Received: 2 March 2021 | Accepted: 9 March 2021

DOI: 10.1002/nem.2158

RESEARCH ARTICLE

WILEY

Industrial control protocols in the Internet core: Dismantling operational practices

Marcin Nawrocki[1] | **Thomas C. Schmidt**[2] | **Matthias Wählisch**[1]

[1]Institute of Computer Science, Freie Universität Berlin, Takustr. 9, Berlin, 14195, Germany

[2]Department Informatik, HAW Hamburg, Berliner Tor 7, Hamburg, 20099, Germany

Correspondence
Marcin Nawrocki, Institute of Computer Science, Freie Universität Berlin, Takustr. 9, Berlin, 14195, Germany.
Email: marcin.nawrocki@fu-berlin.de

Present address
Marcin Nawrocki, Department Informatik, HAW Hamburg, Berliner Tor 7, Hamburg, 20099, Germany

Funding information
Bundesministerium für Bildung und Forschung

Summary

Industrial control systems (ICS) are managed remotely with the help of dedicated protocols that were originally designed to work in walled gardens. Many of these protocols have been adapted to Internet transport and support wide-area communication. ICS now exchange insecure traffic on an inter-domain level, putting at risk not only common critical infrastructure but also the Internet ecosystem (e.g., by DRDoS attacks). In this paper, we measure and analyze inter-domain ICS traffic at two central Internet vantage points, an IXP and an ISP. These traffic observations are correlated with data from honeypots and Internet-wide scans to separate industrial from non-industrial ICS traffic. We uncover mainly *unprotected* inter-domain ICS traffic and provide an in-depth view on Internet-wide ICS communication. Our results can be used (i) to create precise filters for potentially harmful non-industrial ICS traffic and (ii) to detect ICS sending unprotected inter-domain ICS traffic, being vulnerable to eavesdropping and traffic manipulation attacks. Additionally, we survey recent security extensions of ICS protocols, of which we find very little deployment. We estimate an upper bound of the deployment status for ICS security protocols in the Internet core.

1 | INTRODUCTION

Industrial control systems (ICS) are used to monitor and control industrial environments. Deployments can range from a few controllers in a factory to large distributed systems that monitor critical infrastructures. The underlying ICS communication is based on specialized, often proprietary protocols.

Originally, ICS protocols were designed to operate in closed environments, which do not require authentication and encryption. The lack of security features in ICS protocols remained largely unnoticed due to the deployment in isolated (trusted) environments. This changed recently when ICS protocols have been stacked onto IP, enabling the management of ICS controllers via the global Internet. Such communication requires protective measures, either via secure tunnels between trusted domains or end-to-end authentication and encryption. Visible (unencrypted) ICS traffic is particularly dangerous since it is prone to eavesdropping and manipulation attacks. Traffic traces also hint attackers to potentially open ICS services without the need to perform suspicious scans. Figure 1 sketches encrypted and visible traffic flows between ICS. It also shows a passive vantage point and an

Abbreviations: AS, autonomous system; ICS, autonomous system; ISP, Internet service provider; IXP, Internet service provider.

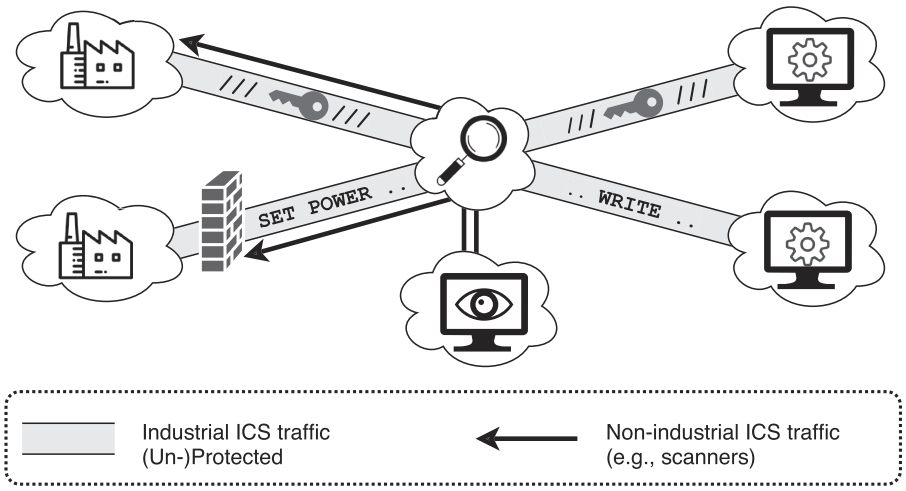


FIGURE 1 Analyzing unprotected ICS protocols

active scanner, which might be blocked by a firewall. Note that the firewall does not help in the case of man-in-the-middle manipulation attacks.

In this paper, we provide the first comprehensive analysis of the visibility of *unprotected* ICS traffic across network domains. In contrast to related work[1,2] which reveals reachable ICS services, we explore the communication of the whole ICS ecosystem, from the ICS controllers to the management stations. We show that ICS systems are controlled remotely without any protective mechanisms, harming both the Internet as well as the industrial infrastructure. Our results attract attention to the insecure usage of ICS protocols and motivate secure ICS deployments based on amendments such as DTLS and encrypted tunnels. To explore the deployment of encrypted ICS traffic, we extend our previous work[3] and provide methods and analysis that reveal limited secure ICS protocols.

In detail, our contributions are the following.

1. We present the first analysis of inter-domain ICS traffic at two central Internet vantage points, an Internet exchange point and an Internet service provider, covering 6 months.
2. We find new unprotected ICS deployments which are undetected by recent scan projects.
3. We classify industrial and non-industrial ICS traffic based on cross-correlations with other data sources such as honeypots.
4. We assess common tools for implementing our proposed methodology to allow for future long-term monitoring and mitigation.
5. We survey recent security extensions of ICS protocols and assess the potential to detect encrypted ICS protocols in Internet traffic.
6. We analyze the deployment status of encrypted ICS protocols seen from the Internet core.

The remainder of this paper is structured as follows. Section 2 presents a taxonomy and related work about ICS protocols. Section 3 introduces our methodology and data sources to identify ICS traffic. Section 4 presents basic properties of ICS traffic seen at the IXP and ISP. Section 5 proposes a method to separate industrial and non-industrial ICS traffic. Section 6 analyzes industrial ICS traffic in detail. Section 7 provides an upper bound of the usage of recent ICS protocol security extensions. Section 8concludes our findings.

2 | BACKGROUND AND RELATED WORK

2.1 | ICS protocol taxonomy

ICS protocols are deployed in four major application areas[2]: (i) process automation, (ii) building management, (iii) smart grids including power plants, and (iv) metering infrastructures, an overview is presented in Table 1. All of these scenarios require security support when the ICS devices are interconnected via untrusted networks.

TABLE 1 Overview of ICS protocols [ND/HD: Normal/Heuristic Dissector, C: Censys, S: Shodan, R: Rapid7, K: Kudelski]

Standard/Protocol	Ports	Use case	Wirshark dissectors	Min. # bytes to identify protocol	Scan projects	Honeypot software	# CVEs
Modbus	502	Process automation	ND	74 B	C/S/K	✓	23
Siemens S7	102	Process automation	HD	93 B	C/S	✓	7
Ethernet/IP	2,221, 2,222, 44,818	Process automation	ND	74 B	S	✓	9
BACnet	47,808–47,823	Building management	ND	46 B	C/S/R/K	✓	7
DNP3	20,000	Smart grids	ND/HD	62 B	C/S	×	39
HART IP	5,094	Process automation	ND	78 B	S	×	6
IEC60870-5-104	2,404	Smart grids	ND	76 B	S	(×)	0
ANSI C12.22	1,153	Metering	ND	n/a	×	×	0
OMRON FINS	9,600	Process automation	ND	54 B	S	×	7
IEC61850 (mms)	102	Smart grids	ND/HD	144 B	×	×	0
Codesys	2,455	Smart grids	×		S	×	20
GE-SRTP	18,245, 18,246	Process automation	×		S	×	7
Niagara Fox	1,911, 4,911	Building management	×		C/S/K	×	5
MELSEC-Q	5,006, 5,007	Process automation	×		S	×	2
ProConOS	20,547	Process automation	×		S	×	1
PCWorx	1,926	Process automation	×		S	×	0
Crimson	789	Process automation	×		S	×	0
ICCP-TASE.2	102	Smart grids	×		×	×	8

The most common use case for ICS protocols is process automation using programmable logic controllers (PLC), which support manufacturing facilities by assisting production. PLCs are configured and queried by ICS protocols. Well-known protocols in this field are Modbus (general industrial networks), Siemens S7 (automobile), Ethernet/IP (time-critical applications), and HartIP (legacy wiring). Equipment and manufacturing facilities also rely on proprietary PLCs that utilize protocols such as Omron, GE-SRTP, Melseq-Q, ProConOS, or PCWorx. The Crimson protocol is used exclusively for human–machine interface (HMI) communication related to Red Lion units.

Remote management of buildings is significantly based on two protocols, BACnet and Niagara Fox. They are deployed to control heating, air-conditioning, lighting, fire detection, and so on. BACnet is used to communicate directly with controlling components. In contrast, Niagara is in use between management workstations, which then subsequently communicate with the controlling components.

Electrical and water companies use protocols such as DNP3, IEC60870-5-104, IEC61850 (goose, mms), Codesys, and ICCP to monitor and automate their power systems. DNP3 is a set of sub-protocols that were released in the early 1990s before the standards IEC60870-5-104 and IEC61850 have been established which became prevalent in this application domain.

Smart meters record the consumption of electric energy and communicate that information to billing centers. The standard protocol for this application in North-America is ANSI C12.22, which delivers measurement data as clear-text tables.

2.2 | A glimpse into ICS protocol security

Vulnerable ICS deployments have been highlighted since several years.[4,5] The first reported incident is an unauthorized manipulation of an ICS which led to a pipeline explosion back in 1982.[6] Although the absolute number of reported ICS incidents is fairly low,[6] a single incident can be hazardous. To understand and improve the protection of ICS deployments, multiple efforts have been undertaken, including (i) the development of honeypots, (ii) Internet-wide scans to find open ICS devices, (iii) the improvement of intrusion detection systems for ICS, and (vi) the modeling and surveying of the ICS ecosystem.

ICS-specific honeypots have been developed[7-10] to understand the origin, frequency, and sophistication of attacks on ICS services. ICS services are popular victims. ICS honeypots receive significantly more requests after being listed on public scanning sites such as Shodan.[11]

Two well-known scan projects, Censys and Shodan, detect globally reachable ICS services.[1,2] Such scan results can be used to asses the security of ICS in individual countries.[12] ICS scans are dominated by few recurrent scanners[13] and captured within few days by honeypot deployments.[1] Mirian et al[2] measured the increase of open ICS services of up to 20 % in 4 months.

Dedicated intrusion detection systems (e.g., for smart meters[14]) and extensions to common IDS tools (e.g., Snort and Bro[15-17]) have been proposed. Valdes[18] introduces an architecture that monitors ICS traffic for irregular patterns. Taking into account recent, distributed ICS deployments, Zhang[19] proposed a distributed multi-layered system.

ICS traffic patterns have been compared with SNMP traffic.[20] Both, ICS protocols and SNMP, show stable, periodical traffic patterns with a small number of constant host changes. However, ICS traffic does not present diurnal patters or self-similar correlations, features known from traditional network protocols.[21] In contrast to our approach, the data for this comparison were collected directly at the corresponding edge-network (traditional network and ICS-facilities). So no protocol classification was necessary.

ICS have been surveyed in several publications introducing historical background, taxonomies, and current security vulnerabilities.[22-27] The number of common vulnerabilities and exposures (CVEs) for ICS implementations grows steadily. Vulnerabilities are often discovered by simple fuzzing techniques.[28,29] Also, the ICS ecosystem requires a secure supply chain.[30] Recent studies show the high DDoS potential of BACnet by analyzing IXP and ISP packet samples over a period of 48 h.[31] Yet still open is a longitudinal analysis of unprotected ICS communication deployed in the global Internet.

2.3 | The problem of unprotected ICS protocols

Most of the common ICS protocols lack protection by design and are susceptible to eavesdropping and traffic manipulation attacks. The only exception is Niagara Fox, which provides authentication. However, authentication alone is

insufficient. Attackers can scout their target and prepare a targeted attack without communicating with the ICS devices at all. Recent malware[32] exploits passive recording of ICS traffic traversing small enterprise routers. Such eavesdropping of unprotected ICS traffic is also possible on the inter-domain level.

Furthermore, it is important to note that infrastructure-based protections such as firewalls or NAT only partially help. They may prevent discovering ICS devices by *active* scanning but do not protect against *passive* listening and spoofed replay attacks.

In this paper, we analyze the highly vulnerable part of the ICS ecosystem; those cases where operators interconnect their systems without any protection. This is challenging because unprotected industrial ICS traffic is suppressed by noise such as scan traffic.

2.4 | ICS scans seen from an internet telescope

To motivate our aim for a detailed classification of ICS traffic, we briefly analyze data from the CAIDA/UCSD network telescope. This data source captures backscatter traffic from randomly spoofed DDoS attacks or Internet-wide scans of the /8 CAIDA/UCSD darknet. Any incoming traffic to the telescope is inter-domain and non-industrial.

Figure 2 shows the daily activity for Modbus (TCP/502), measured at the telescope. There is almost no activity visible until the beginning of 2014. Then, the amount of destination IP addresses that received data on the Modbus port increased by three orders of magnitude. The number of source IP addresses that sent data to the telescope increased by roughly one order of magnitude, indicating scanning from a small set of hosts. The sudden upturn in scan activities can be explained by (i) increased media coverage of ICS systems and (ii) increased research interest and consequently publicly available scan tools. Our observations correlate with the start of the ZMap and Censys projects. We saw no correlation with Shodan, which started to index ICS infrastructures in 2009 and added ICS protocols in 2012.

This brief analysis does not only highlights the increasing interest in ICS protocols but also the need for a careful methodology to analyse ICS traffic.

3 | IDENTIFICATION OF ICS TRAFFIC

Two challenges need to be tackled for analyzing inter-domain ICS traffic. First, we need to reliably identify ICS traffic in global packet traces. Second, we need to distinguish industrial (i.e., transferred by real deployments) from non-industrial ICS traffic (e.g., scanning). In this section, we propose our methodology to solve the first challenge and tackle the second challenge in Section 5.

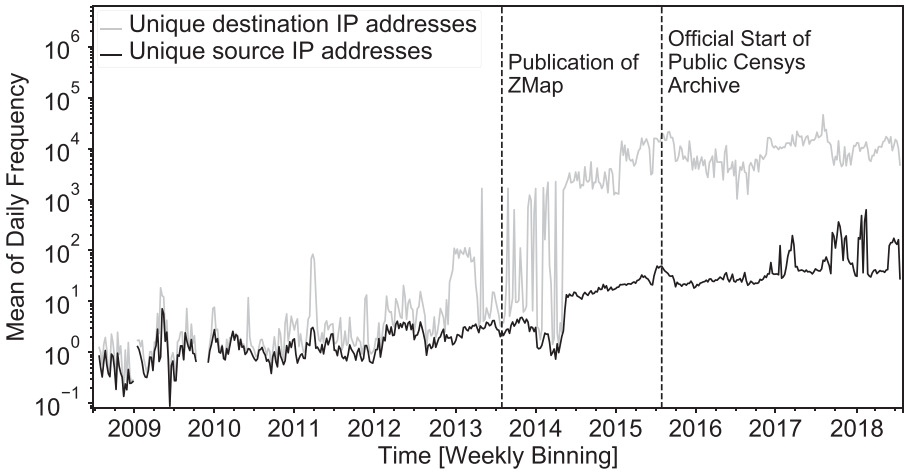


FIGURE 2 Internet-wide scanning of Modbus (TCP/502) observed at the CAIDA network telescope. We highlight research activities around one of the most common ICS scanners

3.1 | Collecting traffic at central internet vantage points

We passively collect traffic at two different Internet vantage points, an IXP and an ISP. The two data source allow us to inspect traffic from two different perspectives, a rich interconnection fabric and an upstream provider.

Internet exchange points (IXP) are centralized network infrastructures where heterogeneous domains intertwine. We receive data from a large, regional IXP from Europe with over 100 member networks with a daily traffic peak of 560 GBit/s. Due to the large traffic volume, flow data are not fully recorded but selectively sampled. We analyze non-anonymized packets collected from October 2017 until April 2018 with a sample rate of $\sim 2^{14}$. The sampled packets are truncated after 128 bytes. Flows from an IXP are inherently inter-domain.

Our second data source is the *Measurement and Analysis on the WIDE Internet* (MAWI) archive.[33] This archive contains daily traces describing 15 min of full traffic captures from a transpacific Internet link between Japan and the United States. We received a private MAWI data set with non-anonymized IP addresses and payload (96 bytes) for the same time range.

Non-anonymized flows allow for mapping with additional meta data, such as autonomous systems (AS).

Please note that we are not allowed to release our data due to privacy constraints.

3.2 | Identifying ICS traffic candidates

We explicitly do not want to implement new traffic classifiers as this conflicts with maintainability and reproducibility on the long-term. Instead, we want to leverage existing tools. We use Wireshark dissectors to find ICS traffic candidates. Half of the ICS protocols can be dissected by Wireshark, as shown in Table 1. Wireshark distinguishes between normal and heuristic dissectors (ND and HD). Normal dissectors identify protocols based on well-known port numbers and check whether the packets comply with simple sanity checks. If they fail, they forward the data to heuristic dissectors which apply pattern matching on protocol fields.

To verify the correctness of the Wireshark dissectors, we apply them on public ground truth data[34] and manually inspect the dissection of packet headers. All dissectors except one work accurately and map operation codes to protocol actions, such as *read* or *write*.

Packet sampling at our vantage points does not store complete packets but only a pre-configured fixed size of the overall packet. This limitation can lead to inaccuracies in identifying the application layer protocol because parts of the corresponding headers are missing. For each protocol“ we reduce the packet length of the ground-truth data byte-wise and detect the minimal packet length required to identify the protocol correctly. All but one protocol dissector require less than 96 bytes, see Table 1. Considering that packets are truncated after 128 bytes at the IXP and 96 Bytes at the ISP, we can identify the ICS traffic candidates reliably.

3.3 | Sanitizing ICS traffic candidates

We do not rely blindly on the Wireshark dissectors. We perform three data sanitizing steps to improve data quality: ❶ We remove tunnel traffic so that we only obtain plain end-to-end traffic. This step mainly excludes ICMP unreachable messages, which encapsulate the original UDP packets. Such backscatter packets are misclassified by Wireshark as ICS traffic. ❷ We remove packets which Wireshark marks as *malformed* or cases in which the dissector reports an *error*. This occurs when the protocol detection of a packet is successful, but the complete dissection fails due to header fields that do not comply with the protocol specification. ❸ We cross-validate our data by applying NDPI,[35] a leading open-source deep packet inspection software. NDPI detects a broad range of protocols, but no ICS protocols. We exclude every packet that NDPI is able to map to a known protocol since we consider such a packet to be a false positive.

In Table 2, we quantify the remaining packets after applying our sanitizing steps. The data are shown relatively to the overall amount of identified ICS packets per vantage point. An 85% of the packets at the IXP are classified malformed, and 48% at the ISP. Wireshark detects ICS protocols although many header fields are set to unspecified values, such as unknown operation codes. This highlights that Wireshark dissectors are rather optimistic and sanitizing is required for a reliable analysis. The removal of packets identified by NDPI accounts for less than 1%, which indicates a very low false-positive rate of our approach. Finally, we compare our approach with a pure port-based detection. Identifying ICS traffic only based on port numbers is not feasible as it leads to significant overestimation.

TABLE 2 Effects of data sanitization process and the ratio of remaining ICS packets by vantage point

	Remaining packets	
	IXP	**ISP**
Sanitizing steps after Wireshark ICS detection	100%	100%
❶ Removal of tunnel packets	99%	99%
❷ Removal of malformed packets	15%	52%
❸ Removal of NDPI fingerprintable packets	14%	51%
Comparison with vanilla approach		
Port-based detection relative to Wireshark	3,950%	1,340%

Abbreviations: ICS, industrial control systems; ISP, Internet service provider; IXP, Internet exchange point.

4 | PROPERTIES OF ICS TRAFFIC

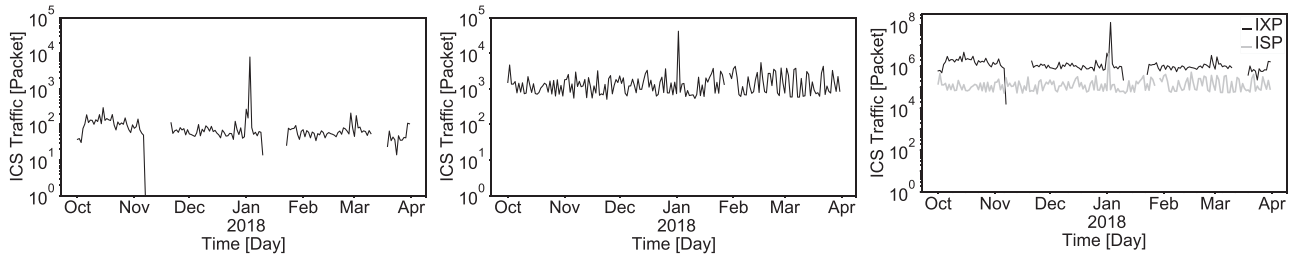
4.1 | Daily patterns and prevalence of inter-domain ICS traffic

During our measurement period, we identified 19k ICS packets at the IXP and 310k ICS packets at the ISP after sanitization. Figure 3 shows the number of daily ICS packets at the IXP and ISP. For better comparison, we consider the different sampling intervals and extrapolate the values (see Figure 3C). The daily ICS traffic at the IXP and ISP is constant apart from one anomalous peak at each vantage point. The traffic peak at the IXP is due to a large number of Ethernet/IP packets (217.5 MB/s traffic peak) during 10 min on January 3, 2018. The destination is a single IP address, and the traffic is sent from several sources located in two AS. The traffic peak at the ISP consists of BACnet messages from 76 source IP addresses to 41,000 destination IP addresses. This event took place 1 day before the IXP peak. We observe uniformly distributed BACnet read messages, which indicates load balancing between scanning nodes. All sources relate to Rapid7 Sonar, a company that performs regular Internet-wide BACnet scans.

Compared to the total traffic volume, ICS inter-domain traffic is low. ICS packets only account for ≈ 0.0001% of all sampled packets at the IXP and ≈ 0.002% at the ISP. However, putting ICS traffic into perspective of well-known non-ICS protocols, ICS traffic is likewise prevalent, which we show in Figure 4. To allow for comparability, this graph visualizes the non-sanitized data set because implementing a sanitization process for non-ICS protocols would be out of scope of this work. This result emphasizes that ICS traffic should not be neglected.

4.2 | ICS message types: Request versus reply

We refer to packets sent to a known ICS port as requests, and packets originating from a known ICS port as replies. Protocols with balanced request-reply ratios are likely to be used in a legitimate way since ICS communication patterns



(A) Internet Exchange Point (IXP) (B) Internet Service Provider (ISP) (C) Extrapolation by sampling rates.

FIGURE 3 Number of inter-domain ICS packets per day at two different vantage points

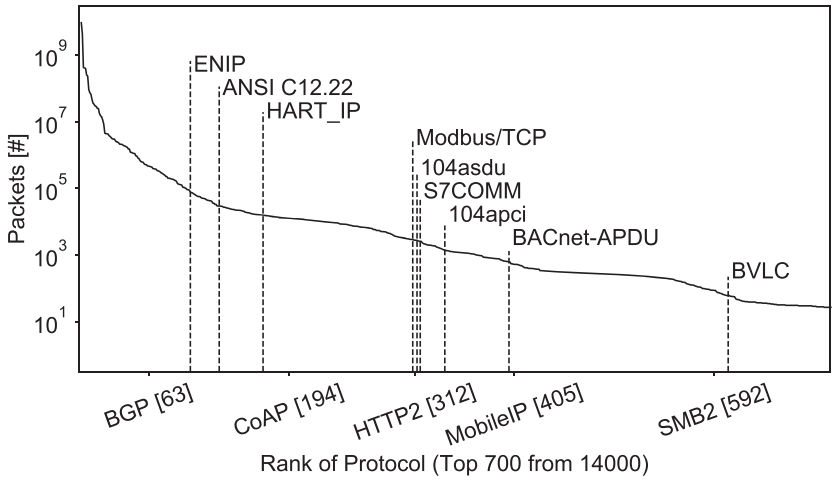


FIGURE 4 Protocols ranked by packet frequency as reported by Wireshark (non-sanitized), observed at a big national IXP during 6 months. ICS protocols are emphasized among some well-known protocols. Ranks are noted in brackets

follow a common client server scheme. Observing significantly enhanced requests may have two reasons: (i) heartbeats sent from sensors to central servers that do not confirm the reception; (ii) scan traffic that reaches hosts which do not offer the corresponding service.

We analyze the ratio of requests and replies per protocol in more detail, check left-hand side of Table 4. We observe a tendency towards requests exceeding replies. Only at the IXP, HartIP and C12.12 show a balanced request-reply ratio. Strikingly, BACnet is very request-heavy across both vantage points. This might be an indication for non-industrial ICS traffic, which we will investigate further in Section 5.

4.3 | ICS traffic sent to and received from AS

To better understand the ICS ecosystem from a networking perspective, we map each source and destination IP address of a sampled packet to autonomous system numbers (ASN). We use daily data from the RIPE RIS project and topological information from the IXP for assigning ASNs.

AS which are the origin of request traffic via multiple ICS protocols host either scanners or heterogeneous ICS monitoring services. In our sanitized data sets, more than 70% of the ASes host nodes that deploy a single ICS protocol, see Figure 5. We find four cases of ASes creating requests for greater than four distinct ICS protocols. Three are eyeball providers, and one is a webhoster. These types of networks are common to connect scanners, which we detect in Section 5.

5 | IDENTIFICATION OF INDUSTRIAL AND NON-INDUSTRIAL ICS TRAFFIC

Separating non-industrial from industrial ICS traffic allows us to identify the vulnerable part of the ICS ecosystem more precisely. We classify ICS traffic at our vantage points as non-industrial if the captured IP addresses belong to scan projects or have been observed at honeypots, as those indicate non-ICS hosts.

5.1 | Filter traffic of common scan projects

Several projects scan for ICS hosts on a regular basis and thus contribute to non-industrial inter-domain ICS traffic. The most common projects are Censys, Shodan, Rapid7, and Kudelski. Censys, Rapid7, and Kudelski publicly document the IP prefixes from which they initiate scans. We use these prefix lists to identify scanners by marking an observed source IP address as scanner if the source IP address is covered by one of the prefixes.

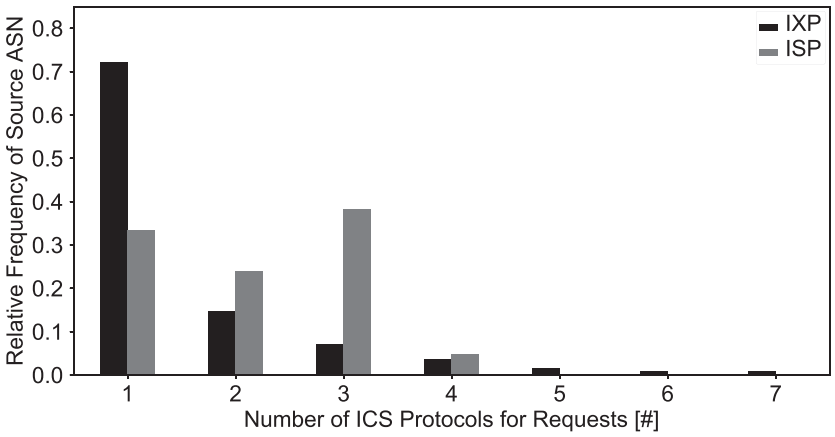


FIGURE 5 Number of ASes sending different ICS protocol requests. Since ICS deployments are rather specific deployment and bound to a single manufacturer, we rate several ICS protocols originating from a single AS as suspicious

TABLE 3 Amount of successful reverse DNS lookups of source IP addresses per scan project

	IXP	ISP
# Unique source IP addresses	1,504	223
# Resolvable Censys IP addresses	105	n/a
# Resolvable Rapid7Labs IP addresses	7	56
# Resolvable Kudelski Sec. IP addresses	0	0
# Resolvable Shodan IP addresses	23	25

Abbreviations: ISP, Internet service provider; IXP, Internet exchange point.

To identify scanners that are not part of the documented IP prefixes, we perform reverse DNS lookups on all source IP addresses captured at our vantage points. By reviewing the assigned names manually, we find Censys, Rapid7, and Shodan scanners (e.g., *pirate.census.shodan.io* and *scanner2.labs.rapid7.com*). Note that we cannot identify any names that relate to Censys at the ISP because Censys performs scans between ≈ 8:00am and ≈ 6:00pm (UTC), whereas the ISP dumps include 15 min packet captures starting at 5:00am (UTC).

Table 3 shows the amount of successful reverse DNS lookups. The IXP and ISP share 86 source IP addresses, predominantly Shodan and Rapid7 scanners. The five most common source IP addresses at the ISP resolve to Shodan names and are located in Quasi Networks, an AS which is also well-known for hosting malicious nodes.[36]

5.2 | Filter traffic of other non-ICS hosts

To account for other hosts that create non-industrial ICS traffic (e.g., attackers), we leverage data from honeypots. Conpot[37] is the de-facto standard ICS honeypot but supports only five ICS protocols, one currently under development. Conpot implements limited variances in responses, which makes it easy to unmask as a honeypot. Thus, we argue to utilize transport layer honeypots in order to measure a broad scope of activities on ICS ports.

We deploy Honeytrap[38] in (i) a university network and (ii) a darknet, a network not offering any public services. Based on these honeypots, we identify suspicious source IP addresses. We create two lists: HP_{all}, which stores all IP addresses observed at the honeypots, and a subset of this list, HP_{ICS}, which stores IP addresses that sent requests to at least one ICS port. HP_{all} consists of 244k IP addresses and HP_{ICS} of only 3,700 IP addresses (1.5%) from 619 ASes. It is worth noting that our honeypots also capture sources of the well-known scan projects. Two hundred and twenty-four IP addresses in HP_{all} are from Censys scanners.

TABLE 4 Relative amount of industrial ICS traffic after applying different filter rules on the observed ICS traffic

	Request response ratio				Traffic share after applying filters							
	# ICS packets		Share of requests		Excluding scanners		Excluding captured honeypot data				Excluding both	
	IXP	ISP	IXP (%)	ISP (%)	IXP (%)	ISP (%)	IXP		ISP		IXP (%)	ISP (%)
							HP_{ICS} (%)	HP_{all} (%)	HP_{ICS} (%)	HP_{all} (%)		
Total	19,060	310,996	81	99	97	46	97	96	15	1.5	96	1.5
BACnet	568	89,922	98	100	15	7	25	11	40	1	10	1
C12.22	1,559	24	63	29	100	100	100	99	100	100	99	100
DNP3	2	2,424	100	100	100	99	0	0	0.4	0.1	0	0
Ethernet/IP	9,171	171,804	94	99	99	75	98	98	5	0.02	98	0
HartIP	126	46,783	62	92	65	9	62	62	9	8	62	8
IEC60870	2,511	13	13	38	100	100	100	99	100	100	99	100
Modbus	2,547	—	95	—	100	—	100	100	—	—	100	—
Siemens	2,576	—	99	—	100	—	100	100	—	—	100	—

Abbreviations: ICS, industrial control systems; ISP, Internet service provider; IXP, Internet exchange point.

We now correlate ICS traffic from our vantage points with the honeypot data. For every observed ICS packet, we check whether the source or destination IP address is present in HP_{all} or HP_{ICS}, see Table 4.

At the IXP, the overlap is minimal, which means that a significant amount of industrial ICS traffic is visible. Ninety-six percent of ICS traffic is industrial based on HP_{all}, 97% based on HP_{ICS}. We perform a comparison per protocol and correlate 506 BACnet packets with HP_{all}, which represent 89% of the total BACnet packets at the IXP. These packets are classified as non-industrial ICS traffic and filtered. The results are stable, even if we only consider HP_{ICS}.

At the ISP, less industrial ICS traffic is visible. Filtering by HP_{all}, we find only 1.5% of the traffic to be industrial. However, the filtering is less effective if we only consider HP_{ICS}, especially for BACnet. The results indicate that it is beneficial to include honeypot information from non-ICS ports.

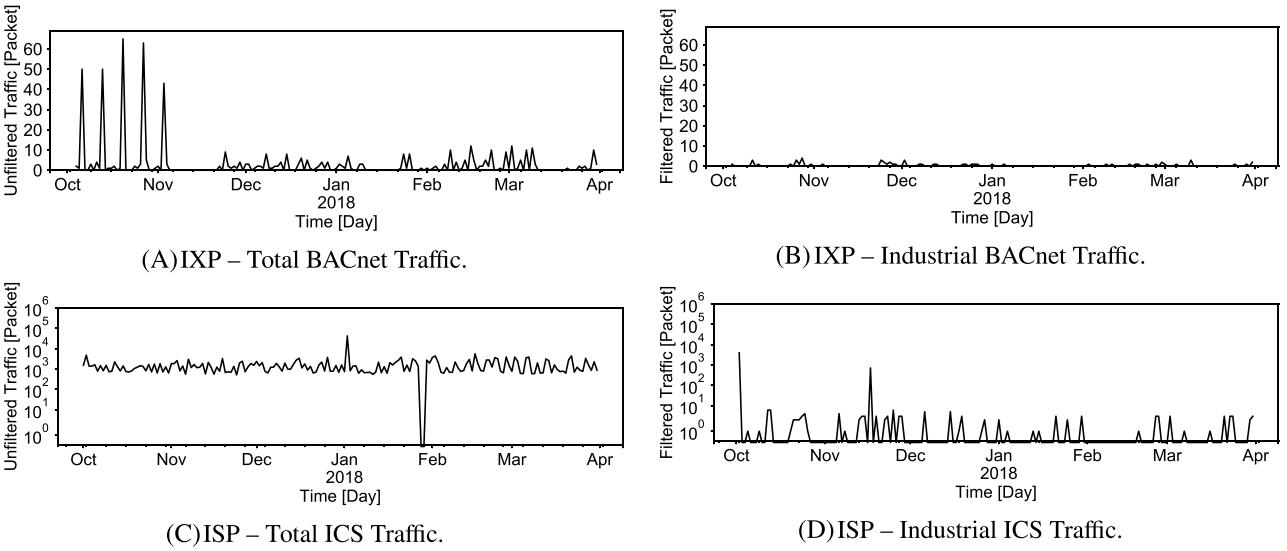
5.3 | Benefits of combining filter rules

To summarize the results from our previous filter steps, we provide an overview of the impact of the different filters. Table 4 shows the relative amount of ICS traffic that remains when traffic from scanners (identified by DNS names and IP prefixes), honeypots, or both is excluded.

While we classify 96% of the traffic at the IXP as industrial, we see only 1.5% of industrial traffic at the ISP. Interestingly, more than half of the traffic at the ISP can already be classified as non-industrial only by excluding public scanners, that is, without maintaining a dedicated infrastructure such as honeypots. However, even though maintaining a honeypot introduces additional complexity, its data are necessary to provide a more complete view on distinguishing industrial and non-industrial traffic.

ICS protocols show similar trends for the share of non-industrial traffic across both vantage points. The substantial difference for Ethernet/IP is caused by a Shodan scan of a complete prefix range at the ISP.

We show the potential effects of filtering non-industrial ICS traffic over 6 months in Figure 6. This enables us to describe the impact of non-industrial traffic over time. At the IXP, we focus on BACnet as this protocol is severely affected by non-industrial activity. We make two observations: (i) at the IXP, non-industrial traffic consists mainly of ephemeral spikes at the beginning of our measurement period. (ii) At the ISP, the non-industrial traffic shows a very constant daily activity. After filtering at both vantage points, we obtain only a few industrial ICS packets per day which allows even for manual inspection of the ICS traffic.



(A) IXP – Total BACnet Traffic.

(B) IXP – Industrial BACnet Traffic.

(C) ISP – Total ICS Traffic.

(D) ISP – Industrial ICS Traffic.

FIGURE 6 Daily amount of all ICS traffic versus industrial ICS traffic visible at the IXP and ISP

TABLE 5 Successful transport and application layer handshakes during Censys scans

	# ICS hosts detected by Censys		
Protocol	**Transport scan**	**Application scan**	
BACnet	31,735	31,154	(98%)
Modbus	8,400,058	126,984	(2%)
Siemens S7	7,202,828	24,946	(0.5%)

Abbreviation: ICS, industrial control systems.

6 | PROPERTIES OF ICS INDUSTRIAL AND NON-INDUSTRIAL TRAFFIC

6.1 | Detecting ICS hosts protected by firewalls

ICS devices might be protected by firewalls which grant access only from specific hosts. We analyze this by comparing IP addresses observed in our passive data with IP addresses of ICS devices revealed by active scans. To reduce overhead on the Internet infrastructure,[39] we do not implement our own active probing but use data from Censys. Censys continuously scans the entire public IPv4 address space fast,[2,40] implements full transport and application layer handshakes,[40] and releases weekly snapshots. We compare three ICS protocols for which we found industrial traffic and which are scanned by Censys during our measurement period: Siemens S7, Modbus, and BACnet.

First, we check how many ICS hosts are detected by Censys on the transport and application layer (see Table 5). Despite many successful transport layer handshakes, Modbus and S7 exhibit a very low success rate on the application layer. We argue that this is related to the use of lower port numbers that are more likely to be used by other applications which listen on the corresponding port. This complies with our previous results which showed that port-based ICS detection is misleading (see Section 3.3).

Now, we compare with ICS hosts observed at our vantage points. We compute the fraction of source or destination IP addresses that have been discovered by Censys (see Table 6) and for which we see communication in our passive data, that is, completely unprotected nodes. At the IXP, 35% of the Modbus and 65% of Siemens destinations are already known because of the transport layer scan. At the ISP, we do not find any correlation, that is, none of the ICS devices that are visible in our ISP traffic data set have been captured by active scans. This is very likely due to port-based access control lists which only allow communication between pre-configured hosts.

We find three source IP addresses that respond to Modbus transport layer scans but do not establish successful application layer sessions based on Censys. However, based on our traffic traces, each of these hosts has sent about 45 Modbus packets. One host is sending packets to a solar energy consulting agency. These results indicate cases of secure ICS services but unprotected ICS traffic.

TABLE 6 Ratio of ICS hosts observed at the IXP and Censys

	% ICS hosts that overlap with Censys	
Host type at IXP	**Transport scan (%)**	**Application scan (%)**
BACnet source	0	0
BACnet destination	0	0
Modbus source	3	0
Modbus destination	35	0
Siemens source	0	0
Siemens destination	65	65

Abbreviations: ICS, industrial control systems; IXP, Internet exchange point.

6.2 | Host stability of industrial ICS traffic

Host stability describes how often a host is visible at our vantage points with respect to an activity span. For each destination IP address in the industrial ICS traffic, we calculate the size of the activity window w in days (i.e., time-lag between first and last day of occurrence) and the number of active days n within this time window.

We assume that as soon as an ICS network is in place an embedded ICS device and an ICS control station will frequently exchange ICS traffic. Furthermore, we assume static assignment of IP addresses to those devices as this will ease operational maintenance (e.g., configuration of firewall rules). Following both assumptions, hosts will achieve high host stability in case of real ICS networks, that is, the same IP address will appear for several days.

The IXP and ISP results differ significantly. In the IXP data set, the most stable host communicates almost every day ($w = 179, n = 146$). In contrast to this, in the ISP data set, hosts communicate less than 4%, relatively to the overall activity span.

To better understand whether stable hosts belong to a real ICS deployment, we map IP addresses to additional meta data: reverse DNS records and `whois` data. Based on this, we find that hosts are operated by a building company (*max-boegl.de*; $w = 179, n = 146$), a trade and transport company (*Handel Uslugi Transport Ewa Cielica*; $w = 159, n = 98$), and a industrial service and consulting company in the field of solar energy (*enerparc.com*; $w = 90, n = 36$). The high number of active days, despite the sampling, indicates a high exchange of messages. Interestingly, these hosts are not marked as ICS hosts by Censys, indicating the role of an ICS monitoring station. In the data set of our transnational ISP, we do not find evidence for ICS companies.

6.3 | Locality of non-industrial traffic

We analyze the locality of industrial and non-industrial ICS traffic. Less local traffic is more likely to be part of Internet-wide scanning activities, whereas some ICS stakeholders may consider locality as reason not to protect (industrial) ICS traffic. We distinguish between topological and geographical locality.

Figure 7 shows a typical inter-domain topology at an IXP. In addition to a source and destination AS, packets may traverse *ingress* and *egress ASes* directly connected at the IXP. ASes which send or receive packets over an IXP member are in the cone of this member. We refer to traffic as IXP local, if the following condition applies:

$$Source\ AS == Ingress\ AS \wedge Egress\ AS == Destination\ AS$$

From a topological point of view, IXP local traffic is more *trustworthy*, because both ASes peer directly with each other (maybe via a route server). In contrast to this, communication from cones is rather expected from Internet-wide scanners, which are located in edge networks. At the IXP, 90% of the BACnet and 40% of the HartIP traffic is non-industrial. Comparing peering transitions for these two protocols with Ethernet/IP, which exhibits only 2% non-industrial traffic, shows a clear distinction, see Table 7. Non-industrial traffic originates only from the cones of the IXP-members, hence is not local at the IXP.

Assuming that critical infrastructures are scanned by malicious hosts and proxies from ASes located in foreign countries, we also check how often traffic is locally bound to a country. We do this by mapping the IP addresses to country

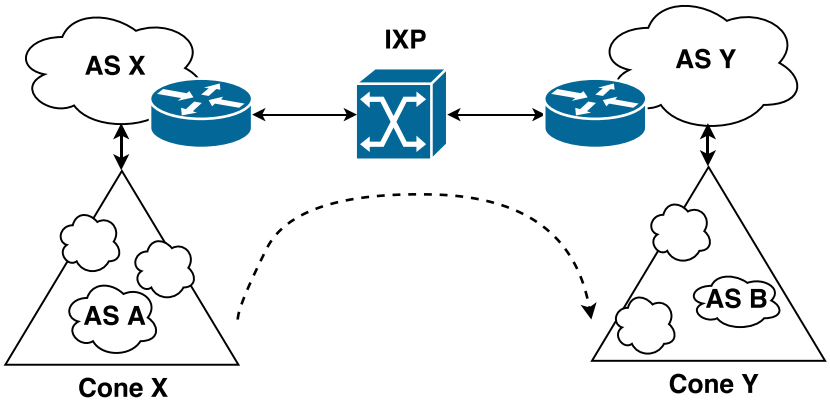


FIGURE 7 Example of cone to cone communication with ingress AS X and egress AS Y

TABLE 7 Relative ratio of traffic transitions for three industrial control systems (ICS) protocols at Internet exchange point (IXP)

	Industrial			Non-industrial		
	BACnet	HartIP	Ethernet/IP	BACnet	HartIP	Ethernet/IP
Member to member	30%	22%	5%	0%	0%	0%
Member to cone	24%	51%	29%	0%	0%	0%
Cone to member	19%	9%	6%	46%	79%	52%
Cone to cone	27%	18%	60%	54%	21%	48%
# Flows	59	78	9,006	509	48	165

Note. Non-industrial traffic originates exclusively from cones and thus is not local.

TABLE 8 Relative ratio of domestic traffic for three industrial control systems (ICS) protocols, compared to the overall traffic of each protocol at the Internet exchange point (IXP)

Industrial			Non-Industrial		
BACnet	HartIP	Ethernet/IP	BACnet	HartIP	Ethernet/IP
29%	24%	1%	0%	0%	0.5%

codes based on MaxMind.[41] If the source and destination IP addresses are located in the same country, we call the traffic *domestic*. Table 8 presents the results of our analysis of domestic traffic. Although industrial traffic is also exchanged across country borders (which might happen in the case of, e.g., global transport companies), there is a clear trend for non-industrial traffic: non-industrial traffic is strictly non-domestic, which highlights globally distributed scanning activities. On the other hand, up to 29% of the industrial traffic is local, which makes it easy to contact and train the ICS network operators in charge.

7 | ENCRYPTED ICS TRAFFIC

7.1 | ICS protocols and (D)TLS extensions

To reduce the attack space of the vulnerable, traditional ICS traffic, ICS protocols have been recently extended to support Transport Layer Security (TLS) and Datagram TLS (DTLS). While TLS works on top of TCP, DTLS works on top of UDP. These extensions ensure three security objectives:

1. integrity, that is, manipulated data is rejected,
2. authenticity, that is, messages from untrusted devices are rejected,

3. authorization, that is, not allowed actions are rejected.

The most recent TLS standard is version `1.3`, which provides major improvements in the areas of security, performance, and privacy. Most strikingly, TLS `1.3` enhances the handshake behavior by encrypting more of the initial negotiation to protect privacy-sensitive data from eavesdroppers. Also, an entire round trip from the connection establishment phase is removed.

We are aware of three ICS protocols that are extended by (D)TLS: Ethernet/IP,[42] DNP3,[43] and Modbus.[44] All of these protocols use a different default transport port in the encrypted version compared to the unencrypted version (see Table 9). Ethernet/IP and Modbus enforce the TLS standard `1.2`. This does not allow TLS downgrades during handshakes, which makes both protocols vulnerable to older TLS-based attacks.[45] DNP3 uses a proprietary security extension called Secure Authentication (SA), in addition to TLS. Please note that DNP3 SA only provides fine-grained device authentication and message integrity.[46,47] Authentication can be performed in either direction (outstation or master) and access control lists allow to enforce roles within an organization. However, as DNP3 SA does not provide encryption, it does not protect from eavesdropping or prevents ICS detection by passive traffic analysis. In this analysis, we only focus on fully encrypted traffic based on TLS and DTLS.

7.2 | Attack vectors for encrypted ICS traffic

7.2.1 | Unencrypted transport headers

ICS traffic can be secured on three different layers, the network layer (IPsec), transport layer (TLS), and within the application (e.g., SA).[47] Extending each protocol based on (D)TLS has the advantage of minimal setup requirements. (D)TLS, however, does not prevent eavesdroppers from dissecting network and transport layer headers. Thus, attackers are able to conduct a port-based analysis, trying to detect ICS deployment. Limiting the traffic to this subset, *i.e.,* focusing on (encrypted) traffic on the respective ports only, reduces computational complexity and it becomes easier for attackers to detect interesting targets. After detecting ICS deployments, attackers can utilize other attack vectors in addition to traffic manipulations to disturb operations, *e.g.,* volumetric DDoS attacks[48] or IP prefix hijacks.

7.2.2 | Machine learning classifiers of ICS traffic

An application protocol can be identified in encrypted traffic even if not only the application but also the transport and network layer are covered, for example, in RDP tunnels.[49] Usually, characteristics that allow for fingerprints are extracted based on statistical analysis. Such methods use a rich training data set and then apply the trained classifier to identify features on a target data set.[50] Machine learning approaches can also be used in the context ICS protocols, for example, DNP3 message types can be identified with high precision in encrypted IPsec tunnels.[51] We discuss, however, that hose approaches conflict with inter-domain traffic analysis as they are challenged by sampled data.

We will now inspect traffic activities on default ports of encrypted ICS protocols. Then, we will evaluate the potential of statistical fingerprints, for example, by machine learning, of ICS traffic at IXPs.

7.3 | Traffic activities on new ICS ports

We now analyze the traffic volume at the IXP for default ports of ICS protocols that support encryption extensions, see Table 9. We do this as a longitudinal study of more than 2 years so that traffic changes due to the specification and

TABLE 9 Security extensions for ICS protocols

ICS protocol	Extension	Default port
Ethernet/IP secure	TLS & DTLS	2,221
DNP3 secure	TLS & SA	19,999
Modbus secure	TLS	802

Abbreviation: ICS, industrial control systems.

market release of the new extensions become noticeable. It is worth noting that we count packets on ports independent of the transport layer payload. Also, we do not exclude non-industrial ICS traffic in order to observe potentially increased scanning activity.

Overall, we did not observe any significant increase of traffic in the last 2 years on the respective ports (see Figure 8). The total number of sampled packets to or from the new ports remains small. Only 0.013% of the total daily packet volume at the IXP can be attributed to ports that belong to encrypted ICS protocols. At the beginning of the measurement period, we observe synchronized valleys on the Ethernet/IP and DNP3 Secure ports. Also, the Ethernet/IP port exhibits one extreme traffic peak in mid 2018. Unfortunately, we could not find any links between these events and see no affiliation to ICS deployments. Based on that, we conclude two findings:

1. The encrypted versions of the ICS protocols have not yet been incorporated by ICS operators and also (potentially malicious) scan projects.
2. A simple port filter allows attackers in the Internet core to reduce the number of potential ICS candidates substantially, that is, the analysis becomes less computationally heavy.

Motivated by the second finding, we now inspect the encrypted application data for specific fingerprints.

7.4 | Application fingerprints at the IXP

We now leverage Wireshark to fingerprint (D)TLS in the detected traffic. Using ground truth data, we verified that Wireshark is able to detect (D)TLS, even in scenarios that include truncated packets. Wireshark detects (D)TLS traffic heuristically, that is, by inspecting (D)TLS record headers for valid content types and TLS versions, which are represented as a 1-complement. At the IXP, we do not find any TLS packet related to the ICS ports, but we do find on average 26 sampled DTLS packets per week (3.7k in total). All these packets include the Ethernet/IP source or destination port, compare Figure 9. Please note that this plot represents the upper bound for Ethernet/IP secure packets which are sent using default configuration.

We try to infer whether the encrypted traffic is indeed Ethernet/IP. Unfortunately, common features used in machine learning to infer application types are not available at the IXP due to aggressive packet sampling and truncation. Such features include the packet inter-arrival times, bi-directional traffic flow analysis, and bit rates.[50] Even though we are challenged by truncated packets, we can still determine the packet sizes of the IP packets by inspecting the IP length field, which is not truncated. In case of ICS packets, we expect that these DTLS packets are smaller compared to all other DTLS traffic. We observe a mostly bi-modal distribution for both traffic types, exhibiting different sizes of classes (170 vs. 200 bytes, 1,250 vs. 1,500 bytes), though, see Figure 10. Overall, if encrypted Ethernet/IP traffic is present in our candidate sets, it remains well hidden and cannot be inferred by statistical and machine learning analysis at the Internet core.

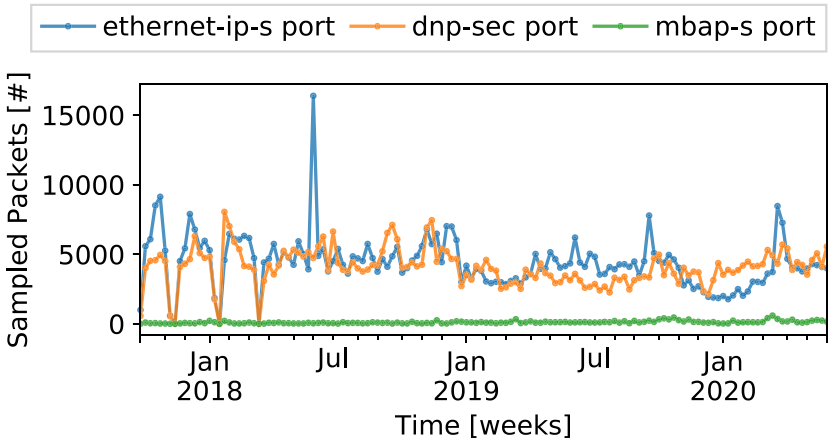


FIGURE 8 Number of packets associated with the ports of ICS protocols with encryption extensions

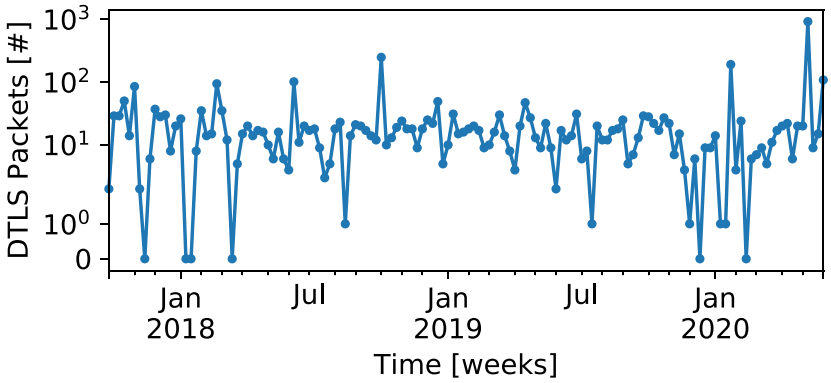


FIGURE 9 Number of DTLS packets associated with the Ethernet/IP secure port

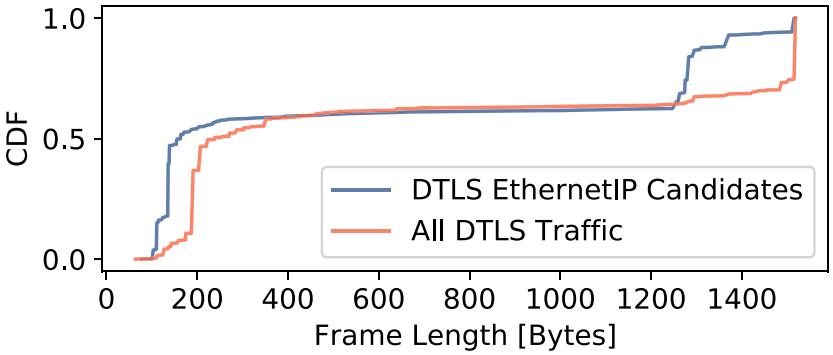


FIGURE 10 Packet sizes for all DTLS packets and Ethernet/IP secure traffic candidates

7.5 | Stable ICS deployments and encrypted traffic

As an additional crosscheck, we test whether encrypted traffic is sent by stable ICS deployments which previously exchanged unencrypted ICS traffic. To this end, we look for IPsec, that is, Encapsulating Security Payloads (ESP), and again DTLS traffic for such deployments. We find no DTLS traffic. We find, however, two ICS deployments, which used an IPsec point-to-point tunnel and exchanged 74k and 260 sampled packets, respectively. To better understand the underlying deployment, we map IP addresses of these packets to their origin AS. One of the tunnel end points is connected to an eyeball provider, the other to an architecture office. Based on these observations, we suspect that the tunnels primarily carry office-related traffic.

8 | CONCLUSION

In this paper, we analyzed the unprotected traffic of protocols that interconnect ICS. Our key results obtained from an IXP and an ISP perspective, that is, the Internet Core, read the following.

8.1 | ICS traffic identification is painful

Common open source tools for traffic classification and analysis do not identify ICN traffic reliably. Due to the limited deployment of ICS protocols, there is a lack of fingerprinting tools. We introduced and explored an advanced but lean approach to detect ICS protocols. Our methodology is based on common Wireshark dissectors, but introduced several sanitizing steps that reduce the number of false positives. Given that we have identified ICS scanners as well as industrial ICS deployments in our traffic traces, we are confident with our true positives.

8.2 | Unprotected ICS traffic is visible at the IXP

After sanitizing our data, we found over 330k ICS packets and one anomalous traffic peak at each vantage point. As Internet-wide ICS scanners operate since several years, it comes as no surprise that inter-domain ICS traffic exists. Hence, we developed a classification mechanism to differentiate between industrial and non-industrial ICS traffic. The 96%-share of unprotected industrial ICS traffic at the IXP is alarming. Since we observe a regional IXP, cooperating companies from the same region might exchange ICS traffic. In contrast, our ISP data represent a transnational link between the United States and Japan, representing the bridge between geographically distributed transit networks. Intuition suggests that distributed ICSs are rather local in deployment. Our results confirm this intuition. With only 1.5% industrial ICS traffic, the ISP is mainly confined to scans.

8.3 | New, stable ICS deployments detected

We isolated (non-) industrial ICS traffic and could classify ICS packets that were exchanged by hosts such as known from the Censys scan project. We also discovered previously undetected ICS devices, though, that belong to real ICS eco-systems. We identified cases of very stable hosts, that is, hosts that exchange ICS traffic regularly. Such hosts are vulnerable to traffic manipulation attacks on a daily basis. We spotted topological features for non-industrial ICS traffic. Such traffic originates at IXP-cones and is not domestic, that is, source and destination are not located in the same country.

8.4 | ICS security extensions are disguised

We present a first study of ICS security extensions at the IXP, with a focus on DTLS traffic. We do not find tangible signs of encrypted ICS traffic. Nonetheless, we present an upper bound for encrypted ICS traffic at the Internet core. First results suggest that (transport layer) ports registered for the security extensions experience substantially less traffic than other ports. This reduces the data size and hence computational complexity for attackers, which attempt to find new ICS deployments.

8.5 | Raising awareness of potential ICS attacks

The insights of this paper help to find unprotected ICS traffic and inform responsible stakeholders for improving protection. They also allow to deploy a long-term monitoring system that can observe malicious inter-domain ICS activities. Solutions already exist (such as SSH tunnel and VPN), but are not yet deployed, leaving ICS data exposed to eavesdropping and traffic manipulation attacks.

8.6 | Future work

In the future, we hope that ICS deployments will upgrade from unprotected configurations to secure ones. Hence, we expect increased traffic volumes on the default ports of the secure ICS protocol variants. This will pave the way for more extensive analysis, including machine learning methodologies which require larger data sets. Overall, observing ICS traffic from the Internet core will remain relevant (i) to quantify malicious scanning activities and (ii) to detect misconfigured ICS deployments, even with security extensions.

ACKNOWLEDGEMENT

Open access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

This work was supported in parts by the German Federal Ministry of Education and Research (BMBF) within the project X-Check.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from third parties. Restrictions apply to the availability of these data, which were used under license for this study. Research data are not shared.

ORCID

Thomas C. Schmidt https://orcid.org/0000-0002-0956-7885
Matthias Wählisch https://orcid.org/0000-0002-3825-2807

REFERENCES

1. Bodenheim RC. Impact of the Shodan computer search engine on internet-facing industrial control system devices. *Tech. Rep*., Wright-Patterson, Air Force Institute of Technology Wright-Patterson AFB OH Graduate School of Engineering and Management; 2014.
2. Mirian A, Ma Z, Adrian D, Tischer M, Chuenchujit T, Yardley T, Berthier R, Mason J, Durumeric Z, Halderman JA, Bailey M. An Internet-wide view of ICS devices. In: 2016 14th annual conference on privacy, security and trust, pst 2016. Institute of Electrical and Electronics Engineers Inc. IEEE; 2016; United States:96-103.
3. Nawrocki M, Schmidt TC, Wählisch M. Uncovering vulnerable industrial control systems from the Internet core. In: Proc. of 17th ieee/ifip network operations and management symposium (noms). IEEE Press IEEE; 2020; Piscataway, NJ, USA.
4. Meixell B, Forner E. Out of control: demonstrating SCADA exploitation. Black Hat USA; 2013.
5. Klick J, Lau S, Marzin D, Malchow J-O, Roth V. Internet-facing PLCs—a new back orifice. Black Hat USA; 2015.
6. Miller B, Rowe D. A Survey SCADA of and critical infrastructure incidents. In: Proceedings of the 1st annual conference on research in information technology. Association for Computing Machinery ACM; 2012; New York, NY, USA:51-56. https://doi.org/10.1145/2380790.2380805
7. Vasilomanolakis E, Srinivasa S, Mühlhäuser M. Did you really hack a nuclear power plant? An industrial control mobile honeypot. In: Communications and network security (cns) conference. IEEE IEEE; 2015:729-730.
8. Wilhoit K. Who's Really Attacking Your ICS Equipment? *Trend Micro Research Paper*; 2013:1. https://www.trendmicro.com.tr/media/wp/whos-really-attacking-your-ics-equipment-whitepaper-en.pdf
9. Winn MM. Constructing cost-effective and targetable ICS honeypots suited for production networks, Wright Patterson, Air Force Institute of Technology Wright-Patterson AFB OH Graduate School of Engineering and Management; 2015.
10. Bernieri G, Conti M, Pascucci F. MimePot: a model-based honeypot for industrial control networks. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC) IEEE; 2019:433-438.
11. Serbanescu AV, Obermeier S, Yu D-Y. ICS threat analysis using a large-scale honeynet. In: Proceedings of the 3rd international symposium for ics & scada cyber security research. BCS Learning & Development Ltd.ACM; 2015; Swindon, GBR:20-30. https://doi.org/10.14236/ewic/ICS2015.3
12. Ceron JM, Chromik JJ, Cardoso de Santanna JJ, Pras A. *Online Discoverability and Vulnerabilities of ICS/SCADA Devices in the Netherlands*. Netherlands: University of Twente; 2019. In opdracht van het Wetenschappelijk Onderzoek en Documentatiecentrum (WODC).
13. Ferretti P, Pogliani M, Zanero S. Characterizing background noise in ICS traffic through a set of low interaction honeypots. In: Proceedings of the acm workshop on cyber-physical systems security & privacy, CPS-SPC'19. Association for Computing Machinery ACM; 2019; New York, NY, USA:51-61. https://doi.org/10.1145/3338499.3357361
14. Berthier R, Sanders WH. Specification-based intrusion detection for advanced metering infrastructures. In: Ieee 17th pacific rim international symposium on dependable computing (prdc). IEEE IEEE; 2011:184-193.
15. Morris T, Vaughn R, Dandass Y. A retrofit network Intrusion Detection System for MODBUS RTU and ASCII Industrial Control Systems. In: 45th hawaii international conference on system science (hicss). IEEE IEEE; 2012:2338-2345.
16. Lin H, Slagell A, Di Martino C, Kalbarczyk Z, Iyer RK. Adapting Bro into SCADA: building a specification-based intrusion detection system for the DNP3 protocol. In: Proceedings of the eighth annual cyber security and information intelligence research workshop, CSIIRW '13. Association for Computing Machinery ACM; 2013; New York, NY, USA. https://doi.org/10.1145/2459976.2459982
17. Bajtoš T, Sokol P, Mézešová T. Multi-stage cyber-attacks detection in the industrial control systems. *Recent Developments on Industrial Control Systems Resilience*. Cham: Springer International Publishing; 2020:151-173. https://doi.org/10.1007/978-3-030-31328-9_8
18. Valdes A, Cheung S. Intrusion monitoring in process control systems. In: 42nd hawaii international conference on system sciences (hicss). IEEE IEEE; 2009:1-7.
19. Zhang Y, Wang L, Sun W, Green II RC, Alam M. Distributed intrusion detection system in a multi-layer network architecture of smart grids. *IEEE Trans Smart Grid*. 2011;2(4):796-808.
20. Barbosa RRR, Sadre R, Pras A. A first look into SCADA network traffic. In: 2012 ieee network operations and management symposium IEEE; 2012:518-521.
21. Barbosa RRR, Sadre R, Pras A. Difficulties in modeling SCADA traffic: a comparative analysis. In: Taft N, Ricciato F, eds. *Passive and Active Measurement*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012:126-135.
22. Igure VM, Laughter SA, Williams RD. Security issues in SCADA networks. *Comput Secur*. 2006;25(7):498-506.
23. Liu J, Xiao Y, Li S, Liang W, Chen CLP. Cyber security and privacy issues in smart grids. *IEEE Commun Surv Tutor*. 2012;14(4):981-997.
24. Ralston PAS, Graham JH, Hieb JL. Cyber security risk assessment for SCADA and DCS networks. *ISA Trans*. 2007;46(4):583-594.

25. Zhu B, Sastry S. SCADA-specific intrusion detection/prevention systems: a survey and taxonomy. In: Proceedings of the 1st workshop on secure control systems (scs), cpsweek 2010 Ptolemy TRUST; 2010.
26. Zhu B, Joseph A, Sastry S. A taxonomy of cyber attacks on SCADA systems. In: 2011 international conference on internet of things and 4th international conference on cyber, physical and social computing. IEEE IEEE; 2011:380-388.
27. Rubio JE, Alcaraz C, Roman R, Lopez J. Current cyber-defense trends in Industrial Control Systems. *Comput Secur*. 2019;87:101561.
28. Shapiro R, Bratus S, Rogers E, Smith S. Identifying vulnerabilities in SCADA systems via fuzz-testing. In: Critical infrastructure protection v. Springer Berlin Heidelberg Springer; 2011; Berlin, Heidelberg:57-72.
29. Devarajan G. Unraveling SCADA protocols: using Sulley Fuzzer. In: Defcon 15 hacking conference. DEF CON Communications, Inc. DEF CON; 2012; Las Vegas.
30. Hou Y, Such J, Rashid A. Understanding security requirements for industrial control system supply chains. In: 2019 ieee/acm 5th international workshop on software engineering for smart cyber-physical systems (sescps) IEEE; 2019:50-53.
31. Gasser O, Scheitle Q, Rudolph B, Denis C, Schricke N, Carle G. The amplification threat posed by publicly reachable BACnet devices. *J Cyber Secur Mobil, River Publ*. 2017;6(1):77-104. https://doi.org/10.13052/jcsm2245-1439.614
32. Talos Intelligence. New VPNFilter Malware Targets at least 500K Networking Devices Worldwide. Blog: https://blog.talosintelligence.com/2018/05/VPNFilter.html; 2018.
33. WIDE MAWI workinggroup, wide traffic archive. Web Archive: http://mawi.wide.ad.jp/
34. User timyardley. ICS security tools, tips, and trade. Git Repository: https://github.com/ITI/ICS-Security-Tools; 2018.
35. NTOP. nDPI. Open and extensible LGPLv3 deep packet inspection library. Website: https://www.ntop.org/products/deep-packet-inspection/ndpi/; 2018.
36. Mursch T. Quasi networks responds as we witness the Death of The Master Needler. Bad Packets Report, online: https://badpackets.net/quasi-networks-responds-as-we-witness-the-death-of-the-master-needler-80-82-65-66-for-now/; 2017.
37. CONPOT ics/scada honeypot, honeynet project. http://conpot.org/
38. Honeytrap transport layer honeypot, honeynet project. https://www.honeynet.org/project/Honeytrap
39. Klick J, Lau S, Wählisch M, Roth V. Towards better Internet citizenship: reducing the footprint of Internet-wide scans by topology aware prefix selection. In: Proc. of acm internet measurement conference (imc) ACM; 2016; New York:421-427.
40. Durumeric Z, Wustrow E, Halderman JA. ZMap: Fast Internet-wide scanning and its security applications. In: Proc. of the 22nd USENIX security symposium. USENIX Association USENIX Assoc.; 2013; Berkeley, CA, USA:605-620. https://www.usenix.org/conference/usenixsecurity13/technical-sessions/paper/durumeric
41. Poese I, Uhlig S, Kaafar MA, Donnet B, Gueye B. IP Geolocation Databases: Unreliable?. *SIGCOMM Comput Commun Rev (CCR)*. 2011;41(2):53-56.
42. Allen-Bradley. EtherNet/IP Secure Communication, 1201 South Second Street, Milwaukee, WI 53204-2496 USA, Rockwell Automation; 2015. https://literature.rockwellautomation.com/idc/groups/literature/documents/um/enet-um003_-en-p.pdf
43. PJM Interconnection. DNP SCADA over Internet with TLS security, 2750 Monroe Boulevard Audubon, PA 19403, PJM Interconnection; 2017. https://www.pjm.com/%2D/media/etools/jetstream/jetstream%2Dguide.ashx%3Fla%3Den
44. Schneider Electric USA. MODBUS/TCP security, Boston, MA, Schneider Electric USA; 2018. https://modbus.org/docs/MB-TCP-Security-v21_2018-07-24.pdf
45. Sheffer Y, Holz R, Saint-Andre P. Summarizing known attacks on transport layer security (TLS) and datagram TLS (DTLS). *RFC*. 7457, LLC 1000 N West Street, Suite 1200 Wilmington, DE 19801 USA., IETF; 2015.
46. Stevens J. Deploying Secure DNP3 (IEEE 1815)—what you need to know. *tech. rep*., 2840 Plaza Pl STE 205, Raleigh, NC 27612, USA, Triangle Microworks; 2016. https://trianglemicroworks.com/docs/default-source/referenced-documents/deploying-secure-dnp3-dtech-2016.pdf
47. Rosborough C, Gordon C, Waldron B. All about eve: comparing DNP3 secure authentication with standard security technologies for SCADA communications. In: Mipsycon 2019 Exelon and Schweitzer Engineering Laboratories, Inc; 2019; Minnesota, USA.
48. Ryba FJ, Orlinski M, Wählisch M, Rossow C, Schmidt TC. Amplification and DRDoS attack defense—a survey and new perspectives. *Technical Report*. arXiv:1505.07892, Open Archive: arXiv.org; 2015. http://arxiv.org/abs/1505.07892
49. Dusi M, Napolitano S, Niccolini S, Longo S. A closer look at Thin-Client connections: statistical application identification for QoE detection. *IEEE Commun Mag*. 2012;50(11):195-202. https://doi.org/10.1109/MCOM.2012.6353701
50. Velan P, Čermák M, Čeleda P, Drašar M. A survey of methods for encrypted traffic classification and analysis. *Int J Netw Manag*. 2015;25(5):355-374.
51. de Toledo TR, Torrisi NM. Encrypted DNP3 Traffic classification using supervised machine learning algorithms. *Mach Learn Knowl Extraction, Multidiscip Digit Publ Inst*. 2019;1(1):384-399.

AUTHOR BIOGRAPHIES

Marcin Nawrocki is a PhD student and research assistant at Freie Universität Berlin, advised by Matthias Wählisch. He supports the Internet Technologies Lab with a special focus on securing the Internet infrastructure. His research includes the longitudinal deployment of honeypots, inter-domain DDoS attack mitigation, as well as the security assessment of industrial control systems.

Thomas C. Schmidt is a professor of Computer Networks and Internet Technologies at Hamburg University of Applied Sciences (HAW), where he heads the Internet Technologies research group (iNET). Prior to moving to Hamburg, he was director of a scientific computer center in Berlin. He studied mathematics, physics, and German literature at Freie Universitaet Berlin and University of Maryland, and received his PhD from FU Berlin in 1993. Since then, he has continuously conducted numerous national and international research projects. He was the principal investigator in a number of EU, nationally funded, and industrial projects as well as visiting professor at the University of Reading, UK. His continued interests lie in the development, measurement, and analysis of large-scale distributed systems like the Internet or its offsprings. He serves as co-editor and technical expert in many occasions and is actively involved in the work of IETF and IRTF, where he co-chaired the SAM RG. Thomas is a co-founder and coordinator of the open source community developing the RIOT operating system—the friendly OS for the Internet of Things.

Matthias Wählisch is an assistant professor of Computer Science at Freie Universität Berlin and heads the Internet Technologies research group. After studying Computer Science and Contemporary German Literature, he received his diploma degree and his doctoral degree (with highest honors) in Computer Science from Freie Universität Berlin. His research and teaching focus on efficient, reliable, and secure Internet communication. This includes the design and evaluation of networking protocols and architectures, as well as Internet measurements and analysis. His efforts are driven by trying to improve Internet communication based on sound research. Matthias is the PI of several national and international projects. Since 2005, he is active within the Internet standardization (IETF/IRTF). His research results have been distinguished multiple times. Among others, he was a winning team member of the IPv6 Application Contest 2009 and received the Young Talents Award of Leibniz-Kolleg Potsdam for outstanding achievements in advancing the Internet. He received the Best of ACM SIGCOMM CCR Award in 2018 and 2019. He co-founded some successful open source projects in the context of Internet of Things (RIOT) and secure Internet routing (e.g., RTRlib), where he is still responsible for the strategic development.

How to cite this article: Nawrocki M, Schmidt TC, Wählisch M. Industrial control protocols in the Internet core: Dismantling operational practices. *Int J Network Mgmt*. 2022;32(1):e2158. doi:10.1002/nem.2158