Авторска справка

на Емил Миланов Колев

Настоящата справка отразява научните приноси в публикациите, представени за участие в конкурс за професор в:

област на висше образование: 4. Природни науки, математика и информатика, професионално направление 4.5. Математика, научна специалност Алгебра и теория на числата (Теория на кодирането) обявен в Държавен вестник, брой 48 от 24.06.2016 г. за нуждите на Института по математика и информатика.

1. Област на изследванията

Теорията на кодирането възниква с фундаменталната работа на Клод Шенон „A Mathematical Theory of Communication “ от 1948 г. Основните задачи, които възникват с развитието на теорията на кодирането, са свързани с намирането на кодове с определени параметри. За решаване на тези задачи се използват знания от различни области на математиката – алгебра, геометрия, комбинаторика, информатика. Разработването на различни методи и подходи за атакуване на проблематиката на теорията на кодирането води до обогатяване и на всяка от използваните области на математиката. От друга страна, успоредно с решаването на основните задачи от теорията на кодирането възникват проблеми, чието решаване представлява чисто математическо предизвикателство.

Основните направления на изследванията са:

1. Изследване на двоични кодове, получени от разширени кодове на Рид-Соломон над поле с характеристика 2.

2. Изследване на функцията $K_q(n, R)$ - минималното естествено число, за което съществува q-ичен код с дължина n и радиус на покритие R. Определяне на минималната мощност на код (двоичен, троичен или смесен двоичен/троичен) със зададена дължина и минимално разстояние.

3. Получаване на точни стойности за мощността на оптимални двоични кодове със зададена дължина и минимално разстояние 3 или 4. Изследване на оптимални линейни кодове.

4. Определяне на оптимални стратегии при задачи за неадаптивно търсене на неизвестен елемент с множества с равни тегла. Двумерно адаптивно търсене. Задачи за адаптивно търсене на неизвестен елемент в граф, когато след всеки въпрос неизвестния елемент променя позицията си в графа.

5. Определяне на нееквивалентните покрития на $\mathbb{F}_3^8$ със сфери с максимален радиус и определяне точната стойност на мощността на оптимално покритие със сфери на $\mathbb{F}_3^9$.

6. Намиране на граници за мощността на кодове с дадена дължина, поправящи определен брой изтривания. Изследване на комбинаторни задачи, свързани с кодове, поправящи изтривания.

Основни приноси в областите на изследване

1. **Изследване на двоични кодове, получени от разширени кодове на Рид-Соломон над поле с характеристика 2**

   Кодовете на Рид-Соломон са циклични кодове с пораждащ полином

   $$g(x) = (x - \alpha^b)(x - \alpha^b) \dots (x - \alpha^{b+D-2}),$$

   където α е примитивен елемент на полето $GF(q)$. Дължината на кода е $N = q^m - 1$, а минималното му разстояние е равно на D. Код на Рид-Соломон с пораждащ полином

   $$g(x) = (x - \alpha)(x - \alpha^2) \dots (x - \alpha^{n-k})$$

   е $[n, k, n - k + 1]$ код и се означава с RS_k. Добавяйки проверка по четност към всяка кодова дума, получаваме разширен код на Рид-Соломон, който означаваме с ERS_k.

   Нека $q = 2^m$ и $(\beta_1, \beta_2, \dots, \beta_m)$ е базис на $GF(2^m)$ над $GF(2)$. Тогава всеки символ от $GF(2^m)$ може да бъде заместен със съответния двоичен вектор с дължина m. По този начин от код на Рид-Соломон (RS_k или ERS_k) се получава двоичен код с дължина $n.m$, който се използва за практически цели (например при запазване на информация в CD-та). Тегловния спектър на получения двоичен код зависи от използвания базис.

   Коригиращите свойства на един код зависят от неговото минимално разстояние и тегловното му разпределение.

   Определени са спектрите на двоичните образи на разширените $[2^m, 5]$ и $[2^m, 4]$ кодове на Рид-Соломон, като е показана връзката между използвания базис и съответния спектър. Като частен случай на конкатенация са разгледани кодовете на Justesen. Определени са възможните тегла на кодове на Justesen при размерности $k = 3, 4, 5, 6$.

   Получените резултати са публикувани в статии номера 1, 4, 6, 7, 10 и 11.

2. **Изследване на функцията $K_q(n, R)$ - минималното естествено число, за което съществува q-ичен код с дължина n и радиус на покритие R. Определяне на минималната мощност на код (двоичен, троичен или смесен двоичен/троичен) със зададена дължина и минимално разстояние**

   Изследвана е задачата за намиране на минималният брой кодови думи за двоични или смесени двоични/троични кодове с дадена дължина и даден радиус на покритие. С

$K(t,b,R)$ означаваме минималният брой кодови думи на код с t троични и b двоични координати и ирадиус на покритие R.

Радиусът на покритие R е важна характеристика на всеки код. Той се дефинира като минималното естествено число, за което кълбетата с центрове кодовите думи и радиус R покриват цялото пространство.

Класифицирани са троичните кодове с дължина 5 и радиус на покритие 1. Оказва се, че всеки такъв код се получава като директно произведение на $GF(3)$ и троичния код на Хеминг с дължина 4. Интересно е да се отбележи, че определянето на $K_3(6,1)$ все още е нерешена задача, като е известно, че $71 \leq K_3(6,1) \leq 73$. Долната граница $71 \leq K_3(6,1)$ е получена с компютърни пресмятания равностойни на 140 години компютърно време.

Намерени са следните граници и точни стойности за оптимални покриващи кодове:

$$K_2(9,1) \geq 57, K_{3,2}(1,5,1) = 16, K_{3,2}(2,4,1) = 20, K_{3,2}(4,2,1) = 36,$$

$$K_{3,2}(5,0,2) = 8, K_{3,2}(4,1,2) = 6, K_{3,2}(2,2R+1,R) = 6,$$

$$K_{3,2}(2,2R-1,R) = 4, K_{3,2}(2,2R,R) = 6.$$

Получените резултати са публикувани в статии номера 2, 3, 5, 12, 16 и 31.

3. **Получаване на точни стойности за мощността на оптимални двоични кодове със зададена дължина и минимално разстояние 3 или 4. Изследване на оптимални линейни кодове**

Изучаването на функцията $A(n,d)$ - най-голямото естествено число, за което съществува двоичен код с дължина n и минимално разстояние d, предизвиква значителен интерес, особено в първите години на развитие на теорията на кодирането.

Долните граници за $A(n,d)$ са винаги конструктивни, т.е. съществуването на (n,M,d) код означава, че $A(n,d) \geq M$. За намирането на подходящи кодове (такива с голяма мощност) се използват различни подходи. При „малки“ стойности на n построяването на оптимален код може да се извърши с помощта на комбинаторни съображения. Оказва се, че с увеличаването на n, трудността на задачата нараства експоненциално, т.е. получаването на кодове с голяма мощност с използването на чисто комбинаторни разсъждения рядко води до намирането на оптимални кодове.

За намиране на горни граници за $A(n,d)$ се използват комбинаторни методи. За доказване, че $A(n,d) < M$ трябва да се докаже, че не съществува (n,M,d) код. Основната трудност е свързването на „локалната“ характеристика минимално разстояние d с „глобалната“ характеристика брой на кодовите думи M.

Стандартният комбинаторен подход за доказване на несъществуване на (n, M, d) код включва подходящо разделяне на кодовите думи на този код на групи в зависимост от някоя от следните характеристики.

Основният резултат е решаването на първия отворен до момента случай – намирането на точните стойности $A(10, 3) = 72$ и $A(11, 3) = 144$, както и определянето на всички 562 нееквивалентни $(10, 72, 3)$ кода и всички 7398 нееквивалентни $(11, 144, 3)$ кода. За целта са класифицирани кодове с по-малка дължина, след което съответните кодове са разширявани и тествани за еквивалентност.

Получените резултати са публикувани в статии номера 13, 14 и 17.

С $[n, k, d]_q$ се бележи линеен код над полето $GF(q)$. Основна задача в теория на кодирането е оптимизирането на един от параметрите n или d при фиксирани останали параметри. Изследвани са основните подходи за получаване на оптимални линейни кодове. Направен е обзор на получените резултати.

Получените резултати са публикувани в статия 15.

4. **Определяне на оптимални стратегии при задачи за неадаптивно търсене на неизвестен елемент с множества с равни тегла. Двумерно адаптивно търсене. Задачи за адаптивно търсене на неизвестен елемент в граф, когато след всеки въпрос неизвестния елемент променя позицията си в графа**

Общата постановка на класическата задача за търсене е следната: Дадено е множество A от което е избран неизвестен за нас елемент x. Можем да задаваме въпроси от вида: принадлежи ли елемента x на избрано от нас подмножество B на A. Отговорът на всеки от въпросите е „да“ или „не“. Целта е да намерим елемента x с възможно най-малко въпроси. Когато елемента x е намерен, казваме, че сме решили задачата за търсене.

В зависимост от начина на задаване на въпросите са възможни следните видове търсене: адаптивно търсене и неадаптивно търсене.

Известно е, че за решаване на класическата задача за търсене са необходими поне $\lceil \log_2 |A| \rceil$ въпроса.

Разглеждаме произволно крайно множество A и функция $w : A \to \mathbb{N}$, наречена *теглова функция* за множеството A. За произволно подмножество B на A по естествен начин се дефинира *тегло на подмножеството* B.

За неизвестен елемент $x \in A$ и дадено естествено число S множествата-въпроси са онези подмножества B на A за които $w(B) = S$. Естественото число S се нарича „добро “ ако съответната задача е решима, т.е. неизвестния елемент x може да бъде намерен с неадаптивно търсене. Естественото число S се нарича *подходящо*, ако неизвестния елемент x може да бъде намерен с минималния възможен брой въпроси.

Основната задача, която е разгледана е намирането на всички добри и подходящи числа при $A = \{a_1, a_2, \ldots, a_{2^n}\}$ и теглова функция от вида:

$$w_h(a_i) = \left[\frac{i-1}{2^h}\right] + 1$$

за $h = 1, 2, \ldots, n$ където $[x]$ е цялата част на x.

За различни стойности на h получаваме различни теглови функции и съответно различни по трудност задачи за търсене. При $h = n, 0, 2^k, n-1$ са намерени всички подходящи числа. При $h = n-2$ са получени граници за подходящите числа.

Разгледана е задачата за търсене на два елемента, като са получени граници за добрите числа.

Изследвана е и задачата, при допускане на грешни отговори. Показана е връзката на тази задача със съществуването на цикличен код с нечетна дължина и минимално разстояние 3.

Разгледана е задачата за адаптивно търсене на неизвестен единичен квадрат в даден правоъгълник. Разрешените въпроси са от вида: принадлежи ли неизвестният квадрат на правоъгълник, чийто горен ляв връх съвдапа с горния ляв връх на дадения правоъгълник.

Да означим с $t(m, n)$ минималният брой въпроси, необходими за намиране на неизвестния елемент. Лесно се доказва, че за $t(m, n)$ има само две възможности – $\lceil \log m + \log n \rceil$ или $\lceil \log m + \log n \rceil + 1$.

Правоъгълник се нарича *разрешим* ако съществува алгоритъм, чрез който неизвестният елемент може да бъде намерен с $\lceil \log m + \log n \rceil$ въпроса. В противен случай правоъгълникът се нарича *неразрешим*.

Доказано, че за всяко m от вида $m = \dfrac{2^{si} - 1}{2^s - 1}$ и всяко n правоъгълник $m \times n$ е разрешим. Намерени са най-малкият неразрешим правоъгълник 11×93 и най-малкият неразрешим квадрат 181×181.

Разгледана е задачата за адаптивно търсене на неизвестен елемент в граф, когато след всеки въпрос неизвестния елемент се премества в съседен връх. Когато с всеки въпрос се проверява дали целта е в някои k върха на графа, задачата се нарича k-търсене.

Характеризирани са всички графи, за които съществува печеливша стратегия при $k = 1$. В случая на двумерена квадратна решетка с четна дължина е определено минималното k, за което съществува печеливша стратегия.

Получените резултати са публикувани в статии номера 18, 19, 20, 21, 22, 23, 24, 25, 26 и 35.

5. **Определяне на нееквивалентните покрития на $\mathbb{F}_3^8$ със сфери и определяне точната стойност на мощността на оптимално покритие със сфери на $\mathbb{F}_3^9$.**

   Разгледана е задачата за намиране на оптимални покрития на $\mathbb{F}_3^n$ за $n \leq 13$ със сфери с радиус n. Това означава, че търсим код C със следното свойство: за всеки елемент $\mathbf{y} \in \mathbb{F}_3^n$ съществува $\mathbf{x} \in C$, за който $\mathbf{d}(\mathbf{x}, \mathbf{y}) = n$. Минималната мощност на такъв код се бележи с $T(n)$. Кодът се нарича оптимален, ако $|C| = T(n)$.

   Редицата от стойностите на $T(n)$ е част от The on-line encyclopedia of integer sequences номер A086676.

   Доказано е, че съществуват две нееквивалентни покрития на $\mathbb{F}_3^8$. Намерена е намерена точната стойност $T(9) = 68$, което води до подобряване на известните граници за $T(n)$, както следва $T(10) \geq 102$, $T(11) \geq 153$, $T(12) \geq 230$ и $T(13) \geq 345$.

   Получените резултати са публикувани в статии номера 27, 28, 29 и 30.

6. **Намиране на граници за мощността на двоични кодове с дадена дължина, поправящи определен брой изтривания. Изследване на комбинаторни задачи, свързани с кодове, поправящи изтривания**

   При предаване на информация освен грешно приемане на даден символ е възможно и загуба на символи. По този начин получателят на съобщението получава по-къса дума и не знае къде точно е станало изтриването на символи. Кодовете, поправящи изтривания възстановяват изпратената кодова дума, при условие, че са станали до определн брой изтривания.

   Доказано е, че оптималния двоичен код с дължина $2t + 3$, поправящ t изтривания е с мощност 6 при $t = 1$ и мощност 5 при $t \geq 2$.

   Получени са точни стойности за мощността на оптимални двоични кодове, поправящи изтривания.

   При решаване на основната задача за намиране на оптимален код поправящ изтривания особено важно е да се определят всички двойки доминантни вектори при конкретно t. Тази задача е решена при $t = 1, 2$.

   Получените резултати са публикувани в статии номера 32, 33 и 34.