha_1, \dots, \alpha_{L+1}, y)) \leqslant q^{-n(L+1)(1-H_q(\rho))}.$$

Par la borne de l'union on déduit donc

$$P(\mathcal{E}) \leqslant \binom{q^k}{L+1} q^n q^{-n(L+1)(1-H_q(\rho))}.$$

Or sait que $\binom{a}{b} \leqslant a^b$ et $k = R \cdot n$, d'où

$$P(\mathcal{E}) \leqslant q^{-n(L+1)[1-H_q(p)-\frac{1}{L+1}-R]} = q^{-n(L+1)[(1-H_q(\rho)-\frac{1}{L}-R)+(\frac{1}{L}-\frac{1}{L+1})]}$$

En supposant

$$1 - H_q(\rho) - \frac{1}{L} - R \geqslant 0$$

et puisque $\frac{1}{L} - \frac{1}{L+1} = \frac{1}{L(L+1)}$, on conclut

$$P(\mathcal{E}) \leqslant q^{-\frac{n}{L}} < 1$$

concluant la preuve.

ii. On va montrer que si C a taux $R \geqslant 1 - H_q(\rho) + \varepsilon \Rightarrow \exists y \in [q]^n$ t.q. $|C \cap \mathcal{B}(y, \rho n)| = q^{\Omega(n)}$.

On fixe code C avec taux $R \geqslant 1 - H_q(\rho) + \varepsilon$, et on fixe $c \in C$.

Choisissons Y aléatoirement sur uniforme$[q]^n$.

$$P(c \in \mathcal{B}(Y, \rho n)) = P(Y \in \mathcal{B}(\mathbf{c}, \rho \mathbf{n})) = \frac{\mathrm{vol}_q(n, \rho n)}{q^n} \geqslant \frac{q^{n(H_q(\rho) - o(1))}}{q^n}$$

$$\mathbb{E}[|C \cap \mathcal{B}(Y, \rho n)|] = \sum_{c \in C} \mathbb{E}[1\{c \in \mathcal{B}(Y, \rho n)\}]$$

$$\mathbb{E}[1\{c \in \mathcal{B}(Y, \rho n)\}] = P(c \in \mathcal{B}(Y, \rho n)) \geqslant q^{-n(1 - H_q(\rho) + o(1))}$$

Puisque $|C| = q^{Rn}$, il suit que

$$\mathbb{E}[|C \cap \mathcal{B}(Y, \rho n)|] \geqslant q^{n(R - 1 + H_q(\rho) - o(1))} = q^{\Omega(n)} \text{ si } R \geqslant 1 - H_q(\rho) + \varepsilon$$

$$\Rightarrow \forall C \exists y : |C \cap \mathcal{B}(y, \rho n)| \geqslant q^{\Omega(n)} \text{ si } R \geqslant 1 - H_q(\rho) + \varepsilon.$$