A Case Study in Power Substation Network Dynamics

DAVID FORMBY, Georgia Institute of Technology
ANWAR WALID, Bell Laboratories
RAHEEM BEYAH, Georgia Institute of Technology

The modern world is becoming increasingly dependent on computing and communication technology to function, but unfortunately its application and impact on areas such as critical infrastructure and industrial control system (ICS) networks remains to be thoroughly studied. Significant research has been conducted to address the myriad security concerns in these areas, but they are virtually all based on artificial testbeds or simulations designed on assumptions about their behavior either from knowledge of traditional IT networking or from basic principles of ICS operation. In this work, we provide the most detailed characterization of an example ICS to date in order to determine if these common assumptions hold true. A live power distribution substation is observed over the course of two and a half years to measure its behavior and evolution over time. Then, a horizontal study is conducted that compared this behavior with three other substations from the same company. Although most predictions were found to be correct, some unexpected behavior was observed that highlights the fundamental differences between ICS and IT networks including round trip times dominated by processing speed as opposed to network delay, several well known TCP features being largely irrelevant, and surprisingly large jitter from devices running real-time operating systems. The impact of these observations is discussed in terms of generality to other embedded networks, network security applications, and the suitability of the TCP protocol for this environment.

ACM Reference format:
David Formby, Anwar Walid, and Raheem Beyah. 2017. A Case Study in Power Substation Network Dynamics. *Proc. ACM Meas. Anal. Comput. Syst.* 1, 1, Article 18 (June 2017), 24 pages.
DOI: http://dx.doi.org/10.1145/3084456

1 INTRODUCTION

In recent years, the confluence of increasingly complex control systems and advancing information technology (IT) has resulted in a new class of networks, cyber-physical systems. These networks in the form of industrial control systems (ICS) and supervisory control and data acquisition (SCADA) systems, are used for a variety of applications ranging from delicate manufacturing processes to critical infrastructures, such as power or water distribution where safe and reliable operation is paramount. The protocols running on these networks were often originally developed for dedicated serial communication lines and later adapted for TCP/IP networks. Although these systems now communicate over networking protocols familiar to the IT world, the true effect of the merging of these two technologies has yet to be thoroughly studied, and as a result, most research in the area of cyber-physical systems is based on artificial testbeds and theoretical assumptions that are not always accurate.

Given the nature of what these systems are actually controlling, detailed knowledge of how they truly behave is necessary for improved operation performance and developing defenses against a new class of cyber attack that could result in physical harm to both equipment and personnel. Intrusion detection algorithms that perform well in lab experiments and are based on assumptions about the target network do not always translate well

DOI: http://dx.doi.org/10.1145/3084456

to real-world ICS networks where the nodes are comprised of a heterogeneous mix of devices of varying ages and the network architecture and configuration are not as clean and straightforward. Without published data on the size and intensity of bandwidth on these networks, security researchers designing these new intrusion detection systems cannot know what kinds of throughput their algorithms must handle. Additionally, detailed understanding of real-world ICS networks can lead to more accurate computer simulations and more efficiently and robustly designed networks in the future.

To date there have been very few works that attempt to provide a detailed characterization of ICS networks, primarily due to the culture of the industry being so resistant to change and the issue of security only recently gaining popularity in the area. For this work, we were given rare access to several live power substation networks and monitored the traffic over an extended period of time to answer important questions about the characteristics of real-world ICS networks. Specifically, the primary goal of this characterization was to examine the accuracy of common assumptions that are made about ICS networks and highlight any unusual behavior including observable security issues. A secondary goal was then to determine, based on the observed traffic, whether TCP is well-matched to ICS networks and whether any modifications could be made to improve the resiliency of remote field devices and efficiency of their low powered hardware. The major contributions of this paper can be summarized as follows:

- The most in-depth traffic characterization of real-world live power substation networks to date
- A study of the stability of this characterization over 2.5 year's time and minor changes in architecture
- A comparison of behavior across multiple substations
- The discovery of several abnormal behaviors in the observed SCADA protocol stacks, configurations, and real-time software
- Suggestions for modifying the TCP protocol to fit the needs of the ICS environment

The remainder of this paper is organized as follows. Related characterization work is discussed in Section 2 followed by background information and an explanation of the experimental setup used in this research in Sections 3 and 4. Section 5 describes in detail the behavior of one substation network over time and Section 6 compares the same behavior across multiple substations. The implications of our findings are discussed in Section 7 and finally our conclusions and future work are summarized in Section 8.

2 RELATED WORK

Most areas of networking research depend on a foundation of knowing the target networks' behavior and traffic patterns, therefore it is crucial that detailed characterizations are regularly conducted. These characterizations can provide insight on how to run a more efficient and reliable network, help create more accurate simulation models, enable OS or browser fingerprinting based on variations in protocol implementations [1] [21], and aid in the design of more precise and effective anomaly based intrusion detection algorithms. Unfortunately, until recently there has been little research published on the characterization of power system networks and industrial control system traffic in general.

However, there has been significant research into various aspects of Internet traffic that provides a solid foundation for this paper. One of the first works to study Internet traffic at a high level was achieved by Vern Paxson in 1999 when he published a study on the end-to-end behavior of bulk TCP transfers across nodes on the Internet. His results included observations on packet loss, out of order deliveries, bottleneck bandwidth, and packet replication, while offering keen insight into the causes of any abnormal behavior seen [15]. The next major milestone in traffic characterization was in 2003 with the proposal of using GPS synchronization to produce more detailed timing analysis. The study 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 [8].

One of the primary purposes of characterization of TCP traffic has always been to study possible techniques for congestion control and more efficient use of bandwidth. As the applications used by the average consumer constantly demand more and more bandwidth, this is still as important as ever. However, this problem does not apply to ICS networks where bandwidth requirements are at fixed low levels and nodes consist almost exclusively of embedded devices. Even though we are becoming more and more reliant on intelligent electronic devices that control physical entities, there has been very little studies on how these specific types of networks perform. One of the earlier related works studied the physical and link layers by simulating the effects of power substation noise on commonly used wireless protocols such as WiFi and Zigbee [20]. Several years later, an in-depth characterization of cellular machine-to-machine traffic was published by Shafiq et al., which contained a small amount of power metering traffic and focused on comparing the behavior of machine-to-machine communication with smart phone traffic over cellular networks by taking measurements of round trip time, packet loss, and temporal patterns [19].

In the specific area of industrial control systems, there has only been limited research. In 2012 a study was published that compared certain traffic characteristics of a water distribution facility with an example IT network dataset and found that the water facility did not exhibit strong diurnal patterns like the IT dataset and that the flow sizes did not quite fit the typical log-normal or Pareto distributions found in other Internet characterizations [3]. Most closely related to this work was a traffic characterization of a power substation network in 2014 that provided an important first look into what these networks look like but only at a very high level [13]. Shortly afterward, a different study focused on characterizing widespread TCP vulnerabilities found in power grid devices [6]. This work differentiates itself from previous work by providing a more detailed characterization over a longer capture period, comparing behavior across multiple substations, and providing unique insight into how these observations can be used to improve TCP for ICS networks.

3 BACKGROUND

In order to better understand the differences between ICS networks and IT networks, some background information on SCADA and power grid networks is required. First, we describe the general operation of a SCADA system in a power distribution network to provide context about where the network traffic capture was taken. Figure 1 illustrates the typical hierarchy where the control center (CC) communicates primarily with the remote terminal unit (RTU) in each substation, and the RTU acts as a middle man between the control center and intelligent electronic devices (IEDs) by collecting data from the IEDs in the field, summarizing the data and reporting it back to the CC when requested. When control operations are required, the CC sends the command to the RTU which forwards it on to the correct IED. The most common SCADA protocols used to perform this kind of communication include Modbus, GOOSE/IEC 61850, and DNP3, all of which collect data in slightly different ways. Modbus, the oldest of the three, relies solely on regular polling of measurement data, while DNP3 can operate in polling mode or spontaneous reporting mode. Finally, GOOSE operates on a publisher-subscriber philosophy by multicasting event data throughout the network.

The primary protocol used in this research was DNP3. The DNP3 protocol was originally developed for serial communication lines so it has its own complex protocol stack illustrated in Figure 2, that either sits directly on a serial line or is encapsulated by the TCP/IP protocol suite. It is important to note that no matter the physical media on which the protocol is deployed, the DNP3 stack still views communication as it would appear on a serial line, as a stream of data. Therefore, the DNP3 data link layer performs frame delimiting, error checking, and optional acknowledgments for reliable transmission. The transport pseudo-layer primarily handles fragmentation of application layer fragments that are too large for the maximum link layer size. Finally, the application layer handles larger sized fragmentation, optional acknowledgments, and a wide variety of flexible functions for SCADA system operation including data collection and control. DNP3 data collection can be achieved through field devices spontaneously reporting important events, or as in the case of this research, the DNP3 master implementing a polling schedule for all devices [2].

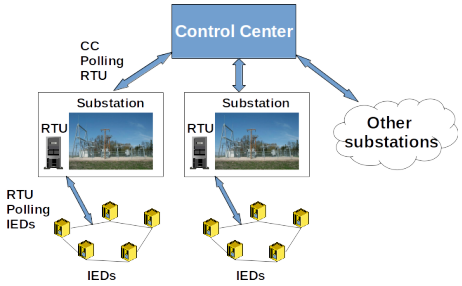


Fig. 1. Typical SCADA hierarchy in a power distribution substation

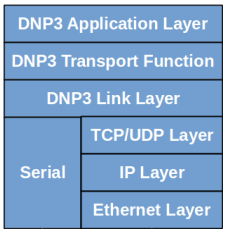


Fig. 2. DNP3 Protocol Stack

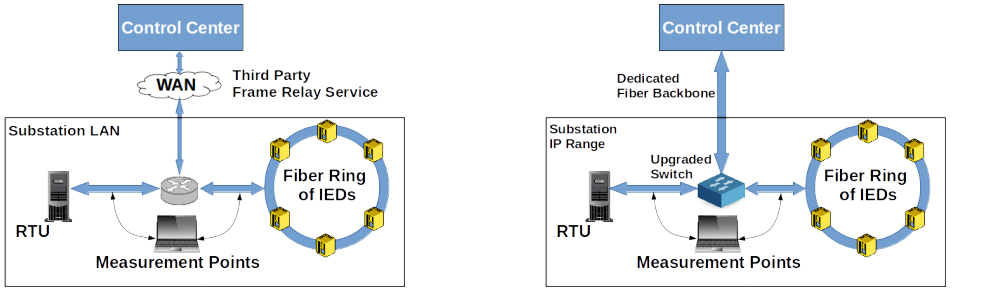
4 EXPERIMENTAL SETUP

The datasets used in this research were captured at medium-voltage distribution substations over the span of two and a half years, with roughly a year gap. During that gap, small changes were made to the configuration and architecture of the network which allows us to study how those changes affected the traffic measurements.

Figure 3a illustrates the architecture from which the first dataset was obtained. Under this architecture, each substation has its own separate LAN, communication between the CC and RTU was over a third-party frame relay service, and DNP3 polling intervals were much slower than in the second dataset. The traffic was captured in the substation close to the RTU so all communication between the CC, RTU, and IEDs in the field could be monitored. The capture setup in Figure 3b is different in that all substations were moved to the same LAN, a new switch was installed that allowed port mirroring, communication between the CC and RTU used a dedicated fiber backbone, and the DNP3 polling intervals were much faster than the first dataset. Table 1 summarizes the basic statistics about the datasets, where datasets A1, A2, and A3 are taken from the same substation and the others are taken from separate substations on the same network. The change from using a third party frame relay service to having a private fiber backbone would theoretically increase the speed and reliability of the communication between the CC and each substation. Furthermore, while data on the old router's processing and switching delay is unavailable, it is relatively safe to assume that the brand new, higher end switch is able to process packets at a higher speed reducing the switching delay for packets on the network. Finally, the change from having each substation on its own LAN to logically combining them on the same LAN should have little effect on the traffic observed at the specified measurement points, except for potentially seeing some traffic from nearby substations.

5 STABILITY OVER TIME

The first half of this characterization studies the evolution of a substation's behavior over a long period of time using datasets A1 and A2 to determine how stable the measurements were.



(a) A1 Network Architecture

(b) Network Architecture for All Other Datasets

Fig. 3. Experimental Setup

Table 1. Description of Datasets

	Start	End	Size [GB]	Nodes
A1	Sept. 2013	Feb. 2014	21.7	204
A2	Jan. 2015	Aug. 2015	146.4	317
A3	Aug. 2015	April 2016	147	517
B	10 Aug. '15	11 Aug. '15	0.34	118
C	April 2016	May 2016	5.7	129
D	April 2016	May 2016	11.7	167

Table 2. Bandwidth statistics averaged over ten-second samples

	A1	A2
μ [kbps]	11.2	58.8
σ [kbps]	7.7	860.2
Max [kbps]	1925	278931
Min [kbps]	1.11	52.1

5.1 High Level Behavior

To first study the network traffic at the highest level of abstraction, patterns in metadata such as timing and bandwidth usage were examined.

Traffic Volume. As explained in Section 1, the two primary functions of any ICS or SCADA system are data acquisition and control. Consequently, the networks are assumed to have very clear communication relationships where the bulk of the traffic is generated from field devices regularly reporting data to the master and the master occasionally sending commands as needed. Although it varies by the exact SCADA protocol used, this is especially true with the DNP3 protocol where there are strict master/slave relationships and the IEDs never communicate with each other. To test whether the bandwith is low and stable, as would be expected from polling traffic, bandwidth samples were taken every ten seconds and summarized in Table 2.

The results confirmed that indeed, the network uses very little bandwidth on average at 11kbps and 58kbps for A1 and A2 respectively of the 100Mbps links and with little variation. The bandwidth usage increased between datasets due to a configuration change that increased the polling frequency of devices for measurements. Another

Table 3. Average Packet Sizes

	A1	A2
μ [bytes]	89.9	96.5
σ [bytes]	91.3	107.6

interesting observation is the maximum bandwidth usage in both datasets appear to be caused by temporary switching loops and subsequent broadcast storms.

As a result of the constant polling for measurements, and the device often having no signficant events to report, the average packet sizes remain rather small, as described in Table 3.

Regularity of Traffic. Popular ICS protocols including DNP3 and Modbus rely on regular polling of measurement data to ensure that the system is operating as intended. DNP3 in particular uses a combination of faster event polls to monitor for exceptions, such as voltages crossing a certain threshold, and slower static polls to monitor the exact status of all measurements over time. These polls are initiated by devices running real-time operating systems and are often configured with the real-time control constraints in mind, i.e. certain measurements have to be updated so often in order maximize efficiency or respond to emergencies in time. Therefore, one might assume that the polling intervals would be very regular with low jitter, but this was not the case. Figure 4 illustrates the event and static polling intervals for all devices in datasets A1 and A2. Note that all polls were initiated by the RTUs running VxWorks, a popular real-time operating system that provides hard real-time guarantees. Minimum and maximum polling intervals varied drastically due to broadcast storms and devices going silent, and even during normal operation polling intervals had standard deviations on the order of several seconds despite the real-time nature of the RTU. To verify that this behavior was not caused by network issues, polling intervals were also calculated by the TCP timestamp ticks in the poll requests coming from the RTU, which were generated by the real-time operating system. Indeed, Figure 5 illustrates how the poll requests being generated by the software running on the real-time operating system in the RTU exhibits similar irregularities.

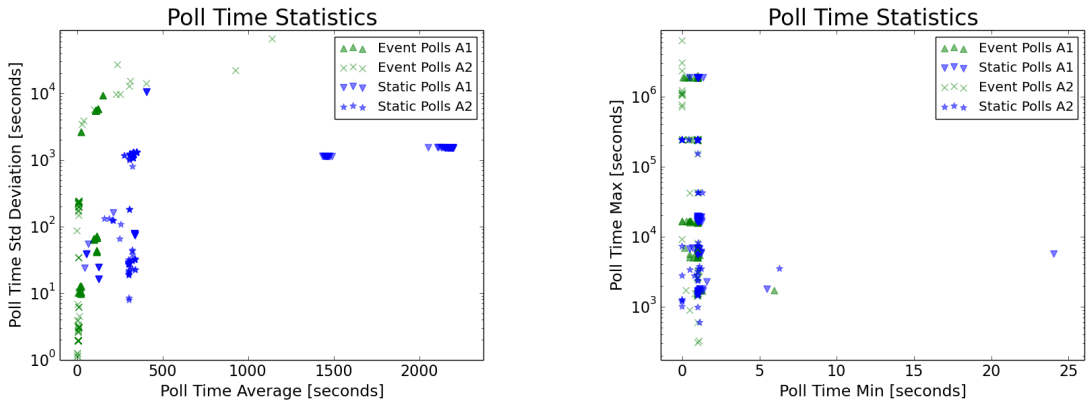


Fig. 4. Polling Intervals

To study another aspect of the regularity of the traffic, the inter-arrival times of packets at the capture point were measured and are shown in Figure 6. The distributions from both datasets appear to be approximately the same, and the most interesting feature is small periodic spikes every 17ms that dampen exponentially. While the

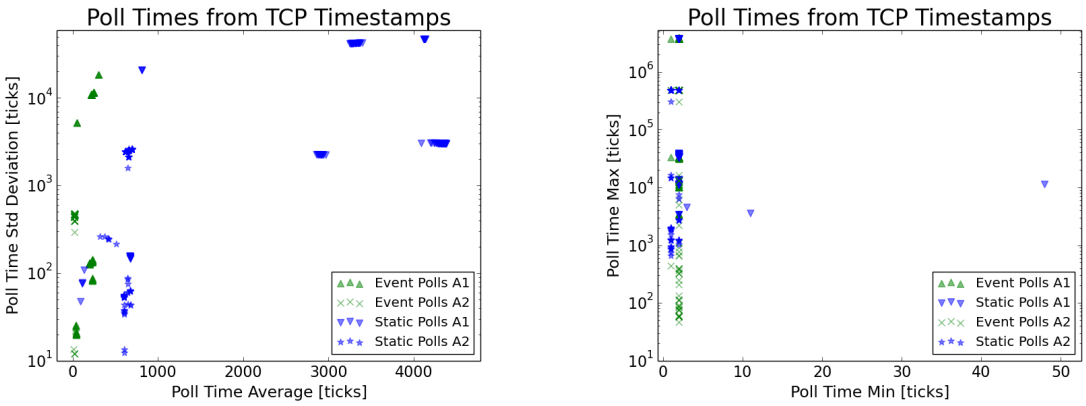


Fig. 5. Polling Intervals from TCP Timestamps

cause for these spikes is still unclear, we speculate that it is most likely originating in the RTU since it initiates all communication with the IEDs. One possible explanation is the use of a very coarse grained timer to decide when to send the next polling request, but other explanations are just as likely to be the true cause. It is also interesting to note that the traffic intensity of the network is so low that inter-arrival times occasionally reach as high as one second.

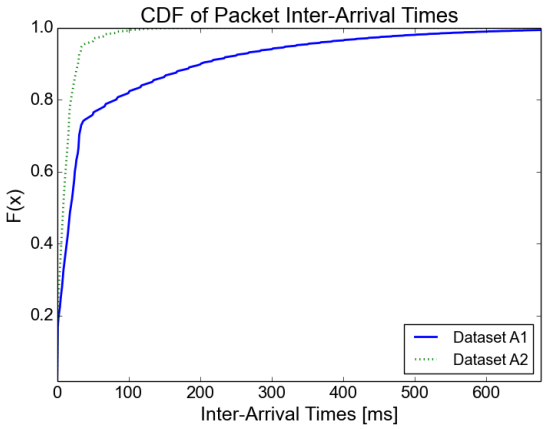


Fig. 6. A1 and A2 packet inter-arrival times at the measurement point

Availability. One interesting aspect that differentiates ICS networks from traditional IT networks is the consequences of node down-time. If a set of nodes is unavailable for a period of time in an IT network, either due to network issues or device issues, revenue and productivity will be lost but no physical harm will be done. However in the ICS environment, every minute that a node is down increases the risk that a critical issue that requires attention will go unnoticed. In the case of the power grid, this could be as innocuous as lost revenue from wasteful energy production when demand has fallen, or as devastating as loss of power to parts of the grid due to increased consumption and lack of available power.

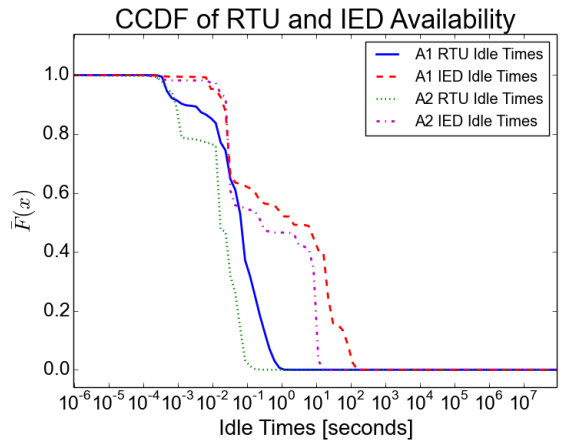


Fig. 7. Idle Times of Devices on the Network

Figure 7 illustrates the idle times for individual devices with the RTU separated from the IEDs. For the first dataset, the distributions show that the RTU never appears to lose availability for long, but surprisingly some IEDs appear to go down for up to 10 days at a time. It is unclear from manual inspection of the network traffic if these devices were intentionally taken out of commission for maintenance or technical problems brought them down. Devices in the second dataset demonstrate similar behavior.

5.2 TCP Level Behavior

In addition to unexpected patterns in high level metadata, unusual behavior was also observed in the common network protocols that have been adapted from the IT world into use for the ICS environment, and it was found that some of the most crucial improvements to TCP that allow modern web traffic today are largely irrelevant in the context of ICS networks.

TCP Flow Duration and Size Although previous work has studied TCP flow sizes for Internet traffic [4] and for water distribution networks [3], data on power grid TCP flow sizes has yet to be published. Studies of Internet traffic suggested that flow durations have long tailed distributions and follow log-normal and Pareto distributions. The characterization of the water facility found similar results for some types of flows, but others exhibited too much variance. The general conclusion that the research was able to draw was that all types of flows observed at the water facility were positively skewed, meaning most values fell within a main body in the distribution but there were a significant number of extreme large values.

Water distribution facilities are another example of industrial control system networks not unlike the power substation networks studied here, so similar results were expected to be found. From a purely theoretical perspective, most flows were expected to be very long lived similar to an extreme persistent HTTP connection, due to the connections staying open for the constant polling for measurements. Few shorter flows were expected to represent the occasional configuration or manual maintenance. However, results obtained from the substation network exhibited high variance due to a variety of configuration issues at the application layer and network layer, illustrated in Figures 8 and 9.

The most striking feature of the A1 distribution is the overwhelmingly large body centered around a 100 second duration. After closer inspection, the cause of this was determined to be a misconfiguration of all devices on one of the circuits in the network. The issue stemmed from the fact that the RTU was configured to poll for

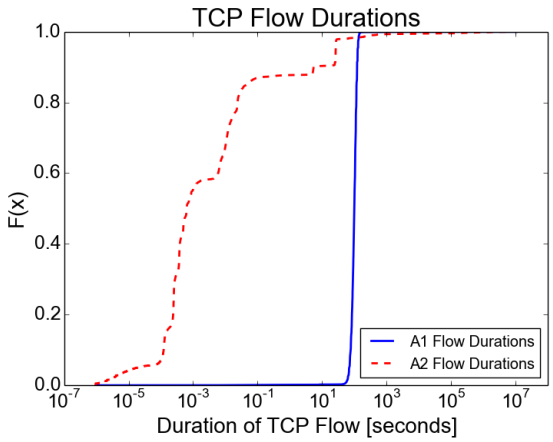


Fig. 8. A1 and A2 TCP Flow Durations

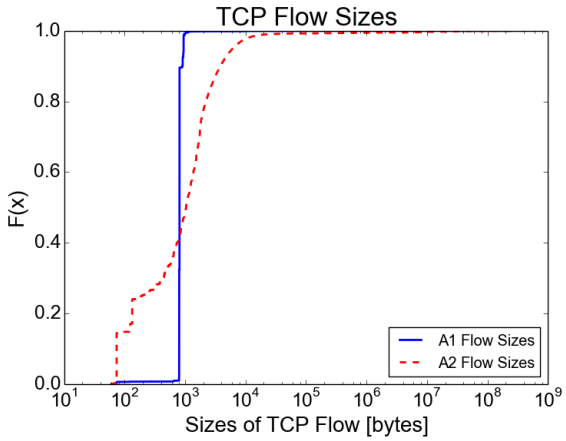
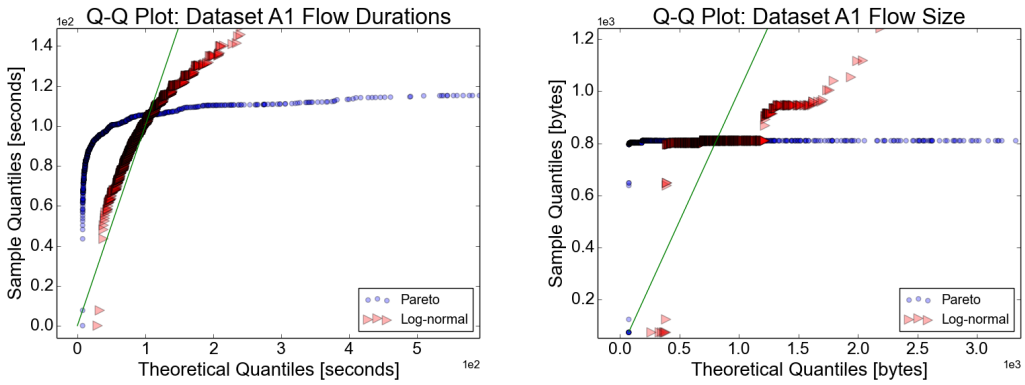


Fig. 9. A1 and A2 TCP Flow Sizes

DNP3 event data every 100 seconds for each device in this circuit, but either the IEDs themselves or the switches in the network were configured to close the connection with a FIN handshake after 30 seconds of being idle. After acknowledging the FIN packet from the IED, the RTU never sends its own FIN to gracefully close the connection and instead waits another 80 seconds, attempts to send another DNP3 poll request and gets a TCP reset flag in response. This type of disagreement in the TCP connection termination was actually so rampant in the network that 99.8% of all flows were terminated with a TCP reset flag instead of the full graceful FIN handshake.

The second most noticeable spike (largely overshadowed in Figure 8 is the flow durations on the order of 10 seconds, corresponding to relatively short flows from maintenance access to various devices. Finally, the relatively large number of extremely short duration flows appeared to mostly come from a brief disruption where the switches were caught in a broadcast loop and UDP messages flooded the network causing connection



(a) A1 Flow Duration Q-Q Plots

(b) A1 Flow Size Q-Q Plots

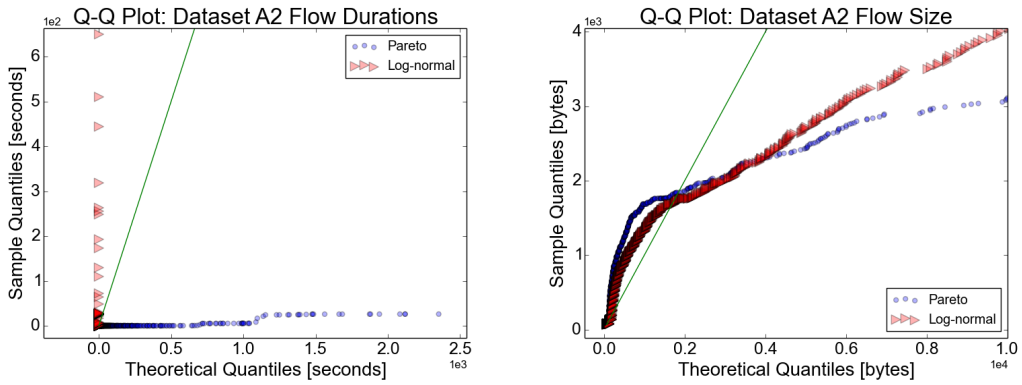
Fig. 10. A1 Q-Q Plots

issues. Other extremely short flows appear to be caused by occasional strange behavior where the RTU becomes temporarily inactive and refuses connection attempts by the control center.

When the network was revisited over a year later in dataset A2, again, large amounts of extremely short flows were a result of another broadcast storm. However, as a result of the faster DNP3 polling, the overwhelming spike at 100 seconds no longer appears because the IEDs are never idle for 30 seconds. Instead, this reveals another strange configuration issue evidenced by the large spike at around 20 seconds. After manual inspection of the traffic, the cause of this new spike appeared to be an implementation issue at the firmware layer. At several points in both datasets, the RTU closes all connections with the IEDs for a short period and refuses connection attempts from the control center. In one of these cases, when the connection is restarted with a specific IED, the IED sends the RTU a DNP3 "Request Link Status" message. The message is acknowledged by the TCP layer at the RTU, but apparently ignored at the application layer. While normal DNP3 data collection is successfully taking place, the IED retries this message several times without getting a reply and finally closes the connection after about 20 seconds apparently with the mistaken belief that the connection is broken at the application layer. The connection is continually restarted following this same pattern throughout the rest of the dataset and is never resolved. As a result of one configuration issue being corrected and another one appearing, the total percentage of TCP flows closed by reset flags in the second dataset was 87.03%.

To compare the distributions with previous work suggesting TCP flows follow log-normal and Pareto distributions, all data was fit to both log-normal and Pareto distributions using Maximum Likelihood Estimation of the parameters. The resulting theoretical distributions were compared with the empirical ones in Q-Q plots (Figures 10 and 11) to qualitatively test for similarities, noting that the closer Q-Q plots resemble the green line $y = x$ the closer their distributions are. Also note that extreme outlier quantiles were omitted to allow for detailed examination of the main body of the distributions. Clearly, the large spikes caused by the observed undesirable behavior prevent the empirical distributions from matching any of the theoretical ones.

Round Trip Times Another notable characteristic of ICS networks is that they consist of devices in fixed locations in the same geographic area as opposed to mobile devices constantly on the move or devices communicating over large distances. Therefore, based on knowledge from traditional IT networks one might think at first that the round trip times (RTTs) would be small and consistent, since they should not be affected by changing locations, different routes over the Internet, or long propagation delays.



(a) A2 Flow Duration Q-Q Plots

(b) A2 Size Duration Q-Q Plots

Fig. 11. A2 Q-Q Plots

As explained in Section 4, the measurement point for our experiments was close to the substation's RTU, therefore to test assumptions about round trip times, RTTs between the RTU and field devices were measured from the dataset by recording the time between the SYN and SYN-ACK packets in the TCP handshake and the RTT between the RTU and IP addresses associated with the control center were calculated from the time between the SYN-ACK and final ACK in the three-way handshake. RTT measurements taken during a brief broadcast storm in the network were discarded due to being extreme outliers (round trip times of greater than four seconds).

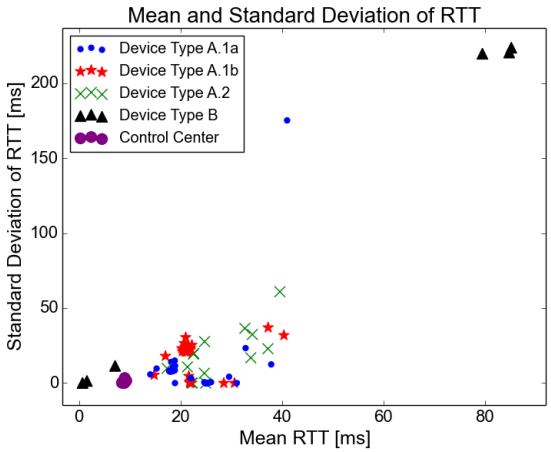


Fig. 12. A1 mean and standard deviation of round trip times for each device

As Figure 12 and Table 4 illustrate, the RTTs are neither small nor consistent. Although this may be counter-intuitive at first, it can easily be explained by the fundamental differences in the make-up of legacy ICS networks and current-day IT networks. The typical IT network consists of large numbers of relatively powerful end devices capable of processing and creating large amounts of data that must be sent over communication lines that

Table 4. RTT Statistics

Between RTU and	μ_{A1} [ms]	σ_{A1} [ms]	μ_{A2} [ms]	σ_{A2}
Type A.1a	18.7	7.4	18.9	22.2
Type A.1b	17.9	8.12	21.1	22.6
Type A.2	18.8	7.13	23.5	47.9
Type B	1.4	1.6	58.7	188
CC	3.2	1.9	9.0	2.94

are heavily constrained by bandwidth and propagation delay. Due to this imbalance of large load generation capability and small network bottleneck size, IT networks would quickly experience congestion collapse without implementing congestion control algorithms [12].

However, in ICS networks the imbalance leans in the opposite direction. The networks consist of low powered embedded devices where it is not uncommon to be running 16-bit microcontroller processors in the tens to low hundreds of MHz range and very limited RAM. The communication distances at most are no bigger than a few miles resulting in propagation delays that are fractions of a millisecond. Additionally, the link bandwidths are typically over-provisioned for availability and reliability reasons (e.g., the observed network only utilized roughly 0.02% of the 100Mbps link). This imbalance suggests that the observed RTTs are largely dominated by the processing time of the end devices rather than the propagation and queuing delay across the links.

Another interesting observation from this comparison is the difference in consistency between the two types of RTTs. Even though the ICS network's traffic intensity was very light, meaning that the network switches should never have been heavily loaded, and the packets always took the same path, the RTTs were surprisingly much more erratic compared to the RTTs over the unpredictable Internet. RTTs measured over the Internet typically have standard deviations on the order of a few milliseconds. By comparison, the RTTs measured in the ICS network were widely varying with standard deviations in the tens and even hundreds of milliseconds, and it was a regular occurrence to have RTTs as large as three seconds.

In order to study whether the cause of the irregularity in RTTs originated in the network or the field devices themselves, we compare the different RTTs in Table 4. The last row contains statistics for communication between the RTU in the substation to IP addresses associated with the control center. This communication is carried wirelessly over a third party frame relay service across a distance (about 20 miles) that is roughly ten times the communication distance between the RTU and the IEDs over fiber. The devices in the control center that communicate with the RTU are all relatively modern PCs capable of quickly processing the TCP handshake, resulting in a standard deviation of RTTs that is expectedly smaller than any of the measured Internet RTTs or other ICS device RTTs. The fact that this link, even with its longer distance and less predictable wireless communication medium, has more stable RTTs than the links between the RTU and field devices further suggests that the embedded device processing time dominates the observed network performance rather than the network infrastructure itself.

Finally, given such strong evidence that the RTT is largely dependent on the processing time of the embedded devices, it suggests that measuring the RTT could supply information about the identity of the device, similar to the cross layer response time fingerprinting methods proposed in [7]. Indeed, Figure 12 shows strong clusters for different physical device types, with Device Types A.1a and A.1b being the same hardware and Type A.2 being only a slightly different model hardware. Furthermore, these device types were not clustered by location and were quite spread out, ruling out any significant effects of propagation delay and switching delay. For example, these

devices ranged from being at the substation close to the measurement point and traversing only one network switch, to being up to a couple miles away traversing upwards of thirty network switches.

With the change in architecture, upgrading of the main switch in the substation, and correction of some of the configuration issues, the RTTs appeared to be less variable, but still exhibited similar means and clusters compared to the first dataset, as illustrated in Figure 13 and Table 4, with samples taken during two broadcast storms omitted due to extreme outliers. In this dataset, the device types seem to be even more heavily clustered based on the mean and variance of the RTTs.

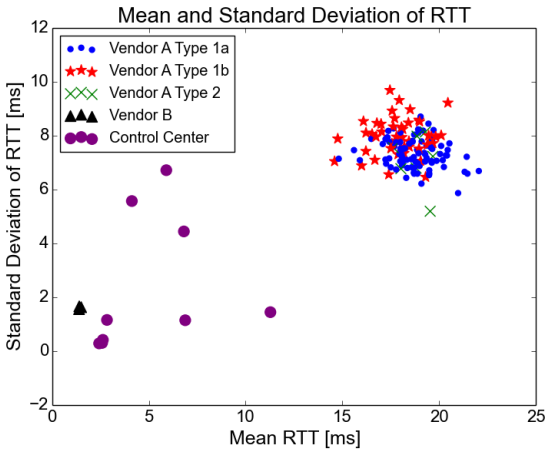


Fig. 13. A2 mean and standard deviation of round trip times for each device

TCP Retransmissions Throughout the duration of a TCP conversation, both ends keep estimates of the RTT from their perspectives in order to calculate a retransmit time-out (RTO) that reduces the chances of wasting bandwidth on retransmissions of data that would have eventually arrived. In Paxson's large scale study [15] involving two captures one year apart of bulk transfers across the Internet, he found that 1% and 2% of the packets observed in the first and second captures respectively were redundant retransmissions. Of these redundant retransmissions, 44% and 17% respectively were deemed to be unavoidable due to the loss of a segment or all acknowledgments of the segment. The majority of the remaining redundant retransmissions could have been avoided using selective acknowledgments (SACK) [14], and only a small percentage (4% and 3%) could have been avoided with a better RTO value.

Given the differences between the local power substation network and the complex Internet, the retransmission performance at the power substation network was expected to be significantly better. The combination of the low datarate (0.02% average traffic intensity), small packet size, and dedicated fiber links at the substation suggests that packet loss should be virtually non-existent and that queuing delays should be small and stable resulting in RTOs that prevent unnecessary retransmissions. However, as the results in the previous section about RTT illustrated, the low-powered embedding processing time appears to have a significant impact on the network performance including retransmissions.

As explained in Section 4, data was collected only at one location in the network so it is impossible to conclusively determine if a data segment and its ACK were actually received by the corresponding ends, as was the case in Paxson's study. However, given the low traffic intensity (0.02%) on the substation network it is highly unlikely that any packets were actually dropped due to congestion rather than just delayed by slow processing.

Table 5. Hourly Retransmission Statistics

	μ_{total}	σ_{total}	μ_{ACK}	σ_{ACK}
A1	0.510%	0.0951%	43.2%	5.77%
A2	1.46%	0.501%	26.6%	2.35%

To elaborate, given that the observed average packet size of 90 bytes has a transmission time of 7μs on a 100Mbps link and that the observed average time between packet arrival for the entire network is significantly greater at 65ms, the outgoing rate at which each switch can push packets on the line is much greater than the incoming rate of arriving packets. Therefore, the average packet should experience roughly zero queuing delay, packet loss should be virtually nonexistent, and almost every observed retransmission could have been avoided.

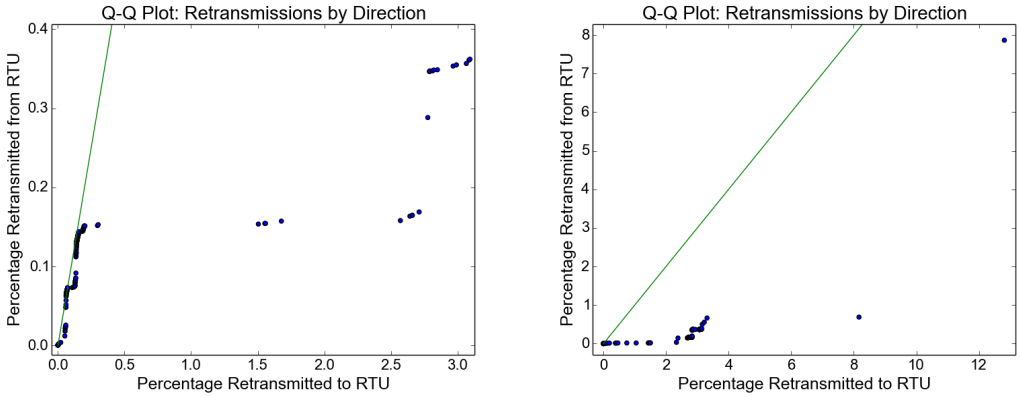
Even though the retransmission performance was expected to be significantly better on the substation network compared to Paxson's study, Table 5 shows only marginal improvements. The substation network had an average hourly retransmission percentage of 0.5% which reached as high as 2.86% during one hour, compared to Paxson's 1% and 2% over the Internet. Of these observed retransmissions 43.2% of them occurred *after* the acknowledgment for the original segment had already traversed the measurement point. Since it is unlikely that the acknowledgment was dropped anywhere in the network, the cause of these avoidable retransmissions had to be due to the sender's RTO expiring before the acknowledgment could reach it.

To examine why the overall retransmission rate was so unexpectedly high, retransmission percentages were calculated based on IP address pairs and classified by direction in the Q-Q plot in Figure 14a. Note that if the retransmission patterns for packets originating from the RTU were equal to those destined to the RTU, then the Q-Q plot would roughly follow the line, $y = x$. Interestingly, the retransmissions were not evenly distributed among device pairs due to some devices having high overall retransmission rates (as high as 15%) and others having much more reasonable rates of around 0.1%. The shape of the Q-Q plot falling so far below the line $y = x$ suggests that packets originating from the IEDs in the field were retransmitted at much higher rates than in the opposite direction. When studied further, it was found that the IP address pairs suffering from higher retransmission rates were located in the same electrical cabinets connected to the same switches, suggesting a possible malfunction or misconfiguration of some of the network switches. Other causes for poor performance could originate from the noisy electrical environment the devices are subjected to, despite being hardened against such conditions.

Another important observation to be made from Figure 14a is the imbalance between directions of the retransmissions. On closer inspection, it was found that the field devices retransmitted much more frequently to the RTU than the RTU retransmitted to the field devices, suggesting two very different measured RTTs or different calculations of the RTO.

The results from Dataset A2 in Figure 14b suggested similar conclusions about uneven retransmissions and offered other interesting observations as well with some IP pairs having a majority of their conversation be retransmissions. When examined manually, this seemed to originate from some IP addresses apparently located in other substations in the network rarely showing up in normal operation but appearing and being retransmitted a number of times during the two broadcast storms. It is also interesting to note that the average retransmission rate increased to 1.45% with the new changes, most likely due to the increased bandwidth usage and more frequent queuing delays.

While it was impossible to accurately measure the RTO for all devices with only one network tap, estimates were made based on the arrival time difference between the original transmission and the subsequent retransmissions. Note that the data in Figure 15 excludes the extreme long tails at 64 seconds in order to offer insight into the



(a) Q-Q plot of A1 retransmission per flow, comparing directions to and from the RTU

(b) Q-Q plot of A2 retransmission per flow, comparing directions to and from the RTU

Fig. 14. Imbalanced direction of retransmissions

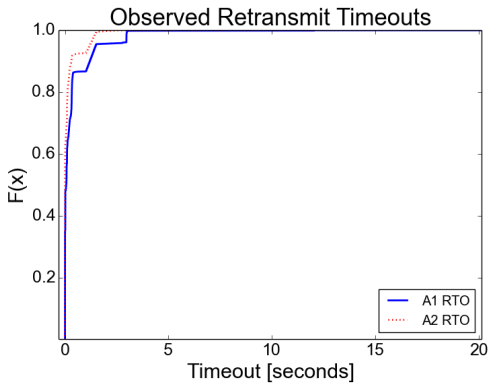


Fig. 15. A1 and A2 RTO Behavior Less than 32 Seconds

minimum RTOs implemented by the different operating systems of the devices on the network. The original specifications for TCP did not clearly define a minimum RTO, but later RFCs recommended a minimum of 1 second or 3 seconds if no RTT measurement has been made yet [16]. This was further updated in 2011 with a modified recommendation to use 1 second if no RTT measurement has been made [17]. Both RFCs state that smaller RTOs may perform better, and it has been noted that several operating systems do use smaller values and that they can increase performance [18].

The clusters around 1 second and 3 seconds appear to primarily come from the RTU, and align with the behavior specified in the RFCs. Most of the retransmissions originating from the RTU have an RTO between 1 and 1.5 seconds, and for the short lived connections discussed in Section 5.2, it appears to use 3 seconds since it never has a chance to get an RTT estimate. The other clusters around 300ms and 200ms align with the observed RTOs used by modern Windows and Linux machines respectively, but the 100ms cluster and large number of RTOs

below 10ms were unexpected. Since it is unlikely that these legacy devices are actually using faster minimum RTOs than more modern OSes, we suspect that packet timing compression is happening somewhere along the network or even in the slow processing of the field device itself.

Although it is difficult to see in the figure, the RTU is also observed to double its RTO as specified in the RFC until it reaches a set maximum at 64 seconds where it retries several more times before finally giving up.

Exact RTO behavior is vendor specific which means distributions from different networks will undoubtedly differ from those presented here. However, two important conclusions can be drawn from this data that *do* generalize to other ICS networks. The first is that due to the use of proprietary application specific OSes and slow update cycles, there is much more inherent diversity in the TCP implementations seen in ICS networks than a typical corporate network consisting primarily of modern Windows and Mac systems. The second conclusion reinforces what was discussed in the previous section in that the slow processing time of the embedding field devices dominates that network performance rather than the network infrastructure.

The results from Dataset A2 in Figure 15 produced similar results.

6 COMPARISON ACROSS SUBSTATIONS

After conducting the previous longitudinal study at a single substation, behavior across multiple substations was examined to determine how general the previous observations were. A third dataset from the first substation was compared with data from three other substations on the power system network.

6.1 High Level Behavior

Comparing differences in behavior at the highest level of abstraction, such as bandwidth usage and packet size, across substations provides insight into how active and connected in the power grid each substation is.

Traffic Volume When similar bandwidth measurements are taken as in Section 5.1, the new results, in Table 6, lead to similar observations. Again, the average bandwidth usage is low in the tens of kilobits per second, and another broadcast storm caused large spikes of traffic in A3.

Table 6. Bandwidth statistics averaged over ten-second samples

	A3	B	C	D
μ [kbps]	48.2	34.1	15.9	33.1
σ [kbps]	232.7	1.44	1.17	1.01
Max [kbps]	224503	46.6	45.2	103.4
Min [kbps]	0	29.2	0	25.5

The average packet sizes in Table 7 were again small due to the polling nature of the traffic, with the only significant differences between substations being the standard deviation of the sizes. This could be explained by the lack of broadcast storms in the shorter datasets B, C, and D, or possibly a sign of how often devices in the field have important events to report back to the master.

Table 7. Packet Sizes

	A3	B	C	D
μ [kbps]	82.2	78.1	78.5	80.0
σ [kbps]	52.9	34.9	37.1	29.2

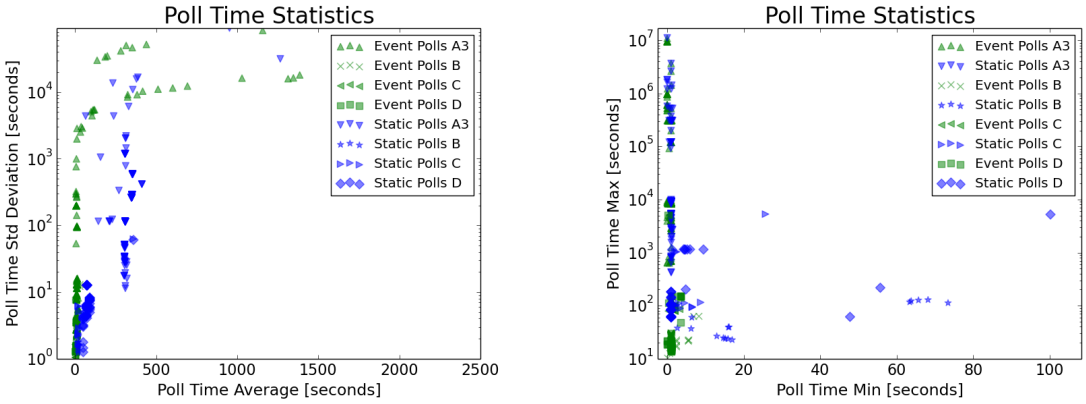


Fig. 16. Polling Intervals

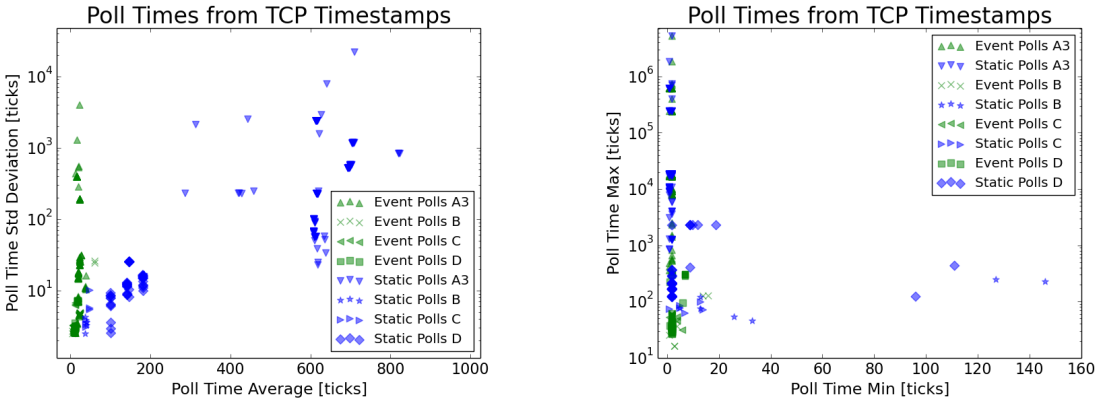


Fig. 17. Polling Intervals from TCP Timestamps

Regularity of Traffic Polling interval jitter was again observed across all substations from both network tap timestamps and timestamps from the RTUs at each substation, illustrated in Figures 16 and 17.

When the inter-arrival times of packets at each of the substations were plotted in Figure 18, again the unexplained spikes at 17ms intervals are present in each one. It can also be noted that the busier substations with more devices on their networks have more short inter-arrival times simply due to the presence of more traffic.

Availability Examining the idle times of the RTUs and IEDs in Figure 19 reveals similar behavior with IEDs going quiet for surprisingly long periods of time. For example, from Dataset A3 a few devices were not heard from for several days while device idle times from Datasets C and D maxed out around 10 minutes of idle time.

6.2 TCP Level Behavior

Measurements made at the TCP level across substations revealed more evidence of configuration issues and confirmed observations made about the round trip times being largely due to processing delay.

TCP Flow Duration and Size Given that the ideal TCP flow for a control system environment would be as long lived as possible to reduce latency, evidence of similar configuration issues can be seen in Figures 20

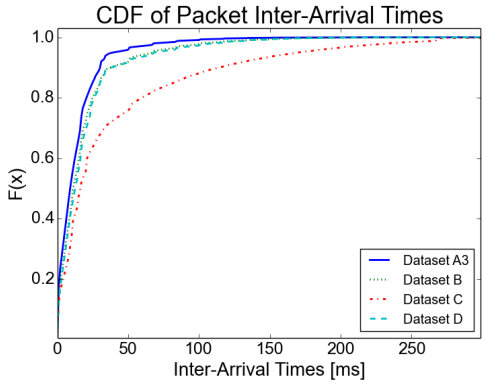


Fig. 18. CDF of Packet Inter-Arrival Times

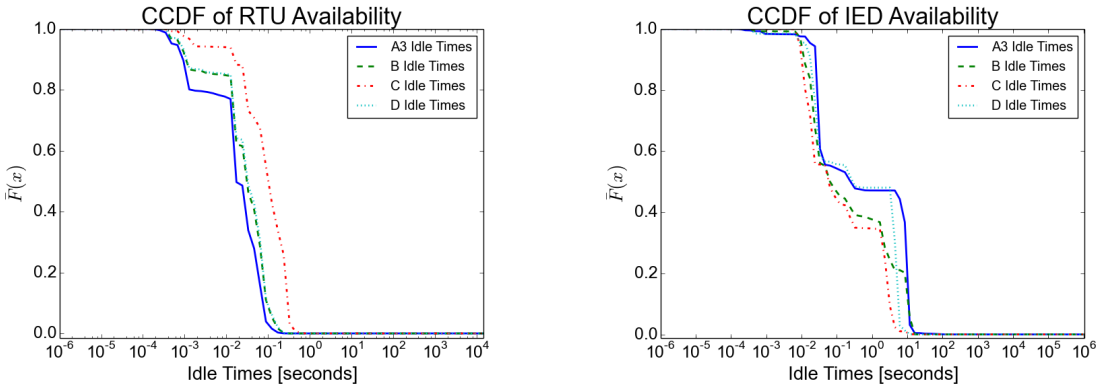


Fig. 19. CCDFs of RTU and IED Idle Times

and 21. For example, the same spike at 20 seconds is still present in Dataset A3, and the other substations in Datasets C and D still are plagued with numerous short duration flows. Plots for Dataset B have been omitted because of the short duration preventing any useful observations. Again all distributions were fit to Pareto and log-normal distributions using MLE and illustrated on Q-Q plots to qualitatively test for distribution fitness. Note that extreme outlier quantiles were omitted again to allow for closer examination of the main body, and again the anomalous behavior prevents the empirical distributions from fitting any of the theoretical ones.

Round Trip Times Due to the shorter capture times, no TCP SYN handshakes were observed in Dataset B, and only certain device types were observed in Datasets C and D. The remaining results in Table 8 and Table 9 do however support the previous observation that RTTs are surprisingly large and variable. Furthermore, the fact that the measurements are so similar with device types even across substation networks further suggests that they are primarily dependent on device processing time rather than network architecture.

TCP Retransmissions Similar high retransmission rates were seen across all substations with the notable exception where the smallest substation, Dataset C, had only 0.298% of all packets retransmitted. Although

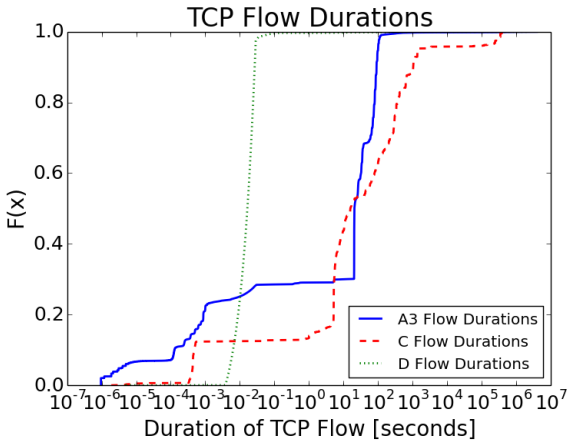


Fig. 20. TCP Flow Durations

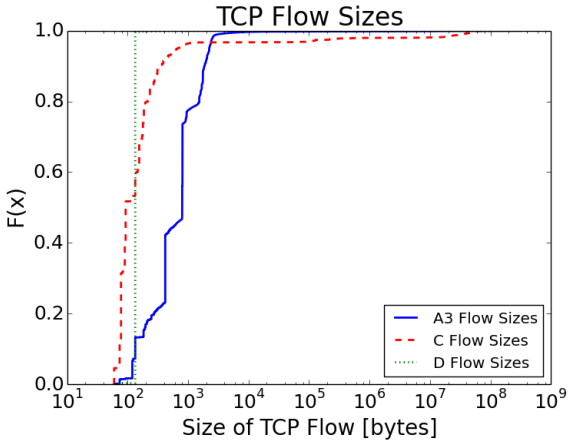


Fig. 21. TCP Flow Sizes

omitted here due to space constraints, retransmissions were again uneven between the master and field devices due to the apparent widely different calculations of the RTO.

Since the same device types are used in all of the substations, they all use the same RTO calculation algorithms resulting in similar RTO distributions in Figure 25, with noticeable clusters at 1 second, 3 seconds, and 300ms.

7 DISCUSSION

Based on the measurements taken during this research, three major areas of discussion and impact arise.

ICS Network and Device Implementations Throughout the course of this research, one of the recurring observations made was that there were widespread implementation issues causing non-ideal network performance. Although the dataset used in this research is not large enough to conclusively determine how many other ICS

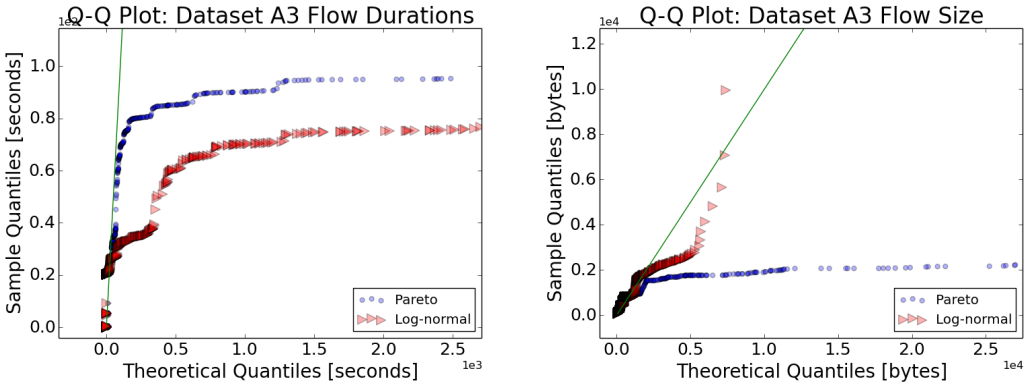


Fig. 22. A3 Q-Q Plots

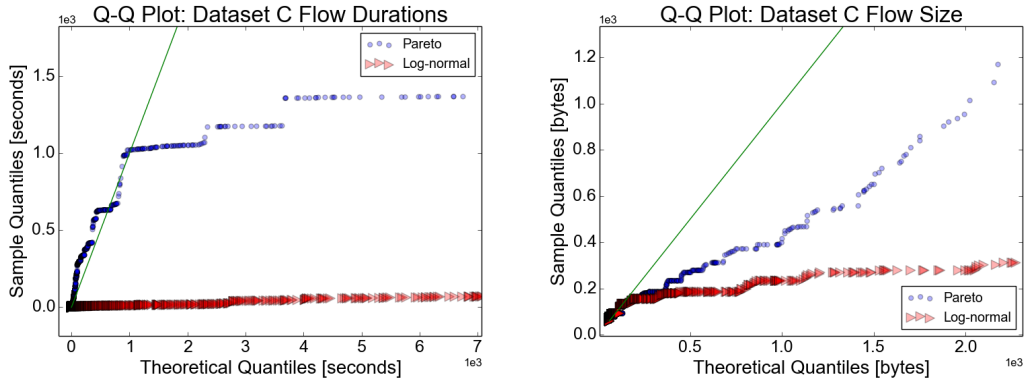


Fig. 23. C Q-Q Plots

Table 8. Round Trip Time Means

Between RTU and	μ_{A3} [ms]	μ_C [ms]	μ_D [ms]
Type A.1a	18.1	-	17.4
Type A.1b	17.4	-	12.5
Type A.2	19.0	20.3	17.2
Type B	1.6	0.4	-
CC	5.4	4.5	7.2

networks suffer similar issues, the problems appear in several popular vendors and in multiple layers of the protocol stack suggesting that it is most likely indeed a widespread problem. Examples of the variety of issues observed in this dataset include overwhelmingly large numbers of TCP connections closing ungracefully, broadcast storms occurring despite the use of the spanning tree protocol, and disagreements in the implementation of DNP3

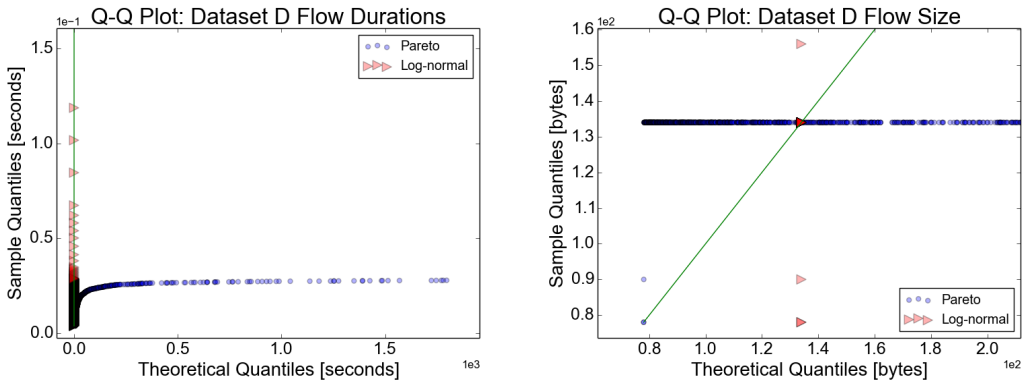


Fig. 24. D Q-Q Plots

Table 9. Round Trip Time Standard Deviations

Between RTU and	σ_{A3} [ms]	σ_C [ms]	σ_D [ms]
Type A.1a	7.35	7.4	8.3
Type A.1b	7.74	8.12	5.8
Type A.2	7.26	7.13	7.6
Type B	2.86	2.86	-
CC	2.8	2.8	1.6

Table 10. Hourly Retransmission Statistics

	μ_{total}	σ_{total}	μ_{ACK}	σ_{ACK}
A3	1.32%	0.33%	24.1%	2.21%
B	0.858%	0.016%	37.5%	0.975%
C	0.298%	0.022%	74.8%	5.25%
D	1.17%	0.027%	28.6%	0.456%

resulting in TCP connections being unnecessarily closed. Other interesting observations revealed surprising amount of jitter in the polling intervals initiated by software running on one of the most popular real-time operating systems in the world.

Since there are clearly efficiency issues in the implementations of the protocol stacks in these devices, it also suggests a strong likelihood of the presence of security vulnerabilities. In fact, a study conducted in 2014 found widespread vulnerabilities in power grid devices [6], and through the course of this research multiple ICS-CERT advisories were also released [10] [11] [9]. The fact that so many devices are still plagued with easily observable vulnerabilities that have been around for decades highlights how vulnerable these networks are and how much more attention should be paid to them.

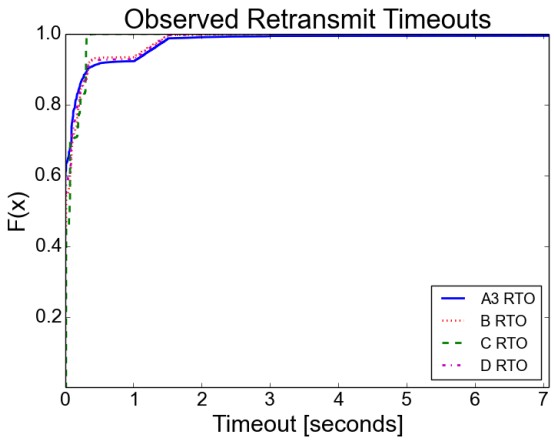


Fig. 25. Retransmit Timeout Distributions

Usage of TCP The second main observation this research offered was that due to the steady low bandwidth requirements of embedded ICS devices, the TCP protocol performs several functions that are completely unnecessary, including congestion control. Combined with the fact that storage space and processing power are both very limited on these devices, it could be very desirable to leave these functions out. The official DNP3 specification even recommends that it may be more efficient to use UDP as opposed to TCP when implemented over LANs where packet reordering is not present. However, if reliable transmission is required, then the DNP3 level acknowledgments and retransmissions would have to be used, which have not been as thoroughly studied as TCP retransmission algorithms. Due to these reasons then it may be desirable to make a compromise between efficiency and functionality and use a stripped down version of TCP/IP like uIP developed by Adam Dunkels [5].

Round Trip Times The final and perhaps the most interesting observation is that traditional RTT estimation, and by extension traditional TCP retransmission algorithms, do not perform as expected in ICS environments due to the slow processing of the embedded devices. This resulted in significant redundant retransmissions and wasted bandwidth that could have been saved if a smarter RTO algorithm was used. Furthermore, as a result of the RTTs being so dependent on processing as opposed to propagation and queuing delay, it appears that measuring RTTs reveals information as to the device's processing power and identity. As the Internet of Things continues to grow with ubiquitous embedded devices, this observation could be crucial to reducing redundant retransmissions and possibly even be leveraged for device fingerprinting and intrusion detection.

Generality Being a case study, by definition, this research only focused on the networks of a single ICS. However, several of the most interesting observations have nothing to do with the specifics of power substations and hold true for any network of embedded devices, including other ICSs. Embedded devices by their nature are designed to perform very specific tasks, which results in limited processing power and memory and equivalently, limited capability for generating large amounts of traffic. Furthermore, when they are running application layer protocols designed to regularly poll for measurements, it virtually guarantees a constant bandwidth usage and packet size. When these devices, which are still often designed to be compatible slow legacy serial links, are networked together with common IT technology capable of speeds of 100Mbps and greater, traditional notions of round trip times and critical concerns for congestion control no longer hold.

8 CONCLUSIONS AND FUTURE WORK

In this paper we present the most detailed characterization of power substation network traffic to date and examine how well the behavior of the target networks align with common assumptions about ICS and SCADA systems. We found that while most assumptions held true, there was also a surprising amount of unexpected behavior including slow and variable round trip times dominated by processing time, relatively high retransmission rates, and polling intervals with large jitter. Finally, evidence suggested that most of the various functions that TCP provides, including congestion control, are largely irrelevant in the ICS environment. These insights and observations are crucial to creating more accurate ICS network simulations and inspiring several areas of new research.

The primary limitation of this work is that its focus was limited to power substation networks at a single company. To conclusively determine whether the observations made here are generalizable to the entire power grid, as well as other industrial control systems, more data needs to be collected from a variety of control systems and networks. However, several of the most interesting observations have nothing to do with the specific network here and apply to any network of embedded devices performing real-time operations. More detailed analysis on round trip times and retransmissions could also be conducted if multiple simultaneous capture points were deployed throughout the network.

Future work will address these limitations and build on this research by collecting data from other control system networks. Other possible areas of further research would primarily involve a detailed study about fine-tuning TCP for embedded ICS devices with a focus on modifying the RTO algorithm. Additionally, the findings with respect to the round trip times being largely device dependent suggest that such measurements could be the basis for the development of new network security applications.

9 ACKNOWLEDGMENTS

We would like to thank the reviewers and our shepherd, Gil Zussman, for their constructive comments and help in improving the quality of this paper.

REFERENCES

[1] Nmap - free security scanner for network exploration & security audits. http://nmap.org/. Accessed 2015-11-23.
[2] Ieee standard for electric power systems communications – distributed network protocol (dnp3). *IEEE Std 1815-2010*, pages 1–775, July 2010.
[3] R. Barbosa, R. Sadre, and A. Pras. A first look into scada network traffic. In *Network Operations and Management Symposium (NOMS), 2012 IEEE*, pages 518–521, April 2012.
[4] A. B. Downey. Lognormal and pareto distributions in the internet. *Computer Communications*, 28(7):790 – 801, 2005.
[5] A. Dunkels. Full tcp/ip for 8-bit architectures. In *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services*, MobiSys '03, pages 85–98, New York, NY, USA, 2003. ACM.
[6] D. Formby, S. S. Jung, J. Copeland, and R. Beyah. An empirical study of tcp vulnerabilities in critical power system devices. In *Proceedings of the 2Nd Workshop on Smart Energy Grid Security*, SEGS '14, pages 39–44, New York, NY, USA, 2014. ACM.
[7] D. Formby, P. Srinivasan, A. Leonard, J. Rogers, and R. Beyah. Who's in control of your control system? device fingerprinting for industrial control system networks. In *2016 Symposium on Network and Distributed System Security (NDSS'16)*, February 2016.
[8] 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*, 17(6):6–16, Nov 2003.
[9] ICS-CERT. Icsa-15-295-01, 2015.
[10] ICS-CERT. Icsa-15-300-01, 2015.
[11] ICS-CERT. Icsa-16-070-01, 2016.
[12] V. Jacobson. Congestion avoidance and control. *SIGCOMM Comput. Commun. Rev.*, 18(4):314–329, Aug. 1988.
[13] S. S. Jung, D. Formby, C. Day, and R. Beyah. A first look at machine-to-machine power grid network traffic. In *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, pages 884–889, Nov 2014.
[14] M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow. Tcp selective acknowledgment options, October 1996. RFC 2018.
[15] V. Paxson. End-to-end internet packet dynamics. *Networking, IEEE/ACM Transactions on*, 7(3):277–292, Jun 1999.

[16] V. Paxson and M. Allman. Computing tcp's retransmission timer, November 2000. RFC 2988.
[17] V. Paxson, M. Allman, J. Chu, and M. Sargent. Computing tcp's retransmission timer, June 2011. RFC 6298.
[18] I. Psaras and V. Tsaoussidis. The tcp minimum rto revisited. In *IFIP Networking*, May 2007.
[19] 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, pages 65–76, New York, NY, USA, 2012. ACM.
[20] 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*, pages 3871–3874, Sept 2007.
[21] M. Zalewski. p0f v3. http://lcamtuf.coredump.cx/p0f3/. Accessed 2015-11-23.