Frequency-aware Generative Models for Multivariate Time Series Imputation

Xinyu Yang[1], Yu Sun[1,✉], Xiaojie Yuan[1], and Xinyang Chen[2,✉]

[1] College of Computer Science, DISSec, Nankai University, China
[2] School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, China

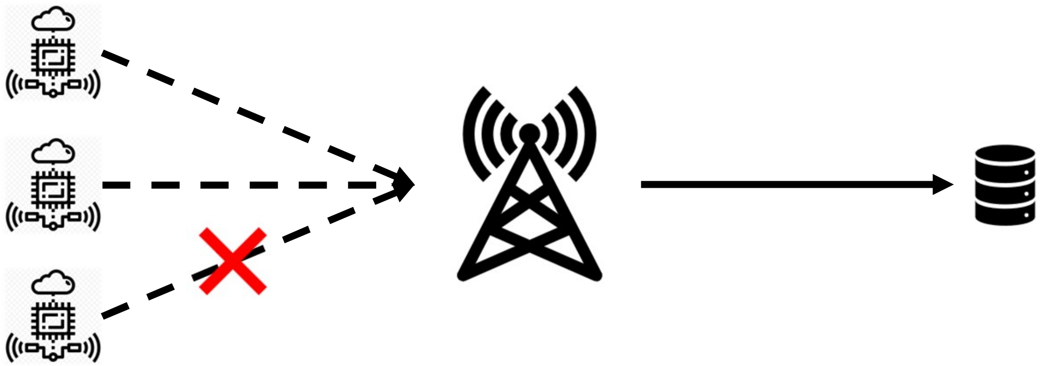
December, 2024

CONTENTS

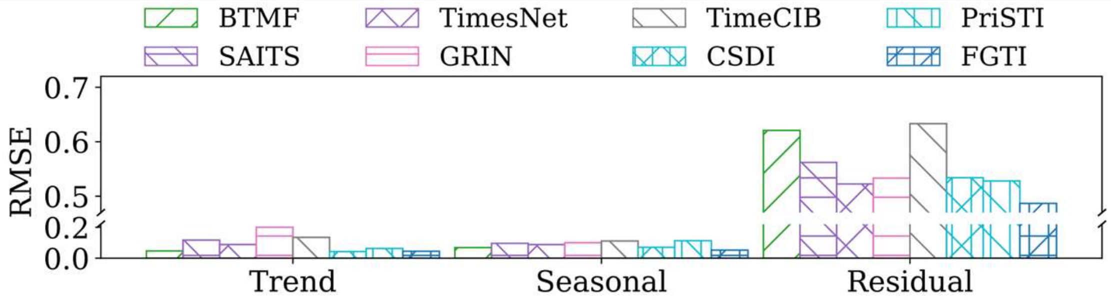
Background

- Multivariate time series appear in many applications
  - e.g. Air Quality, Traffic, Healthcare

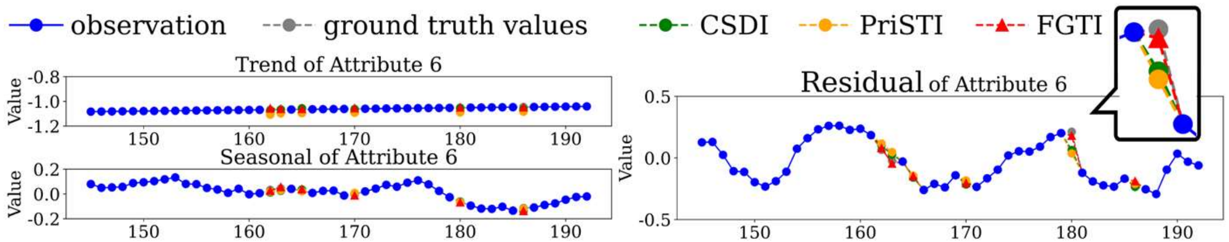
- Time Series data often contain missing values
  - could be harmful for downstream tasks



Background



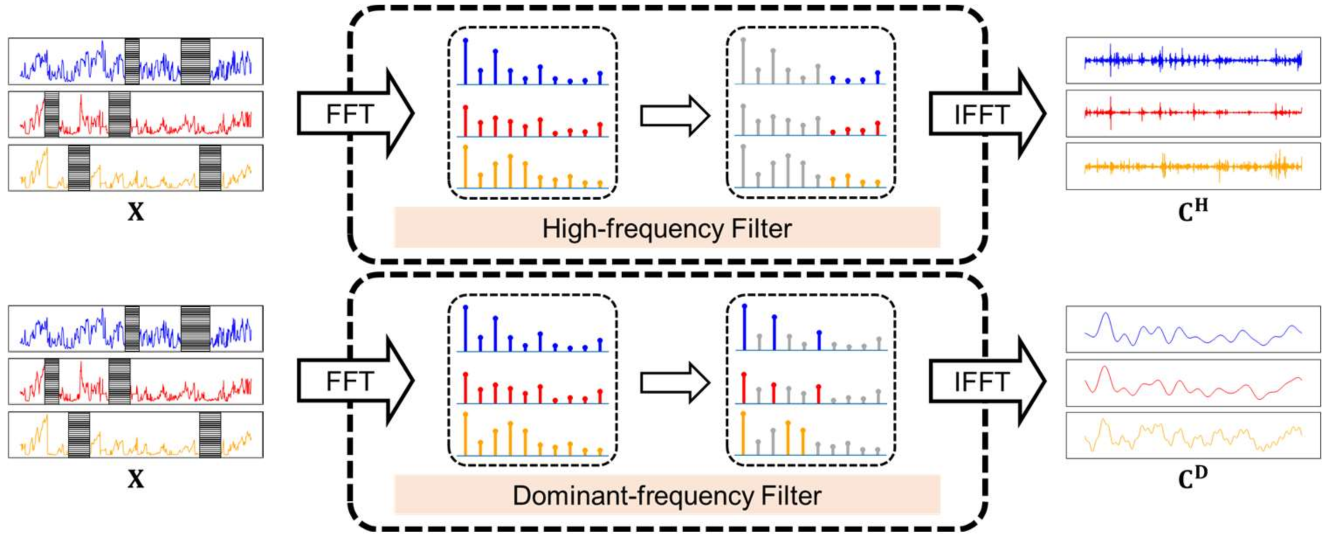
- Time Series data can be decomposed into three terms can be decomposed into three terms
- The imputation error is mainly caused by **Residual** term



- **High-frequency components** are intricately related to Residual
- Existing SOTA imputation methods with deep learning architectures **cannot generalize well** for high frequency components

FGTI Model

- Frequency-domain Condition Filters



- **High-frequency Filter:** Guide the imputation of residual terms
- **Dominant-frequency Filter:** Provide the background structure information to guide the imputation of the trend and seasonal terms

FGTI Model

➢ Implement FGTI with the current advanced generative model

Algorithm 1 Training process

Input: Incomplete time series $\mathbf{X}$, the number of diffusion step T

Output: Optimized denoising network $\epsilon_\theta(\cdot)$

repeat
- $\hat{\mathbf{X}}^0 \leftarrow$ select observed values in $\mathbf{X}$
- $t \sim \text{Uniform}\{1, \ldots, T\}$
- $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
- $\hat{\mathbf{X}}^t \leftarrow \sqrt{\overline{\alpha^t}}\hat{\mathbf{X}}^0 + \sqrt{1 - \overline{\alpha^t}}\epsilon$
- Perform Gradient Descent by $\nabla \mathcal{L}_\theta = \nabla_\theta \left\| \epsilon - \epsilon_\theta \left(t, \hat{\mathbf{X}}^t, \mathbf{X}^{\mathbf{C}}, \mathbf{C}^{\mathbf{H}}, \mathbf{C}^{\mathbf{D}} \right) \right\|^2$

until converged

Algorithm 2 Imputation process

Input: A incomplete time series sample$\mathbf{X}$, the number of diffusion step T, the optimized denoising network $\epsilon_\theta(\cdot)$

Output: Filled missing values $\hat{\mathbf{X}}^0$

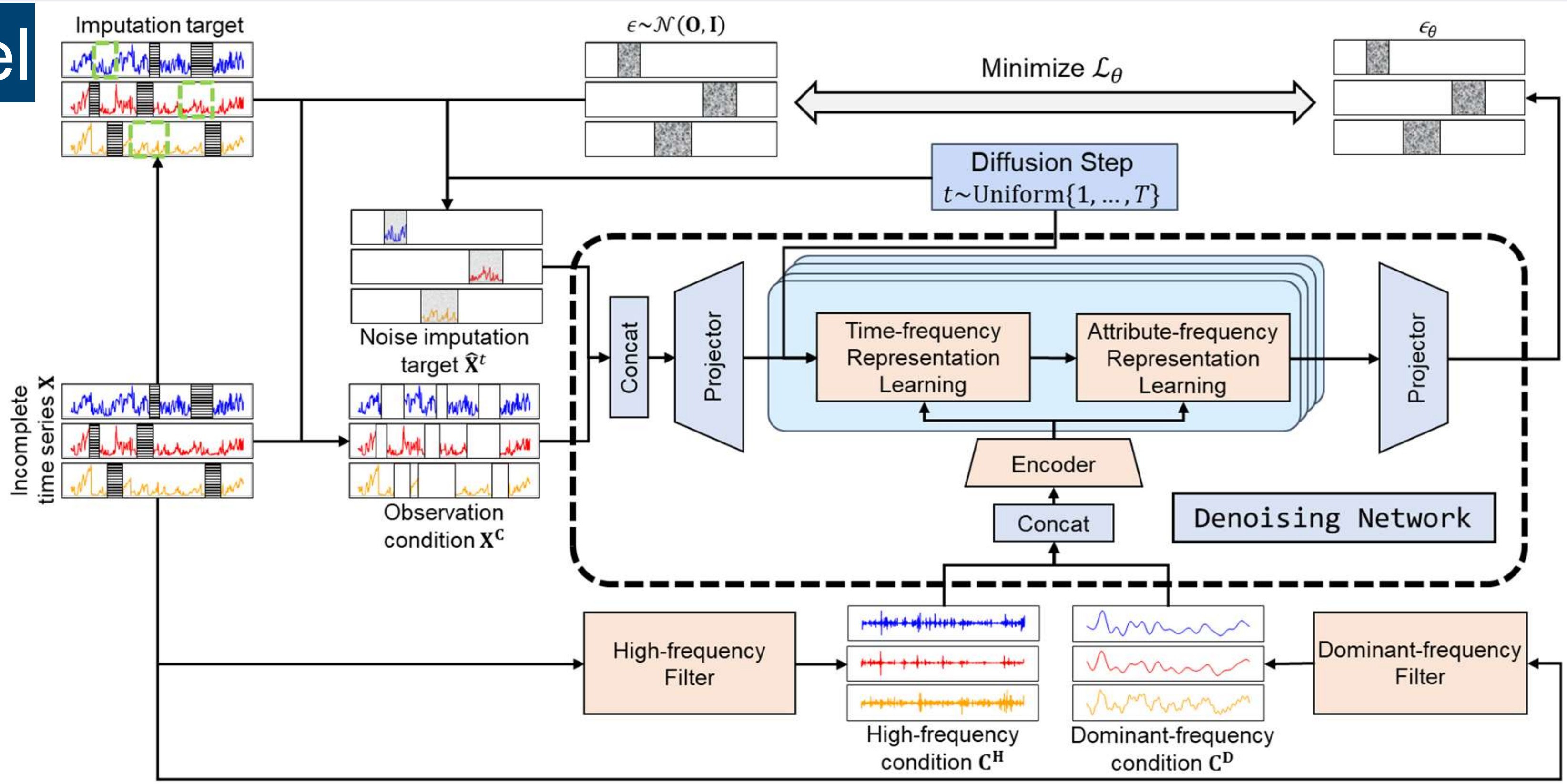
$\hat{\mathbf{X}} \leftarrow$ missing values in $\mathbf{X}$

$\hat{\mathbf{X}}^T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

for $t = T, \ldots, 1$ **do**
- **if** $t > 1$ **then**
  - $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
- **else**
  - $\epsilon \leftarrow \mathbf{0}$
- **end if**
- $$\hat{\mathbf{X}}^{t-1} \leftarrow \frac{1}{\sqrt{\overline{\alpha^t}}} \left[\hat{\mathbf{X}}^t - \frac{\beta^t}{\sqrt{1-\overline{\alpha^t}}} \epsilon_\theta \left(t, \hat{\mathbf{X}}^t, \mathbf{X}, \mathbf{C}^{\mathbf{H}}, \mathbf{C}^{\mathbf{D}} \right) \right] + \sqrt{\frac{(1 - \overline{\alpha^{t-1}})\beta^t}{1 - \overline{\alpha^t}}}\epsilon$$

end for

FGTI Model



➢ **Cross-domain Representation Learning Modules**

Fusing frequency-domain information to capture time dependencies and attribute dependencies by cross-attention mechanism

Experiments

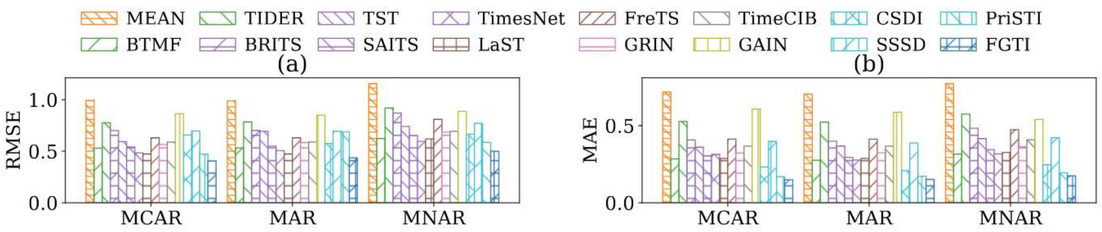
➢ Imputation results by different methods with different missing rates.

Dataset	Miss.	Metric	Mean	BTMF	TIDER	BRITS	TST	SAITS	TimesNet	LaST	FreTS	GRIN	TimeCIB	GAIN	CSDI	SSSD	PriSTI	FGTI
KDD	10%	RMSE	0.993	0.529	0.777	0.700	0.594	0.542	0.484	0.473	0.630	0.565	0.589	0.864	0.459	0.697	0.472	**0.406**
		MAE	0.718	0.285	0.527	0.407	0.360	0.304	0.313	0.287	0.412	0.322	0.367	0.607	0.177	0.397	0.169	**0.149**
	20%	RMSE	1.007	0.554	0.797	0.729	0.740	0.575	0.542	0.532	0.741	0.607	0.613	0.877	0.500	0.701	0.534	**0.451**
		MAE	0.718	0.286	0.531	0.416	0.371	0.310	0.307	0.310	0.489	0.339	0.369	0.606	0.187	0.392	0.180	**0.161**
	30%	RMSE	0.997	0.541	0.783	0.720	0.642	0.574	0.578	0.574	0.796	0.617	0.603	0.870	0.519	0.717	0.547	**0.448**
		MAE	0.717	0.286	0.528	0.420	0.376	0.319	0.357	0.350	0.546	0.360	0.370	0.612	0.199	0.413	0.195	**0.176**
	40%	RMSE	1.001	0.548	0.790	0.734	0.702	0.593	0.648	0.634	0.850	0.650	0.611	0.883	0.569	0.747	0.581	**0.478**
		MAE	0.718	0.287	0.532	0.428	0.387	0.332	0.418	0.393	0.591	0.387	0.372	0.623	0.220	0.435	0.217	**0.205**
Guang.	10%	RMSE	0.799	0.384	0.549	0.481	0.368	0.417	0.400	0.347	0.456	0.466	0.451	0.804	0.306	0.434	0.242	**0.230**
		MAE	0.592	0.252	0.392	0.299	0.249	0.264	0.270	0.244	0.340	0.354	0.300	0.550	0.210	0.293	**0.170**	**0.170**
	20%	RMSE	0.799	0.384	0.537	0.481	0.398	0.415	0.433	0.440	0.602	0.501	0.448	0.804	0.324	0.460	0.324	**0.258**
		MAE	0.592	0.252	0.382	0.300	0.275	0.264	0.303	0.312	0.460	0.385	0.298	0.550	0.220	0.315	0.197	**0.176**
	30%	RMSE	0.799	0.384	0.536	0.485	0.442	0.420	0.481	0.545	0.709	0.542	0.448	0.805	0.364	0.545	0.510	**0.291**
		MAE	0.592	0.252	0.382	0.301	0.312	0.267	0.348	0.388	0.547	0.419	0.298	0.551	0.242	0.384	0.271	**0.202**
	40%	RMSE	0.800	0.385	0.541	0.491	0.540	0.422	0.542	0.637	0.787	0.584	0.449	0.807	0.439	0.622	0.650	**0.356**
		MAE	0.592	0.253	0.387	0.306	0.397	0.270	0.401	0.458	0.611	0.455	0.299	0.554	0.283	0.444	0.381	**0.254**
Phy.	10%	RMSE	0.932	0.630	0.879	0.732	0.632	0.645	0.776	0.768	0.804	0.682	0.697	1.006	0.619	0.875	0.652	**0.580**
		MAE	0.678	0.348	0.605	0.446	0.389	0.371	0.525	0.516	0.540	0.424	0.450	0.747	0.310	0.528	0.369	**0.286**
	20%	RMSE	0.935	0.627	0.889	0.718	0.640	0.641	0.806	0.786	0.825	0.670	0.683	0.988	0.664	0.834	0.638	**0.577**
		MAE	0.675	0.362	0.624	0.451	0.417	0.384	0.569	0.550	0.576	0.434	0.455	0.740	0.335	0.507	0.376	**0.309**
	30%	RMSE	0.934	0.658	0.911	0.734	0.688	0.670	0.849	0.825	0.861	0.695	0.697	0.995	0.805	0.882	0.661	**0.624**
		MAE	0.676	0.382	0.638	0.457	0.452	0.404	0.600	0.578	0.603	0.446	0.459	0.738	0.360	0.545	0.387	**0.336**
	40%	RMSE	0.932	0.677	0.935	0.739	0.732	0.688	0.872	0.850	0.883	0.708	0.698	0.983	0.705	0.904	0.679	**0.669**
		MAE	0.677	0.412	0.658	0.466	0.493	0.431	0.623	0.603	0.626	0.464	0.466	0.729	0.395	0.555	0.406	**0.376**

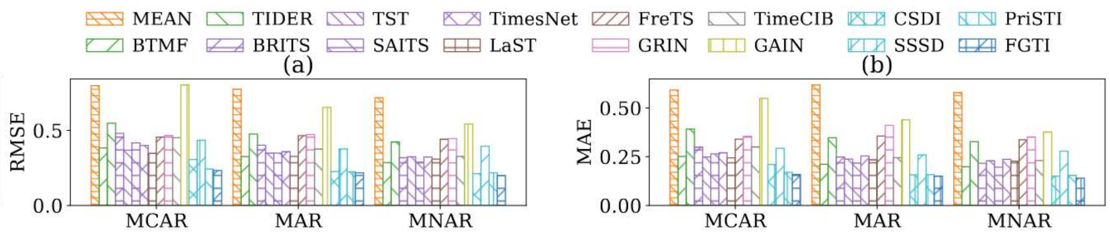
Experiments

➢ Imputation results by different methods with different missing mechanisms (10% missing)

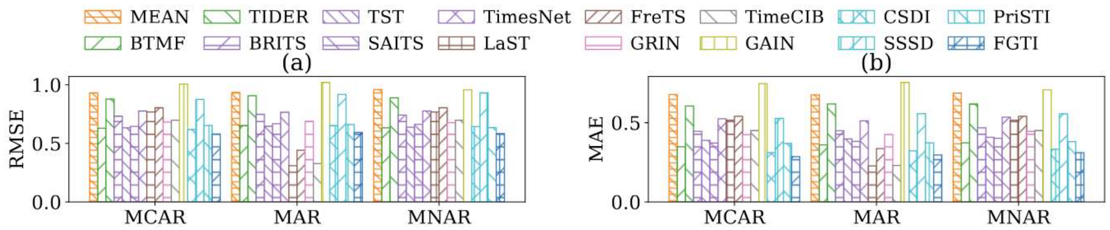
☐ Over KDD dataset



☐ Over Guangzhou dataset

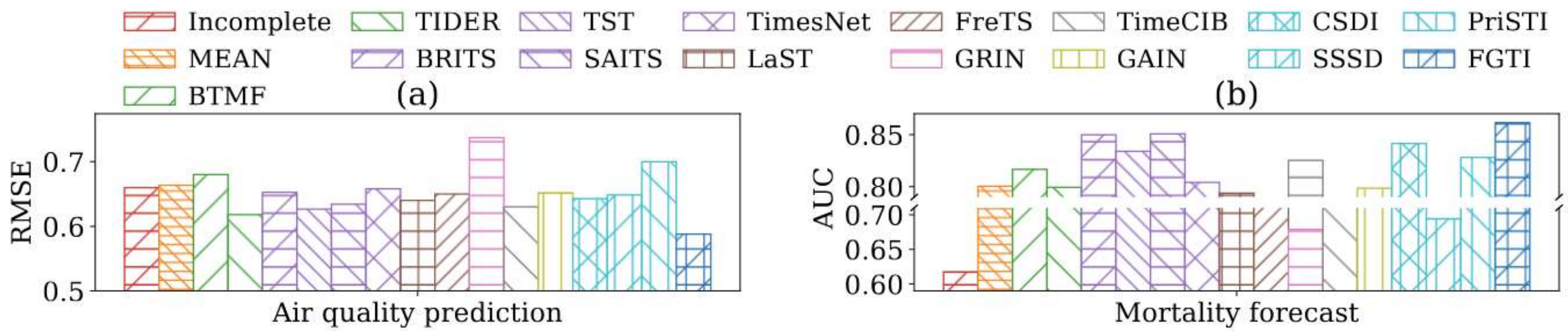


☐ Over PhysioNet dataset



Experiments

➢ Application performance of imputation methods for downstream tasks, over KDD and PhysioNet with real missing values



Conclusion

- We design a frequency-aware generative model FGTI with frequency-domain information integrated by the **high-frequency filter** and the **dominant-frequency filter**, to enhance the awareness of the frequency-domain for imputation.
- We introduce two **cross-domain representation learning modules** that provide models with prior knowledge of intricate frequency-related patterns for missing data imputation.
- We evaluate FGTI on three time series datasets with real-world missing values, which demonstrates the superiority of FGTI in both **imputation accuracy** and **downstream applications**

Thanks for your attention

If you are interested in our work, please check the details in our paper at the NeurIPS 2024.

Email: yangxinyu@dbis.nankai.edu.cn
https://github.com/FGTI2024/FGTI24