# Теория на кодирането и криптография в ИМИ

Цонка Байчева, Петър Бойваленков, Иван Ланджев

Когато се предава сигнал, основните модели го разглеждат като редица от нули и единици. Грешките при предаване, съхранение, четене и т.н. са неизбежни и затова теорията на кодирането се развива отначало като наука за защита на данните. Грешките се коригират и това е причината да имаме добре работещи телефонни връзки (стационарни и безжични), добри GPS-комуникации, добре съхранени данни, добър интернет, за да споменем само най-важните приложения. Кодовете, коригиращи грешки, са навсякъде и вършат своята работа, макар и зад сцената!

Секция Математически основи на информатиката (МОИ) на Института по математика и информатика на БАН е създадена през 1989 г. като формално обединение на учени, занимаващи се с теория на кодирането, някои от тях много отпреди това. От разстоянието на времето може уверено да се каже, че основателят на направлението кодиране в България и първи ръководител на секцията Стефан Додунеков създаде школа, която се утвърди и получи признание на световно ниво.

Ще разгледаме три от направленията, развивани в секция МОИ.

## 1 Универсални граници за енергии на сферични кодове

Темата за получаване и изследване на граници за кодове и дизайни в полиномиални метрични пространства е традиционна за секция МОИ от самото ѝ създаване. В началото на 1990-те по идея и препоръка на акад. Додунеков се създаде група под ръководството на втория автор, която и досега се занимава с изследвания в тази актуална област, пресечна точка на интересите на няколко математически направления.

По-долу ще представим накратко т.нар. универсални граници за кодове върху Евклидови сфери, област, в която резултати, получени от нашата група през 1995-2000 г. получиха признание и интересно развитие през последните пет години.

### 1.1 Линейно програмиране върху $\mathbb{S}^{n-1}$

Евклидовото разстояние между две точки $\mathbf{x} = (x_1, x_2, \dots, x_n)$ и $\mathbf{y} = (y_1, y_2, \dots, y_n)$ върху единичната сфера $\mathbb{S}^{n-1}$ в $n$-мерното Евклидово пространство $\mathbb{R}^n$ се дефинира с равенството

$$d(\mathbf{x}, \mathbf{y}) \coloneqq \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

Съответното скаларно произведение е

$$\langle \mathbf{x}, \mathbf{y} \rangle \coloneqq x_1 y_1 + x_2 y_2 + \cdots + x_n y_n,$$

като лесно се вижда, че $\langle \mathbf{x}, \mathbf{y} \rangle = 1 - d^2(\mathbf{x}, \mathbf{y})/2$.

Ще разгледаме сферични кодове $\mathcal{C} \subset \mathbb{S}^{n-1}$, които са непразни крайни множества от точки от $\mathbb{S}^{n-1}$. Най-важните параметри на един сферичен код са размерността $n$, мощността $|\mathcal{C}| = M$ и максималното скаларно произведение

$$s = s(C) = \max\{(\mathbf{x}, \mathbf{y}) \colon \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}\}.$$

Ще наричаме $\mathcal{C}$ сферичен $(n, M, s)$ код.

Изследването на параметрите на сферичните кодове е интересно за математическия анализ, геометрията, числения анализ, теорията на информацията, теорията на кодирането и др. Ролята на сферичните кодове в кодирането се разглежда в книгите [13, 18, 45, 6] и в много от статиите, цитирани в тези книги.

Една от най-важните задачи в теорията на кодирането е оценяването на величината

$$\mathcal{A}(n,s) := \max\{|\mathcal{C}|: \mathcal{C} \text{ е сферичен } (n,M,s) \text{ код}\}$$

при фиксирани $n$ и $s$. Действително, в много модели сигналите се разполагат като точки върху единична сфера и това предизвиква естествен интерес към сферичните кодовете с много точки (сигнали). Нещо повече, доброто оценяване на параметрите на декодирането изисква добри граници за $\mathcal{A}(n,s)$ и други параметри на сферичните кодове.

Съществува и друга популярна интерпретация на тясно свързаната задача за оценяване на

$$s(n,M) := \min\{s(C): C \subset \mathbb{S}^{n-1},\ |C| = M\},$$

минималното възможно максимално скаларно произведение за код върху $\mathbb{S}^{n-1}$ с фиксирана мощност $M$. Да си представим $n$-мерна планета с $M$ диктатора върху нея. Ние искаме да ги раздалечим максимално, защото иначе те започват да воюват и техните народи страдат. Сериозно, диктаторите са сигналите и планетата може да има и друга метрика.

Линейното програмиране е една от основните техники за оценяване на $\mathcal{A}(n,s)$ и $s(n,M)$. Тя е въведена в теорията на кодирането от Делсарт [14] в началото на 1970-те и развита от много математици (вж. например [29]).

Важна роля в линейното програмиране е отредена на теорията на ортогоналните полиноми.

За $a,b \in \{0,1\}$ означаваме с $\{P_i^{a,b}(t)\}_{i=0}^{\infty}$ полиномите на Якоби $\{P_i^{\alpha,\beta}(t)\}_{i=0}^{\infty}$ [1, Глава 22] с параметри

$$(\alpha,\beta) = \left(a + \frac{n-3}{2}, b + \frac{n-3}{2}\right),$$

нормализирани чрез $P_i^{\alpha,\beta}(1) = 1$. В случая $(a,b) = (0,0)$ получаваме полиномите на Гегенбауер и ще използваме $(n)$ индексиране вместо $0,0$. Нека $r_i := \frac{2i+n-2}{i+n-2}\binom{i+n-2}{i}$. Да означим с $t_i^{a,b}$ най-големия корен на полинома $P_i^{a,b}(t)$, като $t_0^{1,1} = -1$ по дефиниция.

Полиномите на Гегенбауер са ортогонални в интервала $[-1,1]$ с мярка

$$d\mu(t) := \gamma_n (1-t^2)^{\frac{n-3}{2}}\, dt, \quad t \in [-1,1],$$

където $\gamma_n := \Gamma\left(\frac{n}{2}\right)/\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)$ и удовлетворяват рекурентната връзка

$$(i+n-2)P_{i+1}^{(n)}(t) = (2i+n-2)tP_i^{(n)}(t) - iP_{i-1}^{(n)}(t), \quad i = 1,2,\dots,\ P_0^{(n)}(t) = 1,\ P_1^{(n)}(t) = t.$$

Полином $f(t) \in \mathbb{R}[t]$ се нарича положително дефинитен, ако всички коефициенти в развитието му $f(t) = \sum_{i=0}^{k} f_i\, P_i^{(n)}(t)$ по полиномите на Гегенбауер са неотрицателни, т.е. $f_i \geq 0$ за $i = 1,\dots,k$.

**Теорема 1.1** (Граница на линейното програмиране за сферични кодове [15, 24]). *Нека* $n \geq 2$, $s \in [-1,1)$ *и полиномът* $f(t) \in \mathbb{R}[t]$ *е положително дефинитен и е такъв, че* $f(t) \leq 0$ *за* $-1 \leq t \leq s$. *Тогава* $\mathcal{A}(n,s) \leq f(1)/f_0$.

С помощта на подходящи полиноми, за които ще стане дума по-долу, Левенщейн [26] (вж. също [27, 28, 29]) получава следната универсална[1] граница.

**Теорема 1.2** (Граница на Левенщейн). *В сила е*

$$\mathcal{A}(n,s) \leq L_\tau(n,s) := \left(1 - \frac{P_{k-1+\varepsilon}^{1,0}(s)}{P_k^{0,\varepsilon}(s)}\right) \sum_{i=0}^{k-1+\varepsilon} r_i\,, \quad \forall s \in I_\tau . \#(1)$$

*където* $I_\tau$ *е интервалът*

$$I_\tau := \left[t_{k-1+\varepsilon}^{1,1-\varepsilon}, t_k^{1,\varepsilon}\right], \ \tau = 2k-1+\varepsilon, \ \varepsilon \in \{0,1\}.$$

**Пример 1.3** Първите три граници на Левенщейн са $\mathcal{A}(n,s) \leq (s-1)/s$ за $s \in [-1, -1/n]$,

$$A(n,s) \leq \frac{2n(1-s)}{1-ns}$$

за $s \in [-1/n, 0]$ и

$$A(n,s) \leq \frac{n(1-s)(2+(n+1)s)}{1-ns^2}$$

за $s \in \left[0, \frac{\sqrt{n+3}-1}{n+2}\right]$.

Изследването на границите на линейното програмиране е свързано със специален клас от сферични кодове, наречени сферични дизайни.

**Дефиниция 1.4** (Делсарт–Гьоталс–Зайдел [15]). Сферичен код $C \subset \mathbb{S}^{n-1}$ се нарича сферичен $\tau$-дизайн, ако равенството

$$\frac{1}{\mu(\mathbb{S}^{n-1})} \int_{\mathbb{S}^{n-1}} f(x) d\mu(x) = \frac{1}{|C|} \sum_{x \in C} f(x)$$

($\mu(x)$ е обичайната Лебегова мярка) е изпълнено за всички полиноми $f(x) = f(x_1, x_2, \ldots, x_n)$ от степен най-много $\tau$.

Както самата дефиниция подсказва, сферичните дизайни имат многобройни практически приложения. Дизайни върху $\mathbb{S}^2$ с икосаедрална симетрия (геодезични мрежи), въведени през 1968 г. [34] стават важни напоследък в различни области, като например глобално моделиране на океаните [32] и наблюдение на окръжаващата среда [42]. Други дизайни, получени от класически многостени, са вложени в Cosmic Background Explorer project, Google S2 Geometry library, използват се в компютърната графика, 3D сканиране и печатане, дизайн на LEO и GSO сателитни конфигурации.

---

[1] Терминът „универсална“ идва от приложимостта ѝ за всеки код. По-нататък ще видим и други аспекти на универсалността на тази и други граници.

**Теорема 1.5** (Граница на линейното програмиране за сферични дизайни [15]). *Нека* $n \geq 2$, $\tau \geq 1$ *и полиномът* $f(t) \in \mathbb{R}[t]$ *е такъв, че* $f(t) \geq 0$ *за* $-1 \leq t \leq 1$ *и коефициентите в развитието* $f(t) = \sum_{i=0}^{k} f_i P_i^{(n)}(t)$ *удовлетворяват* $f_0 > 0$, $f_i \leq 0$ *за* $i = \tau + 1, \ldots, k$. *Тогава за всеки сферичен* $\tau$*-дизайн* $\mathcal{C} \subseteq \mathbb{S}^{n-1}$ *е изпълнено* $|\mathcal{C}| \geq f(1)/f_0$.

Следващата, също класическа, граница е тясно свързана с границата на Левенщейн.

**Теорема 1.6** (Граница на Делсарт-Гьоталс-Зайдел). *Мощността на сферичен* $\tau$*-дизайн* $\mathcal{C} \subseteq \mathbb{S}^{n-1}$ *удовлетворява*

$$|\mathcal{C}| \geq D(n,\tau) \coloneqq \binom{n+k-2+\varepsilon}{n-1} + \binom{n+k-2}{n-1}, \#(2)$$

*където* $\tau = 2k - 1 + \varepsilon$, $\varepsilon \in \{0,1\}$.

**Теорема 1.7** *Границите* (1) *и* (2) *са свързани с равенствата*

$$L_{\tau-1-\varepsilon}\left(n, t_{k-1-\varepsilon}^{1,1-\varepsilon}\right) = L_{\tau-\varepsilon}\left(n, t_{k-1-\varepsilon}^{1,1-\varepsilon}\right) = D(n, \tau-\varepsilon),\ \varepsilon \in \{0,1\} \#(3)$$

*в краищата на интервалите* $I_\tau$ *(*$\tau = 2k - 1 + \varepsilon$, $\varepsilon \in \{0,1\}$*). В частност, ако* $\mathcal{C} \subseteq \mathbb{S}^{n-1}$ *достига* (2), *той достига и* (1) *в ляв край на интервал* $I_\tau$.

Друга важна задача, която възниква по естествен начин, разглежда потенциалните енергии на сферичните кодове и дизайни.

**Дефиниция 1.8.** За функция $h(t): [-1,1] \to [0,+\infty]$ потенциалната енергия (или $h$-енергията) на сферичния код $\mathcal{C}$ се дефинира с равенството

$$\mathcal{E}_h(\mathcal{C}) \coloneqq \sum_{x,y \in \mathcal{C}, x \neq y} h\left(\langle x, y \rangle\right). \#(4)$$

Да разгледаме задачата за оценяване на величината

$$\mathcal{E}_h(n, M) \coloneqq \inf_{|C|=M} \{E_h(C)\}, \#(5)$$

минималната $h$-енергия на код $\mathcal{C} \subset \mathbb{S}^{n-1}$ с фиксирана мощност $M$. Линейното програмиране работи добре за енергиите. Нещо повече, както беше установено наскоро с участието на втория автор, задачите за оценяване на мощности и енергии са тясно свързани.

**Теорема 1.9** (Граница на линейното програмиране за енергия [43]). *Нека* $n \geq 2$ *и* $M \geq 2$ *са естествени числа и нека* $h: [-1,1] \to [0,+\infty]$. *Нека полиномът* $f(t) \in \mathbb{R}[t]$ *е положително дефинитен и е такъв, че* $f(t) \leq h(t)$ *за* $-1 \leq t < 1$. *Тогава* $\mathcal{E}_h(n, M) \geq M(f_0 M - f(1))$.

От особен интерес са потенциалните енергии за абсолютно монотонни функции (потенциали).

**Дефиниция 1.10.** Функцията $h(t): [-1,1] \to (0,+\infty]$ се нарича абсолютно монотонна в интервала $[-1,1]$, ако $h^{(k)}(t) \geq 0$ за всяко $t \in [-1,1)$ и всяко цяло $k \geq 0$, като $h(1) = \lim_{t \to 1^-} h(t)$.

Примери за абсолютно монотонни потенциали са:

$$h(t) = (2-2t)^{-(n-2)/2} - \text{ потенциал на Нютон,}$$

$$h(t) = (2-2t)^{-s/2} - \text{ потенциал на Рис,}$$

$$h(t) = \exp(2t-2) - \text{ потенциал на Гаус,}$$

$$h(t) = -\log(2-2t) - \text{ логаритмичен потенциал.}$$

Както имената подсказват, задачата е класическа.

**Дефиниция 1.11.** За $(a,b) = (0,0), (1,0)$ и $(1,1)$ полагаме

$$T_k^{a,b}(u,v) = \sum_{i=0}^{k} r_i^{a,b}\, P_i^{a,b}(u) P_i^{a,b}(v),$$

където $r_i^{a,b} = 1/c^{a,b} \int_{-1}^{1} \left(P_i^{a,b}(t)\right)^2 (1-t)^a (1+t)^b d\mu(t)$, $c^{1,0} = c^{0,0} = \gamma_n$ и $c^{1,1} = \gamma_{n+2}$. Нека $\alpha_i$, $i = 0, 1, \dots, k+\varepsilon$, са корените на полинома, използван от Левенщейн за получаването на границата (1), $s = \alpha_{k+\varepsilon}$, $\tau = 2k - 1 + \varepsilon$, $\varepsilon \in \{0,1\}$, и нека

$$\rho_1 = \frac{T_k^{0,0}(s,1)}{T_k^{0,0}(-1,-1) T_k^{0,0}(s,1) - T_k^{0,0}(-1,1) T_k^{0,0}(s,-1)} \quad \text{за } \varepsilon = 1,$$

$$\rho_{i+\varepsilon} = \frac{1}{c^{1,\varepsilon}(1+\alpha_{i+\varepsilon})^{\varepsilon}(1-\alpha_{i+\varepsilon}) T_{k-1}^{1,\varepsilon}(\alpha_{i+\varepsilon}, \alpha_{i+\varepsilon})}, \quad i = 1, 2, \dots, k,$$

През 2016 г. вторият автор заедно с Драгнев, Хардин, Саф и Стоянова получиха следната универсална граница.

**Теорема 1.12** ([10]). *Нека $n \geq 2$ и $\tau = 2k - 1 + \varepsilon \geq 1$, $\varepsilon \in \{0,1\}$, са естествени числа. Нека функцията $h$ е абсолютно монотонна в $[-1,1]$. За всяко $M \in (D(n,\tau), D(n,\tau+1)]$ е изпълнено*

$$\mathcal{E}_h(n,M) \geq M^2 \sum_{i=1}^{k+\varepsilon} \rho_i\, h(\alpha_i). \#(6)$$

*Ако сферичен $(n,M,s)$ код достига границата* (6)*, той е сферичен $\tau$-дизайн и неговите скаларни произведение са точно числата $\alpha_0, \alpha_1, \dots, \alpha_{k+\varepsilon}$.*

Условията за достигане на границите (1) и (6) съвпадат, т.е. сферичен код достига границата на Левенщейн (1) тогава и само тогава, когато достига и (6). В частност, всеки сферичен дизайн, който достига границата на Делсарт-Гьоталс-Зайдел (2), достига и нашата граница (6).

Полиномите на Левенщейн и полиномите, които използваме за получаване на (6), са оптимални в следния смисъл – границите (1) и (6) не могат да бъдат подобрени с полиноми от същата или по-ниска степен. От друга страна, подобрения с полиноми от по-висока степен понякога са възможни и изучаването на този въпрос по отношение на границата на Левенщейн беше започнато от втория автор заедно с Данев и Бумова през 1996 г. [9] (вж. също [8]).

**Теорема 1.13** ([39] за (1), [10] за (6)). *Границите* (1) *и* (6) *не могат да бъдат подобрени чрез използване съответно в Теорема 1.1 и 1.9 на полиноми от по-ниска или същата степен.*

**Теорема 1.14** ([9] за (1), [10] за (6)). *Нека*

$$Q_j^{(n)} := \frac{1}{N} + \sum_{i=1}^{k} \rho_i\, P_j^{(n)}(\alpha_i), \quad j \geq \tau = 2k - 1 + \varepsilon, \ \varepsilon \in \{0,1\}.$$

*Границите (1) и (6) могат да бъдат подобрени (едновременно) с линейно програмиране тогава и само тогава, когато $Q_j^{(n)} < 0$ за някое $j > \tau$.*

Да отбележим още, че границата (6) обобщава, в известен смисъл, работата на Кон и Кумар [12] от 2007 г. върху универсално оптимални сферични кодове.

**Дефиниция 1.15.** Сферичен код $\mathcal{C} \subseteq \mathbb{S}^{n-1}$ се нарича универсално оптимален, ако (слабо) минимизира $h$-енергията измежду всички кодове с $|\mathcal{C}|$ точки върху $S^{n-1}$ и за всеки абсолютно монотонен потенциал $h$.

**Теорема 1.16** ([12]). *Всеки сферичен код, който е сферичен* $(2k-1)$*-дизайн и има точно* $k$ *различни скаларни произведения между различни точки, е универсално оптимален. Кодът* $(4, 120, (1+\sqrt{5})/4)$ *(600-клетката) също е универсално оптимален.*

Всеки код, който достига (1) (следователно и (6)), е универсално оптимален. Не е известно дали има други сферични кодове, освен 600-клетката, които са универсално оптимални, но не достигат (1) и (6).

**Пример 1.17.** Нека $(n, M) = (4,24)$. Добре известният код $D_4$ ($D_4$-коренова система; 24-клетка) е оптимален в смисъл, че реализира дълго търсеното четвърто контактно число 24 [31]. Но този код не е универсално оптимален, въпреки, че енергията му е много близка до границата (6). Например с $h(t) = \frac{1}{2(1-t)}$ имаме енергия $334$, докато (6) дава $333$ (което може да бъде подобрено до $\approx 333.157$).

**Хипотеза 1.18.** *Всеки универсално оптимален код достига граница на линейното програмиране от Теорема 1.9.*

**Пример 1.19.** Стандартна конструкция (вж.[13, Глава 5]) изпраща двоични кодове с дължина $n$ върху сферата $\mathbb{S}^{n-1}$ – координатите 0 и 1 отиват в $\pm 1/\sqrt{n}$, съответно. Лесно се вижда, че скаларното произведение $\langle \overline{x}, \overline{y} \rangle$ на образите върху $\mathbb{S}^{n-1}$ и хеминговото разстояние $d_H(x, y)$ между първообразите са свързани с $\langle \overline{x}, \overline{y} \rangle = 1 - \frac{2d_H(x,y)}{n}$. Прилагайки това изображение върху кодовете на Кердок [25] получаваме сферичен код със скаларни произведения $1, \frac{1}{\sqrt{n}}, 0, -\frac{1}{\sqrt{n}}, -1$ и известно разределение на разстоянията. Този код не е оптимален, но енергията му

$$\mathcal{E}_h(\overline{K}_\ell) = N\left(\left(2^{2\ell+1}-2\right)h(0) + 2^{2\ell}\left(2^{2\ell-1}-1\right)\left(h\left(\frac{1}{2^\ell}\right) + h\left(-\frac{1}{2^\ell}\right)\right) + h(-1)\right).$$

е много близка до границата (6), като енергията и границата съвпадат асимптотично.

# 2 Изследвания на комбинаторни структури, осигуряващи цялостност и сигурност на информацията

Заедно с повсеместното навлизане на информационните технологии и интернет във всички сфери на живота, възниква и въпросът за гарантирането на целостта и сигурността на информацията събирана, разпространявана, съхранявана и използвана от различните приложения основани на тези технологии. Криптологията се занимава с решаването на тези проблеми чрез разработването на алгоритми и протоколи, които използват техники, методи и подходи от различни области на математиката.

Работата в областта на криптографията е сравнително нова за секция МОИ. Тя е свързана с възможността за създаване на ефективни криптографски приложения, базирани на различни комбинаторни структури, които част от членовете на секцията изследват от години, имат

разработен богат инструментариум и получени значими резултати. Те се основават на използването на известни математически свойства на изследваните комбинаторни обекти, допълнени с разработването и използването на ефективни компютърни алгоритми и софтуер. Получените досега резултати са свързани с генериране на субституционни кутии с добри криптографски свойства, разработване на инструмент за автоматичен апостериорен криптоанализ на публични ключове, генерирани с RSA, и на RSA съобщения, криптоанализ на кратки съобщения (по-малко от 75 символа), криптирани с машината за шифроване M-138, оптимизиране на биометричен протокол изискващ Oblivious Transfer. Тук ще представим разработен от нас метод [2] за генериране на оптимални пермутации за обобщени структури на Фейстел.

Блоковите шифри са основни компоненти при изграждането на много криптографски протоколи. Блоковият шифър е детерминистичен алгоритъм, който оперира с групи от битове с фиксирана дължина, наречени блокове, като ги трансформира в съответствие с предварително генериран ключ. Дизайнът на съвременните блокови шифри се базира на концепцията за итеративен шифър създаден и анализиран от Клод Шенон в [37]. Тези шифри осъществяват криптирането чрез последователност от рундове, като всеки рунд използва различен подключ, изведен от основния ключ. Няма известно математическо доказателство, че тези шифри са напълно сигурни, затова, от практическа гледна точка, е достатъчно да се покаже, че шифърът генерира случайна последователност.

Разбъркването и дифузията са свойства на сигурния шифър и са въведени от Шенон в [37]. Разбъркването изисква връзката между ключа и шифрирания текст да е колкото е възможно по-сложна. Дифузията се свързва със зависимостта на изходните битове от входните и от ключа. В шифър с добра дифузия, промяната на един входен бит би трябвало да доведе до промяна с вероятност 1/2 на всеки изходен бит.

Много от блоковите шифри се базират на мрежи на Файстел, като структурата и свойствата на тези шифри са широко използвани в криптографията. Мрежите на Файстел имат първото си комерсиално приложение в шифъра Lucifer на IBM, разработен от Хорст Файстел и Дон Коперсмит през 1973 г. Един $b$-битов шифър на Файстел изпълнява $r$ рунда, които имат идентична структура. Тя се състои от функция на рунда и размяна. Функцията на рунда $f$ съпоставя на $b/2$-битов вход, $b/2$-битов изход, формиран в зависимост от съответния ключ на рунда.

Обобщената структура на Файстел (generalized Feistel structure – GFS) е обобщение на класическата мрежа, като разделя открития текст на $k$ подблока за някое $k > 2$. В [44] е предложена GFS от тип 2, при която трансформацията на Файстел се прилага върху всеки два последователни подблока и след това се прави циклично завъртане на подблоковете. Обобщената структура на Файстел от тип 2 може лесно да се имплементира и примери на такива шифри са RC6 [33], HIGHT [22] и CLEFIA [38]. За съжаление, този тип структура на Файстел има ниска дифузия за големи $k$ и изисква голям брой рундове.

Suzaki и Minematsu [40] предлагат модификация на обобщената структура на Файстел от тип $\pi$ ($GFS_\pi$), която позволява подобряване на дифузията на шифъра.

В нашите изследвания, разглеждаме последователности от квадратни матрици, за да опишем дифузията на блоков шифър базиран на обобщена структура на Файстел от тип $\pi$ ($GFS_\pi$). Важно свойство на всеки шифър е броят на рундовете за криптиране (декриптиране), който е необходим, за да се постигне пълна дифузия (когато всеки подблок зависи от всички входни подблокове). Suzaki и Minematsu конструират пермутации, които водят до най-добрата дифузия за $GFS_\pi$, като прилагат една и съща пермутация на всеки рунд, докато ние търсим по-добра дифузия, като използваме различни пермутации на различните рундове.

За целта на изследването дефинираме редица от целочислени квадратни матрици, която наричаме редица от зависимости; $r$-тата матрица от тази редица представя зависимостта на изходните подблокове от рунд $r$ от подблоковете на открития текст, т.е. разглеждаме $k \times k$ матрицата $M = (m_{ij})$, където $m_{ij} = 1$ ако стойността на $i$-тия подблок зависи от стойността на $j$-тия подблок на открития текст и $m_{ij} = 0$ в противен случай.

Точно една последователност от зависимости отговаря на дадена последователност от пермутации за $GFS_\pi$, докато различни последователности от пермутации с една и съща дифузия могат да се получат от всяка последователност от зависимости. Ето защо, вместо да конструираме пермутации, ние конструираме редици от зависимости с добра дифузия с помощта на компютър. По-точно, ние рекурсивно конструираме редици от зависимости с оптимална дифузия за по-големи $k$, като използваме вече конструираните редици от зависимости за по-малки $k$. Две от рекурсивните конструкции водят до получаване на последователности от зависимости с точно определена дифузия. Останалите са конструкции, които могат да се приложат само при изпълнението на определени условия или са валидни само за някои определени дължини на подблоковете. Така получаваме оптимални последователности от зависимости, с най-добрите известни до момента параметри, за $18 \leq k \leq 2048$. Ограничението 2048 идва единствено от липсата в момента на практически интерес за по-големи стойности на $k$.

Чрез итеративното прилагане на предложените от нас конструкции, получаваме и безкрайни серии от последователности от зависимости за $k = 2^m$, $k = 3.2^m$, $k = 5.2^m$ и $k = 7.2^m$ подблока.

# 3 Линейни кодове над крайни полета и арки в геометрии на Галоа

Голям брой задачи от теория на линейните кодове са по своята същност геометрични. Така методи и резултати от крайните геометрии могат да бъдат използвани за изследване на проблеми от теория на кодирането и обратно. В този раздел ще представим някои интересни задачи, които са привлекли вниманието на изследователите в последните няколко десетилетия.

Един от най-известните примери за връзката между теория на кодирането и геометриите на Галоа е този за МДР-кодовете и арките в крайни проективни геометрии. Класическата монография на МакУилямс и Слоан [30] нарича МДР-кодовете най-завладяващия клас кодове в цялата теория на кодирането. Встрани от чисто математическата им красота тези кодове имат многобройни приложения като например в кодирането на музика в компакт дисковете или в стандартизирания шифър Rijndael [23].

Най-напред ще представим класа на МДР-кодовете (кодове с максимално достижимо разстояние).

**Теорема 3.1** (граница на Сингълтън). *За всеки линеен* $[n, k, d]_q$*-код* $C$ *е в сила неравенството* $d \leq n - k + 1$.

**Дефиниция 3.2.** Линеен код, за който се достига границата на Сингълтън, т.е. $d = n - k + 1$, се нарича код с максимално достижимо разстояние или МДР-код.

Следната теорема отразява фундаменталните свойства на МДР-кодовете и позволява да мислим за тях геометрично в термините на арки в подходящи геометрии на Галоа.

**Теорема 3.3.** *Нека* $C$ *е линеен* $[n, k, d]_q$*-код. Следните условия са еквивалентни:*

*(1)* $C$ *е линеен* $[n, k, n - k + 1]_q$*-код;*

*(2) всеки* $k$ *стълба в коя да е пораждаща матрица на* $C$ *са линейно независими;*

*(3) всеки $n-k$ стълба в коя да е проверочна матрица на $C$ са линейно независисми;*

*(4) $C^{\perp}$ е линеен $[n, n-k, k+1]_q$-код.*

Независимо от МДР-кодовете в геометриите на Галоа се въвежда понятието арка.

**Дефиниция 3.4.** Едно множество от $n$ точки в $\mathrm{PG}(k-1,n)$ наричаме $n$-арка, ако кои да е $k$ от тях са в общо положение (никои три на права, никои четири в равнина и т.н.). Една $n$-арка в $\mathrm{PG}(k-1,q)$ наричаме пълна тогава и само тогава, когато тя не се съдържа в $(n+1)$-арка в $\mathrm{PG}(k-1,q)$.

**Теорема 3.5.** *Множеството $K=\{g_1,\dots,g_n\}$ е $n$-арка в $PG(k-1,q)$ тогава и само тогава, когато $(k\times n)$-матрицата $G=(g_1 \dots g_n)$ е пораждаща матрица на код с параметри $[n,k,n-k+1]_q$.*

Съгласно Теореми 3.3 и 3.5 от съществуването на $n$-арка $K$ в $\mathrm{PG}(k-1,q)$ следва и съществуването на $n$-арка $\widetilde{K}$ в $\mathrm{PG}(n-k-1,q)$. Стандартен пример за $n$-арка в $\mathrm{PG}(k-1,q)$ са нормалните рационални криви.

**Дефиниция 3.6.** Нормална рационална крива в $\mathrm{PG}(k-1,q)$ наричаме всяко множество от точки в $\mathrm{PG}(k-1,q)$, което е проективно еквивалентно на множеството $\{(1,t,\dots,t^{k-1}) \mid t\in \mathbb{F}_q^+\}$, където $\mathbb{F}_q^+=\mathbb{F}_q\cup\{\infty\}$.

Кодовете, асоциирани с нормалните рационални криви, са известни като двойно разширени кодове на Рид-Соломон.

Да означим с $m(k-1,q)$ максималния брой точки в $n$-арка в $\mathrm{PG}(k-1,q)$. Задачата за определяне на точната стойност на $m(k-1,q)$ и задачата за характеризиране на арките, за които тази стойност се достига, е централна както за геометриите на Галоа, така и за теория на кодирането. В термините на теория на кодирането тя се формулира като задача за определяне на максималната дължина на МДР-код с фиксирана размерност $k$. Лесно се показва, че $m(k-1,q)\le q+k-1$.

Добре известно е [7], че

$$m(2,q)=\begin{cases} q+2 & \text{за } q \text{ четно,}\\ q+1 & \text{за } q \text{ нечетно.}\end{cases}$$

Прието е арка с мощност $m(2,q)$ да се нарича овал при нечетно $q$ и хиперовал при четно $q$. Следващата класическа теорема на Б. Сегре [35] характеризира овалите в $\mathrm{PG}(2,q)$, $q$ нечетно. В нея се твърди, че всяко множество от $q+1$ точки в проективна равнина от нечетен ред, никои три от които не са колинеарни, удовлетворява алгебрично уравнение от втора степен.

**Теорема 3.7.** *Всеки овал за $q$ нечетно се състои от рационалните точки на коника.*

За равнини от четен ред $q=2^h$ примери на хиперовали се получават като присъединим към точките на коника техния нуклеус (общата точка на всички допирателни). Хиперовал от този вид се нарича регулярен. Сегре [36] доказва, че за $q=2,4,8$ всеки хиперовал е регулярен. За $h\ge 4$ съществуват и нерегулярни хиперовали. Известни са няколко класа от нерегулярни хиперовали, но проблемът за класифицирането им е един от най-трудните в крайните геометрии. Но в сила е следната частична характеризация [36].

**Теорема 3.8.** *Всеки хиперовал в $PG(2,q)$, $q=2^h$, $h>1$, е проективно еквивалентен на*

$$\mathcal{D}(F)=\{(1,t,F(t)) \mid t\in \mathbb{F}_q\}\cup\{(0,1,0),(0,0,1)\},$$

*където $F$ е пермутационен полином над $\mathbb{F}_q$ от степен, ненадвишаваща $q-2$, за който $F(0)=0, F(1)=1$ и*

$$F_s(X)=\frac{F(X+s)+f(X)}{X}$$

*е пермутационен полином за всяко $s$, удовлетворяващ $F_s(0)=0$.*

Да се върнем към задачата за определяне на точната стойност на $m(k-1,q)$. Предположението, че $m(k-1,q)=q+1$ за всички $k$ и $q$, освен за $q$ четно, $k=3, q-1$, когато $m(k-1,q)=3$, е известно като МДР-хипотеза. Тя е формулирана от Б. Сегре в [35], но малко преди това Буш [11] доказва, че за $\mathrm{PG}(k-1,q)$ $m(k-1,q)=k+1$. Тази стойност се достига тогава и само тогава, когато множеството от точки е еквивалентно на $\{e_1,\dots,e_k,e_1+\cdots+e_k\}$, където $e_1,\dots,e_k$ е базис на $\mathbb{F}_q^k$.

През годините е доказано, че хипотезата е вярна за всички $q\le 27$ и всички $k\le 5$ и $k\ge -3$, както и за $k=6,7,q-4,q-5$ с някои изключения [21]. Максималната стойност $q+1$ (за $k\neq 3,q-1,q$ четно) се достига само за нормални рационални криви (двукратно разширени кодове на Рид-Соломон) с две изключения, конструирани от Глин [19] и Хършфелд [20].

Забележителен напредък по доказване на МДР-хипотезата бе направен в работата на Бол [3]. В нея е доказана следната теорема.

**Теорема 3.9.** *Нека $S$ е такова множество от вектори $S$ във векторното пространство $\mathbb{F}_q^k$, $q=p^h$, където $3\le q-p+1\le k\le q-2$, за което всяко негово подмножество с мощност $k$ е базис. Тогава $S$ е с мощност, ненадхвърляща $q+1$. В случай на равенство $S$ е проективно еквивалентно на нормална рационална крива.*

Важно следствие от тази теорема е, че МДР-хипотезата е вярна в случая на прости полета. На езика на теория на кодирането това означава, че най-дългите кодове от размерност, ненадхвърляща $p$, са двойно разширени кодове на Рид-Соломон, както и че най-дългите МДР-кодове са точно кодове на Рид-Соломон. Съществен напредък в доказването на МДР-хипотезата за непрости полета бе направен в [4], където тя е доказана за „малки“ размерности.

**Теорема 3.10.** *Нека $S$ е такова множество от вектори $S$ във векторното пространство $\mathbb{F}_q^k$, $q=p^h$, $h>1$, където $k\le q-2$, че всяко негово подмножество с мощност $k$ е базис. Тогава $|S|\le q+1$.*

Едно от очевидните следствия на МДР-хипотезата за теория на кодирането е, че дължината на МДР-кодовете е ограничена от мощността на полето. Един клас от кодове, запазващ всички съществени свойства на МДР-кодовете и в същото време съдържащ кодове с два пъти по-голяма дължина, е класът на т.нар. почти-МДР кодове, въведен от Додунеков и третия автор [17]. Този клас съдържа забележителни представители като троичния код на Голей, квадратично остатъчния $[11,6,5]$-код над $\mathbb{F}_4$, разширения квадратично-остатъчен код над същото поле, както и много добри алгебро-геометрични кодове.

Най-естествена е дефиницията на почти-МДР кодовете в термините на обобщените тегла на Хеминг.

**Дефиниция 3.11.** Нека $C$ е линеен $[n,k]_q$-код. Обобщеното тегло на Хеминг $d_r(C)$ (или $r$-то обобщено тегло на Хеминг) се дефинира като минималната мощност на носителя на $[n,r]$ подкод на $C$:

$$d_r(C) = \min\{|\mathrm{supp}D| \mid D \text{ е } [n,r] - \text{подкод на } C\}.$$

За обобщените тегла на Хеминг е в сила обобщената граница на Сингълтън:

$$d_r(C) \leq n - k + r,$$

за всяко $r = 1, \ldots, k$.

**Дефиниция 3.12.** Казваме, че линейният $[n,k]_q$-код е почти-МДР код, ако

$$d_i(C) = n - k + i, \text{за } i = 2, \ldots, k,$$

и

$$d_1(C) = n - k.$$

Не всеки $[n, k, n-k]_q$-код е почти-МДР, но това е вярно за дължини $n > k + q$. Както и при МДР-кодовете, ортогоналният на почти-МДР код е отново почти-МДР код. Спектърът на един почти-МДР код може да бъде изчислен с точност до един параметър [17].

**Теорема 3.13.** *Нека $C$ е $[n,k]_q$ почти-МДР код и нека $(A_i)$ и $A_i{}'$ са съответно спектрите на $C$ и $C^{\perp}$. Тогава за всяко $s \in \{1, \ldots, k\}$ е изпълнено:*

$$A_{n-k+s} = \binom{n}{k-s} \sum_{j=0}^{s-1} (-1)^j \binom{n-k+s}{j} (q^{s-j} - 1) + (-1)^s \binom{k}{s} A_{n-k},$$

$$A'_{k+s} = \binom{n}{k+s} \sum_{j=0}^{s-1} (-1)^j \binom{k+s}{j} (q^{s-j} - 1) + (-1)^s \binom{n-k}{s} A'_k.$$

Както и при МДР-кодовете съществува горна граница за дължината на почти-МДР $[n,k]_q$-код: $n \leq 2q + k$. От класическия геометричен резултат за несъществуване на максимални арки в геометрии над полета с нечетна характеристика [5] тази граница може да бъде подобрена до

$$n \leq 2q + k - 2.$$

Добри почти-МДР кодове се получават от геометрични конструкции. Така например пресечните точки на десетте прави на дезарговата конфигурация в $\mathrm{PG}(2,7)$ се асоциират с $[15,3,12]_7$ почти-МДР код, който лежи на горната граница. Един $[n,k]_q$-код, $q = p^h$, наричаме елиптичен код, ако той е асоцииран с елиптична крива с $n$ рационални точки. Ако $N_q(1)$ означава максималния брой $\mathbb{F}_q$-рационални точки върху елиптична крива, дефинирана над $\mathbb{F}_q$, то от класическия резултат на Уотърхаус [41] имаме, че

$$N_q(1) = \begin{cases} q + \lfloor 2\sqrt{q} \rfloor & \text{ако } p \text{ дели } \lfloor 2\sqrt{q} \rfloor \text{ и } h \geq 3 \text{ е нечетно,} \\ q + \lfloor 2\sqrt{q} \rfloor + 1 & \text{в противен случай.} \end{cases}$$

Следователно почти-МДР $[n,k]_q$-кодове съществуват за всички дължини $n \leq N_q(1)$ и всички размерности $k = 2, \ldots, n-2$. Съществува хипотеза за почти-МДР кодовете, аналогична на МДР-хипотезата, съгласно която най-добрите почти-МДР кодове не се различават много от елиптичните кодове:

За всяка степен на просто $q = p^h$ съществува константа $c$ (независеща от $q$) такава, че максималната дължина на почти-МДР $[n,k]_q$-код удовлетворява $n \leq N_1(q) + c$.

# Литература


[1] M.Abramowitz, I.A.Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965.

[2] T. Baicheva, S. Topalova. Optimal permutation sequences for the improved generalized Feistel, in preparation.

[3] S. Ball, On sets of vectors of a finite vector space in which every subset of a basis is a basis, *J. Europ. Math. Soc.* 14, 2012, 733–748.

[4] S. Ball, J. De Beule, On sets of vectors of a finite vector space in which every subset of a basis is a basis II, *Des. Codes Cryptogr*. 65, 2012, 323–329.

[5] S. Ball, A. Blokhuis, F. Mazzocca, Maximal arcs in Desarguesian planes of odd order do not exist, *Combinatorica* 17, 1997, 31–41.

[6] S. Borodachov, D. Hardin, E. Saff, *Minimal Discrete Energy on Recti_able Sets*, Springer, 2018 (to appear).

[7] R. C. Bose, Mathematical theory of the symmetric factorial design, *Sankhya* 8, 1947, 107–166.

[8] P. Boyvalenkov, Extremal polynomials for obtaining bounds for spherical codes and designs, *Discr. Comp. Geom.* 14, 167–183 (1995).

[9] P. Boyvalenkov, D. Danev, S. Bumova, Upper bounds on the minimum distance of spherical codes, *IEEE Trans. Inform. Theory* 41, 1576–1581 (1996).

[10] P. Boyvalenkov, P. Dragnev, D. Hardin, E. Saff, M. Stoyanova. Universal lower bounds for potential energy of spherical codes, *Constr. Approx.* 44, 2016, 385–415.

[11] K. A. Bush, Orthogonal arrays of index unity, *Ann. Math. Statist.* 23, 1952, 426–434.

[12] H. Cohn, A. Kumar, Universally optimal distribution of points on spheres, *J. Amer. Math. Soc.* 20, 99–148 (2007).

[13] J. H. Conway, N. J. A. Sloane, *Sphere packings, lattices and groups*, Springer-Verlag, New York, 1988.

[14] P. Delsarte, An Algebraic Approach to the Association Schemes in Coding Theory, *Philips Res. Rep. Suppl.* 10, (1973).

[15] P. Delsarte, J.-M. Goethals, J. J. Seidel, Spherical codes and designs, *Geom. Dedicata* 6, 1977, 363–388.

[16] P. Delsarte, V. I. Levenshtein, Association schemes and coding theory, *Trans. Inform. Theory* 44, 2477-2504 (1998).

[17] S. Dodunekov, I. Landgev, On near-MDS codes, *J. Geometry* 50, 1995, 30–43.

[18] T. Ericson, V. Zinoviev, Codes on Euclidean spheres, Elsevier Science B. V., 2001.

[19] D. G. Glynn, The non-classical 10-arc of PG(4,9), *Discrete math.* 59, 1986, 43–51.

[20] J. W. P. Hirschfeld, Rational curves on quadrics over finite fields of characteristic two, *Rend. Mat.* 3, 1971, 772–795.

[21] J. W. P. Hirschfeld, L. Storme, The packing problem in statistics, coding theory and finite projective spaces: update 2001, in: "Developments in Mathematics" 3, Kluwer, Proc. of the Fourth Isle of Thorns Conference, 2001, 201–246.

[22] Hong D., J. Sung, S. Hong, J. Lim, S. Lee, B. Koo, C. Lee, D. Chang, J. Lee, K. Jeong, H. Kim, J. Kim, and S. Chee, HIGHT: A new block cipher suitable for low-resource device, LNCS - CHES, 4249, 46–59 (2006).

[23] W.-A. Jackson, K. M. Martin, M. B. Paterson, Applications of Galois Geometry to Cryptology, Chapter 9 in: Current Research Topics in Galois Geometry, NOVA Sci. Publ., New York, 2011, 213–241.

[24] G. A. Kabatiansky, V. I. Levenshtein, Bounds for packings on a sphere and in space, *Probl. Inform. Transm.* 14, 1-17 (1978).

[25] A. M. Kerdock, A class of low-rate nonlinear binary codes, *Inform. and Control* 20, 1972, 182–187.

[26] V. I. Levenshtein, On bounds for packings in n-dimensional Euclidean space, *Dokl. Akad. Nauk SSSR* 245, 1299-1303, 1979 (in Russian); English translation in Soviet Math. Dokl. 20, 417-421, 1979.

[27] V. I. Levenshtein, Bounds for packings in metric spaces and certain applications, *Probl. Kibernetiki* 40, 1983, 44-110 (in Russian).

[28] V.I.Levenshtein, Designs as maximum codes in polynomial metric spaces, *Acta Appl. Math.* 25, 1992, 1-82.

[29] V. I. Levenshtein, Universal bounds for codes and designs, *Handbook of Coding Theory*, V. S. Pless and W. C. Hu_man, Eds., Elsevier, Amsterdam, Ch. 6, 499_648, (1998).

[30] F. J. MacWilliams, N. J. A. Sloane, The theory of error-correcting codes, North Holland, Nort Holland Math. Library vol. 16, Amsterdam, 1977.

[31] O. Musin, The kissing number in four dimensions. *Ann. of Math.* 168, 1–32 (2008).

[32] T. Ringler, M. Petersen, R. L. Higdonc, D. Jacobsen P. W. Jones, M. Maltrud, A multi-resolution approach to global ocean modeling, *Ocean Modelling* 69, 2013, 211–-232.

[33] Rivest R. L., M. J. B. Robshaw, R. Sidney, and Y. L. Yin, The RC6 block cipher, August 1998. http://people.csail.mit.edu/rivest/pubs/RRSY98.pdf (1998).

[34] P. Sadourny, A. Arakawa, Y. Mintz, Integration of the non-divergent barotropic vorticity equation with an icosahedral-hexagonal grid for the sphere, *Monthly Weather Review* 96 (6), 1968.

[35] B. Segre, Ovals in a finite projective plane, *Canad. J. Math.* 7, 414–416, 1955.

[36] B. Segre, Sui *k*-archi nei piani finiti di caratteristica due, *Rev. Math. Pures Appl.* 2, 289–300, 1957.

[37] Shanon, C. E., Communication theory of secrecy systems, Bell System Technical Journal 28, 656–715 (1949).

[38] Shirai T., K. Shibutani, T. Akishita, S. Moriai, and T. Iwata, The 128-bit block cipher CLEFIA (Extended abstract), LNCS - FSE, 4593, 181–195 (2007).

[39] V. M. Sidelnikov, On extremal polynomials used to estimate the size of codes, *Problems of Information Transmission* **16**, 174–186, (1980).

[40] Suzaki T. and K. Minematsu, Improving the generalized Feistel, LNCS - FSE, 6147, 19–39 (2010).

[41] W. C. Waterhouse, Abelian varieties over finite fields, *Ann. Sci. École. Norm. Sup.* 2(4), 1969, 521–560.

[42] D. White, A. J. Kimerling, W. S. Overton, Cartographic and geometric components of a global sampling design for environmental monitoring, *Cartography and Geographic Information Systems*, 19 (1), 5–22, (1992)

[43] V. A. Yudin, Minimal potential energy of a point system of charges, *Discret. Mat.* 4, 115-121 (1992) (in Russian); English translation: Discr. Math. Appl. 3, 75–81 (1993).

[44] Zheng Y., T. Matsumoto, and H. Imai, On the construction of block ciphers provably secure and not relaying on any unproved hypothesis, Advances in Cryptology - CRYPTO'89, LNCS 435, 461–480 (1989).

[45] C. Zong, Sphere packings, Springer-Verlag, New York, 1999.

**Coding theory and Cryptography in IMI**

Tsonka Baicheva, Peter Boyvalenkov, Ivan Landjev

*Abstract.* We describe old and recent results in three areas of investigations in Department of Mathematical Foundations of Informatics (MFI) of the Institute of Mathematics and Informatics. Recent universal lower bounds on potential energy of spherical codes are presented as natural generalizations and continuations of results of Levenshtein and results of MFI from 1990's. In the second part we explain new results on a cryptography problem about diffusion of certain block ciphers based on generalized Feistel networks. In the third part we describe the relation between coding theory and finite geometry.

проф. дмн Цонка Байчева
проф. дмн Петър Бойваленков
проф. дмн Иван Ланджев
1113 София, ул. Акад. Г. Бончев, бл. 8
Институт по математика и информатика при БАН