

Ceng790 Big Data Analytics

Assignment 4

Introduction to Neo4j

Egemen Berk Galatali
e2099018@ceng.metu.edu.tr

Part I

1. Does imposing this constraint improve the performance in a way other than integrity?

Constraints prevents user from entering wrong kind of data, other than that I could not find any information related to performance improvement

2. How many characters appear in the graph?

107

We can also visualize the graph executing the cypher query below. The visualization is in the Figure 1

```
MATCH grap=(c1:Character)-[weight:INTERACTS]->(c2:Character)
RETURN grap
```

3. Get summary statistics for the minimum, maximum and average number of characters each character has interacted with.

The results are exported as csv files and shared in submission folder.

```
// Maximum and Min statistics
MATCH (c1:Character)-[weight:INTERACTS]->(c2:Character) RETURN c1, count(weight)
ORDER BY(count(weight)) DESC
// Average
MATCH (c1:Character)-[weight:INTERACTS]->(c2:Character) RETURN c1, avg(weight.weight)
```

The code above considers the graph as directed, therefore only out going interactions was counted for each node.

Tyrion has the maximum number of character interaction with 24 many.

4. Find the shortest path from Arya to Ramsay. Consider the possibility of having few paths of same length. List all of them in the result. You can use the functions shortestPath and allShortestPaths provided in Cypher

```
MATCH (source:Character {name: "Arya"}), (dest:Character {name:"Ramsay"})
CALL algo.shortestPath.stream(source, dest, "weight")
YIELD nodeId, cost
RETURN algo.getNodeById(nodeId).name AS name, cost
```

Result is in the Figure 2

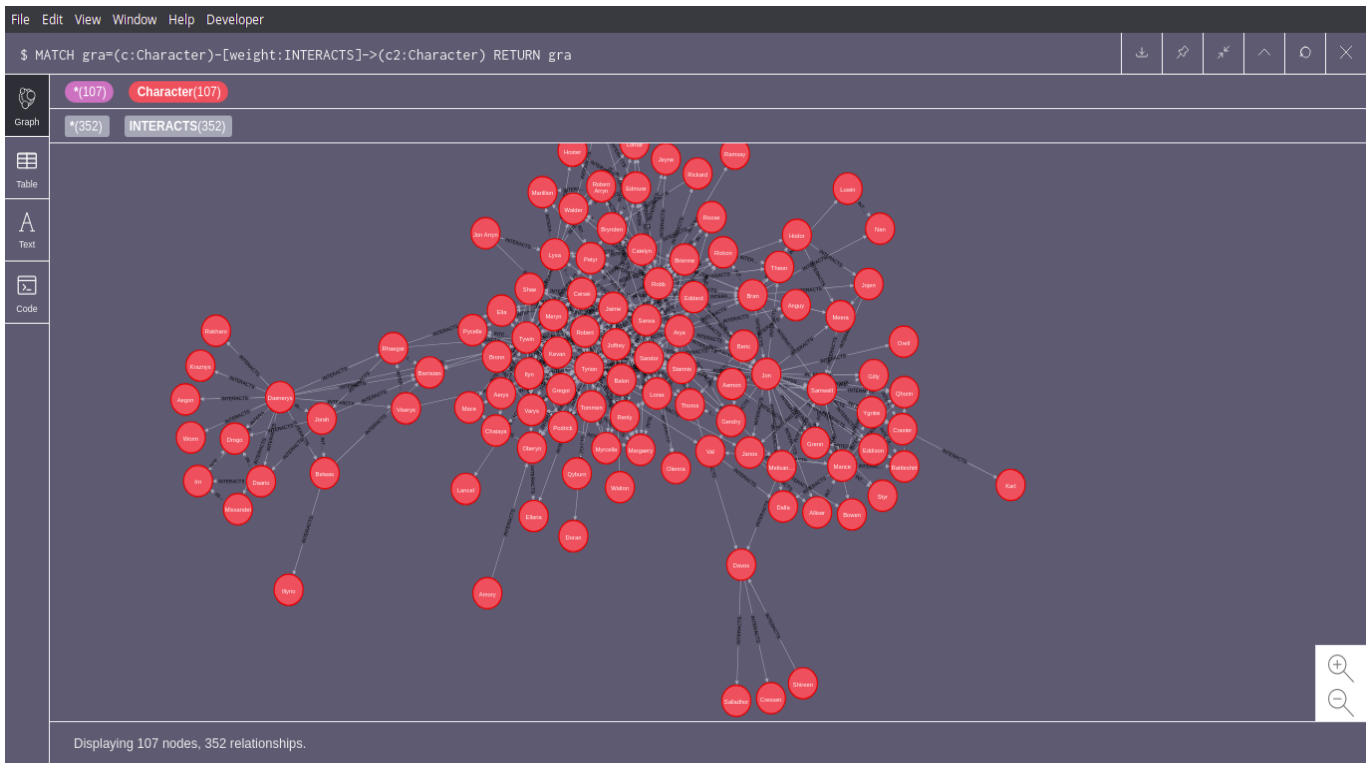


Figure 1: Visualization of the Game Of Thrones Characters

\$ MATCH (source:Character {name: "Arya"}), (dest:Character {name:"Ramsay"}) CALL a1...

	name	cost
Table	"Arya"	0.0
Text	"Roose"	5.0
Code	"Ramsay"	6.0

Figure 2: Shortest Path from Arya to Ramsay

5. What is the longest shortest-path distance between any two characters?

The longest shortest-path is from Salladhor to Illyrio with total weight of 59

```
// All Pair Shortest Path
// Longest Shortest Path between two characters
CALL algo.allShortestPaths.stream('weight', {nodeQuery:'Character', defaultValue:null})
YIELD sourceNodeId, targetNodeId, distance
WITH sourceNodeId, targetNodeId, distance
WHERE algo.isFinite(distance) = true

MATCH (source:Character) WHERE id(source) = sourceNodeId
MATCH (target:Character) WHERE id(target) = targetNodeId
WITH source, target, distance WHERE source <> target

RETURN source.name AS source, target.name AS target, distance
ORDER BY distance DESC
LIMIT 1
```

6. How many characters are there with interaction distance 4 to Cersei Lannister?

There are 5 characters with such conditon. The result is in Figure 3.

```
MATCH (source:Character {name:"Cersei"})
CALL algo.shortestPaths.stream(source, "weight", {nodeQuery:'Character'})
YIELD nodeId, distance
WITH source, nodeId, distance

MATCH (target:Character) WHERE id(target) = nodeId
WITH source, target, distance WHERE source <> target AND distance = 4

RETURN source.name AS source, target.name AS target, distance
ORDER BY distance DESC
```

\$ MATCH (source:Character {name:"Cersei"}) CALL algo.shortestPaths.stream(source, "...			
	source	target	distance
Table	"Cersei"	"Varys"	4.0
	"Cersei"	"Elia"	4.0
	"Cersei"	"Pycelle"	4.0
	"Cersei"	"Bronn"	4.0
	"Cersei"	"Shae"	4.0

Figure 3: Characters with distance of 4 to Cersei

7. Who are Jon Snow's parents?

```
// Jon Snow's parents
MATCH (john:Character )-[relation:RELATIONSHIP]->(relative:Character {name:"Jon"})
RETURN john.name, relation, relative.name
```

Only parent of him is Eddard Stark as father.

8. You can notice that sibling relationships are missing. Create them by finding common parents for the ties we have already in the graph.

We first find characters that have same father or mother then create "SIBLINGS" relationship between them.

```
// Create Siblings
MATCH (ch1:Character)-[relation:RELATIONSHIP]->(ch2:Character)
MATCH (ch3:Character)-[rel:RELATIONSHIP]->(ch4:Character)
WHERE ch1.name = ch3.name AND ch2.name <> ch4.name
CREATE (ch2)-[sib:SIBLINGS {parent: ch1.name}]->(ch4)
RETURN ch2.name, ch4.name, relation.tie, ch1.name
```

9. Are there any children of incestuous relationships?

We first find all the siblings, then we search for the situation in which the siblings are father and mother of a same child.

```
MATCH (father)-[s:SIBLINGS]->(mother),
      (mother)-[r1:RELATIONSHIP {tie:"mother"}]->(child),
      (father)-[r2:RELATIONSHIP]->(child)
RETURN DISTINCT father.name, mother.name
```