ADAPTIVE CUSUM BASED ATTACK DETECTION IN CYBER PHYSICAL SYSTEM

BTECH PROJECT-II REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor's Degree in Manufacturing Science and Engineering of Mechanical Engineering Department

by

Loya Vivek
(Roll number : 20MF10017)

under the supervision of

Prof. Soumyajit Dey

(Dept. of Computer Science & Engineering)

Department of Computer Science and Engineering

Indian Institute of Technology Kharagpur

Spring Semester, 2023-24

April 27, 2024

DECLARATION

I certify that

(a) The work contained in this report has been done by me under the guidance of my supervisor.

(b) The work has not been submitted to any other Institute for any degree or diploma.

(c) I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

(d) Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the thesis and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

Date: April 27, 2024 (Loya Vivek)

Place: Kharagpur (20MF10017)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

KHARAGPUR , INDIA

CERTIFICATE

This is to certify that the project report entitled Adaptive CUSUM-Attack Detection in Cyber Physical System submitted by Loya Vivek (Roll Number:20MF10017) an undergraduate student of Manufacturing Science and Engineering of Mechanical Engineering Department towards partial fulfilment of requirements for the award of degree of Bachelor's Degree in Manufacturing Science and Engineering of Mechanical Engineering, is a record of bona-fide work carried out by him under my supervision and guidance during Spring Semester, 2023-2024.

Professor Soumyajit Dey,

Date :- 27/04/2024 Department of Computer Science and Engineering

Place : Kharagpur Indian Institute of Technology, Kharagpur

Abstract

Name of the student: **Loya Vivek** Roll No: **20MF10017**

Degree for which submitted: **Bachelor of Technology**

Department: **Department of Computer Science and Engineering**

Thesis title: **Adaptive CUSUM Based - Attack Detection in Cyber Physical System**

Thesis supervisor: **Professor Soumyajit Dey**

Month and year of thesis submission: **April 27, 2024**

In advanced technology applications, sensors are instrumental in monitoring and controlling complex systems to ensure their safe operation. However, these sensors are susceptible to attacks that can manipulate their readings, leading to the generation of erroneous data. Detecting such sensor attacks in real-time is crucial to safeguard system safety and prevent critical decisions or actions based on falsified sensor measurements.

Acknowledgements

I express my sincere gratitude to **Prof. Soumyajit Dey**, Computer Science Engineering Department, Indian Institute of Technology, Kharagpur for his supervision and guidance during the project. His valuable advices, encouragement, suggestions and caring nature helped me throughout my project and took me past every obstacle. His guidance was undoubtedly one of the most important reason behind successful completion of this work. I am also thankful to him for extending all kinds of help and for providing the necessary facilities required during this work. I am thankful to my mentor **Mr.Akash Bhattacharya** for his support,patience and guidance throughout the project and his contribution to bring this work to fulfilment.

Contents

1.INTRODUCTION

In advanced technology applications, sensors play a pivotal role in monitoring and controlling complex systems, ensuring their safe operation. Sensors within CPS are indispensable for collecting real-time data to make informed decisions. However, these sensors are vulnerable to attacks that can manipulate their readings, leading to the generation of erroneous data. Detecting such sensor attacks in real-time is crucial to safeguard the integrity of the system and prevent critical decisions or actions based on falsified sensor measurements.

Cyber-physical systems are susceptible to various vulnerabilities due to their interconnected nature and reliance on digital communication. Attackers can exploit these vulnerabilities to manipulate sensor data, compromising the integrity and reliability of the entire system. For instance, an attacker could tamper with temperature sensors in an industrial control system, leading to incorrect temperature readings and potentially causing equipment malfunction or safety hazards.

The consequences of such attacks can be severe, ranging from financial losses to endangering human lives. For example, in a smart grid system, manipulating sensor data could lead to incorrect energy distribution, causing power outages or even damaging critical infrastructure.

To mitigate these risks, it is imperative to implement robust detection mechanisms to identify anomalies in sensor data and take appropriate corrective actions. Early detection of sensor attacks is vital in upholding the reliability and security of advanced technology systems.

Traditional statistical model-based detectors like chi-square and CUSUM are vulnerable to smart attacks, including FDI that manipulate the constant parameters (thresholds and biases). To counter this, our proposed approach involves the adaptive modification of the parameters of the CUSUM detector by learning from the system dynamics, the changes in statistics of the residue, and the system's behaviour under attack and without attack scenarios.

We are experimenting with the approaches using three different controllers.

1.Trajectory Tracking Controller (TTC)

2.Electronic Stability Program (ESP)

3.Suspension Controller

2.BACKGROUND

In previous research a classical CUSUM detector with predefined threshold and bias was used to identify sensor attacks.

System model: We consider a discrete linear time-invariant system model

$$x_{k+1} = Ax_k + Bu_k + v_k,\ y_k = Cx_k + \eta_k,\ u_k = -K\hat{x}_k$$
$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + Lr_k,\ e_k = x_k - \hat{x}_k,\ \hat{y}_k = C\hat{x}_k$$
$$r_k = \bar{y}_k - \hat{y}_k = Ce_k + \eta_k$$

Attack model and distance measure: The attacked model under an additive FDI attack on ak. Based on the difference between predicted output and actual output, residue is calculated

$$\bar{x}_{k+1} = A\bar{x}_k + B\bar{u}_k + v_k,\ \bar{y}_k = C\bar{x}_k + \eta_k + a_k,\ \bar{u}_k = -K\hat{\bar{x}}_k$$
$$\hat{\bar{x}}_{k+1} = A\hat{\bar{x}}_k + B\bar{u}_k + L\bar{r}_k,\ \bar{e}_k = \bar{x}_k - \hat{\bar{x}}_k,\ \hat{\bar{y}}_k = C\hat{\bar{x}}_k$$
$$\bar{r}_k = \bar{y}_k - \hat{\bar{y}}_k = C\bar{e}_k + \eta_k + a_k,\ g_k = \bar{r}_k^T \Sigma^{-1} \bar{r}_k$$

Detector model: We use CUSUM-based detection. The designed CUSUM detector takes Chi square statistics g of r as its input and evolves like below

$$S_k = max(0, S_{k-1} + g_k - b), \qquad if\ S_{k-1} \leq \tau$$
$$S_k = 0, \qquad if\ S_{k-1} > \tau$$

Here, the classical CUSUM-based detector having constant threshold τ and bias b, accumulates gk by calculating CUSUM sequence Sk at every k-th sampling iteration. Whenever Sk exceeds τ , the detector alarms an attack scenario. If the alarm is triggered at (k−1)-th iteration, i.e., Sk−1 > τ , then Sk is reset to 0.

3.MOTIVATION

In our research, we identified a significant vulnerability in the context of FDI (False Data Injection) stealthy attacks. These attacks, characterised by predefined thresholds and biases, demonstrated the capability to circumvent classical CUSUM detectors effectively. Even when the system's state crossed safety boundaries, these attacks remained undetected.To substantiate this finding we have to generate stealthy attacks using mathematical equations such that its CUSUM value is below threshold . A comparative analysis was conducted,giving an attack from the 200th instance and focusing on the mean and variance of z value where we observed with attack mean and variance of z are different from the without attack conditions. For this we have experimented with 3 different controllers with attack and without attack.

Attack equations are as follows:

$$a_k = \begin{cases} -C\bar{e}_k - \eta_k + \Sigma^{\frac{1}{2}}\bar{\psi}_k \text{ for } k = k_a \\ -C\bar{e}_k - \eta_k + \Sigma^{\frac{1}{2}}\bar{q}_k \text{ for } k > k_a \end{cases}$$

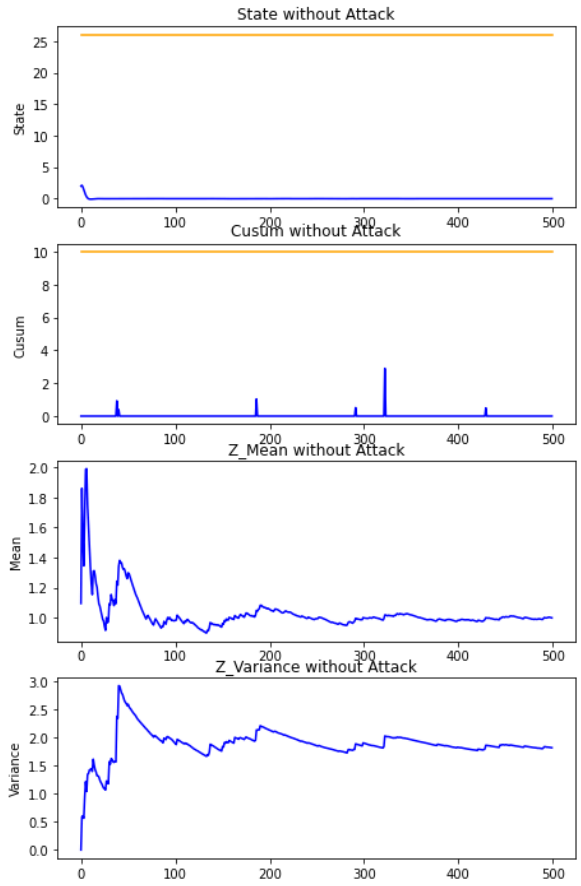
Where $\bar{\psi}_k^T \bar{\psi}_k \leq \tau + b - \bar{S}_{k_a-1}$

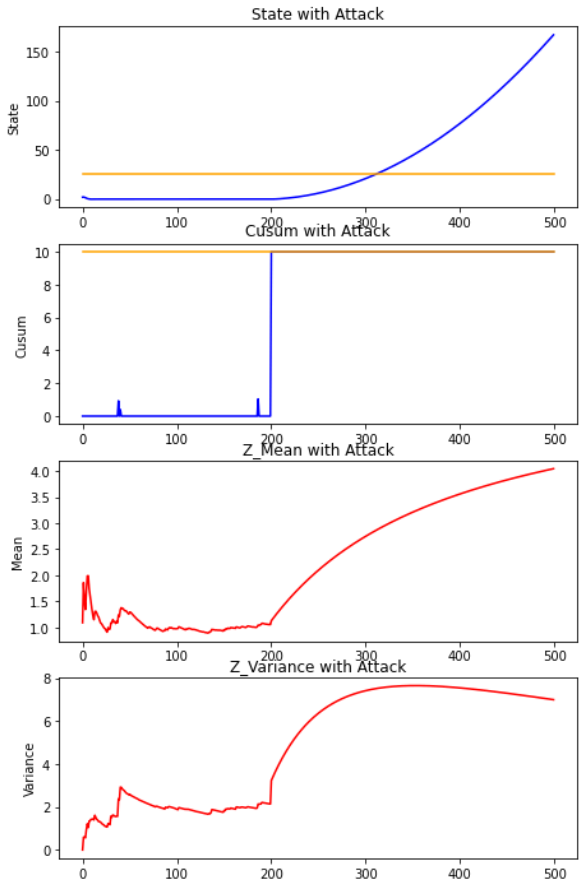
And $\bar{q}_k^T \bar{q}_k \leq b.$

In the analysis of the three controllers, it is evident that even when the state values exceed the safety boundary under attack conditions, the Cumulative Sum (CUSUM) value remains below the threshold, indicating undetected attacks. This observation prompts the exploration of alternative detection criteria. Specifically, we examine the mean and variance of the variable "z" under attack and compare it to the mean and variance of "z" without attack. We find that both the mean and variance of "z" increase significantly under attack, providing a potential detection criterion.

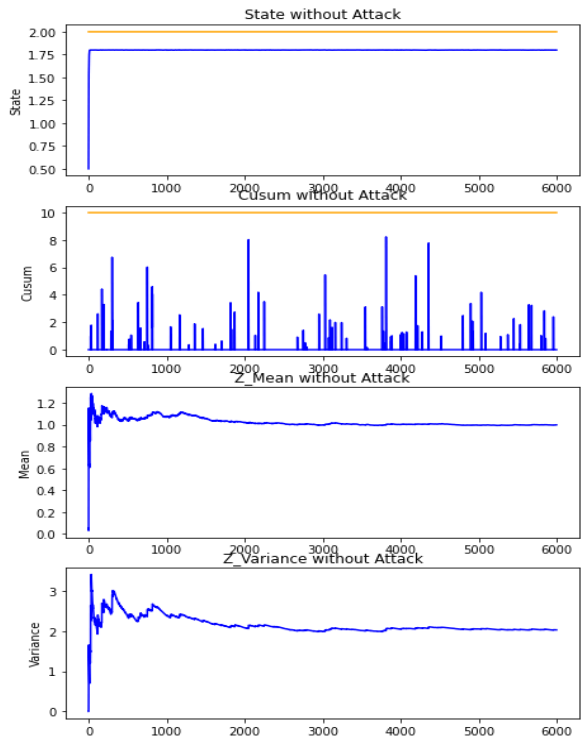
We will show the results with different controllers where the attack is injected from the 200th instance.

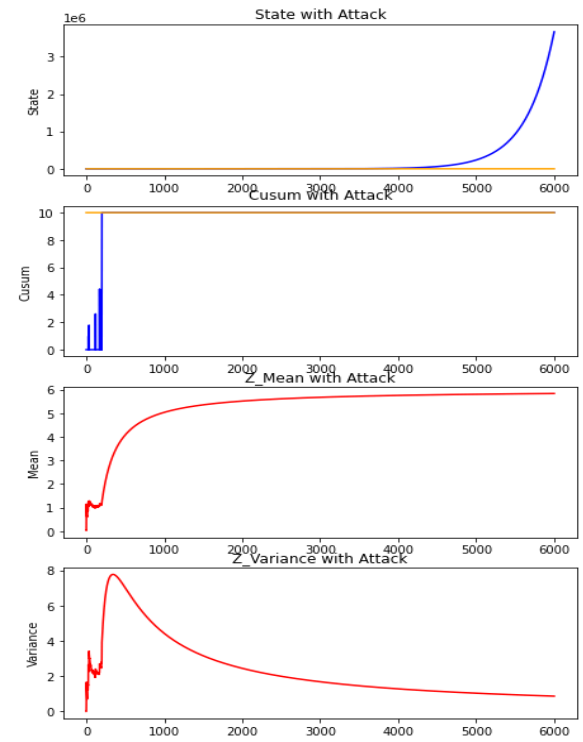
TTC controller:



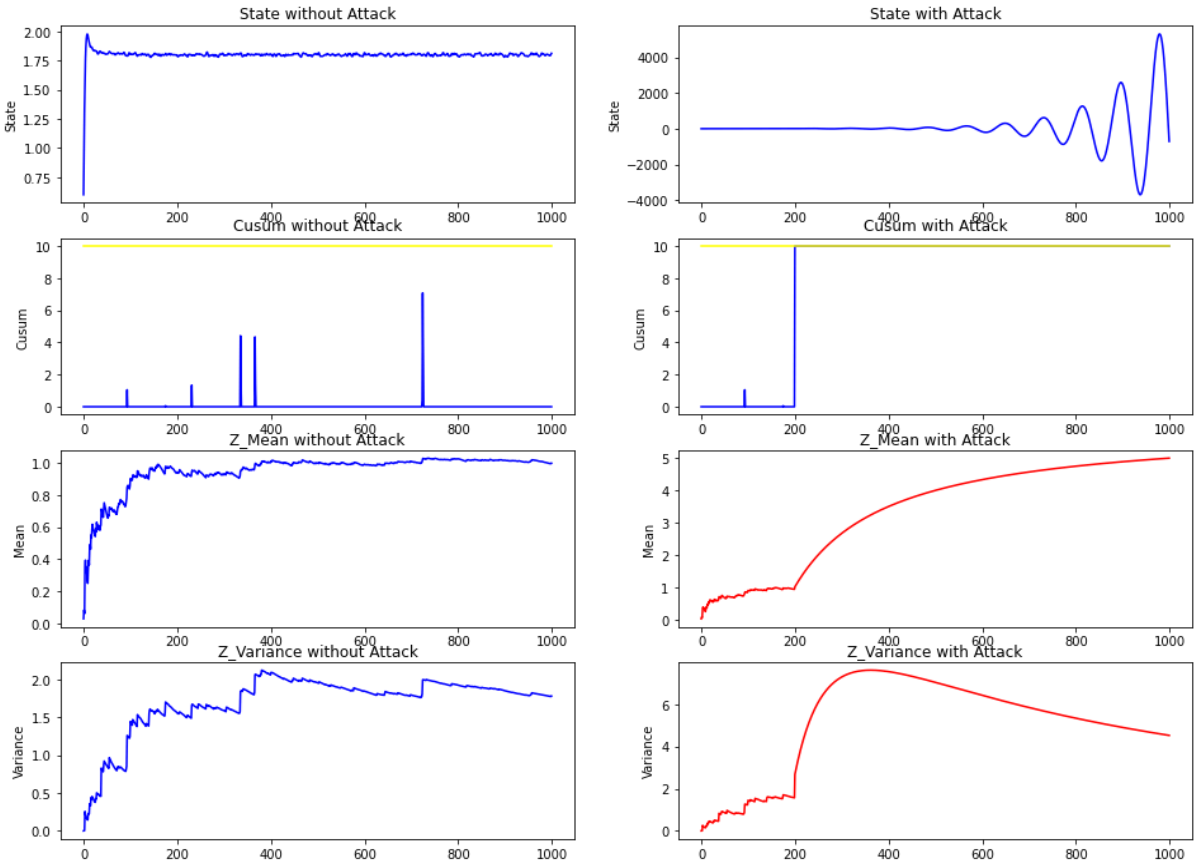


ESP Controller:





Suspension Controller:



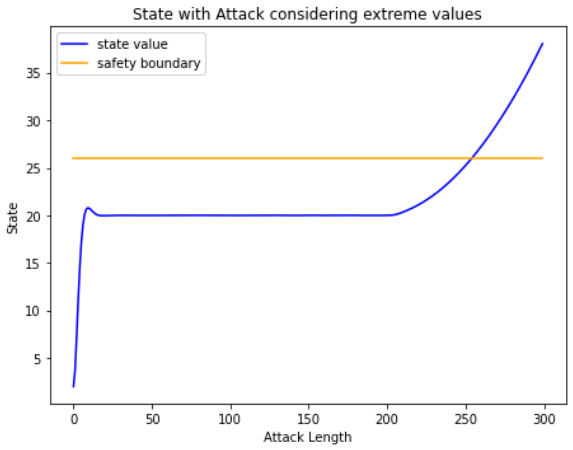
4.METHODOLOGY

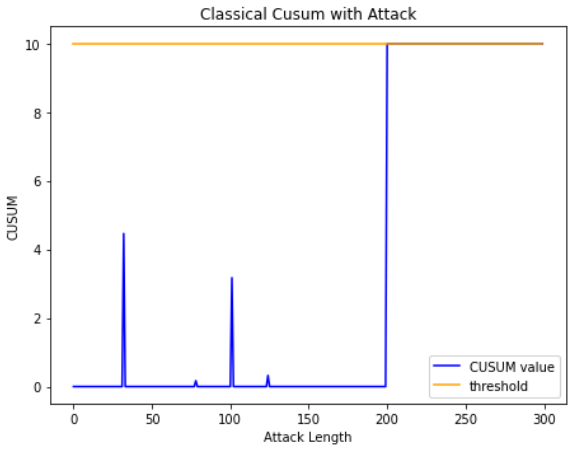
4.1 *State Space Partition Generation*

Since we are trying to find the attack length we find for the most critical attack which will be possible by taking extreme values of $\bar{\psi}_k$ and $\bar{q}_k$. The state values initially less than zero gives most dangerous attack with positive extreme values of $\bar{\psi}_k$ and $\bar{q}_k$ and the state values initially greater than or equal to zero gives most critical attack with negative extreme values of $\bar{\psi}_k$ and $\bar{q}_k$.

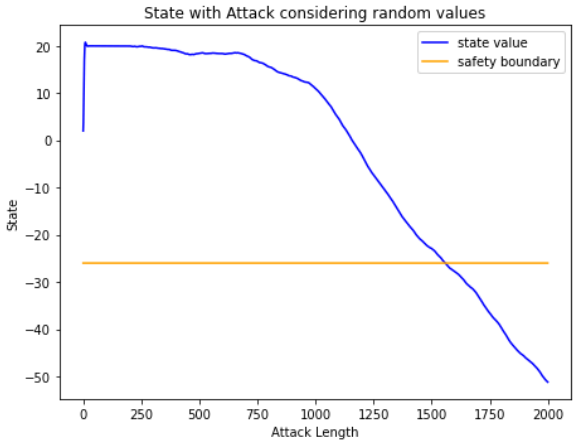
We have experimented with two different conditions one with extreme values of $\bar{\psi}_k$ and $\bar{q}_k$ and other with random values of $\bar{\psi}_k$, $\bar{q}_k$ in a possible range, where attack is injected from 200th instance.

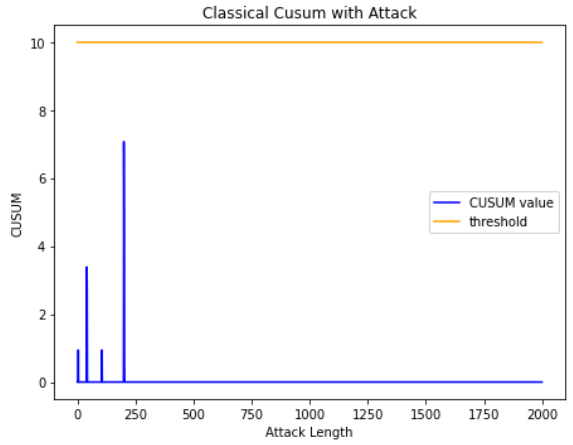
State values under attack at extreme values of $\bar{\psi}_k$ and $\bar{q}_k$ is crossing safety boundary where CUSUM is not exceeding threshold





State values under attack at random values of $\bar{\psi}_k$ and $\bar{q}_k$ is crossing safety boundary where CUSUM is not exceeding threshold

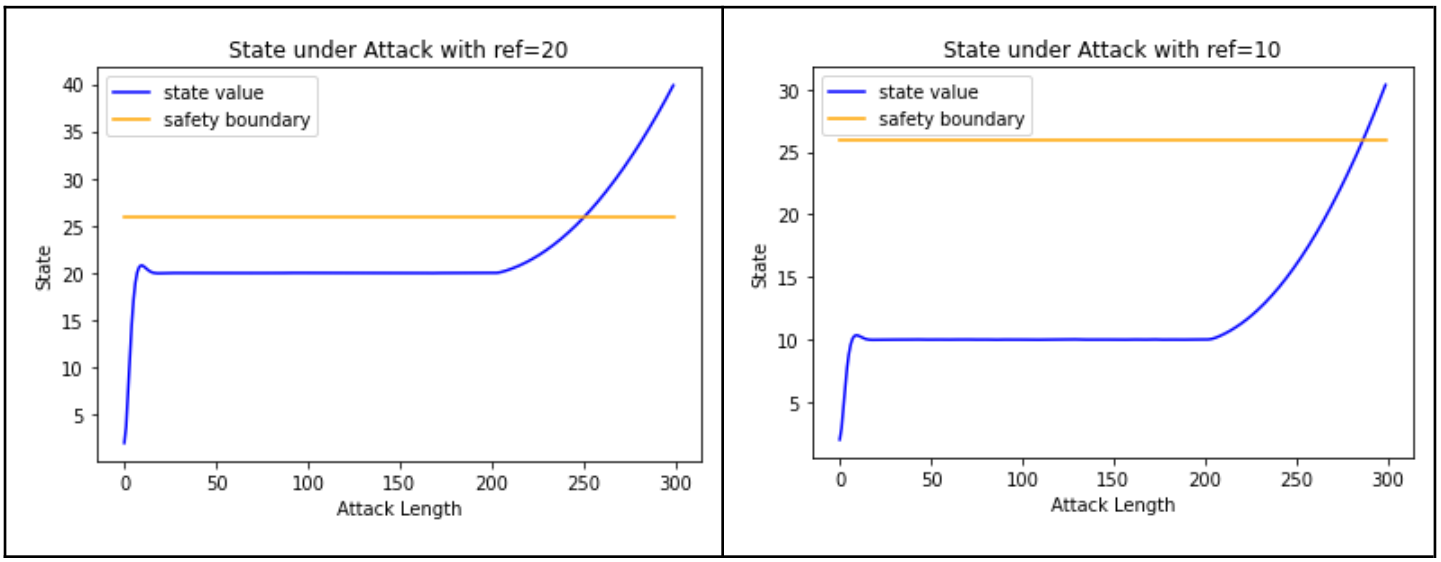




We can observe from result that extreme value attack is vulnerable with attack length 100 than random values of $\bar{\psi}_k$, $\bar{q}_k$ with attack length 1300.

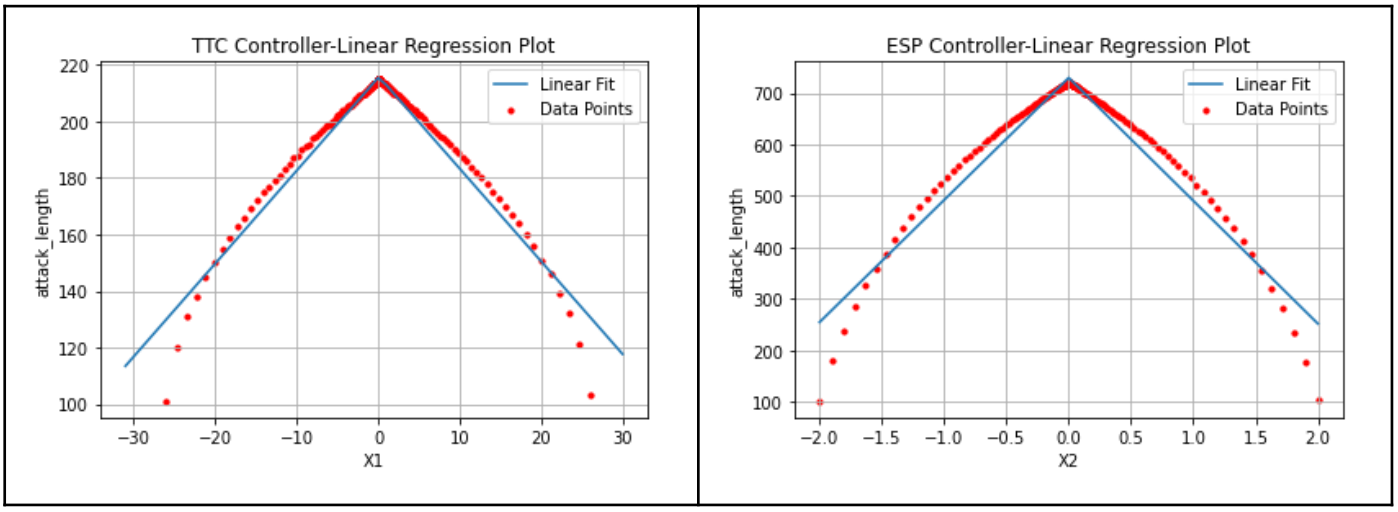
Furthermore, we investigate the relationship between the initial state values and the time it takes for the system to breach the safety boundary.This "attack length"

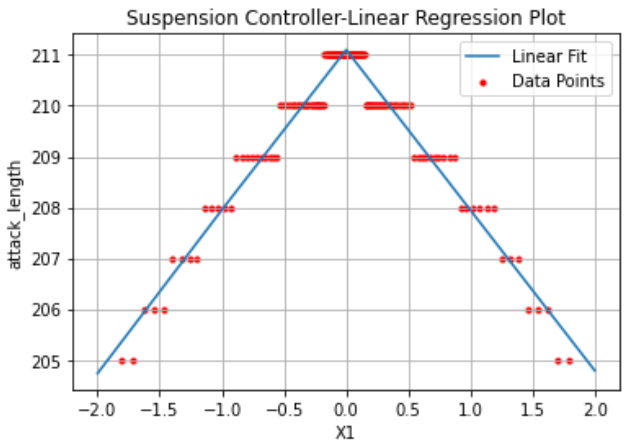
signifies the speed at which the bias should change in response to the attack. For which we experimented with two different reference values, one with Ref=20 where we got attack length 250 and for Ref=10 we got attack length 300. We concluded that the initial value close to the safety boundary has a small attack length compared to initial value far from the safety boundary and we also concluded that for different initial values we get different attack length.



However, determining the attack length in real-time poses challenges due to its complexity. To address this, we employ a linear regression approach to partition the state values, where attacks are generated for different initial values, and the corresponding attack lengths are determined. This method allows us to establish a linear relationship between the initial state values and the attack length, providing a practical means of estimating the attack duration in real-world scenarios.

We will show the results for linear plots for different initial state values with different controllers.





Python code to find attack length for various reference values for TTC Controller.

```
for j in range(150):
    input_len=500
    Xk = np.zeros((2, input_len + 1))
    Xek = np.zeros((2, input_len + 1))
    ek = np.zeros((2, input_len))
    Uk = np.zeros((1,input_len))
    Yk = np.zeros((1,input_len))
    attack= np.zeros((1,input_len))
    rk = np.zeros((1,input_len))
    zk = np.zeros((1,input_len))
    Sk = np.zeros((input_len))
    Xk[:,0] = [2,4]
    Xek[:,0] = [2,4]
    FFG = calculate_FFG(C, G, Gain, F)
    Ref=26*(0.95)**j
    for i in range(input_len):
        ek[:, i] = Xk[:, i] - Xek[:, i]
        Uk[:,i] = -(Gain @ Xek[:, i])+FFG*Ref
        Xk[:, i+1] = F @ Xk[:, i] + G @ Uk[:, i] + n2[:, i]
        if i==100:
            p1=-(t+b-Sk[i-1])**(1/2)
            Table1.loc[j,'X1']=Xk[0,i]
            Table1.loc[j,'X2']=Xk[1,i]
            attack[:,i]=-C@ek[:, i]-n1[:, i]+(covn_r**(1/2))*p1
        elif i>100:
            attack[:,i]=-C@ek[:, i]-n1[:, i]+(covn_r**(1/2))*q1
        Yk[:,i] = C @ Xk[:, i] + n1[:,i] + attack[:,i]
        rk[:,i] = Yk[:,i] - C @ Xek[:, i]
        Xek[:, i+1] = F @ Xek[:, i] + G @ Uk[:,i] + L @ rk[:,i]
        zk[:,i] = rk[:,i].T * (covn_r)**(-1) * rk[:,i]
        mzk.append(zk[:,i])
        if i>100:
            if Xk[0][i]>26 or Xk[0][i]<-26 or Xk[1][i]>31 or Xk[1][i]<-31:
                Table1.loc[j,'Xf1']=Xk[0,i]
                Table1.loc[j,'Xf2']=Xk[1,i]
                Table1.loc[j,'i']=i
                break
```

Other than linear fit plots we can also make a table.Below is an outline of how we generated a table which is mapping different ranges of initial state values to the minimum attack length observed within those ranges:

- Firstly, finding attack length with different state values from minimum to maximum.
- Then for minimum attack length with some difference, let be 5 we divided state values in range.
- Created a table to display the mappings between the ranges of initial state values and their corresponding minimum attack lengths.
- Below are the tables for TTC and Suspension controllers.

TTC Controller:

max	min	length
26	24.7	103
24.7	23.4	122
23.4	22.3	132
22.3	21.2	139
21.2	20.1	146
20.1	19.1	151
19.1	18.1	156
18.1	17.2	160
17.2	15.6	164
15.6	14	170
14	12.7	176
12.7	10.9	180
10.9	9.3	186
9.3	7.6	191
7.6	5.6	196
5.6	3.7	201
3.7	1.7	206
1.7	0	211

max	min	length
-24.7	-26	101
-23.5	-24.7	120
-22.3	-23.5	131
-21.2	-22.3	138
-20.1	-21.2	145
-19.1	-20.1	150
-18.1	-19.1	155
-17.2	-18.1	159
-16.4	-17.2	163
-15.6	-16.4	169
-14	-15.6	175
-12.7	-14	179
-10.9	-12.7	185
-9.3	-10.9	190
-7.6	-9.3	195
-5.6	-7.6	200
-3.5	-5.6	205
-1.5	-3.5	210
0	-1.5	214

Suspension Controller:

max	min	length
1.8	0.2	205
0.2	0	211
-0.6	-1.8	204
0	-0.6	209

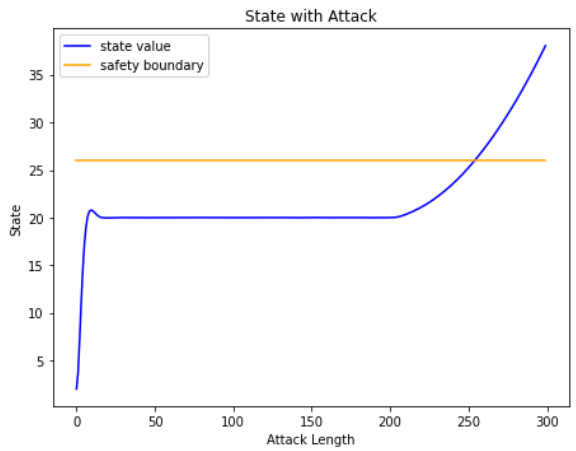
4.2 *Attack Detection:*

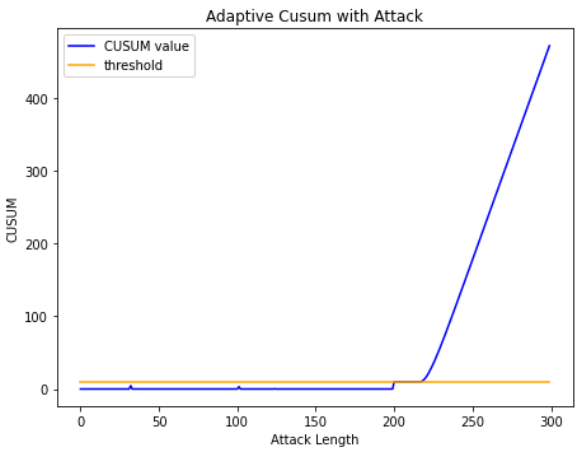
In our detection mechanism, we employ an adaptive approach to adjust the bias when the mean of the variable "z" deviates from the normal range, typically defined as 0.5 to 1.5. Specifically, after the 100th instance, if the mean of "z" exceeds this normal range, we reduce the bias "b" by 1.2 times iteratively. This adaptive adjustment enables us to detect attacks before the system reaches an unsafe state across various controllers.

5.EXPERIMENTAL RESULTS

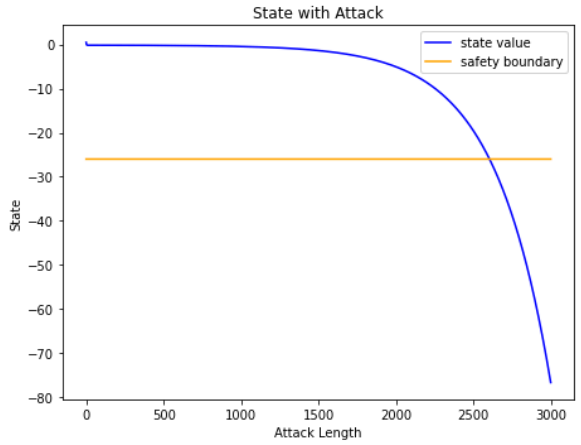
To visualise the effectiveness of our approach, we present plots of the Adaptive Cumulative Sum (CUSUM) values under attack for different controllers. Attacks are given to the system from the 200th instance.These plots illustrate how our adaptive detection mechanism successfully identifies attacks before the system becomes unsafe and we can also observe the false alarm is not generated before the attack injection, thereby highlighting the robustness and efficacy of our approach across diverse controller settings.

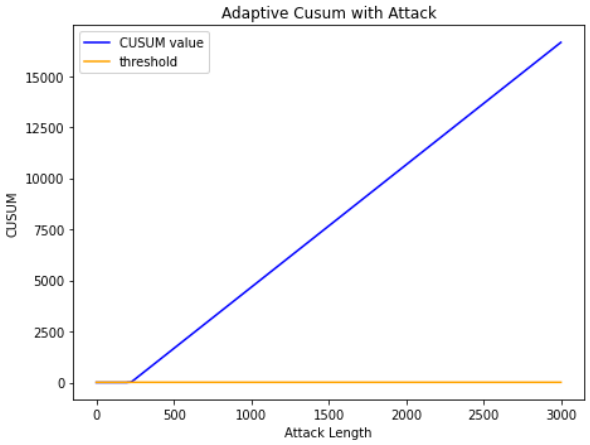
TTC Controller



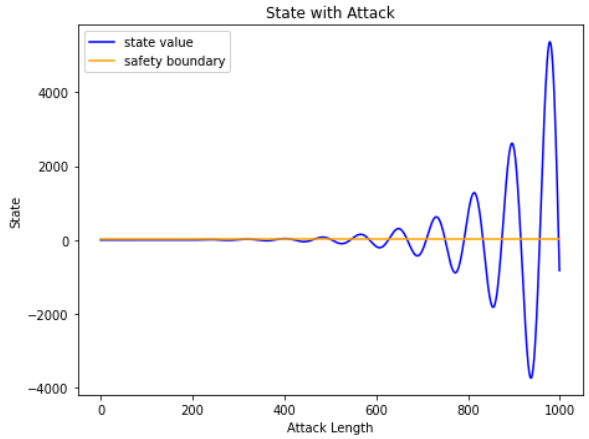


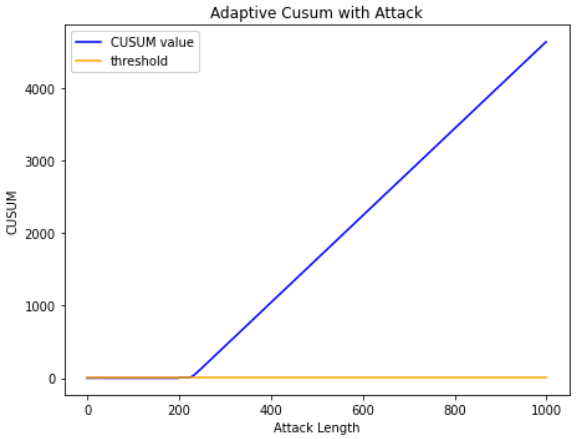
ESP Controller:





Suspension Controller:





6.CONCLUSION

Our investigation into the relationship between initial state values and attack length yielded crucial insights into our system's response dynamics. We found that initial values closer to the safety boundary led to shorter attack lengths, indicating a correlation between proximity to the boundary and the speed of response to attacks. Additionally, varying initial values resulted in different attack lengths, highlighting the sensitivity of system behaviour to initial conditions.

In parallel, our experimentation with an adaptive approach for attack detection showcased promising results. By iteratively adjusting the bias "b" when the mean of the variable "z" deviated from the normal range, we effectively identified attacks before the system reached an unsafe state across multiple controllers. Visual

representations of Adaptive Cumulative Sum (CUSUM) values under attack underscored the robustness and efficacy of this approach.

In conclusion, our findings underscore the significance of both understanding the dynamics of initial state values on attack length and implementing adaptive detection mechanisms. By considering these factors, we can develop more effective defence strategies tailored to our system's characteristics, ultimately enhancing security and resilience in diverse controller settings.

7.FUTURE WORK

In future work, we intend to integrate attack length dynamics into our adaptive attack detection approach to reduce false alarm rates. By establishing a mathematical relationship between attack length and bias adjustment, we aim to develop a more precise and proactive defence mechanism. This strategy will allow us to adjust biases based on observed attack lengths, thereby enhancing our system's ability to differentiate between genuine threats and benign fluctuations. By minimizing false alarms, we can improve the overall reliability and efficiency of our defence system, ultimately contributing to enhanced security in real-world applications.

8.IMPACTS

IMPACTS OF THESIS PROJECT IN REAL LIFE APPLICATION:

In the domain of cyber-physical systems, the potential impact of my work in attack detection is promising and far-reaching. Through the development of an Adaptive CUSUM detector that dynamically adjusts its bias based on the mean of z, this work laid the foundation for enhancing the security and reliability of critical infrastructure. This innovative approach addresses the growing concern of stealthy attacks that have the capacity to bypass traditional detection methods and provides a versatile framework that can adapt to varying detector parameters and initial conditions and offers a proactive means to safeguard interconnected systems.

9.REFERENCES

1. *Carlos Murguia et al. 2019.* On model-based detectors for linear time-invariant stochastic systems under sensor attacks. IET Control Theory & Applications.
2. *Ipsita Koley et al. 2021.* Catch me if you learn: real-time attack detection and mitigation in learning enabled cps. In RTSS. doi: 10.1109/RTSS52674.2021.00023.
3. *Yi Huang, Jin Tang, Yu Cheng, Husheng Li, Kristy A. Campbell, and Zhu Han,* Real-Time Detection of False Data Injection in Smart Grid Networks: An Adaptive CUSUM Method and Analysis.
4. *Ipsita Koley ,Sunandan Adhikary, Soumyajit Dey* An RL-Based Adaptive Detection Strategy to Secure Cyber-Physical Systems.