EPiC Series in Computing

Volume 82, 2022, Pages 63–71

Proceedings of 37th International Conference on Computers and Their Applications

Entropy Analysis for Modbus Traffic over TCP/IP in Industrial Control Systems

Logan Day[1], Tirthankar Ghosh[1], Sikha Bagui[1], Subhash Bagui[1]
[1]University of West Florida
ldd14@students.uwf.edu, tghosh@uwf.edu, bagui@uwf.edu, sbagui@uwf.edu

Abstract

Anomalies in network traffic are usually detected by measuring unexpected deviation from what constitutes a baseline. Several statistical techniques have been proposed to create baselines and measure deviation. However, simply looking at traffic volume to find anomalous deviation may result in increased false positives. Traffic feature distributions need to be created, and deviations need to be measured for these features. An effective approach to finding anomalous deviations starts with entropy analysis on these features. In this paper, we presented an initial entropy analysis on an industrial control system network using selected features with datasets obtained from an HVAC system. We started with the fundamental question: whether a preliminary entropy analysis on Modbus-over-TCP data using only a few TCP/IP features without going into the Modbus traffic itself gives us information about an anomaly in the network. We acknowledge that the initial entropy analysis provides only a starting point that would lead to several questions and investigating relevant issues resulting in an optimal system design and implementation. *

1 Introduction

Industrial Control Systems (ICS) are networks of devices used in critical infrastructure and industrial environments for control of physical processes. These networks typically span a large geographic area, and some examples of such systems are water distribution systems, gas pipelines, and power transmission systems. An ICS can be a large multifaceted infrastructure like Supervisory Control and Data Acquisition (SCADA) system, which collects data and processes it in a centralized environment where it can be viewed and interacted with. There are simpler configurations of industrial control systems as well that are more readily available to any system with lesser changes to an already-

* This work is partially supported by the Office of Naval Research (ONR) under grant number N00014-21-1-2025.

established system. Programmable Logic Controllers (PLC) are widely used in the area of automation and come equipped with programmable memory, various input and output channels, and communication interfaces that make them incredibly valuable.

Industrial Control Systems are fundamentally different from traditional Information Technology (IT) systems. Because ICS provides an interface with physical devices like sensors and actuators, they are categorized in general as Operational Technology (OT). Updates and patching play a pivotal role in the differences in maintainability between IT and OT systems (Cardenas, 2009). An ICS requires real-time availability for their systems to communicate and function, implying that they must be able to identify, diagnose, and respond appropriately to irregular flow of information as quickly as they appear. This necessitates that updates and patching happen as infrequently as possible, making them vulnerable to potential attacks. While this is a drawback that is difficult to circumvent, the design of an intrusion detection system may be simpler for ICS because of their static nature and predictability in communication.

Intrusion Detection and Prevention Systems (IDPS) are designed and deployed to monitor traditional IT systems for signs of undesirable and malicious behavior. A common way to secure a traditional system is to deploy an IDPS to monitor and identify anomalies that take automated action or warrant future investigation. There are two broad classifications of IDPS based on their detection technique – misuse detection or signature-based, and anomaly detection or behavior-based (Angséus & Ekbom, 2017). Signature-based systems are not very effective in an industrial control systems environment where very few known signatures are encountered. An effective way to identify communication anomaly in an ICS is to consider the minimal and maximal values in a given feature within the system and have precautionary measures in place that can react to values that defy the provided range (Angséus & Ekbom, 2017). In other words, measuring the change in entropy or randomness in the communication pattern of an industrial control system is beneficial in detecting a problem.

There are four levels in a SCADA architecture that need to be secured to ensure a stable environment (Koucham, 2019). The lowermost level, or the basic control level, generally consists of sensors and actuators that collect and send information to the upper levels. These machines are found at field sites with Remote Terminal Units. Above the basic control is the supervisory control level consisting of Distributed Control Systems (DCS) servers and PLCs which are accompanied by the Human Machine Interface (HMI) and engineering stations. This level focuses on the global view of the system's control state and operations and collects the information relayed from the first level for analysis. This information is also presented through the HMI to make it more easily accessible to a user reviewing the system's nodes and features. The engineering workstations allow the specification of setpoints and the programming of controllers, which allows for boundaries to be manipulated. The two uppermost levels are commonly grouped together as the "backbone network" and contain servers that are connected to the enterprise IT systems backbone. This level has a variety of purposes, which can most broadly be classified as the allocation and optimization of resources, maintenance, planning, and quality control. Data collected previously is housed within this level in database servers (Koucham, 2019).

Communication characteristics in an ICS setting are also quite different from traditional TCP/IP communication. Sensors and actuators generate a low volume of data that is periodic with short transfer time and low delay (Koucham, 2019). The controllers that accompany these nodes use a communication protocol like Modbus (Koucham, 2019) which is an application protocol that defines the syntax and semantics of the communication and structure. The Protocol Data Unit (PDU) is seven bytes long and consists of the transaction identifier, protocol identifier, length, and unit identifier. Figure 1 shows the Modbus frame and its encapsulation in the TCP header. The transaction identifier is used for transaction pairing when multiple messages are sent and make up two of the seven bytes. The protocol identifier also makes up two bytes and is either empty or padded with zeros to be used for future extensions. The unit identifier is one byte and identifies a remote server located on a non TCP/IP network, and the length

is the byte count of the remaining fields. Modbus is open source making it the most widely used protocol in ICS environments (Koucham, 2019).

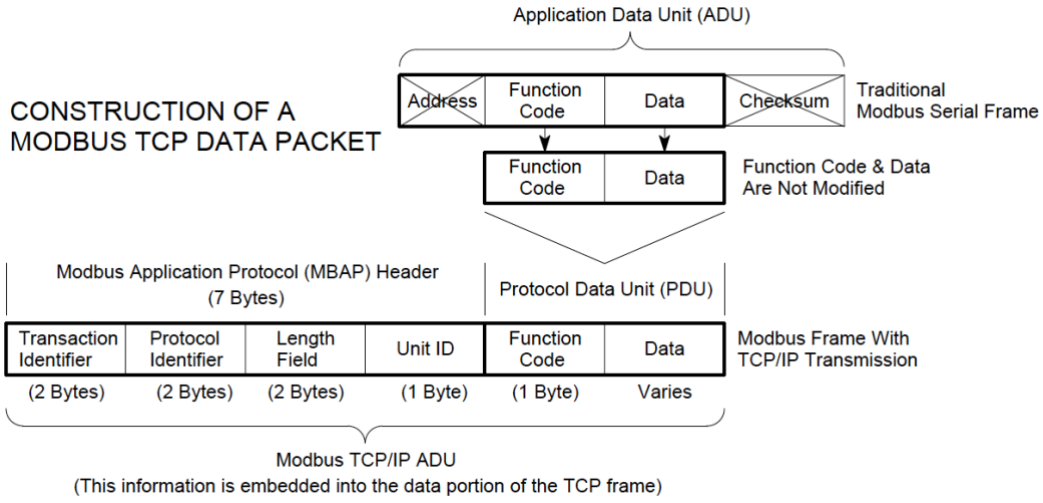

Figure 1: Modbus Frame[†]

In the Supervisory Level, a much higher volume of data is collected than sensors and actuators, and transfer time is restricted to a lesser extent. This level is also much more representative of a traditional IT system, and therefore can be treated as such. Network protocols like OPC DA and OPC UA are commonly utilized in this level, as they cover data access and client/server technology which are important to enable human interaction with the system and the data collected (Koucham, 2019).

Although, because of their static nature and predictability in communication, securing industrial control systems may seem trivial at first, but there are several challenges that arise because of these very characteristics. We previously mentioned that having a periodic system update is often a challenge in ICS environments, and real-time communication poses a challenge for deploying a security solution that adds latency. In addition, because of the operational technology requirements of ICS, physical interaction with the sensors and actuators poses a challenge to secure the system as a whole. Any intrusion detection and prevention systems that need to be designed for these environments must take into account the need for real-time data transfer and must be aware of stringent latency requirements. Hence, adding detection techniques that are resource-intensive may not be the most efficient approach in these ICS environments.

In this paper, we have discussed a preliminary entropy analysis on ICS Modbus over TCP/IP data and an analysis of relative entropy using Kullback-Leibler divergence. The entropy represents the amount of uncertainty that exists in a random variable X. Suppose the random variable X takes on values $x_1, x_2, \cdots, x_n$ with respective probabilities $p(x_1), p(x_2), \cdots, p(x_n)$, then the entropy of the random variable X is given by

[†] Acromag Modbus TCP/IP Technical Reference

$$E(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i)$$

which is equivalent to

$$E(X) = \sum_{i=1}^{n} p(x_i) \log(1/p(x_i))$$

where, $E(X)$ is the entropy value for the selected feature X , $p(x_i)$ is the probability that the feature X takes the value x_i . The quantity $-\log p(x_i)$ represents the surprise evoked if X takes on the value x_i . Thus, the expected amount of surprise after learning X is known as the entropy of the random variable X .

Looking at entropy values of a certain traffic feature over a period of time will provide a basis for possible anomalous deviation. However, if traffic features are analyzed independently and their entropy computed, it may not reflect an anomalous behavior and will lead to high false positives. In that respect, relative entropy analysis needs to be conducted. Relative entropy is computed as below:

$$D(p \,\|\, q) = \sum_{i=1}^{n} p(x_i) \log(p(x_i)/q(x_i))$$

which shows the relative entropy or deviation of the probability distribution $p(x)$ from the probability distribution $q(x)$. $D(p \,\|\, q)$ is called a measure of Kullback-Leibler divergence. The probability $q(x)$ is the probability distribution of normal traffic and $p(x)$ is the probability distribution of malicious traffic.

2 Related Literature

Supervised learning techniques have been proposed by many to detect malicious traffic in industrial control systems networks. In (Eigner, 2018), the authors presented a prototype implementation of an anomaly-detection approach based on Naive Bayes. They ran simulated attacks - Denial of Service (DOS) and Man-in-the-Middle (MITM) - and did a preliminary analysis using the Bayesian classifier. The results were preliminary, and did not justify why and how a Bayesian classifier will be the most appropriate for an ICS network. In (Anthi, 2020), the authors used a supervised learning approach to detect attacks. Their proposed attack detection system not only detects malicious packets but also classifies them as specific attack types. A limited number of features from the Modbus data was selected for classification. In (Goh, 2017), the authors proposed to use Long Short Term Memory Recurrent Neural Network (LSTM-RNN) to detect a sequence of patterns for anomaly detection. It was used as a predictor to model normal behavior and then the cumulative sum method was used to identify malicious behavior. The authors tested the approach with data from only one process, limiting its scope and applicability. In (Valdes, 2009), the authors used pattern-based and flow-based anomaly detection techniques to identify malicious communications. The pattern-based approach used patterns of communicating hosts to identify normal communication, whereas the flow-based approach used

network flow to identify traffic patterns. These approaches typically work well in ICS environments as process-related communications are static and predictable; however they sometimes tend to be resource-intensive as the network scales up. In (Feng, 2017), the authors proposed a two-stage anomaly detection technique using a packet-level signature analysis and a time-series analysis using Long Short Term Memory Neural Network (LSTM -NN). In the signature analysis stage, a signature database for normal behavior was constructed using communication patterns among devices. The database was then passed through a Bloom filter to detect anomalous behavior. The time series analysis involved using LSTM-NN to learn the most likely package signatures from previously seen network packages. The authors used a SCADA dataset obtained from a gas pipeline to test their proposed technique. In (Caselli, 2009), the authors proposed a sequence-aware intrusion detection system that relied on pattern identification of ICS network events, extraction of semantic meanings, and modeling known behaviors over time. They used discrete-time Markov chains to describe several ICS-specific operations and proposed a detection mechanism based on the computation of weighted distance among Markov chain states. In (Yang, 2006), the authors proposed an anomaly detection technique using the auto-associative kernel regression model and statistical probability ratio test and applied the technique on a simulated SCADA network.

Not much work has been done using unsupervised learning to isolate malicious from non-malicious traffic. One persistent problem that remains is threshold selection. In (Almalawi, 2020), the authors proposed an approach called global anomaly threshold to unsupervised detection (GATUD) that is used as an add-on component to improve the accuracy of unsupervised intrusion detection techniques. They used K-means clustering to initially learn two labeled small datasets from the unlabeled data; each dataset represents either normal or abnormal behavior. Then, a set of supervised classifiers were trained to produce an ensemble-based decision-making model that can be integrated into both unsupervised anomaly scoring, and clustering-based intrusion detection approaches to find a global and efficient anomaly threshold.

The work that came close to our proposed approach is in (Berezinski, 2015). The authors proposed an entropy-based network anomaly detector, abbreviated as ANODE, to detect anomalies from network traffic capture. Although the proposed work was not ICS-specific, it provides a very good overall analysis on using entropy for detecting anomalous behavior and is the main motivation behind our approach.

Our paper presents an initial entropy analysis on selected features using datasets obtained from an HVAC system. As ICS sensors and actuators are resource-constrained, and the OT system itself needs more real-time monitoring and response, a computation-intensive intrusion and attack detection method may not be suitable. We acknowledge, however, that the initial entropy analysis provides only a starting point in asking several questions and investigating relevant issues that will lead to a most optimal system design and implementation. We start from the fundamental question: whether a preliminary entropy analysis on Modbus-over-TCP data using only a few TCP/IP features without going into the Modbus traffic itself gives us information about an anomaly in the network. In the following sections, we discuss entropy and relative entropy using Kullback-Leibler divergence and results from our initial analysis of the data.

3 Initial Entropy Analysis For Modbus Over TCP/IP

We based our analysis on a day's worth of Modbus over TCP/IP data collected from an HVAC system at the University of Alabama Huntsville research lab. The data has both malicious and non-malicious traffic representing various stages of the HVAC operations (auto cool, auto heat, normal, and random). Of the various attacks that were simulated, we chose the top three that have been discussed in the literature, namely, Denial of Service (DOS), Man-in-the-middle (MITM), and Reconnaissance.

One of the important considerations in any anomaly detection technique is feature selection. We based our preliminary analysis on only the TCP/IP data with a simple objective of answering the fundamental question: whether a preliminary entropy analysis on Modbus-over-TCP data using only a few TCP/IP features gives us any information about an anomaly in the network. To deduce what features would be relevant in our initial entropy analysis, a correlation matrix was run against the TCP/IP data. The correlation matrix is shown in table 1.

	Packet Size	Inter-Packet Delay	Packet Process Time	Protocol Overhead	Protocol Efficiency	Throughput
Packet Size	1.00	-0.03	-0.44	-0.95	0.95	0.19
IP Delay	-0.03	1.00	0.07	0.02	-0.02	-0.01
PP Time	-0.44	0.07	1.00	0.23	-0.23	-0.04
Protocol Overhead	-0.95	0.02	0.23	1.00	-1.00	-0.20
Protocol Efficiency	0.95	-0.02	-0.23	-1.00	1.00	0.20
Throughput	0.19	-0.01	-0.04	-0.20	0.20	1.00

Table 1: Correlation matrix between features

Three features - packet size, inter-packet delay, and packet process time – were chosen as they had the most significant differences in their relationships with one another. Shannon's entropy, which is a direct measure of the bits needed to store the data in a given variable, was used for entropy computation for each feature for both normal and malicious traffic (DDOS, MITM, and Recon). Table 2 summarizes the results. All three features - inter-packet delay, packet process time, and packet size - have higher entropy values under Denial of Service (DOS) and Man-In-The-Middle (MITM) attacks. Reconnaissance did not have much impact on entropy for packet size and packet process time but has a small increase for the inter-packet delay. This can be explained by the nature of reconnaissance, where probes are sent with a varying time lag. From table 2, it can be seen that entropy can be a potential indicator to alert the system of some anomaly, although it will require further investigation to detect the actual nature of the anomaly as such.

	Packet Size	Inter-Packet Delay	Packet Process Time
Normal	1.987 bits	2.332 bits	2.957 bits
Normal + MITM	2.173 bits	2.488 bits	3.006 bits
Normal + Recon	1.987 bits	2.393 bits	2.957 bits
Normal + DOS	2.066 bits	2.891 bits	3.008 bits

Table 2: Entropy values for three selected features against three attack types

One persistent question in anomaly detection is how much deviation or change is acceptable. To investigate that, we used relative entropy, or Kullback-Leibler divergence, which is a measure of the deviation of one probability distribution from another and is reflective of a realistic threshold that needs to be set to indicate an anomaly in network traffic. We used Kullback-Leibler divergence to measure deviation of malicious traffic distribution from non-malicious traffic distribution. Before analysis, however, it is important to note that these entropy levels are impacted by the volume of the attack; 45.3%, or nearly half, of the data, was of the attack type MITM, or Man in the Middle; Reconnaissance attacks made up 1.2% of the data; and DOS, or Denial-Of-Service attacks, made up 2.1% of the data. Table 3 summarizes the results from the Kullback-Leibler divergence computation. Divergence was computed for each of the malicious traffic categories from normal traffic. Man-In-The-Middle (MITM) attacks had the largest divergence from normal traffic, especially for packet size. This is not unusual given the nature of MITM attacks and the goals they want to accomplish. Reconnaissance traffic had the lowest divergence, which can be because of the very low volume of reconnaissance traffic within the sample (1.2%). DOS attacks had a similar effect on the system as MITM, however divergence for inter-packet delay was much less compared to the other two features. This could be a result of how DOS attacks flood a system and increase packet process times significantly.

	MITM	Recon	DOS
Packe Size KL Divergence	3.176 bits	0 bits	0.952 bits
Inter-Packet Delay KL Divergence	0.949 bits	0.157 bits	0.139 bits
Packet Process Time KL Divergence	0.731 bits	0 bits	0.663 bits

Table 3: Relative entropy using Kullback-Leibler Divergence

It can be inferred from the data that attacks like MITM and DOS can be detected initially by looking at the entropy values of selected features. In contrast, reconnaissance may be undetectable by initial entropy analysis. However, a more intricate multivariate analysis using joint entropy may be effective in detecting reconnaissance in the network. It is important to note that the percentage of each attack in the total flow of recorded traffic plays a significant role in the change of entropy; the more infected traffic within the data, the more likely it is to notice changes compared to normal traffic regardless of what attack is being studied.

4 Conclusion and Future Direction

We presented an initial entropy analysis on an industrial control system network using selected features with datasets obtained from an HVAC system. We acknowledge that the initial entropy analysis only provides a starting point in asking several questions and investigating relevant issues that will lead to optimal system design and implementation. We started from the fundamental question: whether a preliminary entropy analysis on Modbus-over-TCP data using only a few TCP/IP features without going into the Modbus traffic itself gives us information about an anomaly in the network. We based our analysis on a day's worth of Modbus over TCP/IP data collected from an HVAC system at the University of Alabama Huntsville research lab. The data has both malicious and non-malicious traffic representing various stages of the HVAC operations (auto cool, auto heat, normal, and random). Of the various attacks that were simulated, we chose the top three that have been discussed in the literature, namely, Denial of Service (DOS), Man-in-the-middle (MITM), and Reconnaissance. We also used Kullback-Leibler divergence to measure the relative entropy of selected features over normal traffic for each of the malicious activities. Initial analysis shows some promising results regarding entropy and divergence. However, there are several questions that need to be addressed as below.

1. Do this entropy and the divergence values give us a realistic threshold for anomaly detection? These need to be analyzed with more days' worth of data.
2. Do the percentage of attacks traffic have a causal effect on the entropy values? These would also need to be analyzed with more data and more attack traffic.
3. What would be the entropy values for the ModBus traffic features? Do those features give us a more holistic view of an anomaly in the network traffic?
4. Would joint entropies give us a better understanding of network anomaly? The multivariate analysis would be needed to answer this question.

References

Anthi, E., Williams, L., Burnap, P., & Jones, K. (2021). A three-tiered intrusion detection system for industrial control systems. *Journal of Cybersecurity*, *7*(1). doi:10.1093/cybsec/tyab006

Almalawi, A., Fahad, A., Tari, Z., Khan, A. I., Alzahrani, N., Bakhsh, S. T., . . . Qaiyum, S. (2020). Add-On Anomaly Threshold Technique for Improving Unsupervised Intrusion Detection on SCADA Data. *Electronics, 9*(6), 1017. doi:10.3390/electronics9061017

Bereziński, P., Jasiul, B., & Szpyrka, M. (2015). An Entropy-Based Network Anomaly Detection Method. *Entropy*, *17*(4), 2367-2408. doi:10.3390/e17042367

Goh, J., Adepu, S., Tan, M., & Lee, Z. S. (2017). Anomaly Detection in Cyber Physical Systems Using Recurrent Neural Networks. *2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE)*. doi:10.1109/hase.2017.36

Yang, D., Usynin, A., & Hines, J. (2006). Anomaly-Based Intrusion Detection for SCADA Systems.In proc. of the 5. International Topical Meeting on Nuclear Plant Instrumentation Controls, and Human Machine Interface Technology, 2006.

Eigner, O., Kreimel, P., & Tavolato, P. (2018). Attacks on Industrial Control Systems - Modeling and Anomaly Detection. *Proceedings of the 4th International Conference on Information Systems Security and Privacy*. doi:10.5220/0006755405810588

Caltagirone, S. (2018). Industrial Control Threat Intelligence. *Dragos*.

A., Cardenas, Amin, S., Sinopoli, B., Giani, A., Perrig, A., & Sastry, S. (n.d.). Challenges for Securing Cyber Physical Systems. *University of California-Berkeley, Carnegie Mellon University*.

Valdes, A., & Cheung, S. (2009). Communication pattern anomaly detection in process control systems. *2009 IEEE Conference on Technologies for Homeland Security*. doi:10.1109/ths.2009.5168010

Koucham, O. (2018). Intrusion detection for industrial control systems. *Universite Grenoble Aples*. Doi:tel-02108208

Angseus, J., & Ekbom, R. (2017). Network-based Intrusion Detection Systems for Industrial Control Systems. *Chalmers University of Technology*.

Morris, T., & Gao, W. (n.d.). Industrial Control System Traffic Data Sets for Intrusion Detection Research. doi:10.1007/978-3-662-45355-1_5

Drias, Z., Serhrouchni, A., & Vogel, O. (2015). Taxonomy of attacks on Industrial control protocols. *NTDS*. doi:10.1109/NOTERE.2015.7293513

Lin, Q., Verwer, S., Kooji, R., & Mathur, A. (2019). Using Datasets from Industrial Control Systems for Cyber Security Research and Education. *CRITIS 2019*.