Fishy Cyber Attack Detection in Industrial Control Systems
A new approach based on sequence learning LSTM networks

Manikanta Reddy D.[1]
Supervised by Prof. Sandeep Shukla[1]
[1]*Indian Institute of Technology, Kanpur*
manikant@cse.iitk.ac.in, sandeeps@cse.iitk.ac.in

Keywords: Industrial Control Systems, Attack Detection, LSTM, Cyber-Security, Concealment, Machine Learning

Abstract: Cyber attacks have become serious threats to Industrial Control systems too. It becomes important to develop a serious threat defense system against such vulnerabilities. For such process control systems, safety should also be assured apart from security. These critical infrastructures need safeguards to prevent accidents, both natural and artificial, that could potentially be hazardous.
Morita proposed an effective Zone division(Morita et al., 2013), capable of evaluating remote and concealed attacks on the system, coupled with Principal Component Analysis. But the need to analyze the node that has been compromised and to stop any further damages, requires an automated technique.
Illustrating the basic ideas we'll simulate a simple Water plant. We propose a new automated approach based on Long Short Term Memory networks capable of detecting attacks and pin point the location of breach.

1 Introduction

In the age of Internet, cyber attacks have become a major threat. Until recently only private and information centered systems were breached. But now, cyber attacks are a threat to Industrial Systems as Well. Serious security vulnerabilities are patched in regular personal computers and commercial spaces, quite frequently but Industrial control systems are seldom fixed as these patches could lead to new conflicts in the system. This opens up a wide space for attackers to sneak in. A prime example would be of Stuxnet, that sabotaged Iranian uranium enrichment facilities in 2010.

1.1 The Stuxnet malware

Stuxnet utilized existing vulnerabilities in the operating system, along with a good understanding of the PLCs to formulate a malware. Its only target was to manipulate the working of centifuges that enriched Uranium. Such carefully crafted attacks cannot be prevented unless a perfect system is built. Instead it is enough to detect such breaches and then assess the damage. Thus it is important to build a reliable security and safety mechanism to prevent against attacks like Stuxnet.

1.2 Zone Based PCA

Hashimoto proposed a method of securing the information system by dividing the network into "plural zones". By Zone Division (Hashimoto et al., 2013) the probability of detecting possible attacks and accidents is increased.

Conjoined with PCA (Morita et al., 2013), Zone Based PCA can analyze the relationship between the variables in *plural zones* and detect any changes caused by potential concealed breaches and unintended accidents. There can many types of relationships between the variables that are analyzed differently by Zone Based PCA. An on-board safety personnel, is still required, who may notice the change in PCA variables and report an anomaly.

1.3 Sequence Learning

We propose a new approach based on sequence learning algorithms, to detect changes from regular working unlike relationships between variables. The neural network architecture, utilized is capable of learning how the normal functioning of the system looks like and detect if something has changed.

Inorder to illustrate the working of the system, we simulate a simple water plant, that circulates hot water between two tanks.

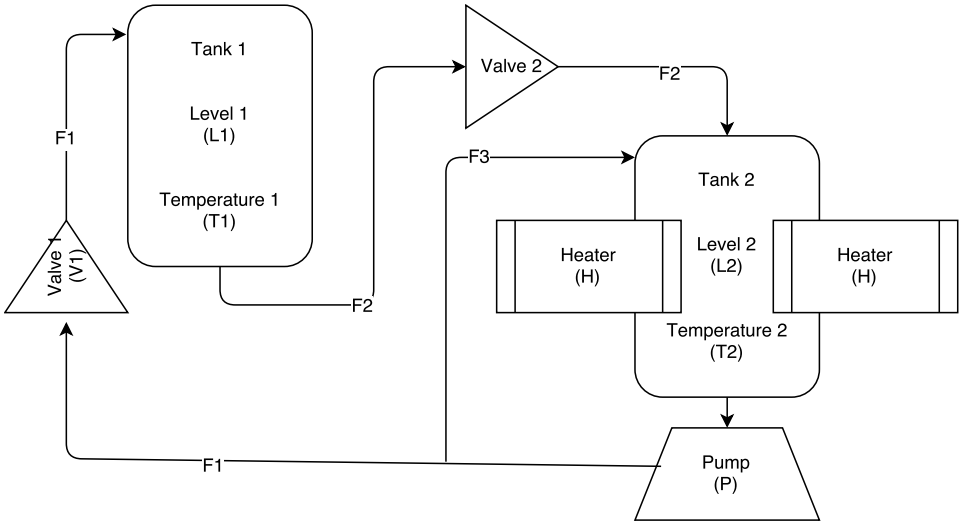

Figure 1: Simulated Water Plant

2 Simulated System

The simulation is a very basic version of plant. In this system water circulates between two tanks (Tank1 and Tank2). The systems contains SCADA and other operators. The plant consists of many sensors and controls.

2.1 Variable in the System

L1 and L2 are levels of water in Tank1 and Tank2 respectively. Similarly T1 and T2 measure the temperatures. The Heater (H) provides heat to increase the temperature of water in Tank2. The Pump (P) pumps the heated water into Tank1. Valves V1 and V2 are controlled to allow water to flow across them. Tank1 is assumed to radiate heat and cool down the water.

We define two kinds of variables in our system. Process variables are the ones that are measured by sensors. Control Variables can be manipulated and change the state of the plant.

Process Variables: L1, L2, T1, T2, F1, F2, F3
Control Variables: V1, V2, H, P

2.2 Network Configuration of the System

The network is divided into two control zones, such that the control of a control variable that directly affects a process variable is in a different zone.

Zone 1: L1, T1, V2, F2, H
Zone 2: L2, T2, V1, F1, F3, P

A control variable in a zone cannot directly affect process variables in the same zone.

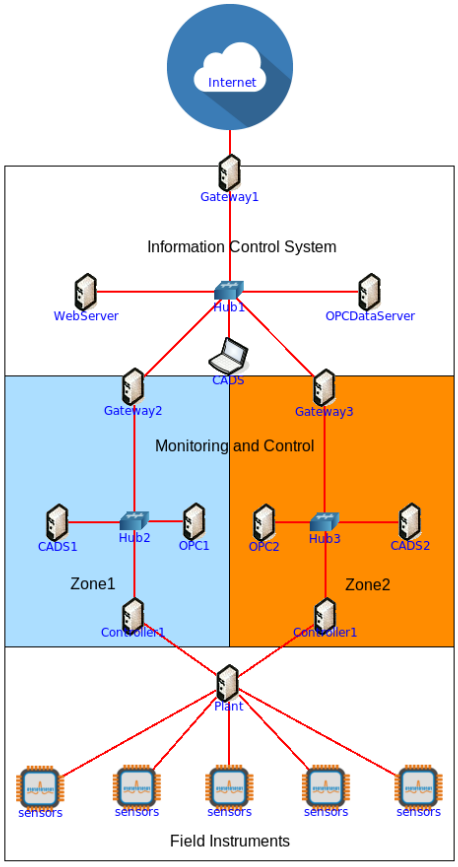

Figure 2: Information Flow Network of the Water Plant
OPC: OLE for Process Control,
CADS: Cyber Attack Detection System

3 Zone Based PCA

Principal Component analysis re-projects the data in study into new space, with coordinates of high variance. Thus the variables with high variance can be maximally noted across them. PCA in a simple sense, brings down the high dimensionality of the data to a smaller number. In our experiments we've considered top 3 coordinates, in decreasing order of variance.

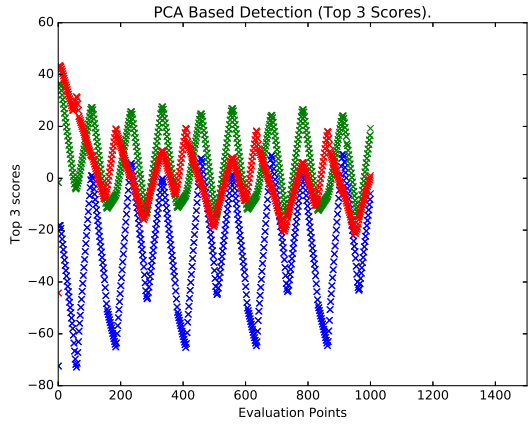

Figure 3: PCA projection of an un-compromised plant.

The order of variance is as follows; blue, green, red in decreasing order. The periodic nature of the plot is due to the way water is circulated. Now we'll simulate a data injection attack with concealment.

The attack manipulates Zone 2 and takes over the control of variables V1 and P. By setting them both to 0, the water in Tank1 doesn't decrease and the temperature of Tank1 increase to the extent where it depressurises.

The attack begins at simulation point 500.

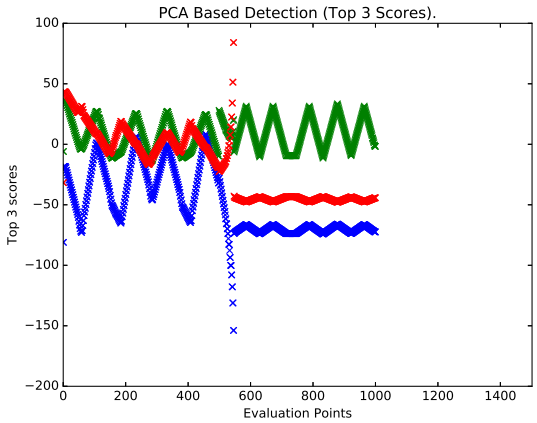

Figure 4: PCA projection of compromised plant. Zone 2 variables are remotely manipulated.

Notice how drastically the projection in blue changes its behaviour. Any on-board personnel can realize such a change of high magnitude and shut down the systems if necessary.

It becomes difficult with increasing types of manipulation to identify what change produces which kind of behavior in the PCA projection.

4 Zone based LSTM network

The PCA method evaluates relationships between variable and raises alarm when the relationships or the dependencies between the variables changes. The relationships can change in many different ways when the number of variables involved is very large.

Instead we should focus on differentiating the modes of running. A normal working state is very different from a compromised state. As the systems tend to repetitive work, it is not hard to notice that there is a pattern to the way an un-compromised plant produces data.

In order to acheive this, we use sequence learning LSTM networks in our method.

4.1 Long Short Term Memory

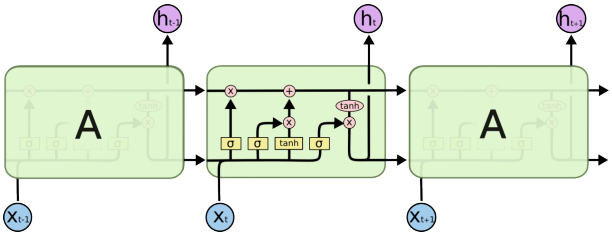

Figure 5: Long Short Term Memory, LSTM, Units in action. Img ref: colah

LSTM(Hochreiter and Schmidhuber, 1997) is a form of Recurrent Neural Network, it remembers what's important and forgets the trivial things. It is capable of learning patterns in data and identify those that don't match.

LSTM achieves this by what we call memory control gates in the unit. There are different gates to control and filter data. The gates decide if what variables at a simulation point are important. By doing so repeatedly, it learns what values of variables to look for to understand the sequence pattern. It then assigns a score , that signifies how strongly it believes in the pattern.

We train an LSTM network with a logged data of the normal functioning of plant, and use it to give a confidence score to the previous 50 points at every

evaluation point. This enables us to know if the current sequence of data generated is in high correlation with data generated during normal functioning.

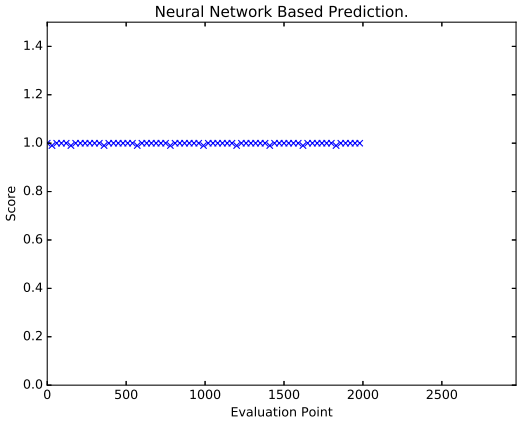

Figure 6: LSTM score of an un-compromised plant.

Normal functioning ensures a score match near close to 1, which is expected, now we compromise zone2 as we did earlier. Both P and V1 are set to 0 here.

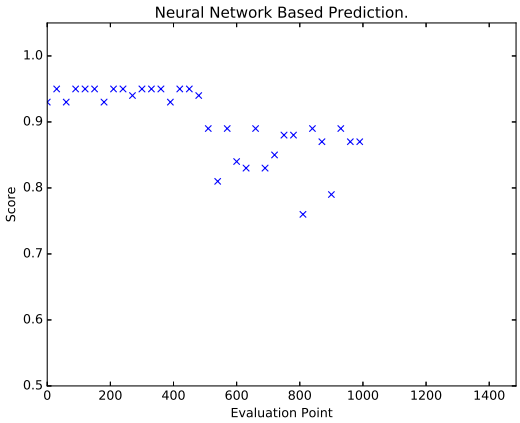

Figure 7: LSTM score of compromised plant.
Zone 2 variables are remotely manipulated and P is set to 0.

It is to be noted that when the zone is compromised, the plant works produces a different pattern of data. Then the LSTM networks gives us a correlation score with the actual pattern of functioning, which in this case, when P is set to 0 averages around a value of 0.85.

When the variable P is set to 1, the pattern of execution produces scores averaging around 0.8.

By thresholding at different values we can easily identify which variable the attacker has modified. We can also detect the zone which has been compromised.

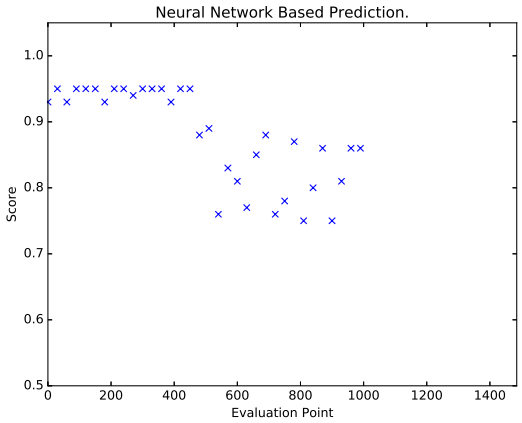

Figure 8: LSTM score of compromised plant.
Zone 2 variables are remotely manipulated and P is set to 1.

A plausible thresholds to detect the variable could be.
P = 0: score $\in$ (0.8, 0.9)
P = 1: score $\in$ (0.75, 0.85)

5 Conclusion

We proposed a new way of detection of attacks in Industrial systems. The method based on LSTM is also capable of diagnosing the attack for points of failure. Sequence based learning and anomaly detection has advantage over PCA based method in this regard. This approach shows an example of interdisciplinary work on implementing machine learning technologies for tackling the problems of Industrial systems and Networks.

REFERENCES

Hashimoto, Y., Toyoshima, T., Yogo, S., Koike, M., Hamaguchi, T., Jing, S., and Koshijima, I. (2013). Safety securing approach against cyber-attacks for process control system. *Computers & Chemical Engineering*, 57:181–186.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.

Morita, T., Yogo, S., Koike, M., Hamaguchi, T., Jung, S., Koshijima, I., and Hashimoto, Y. (2013). Detection of cyber-attacks with zone dividing and pca. *Procedia Computer Science*, 22:727–736.