Enhancing Network Data Into Cyber-Physical Data For Better Attack Detection Performances

Côme Frappé-Vialatoux, Pierre Parrend

Enhancing Network Data Into Cyber-Physical Data For Better Attack Detection Performances

1st Côme Frappé - - Vialatoux *†, 2rd Pierre Parrend*†
*ICube, UMR 7357, Université de Strasbourg, CNRS, 67000 Strasbourg, France
†Laboratoire de Recherche de L'EPITA (LRE), 94270 Le Kremlin-Bicêtre, France
come.frappe-vialatoux@etu.unistra.fr, pierre.parrend@epita.fr

***Abstract*—Critical systems are an essential component of today's society, ensuring needs such as water distribution or power supply. The modernization effort of these infrastructures through a general increase in connectedness allows for better efficiency, monitoring, and safety, but also comes with an increased vulnerability to cyber-attacks. Detecting such cyber-attacks as early and accurately as possible is a hard task for which machine learning (ML) algorithms show promising results, leveraging the huge amount of data that network traffic traces constitute. However, cyber-attack also have measurable impacts on physical systems, but the use of data from such systems is lacking due to a scarcity of available datasets and analysis tools. The recent release of cyber-physical datasets, which captured data from both network communication and physical systems, fills this gap, allowing a joint usage of these two data sources. This paper provides a multi-layer methodology for detection in cyber-physical systems, by combining physical and network data and evaluates its gain in detection performances on multiple ML models.**

***Index Terms*—cyber-security, machine learning, cyber-physical systems**

I. Introduction

As the attacks conducted on critical systems evolve from on-site physical attacks to cyber-attacks [1], the need to develop robust solutions to detect and deter such attacks has become a major research field. ML algorithms have proved to be especially efficient at this task, leveraging the huge amount of data that network traffic traces can generate [2]. However, despite network communications being the primary vector of cyber-attacks, network traffic traces alone cannot reflect cyber-attacks effects on physical processes. This paper demonstrates that the conjunct use of traces from physical and cyber communication systems greatly enhances the detection capability thus increasing the reactivity of security operators by reducing the time to detection for ongoing attacks against cyber-physical systems, as in the example of water distribution network. This paper is organized as follows: first, we show how physical data can capture the effects of cyber-attacks, then we describe our methodology of combining physical and network data. It is followed by a presentation of the effects of this methodology on ML performances and a conclusion.

II. Effects of Cyber Attacks on Physical Systems

Cyber-attacks conducted on critical infrastructures often aim at disrupting their normal operation. Depending on the type of attack, this disruption can be very well captured by the physical data, for instance in the case of Denial of Service (DoS) or False data injection through Man-in-the-middle (MITM), but other attacks such as reconnaissance techniques like scans have no effects at all on physical processes. Amongst critical systems, the water distribution sector has an inherent need to monitor its physical processes because of health stakes [3]. This sector has thus been proficient at releasing open datasets [4] [6] that capture the effects of cyber-attacks on physical processes, as well as cyber-physical datasets [5] [7] that contain both network traces and physical data. As an example, Figure 1 shows the effect of a DoS attack on a tank water level reading from the *Hardware In the Loop* (HITL) dataset [5]. In this graphic we can observe that the DoS effects are both delayed and lingering, reflecting respectively the time until the sensor is incapacitated after the attack started and the time for it to return to a normal state after the attack ended.

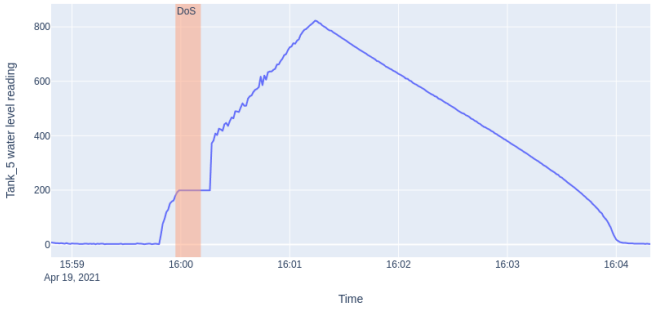

Fig. 1. Effect of a DoS attack on a tank water level reading from HITL testbed

III. Multi-layer approach

As shown in the previous section, physical data contains information that can complement that from network traffic traces. For ML algorithms to leverage these two layers of information, a combination methodology is needed.

A. Methodology

The methodology is as follows: we duplicate each physical data row as many times as there are network communications between this physical data and the consecutive one, thus equalizing the number of data points. Then for each network data row, we add the columns of one of the physical data with the closest inferior time value.

This methodology uses network traces as a base because the time scale of physical processes is much greater than that of network communications, with differences ranging from tenfold to thousandfold in the number of data. By using the data with the finest time granularity as a base for the combined data, we eliminate the need for aggregation and the information loss of such operation. As both network communication and data from physical processes require time information for them to be exploited, using this information for the joint ensures the viability of this methodology.

B. Side effects

In the case of labelled data, this methodology can create two types of inconsistencies in the labels between the physical and network data. The first type is a misalignment of the labels. It happens for attacks with a duration greater than the time granularity of the physical data, because the finer-grained data will start being labelled before the higher-grained data will be updated, and reciprocally for the end of the attack, resulting in the labels in the higher-grained data being updated later at the start and at the end. The second type of inconsistency is the occultation of labels. It happens for attacks with a duration inferior to the grain of the physical data, only if the entirety of the attack span is contained within a single timestep of higher-grained data. Thus, no update will occur on these data during the whole attack duration, resulting in the label appearing only in the finer-grained network data. However, these inconsistencies being characterized and detectable, they can easily be taken into account to avoid biases in the detection.

C. Results

The methodology described in the previous section has been applied to the HITL dataset to combine the physical data and the network traces. Four ML models were trained, namely Decision-tree (CART), Random Forest, XGBoost and Multi-Layer-Perceptron, on the 3 datasets independently: only physical data (around 11k rows), only network data (around 30M rows), and the combined datasets, using the labels from the network dataset. The results shown in Table I demonstrate the benefit of combining the two datasets, with an increase in balanced accuracy of **+22.34%** and **+47.96%** compared respectively to only physical data and only network data. Furthermore, Figure 2 shows the evolution of performance from XGBoost on 8 different metrics, where it can be observed that the combined dataset allows for minor to major improvement in each of these metrics.

TABLE I
CLASSIFIER PERFORMANCE METRICS ON SEPARATED AND JOINED DATASETS

Dataset	Classifier	Accuracy	Balanced Accuracy	MCC	Fit Time	Pred Time
Combined	DecisionTree	**0.9994**	**0.9836**	**0.9987**	2m 11s	0m 0s
	RandomForest	0.8600	0.3195	0.6971	11m 41s	0m 30s
	XGB	0.9993	0.9503	0.9985	1m 52s	0m 5s
	MLP	0.9406	0.6259	0.8747	48m 0s	0m 31s
Physical	DecisionTree	0.9621	0.7515	0.8868	0m 00s	0m 00s
	RandomForest	0.8410	0.2805	0.3833	0m 00s	0m 00s
	XGB	**0.9780**	**0.7602**	**0.9327**	0m 00s	0m 00s
	MLP	0.9160	0.6336	0.7383	0m 01s	0m 00s
Network	DecisionTree	0.8703	**0.5043**	0.7198	0m 27s s	0m 00s
	RandomForest	0.8630	0.3222	0.7045	6m 52s s	0m 30s
	XGB	**0.8717**	0.4978	**0.7251**	1m 02s s	0m 4s
	MLP	0.8545	0.3147	0.6838	24m 07s	0m 16s

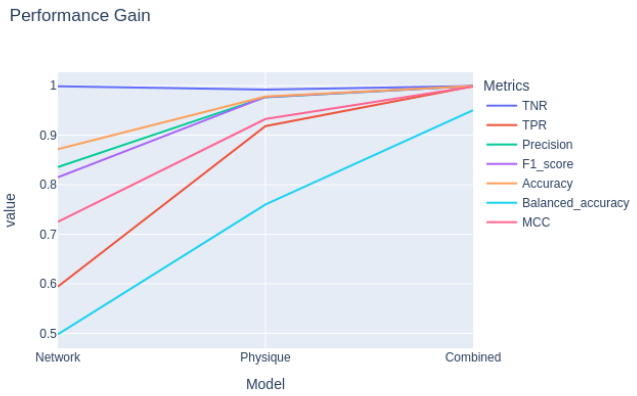

Fig. 2. Effect of joining datasets on multiple metrics

IV. CONCLUSION

Cyber-attacks have measurable impacts on both network traces and physical processes. Considering data from both to train ML models greatly benefits detection performances compared to using only one of them. The combination methodology proposed in this article allows to take advantage of this duality, enabling the training of ML models on both data at the same time. The improvement in detection performance has been assessed on the HITL cyber-physical dataset, with a balanced accuracy 22.34% higher than when using physical data only and 47.96% higher than when using network traces only. Further work on applying this methodology to other cyber-physical datasets and evaluating its effects will allow for greater insights into its benefits.

REFERENCES

[1] A. Hassanzadeh et al., 'A Review of Cybersecurity Incidents in the Water Sector', Journal of Environmental Engineering, Sep. 2019.

[2] Y. Xin et al., "Machine Learning and Deep Learning Methods for Cybersecurity," IEEE Access, vol. 6, pp. 35365–35381, 2018, doi: 10.1109/ACCESS.2018.2836950.

[3] U. Szewzyk, R. Szewzyk, W. Manz, and K.-H. Schleifer, "Microbiological Safety of Drinking Water," Annu. Rev. Microbiol., vol. 54, no. 1, pp. 81–127, Oct. 2000, doi: 10.1146/annurev.micro.54.1.81.

[4] R. Taormina et al., "Battle of the Attack Detection Algorithms: Disclosing Cyber Attacks on Water Distribution Networks," Journal of Water Resources Planning and Management, vol. 144, no. 8, p. 04018048, Aug. 2018, doi: 10.1061/(ASCE)WR.1943-5452.0000969.

[5] L. Faramondi, F. Flammini, S. Guarino, and R. Setola, "A Hardware-in-the-Loop Water Distribution Testbed Dataset for Cyber-Physical Security Testing," IEEE Access, vol. 9, pp. 122385–122396, 2021, doi: 10.1109/ACCESS.2021.3109465.

[6] C. Ahmed, V. Palleti, and A. Mathur, "WADI: a water distribution testbed for research in the design of secure cyber physical systems," Apr. 2017, pp. 25–28. doi: 10.1145/3055366.3055375.

[7] J. Goh, S. Adepu, K. Junejo, and A. Mathur, A Dataset to Support Research in the Design of Secure Water Treatment Systems. 2016.