Kapitola 11

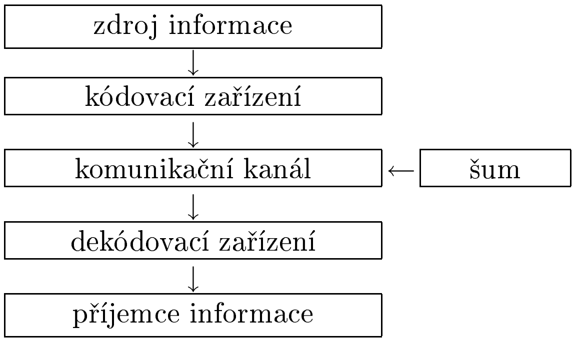
Samoopravné kódy

Teorie kódování se zabývá tím, jak rychle a spolehlivě přenášet informace z jednoho místa na druhé. Mezi její aplikace patří například minimalizace šumu při přehrávání kompaktních disků, přenos finančních informací po telefonních linkách, přenos informací mezi dvěma počítači, mezi pevným diskem a operační pamětí v jednom počítači, přenos informací z telekomunikačních, meteorologických nebo špionážních satelitů, přenos obrázků vzdálených planet a jejich měsíců ze sondy Voyager zpátky na Zem, atd.

Fyzikální prostření, ve kterém je informace přenášena, se nazývá *kanál*. Příkladem kanálů mohou být atmosféra, telefonní linky, elektromagnetické pole, atd. Při přenosu informace kanálem může dojít k jejímu poškození kvůli poruchám fyzikálního prostředí, ve kterém k přenosu dochází. Obecně jim říkáme *šum*. Šum může být způsobený například slunečními skvrnami, bleskem, přehnutím magnetické pásky, meteorickým rojem, přeslechem u telefonních linek, překlepy při psaní, špatnou artikulací, nedoslýchavostí, poškrábáním kompaktního disku, atd., atd. Šum může způsobit, že přijatá zpráva se liší od zprávy vyslané.

Teorie kódování se zabývá problémem jak odhalit a opravit chyby způsobené šumem v komunikačním kanálu. Schematicky můžeme znázornit systém pro přenos informací pomocí obrázku na následující straně. Z našeho pohledu je nejdůležitější částí diagramu šum. Bez něho by teorie kódování nebyla třeba.

Šum můžeme zmenšit pomocí volby vhodného komunikačního kanálu a použitím filtrů, které dokáží šum minimalizovat. Poté, co jsme už z dostupných možností vybrali vhodný komunikační kanál, můžeme obrátit pozornost na konstrukci kódovacího a dekódovacího zařízení.



Konstrukce kódovacího a dekódovacího zařízení sleduje několik cílů:

1. rychlé kódování informace,
2. snadný přenos zakódované zprávy,
3. rychlé dekódování přijaté zprávy,
4. opravu chyb způsobených šumem v kanálu během přenosu zprávy,
5. maximalizaci množství informace přenesené za jednotku času.

Hlavní je čtvrtý z těchto cílů. Problém spočívá v tom, že dosažení čtvrtého cíle není v souladu s pátým cílem, a nemusí být ani příliš v souladu s prvními třemi cíly. Jakékoliv řešení tohoto problému je nutně kompromisem mezi těmito pěti cíly.

Nejběžnější způsob kódování informace je mluvená řeč. Každý přirozený jazyk má v sobě zabudovánu možnost opravy chyb způsobených šumem, kterým může být v tomto případě sbíječka za oknem, špatná artikulace, nedoslýchavost, překlepy, apod. Tato ochrana před šumem je tak přirozená, že si ji ani neuvědomujeme. Odolnost přirozeného jazyka vůči šumu je založena na *redundanci*, nadbytečnosti používaných hlásek pro přenos dané informace. Většině českých vět lze rozumět i když vynecháme všechny samohlásky: zkuste si sami domyslet, co jsem dělal v sobotu večer: “V sbt včr jsm šl s kmrdm n pv”. Obvykle se odhaduje, že redundance přirozeného jazyka je více než 50%. To znamená, že stejné množství informace je možné sdělit s méně než polovinou hlásek.

Nyní si řekneme několik základních předpokladů. Informace budeme přenášet pomocí posloupností složených z čísel 0 a 1. Budeme jim říkat *cifry*. *Slovo* je posloupnost cifer. *Délka slova* je počet cifer ve slově. Tak například 010010101 je slovo délky 9. Slovo přenášíme tak, že vysíláme cifry jednu po druhé *binárním kanálem*. Binární znamená, že používáme pouze dvě cifry

– 0 a 1. Cifry lze přenášet mechanicky, elektricky, magneticky nebo nějak jinak pomocí dvou snadno rozlišitelných impulsů.

Binární kód je nějaká neprázdná množina $\mathcal{C}$ slov. Tak například kód tvořený všemi slovy délky 2 je množina

$$\mathcal{C} = \{00, 01, 10, 11\}.$$

Blokový kód je kód, který obsahuje slova stejné délky. Délka těchto slov se nazývá *délka kódu*. Budeme se zabývat pouze blokovými kódy. Kód pro nás bude vždy *binární blokový kód*. Slova, která patří do daného kódu $\mathcal{C}$, budeme nazývat *kódová slova*. Počet kódových slov v kódu $\mathcal{C}$ budeme označovat $|\mathcal{C}|$.

Další předpoklady se týkají přenosového kanálu. První předpoklad je, že kódové slovo délky n je přijato jako slovo délky n. Žádné cifry se během přenosu neztratí ani nepřibydou. Cifry se mohou během přenosu pouze změnit.

Druhý předpoklad je, že umíme vždy poznat první cifru přenášených slov. To znamená, že pokud používáme kód délky 3 a přijmeme posloupnost 011011101, tak jsme přijali slova 011, 011 a 101.

Třetí a poslední předpoklad je, že šum je *náhodný (random)*. To znamená, že pravděpodobnost, že se nějaká cifra během přenosu změní (poškodí), je stejná pro všechny cifry a není nijak ovlivněná chybami při přenosu sousedních cifer. Opakem náhodných chyb jsou *shlukové chyby (burst errors)*. Shlukové chyby jsou způsobené například škrábnutím kompaktního disku, blýskáním, slunečními erupcemi, apod. Předpoklad náhodnosti chyb není příliš realistický. Teorie kódů odstraňujících náhodné chyby je ale jednodušší. Kódy pro odstraňování shlukových chyb jsou založené na kódech původně vytvořených pro odstraňování náhodných chyb.

V naprosto spolehlivém kanálu, kde není žádný šum, je každá přijatá cifra stejná jako vyslaná cifra. V takových kanálech není třeba informaci kódovat.

Reálně existující kanály vždy nějaký šum obsahují. Binární kanál se nazývá *symetrický*, pokud pravděpodobnost, že vyslaná cifra je správně přijata, nezávisí na této cifře. *Spolehlivost* binárního symetrického kanálu je reálné číslo $p \in [0, 1]$, kde p je pravděpodobnost, že vyslaná cifra je stejná jako cifra přijatá. Číslo $1 - p$ je potom pravděpodobnost, že přijatá cifra je různá od cifry vyslané. Hodnota $p = 1$ znamená, že kanál je absolutně spolehlivý. Hodnota $p = 1/2$ znamená, že přijmeme-li nějakou cifru, nemůžeme vůbec nic říct o tom, jaká cifra byla vyslána. Obě jsou stejně pravděpodobné. Takový kanál je naopak zcela nepoužitelný. U použitelných kanálů vždy je možné předpokládat, že $p \in (1/2, 1)$. Přesnou hodnotu spolehlivosti kanálu je obvykle obtížné určit. Teorie kódování je ale na přesné znalosti spolehlivosti kanálu nezávislá.

Jaké možnosti máme pro *odhalení (detekci)* chyb během přenosu, a jaké pro *opravu (korekci)* těchto chyb? Pokud přijmeme slovo, které není kódovým slovem, tak poznáme, že během přenosu došlo k chybě, přijaté slovo je jiné, než které bylo vysláno. Chybu jsme odhalili. Pokud přijmeme kódové slovo, tak nemůžeme žádnou chybu odhalit, během přenosu nemuselo k žádné chybě dojít.

Pojem opravy chyby je o něco složitější. Tady se odvoláme na schopnost přirozeného jazyka opravovat chyby. Vidíme-li někde napsáno slovo "hospůrka", tak si domyslíme, že to asi mělo být "hospůdka". Dáme přednost slovu "hospůdka" před slovem "rozbuška", protože "hospůdka" se od přijatého slova "hospůrka" liší pouze v jednom písmenu, zatímco "rozbuška" se liší ve čtyřech písmenech. Pokud můžeme předpokládat, že je pravděpodobnější, že písmeno je přijato správně, než že bylo během přenosu poškozeno, tak je náš výběr oprávněný. Pokud bychom přijali slovo "hoshůzka", tak nemáme žádný důvod dát přednost slovu "hospůdka" před slovem "pochůzka". Obě se od přijatého slova "hoshůzka" liší ve dvou písmenech.

Vrátíme-li se k binárním blokovým kódům, tak můžeme usoudit, že kód

$$\mathcal{C}_1 = \{00, 01, 10, 00\}$$

nedokáže odhalit žádnou chybu, protože všechna slova délky 2 jsou kódová slova. Nemůže ani žádnou chybu opravit, protože žádné přijaté slovo není třeba měnit, aby se stalo kódovým slovem.

Upravíme kód $\mathcal{C}_1$ tak, že každé slovo délky 2 zopakujeme třikrát. Dostaneme kód

$$\mathcal{C}_2 = \{000000, 010101, 1010101, 111111\}.$$

Takovému kódu se říká *opakovací kód (repetition code)*. Pokud používáme kód $\mathcal{C}_2$ a přijmeme slovo 110101, tak odhalíme, že při přenosu došlo aspoň k jedné chybě – přijaté slovo není kódové slovo. Kódové slovo 010101 dostaneme z přijatého slova 110101 změnou jedné cifry, zatímco každé jiné kódové slovo vyžaduje změnu nejméně dvou cifer. Říkáme, že slovo 010101 je *nejbližší kódové slovo* k přijatému slovu 110101. Díky předpokladu, že spolehlivost kanálu je $p > 1/2$, lze dokázat, že vyslání kódového slova 010101 je pravděpodobnější, než vyslání kterékoliv jiného kódového slova. Ve skutečnosti, pokud je vysláno jakékoliv jiné slovo $c \in \mathcal{C}_2$ a při přenosu dojde k jedné chybě, pak je c jediné nejbližší kódové slovo k přijatému slovu. Kód $\mathcal{C}_2$ tak dokáže správně opravit jednu chybu.

Upravíme kód $\mathcal{C}_1$ ještě jiným způsobem. Ke každému slovu kódu $\mathcal{C}_1$ přidáme jednu cifru tak, aby byl počet cifer 1 v každém kódovém slově sudý. Dostaneme kód

$$\mathcal{C}_3 = \{000, 011, 101, 110\}.$$

Přidané cifře se říká *cifra pro kontrolu parity (parity-check digit)*. Pokud přijmeme slovo 010 s lichým počtem jednotek, tak odhalíme, že při přenosu došlo k chybě. Každé z kódových slov 110, 000, 011 dostaneme z přijatého slova 010 změnou jedné cifry. Nemáme žádný důvod dát jednomu z těchto tří kódových slov přednost před ostatními. Každé z nich ale bylo vysláno pravděpodobněji, než kódové slovo 101. Kód $\mathcal{C}_2$ tak dokáže jednu chybu odhalit, nedokáže ji ale opravit.

Efekt kontroly parity

Efekt přidání cifry pro kontrolu parity na spolehlivost přenosu informace si ukážeme na příkladu. Budeme předpokládat, že pro přenos informace používáme kód $\mathcal{D}_1$, který je tvořený všemi slovy délky 11. To znamená, že kód nedokáže odhalit žádnou chybu. Dále budeme předpokládat, že spolehlivost použitého kanálu je

$$p = 1 - 10^{-8},$$

a že cifry jsou vysílány rychlostí 10^7 cifer za sekundu.

Pravděpodobnost, že nějaké slovo je přeneseno s jednou chybou, je

$$11p^{10}(1-p),$$

pravděpodobnost, že je přeneseno se dvěma chybami, se rovná

$$\binom{11}{2} p^9 (1-p)^2,$$

tři chyby mají pravděpodobnost

$$\binom{11}{3} p^8 (1-p)^3,$$

a tak dále. Z těchto pravděpodobností je první řádově mnohem větší než součet všech ostatních. Můžeme tedy říct, že pravděpodobnost chybného přenosu slova je přibližně

$$11p^{10}(1-p) \approx \frac{11}{10^8}.$$

To znamená, že za jednu vteřinu se přenese chybně, aniž bychom si toho všimnuli, zhruba

$$\frac{11}{10^8} \cdot \frac{10^7}{11} \approx 0,1 \quad \text{slov.}$$

To je přibližně jedno slovo za 10 sekund, 6 slov za minutu, 360 za hodinu, atd.

Nyní změníme použitý kód $\mathcal{D}_1$ tak, že přidáme cifru kontrolující paritu. Dostaneme tak kód $\mathcal{D}_2$, který je tvořený všemi slovy délky 12, které obsahují sudý počet cifer 1. Všimněte si, že kód $\mathcal{D}_2$ má stejný počet slov, jako původní kód $\mathcal{D}_1$. Při použití tohoto kódu odhalíme každé slovo, které bylo přeneseno s jednou chybou. Chybně přenesené slovo tedy musí obsahovat aspoň dvě (a sudý počet) chybně přenesené cifry, abychom chybný přenos neodhalili. Pravděpodobnost, že slovo je přeneseno se dvěma chybami, je nyní

$$\binom{12}{2} p^{10}(1-p)^2,$$

čtyři chyby mají pravděpodobnost

$$\binom{12}{4} p^{8}(1-p)^4,$$

se šesti chybami je slovo přeneseno s pravděpodobností

$$\binom{12}{6} p^{6}(1-p)^6,$$

atd. Z těchto pravděpodobností je první řádově mnohem větší než ostatní. Pravděpodobnost, že při použití kódu $\mathcal{D}_2$ přijmeme slovo s chybami aniž bychom to odhalili, je tak přibližně

$$\binom{12}{2} p^{10}(1-p)^2 \approx \frac{66}{10^{16}}.$$

Tentokrát za jednu sekundu přeneseme chybně, aniž bychom si toho všimnuli, přibližně

$$\frac{66}{10^{16}} \cdot \frac{10^7}{12} = 5,5 \cdot 10^{-9} \quad \text{slov.}$$

To je zhruba jedno neodhalené chybně přenesené slovo za 2000 dní, téměř 6 let. Cena, kterou jsme zaplatili za takto dramatické zvýšení spolehlivosti, je relativně malá. Přenos informace se zpomalí o 1/12, protože každá dvanáctá přenášená cifra je kontrolní. Nenese žádnou novou informaci, protože ji můžeme dopočítat z jedenácti předcházejících cifer.

Rychlost přenosu informace kódem $\mathcal{C}$ délky n se definuje jako číslo

$$\frac{1}{n}\log_2|\mathcal{C}|.$$

Binárních slov délky n je 2^n, proto $1 \le |\mathcal{C}| \le 2^n$ pro každý binární kód délky n. Rychlost přenosu je tak vždy nějaké číslo mezi 0 a 1. Rovná se 1, pokud kód $\mathcal{C}$ obsahuje všechna slova délky n, a rovná se 0, pokud $\mathcal{C}$ obsahuje jediné slovo. Kód $\mathcal{D}_1$ tak přenáší informace rychlostí 1, kód $\mathcal{D}_2$ rychlostí 11/12. Tato rychlost odpovídá tomu, že z 12 cifer v každém slově kódu $\mathcal{D}_2$ pouze prvních 11 je informačních, zatímco dvanáctá cifra je kontrolní. Podobně rychlost přenosu informace kódy $\mathcal{C}_1$, $\mathcal{C}_2$ a $\mathcal{C}_3$ je postupně 1, 1/3 a 2/3. Ve všech třech případech jsou pouze první dvě cifry v každém kódovém slově informační, zbývající jsou kontrolní. Rychlost přenosu informace kódem tak říká, jaká část cifer v průměru v každém slově tohoto kódu nese informaci.

Úprava kódu $\mathcal{D}_1$ přidáním cifry pro kontrolu parity sice umožní odhalit jednu chybu při přenosu slova, nepomůže ale poznat, která cifra ve slově byla špatně přenesena. Aby to bylo možné zjistit, může být nutné vyslat chybně přijaté slovo ještě jednou. To lze zařídit například tak, že vysílač po každém slově čeká, až dostane z přijímače potvrzení, že slovo bylo přijato správně. To je velmi náročné na čas. Druhou možností je, že přijímač ukládá celou zprávu do paměti a po skončení celého vysílání požádá o opětovné vyslání těch slov, ve kterých odhalil chybu. Tento způsob je zase hodně náročný na velikost pamětí přijímače a vysílače. V některých případech je opětovné vysílání některých slov navíc buď zcela nepoužitelné, například při poslechu hudby z kompaktního disku, nebo příliš nákladné jako třeba při vysílání fotografií vzdálených planet ze sondy Voyager.

Jednoduchým způsobem jak upravit kód tak, aby byl schopen jednu chybu nejen odhalit, ale i opravit, je použít opakovací kód, vyslat každé slovo třikrát po sobě tak, jako to dělá kód $\mathcal{C}_2$. Délka kódových slov se tak prodlouží třikrát, rychlost přenosu informace se sníží na třetinu. Pokud v nějakém kódovém slově dojde k jedné chybě, tak ji opakovací kód nejen odhalí, ale i opraví. Nejméně dvě třetiny slova jsou přenesené správně a jejich porovnáním s třetinou, která se liší, poznáme chybně přijatou cifru v této třetině. Nicméně snížení rychlosti přenosu informace na jednu třetinu je příliš velké a stejné spolehlivosti lze dosáhnout lepším výběrem kódových slov než jak to dělá opakovací kód.

Hammingův kód $(\mathbf{7}, \mathbf{4}, \mathbf{3})$

První samoopravné kódy navrhnul americký matematik Richard W. Hamming krátce po druhé světové válce. Jeho návrh spočíval v kombinaci několika testů na paritu. Ukážeme si to na nejjednodušším případu Hammingových kódů, který se označuje $(7, 4, 3)$.

Tento kód má čtyři informační a tři kontrolní cifry, jeho délka je tedy 7. Kontrolní cifry jsou na prvním, druhém a čtvrtém místě kódového slova, informační cifry jsou na místech 3,5,6,7. Nejdříve si místa cifer ve slově vyjádříme ve dvojkové soustavě:

$$\begin{aligned} 1 &= 001 \\ 2 &= 010 \\ 3 &= 011 \\ 4 &= 100 \\ 5 &= 101 \\ 6 &= 110 \\ 7 &= 111. \end{aligned}$$

Jsou-li nyní dány informační cifry a_3, a_5, a_6, a_7, potřebujeme v kódovém slově

$$. \ . \ a_3 \ . \ a_5 a_6 a_7$$

dopočítat kontrolní cifry a_1, a_2 a a_4. Kontrolní cifra a_1 zajistí, aby byl v kódovém slově sudý počet cifer 1 na místech, která mají ve dvojkovém vyjádření jednotku na místě jednotek, tj. vpravo. To znamená, že

$$a_1 + a_3 + a_5 + a_7 = 0, \quad \text{neboli} \quad a_1 = a_3 + a_5 + a_7.$$

Podobně kontrolní cifra a_2 zajistí, aby byl v každém kódovém slově sudý počet cifer 1 na místech, která mají ve dvojkovém vyjádření jednotku namístě desítek, tj. uprostřed. Proto

$$a_2 + a_3 + a_6 + a_7 = 0, \quad \text{neboli} \quad a_2 = a_3 + a_6 + a_7.$$

A nakonec dopočítáme kotrolní cifru a_4 tak, aby byl v každém slově sudý počet cifer 1 na místech, která mají ve dvojkovém vyjádření jednotku na místě stovek, tj. vlevo:

$$a_4 + a_5 + a_6 + a_7 = 0, \quad \text{neboli} \quad a_4 = a_5 + a_6 + a_7.$$

Ve všech třech rovnostech je vždy jediná kontrolní cifra a zbývající tři jsou informační. Každá rovnost tak umožňuje jednoznačně vypočítat kontrolní cifru.

Každé kódové slovo $a_1a_2a_3a_4a_5a_6a_7$ Hammingova kódu $\mathcal{H}$ je tak řešením následující soustavy tří rovnic o sedmi neznámých nad dvouprvkovým tělesem $\mathbf{Z}_2$:

$$\begin{aligned} a_1 + a_3 + a_5 + a_7 &= 0 \\ a_2 + a_3 + a_6 + a_7 &= 0 \\ a_4 + a_5 + a_6 + a_7 &= 0. \end{aligned}$$

A naopak, každé řešení této soustavy je prvkem kódu $\mathcal{H}$. Hammingův kód tak můžeme definovat jako množinu všech řešení této soustavy. Matice této homogenní soustavy

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

má hodnost 3, protože je v redukovaném řádkově odstupňovaném tvaru.

A jak je to se schopností Hammingova kódu odhalit a opravit jednu chybu? Pokud přijmeme slovo $b_1b_2b_3b_4b_5b_6b_7$, dosadíme cifry $b_1, \dots, b_7$ do rovnic definujících Hammingův kód a spočítáme čísla $e_1, e_2, e_3 \in \mathbf{Z}_2$

$$\begin{aligned} b_1 + b_3 + b_5 + b_7 &= e_1 \\ b_2 + b_3 + b_6 + b_7 &= e_2 \\ b_4 + b_5 + b_6 + b_7 &= e_3. \end{aligned}$$

Pokud je $e_1 = e_2 = e_3 = 0$, je slovo $b_1b_2b_3b_4b_5b_6b_7$ kódovým slovem. Pokud je aspoň jedno z čísel e_1, e_2, e_3 různé od 0, tak při přenosu došlo k chybě. Je-li $e_1 = 1$, pak k chybě došlo na některém z míst, která mají ve vyjádření ve dvojkové soustavě 1 na místě jednotek, tj. u jedné z cifer b_1, b_3, b_5, b_7. Je-li rovněž $e_2 = 1$, pak je chyba také u jedné z cifer na místech, která mají ve dvojkové soustavě 1 na místě desítek, tj. u cifer b_2, b_3, b_6, b_7. Pokud je naopak $e_3 = 0$, tak víme, že k chybě nedošlo na žádném z míst, která mají ve dvojkové soustavě 1 na místě stovek, tj. chyba musí být u některé z cifer na místech, která mají cifru 0 na místě stovek, tj. u některé z cifer b_1, b_2, b_3. Chybně přenesená je tedy cifra b_3. Všimněte si, že místo $3 = e_3e_2e_1 = 011$ při vyjádření ve dvojkové soustavě. Změníme-li cifru b_3, bude nové slovo vyhovovat všem třem rovnicím definujícím Hammingův kód, bude tedy kódovým slovem.

Postup z předchozího odstavce vede k odhalení chybně přeneseného místa zcela obecně. Přijmeme-li slovo $b_1b_2b_3b_4b_5b_6b_7$, spočítáme cifry e_1, e_2, e_3. Pokud jsou všechny rovné 0, je $b_1b_2b_3b_4b_5b_6b_7$ kódové slovo. Pokud nejsou všechny tři cifry e_1, e_2, e_3 rovné 0, pak číslo $e_3e_2e_1$ vyjadřuje ve dvojkové soustavě jednoznačně určenou polohu cifry, jejíž změna udělá z přijatého slova $b_1b_2b_3b_4b_5b_6b_7$ kódové slovo. Hammingův kód tak nejenom odhalí, ale také opraví jednu chybu. Rychlost přenosu informace Hammingovým kódem je $4/7$. To je výrazné zrychlení oproti rychlosti $1/3$, které lze dosáhnout při použití opakovacího kódu, který také dokáže správně opravit jednu chybu.

Nyní začneme budovat souvislosti mezi lineární algebrou a teorií samoopravných kódů. Binární slova $b_1 \cdots b_n$ délky n budeme považovat za prvky aritmetického vektorového prostoru $\mathcal{Z}_2^n$ dimenze n nad dvouprvkovým tělesem $\mathbf{Z}_2 = \{0, 1\}$. Blokový binární kód délky n je potom podmnožina $\mathcal{C} \subseteq \mathcal{Z}_2^n$.

Definice 11.1 *Kód $\mathcal{C} \subseteq \mathcal{Z}_2^n$ délky n se nazývá* lineární*, jestliže $\mathcal{C}$ je podprostor prostoru $\mathcal{Z}_2^n$.* Dimenze *lineárního kódu je dimenze* $\dim \mathcal{C}$ *podprostoru $\mathcal{C}$.*

Protože počet prvků vektorového prostoru dimenze k nad dvouprvkovým tělesem $\mathbf{Z}_2$ se rovná 2^k, rychlost přenosu informace lineárním kódem dimenze k a délky n se rovná k/n.

Prakticky všechny používané kódy jsou lineární. Lineární kód je jednoduché popsat, stačí zadat nějakou jeho bázi. Bázi lineárního kódu zapisujeme do řádků matice.

Definice 11.2 *Je-li dán lineární kód $\mathcal{C} \subseteq \mathcal{Z}_2^n$ dimenze k, pak matice $\mathbf{C} = (a_{ij})$ tvaru $k \times n$ s prvky z tělesa $\mathbf{Z}_2$ se nazývá* generující matice *kódu $\mathcal{C}$, jestliže $\mathcal{C} = \mathcal{R}(\mathbf{C})$.*

V generující matici kódu jsou tedy řádky lineárně nezávislé. Matice

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

je generující maticí Hammingova kódu $(7, 4, 3)$. Sestrojili jsme ji tak, že jsme postupně zakódovali posloupnosti 1000, 0100, 0010 a 0001. Kvůli třetímu, pátému, šestému a sedmému sloupci jsou řádky matice $\mathbf{H}$ lineárně nezávislé. Dimenze Hammingova kódu $(7, 4, 3)$ je 4, neboť jsme si uvedli, že jeho prvky jsou všechna řešení soustavy tří rovnic o 7 neznámých a matice této soustavy má hodnost 3. Podle Důsledku 6.16 má podprostor všech řešení takové soustavy dimenzi rovnou 4. Vynásobíme-li generující matici $\mathbf{C}$ nějakého lineárního kódu $\mathcal{C}$ dimenze k zleva nějakou regulární maticí $\mathbf{P}$ řádu k, pak součin $\mathbf{PC}$ je také generující matice kódu $\mathcal{C}$ podle Úlohy 6.4. Každý lineární kód má tak mnoho různých generujících matic.

To, patří-li nějaký vektor $\mathbf{v} \in \mathcal{Z}_2^7$ do Hammingova kódu $\mathcal{H}$, můžeme ověřit dvěma způsoby. Protože $\mathcal{H} = \mathcal{R}(\mathbf{H})$, stačí ověřit, že $\mathbf{v} \in \mathcal{R}(\mathbf{H}) = \mathcal{L}(\mathbf{H}_{1*}, \mathbf{H}_{2*}, \mathbf{H}_{3*}, \mathbf{H}_{4*})$. Musíme tedy ověřit, že vektor $\mathbf{v}$ je lineární kombinací řádků generující matice kódu $\mathcal{H}$, tj. že $\mathbf{v} = \mathbf{uH}$ pro nějaký vektor $\mathbf{u} \in \mathcal{Z}_2^4$. Souřadnice vektoru $\mathbf{u} \in \mathcal{Z}_2^4$ najdeme jako řešení vhodné soustavy lineárních rovnic o čtyřech neznámých.

Máme-li nyní dány čtyři informační cifry a_3, a_5, a_6, a_7 a chceme-li dopočítat kontrolní cifry a_1, a_2, a_4, označíme $\mathbf{u} = (a_3, a_5, a_6, a_7) \in \mathcal{Z}_2^4$ a spočítáme součin

$$\mathbf{uH} = (b_1, b_2, b_3, b_4, b_5, b_6, b_7).$$

Z tvaru matice $\mathbf{H}$ ihned dostaneme, že $b_3 = a_3$, $b_5 = a_5$, $b_6 = a_6$ a $b_7 = a_7$. Protože $\mathbf{uH} \in \mathcal{H}$, je vektor $\mathbf{uH}$ kódovým slovem, jehož informační symboly se rovnají pořadě a_3, a_5, a_6, a_7.

To, je-li vektor $\mathbf{v} = (a_1, a_2, \ldots, a_7) \in \mathcal{H}$, můžeme ověřit také tak, že zjistíme, vyhovuje-li rovnicím definujícím Hammingův kód. Označíme-li

$$\mathbf{D} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

matici transponovanou k matici soustavy definující Hammingův kód, pak platí

$$\mathbf{v} \in \mathcal{H} \quad \text{právě když} \quad \mathbf{vD} = \mathbf{0}.$$

Druhý způsob je mnohem jednodušší. Stačí spočítat součin $\mathbf{vD}$, namísto řešení soustavy lineárních rovnic. Zkuste si spočítat, kolik operací každá z těchto dvou metod vyžaduje. Matice $\mathbf{D}$ umožňuje nejen ověřit, platí-li pro přijaté slovo $\mathbf{v} \in \mathbf{H}$, ale také najít chybně přenesenou cifru, pokud $\mathbf{vD} \neq \mathbf{0}$. V takovém případě je chybně přenesená cifra na místě $i \in \{1, 2, \ldots, 7\}$ právě když vektor $\mathbf{vD}$ se rovná i-tému řádku matice $\mathbf{D}$. Rozmyslete si proč. Všimněte si, že řádky matice $\mathbf{D}$ tvoří všechny nenulové vektory prostoru

$\mathbf{Z}_2^3$. Matice $\mathbf{D}$ se nazývá *kontrolní matice* Hammingova kódu $(7, 4, 3)$. Tyto úvahy vedou k následující definici.

Definice 11.3 *Je-li $\mathcal{C}$ lineární kód dimenze k a délky n, pak matici $\mathbf{D}$ tvaru $n \times (n-k)$ nazýváme* kontrolní matice *kódu $\mathcal{C}$, platí-li*

$$\mathbf{v} \in \mathcal{C} \quad \text{právě když} \quad \mathbf{vD} = \mathbf{0}, \quad \text{tj. platí-li} \quad \mathcal{C} = \mathcal{M}(\mathbf{D}),$$

kde $\mathcal{M}(\mathbf{D})$ je levý nulový prostor matice $\mathbf{D}$.

Úloha 11.1 *Předpokládáme, že $\mathcal{C}$ je lineární kód dimenze k a délky n a $\mathbf{D}$ je kontrolní matice kódu $\mathcal{C}$. Dokažte, že*

- *hodnost matice $\mathbf{D}$ je $n-k$, tj. sloupce kontrolní matice jsou lineárně nezávislé,*
- *je-li $\mathbf{Q}$ regulární matice řádu $n-k$, pak součin $\mathbf{DQ}$ je také kontrolní matice kódu $\mathcal{C}$.*

Řešení. Z Definice 11.3 vyplývá, že $\mathcal{C} = \mathcal{M}(\mathbf{D})$. Podle Tvrzení 6.17 pak platí

$$k = \dim \mathcal{C} = \dim \mathcal{M}(\mathbf{D}) = n - r(\mathbf{D}), \quad \text{tj.} \quad r(\mathbf{D}) = n - k.$$

Abychom dokázali druhou část úlohy, stačí si připomenout, že elementární řádkové úpravy nemění nulový prostor matice. Přechodem k transponovaným maticím dostaneme, že elementární sloupcové úpravy nemění levý nulový prostor matice. Vynásobit nějakou matici zprava regulární maticí znamená provést posloupnost elementárních sloupcových úprav. Proto

$$\mathcal{C} = \mathcal{M}(\mathbf{D}) = \mathcal{M}(\mathbf{DQ}).$$

□

Každý lineární kód má proto také mnoho různých kontrolních matic. Následující tvrzení udává nutnou a postačující podmínku pro to, aby dvě matice $\mathbf{C}$ a $\mathbf{D}$ byly generující a kontrolní maticí nějakého lineárního kódu.

Tvrzení 11.4 *Předpokládáme, že $\mathbf{C}$ je matice tvaru $k \times n$ a $\mathbf{D}$ je matice tvaru $n \times (n-k)$, obě s prvky z dvouprvkového tělesa $\mathbf{Z}_2$. Potom je ekvivalentní*

- *matice $\mathbf{C}$ a $\mathbf{D}$ jsou generující a kontrolní matice nějakého lineárního kódu $\mathcal{C}$ dimenze k a délky n,*

- *platí* $r(\mathbf{C}) = k$, *tj. řádky matice* $\mathbf{C}$ *jsou lineárně nezávislé,* $r(\mathbf{D}) = n-k$, *tj. sloupce matice* $\mathbf{D}$ *jsou lineárně nezávislé, a* $\mathbf{CD} = \mathbf{0}$.

Důkaz. Je-li $\mathbf{C}$ generující matice kódu $\mathcal{C}$ dimenze k, tvoří řádky matice $\mathbf{C}$ bázi podprostoru $\mathcal{C} \subseteq \mathbf{Z}_2^n$, který má dimenzi k. Proto jsou řádky matice $\mathbf{C}$ lineárně nezávislé a tedy $r(\mathbf{C}) = k$. Rovnost $r(\mathbf{D}) = n - k$ pro libovolnou kontrolní matici $\mathbf{D}$ lineárního kódu dimenze k a délky n jsme dokázali v Úloze 11.1. Protože $\mathbf{D}$ je kontrolní matice kódu $\mathcal{C}$ a $\mathbf{C}_{i*} \in \mathcal{R}(\mathbf{C}) = \mathcal{C}$ pro každé $i = 1, \ldots, k$, plyne odtud podle Tvrzení 3.7, že $(\mathbf{CD})_{i*} = \mathbf{C}_{i*}\mathbf{D} = \mathbf{0}$ pro každé $i = 1, \ldots, k$, tj. $\mathbf{CD} = \mathbf{0}$.

Abychom dokázali opačnou implikaci, definujeme lineární kód $\mathcal{C} = \mathcal{R}(\mathbf{C})$. Protože $r(\mathbf{C}) = k$, jsou řádky matice $\mathbf{C}$ lineárně nezávislé, tvoří proto bázi kódu $\mathcal{C}$ a matice $\mathbf{C}$ je proto generující maticí kódu $\mathcal{C}$. Vzhledem k předpokladu $\mathbf{CD} = \mathbf{0}$ platí $\mathbf{C}_{i*} \in \mathcal{M}(\mathbf{D})$ pro každé $i = 1, \ldots, k$ a proto také $\mathcal{C} = \mathcal{L}(\mathbf{C}_{1*}, \ldots, \mathbf{C}_{k*}) \subseteq \mathcal{M}(\mathbf{D})$. Protože $r(\mathbf{D}) = n - k$, platí podle Tvrzení 6.17, že $r(\mathcal{M}(\mathbf{D})) = n - r(\mathbf{D}) = k$. Každá báze podprostoru $\mathcal{C}$ je proto také bází levého nulového prostoru matice $\mathbf{D}$, a proto $\mathcal{C} = \mathcal{M}(\mathbf{D})$. Matice $\mathbf{D}$ je tak kontrolní maticí lineárního kódu $\mathcal{C}$. □

Každý lineární kód $\mathcal{C} \subseteq \mathbf{Z}_2^n$ můžeme zadat dvěma způsoby – buď pomocí generující matice nebo pomocí kontrolní matice. Tyto matice nejsou kódem $\mathcal{C}$ jednoznačně určené. Stejně tak ani jedna z těchto matic neurčuje jednoznačně tu druhou. Budeme se nyní zabývat otázkou, jak rychle najít kontrolní matici $\mathbf{D}$ kódu $\mathcal{C}$, známe-li jeho generující matici $\mathbf{C}$, a naopak jak najít generující matici $\mathbf{C}$, známe-li kontrolní matici $\mathbf{D}$.

Nejjednodušší případ nastává, je-li generující matice ve speciálním tvaru.

Definice 11.5 *Generující matice* $\mathbf{C}$ *lineárního kódu* $\mathcal{C}$ *dimenze* k *a délky* n *je ve* standardním tvaru, *platí-li*

$$\mathbf{C} = \begin{pmatrix} \mathbf{I}_k & \mathbf{X} \end{pmatrix},$$

kde $\mathbf{X}$ *je nějaká matice tvaru* $k \times (n-k)$ *s prvky z tělesa* $\mathbf{Z}_2$. *Kód* $\mathcal{C}$ *se nazývá* systematický kód, *pokud má nějakou generující matici ve standardním tvaru.*

Cvičení 11.1 *Dokažte, že je-li* $\mathbf{C} = \begin{pmatrix} \mathbf{I}_k & \mathbf{X} \end{pmatrix}$ *generující matice systematického kódu* $\mathcal{C}$ *dimenze* k *a délky* n *ve standardním tvaru a* $\mathbf{u} \in \mathbf{Z}_2^k$, *pak*

$$\mathbf{uC} = \begin{pmatrix} \mathbf{u} & \mathbf{uX} \end{pmatrix}.$$

To znamená, že v systematickém kódu dimenze k a délky n je v každém kódovém slově prvních k cifer informačních a po nich následují kontrolní cifry.

Tvrzení 11.6 *Je-li $\mathbf{C} = \begin{pmatrix} \mathbf{I}_k & \mathbf{X} \end{pmatrix}$ generující matice systematického kódu $\mathcal{C}$ dimenze k a délky n ve standardním tvaru, pak matice*

$$\mathbf{D} = \begin{pmatrix} \mathbf{X} \\ \mathbf{I}_{n-k} \end{pmatrix}$$

je kontrolní matice kódu $\mathcal{C}$.

Naopak, je-li

$$\mathbf{D} = \begin{pmatrix} \mathbf{X} \\ \mathbf{I}_{n-k} \end{pmatrix}$$

kontrolní matice lineárního kódu $\mathcal{D}$ dimenze k a délky n, pak matice $\mathbf{C} = \begin{pmatrix} \mathbf{I}_k & \mathbf{X} \end{pmatrix}$ je generující matice kódu $\mathcal{D}$.

Důkaz. Přímým výpočtem dostaneme, že

$$\mathbf{CD} = \begin{pmatrix} \mathbf{I}_k & \mathbf{X} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{I}_{n-k} \end{pmatrix} = \mathbf{I}_k\mathbf{X} + \mathbf{X}\mathbf{I}_{n-k} = \mathbf{X} + \mathbf{X} = \mathbf{0},$$

protože těleso $\mathbf{Z}_2$ má charakteristiku 2.

Vzhledem k přítomnosti bloku $\mathbf{I}_{n-k}$ v matici $\mathbf{D}$ je hodnost $r(\mathbf{D}) = n-k$. Podle Tvrzení 11.4 je tak $\mathbf{D}$ kontrolní matice kódu generovaného maticí $\mathbf{C}$, tj. kódu $\mathcal{C}$.

Tvrzení z druhého odstavce dokážeme stejně. Protože $r(\mathbf{C}) = k$, je $\mathbf{C}$ generující maticí kódu, jehož kontrolní maticí je $\mathbf{D}$, tj. kódu $\mathcal{D}$. □

Ne každý kód je ovšem systematický a má generující matici ve standardním tvaru. Z hlediska teorie kódování je ale možné omezit se pouze na systematické kódy. V libovolném lineárním kódu stačí změnit pořadí cifer v kódových slovech tak, aby informační cifry byly na počátku kódových slov a informační cifry až po nich.

Definice 11.7 *Dva lineární kódy $\mathcal{C}$ a $\mathcal{C}'$ dimenze k a délky n se nazývají* ekvivalentní, *pokud existuje generující matice $\mathbf{C}$ kódu $\mathcal{C}$ taková, že změnou pořadí sloupců v matici $\mathbf{C}$ dostaneme generující matici $\mathbf{C}'$ kódu $\mathcal{C}'$.*

Tvrzení 11.8 *Každý lineární kód $\mathcal{C}$ je ekvivalentní nějakému systematickému kódu.*

Důkaz. Vezmeme libovolnou generující matici kódu $\mathcal{C}$ a pomocí elementárních řádkových úprav ji převedeme do redukovaného řádkově odstupňovaného tvaru $\mathbf{C}$. Poté přeházíme sloupce v matici $\mathbf{C}$ tak, aby první pivot ležel v prvním sloupci, druhý pivot ve druhém sloupci, atd. Dostaneme tak matici $\mathbf{C}' = \begin{pmatrix} \mathbf{I}_k & \mathbf{X} \end{pmatrix}$, která je ve standardním tvaru a je proto generující maticí systematického kódu, který je ekvivalentní kódu $\mathcal{C}$. □

Hammingova vzdálenost

Důkazy schopnosti kódů odhalovat a opravovat některé chyby jsou založené na následující jednoduché, ale důležité definici.

Definice 11.9 *Jsou-li $\mathbf{u}, \mathbf{v} \in \mathbf{Z}_2^n$ dvě binární slova stejné délky n, pak definujeme jejich* Hammingovu vzdálenost $d(\mathbf{u}, \mathbf{v})$ *jako počet souřadnic, ve kterých se oba vektory liší.* Hammingova váha $wt(\mathbf{u})$ *slova $\mathbf{u} \in \mathbf{Z}_2^n$ se rovná počtu souřadnic ve slově $\mathbf{u}$ rovných 1. Je-li $\mathcal{C}$ kód obsahující aspoň dvě slova, pak definujeme* Hammingovu vzdálenost $d(\mathcal{C})$ *kódu $\mathcal{C}$ jako nejmenší ze vzdáleností $d(\mathbf{u}, \mathbf{v})$, kde $\mathbf{u} \neq \mathbf{v}$ jsou libovolné dva různé prvky kódu $\mathcal{C}$.*

Tak například $d(01011, 00111) = 2$ a $d(00111, 00111) = 0$. Podobně $wt(10101) = 3$ a $wt(00000) = 0$. Snadno si sami dokážete jednoduchou ale důležitou rovnost

$$d(\mathbf{u}, \mathbf{v}) = wt(\mathbf{u} + \mathbf{v})$$

pro libovolné dva vektory $\mathbf{u}, \mathbf{v} \in \mathbf{Z}_2^n$. Také všechny důkazy v následujícím cvičení jsou jednoduché.

Cvičení 11.2 *Dokažte, že pro každé tři vektory $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{Z}_2^n$ a libovolný skalár $a \in \mathbf{Z}_2$ platí*

1. $0 \leq wt(\mathbf{v}) \leq n$,
2. $wt(\mathbf{0}) = 0$,
3. *je-li* $wt(\mathbf{v}) = 0$, *pak* $\mathbf{v} = \mathbf{0}$,
4. $d(\mathbf{v}, \mathbf{v}) = 0$,
5. *je-li* $d(\mathbf{v}, \mathbf{w}) = 0$, *pak* $\mathbf{v} = \mathbf{w}$,
6. $d(\mathbf{v}, \mathbf{w}) = d(\mathbf{w}, \mathbf{v})$,
7. $wt(\mathbf{v} + \mathbf{w}) \leq wt(\mathbf{v}) + wt(\mathbf{w})$,

8. $d(\mathbf{u},\mathbf{w}) \leq d(\mathbf{u},\mathbf{v}) + d(\mathbf{v},\mathbf{w})$ (trojúhelníková nerovnost),

9. $wt(a\mathbf{v}) = a \cdot wt(\mathbf{v})$,

10. $d(a\mathbf{v}, a\mathbf{w}) = a \cdot d(\mathbf{v},\mathbf{w})$,

11. je-li $\mathcal{C}$ *lineární kód, pak* $d(\mathcal{C}) = \min\{wt(\mathbf{v}) : \mathbf{0} \neq \mathbf{v} \in \mathcal{C}\}$.

S použitím definice Hammingovy vzdálenosti můžeme formálně definovat, co to znamená, že nějaký kód odhalí nebo opraví nějakou chybu.

Definice 11.10 *Libovolné binární slovo* $\mathbf{e} \in \mathbf{Z}_2^n$ *délky* n *nazýváme* typ chyby. *Je-li* $\mathcal{C} \subseteq \mathcal{Z}_2^n$ *kód délky* n*, pak říkáme, že kód* $\mathcal{C}$ dokáže odhalit chybu *typu* $\mathbf{e} \neq \mathbf{0}$*, jestliže pro každé kódové slovo* $\mathbf{v} \in \mathcal{C}$ *platí* $\mathbf{v} + \mathbf{e} \notin \mathcal{C}$.

Říkáme, že kód $\mathcal{C}$ dokáže odhalit t *chyb, jestliže dokáže odhalit každou chybu* $\mathbf{e} \neq \mathbf{0}$ *s váhou* $wt(\mathbf{e}) \leq t$.

Dále říkáme, že $\mathcal{C}$ dokáže opravit chybu *typu* $\mathbf{e} \neq \mathbf{0}$ *jestliže pro každé slovo* $\mathbf{v} \in \mathcal{C}$ *platí*

$$d(\mathbf{v}, \mathbf{v}+\mathbf{e}) < d(\mathbf{v}+\mathbf{e}, \mathbf{w}) \quad \text{pro libovolné slovo} \quad \mathbf{w} \in \mathcal{C} \quad \text{různé od} \quad \mathbf{v}.$$

Konečně říkáme, že kód $\mathcal{C}$ dokáže opravit t *chyb, jestliže dokáže opravit každou chybu* $\mathbf{e} \neq \mathbf{0}$ *s váhou* $wt(\mathbf{e}) \leq t$.

Věta 11.11 *Má-li kód* $\mathcal{C}$ *délky* n *vzdálenost* $d(\mathcal{C}) = d$*, pak dokáže odhalit všechny chyby typu* $\mathbf{e} \neq \mathbf{0}$ *s váhou* $d-1$ *a existuje aspoň jedna chyba s váhou* d*, kterou odhalit nedokáže. Takový kód tak dokáže odhalit* $d-1$ *chyb a nedokáže odhalit* d *chyb.*

Důkaz. Je-li $0 < wt(\mathbf{e}) < d$ pro nenulový vektor $\mathbf{e} \in \mathbf{Z}_2^n$ a $\mathbf{v} \in \mathcal{C}$, pak $d(\mathbf{v}, \mathbf{v}+\mathbf{e}) = wt(\mathbf{e}) < d$. Protože $d = d(\mathcal{C})$, není vektor $\mathbf{v}+\mathbf{e}$ prvkem kódu $\mathcal{C}$.

Jsou-li $\mathbf{v}, \mathbf{w} \in \mathcal{C}$ dvě kódová slova taková, že $d(\mathbf{v},\mathbf{w}) = d = d(\mathcal{C})$, pak definujeme $\mathbf{e} = \mathbf{v}+\mathbf{w}$. Platí $wt(\mathbf{e}) = wt(\mathbf{v}+\mathbf{w}) = d$ a kód $\mathcal{C}$ neodhalí chybu typu $\mathbf{e}$, neboť $\mathbf{v}+\mathbf{e} = \mathbf{w} \in \mathcal{C}$. □

Ve formulaci následující věty se objevuje označení $\lfloor x \rfloor$ pro reálné číslo x. Symbol $\lfloor x \rfloor$ označuje *celou část* čísla x, tj. největší celé číslo menší nebo rovné x. Například $\lfloor 3/2 \rfloor = \lfloor 1 \rfloor = 1$, $\lfloor -(3/2 \rfloor = -2$, atd.

Věta 11.12 *Má-li kód* $\mathcal{C}$ *délky* n *vzdálenost* $d(\mathcal{C}) = d$*, pak dokáže opravit všechny chyby s váhou menší nebo rovnou* $\lfloor (d-1)/2 \rfloor$ *a existuje aspoň jedna chyba s váhou* $1 + \lfloor (d-1)/2 \rfloor$*, kterou kód* $\mathcal{C}$ *odhalit nedokáže. Kód* $\mathcal{C}$ *tak dokáže opravit* $\lfloor (d-1)/2 \rfloor$ *chyb a nedokáže opravit* $1 + \lfloor (d-1)/2 \rfloor$ *chyb.*

Důkaz. Předpokládáme, že $\mathbf{e} \neq \mathbf{0}$ je typ chyby s váhou $wt(\mathbf{e}) \leq (d-1)/2$. Nyní vezmeme libovolná dvě různá kódová slova $\mathbf{v}, \mathbf{w} \in \mathcal{C}$. Potom platí

$$d(\mathbf{w}, \mathbf{v}+\mathbf{e}) + d(\mathbf{v}+\mathbf{e}, \mathbf{v}) \geq d(\mathbf{w}, \mathbf{v}) \geq d.$$

Dále platí

$$\begin{aligned} d(\mathbf{w}, \mathbf{v}+\mathbf{e}) + wt(\mathbf{e}) &\geq 2wt(\mathbf{e}) + 1 \\ d(\mathbf{w}, \mathbf{v}+\mathbf{e}) &\geq wt(\mathbf{e}) + 1 \\ &\geq d(\mathbf{v}, \mathbf{v}+\mathbf{e}) + 1, \end{aligned}$$

neboť $wt(\mathbf{e}) = d(\mathbf{v}+\mathbf{e}, \mathbf{v})$ a $2wt(\mathbf{e}) + 1 \leq d$. Tím je dokázáno, že kód $\mathcal{C}$ opraví chybu typu $\mathbf{e}$.

Existují dvě kódová slova $\mathbf{v}, \mathbf{w} \in \mathcal{C}$, pro která platí $d(\mathbf{v}, \mathbf{w}) = d$. Vytvoříme typ chyby $\mathbf{e}$ tak, že v součtu $\mathbf{v} + \mathbf{w}$ změníme $d - 1 - \lfloor (d-1)/2 \rfloor$ souřadnic 1 na 0. Potom

$$\begin{aligned} d(\mathbf{v}, \mathbf{v}+\mathbf{e}) &= wt(\mathbf{e}) = 1 + \lfloor (d-1)/2 \rfloor, \quad \text{a} \\ d(\mathbf{w}, \mathbf{v}+\mathbf{e}) &= wt(\mathbf{w}+\mathbf{v}+\mathbf{e}) = d(\mathbf{v}+\mathbf{w}, \mathbf{e}) \\ &= d - (1 + \lfloor (d-1)/2 \rfloor). \end{aligned}$$

Je-li d liché číslo, tj. $d = 2t + 1$, potom

$$\begin{aligned} d(\mathbf{v}, \mathbf{v}+\mathbf{e}) &= wt(\mathbf{e}) = 1 + (2t)/2 = 1 + t, \quad \text{a} \\ d(\mathbf{w}, \mathbf{v}+\mathbf{e}) = 2t + 1 - (1+t) = t, \end{aligned}$$

takže $d(\mathbf{v}, \mathbf{v}+\mathbf{e}) > d(\mathbf{w}, \mathbf{v}+\mathbf{e})$. A je-li d sudé číslo, tj. $d = 2t$, pak

$$\begin{aligned} d(\mathbf{v}, \mathbf{v}+\mathbf{e}) &= 1 + \lfloor t - 1/2 \rfloor = t \quad \text{a} \\ d(\mathbf{w}, \mathbf{v}+\mathbf{e}) = 2t - t = t. \end{aligned}$$

V obou případech platí $d(\mathbf{v}, \mathbf{v}+\mathbf{e} \geq d(\mathbf{w}, \mathbf{v}+\mathbf{e})$ a proto $\mathbf{v}+\mathbf{e}$ není blíže ke kódovému slovu $\mathbf{v}$ než ke kódovému slovu $\mathbf{w}$. Kód $\mathcal{C}$ tak nedokáže opravit chybu typu $\mathbf{e}$.

Proto kód $\mathcal{C}$ (jehož vzdálenost je d) nedokáže opravit $1 + \lfloor (d-1)/2 \rfloor$ chyb, dokáže ale opravit $\lfloor (d-1)/2 \rfloor$ chyb. □

Vzdálenost lineárního kódu $\mathcal{C}$ můžeme velmi dobře zjistit pomocí kontrolní matice $\mathbf{D}$.

Věta 11.13 *Je-li* $\mathbf{D}$ *kontrolní matice lineárního kódu* $\mathcal{C}$*, pak platí, že vzdálenost kódu* $\mathcal{C}$ *se roná* d *právě tehdy když je libovolných* $d-1$ *různých řádků matice* $\mathbf{D}$ *lineárně nezávislých a existuje* d *řádků v matici* $\mathbf{D}$*, které jsou lineárně závislé.*

Důkaz. Je-li $\mathbf{v}$ libovolné slovo, pak $\mathbf{vD}$ je lineární kombinací přesně $wt(\mathbf{v})$ řádků matice $\mathbf{D}$. Je-li tedy $\mathbf{0} \neq \mathbf{v} \in \mathcal{C}$, pak platí $\mathbf{vD} = \mathbf{0}$ a proto existuje $wt(\mathbf{v})$ řádků matice $\mathbf{D}$, které jsou lineárně závislé.

Pokud je vzdálenost kódu $\mathcal{C}$ rovna d, pak existuje vektor $\mathbf{0} \neq \mathbf{v} \in \mathcal{C}$, který má váhu $wt(\mathbf{v}) = d$ a proto také existuje d lineárně závislých řádků matice $\mathbf{D}$. Kromě toho pro každý nenulový vektor $\mathbf{w}$ s váhou $wt(\mathbf{w}) \leq d-1$ platí $\mathbf{w} \notin \mathcal{C}$, tj. $\mathbf{wD} \neq \mathbf{0}$, tj. každých $d-1$ řádků matice $\mathbf{D}$ je lineárně nezávislých.

Opačná implikace se dokáže podobně $\square$

Nyní si řekneme, jak jsou definovány obecné Hammingovy kódy.

Definice 11.14 *Pro přirozené číslo $r \geq 2$ definujeme* Hammingův kód *délky $2^r - 1$ jako lineární kód, jehož kontrolní matice $\mathbf{D}$ je tvořena všemi nenulovými vektory délky r. Délka tohoto kódu je zřejmě $2^r - 1$.*

Věta 11.15 *Hammingův kód délky $2^r - 1$ dokáže opravit jednu chybu.*

Důkaz. Stačí ukázat, že vzdálenost Hammingova kódu je 3, tj. libovolné dva řádky kontrolní matice $\mathbf{D}$ jsou lineárně nezávislé, a dále že existují tři řádky, které lineárně závislé jsou. To je ale zřejmé, protože libovolné dva různé nenulové vektory $\mathbf{v}, \mathbf{w}$ téže délky nad dvouprvkovým tělesem jsou lineárně nezávislé. Tři vektory $\mathbf{v}, \mathbf{w}, \mathbf{v} + \mathbf{w}$ lineárně závislé jsou. $\square$