

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

- 1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
- 2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
- 3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

A UNIFIED FRAMEWORK FOR PATH TRAVEL TIME PREDICTION USING HETEROGENEOUS TRAFFIC DATA AND WEATHER INFORMATION

LI ANG

PhD

The Hong Kong Polytechnic University 2025

The Hong Kong Polytechnic University Department of Civil and Environmental Engineering

A unified framework for path travel time prediction using heterogeneous traffic data and weather information

Li Ang

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

May 2024

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

Signature: Name of student: Ang LI

Abstract

Advanced Traveler Information Systems (ATIS) usually offer information on travel times for specific paths or routes. They aim to impart timely information to road users and assist them in making route choices under uncertainties in the near future, especially in road networks with frequent adverse weather conditions. Recent research has explored the use of diverse traffic data sources for predicting path travel times in the current and future time intervals. These traffic data include both real-time and historical data from different sources, in which the former is collected on the current day, and the latter is gathered before the current day. There are three challenges integrating them and relevant weather information for path travel time prediction in the current and future time intervals.

First, some traffic data are sampled at high frequency (say, once every 1 or 2 minutes) due to the requirement of practical applications in ATIS. Consequently, the sample size per time interval is insufficient to provide reliable information for removing outliers from real-time data. Moreover, ground truth on path travel times is difficult to collect with high cost and limited samples for field surveys (e.g., floating car surveys). Collecting these ground truths is more suitable for validation than model training.

Second, existing ATIS generally disseminate the predicted average path travel times in the current time intervals for all vehicle classes in reality. However, the observed path travel times of a significant proportion of vehicles (i.e., private cars) may deviate substantially from the average path travel times. It is specifically true when many other vehicles (e.g., buses and goods vehicles) travel with private cars on the same road. There is a need to integrate traffic data to predict multi-class path travel times. Additionally, different traffic sensors may furnish heterogeneous traffic data (e.g., travel time, flow, speed, etc.), which complicates the path travel time prediction for different vehicle classes.

Third, in cities with frequent rainfall, the rainfall intensity can significantly impact the accuracy of travel time predictions. Existing studies have used historical rainfall intensity data to predict path travel times. However, previous studies may not fully

consider the temporal relationships between rainfall intensity data and predicted path travel times. Moreover, less attention has been given to the weather forecast information, which can be further investigated as adverse weather can affect the travel behavior of road users (e.g., departure time and route choices). Addressing the usage of weather information is crucial for improving the performance of ATIS on path travel time prediction under varying traffic and weather conditions.

Based on the above challenges, this thesis seeks to propose a unified framework for path travel time prediction in ATIS, offering the following three key contributions:

Firstly, the proposed unsupervised algorithm is designed to filter limited real-time automatic vehicle identification (AVI) data without relying on ground truth for training. Real-time AVI data can be limited due to the high frequency of collection. Contrastingly, historical data contain adequate information on variations of path travel times on each time interval. This type of variation helps to indicate the typical traffic conditions by different time of day. Therefore, the proposed unsupervised algorithm goes beyond traditional filtering methods (relying purely on real-time AVI data) by incorporating day-to-day variations of path travel times. It consequently offers valuable insights for data filtering, particularly when real-time AVI data is limited.

The second contribution involves the development of a novel model for multi-class path travel time prediction in the current time interval. The proposed prediction model effectively utilizes heterogeneous traffic data from various types of traffic sensors. This prediction model incorporates the temporal relationships of path travel times across different vehicle classes inferred from multi-source traffic data. It allows for the fusion of traffic information from diverse traffic data sources. As a result, the proposed prediction model can provide satisfactory predicted path travel times by vehicle class in the current time interval.

The third contribution arises with a new model that considers weather information to predict path travel times in future time intervals. This thesis proposes a modeling framework to further capture the relationship between predicted path travel times and weather information. Therefore, the proposed modeling framework can help describe the dynamics of predicted path travel times under future rainy conditions. Additionally, the proposed modeling framework distinguishes the effects of weather information under different traffic conditions and various rainfall categories. Hence, it can ultimately enhance the prediction accuracy.

The empirical evidence from real-world traffic data in Hong Kong has demonstrated the effectiveness of the proposed unified framework for path travel time prediction. Three key contributions have been justified with corresponding case studies or numerical experiments in this thesis.

Firstly, the case study conducts sensitivity tests using different sampling rates of AVI data. It reveals that the proposed unsupervised algorithm robustly surpasses the existing filtering algorithms without using ground truth for training.

The second contribution is confirmed using multiple sources of traffic data gathered on an urban expressway in Hong Kong. It shows that the prediction accuracy of path travel times by vehicle class in the current time interval is significantly improved when a proper combination of data sources is selected for training. The proposed prediction model can output the multi-class path travel times with satisfactory performance.

Lastly, the empirical tests illustrate that the proposed modeling framework, considering the weather information, achieves a higher accuracy of predicted path travel times in future time intervals. It outperforms the other benchmarks on a dataset collected in Hong Kong, a city with abundant rainfall throughout the year.

These three significant contributions in this thesis are properly justified to support the proposed unified framework as a valuable platform for further research in the development of various ATIS.

 $(A_{i}, A_{i}) \in \{A_{i}, A_{i}\}$

Publications Arising from the Thesis

Journal papers:

- Li, A., Lam, W.H.K., Ma, W., Chow, A.H.F., Wong, S., Tam, M.L., 2023. Filtering Limited Automatic Vehicle Identification Data for Real-Time Path Travel Time Estimation Without Ground Truth. *IEEE Transactions on Intelligent Transportation Systems* 25, 4849–4861. (Chapter 3 in the thesis) Weblink: https://ieeexplore.ieee.org/document/10337760
- Li, A., Lam, W.H.K., Ma, W., Wong, S.C., Chow, A.H.F., Tam, M.L, 2024. Realtime estimation of multi-class path travel times using multi-source traffic data. *Expert Systems with Applications* 237, 121613. (Chapter 4 in the thesis) Weblink: https://www.sciencedirect.com/science/article/pii/S0957417423021152
- Li, A., Lam, W.H.K., Ma, W., Tam, M.L., 2024. A novel modeling framework for real-time prediction of path travel times using both traffic and weather data. *Transportmetrica A: Transport Science* 1–34. (Chapter 5 in the thesis) Weblink: https://www.tandfonline.com/doi/full/10.1080/23249935.2024.2419499
- Li, A., Lam, W.H.K., Tam, M.L., Zhong, R.X., Ma, W., 2022. Prediction of travel time on urban road links with and without point detectors. *Asian Transport Studies* 8, 100081. In: The 14th International Conference of the Eastern Asia Society for Transportation Studies (EASTS), Japan, September 12-15. Best Paper Award for Methodological Development

Weblink: https://www.sciencedirect.com/science/article/pii/S218555602200027X

Conference papers:

- Li, A., Lam, W.H.K., 2019. Estimation of travel times on urban roads with and without detected data. In: 24th International Conference of Hong Kong Society for Transportation Studies, Hong Kong, China, pp. 4.
- Li, A., Lam, W.H.K., Zhong, R.X., Ma, W. 2021. Estimation of travel times on urban roads with and without detected data. In: 2021 International Symposium on Transportation Data and Modeling, Virtual
- Li, A., Lam, W.H.K., 2022. Multi-class path travel time estimation using multisource traffic sensor data on urban roads. In: 26th International Conference of Hong Kong Society for Transportation Studies, Hong Kong, China, pp. 5.

Acknowledgements

I wish to express my profound appreciation to my chief supervisor, Prof. William H. K. Lam, for his invaluable mentorship, unwavering support, and thoughtful advice during my academic journey in Hong Kong. He steered me towards adopting the appropriate researcher's mindset, encouraging continuous exploration within my field of knowledge. Completing this thesis would have been unattainable without his motivation and wise contributions.

My heartfelt thanks go to Dr. Wei Ma for his generous support on both technical matters and his ongoing guidance on writing. I am also thankful to Prof. Anthony Chen for his enlightening lectures that directed me toward optimization and network modeling studies. I must acknowledge Prof. Renxin Zhong for his constructive advice and selfish help in my research, which I deeply appreciate. I also express my gratitude to Prof. Bi Yu Chen, Prof. Hu Shao, and Dr. Xiao Fu for providing me with insightful suggestions on conducting research. I would like to thank Prof. Seungjae Lee and Dr. Zhiwei Wang, and Dr. Shu-Chien Hsu for serving on my thesis committee and for their valuable suggestions and encouragement.

I owe a debt of gratitude to both my former and current colleagues and friends at The Hong Kong Polytechnic University. This includes, but is certainly not limited to, Dr. Karen Tam, Dr. Julio Ho, Dr. Hao Fu, Dr. Cheng Zhang, Dr. Xiaomeng Shi, Dr. Chaoyang Shi, Dr. Vo Dang Khoa, and Ms. Lu Chen.

Most importantly, I cannot find the right words in any language I know to fully express my gratitude towards my parents for their love and provision.

Table of Contents

Abstract IV	
Publications Arising from the ThesisVII	
AcknowledgementsIX	
Table of Contents X	
List of FiguresXII	
List of TablesXIV	
List of NotationsXVII	
1. Introduction	
1.1 Background1-1	
1.2 Problem Statement1-3	
1.3 Objectives and Scope of the Thesis1-5	
1.4 Thesis Organization1-7	
2. Literature Review	
2.1 Basic Concept2-1	
2.2 Categories of Traffic Sensors for Data Collection2-8	
2.3 Types of Traffic Data for Path Travel Time Prediction2-21	
2.4 Review of Filtering Algorithms for Different Types of Traffic Data	
2-26	
2.5 Review of Path Travel Time Prediction Models2-26	
2.6 Summary	
3. Filtering Limited Automatic Vehicle Identification Data without Ground Truth	
on Path Travel Times	
3.1 Initial Considerations	
3.2 Problems under Investigation	
3.3 Techniques for Filtering AVI Data	
3.4 Case Studies	
3.5 Concluding Remarks	
4. Prediction of Multi-class Path Travel Times Using Multi-source Traffic Data	
4.1 General	
4.2 Research Problems	

4.3 Prediction of Multi-class Path Travel Times
4.4 Numerical Experiments
4.5 Findings
5. Prediction of Path Travel Times Using Weather Forecast and Historical Rainfall
Intensity
5.1 Preliminary Overview
5.2 Research Questions
5.3 Methodology
5.4 Model Validation
5.5 Discussions
6. Conclusions and Further Studies
6.1 Highlights of Main Contributions of the Thesis
6.2 Discussions of Key Findings6-5
6.3 Recommendations for Future Research
Appendix A
Appendix B
Appendix C
References

List of Figures

Figure 1.1 Definitions of instantaneous path travel times and experienced path
travel times1-2
Figure 1.2 Thesis structure1-8
Figure 2.1 A schematic example of various traffic sensing systems commonly
used in ATIS2-25
Figure 3.1 Example of invalid AVI data in an AVI system
Figure 3.2 Framework of the proposed unsupervised algorithm
Figure 3.3 An illustrative example of the rolling horizon scheme
Figure 3.4 Illustration of the JTIS in Hong Kong
Figure 3.5 Overview of the two study paths in Hong Kong in Chapter 3
Figure 3.6 Sampling rates of (a) valid real-time AVI data, and (b) real-time
AVI data on both study paths
Figure 3.7 Filtering performance of U1 and U2 on study path 2 by time of
day
Figure 3.8 Sensitivity test with various sampling rates of real-time AVI data on
study path 2
Figure 3.9 Sensitivity test of sampling rates of valid real-time AVI data on
U1 for study path 1 (left) and study path 2 (right)
Figure 3.10 Comparison of day-to-day variance of the mean of predicted path
travel times from the proposed model using AVI data $Var(\mu_{T^{s_A}}(d))$ against
day-to-day variance of ground truth on mean of path travel times
$Var(\mu_{T^{s_G}}(d))$ by 2-minute intervals on the study path 1
Figure 4.1 Overview of the study path and example of data for one vehicle in
Chapter 4
Figure 4.2 CDF of the sample size of valid AVI and GPS data used in Chapter 4
Figure 4.3 CDF plot of path travel times by vehicle class using AVI data4-7
Figure 4.4 The framework of the proposed prediction model
Figure 4.5 CDF plots of prediction errors for (a) private cars, (b) goods vehicles,
(c) other vehicles, and (d) all vehicles

Figure 4.6 The comparison between predicted path travel times for overall vehicle
class against ground truth on path travel times4-31
Figure 4.7 CDFs of prediction errors for all vehicles with and without considering
the covariance of path travel times between vehicle classes
Figure 5.1 Average annual rainfall (mm) in major Pacific Rim cities5-2
Figure 5.2 CDF of 2-min and hourly rainfall intensity data in 2018
Figure 5.3 The proposed modeling framework for path travel time prediction in
the near future
Figure 5.4 The selected path for the empirical tests5-19
Figure 5.5 CDF of correctness and BS of weather forecast
Figure 5.6 CDF of MAPE for predicting path travel times 1-hour ahead for study
path 1 using different input data5-25
Figure 5.7 Comparison of normalized cross-correlation coefficient of (a) rainfall
intensity data and path travel times $(\rho_{p,r}(t))$; (b) FRA and path travel times
$(\rho_{p,\hat{r}}(t))$
Figure 5.8 The effects of weather forecast correctness on the prediction accuracy
(in terms of MAPE) of path travel times by 2-minute intervals in the next
hour
Figure 5.9 CDF of MAPE for predicting 1-hour ahead on path travel times using
different sources of traffic data as input for study path 15-40
Figure A.1 511 Travel Information (USA)A-2
Figure A.2 Google Maps (Google company)A-3
Figure A.3 JTIS (Hong Kong)
Figure C.1 Data format for AVI dataA-8
Figure C.2 Data format for point sensor data
Figure C.3 Data format for GPS data

List of Tables

Table 1.1 Summary table of different types of temporal covariance of path travel
times in this thesis1-6
Table 2.1 ATIS across the world 2-3
Table 2.2 Common features in ATIS 2-5
Table 2.3 Differences between path travel time, link travel time, and network
travel time2-7
Table 2.4 Various traffic sensors and their practical applications2-9
Table 2.5 Description summary of AVI sensors 2-15
Table 2.6 Summary of different technologies applied to point sensors with their
advantages and limitations2-20
Table 2.7 Summary of different types of traffic data 2-24
Table 2.8 Summary of traditional models for path travel time prediction2-35
Table 2.9 Summary of traditional models for path travel time prediction2-37
Table 3.1 Summary of previous related studies for filtering AVI data in the path
travel time prediction
Table 3.2 Summary of two study paths 3-26
Table 3.3 Comparison of prediction performance with respect to mean/standard
deviation of predicted path travel times
Table 3.4 Comparison of prediction performance on the mean of predicted path
travel times when historical ground truth is used for training3-33
Table 3.5 The average computational time of different filtering algorithms3-34
Table 3.6 The prediction performance of the proposed unsupervised algorithm on
study path 2 under different traffic conditions
Table 3.7 The prediction performance of the proposed unsupervised algorithm
with a fixed sample size of real-time AVI data
Table 3.8 The effects of short-spacing intersections and frontage access on
predicted path travel times
Table 3.9 Sensitivity test with different percentages of historical ground truth
removed in 2017 data on study path 2
Table 3.10 Sensitivity test of the percentage deviation of the results of S1 from
the optimum threshold

Table 3.11 Impacts of prediction horizon (Δt), rolling horizon (δ), and rolling
step (Δ) on prediction accuracy (in terms of MAPE%) for the study path 2
Table 3.12 Prediction performance of the proposed unsupervised algorithm (in
terms of MAPE (%)/MAE (min))
Table 4.1 Traffic characteristics of the study path4-4
Table 4.2 P-values of the Kolmogorov-Smirnov test on path travel times by
vehicle class ($\alpha = 0.05$)4-8
Table 4.3 Recent related studies using the GMM 4-10
Table 4.4 Summary of input and output in the experiments of Section 4.4.2
Table 4.5 Path travel time prediction of the overall class using various data
sources4-25
Table 4.6 Comparison of predicted path travel time by vehicle class
Table 4.7 Benchmark comparison of path travel time prediction models of the
overall class4-33
Table 4.8 Prediction performance of different models on predicted path travel
times by vehicle types
Table 4.9 Benchmark comparison of path travel time prediction models of overall
class on weekends
Table 4.10 Effects of the proportions of AVI, GPS, and point sensor data used for
training on prediction performance of path travel times (in terms of MAPE).
Table 5.1 Related studies utilizing rainfall intensity/weather forecast data for
travel time prediction5-4
Table 5.2 Availability of different input data 5-23
Table 5.3 Prediction performance of proposed modeling framework with different
input data
Table 5.4 Online updating performance of different updating models on study
path 15-26
Table 5.5 Maximum errors of online updating performance of different updating
models on study path 15-26
Table 5.6 Offline training performance of different offline training models on

XV

study path 1
Table 5.7 Maximum errors of offline training performance of different offline
training models on study path 1
Table 5.8 Offline training performance of proposed offline training model using
rainfall intensity data with different frequencies for study path 15-33
Table 5.9 Maximum errors of proposed offline training model using rainfall
intensity data with different frequencies for study path 15-34
Table 5.10 Detailed information of four scenarios
Table 5.11 Prediction performance of the proposed modeling framework on study
path 1 at the time interval $t + 1$ using c1. (in terms of MRPE (%))5-35
Table 5.12 The prediction performance of the proposed modeling framework with
a fixed sample size of real-time AVI data5-37
Table 5.13 Performance of the proposed modeling framework for predicting path
travel times 1-hour ahead using hourly FRA and POP data with different
strategies5-38
Table 5.14 Different modes on effects of rainfall on path travel times5-39
Table 5.15 Online updating performance of proposed modeling framework under
different rainfall categories and levels of service for study paths 1 and 2
Table B.1 Photographs and weblinks for different AVI and point sensorsA-5

List of Notations

Unless otherwise specified, the notations used throughout the thesis are listed below. *Abbreviations*

ALPR	Automatic license p	late recognition		
APTN	Attention-based per	iodic-temporal n	eural network	
ATIS	Advanced Traveler	Information Sys	tems	
AVI	Automatic vehicle i	dentification		
BIC	Bayesian Information	on Criterion		
BS	Brier score			
CDF	Cumulative distribu	tion function		
CNN	Convolutional neura	al network		
DBN	Deep belief network	X		
DSRC	Dedicated short-ran	ge communication	ons	
FPCA	Functional principal	l component ana	lysis	
GMM	Gaussian Mixture M	Iodel		
GPS	Global Positioning	System		
JTIS	Journey	Time	Indication	System
	(https://www.td.gov	v.hk/en/transport	_in_hong_kong/its/its	_achiev
	ements/journey_tim	e_indication_sys	stem_/index.html)	
LOS	Level of service			
LSTM	Long short-term me	emory		
MAC	Media access contro	ol		
MAE	Mean absolute error	ſ		
MAPE	Mean absolute perce	entage error		
MaxAPE	Maximum absolute	percentage error	•	
MaxAE	Maximum absolute	error		
PACE	Principal analysis by	y conditional exp	pectation	
p.d.f.	Probability density	function		
POP	Probability of precip	pitation		
RFID	Radio-frequency ide	entification		
RMSE	Root mean square e	rror		
RNN	Recurrent neural ne	twork		

RTMS	Remote traffic microwave sensors
SMPS	Speed Map Panels System
	(https://www.td.gov.hk/en/transport_in_hong_kong/its/its_achiev
	ements/speed_map_panels/index.html)
V2I	Vehicle-to-infrastructure
V2V	Vehicle-to-vehicle
Sets	
L	Set for rainfall categories, indexed by $l = 1,, L $
S	Set of traffic data from multiple sources (different types of traffic
	sensors)
S_m	m-th category of observed path travel times partitioned by traffic
	conditions, $S = S_1 \cup S_2 \cup \cup S_m \cup \cup S_{ M }$, where M is the
	number of categories determined by statistical distributions of
	observed path travel times for different traffic conditions
M_p	Set of modes that represent the distinguished impacts of forecasted
	rainfall amount (FRA) on predicted path travel times by different
	levels of service and rainfall categories for path p , indexed by
	$m_p = 1, \dots, \left M_p \right $
D	The set of days with historical AVI data and historical ground truth
	on path travel time (if it is available), indexed by $d = 1, 2,, D $
K _s	Set of vehicle classes that can be available from data source s_A
	(AVI sensors) with the number of vehicle classes $ K_{s_A} $ that can be
	available from data source s_A
K_{s_B}	Set of vehicle classes that can be available from data source s_B
	(GPS sensors)
LOS_p	Set for levels of service for describing the traffic conditions path
	p
ψ	Set of parameters in the Gaussian mixture model
Variables used i	n model formulation
p	Path <i>p</i> monitored by traffic sensors
М	Number of categories determined by statistical distributions of

observed path travel times for different traffic conditions

S _A	Data source from AVI sensors; $s_A \in S$
S _B	Data source from GPS sensors; $s_B \in S$
S_W	Data source from point sensors; $s_W \in S$
S _G	Data source from SMPS or JTIS (i.e., ground truth used in the
	thesis); $s_G \in S$
τ	Timestamp of observations collected from traffic sensors
d	Historical day with traffic data or weather information
t	Time interval
t_0	Current time interval
U(t)	Upper bound of validity window for filtering real-time AVI data
	at time interval t
L(t)	Lower bound of validity window for filtering real-time AVI data
	at time interval t
δ	Length of study horizon where traffic data/weather information on
	the current day are modeled/considered
Δ	Rolling step or updating interval or prediction step in the rolling
	horizon scheme, which is the forward step after filtering or
	prediction is performed
$\mathcal{Y}_{i,d,p}^{s_A,k}$	<i>i</i> -th observed path travel time for path p of vehicle class k for
	from data source s_A on day d
$\tau^{s_A,k}_{i,d,d_p,p}$	Timestamp of the <i>i</i> -th observed path travel time at AVI sensor
r -	location d_p for path p of vehicle class k from data source s_A on
	day <i>d</i>
$\varepsilon_{i,d,p}^{s_A,k}$	Measurement error of i -th observation for path p of vehicle class
	k from data source s_A on day d
$T^{s_A}(d)$	Path travel time function from s_A as function of day d
$\mu_{T^{s_A}}(d)$	Mean of path travel time function from s_A as a function of day d
$T^{s_A}(t)$	Path travel time function from s_A as function of time interval t
$\mu_{T^{s_A}}(t)$	Mean of path travel time function from s_A as a function of time
	interval t
$K^{s_A,D}$	Number of functional principal components from s_A for $ D $ days
$\phi_k^{s_A,D}(d)$	Eigenfunction of k^{th} functional principal component from s_A from

set D as a function of day d

$\xi_k^{s_A,D}$	Score/weight of k^{th} functional principal component from s_A from
	set D
$\Sigma_t^{s_A}(d_i, d_j)$	Day-to-day covariance of path travel times at time interval t from
	data source s_A between day d_i and d_j , for $i, j \in D$
$\lambda_k^{s_A,D}$	Eigenvalue of k^{th} functional principal component from s_A from set
n.	D
N _{d,\delta}	Number of samples within the study horizon δ on day d
κ_{s_A}	Kernel function for calibrating covariance function of path travel
	time from data source s_A
h _{SA}	Bandwidth for calibrating covariance function of path travel time
	from data source s_A
D^*	Set of days after sample selection, indexed by $d = 1, 2,, D^* $
H^*	Maximal threshold of the path travel time covariance between
	different days used in sample selection
$\Sigma_d^{s_A}(t_a,t_b)$	Within-day covariance of path travel times on day d from data
	source s_A between time interval t_a and t_b , for $a, b \in \delta$
$\lambda_k^{s_A,\delta}$	Eigenvalue of k^{th} functional principal component from s_A during
	rolling horizon with length δ , where traffic data on the current day
	are considered
$\phi_k^{s_A,\delta}(t)$	Eigenfunction of k^{th} functional principal component from s_A at
	time interval t during rolling horizon length δ , where traffic data
	on the current day are considered
$\xi_k^{s_A,\delta}$	Score/weight of k^{th} functional principal component from s_A
	during rolling horizon with length δ , where traffic data on the
	current day are considered
$K^{s_A,\delta}$	Number of functional principal components from s_A during
	rolling horizon with length δ , where traffic data on the current day
	are considered
$T^{s_G}(t)$	Ground truth on path travel time from data source s_G at time
	interval t
$\mu_{T^{s_{G}}}(t)$	Mean of ground truth on path travel time from data source s_G at

time interval t

$Var(\mu_{T^{s_A}}(d))$	Variance of the predicted mean of path travel times by 2-minute
	intervals on day d from data source s_A

- $Var(\mu_{T^{s_G}}(d))$ Variance of ground truth on the mean of path travel times by 2minute intervals on day *d* from data source s_G
- $\Sigma_d^{s_G}(t_a, t_b)$ Covariance of ground truth on path travel times on day d from data source s_G between time intervals t_a and t_b , for $a, b \in \delta$
- $\xi_k^{s_G}$ Score/weight of k^{th} functional principal component from data source s_G
- $\phi_k^{s_G}(t)$ Eigenfunction of k^{th} functional principal component from data source s_G at time interval t
- $\lambda_k^{s_G}$ Eigenvalue of k^{th} functional principal component from data source s_G
- $K^{s_G,\delta}$ Number of functional principal components from data source s_G during rolling horizon with length δ , where traffic data on the current day are considered
- κ_{s_G} Kernel function for calibrating conditional function of ground truth on path travel time from data source s_G
- h_{s_G} Bandwidth for calibrating conditional function of ground truth on path travel time from data source s_G
- o_p Origin of path p
- d_p Destination of path p
- x^{o_p} Location of the origin o_p of path p

 x^{d_p} Location of the destination d_p of path p

 $x_{i,p}$ Location of *i*-th nearby weather station or point sensor for path p

$$r_p(x,t)$$
 The rainfall intensity data on location x along path p at time interval t

- $\hat{r}_p(x, t_0 + \Delta t)$ Forecasted rainfall amount for location x at time interval $t_0 + \Delta t$ along path p
- Δt Time ahead of current time interval t_0 (prediction horizon)
- $C_{\hat{r}_p(x,t_0+\Delta t)} \qquad \text{The correctness of FRA } \hat{r}_p(x,t_0+\Delta t), \text{ which is FRA for location}$ x at time interval $t_0 + \Delta t$ along path p

$\widehat{P}_{p,l}(x,t_0$	Probability of precipitation (POP) for forecasting rainfall category	
$+ \Delta t$)	<i>l</i> for location <i>x</i> along path <i>p</i> at time interval $t_0 + \Delta t$	
$o_{p,l}(x,t)$	Observed frequency of rainfall at rainfall category l for location x	
	along path p at time interval t	
$C_{\hat{P}_{p,l}(x,t_0+\Delta t)}$	The correctness of POP $\hat{P}_{p,l}(x, t_0 + \Delta t)$, which is POP for	
	forecasting rainfall category l for location x along path p at time	
	interval $t_0 + \Delta t$	
$\hat{T}_{t_0+\Delta t,p}$	Predicted path travel times for path p at time interval $t_0 + \Delta t$	
$\tilde{T}_{t+\Delta t,p}$	Offline predicted path travel times of path p for Δt ahead of time	
	interval t	
$n_{t_0,p}$	Sample size of individual path travel times collected at current	
	time interval t_0 for path p from data source s_A	
N_p	Number of nearby rainfall stations for path <i>p</i>	
$\lambda_{i,p}$	Kriging weight for i -th nearby rainfall station for path p in the	
	kriging model for $r_p(x, t)$ (i.e., the rainfall intensity data on	
	location x along path p at time interval t)	
$D(x_{i,p})$	External drift for i -th nearby rainfall station for path p in the	
	kriging model for $r_p(x, t)$ (i.e., the rainfall intensity data on	
	location x along path p at time interval t)	
μ_1, μ_2	Lagrange parameters for spatial interpolation accounting for two	
	constraints on λ_i in the kriging model for $r_p(x, t)$ (i.e., the rainfall	
	intensity data on location x along path p at time interval t)	
γ	Similarity between rainfall intensity data of two locations in the	
	kriging model for $r_p(x,t)$ (i.e., the rainfall intensity data on	
	location x along path p at time interval t)	
$LOS_{i,p}$	<i>i</i> -th level of service for describing the traffic conditions on the	
	road, $LOS_{i,p} \in LOS_p$ for path p	
$v_{f,p}$	Free-flow travel speed for path <i>p</i>	
$\delta_{LOS_{i,p}}$	The ratio between $v_{f,p}$ (free-flow travel speed for path p) and	
	$v_{LOS_{i,p}}$ (the threshold of average path speed for $LOS_{i,p}$ for path p)	
AC(x,t)	Rainwater accumulation at location x at time interval t	

XXII

rd(x)	Rate of rainwater drainage on location x	
$AC_0(x)$	Threshold of rainwater accumulation at location x	
<i>S</i> ₁	Data source of $r_p(x, t)$ for $t = t_0$	
\tilde{S}_1	Data source of $r_p(x, t)$ for $t < t_0$	
<i>S</i> ₂	Data source of $T_{t,p}$ for $t = t_0$	
\tilde{S}_2	Data source of $T_{t,p}$ for $t < t_0$	
<i>S</i> ₃	Data source of $\hat{P}_{p,l}(x, t_0 + \Delta t)$ for $t = t_0$	
<i>S</i> ₄	Data source of $\hat{r}_p(x, t_0 + \Delta t)$ for $t = t_0$	
$y_{i,d,p}^{\{s_A,s_W\},1}$	<i>i</i> -th element of vectors of observed path travel times $\boldsymbol{Y}_{p}^{\{s_{A},s_{W}\}}$ with	
	vehicle class 1 from data sources s_A and s_W for path p on day d	
BIC_{model}	BIC value of the clustering model	
\hat{L}_{model}	The maximized value of the likelihood function for the clustering model	
n	Sample size of multi-source traffic data used in the clustering model	
π_m	Probability of the m -th category by traffic conditions (e.g., probability of free-flow condition when $m = 1$)	
$v_{x,t}$	Average spot speed at location x in time interval t from data source s_c	
$z_i^k(t+j\Delta g)$	Trajectory of the <i>i</i> -th vehicle of class <i>k</i> : where Δg is the sampling time interval of data source s_B , $j = 1, 2,, J$ is the index of sampling points for an individual vehicle trajectory within path <i>p</i> , with the corresponding location $x_{i,j}^k$, speed measurement $v_{i,j}^k$ and timestamp $\tau_{i,j}^k$	
BS	Brier score used to measure the accuracy/correctness of probabilistic predictions (Wu et al., 2019; Zhu et al., 2022a)	
$P_{i,j,p}$	Transitional probability from mode <i>i</i> to <i>j</i> for path $p, i, j \in M_p$	
$w_{t,p}^{i,j}$	Normalized probability for mode i to j for path p at time interval	
	$t, \iota, j \in M_p$	
$c_{t,p}^{J}$	Normalization factor for mode j for path p at time interval $t, j \in M_p$	

XXIII

$\hat{P}_{p,l}(x,t_0)$	The probability of precipitation for forecasting rainfall category l		
$+ \Delta t$)	for location x along path p at time interval $t_0 + \Delta t$ along path p		
	with a certain probability of transition from mode i to mode j at		
	time interval $t_0, i, j \in M_p$		
$r_p(t)$	Average rainfall intensity for path p at time interval t		
$T_{p,j}^{s_A}(t)$	Historical path travel time for path p from data source s_A in		
	function of time interval t under mode $j, j \in M_p$		
$\mu_{p,j}$	Expected rate of change of path travel times over time for path p		
	under mode $j, j \in M_p$		
$\sigma_{p,j}$	The noise of path travel times for path <i>p</i> under mode $j, j \in M_p$		
ϵ_p	Relaxation rate to the historical mean of rainfall intensity for path		
	p		
$\rho_{p,r}(t)$	The normalized cross-correlation coefficient of path travel times		
	and rainfall intensity data for path p		
$\rho_{p,\hat{r}}(t)$	The normalized cross-correlation coefficient of path travel times		
	and FRA for path p		
β_m	Coefficient of the linear effect within category m		
Parameters			
θ	Frequency of rainfall intensity data (i.e., 2 minutes used in the		
	thesis)		
$P_{\boldsymbol{Y}_{p}^{\{s_{A},s_{W}\}}}$	Dimension of vector of observed path travel times $\boldsymbol{Y}_{p}^{\{s_{A},s_{W}\}}$ (with		
	vehicle classification) for path p from data sources s_A and s_W		
$D_{\boldsymbol{T}_{p}^{\{s_{A},s_{W}\}}}$	Dimension of vector of predicted path travel time $T_p^{\{s_A, s_W\}}$ for path		
	p based on data sources s_A and s_W		
$ ho_0$	Threshold for $\rho_{p,r}(t)$ (normalized cross-correlation coefficient of		
	path travel times and rainfall intensity data for path p)		
npar _{model}	The number of parameters in the clustering model		
$v_{LOS_{i,p}}$	The threshold of average path speed for $LOS_{i,p}$ for path p		
Vectors and matrices			
$\boldsymbol{Y}_p^{\{s_A,s_W\}}$	Vector of observed path travel times (with vehicle classification)		

for path p from data sources s_A and s_W with dimension of

 $P_{T_p^{\{s_A,s_W\}}}$

$\boldsymbol{Y}_{t,p}^{\{s_A,s_W\}}$	Vector of observed path travel times (with vehicle classification)			
	for path p from data sources s_A and s_W for time interval t			
$T_p^{\{s_A,s_W\}}$	Vector of predicted path travel time for path p based on data			
	sources s_A and s_W .			
$T_{p,t}^{\{s_A,s_W\}}$	Vector of predicted path travel time for path p based on data			
	sources s_A and s_W . for time interval t			
$\boldsymbol{\mu}_{\boldsymbol{Y}^{\{s_A,s_W\}}}$	Mean of travel time measurements (with vehicle classification)			
- <i>m,p</i>	$\boldsymbol{Y}_{m,p}^{\{s_A,s_W\}}$ from set $\{s_A, s_W\}$ within category <i>m</i> for path <i>p</i>			
$\boldsymbol{\Sigma}_{\boldsymbol{V}^{\{s_A,s_W\}}}$	Covariance of travel time measurements (with vehicle			
" m,p	classification) $\boldsymbol{Y}_{m,p}^{\{s_A,s_W\}}$ from set $\{s_A,s_W\}$ within category m for			
	path <i>p</i>			
$\mu_{\widehat{T}_{m,p}^S}$	Mean of predicted path travel time (with vehicle classification)			
·/r	$\widehat{T}^{S}_{m,p}$ within category m for path p based on set S			
$\boldsymbol{\varphi}_{P_{\boldsymbol{Y}_{p}^{\{s_{A},s_{W}\}}}}$	Gaussian random vector with $P_{Y_p^{\{s_A, s_W\}}}$ variates			
$oldsymbol{arphi}_{p}_{T_{p}^{\{s_{A},s_{W}\}}}$	Gaussian random vector with $D_{T_p^{\{s_A, s_W\}}}$ variates			
$T_p^{\{s_A,s_W\}}(t)$	Vector of path travel times (with vehicle classification) at time			
r	interval t for path p from data sources s_A and s_W			
$T_{p}^{\{s_{A},s_{W}\},+}(t)$	Vector of the posterior estimate (with vehicle classification) of the			
r	mean of path travel time at time interval t for path p from data			
	sources s_A and s_W			
$T_{p}^{\{s_{A},s_{W}\},-}(t)$	Vector of the prior estimate (with vehicle classification) of the			
·	mean of path travel time at time interval t for path p from data			
	sources s_A and s_W			
$P_{p}^{\{s_{A},s_{W}\},+}(t)$	Matrix of the posterior estimate (with vehicle classification) of the			
$\sum_{k=1}^{\{s_A, s_W\}, +} (t)$	within-day/day-to-day covariance at time interval t for path p			
p (c)	from data sources s_A and s_W			
$\boldsymbol{P}_p^{\{s_A,s_W\},-}(t)$	Matrix of the prior estimate (with vehicle classification) of the			
$\Sigma_{m}^{\{s_A,s_W\},-}(t)$	within-day/day-to-day covariance at time interval t for path p			
<i>i</i> - <i>p</i> (*)	from data sources s_A and s_W			

$\Sigma_{t,d}^{s_A,+}(k,k')$	Matrix of the posterior estimate of the covariance of observed path
	travel times between vehicle classes k and k' at time interval t for
	path p from data sources s_A
$\Sigma^{s_{A},-}_{t,d}(k,k')$	Matrix of the prior estimate of the covariance of observed path
	travel times between vehicle classes k and k' at time interval t for
	path p from data sources s_A
G ₁	Updating matrix of estimates of mean and covariance of path
	travel times (with vehicle classification)
G ₂	Updating matrix of estimates of within-day and day-to-day
	covariance of path travel times of the same vehicle class
G ₃	Updating matrix of estimates of covariance of path travel times
	between vehicle classes
\boldsymbol{x}_p	State vector for $\hat{r}_p(x, t_0 + \Delta t)$, $r_p(x, t_0)$, and $T_{t_0, p}$ for path p
μ_p	Vector of mean of state vector $\boldsymbol{x}\boldsymbol{x}$ for path p
$oldsymbol{Q}_p$	Covariance matrix of state vector xx for path p
$\widehat{x}_{t,p}^{+0,j}$	Mixed initial vectors for state vector xx at time interval t for
	mode <i>j</i> for path $p, j \in M_p$
$\widehat{\pmb{x}}_{t,p}^{+i}$	Mixed vectors for state vector xx at time interval t for mode i for
	path $p, i \in M_p$
$\widehat{\boldsymbol{Q}}_{t,n}^{+0,j}$	Mixed initial covariance matrix of state vector xx at time interval
0)2	t for mode j for path $p, j \in M_p$
K	Kalman gain in the Kalman filter
$\boldsymbol{z}_{t,p}$	Observation vector for path travel times at time interval t for path
	p
Н	Operator for updating vector and covariance matrix
R	Matrix for errors in the system
$\boldsymbol{\xi}_{T_{p,i}^{s_A}}(t_i)$	<i>i</i> -th eigenfunction vector of $T_{p,j}^{s_A}(t)$, which is historical path travel
1.9	time for path p from data source s_A in function of time interval t
	under mode $j, j \in M_p$
$\boldsymbol{\xi}_{r_p}(t_i)$	<i>i</i> -th eigenfunction vector of $r_p(t)$, which is average rainfall
	intensity for path p at time interval t

XXVI

1. Introduction

1.1 Background

The rapid development of advanced traveler information systems (ATIS) assists travelers in understanding traffic conditions promptly to make better travel choices (Tam and Lam, 2008). Travel time plays a critical role in ATIS. It is used as a metric to evaluate the effectiveness and efficiency of transportation networks and disseminated through ATIS to travelers. Path travel time of vehicles is the focus of this thesis, where a path is defined as an alternating sequence of nodes and links connecting an origin and destination pair of interest in a road network. In order to help road users choose the best route for their travel, the ATIS could provide traffic information, such as the predicted path travel times in the current and/or future time intervals (Mori et al., 2015).

There are two types of path travel times: experienced path travel times and instantaneous path travel times. The former is the sum of travel time on each segment along the path at the same time interval (Zhong et al., 2017). The latter is the actual travel time of a vehicle and cannot be measured until a traveler completes the trip. Figure 1.1 gives the trajectory of a vehicle and distinguishes these two terms.

Experienced path travel time is the actual, realized travel time that a vehicle could experience along the path. In Figure 1.1, the experienced path travel time of the vehicle is 8 min. The instantaneous path travel time is the sum of travel time on each road segment at the same 2-minute interval. It is not the actual travel time of any vehicle traveling on the path. For example, the instantaneous path travel time at 8:30-8:32 is 6.5 minutes, which is the sum of 2 minutes, 3 minutes, and 1.5 minutes on each road segment, respectively. It can be updated regularly (e.g., every 2 minutes) and is easier to disseminate via ATIS to road users. It is noted that the path travel times mentioned in this thesis are instantaneous path travel times for real-time ATIS applications (Tam et al., 2008; Tam and Lam, 2011).



Figure 1.1 Definitions of instantaneous path travel times and experienced path travel times

Prediction of instantaneous path travel time can be focused on the current time interval (t_0) and future time intervals $(t_1, t_2, ...)$. If t_0 is 8:30-8:32 in Figure 1.1, the research output in this thesis is the predicted path travel time $\hat{T}_{t_0,p}$ on path p in the current time interval t_0 , and predicted path travel time on path p in the future time interval (e.g., $\hat{T}_{t_1,p}$ at the future time interval t_1). The former should be close to 2+3+1.5=6.5 min, while the latter should be close to 2.1+3.4+1.5=7 min. These instantaneous path travel times are to be predicted in this thesis.

This thesis focuses on predicting instantaneous path travel time in the current and future time intervals. This topic has been studied by both academics (Chen et al., 2020; Zhong et al., 2017) and commercial companies such as Google (Derrow-Pinion et al., 2021) as a state-of-practice method for providing instantaneous traffic information in ATIS (Chen and Rakha, 2014). However, the prediction accuracy still needs improvement, especially with different sources of traffic data (Bai et al., 2018; Liu et al., 2017; Xu and Liu, 2021; Sheng et al., 2023).

Path travel times of vehicles are of paramount importance to both road users and transportation management authorities. For road users, accurate and reliable path travel times can benefit trip planning and route selection, estimating their arrival times, real-time navigation, etc. From the perspective of transportation management authorities, it is crucial to understand the path travel times, as they are key components of traffic management, infrastructure planning, and policy-making. The path travel times can directly expose the bottlenecks and congested areas in the road networks. With the path travel time information, transportation management authorities could select measures and plan more proficiently to alleviate traffic congestion and improve the performance of road networks.

In summary, predicting path travel times in the current and future time intervals is an essential task for ATIS and has been studied for over two decades (Yang and Qian, 2019). They rely on collecting multi-source traffic data, including both real-time and historical data. The former is obtained on the current day, and the latter is collected before the current day. Both are to be integrated and utilized in this thesis.

1.2 Problem Statement

This thesis studies the instantaneous path travel time prediction problems for ATIS using multiple sources of traffic data and weather information. Three research problems are investigated here. The first one is effectively filtering traffic data to predict path travel times in current time intervals. The second problem is multi-class path travel time prediction in current time intervals. The third problem is integrating with weather information for path travel time prediction in future time intervals. The following are detailed descriptions of each of these three problems.

An efficient and effective filtering algorithm is urgently needed for path travel time prediction in the context of ATIS. When ATIS provide low-frequency predicted path travel time (say, per 30 minutes or 1 hour), the sample sizes of traffic data at comparatively low frequency should be sufficient for path travel time prediction.

However, the traffic data may be insufficient when the frequency of traffic data is higher. It is reported (Dion and Rakha, 2006) that two or three observations of automatic vehicle identification (AVI) data could be collected every 2 minutes in the metropolitan area (i.e., San Antonio). For the case in Hong Kong, the sampling rate of AVI data per 2-minute interval is even lower. Only AVI data from commercial vehicles are allowed to be collected due to privacy issues in Hong Kong. Besides, low traffic demand during some periods (e.g., midnight) also contributes to a lower sample size. As outliers are to be removed, it is quite challenging to effectively filter out outliers and predict path travel times using limited AVI data.

The predicted path travel time provided by existing ATIS is generally the average path travel times (or the mean) for the road users. However, a noticeable proportion of drivers (e.g., drivers of private cars) may experience another path travel time, which deviates significantly from the average path travel times. Therefore, it is demanding to predict path travel times by different vehicle classes. As single-source traffic data is inadequate for this challenging task, multi-source traffic data should be considered. Furthermore, vehicle class information is not always available for all traffic data sources at high frequency. Therefore, it is difficult to efficiently integrate these traffic data for multi-class path travel time prediction.

The predicted path travel times are greatly needed by road users from ATIS. The typical models for travel time prediction analyze the traffic data collected in the present and past. However, non-traffic data should be further utilized to improve the prediction accuracy of existing prediction models. For areas with frequent rainfall, the weather information can be useful for path travel time prediction. It is worthwhile investigating the relationships between weather information and predicted path travel times.

1.3 Objectives and Scope of the Thesis

This thesis endeavors to establish a comprehensive framework designed for predicting path travel times in the current and future time intervals. It integrates heterogeneous traffic data streams and meteorological information for ATIS, as depicted in Section 1.1. For a path in the research problem presented in Section 1.2, the path travel times are correlated with relevant variables relating to time.

The temporal correlation relationships include: 1) path travel times in different time intervals within the same day, 2) path travel times in the same time interval on different days, 3) path travel times by different vehicle classes, and 4) path travel times and weather information. The first three temporal correlation relationships are captured through temporal covariance in this thesis for predicting path travel times, while the fourth is modeled by introducing the cross-correlation coefficient.

This thesis describes three types of temporal covariance of path travel times by 2minute intervals. Table 1.1 gives a summary of these types for clarification. The first one is the **within-day covariance** of path travel times ($\Sigma_d^{s_A}(t_a, t_b)$), which is withinday covariance of path travel times on day d from data source s_A between time interval t_a and t_b , for $a, b \in \delta$). The second one is the **day-to-day covariance** of path travel times ($\Sigma_t^{s_A}(d_i, d_j)$), which is day-to-day covariance of path travel times at time interval t from data source s_A between day d_i and d_j , for $i, j \in D$). The third one is the temporal covariance of path travel times **between vehicle class** (($\Sigma_{t,d}^{s_A}(k, k')$), which is temporal covariance of path travel times between vehicle class k and k' at time interval t on day d from data source s_A , for $k, k' \in K_{s_A}$).

Considering the temporal covariance of path travel times by 2-minute intervals is essential because it provides a clear explanation for the variations in travel times with physical meanings. The temporal covariance can effectively capture the strong correlations between travel times under recurrent traffic conditions across specific periods (e.g., 8:00-8:02 am) but on different weekdays.

Temporal covariance of path travel times*	Notation	Description
1. within-day covariance	$\Sigma_d^{s_A}(t_a,t_b)$	within-day covariance of path travel times on day <i>d</i> from data source s_A between time interval t_a and t_b , for $a, b \in \delta$
2. day-to-day covariance	$\Sigma_t^{s_A}(d_i,d_j)$	day-to-day covariance of path travel times at time interval <i>t</i> from data source s_A between day d_i and d_j , for $i, j \in D$
3. temporal covariance between vehicle class	$\Sigma_{t,d}^{s_A}(k,k')$	temporal covariance of path travel times between vehicle class k and k' at time interval t on day d from data source s_A , for $k, k' \in K_{s_A}$

 Table 1.1 Summary table of different types of temporal covariance of path travel

 times in this thesis

*The descriptions of 2-minute intervals are omitted for simplicity.

These relationships indicate the interdependence and interactions between various traffic parameters and weather conditions that can significantly influence travel times. The thesis comprehensively considers these relationships for predicting path travel times in the current and future time intervals. The scope of this study fits in with the operational aspect in that the predicted travel times of the path studied at the current and future time intervals can be acquired for ATIS.

The objectives of research in this thesis are:

- To effectively capture and model various temporal covariance of path travel times by 2-minute intervals, advanced statistical and machine learning techniques will be employed. A comprehensive framework is proposed to explore the potential of various modeling approaches to accurately represent the stochastic nature of path travel times.
- 2. The integration of multi-source data presents a unique opportunity to enhance the granularity and robustness of path travel time prediction. Traffic data sources may include, but are not limited to, point sensors (referring to as video-

based cameras in this thesis), navigation satellite system sensors (e.g., Global Positioning System (GPS) sensors), and AVI sensors. Concurrently, weather information, such as rainfall intensity data and weather forecasts, will be incorporated to account for their impact on traffic conditions.

- 3. The temporal covariance of path travel times by 2-minute intervals listed in Objective 1 among these data sources given in Objective 2 will be meticulously analyzed to understand the extent to which each variable contributes to the variability in path travel times. By identifying and quantifying these relationships, the framework aims to improve the precision of path travel time predictions, particularly under varying traffic and weather conditions.
- 4. The framework will provide timely and actionable information to traffic managers and road users through ATIS. The predicted travel time for a path of interest in the current time interval (studied in Chapters 3 and 4) and future time intervals (investigated in Chapter 5) can facilitate more efficient traffic management and enhance the travel experience through informed decision-making.

1.4 Thesis Organization

This thesis consists of six chapters, the structure of which is shown in Figure 1.2. Chapter 1 provides a brief introduction and the objectives of the thesis. Chapter 2 extensively reviews the previous literature on path travel times, traffic data, data filtering, and path travel time prediction models.

The core of this thesis includes Chapters 3, 4, and 5. Firstly, traffic data should be filtered after collection. Chapter 3 investigates the filtering of limited AVI data and effectively uses them for path travel time prediction in the current time interval. Chapter 4 extensively considers the multi-source traffic data for predicting path travel times by vehicle class in the current time interval. In addition, Chapter 5 incorporates non-traffic data (i.e., weather information), including weather forecasts and historical rainfall intensity data, to predict path travel times in future time intervals. It integrates with non-traffic data sources to improve prediction accuracy. Finally, a conclusion and suggestions for further study are given in Chapter 6.



Figure 1.2 Thesis structure

The interrelationships between these chapters are presented as follows.

• Chapter 1 of this thesis lays the necessary background and context for the three research problems being addressed in this thesis. It outlines the research problem, objectives, and scope of the thesis, which will be illustrated in the subsequent chapters. Additionally, this chapter gives the structure and

organization of the thesis to provide a better understanding of the remaining chapters.

- In Chapter 2, a comprehensive overview of previous works related to Chapters 3, 4, and 5 is shown through a comprehensive literature review. This review helps to identify the relevant research gaps and challenges in existing literature. It also highlights the motivation for the research that will be introduced in Chapters 3, 4, and 5. The contributions of Chapters 3, 4, and 5 are distinguished based on the literature review presented in Chapter 2.
- In Chapter 3, a novel filtering algorithm is displayed for limited AVI data. This algorithm is essential during the data preprocessing stage for accurately predicting path travel time in the current time interval, as illustrated in Chapter 4.
- Chapter 4 demonstrates a modified model for predicting multi-class path travel time using multi-source data. The data used in Chapter 4 requires data preprocessing using the methodology explained in Chapter 3. The resulting predicted path travel times serve as the foundation for addressing the path travel time prediction in future time intervals considering weather information, which will be tackled in Chapter 5.
- Chapter 5 focuses on using weather information to improve path travel time prediction in future time intervals. This prediction model is further extended based on the results obtained from Chapters 3 and 4.
- Chapters 3, 4, and 5 model different types of temporal relationships of path travel times, including within-day and day-to-day covariance of path travel times and path travel time covariance between different vehicle classes.
- Chapter 6 provides concluding remarks for the above chapters and summarizes the contributions of the core Chapters 3, 4, and 5. Finally, recommendations for further study related to Chapters 3, 4, and 5 are given in Chapter 6.
2. Literature Review

2.1 Basic Concept

2.1.1 ATIS

ATIS have been developed with technological advances in collecting and disseminating real-time traffic information (Mori et al., 2015). ATIS diffuse necessary messages to road users and administrators for decision-making. It disseminates real-time information, including traffic and weather conditions, alternative paths or routes, temporary road works, on-street parking, etc. (Ng et al., 1995).

Table 2.1 gives a summary of ATIS across the world. Table 2.1 shows that the deployment of ATIS is widespread, particularly for smart city development. Moreover, the predicted path travel times (instantaneous travel time) are key information available across all ATIS (Lee et al., 2006), as listed in Table 2.1. The existing prediction methods from commercial companies (e.g., Google) will also be compared with the proposed prediction model in Chapter 4.

Appendix A further provides detailed information on typical ATIS across the world. In Appendix A, it is found that accurate predicted path travel times in the current and future time intervals are of urgent need for ATIS in terms of offering both route guidance and an overview of traffic conditions. This is one of the motivations behind this thesis. Besides, it is also observed that these ATIS only provide the average predicted path travel times while there is a research gap to predict multi-class path travel times, which will be further elaborated in Section 4.1.1. It is also seen that weather information is available in some ATIS (e.g., 511 Travel Information in the USA). It may be worthwhile integrating them to improve the performance of predicted path travel times, as explained in detail in Section 5.1.1.

In addition, some ATIS have weather information (including rainfall intensity and weather forecast, as shown in Appendix A) that may help road users make their travel

choices properly. This indicates that non-traffic information, like adverse weather, can also affect traffic conditions and travel behavior. This thesis also considers the effects of weather information on path travel time prediction.

Table 2.1 ATIS across the world

Service providers	Name of ATIS	Related websites	Features	Technology used	Coverage area
USA	511 Travel Information	https://ops.fhwa.dot.gov /travelinfo/about/about5 11.htm	Instantaneous path travel times, weather, route planning	Telephone, Web, Mobile apps	Nationwide, with regional variations
Google company	Google Maps	https://www.google.co m/maps/	Instantaneous path travel times, route planning	Web, Mobile apps	Worldwide
Canada	Traveler Information Services	https://travel.gc.ca/trave lling/advisories	Traffic updates, border wait times, road conditions	Web, Mobile apps	Nationwide, with provincial systems
South Korea	Hi-pass	https://www.hipass.co.k r/main.do	Toll collection, traffic information	Radio-frequency identification, Mobile apps	Nationwide
Japan	Vehicle Information and Communication System	https://www.vics.or.jp/e n/	Traffic congestion, accidents, parking	Radio, Infrared beacons, navigation	Major urban areas

				systems	
Australia	Live Traffic NSW	https://www.livetraffic.	Traffic incidents,	Web, Mobile	State-specific
	(New South Wales)		Traffic conditions,		
China	eMapGo	http://www.emapgo.co m.cn/	route planning, navigation	web, Mobile apps	Nationwide
India	iTraffic	https://www.indiamart.c om/itrafficexports/profil e.html	Traffic alerts, route planning, congestion maps	Mobile apps, SMS	Major cities
Brazil	Companhia de	https://www.cetsp.com.	Traffic conditions,	Web, Mobile	São Paulo and
DIazii	Engenharia de Tráfego	br/	incidents	apps	other major cities
South Africa	South African National Roads Agency Limited	https://www.nra.co.za/	Traffic updates, road conditions, toll information	Web, Mobile apps	Nationwide
Hong Kong	Journey Time Indication System (JTIS)	https://www.hkemobilit y.gov.hk/en/traffic- information/live/jt	Instantaneous path travel times	Roadside markers, web, mobile apps	Major routes of the urban area

Table 2.2 lists the common features of ATIS. Travel time is the most popular feature as it is intuitive and easily understood by road users. Adverse weather significantly affects traffic conditions and influences path travel times. Hence, weather information usually raises significant concerns among road users. The thesis mainly focuses on the methodology for travel time prediction, considering different issues in practice. The rest of the other variables are recommended for investigation in further study.

Features	Description
Path travel times	Travel time of vehicles to traverse a path, which is an alternating sequence of nodes and links connecting an origin and destination pair of interest in a road network.
Weather	Updates on weather patterns influencing the road network, such
conditions	as tog, rain, snow, and ice.
Traffic	Real-time data on traffic flow, congestion levels, and vehicle
conditions	speeds on various road segments.
Incident reports	Information on accidents, road closures, construction work, and
mendent reports	other events that may impact travel.
Road conditions	Status of the road surface, including potholes, ice, or debris that
Road conditions	could alter traffic conditions.
Route guidance	Recommendations for the best routes to take, considering
Route guidance	current traffic conditions and user preferences.
Parking	Availability of parking spaces at destinations or along the route,
information	including pricing and restrictions.
Public transit	Schedules, routes, and service status for buses, trains, and other
information	public transportation options.
Toll	I anothing of talls according to a discussion of a strength and a
information	Locations of tons, associated costs, and payment options.
Fuel prices	Information on fuel prices at different service stations along the route.
Charging	For electric vehicles, the locations and availability of charging
station locations	stations.

Table 2.2 Common features in ATIS

Rest areas and	Locations of rest areas, restaurants, restrooms, and other
amenities amenities along the route.	
Emergency	Information on the nearest hospitals, police stations, and
services	roadside assistance services.
Troval alarta	Notifications about significant events that could impact travel,
Traver alerts	such as major public events or severe weather warnings.
Multimodal	Information on alternative transportation modes, such as bike-
options	sharing or ride-hailing services.

2.1.2 Travel time

Travel time is a fundamental concept in the field of transportation, representing the duration required to move from one location to another. It serves as a critical metric for assessing the efficiency and performance of transportation systems, influencing both individual travel choices and broader transportation planning and policy decisions. Travel time is not only a key factor in the daily lives of commuters but also a crucial parameter for engineers, urban planners, and policymakers aiming to design, evaluate, and improve transportation networks.

In road networks, the travel time of vehicles is also known as vehicular travel time. This measure is crucial for understanding and analyzing the efficiency and performance of road networks and planning and managing traffic flows. According to the purposes of different applications, there are three types of vehicular travel times to be distinguished, i.e., path travel times, link travel times, and network travel times.

Table 2.3 gives detailed explanations of these travel times for clarification. In the context of negotiation systems and ATIS, path travel time is a better indicator for providing road users and authorities with the most updated and relevant traffic information on their chosen path. As introduced in Section 1.3, this thesis is intent on predicting path travel time using multiple sources of traffic data.

	Description	Factors influencing it	Typical application areas
Path Travel Time	The total time taken to travel from an origin to a destination along a specific route or path , including all links and intersections along the way.	Traffic congestion, traffic control devices, road type, weather conditions, incidents, and driver behavior.	Route planning, navigation systems, transportation modeling, and travel demand analysis.
Link Travel Time	The time required to travel from one end of a roadway segment or link to the other, not including the time spent at intersections or junctions.	Link length, speed limits, link capacity, traffic density, presence of traffic signals or stop signs.	Traffic simulation, network analysis, performance measurement, and congestion management.
Network Travel Time	The time to travel across a network from one point to another, considering all possible paths and the overall conditions of the network.	Network topology, overall traffic conditions, distribution of congestion, incidents, and traffic management strategies.	Network optimization, system-wide traffic studies, strategic planning, and emergency response planning.

Table 2.3 Differences between path travel time, link travel time, and network travel

time

2.1.3 Covariance of travel time

There are temporal and spatial covariance of travel times studied in the past decades (Chan et al., 2009; Tani et al., 2020; Fu et al., 2022). The former pertains to the variability of travel times across different periods for the same route or network. It captures how travel times fluctuate over time, reflecting the dynamic nature of traffic conditions. The latter one, on the other hand, deals with the correlation of travel times

across different routes or locations within a network. It provides insights into how travel times in one part of the network relate to those in another (Pan et al., 2013).

Considering the spatial covariance of path travel times is advantageous for capturing the interdependencies between different routes and understanding congestion propagation. It can enhance the predictive accuracy of travel time in complex traffic networks. However, it can increase model complexity and data sparsity issues (Chan et al., 2009; Ma et al., 2018a). In contrast, temporal covariance is beneficial for short-term forecasting and capturing dynamic traffic patterns over time, but it can overlook spatial factors (Li et al., 2012; Zhong et al., 2017). Generally, integrating both spatial and temporal covariances can provide a more holistic and precise approach to travel time prediction.

As introduced in Section 1.3, three types of temporal covariance of path travel times are mainly used for predicting path travel times. In this thesis, the study paths are mainly trunk roads or expressways. They have fewer alternative routes compared to other road types. Hence, the spatial covariance of path travel times between alternative paths is insignificant and not the primary focus of this study. Some papers considered spatial covariance of link travel times frequently (Li et al., 2012; Stathopoulos and Karlaftis, 2001; Tam and Lam, 2011) instead of path travel times. The prediction of link travel times considering their spatial covariance is recommended for further study.

2.2 Categories of Traffic Sensors for Data Collection

Traffic sensors play an important role in ATIS for obtaining different types of traffic data. Multi-source traffic data are collected from various traffic sensors. Each of these traffic sensors may provide at least one data source for path travel time information. According to the categories of traffic sensors set by Mori et al. (2015), point-to-point sensors and point sensors are used to gather various traffic data. They are thus integral components of ATIS. Point-to-point, such as AVI and GPS (similarly Bei Dou) sensors measure the travel times of vehicles passing through a specific road section. In contrast, point sensors detect traffic conditions (e.g., spot speed and flow) at specific locations installed with the sensors.

There are several categories of traffic sensors. The difference in measurement type leads to various practical applications of these traffic sensors, as shown in Table 2.4. It should be noted in the following table that the traffic sensors are used in the thesis. The point sensor refers to the video-based cameras used to collect point speed data in the remaining contents of the thesis.

Categorization	Traffic sensor	Measurement type	Practical application
Automatic vehicle identification sensor	Radio- frequency identification Automatic license plate recognition	Point-to-point information for identified vehicles (e.g., travel time and flow data).	Tolling; vehicle tracking; access control (e.g., parking) Vehicle classification; traffic count estimation
Point sensor	Video-based camera	Vehicle counts and point speed	Detecting vehicles across multiple lanes, classifying them based on length, and providing data on vehicle presence, flow rate, occupancy, and speed for each vehicle class.
Navigation satellite system sensor	Global positioning system	Semi trajectory of monitored vehicles (e.g., speed and location)	Navigation; fleet tracking; traffic management; public transportation

Table 2.4 Various traffic sensors and their practical applications

The details of the technology adopted for traffic sensors are as follows.

2.2.1 AVI sensors

There are various AVI sensors, such as radio-frequency identification (RFID) tag readers, automatic license plate recognition (ALPR) cameras, Bluetooth media access control (MAC) address readers, infrared sensors, barcode scanners, and dedicated short-range communications (DSRC) sensors. The subsequent paragraph will illustrate the corresponding technology item by item.

RFID technology utilizes radio waves to recognize and track objects of interest. It consists of a small chip or tag attached to the object of interest, and a reader emitting radio waves to communicate with the tag. When the tag comes within range of the reader, it transmits its unique identification information, allowing the reader to identify and track the object. The RFID tag readers to be mentioned in Chapter 3 are installed at the roadside, while the RFID tags are equipped at the front of vehicles. Therefore, vehicles with RFID tags can be tracked by RFID tag readers.

RFID has four benefits. First, RFID systems can automatically detect and track items without manual scanning, leading to increased efficiency and reduced labor costs. Second, RFID technology provides accurate and reliable data capture, minimizing errors associated with manual data entry. Third, RFID tags can be attached to various objects, including assets, inventory, and vehicles, making the technology suitable for various industries and applications. Fourth, RFID systems can be integrated with access control systems to enhance security and prevent unauthorized access to restricted areas or assets.

RFID technology has many applications, including automatic vehicle identification, inventory management, access control, and supply chain management. It offers several advantages, such as the ability to read multiple tags simultaneously, work in harsh environments, and track items without line-of-sight. Due to its efficiency and accuracy in tracking and recognizing objects, RFID technology has become increasingly popular in various industries.

ALPR is a technology that uses optical character recognition to read and recognize images of license plates on vehicles automatically. It typically consists of cameras, software, and databases that work together to detect and track vehicles based on their license plate information.

ALPR systems use specialized cameras to capture high-quality license plate images,

even in varying lighting and weather conditions. These cameras are often mounted on fixed structures such as poles or gantries. The captured images of license plates are processed using optical character recognition software. The software can recognize and extract the alphanumeric characters from the plate. This allows the system to convert the visual data into machine-readable text. The system can also be integrated with other databases to provide additional context, such as vehicle registration information. They are to be used in Chapters 4 and 5.

ALPR technology has a wide range of applications, including law enforcement, toll collection, parking management, and traffic monitoring. It offers several benefits, such as the ability to read license plate data, automate the process of spotting vehicles, and enhance security and surveillance capabilities quickly and accurately. In law enforcement, ALPR technology can discover stolen vehicles, locate vehicles associated with criminal activity, and enforce traffic laws. In toll collection and parking management, ALPR systems can automate the detection and billing of vehicles, improving efficiency and reducing the need for manual intervention.

Bluetooth technology is a wireless communication standard that allows electronic devices to link and exchange data over short distances. It operates on the 2.4 to 2.485 GHz frequency band and is commonly used for connecting devices such as smartphones, tablets, laptops, and peripherals like keyboards, mice, and headphones. Bluetooth technology has evolved over the years, with the latest version being Bluetooth 5.2, offering improved range, speed, and data capacity.

Bluetooth technology for AVI purposes involves using Bluetooth-enabled devices to recognize and track vehicles. Bluetooth is a wireless technology that enables short-range communication between devices, making it suitable for AVI applications in various scenarios. Bluetooth sensors can detect Bluetooth-enabled vehicles by identifying vehicles' MAC addresses. Bluetooth technology offers a balance between range and accuracy. Hence, it enables applications where vehicle identification is within a specific proximity. The range can be adjusted based on the specific requirements of the AVI system.

Bluetooth technology can be used for electronic toll collection, parking management, and access control systems. It allows for seamless vehicle identification and tracking without the need for physical interaction, enhancing the efficiency of AVI processes.

Infrared technology is a wireless communication method that uses infrared light to transmit data between devices. Infrared light is a type of electromagnetic radiation with wavelengths longer than visible light but shorter than radio waves. Infrared technology has been widely used in various applications, including remote controls, data transmission, and sensing. Infrared sensors can be used to detect the presence of vehicles as they approach a specific point, such as a toll booth or a parking gate. When a vehicle interrupts the infrared beam, the sensor registers the presence of the vehicle, triggering the AVI system to capture and process the vehicle's identification information.

Infrared technology transmits data wirelessly between devices, such as remote controls for TVs, DVD players, and other consumer electronics. It can also be used for shortrange communication between devices, such as smartphones, tablets, and laptops. Infrared communication typically requires a direct line of sight between the transmitting and receiving devices. This means that obstacles such as walls or objects can block data transmission. However, some infrared systems use reflection or bouncing of the infrared signal to overcome this limitation. Infrared communication can offer a degree of security, as the signal is less likely to be intercepted by unauthorized devices outside the line of sight. However, it is still important to implement encryption and other security measures for sensitive data transmission.

Infrared technology has been widely used in consumer electronics for remote control applications. It allows users to wirelessly control devices from a distance, making it a convenient and widely adopted technology for home entertainment systems and other appliances. While infrared technology offers advantages such as low power consumption and low cost, it has limitations related to its line-of-sight requirement and relatively short range compared to other wireless technologies like Bluetooth or Wi-Fi.

Barcode scanning is a technology that involves using optical scanners or cameras to read and decode information stored in barcodes. Barcodes are visual representations of data consisting of parallel lines or geometric patterns, and they are widely used for product identification, inventory management, and various other applications. Barcode scanning can also be used for vehicle access control, parking management, and toll collection. Barcodes on vehicle permits or tickets can be scanned at entry and exit points to detect and track vehicles as they pass through AVI checkpoints.

There are several barcodes, including linear barcodes (UPC and EAN codes) and twodimensional barcodes (QR codes and Data Matrix codes). Each type of barcode has its own structure and encoding method. Barcode scanning can be performed using dedicated handheld scanners, mobile devices with built-in cameras, or stationary scanners integrated into retail checkout counters, warehouses, and other locations. These devices capture the barcode image and use software to decode the information. When a barcode is scanned, the scanner captures the visual pattern of the barcode and converts it into a digital signal. The encoded data, such as product codes, serial numbers, or other information, is then extracted from the barcode and transmitted to a computer or database for processing.

Barcode scanning is widely used in retail for inventory management, point-of-sale transactions, and product tracking. It is also used in logistics and supply chain management to track shipments and improve order accuracy. Additionally, barcodes are used in healthcare, ticketing, asset tracking, and other industries. Barcode scanning offers several advantages, including speed and accuracy in data capture, ease of implementation, and cost-effectiveness. It provides a standardized method for identifying and tracking items, reducing errors, and improving operational efficiency.

However, barcodes require a direct line of sight between the barcode scanner and the barcode label. Moreover, barcodes can be affected by environmental factors such as dirt, damage, or weather conditions. Besides, traditional 1D barcodes have limited data capacity, which may not be sufficient for storing extensive vehicle identification information. While 2D barcodes can store more data, they still have limitations compared to other identification technologies, such as RFID or license plate

recognition. These limitations make them less suitable for high-speed, high-traffic, or challenging environmental conditions.

DSRC is a wireless communication technology designed for vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. It operates in the 5.9 GHz band and is specifically developed for automotive and transportation applications. DSRC enables vehicles and roadside infrastructure to exchange information, such as safety warnings, traffic management data, and other relevant information to improve road safety and traffic efficiency.

DSRC technology is subject to regulatory standards and requirements, with specific protocols and communication standards defined to ensure interoperability and compatibility across different vehicle and infrastructure implementations. Moreover, DSRC technology is expected to be deployed in autonomous vehicles, as it provides a means for vehicles to communicate with each other and infrastructure to support cooperative driving and decision-making.

DSRC offers high-speed, low-latency communication with robust security features, making it well-suited for specific AVI applications. However, its limitations in range, line-of-sight requirements, and infrastructure costs should be carefully considered when evaluating its suitability for specific AVI deployments.

Table 2.5 gives a description summary of these AVI sensors. In general, these technologies play a crucial role in modern traffic management systems, offering unique benefits for tracking, identifying, and analyzing vehicle movements. However, their effectiveness can vary based on the application, environmental conditions, and the existing infrastructure. Hence, the sample size and accuracy of AVI data are different for each technology. Appendix B gives photographs and weblinks for different AVI sensors for further information.

Technology	Key Features	Advantages	Limitations
	Use radio waves to	-High read accuracy	- Needs vehicles to
RFID tag	read data from RFID	-Fast processing	have RFID tags
readers	tags attached to	-Work in various	- Limited read
	vehicles.	weather conditions	range
ALPR cameras	Use optical character recognition on images to read vehicle registration plates.	-Can track vehicles without requiring pre-installed tags -Useful for law enforcement and toll collection	 Susceptible to errors due to poor lighting or obstructions Privacy concerns
Bluetooth MAC address readers	Identify and record the MAC addresses of Bluetooth devices in passing vehicles.	-Obtain travel times and traffic flow -Non-intrusive	 Limited range Privacy concerns regarding tracking
Infrared sensors	Detect vehicle presence and count vehicles by emitting and detecting infrared light.	-Work in various lighting conditions -Low maintenance	 Limited to presence detection and counting Can be affected by environmental conditions
Barcode scanners	Read barcodes (e.g., on tickets or passes) for access control and payment processing. Enable communication	-Quick processing of vehicles with barcoded passes - Accurate	 Demands vehicles to have a barcode Limited to controlled access points Requires DSRC
DSRC sensors	between vehicles and roadside infrastructure using short-range radio signals.	 Supports V2I communication Can enhance safety and traffic efficiency 	equipment in vehicles and on infrastructure - Limited range

Table 2.5 Description summary of AVI sensors

For identifiers that are installed at vehicles, the sample size of AVI data is constrained by the number of identifiers. For example, not all vehicles are installed with RFID tags, while all vehicles have license plates. Therefore, the sample size of AVI data collected by RFID technology can be less than ALPR technology. For identifiers that commonly appear in both vehicles and people (e.g., Bluetooth), the corresponding sensors can gather records from both vehicles and people. Hence, the sample size can be quite large, but only a small number belong to vehicles on the road. Therefore, the models for predicting path travel times fully consider this issue.

The choice of AVI technology in ATIS often depends on the balance between the need for accuracy, privacy considerations, cost, and the specific traffic management or monitoring system requirements. Furthermore, it is worthy of mentioning RFID and ALPR technology in the existing AVI system of Hong Kong. This thesis mainly uses AVI data from these two technologies. RFID technology is applied in the JTIS and the current tolling system of urban road networks in Hong Kong. ALPR cameras are deployed in the Speed Map Panels System (SMPS) for data collection.

Due to their properties, ALPR cameras can gather more data than RFID technology. It is noted that RFID technology is used in Chapters 3 and 5, while ALPR technology is adopted in Chapters 4 and 5. Both JTIS and SMPS provide the average predicted path travel time of all vehicles, and the predicted results have been validated to exhibit satisfactory performance (Tam and Lam, 2011b, 2013). Therefore, they are used for validation purposes in the thesis.

2.2.2 Point sensors

Point sensors are set in fixed points on roads and collect traffic variables for specific locations. Traditional point sensors are inductive loop sensors embedded in the road surface. They are designed to detect the presence of vehicles by measuring changes in inductance as a vehicle passes over the loop.

The inductive loop sensor consists of a loop of wire embedded in the pavement, typically in a rectangular or diamond shape. When an alternating current is passed

through the loop, it generates a magnetic field. When a vehicle with a metal component, such as the body of the car, passes over the loop, it disrupts the magnetic field, causing a change in inductance. This change is then detected by the sensor, indicating the presence of a vehicle.

Inductive loop sensors are known for their reliability and accuracy in detecting vehicles, making them a popular choice for traffic management systems. When properly installed and maintained, inductive loop sensors are durable and can withstand heavy traffic and harsh environmental conditions. Moreover, they are relatively cost-effective compared to other vehicle detection technologies, making them a practical choice for many traffic management systems.

There are single-loop sensors and double-loop sensors (Mori et al., 2015). Single-loop inductive sensors consist of a single loop of wire embedded in the road surface. They are primarily used for essential vehicle detection, such as detecting the presence of a vehicle at a traffic signal or intersection. Single-loop sensors are typically used for simple vehicle presence detection and are not as effective in providing detailed information about the size or speed of the vehicle.

Double-loop inductive sensors consist of two separate loops of wire embedded in the road surface, typically arranged in a rectangular or diamond shape. They are used for more advanced vehicle detection and can provide additional information such as vehicle length, speed, and direction of travel. Double loop sensors are often used for more complex traffic management systems, such as traffic signal control, vehicle classification, and traffic flow monitoring.

Besides, acoustic sensors use sound waves to detect and monitor various physical phenomena. In the context of transportation and traffic management, acoustic sensors are used to discover and analyze vehicle presence, movement, and traffic patterns. These sensors operate by emitting sound waves and then analyzing the reflected waves to determine the presence and movement of vehicles. When a vehicle passes through the area monitored by the acoustic sensor, it creates disturbances in the sound waves, which are then detected and inspected to furnish information about the vehicle's speed,

direction, and size.

Acoustic sensors are commonly used in traffic management systems for applications such as traffic monitoring, vehicle counting, and speed measurement. They are often deployed in urban areas, highways, and tunnels to gather real-time traffic flow data and aid in traffic management and control.

One of the advantages of acoustic sensors is their ability to impart accurate and reliable data in various weather conditions, including rain, fog, and snow, making them suitable for use in diverse environments. Additionally, acoustic sensors are non-intrusive and do not require physical contact with vehicles, which can be advantageous in specific traffic monitoring applications.

Remote traffic microwave sensors (RTMS) are advanced traffic monitoring devices that utilize microwave technology to detect vehicles on roadways. These sensors are designed to provide real-time data on traffic volume and speed, making them valuable tools for traffic management and control. RTMS operate by emitting microwave pulses and analyzing the reflections to spot and monitor vehicles. When a vehicle passes through the monitored area, it causes a disturbance in the microwave signals, which can be analyzed to determine the vehicle's speed, length, and presence. RTMS can also supply traffic density, occupancy, and flow patterns.

RTMS are typically installed along highways, urban streets, and intersections, where they can gather comprehensive traffic data without physical contact with vehicles. RTMS devices are often used by transportation agencies and city planners to optimize traffic flow, improve safety, and make informed decisions regarding road infrastructure and traffic management.

One of the key advantages of RTMS is its ability to furnish accurate and reliable traffic data in various weather conditions and lighting environments. Additionally, RTMS sensors are non-intrusive and do not need physical installation in the road surface, making them easier to deploy and maintain compared to traditional in-road sensors.

In recent years, video-based cameras have been used as point sensors. These cameras are equipped with sophisticated technology to read and evaluate real-time traffic data, providing valuable insights for traffic engineering, urban planning, and infrastructure management. The camera system typically consists of high-resolution cameras, image-processing software, and communication interfaces. These cameras are strategically installed at intersections, highways, and urban roadways to capture video footage of traffic movements, vehicle counts, and other relevant data.

The cameras use advanced image processing algorithms to detect and track vehicles, measure traffic flow, and analyze traffic patterns. Information on vehicle speed and occupancy, traffic violations, and incidents are stored in the database. The data collected by cameras is used to optimize traffic signal timing, improve traffic flow, and enhance roadway safety.

One key advantage of these cameras is their ability to provide accurate and reliable traffic data. They are robust in various environmental conditions, including day and night, adverse weather, and varying light conditions. The cameras can capture high-quality images and videos, allowing for detailed traffic behavior and pattern analysis. In this thesis, the point sensors deployed in JTIS and SMPS are **video-based cameras**.

Table 2.6 summarizes the different technologies applied to point sensors. In practice, the choice of technology often depends on the specific requirements of the traffic monitoring project, including the level of detail needed, environmental conditions, installation and maintenance costs, and potential scalability. Appendix B provides photographs and weblinks for various point sensors for illustration. This thesis uses point sensor data from video-based cameras to predict path travel times.

Though point sensor data have no vehicle class information, they can gather the average speed of all vehicles that pass through the locations of point sensors in a specific time interval. As shown in Table 2.6, the video-based cameras enable rich data collection, which supports multi-class path travel time prediction, as illustrated in Section 4.3.1.

Technology	Characteristics	Advantages	Limitations
Single-loop sensors	Consist of a single loop of wire embedded in the roadway surface, connected to a sensor.	- Simple installation - Cost-effective - Reliable for vehicle detection	 Limited data (primarily presence and count) Can be damaged by roadway maintenance
Double- loop sensors	Comprise two closely spaced loops of wire embedded in the roadway	 Differentiate between vehicle classes More accurate speed and length measurement 	 More expensive to install than single- loop Susceptible to similar physical damage
Acoustic sensors	Use microphones to detect sound waves generated by vehicles.	 Non-intrusive installation Perform well in a variety of weather conditions 	 Noise pollution can influence accuracy May need calibration for different environments
Remote traffic microwave sensors (RTMS)	They can be mounted on poles or overhead structures.	 Provide speed, volume, and occupancy Non-intrusive and unaffected by lighting or weather conditions 	 More expensive than inductive loop sensors Requires line-of- sight to the target area
Video- based cameras (used in the thesis)	Capture and analyze traffic video for vehicle detection using image processing.	 Versatile and can collect counts, speed, and behavior analysis Flexible installation options 	 Need significant processing power Performance can be changed by lighting, weather, and occlusion

 Table 2.6 Summary of different technologies applied to point sensors with their advantages and limitations

2.2.3 Navigation satellite system sensors

GPS and BeiDou are popular satellite-based positioning systems. They provide accurate location and timing information for various applications, including navigation, mapping, and geospatial data collection. GPS is a satellite navigation system developed by the United States government. GPS sensors receive signals from a network of satellites orbiting the Earth. They use the timing and positioning information from these satellites to determine the receiver's location, velocity, and time. GPS sensors are widely used in automotive navigation systems, smartphones, aviation, marine navigation, surveying, and other location-based applications (Jang et al., 2023). They capture accurate positioning information, typically with an accuracy of a few meters, and are essential for modern navigation and location-based services.

BeiDou navigation satellite system is a global satellite navigation system established by China. Similar to GPS, BeiDou sensors receive signals from a constellation of satellites to determine the receiver's position, velocity, and timing. The BeiDou system provides global coverage and is designed to offer high-precision positioning services. It is employed for various applications, including transportation, surveying, mapping, and timing synchronization. BeiDou sensors are widely used in China and other countries, which offers an alternative or complementary positioning solution to GPS.

GPS and BeiDou sensors are crucial in providing accurate and reliable positioning and timing information. They contribute to a wide range of applications, including navigation, transportation, surveying, and location-based services. These satellitebased positioning systems have become integral components of modern technology. They enable precise and efficient location determination across various industries and sectors.

2.3 Types of Traffic Data for Path Travel Time Prediction

Several types of traffic data are gathered from different categories of traffic sensors, as elaborated in Section 2.2. These traffic data can be used for path travel time prediction, which will be explained in this section.

2.3.1 Travel time data

The timestamps of vehicles entering and leaving the study path can be collected from AVI sensors. The difference between these timestamps is the observed path travel time, as mentioned in Section 2.2.1. It is mostly used for path travel time prediction. Over the past two decades, AVI data has been increasingly explored for use in ATIS. Through AVI sensors, a vehicle passing an AVI sensor with its specific identifiers (e.g., RFID tags for RFID tag readers and license plate numbers for ALPR cameras) and the corresponding timestamp can be recorded. These data from successive AVI sensors are matched to the vehicle and used to calculate its travel time (Zhou and Mahmassani, 2006; Ahmed and Abdel-Aty, 2012; Chow et al., 2014; Soriguera and Martinez-DIaz, 2021), which is denoted as AVI data.

AVI data contains the timestamps of vehicles entering and leaving a path equipped with AVI sensors at both ends. The observed path travel time is calculated as the difference between these two timestamps. Appendix C shows a sample data format of AVI data. Figure 2.1 in Section 2.3.4 will further give a graphical illustration of timestamps with symbols and their relationships with observed path travel times. The locations of AVI sensors are also called checkpoints, while travel time between adjacent checkpoints can be obtained from AVI data (Qi et al., 2024). The vehicle class information is also contained in AVI data. When travel time distribution varies by vehicle class, it is interesting to predict path travel times by vehicle class.

2.3.2 Speed and flow data

The speed and flow data collected from point sensors (also referred to as point sensor data in the thesis), as explained in Section 2.2.2, can be used to acquire the observed path travel times indirectly. Point sensor data are most widely used in ATIS (Mori et al., 2015) to provide traffic measurements regarding flow, occupancy, and speed. The spot speed as well as vehicular flow data have a wide application for ATIS, including traffic speed prediction, travel time prediction, modeling of travel time distribution, travel time prediction, etc. (Han et al., 2010; Zhang, 2006; Chalumuri and Yasuo, 2014; Liu et al., 2005; Ye et al., 2011).

From the perspective of path travel time prediction, owing to the many vehicles that

pass through at locations with point sensors, a database may store speed data and flow at an aggregate level (i.e., the average speed and flow in a fixed time interval). Point sensor data cannot track individual vehicles unless other data sources are used simultaneously (e.g., Hyun et al. (2017), which tracked the trunks with point sensor data and weigh-in-motion data). Vehicle class information is not always available for point sensor data in practice, although speed data are indirectly converted to the path travel time of vehicles for prediction (Soriguera and Martinez-Diaz, 2021; Yildirimoglu and Geroliminis, 2013). Appendix C gives a sample data format of point sensor data used in this thesis.

2.3.3 Navigation satellite system data

GPS data are the backbone of several location-based ATIS applications, such as route guidance and map services (Sharath et al., 2019). The trajectories of each probe vehicle that can be obtained from GPS sensors, which comprise data of the vehicle speed, acceleration, and location, are gathered every few seconds to every few minutes. A vehicle's trajectory can be tracked from its entry point to its exit point on a path or corridor, which determines the vehicle's path travel time.

Owing to the uncertain sampling rate and the sparsely distributed sampling points, both the spot speeds reported by vehicles and the converted travel times of short segments require further investigation (Hofleitner et al., 2012; Ma et al., 2022; Moreira-Matias et al., 2016; Tang et al., 2018; Yin et al., 2015; Zhong et al., 2017). Vehicle class information is also available for GPS data as such data are generally collected among a particular class of vehicles. The sample data format of GPS data used in Chapter 4 is provided in Appendix C.

Similar to GPS, BeiDou Navigation Satellite System is also one of the four largest Global Navigation Satellite Systems in the world. The license plate number, trajectory, vehicle class, and speeds can be acquired for vehicles installed with BeiDou Navigation Satellite System. Numerous types of research have been conducted using BeiDou data in the transportation field, including traffic prediction (Wei et al., 2019; Zhao et al., 2019), driving behavior analysis (Sun et al., 2016; Yang et al., 2021a), lane changing studies (Ma et al., 2021, 2022).

2.3.4 Summary of traffic sensors and traffic data

The summary table of traffic sensors is given in Table 2.7. Both AVI and navigation satellite system data sensors directly gauge the travel times of vehicles between two specified locations. In contrast, point sensors indirectly provide the travel times of vehicles through speed measurements. The AVI data offers the experienced path travel times of vehicles with the vehicle class information at the stationary locations installed with AVI readers. However, as depicted in Section 2.3.1, the sample size of the AVI data per time interval used in the thesis (i.e., 2-minute) can be small for the prediction of path travel times in the current and future time intervals.

Traffic data	Data	Advantages	Disadvantages
		Experienced path travel	
	Timestamps	time; vehicle class	Low sampling rate, e.g.,
	of vehicles,	information of all	small sample size per
AVI data	vehicle	vehicles (tolling purposes	time interval (Dion and
	class	requiring vehicle class	Rakha, 2006)
		information)	
Daint gangan	Spotsmood	Average aread of	No vehicle class
	spot speed,	Average speed of	information (Soriguera
data	Ilow	venicles	and Robusté, 2011a)
Novigation	Timestamps	Vahiala alass information	Trajectories need more
Navigation	of vehicles,	eveilable for tracked	data processing
satellite system data	vehicle		procedures (Wang et
	class, speed	venicles	al., 2021a)

Table 2.7 Summary of different types of traffic data

Navigation satellite system datasets consist of mobile data points compared to stationary AVI data. However, the trajectory of vehicles tracked with GPS data is not exactly the same as the selected path installed with AVI sensors at both ends. Point sensor data have accurate speed measurements for all vehicles passing through locations equipped with point sensors, but the spatial features of vehicle trajectories are unavailable from point sensors. Furthermore, vehicle class information is also unobtainable from point sensor data. These three data sources should be fully utilized when performing multi-class path travel time prediction, which is to be illustrated in Section 4.3.

Figure 2.1 presents a schematic example of various traffic sensing systems commonly used in ATIS. A GPS-equipped truck with an AVI identifier travels along path p on day d. A pair of AVI sensors is installed at origin x^{o_p} and destination x^{d_p} of path p. When this truck travels along this path, both AVI sensors and GPS sensors will track this truck with the identification number (say 10001 for AVI sensor and 18000 for GPS sensor).



Figure 2.1 A schematic example of various traffic sensing systems commonly used in ATIS

For the AVI sensor, the observed path travel time of vehicle 10001 (i.e., 19.2 min) can be obtained by the difference between these two timestamps. The timestamps of vehicle 10001 with vehicle class 2 (which represents truck) traversing path p on day d at location x^{o_p} and x^{d_p} are collected from AVI sensors (i.e., $\tau_{10001,d,x^{o_p},p}^{S_{A},2}$ and $\tau_{10001,d,x^{d_p},p}^{S_{A},2}$) and shown in Figure 2.1. It should be noted that AVI sensors in this example are ALPR cameras. They will be used in numerical experiments in Chapter 4 and empirical tests in Chapter 5 correspondingly. For GPS sensors, three observed data points are allocated on path p, with the location, timestamp, vehicle ID by GPS sensors (i.e., 18000), and the instantaneous speed of vehicle 18000. The trajectory of vehicle 18000 captured by the GPS sensor can indirectly infer the experienced travel time of vehicle 18000 on path p.

For point sensors (video-based cameras), in the studied time interval t (i.e., 16:18:00-16:20:00, indexed by 1), there are 29 and 33 vehicles pass through two point sensors, respectively. Their average speed is 67 km/h. Obviously, the studied truck is not measured by these two point sensors at time interval t = 1. However, the predicted path travel time from point sensors at time interval t = 1 can still provide a reference to verify the validity of observation 10001 in the database of AVI sensors and observation 18000 in the dataset from GPS sensors.

In summary, the AVI data are sparse over time intervals, while they are regularly collected for the point sensor data. The GPS data are randomly distributed in time and space. In this thesis, all information will be integrated to predict path travel times.

2.4 Review of Filtering Algorithms for Different Types of Traffic Data

For different types of traffic data, invalid data or outliers are to be filtered out. It is essential as the amount of these invalid data or outliers can be significantly large. The predicted path travel times would be biased without removing these data. Fixed-value filters and float filters are designed for AVI data collected from different road types. The fixed-value filter is generally applied for point sensor data. The filtering process is relatively simple as no data identification problem is encountered. GPS data needs more spatial considerations as the locations of GPS sensors are dynamic compared with AVI and point sensors. Map-matching algorithms are applied with spatiotemporal constraints on GPS data.

Data-filtering and outlier-detection algorithms have been developed for other traffic variables, including flow (Li et al., 2015) and speed (Chakraborty et al., 2019). These

algorithms assume that most data are valid, and hence, they remove only small portions of invalid data (Chen et al., 2010). However, a large proportion of invalid data can exist in practice. For example, for AVI data with low sampling rates, its distribution can be more scattered and varied. Thus, a large proportion of AVI data may be invalid. Furthermore, the occurrence of longer travel times by path is more frequent when traffic is congested. It is challenging to distinguish invalid data from comparatively long travel times by path under this scenario (Shang et al., 2022).

2.4.1 Filtering algorithms for AVI data

The filtering algorithms for various types of AVI data can be different. As Bluetooth sensors can also collect data from passengers in vehicles, Bluetooth data is less accurate with large sample sizes. On the contrary, RFID and ALPR data are more accurate with comparatively small sample sizes. Bluetooth and Wi-Fi data share similar filtering algorithms, as both have large sample sizes while the accuracy is relatively low (Pu et al., 2021; Ghavidel et al., 2022).

For the RFID and ALPR data, generally, both fixed value and float filters can be applied for AVI data (Chen et al., 2022). The maximum and minimum speed limits can be used to derive the upper and lower bounds of path travel times. The mean, median, standard deviation, and percentiles of measured AVI data can also be used to construct the time window for filtering. The time window for selecting valid AVI data is called the validity window (Tam and Lam, 2008). These models have been evaluated respectively (Asqool et al., 2021). The related models will be introduced in detail in Section 3.1.1.

As existing filtering algorithms only use real-time AVI data, the resultant validity windows lack rigorous mathematical guarantees, particularly for limited real-time AVI data with low sampling rates. Therefore, existing filtering algorithms for AVI data may not be effective. It should be noted that there are several reasons for the low sampling rates of real-time AVI data. First, real-time AVI data is collected at high frequency (e.g., gathered every two minutes). Second, only AVI data from commercial vehicles can be collected due to privacy issues in some cities. Third, the traffic demand during some periods (e.g., midnight) is low. The sample size of real-time AVI data may not be

sufficient for effectively selecting the outliers.

2.4.2 Filtering algorithms for point sensor data

The speed measurements are most widely used from point sensor data for path travel time prediction. The point speed data for specific locations can have fewer outliers/invalid data unless the sensor fails. A fixed threshold concerning the free-flow speed of the location of point sensors is usually applied to screen out extreme speed measurements. Moreover, some pieces of literature assume that there are no systemic errors in GPS data, as they are gathered from thousands of devices. On the contrary, the breakdown of point sensors can affect the accuracy of point sensor data for a long time. Therefore, GPS data can be regarded as ground truth and used to filter out obviously problematic point sensor data (Li et al., 2016).

2.4.3 Filtering algorithms for navigation satellite system data

As introduced in Section 2.3.3, various types of information are collected from navigation satellite system sensors. Therefore, each type of information can be used to filter out invalid records (e.g., location and speed). For individual speed measurements from navigation satellite system sensor data, the acceleration and deceleration speeds have been considered for screening out those valid speed measurements (Rim et al., 2016). Speed changes exceeding the given maximum acceleration and deceleration speed speed can be regarded as outliers.

Navigation satellite system data without valid location state, vehicle status, and vehicle identification number are filtered out (Chen et al., 2021). The map-matching algorithms identify the invalid vehicle locations that are out of the scope of the studied path travel times (Chen et al., 2021). The related path travel time measurements on these records are also removed. Moreover, as both AVI and GPS data enable the collection of direct travel times, the filtering algorithms for AVI data are sometimes applied to GPS data (Yuan et al., 2023).

Similar to GPS/BeiDou data, mobile phones can also provide speed measurements of vehicles by detecting signals of passengers/drivers in the vehicles. There are ping-pong data resulting from overlapping base station signals (Wu et al., 2023). The density peak

clustering algorithm has been applied to filter out invalid mobile phone data (Wu et al., 2023).

2.5 Review of Path Travel Time Prediction Models

After data filtering as described in Section 2.4, these traffic data can be used to predict the travel times of the desired path. Studies on the prediction of path travel times have focused on predicting the mean travel time of the whole path instead of the mean travel time of each individual road segment or link along the path (Sun et al., 2022a). Research on path travel time prediction can be categorized into those conducted at the strategic level and those performed at the operational level.

The path travel time prediction at the strategic level obtains the travel time in an offline manner, focusing on the sensor location problem (Salari et al., 2019; Shao et al., 2021; Sun et al., 2022b) or sensor replacement problem (Manco et al., 2017; Zhu et al., 2017), where the path travel time is impacted by the constraints or objectives of the models used to determine the locations of different types of sensors. In contrast, path travel time prediction at the strategic level analyzes historical traffic data to predict the traffic states and travel time distributions on the road networks over the long term (Fu et al., 2019; Laña et al., 2019; Nantes et al., 2015; Qin et al., 2020; Yun et al., 2019a).

For freeways with no exits or entrances, traffic flow modeling the traffic conditions and derives the corresponding path travel times (Ngoduy et al., 2006; Celikoglu, 2013a; Nantes et al., 2015). For urban roads with exits and entrances along the path, the path flow is difficult to capture with a limited number of traffic sensors. Therefore, point sensors, which measure accurate flow, contribute less than point-to-point sensors like AVI and GPS sensors to the prediction of the path travel time.

2.5.1 Review of path travel time prediction models using AVI data

As AVI data provide direct path travel time information for individual vehicles, they can be preprocessed to eliminate outliers and thus used to forecast travel times on monitored paths or road segments. Several algorithms have been developed and applied to predict travel times using AVI data, e.g., Dion and Rakha (2006), Mouskos

et al. (1998), Southwest Research Institute (1998), Tam and Lam (2008), and TranStar (2021) used RFID-based AVI data, Haghani et al. (2010) used Bluetooth-based AVI data, Park and Kim (2018) used dedicated short-range communication data, and Ma and Koutsopoulos (2010) used AVI data from ALPR cameras. It should be noted that the filtering algorithms mentioned in Section 2.4.1 usually present the predicted path travel time as they need to compare with ground truth for performance evaluation.

The determination of a validity window for screening outliers has been continuously studied in the past two decades. The TransGuide algorithm (Southwest Research Institute, 1998) initially set the fixed threshold so that any observation with a percentage deviation larger than the fixed threshold would be filtered out. This is more appropriate for rural roads because there are fewer congestion cases on rural roads than on arterial roads. For arterial roads that have complicated traffic conditions, it is limited to capturing the characteristics of traffic conditions.

Dion and Rakha (2006) further introduced the machisum by looking back at three consecutive outliers in the previous time intervals. If they increase or decrease with clear trends, the validity windows will be enlarged in case the traffic conditions change. This approach has been followed by Tam and Lam (2008) and Ma and Koutsopoulos (2010). Park and Kim (2018) further analyzed the distributions of path travel times collected from AVI data to improve the filtering performance further. However, these filtering algorithms for AVI data are less effective when real-time AVI data are limited. This will be further elaborated in Section 3.1.1.

2.5.2 Review of path travel time prediction models using navigation satellite system sensor data

Some studies have investigated the use of GPS data for travel time prediction. As GPS data offer more vehicle location data than AVI data, predicted travel time can be more easily obtained from GPS data (after map-matching) than from AVI data (Gong et al., 2015; Zhong et al., 2020). The trajectory-based model can be applied to GPS data to track vehicle trajectories with deep learning models (Zhu et al., 2022b). The traffic-flow theory-based model can also be applied to GPS data once the road information

(e.g., signal timing) is known (Hiribarren and Herrera, 2014).

Apart from filtering algorithms used for AVI data, map-matching algorithms designed for GPS data have been studied for decades. Simple geometric techniques are applied to complex models (Quddus et al., 2007). Others include geometrical algorithms, topological map-matching algorithms, Kalman filter-based algorithms, hidden Markov models, and Frechet distance-based algorithms. These algorithms have been compared and evaluated (Singh et al., 2023). It was found that topological algorithms outperformed the other algorithms.

In this thesis, GPS data is mainly used to predict path travel times by vehicle class. The availability of vehicle class information, as depicted in Section 2.3.4, ensures that the temporal covariance of path travel times between vehicle classes by 2-minute intervals can be obtained from GPS data to improve prediction performance.

2.5.3 Review of path travel time prediction models using point sensor data

On the one hand, Lighthill–Whitham–Richards theory can be applied when paths are controlled. Flow conservation equations and traffic dynamics are investigated using flow and speed data collected from point sensors (Celikoglu, 2013b). On the other hand, the speed interpolation models can be applied along the study path so that speed fields along the path can be generated. The path travel time prediction models using point sensor data from point sensors are illustrated in the following paragraphs.

It has been pointed out that the kind of point speed data from point sensor data may not be suitable for quantifying the average speed of a road section (Soriguera and Robusté, 2011b). In view of this, various travel time prediction models based on spot speed data from point sensors have been proposed previously in a heuristic and practical manner. Since the spatial speed propagation of vehicles cannot be captured by point sensors, a constant or linear relationship of speeds along the road section between the two detectors is usually assumed.

The most common approach is the instantaneous travel time prediction model, which simultaneously calculates the link travel times at these two adjacent point sensors. In

the past years, the time slice model, dynamic slice model, and linear model have been developed on the basis of different time intervals that are used for travel time prediction (Bajwa et al., 2003; Cortés et al., 2002; Lint and Zjipp, 2003). These existing approaches directly use the spot speed or point speed data to predict link travel times with different assumptions.

Li et al. (2006) compared four categories of speed-based travel time prediction models, including the instantaneous model, time slice model, dynamic time slice model, and linear model. They concluded that the last two models provided better-predicted path travel times for cases when detectors are spaced at larger distances. They also found that the results of these four speed-based models would tend to underestimate the travel times on the measured road sections.

For data-driven approaches based on spot speed data, some scholars discovered the relationship between speed measurements at both point sensors and travel times between AVI sensors through supervised learning techniques. Distance, speed measurement, flow, occupancy, and other parameters are used in data-driven approaches to investigate travel time prediction problems.

Cherrett et al. (2002) applied a neural network using average loop-occupancy time per vehicle, average time gap between vehicles, and percentage occupancy as input for travel time prediction. Tang et al. (2016) applied an evolving fuzzy neural network to predict the link travel times between two adjacent detectors. The input variables for their neural network training are traffic volume, occupancy, and speed measured by the loop detectors. Recently, Lu et al. (2017) applied the clustering algorithm to find similar traffic patterns, followed by support vector regression to predict freeway travel times.

In Chapter 4, the point sensor data is used to enrich the information on traffic conditions for path travel time prediction by different vehicle classes. Though point sensor data has no vehicle class information, it provides the average speed of all vehicles. Hence, the means of path travel times of all vehicles can be obtained from point sensor data. As the model choice of using point sensor data for path travel time

prediction is not the major contribution of this thesis, it is briefly introduced in Section 4.3.1.

2.5.4 Review of path travel time prediction models using multi-source traffic data

The traditional fusion model is a weighted average (Zhu et al., 2018), which gives different weightings/scores to each data source. It is straightforward, while the weighting scores are not flexible once they are determined by historical traffic data. A widely used fusion model is an extended generalized Treiber-Helbing filter, which can fuse different sources of traffic data for path travel time prediction. Jiang et al. (2017) fused GPS and point sensor data to forecast path travel times. Other data fusion models have been adopted for multi-source traffic data. Kalman filtering was used to combine GPS and point sensor data (Anusha et al., 2012). Zhu et al. (2018) connected GPS, mobile phone data, and point sensor data with ANN. Shi et al. (2017) fused point sensor data using Dempster-Shafer's theory to foretell path travel times for urban roads.

To categorize these fusion models, three levels of fusion can be concluded: data, feature, and decision. Based on these three levels, the Kalman filter, neural network, Dempster-Shafer, Fuzzy Logic, joint probabilistic data association, software agent, Bayesian, and hybrid algorithms have been developed to fuse these traffic data for path travel time prediction (Kashinath et al., 2021).

As introduced in Section 2.3.4, this thesis aims to predict path travel times by vehicle class using multi-source traffic data. It is still a challenge to integrate these traffic data with different characteristics for path travel time prediction.

2.5.5 Classification of path travel time prediction models

There are statistical models and machine learning models for path travel time prediction. Table 2.8 lists their descriptions, advantages, and limitations. Table 2.8 shows that these models can capture typical traffic patterns with relatively low computational costs. However, sometimes, the traffic conditions can suddenly change due to adverse weather, or the data amount is insufficient for traditional models to give

a clear picture of variations in path travel times.

In recent years, more deep learning models have been applied in the area of path travel time prediction, which is summarized in Table 2.9. They further specified components within the neural works (e.g., understand traffic data as an image; extract features from time-series data) to have outperformance dealing with traffic data. However, the computational costs of these models are usually more expensive. Hence, there is a trade-off between accuracy and computation cost for model selection, depending on users' preference for evaluating prediction models. Some commonly used deep learning models will be used as benchmarks for comparison in the later chapters, such as long short-term memory (LSTM) networks. The proposed models in the thesis ensure the accuracy and efficiency of predicted path travel times.

Model Type	Description	Key Inputs	Strengths	Limitations
Historical average models	Use historical travel time data to predict future travel times, often assuming that traffic patterns are repetitive and consistent.	Historical travel time data, time of day, day of week.	Simple to implement, require minimal data processing.	May not account for non-recurrent events or anomalies.
Time-series analysis models Regression models	Apply statistical methods to analyze and forecast travel times based on observed time-series data. Use statistical techniques to establish relationships between travel time and various influencing factors.	Historical travel time data, temporal patterns, seasonality. Traffic volume, weather conditions, road incidents, time of day.	Capture trends and seasonal variations. Incorporate multiple variables affecting travel times.	May not respond well to sudden changes in traffic conditions. Needs accurate and comprehensive data for model calibration.
Machine learning models	Employ algorithms that learn from data to make predictions, including neural networks, support vector machines, and decision trees.	Historical travel time data, traffic flow, weather conditions, and special events.	Handle complex nonlinear relationships and interactions.	May call for large datasets and significant computational resources.

Table 2.8 Summary of traditional models for path travel time prediction

	Utilize recursive algorithms to	Real-time traffic data,	A dant to real time	Complexity in
IZ 1 C [*] 1, 1 1	forecast travel times by continuously	historical travel time	Adapt to rear-time	implementation and
Kaiman intering models	updating predictions with real-time	data, and traffic flow		tuning of the model
	data.	models.	conditions.	parameters.
Simulation-based models	Create detailed representations of traffic flow to simulate and predict travel times under various scenarios.	Road network characteristics, traffic demand, control strategies.	Model specific scenarios and interventions.	Computationally intensive and demand detailed input data.
Hybrid models	Combine features of different models to leverage their strengths and mitigate their weaknesses.	Varies based on the combination of models used (e.g., machine learning with time-series analysis).	Improve accuracy and reliability of predictions.	May be complex to develop and require diverse datasets.
Deep learning models	Description	Key inputs	Strengths	Limitations
---	--	--	--	--
Recurrent neural network (RNN)	Designed to handle sequential data, RNNs are particularly suited for time- series prediction tasks.	Historical travel time data, traffic flow, and temporal patterns.	Good at capturing temporal dependencies in time-series data.	Can struggle with long- term dependencies due to vanishing gradient problems.
Long short-term memory (LSTM) Networks	A type of RNN that can learn long- term dependencies using memory cells and gates.	Historical travel time data, traffic flow, temporal patterns, and weather conditions.	Better at capturing long- term temporal relationships than standard RNN.	More complex and computationally intensive than RNN.
Gated recurrent units	Similar to LSTMs but with a simplified structure, gated recurrent units are another type of RNN that can capture temporal dependencies.	Historical travel time data, traffic flow, temporal patterns, and weather conditions.	Require fewer parameters than LSTMs, making them faster to train.	May be less expressive than LSTMs for specific tasks.
Convolutional neural network (CNN)	Typically used for image processing, CNN can also be applied to traffic data represented in grid-like structures.	Spatial representation of traffic data.	Good at capturing spatial dependencies in grid-like data representations.	Not inherently designed for sequential time-series data.

Table 2.9 Summary of traditional models for path travel time prediction

Graph convolutional networks	Extend CNNs to work with graph- structured data, making them suitable for road network representations.	Road network graph, traffic flow, historical travel time data.	Attain spatial dependencies in non- Euclidean data like road networks.	May require significant domain knowledge to structure data appropriately.	
Deep belief network (DBN)	A generative model composed of multiple layers of stochastic, latent variables. Used for unsupervised learning of efficient data encodings,	Historical travel time data, traffic flow, and other relevant features. Historical travel time data_traffic flow_and	Learn to represent complex distributions of data. Learn compressed representations of data,	Training can be challenging and time- consuming. Mainly used for dimensionality reduction	
	autoencoders can help in feature extraction for travel time prediction.	other relevant features.	useful for feature extraction. Leverage the strengths	rather than direct prediction.	
Ensemble models	Combine predictions from multiple deep learning models to improve overall performance.	Outputs from various deep learning models.	of individual models to enhance prediction accuracy.	Increased complexity and computational cost due to multiple models.	

2.6 Summary

As discussed in the above sections of Chapter 2, ATIS urgently need accurate and reliable predicted path travel times in the current and future time intervals. Relevant existing works have made numerous contributions to this problem, applying various models and multiple sources of traffic data. However, some challenges remain to be tackled and will be studied in this thesis.

First, AVI data are commonly used for path travel time prediction, as introduced in Section 2.3.1. There are existing filtering algorithms for extracting outliers from realtime AVI data. However, their performance worsens when the real-time AVI data sampling rate is low, as explained in Section 1.2. On the one hand, historical traffic data can supplement the traffic information and help to improve the filtering performance. On the other hand, it can be difficult to collect sufficient ground truth in practice for model training. Thus, an unsupervised filtering algorithm for real-time AVI data without using ground truth for training is to be proposed in Chapter 3.

Second, existing ATIS typically disseminate the average predicted path travel times for road users in Table 2.1. However, the observed path travel times of partial vehicles (e.g., private cars) can deviate significantly from the average path travel times, which is to be illustrated in Figure 4.3 in Section 4.1. There is a research gap to predict path travel times by vehicle class. Besides, there are various types of traffic data from different traffic sensors, as introduced in Sections 2.1.3 and 2.3. However, their data formats and the traffic information contained in each data source are heterogeneous, as shown in Appendix C. Hence, it is necessary to integrate them effectively and efficiently to forecast path travel times. Therefore, a novel prediction model for multiclass path travel time prediction using multi-source traffic data is raised in Chapter 4.

Third, weather information is noticeable to be provided in some ATIS, as mentioned in Table 2.1 and Table 2.2. For cities with frequent rainfall, it is important to incorporate rainfall-related weather information to enhance the performance of predicted path travel times. This information consists of rainfall intensity data collected from the past and weather forecasts referring to future weather conditions. Therefore, existing prediction models can be further extended to a more comprehensive modeling framework, which ultimately utilizes rainfall-related weather information. To solve this problem, Chapter 5 proposes a modeling framework for forecasting path travel times in the future time intervals. It incorporates different sources of weather information for path travel time prediction.

3. Filtering Limited Automatic Vehicle Identification Data without Ground Truth on Path Travel Times

3.1 Initial Considerations

Section 2.4 has presented the importance of filtering out outliers/invalid data before path travel time prediction. The relevant data filtering algorithms for traffic data have also been illustrated in Section 2.4. However, in practice, the real-time AVI data can be limited for several reasons (e.g., technology and frequency used for data collection, privacy issues, periods of collection). Different from Bluetooth data, as mentioned in Section 2.4.1, RFID and ALPR data could provide high-quality but limited AVI data per time interval. Given the challenges of path travel time prediction with these limited high-quality AVI data in practice, this chapter investigates various data filtering algorithms. It develops a novel filtering algorithm for limited AVI data for path travel time prediction in the current time interval.

The rest of this chapter is organized as follows. The background, motivation, and contributions of this chapter are presented in Section 3.1. The problem under investigation in this chapter is displayed in Section 3.2. A novel filtering algorithm is proposed in Section 3.3. The case studies are conducted to reveal the merits of the proposed filtering algorithm in Section 3.4. Finally, Section 3.5 provides the concluding remarks in this chapter.

3.1.1 Background

The AVI data has been categorized by the technology used for data collection in Section 2.3.1. There is an alternative classification criterion. According to the uniqueness of the identifier of each vehicle, there are two types of AVI data. On the one hand, Bluetooth sensors can gather numerous AVI data. However, the MAC address detected by Bluetooth sensors can be provided by either vehicles, passengers within the exact vehicle, or even pedestrians on the roadside through their mobile devices. Therefore, the AVI data obtained in this case is invalid.

On the other hand, for AVI technologies requiring identifier information (e.g., RFID technology uses electromagnetic fields to automatically recognize RFID tags; RFID-tagged vehicles are detected when they pass RFID tag readers fixed at the roadside or under footbridges), AVI data are gathered accurately, but the sampling rate is very few in a relatively short time interval due to privacy issues.

In this chapter, it is worthwhile to distinguish between accurate and valid AVI data. Under these circumstances, the observed path travel times derived from an AVI system can be outliers. The corresponding AVI data are regarded as invalid AVI data, which novel filtering algorithms must remove to extract valid AVI data for use in path travel time prediction. Figure 3.1 illustrates one scenario in which AVI data from an AVI system may be invalid. Vehicle B travels to a shopping mall after being detected by an AVI sensor at the origin. Hence, the experienced path travel time of vehicle B is much longer than that of vehicle A. Therefore, the AVI data from vehicle B (which contains the timestamps with a larger font size in Figure 3.1) is invalid.



Figure 3.1 Example of invalid AVI data in an AVI system

Accurate AVI data refers to data collected from AVI technologies with specific identifiers for vehicles (e.g., RFID and ALPR). The identifier for RFID technology is

the RFID tag, which is fixed at the front of each monitored vehicle. The license plate number of vehicles is used for ALPR. These two identifiers can ensure that observed path travel times from AVI data are from vehicles. On the contrary, Bluetooth MAC addresses are available for vehicles and passengers in the vehicles. Therefore, if one vehicle travels along the study with four passengers holding Bluetooth-enabled cell phones, Bluetooth sensors will obtain five records. It can be challenging to distinguish five vehicles and five Bluetooth devices. Therefore, this type of AVI data is inaccurate.

The availability of identifier information in the database depends on the corresponding privacy issues concerned by different cities (Zhu et al., 2020; Xia et al., 2022). In Hong Kong, only AVI data on commercial vehicles are available for collection. The sampling rate is low without the collection of AVI data on private cars. Therefore, it is challenging to filter AVI data at relatively low sampling rates and use them to predict path travel times.

Some AVI data from RFID tag readers and ALPR cameras may be inappropriate for path travel time prediction. Similar to the data cleaning process of GPS data that can accurately capture the trajectory of vehicles for travel time prediction (Correa and Ozbay, 2022; Gao et al., 2022; Wang et al., 2022; Wang et al., 2022; Zhou et al., 2023; Zhu et al., 2022b), AVI data also need data preprocessing before path travel time prediction.

As discussed by researchers such as Chow et al. (2014) and Robinson and Polak (2006), errors may arise from vehicles being misidentified, stopping en-route (e.g., see Figure 3.1), or choosing unusually long routes (e.g., detours) between two locations that are equipped with AVI sensors. Thus, invalid data (or outliers) are most often obtained if AVI sensors (i) are far apart, implying that vehicle detours or stops are more frequent, or (ii) contain many short-spacing intersections and frequent frontage access (which explains why it can be more difficult to acquire valid AVI data from urban roads than from freeways).

Apart from the categorization by different AVI technologies, as mentioned in Section 2.3.1, AVI data can also be categorized as real-time AVI data or historical AVI data,

depending on when it is gathered. Real-time AVI data are obtained on the current day, while historical AVI data are collected on previous days. Both these two data categories comprise valid and invalid AVI data. As real-time AVI data are generally used for path travel time prediction (Kwong et al., 2009; Chen et al., 2017; Zhan et al., 2020), it is critically important to remove invalid real-time AVI data by novel filtering algorithms to enable path travel time prediction.

Furthermore, for AVI technologies based on identified information in cities with privacy issues, the AVI data is accurate with fewer samples. It is more challenging to distinguish invalid real-time AVI data from limited, accurate real-time AVI data with a low sampling rate. Therefore, this chapter focuses more on the latter.

There are some existing offline algorithms for the filtering of historical AVI data. These algorithms are devoted to the data clustering or modeling of travel time distributions using a large amount of historical AVI data (Kazagli and Koutsopoulos, 2013; Yun et al., 2019b; Duan et al., 2020; Qin et al., 2020; Washington et al., 2020). However, these algorithms lack sufficient computation time to generate validity windows for filtering real-time AVI data. Consequently, various data-filtering algorithms have been developed to screen out invalid real-time AVI data in various ATIS.

Table 3.1 gives the summary of previous related studies for filtering AVI data in the path travel time prediction. First, it is observed in Table 3.1 that all online filtering algorithms can update in a high-frequency manner, with the updating interval (or rolling step to be introduced in Section 3.3.5) Δ ranging from 2 to 15 minutes. Second, the sample size per updating interval Δ is relatively small (e.g., 0-2 samples per updating interval Δ in this chapter). A detailed description of the sampling size of these data will be given in Section 3.4.1. It is a challenge to provide reliable validity windows with insufficient AVI data.

Third, the distance between successive AVI sensors is rather large. As a result, more outliers/invalid AVI data are collected. When this distance increases, the proportion of valid real-time AVI data is smaller, as depicted in Section 3.1.1. Based on the features of AVI data, the validity windows for data filtering are generated. They consist of a

center point and the width of the validity window in general. The dynamic validity windows are determined by the first-order (mean/median) and second-order (variance) properties of real-time AVI data in previous literature. The fixed threshold of $\pm 20\%$ from the mean of observed path travel times from AVI data is adopted by Southwest Research Institute (1998), Mouskos et al. (1998), and TranStar (2021).

Besides, more parameters are introduced to control the center and width of validity windows, considering the weighting of data between previous time intervals (i.e., $t < t_0$) and current time interval t_0 and factors allowing validity windows to be more flexible based on sample sizes. The parameters used in these models/algorithms need calibration from historical AVI data.

Moreover, some previous algorithms (Dion and Rakha, 2006; Tam and Lam, 2008) have taken into account transition traffic conditions. The width of validity windows should be enlarged when the traffic is more congested especially. They checked previous successive time intervals to evaluate the trend of experienced path travel times. These algorithms generally perform well when AVI data is adequate. However, when real-time AVI data is rather limited within a relative short time interval, there may be insufficient real-time information to construct a satisfying dynamic validity window for covering most of the valid AVI data. It indicates that the performance of existing filtering algorithms depends largely on real-time AVI data are limited.

Therefore, there is a need for a novel filtering algorithm capable of effectively extracting real-time AVI data, especially when they are limited. The use of historical AVI data, including both valid and invalid AVI data collected in previous days, can be beneficial in this regard. This chapter extends to consider the variations of path travel time from historical AVI data into the algorithms so that the performance can still be maintained when real-time information is insufficient.

	Updating interval	Sample size	Distance between	Turne of AVI	Validity window	
Related studies	of validity	per updating	two AVI sensors	Type of Avi		
	window Δ (min)	interval	(km)	sensors		
Park and Kim (2018)	5 N	More than 10	0.9–3.7	5.8 GHz	Distribution center of real-time	
			0.9 5.7	DSRC sensors	data ¹	
Ma and Koutsopoulos (2010)	2 and 5	1-3	-	ALPR cameras		
Tam and Lam (2008)	2 1-2	()	RFID tag	Mean and variance of real-time		
		1-2	0.2	readers	data and transition	
Dian and Baltha (2006)	2 2-3	4.0 and 1.0	RFID tag	identification		
Dion and Kakna (2000)		2-3	4.0 and 1.9	readers		
Mouskos et al. (1998);				R FID tag	Mean of real-time data and	
Southwest Research Institute	0.5, 2, and 15	-	-	readars	fixed threshold of window	
(1998); TranStar (2021)				Teauers	width	
This chapter	2 0.2	4.3, 4.5, and	RFID tag	Conditional mean and variance		
This chapter	2	0-2	9.2	readers	of real-time and historical data	

Table 3.1 Summary of previous related studies for filtering AVI data in the path travel time prediction

¹ The real-time data refers to the AVI data collected on the current day.

The effect of complex network structures in urban areas has not been investigated in most previous studies on filtering AVI data. Attention has been mainly given to freeways (Diaz et al., 2016), which have relatively simple network topologies and very few entries and exits between pairs of AVI sensors. Moreover, the numerous entries, exits, and bus stops along the urban study paths (as used in the case study) may indicate that valid real-time AVI data can be limited, which adversely affects the performance of existing filtering algorithms used for path travel time prediction.

When valid real-time AVI data is limited, it is worthwhile to model temporal covariances of path travel times by 2-minute intervals at different time intervals and on different days from historical AVI data. They are significantly beneficial for filtering out invalid real-time AVI data and for path travel time prediction.

A novel filtering algorithm is proposed to filter out invalid real-time AVI data for path travel time prediction without ground truth for training purposes. As no ground truth is used for training, it is also referred to as the proposed unsupervised algorithm in this chapter.

The proposed unsupervised algorithm is particularly useful when privacy policies prohibit the availability of many valid AVI data from privately owned vehicles (e.g., Hong Kong only allows the collection of AVI data from commercial vehicles. These commercial vehicles include goods vehicles, non-franchised and franchised buses, and private cars owned by commercial companies, which account for approximately 19% of the total vehicle fleet in Hong Kong¹) for utilization in the development of various ATIS.

Furthermore, most existing filtering algorithms use simple first-order central tendency measures, such as observed mean or median values, of AVI data. In contrast, the proposed unsupervised algorithm considers both first- and second-order statistical properties of AVI data via a functional principal component analysis (FPCA). The mean and standard deviation of predicted path travel times by FPCA can help to

 $^{{}^{1}}https://www.td.gov.hk/en/transport_in_hong_kong/transport_figures/index.html$

construct a dynamic validity window for filtering out invalid real-time AVI data for path travel time prediction on urban arterials.

FPCA is a statistical tool for functional data analysis that uses advanced feature approximation techniques. It has received increasing attention in recent related studies, as it can be used for analyzing highly stochastic data. For example, Chiou (2012) proposed an FPCA model to predict traffic flows, and Guardiola et al. (2014) and Wagner-Muns et al. (2018) used FPCA to identify and monitor traffic patterns. In addition, Chiou et al. (2021) applied FPCA to model the variability and reliability of freeway travel times. Furthermore, Chen and Müller (2014) performed FPCA of GPS data to forecast vehicle speed distributions. Moreover, Zhong et al. (2017) further highlighted the merits of FPCA on path travel time predictions under abnormal traffic conditions.

The FPCA model regards the path travel time as a stochastic process (Celikoglu, 2013b; Zhong et al., 2017, 2020). In this chapter, the FPCA model has been extended to generate temporal covariances of path travel times by 2-minute intervals. These relationships are then used to develop the proposed unsupervised algorithm for filtering limited but accurate real-time AVI data, which enables the prediction of path travel times without ground truth for training purposes.

3.1.2 Contributions

In general, the major contributions of this chapter are summarized into the following three categories.

C3.1 A novel unsupervised algorithm is proposed, with the usage of historical AVI data without using historical ground truth for training purposes, for constructing dynamic validity windows to filter out invalid real-time AVI data from limited real-time AVI data.

C3.2 A FPCA-based model is adapted to consider both the historical and real-time AVI data for modeling their temporal covariances of path travel times by 2-minute intervals at different time intervals and on different days. Both mean and standard deviation of the predicted path travel times are provided and used for improving of the filtering performance of real-time AVI data.

C3.3 Sensitivity tests are conducted to examine the effects of different sampling rates of the real-time AVI data or the valid real-time AVI data only in order to verify the robustness of the proposed unsupervised algorithm without or with the use of the ground truth for training purposes.

3.2 Problems under Investigation

In this chapter, any given path p with two AVI sensors at both ends is studied. As multisource traffic data and multi-class path travel times are to be mentioned in Chapter 4, the data source and vehicle class are also defined in this chapter to maintain consistency. In this setting, *i*-th observed path travel time for path p of vehicle class kfor from data source s_A on day d is denoted as $y_{i,d,p}^{s_A,k}$. $\tau_{i,d,d_p,p}^{s_A,k}$ is the timestamp of the *i*-th observed path travel time at AVI sensor location d_p for path p of vehicle class kfrom data source s_A on day d. The other symbols in this chapter omit the notation of vehicle class, as this chapter only considers the overall vehicles as one class. The set of days with historical AVI data is denoted as D. The assignment of d from set D.

The proposed unsupervised algorithm aims to provide a dynamic validity window for screening out invalid real-time AVI data. The dynamic validity window consists of the upper bound U(t) and the lower bound L(t) for each time interval t. In the proposed unsupervised algorithm, the available data are the real-time AVI data before time interval t on the current day and historical AVI data.

3.3 Techniques for Filtering AVI Data

3.3.1 Proposed unsupervised algorithm

Figure 3.2 presents the framework of the proposed unsupervised algorithm. There are two stages: online filtering and offline training. In each of these two stages, there are five different steps in the proposed framework with the corresponding equation numbers shown at each of these five steps. The following paragraphs give detailed descriptions of these two stages and their corresponding five steps.

Stage 1 involves offline training, which uses historical AVI data for the development of the trained FPCA. Stage 2 concerns the online filtering of real-time AVI data, in which the trained FPCA models are used to construct dynamic validity windows to screen out invalid real-time AVI data.

As the backbone of the methodology framework, FPCA models are trained to map the predictor to the response (Zhong et al., 2017). Thus, the eigenfunctions and principal components must be trained for the predictor and the response. Then, the conditional distributions of the response based on the predictor can be obtained and represented by the trained eigenfunctions and principal components (Yao et al., 2005; Müller and Yao, 2008; Chen and Müller, 2014).



Figure 3.2 Framework of the proposed unsupervised algorithm

In the proposed unsupervised algorithm, the historical AVI data is considered as the predictor, and the response is the offline predicted path travel time from sufficient historical AVI data using existing filtering algorithms, such as TransGuide algorithm

(Southwest Research Institute, 1998). They are preliminary predicted travel time shown in Step 1 without the use of ground truth. Therefore, the i^{th} preliminary predicted travel time is denoted as $y_{i,d,p}^{s_G,k}$.

It is as shown in Section 3.3.2 that selecting appropriate training set is performed in Step 2 (shown in the orange box). It is based on the $\Sigma_t^{s_A}(d_i, d_j)$, which is the day-today covariance of path travel times by 2-minute intervals at time interval t from data source s_A between day d_i and d_j , for $i, j \in D$. Afterward, the learning training sets are proceeded with modeling of the within-day covariance of path travel times by 2-minute intervals ($\Sigma_d^{s_A}(t_a, t_b)$), which is within-day covariance of path travel times on day dfrom data source s_A between time interval t_a and t_b , for $a, b \in \delta$) in Step 3. Dynamic validity windows are constructed based on the mean and standard deviation of the predicted path travel times by 2-minute intervals provided in Step 4. Dynamic validity windows can filter out invalid real-time AVI data in a rolling horizon scheme in Step 5.

An illustrative example is given for illustration of using $\Sigma_t^{s_A}(d_i, d_j)$ in Step 2 (in the orange box) based on AVI data collected at the selected path in 2017. For each 2-minute interval, Step 2 is required to screen out some historical days that are less relevant to the specific traffic conditions on the current day. There is a total of 299 historical weekdays with AVI data collected in 2017. However, after considering $\Sigma_t^{s_A}(d_i, d_j)$, the number of historical days $|D^*|$ varies from 253 to 279 after Step 2 for each time interval *t* due to different traffic conditions by time of day.

3.3.2 Selecting appropriate training set

Historical AVI data may reflect different traffic patterns due to the changing traffic demand and network supply (e.g., incidents and sensor failures). If the traffic patterns are different from that of the current day, then those historical AVI data may provide little useful information for constructing the current day's dynamic validity windows. Accordingly, historical AVI data that contain similar traffic patterns to the current day are selected for filtering real-time AVI data.

To this end, the day-to-day covariances of path travel times by 2-minute intervals are modeled by FPCA to reflect the similarities of traffic patterns across multiple days. At timestamp τ , historical AVI data at time $\tau_{i,d,d_p,p}^{s_A,k} \in [\tau - \delta, \tau]$ is considered, where the δ is the length of the study horizon and it is the unit for the rolling horizon scheme presented later.

The observed path travel time is the sum of travel time and measurement error $\varepsilon_{i,d}$, and are given by Eq. (3.1):

$$y_{i,d,p}^{s_{A},k} = \mu_{T^{s_{A}}}(d) + \sum_{k=1}^{K^{s_{A},D}} \xi_{k}^{s_{A},D} \phi_{k}^{s_{A},D}(d) + \varepsilon_{i,d,p}^{s_{A},k}$$
(3.1)

where $\mu_{T^{s_A}}(d)$ is the mean function of travel times from AVI data on day d, which is $T_{d,i}^{s_A,p}$; $\xi_k^{s_A,D}$ is the score of the k^{th} functional principal component; $\phi_k^{s_A,D}(d)$ is the eigenfunction of the k^{th} functional principal component from AVI data on day d for |D| days according to the Karhunen-Loève representation; and $K^{s_A,D}$ is the number of functional principal components from AVI data for |D| days, where $\varepsilon_{i,d,p}^{s_A,k}$ represents the measurement error of *i*-th observed path travel time for path p of vehicle class k from data source s_A (AVI data) on day d.

Eq. (3.2) assumes that the path travel time in $[\tau - \delta, \tau]$ is continuous in *d*, and the corresponding path travel time function from the AVI data $T^{s_A}(d)$ is obtained from Eq.

(3.2):

$$T^{s_A}(d) = \mu_{T^{s_A}}(d) + \sum_{k=1}^{K^{s_A,D}} \xi_k^{s_{A,D}} \phi_k^{s_{A,D}}(d)$$
(3.2)

where the function $\mu_{T^{s_A}}(d)$ is expressed by

$$\mu_{T^{s_A}}(d) = E(T^{s_A}(d)) \tag{3.3}$$

The day-to-day covariance of path travel times from s_A between day d_1 and d_2 is denoted by $\Sigma_t^{s_A}(d_i, d_j)$ and is obtained by Eq. (3.4), as below:

$$\Sigma_t^{s_A}(d_i, d_j) = \sum_{k=1}^{K^{s_A, D}} \lambda_k^{s_A, D} \phi_k^{s_A, D}(d_1) \phi_k^{s_A, D}(d_2)$$
(3.4)

where $\lambda_k^{s_{A,D}}$ is the eigenvalue of the k^{th} functional principal component from AVI data.

It is assumed that the weighting or score of the functional principal component $\lambda_k^{S_A,D}$ has the statistical properties given by Eqs. (3.5) and (3.6), as below:

$$E\left(\xi_k^{s_{A,D}}\right) = 0 \tag{3.5}$$

$$Var(\xi_k^{s_A,D}) = \lambda_k^{s_A,D}$$
(3.6)

The covariance $\Sigma_t^{s_A}(d_i, d_j)$ is derived by solving the following minimization objective (3.7) for the AVI data:

$$\min_{\beta_{0},\beta_{1},\beta_{2}} \sum_{1 \le d_{3} \le d_{4}}^{|D|} \sum_{i=1}^{N_{d,\delta}} \kappa_{c} \left(\frac{d_{3}-d_{1}}{h_{c}}\right) \kappa_{c} \left(\frac{d_{4}-d_{2}}{h_{c}}\right) \cdot \left(\widehat{Cov}\left(T_{d_{3},i}^{S_{A}}, T_{d_{4},i}^{S_{A}}\right) - \beta_{0} - \beta_{1}(d_{3}-d_{1})\right)^{2} \quad (3.7)$$

where $\widehat{Cov}(T_{d_3,i}^{s_A}, T_{d_4,i}^{s_A})$ represents the predicted travel time covariance between day d_3 and d_4 , the estimates of the model coefficients $\beta_0, \beta_1, \beta_2$ are dependent on days d_1 and d_2 , and $N_{d,\delta}$ is the number of samples within the study horizon δ on day d. The estimates of β_0 are denoted as $\hat{\beta}_0(d_1, d_2)$ and an estimate of $\Sigma_t^{s_A}(d_1, d_2)$ is obtained from $\hat{\Sigma}_t^{s_A}(d_1, d_2) = \hat{\beta}_0(d_1, d_2)$. Moreover, κ_{s_A} is a kernel function in which h_{s_A} is the bandwidth that enables calibration of the covariance function.

Referring to the covariance function of path travel times for different days $\hat{\Sigma}_t^{s_A}(d_1, d_2)$,

the samples with larger covariance values are selected and used to calibrate the model. D^* is the set of days after sample selection, which is determined by Eq. (3.8):

$$D^* = \left\{ d \left| \left| \hat{\Sigma}_t^{s_A}(d_1, d_2) \right| \ge \Sigma^*, d \in D \right\}$$
(3.8)

where Σ^* is the threshold of the path travel time covariance between different days.

3.3.3 Learning training set

Two FPCA models are utilized to model temporal covariances of path travel times by 2-minute intervals at different time intervals. The first FPCA model is based on the predictor (i.e., historical AVI data). The second FPCA model is based on the responses, which are preliminary predicted travel time in the proposed unsupervised algorithm. The historical AVI data $x_{i,d}$ is modeled in Eq. (3.9), as follows:

$$y_{i,d,p}^{s_{A},k} = \mu_{T^{s_{A}}}(t) + \sum_{k=1}^{K^{s_{A},\delta}} \xi_{k}^{s_{A},\delta} \phi_{k}^{s_{A},\delta}(t) + \varepsilon_{i,d,p}^{s_{A},k} d \in D^{*}$$
(3.9)

where $\mu_{T^{s_A}}(t)$ is the mean function of the observed path travel times at time interval t; $\xi_k^{s_A,\delta}$ represents the score/weight of the kth functional principal component; $\phi_k^{s_A,\delta}(t_{d,i}^{s_A})$ is the eigenfunction of the kth functional principal component from AVI data at time $t_{d,i}^{s_A}$; K^T is the number of functional principal components from AVI data during study horizon T.

Analogously, the path travel time function based on AVI data $T^{S_A}(t)$ can be described as Eq. (3.10):

$$T^{s_A}(t) = \mu_{T^{s_A}}(t) + \sum_{k=1}^{K^{s_A,\delta}} \xi_k^{s_A,\delta} \phi_k^{s_A,\delta}(t_{d,i}^{s_A})$$
(3.10)

where $\mu_{T^{s_A}}(t)$ is given by Eq. (3.11), as below:

$$\mu_{T^{s_A}}(t) = E(T^{s_A}(t)) \tag{3.11}$$

 $\Sigma_d^{s_A}(t_a, t_b)$ is denoted as the within-day covariance of path travel times from s_A between time t_1 and t_2 in Eq. (3.12), as below:

$$\Sigma_{d}^{s_{A}}(t_{a},t_{b}) = \sum_{k=1}^{K^{s_{A},\delta}} \lambda_{k}^{s_{A},\delta} \phi_{k}^{s_{A},\delta}(t_{1}) \phi_{k}^{s_{A},\delta}(t_{2})$$
(3.12)

Again, the weighting/score of functional principal components has the same statistical properties as shown in Eqs. (3.5) and (3.6).

If a response $y_{i,d,p}^{s_G,k}$ is available at time interval *t* on day *d*, Eq. (3.9) can be expressed as Eq. (3.13):

$$y_{i,d,p}^{s_G,k} = \mu_{T^{s_G}}(t) + \sum_{k=1}^{K^{s_G,\delta}} \xi_k^{s_G} \phi_k^{s_G}(t), \, d \in D^*$$
(3.13)

where $\mu_{T^{s_G}}(t)$ is the mean function of responses over study horizon T; $\xi_k^{s_G}$ is the score/weight of the k^{th} functional principal component of the responses; $\phi_k^{s_G}(t)$ is the eigenfunction of the k^{th} functional principal component of the responses at time interval t; and $K^{s_G,\delta}$ is the number of functional principal components of the responses during study horizon δ .

Correspondingly, the path travel time function based on the responses can be expressed as Eq. (3.14), in below:

$$T^{s_G}(t) = \mu_{T^{s_G}}(t) + \sum_{k=1}^{K^{s_G,\delta}} \xi_k^{s_G} \phi_k^{s_G}(t)$$
(3.14)

where $\mu_{T^{s_G}}(t)$ given by Eq. (3.15), as follows:

$$\mu_{T^{s_{G}}}(t) = E(T^{s_{G}}(t)) \tag{3.15}$$

 $\Sigma_d^{s_G}(t_1, t_2)$ is denoted as the within-day covariance of path travel times for responses between time t_1 and t_2 during study horizon δ , as below:

$$\Sigma_{d}^{s_{G}}(t_{1},t_{2}) = \sum_{k=1}^{K^{s_{G},\delta}} \lambda_{k}^{s_{G}} \phi_{k}^{s_{G}}(t_{1}) \phi_{k}^{s_{G}}(t_{2})$$
(3.16)

where $\lambda_k^{s_G}$ is the eigenvalue of the k^{th} functional principal component of responses.

The predictors $y_{i,d,p}^{s_A,k}$ and responses $y_{i,d,p}^{s_G,k}$ as shown in Section 3.3.1 can be used to calibrate the above-described FPCA-based models. The details of the procedure for calibrating mean functions, covariance functions, and functional principal components (including weighting/score and eigenfunctions) are available in the literature (Yao et al., 2005; Müller and Yao, 2008; Chen and Müller, 2014; Zhong et al., 2017). The number of functional principal components is generally determined by applying one of the following three methods: the fraction of variance explained, the Akaike information criterion, or the Bayesian information criterion.

3.3.4 Constructing dynamic validity window

The principal analysis by conditional expectation (PACE) is now formulated for the FPCA models presented in the previous section, for use in data filtering. The objective is to relate the models derived from the predictors and the responses via the method of additive models (Müller and Yao, 2008; Chen and Müller, 2014; Zhong et al., 2017). Specifically, the conditional distributions of the responses derived from the AVI data are adopted. The advantage of this PACE approach is its superiority over other approaches under the Gaussian assumption (Ji and Müller, 2017).

Application of the functional additive model (Müller and Yao, 2008) provides the conditional model as below:

$$E(T^{s_G}(t)|T^{s_A}(t)) = \mu_{T^{s_G}}(t) + \sum_{q=1}^{K^{s_G,\delta}} \left(\sum_{k=1}^{K^{s_A,\delta}} E\left(\xi_q^{s_G,\delta}|\xi_k^{s_A,\delta}\right)\right) \phi_q^{s_G,\delta}(t)$$
(3.17)

Similar to the calibration procedure adopted in the general FPCA model, $f_{qk}(\xi_k^{s_A,\delta}) = E\left(\xi_q^{s_G,\delta} | \xi_k^{s_A,\delta}\right)$ on each day d, $f_{qk}(\xi)$ can be obtained by minimizing the following expression with respect to γ_0 and γ_1 :

$$\min_{\gamma_0,\gamma_1} \sum_{d \in D^*} \kappa_f \left(\frac{\hat{\xi}_{k,d}^{s_A,\delta} - \xi}{h_f} \right) \left[\hat{\xi}_{k,d}^{s_G,\delta} - \gamma_0 - \gamma_1 \left(\xi - \hat{\xi}_{k,d}^{s_A,\delta} \right) \right]^2$$
(3.18)

where $\hat{\xi}_{k,d}^{s_A,\delta}$ and $\hat{\xi}_{k,d}^{s_G,\delta}$ are the estimated $\xi_k^{s_A,\delta}$ and $\xi_k^{s_G,\delta}$, respectively, on each day d. This leads to $\hat{f}_{qk}(\xi) = \hat{\gamma}_0(\xi)$. Moreover, the conditional covariance function is acquired by Eq. (3.19):

$$Cov(T^{s_{G}}(t_{1}), T^{s_{G}}(t_{2})|T^{s_{A}}(t)) = \sum_{q=1}^{K^{s_{G},\delta}} var(\xi_{q}^{s_{G},\delta}|T^{s_{A}}(t)) \phi_{q}^{s_{G},\delta}(t_{1}) \phi_{q}^{s_{G},\delta}(t_{2})$$
(3.19)

By using the property of variance, $\operatorname{var}\left(\left.\xi_{q}^{s_{G},\delta}\right|T^{s_{A}}(t)\right)$ can be further expanded such that Eq. (3.19) can be rewritten as Eq. (3.20):

$$Cov\left(T^{s_{G}}(t_{1}), T^{s_{G}}(t_{2}) \middle| T^{s_{A}}(t)\right)$$

$$= \sum_{q=1}^{K^{s_{G},\delta}} \left[var(\xi_{q}^{s_{G},\delta}) + \sum_{k=1}^{K^{s_{A},\delta}} E\left(\left(\xi_{q}^{s_{G},\delta}\right)^{2} - var(\xi_{q}^{s_{G},\delta}) \middle| \xi_{k}^{s_{A},\delta}\right) - E^{2}(\xi_{q}^{s_{G},\delta} \middle| \xi_{k}^{s_{A},\delta}) \right]$$

$$\cdot \phi_{q}^{s_{G},\delta}(t_{1}) \phi_{q}^{s_{G},\delta}(t_{2})$$

$$= H^{s_{G}}(t_{1},t_{2}) + \sum_{q=1}^{K^{s_{G},\delta}} \sum_{k=1}^{K^{s_{A},\delta}} \left[g_{q_{k}}(\xi_{k}^{s_{A},\delta}) - f_{q_{k}}^{2}(\xi_{k}^{s_{A},\delta})\right] \phi_{q}^{s_{G},\delta}(t_{1}) \phi_{q}^{s_{G},\delta}(t_{2})$$

$$= H^{s_{G}}(t_{1},t_{2}) + \sum_{q=1}^{K^{s_{G},\delta}} \sum_{k=1}^{K^{s_{A},\delta}} \left[g_{q_{k}}(\xi_{k}^{s_{A},\delta}) - f_{q_{k}}^{2}(\xi_{k}^{s_{A},\delta})\right] \phi_{q}^{s_{G},\delta}(t_{1}) \phi_{q}^{s_{G},\delta}(t_{2})$$

$$(3.20)$$

where $g_{qk}(\xi_k^{s_A,\delta})$ is given by Eq. (3.21), as follows:

$$g_{qk}\left(\xi_{k}^{s_{A},\delta}\right) = E\left[\left(\xi_{q}^{s_{G},\delta}\right)^{2} - \operatorname{var}\left(\xi_{q}^{s_{G},\delta}\right) \left|\xi_{k}^{s_{A},\delta}\right]$$
(3.21)

By setting $f_{qk}(\xi_k^{s_A,\delta}) = \hat{f}_{qk}(\xi_k^{s_A,\delta})$, an estimate of $g_{qk}(\xi_k^{s_A,\delta})$ can be further acquired by minimizing objective (3.22) with respect to η_0 and η_1 :

$$\min_{\eta_{0},\eta_{1}} \sum_{d \in D^{*}} \kappa_{g} \left(\frac{\hat{\xi}_{k,d}^{s_{A},\delta} - \xi_{k}^{s_{A},\delta}}{h_{g}} \right) \left[\hat{\xi}_{q,d}^{s_{G},\delta^{2}} - \operatorname{var}\left(\hat{\xi}_{q,d}^{s_{G},\delta} \right) - \eta_{0} - \eta_{1} \left(\xi_{k}^{s_{A},\delta} - \hat{\xi}_{k,d}^{s_{A},\delta} \right) \right]^{2}$$
(3.22)

which leads to $\hat{g}_{qk}\left(\xi_k^{s_A,\delta}\right) = \hat{\eta}_0\left(\xi_k^{s_A,\delta}\right).$

The conditional mean of responses based on path travel times derived from AVI data and the conditional covariance of responses based on travel times from AVI data can be modeled as Eqs. (3.23) and (3.24), respectively:

$$\hat{E}(T^{s_G}(t)|T^{s_A}(t)) = \mu_{T^{s_G}}(t) + \sum_{q=1}^{K^{s_G,\delta}} \left(\sum_{k=1}^{K^{s_A,\delta}} \hat{f}_{qk}(\xi_k^{s_A,\delta}) \right) \hat{\phi}_q^{s_G,\delta}(t)$$
(3.23)

$$\widehat{Cov}\left(T^{s_{G}}(t_{1}), T^{s_{G}}(t_{2}) \middle| T^{s_{A}}(t)\right) = \sum_{q=1}^{K^{s_{G},\delta}} \left(\operatorname{var}\left(\hat{\xi}_{q}^{s_{G},\delta}\right) + \sum_{k=1}^{K^{s_{A},\delta}} \left(\hat{g}_{qk}\left(\xi_{k}^{s_{A},\delta}\right) - \hat{f}_{qk}^{2}\left(\xi_{k}^{s_{A},\delta}\right)\right)\right) \hat{\phi}_{q}^{s_{G},\delta}(t_{1}) \hat{\phi}_{q}^{s_{G},\delta}(t_{2})$$

$$(3.24)$$

The conditional mean and covariance function of responses derived from path travel times determined from AVI data can be obtained from Eqs. (3.23) and (3.24) by learning from historical information on the predictors and responses.

Proposition 3.1 presents the uniform convergence properties of the conditional model of path travel time prediction. The conditional mean and covariance of predicted path travel times are accurate when the number of observations of AVI data $|D^*| \rightarrow +\infty$. If more principal components are considered (i.e., $K^{s_A,\delta}, K^{s_G,\delta}$ is large), more data samples are required.

Proposition 3.1. (The uniform convergence of the conditional modeling of path travel time)

Suppose that the number of travel time data $|D^*| \to +\infty$ and the path travel time on each day in D^* are i.i.d., and that the mean $\hat{E}(T^{s_G}(t)|T^{s_A}(t))$ and the covariance $\widehat{Cov}(T^{s_G}(t_1), T^{s_G}(t_2)|T^{s_A}(t))$ in the calibrated conditional model of travel time in Eqs. (3.23) and (3.24) approximate the actual conditional mean and covariance with the error rate $O_p\left(\frac{K^{s_A,\delta}K^{s_G,\delta}}{\sqrt{|D^*|}}\right)$. Thus, mathematically Eqs. (3.25) and (3.26) are presented:

$$\begin{split} \sup_{t\in\delta} \left| \hat{E} \left(T^{s_{G}}(t) \left| T^{s_{A}}(t) \right) - E \left(T^{s_{G}}(t) \left| T^{s_{A}}(t) \right) \right| &= O_{p} \left(\frac{K^{s_{A},\delta}K^{s_{G},\delta}}{\sqrt{|D^{*}|}} \right) \\ \widehat{Cov} \left(T^{s_{G}}(t_{1}), T^{s_{G}}(t_{2}) \left| T^{s_{A}}(t) \right) \right) \\ &= \sup_{t\in\delta} \left| \left(T^{s_{G}}(t_{1}), T^{s_{G}}(t_{2}) \left| T^{s_{A}}(t) \right) - Cov \left(T^{s_{G}}(t_{1}), T^{s_{G}}(t_{2}) \left| T^{s_{A}}(t) \right) \right) \right| \\ &= O_{p} \left(\frac{K^{s_{A},\delta}K^{s_{G},\delta}}{\sqrt{|D^{*}|}} \right) \end{split}$$
(3.25)

It is straightforward to prove that the following (3.27) and (3.28) hold:

$$\sup_{t\in\delta} \left| \hat{f}_{qk} \left(\xi_k^{s_A, \delta} \right) - f_{qk} \left(\xi_k^{s_A, \delta} \right) \right| = O_P \left(\frac{1}{\sqrt{|D^*|}} \right)$$
(3.27)

$$\sup_{t\in\delta} \left| \hat{g}_{qk} \left(\xi_k^{s_A,\delta} \right) - g_{qk} \left(\xi_k^{s_A,\delta} \right) \right| = O_P \left(\frac{1}{\sqrt{|D^*|}} \right)$$
(3.28)

Then, for the conditional mean, it is derived that:

$$\begin{split} \sup_{t\in\delta} \left| \hat{E}(T^{s_{G}}(t)|T^{s_{A}}(t)) - E(T^{s_{G}}(t)|T^{s_{A}}(t)) \right| \\ &= \sup_{t\in\delta} \left| \left(\hat{\mu}_{T^{s_{G}}}(t) + \sum_{q=1}^{K^{s_{G},\delta}} \left(\sum_{k=1}^{K^{s_{A},\delta}} \hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) \right) \hat{\phi}_{q}^{s_{G},\delta}(t) \right) - \left(\mu_{T^{s_{G}}}(t) + \sum_{q=1}^{K^{s_{G},\delta}} \left(\sum_{k=1}^{K^{s_{A},\delta}} \hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) \right) \phi_{q}^{s_{G},\delta}(t) \right) \right| = \\ \sup_{t\in\delta} \left| \left(\hat{\mu}_{T^{s_{G}}}(t) - \mu_{T^{s_{G}}}(t) \right) + \sum_{q=1}^{K^{s_{G},\delta}} \left(\sum_{k=1}^{K^{s_{A},\delta},\delta} \left(\hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) \right) \right) \phi_{q}^{s_{G},\delta}(t) - \sum_{q=1}^{K^{s_{G},\delta}} \left(\sum_{k=1}^{K^{s_{A},\delta},\delta} \left(\hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) \right) \right) \phi_{q}^{s_{G},\delta}(t) - \sum_{q=1}^{K^{s_{G},\delta}} \left(\sum_{k=1}^{K^{s_{A},\delta},\delta} \left(\hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) \right) \right) \phi_{q}^{s_{G},\delta}(t) - \sum_{q=1}^{K^{s_{G},\delta}} \left(\sum_{k=1}^{K^{s_{A},\delta},\delta} \left(\hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) \right) \right) \phi_{q}^{s_{G},\delta}(t) - \sum_{q=1}^{K^{s_{G},\delta},\delta} \left(\sum_{k=1}^{K^{s_{A},\delta},\delta} \left(\hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) \right) \right) \phi_{q}^{s_{G},\delta}(t) - \sum_{q=1}^{K^{s_{G},\delta},\delta} \left(\sum_{k=1}^{K^{s_{A},\delta},\delta} \left(\hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) \right) \right) \phi_{q}^{s_{G},\delta}(t) - \sum_{q=1}^{K^{s_{G},\delta},\delta} \left(\sum_{k=1}^{K^{s_{A},\delta},\delta} \left(\sum_{k=1}^{K^{s_{G},\delta},\delta} \left(\sum_{k=1}^{K^{s_{G},\delta},\delta} \left(\hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) - f_{qk}(\xi_{k}^{s_{A},\delta}) \right) \right) \left(\hat{\phi}_{q}^{s_{G},\delta}(t) - \phi_{q}^{s_{G},\delta}(t) \right) \right| \\ + \sum_{q=1}^{K^{s_{G},\delta}} \left(\sum_{k=1}^{K^{s_{G},\delta},\delta} \left(\hat{f}_{qk}(\xi_{k}^{s_{A},\delta}) - f_{qk}(\xi_{k}^{s_{A},\delta}) \right) \right) \phi_{q}^{s_{G},\delta}(t) \right| \\ = O_{p} \left(\frac{1}{\left| D^{*} \right|} \right) + O_{p} \left(\frac{K^{s_{A},\delta}K^{s_{G},\delta}}{\left| D^{*} \right|} \right) + O_{p} \left(\frac{K^{s_{A},\delta}K^{s_{G},\delta}}{\left| D^{*} \right|} \right) = O_{p} \left(\frac{K^{s_{A},\delta}K^{s_{G},\delta}}{\left| D^{*} \right|} \right)$$

By using the same decomposition techniques for the conditional covariance, it is obtained that: $\sup_{t\in\delta} |\widehat{Cov}(T^{s_G}(t_1), T^{s_G}(t_2)|T^{s_A}(t)) - Cov(T^{s_G}(t_1), T^{s_G}(t_2)|T^{s_A}(t))|$

$$= \sup_{t \in \delta} \left| \sum_{q=1}^{K^{S_G,\delta}} \left(var(\xi_k^{S_G,\delta}) + \sum_{k=1}^{K^{S_A,\delta}} \left(\hat{g}_{qk}(\xi_k^{S_A,\delta}) - \hat{f}_{qk}^2(\xi_k^{S_A,\delta}) \right) \right) \hat{\phi}_q^{S_G,\delta}(t_1) \hat{\phi}_q^{S_G,\delta}(t_2) \right|$$
(3.30)
$$- \sum_{q=1}^{K^{S_G,\delta}} \left(var(\xi_q^{S_G,\delta}) - \sum_{k=1}^{K^{S_A,\delta}} \left(g_{qk}(\xi_k^{S_A,\delta}) - f_{qk}^2(\xi_k^{S_A,\delta}) \right) \right) \phi_q^{S_G,\delta}(t_1) \phi_q^{S_G,\delta}(t_2) \right|$$

$$\begin{split} & \leq \sup_{t \in \delta} \left| \sum_{q=1}^{K^{S_G,\delta}} var(\xi_k^{s_G,\delta}) - \sum_{q=1}^{K^{S_G,\delta}} var(\xi_q^{s_G,\delta}) \right| \\ & + \sup_{t \in \delta} \left| \sum_{q=1}^{K^{S_G,\delta}} \sum_{k=1}^{K^{S_G,\delta}} \left(\hat{g}_{qk}(\xi_k^{s_A,\delta}) - \hat{f}_{qk}^2(\xi_k^{s_A,\delta}) \right) \hat{\phi}_q^{s_G,\delta}(t_1) \hat{\phi}_q^{s_G,\delta}(t_2) \right| \\ & - \sum_{q=1}^{K^{S_G,\delta}} \sum_{k=1}^{K^{S_A,\delta}} \left(g_{qk}(\xi_k^{s_A,\delta}) - f_{qk}^2(\xi_k^{s_A,\delta}) \right) \hat{\phi}_q^{s_G,\delta}(t_1) \hat{\phi}_q^{s_G,\delta}(t_2) \right| \\ & \leq O_p\left(\frac{K^{S_G,\delta}}{\sqrt{|D^*|}} \right) + O_p\left(\frac{K^{S_A,\delta_K^S_G,\delta}}{\sqrt{|D^*|}} \right) + O_p\left(\frac{K^{S_A,\delta_K^S_G,\delta}}{\sqrt{|D^*|}} \right) \\ & \leq O_p\left(\frac{K^{S_A,\delta_K^S_G,\delta}}{\sqrt{|D^*|}} \right) \end{split}$$

Based on the conditional mean and covariance function of predicted path travel times, U(t) and L(t) as upper and lower bounds the dynamic validity windows are from Eqs. (3.31) and (3.32):

$$U(t) = \hat{E}(T^{s_G}(t)|T^{s_A}(t)) + Z_{\frac{\alpha}{2}} \cdot \widehat{Cov}(T^{s_G}(t_1), T^{s_G}(t_2)|T^{s_A}(t))^{\frac{1}{2}}$$
(3.31)

$$L(t) = \hat{E}(T^{s_G}(t)|T^{s_A}(t)) - Z_{\frac{\alpha}{2}} \cdot \widehat{Cov}(T^{s_G}(t_1), T^{s_G}(t_2)|T^{s_A}(t))^{\frac{1}{2}}$$
(3.32)

The invalid real-time AVI data can then be filtered out if they are not falling within the dynamic validity window [L(t), U(t)]. At each time interval t, the validity window is updated based on the rolling horizon scheme to be introduced in the following Section 3.3.5.

3.3.5 Rolling horizon scheme

The rolling horizon scheme is adopted as previous real-time applications (Pan et al., 2013; Zhong et al., 2017). The dynamic validity windows governed by U(t) and L(t) (which are determined from the proposed unsupervised algorithm) are updated in each time interval t^* , when these new real-time AVI data are streamed for filtering. The filtering framework generates the dynamic validity windows for each rolling step Δ (i.e., 2 min) using the flexible and adaptive rolling horizon (study horizon) δ . In contrast, most existing data filtering algorithms adopted a fixed rolling horizon for their applications (Pan et al., 2013; Zhong et al., 2017).

There are three factors to be distinguished: prediction horizon Δt , which is the time range into the future that the model is predicting; rolling step Δ , which is the interval

at which the model updates its predictions; rolling horizon δ , which the time range to consider real-time data for prediction.

An illustrative example is given in Figure 3.3. The green box represents the chosen rolling horizon (10 min), the red box illustrates the pre-determined prediction horizon (10 min), and the blue box shows the selected rolling step (2 min). For path 1 (urban arterial) with free-flow travel time of 8.4 min, the upper and lower bounds of the dynamic validity window are 23.89 min and 12.96 min, respectively at 10:00-10:02 (t_0) on January 10th, 2018 (Wednesday). However, for the next 2-minute interval (t_1) . They change to 23.13 min and 12.54 min, respectively. The effects of the lengths of rolling step, prediction horizon, and rolling horizon on the prediction accuracy of the path travel times will be elaborated in the following paragraphs.



Figure 3.3 An illustrative example on the rolling horizon scheme

The following discussion is concerned with the impacts of these three factors on the prediction performance on the path travel times.

A longer prediction horizon (Δt) may lead to decreased accuracy due to the increasing likelihood of unforeseen events affecting the system state over time. This can include incidents, changes in weather, or other dynamic factors that are difficult to predict far

in advance. Conversely, a shorter prediction horizon might limit the predictive model's ability to forecast future conditions effectively.

A larger rolling step (Δ) in the prediction model implies that the prediction will be updated less frequently, which can result in slower response to real-time changes in the traffic conditions. This delay can cause the predictions to become less accurate as they do not reflect the most current traffic condition. If the rolling step is too short, it would however increase the computational load.

A longer rolling horizon (δ) can lead to accumulated errors in the predictions, as each successive prediction is based on the previous one, potentially magnifying any initial inaccuracies. On the contrary, A shorter rolling horizon might not fully utilize real-time data on the current day for the predictions, resulting in less stable and potentially less accurate forecasts.

3.4 Case Studies

The proposed unsupervised algorithm is examined in case studies of two selected paths using real-world data collected from the Hong Kong urban road network.

3.4.1 Traffic data

The historical ground truth on path travel times is obtained from the Hong Kong JTIS, the predicted path travel time of which has been independently validated using floating car survey data (Tam and Lam, 2008, 2011a). The predicted path travel times provided by the JTIS are instantaneous travel times. Figure 3.4 depicts an example of the real-time information supplied by the JTIS (Hong Kong Transport Department, 2023). It is seen in Figure 3.4 that the different traffic sensors have been deployed in JTIS to predict path travel times. They are regarded as ground truth in the case studies.



Figure 3.4 Illustration of the JTIS in Hong Kong

The numbers shown in the digital signs are journey times (or path travel times) in minutes from the locations of these signs to the exits of the corresponding road tunnels crossing Victoria Harbor in Hong Kong. The colors of the digits in the display panel represent the congestion levels of each route: red digits indicate congested traffic (<25 km/h), yellow digits imply slow traffic (25–50 km/h), and green digits reveal free-flowing traffic (>50 km/h).

As there are a limited number of AVI sensors (RFID tag readers) in the JTIS, and the average distance between these sensors is relatively long, the rates at which AVI data are sampled in the JTIS are very low. Accordingly, some point sensors are also deployed in the JTIS to provide additional data at selected locations along major paths in urban areas. These point sensors collect the point speed data of vehicles traveling along the major paths.

The combination of AVI and point sensor data enables the JTIS to generate updated predicted path travel time along major routes in Hong Kong urban areas once every 2

min (Tam and Lam, 2011a). As reported, independent floating car surveys have confirmed the validity of JTIS path travel times (Tam and Lam, 2008, 2011a). Hence, they are regarded as the ground truth for this study.

3.4.2 Set-ups in the case studies

Case studies on two selected paths in the Hong Kong urban road network are performed using real-world data. Figure 3.5 and Table 3.2 show the locations and characteristics of these two selected paths, respectively.

Study path 1 is 9.2 km long and connects the Island Eastern Corridor on Hong Kong Island to the Western Harbor Crossing in Kowloon; its free-flow path travel time is 8.4 min. Study path 2 is 8.8 km long and links Gascoigne Road and the entry of the Eastern Harbor Crossing; its free-flow travel time is 7.9 min. It is seen in Figure 3.5 that study path 2 has 3 AVI sensors, compared with study path 1 with 2 AVI sensors. An additional AVI sensor 4 is installed in the middle of study path 2 to collect more AVI data in order to enable the prediction of path travel times.



Figure 3.5 Overview of the two study paths in Hong Kong in Chapter 3

	Study path 1	Study path 2
Number of AVI sensors	Two	Three
Road type	Urban arterials with 21 bus stops and one signalized junction	Urban arterials with 19 bus stops only
Path length (km)	9.2	8.8
Number of bus stops	20	8
Number of entries	13	13
Number of exits	13	11
Number of short-spacing intersections	18	11
Free-flow travel time (min)	8.4	7.9
Speed limits (km/h)	70 (31%), 50 (18%), 60 (19%), 80 (32%)	70 (58%), 50 (20%), 80 (22%)
Number of point sensors	Seven	Five

Table 3.2 Summary of two study paths

These two study paths differ primarily in the number of AVI sensors and bus stops in Table 3.2. In addition, there is a signalized intersection on study path 1 but not on study path 2. The study paths contain several bus stops and frontage access with entries and exits. These site characteristics can lead to very few valid real-time AVI data available for path travel time prediction.

Figure 3.6(a) shows the low sampling rates of valid real-time AVI data from both paths. Based on descriptions of Dion and Rakha (2006), low sampling rates refer to as representatively two or three AVI data per 2-minute interval. However, in the case studies, as shown in Figure 3.6(a), there are only 12% and 30% of 2-minute intervals with no less than two valid real-time AVI data on study paths 1 and 2, respectively.







(b)

Figure 3.6 Sampling rates of (a) valid real-time AVI data, and (b) real-time AVI data on both study paths

It can be seen Figure 3.6(a) and Table 3.2 that the existence of signalized intersections and more bus stops on study path 1 further decreases the sampling rates of valid real-

time AVI data. Moreover, there are more than 50% of 2-minute intervals without any valid real-time AVI data or real-time AVI data from Figure 3.6(a) and Figure 3.6(b). The latter consists of both valid and invalid AVI data.

The AVI data and JTIS ground truth collected on all weekdays in 2017 and January 2018 are used in these two case studies. Public holidays and days with adverse weather and incidents are excluded. Hence, data of 299 days in 2017 are employed for training. Data from January 8th to 12th in 2018 are adopted for testing and evaluation for the rest of the case studies unless other specifications. The chosen rolling step is 2 min, and the confidence level for the dynamic validity window is 90%.

3.4.3 Results

The performance of the proposed unsupervised filtering algorithm is compared with that of the three corresponding existing algorithms that are commonly used in practice for filtering real-time AVI data. The algorithm developed by Southwest Research Institute (1998) is used to generate preliminary predicted path travel time for historical AVI data.

The proposed unsupervised algorithm is denoted as U1, and the other three corresponding existing algorithms are existing algorithms in practice for filtering realtime AVI data, which have been used successfully for decades in various ATIS (Southwest Research Institute, 1998; Dion and Rakha, 2006; Ma and Koutsopoulos, 2010). The algorithm of Dion and Rakha (2006) is denoted as U2, that of Ma and Koutsopoulos (2010) as U3, and that of Southwest Research Institute (1998) as U4. These existing algorithms use real-time AVI data, while U1 utilizes both real-time and historical AVI data.

The mean absolute error (MAE) and the mean absolute percentage error (MAPE),

which are given by Eqs. (3.33) and (3.34), respectively, are used to evaluate the filtering performance of algorithms with respect to the JTIS ground truth.

$$MAE = \frac{1}{T} \sum_{t=1}^{T} |T_t - \hat{T}_t|$$
(3.33)

$$MAPE = \frac{100}{T} \sum_{t=1}^{T} \frac{|T_t - \hat{T}_t|}{T_t}$$
(3.34)

where T_t are true values, and \hat{T}_t are predicted values.

Figure 3.7 illustrates the contribution of using historical AVI data for filtering out invalid real-time AVI data. It is observed in Figure 3.7 from the left black circle that when real-time AVI data is rather limited in the early morning, the filtering window of U1 is more appropriate than U2 as the variations of path travel times are not significant (traffic demand is low during early morning). Moreover, there is a chance that the transition between congestion and free-flow conditions can hardly be recognized properly by the existing filtering algorithms (e.g., U2, Dion's algorithm, which has already considered the transition recognition of the real-time traffic conditions by looking back at real-time AVI data in consecutive preceding time intervals).



Figure 3.7 Filtering performance of U1 and U2 on study path 2 by time of day

As shown in Figure 3.7, the right black circle indicates that U2 fails to select valid real-time AVI data. In contrast, U1 with the use of historical AVI data performs well in filtering the limited real-time AVI data. The day-to-day covariance of path travel times modeled in Eq. (3.4) can help to recognize traffic conditions by time of day. It is also observed that most of the relevant ground truth is captured within the dynamic validity windows resulting from U1 throughout the day.

Table 3.3 compares the filtering performance of the proposed unsupervised algorithm with benchmarks with respect to the mean/standard deviation of predicted path travel times. U1 outperforms the other three existing unsupervised algorithms from both aspects. For the mean of path travel times, the MAPE of U1 is 19.3% for study path 1

and 16.1% for study path 2. For the standard deviation of path travel times, the MAE values of U1 are 0.61 min and 0.52 min for study paths 1 and 2, respectively. The comparison of results between U1 and the other three existing unsupervised algorithms provides evidence to support the contribution of making use of historical AVI data.

	Study path 1			Study path 2				
	Mean		Standard		Mean		Standard	
Algorithms			deviation				deviation	
	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE
	(%)	(min)	(%)	(min)	(%)	(min)	(%)	(min)
U1 (proposed								
unsupervised	19.3	2.94	15.6	0.61	16.1	2.53	13.2	0.52
algorithm)								
U2 (Dion's	20.2	3 14	193	0.67	18.2	2 74	16.9	0.56
algorithm)	20.2	5.14	17.5	0.07	10.2	2.74	10.9	0.50
U3 (Median-	21.4	21.4 3.32	22.4 0.81	19.4	2.80	18.2	0.63	
based filter)				0.01	17.4	2.07	10.2	0.05
U4	22.1	3 49	27.1	0.94	20.1	3.01	20.4	0.68
(TransGuide)	<i>44</i> .1	J.TJ	2/.1	0.74	20.1	5.01	20.7	0.00

 Table 3.3 Comparison of prediction performance with respect to mean/standard

 deviation of predicted path travel times

3.4.4 Sensitivity analysis

In the real world, historical ground truth data on path travel times are assumed to be available (e.g., existing path travel times obtained from existing ATIS and samples collected independently from floating car surveys). It is a special case of research problems when historical ground truth is ready for training purposes. Under this scenario, $y_{i,d,p}^{s_g,k}$ denotes the *i*th ground truth on path travel time in Step 1 of Figure 3.2.

With the use of historical ground truth for training purposes, S1 represents the proposed unsupervised algorithm under this scenario. Three existing advanced supervised learning algorithms are selected for benchmark comparison. The LSTM NN in Ouyang et al. (2020) is denoted as S2. The LSTM encoder-decoder model in Wang et al. (2021b) is denoted as S3. The attention-based periodic-temporal neural network (APTN) in Shi et al. (2021) is denoted as S4. It should be noted historical ground truth in 2017 (299 weekdays) is used for training.

In contrast to other neural networks that have black-box procedures and nonexplanatory performance, the FPCA model explicitly describes the temporal covariances of path travel times by 2-minute intervals at different time intervals and on different days (Zhong et al., 2017). Moreover, as the FPCA model enables a better understanding of trends, it can be used to quantify the uncertainties of valid real-time AVI data with low sampling rates, particularly those data that are scattered and timevarying. The input data used in the sensitivity analysis include historical AVI data, historical JTIS ground truth, and real-time AVI data.

Table 3.4 gives the comparison results on the mean of predicted path travel times under this scenario. The MAPE of S1 is 11.4% for study path 1 and 5.1% for study path 2. It is found in Table 3.4 that S1 performs better than the other benchmarks. Besides, it should be noted that other benchmarks can only provide the mean of predicted travel times.

However, S1 can also produce the corresponding results and the standard deviation of predicted path travel times. In the case study, the corresponding MAPE and MAE are 12.7% and 0.5 min for study path 1 and 9.5% and 0.36 min for study path 2. These results can demonstrate the contribution of using FPCA to capture the temporal covariances of path travel times by 2-minute intervals at different time intervals and
on different days for data filtering and path travel time prediction.

	Study p	oath 1	Study p	bath 2
Algorithms	MAPE	MAE	MAPE	MAE
	(%)	(min)	(%)	(min)
S1 (proposed unsupervised				
algorithm with the use of historical	11.4	1.79	5.1	0.81
ground truth)				
S2 (LSTM NN)	15.3	2.38	7.5	0.95
S3 (encoder-decoder model)	14.1	2.21	6.6	0.88
S4 (periodic-temporal NN)	12.5	1.96	6.4	0.87

 Table 3.4 Comparison of prediction performance on the mean of predicted path travel

 times when historical ground truth is used for training

The average computational times of the proposed unsupervised and benchmarks without (U1-U4) and with (S1-S4) ground truth are provided. All case studies are conducted on a standard computer with an AMD Ryzen 5 5600X processor (3.7 GHz, six cores), as shown in Table 3.5. The average computational time for dynamic validity windows and predicted path travel time varies from 0.07 to 0.63 min. It is found in Table 3.5 that the U1 is applicable for online ATIS applications; that is, U1 can filter the real-time AVI data collected at about each 1.5-minute time interval and then rapidly (within 0.55 min) generate the predicted path travel times.

Algorithms		The average computational time for predicted travel		
		time and filtering windows for each time interval		
		(min)		
		Study path 1	Study path 2	
	U1	0.52	0.51	
Unsupervised	U2	0.13	0.2	
algorithms	U3	0.07	0.17	
	U4	0.1	0.15	
	S1	0.55	0.52	
Supervised	S2	0.63	0.57	
algorithms	S3	0.5	0.6	
	S4	0.52	0.55	

Table 3.5 The average computational time of different filtering algorithms

Table 3.3 gives the overall errors (in terms of MAPE and MAE) of predicted path travel times by the proposed unsupervised algorithm. However, the detailed prediction errors should be further elaborated as the sample sizes of real-time AVI data and historical AVI data can be varied under different traffic conditions. Therefore, an additional experiment has been conducted on study path 2 to show the performance on predicted path travel times from unsupervised algorithm under different traffic conditions.

The set for levels of service for describing the traffic condition on path p is denoted as LOS_p . The threshold of average speed for the *i*-th level of service $LOS_{i,p}$ on path p is $v_{LOS_{i,p}}$, while the free-flow travel speed on path p is $v_{f,p}$, the corresponding ratio for $LOS_{i,p}$ on path p is:

$$\delta_{LOS_{i},p} = \frac{v_{f,p}}{v_{LOS_{i,p}}} \tag{3.35}$$

Table 3.6 presents the findings of this experiment and shows the sample sizes of both historical and real-time AVI data. The sample sizes of both historical and real-time AVI data are smallest under LOS A&B (0.6 and 1.1 for real-time and historical AVI data, respectively). As a result, the employment of historical AVI data yields the most significant enhancement in prediction accuracy, with a maximum reduction in MAPE of 7.9%. Moreover, the MAPE is reduced to 18.4%, which meets the requirement proposed by Tam and Lam (2008). It implies that though the sample size of historical data is also insufficient when traffic is light (i.e., LOS A&B), their influence on the accuracy of predicted path travel times can be significant.

However, when levels of service are C&D, the prediction accuracy with and without the incorporation of historical AVI data in the offline training stage is nearly identical, with a minimal MAPE difference of 0.7% in Table 3.6. This is because the sample size of real-time AVI data is relatively sufficient (2.9). These observations underscore the critical role of historical AVI data in enhancing performance on predicted path travel times, particularly during uncongested traffic conditions (i.e., LOS A&B).

Traffic condition	Samp AVI d minut	le size of lata per 2- e interval	With his AVI d the of training	storical ata in fline g stage	With historic data in offline t sta	nout al AVI n the raining ge	Perforr improv aft consid historica dat	nance ement er ering al AVI ta
	Real-							
	time	Historical	MAPE	MAE	MAPE	MAE	MAPE	MAE
	AVI	AVI data	(%)	(min)	(%)	(min)	(%)	(min)
	data							
LOS	0.6	11	18.4	2 76	263	3 38	7.9	0.62
A&B	0.0	1.1	10.1	1.70	20:5	5.50	1.	0.02
LOS	2.0	3 7	14.5	2 35	15.2	2 12	0.7	0.07
C&D	2.9	5.2	14.5	2.35	13.2	2.42	0.7	0.07
LOS	1 /	17	15.2	2.44	10.0	2.01	2.6	0.27
E&F	1.4	1./	13.3	2.44	18.9	2.81	3.0	0.37
Overall	1.2	1.4	16.1	2.53	20.1	3.01	4.0	0.48

Table 3.6 The prediction performance of the proposed unsupervised algorithm onstudy path 2 under different traffic conditions

Therefore, further sensitivity tests with pre-determined and fixed sample sizes are carried out to investigate the effects of traffic conditions on the performance of predicted path travel times. As traffic data under LOS A&B have the smallest sample sizes, the sample sizes of real-time AVI data for path travel time prediction under other traffic conditions are also reduced to 0.6 per 2-minute interval. Moreover, the comparison excludes the involvement of historical AVI data to eliminate their complicated impact. The results of these further sensitivity tests are shown in Table 3.7.

Traffic condition	MAPE (%)	MAE (min)
LOS A&B	26.3	3.38
LOS C&D	36.4	4.01
LOS E&F	30.1	3.86

 Table 3.7 The prediction performance of the proposed unsupervised algorithm with a fixed sample size of real-time AVI data

It is found that the prediction performance under LOS A&B (which is the worst in Table 3.6) is the best. This is due to the relatively low variability of path travel times under uncongested LOS A&B conditions. However, the prediction performance under all traffic conditions ranging from LOS A to LOS F is unsatisfactory, with MAPE being over 20% if the real-time AVI data is only 0.6 sample per 2-minute interval. This finding highlights the importance of considering historical data in the offline training stage when the real-time AVI data is insufficient.

Furthermore, as both study paths in Table 3.2 have different numbers of short-spacing intersections and frontage access, their impact on the travel time calculation process can be evaluated. AVI data of these study paths with similar weather and traffic conditions are selected. The widths of their validity windows are to be compared. However, the difference in path distance may have a significant impact on their results. Therefore, the metric is the ratio between the number of short-spacing intersections and path distance, which compares two study paths. Similarly, the ratio between the number of frontage accesses (entries and exits) and path distance is calculated. The results are given in Table 3.8 below.

Study paths	Number of short-spacing intersections per kilometer	Frequent frontage access per kilometer	MAPE of the predicted mean of path travel times (%)	Average width of validity window (min)
1	2.0	2.8	19.3	5.1
2	1.3	2.7	16.1	4.6

 Table 3.8 The effects of short-spacing intersections and frontage access on predicted

 path travel times

It is found in Table 3.8 that study path 1 has more short-spacing intersections per kilometer than study path 2. Consequently, study path 1 has a lower prediction accuracy and wider validity window than study path 2. The risk of identifying wrong valid AVI data may be higher for paths with many short-spacing intersections (Ban et al., 2010; Van Hinsbergen et al., 2012; Elfar et al., 2018). However, further study should be carried out on various study paths with different road types in urban areas to investigate the effects of frequent frontage access on path travel time prediction.

Another sensitivity analysis is conducted to examine the effect of sampling rates of real-time AVI data on the performance of the proposed unsupervised algorithm. As study path 2 has more AVI sensors than study path 1, the range of sampling rate for study path 1 is greater. Moreover, it can be used to investigate the problem of filtering AVI data under different sensor failure scenarios for further study. Therefore, study path 2 is used for sensitivity analysis to examine the effects of sampling rates. The results are shown in Figure 3.8.



Percentage of time intervals with no less than 2 real -time AVI data

Figure 3.8 Sensitivity test with various sampling rates of real-time AVI data on study

path 2

There are two observations in Figure 3.8, Firstly, U1 differs from U2-U4 in that it incorporates the day-to-day covariance of path travel times by 2-minute intervals $\Sigma_t^{s_A}(d_i, d_j)$. U1 enables the identification of traffic condition patterns, such as regular periods of congestion at certain times of the day or on specific days from the historical AVI data. The consideration of $\Sigma_t^{s_A}(d_i, d_j)$ can improve the prediction accuracy of 3-39

path travel times.

Secondly, while U2, U3, and U4 rely only on real-time AVI data as their input, U1 makes full use of both the real-time and historical AVI data. When real-time AVI data is insufficient to construct the dynamic validity window (e.g., particularly when there is either 0 or 1 sample per 2-minute interval), historical AVI data can contribute significantly to U1 for improving the prediction accuracy of the path travel times.

To further investigate the effects of sampling rates of valid real-time AVI data on the proposed unsupervised algorithm, another sensitivity test is carried out. 30 out of 299 weekdays in 2017 are randomly segregated from the original training set and used as the new validation set. The performance of the proposed unsupervised algorithm on both study paths is provided in Figure 3.9.



Validation set: random 30 weekdays in 2017

Figure 3.9 Sensitivity test of sampling rates of valid real-time AVI data on U1 for study path 1 (left) and study path 2 (right)

It is noted in Figure 3.9 that when the sampling rate of valid real-time AVI data is no less than two valid AVI data per 2-minute interval, the performance of U1 is similar

on different datasets (95% of the absolute percentage errors less than 15.2% and 14.9% for study paths 1 and 2, respectively). It demonstrates the generality and robustness of the proposed unsupervised algorithm. In general, it is found in Figure 3.9 that 95% of the absolute percentage errors of the predicted results are less than 20%.

Additionally, as shown in Figure 3.6(a), study path 2 with more AVI sensors would have a higher percentage of 2-minute intervals with no less than two valid data than that of study path 1. Hence, the performance of both S1 and U1 for study path 2 is better than that of study path 1, as shown in Table 3.3 and Table 3.4. The same finding can also be obtained in Figure 3.9, even though the validation dataset is different in Figure 3.9, Table 3.3, and Table 3.4.

As the case study is carried out using accurate but limited real-time AVI data, it is worthwhile to discuss the performance of U1 on inaccurate real-time AVI data with more samples (e.g., Bluetooth data). It is assumed that this type of AVI data has a much lower percentage of valid real-time AVI data. Therefore, the performance will deteriorate due to the extremely low sampling rate of valid real-time AVI data for U1. Further study should be conducted in the future if this type of AVI data is available.

A sensitivity test is performed to assess the effect of historical ground truth data by reducing the number of days with historical ground truth data. A percentage varying from 0% to 100% of historical ground truth data is available for training to evaluate the performance of S1. The result is displayed in Table 3.9. When there is no ground truth available for training, U1 has better performance (83%, as indicated in Figure 3.8) than S1 (71%). Furthermore, the percentage of absolute percentage errors less than 20% is reduced to 83% or lower if less than 50% of the historical ground truth is used for training purposes. This implies that U1 is better than S1 in practice, particularly when less than half of the historical ground truth on path travel time is available to

filter real-time AVI data and path travel time prediction.

Table 3.9 Sensitivity test with different percentages of historical ground truthremoved in 2017 data on study path 2

Percentage of historical ground truth	Percentage of absolute percentage
available in 2017 data (%)	errors less than 20% (%)
0	71
10	76
30	78
50	83
70	86
90	90
100	93

With reference to the above Eq. (3.8), the sensitivity test of threshold Σ^* is conducted. The relevant results are given in Table 3.10, in which the optimum thresholds of Σ^* for study paths 1 and 2 are 10.8 and 9.6 min², respectively.

Table 3.10 Sensitivity test of the percentage deviation of the results of S1 from the

optimum threshold

		Percenta	age of deviati	on from
		optin	num threshold	d (%)
		0	10	20
Value of threshold	Study path 1	10.8	11.9	13
(min^2)	Study path 2	9.6	10.6	11.5
MAPE (%)	Study path 1	11.4	12.6	13.4
	Study path 2	5.1	7.8	9.5
Percentage of absolute	Study path 1	83	78	74
percentage errors less	2000) point 1		,	
than 20% (%)	Study path 2	93	86	82

For study path 1, only 74% of absolute percentage errors are less than 20% when there is a 20% deviation from the optimum threshold. Σ^* can also be an annual average figure, as it is based on weekday data in 2017 (excluding public holidays and days with adverse weather and incidents) to capture the seasonal variation of path travel times. Moreover, the optimum threshold is based on the current dataset, but Σ^* may deviate from the actual optimum threshold, as the latter will depend on the updated dataset.

Previous contents have validated the mean of path travel times by 2-minute intervals. Figure 3.10 compares the day-to-day variance of the mean of predicted path travel times obtained from the proposed model using AVI data $(Var(\mu_{T^{s_A}}(d)), d \in D, which$ is the variance of the predicted mean of path travel times by 2-minute intervals on day d from data source s_A) against day-to-day variance of ground truth on mean of path travel times $(Var(\mu_{T^{s_G}}(d)), d \in D, which is the variance of ground truth on mean of$ path travel times by 2-minute interval on day <math>d from data source s_G) on the study path 1.

In general, it is found in Figure 3.10 that the latter one is smaller than the former one. The 90 percentiles of day-to-day variance of mean of path travel times by 2-minute intervals are 28.1*min*² and 37 *min*² for predicted sample mean and ground truth of population mean, respectively. This is because the former regards variations of the population mean, while the latter refers to variations of the sample mean based on AVI data. The variations of sample mean should include variations of population mean, as AVI sensors only collect timestamps of commercial vehicles. Detailed validation requires the collection of ground truth of this covariance for further study.



Figure 3.10 Comparison of day-to-day variance of the mean of predicted path travel times from the proposed model using AVI data $Var(\mu_{T^{S}A}(d))$ against day-to-day variance of ground truth on mean of path travel times $Var(\mu_{T^{S}G}(d))$ by 2-minute intervals on the study path 1

Different combinations of Δt , Δ , and δ are implemented in a sensitivity test to investigate their impacts on the accuracy of predicted path travel times. The values of Δ are 2 min, 10 min, 30 min, and 60 min, while the values of Δt and δ are 10 min, 30 min, and 60 min. There is a constraint regarding real-time implementation of prediction models, i.e., $\Delta t \geq \Delta$. For each combination, these three factors are fixed during the prediction process. The tests are conducted on study path 2. Table 3.11 gives the corresponding results.

Rolling step	Cotocom	Rolling horizon	Prediction	
Δ (min)	Calegory	δ (min)	horizon Δt (min)	$\mathbf{MAPE}(70)$
		30	30	17.2
	$\delta = \varDelta t$	60	60	18.4
		10	10	21.5
		30	60	17.9
2	$\delta < \Delta t$	10	60	20.1
		10	30	20.8
		60	30	17.6
	$\delta > \Delta t$	60	10	18.9
		30	10	19.2
		30	30	26.8
	$\delta = \varDelta t$	60	60	28.7
		10	10	33.5
		30	60	27.9
10	$\delta < \Delta t$	10	60	31.4
		10	30	32.4
		60	30	27.5
	$\delta > \Delta t$	60	10	29.5
		30	10	30.0
	$\delta = \Lambda t$	30	30	34.9
	0 – 11	60	60	37.3
30		30	60	36.3
50	$\delta < \Delta t$	10	60	40.8
		10	30	42.2
	$\delta > \Delta t$	60	30	35.7
	$\delta = \Delta t$	60	60	48.5
60	δ < Λt	30	60	48.2
	$0 \leq \Delta t$	60	60	53.1

Table 3.11 Impacts of prediction horizon (Δt), rolling horizon (δ), and rolling step (Δ) on prediction accuracy (in terms of MAPE%) for the study path 2

It is seen in Table 3.11 that the best accuracy of predicted path travel times is achieved when Δt , δ , and Δ are 30 min, 30 min, and 2 min (MAPE=17.2%). On the contrary, the largest MAPE appears when all three factors are 60 min (MAPE=53.1%). It is also found that MAPEs are 26.8%, 34.9%, and 48.5% when values of Δ are 10 min, 30, and 60 min, respectively. It implies that the detrimental effect on the prediction accuracy is greater for longer Δ . Moreover, it is observed in Table 3.11 that the accuracy is highest under the category when $\delta = \Delta t$. It suggests that the values of δ and Δt should be the same for the prediction of path travel times in further study.

Further analysis is carried out to investigate the effects of missing data on filtering performance. Five scenarios are generated by extracting varying proportions of online and historical AVI data. They are (i) 50% of data missing in the collected online AVI data at the certain time interval; (ii) all data missing at the certain time interval; (iii) 50% of data missing in the collected historical AVI data at the certain time interval; (iv) all data missing in the collected historical AVI data at the certain time interval; (v) full data (without deduction of sample size). 30 2-minute intervals from congested and uncongested periods are selected for investigation of the effects of missing data on filtering performance, respectively. The prediction performance of path travel times under these five scenarios on study path 1 is summarized in Table 3.12.

It is found in Table 3.12 that Scenario (i) has the least adverse impact on prediction accuracy (MAPE=12.4%) compared with the full data scenario (v) (MAPE=10.8%). Furthermore, as the original dataset has limited online AVI data, the missing data issue of online AVI data has less impact on prediction accuracy than that of historical AVI data. Moreover, when all historical data are missing in Scenario (iv), the resulting MAPE is 22.4%, which is unsatisfactory and motivates the need to collect more ground truth data for training purposes.

Scenarios	MAPE (%)	MAE (min)	RMSE (min)
(i)	12.4	1.9	2.1
(ii)	14.3	2.3	2.6
(iii)	17.7	2.7	2.9
(iv)	22.4	3.5	3.7
(v)	10.8	1.5	1.7

Table 3.12 Prediction performance of the proposed unsupervised algorithm (in terms of MAPE (%)/MAE (min))

(i) 50% of data missing in the collected online AVI data at a certain time interval; (ii) all data missing at a certain time interval; (iii) 50% of data missing in the collected historical AVI data at a certain time interval; (iv) all data missing in the collected historical AVI data at the certain time interval; and (v) full data.

Traffic accidents can increase path travel times significantly. To investigate the impacts of traffic accidents on the filtering performance for path travel time prediction, AVI data on days with accidents in 2018 are further analyzed. It is found that the increment of path travel times varies from 3% to 15% after filtering. Regarding outliers filtered out by the proposed unsupervised algorithm, they may be affected by accidents or detours of vehicles. These outliers need to be further distinguished with detailed trajectories of accident vehicles in further study.

3.5 Concluding Remarks

This chapter proposes a novel unsupervised algorithm (U1) for filtering limited but accurate real-time AVI data without ground truth for training. Instead, it makes use of real-time AVI data gathered on the current day and historical AVI data collected on previous days. The temporal covariances of path travel times by 2-minute intervals at different time intervals and on different days are explicit. Both mean and standard

deviation of the predicted path travel times are provided and used for constructing the validity window for filtering real-time AVI data.

As FPCA can effectively reduce the dimension of high-variability data, the corresponding PACE approach is used to construct the dynamic validity windows in the proposed unsupervised algorithm. The real-time dynamic validity windows are generated via a rolling horizon scheme. Furthermore, the asymptotic properties of U1 have been theoretically proven to confirm their ability to generate reliable dynamic validity windows for real-time AVI data filtering and path travel time prediction.

The performance of U1 is compared respectively with three existing data filtering algorithms in the case studies using real-world data collected from two selected paths in the Hong Kong urban road network. The comparison between U1 and U2 on filtering performance by time of day demonstrates the merit of using historical AVI data. It is also found that U1 surpasses the existing algorithms in terms of both mean and standard deviation of predicted path travel times.

A sensitivity analysis is conducted for a special case where ground truth is available for training. The proposed unsupervised algorithm with ground truth for training (namely S1) outperforms other benchmarks in terms of both mean and standard deviation of predicted path travel times. It illustrates the merit of using FPCA for modeling temporal covariances of path travel times by 2-minute intervals at different time intervals and on different days. Another sensitivity test is also performed to reveal the merits of using historical AVI data when real-time AVI data is sampled at a very low rate. U1 performs much better than the other three benchmarks in terms of the probabilities of absolute percentage errors of the predicted results, which are less than 20% (i.e., 83% against 56%-60%). The sensitivity test on sampling rates of valid real-time AVI data demonstrates the generality and robustness of the proposed unsupervised algorithm. When there are no less than two valid real-time AVI data per 2-minute interval, there is a 95% probability of generating absolute percentage errors of less than 20%. Moreover, the performance of U1 on different study paths is similar under this scenario, which implies that U1 is generalized with robust performance. The expected worsened performance on inaccurate real-time AVI data with more samples (e.g., Bluetooth) is also discussed with the assumed lower sampling rate of valid real-time AVI data. Filtering of this type of AVI data is suggested for further study if this dataset is available.

Moreover, an additional sensitivity test is carried out to show the advantage of U1. The percentage of absolute percentage errors less than 20% is reduced to 83% or lower if less than 50% of the historical ground truth is used for training purposes. It implies that U1 is better than S1 in reality when less than half of the historical ground truth on path travel time is available for filtering real-time AVI data and path travel time prediction.

In this chapter, only a single traffic data source (i.e., AVI data) has been considered for path travel time prediction. However, there can be several traffic sensors allocated on the road networks in practice with various travel time data, as presented in Figure 2.1. Moreover, vehicle type information is available for AVI data, but this chapter has not addressed this issue, as mentioned in Section 3.2. It is worthwhile to make full use of AVI data, together with other traffic data sources, for multi-class path travel time prediction in the following Chapter 4.

Chapter 3 mainly focuses on limited AVI data with high accuracy for capturing vehicular travel times (e.g., RFID and ALPR data). For AVI data with less accuracy but much more sample sizes (e.g., Bluetooth data), the filtering algorithms can be

further studied in the future.

In addition, U1 could benefit from the integration of various traffic-related data types to enhance its data filtering capabilities. These data types include weather conditions, traffic accidents, and the schedule of construction works. Furthermore, vehicular flow data, bus frequencies, signal timing, short spacing between intersections, the number of frequent frontage access points (i.e., entries and exits), and road types could also be incorporated into U1.

The weather information (e.g., historical rainfall intensity and weather forecast) has been investigated in Chapter 5 to improve the prediction performance of path travel times. It is interesting to study whether it will affect filtering performance. As both AVI sensors and point sensors are deployed in the JTIS, it is interesting to explore the sensor-location problems and trade-offs of these two categories of traffic sensors. Similarly, it is also interesting to extend U1 to examine the effects of sensor failure on data from multiple AVI sensors at urban road corridors by considering network topology and measurement errors.

As mentioned in Section 2.4, after filtering outliers/invalid data, path travel time prediction can be carried out using valid AVI data. In the following Chapter 4, the output of this chapter (i.e., the filtered real-time AVI data) will be the input for path travel time prediction by vehicle class. The vehicle type information is available from AVI data, as illustrated in Section 2.2. This information will be critical for the following chapter, as Chapter 4 aims to predict path travel times by different vehicle classs. The evidence of distinct distributions of path travel times by vehicle class relies on the availability of AVI data after the filtering process. Furthermore, the filtered AVI data will also be used for path travel time prediction in Chapter 5.

4. Prediction of Multi-class Path Travel Times Using Multisource Traffic Data

4.1 General

After data filtering described in Chapter 3, the filtered traffic data can be used to predict path travel times for ATIS. In practice, most of the ATIS provide average travel times (i.e., mean of path travel times) of all vehicles on selected paths in real time on a regular basis. These predicted path travel times have been validated with ground truth in the literature, as shown in Section 2.5. However, the path travel times of different vehicles could vary widely under different traffic conditions. There is a need to consider the differences in vehicle classes for path travel time prediction.

This chapter proposes a novel prediction model that models temporal covariances of path travel times between vehicle classes by 2-minute intervals for predicting multiclass path travel times. It uses multi-source traffic data collected from various types of sensors. The proposed prediction model is examined with numerical experiments of a selected urban expressway in Hong Kong with data obtained from multiple sources. The predicted path travel times by vehicle class in the experiments demonstrate the merits and performance of the proposed prediction model.

The rest of this chapter is organized as follows. Section 4.1 starts with the basic information and the motivation, together with the constructive contributions of this chapter. The research problems are presented in Section 4.2. Section 4.3 provides the details of the proposed prediction model for multi-class path travel time prediction. The results of numerical experiments are shown in Section 4.4. Finally, the findings of the research in this chapter are given in Section 4.5.

4.1.1 Basic information

The traffic conditions on a path, which can be defined as an alternating sequence of nodes and links connecting an origin and destination pair, as mentioned in Section 1.1, have received much attention in the field of transportation research. The travel times spent by vehicles on their designated path are referred to as path travel times in this thesis. Path travel time information can assist road users in route selection. The information can also help road authorities manage traffic conditions on roads, especially urban roads. However, the path travel times change over time owing to variations in traffic demand and supply on the roads (Shao et al., 2013; Han et al., 2018). It is thus necessary to accurately predict the path travel times using traffic information gathered on road networks.

As exhibited in Figure 4.1, a selected expressway is used to illustrate different sensors along the study. It is a major route from Tuen Mun New Town to Tsuen Wan New Town. Table 4.1 gives the detailed traffic characteristics of the study path. The chosen path is **17.8 km** long and has a free-flow path travel time of **14.3 min**. It can be observed that the study path is a representative route in Hong Kong, with many bus stops, entries, and exits along the path. The speed limits vary between 70 km/h and 80 km/h on the study path.

ALPR technology introduced in Section 2.2.1 is adopted for the collection of traffic data on the study path. Compared with the RFID technology used in Chapter 3, ALPR can assemble more AVI data. Video-based cameras as point sensors are installed along the path. Figure 4.1 presents an example of data obtained for one vehicle. If a truck is identified by both AVI and GPS sensors, the trajectories are constructed using the corresponding AVI and GPS data.



Figure 4.1 Overview of the study path and example of data for one vehicle in Chapter 4

There are four remarks worth making in Figure 4.1. First, the identification numbers in the AVI system and GPS are non-identical for the same vehicle. Second, the trajectory of the truck from GPS data required extrapolation as there are not necessarily GPS data corresponding exactly to both the origin and destination. Third, there are two constructed trajectories for the same vehicle with contrasting observed path travel times. They are 19.2 min and 18.8 min from AVI and GPS data, respectively. This indicates that AVI and GPS data cannot be used simultaneously owing to the double-counting problem. Fourth, point sensor data are not shown in Figure 4.1 as they are aggregated at 2-minute intervals, as presented in Figure 2.1. The individual speeds of a truck captured by point sensors are not stored in the point sensor system.

Road type	Expressway
Path length (km)	17.8
Number of bus stops	20
Number of entries along the study path (e.g., slip road and frontage access)	17
Number of exits along the study path (e.g., slip road and frontage access)	19
Free-flow travel time (min)	14.3
Speed limits (km/h)	70 (47%), 80 (53%)

Table 4.1 Traffic characteristics of the study path

Given the features of different traffic sensors introduced in Section 2.3.4, there are three challenges in dealing with multi-source data in this study. First, there is limited data for valid AVI and GPS data for path travel time prediction. Figure 4.2 illustrates the cumulative distribution function (CDF) plots of the sample sizes of the valid AVI and GPS data after data filtering from the study path. It is noted that GPS data are generally processed on several short links so that the allocation of data can be more accurate (Zhong et al., 2017). There are 67 links along the study path. Therefore, the sample size of GPS data is counted per link.



(a) CDF of the sample size of valid AVI data



(b) CDF of the sample size of valid GPS data



It can be seen in Figure 4.2(a) that over 55% of the 2-minute intervals had **no** valid AVI data, and 80% of the 2-minute intervals had no more than two samples of all vehicles. Figure 4.2(b) shows that 64% of 2-minute intervals had no more than two samples per link (67 links for the study path). It was reported in the literature (Dion and Rakha, 2006) that two or three observations per 2-minute interval were considered a low sampling rate with limited valid AVI or GPS data for path travel time prediction.

Second, there are difficulties in predicting path travel times by vehicle class. The sample size is rather limited for each vehicle class. For vehicle classes with extremely low sample sizes, there is a need to combine them into a new class for model training. Moreover, from Table 2.7 in Section 2.3.4, it is noted that not all traffic sensors (such as point sensors in this study) cover vehicle class information of all vehicle classes. Third, due to the limited data and incomplete vehicle class information, it is difficult to consider the covariance of path travel times between vehicle classes.

Apart from variations of path travel times over time, as shown in Figure 3.7 and explained in Section 3.4.3, the path travel times can be distinct by vehicle class. Figure 4.3 gives the CDF plots of observed travel times on the study path by vehicle class from AVI data. In the dataset used for analysis, the percentages for private cars, goods vehicles, and other vehicles (such as buses and coaches) are 24.8%, 56.1%, and 19.1%, respectively.



Figure 4.3 CDF plot of path travel times by vehicle class using AVI data

It is noted in Figure 4.3 that 50% of path travel times are less than 13.8 min, 15.1 min, 16.6 min and 15.0 min for private cars, goods vehicles, other vehicles, and all vehicles respectively. Thus, over 50% (56.1%) of drivers (i.e., drivers of goods vehicles) receive an underestimated path travel times with 10.7% deviation from observed path travel times. The majority of drivers hence should receive more accurate predicted path travel time.

Moreover, private cars are comparatively faster than good vehicles and much faster than other vehicles as over 95% of other vehicles are regular buses which need to stop at stops along the study path. Moreover, the standard deviation of path travel times by vehicle class should be different inferred from Figure 4.3. As this chapter mainly studies the mean of path travel times by vehicle class, the deviation of multi-class path travel times should be further investigated in the future. Table 4.2 presents the p-value of the Kolmogorov–Smirnov test at the 5% significance level. As all p-values between vehicle classes (with the percentage of samples) are much less than 0.05, the path travel time distributions of these vehicle classes are significantly different. Hence, the empirical evidence supports that there is a need to predict path travel times by vehicle class for ATIS development.

Table 4.2 P-values of the Kolmogorov-Smirnov test on path travel times by vehicle

Vehicle	Vehicle classes				
classes	Private cars (24.8%)	Goods vehicles (56.1%)	Other vehicles (19.1%)	All vehicles (100%)	
Private cars	1	7.2×10 ⁻¹⁷	1.4×10 ⁻¹⁹⁹	1.72×10 ⁻³³	
Goods vehicles	7.2×10 ⁻¹⁷	1	2.8×10 ⁻¹⁰⁹	2.3×10 ⁻¹⁵	
Other vehicles	1.4×10 ⁻¹⁹⁹	2.8×10 ⁻¹⁰⁹	1	3.9×10 ⁻¹⁸⁴	
All vehicles	1.72×10 ⁻³³	2.3×10 ⁻¹⁵	3.9×10 ⁻¹⁸⁴	1	

class ($\alpha = 0.05$)

However, the path travel time information is usually provided for all vehicles as a single class (average path travel times for all vehicles) instead of a multi-class (multiple vehicle classes). Previous studies predicted the path travel times of all vehicles using data from either AVI sensors or point sensors, as presented in Sections 2.5.1 and 2.5.3. The sample size of a single traffic data source may only afford to predict path travel time for a single class. The relatively low sampling rates of the single source of traffic data (as shown in Figure 4.2) motivated the use of multi-source

data, especially for multi-class travel time prediction. Furthermore, although the point sensor dataset had no information on the vehicle class, it enriches the data amount which would help to improve the prediction accuracy.

Besides, current path travel time prediction models based on AVI data have only provided the average path travel times of all vehicles. As AVI sensors are more expensive than point sensors, AVI sensors are spaced at greater intervals than point sensors. Hence, for each 2-minute interval, fewer than three samples are usually collected from an AVI system. After removing outliers using outlier detection algorithms (Chen et al., 2010), the size of valid samples is smaller. Therefore, the combined application of AVI, GPS, and point sensor data is attractive for multi-class path travel time prediction. Either AVI or GPS data, together with point sensor data, are used to enrich the information on traffic conditions on an urban road for this problem.

Hence, there are three major challenges to be tackled. First, considering the different features of the three types of traffic data, it is a challenge to evaluate the appropriate usage of each data source for predicting path travel times. The simultaneous use of AVI and GPS data may cause a double-counting problem. It can lead to an inaccurate path travel time distribution and should be avoided.

Second, as the availability of vehicle class information varies, it is a challenging task to predict path travel times by vehicle class using a unified modeling framework, instead of providing the average path travel times for all vehicles. Third, it is challenging to model various types of temporal covariance of path travel times (especially the temporal covariance of path travel times between vehicle classes) to improve the prediction accuracy of the modeling framework. The Gaussian mixture model (GMM) is a powerful tool that has been widely used in analyzing the temporal covariance of path travel times. Recent studies have adopted the GMM for travel time prediction using one or multiple sources of traffic data, as shown in Table 4.3. These studies clustered data by traffic conditions without considering vehicle classes due to the limited availability of vehicle class information. The modified GMM presented in this chapter extends to taking into account the temporal covariance of path travel times between vehicle classes to enhance the prediction accuracy of path travel times by vehicle class.

Related studies	Data	Categories (clusters)
Yang et al. (2018)	AVI data (Bluetooth data)	Free-flow, saturated, or oversaturated
Ramezani and Geroliminis (2012)	Probe vehicle data	No more than three clusters
Wang et al. (2021b)	AVI data (ALPR data)	Two or three components decided by the data
Mil and Piantanakulchai (2018)	Loop detector, AVI, and GPS data	Free-flow; congestion; transition
This chapter	AVI, GPS, and point sensor data	Clustering using Bayesian Information Criterion (BIC) by vehicle class

Table 4.3 Recent related studies using the GMM

Other machine-learning models have also frequently been applied to transportation problems. Two of them are introduced as benchmarks and compared against GMM in numerical experiments. The first is the LSTM model, which is one of the hot technologies in deep learning. It belongs to recurrent neural networks (Du et al., 2022a). The special design of the memory block solves the problem of long-term dependence compared to traditional recurrent neural networks (Du et al., 2022b). The LSTM is commonly used in traffic prediction problems (Ku et al., 2021; Yang et al., 2021b; Ouyang et al., 2020; Zhao et al., 2019).

The second is the attention-based periodic-temporal neural network (APTN), which is a branch of encoder-decoder networks. The attention mechanism is proposed to solve the problem of unsatisfactory performance of traditional encoder-decoder networks when the input sequence is too long. The temporal attention selects the most relevant input features while the spatial attention correlates the local node to the entire graph (Shi et al., 2021).

4.1.2 Constructive contributions of this chapter

This chapter studies multi-class path travel time prediction with the use of either AVI or GPS data, together with point sensor data. The proposed prediction model in this chapter extends existing works by providing the following constructive contributions.

C4.1 The proposed prediction model can predict the path travel times of different vehicle classes using the appropriate combination of traffic sensor data with satisfactory performance.

C4.2 The proposed prediction model improves prediction accuracy by studying various types of temporal covariances of path travel times by 2-minute intervals, especially the temporal covariance of path travel times between vehicle classes by 2-minute intervals.

C4.3 The robustness of the proposed prediction model is tested independently in a real-

world numerical experiment and shown to be satisfactory.

4.2 Research Problems

Consider a path p with a pair of AVI sensors installed at both ends. The travel time of road corridors between two sensors o_p and d_p can be available from a set of multi-source traffic data S, which consists of AVI, GPS, and point sensor data (defined as multi-sources s_A , s_B , and s_W , respectively). The observed path travel times can be extracted from AVI data directly, whereas it is necessary to preprocess the GPS and point sensor data to obtain path travel times indirectly.

Vehicle class information can be available from the AVI and GPS data. Vehicle class information for all vehicles is available in AVI data. K_{s_A} denotes the set of vehicle classes that can be available from data source s_A (AVI sensors). The GPS sensors only cover the partial vehicle class information of commercial vehicles; i.e., mainly goods vehicles. The set of vehicle classes available from data source s_B (GPS sensors) is denoted K_{s_B} ($K_{s_B} \subseteq K_{s_A}$). Point sensors have no information on vehicle class but with traffic flow and speed data for all vehicles.

The available data are $y_{i,d,p}^{s_A,k}$, which is *i*-th observed path travel time for path *p* of vehicle class *k* for from data source s_A on day *d*. It is consistent with the observed path travel time denoted in Section 3.2. Moreover, the observed path travel time from point sensors and GPS sensors rely on the average spot speed at location *x* in time interval $t (v_{x,t})$, and trajectory of vehicle *i* of vehicle class $k (z_i^k (t + j\Delta g))$, where Δg is the sampling time interval of the data source s_B (GPS) and j = 1, ..., J is the index of sampling points for an individual vehicle trajectory within path *p*, with the corresponding location $x_{i,j}^k$, speed measurement $v_{i,j}^k$, and timestamp $\tau_{i,j}^k$. This chapter

predicts the path travel times for different vehicle classes with the proper usage of the three types of traffic sensor data.

4.3 Prediction of Multi-class Path Travel Times

The multi-source data used in this study are presented as follows. Three types of traffic sensor data are used: AVI (using ALPR technology), GPS, and point sensor data. For privacy reasons, path travel time with vehicle class information is automatically saved in an AVI dataset without vehicle identity information. GPS data only covers a limited number of vehicle classes. In the setting, only AVI data are collected and updated.

Based on these traffic sensor data, there are three reasons for classifying vehicles in the methodology. First, there is a need to distinguish predicted path travel time by vehicle class as they distribute significantly distinctly, as shown in Figure 4.3. Second, the vehicle classification design can help to improve the prediction accuracy of path travel time, which will be introduced later. Third, the proposed prediction model can be generalized to predict path travel times by vehicle classes, while the existing approach for predicting the average travel times for all vehicles on the selected path can be viewed as a special case of this study.

Based on the three types of data sources considered in this study, the modeling framework for multi-class travel time prediction on the selected path consists of an offline prediction stage and an online prediction stage to satisfy the needs of various ATIS. Figure 4.4 presents the framework of the proposed prediction model for the prediction of multi-class path travel times. In the offline prediction stage, the predicted path travel times and the various types of temporal covariance of path travel times (in matrix format) by vehicle class are obtained using the modified GMM (which is shown later).



Figure 4.4 The framework of the proposed prediction model

There are three types of temporal covariance of path travel times in this chapter. The first is the within-day covariance of path travel times, which takes into account the variations in the path travel times within a day. The second is the day-to-day covariance of path travel times, which is based on the variations between the path travel times in a certain time interval on different days. These two types of temporal covariance functions of path travel times by 2-minute intervals have been considered in the proposed filtering algorithm in Section 3.3. The third is the covariance of path travel times between vehicle classes in a specific time interval. In contrast to the previous models, the consideration of these three types of covariance relationships allows for more explicit modeling of path travel times over time by vehicle class.

In the online prediction stage, the mean and temporal covariance of path travel times by vehicle class obtained in offline prediction can be regarded as prior information. Prior information is updated when real-time AVI data are gathered and uploaded to the system database. The posterior of predicted path travel time for multiple vehicle classes is the output of the model that is disseminated to road users and transportation management authorities.

4.3.1 Offline prediction

Based on the distinct features of traffic data, the travel time must be collected either directly or indirectly from various data sources. The $y_{i,d,p}^{s_A,k}$, *i*-th observed path travel time for path *p* of vehicle class *k* for from data source s_a on day *d* is defined as the difference between the timestamps of this vehicle entering and leaving the path:

$$y_{i,d,p}^{s_A,k} = \tau_{i,d,d_p,p}^{s_A,k} - \tau_{i,d,o_p,p}^{s_A,k}$$
(4.1)

where $\tau_{i,d,d_p,p}^{s_A,k}$ and $\tau_{i,d,o_p,p}^{s_A,k}$ are the timestamps of *i*-th observed path travel time at AVI sensor location d_p (detestation) and o_p (origin) for path *p* of vehicle class *k* from data source s_A on day *d*.

Point sensor data contain the average spot speed of vehicles at different locations along path p. The speed-based model (Li et al., 2006) is adopted to convert spot speed data to the path travel time as follows:

$$y_{i,d,p}^{s_{W},0} = \sum \frac{2(x_{i+1,p} - x_{i,p})}{v_{x_{i+1,p},t} + v_{x_{i,p},t}}$$
(4.2)

where $x_{i,p}$ is the location of *i*-th point sensor ordered along the study path *p*. As point sensor data have no vehicle class information. Vehicle class *k* equals 0, which represents the overall vehicle class.

GPS sensors provide the trajectory of the *i*-th vehicle of class $k: z_i^k(t + j\Delta g)$. As the captured trajectories may not cover all of the target path, extrapolation is conducted using:

$$y_{i,d,p}^{s_{B,k}} = \frac{\left(\sum_{j=1}^{J-1} t_{i,j+1}^{k} - t_{i,j}^{k}\right) \left(x^{d_{p}} - x^{o_{p}}\right)}{\sum_{j=1}^{J-1} x_{i,j+1}^{k} - x_{i,j}^{k}}$$
(4.3)

where $x^{d_p} - x^{o_p}$ is the distance between the origin and destination of the path p.

It is assumed that observed path travel times from different sources, such as AVI sensors, point sensors, and GPS sensors, can be partitioned into M categories by traffic conditions:

$$S = \Omega_1 \cup \dots \cup \Omega_M \tag{4.4}$$

where the set of travel time data is denoted Ω . For any two categories m and m',

$$\Omega_m \cap \Omega_{m'} = \emptyset \tag{4.5}$$

The value of M categories is determined by the statistical distributions of observed path travel times. The distributions of path travel times consist of different traffic conditions. The category for each traffic condition can be determined by clustering data with similar path travel time measurements. Some related papers directly selected the value of M by observing traffic conditions (Mil and Piantanakulchai, 2018; Yang et al., 2018). There are speed or travel time thresholds to partition traffic conditions. However, It can be observed that the number of components within a multimodal path travel time distribution can be distinct for various paths in the empirical study of Wang et al. (2021b). Therefore, it is worthwhile to generalize the model by clarifying the principles for setting the number of categories.

In this chapter, the value of *M* is decided by applying BIC (Schwarz, 1978), which has been commonly employed for model selection (Ma et al., 2018; Zhong et al., 2017). The BIC can be obtained by $BIC_{model} = -2 \ln \hat{L}_{model} + npar_{model} \ln n$. \hat{L}_{model} is the maximized value of the likelihood function for the clustering model. $npar_{model}$ denotes the number of parameters in the clustering model. *n* represents the sample size used in the clustering model. For a series of candidate clustering models with different values of *M*, the one with the minimum value of BIC_{model} is preferred. As there are multiple data sources (set *S*) providing travel time data, the vectors of observed path travel times $Y_p^{\{s_A, s_W\}}$ for path *p* are denoted as:

$$\boldsymbol{Y}_{p}^{\{s_{A},s_{W}\}} = \left\{ y_{i,d,p}^{s_{A},1}, y_{i,d,p}^{s_{A},2}, \dots, y_{i,d,p}^{s_{A},|K_{s_{A}}|}, y_{i,d,p}^{s_{W},0} \right\}^{T}$$
(4.6)

if both AVI data and point sensor data are used simultaneously. $y_{i,d,p}^{\{s_A,s_W\},1}$ is *i*-th element of observed path travel times $Y_p^{\{s_A,s_W\}}$ with vehicle class 1 from data sources s_A and s_W for path p on day d. The dimension $P_{Y_n^{\{s_A,s_W\}}}$ of $Y_p^{\{s_A,s_W\}}$ is:

$$P_{Y_p^{\{s_A, s_W\}}} = |K_{s_A}| + 1 \tag{4.7}$$

where $|K_{s_A}|$ is the number of vehicle classes that can be available from data source s_A (AVI sensors). It should be noted that the value of $|K_{s_A}|$ should be determined after balancing the number of vehicle classes and the sample size within each vehicle class. The sample size checking for each vehicle class in Tam and Lam (2011b) is followed. On the one hand, the classes with insufficient sample sizes should be merged into one class for model training. On the other hand, the significance tests (e.g., Kolmogorov–Smirnov test) can be applied to ensure there are significant differences between vehicle classes, as shown in Table 4.2.

The observed path travel time $Y_p^{\{s_B, s_W\}}$ for path p are expressed as:

$$\boldsymbol{Y}_{p}^{\{s_{B},s_{W}\}} = \left\{ y_{i,d,p}^{s_{B},1}, y_{i,d,p}^{s_{B},2}, \dots, y_{i,d,p}^{s_{B},|K_{s_{B}}|}, y_{i,d,p}^{s_{W},0} \right\}^{I}$$
(4.8)

if both GPS data and point sensor data are adopted simultaneously. The dimension $P_{Y_p^{\{s_B,s_W\}}}$ of $Y_p^{\{s_B,s_W\}}$ is:

$$P_{Y_p^{\{s_B, s_W\}}} = |K_{s_B}| + 1 \tag{4.9}$$

where $|K_{s_B}|$ is the number of vehicle classes that can be available from data source s_b (GPS sensors). The determination of $|K_{s_B}|$ is the same as that of $|K_{s_A}|$. It should be noted that the dimension of input is decided by the most appropriate scheme for data usage with consideration of double-counting problems instead of a simple summation of the number of unique vehicle classes monitored by all data sources.

The data sources s_A and s_W are used as an example to illustrate the usage of multisource data (same procedure can be applied for data sources s_B and s_W) in the following modeling framework. $T_p^{\{s_A, s_W\}}$ is denoted as the predicted path travel time for path p based on data sources s_A and s_W . The dimension of $T_p^{\{s_A, s_W\}}$ is $D_{T_p^{\{s_A, s_W\}}}$.

For each time interval *t*, the probability density function (p.d.f.) of $(\mathbf{Y}_{p,t}^{\{s_A,s_W\}}, \mathbf{T}_{p,t}^{\{s_A,s_W\}})$ is:

$$f\left(\mathbf{Y}_{p,t}^{\{s_{A},s_{W}\}}, \mathbf{T}_{p,t}^{\{s_{A},s_{W}\}}\right) = \sum_{m=1}^{M} \pi_{m} f\left(\mathbf{Y}_{p,t}^{\{s_{A},s_{W}\}} \middle| \Omega_{m}\right) f\left(\mathbf{T}_{p,t}^{\{s_{A},s_{W}\}} \middle| \mathbf{Y}_{p,t}^{\{s_{A},s_{W}\}}, \Omega_{m}\right)$$
(4.10)

where π_m is the probability of the *m*-th category of traffic condition, $f\left(Y_{p,t}^{\{s_A,s_W\}} \middle| \Omega_m\right)$ is the p.d.f of $Y_{p,t}^{\{s_A,s_W\}}$ in category *m* from data sources s_A and s_W , and $f\left(T_{p,t}^{\{s_A,s_W\}} \middle| Y_{p,t}^{\{s_A,s_W\}}, \Omega_m\right) f\left(y_t \middle| x_t, \Omega_m\right)$ is the conditional p.d.f. of $T_{p,t}^{\{s_A,s_W\}}$ given $Y_{p,t}^{\{s_A,s_W\}}$ in category *m*. It is assumed that $Y_{p,t}^{\{s_A,s_W\}} \middle| \Omega_m$ follows a *P*-variate normal distribution, with mean vector $\boldsymbol{\mu}_{Y_{p,t}^{\{s_A,s_W\}}}$ and covariance matrix $\boldsymbol{\Sigma}_{Y_{p,t}^{\{s_A,s_W\}}}$, for m = $1, \dots, M$.

Assuming linear effects on $T_{p,t}^{\{s_A,s_W\}}$ in the *m*-th category with the vector of linear effects β_m , it can be written as:

$$\boldsymbol{\mu}_{\boldsymbol{Y}_{m,p}^{\{s_{A},s_{W}\}}}\left(\boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}};\boldsymbol{\beta}_{m}\right) = \boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}}\boldsymbol{\beta}_{m}$$
(4.11)
where $\boldsymbol{\mu}_{Y_{m,p}^{\{s_A,s_W\}}}\left(Y_{p,t}^{\{s_A,s_W\}};\beta_m\right)$ is the conditional expectation of $TT_{p,t}^{\{s_A,s_W\}}|Y_{p,t}^{\{s_A,s_W\}}=$ $Y_{p,t}^{\{s_A,s_W\}}, \Omega_m Y_t|(X_t = \boldsymbol{x}_t, \Omega_m.$ It is noted that the value of β_m varies for diverse traffic conditions. Hence, the p.d.f. in Eq. (4.10) is written as:

$$f\left(\boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}}, \boldsymbol{T}_{p,t}^{\{s_{A},s_{W}\}}; \boldsymbol{\psi}\right) = \sum_{m=1}^{M} \frac{\pi_{m} \,\boldsymbol{\varphi}_{P}}{\varphi_{p}^{\{s_{A},s_{W}\}}} \left(\boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}}; \boldsymbol{\mu}_{Y_{m,p}^{\{s_{A},s_{W}\}}}, \boldsymbol{\Sigma}_{Y_{m,p}^{\{s_{A},s_{W}\}}}\right)$$
(4.12)
$$= \sum_{m=1}^{M} \frac{\varphi_{P}}{\varphi_{P}} \left(\boldsymbol{T}_{p,t}^{\{s_{A},s_{W}\}} \left| \boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}}; \boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}} \beta_{m}, \boldsymbol{\Sigma}_{T_{m,p}^{\{s_{A},s_{W}\}}}\right)$$

where $\boldsymbol{\psi}$ denotes the set of parameters in the GMM. Furthermore, both $\boldsymbol{\varphi}_{P_{Y_{p}^{\{S_{A},S_{W}\}}}}$ and $\boldsymbol{\varphi}_{D_{T_{p}^{\{S_{A},S_{W}\}}}}$ are Gaussian random vectors with $P_{Y_{p}^{\{S_{A},S_{W}\}}}$ and $D_{T_{p}^{\{S_{A},S_{W}\}}}$ variates. As it is assumed that $Y_{p,t}^{\{S_{A},S_{W}\}} |\Omega_{m} \sim N_{P_{Y_{p}^{\{S_{A},S_{W}\}}}} \left(\boldsymbol{\mu}_{Y_{m,p}^{\{S_{A},S_{W}\}}}, \boldsymbol{\Sigma}_{Y_{m,p}^{\{S_{A},S_{W}\}}} \right)$, for m = 1, ..., M, if the p.d.f. of $Y_{p,t}^{\{S_{A},S_{W}\}} |\Omega_{m}$ is not dependent on category m (which means that $\boldsymbol{\varphi}_{P_{Y_{p}^{\{S_{A},S_{W}\}}}} \left(Y_{p,t}^{\{S_{A},S_{W}\}}; \boldsymbol{\mu}_{Y_{m,p}^{\{S_{A},S_{W}\}}}, \boldsymbol{\Sigma}_{Y_{m,p}^{\{S_{A},S_{W}\}}} \right) = \boldsymbol{\varphi}_{P_{Y_{p}^{\{S_{A},S_{W}\}}}} \left(Y_{p,t}^{\{S_{A},S_{W}\}}; \boldsymbol{\mu}_{Y_{p}^{\{S_{A},S_{W}\}}}, \boldsymbol{\Sigma}_{Y_{p,t}^{\{S_{A},S_{W}\}}} \right)$ for m = 1, ..., M), Eq. (4.12) can be rewritten as:

$$f\left(\boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}}, \boldsymbol{T}_{p,t}^{\{s_{A},s_{W}\}}; \boldsymbol{\psi}\right) = \boldsymbol{\varphi}_{P}\left(\boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}}; \boldsymbol{\mu}_{\boldsymbol{Y}_{p}^{\{s_{A},s_{W}\}}}, \boldsymbol{\Sigma}_{\boldsymbol{Y}_{p}^{\{s_{A},s_{W}\}}}\right)$$

$$\prod_{d=1}^{D} \sum_{m_{d}=1}^{M_{d}} \pi_{m_{d}} \boldsymbol{\varphi}_{m_{d}}\left(T_{i,p,t}^{\{s_{A},s_{W}\}} \middle| \boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}}; \boldsymbol{Y}_{p,t}^{\{s_{A},s_{W}\}} \beta_{m_{d}d}, \sigma_{m_{d}d}^{2}\right)$$
(4.13)

where $T_{i,p,t}^{\{s_A,s_W\}}$ is the *i*-th element of $T_{p,t}^{\{s_A,s_W\}}$. Moreover, the conditional distribution of $T_{i,p,t}^{\{s_A,s_W\}} | Y_{p,t}^{\{s_A,s_W\}} = Y_{p,t}^{\{s_A,s_W\}}$ varies among M_d disjoint categories.

In the real world, the sample size of traffic data for a specific vehicle class can be limited, as shown in Figure 4.2, especially when the frequency of the time interval is high (e.g., 2-minute interval). For vehicle classes with an insufficient sample size of real-time traffic data, the covariance of path travel times between vehicle classes can be considered. The modeling of this type of covariance connects the vehicle classes with and without enough sample size. It can contribute to the improvement of the prediction accuracy of the model. For any two vehicle classes k and k' ($k \neq k'$ $k \neq k'$), the covariance of path travel times is modeled as:

$$\Sigma_{t,d}^{\{s_A,s_W\}}(k,k') = Cov\left(Y_{p,t}^{\{s_A,s_W\},k}, Y_{p,t}^{\{s_A,s_W\},k'}\right)$$

$$= \sum_{m=1}^{M} \pi_m \left(\Sigma_{t,d}^{\{s_A,s_W\}}(k,k') + \mu_{Y_{p,t}^{\{s_A,s_W\},k}} \mu_{Y_{p,t}^{\{s_A,s_W\},k'}}\right)$$

$$- \sum_{m=1}^{M} \pi_m \mu_{Y_{p,t}^{\{s_A,s_W\},k}} \sum_{m=1}^{M} \pi_m \mu_{Y_{p,t}^{\{s_A,s_W\},k'}}$$
(4.14)

4.3.2 Online prediction

After training the proposed GMM, the offline predicted mean and various types of covariance of the path travel time by vehicle class are used for online prediction, where $T_p^{\{s_A, s_W\}^-}(t) = E(Y_{p,t}^{\{s_A, s_W\}}), \Sigma_p^{\{s_A, s_W\}^-}(t)$, and $\Sigma_{t,d}^{\{s_A, s_W\}^-}(k, k')$ are obtained from Eqs. (4.13) and (4.14). The superscript "–" indicates offline predicted path travel times based on Eqs. (4.1)-(4.14). With the updated observations from AVI and point sensor data, an online prediction using the Kalman Filter is performed as shown in Eqs. (4.15)-(4.17). The updated predicted path travel time are:

$$\boldsymbol{T}_{p}^{\{s_{A},s_{W}\}+}(t) = \boldsymbol{T}_{p}^{\{s_{A},s_{W}\}-}(t) + \boldsymbol{G}_{1}\left(\boldsymbol{Y}_{p}^{s_{A}}(t) - \boldsymbol{H}\boldsymbol{T}_{p}^{\{s_{A},s_{W}\}-}(t)\right)$$
(4.15)

where G_1 is the updating matrix in the Kalman Filter, $Y_p^{s_A}(t)$ is the vector of real-time measurements of path travel times at time interval t from data source s_A for path p, His the mapping matrix that connects $Y_p^{s_A}(t)$ and $T_p^{\{s_A,s_W\}^-}(t)$, and $T_p^{\{s_A,s_W\}^+}(t)$ is the vector of updated predicted path travel time based on data sources s_A and s_W , containing path travel times by vehicle class at time interval t.

The updated covariance matrix of path travel times is:

$$\boldsymbol{P}_{p}^{\{s_{A},s_{W}\}^{+}}(t) = \boldsymbol{P}_{p}^{\{s_{A},s_{W}\}^{-}}(t) + \boldsymbol{G}_{1}\boldsymbol{H}\boldsymbol{P}_{p}^{\{s_{A},s_{W}\}^{-}}(t)$$
(4.16)

where $P_p^{\{s_A, s_W\}^+}(t)$ is the updated covariance of predicted path travel time for path p, which will be further detailed in the form of within-day temporal covariance matrices $\Sigma_p^{\{s_A, s_W\}^+}(t)$ and the covariance of path travel times between vehicle classes $\Sigma_{t.d}^{\{s_A, s_W\}^+}(k, k')$. The updating matrix in the updating process is:

$$\boldsymbol{G}_{1} = \boldsymbol{P}_{p}^{\{s_{A}, s_{W}\}^{-}}(t)\boldsymbol{H}^{T} \left(\boldsymbol{H}\boldsymbol{P}_{p}^{\{s_{A}, s_{W}\}^{-}}(t)\boldsymbol{H}^{T} + \boldsymbol{R}\right)^{-1}$$
(4.17)

where \mathbf{R} is the covariance matrix of error. The different covariance matrices in Eqs. (4.16) and (4.17) are described in Eqs.(4.18)-(4.21) in detail. The various types of temporal covariance of path travel times need to be updated when new data are collected. For within-day and day-to-day covariance of path travel times of the same vehicle class:

$$\boldsymbol{G}_{2} = \boldsymbol{\Sigma}_{p}^{\{s_{A}, s_{W}\}^{-}}(t) \boldsymbol{H}^{T} \left(\boldsymbol{H} \boldsymbol{\Sigma}_{p}^{\{s_{A}, s_{W}\}^{-}}(t) \boldsymbol{H}^{T} + \boldsymbol{R} \right)^{-1}$$
(4.18)

where the within-day and day-to-day covariance matrices are updated according to

$$\Sigma_{p}^{\{s_{A},s_{W}\}^{+}}(t) = \Sigma_{p}^{\{s_{A},s_{W}\}^{-}}(t) + G_{2}H\Sigma_{p}^{\{s_{A},s_{W}\}^{-}}(t)$$
(4.19)

As both types of temporal covariance matrices are updated using Eqs. (4.18) and (4.19), the model trained in an offline manner can be used to update the covariance relationships in online prediction. Similarly, for the covariance of the path travel times between vehicle classes,

$$\boldsymbol{G}_{3} = \boldsymbol{\Sigma}_{t,d}^{\{s_{A},s_{W}\}^{-}}(k,k')\boldsymbol{H}^{T} \left(\boldsymbol{H}\boldsymbol{\Sigma}_{t,d}^{\{s_{A},s_{W}\}^{-}}(k,k')\boldsymbol{H}^{T} + \boldsymbol{R}\right)^{-1}$$
(4.20)

where the temporal covariance of path travel times between vehicle classes by 2minute intervals are updated according to

$$\Sigma_{t,d}^{\{s_A,s_W\}+}(k,k') = \Sigma_{t,d}^{\{s_A,s_W\}-}(k,k') + G_3 H \Sigma_{t,d}^{\{s_A,s_W\},-}(k,k')$$
(4.21)

With the application of Eqs. (4.20) and (4.21)the covariance of the path travel time between vehicle classes is updated using real-time AVI data. Hence, the predicted path travel times by vehicle classes are updated using Eq. (4.15) and provided to road users

and transportation management authorities.

4.4 Numerical Experiments

Numerical experiments using actual data collected from Hong Kong roads are conducted to evaluate the accuracy of the proposed prediction model, and to perform a sensitivity analysis for assessing the impacts of using different amount of data from the three data sources considered in this chapter.

4.4.1 Introduction of study sites

Various sources of traffic data have been preprocessed before training in the proposed prediction model. The outliers that are extremely large observed path travel times due to measurement errors or alternative path selection are removed using a modified algorithm based on the filtering algorithm of Dion and Rakha (2006). The mapmatching algorithm is applied for GPS data to identify vehicles on the study path and remove the outliers. The speed from point sensor data is also smoothed. The extreme values of speed that exceed the speed limit too much are eliminated as well for a better prediction performance of the proposed prediction model.

As the study path is a major route in Hong Kong's road network, the predicted path travel time of all vehicles (without vehicle classification) could be obtained from the SMPS, which provides the instantaneous path travel times of several major routes in Hong Kong.

The weekday AVI, GPS, and point sensor data from May to July 2018, excluding public holidays, are used in numerical experiments. The data of the last five weekdays in July 2018 (i.e., July 25, 26, 27, 30, and 31, 2018) are selected for validation of the predicted path travel time. The data of remaining weekdays in May to July 2018 are

trained for offline prediction. The MAPE and MAE are adopted for model evaluation.

4.4.2 Performance evaluation

In this numerical experiment, the vehicles are divided into four classes: (1) an overall class (vehicles of all classes), (2) private cars, (3) goods vehicles (vehicles that can be identified by GPS sensors), and (4) other vehicles. These four classes have distinct distributions, as shown in Figure 4.3. Table 4.4 provides a summary of input and output in the following experiments to avoid confusion. In summary, Figure 4.5 compares the results of predicted path travel time by vehicle class, as no ground truth by vehicle class is available. The predictions from Dion and Rakha (2006) using AVI data are compared. The other figures and tables show the validation predicted results on overall path travel time.

Results	Iı	nput s	et	Output vehicle class of path travel time				Validation	Comparison
(by the presenting order)	A	G	Р	Overall	Private cars	Goods vehicles	Other vehicles	set	set
Table 4.5	\checkmark	\checkmark	\checkmark	\checkmark	-	-	-	\checkmark	-
Table 4.6	\checkmark	-	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	-	\checkmark
Figure 4.5(a)	\checkmark	-	\checkmark	-	\checkmark	-	-	-	\checkmark
Figure 4.5(b)	\checkmark	\checkmark	\checkmark	-	-	\checkmark	-	-	\checkmark
Figure 4.5(c)	\checkmark	-	\checkmark	-	-	-	\checkmark	-	\checkmark
Figure 4.5(d)	\checkmark	-	\checkmark	\checkmark	-	-	-	-	\checkmark
Figure 4.6, Figure 4.7, Table 4.7, Table 4.8, Table 4.9	\checkmark	-	\checkmark	\checkmark	-	-	-	\checkmark	-
Table 4.10	-	\checkmark	\checkmark	\checkmark	-	-	-	\checkmark	-

Table 4.4 Summary of input and output in the experiments of Section 4.4.2

1. A denotes AVI data (S_A); 2. P is point sensor data (S_C); 3. G is short for GPS data (S_B). 4. Validation set refers to path travel times from the SMPS 5. Comparison set contains predicted results based on A.

Although the proposed prediction model can fuse different types of traffic data in a vector form, the sample sizes across these traffic data are different. Thus, the implicit weightings on these data sources can be varied. AVI sensors can collect the path travel times of vehicles when the vehicle journey (along this path) is completed. Similar path travel times can be obtained by analyzing the trajectories of vehicles (point speed and location) from GPS data.

Therefore, a comparison can be made on the number of vehicles per 2-minute interval (defined as average sample size) from different traffic data, as shown in Table 4.5. It is observed in Table 4.5 that the average sample size of point sensor data is the largest (21.2 per 2-minute interval). Consequently, the combination of AVI and point sensor data demonstrates optimal performance on predicted path travel times (MAPE=6.9%), as depicted in Table 4.5.

Data	*Average sample size per 2-minute interval	MAPE (%)	MAE (min)	
AVI data	1.2	9.4	1.5	
GPS data	0.2	12.8	2.1	
AVI + point sensor data	*01.0	6.9	1.0	
GPS + point sensor data	21.2	11.3	1.9	

Table 4.5 Path travel time prediction of the overall class using various data sources

^{*}The column of average sample size shows the number of vehicles per 2-minute interval observed from different traffic data.

**Average sample size per 2-minute interval of point sensor data is presented here.

Apart from the impacts of sample size, the coverage of different vehicle types can contribute to the prediction accuracy of path travel times. AVI data have path travel times of all vehicle types, while GPS data in this study are mainly obtained from goods

vehicles of commercial companies. As a result, it is found in Table 4.5 that AVI data weighs much more than GPS and point sensor data. The MAPEs are 9.4% for using AVI data and 11.3% for using GPS data and point sensor data simultaneously. However, the concurrent use of AVI and GPS data can lead to a double-counting issue, as discussed in Section 4.1. As a result, utilizing all AVI, GPS and point sensor data together only achieves an MAPE of 9.1% and an MAE of 1.4 minutes.

The differing results between AVI and GPS technologies are primarily raised by the difference in percentages of shares of vehicles (Deng et al., 2013). In Chapter 4, the AVI technology used for the case study is the ALPR technology, which can identify all commercial vehicles by their license plates, while the private car data cannot be used in this study due to privacy issues (the number of vehicles per 2-minute interval by AVI data is 1.2 in Table 4.5).

On the contrary, GPS data are only obtained from several specific commercial companies who are willing to share the data for this study. As a result, the number of vehicles per 2-minute interval by GPS data is 0.2 in Table 4.5. Only path travel times of goods vehicles are available from GPS data. Therefore, there is a significant gap in the percentages of shares of vehicles between AVI and GPS technologies used in the case study of Chapter 4.

For the predicted path travel times by vehicle class, a comparison is conducted to evaluate the difference between the predicted results from the proposed prediction model and predicted path travel times from AVI data using the previous prediction model. The results are shown in Table 4.6. The MAPEs for different vehicle classes are within 10%. The results of these two models are closest for predicting the path travel times of private cars (MAPE = 5.8% and MAE = 0.8 min).

In Table 4.6, the proposed prediction model's performance is most closely aligned with the estimates based on AVI data for predicting path travel times of private cars. This is attributed to the fact that the variability in path travel times for private cars is lower as compared to that for goods vehicles and other vehicle classes. Unlike private cars, goods vehicles and other classes encompass a diverse range of vehicles, including heavy trucks, concrete mix trucks, light goods vans, and ambulances. The prediction of their path travel times is inherently more challenging due to this diversity and complexity.

 Table 4.6 Comparison of predicted path travel time by vehicle class

Vehicle Class	MAPE (%)	MAE (min)	
Overall (100%)	7.2	1.1	
Private cars (24.8%)	5.8	0.8	
Goods vehicles (56.1%)	8.9	1.5	
Others (19.1%)	9.5	1.7	

The CDF plots of the prediction errors are given in Figure 4.5, with the prediction errors for private cars, goods vehicles, other vehicles, and all vehicles presented in Figure 4.5(a), Figure 4.5(b), Figure 4.5(c), and Figure 4.5(d), respectively. Figure 4.5(a) and Figure 4.5(c) show the distributions of absolute percentage errors in the predicted path travel time for private cars and other vehicles. The median and 95% of MAPE are 7.8% and 15.7% for private cars and 8.1% and 17.1% for other vehicles.



(b) Goods vehicles





A denotes AVI data; 2. P is point sensor data; 3. G is short for GPS data
 Figure 4.5 CDF plots of prediction errors for (a) private cars, (b) goods vehicles, (c) other vehicles, and (d) all vehicles

For goods vehicles, the results from the AVI and point sensor data and those obtained with GPS and point sensor data using the proposed prediction model are compared with the results acquired from the AVI data using the previous prediction model. In Figure 4.5(b), the median and 95% of MAPE are 7.4% and 17.1% if AVI and point sensor data are utilized (against 10.3% and 25% using GPS and point sensor data). $T_{p,t}^{\{s_A,s_W\},2}$ using AVI and point sensor data are closer to $\hat{T}_{p,t}^{s_A,2}$ using AVI data than $\hat{T}_{p,t}^{\{s_B,s_W\},2}$ based on GPS and point sensor data.

The CDF plot for all vehicles is demonstrated in Figure 4.5(d). Different from the other three figures, the validation results are shown using SMPS data as ground truth. The median and 95% of MAPE are 8.3% and 17.9%. Additionally, Figure 4.6 presents a comparison between the predicted path travel times and the actual data for five specific weekdays in 2018: July 25, 26, 27, 30, and 31. They are five typical weekdays with significant rainfall. Therefore, the impacts of rainfall on the accuracy of predicted path travel times can be evaluated.

The R-squared value for this analysis is determined to be 0.77, indicating a strong fit. A substantial 96.7% of the predicted travel times have absolute percentage errors below 20%, surpassing the benchmark set by Tam and Lam (2011b, 2013), who stipulated that at least 95% of predictions should fall within this error margin. The visual assessment of Figure 4.6 confirms that the predictive performance is satisfactory, aligning with the established requirements.



Figure 4.6 The comparison between predicted path travel times for overall vehicle class against ground truth on path travel times

To test the effects of covariance of path travel times between vehicle classes, the performances of the proposed prediction model with and without consideration of path travel times between vehicle classes are compared. For the former case, the MAPE and MAE for predicted path travel time for all vehicles are 6.9% and 1.0 min, respectively. For the latter case, the MAPE and MAE are 10.1% and 1.6 min, respectively. The corresponding CDF plots are given in Figure 4.7. The median and 95% of MAPE are 4.8% and 12.9% against 8.5% and 17.4% with/without considering the covariance of path travel times between vehicle classes. The median of errors is reduced to 56% (4.8%/8.5%) of the original one after considering this covariance. Thus, it can improve the prediction accuracy of path travel times.



Figure 4.7 CDFs of prediction errors for all vehicles with and without considering the covariance of path travel times between vehicle classes

Furthermore, recent machine-learning models are selected as benchmarks for comparison; i.e., the GMM proposed by Mil and Piantanakulchai (2018), the APTN proposed by Shi et al. (2021), the LSTM model proposed by Yang et al. (2021b), and the methods from Google Maps (Derrow-Pinion et al., 2021). These machine-learning models are recently popular with encouraging performance in the field of travel time prediction problems. They are the most updated and related to the research in this chapter.

Furthermore, these path travel time prediction models have been evaluated and benchmarked in recent related studies (Casabianca et al., 2021; Monje et al., 2022; Nie

et al., 2022). Additionally, the existing prediction method adopted by Google Maps has been compared (Derrow-Pinion et al., 2021). They are chosen as both $\Sigma_d^{S_A}(t_a, t_b)$ and $\Sigma_t^{S_A}(d_i, d_j)$ are implicitly considered in their designed neural networks. The best combination of data sources (i.e., AVI + point sensor data) available in this study is selected for comparison. Table 4.7 gives their performance results on path travel time prediction for the overall class. The proposed prediction model (MAPE = 6.9% and MAE = 1.0 min) outperformed the other benchmark models, including the one recently adopted by Google Maps (MAPE = 9.4% and MAE = 1.5 min).

Table 4.7 Benchmark comparison of path travel time prediction models of the overall

class

Models	MAPE (%)	MAE (min)	
Proposed prediction model (1)	6.9	1.0	
GMM (2)	9.1	1.4	
APTN (3)	10.6	1.7	
LSTM (4)	12.1	2.0	
Google Maps (5)	9.4	1.5	

(1) Proposed prediction model;
 (2) Mil and Piantanakulchai (2018);
 (3) Shi et al. (2021);
 (4) Yang et al. (2021b);
 (5) Derrow-Pinion et al. (2021).

Table 4.8 presents the performance of various models in predicting path travel times for different vehicle types. It is important to clarify that the term "performance" here denotes the validation results for the overall vehicle class, as well as comparative results for other vehicle classes. This is because ground truth data is only available for the path travel times of all vehicles (i.e. overall vehicle class), which is used for validation purposes.

Prediction models	All vehicles (Overall) (100%) (Validation)	Private cars (24.8%) (Comparison)	Goods vehicles (56.1%) (Comparison)	Other vehicles (19.1%) (Comparison)
Proposed prediction model (1)	6.9	5.8	8.9	9.5
GMM (2)	9.1	8.3	13.2	12.6
APTN (3)	10.6	10.2	11.0	13.8
LSTM (4)	12.1	11.5	12.9	14.4
Google Maps (5)	9.4	8.5	11.3	12.9

Table 4.8 Prediction performance of different models on predicted path travel times

by vehicle types

(1) Proposed prediction model; (2) Mil and Piantanakulchai (2018); (3) Shi et al.
 (2021); (4) Yang et al. (2021b); (5) Derrow-Pinion et al. (2021).

As observed in Table 4.8, the proposed prediction model demonstrates superior performance compared to the other models, both in terms of the overall vehicle class (MAPE of 6.9%) and the specific vehicle classes (5.8%, 8.9%, 9.5% for private cars, goods vehicles, and other vehicles, respectively). It is also found in Table 4.8 that private cars have the best comparison results, which is consistent with the results provided in Table 4.6. This is because the variability in path travel times for private cars is the lowest, as explained in Table 4.6.

Furthermore, the robustness of the proposed prediction model is tested with additional independent datasets collected from AVI, GPS, and point sensors during weekends from May to July 2018 (26 days). Other models are also compared. The results are given in Table 4.9. The proposed prediction model is robust and had the best performance (MAPE = 8.4% and MAE = 1.3 min) among the models.

Models	MAPE (%)	MAE (min)
Proposed prediction model (1)	8.4	1.3
GMM (2)	10.4	1.7
APTN (3)	16.2	2.5
LSTM (4)	18.1	2.8
Google Maps (5)	13.2	2.0

Table 4.9 Benchmark comparison of path travel time prediction models of overall

class on weekends

(1) Proposed prediction model;
 (2) Mil and Piantanakulchai (2018);
 (3) Shi et al. (2021);
 (4) Yang et al. (2021b);
 (5) Derrow-Pinion et al. (2021).

The sensitivity test on the proportion of data used in model training is conducted to evaluate the robustness of the proposed prediction model. Its effectiveness can be verified under scenarios when some traffic data are missing. On the one hand, the sample size of data may be less than expected for several reasons, including paths equipped with fewer sensors, sensor failure, accidents, bad weather conditions, etc. This test can verify the generality of the proposed prediction model under various scenarios.

On the other hand, prediction accuracy is usually required for ATIS in practice. This test can prove the robustness of the proposed prediction model with the expected

prediction accuracy when the amount of data is reduced. Moreover, as multiple data sources may have distinct effects on prediction accuracy, this test can identify the sensitivity of prediction accuracy to different data sources.

As the sample size of different data sources is shown in Figure 4.2, the effect of the data volume on prediction accuracy can also be examined. This analysis systematically eliminated varying proportions (ranging from 50% to 100%) of AVI, GPS, and point sensor data for model training purposes. The resulting accuracy of the predicted path travel times, quantified by MAPE, is detailed in Table 4.10. It is found that the predictive accuracy of path travel times declines more rapidly with the volume reduction of AVI data compared to the volume reduction of point sensor data.

Table 4.10 Effects of the proportions of AVI, GPS, and point sensor data used for

Percentage of	Percentage	of AVI data	Percentage of GPS data		
point sensor data	100%	50%	100%	50%	
100%	6.9%	30.7%	11.3%	34.7%	
50%	8.7%	25.2%	14.1%	41.4%	

training on prediction performance of path travel times (in terms of MAPE).

For example, it can be seen in Table 4.10 that reducing 50% of AVI data and point sensor data for training purposes results in respective MAPE of 25.2% and 8.7%. The rates of change of MAPEs are 0.37 and 0.04 for AVI and point sensor data, respectively. The accuracy of predicted path travel times is thus more sensitive to AVI data than point sensor data.

Similarly, the MAPEs are 34.7% and 14.1%, by 50% of reduction on the GPS data and point sensor data for training. The rates of change of MAPE are 0.47 and 0.06 for GPS and point sensor data, respectively. It is concluded that the accuracy of predicted path travel times is more sensitive to GPS data than to point sensor data. Furthermore, regarding the rates of change of MAPE for AVI and GPS data (0.37 and 0.47), it illustrated that the accuracy of predicted path travel times is more sensitive to GPS data than AVI data.

4.5 Findings

Considering the temporal covariance of the path travel times by vehicle class, this chapter proposes a novel prediction model to use multi-source traffic data for the prediction of multi-class path travel times. The AVI data and GPS data are combined with point sensor data for predicting path travel times for different vehicle classes. The proposed prediction model includes various types of temporal covariance of the path travel times (especially the covariance of path travel times between vehicle classes) to improve the prediction accuracy of the path travel times by vehicle class.

The proposed prediction model is validated using multi-source traffic data in Hong Kong. For validation, the results of the proposed prediction model are compared with those using four existing machine-learning models. The proposed prediction model outperforms the benchmark models in the prediction of the path travel times. The proposed prediction model has a MAPE of 6.9%, whereas the worst performing benchmark model has a MAPE of 14.1%.

Moreover, the predicted path travel times for four vehicle classes, namely all vehicles, private cars, goods vehicles, and others, are provided. The predicted travel times of private cars are closest to the predicted results obtained using the previous AVI databased prediction model (MAPE = 5.8%). The experiments also indicate that consideration of the covariance of path travel times between vehicle classes can improve the prediction accuracy. (MAPE reduces from 10.1% to 6.9%, and MAE

drops from 1.6 min to 1.0 min).

The comparison of the distributions of the prediction errors by vehicle class highlights that the error distributions vary appreciably. It demonstrates the necessity of multiclass path travel time prediction. The MAPE values at the 95 percentiles vary from 15.7% (private cars) to 25% (goods vehicles). Additionally, the satisfactory performance of the predicted path travel time of the overall vehicle class confirms the ability of the proposed prediction model to predict the overall average path travel times. The validation results based on different independent datasets demonstrate the robustness of the proposed prediction model.

The proposed prediction model had a MAPE of less than 10% with another dataset, whereas the MAPEs of the benchmark models exceeded 10%. A sensitivity test on the proportion of data used in model training shows the importance of using various data sources. For experiments with AVI and point sensor data, the rates of change of MAPE for AVI data and point sensor are 0.37 and 0.04, respectively.

For experiments with GPS and point sensor data, the rates of change of MAPE for GPS data and point sensor are 0.47 and 0.06, respectively. The prediction accuracy of the overall path travel times is thus more sensitive to AVI or GPS data than to point sensor data. Furthermore, by comparing the rates of change of MAPE for AVI and GPS data, the prediction accuracy of the overall path travel times is more sensitive to GPS data than to AVI data.

To further enhance the accuracy of the prediction by vehicle class under non-recurrent conditions, other data sources, such as weather and accident data, should be considered. Moreover, more advanced models for the preprocessing of multiple data sources could be developed and applied. More large-scale road networks can be further studied with

other types of traffic data sources. An efficient solution model should be further proposed to enable the prediction of multi-class travel times on various road links along different paths within the same network. Finally, the deviation of multi-class path travel times should be further investigated to improve path travel time prediction in the future.

As mentioned in Section 4.1, the predicted path travel times based on AVI data are obtained from Chapter 3 after data filtering. In the following Chapter 5, the output of this chapter (i.e., the predicted path travel times in the current time interval) will be the input for path travel time prediction in future time intervals.

5. Prediction of Path Travel Times Using Weather Forecast and Historical Rainfall Intensity

5.1 Preliminary Overview

5.1.1 Background information

Chapter 3 filters traffic data with the proposed unsupervised algorithm. Chapter 4 conducts multi-class path travel time prediction in the current time interval using the filtered traffic data. This chapter extends to predict path travel times in the future time intervals. For pluvial cities, the impacts of rainfall have been considered in previous studies on path travel time prediction. However, the effects are analyzed from the collected historical rainfall intensity data. For predicting path travel times in future time intervals, it is worthwhile utilizing the rainfall conditions in the near future (i.e., weather forecast) to improve the prediction performance.

Weather forecast information, particularly to cities with abundant precipitation, plays an important role for road users to make selections on their departure times and/or transportation modes etc. The probability of precipitation (POP) provides the likelihood of measurable precipitation at a specific location within a specified period. The forecasted rainfall amount (FRA) is the quantity of rainfall to occur at a specific location within a specified period. It is noted that the weather forecasts in this chapter refer to POP and/or FRA. Both inform road users of possible future road conditions. Similarly, weather forecasts should be considered to improve the accuracy of predicted path travel times.

However, the correctness of weather forecast information can be a concern, especially

for metropolitan cities with frequent rainfall (e.g., Hong Kong). Hong Kong has the highest average annual rainfall of 2431 mm among major Pacific Rim cities, as shown in Figure 5.1. Furthermore, there are around 140 rainy days per year in Hong Kong (from the World Weather Information Services (http://www.worldweather.org/)). Under these circumstances, inaccurate weather forecasts may cause adverse impacts (e.g., untimely weather signals during rainy days can even lead to casualties). Hence, there is a research gap to examine the effect of weather forecast correctness on the accuracy of path travel time prediction.



Figure 5.1 Average annual rainfall (mm) in major Pacific Rim cities

Apart from using weather forecasts, it is vital to fully investigate the effects of rainfall intensity data on path travel times. Due to the limitations of collecting and storage technology, previous studies mainly use hourly rainfall intensity to analyze these effects. However, the amount of rainfall intensity can be biased when the frequency for predicted path travel times is less than 1 hour. Figure 5.2 gives the CDF of 2-min and hourly rainfall intensity data. With the result of the K-S test, it is evident that the

rainfall intensity data with 2-min and hourly frequency have statistically different distributions. As this chapter aims to predict path travel times every 2 minutes, there is a research gap in incorporating high-frequency rainfall intensity data into the prediction model.



Figure 5.2 CDF of 2-min and hourly rainfall intensity data in 2018

The effects of rainfall on path travel times can also depend on traffic conditions. In this chapter, the traffic condition is described by the level of service (LOS). The LOS is defined by the Highway Capacity Manual (HCM, 2016) using letters A through F to represent different traffic conditions on the road. The LOS A stands for free-flow condition (best), while the LOS F is referred to as congested (worst) condition. This chapter also investigates the prediction accuracy under different levels of service.

Previous works attempted to use rainfall intensity/weather forecast for travel time

prediction in the near future. They generally considered weather data (e.g., rainfall intensity, weather condition (cloudy, sunny, snowy, rainy) (Qiao et al., 2016; Li, Wang, and Xiong, 2021), temperature, humidity, visibility, and wind speed (Yang and Qian, 2019; Walch, Neubauer, and Schildorfer, 2023)) as one variable (Yang et al., 2013; Qiu et al., 2016; Gazder and Ratrout, 2018; Wang et al., 2018; Wirtgen et al., 2022) or feature (Yu et al., 2010; Thakuriah and Tilahun, 2013; Nair et al., 2019; Sadeghi-Niaraki et al., 2020; Xue et al., 2020; Petelin, Hribar, and Papa, 2023) in their models. Table 5.1 summarizes the previous related studies utilizing rainfall intensity/weather forecast data for travel time prediction.

Related studies	Role of weather forecast (if any)	Frequency of rainfall intensity data θ (if any)	Prediction model	Prediction horizon Δt	Prediction step Δ
Thakuriah and Tilahu (2013)	Probability of each weather condition as a feature	1 hour	Support vector regression	30 min	5 min
Zhang et al. (2018)	Distinguish between rainfall and normal days	15 min	Weather- correction model	15 min	15 min
Harper, Qian, and Samaras (2021)	-	15 min	LASSO linear regression, support vector regression, random forest	15~30 min	15 min
This chapter	Correlate with path travel times	2 min	Two-stage modeling framework	1 hour/1 day/1 week	2 min

Table 5.1 Related studies utilizing rainfall intensity/weather forecast data for travel

time prediction

There are several findings from Table 5.1. First, a few papers use weather forecasts for travel time prediction, as it can be a good supplement for travel time prediction, especially in pluvial cities. However, these papers regard weather forecasts as a variable in the machine-learning model (Thakuriah and Tilahu, 2013) or an indicator of rainy days (Zhang et al., 2018). The relationship between weather forecasts and predicted path travel times can be further investigated to enhance prediction precision.

Second, the frequency of rainfall intensity data (which is denoted as θ) is 1 hour (low frequency) in some of previous papers (i.e., only hourly rainfall amount per hour is available). To predict traffic speeds once every 2 minutes (at higher frequency), it was assumed implicitly that the rainfall intensity is constant and uniform for all time intervals during each hour (Jia et al., 2017). The assumption can be relaxed when 2-minute rainfall intensity data is available.

Though different types of prediction models are adopted for travel time prediction using rainfall intensity/weather forecast data, the inappropriate use of weather data can worsen the prediction accuracy. It was found that additional rainfall input adversely affected the accuracy of typical models (e.g., ARIMA), as they were less effective to find the relationships between speed and rainfall intensity data (Jia et al., 2017). Therefore, it is vital to propose a modeling framework which fully utilizes weather data to enhance the prediction accuracy.

Third, the prediction horizon in previous studies varies from 2 minutes to 1 week with corresponding prediction steps (between 2 minutes and 1 hour), this chapter predicts path travel times one-hour/one-day/one-week ahead once every 2 minutes.

5.1.2 Contributions

In summary, the proposed modeling framework in this chapter extends the previous

related work by providing the four key contributions below:

C5.1 A novel two-stage modeling framework is proposed for predicting path travel times in the near future. Real-time weather forecasts are used to update the offline predicted path travel times based on historical rainfall intensity data.

C5.2 In the online updating stage, a modified Kalman filter is proposed to update the offline predicted path travel times with consideration of the normalized cross-correlation coefficient between the real-time weather forecasts and predicted path travel times, which outperforms the other benchmark updating models. Besides, the variations of this normalized cross-correlation coefficient under different levels of service are established. The effects of weather forecast correctness on prediction accuracy are also examined.

C5.3 In the offline prediction stage, an improved offline training model is proposed to use high-frequency rainfall intensity data, considering the normalized crosscorrelation coefficient of path travel times and rainfall intensity data under different levels of service. It can output the offline predicted path travel times with higher quality before online updating compared with the one from the other benchmark offline training models. In addition, the variations of this normalized cross-correlation coefficient under different levels of service are also studied. The advantages of using these high-frequency data are presented and discussed.

C5.4 The real-world dataset in Hong Kong is used to validate the proposed modeling framework. The effects of rainfall categories and levels of service on the accuracy of predicted path travel times are also assessed. Furthermore, its applicability is verified with and without using the ground truth on path travel times as input for the proposed modeling framework.

5.2 Research Questions

Consider a path p with a pair of AVI sensors installed at both ends (i.e., o_p and d_p). The *i*-th observed path travel time for path p of vehicle class k for from data source s_A on day d is denoted as $y_{i,d,p}^{s_A,k}$, which is consistent with Sections 3.2 and 4.2. For the prediction of path travel times at the current time interval t_0 , there are different types of input for the proposed modeling framework.

First, the rainfall intensity data on location x along path p at time interval t can be represented by $r_p(x, t)$ and is available for $t \le t_0$, with the frequency of θ ($\theta = 2$ minutes in this chapter). Second, the FRA for location x along path p at time interval $t_0 + \Delta t$ is $\hat{r}_p(x, t_0 + \Delta t)$, where Δt is the time ahead of current time interval t_0 (prediction horizon). Third, assuming that the set for rainfall categories is L, the POP for forecasting rainfall category l for location x along path p at time interval $t_0 + \Delta t$ along path p is denoted as $\hat{P}_{p,l}(x, t_0 + \Delta t)$, for $l \in L$.

As the weather forecasts are not 100% correct, the accuracy or correctness of weather forecasts is also considered in the proposed modeling framework. There are two types of weather forecast correctness. They are used to describe the accuracy of received weather forecast (i.e., the correctness of POP for forecasting rainfall category *l* for location *x* along path *p* at time interval $t_0 + \Delta t$ is $C_{\hat{P}_{p,l}(x,t+\Delta t)}$. The correctness of FRA for location *x* along path *p* at time interval $t_0 + \Delta t$ is $C_{\hat{r}_p(x,t)}$. The equations for obtaining $C_{\hat{P}_{p,l}(x,t+\Delta t)}$ and $C_{\hat{r}_p(x,t)}$ are given as follows:

$$C_{\hat{P}_{p,l}(x,t)} = \frac{100\%}{\Delta t} \sum_{t=1}^{\Delta t} \left(1 - \frac{\left| \hat{P}_{p,l}(x,t) - o_{p,l}(x,t) \right|}{o_{p,l}(x,t)} \right)$$
(5.1)

$$C_{\hat{r}_p(x,t)} = \frac{100\%}{\Delta t} \sum_{t=1}^{\Delta t} \left(1 - \frac{\left| \hat{r}_p(x,t) - r_p(x,t) \right|}{r_p(x,t)} \right)$$
(5.2)

where $o_{p,l}(x,t)$ and $r_p(x,t)$ are the observed frequency of rainfall at rainfall category l, and observed rainfall amount for location x along path p at time interval t, respectively. Δt is the prediction horizon to be assessed.

For example, when there are 30 2-minute intervals with weather forecast to be evaluated, $\Delta t = 30$. There are 30 values of $\hat{P}_{p,l}(x,t)$ (for each rainfall category l) and $\hat{r}_p(x,t)$ with the observed $r_p(x,t)$ and $o_{p,l}(x,t)$ (for each rainfall category l). The correctness can be calculated correspondingly. $C_{\hat{r}_p(x,t)}$ represents the accuracy of FRA, while $C_{\hat{F}_{p,l}(x,t)}$ measures the accuracy of POP under each rainfall category.

In this chapter, both $y_{i,d,p}^{s_A,k}$ and $y_{i,d,p}^{s_G,k}$ are available as studies in Chapters 3 and 4. With these available data, the proposed modeling framework is going to predict path travel times for path p at time interval $t_0 + \Delta t$ (i.e., $\hat{T}_{t_0+\Delta t,p}$), with the prediction step of Δ ($\Delta = 2$ minutes in this chapter).

5.3 Methodology

Compared with previous studies, this chapter extends to using the weather forecast in predicting path travel times. Therefore, an online updating stage is needed for updating the predicted path travel times obtained from the relationships between rainfall intensity data and path travel times. It is therefore a two-stage prediction model presented in this chapter.

5.3.1 Proposed modeling framework

Figure 5.3 gives the modeling framework for predicting path travel times. For the current time interval t_0 , there are two stages. In stage 1, four types of real-time data are involved in online updating. The first is the rainfall intensity data collected at t_0 including $r_p(x, t_0)$ (from data source S_1). The second is the observed path travel time from AVI sensors $y_{i,d,p}^{S_A,k}$ (from data source S_2). The third is the POP for Δt ahead, $\hat{P}_{p,l}(x, t_0 + \Delta t)$ (from data source S_3) and the fourth is the FRA $\hat{r}_p(x, t_0 + \Delta t)$ (from data source S_4) for Δt ahead. With input of these four real-time data and results from Stage 2, this stage can provide the final output $\hat{T}_{t+\Delta t,p}$. Both the prediction on withinday ($\Delta t \leq 24h$) and day-to-day ($\Delta t > 24h$) path travel times are the output.



Figure 5.3 The proposed modeling framework for path travel time prediction in the

near future

In stage 2, the proposed offline training model uses historical data $r_p(x, t)$ (from data source S_1) and $y_{i,d,p}^{s_A,k}$ for $t < t_0$ (from data source S_2) for calibration and/or training. Regarding the observed path travel time $T_p^{s_A}(t)$, the normalized cross-correlation coefficient of $r_p(x, t)$ and $T_p^{s_A}(t)$ is explicitly modeled for better quality of offline predicted travel times $\tilde{T}_{t+\Delta t,p}$ on path p for Δt ahead of time interval t.

5.3.2 Online updating

5.3.2.1 Real-time data

The real-time data refers to the available data collected on the current day, including both real-time traffic data and real-time rainfall intensity data, and the real-time weather forecast available on t_0 . The latter consists of real-time FRA and real-time POP. The observed path travel time $y_{i,d,p}^{s_A,k}$ is defined as the difference between the arrival and departure times of the *i*-th vehicle for path *p*:

$$y_{i,d,p}^{s_A,k} = \tau_{i,d,d_p,p}^{s_A,k} - \tau_{i,d,o_p,p}^{s_A,k}$$
(5.3)

where $\tau_{i,d,d_p,p}^{s_A,k}$ and $\tau_{i,d,o_p,p}^{s_A,k}$ are the arrival and departure time of the *i*-th vehicle of vehicle class k for path p from data source s_A on day d. As vehicle class is not the main contribution of this study, k = 0 (represents overall vehicle class) in this chapter.

The value of $T_{p,t}^{s_A}$ is available in the AVI system at current time interval t_0 for $t_0 - 1 \le \tau_{i,d,d_p,p}^{s_A,k} \le t_0$. The path travel time from data source s_A can be calculated as:

$$T_{p,t_0}^{s_A} = \frac{\sum_{i=1}^{n_{t_0}} y_{i,d,p}^{s_A,k}}{n_{p,t_0}}$$
(5.4)

where n_{p,t_0} is sample size of individual path travel times for path p collected at current time interval t_0 from data source s_A .

The impact of actual rainfall intensity on the travel time of path p is investigated and modeled in this chapter. Along the study path p, there can be several nearby rainfall stations that provide the actual rainfall intensity data, $r_p(x, t_0)$ can be obtained using external drift kriging (Kebaili Bargaoui and Chebbi, 2009; Shehu et al., 2023) as follows:

$$r_p(x, t_0) = \sum_{i=1}^{N_p} \lambda_{i,p} r_p(x_{i,p}, t_0)$$
(5.5)

where N_p is the number of nearby rainfall stations and $x_{i,p}$ is the location of *i*-th nearby rainfall station, with the corresponding kriging weight $\lambda_{i,p}$ for path *p*.

To determine the value of λ_i , the external drift $D(x_{i,p})$ is introduced in the kriging system as:

$$\sum_{i=1}^{N_p} \lambda_{j,p} \gamma (x_{j,p} - x_{i,p}) + \mu_1 + \mu_2 D(x_{j,p}) = \gamma (x_{j,p} - x), j = 1, \dots, N_{nb,p}$$
(5.6)

$$\sum_{i=1}^{N_p} \lambda_{i,p} = 1 \tag{5.7}$$

$$\sum_{i=1}^{N_p} \lambda_{i,p} Y(x_{i,p}) = Y(x)$$
(5.8)

where μ_1 and μ_2 are Lagrange parameters for spatial interpolation accounting for two constraints on $\lambda_{i,p}$.

The kriging variance is then derived as:

$$\sigma^{2} = \gamma(0) - \sum_{i=1}^{N_{p}} \lambda_{j,p} \gamma (x_{j,p} - x_{i,p}) - \mu_{1} - \mu_{2} Y(x)$$
(5.9)

The variogram is adopted to evaluate the covariance, which involves fitting a mathematical model to the empirical semivariogram of both the primary and external drift variables. It is:

$$\gamma_{\delta}(h) = w(\delta) * \left[1.5 * \left(\frac{h}{a(\delta)}\right) - 0.5 * \left(\frac{h}{a(\delta)}\right)^3$$
(5.10)

where δ is the duration of rainfall and $w(\delta)$ is the highest variance with provided rainfall intensity data. $a(\delta)$ refers to the distance over which the data are correlated. The least square method to fit the variogram and hence obtain the kriging weight λ_i .

5.3.2.2 Weather forecast and correctness

Weather forecasts should not be generated if they have been backed up. Actual weather forecasts have only been used for indication of normal and rainfall days (Zhang et al., 2018). In general, it is available at the current time interval on the current day. Unfortunately, the weather forecasts for historical days have not been stored by the Hong Kong Observatory. Traffic data and rainfall intensity data in 2018 are available in the dataset. However, forecast data in 2018 is unavailable. Hence, the weather forecast data from HK Observatory in September 2023 are extracted to acquire its correctness for generating weather forecasts in 2018.

The weather forecast downloaded from HK Observatory in September 2023 is used to work out the weather forecast correctness. It is then compared with results from previous literature. It is found that the weather forecast accuracy is similar as reported in other previous related studies (Wu et al., 2019; Zhu et al., 2022a). The average values of the Brier score (BS) range from 0.20 to 0.28. BS (denoted as *BS*) can be obtained from:

$$BS = \frac{1}{\Delta t} \sum_{t=1}^{\Delta t} \left(\hat{P}_{p,l}(x,t) - o_{p,l}(x,t) \right)$$
(5.11)

where $o_{p,l}(x, t)$ is the observed frequency of rainfall at rainfall category l for location x along path p at time interval t. Δt is the prediction horizon to be evaluated.

The proposed updating model extends to consider the weather forecast to improve the prediction accuracy of path travel times. With relevant findings from previous papers

(Sun et al., 2023; Lyu et al., 2023), it is worthwhile to examine the effects of weather forecast and their correctness on the accuracy of predicted path travel times.

Moreover, the effects of weather forecasts can be varied under different levels of service on the study path. The criterion for level of service includes density, speed, maximum volume/capacity ratio, and maximum service flow rate. Inspired by the approach of Wilby et al. (2022), the thresholds of average speeds for distinguishing different levels of service together with free-flow travel speed are taken to obtain the ratio between these two variables. LOS_p obtained by Eq. (3.35) is also used in this chapter.

5.3.2.3 Hybrid model

A hybrid model combining the interacting multiple model and cubature Kalman filter has been proposed to update the offline predicted path travel times. Previous research has studied the impacts of rainfall intensity by different rainfall categories (Lam et al., 2008; Li et al., 2016). It can be extended to assume that the effects of rainfall intensity on traffic are distinguished under different modes of rainfall categories and levels of service. Therefore, for each combination of the specific rainfall category and level of service, there is a corresponding mode for modeling the concrete impacts. The rainfall category and level of service function as the boundary conditions of these effects.

There are different modes regarding levels of service and rainfall categories. The set of these modes with the rainfall effects on travel times of path p is denoted as M_p . Therefore, there are $|M_p|$ modes of distinct rainfall effects in the proposed modeling framework. Due to the uncertainties of weather forecast and path travel times, the transitional probability from mode i to mode j for path p exists for $i, j \in M_p$. It is denoted as $P_{i,j,p}$. For each time interval t, the normalized probability $w_{t,p}^{i,j}$ for mode ito j is:

$$w_{t,p}^{i,j} = \frac{1}{c_{t,p}^{j}} w_t^{i} P_{i,j,p}, i, j \in M_p$$
(5.12)

where $c_{t,p}^{j}$ is the normalization factor which can be obtained by:

$$c_{t,p}^{j} = \sum_{i=1}^{|M_{p}|} w_{t,p}^{i} P_{i,j,p}$$
(5.13)

Although the FRA is more informative than POP, it is worthwhile to integrate both of them to improve predicted path travel times. With the use of the Bayes' theorem, the following equation holds:

$$w_{t_{0}}^{\left(i,j\middle|\hat{P}_{p,l}(x,t_{0}+\Delta t)\right)} = \frac{\left(C_{\hat{P}_{p,l}(x,t_{0}+\Delta t)}\hat{P}_{p,l}(x,t_{0}+\Delta t)\middle|i,jw_{t_{0}-1}^{i,j}\right)}{C_{\hat{P}_{p,l}(x,t_{0}+\Delta t)}\hat{P}_{p,l}(x,t_{0}+\Delta t)}$$
(5.14)

where the POP for forecasting rainfall category l for location x along path p at time interval $t_0 + \Delta t$ along path p with a certain probability of transition from mode i to mode j at time interval t_0 is $\hat{P}_{p,l}(x, t_0 + \Delta t)|i, j$. It can be identified at time interval t_0 .

It is assumed that the state vector \mathbf{x}_p includes m variables (e.g., m = 3 in the proposed updating model including $\hat{r}_p(x, t_0 + \Delta t), r_p(x, t_0)$, and T_{t_0, p, s_A} for path p) in the system, where multivariate Gaussian distribution is presumed for these m variables. Therefore, $\mathbf{x}_p \sim N(\boldsymbol{\mu}_p, \boldsymbol{Q}_p)$, where $\boldsymbol{\mu}_p$ is the vector of mean and \boldsymbol{Q}_p is the covariance matrix of these m variables for path p. With distinguished effects of different FRA and POP, the mixed initial state is:

$$\widehat{\mathbf{x}}_{t_0+\Delta t,p}^{+0,j} = \sum_{i=1}^{|M_p|} w_{t_0}^{(i,j|\widehat{P}_{p,l}(x,t_0+\Delta t))\widehat{\mathbf{x}}_{t_0+\Delta t,p}^{+i}}$$
(5.15)

where $\widehat{\mathbf{x}}_{t_0,p}^{+0,j}$ is the updated initial state for mode j at time interval t_0 for path p. Correspondingly, the updated covariance matrix $\widehat{\mathbf{Q}}_{t,p}^{+0,j}$ for mode j is:

$$\widehat{\boldsymbol{Q}}_{t_{0}+\Delta t,p}^{+0,j} = \sum_{i=1}^{|M_{p}|} w_{t_{0}}^{\left(i,j\middle|\widehat{P}_{p,l}(x,t_{0}+\Delta t)\right)\left[\widehat{\boldsymbol{q}}_{t_{0}+\Delta t,p}^{+j}+\left(\widehat{\boldsymbol{x}}_{t_{0}+\Delta t,p}^{+i}-\widehat{\boldsymbol{x}}_{t_{0}+\Delta t,p}^{+0,j}\right)\left(\widehat{\boldsymbol{x}}_{t_{0}+\Delta t,p}^{+i}-\widehat{\boldsymbol{x}}_{t_{0}+\Delta t,p}^{+0,j}\right)^{T}\right]$$
(5.16)

In general, the Kalman filter corrects or updates the state after observation is collected. When observation vector $\mathbf{z}_{t_0,p} = \left[C_{\hat{r}_p(x,t_0+\Delta t)}, \hat{r}_p(x,t_0+\Delta t), r_p(x,t_0), T_{t_0,p,s_a} \right]^T$ is available, an observation operator **H** is applied as follows.

$$\widehat{\mathbf{x}}_{t_0+\Delta t,p}^{+j} = \widehat{\mathbf{x}}_{t_0-1+\Delta t,p}^{+0,j} + K \big(\mathbf{z}_{t_0,p} - H \widehat{\mathbf{x}}_{t_0-1+\Delta t,p}^{+0,j} \big)$$
(5.17)

where **K** denotes the Kalman gain. It should be noted that the offline predicted path travel times $\tilde{T}_{t+\Delta t,p}$ is an element in $\hat{x}_{t_0-1+\Delta t,p}^{+0,j}$ while the predicted path travel time $\hat{T}_{t+\Delta t,p}$ is an element in $\hat{x}_{t_0+\Delta t,p}^{+j}$ after updating. The state covariance matrix $\hat{Q}_{t,p}^{+0,j,post}$ can also be updated below:

$$\widehat{\boldsymbol{Q}}_{t_0+\Delta t,p}^{+j} = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\widehat{\boldsymbol{Q}}_{t_0+\Delta t-1,p}^{+0,j}$$
(5.18)

where I is the identity matrix. The K can be obtained using Eq. (5.19) below:

$$\boldsymbol{K} = \boldsymbol{\widehat{Q}}_{t_0 + \Delta t}^{+0,j} \boldsymbol{H}^{\mathrm{T}} \left(\boldsymbol{H} \boldsymbol{\widehat{Q}}_{t_0 + \Delta t}^{+0,j} \boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R} \right)^{-1}$$
(5.19)

The error covariance matrix \boldsymbol{R} is supposed to be unbiased.

5.3.3 Offline training

5.3.3.1 Proposed offline training model

The historical observed path travel time $y_{i,d,p}^{s_A,k}$ is modeled as a function of time interval t, i.e., $T_p^{s_A}(t)$. Moreover, it is assumed that the effects of rainfall intensity on path travel times by different rainfall categories and levels of service are similar in the online updating stage.

For each mode $j \in M_p$, the Geometric Brownian Motion model is introduced to
represent the evolution of path travel times.

$$dT_{p,j}^{s_A}(t) = T_{p,j}^{s_A}(t)\mu_{p,j}dt + T_{p,j}^{s_A}(t)\sigma_{p,j}dW(t)$$
(5.20)

where $\mu_{p,j}$ is the expected rate of change of path travel times over time for path p under mode j. $\sigma_{p,j}$ is the noise of path travel times for path p under mode j. dW(t) is the infinitesimal Wiener Process. The path travel times $T_{p,j}^{S_A}(t)$ under mode j is further modeled to distinguish the various effects of rainfall intensity under different rainfall categories and levels of service.

The interpolated rainfall intensity data $r_p(x, t)$ is available for any given location x from Eqs. (5.3)-(5.10). To correlate the path travel times, the average rainfall intensity for the path p can be derived as:

$$r_p(t) = \frac{1}{x^{d_p} - x^{o_p}} \int_{x^{o_p}}^{x^{d_p}} r_p(x, t) dx$$
(5.21)

The exponential Ornstein-Uhlenbeck process has been used to describe the evolution of rainfall intensity data with high frequency (i.e., 2 minutes).

$$dr_p(t) = r_p(t)\epsilon_p\left(\theta - \ln r_p(t)\right)dt + r_p(t)\sigma dW(t)$$
(5.22)

where ϵ_p is the relaxation rate to the historical mean of rainfall intensity for path p. It is assumed that there should be a normalized cross-correlation coefficient between rainfall intensity and path travel times $\rho_{p,r}(t)$. It is used to identify the time periods when path travel times are affected by rainfall. The following equation can be written as:

$$\rho_{p,r}(t) = \frac{E\left[\left(T_p^{s_A}(t) - \mu\left(T_p^{s_A}(t)\right)\right)\left(r_p(t) - \mu\left(r_p(t)\right)\right)\right]}{\sigma\left(T_p^{s_A}(t)\right)\sigma\left(r_p(t)\right)}$$
(5.23)

where $\rho_{p,r}(t)$ is normalized cross-correlation coefficient between rainfall intensity and path travel times at time interval t. Similarly, the normalized cross-correlation coefficient between FRA and path travel times can be obtained by:

$$\rho_{p,\hat{r}}(t) = \frac{E\left[\left(T_p^{s_A}(t) - \mu\left(T_p^{s_A}(t)\right)\right)\left(\hat{r}_p(t) - \mu\left(\hat{r}_p(t)\right)\right)\right]}{\sigma\left(T_p^{s_A}(t)\right)\sigma\left(\hat{r}_p(t)\right)}$$
(5.24)

where $\rho_{p,\hat{r}}(t)$ is normalized cross-correlation coefficient between FRA and path travel times at time interval t.

Due to the unavailability of high frequency data on rainfall intensity, the hourly-based rainfall intensity data is usually studied and used for modeling their effects on traffic state variables in the previous related works (Shao et al., 2008; Ryu, Kim, and Kim, 2020) and Table 5.1. The temporal evolution of rainfall in this situation can hardly be captured using low frequency rainfall intensity data.

Both duration and temporal evolution of rainfall affect rainwater accumulation on the road, which causes a reduction in road capacity and vehicular speed (especially speed at capacity). Hence, high-frequency rainfall intensity data can help to model the effect of rainfall intensity on path travel time prediction in the near future. If rainfall begins at t_1 and ends at t_2 on location x, with rainfall intensity $r_p(x,t)$, the rainwater accumulation AC(x,t) can be obtained by:

$$AC(x,t) = \int_{t_1}^{t_2} |r_p(x,t) - r_p d(x)| dt$$
(5.25)

where $r_p d(x)$ is the rate of rainwater drainage at path p on location x.

It is assumed that only rainfall intensities with significant rainwater accumulation on the road have impacts on the speed of vehicles. Therefore, path travel times during time intervals with significant rainwater accumulation (which is larger than the threshold $A_0(x)$) are affected by rainfall intensities.

$$t \in \{t^* | AC(x, t^*) \ge AC_0(x)\}$$
(5.26)

With the above consideration, there should be a threshold normalized cross-correlation

coefficient ρ_0 . Any time interval with a significant normalized cross-correlation coefficient between its rainfall intensity and path travel time should be considered as follows:

$$t \in \{t^* | \rho_{p,r}(t^*) \ge \rho_0\}$$
(5.27)

Therefore, the offline predicted path travel times for time interval $t + \Delta t$ should be:

$$\tilde{T}_{t+\Delta t,p} = T_{p,j}^{s_A}(t+\Delta t), j \in M_p$$
(5.28)

5.3.3.2 Historical data

The historical data used in Section 5.3.3.1 above refers to the available data collected before the current day, including both historical rainfall intensity data $r_p(x, t)$ and historical observed path travel time $y_{i,d,p}^{s_A,k}$. The derivations for real-time data in Section 5.3.2.1 can be applied to acquire historical data with Eqs. (5.3)-(5.10).

5.4 Model Validation

Four empirical tests on two selected expressways are conducted for model validation. The first test aims to highlight the advantages of using both the weather forecast and historical rainfall intensity data for path travel time prediction, as claimed in contribution 1. The second test shows the superiority of the online updating stage over other benchmarks with consideration of the normalized cross-correlation coefficient between the weather forecasts and predicted path travel times under different levels of service. It justifies contribution 2 in the online updating.

The third test is conducted to illustrate the merits of considering various effects of rainfall intensity on path travel times by different rainfall categories and levels of service, as well as the usage of high-frequency rainfall intensity data. Contribution 3 is presented through this test. The fourth test verifies the applicability of the proposed modeling framework without ground truth on path travel times as input. Contribution

4 has been demonstrated in this test.

5.4.1 Dataset description and preprocessing

The empirical tests are carried out on two selected expressways with nearby weather stations in Hong Kong (Figure 5.4). Study path 1 is a major route from Tuen Mun New Town to Tsuen Wan New Town. The chosen study path is 17.8 km long with a free-flow path travel time (based on the free-flow travel speed of 75 km/h) of 14.3 min. Study path 2 is 9.2 km long. It connects the Island Eastern Corridor on Hong Kong Island to the Western Harbor Crossing in Kowloon. The length of study path 2 is 9.2 km, and the corresponding free-flow path travel time is 8.8 min. It is noted that study path 2 has a signal near the destination while study path 1 has no signals.



Figure 5.4 The selected path for the empirical tests

As both study paths are major routes in Hong Kong's road network, the predicted path travel times have been worked out in the existing ATIS allocated in Hong Kong. The predicted path travel time for study paths 1 and 2 can be obtained from the SMPS and JTIS, respectively. Both SMPS and JTIS have been validated with satisfactory results

(Tam and Lam, 2013). Hence, in this chapter, it is regarded as ground truth for path travel times, which is $y_{i,d,p}^{s_G,k}$ for k = 0 as used in Chapter 3.

There are three weather stations³ nearby each of both study paths in Hong Kong $(N_{nb,p} = 3)$. The rainfall intensity data collected from these three weather stations are analyzed in this chapter. The graduation of rainfall intensity of three weather stations has been unified to be 0.5mm/h after data processing. The observed path travel times and rainfall intensity data of weekdays in 2018, excluding public holidays (i.e., 214 weekdays for the whole year of 2018), are used in this test. Five rainy weekdays out of the 214 weekdays (from Monday to Friday) in 2018 are selected for validation of the path travel time prediction according to the representative rainfall amount and traffic conditions.

As the historical weather forecast data in 2018 is unavailable, they are generated with the following two assumptions: A1: It is assumed that the distribution of weather forecast correctness in September 2023 remains the same as in 2018. A2: The relationship between the level of service and weather forecast correctness in September 2023 stays the same as in 2018. For example, when level of service C is matching to the weather forecast correctness of 85% obtained in 2023, a value of 85% as weather forecast correctness is assumed for periods in 2018 when the traffic condition is at level of service C.

The data on the level of service and weather forecast correctness collected in September 2023 are used in the artificial neural network to calibrate their relationship. 80% of data in September 2023 are used for calibration, with the remaining for testing. The relationship is verified using the testing data with satisfactory simulation

³ https://www.hko.gov.hk/en/cis/stn.htm

performance. This relationship is then applied with reference to the level of service data in 2018 to generate the weather forecast correctness in 2018 by time of day. The correctness and BS of POP and FRA during September 2023 are shown in Figure 5.5. The medians for correctness and BS of POP and FRA are 85% and 0.14, as well as 79% and 0.19, respectively.



Figure 5.5 CDF of correctness and BS of the weather forecast

The remaining 218 weekdays in 2018 are used as training sets. The MAPE, MAE, and root mean square error (RMSE) are adopted for evaluation of prediction performance. The results on MAPE, MAE and RMSE are obtained using:

$$MAPE = \frac{1}{\Delta t} \sum_{i=1}^{\Delta t} \left| \frac{\hat{T}_{t_0+i,p} - T_{t_0+i,p}^{s_G}}{T_{t_0+i,p}^{s_G}} \right|$$
(5.29)

$$MAE = \frac{1}{\Delta t} \sum_{i=1}^{\Delta t} \left| \hat{T}_{t_0+i,p} - T_{t_0+i,p}^{s_G} \right|$$
(5.30)

$$RMSE = \sqrt{\frac{1}{\Delta t} \sum_{i=1}^{\Delta t} \left(\hat{T}_{t_0+i,p} - T_{t_0+i,p}^{s_G} \right)^2}$$
(5.31)

where three values of Δt are selected to evaluate the proposed modeling framework under different prediction horizons. They are in terms of the number of intervals in each of these three prediction horizons; namely 30 intervals for 1-hour ahead, 720 and 5040 intervals for 1-day ahead and 1-week ahead, respectively.

Besides, the maximum errors are observed to assess the worst performance of the proposed modeling framework. The maximum absolute percentage error (MaxAPE) and maximum absolute error (MaxAE) are used as follows:

$$MaxAPE = \left\{ Max \left| \frac{\hat{T}_{t_0+i,p} - T_{t_0+i,p}^{s_G}}{T_{t_0+i,p,s_g}} \right| \left| i = 1, \dots, \Delta t \right\}$$
(5.32)

MaxAE =
$$\left\{ Max \left| \hat{T}_{t_0+i,p} - T^{s_G}_{t_0+i,p} \right| \left| i = 1, ..., \Delta t \right\}$$
 (5.33)

5.4.2 Performance of the proposed modeling framework

Different input data (denoted as c_1, c_2, c_3, c_4 , and c_5) are prepared to investigate the effects of these data on the accuracy of predicted path travel times. The availability of these input data is shown in Table 5.2.

Table 5.3 presents the prediction performance of the proposed modeling framework using these data. It is observed in Table 5.3 that c_1 has the best performance for predicting 1-hour ahead on path travel times (MAPE is 6.9% and 6.1% for study path 1 and 2, respectively) while c_5 has the worst performance of predicted path travel times 1-week ahead (MAPE is 20.6% and 15.7% for study path 1 and 2, respectively).

		Rea	l-time data			Historical data		
Input	Rainfall intensity data $r_p(x, t_0)$	Observed path travel time $y_{i,d,p}^{s_A,k}$	FRA $\hat{r}_p(x, t_0 + \Delta t)$	POP $\hat{P}_{p,l}(x, t_0 + \Delta t)$	Number of data sources	Rainfall intensity data $r_p(x,t)$	Observed path travel time $y_{i,d,p}^{s_A,k}$	
<i>c</i> ₁	\checkmark	\checkmark			4			
<i>c</i> ₂	\checkmark				3			
<i>C</i> ₃	\checkmark				3	All available		
<i>C</i> ₄	\checkmark				2			
<i>C</i> ₅					1			

Table 5.2 Availability of different input data

Table 5.3 Prediction performance of the proposed modeling framework with different

input data

			Δt (time intervals)									
Dath	Innut data	1-hour ahead (30)			1-day ahead (720)			1-week ahead (5040)				
1 411	input data	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE		
		(%)	(min)	(min)	(%)	(min)	(min)	(%)	(min)	(min)		
	<i>c</i> ₁	6.9	0.7	1.0	10.3	1.8	1.9	14.6	2.8	2.8		
Study	<i>C</i> ₂	9.8	1.7	1.8	12.9	2.4	2.5	16.1	3	3.2		
nath 1	<i>C</i> ₃	11.8	2.1	2.2	14.1	2.7	2.8	16.7	3	3.4		
paul 1	<i>C</i> ₄	12.2	2.3	2.5	15.4	2.8	3.1	18.5	3.9	4.0		
	<i>C</i> ₅	15.2	2.8	3.1	17.8	3.8	3.9	20.6	4.3	4.7		
	<i>c</i> ₁	6.1	0.5	0.6	7.7	1.3	1.3	11.1	1.9	2		
Study	<i>C</i> ₂	7.4	1.2	1.3	9.9	1.7	1.9	11.7	2	2.1		
path 2	<i>C</i> ₃	9.1	1.5	1.7	10.4	1.8	2	11.9	2.1	2.2		
	C ₄	9.5	1.6	1.8	10.9	1.9	2	13.4	2.6	2.7		
	<i>C</i> ₅	10.7	1.9	2	13	2.5	2.6	15.7	2.8	3.2		

With reference to performance on study path 1, there are several observations. Comparing c_1 and c_5 , both types of weather forecast information (including FRA and POP) and historical rainfall intensity data can significantly achieve better performance of predicted path travel times in the near future (MAPE reduced from 15.2% to 6.9% for 1-hour ahead prediction of path travel times, which is an accuracy improvement of $\frac{15.2\%-6.9\%}{15.2\%} = 55\%$).

Comparing results from c_2 and c_3 , it is found in Table 5.3 that the FRA is more useful than POP for enhancing the accuracy of predicted path travel times. It should also be noted that historical rainfall intensity data plays an important role in improving the offline prediction accuracy by modeling the normalized cross-correlation coefficient between rainfall intensity and path travel times from results of c_4 and c_5 .

Figure 5.6 presents the CDF plots of MAPE for predicting 1-hour ahead on path travel times for study path 1. The significant difference in median of error distributions between c_1 and c_5 (i.e., 15.4%-7.8%=7.6%) again emphasizes the advantage of using both weather forecast information and historical rainfall intensity data. Hence, the subsequent tests select c_1 as input data unless further specification.



Figure 5.6 CDF of MAPE for predicting path travel times 1-hour ahead for study path 1 using different input data

Furthermore, the CDF of c_1 has a shorter tail (i.e., less extreme errors) than that of c_5 . It illustrates the stable performance of the proposed modeling framework with the use of both weather forecast and historical rainfall intensity data. The normalized crosscorrelation coefficient between path travel times and weather forecasts (modeled in the online updating stage) and rainfall intensity (modeled in the offline training stage) by different levels of service ensures stable performance. These two types of crosscorrelation coefficients will be discussed in the following tests.

R1 is used to represent online updating in the proposed modeling framework. The performance of R1 is compared with the other benchmark updating models using data collected from study path 1. There are three benchmarks. The first is the Extended

fractional singular KF (EFSKF, R2 stands for this model). The second is the Unscented Kalman filter (UKF, signified as R3), and the third is the Ensemble Kalman filter (EKF, represented by R4). They are also compared in Table 5.4 and Table 5.5 (Nerini et al., 2019; Trinh et al., 2022; Nosrati et al., 2023).

		Δt (time intervals)									
Updating	1-hour ahead (30)			1-day ahead (720)			1-week ahead (5040)				
models	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE		
	(%)	(min)	(min)	(%)	(min)	(min)	(%)	(min)	(min)		
R1	6.9	0.7	1.0	10.3	1.8	1.9	14.6	2.8	2.8		
R2	10.1	1.8	1.9	12.9	2.4	2.5	15.7	2.9	3.1		
R3	11.0	1.9	2.0	14.3	2.7	2.8	16.9	3.1	3.5		
R4	11.9	2.3	2.4	15.0	2.8	3.0	18.4	3.9	4.0		

Table 5.4 Online updating performance of different updating models on study path 1

R1: online updating of proposed modeling framework; R2: EFSKF; R3: UKF; R4: EKF

Table 5.5 Maximum errors of online updating performance of different updating

model	ls on	stud	y pat	h 1
-------	-------	------	-------	-----

	Δt (time intervals)									
Updating	1-hour al	head (30)	1-day ah	ead (720)	1-week ahead (5040)					
models	MaxAPE	MaxAE	MaxAPE	MaxAE	MaxAPE	MaxAE				
	(%)	(min)	(%)	(min)	(%)	(min)				
R1	10.4	1.8	15.5	2.8	21.9	4.5				
R2	16.2	3.0	20.6	4.3	25.1	4.9				
R3	18.7	3.9	24.3	4.8	28.7	5.5				
R4	17.9	3.8	22.5	4.5	27.6	5.2				

R1: online updating of proposed modeling framework; R2: EFSKF; R3: UKF; R4: EKF

The proposed modeling framework achieves the best performance on updating (MAPE=6.9% in Table 5.4 and MaxAPE=10.4% in Table 5.5), with its superiority in dealing with non-linear systems. Besides, the normalized cross-correlation coefficient between predicted path travel times and the weather forecast is captured in the online updating stage and contributes to the prediction accuracy.

The normalized cross-correlation coefficient $\rho_{p,r}(t)$ (or $\rho_{p,\hat{r}}(t)$) of path travel times and rainfall intensity data (or FRA) at time interval *t* can identify the time periods when path travel times are affected by rainfall. Figure 5.7 compares $\rho_{p,r}(t)$ and $\rho_{p,\hat{r}}(t)$ under different traffic conditions in Figure 5.7(a) and Figure 5.7(b), respectively.



Figure 5.7 Comparison of normalized cross-correlation coefficient of (a) rainfall intensity data and path travel times ($\rho_{p,r}(t)$); (b) FRA and path travel times ($\rho_{p,\hat{r}}(t)$)

There are two observations from two figures. First, both $\rho_{p,r}(t)$ and $\rho_{p,\hat{r}}(t)$ are larger when the traffic condition is more congested. It is seen that the values are the largest when LOS is E&F in Figure 5.7(a) and 5.8(b). Second, the magnitude of $\rho_{p,r}(t)$ in Figure 5.7(a) is larger than $\rho_{p,\hat{r}}(t)$ in Figure 5.7(b). The maximum value is 0.93 in 5-27 Figure 5.7(a) and 0.62 in Figure 5.7(b). This implies that rainfall intensity data has a greater impact on predicted path travel times. The use of rainfall intensity data is more essential to path travel time prediction.

This is because the impacts on rainfall under congested traffic are more severe than in uncongested traffic. Lam et al. (2013) also pointed out that the percentage of reduction on speed at capacity is larger than that on free-flow speed under rainfall. The variations of the normalized cross-correlation coefficient between path travel times and weather forecast/historical rainfall intensity have been captured and modeled in Eqs. (12)-(16) and (23). This normalized cross-correlation coefficient distinguished by different levels of service helps to improve prediction accuracy.

The weather forecast correctness plays a crucial role in the quality of traffic prediction. This is because weather forecasts may not always be completely accurate (Thakuriah and Tilahun, 2013). Hence, the effects of weather forecast correctness on path travel time prediction accuracy are examined in this chapter using different percentages of reduction on the correctness of weather forecast based on c_1 . Figure 5.8 illustrates the effects of weather forecast correctness on the prediction accuracy (in terms of MAPE) of path travel times by 2-minute intervals in the next hour.



Figure 5.8 The effects of weather forecast correctness on the prediction accuracy (in terms of MAPE) of path travel times by 2-minute intervals in the next hour

There are two observations in Figure 5.8. First, there is a distinct difference between maximum point \mathcal{A} (22.3% when the medians of correctness for FRA and POP are 51% and 47%) and minimum point \mathcal{B} (6.9% for data combination c_1 without reduction on weather forecast correctness) of MAPE in Figure 5.8 (22.3%-6.9%=15.4%). It shows that weather forecast correctness significantly affects the accuracy of predicted path travel times.

When weather forecasts are inaccurate (e.g., the medians of correctness for FRA and POP are 51% and 47% when 40% of correctness for FRA and POP are reduced), the performance is even worse than data combination c_5 (MAPE=15.2%) when only path travel time observations are used for prediction.

Second, the rate of change for correctness of POP (1.2) is much higher than that of FRA (0.4). It indicates that the prediction accuracy is more sensitive to the correctness of POP than that of FRA. As the effects of rainfall are modeled by different rainfall categories, the inaccurate POP has a higher chance of leading to a wrong rainfall category with mismatched relationships on the actual conditions. Therefore, the prediction accuracy is more sensitive to the correctness of POP.

Two strategies for simulation of weather forecast data are adopted by Thakuriah and Tilahun (2013) to regenerate the POP. The simulated FRA is then obtained from the mean value of each category corresponding to POP. Strategy (\mathbb{A}) assumes the equal chance of weather condition for each time interval, while strategy (\mathbb{B}) derives the conditional probability of each weather condition based on progression of historical data. Both strategies can be applied to the dataset used in this chapter for comparison.

Strategy (A) assumes an equal chance of occurrence of weather events as the weather forecast. If there are N events, the resultant POP is:

$$POP = \frac{1}{N} \tag{5.34}$$

Two counts are considered for strategy (\mathbb{B}); the first is counting the instances when *i* occurs at time interval *t* and *j* occurs at time $t + \Delta t$ (call it f_1). The second is the count when *i* occurs at time interval *t* but without followed by *j* at $t + \Delta t$ (call it f_2). Then the probability of occurrence of *j* when *i* occurs can be obtained by:

$$POP_{i,j} = \frac{f_1}{f_1 + f_2}$$
(5.35)

Strategy (\mathbb{A}) employs an equal probability distribution for each rainfall category and calculates the average FRA in the dataset. This approach assumes that each category has an equal likelihood of occurring. Strategy (\mathbb{B}), on the other hand, incorporates

conditional probability to update both the POP and the FRA. This means that the forecast is adjusted based on the likelihood of specific conditions leading to certain rainfall categories.

The comparison of the correctness of the POP and FRA between strategy (\mathbb{A}) and strategy (\mathbb{B}) is presented in Figure 5.8. The MAPEs for strategy (\mathbb{A}) and strategy (\mathbb{B}) are 13.4% and 8.1% respectively. This indicates that strategy (\mathbb{B}) exhibits superior performance compared to strategy (\mathbb{A}), which is consistent with the observation reported by Thakuriah and Tilahun (2013).

However, it is important to note that the prediction performance can be further enhanced by increasing the correctness of the weather forecast. This can be achieved by collecting weather forecast data either manually or automatically and following the assumptions outlined in Section 5.4.1. By considering this additional information and adhering to the specified assumptions, the accuracy of the weather forecast data is further improved from 8.1% (strategy (\mathbb{B})) to 6.9% (strategy in this chapter). It is also recommended to update the strategies for generating weather forecast data in further study.

Three of the most commonly-used machine-learning models are selected in the offline training as benchmarks for comparing with offline training model in the proposed modeling framework (O1); i.e., APTN proposed by Shi et al. (2021) (listed as O2), CNN model (Dunne and Ghosh, 2013) (represented as O3) and the LSTM model proposed by Yang et al. (2021b), (denoted as O4) used for traffic state prediction.

The average performance of the proposed prediction model in path travel time prediction is given in Table 5.6 and Table 5.7. The proposed modeling framework outperforms the other offline training models in terms of offline prediction (MAPE=12.8% in Table 5.6 and MaxAPE=19.2% in Table 5.7). It is because of the consideration of normalized cross-correlation coefficient between path travel times and rainfall intensity data by different levels of service as presented in Figure 5.7.

path 1												
Offline		Δt (time intervals)										
training model	1-hour ahead (30)			1-day ahead (720)			1-week ahead (5040)					
	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE			
	(%)	(min)	(min)	(%)	(min)	(min)	(%)	(min)	(min)			
01	12.8	2.4	2.5	15.4	2.8	3.1	18.9	4.0	4.1			
02	13.7	2.6	2.7	16.2	3	3.2	19.8	4.1	4.4			
03	13.9	2.6	2.7	16.5	3	3.3	20.1	4.2	4.5			
04	14.1	2.7	2.8	16.9	3.1	3.5	20.2	4.2	4.5			

Table 5.6 Offline training performance of different offline training models on study

O1: offline training in the proposed modeling framework; O2: APTN; O3: CNN; O4: LSTM

Table 5.7 Maximum errors of offline training performance of different offline

training	model	s on	study	path	1
i annig	mouen	0 011	Stady	pum	-

Offling	Δt (time intervals)									
training	1-hour al	nead (30)	1-day ah	ead (720)	1-week ahead (5040)					
model	MaxAPE	MaxAE	MaxAPE	MaxAE	MaxAPE	MaxAE				
	(%)	(min)	(%)	(min)	(%)	(min)				
01	19.2	4.0	23.1	4.6	28.4	5.4				
O2	21.2	4.4	25.4	4.9	30.3	5.7				
03	21.9	4.5	25.9	5.0	31.7	5.9				
04	23.6	4.7	28.1	5.3	34.2	7.2				

O1: offline training in the proposed modeling framework; O2: APTN; O3: CNN; O4: LSTM

The usage of high-frequency rainfall intensity data is a prerequisite for uncovering the normalized cross-correlation coefficient shown in Table 5.8 and Table 5.9. A range of frequency from 2 minutes (high frequency) to 1 hour (low frequency) for rainfall intensity data has been used as input for the proposed offline training model. The high-frequency rainfall intensity data as input can lead to higher prediction accuracy (MAPE=12.8% in Table 5.8 and MaxAPE=19.2% in Table 5.9 for the proposed offline training model) compared with low-frequency data (i.e., MAPE=21.7% in Table 5.8 and MaxAPE=27.6% in Table 5.9). It is consistent with the findings in Harper, Qian, and Samaras (2021) and highlights the merit of using high-frequency rainfall intensity data.

Frequency		Δt (time intervals)								
of rainfall	1-hour ahead (30)			1-day ahead (720)			1-week ahead (5040)			
intensity	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	
data (min)	(%)	(min)	(min)	(%)	(min)	(min)	(%)	(min)	(min)	
2	12.8	2.4	2.5	15.4	2.8	3.1	18.9	4.0	4.1	
10	15.2	2.7	2.9	22.5	4.4	18.5	3.9	4.1	23.8	
30	18.5	3.9	4.0	20.8	4.4	4.8	23.4	4.7	5.1	
60	21.7	4.5	4.9	23.1	4.6	5.1	25.3	4.9	5.3	

 Table 5.8 Offline training performance of proposed offline training model using

 rainfall intensity data with different frequencies for study path 1

		Δt (time intervals)								
Offline training	1-hour ahead (30)		1-day ah	ead (720)	1-week ahead (5040)					
model	MaxAPE	MaxAE	MaxAPE	MaxAE	MaxAPE	MaxAE				
	(%)	(min)	(%)	(min)	(%)	(min)				
2 minutes	19.2	4.0	23.1	4.6	28.4	5.4				
30 minutes	25.3	4.9	29.4	5.6	32.3	6.9				
1 hour	27.6	5.2	33.8	7.0	38.4	7.6				

Table 5.9 Maximum errors of proposed offline training model using rainfall intensitydata with different frequencies for study path 1

It is also found that the prediction errors changed over different rainfall categories and levels of service. Regarding different rainfall categories, the categorization is based on historical rainfall intensity data, and the boundaries for distinguishing light rain, moderate rain, and heavy rain are 0.8mm/h and 6.5mm/h, which is similar to the boundaries set by the previous studies on the effects of rainfall on speed and capacity reduction (Lam et al., 2013). Besides, it has been indicated that travel demand may be affected by the traffic environment under different weather conditions (Lam et al., 2008). Similarly, the predicted path travel times can also be influenced by different levels of service under various rainfall categories.

Weather conditions, particularly the distinction between rainy and non-rainy conditions, significantly impact on the predicted path travel times. Subsequently, traffic conditions, assessed in terms of levels of service, also play a crucial role in affecting these travel times. Moreover, initiation and termination of rainfall may cause significant impacts on predicted path travel times.

Therefore, four scenarios are investigated for predicting path travel times of study path 1 at time interval t + 1 depending on the rainy or non-rainy condition during two successive time intervals t and t + 1. They are: (a) no rain at time interval t, rain at 5-34 time interval t + 1; (b) rain at time interval t, no rain at time interval t + 1; (c) rain at time intervals t and t + 1; (d) no rain at time intervals t and t + 1. Table 5.10 provides detailed information on four scenarios.

Scenarios	Time in	nterval t	Time interval $t + 1$		
	Rain	No rain	Rain	No rain	
(a)		\checkmark	\checkmark		
(b)	\checkmark			\checkmark	
(c)	\checkmark		\checkmark		
(d)		\checkmark		\checkmark	

Table 5.10 Detailed information of four scenarios

Data combination of c1 (using AVI data, rainfall intensity data, and real-time weather forecasts as input) are used for four scenarios. Table 5.11 gives the results of four scenarios regarding the accuracy of predicted path travel times (in terms of mean relative percentage error (MRPE)).

Table 5.11 Prediction performance of the proposed modeling framework on study path 1 at the time interval t + 1 using c1. (in terms of MRPE (%))

	Weather	condition at	Traffic condition			
Scenario	Time	Time interval	I OS A&B	$I \cap S \cap \mathcal{E}$ D	LOS E&F	
	interval <i>t</i>	t + 1	LOS ACD	LOS COLD		
(a)	No rain	Rain	3.7	2.1	4.3	
(b)	Rain	No rain	6.1	5.3	6.8	
(c)	Rain	Rain Rain		3.4	5.2	
(d)	No rain	No rain	5.3	4.5	5.9	

First, it is noted in Table 5.11 that all MRPEs are positive, which indicates that the

proposed modeling framework overestimates the path travel times than the actual path travel times from SMPS. It is because the consideration of $\Sigma_d^{s_a}(t_a, t_b)$ and $\Sigma_t^{s_a}(d_i, d_j)$ may contain some intervals that are less relevant to study time intervals. Consequently, the usage of abundant data may cause overfitting problems in the path travel time prediction (Ricard et al., 2022). Therefore, the results are overestimated. Second, the MRPE is smallest (2.1%) under LOS C&D for Scenario (a) and largest (6.8%) under LOS E&F for Scenario (b). The differences in the overestimation results imply that it is worthwhile investigating the impact of rainfall on driver behavior in further study.

Further sensitivity tests with pre-determined and fixed sample sizes are also conducted to assess the effects of traffic conditions on the performance of predicted path travel times, as investigated in Section 3.4. For each scenario with different weather conditions at two successive 2-minute intervals, the sample size of real-time AVI data is reduced to that under LOS A&B (i.e. smallest sample size under uncongested condition). The results are provided in Table 5.12.

It can be seen in Table 5.12 that when the weather conditions at two successive 2minute intervals t and t + 1 are fixed, predicted path travel times under LOS A&B have the smallest MRPE (3.7%). Hence, it is easier to predict path travel times under LOS A&B, which is consistent with the findings shown in Table 3.7 (due to the relatively low variability of path travel times under LOS A&B) in Section 3.4.

	Weather condition at		Traffic condition			
Scenario	Time	Time	LOS A&B	LOS C&D	LOS E&F	
	interval t	interval t+1	20011002	202 0002		
(a)	No rain	Rain	3.7	4.8	6.1	
(b)	Rain	No rain	6.1	7.6	8.8	
(c)	Rain	Rain	4.1	5.8	6.5	
(d)	No rain	No rain	5.3	7.3	6.8	

 Table 5.12 The prediction performance of the proposed modeling framework with a fixed sample size of real-time AVI data

The preceding experiments presupposed the availability of POP and FRA data every 2 minutes. However, for daily or weekly forecasts, the frequency may not be as frequent as assumed in this study, with hourly POP/FRA data being the norm.

To address this discrepancy, three strategies have been implemented to increase the frequency of POP/FRA data: first, the standard practice of evenly distributing hourly POP/FRA data across minutely intervals was adopted, a common approach when rainfall intensity data is less frequent than desired (Jia et al., 2017). Second, linear interpolation was applied to the hourly POP/FRA data to trace the rainfall amount trend between two consecutive hourly intervals. Third, machine learning algorithms were employed to reveal the underlying correlations (Shi et al., 2022; Wang et al., 2018; Yang et al., 2018) between hourly and minutely FRA, offering a more detailed insight into the effects of various rainfall categories.

Consequently, it is pertinent to examine the performance of the proposed modeling framework for predicting path travel times 1-hour ahead when only hourly FRA data is accessible. Table 5.13 presents the framework's performance in predicting 1-day ahead path travel times using hourly FRA and POP data. It is observed in Table 5.13 that the poorest performance on predicted path travel times occurs under the 5-37

assumption of a uniform distribution of hourly FRA/POP data, with a MAPE of 17.1%, which still remains within an acceptable range of 20% error. This result supports the adaptability of the proposed modeling framework for varying frequencies of weather forecast data.

Stratagiog	MAPE (%)	MAE (min)	RMSE	MaxAPE	MaxAE
Strategies			(min)	(%)	(min)
1. Uniform	17.1%	3.2	3.5	30.1	5.6
2. Linear	16.3%	3.1	3.1	28.6	5.5
3. Machine	12 10/	2.5	2.6	20.4	4.1
learning	15.1%				

Table 5.13 Performance of the proposed modeling framework for predicting path travel times 1-hour ahead using hourly FRA and POP data with different strategies

The modes on effects of rainfall on path travel times are defined in Table 5.14. With this reference, Table 5.15 displays the prediction performance of the proposed modeling framework under different rainfall categories and levels of service for both study paths. It is observed in Table 5.15 that the proposed modeling framework performs better during heavy rain than light and moderate rain while having the highest prediction accuracy in the dry condition.

It is also noted in Table 5.15 that the proposed modeling framework has smallest errors under unstable operation or congestion (MAPE=5.9% under dry condition) and free-flow/unimpeded (MAPE=5.8% under dry condition) for study paths 1 and 2, respectively. The largest errors occur under stable or approaching unstable operation under moderate rain (MAPE=9.4% for study path 1 and MAPE=8.9% for study path 2) for predicting path travel times 1-hour ahead.

	Level of service			
Rainfall categories	A&B (Free- flow /unimpeded) $\delta_{LOS_i,p} \leq 149\%$	C&D (Stable /approaching unstable) $149\% < \delta_{LOS_i,p} \leq 250\%$	E&F (Unstable/congestion) $\delta_{LOS_i,p} > 250\%$	
Dry	D1	D2	D3	
Light rain	L1	L2	L3	
Moderate rain	M1	M2	M3	
Heavy rain	H1	H2	H3	

Table 5.14 Different modes on effects of rainfall on path travel times

Table 5.15 Online updating performance of proposed modeling framework under different rainfall categories and levels of service for study paths 1 and 2

	Study path 1			Study path 2		
Modes	MAPE	MAE	RMSE	MAPE	MAE	RMSE
	(%)	(min)	(min)	(%)	(min)	(min)
D1	6.1	0.5	0.7	5.8	0.4	0.5
D2	7.5	0.7	1.0	7.1	0.8	0.9
D3	5.9	0.5	0.6	6.1	0.5	0.6
L1	7.4	0.7	1.0	6.4	0.6	0.7
L2	8.6	1.0	1.3	8.2	1.4	1.5
L2	7.3	0.7	1	6.8	0.7	0.8
M1	7.6	0.7	1.1	6.8	0.7	0.9
M2	9.4	1.6	1.7	8.9	1.5	1.6
M3	7.8	0.7	1.1	7.2	1	1.1
H1	7.2	0.7	0.9	6.2	0.5	0.7
H2	8.1	0.8	1.2	7.5	1.2	1.3
Н3	6.9	0.7	1.1	6.5	0.6	0.7

The limited AVI data (from ALPR) are used as the input in the previous tests. The

filtering of AVI data described in Li et al. (2023) and Li et al. (2024) has been adopted. To test the effectiveness of the filtering algorithms, both $y_{i,d,p}^{s_A,k}$ and $y_{i,d,p}^{s_G,k}$ are used as input for predicting path travel times 1-hour ahead. Figure 5.9 presents the comparison of CDF plots for study path 1. It is observed in Figure 5.9 that over 90% of predicted path travel times have less than 16.7% of absolute percentage errors provided by the proposed modeling framework. It is satisfactory without using SMPS data (ground truth) as input, which verifies the applicability of the proposed modeling framework.



Figure 5.9 CDF of MAPE for predicting 1-hour ahead on path travel times using different sources of traffic data as input for study path 1

5.5 Discussions

A novel two-stage modeling framework is proposed for the prediction of path travel times in the near future using both the weather forecast and historical rainfall intensity data. Traditionally, only the weather data (mainly rainfall intensity data) is used as an explanatory variable or feature in the traffic prediction models. However, weather forecasts at the current time interval can be a good supplement for predicting path travel times.

Both FRA and POP are utilized and integrated into the proposed modeling framework. The offline training stage uses historical rainfall intensity data to provide offline predicted path travel times. Afterwards, the online updating stage performs the prediction of path travel times with consideration of weather forecast correctness and different levels of service.

The proposed modeling framework is validated in the empirical tests with independent data collected on two selected major expressways in Hong Kong. First, the results of c_5 (path travel times with both weather forecasts and rainfall intensity data as input) outperform that of c_1 (only path travel times as input) significantly for study path 1, with an accuracy improvement of 55% as shown in Table 5.3. It shows the contribution of using both weather forecasts and historical rainfall intensity data that would notably enhance the accuracy of predicted path travel times.

Second, it is found in Figure 5.7 that the normalized cross-correlation coefficient between weather forecast/historical rainfall intensity and path travel times is distinct by different levels of service. The values of the normalized cross-correlation coefficient are larger when traffic is more congested. With this consideration, the proposed online updating stage surpasses other online updating models.

As the weather forecast correctness is of importance to prediction accuracy, the sensitivity test on weather forecast correctness has been conducted. It is observed in Figure 5.8 that a reduction of 40% on weather forecast correctness can increase the MAPE from 6.9% to 22.3% during online updating stage. It provides empirical

evidence to illustrate the significant impacts of weather forecast correctness on prediction accuracy.

Third, the offline training stage of the proposed modeling framework yields the best performance (MAPE=12.8% and MaxAPE=19.2% for study path 1) compared with other offline training models in Tables 5.6 and 5.7. The offline training stage considers the normalized cross-correlation coefficient between path travel times and historical rainfall intensity data by different levels of service. Hence, it has the best quality of offline predicted path travel times, with the usage of historical rainfall intensity data. Furthermore, the sensitivity test on the frequency of rainfall intensity data reveals the merits of using high-frequency rainfall intensity data. The MAPE of predicted path travel times stage reduces from 21.7% to 12.8% in Table 5.8 when the frequency of rainfall intensity data is changed from 1 hour to 2 minutes.

Fourth, the proposed modeling framework performs the best under unstable operation or congestion and dry condition for study path 1 (mode D3 with MAPE=5.9%) and free-flow/unimpeded and dry condition for study path 2 (mode D1 with MAPE=5.8%) and has largest errors under stable or approaching unstable operation (mode M2 with MAPE=9.4% for study path 1 and MAPE=8.9% for study path 2) for predicting path travel times 1-hour ahead in Table 5.15. It indicates that the accuracy of predicted path travel times varies under different rainfall categories and levels of service. It motivates further study on the improvement of predicted path travel times under each mode.

Over 90% of predicted path travel times has less than 16.7% of absolute percentage errors, without using SMPS data as input for study path 1 in Figure 5.9. Hence, the applicability of proposed modeling framework can be verified with and without using ground truth of path travel times as input for the proposed modeling framework.

To further improve the accuracy of the prediction results under non-recurrent conditions, other data sources, such as accident data, should be considered. Traffic accidents on the study path can increase the path travel times significantly (Mil and Piantanakulchai, 2018; Ma et al., 2019; Zhong et al., 2020). To investigate the impact of traffic accidents on the filtering performance for path travel time prediction, AVI data on days with accidents in 2018 are further analyzed. For example, the path travel time during 9:00-10:00 on March 18th, 2018 (Tuesday, with accident data) is compared with the same time period on previous Tuesdays.

It is found that the increment of path travel times is 5.3% after filtering. Regarding outliers filtered out by the proposed unsupervised algorithm, they may be affected by accidents or detours of vehicles. These outliers need to be further distinguished with detailed trajectories of accident vehicles in further study. Furthermore, as the impacts of rainfall can be various for different road types (Zhang et al., 2018), it is worthwhile to identify their distinguished effects by different road types in the further study.

Additionally, travel behaviors (such as departure time and route choices) can change after receiving the predicted path travel times. The corresponding impacts would adversely adjust the path travel times under each scenario with the specific rainfall category and level of service. They should be modeled in further study. Moreover, the modeling framework proposed in this chapter only uses AVI data. Multi-source traffic data could be integrated to predict path travel times in future time intervals by vehicle class, as presented in Chapter 4. The link travel times along the study path can also be predicted in future studies.

6. Conclusions and Further Studies

This thesis conducts an in-depth exploration of the instantaneous path travel time prediction problems, utilizing heterogeneous traffic data and weather information (including historical weather data and updated weather forecasts). It primarily encompasses three key areas of study. The first area of focus is the filtering of AVI data, particularly when the ground truth is always unavailable for training purposes in practice. The proposed unsupervised algorithm is impactful as it offers a novel perspective on data filtration, confronting the challenge of lacking ground truth for training (as detailed in Section 3.1.1).

The second area of study involves multi-class path travel time prediction. This research area is crucial as the existing advanced traveler information systems (ATIS) only provide the average path travel times of all vehicles. However, the path travel times of certain vehicle types (e.g., private cars), which account for a significant proportion, may differ greatly from the average path travel times of all vehicles (see Figure 4.3 for detailed illustration) under some circumstances. The proposed prediction model fills the research gap by providing a comprehensive understanding of multi-class path travel times.

The third area of research in this thesis is the prediction of path travel times in the near future, using weather forecast data to update the offline predicted path travel times. This research area is specifically pertinent given the increasing impacts of weather information in path travel time prediction. It is because the path travel times are affected by rainfall from both demand (travel behavior of road users) and supply (road capacity and free flow speed/speed at capacity) aspects (as explained in Section 5.1.1).

For cities with frequent rainfall, weather forecasts can be beneficial for predicting path

travel times. The weather forecasts on rainfall considered in this study can further contribute to improving the prediction accuracy of path travel time prediction in future time intervals.

6.1 Highlights of Main Contributions of the Thesis

The highlights of main contributions of Chapter 3 are listed below:

In connection with C3.1 in Section 3.1.2, in order to effectively filter limited real-time AVI data, the proposed unsupervised algorithm uses collected historical AVI data as supplementary information. It helps to improve the filtering performance of limited real-time AVI data without using ground truth for training purposes.

Regarding C3.2 in Section 3.1.2, a comprehensive dynamic validity window should be determined by not only the within-day covariance of path travel times but also the day-to-day covariance of path travel times. Hence, historical AVI data offers substantial contributory assistance in acquiring the latter. The proposed unsupervised algorithm adapts FPCA to capture both with-day and day-to-day covariance of path travel times (as illustrated in Sections 3.3.2 and 3.3.3). Both the mean and standard deviation of the path travel times are predicted and used to construct the dynamic validity window. It has been validated to outperform the existing filtering algorithms in filtering the limited real-time AVI data.

In relation to C3.3 in Section 3.1.2, the applicability of the proposed unsupervised algorithm is examined with different sample sizes of real-time AVI data with different AVI sensors along different paths in urban areas. Chapter 3 also studies the effects of different sampling rates of the real-time AVI data on the performance of data filtering in Section 3.4.4. They are used to verify the robustness of the proposed unsupervised

algorithm.

The highlights of three major contributions in Chapter 4 are presented in the following:

With respect to C4.1 in Section 4.1.2, the proposed prediction model predicts path travel times by different vehicle classes with satisfactory performance. The proposed prediction model makes proper use of multi-source traffic data from AVI, GPS, and point sensors to solve this problem.

In light of C4.2 in Section 4.1.2, with multi-source traffic data, the proposed prediction model extends to consider the temporal covariance of path travel time between vehicle classes (modeled in Section 4.3). This covariance can help enhance the prediction accuracy when real-time traffic data is insufficient for path travel time prediction of a specific vehicle class.

With reference to C4.3 in Section 4.1.2, robustness is a key indicator for accessing the path travel time prediction model. Experiments are conducted to verify the robustness of the proposed prediction model using data collected from a major expressway in Hong Kong. The satisfactory results demonstrate its robustness (in Section 4.4.2).

The following paragraphs summarize the highlights of main contributions of Chapter 5:

Regarding C5.1 in Section 5.1.2, in order to improve the prediction of path travel times, this study extends to investigating real-time weather forecasts and historical rainfall intensity data. The latter is used to provide the offline predicted path travel times before updating with weather forecasts and real-time traffic data. Both online updating and offline prediction stages are considered in the proposed modeling framework

(introduced in Section 5.1.1).

In connection with C5.2 in Section 5.1.2, the impacts of rainfall on path travel times can be distinct under different rainfall categories and traffic conditions. In the online updating stage, these boundary conditions (or modes in section and/or equation numbers) distinguished impacts are modeled in the modified Kalman filter (in Section 5.3.2). It models the normalized cross-correlation coefficient between real-time weather forecasts and predicted path travel times under different rainfall categories and traffic conditions. The effects of weather forecast correctness on prediction accuracy are also examined.

With respect to C5.3 in Section 5.1.2, a higher frequency of rainfall intensity data (say, once every 2 minutes) helps to capture the detailed impacts of rainfall on path travel time prediction, as claimed in C5.2. The offline prediction stage captures the normalized cross-correlation coefficient between path travel times and rainfall intensity data. With this information, it can predict offline path travel times with higher quality. This normalized cross-correlation coefficient is tested under different levels of service in Section 5.4.2. The merits of using high-frequency rainfall intensity data are also illustrated in the numerical examples (see Table 5.8).

In regards to C5.4 in Section 5.1.2, the proposed modeling framework is examined on the real-world dataset in Hong Kong. It presents the impacts on the prediction accuracy by different rainfall categories and levels of service (traffic conditions). The applicability of the proposed modeling framework is also demonstrated with different types of input (with and without ground truth on path travel times).

6.2 Discussions on Key Findings

The three major findings for each of Chapters 3, 4, and 5 are summarized respectively with reference to their corresponding key contributions:

F3.1: In relation to C3.1, the proposed unsupervised algorithm (U1) is compared with other existing filtering algorithms. It is found in Figure 3.8 that U1 has 83% of absolute percentage errors of predicted results less than 20%, while the statistic varies from 56% to 60% for other benchmarks. It demonstrates the advantages of using historical AVI data but without the ground truth.

F3.2: In connection with C3.2, if the partial or complete set of ground truth can be obtained by surveys, U1 can be extended to a supervised algorithm with ground truth for training (S1). However, it is observed in Table 3.9 that the percentage of absolute percentage errors less than 20% is reduced to 83% or lower if less than 50% of the historical ground truth is used for training purposes. It suggests that when there is inadequate historical ground truth collected (i.e., less than 50%), U1 is still better than S1 in practice for filtering real-time AVI data and path travel time prediction.

F3.3: With reference to C3.3, it is noted in Figure 3.9 that when the sampling rate of real-time AVI data within the validity window is no less than two valid AVI data per 2-minute interval, U1 is tested to be robust on different paths (15.2% and 14.9% of 95 percentile of absolute percentage errors for study paths 1 and 2, respectively). It verifies the robustness of U1.

F4.1: The proposed prediction model outperforms other benchmark prediction models in terms of MAPE of overall predicted path travel time (i.e., 6.9% of the proposed prediction model against 9.1%-14.1% of others) in Table 4.7.

F4.2: The use of muti-source traffic data is proved to improve the prediction accuracy in terms of MAPE of overall predicted path travel time (i.e., 6.9% of using AVI and point sensor data against 12.8% of using GPS data) in Table 4.5.

F4.3: The trade-off of using different traffic data has been investigated. The accuracy of predicted path travel time is more sensitive to AVI or GPS data than point sensor data, as shown in Table 4.10. Compared with point sensor data, the cost of collecting AVI or GPS data is larger. However, they contribute more to the prediction accuracy. Therefore, it is a trade-off between cost and accuracy for planners to allocate traffic sensors on the road network.

F5.1: The prediction accuracy is significantly enhanced after considering weather forecast and rainfall intensity data in Hong Kong (i.e., MAPE for predicted path travel times for 1-hour ahead reduced from 15.2% to 6.9%) in Table 5.3.

F5.2: The usage of high-frequency rainfall intensity data improves the prediction accuracy in terms of MAPE for predicted path travel times for 1-hour ahead (i.e., 12.8% for 2-minute rainfall intensity data against 21.7% for hourly rainfall intensity data in the offline prediction stage) in Table 5.8.

F5.3: The applicability of the proposed modeling framework has been verified. It is observed in Figure 5.9 that over 90% of predicted path travel times has less than 16.7% of absolute percentage errors without input of the SMPS data as the ground truth.

6.3 Recommendations for Future Research

In Chapter 3, the ground truth is not used in the proposed unsupervised algorithm for

training as it is usually unavailable in practice. If it can be obtained partially by observation surveys, it is worthwhile developing a strategy for collecting the representative ground truth with limited budgets. Moreover, this chapter only considered the limited, highly accurate AVI data for predicting path travel times. However, in reality, there are multiple traffic sensors deployed on road networks, providing various travel time data that should be utilized in the future. It is also interesting to explore the sensor-location problems and trade-offs of different types of traffic sensors in the multi-modal transport networks.

Additionally, other AVI data with larger sample sizes but lower accuracy (e.g., Bluetooth data) could be further investigated. Furthermore, other types of trafficrelated data, including the built environment, other weather information (e.g., wind speed), traffic accidents, construction works, vehicular flow data, bus frequencies, signal timing, and road types, etc., could be integrated into the proposed unsupervised algorithm to improve the performance of data filtering.

Likewise, it is interesting to expand the proposed unsupervised algorithm to study the impact of sensor failures on data from multiple AVI sensors in urban road corridors, taking into account network topology and measurement errors.

In Chapter 4, there is potential for the exploration of larger-scale road networks with diverse types of traffic data sources. It is also recommended to consider additional data sources such as weather and accident data. Besides, the application of more advanced prediction models for preprocessing multiple data sources could significantly enhance the accuracy of predictions. It is also crucial to develop an efficient solution method that enables the prediction of multi-class travel times on various road links along different paths within the same network.

In Chapter 5, it is recommended to retain historical weather forecasts for validation of their impact on path travel time prediction purposes. It is also advisable to incorporate additional traffic data sources and non-traffic data, such as incident data, into the proposed modeling framework. Furthermore, given the varying impacts of rainfall on different road types (Zhang et al., 2018), it would be valuable to investigate and differentiate their effects in further studies.

Moreover, the empirical studies presented in Chapters 3, 4, and 5 are based on traffic data collected in Hong Kong. When contrasted with the rich data amount available in other cities or countries such as Singapore, the accessibility of traffic data in Hong Kong is however constrained by privacy concerns. Consequently, the sample size of the traffic data for ATIS development in Hong Kong is notably smaller compared to datasets from other places. It would be insightful to conduct future research that benchmarks the Hong Kong dataset with those from other cities with rich data amounts for path travel time prediction.

Additionally, the proposed modeling framework could be further extended to facilitate the prediction of path travel times for different vehicle classes. Moreover, travel behaviors (such as departure time and route choices) can also be affected by the predicted path travel times. The interactions between predicted path travel times and travel behaviors under different weather and traffic conditions should be further studied.
Appendix A

Typical ATIS across the World

In this appendix, three typical ATIS are presented, including 511 Travel Information (USA), Google Maps (Google company), and JTIS (Hong Kong), as indicated in Table 2.1. It is worthwhile mentioning that all these ATIS provide instantaneous travel time. Overall, they identify a critical need for accurate predictions of path travel times in both current and future time intervals within ATIS. It serves to enhance route guidance and provides a comprehensive overview of traffic conditions. This requirement forms a primary motivation for the research presented in this thesis.

Furthermore, it is noted that existing ATIS predominantly offer predicted average path travel times, revealing a research gap in the predicting multi-class path travel times. This gap will be extensively discussed in Section 4.1.1. Additionally, the availability of weather information in certain ATIS (e.g., the 511 Travel Information system in the USA) suggests a potential avenue for enhancing the accuracy of path travel time predictions. The integration of weather data and its implications for improving the performance of travel time predictions will be thoroughly examined in Section 5.1.1.

Figure A.1 gives the interface of 511 Travel Information. It is a nationwide program in USA. It is seen in Figure A.1 that both path distance and predicted path travel times in the current time interval are given. Detailed route guidance based on predicted path travel times is also available. It confirms that the predicted path travel time plays an important role in ATIS.



Figure A.1 511 Travel Information (USA) https://ops.fhwa.dot.gov/travelinfo/about/about511.htm

Apart from path travel time information, both rainfall intensity data and weather forecast (in terms of weather conditions for 1-week ahead) are given by 511 Travel Information in Figure A.1. As adverse weather can influence travel behavior, weather information should be fully investigated for improving the performance of predicted path travel times. It motivates the study in Chapter 5.

Figure A.2 provides two cases using Google Maps. It presents the path distance (5.3 km in this example). Besides, it will provide the predicted path travel times in the current travel time if the option is to leave now (e.g., 10 minutes in the example). It will display predicted path travel times in future time intervals if the option is scheduling a trip in the future (e.g., 8-18 minutes, shown in Figure A.2). It is concluded

that predicted path travel times in both current and future time intervals are needed in ATIS, which motives this thesis.



Figure A.2 Google Maps (Google company) https://www.google.com/maps/

de A.3 shows the predicted path travel times in the current time intervals of major routes provided by JTIS. They are updated every 2 minutes. Details of JTIS are presented in Section 3.4.1.



Figure A.3 JTIS (Hong Kong)

https://www.td.gov.hk/en/transport_in_hong_kong/its/its_achievements/journey_time __indication_system_/index.html

Appendix B

Photographs and Weblinks on Various Traffic Sensors

This appendix is an extension of Tables 2.5 and 2.6. Table B.1 further shows photographs and weblinks for different AVI and point sensors. With the use of various traffic sensors, the traffic data are heterogeneous and hence need to be integrated effectively, as illustrated in detail in Section 4.1.

Tra	affic sensors	Photographs	Weblinks
	RFID tag readers		https://www.hk- rfid.com/passive-rfid- reader
	ALPR cameras		https://sls.eff.org/techno logies/automated- license-plate-readers- alprs
AVI sensors	Bluetooth MAC address readers		https://www.elefinetech .com/long-range-rfid- bluetooth-reader-with- stickers-for-parking- access-control-system/
	Infrared sensors		https://blog.iceslicer.co m/road-surface- temperature-sensors

Table B.1 Photographs and weblinks for different AVI and point sensors

AVI	Barcode scanners	https://www.barcodehq. com/vinbarcode.html
sensors	DSRC sensors	https://www.grs.com.nl/ dsrc-remote- tachograph-monitoring- systems
Point	Single-loop sensors	https://www.proconel.c om/products/single- channel-loop-detectors- parking/
sensors	Double-loop sensors	https://www.nobleled.c om/vehicl-loop- detector/dual-channel- loop- detector/NBLD206.htm l

	Acoustic sensors		https://hackernoon.com/ survey-on-acoustic- sensors-in-self-driving- cars
Point sensors	RTMS		http://www.invisiblebox es.info/remote-traffic- microwave-radar/
	Video-based cameras	E BLUFFCREEK	https://www.econolite.c om/solutions/sensors/au toscope/

Appendix C

Sample Data Format of Traffic Data

The sample data formats for AVI data, point sensor data, GPS data, and JTIS/SMPS data are given in this appendix. Though there are two sources of AVI data including ALPR cameras (used in Chapters 4 and 5) and RFID tag readers (adopted in Chapters 3 and 5), the data format for AVI data is the same.

It is seen in Figure C.1 that timestamps of vehicles at origin and destination are available. The collected AVI data has no vehicle ID due to privacy issues in Hong Kong. More details on privacy issues in Hong Kong can be found in Section 3.1.1. Contrastingly, the observed path travel time (in the unit of seconds) is given directly in the AVI data (Column 6). It is obtained by the difference between arrival time (Column 2) and departure time (Column 1). For example, the first record in Figure C.1 provides the observed path travel times of 1514s, which is the difference between two timestamps, 2018/1/1 0:00 and 2017/12/31 23:34.

Besides, the vehicle class ID presents the vehicle class information in Column 7 in Figure C.1. For example, the vehicle class ID for the first record is 7, representing it as a public light bus. Vehicle class information is to be used in Section 4.4 for predicting path travel times by vehicle class.

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7
RT_FROM_TIME	RT_TO_TIME	RT_FROM_LOC	RT_TO_LOC	RT_OD_PAIR	RT_TRAVEL_TIME	RT_VEHICLE_CLASS
2017/12/31 23:34	2018/1/1 0:00	TMR3LSD1Z1	TMR14LSD2Z1	TMR3-TMR14	1514	6
2018/1/1 0:01	2018/1/1 0:05	TMR19LSD2Z1	TMR23ALSD1Z2	TMR19-TMR23A	255	6
2017/12/31 23:46	2018/1/1 0:04	TMR3LSD1Z2	TMR14LSD2Z1	TMR3-TMR14	1060	7
↑	1	Ť	Ť	Ť	1	↑
Timestamp of departure	Timestamp of arrival	AVI sensor ID at origin	AVI sensor ID at destination	Path ID	Observed path travel time	Vehicle class ID

Figure C.1 Data format for AVI data

Figure C.2 shows the data format for point sensor data. The record is available once every 2 minutes. The sample size of detected vehicles by point sensors (video-based cameras) is given. Both time mean speed and space mean speed of vehicles in the past 2-minute interval are available (both in the unit of km/h). For example, the 2-minute interval of the first record in the following photograph is 00:00:10-00:02:10 on January 1, 2018 (Column 5). There are 10 vehicles captured by point sensors (Column 4). The arithmetic mean of spot speeds of these 10 is 99 km/h (time mean speed). As they are originally used to provide the link travel time, they are converted to space mean speed, which is 98.2 km/h (Column 3). As point sensor data are stored at an aggregated level, the individual speed of each vehicle is unavailable in the dataset.

Column 1	Column 2	Column 3	Column 4	Column 5
LINK_ID	SPEED	SPACE_SPEED	SAMPLE	TIMEADDED
991048-992139	99	98.2055351	10	2018-01-01 00:02:10
991048-992139	96.75	95.70560107	16	2018-01-01 00:04:10
991048-992139	98.875	98.18206541	8	2018-01-01 00:06:10
Point sensor ID	Time mean speed	Space mean speed	Sample si detected ve per time in	ze of Time interval hicles terval

Figure C.2 Data format for point sensor data

In Figure C.3, each record of GPS data contains the vehicle ID (Column 1), position of the vehicle (Columns 2 and 3), spot speed of the vehicle in units of km/h (Column 4), as well as the timestamp of this record (Column 5). For example, the first record measures vehicle 90542 (Column 1) has a spot speed of 13 km/h at location (815727.1, 828082.2) with a timestamp of 00:00:22 May 2, 2018. Moreover, there are no direct relationships between two consecutive records in GPS data. For example, two records in Figure C.3 have different vehicle IDs (90542 and 111757). Thus, the map-matching and reidentification are necessary for GPS data. As GPS data used in Chapter 4 are

purchased from commercial companies and captured from goods vehicles, the vehicle class information for these GPS data is also available.

VehicleID	X	Y	Speed	Clock
90542	815727.1	828082.2	13	02-05-2018 00:00:22
111757	825050.3	825180.4	72	02-05-2018 00:00:26
Ť	Ť	↑		Ť

Figure C.3 Data format for GPS data

Predicted path travel times provided by either JTIS or SMPS are regarded as ground truth in this thesis. Similar to point sensor data, they provide predicted path travel times (in units of minutes) in the current time interval every 2 minutes. After identification of the studied path, these predicted path travel times can be used directly for validation in Section 3.4, Section 4.4, and Section 5.4. Figure C.4 gives the corresponding data format. Take the first record in Figure C.4 as an example; for a path with id SJ5-TWTM (Column 1), the predicted path travel time for the current time interval 0:00-0:02 January 1, 2018 (Column 3) is 13.87 minutes.



Figure C.4 Data format for JTIS/SMPS data

References

- Ahmed, M.M., Abdel-Aty, M.A., 2012. The viability of using automatic vehicle identification data for real-time crash prediction. *IEEE Trans. Intell. Transp. Syst.* 13, 459–468.
- Anusha, S.P., Anand, R.A., Vanajakshi, L., 2012. Data Fusion Based Hybrid Approach for the Estimation of Urban Arterial Travel Time. *J. Appl. Math.* 2012, 1–17.
- Asqool, O., Koting, S., Saifizul, A., 2021. Evaluation of Outlier Filtering Algorithms for Accurate Travel Time Measurement Incorporating Lane-Splitting Situations. *Sustainability* 13, 13851.
- Ban, X.J., Li, Y., Skabardonis, A., Margulici, J.D., 2010. Performance evaluation of travel-time estimation methods for real-time traffic applications. *J. Intell. Transp. Syst. Technol. Planning, Oper.* 14, 54–67.
- Casabianca, P., Zhang, Y., Martínez-García, M., Wan, J., 2021. Vehicle Destination Prediction Using Bidirectional LSTM with Attention Mechanism. *Sensors* 21, 8443.
- Celikoglu, H.B., 2013a. Reconstructing freeway travel times with a simplified network flow model alternating the adopted fundamental diagram. *Eur. J. Oper. Res.* 228, 457–466.
- Celikoglu, H.B., 2013b. Flow-based freeway travel-time estimation: A comparative evaluation within dynamic path loading. *IEEE Trans. Intell. Transp. Syst.* 14, 772–781.
- Chakraborty, P., Hegde, C., Sharma, A., 2019. Data-driven parallelizable traffic incident detection using spatio-temporally denoised robust thresholds. *Transp. Res. Part C Emerg. Technol.* 105, 81–99.
- Chalumuri, R.S., Yasuo, A., 2014. Modelling travel time distribution under various uncertainties on hanshin expressway of Japan. *Eur. Transp. Res. Rev.* 6, 85–92.
- Chan, K., Lam, W., Tam, M., 2009. Real-Time Estimation of Arterial Travel Times

with Spatial Travel Time Covariance Relationships. *Transp. Res. Rec. J. Transp. Res. Board* 2121, 102–109.

- Chen, C.-M., Liang, C.-C., Chu, C.-P., 2020. Long-term travel time prediction using gradient boosting. *J. Intell. Transp. Syst.* 24, 109–124.
- Chen, H., Rakha, H.A., 2014. Real-time travel time prediction using particle filtering with a non-explicit state-transition model. *Transp. Res. Part C Emerg. Technol.* 43, 112–126.
- Chen, K., Müller, H.G., 2014. Modeling conditional distributions for functional responses, with application to traffic monitoring via GPS-enabled mobile phones. *Technometrics* 56, 347–358.
- Chen, M., Yu, G., Chen, P., Wang, Y., 2017. A copula-based approach for estimating the travel time reliability of urban arterial. *Transp. Res. Part C Emerg. Technol.* 82, 1–23.
- Chen, M.Y., Chiang, H. Sen, Yang, K.J., 2022. Constructing Cooperative Intelligent Transport Systems for Travel Time Prediction With Deep Learning Approaches. *IEEE Trans. Intell. Transp. Syst.* 23, 16590–16599.
- Chen, S., Wang, W., Zuylen, H. Van, 2010. A comparison of outlier detection algorithms for ITS data. *Expert Syst. Appl.* 37, 1169–1178.
- Chen, Y., Chen, C., Wu, Q., Ma, J., Zhang, G., Milton, J., 2021. Spatial-temporal traffic congestion identification and correlation extraction using floating car data. *J. Intell. Transp. Syst.* 25, 263–280.
- Cherrett, T., McLeod, F., Bell, H., McDonald, M., 2002. Journey time estimation using single inductive loop detectors on non-signalised links. J. Oper. Res. Soc. 53, 610–619.
- Chiou, J.M., 2012. Dynamical functional prediction and classification, with application to traffic flow prediction. *Ann. Appl. Stat.* 6, 1588–1614.
- Chiou, J.M., Liou, H.T., Chen, W.H., 2021. Modeling Time-Varying Variability and Reliability of Freeway Travel Time Using Functional Principal Component

Analysis. IEEE Trans. Intell. Transp. Syst. 22, 257-266.

- Chow, A.H.F., Santacreu, A., Tsapakis, I., Tanasaranond, G., Cheng, T., 2014. Empirical assessment of urban traffic congestion. *J. Adv. Transp.* 48, 1000–1016.
- Correa, D., Ozbay, K., 2022. Urban path travel time estimation using GPS trajectories from high-sampling-rate ridesourcing services. J. Intell. Transp. Syst. 28, 267– 282.
- Cortés, C.E., Lavanya, R., Oh, J.-S., Jayakrishnan, R., 2002. General-Purpose Methodology for Estimating Link Travel Time with Multiple-Point Detection of Traffic. *Transp. Res. Rec. J. Transp. Res. Board* 1802, 181–189.
- Deng, W., Lei, H., Zhou, X., 2013. Traffic state estimation and uncertainty quantification based on heterogeneous data sources: A three detector approach. *Transp. Res. Part B Methodol.* 57, 132–157.
- Derrow-Pinion, A., She, J., Wong, D., Lange, O., Hester, T., Perez, L., Nunkesser, M., Lee, S., Guo, X., Wiltshire, B., Battaglia, P.W., Gupta, V., Li, A., Xu, Z., Sanchez-Gonzalez, A., Li, Y., Velickovic, P., 2021. ETA Prediction with Graph Neural Networks in Google Maps. *Int. Conf. Inf. Knowl. Manag. Proc.* 3767– 3776.
- Diaz, J.J.V., Rodriguez Gonzalez, A.B., Wilby, M.R., 2016. Bluetooth Traffic Monitoring Systems for Travel Time Estimation on Freeways. *IEEE Trans. Intell. Transp. Syst.* 17, 123–132.
- Dion, F., Rakha, H., 2006. Estimating dynamic roadway travel times using automatic vehicle identification data for low sampling rates. *Transp. Res. Part B Methodol.* 40, 745–766.
- Du, L., Gao, R., Suganthan, P.N., Wang, D.Z.W., 2022a. Graph ensemble deep random vector functional link network for traffic forecasting. *Appl. Soft Comput.* 131, 109809.
- Du, L., Gao, R., Suganthan, P.N., Wang, D.Z.W., 2022b. Bayesian optimization based dynamic ensemble for time series forecasting. *Inf. Sci. (Ny).* 591, 155–175.

- Duan, P., Mao, G., Kang, J., Huang, B., 2020. Estimation of Link Travel Time Distribution with Limited Traffic Detectors. *IEEE Trans. Intell. Transp. Syst.* 21, 3730–3743.
- Dunne, S., Ghosh, B., 2013. Weather Adaptive Traffic Prediction Using Neurowavelet Models. *IEEE Trans. Intell. Transp. Syst.* 14, 370–379.
- Elfar, A., Xavier, C., Talebpour, A., Mahmassani, H.S., C, C., Talebpour, A., Mahmassani, H.S., Xavier, C., Talebpour, A., Mahmassani, H.S., 2018. Traffic shockwave detection in a connected environment using the speed distribution of individual vehicles. *Transp. Res. Rec.* 2672, 203–214.
- Fu, F., Qian, W., Dong, H., 2019. Estimation of Route Travel Time Distribution with Information Fusion from Automatic Number Plate Recognition Data. J. Transp. Eng. Part A Syst. 145, 1–14.
- Fu, H., Lam, W.H.K., Shao, H., Ma, W., Chen, B.Y., Ho, H.W., 2022. Optimization of multi-type sensor locations for simultaneous estimation of origin-destination demands and link travel times with covariance effects. *Transp. Res. Part B Methodol.* 166, 19–47.
- Gao, R., Sun, F., Xing, W., Tao, D., Fang, J., Chai, H., 2022. CTTE: Customized Travel Time Estimation via Mobile Crowdsensing. *IEEE Trans. Intell. Transp. Syst.* 23, 19335–19347.
- Gazder, U., Ratrout, N.T., 2018. Determining factors affecting cross-border transport on King Fahd Causeway and comparison with urban travel. *Case Stud. Transp. Policy* 6, 737–742.
- Ghavidel, M., Khademi, N., Bahrami Samani, E., Kieu, L.-M., 2022. A Random Effects Model for Travel-Time Variability Analysis Using Wi-Fi and Bluetooth Data. J. Transp. Eng. Part A Syst. 148.
- Gong, Q., Adams, T.M., Wang, X.B., 2015. Estimating link travel time with sparse GPS data on highway corridors. *Transp. Res. Rec.* 2477, 7–17.
- Guardiola, I.G., Leon, T., Mallor, F., 2014. A functional approach to monitor and

recognize patterns of daily traffic profiles. *Transp. Res. Part B Methodol.* 65, 119–136.

- Haghani, A., Hamedi, M., Sadabadi, K.F., Young, S., Tarnoff, P., 2010. Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors. *Transp. Res. Rec.* 2160, 60–68.
- Han, J., Polak, J.W., Barria, J., Krishnan, R., 2010. On the estimation of space-meanspeed from inductive loop detector data. *Transp. Plan. Technol.* 33, 91–104.
- Han, L., Sun, H., Wang, D.Z.W., Zhu, C., 2018. A stochastic process traffic assignment model considering stochastic traffic demand. *Transp. B Transp. Dyn.* 6, 169–189.
- Harper, C.D., Qian, S., Samaras, C., 2021. Improving Short-Term Travel Speed Prediction with High-Resolution Spatial and Temporal Rainfall Data. J. Transp. Eng. Part A Syst. 147, 1–16.
- Hiribarren, G., Herrera, J.C., 2014. Real time traffic states estimation on arterials based on trajectory data. *Transp. Res. Part B Methodol.* 69, 19–30.
- Hofleitner, A., Herring, R., Bayen, A., 2012. Arterial travel time forecast with streaming data: A hybrid approach of flow modeling and machine learning. *Transp. Res. Part B Methodol.* 46, 1097–1122.
- Hong Kong Transport Department, 2023. Journey Time Indication System.
- Hyun, K. (Kate), Tok, A., Ritchie, S.G., 2017. Long distance truck tracking from advanced point detectors using a selective weighted Bayesian model. *Transp. Res. Part C Emerg. Technol.* 82, 24–42.
- Jang, Y., Ku, D., Lee, S., 2023. Pedestrian mode identification, classification and characterization by tracking mobile data. *Transp. A Transp. Sci.* 19, 2008044.
- Ji, H., Müller, H.G., 2017. Optimal designs for longitudinal and functional data. J. R. Stat. Soc. Ser. B Stat. Methodol. 79, 859–876.
- Jia, Y., Wu, J., Ben-Akiva, M., Seshadri, R., Du, Y., 2017. Rainfall-integrated traffic speed prediction using deep learning method. *IET Intell. Transp. Syst.* 11, 531–

536.

- Jiang, Z., Chen, X. (Michael), Ouyang, Y., 2017. Traffic state and emission estimation for urban expressways based on heterogeneous data. *Transp. Res. Part D Transp. Environ.* 53, 440–453.
- Kashinath, S.A., Mostafa, S.A., Mustapha, A., Mahdin, H., Lim, D., Mahmoud, M.A., Mohammed, M.A., Al-Rimy, B.A.S., Fudzee, M.F.M., Yang, T.J., 2021. Review of data fusion methods for real-time and multi-sensor traffic flow analysis. *IEEE Access* 9, 51258–51276.
- Kazagli, E., Koutsopoulos, H.N., 2013. Estimation of Arterial Travel Time from Automatic Number Plate Recognition Data. *Transp. Res. Rec. J. Transp. Res. Board* 2391, 22–31.
- Kebaili Bargaoui, Z., Chebbi, A., 2009. Comparison of two kriging interpolation methods applied to spatiotemporal rainfall. *J. Hydrol.* 365, 56–73.
- Ku, D., Na, S., Kim, J., Lee, S., 2021. Real-time taxi demand prediction using recurrent neural network. *Proc. Inst. Civ. Eng. - Munic. Eng.* 174, 75–87.
- Kwong, K., Kavaler, R., Rajagopal, R., Varaiya, P., 2009. Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors. *Transp. Res. Part C Emerg. Technol.* 17, 586–606.
- Lam, W.H.K., Shao, H., Sumalee, A., 2008. Modeling impacts of adverse weather conditions on a road network with uncertainties in demand and supply. *Transp. Res. Part B Methodol.* 42, 890–910.
- Lam, W.H.K., Tam, M.L., Cao, X., Li, X., 2013. Modeling the Effects of Rainfall Intensity on Traffic Speed, Flow, and Density Relationships for Urban Roads. J. *Transp. Eng.* 139, 758–770.
- Laña, I., Lobo, J.L., Capecci, E., Del Ser, J., Kasabov, N., 2019. Adaptive long-term traffic state estimation with evolving spiking neural networks. *Transp. Res. Part C Emerg. Technol.* 101, 126–144.
- Lee, S., Lee, Y.I., Cho, B., 2006. Short-term travel speed prediction models in car

navigation systems. J. Adv. Transp. 40, 122-139.

- Li, A., Lam, W.H.K., Ma, W., Chow, A.H.F., Wong, S., Tam, M.L., 2024a. Filtering Limited Automatic Vehicle Identification Data for Real-Time Path Travel Time Estimation Without Ground Truth. *IEEE Trans. Intell. Transp. Syst.* 25, 4849– 4861.
- Li, A., Lam, W.H.K., Ma, W., Wong, S.C., Chow, A.H.F., Lam Tam, M., 2024b. Realtime estimation of multi-class path travel times using multi-source traffic data. *Expert Syst. Appl.* 237, 121613.
- Li, H., Wang, Q., Xiong, W., 2021. New Model of Travel-Time Prediction Considering Weather Conditions: Case Study of Urban Expressway. J. Transp. Eng. Part A Syst. 147, 1–8.
- Li, J., Lam, W.H.K., Li, X., 2016. Modeling the Effects of Rainfall Intensity on the Heteroscedastic Traffic Speed Dispersion on Urban Roads. J. Transp. Eng. 142, 1–9.
- Li, L., Su, X., Wang, Y., Lin, Y., Li, Z., Li, Y., 2015. Robust causal dependence mining in big data network and its application to traffic flow predictions. *Transp. Res. Part C Emerg. Technol.* 58, 292–307.
- Li, M., Chen, X. (Michael), Ni, W., 2016. An extended generalized filter algorithm for urban expressway traffic time estimation based on heterogeneous data. J. Intell. Transp. Syst. Technol. Planning, Oper. 20, 474–484.
- Li, Q., Lam, W.H.K., Tam, M.L., 2012. Vehicle Travel Time Prediction in Spatio-Temporal Space. *Appl. Mech. Mater.* 253–255, 1662–1665.
- Li, R., Rose, G., Sarvi, M., 2006. Evaluation of Speed-Based Travel Time Estimation Models. J. Transp. Eng. 132, 540–547.
- Lint, J., Zjipp, N., 2003. An Improved Travel Time Estimation Algorithm Using Dual Loop Detectors. TRB 82nd Annu. Meet.
- Liu, H., Van Zuylen, H.J., Van Lint, H., Chen, Y., Zhang, K., 2005. Prediction of urban travel times with intersection delays, in: IEEE Conference on Intelligent

Transportation Systems, Proceedings, ITSC. IEEE, pp. 1062–1067.

- Lu, L., Wang, J., He, Z., Chan, C.-Y., 2017. Real-time estimation of freeway travel time with recurrent congestion based on sparse detector data. *IET Intell. Transp. Syst.* 12, 2–11.
- Lyu, C., Song, Y., Yang, D., Wang, W., Ge, Y., Wang, L., 2023. Online prediction for heat generation rate and temperature of lithium-ion battery using multi-stepahead extended Kalman filtering. *Appl. Therm. Eng.* 231, 120890.
- Ma, J., Chan, J., Rajasegarar, S., Leckie, C., 2022. Multi-attention graph neural networks for city-wide bus travel time estimation using limited data. *Expert Syst. Appl.* 202, 117057.
- Ma, T., Zhou, Z., Antoniou, C., 2018a. Dynamic factor model for network traffic state forecast. *Transp. Res. Part B Methodol.* 118, 281–317.
- Ma, T., Zhou, Z., Antoniou, C., 2018b. Dynamic factor model for network traffic state forecast. *Transp. Res. Part B Methodol.* 118, 281–317.
- Ma, X., Koutsopoulos, H., 2010. Estimation of the automatic vehicle identification based spatial travel time information collected in Stockholm. *IET Intell. Transp. Syst.* 4, 298–306.
- Ma, Y., Lv, Z., Zhang, P., Chan, C.-Y., 2021. Impact of lane changing on adjacent vehicles considering multi-vehicle interaction in mixed traffic flow: A velocity estimating model. *Phys. A Stat. Mech. its Appl.* 566, 125577.
- Ma, Y., Yin, B., Chen, K., Zhang, P., Chan, C., 2022. Multi-Vehicle Interactive Lane-Changing Velocity Change Model Based on Potential Energy Field. *Transp. Res. Rec. J. Transp. Res. Board* 2676, 306–323.
- Manco, G., Ritacco, E., Rullo, P., Gallucci, L., Astill, W., Kimber, D., Antonelli, M., 2017. Fault detection and explanation through big data analysis on sensor streams. *Expert Syst. Appl.* 87, 141–156.
- Mil, S., Piantanakulchai, M., 2018. Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions. *Appl. Soft*

Comput. J. 72, 65–78.

- Monje, L., Carrasco, R.A., Rosado, C., Sánchez-Montañés, M., 2022. Deep Learning XAI for Bus Passenger Forecasting: A Use Case in Spain. *Mathematics* 10, 1428.
- Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J., Damas, L., 2016. Time-evolving O-D matrix estimation using high-speed GPS data streams. *Expert Syst. Appl.* 44, 275–288.
- Mori, U., Mendiburu, A., Álvarez, M., Lozano, J.A., 2015. A review of travel time estimation and forecasting for Advanced Traveller Information Systems. *Transp. A Transp. Sci.* 11, 119–157.
- Mouskos, K.C., Niver, E., Pignataro, L.J., 1998. Transmit System Evaluation. *Database* 1–170.
- Müller, H.G., Yao, F., 2008. Functional additive models. J. Am. Stat. Assoc. 103, 1534–1544.
- Nair, R., Hoang, T.L., Laumanns, M., Chen, B., Cogill, R., Szabó, J., Walter, T., 2019. An ensemble prediction model for train delays. *Transp. Res. Part C Emerg. Technol.* 104, 196–209.
- Nantes, A., Ngoduy, D., Miska, M., Chung, E., 2015. Probabilistic travel time progression and its application to automatic vehicle identification data. *Transp. Res. Part B Methodol.* 81, 131–145.
- Nerini, D., Foresti, L., Leuenberger, D., Robert, S., Germann, U., 2019. A reducedspace ensemble kalman filter approach for flow-dependent integration of radar extrapolation nowcasts and NWP precipitation ensembles. *Mon. Weather Rev.* 147, 987–1006.
- Ng, L., Barfield, W., Mannering, F., 1995. A survey-based methodology to determine information requirements for advanced traveler information systems. *Transp. Res. Part C Emerg. Technol.* 3, 113–127.
- Ngoduy, D., Hoogendoorn, S.P., Van Zuylen, H.J., 2006. Continuum Traffic Model for Freeway with On- and Off-Ramp to Explain Different Traffic-Congested

States. Transp. Res. Rec. J. Transp. Res. Board 1965, 90-102.

- Nie, T., Qin, G., Sun, J., 2022. Truncated tensor Schatten p-norm based approach for spatiotemporal traffic data imputation with complicated missing patterns. *Transp. Res. Part C Emerg. Technol.* 141, 103737.
- Nosrati, K., Belikov, J., Tepljakov, A., Petlenkov, E., 2023. Extended fractional singular kalman filter. *Appl. Math. Comput.* 448, 127950.
- Ouyang, Q., Lv, Y., Ma, J., Li, J., 2020. An LSTM-Based Method Considering History and Real-Time Data for Passenger Flow Prediction. *Appl. Sci.* 10, 3788.
- Pan, T.L., Sumalee, A., Zhong, R.X., Indra-Payoong, N., 2013. Short-term traffic state prediction based on temporal-spatial correlation. *IEEE Trans. Intell. Transp. Syst.* 14, 1242–1254.
- Park, H., Kim, Y., 2018. Model for Filtering the Outliers in DSRC Travel Time Data on Interrupted Traffic Flow Sections. *KSCE J. Civ. Eng.* 22, 3607–3619.
- Petelin, G., Hribar, R., Papa, G., 2023. Models for forecasting the traffic flow within the city of Ljubljana. *Eur. Transp. Res. Rev.* 15, 30.
- Pu, Z., Zhu, M., Li, W., Cui, Z., Guo, X., Wang, Y., 2021. Monitoring Public Transit Ridership Flow by Passively Sensing Wi-Fi and Bluetooth Mobile Devices. *IEEE Internet Things J.* 8, 474–486.
- Qi, W., Rao, B., Fu, C., 2024. A Novel Filtering Method of Travel-Time Outliers Extracted from Large-Scale Traffic Checkpoint Data. J. Transp. Eng. Part A Syst. 150, 1–13.
- Qiao, W., Haghani, A., Shao, C.F., Liu, J., 2016. Freeway path travel time prediction based on heterogeneous traffic data through nonparametric model. J. Intell. Transp. Syst. Technol. Planning, Oper. 20, 438–448.
- Qin, W., Ji, X., Liang, F., 2020. Estimation of urban arterial travel time distribution considering link correlations. *Transp. A Transp. Sci.* 16, 1429–1458.
- Qiu, H., Li, R., Liu, H., 2016. Integrated model for traffic flow forecasting under rainy conditions. J. Adv. Transp. 50, 1754–1769.

- Quddus, M.A., Ochieng, W.Y., Noland, R.B., 2007. Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transp. Res. Part C Emerg. Technol.* 15, 312–328.
- Ramezani, M., Geroliminis, N., 2012. On the estimation of arterial route travel time distribution with Markov chains. *Transp. Res. Part B Methodol.* 46, 1576–1590.
- Ricard, L., Desaulniers, G., Lodi, A., Rousseau, L.-M., 2022. Predicting the probability distribution of bus travel time to measure the reliability of public transport services. *Transp. Res. Part C Emerg. Technol.* 138, 103619.
- Rim, H., Park, S., Oh, C., Park, J., Lee, G., 2016. Application of locally weighted regression-based approach in correcting erroneous individual vehicle speed data. *J. Adv. Transp.* 50, 180–196.
- Robinson, S., Polak, J., 2006. Overtaking rule method for the cleaning of matched license-plate data. *J. Transp. Eng.* 132, 609–617.
- Ryu, S., Kim, D., Kim, J., 2020. Weather-Aware Long-Range Traffic Forecast Using Multi-Module Deep Neural Network. *Appl. Sci.* 10, 1938.
- Sadeghi-Niaraki, A., Mirshafiei, P., Shakeri, M., Choi, S.-M., 2020. Short-Term Traffic Flow Prediction Using the Modified Elman Recurrent Neural Network Optimized Through a Genetic Algorithm. *IEEE Access* 8, 217526–217540.
- Salari, M., Kattan, L., Lam, W.H.K., Lo, H.P., Esfeh, M.A., 2019. Optimization of traffic sensor location for complete link flow observability in traffic network considering sensor failure. *Transp. Res. Part B Methodol.* 121, 216–251.
- Schwarz, G., 1978. Estimating the Dimension of a Model. Ann. Stat. 6, 461–464.
- Shang, Y., Li, X., Jia, B., Yang, Z., Liu, Z., 2022. Freeway Traffic State Estimation Method Based on Multisource Data. J. Transp. Eng. Part A Syst. 148, 1–14.
- Shao, H., Lam, W.H.K., Sumalee, A., Chen, A., 2013. Journey time estimator for assessment of road network performance under demand uncertainty. *Transp. Res. Part C Emerg. Technol.* 35, 244–262.

Shao, H., Lam, W.H.K., Tam, M.L., Yuan, X.M., 2008. Modelling rain effects on risk-

taking behaviours of multi-user classes in road networks with uncertainty. *J. Adv. Transp.* 42, 265–290.

- Shao, M., Xie, C., Sun, L., 2021. Optimization of network sensor location for full link flow observability considering sensor measurement error. *Transp. Res. Part C* 133, 103460.
- Shehu, B., Willems, W., Stockel, H., Thiele, L.-B., Haberlandt, U., 2023. Regionalisation of rainfall depth–duration–frequency curves with different data types in Germany. *Hydrol. Earth Syst. Sci.* 27, 1109–1132.
- Shi, C., Chen, B.Y., Lam, W.H.K., Li, Q., 2017. Heterogeneous Data Fusion Method to Estimate Travel Time Distributions in Congested Road Networks. *Sensors* 17, 2822.
- Shi, R., Mo, Z., Huang, K., Di, X., Du, Q., 2022. A Physics-Informed Deep Learning Paradigm for Traffic State and Fundamental Diagram Estimation. *IEEE Trans. Intell. Transp. Syst.* 23, 11688–11698.
- Shi, X., Qi, H., Shen, Y., Wu, G., Yin, B., 2021. A Spatial–Temporal Attention Approach for Traffic Prediction. *IEEE Trans. Intell. Transp. Syst.* 22, 4909–4918.
- Singh, S., Singh, J., Goyal, S.B., El Barachi, M., Kumar, M., 2023. Analytical Review of Map Matching Algorithms: Analyzing the Performance and Efficiency Using Road Dataset of the Indian Subcontinent. *Arch. Comput. Methods Eng.* 30, 4897– 4916.
- Soriguera, F., Martinez-DIaz, M., 2021. Freeway Travel Time Information from Input-Output Vehicle Counts: A Drift Correction Method Based on AVI Data. *IEEE Trans. Intell. Transp. Syst.* 22, 5749–5761.
- Soriguera, Francesc, Robusté, F., 2011. Requiem for freeway travel time estimation methods based on blind speed interpolations between point measurements. *IEEE Trans. Intell. Transp. Syst.* 12, 291–297.
- Soriguera, F., Robusté, F., 2011. Highway travel time accurate measurement and shortterm prediction using multiple data sources. *Transportmetrica* 7, 85–109.

- Southwest Research Institute, 1998. Automated Vehicle Identification Model Deployment Initiative System Design Document. San Antonio, Texas.
- Stathopoulos, A., Karlaftis, M., 2001. Temporal and spatial variations of real-time traffic data in urban areas. *Transp. Res. Rec.* 135–140.
- Sun, R., Han, K., Hu, J., Wang, Y., Hu, M., Ochieng, W.Y., 2016. Integrated solution for anomalous driving detection based on BeiDou/GPS/IMU measurements. *Transp. Res. Part C Emerg. Technol.* 69, 193–207.
- Sun, T., Zhao, K., Zhang, C., Chen, M., Yu, X., 2022. PR-LTTE: Link travel time estimation based on path recovery from large-scale incomplete trip data. *Inf. Sci.* (Ny). 589, 34–45.
- Sun, W., Shao, H., Wu, T., Shao, F., Fainman, E.Z., 2022. Reliable location of automatic vehicle identification sensors to recognize origin-destination demands considering sensor failure. *Transp. Res. Part C Emerg. Technol.* 136, 103551.
- Sun, Y., Tian, X., Bao, W., Qu, S., Li, Q., Chen, Y., Zhou, Z., Shi, P., 2023. Improving Treatment of Noise Specification of Kalman Filtering for State Updating of Hydrological Models: Combining the Strengths of the Interacting Multiple Model Method and Cubature Kalman Filter. *Water Resour. Res.* 59, e2022WR033635.
- Tam, M.L., Lam, W.H.K., 2008. Using automatic vehicle identification data for travel time estimation in Hong Kong. *Transportmetrica* 4, 179–194.
- Tam, M.L., Lam, W.H.K., 2011a. Application of automatic vehicle identification technology for real-time journey time estimation. *Inf. Fusion* 12, 11–19.
- Tam, M.L., Lam, W.H.K., 2011b. Validation of Instantaneous Journey Time Estimates: A Journey Time Indication System in Hong Kong. Proc. East. Asia Soc. Transp. Stud. 2011, 336.
- Tam, M.L., Lam, W.H.K., 2013. Validation of ATIS journey time and traffic speed estimates by floating car survey. J. East. Asia Soc. Transp. Stud. 10, 131–146.
- Tam, M.L., Lam, W.H.K., Chen, B., Chan, K.S., Wong, S.C., 2008. Using Automatic Vehicle Identification Data for Investigation of Rain Effects on Vehicular Travel

Speeds and Travel Choice Behaviour.

- Tang, J., Zou, Y., Ash, J., Zhang, S., Liu, F., Wang, Y., 2016. Travel Time Estimation Using Freeway Point Detector Data Based on Evolving Fuzzy Neural Inference System. *PLoS One* 11, e0147263.
- Tang, K., Chen, S., Khattak, A.J., 2018. Personalized travel time estimation for urban road networks: A tensor-based context-aware approach. *Expert Syst. Appl.* 103, 118–132.
- Tani, R., Kato, T., Uchida, K., 2020. A method for structuring stochastic travel time by using risk premiums of stochastic link flow. *Transp. A Transp. Sci.*
- Thakuriah, P. (Vonu), Tilahun, N., 2013. Incorporating Weather Information into Real-Time Speed Estimates: Comparison of Alternative Models. J. Transp. Eng. 139, 379–389.
- TranStar, H., 2021. TranStar Description. < http://traffic.houstontranstar.org>
- Trinh, X.-S., Ngoduy, D., Keyvan-Ekbatani, M., Robertson, B., 2022. Incremental unscented Kalman filter for real-time traffic estimation on motorways using multi-source data. *Transp. A Transp. Sci.* 18, 1127–1153.
- Van Hinsbergen, C.P.I.J., Schreiter, T., Zuurbier, F.S., Van Lint, J.W.C., Van Zuylen, H.J., 2012. Localized extended kalman filter for scalable real-time traffic state estimation. *IEEE Trans. Intell. Transp. Syst.* 13, 385–394.
- Wagner-Muns, I.M., Guardiola, I.G., Samaranayke, V.A., Kayani, W.I., 2018. A Functional Data Analysis Approach to Traffic Volume Forecasting. *IEEE Trans. Intell. Transp. Syst.* 19, 878–888.
- Walch, M., Neubauer, M., Schildorfer, W., 2023. Floating Car Data–Based Short-Term Travel Time Forecasting with Deep Recurrent Neural Networks Incorporating Weather Data. J. Transp. Eng. Part A Syst. 149, 04023035.
- Wang, C., Zhao, F., Zhang, H., Luo, H., Qin, Y., Fang, Y., 2022. Fine-Grained Trajectory-Based Travel Time Estimation for Multi-City Scenarios Based on Deep Meta-Learning. *IEEE Trans. Intell. Transp. Syst.* 23, 15716–15728.

- Wang, P., Huang, Z., Lai, J., Zheng, Z., Liu, Y., Lin, T., 2021. Traffic Speed Estimation Based on Multi-Source GPS Data and Mixture Model. *IEEE Trans. Intell. Transp. Syst.* 23, 10708–10720.
- Wang, Y., Bie, Y., An, Q., 2018. Impacts of Winter Weather on Bus Travel Time in Cold Regions: Case Study of Harbin, China. J. Transp. Eng. Part A Syst. 144, 1– 9.
- Wang, Z., Liang, M., Delahaye, D., 2018. A hybrid machine learning model for shortterm estimated time of arrival prediction in terminal manoeuvring area. *Transp. Res. Part C Emerg. Technol.* 95, 280–294.
- Wang, Z., Su, X., Ding, Z., 2021. Long-Term Traffic Prediction Based on LSTM Encoder-Decoder Architecture. *IEEE Trans. Intell. Transp. Syst.* 22, 6561–6571.
- Washington, S., Karlaftis, M., Mannering, F., Anastasopoulos, P., 2020. Statistical and Econometric Methods for Transportation Data Analysis. Chapman and Hall/CRC.
- Wei, X., Li, J., Yuan, Q., Chen, K., Zhou, A., Yang, F., 2019. Predicting Fine-Grained Traffic Conditions via Spatio-Temporal LSTM. *Wirel. Commun. Mob. Comput.* 2019, 1–12.
- Wilby, M.R., Gonzalez, A.B.R., Pozo, R.F., Vinagre Diaz, J.J., 2022. Short-Term Prediction of Level of Service in Highways Based on Bluetooth Identification. *IEEE Trans. Intell. Transp. Syst.* 23, 142–151.
- Wirtgen, C., Kowald, M., Luderschmidt, J., Hünemohr, H., 2022. Multivariate Demand Forecasting for Rental Bike Systems Based on an Unobserved Component Model. *Electronics* 11, 4146.
- Wu, L., Shou, G., Xie, Z., Jing, P., 2023. Mobile Phone Data Feature Denoising for Expressway Traffic State Estimation. Sustain. 15, 1–15.
- Wu, Y., Yang, X., Zhang, W., Kuang, Q., 2019. Mixture probabilistic model for precipitation ensemble forecasting. Q. J. R. Meteorol. Soc. 145, 3516–3534.
- Xia, D., Zheng, L., Tang, Y., Cai, X., Chen, L., Liu, W., Sun, D., 2022. Link-Based Traffic Estimation and Simulation for Road Networks Using Electronic

Registration Identification Data. IEEE Trans. Veh. Technol. 71, 8075-8088.

- Xue, Y., Feng, R., Cui, S., Yu, B., 2020. Traffic Status Evolution Trend Prediction Based on Congestion Propagation Effects under Rainy Weather. J. Adv. Transp. 2020, 1–14.
- Yang, J.-Y., Chou, L.-D., Tung, C.-F., Huang, S.-M., Wang, T.-W., 2013. Average-Speed Forecast and Adjustment via VANETs. *IEEE Trans. Veh. Technol.* 62, 4318–4327.
- Yang, J.M., Peng, Z.R., Lin, L., 2021. Real-time spatiotemporal prediction and imputation of traffic status based on LSTM and Graph Laplacian regularized matrix factorization. *Transp. Res. Part C Emerg. Technol.* 129, 103228.
- Yang, S., Chen, M., Wu, Y.J., An, C., 2018. Machine Learning Approach to Decomposing Arterial Travel Time Using a Hidden Markov Model with Genetic Algorithm. J. Comput. Civ. Eng. 32, 1–11.
- Yang, S., Qian, S., 2019. Understanding and Predicting Travel Time with Spatio-Temporal Features of Network Traffic Flow, Weather and Incidents. *IEEE Intell. Transp. Syst. Mag.* 11, 12–28.
- Yang, Y., Yan, J., Guo, J., Kuang, Y., Yin, M., Wang, S., Ma, C., 2021. Driving Behavior Analysis of City Buses Based on Real-Time GNSS Traces and Road Information. *Sensors* 21, 687.
- Yao, F., Müller, H.-G., Wang, J.-L., 2005. Functional Data Analysis for Sparse Longitudinal Data. J. Am. Stat. Assoc. 100, 577–590.
- Ye, D.J., Tam, M.L., 2011. Automatic Incident Detection for Urban Roads in Hong Kong. J. East. Asia Soc. Transp. Stud. 9, 1897–1912.
- Ye, Y., Zhu, Y., Markos, C., Yu, J.J.Q., 2022. CatETA: A Categorical Approximate Approach for Estimating Time of Arrival. *IEEE Trans. Intell. Transp. Syst.* 23, 24389–24400.
- Yildirimoglu, M., Geroliminis, N., 2013. Experienced travel time prediction for congested freeways. *Transp. Res. Part B Methodol.* 53, 45–63.

- Yin, K., Wang, W., Bruce Wang, X., Adams, T.M., 2015. Link travel time inference using entry/exit information of trips on a network. *Transp. Res. Part B Methodol*. 80, 303–321.
- Yu, Bin, Yang, Z.-Z., Chen, K., Yu, Bo, 2010. Hybrid model for prediction of bus arrival times at next station. J. Adv. Transp. 44, 193–204.
- Yuan, S., Zhao, K., Xu, Z., 2023. Adaptive Gaussian mixture model for identifying outliers in historical route travel times. *IET Intell. Transp. Syst.* 17, 2458–2473.
- Yun, M., Qin, W., Yang, X., Liang, F., 2019a. Estimation of urban route travel time distribution using Markov chains and pair-copula construction. *Transp. B Transp. Dyn.* 7, 1521–1552.
- Yun, M., Qin, W., Yang, X., Liang, F., 2019b. Estimation of urban route travel time distribution using Markov chains and pair-copula construction. *Transp. B* 7, 1521–1552.
- Zhan, X., Li, R., Ukkusuri, S. V., 2020. Link-based traffic state estimation and prediction for arterial networks using license-plate recognition data. *Transp. Res. Part C Emerg. Technol.* 117, 102660.
- Zhang, J., Song, G., Gong, D., Gao, Y., Yu, L., Guo, J., 2018. Analysis of rainfall effects on road travel speed in Beijing, China. *IET Intell. Transp. Syst.* 12, 93– 102.
- Zhang, W., 2006. Freeway travel time estimation based on spot speed measurements.
- Zhao, J., Gao, Y., Bai, Z., Wang, H., Lu, S., 2019. Traffic Speed Prediction Under Non-Recurrent Congestion: Based on LSTM Method and BeiDou Navigation Satellite System Data. *IEEE Intell. Transp. Syst. Mag.* 11, 70–81.
- Zhong, R.X., Luo, J.C., Cai, H.X., Sumalee, A., Yuan, F.F., Chow, A.H.F., 2017. Forecasting journey time distribution with consideration to abnormal traffic conditions. *Transp. Res. Part C Emerg. Technol.* 85, 292–311.
- Zhong, R.X., Xie, X.X., Luo, J.C., Pan, T.L., Lam, W.H.K., Sumalee, A., 2020. Modeling double time-scale travel time processes with application to assessing

the resilience of transportation systems. *Transp. Res. Part B Methodol.* 132, 228–248.

- Zhou, W., Xiao, X., Gong, Y.-J., Chen, J., Fang, J., Tan, N., Ma, N., Li, Q., Hua, C., Jeon, S.-W., Zhang, J., 2023. Travel Time Distribution Estimation by Learning Representations Over Temporal Attributed Graphs. *IEEE Trans. Intell. Transp. Syst.* 24, 1–13.
- Zhou, X., Mahmassani, H.S., 2006. Dynamic origin-destination demand estimation using automatic vehicle identification data. *IEEE Trans. Intell. Transp. Syst.* 7, 105–114.
- Zhu, K., Zhang, C., Xue, M., Yang, N., 2022. Predictability and skill of convectionpermitting ensemble forecast systems in predicting the record-breaking "21·7" extreme rainfall event in Henan Province, China. Sci. China Earth Sci. 65, 1879– 1902.
- Zhu, L., Guo, F., Polak, J.W., Krishnan, R., 2018. Urban link travel time estimation using traffic states-based data fusion. *IET Intell. Transp. Syst.* 12, 651–663.
- Zhu, N., Ma, S., Zheng, L., 2017. Travel time estimation oriented freeway sensor placement problem considering sensor failure. J. Intell. Transp. Syst. Technol. Planning, Oper. 21, 26–40.
- Zhu, Y., He, Z., Sun, W., 2020. Network-Wide Link Travel Time Inference Using Trip-Based Data from Automatic Vehicle Identification Detectors. *IEEE Trans. Intell. Transp. Syst.* 21, 2485–2495.
- Zhu, Y., Ye, Y., Liu, Y., Yu, J.J.Q., 2022. Cross-Area Travel Time Uncertainty Estimation From Trajectory Data: A Federated Learning Approach. *IEEE Trans. Intell. Transp. Syst.* 23, 24966–24978.