Digital Twin-based Anomaly Detection in Cyber-physical Systems (ATTAIN)

Qinghua Xu

Dept. of Engineering Complex Software Systems

Simula Fornebu

Inria-Simula Workshop
March 16, 2021

*Q. Xu, S. Ali, and T. Yue "Digital Twin-based Anomaly Detection in Cyber-physical Systems.", Accepted at ICST 2021.

simula

CONTENTS

simula

Background

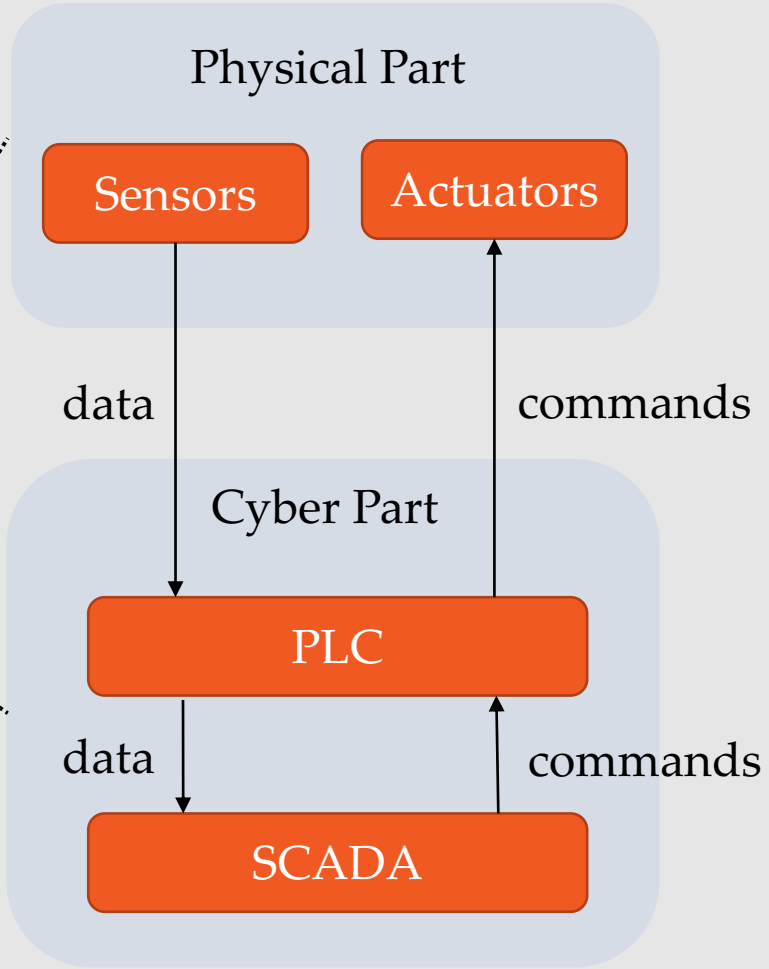
Cyber-Physical System

Drone

Self-driving car

Water treatment plant

Cyber-physical System



simula

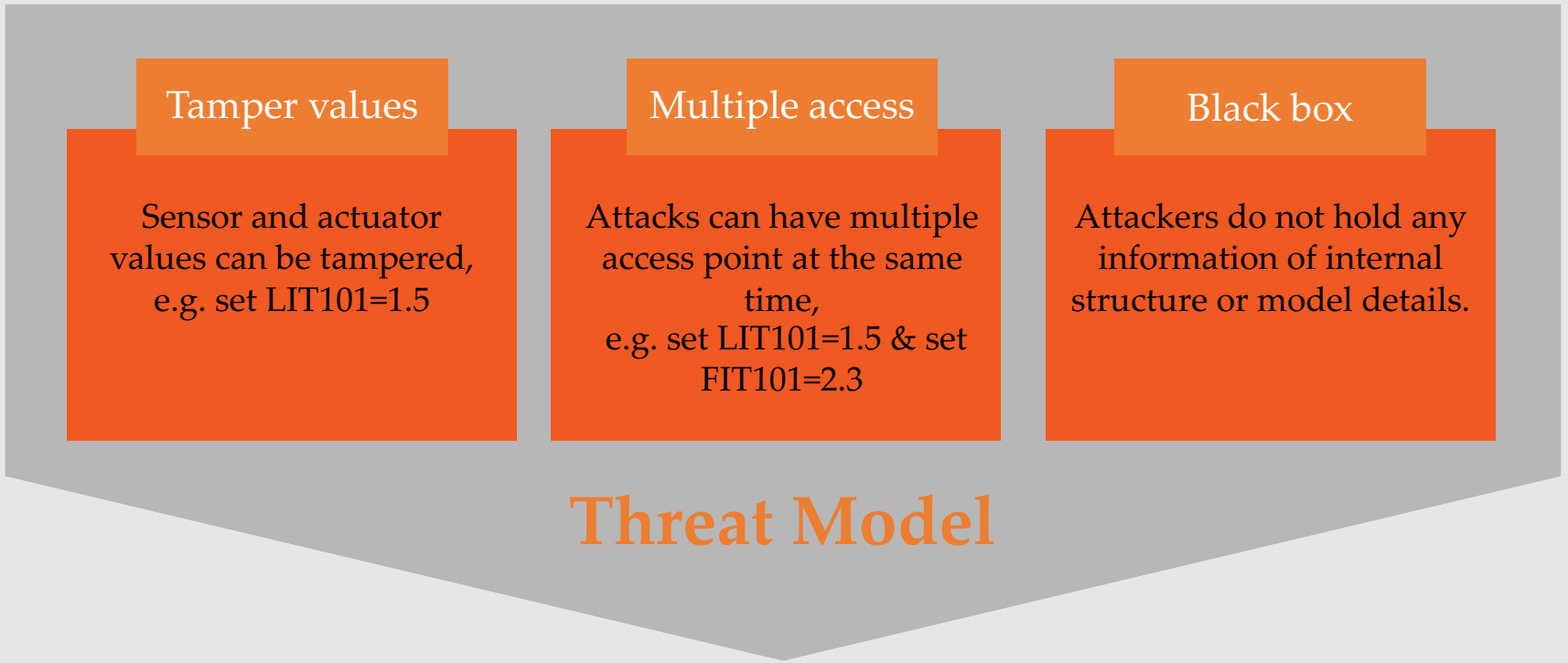
Background Threats

Broader Threats

simula

Background

Threat Model & Task Definition



Find anomaly in real time in CPS

simula

Background

Example

Time	FIT101	LIT101	MV101	P101	P102	Label
10:00:00	2.43	522.84	2	2	1	Normal
10:00:01	2.45	522.88	2	2	1	Normal
…	…	…	…	…	…	…
10:29:13	2.44	816.84	2	1	1	Normal
10:29:14	2.49	817.67	2	1	1	Attack
10:29:15	2.54	817.94	2	1		Attack
…	…	…	…	…	…	…
10:44:53	6e-4	869.72	1	2	1	Attack

FIT101

LIT101

MV101

P101

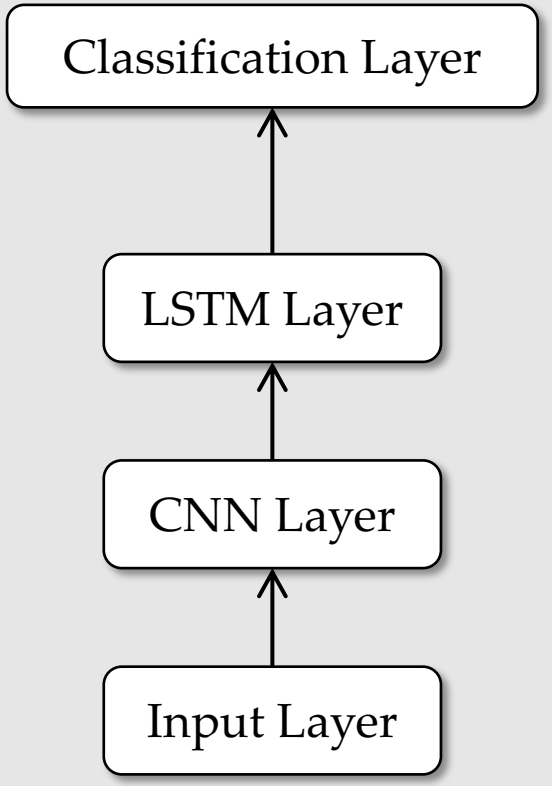
P102

simula

Background

Literature & Challenges

Existing Methods



Main Challenges

1. Spatial features are not well captured
2. Lack of sufficient labeled data
3. Unable to learn during runtime

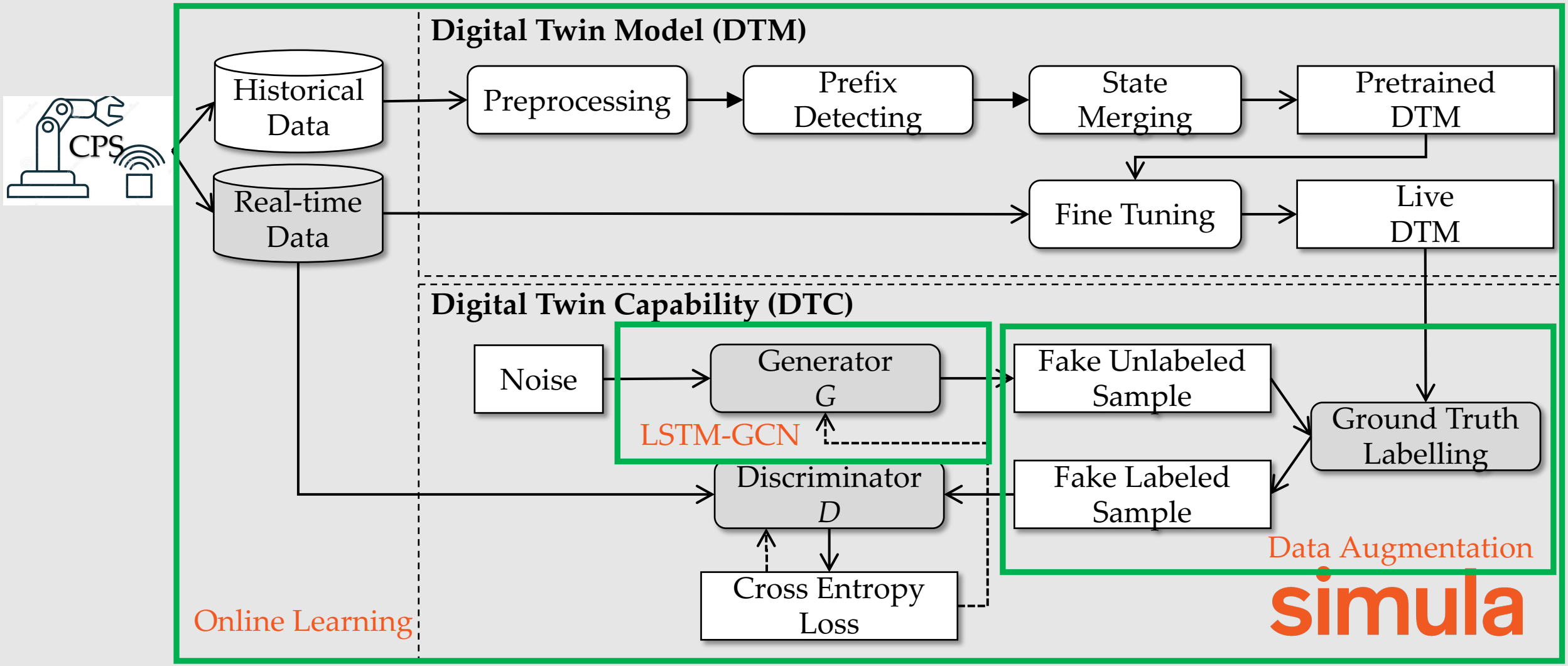
simula

Methodology Overview

- **Digital twin model** is a virtual replica or live model of CPS
- **Digital twin capability** is the functionality of a digital twin

simula

Methodology Details



Experiment

Case Studies

Battle Of The Attack Detection Algorithms (BATADAL)

Water Distribution (WADI)

Secure Water Treatment (SWaT)

simula

RQ1

How effective is our anomaly detector as compared to the literature?

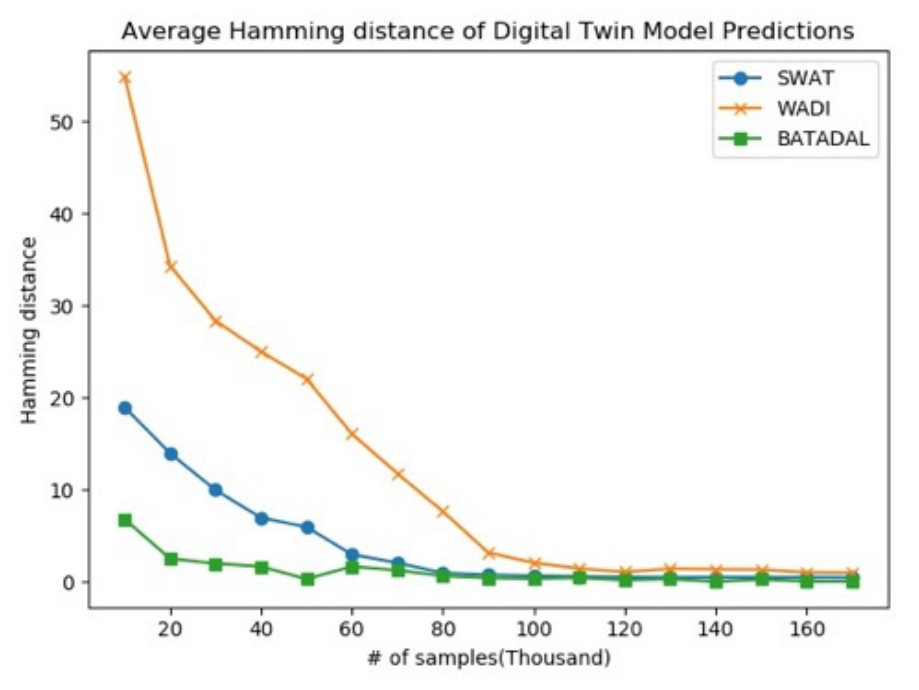
	SWaT			WADI			BATADAL		
Model	P	R	F1	P	R	F1	P	R	F1
LSTM-CUSUM	0.907	0.677	0.775	0.614	0.697	0.659	0.657	0.721	0.687
MAD-GAN	0.961	0.942	0.951	0.432	0.952	0.594	0.529	0.962	0.683
ATTAIN (without signal)	0.922	0.954	0.937	0.524	0.782	0.627	0.553	0.774	0.645
ATTAIN	0.959	0.992	0.975	0.665	0.844	0.744	0.722	0.763	0.742

ATTAIN outperforms LSTM-CUSUM and MAD-GAN for almost all metrics on all the three datasets, with particularly good performance in terms of precision.

simula

RQ2

How realistic is our digital twin model?



- **State prediction:** Hamming distance converges after training for 80,000 samples
- **Outlier Detection:** Accuracy on SWaT, WADI, and BATADAL are 0.82, 0.69, and 0.74

simula

RQ3

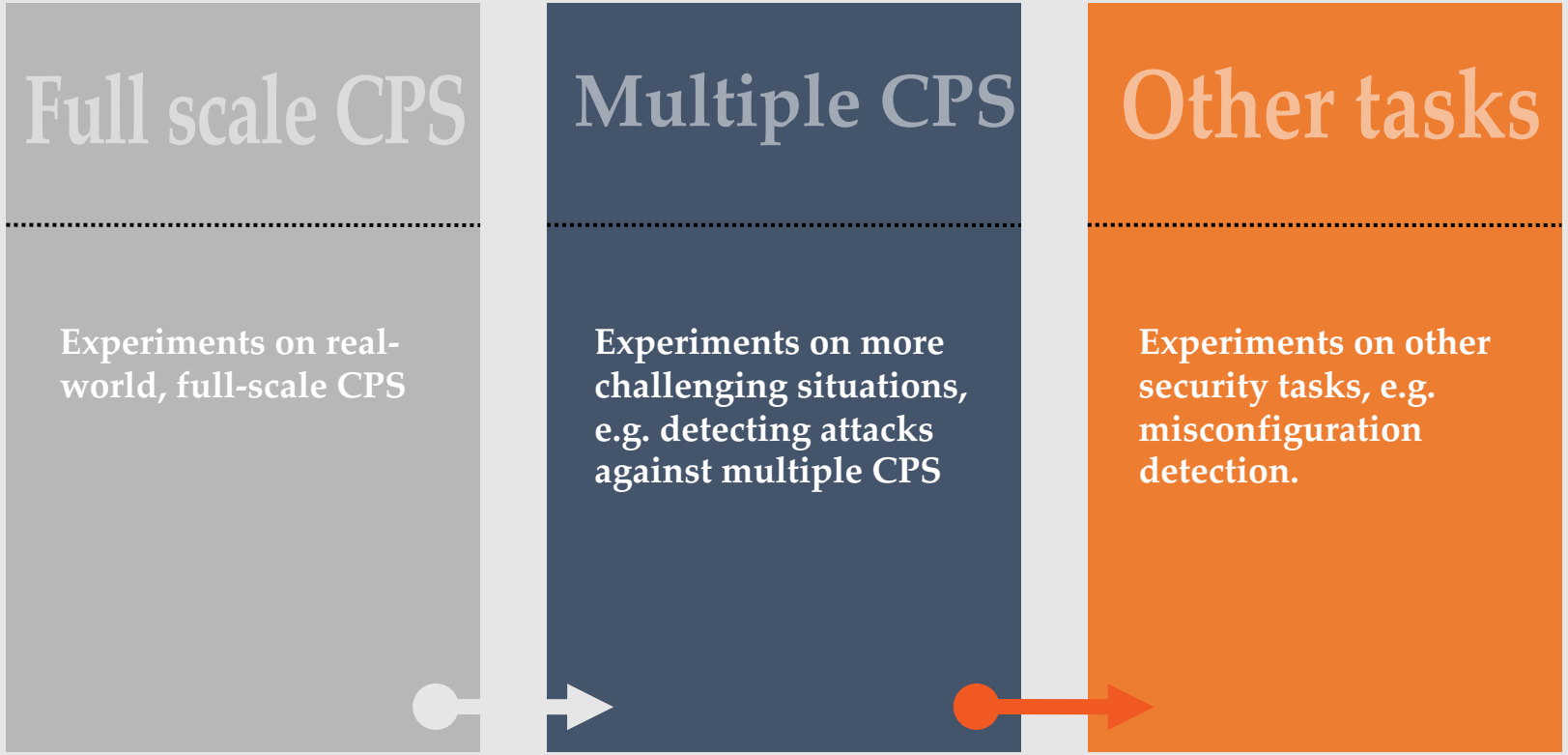
Is using DT effective in detecting anomalies as compared to not using it?

	SWaT			WADI			BATADAL		
Model	P	R	F1	P	R	F1	P	R	F1
LSTM-CUSUM	0.907	0.677	0.775	0.614	0.697	0.659	0.657	0.721	0.687
MAD-GAN	0.961	0.942	0.951	0.432	0.952	0.594	0.529	0.962	0.683
ATTAIN (without signal)	0.922	0.954	0.937	0.524	0.782	0.627	0.553	0.774	0.645
ATTAIN	0.959	0.992	0.975	0.665	0.844	0.744	0.722	0.763	0.742

ATTAIN with signals from the digital twin model improves the F1 score by more than 10% on the SWaT and BATADAL datasets when compared with ATTAIN without signals

simula

Future work



simula