Совершенные раскраски и дизайны

Для чего они нужны?

Совершенные структуры в кодировании и криптографии

Совершенная раскраска

Пусть $\{1, \ldots, k\}$ — множество из k цветов. Тогда раскраской вершин графа $G = (V, E)$ называется функция $f : V \to \{1, \ldots, k\}$.

Для каждой вершины $x \in V(G)$ определим окрестность вершины (единичную сферу с центром в вершине) следующим образом: $U_1(x) = \{y \in V \mid \{x, y\} \in E\}$. Множество $U_1(x)$ состоит из вершин смежных с вершиной x. $B_1(x) = \{x\} \cup U_1$ — шар с центром в вершине.

Пусть $f(x) = i$, обозначим $s_{ij}(x) = |U_1(x) \cap f^{-1}(j)|$.

Определение

Раскраска называется совершенной, если величина $s_{ij}(x) = s_{ij}$ не зависит от выбора вершины x цвета i.

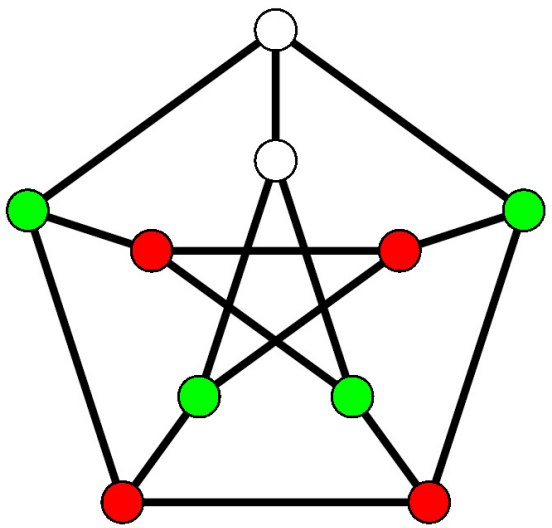
Совершенная раскраска

Определение

Матрицей параметров совершенной раскраски называется матрица

$$S_{k\times k} = (s_{ij})$$

Совершенная раскраска графа Петерсена, порядок цветов: белый, зелёный, красный.



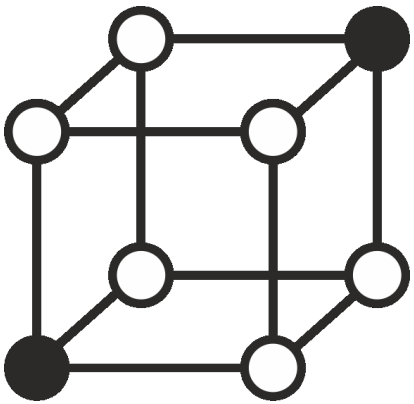
$$S = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

Совершенный код

Определение

Подмножество C вершин графа называется совершенным кодом, если все вершины графа разбиваются на шары с центрами в вершинах из C, т. е. $V = \bigcup\limits_{x \in C} B_1(x)$ и $B_1(x) \cap B_1(y) = \varnothing$, если $x, y \in C$.

Для r-регулярного графа это условие означает, что раскраска кодовых вершин в один цвет, а некодовых – в другой является совершенной с матрицей параметров $\begin{pmatrix} 0 & r \\ 1 & r-1 \end{pmatrix}$.



$$S = \begin{pmatrix} 0 & 3 \\ 1 & 2 \end{pmatrix}$$

Совершенный код

Рассмотрим следующую модель передачи сообщений по каналу связи. Каждому возможному сообщению поставим в соответствие вершину графа. Соединим рёбрами две вершины u и v, если при передаче по каналу связи какая-либо ошибка может превратить сообщение u в сообщение v.

Определение

Подможество вершин графа называется кодом, исправляющим ошибку, если шары радиуса 1 с центрами в кодовых вершинах не пересекаются.

Совершенный код

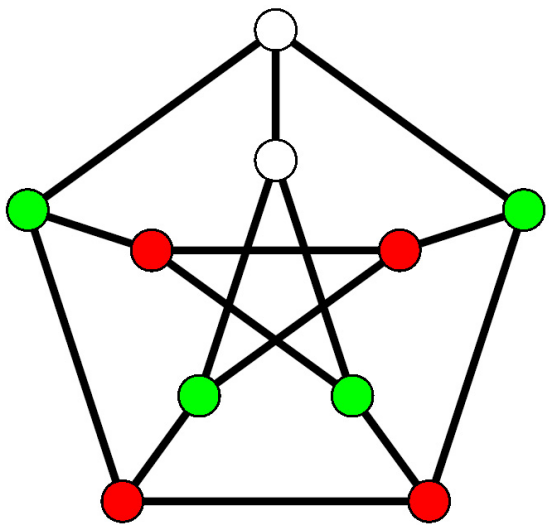
Будем передавать только те сообщения, которые соответствуют вершинам кода. Тогда после любой ошибки на приёмном конце канала связи можно выяснить какое сообщение было передано. Чем больше мощность множества кодовых вершин, тем больше информации можно передать по каналу связи за единицу времени. Пусть все шары радиуса 1 в графе G имеют одинаковую мощность $r+1$, т. е. граф G — r-регулярный. Тогда для мощности кода C, исправляющего ошибку, справедливо неравенство $(r+1)|C| \leq |V(G)|$. Это неравенство носит название границы Хэмминга мощности кода. Нетрудно видеть, что код имеет максимальную мощность, когда граф полностью разбивается на шары радиуса 1 с центрами в кодовых вершинах, т.е. когда код совершенный.

Независимое множество

Определение

Подможество вершин графа называется независимым, если оно не содержит смежных вершин.

Если $s_{ii} = 0$, то i-ый цвет совершенной раскраски является независимым множеством.

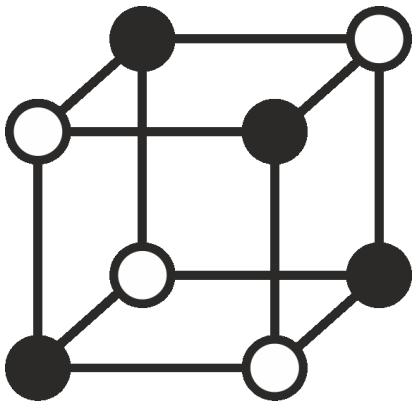


$$S = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

Независимое множество

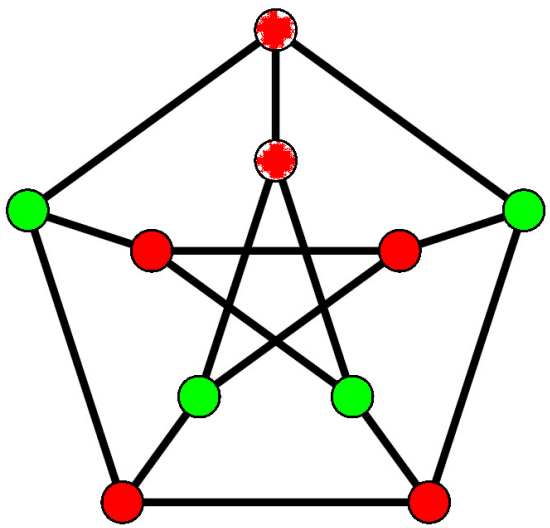
Теперь предположим, что мы всегда можем запросить повторную передачу испорченного сообщения. Тогда нам достаточно знать, что при передачи сообщения произошла ошибка. Если передавать по каналу связи только те сообщения, которые соответствуют независимым вершинам, то ошибочное сообщение всегда распознаётся. В r-регулярном графе независимое множество не может содержать больше половины вершин графа. А если в графе найдётся независимое множество такой мощности, то вторая половина вершин также оказывается независимым множеством и раскраска двух долей графа в разные цвета является совершенной с матрицей

параметров $\begin{pmatrix} 0 & r \\ r & 0 \end{pmatrix}$.



Независимое множество

Хоффманом доказана верхняя оценка $\frac{\lambda|V(G)|}{\lambda - r}$ мощности независимого множества для r-регулярных графов (λ — минимальное собственное число матрицы смежности графа). При достижении этой оценки характеристическая функция независимого множества оказывается совершенной раскраской. В графе Петерсена $\lambda_{min} = -2$.



$$S = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$$

Латинская раскраска

Определение

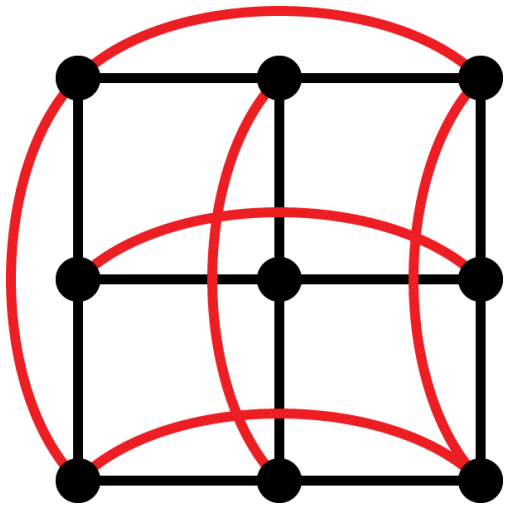
Латинским квадратом называется таблица $k \times k$, заполненная k различными символами так, чтобы в каждой линии (строке или столбце) любой символ встречался по одному разу.

0	2	1
1	0	2
2	1	0

0	2	3	1
3	1	0	2
1	3	2	0
2	0	1	3

Латинская раскраска

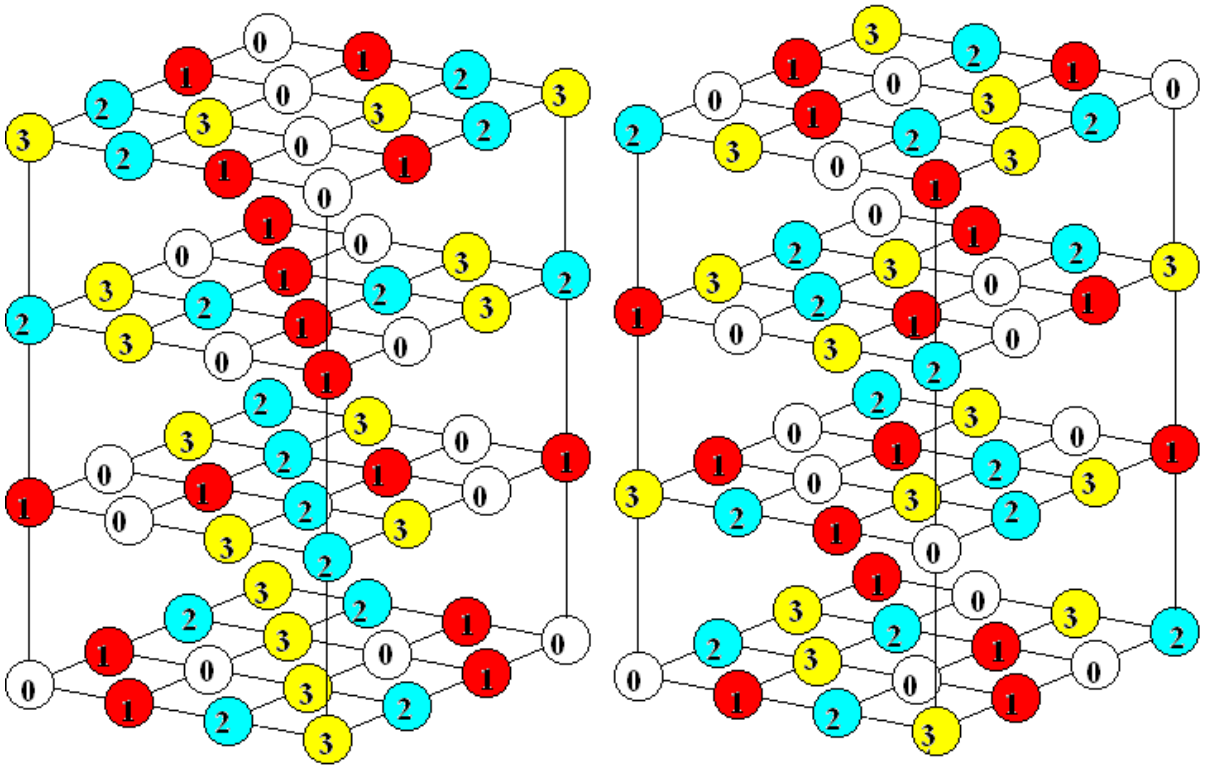
Каждую клетку таблицы будем считать вершиной графа, две вершины — смежными, если они лежат на одной линии. Полученный граф называется графом Хэмминга $H(2, k)$ (здесь k — размер таблицы, 2 — её размерность). Граф $H(2, 3)$.



Будем считать цветом вершины символ в клетке таблице. Тогда латинскому квадрату соответствует совершенная раскраска графа $H(2, 3)$ с матрицей параметров $S = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$.

Латинская раскраска

Аналогичным образом можно определить латинский куб, которому соответствует совершенная раскраска графа $H(3, k)$ с матрицей параметров $S = \begin{pmatrix} 0 & 3 & 3 & 3 \\ 3 & 0 & 3 & 3 \\ 3 & 3 & 0 & 3 \\ 3 & 3 & 3 & 0 \end{pmatrix}$.



Латинская раскраска

Латинский n-мерный куб можно представить как таблицу значений функции $f(x_1, \ldots, x_n)$.
Рассмотрим задачу передачи сообщения по каналу связи со стиранием, т. е. при передаче сообщение $(a_1, a_2, \ldots, a_n, f(a_1, \ldots, a_n))$ преобразуется в сообщение вида $(a_1, *, a_3, \ldots, a_n, f(a_1, \ldots, a_n))$, где символ $*$ заменяет утраченную букву. При потере только одного символа мы знаем линию и символ. Значит по определению латинского гиперкуба мы можем восстановить утерянную координату.
Рассмотрим задачу распределённого хранения информации. Пусть информация хранится на n серверах и нужно уметь восстанавливать всю информацию при поломке одного сервера. В этом случае, информацию удобно представлять в виде слов $(a_1, a_2, \ldots, a_{n-1}, f(a_1, \ldots, a_{n-1}))$ так, что i-ый символ кодового слова хранится на i-ом сервере. Тогда потеря сервера эквивалентна стиранию одной буквы в каждом кодовом слове.

Комбинаторные дизайны

Рассмотрим множество из n элементов. Блоками, точнее k-блоками будем называть любое его k-элементное подмножество.

Определение

Комбинаторным дизайном (точнее t-дизайном) с параметрами $t-(n,k,\lambda)$ называется такой набор блоков, что любое t-элементное подмножество содержится ровно в λ k-блоках.

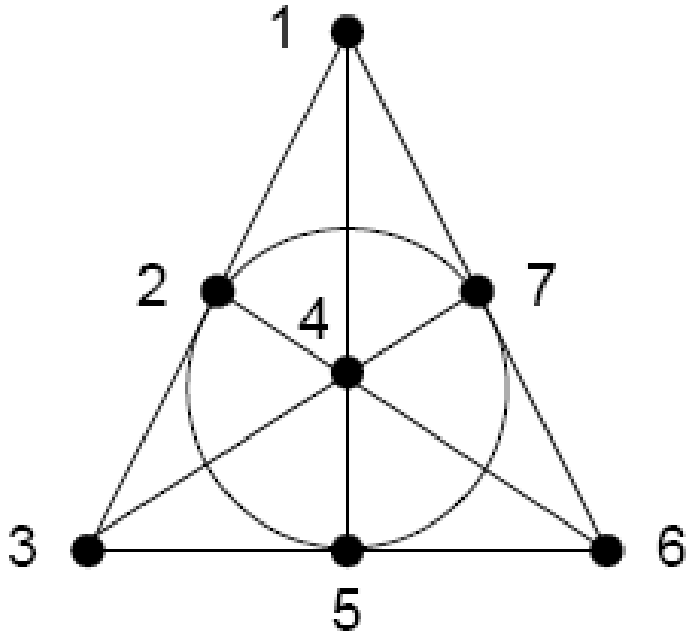
Примером $2-(7,3,1)$ дизайна является плоскость Фано.

{1,2,3}, {1,4,5}, {1,6,7}, {2,4,6}, {2,5,7}, {3,4,7}, {3,5,6}

Комбинаторные дизайны

{1,2,3}, {1,4,5}, {1,6,7}, {2,4,6}, {2,5,7}, {3,4,7}, {3,5,6}

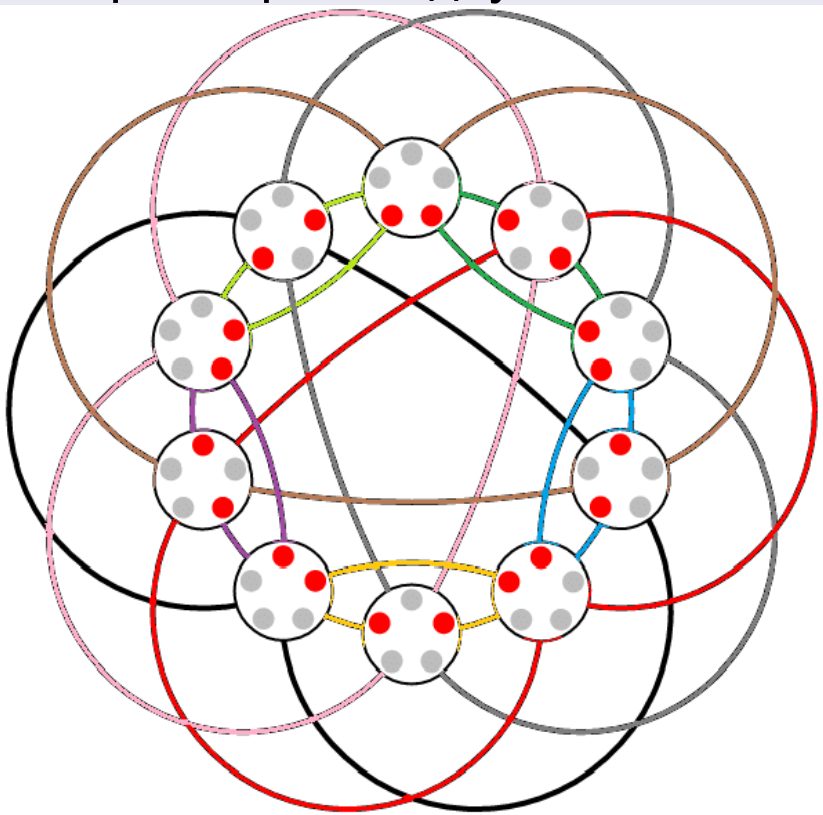
$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$



Определение

Графом Джонсона $J(n, k)$ называется граф, вершинами которого являются все двоичные наборы длины n и веса k. Две вершины графа соединены ребром, если расстояние Хэмминга (число несовпадающих элементов) между соответствующими им двоичными наборами равно двум.

$J(5, 2)$



Утверждение

Характеристическая функция дизайна D с параметрами $(k-1)$-(n,k,λ) является совершенной раскраской графа $J(n,k)$ с матрицей параметров $S = \begin{pmatrix} k(\lambda-1) & k(n-k-\lambda+1) \\ k\lambda & k(n-k-\lambda) \end{pmatrix}$.

При $\lambda = 1$ $S = \begin{pmatrix} 0 & k(n-k) \\ k & k(n-k-1) \end{pmatrix}$.

Пример

Характеристическая функция плоскости Фано с параметрами 2-$(7,3,1)$ является совершенной раскраской графа $J(7,3)$ с матрицей параметров $S = \begin{pmatrix} 0 & 12 \\ 3 & 9 \end{pmatrix}$.

Комбинаторные дизайны

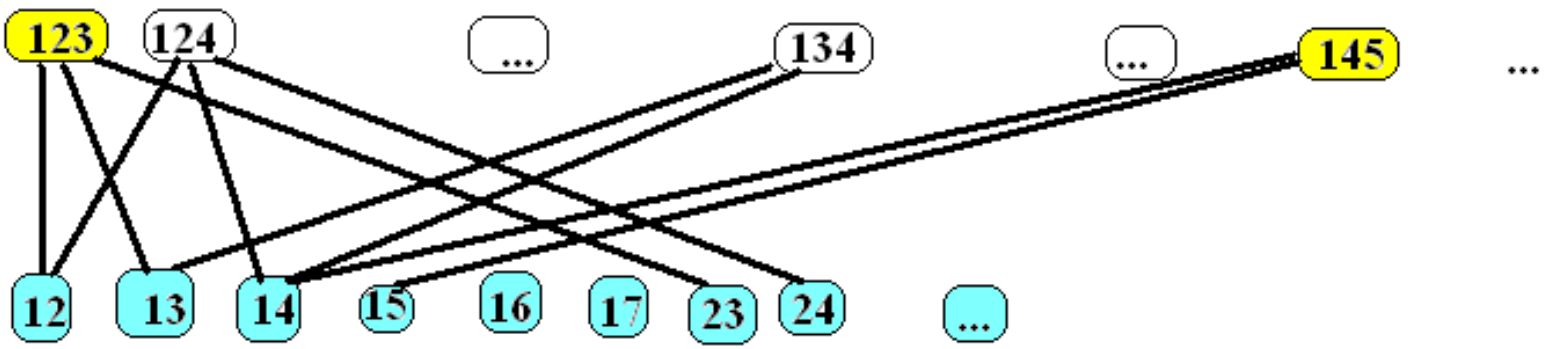
Рассмотрим задачу оценки экспертами заявок. Пусть имеется n заявок, каждый эксперт рассматривает k заявок так, что каждый набор из t заявок рассматривается λ экспертами. Тогда эксперт – это k-блок, а распределение заявок по экспертам – это комбинаторный дизайн с параметрами $t-(n,k,\lambda)$.

Плоскость Фано, дизайн с параметрами $2-(7,3,1)$. E_i — эксперты, Z_j — заявки.

$$\begin{pmatrix} & Z_1 & Z_2 & Z_3 & Z_4 & Z_5 & Z_6 & Z_7 \\ E_1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ E_2 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ E_3 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ E_4 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ E_5 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ E_6 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ E_7 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

Комбинаторные дизайны

Совершенными раскрасками в графах Джонсона $J(n, k)$ являются дизайны с параметрами t-(n, k, λ) при $t = k - 1$.
Теперь рассмотрим произвольные t, $t < k$.
Определим двудольный граф $B(n, k, t)$. Вершинами одной доли графа $B(n, k, t)$ будут k-блоки, а вершинами другой доли t-блоки. Вершина — k-блок соединена с вершиной — t-блоком,



Утверждение

Покрасим все t-блоки в 1й цвет, k-блоки из дизайна t-(n, k, λ) во 2й, оставшиеся k-блоки в 3й. Полученная раскраска является совершенной раскраской графа $B(n, k, t)$ с матрицей параметров $S = \begin{pmatrix} 0 & \lambda & \frac{(n-t)!}{(k-t)!(n-k)!} - \lambda \\ \frac{k!}{(k-t)!t!} & 0 & 0 \\ \frac{k!}{(k-t)!t!} & 0 & 0 \end{pmatrix}$.

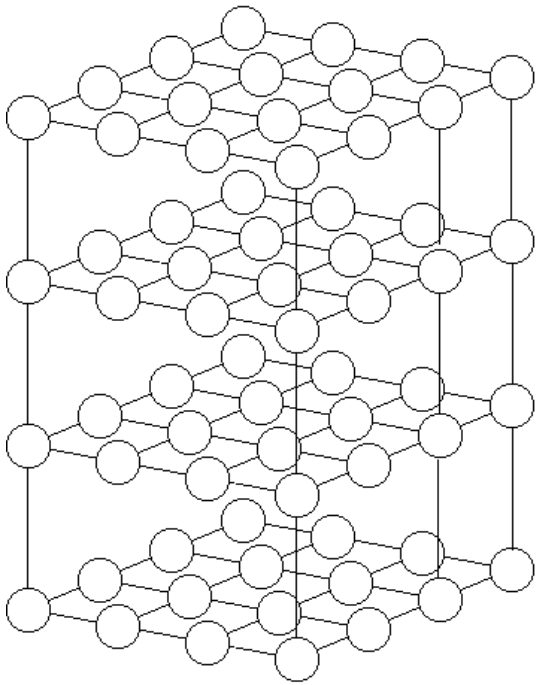
Плоскость Фано в $B(7, 3, 2)$ определяют раскраску с матрицей параметров $S = \begin{pmatrix} 0 & 1 & 5 \\ 3 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix}$.

МДР-коды, трансверсальные дизайны, GDD

Гиперкубом (графом Хэмминга) $H(n, q)$ называется граф, вершинами которого являются q-ичные вектора длины n. Два вектора соединены ребром, если они отличаются в одной координате.

Множество вершин гиперкуба $H(n, q)$, у которых зафиксировано значение нескольких координат называется гранью. Размерность грани равна числу свободных координат.

$H(3, 4)$



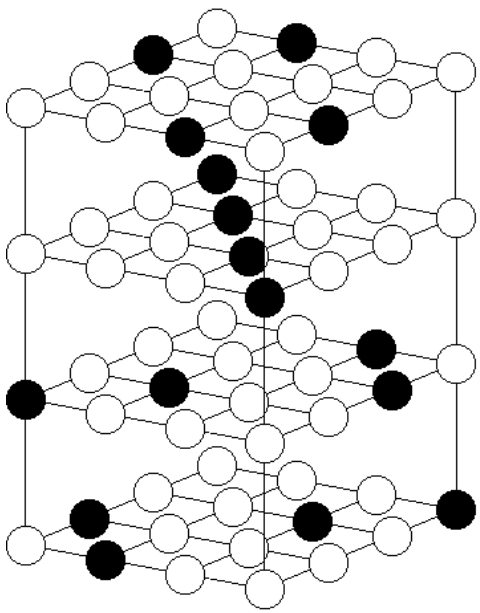
МДР-коды, трансверсальные дизайны, GDD

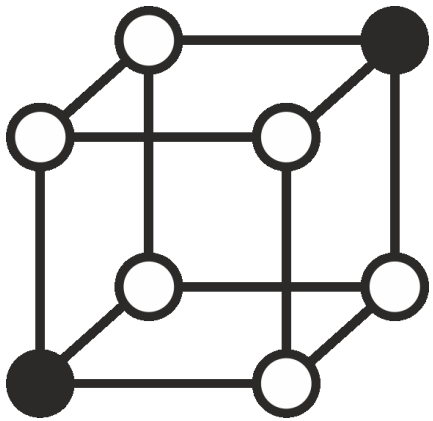
Определение

Множество $C \subset H(n, q)$ называется МДР-кодом с расстоянием d, если $|C \cap \Gamma| = 1$ для любой $(d-1)$-мерной грани Γ.

Граница Синглтона

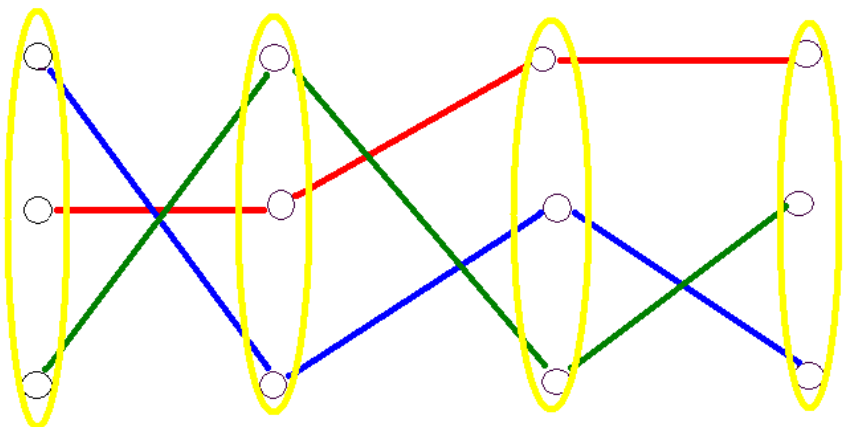
Код C с кодовым расстоянием d является МДР-кодом тогда и только тогда когда $|C| = q^{n-d+1}$.





МДР-коды, трансверсальные дизайны, GDD

Пусть имеется n групп по q элементов каждая. Можно считать, что все группы состоят из одинаковых элементов $Q_q = \{0, 1, \ldots, q-1\}$. Блоками будем называть наборы из n элементов по одному из каждой группы. Набор блоков — это GDD с параметрами $t-(n, n, \lambda)_q$, если каждый набор из t элементов (не более одного из каждой группы) содержится в λ блоках.

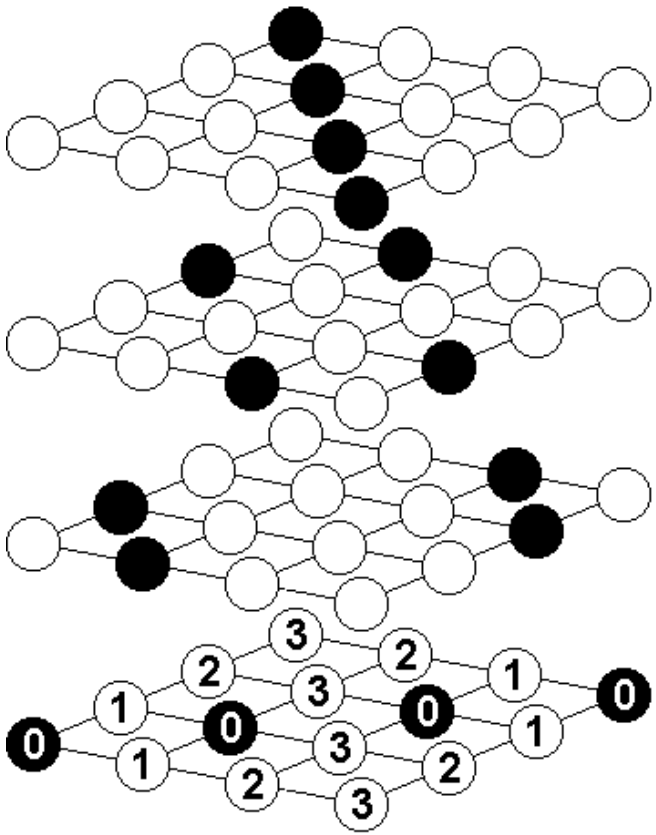


Утверждение

GDD с параметрами $t-(n, n, 1)_q$ эквивалентен МДР-коду в $H(n, q)$ с кодовым расстоянием $n-t+1$.

МДР-коды, трансверсальные дизайны, GDD

Латинский квадрат размера $q \times q$ можно рассматривать как таблицу значений некоторой функции $f : Q_q^2 \to Q_q$, $Q_q = \{0, 1, \ldots, q-1\}$.
Множество $C = \{(x, y, f(x, y)) : x, y \in Q_q\}$ является МДР-кодом с расстоянием 2.



Латинский k-мерный куб порядка q можно рассматривать как таблицу значений некоторой функции $f : Q_q^k \to Q_q$,
$Q_q = \{0, 1, \ldots, q-1\}$.
Множество $C = \{(x_1, x_2, \ldots, x_k, f(x_1, \ldots, x_k)) : x_i \in Q_q\}$
является МДР-кодом с расстоянием 2.

Определение

Два латинских квадрата f и g называются ортогональными, если все пары $(f(x_1, x_2), g(x_1, x_2))$ при $(x_1, x_2) \in Q_q^2$, различны.

Попарно ортогональные латинские квадраты порядка 4.

0	**2**	3	1
3	1	0	**2**
1	3	**2**	0
2	0	1	3

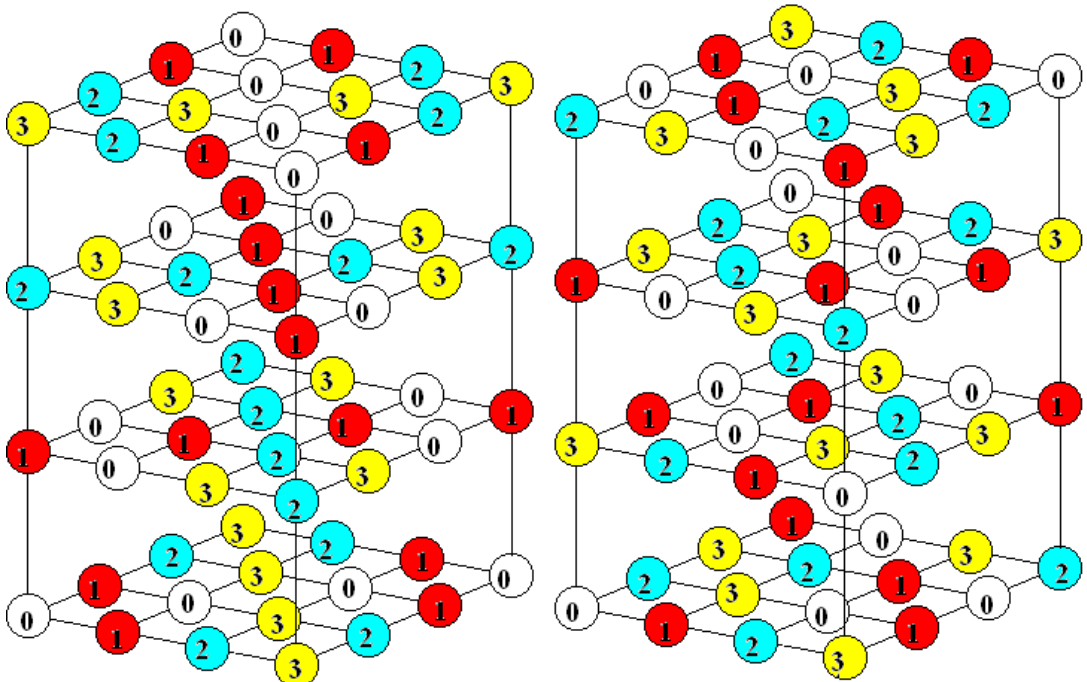
0	**2**	3	1
2	0	1	**3**
3	1	**0**	2
1	3	2	0

0	**2**	3	1
1	3	2	**0**
2	0	**1**	3
3	1	0	2

МДР-коды, трансверсальные дизайны, GDD

Определение

Два латинских куба f и g называются ортогональными, если их одинаковые грани содержат ортогональные латинские квадраты, три латинских куба f, g, h называются ортогональными, если они попарно ортогональны и все тройки $(f(x_1, x_2, x_3), g(x_1, x_2, x_3), h(x_1, x_2, x_3))$ при $(x_1, x_2, x_3) \in Q_q^3$, различны.



МДР-коды, трансверсальные дизайны, GDD

Утверждение

Пусть $f_1, f_2, \ldots, f_k$ набор попарно ортогональных латинских квадратов. Тогда множество
$C = \{(x, y, f_1(x, y), f_2(x, y), \ldots, f_k(x, y)) : x, y \in Q_q\}$ является МДР-кодом с расстоянием $k + 1$.

Утверждение

Пусть $f_1, f_2, \ldots, f_k$ набор ортогональных латинских кубов. Тогда множество
$C = \{(x, y, z, f_1(x, y, z), f_2(x, y, z), \ldots, f_k(x, y, z)) : x, y, z \in Q_q\}$ является МДР-кодом с расстоянием $k + 1$.

МДР-коды, трансверсальные дизайны, GDD

Определим двудольный граф $\Gamma(n, k, q)$. Вершинами одной доли графа $\Gamma(n, k, q)$ будут k-мерные грани в $H(n, q)$, а вершинами другой доли — вершины из гиперкуба $H(n, q)$. Вершина — k-грань соединена с вершиной гиперкуба, если вершина гиперкуба содержится в грани.

Утверждение

Покрасим все грани в 1й цвет, вершины из МДР-кода в $H(n, q)$ с расстоянием $k + 1$ во 2й цвет, оставшиеся вершины в 3й. Полученная раскраска является совершенной раскраской графа $\Gamma(n, k, q)$ с матрицей параметров $S = \begin{pmatrix} 0 & 1 & q^k - 1 \\ \frac{n!}{(n-k)!k!} & 0 & 0 \\ \frac{n!}{(n-k)!k!} & 0 & 0 \end{pmatrix}$.

Вершина $(a_1, a_2, \ldots, a_n)$ видит гиперграни вида $(a_1, *, \ldots, a_n)$, где k элементов стёрты.

МДР-коды, трансверсальные дизайны, GDD

МДР-код можно использовать для решения задачи разделения ключа. Задача состоит в следующем. Пусть имеется коллектив из $n-1$ участников. Требуется так разделить между участниками информацию о секретном ключе, чтобы любые m участников вместе могли узнать секретный ключ, а любые $m-1$ участников не только не могли, но и даже не имели бы никакой информации о секретном ключе.

Пусть имеется МДР-код длины n с кодовым расстоянием $n-m+1$. Сам МДР-код считается общеизвестным. Рассмотрим произвольное кодовое слово $(a_1, a_2, \ldots, a_n)$. В качестве индивидуального ключа i-го участника возьмём $a_i \in Q_q$, а в качестве секретного ключа $a_n \in Q_q$. Какой бы ни была группа из m участников мы можем определить $n-m$ мерную грань, в которой лежит слово $(a_1, a_2, \ldots, a_n)$. По определению МДР-кода в грани лежит только одно слово из кода. Значит мы нашли a_n.

Если в группе только $m-1$ участник, то известны значения только $m-1$ переменной. Тогда нам известна $(n-(m-1))$-мерная грань. В ней лежит q вершин из МДР-кода, у которых координата x_n принимает по одному разу все возможные q значений.