

Diplomová práce



České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra kybernetiky

Analýza variability srdečního rytmu

Tomáš Grosman

Květen 2016

Vedoucí práce: Ing. Jakub Parák

České vysoké učení technické v Praze
Fakulta elektrotechnická

Katedra kybernetiky

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: Bc. Tomáš Grosman

Studijní program: Biomedicínské inženýrství a informatika

Obor: Biomedicínská informatika

Název tématu: Analýza variability srdečního rytmu

Pokyny pro vypracování:

1. Seznamte se s metodikou analýzy stacionarity signálů a HRV analýzou.
2. Naprogramujte vhodný algoritmus pro analýzu HRV. Zkuste použít autoregresní a kvestrální model a optimalizujte jejich parametry.
3. Otestujte daný algoritmus pro různé parametry na RR a ACC záznamech, otestujte stacionaritu jednotlivých segmentů. Na rozdělení segmentů použijte prahování.
4. Ověřte závislosti mezi stacionaritou RR a ACC záznamů.
5. Diskutujte možnost odhadu stacionarity v reálném čase v závislosti na záznamu z akcelerometru.

Seznam odborné literatury:

- [1] A. N. Kalinichenko, M. I. Nilicheva, S. V. Klasheva, O. D. Yourieva, O. V. Mamontov: Signal Stationarity Assessment for the Heart Rate Variability Spectral Analysis, ISSN 0276-6574, Computers in Cardiology 2008; 35:965-968
- [2] J. Aitchison, I. R. Dunsmore: Statistical Prediction Analysis, ISBN-13: 978-0521298582, Cambridge University Press, New York, 1980
- [3] S. W. Smith: The Scientists and Engineer's Guide to Digital Signal Processing. California Technical Publishing, 1999. <http://www.dspguide.com>
- [4] A. V. Oppenheim, R.W. Schaffer: Discrete-Time Signal Processing. Prentice-Hall, 2009, ISBN-13: 978-0131988422
- [5] M. H. Hayes: Statistical digital signal processing and modeling, John Wiley, 1996

Vedoucí diplomové práce: Ing. Jakub Parák

Platnost zadání: do konce letního semestru 2016/2017

L.S.

prof. Dr. Ing. Jan Kybic
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 14. 12. 2015

Poděkování / Prohlášení

Za odborné vedení, cenné rady a doporučení bych na tomto místě rád poděkoval panu Ing. Jakubu Parákovi, který byl vedoucím mé diplomové práce.

Dále bych chtěl také poděkovat panu prof. Ing. Pavlu Sovkovi, CSc. za užitečné rady a konzultace ohledně diplomové práce.

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne

.....
Podpis autora práce

Abstrakt / Abstract

Tato práce je zaměřena na vytvoření algoritmu pro porovnání závislosti stacionárních úseků v záznamech RR intervalů se stacionárními úseky záznamů z akcelerometru. Tento problém byl řešen pomocí vypočtení autoregresních koeficientů v jednotlivých segmentech záznamů. Autoregresní koeficienty byly převedeny na kepstrální koeficienty, ze kterých byla vypočtena kepstrální vzdálenost. Stacionární úseky v těchto záznamech byly určeny v místech, kde byla nízká kepstrální vzdálenost. U těchto úseků musela být potvrzena stacionarita pomocí vhodného statistického testu. V této práci byla potvrzována stacionarita pomocí znaménkového testu nebo ANOVA testu. Výsledky ukazují, že kepstrální vzdálenost je v podobných úsecích nízká. Výsledky nepotvrzují, že by existovala přímá závislost mezi stacionárními úseky v obou typech dat, díky tomu, že zmíněné testy na stacionaritu nepotvrzují úseky ve stejných místech. Pro možnost odzkoušení vytvořených metod na vlastních datech byla vytvořena grafická aplikace.

Klíčová slova: variabilita srdečního rytmu, RR interval, akcelerometr, záznam signálu, AR model, kepstrum, kepstrální vzdálenost, ANOVA, znaménkový test, stacionarita

This thesis is focused on creating an algorithm for comparing dependency between stationary sections in the records of RR intervals and stationary sections of the records from the accelerometer. This problem was solved with computing of the autoregressive coefficients in each segment of the records. Autoregressive coefficients were converted to cepstral coefficients, from which was counted cepstral distance. Stationary sections in these records were identified in places, where was low cepstral distance. In this sections had to be confirmed stationarity by suitable statistical test. In this thesis was stationarity confirmed with sign test or ANOVA test. The results show, that cepstral distance is low in the similar sections. The results not confirm, that exists direct dependence between stationary parts in both data types, due to the fact, that the mentioned tests for the stationarity don't confirm section in same parts. For the possibility of testing created methods on own data was created graphical application.

Keywords: heart rate variability, RR interval, accelerometer, signal record, AR model, cepstrum, cepstral distance, ANOVA, sign test, stationarity

Obsah /

Úvod	1
1 Srdeční rytmus a stacionarita signálu	2
1.1 Srdeční tep	2
1.2 Analýza srdečního rytmu	3
1.2.1 Metody v časové oblasti ...	4
1.2.2 Metody ve frekvenční oblasti	7
1.3 Stacionarita signálu	9
1.4 Analýza stacionarity signálu ...	9
1.4.1 Znaménkový test	9
1.4.2 Analýza rozptylu spojená s Bartlettovým testem	9
1.4.3 AR model	10
1.4.4 Použití AR modelu	11
1.4.5 Képstrum	12
1.4.6 Převod koeficientů	12
1.4.7 Použití stacionarity signálu	13
2 Cíle práce	14
3 Návrh řešení	15
4 Realizace řešení	17
4.1 Ověření převodu koeficientů ...	17
4.1.1 Převod z 1. řádu AR modelu na 3. řád kepstra	17
4.1.2 Převod z 2. řádu AR modelu na 4. řád kepstra	18
4.1.3 Převod z 3. řádu AR modelu na 5. řád kepstra	19
4.2 Ověření kepstrální vzdálenosti	20
4.2.1 AR model 1. řádu	20
4.2.2 AR model 3. řádu	21
4.2.3 AR model 5. řádu	23
4.3 Nahrávání záznamu	25
4.4 Rozdělení dat záznamů	26
4.5 Porovnání vzdáleností	28
4.6 Označení možných stacionárních úseků	30
4.7 Potvrzení stacionarity v úsecích	31
4.7.1 První záznam s manuálními úseky	31
4.7.2 Druhý záznam s manuálními úseky	33
4.8 Porovnání úseků ze dvou typů dat	34
4.9 Grafická aplikace	35
5 Výsledky	37
5.1 Statický práh	38
5.1.1 Nižší práh	38
5.1.2 Vyšší práh	41
5.2 Dynamický práh	44
5.2.1 Nižší práh	45
5.2.2 Vyšší práh	47
5.3 Uložení dalších výsledků	50
6 Zhodnocení výsledků	51
7 Diskuze	52
Závěr	54
Literatura	55
A Seznam symbolů a zkratk	57
B Obsah příloženého CD	59
C Histogramy při jednotlivých aktivitách	61

Tabulky / Obrázky

1.1. Přehled parametrů získaných ze statistických metod v časové oblasti	4	1.1. EKG křivka	3
1.2. Přehled parametrů získaných z geometrických metod v časové oblasti	6	1.2. RR interval	4
1.3. Přehled parametrů ve frekvenční oblasti u krátkodobých záznamů	7	1.3. Rozdělení hustoty pravděpodobnosti	6
1.4. Přehled parametrů ve frekvenční oblasti u dlouhodobých záznamů	8	1.4. Diagram procesu analýzy srdečního rytmu	8
4.1. Tabulka výsledků znaménkových testů u prvního záznamu .	33	4.1. 1. frekvenční charakteristika AR modelů	20
4.2. Tabulka výsledků ANOVA testů u prvního záznamu	33	4.2. 1. filtrovaný signál.....	21
4.3. Tabulka výsledků znaménkových testů u druhého záznamu	34	4.3. 1. kepstrální vzdálenost	21
4.4. Tabulka výsledků ANOVA testů u druhého záznamu	34	4.4. 2. frekvenční charakteristika AR modelů	22
5.1. Tabulka prvních výsledků na 1 záznamu	40	4.5. 2. filtrovaný signál.....	22
5.2. Tabulka prvních výsledků na všech záznamech	40	4.6. 2. kepstrální vzdálenost	23
5.3. Tabulka druhých výsledků na 1 záznamu.....	41	4.7. 3. frekvenční charakteristika AR modelů	23
5.4. Tabulka druhých výsledků na všech záznamech	41	4.8. 3. filtrovaný signál.....	24
5.5. Tabulka třetích výsledků na 1 záznamu	43	4.9. 3. kepstrální vzdálenost	24
5.6. Tabulka třetích výsledků na všech záznamech	43	4.10. Záznam jednoho měření	26
5.7. Tabulka čtvrtých výsledků na 1 záznamu.....	44	4.11. Histogram RR intervalů	26
5.8. Tabulka čtvrtých výsledků na všech záznamech	44	4.12. Histogram RR intervalů po interpolaci	27
5.9. Tabulka pátých výsledků na 1 záznamu	46	4.13. Histogram všech os z akcelerometru	27
5.10. Tabulka pátých výsledků na všech záznamech	46	4.14. Histogram všech os dohromady z akcelerometru.....	28
5.11. Tabulka šestých výsledků na 1 záznamu	47	4.15. Vzdálenost koeficientů AR modelu.....	29
5.12. Tabulka šestých výsledků na všech záznamech	47	4.16. Kepstrální vzdálenost 1	29
		4.17. Kepstrální vzdálenost 2	30
		4.18. První záznam s manuálně vybranými úseky	32
		4.19. Druhý záznam s manuálně vybranými úseky	33
		4.20. Grafická aplikace	35
		4.21. Běh grafické aplikace.....	36
		5.1. Záznam použitého měření	37
		5.2. První nalezené úseky.....	39
		5.3. První potvrzené úseky	39
		5.4. Druhé nalezené úseky	40
		5.5. Druhé potvrzené úseky	41
		5.6. Třetí nalezené úseky	42
		5.7. Třetí potvrzené úseky	42
		5.8. Čtvrté nalezené úseky.....	43
		5.9. Čtvrté potvrzené úseky	44
		5.10. Páté nalezené úseky	45

5.13. Tabulka sedmých výsledků na 1 záznamu	49	5.11. Páté potvrzené úseky	45
5.14. Tabulka sedmých výsledků na všech záznamech	49	5.12. Šesté nalezené úseky	46
5.15. Tabulka osmých výsledků na 1 záznamu	50	5.13. Šesté potvrzené úseky	47
5.16. Tabulka osmých výsledků na všech záznamech	50	5.14. Sedmé nalezené úseky	48
B.1. Přehled souborů a složek na přiloženém CD	59	5.15. Sedmé potvrzené úseky	48
		5.16. Osmé nalezené úseky	49
		5.17. Osmé potvrzené úseky	50
		C.1. Příloha histogram 1	61
		C.2. Příloha histogram 2	62
		C.3. Příloha histogram 3	62
		C.4. Příloha histogram 4	63
		C.5. Příloha histogram 5	63
		C.6. Příloha histogram 6	64



Úvod

Tématem této práce je analýza variability srdečního rytmu. Tato analýza je zaměřena na vytvoření algoritmu pro nalezení stacionárních úseků v signálu, který monitoruje srdeční tep, a tyto stacionární úseky jsou porovnány se signály, které monitorují pohyb. Pro tento účel byly pořízeny dlouhodobé 24hodinové záznamy, které monitorovaly oba typy zmíněných signálů synchronně. Celý algoritmus se opírá o autoregresní analýzu a kepstrem signálu.

V první části této práce je popsán srdeční tep a možnosti analýzy variability srdečního rytmu. Dále je popsána stacionarita signálu a její zjištění a jak je tato stacionarita spojená s autoregresní analýzou a kepstrem signálu.

V dalších částech poté jsou popsány podrobně všechny použité metody, které vedly ke konečné analýze. Kromě popsání těchto metod jsou také tyto metody odzkoušeny na jednodušších datech, aby se dala ověřit jejich funkčnost.

Aby bylo možné metody použít efektivně, tak bylo vytvořeno jednoduché grafické rozhraní, ve kterém je možné nastavovat jednotlivé parametry. Funkčnost této aplikace je zobrazena za vysvětlením použitých metod.

Po představení grafické aplikace jsou zobrazeny jednotlivé výsledky pro různá nastavení parametrů a z těchto výsledků jsou poté vyvozeny určité závěry a možnosti rozšíření do budoucna.

Všechny použité algoritmy byly naprogramovány ve vývojovém prostředí MATLAB, které je zároveň i skriptovacím programovacím jazykem. Konkrétně byl použit MATLAB verze 2015b.

Kapitola 1

Srdeční rytmus a stacionarita signálu

Tato kapitola se dá rozdělit na dvě části. V první části se tato kapitola věnuje srdečnímu tepu a způsobům analýzy srdečního rytmu. Ve druhé části se tato kapitola věnuje stacionaritě signálu a metodám zjištění stacionarity signálu.

1.1 Srdeční tep

Srdce je důležitý orgán, který funguje jako pumpa, která pohání krev pomocí cév do jednotlivých částí těla. Srdce se skládá ze 4 dutin. Jedná se o 2 síně a 2 komory. Jednotlivé srdeční dutiny se při srdečním tepu mohou nacházet ve dvou různých stavech. Jedná se o systolický a diastolický stav.

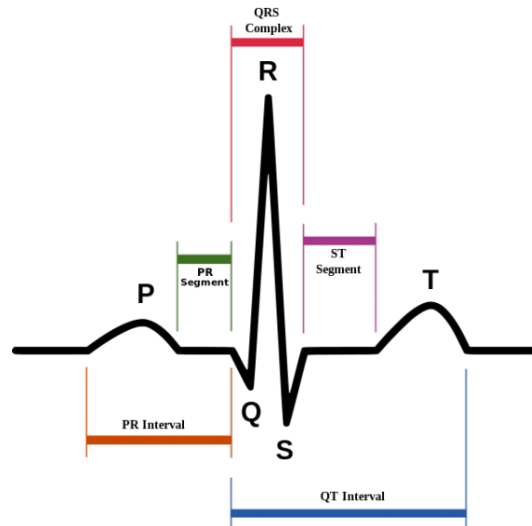
Při systolickém stavu se srdce stlačí a krev se z komor dostává do krevního oběhu. Při diastolickém stavu naopak dochází k relaxaci srdečního svalu. Během této relaxační fáze se plní síně a komory krví.

Samotný srdeční cyklus se skládá ze tří fází [1]. Konkrétně se jedná o systolu síní, systolu komor a relaxační fázi.

Během systoly síní se ze síní po otevření chlopní vypustí krev do komor. Během systoly komor se otevřou chlopně komor a krev se vypustí do velkého a malého krevního oběhu. Z pravé komory se vypustí krev do malého krevního oběhu, který vede přes plíce, z levé komory se vypustí krev do velkého krevního oběhu, který vede do zbytku těla. Při relaxační fázi se do všech 4 dutin dostává krev z žil. Komory se zaplní ze 75 % během této fáze. Doplnění komor dochází při fázi systoly síní.

Normální srdeční frekvence u člověka se pohybuje přibližně v rozmezí 50 - 70 úderů za minutu [2], pokud je člověk v klidu. Srdeční frekvence je závislá na věku a také na sportovní výdrž člověka. Čím je člověk trénovanější, tím by měl mít klidovou srdeční frekvenci nižší. Srdeční frekvence je řízena sinoatriálním uzlem, který reaguje na aktivitu nervů ze sympatika a parasympatika. Sympatikus zvyšuje srdeční frekvenci, zatímco parasympatikus snižuje srdeční frekvenci.

Záznam srdeční aktivity se jmenuje elektrokardiogram. Samotná metoda záznamu srdeční aktivity se nazývá elektrokardiografie (dále EKG). Při jednom srdečním tepu se postupně v EKG záznamu nachází P vlna, QRS komplex, ST segment a vlna T, viz obrázek 1.1. V následující části jsou popsány tyto jednotlivé vlny a kmity (zdroj [3]).



Obrázek 1.1. EKG křivka, zdroj [4]

Při vlně P dochází k depolarizaci obou síní. Tato vlna obvykle trvá cca 80 ms. Při této vlně dochází k šíření elektrického impulsu z SA uzlu v pravé síni až k síňokomorové přepážce. Zároveň se elektrický impuls šíří Brachmanovým svazkem do levé síně.

QRS komplex je složen z kmitů Q, R a S. Při tomto komplexu dochází k šíření akčního potenciálu v srdečních komorách. Tento komplex trvá přibližně 80 až 120 ms. Elektrický impuls se šíří zleva doprava mezikomorovým septem. Po aktivaci septa dochází k aktivaci v levé komoře. Aktivaci pravé komory lze vidět na EKG křivce pouze za některých podmínek.

Za QRS komplexem se nachází ST segment, kde dochází k depolarizaci komor. ST segment je následován vlnou T, která vyjadřuje repolarizaci komor. Tato vlna trvá obvykle maximálně 200 ms.

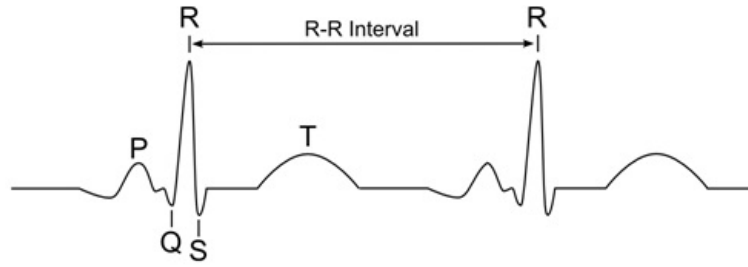
1.2 Analýza srdečního rytmu

Analýza srdečního rytmu se zabývá tím, jestli je srdeční rytmus správný. Tedy hlavní je, jestli není příliš odchýlena tepová frekvence od normálních hodnot. Existují dvě možné arytmie, konkrétně se jedná o bradyarytmii a tachyarytmii [5].

Bradyarytmie znamená pomalý srdeční rytmus. Existují různé typy bradyarytmií. Tento problém je nejhorší při zátěži, protože pokud není schopno srdce zrychlit svůj rytmus při vyšší zátěži, tak se pak člověku může udělat nevolno. V takovém případě může člověk také trpět vyšší unaveností, pokud musí provádět vyšší fyzickou zátěž a při velmi nízkém srdečním tepu i bez této zátěže. Dokonce může docházet občas i ke ztrátě vědomí.

Naopak tachyarytmie znamená zrychlený srdeční rytmus. Častým příznakem u lidí trpících tímto problémem je příliš rychlé bušení srdce. Většinou se jedná o problém vzniku extrasystoly. Jedná se o předčasně přicházející stahy z míst abnormální tvorby vzruchu v jednotlivých dutinách srdce. Extrasystoly mohou jít rychle za sebou. V takovém případě jsou pocíťovány jako nepravidelné bušení srdce. Při této poruše často mají lidé pocit tísně, kratšího dechu. Může ale také docházet ke krátkodobé ztrátě vědomí.

Základem analýzy srdečního rytmu jsou vzdálenosti RR intervalů (obrázek 1.2). Jedná se o časový rozdíl mezi jednotlivými špičkami vln R v EKG signálu. Občas tento interval bývá označován také jako NN interval.



Obrázek 1.2. RR interval, zdroj [6]

Srdeční rytmus se dá analyzovat jak v časové, tak ve frekvenční oblasti. Kde se v časové oblasti dají metody dále dělit na statistické a geometrické metody. Ve frekvenční oblasti záleží na délce záznamu. Zde se metody dělí na vlastnosti vypočtené v krátkých záznamech a na vlastnosti vypočtené v dlouhých záznamech. (viz [7])

1.2.1 Metody v časové oblasti

Z delších časových nahrávek srdeční aktivity (24 hod. měření) se dají vypočítat parametry na základě komplexních statistických metod v časové oblasti. Všechny metody vycházejí z RR intervalu. Získávané parametry jsou vypsány v tabulce 1.1.

Název	Jednotka	Popis
<i>SDNN</i>	ms	směrodatná odchylka RR intervalů
<i>RMSSD</i>	ms	druhá odmocnina průměrného rozdílu přilehlých RR intervalů
<i>SDANN</i>	ms	směrodatná odchylka průměrných RR intervalů vypočítaných v segmentu určité délky, např. 5 min dlouhý segment
<i>SDNN index</i>	ms	průměr všech směrodatných odchylek RR intervalů ze všech segmentů
<i>SDSD</i>	ms	směrodatná odchylka rozdílu mezi přilehlými RR intervaly
<i>NN50</i>		počet RR intervalů, které se oproti následujícím liší o 50 a více ms
<i>pNN50</i>	%	počet NN50 intervalů vydělený celkovým počtem RR intervalů v záznamu

Tabulka 1.1. Přehled parametrů získaných ze statistických metod v časové oblasti

SDNN se počítá následovně (zdroj [8]):

$$SDNN = \sqrt{\frac{1}{N-1} \sum_{n=1}^N [RR(n) - \overline{RR}]^2} \quad [ms] \quad (1)$$

kde N je počet RR intervalů, RR znamená RR interval a \overline{RR} je průměr RR intervalů počítaný jako:

$$\overline{RR} = \frac{1}{N} \sum_{n=1}^N RR(n) \quad [ms] \quad (2)$$

RMSSD se počítá následovně (zdroj [9]):

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{n=2}^N [RR(n) - RR(n-1)]^2} \quad [ms] \quad (3)$$

kde N je počet RR intervalů a RR je čas RR intervalu.

SDANN se počítá následovně (zdroj [10]):

$$SDANN = \sqrt{\frac{1}{N-1} \sum_{n=1}^N [MRR(n) - \overline{MRR}]^2} \quad [ms] \quad (4)$$

kde N je počet segmentů, do kterých je signál rozdělen, MRR je průměr RR intervalů v daném segmentu a \overline{MRR} je průměr průměrů RR intervalů ze všech segmentů.

SDNN index se počítá následovně (zdroj [10]):

$$SDNN \text{ index} = \frac{1}{N} \sum_{n=1}^N STDRR(n) \quad [ms] \quad (5)$$

kde N je počet segmentů, do kterých byl signál rozdělen a $STDRR$ je směrodatná odchylka RR intervalů v daném segmentu, ta se počítá z rovnice pro SDNN (1).

SDSD se počítá následovně:

$$SDSD = \sqrt{\frac{1}{N-1} \sum_{n=1}^N [DRR(n) - \overline{DRR}]^2} \quad [ms] \quad (6)$$

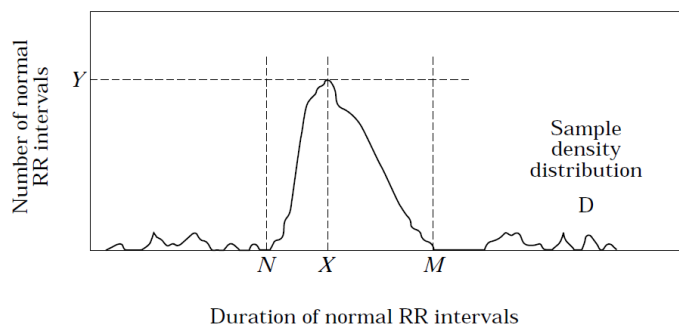
kde N je počet rozdílů RR intervalů, DRR je rozdíl 2 RR intervalů a \overline{DRR} je průměr všech rozdílů po sobě jdoucích RR intervalů. U výpočtu tohoto parametru je důležité zvolit, jakým způsobem se bude rozdíl dvou RR intervalů počítat. Je možnost buď vypočítat absolutní hodnoty nebo vypočítat odečet předchozího od následujícího, či obráceně. O tento fakt výpočtu rozdílu RR intervalů se opírá také počítání NN50 a pNN50 parametrů.

Pro geometrické metody je dobré použít geometrický vzor, ze kterého vycházejí, jako např. hustotu pravděpodobnosti rozložení RR intervalů. Pro hustotu rozdělení je potřeba vybrat velikost pro každé pásmo histogramu, do kterého v rozložení spadají jednotlivé RR intervaly. Dobrá velikost jednoho pásma pro histogram u RR intervalů je přibližně 8 ms. Parametry získané pomocí geometrických metod jsou popsány v tabulce 1.2.

Název	Jednotka	Popis
<i>HRV trojúhelníkový index</i>		Celkový počet RR intervalů, který je vydělen velikostí nejvyššího pásma v histogramu
<i>TINN</i>	ms	základní šířka při aproximaci trojúhelníkové distribuce RR intervalů vycházející z histogramu
<i>Index diferencí</i>	ms	rozdíl mezi šířkami histogramů u přilehlých RR intervalů v závislosti na parametrech
<i>Logaritmický index</i>		koeficient φ z negativní exponenciální křivky $k \cdot e^{\varphi t}$ z absolutních rozdílů přilehlých RR intervalů

Tabulka 1.2. Přehled parametrů získaných z geometrických metod v časové oblasti

Z popsaných parametrů jsou důležité a používané *HRV trojúhelníkový index* a *TINN*. Oba parametry vychází z rozdělení hustoty pravděpodobnosti, která se aproximuje histogramem. Aproximace distribuční funkce vypadá tak, jak je znázorněno na obrázku 1.3.



Obrázek 1.3. Rozdělení hustoty pravděpodobnosti, zdroj [7]

HRV trojúhelníkový index se dá z histogramu vypočítat jako:

$$HRVti = \frac{CRR}{Y} \quad (7)$$

kde $HRVti$ je *HRV trojúhelníkový index*, CRR je celkový počet RR intervalů a Y vychází z obrázku 1.3.

Pro *TINN* se stanoví místa M a N z obrázku 1.3 a z těch se vytvoří funkce $q(t)$ tak, že $q(t) = 0$ pro $t \leq N$ a $t \geq M$ a $q(X) = Y$ a integrál

$$\int_0^{\infty} (D(t) - q(t))^2 dt \quad (8)$$

je minimální přes všechny výběry hodnot N a M . *TINN* se udává v ms a vychází jako hodnota $TINN = M - N$.

1.2.2 Metody ve frekvenční oblasti

Metody ve frekvenční oblasti vycházejí ze spektrální výkonové hustoty, která se získá z vypočteného spektra signálu. Spektrální výkonová hustota se stejně jako předchozí metody provádí na záznamu RR intervalů. U metod v časové oblasti nebylo potřeba provádět žádnou změnu v záznamu RR intervalů. Pro spektrum je potřeba provádět interpolaci před počítáním vlastností, protože záznam RR intervalů není vzorkován pravidelně. Tedy vzorkovací perioda není konstantní. Pro práci se spektrem je potřeba mít časovou vzdálenost jednotlivých vzorků stejnou.

Spektrální výkonová hustota se dá spočítat pomocí parametrických nebo neparametrických metod. Parametrické metody vycházejí ze získání parametrů autoregresního (AR) modelu, modelu klouzavých součtů (MA) nebo autoregresního modelu klouzavých součtů (ARMA). Základem neparametrických metod je výpočet spektra pomocí fourierovy transformace. Parametrické modely mají hladší spektrum, oproti tomu jsou ale časově náročnější než neparametrické metody.

Ze spektrální výkonové hustoty se poté zjišťují jednotlivé parametry. Vypočítané parametry jsou závislé na tom, jak dlouhý je záznam RR intervalů, ze kterých se zjišťují. Pro krátké záznamy se zjišťují rozdílné parametry než pro dlouhotrvající záznamy.

Krátkodobými záznamy jsou myšleny záznamy trvající přibližně od 2 do 5 minut. Parametry zjišťované z krátkodobých záznamů jsou popsány v tabulce 1.3.

Název	Jednotka	Popis
5 min. celkový výkon	ms ²	rozptyl RR intervalů v dočasném segmentu
<i>VLF</i>	ms ²	výkon ve velmi nízkém frekvenčním pásmu
<i>LF</i>	ms ²	výkon v nízkém frekvenčním pásmu
<i>LF norm</i>	n. u.	normovaný <i>LF</i> výkon
<i>HF</i>	ms ²	výkon ve vyšším frekvenčním pásmu
<i>HF norm</i>	n. u.	normovaný <i>HF</i> výkon
<i>LF/HF</i>		poměr <i>LF</i> vůči <i>HF</i>

Tabulka 1.3. Přehled metod ve frekvenční oblasti u krátkodobých záznamů, zkratka n. u. značí normovanou jednotku

Celkový výkon se počítá ve frekvenčním pásmu přibližně do 0,4 Hz. U *VLF* se počítá výkon ve frekvenčním pásmu do 0,04 Hz. U *LF* se počítá výkon ve frekvenčním pásmu od 0,04 Hz do 0,15 Hz. U *HF* se počítá výkon ve frekvenčním pásmu od 0,15 Hz do 0,4 Hz.

Normovaná hodnota *LF* se počítá následovně:

$$LF \text{ norm} = \frac{LF}{(TP - VLF)} \cdot 100 \quad [n.u.] \quad (9)$$

kde *TP* je celkový výkon ve spektru ve vstupním záznamu.

Normovaná hodnota *HF* se počítá následovně:

$$HF \text{ norm} = \frac{HF}{(TP - VLF)} \cdot 100 \quad [n.u.] \quad (10)$$

Poměr LF vůči HF se počítá následovně:

$$LF/HF = \frac{LF}{HF} \quad (11)$$

tato rovnice je bezrozměrná, tedy výsledek nemá žádnou jednotku.

Dlouhodobé záznamy mají většinou délku přibližně 24 hodin. Parametry zjišťované z dlouhodobých záznamů jsou popsány v tabulce 1.4.

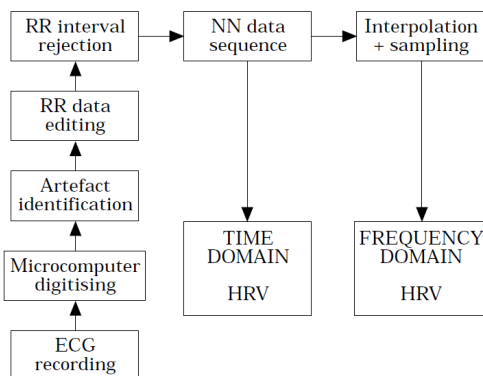
Název	Jednotka	Popis
Celkový výkon	ms^2	rozptyl všech RR intervalů
ULF	ms^2	výkon v ultra nízkém frekvenčním pásmu
VLF	ms^2	výkon ve velmi nízkém frekvenčním pásmu
LF	ms^2	výkon v nízkém frekvenčním pásmu
HF	ms^2	výkon ve vyšším frekvenčním pásmu
a		sklon lineární interpolace spektra v log-log škále

Tabulka 1.4. Přehled metod ve frekvenční oblasti u dlouhodobých záznamů

Celkový výkon se znovu počítá ve frekvenčním pásmu do 0,4 Hz. U ULF se počítá výkon ve frekvenčním pásmu do 0,003 Hz. U VLF se počítá výkon ve frekvenčním pásmu od 0,003 Hz do 0,04 Hz. U LF se počítá výkon ve frekvenčním pásmu od 0,04 Hz do 0,15 Hz. U HF se počítá výkon ve frekvenčním pásmu od 0,15 Hz do 0,4 Hz. Parametr a se počítá ve frekvenčním pásmu do 0,04 Hz.

U dlouhodobých záznamů se často zjišťují parametry ve frekvenční oblasti tak, že se signál rozdělí na několik úseků (např. 5 minutových) a v každém z těchto úseků se zjišťují tyto parametry zvlášť. Po získání parametrů ze všech těchto úseků se z těchto hodnot získá průměr. Tento proces se provádí, protože 24 hodinové záznamy nejsou po celou dobu stacionární.

Celý proces, jak probíhá analýza srdečního rytmu, je zobrazen diagramem na obrázku 1.4.



Obrázek 1.4. Diagram procesu analýzy srdečního rytmu, zdroj [7]

Na diagramu (obrázek 1.4) je nejdůležitější část, která ukazuje, že před provedením spektrální analýzy je potřeba signál interpolovat. Při interpolaci je důležité určit vzorkovací frekvenci RR intervalů. Díky tomu, že záznam RR intervalů má důležité frekvence do 0,4 Hz, tak není potřeba volit nějakou vysokou vzorkovací frekvenci. Takže může stačit vzorkovací frekvence od 1 do 5 Hz.

1.3 Stacionarita signálu

Existují dvě různé definice stacionarity signálu, striktní stacionarita a stacionarita druhého řádu (někdy označována jako slabší stacionarita)[11]. Striktně stacionární signál má vlastnosti, že se jeho hustota pravděpodobnosti nemění v čase a tím pádem se nemění ani žádné jeho statistické vlastnosti. Jelikož se jedná o příliš přísnou definici stacionarity, tak se používá hlavně stacionarita druhého řádu. Pro stacionaritu druhého řádu platí, že v čase se nemění pouze střední hodnota, směrodatná odchylka a autokorelace signálu. U autokorelace záleží pouze na tom, o kolik vzorků vůči sobě signál posouváme. V celé práci se tedy stacionárním signálem rozumí signál, který splňuje definici stacionarity druhého řádu.

Signály můžeme dělit na takové, které jsou celé stacionární, stacionární po částech nebo nestacionární. Celý stacionární signál znamená, že po celou dobu trvání má neměnnou střední hodnotu, směrodatnou odchylku a autokorelaci. U signálu, který je stacionární po částech nalezneme určité úseky významné délky, které jsou celé stacionární, ale zbytek signálu je nestacionární. Nestacionární signál nespĺňuje výše popsané vlastnosti po celou dobu svého trvání.

1.4 Analýza stacionarity signálu

Stacionaritu signálu je možné testovat několika způsoby, proto je dobré zvolit ideální způsob pro daná data. Některé testy pro analýzu stacionarity vychází z předpokladů charakteru signálu a jeho rozdělení (parametrické testy), jiné testy naopak ne (neparametrické testy).

1.4.1 Znaménkový test

Jednou z možností je použití testu, který je založen na znaménkovém neparametrickém testu, který je popsán například v [12] nebo v [13]. Hlavním parametrem signálu, který je testován, je výkon.

Celý test probíhá tak, že se signál rozdělí na segmenty a v každém segmentu se spočítá výkon. Testuje se hypotéza, že výkony jsou nezávislé náhodné proměnné. Při správnosti hypotézy jsou posloupnosti výkonů náhodné a neexistuje mezi nimi žádný trend. Očekává se, že změny znamének vůči střední hodnotě posloupnosti vychází z binomického rozdělení s parametry N a 0.5, kde N je počet segmentů (viz [14] – dodatek B).

Analýza probíhá tak, že se od všech výkonů odečte střední hodnota a pak se zjišťuje, kolikrát dojde k průchodu nulou, neboli kolikrát přejdeme z kladných hodnot do záporných, či obráceně. Navíc platí pravidlo, že při velkém počtu segmentů (20 a více) se dá nahradit binomické rozdělení normovaným normálním rozdělením. Tento převod se provádí následovně:

$$U = \frac{Yp - N}{\sqrt{N}} \quad (12)$$

kde Yp je počet průchodů nulou a N je počet segmentů.

1.4.2 Analýza rozptylu spojená s Bartlettovým testem

Na základě definování stacionarity se nabízí varianta, aby bylo kontrolováno, jestli se nemění střední hodnota a směrodatná odchylka. Pro tuto kontrolu lze použít jednofaktorovou analýzu rozptylu (dále ANOVA) a Bartlettův test (viz [15]). U obou těchto postupů je potřeba rozdělit signál, který považujeme za stacionární, na několik segmentů,

u kterých pak porovnááme, jestli mají statisticky stejnou střední hodnotu a směrodatnou odchylku.

Jednofaktorová ANOVA testuje nulovou hypotézu, že střední hodnota je ve všech skupinách stejná. Pokud není střední hodnota ve všech skupinách stejná, pak platí alternativní hypotéza a nulovou hypotézu můžeme zamítnout na zvolené hladině. Tento test je závislý na velikosti poměru F , ten se vypočítá následovně:

$$F = \frac{MS_A}{MS_E} \quad (13)$$

Kde MS_A je meziskupinový rozptyl a MS_E je vnitroskupinový rozptyl, tyto hodnoty se počítají následovně:

$$MS_A = \frac{\sum_{j=1}^a n_j (\bar{X}_j - \bar{X})^2}{a - 1} \quad (14)$$

$$MS_E = \frac{\sum_{j=1}^a \sum_{i=1}^{n_j} (X_{j,i} - \bar{X}_j)^2}{\sum_{j=1}^a (n_j - a)} \quad (15)$$

kde a je počet skupin, n_j je počet vzorků v jedné skupině, \bar{X}_j je průměr jednotlivých skupin a \bar{X} je celkový průměr. Hodnota F se pak porovnává s kritickou hodnotou. Tato hodnota má Fisherovo rozdělení s $a - 1$ stupni volnosti v čitateli a $\sum_{j=1}^a (n_j - a)$ stupni volnosti ve jmenovateli. Popis tohoto testu vychází z reference [16].

ANOVA test je použit i v jiném testu na stacionaritu a to konkrétně u Priestlyho a Subba Raova testu ([17]), který pracuje tak, že se spočítá spektrum v každém segmentu, na který je signál rozdělen a pak se počítá ANOVA test mezi těmito spektry.

Bartlettův test testuje nulovou hypotézu, že směrodatné odchylky jsou ve všech skupinách stejné. Zde se počítá kritická hodnota T následovně:

$$T = \frac{(N - k) \ln s_p^2 - \sum_{i=1}^k (N_i - 1) \ln s_i^2}{1 + \frac{1}{3(k-1)} \left(\left(\sum_{i=1}^k \frac{1}{N_i - 1} \right) - \frac{1}{N - k} \right)} \quad (16)$$

kde N je počet všech vzorků, N_i jsou velikosti každé skupiny, s_i^2 je rozptyl každé skupiny, k je počet segmentů a s_p^2 je spojený rozptyl, který se počítá:

$$s_p^2 = \frac{\sum_{i=1}^k (N_i - 1) s_i^2}{N - k} \quad (17)$$

Hodnota T se porovnává s kritickou hodnotou z rozdělení χ^2 s $k - 1$ stupni volnosti. Popis tohoto testu vychází z reference [18].

1.4.3 AR model

Další možností pro zjišťování stacionarity signálu je zjišťování koeficientů autoregresního (AR) modelu a porovnávání změn vlastností tohoto modelu v jednotlivých částech signálu. Pro popis AR modelu a kepstra se vychází z knihy [19].

Samotný autoregresní proces pro generování signálu je generován IIR filtrem řádu p , který neobsahuje dopředné vazby

$$x[n] = - \sum_{k=1}^p a_k x[n - k] + u[n] \quad (18)$$

kde x je výsledný signál generovaný tímto filtrem a u je bílý šum. AR model obsahuje pouze póly, takže bývá označován jako „all-pole“ model.

Kromě modelování takového signálu je možné také z analyzovaného signálu odhadnout koeficienty AR modelu. Model předvídá n -tý vzorek analyzovaného signálu následujícím způsobem:

$$\hat{s}[n] = - \sum_{k=1}^p a_k s[n-k] \quad (19)$$

kde p je řád modelu. Pro správné nalezení příslušných koeficientů k danému signálu je potřeba minimalizovat chybu odhadu. Musí být předvídané hodnoty blízké skutečným hodnotám v signálu. Proto je definován chybový signál následujícím způsobem:

$$e[n] = s[n] - \hat{s}[n] = s[n] + \sum_{k=1}^p a_k s[n-k] = \sum_{k=0}^p a_k s[n-k] \quad (20)$$

Určení koeficientů a_k pak vychází z minimalizování předchozí rovnice. Nejčastější metodou pro výpočet je Levinsonův-Durbinův algoritmus, který využívá Youle-Walkerovy rovnice a v MATLABu je implementován ve funkci *lpc*. Dalšími metodami jsou Itakurův algoritmus nebo Burgův algoritmus (v MATLABu implementován jako funkce *arburg*).

Kromě výpočtu koeficientů AR modelu je také důležitý výpočet výkonu chyby predikce, který je označován jako P_{min} . Pokud počítáme s tím, že se autoregresní model budí bílým šumem, pak se dá z výkonu chyby predikce vypočítat koeficient zesílení G , jehož vztah k výkonu chyby predikce je následující:

$$G = \sqrt{P_{min}} \quad (21)$$

1.4.4 Použití AR modelu

Samostatná analýza stacionarity signálu se pak provádí tak, že se signál rozdělí na několik segmentů a v těch se vypočítají koeficienty AR modelu a poté se zjistí, jestli jsou AR modely v po sobě následujících segmentech podobné. Toho využívá článek [20], kde se k výpočtu rozdílu AR modelů v jednotlivých segmentech používá následující rovnice:

$$d_A(k) = \sum_{p=1}^P [a_{p,k}^2 - a_{p,k-1}^2], k = 2, \dots, K \quad (22)$$

kde P je řád modelu a K je počet segmentů. Tato rovnice je ovšem nejspíše vzhledem k výsledkům špatně interpretována a mělo by se jednat pouze o rovnici, která počítá euklidovskou vzdálenost koeficientů ve dvou po sobě jdoucích segmentech, která se počítá následovně:

$$d_A(k) = \sqrt{\sum_{p=1}^P (a_{p,k} - a_{p,k-1})^2}, k = 2, \dots, K \quad (23)$$

Ovšem vzdálenost koeficientů dvou AR modelů nemusí být úplně využitelná, protože samostatné koeficienty nevyjadřují charakteristiku modelu. Proto se dá využít metody, která vychází z článku [21], kde se zjišťuje, jestli póly AR modelu jsou blízké a tím pádem spektra AR modelů jsou podobná. Pro měření vzdáleností spektra se dá využít například logaritmické spektrální vzdálenosti (zdroj [22]):

$$LSD = \sqrt{\frac{1}{B} \int_0^B [10 \log S(\omega) - 10 \log \hat{S}(\omega)]^2 d\omega} \quad (24)$$

kde S a \hat{S} jsou výkony spektra, B je šířka pásma spektra a ω je frekvence ve spektru. Pro takový odhad je ovšem potřeba z AR modelů zjistit spektrum a pak porovnat spektra mezi sebou. Logaritmická spektrální vzdálenost se dá ovšem aproximovat kepstrální vzdáleností, ke které stačí vypočítat pouze kepstrální koeficienty z koeficientů AR modelu.

1.4.5 Kepstrum

Myšlenka kepstrální analýzy vychází z převodu konvoluce na součet. Máme posloupnost $x[n]$, která vyjadřuje konvoluci dvou posloupností. Tato posloupnost se převede na posloupnost $\hat{x}[n]$, která vyjadřuje součet dvou jiných posloupností. Slovo kepstrum je přesmyčkou slova spektrum, takže kepstrum nevyjadřuje ani samostatný signál, ale ani jeho spektrum. Celá transformace z časové oblasti do kepstrální oblasti se dá popsat základním vztahem:

$$\hat{x}[n] = D\{x[n]\} = Z^{-1}\ln(Z\{x[n]\}) \quad (25)$$

kde D vyznačuje nelineární transformaci, Z vyznačuje z-transformaci a Z^{-1} vyznačuje inverzní z-transformaci. Jedná se pouze o základní vztah, existují i další vztahy, kdy se například z-transformace nahradí diskretní fourierovou transformací a inverzní z-transformace se nahradí inverzní diskretní fourierovou transformací.

Velmi důležitá je kepstrální vzdálenost, která je výhodná. Výhoda spočívá v tom, že se jedná o Euklidovskou vzdálenost kepstrálních koeficientů. Kepstrální vzdálenost dvou kepster mezi sebou se počítá následovně:

$$d \approx 4.34 \sqrt{(c[0] - \hat{c}[0])^2 + 2 \sum_{k=1}^L (c[k] - \hat{c}[k])^2} \quad [dB] \quad (26)$$

Pokud ovšem chceme znát jen rozdíl tvaru spekter, tak je dobré použít variantu výpočtu s potlačením vlivu energie. V takovém případě se používá vztah bez nultého kepstrálního koeficientu:

$$d \approx 4.34 \sqrt{2 \sum_{k=1}^L (c[k] - \hat{c}[k])^2} \quad [dB] \quad (27)$$

1.4.6 Převod koeficientů

Kepstrum jako takové může být komplexní, výkonové nebo reálné. Zde se budeme zabývat reálným kepstrem, protože koeficienty reálného kepstra se dají získat z koeficientů AR modelu a to tak, že z konečného počtu koeficientů AR modelu je možné v podstatě získat nekonečné množství kepstrálních koeficientů. Pro výpočet nultého kepstrálního koeficientu platí vztah:

$$c[0] = \frac{1}{2} \ln(\alpha) = \ln(G) \quad (28)$$

kde α označuje výkon chyby predikce a G je koeficient zesílení. Pro výpočet dalších kepstrálních koeficientů z koeficientů AR modelu se použijí následující vztahy:

$$c[n] = -a_n - \frac{1}{N} \sum_{k=1}^{n-1} (n-k)a_k c[n-k], \quad \text{pro } n \leq p \quad (29)$$

kde $c[n]$ je n -tý kepstrální koeficient, a jsou koeficienty AR modelu a p je řád AR modelu. Předchozí vztah je možné extrapolovat na výpočet dalších kepstrálních koeficientů:

$$c[n] = -\frac{1}{N} \sum_{k=1}^p (n-k)a_k c[n-k], \quad \text{pro } n > p \quad (30)$$

Z poslední rovnice tedy vychází skutečnost, že je možné počítat neomezené množství kepstrálních koeficientů na základě koeficientů AR modelu.

■ 1.4.7 Použití stacionarity signálu

Zjištění stacionarity signálu bylo použito v již zmíněných článcích [20] a [21], kde v obou článcích byl použit AR model. Další použití nestacionárních a stacionárních úseků se nachází v článku [23], kde se porovnává pulsní variabilita se srdeční variabilitou. Další články, které se zabývají stacionaritou spojenou se srdečním rytmem, jsou [24] a [25].

Další možnosti, jak analyzovat stacionaritu signálu, jsou popsány například v referenci [26].

Kapitola 2

Cíle práce

Cílem práce je nastudování metodiky analýzy stacionarity signálu společně s analýzou změny srdečního rytmu. Konkrétně se jedná o seznámení se s parametry a možnostmi, které se zjišťují ze záznamu měření srdečního rytmu.

Na základě těchto znalostí je cílem naprogramovat vhodný algoritmus pro analýzu změny srdečního rytmu. Tento algoritmus se opírá o autoregresní a keprstrální model. U těchto modelů je dále důležité optimalizovat parametry.

Tento algoritmus je důležité otestovat na záznamech posloupností RR intervalů a na záznamech z akcelerometru, kdy je důležité testovat různé nastavení parametrů algoritmu. Dále je potřeba na základě vypočtených modelů otestovat stacionaritu jednotlivých segmentů. Pro rozdělení těchto segmentů je použito prahování.

Ze získaných stacionárních úseků v záznamech RR intervalů i v záznamech z akcelerometru je důležité ověřit, zda existuje nějaká závislost mezi těmito úseky. Tedy jestli se stacionární úseky vyskytují přibližně ve stejných částech v obou typech záznamů. Konkrétně se jedná o zjištění závislosti změny četnosti pohybu na změnu srdečního rytmu.

Na základě výsledků a zjištěných poznatků z těchto záznamů je vhodné diskutovat možnost odhadu stacionarity v reálném čase v závislosti na záznamu z akcelerometru.

Kapitola 3

Návrh řešení

Základem celého řešení bylo pořízení 24hodinových záznamů, které obsahovaly časy mezi srdečními stahy (tedy záznamy RR intervalů) a zároveň obsahovaly informaci o změně pohybu na základě akcelerometru. Ke každému takovému záznamu bylo potřeba vytvořit anotace aktivit podle záznamů, které byly během měření prováděny. Na těchto záznamech byla provedena analýza pomocí výpočtu koeficientů AR modelu, ze kterých se poté vypočítaly kepstrální koeficienty. Z těchto koeficientů poté byla vypočítána kepstrální vzdálenost mezi po sobě následujícími segmenty a segmenty, které mezi sebou měly malou kepstrální vzdálenost, byly označeny za stacionární. Úseky, které byly označeny za stacionární, byly porovnány, jestli jsou v obou typech dat přibližně ve stejných místech. Kromě provedení samotného nalezení těchto úseků bylo ještě dobré ověřit, jestli jsou úseky skutečně stacionární pomocí dalších používaných metod pro zjištění stacionarity signálu.

Pro použití těchto metod bylo potřeba ověřit správnost přepočtu koeficientů AR modelu na kepstrální koeficienty. Pro toto ověření byly navrženy AR modely 1., 2. a 3. řádu a ručně se z nich vypočítaly kepstrální koeficienty. Tyto výsledky se poté porovnal s koeficienty, které byly vypočítány naprogramovanou funkcí pro převedení těchto koeficientů.

Po ověření správnosti přepočtů koeficientů bylo dále potřeba zjistit, jestli skutečně kepstrální vzdálenost může něco říct o změně charakteru signálu. Pro tento výpočet byl vygenerován bílý šum, který byl poté filtrován AR modelem tak, aby tento model měl v první polovině odlišné vlastnosti než v druhé polovině. Očekávalo se tedy, že zhruba v polovině signálu bude výrazná změna kepstrální vzdálenosti. Tento postup bylo vhodné provést u modelů 1., 3. a 5. řádu, aby bylo poznat, že je tato změna u jakéhokoliv řádu výrazná právě v místě, kde se změni vlastnosti signálu.

Pro použití výpočtu kepstrální vzdálenosti na získaných záznamech bylo potřeba ze záznamů odstranit místa, která nemají v takovém záznamu být. Konkrétně se jednalo u obou typů záznamu o úseky, ve kterých se neprovádělo měření a u záznamů RR intervalů bylo potřeba odstranit extrasystoly. Po tomto odstranění bylo potřeba ještě interpolovat záznam RR intervalů tak, aby se jednalo o signál, který je vzorkován pravidelně. U akcelerometru se nedá očekávat, že by měřil pohyb pouze v jednom směru, takže bylo potřeba sloučit rozumně data ze všech os dohromady.

Další částí postupu bylo zjištění rozdělení dat RR záznamů a rozdělení dat záznamů z akcelerometru. Pro tento způsob bylo vhodné vykreslení histogramu, který dokáže zobrazit charakter dat nejlépe. Navíc se tím dalo ověřit, že záznamy RR intervalů vychází z předpokládaného rozdělení. Předpokládaným rozdělením je směs více normálních jednorozměrných rozdělení (viz [27]). U akcelerometru je složitější určit, z jakého rozdělení data pochází, protože záleží na umístění akcelerometru a také na osobě, která ho používá. Toto mínění vychází z toho, že na základě histogramu záznamu dat z akcelerometru při chůzi se dá rozpoznat, kdo tento přístroj používá (viz [28]). Na základě toho se tedy dá tvrdit, že neexistuje nějaký jednotný model pro data z akcelerometru, jako pro záznam RR intervalů.

Po zjištění těchto vlastností byl proveden výpočet koeficientů AR modelů a kepstrálních koeficientů na reálných záznamech obou typů dat. Na základě výsledků se dalo zjistit, jaká je přibližně vhodná délka jednoho segmentu pro výpočet koeficientů a jaký řád AR modelů a kepstra je ideální. Vzhledem k zobrazeným spektrům v článku [7], by měl stačit 5. řád, ovšem pro lepší odhad je vždy lepší volit o něco vyšší řád.

Po zjištění koeficientů byl proveden výpočet kepstrální vzdálenosti a euclidovské vzdálenosti koeficientů AR modelu a také výpočet vzdálenosti podle článku [20]. Tento postup byl zvolen proto, aby se dalo potvrdit, že kepstrální vzdálenost je pro porovnávání podobnosti AR modelů nejsprávnější. Dále byla porovnána vypočtená kepstrální vzdálenost s nulovým koeficientem se vzdáleností vypočtenou bez nulového koeficientu. Tyto vzdálenosti pak mezi sebou byly porovnány.

Dalším krokem bylo určení stacionárních úseků v naměřených záznamech. Tyto úseky byly určeny podle kepstrální vzdálenosti a aktuálního zvoleného prahu. Byl použit dynamický i statický práh a tyto dva použité prahy byly porovnány mezi sebou.

Pro potvrzení jestli je tato metoda dobrá bylo důležité zvolit vhodnou metodu pro ověření stacionarity v úsecích, které jsou označeny za stacionární. Pro toto potvrzení byly použity 2 rozdílné testy. Jednou se jednalo o znaménkový test výkonu a podruhé o ANOVA test, který byl spojený s Bartlettovým testem (dále označován ANOVA test). Tyto testy byly použity i samostatně na manuálně zvolené 15 minutové úseky z dat, podle toho, jestli se u tohoto úseku očekává, že je stacionární nebo ne. Manuálně zvolené úseky byly vhodné pro ladění parametrů.

Po testu stacionarity byly použity dvě metody pro porovnání, jestli jsou určené stacionární úseky v obou typech dat na jednom záznamu ve stejném čase nebo v rozdílném čase. Tyto metody se opírají o výpočet velikosti překryvu mezi jednotlivými úseky.

Poslední částí bylo vytvoření grafické aplikace, která používá jednotlivé metody pro konkrétní záznamy. U grafické aplikace bylo důležité, aby byla možnost nastavovat jednotlivé parametry manuálně.

Kapitola 4

Realizace řešení

Pro možnost realizovat navržené řešení bylo potřeba nahrát 24hodinové záznamy, které obsahovaly posloupnosti RR intervalů a změny pohybu pomocí akcelerometru. Pro vypracování této práce bylo použito 8 záznamů. Ještě než ale byly jednotlivé zmíněné metody použity na tyto záznamy, tak bylo potřeba ověřit převod koeficientů z AR modelu na kepstrální koeficienty a také bylo potřeba zjistit, zda na umělých datech je kepstrální vzdálenost rozdílná, pokud se změní vlastnosti signálu.

4.1 Ověření převodu koeficientů

Pro převody koeficientů AR modelu na kepstrální koeficienty byly použity rovnice (29) a (30) z kapitoly 1.4.6. Pro toto ověření byly použity převody z 1. řádu AR modelu na 3. řád kepstra, z 2. řádu AR modelu na 4. řád kepstra a z 3. řádu AR modelu na 5. řád kepstra. Pro tento převod byly zvoleny ve všech případech 2 různé modely. Všechny zadané parametry modelů jsou vymyšlené, tedy nejsou odnikud předem vypočítané. V MATLABu byla pro tento převod vytvořena funkce *ar2cepstra*.

4.1.1 Převod z 1. řádu AR modelu na 3. řád kepstra

Koeficient prvního AR modelu pro ověření byl $a_1 = 0.9$. Ruční přepočítání koeficientů kepstra 3. řádu vypadal následovně:

$$\begin{aligned}c[1] &= -a_1 - 0 = -0.9 \\c[2] &= -\frac{1}{2} \sum_{k=1}^1 (2-k)a_k c[2-k] = -\frac{1}{2}(0.9 \cdot (-0.9)) = 0.405 \\c[3] &= -\frac{1}{3} \sum_{k=1}^1 (3-k)a_k c[3-k] = -\frac{1}{3}(2 \cdot 0.9 \cdot 0.405) = -0.243\end{aligned}\quad (1)$$

Výstupem funkce pro přepočítání koeficientů při zmíněných parametrech byly kepstrální koeficienty (seřazeny od 1. po poslední koeficient, stejně i u ostatních výsledků):

$$c[1] = -0.9, c[2] = 0.405, c[3] = -0.243$$

Koeficient druhého AR modelu pro ověření byl $a_1 = 0.5$. Ruční přepočítání koeficientů kepstra 3. řádu vypadal následovně:

$$\begin{aligned}c[1] &= -a_1 - 0 = -0.5 \\c[2] &= -\frac{1}{2} \sum_{k=1}^1 (2-k)a_k c[2-k] = -\frac{1}{2}(0.5 \cdot (-0.5)) = 0.125 \\c[3] &= -\frac{1}{3} \sum_{k=1}^1 (3-k)a_k c[3-k] = -\frac{1}{3}(2 \cdot 0.5 \cdot 0.125) = -0.04167\end{aligned}\quad (2)$$

Výstupem funkce pro přepočtení koeficientů při zmíněných parametrech byly keprální koeficienty:

$$c[1] = -0.5, c[2] = 0.125, c[3] = -0.0417$$

4.1.2 Převod z 2. řádu AR modelu na 4. řád kepra

Koeficienty prvního AR modelu pro ověření byly $a_1 = 0.9, a_2 = 0.25$. Ruční přepočtení na koeficienty kepra 4. řádu vypadal následovně:

$$\begin{aligned} c[1] &= -a_1 - 0 = -0.9 \\ c[2] &= -0.25 - \frac{1}{2} \sum_{k=1}^1 (2-k)a_k c[2-k] = -0.25 - \frac{1}{2}(0.9 \cdot (-0.9)) = 0.155 \\ c[3] &= -\frac{1}{3} \sum_{k=1}^2 (3-k)a_k c[3-k] = -\frac{1}{3}(2 \cdot 0.9 \cdot 0.155 + 0.25 \cdot (-0.9)) = -0.018 \\ c[4] &= -\frac{1}{4} \sum_{k=1}^2 (4-k)a_k c[4-k] = -\frac{1}{4}(3 \cdot 0.9 \cdot (-0.018) + 2 \cdot 0.25 \cdot 0.155) = \\ &= -0.00722 \end{aligned} \quad (3)$$

Výstupem funkce pro přepočtení koeficientů při zmíněných parametrech byly keprální koeficienty:

$$c[1] = -0.9, c[2] = 0.155, c[3] = -0.018, c[4] = -0.0072$$

Koeficienty druhého AR modelu pro ověření byly $a_1 = 0.5, a_2 = 0.2$. Ruční přepočtení na koeficienty kepra 4. řádu vypadal následovně:

$$\begin{aligned} c[1] &= -a_1 - 0 = -0.5 \\ c[2] &= -0.2 - \frac{1}{2} \sum_{k=1}^1 (2-k)a_k c[2-k] = -0.2 - \frac{1}{2}(0.5 \cdot (-0.5)) = -0.075 \\ c[3] &= -\frac{1}{3} \sum_{k=1}^2 (3-k)a_k c[3-k] = -\frac{1}{3}(2 \cdot 0.5 \cdot (-0.075) + 0.2 \cdot (-0.5)) = 0.05833 \\ c[4] &= -\frac{1}{4} \sum_{k=1}^2 (4-k)a_k c[4-k] = -\frac{1}{4}(3 \cdot 0.5 \cdot 0.05833 + 2 \cdot 0.2 \cdot (-0.075)) = \\ &= -0.01437 \end{aligned} \quad (4)$$

Výstupem funkce pro přepočtení koeficientů při zmíněných parametrech byly keprální koeficienty:

$$c[1] = -0.5, c[2] = -0.075, c[3] = 0.0583, c[4] = -0.0144$$

■ 4.1.3 Převod z 3. řádu AR modelu na 5. řád kepra

Koeficienty prvního AR modelu pro ověření byly $a_1 = 0.9, a_2 = 0.25, a_3 = 0.4$. Ruční přepočítání na koeficienty kepra 5. řádu vypadal následovně:

$$\begin{aligned}
 c[1] &= -a_1 - 0 = -0.9 \\
 c[2] &= -0.25 - \frac{1}{2} \sum_{k=1}^1 (2-k)a_k c[2-k] = -0.25 - \frac{1}{2}(0.9 \cdot (-0.9)) = 0.155 \\
 c[3] &= -0.4 - \frac{1}{3} \sum_{k=1}^2 (3-k)a_k c[3-k] = \\
 &= -0.4 - \frac{1}{3}(2 \cdot 0.9 \cdot 0.155 + 0.25 \cdot (-0.9)) = -0.418 \\
 c[4] &= -\frac{1}{4} \sum_{k=1}^3 (4-k)a_k c[4-k] = \\
 &= -\frac{1}{4}(3 \cdot 0.9 \cdot (-0.418) + 2 \cdot 0.25 \cdot 0.155 + 0.4 \cdot (-0.9)) = 0.3528 \\
 c[5] &= -\frac{1}{5} \sum_{k=1}^4 (5-k)a_k c[5-k] = \\
 &= -\frac{1}{5}(4 \cdot 0.9 \cdot 0.3528 + 3 \cdot 0.25 \cdot (-0.418) + 2 \cdot 0.4 \cdot 0.155) = -0.2161 \quad (5)
 \end{aligned}$$

Výstupem funkce pro přepočítání koeficientů při zmíněných parametrech byly keprální koeficienty:

$$c[1] = -0.9, c[2] = 0.155, c[3] = -0.418, c[4] = 0.3528, c[5] = -0.2161$$

Koeficienty druhého AR modelu pro ověření byly $a_1 = 0.5, a_2 = 0.2, a_3 = 0.05$. Ruční přepočítání na koeficienty kepra 5. řádu vypadal následovně:

$$\begin{aligned}
 c[1] &= -a_1 - 0 = -0.5 \\
 c[2] &= -0.2 - \frac{1}{2} \sum_{k=1}^1 (2-k)a_k c[2-k] = -0.2 - \frac{1}{2}(0.5 \cdot (-0.5)) = -0.075 \\
 c[3] &= -0.05 - \frac{1}{3} \sum_{k=1}^2 (3-k)a_k c[3-k] = \\
 &= -0.05 - \frac{1}{3}(2 \cdot 0.5 \cdot (-0.075) + 0.2 \cdot (-0.5)) = 0.00833 \\
 c[4] &= -\frac{1}{4} \sum_{k=1}^3 (4-k)a_k c[4-k] = \\
 &= -\frac{1}{4}(3 \cdot 0.5 \cdot 0.00833 + 2 \cdot 0.2 \cdot (-0.075) + 0.05 \cdot (-0.5)) = 0.01064 \\
 c[5] &= -\frac{1}{5} \sum_{k=1}^4 (5-k)a_k c[5-k] = \\
 &= -\frac{1}{5}(4 \cdot 0.5 \cdot 0.01064 + 3 \cdot 0.2 \cdot 0.00833 + 2 \cdot 0.05 \cdot (-0.075)) = -0.00375 \quad (6)
 \end{aligned}$$

Výstupem funkce pro přepočítání koeficientů při zmíněných parametrech byly kepstrální koeficienty:

$$c[1] = -0.5, c[2] = -0.075, c[3] = 0.0083, c[4] = 0.0106, c[5] = -0.0038$$

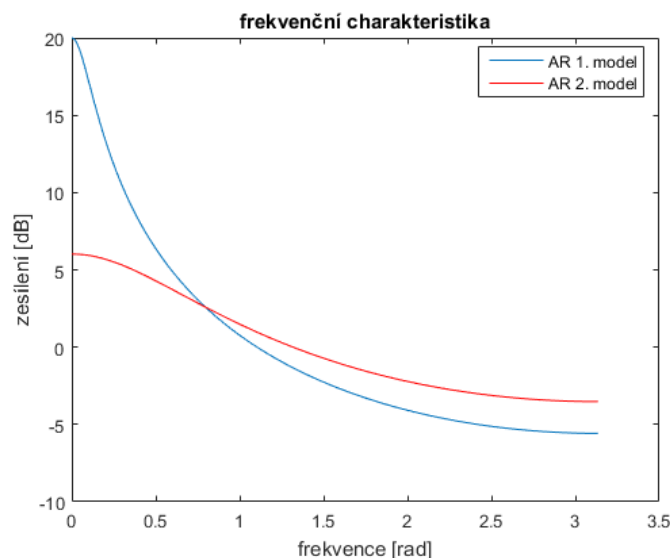
Z výsledků je tedy patrné, že funkce pro převod koeficientů fungovala správně. Celý tento proces ověření převodu je uložen v samotném skriptu, který je pojmenován *testAR2Cepstra*.

4.2 Ověření kepstrální vzdálenosti

Pro ověření, jestli je výpočet kepstrální vzdálenosti schopen zachytit změny vlastností signálu, byl zvolen postup, kdy se vygenerovalo 1000 vzorků signálu bílého šumu. Tento signál byl poté filtrován dvěma AR modely, kdy každá filtrace byla provedena na původní signál a výsledné filtrované signály se poté spojily za sebe. Poté byl signál rozdělen na segmenty o délce 100 vzorků a v každém segmentu byly vypočítány koeficienty AR modelu. Tyto koeficienty pak byly převedeny na kepstrální koeficienty a z kepstrálních koeficientů pak byla spočítána kepstrální vzdálenost u všech po sobě následujících segmentů. Řád AR modelu a kepstra byl v tomto případě stejný. Pro dobré ověření byly na filtraci použity AR modely 1., 3. a 5. řádu. Stejně jako v předchozí části i zde všechny parametry AR modelů byly vymyšlené. Pro výpočet kepstrální vzdálenosti byly použity rovnice (26) a (27) z kapitoly 1.4.5.

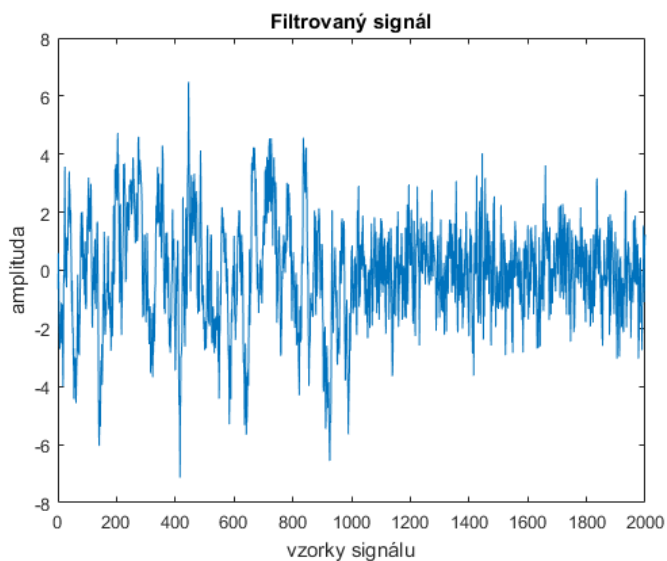
4.2.1 AR model 1. řádu

U modelů prvního řádu byla první polovina signálu filtrována modelem s koeficientem $a_1 = -0.9$. Druhá polovina signálu byla filtrována modelem s koeficientem $a_1 = -0.5$. Frekvenční charakteristiky těchto filtrů jsou zobrazeny na obrázku 4.1.



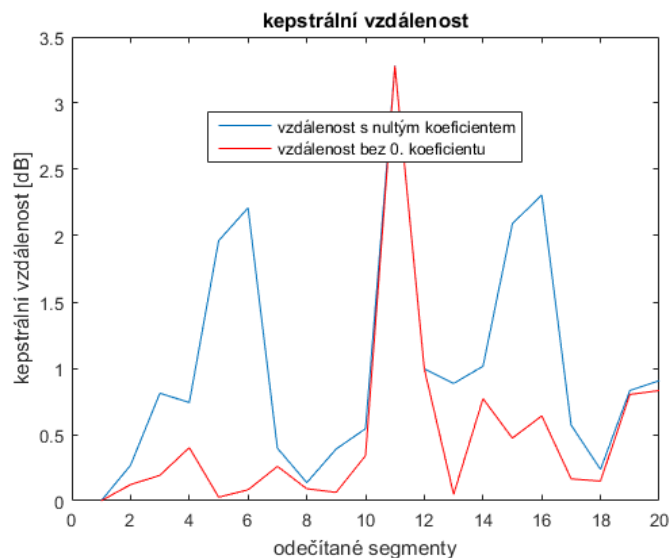
Obrázek 4.1. 1. frekvenční charakteristika AR modelů

Z obrázku 4.1 lze odvodit, že oba filtry mají rozdílné vlastnosti. Že jsou skutečně rozdílné, je možné ověřit na průběhu signálu po této filtraci, který je zobrazen na obrázku 4.2. První polovina signálu byla filtrována prvním modelem, druhá polovina byla filtrována druhým modelem.



Obrázek 4.2. 1. filtrovaný signál

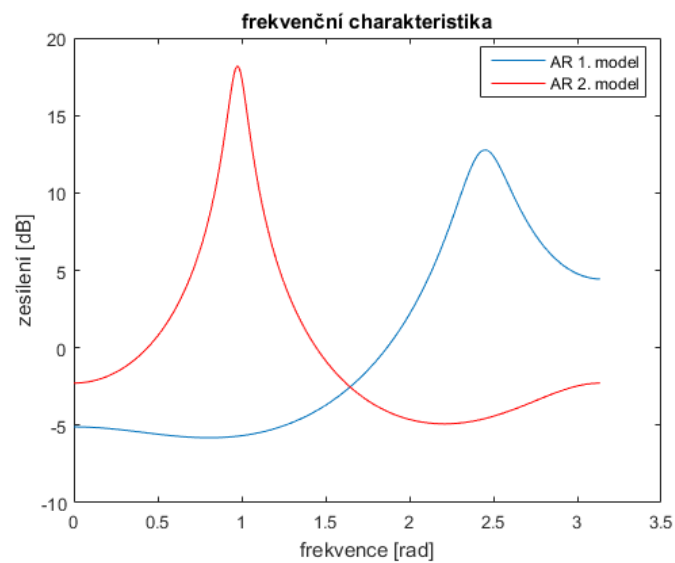
Na obrázku 4.3 je zobrazen průběh kepstrální vzdálenosti.



Obrázek 4.3. 1. kepstrální vzdálenost

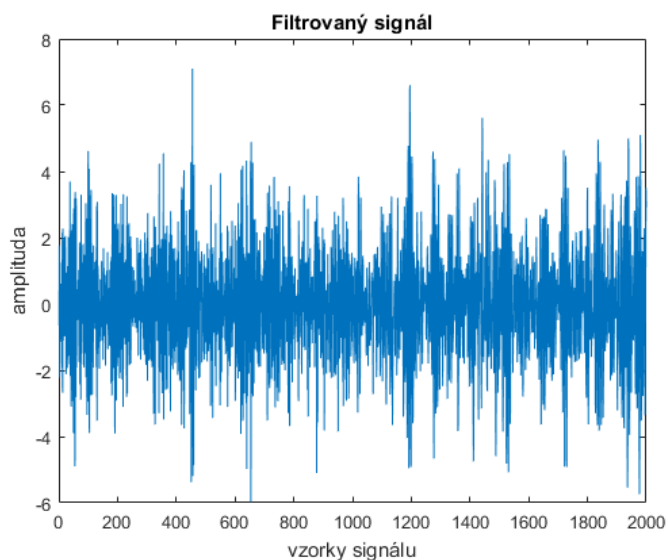
4.2.2 AR model 3. řádu

U modelů třetího řádu byla první polovina signálu filtrována modelem s koeficienty $a_1 = -0.9, a_2 = 0.2, a_3 = -0.3$. Druhá polovina signálu byla filtrována modelem s koeficienty $a_1 = -0.5, a_2 = 0.3, a_3 = 0.5$. Frekvenční charakteristiky těchto filtrů jsou zobrazeny na obrázku 4.4.



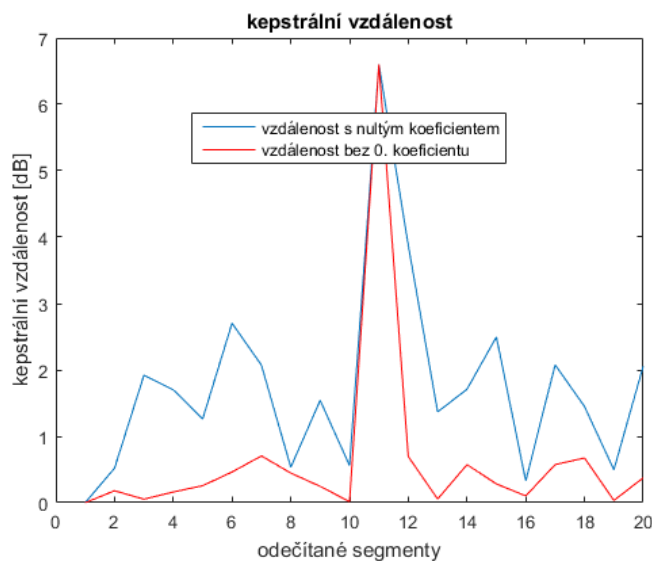
Obrázek 4.4. 2. frekvenční charakteristika AR modelů

Z obrázku 4.4 lze odvodit, že tyto filtry mají více rozdílnou frekvenční charakteristiku než filtry 1. řádu. Filtrovaný signál je zobrazen na obrázku 4.5. První polovina signálu byla filtrována prvním modelem, druhá polovina byla filtrována druhým modelem.



Obrázek 4.5. 2. filtrováný signál

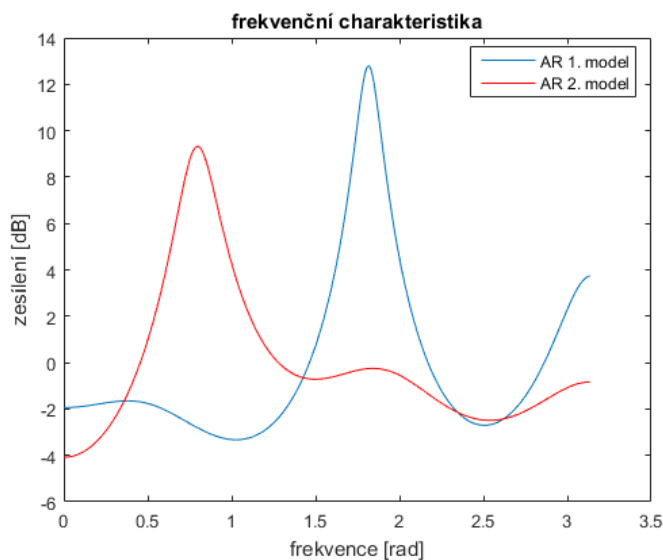
Na obrázku 4.6 je zobrazen průběh kepstrální vzdálenosti.



Obrázek 4.6. 2. kepstrální vzdálenost

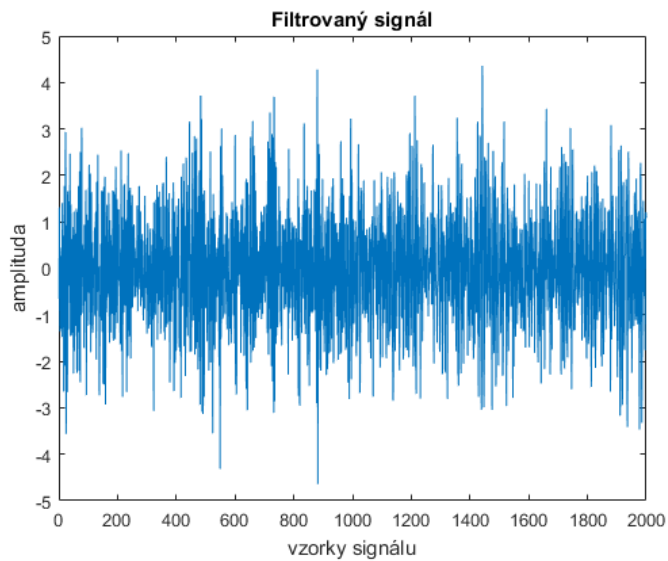
4.2.3 AR model 5. řádu

U modelů pátého řádu byla první polovina signálu filtrována modelem s koeficienty $a_1 = 0.2, a_2 = 0.25, a_3 = -0.15, a_4 = -0.3, a_5 = 0.25$. Druhá polovina signálu byla filtrována modelem s koeficienty $a_1 = -0.2, a_2 = 0.15, a_3 = 0.25, a_4 = 0.2, a_5 = 0.2$. Frekvenční charakteristiky těchto filtrů jsou zobrazeny na obrázku 4.7.



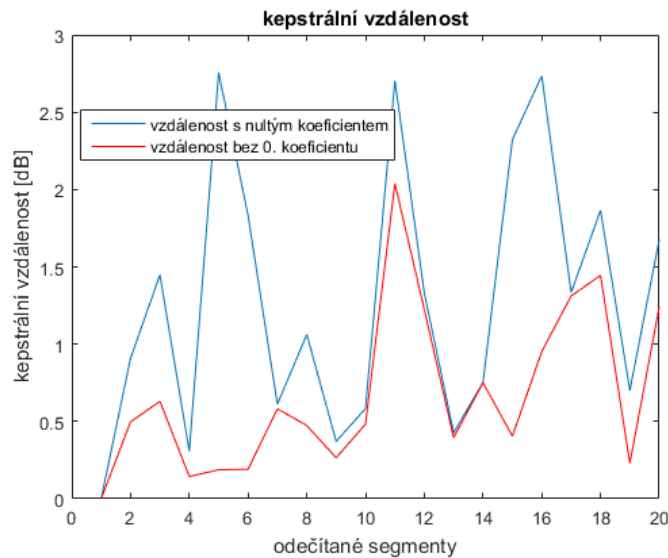
Obrázek 4.7. 3. frekvenční charakteristika AR modelů

Z obrázku 4.7 lze odvodit, že tyto filtry mají více rozdílnou frekvenční charakteristiku než filtry 3. řádu. Filtrovaný signál je zobrazen na obrázku 4.8. První polovina signálu byla filtrována prvním modelem, druhá polovina byla filtrována druhým modelem.



Obrázek 4.8. 3. filtrovaný signál

Na obrázku 4.9 je zobrazen průběh kepstrální vzdálenosti.



Obrázek 4.9. 3. kepstrální vzdálenost

Ze všech zobrazených grafů kepstrálních vzdáleností (obrázky 4.3, 4.6 a 4.9) je zřejmé, že u těchto signálů lépe pozná významnou změnu signálu kepstrální vzdálenost bez nulového koeficientu. U kepstrální vzdálenosti s nulovým koeficientem z výsledků vychází, že se velmi mění i v místech, kde by se měnit neměla. Každý AR model byl počítán pro kepstrální vzdálenost ze segmentů o délce 100 vzorků, takže mohlo dojít k chybě odhadu, proto nakonec byl ve výsledcích použit pro porovnání na datech i nulý koeficient. Postup v této kapitole sloužil pouze k ověření, že při velké změně AR modelu skutečně bude i vysoká kepstrální vzdálenost, což bylo potvrzeno u obou typů kepstrálních vzdáleností.

4.3 Nahrávání záznamu

Pro nahrání požadovaného záznamu byl použit přístroj Firstbeat Bodyguard 2¹). Toto zařízení je schopno zároveň nahrávat jak zrychlení, tak posloupnost RR intervalů. V této práci bylo použito 8 24hodinových záznamů. Všechny tyto záznamy byly pořízeny na jedné osobě. U všech záznamů byla zvolena u akcelerometru vzorkovací frekvence 50 Hz. Samostatný přístroj měří EKG záznam, ze kterého vypočítává RR intervaly. Pro získání EKG záznamu používá přístroj vzorkovací frekvenci 1 kHz. Původní záznam RR intervalů je uložen v souboru formátu *SDF* a záznam z akcelerometru je uložen v souboru formátu *CSV*.

Ještě před použitím těchto záznamů bylo potřeba upravit záznamy RR intervalů, protože původní záznamy obsahovaly artefakty a úseky, ve kterých se neměřilo. Pro odstranění artefaktů byl použit software od výrobce, který odstraňuje tato nevhodná místa v záznamu. Algoritmus výrobce vrací záznam v souboru formátu *CSV*. Tento soubor obsahuje v prvním sloupci původní záznam, ve druhém sloupci upravený záznam a ve třetím sloupci číslo artefaktu, který byl odstraněn.

Pro odstranění úseků, ve kterých se neměřilo, byla vytvořena funkce *deleteNonMeasureParts*. Tato funkce vychází z upravených dat algoritmem od výrobce. Část, ve které se neměřilo, je považována také za artefakt a ten je označen v těchto záznamech číslem 2. Pokud je tento artefakt obsažen v záznamech delší dobu za sebou, tak se odstraní celý tento úsek a čas začátku a konce takového úseku se uchová.

Po odstranění úseků, kde se neměřilo, byla použita funkce *detectEctopicHRBeats4*, která odstraňuje extrasystoly, které nebyly odstraněny v předchozí části. Tato funkce vychází z článku [29].

Po odstranění artefaktů a neměřených úseků bylo třeba, pro použití na výpočet koeficientů AR modelu, signál RR intervalů interpolovat. Byla použita lineární interpolace, která je implementována v MATLABu ve funkci *interp1*.

Signál zrychlení z akcelerometru obsahuje zrychlení ve 3 osách, které jsou zde označeny *X*, *Y* a *Z*. U všech os bylo nejdříve potřeba odstranit ty části záznamu, které byly odstraněny i ze záznamu RR intervalů. Pro výpočet koeficientů AR modelu bylo potřeba tyto 3 osy spojit dohromady tak, aby zůstala původní informace zachována. Pro toto spojení os byl použit následující výpočet:

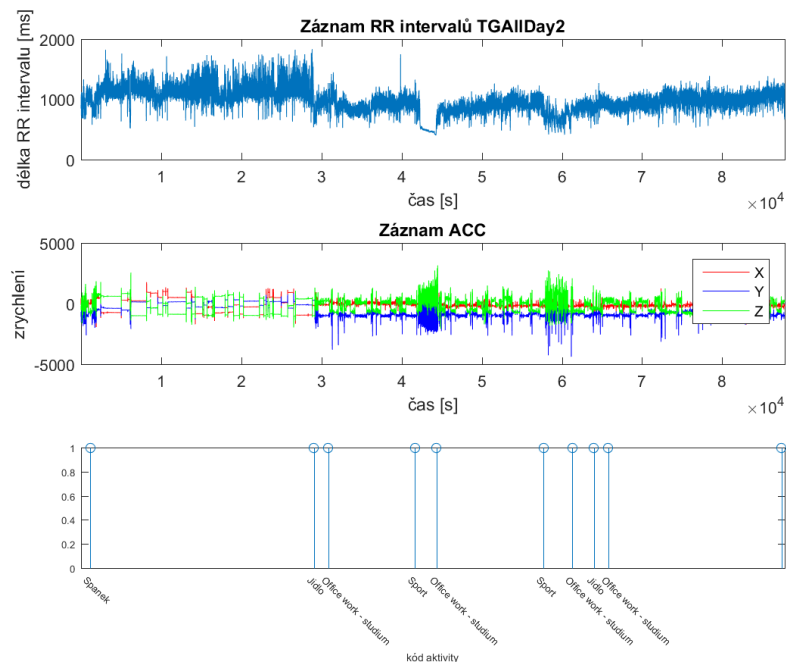
$$V(t) = \sqrt{(X(t)^2) + (Y(t)^2) + (Z(t)^2)} \quad (7)$$

Jedná se tedy o spojení os, které dosahuje vyšších hodnot, pokud nastane nějaký pohyb. Naopak, pokud je člověk v klidu, tak se ustálí na nízké hodnotě a nemění se.

Kromě těchto záznamů bylo potřeba zaznamenat seznam aktivit, které byly prováděny během jednoho 24hodinového měření. Protokol aktivit pro jedno měření byl uložen v souboru formátu *CSV*, kde byl vždy určen čas začátku v prvním sloupci, čas konce aktivity ve druhém sloupci a ve třetím sloupci byl uveden číselný kód aktivity. Všechny časy byly uloženy ve formátu *hodiny:minuty* podle toho, kdy byly během dne prováděny. Tedy ne v čase od začátku měření, ale v aktuálním čase s tím, že se jednalo o čas, který se vztahoval k nejbližší čtvrt hodině, kdy byla aktivita prováděna. Pro možnost převedení kódu aktivit na text, tak byl vytvořen další soubor formátu *CSV*. V tomto souboru se vyskytuje na řádce v prvním sloupci textové označení aktivity a ve druhém sloupci kód aktivity.

Výsledný záznam po odstranění artefaktů a úseků bez měření je zobrazen na obrázku 4.10 (zobrazeny všechny 3 osy, nikoliv jejich součet, se kterým se poté počítalo).

¹) <http://shop.firstbeat.com/all-products/bodyguard.html>

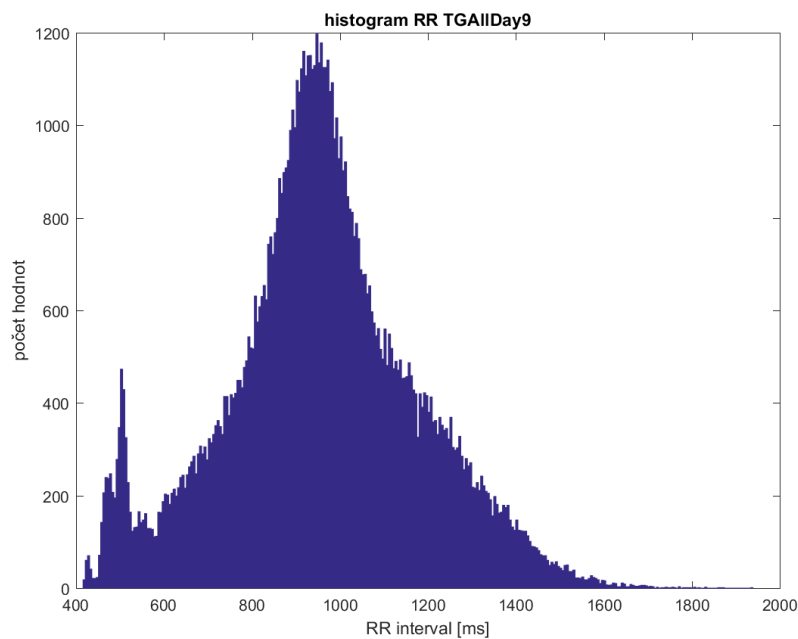


Obrázek 4.10. Záznam jednoho měření

4.4 Rozdělení dat záznamů

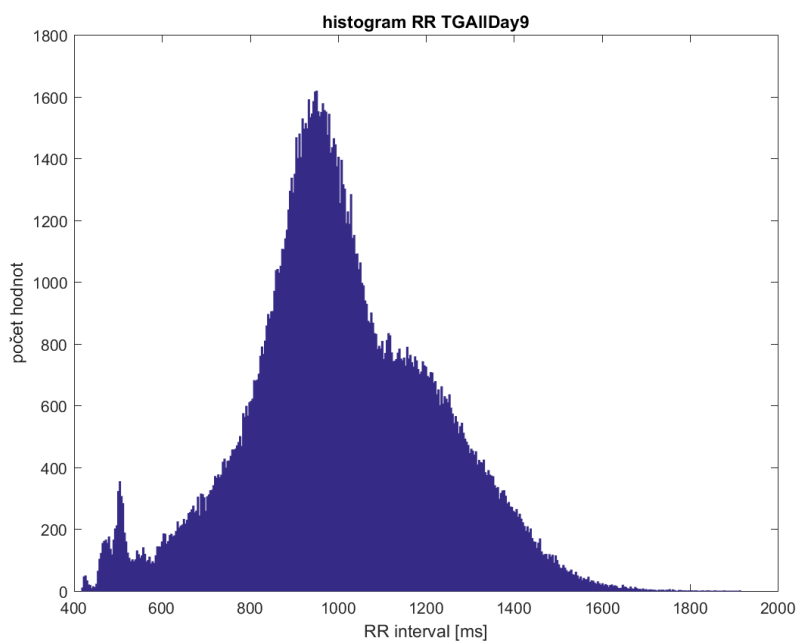
Pro zjištění rozdělení dat RR intervalů a dat z akcelerometru bylo potřeba vykreslit histogramy. Pro grafy histogramů bylo vybráno jedno referenční měření, protože histogramy u ostatních měření se příliš nelišily.

Histogram RR intervalů před interpolací je zobrazen na obrázku 4.11.



Obrázek 4.11. Histogram RR intervalů

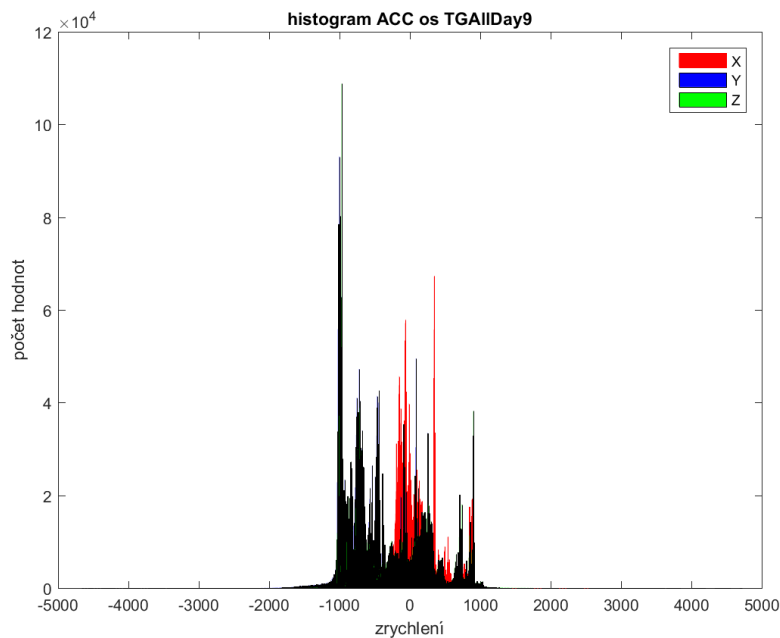
Z obrázku 4.11 lze poznat, že se jedná o směs normálních rozdělení. Histogram RR intervalů po provedení interpolace je zobrazen na obrázku 4.12.



Obrázek 4.12. Histogram RR intervalů po interpolaci

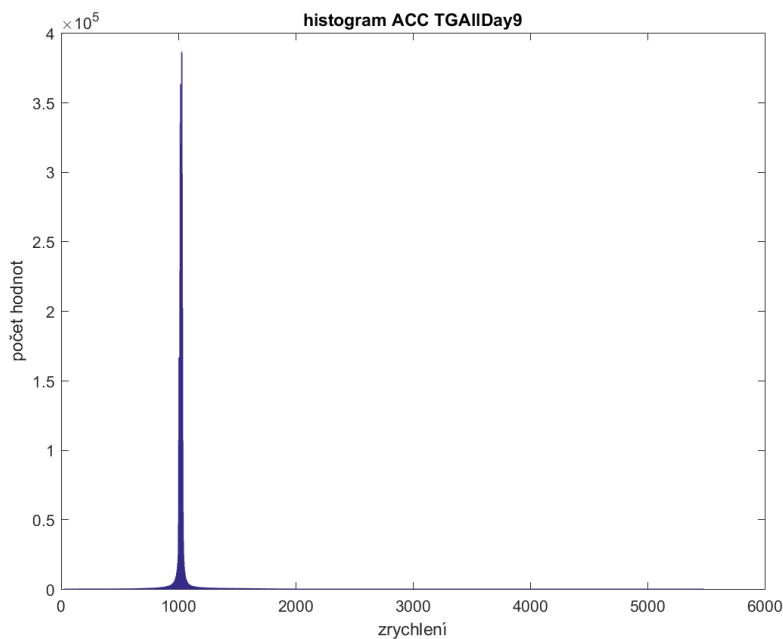
Je patrné, že histogram se změní pouze v některých částech, ale i tak histogram odpovídá směsi normálních rozdělení.

Histogram všech 3 os z akcelerometru je zobrazen na obrázku 4.13.



Obrázek 4.13. Histogram všech os z akcelerometru

Z obrázku 4.13 je možné vypořadovat, že se všechny 3 osy překrývají. Histogram po spojení všech 3 os dohromady je zobrazen na obrázku 4.14.



Obrázek 4.14. Histogram všech os dohromady z akcelerometru

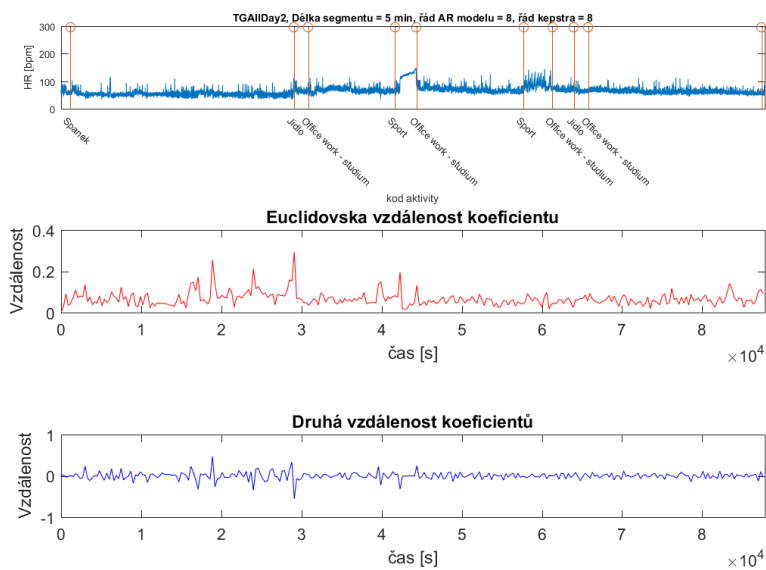
Z obrázku 4.14 je viditelné, že po spojení všech 3 os výrazně převažuje jedna hodnota. U všech os zvláště je sice histogram trochu širší, ale také i tam se vyskytují hodnoty, které značně převažují v každé ose. Převažující hodnota zobrazuje gravitační zrychlení.

4.5 Porovnání vzdáleností

Po získání záznamů byl vždy signál rozdělen na několik segmentů stejné délky a v každém segmentu byly vypočítány koeficienty AR modelu. Z těchto koeficientů byly vypočítány koeficienty kepstra. Pro správný výpočet bylo potřeba optimalizovat délku segmentu signálu, řád AR modelu a řád kepstra. Pro tento účel je podle článku [7] nejlepší používat segmenty délky 5 minut, proto byly v celé práci vypočteny koeficienty z úseků dlouhých 5 minut.

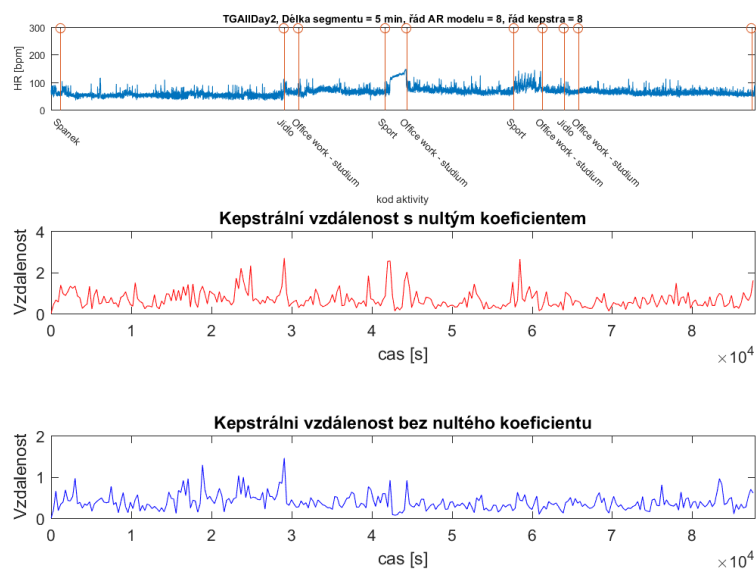
Po vypočítání těchto koeficientů bylo potřeba zjistit, o kolik se vlastnosti navazujících segmentů liší. Pro tento účel byla zvolena kepstrální vzdálenost, která vychází z rovnic (26) a (27) v kapitole 1.4.5. Výsledky kepstrální vzdálenosti byly porovnány s výsledky vzdáleností vypočtenými z rovnic (22) a (23) v kapitole 1.4.4.

Na obrázku 4.15 je zobrazena ukázka použití Euclidovské vzdálenosti a vzdálenosti z článku [20]. Na obrázku je nejdříve zobrazen graf srdečního rytmu a pod ním jsou dva grafy, kdy první zobrazuje Euclidovskou vzdálenost koeficientů a druhý graf zobrazuje vzdálenost ze zmíněného článku. Řád modelu byl v tomto případě 8.



Obrázek 4.15. Vzdálenost koeficientů AR modelu

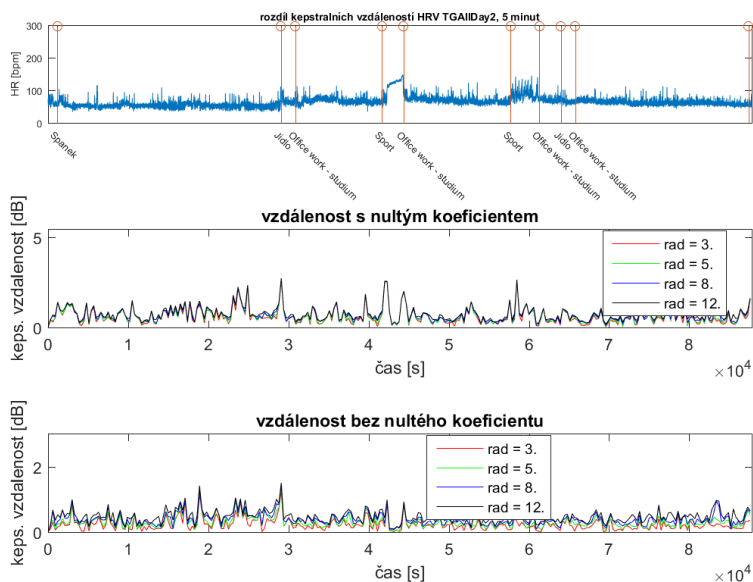
Z grafů na obrázku 4.15 lze poznat, že Euclidovská vzdálenost je rozhodně lepší než vzdálenost z článku [20], ale také je potřeba podívat se na kepstrální vzdálenost. Na obrázku 4.16 jsou zobrazeny grafy, které obsahují kepstrální vzdálenost s nultým koeficientem a bez nultého koeficientu. Byly použity stejné parametry jako u obrázku 4.15.



Obrázek 4.16. Kepstrální vzdálenost 1

Z grafů na obrázku 4.16 se dá vypořádat, že Euclidovská vzdálenost v mnoha místech aproximuje kepstrální vzdálenost bez nultého koeficientu. Euclidovská vzdálenost by správně neměla být používána u AR modelu, protože důležité je umístění pólů,

nikoliv velikost samostatných koeficientů. Navíc Euklidovská vzdálenost je schopna porovnat pouze AR modely stejného řádu. Díky této skutečnosti byla u všech výpočtů používána pouze keprální vzdálenost, kde se dají dva AR modely rozdílného řádu převést na keprum stejného řádu. Na obrázku 4.17 je zobrazen rozdíl mezi jednotlivými řády AR modelu při výpočtu keprální vzdálenosti.



Obrázek 4.17. Keprální vzdálenost 2

Z grafů na obrázku 4.17 je možné vypořozovat, že energie signálu u keprální vzdálenosti s nultým koeficientem má velký vliv, zatímco u keprální vzdálenosti bez nultého koeficientu záleží na řádu AR modelu. Vzhledem k předchozím grafům vychází jako nejlepší varianta použití AR modelu alespoň 8. řádu. Zároveň kvůli výpočetní rychlosti není dobré používat příliš vysoké řády modelu.

4.6 Označení možných stacionárních úseků

Dalším krokem bylo označení možných stacionárních úseků na základě keprální vzdálenosti. Pro úseky, které jsou označovány, jako stacionární musí platit, že mají nízkou keprální vzdálenost. K označení úseků byl použit práh keprální vzdálenosti. Pokud byla keprální vzdálenost menší než zvolený práh, tak byl úsek označen za stacionární, pokud byla naopak keprální vzdálenost vyšší, tak nebyl takový úsek označen za stacionární. Pokud byla keprální vzdálenost nižší než zvolený práh několikrát za sebou, pak byly jednotlivé segmenty, spojeny do jednoho úseku. Úsek pak byl označen jako stacionární celý, tedy pokrýval všechny segmenty, mezi kterými byla nižší keprální vzdálenost než práh, za sebou. Práh v celé této práci byl zvolen na základě následujícího vztahu:

$$T = \text{median}(dC) + a \cdot \text{std}(dC) \quad (8)$$

kde T je práh, dC je keprální vzdálenost, a je volitelný koeficient, median je funkce medianu a std je funkce pro výpočet směrodatné odchylky. Směrodatná odchylka je přičítána k mediánu ve zvoleném poměru. Koeficient a by neměl být záporný.

Tento práh byl použit dvojnásobným způsobem. Prvním způsobem bylo použití statického prahu a druhým způsobem bylo použití dynamického prahu. U dynamického prahu platí, že se bere aktuální hodnota T z předchozích několika vzorků a ne z celého měření jako u statického prahu.

4.7 Potvrzení stacionarity v úsecích

Předchozí postup sloužil pouze k nalezení možných stacionárních úseků. Po použití předchozího popsání postupu bylo potřeba potvrdit, jestli se skutečně jedná o stacionární úseky. Pro toto potvrzení byly použity nezávisle na sobě dva testy. Byl použit znaménkový test a ANOVA test. Oba tyto testy jsou popsány v kapitole 1. Pro správné použití těchto metod bylo potřeba najít vhodné parametry a také bylo potřeba zjistit, jestli jsou takové metody vůbec vhodné pro naměřené záznamy.

Pro ověření, jestli se jedná o vhodné metody, byly manuálně vybrány úseky dat o délce 15 minut, které měly charakteristické vlastnosti. U těchto úseků se buď dalo očekávat, že se jedná o stacionární, nebo ne. Tyto úseky byly vybrány jak ze záznamů RR intervalů, tak ze záznamů z akcelerometru ve stejných místech.

Celkový postup pro znaménkový test u akcelerometru i u RR intervalu spočíval v tom, že tento 15 minutový úsek byl rozdělen na segmenty o určité délce. V každém segmentu byl poté vypočítán výkon a na posloupnost výkonů byl použit znaménkový test. U RR intervalů byl znaménkový test navíc použit na vypočítané parametry $RMSSD$ a $SDNN$, které byly vypočítány v každém segmentu. U metody znaménkového testu s výkonem byl použit interpolovaný signál RR intervalů, u výpočtu parametrů $RMSSD$ a $SDNN$ byl použit původní (neinterpolovaný) signál RR intervalů. U výpočtu výkonu byl použit interpolovaný signál, protože by každý segment měl obsahovat stejně dlouhý časový interval pro správné výsledky.

Alternativou byl ANOVA test. U tohoto testu byl, stejně jako u znaménkového testu, každý 15 minutový úsek rozdělen na několik segmentů o stejné délce. Tyto segmenty pak byly porovnány ANOVA testem mezi sebou. U ANOVA testu byl použit původní signál RR intervalů, protože interpolace mírně změní původní vlastnosti dat.

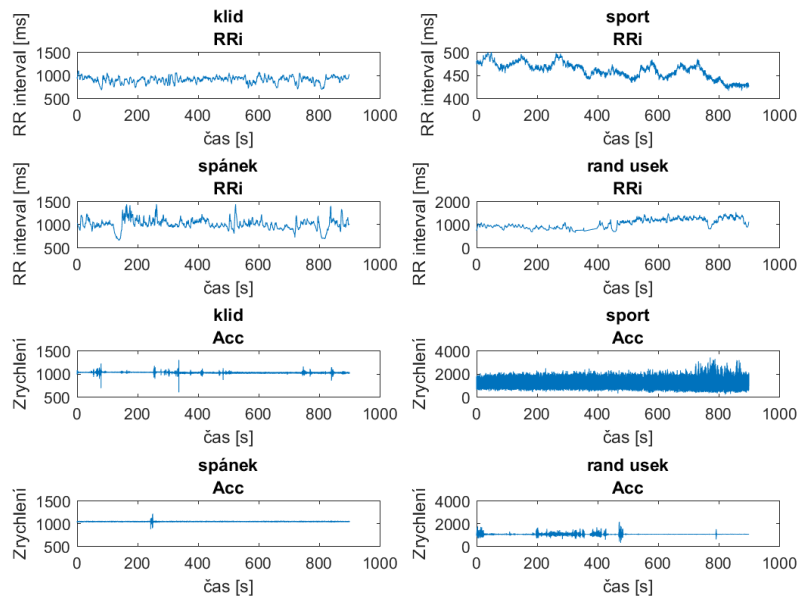
Byly zde použity 2 24hodinové záznamy, ze kterých byl vždy vybrán úsek, kde byl sport, úsek klidu, úsek ve spánku a náhodný úsek, kde se neočekává, že by byl stacionární. U ostatních se očekává, že by měly být stacionární.

U obou zmíněných testů byla vždy velikost segmentu pro jeden typ dat stejná. U akcelerometru se jednalo o 2500 vzorků, u RR intervalů se jednalo o 100 vzorků. Oba statistické testy (znaménkový test a ANOVA test) byly testovány na hladině významnosti $\alpha = 0.05$.

Výsledky jednotlivých metod jsou zapsány v tabulkách, kde pokud je ve výsledku *ano*, tak to znamená, že úsek byl konkrétní metodou označen za stacionární. Pokud je u výsledku *ne*, tak úsek nebyl konkrétní metodou označen za stacionární. U první tabulky je v prvním sloupci zobrazen výsledek znaménkového testu na $RMSSD$ parametr, ve druhém sloupci výsledek znaménkového testu na $SDNN$ parametr, ve třetím sloupci výsledek znaménkového testu na výkon záznamu RR intervalu a v posledním sloupci výsledek znaménkového testu na výkon ze záznamu dat z akcelerometru. U druhé tabulky je v prvním sloupci výsledek ANOVA testu na záznamu RR intervalů a ve druhém sloupci výsledek ANOVA testu na záznamu z akcelerometru.

4.7.1 První záznam s manuálními úseky

Úseky z prvního záznamu jsou zobrazeny na obrázku 4.18.



Obrázek 4.18. První záznam s manuálně vybranými úseky

Tabulka výsledků znaménkových testů pro tento záznam vypadala následovně:

Typ aktivity	Znam. test RMSSD	Znam. test SDNN	Znam. test výkon RR	Znam. test výkon Acc
Klid	ano	ano	ano	ano
Sport	ano	ano	ano	ne
Spánek	ano	ano	ano	ano
Náhodný úsek	ano	ano	ne	ano

Tabulka 4.1. Tabulka výsledků znaménkových testů u prvního záznamu

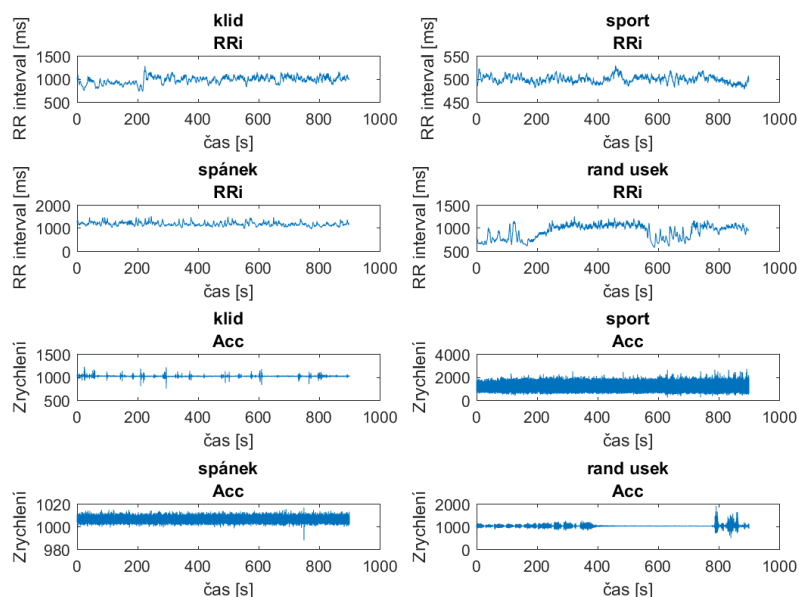
Tabulka výsledků ANOVA testů u tohoto záznamu vypadala následovně:

Typ aktivity	ANOVA test RR	ANOVA test Acc
Klid	ne	ne
Sport	ne	ne
Spánek	ne	ne
Náhodný úsek	ne	ne

Tabulka 4.2. Tabulka výsledků ANOVA testů u prvního záznamu

4.7.2 Druhý záznam s manuálními úseky

Úseky z druhého záznamu jsou zobrazeny na obrázku 4.19.



Obrázek 4.19. Druhý záznam s manuálně vybranými úseky

Tabulka výsledků znaménkových testů pro tento záznam vypadala následovně:

Typ aktivity	Znam. test RMSSD	Znam. test SDNN	Znam. test výkon RR	Znam. test výkon Acc
Klid	ano	ano	ano	ano
Sport	ano	ano	ano	ano
Spánek	ano	ano	ano	ano
Náhodný úsek	ano	ano	ne	ne

Tabulka 4.3. Tabulka výsledků znaménkových testů u druhého záznamu

Tabulka výsledků ANOVA testů u tohoto záznamu vypadala následovně:

Typ aktivity	ANOVA test RR	ANOVA test Acc
Klid	ne	ne
Sport	ne	ne
Spánek	ne	ano
Náhodný úsek	ne	ne

Tabulka 4.4. Tabulka výsledků ANOVA testů u druhého záznamu

Z výsledků lze poznat, že ANOVA test je velmi kritický a není nejspíše vhodným testem pro ověření stacionarity, protože tento test očekává, že data pochází z určitého rozdělení. Proto tento test u výsledků nebyl použit. U znaménkového testu je důležité, aby každý segment byl dostatečně dlouhý, protože odhad výkonu může být chybný, pokud je segment příliš krátký. Chyba odhadu výkonu se snižuje s délkou signálu, ze kterého je odhadován. U výpočtu parametrů náležících pouze RR intervalům (*RMSSD* a *SDNN*) znaménkový test potvrdil stacionaritu vždy, proto tento postup nebyl dále použit. Druhým důvodem bylo, že parametry *RMSSD* a *SDNN* nelze vypočítat z akcelerometru.

4.8 Porovnání úseků ze dvou typů dat

Nakonec bylo porovnáno, jestli se vyskytují jednotlivé nalezené stacionární úseky přibližně ve stejných místech v obou typech záznamů. Tedy jak v záznamu RR intervalů, tak v záznamu z akcelerometru. Pro toto porovnání byly navrženy 2 metody.

První metoda vycházela z toho, že je hledán překryv úseků, které začínají přibližně ve stejném čase. Přibližně stejný čas znamená +/- 10 sekund. U této metody se počítá překryv pouze těmi úseky, které začínají přibližně ve stejném čase.

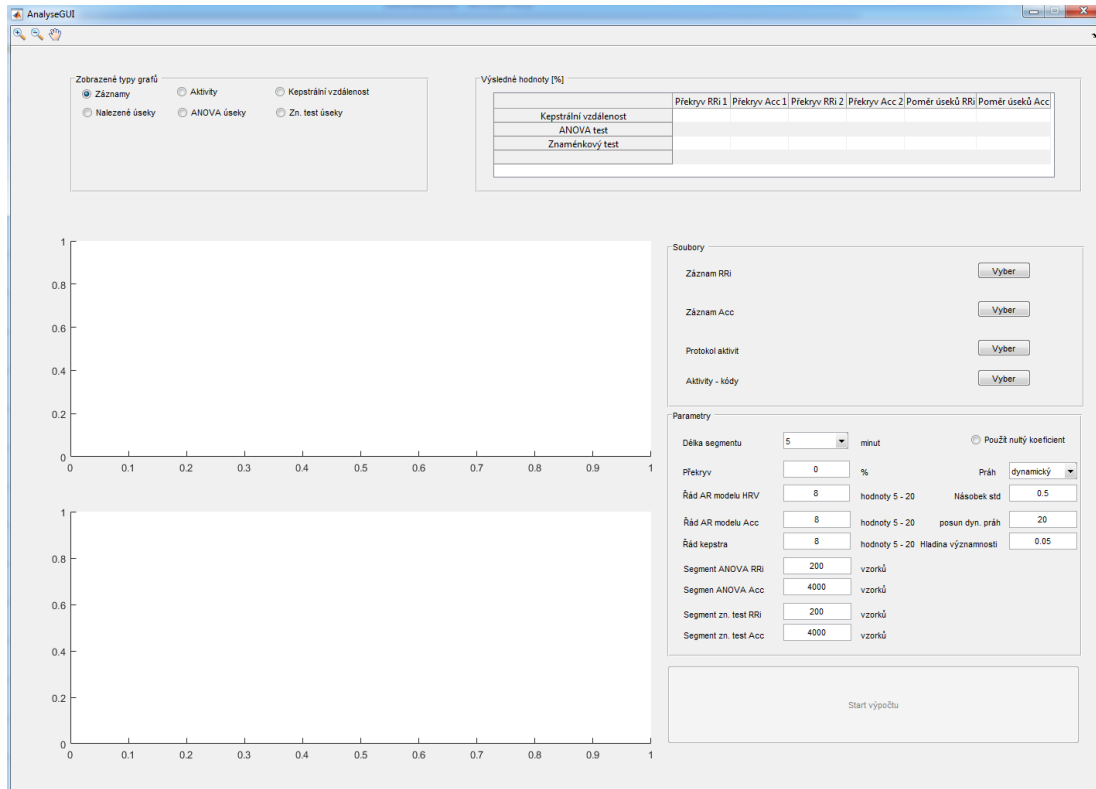
Druhá metoda počítala celkový překryv. Jednalo se o výpočet, z jak velké části jsou překryty všechny úseky jednoho typu dat všemi úseky z druhého typu dat, u kterých nezáleželo, jaký měly počáteční čas.

Další metodou byl výpočet celkového poměru stacionárních úseků vůči celému záznamu. Metoda počítala, jak velkou část pokrývají stacionární úseky v celém záznamu.

Kromě těchto metod bylo také dobré podívat se v jednotlivých záznamech na distribuci délky úseků u jednotlivých vyznačených aktivit. Zde každý úsek, který zasahuje do konkrétní označené aktivity, byl vybrán a pak byly vykresleny grafy, které obsahují histogramy časů jednotlivých úseků. Histogramy časů jednotlivých úseků jsou v příloze C.

4.9 Grafická aplikace

Pro snadnou analýzu dat byla vytvořena jednoduchá grafická aplikace v MATLABu, ve které je možné po nahrání souborů analyzovat data podle popsaných metod. Vzhled aplikace po spuštění je zobrazen na obrázku 4.20.



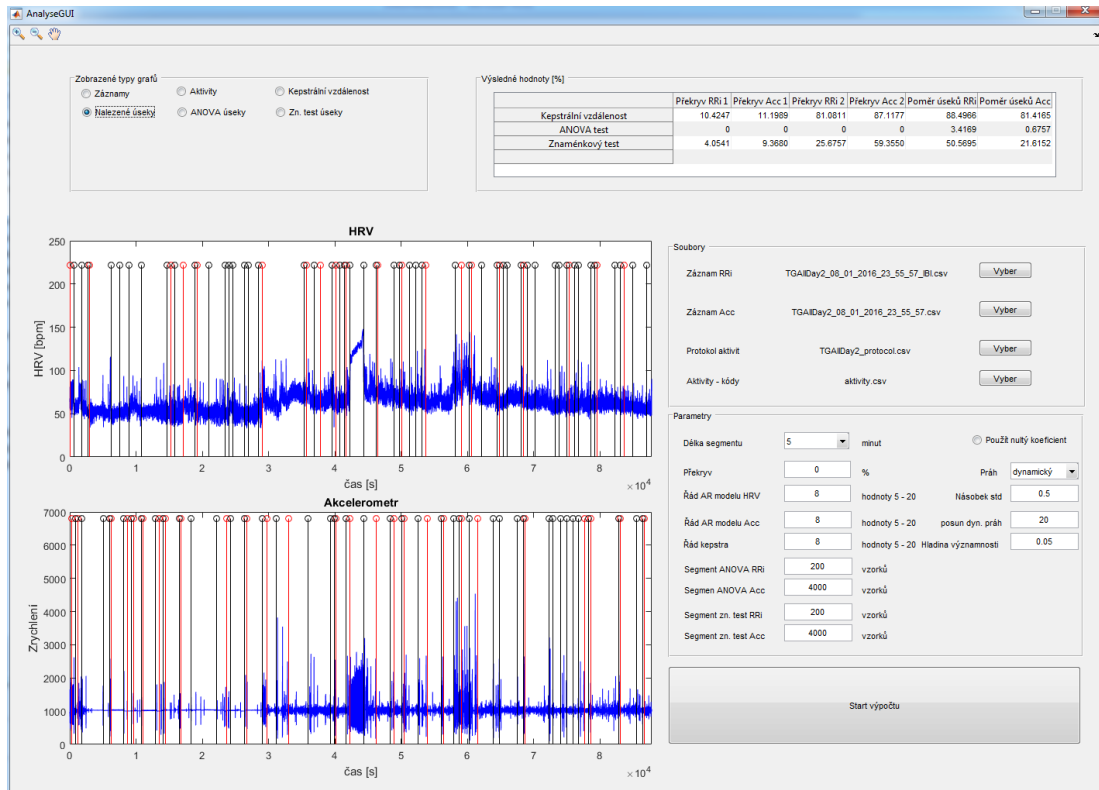
Obrázek 4.20. Grafická aplikace

U této aplikace se nahoře vpravo zobrazují výsledky překryvu jednotlivých úseků při jednotlivých metodách. Všechny tyto výsledky jsou zobrazeny v procentech. Pod těmito výsledky je potřeba načíst cestu k souborům, které jsou potřeba, aby bylo možné spustit analýzu. Pod nahráním těchto dat se dají nastavit jednotlivé parametry, které jsou důležité a dají se měnit. Po kliknutí na tlačítko vpravo dole je potřeba počkat, než se vygenerují výsledky, které se vykreslují v levé části aplikace. Vlevo nahoře je možné vybrat, co přesně se má vykreslovat za data.

Na rozdíl od výsledků uvedených v následující kapitole, je v aplikaci možné také vidět výsledek po použití ANOVA testu. Jak je zřejmé, tak je možné nastavovat velké množství parametrů a zároveň je možné vidět několik různých výsledků. Dále je také možné vidět samostatná původní data a také je možné vidět samostatnou kepstrální vzdálenost.

V aplikaci je například možné pro každý typ záznamu zvolit jiný řád AR modelu. Řád kepra ovšem musí zůstat stejný pro oba typy záznamů, protože jinak by nebylo možné vypočítat kepstrální vzdálenost. Dále je možné zvolit například hladinu významnosti pro statistické testy (ANOVA a znaménkový test) nebo třeba počet vzorků kepstrální vzdálenosti, ze kterého se počítá dynamický práh. Pro možnost spuštění je potřeba mít uvedené cesty ke všem typům souborů, které je potřeba použít. Tedy záznam RR intervalů, záznam z akcelerometru, protokol aktivit pro tento záznam a převodní tabulku u aktivit mezi kódem a názvem aktivity.

Po spuštění výpočtu u této aplikace se v levé části zobrazí 2 grafy. Výstup po provedení výpočtu a zobrazení výsledků u aplikace může vypadat například jako na obrázku 4.21.



Obrázek 4.21. Běh grafické aplikace

Pro spuštění této aplikace je potřeba mít nainstalovaný MATLAB s podporou grafického uživatelského rozhraní. Krátká dokumentace k aplikaci je na přiloženém CD (viz příloha B).

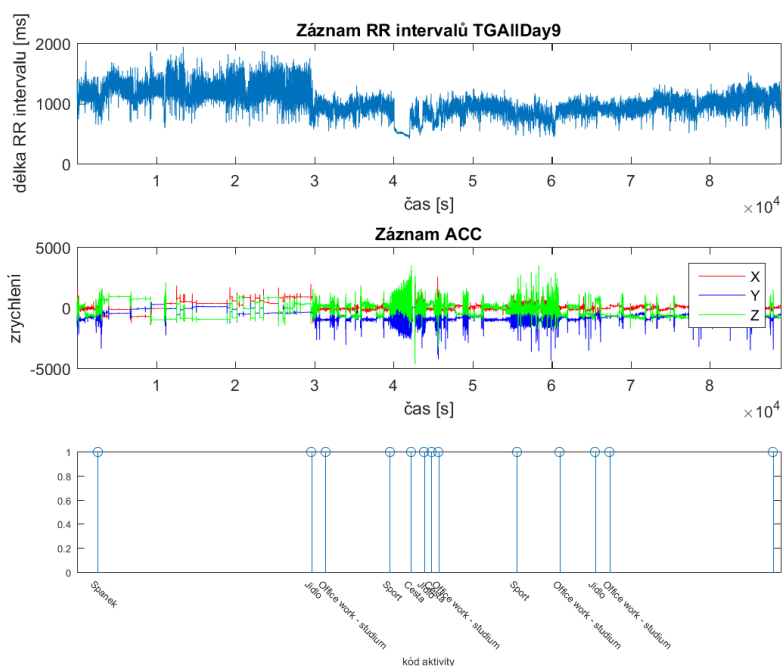
Kapitola 5

Výsledky

Výsledky zobrazují nalezené úseky, které jsou označené za stacionární a potvrzené úseky, které jsou stacionární. Dále zobrazují velikosti jednotlivých překryvů u úseků záznamů RR intervalů a úseků ze záznamů z akcelerometru. U všech výsledků bylo použito prahování kepstrální vzdálenosti na základě předchozí kapitoly. Pro potvrzení stacionarity jednotlivých úseků byl vybrán znaménkový test, protože podle popisu v předchozí kapitole je méně kritický než ANOVA test a zároveň nepředpokládá, že data vychází z konkrétního rozdělení.

Výsledky srovnávají dynamický se statickým prahem, rozdíl u velikosti prahu a rozdíl u použití kepstrální vzdálenosti s nulovým koeficientem a bez nulového koeficientu.

Pro jednotlivé obrázky výsledků v této kapitole byly vždy vybrány výsledky vypočtené ze záznamu, který je zobrazen na obrázku 5.1. Výsledky z konkrétního záznamu jsou zobrazeny proto, aby se dala jednotlivá nastavení parametrů dobře porovnat graficky.



Obrázek 5.1. Záznam použitého měření

U výsledků při jednom nastavení parametrů je nejdříve zobrazen obrázek, který ukazuje grafy nalezených stacionárních úseků pomocí prahování kepstrální vzdálenosti. Na všechny nalezené úseky pomocí prahování kepstrální vzdálenosti byl použit znaménkový test. Grafy potvrzených úseků pomocí znaménkového testu jsou zobrazeny na druhém obrázku u konkrétního nastavení parametrů. Znaménkový test byl vybrán, protože podle popisu v předchozí kapitole je méně kritický než ANOVA test a zároveň nepočítá, že data vychází z konkrétního rozdělení.

Na zobrazených grafech výsledků vyznačuje červená čára začátek úseku a černá konec úseku. Pokud jsou zobrazeny 2 černé čáry za sebou, tak to znamená, že v místě, kde končí jeden úsek, začíná další úsek. Dochází k tomu proto, že pokud budeme mít například celkově 4 segmenty a mezi 1. a 2. segmentem bude vzdálenost malá, mezi 2. a 3. segmentem bude naopak vzdálenost velká a mezi 3. a 4. segmentem bude vzdálenost malá, tak budeme mít 2 úseky, kde jeden začíná na začátku 1. segmentu a končí na konci 2. segmentu. Druhý úsek začíná na začátku 3. segmentu a končí na konci 4. segmentu. Ale konec 2. segmentu je časově téměř ve stejném místě jako začátek 3. segmentu. Jde o posunutí pouze o 1 vzorek signálu, což při zobrazení 24hodinového záznamu není téměř poznat. Kdyby byl zvolen překryv segmentů, tak by se mohl nacházet překryv i u nalezených úseků.

Za obrázky, které zobrazují nalezené a potvrzené stacionární úseky, se nachází dvě tabulky. První tabulka ukazuje výsledky ze záznamu, který je zobrazen na obrázku 5.1. Druhá tabulka zobrazuje výsledky na všech použitých záznamech při stejném nastavení parametrů.

U všech tabulek výsledků se vycházelo z kapitoly 4.8. Překryv 1 v tabulkách vyznačuje první popsanou metodu, tedy překryvy úseků jednoho typu dat pouze úseky druhého typu dat, které začínají přibližně ve stejném čase. Překryv 2 vyznačuje celkový překryv úseků jednoho typu dat od druhého a celý překryv znamená, jak moc je celkový signál překryt stacionárními úseky u jednoho typu dat. Celý překryv je v tabulkách označen jako celý Acc nebo celý HRV. Pokud je označen sloupec například Překryv Acc 1, tak to znamená, že se kontroluje, z jaké části jsou překryty stacionární úseky z akcelerometru stacionárními úseky z RR intervalů na základě první metody překryvu. HRV vyznačuje výsledky ze záznamu srdečního rytmu (RR intervalů). Acc vyznačuje výsledky ze záznamu z akcelerometru.

Pro výpočet koeficientů AR modelů byla délka jednoho segmentu u obou typů dat dlouhá 5 minut. Překryv segmentů byl nulový. Znaménkový test byl testován na hladině významnosti $\alpha = 0.05$. Pro znaménkový test délka jednoho segmentu v nalezeném úseku byla 100 vzorků u záznamu RR intervalů a 2500 vzorků u záznamu z akcelerometru. U všech výsledků byl zvolen 8. řád keprsta i AR modelu.

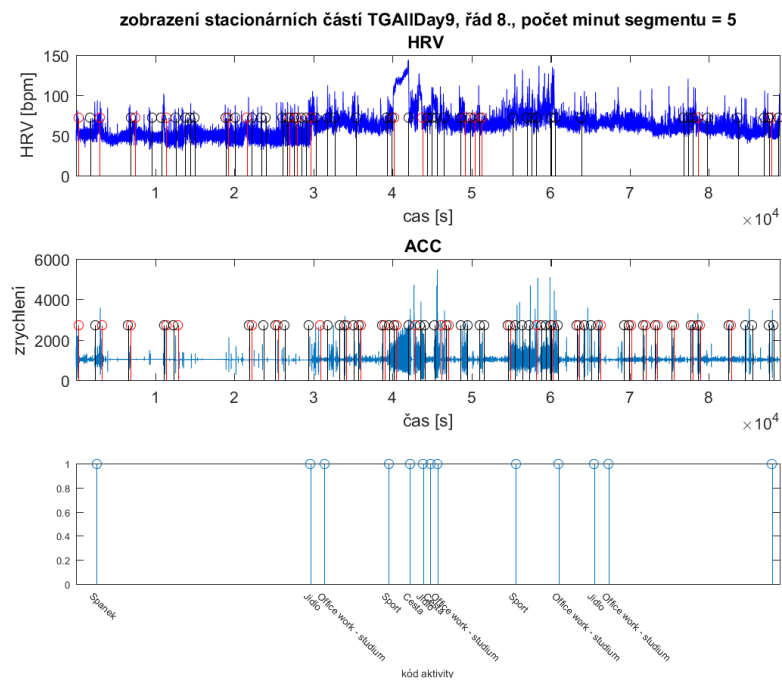
5.1 Statický práh

U statického prahu byla nastavena jedna hodnota z celého měření podle rovnice (8) z kapitoly 4.6. Tento práh byl stejný pro všechny vzorky vypočtené keprstrální vzdálenosti. Měnil se pouze parametr, který určuje násobek směrodatné odchylky. Zobrazeny jsou výsledky u 2 nastavených prahů a u každého prahu je postupně použit výsledek s použitým nultým koeficientem a výsledek bez použitého nultého koeficientu.

5.1.1 Nižší práh

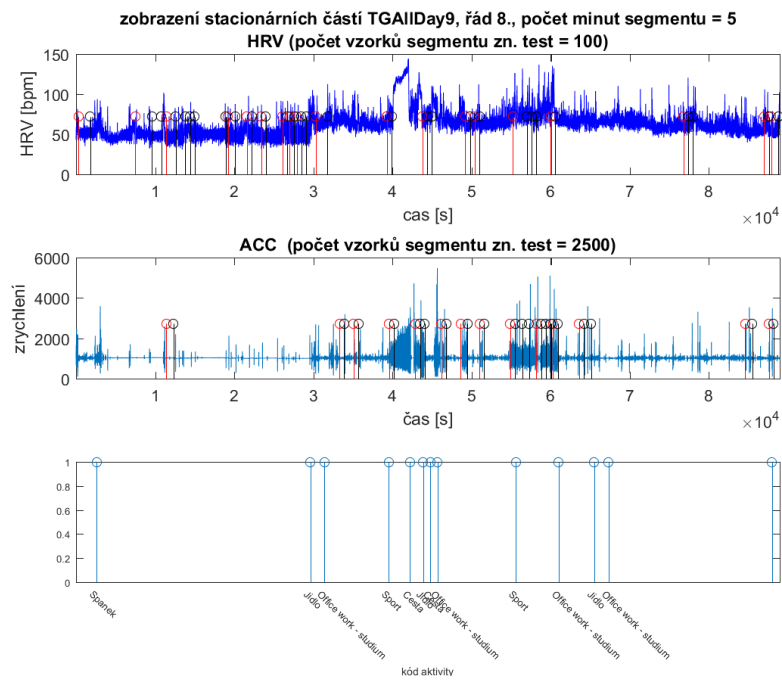
U nižšího prahu byl parametr a nastaven na hodnotu 0.5.

U následujících výsledků byla použita keprstrální vzdálenost s nultým koeficientem. Na obrázku 5.2 jsou zobrazeny nalezené stacionární úseky podle prahu keprstrální vzdálenosti.



Obrázek 5.2. První nalezené úseky

Obrázek 5.3 zobrazuje potvrzené stacionární úseky znaménkovým testem.



Obrázek 5.3. První potvrzené úseky

Tabulka překryvů vychází následovně:

Test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	6.5 %	6.5 %	87.8 %	88 %	88.4 %	86.4 %
Znam. test	3.3 %	6.5 %	16.3 %	32.6 %	31 %	15.2 %

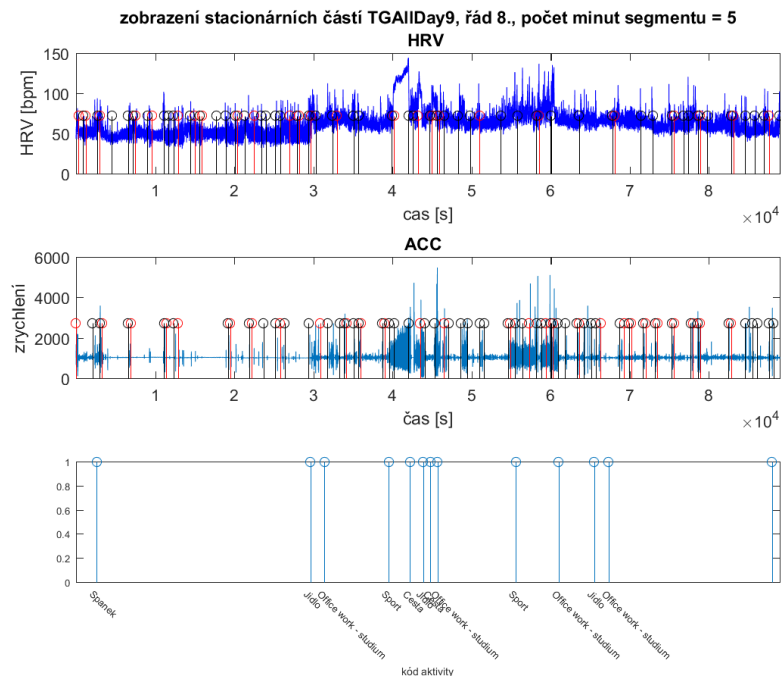
Tabulka 5.1. Tabulka prvních výsledků na 1 záznamu

Tabulka pro všechny záznamy při těchto parametrech vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	15 %	15.5 %	86.6 %	88.6 %	89 %	85.3 %
Znam. test	4.4 %	10.7 %	13.8 %	35.9 %	40.8 %	15.5 %

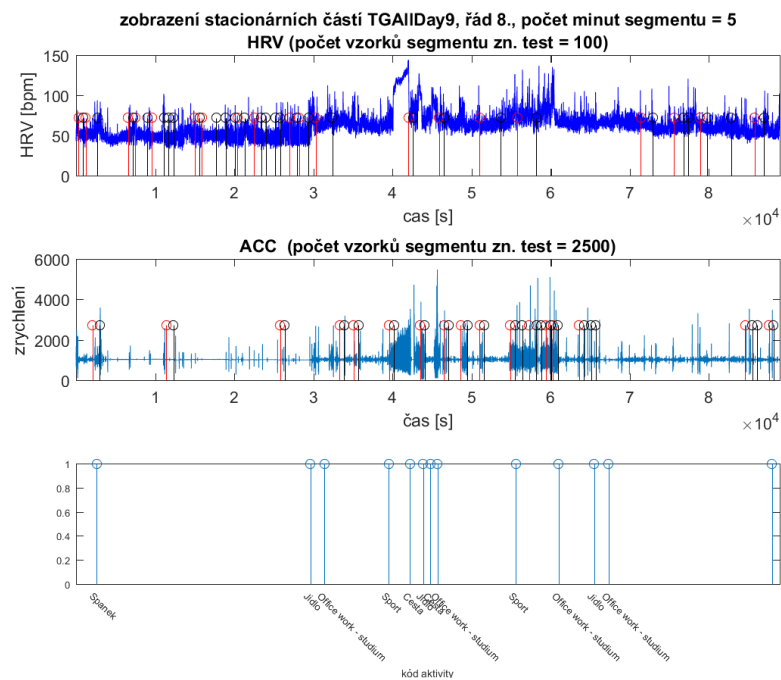
Tabulka 5.2. Tabulka prvních výsledků na všech záznamech

U následujících výsledků byla použita kepstrální vzdálenost bez nultého koeficientu. Na obrázku 5.4 jsou zobrazeny nalezené stacionární úseky podle prahu kepstrální vzdálenosti.



Obrázek 5.4. Druhé nalezené úseky

Obrázek 5.5 zobrazuje potvrzené stacionární úseky znaménkovým testem.



Obrázek 5.5. Druhé potvrzené úseky

Tabulka překryvů vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	12.1 %	12.1 %	86 %	86 %	86.5 %	84.8 %
Znam. test	1.7 %	3.9 %	11.2 %	25.5 %	39.1 %	16.9 %

Tabulka 5.3. Tabulka druhých výsledků na 1 záznamu

Tabulka pro všechny záznamy při těchto parametrech vychází následovně:

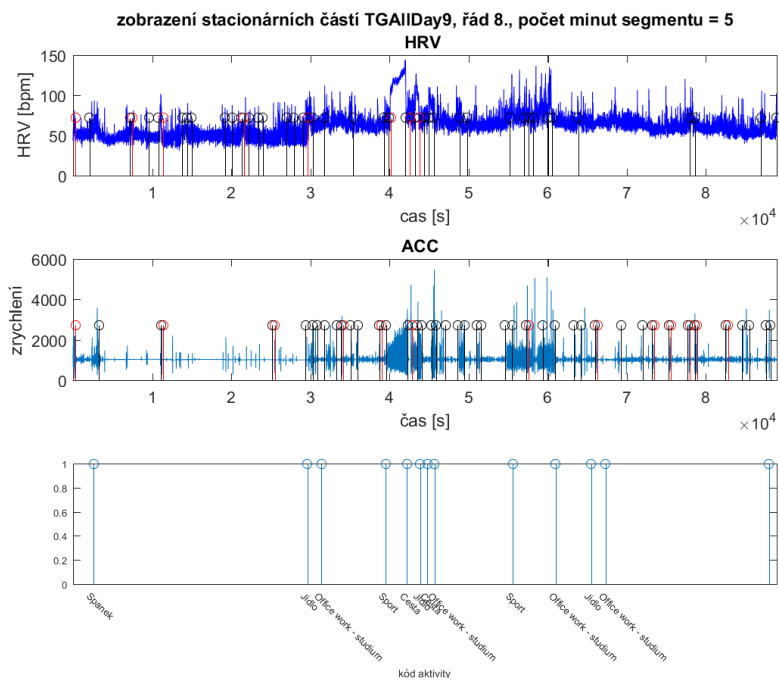
test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	12.6 %	12.6 %	84.7 %	84.4 %	85.6 %	84.1 %
Znam. test	3.1 %	7.8 %	14.7 %	37 %	38.1 %	15.2 %

Tabulka 5.4. Tabulka druhých výsledků na všech záznamech

5.1.2 Vyšší práh

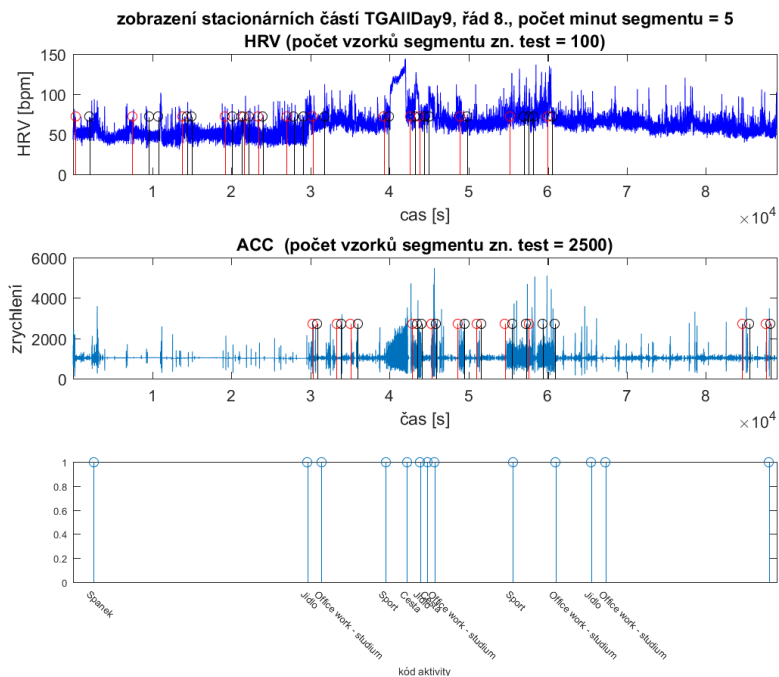
U vyššího prahu byl parametr a nastaven na hodnotu 1.

U následujících výsledků byla použita kepstrální vzdálenost s nulovým koeficientem. Na obrázku 5.6 jsou zobrazeny nalezené stacionární úseky podle prahu kepstrální vzdálenosti.



Obrázek 5.6. Třetí nalezené úseky

Obrázek 5.7 zobrazuje potvrzené stacionární úseky znaménkovým testem.



Obrázek 5.7. Třetí potvrzené úseky

Tabulka překryvů vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	8.8 %	8.7 %	96.1 %	95.2 %	96 %	95 %
Znam. test	6 %	9.3 %	25.4 %	39.5 %	22.6 %	14.2 %

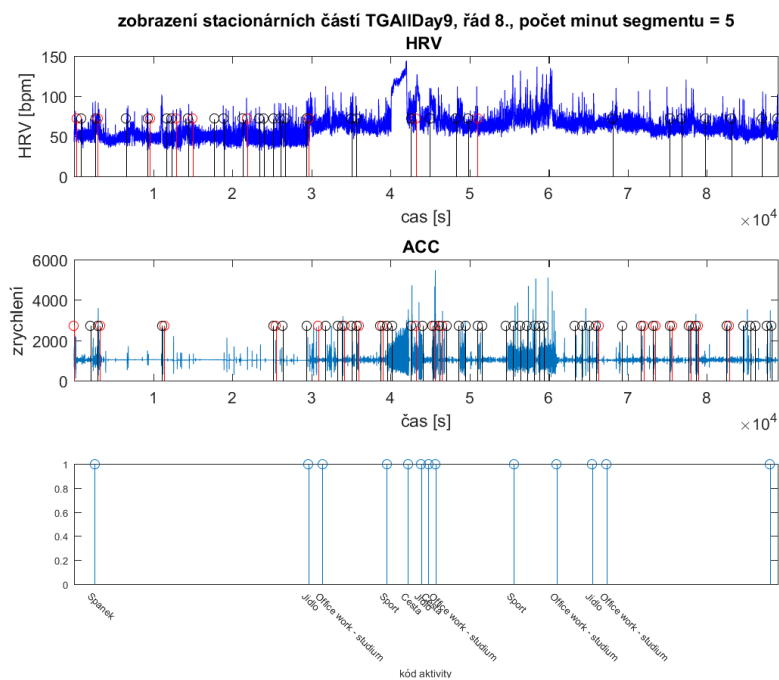
Tabulka 5.5. Tabulka třetích výsledků na 1 záznamu

Tabulka pro všechny záznamy při těchto parametrech vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	10.7 %	10.9 %	92.8 %	94.4 %	95.6 %	92.1 %
Znam. test	3.2 %	5.9 %	10.7 %	21.3 %	25.5 %	12.5 %

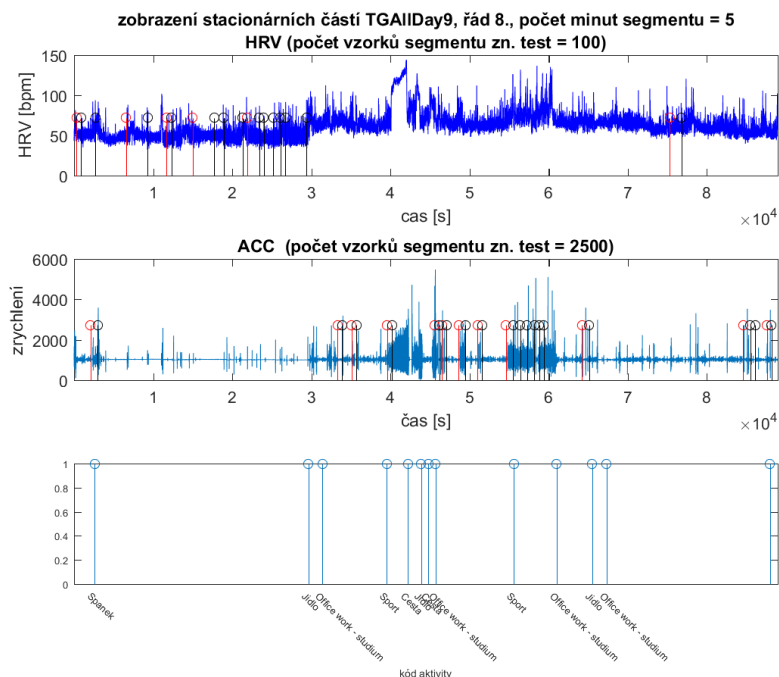
Tabulka 5.6. Tabulka třetích výsledků na všech záznamech

U následujících výsledků byla použita kepstrální vzdálenost bez nultého koeficientu. Na obrázku 5.8 jsou zobrazeny nalezené stacionární úseky podle prahu kepstrální vzdálenosti.



Obrázek 5.8. Čtvrté nalezené úseky

Obrázek 5.9 zobrazuje potvrzené stacionární úseky znaménkovým testem.



Obrázek 5.9. Čtvrté potvrzené úseky

Tabulka překryvů vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	2.5 %	2.5 %	93.2 %	93.7 %	94.6 %	92.4 %
Znam. test	0 %	0 %	28.6 %	4.5 %	23.6 %	14.6 %

Tabulka 5.7. Tabulka čtvrtých výsledků na 1 záznamu

Tabulka pro všechny záznamy při těchto parametrech vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	7.3 %	7.4 %	92 %	92.7 %	94.4 %	91.8 %
Znam. test	1.5 %	3.1 %	7 %	17.5 %	26.8 %	12.1 %

Tabulka 5.8. Tabulka čtvrtých výsledků na všech záznamech

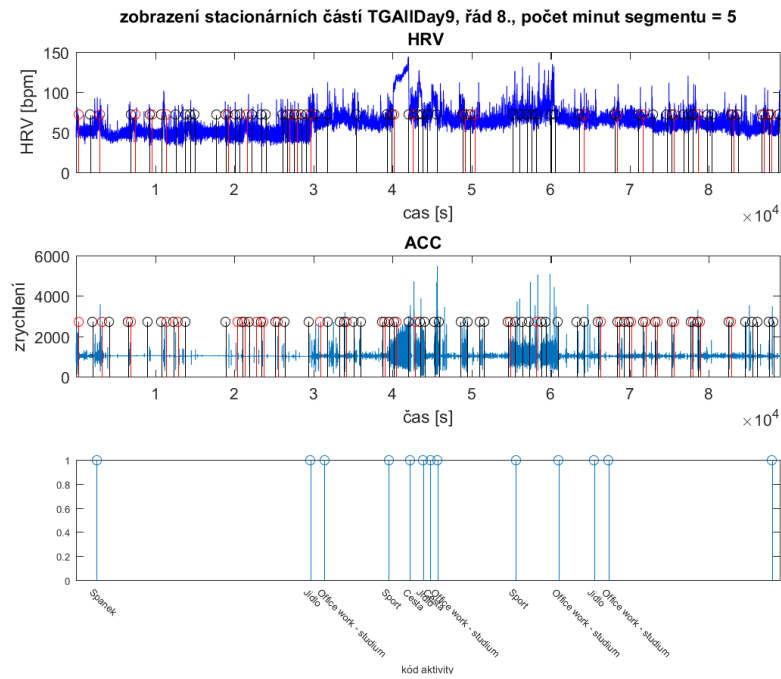
5.2 Dynamický práh

U dynamického prahu byla hodnota prahu vypočtena z předchozích 20 vzorků vypočtené kepstrální vzdálenosti. Pro výpočet prahu byla použita stejná rovnice jako u statického prahu. U prvních 20 vzorků kepstrální vzdálenosti byl práh stejný a byl vypočítán z prvních 20 vzorků kepstrální vzdálenosti.

5.2.1 Nižší práh

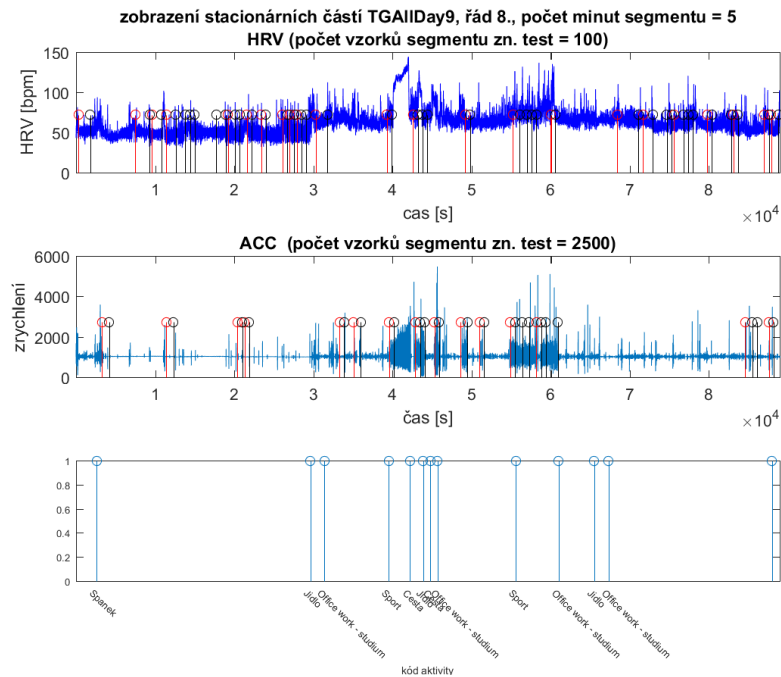
U nižšího práhu byl parametr a nastaven na hodnotu 0.5.

U následujících výsledků byla použita kepstrální vzdálenost s nultým koeficientem. Na obrázku 5.10 jsou zobrazeny nalezené stacionární úseky podle prahu kepstrální vzdálenosti.



Obrázek 5.10. Páté nalezené úseky

Obrázek 5.11 zobrazuje potvrzené stacionární úseky znaménkovým testem.



Obrázek 5.11. Páté potvrzené úseky

Tabulka překryvů vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	10.3 %	10.5 %	87.4 %	88.9 %	88 %	84.8 %
Znam. test	2.3 %	5.8%	16.2 %	40.3 %	43.8 %	17.2 %

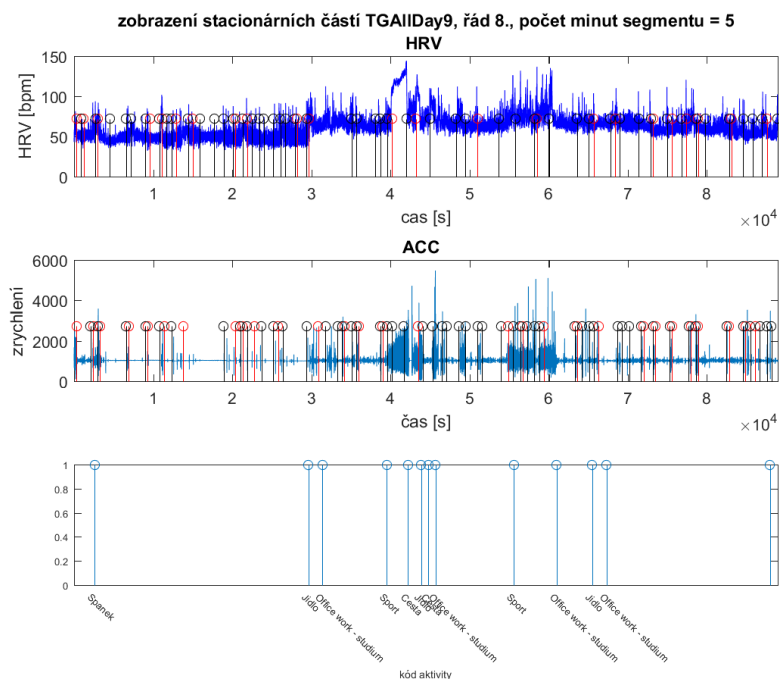
Tabulka 5.9. Tabulka pátých výsledků na 1 záznamu

Tabulka pro všechny záznamy při těchto parametrech vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	15.1 %	15.5 %	85.3 %	87.5 %	87.4 %	83.6 %
Znam. test	4.5 %	13.3 %	15.2 %	44.4 %	47.3 %	15.8 %

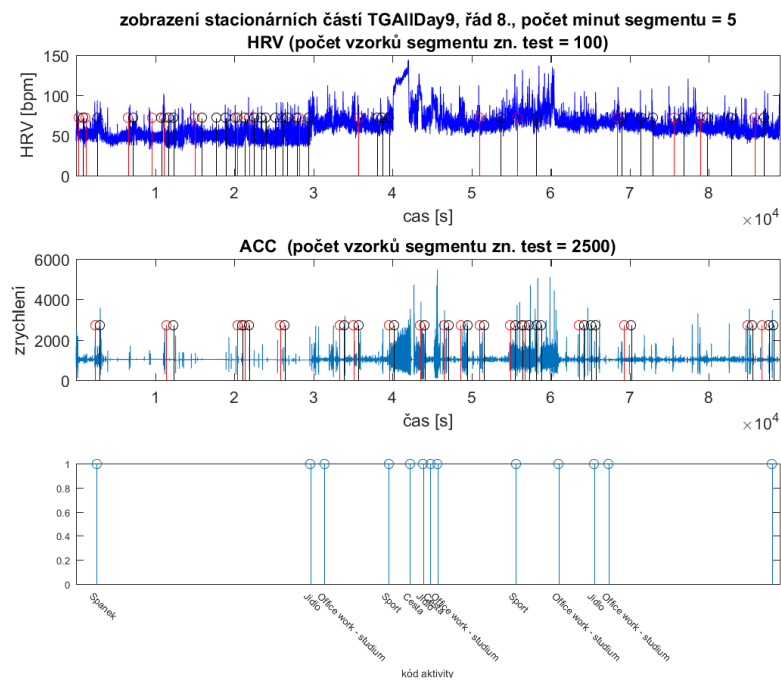
Tabulka 5.10. Tabulka pátých výsledků na všech záznamech

U následujících výsledků byla použita kepstrální vzdálenost bez nultého koeficientu. Na obrázku 5.12 jsou zobrazeny nalezené stacionární úseky podle prahu kepstrální vzdálenosti.



Obrázek 5.12. Šesté nalezené úseky

Obrázek 5.13 zobrazuje potvrzené stacionární úseky znaménkovým testem.



Obrázek 5.13. Šesté potvrzené úseky

Tabulka překryvů vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	9.4 %	9.7 %	83.6 %	87.1 %	86.5 %	81.5 %
Znam. test	3.1 %	7.3 %	16.5 %	38.2 %	42.8 %	18.2 %

Tabulka 5.11. Tabulka šestých výsledků na 1 záznamu

Tabulka pro všechny záznamy při těchto parametrech vychází následovně:

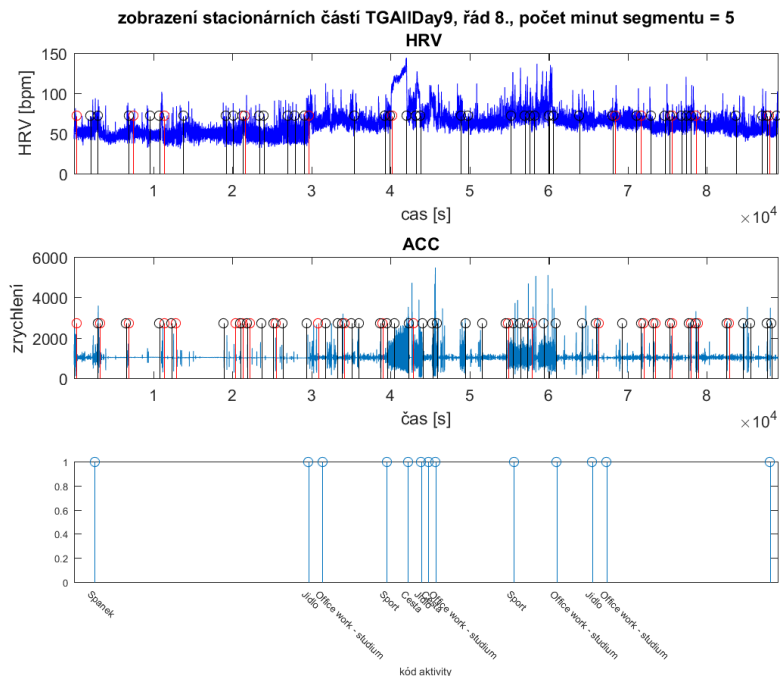
test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	12 %	12.4 %	83 %	85.5 %	85.9 %	81.8 %
Znam. test	5.1 %	11.8 %	17.2 %	43.2 %	41.2 %	16.3 %

Tabulka 5.12. Tabulka šestých výsledků na všech záznamech

5.2.2 Vyšší práh

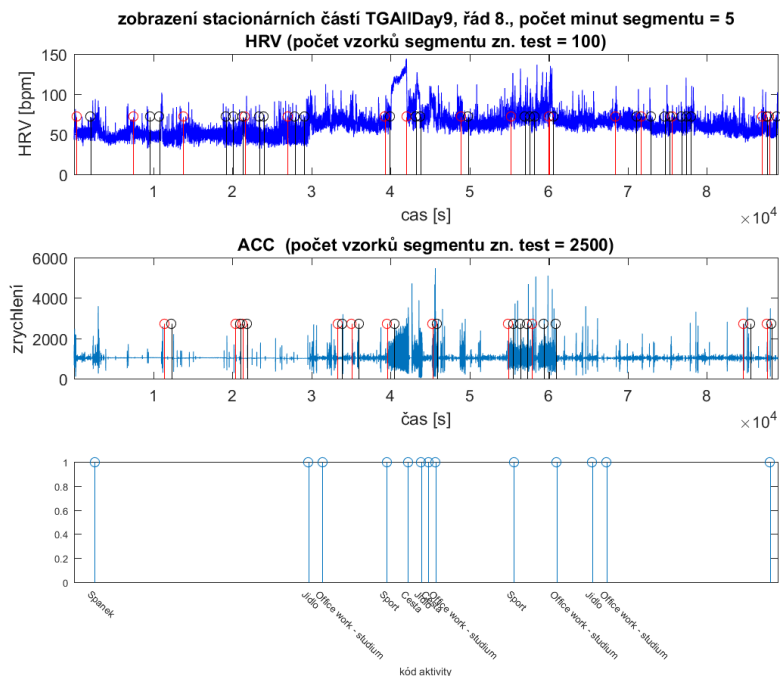
U vyššího prahu byl parametr a nastaven na hodnotu 1.

U následujících výsledků byla použita keprstrální vzdálenost s nulovým koeficientem. Na obrázku 5.14 jsou zobrazeny nalezené stacionární úseky podle prahu keprstrální vzdálenosti.



Obrázek 5.14. Sedmé nalezené úseky

Obrázek 5.15 zobrazuje potvrzené stacionární úseky znaménkovým testem.



Obrázek 5.15. Sedmé potvrzené úseky

Tabulka překryvů vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	4.6 %	4.9 %	89.6 %	93.9 %	94.4 %	88.4 %
Znam. test	0 %	0 %	13.2 %	37.5 %	38.4 %	13.2 %

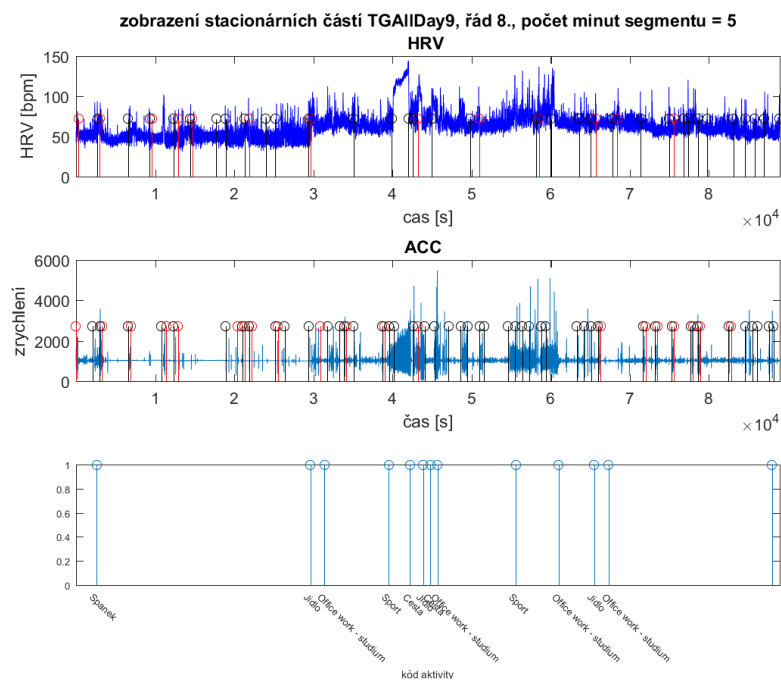
Tabulka 5.13. Tabulka sedmých výsledků na 1 záznamu

Tabulka pro všechny záznamy při těchto parametrech vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	11.7 %	11.8 %	91.1 %	92.9 %	93.9 %	90.3 %
Znam. test	3.8 %	9.4 %	13.1 %	36.9 %	36.5 %	12.6 %

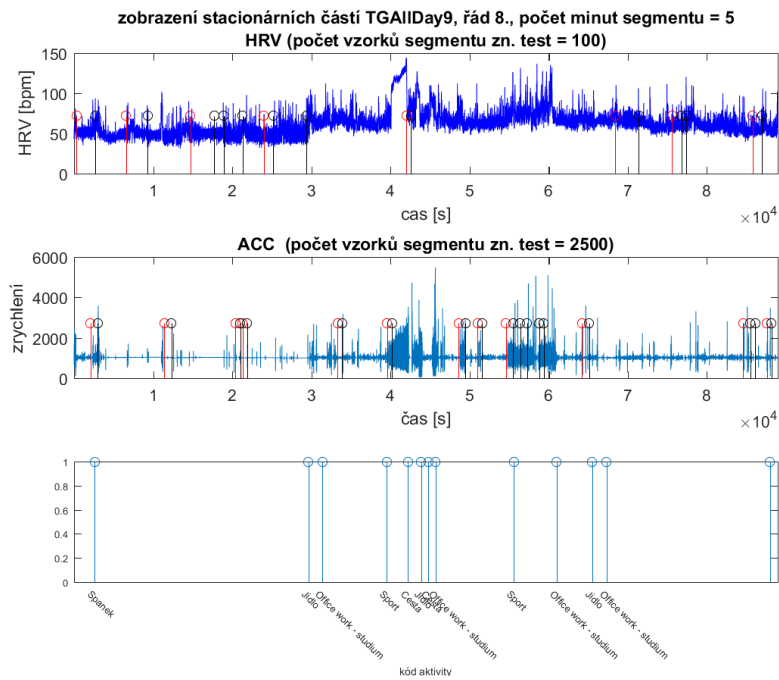
Tabulka 5.14. Tabulka sedmých výsledků na všech záznamech

U následujících výsledků byla použita kepstrální vzdálenost bez nultého koeficientu. Na obrázku 5.16 jsou zobrazeny nalezené stacionární úseky podle prahu kepstrální vzdálenosti.



Obrázek 5.16. Osmé nalezené úseky

Obrázek 5.17 zobrazuje potvrzené stacionární úseky znaménkovým testem.

**Obrázek 5.17.** Osmé potvrzené úseky

Tabulka překryvů vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	10.6 %	10.7 %	91.3 %	92.3 %	92.6 %	89.7 %
Znam. test	0 %	0 %	6.3 %	11.1 %	26.6 %	14.9 %

Tabulka 5.15. Tabulka osmých výsledků na 1 záznamu

Tabulka pro všechny záznamy při těchto parametrech vychází následovně:

test	Překryv HRV 1	Překryv Acc 1	Překryv HRV 2	Překryv Acc 2	Celý HRV	Celý Acc
Kepst. vzd.	9.9 %	10.1 %	89.6 %	91.8 %	93 %	89 %
Znam. test	1.8 %	5.5 %	8.2 %	24.3 %	32.7 %	11.2 %

Tabulka 5.16. Tabulka osmých výsledků na všech záznamech

5.3 Uložení dalších výsledků

Distribuce délky časových úseků je uložena pro každé nastavení v příloze CD. Některé zajímavé histogramy jsou zobrazeny v příloze C. Na CD se dále nachází výsledky ze všech použitých záznamů. U tabulek překryvů hodnota -1 znamená, že neexistovaly úseky pro porovnání. Tedy nebyl nalezen nebo potvrzen žádný nalezený úsek alespoň v jednom ze dvou typů dat u konkrétního záznamu. Na CD se nachází výsledky pro 12. řád AR modelu a kepstra.

Kapitola 6

Zhodnocení výsledků

Výsledky z kapitoly 5 ukazují, že znaménkový test zamítá úseky, které jsou až příliš dlouhé, takže většinou jsou potvrzené jen kratší úseky. Z tohoto důvodu je dobré volit spíše nižší práh než vyšší práh. Je patrné, že čím vyšší práh bude, tím delší budou nalezené úseky, které jsou označeny jako stacionární. Při vyšším práhu je stacionarita potvrzena jen u malého množství úseků. Daleko více úseků je potvrzených u záznamů z RR intervalů než u záznamů z akcelerometru. U dynamického práhu nižší hodnoty se jedná přibližně o polovinu potvrzených stacionárních úseků RR intervalů.

Dále je patrné, že u překryvu 2. typu vychází daleko vyšší hodnoty, než u překryvu 1. typu. Tedy jednotlivé stacionární úseky nezačínají u obou typů dat často ve stejných místech a zároveň jsou tyto úseky jen v malém množství případů stejně dlouhé. Z toho plyne, že se nedá očekávat, že v místě, kde je začíná stacionární úsek z jednoho typu dat, bude začínat stacionární úsek druhého typu dat. U překryvu 2. typu naopak tato závislost existuje, ale pouze u nalezených úseků pomocí prahování kepstrální vzdálenosti. Po potvrzení úseků znaménkovým testem vychází u překryvu 2. typu velmi nízké hodnoty, i když stále jsou tyto hodnoty o mnoho vyšší než u překryvu 1. typu.

Dále je patrné, že použití dynamického práhu vykazuje povětšinou lepší výsledky než statický práh, protože u statického je problém s tím, že pokud budou v nějaké oblasti hodně vysoké hodnoty, tak může být práh až příliš vysoko, a proto nezaregistruje menší změny v kepstrální vzdálenosti. Zatímco dynamický práh dokáže reagovat dobře na jakékoliv změny a pak při stejném parametru a bude více úseků označených za stacionární než u statického práhu. Navíc tyto úseky jsou většinou kratší, proto je větší pravděpodobnost, že tyto úseky budou potvrzeny znaménkovým testem.

Rozdíl u výsledků s nultým koeficientem a bez nultého koeficientu není příliš významný, takže se dá použít kepstrální vzdálenost s nultým i bez nultého koeficientu. Z výsledků lze poznat, že po použití na reálných datech a deších úsecích vychází kepstrální vzdálenost bez nultého koeficientu rozumněji, než u použití na uměle vytvořených záznamech v kapitole 4.2.

Ze všech předchozích faktů tedy vychází, že pomocí kepstrální vzdálenosti je shoda mezi stacionárními úseky v obou typech dat u celkového překrytí. Důležitý je také výsledek, že nebývají potvrzeny stacionární úseky na podobných místech pomocí znaménkového testu u obou typů dat. Současně je viditelné, že čím delší je úsek, tím menší je pravděpodobnost, že by skutečně byl tento úsek stacionární.

Kapitola 7

Diskuze

Do budoucna by bylo dobré pokusit se najít závislost mezi nestacionárními úseky. Jedná se o úseky, kde vychází vysoká kepstrální vzdálenost. Otázka je, jestli by se výsledky nějak výrazně lišily oproti zobrazeným výsledkům v této práci.

V budoucnu by bylo možná dobré použít i jiný test na ověření stacionarity, protože zmíněný ANOVA test je příliš kritický a znaménkový test není kritický jen v případě, kdy se počítá z malého množství segmentů. Znaménkový test potvrzuje hlavně kratší nalezené úseky. Pokud je ale znaménkový test proveden na velkém množství segmentů, tak je kritický také, i když stále méně než ANOVA test. K tomuto problému dochází i proto, že v každém segmentu při výpočtu výkonu vzniká určitá výpočetní chyba, která se zmenšuje s větším počtem vzorků, ze kterých je výkon počítán. Proto je pak možné, že při větším množství segmentů se tato chyba projeví více a zkrusluje dosažené výsledky. Existují i jiné stacionární testy, které jsou zmíněny v kapitole 1.4, které by bylo možné použít. I když například u použití testu, který je založen na spektrální výkonové hustotě by docházelo ke stejnému problému jako u znaménkového testu s tím, že by bylo malé množství vzorků v každém segmentu a výpočet spektrální výkonové hustoty by nebyl přesný.

Jelikož ne všechny zobrazené výsledky vychází ze stejného řádu AR modelu u obou typů záznamů, tak by bylo určitě vhodné podívat se na chování vytvořených metod v době, kdy mají oba typy záznamů jiný řád AR modelu. Řád kepstra ovšem musí zůstat vždy stejný u obou typů dat. Takhle varianta je možná u vytvořené grafické aplikace.

Všechny výsledky navíc byly použity bez překryvu segmentů, takže by bylo dobré se zaměřit na chování metod při použití překrytí jednotlivých segmentů. I tato možnost je přidána do grafické aplikace.

Další možností pro detekování změn v těchto datech by bylo použití jiných metod než kepstrální vzdálenosti, protože kepstrální vzdálenost vychází hlavně z řečových signálů. Těžko je možné přesně určit, jestli je nejvhodnější pro tento typ dat, ale několikrát už byl AR model na tento typ dat použit, takže se jedná o jednu z možností. Jako alternativa se nabízí například kontrola změny průměru a směrodatné odchylky v segmentech za sebou. U záznamů RR intervalů je možno se zaměřit také na parametry, které jsou počítány speciálně ze srdečního rytmu (viz kapitola 1.2).

Kromě těchto zmíněných postupů by bylo také dobré podívat se na vlastnosti signálů ve stacionárních a nestacionárních úsecích. Zde dochází k problému, že pokud těchto úseků bude příliš, tak vyjde velké množství dat, ve kterém bude složité se vyznat.

Další možností je také u záznamů z akcelerometru, místo metody pro spojení 3 os dohromady, použít například analýzu hlavních komponent a zjistit, jestli se výsledky budou výrazně lišit nebo ne. Také by mohlo být dobré používat AR model na každou osu zvlášť, zde ale spočívá problém v tom, že budou ve všech 3 osách nejspíš vycházet úseky v jiných časových oblastech a pak by bylo těžké porovnat, která osa je tedy nejlepší pro konečné výsledky. Navíc když se měřená osoba pohne, tak nemusí být zákonitě pohyb

ve všech osách stejný, což bylo vidět na jednotlivých záznamech, a proto byla použita zmíněná metoda pro spojení všech os dohromady.

Seznam aktivit byl u každého měření relativně malý, proto by možná bylo lepší použít i více zaznamenaných aktivit nebo mít větší počet změn během jednoho měření, a to z toho důvodu, aby jeden typ aktivity vyloženě nepřevažoval nad zbytkem. Také by bylo dobré porovnat výsledky na více naměřených osobách, protože výsledky nemusí vycházet pro každou osobu stejně.

Celkově by bylo možné použít tento algoritmus v reálném čase, kdy by byl používán dynamický práh u kepstrální vzdálenosti. Pro tento účel by ale bylo vhodné brát segmenty kratší než 5 minut. V této práci pro všechny výsledky byl použit segment délky 5 minut. Z důvodu, aby byl dynamický práh použitelný, tak by bylo dobré mít přibližně 20 naměřených segmentů, mezi kterými se dá vypočítat kepstrální vzdálenost. U segmentu délky 5 minut by ovšem byl problém v tom, že buď by byla potřeba velký překryv segmentů, nebo bez překryvu by se dynamický práh musel počítat ze záznamu dlouhého téměř 2 hodiny. Zde ale bylo 5 minut vybíráno vzhledem k tomu, že u analýzy srdečního rytmu je to obecně doporučená délka segmentu. U akcelerometru by tedy nemusel být potřeba tak dlouhý čas pro 1 segment.

Závěr

V práci byl vytvořen algoritmus pro analýzu srdeční variability na základě použití autoregresního a kepstrálního modelu, kdy v kapitole 4 je potvrzeno, že algoritmus skutečně funguje. Tento algoritmus byl otestován pro různé parametry tak, aby se dala zjistit vhodně stacionarita jednotlivých segmentů dat.

Z výsledků vychází, že mezi nalezenými úseky, kde je nízká kepstrální vzdálenost, je určitá spojitost u obou typů dat, protože se úseky překrývají. Po provedení testu na stacionaritu pomocí znaménkového testu, už tato závislost není zřejmá, protože jsou potvrzeny úseky v jiných časových oblastech u obou typů záznamů. Znaménkový test je kritický u delších úseků, kratší úseky ve většině případů prochází. Druhý použitý test, ANOVA test, na stacionaritu je daleko více kritický. ANOVA test není vhodný pro použitý typ záznamů v této práci.

Z výsledků dále vychází, že je lepší používat dynamický práh u kepstrální vzdálenosti. Tento práh má navíc výhodu v tom, že je možné ho použít v reálném čase. Tedy se dá použít za běhu měření srdeční aktivity a pohybu na rozdíl od statického prahu u kepstrální vzdálenosti.

Kromě vytvořeného algoritmu také byla vytvořena jednoduchá grafická uživatelská aplikace, ve které je možné nastavovat různé parametry a zjišťovat výsledky.

Literatura

- [1] Tim Taylor. *Heart*.
<http://www.innerbody.com/image/card01.html>.
- [2] Jindřich Špinar a Jiří Vítovec. Tepová frekvence a kardiovaskulární onemocnění. *Interní medicína pro praxi*. 2009, 7 (11), 315-318.
- [3] Pavel Osmančík. *EKG UČEBNICE*.
<http://www.ucebnice-ekg.cz/index.php>.
- [4] Jan Habásko. *Elektrokardiogram (EKG)*.
<http://www.myokarditida.cz/diagnostika-a-lecba/diagnostika/elektrokardiogram-ekg/>.
- [5] Josef Kautzner. *Poruchy srdečního rytmu – arytmie*.
<http://www.ikem-kardiologie.cz/cs/pro-pacienty/co-u-nas-lecime/poruchy-srdecniho-rytmu--arytmie/>.
- [6] Adam Thompson. *The Electrocardiogram - Part V*.
<http://paramedicine101.blogspot.cz/2009/09/electrocardiogram-part-v.html>.
- [7] Marek Malik. Heart rate variability. *Annals of Noninvasive Electrocardiology*. 1996, 1 (2), 151–181. DOI 10.1111/j.1542-474X.1996.tb00275.x.
- [8] Hui-Min Wang a Sheng-Chieh Huang. SDNN/RMSSD as a Surrogate for LF/HF: A Revised Investigation. *Modelling and Simulation in Engineering*. 2012, 2012 DOI 10.1155/2012/931943.
- [9] BIOPAC Systems inc. *RMSSD for HRV Analysis*.
<https://www.biopac.com/application/ecg-cardiology/advanced-feature/rmssd-for-hrv-analysis/>.
- [10] Christoph Guger. *Heart-rate Variability presentation*.
http://www0.cs.ucl.ac.uk/research/vr/Projects/Presencia/ConsortiumPublications/graz_papers/HRVanalysis.pdf.
- [11] G. P. Nason. *Stationary and non-stationary time series*. In: *Statistics in Volcanology*. 2010. 11.
- [12] Jiří Anděl. *Matematická statistika*. SNTL, Praha, 1985.
- [13] Mirko Navara. *Pravděpodobnost a matematická statistika*. 1. vydání. vydavatelství ČVUT, Praha, 2007. ISBN 978-80-01-03795-9.
- [14] Pavel Sovka a Jan Uhlíř. *Číselkové zpracování signálů*. 2. přepracované vydání. vydavatelství ČVUT, Praha, 2002. ISBN 80-01-02613-6.
- [15] Richard Shiavi. *Biomedical Engineering: Introduction to Applied Statistical Signal Analysis: Guide to Biomedical and Electrical Engineering Applications*. 3. vydání. Academic Press, Burlington, US, 2010.
- [16] David M. Lane. *Online Statistics Education: An Interactive Multimedia Course of Study*.
http://onlinestatbook.com/2/analysis_of_variance/ANOVA.html.

- [17] M. B. Priestley a T. Subba Rao. A Test for Non-Stationarity of Time-Series. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1969, 31 (1),
- [18] NIST/SEMATECH. *e-Handbook of Statistical Methods*.
<http://www.itl.nist.gov/div898/handbook/eda/section3/eda357.htm>.
- [19] Pavel Sovka a Petr Pollák. *Vybrané metody číslicového zpracování signálů*. 2. přepracované vydání. vydavatelství ČVUT, Praha, 2003. ISBN 80-01-02821-6.
- [20] A. N. Kalinichenko, M. I. Nilicheva, S. V. Khasheva, O. D. Yurieva a O. V. Mamonov. Signal stationarity assessment for the heart rate variability spectral analysis. 2008, 965-968. DOI 10.1109/CIC.2008.4749204.
- [21] M. E. Gomes, A. V. Souza, H. M. Guimaraes a L. A. Aguirre. Investigation of determinism in heart rate variability. *Chaos: An Interdisciplinary Journal of Non-linear Science*. 2000, 10 (2), DOI 10.1063/1.166507.
- [22] Ravi Ramachandran a Richard Mammone. *Modern Methods of Speech Processing*. Springer Science & Business Media, 2012. ISBN 9781461522812.
- [23] E Gil, M Orini, R Bailón, J M Vergara, L Mainardi a P Laguna. Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions. *Physiological Measurement*. 2010, 31 (9), 1271.
- [24] MIKIO TAKADA, TAKESHI EBARA, YASUSHI SAKAI a YASUHIRO KUWANO. STATIONARITY OF THE HEART RATE VARIABILITY BY ACCELERATION PLETHYSMOGRAPHY: SHORT-TERM MEASUREMENTS OF HEALTHY YOUNG MALES IN DAILY LIFE. *Journal of Human Ergology*. 2009, 38 (2), 41-50. DOI 10.11183/jhe.38.41.
- [25] M.-K. Yum, J.-T. Kim a H.-S. Kim. Increased non-stationarity of heart rate during general anaesthesia with sevoflurane or desflurane in children. *British Journal of Anaesthesia*. 2008, 100 (6), 772-779. DOI 10.1093/bja/aen080.
- [26] Arthur Charpentier. *Unit root tests*.
<http://freakonometrics.hypotheses.org/12729>.
- [27] Tommaso Costa, Giuseppe Boccignone a Mario Ferraro. Gaussian Mixture Model of Heart Rate Variability. *PLoS ONE*. 2012, 7 (5), 1-9. DOI 10.1371/journal.pone.0037731.
- [28] Davrondzhon Gafurov, Kirsi Helkala a Torkjel Søndrol. Biometric Gait Authentication Using Accelerometer Sensor. *Journal of Computers*. 2006, 1 (7),
- [29] J. Mateo a P. Laguna. Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Transactions on Biomedical Engineering*. 2003, 50 (3), 334-343. DOI 10.1109/TBME.2003.808831.

Příloha A

Seznam symbolů a zkratek

V této příloze je vypsán seznam použitých zkratek a symbolů.

EKG	elektrokardiogram nebo elektrokardiograf
ACC	akcelerometr
ANOVA	analýza rozptylu
AR	autoregrese, často ve spojení AR model, který vyznačuje autoregresní model
MA	moving average, klouzavé součty
HRV	variabilita srdečního rytmu
MATLAB	Matrix Laboratory - software a skriptovací jazyk hlavně pro práci s maticemi
SA	ve spojení SA uzel - sinoatriální uzel
n. u.	norm unit, normovaná jednotka
QRS	často ve spojení QRS komplex, označení pro stah srdce v EKG záznamu
RR	často ve spojení RR interval, označení pro vzdálenost dvou špiček QRS komplexů
ST	ve spojení ST segment - depolarizace komor
NN	ve spojení NN interval - jiné označení pro RR interval
SDNN	směrodatná odchylka RR intervalů
RMSSD	druhá odmocnina průměrného rozdílu přilehlých RR intervalů
SDANN	směrodatná odchylka průměrných RR intervalů vypočítaných v segmentu určité délky
SDSD	směrodatná odchylka rozdílů mezi přilehlými RR intervaly
NN50	počet RR intervalů, které se oproti následujícím liší o 50 a více ms
pNN50	počet NN50 intervalů vydělený celkovým počtem RR intervalů v záznamu
TINN	základní šířka při aproximaci trojúhelníkové distribuce RR intervalů vycházející z histogramu
P	P vlna u EKG záznamu
T	T vlna u EKG záznamu
ms	milisekunda
φ	řecké písmeno fi
e	Eulerovo číslo
Hz	Hertz - jednotka frekvence
ARMA	autoregresní model klouzavých součtů
V LF	výkon ve velmi nízkém frekvenčním pásmu
LF	výkon v nízkém frekvenčním pásmu
HF	výkon ve vyšším frekvenčním pásmu
ULF	výkon v ultra nízkém frekvenčním pásmu
LSD	logaritmická spektrální vzdálenost
SDF	formát souboru - SQL Server Compact Edition Database File
CSV	formát souboru - comma separated values
α	hladina významnosti statistického testu

Příloha B

Obsah přiloženého CD

V této příloze je vypsán seznam složek a souborů v hlavní složce na přiloženém CD. K vypsáným složkám a souborům je přidán krátký komentář.

Soubor/Složka	Popis
info.pdf	Soubor s informacemi o jednotlivých souborech a složkách
dokumentaceGUI.pdf	Soubor s dokumentací ke grafické aplikaci
Tomas_Grosman_Diplomova_Prace_2016.pdf	Soubor celé vypracované diplomové práce
results/	Složka s jednotlivými výsledky
figures/	Složka se zajímavými grafy z práce
scripts/	Složka, která obsahuje naprogramované skripty a jednotlivá použitá data

Tabulka B.1. Přehled souborů a složek na přiloženém CD v hlavní složce

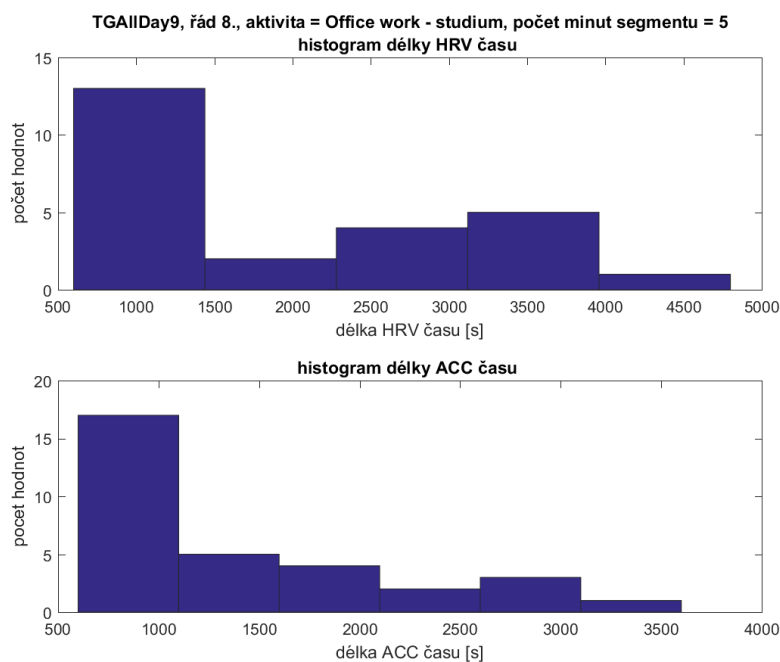
Příloha C

Histogramy při jednotlivých aktivitách

V této příloze jsou některé zajímavé histogramy délky časů stacionárních úseků u jednotlivých aktivit.

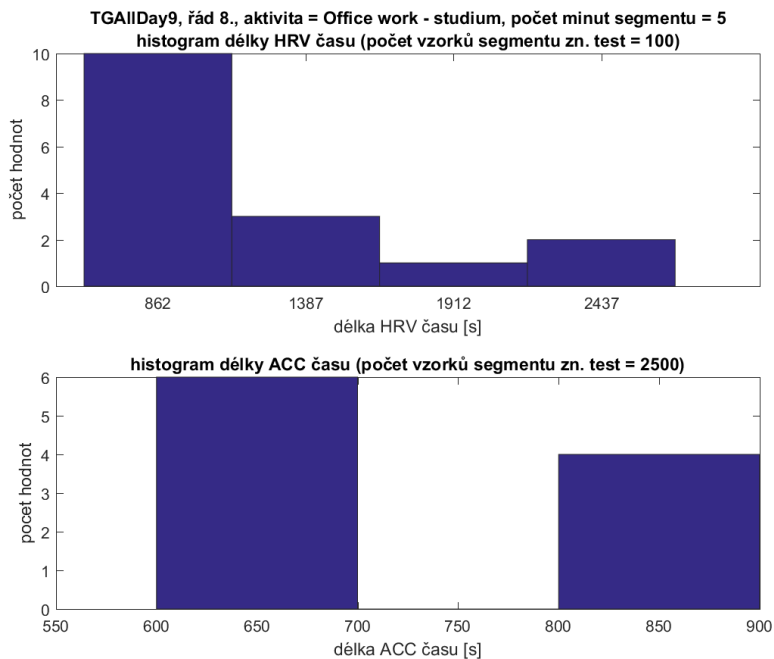
Na následujících histogramech je použit dynamický práh pro kepstrální vzdálenost s nultým koeficientem, parametr a se rovná hodnotě 0.5. Byl zvolen 8. řád kepstra a AR modelu. Pro znaménkový test byl segment o délce 100 vzorků u RR intervalů a 2500 vzorků u akcelerometru.

Na obrázku C.1 je zobrazen histogram úseků po použití dynamického prahu u malé fyzické aktivity (práce, studium).



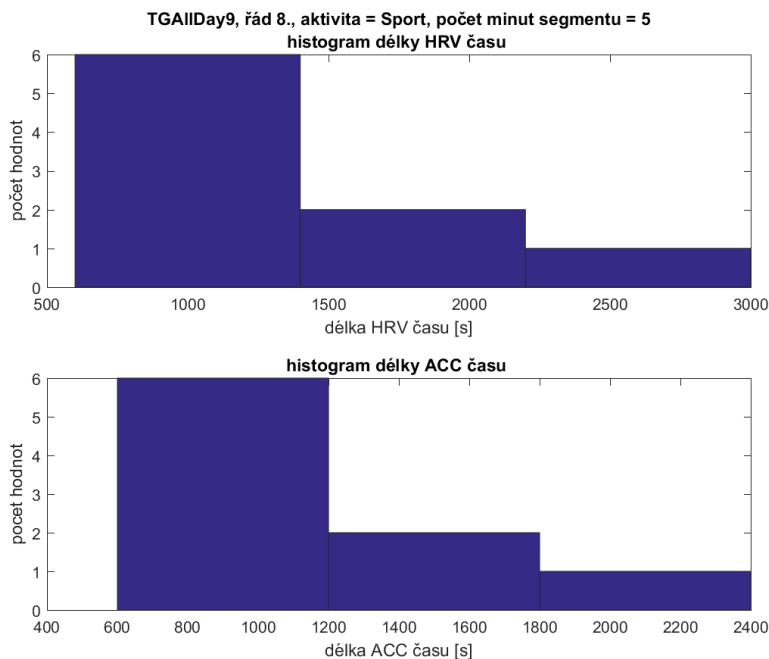
Obrázek C.1. Histogram úseků po použití dynamického prahu u malé fyzické aktivity

Na obrázku C.2 je zobrazen histogram úseků po použití znaménkového testu u malé fyzické aktivity (práce, studium).



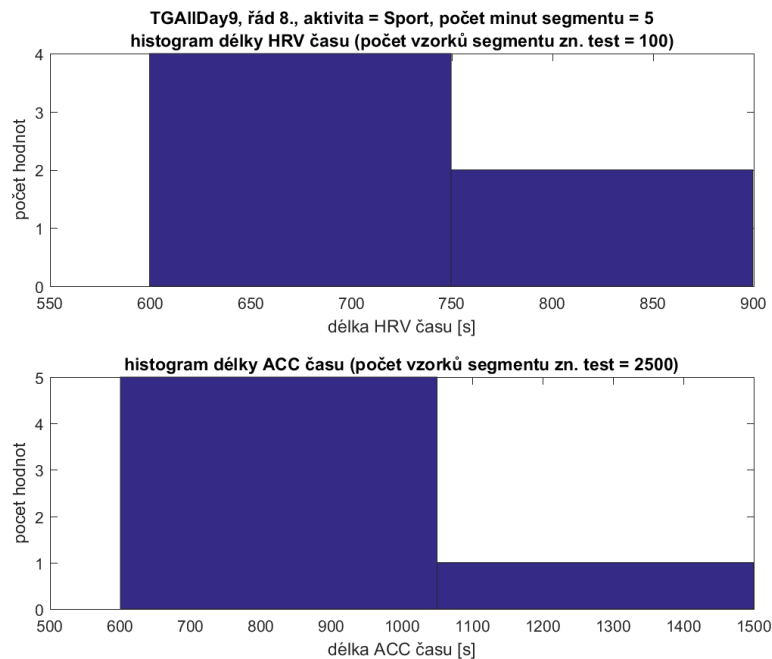
Obrázek C.2. Histogram úseků po použití znaménkového testu u malé fyzické aktivity

Na obrázku C.3 je zobrazen histogram úseků po použití dynamického prahu u sportu.



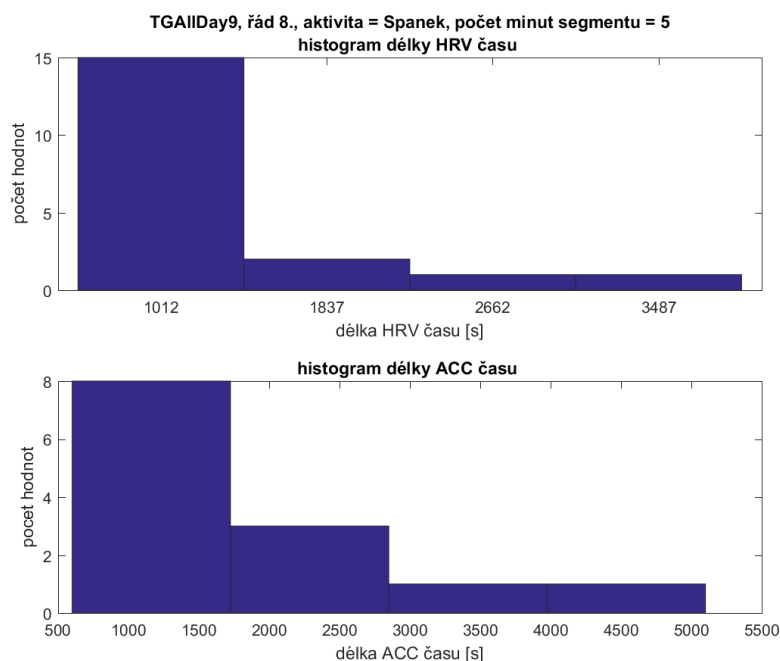
Obrázek C.3. Histogram úseků po použití dynamického prahu u sportu

Na obrázku C.4 je zobrazen histogram úseků po použití znaménkového testu u sportu.



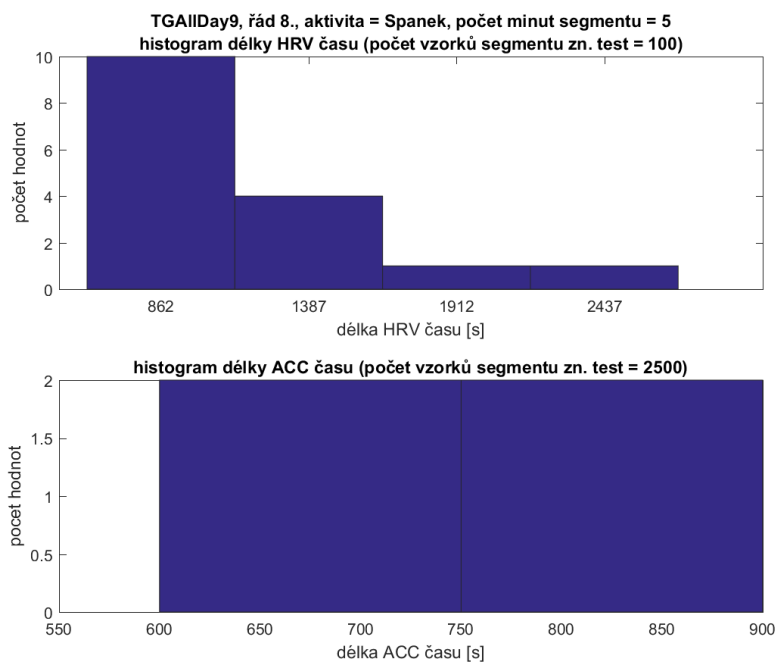
Obrázek C.4. Histogram úseků po použití znaménkového testu u sportu

Na obrázku C.5 je zobrazen histogram úseků po použití dynamického prahu ve spánku.



Obrázek C.5. Histogram úseků po použití dynamického prahu ve spánku

Na obrázku C.6 je zobrazen histogram úseků po použití znaménkového testu ve spánku.



Obrázek C.6. Histogram úseků po použití znaménkového testu ve spánku