Bayesian Network for Traffic Management Application: Estimated the Travel Time

Ahmed Derbel
Unit Research, Applied Economics URECA
Faculty of Economic Sciences and Management of Sfax
Sfax, Tunisia
derbelamd@gmail.com

Younes Boujelbene
Unit Research, Applied Economics URECA
Faculty of Economic Sciences and Management of Sfax
Sfax, Tunisia
boujelbene.younes@yahoo.fr

Abstract— **The notion of travel time is a simple and necessary information for commuters, the intelligent traffic management allows to reduce congestion both temporally and spatially. In this context, this article focuses on the problem of estimating the route time where highway traffic is characterized by several macroscopic variables such as flow, intensity and velocity. We propose a probabilistic model based on Dynamic Bayesian Networks (DBN) which helps us to combine these variables. This model represents a technique for analyzing the traffic to obtain a knowledge model that evolves with time. We estimated, predicted and monitoring travel time given a traffic situation. The travel time is an element that is not directly observable so we used the technique of Hidden Markov Model (HMM) to assure the prediction of travel time in different traffic situations, especially in critical cases where the road traffic is limited.**

Keywords— Bayesian Network; Intelligent Transport System; Hidden Markov Model; Decision Support System; Artificial Intelligence;

I. Introduction

The estimation of travel time is an important issue in the field of road information especially in a situation of saturation or congestion for a path. Knowledge and prediction of traffic helps us to exploit the networks operate and improve the quality of service for transport users, such as providing dynamic information on the state of traffic (is it raining? Is the visibility clear enough? Is there an (Accidents, event)? These are all vital and important information regarding automatic steering and dynamic. The development and implementation of a computerized decision support system has become a major concern for all researchers and managers of the road traffic. The establishment of a system is able to process data and provide advances functions to traffic managers. This system allows users to take appropriate decisions and distribute these information's.

Our goal is to estimate and predict the travel time from a traffic state, we've built a probabilistic model based on Bayesian networks, the travel time is not directly observable, it depends on the traffic conditions, for example if the traffic is flowing we observe a fast route time indication and vice versa. The estimation from deterministic models remains very difficult, for this reason, we use probabilistic models based on the concept of the Hidden Markov Model or HMM for prediction, filtering and smoothing the travel time given a traffic situation.

II. Literature Review

The estimated of travel time is an active area, several techniques emerged and some researchers have estimated travel time from fixed sensors, other from time experienced by vehicles tracers' course. The sensors of magnetic loops were introduced in the field of traffic and theses sensors are the most used in the analysis of traffic [1, 2]. These data are almost complete because they take into consideration all the vehicles that travel the road network.

There are two types of approaches a deterministic approach and stochastic approach:

Deterministic approaches: Deterministic modeling is defined as a function over time, and he explained by the deterministic mathematical models we mentioned for example method of Nam and Drew [3], the method of Oh et al, [4], the method of Bonvalet and Robin BRP [5] and method of Buisson and Lessort [6]. All these models are used as indicators to estimate the travel time. But, they are tainted by measurement risk on the detection and counting,

Stochastic approaches: The random certain fluctuations quantities are often a reflection of a stochastic process. In the literature, there are several approaches that use data from fixed sensors, Rice and Van Zwet [7] observed a linear regression between the future travel time and the current travel time. Methods Hamed [8] and Davis and al, [9] are based on the analysis of chronologies series ARIMA[1]for the prediction of traffic parameters and methods Dailey and Hage [10, 11] are used to estimate the travel time by the filtering method, for example the method kalman filter.

The disadvantages of this approach that it is necessary to find a large number of vehicle test to estimate the travel time, and this approach based on the estimated travel time per serving (section by section) and a time interval with a preliminary study.

In this article, we will attempt to estimate the travel time in a highway from an HMM, we have chosen a traffic condition so we can estimate the travel time, the traffic conditions dependents on several variables, so the DBN is used for the management of uncertainty and to set the best models of knowledge. The data we have are temporal and multi-dimensional, the decision to be taken depends on the current state of traffic, for this reason we used dynamic DBN as an

[1] : Auto Regression Integrated Moving Average

978-1-4799-8172-4/15/$31.00 ©2015 IEEE

analytical technique for knowledge models that evolve over time.

Our work deals with 3 sections:

- Section §3: we defined the technique of Bayesian Network to combine the different macroscopic variables of road traffic.
- Section §4: these variables are temporal and multidimensional; we used the technique of Dynamic Bayesian Network to evolve this data in overtime and to make a proper decision with a simulator.
- Section §5: we defined the Hidden Markov Model and theses inferences, having used the different dynamic programming algorithm.

III. BAYESIAN NETWORK

A graphic pattern is a family of probability distribution defined in terms of graph. It consist a set of nodes and arcs. Nodes generally represent random variables and arcs represent dependencies between variables. A BN is a union between probability theory and graph theory, the use of a Bayesian model allows us to predict, to control, to simulate an observed behavior, to diagnose causes and to take appropriate decisions. Moreover, the application areas are very diverse; we are interested in the intelligent transportation system (ITS), including an application that determines the state of movement and speed float from uncertain data, which are generally derived from sensors. The sensors provide information about the environment for example the speed and intensity of vehicles; we have combined this information to traffic management and to know the real road situations.

A. Basic Concept

A Bayesian network $\beta = \{G, P\}$ is defined by a directed graph, a probability space and a set of random variable. The graph $G=(X, E)$ is a circuit in (X) represents all nodes or vertices and (E) is the set of arcs. The probability space is $P= (\Omega, p)$ and Ω is the universe probability and p represents all the random variables: $X=\{x_1, x_2,\ldots,x_n\}$ such as

$$p(x_1,\ldots,x_n) = \prod_{i=1}^{n} p(x_i / Pa(x_i)) \qquad (1)$$

With $Pa(x_i)$ is the set of all parent nodes X_i in a graph G.

B. Establishment of a Bayesian Network

As we saw earlier in the definition, the complete specification of a BN needs to specify one hand, the structure (directed acyclic graph) and secondly the parameter. The extension of an arc can be analyzed in influence terms of one variable on another. To explain how the BN can be used in traffic, an application will be processed. We define the random variables in our application:

• Visibility: it is to say that the clear field or not of view in driving.

• Event: the possibility of having an accident or event or other event that prevents the movement of vehicles.

• Weather conditions: know the meteorological conditions along the route (weather and / or wet).

• Density (Q): the number N of vehicles passing over a period ΔT at a point x, based on the duration of the period, $Q=N/\Delta T$, is expressed as (veh/h or veh/s).

• Intensity (K): also called density is the number of N cars between x and $x +\Delta X$ at time t, based on track length, $K=N/\Delta X$, is expressed as (veh/km or veh /m).

They are data that comes from the sensors, there are two types of information, static data (data were derived from sensors that are installed in the road such as a magnetic loop) and dynamic data (surveillance cameras). The traffic management is an application of intelligent transportation systems (ITS) which allows to adopt proactive measures in real time and merge the various information to transmit in several scope (message board, media , ..., etc.). For further explain and clarify our probabilistic model “The Road Traffic Management”, we need other random variables:

• Circulation state: it determines the current state of circulation, is flowing through the good or bad circulation conditions.

• Traffic flow (TF): it determines the traffic flow, what is the average vehicles speed present at time t on a given route length.

Three random variables (Visibility, Event, weather conditions) estimated the probability of the circulation state. The density and Intensity estimated the traffic flow probability. The three variables (W, Q, and K) are causes nodes, the circulation state and the traffic flow, are variables of effect. For example, the knowledge that I have on (Visibility, Event, weather conditions) determines the knowledge of the traffic state.

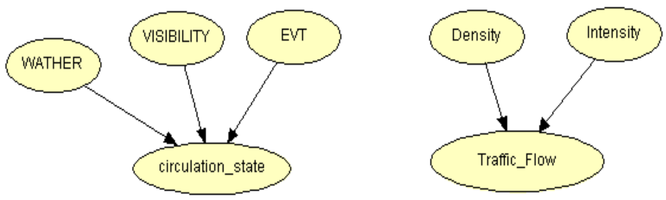


Fig 1: Random variables

BN is used by highlighting some nodes, we calculated the probabilities for all the unknown nodes, and this process is called probabilistic inference. The main use of a BN is to calculate the probabilities given a posteriori observed event, for example the probability of the circulation state is calculated as follows:

$$p(C / E, V, W) = \frac{p(C, E, V, W)}{p(E, V, W)}$$

This method is not very convenient for BN with numerous variables [12] and conditional independence can be exploited to simplify the calculation, and there are many probabilistic algorithms for BN and more details can be found in [13].

To build a complete BN and to the management of road traffic, two random variables appeared as:

• Traffic situation (*TS*): this and a variable that determines the current traffic situation, there are 4 positions {fleet, heavy, blocked, saturated}.

• Individual Speed (*S_IND*): it can adopt as knowing the circulation state and the flow speed. It is a variable that determines the authorized traffic speed. It divided into 4 intervals {[0-50], [50-70], [70-90] and [90-110]; km/h}.

The figure below shows an example the BN of the road traffic management.

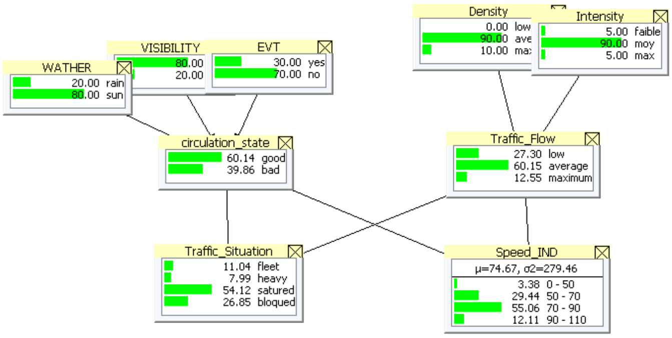
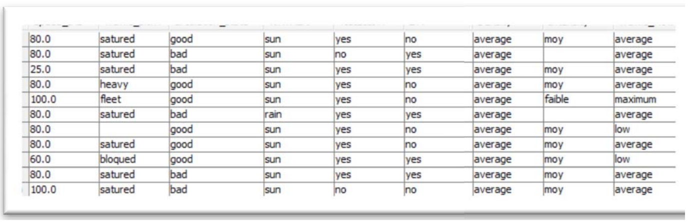


Fig 2: BN of Traffic Management

C. Learning Settings

The parameter learning is to assume that the network structure is fixed and therefore we can determine the conditional probabilities of each variable giving the case when all the variables are observed and discrete. The most commonly used method for estimating the parameters is the statistical estimate the probability of an event in the database, this method is called maximum likelihood [14], and there are other approaches known as maximum a posteriori MAP and experience a posteriori EAP [15], the interests of these two approaches is to improve the statistical estimation. In the case of incomplete data, there are several algorithms to estimate missing data from a database, the most widely approach used is the EM [16] algorithm, it is an iterative algorithm which has been developed by Dempster, to start we initialize the parameters to zero and estimate the missing values from the current settings, then calculate p (x missing / x measured) in the BN and its inferences. For the build the tables of probability, we worked on a database contains a 100 observations for a road traffic (figure 3 below), also we used the EM algorithm to calculate and estimate all the missing information and determined the all probability conditional for our model.



80.0	satured	good	sun	yes	no	average	moy	average
80.0	satured	bad	sun	no	yes	average		average
25.0	satured	bad	sun	yes	yes	average	moy	average
80.0	heavy	good	sun	yes	no	average	moy	average
100.0	fleet	good	sun	yes	no	average	faible	maximum
80.0	satured	bad	rain	yes	yes	average		average
80.0		good	sun	yes	no	average	moy	low
80.0	satured	good	sun	yes	no	average	moy	average
60.0	bloqued	good	sun	yes	yes	average	moy	low
80.0	satured	bad	sun	yes	yes	average	moy	average
100.0	satured	bad	sun	no	no	average	moy	average

Fig 3: Database with incomplete information

IV. DYNAMIC BAYESIAN NETWORKS

A Dynamic Bayesian Networks (DBN) is an extension of a static BN used to represent the evolution of random variables according to a discrete sequence. The term dynamic is characterized only by the system and not the network that does not change over time. The DBN extend the representation of BN dynamic process, it contains T random vectors of variables $X[t]=X_1[t],\ldots,Xn[t]$.

$$X[t] = X_1[t],\ldots,Xn[T] = \prod_{i=1}^{T}\prod_{i=1}^{n} p\left(X_i[t]/Pa(X_i[t]\right) \quad (2)$$

The environment temporal dynamic model is the transition of a portion time to the other and they are connected by arc. The principle is based on the stationary principle that is to say, the structure and parameters of the DBN are repeated and the JPD is encoded by the initial states with the transmissions states.

$$P(X_1[t],\ldots,Xn[T]) = P_1(X_1[t])\prod_{t=2}^{T-1} P->(X_i[t]/X_i[t-1]) \quad (3)$$

The objective of our DBN is to predict the authorized traffic speed and traffic conditions, by calculating the probability of a future event (*C_ Expected*, *F_ Expected*) given observations (*Circulation state*, *traffic flow*) in the past and over time. So the role of DBN is used to predict the evolution of the traffic indicators in the near future to *t*+1, *t*+2 …,*t*+*h* where *h* is the prediction of horizon.

The figure below illustrates the construction of a DBN with 3 instances of time (*t*-1, *t*, *t*+1).

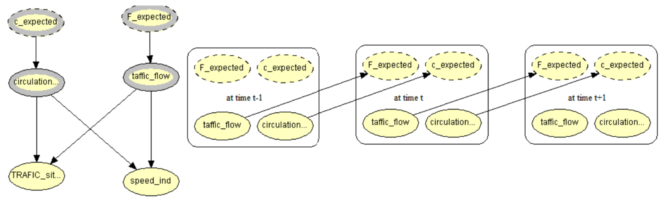


Fig 4: DBN with 3 instances of time

The DBN represent a tool for decision support, it can calculate the probability of states based on the knowledge available, we used the software Huguin [17] as an intuitive graphical interface to simulate dynamic situations and this software provides a set of tools for a constriction IDSS [2] based on models characterized by uncertainty.

The table: 1 below illustrates the results of evolution the DBN in overtime using maximum likelihood criteria decision.

Tableau 1: Results of DBN Simulations

Time (*t*-1)	Time (*t*)	Time (*t*+1)
TS=saturated	*TS =blocked*	*TS =blocked*
S_IND =50-70	*S_IND =0-50*	*S_IND =0-50*
TF=average	*TF=low*	*TF=low*

The results of the dynamic simulation indicate, that the traffic state have changed their behavior, it spent a situation of saturation a time (*t*-1) to a blocking situation at time (*t*, *t* + 1),

[2] Interactive Decision Support System

that is i.e., the road traffic has attempted a road congestion state, this finding is justified by the speed and the traffic flow that becomes weak over time. From this result, we tried to continue the journey time knowing the different situations of traffic, for this reason, we used a probabilistic model based on Hidden Markov Model for the prediction, filtering and smoothing with a travel time given a traffic situation

V. Hidden Markov Model

The DBN is extensions of Bayesian Network representing the temporal and spatial evolution. The Hidden Markov Model or HMM is a special case of the DBN and it is a statistical Markov model, in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. The state is not directly visible to the observer, and the state transition probabilities are the only parameters.

The technical of Hidden Markov models are used extensively including in the divers fields for example; Pattern recognition, artificial intelligence or automatic natural language processing and speech recognition.

Our application is based on tracking the Travel Time (TT: hidden state) giving the Traffic Situation conditions (TS: observations state).

The diagram below shows the architecture of an instantiated HMM. Each oval shape represents a random variable with the $(TT_t)_{1\le t\le T}$ modeling hidden states. The random variable $(TS_t)_{1\le t\le T}$ are represent the observations. The arrows in the diagram denote the conditional dependencies.

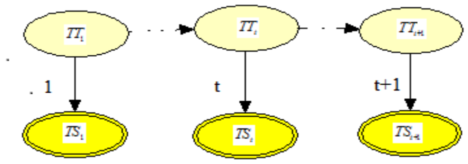


Fig 5: HMM architecture

This model has undergone for three essential elements:

- First, it is assumed, that the dynamic change is caused by a Markov process, i.e. the current state depends only on a finite previous states numbers. For example, if observed TT_{t-1}, TT_t is independent of the other wholes random variables.
- Second, it is assumed, that the observation is generated only by the current state, for example knowing the hidden state TT_t, the distribution of TS_t is independent for all random variables $\{TT_{t-2}\dots,TT_1\}$ and $\{TS_{t-2}\dots,TS_1\}$.
- Finally, it is also assumed, the dynamic change is caused by a stationary process; i.e. the distribution of probability do not change over time, in other words, all $p(TT_t / Pa(TT_t))$ are the same for all values t.

Using the HMM technique, we can compute the joint probabilities of all states. This joint probability is generated by a hidden sequence with an observed sequence

$$p(TS_{1:T},TT_{1:T}) = p(TT_1)p(TS_1/TT_1)\prod_{t=2}^{T} p(TT_t/TT_{t-1})p(TS_t/TT_t) \quad (4)$$

Also, we can compute the marginal probability for an observation sequence

$$p(TS_{1:T}) = \sum_{h1:T} p(TT_{1:T}=t_{1:T})\,p(TS_{1:T}/TT_{1:T}=t_{1:T}) \quad (5)$$

The same sequence observed can be produced by several hidden sequences, so there are an exponential number of hidden sequences possible, if we take the example 1 hidden state takes 2 values, so I'll have 2^T possible sequences for the hidden state, this is a very naive and inefficient calculation. We used the concept of dynamic programming to better exploit the structure of an HMM.
The idea is to break the calculation by structured intermediate calculations in a table.
So we define a table **α** with *i* represent the tidy (possible state) and *t* represent the column (evolution of the time).
The initial values by definition will be calculated as follows:

$$\alpha(i,1) = p(TS_1=s_1/TT_1=i)\sum_j p(TS_1=s_1/TT_1=i)\,p(TT_1=i)$$

We also define a recursion:

$$\alpha(i,t+1) = p(TS_{t+1}=s_{t+1}/TT_{t+1}=i)\sum_j p(TT_{t+1}=i/TT_t=j).\alpha(j,t)$$

Where, *j* is the possible state of hidden values.

Once the table is treated α, we easily obtain the marginal probability of an observed sequence:

$$p(TS_{1:T}=s_{1:T}) = \sum_j \alpha(j,t)$$

A. Inference

In our case, the result of the DBN simulation provides us two observations; a situation of saturation at time (*t*-1) and two blocking situations corresponding to the congestion at the instant (*t*) and (*t* + 1) (section §4; table 1). For calculate the inferences we must have 3 matrices distributions. First, we sampled an initial value for the hidden state Travel time is an initial distribution matrix (Table 4). Then we set the probability of the hidden state power knowing the previous hidden state, a transition matrix (Table 3). Finally, we have also defined the probabilities of the observations given the hidden state, which is an observation matrix (Table 2).

We decoded the traffic state, TS_1=*saturated*, TS_2= *blocked*, TS_3 =*blocked* with a travel time, it depends with 2 possible situations, TT_t= {saturated, blocked}.
We proposed the following probability distributions:

Tableau 2: Observation Model

	TT=saturated	*TT=blocked*
p(TS_1= saturated /TT)	0,9	0,1
p(TS_1= blocked /TT)	0,1	0,9

Tableau 3: Transition Model

	TT_t=saturated	TT_{t-1}= blocked
$p(TT_1$=saturated/$TT_{t-1})$	0,3	0,2
$p(TT_1$= blocked /$TT_{t-1})$	0,7	0,8

Tableau 4: Initial Distribution

	TT_1= saturated	TT_1= blocked
$p(TT_1)$	0,5	0,5

For this model, several learning techniques and inferences have been proposed, including the types of inference we defined:

Filtering: The task is to compute, given the model's parameters and a sequence of observations, the distribution over hidden states of the last latent variable at the end of the sequence, i.e. the posterior distribution of the latest hidden variable $p(TT_t /TS_1, TS_2, TS_3)$ to know the latest travel time from a traffic condition. This problem can be handled efficiently using the forward algorithm[18].

$$p(TT_i = i / TS_{1:T} = s_{1:T}) = \frac{p(TT_i = i, TS_{1:T} = s_{1:T})}{\sum_i p(TT_i = i, TS_{1:T} = s_{1:T})}$$

Smoothing: This is similar to filtering but asks about the distribution of a latent variable somewhere in the middle of a sequence, i.e. to compute the posterior distribution on a past state $p(TT_{1:N} /TS_1, TS_2, TS_3)$ for some $N<t$. That it is to say, the probability of the travel time from a traffic condition in the past. The forward-backward algorithm [18] is an efficient method for computing the smoothed values for all hidden state variables.

We calculate the posterior distribution on a past state, the calculation of $\alpha(i,t)$ gives a scanning from left to right, you can do the same thing but from right to left we define a new table

$$\beta(i,t) = p(TS_{t+1:T} = s_{t+1:T} / TT_t = i)$$

The initial states $\beta(i,t)$ =1

We defined a recursion:

$$\beta(i,t-1) = \sum_j p(TS_t = s_t / TT_t = j) p(TT_t = j / TT_{t-1} = i)$$

Once the table is calculated β is easily obtained the marginal probability:

$$p(TS_{1:T} = s_{1:T}) = \sum_j \beta(j,1) p(TS_1 = s_1 / TT_1 = j) p(TT_1 = j)$$

The Table $\alpha(i,t)$ and $\beta(i,t)$can also be used for smoothing in any hidden state for any moment.

$$p(TT_k = i / TS_{1:T} = s_{1:T}) = \alpha(i,1)\beta(i,t) / \Upsilon \quad ,$$

With is a normalization factor that is corresponds to a sum of i smoothing the results at time t=2 and t=1.

Most likely explanation: The task, unlike the previous two, asks about the joint probability of the entire sequence of hidden states that generated a particular observations sequence.

This task requires finding a maximum over all possible state sequences, and can be solved efficiently by the Viterbi algorithm [18].

We define the initial values :

$$\alpha^*(i,1) = p(TS_1 = s_1 / TT_1 = i) p(TT_1 = i)$$

The recursion is:

$$\alpha^*(i,t+1) = p(TS_{t+1} = s_{t+1} / TT_{t+1} = i) \max_j p(TT_{t+1} = i / TT_t = j) \alpha^*(j,t)$$

Prediction: We calculate the posterior distribution of a future state $p(TT_{t+k} / TS_{1:t})$ with $k \geq 0$. In other words, we determined the likelihood of travel time in k days.

The table $\alpha(i,1)$ can be used to infer the distribution of the prediction, we define a new table

$$\Pi(i,k) = p(TT_{t+k} = i / TS_{1:t} = s_{1:t})$$

The initial values are calculated as follows:

$$\Pi(i,0) = \frac{\alpha(i,t)}{\sum_j \alpha(j,t)}$$

The recursion is:

$$\Pi(i,k+1) = \sum_j p(TT_{t+k+1} = i / TT_{t+k} = j) \Pi(j,k)$$

We can also make the prediction of TS_{t+k}

$$p(TS_{t+k} = s / TS_{1:t} = s_{1:t}) = \sum_j p(TS_{t+k} = s) / TT_{t+k} = j) \Pi(j,k)$$

VI. Results and Discussion

The table 5 and the table 6, illustrates the results of calculating the smoothing and filtering. The probabilities filtering calculations show that the current TT_3 hidden state given all observations sequences is a situation of saturation, since $p(TT_3$=saturated/TS_1= saturated, TS_2=blocked, TS_3=blocked) has a higher probability compared with

$p(TT_3$=blocked/TS_1=saturated, TS_2= blocked, TS_3= blocked)

Thus, the smoothing results show that TT_1 and TT_2 are saturation situations, since they have a higher probability compared to the blocking situation.

Tableau 5: Table α(i,t) for Filtering

i/t	1	2	3
Saturated	0,45	0,13	0,04
Blocked	0,05	0,03	0,01

Tableau 6: Table β(i,t) for Smoothing

i/t	1	2	3
Saturated	0,48	0,66	1
Blocked	0,54	0,74	1

The table 7 illustrates the calculation of the prediction probabilities. These results show that the states observed traffic and hidden travel time states are blocking situations. This means in the future. The traffic will become a critical situation, what we call the phenomenon of congestion; this observation is made by an increase in transport demand with slower flow rate. This may be due to an increase in traffic or to a decrease in channel capacity. This finally resulted in the complete cessation of circulation.

Tableau 7: Table π(i,k) for Prediction

i/k	0	1	2
Saturated	0,77	0,27	0,22
Blocked	0,23	0,73	0,78

The notion of travel time is vital information for policy maker's transportation. It is also a necessary element in choosing the means of transport. We can estimate this information to determine the shortest way according to character speed of a trip.

VII. Conclusion And Future Work

In this work, we proposed an approach that helps to estimate the travel time from a traffic state, mainly in the state of congestion. We used the BN to manage uncertainty and secure the best knowledge models by combining the macroscopic random variables of road traffic. We have also developed a probabilistic model based on Hidden Markov Model for the prediction, filtering and smoothing a travel time given a traffic situation. We will test our model with a software simulation to determine the strong and the weak point, also, we can compare with other similar method to improve the effectiveness.

The perspective as we plan to implement other estimation models and adopt other approaches to improve the dynamic learning.

References

[1] M. Papageorgiou, J-M. Blosseville and H. Haj-Sale, "Dynamic modeling, assignment, and route guidance in traffic networks", Transportation Research part B, vol. 24B, n° 6, pp. 471-495, 1990.

[2] D. Middleton, D. Gopalakrishna and M. Raman, "Advances in traffic data collection and management", Technical Report BAT- 02-006, Texas Transportation Institute, Cambridge systematics, Inc., Washngton, DC, December, 2002.

[3] D.H. Nam and D.R Drew, "Automatic measurement of traffic variables for intelligent transportation systems applications", in Transportation Research Part B: Methodological, n° 33, pp. 437- 457, 1999.

[4] J.S. Oh, R. Jayakrishan and W. Rechker, " Section travel time estimation from point detection data", in proceedings of the 82th Annual Meeting of Transportation Research Board, Washington, D.C., U.S.A, 2003.

[5] F. Bonvalet and Y. Robin-Prévallée, "Mise au point d'un indicateur permanent des conditions de circulation en Île-de- France", Transport Environnement Circulation, pp. 84-85, 1987.

[6] C. Buisson and JB. Lessort, "Comprendre le trafic routier, méthodes et calculs", Edition Certu, 2010.

[7] J. Rice and E. Van Zwet,"A simple and effective method for predicting travel times on freeways", IEEE, Transactions on Intelligent Transportation Systems, n° 5, pp. 200-207, 2004.

[8] M.M. Hamed, H.R. Al-Masaeid and Z.M. Bani Said., "Short-term prediction of traffic volume in urban arterials", Journal of Transportation Engineering, pp. 249-254, 1995.

[9] G.A. Davis, N.L. Nihan, M.M. Hamed and L.N. Jacobson, "Adaptive forecasting of freeway traffic congestion", Transportation Research Record, 1287, pp. 29-33, 1990.

[10] D.J. Dailey, "A statistical algorithm for estimating speed from single loop volume and occupancy measurements", in Transportation Research Part B: Methodological, 33, pp. 313-322, 1999.

[11] RM. Hage, D. Bétaille, F. Peyret and D. Meizel, "Unscented Kalman filter for urban network travel time estimation", 15th meeting of the EURO Working Group on Transportation, Paris, Septembe 2012.

[12] D. Heckerman, "A tutorial on learning with Bayesian Networks", Technical report, Microsoft Research, Advanced Technology Devision: Microsoft Corporation, Redmond, US, 1995.

[13] S. K. Andersen, K. G. Olesen, F. V. Jensen and F. Jensen, "Hugin – a shell for building Bayesian belief universes for expert systems", In Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, pages 1080-1085, Detroit, Michigan, Aug. 20-25, 1989.

[14] J. Aldrich, "R.A. Fisher and the making of maximum likelihood 1912-1922", Statistical Science, vol. 12, n° 3, 1997, p. 162-176 .

[15] A. Dawid, "Applications of a general propagation algorithm for probabilistic expert systems Statistics and Computing", 2:25–36, 1992.

[16] A. Dempster, N. Laird and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", Journal of the Royal Statistical Society, B 39: 1-38, 1977.

[17] P. Spirtes, C. Glymour and R. Scheines, " Causation, Prediction, and Search", *MIT Press*, Adaptive Computation and Machine Learning, second edition, 2000.

[18] BH. Juang. and L. Rabiner, " An introduction to hidden Markov models", ASSP Magazine IEEE, Volume: 3, Issue: 1, 4 – 16, 1986.