C-GATS: Conditional Generation of Anomalous Time Series

Vikramank Singh, Abishek Sankararaman, Balakrishnan (Murali) Narayanaswamy, Zhao Song

Amazon Web Services (AWS) AI Labs

NeurIPS 2022 SyntheticData4ML Workshop

Anomalies in real-world are rare.

Anomaly Detection (AD) is fundamental to many real-world applications

Fraud Detection

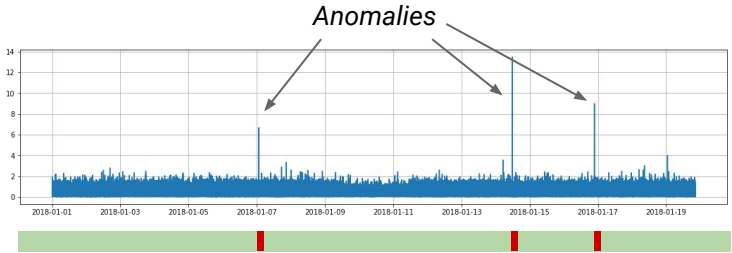
Performance Monitoring

Health Tracking

Let X : Anomalous time series, Y: Labels

X in real-world are rare

Obtaining Y is hard.



Supervised AD models need large amounts of (X,Y) for good performance.

Q. Can we generate these (X,Y) pairs synthetically to improve performance of supervised AD models?

C-GATS: Framework

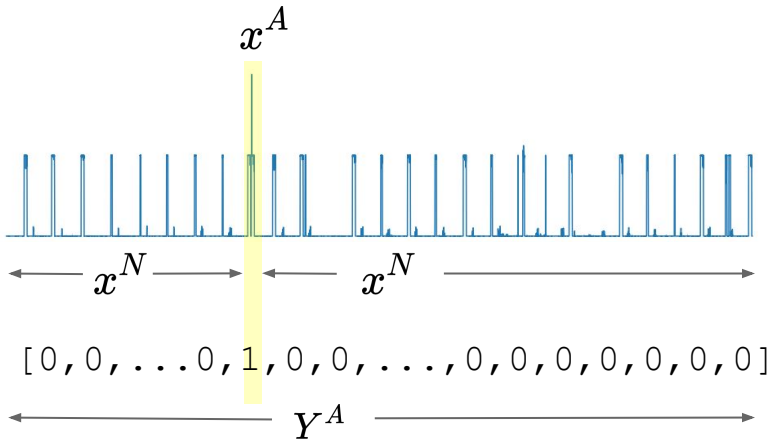
$$X^A \equiv x^A \cup x^N$$

$$p(X^A) = \underbrace{p(x^N)}_{\text{modeled using } p(X^N)} \cdot \underbrace{p(x^A)}_{\text{modeled using } p(X^A|X^N,Y^A)} \cdot \underbrace{p(Y^A)}_{\text{modeled using } P(Y^A)}$$

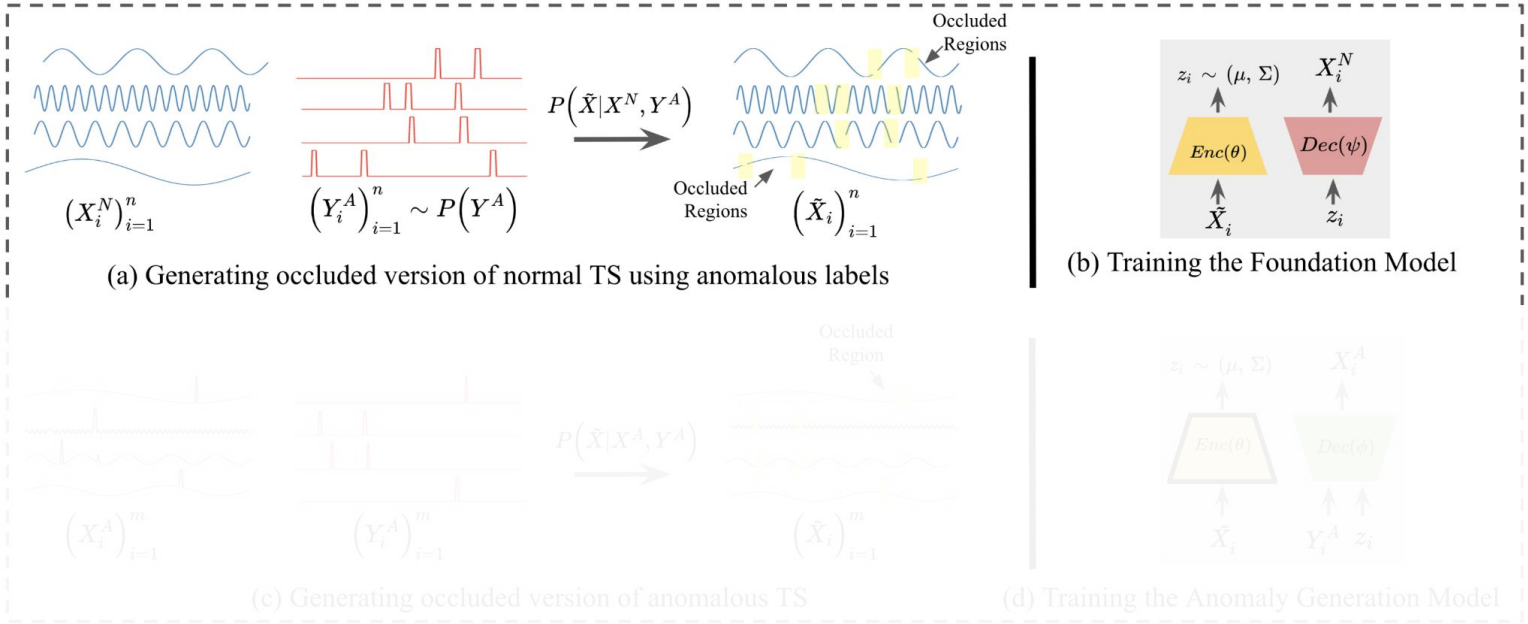
Distribution of normal data

Distribution of anomalous data

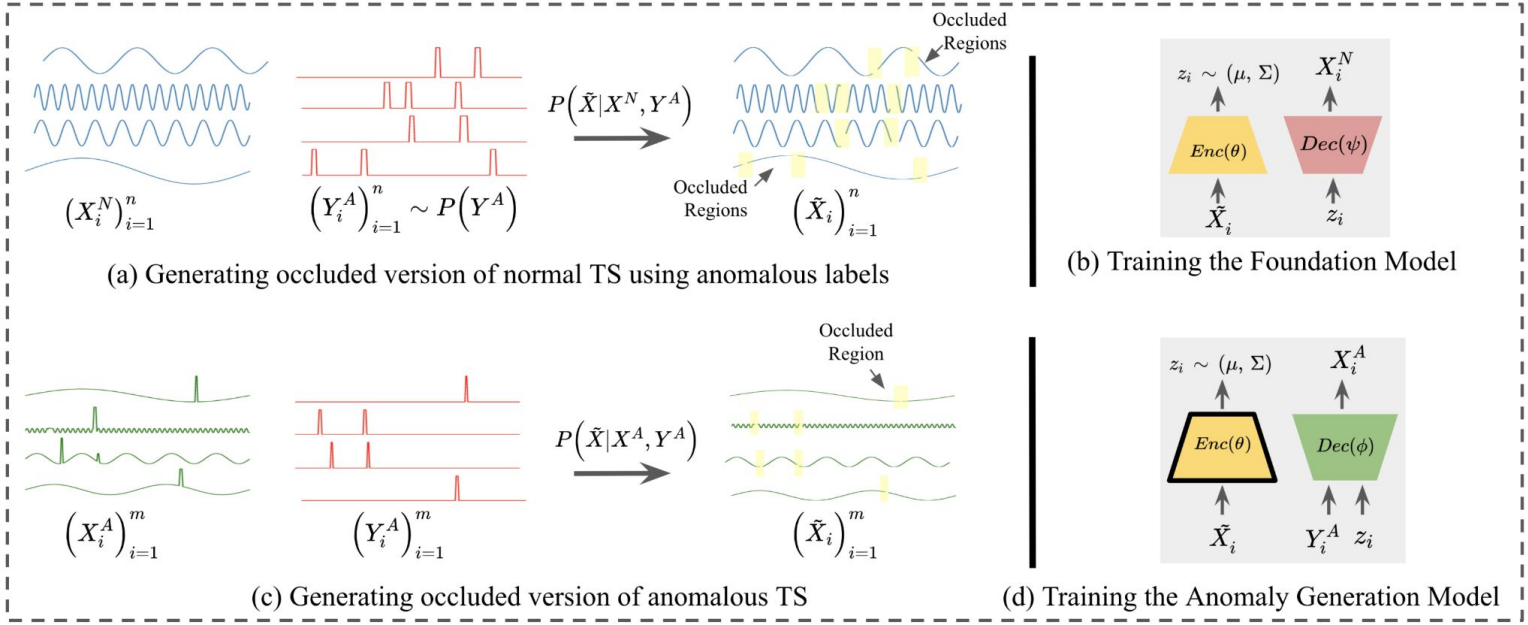
Attributes of the anomaly (eg: position, type, etc)



C-GATS: Framework (Stage - I)



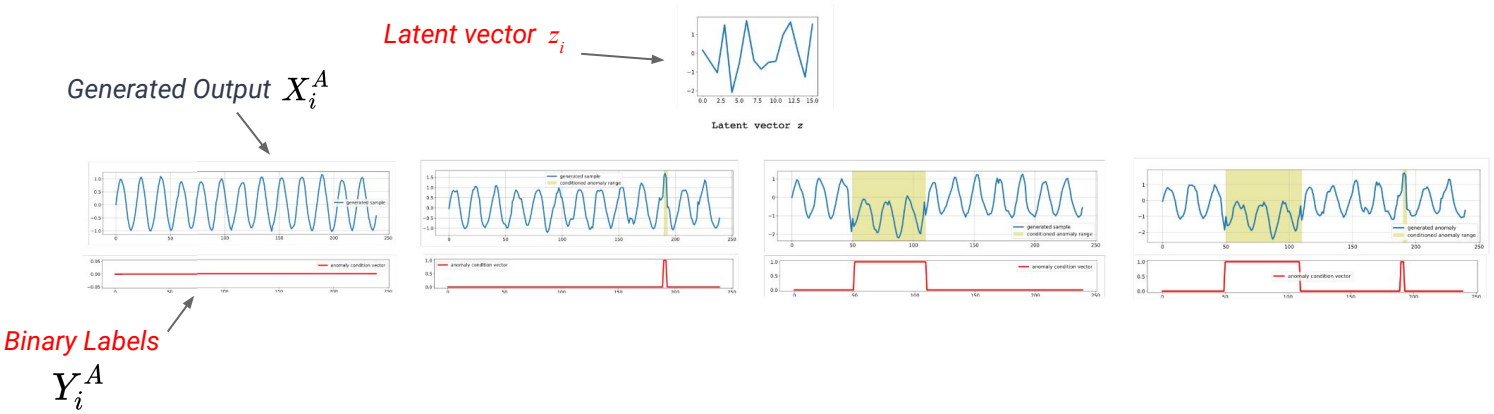
C-GATS: Framework (Stage - I)

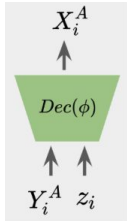


C-GATS: Qualitative Results

Baselines

7 Data Augmentation methods
8 Anomaly Detection algorithms
6 different real and synthetic datasets



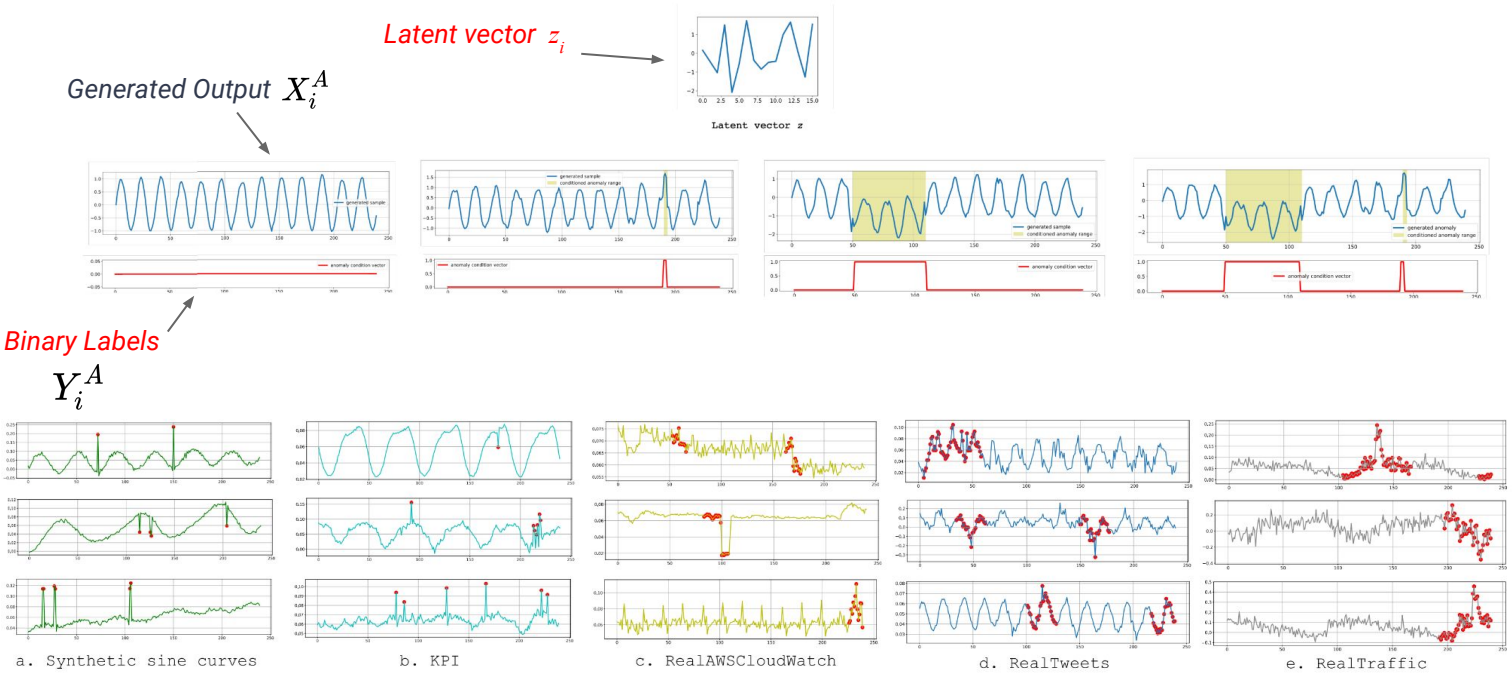


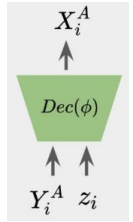
Inference Network

C-GATS: Qualitative Results

Baselines

7 Data Augmentation methods
8 Anomaly Detection algorithms
6 different real and synthetic datasets





Inference Network

C-GATS: Quantitative Results

Baselines

7 Data Augmentation methods
8 Anomaly Detection algorithms
6 different real and synthetic datasets

AD Algorithm	TRTR			T(R+S)TR			TRTS			TSTR		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
RobustTAD	0.57	0.53	0.55	0.69	0.66	0.67	0.50	0.45	0.48	0.65	0.64	0.64
SR-CNN	0.55	0.51	0.54	0.61	0.55	0.58	0.47	0.43	0.44	0.60	0.55	0.57
NCAD	0.60	0.68	0.64	0.70	0.69	0.69	0.51	0.54	0.52	0.65	0.73	0.68

Comparing performance (F1 scores) of a supervised SOTA AD algorithm on a real held-out test set when trained on Real vs Synthetic dataset

C-GATS: Quantitative Results

Baselines

7 Data Augmentation (DA) methods
8 Anomaly Detection (AD) algorithms
6 different real and synthetic datasets

AD Algorithm	TRTR			T(R+S)TR			TRTS			TSTR		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
RobustTAD	0.57	0.53	0.55	0.69	0.66	0.67	0.50	0.45	0.48	0.65	0.64	0.64
SR-CNN	0.55	0.51	0.54	0.61	0.55	0.58	0.47	0.43	0.44	0.60	0.55	0.57
NCAD	0.60	0.68	0.64	0.70	0.69	0.69	0.51	0.54	0.52	0.65	0.73	0.68

Comparing performance (F1 scores) of a supervised SOTA AD algorithm on a real held-out test set when trained on Real vs Synthetic dataset

Algorithm	Dataset	w/o Aug.	Mixup *(alpha=0.2)*	CutMix	Combination *(Scaling, Jitter, Permute,TimeWarp)*	RCGAN	C-GATS
RobustTAD	Synthetic Sines	0.74 ±0.005	0.74 ±0.007	0.27 ±0.003	0.75 ±0.006	0.23 ±0.002	**0.75 ±0.004**
	KPI	0.55 ±0.011	0.56 ±0.008	0.23 ±0.010	0.59 ±0.003	0.19 ±0.009	**0.67 ±0.003**
	RealTweets	0.54 ±0.021	0.53 ±0.011	0.32 ±0.026	0.57 ±0.009	0.18 ±0.041	**0.57 ±0.010**
	RealTraffic	0.67 ±0.026	0.68 ±0.029	0.37 ±0.020	0.69 ±0.012	0.27 ±0.071	**0.73 ±0.015**
	RealAWSCloudWatch	0.34 ±0.262	0.34 ±0.089	0.06 ±0.344	**0.36 ±0.218**	0.04 ±0.181	0.34 ±0.192
	ArtificialWithAnomaly	0.36 ±0.275	0.38 ±0.118	0.33 ±0.229	0.38 ±0.136	0.05 ±0.099	**0.41 ±0.107**
SR-CNN	Synthetic Sines	0.73 ±0.009	0.72 ±0.002	0.18 ±0.009	0.74 ±0.001	0.16 ±0.010	**0.74 ±0.003**
	KPI	0.54 ±0.008	0.55 ±0.005	0.19 ±0.009	0.56 ±0.003	0.17 ±0.011	**0.58 ±0.007**
	RealTweets	0.46 ±0.027	0.45 ±0.019	0.24 ±0.033	0.51 ±0.009	0.19 ±0.008	**0.53 ±0.009**
	RealTraffic	0.64 ±0.021	0.65 ±0.017	0.33 ±0.019	0.65 ±0.028	0.26 ±0.035	**0.66 ±0.020**
	RealAWSCloudWatch	0.25 ±0.319	0.25 ±0.283	0.04 ±0.081	**0.28 ±0.193**	0.01 ±0.198	0.27 ±0.199
	ArtificialWithAnomaly	0.28 ±0.019	0.30 ±0.174	0.24 ±0.111	0.32 ±0.081	0.12 ±0.059	**0.34 ±0.011**
NCAD	Synthetic Sines	0.77 ±0.008	0.76 ±0.001	0.30 ±0.003	0.78 ±0.019	0.27 ±0.071	**0.79 ±0.001**
	KPI	0.64 ±0.006	0.66 ±0.003	0.24 ±0.009	0.67 ±0.010	0.21 ±0.004	**0.69 ±0.001**
	RealTweets	0.57 ±0.019	0.55 ±0.014	0.32 ±0.028	0.58 ±0.010	0.20 ±0.085	**0.61 ±0.009**
	RealTraffic	0.69 ±0.029	0.69 ±0.033	0.43 ±0.098	0.71 ±0.051	0.28 ±0.171	**0.71 ±0.019**
	RealAWSCloudWatch	0.35 ±0.219	0.34 ±0.118	0.11 ±0.092	**0.37 ±0.071**	0.07 ±0.215	0.35 ±0.111
	ArtificialWithAnomaly	0.39 ±0.101	0.42 ±0.213	0.36 ±0.128	0.42 ±0.098	0.13 ±0.023	**0.45 ±0.091**

Comparing how different DA methods impact the performance of different AD algorithms on various real-world datasets