

COMPUTER ENGINEERING THESIS GUIDELINES

A Thesis

Presented to

The Faculty of the Department of Computer Engineering
San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

David C. Anastasiu

December 2019

© 2019

David C. Anastasiu

ALL RIGHTS RESERVED

The Designated Thesis Committee Approves the Thesis Titled

COMPUTER ENGINEERING THESIS GUIDELINES

by

David C. Anastasiu

APPROVED FOR THE DEPARTMENT OF COMPUTER ENGINEERING

SAN JOSÉ STATE UNIVERSITY

December 2019

Advisor Name, Ph.D.	Department of Computer Engineering
---------------------	------------------------------------

Committee Member Name, Ph.D.	Department of Committee Member
------------------------------	--------------------------------

Committee Member Name, Ph.D.	Department of Committee Member
------------------------------	--------------------------------

ABSTRACT

COMPUTER ENGINEERING THESIS GUIDELINES

by David C. Anastasiu

This “thesis” provides a visual example of how a San José State University (SJSU) Computer Engineering (CMPE) department thesis should be formatted. Our LaTeX thesis template follows the IEEE Computer Society (compsoc) journal style in most respects, yet it complies with the SJSU Graduate and Undergraduate Programs (GUP) thesis guidelines where those guidelines supersede our chosen publication style. Note that the latest GUP guidelines always take precedence over this template or style. If conflicts arise, please inform the template maintainer so they can be promptly corrected. In general, students should thoroughly read the latest Graduate Studies thesis guidelines before writing their thesis. The guidelines contain a lot of important information beyond style requirements, including directions for tone of language in the thesis, suggestions for correcting common language-related errors, and procedures and deadlines for submitting the thesis. The LaTeX template project that this thesis is part of was created as a way to simplify GUP thesis style compliance for CMPE and College of Engineering students. Beyond style information, this document will also provide instructions to properly use the LaTeX style and some general instructions for doing research and writing LaTeX documents.

ACKNOWLEDGMENTS

I would like to thank my advisor, for being the greatest advisor in the world.

I would also like to thank Cheryl R. Cowan, SJSU Graduate Studies Associate, for assisting me in creating this LaTeX template and ensuring that it accurately follows the Graduate Studies thesis style requirements.

Last, but not least, I would like to thank my family. Without their support, I would have never been able to complete this thesis or pursue my dreams.

TABLE OF CONTENTS

List of Tables	vii
List of Figures	viii
1 Introduction.....	1
2 Literature Review	3
2.1 How to Write a Literature Review	3
2.2 Activity Recognition.....	3
2.3 Multivariate Time Series Segmentation	5
2.4 Where to Find More Information About Literature Reviews	6
3 Sections and Headings	7
3.1 Second Section Level.....	7
3.1.1 Third Section Level	8
3.1.1.1 Fourth section level	8
4 Example Sections	9
5 Mini-introduction to LaTeX	10
5.1 Equations	10
5.2 Figures.....	10
5.3 Tables	11
5.4 Algorithms	12
5.5 Location of Floats	13
5.6 Citations	14
6 Writing Advice	16
7 Future Work	17
8 Conclusions	18
Literature Cited.....	19
Appendix A: What Should Be Included in Appendices	22
A.1 Subsections in Appendices	22
Appendix B: Title of Another Appendix	24

LIST OF TABLES

Table 1.	Dataset Characteristics	11
Table 2.	Graph and Projection Notation	13

LIST OF FIGURES

Fig. 1. Project architecture.	11
Fig. 2. Distribution of sample non-zeros.	23
Fig. 3. Speedup comparison of our method against baselines.....	24

1 INTRODUCTION

“SJSUthesis” is a LaTeX class created by David C. Anastasiu for Masters thesis authors from the College of Engineering (CoE) at San José State University (SJSU). It follows the current Graduate and Undergraduate Programs (GUP) thesis guidelines, currently found at <http://www.sjsu.edu/gup/gradstudies/thesis/index.html>, and attempts to simplify the formatting of theses for students. Students should, however, read the current GUP requirements carefully and rectify any style discrepancies before submitting their theses. Additionally, note that GUP is very strict on grammar and spelling mistakes.

The present “thesis” serves as a visual example of what the finished thesis product should look like. In addition, it provides some guidance on how to use the SJSUthesis style, the provided LaTeX template project, and LaTeX in general to produce a quality thesis that complies with GUP requirements. LaTeX, in combination with the provided style, automate many of the style aspects of the thesis, as long as students do not modify the style or introduce new style elements manually in the text. For example, one can make all section titles bold by simply putting the text within a `\textbf` block. However, that would break the style and cause the thesis to be returned for edits by GUP. Students are thus, once again, urged to refrain from style modifications and to double-check style compliance against the GUP thesis guide before submitting their theses.

Throughout the thesis, we will make use of tables and figures from some of my own papers and my students’ theses. These tables and figures will likely be taken out of context and should not be considered relevant to the topic of this thesis. They are simply provided as examples for how to properly include these elements in a thesis document.

The remainder of this thesis is structured as follows. First, in Chapter 2, we will introduce the reader to the task of reviewing literature that is relevant to the chosen topic of study, with the goal of identifying the current state-of-the-art method or system that potentially solves the problem at hand. This is the basis of all research projects, as we

should not “re-invent the wheel” and try not to repeat mistakes made by others. Chapter 4 introduces an example of other chapters one may include in the thesis. Chapter 5 serves as a brief introduction to writing documents in the TeX macro language. We provide a list of writing tips, generally collected from correcting student theses over the years, in Chapter 6. Finally, Chapter 7 describes to-do items envisioned for this project and Chapter 8 concludes this thesis.

2 LITERATURE REVIEW

2.1 How to Write a Literature Review

This chapter will contain some practical advice for starting the research journey by finding the state-of-the-art for the project at hand. However, until we have time to write this chapter, we are including some related works from an in-progress project.

In this chapter, we review some supervised and unsupervised techniques that are used to identify human activities using mobile sensor data. Additionally, we review some methods that can help identify appropriate features that need to be extracted from the raw sensor data to better identify activity patterns.

2.2 Activity Recognition

Appropriate datasets are hard to find for human activity recognition (HAR) research, which lead many authors to collect their own data using several sensors. Chen and Shen [1] presented performance analysis based on the placement of smartphones on various parts of a subject's body. In their data collection phase, they captured data from a combination of different sensors, e.g., only accelerometer versus combined accelerometer and gyroscope, from smartphones attached to the forearm, abdomen, or legs of the subject, with the ultimate goal of classifying the type of activity being performed. Su et al. [2] computed features in two different domains: time and frequency. In the time domain, they calculated mean, max-min, standard deviation, correlation, and signal magnitude area (SMA). In the frequency domain, they calculated energy, entropy, binned distribution, and the time difference between peaks. They tested a number of supervised learning methods for activity recognition, such as using decision tree, naïve Bayes, and support vector machine (SVM) classification models.

Anguita et al. [3] focused on labeled data collection using built-in smartphone sensors (accelerometer, gyroscope at 50Hz) with six predefined activity labels. To validate the quality of the dataset, they used the SVM classifier, after filtering noise, and feature

mapping to a randomly-split test/train dataset. Kim et al. [4] summarized pattern discovery in supervised HAR approaches, including hidden Markov model (HMM), linear chain conditional random fields (CRF), and skip-chain CRF (SCCRF). Several approaches have been devised for unsupervised learning for human activity recognition. Bhattacharya et al. [5] and Li et al. [6] compared sparse encoding and principle component analysis (PCA) for activity recognition from unlabeled sensor data and found sparse encoding to be better. In [6], Li et al. also compared denoising autoencoder (DAE) and fast fourier transform (FFT) techniques for feature extraction and found sparse autoencoder to provide better results than other techniques.

Chamroukhi et al. [7] described a method to identify hidden discrete logistic processes that maximize the likelihood of data being observed, which they trained via a dedicated expectation maximization (EM) algorithm. Similarly, Kwon et al. [8] compared various clustering algorithms for labeling a mixed bag of known and unknown activities. Results showed that given a number of activities, GMM seems to outperform k -means and hierarchical clustering for a given known number of activities, but DBSCAN [9] proved the best choice when the number of activities is unknown.

Wang et al. [10] discussed linear dynamical systems for human activity recognition, which uses k -medoids clustering and a bag-of-systems (BoS) to predict activities involving motion, such as jumping and running. Ronao and Cho [11] proposed using a deep convolutional neural networks (DCNN) to extract features from the time series data. They found that, as the number of layers increases, the classification performance improved but the complexity of derived features decreased. Analysis of additional algorithms showed that DCNNs with hand designed features yield better results. Gjoreski and Roggen [12] proposed a method called “Unsupervised Online Activity Discovery Using Temporal Behavior Assumption” (UnADevs) based on an online clustering which includes the temporal information of the occurring activities. It is well suited to identify

clusters of repeating activities while keeping track of the time interval of discovered activity clusters.

Most of the works in the area of human activity recognition involves smartphones, which have many built-in sensors which can generate diverse signal shapes given the phone's position. These signals can be used to recognize many different human activities. Existing works have used many different combinations of the sensors and machine learning techniques, but are primarily focused on identifying base activities (sitting, running), while our method is designed to characterize human activity patterns and identify how these patterns change over time.

2.3 Multivariate Time Series Segmentation

Time series segmentation has been an area of interest for many research communities, including signal processing, pattern recognition, machine learning and language processing. Abonyi et al. [13] presented fuzzy clustering as one such approach but it requires clusters to be contiguous in time. It uses probabilistic PCA models to measure homogeneity of the segments and fuzzy sets. To overcome this limitation, Wang et al. proposed an improved unsupervised method [14] that can automatically determine the optimal segmentation order.

Dynamic programming (DP) algorithms have been widely explored to automatically segment multivariate time series [15], [16]. In one example, Anastasiu et al. [16] proposed an algorithm for optimal cross-user segmentation for identifying how users use their computers and how that usage changes over time. Our method is similar in scope, but focused on automatic discovery of human activity patterns. A similar dynamic programming based approach was proposed by Guo et al. [17]. It initially applies fuzzy clustering to predefined segments and uses dynamic time warping to determine distances between non-equal length series. Segments are then iteratively updated through a segmentation objective function optimized using dynamic programming.

2.4 Where to Find More Information About Literature Reviews

This section will contain some relevant references on this topic.

3 SECTIONS AND HEADINGS

A thesis is generally broken down into sections, each section encompassing a different area of the work. For example, the first section in a thesis is generally an “Introduction”, which lays out what the research is about. The top-level sections in the thesis are called chapters and always start at the top of a page. When referring to a top-level section, one should mention it as Chapter 1, where the number is the section/chapter number. The numbers, as with all references in the thesis, should never be written out. Rather, use LaTeX reference macros, which will be automatically replaced with appropriate numbers when compiling the thesis document. Chapter 5, as well as the source code of this document, has many examples of references.

A thesis section can be further broken down into subsections. One should use at most 4 section levels in a thesis. This paragraph is in the 1st section level, “section”. We will create 3 more subsection levels. The 2nd section level is created using the “subsection” macro. The 3rd section level is created using the “subsubsection” macro. However, the 4th section level is created using the “paragraph” macro. Note that the IEEE paragraph style is slightly different than the section styles, as the paragraph caption will be in-line with the text of the section and will be automatically ended with a colon. Students should note that more than 3 subheadings (4 section levels) becomes confusing to the reader and should thus avoid further breakdown of paragraph sections.

Note that the section header is written in title case, though it appears in the text as uppercase. Students should not deviate from using title case, as it will change the appearance of the section title in the table of contents. The section title should be written in title case.

3.1 Second Section Level

This is an example of the 2nd level of a section. The subsection title should be written in title case.

3.1.1 Third Section Level

This is an example of the 3rd level of a section. The subsubsection title should be written in title case.

3.1.1.1 Fourth section level: This is an example of the 4th level of a section. Note that the paragraph title should be written in sentence case, unlike the higher level titles.

4 EXAMPLE SECTIONS

In this chapter we discuss a possible set of chapters for a thesis. The example is somewhat specific to the author's field of research (Data Science). Students should, in general, consult with their adviser on the list of chapters that they should include in their thesis.

5 MINI-INTRODUCTION TO LATEX

This chapter needs to be filled out. For now, it contains some very basic examples.

5.1 Equations

Equations and mathematical formulas can be included in-line, within the text, for example when explaining that $\|\mathbf{x}\|_2$ signifies the ℓ_2 norm of a vector, as long as the included math formula or symbols are short (less than half a line). For longer formulas, or formulas that need to be referenced later in the text, one should use the `equation` macro or related macros, e.g., `align`. For example, the dot-product of two vectors, \mathbf{x} and \mathbf{y} , is defined as

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{j=1}^m x_j \times y_j. \quad (1)$$

Note that equations can be assigned labels and referenced throughout the text. However, unlike figures and tables (see next sections), the equation should be defined prior to its reference. Now that we can defined the formula for the vector dot-product, we can now refer to Equation 1 to remind readers of its definition.

5.2 Figures

Figures should contain a caption at the bottom. The caption is a sentence or phrase that describes the content of the figure. As such, it should start with a capital letter and end with a period. The caption is not a title, so subsequent words after the first word should not be capitalized unless they are proper names. A figure should be introduced in the text prior to its inclusion in the text. When referring to the figure, make use of the `\figurename` macro, provided by the IEEE style, to ensure style compliance. Additionally, use a non-breaking space character to connect the Fig. text with the number assigned to the figure reference. This will ensure that Fig. and the reference will be included on the same line. For example, Fig. 1 introduces our project architecture. Note

that font sizes in figures and tables may be smaller than the rest of the thesis. However, they should be readable.

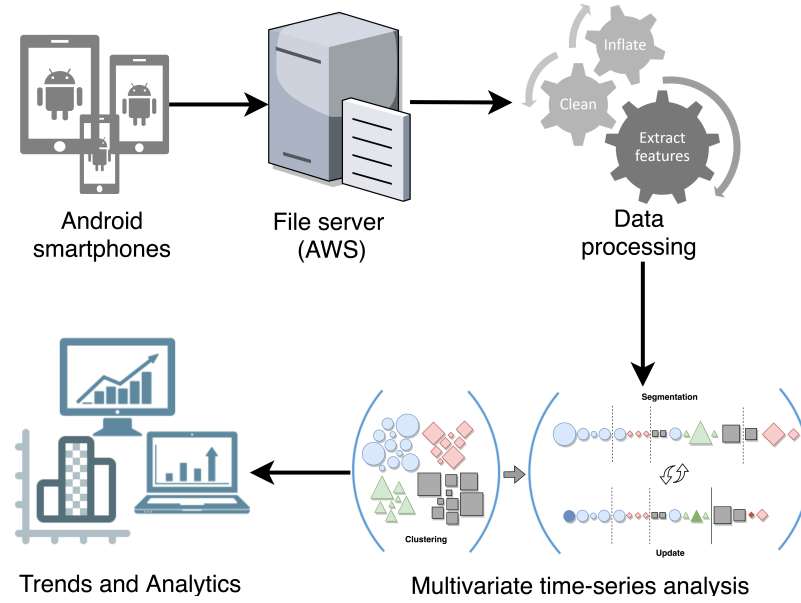


Fig. 1. Project architecture.

5.3 Tables

Tables should contain a caption at the top. Unlike figures, the caption in the table is the table title, and it should be capitalized accordingly. The table should be presented in the text prior to its inclusion in the text. Moreover, references to tables should follow the same directions as those for figures, except use the `\tablename` macro. For example, Table 1 introduces statistics about the dataset used in our analysis.

Table 1
Dataset Characteristics

Name	Duration (h)	Activities	Subjects	Sensors	Frequency (Hz)
JSI-ADL	18.8	14	10	2	50
REALDISP	9	33	17	2	50
UCI-HAR	1.5	6	30	2	50
Ours	1500	14	24	2	20

5.4 Algorithms

LaTeX provides powerful macros for writing algorithms. Note that labels can be included within the algorithm, which allows the writer to refer to specific lines in the algorithm. The algorithm should be introduced first. Note that the label referring to the whole algorithm is placed right after the caption. Similar to a table, the caption is a title and should be capitalized accordingly. For example, Algorithm 1 describes the process for producing a normalized 2-hop left-projection of a bipartite graph described by the vertex sets V^1 and V^2 and the edge set E . Long algorithms (greater than half a page) should be relegated to the Appendix.

Algorithm 1 The Normalized 2-hop Left-Projection Algorithm

```

1: procedure PROJECT( $V^1, V^2, E$ )                                ▷ Construct the left projection  $G^1$ 
    // Initialize and compute degree vectors.
2:   for  $(v_i^1, v_j^2) \in E$  do                                     ▷ Count neighbors
3:      $\lambda_n(i) \leftarrow \lambda_n(i) + 1$ 
4:      $\lambda_m(j) \leftarrow \lambda_m(j) + 1$ 
    // SpGEMM
5:   for  $v_i^1 \in V^1$  do
6:      $acc(\cdot) \leftarrow 0$                                          ▷ Initialize output row accumulator data structure
7:     for  $v_k^2 \in \Gamma(v_i^1)$  do
8:        $l \leftarrow (1/\lambda_n(i)) \times w_{i,k} \times (1/\lambda_m(k))$ 
9:       for  $v_j^1 \in \Gamma(v_k^2)$  do
10:        if  $v_j^1 \neq v_i^1$  then                                     ▷ Avoid computing output diagonal
11:           $acc(j) \leftarrow acc(j) + l \times w_{j,k}$ 
12:       $C(i, \cdot) = acc(\cdot)$ 
    // Normalize rows of  $C$ 
13:   for  $i = 1, \dots, |V^1|$  do
14:      $n_i \leftarrow 0$ 
15:     for  $j = 1, \dots, |V^2|$  s.t.  $C(i, j) > 0$  do
16:        $n_i \leftarrow n_i + \text{abs}(C(i, j))$ 
17:     for  $j = 1, \dots, |V^2|$  s.t.  $C(i, j) > 0$  do
18:        $C(i, j) \leftarrow C(i, j)/n_i$ 
19:   return  $C$                                                     ▷ Edge set weight matrix for  $G^1$ 

```

After introducing the algorithm, one can then refer to sections of the algorithm in the process of describing it in the text. For example, note that line 11 denotes the accumulation of projection weights for a vertex v_i^1 in V^1 . For theses that contain many

mathematical symbols, it is a good idea to provide the reader with a reference for the meaning of these symbols. This table is generally included in the preliminary chapters, after the introduction. We will include an example here. Table 2 provides a reference for notation used throughout the paper that the algorithm above is included in.

Table 2
Graph and Projection Notation

Symbol	Meaning
G	A graph topology.
V^1, V^2	The left and right node partitions in a bipartite graph.
E	The set of edges in G .
$ X $	The cardinality of set X .
$v_i^1 \in V^1$	A vertex in the left node partition V^1 .
$v_j^2 \in V^2$	A vertex in the right node partition V^2 .
$(v_i^1, v_j^2, w_{i,j}) \in E$	The edge connecting nodes v_i^1 and v_j^2 with weight $w_{i,j}$ in graph G .
$\Gamma(v_i^1)$	The set of vertices in V^2 adjacent to v_i^1 , i.e., its neighborhood.
λ_n, λ_m	Node degree vectors for the left and right partitions.
$A \in \mathbb{R}^{n \times m}$	The edge weight matrix for the bipartite graph G ; $A(i, j) = w_{i,j}$.
B	Weighted adjacency matrix for graph G .
$G^1 = (V^1, E^1)$	The left projection graph.
$G^2 = (V^2, E^2)$	The right projection graph.
$G^b = (V, E^b)$	The biprojection graph.
$(v_i^1, v_j^1, s_{i,j}^1) \in E^1$	The edge connecting v_i^1 and v_j^1 with weight $s_{i,j}^1$ in the left projection graph.
W^1	Edge weight matrix for the left projection graph; $W^1(i, j) = s_{i,j}^1$.
W^2	Edge weight matrix for the right projection graph; $W^2(i, j) = s_{i,j}^2$.
W	Weighted adjacency matrix for the bi-projection graph of G .

5.5 Location of Floats

Floats (e.g., algorithms, tables and figures) should not follow the IEEE standard of being placed at the top or bottom of the page. As per the GUP thesis guidelines, they must appear right after their reference and not break any paragraphs, i.e., they should be placed in the text, after the paragraph in question. Use the location hint `[!htb]` to let LaTeX know the float should be placed in between paragraphs. If the float takes up an entire page and the last paragraph on the current page does not fit in its entirety, it should be moved after the float page to avoid breaking the continuity of the paragraph, leaving some blank lines at the end of the current page.

Pages that contain only floats will have floats vertically centered, which will produce additional blank lines between the floats in the page or between a float and the top and bottom of the page. The thesis guidelines specifically prohibit additional white space in the thesis. This case and the one discussed in the previous paragraph are the only acceptable reasons for having extra blank lines in the text.

5.6 Citations

When citing papers, it is advisable to collect BibTeX records for the papers that you wish to cite, which will be automatically processed into the correct citation style by the LaTeX/BibTeX compiler. Citing a work by Anastasiu and Karypis [18], for example, is as simple as writing “Anastasiu and Karypis~\cite{anastasiu-dsaa2016},” where *anastasiu-dsaa2016* is the key associated with the paper’s BibTeX record in the *references.bib* file. Note the non-breaking space character (~) between the word before the citation and the \cite command. It gets translated into a space, and does not allow LaTeX to print the citation on a different line than the connecting word. This, of course, does not work if you put a space between the word and the non-breaking space character, e.g., “Karypis ~\cite{anastasiu-dsaa2016}”.

When citing papers with a single author, one should refer to the work by citing the author’s last name. For example, Anastasiu [19] used a combination of heuristic candidate selection and search space filtering to efficiently identify approximate nearest neighbors. When the work has two authors, both authors should be referenced by their last name, as in the previous example of the work by Anastasiu and Karypis [18]. Finally, if there are more than two authors, one should reference the first author’s last name, followed by the phrase *et al.*, which is short for the Latin *et alia* and means “and others”. Note the placement of the period. As an example, Kapoor et al. [20] discuss the benefits of competitive active learning in engineering education. Note that it is not necessary to mention authors by last name when listing multiple works as examples or evidence of a

statement. For example, many recent works have developed filtering-based methods for nearest neighbor graph construction [16], [21]–[23].

6 WRITING ADVICE

This chapter contains a list of common writing mistakes that students tend to make when writing theses. Students should read the list carefully, in addition to the examples provided by GUP in their thesis guidelines, and guard against making these mistakes in their theses.

7 FUTURE WORK

First of all, we plan to finish this template example and, once the style/template has been approved by GUP, disseminate it to others in the CoE. We plan to invite contributions from other CMPE and CoE faculty in improving this document, with the goal to include the most useful information to help students succeed in writing their thesis.

8 CONCLUSIONS

This macro example thesis was written as a guide to CMPE and CoE students for writing quality theses in LaTeX. The provided template, along with this guide, should make the task of complying with GUP formatting requirements much easier for students. The CMPE LaTeX thesis template and this document are a work in progress. The reader is invited to participate in improving both via Git pull requests. The project can be found at https://github.com/davidanastasiu/thesis_template-SJSU_CMPE.

Literature Cited

- [1] Y. Chen and C. Shen, "Performance analysis of smartphone-sensor behavior for human activity recognition," *IEEE Access*, vol. 5, pp. 3095–3110, 2017.
- [2] X. Su, H. Tong, and P. Ji, "Activity recognition with smartphone sensors," *Tsinghua Science and Technology*, vol. 19, pp. 235–249, June 2014.
- [3] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, ESANN 2013, 01 2013.
- [4] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *IEEE Pervasive Computing*, vol. 9, pp. 48–53, Jan 2010.
- [5] S. Bhattacharya, P. Nurmi, N. Hammerla, and T. Pltz, "Using unlabeled data in a sparse-coding framework for human activity recognition," *Pervasive and Mobile Computing*, vol. 15, pp. 242 – 262, 2014. Special Issue on Information Management in Mobile Applications Special Issue on Data Mining in Pervasive Environments.
- [6] Y. Li, D. Shi, B. Ding, and D. Liu, "Unsupervised feature learning for human activity recognition using smartphone sensors," in *Mining Intelligence and Knowledge Exploration* (R. Prasath, P. O'Reilly, and T. Kathirvalavakumar, eds.), (Cham), pp. 99–107, Springer International Publishing, 2014.
- [7] F. Chamroukhi, S. Mohammed, D. Trabelsi, L. Oukhellou, and Y. Amirat, "Joint segmentation of multivariate time series with hidden process regression for human activity recognition," *Neurocomputing*, vol. 120, pp. 633 – 644, 2013. Image Feature Detection and Description.
- [8] Y. Kwon, K. Kang, and C. Bae, "Unsupervised learning for human activity recognition using smartphone sensors," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6067 – 6074, 2014.
- [9] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, pp. 226–231, AAAI Press, 1996.

- [10] W. Wang, H. Liu, L. Yu, and F. Sun, “Human activity recognition using smart phone embedded sensors: A linear dynamical systems method,” in *2014 International Joint Conference on Neural Networks (IJCNN)*, pp. 1185–1190, July 2014.
- [11] C. A. Ronao and S.-B. Cho, “Deep convolutional neural networks for human activity recognition with smartphone sensors,” in *Neural Information Processing* (S. Arik, T. Huang, W. K. Lai, and Q. Liu, eds.), (Cham), pp. 46–53, Springer International Publishing, 2015.
- [12] H. Gjoreski and D. Roggen, “Unsupervised online activity discovery using temporal behaviour assumption,” in *Proceedings of the 2017 ACM International Symposium on Wearable Computers, ISWC ’17*, (New York, NY, USA), pp. 42–49, ACM, 2017.
- [13] J. Abonyi, B. Feil, S. Nemeth, and P. Arva, “Modified gath–geva clustering for fuzzy segmentation of multivariate time-series,” *Fuzzy Sets Syst.*, vol. 149, pp. 39–56, Jan. 2005.
- [14] N. Wang, X. Liu, and J. Yin, “Improved gath–geva clustering for fuzzy segmentation of hydrometeorological time series,” *Stochastic Environmental Research and Risk Assessment*, vol. 26, pp. 139–155, Jan 2012.
- [15] H. Guo, X. Liu, and L. Song, “Dynamic programming approach for segmentation of multivariate time series,” *Stochastic Environmental Research and Risk Assessment*, vol. 29, pp. 265–273, Jan 2015.
- [16] D. C. Anastasiu, A. M. Rashid, A. Tagarelli, and G. Karypis, “Understanding computer usage evolution,” in *31st IEEE International Conference on Data Engineering, ICDE 2015*, pp. 1549–1560, 2015.
- [17] H. Guo and X. Liu, “Dynamic programming-based optimization for segmentation and clustering of hydrometeorological time series,” *Stochastic Environmental Research and Risk Assessment*, vol. 30, pp. 1875–1887, Oct 2016.
- [18] D. C. Anastasiu and G. Karypis, “Efficient identification of tanimoto nearest neighbors,” in *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, DSAA ’16, pp. 156–165, 2016.
- [19] D. C. Anastasiu, “Cosine approximate nearest neighbors,” in *Data Science – Analytics and Applications*, iDSC 2017, (Wiesbaden), pp. 45–50, Springer Fachmedien Wiesbaden, 2017.

- [20] M. Kapoor, S. Hua, and D. C. Anastasiu, “Improving student motivation through competitive active learning,” in *2018 IEEE Frontiers in Education Conference*, FIE 2018, IEEE, October 2018.
- [21] D. C. Anastasiu and G. Karypis, “L2ap: Fast cosine similarity search with prefix l-2 norm bounds,” in *The 30th IEEE International Conference on Data Engineering*, ICDE 2014, pp. 784–795, 2014.
- [22] D. C. Anastasiu and G. Karypis, “Pl2ap: Fast parallel cosine similarity search,” in *Proceedings of the 5th Workshop on Irregular Applications: Architectures and Algorithms, in conjunction with SC’15*, IA3 2015, (New York, NY, USA), pp. 1–8, ACM, 2015.
- [23] D. C. Anastasiu and G. Karypis, “Fast parallel cosine k-nearest neighbor graph construction,” in *2016 6th Workshop on Irregular Applications: Architecture and Algorithms (IA3)*, pp. 50–53, Nov 2016.

Appendix A

WHAT SHOULD BE INCLUDED IN APPENDICES

Appendices should contain information that is too lengthy to be included in the thesis chapters but further support the conclusions of the thesis. For example, one could include additional experimental results in the form of tables or additional figures. Each appendix should start with a paragraph introducing the items being presented. Additionally, each table or figure should be preceded by a paragraph that explains the data being presented and the conclusions that can be drawn from them. Please note that data already presented in the main thesis should not be repeated in an appendix. One cannot, for example, include the results of an experiment as a figure in the thesis and again as a table in the appendix. One can, however, include detailed results in tables in the appendix and a summary figure in the thesis portraying only the “best” results.

A.1 Subsections in Appendices

It is possible, and advisable, to use (multiple levels of) subsections in an Appendix. The main Appendix sections may be optionally titled. They will automatically be assigned Appendix A, Appendix B, etc. If a title is present, it will appear on the next line in the appendix and on the same line in the TOC. Sub-sections will be listed in the TOC as A.1, A.2, etc.

Figures, such as Fig. 2, should only be referenced in the appendix. The thesis content can point to an appendix but not specifically reference a table or figure in the appendix.

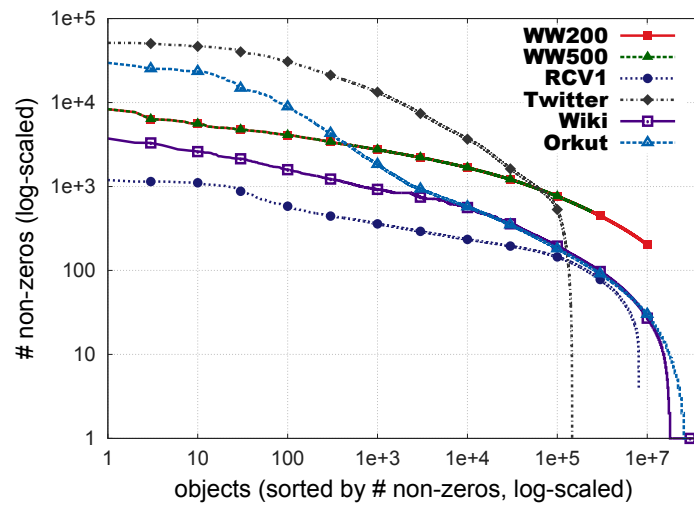


Fig. 2. Distribution of sample non-zeros.

Appendix B

TITLE OF ANOTHER APPENDIX

Different appendices should cover unrelated topics. If the appendices are on the same topic, they should just be sub-sections within the same appendix.

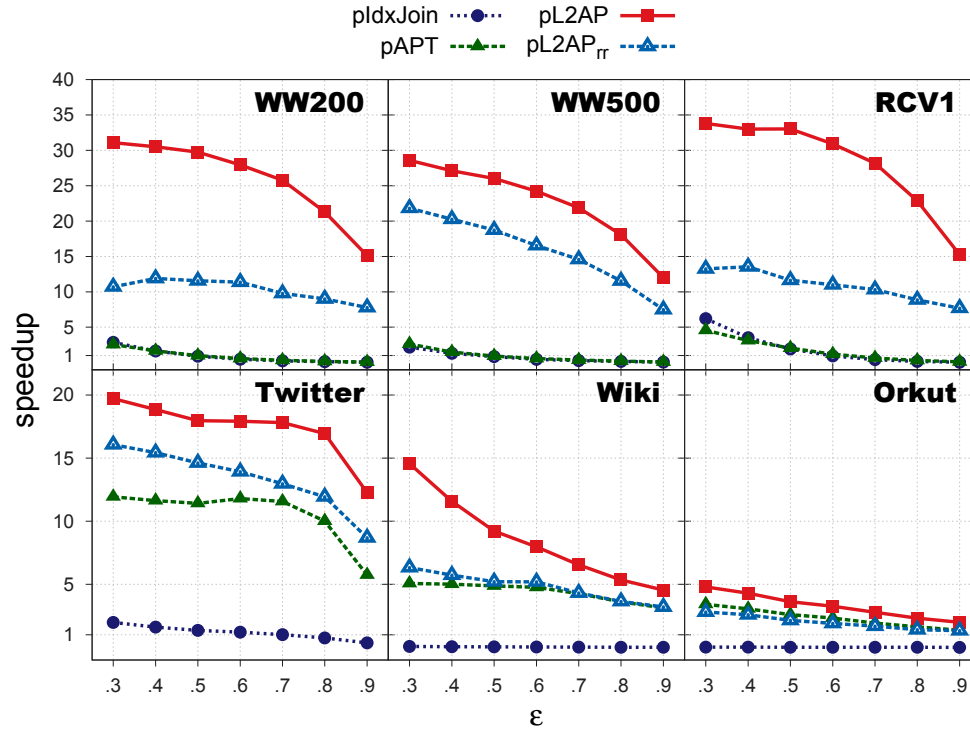


Fig. 3. Speedup comparison of our method against baselines.