Los Angeles Traffic Prediction Network

Cody West
CS5368

Full SAMIAM network is available at: https://github.com/cww2697/LA_Traffic_Bayesian_Net

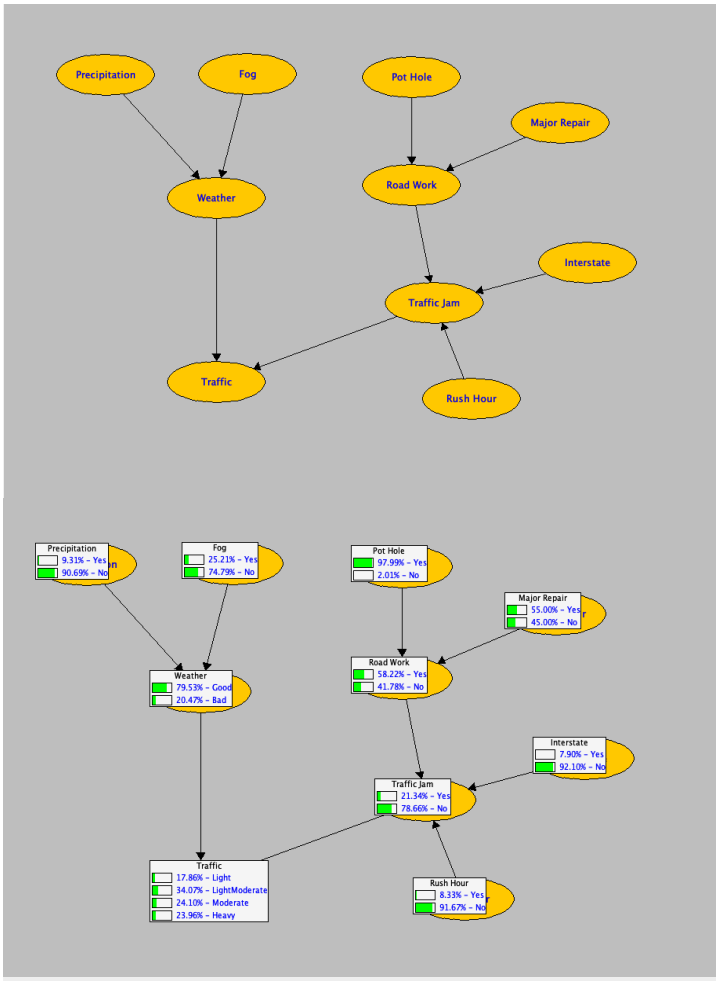
---

Network

The network I have designed attempts to predict the amount of traffic on streets in Los Angeles, CA. The location of the domain needed to be fixed to have variables with realistic values, since many of the values probability will differ based on the location where the experiment is being performed. The variances in these variables will be described later in this paper.

Below are the diagrams of the network I have designed to test, as well as the network with the non fixed probabilities of each condition shown.
The network is modeled with the following variables and statistics:

Query Variables: Traffic, Weather, Traffic Jam, Road Work
Evidence Variables: Precipitation, Fog, Pot Hole, Major Repair, Interstate, Rush Hour



Nodes:	10
Edges:	9
Min domain cardinality:	2
Max domain cardinality:	4

Variable Descriptions and Conditional Probability Tables

Query Variables

Traffic: Traffic is listed as a query variable since the value of Traffic is what the network is trying to find. Traffic shows the traffic load on a street based on the probabilities of both Weather and Traffic Jam. Traffic can be assigned four different values that are based on the probabilities of Traffic Jam and Weather. The values that can be assigned are **light traffic**, **light-moderate traffic**, **moderate traffic**, and **heavy traffic**. Heavy traffic is based on the probability of **Traffic Jam** and **Weather** being in the worst case scenario. Moderate Traffic is the possibility of at least one being in the worst case. **Light-Moderate** is the possibility of exactly one being in the worst case, and **Light** is the possibility of neither being in the worst case. These values represent the amount of traffic on the street at any given time.

Weather	Good		Bad	
Traffic Jam	Yes	No	Yes	No
Light	0.0075	0.275	0.03	0.025
LightModerate	0.2675	0.45	0.04	0.075
Moderate	0.275	0.2675	0.25	0.1
Heavy	0.45	0.0075	0.68	0.8

Weather: Weather is a query variable because its value depends on Precipitation and Fog. This variable has the potential to be expanded to include other weather based factors such as wind, smog, etc. these expanded factors would increase the accuracy of being able to decide if weather is listed as good or bad. The values that can be assigned are either **good** or **bad**, Good weather is based on the probability of Precipitation and Fog both being 0%, Bad weather is based on the probability of both being 100%. These values both decrease as they reach 50% on earth evidence variable. These values represent the current weather on a binary scale.

Precipitation	Yes		No	
Fog	Yes	No	Yes	No
Good	0.6782	0.9304	0.2286	0.9765
Bad	0.3218	0.0696	0.7714	0.0235

Traffic Jam: Traffic Jam is a query variable, since the value relies on the value of **Rush Hour**, **Road Work**, and **Interstate**. Traffic Jam represents the probability of the roads being at standstill traffic. The value can then be used to determine the traffic levels on the roads. The values that can be held are binary, and either hold **Yes** or **No**.

Rush Hour	Yes				No			
Road Work	Yes		No		Yes		No	
Interstate	Yes	No	Yes	No	Yes	No	Yes	No
Yes	0.9941	0.0059	0.920949	0.779442	0.920949	0.220558	0.220558	0.079051
No	0.0059	0.9941	0.079051	0.220558	0.079051	0.779442	0.779442	0.920949

Road Work: Rush hour's values relies on the probabilities gained from **Pot Hole** and **Major Repair**. The rush hour node represents the probability that there will be a construction zone on the streets of Los Angeles. Since, typically, pot holes are minor the affect on this node are reduced, while major repairs make up a higher percentage of the probability.

Pot Hole	Yes		No	
Major Repair	Yes	No	Yes	No
Yes	0.9999	0.0735	0.9815	0.0015
No	0.0001	0.9265	0.0185	0.9985

Evidence Variables

Rush Hour: The probability of Rush hour is based on the assumption of Rush Hour being 7:30 a.m. to 8:30 a.m. and 4:30 p.m. to 5:30 p.m. These factors can not be changed since it is an observed value and there are twenty-four hours in a day. Rush Hour represents the state of the current time being in one of those two ranges or not.

Yes	0.0833
No	0.9167

Precipitation: The probability of Precipitation is based on the average number of days of precipitation in a year for Los Angeles, CA. This value, while it can change, is going to be handled as a static since many years of data would be needed to change the data. The value is based on data obtained from the National Weather Service Forecast Office (36 days per year).

Yes	0.0931
No	0.9069

Fog: The probability of Fog is based on the average number of days of fog in a year for Los Angeles, CA. This value, while it can change, is going to be handled as a static since many years of data would be needed to change the data, much like **Precipitation**. The value is based on data obtained from the National Weather Service Forecast Office (92 days per year).

Yes	0.2521
No	0.7479

Pot Hole: The values in Pot Hole represent the probability of encountering a pot hole during a drive. This variable was used as pot holes will slow down traffic, with many people either slowing down or swerving to avoid them. This value is an inferred value as pot holes are extremely common and it would be almost impossible to count each pot hole and find the average number per mile.

Yes	0.9799
No	0.0201

Major Repair: Major repairs represents the need for major repairs in the road. Typically these types of problems in a road are rare and as such are represent by lower values for probability. This value is also an inferred value since much like potholes we are unable to calculate the exact value.

Yes	0.55
No	0.45

Interstate: The value of interstate represents the likelihood of the street in question being a highway. These values are estimates based on the total miles of road in Los Angeles and the total mileage of Highway.

Yes	0.079
No	0.921

Test Cases

Test Case 1 (Major Repair = Yes):

Test case 1 sets the value of the major repair variable to Yes. This affects the values of Road Work, Traffic Jam, and Traffic.

The values of the effected variables values have been provided.

Road Work:

Yes: 99.95%

No: 0.05%

Traffic Jam:

Yes: 25.98%

No: 74.02%

Traffic:

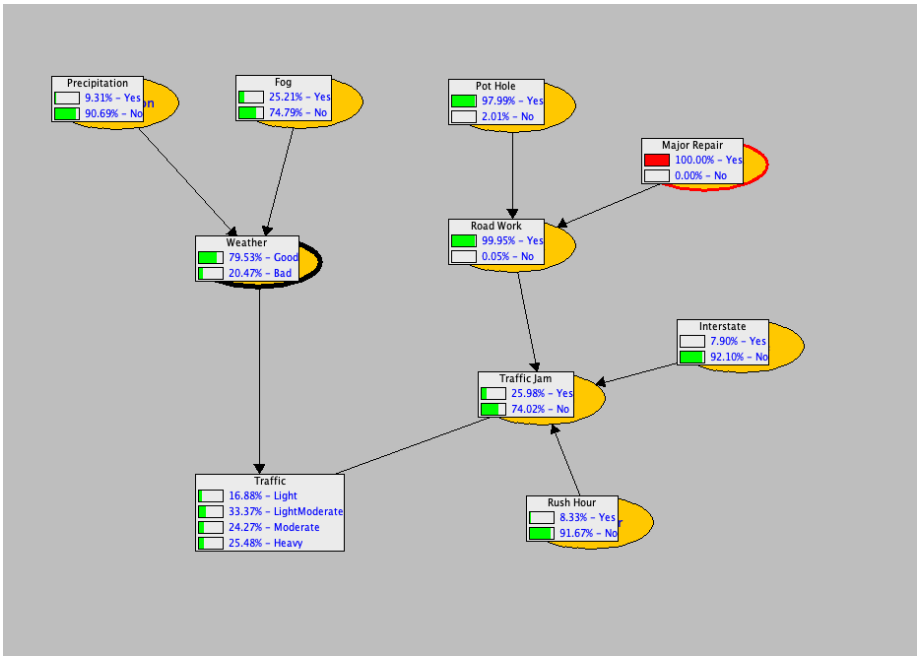
Light: 16.88%

Light-Moderate: 33.37%

Moderate: 24.27%

Heavy: 25.48%

These values are accurate based on estimations since there is a significant chance that even though there is road work, Interstate and Rush Hour are much more likely to be the **No** value, thus making it more likely that there is no traffic jam. This in turn makes light or light-moderate significantly more likely than the two higher values. The following image shows the full network after the test case has been set, along with the probabilities for all variables.



Test Case 2 (Interstate = Yes):

Test case 2 sets the Interstate variable to the Yes value, and therefore affects the values of Traffic Jam.

For test case 2 the values of Traffic and Traffic Jam have been provided.

Traffic:

Light: 8.49%

Light-Moderate: 27.33%

Moderate: 25.73%

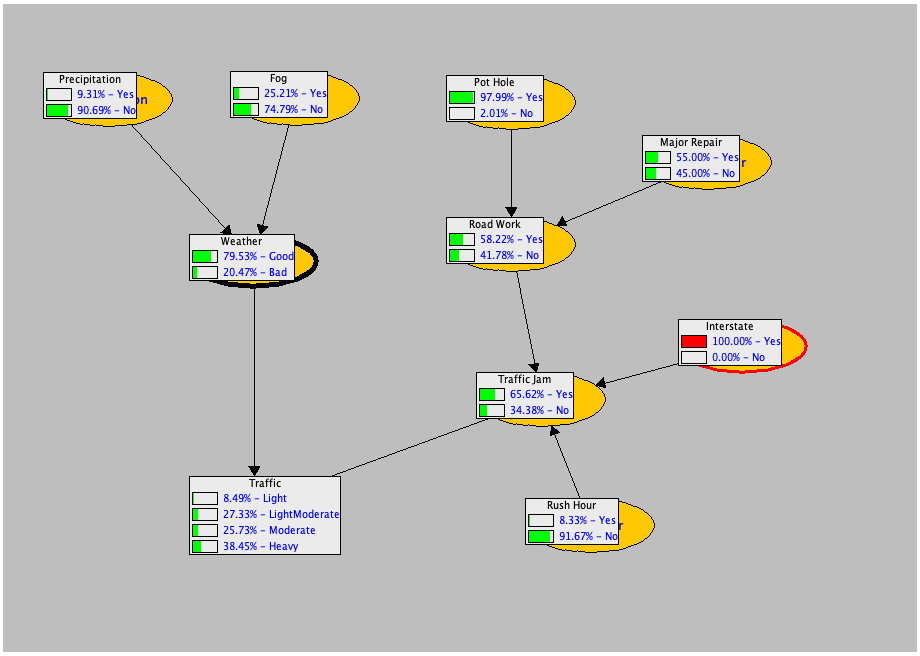
Heavy: 38.45%

Traffic Jam:

Yes: 65.62%

No: 34.38%

These values show that while all values have shifted towards heavy traffic due to the significant increase in p(traffic jam = yes) increase, traffic remaining light and light-moderate is still a major option. This makes sense on a highway, since there is an increase in road capacity on these types of streets. The following image shows the full network after the test case has been set, along with the probabilities for all variables.



Test Case 3 (Weather = Bad):

Test case 3 exhibits the explaining away phenomenon. This is shown by the overall decrease in chance in precipitation and increase in fog if the weather is a 100% chance for Bad Weather.

The overall Traffic probabilities after assigning weather to Bad are listed as:

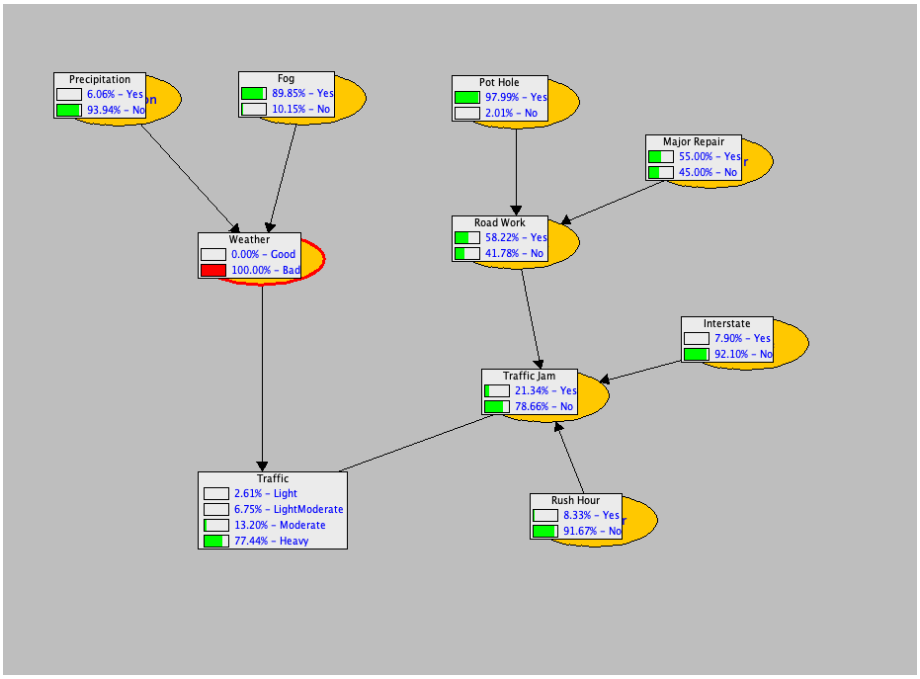
Light: 2.61%

Light-Moderate: 6.75%

Moderate: 13.20%

Heavy: 77.44%

These probabilities show that the Bayesian network is performing as it should. Since when the weather is in the negative state (**BAD**) the chance of having heavy traffic is much higher than in other scenarios. The following image shows the full network after the test case has been set, along with the probabilities for all variables.



Sources

Service, N. (2002, January 01). Weather Prediction Center (WPC). Retrieved November 04, 2020, from https://www.wpc.ncep.noaa.gov/

Total Cloudy and Foggy Days at US Cities a Year. (n.d.). Retrieved November 04, 2020, from https://www.currentresults.com/Weather/US/cloud-fog-city-annual.php

Miles of Public Roads Los Angeles County. (n.d.). Retrieved November 04, 2020, from http://www.laalmanac.com/transport/tr01.php