A Bayesian Approach to Understanding the Influence of Traffic Congestion given the Road Structure

Darren Blackwell and Ritesh Ajoodha

The University of the Witwatersrand, Johannesburg, South Africa,
1708102@students.wits.ac.za, ritesh.ajoodha@wits.ac.za
https://www.wits.ac.za/

Abstract. Traffic congestion analysis is a widely researched field, resulting in studies that have become increasingly more accurate by using machine learning paradigms. Paradigms used by others primarily include artificial neural networks (ANNs) or deep belief networks (DBNs) which provide accurate results, however, these models are uninterpretable. These models place a large focus on more traditional features such as weather conditions, time, and road type which allows for accurate predictions however, falter when learning a model description of this problem to learn the influence of congestion between roads.
A Bayesian network is a probabilistic model that will allow us to both incorporate prior knowledge, and understand the relationships within the model. This research proposes the use of a Bayesian network in combination with road connections (implemented as features) to determine the influence that road structure has on congestion. This will allow our model to make predictions of congestion based on surrounding roads' congestion.

1 Introduction

The need for accurate traffic prediction has become prevalent in modern times. Motor traffic influences the economic, social, and environmental aspects of society [8]. For this reason, the field of traffic analysis has been explored quite thoroughly and, with advancements in machine learning, have provided highly accurate and reliable results. These results have become so reliable, that software is able to predict elapsed travel time accurately between two locations with traffic congestion between roads.

With the field stacked in favor of making accurate predictions, it is also vital to understand the development of traffic and how it fluctuates. More traditional studies focus entirely on general trends that occur with traffic, however as highlighted in [6], the factors that influence traffic are unlikely to change. This is the main reason why traffic congestion can be predicted so accurately. These traditional methods of traffic prediction unfortunately falter in predicting how traffic collects and disperses after irregular events such as accidents occur.

Although it is important to make accurate predictions in traffic congestion, there is a neglect in attempts trying to understand how traffic propagates or develops in a network. This research attempts to provide insight in these areas of neglect by using a more interpretable paradigm, as well as using a different approach on which features to use. By focusing purely on the structure of the road network rather than other influencing factors we will attempt to learn how traffic congestion in some areas affect traffic congestion in other areas in both direct and indirect manners. This leads to the research question "How much does traffic congestion on a collection of roads influence the traffic congestion on another collection of roads given the state of congestion in the initial collection within a road network?"

The main inspiration behind this research is related to Braess's paradox, which is simply the observation that closing roads in a road-network will improve the traffic flow of the system. This provides our research with a reason to investigate how much the roads influence each other in a network. By investigating the influences in this manner, we could optimistically hope for contributions in better traffic management systems, improving traffic flow overall as well as understand how traffic propagates when irregular events occur.

Traditionally, traffic congestion is researched with respect to factors such as weather, time and road type. This research proposes traffic analysis with respect to the structural layer component of traffic congestion. Investigating the influence particular roads have on other roads behaviour, could provide better insight to how sporadic changes in particular locations, affects surrounding areas.

The final results of the research provide compelling evidence that supports Braess's paradox but also extends this by identifying how much roads influence each other. The resulting model has the ability to make relatively accurate predictions given evidence of traffic congestion on other roads.

2 Related Work

2.1 Underlying structure of traffic

The ability to decompose traffic into multiple sources allows us to evaluate the influence each source has on the entire entity. Thus we can break down traffic into three main sources, structural, behavioral, and external influences. For the purpose of this research, only the structural layer is investigated.

Structural Layer The structural layer encompasses the physical layout of the road, road type, and traveling direction and should not consider the quality of the road itself. It is highlighted in [3] how the impact of road type particularly in freeways and highways affects traffic flow. Although road structure is unlikely to change, a popular concept of mitigating traffic congestion would be to physically add more roads or somehow increase the surface area that allows motor vehicles to travel. However, Braess' paradox shows how closing particular roads can improve the overall flow of traffic. The importance of this is that the structure of the road influences the creation and elimination of traffic congestion. This is a powerful tool when predicting congestion in areas where traffic anomalies occur in specific locations since structure influence is independent of historical behavior.

An example of a direct influence of traffic would be that of a 'grid-lock', you are unable to travel in a specific direction since cars in that direction are also currently unable to move, this is probably the most notable type of traffic congestion. Inverse influence of traffic would be that in a particular area, point A, where there is a 'grid-lock' of cars, fewer cars would be able to travel to point B consequently improving the flow at point B.

2.2 Structure Learning

For Bayesian Networks, it is not always the case that we know how the features relate to one another. If this is the case, we make use of the process of structure learning, which learns the structure (Directed-Acyclic-Graph) of the Bayesian Network from the data.

Hill-Climb The Hill-Climb algorithm is a heuristic search algorithm that is used to estimate the structure of the network by choosing the network with the optimal score, based on the scoring function.[1]

Scoring Functions Scoring a network requires a scoring function. Bayesian scoring functions computes the prior distribution of the data and attempts to maximize the posterior probability. Information-theoretic scoring functions relate to the compression that can be achieved over a data set. Hence why the BIC score (Information-theoretic) penalizes graphs that have a higher complexity [2]. Examples of scoring functions used in this paper:

- K2 (Bayesian Scoring)
- BDEU (Bayesian Scoring)
- BIC (Bayesian Information Criterion) (Information-theoretic scoring)

2.3 Machine learning in traffic prediction

Machine learning over recent years has developed quite rapidly due to its ability to learn and predict outcomes accurately in fields that might seem difficult for humans to interpret. Varying implementations of these models below have been used in the literature to investigate their research hypothesis. In conjunction with these models and methods, it is important to disclose how each model's accuracy is measured, all these models use an evaluation metric such as the mean absolute error(MAE), the mean absolute percentage error (MAPE) and the root mean square error (RMSE). These are popular metrics for determining how accurately a model predicts an output. In this section we discuss the Bayesian approaches to predicting traffic.

Bayesian Inference Bayesian inference is a statistical method of understanding a problem area using prior information. Statistical methods are often less accurate than neural networks but can produce a more intuitive understanding of the influences of the data. This can be seen in [9] where Bayesian inference is used to predict the evolution of a driver's route choice daily. Using a Markov chain Monte Carlo algorithm the paper was able to determine interpretable results in determining route changes. Examples of inference include the average treatment effect (ATE, a causal inference procedure) that shows the average effect that an intervention variable has on a target variable. Variable elimination and belief propagation (junction tree algorithm), which are both exact inference algorithms that can determine the influence that multiple different variables have on each other.

Bayesian Networks Little traffic prediction methods have used the Bayesian Network paradigm since generally there is a large focus on how accurately one can predict congestion. This accuracy leads to studies leaning more towards Neural Networks which provide greater accuracy, however lack in interpretable results. [5] Uses Bayesian networks in the attempt to model crashes on freeways. The model used allowed for accuracies such as 82 percent to be determined in the best case, with a 20 percent false alarm rate. This study did not provide incredibly high accuracy in comparison to other models, but produced respectable predictions with the ability to detect inference between variables. [7] provides valuable information on how to build a Bayesian network to predict traffic within area. This research provides valuable information to this research however lacks in both experimental results or how accuracies were obtained and seems purely theoretical.

3 Method and Data

3.1 Method

The research question is studied experimentally. Our dataset comprises of FCD data. A Bayesian network will be used to model the conditional independencies between roads that manifest in the data. The Bayesian network is constructed such that each feature represents a particular road that can take the form of 6 congestion states, i.e. X = $\{c_0, c_1, c_2, c_3, c_4, c_5\}$. The process of determining the Bayesian Network was completed via structure learning. Where different scoring metrics were used with the Hill-Climb algorithm. The learned networks were compared against each other to determine which model had the optimal score. In particular, we compare the score of the learned model against a model determined by the physical road network (with cycles removed) in addition the same physical road structure is used as a starting 'skeleton' for our structure learning algorithm. Variable elimination is then applied to the optimal model to determine the influence particular roads have on other roads congestion states.

3.2 Data

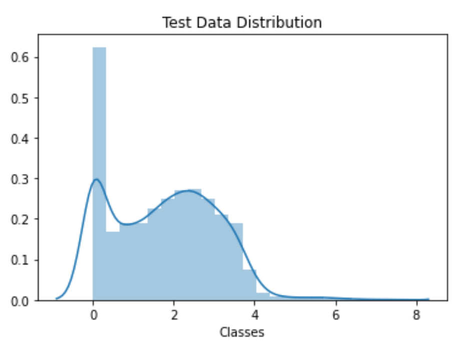
For this research we will be using FCD (Floating car data) obtained from [4], particularly a subset of FCD relating to flow. This data provides us with accurate measures of congestion between a link of two points. Each link (physical road) contains a floating point 'jam factor' feature which indicates how congested the road is on a scale from 0 to 10. (10 being idle traffic, 0 indicating free-flow).

4 Process and Results

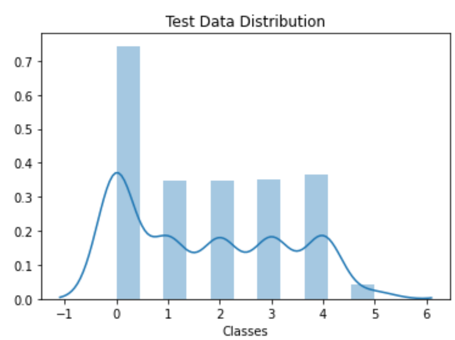
4.1 Data Pre-Processing

Figure 1 represents the distribution of the test data that was used in the experiments. Although the congestion value has a maximum of 10 (generally closed roads or accidents), it is around congestion level 4 that seems subjectively like a 'grid-lock'.

In order to accommodate for different distributions of data with respect to their areas, it is important to maintain the histograms general shape when trying to sort it into 6 different classes. For this we use a technique called histogram equalization (generally used in image processing). While trying to evenly distribute the data amongst the specified classes, it still maintains the 'integrity' of the data (i.e. congestion levels remain the same relative to each other).



1a) Distribution of traffic values over two days in Johannesburg city center



1b) Distribution of traffic values over two days in Johannesburg city center after histogram equalization and binning the congestion values into separate classes

When comparing figure **1a** to figure **1b**, we find that there is still a bias towards the lower levels of congestion, and there is still a 'dip' closer to the higher levels of congestion since histogram equalization maintains the 'integrity' of the data relative to each other, with a more even distribution over the full range of values it can embrace. The 'bump' that's originally in figure two has been stretched out over the center classes.

4.2 Structure learning and Scoring models

An integral portion of this research involved learning the structure of our Bayesian Network. For the purpose of this research, 7 models are compared against each other. The first model is the road structure, this structure is formed as a direct result of how the roads are physically connected. This however contains cycles, which are removed using [10]. The remaining 6 models are formed from using the Greedy-Hill climb algorithm with three different scoring functions (BIC, K2 and BDEU) to be optimized. Three of the 6 models use the road structure as a starting skeleton for the greedy hill climb algorithm. The following results were recorded:

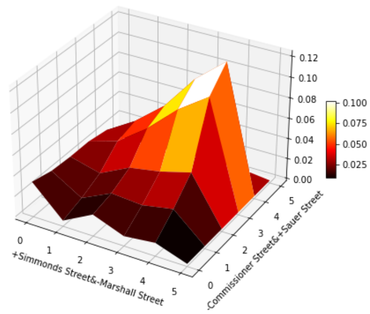
Model Type	BIC Score	K2 Score	BDEU Score
Road Structure	-3120495.6	-3079.8	-3257.2
HC_BIC_Empty	-3819.1	-3750.5	-3761.0
HC_BDEU_Empty	-4086.3	-3606.0	-3623.2
HC_K2_Empty	-5867.5	-3572.0	-3728.3
HC_BDEU_Structure	-1839864.4	-3755.5	-3882.7
HC_BIC_Structure	-3819.1	-3750.5	-3761.0
HC_K2_Structure	-19171.3	-3596.2	-3894.1

Table 1. Table showing the learned models along with their corresponding scores across three different metrics.

From the results obtained, the model with the 'best' score is difficult to determine since we have no ground truth to compare it to. Whilst the road structure has the best K2 and BDEU score, it has an extremely low BIC Score since BIC score penalizes model complexity, this would make processes such as variable elimination not viable. The model that was chosen for the remainder of the results was the 'HC_BDEU_Empty' model since it has a relatively low BIC Score, and has the highest average between all three scores.

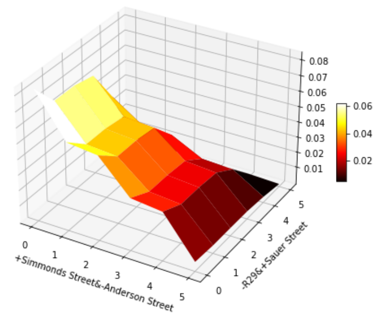
4.3 Variable Elimination

Variable elimination (an exact inference procedure) can perform more advanced queries over multiple different features in a network. The purpose here is to demonstrate a more direct influence between two roads alone. The following figures (Figures 2a to 2f) show the probabilities (z-axis) of a single roads' congestion (x-axis), given evidence of the congestion on another road (y-axis).



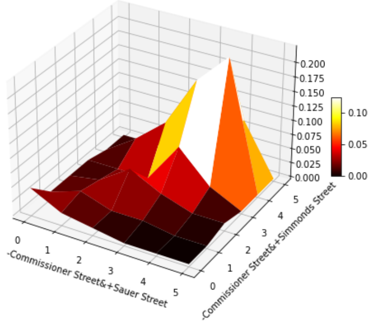
2a)

The most common type of relation between two roads, phi(z axis) increases as the both x and y increase their congestion class. This can be interpreted as whilst one roads congestion increases, so does the other. Resulting in this 'positive' gradient.



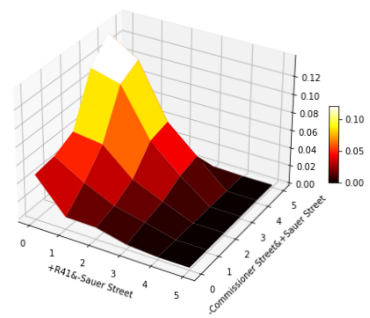
2b)

This case is the same concept as **2a**, except with a higher distribution of classes closer to 0. The two axis however still correlate to each other in a positive manner.



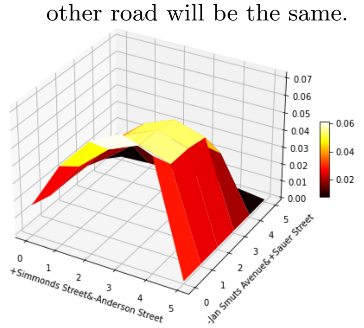
2c)

Similar to Figure **2a**. It has a 'positive' slope, however there are peaks forming where the two classes are equal. This is a result of traffic congestion being relatively the same for the two roads. Interpreting it could be viewed as, if I experience the traffic level on this road, it is likely the other road will be the same.



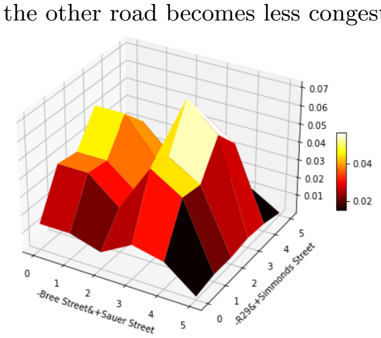
2d)

This category corresponds to a negative ATE. Although not as frequent as other cases,it is a perfect example of Braess's paradox. As congestion increases (for example along the y axis), there is an inverse relation between the classes. i.e. if we increase congestion on the one road, the other road becomes less congested.



2e)

The elongated arch shape is a result of the two roads being relatively independent. As the one axis increases, the other axis *generally* remains at the apex of the arch.(In this case as the y-axis increases, the x-axis remains at class 2). Other forms similar to this exist (such as half of the arch).



2f)

The irregular shape in this figure is similar to that of the arch shape but with a crevice in the center, forming two peaks on either side. These peaks have the highest probability of occurrence, suggesting that there would be uncertainty inferring traffic levels for variables that axis follow parallel the crevice.

5 Conclusion

In this paper I have presented a Bayesian Network approach to understanding how different roads congestion levels influence each other by allowing us to extract inferences by querying over a sub-set of particular roads. This gives valuable insight to *what* causes traffic as opposed to making accurate predictions. The models learned in this paper can also be used to predict traffic congestion levels, however provide a relatively low accuracy when trying to predict congestion. Varying from 50% on average to 80-90% accuracy depending on the road. This is potentially due to the histogram equalization of the data (By removing a constant congestion state bias, prediction accuracy will decrease) and excluding features such as weather. In conclusion, using Bayesian Networks to determine the influence roads have on each other provided novel results using variable elimination. This research will also hopefully contribute towards designing smart traffic management systems by distributing congestion in a manner that optimizes roads that have little influence on each other, or an indirect influence on each other, increasing flow. The applications of this paper could also provide interesting results when investigating other types of networks rather than just that of road traffic.

Acknowledgements

This work is based on the research supported in part by the National Research foundation of South Africa (Grant number: 121835).

References

1. Ankur Ankan and Abinash Panda. pgmpy: Probabilistic graphical models using python. In *Proceedings of the 14th Python in Science Conference (SCIPY 2015)*. Citeseer, 2015.
2. Alexandra M. Carvalho. Scoring functions for learning bayesian networks, 2009.
3. W. Elleuch, A. Wali, and AM. Alimi. Towards an efficient traffic congestion prediction method based on neural networks and big gps data. In *IIUM Engineering Journal*, volume Vol. 20, pages 108–118, Jun. 2019.
4. Here Traffic. https://developer.here.com/cn/documentation/traffic/topics_v6.1/example-flow-intro.html.
5. M. Hossain and Y. Muromachi. In *A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways*, volume 45, pages 373–381, July 2011.
6. S. Kolidakis, G. Botzoris, V. Profillidis, and P Lemonakis. Road traffic forecasting - a hybrid approach combining artificial neural network with singular spectrum analysis. In *Economic Analysis and Policy*, number 64, pages 159–171, Aug. 2019.
7. Y. Liu, X. Feng, and Q. Wang. In *Prediction of Urban Road Congestion Using a Bayesian Network Approach*, pages 671–678, 2014.
8. I. Papamichail and N. Bekiaris-Liberis. In *Motorway traffic flow modelling, estimation and control with vehicle automation and communication systems*, Sept. 2019.
9. K. Parry and M. Hazelton. Bayesian inference for day-to-day dynamic traffic models. In *Transportation research Part B*, pages 104–115, Sep. 2013.
10. Jiankai Sun, Deepak Ajwani, Patrick K. Nicholson, Alessandra Sala, and Srinivasan Parthasarathy. Breaking cycles in noisy hierarchies. In *Proceedings of the 2017 ACM on Web Science Conference*, WebSci '17, pages 151–160, New York, NY, USA, 2017. ACM.