Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií

FIIT-5212-102927

Miroslav Hájek

Spracovanie dát generovaných senzorovou IoT sieťou

Bakalárska práca

Vedúci práce: Ing. Marcel Baláž, PhD.

Máj 2022

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií

FIIT-5212-102927

Miroslav Hájek

Spracovanie dát generovaných senzorovou IoT sieťou

Bakalárska práca

Študijný program:	Informatika
Študijný odbor:	Informatika
Miesto vypracovania:	Ústav počítačového inžinierstva a aplikovanej informatiky
Vedúci práce:	Ing. Marcel Baláž, PhD.
Pedagogický vedúci:	Ing. Jakub Findura

Máj 2022

Slovenská technická univerzita v Bratislave
Ústav počítačového inžinierstva a aplikovanej informatiky

Fakulta informatiky a informačných technológií
Akademický rok: 2021/2022
Evidenčné číslo: FIIT-5212-102927

ZADANIE BAKALÁRSKEJ PRÁCE

Študent:	**Miroslav Hájek**
ID študenta:	102927
Študijný program:	informatika
Študijný odbor:	informatika
Vedúci práce:	Ing. Marcel Baláž, PhD.
Vedúci pracoviska:	Ing. Katarína Jelemenská, PhD.
Pedagogický vedúci práce:	Ing. Jakub Findura

Názov práce: **Spracovanie dát generovaných senzorovou IoT sieťou**

Jazyk, v ktorom sa práca vypracuje: slovenský jazyk

Špecifikácia zadania:

Senzorové IoT siete sa stali bežnou súčasťou rôznych priemyselných procesov. Ich primárnou úlohou je zbieranie rôznorodých dát z prostredia, ich ukladanie a vyhodnocovanie v reálnom čase. Analyzovanie a vyhodnocovanie dát pri nepredržitom monitorovaní už len z malého množstva senzorov predstavuje veľkú vyzvu. Senzory produkujú veľké množstvo dát a anomálie nemusia byť na prvý pohľad detegovateľné. Cieľom projektu je analyzovať dáta zachytené senzorovou sieťou. Analyzovať algoritmy na ich ukladanie a spracovanie. Analyzujte jednotlivé úrovne senzorovej siete a identifikujte miesta, kde by sa dali dáta čiastočne spracovať. Na základe analýzy navrhnite spôsob ukladania a spracovania dát, prípadne optimalizáciu toku dát pre existujúcu senzorovú sieť. Vaše riešenie implementujte a otestujte jeho funkčnosť.

Rozsah práce: 40

Termín odovzdania bakalárskej práce:	16. 05. 2022
Dátum schválenia zadania bakalárskej práce:	23. 11. 2021
Zadanie bakalárskej práce schválil:	doc. Ing. Valentino Vranić, PhD. – garant študijného programu

Čestné prehlásenie

Čestne vyhlasujem, že som túto prácu vypracoval samostatne, na základe konzultácií a s použitím uvedenej literatúry.

V Bratislave, 1.5.2022 ....

Miroslav Hájek

Poďakovanie

Text poďakovania

Anotácia

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií

Študijný program:	Informatika
Autor:	Miroslav Hájek
Bakalárska práca:	Spracovanie dát generovaných senzorovou IoT sieťou
Vedúci bakalárskej projektu:	Ing. Marcel Baláž, PhD.
Pedagogický vedúci:	Ing. Jakub Findura

Máj 2022

Stručná charakteristika zadania bakalárskeho projektu ale predovšetkým výsledkov bakalárskeho projektu v slovenskom a anglickom jazyku každá v rozsahu max. 1 strany A4 (hlavička + cca 150-200 slov).

Annotation

Slovak University of Technology Bratislava

Faculty of Informatics and Information Technologies

Degree course:	Informatics
Author:	Miroslav Hájek
Bachelor's Thesis:	Data Processing for Sensor IoT Network
Supervisor:	Dr. Marcel Baláž
Departmental advisor:	Jakub Findura

2022, May

Annotation text in English, 150-200 words.

Obsah

1 Úvod

2 Analýza

2.1 Monitorovanie vibrácií a šoku

Vibrácie sú periodickým kmitaním hmoty okolo rovnovážnej polohy vznikajúce excitáciou látky, ktorej je dodaná potenciálna energia, a zo zákona zachovania energie je následne premieňaná na kinetickú energiu. V realite dochádza pôsobením trenia k útlmu voľného oscilačného pohybu s časom a pohybová energia sa uvoľňuje v podobe tepelnej alebo akustickej emisie do okolitého prostredia. Častejšie ako presné harmonické kmity sú pozorované náhodné vibrácie, ktorých vývoj nevieme dopredu predvídať. Naproti tomu šok, alebo aj prechodový jav, je náhle uvoľnenie kinetickej energie krátkeho trvania oproti prirodzenej oscilácii systému.

Význam a dôležitosť sledovania vibrácií spočíva v ich výskyte u každého mechanického zariadenia a je zapríčinená pohybom jednotlivých súčiastok a trením v ložiskách. Ich nadmerná prítomnosť býva spôsobená opotrebením dielov stroja alebo nevyvážením rotačných častí, zakliesňovaním ozubených kolies, ako dôsledkoch iných technických defektov. V prevažnej väčšine prípadov ide o nežiaduci jav nakoľko zakladá zníženiu účinnosti so zvýšením hlučnosti ako vedľajšiemu produktu.

Ďalšou oblasťou hojnej prítomnosti vibrácií je preprava osôb alebo tovaru cestnými a železničnými dopravnými prostriedkami, kde sú zapríčinené nerovnosťami povrchu vozovky alebo koľaje v bode styku s kolesami vozidla. Na zvýšenie ovládateľnosti vozidla a komfortu pasažierov sú kabíny odpružené od kolies tlmičmi. Lietadlá sú zasa pod vplyvom trenia vzduchu s trupom a krídlami konštrukcie, ktoré je ďalej zosilnené vzdušnými prúdmi a turbulenciami.

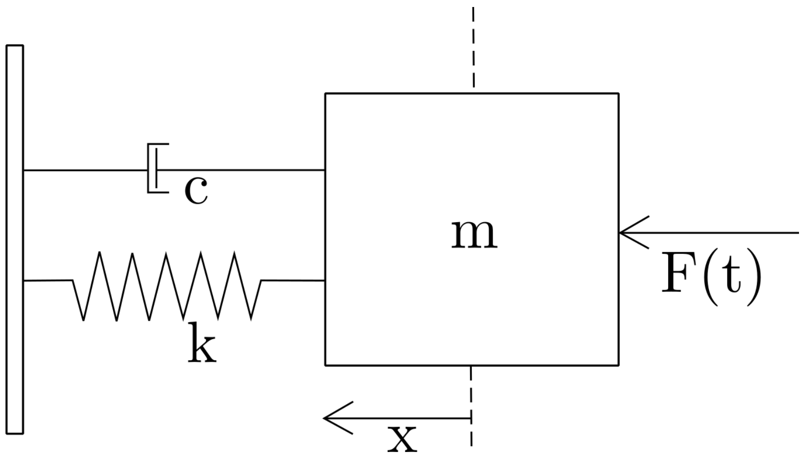
Druhým významným faktorom podieľajúci sa na tvorbe vibrácii je aparát, ktorý uvádza vozidlo do pohybu alebo zastavuje, čiže hnací najčastejšie spaľovací, dieselový alebo elektrický motor a brzdový systém. Jedná sa najmä o vplyv pohybu piestov, alebo rotora u elektrických vozidiel, a prenosu otáčavého pohybu motora cez oje hriadeľa na nápravy. ABS brzdový systém prítomný pri väčšine automobilov zabraňujú šmyku striedavým zomknutím a uvoľňovaním brzdových kotúčov, čo má tiež vplyv na podmienky počas jazdy.

Detekciou nežiaducich vibrácií v preprave sa dokáže zabezpečiť aj bezpečnosť pasažierov včasnou výmenou súčiastky, ktorá by ovplyvnila prevádzkyschopnosť v kritických momentoch. Ich eliminácia dokáže predísť nenávratnému poškodeniu krehkých materiálov alebo znehodnoteniu reaktívnych substancií, či dokonca ich aktivácii v prípade výbušnín a pyrotechniky.

V neposlednom rade sú vibrácie súčasťou potenciálne nebezpečných prírodných úkazov a ich správna identifikácia má za následok varovania pre preventívnu evakuáciu obyvateľstva v oblasti, ktoré bude zasiahnutá zemetrasením, či erupciou sopky vedúcimi k ohrozenia zdravia osôb a poškodenia majetku.

2.1.1 Meranie fyzikálnej veličiny akcelerácie

Pohyb mechanického systému vystaveného vonkajším silám sa nazýva odozva, ktorej správanie opisuje zjednodušený model s jedným stupňom voľnosti (1DOF) kmitajúceho telesa s pružinou a tlmičom [1].



Obr. 2.1: Model oscilujúceho systému s pružinou a tlmičom

Pri pôsobení vonkajšej sily F na hmotu upevnenú na pružine vznikajú nútené vibrácie, ktoré ju vychyľujú z rovnovážnej polohy. Uvedená sila je charakterizovaná druhým Newtonovým zákonom v tvare $F = ma$, kde m je hmotnosť telesa a a predstavuje zrýchlenie. V protismere pôsobí sila vyvolaná pružinou $F_s = -kx$ a tlmiacim členom $F_d = -cv$, kde k je tuhosť pružiny ovplynená jej konštrukciou, c je tlmiaci koeficient, x je vychýlenie z rovnovážneho stavu, a v rýchlosť vychýlenia.

Fyzickým obmedzením telesa, ktorým je viazaný na pevnú podložku dochádza pri zanedbaní deformácie k takmer zaručenému návratu do rovnovážnej polohy a to nám umožňuje merať intenzitu vibrácií cez zrýchlenie ťažidla. Výslednú silu v jednom smere získame sčítaním síl podieľajúcich sa na dynamike telesa.

$$F(t) = ma - cv - kx \tag{2.1}$$

Pri použití trojosového akcelerometra, kedy sú evidované všetky tri priestorové súradnice časovo-premennej akcelerácie dostávame nasledujúcu rovnicu vo vektorovom tvare:

$$\vec{a}(t) = \frac{\vec{F}(t)}{m} \tag{2.2}$$

Magnitúda akcelerácie s troma súradnicami je daná L_2 normou vektora $\vec{a} = (a_x, a_y, a_z)$:

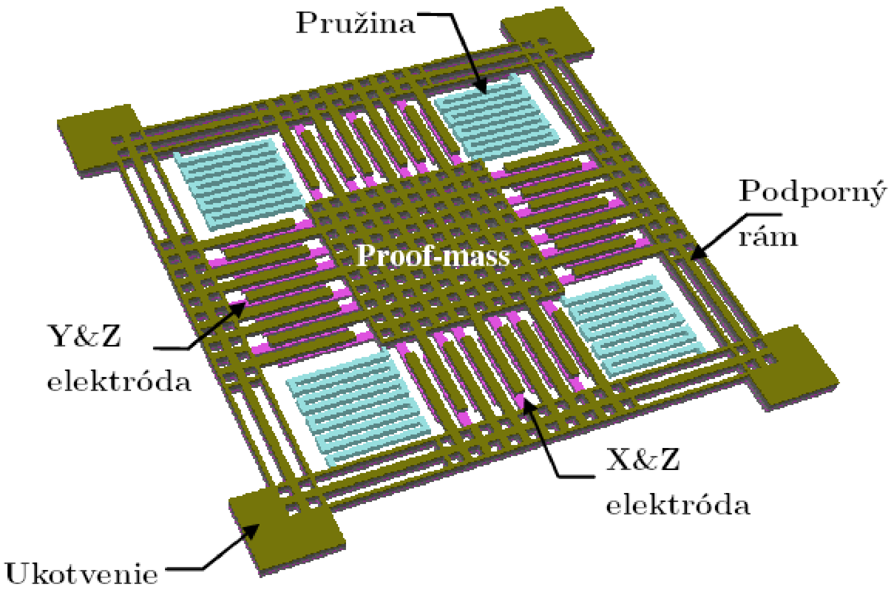
$$|a| = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{2.3}$$

2.1.2 MEMS kapacitný akcelerometer

Bežné inerciálne senzory na meranie zrýchlenia priamočiareho, ale aj rotačného pohybu (gyroskop), sa vyrábajú technológiou *MEMS – mikromechanický systém*, kedy je celé zariadenie vrátane všetkých mechanických súčastí umiestnené na kremík procesom mikrovýroby vo viacerých vrstvách. Sila spôsobujúca zrýchlenie je potom meraná vychýlením vstavanej odpruženej hmoty vzhľadom na pevné elektródy, ktoré môžu byť usporiadané jednostranne alebo ako diferenčný pár [2].

Pri diferenčnom páre spôsobí pohyb doštičky ťažidla medzi elektródami zmenu kapacít a ich rozdielom je možné zistiť aplikovanú silu a cez uvedený vzťah zrýchlenie. Na zvýšenie celkovej kapacity sa používa viacero párov elektród zapojených paralelne. Pred prevodom na číslicový signál musí napäťová úroveň zo senzora prejsť úpravou zahŕňajúcou nábojovocitlivý predzosilňovač, osovú demoduláciu a anti- aliasingové filtrovanie.

Viacosové akcelerometre vyžadujú viaceré opísané štruktúry orientované kolmo



Obr. 2.2: Mikroštruktúra 3DOF MEMS kapacitného akcelerometra [3]

na seba, podľa obr. 2.2, s ohľadom na počet vyžadovaných stupňov voľnosti, pričom v skutočných senzoroch vždy existuje aspoň minimálna závislosť medzi osami rádovo najviac v jednotkách percent. Teplota ovplyvňuje citlivosť MEMS akcelerometrov len nepatrne v stotinách percenta na stupeň Celzia.

Akcelerometre sa odlišujú v niekoľkých dôležitých vlastnostiach, ktoré zvyknú byť nastaviteľné vo výrobcom stanovenom rozsahu prípustných hodnôt s príslušnými toleranciami [4].

Citlivosť stanovuje najmenšiu rozlíšiteľnú zmenu v odčítanom napätí ku zmene externého pohybu respektíve zrýchlenia. Uvádza sa v jednotkách mV/g (milivolt na tiažové zrýchlenie) pri analógovom výstupe, alebo mg/LSB (mili-g na najmenej významový bit). pri senzoroch so vstavaným analógovo-digitálnym prevodníkom. Jednotka mg/LSB vyjadruje o koľko sa zmení zrýchlenie keď zvýšime alebo ponížime binárne číslo na výstupe o jedna. Niekedy sa namiesto citlivosti uvádza mierka pre presnosť ako prevrátená hodnota citlivosti v LSB/g. Tiažové zrýchlenie g sa mierne líši podľa zemepisnej šírke, ale stanovený prepočet na jednotky SI je $1g = 9.80665\,m/s^2$ [1].

Dynamický rozsah sa uvádza v tiažovom zrýchlení g. Hovorí o najmenšej a najväčšej rozlíšiteľnej hodnote zrýchlenia nad úrovňou ktorej už dochádza k skresleniu signálu orezaním špičiek. S nevyhnutnými drobnými nepresnosťami výroby mik-

[1] `https://physics.nist.gov/cgi-bin/cuu/Value?gn|search_for=acceleration`

romechaniky je tzv. *zero-g napätie* popisujúce odchýlku skutočného od ideálneho výstupu, keď na sústavu nepôsobí žiadne zrýchlenie. Za ideálnych okolností bez pohybu na vodorovnom povrchu namerajú osi x a y zrýchlenie $0g$, zatiaľčo na z pôsobí $1g$. Očakávaním je nulová hodnota výstupného napätia a tým aj výstupného registra.

Šírka pásma senzora v *Hz* predurčuje rozsah frekvencie vibrácií, ktoré je možné zachytiť. Podmienená je zvolenou početnosťou čítania akcelerácie za sekundu, čiže vzorkovacou frekvenciou. Stanovuje sa tiež nastaviteľným parameterom *ODR* (Output Data Rate) - výstupný dátový tok, pričom šírka pásma je spravidla polovicou ODR. Menej uvádzanou vlastnosťou býva *frekvenčná odozva* senzora, ktorá určuje o koľko sa v rámci tolerancie odlišuje skutočná citlivosť od referenčnej pre zodpovedajúcu frekvenciu vibrácii.

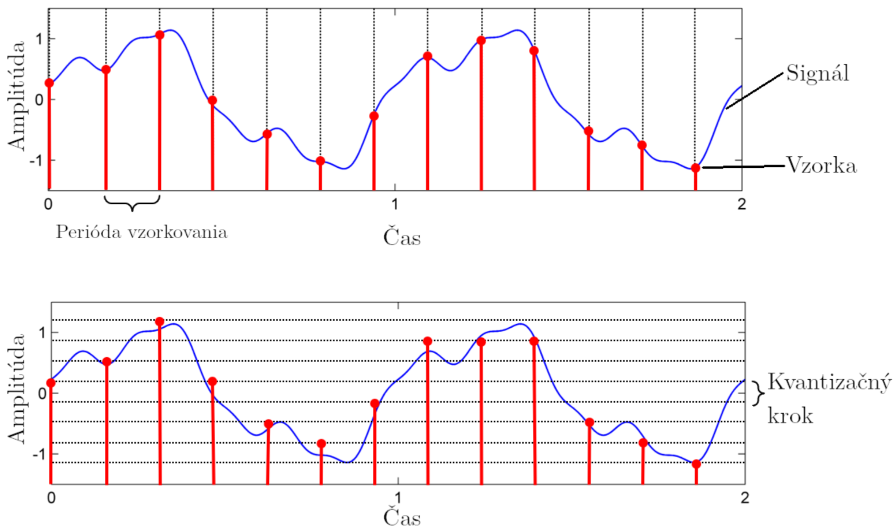
Na meranie zrýchlenia má nevyhnutný vplyv šum zapríčinený Brownovým pohybom a nedokonalosťou skutočných materiálov v štruktúre akcelerometra. Intenzita šumu rastie inverznou odmocninou so šírkou pásma, čiže s častejším meraním získavame menšiu presnosť. Pri dostatočnom odstupe signálu od šumu, $SNR = P_{signal}/P_{um}$, umožňuje hardvér akcelerometra vzorkovať amplitúdy až nad stanovený prah generovaním prerušenia, čím sa dokáže efektívne zbaviť nevýznamných fluktuácií.

2.1.3 Analógovo-digitálny prevodník

Spojitá napäťová úroveň transformuje analógovo-digitálny (A/D) prevodník pre spracovanie digitálnym systémom do množiny diskrétnych hodnôt. Vstupný signál najprv prechádza fázou vzorkovania, kedy sa vzorky zaznamenávajú v pravidelných intervaloch. Počet vzoriek odčítaných za sekundu je vyjadrený vzorkovacou frekvenciou f_s v Hz. Časový rozdiel medzi vzorkami, nazývaný perióda vzorkovania, je prevrátenou hodnotou vzorkovacej frekvencie $T_s = \frac{1}{f_s}$. Pre presnú rekonštrukciu pásmovo obmedzeného signálu v hraniciach $[-f_{max}; f_{max}]$ je nevyhnuté podľa *Nyquist-Shannonovej vety* o vzorkovaní, aby vzorkovacia frekvencia bola najmenej dvojnásobkom maximálnej frekvencie snímaného signálu.

$$f_s \geq 2 \cdot f_{max} \tag{2.4}$$

Každej vzorke je následne v procese kvantovania priradená diskrétna hodnota s konečným počtom n bitov, ktorá je najbližšia možná ku skutočnej hladine analógového vstupu. Dochádza pritom k istému zaokrúhľovaniu z dôvodu nepresnosti vyjadrenia spojitej domény amplitúd diskrétnym číslom. Tento jav označujeme ako kvantizačný šum, ktorý je najviac polovicou z maximálnej rozlíšiteľnej zmeny signálu a trpia nim všetky existujúce A/D prevodníky.



Obr. 2.3: Digitalizácia signálu v analógovo-digitálnom prevodníku [5]

Prevodníky integrované priamo s inerciálnymi jednotkami sa vyhotovujú v rozlíšeniach 12, 16 alebo 20 bitov. Umožňujú tak pripojiť akcelerometer rovno na sérové zbernice *SPI* alebo *I2C*. Všeobecne platí, že pri n bitoch je k dispozícii 2^n rozličných čísel. Kódovaním v dvojkovom doplnku pre zachytenie záporných hodnôt sa uvažuje s intervalom $[-2^{\frac{n}{2}}; 2^{\frac{n}{2}} - 1]$.

Napríklad pri 12-bitovom A/D prevodníku s referenčným napätím $3.3V$ je teoreticky najmenšia rozlíšiteľná zmena na najmenej významový bit $3.3V/2^{12} = 0.81mV$. Ak je rozhranie senzora priamo vybavené analógovým výstupom nič nebráni v použití presnejšieho prevodníka, napriek tomu najmenší merateľný dielik je zdola stále ohraničený citlivosťou akcelerometra.

Číslicová hodnota v dvojkovom doplnku získanú konverziou $\hat{x}$ je prepočítaná na štandardné fyzikálne jednotky pre zrýchlenie, a v m/s^2). R prestavuje nastavený dynamický rozsah v jednotkách g a n je počet bitov A/D prevodníka.

$$a = \hat{x} \cdot ((R \cdot g)/2^{n/2}) \tag{2.5}$$

Na základe už zmieneného ohľadom vlastností MEMS akcelerometrov, presnejší prevod dosiahneme zužitkovaním deklarovanej citlivosti senzora pri danom dynamickom rozsahu S_R udávaného v mg/LSB.

$$a = \hat{x} \cdot (S_R \cdot g)/1000 \tag{2.6}$$

2.1.4 Vlastnosti bežných akcelerometrov

Na ilustráciu uvádzame parametre zvolených najrozšírenejších typov akcelerometrov. Akcelerometer LSM9DS1 [6] umožňuje cez zbernicu SPI alebo I2C zvoliť zo štyroch dynamických rozsahov, pričom každé rozpätie sa vyznačuje svojou citlivosťou. Zvolením menšieho dynamického rozsahu zvýšime citlivosť. LSM9DS1 funguje pri rozsahoch ±2g, ±4g a ±8g a ±16g, postupne s citlivosťami 0.061 mg/LSB, 0.122 mg/LSB, 0.244 mg/LSB, 0.732 mg/LSB. Výstupný dátový tok (ODR) je možné nastaviť na 10Hz, 50Hz, 119Hz, 238Hz, 476Hz a najvyššie na 952 Hz. Navzorkované hodnoty sú ukladané do 16-bitového výstupného registra v dvojkovom doplnku.

Nízkoenergetický 3DOF MEMS akcelerometer ADXL362 [7] so spotrebou $2\,\mu A$ pri 100Hz disponuje rozsahmi ±2g, ±4g a ±8g s citlivosťami 1, 2 a 4 mg/LSB. Dostupné vzorkovacie frekvencie 12-bitového A/D prevodníka sú $12.5 - 400$Hz v 8 krokoch vždy po násobkoch predošlého kroku. Pre rýchlejšie čítanie pri nižšom rozlíšení dokáže senzor zakódovať dáta do 8-bitového registra.

Vyrábajú sa tiež akcelerometre s väčšími dynamickými rozsahmi a nízkym šumom, ide napríklad o ADXL356 a ADXL357 [8] so škálami ±10g, ±20g a ±40g s citlivosťou $0,019$ mg/LSB po $0,078$ mg/LSB a rozlíšením A/D prevodníka 20 bitov pri ODR $4 - 4000$Hz. ADXL357 ponúka priamo analógové výstupy s citlivosťou $20 - 80$ mV/g pri napájaní 3.3 V.

2.1.5 Odvodzovanie rýchlosti a dráhy zo zrýchlenia

Meranie akcelerácie umožňuje zároveň nepriamo získať ďalšie údaje o pohybe celkovom v priestore ako aj spôsobenom vibráciami. Zrýchlenie $\vec{a}$ je definované ako časová zmena rýchlosti $\vec{v}$, zatiaľčo rýchlosť je časovou zmenou polohy $\vec{r}$. Na pozorovanie prechodových javov alebo na vyjadrenie miery plynulosti pohybu slúži ryv $\vec{j}$, ktorý je časovou zmenou akcelerácie. Pokiaľ nie sú známe počiatočné podmienky v okamihu začiatku snímania akcelerácie, budú hodnoty veličín relatívne vzhľadom na štart záznamu. Kinematika v diskrétnom čase je potom opísaná nasledujúci rovnicami, kde Δ je operátor diferencie $\Delta t = t(i) - t(i-1)$:

$$\vec{v} = \frac{\Delta \vec{r}}{\Delta t}; \quad \vec{a} = \frac{\Delta \vec{v}}{\Delta t}; \quad \vec{j} = \frac{\Delta \vec{a}}{\Delta t} \tag{2.7}$$

Vyjadrenie neznámych premenných vzhľadom na akceleráciu spočíva v prenásobení rovníc členom Δt, čím sa získajú vzťahy pre okamžitú dráhu a okamžitú rýchlosť. Spočítaním čiastkových okamžitých rýchlostí na intervale dostaneme celkovú rýchlosť a rovnaký úsudok platí pre polohu. V spojitom čase, keď by vzorkovacia perióda bola nekonečne krátka, dochádza naproti tomu k integrovaniu funkcie akcelerácie. Dostávame, že rýchlosť je integrálom zrýchlenia a poloha je dvojným integrálom zrýchlenia:

$$\vec{v}(t) = \vec{a_0} + \int \vec{a}(t)\,\mathrm{dt} \tag{2.8}$$

$$\vec{r}(t) = \vec{r_0} + \vec{v_0}t + \iint \vec{a}(t)\,\mathrm{dt} \tag{2.9}$$

2.1.6 Numerická kvadratúra

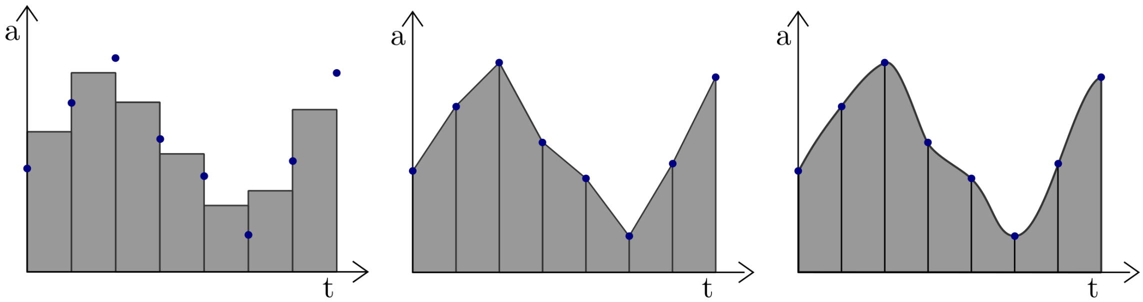
Približný výpočet určitého integrálu funkcie akcelerácie je založený na geometrickej interpretácii integrálu ako plochy pod krivkou. Hovoríme vtedy o probléme numerickej kvadratúry, ktorý navrhuje nahradiť pôvodný integrand interpolačným polynómom [9]. Rád polynómu n implicitne stanoví priebeh funkcie medzi ekvidištantnými vzorkami a má dopad na presnosť aproximácie. Najčastejšie sa používajú konštantný ($n = 0$), lineárny ($n = 1$) alebo kvadratický ($n = 2$) polynóm, podľa toho rozlišujeme obdĺžnikové pravidlo (vzorec 2.10), lichobežníkové pravidlo (vzorec 2.11) a Simpsonovo pravidlo (vzorec 2.12).

$$v(t_i) = T_s \cdot a\left(\frac{t_i + t_{i-1}}{2}\right) \tag{2.10}$$

$$v(t_i) = \frac{T_s}{2} \cdot [a(t_i) + a(t_{i-1})] \tag{2.11}$$

$$v(t_i) = \frac{T_s}{3} \cdot [a(t_{2i}) + 4a(t_{2i-1}) + a(t_{2i-2})] \tag{2.12}$$

Pri *obdĺžnikovom pravidle* (obr. 2.4a) nepripúšťame zmenu hodnoty zrýchlenia medzi vzorkami a okamžitú rýchlosť, čiže plochu, odhadneme ako dĺžku intervalu vzorkovania vynásobenú priemerom výšok dvoch následných pozorovaní. Interpolačný polynóm je konštantná funkcia. *Lichobežníkové pravidlo* (obr. 2.4b) uvažuje s lineárnou zmenou veličiny medzi meraniami, preto interpoluje priamkou. *Simpsonovo pravidlo* (obr. 2.4c) sa snaží o ešte tesnejší odhad s využitím kvadratickej funkcie. Každé kvadratúrne pravidlo sa síce vyznačuje presne vyčísliteľnou chybovosťou, ale k tomu je nevyhnutné poznať analytické vyjadrenie vibrácií, čo dáva realistický odhad len pri čisto periodických kmitoch.



(a) Obdĺžnikové (b) Lichobežníkové (c) Simpsonovo

Obr. 2.4: Porovnanie pravidiel numerickej integrácie

Priama integrácia zašumeného signálu zrýchlenia vedie k neskutočnému driftu, ktorý je ešte zvýraznený dvojitou integráciou pri odvodzovaní relatívneho posunutia. Dochádza k zosilneniu nízkych a potlačeniu vyšších frekvencií, čím sa začne dominovať neexistujúci trend vo výstupných dátach. Očakávané oscilujúce správanie vychýlenia u vibrácií so zväčšujúcim sa počtom sčítancov pri rekurentnom výpočte zaniká. Na zlepšenie stability integrátora sa uplatňuje korekcia cez obálky [10].

Najprv je na vstupnom signále vykonaná zvoleným pravidlom numerická kvadratúra, ktorá môže byť realizovaná na krátkych úsekoch funkcie, aby sa predišlo pretečeniu pri výraznej akumulácií odklonu. Prichádza k identifikácií lokálnych extrémov, či už maxím, respektíve miním (pozri 2.2.4) a ich interpoláciou s kubickou B-spline sa sformuje horná $e_u(t)$, respektíve dolná obálka signálu $e_d(t)$. Obálky sú spriemerované $\bar{e}(t)$, čím vznikne odhad trendovej krivky, ktorá je od už integrovaného signálu odčítaná $g(t) = f(t) - \bar{e}(t)$. V prípade výpočtu polohy je možné aplikovať uvedený postup kaskádovito, čiže rovnako ako akcelerácia je aj signál rýchlosti opäť integrovaný a korigovaný obálkami. Trend je rovnako tak odstrániteľný filtrom na potlačenie prítomnosti jednosmernej zložky v podobe hornej priepuste (pozri 2.3.4).

2.2 Metódy analýzy signálu v časovej doméne

Pozorovania veličiny predstavujú udalosti merané sekvenčne v čase, kde je s každou obdržanou hodnotou x_i viazaná unikátna časová značka t_i. Postupnosť jednotlivých čítaní je jednorozmerný časový rad znázoriteľný ako usporiadaná množina dvojíc pečiatky rastúcej v čase a nasnímanej úrovne: $T = \{(t_1, x_1), (t_2, x_2), \ldots, (t_n, x_n)\}$. Vzorkovaním v pravidelných intervaloch stačí uvažovať namiesto časových značiek o celočíselných indexoch, ktoré určujú pozíciu prvkov vo vektore pozorovaní: $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$. Keď sú súčasne zaznamenávané viaceré dátové body hovoríme o matici pozorovaní $X_{n,m}$, či o viacrozmernom časovom rade.

Pri veľkom objeme prichádzajúcich vzoriek produkované senzormi, nie je uskutočniteľné ich úplné uchovanie ani spracovanie celkého dátového toku naraz. Častokrát by stratégia neuváženého odkladania viedla k plýtvaniu zdrojov a zbytočnému archivovaniu údajov s nízkou informačnou hodnotou. Vhodnejšie je agregovanie toku údajov podľa preddefinovaného zmysluplného kritéria, ktoré by umožňovalo zachytiť významné rysy a prompte zodpovedať na vyžadované dopyty. Napríklad zrýchlenie vozidla v konkrétnom okamihu nebude až tak podstatné v porovnaní so znalosťou trvania úsekov pridávania alebo brzdenia, či priemernej prudkosti s akou tieto aktivity boli činené.

2.2.1 Prúdové algoritmy

Priamočiarou realizáciou agregácie je nahliadať na prvky časového radu postupne ako prichádzajú. Prúdové algoritmy pôsobiace v reálnom čase, a teda neschopné vidieť finálny vektor vzoriek vstupu sa vyznačujú vlastnosťou, že vyprodukujú len na základe takého čiastkového vstupu parciálny výsledok platný pre dosiaľ sa vyskytnutú podmnožinu.

Online algoritmus spracúvajúci neprestajný potenciálne nekonečne sa rozširujúcu vstupnú sadu sa ideálne vyznačuje sublineárnou alebo polylogaritmickou časovou zložitosťou spracovanie jednotlivých prvkov, celkové spracovanie, a sú žiaduce sublineárne pamäťové nároky [11].

Za ideálnych okolností by sa mal online algoritmus učiť kontinuálne bez ukladania predošlých bodov a detekcií. V rozhodnutiach algoritmu sú zahrnuté informácie o všetkých predošlých bodoch do terajšieho rozhodnutia. Mal by mať schopnosť sa adaptovať dynamickému prostrediu, v ktorom pôsobí, bez nutnosti manuálnych úprav parametrov modelu. Zároveň je žiaduce minimalizovať falošné pozitíva a negatíva pri detekcii udalostí. V turniketovom modeli sa pri online agregácii bod x_i načítaný v prúde, zaráta do každého relevantného počítadla, a môže byť ihneď zabudnutý [11].

2.2.2 Posuvné a rozširujúce sa okná

Časový rad $(x_i)_{i=0}^{n}$ s dĺžkou n môže byť pre účely výpočtu sumárnych štatistík rozdelený oknovou funkciou $\mathcal{W}_{l,d}$ na podpostupnosti nazývané okná.

Posuvné okná („rolling window") majú spravidla konštatnú dĺžku l menšiu ako celkovú veľkosť radu a sú aplikované s krokom odstupu d pozorovaní. Rad pozorovaní pozostáva z $(n-(l-1))/d$ okien [12]. Prirodzene sa posuvné okná objavujú pri manipulácii s vyrovnávacou pamäťou, ktoré sa využívajú pri blokovom prenose z adaptéra senzora do hlavnej pamäte. Vtedy sa veľkosť bloku sa rovná posunu $l = d$.

Rozširujúce sa okná („expanding window") nachádzajú uplatnenie v menej prípadoch, spravidla sa jedná o inkrementálny odhad globálnej štatistiky, ktorá má zmysel prevažne pri sledovaní stabilného javu. [13]. Okno začína na stanovenej minimálnej veľkosti a s pribúdajúcim počtom bodov ich zahŕňa, čím sa zväčšuje. Dostavuje sa rovnaký výsledok ako v turniketovom modeli, ale navyše neplatí obmedzenie zahadzovania už započítaných údajov, ale dokážeme teoreticky operovať so všetkými vzorkami vrámci okna.

2.2.3 Číselné charakteristiky štatistického rozdelenia

Náhodné vibrácie vyskytujúce sa pri skutočných materiáloch sú stochastický proces, ktorý tvorí sekvencia časovo indexovaných náhodných premenných. Časový rad predstavuje realizáciu tohto stochastického procesu $\mathbf{Y} = (X_1, X_2, ..., X_n)^T$, kde X_t je náhodná premenná so svojím rozdelením pravdepodobnosti. Všeobecne sa pri ideálnych stacionárnych otrasoch predpokladá, že premenné pochádzajú z unimodálnej Gaussovej distribúcie: $X_t \sim N(\mu, \sigma^2)$ [1].

Sumárna deskripcia nameraného deja pre extrakciu typických čŕt konkrétnych pozorovaných situácií sa uskutočňuje viacerými štatistikami $h(X_1, X_2, ..., X_n)$ zostručňujúcimi opis funkcie hustoty rozdelenie. Na rozmiestnenie hodnôt meraní v priebehu časového úseku sa nazerá z pohľadu polohy, rozptýlenosti a tvaru. Rozsah oboru hodnôt je amplitúda špička-špička („peak-to-peak”), ktorá je rozdielom maximálnej a minimálnej úrovne, údaj známy tiež ako variančné rozpätie [14].

$$x_{pp} = \max_{t \in \mathcal{W}}\{x_t\} - \min_{t \in \mathcal{W}}\{x_t\} \tag{2.13}$$

Priemernú energiu obsiahnutú v signále predstavuje štvorec efektívnej amplitúdy RMS a určí sa ako kvadratický priemer pozorovaní:

$$x_{rms} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} x_t^2} \tag{2.14}$$

Mierami polohy rozdelenia pozorovaní sú stredná hodnota, informujúca o centre hodnôt veličiny, a kvantily rozkladajúce usporiadaný vektor pozorovaní na určený počet rovnakých skupín. Nevychýleným bodovým odhadom strednej hodnoty je *výberový priemer*, ktorý je zároveň amplitúdou jednosmernej zložky signálu:

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t \tag{2.15}$$

Najvýznamnejším kvantilmi sú kvartily Q_q vytvárajúce štyri rovnako veľké časti z pôvodných dát, konkrétne dolný kvartil Q_1 oddelí 25% najmenších údajov, medián Q_2 predelí zoradené údaje na polovicu a horný kvartil Q_3 zahrnie 75% nižších hodnôt. Hľadaný kvartil je k-ty najmenší prvok v utriedenom zozname meraní, pričom podľa želaného kvartilu q a počtu pozorovaní je $k = \lceil n \cdot (1/q) \rceil$.

Zistenie k-teho najmenšieho prvku s časovou zložitosťou $\mathcal{O}(n \log n)$ umožňuje ľubovoľný lepší triediaci algoritmus napríklad triedenie zlučovaním (merge sort). Algoritmus *Quickselect* dokáže taký prvok objaviť v čase $\mathcal{O}(n)$. V každom kroku vyberie náhodný deliaci bod (pivot) a preskupí k sebe hodnoty menšie ako pivot naľavo a väčšie ako pivot napravo. Najmenší prvok následne hľadá v časti, kde zostalo viac ako k prvkov. Pokiaľ došlo k deleniu zoznamu, že pivot zaujme presne k-tu pozíciu prehľadávanie je ukončené a pivot prehlásený za riešenie. Nesprávnym výberom pivota môže v najhoršom prípade dôjsť až k zložitosti $\mathcal{O}(n^2)$, čomu sa predchádza výberom pivota cez medián mediánov.

Sústreďovanie realizácie veličiny, respektíve jej rozptýlenosť okolo strednej hodnoty vieme opísať viacerými štatistikami ako sú *výberový rozptyl* (2.16), ktorej odmocninou dostaneme smerodajnú odchýlku, *priemerná absolútna odchýlka* (2.17), *mediánová absolútna odchýlka* (2.18) a *medzikvartilové rozpätie* (2.19) [14]. Priemerná absolútna odchýlka je upraviteľné o mieru centrálnej tendencie, ktorou okrem priemeru môže byť aj medián alebo modus. Vyvarovanie sa príliš extrémnym a vychýlením hodnotám docielime zapojením práve mediánu do štatistík absolútnej odchýlky, rovnako tak to dosiahneme medzikvartilovým rozpätím obmedzením sa na 50% centrálnych dát.

$$s^2 = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \bar{x})^2 \tag{2.16}$$

$$d = \frac{1}{n}\sum_{t=1}^{n}(|x_t - \bar{x}|) \tag{2.17}$$

$$\mathrm{MAD} = med(|x_t - med(\mathbf{x})|) \tag{2.18}$$

$$IQR = Q_3 - Q_1 \tag{2.19}$$

Numericky stabilné bežiace štatistiky priemeru a smerodajnej odchýlky sa udržiavajú cez rekurentné rovnice *Welfordovho algoritmu* [15]. M_1 je aktuálna priemerná hodnota údajov v toku a S_1 je počítadlo pre rozptyl, z ktorého je v ktoromkoľvek okamihu získateľná smerodajná odchýlka súboru σ:

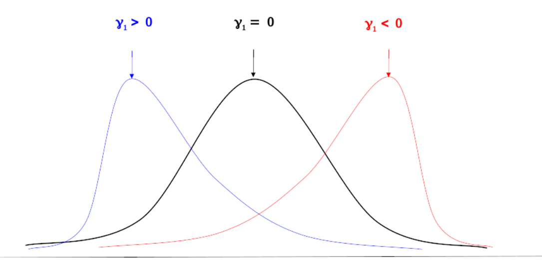
$$M_1 = x_1; \quad M_k = M_{k-1} + \frac{(x_n - M_{k-1})}{k} \tag{2.20}$$

$$S_1 = 0; \quad S_k = S_{k-1} + (x_k + M_{k-1})(x_k + M_k) \tag{2.21}$$

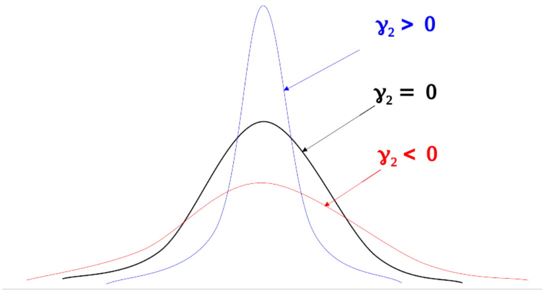
$$\sigma = \sqrt{S_n/(n-1)} \tag{2.22}$$

Tvar distribúcie náhodnej premennej opisujú centrálne momenty šikmosť a špicatosť. Šikmosť $\gamma_1 = \mu^3/\sigma^3$ udáva skosenie rozdelenia, pričom platí že záporná šikmosť značí dlhší ľavý chvost a modus funkcie hustoty sa prevažuje napravo. Zatiaľ čo u kladnej šikmosti je to naopak (obr. 2.5a).

Špicatosť $\gamma_2 = \mu^4/\sigma^4 - 3$ (obr. 2.5b) porovnáva rozdelenie pozorovaní so strmosťou krivky normálneho rozdelenia, čiže viacej realizácií leží bližšie alebo ďalej od strednej hodnosti. Kladná špicatosť signalizuje strmejšiu a záporná zasa sploštenejšiu distribúciu. Platí, že μ_n je priemerom hodnôt $(x_t - \bar{x})^n$.



(a) Šikmosť



(b) Špicatosť

Obr. 2.5: Dopad šikmosti a špicatosti na histogram distribúcie

Závislosť dvojíc veličín sa vyjadruje *kovariancia* $cov(\mathbf{x}, \mathbf{y})$ a *korelácia* $\rho(\mathbf{x}, \mathbf{y})$. U vektora akcelerácie nás bude napríklad zaujímať vzájomná korelácia medzi osami pohybu: $\rho(\vec{x}, \vec{y})$, $\rho(\vec{x}, \vec{z})$, $\rho(\vec{y}, \vec{z})$ upozorňujúca na diagonálny pohyb alebo podobné budenie v oboch korelovaných smeroch a tým umožňujúce redukciu údajov z dôvodu redundancie. Kovariancia je daná strednou hodnotou súčinu odchýlky od priemeru zodpovedajúcej premennej (vzťah 2.23). Normovaním kovariancie smerodajnými odchýlkami veličín získame *Pearsonov korelačný koeficient* (vzťah 2.24), ktorý je z intervalu $[-1; 1]$. Hodnota koeficientu -1 značí nepriamu lineárnu závislosť a $+1$ priamu závislosť.

$$\operatorname{cov}(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{t=1}^{n} (x_t - \bar{x})(y_t - \bar{y}) \tag{2.23}$$

$$\rho(\mathbf{x}, \mathbf{y}) = \frac{\operatorname{cov}(\mathbf{x}, \mathbf{y})}{\sigma_x \sigma_y} \tag{2.24}$$

2.2.4 Algoritmy na rozpoznávanie špičiek

Detekcia udalostí a významných zmien signálového priebehu sa spolieha na hodnovernú identifikáciu špičiek amplitúdy. Dôležitými indikátormi pre celkovú charakterizáciu javu slúži potom časová pozícia špičky v rámci prúdu, výška prejavujúca sa získanou úrovňou, šírka obsahujúca údaj o trvaní, či plocha stvárňujúca energiu.

Ekvivalentne sa špičky z matematického hľadiska stotožňujú s lokálnymi extrémami funkcie, čo sú maximá (vrcholy) a minimá (údolia). Podľa definície je lokálne maximum t_0 bodom, ktorý má vyššiu funkčnú hodnotu ako všetky ostatné body na intervale $t_0 \in I$ (2.25), lokálne minimum má naopak najnižšiu hodnotu na intervale

(2.26) [16].

$$f[t_0] \geq f[t], \forall t \in I \tag{2.25}$$

$$f[t_0] \leq f[t], \forall t \in I \tag{2.26}$$

Kľúčové pre spoľahlivé určenie extrémov je práve interpretácia intervalu I v algoritmoch, ktoré zastupujú rozličné potreby korektného vyhodnotenia. Jediné minimum a maximum sa dosiahne zvolením celej dĺžky záznamu za interval, čím sa stratia dočasné disturbancie. Na druhej stane prílišným skrátením intervalu sa skoro všetky vzorky budú javiť ako náhle zmeny.

Skutočné signály sa potýkajú so šumom, ktorý sťažuje odlíšenie pravej tendencie od krátkodobých výkyvov. Pred samotným procesom hľadania špičiek je preto aplikovaný vyhladzovací filter, v prípade potreby aj opakovane na už vyhladení signál. Najčastejšie sa jedná o filter kĺzavého priemeru, Savitzky–Golay alebo Gaussov filter [17]. Filtrovanie sa realizuje diskrétnou jednorozmernou konvolúciou vstupného signálu a masky filtra $y[n] = x[n] * w[n]$, ktorá býva hardvérovo akcelerovaná inštrukciami „fused multiply-add"[2].

Detekcia špičiek prahovou úrovňou

Za predpokladu, že priebeh meranej veličiny sa vyznačuje krátkymi impulzmi s viac-menej pravidelnou amplitúdou je priamočiarou metódou na odlíšenie špičiek od hladín nízkej aktivity určenie prahu θ, ktorý zaregistruje všetky väčšie hodnoty. Lokálne extrémy sú potom vzorky signálu spĺňajúce podmienku:

$$|f[t]| \geq \theta \tag{2.27}$$

Určenie takejto hraničnej hladiny prebieha zväčša empiricky alebo na základe heuristík, ktoré so sebou nesú domnienku o vlastnostiach priebehu pozorovaní. Uspokojivými odhadom za určitých okolností môžu byť prahy θ: viac ako priemer s toleranciou, horné 3/4 celkového nedávneho rozsahu hodnôt, či dokonca viac ako k smerodajných odchýlok. Odlišné nazeranie na prahovú hodnotu spočíva v jej nastavení pre rozpoznanie vzájomnej korelácie signálu a masky zodpovedajúcej tvaru

[2] https://developer.arm.com/documentation/102198/0200/Convolution

impulzu. Táto úvaha sa opiera o to, že impulz musí byť dostatočne pravidelný, aby bol nezameniteľne odlíšiteľný.

Význačnosť vrchola spomedzi susedov

Doplnkom ku rozpoznávaniu špičiek podľa absolútnej prahovej úrovne je porovnávanie bodov na obe strany od preskúmavaného vrchola, čím zistíme relatívnu významnosť extrému pre najbližšie susedstvo. Aby bola hodnota na danej pozícii t označená za špičku v okolí pozostávajúcom z k priľahlých bodov, musí byť v porovnaní so všetkými väčšia. Pre okrajové dátové body $f[0]$ a $f[n]$ dochádza k porovnaniu iba z jednej strany [16].

$$f[t-i] < f[t] > f[t+i], \quad \forall i \in 1, 2, ..., k \tag{2.28}$$

Algoritmus č.1 „najvyšší spomedzi susedov" [3] prechádza postupne pozorovania veličiny zo zoznamu y a ku kandidátnej špičke na indexe i preveruje najbližších k hodnôt na obe strany, ak existujú. Keď po preskúmaní zostáva $y[i]$ najväčšou

Algoritmus 1 Najvyšší spomedzi susedov

```
function FIND_PEAKS_NEIGHBOURS(y, k, ε, h_rel, h)
    peaks ← []                                   ▷ Zoznam indexov nájdených špičiek v signále y
    for i ← 0 to length(y) do
        if h ≠ null and |y[i]| < h then          ▷ Preskoč príliš nízke magnitúdy
            continue
        end if
        possible_peak ← true
        a ← max(i − k, 0)
        b ← min(i + k, length(y))
        for j ← a to b do                        ▷ Porovnaj špičku s bodmi v susedstve
            if i ≠ j and y[j] − y[i] > ϵ then
                possible_peak ← false            ▷ Kopec nie je dostatočne strmý
            end if
        end for
        if possible_peak = true and y[i] − min(y[a], y[b]) > h_rel then
            peaks ← peaks + [j]                  ▷ Kandidát je prehlásený za špičku
        end if
    end for
    return peaks
end function
```

[3] https://terpconnect.umd.edu/~toh/spectrum/PeakFindingandMeasurement.htm

hodnotou spomedzi susedov v rozmedzí $[a; b]$, za tolerancie strmosti stúpania ε medzi pozorovaniam, a súčasne je relatívna výška vrcholu väčšia oproti nižšiemu okraju než nastavený parameter h_{rel} potom je kandidátny bod prehlásený za skutočnú špičku a pridaný do zoznamu *peaks*.

Súčasťou algoritmu je tiež preskočenie hodnôt, ktoré nespĺňajú základný predpoklad pre absolútnu amplitúdu h. Časová zložitosť pre rozhodnutie o jednej špičke je lineárna v závislosti od veľkosti posuvného okna uvažovaného susedstva $\mathcal{O}(2k)$.

Algoritmus prechodu nulou do záporu

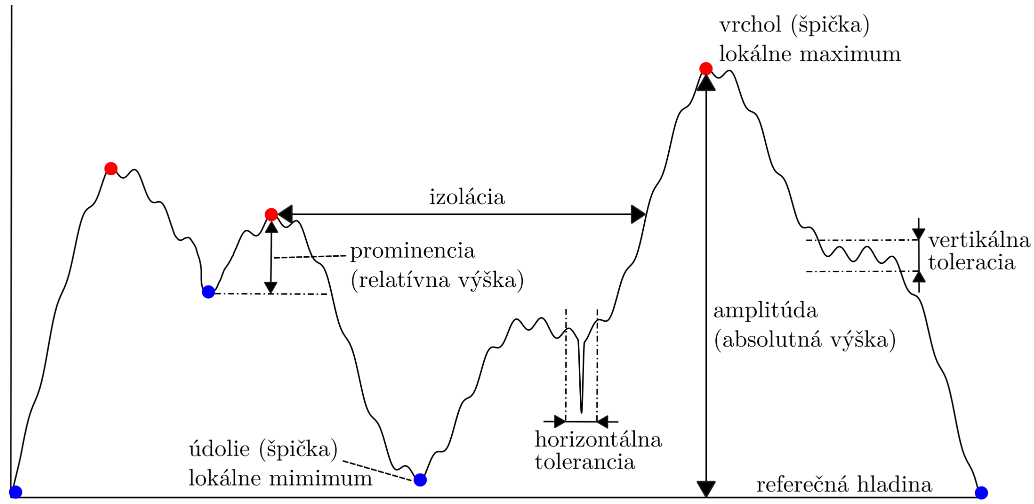
Pomyselné vrcholy a údolia v zosnímaných hodnotách sú miestom, kde sa mení smer úrovní amplitúdy zo stúpania na klesanie alebo z klesania na stúpanie, čím na pomedzí týchto opozitných trendov vzniká stacionárny bod, kde je prvá diferencia nulová: $\Delta f[i] = 0$. V lokálnom maxime dochádza súčasne k zmene znamienka prvej diferencie z kladného na záporné. Prudkosť kopca vyplýva z absolútnej hodnoty diferencie.

Viacnásobné vyhladenie signálu predom je nesmierne dôležité, pretože algoritmus č.2 „prechodu nulou do záporu" (Negative Zero-Crossing) je nesmierne citlivý na zákmity a nesprávneby ich považoval za špičky. Zvýšenie odolnosti proti takýmto tendenciám sa dosahuje dlhšou sečnicou spájajúcou bod i s k-tou vzorkou vedľa, ktorá sa použije namiesto diferencie s jednotkovým krokom.

Algoritmus 2 Prechod prvej derivácie nulou do záporu

function FIND_PEAKS_ZERO_CROSSING(y, k, ε, $slope$)
    $peaks \leftarrow []$
    for $i \leftarrow k$ **to** $length(y) - k$ **do**
        if $(|y[i+k] - y[i-k]| \leq \epsilon$ **and**
                $(y[i+k] - y[i]) - (y[i] - y[i-k]) < 0$ **and**
                $|(y[i+k] - y[i]) - (y[i] - y[i-k])| > slope)$ **then**
            $peaks \leftarrow peaks + [i]$
        end if
    end for
    return $peaks$
end function

Označenie kandidátneho bodu za špičku v zozname hodnôt y stojí na teda troch kritériách. Sklon sečnice sa musí v rámci tolerancie ε blížiť nule, rozdiel prvých diferencií $\Delta y[i+k] - \Delta[i]$ musí byť záporný a veľkosť rozdielu diferencií prekračuje prahovú strmosť kopca *slope*, kde leží uvažovaný vrchol. Časová zložitosť pre jednu špičku je $\mathcal{O}(1)$.



Obr. 2.6: Topografia priebehu signálu

Algoritmus horského turistu

Zanesením do grafu pripomína priebeh funkcie kmitajúceho deja členité pohorie. Na problém rozhodovania sa o tom, či danú lokalitu považovať za vrchol možno nahliadať z pohľadu chodca cestujúceho po krivke z lineárne interpolovaných vzoriek. V princípe ide myšlienkou o jednoduchý stavový automat sledujúci aktuálny stav terénu a konajúci rozhodnutia na základe predošlej skúsenosti v intenciách rozhodovacích pravidiel.

Algoritmus č.3 horského turistu na začiatku púte z počiatočných bodov zistí, ktorým z dvoch vertikálnych smerov sa krivka uberá. V prípade, že po druhom kroku dôjde k zmene smeru zapíše sa indikácia možného spádu kopca. Výchylka môže byť v dôsledku neprekročenia prahových úrovní v horizontálnej (*hole*) a vertikálnej (*tolerance*) osi ignorovaná, lebo ani na lesnom chodníku sa nepovažuje každá jama alebo vydutie za horu.

Domnelý vrchol je označený za lokálne maximum, keď spĺňa parametre pre topografické vlastnosti minimálnej akceptovateľnej prominencie a izolácie (obr. 2.6. Prominencia znamená relatívnu výšku oproti predošlej navštívenej doline. Izolácia vyčísľuje vzdialenosť k najbližšiemu skoršiemu vrcholu. Podobný algoritmus už existuje v literatúre [18], avšak nižšie prezentovaný pseudokód je oproti nemu zjednodušený a doplnený o požadované tolerancie.

Algoritmus 3 Algoritmus horského turistu

```
function FIND_PEAKS_HILL_WALKER(y, tolerance, hole, prominence, isolation)
    peaks ← []
    i_change ← 0
    y_valley ← 0
    possible_change ← false
    uphill ← (y[1] − y[0]) ≥ 0
    for i ← 1 to length(y) do
        y_step ← y[i] − y[i − 1]
        slope ← y_step ≥ 0
        if possible_change = false and uphill ≠ slope then
            possible_change ← true            ▷ Označenie potenciálneho extrému
            i_change ← i − 1
        else if possible_change = true and uphill = slope then
            possible_change ← false           ▷ Potenciálny extrém bol zachvením
        end if
        if (possible_change = true
                and uphill ≠ slope
                and |i − i_change| > hole
                and |y[i] − y[i_change]| > tolerance) then
            posible_change ← False            ▷ Významný lokálny extrém potvrdený
            prev_uphill ← uphill
            uphill ← slope
            if prev_uphill = false and uphill = true then
                y_valley ← y[i_change]        ▷ Nájdené údolie
            else if (prev_uphill = true
                    and uphill = false
                    and |y[i − hole] − y_valley| > prominence)
                    and |y[i − hole] − y[last(peaks)]| > isolation) then
                y_peak ← y[i_change]          ▷ Skutočný vrchol identifikovaný
                peaks ← peaks + [i_change]
            end if
        end if
    end for
    return peaks
end function
```

2.2.5 Metriky pre binárny klasifikátor

Uviedli sme tri rozdielne rovnocenné prístupy odhalenia špičiek uplatniteľné v nasadení online na sústavný prúd údajov separovaný do blokov posuvným oknom. Rozobrali sme algoritmus opierajúci sa o porovnávanie striktne usporiadaných susedov na obe stany, algoritmus využívajúci sklon sečníc vychádzajúc s vlastností prvej derivácie a napokon stavový automat odvolávajúci sa na sekvenčne preskúmavanú topografiu krivky grafu. Spoločným rysom zmienených techník je vykonávanie binárne zaradenie pre každú vzorku, či sa sa nachádza alebo nenachádza na aktuálnej pozícií vrchol.

Rozhodnutie môže viesť k správnemu P alebo nesprávnemu N riešeniu vzhľadom na objektívnu pravdu sprostredkovanú anotovanými dátami. Keď sa kategorizácia zhoduje s realitou dostávame skupiny skutočne pozitívnych TP a skutočne negatívnych TN. V prípade, že sa klasifikátor pomýli vyjde buď chyba prvého rádu FP, kedy registrujeme neexistujúcu špičku, alebo chyba druhého rádu FN, kedy ju prehliadneme. Umiestnením počtov charakteru rozhodnutí do tabuľky vzniká matica zámen.

Úspešnosť klasifikačných algoritmov pre ich vzájomné porovnanie kvantifikujú viaceré metriky. Na odladenie parametrov vplývajúcich na náchylnosť preferovať kladné alebo záporné výsledky sa vzťahuje *prevalencia* výskytu očakávaného javu: $P/(P+N)$ v samotných dátach. Snahou rozhodovania je maximalizovať senzitivitu a špecifickosť výsledkov algoritmu. *Senzitivita* (2.29) udáva koľko bodov, ktoré sú prehlásené za špičky naozaj je špičkami. *Špecifickosť* (2.30) sa zameriava na potvrdenie, aké množstvo pozorovaní nepovažovaných za špičky, nie sú nimi aj skutočne.

$$TPR = \frac{TP}{P} = \frac{TP}{TP+FN} \tag{2.29}$$

$$TNR = \frac{TN}{N} = \frac{TN}{TN+FP} \tag{2.30}$$

Presnosť určenia lokálneho extrému sa skladá z správnosti (2.31) a precíznosti (2.32), ktoré je rovnako žiaduce dosahovať čo najbližšie sto percentám, pri nízkej chybovosti (2.33), čiže nízkeho počtu falošných poplachov.

$$PPV = \frac{TP}{TP + FP} \tag{2.31}$$

$$ACC = \frac{TP + TN}{P + N} \tag{2.32}$$

$$FPR = \frac{FP}{FP + TN} \tag{2.33}$$

Štandardným nástrojom na vyjadrenie kvality binárneho klasifikátora je *ROC krivka* zakresľujúca senzitivitu TPR vo zvislom smere voči vodorovnej chybovosti FPR. ROC vytvoríme graduálnym posúvaním prahu pre klasifikáciu prostredníctvom parametrov algoritmu. Použiteľný algoritmus sa vyznačuje vypuklou krivkou smerom k ľavému hornému rohu nad diagonálou, ktorá by sprevádzala počínanie náhodného rozhodovania. Dokonalá metóda pri dosahovala stopercentnú senzitivitu za nulovej chyby. Vyjadrením plochy pod ROC krivkou je miera AUC, ktorá umožňuje približné číselné porovnanie rôznych získaných kriviek.

2.3 Frekvenčná a časovo-frekvenčná analýza signálu

Vibrácie prítomné v získanom zázname zrýchlenia sa prejavujú s variabilnou pravidelnosťou a cyklickým opakovaním, ktoré vyplýva z povahy monitorovaného pohybu.

[19] vlastnosti frekvenčného spektra, decibele, spektrogram, odstup od šumu (SNR), power spectrum density (dbFS), spektrálny analyzátor

2.3.1 Fourierová transformácia

Diskrétna fourierová transformácia mapuje signál dĺžky N do množiny N diskrétnych frekvenčných komponentov. [20]

$$X = \mathbf{W}x; W_{nk} = e^{-i\frac{2\pi}{N}nk} = W_N^{nk} \tag{2.34}$$

Inverzná transformácia

$$x = \frac{1}{N}\mathbf{W}^H X \tag{2.35}$$

Integrálne transformácie: Fourierová transformácia (CFT, DFT), Kosínusová transfomácia (MDCT), [21] [22]

$$\mathcal{F} : X(\omega) = \int_{-\infty}^{+\infty} x(t) \cdot e^{-i\omega t} \mathrm{dt} \tag{2.36}$$

2.3.2 Algoritmus FFT pre DFT a DCT

Opis DIT radix-2 FFT algoritmus komplexných, reálny, pre MDCT [23] konštantný digitalizačný krok bez prerušení a navyše dĺžka radu musí byť celočíselná mocnina dvoch (s.29)

prevádza konečnú sekvenciu dát na súčet kosínov oscilujúcich na rôznych frekvenciách

$$X[m] = \sum_{n=0}^{N-1} x[n] \cdot e^{-i2\pi nm/N} \tag{2.37}$$

$$X[m] = \sum_{n=0}^{N-1} x[n] \cdot [\cos(2\pi nm/N) - i \cdot \sin(2\pi nm/N)] \tag{2.38}$$

Frekvenčné rozlíšenie

$$\Delta f = \frac{f_s}{N} \tag{2.39}$$

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right] \qquad k = 0, \ \ldots \ N-1$$

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\, frac\pi N \ (n+1/2)\,(k+1/2)\right] \qquad k = 0, \ \ldots \ N-1$$

$$X_k = \sum_{n=0}^{2N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}+\frac{N}{2}\right)\left(k+\frac{1}{2}\right)\right]$$

2.3.3 Oknové funkcie

the DFT implicitly assumes that the signal is periodic, i.e. that the time series of length N repeats itself infinitely in a cyclic manner. If the frequency of the sinusoidal input signal is not an exact multiple of the frequency resolution fres , i.e. does not fall in the exact center of a frequency bin, this assumption is not true, and the DFT will 'see' a discontinuity between the last sample and the first sample due to the cyclic continuation. That discontinuity spreads power all across the spectrum. [24]

Pri FT dostávame frekvenčné zložky za prítomné v celkom priebehu v rôznych časových okamihoch, keby sme chceli dosiahnuť lepšiu lokalizáciu v čase, tak prvý pokus ponúka Gáborova transformácia. Princíp neurčitosti: sústredenosť vo frekvenčnej domény vyvoláva neurčitosť v čase a symetricky naopak.

Dochádza k vynásobenie zdrojového signálu koeficientmi okna pre odstránenie hraničných efektov, ktoré spôsobujú/vnášajú neexistujúce periodické komponenty, na okraji okna preto znižuje úroveň na nulu. Snahou je dokonalé izolovanie frekvenčného pásma, ale pri FT dostávame sinc. Prehľad okien - symetrické okolo stredu

$$\text{Obdĺžníkové (rovnomerné) okno} w(n) = 1,\ n = 0, 1, ..., N-1 \tag{2.40}$$

$$\text{Trojuholníkové (Barlett) okno} w(n) = \begin{cases} \frac{n}{N/2}; n = 0, \dots, N/2 \\ 2 - \frac{n}{N/2}; n = N/2 + 1, \dots, N-1 \end{cases} \tag{2.41}$$

$$\text{Hanningovo okno} w(n) = 0.5 - 0.5\cos(2\pi n/N) \tag{2.42}$$

$$\text{Hammingovo okno} w(n) = 0.54 - 0.46\cos(2\pi n/N) \tag{2.43}$$

$$\text{Blackman-Harrisovo okno} w(n) = 0.35875 - 0.48829\cos(z) + 0.14128 cos(2z) - 0.01168 cos(3z), z \tag{2.44}$$

Efekt oknových funkcií na spektrálny leakage efekt (únik spektra) výhodné percentá prekryvu FT [25] [24] Priemerovanie a prekryv: - Amplitude Flatness (AF) - Power Flatness (PF) - Overlap Correlation (OC)

2.3.4 Filtre s konečnou impulznou odozvou

Roziel medzi FIR a IIR, Dolná pripusť, pásmová priepusť, horná pripusť, Konvolúcia a konvolučné jadro, konvolučná veta, účel: identifikácia prítomnosti známej frekvencie v signále akcelerácie

$$y(n) = \sum_{k=0}^{D_y} x(k) \cdot h(n-k) = x(n) * h(n) \tag{2.45}$$

Prenosová funkcia - koeficienty filtra

Detektor obálok

[4] [5] - Square-Law / Full-Wave - umocnenie dvomi (alebo absolútna hodnota) - amplifikácia faktorom 2 (zostane iba horná polovica) - decimácia 1.5 (podvzorkovanie) - dolná priepusť FIR - druhá odmocnina pre odstránenie skreslenia spôsobeného úvodným umocnením -

2.4 Senzorová sieť

Nízko-energetické zariadenia komunikujúce cez odľahčené sieťové protokoly so snahou spracovania v reálnom čase a ponechaním najdôležitejších informácii dolovaním z veľkého množstva zdrojových dát. Cloud / Hmlové počítanie. okrajový a sink (zoskupujúci) uzol

Vlastnosti senzorovej siete

- Autokonfigurácia senzora - reakcia na zmeny v sieti a prostredia pôsobenia
- Škálovateľnosť - veľké množstvo senzorov so spoločným účelom a schopnosťou vzájomnej kooperácie a interoperability.
- Odolnosť voči chybám - v prípade pridania alebo odobratia uzla budú spojenia bez prerušenia.
- Energeticky efektívna komunikácia uzlov - s upravenými protokolmi štandardného sieťového zásobníka
  - Event-driven - stály zber dát a reakcia na náhle zmeny. posielajú údaje až po prekročený kritického prahu
  - Query-driven - zbierajú údaje iba po prijatí dopytu od používateľa
  - Time-driven - pravidelne odosielajú údaje sinku. vzorkovaciu frekvenciu volí sink

[26]

Spracovanie toku informácií (IFP - Information flow processing) - nástroj na včasné spracovanie dát ako tečie z periférií do centra systému. Snahou je ukladanie

[4] https://www.mathworks.com/help/dsp/ug/envelope-detection.html
[5] https://www.dsprelated.com/showarticle/938.php

agregovaných štatistík, napr. detektor požiaru za použitia čidiel teploty a dymu nepotrebuje ukladať jednotlivé merania, lebo sú samo o sebe nepodstatné. Keď nastane varovná situácia, je potrebné aby tá obsahoval všetky údaje na lokalizáciu ohniska.

CEP - Complex event processing - spracúva toky udalosti zo zdrojov reálneho sveta na základe aplikovania aktívnych pravidiel stanovených správcami systému a poupraví toky do komplexnejšieho výstupu. Pravidlá sú v tvare Event-Condition-Action (ECA).

- Udalosť - definuje zdroje ako generátory udalostí
- Podmienka - uvažuje ktorá časť udalosti bude braná do úvahy pri spracovaní, napr. môže ísť o prekročenie prahu
- Akcia - aká sada úloh má byť vykonaná pri detekcii udalosti

[27]

Súčasti senzorovej jednotky: Zberná jednotka, Výpočtová jednotka, Komunikačná jednotka, Napájacia jednotka

Obmedzenia na senzorové uzly

- Spotreba energie - energetická autonómia uzlov vo WSN umožňuje nasadzovanie zariadení do odľahlých miest pre využitie v inteligentných mestách alebo na účely ochrany prírody. životnosť senzorovej jednotky je ohraničená kapacitou batérie.
- Dosah komunikácie - Senzory disponujú obmedzenou energiou na vysielanie a dosah je negatívne ovplyvnený silou signálu na anténe. Z toho vyplývajú aj nižšie prenosové rýchlosti.
- Výpočtový výkon a úložisko - Nízka taktovacia frekvencia procesora v megaherzoch a veľkosti pracovných pamätí v stovkách kilobajoch alebo megabajtoch.

[28]

Bluetooth, BLE - IEEE 802.15.1 bezdrôtová technológia s krátkym dosahom spravidla do 10 m (0,5mW pri 0.5m, 1mW pri 1m, 2.5mW pre 10m, 10mW pre

20m) na ultra krátkych vlnách 2.402 GHz - 2.48 GHz. SDP, RFCOMM náhradou za sériové porty + L2CAP. (symetrická komunikácia) `https://www.bluetooth.com/specifications/specs/rfcomm-1-2/` Wifi - 802.11 (klient-server cez prístupový bod) - TCP/IP sieťový zásobník, CoAP, MQTT na 6LoWPAN LoRa 863-870Hz IEEE802.15.4e (Smart Mesh IP) Time Synchronized Mesh Protocol (TSMP) The TSMP includes a Time Slotted Channel Hopping (TSCH) media access layer (MAC). TSCH works by dividing time into 'slots', and providing a mechanism to map time slots to channels with a pre-assigned hopping sequence. Zigbee

Literatúra

1. BROCH, Jens Trampe. *Mechanical Vibration and Shock Measurements*. 2. vyd. Brüel & Kjær, 1984. ISBN 87-87355-3-4-5.
2. MOHAMMED, Zakriya; ELFADEL, Ibrahim (Abe) M.; RASRAS, Mahmoud. Monolithic Multi Degree of Freedom (MDoF) Capacitive MEMS Accelerometers. *Micromachines*. 2018, roč. 9, č. 11. ISSN 2072-666X. Dostupné z DOI: `10.3390/mi9110602`.
3. TSAI, Ming-Han; LIU, Yu-Chia; SUN, Chih-Ming; WANG, Chuanwei; CHENG, Chun-Wen; FANG, Weileun. A 400×400µm2 3-axis CMOS-MEMS accelerometer with vertically integrated fully-differential sensing electrodes. In: *2011 16th International Solid-State Sensors, Actuators and Microsystems Conference*. 2011, s. 811–814. Dostupné z DOI: `10.1109/TRANSDUCERS.2011.5969154`.
4. DADAFSHAR, Majid. *Accelerometer and Gyroscopes Sensors: Operation, Sensing, and Applications*. 2014.
5. MÜLLER, Meinard. *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*. 1st. Springer, Inc., 2015. ISBN 978-3-319-21945-5. Dostupné z DOI: `10.1007/978-3-319-21945-5`.
6. *iNEMO inertial module: 3D accelerometer, 3D gyroscope, 3D magnetometer*. 2015. Č. LSM9DS1. Rev. 3.
7. *Micropower, 3-Axis, ±2 g/±4 g/±8 g Digital Output MEMS Accelerometer*. 2019. Č. ADXL362. Rev. F.
8. *Low Noise, Low Drift, Low Power, 3-Axis MEMS Accelerometers*. 2020. Č. ADXL356. Rev. A.

9. QUARTERONI, Alfio; SACCO, Riccardo; SALERI, Fausto. Numerical Mathematics. In: Springer Inc., 2000, kap. 9. Numerical Integration, s. 371–398. ISBN 0-387-98959-5.

10. YANG, Yanli; ZHAO, Yanfei; KANG, Dali. Integration on acceleration signals by adjusting with envelopes. *Journal of Measurements in Engineering,* 2016, roč. 4, s. 117–121.

11. MUTHUKRISHNAN, S. Data Streams: Algorithms and Applications. *Foundations and Trends in Theoretical Computer Science.* 2005, roč. 1, č. 2, s. 117–236. ISSN 1551-305X. Dostupné z DOI: 10.1561/0400000002.

12. PAJUREK, Tomáš. *Online Anomaly Detection in Time-Series.* Fakulta informačních technologií, České vysoké učení technické v Praze, 2018. Dostupné tiež z: https://dspace.cvut.cz/bitstream/handle/10467/76417/F8-DP-2018-Pajurek-Tomas-thesis.pdf. Dipl. pr.

13. NIELSEN, Aileen. *Practical Time Series Analysis: Prediction with Statistics and Machine Learning.* O'Reilly Media, 2019. ISBN 978-1-492-04165-8.

14. NEUBAUER, Jiří; SEDLÁČEK, Marek; KŘÍŽ, Oldřich. *Základy statistiky - Aplikace v technických a ekonomických oborech.* 1. vyd. Grada Publishing, a.s., 2012. ISBN 978-80-247-4273-1.

15. KNUTH, Donald E. The Art of Computer Programming. In: 2. vyd. Addison-Wesley, 1981, zv. 2, kap. 4.2.2, s. 216. ISBN 0-201-03822-6.

16. SCHNEIDER, Roger. Survey of Peaks/Valleys identificationin Time Series. 2011. Department of Informatics, University of Zürich.

17. YANG, Chao; HE, Zengyou; YU, Weichuan. Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis. *BMC bioinformatics.* 2009, roč. 10, s. 4. Dostupné z DOI: 10.1186/1471-2105-10-4.

18. ARGÜELLO-PRADA, Erick Javier. The mountaineer's method for peak detection in photoplethysmographic signals. *Revista Facultad de Ingeniería Universidad de Antioquia.* 2019, č. 90, s. 42–50. Dostupné z DOI: 10.17533/udea.redin.n90a06.

19. WEI, William W.S. Time Series Analysis: Univariate and Multivariate Methods. In: 2. vyd. Pearson Education, Inc., 2006, kap. 13, s. 289–318. ISBN 0-321-32216-9.

20. PRANDONI, Paolo; VETTERLI, Martin. *Signal Processing for Communications*. EPFL Press, 2008. ISBN 978-2-940222-20-9.

21. AHMED, Nasir; NATARAJAN, T; RAO, Kamisetty R. Discrete cosine transform. *IEEE transactions on Computers*. 1974, roč. 100, č. 1, s. 90–93.

22. KOVAČEV, Radovan. *Časově-frekvenční analýza signálu*. Fakulta informačních technologií, Vysoké učení technické v Brně, 2012. Dipl. pr.

23. CHU, Eleanor; GEORGE, Alan. *Inside the FFT Blackbox: Serial and Parallel Fast Fourier Transform Algorithms*. CRC Press LLC, 2000. Computational Mathematics. ISBN 0-8493-0270-6.

24. HEINZEL, G.; RÜDIGER, A.; SCHILLING, R. Spectrum and spectral density estimation by the Discrete Fourier transform (DFT), including a comprehensive list of window functions and some new at-top windows. In: 2002.

25. LYONS, Richard G. *Understanding Digital Signal Processing*. 3. vyd. Pearson Education, Inc., 2011. ISBN 978-0-13-702741-5.

26. MATIN, M.A.; ISLAM, M.M. Overview of Wireless Sensor Network. In: MATIN, Mohammad A. (ed.). *Wireless Sensor Networks*. IntechOpen, 2012, kap. 1, s. 1–22. ISBN 978-953-51-0735-4. Dostupné z DOI: 10.5772/49376.

27. CUGOLA, Gianpaolo; MARGARA, Alessandro. Processing Flows of Information: From Data Stream to Complex Event Processing. *ACM Computing Surveys*. 2012, roč. 44, č. 3. ISSN 0360-0300. Dostupné z DOI: 10.1145/2187671.2187677.

28. DJEDOUBOUM, Asside; ARI, Ado; GUEROUI, Abdelhak; MOHAMADOU, Alidou; ALIOUAT, Zibouda. Big Data Collection in Large-Scale Wireless Sensor Networks. *Sensors*. 2018, roč. 18. Dostupné z DOI: 10.3390/s18124474.