Traffic Modeling and Identification using a Self-adaptive Fuzzy Inference Network

Sau Wai Tung
Institute for Infocomm Research
Agency for Science, Technology and Research (A*STAR)
Singapore
Email: tungsw@i2r.a-star.edu.sg

Chai Quek
Centre for Computational Intelligence (C2i)
School of Computer Engineering
Nanyang Technological University
Singapore
Email: ashcquek@ntu.edu.sg

Cuntai Guan
Institute for Infocomm Research
Agency for Science, Technology and Research (A*STAR)
Singapore
Email: ctguan@i2r.a-star.edu.sg

Abstract—Traffic modeling and identification is an important aspect of traffic control today. With an increase in the demands on today’s transportation network, an efficient system to model and understand the changes in the network is necessary for policy makers to make timely decisions which affect the overall level of service experienced by commuters. This paper proposes a novel approach to traffic modeling and identification using a Self-adaptive Fuzzy Inference Network (SaFIN). The study is performed on a set of real world traffic data collected along the Pan Island Expressway (PIE) in Singapore. By applying a hybrid fuzzy neural network in the traffic modeling task, SaFIN is able to capitalize on the functionalities of both the fuzzy system and the neural network to (1) provide meaningful and intuitive insights to the traffic data, and (2) demonstrate excellent modeling and identification capabilities for highly nonlinear traffic flow conditions.

Index Terms—Traffic modeling, fuzzy system, neural network, SaFIN.

I. INTRODUCTION

A pressing force for changes and adjustments in a transportation system is an increase in the demands on the system caused by an increase in the number of motor vehicles. This is a direct consequence of growing affluence, where vehicle ownership tends to rise with increasing income [1]. Subsequently, the performance of a transportation system is evaluated by its ability to move the increase influx of people, goods and equipments from one place to another place in an effective and reliable manner. On the other hand, traffic congestion on a highway occurs when the traffic demand exceeds the operational capacity of the highway. It is one of the most serious challenges threatening the effectiveness of a transportation system, since traffic congestion can lead to other associated environmental and economical issues, such as reduced capacity, reduced safety, increased gas emissions, increased fuel consumption, and increased travel times leading to productivity losses [2].

A most direct way to access the traffic flow conditions is to measure the level of service that is experienced by commuters [3]. This means that a high level of service is expected when the forecasted traffic flow is small; while the expected level of service drops when the forecasted traffic flow is huge. In such a case, pre-emptive measures can then be adopted to improve the level of service and prevent the onset of congestion when a clearer understanding of the traffic conditions is available (for example, raising the tolls on a road during peak hours to limit the number of vehicles, or increasing the gasoline tax to reduce the overall number of vehicles on the roads). Thus, a model that provides a high level of forecasting accuracy, together with a high level of interpretability, is desired.

This paper proposes a novel approach to traffic modeling and identification using a Self-adaptive Fuzzy Inference Network (SaFIN) [4]. The objective is two folds; namely, (1) to provide valuable and meaningful insights to the numerical raw traffic data; and (2) to provide an efficient way to model the traffic flow conditions. Since SaFIN is a hybrid fuzzy neural network, it possesses the functionalities of the two individual systems. This means that the system possess high-level human-like reasoning mechanism where information can be expressed as a set of human-interpretable linguistic IF-THEN fuzzy rules; while concurrently utilizing low-level neural learning for the modeling and identification of the system.

The rest of the paper is organized as follows. Section II presents the hybrid system SaFIN for traffic modeling, where the architecture and the reasoning mechanisms in SaFIN are discussed. This is followed by a description of the highway traffic flow data in Section III. Section IV presents the experimental results achieved. Lastly Section V concludes the paper.

II. SAFIN: SELF-ADAPTIVE FUZZY INFEERENCE NETWORK

SaFIN is a five layers fuzzy neural network as shown in Fig. 1. Layer 1 consists of the P number of input dimensions \( I_p \) for \( p = 1 \ldots P \). Layer 2 represents the antecedent nodes \( A_{r_k} \), where there are \( J_r \) number of fuzzy clusters in \( I_r \). Layer 3 consists of the rule nodes \( R_k \) for \( k = 1 \ldots K \). Layer 4 represents the consequent nodes \( C_{q_k} \), where there are \( L_q \) number of fuzzy clusters in an output dimension \( O_q \). Finally layer 5 is the \( Q \) number of output dimensions. As seen, the input vector is denoted as \( x = (x_1, \ldots, x_P) \); the corresponding desired output vector is denoted as \( d = (d_1, \ldots, d_Q) \); and the computed output vector is denoted as \( y = (y_1, \ldots, y_Q) \).
Layer 3 of SaFIN encrypts the rulebase of the system where each rule node encodes an IF-THEN Mamdani-type fuzzy rule [5] given as in (1):

\[
R_k : \text{IF } x_1 \text{ is } A_{1_k}^{(h)} \text{ and } \ldots \text{ and } x_p \text{ is } A_{p_k}^{(h)} \text{ THEN } y_1 \text{ is } C_{1_k}^{(l)} \text{ and } \ldots \text{ and } y_q \text{ is } C_{q_k}^{(l)}
\]  

(1)

where \(A_{j_k}^{(h)}\) (resp. \(C_{j_k}^{(l)}\)) is the \(j\)-th antecedent (resp. \(l\)-th consequent) node associated with the \(p\)-th input (resp. \(q\)-th output) dimension that is connected to the rule node \(R_k\). Each antecedent and consequent node is then defined as a Gaussian membership function: \(\mu(c, \sigma; x) = e^{-((x-c)^2)/2\sigma^2}\) such that \(c\) and \(\sigma\) are the centre and width of the function respectively.

There are two main components to the design of SaFIN: namely, the fuzzy partitioning of the input-output dimensions and the rule generation procedure. Initially, there are neither fuzzy partitionings in the input-output spaces nor fuzzy rules in the system, i.e., there are no nodes in hidden layers 2–4. A localized learning of the fuzzy labels using the single-pass Categorical Learning Induced Partitioning (CLIP) technique [4], inspired from the behavioral category learning process demonstrated in humans, is first carried out. This enhances the interpretability of SaFIN such that the resultant fuzzy clusters are highly ordered. New clusters are incorporated into the system when the knowledge extracted from the incoming training tuple is novel as compared to existing clusters in the system; while refinements are made to existing clusters to incorporate the new knowledge. The initialization performed in an input dimension \(p\) is described as in (2):

\[
c_{j_p} = x_p \\
\sigma_{j_p} = R\left(\sqrt{\frac{\text{min}_p - x_p^2}{\log n}}, \sqrt{\frac{\text{max}_p - x_p^2}{\log n}}\right)
\]

(2)

where \(R(\sigma_1, \sigma_2) := \frac{1}{2} [\sigma_1 + \sigma_2]\) defines a regulator function, and \(\alpha\) is the minimum membership threshold. The boundary for the domain is given as \([\text{min}_p, \text{max}_p]\). Subsequently, the formation of a new fuzzy cluster \(A_{j_p(t)+1}\) in the input dimension \(p\) can be described using (3):

\[
c_{j_p(t)+1} = x_p \\
\sigma_{j_p(t)+1} = \begin{cases} 
\sigma^R & \text{if } j_p^L = \text{NULL} \\
\sigma^L & \text{if } j_p^R = \text{NULL} \\
R(\sigma^R, \sigma^L) & \text{otherwise}
\end{cases}
\]

(3)

where

\[
\sigma^R = R\left(-\frac{(c_{j_p} - x_p)^2}{\log n}, \sigma_{j_p}^{(l)}(t)\right)
\]

\[
\sigma^L = R\left(-\frac{(c_{j_p} - x_p)^2}{\log n}, \sigma_{j_p}^{(l)}(t)\right)
\]

The immediate left and right neighbours of the newly created cluster are denoted as \(A_{j_p}^L\) and \(A_{j_p}^R\) respectively, where

\[
j_p^L = \begin{cases} 
\text{NULL} & \text{if } c_{j_p} \geq x_p \\
\arg\min_{x_{j_p}<x_p} r_j(x_j - x_p) & \text{for } 1 \leq j_p \leq j_p(t)
\end{cases}
\]

\[
j_p^R = \begin{cases} 
\text{NULL} & \text{if } c_{j_p} \leq x_p \\
\arg\min_{x_{j_p}>x_p} r_j(x_j - x_p) & \text{for } 1 \leq j_p \leq j_p(t)
\end{cases}
\]

Refinements are made to immediate left and right neighbours of the newly created cluster as follows: \(\sigma_{j_p}^{(l)}(t+1) = \sigma_{j_p}^{(l)}(t)\) for \(j_p^L \neq \text{NULL}\) and \(j_p^R \neq \text{NULL}\). The same fuzzy clustering process is performed for each output dimension.

The second key component in the design of SaFIN is the formulation of the rulebase. Firstly, a fuzzy rule is formulated to capture the knowledge from each of the incoming training tuples. A fuzzy rule \(R^*\) is created for an incoming training tuple such that the antecedents and consequents of \(R^*\) are \(\{A_{j_p}^L\}_p\) and \(\{C_{j_p}^R\}_q\) respectively, where \(A_{j_p}^L\) is the best matched fuzzy cluster in the \(p\)-th input dimension and \(C_{j_p}^R\) is the best matched fuzzy cluster in the \(q\)-th output dimension. The best matched cluster of an input (and output) dimension is given as follows: \(j_p^R = \arg\max_{j_p \in \{C_{j_p}^R\}_q} r_j(x_j - x_p)\). If \(R^*\) is novel, it is inserted into the rulebase with an initial weight of 1. Else, the weight of the identical fuzzy rule with the same antecedent and consequent segments in the existing rulebase is increased by 1. Finally, conflicting rules with lower weights are deemed as outliers and are deleted from the system. This approach ensures that SaFIN maintains a unique and consistent rulebase that is able to provide a most aptly description to the application problem.

As seen from Fig. 1, the reasoning process of SaFIN is represented by solid arrows where the input vector \(x\) is presented to the system at layer 1. The system then performs inference based on the input vector by propagating the information through layers 2 to 4. Consequently, the system produces a computed output vector \(y\) at layer 5. To ensure a logically tractable reasoning mechanism, the most commonly adopted...
Compositional Rule of Inference (CRI) [6] is employed in SaFIN. The generic operations for SaFIN are defined as follows. The activation functions of each layer $M \in \{1 \ldots 5\}$ are denoted as $a^{(M)}$, and the corresponding output for an arbitrary node is denoted as $o$:

**Layer 1:**

$$o_p = a^{(1)}(x_p) = x_p.$$  

**Layer 2:**

$$o_p = a^{(2)}(o_p) = \mu_{x_p}(E_{x_p}\cdot\sigma_{x_p}; x_p).$$

**Layer 3:**

$$o_k = a^{(3)} \left( a^{(4)}(o_1^{(4)}, \ldots, o_{Lm}^{(4)}), a^{(5)}(o_1^{(5)}, \ldots, o_{Ln}^{(5)}) \right) = \min_{p \in \{1 \ldots P\}} o_p^{(3)}.$$  

**Layer 4:**

$$a_k = a^{(4)} \left( a^{(4)}(o_1^{(4)}, \ldots, o_{Lm}^{(4)}), a^{(5)}(o_1^{(5)}, \ldots, o_{Ln}^{(5)}) \right) = \max_{k \in \{1 \ldots K\}} a_k^{(4)}.$$  

**Layer 5:**

$$y_q = o_q = a^{(5)}(o_1, \ldots, o_{Ln}) = \frac{\sum_{l \in \{1 \ldots Lm\}} q_l \cdot \mu_l \cdot \sigma_l}{\sum_{l \in \{1 \ldots Ln\}} q_l \cdot \sigma_l}.$$  

III. DATA COLLECTION AND EXPERIMENTAL SETUP

The set of highway traffic flow data employed is described in this section. Data was collected from a five lanes section along the Pan Island Expressway (PIE), Singapore, in both the east-bound and west-bound directions towards Changi and Jurong respectively. In this paper, only the east-bound direction is considered, where the data collection is performed at site 29 located at exit 15 on the expressway. Data samples were collected from inductive loop detectors installed beneath the road surface. Since 1996, the Land Transport Authority has pre-installed such detectors along major roads in Singapore to facilitate the collection of traffic flow data.

Fig. 2 shows a picture of the location where the data was collected. There are a total of five lanes; namely, three straight lanes for the main traffic (lanes 1–3), and two exit lanes (lanes 4–5). Only data from the three straight lanes are used in this paper. They are correspondingly denoted as L1, L2, and L3. Each data sample has four attributes; namely, the time $t$ at which the traffic flow data was measured, and the traffic flow densities for the three straight lanes during time $t$. Subsequently, SaFIN is employed to model the traffic flow trend. After which, the trained model is used to predict traffic flow density of a lane (L1, L2 or L3) at the time $t + \tau$ for $\tau = 5, 15, 30, 45$ and 60 minutes.

Fig. 3 shows the traffic flow data for lanes L1–L3 spanning over a period of six days from 5-th to 10-th September 1996. The data is divided into three cross-validation groups, hereby denoted as CV1, CV2 and CV3 respectively. The training data for each cross-validation group is extracted accordingly from the period labelled in Fig. 3.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

This section presents the modeling performances of SaFIN for the highway traffic flow data. The effectiveness of SaFIN is evaluated in terms of the mean squared error MSE and the Pearson correlation coefficient $R$ achieved between the computed output of the model and the actual traffic flow density of the highway. These measures help to determine the suitability of employing SaFIN for the detection of changes in the demands on the highway usage throughout the entire experimental duration (i.e. the duration of interest). The performance of SaFIN is subsequently benchmarked against the following models; namely, Hebb-R-R [7]; RSTOP [8]; MLP (with a configuration of 4 input nodes, 10 hidden nodes and 1 output node); GenSoFNN [9], EFuNN [10]; DENFIS [11]; and eFSM [12].

Fig. 4 illustrates the identified fuzzy clusters in the time domain and those for lanes L1–L3 for the training set of CV1 when $\tau = 5$ minutes. The distributions of the raw numerical data are also shown for each of the three lanes in the figure. The identified fuzzy clusters are well-ordered where clear distinct semantic fuzzy meanings can subsequently be attached. As seen, the fuzzy clusters identified in SaFIN for lanes L1–L3 coincide with the peaks of the distributions as marked by...
the dotted circles. Although there are very low distributions of data closer to the upper bounds of the three lanes, it is observed that the distributions immediately before the peaks are closest at the zero mark. Hence, SaFIN identifies a cluster near the upper bound of each lane to cater for this peak. This figure demonstrates the tailored approach adopted in SaFIN for the fuzzy partitioning, where the numbers and positions of fuzzy clusters identified correspond to the distributions of the original numerical data. Subsequently, this allows traffic officers and policy makers to have an intuitive grasp of the knowledge embedded in the raw numerical data.

Using the semantic fuzzy labels in Fig. 4, twenty-three fuzzy rules are identified by SaFIN as listed in Table I. From the table, four rules correspond to the morning hours, fourteen rules correspond to the afternoon period, and five rules describe the night time traffic. During the early morning and late night hours, the amount of traffic flowing through lanes L1–L3 range from Low to Medium. This means that, there is only light to moderate amount of traffic on the highway during these periods. Comparing this observation to the training set of CV1 in Fig. 3, this is indeed the case. Subsequently, the predicted traffic demand on the highway during these periods is moderately low. This corresponds to the usual work-life cycle of most people, where majority of the highway users are resting during the early morning and late night periods. On the other hand, most of the identified rules describe traffic flow for the afternoon period, i.e., during the day when most people travel for work and errands. Within this period, the traffic flow in lanes L1–L3 can range from Low to High. This means that variations in the highway usage can be observed throughout the day period, where half of the fourteen rules describe the peaks hours of the day when people

<table>
<thead>
<tr>
<th>Rule</th>
<th>Time t</th>
<th>L1(t)</th>
<th>L2(t)</th>
<th>L3(t)</th>
<th>L1(t + 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Morning</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>Morning</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>3</td>
<td>Morning</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>4</td>
<td>Morning</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>5</td>
<td>Afternoon</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>6</td>
<td>Afternoon</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>7</td>
<td>Afternoon</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>8</td>
<td>Afternoon</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>9</td>
<td>Afternoon</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>10</td>
<td>Afternoon</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>11</td>
<td>Afternoon</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>12</td>
<td>Afternoon</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>13</td>
<td>Afternoon</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>14</td>
<td>Afternoon</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>15</td>
<td>Afternoon</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>16</td>
<td>Afternoon</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>17</td>
<td>Afternoon</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>18</td>
<td>Afternoon</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>19</td>
<td>Night</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>20</td>
<td>Night</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>21</td>
<td>Night</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>22</td>
<td>Night</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>23</td>
<td>Night</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
</tr>
</tbody>
</table>

L–Low; M–Medium; H–High
rush to/off work (i.e. rules *13–*18), with the remaining describing the day hours when people are already at work. Subsequently, the predicted traffic demand on the highway during the day period is moderately high (as compared to that during the early morning and late night hours). Hence, the set of identified fuzzy rulebase in SaFIN provides a logical, sound, and intuitive description to the highway traffic flow, where the demands on the highway usage for different hours of the day can be effectively accounted for.

The recall and prediction performances of SaFIN for lane L1 when \( \tau = 5 \) minutes are shown in Fig. 5. The results are shown for the three cross-validation classes CV1–CV3. As seen, the recall performances on the training sets for CV1–CV3 (i.e. Fig. 5(a),(c),(e)) are excellent, with more than 0.9 achieved for the R values. To illustrate that SaFIN has learnt from the training sets to generalize the underlying knowledge to the entire duration of interest, the trained model is subsequently employed on the testing datasets. The prediction performances of SaFIN on the unseen data for CV1–CV3 are shown in Fig. 5(b),(d),(f) respectively. Compared to the recalling performances, the prediction performances of SaFIN have deteriorated slightly with larger reported MSE values, and smaller reported R values. Despite that, SaFIN is able to perform excellent prediction on the testing datasets, where the peaks and the valleys of the traffic flow densities (i.e. the increase and decrease in demands on the highway usage) can be clearly identified. This result demonstrates the superior modeling and generalization abilities of SaFIN such that it is able to provide a good prediction of the highway traffic flow density, where the changes in demands on the highway usage are effectively accounted for.

The prediction performances of SaFIN are benchmarked against MLP and existing neuro-fuzzy systems, and the consolidated highway traffic flow prediction results are shown in Fig. 6. The average R value and the average MSE value from the three cross-validation groups CV1–CV3 for each prediction horizon are plotted with respect to the lanes L1–L3. As seen, SaFIN is among the top performers in this highway traffic flow prediction application such that it is able to consistently achieve one of the highest average R and the lowest average MSE under different time horizons. This is particularly prominent when \( \tau = 60 \) minutes where SaFIN is ranked either the first or second positions for all three lanes L1–L3; while most of the benchmarking models have greater errors due to a longer time lag in the prediction horizon. Through this benchmarking with other existing models, SaFIN has demonstrated a consistently excellent performance such that it is able to perform good prediction within the various time periods of interest.

Table II shows the average performances of all the models for this highway traffic flow density modeling task. As clearly shown, SaFIN demonstrates superior modeling potential, second only to Hebb-R-R, in terms of the average benchmarking measures achieved. Despite employing a time-consuming and computationally intensive iterative post-training phase to recursively identify a reduced set of fuzzy rules with the aim of a good accuracy, Hebb-R-R performs only marginally better than SaFIN. Comparatively, the performance of SaFIN is much more consistent and stable as shown by the small standard deviations about the average benchmarking indexes. Although GenSoFNN achieves a lower standard deviation in the R value, it should be noted that the average performance of GenSoFNN is among the poorest under both the benchmarking measures. This result demonstrates the excellent modeling potential of SaFIN for the highway modeling task; while maintaining a highly consistent and stable performance under varying time horizons (i.e. different time periods of interest).

### Table II

<table>
<thead>
<tr>
<th>Model</th>
<th>Average R (± Std. Dev.)</th>
<th>Average MSE (± Std. Dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hebb-R-R</td>
<td>0.864 (± 0.046)</td>
<td>0.114 (± 0.042)</td>
</tr>
<tr>
<td>RSPOP</td>
<td>0.834 (± 0.041)</td>
<td>0.146 (± 0.038)</td>
</tr>
<tr>
<td>MLP (4-10-1)</td>
<td>0.847 (± 0.065)</td>
<td>0.130 (± 0.055)</td>
</tr>
<tr>
<td>GenSoFNN</td>
<td>0.813 (± 0.028)</td>
<td>0.164 (± 0.037)</td>
</tr>
<tr>
<td>EFBNN</td>
<td>0.798 (± 0.050)</td>
<td>0.189 (± 0.041)</td>
</tr>
<tr>
<td>DENFIS</td>
<td>0.831 (± 0.051)</td>
<td>0.153 (± 0.054)</td>
</tr>
<tr>
<td>eFSM</td>
<td>0.840 (± 0.043)</td>
<td>0.154 (± 0.040)</td>
</tr>
<tr>
<td>SaFIN</td>
<td>0.862 (± 0.043)</td>
<td>0.118 (± 0.037)</td>
</tr>
</tbody>
</table>

This paper demonstrates the suitability and effectiveness of employing SaFIN for the modeling of highway traffic flow density. As the amount of traffic flow on the highway is directly related to the service of quality experienced by commuters, it is important to have an efficient system to perform a good and reliable traffic flow prediction. Subsequently, adjustments can be made to existing traffic planning policies when the forecasted demand on the highway exceeds the operational capacity of the highway (i.e. when traffic congestion occurs) such that a comfortable level of service can be provided. This is possible when a clear understanding of the traffic conditions is available.

Experimental results based on real-world traffic flow data collected from the Pan Island Expressway in Singapore are presented. Results showed that SaFIN is able to provide intuitive insights to the raw numerical data by providing semantic fuzzy labels and fuzzy rules to explain the traffic flow density on the highway throughout the day, i.e., the changes in the demands on the highway usage throughout the duration of interest can be effectively accounted for in SaFIN. The modeling and generalization abilities of SaFIN are then compared against MLP and some existing neuro-fuzzy systems. Results showed that SaFIN is able to achieve superior performances with varying prediction horizons, while most of the benchmarking models are not able to handle the larger time lag in the prediction. In addition, SaFIN has also achieved a more consistent and stable performance for the highway traffic flow prediction when compared to the other models.

### V. Conclusion

This paper demonstrates the suitability and effectiveness of employing SaFIN for the modeling of highway traffic flow density. As the amount of traffic flow on the highway is directly related to the service of quality experienced by commuters, it is important to have an efficient system to perform a good and reliable traffic flow prediction. Subsequently, adjustments can be made to existing traffic planning policies when the forecasted demand on the highway exceeds the operational capacity of the highway (i.e. when traffic congestion occurs) such that a comfortable level of service can be provided. This is possible when a clear understanding of the traffic conditions is available.

Experimental results based on real-world traffic flow data collected from the Pan Island Expressway in Singapore are presented. Results showed that SaFIN is able to provide intuitive insights to the raw numerical data by providing semantic fuzzy labels and fuzzy rules to explain the traffic flow density on the highway throughout the day, i.e., the changes in the demands on the highway usage throughout the duration of interest can be effectively accounted for in SaFIN. The modeling and generalization abilities of SaFIN are then compared against MLP and some existing neuro-fuzzy systems. Results showed that SaFIN is able to achieve superior performances with varying prediction horizons, while most of the benchmarking models are not able to handle the larger time lag in the prediction. In addition, SaFIN has also achieved a more consistent and stable performance for the highway traffic flow prediction when compared to the other models.
Fig. 5. Recall and prediction of traffic density on lane L1 when $\tau = 5$ minutes.

Fig. 6. Traffic flow prediction results.
REFERENCES


