



# BRITS: Bidirectional Recurrent Imputation for Time Series

WEI CAO, DONG WANG, JIAN LI, HAO ZHOU, LEI LI, YITAN LI(NeurIPS,2018)

2025.05.15  
발표자: 박기웅

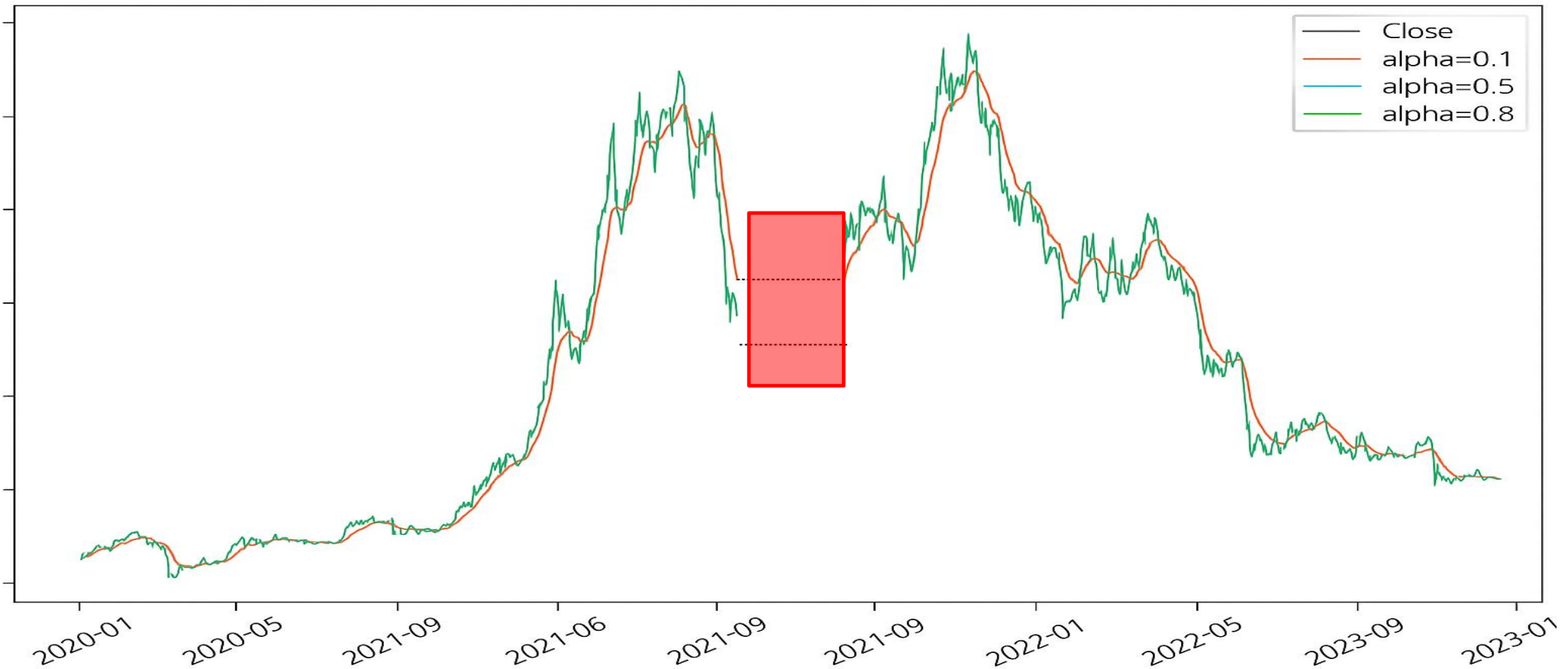
# CONTENTS

- 1. 배경 지식
- 2. 연구 배경 및 문제 정의
- 3. 핵심 아이디어
- 4. 모델 구조
- 5. 결과
- 6. 한계 및 의의

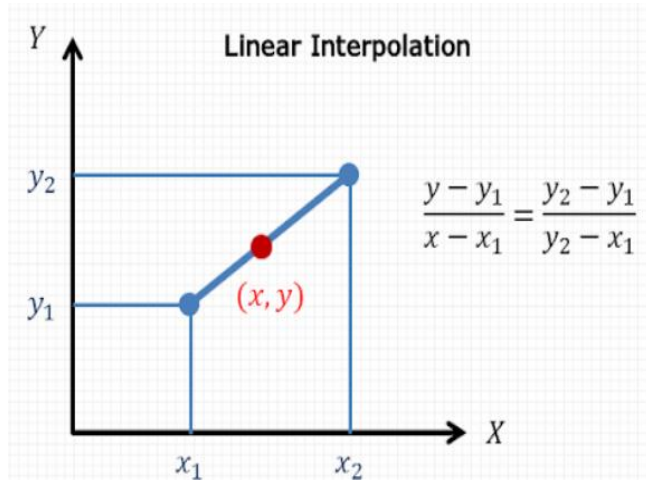
## 1. 배경 지식 - Time Series Interpolation

### 보간법(Interpolation)

주어진 데이터 포인트의 값을 사용하여, 두 개 이상의 데이터 포인트 사이에서 누락된 값을 추정하는 기법

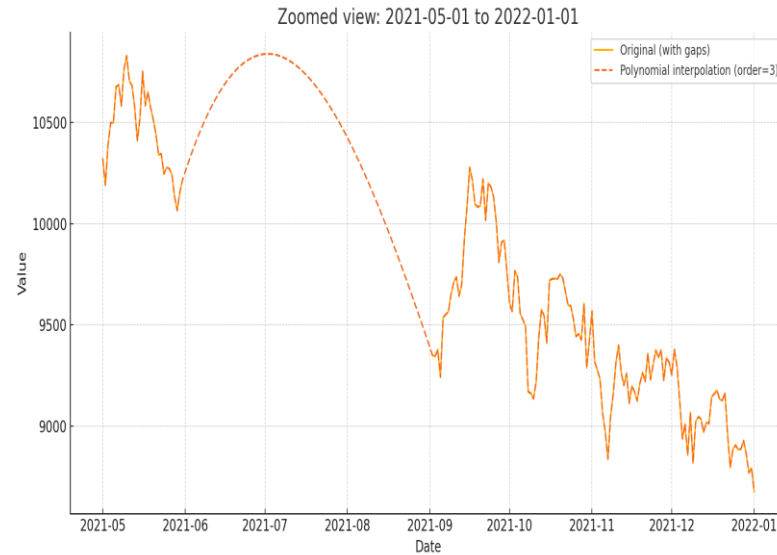


## 1. 배경 지식 - 기존의 Time Series Imputation(수학적)



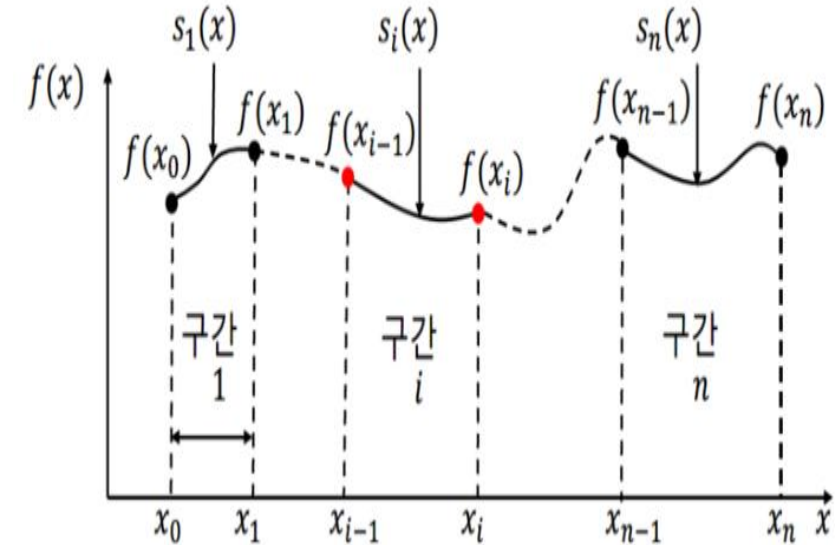
Linear Interpolation

누락 값을 직선 방정식을 사용하여 추정(두 데이터 포인트 간 선형 관계를 가정)



Polynomial Interpolation

다항식을 사용하여 데이터 포인트 간의 곡선 관계를 근사

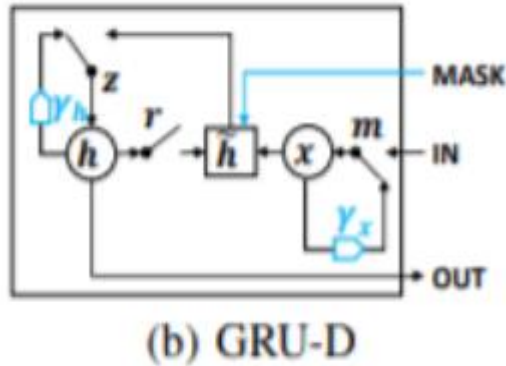


Spline Interpolation

보간 대상이 되는 구간을 여러개의 구간으로 나누고, 이를 저차 다항식인 스플라인 함수로 구성하여 해결하는 방식으로, 각 구간마다 서로 다른 스플라인 함수를 가진다.

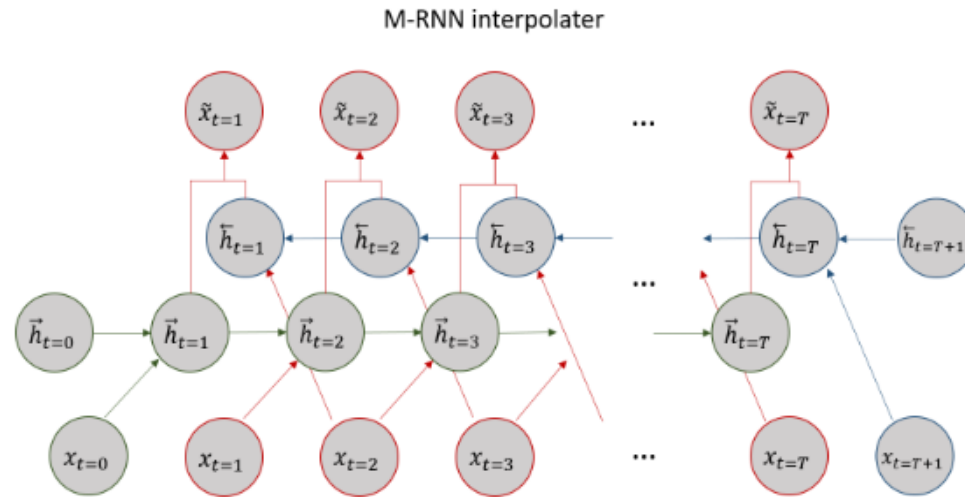
+ 평균 대체, 이전값 대체, 중앙값 대체 등등..

## 1. 배경 지식 - 기존의 Time Series Imputation(RNN)



GRU-D (2016)

- Decay rate를 이용하여 결측치를 대체하고, 결측치를 채운 값이라는 표시인 mask를 학습과정에 포함하여 학습 진행
- 마지막으로 관측된 값 & 평균으로 결측치를 대체하므로, 절대적인 평균(체온, 맥박, 혈압)이 있는 데이터에만 사용 가능.



M-RNN (2017)

- 각 변수마다 전·후방 RNN을 돌려 과거와 미래 정보를 동시에 끌어와 결측 구간을 보간한다.
- 고차원 및 장기 시계열에서 훈련시간 매우 급증
- 규칙적인 시간간격을 가정
- 오차 누적
- 같은 시점의 다른 변수들을 읽는 별도 RNN을 더해 상관관계를 보완하고, 두 예측을 게이트로 합성해 최종 값을 만든다.

## 2. 연구 배경 및 문제 정의

기존 연구의 문제점

- 통계 & RNN 기반 보간은 결측치에 대한 강력한 가정이 필요하다. (특수한 상황에서만 가능)
- 결측값을 단순 파라미터처럼 다뤄, 역전파로 직접 갱신할 수 없음
- 대부분의 결측치는 MCAR(완전 무작위 결측) 인 경우가 많기에 한계가 있음

## 3. 핵심 아이디어

결측값 자체를 학습 변수로 취급 + Bidirectional 순환

장점

1. 어떠한 가정도 필요없이 결측값을 바로 imputation 가능
2. 시계열 데이터에서 여러개의 결측치를 다룰 수 있다.
3. 길이가 짧은 결측 구간에 대해서 높은 정확도로 예측 가능

## 4. 모델 구조

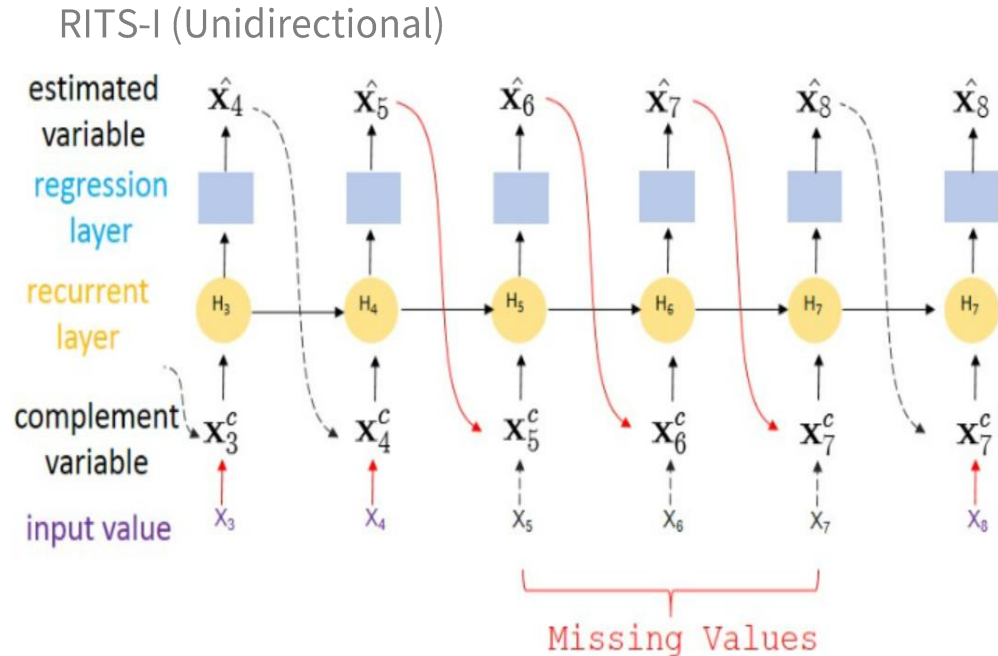


Figure 2: Imputation with unidirectional dynamics.

- 각각의 time step( $x_3, x_4, \dots$ )에서  $t-1$  step의 값을 이용하여  $t$ 의 값을 추정할 수 있다.
- 그러나,  $t$ 가 5, 6, 7 (결측 구간) 이면, 실제 값이 없기 때문에 error를 즉시 계산 할 수 없다. 따라서  $x_8$ 에서는  $x_5, x_6, x_7$ 의 지연된 에러를 얻는다.

$$\hat{x}_t = W_x h_{t-1} + b_x, \quad (1)$$

$$x_t^c = m_t \odot x_t + (1 - m_t) \odot \hat{x}_t, \quad (2)$$

$$\gamma_t = \exp\{-\max(0, W_\gamma \delta_t + b_\gamma)\}, \quad (3)$$

$$h_t = \sigma(W_h[h_{t-1} \odot \gamma_t] + U_h[x_t^c \odot m_t] + b_h), \quad (4)$$

$$l_t = \langle m_t, \mathcal{L}_e(x_t, \hat{x}_t) \rangle. \quad (5)$$

- 이를 해결하기 위해 위의 식을 사용한다.
- 단방향 GRU + 시간지연( $\Delta t$ ) 임베딩으로, 이전 시점 상태를 보정하면서 결측치를 직전 예측치로 임시 대체한다.

## 4. 모델 구조

RITS-I의 문제점

- RITS-I에서 결측치의 추정된 에러는 다음 관측이 있을때 까지 지연된다.
- 이러한 에러 지연은 모델 수렴을 느리게하고 학습을 비효율적으로 하며 bias exploding 문제를 일으키기도 한다.

BRITS-I (Bidirectional)

$$\ell_t^{cons} = \text{Discrepancy}(\hat{\mathbf{x}}_t, \hat{\mathbf{x}}'_t)$$

- 이를 해결하기 위해 , consistency loss를 도입하여 각 단계의 예측을 양방향으로 일관되게 시행한다.
- Ex ) bidirectional recurrent dynamics에서는  $x_4$ 가 반대로  $x_5 \sim x_7$ 에 영향을 받는다. 따라서 5번째 step의 에러가 8번째 step의 forward direction뿐만 아니라 4번째 step의 backward direction에도 영향을 받게 된다.
- 그러면 sequence도 forward, backward 두개가 생기고 그거에 따른 loss sequence 도 두개가 생긴다.
- 최종적인 추정 loss는 forward loss, backward loss, consistency loss를 모두 합쳐서 얻는다.

순방향/역방향 두개의 RITS를 공유 가중치로 학습하고, 추가 제약을 통해 과거,미래 패턴 모두 활용 -> 더 정확



## 5. 결과

Table 1: Performance Comparison for Imputation Tasks (in MAE(MRE%))

Method		Air Quality	Health-care	Human Activity
Non-RNN	Mean	55.51 (77.97%)	0.720 (100.00%)	0.767 (96.43%)
	KNN	29.79 (41.85%)	0.732 (101.66%)	0.479 (58.54%)
	MF	27.94 (39.25%)	0.622 (87.68%)	0.879 (110.44%)
	MICE	27.42 (38.52%)	0.634 (89.17%)	0.477 (57.94%)
	ImputeTS	19.58 (27.51%)	0.390 (54.2%)	0.363 (45.65%)
	STMVL	12.12 (17.40%)	/	/
RNN	GRU-D	/	0.559 (77.58%)	0.558 (70.05%)
	M-RNN	14.24 (20.43%)	0.451 (62.65%)	0.248 (31.19%)
Ours	RITS-I	12.73 (18.32%)	0.395 (54.80%)	0.240 (30.10%)
	BRITS-I	11.58 (16.66%)	0.361 (50.01%)	0.220 (27.61%)
	RITS	12.19 (17.54%)	0.300 (41.89%)	0.248 (31.21%)
	<b>BRITS</b>	<b>11.56 (16.65%)</b>	<b>0.281 (39.14%)</b>	<b>0.219 (27.59%)</b>

- 결측 비율 50% 이상 상황에서도 성능 붕괴가 적다
- 추론속도는 GRU-D 수준으로 실용적이나, 성능이 21% 정도 개선되었다.

## 6. 한계 및 의의

- 긴 시계열( $l > 1k$ )에서 GRU 메모리 한계. → 구간별 Attention (2021 Interpolation Transformer)로 확장 연구.
- 외생 변수, 그래프 구조 활용 미흡 → 2022 GraphBRITS에서 개선.
- 확률적 불확실성 추정은 제한적 → 2021 CSDI, 2024 DiffImp 등이 디퓨전 기반으로 보완.
- 길이가 길어질 수록 기존의 정보가 희미해져서 정보 손실 -> 2019 NAOMI에서 보완

