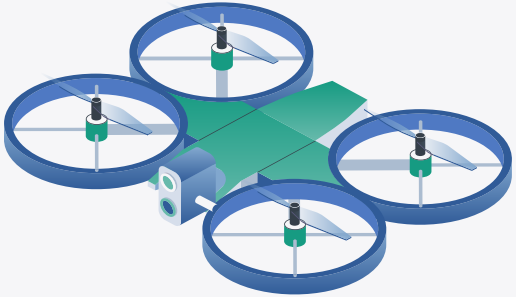
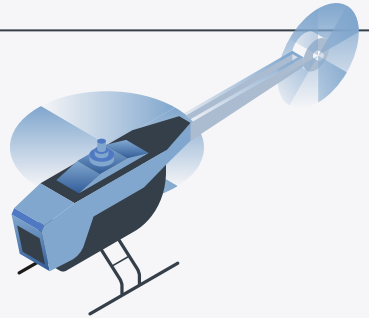


DATASCI 205: Project 3

Daniel Bostwick, Michelle Cheung, Ethan Wang, and Nikita Chauhan

TABLE OF CONTENTS



MongoDB

Flexible Database System that stores JSON-like data structures

Redis

Efficient database system for real-time notifications and increased data-retrieval performance

NEO4j

A graph database management system data elements Neo4j stores are nodes, edges connecting them, and attributes of nodes and edges.

BUSINESS CASE #1

MongoDB

Redis

In order to both conserve time and resources while providing the most efficient method of delivery to customers, we advise a business case in which we use these two technologies to help achieve optimized coordination between BART trains and delivery drones.

MongoDB - Why Is It Better

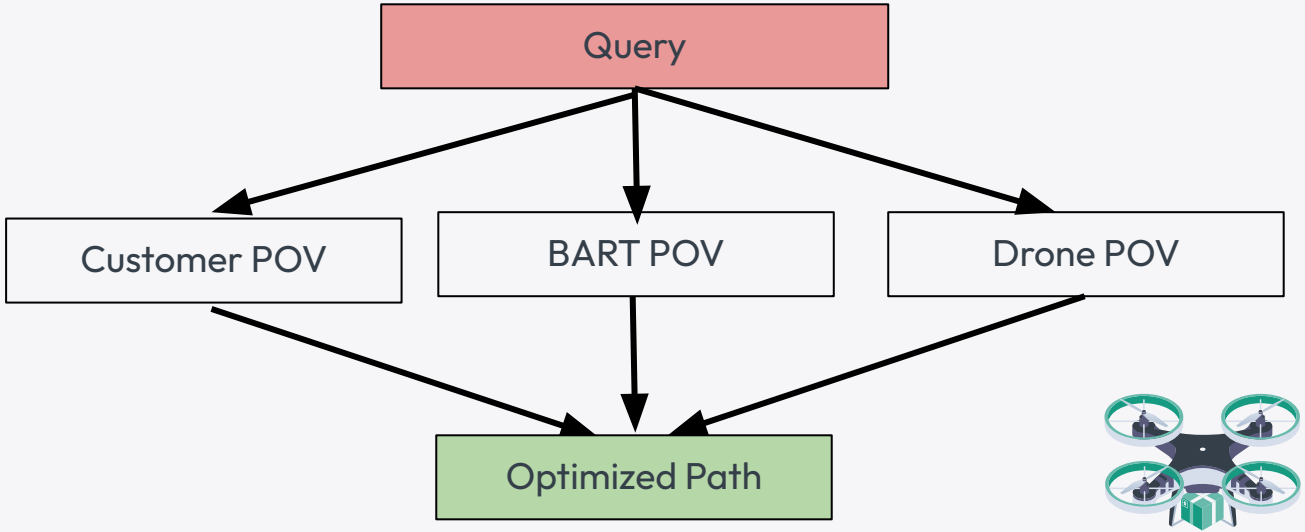
Features

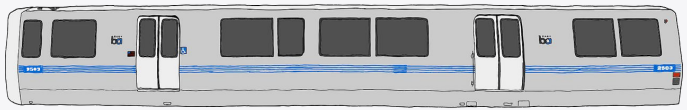
- POV functionality
- Geo-spatial query capabilities
- Real-time Data Updates

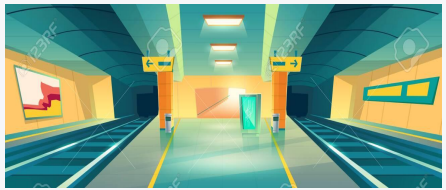
Applications

- Integration of multiple POVs to streamline efficient queries
- Consistent live updates of geo-spatial data can provide efficient algorithms to coordinate BART delivery and drone pick up

MongoDB - Process







MongoDB – Benefits

- Saved cost on human resources for delivery drivers
- Some cases, delivery time can be shortened by avoiding local traffic during rush hours
- Environmentally friendly, reduce pollution

Redis - Uses

Improve Delivery Process

Ability to Notify in Real-Time

Cache Frequently Accessed Data

Increase Data-Retrieval Process

Redis - Inputs

Improve Delivery Process

Order Status
Drone Movements and Location
Delivery Confirmations

Cache Frequently Accessed Data

Restaurant and Pickup Locations
Product Inventory

Redis - Design

Improve Delivery Process

Publisher and Subscriber

Cache Frequently Accessed Data

Customer and Notification Channel

Redis - Why Is It Better?

Improve Delivery Process

Messaging

Cache Frequently Accessed Data

Key-value data storage model

Business Case #2 - Neo4j

Scenario

AGM is experiencing delays in its delivery service due to unpredictable traffic conditions and the static nature of delivery scheduling, which does not account for the variability of urban transit and congestion.

Algorithm Applications

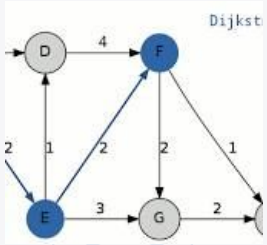
Explore the application and business impact of the following algorithms:

- Yen's Algorithm
- Betweenness Centrality
- Louvain Algorithm

Yen's Algorithm

Application

Identifying multiple shortest paths for alternate delivery routes when there is a disruption.

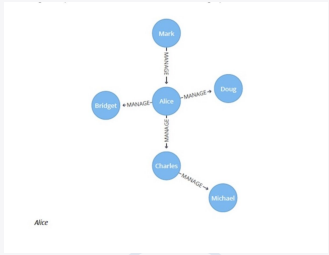


Business Impact

Ensures robustness of AGM's delivery leading to maintaining reliable service and customer satisfaction.

```
---------------------------------
Path 1:
From:  depart SFO
To:  arrive OAK
Total Cost:  3882
Minutes:  64.7
Path:  depart SFO -> yellow SFO -> yellow San Bruno -> yellow South San Francisco -> yellow Colma -> yellow Daly Ci
ty -> yellow Balboa Park -> green Balboa Park -> green Glen Park -> green 24th Street Mission -> green 16th Street
Mission -> green Civic Center -> green Powell Street -> green Montgomery Street -> green Embarcadero -> green West
Oakland -> green Lake Merritt -> green Fruitvale -> green Coliseum -> gray Coliseum -> gray OAK -> arrive OAK
---------------------------------
```

Betweenness Centrality



Application

Pinpointing Key Transit Hubs with the BART network.

	name	betweenness
0	yellow MacArthur	6531.833333
1	orange Coliseum	5765.666667
2	orange MacArthur	5585.833333
3	yellow Rockridge	5509.000000
4	orange Lake Merritt	4998.000000
...	...	...
209	arrive Warm Springs	0.000000
210	depart West Dublin	0.000000
211	arrive West Dublin	0.000000
212	depart West Oakland	0.000000
213	arrive West Oakland	0.000000

Business Impact

These key hubs should have enhanced facilities to manage heavy traffic flow enhancing overall responsiveness of transit hub.

Louvain Algorithm

Application

Strength in detecting communities within networks.

Business Impact

Create express delivery zones for these communities.

	name	community	intermediate_community
0	arrive 16th Street Mission	7	[3, 7, 7]
1	arrive 24th Street Mission	7	[7, 7, 7]
2	blue 16th Street Mission	7	[3, 7, 7]
3	blue 24th Street Mission	7	[7, 7, 7]
4	depart 16th Street Mission	7	[3, 7, 7]
...	...	...	...
209	orange Berryessa	91	[17, 57, 91]
210	orange Fremont	91	[41, 91, 91]
211	orange Milpitas	91	[57, 57, 91]
212	orange Union City	91	[91, 91, 91]
213	orange Warm Springs	91	[95, 57, 91]

Neo4j - Why is it Better?

Complex Joins

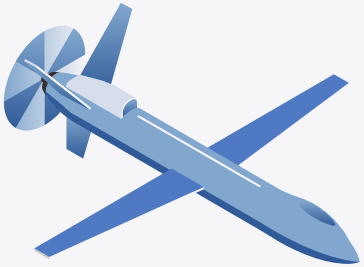
Inefficient queries that are slow to execute as the size of the network grows.

Inadequate for Dynamic Queries

Adjusting routes on-the-fly based on live data would be slow

Scalability Issues

Neo4j's schema-less nature allows for easier scalability.



Thanks!