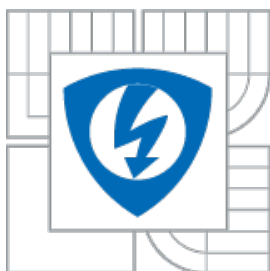




**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**  
BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH  
TECHNOLOGIÍ  
ÚSTAV TELEKOMUNIKACÍ**

**FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION  
DEPARTMENT OF TELECOMMUNICATIONS**

# **KLASIFIKACE PŘENOSOVÝCH KANÁLŮ NA ZÁKLADĚ ANALÝZY ŘEČOVÉHO SIGNÁLU**

**CLASSIFICATION OF TRANSMISSION CHANNELS BASED ON SPEECH SIGNAL ANALYSIS**

**DIPLOMOVÁ PRÁCE**  
MASTER'S THESIS

**AUTOR PRÁCE**  
AUTHOR

**BC. JOSEF BÁŇA**

**VEDOUCÍ PRÁCE**  
SUPERVISOR

**ING. HICHAM ATASSI**

BRNO 2013

## **VLOŽIT ORIGINAL ZADÁNÍ**

(při odevzdávání více kusů závěrečné práce, obsahuje originál zadání jen jedna z nich, ostatní obsahují pouze jeho okopírovanou kopii)

## **ANOTACE**

Práce zkoumá vliv charakteristických vlastností přenosového kanálu VoIP na řečové parametry. V práci je hledán způsob emulace vlastností vybraného přenosového kanálu VoIP a vytvořeno emulační síťové prostředí. Ve vývojovém prostředí Matlab je vytvořeno několik skriptů sloužících k úpravě a automatickému rozdělení souvislé nahrávky na dílčí části shodné se originálními promluvami, před průchodem přenosovým kanálem. Jsou vytvořeny databáze řečových nahrávek ovlivněných vybranými parametry - jitter, šířka pásma, ztráty. V rámci těchto databází jsou hledány příznaky, na kterých se nejvíce projevují parametry přenosového kanálu. Pomocí korelace jsou vybrány příznaky, které jsou nejvhodnější k automatizovanému určení úrovně charakteristických vlastností přenosového kanálu jako je jitter, ztráty a šířka pásma.

## **KLÍČOVÁ SLOVA**

Řečový příznak, VoIP, přenosový kanál, emulace sítě WAN, jitter, šířka pásma, ztráty.

## **ABSTRACT**

The thesis examines the impact of VoIP transmission channel characteristics on speech parameters. It seeks ways of emulating properties of a selected VoIP transmission channel and creating a network emulation environment. Several scripts have been created in the Matlab development environment and used to modify and divide a continuous recording into parts identical to the original speeches, before passing through the transmission channel. Subsequently a database of speech recordings is created, as affected by selected characteristics - jitter, bandwidth, loss. Within these databases, symptoms are sought as the most evident characteristics of the transmission channel. Using correlation, symptoms are selected that are best suited for automated determination of the properties of transmission channel characteristics such as jitter, loss and bandwidth.

## **KEYWORDS**

Speech parameter, VoIP, transmission channel, WAN network emulation, jitter, bandwidth, loss.

BÁŇA, J. *Klasifikace přenosových kanálů na základě analýzy řečového signálu*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací, 2013 37 s., 7 s. příloh. Diplomová práce. Vedoucí práce: Ing. Hicham Atassi.

## **PROHLÁŠENÍ**

Prohlašuji, že svou diplomovou práci na téma Klasifikace přenosových kanálů na základě analýzy řečového signálu jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne .....

.....

(podpis autora)

## **PODĚKOVÁNÍ**

Děkuji vedoucímu diplomové práce Ing. Hichamu Atassimu za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce. Děkuji manželce za podporu a trpělivost.

V Brně dne .....

.....

(podpis autora)

Vložit poděkování SIX

# OBSAH

<b>Seznam obrázků</b>	<b>ix</b>
<b>Seznam tabulek</b>	<b>x</b>
<b>Úvod</b>	<b>1</b>
<b>1 Přenosové kanály</b>	<b>2</b>
1.1 Pevná linka.....	2
1.1.1 Místní spojení .....	2
1.1.2 Vzdálené spojení.....	3
1.2 GSM.....	3
1.3 VoIP .....	3
<b>2 Voice over ip</b>	<b>4</b>
2.1 Spojení a přenos.....	4
2.1.1 SIP.....	5
2.2 Audio kodeky.....	5
2.2.1 G.722.....	5
2.2.2 G.722.2 (WB-AMR).....	6
2.2.3 Silk.....	6
2.2.4 Speex.....	6
2.3 Vlastnosti a parazity.....	7
<b>3 Nalezení metody emulace síťových parazit</b>	<b>8</b>
3.1 WANem.....	8
3.2 Zapojení sítě s emulátorem.....	10
3.3 Použitý software .....	11
3.3.1 VLC .....	11
3.3.2 Blink.....	11
3.3.3 GoldWave .....	12
3.3.4 Matlab .....	12
<b>4 Databáze</b>	<b>13</b>
4.1 Předúprava databází.....	14

4.1.1	Zjištění délky originálních promluv .....	14
4.1.2	Normalizace okrajů databází .....	14
4.2	Rozdělení databáze na dílčí skladby .....	15
4.2.1	Algoritmus dělení skladeb podle původní délky .....	15
4.2.2	Ruční rozdělení databáze .....	16
4.3	Jitter .....	16
4.4	Šířka pásma .....	17
4.5	Ztráty .....	17
<b>5</b>	<b>Metodika rozpoznání parametrů přenosového kanálu</b>	<b>19</b>
5.1	Rozhodovací stromy .....	19
5.2	Příznaky .....	20
5.3	Výběr na základě korelace .....	20
5.3.1	Jitter .....	23
5.3.2	Ztráty .....	27
5.3.3	Šířka pásma do 1024 kb/s .....	32
<b>6</b>	<b>Závěr</b>	<b>37</b>
	<b>Literatura</b>	<b>38</b>
	<b>Seznam symbolů, veličin a zkratk</b>	<b>40</b>
	<b>Seznam příloh</b>	<b>41</b>



## SEZNAM OBRÁZKŮ

Obr. 3.1	Grafické uživatelské rozhraní programu WANem .....	9
Obr. 3.2	Zapojení systému s emulátorem.....	10
Obr. 4.1	Rozdělené skladby s posunutím (Zobrazeno v programu GoldWave). .....	16
Obr. 4.2	Výřez slova „close“ při ztrátách 5% (nahore) a 60% (dole). .....	18
Obr. 5.1	Znázornění korelace reálných hodnot a regresně vypočtených hodnot .....	22
Obr. 5.2	Korelace pro příznak SpeFe .....	25
Obr. 5.3	Střední odchylky pro příznak SpeFe – zvyšování nOf. ....	25
Obr. 5.4	Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf. .....	26
Obr. 5.5	JT - Střední odchylka pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf .....	27
Obr. 5.6	ZT - Korelace pro příznak SpeFe – zvyšování nOf. ....	29
Obr. 5.7	ZT - Střední odchylky pro příznak SpeFe – zvyšování nOf. ....	30
Obr. 5.8	Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf. .....	31
Obr. 5.9	Střední odchylka pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf. ....	32
Obr. 5.10	SP Korelace pro příznak SpeFe – zvyšování nOf.....	34
Obr. 5.11	SP - Střední odchylky pro příznak SpeFe .....	34
Obr. 5.12	Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf. .....	35
Obr. 5.13	Střední odchylka pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf .....	36
Obr. C.6.1	Korelace příznaku SpectralFeatures pro šířku pásma do 256 kb/s.....	47
Obr. C.6.2	Střední odchylky příznaku SpectralFeatures pro šířku pásma do 256 kb/s.	48

## SEZNAM TABULEK

Tab. 5.1	JT - Korelace pro jednotlivé příznaky .....	23
Tab. 5.2	JT - Korelace pro kombinace dvou příznaků .....	23
Tab. 5.3	JT - Korelace pro vybrané kombinace příznaků .....	24
Tab. 5.4	JT- Korelace pro příznak SpeFe – zvyšování nOf .....	24
Tab. 5.5	JT- Korelace pro kombinaci příznaků SpeFe a MFCC – zvyšování nOf. ...	26
Tab. 5.6	JT - Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf. ....	26
Tab. 5.7	ZT - Korelace pro jednotlivé příznaky.....	27
Tab. 5.8	ZT - Korelace pro kombinace dvou příznaků .....	28
Tab. 5.9	ZT - Korelace pro vybrané kombinace příznaků .....	28
Tab. 5.10	ZT - Korelace pro příznak SpeFe – zvyšování nOf. ....	29
Tab. 5.11	ZT - Korelace pro kombinaci příznaků SpeFe a MFCC – zvyšování nOf. .	30
Tab. 5.12	ZT Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf. ....	31
Tab. 5.13	SP - Korelace pro jednotlivé příznaky .....	32
Tab. 5.14	SP - Korelace pro kombinace dvou příznaků .....	33
Tab. 5.15	SP - Korelace pro vybrané kombinace příznaků.....	33
Tab. 5.16	SP - Korelace pro příznak SpeFe – zvyšování nOf.....	34
Tab. 5.17	SP Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf. ....	35
Tab. C.1	Korelace SpectralFeatures pro Šířku pásma do 256 kb/s.....	47

# ÚVOD

Tato práce zkoumá vliv přenosových kanálů na řečové parametry – příznaky. Zkoumá charakteristické vlastnosti přenosových kanálů, jako jsou pevná linka, GSM a VoIP. Zabývá se hlubší analýzou a aplikací těchto poznatků na vybraný přenosový kanál.

V práci je hledán způsob emulace charakteristických vlastností vybraného přenosového kanálu a poté vytvořeno emulační síťové prostředí. Dále jsou za použití tohoto prostředí vytvořeny souvislé nahrávky. Je vytvořen program v prostředí Matlab k automatickému rozdělení těchto nahrávek na dílčí úseky odpovídající promluvám z originální databáze a vytvořeny databáze řečových nahrávek ovlivněných vybranými parametry. Pomocí těchto databází jsou hledány příznaky, na kterých se nejvíce projevují parametry přenosového kanálu. Pomocí korelace jsou vybrány příznaky, pomocí kterých, by bylo možné automatizovaně určit charakteristické vlastnosti přenosového kanálu.

# 1 PŘENOSOVÉ KANÁLY

Práce se měla zabývat prozkoumáním hypotézy, že každý přenosový kanál má specifické vlastnosti, které by ovlivňovaly řečové příznaky natolik, že by se pomocí určitých příznaků daly přenosové kanály detekovat. Základními zvolenými přenosovými kanály byly pevná linka – místní spojení, pevná linka – vzdálené spojení, VoIP a GSM síť. Dále jsou popsány základní charakteristiky dílčích přenosových kanálů.

## 1.1 Pevná linka

Pevná linka je komutační síť s přepojováním okruhů. Jedná se o nejdéle používaný přenosový kanál řečového signálu. Počíná analogovým systémem vedeným dvojlinkou, kdy hovory spojoval mechanicky člověk – spojovatel, přes vývoj několika generací ústředen, přechod z pulsní volby na tónovou, až po digitální systémy v páteřní části spojené optickými kabely, umožňující multiplexovaný přenos mnoha telefonních hovorů najednou po jediném vláknu.

Dále jsou uvedeny charakteristické vlastnosti sítě s přepojováním okruhů. Komunikace se odehrává po předem sestaveném okruhu (spojení) mezi zdrojovou a cílovou stanicí. Všechna data sledují jednu předem danou cestu sítí a jsou doručena ve správném pořadí. Komunikace v těchto sítích probíhá vždy spolehlivě, se spojením. Přidělování kanálů lze provádět multiplexováním, aby se sdílela kapacita systému v čase. Kmitočtový rozsah zvukového signálu je ořezaný na pásmo 300 – 3400 Hz.

### 1.1.1 Místní spojení

Místní spojení je nejčastěji realizováno pomocí kabelů dvojlinky, nebo strukturované kabeláže uvnitř samostatné telefonní sítě, a propojené pomocí pobočkové ústředny. V České republice je místní (podniková) síť velmi často realizována pomocí analogového systému.

### **1.1.2 Vzdálené spojení**

Vzdálené spojení pevné linky může mít obdobné charakteristiky, jako místní pevná linka, s tím, že hovor mezi účastníky sídlícími v různých městech je propojen digitální telefonní sítí patřící telefonní společnosti, a spojení může být realizováno pomocí několika telefonních ústředen propojených komunikační trasou. Komunikační trasa může být realizována pomocí různých technologií, v České republice jsou to nejčastěji optické kabely.

## **1.2 GSM**

Jedná se o buňkovou síť, fungující na několika radiových frekvencích, kdy se telefony připojují do sítě prostřednictvím nejbližší buňky.

Nejvíce užívaný zvukový kodek je AMR (Adaptive Multi-Rate audio codec). Jedná se o standardizované adaptivní schéma komprese audio dat, optimalizované pro kódování řeči. Zvukový signál je filtrován na pásmo 200 Hz až 3400 Hz. Vzorkovací frekvence zvukového signálu je standardně 8 kHz (za použití 13-bitů). AMR využívá 8 základních kompresních kodeků Přenosové rychlosti 12.2, 10.2, 7.95, 7.40, 6.70, 5.90, 5.15 a 4.75 kbit/s. [1]

Od roku 2009 je v ČR v 3G sítích (UMTS) používán širokopásmový kodek WB-AMR, standardizován s ITU označením G.722.2, se zvukovým signálem filtrovaným na širší pásmo 50-7000 Hz. [4]

## **1.3 VoIP**

Tato technologie využívá standardního protokolu IP pro přenos digitalizovaného hlasu po síti ethernet, zapouzdřeného do těl UDP datagramů. Přestože počátky této technologie sahají až k síti ARPANET, v České republice i ve světě se volání přes internet začíná více používat a mohutněji rozšiřovat až v posledních deseti letech při nástupu programu Skype. O VoIP můžeme říci, že pro přenos hlasu využívá techniku přepojování paketů, jedná se tedy v obecném pohledu o nespojovou a nespolehlivou službu. Více je přenos hlasu po IP sítích popsán v samostatné kapitole č. 2.

## 2 VOICE OVER IP

Oproti původnímu tématu zkoumání vlastností přenosových kanálů jsme se rozhodli zabývat hlubším a zajímavějším zkoumáním specifických vlastností přenosu hlasu po ethernetové síti, neboli VoIP (Voice over IP).

V práci jsou dále zkoumána specifika internetové telefonie, typické přenosové vlastnosti sítě a parazity ovlivňující výsledný přenesený hlasový záznam. Dále práce zkoumá, jak tyto vlastnosti ovlivňují hlasové příznaky a zda pomocí podrobnějšího vědeckého zkoumání je možné zjistit, případně automaticky detekovat z neznámé nahrávky bližší údaje o přenosovém kanálu.

V této práci byl použit spojovací a signalizační protokol SIP a zvukový signál byl kódován/dekódován standardizovaným kodekem G.722. Proč byly použity právě tyto parametry, je uvedeno v následujících podkapitolách.

### 2.1 Spojení a přenos

Technologie VoIP využívá standardního protokolu IP pro přenos digitalizovaného hlasu po síti ethernet, zapouzdřeného do těl UDP datagramů. O VoIP můžeme říci, že pro přenos hlasu využívá techniku přepojování paketů, jedná se tedy v obecném pohledu o nespojovou a nespolehlivou službu. Proto nutnými podmínkami pro srozumitelné a spolehlivé spojení se zabývá tzv. QoS (Quality of Service), nebo-li zajištění kvality služeb. [14]

Obecně používaných je několik standardizovaných i zatím nestandardizovaných a přesto rozšířených spojovacích protokolů, kdy k nejběžnějším patří SIP, dále ustupující, ale stále užívaný H.323, otevřené protokoly IAX a IAX2 a dále lze zmínit firemní protokoly Skinny (CISCO) a HFA (Siemens). [15]

### **2.1.1 SIP**

SIP je velmi jednoduchý, textově orientovaný protokol, který využívá známých principů internetových protokolů SMTP nebo HTTP. Tento protokol pouze zajišťuje sestavení spojení, dohodnutí přenosových parametrů a signalizaci; samotný přenos zajišťuje standardní RTP protokol zapouzdřený nejčastěji do UDP datagramů. Lze však využít i TCP paketů. SIP protokol se stále rozvíjí a lze jej využít k více než jen IP telefonii, použit lze například k zjišťování dostupnosti uživatele, instant messagingu a dalším aplikacím. Vznikl jako odezva na velmi složitý protokol H.323. [2]

## **2.2 Audio kodeky**

K digitalizaci a kompresi audio signálu lze využít většinu standardizovaných audio kodeků, jako jsou G.711 (PCM – Pulsně Kódová Modulace), G.726 (ADPCM), GSM kodek, G.722, G.723.1 (TrueSpeech 6.3), AMR, WB-AMR, Speex, Silk, a nově vzniklý otevřený protokol Opus. Cílem těchto kodeků je co nejdříve zachytit vstupní signál a přitom využít co nejméně bitů digitalizované informace. Po příjmu signálu a jeho dekompresi bychom v ideálním případě neměli sluchově rozeznat původní signál od digitalizovaného.

Jak bude v další kapitole uvedeno, v realizaci virtuální sítě byl nakonec použit audio kodek G.722, i když k zřejmě nejrozšířenějším audio kodekům ve VoIP telefonii je kodek Silk, který používá program Skype. G.722 byl použit kvůli implementaci v softwaru, jenž umožňoval realizaci zapojení emulované sítě.

Blíže si uvedeme vlastnosti některých nejpoužívanějších audio kodeků.

### **2.2.1 G.722**

Jedná se o standardizovaný, dříve proprietární širokopásmový kodek, původně používaný v telefonních linkách. Lze jej přenášet přes B kanál ISDN 64. Existuje poměrně dlouho, od roku 1988. Po 20 letech tomuto kodeku vypršela licence a je volně k použití. Šířka pásma kódovaného zvuku u kodeku G.722 je 50 – 7000 Hz oproti staršímu G.711 s šířkou zvukového pásma 300 – 3400 Hz a stejnými nároky na šířku

přenosového pásma. Vzorkovací frekvence u G.722 je 16 kHz (za použití 14 bitů) Přenosové rychlosti jsou 48, 56 a 64 kbit/s. [3]

### **2.2.2 G.722.2 (WB-AMR)**

Nejedná se o vylepšenou verzi G.722, jak by se mohlo zdát. WB-AMR je samostatný kodek, vyvinutý společnostmi Nokia a VoiceAge, jenž nese standardizované ITU označení G.722.2. Postaven je na adaptivní změně kódování podobné jako ACELP. Z principu vychází název WB-AMR (Adaptive Multi-Rate Wideband). Nejvíce se využívá zřejmě v mobilních sítích a to především mobilních sítích třetí generace. Tento kodek používá vzorkovací frekvenci 16 kHz (za použití 14 bitů) a přenosové rychlosti 6,6 , 8,85, 12,65 kb/s (poskytující stejnou kvalitu jako přenosová rychlost 56kb/s u G.722), 14.25, 15.85, 18.25, 19.85, 23.05, 23.85 kbit/s (poskytující kvalitu jako za použití rychlosti 64 kbit/s u G.722) a 64 kbit/s. Tento kodek tedy dosahuje stejné kvality jako G.722 dosahuje ale 2,5 násobné efektivity. [4]

### **2.2.3 Silk**

Nový kodek s dynamicky se měnící (dle rychlosti připojení, odolnosti ztráty paketů) vzorkovací frekvencí 8, 12, 16 nebo 24 kHz a přenosovou rychlostí od 6 do 40 kbit/s. Tento kodek začal být veřejně používán od roku 2009 v programu Skype. Nahradil tak v tomto programu dříve používaný SVOPC. V současnosti začíná být užíván také chatovacích programech integrovaných do on-line počítačových her. Dále se na českém trhu objevují telefonní aparáty pro VoIP podporující tento kodek. V porovnání s výše uvedenými kodeky může Silk dosáhnout lepší komprese při srovnatelné kvalitě přenášeného zvuku, a zároveň podporuje adaptivní úpravu kvality kódovaného signálu čímž může snížit síťovou zátěž při zahlcení přenosového kanálu. Je plánováno otevření licence pro volné komerční využití. [5]

### **2.2.4 Speex**

Jedná se o komunitně vyvinutý open-source audio kodek, vytvořený pro přenos hlasu, nevázaný na patenty, šířený a použitelný pod revidovanou BSD licencí.

Užívá 3 vzorkovací frekvence pro digitalizaci řeči a to 8 kHz pro úzkopásmové



přenosové kanály, 16 kHz pro širokopásmové přenosové kanály a 32 kHz pro ultra-širokopásmové přenosové kanály v jednom datovém proudu. Přenosové rychlosti jsou rozděleny v deseti stupních od 2 do 44 kb/s. [6]

## **2.3 Vlastnosti a parazity**

Základními parametry ovlivňujícími přenos audio signálu v IP síti za použití SIP protokolu jsou šířka pásma, ztráty, zpoždění UDP datagramů, jitter (kolísání zpoždění), poškození paketů, změna pořadí odeslaných a doručených datagramů.

## **3 NALEZENÍ METODY EMULACE SÍŤOVÝCH PARAZIT**

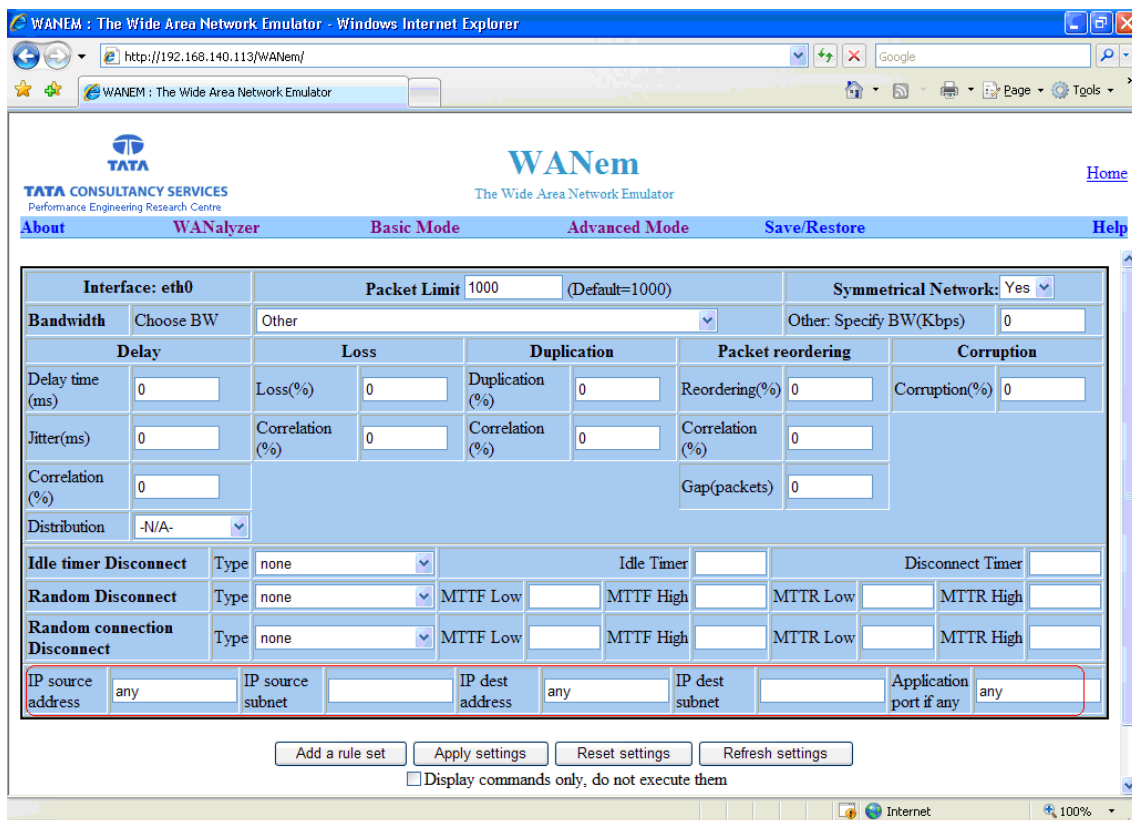
Většina programů se zabývá simulací rozsáhlých sítí, ale není pomocí nich možné přenést nahrávku a zároveň emulovat konkrétní hodnoty dílčích vlastností sítě, jako třeba ztráty paketů. Pro účely této práce byla vytvořena zkušební virtuální síť, v které by se daly dostatečně rychle vytvořit a různě obměňovat síťová zapojení a dílčí podmínky. Toto virtuální prostředí bylo vytvořeno v programu VirtualBox. K dosažení specifických vlastností byl hledán potřebný emulátor, který by umožňoval napodobení parazit a vlastností rozsáhlých sítí. Po vyzkoušení několika variant byl zvolen program WANem. Tento program je blíže popsán v samostatné podkapitole. Při realizaci se vyskytly problémy s přenosem zvukové nahrávky na vstupu virtuálního počítače a hlavně se zaznamenáním zvukové nahrávky na výstupu virtuálního počítače, po průchodu ovlivněným přenosovým kanálem. Program Virtual Box totiž mapuje celou zvukovou kartu fyzického přístroje, na kterém běží přístroje virtuální. Nepodařilo se ovlivnit nastavení tak, aby bylo dosaženo potřebného výsledku pro simulace. Patrně by po úpravě a doprogramování do virtualizačního programu, jež je s otevřeným kódem, bylo možné výstupy přesměrovat. Toto řešení bylo shledáno jako časově nepřijatelné a od virtualizace bylo odstoupeno. Namísto virtualizovaného prostředí byla zapojena malá fyzická laboratorní síť. Tato varianta snadno umožňovala připojení zvukového výstupu libovolného zařízení (v našem případě počítače) na vstup počítače s programem zajišťujícím přenos a kódování zvukové informace přes ethernet.

### **3.1 WANem**

WANem je emulátor sítě typu WAN. Umožňuje emulovat vlastnosti a parazity sítě WAN v laboratorních podmínkách sítě LAN. Program WANem může sloužit jako transparentní přenosová brána, simulující konkrétní síťové charakteristiky, jako například šířka pásma, ztráty paketů, zpoždění, či změna pořadí paketů, Jitter a podobně. Lze jej použít k simulaci síťového provozu přenosu hlasu - VoIP (voice over

IP), či dat. Tento program je vyvíjen jako OpenSource (software s otevřeným kódem) pod obecně známou licencí GPL verze 2. A za podmínek této licence je volně šiřitelný.

Samotné prostředí je postaveno na upravené linuxové distribuci Knoppix a je ke stažení v podobě tzv. Live CD. Po nahrání linuxové distribuce do RAM paměti počítače je ovládání dostupné pomocí intuitivního webové orientovaného grafického rozhraní (Obr. 3.1) a to jak z lokálního počítače, tak i z jiného počítače na stejné síti. [7]

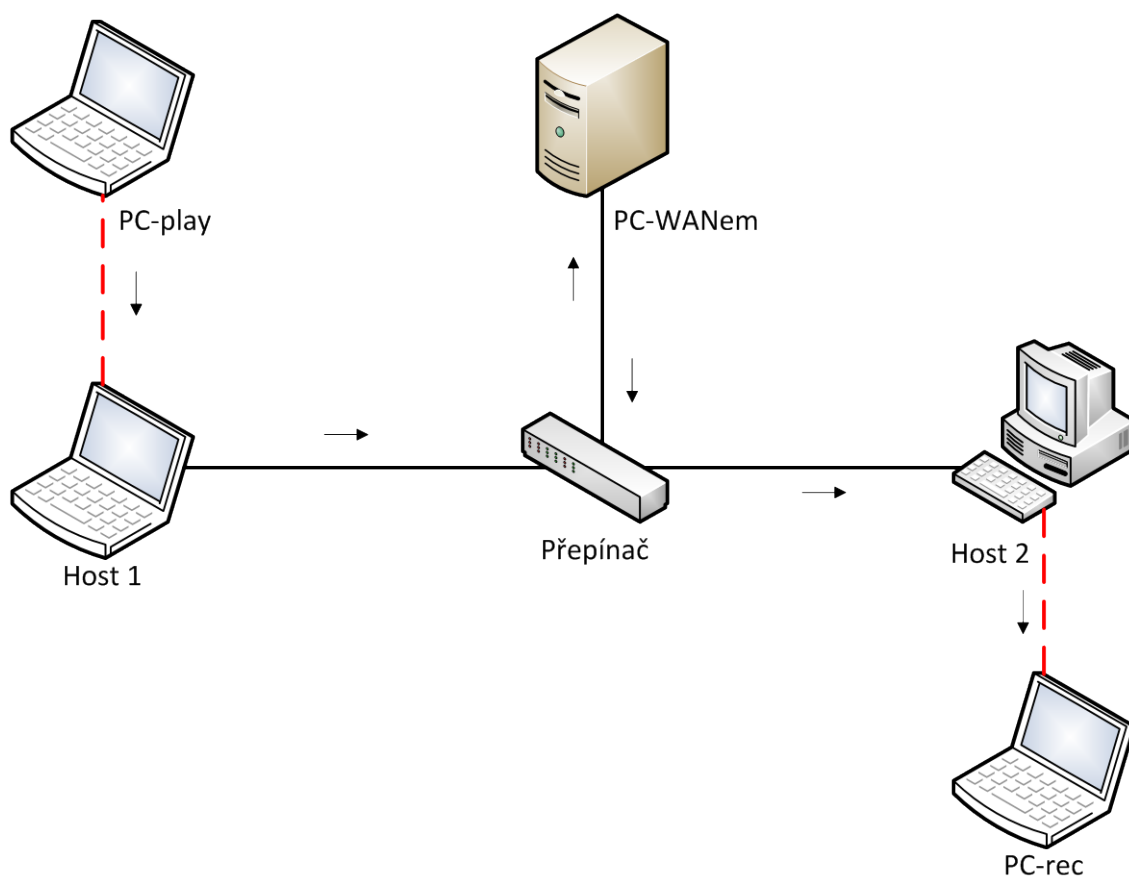


Obr. 3.1 Grafické uživatelské rozhraní programu WANem

V našem případě, byla tato aplikace použita k emulaci vlivu parazit sítě na přenášenou hlasovou informaci databáze TIMIT. Přenesené zvukové informace ovlivněné vybranými vlastnostmi sítě byly po přenosu zaznamenány. Tyto záznamy nazýváme souhrnně databáze.

## 3.2 Zapojení sítě s emulátorem

Použité laboratorní zapojení emulované sítě je znázorněno na obrázku (Obr. 3.2). Na PC-play jsou přehrávány promluvy databáze TIMIT s pípnutím; zvukový výstup počítače PC-play je zapojen do vstupu počítače Host 1. Na počítači Host 1 je nainstalován program Blink, pomocí nějž je zvuková informace kódována kodekem G.722 a přenášena po síti ethernet na počítač Host 2. Na počítači Host 2 je opět nainstalován program Blink a informace jsou dekodovány.



Obr. 3.2 Zapojení systému s emulátorem

Spojení mezi počítačem Host 1 a Host 2 je vedeno přes počítač PC-WANem, který emuluje přenosové parametry. Spojení přes počítač WANem je dosaženo přidáním záznamu do směrovacích tabulek na počítači Host 1 a Host 2.

Záznamy zavedené do směrovacích tabulek:

Host 1

```
route add 192.168.140.14 mask 255.255.255.255 192.168.140.20
```

Host 2

```
route add 192.168.140.12 mask 255.255.255.255 192.168.140.20
```

Běžně je spojení mezi klientskými programy využívajícími telefonie VoIP směrováno přes servery daného protokolu. Kvůli standardní komunikaci přes servery by emulace v našem zapojení, využívajícím program WANem, neovlivňovala přenášené zvukové informace. Právě proto byl použit program Blink, který umožňuje jak standardní spojení přes servery, nebo přímou komunikaci mezi klienty. K nalezení dalších klientů ve stejném síťovém rozsahu vysílá program Blink do sítě zprávy protokolu bonjour. Řečová informace je zakódována kodekem G7.22 a spojení mezi počítači zajištěno pomocí protokolu SIP. Z audio-výstupu počítače Host 2 je pomocí jack konektoru připojen vstup počítače PC-rec a přenesená zvuková informace je zaznamenána do formátu wav pomocí programu All2WAV Recorder. Přerušovanými čarami je značeno propojení audiokabelem se zakončením jack. Plnými čarami je znázorněno spojení kabelem strukturované kabeláže se zakončením RJ45.

## 3.3 Použitý software

### 3.3.1 VLC

Jedná se o multimedialní open-source přehrávač podporující dekódování téměř všech známých audio a video formátů. Nevyžaduje samostatné instalování kodeků, vše potřebné je implementováno v samotném programu. [8]

### 3.3.2 Blink

Jednoduchý klient pro VoIP telefonii využívající ke spojení protokol SIP, podporující audio kodeky G.722 a Speex, PSTN a další. Tento klient obsahuje podporu protokolu bonjour, pomocí kterého vyhledává další klienty na stejné síti, bez nutnosti SIP účtu a

vázanosti na SIP providera. Spojení je tedy možné realizovat pár vteřin po spuštění více klientů na stejné síti. [9]

### **3.3.3 GoldWave**

Velmi spolehlivý stabilní audio editor. Umožňuje záznam, přehrávání, úpravu, snadné dělení i spojování audio nahrávek, přidávání zvukových efektů a ukládání nahrávek do mnoha komprimovaných i bezkompresních formátů. [10]

### **3.3.4 Matlab**

Multiplatformní interaktivní programové prostředí obsahující vlastní programovací jazyk stejného jména. Původně byl vyvinut především pro matematické operace, v současnosti má široké využití, nejvíce je zřejmě využíván ve vědecko-výzkumných aplikacích. Umožňuje snadnou práci, počínaje aplikací matematických operací, přes počítačovou simulaci, po vývoj vlastních aplikací. [11]

## 4 DATABÁZE

Aby nebylo vytvářeno něco, co již existuje, a mohli jsme se rychle zaměřit na zkoumanou problematiku, byla použita již vytvořená databáze audio nahrávek TIMIT. Databáze TIMIT obsahuje mnoho promluv od mnoha mluvčích, celkově přibližně 5400. Z této databáze bylo vybráno 300 různých promluv od co nejvíce mluvčích, rovnoměrně zastoupených jak ženami, tak muži. Promluvy byly přejmenovány tak, aby názvem nahrávky bylo sudé číslo. Lichá čísla tvořila stále stejná nahrávka, obsahující navíc pípnutí. Pípnutí bylo vytvořeno složením fázově posunutých sinusových signálů s vysokou amplitudou, jednoho o kmitočtu 2000Hz a druhého o kmitočtu 2500Hz. Na vstupní straně byly vždy přehrávány promluvy ve stejném pořadí, proložené pípnutími.

Po průchodu přenosovým kanálem a ovlivněním řečového signálu emulovanou parazitou, jako jsou ztráty, či parametrem, jako je omezení šířky pásma, vznikla tzv. ovlivněná databáze. Po zaznamenání upravené databáze, která je tvořena třemi sty různými promluvami, bylo nutno pro další zpracování tyto promluvy ze souvislé nahrávky rozdělit na dílčí promluvy. Právě k tomu byla mezi jednotlivé nahrávky vložena pípnutí, která by zpětné rozdělení na dílčí promluvy podstatně usnadnila. Byl vytvořen algoritmus dělení skladeb, využívající detekce začátku pípnutí. Pokud program detekoval vzorek dostatečně vysoké, námi zvolené intenzity, označil úsek o délce tři sekund. A to 1s tomuto vzorku předcházející a 2s za tímto vzorkem následující. Označené úseky byly brány jako pípnutí a smazány. Mezery mezi těmito úseky byly označeny jako promluva a postupně uloženy pod sudými čísly od 2 do 600.

Bohužel se při práci ukázalo, že pro automatické rozdělení mnoha zaznamenaných databází je tento algoritmus nepoužitelný. Při vysokých ztrátách, při velkém kolísání zpoždění a podobně, dochází k tvorbě krátkých úseků vysoké intenzity, které jsou poté detekovány jako pípnutí. K takovýmto parazitám, znějícím při poslechu jako lupnutí či praskání, docházelo již na začátku zpracované databáze, a poté, co byly vymazány úseky, které ve skutečnosti obsahovaly promluvy, docházelo při dalším zpracování k chybnému označování dalších úseků a výsledná data ztratila relevanci. Proto bylo nutné použít dělení na jiném principu. O použitém algoritmu je více uvedeno v samostatné podkapitole .

## 4.1 Předúprava databází

Pro samotné dělení databáze na dílčí skladby, tedy promluvy a pípnutí, bylo v programu Matlab naprogramováno několik skriptů.

### 4.1.1 Zjištění délky originálních promluv

Pro rozdělení databáze na promluvy a pípnutí, využívající původní délky pípnutí a promluv, bylo nejprve nutné zjistit jejich původní délky. K zjištění délky byl naprogramován skript `Length_of_waves.m` (Příloha A.1). Tento skript nejprve načte jména všech souborů s příponou `.wav` v aktuální složce. Poté jsou dílčí promluvy načteny do proměnné  $f$  a v dalším kroku je délka dané promluvy uložena do pole  $nsp(i)$ . Takto je v poli  $nsp$  uložena délka s počtem vzorků. Tuto délku budeme používat při dalším zpracování. Kdybychom chtěli získat délku dílčí skladby v sekundách, získáme ji tak, že podělíme délku ve vzorcích vzorkovací frekvencí  $fs$ , což je v našem případě 16 000 Hz. Délky v sekundách poté uložíme do vektoru  $times$ . Takto získaná data z vektoru  $nsp$  uložíme do textového souboru `nsp.txt`, odkud budeme tato data načítat při dalším zpracování. Celkovou délku originální databáze zjistíme sečtením dílčích délek promluv a pípnutí, v našem případě tedy provedení sumy vektoru  $nsp$ .

### 4.1.2 Normalizace okrajů databází

Pro rozdělení skladeb algoritmem využívajícím původní délky je třeba vysoká přesnost. K dosažení požadované přesnosti je nutná další normalizovaná úprava zaznamenané databáze před jejím samotným dělením. K tomuto byl naprogramován v prostředí Matlab skript `orezDB.m` (Příloha A.2A.2). Tento skript od začátku prochází jednotlivé vzorky, dokud nenarazí na vzorek s hodnotou vyšší než 0,4. Poté se přesune o 16 036 vzorků zpět k začátku a označí tento vzorek jako první. Vše před ním je vymazáno. Díky tomuto má každá databáze jednu vteřinu ticha a poté následují promluvy. Obdobného mechanismu je využito i k určení konce databáze a za poslední promluvou jsou tři sekundy ticha a vše další je ořezáno. Pro výslednou kontrolu, že vše proběhlo v pořádku, je do konzole Matlabu vepsána původní a nová délka databáze.

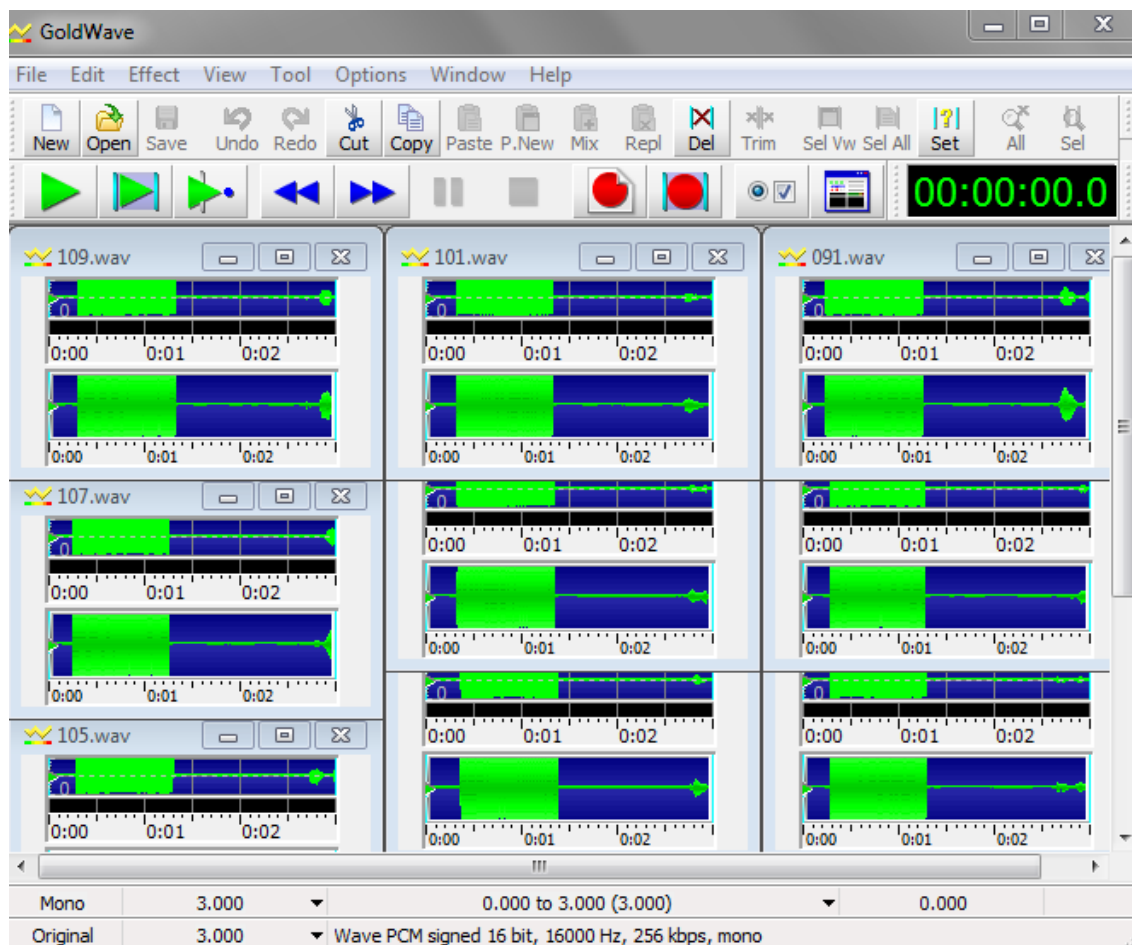


## 4.2 Rozdělení databáze na dílčí skladby

Pro rozdělení většiny databází byl použit skript naprogramovaný v programu Matlab. Ve vybraných případech však bylo nutno databázi rozdělit ručně. Blíže k problematice se dozvíme v následujících podkapitolách.

### 4.2.1 Algoritmus dělení skladeb podle původní délky

Po předchozím předzpracování je pomocí naprogramovaného skriptu PartingDB.m, viz. A.3, databáze dělena na dílčí promluvy a pípnutí a pojmenována originálním názvem. Aby nebylo nutno při zpracování každé databáze měnit jméno databáze v několika místech skriptu, je vždy na začátku jméno databáze uloženo do proměnné *dbname*. Nejdříve je načten wav soubor zpracovávané databáze, poté je určena její délka. Dále jsou do proměnné *nsp* načteny délky dílčích promluv a pípnutí z textového souboru *nsp.txt*. Poté je vypočten rozdíl délky databáze prošlé přenosovým kanálem a originální databáze. Dále je vypočten rozdíl těchto databází vyjádřený v procentech a také v sekundách. Tyto údaje jsou vypsány do konzolového okna Matlabu. Takto bylo zjištěno, že databáze nahraná po průchodu přenosovým kanálem je běžně o cca 2 minuty delší, než databáze originální. Tento nárůst délky je způsoben zpožděním při kódování hlasu programem Blink na odesílací straně a taktéž dekódováním na straně přijímací. Další drobné zpoždění zřejmě vzniká při přechodu mezi jednotlivými skladbami přehrávanými za sebou v programu VLC. Abychom kompenzovali rozdílnost délek databází, je vypočteno průměrné zpoždění *avgdelay* mezi dílčími skladbami. Toto zpoždění je vypočteno jako podíl délky rozdílu databází a 601, což je počet promluv včetně pípnutí. Tento podíl je pak ještě násoben koeficientem *koef*. Výsledek je poté zaokrouhlen na celé číslo. Výsledkem počet vzorků průměrného zpoždění mezi skladbami. Koeficient *koef* byl do celého algoritmu přidán z potřeby ručního zásahu. Po prohlédnutí rozdělených promluv v programu GoldWave byl koeficient *koef* poupraven tak, aby kompenzoval posunutí způsobené zpožděním při kódování/dekódování a přenosem po síti. Posunutí můžeme vidět na Obr. 4.1.



Obr. 4.1 Rozdělené skladby s posunutím (Zobrazeno v programu GoldWave).

## 4.2.2 Ruční rozdělení databáze

Někdy je příliš obtížné naprogramovat adaptibilní program řešící nepravidelně se měnící vstupní podmínky, jako je příliš velké kolísání zpoždění a ztráty rozsáhlých úseků nahrávky. Proto bylo v případě jedné databáze nutné rozdělit ji manuálně. Zobrazování, poslech a ukládání vybraných úseků databáze probíhalo v programu GoldWave, který je určen právě ke zpracování audio nahrávek.

## 4.3 Jitter

Pro tento parametr bylo vytvořeno 7 databází s jittrem 5, 10, 20, 30, 40, 55 a 75 ms. Vzhledem k tomu, že jitter je kolísání zpoždění, u všech databází v této skupině bylo zvoleno zpoždění 40ms. Tato hodnota odpovídá typickému zpoždění ve WAN sítích

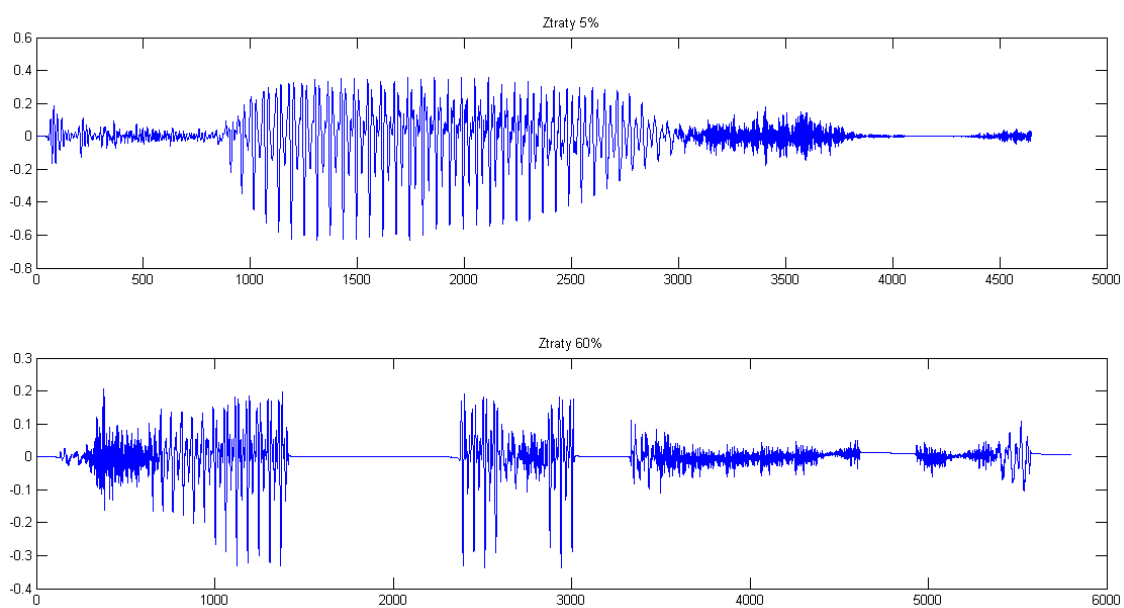
[11]. Hodnoty jitteru byly vybrány tak, aby maximální doporučená hodnota pro tento parametr ve WAN sítích se blížila mediánu vybraných hodnot. Doporučený maximální jitter ve WAN sítích je 30ms.[11] Každá databáze v této skupině obsahuje 120 promluv.

## 4.4 Šířka pásma

Pro tento parametr bylo vytvořeno 7 databází se šířkou pásma 1024, 512, 256, 128, 96, 48 a 24 kb/s. Hodnoty byly odstupňovány v tomto intervalu s předpokladem, že šířka pásma 1024 kb/s bude dostatečná pro to, aby nedocházelo k žádným ztrátám a zpoždění. A naopak u nahrávek s šířkou pásma pod 128 kb/s začnou být ztráty pozorovatelné při vykreslení nahrávky, a s dalším zmenšováním hodnot budou i slyšitelné. U databází s šířkou pásma 48 a 24 kb/s nastal problém s automatickým dělením skladeb pomocí skriptu. Tento skript nebyl použitelný z důvodu velké nelinearity zpoždění a kolísání ztrát u takto úzkého přenosového pásma. Databáze se šířkou pásma 48kb/s byla rozdělena ručně a k další práci bylo použito pouze 6 databází. Každá databáze v této skupině obsahuje 79 promluv. Některé promluvy byly vynechány, protože nebyly zastoupeny ve všech databázích. Především v databázích s nejmenší šířkou pásma docházelo často ke ztrátě podstatné části promluvy, nebo dokonce ke ztrátě celé promluvy. A to cca každé deváté.

## 4.5 Ztráty

Pro tento parametr bylo vytvořeno 7 databází se ztrátami 5, 10, 20, 30, 40, 50 a 60%. Tyto hodnoty byly vybrány tak, aby databáze s nejvyššími ztrátami byly již snadno pozorovatelné při vykreslení. Na obrázku (Obr. 4.2) je vidět výřez slova „close“ z promluvy č. 58, v horní části se ztrátami 5% a v dolní části 60%. Při ztrátách šedesáti procent paketů nejsou po dekodování na přijímací straně zrekonstruovány podstatné části slova, také je slovo o téměř 1000 vzorků delší kvůli zpoždění. Každá databáze v této skupině obsahuje 120 promluv.



Obr. 4.2 Výřez slova „close“ při ztrátách 5% (nahore) a 60% (dole).

# 5 METODIKA ROZPOZNÁNÍ PARAMETRŮ PŘENOSOVÉHO KANÁLU

## 5.1 Rozhodovací stromy

Klasifikace (ohodnocení) zařazuje data do předdefinovaných skupin (tříd), algoritmus se učí charakteristiku tříd na trénovací množině a to na datech, pro které známe jejich příslušnost ke skupinám [17]

Rozhodovací stromy mají široké využití, v případě této práce byly využity jako jeden z možných klasifikačních modelů. Rozhodovací strom lze definovat jako strom, kde každý nelistový uzel stromu, představuje test na hodnotu atributu a větve vedoucí z tohoto uzlu, představují možné výsledky tohoto testu. Klasifikace probíhá cestou záznamu od kořene k listu stromu. V každém kroku proběhne test a cesta dále pokračuje po větvi výsledku testu. Jakmile záznam dojde do listového uzlu, je oklasifikován třídou, kterou identifikuje hodnota listu stromu. [18]

Konstrukci rozhodovacího stromu z trénovacích vzorků  $S$  podle atributů  $D$  provádí algoritmus (převzatý z [18]):

1. vytvoř uzel  $N$ ,
2. pokud jsou vzorky v  $S$  stejné třídy  $C$ , tak ohodnot'  $N$  identifikátorem  $C$  a skonči,
3. pokud je  $D$  prázdné, tak ohodnot'  $N$  identifikátorem třídy převládající v  $S$  a skonči,
4. ohodnot'  $N$  atributem  $t$  z  $D$ , který má „největší informační zisk“,
5. pro každou hodnotu  $v_i$  atributu  $t$  vytvoř z  $N$  hranu ohodnocenou podmínkou  $t=v_i$  do uzlu vzniklého spuštěním tohoto algoritmu s podmnožinou trénovacích vzorků  $S$ , které vyhovují uvedené podmínce, a podmnožinou atributů  $D$  bez atributu  $t$ .

Kvůli možnému neoptimálnímu tvaru, nevhodnému pro rychlou konkrétní klasifikaci

jsou rozhodovací stromy tzv, ořezávány (tree-pruning). Používají se 2 způsoby ořezávání a to při a po konstrukci rozhodovacího stromu.

Ořezání při konstrukci stromu proběhne předčasným ukončením některých větví rozhodovacího stromu již při tvorbě stromu. Důvodem k takovému ukončení konstrukce větví stromu, může být dostatečně velká pravděpodobnost, náležitosti dat do určité třídy. Problém je v určení hranice dostatečné pravděpodobnosti.

Ořezání po konstrukci stromu, kdy jsou některé větve již vytvořeného stromu odstraněny a nahrazeny rovnou listy a to takovými, které byly ohodnoceny třídou s největší četností listů daného podstromu, ořezané větve. Ořezávání po konstrukci stromu poskytuje lepší výsledky, ale je více výpočetně náročné.[18]

V práci bylo tedy pro klasifikaci využíváno rozhodovacích stromů s metodou ořezání stromu až po jeho konstrukci.

## 5.2 Příznaky

Z různých skupin příznaků byly vybrány některé reprezentující příznaky a ty byly dále testovány. Ze skupiny Perceptual spectral features (percepční spektrální funkce) byl vybrán příznak LMCC a příznak MFCC. Ze skupiny Energies and zcr (Zcr a energie) byly použity pro další práci příznaky Temporal Energy (v textu zkrácený zápis TpEn) a Teager energy operator (v textu zkrácený zápis TEO). Skupina Spectral features (spektrální funkce) je v textu vedena jako samostatný příznak.

## 5.3 Výběr na základě korelace

Na základě korelace byly vybrány příznaky. Korelaci dvou parametrů (příznaků) lze popsat jako závislost. Závislost znamená, že hodnoty jednoho příznaku odpovídají přímo úměrně (kladná korelace) nebo neúměrně (záporná korelace) hodnotám druhého příznaku. Pokud jsou hodnoty obou příznaků nezávislé, bude korelace blízká nule.

Pomocí programu Hila vytvořeném v programu Matlab, byly vypočítána korelační koeficienty  $R$ , normované od -1 do 1, kde příznaky blízkí se absolutní hodnotě 1 jsou nejvíce korelovány a příznaky blízkí se 0 jsou nejméně korelovány.

Rovnice pro výpočet korelačního koeficientu R:

$$R = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}}, [16] \quad (5.1)$$

kde  $x$  a  $y$  jsou testované veličiny.

Koeficient determinace  $R^2$  se vypočítá podle vztahu [16]:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y(x_i))^2}{\sum_{i=1}^n (y_i - \langle y_i \rangle)^2} \quad (5.2)$$

Kde  $x$  je x-ová souřadnice datového bodu,

$y$  je y-ová souřadnice datového bodu,

$n$  je počet datových bodů,

$\langle y_i \rangle$  se vypočítá podle následujícího vztahu:

$$\langle y_i \rangle = \frac{\sum_{i=1}^n y_i}{n} \quad (5.3)$$

Podle hodnoty Indexu (koeficientu) korelace určuje míru závislosti. Když bude mít Index (koeficient) korelace hodnoty:

$R^2 = 0,0 - 0,3$       jedná se o slabou závislost,

$R^2 = 0,3 - 0,7$       jedná se o průměrnou závislost,

$R^2 = 0,7 - 0,9$       jedná se o silnou závislost,

$R^2 = 0,9 - 1,0$       jedná se o velmi silnou závislost.

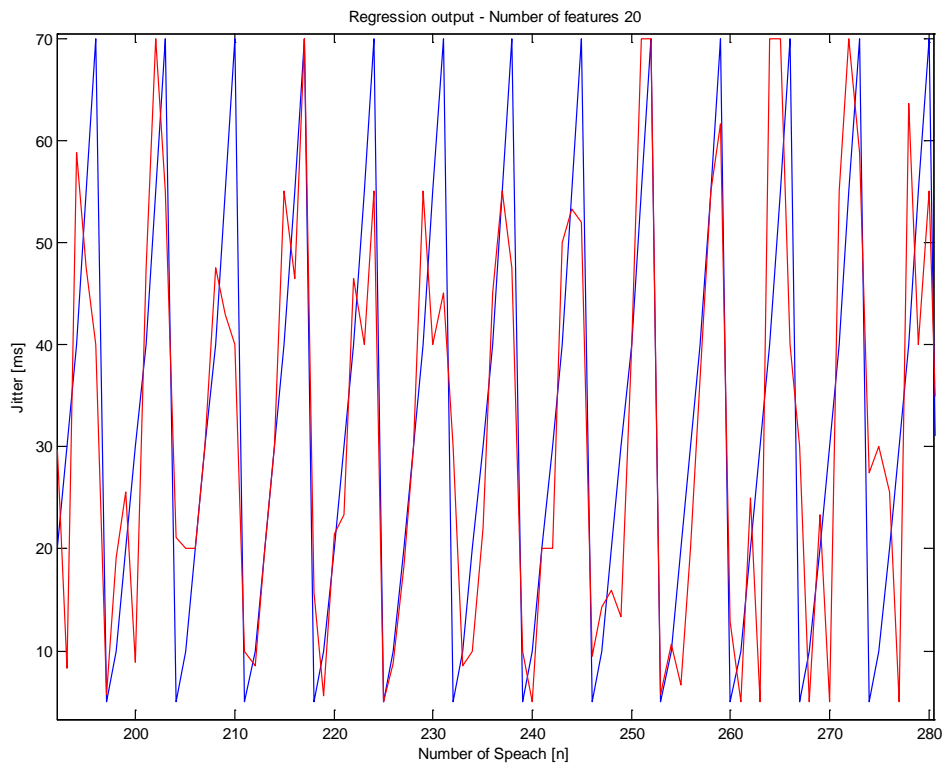
Střední absolutní odchylka je popsána vztahem [13]:

$$\text{Std. absolutní. odchylka} = \frac{1}{n} \sum |x_i - \bar{x}|. \quad (5.4)$$

Střední kvadratická odchylka je popsána vztahem [13]:

$$\text{Std. kv. odchylka} = \frac{1}{n} \sum (x_i - \bar{x})^2. \quad (5.5)$$

Dále jsou již tabulky a grafy testovaných příznaků a jejich korelace reálné hodnoty s regresně vypočtenou (odhadnutou) hodnotou. Detail průběhu reálných hodnot a regresně vypočtených hodnot pro databáze s jitterem s počtem funkcí (Number of features) 20 je vidět na Obr. 5.1 vykresleném v programu Hila.



Obr. 5.1 Znáznornění korelace reálných hodnot a regresně vypočtených hodnot

V tabulkách jsou výsledky výpočtů korelačních koeficientů  $R$ , jejich druhá mocnina  $R^2$  (nebo-li koeficienty determinace), střední kvadratická odchylka a střední absolutní odchylka pro dílčí příznaky daných 3 parametrů, z nichž každý má 7



klasifikačních skupin.

### 5.3.1 Jitter

V tabulkách je vždy červeným textem označen příznak s nejvyšší korelací nebo kombinace příznaků mající pozitivní vliv na dosaženou korelaci.

Pro parametr jitter (kolísání zpoždění) byly nejprve vypočítány hodnoty korelace a dalších parametrů pro samostatné příznaky jak je vidět v Tab. 5.1

Tab. 5.1 JT - Korelace pro jednotlivé příznaky

JT [ms]	MFCC	LFCC	SpFe	TEO	TpEn
R <sup>2</sup>	0,61	0,58	0,65	0,00	0,59
R	0,81	0,79	0,83	0,23	0,80
Střed. kv. odchylka	13,68	14,11	12,66	26,06	13,94
Střed. abs. odchylka	10,60	11,11	9,59	20,32	10,90

Poté byly vypočítány korelace pro kombinace vždy dvou příznaků (Tab. 5.2).

Tab. 5.2 JT - Korelace pro kombinace dvou příznaků

JT [ms]	MFCC LFCC	LFCC SpFe	LFCC TEO	LFCC TpEn	SpFe MFCC	SpFe TEO	SpFe TpEn	MFCC TEO	MFCC TpEn	TEO TpEn
R <sup>2</sup>	0,57	0,62	0,58	0,58	0,65	0,65	0,65	0,61	0,61	0,56
R	0,80	0,81	0,79	0,79	0,83	0,83	0,83	0,81	0,81	0,79
Střed. kv. odchylka	13,76	13,41	14,11	14,11	12,66	12,66	12,66	13,68	13,68	14,06
Střed. abs. odchylka	10,58	10,21	11,11	11,11	9,59	9,59	9,59	10,60	10,60	10,89

Následně byly vypočítány korelace všech pěti příznaků. Poté byly příznaky separovány, na základě zkušeností (výsledků z tabulek Tab. 5.1 a Tab. 5.2), které příznaky neovlivňují výsledek korelace a nebo jej ovlivňují negativně. Byly ponechány jen kombinace příznaků, jež pozitivně ovlivňují výsledek korelace. Výsledky tohoto postupu jsou zaznamenány do Tab. 5.3.

Tab. 5.3 JT - Korelace pro vybrané kombinace příznaků

JT [ms]	All5	SpFe MFCC LFCC TpEn	SpFe LFCC TpEn	SpFe MFCC LFCC
R <sup>2</sup>	0,62	0,62	0,62	0,62
R	0,81	0,81	0,81	0,81
Střed. kv. odchylka	13,41	13,41	13,41	13,41
Střed. abs. odchylka	10,21	10,21	13,41	10,21

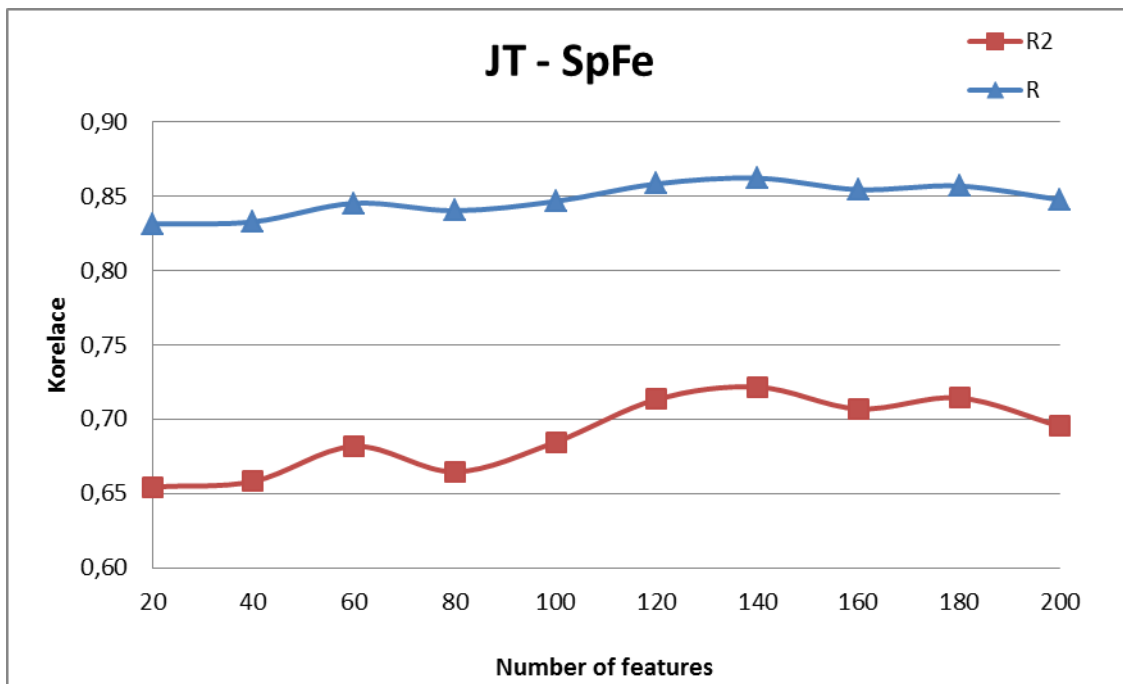
Jak je vidět v tabulkách Tab. 5.1, Tab. 5.2 a Tab. 5.3., nejvyšší korelace dosahuje vždy příznak SpFe (Spectral Features), ať už samostatný nebo v kombinaci s jiným příznakem pozitivně ovlivňujícím výsledek. Ze všech výpočtů byly vybrány tři nejlépe ovlivňující příznaky nebo jejich kombinace a počítány pro ně korelace se zvyšujícím se nOf, výsledky jsou v tabulkách

Vypočítané výsledky korelačních koeficientů a jejich střední odchylky pro nejvíce korelující samostatný příznak SpFe (Spektral koeficient) s rostoucím nOf (number Of features) jsou zaznamenány v tabulce Tab. 5.4 a poté vyneseny do grafů Obr. 5.2 Korelace a Obr. 5.2 Střední odchylky.

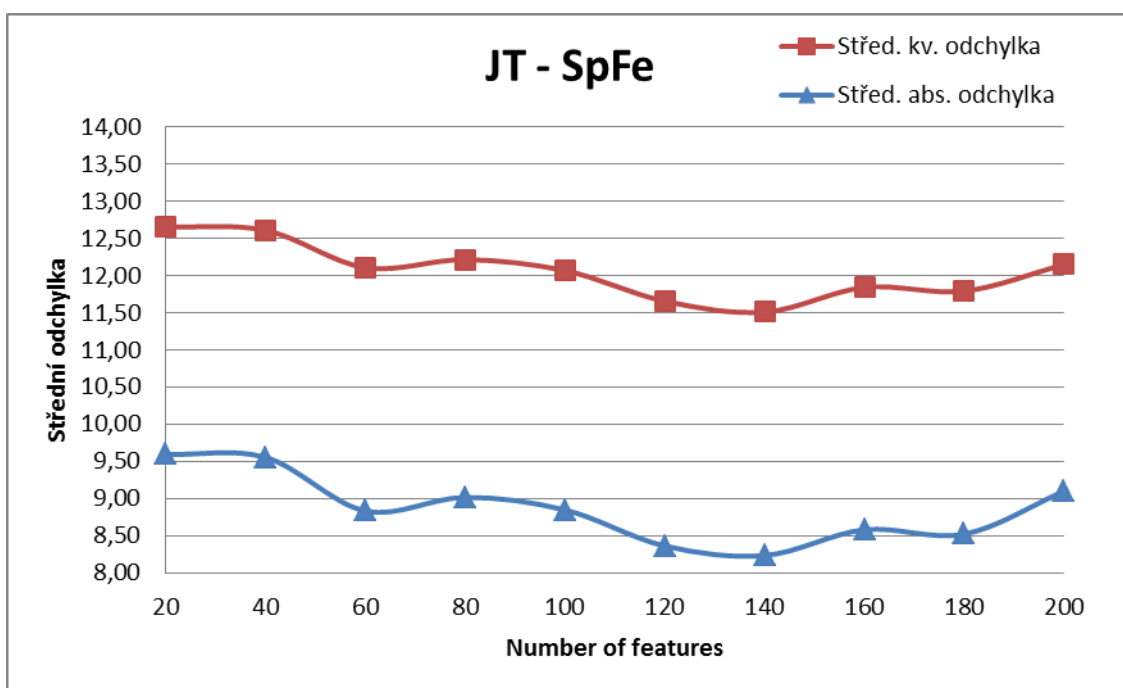
Tab. 5.4 JT- Korelace pro příznak SpFe – zvyšování nOf

JT [ms] SpFe (nOf)	20	40	60	80	100	120	140	160	180	200
R <sup>2</sup>	0,65	0,66	0,68	0,66	0,68	0,71	0,72	0,71	0,71	0,70
R	0,83	0,83	0,85	0,84	0,85	0,86	0,86	0,85	0,86	0,85
Střed. kv. odchylka	12,66	12,61	12,11	12,21	12,07	11,66	11,51	11,84	11,79	12,15
Střed. abs. odchylka	9,59	9,55	8,83	9,02	8,84	8,36	8,23	8,58	8,52	9,10

Vypočítané výsledky korelačních koeficientů a jejich střední odchylky pro nejvíce korelující kombinaci dvou příznaků SpFe (Spektral koeficient) a MFCC s rostoucím nOf (number Of features), jsou zaznamenány v tabulce Tab. 5.5.



Obr. 5.3 Korelace pro příznak SpeFe.



Obr. 5.4 Střední odchylky pro příznak SpeFe – zvyšování nOf.

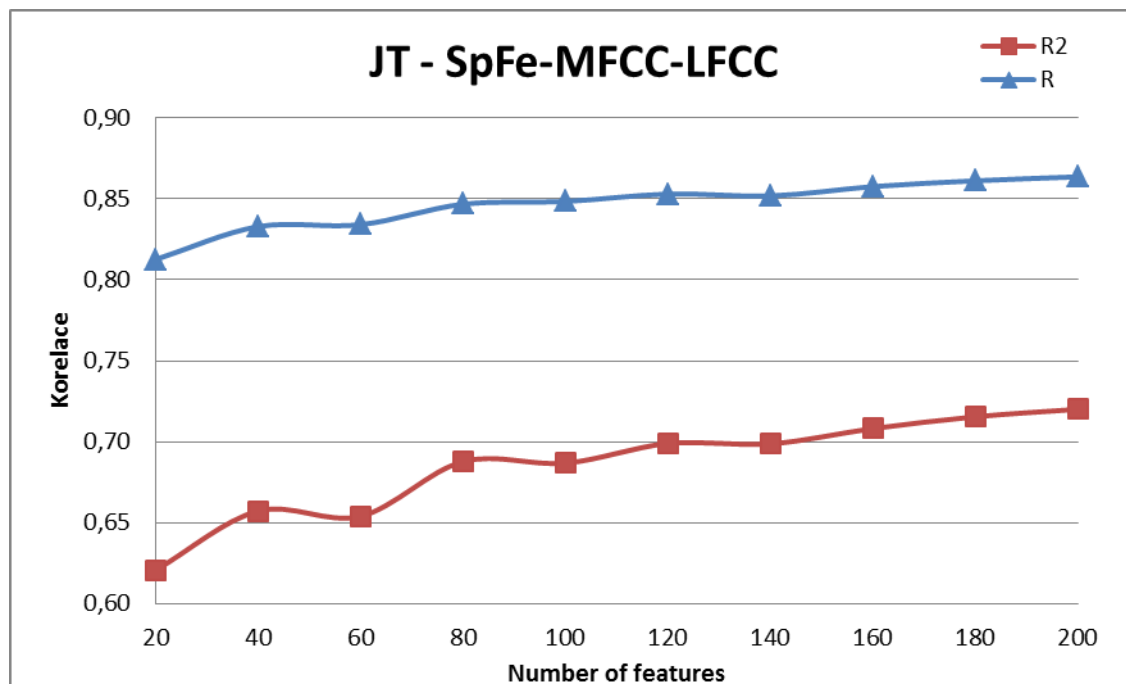
Tab. 5.5 JT- Korelace pro kombinaci příznaků SpeFe a MFCC – zvyšování nOf.

JT [ms] SpFe-MFCC (nOf)	20	40	60	80	100	120	140	160	180	200
R <sup>2</sup>	0,65	0,66	0,68	0,66	0,68	0,71	0,72	0,71	0,71	0,70
R	0,83	0,83	0,85	0,84	0,85	0,86	0,86	0,85	0,86	0,85
Střed. kv. odchylka	12,66	12,61	12,11	12,21	12,07	11,66	11,51	11,84	11,79	12,15
Střed. abs. odchylka	9,59	9,55	8,83	9,02	8,84	8,36	8,23	8,58	8,52	9,10

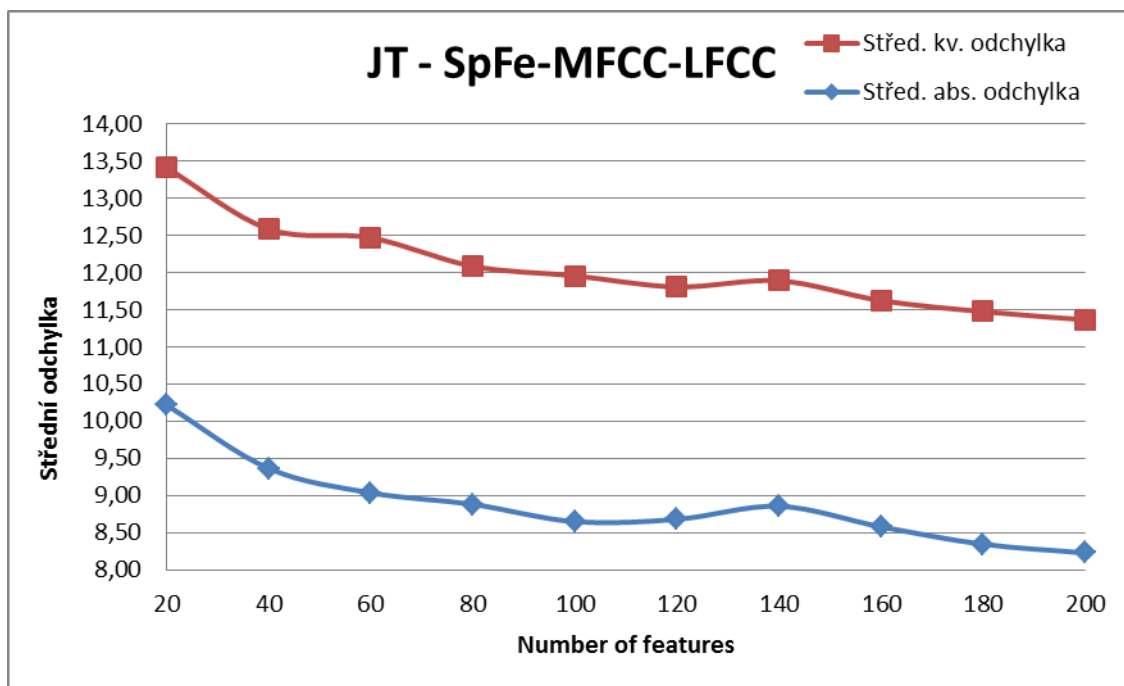
Vypočítané výsledky korelačních koeficientů a jejich střední odchylky pro nejvíce korelující kombinaci tří příznaků SpFe (Spektral koeficient), MFCC a LFCC s rostoucím nOf (number Of features), jsou zaznamenány v tabulce Tab. 5.6 a poté vyneseny do grafů Obr. 5.4 Korelace a Obr. 5.5 Střední odchylky.

Tab. 5.6 JT - Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf.

JT [ms] SpFe (nOf)	20	40	60	80	100	120	140	160	180	200
R <sup>2</sup>	0,65	0,66	0,68	0,66	0,68	0,71	0,72	0,71	0,71	0,70
R	0,83	0,83	0,85	0,84	0,85	0,86	0,86	0,85	0,86	0,85
Střed. kv. odchylka	12,66	12,61	12,11	12,21	12,07	11,66	11,51	11,84	11,79	12,15
Střed. abs. odchylka	9,59	9,55	8,83	9,02	8,84	8,36	8,23	8,58	8,52	9,10



Obr. 5.5 Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf.



Obr. 5.6 JT - Střední odchylka pro kombinaci příznaků SpFe a MFCC a LFCC – zvyšování nOf

Nejvyšší hodnot korelačních koeficientů pro jitter dosahuje kombinace příznaků SpFe, MFCC a LFCC, kdy korelační koeficient se zvyšuje s použitým množstvím nOf. Zároveň se se zvyšujícím množstvím použitých funkcí snižují střední absolutní i kvadratická odchylka. Na úroveň korelace naopak nemá vliv příznak TEO

### 5.3.2 Ztráty

Pro parametr ztráty byly nejprve vypočítány hodnoty korelace a dalších parametrů pro samostatné příznaky jak je vidět v Tab. 5.7.

Tab. 5.7 ZT - Korelace pro jednotlivé příznaky

ZT [%]	MFCC	LFCC	SpFe	TEO	TpEn
$R^2$	0,40	0,15	0,59	0,00	0,00
R	0,71	0,59	0,80	0,28	0,47
Střed. kv. odchylka	14,12	16,77	11,68	21,74	18,89
Střed. abs. odchylka	10,44	12,47	8,47	17,23	14,56

Poté byly vypočítány korelace pro kombinace vždy dvou příznaků (Tab. 5.8).

Tab. 5.8 ZT - Korelace pro kombinace dvou příznaků

ZT [%]	MFCC LFCC	LFCC SpFe	LFCC TEO	LFCC TpEn	SpFe MFCC	SpFe TEO	SpFe TpEn	MFCC TEO	MFCC TpEn	TEO TpEn
R <sup>2</sup>	0,41	0,59	0,15	0,15	0,59	0,59	0,59	0,40	0,40	0,00
R	0,72	0,80	0,59	0,59	0,80	0,80	0,80	0,71	0,71	0,50
Střed. kv. odchylka	14,07	11,78	16,77	16,77	11,76	11,68	11,68	14,12	14,12	18,33
Střed. abs. odchylka	10,13	8,56	12,47	12,47	8,15	8,47	8,47	10,44	10,44	13,86

Následně byly vypočítány korelace všech pěti příznaků. Poté byly příznaky separovány, na základě zkušeností (výsledků z tabulek Tab. 5.7 a Tab. 5.8), které příznaky neovlivňují výsledek korelace a nebo jej ovlivňují negativně. Byly ponechány jen kombinace příznaků, jež pozitivně ovlivňují výsledek korelace. Výsledky tohoto postupu jsou zaznamenány do Tab. 5.9.

Tab. 5.9 ZT - Korelace pro vybrané kombinace příznaků

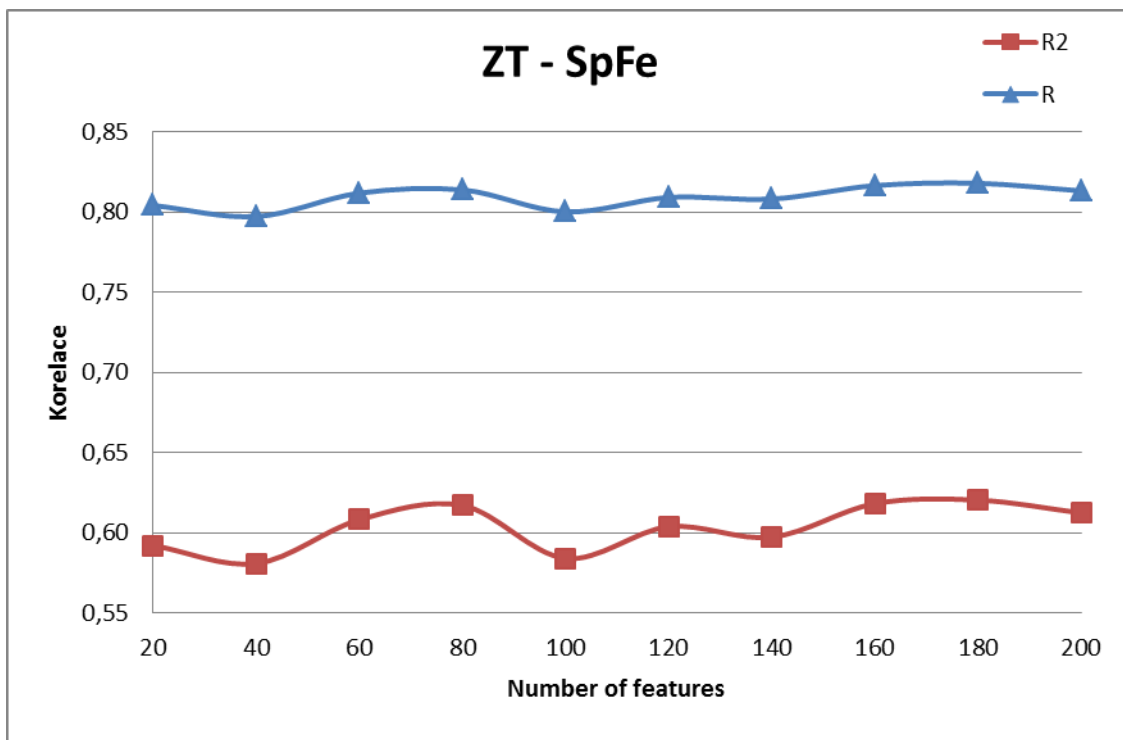
ZT [%]	All5	SpFe MFCC LFCC TpEn	SpFe LFCC TpEn	SpFe MFCC LFCC
R <sup>2</sup>	0,61	0,61	0,59	0,61
R	0,82	0,82	0,80	0,82
Střed. kv. odchylka	11,36	11,36	11,78	11,36
Střed. abs. odchylka	7,85	7,85	8,56	7,85

Jak je vidět v tabulkách Tab. 5.7, Tab. 5.8 a Tab. 5.9., nejvyšší korelace dosahuje vždy příznak SpFe (Spectral Features), ať už samostatný nebo v kombinaci s jiným příznakem pozitivně ovlivňujícím výsledek. Ze všech výpočtů byly vybrány tři nejlépe ovlivňující příznaky nebo jejich kombinace a počítány pro ně korelace se zvyšujícím se nOf, výsledky jsou v tabulkách

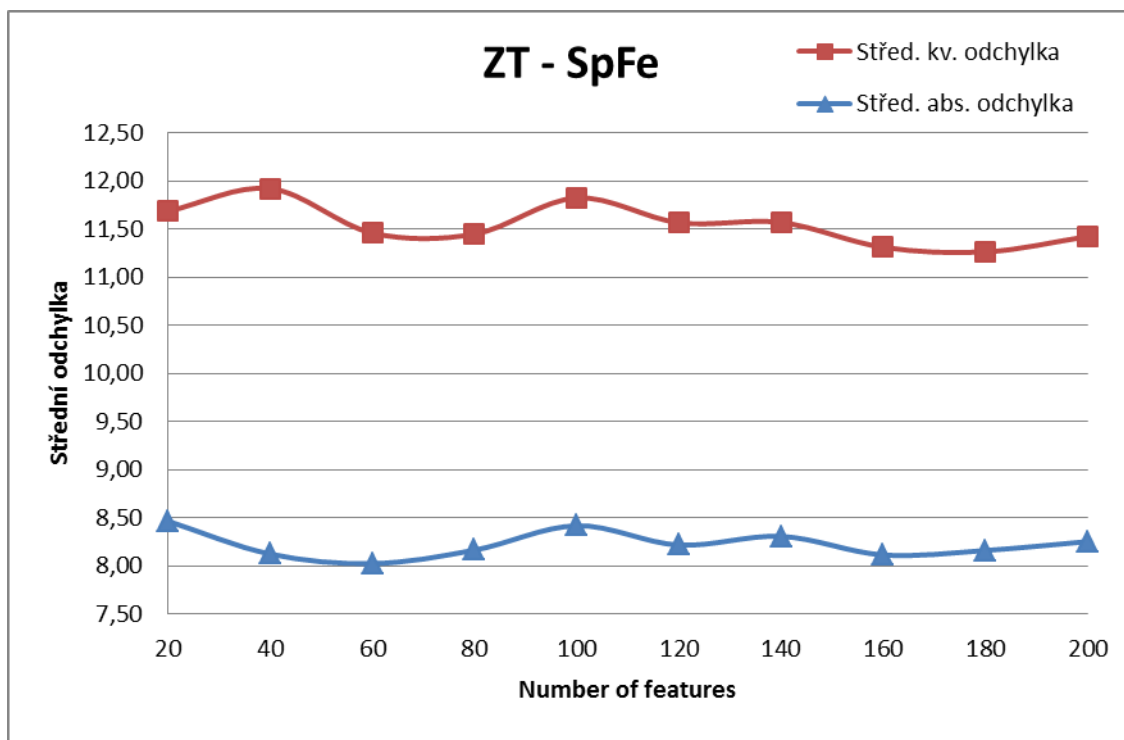
Vypočítané výsledky korelačních koeficientů a jejich střední odchylky pro nejvíce korelující samostatný příznak SpFe (Spektral koeficient) s rostoucím nOf (number Of features) jsou zaznamenány v tabulce Tab. 5.10 a poté vyneseny do grafů Obr. 5.6 Korelace a Obr. 5.7 Střední odchylky.

Tab. 5.10 ZT - Korelace pro příznak SpeFe – zvyšování nOf.

ZT [%] - SpFe (nOf)	20	40	60	80	100	120	140	160	180	200
$R^2$	0,59	0,58	0,61	0,62	0,58	0,60	0,60	0,62	0,62	0,61
R	0,80	0,80	0,81	0,81	0,80	0,81	0,81	0,82	0,82	0,81
Střed. kv. odchylka	11,68	11,92	11,46	11,45	11,82	11,57	11,57	11,31	11,26	11,42
Střed. abs. odchylka	8,47	8,12	8,02	8,16	8,42	8,22	8,31	8,11	8,16	8,25



Obr. 5.7 ZT - Korelace pro příznak SpeFe – zvyšování nOf.



Obr. 5.8 ZT - Střední odchylky pro příznak SpeFe – zvyšování nOf.

Vypočítané výsledky korelačních koeficientů a jejich střední odchylky pro nejvíce korelující kombinaci dvou příznaků SpFe (Spektral koeficient) a MFCC s rostoucím nOf (number Of features), jsou zaznamenány v tabulce Tab. 5.11.

Tab. 5.11 ZT - Korelace pro kombinaci příznaků SpeFe a MFCC – zvyšování nOf.

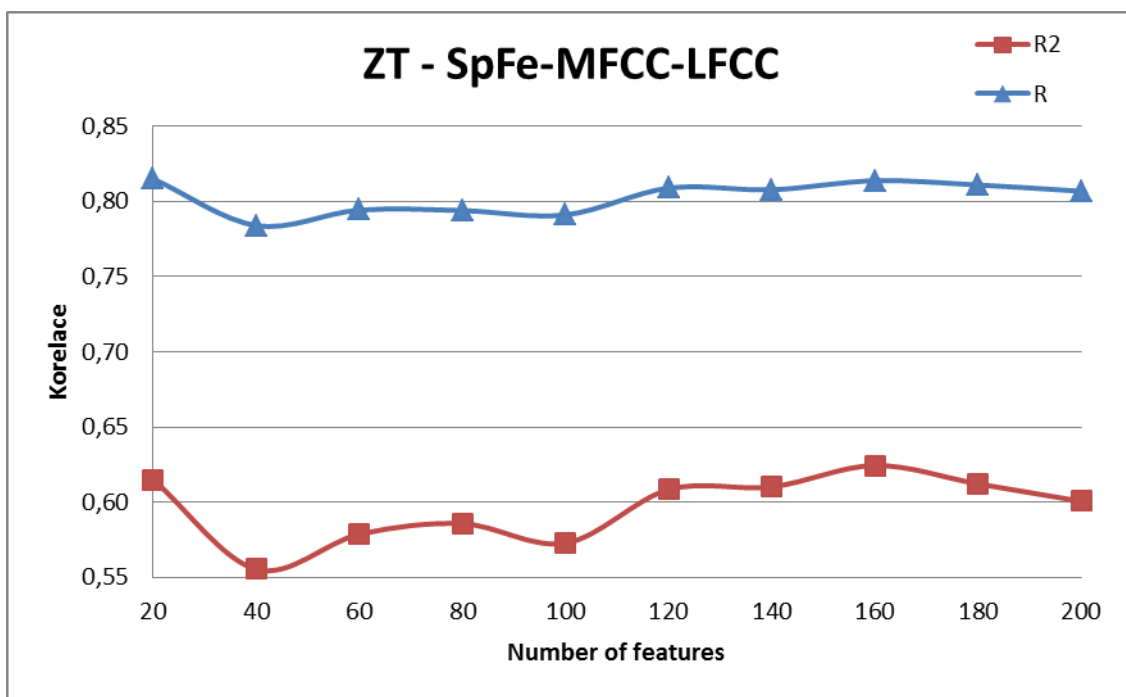
ZT [%] - SpFe-MFCC (nOf)	20	40	60	80	100	120	140	160	180	200
$R^2$	0,59	0,59	0,59	0,58	0,59	0,58	0,56	0,57	0,56	0,56
R	0,80	0,80	0,80	0,79	0,80	0,79	0,78	0,79	0,79	0,79
Střed. kv. odchylka	11,76	11,87	12,01	11,99	11,80	12,11	12,42	12,20	12,23	12,27
Střed. abs. odchylka	8,15	8,22	8,09	8,16	8,17	8,44	8,61	8,35	8,53	8,52

Vypočítané výsledky korelačních koeficientů a jejich střední odchylky pro nejvíce korelující kombinaci tří příznaků SpFe (Spektral koeficient), MFCC a LFCC s rostoucím nOf (number Of features), jsou zaznamenány v tabulce Tab. 5.12 a poté vyneseny do grafů Obr. 5.8 Korelace a Obr. 5.9 Střední odchylky.

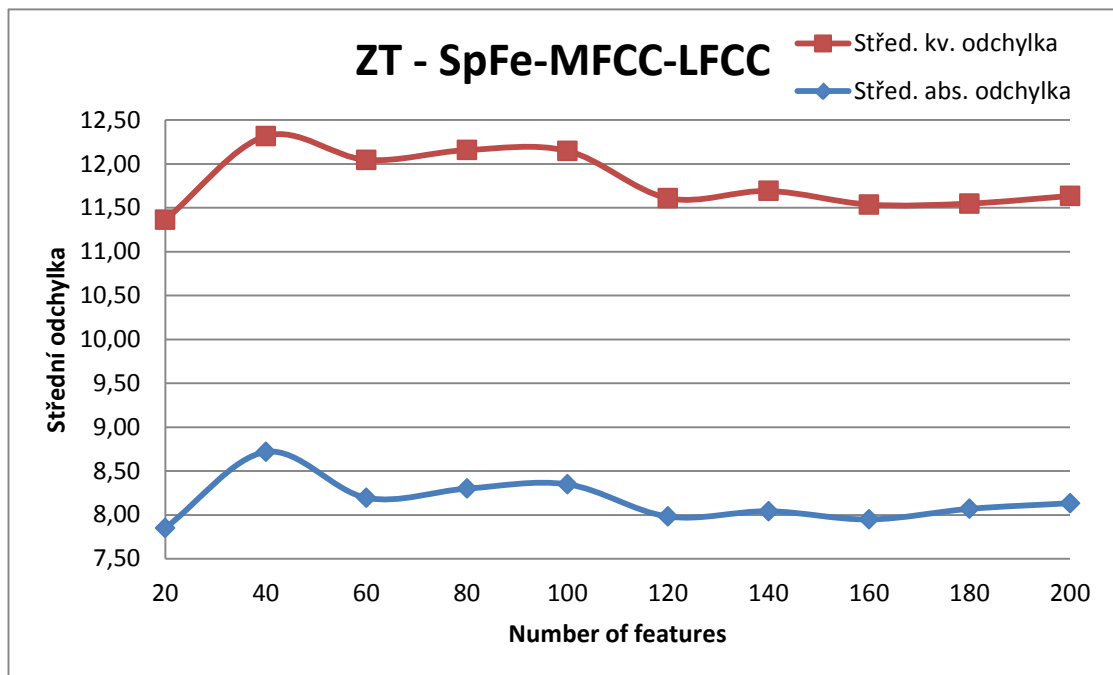


Tab. 5.12 ZT Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf.

ZT [%] - SpFe-MFCC-LFCC (nOf)	20	40	60	80	100	120	140	160	180	200
R <sup>2</sup>	0,61	0,56	0,58	0,59	0,57	0,61	0,61	0,62	0,61	0,60
R	0,82	0,78	0,79	0,79	0,79	0,81	0,81	0,81	0,81	0,81
Střed. kv. odchylka	11,36	12,32	12,05	12,16	12,15	11,61	11,69	11,54	11,55	11,64
Střed. abs. odchylka	7,85	8,72	8,19	8,30	8,35	7,98	8,04	7,95	8,07	8,13



Obr. 5.9 Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf.



Obr. 5.10 Střední odchylka pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf.

### 5.3.3 Šířka pásma do 1024 kb/s

Pro parametr Šířka pásma byly nejprve vypočítány hodnoty korelace a dalších parametrů pro samostatné příznaky jak je vidět v Tab. 5.13.

Tab. 5.13 SP - Korelace pro jednotlivé příznaky

SP [kb/s]	MFCC	LFCC	SpFe	TEO	TpEn
$R^2$	0,00	0,00	0,00	0,00	0,00
R	0,30	0,21	0,21	0,26	0,29
Střed. kv. odchylka	395,20	405,86	410,64	399,10	392,44
Střed. abs. odchylka	276,20	286,82	280,13	275,93	267,44

Poté byly vypočítány korelace pro kombinace vždy dvou příznaků.

Tab. 5.14 SP - Korelace pro kombinace dvou příznaků

SP [kb/s]	MFCC LFCC	LFCC SpFe	LFCC TEO	LFCC TpEn	SpFe MFCC	SpFe TEO	SpFe TpEn	MFCC TEO	MFCC TpEn	TEO TpEn
R <sup>2</sup>	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
R	0,25	0,19	0,21	0,21	0,16	0,21	0,21	0,30	0,30	0,30
Střed. kv. odchylka	403,8	412,9	405,9	405,9	421,5	410,6	410,6	395,2	395,2	395,2
Střed. abs. odchylka	276,4	293,7	286,8	286,8	290,4	280,1	280,1	276,2	276,2	276,2

Následně byly vypočítány korelace všech pěti příznaků. Poté byly příznaky separovány, postupem na základě zkušeností, z předchozích kapitol. Výsledky tohoto postupu jsou zaznamenány do Tab. 5.15.

Tab. 5.15 SP - Korelace pro vybrané kombinace příznaků

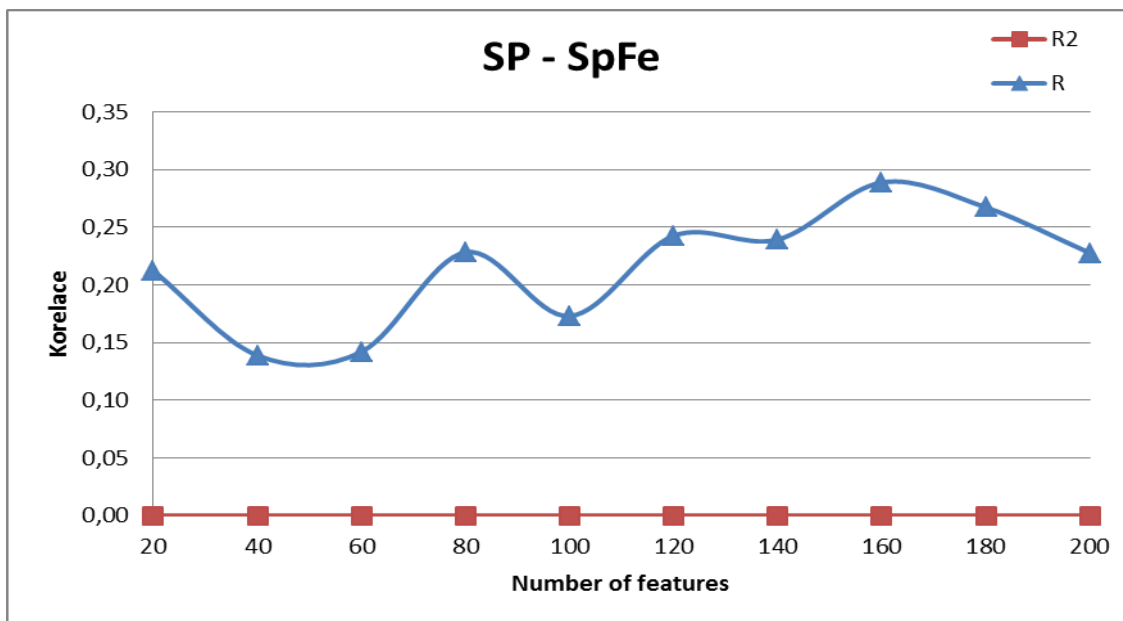
SP [kb/s]	All5	SpFe MFCC LFCC TpEn	SpFe LFCC TpEn	SpFe MFCC LFCC
R <sup>2</sup>	0,00	0,00	0,00	0,00
R	0,32	0,32	0,19	0,32
Střed. kv. odchylka	389,91	389,91	412,92	389,91
Střed. abs. odchylka	266,21	266,21	293,67	266,21

Jak je vidět z prvních třech tabulek této podkapitoly, nejvyšší korelace tentokrát nedosahuje příznak SpFe (Spectral Features). Přesto bylo pokračováno v extrakci a zvyšování nOf stejných kombinací příznaků, jako v předchozích 2 podkapitolách, z důvodu porovnatelnosti výsledků.

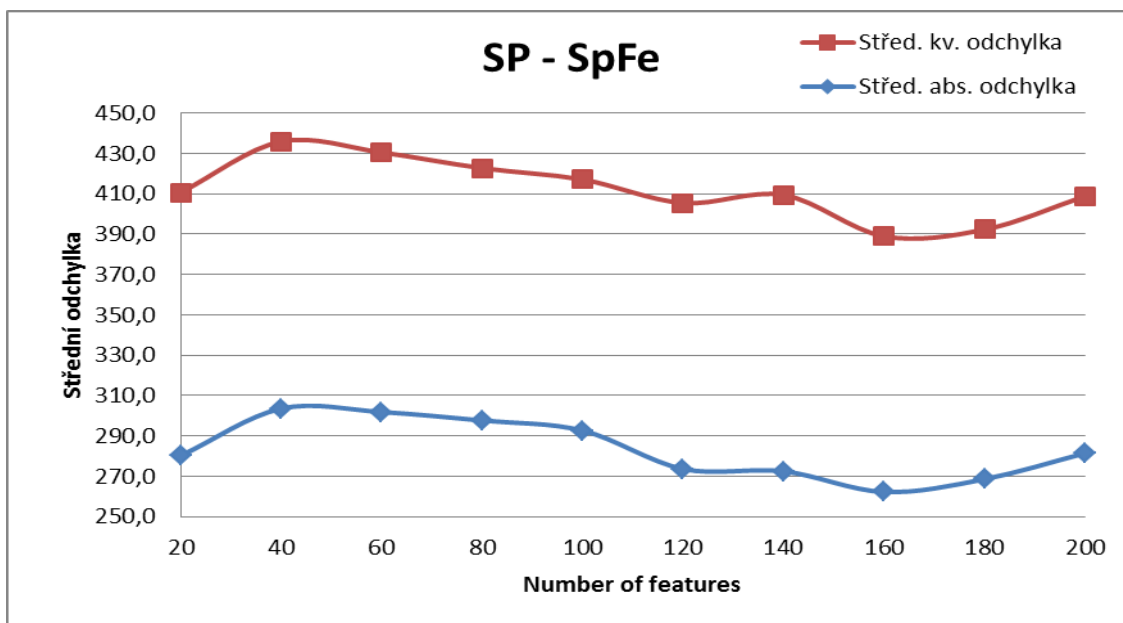
Vypočítané výsledky korelačních koeficientů a jejich střední odchylky pro samostatný příznak SpFe (Spektral koeficient) s rostoucím nOf (number Of features) jsou zaznamenány v tabulce Tab. 5.16 a poté vyneseny do grafů Obr. 5.10 Korelace a Obr. 5.11 Střední odchylky.

Tab. 5.16 SP - Korelace pro příznak SpeFe – zvyšování nOf.

SP [kb/s] SpFe (nOf)	20	40	60	80	100	120	140	160	180	200
R <sup>2</sup>	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
R	0,21	0,14	0,14	0,23	0,17	0,24	0,24	0,29	0,27	0,23
Střed. kv. odchylka	410,6	435,9	430,5	422,7	417,1	405,5	409,4	389,0	392,3	408,8
Střed. abs. odchylka	280,1	303,5	301,8	297,6	292,6	273,6	272,5	262,2	268,6	281,4



Obr. 5.11 SP Korelace pro příznak SpeFe – zvyšování nOf.

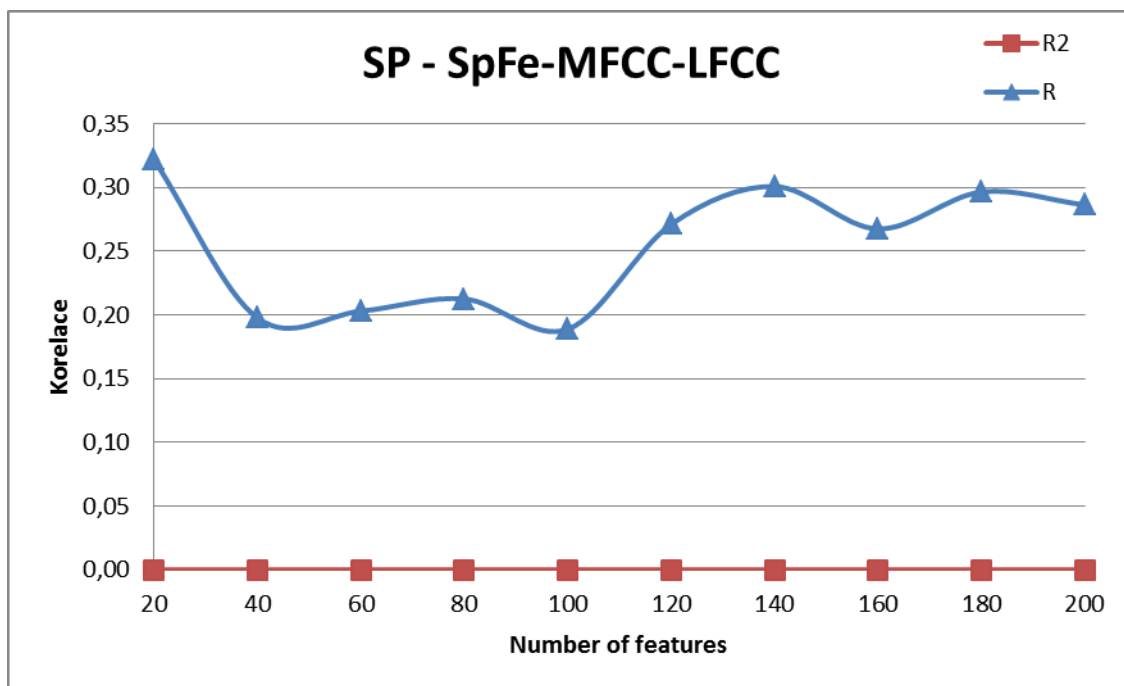


Obr. 5.12 SP - Střední odchylky pro příznak SpeFe

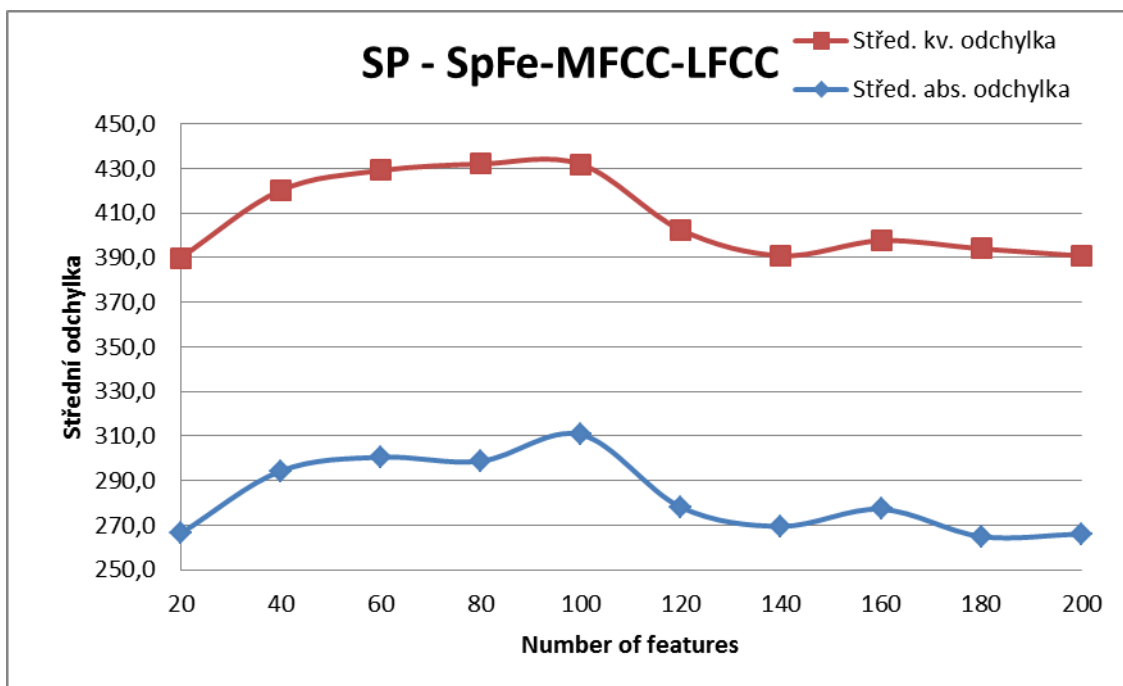
Vypočítané výsledky korelačních koeficientů a jejich střední odchylky pro kombinaci tří příznaků SpFe (Spektral koeficient), MFCC a LFCC s rostoucím nOf (number Of features), jsou zaznamenány v tabulce Tab. 5.17 a poté vyneseny do grafů Obr. 5.12 Korelace a Obr. 5.13 Střední odchylky.

Tab. 5.17 SP Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf.

SP [kb/s] SpFe-MFCC-LFCC (nOf)	20	40	60	80	100	120	140	160	180	200
R <sup>2</sup>	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
R	0,32	0,20	0,20	0,21	0,19	0,27	0,30	0,27	0,30	0,29
Střed. kv. odchylka	389,9	420,2	429,1	432,0	431,9	402,4	390,9	397,6	394,0	390,8
Střed. abs. odchylka	266,2	294,3	300,5	298,7	310,6	278,1	269,5	277,2	264,9	266,1



Obr. 5.13 Korelace pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf.



Obr. 5.14 Střední odchylnka pro kombinaci příznaků SpeFe a MFCC a LFCC – zvyšování nOf

## 6 ZÁVĚR

V práci byly prozkoumány základní řečové vlastnosti. Dále byly prozkoumány přenosové kanály jako je pevná linka – místní spojení, pevná linka – vzdálené spojení, GSM a IP telefonie. Práce se zaměřuje na přenosové vlastnosti a především zvukové charakteristiky kanálu VoIP. Byla nalezena metoda emulace síťových parazit a s použitím emulačního prostředí WANem bylo realizováno síťové zapojení. Díky tomuto zapojení umožňující emulaci vlastností rozsáhlých sítí WAN mohlo být vytvořeno velké množství souvislých nahrávek trénovačích databází. Tyto nahrávky bylo třeba vhodně upravit a rozdělit. K tomu byly v prostředí Matlab vytvořeny skripty zajišťující automatizované dělení promluv na úseky ekvivalentní původní datavabázi.

Byly vytvořeny tři skupiny databází: ZT - ztráty [%], JT – jitter [ms] a SP – šířka (přenosového) pásma [kb/s]. Skupina ZT obsahuje 7 databází, a každá z databází tvoří jinou třídu úrovně ztrát. Dílčí databáze ZT jsou tvořeny vždy 120-ti promluvami. Skupina JT obsahuje 7 databází, a každá z databází tvoří jinou třídu úrovně jitteru. Dílčí databáze JT jsou tvořeny vždy 120-ti promluvami. Poslední skupinou je šířka přenosového pásma. Skupina SP má dvě varianty. V první variantě obsahuje 5 databází, a každá z databází tvoří jinou šířku pásma 96, 128, 256, 512 a 1024 kb/s, dílčí databáze obsahují 118 promluv. Druhá varianta SP obsahuje 6 databází s šířkami pásma 48, 96, 128, 256, 512 a 1024. Dílčí databáze obsahují pouze 79 promluv. Kvůli, nemožnému automatizovanému dělení celistvé nahrávky na dílčí promluvy a díky ztrátám velkých zvukových úseků při šířce přenosového pásma 48 kb/s a 24kb/s. Databáze s šířkou pásma 48 kb/s musela být rozdělena manuálně a získáno bylo pouze 79 promluv.

Z výsledků vyplývá, že k rychlému určení hodnoty jitteru či ztrát z nahrávky dostačuje využití příznaku SpectralFeatures (spektrálních charakteristik). U určení hodnot šířky pásma je zapotřebí vytvořit databáze s šířkou pásma nejvíce 256kb/s. Je však problematické takovéto databáze rozdělit na dílčí promluvy.

Výsledků dosažených v této práci lze aplikovat například při policejním vyšetřování, kdy policie získá neznámou nahrávku, řečovou analýzou určí přenosový kanál a potřebuje rychle zjistit bližší informace o tomto kanálu.

# LITERATURA

- [1] GSM tutorial: GSM Audio Codec / Vocoder. ADRIO COMMUNICATIONS LTD. *Radio-Electronics.com: Cellular telecoms* [online]. [cit. 2013-06-02]. Dostupné z: [http://www.radio-electronics.com/info/cellularcomms/gsm\\_technical/audio-codecs-vocoders-amr-celp.php](http://www.radio-electronics.com/info/cellularcomms/gsm_technical/audio-codecs-vocoders-amr-celp.php).
- [2] SOVA, Lukáš. *VOIP: HLASOVÁ KOMUNIKACE V IP SÍTÍCH, VOIP GSM*. České Budějovice, 2006. Absolventská práce. Soukromá vyšší odborná škola a Obchodní akademie s.r.o. Vedoucí práce PaedDr. Petr Pexa.
- [3] ITU-T. *ITU-T Recommendation G.722: 7 kHz audio-coding within 64 kbit/s*. Geneva: International Telecommunication Union, 2012. Dostupné z: <http://www.itu.int/rec/T-REC-G.722-201209-1/en>.
- [4] ITU-T. *ITU-T Recommendation G.722.2: Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*. Geneva: International Telecommunication Union, 2003. Dostupné z: <http://www.itu.int/rec/T-REC-G.722.2-200307-1/en>.
- [5] SILK Speech Codec. KUNEŠ, Jan, ASTROM, Henrik a VOS, Koen. SKYPE LIMITED. *Skype: Skype developers* [online]. 2010 [cit. 2013-06-01]. Dostupné z: <http://dev.skype.com/silk>.
- [6] XIPH.ORG. *Speex: Speex: A Free Codec For Free Speech* [online]. 2008 [cit. 2013-06-02]. Dostupné z: <http://www.speex.org>.
- [7] WANem: The Wide Area Network emulator. TATA CONSULTANCY SERVICES LTD. WANem [online]. [cit. 2013-05-19]. Dostupné z: <http://wanem.sourceforge.net>.
- [8] THE VIDEOLAN NON-PROFIT ORGANISATION. *VLC media player* [online]. 2001, 2013 [cit. 2013-06-01]. Dostupné z: <http://www.videolan.org>.
- [9] AG PROJECTS. *Blink: A state of the art, easy to use SIP client Available for Mac, Windows and Linux* [online]. 2009, 2013 [cit. 2013-06-01]. Dostupné z: <http://icanblink.com>.
- [10] GOLDWAVE INC. *GoldWave* [online]. 1993, 2013 [cit. 2013-06-01]. Dostupné z: <http://www.goldwave.com>.
- [11] MATHWORKS. *Matlab: The Language of Technical Computing* [online]. 1994, 2013 [cit. 2013-06-01]. Dostupné z: <http://www.mathworks.com/products/matlab>.
- [12] SZIGETI, Tim a Christina HATTINGH. *Quality of Service Design Overview*. [online]. Cisco Press, 17.12.2004 [cit. 2013-05-26]. Dostupné z: <http://www.ciscopress.com/articles/article.asp?p=357102>.
- [13] I. Růžičková, B. Fajmon: *Matematika 3.*, Elektronický text FEKT, Brno 2005. (cs)
- [14] PUŽMANOVÁ, Rita. *Věčné téma: přepojování okruhů či paketů? (1)* [online]. 18. dubna 2006, [cit. 2013-05-26]. Dostupné z: <http://www.svetsiti.cz/clanek.asp?cid=Vecne-tema-prepojovani-okruhu-ci-paketu-1-1842006>.
- [15] BAZALA, David. *Telekomunikace*. 1. vyd. Praha: BEN - technická literatura, 2006, 222 s. ISBN 80-730-0201-9.
- [16] Statistika. SYNEK, V. *UNIVERZITA JANA EVANGELISTY PURKYNĚ V ÚSTÍ NAD LABEM: FAKULTA ŽIVOTNÍHO PROSTŘEDÍ* [online]. 2013 [cit. 2013-06-03]. Dostupné



z: <http://fzp.ujep.cz/~synek/# Skripta>

- [17] Klasifikace: DMINA. In: *EAmos* [online]. 2009 [cit. 2013-06-03]. Dostupné z: [http://eamos.pf.jcu.cz/amos/kat\\_inf/externi/kat\\_inf\\_21586/files/studijni\\_texty](http://eamos.pf.jcu.cz/amos/kat_inf/externi/kat_inf_21586/files/studijni_texty)
- [18] RYCHLÝ, Marek. *Klasifikace a predikce* [online]. Brno, 2005 [cit. 2013-06-03]. Dostupné z: <http://www.fit.vutbr.cz/~rychly/public/docs/classification-and-prediction/xhtml/classification-and-prediction.xhtml>. Nepublikovaná práce. VUT v Brně.

## SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

VoIP	Voice over IP, přenos hlasu po ethernetové síti.
Open-source	Software s otevřenými zdrojovými kódy.
JT	jitter [ms](kolísání zpoždění)
ZT	ztráty [%]
SP	šířka pásma [kb/s]
SpFe	Spectral Features (spektrální funkce)
TpEn	Temporal Energy
TEO	Teager energy operator
nOf	number Of features

# SEZNAM PŘÍLOH

<b>A</b>	<b>Kódy v programu Matlab</b>	<b>42</b>
A.1	Length_of_waves.m.....	42
A.2	orezDB.m.....	43
A.3	PartingDB.m .....	44
<b>B</b>	<b>Databáze nahrávek</b>	<b>46</b>
B.1	ZT - ztráty .....	46
B.2	SP –šířka pásma.....	46
B.3	JT - jitte.....	46
<b>C</b>	<b>Šířka pásma do 256 kb/s</b>	<b>47</b>

# A KÓDY V PROGRAMU MATLAB

## A.1 Length\_of\_waves.m

---

```
%Zjisteni delky skladeb (pocet vzorku na skladbu) a nahrani do pole
nsp
filelist=dir('*.wav');
for i=1:numel(filelist)
[f,fs] = wavread(filelist(i).name);
nsp (i) = (length(f));
times (i) = (length(f)/fs); %vypocet delky skladby v sekandach
end
dbtimeOrig=sum(nsp);
fprintf('Delka originalni databaze je\n%d\n', dbtimeOrig);
%ulozeni poctu vzorku dilcich skladeb do souboru nsp.txt
fid2 = fopen('nsp.txt', 'w');
fprintf(fid2, '%1.0f\n', nsp);
fclose(fid2);

% v pripade potreby ulozeni delky skladeb v sekundach
% fid = fopen('times.txt', 'w');
% fprintf(fid, '%1.7f\n', times);
% fclose(fid);
```

---

## A.2 orezDB.m

.

---

```
% Nacteni jmena databaze s nazvem db131.wav do promene dbname
dbname = 'db131.wav';

%% nacteni databaze db se jmenem ulozenym v dbname
[db, Fs] = wavread(dbname);
dbtimeTemp = length(db);

% urceni zacatku databaze db pomoci detekce urovne ampl
ampl=0.4; %nastaveni detekovane urovne signalu ampl
m = 1;
time = 1;
    while db(m) < ampl
        m=m+1;
    end
    time = m-Fs-36;
    startDB=time;
fprintf('startDB = %d, m=%d\n', startDB,m);

% urceni konce databaze db
z=dbtimeTemp;
    while db(z) < ampl
        z=z-1;
    end
    endDB = z+3*Fs;
fprintf('endDB = %d\n', endDB);

wavwrite(db(startDB:endDB),Fs,dbname);
dbtime= endDB-startDB;
% zaverena kontrola vypsanim nove delky databaze do konzole
matlabu
fprintf('dbtime %d\n', dbtime);
clear;
```

---

## A.3 PartingDB.m

---

```
% nacteni jmena databaze db112.wav do promene dbname
dbname = './db114od22.wav';
koef=0.933;

% standardne nasob cislem 0.933, pracuj pouze s 3-mi desetiny
mistry
% skladba (v cacovem vyseku) ujizdi do leva, zmensuj cislo
% skladba (v cacovem vyseku) ujizdi doprava, zvetsuj cislo

dbtimeOrig = 29607112; % delka originalni databaze (pocet vzorku)
% nacteni nahrane databaze db s jemnem ulozenym v dbname
[db, Fs] = wavread(dbname);
% nacteni delky (pocet vzorku) databaze zpracovavane databaze db
dbtime = length(db);
% nacteni delky dilcich skladeb originalni databze TIMIT
fid2 = fopen('nsp.txt', 'r');
nsp = fscanf(fid2, '%g');
fclose(fid2);

% vypocet rozdilu delky a prumerneho zpozdeni oproti originalu
rozdilDB = (dbtime-dbtimeOrig);
procent = (rozdilDB/dbtime*100);
sec = (rozdilDB/Fs);
avgdely = floor(rozdilDB/length(nsp)*koef);

% vypsani statistik do kozole matlabu
fprintf('\ndbtime %d\n', dbtime);
fprintf('origDB %d\n', dbtimeOrig);
fprintf('    DB %d\n', dbtime);
fprintf('rozdil  %d, %1.2f%%, %1.1fs = %1.1fmin\n', rozdilDB,
procent, sec, sec/60);
fprintf('avgdely  %d, koef  %1.3f \n\n', avgdely, koef);
time = 1;

% rozdeleni databaze na dilci skladby
name = 001;
for i=1:length(nsp)
    tstart = time;
```

```
tend = time + nsp (i);
    wavwrite(db(tstart:tend),Fs,num2str(name,'%0.3i'));
time = tend + avgdely;
name = name + 1;
end

% roztrideni tozsekanych skladeb na na nahravky a pipnuti
mkdir rozsekane;
movefile('*.wav', 'rozsekane');
cd rozsekane;
movefile(dbname, '../');
mkdir 'pip';
mkdir 'db';
movefile('*2.wav', 'db');
movefile('*4.wav', 'db');
movefile('*6.wav', 'db');
movefile('*8.wav', 'db');
movefile('*0.wav', 'db');
movefile('*1.wav', 'pip');
movefile('*3.wav', 'pip');
movefile('*5.wav', 'pip');
movefile('*7.wav', 'pip');
movefile('*9.wav', 'pip');
cd ../;
beep;
clear;
```

---

## **B DATABÁZE NAHRÁVEK**

Příložené CD obsahuje vytvořenou databázi nahrávek.

### **B.1 ZT - ztráty**

### **B.2 SP –šířka pásma**

### **B.3 JT - jítte**

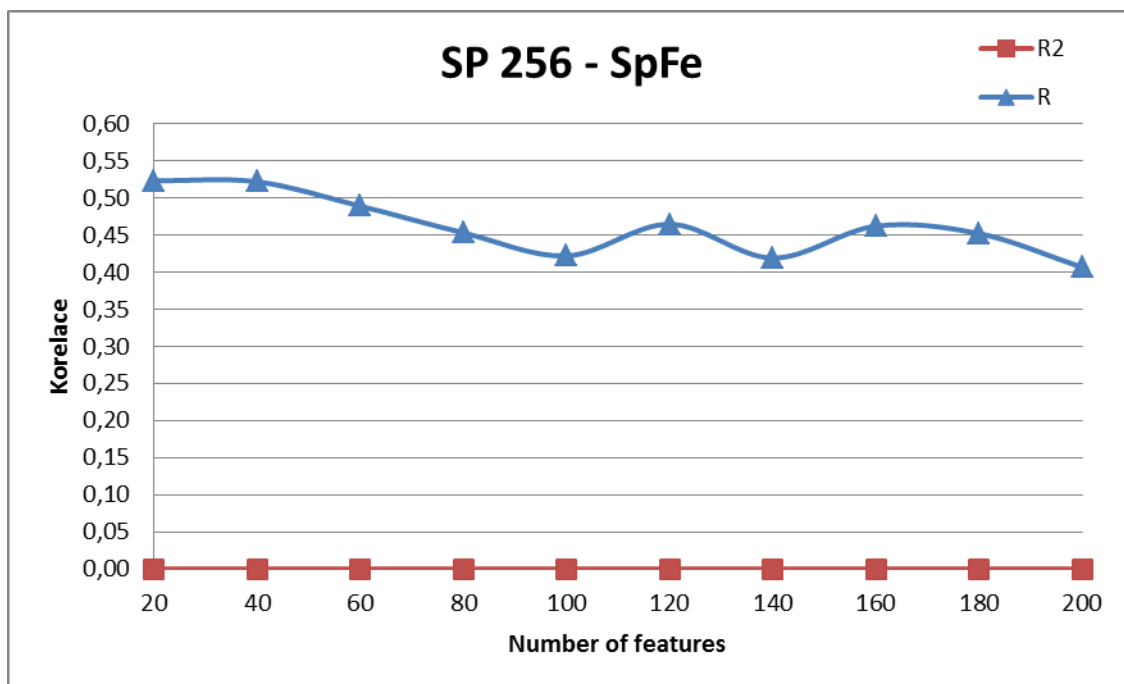


## C ŠÍŘKA PÁSMO DO 256 KB/S

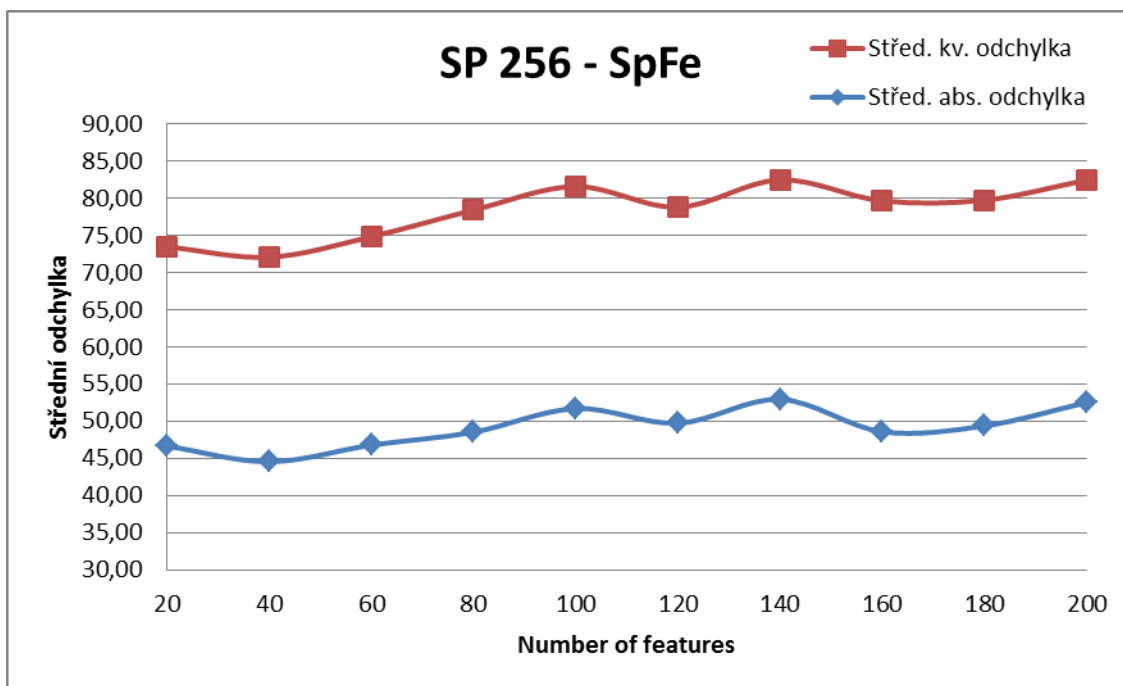
Tab. C.1 Korelace SpectralFeatures pro Šířku pásma do 256 kb/s

SP 256 [kb/s] SpFe (nOf)	20	40	60	80	100	120	140	160	180	200
R <sup>2</sup>	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
R	0,52	0,52	0,49	0,45	0,42	0,46	0,42	0,46	0,45	0,41
Střed. kv. odchylka	73,48	72,07	74,82	78,41	81,58	78,80	82,47	79,64	79,74	82,45
Střed. abs. odchylka	46,65	44,55	46,78	48,52	51,70	49,73	52,94	48,58	49,37	52,53

..



Obr. C.6.1 Korelace příznaku SpectralFeatures pro Šířku pásma do 256 kb/s



Obr. C.6.2 Střední odchylnky příznaku SpectralFeatures pro šířku pásma do 256 kb/s