2018 IEEE Intelligent Vehicles Symposium (IV)
Changshu, Suzhou, China, June 26-30, 2018

Modeling and Predicting Vehicle Motion Activities by Using And-Or Graph

Shuofeng Wang, Li Li, *Fellow, IEEE*, Nan-Ning Zheng, *Fellow, IEEE*, and Dongpu Cao

***Abstract*—The ability of modeling and predicting vehicle motion activities is important for automated vehicles. In this paper, we propose an And-Or Graph based model to give a simple and clear description of motion activities. Compared to other models, this new model relaxes the Markov property requirement in transition between activities and is thus more flexible. The parameters of this model can be easily learned from data. Using the trained new model, we can predict the on-going motion activity label and its corresponding probability. Experiments show that a high prediction accuracy (97%) can be achieved by this new model.**

***Index Terms*—And-Or Graph, vehicle motion activities**

I. INTRODUCTION

Modeling and predicting vehicle motion activities is one of the key function of automated vehicles [1] [2]. For example, to make an overtaking requires a human driver to correctly recognize and predict the positions and movements of neighboring vehicles, so as to avoid a collision. This requires an online algorithm to describe the activities of vehicles.

As pointed in [3], existing studies can be roughly categorized into three kinds: Physics-based models, Maneuver-based models and Interaction-aware models.

Physics-based models apply a certain simple physical law to describe the intention level of collision avoidance for each vehicle [4]. The short-term prediction of future motions for vehicles can be calculated by simulating the given evolution models for a certain time interval. Physics-based models have the highest degree of abstraction and the lowest degree of calculation. However, they are sometimes oversimplified to fully describe some complex interactions between vehicles and are unreliable for long-term prediction.

Maneuver-based models further consider that the future motion of a vehicle with respect to the maneuver that the driver intends to perform [5]. Each driver will make a series of maneuvers executed independently from other drivers. Once we identify the maneuver intention of a driver, the future motion of the corresponding vehicle will strictly match this maneuver. Such assumptions sometimes fail, since drivers may adapt their maneuvers from time to time, due to the influence of other drivers.

Interaction-aware motion models further consider the interactions between every two vehicles so as to provide situation-aware modeling and a more reliable prediction of vehicle motions [6]. However, there are few interaction-aware motion models in the literature, because of the following difficulties in building such models.

First, we need to establish more realistic models that fit with empirical observations. Many models are artificially designed and thus may contain some bias. Along with the development of sensing (e.g. vision, radar, laser scanner, etc.) technologies, we had accumulated "Big Data" recording how human drivers finished driving tasks. It is expected to design an appropriate model that directly accepts these data so as to learn how human drivers interacts with each other.

Second, we need to bring more flexibility into the model. Some existing models assumed that the series of maneuvers can be characterized via Markov processes [7]–[9]. However, this assumption may not always hold in practice. How to relax this assumption received increasing interests.

Third, we need to enable the model to easily self-explain its recognition and prediction results. Since the vehicle is moving forward, we should be able to keep an online updating of the recognition and prediction results.

In order to solve these important problems, we proposed an And-Or graph (AOG) [10] based model to give a simple and a clear description of motion activities.

AOG is a stochastic hierarchical graph model. AOG has been used to deal with several computer vision tasks such as online object tracking [11], inferring human pose [12], and modeling human-object interactions [13]. Previous studies show that And-Or graph is powerful to represent a large number of part object configurations and to represent the spatio-temporal relations. The vehicle motion activity is always spatio-temporal related.

The design of our model has addressed the vehicle-vehicle interactions and the non-Markov property in transition between activities. The parameters of this model can be easily learned from data. Using the trained new model, we can predict the on-going motion activity label and its corresponding probability. Experiments show that a high prediction accuracy (97%) can be achieved by this new model.

Fig. 1 summarizes the data processing procedure, the encountered difficulties and the corresponding solving algorithm studied in this paper.

This work was supported in part by the National Natural Science Foundation of China under Grant 61533019, the Beijing Municipal Science and Technology Commission Program under Grant D171100000317002, the Beijing Municipal Commission of Transport Program under Grant ZC179074Z. (*Corresponding author: Li Li.*)

S. Wang is with Department of Automation, Tsinghua University, Beijing 100084, China.

L. Li is with Department of Automation, BNRist, Tsinghua University, Beijing, China 100084. (Email: li-li@tsinghua.edu.cn)

N.-N. Zheng is with Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China.

D. P. Cao is with the Mechanical and Mechatronics Engineering Department at the University of Waterloo, N2L 3G1, Canada.

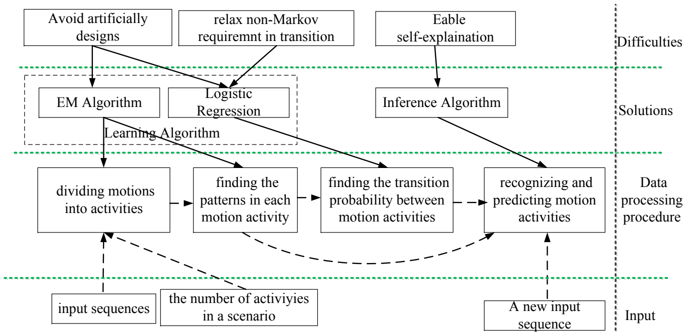


Fig. 1. The flowchart of our model

To give a better explanation of our new finding, the rest of this paper is organized as follows. The relative notations used in the paper are provided in Tab, I. *Section II* briefly formulates the problem studied in this paper. *Section III* introduces the structure of our model. In *Section IV*, we introduce the approaches for learning the parameters of our model. In *Section V*, we introduce the approaches for modeling and predicting vehicle motion activities based on our model. Experiment results are shown in *Section VI*. Finally, we present our conclusions in *Section VII*.

II. PROBLEM FORMULATION

In this section, we describe the input data and tasks in the modeling and predicting problems of vehicle motion activities.

A. Input

The input of this problem are n sampled sequences of vehicle motion states and vehicle-vehicle interactions collected in a driving scenario E, and a pre-selected number K_E. The input sequences are denoted as $\{\boldsymbol{q}_i | i = 1, 2, 3, ...n\}$. Each vector $\boldsymbol{q}_i = [q_{i1}, q_{i2}, ...q_{i\tau_i}]$ consists of element q_{it} as the vehicle motion states and vehicle-vehicle interactions at time t, and τ_i is the sequence length for $\boldsymbol{q}_i$. K_E is the number of motion activities in driving scenario E.

In this paper, the vehicle motion states refer to lateral velocity, longitudinal velocity, lateral acceleration and longitudinal acceleration of the vehicle. In this paper, the vehicle-vehicle interactions refer to the relative lateral velocity, longitudinal velocity, lateral acceleration and longitudinal acceleration to surrounding vehicles. For example, in a lane-changing scenario, the vehicle-vehicle interactions refers to the longitudinal velocity, longitudinal velocity, lateral acceleration and longitudinal acceleration relative to the preceding vehicle in the same lane.

The input sequences can be obtained from various sensors and transmitted by Vehicle-to-Vehicle (V2V) or Vehicle-to-Road (V2R) communication in practice [14].

B. Tasks

Given the input sequence, we have three tasks as follows.

TABLE I
NOTATION

Symbols	Meaning
t	time
M_t	vehicle motion state at time t, such as the lateral velocity, longitudinal velocity, lateral acceleration and longitudinal acceleration
I_t	vehicle-vehicle interaction at time t, such as the lateral velocity, longitudinal velocity, lateral acceleration and longitudinal acceleration relative to the surrounding vehicles
$\boldsymbol{q}_i$	$\boldsymbol{q}_i = (q_{i1}, q_{i2}, ...)$ the ith sequence of input data
q_{it}	the vehicle motion state and vehicle-vehicle interaction at time t in $\boldsymbol{q}_i$
S_t	$S_t = (M_t, I_t)$ is the vehicle motion state and vehicle-vehicle interaction at time t
$\boldsymbol{S}$	$\boldsymbol{S} = (S_1, S_2, ...)$ is a sequence of vehicle motion states and vehicle-vehicle interactions to be interpreted by AOG
E	a driving scenario
K_E	the number of motion activities in scenario E
A_k	the kth motion activity in a scenario
Ω_E	$\Omega_E = \{A_1, A_2, ..., A_{K_E}\}$ is the set of motion activities in scenario E
l_t	the vehicle motion activity label for S_t
$\boldsymbol{L}$	$\boldsymbol{L} = (l_1, l_2, ...)$ is a sequence of motion activity labels interpreted for $\boldsymbol{S}$
$p(X)$	the probability of X
$p(X\|Y)$	the probability of X given Y
$\mu^M_{A_k}$, $\Sigma^M_{A_k}$	the mean and covariance of vehicle motion states for motion activity A_k
$\mu^I_{A_k}$, $\Sigma^I_{A_k}$	the mean of vehicle-vehicle interactions for motion activity A_k
d_k	the duration of activity A_k
z_k	a boolean number denoting whether the transition from activity A_k to activity A_{k+1} is happened
α_k, β_k	the parameters in logistic sigmoid function for motion activity A_k
R_{ik}	a latent variable denoting the transition moment from activity A_k to activity A_{k-1} for q_i
$\boldsymbol{R}_i$	$\boldsymbol{R}_i = (R_{i1}, R_{i2}, ...R_{iK_E})$ is a latent vector which denotes the activity transition moments for $\boldsymbol{q}_i$
G^m_t	the mth AOG candidate given $S_1, S_2, ...S_t$
pG^m_t	$pG^m_t = p(G^m_t \| S_1, S_2, ...S_t)$ is the posterior probability of the AOG candidate G^m_t given $S_1, S_2, ...S_t$
G^*_t	the optimal AOG candidate given $S_1, S_2, ...S_t$

1) Dividing Driving Scenario into Activities: As shown in [15], a vehicle needs to finish a series of driving-tasks successively to pass a driving scenario. Since an individual driving-task may exist in lots of driving scenarios, the analysis of an individual driving-task can help understand different driving scenarios. So, researchers are interested in studying how to divide the motions in driving scenario E into K_E activities in a natural way.

We denote the motion activities in driving scenario E as Ω_E. $\Omega_E = \{A_1, A_2, ..., A_{K_E}\}$, where $A_1, A_2, ..., A_{K_E}$ are the motion activity labels in driving scenario E. For example, in lane-changing scenario, the motions of vehicle can be divided into 3 motion activities called “following in the current lane”, “approaching the objective lane” and “following in the objective lane”; see Fig.2.

Finishing this task, we will get a model which is able to determine the activity transition moments along the input sequence of vehicle motion states. The detailed model and algorithm to finish this task will be presented in *Section IV*.

2) Recognizing the Motion Activities: Given a new sequence of vehicle motion states and vehicle-vehicle interactions denoted as $\boldsymbol{S} = (S_1, S_2, ...S_\tau)$, we need to give a sequence of motion activity labels to describe $\boldsymbol{S}$. $S_t = (M_t, I_t)$ where M_t is the vehicle motion state and I_t is the vehicle-vehicle interaction at time t.

Finishing this task, we will get a sequence of motion activity labels denoted as $\boldsymbol{L}$. $\boldsymbol{L} = (l_1, l_2, ..., l_\tau)$, in which element $l_t \in \Omega_E$ labels the vehicle motion activity at time t. The detailed model and algorithm to finish this task will be presented in *Section V*.

3) Predicting the Motion Activity: Given a new sequence of vehicle motion states and vehicle-vehicle interactions denoted as $\boldsymbol{S} = [S_1, S_2, ...S_\tau]$, we need to predict the motion activity that the vehicle may take at time $\tau + 1$.

To make the prediction can be self-explained, not only the motion activity label $l_{\tau+1}$ but also the the corresponding probability should be output. The detailed model and algorithm to finish this task will be presented in *Section V*.

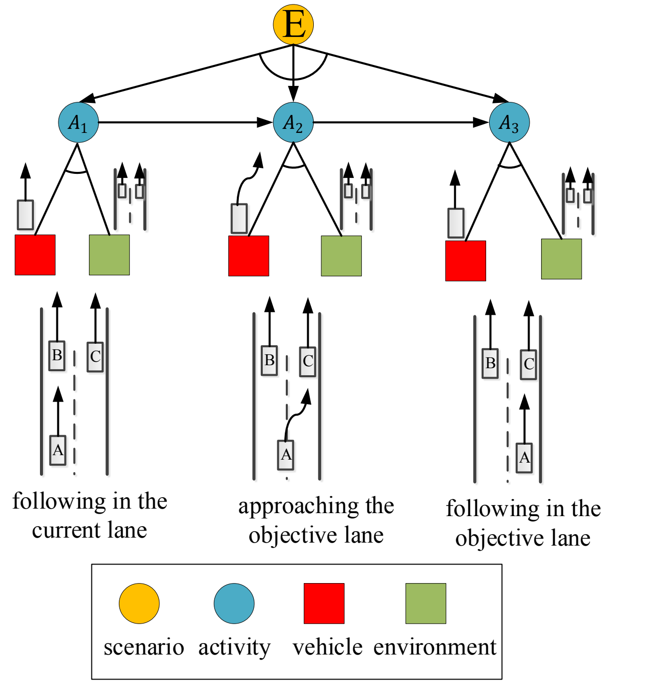


Fig. 2. Illustration of the And-or Graph in lane-changing scenario.

III. Structure of Our Model

A. Graphical Structure of Our Model

Our model is based on And-Or Graph [13]. Let $\boldsymbol{S}$ be a sample in $\{\boldsymbol{q}_i | i = 1, 2, 3, ...n\}$, $\boldsymbol{q}_i = \{q_{i1}, q_{i2}, ...q_{i\tau_i}\}$. $\boldsymbol{S} = (S_1, S_2, ..., S_\tau)$, where $S_t = (M_t, I_t)$. τ is the sequence length of $\boldsymbol{S}$. M_t is the vehicle motion state at time t. I_t is the vehicle-vehicle interaction at time t. $\boldsymbol{S}$ is represented by a hierarchical And-Or graph $G =< E, L >$, where $\boldsymbol{L} = (l_1, l_2, ..., l_\tau)$ is a sequence of vehicle motion activity labels. $l_t \in \Omega_E$ is the vehicle motion activity label at time t. $\Omega_E = \{A_1, A_2, ..., A_{K_E}\}$ is the activity labels set of scenario E.

We formulate driving scenario E and motion activities $A_1, A_2, ..., A_{K_E}$ as And-Nodes. An And-Node represents decomposition operation (decomposing the complex driving scenario into several activities, decomposing the activities into vehicle motion states and vehicle-vehicle interactions).

We formulate vehicle motion states and vehicle-vehicle interactions as Or-Nodes. An Or-Node represents the diversity of vehicles motion states (such as velocity, acceleration, and steering angle) and vehicle-vehicle interactions (such as vehicles' relative speed and acceleration to surrounding vehicles).

Fig. 2 gives an illustration of And-or Graph in lane-changing scenario, in which motions in the lane-changing scenario are divided into 3 motion activities. Correspondingly, we have three And nodes that denote for each motion activities, respectively.

B. Probabilistic Structure of Our Model

Given an $\boldsymbol{S}$, the posterior probability of a hierarchical AOG is defined as:

$$\begin{aligned} p(G|V) &\propto p(V|G) \\ &= p(S_1, S_2, ..., S_\tau | l_1, l_2, ..., l_\tau) \\ &= \prod_{t=1}^{t=\tau} p(S_t|l_t) p(l_t | l_1, l_2, ...l_{t-1}) \end{aligned} \tag{1}$$

where $p(S_t|l_t)$ indicates the likelihood of vehicle motion state and vehicle-vehicle interaction. We assume that the vehicle motion state is independent with the vehicle-vehicle interaction given vehicle motion activity. Thus,

$$p(S_t|l_t) = p(M_t|l_t) p(I_t|l_t) \tag{2}$$

Both $p(M_t|l_t)$ and $p(I_t|l_t)$ are assumed to follow a multi-dimensional Gaussian distribution for each vehicle motion activity. $p(M_t|l_t)$ and $p(I_t|l_t)$ can be assumed to follow other distributions. Here, we choose multi-dimensional Gaussian distribution for computational convenience in learning distribution parameters (This will be shown in *Section IV*). Tests show that such assumption is able to describe the driving scenarios observed in practice. Thus,

$$p(M_t|l_t) = N(M_t; \mu^M_{l_t}, \Sigma^M_{l_t}) \tag{3}$$

$$p(I_t|l_t) = N(I_t; \mu^I_{l_t}, \Sigma^I_{l_t}) \tag{4}$$

$\mu^M_{l_t}$, $\Sigma^M_{l_t}$, $\mu^I_{l_t}$ and $\Sigma^I_{l_t}$ are respectively the mean and the covariance. We will describe the learning methods for these parameters in *Section IV*.

$p(l_t|l_1, l_2, ...l_{t-1})$ indicates the transition probability between vehicle motion activities. Let A_{k-1} and A_k be two consecutive vehicle motion activities in a driving scenario. The transition probability from A_{k-1} to A_k depends on the duration of A_{k-1}. Therefore, the transition probability does not have the Markov property. We denote the duration of A_{k-1} as d_{k-1}. The larger d_{k-1} is, the larger $p(l_t = A_k|l_{t-1} = A_{k-1}, d_{k-1})$ is. The transition probability is assumed to follow a logistic sigmoid function as shown in Equation. (5). The transition probability can be assumed to follow other bounded functions. Here, we choose the logistic sigmoid function for two reasons. First, the logistic sigmoid function is powerful to depict the nonlinear decision boundaries. Therefore, we think it is powerful to give a boundary depicting when the transition will happen. Second, the learning of parameters in logistic sigmoid function can be formulated to a Logistic Regression problem which is easy to solve.

$$
\begin{aligned}
p(l_t = A_k|l_{t-1} = A_{k-1}, d_{k-1}) &= \frac{\alpha_{k-1}d_{k-1} + \beta_{k-1}}{1 + exp(\alpha_{k-1}d_{k-1} + \beta_{k-1})} \\
p(l_t = A_{k-1}|l_{t-1} = A_{k-1}, d_{k-1}) &= \frac{1}{1 + exp(\alpha_{k-1}d_{k-1} + \beta_{k-1})}
\end{aligned} \tag{5}
$$

α_{k-1} and β_{k-1} are parameters in logistic sigmoid function. We will describe the learning methods for these parameters in *Section IV*.

IV. Learning of Our Model

In this section, we describe how to learn model parameters from data. First, we describe the EM algorithm for learning the mean and covariance in multi-dimensional Gaussian distributions mentioned in Equation (3)-(4). Specially, the latent vector $\boldsymbol{R}_i$ introduced in the EM algorithm denotes the activity transition moments for $\boldsymbol{q}_i$. Second, we explain how to manipulate parameters according to Equation (5).

A. Learning Patterns for Motion Activities

As we have assumed that both the likelihood of vehicle motion state and vehicle-vehicle interaction in a motion activity follow a multi-dimensional Gaussian distribution, learning patterns for motion activities is equivalent to learn the mean and covariance in multi-dimensional Gaussian distributions mentioned in Equation (3)-(4).

To learn these parameters, we first introduce a latent vector $\boldsymbol{R}_i = (R_{i1}, R_{i2}, ...R_{iK_E})$ which denotes the activity transition moments for $\boldsymbol{q}_i$, and then an EM algorithm is implemented. K_E is a pre-selected number which denotes the number of activities in the scenario E. R_{ik} denotes the transition moment from activity A_k to activity A_{k-1} for $\boldsymbol{q}_i$. $R_{ik} = \tau_i$ where τ_i is the sequence length of $\boldsymbol{q}_i$.

An illustration of transition moment is shown in Fig. 3. In Fig. 3, the scenario contains 3 activities. The input data in the time interval $[1, R_{i1}]$, $[R_{i1} + 1, R_{i2}]$, $[R_{i2} + 1, R_{i3}]$ should be respectively assigned to activity A_1, A_2 and A_3.

In the E-Step of EM algorithm, the latent variable $\{R_i|i = 1, 2, 3, ...n\}$ is estimated based on current mean and covariance, using the IAP method proposed by [13]. In the M-Step of EM algorithm, the mean and covariance for each motion activities are estimated based on the current latent variable $\{R_i|i = 1, 2, 3, ...n\}$.

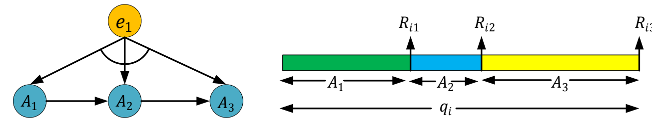


Fig. 3. An illustration of transition moment

B. Learning Temporal Relations for Transition Probability

According to Equation (5), learning temporal relations for transition probability is equivalent to learn the parameters in logistic sigmoid function mentioned in Equation (5).

To this end, we introduce a boolean z_k to denote whether the transition from activity A_k to activity A_{k+1} happens, and then formulate the learning for these parameters into a Logistic Regression problem.

Since we have obtained the transition moment $\{\boldsymbol{R}_i|i = 1, 2, 3, ...n\}$, we can compute the Duration-Transition pair samples (d_k, z_k), where $k \in [1, K_E - 1]$, d_k is the duration of A_k. For the input sequence q_i, the transition from activity A_1 to activity A_2 would not happen, until $d_1 = R_{i1}$, and the transition from activity A_2 to activity A_3 would not happen, until $d_2 = R_{i2} - R_{i1}$, and so on.

According to Equation (5), we have

$$
\begin{aligned}
p(z_{k-1} = 1|d_{k-1}) &= \frac{exp(\alpha_{k-1}d_{k-1} + \beta)}{1 + exp(\alpha_{k-1}d_{k-1} + \beta)} \\
p(z_{k-1} = 0|d_{k-1}) &= \frac{1}{1 + exp(\alpha_{k-1}d_{k-1} + \beta)}
\end{aligned} \tag{6}
$$

Thus the learning of parameters α_{k-1} and β_{k-1} can be formulated into a Logistic Regression (LR) problem. The objective function of LR problem is convex and can be solved by gradient-based methods.

V. Inference of Our Model

In this section, we describe the approach for recognizing and predicting the vehicle's motion activities given $\boldsymbol{S}$, where $\boldsymbol{S} = (S_1, S_2, ...)$ is a sequence of vehicle motion states and vehicle-vehicle interactions. Here, we assume that the model parameters (mean and covariance in multi-dimensional Gaussian distributions, and parameters in logistic function) have been learned in *Section IV*.

To predict the activity of a vehicle along time, we will generate a number of candidates And-Or Graph along time. We show an illustration of AOG online inference algorithm in Fig. 4. In the illustration, six motion activities in two scenarios are considered, as shown in Fig. 4(a). The solid arrows in Fig. 4(b) represent all the possible AOG for input sequence in time interval $[1, t-1]$. In other words, the solid arrows in Fig. 4(b) represent all the possible $\boldsymbol{L} = (l_1, l_2, ...l_{t-1})$ given $S_1, S_2, ...S_{t-1}$. The red solid arrow and the black solid arrow represent two AOG candidates for input sequence in time interval $[1, t-1]$. The red dotted arrow and the black dotted arrow represent all the possible motion activities for input data at time t. Thus, four AOG candidates are obtained for input sequence in time interval $[1, t]$. Then, the AOG candidates with lower posterior probability are removed. The same process goes for input data at time $t + 1$ and so on.

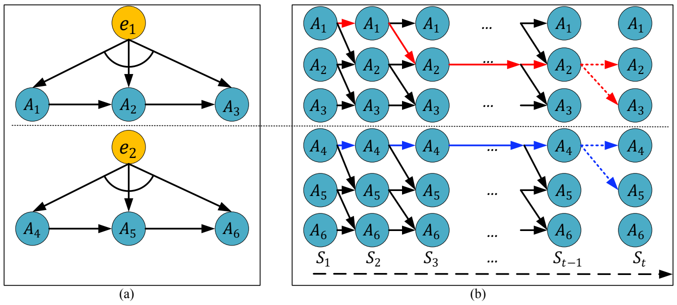


Fig. 4. An illustration of AOG online inference algorithm. (a) An example of vehicle motion activities in two driving scenarios. (b) inferring AOG time-step by time step

Recognizing the motion activities is equivalent to compute the And-Or Graph which interprets input S the best. It can be solved by maximize the posterior probability of G:

$$G^* = \underset{G}{\operatorname{argmax}}\, p(G|V) \tag{7}$$

To solve the Equation (7), an online inference algorithm for AOG is proposed, as shown in Algorithm 1. The AOG is inferred time-step by time-step. For each time-step, our algorithm will execute 3 steps. First, several candidate vehicle motion activities are proposed for the current time-step based on the AOG candidates for the past. Second, the likelihood and transition probability are computed. Third, AOG candidates with lower posterior probability are removed.

Given G^*, the future vehicle motion activity can be predicted by transition probability from the current motion activity to the next motion activity, as shown in the Equation (5).

Algorithm 1: OnlineUpdatingAOG

Input: S_t, $G^1_{t-1}, G^2_{t-1}, ...G^M_{t-1}$, $pG^1_{t-1}, pG^2_{t-1}, ...pG^M_{t-1}$, N

S_t : the input vehicle motion state and vehicle-vehicle interaction data at time t.

$G^1_{t-1}, G^2_{t-1}, ...G^M_{t-1}$: the AOG candidates for the input data in time interval $[1, t-1]$.

$pG^1_{t-1}, pG^2_{t-1}, ...pG^M_{t-1}$: the posterior probability of the AOG candidates, $pG^m_{t-1} = p(G^m_{t-1}|S_1, S_2, ...S_{t-1})$ for $1 \leq m \leq M$.

N: the threshold for AOG candidates numbers.

Output: $G^1_t, G^2_t, ...G^N_t$, $pG^1_t, pG^2_t, ...pG^N_t$, G^*_t

$G^1_t, G^2_t, ...G^N_t$: the AOG candidates for the input data in time interval $[1, t]$.

$pG^1_t, pG^2_t, ...pG^N_t$: the posterior probability of the AOG candidates, $pG^m_t = p(G^m_t|S_1, S_2, ...S_t)$ for $1 \leq m \leq N$.

G^*_t: the most likely AOG for the input data in time interval $[1, t]$.

```
num=0;
for i = 1 to M do
    Let l_{t-1} = A_{k-1} be the activity label for the input
    data at time t-1 in AOG G^i_{t-1} . Let d_{k-1} be the
    duration of A_{k-1};
    for l_t = A_{k-1}, A_k do
        num = num + 1
        compute p_1 = p(l_t|l_{t-1}, d_{k-1}), p_2 = p(S_t|l_t)
        update AOG candidates G^{num}_t = G^i_{t-1} + l_t
        update posterior probability
        pG^{num}_t = pG^i_{t-1} * p_1 * p_2
    end
end
keep N AOG with the highest posterior probability in as
AOG candidates;
G*_t = argmax_{G^i_t} pG^i_t, i = 1, 2, ...N
return G^1_t, G^2_t, ...G^N_t, pG^1_t, pG^2_t, ...pG^N_t, G*_t
```

VI. Experiments

To test the performance of our model in modeling and predicting motion activities, an intersection driving scenario is simulated using VISSIM. On average, a vehicle takes 27 seconds to pass this intersection. We take the driving scenario “turning right at the intersection” for example. 400 vehicle trajectories are collected in this scenario. The corresponding velocity and acceleration are also recorded. The lateral velocity and longitudinal velocity are used as the input of our model. K_E is set to 3. The time-step in the experiment is 1 second. The trajectories are used to manually calibrate the ground truth of vehicles motion activities. 200 data are taken as train data to learn the motion patterns in each motion activities, and the other 200 data are used as test data.

We take the driving scenario “turning right at the intersection” for example. As shown in Fig. 5, the learned AOG divides motions in this scenario into 3 motion activities, called “approaching the intersection”, “turning left”, and “leaving the intersection”. Fig 5(b) shows how the predicted activity (which

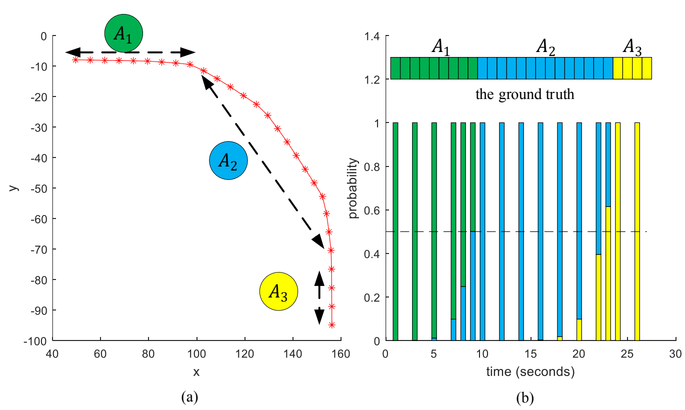


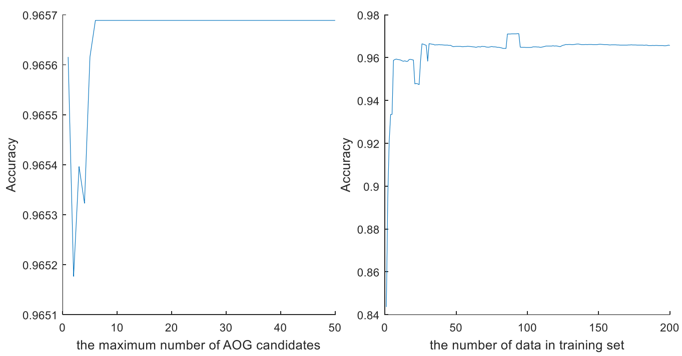
Fig. 5. (a) An illustration of the vehicle trajectory. (b) The predicted activity which the vehicle will take in the next time-step.

the vehicle will take in the next time-step) evolves along time and the ground truth of vehicle activities.

To measure the performance of our model, we define the prediction $Accuracy$ as follows. For each test sequence $\boldsymbol{q}_i = (q_{i1}, q_{i2}, ...q_i\tau_i)$, we successively take $\boldsymbol{S} = (q_{i1}, q_{i2}...q_{it})$ as input to predict l_{t+1}. l_{t+1} is the activity that the vehicle may take as the time $t+1$. The activity with a probability larger than 0.5 is taken as the final predicted activity. Then, for $\boldsymbol{q}_i$, we give $\tau_i - 1$ predictions. The number of correct prediction is denoted as c_i and the number of wrong prediction is $\tau_i - 1 - c_i$. The $Accuracy$ of our model is:

$$Accuracy = \frac{\sum_{i=1}^{i=200} c_i}{\sum_{i=1}^{i=200} \tau_i - 1} \quad (8)$$

Obviously, the $Accuracy$ defined here depends on both the number of data in training set and the parameter N in Algorithm 1 which denotes the maximum number of AOG candidates in inference.



(a) (b)

Fig. 6. The influence of the maximum number of AOG candidates and the size of training dataset

Fig.6(a) shows the variation of the $Accuracy$ with respect to the maximum number of AOG candidates in inference, where the number of data in training set is set as 200. We can see that, if the maximum number of AOG candidates in inference is larger than 6, the inference accuracy reaches the saturation level around 97%.

Fig.6(b) shows the variation of the $Accuracy$ with respect to the number of data in training set, where the maximum number of AOG candidates in inference is set as 10. We can see that, if the number of data in training set is larger than 100, the inference accuracy reaches the saturation level around 97%.

So, experiments show that the prediction accuracy of our model in intersection driving scenario is around 97%, when we maintain 6 candidate AOG in the pool and at least 100 training samples. This indicates that the effectiveness of the new model in modeling and predicting vehicle motion activities.

VII. Conclusion

In this paper, we propose a model based on And-or Graph to predict the vehicle motion activities. Our model can not only predict the motion activity but also give the corresponding probability, therefore it is more fault-tolerant than other models which only give a predicted motion activity. Experiments based on simulating data at an intersection show a high prediction accuracy (97%) of our model.

The application of our model also relies on the prior knowledge of driving scenarios. In the future work, we will summarize all the possible driving scenarios that a vehicle may encounter and check the performance of the proposed model.

References

[1] N. N. Zheng, S. Tang, H. Cheng, Q. Li, G. Lai, and F. Y. Wang, "Toward intelligent driver-assistance and safety warning systems," *Intelligent Systems IEEE*, vol. 19, no. 2, pp. 8–11, 2004.

[2] L. Li, D. Wen, N.-N. Zheng, and L.-C. Shen, "Cognitive cars: A new frontier for adas research," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 395–407, 2012.

[3] S. Lefèvre, D. Vasquez, and C. Laugier, "A survey on motion prediction and risk assessment for intelligent vehicles," *Robomech Journal*, vol. 1, no. 1, p. 1, 2014.

[4] K. Kim, B. Kim, K. Lee, B. Ko, and K. Yi, "Design of integrated risk management-based dynamic driving control of automated vehicles," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 57–73, 2017.

[5] M. Schreier, V. Willert, and J. Adamy, "An integrated approach to maneuver-based trajectory prediction and criticality assessment in arbitrary road environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 10, pp. 2751–2766, 2016.

[6] M. Bahram, C. Hubmann, A. Lawitzky, M. Aeberhard, and D. Wollherr, "A combined model- and learning-based framework for interaction-aware maneuver prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 6, pp. 1538–1550, 2016.

[7] D. Vasquez, T. Fraichard, and C. Laugier, "Growing hidden markov models: An incremental tool for learning and predicting human and vehicle motion," *The International Journal of Robotics Research*, vol. 28, no. 11-12, pp. 1486–1506, 2009.

[8] S. Sivaraman, B. Morris, and M. Trivedi, "Learning multi-lane trajectories using vehicle-based vision," in *Proceedings of the 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2011, pp. 2070–2076.

[9] T. Streubel and K. H. Hoffmann, "Prediction of driver intended path at intersections," in *Proceedings of the 2014 IEEE Intelligent Vehicles Symposium*. IEEE, 2014, pp. 134–139.

[10] S.-C. Zhu and D. Mumford, "A stochastic grammar of images," *Foundations and Trends® in Computer Graphics and Vision*, vol. 2, no. 4, pp. 259–362, 2007.

[11] Y. Lu, T. Wu, and S.-C. Zhu, "Online object tracking, learning and parsing with and-or graphs," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3462–3469.

[12] S. Park, X. Nie, and S.-C. Zhu, "Attribute and-or grammar for joint parsing of human pose, parts and attributes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[13] P. Wei, Y. Zhao, N. Zheng, and S.-C. Zhu, "Modeling 4d human-object interactions for joint event segmentation, recognition, and object localization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1165–1179, 2017.

[14] L. Li, D. Wen, and D. Yao, "A survey of traffic control with vehicular communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 425–432, 2014.

[15] L. Li, W. L. Huang, Y. Liu, N. N. Zheng, and F. Y. Wang, "Intelligence testing for autonomous vehicles: A new approach," *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 2, pp. 158–166, 2017.