MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks

Dan Li1, Dacheng Chen1, Lei Shi1, Baihong Jin2, Jonathan Goh3, and See-Kiong Ng1

nstitute of Data Science, National University of Singapore, 3 Research Link Singapore 117602
2 Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720 USA
3 ST Electronics (Info Security) Pte Ltd, 100 Jurong East Street 21 Singapore 609602

Cham: Springer International Publishing, 2019.

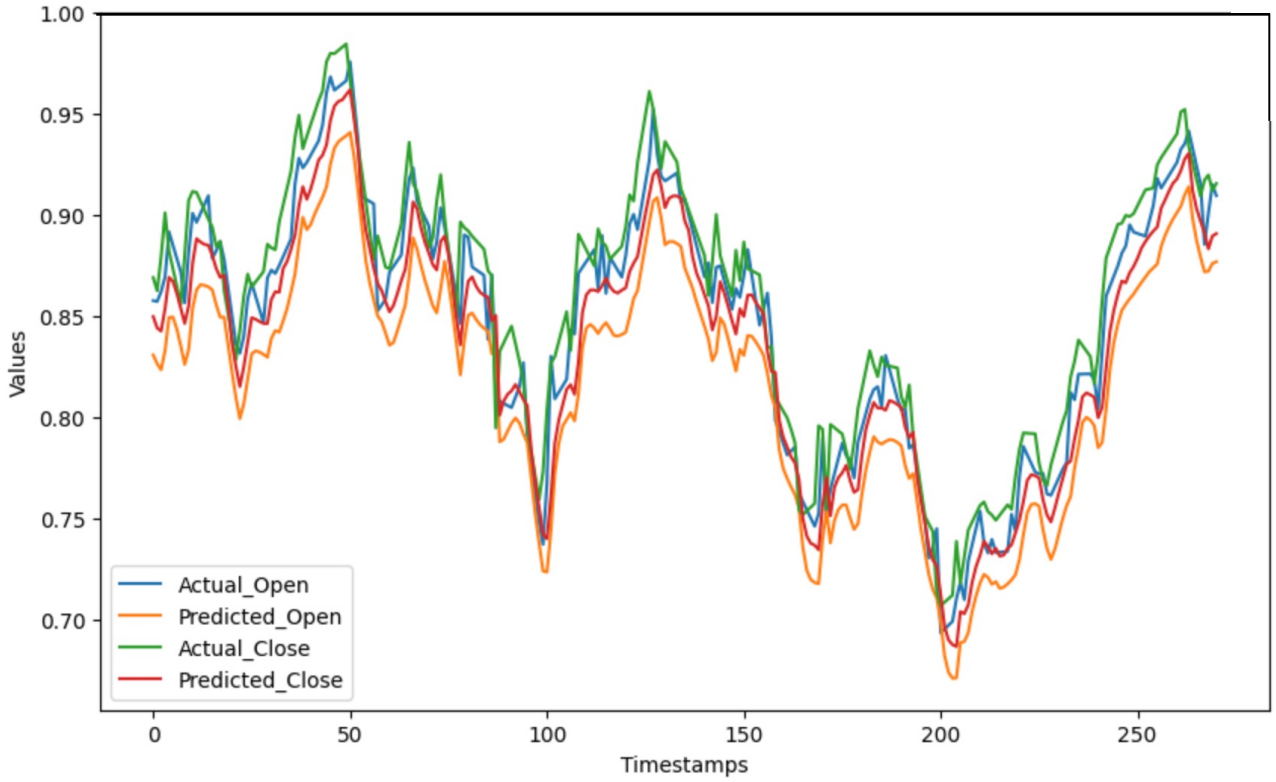
Graduate School
Department of Computer Science and Engineering
Chung-Ang University
DMAIS Lab
Master Course Park Sae Joon

CONTENTS

- Backgrond
- Why?
- Previous Work
- MAD-GAN
- Experiments
- Conclusions

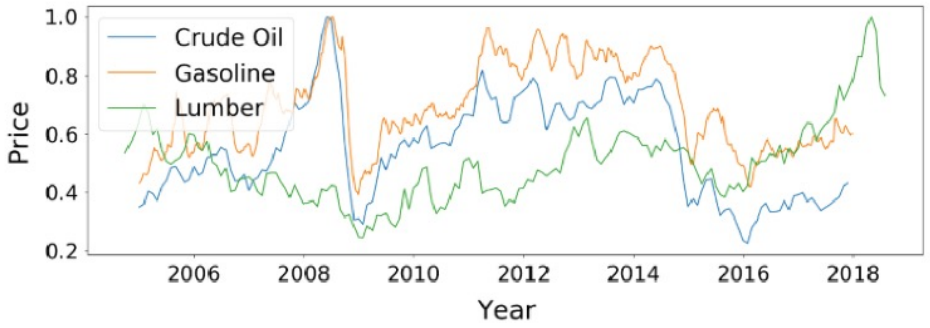
BACKGROUND

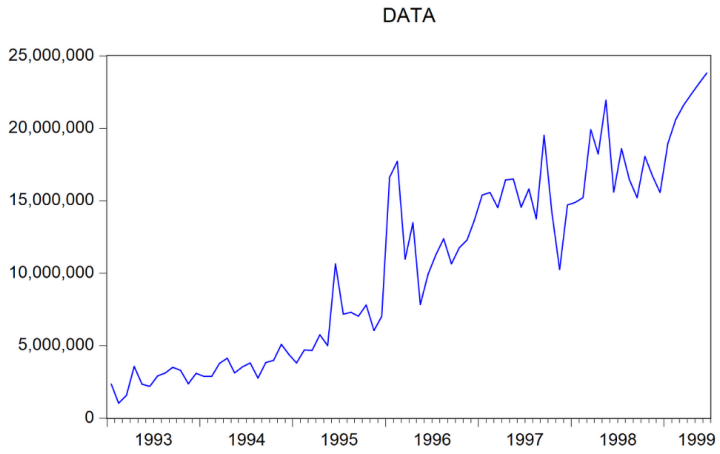
- **Time Series**
  - 일정한 시간 동안 수집 된 일련의 순차적으로 정해진 데이터 세트의 집합



BACKGROUND

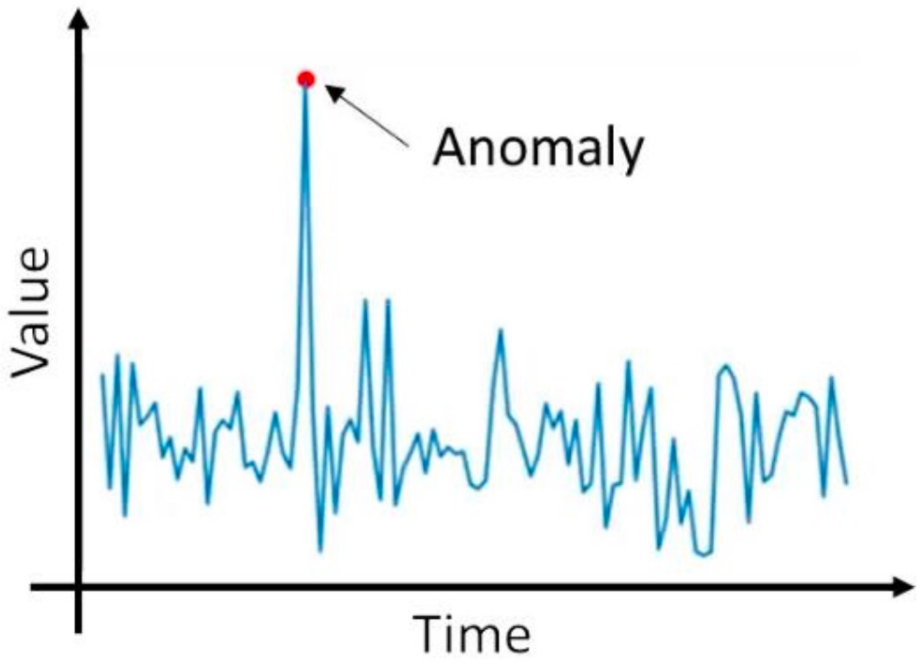
- **Multivariate & Univariate Time Series**





BACKGROUND

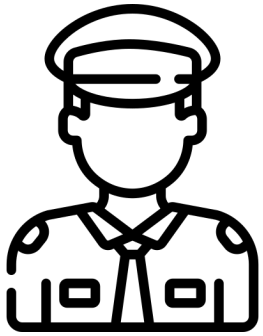
- **Anomaly Detection**
  - 데이터 세트에서 정상적인 패턴과 일치하지 않는 비정상적이거나 예외적인 패턴을 식별하는 것



BACKGROUND

- **Generate Adversarial Nets**
  - Generative model과 Discriminative model로 이루어짐
  - 두 모델이 adversarial training을 통해 학습을 진행

VS



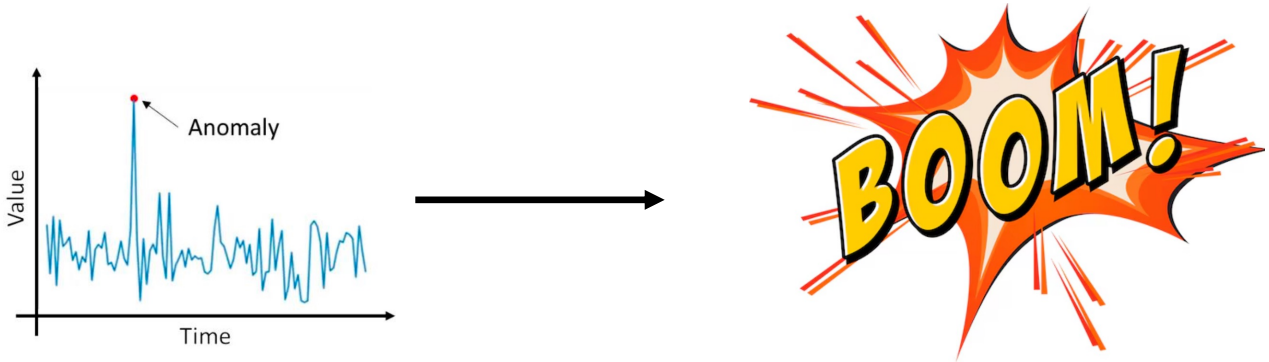
Generative model

Discriminative model

$$\min_G \max_D V(D,G) = \mathcal{E}_{x \sim p_{data}(X)} [\log D(x)] + \mathcal{E}_{z \sim p_z(Z)} [\log(1 - D(G(z)))]$$

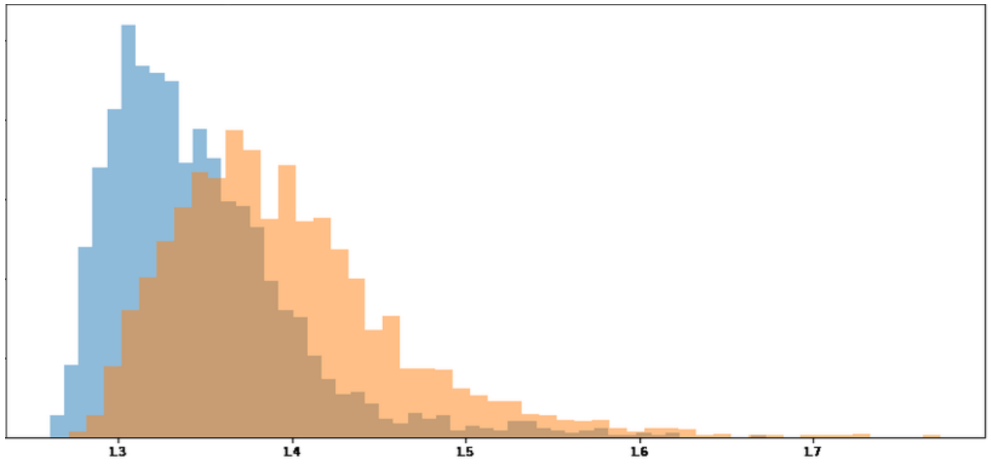
WHY?

- **Why Anomaly Detection in time series is important?**
  - 다양한 센서와 같은 것들이 네트워크를 통해 연결돼 실시간으로 방대한 데이터 생성
  - 시계열 데이터는 Temporal Dependency를 가짐
  - CPS데이터는 동적이고, 정상 패턴이 변경될 수 있음
  - 이 데이터를 실시간으로 분석하는 것이 중요
  - 센서와 같은 것들이 조금의 문제가 생길 경우 전체에 심각한 영향을 줄 수 있음
  - 신뢰성과 보안성을 유지하기 위한 필수 요소

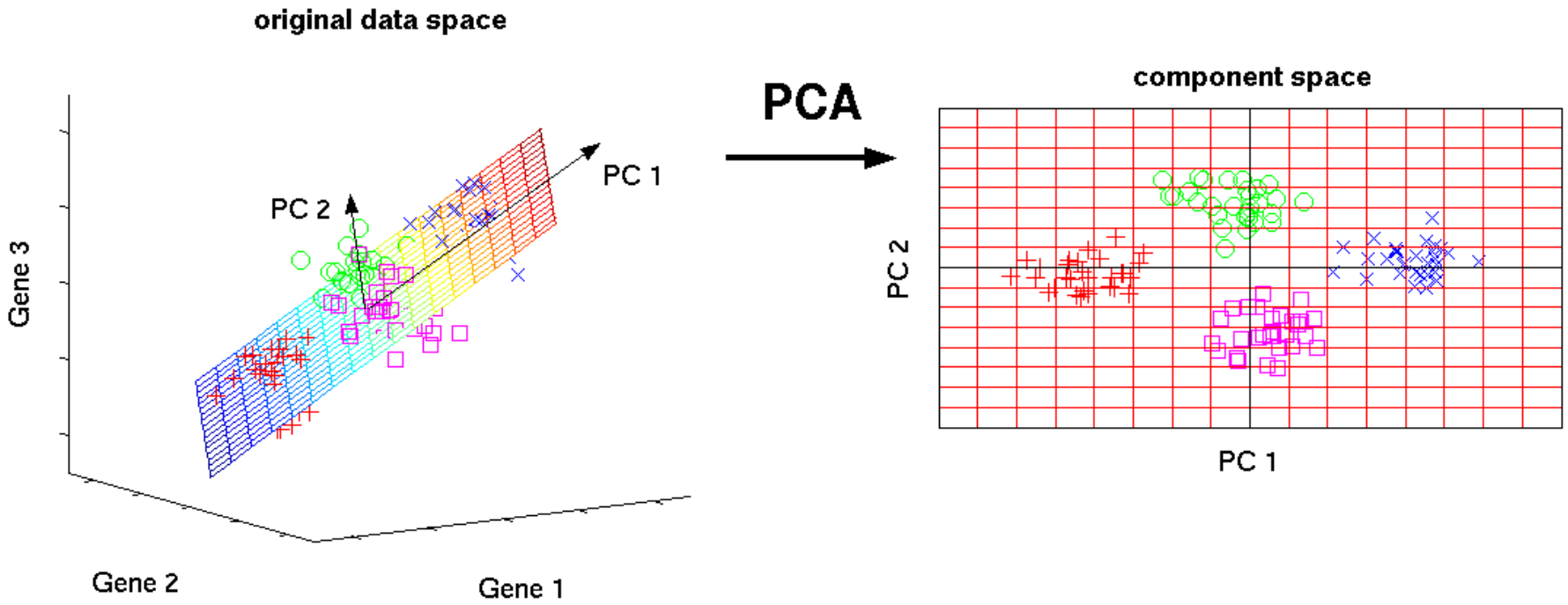


Reconstruction Based TSAD

- **Reconstruction based TSAD**
  - 실제 데이터와 Reconstruction된 데이터를 비교하여 anomaly를 찾아내는 방법
  - Anomaly면 실제 데이터와 차이가 클 것
  - Unsupervised Learning이 가능함
  - 시계열 데이터의 Temporal Dependency 반영 필요



PREVIOUS WORK

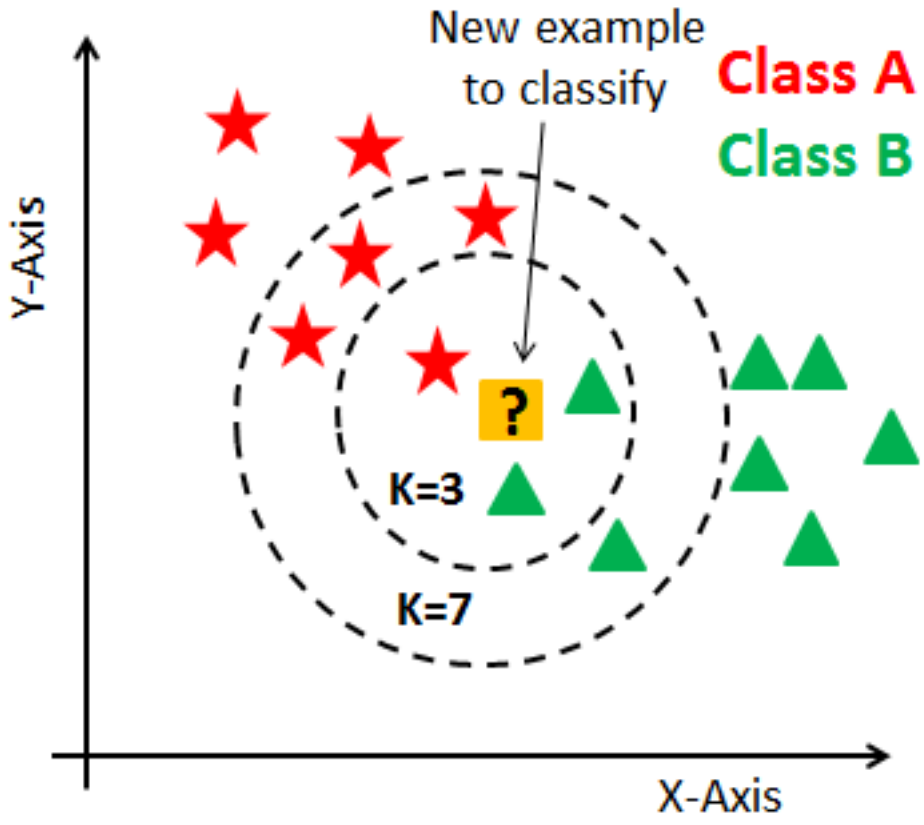


PREVIOUS WORK

- **Linear Based Model Methods**
  - 대표적으로 PCA, PLS등이 있음
  - 정상데이터의 차원을 축소
  - Principle Component를 구한 뒤, PC로 이루어진 공간을 기준으로 정상, 비정상 판단

- **Limits**
  - PCA의 경우 변수간 연관성이 높은 자료를 요구
  - PLS는 다변량 가우시안 분포를 가정해야함
  - 적용할 수 있는 데이터가 한정적

PREVIOUS WORK

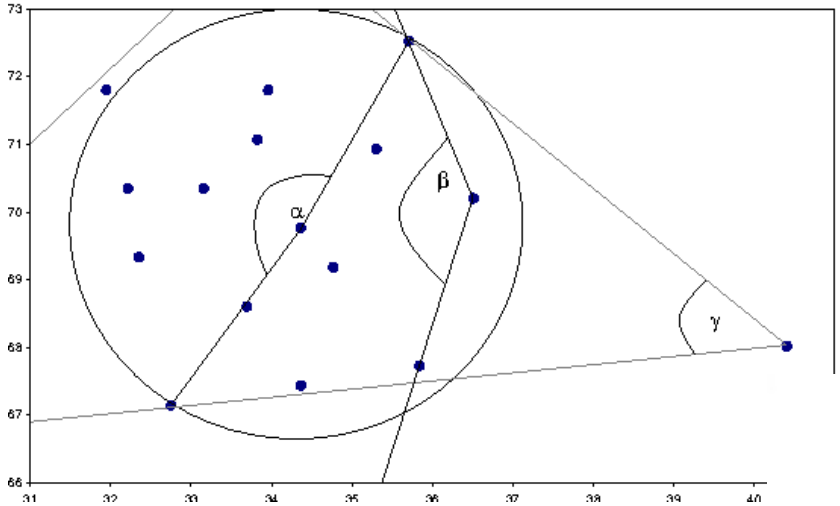


PREVIOUS WORK

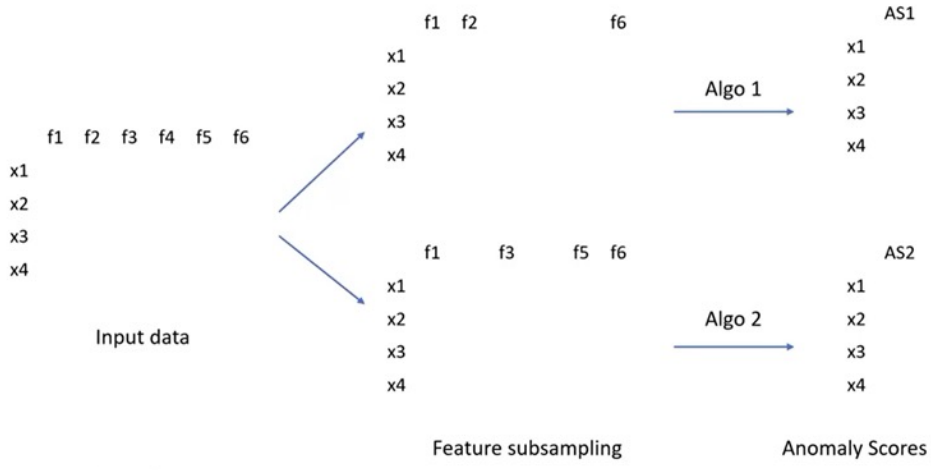
- **Distance Based Methods**
  - KNN, CBLOF등 거리 기반 이상치 탐지 기법
  - 평균 거리 혹은 가중 거리 기반으로 이상 탐지 점수를 할당
  - 데이터 클러스터링을 수행한 후, 특정 클러스터에 속하지 않는 이상치를 탐지

- **Limits**
  - 이상 탐지 성능이 데이터 분포에 따라 달라짐
  - 이상 지속 시간과 이상 개수를 사전에 알지 못하면 성능이 저하될 가능성이 있음
  - CPS 시스템에서는 이상 패턴이 동적으로 변화해 적절하지 않을 수 있음

PREVIOUS WORK



Feature bagging for anomaly detection

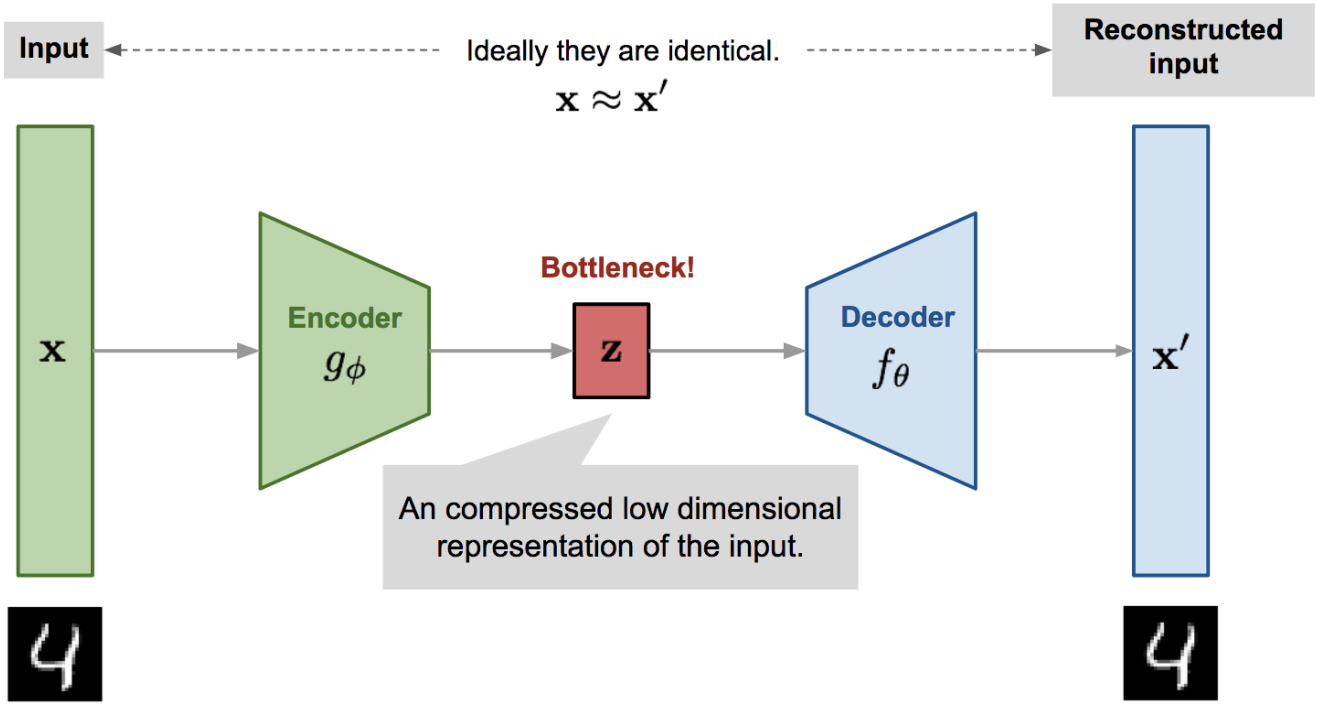


PREVIOUS WORK

- **Probabilistic & Density Estimation-Based**
  - ABOD, FB등이 있음
  - 확률 기반 방법론들은 데이터의 분포에 집중
  - 변수간 상관성만 고려

- **Limits**
  - 시간적인 상관성을 고려하지 않음
  - 다변량 시계열 데이터에 적합하지 않음

PREVIOUS WORK



PREVIOUS WORK

- **Deep Learning Based Methods**
  - Auto Encoder, DAGMM, LSTM Encoder-Decoder등이 있음
  - 다변량 이상 탐지 분야에서 좋은 성능을 보임

- **Limits**
  - Reconstruction loss만 사용해 이상을 탐지
  - 일부 이상치는 실제 데이터의 특징과 유사할 수 있음
  - 따라서 Reconstruction loss만 사용하면 불완전함

MAD-GAN

- **Definition**
  - Multivariate Anomaly Detection for Time Series Data with GAN

- **Characteristic**
  - Generator와 Discriminator를 두개의 LSTM-RNN으로 구성함
  - LSTM-RNN을 채용함으로써 temporal dependency를 반영함
  - Generator가 시계열 데이터를 만듬
  - DR-SCORE라는 고유의 loss function을 사용하여 Reconstruction, Discriminate loss를 둘다 사용

MAD-GAN

- **Architecture**

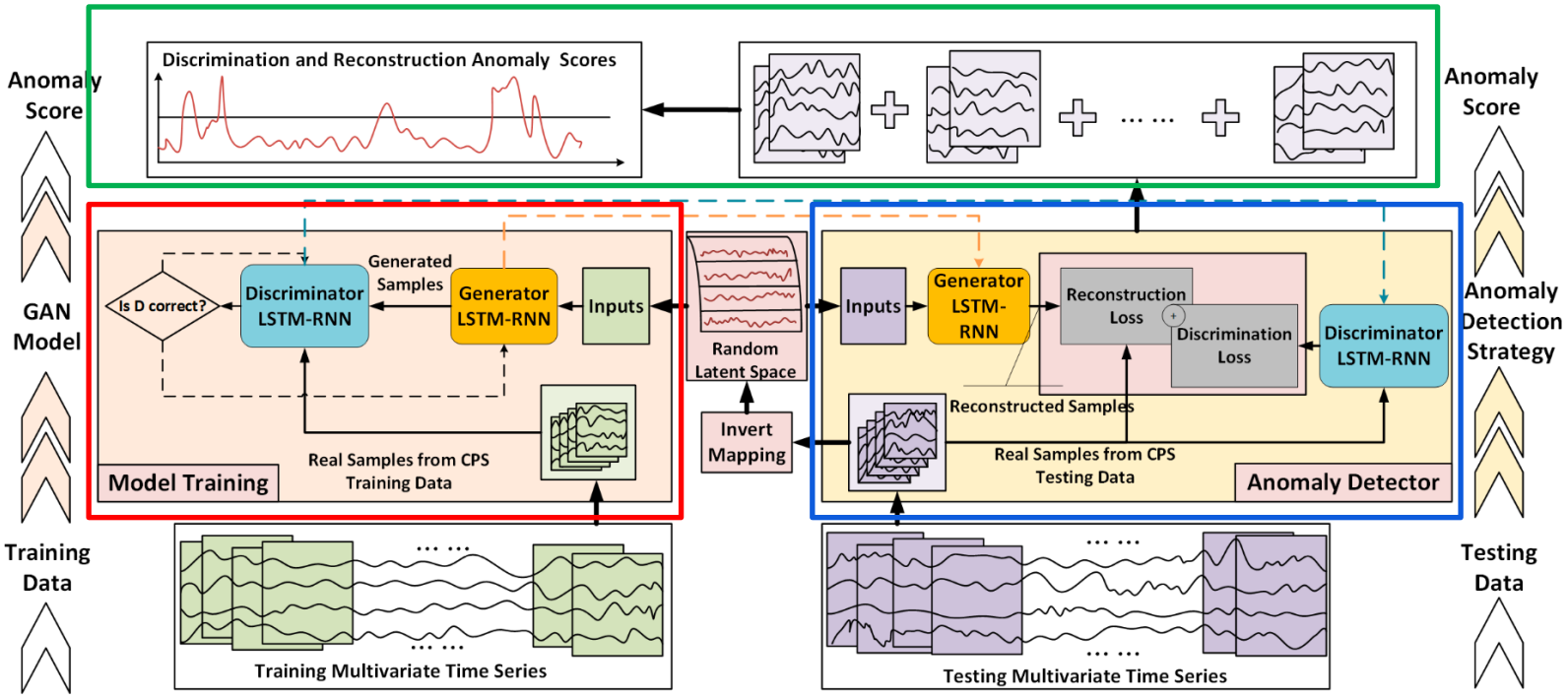
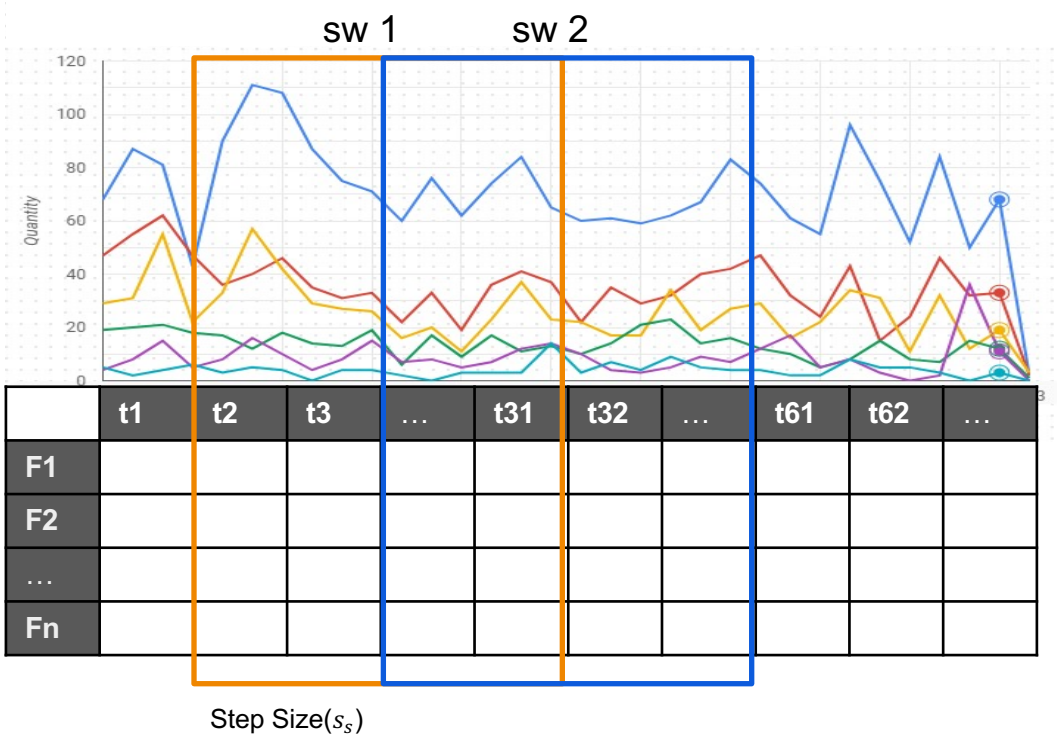
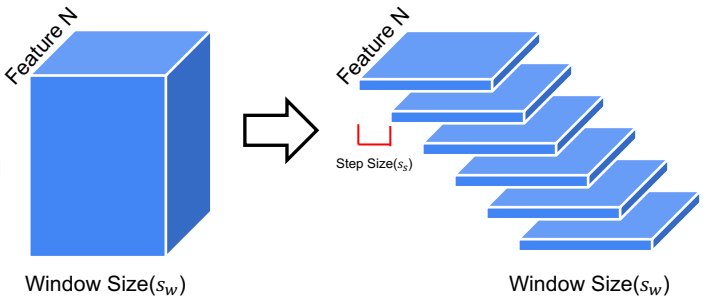


Fig. 1: MAD-GAN: Unsupervised GAN-based anomaly detection. On the left is a GAN framework in which the generator and discriminator are obtained with iterative adversarial training. On the right is the anomaly detection process where both the GAN-trained discriminator and generator are applied to compute a combined anomaly score based on discrimination and reconstruction.

MAD-GAN

- **Data Preprocessing**



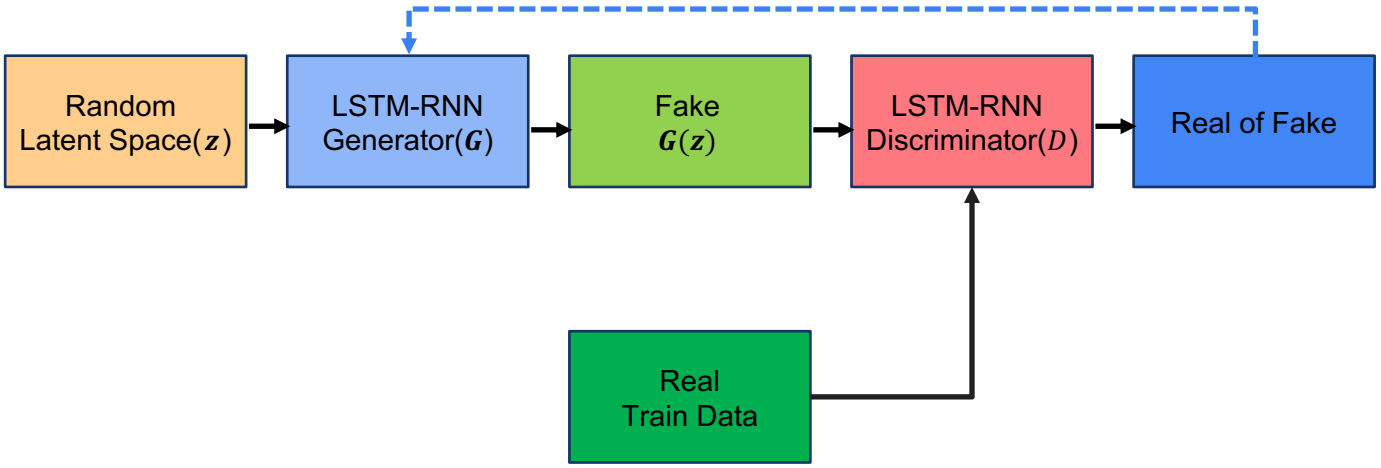


$$s_w = 30 \times i, i = 1, 2, ..., 10.$$

MAD-GAN

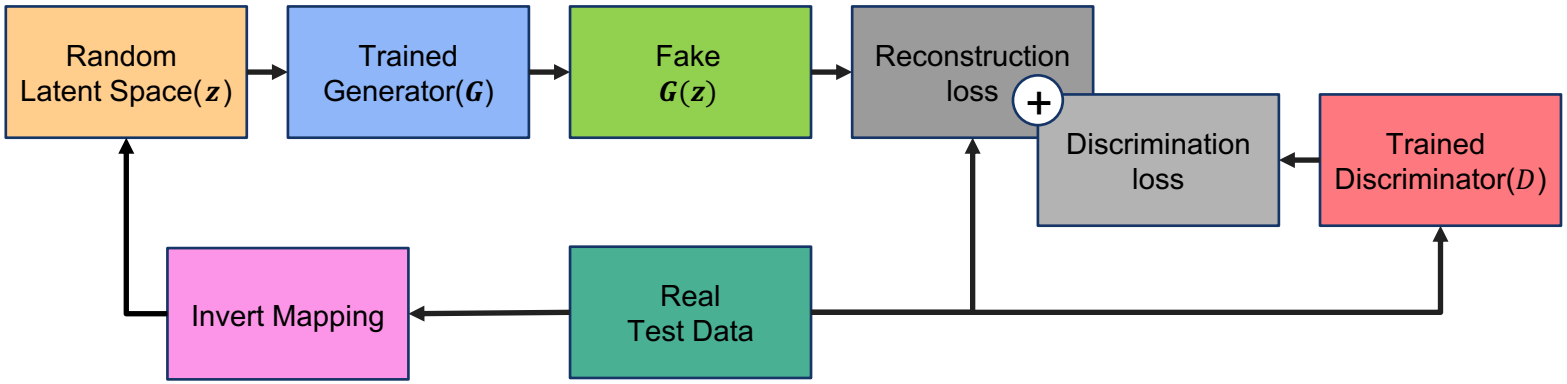
- **Train**

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$



MAD-GAN

- **Test**


Random
Latent Space(z)
Trained
Generator(G)
Fake
$G(z)$
Reconstruction
loss
+
Discrimination
loss
Trained
Discriminator(D)
Invert Mapping
Real
Test Data

MAD-GAN

- **DR Score**
  - 테스트 데이터 세트의 Sliding Window에서 아래와 같이 라벨링을 함

$$A_t^{tes} = \begin{cases} 1, if\ H\left(DRS_t, 1\right) > \tau \\ 0, \qquad\qquad else \end{cases}$$

    - 0은 정상데이터, 0이 아닌 데이터는 anomaly가 찾아진 것을 의미함
    - τ는 미리 정해두는 threshold
  - 최적의 Latent Space의 Z를 찾기 위해 아래와 같은 Loss function을 사용

$$\min_{Z^k} Er(X^{tes}, G_{rnn}(Z^k)) = 1 - Simi(X^{tes}, G_{rnn}(Z^k))$$

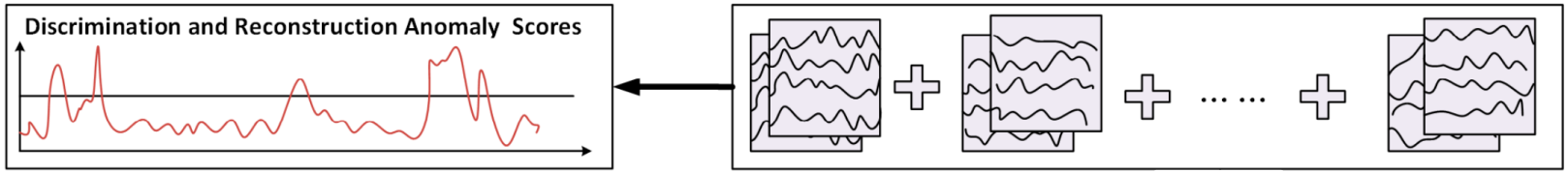
  - 충분한 반복 후에는 에러가 충분히 작을 것임 아래는 Reconstruction loss

$$Res(X_t^{tes}) = \sum_{i=1}^{n} \mid x_t^{tes,i} - G_{rnn}(Z_t^{k,i}) \mid$$

MAD-GAN

- **DR Score**

$$L_t^{tes} = \lambda \mathrm{Res}(\mathrm{X_t^{tes}}) + (1-\lambda)\mathrm{D_{rnn}}(\mathrm{X_t^{tes}})$$

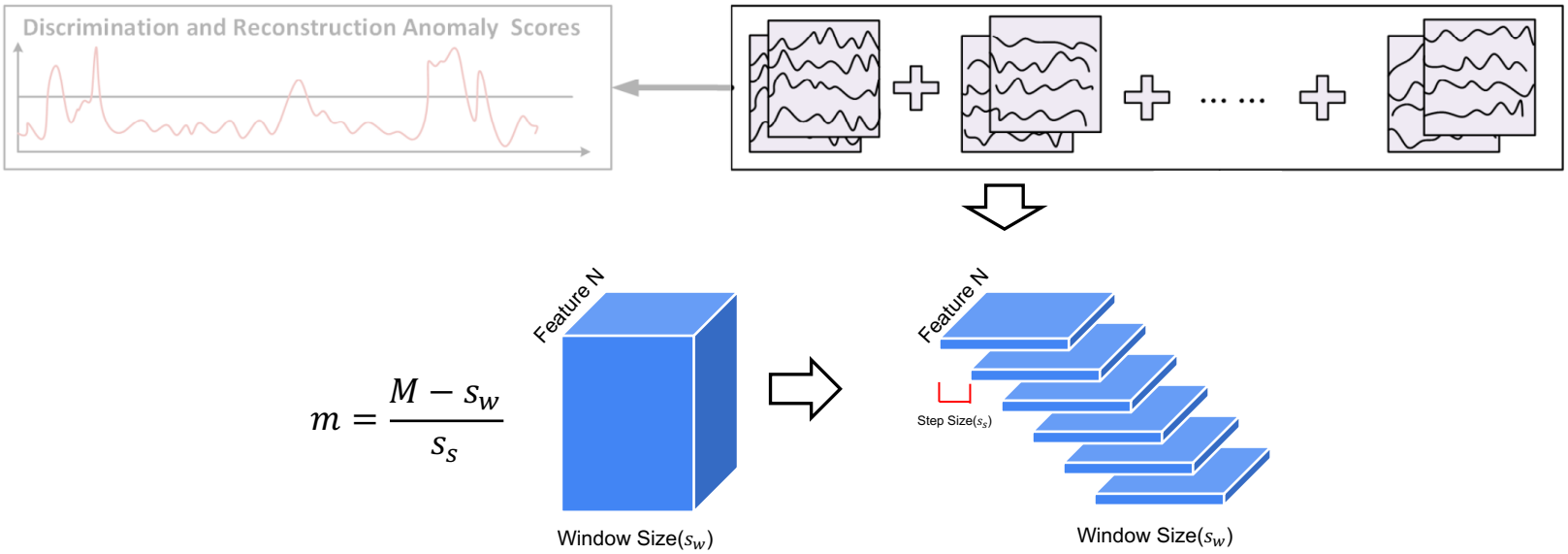


MAD-GAN

- **DR Score**

$$DRS_t = \frac{\sum_{j,s \in \{j+s=t\}} L_{j,s}}{lc_t}$$

$$lc_t = count(j, s \in \{j + s = t\}) \quad \text{s: window size, j: slide 횟수}$$

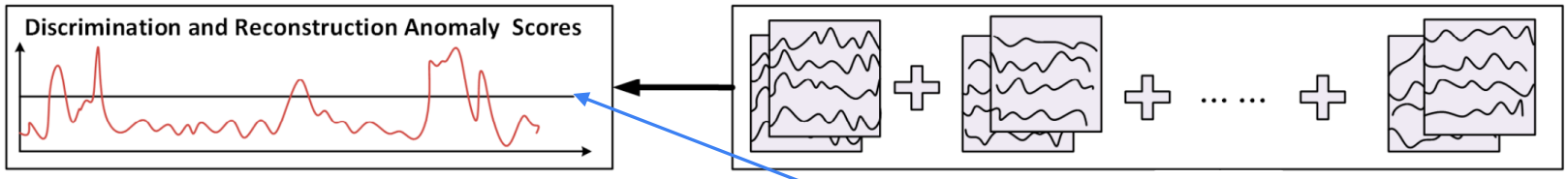


MAD-GAN

- **DR Score**

$$DRS_t = \frac{\sum_{j,s \in \{j+s=t\}} L_{j,s}}{lc_t}$$

$$lc_t = count(j, s \in \{j + s = t\}) \quad \text{s: window size, j: slide 횟수}$$



$$A_t^{tes} = \begin{cases} 1, if\ H\left(DRS_t, 1\right) > \tau \\ 0, \qquad else \end{cases}$$

MAD-GAN

Algorithm 1 LSTM-RNN-GAN-based Anomaly Detection Strategy

loop
  if epoch within number of training iterations **then**
    for the k^{th} epoch **do**
      Generate samples from the random space:
      $Z = \{z_i, i = 1, ..., m\} \Rightarrow G_{rnn}(Z)$
      Conduct discrimination:
      $X = \{x_i, i = 1, ..., m\} \Rightarrow D_{rnn}(X)$
      $G_{rnn}(Z) \Rightarrow D_{rnn}(G_{rnn}(Z))$
      Update discriminator parameters by minimizing(descending) D_{loss}:
      $\min \frac{1}{m} \sum_{i=1}^{m} [-\log D_{rnn}(x_i) - \log(1 - D_{rnn}(G_{rnn}(z_i)))]$
      Update discriminator parameters by minimizing(descending) G_loss :
      $\min \sum_{i=1}^{m} \log(-D_{rnn}(G_{rnn}(z_i)))$
      Record parameters of the discriminator and generator in the current iteration.
    end for
  end if
  for the lth iteration **do**
    Mapping testing data back to latent space:
    $Z^k = \min_Z Er(X^{tes}, G_{rnn}(Z^i))$
  end for
  Calculate the residuals:
  $Res = \mid X^{tes} - G_{rnn}(Z^k) \mid$
  Calculate the discrimination results:
  $Dis = D_{rnn}(X^{tes})$
  Obtain the combined anomaly score:
  for k,j and s in ranges **do**
    if j+s=k **then**
      $R = \lambda Res + (1 - \lambda) Dis$
      $DRS_k = \frac{\sum R_{j,s}}{L_k}$
    end if
  end for
end loop

EXPERIMENTS

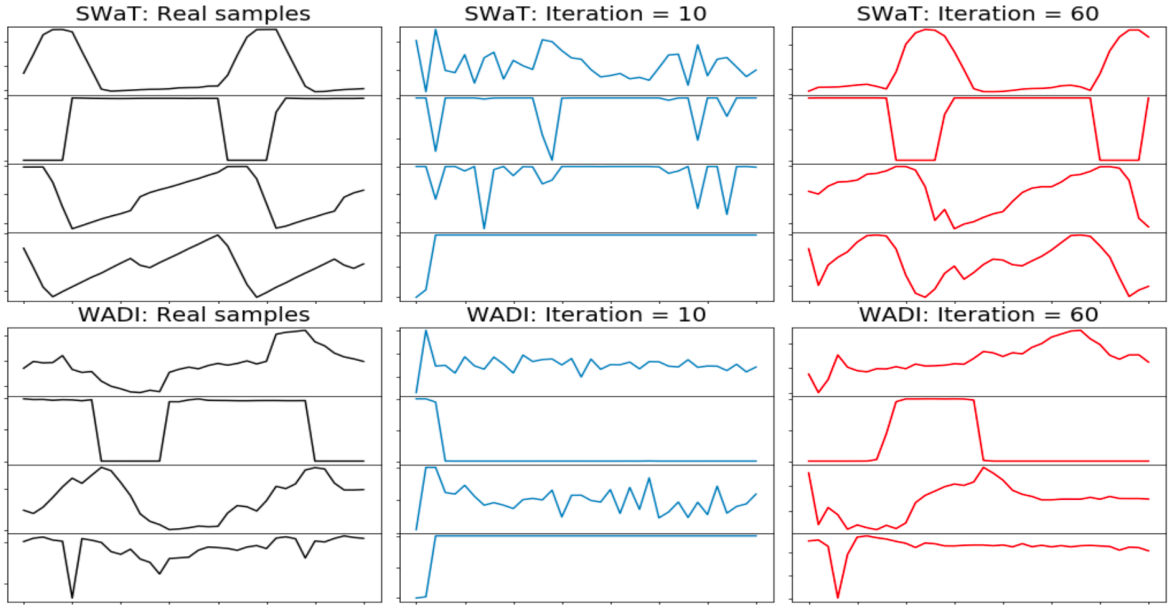


Fig. 2: Comparison between generated samples at different traning stages: GAN-generated samples at early stage are quite random while those generated at later stages almost perfectly took the distribution of original samples. Note that we only plot four variables for each dataset as visualization examples.

- **Generate Data**
  - SWaT과 WADI에서 데이터 생성의 결과를 보여줌
  - 초기 학습 단계에선 랜덤한 데이터가 생성, 학습이 진행될수록 실제 데이터 분포와 유사
  - Maximum Mean Discrepancy (MMD)지표로 평가

EXPERIMENTS

- **MMD**
  - Maximum Mean Discrepancy (MMD)
  - GAN모델이 학습데이터의 분포 학습을 평가하는 지표

$$MMD(Z_j, X_{tes}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i}^{n} K(Z_i^k, Z_j^k) - \frac{2}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} K(Z_i^k, X_j^{tes}) + \frac{1}{m(m-1)} \sum_{i=1}^{m} \sum_{j \neq i}^{m} K(X_i^{tes}, X_j^{tes})$$

EXPERIMENTS

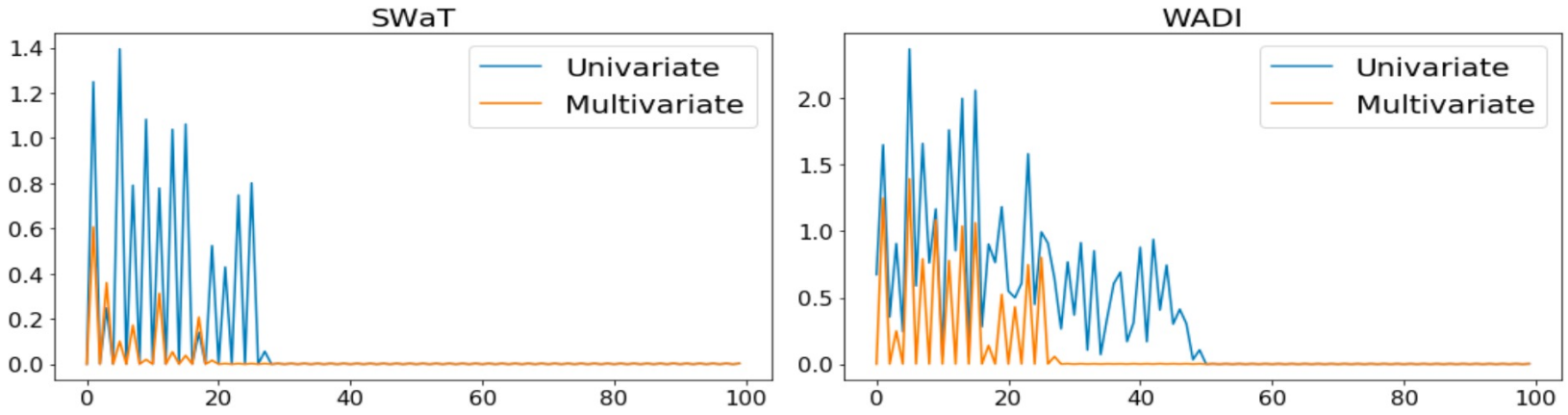


Fig. 3: MMD: generation for multiple time series v.s. generation single time series.

- **MMD**
  - 30~50번의 학습 후 값이 수렴하여 정상 데이터의 분포를 학습
  - Multivariate Data에서 Univariate Data보다 학습이 빠르고 성능이 우수

EXPERIMENTS

Table 2: Anomaly Detection Metrics for Different Datasets

Datasets	Methods	Pre	Rec	F_1
SWaT	PCA	24.92	21.63	0.23
	KNN	7.83	7.83	0.08
	FB	10.17	10.17	0.10
	AE	72.63	52.63	0.61
	EGAN	40.57	67.73	0.51
	MAD-GAN*	**99.99**	54.80	0.70
	MAD-GAN**	12.20	**99.98**	0.22
	MAD-GAN***	98.97	63.74	**0.77**
WADI	PCA	39.53	5.63	0.10
	KNN	7.76	7.75	0.08
	FB	8.60	8.60	0.09
	AE	34.35	34.35	0.34
	EGAN	11.33	37.84	0.17
	MAD-GAN*	**46.98**	24.58	0.32
	MAD-GAN**	6.46	**99.99**	0.12
	MAD-GAN***	41.44	33.92	**0.37**
KDDCUP99	PCA	60.66	37.69	0.47
	KNN	45.51	18.98	0.53
	FB	48.98	19.36	0.28
	AE	80.59	42.36	0.55
	EGAN	92.00	95.82	**0.94**
	MAD-GAN*	**94.92**	19.14	0.32
	MAD-GAN**	81.58	**96.33**	0.88
	MAD-GAN***	86.91	94.79	0.90

Rows GAN-AD* list results chosen by best Precision.
Rows GAN-AD** list results chosen by best Recall.
Rows GAN-AD*** list results chosen by best F_1.

EXPERIMENTS

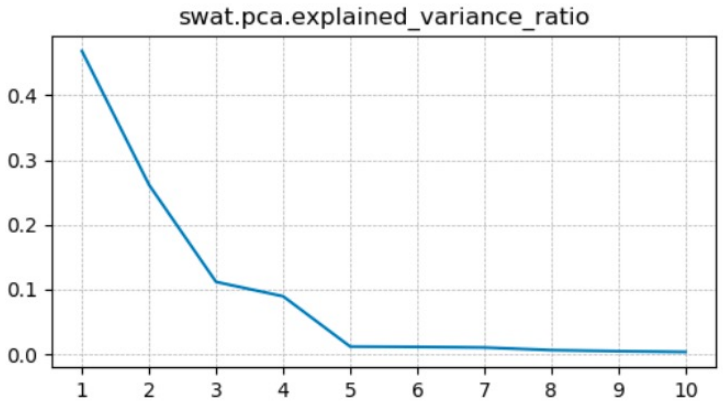
Table 3: Anomaly Detection Metrics of MAD-GAN at Different PC Resolutions

EM	PC=1	PC=2	PC=3	PC=4	PC=5	PC=6	PC=7	PC=8	PC=9	PC=10	ALL
Pre	16.90	17.76	18.57	14.53	26.56	24.39	13.37	14.78	13.60	13.83	13.95
Rec	72.03	95.34	92.76	91.16	95.27	81.25	91.87	92.02	95.06	96.03	92.96
F_1	0.24	0.23	0.23	0.23	0.37	0.22	0.22	0.22	0.23	0.23	0.23

EM: evaluation metrics. Sub-sequence length equals to 30.

- **PCA**
  - PCA를 활용해 입력 데이터의 차원을 줄임
  - 모든 변수를 사용하지 않고, 일부 주요 변수만 선택하면 계산 속도를 줄일 수 있음
  - PC가 1부터 4인 경우 Recall은 대체적으로 높지만 Precision이 낮음
  - PC가 5인 경우 Precision이 향상돼 이전보다 오탐비율이 줄었고, Recall과 F1스코어도 높음
  - PC가 6이상인 경우 Recall은 계속 90이상으로 유지되지만 Precision이 다시 낮아짐

EXPERIMENTS



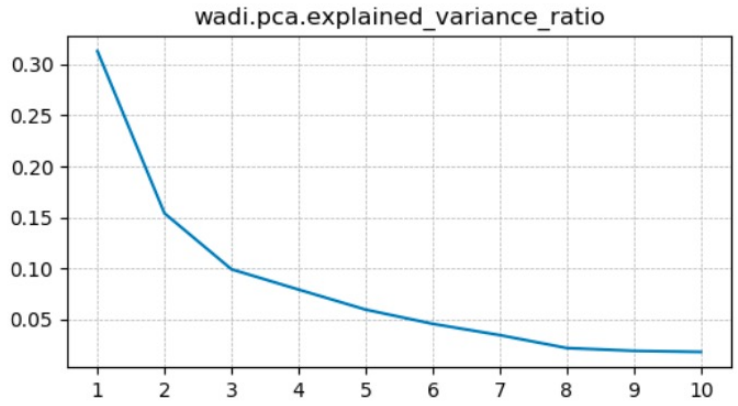
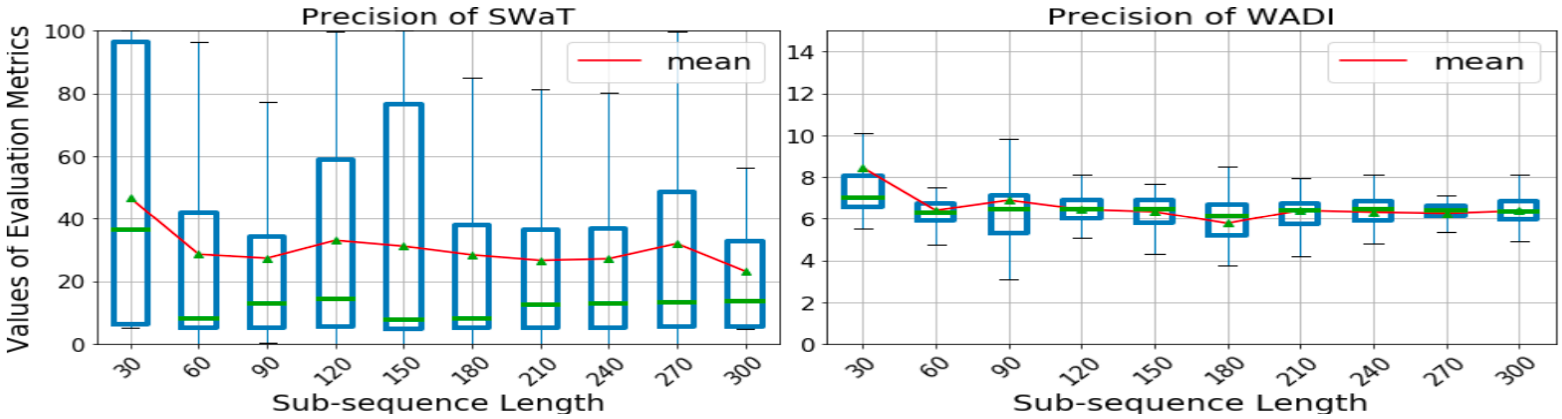


Fig. 4: Variance Ratio of Principal Component for the SWaT and WADI data.

- **Principal Component for the SWaT and WADI**
  - 각 데이에서 비중이 높은 PC를 나타낸 그래프

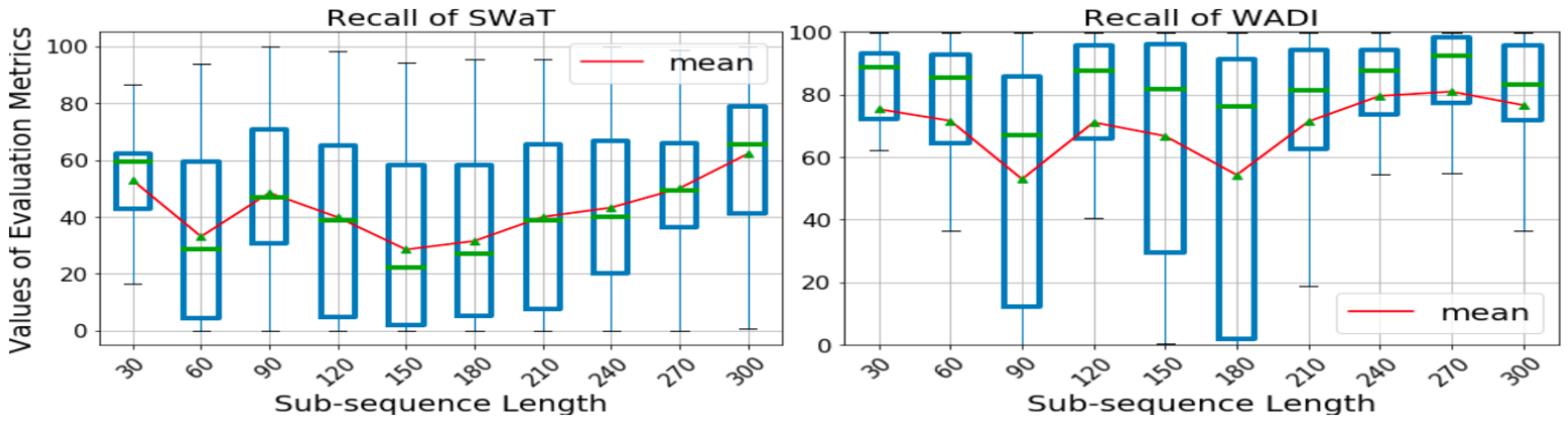
EXPERIMENTS

- **Precision with Sliding Window Size**
  - SWaT에서는 SW사이즈가 30일 때 가장 높은 Precision을 기록
  - WADI에서도 30일 때 가장 높은 스코어를 기록했지만 모든 값들이 낮은 값을 기록

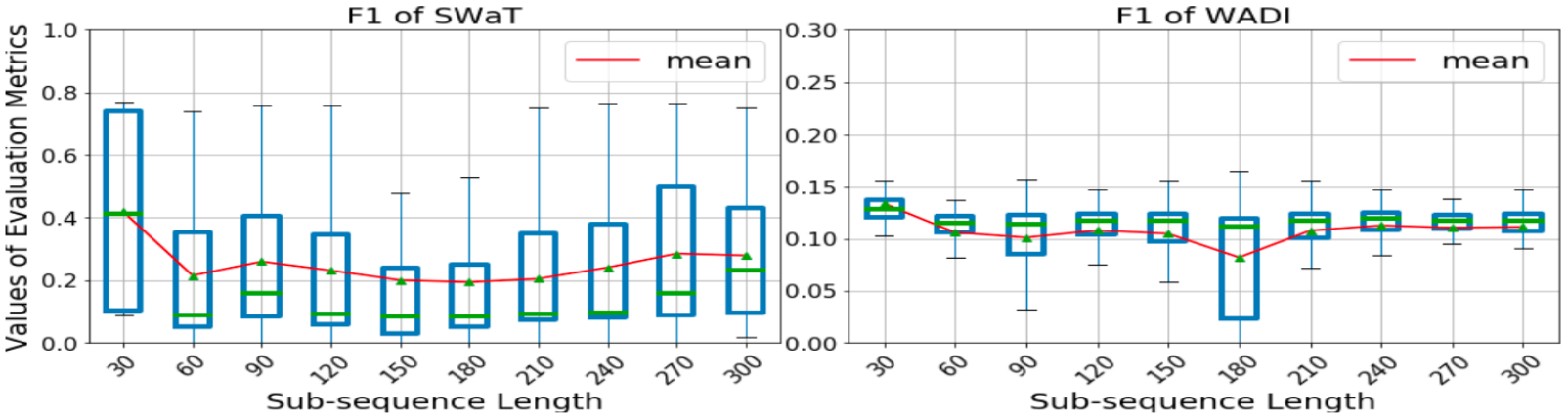


EXPERIMENTS

- **Recall With Sliding Window Size**



- **F1 score with Sliding Window Size**



EXPERIMENTS

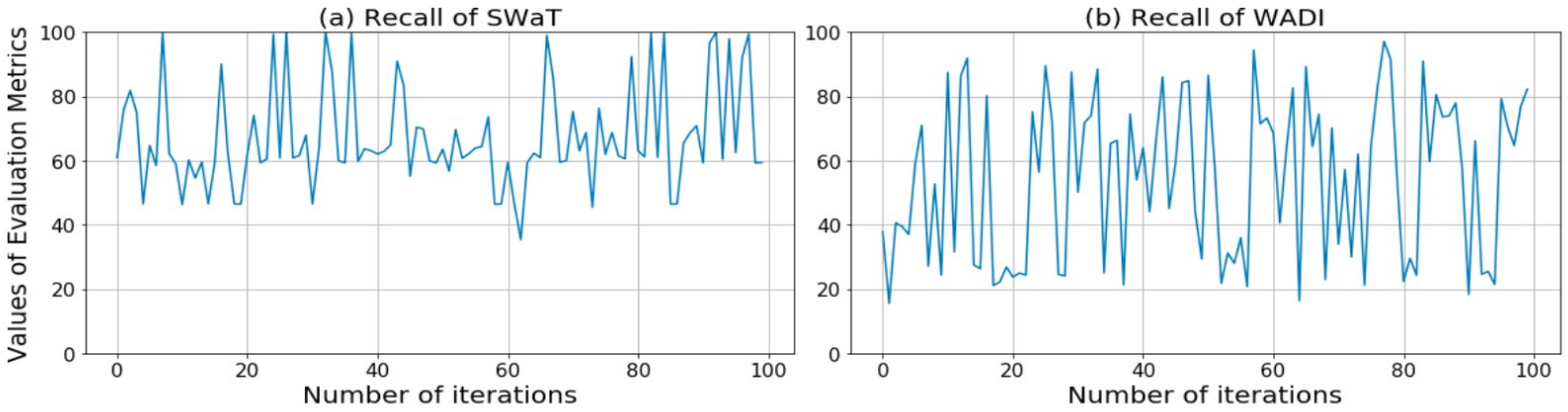


Fig. 6: Values of evaluation metrics as a function of the iteration epochs for the SWaT and WADI datasets when the sub-sequence length $s_w = 180$.

Conclusions

- **MAD-GAN**
  - TSAD분야에서 GAN을 접목시킨 첫번째 논문
  - Reconstruction based model
  - 데이터를 생성하는 Generator와 판별하는 Discriminator로 구성

- **Pros**
  - Reconstruction Loss만 사용하지 않고 Discriminative Loss도 사용해 불완전함 해소
  - GAN의 TSAD분야 속 가능성을 제시

- **Cons**
  - 기존 모델들 보다 좋은 성능을 지속적으로 보이지 못함
  - SWaT보다 WADI에서 성능이 떨어졌으며, 일관적이지 못한 성능을 보임
  - 결과의 일관성이 떨어짐