

Table 1: Description of the key notations.

Notation	Descriptions
\mathbf{X}_t	Original input series
\mathbf{Y}_t	Target output series
N	Number of variables in the series
L	Length of the historical window
T	Length of the prediction window
$\mathbf{X}_{t,:}$	Time series collected at time step t
$\mathbf{X}_{:,n}$	Entire time series of of each variable indexed by n
$G = (\mathbf{V}, \mathbf{E})$	Graph with node set V and edge set E
G_{strong}	Strong correlation graph
G_{weak}	Weak correlation graph
$f = \{f_1, \dots, f_k\}$	Multi-scale representations in MTS
$\mathbf{A}_{\text{strong}}^k$	Adjacency matrix of the strong correlation graph at scale k
$\mathbf{A}_{\text{weak}}^k$	Adjacency matrix of the weak correlation graph at scale k
$\mathbf{E}_{i \in \{m,d,w,h,t\}}$	Embeddings for month, day, week, hour, and minute
\mathbf{H}_{emb}	Embedding of the original input series
\mathbf{X}_f	Fast Fourier Transform of \mathbf{H}_{emb}
\mathbf{F}	Overall amplitude measure
p_i	Period corresponding to different scales
\mathbf{X}^i	The i -th reshaped time series for time scale i
$\mathbf{E}_1^l, \mathbf{E}_2^l$	Learnable parameters at layer l for source and target node embeddings
\mathbf{A}^l	Adjacency matrices at layer l
$\mathbf{A}_{\text{strong}}^l, \mathbf{A}_{\text{weak}}^l$	Strong and Weak correlation matrices at layer l
$\mathbf{E}_{\text{index}}^l$	Edge index matrix at layer l indicating connections
$\mathbf{A}_{\text{edge}}^l$	Edge attribute matrix at layer l describing edge features
\mathbf{G}_{data}	Contains node and edge features
\mathbf{B}^l	All-zero batch tensor
r	Source node of an edge
c	Target node of an edge
\mathbf{E}_{rep}	Node features from source and target node indices
\mathbf{S}_{edge}	Computed edge scores
\mathbf{A}_f^l	Fused adjacency matrix at layer l
$\mathbf{X}_{\text{out}}^i$	The output after fusion at scale i