

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií

FIIT-1234-98765

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-1234-98765

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

Bakalársky projekt

Akademický rok:	2021/2022
Vedúci projektu:	Ing. Marcel Baláž, PhD.
Pedagogický vedúci:	Ing. Jakub Findura
Pridelený študentovi:	Miroslav Hájek
Študijný program:	Informatika
Miesto vypracovania:	Ústav počítačového inžinierstva a aplikovanej informatiky, FIIT STU

Spracovanie dát generovaných senzorovou IoT sieťou

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 nepretržitom monitorovaní už len z malého množstva senzorov predstavuje veľkú výzvu. Sensory 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 navrhnete spôsob ukladania a spracovania dát, prípadne optimalizáciu toku dát pre existujúcu senzorovú sieť. Vaše riešenie implementujte a otestujte jeho funkčnosť.

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

Pod'akovanie

Text pod'akovania

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: Processing of Data Generated by the Sensor IoT Network

Supervisor: Ing. Marcel Baláž, PhD.

Departmental advisor: Ing. Jakub Findura

2022, May

Annotation text in English, 150-200 words.

Obsah

1	Úvod	1
2	Analýza	3
2.1	Podsekcia	3
3	Opis riešenia	5
4	Zhodnotenie	7
A	Technická dokumentácia	
B	Harmonogram práce	
C	Obsah digitálneho média	

1 Úvod

Závěrečná správa (= finálny dokument bakalárskej práce) musí mať rozsah minimálne 20 normovaných strán (60 znakov na riadok, 30 riadkov na stranu). Toto platí pre hlavný obsah, teda obsah bez príloh a technickej dokumentácie.

Závěrečná správa o riešení bakalárskeho projektu by mala obsahovať súčasti uvedené nižšie. Presné pokyny sú súčasťou zadania bakalárskeho projektu, konzultujte s vedúcim bakalárskeho projektu.

Obsahové (aj formálne) požiadavky na priebežnú správu k riešeniu bakalárskeho projektu odovzdávané v rámci BP I sú stanovené rovnako po konzultácii s vedúcim bakalárskeho projektu a spravidla sú podmnožinou požiadaviek na záverečnú správu).

2 Analýza

Táto časť bakalárskeho projektu má:

- poskytovať obraz o stave riešenia daného problému známeho z preštudovanej literatúry (nielen informácie z prednášok, prípadne skrípt a katalógov),
- porovnanie podobných riešení, ich kategorizáciu s uvedením charakteristických atribútov atď., podľa charakteru bakalárskeho projektu
- zdôvodnenie voľby spôsobu riešenia a stručný opis celkového spôsobu riešenia (napr. v opise sa treba sústrediť na prípadné modifikácie použitých štandardných metodík a ich zdôvodnenie z hľadiska splnenia cieľov projektu)

2.1 Podsekcia

Text

3 Opis riešenia

Táto časť bakalárskeho projektu obsahuje opis výsledkov riešenia jednotlivých etáp projektu. V prípade, že záverečný projekt nerieši všetky etapy, malo by byť v príslušnej časti uvedené kto, resp. kde sa príslušná etapa rieši/riešila/bude riešiť.

Typické etapy riešenia pri tvorbe softvérového systému:

- špecifikácia požiadaviek
- návrh
- implementácia (ak to zadanie požaduje)
- overenie riešenia

Podľa možností treba vychádzať zo známych prístupov (napr. pri softvérových projektoch štruktúrovaný alebo objektovo orientovaný prístup) a techník (napr. blokové schémy, vývojové diagramy, UML, entito-relačné diagramy atď.). Táto časť práce závisí od konkrétneho zadania. Je dôležité prezentovať návrhové rozhodnutia, alternatívy, ktoré sa zvažovali pri riešení a samotný návrh riešenia zadaného problému. Štruktúrovanie textu tejto časti DP by malo vychádzať zo zadanej úlohy, ktorá sa rieši. Najmä v tejto časti študent preukazuje tvorivý prístup k riešeniu problémov a kritické myslenie.

4 Zhodnotenie

Hlavné výsledky práce, prípadne porovnanie s inými prístupmi, možné smery ďalšieho rozvíjania. Tu sa musí presne špecifikovať, čo je pôvodné a čo riešiteľ prebral.

Literatúra

1. LAW, Yee Wei; CHATTERJEA, Supriyo; JIN, Jiong; HANSELMANN, Thomas; PALANISWAMI, Marimuthu. Energy-efficient data acquisition by adaptive sampling for wireless sensor networks. In: *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*. Association for Computing Machinery, 2009, s. 1146–1151. IWCMC '09. ISBN 978-1-60558-569-7. Dostupné z DOI: 10.1145/1582379.1582631.
2. NIELSEN, Aileen. *Practical Time Series Analysis: Prediction with Statistics and Machine Learning*. O'Reilly Media, 2019. ISBN 978-1-492-04165-8.
3. WEI, William W.S. *Time Series Analysis: Univariate and Multivariate Methods*. 2. vyd. Pearson Education, Inc., 2006. ISBN 0-321-32216-9.
4. GIANNONI, Federico; MANCINI, M.; MARINELLI, Federico. Anomaly Detection Models for IoT Time Series Data. *ArXiv*. 2018, roč. abs/1812.00890. Dostupné tiež z: <https://arxiv.org/pdf/1812.00890.pdf>.
5. 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.
6. 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.
7. 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.

8. AGGARWAL, Charu C. *Outlier Analysis*. 2. vyd. Springer Publishing Company, Inc., 2016. ISBN 978-3-319-47578-3.
9. MEHROTRA, Kishan G.; MOHAN, Chilukuri K.; HUANG, HuaMing. *Anomaly Detection Principles and Algorithms*. Springer Publishing Company, Inc., 2017. ISBN 978-3-319-67524-4.
10. 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.
11. GUPTA, Manish; GAO, Jing; AGGARWAL, Charu C.; HAN, Jiawei. Outlier Detection for Temporal Data: A Survey. *IEEE Transactions on Knowledge and Data Engineering*. 2014, roč. 26, č. 9, s. 2250–2267. Dostupné z DOI: 10.1109/TKDE.2013.184.
12. ZHANG, Yang; MERATNIA, Nirvana; HAVINGA, Paul. Outlier Detection Techniques for Wireless Sensor Networks: A Survey. *IEEE Communications Surveys Tutorials*. 2010, roč. 12, č. 2, s. 159–170. Dostupné z DOI: 10.1109/SURV.2010.021510.00088.
13. BLÁZQUEZ-GARCÍA, Ane; CONDE, Angel; MORI, Usue; LOZANO, Jose Antonio. A review on outlier/anomaly detection in time series data. *CoRR*. 2020, roč. abs/2002.04236. Dostupné tiež z: <https://arxiv.org/abs/2002.04236>.
14. BRAEI, Mohammad; WAGNER, Sebastian. Anomaly Detection in Univariate Time-series: A Survey on the State-of-the-Art. *CoRR*. 2020, roč. abs/2004.00433. Dostupné tiež z: <https://arxiv.org/abs/2004.00433>.
15. BREUNIG, Markus M.; KRIEGEL, Hans-Peter; NG, Raymond T.; SANDER, Jörg. LOF: Identifying Density-Based Local Outliers. *SIGMOD Rec.* 2000, roč. 29, č. 2, s. 93–104. ISSN 0163-5808. Dostupné z DOI: 10.1145/335191.335388.
16. PRANDONI, Paolo; VETTERLI, Martin. *Signal Processing for Communications*. EPFL Press, 2008. ISBN 978-2-940222-20-9.

17. GURALNIK, Valery; SRIVASTAVA, Jaideep. Event Detection from Time Series Data. In: 1999, s. 33–42. Dostupné z DOI: 10.1145/312129.312190.
18. TRAN, Duc-Tan; LUU, Ha; NGUYEN, Tien Anh. A real-time vibration monitoring for vehicle based on 3-DOF MEMS accelerometer. In: 2010.
19. BASSEVILLE, Michèle; NIKIFOROV, Igor. *Detection of Abrupt Change Theory and Application*. 1993. ISBN 0-13-126780-9.
20. HOCHENBAUM, Jordan; VALLIS, Owen S.; KEJARIWAL, Arun. Automatic Anomaly Detection in the Cloud Via Statistical Learning. *CoRR*. 2017, roč. abs/1704.07706.
21. ARENAS-GARCÍA, Jerónimo; GÓMEZ-VERDEJO, Vanessa; NAVIA-VÁZQUEZ, Ángel. RLS Adaptation of One-Class SVM for Time Series Novelty Detection. 2004.
22. O'REILLY, Colin; GLUHAK, Alexander; IMRAN, Muhammad Ali; RAJASEGARAR, Sutharshan. Anomaly Detection in Wireless Sensor Networks in a Non-Stationary Environment. *IEEE Communications Surveys Tutorials*. 2014, roč. 16, č. 3, s. 1413–1432. Dostupné z DOI: 10.1109/SURV.2013.112813.00168.
23. KEOGH, E.; LIN, J.; TRUPPEL, W. Clustering of time series subsequences is meaningless: implications for previous and future research. In: *Third IEEE International Conference on Data Mining*. 2003, s. 115–122. Dostupné z DOI: 10.1109/ICDM.2003.1250910.
24. HYNDMAN, Rob J.; WANG, Earo; LAPTEV, Nikolay. Large-Scale Unusual Time Series Detection. In: *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*. 2015, s. 1616–1619. Dostupné z DOI: 10.1109/ICDMW.2015.104.
25. LIU, Jiang; WANG, Pengcheng; TIAN, Xincheng. Vibration displacement measurement based on three axes accelerometer. In: *2017 Chinese Automation Congress (CAC)*. 2017, s. 2374–2378. Dostupné z DOI: 10.1109/CAC.2017.8243173.

26. VARANIS, Marcus; SILVA, Anderson; MERELES, Arthur; PEDERIVA, Robson. MEMS accelerometers for mechanical vibrations analysis: a comprehensive review with applications. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*. 2018, roč. 40. Dostupné z DOI: 10.1007/s40430-018-1445-5.
27. AHMED, Nasir; NATARAJAN, T; RAO, Kamisetty R. Discrete cosine transform. *IEEE transactions on Computers*. 1974, roč. 100, č. 1, s. 90–93.
28. DADAFSHAR, Majid. *Accelerometer and Gyroscopes Sensors: Operation, Sensing, and Applications*. 2014.
29. VERMA, Nishchal K.; GUPTA, Rajat; SEVAKULA, Rahul K.; SALOUR, Al. Signal transforms for feature extraction from vibration signal for air compressor monitoring. In: *TENCON 2014 - 2014 IEEE Region 10 Conference*. 2014, s. 1–6. Dostupné z DOI: 10.1109/TENCON.2014.7022275.
30. POPIVANOV, I.; MILLER, R.J. Similarity search over time-series data using wavelets. In: *Proceedings 18th International Conference on Data Engineering*. 2002, s. 212–221. Dostupné z DOI: 10.1109/ICDE.2002.994711.
31. GHOSH, Ripul; AKULA, Aparna; KUMAR, Satish; SARDANA, Hk. Time–frequency analysis based robust vehicle detection using seismic sensor. *Journal of Sound and Vibration*. 2015, roč. 346, s. 424–434. Dostupné z DOI: 10.1016/j.jsv.2015.02.011.
32. SHIM, JaeSeok; KIM, GeoYoung; CHO, ByungJin; KOO, JeongSeo. Application of Vibration Signal Processing Methods to Detect and Diagnose Wheel Flats in Railway Vehicles. *Applied Sciences*. 2021, roč. 11, č. 5. ISSN 2076-3417. Dostupné z DOI: 10.3390/app11052151.
33. HAN, Sangbo. Measuring displacement signal with an accelerometer. *Journal of Mechanical Science and Technology*. 2010, roč. 24, s. 1329–1335. Dostupné z DOI: 10.1007/s12206-010-0336-1.
34. 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.

35. 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.
36. 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.
37. CHOWDHURY, Ranak Roy; ADNAN, Muhammad Abdullah; GUPTA, Rajesh K. Real-Time Principal Component Analysis. *ACM/IMS Trans. Data Sci.* 2020, roč. 1, č. 2. ISSN 2691-1922. Dostupné z DOI: 10.1145/3374750.
38. VETTERLI, Martin; KOVAČEVIC, Jelena. *Wavelets and Subband Coding*. Prentice-Hall, Inc., 1995. ISBN 978-0-13-097080-0.
39. PALSHIKAR, Girish. Simple Algorithms for Peak Detection in Time-Series. 2009.
40. BROCH, Jens Trampe. *Mechanical Vibration and Shock Measurements*. 2. vyd. Brüel & Kjær, 1984. ISBN 87-87355-3-4-5.
41. LYONS, Richard G. *Understanding Digital Signal Processing*. 2. vyd. Pearson Education, Inc., 2004. ISBN 0-13-108989-7.
42. 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.
43. MUÑOZ, Arrate; ERTLÉ, Raphaël; UNSER, Michael. Continuous wavelet transform with arbitrary scales and $O(N)$ complexity. *Signal Processing*. 2002, roč. 82, č. 5, s. 749–757. ISSN 0165-1684. Dostupné z DOI: [https://doi.org/10.1016/S0165-1684\(02\)00140-8](https://doi.org/10.1016/S0165-1684(02)00140-8).
44. NIST/SEMATECH. *e-Handbook of Statistical Methods: Generalized ESD Test for Outliers*. 2012. Dostupné z DOI: <https://doi.org/10.18434/M32189>.

Príloha A: Technická dokumentácia

Prílohy dopĺňajú hlavnú časť práce. Obsahujú napríklad podrobné informácie k jednotlivým etapám riešenia projektu. Typicky sa tu uvádza aj podstatná časť technickej dokumentácie. Pozor, prílohy nesmú obsahovať také informácie, ktoré sú pre pochopenie práce kľúčové. Tie musí obsahovať hlavná časť práce, ktorá musí byť úplná, celistvá.

Súčasťou príloh nie je len textový obsah, ale aj ďalšie artefakty, ktoré sú výsledkom projektu, napr. počítačový kód, dátové vzorky, vedecký článok či plagát. Zvláštnu pozornosť venujte tým artefaktom, ktoré sú potrebné pre replikovateľnosť postupov opisovaných v práci (napr. aby mohol oponent pri vyhodnocovaní práce zopakovať uvádzané postupy a prísť k rovnakým záverom).

Digitálne artefakty sa prikladajú na elektronickom médiu. K akémukoľvek digitálnemu obsahu treba uviesť v dokumente priebežnej či záverečnej správy bakalárskej/diplomovej práce primeraný textový opis, preto nezabudnite digitálne médium zdokumentovať. Prínajmenšom medzi prílohy zaraďte kapitolu „Obsah elektronického média“. Na prílohy sa nezabudnite z hlavnej časti práce primerane odkazovať.

Obsah technickej dokumentácie závisí od povahy riešeného problému. Uvádza sa technická dokumentácia k systému (počítačový, softvérový), ktorý bol vytvorený v rámci riešenia projektu (ak sa toto v zadaní požadovalo). Samotný obsah a rozsah závisí aj od účelu vytvoreného systému (produkt, experimentovanie a pod.) V prípade softvérového systému technická dokumentácia spravidla obsahuje časti v náväznosti na etapy tvorby softvérového systému:

- dokumentáciu k etape špecifikácie požiadaviek
- dokumentáciu k etape návrhu projektu
- dokumentáciu k implementácii

- v prípade, že súčasťou riešenia sú programy, dokumentáciu k implementácii tvoria zdrojové texty programov
- v prípade, že súčasťou riešenia je návrh zariadenia, dokumentáciu k implementácii tvorí technická dokumentácia (schémy zapojenia, návrh dosiek plošných spojov, schémy rozmiestnenia súčiastok, zoznam použitých súčiastok, opis konektorov atď.)
- dokumentáciu k overeniu riešenia
- dokumentáciu k používaniu a údržbe (návod na použitie a údržbu projektu)

Príloha B: Harmonogram práce

B.1 Zimný semester

B.2 Letný semester

Príloha C: Obsah digitálneho média

Evidenčné číslo práce v informačnom systéme: FIIT-1234-98765

Obsah digitálnej časti práce (archív ZIP):

Názov odovzdaného archívu: ...zip