Modelling Vehicular Traffic using Hidden Markov Model

Team Members:
15BCE0589 (Lakkysetty Nikhil)
15BCE0537 (Kamisetty Yeshwanth)
15BCE0624 (Pabba Yeshwanth)

Report submitted for the
Final Project Review of

Course Code: CSE3013 – Artificial intelligence

Slot: B2 + TB2
Professor: Dr.W.B.Vasantha

VIT
UNIVERSITY
(Estd. u/s 3 of UGC Act 1956)

Abstract:

Traffic congestion has become a severe problem in metropolises, resulting in widespread wastage of time and energy. Traffic monitoring and estimation is an important method for obtaining information on traffic conditions; thus, it plays a vital role in reducing traffic congestion. Static sensors inductive loop detectors, video cameras, etc., deployed at fixed locations on roads, are used to detect traffic state (e.g., flow velocity and traffic density). However, it is difficult for these traditional approaches to cover all roads because they involve extensive infrastructure deployment and high maintenance costs.

In order to overcome the problem, this paper proposes a hidden Markov model (HMM)-based traffic estimation model, in which the traffic condition on a road segment is considered as a hidden state that can be estimated according to the conditions of road segments having similar traffic characteristics. First, we identified a small segment (say 2 or 3 Kilometres) of a road that we want to study. We collected some data about the above chosen route by reading and saving the map every few seconds and analysing this offline. Divide the chosen route into small segments (as provided by the map). For each segment, find the traffic condition (jam/slow/fast) by "detecting" the colour from the map. After collecting data for sufficient duration (say 24 hours), save the data to the hard disk for processing offline. Thus, we can say that observed states are the colours from the map. Finally, we tried to predict traffic situation at a future using hidden Markov model.

In brief, we obtained the estimated congestion for the succeeding interval. And also we calculated the accuracy for multiple runs. Thus, we can model the vehicular traffic by suggesting the alternate path.

1. Introduction:

When cities prosper and begin to outgrow their infrastructure, traffic often grinds to a halt. The problem of traffic congestion has been increasingly severe in metropolises and results in wastage of time and energy. Traffic monitoring and estimation is gaining vital importance in recent years. Meticulous planning of cities that optimally distributes the resources and infrastructure across the area has to be undertaken. The HMM model can be used to predict the traffic situation. HMM is a doubly stochastic process with an underlying stochastic process that is not observable, but can only be observed through another set of stochastic processes that produce the sequence of observed symbols. Usually HMM is used in machine learning and pattern recognition fields such as speech recognition systems and molecular biology computational system. For example, we can find sequence of weather of a day, without knowing the actual weather first, based on number of ice creams eaten by certain person on corresponding day.

The forward algorithm is an inference algorithm for hidden Markov models which computes the posterior marginals of all hidden state variables given a sequence of observations, computes, for all hidden state variables ,the distribution . This inference task is usually called smoothing. The algorithm makes use of the principle of dynamic programming to compute efficiently the values that are required to obtain the posterior marginal distributions in two passes. The first pass goes forward in time while the second goes backward in time; hence the name forward-backward algorithm.

Thus, HMM model can be used to estimate the traffic congestion. By knowing this information, we could suggest alternate paths and thus regulate the traffic.

2. Literature Review Summary Table

Authors and Year (Reference)	*Title (Study)*	*Concept / Theoretical model/ Framework*	*Methodology used/ Implementation*	*Dataset details/ Analysis*	*Relevant Finding*	*Limitations/ Future Research/ Gaps identified*
Bingnan Jiand and Yunsi Fei 2015	*Traffic and Vehicle Speed Prediction with Neural Network and Hidden Markov Model in Vehicular Networks*	*Neural Networks to predict the traffic speed*	*Baum-Welch algorithm*	*Historical traffic data*	*Hidden Markov model to predict traffic*	*To perform better than KDE method*
Porikli.F, Li.X 2004	*Traffic Congestion Estimation Using HMM Models Without Vehicle Tracking*	*Gaussian Mixture Hidden Markov Models(GM-HMM)*	*Low-lattency traffic congestion estimation algorithm*	*MPEG video data*	*Training the traffic data set into patterns like empty, mild congestion, heavy congestion, stopped etc*	*To achieve more accuracy under illumination conditions like sunny, cloudy and different camera setup.*

Xiaomeng Wang, Ling Peng, Tianhe Chi, Mengzhu Li, Jing Shao 2015	*A Hidden Markov Model for Urban-Scale Traffic Estimation Using Floating Car Data*	*Hidden Markov based traffic estimation model*	*Algorithm based on clustering and pattern mining*	*Floating Car Data*	*Clustering to find clusters with road segments having similar traffic characteristics*	*To achieve more efficiency than model based on adjacency relationships for traffic estimation.*
Bowu Zhang, Kai Xing, Xiuzhen Cheng, Liusheng Huang and Rongfang Bie 2012	*Traffic Clustering and Online Traffic Prediction in Vehical Networks: A Social Influence Perspective*	*Online traffic Clustering Technique*	*Neural Network based traffic prediction algorithm*	*Historical traffic data*	*Predicting traffic conditions cluster by cluster*	*To measure the performance of the algorithm when large data is taken.*
Constantinos Antoniou, Haris N. Koutsopoulos and George Yannis 2007	*Traffic state prediction using Markov chain models*	*Markov Chain model*	*Model based Clustering*	*Irvine data set*	*Clustering algorithm to find clusters with road segments having similar traffic characteristics*	*To achieve automatic accident detection*
Qing-Jie Kong, Qiankun Zhao, Chao	*Efficient Traffic State Estimation for*	*Global Positioning System (GPS)*	*Curve fitting based method*	*Urban road network data*	*Dividing the path and predicting for each segment*	*To achieve accuracy for large*

Wei and Yuncai Liu 2013	*Large Scale Urban Road Networks*					*data set with more GPS vehicles*
Zhitang Chen, Jiayao Wen and Yahui Geng 2016	*Predicting Future Traffic using Hidden Markov Models*	*HMM based on kernel bayes rule*	*Recurrent Neural network algorithm*	*Real time data*	*Using HMM to determine the unknown states*	*To apply theproposed framework to real network monitoring and traffic engineering*
Xiaokun Li andFaith M. Porikli 2004	*A Hidden Markov Model Framework For Traffic Event Detection using video features*	*Gaussian Mixture Hidden Markov Model (GMHMM)*	*Discrete Cosine Transform, Viterbi algorithm*	*MPEG video data*	*Training the traffic data set into patterns like empty, mild congestion, heavy congestion, stopped etc*	*To extend the system for detection of similar traffic events.*
Evan Tan, Jing Chen 2007	*Vehicular traffic Density Estimation via Statistical Methods with Automated State Learning*	*Clustering scheme Autoclass, HMM*	*Feature extraction, ROI, Local Binary Pattern*	*Video data set*	*Clustering algorithm to find clusters with road segments having similar traffic characteristics*	*To improve more accuracy than GHMMM model.*

Lin LI, Jianing HU, Qiao HUANG, Jingyan SONG *2006*	*A Fuzzy Hidden Markov Model for Traffic Status Classification* *Based on Video Features*	*Fuzzy C-Means Hidden Markov Model* *(FCM-HMM)*	*Discrete Cosine Transform* *(DCT)*	*Real Time Video data*	*Traffic status is classified into different states. HMM is used to define states.*	*Achieved memory resource and computing speed but developing for high accuracy results*

3. Objective of the project:

- Collection and analysis of traffic data using Google Maps.
- Prediction of traffic congestion using hidden Markov model and Forward-Backward algorithm which is used for smoothing.
- Calculate the accuracy for multiple runs.
- Understand the pros and cons of this model and its advantage over other models.

4. Innovation component in the project:

An efficient traffic management system needs accurate traffic condition information. Since a modest system may consist of hundreds of video cameras, the computational complexity is another important consideration. Such systems also require maximum automation to decrease the burden on human operators as well. Most existing vision systems for monitoring road traffic relied on stationary cameras and vehicle tracking.

In this paper, we proposed an accurate, computationally simple, lighting and camera setup independent method. We used hidden Markov model to estimate the traffic congestion. We identified a small segment of a road and detect the color of the segment (denoting traffic congestion).

After this, we fed the data to build Hidden Markov Model. Finally, we applied forward algorithm and predicted the traffic congestion for the succeeding time intervals. This is the most innovation component in our project.

5. Work to be done and implementation

Methodology to be adapted:

- Identify a small segment of a road (3 to 4 Km). Find such a route using Google maps. Read and save the map for offline analysis. Detect the color of the segment (denoting traffic congestion).
- Reiterate the process for several segments across the city and save the data into a database.
- Feed the data to build Hidden Markov Model.
- And apply forward algorithm and predict the traffic congestion for the succeeding time intervals.
- Take snapshots of a road segment in Google Maps at regular intervals (say 15 mins) for 24 hrs.
- Apply suitable image processing techniques to extract colour codes from the image.
- Feed the sequences to the already implemented HMM model.
- Obtain the estimated congestion for the succeeding interval.
- Calculate the accuracy for multiple runs.

Dataset used:

a) We are getting the real-time data from using google maps. We detect the color of the segment which indicates traffic congestion. We take snapshots of a road segment at regular intervals and process this images using image processing techniques.

b) Yes, this project is based on other reference project.

c) This project used hidden Markov model which differs from reference project using FCD model.

Tools used:

Python interpreter, Spyder(python2.7), matlab.

6. Screenshots and demo:

Code for extracting colours(observations):

```
ssno=0
pixsum=[0]*119
for j in range(0,119):
      filename='Screenshots/1/'
      ssno=ssno+1
      filename=filename+str(ssno)
      im = Image.open(filename) #Can be many different formats.
      pix = im.load()

      #SEG1
      for i in range(0,5):
            P = pix[x[i],y[i]]
            B = sum(P)
            pixsum[j]=pixsum[j]+B
            M=[0,0,0,0]
            #print B
            #print P #Get the RGBA Value of the a pixel of an image
            if B in range(180,215):
                  M[0]=M[0]+1
            elif B in range(215,270):
                  M[1]=M[1]+1
            elif P[0]>200:
                  M[2]=M[2]+1
            else:
                  M[3]=M[3]+1
            a=max(M[0],M[1],M[2],M[3])
      if a==M[0]:
            OBSERVED_STATE[j]=OBSERVED_STATE[j]+'M'
      elif a==M[1]:
            OBSERVED_STATE[j]=OBSERVED_STATE[j]+'R'
      elif a==M[2]:
            OBSERVED_STATE[j]=OBSERVED_STATE[j]+'O'
      elif a==M[3]:
            OBSERVED_STATE[j]=OBSERVED_STATE[j]+'G'
```

Output for dataset1:

```
F:/WINTER_SEMESTER_2017-2018/AI/AIPROJECT/PythonHMM')
['MMMM', 'MROM', 'MROM', 'MROM', 'OGRR', 'OGRR', 'OGRR', 'RRRM', 'RRRM', 'MRMR', 'MRRM',
'MRRM', 'MMRM', 'MMRM', 'MMRM', 'RRRR', 'RRRR', 'RRRR', 'RMMR', 'MMRM', 'MMRM', 'MMRM',
'MMRM', 'MMRR', 'MMRR', 'MMRR', 'MMRR', 'RMRR', 'RRRM', 'RRRM', 'RRRM', 'MMRM', 'MMRM',
'MMRM', 'MMRM', 'MMRM', 'ROOR', 'OOOM', 'OOOM', 'ROOM', 'ROOM', 'ROOM', 'OGRM', 'OGRM',
'OGRM', 'MRMM', 'MRRM', 'MRRM', 'MRRM', 'MRRM', 'ROOO', 'ROOO', 'ROOO', 'RORM', 'RORM']
```

Output for dataset2:

```
['OORM', 'OORM', 'RGOM', 'RGOM', 'RGOM', 'RRRM', 'RRRM', 'RRRM', 'OMOG', 'OMOG', 'OOOG',
'OOOG', 'OOOG', 'OGOG', 'OGOG', 'OGOG', 'ORRG', 'ORMM', 'MRRO', 'MRRO', 'MRRO', 'RROO',
'RROO', 'RROO', 'OGGG', 'OGGG', 'OOMO', 'OORM', 'OORM', 'MRRM', 'MRRM', 'MRRM', 'RROM',
'RROM', 'RROM', 'OGOR', 'OGOM', 'OGRM', 'OGRM', 'OGRM', 'MRRR', 'MRRR', 'MRRR', 'MRRO',
'MRRO', 'GGOO', 'GGOO', 'GGOO', 'GGOO', 'GGOO', 'GGOO', 'GRRM', 'GRRM', 'MMRR', 'MMRR',
'MMRR', 'RGOO', 'RGOO', 'RGOO', 'OOGO', 'OOGO', 'OOGO', 'GORR', 'GORR', 'RMRR', 'RMRR',
'RMRR', 'MMRR', 'MMRR', 'MMRR', 'RGGM', 'RGGR', 'GGGO', 'GGGO', 'GGGO', 'RRGO', 'RRGO',
'RRGO', 'MMOO', 'MMOO', 'MMOO', 'MMOO', 'MMOO', 'MRRO', 'MRRO', 'MRRO', 'MMRO', 'MMRO',
'MMRO', 'OMGG', 'OMGG', 'OOGG', 'OOGG', 'OOGG', 'RRRG', 'RRRG', 'RRRG', 'MMRG', 'MMRG',
'OMRG', 'OMRG', 'OMRG', 'ROOG', 'ROOG', 'ROOG', 'GGGG', 'GGGG', 'RRRO', 'RRRG', 'RRRG',
'MMRG', 'MMRG', 'MMRG', 'RRRG', 'RRRG', 'RRRG', 'OOOG', 'OOOG', 'OGGG']
```

Code for forward algorithm:

```
def evaluate(self, sequence):
    """
    Use the `forward algorithm

    """
    length = len(sequence)
    if length == 0:
        return 0

    prob = 0
  predicted_symbol = 'RRRR'
  predicted_state = '1'
    alpha,index = self._forward(sequence)
  maximum = 0
    for state_to in self._states:
        prob = 0
        for state_from in self._states:
            prob += alpha[index - 1][state_from] * \
                self.trans_prob(state_from, state_to)
      for symbols in self._symbols:
            if(prob * self.emit_prob(state_to, symbols) > maximum):

           maximum = prob * self.emit_prob(state_to, symbols)

           predicted_symbol = symbols

           predicted_state = state_to
    print "The hidden state predicted for the next interval is: "
    return (predicted_state)

def decode(self, sequence):
```

Output for dataset1:

```
MILDLY CONGESTED': 6.926012531170814e-51, 'MODERATELY CONGESTED': 0.0, 'HEAVILY CONGESTE
D': 0.0}, {'MILDLY CONGESTED': 2.597254699189055e-51, 'MODERATELY CONGESTED': 0.0, 'HEA
VILY CONGESTED': 0.0}, {'MILDLY CONGESTED': 9.739705121958957e-52, 'MODERATELY CONGESTED
': 0.0, 'HEAVILY CONGESTED': 0.0}, {'MILDLY CONGESTED': 0.0, 'MODERATELY CONGESTED': 2.0
291052337414493e-53, 'HEAVILY CONGESTED': 6.24340071920446e-54}, {'MILDLY CONGESTED': 0.
0, 'MODERATELY CONGESTED': 1.410162040842072e-54, 'HEAVILY CONGESTED': 2.382649787348065
e-55}]
The hidden state predicted for the next interval is:
MODERATELY CONGESTED
>>> |
```

Output for dataset2:

```
HEAVILY CONGESTED': 0.0}, {'MILDLY CONGESTED': 0.0, 'MODERATELY CONGESTED': 4.1093362310
334416e-152, 'OPEN': 0.0, 'HEAVILY CONGESTED': 0.0}, {'MILDLY CONGESTED': 0.0, 'MODERATE
LY CONGESTED': 0.0, 'OPEN': 5.593470368920293e-154, 'HEAVILY CONGESTED': 0.0}, {'MILDLY
CONGESTED': 0.0, 'MODERATELY CONGESTED': 0.0, 'OPEN': 7.23293582187969e-155, 'HEAVILY CO
NGESTED': 0.0}, {'MILDLY CONGESTED': 0.0, 'MODERATELY CONGESTED': 0.0, 'OPEN': 5.6117605
5145838e-156, 'HEAVILY CONGESTED': 0.0}]
The hidden state predicted for the next interval is:
OPEN
>>> |
```

7. Results and discussions:

Firstly, the colour of the snapshots for each segment are extracted using image processing techniques to get the sequence of observations. With these sequence of observations, we estimated the traffic congestion using hidden Markov model and forward algorithm. In the above output for the dataset1, the probability for being in ‘moderately congested’ state is high compared to other hidden states. Thus, hidden state predicted would be ‘moderately congested’. Similarly, for the dataset2, the hidden state predicted for next interval would be ‘open’.

8. Conclusion:

This paper presented an effective HMM based model to estimate traffic congestion using forward algorithm. A real time dataset is collected from google maps for one day under ideal conditions (no traffic delay due to accidents, rallies, VIP convoys.,etc) and extracted using image processing techniques. For this purpose, we have chosen a Banglore road as an area of our interest. There are four hidden states (mildly congested, moderately congested, heavily congested, open) in our project. We used forward algorithm to get the most probable hidden state using the sequence of observations which were extracted from snapshots. Thus, by predicting the hidden state for next interval we estimated the traffic congestion in that particular road under ideal conditions. By knowing this information, we could suggest an alternate path and thus regulate the traffic.

9. References

[1] Bingnan Jiand and Yunsi Fei," Traffic and Vehicle Speed Prediction with Neural Network and Hidden Markov Model in Vehicular Networks", IEEE Intelligent Vehicles Symposium (IV) June 28 - July 1, 2015. COEX, Seoul, Korea.

[2] Porikli.F, Li.X," Traffic Congestion Estimation Using HMM Models Without Vehicle Tracking", TR2004-019 March 2004.

[3] Xiaomeng Wang, Ling Peng, Tianhe Chi, Mengzhu Li, Jing Shao," A Hidden Markov Model for Urban-Scale Traffic Estimation Using Floating Car Data", PLoS One. 2015; 10(12): e0145348,2015 Dec 28.

[4] Bowu Zhang, Kai Xing, Xiuzhen Cheng, Liusheng Huang and Rongfang Bie," Traffic Clustering and Online Traffic Prediction in Vehical Networks: A Social Influence Perspective", Proceedings IEEE INFOCOM 2012.

[5] Constantinos Antoniou, Haris N. Koutsopoulos and George Yannis,” Traffic state prediction using Markov chain models”, Proceedings of the European Control Conference 2007 Kos, Greece, July 2-5, 2007

[6] Qing-Jie Kong, Qiankun Zhao, Chao Wei and Yuncai Liu, “Efficient Traffic State Estimation for Large Scale Urban Road Networks”, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 14, NO. 1, MARCH 2013.

[7] Zhitang Chen, Jiayao Wen and Yahui Geng,” Predicting Future Traffic using Hidden Markov Models “2016 IEEE 24th International Conference on Network Protocols (ICNP).

[8] Xiaokun Li andFaith M. Porikli, “A Hidden Markov Model Framework For Traffic Event Detection usinf video features“,2004 International Conference on Image Processing (ICIP).

[9] *Evan Tan, Jing Chen, “Vehicular traffic Density Estimation via Statistical Methods with Automated State Learning”* 978-1-4244-1696-7/07/$25.00 ©2007 IEEE

[10] Lin LI, Jianing HU, Qiao HUANG, Jingyan SONG,” A Fuzzy Hidden Markov Model for Traffic Status Classification Based on Video Features” IMACS Multiconference on "Computational Engineering in Systems Applications"(CESA).