用于多元时间序列异常检测的图混合专家模型和记忆增强路由器

汇报人: 韦浩文

时间：2025-3-16

目录

1 背景

背景

现有研究

多元时间序列（MTS）异常检测指的是从多个相互关联的时间序列组成的数据中识别异常

基于GNN的方法广泛用于MTS异常检测

现有的方法的共同局限性

- 只利用GNN最后一层的输出来进行异常估计，忽略了中间层信息
- 所有节点共享相同的聚合机制，忽略了节点在特征和邻域上的差异

背景

本研究

Graph-MoE

全面利用多层GNN的所有中间信息

- 混合专家模块-层内聚合
- 记忆增强路由器-层间聚合
- 即插即用

2

方法

Graph-MoE

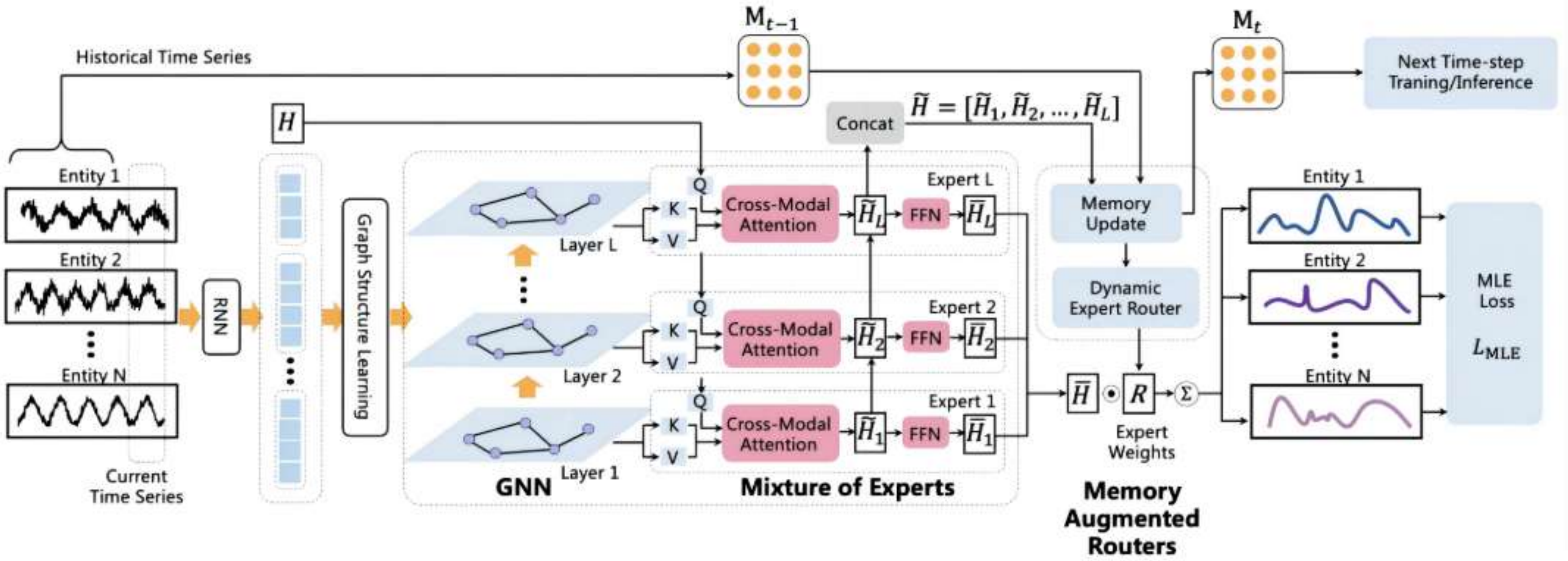
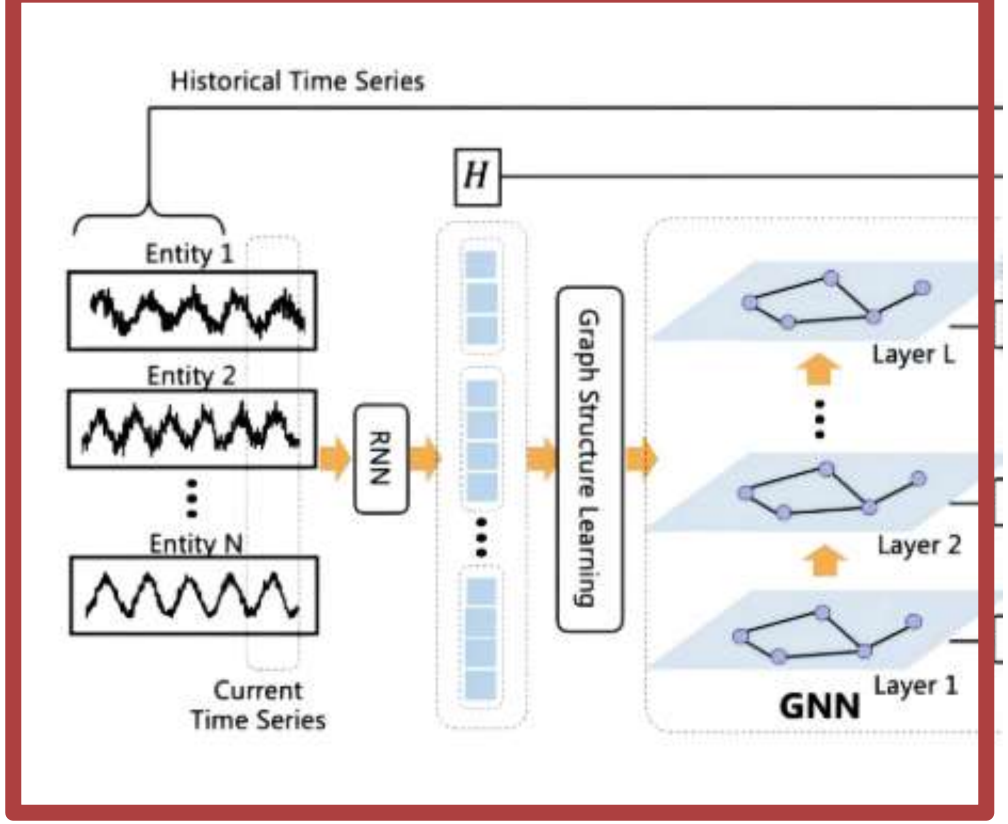


Figure 2: The overview of our proposed Graph-MoE network. It mainly consists of 1) graph mixture of experts and 2) memory-augmented routers.

Graph-MoE



$$\bar{x}_i = \frac{x_i - \mathrm{mean}(x_i)}{\mathrm{std}(x_i)}$$

x^c stands for $x^{cS:cS+T}$

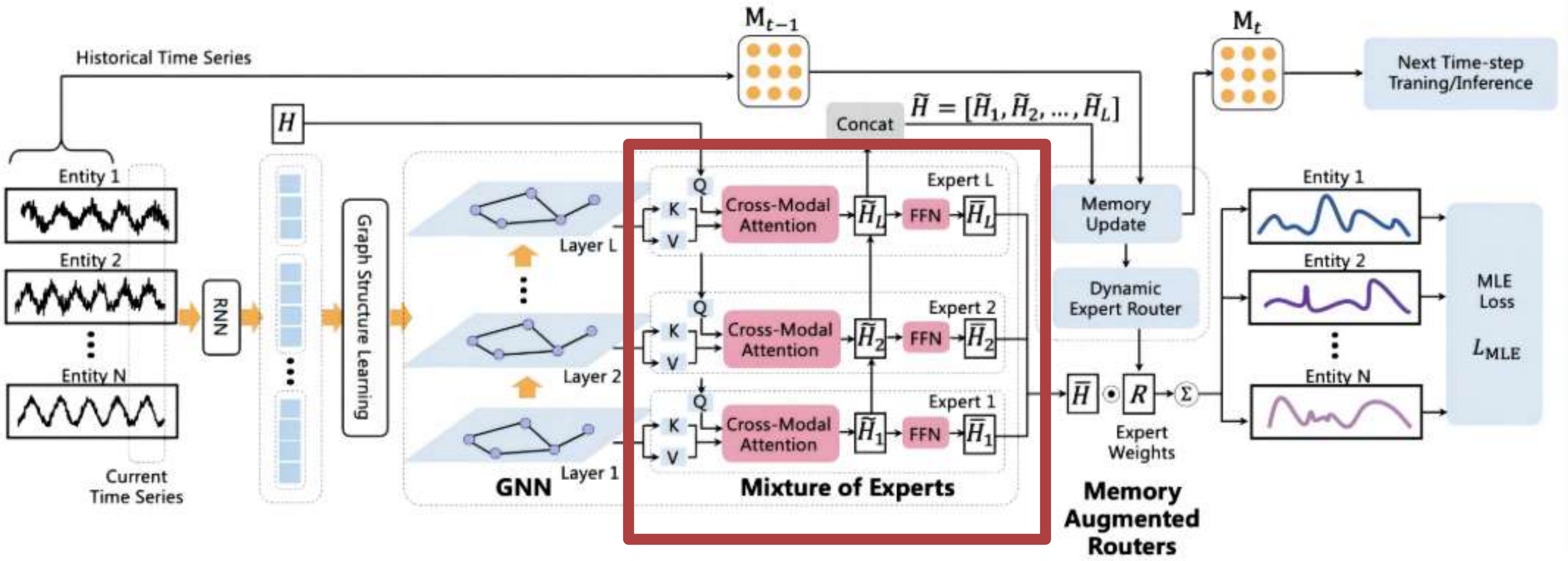
$$H_k^t = \mathrm{RNN}(x_k^t, H_k^{t-1})$$

$$e_{ij}^c = (\phi_e^1(x_i^c)) \cdot (\phi_e^2(x_j^c))^\top$$

$$a_{ij}^c = \frac{\exp(e_{ij}^c)}{\sum_{j=1}^K \exp(e_{ij}^c)}$$

$$H_t^l = \mathrm{ReLU}(A^c H_t^{l-1} W_1 + H_{t-1}^{l-1} W_2) \cdot W_3$$

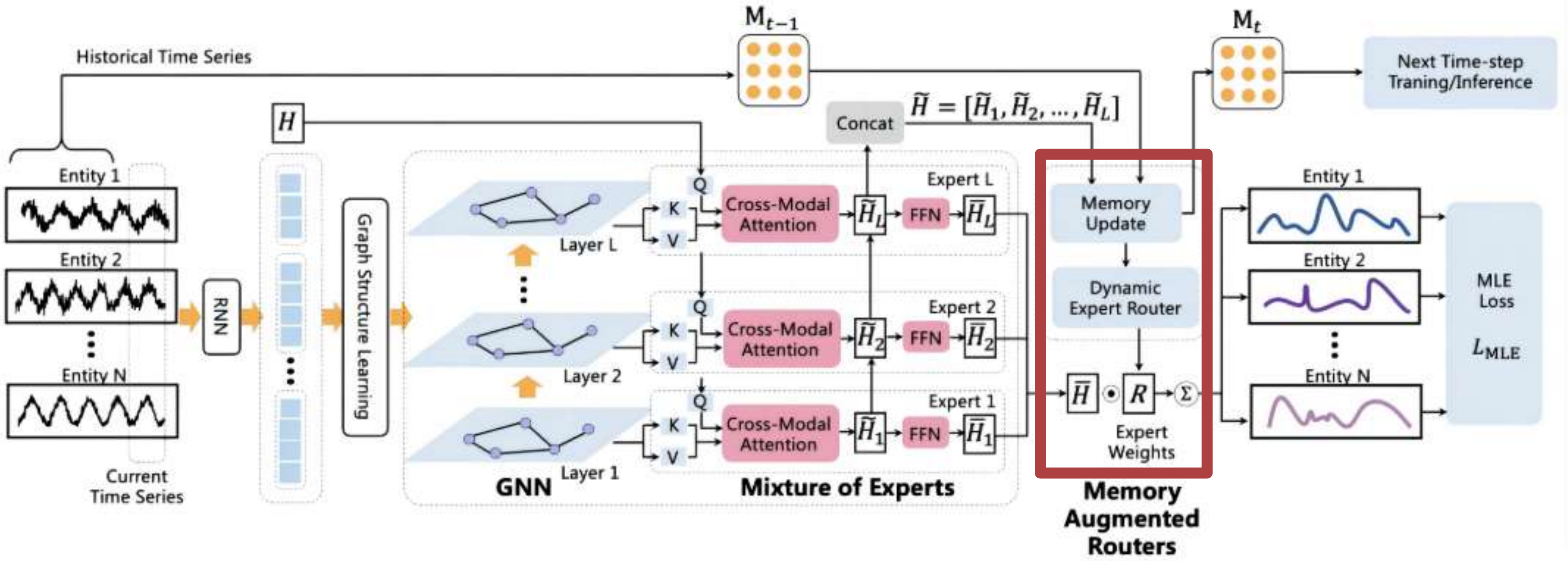
Graph-MoE——图混合专家



$$\tilde{H}^l = \mathrm{LN}(\mathrm{Attention}(H, H^l))|_{Q:H,\{K,V\}:H^l}$$

$$\bar{H}^l = \mathrm{FFN}_l(\tilde{H}^l) = \phi_l^2\Big(\mathrm{ReLU}\big(\phi_l^1(\tilde{H}^l)\big)\Big)$$

Graph-MoE——记忆增强路由器



$$\tilde{\mathbf{H}} = [\tilde{H}^1, \tilde{H}^2, \ldots, \tilde{H}^L].$$

Graph-MoE——记忆增强路由器

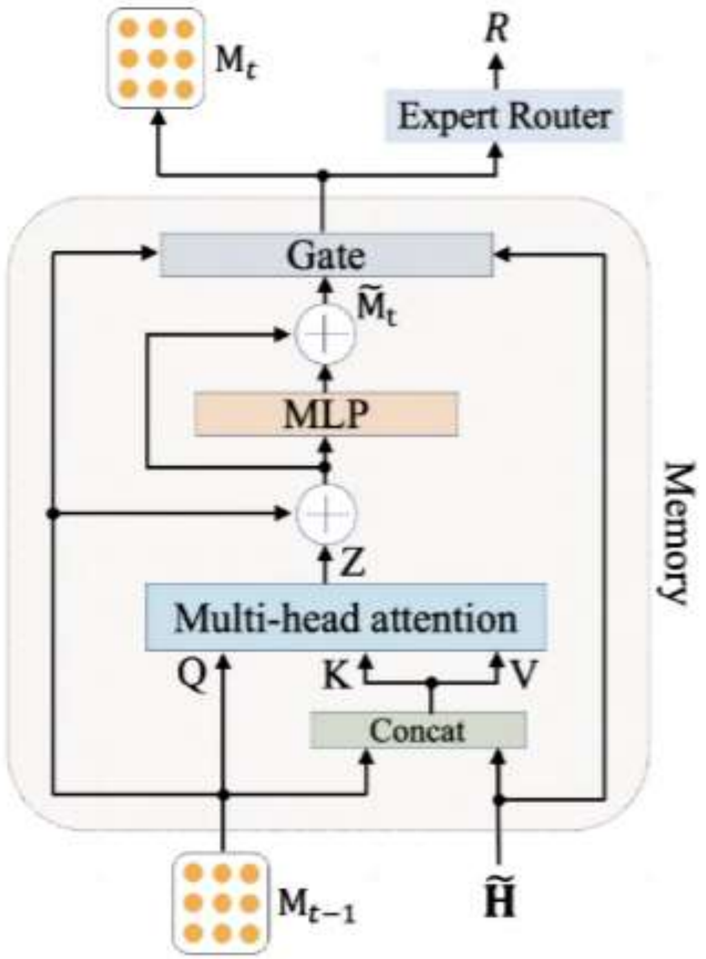


Figure 3: The framework of our memory-augmented routers.

$$Y = [M_{t-1}; \tilde{\mathbf{H}}_t],$$
$$Z = \text{Attention}(M_{t-1}, Y)|_{Q:M_{t-1},\{K,V\}:Y}$$

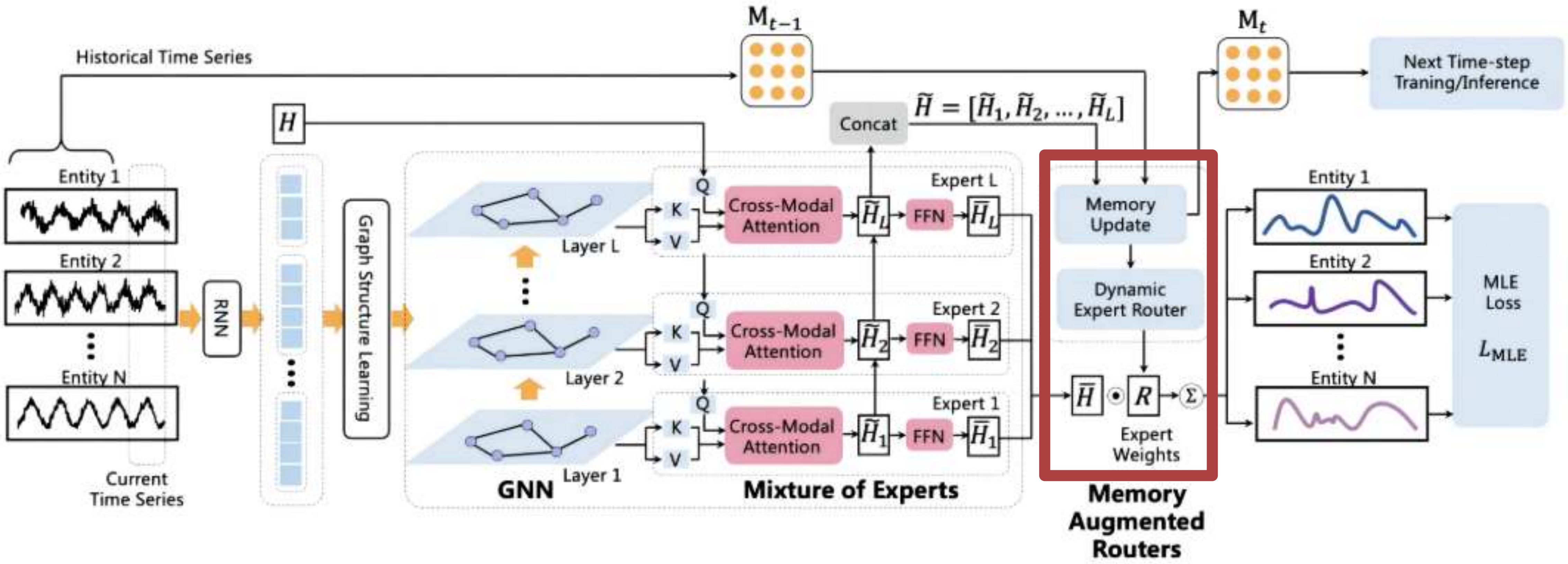
$$\tilde{M}_t = \phi_M(Z + M_{t-1}) + Z + M_{t-1}$$

$$G_t^f = \tilde{\mathbf{H}}_t W^f + \tanh(M_{t-1}) \cdot U^f$$
$$G_t^i = \tilde{\mathbf{H}}_t W^i + \tanh(M_{t-1}) \cdot U^i,$$

$$M_t = \sigma(G_t^f) \odot M_{t-1} + \sigma(G_t^i) \odot \tanh(\tilde{M}_t)$$

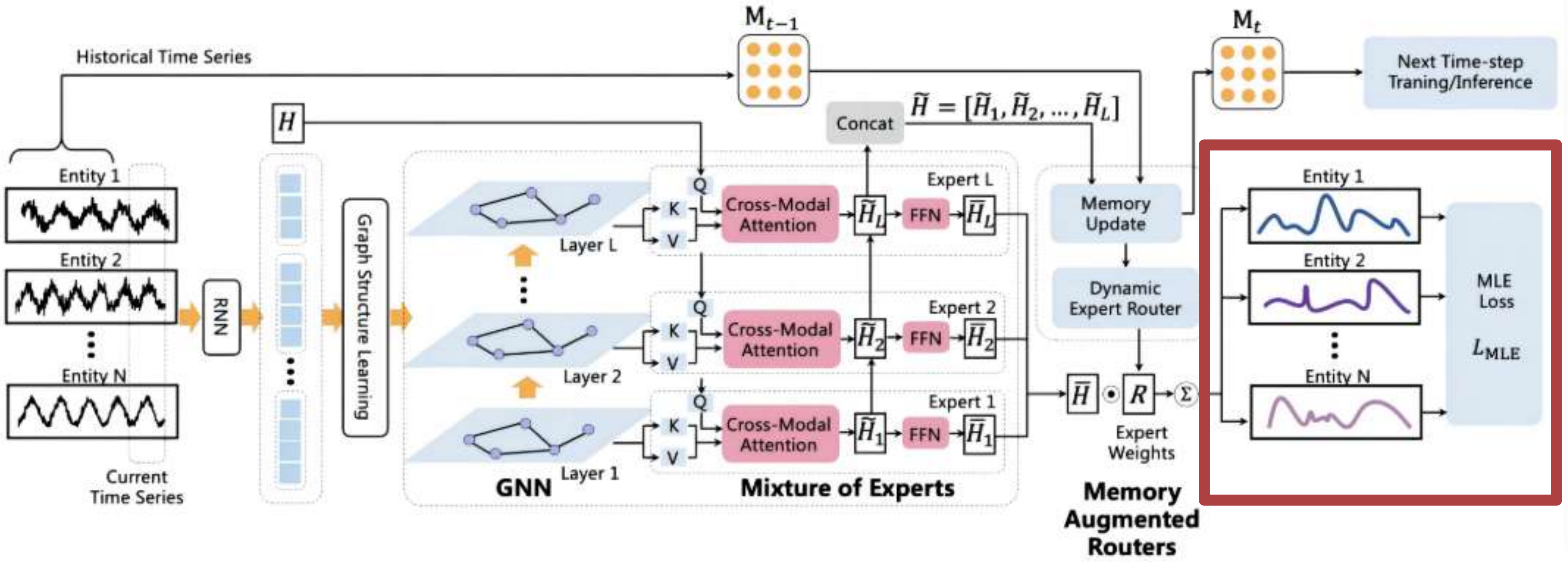
$$R = \text{Softmax}(\phi_R(M_t))$$

Graph-MoE——记忆增强路由器



$$C = \sum_{l=1}^{L} R_l \cdot f_l(H^l) = \sum_{l=1}^{L} R_l \cdot \bar{H}^l$$

Graph-MoE



$$z = f_\theta(x|C) \qquad \theta^* = \arg\max_\theta(\log(P_{\mathcal{Z}}(f_\theta(x|C)) + \log(\left|\det\frac{\partial f_\theta(x)}{\partial x^{\mathrm{T}}}\right|)) \qquad P_{\mathcal{X}}(x) = P_{\mathcal{Z}}(z)\left|\det\frac{\partial z}{\partial x^{\mathrm{T}}}\right|.$$

3 实验

实验-数据集、评估指标

数据集

1.SwaT 51个传感器，4天内的正常操作和41次异常操作

2.WADI 123个传感器和执行器，数据采样频率为每秒一次

3.PSM 25个特征，8周的服务器节点数据

4.MSL 55个传感器和执行器

5.SMD 38个特征，5周的互联网公司数据

评估指标

- 窗口级别的异常检测
- AUROC

  -ROC曲线下的面积

实验-主要比较

Table 1: Comparison with the state-of-the-art methods in anomaly detection on five challenging datasets, *i.e.*, SWaT, WADI, PSM, MSL, and SMD. The best results are highlighted in bold.

Method	Venue	Datasets				
		SWaT	WADI	PSM	MSL	SMD
DeepSVDD (Ruff et al. 2018)	ICML2018	66.8±2.0	83.5±1.6	67.5±1.4	60.8±0.4	75.5±15.5
DAGMM (Zong et al. 2018)	ICLR 2018	72.8±3.0	77.2±0.9	64.6±2.6	56.5±2.6	78.0±9.2
ALOCC (Sabokrou et al. 2020)	TNNLS 2020	77.1±2.3	83.3±1.8	71.8±1.3	60.3±0.9	80.5±11.1
DROCC (Goyal et al. 2020)	ICML 2020	72.6±3.8	75.6±1.6	74.3±2.0	53.4±1.6	76.7±8.7
DeepSAD (Ruff et al. 2020)	ICLR 2020	75.4±2.4	85.4±2.7	73.2±3.3	61.6±0.6	85.9±11.1
USAD (Audibert et al. 2020)	KDD 2020	78.8±1.0	86.1±0.9	78.0±0.2	57.0±0.1	86.9±11.7
GANF (Dai and Chen 2022)	ICLR 2022	79.8±0.7	90.3±1.0	81.8±1.5	64.5±1.9	89.2±7.8
MTGFlow (Zhou et al. 2023)	AAAI 2023	84.8±1.5	91.9±1.1	85.7±1.5	67.2±1.7	91.3±7.6
Ours (Graph-MoE)		**87.2±1.3**	**94.2±0.8**	**88.0±0.7**	**72.1±1.1**	**93.3±5.6**

➢ 基线方法

半监督方法：DeepSAD、DROCC

无监督方法：DeepSVDD、ALOCC、USAD、DAGMM、GANF、MTGFlow

实验-消融研究

Table 2: The results of ablation studies on the SWaT dataset to discuss the number of experts, which is the important hyperparameter of our method.

# of experts	Datasets	
	SWaT	WADI
1	86.2±0.7	92.5±1.2
2	86.8±0.4	93.4±0.6
3	**87.2±1.3**	**94.2±0.8**
4	85.6±2.1	92.6±0.7

Table 3: The results of ablation studies on SWaT and WADI datasets to discuss the effectiveness of our proposed components. MoE and MAR stand for the mixture of experts and memory-augmented routers, respectively.

Components		SWaT	WADI
MoE	MAR		
×	×	85.5±1.1	92.2±0.2
×	✓	86.2±1.2	92.9±1.0
✓	×	86.8±1.2	93.5±0.4
✓	✓	**87.2±1.3**	**94.2±0.8**

Table 4: We integrate our Graph-MoE into the three baseline methods of GANF, MTGFlow, and USD, and the results show the superiority of our model.

Methods	**SWaT**	**Improvement(Δ%)**
GANF	79.8±0.7	
GANF+Graph-MoE	**82.6±0.7**	Δ%=3.5
MTGFlow	84.8±1.5	
MTGFlow+Graph-MoE	**87.2±1.3**	Δ%=2.8
USD	90.2±0.9	
USD+Graph-MoE	**92.3±1.4**	Δ%=2.3

➢ 消融研究

1.Graph-MoE层数

2.混合专家（MoE）和记忆增强路由器（MAR）的有效性

3.即插即用特性

4 总结

总结

- **总结:** 为了应对现有基于GNN的MTS异常检测方法的局限性，本研究提出了一种基于无监督学习的MTS异常检测模型——Graph-MoE。 Graph-MoE通过引入混合专家模块实现节点特征的动态层内聚合，通过记忆增强路由器完成不同GNN层的动态层间聚合，实现了检测性能的提升。

谢谢大家!

汇报人: 韦浩文

时间：2025-3-16