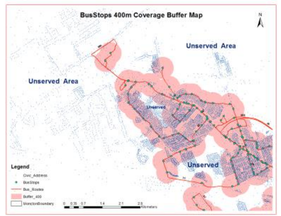
How can graph databases improve transit systems?

Hung Cao, Ikechukwu Maduako, Emerson Cavalheri, Ryan Brideau, Monica Wachowicz
{hcao3, imaduako, e.cavalheri, Ryan.Brideau, monicaw}@unb.ca
People in Motion Lab, University of New Brunswick

Motivation

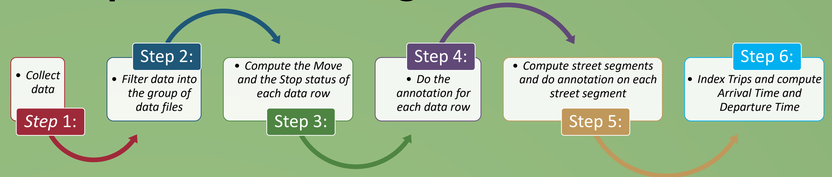
The metropolitan area of Greater Moncton is the fastest growing census metropolitan area in eastern Canada. We are working with CODIAC transit to promote efficient transportation services and reduce private car dependency.

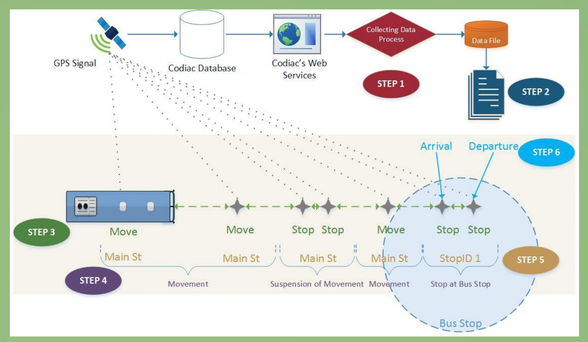
Bus Stops Coverage

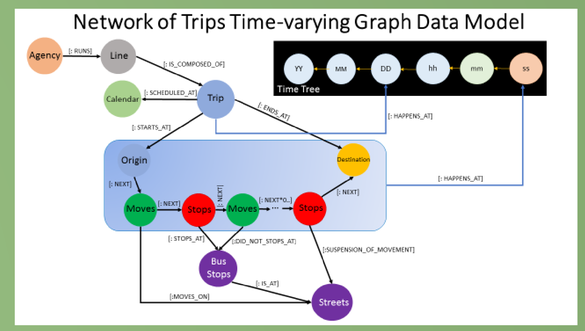


CODIAC Transpo currently operates 30 regular routes Monday to Saturday, some of which provide additional evening and Sunday services. The map shows the 400m catchment areas of all bus stops in Moncton.

Graph Database Design



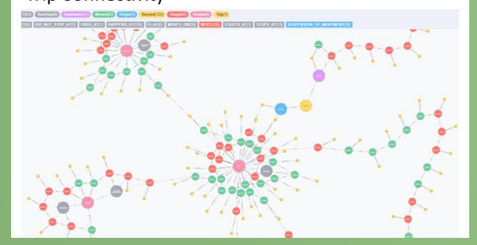




The graph database contains the geographical location of each bus every 5 seconds during a period of 2 weeks. Approximately 900,000 nodes and 4.5 million relationships have been created for the 30 regular routes.

Queries

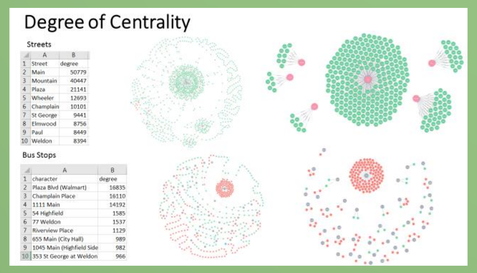
Trip Connectivity



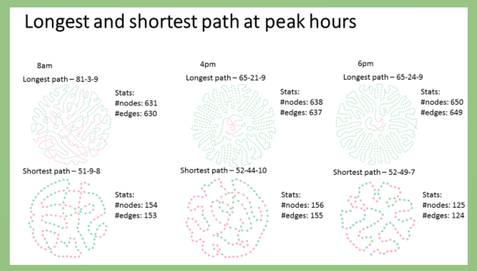
The trip connectivity graph above shows one specific trip taken by a bus. The nodes of the network represent the actual stops and moves that have occurred during this trip, as well as their geographical location type (e.g. the street segment, bus stop, or street intersection).

The degree centrality graphs below are showing the nodes with the highest number of connections. The graph database allow us to retrieve the busiest streets and bus stops in the network according to the centrality measure. The bus stops located at the Plaza is busiest bus stop and Main street is the busiest street in the network.

Degree of Centrality

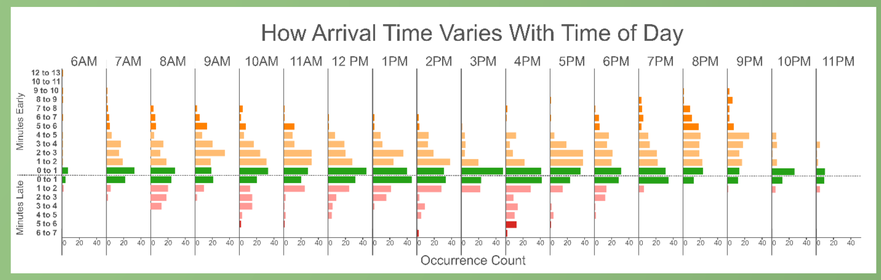


Longest and shortest path at peak hours



The graphs above show the longest and shortest trips over two weeks. We have used the longest and shortest path queries to count the number edges for all trips and find the longest and shortest trips by hour. It can be seen that the number of edges vary with time, space and traffic volume.

Data Visualization



This chart compares the actual arrival time to the expected arrival time by hour of day for a single bus route. It uses one week of data, from June 3rd to June 9th, and compares the arrival times for every stop along the route. At times of day where there is lower ridership, the charts skew toward early arrivals since the buses do not pause at the stops without riders, causing them to be early for the following stops.

Conclusions

This research project has developed a graph database model to provide easy-to-access information on transit options, incentivize the use of transit, and explore different transport capacities. Graph databases are unique and novel because they provide different stakeholders with a new set of policy education, exploration, and management options. The Black Arcs anticipates civic engagement tools as a key component to further explore the use of graph databases in transit planning.