Approximate Support Recovery: Bounds

Лизюра Дмитрий[1]
Руководитель: Фролов Алексей Андреевич[2]

[1]Московский физико-технический институт

[2]Сколковский институт науки и технологий (Сколтех)

18 марта 2025 года

Постановка задачи

Неформально

Имеется сервер и N пользователей, но в каждый момент времени не более $K_\alpha << N$ пользователей активно. Активные пользователи посылают серверу сообщения, однако они “склеиваются“, а ещё к ним добавляется шум. Задача сервера — расшифровать исходные сообщения. Наша задача — минимизировать длину кодовых слов.

Постановка задачи

В общем случае (определение 1 из [1])

Пусть:

- M — количество кодовых слов,
- n — длина сообщения,
- $P : (\mathcal{X}^n)^{K_\alpha} \to \mathcal{Y}^n$ — канал передачи сообщений,
- ε — вероятность ошибки,
- $f : [M] \to \mathcal{X}^n$ — кодировщик,
- $g : \mathcal{Y}^n \to \binom{[M]}{K_\alpha}$ — декодировщик,
- $W_1, \dots, W_{K_\alpha} \sim U[M]$ — посылаемые пользователями сообщения,
- $E_j = \{W_j \notin g(P(f(W_1), \dots, f(W_{K_\alpha})))\} \cup \{W_j = W_i \text{ для } i \neq j\}$ — событие ошибки j–ого пользователя.

Пара (f, g) называется (N, M, n, ε)–*кодом с произвольным доступом*, если выполнено

$$\frac{1}{K_\alpha} \sum_{j=1}^{K_\alpha} \mathbb{P}[E_j] \leq \varepsilon.$$

Постановка задачи

В частном случае

Новые ограничения:

- $\mathcal{X} = \mathcal{Y} = \mathbb{R}$.
- $P(X_1, \ldots, X_{K_\alpha}) = X_1 + \cdots + X_{K_\alpha} + Z,\ Z \sim \mathcal{N}(0, I_n)$.

Постановка задачи

В частном случае

Новые ограничения:

- $\mathcal{X} = \mathcal{Y} = \mathbb{R}$.
- $P(X_1, \ldots, X_{K_\alpha}) = X_1 + \cdots + X_{K_\alpha} + Z$, $Z \sim \mathcal{N}(0, I_n)$.
- Существует константа P, такая что для всех j выполнено $\|f(j)\|_2^2 \leq nP$ п.н.

Постановка задачи

В частном случае

Новые ограничения:

- $\mathcal{X} = \mathcal{Y} = \mathbb{R}$.
- $P(X_1, \dots, X_{K_\alpha}) = X_1 + \dots + X_{K_\alpha} + Z$, $Z \sim \mathcal{N}(0, I_n)$.
- Существует константа P, такая что для всех j выполнено $\|f(j)\|_2^2 \leq nP$ п.н.

Удельной энергией кода называется величина $\mathcal{E} = \frac{nP}{2\log_2 M}$. Задача — получить оценку на $\inf\{\mathcal{E}\}$ по всем $(N, M, \varepsilon, \mathcal{E}, K_\alpha)$–кодам с произвольным доступом.

Первая оценка [1]

Theorem 1. *Fix $P' < P$. There exists an (M, n, ϵ) random-access code for K_a-user GMAC satisfying power-constraint P and*

$$\epsilon \le \sum_{t=1}^{K_a} \frac{t}{K_a} \min(p_t, q_t) + p_0\,, \tag{3}$$

where

$$p_0 = \frac{\binom{K_a}{2}}{M} + K_a \mathbb{P}\left[\frac{1}{n}\sum_{j=1}^{n} Z_j^2 > \frac{P}{P'}\right], \tag{4}$$

$$p_t = e^{-nE(t)}, \tag{5}$$

$$E(t) = \max_{0 \le \rho, \rho_1 \le 1} -\rho\rho_1 t R_1 - \rho_1 R_2 + E_0(\rho, \rho_1)$$

$$E_0 = \rho_1 a + \frac{1}{2}\log(1 - 2b\rho_1)$$

$$a = \frac{\rho}{2}\log(1 + 2P' t\lambda) + \frac{1}{2}\log(1 + 2P' t\mu) \tag{6}$$

$$b = \rho\lambda - \frac{\mu}{1 + 2P' t\mu}, \quad \mu = \frac{\rho\lambda}{1 + 2P' t\lambda} \tag{7}$$

$$\lambda = \frac{P' t - 1 + \sqrt{D}}{4(1 + \rho_1\rho)P' t}, \tag{8}$$

$$D = (P' t - 1)^2 + 4P' t \frac{1 + \rho\rho_1}{1 + \rho}$$

$$R_1 = \frac{1}{n}\log M - \frac{1}{nt}\log(t!) \tag{9}$$

$$R_2 = \frac{1}{n}\log\binom{K_a}{t} \tag{10}$$

$$q_t = \inf_{\gamma} \mathbb{P}[I_t \le \gamma] + \exp\{n(tR_1 + R_2) - \gamma\}$$

Идея доказательства:

- Зафиксировать $c_1, \ldots, c_M \sim \mathcal{N}(0, P')$.
- $f(j) = c_j$.
- $g(y) = \arg\min_{S \subset [M]} \left|\sum_{j \in S} c_j - y\right|$.

Вторая оценка [2]

Зафиксируем $\mu = \frac{K_\alpha}{N}$ и определим

$$\mathcal{E}^*(M, \mu, \varepsilon) = \lim_{N \to \infty} \inf\{\mathcal{E} \mid \text{Существует } (N, M, \varepsilon, \mathcal{E}, \mu N) \text{ код}\}.$$

Тогда

Theorem 3 (Bound-MAC). *Fix M, μ and ϵ. Then*

$$\mathcal{E}^*(M, \mu, \epsilon) \leq \inf \frac{b^2}{2 \log_2 M},$$

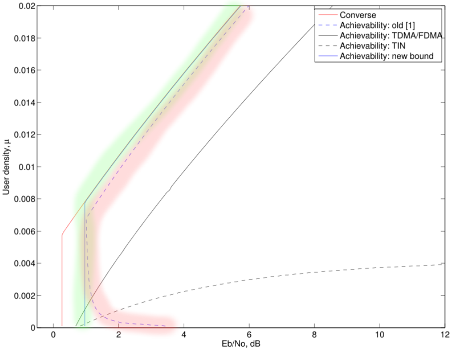
where infimum is over all $b > 0$ such that for all $\theta \in [\epsilon, 1]$ it holds that

$$\theta\mu \ln M + \mu h(\theta) < \max_{\lambda \geq 0} \{ \frac{1}{2} \ln(1 + 2b^2\theta\mu\lambda) + \lambda \frac{\psi(b, \theta, \mu)^2}{1 + 2b^2\theta\mu\lambda} - \lambda \}, \qquad (13)$$

where ψ is defined in (12).

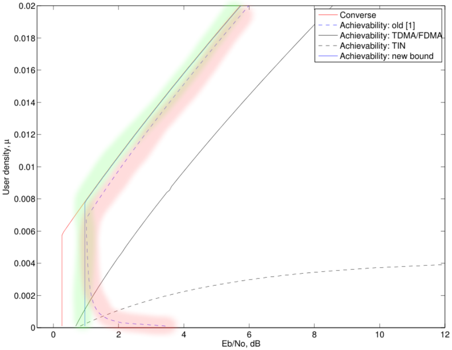
Сравнение оценок

Оценки на $\mathcal{E}^*$ для разных μ при $M = 2^{100}$ и $\varepsilon = 10^{-3}$ [2].



Сравнение оценок

Оценки на $\mathcal{E}^*$ для разных μ при $M = 2^{100}$ и $\varepsilon = 10^{-3}$ [2].



Вторая оценка гораздо лучше, но можно ли её модифицировать для $K_\alpha \not\to \infty$?

Ссылки

Y. Polyanskiy, "A perspective on massive random-access," *IEEE Information Theory (ISIT), 2017.*

I. Zadik, Y. Polyanskiy, and C. Thrampoulidis, "Improved bounds on Gaussian MAC and sparse regression via Gaussian inequalities," *IEEE International Symposium on Information Theory (ISIT), 2019.*