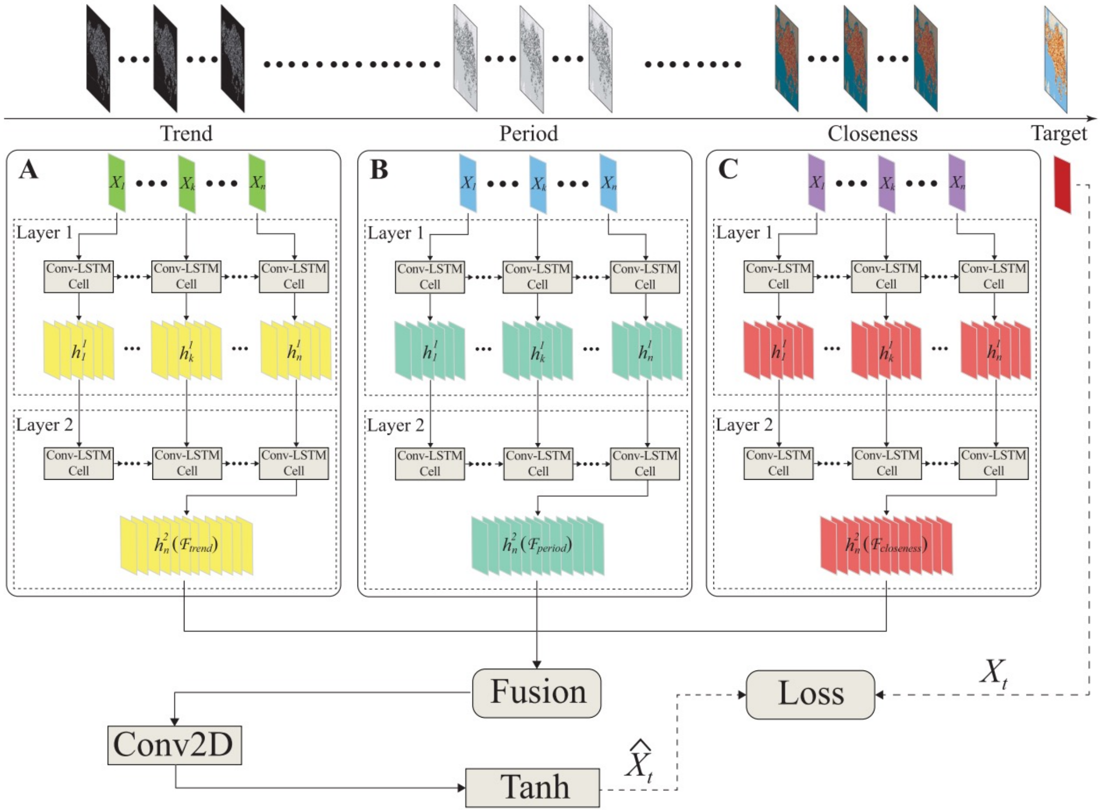



Fig. 2. The architecture of Spatial-Temporal Memory Network. Three independent Conv-LSTM modules, $A$, $B$, and $C$ are utilized to capture spatial-temporal features from corresponding historical periods: $Trend$, $Period$ and $Closeness$. Each Conv-LSTM module contains two stacked layers for extracting deep-level spatial-temporal correlations. The fusion module is to merge three spatial-temporal features for final forecasting. Then a $Conv2D$ is to reduce the dimension, and a function $Tanh$ is used to activate a non-linear process. $\hat{X}_t$ and $X_t$ denote the prediction result and the ground truth, respectively. The backpropagation is implemented in loss module to adjust the weights.

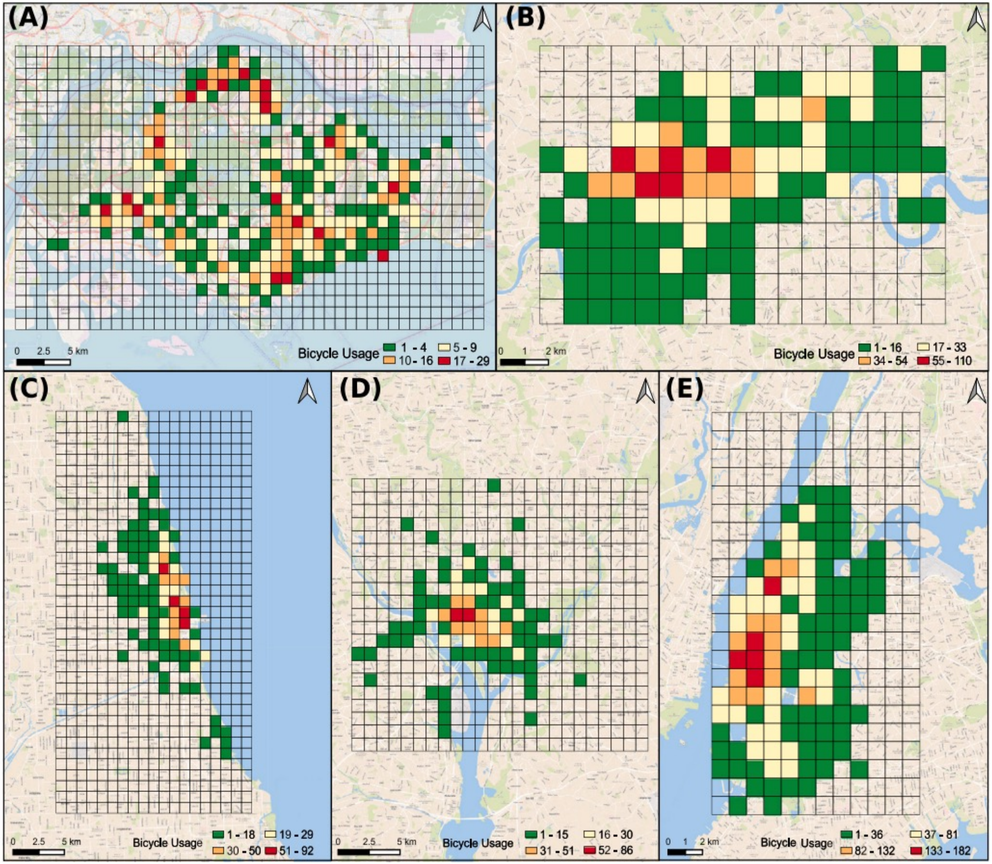



**Fig. 1.** Spatial distribution of bicycle usage in five cities. **(A)** Singapore during 5–6 PM on August 25th, 2017; **(B)** London during 5–6 PM on August 24th, 2019; **(C)** Chicago during 1–2 PM on August 25th, 2019; **(D)** Washington, D.C. during 5–6 PM on August 25th, 2019; **(E)** New York during 5–6 PM on August 25th, 2019.

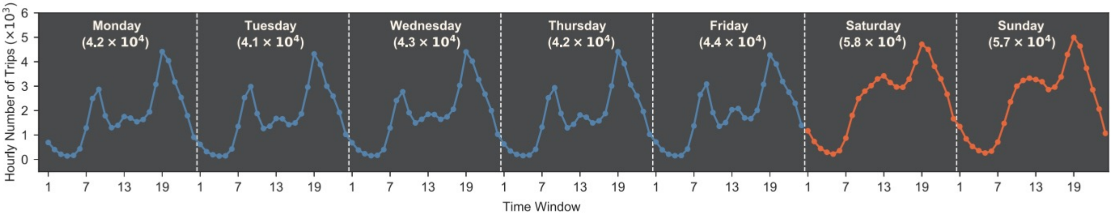
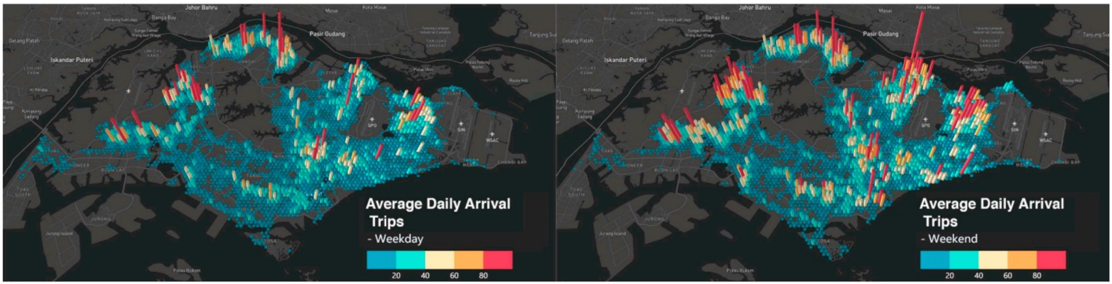



**Fig. 1.** Hourly number of trips averaged by day of week.



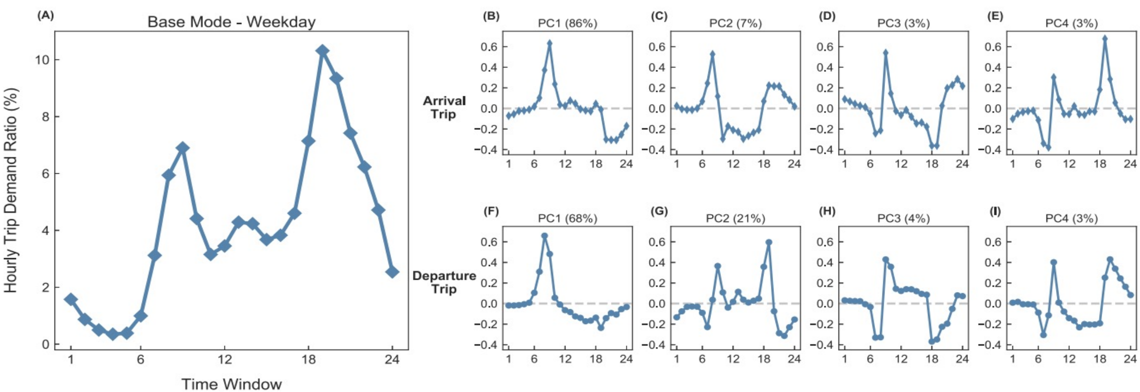
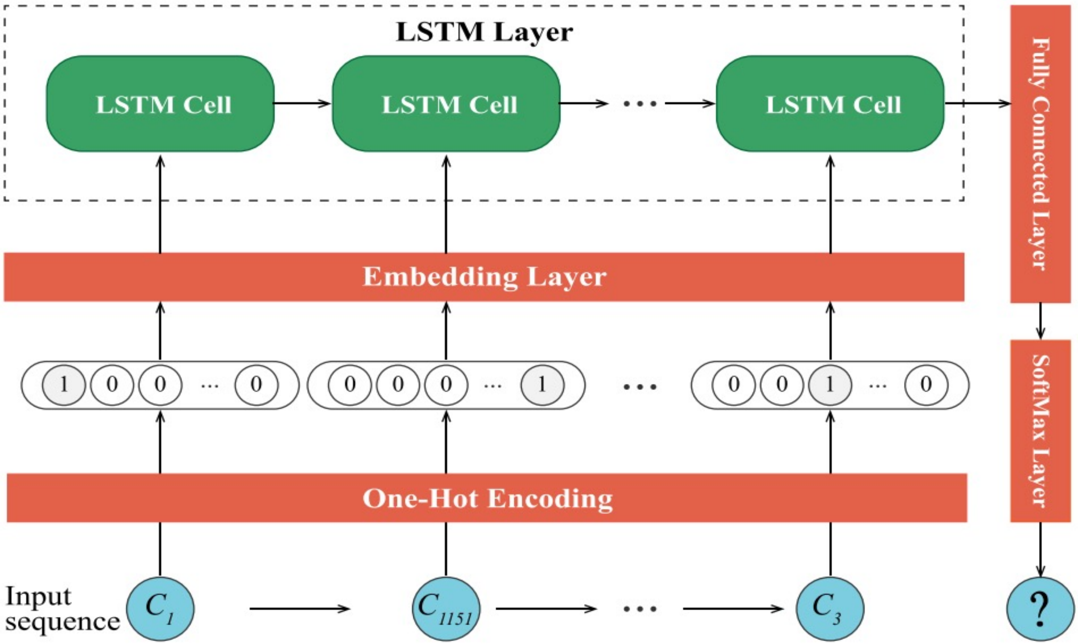
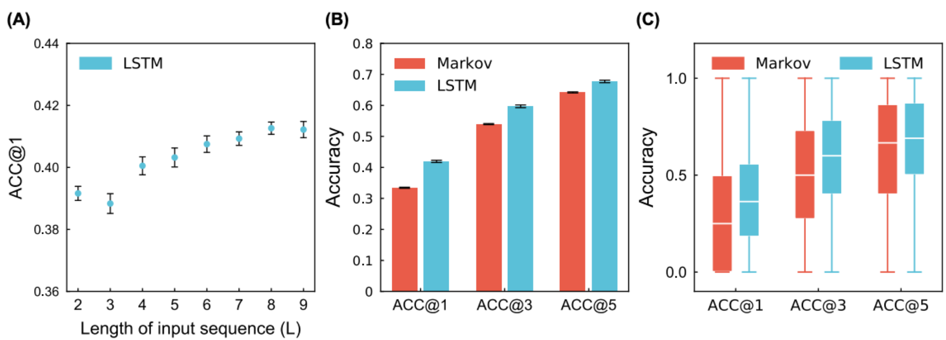
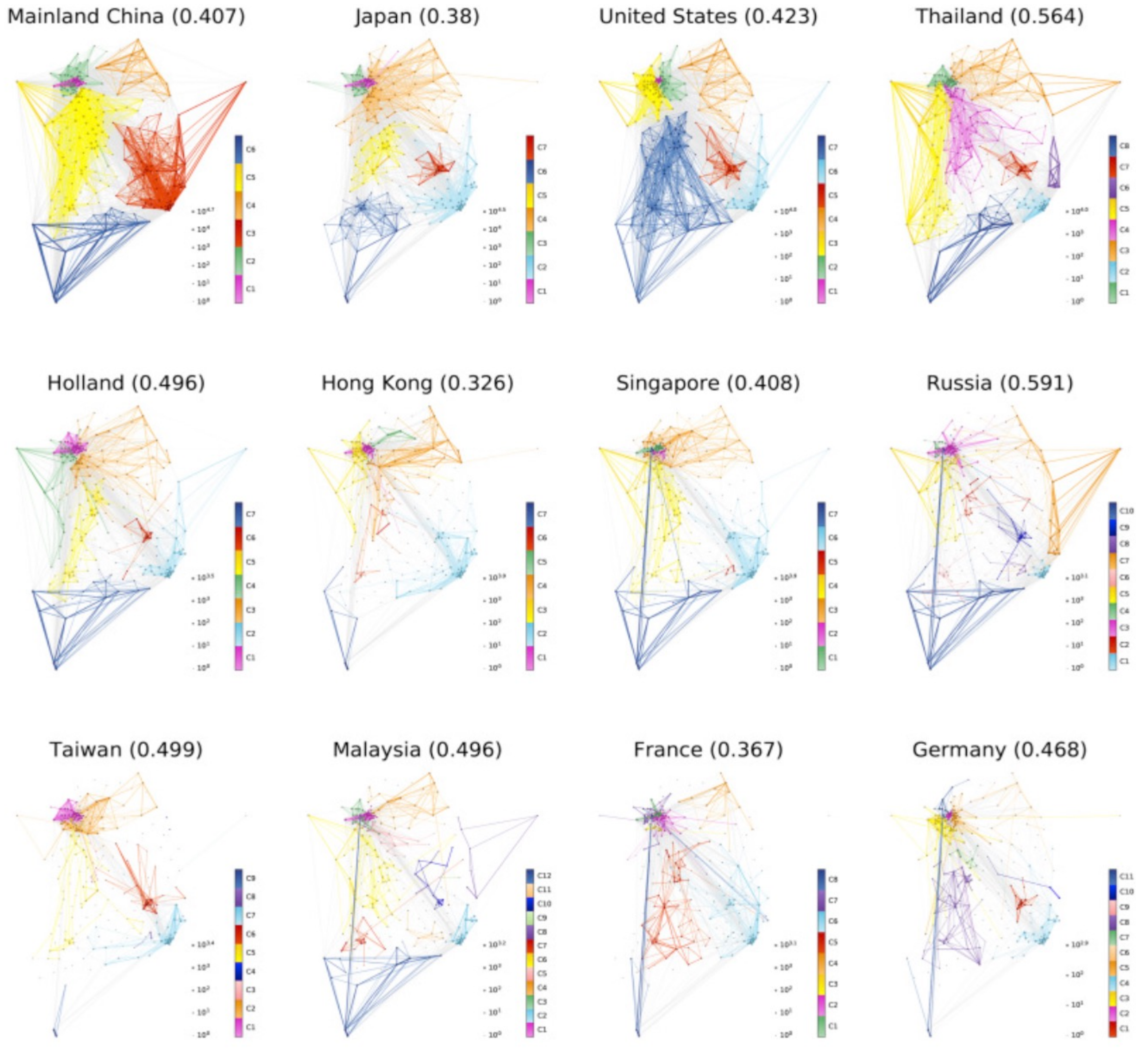




**Fig. 3.** Results of eigendecomposition — weekday.


LSTM Layer
LSTM Cell
LSTM Cell
LSTM Cell
Fully Connected Layer
Embedding Layer
SoftMax Layer
One-Hot Encoding
Input sequence
$C_1$
$C_{1151}$
$C_3$
?


(A)
LSTM
ACC@1
Length of input sequence (L)
(B)
Markov
LSTM
Accuracy
ACC@1
ACC@3
ACC@5
(C)
Markov
LSTM
Accuracy
ACC@1
ACC@3
ACC@5


Mainland China (0.407)
Japan (0.38)
United States (0.423)
Thailand (0.564)
Holland (0.496)
Hong Kong (0.326)
Singapore (0.408)
Russia (0.591)
Taiwan (0.499)
Malaysia (0.496)
France (0.367)
Germany (0.468)