

Graph Databases Study Points Assignment

Objectives

To practice applying graph structure as a database model and implementing graph theory as a query engine, with references to the Curriculum requirements.

Tasks

1. Select a business or social practice domain, where the associated data is highly related (for example, education and training, sales, supply chain, human relations, financial transactions, media sources, medical treatment, software services, computer networks, social events, or similar).
2. Collect large volumes of relevant data from public sources in Internet.
3. Design a graph model of the data and a database schema of it. Consider arguments about separation of concerns, while making choices of entities, relations and attributes. Consider comparing alternative designs in regards with the performance optimisation.
4. Create a graph database and populate it with the data. Avoid duplications.
5. Retrieve information from the database by searching, seeking and mining operations formulated in Cypher. There must be at least one graph-algorithms-based operation designed, developed and explained by each individual student.
6. Create a simple client application in a programming language of your choice. The application should serve as platform for testing and validation of the queries.
7. Provide in writing, each student individually, answers to the following questions:
 - a. What are the advantages and disadvantages of using graph databases and which are the best and worse scenarios for it?
 - b. How would you code in SQL the Cypher statements you developed for your graph-algorithms-based query, if the same data was stored in a relational database?
 - c. How does the DBMS you work with organizes the data storage and the execution of the queries?
 - d. Which methods for scaling and clustering of databases you are familiar with so far?
8. Store the data, the queries, the application code, snapshots from the database, and the answers of the questions in a Github repository and upload the link to it in Peegrade.

Notes

1. This is group assignment. The recommended size of a group is 1-3 students.
2. While the design and implementation can be group work, p.p. 5 and 7 contain individual responses signed by the name of the creator.
3. The deadline for submission the solutions in Peergrade is 9:00 on 11/04/23. No peer grading is expected afterwards.
4. The solutions bring 20 study points to their authors.

Have fun and a happy Easter holiday!
the instructor