1. Stiahnite a importujte databázu

Vytvoril som databázový server s názvom: tweets, heslom: 123456789 a verziou 4.4.5. Potom som klikol na tri bodky, otvoril terminál a zadal som príkaz cd bin. Následne som použil tento command

neo4j-admin load --from="C:\Users\pplev\Downloads\tweets2022-4\_4.dump" --database=tweets --force –verbose

Skontroloval som či sa do priečinku spomenutého nižšie pridala databáza tweets ak áno pokračujem ďalej ak nie skúsim import znova dokiaľ tam tá databáza nebude.

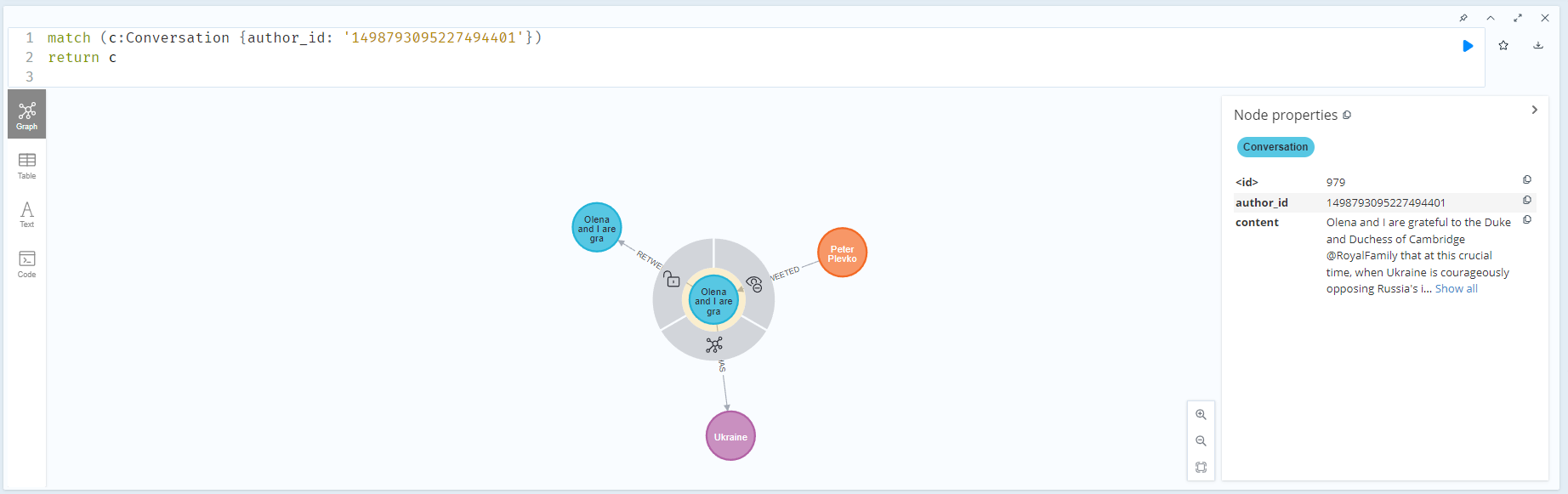
C:\Users\pplev\.Neo4jDesktop\relate-data\dbmss\dbms-a5e0a6ab-5d0f-4977-a9f5-f1e187356303\data\databases

Ďalší postup, pustil som start server v Neo4j GUI. Klikol som na create database vytvoril som db s názvom tweets opäť celé použitím GUI. Následne som server reštartoval a vypol celé Neo4j. Teraz ho znova zapnem a databáza bola importnutá

1. Vytvorte volaním iba jednej query nového Autora s Vašim menom (vyplňte minimálne name a username) a tým istým volaním vytvorte tweet ktorý bude obsahovať aspoň jeden vami vybraný Hashtag (vzťah :HAS) a ktorý bude retweetovať najretweetovanejší tweet Vladimíra Zelenského{username:"ZelenskyyUa"}.

Obrázok, na ktorom je text

Automaticky generovaný popis



Použitá query:

MATCH (a: Author {username: "ZelenskyyUa"})-[:TWEETED]->(c: Conversation)

MATCH (rt: Conversation)-[r:REPLIED\_TO]->(c: Conversation)

WITH c, size(collect((c)<-[:REPLIED\_TO]-(rt))) AS relations

ORDER BY relations DESC LIMIT 1

match (h:Hashtag {tag: 'Ukraine'})

CREATE (me: Author {name: 'Peter Plevko', username: 'Kuntox', id:'1498793095227494401'})

CREATE(myConversation:Conversation {content: c.content, author\_id: me.id})

CREATE (myConversation)-[rtc:RETWEETED]->(c)

CREATE (myConversation)-[:HAS]->(h)

CREATE (me)-[t:TWEETED]->(myConversation)

Najprv som si selectol všetky tweety z conversations ktorých autorom je username s menom ‘ZelenskyyUa’. Následne som pre tieto nájdene konverzácie spočítal ich vzťah REPLIED\_TO. Dal som order by podľa tohto vzťahu a zoradil od najväčšieho po najmenší a vybral ten najväčší. Pomocou with som si tieto údaje preniesol ďalej. Matchol som hastag podľa tagu ‘Ukraine’. Vytvoril som autora s mojim menom a usernejmom a ídéčkom. Vytvoril som konverzáciu(tweet) s rovnakým obsahom ako je ta Zelenského. Nastavil som vzťah, že moja konverzácia retweetuje Zelenského tweet. Vytvoril som vzťah pre moju konverzáciu, že má hastag Ukraine. Nastavil som, že môj autor tweetoval danú konverzáciu (tweet).

Pomocou tejto query som skontroloval či bol mnou tweetovany tweet naozaj vytvorený:

match (c:Conversation {author\_id: '1498793095227494401'})

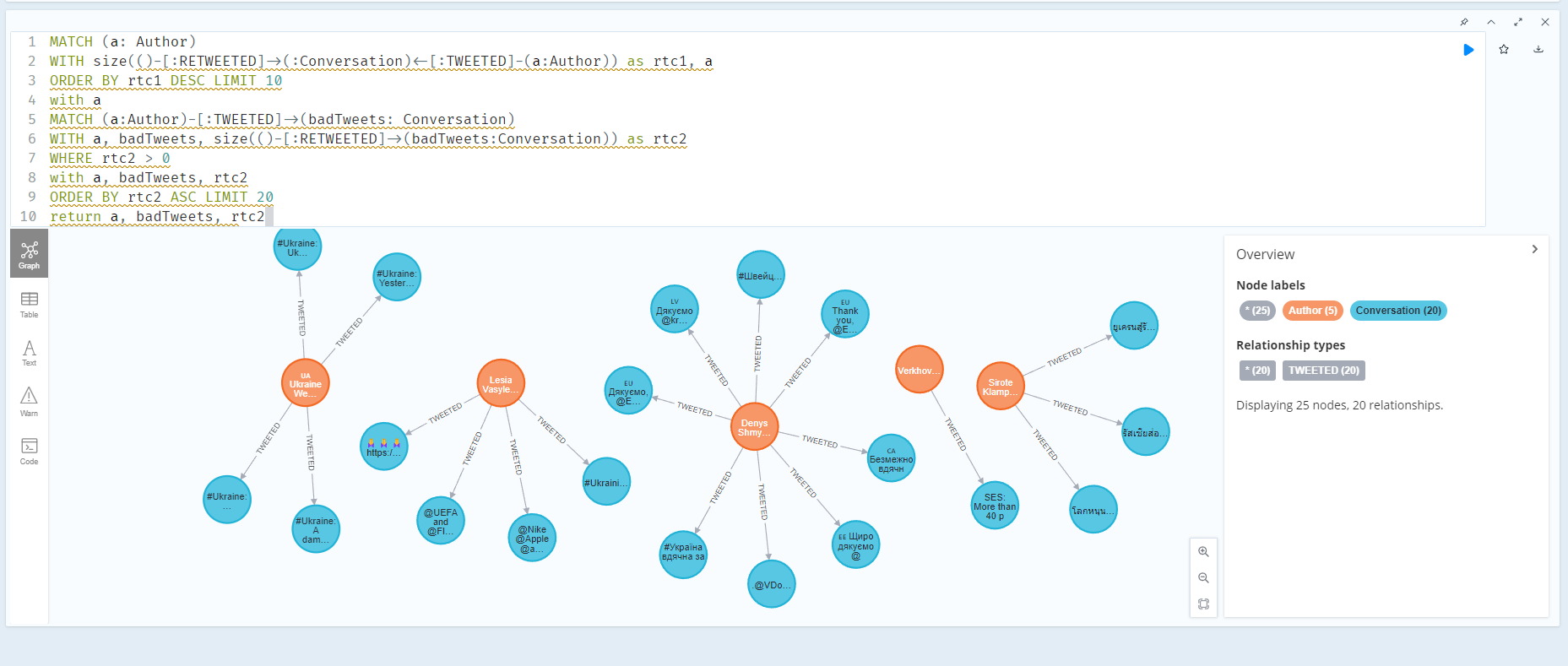
return c

Keďže query som robil na viac pokusov pomocou tohto som neúspešne pokusy mazal:

match (c:Conversation {author\_id: '1498793095227494401'})

detach delete c

1. Vyhľadajte zlyhania influencerov. Vyhľadajte 20 najmenej retweetovanych tweetov od Accountov, ktoré sú na prvých 10 miestach v celkovom počte retweetov. Aj keď taký príklad v datasete nie je, počítajte aj s prípadom, ak by niektorý tweet bol retweetnutý 0 krát.



Query:

MATCH (a: Author)

WITH size(()-[:RETWEETED]->(:Conversation)<-[:TWEETED]-(a:Author)) as rtc1, a

ORDER BY rtc1 DESC LIMIT 10

with a

MATCH (a:Author)-[:TWEETED]->(badTweets: Conversation)

WITH a, badTweets, size(()-[:RETWEETED]->(badTweets:Conversation)) as rtc2

WHERE rtc2 > 0

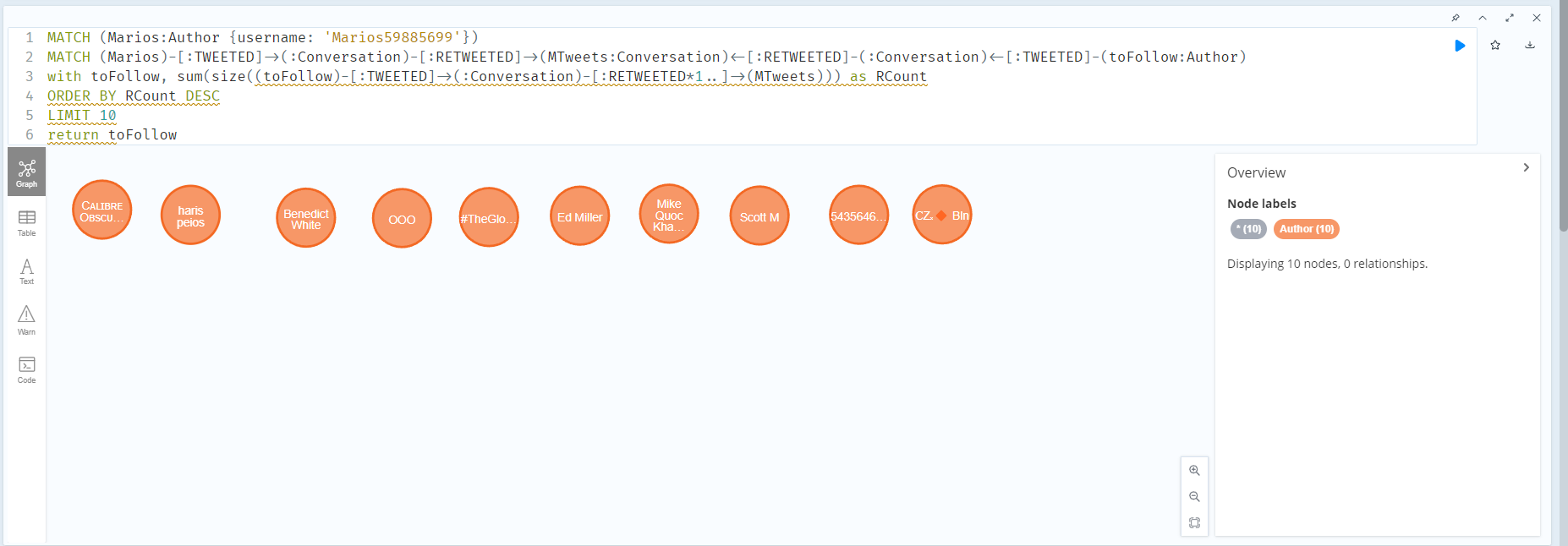
with a, badTweets, rtc2

ORDER BY rtc2 ASC LIMIT 20

return a, badTweets, rtc2

Najprv som selectol 10 užívateľov ktorý majú najviac retweetovane konverzácie(tweety), takých ktorý majú retweeted iný ako null. Potom som pozrel všetky konverzácie(tweety) týchto používateľov a našiel tweety ktoré majú najmenej retweetov. Podľa tohto počtu som ich zoradil vzostupne to znamená od najmenšieho po najväčšieho a vybral prvých 20.

1. Odporučte používateľovi (username: Marios59885699) followovanie ďalších authorov na základe zhody v retweetovaní rovnakých tweetov: Vyhľadajte 10 autorov, ktorí retweetli najviac tweetov rovnakych, ako používateľ Marios59885699. Počítajú sa aj retweety tweetov, ktoré retweetujú rovnaký tweet.



Query:

MATCH (Marios:Author {username: 'Marios59885699'})

MATCH (Marios)-[:TWEETED]->(:Conversation)-[:RETWEETED]->(MTweets:Conversation)<-[:RETWEETED]-(:Conversation)<-[:TWEETED]-(toFollow:Author)

with toFollow, sum(size((toFollow)-[:TWEETED]->(:Conversation)-[:RETWEETED\*1..]->(MTweets))) as RCount

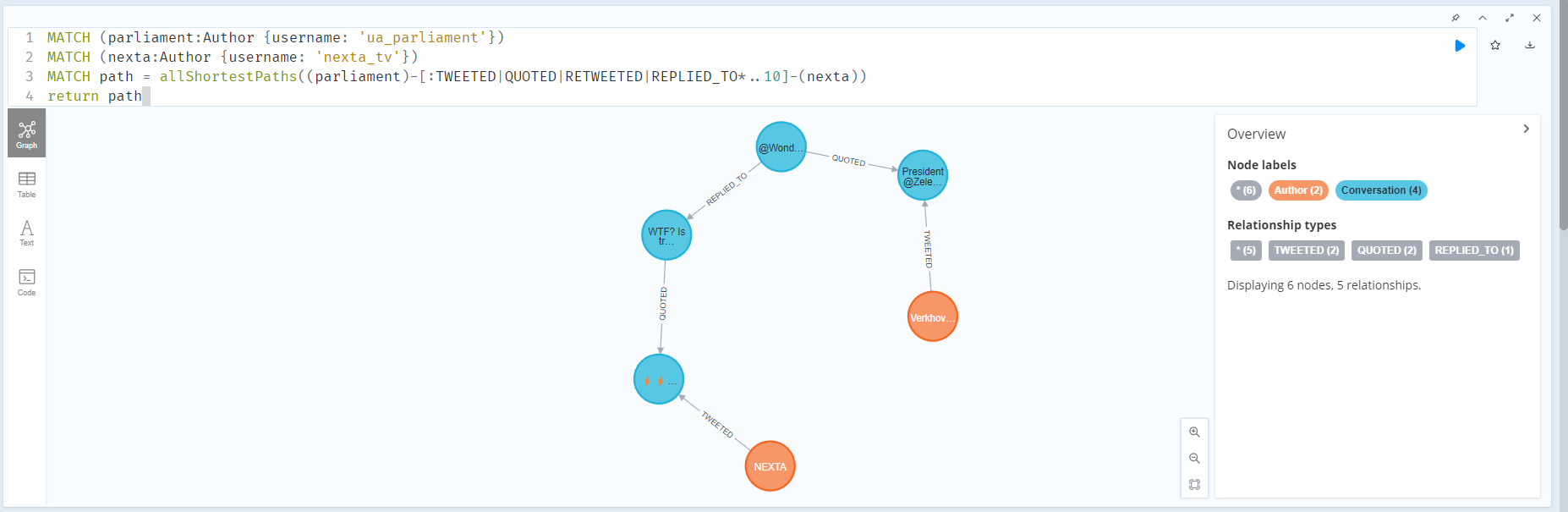
ORDER BY RCount DESC

LIMIT 10

return toFollow

Najprv som si našiel authora s username: ‘Marios59885699’. Následne som našiel všetky tweety ktoré retweetuju rovnaké tweety ako marios a naviazal nane Authora. Potom som cez size a sum zistil ktoré acounty majú najviac rovnakých retweetov ako marios. Zoradil som od najviac rovnakých až po najmenej rovnakých a limitoval na prvých 10.

1. Nájdite najkratšie cesty medzi Ukrajinským parlamentom (username: “ua\_parliament”) a NextaTV (username: “nexta\_tv”) cez vzťahy TWEETED, RETWEETED, REPLIED\_TO a QUOTED. Hľadajte do hĺbky maximálne 10. (hint: allShortestPaths)



Query:

MATCH (parliament:Author {username: 'ua\_parliament'})

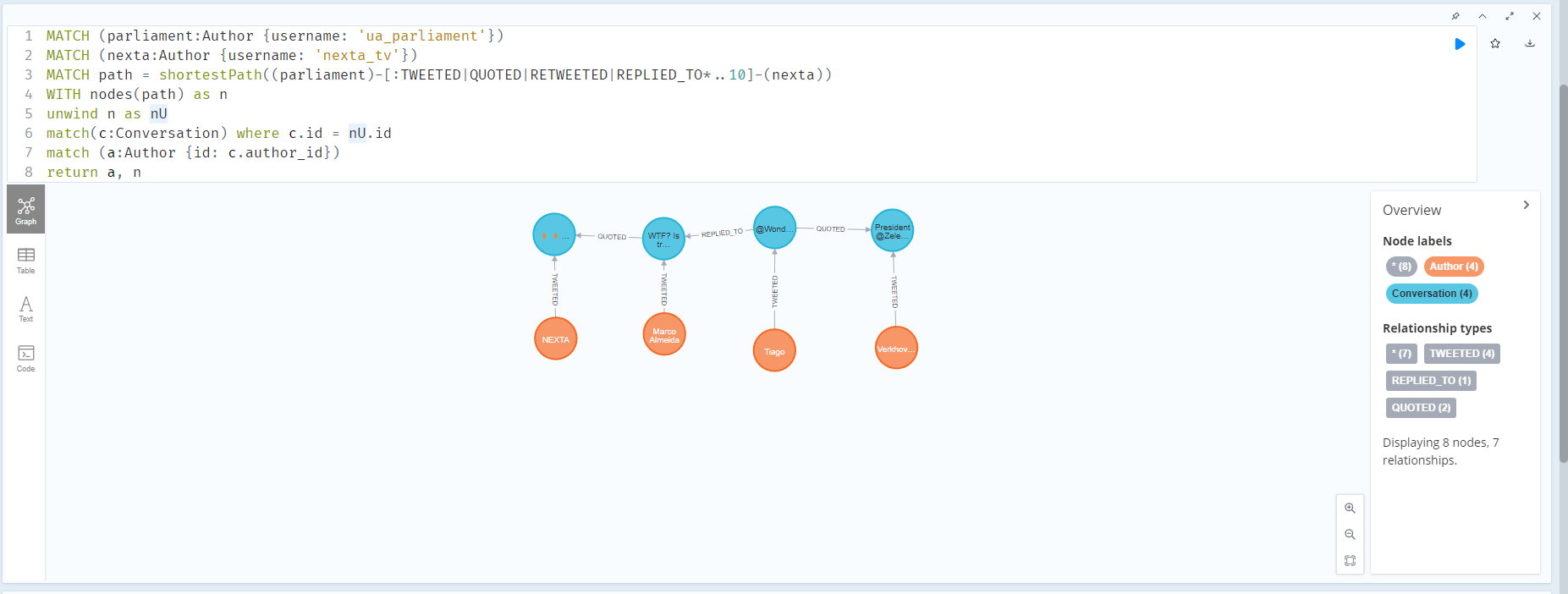
MATCH (nexta:Author {username: 'nexta\_tv'})

MATCH path = allShortestPaths((parliament)-[:TWEETED|QUOTED|RETWEETED|REPLIED\_TO\*..10]-(nexta))

return path

Matchol som authora s usernamom ‘ua\_parliament’ a autora s usernamom 'nexta\_tv'. Následne som použil all shortest paths na nájdenie všetkých ciest medzi nimi kde som do špecifikoval aby sa použili iba vzťahy TWEETED QUOTED RETWEETED REPLIED\_TO a išlo sa max do hĺbky 10. Následne som všetky tieto cesty vrátil avšak našlo iba jednu cestu.

1. Upravte query z predchádzajúcej úlohy, aby vrátila iba nájdenú najkratšiu cestu a jednotlivé uzly typu Conversation v nej aj spolu z autorom, ktorý ich tweetol. (hint: UNWIND)



Query:

MATCH (parliament:Author {username: 'ua\_parliament'})

MATCH (nexta:Author {username: 'nexta\_tv'})

MATCH path = shortestPath((parliament)-[:TWEETED|QUOTED|RETWEETED|REPLIED\_TO\*..10]-(nexta))

WITH nodes(path) as n

unwind n as nU

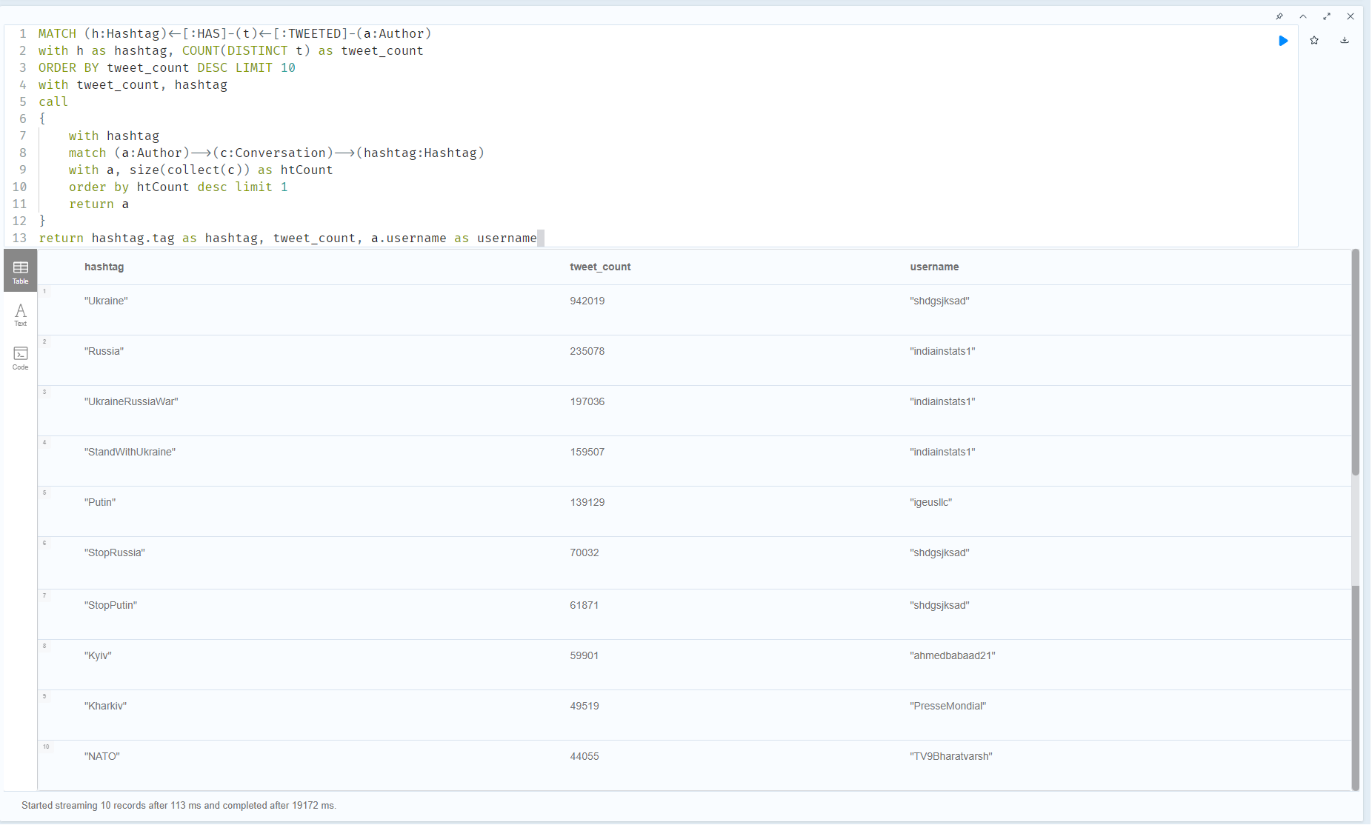
match(c:Conversation) where c.id = nU.id

match (a:Author {id: c.author\_id})

return a, n

Na začiatok som si matchol authora s usernamom ‘ua\_parliament’ a potom som matchol authora s username ‘nexta\_tv’. Našiel som najkratšiu cestu medzi týmito dvoma userami. V shortest path som nastavil aby používalo iba vzťahy TWEETED QUOTED RETWEETED REPLIED\_TO a išlo sa len do hĺbky 10. Shorest path mi vrátilo list z najkratšou cestou. Tento list som pomocou príkazu unwind rozbil na uzly. Následne som matchol konverzácie podľa id. Z týchto konverzácii som matchol ich autora a vrátil si týchto nájdených authorov a cestu v uzloch.

1. Vypíšte 10 najpoužívanejších hashtagov v datasete aj s celkovým počtom použití a autorom, ktorý daný hashtag najviac krát použil. (hint: collect)



Query:

MATCH (h:Hashtag)<-[:HAS]-(t)<-[:TWEETED]-(a:Author)

with h as hashtag, COUNT(DISTINCT t) as tweet\_count

ORDER BY tweet\_count DESC LIMIT 10

with tweet\_count, hashtag

call

{

with hashtag

match (a:Author)-->(c:Conversation)-->(hashtag:Hashtag)

with a, size(collect(c)) as htCount

order by htCount desc limit 1

return a

}

return hashtag.tag as hashtag, tweet\_count, a.username as username

Na začiatok som si našiel všetky hastagy ktoré existujú. Potom som spočítal koľko krát sa nachádza hastag v tweete zoradil ich zostupne a zobral 10 najviac vyskytujúcich sa. Pomocou funkcie call som pre každý tento hashtag zistil autora ktorý ho použil. Následne som týchto authorov zoradil podľa toho koľko krát daný hashtag použili a vrátil autora ktorý ho použil najväčší počet krát. Na konci som len vrátil hodnoty potrebne pre výpis