COMP9321 Semester 2, 2017

Assignment 3: Social Media - UNSWBook

Semantic Web + Performance Evaluation + Video

Group Assignment - 15 Marks (Same Groups as Assignment 2)

Aims

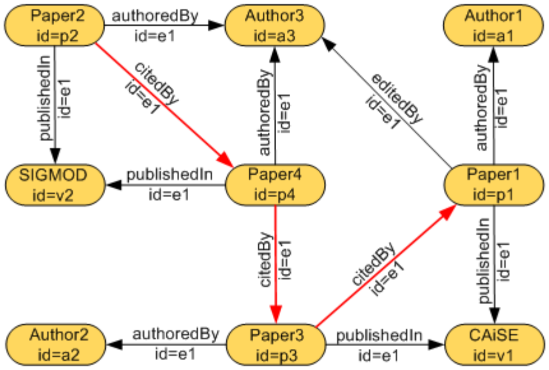
This project aims to give students hands-on experience in:

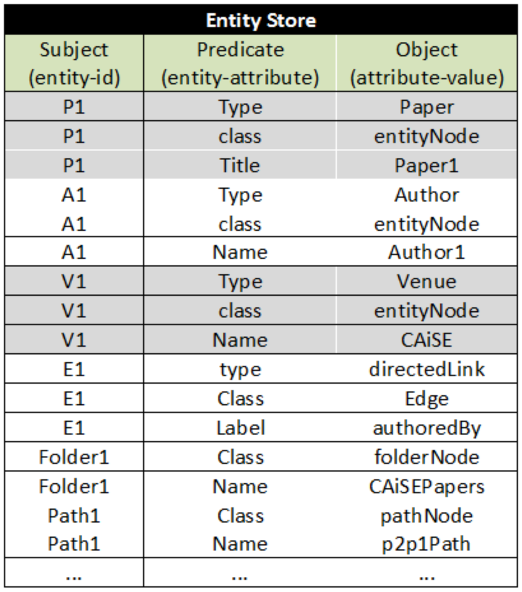
- Designing and implementing a fully-functional Web application.
- Using well-known application patterns such as MVC.
- Using JDBC and related technologies to maintain persistence of user data.

For this Assignment you will extend assignment 2. This is a **group assignment**: same group members as in assignment 2.

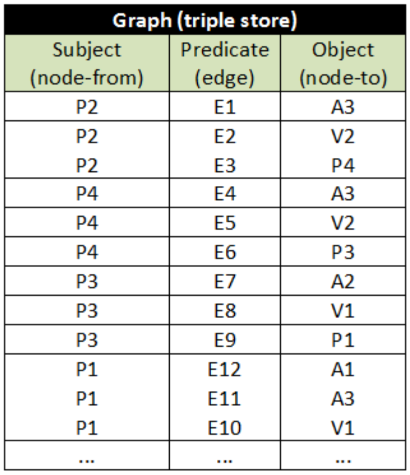
Part1: Graph Store and Visualization

1. Create two triple tables to store the entities (minimum: People and Messages) and the relationships (minimum: "person--(friendOf)-->person", "person--(posted)-->message", and "person--(liked)-->message") among them.
2. Support the following queries on the graph nodes (e.g. People) attributes (e.g. Name, Gender, DoB): (1) search for an people; (2) search for a messages; and (3) search for friends of friends;
3. Visualize the whole graph and the search results (as a graph). For the visualiztion, you can use existing APIs such as "https://d3js.org/" and "http://visjs.org/".





Entity Store		
Subject (entity-id)	Predicate (entity-attribute)	Object (attribute-value)
P1	Type	Paper
P1	class	entityNode
P1	Title	Paper1
A1	Type	Author
A1	class	entityNode
A1	Name	Author1
V1	Type	Venue
V1	class	entityNode
V1	Name	CAiSE
E1	type	directedLink
E1	Class	Edge
E1	Label	authoredBy
Folder1	Class	folderNode
Folder1	Name	CAiSEPapers
Path1	Class	pathNode
Path1	Name	p2p1Path
...	...	...



Graph (triple store)		
Subject (node-from)	Predicate (edge)	Object (node-to)
P2	E1	A3
P2	E2	V2
P2	E3	P4
P4	E4	A3
P4	E5	V2
P4	E6	P3
P3	E7	A2
P3	E8	V1
P3	E9	P1
P1	E12	A1
P1	E11	A3
P1	E10	V1
...	...	...

Part2: Data Curation

The curation APIs are available as an open source project on GitHub. https://github.com/unsw-cse-soc/Data-curation-API

Please Refere to Assignment 1.

For this part of the assignment, you will need to extract keywords and named entities (similar to assignment 1) from the messages. If the message contains any keywords related to Bullying (e.g. see HERE) then the system admin should recieve a notification (by email and also in the user activity report - see Assignment 2).

Part3: Performance Evaluation

Using JMeter and PSI Probe, conduct a performance analysis test on your application (refer to lab notess).

We'll consider the Web application developed for Assignment 3. For your simulation, you should consider the following scenarios in the order listed below:

1. A user logs into the system.
2. A user search for friends.
3. A user add a friend.
4. A user post a message.
5. Etc...

There should be a delay of 5 seconds between each user action.

1. The Client-Side round trip time/response time for executing the operation: Use JMeter to simulate user interaction and experiment with 5,10,20,25,40,50,75,80,100,200 users and so on with one minute as ramp-up period. If your webapp can handle 200 users in 1 minute, increase the number of users in steps of 200 until the median response time suffers and increases to a large extent (up to 3 times the normal response time).

   (**Note:** If you don't see the response times increasing, then your container may be dropping requests. Look at the error rate in Jmeter. Stop the test if the error rate is > 2%)

2. For each set of users, note the median response time and the 90-percentile response time in JMeter. Draw a graph using with these response times on y-axis and number of users on x-axis on a log-scale.

   (**NB1: Important** Restart the application between each test/set of users)

   (**NB2: Important** Always load the application in a browser before starting a test. The first request takes the longest time and it biases your results.)

3. Plot the client-side throughput data against the number of users
4. Use Little's Law to compute the number of requests that are in your web application for each test
5. On the server side,
   1. Note down the average response time for each set from PSI probe.
   2. Compute the throughput (requests/sec) for each set of users. Draw a graph of throughput vs number of users.
   3. Find the utilization of your Tomcat container and your system for each test
6. Compute the average number of visits to the database per request (total number of database calls/number of requests). Then, use Forced Flow Law to compute the throughput of your database.

You can use the lab machines or your own laptops to do the performance analysis. Most importantly, **you must run JMeter and your application on separate machines** .

Written Report Format

Each group should create a report with the following sections:

1. Introduction
2. Application Details (1 **marks**)
    - The anatomy of the booking operation - the parts of your application that are involved, how many database calls. Use a diagram to illustrate this.
    - Search parameters used and implications of these (how many results will be returned, calls to DB, etc.).
3. Test Setup (1.5 **marks**)
    - Server details : Hardware (CPU, RAM, HDD), Operating environment, Processes running on the machine, Database used and the number of connections in the connection pool.
    - Client details : Hardware, Details of test plan, Are server and client on same network?
    - How was the experiment constructed?
4. Test Results and Discussion (1.5 **marks**)
    - Graphs...
    - Analysis of how your application has performed, what is the maximum number of users you can handle, how can you improve performance?
    - Which step in the operation takes the most amount of time?
    - If there's a slow down why? If performance is good, why?
5. Conclusion

The report should be in Arial 11 pt. with headings in Arial 12 pt. bold with 1 inch margins on all sides. A template is given here. Latex users can use Helvetica or Arial package.**Please stick to this format** . It is expected that the report would have upto 10 pages, though it will be judged on content and not on length. The report **must be submitted as a PDF file**.

Part4: Solution Video for Assignment 2 or 3

The third part of the assignment aim to give students the confident to present their technical as well as teaching capabilities. Each group will release the solution to the second/third assignment by preparing a video with a full walk-through. The video should be uploaded on the youtube. You should use the lecture's first slide and add you group member's name there. This slide will be shown in the begining of the video. The best video will be added to the course material, under the assignment section.

NOTICE: The first slide on the video should have the similar format to THIS VIDEO.

Submissions

The due date for this assignment is (end of Week 12), **Sunday, Oct 22 2017, 23:59:59.**

DEMO: Demo will be held during the lab times in week 13. You should use the same .war file that you submitted for your demo. Each demo will be around 10-15 mins long.

Submission Instructions

1. Name your report as groupXX.pdf where XX is your group id. **The last page of the report should illustrates the youtube link** (for Part4 of the assignment).
2. email the war file (part 1 and 2) + PDF file (part 3) + the video youtube link (part 4) to the following email address:

```
3. To: unsw.cse.comp9321@gmail.com
4. CC: All your group members
5. From: [Your unsw email address]
6. Subject: COMP9321-Ass3-S2_17
7. Body:
8.   [group id]
9.   For-each Group member: [Student Number]-[Student Name]
```

Marking

This assignment is worth **15 marks** and will be marked as shown above (**part1: 5 marks, part2: 3 marks, part3: 4 marks, and part 4: 3 marks**). There will be no other marking guide supplied for this assignment.

Here is an example of an excellent report that was produced from one of the previous semester. (Obviously, their test-cases and requirements were different from what is asked here, but they provided an in-depth analysis of their application.)

Late Submissions and Penalties

Late submissions will lose **1 mark for every day** the assignment is delayed.

Important

Please use the message board for resolving doubts.