Predicting Traffic Breakdown in Urban Expressways Based on Simplified Reservoir Computing

Taiga Okamoto 1, Hiroyasu Ando 1, Kentaro Wada 1, Risa Mukai 2, Yoshifumi Nishiumi 2, Dai Tamagawa 2

1 Systems and Information Engineering University of Tsukuba
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan
2 Hanshin Expressway Co., LTD
3-2-4 Nakanoshima, Kita-ku, Osaka 530-0005, Japan
s1920552@s.tsukba.ac.jp, ando@sk.tsukuba.ac.jp

Abstract
Due to the development of information technology, several studies have been conducted on short-term traffic forecasting in intelligent transportation systems during the last few decades. Specifically, predicting traffic congestion plays a crucial role in smoothing traffic flow, so that several models for predicting it have been proposed. This study proposed a simple prediction model based on reservoir computing for occurrence of traffic congestion and compared it with two conventional models, namely the autoregressive model as a baseline and the convolutional neural network based deep learning model. We use data obtained every five minutes for four years from traffic detectors installed on the Hanshin Expressway in Japan. The prediction results of the three models were evaluated through several indices. The accuracy of the proposed model was slightly lower than that of the deep learning-based approach. Moreover, the training time of the proposed model was less than 2 s, which is 1900 times shorter than the deep learning model, approximately 45 minutes. These results indicate that the computational cost can be remarkably reduced through the proposed method while maintaining a certain accuracy level.

Introduction
In many years, traffic congestion has been a serious problem with several issues remaining to be addressed, such as increased travel time and environmental impact by the emission of GHG, as well as blocking emergency vehicles. Due to these problems, in the research field of intelligent transportation systems (ITS), short-term traffic prediction has become a hot topic; thus, several studies have been conducted in this regard. Concerning short-term traffic forecasting, various targets are studied (Vlahogianni, Karlaftis, and Golias 2014), ranging from traffic flow (Polson and Sokolov 2017; Yao et al. 2019), speed (Ma et al. 2015a) to travel time (Zhang and Haghani 2015). Moreover, it is also important to directly forecast traffic congestion, which is the congestion prediction in the short-term future. If occurrence of traffic congestion, called traffic breakdown, can be predicted long enough in advance, the prediction can play an important role by providing drivers with information to avoid congestion. Consequently, traffic congestion can be eliminated. From the service perspective, car navigation systems accompanied by a congestion prediction model can provide the best route in terms of travel time. As avoiding traffic congestion is one of the most important ITS topics, various models have been proposed, such as pheromone communications (Ando et al. 2006) and spatial econometric models, focusing on congestion propagation on adjacent links (Hu, Kaparias, and Bell 2009). Moreover, several deep learning models for predicting traffic congestion and flow have been developed (Ma et al. 2015b, 2017; Zhang et al. 2019).

Conventional deep learning based algorithms require high computational power for predicting traffic congestion with high precision. Therefore, in this study, we focus on obtaining a computational low-cost prediction algorithm. Specifically, we propose a simplified reservoir computing (SRC) as a new prediction methodology for traffic breakdown. Reservoir computing (RC) (Lukoševičius and Jaeger 2009) is an extension of neural networks developed after recurrent neural networks (RNNs), including the echo state network (ESN) (Jaeger 2001) and the liquid state machine (LSM) (Maass, Natschläger, and Markram 2002). RC consists of two components: the reservoir, which is equivalent to the hidden layer of an RNN, and the readout wires to the output units. The difference from the RNN is that RC learns only the readout weights, and the other weights are fixed at random. The reservoir part maps the input information into a high-dimensional feature space with nonlinearity. This nonlinear map makes it possible to create efficient learning algorithms, such as linear regression, to reduce the computational cost significantly. In recent years, with the development of RC, considerable research has been conducted in the sense of physical RC, where the reservoir is replaced with physical systems instead of random neural networks. The idea is that physical phenomena with some inputs are so complex that the observed dynamics provide a nonlinear high-dimensional map. Several studies have been developed in recent years (Tanaka et al. 2019), such as the use of water surfaces (Fernando and Sojakka 2003) and soft materials (Nakajima et al. 2014) as reservoirs. These approaches directly use the dynamics of physical systems as a computational resource, holding a great promise in improving energy efficiency and reducing computational resources. In this paper, we apply the concept of physical RC to road traffic dynamics with an appropriate modification. The effectiveness...
of the proposed method in predicting traffic breakdown is shown by comparing the performance of the proposed algorithm with that of the AR model and the convolutional neural network (CNN) model in several indices. The data were obtained from traffic detectors installed at 100 locations on the Hanshin Expressway in Japan from April 2015 to March 2019.

Predicting traffic breakdown

This section describes the details of traffic congestion prediction on the Hanshin Expressway. In this study, we developed a binary classification model to predict whether the traffic congestion occurs after 10 min from the time of the prediction regarding the No. 11 inbound line of the Hanshin Expressway, which is a single lane (See Fig. 1). The data used to train the prediction model was sampled every five minutes from detectors installed at 100 locations. We divided the obtained data into two parts. The first part was used for training the model, considered data from April 1, 2015, to November 31, 2017. The second part was used for testing, considered data from April 1, 2018, to March 31, 2019. The following criteria were considered from a practical perspective to obtain an accurate prediction.

- Time difference between predicted times ahead to the actual beginning of traffic congestion is within 30 minutes (early prediction)
- Traffic congestion begins five minutes earlier than the predicted time (slightly later prediction)

These criteria were considered to be sufficiently useful in predicting the signs of congestion before it occurs. Moreover, to evaluate the performance accuracy of the model, the following model evaluation indicators were used in addition to the conventional accuracy: (i) the learning time representing the computational cost of the model, (ii) the inference time representing the real-time feasibility of predicting the occurrence of congestion, and (iii) the number of parameters representing the complexity of the model. Note that we use F1 score for the conventional accuracy instead of AUC-ROC, because the threshold 10 mins plays a significant role for practical prediction.

Models

In this section, we explain in detail the learning and prediction methods for three models relying on the (I) AR model, (II) CNN model, and the proposed (III) simplified reservoir computing.

AR Model

We considered the AR model as a baseline. This model predicts traffic breakdown using only the data obtained from the target detector, that is one dimensional model, which can be written as follows:

$$\hat{y}(t) = c_{bias} + \sum_{i=0}^{I} \sum_{j=0}^{J} c^j(t-i)x^j(t-i)$$  \quad (1)

where $x^j(t)$ represents the data at time $t$ with data type $j$, and $c$ is the weight for each data $x$. The data type obtained from the detector includes aspects such as the traffic volume, average speed, and vehicle occupancy. The number of data type is $J$. The order of AR model is $I$. The teacher data $y$ takes 0 when traffic is not congested and 1 when congested. The estimated $\hat{y}$ takes continuous values. As a baseline, we considered the 285th order AR model. This was the most accurate AR model of order 288(one day ago) from the first order.

The teacher data $y$ used in the training phase is not only when traffic congestion begins but also during traffic congestion. The reason is that if the time of the beginning of traffic congestion is only used as teacher data, the training data is unbalanced. In this case, learning is not successful. Moreover, this task is a binary classification; thus, the binary prediction variable $\hat{u}(t)$ is calculated as follows:

$$\hat{u}(t) = \begin{cases} 1 & (\hat{y}_u \geq 0.5) \\ 0 & (\hat{y}_u < 0.5) \end{cases}$$  \quad (2)

Because the data of whether the traffic is congested or not is used for training, that is, $y \in \{0, 1\}$, the prediction outputs should also be whether the traffic is congested or not 10 min after the predicted moment. Namely, the output $\hat{u}$ is represented by 0 (not congested) or 1 (congested). The threshold value of 0.5 in Equation 2 does not have a significant effect on the F1 value even if it is changed slightly, so the threshold value of 0.5 was used this study. Therefore, the timing when traffic congestion begins is defined as follows:

- The first time point that is predicted to be $\hat{u}(t) = 1$.
- The predicted $\hat{u}(t)$ in the previous 30 minutes from the first appearance of 1 is 0, i.e. $\hat{u} = \{\ldots 0000001 \ldots\}$.

The timing when these two conditions are satisfied is defined as the predicted beginning time of congestion.

Deep Learning Model

The CNN, which is often used for image recognition, was used as a deep learning model in this study. It is important to incorporate local information regarding the spatial distance...
Table 1: Overall assessment in each model (The AR model uses only one point, while the other models use all 44 points.)

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1 value</th>
<th>Number of parameters</th>
<th>Training Time(s)</th>
<th>Inference Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR (285th order)</td>
<td>0.473</td>
<td>0.439</td>
<td>0.455</td>
<td>1,157</td>
<td>0.461</td>
<td>0.014</td>
</tr>
<tr>
<td>CNN</td>
<td>0.802</td>
<td>0.697</td>
<td>0.746</td>
<td>3,100,988</td>
<td>2688</td>
<td>61</td>
</tr>
<tr>
<td>SRC</td>
<td>0.734</td>
<td>0.713</td>
<td>0.723</td>
<td>17,644</td>
<td>1.41</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 2: Evaluation of each model at 9 km post from the origin.

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1 value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR (285th order)</td>
<td>0.5159</td>
<td>0.630</td>
<td>0.567</td>
</tr>
<tr>
<td>CNN</td>
<td>0.718</td>
<td>0.818</td>
<td>0.763</td>
</tr>
<tr>
<td>SRC</td>
<td>0.758</td>
<td>0.837</td>
<td>0.796</td>
</tr>
</tbody>
</table>

between the neighboring detectors and temporal changes in the traffic volume around the target to predict traffic congestion at a particular point in time. Here, we followed the prediction process presented in (Mukai et al. 2019). The data from 44 vehicle detectors installed on the side of the passing lane in the target interval were used for the training. The input corresponded to an image of 44 locations at 6 points of time (30 min) with one channel of image size, that is, a space-time relationship. So we used seven channels, that means seven data types as the input. Unlike in the AR model, the teacher data for training are flagged only when traffic congestion occurs. Moreover, the definition of the occurrence of traffic congestion was the same as in the previous subsection. Here, we randomly assigned the model structure and initial parameter values for 149 trials and adopted a model with seven convolutional layers and four total coupling layers (for the output unit) using Leaky ReLU for activation function that had the highest accuracy.

Simplified Reservoir Computing (SRC)

Finally, we consider a simplified reservoir computing (SRC) for predicting traffic dynamics. The SRC does not consider an external input, and thus, the reservoir units have only internal feedback. In addition, the SRC prediction target is also inside the system. Similar to this concept, RC has been recently applied to road traffic (Ando and Chang 2019; Chen et al. 2020). Accordingly, the detectors on the highway were considered as reservoir units to predict traffic breakdown. The reservoir units exploited were not limited within the No. 11 inbound line but also included the detectors on the outbound line of No. 11, which is in the opposite lane. The reason for expanding reservoir units is that road traffic is highly complex, and to create the higher dimensional mapped space, that is, the dimension becomes almost twice. These can be described as follows: \( X(t) = [1; x_1^1(t), ..., x_n^1(t); x_1^2(t), ..., x_n^2(t); ..., x_1^J(t), ..., x_n^J(t)] \), where \( x_n^j(t) \) represents the data obtained from the detector \( n \) at time \( t \) with data type \( j \in \{1, ..., J\} \). [\( \cdot \cdot \cdot \)] represents a vector concatenation. The output is calculated as follows:

\[
\hat{y}(t + \tau) = W^{out}X(t)
\]

The weight matrix \( W^{out} \) was estimated by ridge regression using the teacher data \( y \). \( \tau \) is forecasting horizon. Similar to the AR model, the SRC model also uses \( y(t) = 1 \) during traffic congestion as the teacher data for training. Thus, the prediction \( \hat{u}(t) \) was obtained similar to Equation (2)

Results

Tables 1 and 2 show the results from evaluating the three models illustrated above for precision, recall, and F1 value. Table 1 shows the results of summing 44 confusion matrices with respect to detectors. The baseline model (6th order AR model) has an F1 value of 0.278, and the 288th order AR model has increased accuracy of 0.455. However, this result is still lower than that of the CNN and SRC models. The CNN model has the highest F1 value, 0.746, among the three models. As for the SRC model, the Recall shows the highest accuracy of 0.713. The F1 value of 0.723 was slightly lower than that of CNN. Note that the number of parameters for the SRC model is approximately 200 times smaller than that of the CNN, which implies that the SRC model has the same level of accuracy despite the simplicity of the model. Due to this result, we can say that when SRC is almost similar to that of the CNN in terms of accuracy even though the number of parameters is much smaller. This difference affects the actual computation time for training and inference. As shown in Table 1, the training time for CNN is approximately 1900 times as long as that for SRC. The time for inference is also very different from each other. We used a single CPU of Intel(R) Core i9-9940X for computing the SRC, and the GPU of NVIDIA(R) GeForce GTX 1080 Ti 11GB for the calculation of the CNN. Moreover, Table 2 lists the results for one confusion matrix at 9 km from the origin. As seen in the table, the accuracy of the SRC model is better than the CNN model in terms of the F1 value, 0.763 for CNN and 0.796 for SRC. In other words, the accuracy of the SRC is better than that of the CNN model at some locations, which implies that the SRC is effective in some cases.

Discussion

In this study, we proposed a simplified reservoir computing method along with AR models and a CNN model for a 10-minute prediction task for the occurrence of traffic congestion on the Hanshin Expressway. While the 6th and 288th order AR models were not suitable for short-term traffic congestion prediction, the CNN model could predict with a certain degree of accuracy. The reason is that CNN has the property of capturing local information in both temporal and spatial aspects. In addition, although the computation method for the SRC model is the same as the multidimensional AR models, enabled sufficiently accurate prediction by the usage of information from detectors in the opposite direction to the target. This is under the assumption of that there exists nonlinear dynamics behind the observed time se-
ries, which is different approach from statistical AR models. Multiple observations may provide a high-dimensional nonlinear map of the target time series. Due to the complexity of the traffic dynamics obtained from the highway detectors, the SRC was found to be an effective method of prediction in terms of accuracy and computational cost. Furthermore, we have used the globally optimized parameter in table 2, but it has been found that the accuracy can be slightly improved if the parameter is optimized individually at each location. Setting $\lambda$ to 500 results in an F1 value of 0.832 at 9km post from the origin. And if data gathered by more detectors are much larger than those used in this study, the SRC can reduce the computational cost because of the relatively small number of parameters compared to deep learning. Therefore, it may also have great power in training with much higher dimensional data obtained from the entire highway network.

Finally, predicting traffic congestion must respond to varying traffic conditions in real time. Therefore, it is necessary to update the models to adapt to such a changing environment. To this end, deep learning can use online training algorithms. However, it takes time to train a model with a large number of parameters without high-performance computers. For instance, for the current task, training with the GPU took approximately 45 min for the CNN, while updating the models to adapt to such a changing environment. To this end, deep learning can use online training algorithms. However, it takes time to train a model with a large number of parameters without high-performance computers. For instance, for the current task, training with the GPU took approximately 45 min for the CNN, while updating $W^{out}$ in the SRC was less than 2 s with a single CPU. These results imply that the real-time performance of SRC is more suitable than that of deep learning. In other words, if it is possible to update $W^{out}$ optimally at any particular moment for the prediction of traffic congestion where real-time performance is important and very high accuracy is not required. Thus, the SRC will be very effective. Of course, with the current SRC, it may be difficult to achieve the same level of accuracy of prediction as a more optimized deep learning model. In the future, it is necessary to verify the effectiveness of the SRC model with data from public roads and other highways.

Acknowledgments

This work was partially supported by JST MIRAI No. JPMJMI19B1, and JSPS KAKENHI No. 19K12198 and 17H03280.

References


