Applied Parallel Coordinates for ICS Network Traffic in Simulated Nuclear Power Plant Environments

Chaechang Lee* and Soomin Lim

Korea Institute of Nuclear Nonproliferation and Control, 1534 Yuseong-daero, Yuseong-gu, Daejeon, Republic of Korea

*chiching@kinac.re.kr, s2min@kinac.re.kr

1. Introduction

A visualization is an important approach to helping understand what data means and discover data values. Parallel coordinates are a useful method to visualize high-dimensional geometry and analyzing multivariate data. For those inherent characteristics, they were already studied well to apply in IT environment in order to detect unknown large-scale Internet attacks including Internet worms, DDoS attacks, and network scanning activities [1]. This article presents differences of the application of parallel coordinates between IT environment and Instrument and Control System (ICS) network, especially nuclear facilities, effectivenesses of this method, and practical consideration to apply them to a real site.

2. Visualization Description

2.1 Materials

A main material used in this study is a Distributed Control System(DCS), OPERASYSTEM-1400, developed by Woori technology Inc. The DCS localized digital I&C system in Korea is used as a platform for non-safety control systems of Advanced Power Reactor 1400 (APR-1400) in nuclear power plants [2].

In this study, it receives digital input signals and sends digital output signals from and to a simulator which plays a role as controllers and actuators in nuclear facilities. It is also connected to other two systems via a local-area network (LAN): One is a Human-Machine Interface (HMI) PC for controlling and monitoring DCS, the other is workbench PC for building and uploading a program logic to DCS. Fig. 1 shows the LAN structure used in this study.

For an example of the data capture of the network traffics we used Wireshark with a promiscuous mode at the HMI PC. R, a free software environment for statistical computing and graphics, was used to visualize to parallel coordinate plots from the captured traffics with its associated libraries such as MASS.

Fig. 1. LAN structure used in this study.

2.2 Method

To draw parallel coordinates in this study we used 5 parallel lines to represent captured traffics source and destination IP address, destination port, a length of a packet, and its protocol. A set of the above five features extracted from the captured traffic is represented as a polyline with vertices on the parallel axes. Two sets which have same features from the traffic data was not drawn as one polyline but two lines distinguishable from each other. It helps to find out its scales and patterns with frequencies of the packets. In addition, we assign a different color to each source IP address to find out the relationship with two axes even if they are not adjacent.

3. Result and Discussion

Fig. 2 shows the parallel coordinates with the method mentioned above.

3.1 Effectiveness of this Method

The effectiveness of this method is to make easy to analyze the network traffic patterns even if it uses proprietary protocols. A lot of systems in the industry sites, including the DCS we used in this study, uses proprietary protocols. Since they are unknown to publics, it is difficult to understand what values in a payload mean though all packets are deeply captured. Most proprietary protocols, however, use Ethernet network based on TCP/IP like IT protocols. By plotting parallel coordinates with 3rd layer's information, the traffic patterns of them can be

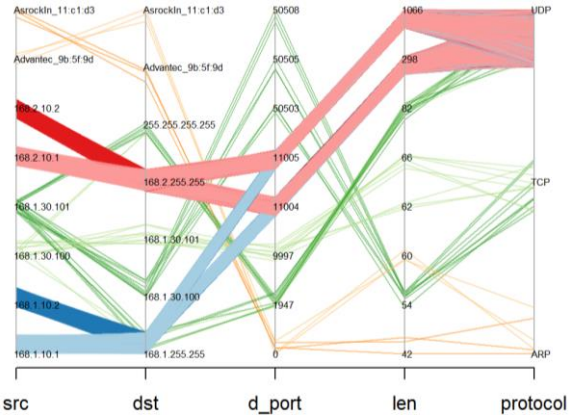

Fig. 2. The applied parallel coordinates on simulated ICS network traffic.

revealed as for where they are going from and to, and how much their amounts are.

Assigning different colors to each source IP addresses helps to find out the relationship between two traffic features even not adjacency axes. In a case of plotting with one color, the parallel coordinate should be analyzed with whole shapes, since it shows the relationship only between adjacent axes.

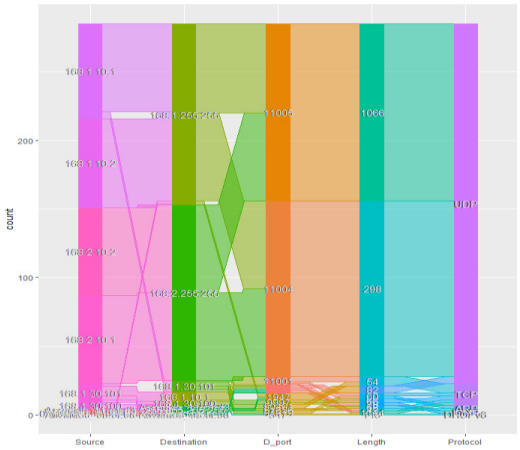

Fig. 3. An example of ambiguous parallel coordinates.

Parallel coordinates with this method reveal the main stream of the ICS traffic in its network. A thick line represents a large amount of the traffic with the same pattern. A thin line means that the traffic pattern is an uncommon occurrence. With the results, the specific traffic patterns can be used as a whitelist for outlier detections in the ICS network that its network traffic patterns are somewhat static.

To use this method, network traffics occurred in an ICS network should be visualized to identify its scale and one packet should be represented as a polyline. Besides, any two packets which have same traffic features should not be overlapped into a polyline when plotted. If scales in a parallel coordinate are visually ambiguous, it is not useful to get any information from the plot. Fig. 3 shows an example of ambiguous parallel coordinates.

3.2 Differences IT and ICS

In a case of a network with IT systems, especially connected to the internet network, there are too many hosts that represent source and destination IP addresses to assign colors that are distinguishable. In other words, parallel coordinates plotted on IT networks do not tell us what a destination port is, what a protocol service is, and what a length is a packet from one source IP address. They only show us the whole shapes in order to detect an outline of network compromises with a broad view.

On the other hand, parallel coordinates plotted on ICS networks have relatively small number of hosts. They can reveal which features are related to a source IP address with a colored line. It also helps to find out an overview of the whole network situation without detail specifications of proprietary protocols.

4. Conclusion

In this study, we visualize network traffics in a DCS network, simulated nuclear power plant environments with the parallel coordinate method. The network traffics have proprietary protocols and we can get some intuitive information with this method. As a next work, we plan to collect characteristics on parallel coordinates when there is any kind of abnormal behavior in the network and develop visualized signatures to detect the network compromise.

REFERENCES

[1] Choi, Hyunsang, Heejo Lee, and Hyogon Kim. “Fast detection and visualization of network attacks on parallel coordinates.” computers & security 28.5 (2009): 276-288.

[2] Kwon, Kee-Choon, and Myeong-Soo Lee. “Technical review on the localized digital instrumentation and control systems.” Nuclear engineering and technology 41.4 (2009): 447-454.