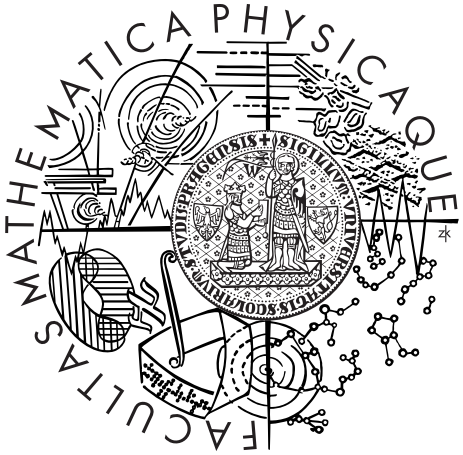
Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Jan Butora

Kapacita diskrétního kanálu

Katedra algebry

Vedoucí bakalářské práce:	doc. Mgr. Štěpán Holub, Ph.D. et Ph.D.
Studijní program:	Matematika
Studijní obor:	Matematické metody informační bezpečnosti

Praha 2015

Na tomto místě bych chtěl poděkovat vedoucímu mé bakalářské práce, doc. Mgr. Štěpánu Holubovi, Ph.D. et Ph.D., za mnoho přínosných rad, kterými přispěl k tvorbě tohoto textu. Hlavně však chci poděkovat mé rodině, která mne při studiu a tvorbě této práce vždy naplno podporovala.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V Praze dne 18. května 2015 Jan Butora

Název práce: Kapacita diskrétního kanálu

Autor: Jan Butora

Katedra: Katedra algebry

Vedoucí bakalářské práce: doc. Mgr. Štěpán Holub, Ph.D. et Ph.D., Katedra algebry

Abstrakt: V předložené práci představíme teorii kapacity diskrétního kanálu, kterou v roce 1948 publikoval C.E. Shannon a odstartoval tak éru matematické teorie informace. Nejprve si řekneme, jak vlastně můžeme informaci měřit a popíšeme komunikační systémy. Poté se zaměříme na diskrétní kanály bez šumu a dokážeme větu, která říká, jak spočítat kapacitu takových kanálů. Shannon ve svém důkazu použil několik netriviálních poznatků z rekurentních posloupností a my si ji proto dokážeme podrobně. Nakonec si na příkladě ukážeme jak tuto větu použít pro výpočet kapacity kanálu.

Klíčová slova: diferenční rovnice, generující funkce, kapacita kanálu

Title: Discrete Channel Capacity

Author: Jan Butora

Department: Department of Algebra

Supervisor: doc. Mgr. Štěpán Holub, Ph.D. et Ph.D., Department of Algebra

Abstract: This Bachelor thesis introduces and examines C.E. Shannon's discrete channel capacity theory, which was first published in 1948 as one of the founding studies in the field of mathematical information theory. In the first place, possible way of information measurement is presented and communication systems are described. Additionally, emphasis is given to discrete noiseless channel and the theorem on calculating the capacity of such channels is examined and proven. Shannon's proof is examined in detail as it contains several non-trivial results in finite differences. Finally, calculation of channel capacity using the theorem is shown in practice.

Keywords: difference equations, generating function, channel capacity

Obsah

Kapitola 1

Úvod

Od poloviny minulého století bylo vynalezeno několik způsobů zpracování a přenosu informace. Díky tomu se zvýšil zájem o teorii informace, jejíž základy jsou popsány v (Shannon, 1948), z čehož budeme v následující sekci převážně vycházet. V této práci se se zmíněnou teorií okrajově seznámíme a ukážeme si, jak ji prakticky využít v teorii kapacity diskrétního kanálu.

Základním problémem komunikace je rekonstrukce zprávy, ať už přesně nebo přibližně, která byla vybrána na jiném místě. Běžně mívají zprávy *význam*, tedy něco sdělují a jsou na sobě závislé dle nějakého systému s jistými fyzikálními či statistickými vlastnostmi. Tyto vlastnosti komunikace jsou z inženýrského hlediska irelevantní. Důležitým aspektem je, že pro aktuální zprávu se jedná o zprávu *vybranou z množiny možných zpráv*. Systém musí být navrhnut tak, aby pracoval pro každý možný výběr zprávy, nejen pro tu jednu zvolenou, jelikož ta při návrhu není známá.

Pokud jsou všechny zprávy v konečné množině stejně pravděpodobné, tak počet zpráv v této množině (nebo jakoukoli monotónní funkci tohoto počtu) můžeme při zvolení jedné zprávy chápat jako míru poskytnuté informace. V této práci budeme pro míru informace používat logaritmickou míru. Tato míra se někdy nazývá Hartleyova míra informace a dospěl k ní R. V. L. Hartley (viz Hartley, 1928).

Logaritmická míra je vhodná z několika důvodů:

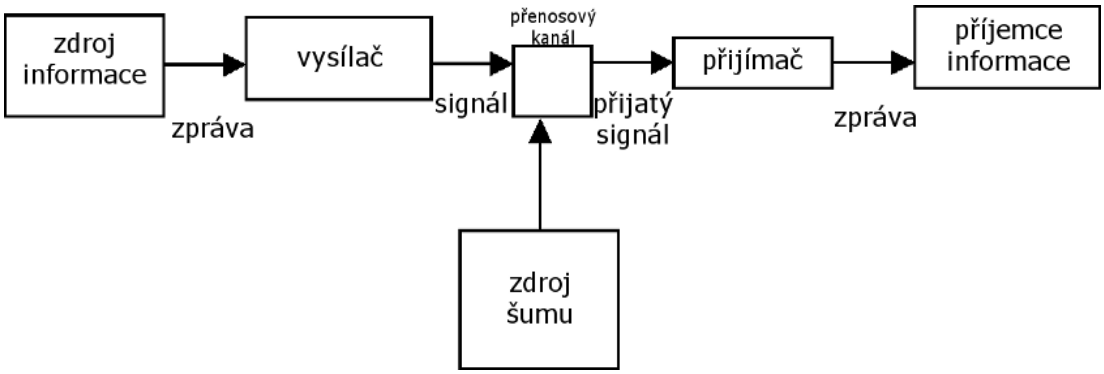
1. Je prakticky užitečná. Technické parametry jako čas, šířka pásma, počet relé, atd. se mění lineárně s logaritmem počtu možností. Například po přidání jednoho relé do skupiny se zdvojnásobí počet možných stavů. Přidává jedničku k logaritmu při základě dva tohoto počtu. Zdvojnásobením času zhruba umocníme počet možných zpráv na druhou, neboli zdvojnásobíme logaritmus, atd.

2. Z (1) lze říci, že je blízko našemu chápání míry, neboť jsme zvyklí měřit hodnoty lineárním porovnáním.

3. Je matematicky příjemnější. Mnoho limitních operací je v logaritmu počtu možností jednoduchých, avšak jinak se tyto operace stávají složitými.

Výběr základu logaritmu odpovídá výběru jednotky pro měření informace. Pokud je základ dva, pak výsledné jednotky jsou dvojkové číslice neboli *bity*[1].

[1] Anglicky - binary digits = bits



Obrázek 1.1: Schéma obecného komunikačního systému

Zařízení s dvěmi pevnými pozicemi, jako například relé, může uchovat jeden bit informace. N takových zařízení může uchovat N bitů, protože počet možných stavů je 2^N a $\log_2 2^N = N$. Pokud je zvolen základ 10, pak můžeme jednotky nazvat desítkové číslice. Jelikož

$$\log_2 M = \log_{10} M / \log_{10} 2 = 3.32 \log_{10} M,$$

je desítková číslice zhruba $3\frac{1}{3}$ bitů. V této práci budeme pracovat pouze s logaritmem o základu dva a budeme jej značit log.

Komunikačním systémem budeme rozumět typ systému znázorněného na obrázku 1.1. Je tvořen v podstatě pěti částmi:

1. *Zdroj informace* produkující zprávu nebo posloupnost zpráv, které mají být sděleny příjemci. Zpráva může být různých typů:
   (a) Posloupnost písmen jako u telegrafu či dálnopisu.
   (b) Jedna funkce času $f(t)$ jako u rádia či telefonu.
   (c) Funkce času a dalších proměnných jako u černobílé televize - zde si můžeme zprávu představit jako funkci $f(x, y, t)$ dvou souřadnic a času neboli světelnou intenzitu v bodě (x, y) a čase t na obrazovce.
   (d) Dvě či více funkcí času, řekněme $f(t), g(t), h(t)$ - toto je případ tzv. 3D zvuku.
   (e) Několik funkcí několika proměnných - v barevných televizích se zpráva skládá ze tří funkcí $f(x, y, t), g(x, y, t), h(x, y, t)$, kde každá funkce přenáší červenou, zelenou nebo modrou barvu (RGB model). Obdobně, několik černobílých televizí přenáší zprávy složené z několika funkcí tří proměnných.
   (f) Objevují se i různé kombinace, například televize se zvukem.
2. *Vysílač*, který nějakým způsobem vytvoří ze zprávy signál vhodný na přenos přes kanál. U telefonu jde pouze o změnu akustického tlaku na elektrický proud. V telegrafii máme kódovací mechanismus, který přes kanál posílá posloupnost teček, čárek a mezer odpovídající zprávě. V PCM[2] musí být různé

[2]PCM - Pulzně kódová modulace je metoda převodu analogového zvukového signálu na signál digitální.

zvukové funkce navzorkovány, zkomprimovány, kvantovány, zakódovány a nakonec řádně promíchány, aby vytvořily signál. Televizní a frekvenční modulace jsou další příklady komplexních operací, kterými lze ze zprávy získat signál.

3. *Kanál* je pouze médium užité pro přenos signálu od vysílače k přijímači. Může jít o pár vodičů, koaxiální kabel, paprsek světla, pásmo rádiových frekvencí, atd.

4. *Přijímač* běžně provádí inverzní operaci k operaci vysílače, čímž z přijatého signálu získá zprávu.

5. *Příjemce informace* je osoba (nebo věc), které byla zpráva zamýšlena.

Chtěli bychom vzít v úvahu některé obecné problémy komunikačních systémů. Abychom to dokázali, tak nejprve potřebujeme matematicky popsat různé vyskytující se prvky. Komunikační systémy můžeme zhruba rozdělit do tří kategorií: diskrétní, spojité a smíšené. Diskrétním systémem budeme mít na mysli takový systém, ve kterém jak zpráva, tak i signál je posloupnost diskrétních symbolů. Typickým příkladem je telegrafie, kde zpráva je posloupnost písmen a signál posloupnost teček, čárek a mezer. Ve spojitém systému můžeme na zprávu i signál nahlížet jako na spojitou funkci, například rádio či televize. Ve smíšeném systému se vyskytují diskrétní i spojité proměnné, například PCM přenos řeči.

V této práci se budeme zabývat pouze diskrétním případem. Ten má aplikace nejen v teorii komunikace, ale i v teorii výpočetních strojů, návrhu telefonních ústředen a dalších oblastech. Navíc tvoří diskrétní případ základ pro spojitý a smíšený typ.

Kapitola 2

Matematické předpoklady

2.1 Charakteristický polynom

V této kapitole probereme několik poznatků z rekurentních posloupností. Čerpat přitom budeme z (Kazda, 2005).

Mějme posloupnost čísel $a_0, a_1, \ldots$ zadanou rekurentní rovnicí s konstantními koeficienty

$$c_k a_{n+k} + c_{k-1} a_{n+k-1} + \ldots + c_0 a_n = 0, \tag{2.1}$$

kde c_i jsou obecně nějaká komplexní čísla (to jsou ony konstantní koeficienty), $k \in \mathbb{N}$. Co se s takovou rovnicí dá dělat? Zkusme hledat řešení ve tvaru $a_n = \lambda^n$. Má být

$$\begin{aligned} c_k a_{n+k} + c_{k-1} a_{n+k-1} + \ldots + c_0 a_n &= 0 \\ c_k \lambda^{n+k} + c_{k-1} \lambda^{n+k-1} + \ldots + c_0 \lambda^n &= 0\,. \end{aligned}$$

Pro $\lambda \neq 0$ je to ekvivalentní s rovnicí:

$$c_k \lambda^k + c_{k-1} \lambda^{k-1} + \ldots + c_0 = 0\,. \tag{2.2}$$

Polynomu na levé straně říkáme *charakteristický polynom soustavy*. Vidíme, že pro $\lambda \neq 0$ je $a_n = \lambda^n$ řešením rovnice (2.1), právě když λ je řešením (2.2).

Co když je nula řešením charakteristického polynomu? Potom lze snadno nahlédnout, že $a_0 = 1, a_1 = a_2 = \ldots = 0$ je řešením rovnice.

Pokud má charakteristický polynom pouze jednoduché kořeny, lze dokázat, že z těchto nalezených řešení (nazývají se *fundamentální systém*) lze právě jedním způsobem pomocí lineárních kombinací poskládat každé jiné řešení.

Poznámka. Rekurentní rovnice se někdy nazývají diferenční rovnice.

Příklad. Mějme rekurentní rovnici $a_{n+2} - 5a_{n+1} + 6a_n = 0$. Kořeny charakteristického polynomu jsou 3 a 2 (kořen 0 vynecháváme). To jest

$$\begin{aligned} 2^{n+2} - 5 \cdot 2^{n+1} + 6 \cdot 2^n &= 0 \\ 3^{n+2} - 5 \cdot 3^{n+1} + 6 \cdot 3^n &= 0\,. \end{aligned}$$

Pak obecné řešení vypadá jako $a_n = A \cdot 2^n + B \cdot 3^n$, kde $A, B \in \mathbb{C}$ jsou libovolná čísla. Pokud máme počáteční podmínku $a_0 = 0, a_1 = 1$, vyřešíme soustavu rovnic

$$\begin{aligned} A \cdot 2^0 + B \cdot 3^0 &= 0 \\ A \cdot 2^1 + B \cdot 3^1 &= 1 \end{aligned}$$

a dostaneme $A = -1, B = 1$.

2.2 Generující funkce

Definice 1. *Mějme posloupnost čísel $a_0, a_1, \ldots$. Pak generující funkcí této posloupnosti nazveme funkci*

$$G(x) = \sum_{n=0}^{\infty} a_n x^n \,.$$

Symbolem $[x^n]G(x)$ budeme rozumět číslo a_n.

Tuto definici zavedli (Bender a Williamson, 2006).

Příklad. Snadno nahlédneme, že $\frac{1}{1-qx} = \sum_{i=0}^{\infty} q^i x^i$, neboť

$$1 = (1 - qx)(1 + qx + q^2x^2 + q^3x^3 + \ldots) \,.$$

Můžeme tedy říci, že $\frac{1}{1-qx}$ je generující funkce posloupnosti $\{1, q, q^2, \ldots\}$.

2.3 Pomocná tvrzení

Bude se nám později v důkazu věty hodit následující lemma:

Lemma 1.

$$\frac{1}{(1-y)^k} = \sum_{n=0}^{\infty} \binom{n+k-1}{k-1} y^n \,.$$

Nejprve si ale dokážeme

Lemma 2.

$$\sum_{i=0}^{m} \binom{k+i}{i} = \binom{k+m+1}{m} \,.$$

Důkaz: Dokážeme indukcí dle m. Pro $m = 0$ tvrzení zřejmě platí. Předpokládejme, že platí pro $m > 0$ a dokažme tvrzení pro $m + 1$:

$$\begin{aligned} \sum_{i=0}^{m+1} \binom{k+i}{i} &= \sum_{i=0}^{m} \binom{k+i}{i} + \binom{k+m+1}{m+1} \\ &= \binom{k+m+1}{m} + \binom{k+m+1}{m+1} = \binom{k+m+2}{m+1} \,. \end{aligned}$$

❑

Důkaz: [Lemma 1] Opět použijeme indukci, tentokrát dle k. Z příkladu za definicí generující funkce víme, že pro $k = 1$ výraz platí. Nechť tedy platí pro nějaké $k > 1$. Dokážeme, že platí pro $k + 1$.

$$\frac{1}{(1-y)^{k+1}} = \frac{1}{1-y} \cdot \frac{1}{(1-y)^k} = \sum_{m=0}^{\infty} y^m \cdot \sum_{n=0}^{\infty} \binom{n+k-1}{k-1} y^n$$
$$= \sum_{m,n\geq 0} \binom{n+k-1}{k-1} y^{m+n} \overset{(1)}{=} \sum_{j=0}^{\infty} \left(\sum_{n=0}^{j} \binom{n+k-1}{n} \right) y^j$$
$$\overset{(2)}{=} \sum_{j=0}^{\infty} \binom{k+j}{j} y^j = \sum_{j=0}^{\infty} \binom{k+j}{k} y^j \,.$$

V rovnosti (1) jsme provedli substituci $m+n=j$, to je $m=j-n$. Suma přes n končí v j, protože $m \geq 0$ je ekvivalentní $j-n \geq 0$ tj. $n \leq j$. V rovnosti (2) jsme využili Lemma 2.

❑

Kapitola 3

Diskrétní systémy bez šumu

3.1 Diskrétní kanál bez šumu

Nyní si na základě (Shannon, 1948) ukážeme, co to vlastně kapacita diskrétního kanálu je a jak ji v některých případech spočítat.

Dálnopis a telegraf jsou dva jednoduché příklady diskrétního kanálu pro přenos informace. Obecně budeme diskrétním kanálem rozumět systém, ve kterém mohou být posílány posloupnosti prvků z konečné množiny znaků $S_1, \ldots, S_n$. Takovým posloupnostem budeme někdy pro jednoduchost říkat slova. Předpokládejme, že každý symbol S_i má jistou dobu trvání t_i jednotek času (ne nutně stejnou pro různá S_i, například tečky a čárky v telegrafii). Není nutné, abychom byli schopni poslat všechny možné posloupnosti S_i, mohou být povoleny jen některé. To budou možné signály kanálu. Tedy předpokládejme, že v telegrafii máme znaky:

1. Tečka, velikosti 2, tedy s dobou trvání 2 jednotky času.
2. Čárka, velikosti 4.
3. Mezera mezi písmeny, velikosti 3.
4. Mezera mezi slovy, velikosti 6.

Je třeba zavést i písmennou mezeru, protože dva znaky za sebou mohou reprezentovat dvě stejná písmena nebo jedno úplně jiné písmeno. Na povolených posloupnostech mějme další omezení takové, že za sebou nemáme více mezer. Nyní se budeme zabývat otázkou, jak rychle můžeme informace přes takový kanál posílat.

U dálnopisu, jehož symboly mají všechny stejnou dobu trvání a libovolná posloupnost 32 symbolů je povolena, je odpověď jednoduchá. Každý symbol představuje pět bitů informace. Pokud systém pošle n symbolů za sekundu, tak přirozeně řekneme, že přes kanál pošle $5n$ bitů za sekundu. To neznamená, že by dálnopisným kanálem informace proudily stále takovou rychlostí - je to pouze největší možná rychlost a zda-li aktuální rychlost dosáhne svého maxima závisí na zdroji informace, který kanál využívá, což Shannon dokázal, ale my se tím zabývat nebudeme.

V obecnějším případě s různými délkami symbolů a omezeními na povolené posloupnosti uvažujme následující definici.

Definice 2. *Kapacita C diskrétního kanálu je*

$$C = \lim_{T\to\infty} \frac{\log N(T)}{T}$$

kde $N(T)$ je počet povolených signálů v čase T.

Snadno nahlédneme, že u dálnopisu dostáváme již zmíněný výsledek, neboť

$$C = \lim_{T\to\infty} \frac{\log N(T)}{T} = \lim_{T\to\infty} \frac{\log 32^{n\cdot T}}{T} = \lim_{T\to\infty} \frac{T\cdot n\cdot \log 32}{T} = n\cdot \log 32 = 5n$$

bitů za sekundu, jelikož za jednu sekundu můžeme poslat n symbolů a každý symbol může nabýt všech 32 hodnot, tedy $N(T) = 32^{n\cdot T}$. Předpokládejme, že všechna slova složená ze znaků $S_1, \ldots, S_n$ jsou povolena a tyto znaky mají doby trvání $t_1, \ldots, t_n$. Jaká bude kapacita kanálu? Pokud $N(t)$ představuje počet slov s dobou trvání t, dostáváme

$$N(t) = N(t-t_1) + N(t-t_2) + \ldots + N(t-t_n).$$

Celkový počet slov je roven součtu počtu slov končících na $S_1, \ldots, S_n$ a ty jsou zřejmě $N(t-t_1), \ldots, N(t-t_n)$. Máme charakteristický polynom rekurentní soustavy

$$r^t = r^{t-t_1} + r^{t-t_2} + \ldots + r^{t-t_n}\,.$$

Což je pro $r \neq 0$ ekvivalentní polynomu

$$1 = r^{-t_1} + r^{-t_2} + \ldots + r^{-t_n},$$

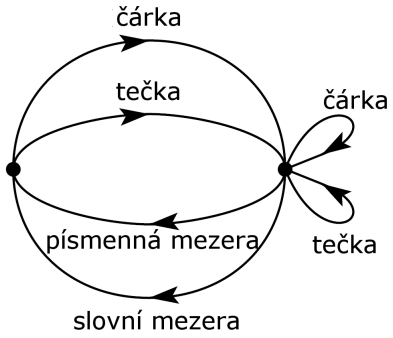
a tedy je $N(t)$ pro velká t rovno r_0^t, kde r_0 je největším reálným řešením tohoto polynomu, jak vyjde najevo později. Tudíž $C = \log r_0$.

V případě, že na povolených slovech máme nějaká omezení, můžeme získat diferenční rovnici tohoto typu a zjistit C z charakteristické rovnice. Ve výše zmíněném telegrafním případě můžeme rozdělit počet slov délky t podle posledního a předposledního znaku. Jestliže slovo končí na tečku, stačí nám vzít slova délky $t-2$ a přidat k nim tečku (protože tečka má délku 2). Podobně pro slova končící na čárku stačí vzít slova délky $t-4$. Pak je třeba rozlišit případy, kdy poslední znak je mezera, jelikož předposlední znak nemohl být mezera, tj. pokud slovo končí dvojicí tečka, písmenná mezera - počítáme se slovy délky $t-5$, pro slova končící dvojicí čárka, písmenná mezera - počítáme se slovy délky $t-7$. Obdobně pro slova končící slovní mezerou počítáme se slovy délek $t-8$ (předposlední znak byl tečka) a $t-10$ (předposlední znak byl čárka). Dostáváme tedy rekurentní rovnici:

$$N(t) = N(t-2) + N(t-4) + N(t-5) + N(t-7) + N(t-8) + N(t-10)\,.$$

Tedy $C = \log \mu_0$, kde μ_0 je největší reálný kořen charakteristické rovnice $1 = \mu^{-2} + \mu^{-4} + \mu^{-5} + \mu^{-7} + \mu^{-8} + \mu^{-10}$. Po vyřešení dostaneme $C = 0{,}539$. Obecně ovšem můžeme mít na povolených slovech taková omezení, že nebudeme schopni určit diferenční rovnici. Jak potom spočítat kapacitu kanálu?

Mějme možné stavy systému $a_1, a_2, \ldots, a_m$. V každém stavu mohou být vyslány pouze některé znaky z množiny $S_1, \ldots, S_n$ (různé znaky pro různé stavy). Při



Obrázek 3.1: Grafická reprezentace omezení na telegrafních symbolech

vyslání jednoho znaku se s ohledem na předchozí stav a na zrovna vysílaný znak stav změní. Telegrafie je toho jednoduchým příkladem. Máme dva stavy, které závisí na tom, zda poslední vyslaný znak byl či nebyl mezera. Pokud ano, pak můžeme poslat pouze tečku nebo čárku a stav se vždy změní. Pokud ne, tak můžeme poslat cokoli a stav se změní pokud pošleme mezeru, jinak zůstane stejný. Tyto konkrétní podmínky můžeme znázornit grafem v obrázku 3.1.

Zvýrazněné vrcholy reprezentují stavy a křivky indukují možné znaky ve stavu a další stav. Pokud jsme schopni podmínky na povolená slova popsat takovýmto způsobem, pak (Shannon, 1948) zavedl následující větu.

3.2 Výpočet kapacity diskrétního kanálu

Věta 3. *Nechť $b_{ij}^{(s)}$ je doba trvání s-tého symbolu, který je povolen ve stavu i a vede do stavu j. Potom kapacita kanálu C je rovna* $\log W$*, kde W je největší reálný kořen rovnice s determinantem:*

$$\left| \sum_s y^{-b_{ij}^{(s)}} - \delta_{ij} \right| = 0$$

a δ_{ij} je Kroneckerovo delta.

Například v telegrafním případě (obr. 3.1) dostáváme takovýto determinant:

$$\begin{vmatrix} -1 & (y^{-2}+y^{-4}) \\ (y^{-3}+y^{-6}) & (y^{-2}+y^{-4}-1) \end{vmatrix}$$

přičemž za stav 1 bereme z obrázku vrchol nalevo a za stav 2 vrchol napravo.

Důkaz: Nechť $N_i(L)$ je počet slov s dobou trvání L končících ve stavu i a označme množinu stavů $\{1, 2, \ldots, l\}$. Bez újmy na obecnosti můžeme předpokládat, že při začátku komunikace se systém nachází ve stavu 1. Z toho dostáváme počáteční podmínky $N_1(0) = 1, N_2(0) = \ldots N_l(0) = 0$, protože se v čase 0 dostaneme do stavu 1 při začátku komunikace. Dále můžeme dodefinovat $N_j(L) = 0$ pro $L < 0$.

Potom máme pro $1 < j \leq l$:

$$N_j(L) = \sum_{i=1}^{l} \sum_s N_i(L - b_{ij}^{(s)})$$
$$N_1(L) = \sum_{i=1}^{l} \sum_s N_i(L - b_{i1}^{(s)}) + a_1(L), \tag{3.1}$$

kde $a_1(0) = 1$, protože pro $L = 0$ je $L - b_{ij}^{(s)} < 0$, tedy $N_i(L - b_{i1}^{(s)}) = 0$ (přitom $N_1(0) = 1$) a $a_1(L) = 0$ pro $L \neq 0$. To je diferenční rovnice, označme tedy

$$\xi_j(x) = \sum_{L=0}^{\infty} N_j(L) x^L$$

generující funkci posloupnosti $\{N_j(0), N_j(1), \ldots\}$. Ze vztahu (3.1) pak plyne

$$\begin{aligned}
\xi_j(x) &= \sum_{L=0}^{\infty} N_j(L) x^L = \sum_{L=0}^{\infty} \left(\sum_{i=1}^{l} \sum_s N_i(L - b_{ij}^{(s)}) \right) x^L \\
&= \sum_{L=0}^{\infty} \sum_{i=1}^{l} \sum_s \left(N_i(L - b_{ij}^{(s)}) x^L \right) \overset{(1)}{=} \sum_{L=0}^{\infty} \sum_{i=1}^{l} \sum_s N_i(L) x^{L + b_{ij}^{(s)}} \\
&= \sum_{i=1}^{l} \sum_{L=0}^{\infty} N_i(L) x^L \sum_s x^{b_{ij}^{(s)}} = \sum_{i=1}^{l} \xi_i(x) \sum_s x^{b_{ij}^{(s)}} .
\end{aligned}$$

Vztah (1) platí, kvůli $N_j(L) = 0$ pro $L < 0$. Odsud dostáváme pro $j \neq 1$:

$$\xi_j(x) = \sum_{i=1}^{l} \xi_i(x) \sum_s x^{b_{ij}^{(s)}}$$

a stejným postupem získáme:

$$\xi_1(x) = \sum_{i=1}^{l} \xi_i(x) \sum_s x^{b_{i1}^{(s)}} + 1 .$$

Sestavíme našich l rovnic do soustavy:

$$\begin{pmatrix} \xi_1(x) \\ \xi_2(x) \\ \vdots \\ \xi_l(x) \end{pmatrix} = \begin{pmatrix} \sum_s x^{b_{11}^{(s)}} & \cdots & \sum_s x^{b_{l1}^{(s)}} \\ \vdots & \ddots & \vdots \\ \sum_s x^{b_{1l}^{(s)}} & \cdots & \sum_s x^{b_{ll}^{(s)}} \end{pmatrix} \cdot \begin{pmatrix} \xi_1(x) \\ \xi_2(x) \\ \vdots \\ \xi_l(x) \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Označme matici vyskytující se na pravé straně jako X a vektor na levé straně $\vec{\xi}$. Pak tuto soustavu můžeme napsat do rovnice

$$\vec{\xi} = X \cdot \vec{\xi} + \vec{e_1}$$

kde $\vec{e_1}$ je vektor kanonické báze l-dimenzionálního prostoru. Označíme-li I_l jednotkovou matici řádu l, pak je tato rovnice ekvivalentní s

$$\begin{aligned}
I_l \cdot \vec{\xi} &= X \cdot \vec{\xi} + \vec{e_1} \\
I_l \cdot \vec{\xi} - X \cdot \vec{\xi} &= \vec{e_1} \\
(I_l - X) \cdot \vec{\xi} &= \vec{e_1}
\end{aligned}$$

Označme matici $I_l - X$ symbolem $\hat{X}$. Pak dle Cramerova pravidla pro řešení systému lineárních rovnic dostáváme, že

$$\xi_i(x) = \frac{|\hat{X}_i|}{|\hat{X}|}$$

kde $|\hat{X}|$ značí determinant matice $\hat{X}$ a $\hat{X}_i$ je matice vzniklá z $\hat{X}$ nahrazením i-tého sloupce vektorem $\vec{e_1}$. Tím jsme získali racionální funkci

$$\xi_i(x) = \frac{P_i(x)}{Q(x)}, \tag{3.2}$$

kde P_i a Q jsou polynomy reprezentující dané determinanty. Pro Q platí $Q(0) = 1$, tedy není nulový a navíc 0 není jeho kořenem. To plyne z tvaru matice $\hat{X}$:

$$\hat{X} = \begin{pmatrix} 1 - \sum_s x^{b_{11}^{(s)}} & -\sum_s x^{b_{21}^{(s)}} & \cdots & -\sum_s x^{b_{l1}^{(s)}} \\ -\sum_s x^{b_{12}^{(s)}} & 1 - \sum_s x^{b_{22}^{(s)}} & \cdots & -\sum_s x^{b_{l2}^{(s)}} \\ \vdots & & \ddots & \vdots \\ -\sum_s x^{b_{1l}^{(s)}} & \cdots & \cdots & 1 - \sum_s x^{b_{ll}^{(s)}} \end{pmatrix}.$$

Dále zřejmě platí $deg(P_i) \leq deg(Q)$, protože P_i vznikne z matice, která má v i-tém sloupci vektor $\vec{e_1}$, tedy je nulový až na první složku, kterou můžeme chápat jako polynom stupně 0. Z (3.2) už určíme, jak se chovají členy $[x^L]\xi_i(x)$, tj. $N_i(L)$. Předpokládejme, že $deg(P_i) < deg(Q)$. Pokud by platilo $deg(P_i) = deg(Q)$, najdeme polynom R_i menšího stupně tak, že $aQ + R_i = P_i, a \in \mathbb{Q}$. Potom můžeme pracovat s funkcí $a + \frac{R_i(x)}{Q(x)}$ a postupujeme obdobně jako v následujících úvahách. Nyní můžeme použít rozklad na parciální zlomky. Označme $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{C}$ všechny různé kořeny polynomu Q (o těch už víme, že jsou nenulové) a $r_1, r_2, \ldots, r_k \in \mathbb{N}$ jejich násobnosti. Pak z rozkladu na parciální zlomky dostáváme:

$$\xi_i(x) = \sum_{j=1}^{k} \sum_{s=1}^{r_j} \frac{c_{\alpha_j,s}^{(i)}}{(x - \alpha_j)^s} = \sum_{j=1}^{k} \sum_{s=1}^{r_j} \frac{(-1)^s c_{\alpha_j,s}^{(i)}}{\alpha_j^s} \cdot \frac{1}{(1 - \frac{x}{\alpha_j})^s} \tag{3.3}$$

pro vhodná $c_{\alpha_j,s}^{(i)} \in \mathbb{C}$. Dle Lemmatu 1 platí:

$$\frac{1}{(1 - \frac{x}{\alpha_j})^s} = \sum_{n=0}^{\infty} \binom{s+n-1}{s-1} \alpha_j^{-n} x^n,$$

a tedy

$$\xi_i(x) = \sum_{n=0}^{\infty} \sum_{j=1}^{k} \sum_{s=1}^{r_j} \frac{(-1)^s c_{\alpha_j,s}^{(i)}}{\alpha_j^s} \binom{s+n-1}{s-1} \alpha_j^{-n} x^n \,.$$

Označme pro každé $j = 1, \ldots, k$:

$$S_j = \sum_{s=1}^{r_j} \frac{(-1)^s c_{\alpha_j,s}^{(i)}}{\alpha_j^s} \binom{s+n-1}{s-1} \alpha_j^{-n}$$

pak máme pro pevné i a n, že

$$[x^n]\xi_i(x) = \sum_{j=1}^{k} S_j \,.$$

Nechť α_m má mezi všemi kořeny Q nejmenší absolutní hodnotu. Označme $W = \frac{1}{\alpha_m}$ a pro $j \neq m$ označme $q_j = \frac{1}{W\alpha_j} \left(= \frac{\alpha_m}{\alpha_j}\right)$. Pak určitě platí $|q_j| < 1$. Všimněme si, že pro $j \neq m$ platí

$$\begin{aligned}\lim_{n\to\infty} \binom{s+n-1}{s-1} \cdot \left(\frac{1}{W\alpha_j}\right)^n &= \lim_{n\to\infty} \frac{(n+1)(n+2)\dots(n+s-1)}{(s-1)!} \cdot q_j^n \\ &= \lim_{n\to\infty} \underbrace{\frac{\left(1+\frac{1}{n}\right)\left(1+\frac{2}{n}\right)\dots\left(1+\frac{s-1}{n}\right)}{(s-1)!}}_{\to \frac{1}{(s-1)!}} \cdot n^{s-1} q_j^n \,.\end{aligned}$$

Ukážeme, že

$$\lim_{n\to\infty} n^{s-1} q_j^n = 0\,. \tag{3.4}$$

Položme $p_j = \frac{1}{q_j}$, tedy $|p_j| > 1$. Pak pro $\epsilon > 0, s > 1$

$$\begin{aligned} |n^{s-1}| \cdot |q_j^n| &< \epsilon \\ n^{s-1} &< |p_j|^n \epsilon \\ (s-1)\log n &< n \log |p_j| + \log \epsilon \\ \frac{\log |p_j|}{s-1} &> \frac{\log n}{n} - \frac{\log \epsilon}{n(s-1)} \end{aligned}$$

ovšem $\frac{\log \epsilon}{n(s-1)}$ i $\frac{\log n}{n}$ klesá pro velká n k 0 (např. dle l'Hospitalova pravidla). Pro $s = 1$ je $n^{s-1} = 1$, a tak se v předcházející úvaze stačí omezit na $s > 1$. Takže z definice (komplexní) limity (3.4) platí. Odtud plyne

$$\lim_{n\to\infty} \binom{s+n-1}{s-1} \cdot \left(\frac{1}{W\alpha_j}\right)^n = 0$$

tedy platí také

$$\lim_{n\to\infty} \frac{S_j}{W^n} = \sum_{s=1}^{r_j} \lim_{n\to\infty} \underbrace{\frac{(-1)^s c^{(i)}_{\alpha_j,s}}{\alpha_j^s}}_{\in\mathbb{C}} \cdot \binom{s+n-1}{s-1} \cdot (W\alpha_j)^{-n} = 0\,.$$

První rovnost platí dle aritmetiky limit, neboť výraz na pravé straně je definován. Nyní se podívejme na limitní chování S_m. Připomeňme, že r_m je násobnost kořene α_m. Opět s využitím aritmetiky limit máme:

$$\begin{aligned}\lim_{n\to\infty} \frac{S_m}{n^{r_m-1}W^n} &= \sum_{s=1}^{r_m} \lim_{n\to\infty} \frac{(-1)^s c^{(i)}_{\alpha_m,s}}{\alpha_m^s} \cdot \frac{\binom{s+n-1}{s-1}}{n^{r_m-1}} \cdot \left(\frac{\alpha_m}{\alpha_m}\right)^n \\ &= \lim_{n\to\infty} \frac{(-1)^{r_m} c^{(i)}_{\alpha_m,r_m}}{\alpha_m^{r_m}} \cdot \frac{\binom{r_m+n-1}{r_m-1}}{n^{r_m-1}} = \frac{c^{(i)}_{\alpha_m,r_m}}{(-\alpha_m)^{r_m} \cdot (r_m-1)!} \in \mathbb{C}\,.\end{aligned}$$

Členy pro $s < r_m$ ze sumy vypadly, protože $\binom{s+n-1}{s-1}$ se v limitě chová jako $\frac{n^{s-1}}{(s-1)!}$ a po vydělení n^{r_m-1} tedy klesá k 0. Označme ještě poslední výraz jako A_i. Celkem tedy platí

$$\lim_{n\to\infty} \frac{[x^n]\xi_i(x)}{n^{r_m-1} \cdot W^n} = \sum_{j=1}^{k} \lim_{n\to\infty} \frac{S_j}{n^{r_m-1} \cdot W^n} = \lim_{n\to\infty} \frac{S_m}{n^{r_m-1} \cdot W^n} = A_i\,.$$

Odtud už jde vidět, že pro $n \to \infty$ platí:

$$[x^n]\xi_i(x) = A_i \cdot n^{r_m - 1} \cdot W^n \,. \tag{3.5}$$

Stojí za zmínku, že (Shannon, 1948) vzal jako známý fakt z rekurentních posloupností pro $n \to \infty$ vztah:

$$N_i(n) = B_i W^n,$$

pro nějaká B_i a nějaké W (tj. vše zatím dokázané). Čísly B_i se potom vůbec nezabýval a W pak nalezl jiným způsobem. My jsme zmíněné B_i $(= A_i \cdot n^{r_m-1})$ a W nalezli přesně.

Z (3.5) vyplývá, že $A_i \neq 0$, protože $[x^n]\xi_i(x) \in \mathbb{N}$ (je to počet slov). Výraz (3.5) skutečně platí pro všechna $i = 1, \ldots, l$, neboť pro každé i má generující funkce $\xi_i(x)$ ve vyjádření racionální funkcí (3.2) ve jmenovateli stejný polynom Q, tedy i kořen α_m, se kterým jsme pracovali, je pro všechna i stejný. Máme, že $[x^n]\xi_i(x) = N_i(n)$, zřejmě celkový počet slov v čase n získáme jako $N(n) = \sum_{i=1}^{l} N_i(n)$. Tedy pro kapacitu kanálu platí

$$\begin{aligned} C &= \lim_{n\to\infty} \frac{\log \sum_i W^n \cdot A_i \cdot n^{r_m-1}}{n} = \lim_{n\to\infty} \left(\frac{n \log W}{n} + \frac{(r_m - 1)\log n}{n} + \frac{\log \sum_i A_i}{n} \right) \\ &= \log W \,. \end{aligned}$$

Může se ovšem stát, že kapacita C není definována. Například pokud polynom $Q(x)$ je konstantní, pak nemá žádné kořeny, a tedy naše W vůbec neexistuje. Pak ale nemá smysl kapacitu počítat, neboť by to znamenalo, že z nějakého stavu systému nemůžeme poslat žádný znak. Mějme tedy, že kapacita kanálu je počet bitů, které jsme schopni poslat za jednu sekundu. Pak můžeme předpokládat, že je kapacita reálné nezáporné číslo. Odtud plyne, že W, tedy i α_m jsou kladná reálná čísla. Pak z (3.5) platí, že A_i jsou reálná kladná čísla, tudíž si v předchozím výpočtu vystačíme s reálným logaritmem. α_m je tedy nejmenší kořen polynomu Q a navíc je reálný. W je tedy největší kladný reálný kořen funkce $R(y)$, která vznikne z $Q(x)$ substitucí $y = x^{-1}$. $R(y)$ je zřejmě determinant matice Y vzniklé z $\hat{X}$ stejnou substitucí. W je určitě i největší reálný kořen determinantu matice $-Y$ (případná změna znaménka nemění kořeny) a ten je roven determinantu matice transponované $\hat{Y} = -Y^T$.

$$det(\hat{Y}) = \begin{vmatrix} \sum_s y^{-b_{11}^{(s)}} - 1 & \sum_s y^{-b_{12}^{(s)}} & \cdots & \sum_s y^{-b_{1l}^{(s)}} \\ \sum_s y^{-b_{21}^{(s)}} & \sum_s y^{-b_{22}^{(s)}} - 1 & \cdots & \sum_s y^{-b_{2l}^{(s)}} \\ \vdots & & \ddots & \vdots \\ \sum_s y^{-b_{l1}^{(s)}} & \cdots & \cdots & \sum_s y^{-b_{ll}^{(s)}} - 1 \end{vmatrix} .$$

W je tedy skutečně největší reálný kořen rovnice

$$\left| \sum_s y^{-b_{ij}^{(s)}} - \delta_{ij} \right| = 0 \,.$$

❑

3.3 Aplikace

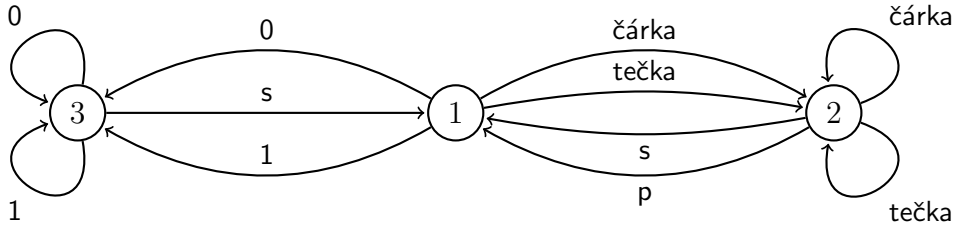
Předvedeme si pro přehlednost některé kroky předchozího důkazu na následujícím umělém příkladě. Uvažujme množinu znaků:

- Tečka, velikosti 2, tedy s dobou trvání 2 jednotky času (pro jednoduchost uvažujme sekundy).
- Čárka, velikosti 4.
- Mezera mezi písmeny, velikosti 3.
- Mezera mezi slovy, velikosti 6.
- 0 velikosti 5.
- 1 velikosti 7.

Můžeme si naši abecedu představit jako telegrafii, kde písmena tvoříme posloupnostmi teček a čárek a navíc čísla tvoříme pomocí nul a jedniček, tj. tvoříme binární zápis čísla. Mějme omezení taková, že nemůžeme poslat dvě mezery za sebou a nemůžeme posílat nuly a jedničky vedle teček či čárek. Navíc posloupnosti nul a jedniček určují právě jedno číslo a tedy nemá smysl mezi těmito symboly posílat písmennou mezeru. To nám rozdělí náš komunikační systém do tří stavů:

1. Byla poslána mezera.
2. Byla poslána tečka nebo čárka.
3. Byla poslána nula nebo jednička.

Graficky si tento systém lze představit jako na obrázku 3.2.



Obrázek 3.2: Grafická reprezentace upravené telegrafní abecedy
s - slovní mezera, p - písmenná mezera

Nechť $\xi_i(x)$ je opět generující funkce posloupnosti $\{N_i(0), N_i(1), \ldots\}$. Zafixujme $i = 1$, ostatní případy se provedou obdobně. Najdeme vyjádření generující funkce racionální funkcí

$$\xi_1(x) = \frac{P_1(x)}{Q(x)}\,.$$

$$Q(x) = \begin{vmatrix} 1 & -x^3 - x^6 & -x^6 \\ -x^2 - x^4 & 1 - x^2 - x^4 & 0 \\ -x^5 - x^7 & 0 & 1 - x^5 - x^7 \end{vmatrix}$$
$$P_1(x) = \begin{vmatrix} 1 & -x^3 - x^6 & -x^6 \\ 0 & 1 - x^2 - x^4 & 0 \\ 0 & 0 & 1 - x^5 - x^7 \end{vmatrix}.$$

Takže

$$\xi_1(x) = \frac{x^{11} + 2x^9 - x^5 - x^4 - x^2 + 1}{2x^{17} + 4x^{15} + x^{14} + x^{13} + 2x^{12} + 2x^9 - x^8 - x^7 - 2x^5 - x^4 - x^2 + 1}.$$

Mathematica nám říká, že všechny kořeny polynomu $Q(x)$ mají násobnost 1 a kořen s nejmenší absolutní hodnotou je $\alpha_m \doteq 0{,}686575$. Takže $W = \alpha_m^{-1} \doteq 1{,}4565$. Ukázali jsme, že pro $n \to \infty$ platí

$$[x^n]\xi_1(x) = A_1 \cdot n^{r_m - 1} \cdot W^n = A_1 \cdot 1{,}4565^n$$

pro nějaké kladné reálné A_1. Všimněme si, že abychom určili A_1 nemusíme hledat celý rozklad na parciální zlomky, vystačíme si s příslušným $c^{(1)}_{\alpha_m,1}$, jelikož

$$A_1 = \frac{c^{(1)}_{\alpha_m,r_m}}{(-\alpha_m)^{r_m} \cdot (r_m - 1)!} = \frac{c^{(1)}_{\alpha_m,1}}{-\alpha_m}.$$

Protože má $Q(x)$ pouze jednoduché kořeny, tak z (3.3) platí

$$\xi_1(x) = \sum_{j=1}^{17} \frac{c^{(1)}_{\alpha_j,1}}{x - \alpha_j}$$
$$(x - \alpha_m)\xi_1(x) = c^{(1)}_{\alpha_m,1} + \sum_{j \neq m} c^{(1)}_{\alpha_j,1} \cdot \frac{x - \alpha_m}{x - \alpha_j}$$
$$\lim_{x \to \alpha_m} (x - \alpha_m)\xi_1(x) = c^{(1)}_{\alpha_m,1}.$$

Opět pomocí *Mathematicy* zjistíme, že $c^{(1)}_{\alpha_m,1} \doteq -0{,}055$ a tedy

$$A_1 = \frac{c^{(1)}_{\alpha_m,1}}{-\alpha_m} = 0{,}0808$$

takže

$$[x^n]\xi_1(x) = 0{,}0808 \cdot 1{,}4565^n.$$

Stejným postupem nalezneme

$$A_2 = 0{,}1129$$
$$A_3 = 0{,}0109$$

a tedy pro $t \to \infty$ dostáváme, že počet slov v čase t je

$$N(t) = [x^t]\xi_1 + [x^t]\xi_2 + [x^t]\xi_3 = 1{,}4565^t \cdot (0{,}0808 + 0{,}1129 + 0{,}0109)$$
$$= 0{,}2046 \cdot 1{,}4565^t.$$

t	$N(t)$	$\frac{\log N(t)}{t}$
20	378	0,428
50	$3 \cdot 10^7$	0,497
100	$4{,}4 \cdot 10^{15}$	0,52
1000	$4{,}2 \cdot 10^{162}$	0,54
10000	$2{,}6 \cdot 10^{1632}$	0,542

Tabulka 3.1: Počet slov N v čase t

Toto vyjádření nám umožňuje spočítat přibližný počet slov v libovolném čase, pro ukázku jsme spočetli několik hodnot viz. tabulka 3.1.

Dále jsme ukázali, že kapacita kanálu je pak $C = \log W = 0{,}5425$ bitů za sekundu, tedy přidáním nuly a jedničky do komunikačního systému se kapacita kanálu mírně zvýšila, protože v sekci 3.1 nám v příkladu vyšla kapacita $C = 0{,}539$. Z tabulky 3.1 lze skutečně vidět, že pro rostoucí t se v třetím sloupci blížíme tomuto číslu. Ověříme, že W splňuje podmínku z Věty 3. Napíšeme si zadaný determinant a najdeme největší kořen

$$0 = \begin{vmatrix} -1 & x^{-2}+x^{-4} & x^{-5}+x^{-7} \\ x^{-3}+x^{-6} & x^{-2}+x^{-4}-1 & 0 \\ x^{-6} & 0 & x^{-5}+x^{-7}-1 \end{vmatrix}$$

$$1 = x^{-2}+x^{-4}+2x^{-5}+x^{-7}+x^{-8}-2x^{-9}-2x^{-12}-x^{-13}-x^{-14}-4x^{-15}-2x^{-17}.$$

Zjistíme, že největší reálný kořen této rovnice je 1,4565 což je opravdu W.

Kapitola 4

Závěr

V této práci jsme předvedli základy teorie informace a komunikace. Ukázali jsme, že za obecných podmínek v diskrétním kanálu roste logaritmus počtu možných signálu lineárně v čase, tedy se opět ukázalo, že logaritmická míra je vhodná. Dále jsme ukázali (a dokázali) návod jak spočítat kapacitu diskrétního kanálu bez šumu s obecným typem omezení na možná slova. Jak už jsme několikrát zmínili, Shannon důkaz zkoumané Věty lehce odbyl, což byl hlavní důvod vzniku této práce. Takový návod se nám může hodit v případech, kdy šum v kanále není vůbec nebo je zanedbatelný. V předchozím uměle vytvořeném příkladu jsme našli drobné vylepšení Shannonem zavedené telegrafní abecedy, neboť nám vyšla větší kapacita kanálu. Navíc jsme v příkladu ukázali, jak spočítat přibližný počet možných slov v libovolném čase, což se někdy také může hodit.

Literatura

Bender, E. A. a Williamson, S. G. (2006). *Foundations of Combinatorics with Applications*. Dover. ISBN 0-486-44603-4.

Hartley, R. V. L. (1928). Transmission of Information. page 535.

Kazda, S. (2005). Rekurentní rovnice. pages 4–6.

Shannon, C. E. (1948). A Mathematical Theory of Communication. **27**, 379–423, 623–656.

Seznam obrázků

Seznam tabulek