A First Look at Machine-to-Machine Power Grid Network Traffic

Sang Shin Jung†, David Formby†, Carson Day‡, and Raheem Beyah†
†Communications Assurance and Performance (CAP) Group
‡National Electric Energy Testing, Research and Application Center (NEETRAC)
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332
Email: {sangsin, djformby}@gatech.edu, carson.day@neetrac.gatech.edu, raheem.beyah@ece.gatech.edu

***Abstract*—The purpose of network traffic characterization is to explore unknown patterns in different types of network communications to help improve many aspects of the network. While many previous studies have explored the characterization of many different networks (e.g., university networks), the power grid network (and other SCADA networks) characterization has not yet been studied. In this paper, we provide a characterization of the power grid network to answer questions like the following: i) how stable is the communication based on configurations?; ii) are there different observable traffic patterns in different vendor equipment?; iii) are there trends in the network traffic?; iv) can information be gathered from the traffic characterization to help secure the power grid network? To address these questions, we have collected power grid network traffic in a live substation for two months and conducted an empirical study to identify network traffic behaviors in the live substation. Our empirical study shows different behaviors between the devices and vendors when they communicate with each other.**

I. Introduction

The power grid network has been around for a long while and helps utility providers manage power distribution during normal and abnormal (e.g., line outage) operations. The power grid SCADA network consists of different types of machines: human machine interfaces (HMIs), historians, front-end processors (FEPs) (SCADA masters in general), remote terminal units (RTUs), and intelligent electronic devices (IEDs) or relays. Each type of machine has a different role in the power grid network, but they communicate with the same protocol or a compatible protocol. The majority of the communication in the power grid network is between machines. For instance, IEDs report voltage levels or current levels to an RTU by its request at various polling intervals, or an unsolicited report is generated as a result of a special event. RTUs also report all the aggregated data from IEDs to a FEP. Figure 1 shows the communication among a FEP, an RTU, and IEDs in a simplified power grid network.

Previous studies on the power grid have relied on simulations or testbeds for different purposes (e.g., developing security applications). For instance, Liu et al. illustrate false data injection attacks that would forge state estimation values from the power grid devices [1]. The authors use test systems: IEEE 9-bus, 14-bus, 30-bus, 118-bus, and 300-bus test systems [2], but these systems do not precisely mimic real behaviors in the power grid network. Hines et al. use the IEEE 300-bus test system [2] to measure a centrality in power grid networks [3]. These previous studies have just assumed that power grid networks would have a similar behavior to traditional networks. However, according to our empirical observations, behaviors different from those in traditional networks exist in power grid networks. The characterization of real traffic from power grid networks has not yet received much attention. To the best of our knowledge, our empirical study is the first characterization work on power grid networks.

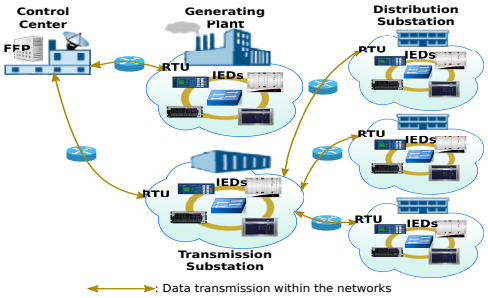


Fig. 1: A simplified power grid network.

Our study investigates the behaviors of power grid networks. In particular, we address the different patterns in frame sizes and transport layer protocol behaviors. For instance, the behaviors characterized are how many connections each device has, what major protocols the devices use, and what types of connections are prevalent in each device type (passive connections or active connections, e.g., requests). In addition, we investigate the transmission control protocol (TCP) in depth to find the different device behaviors from those in traditional networks, these behaviors include: the number of connections by different initiators, port numbers, initial sequence numbers, and initial window sizes frequently used by specific types of devices. Our study aims to answer fundamental questions which have not been addressed by looking at real traffic in a power grid network. Our main contributions are the following:

- We collect network traffic from a live substation in a power grid network.
- We conduct an empirical study to find different behaviors in the power grid networks.
- We address fundamental questions regarding the network traffic behaviors from our observations.
- We provide insight into how this characterization can be leveraged for security applications. We further point out a

flaw in the TCP/IP protocol stack of the devices on the substation network.

The rest of this paper is organized as follows. We review related works in Section II. In Section III, we describe the data set and its device categorization that we use in this study. In Section IV, we present the details of our network characterization with the analysis of frame sizes and TCP features. In Section V, we briefly discuss how security applications in the power grid networks can leverage our empirical observations. We conclude our work and discuss future work in Section VI.

II. Related Work

Network traffic characterization serves a variety of useful purposes, including providing statistics to help networks run more efficiently, creating more accurate software models, device fingerprinting based on variations in protocol implementations, and designing more precise and effective anomaly-based intrusion detection algorithms. Unfortunately, up to this point there has been little to no research published on the characterization of power grid networks and SCADA system traffic in general.

However, since the rise of the Internet in the mid 1990s, there has been extensive research done on measuring different aspects of this massive global infrastructure which provides a precedent for how to produce a useful traffic characterization. One of the first serious attempts at getting a big picture view of the behavior of Internet traffic was accomplished by Vern Paxson in 1999 when he published a study on the end-to-end behavior of bulk TCP transfers across dozens of Internet sites. His research measured such characteristics as packet loss, out of order deliveries, bottleneck bandwidth, and packet replication, while offering keen insight into the causes of any abnormal behavior seen [4]. Another study was published a few years later in 2003 that proposed a new tool for Internet characterization and measured similar characteristics over a major Internet backbone. This work used GPS synchronization to produce a more detailed timing analysis and also found that the traffic content flowing through the Internet had shifted to being a majority of file sharing and media streaming as opposed to simple web sites [5].

Looking into Internet behavior from the perspective of the average consumer, a study was published in 2007 that measured and compared bandwidth, round trip time, packet loss, and queue policies for a large sample set of DSL and cable subscribers in North America and Europe. With these measurements the authors were able to point out how Internet performance is affected by the ISP and differs from what was suggested by previous studies [6]. Shortly afterward, another study was conducted that looked into further detail about the Internet performance for DSL customers from a major European ISP [7]. Shifting from a focus on the Internet in the average household, Bensen et al. published a study in 2010 that provided a unique insight into the traffic behavior of large data centers. Their work highlighted the differences between typical Internet traffic and that seen in data centers, and suggested how this kind of information could be used to manage such data centers more efficiently [8].

In the field of power grid communications, there has not yet been any characterization research at the network level or above, but studies have been conducted to simulate the effects of power substation noise on commonly used link layer protocols such as WiFi and Zigbee [9]. While not specifically in the area of power grid, the most closely related work to the research presented here comes from an in-depth characterization of cellular machine-to-machine traffic published in 2012 by Shafiq et al., which focused on comparing the behavior of machine-to-machine communication with smart phone traffic over cellular networks by taking measurements on round trip time, packet loss, spectral usage, and temporal patterns [10].

This paper attempts to begin to fill the void in characterization research of control system networks by providing a first look at the behavior of live substation traffic.

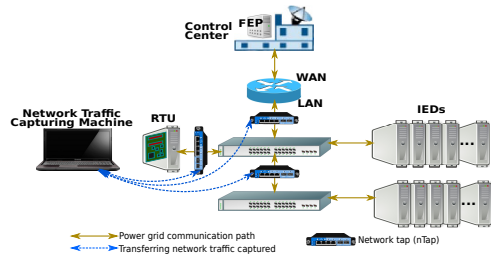


Fig. 2: Network capture environment in a live substation.

III. Data Set

In this section, we present the data set used in our characterization and describe an overview of the data set from the live substation.

We installed our network traffic monitoring system in the live substation as shown in Figure 2. The network environment in which we capture network traffic consists of mainly an RTU and IEDs connected to network switches and a router. In the live substation, the RTU has frequent communication with IEDs and the FEP at the control center. IEDs have periodic communication with the RTU and sometimes have communication with the FEP at the control center as well, but not frequently.

We first summarize the data set that was captured over 2 months (about 16GB) in Table I. In total, our data set consisted of 158 devices including the RTU, IEDs and some devices (e.g., the FEP) from the WAN through the router. We categorize the devices in vendor types by media access control (MAC) addresses. Seven different vendors (VDs), which are well-known in the power grid industry, provide most of the devices: VDs 1-3 and VDs 5-6 for IEDs, VD4 for the router, and VD7 for the RTU. Each vendor has a different number of devices running the TCP/IP stack and application protocols for distributed network protocol (DNP) version 3, HTTP, and SMTP, except with some unidentified protocols used by VD3 and VD5. VD4 represents devices communicating from a control center through the router.

IV. Power Grid Communication Characteristics

From looking at just a few high-level attributes, the data set shows significant variations in network traffic between different vendor types. In this section, we further examine the variations in frame sizes of all traffic and TCP connections (which is the major transport protocol used in the live substation).

TABLE I: The summary of power grid network traffic for two months (VD refers to vendor types).

Vendor types	# of devices	Device types	Incoming packets	Outgoing packets	Transport layer	Applications
VD1	1	IED	106995 (0.14%)	681072 (0.9%)	TCP/UDP	Proprietary (Port: 5000)
VD2	6	IED	4214765 (5.8%)	4209307 (5.7%)	TCP	DNP (Port: 20000)
VD3	3	IED	12 (0.000016%)	12 (0.000016%)	Unknown	Unknown
VD4	13	Router	2262567 (3.1%)	2982122 (4.0%)	TCP/UDP	SMTP (Port: 25)
VD5	1	IED	845 (0.001%)	831 (0.001%)	TCP/UDP	Unknown
VD6	133	IED	31338039 (42.9%)	28798859 (38.7%)	TCP	DNP (Port: 20000)
VD7	1	RTU	34979400 (47.9%)	37626528 (50.6%)	TCP	DNP, HTTP (20000, 80)
Total	158	-	72902623	74298731	-	-

A. Variations in Frame Sizes

We measured the frame sizes of all traffic from each vendor over two months and found the minimum (min), maximum (max), average (μ), and standard deviation (σ) for each day of the week and for each hour of the day. In Figure 3, we first present the frame sizes by day of the week. As shown in the figure, VDs 1-2, VD4, and VDs 6-7 have consistent min, μ, and σ frame sizes each day of the week. The max sizes for VD1, VD4, and VD6 are distinct from day to day whereas VD2 and VD7 have consistent max sizes each day. VD3 has the same min, max, and μ of the frame sizes on Thu. and Fri., which implies that traffic (about 70 bytes frames) as shown in the figure occurs every Thu. and Fri. regularly. VD5 has one spike (about 350 bytes more) on Mon. that is 250 bytes more than the rest of the days (about 100 bytes).

In addition to the day-of-week patterns, we present hour-of-day patterns in Figure 4. The hour-of-day patterns for VD1, VD4, and VD6 are similar to those in the day-of-week representation: consistent sizes of min, μ, and σ and irregular max sizes. VD2 and VD7 still show consistent patterns of min, max, μ, and σ for each hour of the day. The 70 bytes of the frame sizes on Thu. and Fri. from VD3, as shown in Figure 3, are shown in exact hours between 14:00 and 15:00 of each day. The spike with extra 350 bytes on Mon. from VD5 as shown in Figure 3 is also shown in exact hours between 9:00 and 10:00 of the day.

B. Variations in TCP Connections

In addition to studying the variations of frame sizes from different vendors, we also examined connection initiators (e.g., those who always request a connection to other devices) and found that only a few of the devices initiated the connections in the live substation. In Figure 5, we present the three vendor types that initiate all of the connections by day-of-week and hour-of-day. Only VDs 4-5 and VD7 (a total of 15 devices as shown in Table I) are the TCP connection initiators to the other types (143 devices). The connections are mostly established by VD7 with a large difference between its 4000 times per day and VD4's 300 times per day. VD5 only made 33 connections between 9:00 and 10:00 on Mon. Regardless of the number, the patterns of the initiations per day and per hour are fairly constant, except a few spots between 7:00 and 14:00 shown on VD4's hour-of-day figure and a few spots shown on Thu., Fri., and Sat. on VD7's day-of-week figure.

In addition to the number of accumulated connections by the initiators, we also present the min, max, μ, and σ of the connections in Figure 6.

C. Variations in TCP Ephemeral Port Number

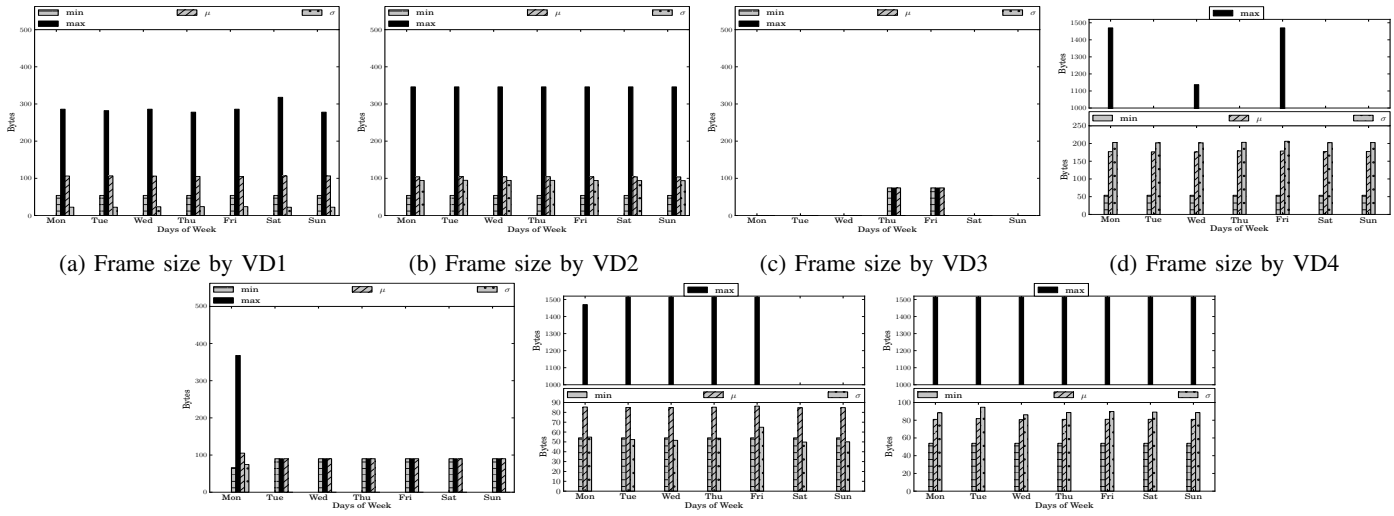
Another important variable to characterize is TCP ephemeral port numbers that each vendor randomly chooses for connection initiation. In Figure 7, we present the source port numbers of connection initiators. Even though we do not discuss the destination port numbers in detail here, it is still important to note that the destination port numbers of 25, 80, 1024, and 5000 are used by VD1, VD4, and VDs 6-7 respectively; in addition, the destination port number of 20000 is used by all types except VD4. Figure 7 shows that VD5 uses the port numbers from 1046 to 1077 evenly at about 3% each. Similar to VD5, VD7 evenly uses the port numbers at about 1% each, but VD7 has a wider range (between 1000 and 5000) than that of VD5. VD4 has two distinct clusters: one cluster has the port numbers from 1025 to 5534 at about 14% each; the other cluster has the port numbers from 49338 to 65443 with less than 2% each.

D. Variations in TCP Initial Sequence Number

We also examine TCP initial sequence numbers (ISNs) of both peers of connections (i.e., the first sequence number of TCP SYN and SYN/ACK packets). In Figure 8, TCP ISNs for all vendor types except VD3 (VDs 1-2 and VDs 4-7) are shown. The reason VD3 is not shown here is that traffic from VD3 does not use TCP. VDs 1-2 have smaller ISNs (ranging up to $2 * 10^8$ and $7 * 10^8$ respectively) than those of the other types (VDs 4-7) (ranging up to $4.5 * 10^9$).

One important finding of this research was the discovery of several significant clusters in the distribution of ISNs for VD1[1], indicating a clearly non-random generation algorithm. Not enough samples were collected from VD2 and VD5 to draw any significant conclusions, but the distributions of ISNs for VD4 and VDs 6-7 all appear to be roughly even and suggest the proper use of a random number generator.

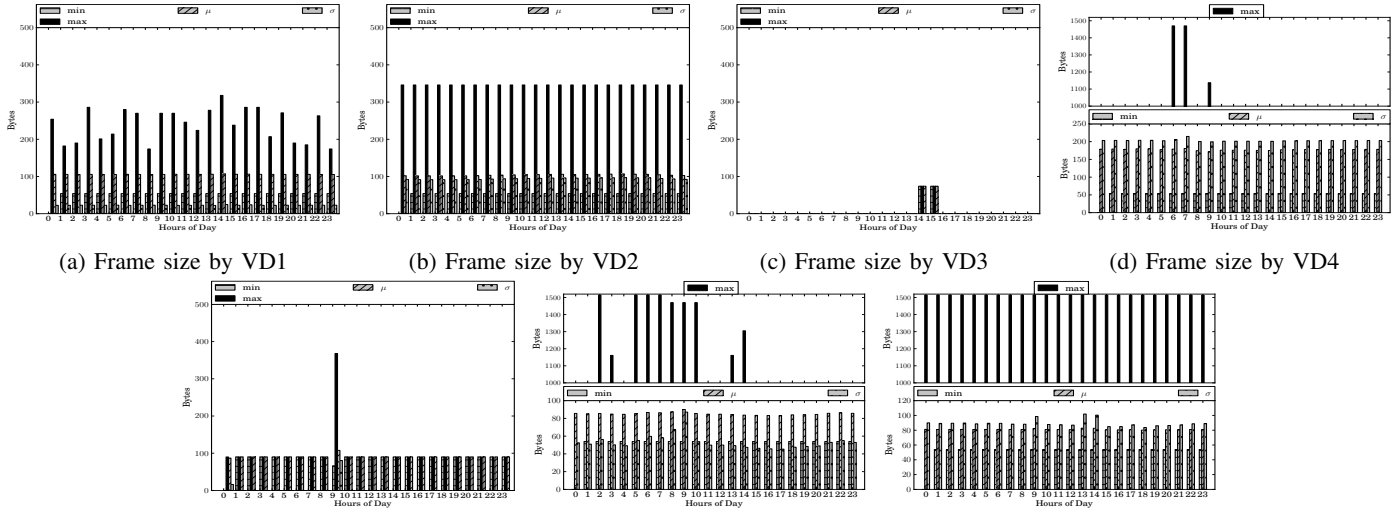
[1]We confirmed this finding as a security problem with the vendor via the provider of the data set. We are now in process of confirming an ICS-CERT Vulnerability Report.



(a) Frame size by VD1 (b) Frame size by VD2 (c) Frame size by VD3 (d) Frame size by VD4

(e) Frame size by VD5 (f) Frame size by VD6 (g) Frame size by VD7

Fig. 3: Frame sizes by VDs 1-7 in day-of-week.



(a) Frame size by VD1 (b) Frame size by VD2 (c) Frame size by VD3 (d) Frame size by VD4

(e) Frame size by VD5 (f) Frame size by VD6 (g) Frame size by VD7

Fig. 4: Frame sizes by VDs 1-7 in hour-of-day.

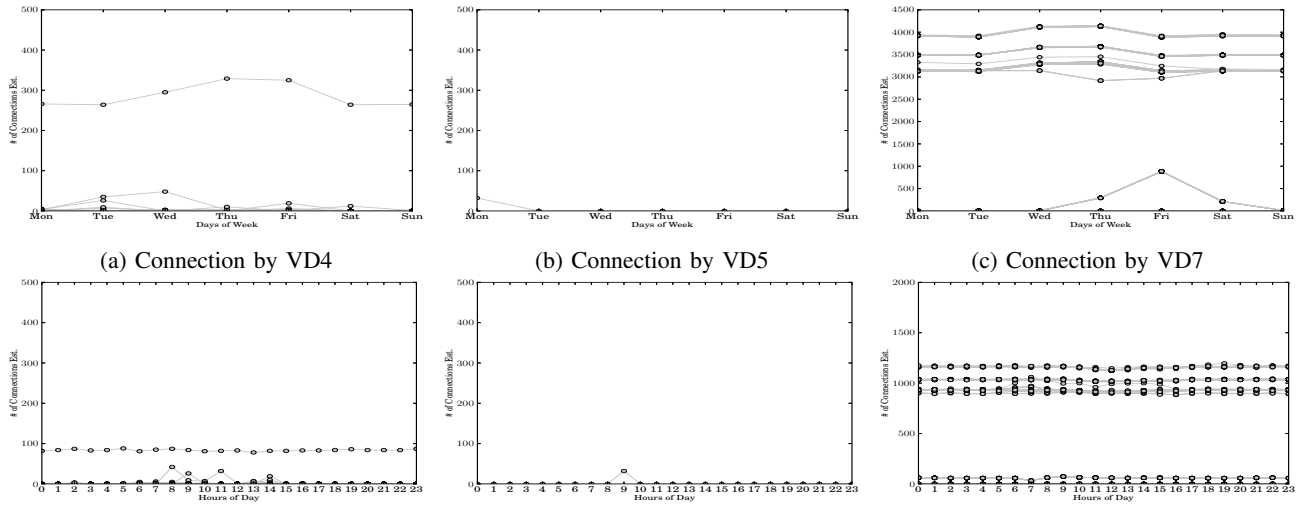
E. Variations in TCP Initial Window Size

The last variation that we examine in this paper is TCP initial window sizes (IWSs) of both peers of connections (i.e., window sizes advertised at the beginning of a TCP connection). The TCP IWS is an indication of the receiver's buffer for that TCP instance. In Figure 9, we present six vendor types, with most using different IWSs: VDs 1-2 and VD5 use only one constant IWS whereas VD4 and VDs 6-7 use a few different IWSs. It should be noted that VD1 and VD5 have the same IWSs (5.8KB). VD2 uses a very small IWS (1.4KB) as one constant value. The IWSs from VD6 are mostly 8.6KB bytes, but never appear to exceed 10KB. VD4 uses a couple of different IWSs: 8KB, 64KB, and 32KB in between, but the majority of the time, the IWS is over 64KB. VD7 has two significant groups using 4KB and 16KB of IWSs: 4KB is rarely used.

V. Discussion

In the previous section, we presented the results and observations of the variations in frame sizes, TCP ephemeral source port numbers, ISNs, and IWSs found in our empirical data. In this section, we highlight the importance of the practical benefits gained by conducting measurement-based studies to the field of security.

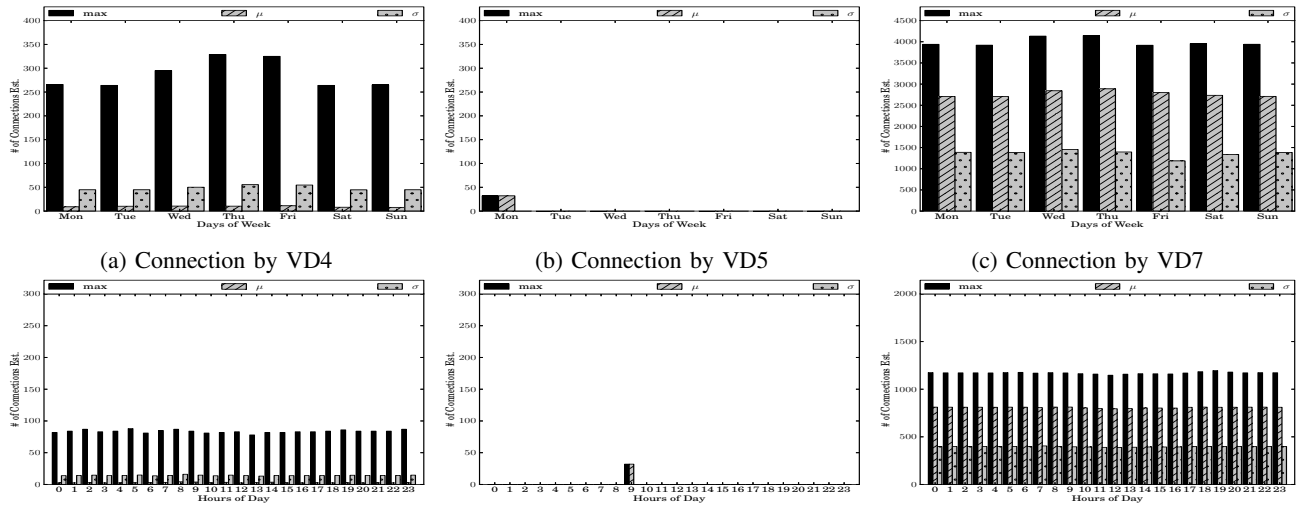
A number of security applications, particularly firewalls, network intrusion detection systems (NIDSs), and device fingerprinting, rely on differences in network traffic characteristics presented in this work. For example, network mapping tools such as nmap and p0f [11], [12], could use the variations found in TCP IWSs in Figure 9, TCP ISNs in Figure 8, and other measurements to perform fingerprinting on power control system devices to determine OS versions as well as product vendors and models.



(a) Connection by VD4 (b) Connection by VD5 (c) Connection by VD7

(d) Connection by VD4 (e) Connection by VD5 (f) Connection by VD7

Fig. 5: Total number of connections initiated by VDs 4-5 and VD7 day-of-week (top) and hour-of-day (bottom). Each line represents a pair of connections.



(a) Connection by VD4 (b) Connection by VD5 (c) Connection by VD7

(d) Connection by VD4 (e) Connection by VD5 (f) Connection by VD7

Fig. 6: Min, max, μ, and σ of connections initiated by VDs 4-5 and VD7 in day-of-week (top) and hour-of-day (bottom).

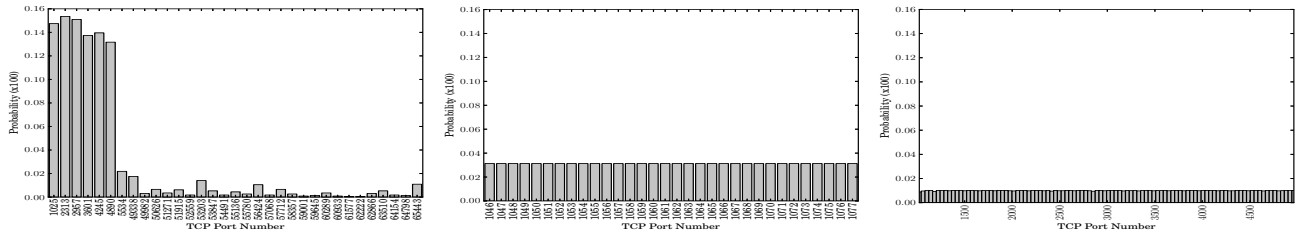
Another excellent example of the application of our results to the field of security can be seen by studying the distribution of ISNs for VD1, found in Figure 8. As noted before, there are significant clusters in the distribution suggesting a non-random generation algorithm. This unique property could be used not only to identify this device's vendor, but also perform one of the TCP sequence number attacks summarized by Bellovin in [13]. Specifically, due to the non-random nature of the generation of these ISNs, an attacker has a significant advantage in guessing the next ISNs to be used. This allows the attacker hijack a TCP connection (e.g., TCP sequence prediction attack).

VI. Conclusion and Future Work

In this paper, we performed an empirical study of network traffic that we collected from a live substation. We examined the variations in frame sizes to present the existence of diurnal patterns. We have also studied various features at the transport layer. Variations exist in the number of TCP connections and their particular initiators, TCP source port numbers used by the connection initiators, TCP ISNs, and TCP IWSs.

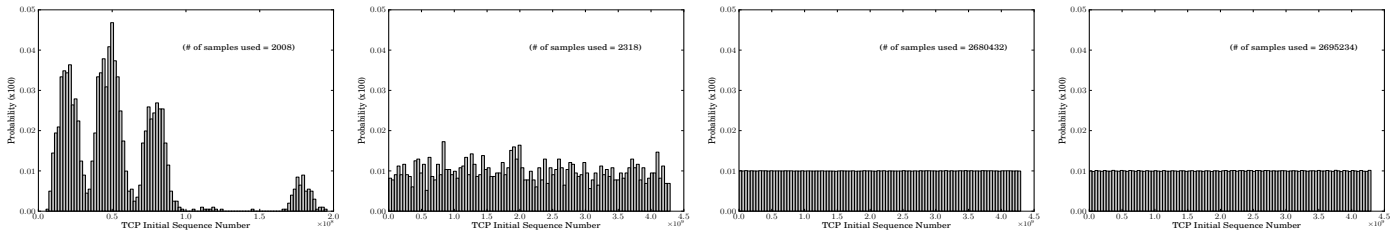
The number of TCP connections also suggests the existence of diurnal patterns. Only three vendor types with a total of 15 devices initiate the connections to the rest of 143 devices. TCP source port numbers are chosen in different ways by each vendor type: one chooses sequential numbers between 1000 and 1075; another chooses evenly between 1000 and 5000; the other chooses from two different groups of port numbers (less than 5000 or greater than 50000). TCP ISNs show variations in the ISNs ranges by different vendor types. TCP IWSs are also different from vendor to vendor: one uses a constant IWS all the time; the other has a few preferable IWSs used more than 80%.

We believe that our empirical observations can help improve



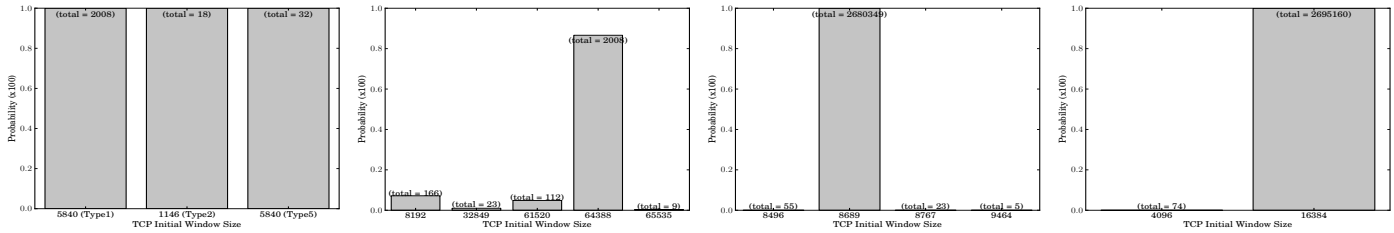
(a) Src ports by VD4 (b) Src ports by VD5 (c) Src ports by VD7

Fig. 7: Source port numbers used by VDs 4-5 and VD7.



(a) ISN by VD1 (b) ISN by VD4 (c) ISN by VD6 (d) ISN by VD7

Fig. 8: TCP ISNs used by VD1, VD4, and VDs 6-7.



(a) IWS by VD1-2 and VD5 (b) IWS by VD4 (c) IWS by VD6 (d) IWS by VD7

Fig. 9: TCP IWSs used by VDs 1-2 and VDs 4-7.

security applications (e.g., IDS). The large data set size requires a significant amount of computation time to process, so for future work we plan to finish verifying our measurements on TCP flow durations and round trip times. Additionally, we will perform a more in-depth study to find variations in traffic over a longer time period. We also plan to study application layer protocols (e.g., DNP) characteristics in detail.

ACKNOWLEDGMENTS

This work was partly supported by DARPA-N10AP20022 and NEETRAC.

REFERENCES

[1] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proceedings of the 16th ACM conference on Computer and communications security*, CCS '09, (New York, NY, USA), pp. 21–32, ACM, 2009.

[2] "Power Systems Test Case Archive." http://www.ee.washington.edu/research/pstca/.

[3] P. Hines and S. Blumsack, "A Centrality Measure for Electrical Networks," in *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, p. 185, jan. 2008.

[4] V. Paxson, "End-to-end Internet packet dynamics," *Networking, IEEE/ACM Transactions on*, vol. 7, pp. 277–292, Jun 1999.

[5] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, and S. Diot, "Packet-level traffic measurements from the Sprint IP backbone," *Network, IEEE*, vol. 17, no. 6, pp. 6–16, 2003.

[6] M. Dischinger, A. Haeberlen, K. P. Gummadi, and S. Saroiu, "Characterizing residential broadband networks," in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, IMC '07, (New York, NY, USA), pp. 43–56, ACM, 2007.

[7] G. Maier, A. Feldmann, V. Paxson, and M. Allman, "On dominant characteristics of residential broadband internet traffic," in *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, IMC '09, (New York, NY, USA), pp. 90–102, ACM, 2009.

[8] T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, IMC '10, (New York, NY, USA), pp. 267–280, ACM, 2010.

[9] Q. Shan, I. Glover, P. Moore, I. Portugues, R. Watson, and R. Rutherford, "Performance of Zigbee in Electricity Supply Substations," in *Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on*, pp. 3871–3874, Sept 2007.

[10] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, and J. Wang, "A First Look at Cellular Machine-to-machine Traffic: Large Scale Measurement and Characterization," in *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '12, (New York, NY, USA), pp. 65–76, ACM, 2012.

[11] "nmap security scanner." http://nmap.org/.

[12] "pOf v3." http://lcamtuf.coredump.cx/p0f3/.

[13] S. M. Bellovin, "A Look Back at Security Problems in the TCP/IP Protocol Suite," in *Proceedings of the 20th Annual Computer Security Applications Conference*, ACSAC '04, (Washington, DC, USA), pp. 229–249, IEEE Computer Society, 2004.