Data Science Seminar 2018: Graph Embedding

Mourad Khayati and Dingqi Yang

https://diuf.unifr.ch/main/xi/ds-xi/index.html

September 25, 2018

UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

eXascale Infolab

Contents

Lecturers

- **Mourad Khayati**:
  - Senior researcher @ UNIFR since 06/2015
  - Ph.D. from University of Zürich
  - Research interests: Time Series, Matrix decomposition techniques and Recovery of missing values

- **Dingqi Yang**:
  - Senior researcher @ UNIFR since 04/2015
  - Ph.D. from Université Pierre et Marie Curie Paris VI
  - Research interests: Social media data mining and Pervasive big data analytics.

- **Rana Hussein**:
  - PhD student @ UNIFR since 05/2017
  - Research interests: distributed database systems and machine learning.

Goals of the Seminar

- Gather in-depth knowledge of an advanced topic/paper in data science: graph embeddings
- Focus on one research paper/topic.
- Learn how to critically read and study a research paper.
- Describe a paper in a report, present it in front of an audience and empirically reevaluate the proposed techniques.

Contents

Requirements and Organization

- Good understanding of algorithms, data structures and complexity.
- Write a report (first version and final version) of min 6 pages each describing the selected paper.
- Make a presentation of 20 min +10 min demo.
- The seminar participation will be limited to **10 people**.
- Presentations will be split into two seminar sessions by the lecturer. Participants **must** attend **both sessions**.

Contents

Final Grade

- Quality of the first and the last versions of the report.
- Quality of the presentation (including Q&A).
- Reproducibility experiments.
- Active seminar participation.

Contents

How to write the report

- The report[1] should reflect your understanding of the paper and not a rephrasing/summary of the paper.
- The structure of report might not follow the same structure of the paper.
- Explore extreme cases: the report might contain your own small running example and counterexample (that illustrate how the proposed solution works).
- Replot figures that describe the proposed technique, redefine complex formulas, add algorithmic description, etc.

[1] https://cs.stanford.edu/~rishig/courses/ref/paper-reading-technical.pdf

How to prepare your presentation

- The presentation[2] should explain the proposed contribution(s) and should not summarize the report/paper: what do I need to describe in order to explain the proposed solution?
- The presentation should follow a coherent structure and not necessarily the structure of the paper.
- An animated (toy) example helps to better explain the steps of the proposed solution.

[2] http://matt.might.net/articles/academic-presentation-tips

How to reproduce the experiments

- Rerun the efficiency and precision experiments **of the proposed technique(s)** introduced in the paper.
- Use a dataset from the paper or a similar one (similar properties and size) from other sources.
- A 10 min demo of the tool will be given after the presentation.

Contents

How to proceed/1

- Enroll to the seminar via `http://ilias.unibe.ch` (if not done yet).
- Select a paper among the set of proposed papers **by October 5th** via: https://doodle.com/poll/pxxrf3vnwh5et3ec .
- The papers will be distributed in FCFS basis.
- Enter your time preferences for the two sessions of the Data Science seminar (the poll will be sent on ILIAS).

How to proceed/2

- Prepare the first draft of the report and send it two weeks before your talk.
- Meet with the lecturer 1 week before your talk.
- Participate actively in the discussions of the other talks.
- Send your final report before the final deadline (hard deadline).
- All the deadlines will be posted on the seminar website: `https://diuf.unifr.ch/main/xi/ds-xi/index.html`. Stay tuned!