به نام خدا

کدینگ کانال

- مفاهیم پایه

- فاصله همینگ Hamming Distance

- فاصله همینگ حداقل (کمترین فاصله همینگ) کلمات کد $C(n,k)$

برای کد $C(n,k)$ با 2^k کلمه کد c_0، c_1، ...، c_{2^k-1} (در حالت باینری)
عبارتست از کمترین فاصله همینگ کلمات کد. به عبارت دیگر

$$d_{min} = \operatorname{Min} \left\{ d_H(\underline{c}_i, \underline{c}_j), \forall\, c_i \neq c_j \right\}$$

خواهیم دید که برای یک کد $C(n,k)$، d_{min} معیاری از قابلیت تشخیص و تصحیح خطا را به دست می دهد.

مثال: برای کد کانال به صورت $C = \{\overbrace{000}^{\underline{c}_0}, \overbrace{011}^{\underline{c}_1}, \overbrace{101}^{\underline{c}_2}, \overbrace{110}^{\underline{c}_3}\}$ کهریس

فاصله همینگ کلمات کد را به دست بیاورید

با توجه به اینکه 4 کلمه کد داریم، تعداد حالت‌های ممکن برای پیام‌های ورودی برابر 4 است. در حالت باینری

تعداد پیامها $=$ تعداد کلمات کد $= 2^{k} = 4 \longrightarrow k = 2$

طول کلمات کد $= n = 3 \implies C(n,k) = C(3,2)$

پیامها	کلمات کد	
00	000	c_0
01	011	c_1
10	101	c_2
11	110	c_3

$d_H(c_0, c_1) = 2$

$d_H(c_0, c_2) = 2$

$d_H(c_0, c_3) = 2$

$d_H(c_1, c_2) = 2$

$d_H(c_1, c_3) = 2$

$d_H(c_2, c_3) = 2$

$d_{min} = 2$

مثال - برای کد زیر کمترین فاصله همینگ کلمات کد را به دست بیاورید

$$C = \{ \overbrace{0000}^{\underline{c}_0}, \overbrace{0101}^{\underline{c}_1}, \overbrace{1011}^{\underline{c}_2}, \overbrace{1110}^{\underline{c}_3} \} \quad , \quad C(4,2)$$

کلمه پیام	کلمه کد
00	0000
01	0101
10	1011
11	1110

$$d_H(\underline{c}_0, \underline{c}_1) = 2$$

$$d_H(\underline{c}_0, \underline{c}_2) = 3$$

$$d_H(\underline{c}_0, \underline{c}_3) = 3 \quad \longrightarrow d_{min} = 2$$

$$d_H(\underline{c}_1, \underline{c}_2) = 3$$

$$d_H(\underline{c}_1, \underline{c}_3) = 3$$

$$d_H(\underline{c}_2, \underline{c}_3) = 2$$

وزن همینگ یک بردار $\underline{v} = (v_0, v_1, \dots, v_{n-1})$ برابر تعداد مؤلفه‌هایی است که

مؤلفه‌های غیرصفر دارند. به عبارت دیگر

$$W(\underline{v}) = \sum_{i=0}^{n-1} I(v_i)$$

وزن همینگ

که در آن

$$I(v_i) = \begin{cases} 0 & v_i = 0 \\ 1 & v_i \neq 0 \end{cases}$$

به عنوان مثال: وزن همینگ بردارهای زیر را به دست بیاورید:

$$\underline{x} = (1,0,0,1,1,0,1) \Longrightarrow w(\underline{x}) = 4$$

$$\underline{y} = (2,-1,0,0,3,-2,1) \Longrightarrow w(\underline{y}) = 5$$

بین فاصله همینگ دو بردار و وزن همینگ آنها ارتباط وجود دارد.

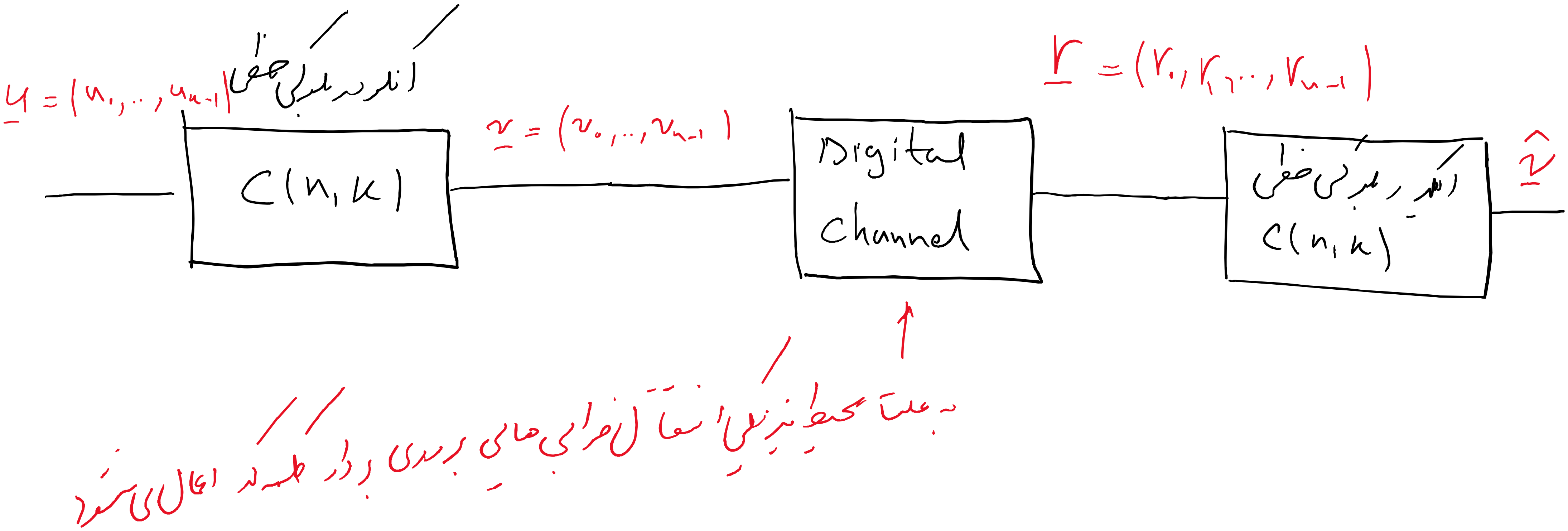
برای دو بردار $\underline{a} = (a_0, a_1, \ldots, a_{n-1})$ ، $\underline{b} = (b_0, \ldots, b_{n-1})$ داریم

$$d_H(\underline{a}, \underline{b}) = w(\underline{a} - \underline{b}) = w(\underline{b} - \underline{a})$$

همان طور که گفتیم برای یک کد $C(n,k)$ ، کمترین فاصله همینگ کلمات کد ، معیاری از توانایی تشخیص و تصحیح خطا را به دست می دهد.

قضیه: برای هر کد $C(n,k)$ با کمترین فاصله همینگ کلمات کد برابر d_{min} ، این کد تا قابلیت تشخیص $s = d_{min} - 1$ خطا، و تا قابلیت تصحیح $t = \left\lfloor \frac{d_{min}-1}{2} \right\rfloor$ خطا را دارد.

این قضیه برای تمامی کدها برقرار است. در ادامه یک اثبات شهودی برای حالتی که کد $C(n,k)$ یک کد بلوکی خطی باشد ارائه می‌کنیم. امیدواریم این دلیل نیز به صورت تصویری قابل فهم باشد.



$$\underline{r} = \underline{v} + \underline{e}$$

$\underline{e} = (e_0, e_1, \dots, e_{n-1})$,
$e_i = 0$ در مکان i ام خطا نداریم
$e_i \neq 0$ در مکان i ام خطایی رخ داده است

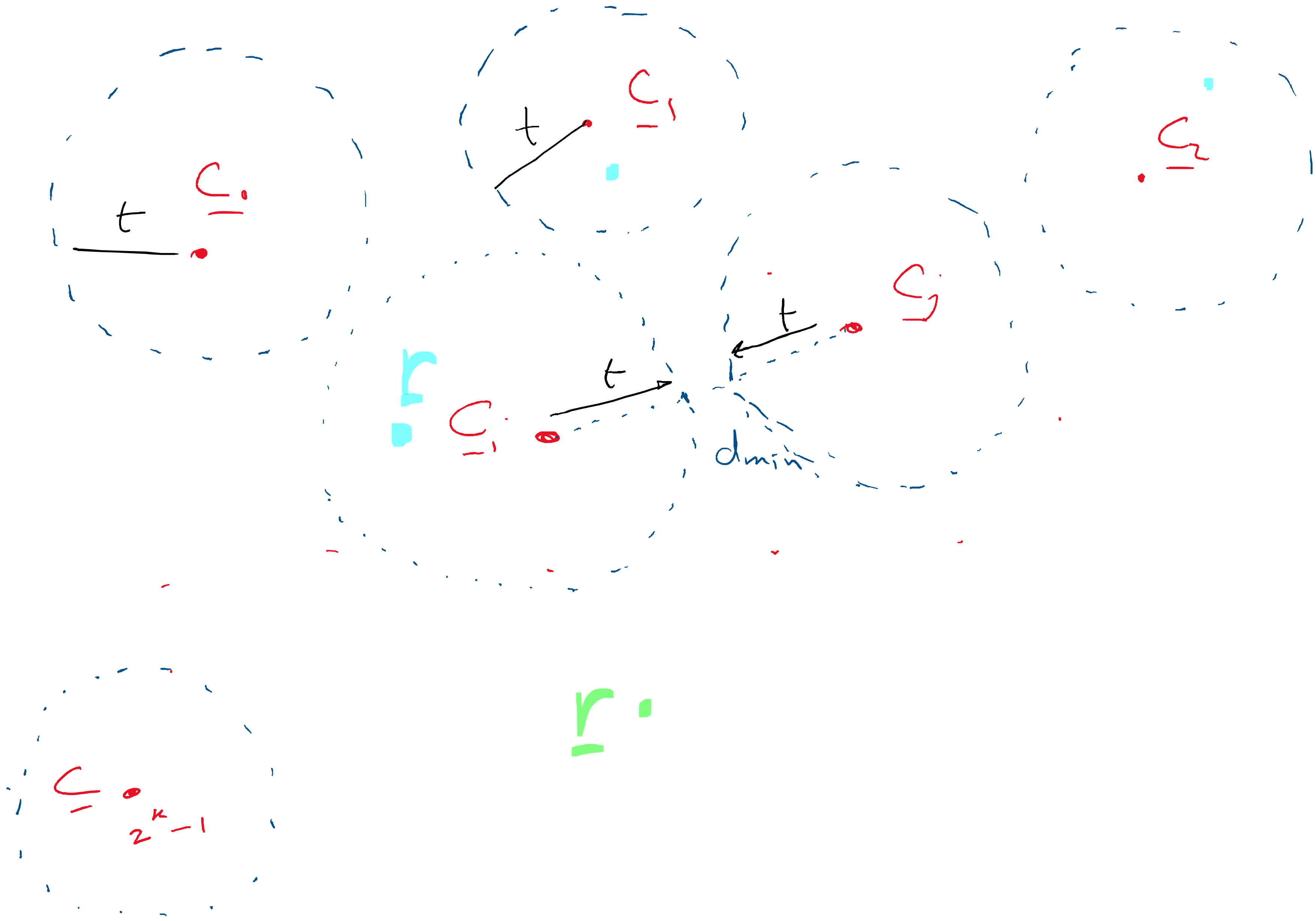
دیکدر را دریافت بردار y ، بر اساس معیارهای مانند بیشترین شباهت یا بیشترین احتمال
پسین یا کمترین احتمال خطا ، کلمه کدی را پیدا می کند که بیشترین شباهت را به y داشته باشد
آن را به عنوان کلمه کد ارسالی بازیابی می کند ($\hat{v}$) . چون ممکن است عملیات
بازیابی با خطا همراه باشد $\hat{v}$ دقیقاً همان v نباشد، کلمه بازیابی شده را با
$\hat{v}$ نمایش داده ایم. پس از آن با توجه به $\hat{v}$ ، دیکدر پیام متناظر با این کلمه کد را
به عنوان پیام ارسالی بازیابی می کند. ($\hat{u}$) که ممکن است با u یکسان نباشد که
به آن خطای دیکدینگ می گوییم. این خطا قابل ارزیابی و محاسبه است و باید به میزان دلخواه کم باشد.

به این ترتیب با دریافت بردار $\underline{r}$ ، اگر بردار $\underline{r}$ با هیچ یک از کلمات کد $C(n,k)$ یکسان نباشد، دریافتگر متوجه می‌شود که خطایی در کانال رخ داده است.

با توجه به اینکه کمترین فاصله همینگ کلمات کد برابر d_{min} است، اگر در کانال به تعداد $d_{min}-1$ یا کمتر خطا رخ دهد، به هیچ عنوان امکان ندارد که یک کلمه کد ارسالی به کلمه کد دیگری از $C(n,k)$ تبدیل شود. بنابراین دریافتگر متوجه رخ دادن خطا می‌شود.

$\Leftarrow$ قابلیت تشخیص خطای کد برابر $d_{min}-1$ است.

* در بحث دیکدینگ، الگوهای خطایی وجود دارند که اگر این الگوهای خطا در کانال رخ بدهند، باعث می‌شود که یک کلمه کد به یک کلمه دیگر از $C(n,k)$ تبدیل شود و دیکدر نتواند رخ دادن خطا تشخیص بدهد. به این خطاها، خطاهای غیرقابل تشخیص (Undetectable errors) گفته می‌شود که باعث ایجاد خطای دیکدینگ می‌شوند. احتمال وقوع چنین خطاهایی نیز معمولا کاملا ناچیز است.



$\underline{c}_0$
t
$\underline{c}_1$
t
$\underline{c}_2$
$\underline{c}_j$
t
r
$\underline{c}_i$
t
dmin
$\underline{c}_{2^k-1}$
r'

$$t = \left\lfloor \frac{d_{min}-1}{2} \right\rfloor \implies t \leq \frac{d_{min}-1}{2}$$

$$\implies d_{min} \geq 2t+1 \qquad (1)$$

می‌دانیم که عملیات تصحیح خطا براساس بیشترین شباهت یا بیشترین احتمال یعنی با کمترین احتمال خطا انجام می‌شود. اگر در فضای n بعدی، به دور هر یک از کلمات کد، یک ...

ابرکره یا یک کره n بعدی (هایپرکره) به شعاع t رسم کنیم، طبق رابطه (1) این کره‌ها هیچ همپوشانی با هم ندارند. از طرف دیگر تمامی بردارهای n تایی که فاصله همینگ آنها از کد

c_i کمتر یا مساوی t است، درون کره به مرکز c_i قرار می‌گیرند. به عبارت

دیگر تمامی بردارهایی که در کره c_i قرار دارند، بیشترین شباهت (کمترین فاصله [illegible]

را با c_i نسبت به سایر کلمات کد $C(n,k)$ خواهند داشت. بنابراین می‌توانیم

دیکدینگ را براساس این کره‌ها انجام دهیم. به همین دلیل به این کره‌ها، کره‌های

دیکدینگ گفته می‌شود. به این ترتیب با دریافت بردار y اگر بردار y در کره

دیکدینگ مربوط به c_i قرار بگیرد، دیکدر c_i را به عنوان کلمه کد ارسالی بازیابی

می‌کنند $\Longleftarrow$ که $C(n,k)$ با کمترین فاصله همینگ d_{min} ، تعداد خطاها را بطور قطع کاملاً تصحیح کنند

* اگر بردار r در کره دیکدینگ هیچ کلمه کدی از $C(n,k)$ قرار نگیرد، دیکدر نمی‌تواند تصمیم بگیرد که کدام یک از کلمات کد ارسال شده است. در این صورت می‌گوییم که دیکدر دچار failure شده است. به این نوع دیکدر، دیکدر غیرکامل Incomplete یا دیکدر فاصله محدود Bounded Distance گفته می‌شود.

به این ترتیب در روند دکدینگ حالت های زیر را خواهیم داشت.

1) اگر $\hat{\underline{v}} = \underline{c}_i = \underline{v}$ باشد، به این معنا است که اگر خطای رخ داده در کانال
به صورتی است که بردار دریافتی $\underline{r}$ درون کره مربوط به $\underline{c}_i = \underline{v}$ قرار گرفته است
دکدینگ به درستی کلمه کد ارسالی را بازیابی کرده است (خطای قابل تصحیح رخ داده است
یعنی خطای کمتر یا مساوی t)

2) اگر $\hat{\underline{v}} = \underline{c}_j \neq \underline{v}$ به این معنا است که اگر خطا رخ داده در کانال به گونه ای بوده است

که باعث شده تا بردار دریافتی $\underline{r}$ درون کره دیکدینگ کلمه کد دیگری مانند $\underline{c}'$ قرار بگیرد و دیکدر به اشتباه $\underline{c}'$ را بازیابی می‌کند که باعث ایجاد خطای دیکدینگ می‌شود (الگوی خطای قابل تشخیص ولی غیرقابل تصحیح رخ داده است)

3) $\hat{\underline{v}} = ?$ یعنی حالتی که بردار دریافتی $\underline{r}$ در کره دیکدینگ هیچ‌یک از کلمات کد $C(n,k)$ قرار نگرفته است و در نتیجه دیکدر نمی‌تواند بردار دریافتی را بازیابی کند و دچار failure می‌شود. که ایجاد خطای دیکدینگ می‌کند.

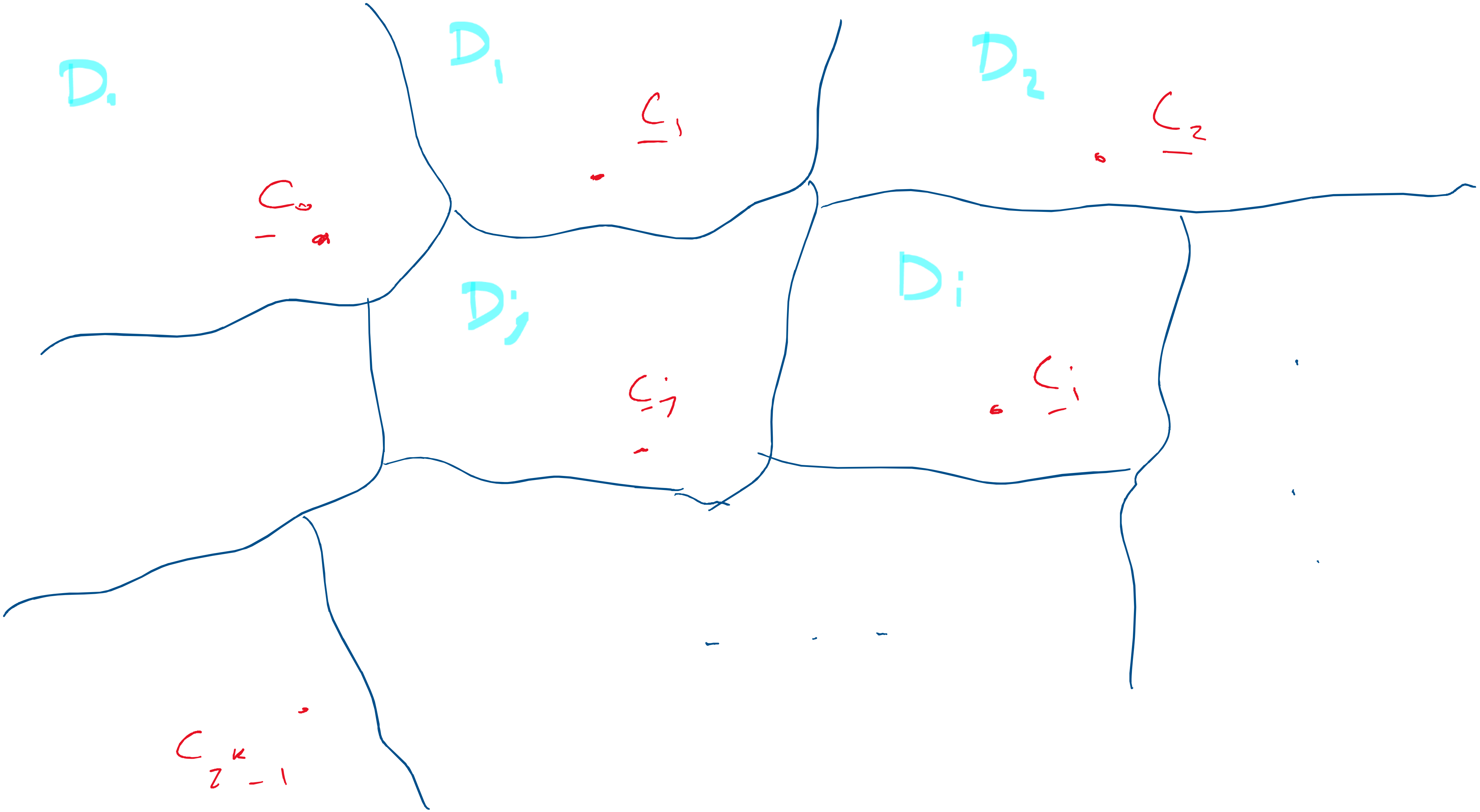
برای اینکه حالت failure را نداشته باشیم، می‌توانیم به جای کره‌های دیگری، از نواحی دیگری استفاده کنیم. به طور کلی در این روش، اساس کار بر این است که بر پایه معیارهای مانند بیشترین شباهت (کمترین فاصله همینگ) یا بیشترین احتمال بین یا کمترین احتمال خطا، فضای مربوط به کد $C(n,k)$ را به نواحی غیرهمپوشان D_0 و D_1 و ... و D_{2^k-1} تقسیم‌بندی کنیم به طوری که ناحیه‌ی D_i مربوط به کلمه‌ی کد c_i، شامل همه‌ی تمام نقاطی از فضا باشد که بیشترین شباهت را با c_i دارند.

$$D_i = \left\{ \underline{r} \in \mathbb{R}^n \;\middle|\; d_H(\underline{r}, \underline{c}_i) \leq d_H(\underline{r}, c_j) \right\} \quad \forall\, \underline{c}_j \neq c_i$$

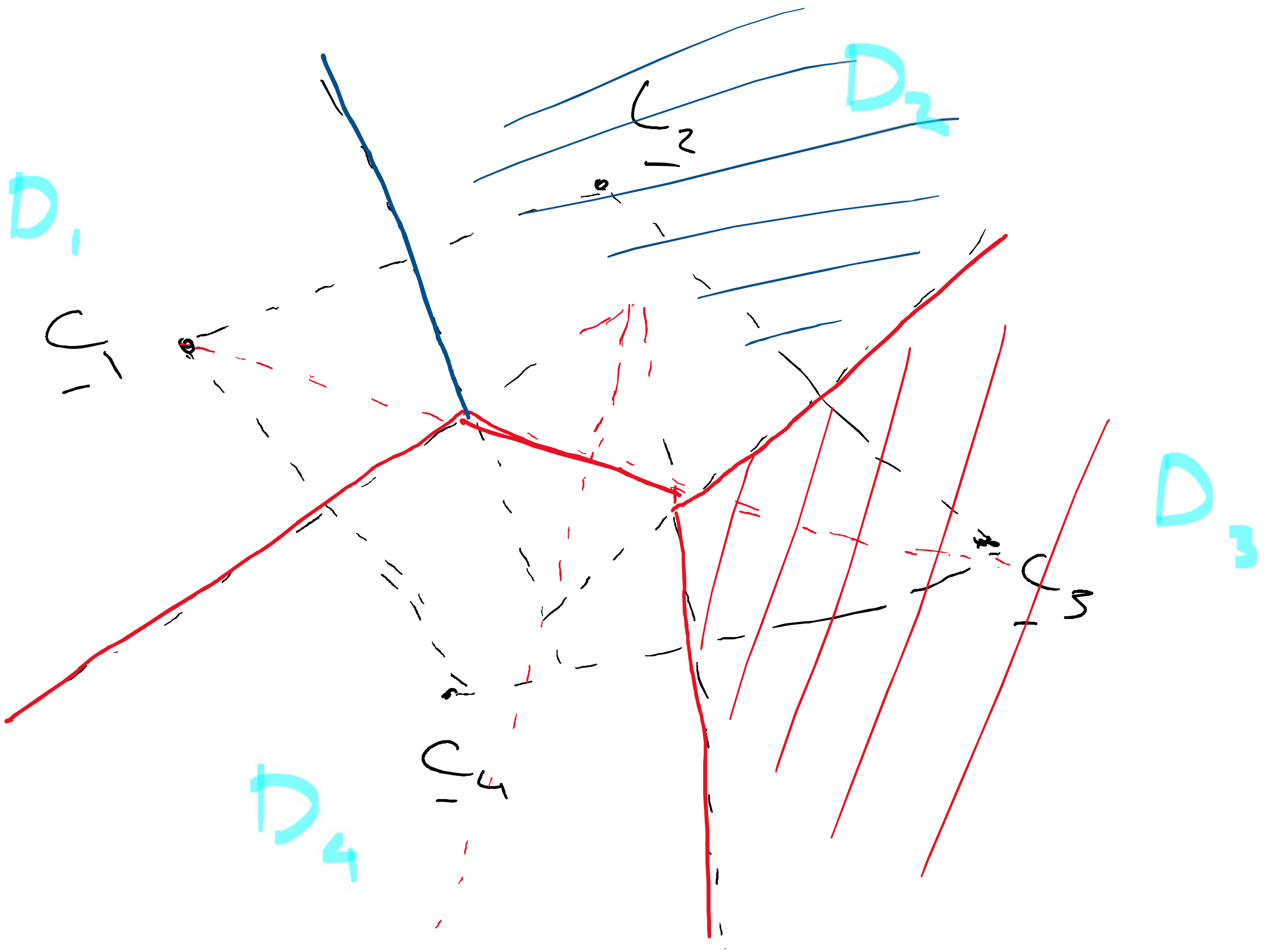
بیشترین شباهت یا کمترین فاصله همینگ

اگر بردار دریافتی $\underline{r}$ جز ناحیه‌ی D_i باشد دیکدر $\underline{c}_i$ را به عنوان کلمه کد ارسالی بازیابی می‌کند و failure نخواهیم داشت.

این نوع دیکدر را دیکدر کامل می‌گوییم.



D_0
D_1
D_2
D_j
D_i
c_0
c_1
c_2
c_j
c_i
c_{2^k-1}



D1
D2
D3
D4
c1
c2
c3
c4

* به این ترتیب می‌توانیم مسئله دیکدینگ را معادل مسئله انتخاب 2^k بردار

از بین 2^n بردار ممکن در فضای n بعدی در نظر بگیریم، به طوری که d_{min}

بیشترین مقدار ممکن باشد (با در نظر گرفتن کمترین احتمال در بیشترین نرخ ممکن)

* در مورد کدهای بلوکی خطی، نشان داده می‌شود که $C(n,k)$ یک زیرفضای k بعدی

روی فضای n بعدی (بردارهای به طول n) است.

سوالی که مطرح است، این است آیا حدی برای d_{min} یا به طور معادل برای t قابل بیان هست یا خیر؟

این حد را با ابزاری به عنوان کران بسته بندی کروی، بیان می‌شود.

تعداد بردارهای n تایی که درون کره‌ی دلخواهی قرار دارند

$$= \overbrace{\binom{n}{0}}^{\text{خود کلمه } c_i} + \underbrace{\binom{n}{1}}_{\text{تعداد بردارهایی که فاصله همینگ آنها با } c_i \text{ برابر 1 است}} + \underbrace{\binom{n}{2}}_{\text{که فاصله همینگ برابر 2}} + \cdots + \underbrace{\binom{n}{t}}_{\text{با فاصله همینگ برابر } t}$$

$$\Longrightarrow \underbrace{2^k}_{\text{تعداد کلمات کد}} \left[\binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{t}\right] \leq \underbrace{2^n}_{\text{کل اعضای فضای } n \text{ بعدی}}$$

$$\Longrightarrow \binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{t} \leq 2^{n-k}$$

Sphere packing Bound

کران بسته کروی

(کران برای t)

* اگر برای یک کد $C(n,k)$ کران بسته بندی کروی در حالت تساوی برقرار باشد،
به معنی این است که تمامی اعضای فضای n بعدی در کره های دیکدینگ قرار دارند و
در نتیجه دیکدر دچار failure نمی شود. به عبارت دیگر بین کره های دیکدینگ و
نواحی دیکدینگ در این کدها، هیچ نقطه ای بی برچسب نیست و تفاوتی ندارد که دیکدینگ
را براساس کره های دیکدینگ انجام دهیم یا براساس نواحی دیکدینگ. به این گونه کدها
کدهای کامل گفته می شود.
Complete Codes

در ادامه می خواهیم روند تشکیل گروه ها در الگوی دکدینگ را برای یک کد $C(n,k)$ با فاصله همینگ حداقل d_{min} پیدا کنیم.

$\underline{v}_0 = \underline{0}$	$\underline{v}_1$	$\underline{v}_2$	$\cdots$	$\underline{v}_{2^k-1}$
$\underline{e}_1$	$\underline{v}_1 + \underline{e}_1$	$\underline{v}_2 + \underline{e}_1$		$\underline{v}_{2^k-1} + \underline{e}_1$
$\underline{e}_2$	$\underline{v}_1 + \underline{e}_2$	$\underline{v}_2 + \underline{e}_2$		$\underline{v}_{2^k-1} + \underline{e}_2$
$\vdots$				
$\underline{e}_{2^{n-k}-1}$	$\underline{v}_1 + \underline{e}_{2^{n-k}-1}$	$\cdots$		$\underline{v}_{2^{k-1}} + \underline{e}_{2^{n-k}-1}$