School of Information Technology
Department of Computer Science

COS326 Database Systems
Practical 9 2024

Release Date: 14 October 2024
Submission Date: 20 October 2024 @ 23:59Hrs
Lecturer: Mr S.M Makura

Total: 50 Marks

Objectives

1. Get exposure to the Neo4j graph DBMS.
2. Learn how to create and use a graph for a Neo4j database.
3. Appreciate the differences between SQL and NoSQL databases.

You are expected to have completed the Movie tutorial on Neo4j before you start on this practical exercise. When you are done, you may start working on this practical and do the following:

1. You must submit the following files:
   a. ***Task1Queries.txt*** with all the CREATE and MATCH queries for Task 1
   b. ***Task2Queries.txt*** with all the MATCH queries for Task 2
2. Compress the above documents into an archive and upload it to ClickUP **before** the due date/time. The file name for the archive must have your student number as part of the file name, e.g. **uxxxxxxxx-prac9.zip** or **uxxxxxxxx-prac9.tar.gz** where uxxxxxxxx is your student number.
3. Book for a demo session via Discord to demo the practical.

NO LATE submissions will be accepted after the submission date and time has lapsed. Do not wait till the last minute to submit and start giving excuses that you faced technical challenges when you tried to submit.

Scenario: Cybersecurity Threat Intelligence Graph

In this practical, you will **model a cybersecurity threat intelligence system**. The system tracks **threat actors**, **cyber incidents**, and **mitigation strategies**. **Threat actors** can launch multiple incidents, and **incidents** can be mitigated by specific **strategies**.

Entity Descriptions:

Entity	Properties	Relationships	With Entity	Relationship Property
ThreatActor	`name`: APT29, `origin`: Russia, `motivation`: Espionage	LAUNCHED	Incident, Phishing Campaign	`date`: 2023-09-01
ThreatActor	`name`: Lazarus Group, `origin`: North Korea, `motivation`: Financial Gain	LAUNCHED	Incident, Ransomware Attack	`date`: 2023-07-15
Incident	`type`: Phishing Campaign, `severity`: High	MITIGATED_BY	Strategy, User Awareness Training	`effectiveness`: 85%
Incident	`type`: Ransomware Attack, `severity`: Critical	MITIGATED_BY	Strategy, Network Segmentation	`effectiveness`: 90%
Strategy	`name`: User Awareness Training, `type`: Preventive			
Strategy	`name`: Network Segmentation, `type`: Containment			

Task 1: Create and Query a Graph Database

[25 Marks]

1. **Create a Neo4j Database:** Name it **ThreatIntel.graphdb**.
2. **Use Cypher to Create the Graph** using the provided data:

Required Queries:

a) **Write Cypher statements** to create the graph:

i. Create **ThreatActor**, **Incident**, and **Strategy** nodes and **LAUNCHED** relationships. (8 Marks)

ii. Show the current **nodes and relationships** in the database. (2 Marks)

iii. Create **MITIGATED_BY** relationships between incidents and strategies. (6 Marks)

iv. Show the updated contents of the graph. (2 Marks)

b) **Write Cypher queries** to answer the following:

i. List all **unique node labels**. (1 Mark)

ii. List all **threat actors**, sorted by **name**. (1 Mark)

iii. List all **incidents** by **severity**, in descending order. (1 Mark)

iv. List the **relationship types** present in the graph. (1 Mark)

v. List all **threat actors** and the **incidents they launched**. (1 Mark)

vi. List all **incidents** and their corresponding **mitigation strategies**. (1 Mark)

vii. Find the **most effective strategy** for mitigating incidents. (1 Mark)

Task 2: Aggregation and Path Queries

[25 Marks]

1. **Path and Other Queries:**

a) Find **incidents** that are **1 or 2 links** away from **APT29**. (3 Marks)

b) Show the **nodes in the shortest path** from the threat actor 'APT29' to the strategy 'User

Awareness Training'. (3 Marks)

c) Report whether each **incident** has an associated **mitigation strategy**. (4 Marks)

d) For all paths of **length 2**, list the **node names** and **path length** (3 Marks)

2. **Aggregation Queries:**

a) **Count the number of nodes** in the graph. (3 Marks)

b) **Count the number of incidents** launched by each threat actor. (3 Marks)

c) **Count the number of mitigation strategies** applied to incidents. (3 Marks)

d) **Identify the threat actor** responsible for the most incidents. (3 Marks)