

Received September 9, 2019, accepted October 2, 2019, date of publication October 21, 2019, date of current version November 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2948793

Trends and Detection Avoidance of Internet-Connected Industrial Control Systems

DAVID HASSELQUIST¹, ABHIMANYU RAWAT², AND ANDREI GURTOV¹²

¹Linköping University, 581 83 Linköping, Sweden

²ADIT-IDA, Linköping University, 581 83 Linköping, Sweden

Corresponding author: Andrei Gurtov (andrei.gurtov@liu.se)

This work was supported by the CENIIT under Project 17.01.


ABSTRACT The search engine Shodan crawls the Internet for, among other things, Industrial Control Systems (ICS). ICS are devices used to operate and automate industrial processes. Due to the increasing popularity of the Internet, these devices are getting more and more connected to the Internet. These devices will, if not hidden, be shown on Shodan. This study uses Shodan, together with data found by other researchers to plot the trends of these ICS devices. The studied trends focus on the country percentage distribution and the usage of ICS protocols. The results show that all studied countries, except the United States, have decreased their percentage of world total ICS devices. We suspect that this does not represent the real story, as companies are getting better at hiding their devices from online crawlers. Our results also show that the usage of old ICS protocols is increasing. One of the explanations is that industrial devices, running old communication protocols, are increasingly getting connected to the Internet. In addition to the trend study, we evaluate Shodan by studying the time it takes for Shodan to index one of our devices on several networks. We also study ways of avoiding detection by Shodan and show that, by using a method called port knocking, it is relatively easy for a device to hide from Shodan, but remain accessible for legitimate users.

INDEX TERMS Trends, Shodan, avoidance, industrial control systems, ICS, SCADA, security, Internet of Things, IoT.

I. INTRODUCTION

Shodan is a search engine for devices connected to the Internet [1]. While a standard search engine, such as Google or Bing, indexes only content on the web, Shodan indexes all kind of devices such as Industrial Control Systems (ICS), web cameras and refrigerators. Shodan is publicly available and can be used as a tool for detecting vulnerable devices, which in turn can be exploited if the user has malicious intent. In the same way, it can also be used by a system administrator as a helpful tool for improving network security.

Since devices found on the Internet can be reached from anywhere by anyone, the risk of attacks that exploit vulnerable devices increases. The consequences also differ depending on what kind of device is subjected to a successful attack. If an ICS device is the subject of a successful attack, the consequences could be critical. Many ICS devices run on outdated protocols that were never intended to be connected to the Internet, which makes them especially

The associate editor coordinating the review of this manuscript and approving it for publication was Yassine Maleh .

vulnerable [2], [3]. Because ICS devices are both commonly vulnerable and could have critical consequences if subjected to successful attacks, this type of device is the focus of this paper.

This paper introduces the search engine Shodan, ICS and common protocols in use with these. We study the trends of the number of ICS devices and ICS protocols, both in Sweden and worldwide. This is done by comparing Shodan search results collected over time, together with results gathered by other studies using the same search queries. The following trends are in focus of our study:

- Number of ICS devices in Sweden
- Number of ICS devices worldwide
- Country percentage of ICS devices in the world
- Device usage of ICS protocols in Sweden
- Device usage of ICS protocols worldwide

We show that, according to Shodan, the total number of ICS devices are fluctuating, but have remained the same for the last few years. For the studied countries, all countries except the United States have decreased their percentage of

world total ICS devices. Since the world is getting more connected and industrialized, we believe the number presented by Shodan does no longer represent the whole picture and that the real number is increasing. This is because the security awareness in the world is increasing, and many companies are enhancing their infrastructure and limiting their device exposure. We support our claim by showing one simple method that can be used to avoid detection from the Shodan crawlers. We also evaluate Shodan by studying the time Shodan takes to index different networks. Also, due to the increasing popularity of IPv6 and the difficulties of scanning the IPv6 network range, devices might be harder to detect. The trend of decreased exposure of ICS devices will most likely continue as IPv6, and the internet security field is getting more attention.

Furthermore, we compare the country percentage of ICS devices to the GDP share of these countries and find that there is a correlation between a country's percentage of ICS devices and their GDP value. A high GDP value often corresponds to a high number of ICS devices, with a few exceptions such as for China. These countries might be blocking parts of the internet or using an experimental internet for their industrial devices that are separated from the public internet.

The rest of the paper is organized as follows. Section II presents related work and Section III covers the background theory of Shodan, port knocking and ICS. Section IV covers the data collection and Shodan experiment setup together with a methodology evaluation. These results are later presented in Section V and discussed in Section VI. Lastly, the conclusions and future work are presented in Section VII.

II. RELATED WORK

ICS devices have been in use for more than two decades now, and their running software did not have the security as the highest priority in the first place. With the advent of millions of Internet of Things (IoT) devices out there and widespread internet connectivity, IoT devices have an open attack surface. Protocols related attacks have been prevalent in the internet history since the protocol standards are widespread and mostly open implementations [4].

In the article by Abe *et al.* [5], a honeypot based scheme, has been devised which enables security administrators to actively monitor the attacks over the end services. The proposed scheme maintains the listings of valid and invalid actions a user can perform and checks if the user has queried for invalid action, then it is reported. The method scrutinizes packet payload to verify user actions over the Modbus protocol. Honeypots in general are expensive [6] in the real world deployments, hence less common in the IoT ecosystem.

Similar protocol-based attacks are addressed in work by Hong *et al.* [7]. Whitelist-based attack detection technology has been discussed where the data manipulation in the protocol packets can result in system anomalies. It showcases the classical case of replay attacks based on rudimentary protocols which have not been developed actively to cope up with

the advanced attacks. The paper concludes with a machine learning based solution to handle replay based attacks.

A study by Scarfo [8] briefly indicates the connection between the developed economies and the IoT business by the year 2025. The study also highlights the attack cost, which is levied over the businesses/countries. Latest attacks such as Ransomware not only cripple the businesses but also impact the countrywide economy. The study still lacks to provide a holistic view concerning the businesses and the respective countries.

In work by Lee *et al.* [9] a detection method which runs against the popular IP device search engines like Shodan and Censys has been devised. A stateful detection approach has been used with the rules such as SYN, banner grabbing, and their combination. While these IP device search engines have not openly published their scanning methods, it's another attempt to do so. The approach is to detect and not to prevent the scanning of IP devices, to counter the malicious attacker's prevention is required.

III. BACKGROUND

This section contains the necessary background information about Shodan, port knocking, and ICS that is needed for the scope of this paper.

A. SHODAN

Shodan is a search engine that can be used to search for devices connected to the Internet [1]. These devices can be anything from a router to an ICS. Shodan also provides an option to apply a variety of filters to make searches more specific. These filters include country, port, product, and category, among others. An example of a search query is `port:502 country:"se" category:ics`, which will search for all ICS devices in Sweden using port number 502.

The way Shodan works is by having servers located all over the world that crawl the Internet for accessible devices all hours around the clock [10]. The crawlers generate a random IPv4 address and a random port before invoking that specific service for banner information [10].

The banner that the crawler grabs contain information in the form of text that describes a service on a device. Along with the banner, Shodan also collects information about the device's geographical location, hostname, operating system and more [10]. Those devices that are identified by the crawlers and the information gathered regarding them can be seen by using the Shodan API or the Shodan website [10], [11]. A search made with the API is more detailed and contains, among other things, the last detection time of each service independently. With the API, it is also possible to request a scan to be made, e.g., on a specific IP-address.

B. PORT KNOCKING

Port knocking is a method that can be used to access a server that has no open ports [12], [13]. It is a form of one way host-to-host communication where information is flowing to closed ports at the end device. The server uses a monitoring

daemon to intercept the traffic on the closed ports and will silently process the data. There are variants of port knocking methods, using different data requirements and sequence on the data received. These requirements can be seen as a secret. If the requirements are fulfilled, the server will modify the firewall rules and allow the IP-address of the knocking user to connect to the server [12], [13]. Later, another knocking sequence can be used to close the port, or the port knocking can be configured to close after a certain amount of time has passed automatically.

C. INDUSTRIAL CONTROL SYSTEMS

An Industrial Control System (ICS) is a general term for different types of control systems used to operate and/or automate industrial processes. An ICS often consists of combinations of control components that work together to achieve some sort of industrial objective. The most common type of ICS are *Supervisory Control and Data Acquisition* (SCADA) systems [14].

SCADA systems capabilities are focused on control at a supervisory level, and the systems are composed of devices that are distributed in various locations. These devices are often programmable logic controllers (PLC). The main purpose of using SCADA is for control of field sites through a centralized control system and for monitoring. Example uses of SCADA can be monitoring a local environment for alarm conditions or performing local operations such as opening or closing of valves through field devices [14].

D. ICS PROTOCOLS

There are many ICS protocols using different methods and ports for communication with other interconnected devices [2]. Different protocols are used for different purposes. This section gives a brief overview of the common ICS protocols. There are many other popular ICS protocols, such as Ethernet/IP, OMRON FINS, Mitsubishi MELSEC-Q, and Automated Tank Gauge. However, these are outside the scope of this study. Table 1 shows an overview of the presented protocols and their ports used for communication.

TABLE 1. ICS protocols and ports.

Protocol	Port
Modbus	502
MQTT	1883
Niagara Fox	1911, 4911
BACnet	47808
DNP3	20000

1) MODBUS

Modbus was created in 1979 and is used for serial communication with PLC devices [15]. It has become widely adopted as a de-facto industrial standard in the ICS world. Since Modbus is very old and was invented before the Internet became widely used, it was not built with security in mind. This means that there is no built-in encryption, integrity checks

or authentication [15], [16]. There have been many attacks performed against this protocol. Huitsing *et al.* [17] identified 59 attack instances and 20 distinct attacks performed. Furthermore, they identify over 100 attack instances performed on variations of the Modbus protocols. Some typical attacks that Modbus is susceptible to are man-in-the-middle, spoofing, and replay attacks.

2) MQTT

Message Queuing Telemetry Transport (MQTT) is a lightweight communication protocol designed to be used by devices having low bandwidth and high latency on unreliable networks [18]. It was invented in 1999 and had since been widely adopted among the industry. MQTT uses a publish/subscribe architecture, meaning that ICS devices that implement this protocol can send and publish data to a specific server. Clients who have subscribed to this device can then be notified by the server when new data is available. To keep the protocol lightweight and simple, encryption is not built in and handled by the protocol itself. There is only an optional authentication by a user name and password, which is passed together with the MQTT packet [18].

3) NIAGARA FOX

Niagara Fox protocol is most commonly used in building automation systems such as offices, libraries, and universities [2]. It is a part of the Niagara framework from Tridium and is also known as Tridium Fox. Niagara AX systems often use it for communication between the different stations and the central machine. Niagara Fox supports the use of SSL for secure communications between these devices. The security model of Niagara Fox resembles the classical mandatory access control and is based on the concept of Users, Permissions, and Categories [19]. The protected objects are grouped into categories, and users are given a set of permissions in each category.

4) BACnet

Building Automation and Control Networks (BACnet) is a communication protocol used for building automation and control networks. The development began in 1987 and has since then become protocol standard. BACnet is used to provide control automation in buildings, such as ventilation, fire detection system, heating, and access control [2].

5) DNP3

The communication protocol DNP3 was developed in 1993 and is now an IEEE standard and a commonly used protocol for ICS devices. It is often used between automated devices and implemented in many critical infrastructure applications, such as the electricity sector [3], [16]. Similar to the Modbus protocol, it is designed to be reliable and does not enforce security such as encryption, integrity checks, or authentication. To achieve security with this protocol, additional security-specific hardware or security features in the application have to be added [16]. Compared to Modbus,

DNP3 is more robust, efficient, and interoperable. However, this comes as a cost of higher complexity.

IV. METHODOLOGY

This section covers how searching in Shodan was performed, the search queries used, the Shodan auto indexing and avoidance experiment and an evaluation of the methodology.

A. SEARCH USING SHODAN

Searching with the Shodan search engine was performed by entering the search queries into the free text field of Shodan’s web interface. The search queries that were used can be seen in Table 2.

TABLE 2. Shodan search queries.

Target results	Search query
ICS devices worldwide	category:ics
ICS devices country X	category:ics country:"X"
Modbus worldwide	port:502 category:ics
Modbus in Sweden	port:502 country:"se" category:ics
BACnet worldwide	port:47808 category:ics
BACnet in Sweden	port:47808 country:"se" category:ics
Niagara Fox worldwide	port:1911,4911 product:Niagara category:ics
Niagara Fox in Sweden	port:1911,4911 product:Niagara country:"se" category:ics
DNP3 worldwide	port:20000 category:ics
DNP3 in Sweden	port:20000 country:"se" category:ics
MQTT in Sweden	port:1883 country:"se"

B. SHODAN AUTO INDEXING

To evaluate Shodan, a study has been done to see which networks that are indexed, the time it takes for Shodan to index these networks and the requests that Shodan makes to these networks. Three devices have been used for this experiment: a Raspberry Pi 3B+, a 4G-router, and a LAN-router. The 4G router was also able to downgrade itself and force the network to use either only 2G or 3G. Both routers were set to reply on ICMP echo requests (pings) and had port forwarding for port 80 set to the router’s local IP-address of the Raspberry Pi. On the Raspberry Pi, a web server was running Python SimpleHTTPServer containing only a simple HTML web page shown in Figure 1. Four different networks were used: 2G, 3G, 4G, and a residence network using Bahnhof, a popular Internet Service Provider (ISP) in Sweden.

In order to host a public web server, a public IP is required. Bahnhof partly uses Carrier-grade NAT in their network, resulting in residence networks getting a private Wide Area Network (WAN) IP-address. We have noticed that this is a typical case in Sweden [20], most likely due to the IPv4 address exhaustion. After contacting the ISP, the residence network was given a public IP-address, making it possible to access the web server from the Internet. To detect when Shodan indexes the web server, Selenium was used to crawl the Shodan search results once every hour. The time of

```
<html>
<header>
  <title>Shodan test</title>
</header>
<body>
  <h1>Hello World!</h1>
  <h2>Shodan test</h2>
</body>
</html>
```

FIGURE 1. HTML code of the web page.

the indexing, as well as the requests being sent to the web server at that time, were recorded.

C. SHODAN AVOIDANCE

To evaluate our results, we studied ways to avoid detection by the Shodan indexing. The simplest and most straight forward way was found to be port knocking. In addition to the SimpleHTTPServer web server running, SSH connection to the device using the default port of 22 was enabled. The avoidance setup was done by using Shodan on-demand scanning and the same hardware as the Shodan auto indexing, described in Section IV-B. In our setup, the following three random ports were chosen as the opening knock sequence for the SSH port: 7336, 8710 and 8905. The user had to send packets with SYN flag set to these port in the correct sequence under a total of five seconds. If this were done, the server would grant this specific host IP SSH access over port 22. For this setup, the ports 22, 80, 7336, 8710 and 8905 were forwarded from the router to the device.

The Shodan API had to be used since the web interface does not show the last detection time for each service port independently. The methodology for the avoidance study consists of six on-demand scans. These scans, together with the web server and SSH status, are shown in Table 3. The web server was used to verify that Shodan had performed a scan on the IP-address and was therefore always active.

The first on-demand scan verified that Shodan detected both our web server and the SSH port. Later, before the second scan, the SSH port was turned off. This was done to verify that Shodan updates the latest detection time on the web server, however, does not do this on the SSH port, meaning that it had not detected the SSH port. Before the third scan, the SSH port was turned on again, verifying that Shodan

TABLE 3. Shodan on-demand scans.

Scan	Web server	SSH
1	Active	Active
2	Active	Inactive
3	Active	Active
4	Active	Active with port knocking
5	Active	Active with port knocking
6	Active	Active with port knocking

detects both the web server and SSH port again, updating their last detection time. Lastly, SSH was kept on but is now protected by port knocking and only accessible to users knowing the knock secret. Scans 4-6 are to verify that Shodan has scanned our network, but only detected the web server.

D. EVALUATION OF METHODOLOGY

The service chosen for the Shodan auto indexing experiment was HTTP over port 80. However, Industrial Control Systems often use other services than HTTP to communicate and transfer data. This choice could be criticized as the detection time by Shodan could vary depending on which port and service being used. The experiment could be improved by using a real ICS device together with an ICS protocol, such as Modbus. If Shodan prioritizes some ports or services, this would most likely have an impact on the detection time.

In the case of Shodan auto indexing, Selenium was used to periodically crawl the Shodan web search results to see if Shodan had indexed our device. We used Selenium instead of other tools such as *curl* to simulate user interaction. However, if Shodan would record these interactions and queries and base their crawling algorithm on user searches, this could have an impact on our results. One alternative to using the web interface of Shodan is to use the Shodan API. However, using the API does not exclude the possibility that Shodan dynamically adjusts their crawling algorithm based on user queries.

In order to improve the reliability of our Shodan avoidance experiment, more searches and scans, especially on several networks using several devices, could have been made to confirm better the results that were collected. It would also have been interesting to find some other method for avoiding detection by the Shodan crawlers and comparing it to port knocking, both in terms of effectiveness and ease of use.

V. RESULTS AND COLLECTED DATA

This section contains the results and trends that were collected from other studies combined with our data collection. ICS data from 2013 have been collected from a research project at Aalto University [21], [22]. ICS data from 2015 and 2016 have been collected from a recent study regarding an Internet-wide view of ICS devices [23]. ICS data from 2016 have been gathered from a study of ICS [24]. ICS data from 2017, 2018, and 2019 comes from our research department at the University, and the data from 2017 have also been used in a previous study [25]. GDP values have been collected from the World Bank Group [26]. Lastly, in this section, the results from the Shodan auto indexing and avoidance experiment are presented.

A. ICS COUNTRY PERCENTAGE TRENDS

This section covers the ICS country trends that can be seen from the searches performed with the Shodan search engine and the data from other studies. Figure 2, 3 and 4 shows, for the selected countries in the study, the country's percentage of the total ICS devices worldwide over the years 2013 to 2019.

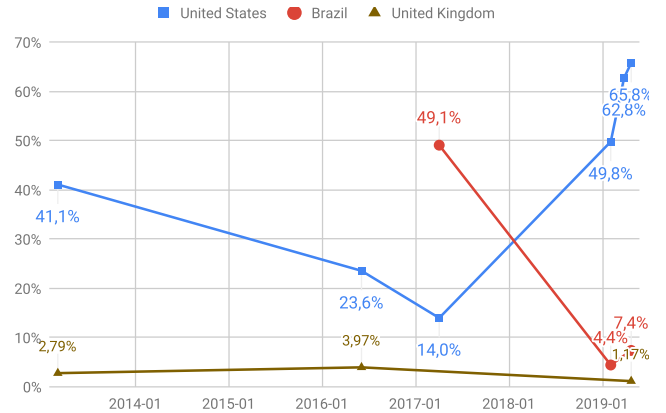


FIGURE 2. US, Brazil and UK percentage of ICS devices worldwide.

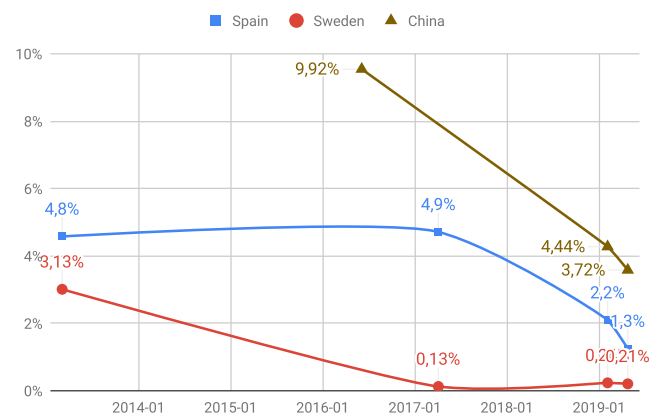


FIGURE 3. Spain, Sweden and China percentage of ICS devices worldwide.

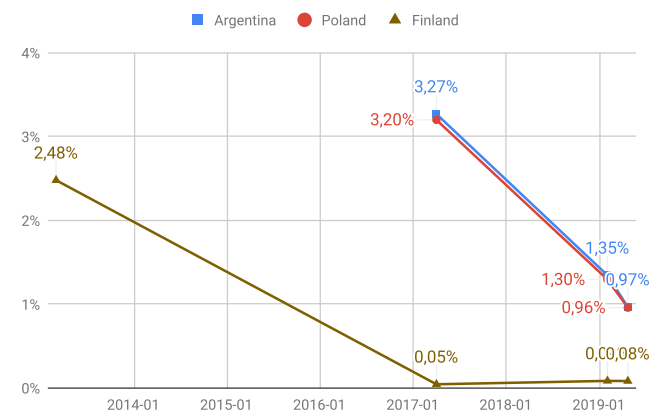


FIGURE 4. Argentina, Poland and Finland percentage of ICS devices worldwide.

An increase in the percentage of world ICS devices can be seen for the United States since 2013, however with a large decrease near 2017, while a decrease can be seen for the rest of the countries. The total number of ICS devices worldwide and in Sweden can be seen in Table 4. Here, it is shown that the number of ICS devices worldwide was quite stable in the last few years, while in Sweden it has almost doubled. When looking at the last few months, some fluctuations can

TABLE 4. Number of ICS devices.

Date	Worldwide	Sweden
2017-04	2,280,652	2,965
2019-02	2,276,259	5,539
2019-03	2,545,414	5,311
2019-04	2,702,311	5,726
2019-06	2,665,804	7,292

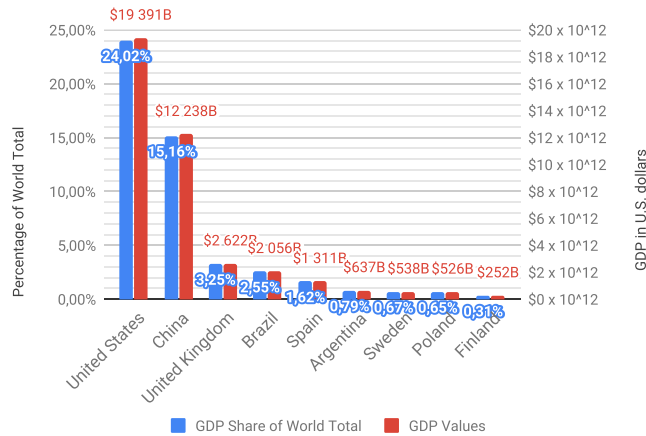


FIGURE 5. 2017 GDP values and ICS percentage.

be observed. However, the increasing trend of ICS devices is clear.

A comparison between the number of ICS devices for a country and that country’s GDP is also interesting to explore. Therefore the 2017 GDP for the study’s selected countries can be seen in Figure 5, both in the percentage of world total GDP and value in U.S. dollars. As seen in Figure 5 compared to Figure 2, 3 and 4, the countries with the most ICS devices are generally the ones with the highest GDP. The country that significantly stands out from this is China. China has the second highest GDP value, with over 15% of total GDP worldwide. However, China only stands for a few percentages of the ICS devices.

B. ICS PROTOCOL TRENDS

Figure 6 shows how many devices in Sweden which used the protocols Modbus, MQTT, BACnet, Niagara Fox, and DNP3 over the years 2017 to 2019. Usage of all these protocols has increased since 2017, especially MQTT, which has almost doubled in usage. The number of ICS devices which uses the DNP3 protocol worldwide has significantly increased since 2018 and can be seen in Figure 7. Total number of ICS devices worldwide that used the protocols Modbus, BACnet, and Niagara fox between the years 2016 and 2019 can be seen in Figure 8. All these protocols have seen an increase in usage as well, but not as significant as the DNP3 protocol.

C. SHODAN AUTO INDEXING

The time taken for Shodan to index the web server on different networks is shown in Table 5. At the time of the Shodan indexing, we also recorded the requests being made to our

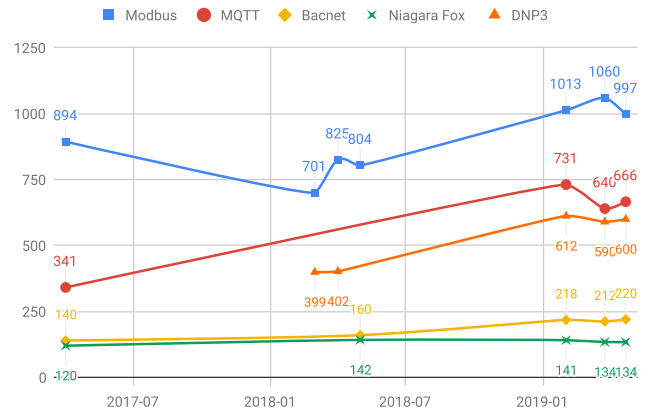


FIGURE 6. Device usage of ICS protocols in Sweden.

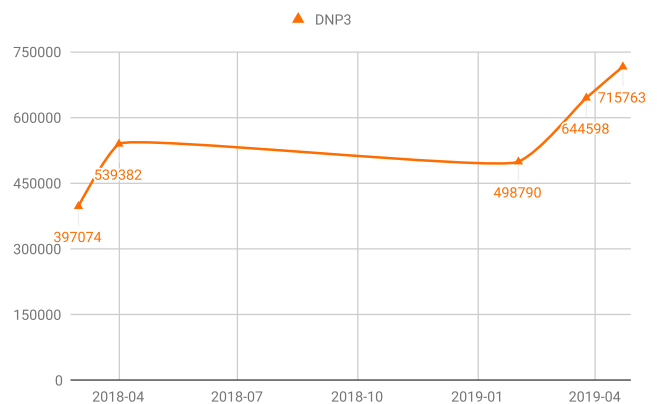


FIGURE 7. Device usage of DNP3 protocol worldwide.

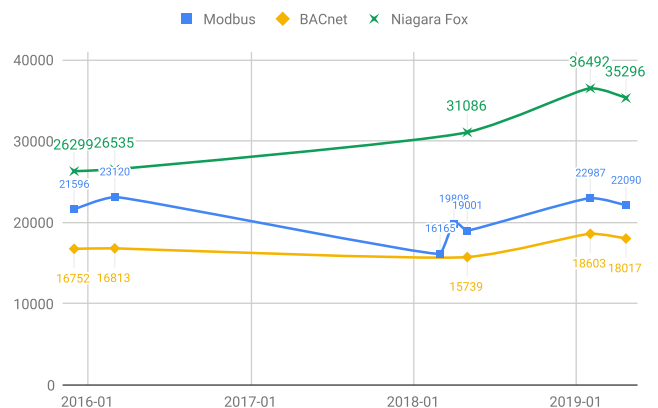


FIGURE 8. Device usage of ICS protocols worldwide.

server. The IP-address of Shodan is shown in Table 5. For each detection, five HTTP GET requests were made to our server, requesting the index page, robots.txt, sitemap.xml, .well-known/security.txt and favicon.ico. The HTTP requests were identical for both the 3G and 4G networks. For the 2G and the residence network, Shodan did not manage to index the device under seven days.

D. SHODAN AVOIDANCE

As can be seen in Table 3, six on-demand-scans were made. Every scan verified the expected behavior. The first scan

TABLE 5. Shodan indexing.

Network	Network IP	Shodan IP	Time
2G	80.170.92.xxx	-	> 7 days
3G	176.68.52.xxx	89.248.167.131	6 hours
4G	5.241.58.xxx	89.248.172.16	23 hours
Residence	81.170.152.xxx	-	> 7 days

detected both the web server and the SSH port. The second scan detected only the web server. The third scan detected, just like the first scan, both the server and the SSH port. Later, SSH was kept on but hidden with port knocking. The scans 4-6 showed that Shodan failed to index the SSH port, even with three attempts. This shows us that port knocking can be used as a way to avoid detection by Shodan.

VI. DISCUSSION

This section analyzes the collected data and results regarding ICS trends, and Shodan auto indexing.

A. ICS TRENDS

As seen in the results, most countries have a lower percentage of the total ICS devices in the world in their latest data collection point compared to their first data collection point. The exception is the United States. The industry in all of these countries is still growing, especially in China and Brazil, which are rapidly increasing. More industry generally leads to more ICS devices, but the results gathered in this study do not show this. This could be explained by that the knowledge about limiting device exposure are increasing, and that companies using ICS devices are getting better at hiding them from Shodan and similar services. The exception of the United States could be explained either by that their industry is growing so fast that they can not keep up with hiding all the devices, or that they simply do not prioritize hiding their devices because they have implemented some other kind of security in them.

There are multiple ways of avoiding detection by Shodan, and in this study, we have given an example of an easy way of doing so. If it is the case that companies are using this, or similar techniques for hiding from Shodan, the real number of ICS devices in the studied countries could very well be increasing, which indicates that the results gathered from Shodan could be misleading. That may further question Shodan's reliability and usefulness in such searches and cases. An alternative explanation to companies getting better at hiding their devices from Shodan is a country-wide blocking. However, according to Matherly [10], this is not the case since Shodan prevents this by placing their servers in different parts of the world.

China's GDP value stands out from the rest of the studied countries compared to the country's total number of ICS devices. The reason for this could be that China has a much larger population than any of the other countries, which might affect the result. It would have been interesting to compare the number of ICS devices of the countries with GDP per capita and see if China still would stand out from the rest

of the countries. Another reason could be that companies in China are simply better at hiding their devices from Shodan compared to other countries.

With regards to the total number of ICS devices in the world, there has not been much of a change between 2017 and February 2019. We believe that this does not represent the real number. The industry has expanded in these years and so should the number of ICS devices. One explanation for this finding is that the number of ICS devices is increasing, but more devices are being hidden from detection. The results also show that from February 2019 to April 2019, a steady increase of around 400 000 devices, almost a 15 percent increase, has occurred. The reason for this could be that the number of active ICS devices depends on seasonal changes. For example, more devices could be activated and detected by Shodan during the summer and then deactivated and not be detected by Shodan during the winter.

B. SHODAN AUTO INDEXING

The speed of crawling the Internet depends mainly on the network connection and software implementation. Durumeric *et al.* [27] introduce a network scanner called ZMap, specifically designed to perform Internet-wide scans and capable of scanning a port on the entire IPv4 address space under 45 minutes from a single machine, approaching the theoretical maximum speed of gigabit Ethernet. We believe, therefore, that it is reasonable that Shodan was able to index two of our devices under 24 hours. The question that arises is why the 2G and the residence network was unable to get indexed in under seven days, despite having the same setup as the indexed networks. Perhaps this could be because Shodan recognizes it to be a honeypot and does not index the device. According to Shodan, honeypots are detected and still listed but given a tag [10], so this does not seem reasonable. Bodenheimer *et al.* [11] performed a similar study, studying the time Shodan required to successfully identify their PLC devices and found that it took up to 19 days for Shodan to successfully index and identify their four devices. According to Shodan themselves, the entire Internet is being crawled at least once a month [28], which have led us to believe that in some cases, Shodan requires more than seven days to index the device. This means that all of our devices would have most likely been indexed if they had stayed on for longer.

VII. CONCLUSION

According to Shodan, the number of ICS devices connected to the Internet have remained roughly the same for the last few years. Out of all studied countries, only the United States have increased their percentage of world total ICS devices. As we are moving towards a more connected industry, the total number of ICS devices should be increasing. Our results do not show this, and the reason for this may be because many countries and organizations are hiding their ICS devices from search engines like Shodan. The security awareness in the world has increased for the last few years, and we believe

that many organizations are enhancing their infrastructure and limiting their device exposure. We support our claim by evaluating Shodan and showing that a simple method such as port knocking can be used to avoid detection from Shodan.

This is an IoT era, and as indicated by the trends in the graphs, there may be a possibility that countries like China and Russia are hiding their devices majorly due to the security concerns as their deployment is directly into the country's infrastructure. To reason this possibility, there is "great Chinese firewall" which we are aware used for internet censorship already. Meanwhile, Russia is already experimenting with its Intranet, i.e., disconnection itself from the legacy Internet¹, this Intranet can help these devices going off the radar of IP device search engines. For countries like the USA where the IP devices are on a surge can be interpreted as economic strength in the developed nations. These nations have adequate funding to carry the IoT research on a large scale, and with US universities, it can be very much possible that such research devices might be getting detected. IoT devices using the IPv6 also have a higher chance of getting undetected as through random selection IPv6 range is quite extensive.

Many ICS devices connected to the Internet are still using old protocols, and the popularity of all protocols studied in this study has increased. This is likely because old devices that are running these protocols are getting connected. As this study has shown, there are simple and effective methods of blocking Shodan from detecting a device. If this, or similar practices, were to be used more, the usefulness of Shodan might decrease.

For future work, it would be interesting to continue with the ICS trends, and see the future development over the coming years for the number of ICS devices, both for specific countries and worldwide, as well as the protocols studied. This could give a better inclination to whether our theory about if Shodan still is a reliable device search engine is true or not. It would also be interesting to look at Shodan's historical data, provided by Shodan to users with Enterprise Data License, and compare those to the sources we have used in this study. With the historical data, a more thorough analysis could be made, and more clear trends might be seen.

ACKNOWLEDGMENT

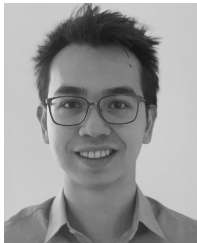
The authors would like to thank Andreas Lundquist at Linköping University for his help to carry out this study and other researchers at Linköping University who gathered data and conducted a similar study over past two years, making this paper possible.

REFERENCES

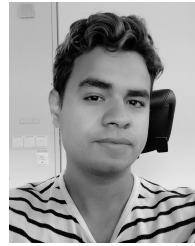
- [1] Shodan. *Shodan*. [Online]. Available: <https://www.shodan.io/>
- [2] Shodan. *Industrial Control Systems*. [Online]. Available: <https://www.shodan.io/explore/category/industrial-control-systems>
- [3] I. A. Siddavatam and F. Kazi, "Security assessment framework for cyber physical systems: A case-study of DNP3 protocol," in *Proc. IEEE Bombay Sect. Symp.*, Sep. 2015, pp. 1–6.
- [4] A. Wright, "Mapping the Internet of Things," *Commun. ACM*, vol. 60, no. 1, pp. 16–18, Jan. 2017.
- [5] S. Abe, Y. Tanaka, Y. Uchida, and S. Horata, "Developing deception network system with traceback honeypot in ICS network," *SICE J. Control, Meas., Syst. Integr.*, vol. 11, no. 4, pp. 372–379, 2018.
- [6] K. Li, J. You, H. Wen, H. Li, and L. Sun, "Collaborative intelligence analysis for industrial control systems threat profiling," in *Proc. Future Technol. Conf.* Cham, Switzerland: Springer, 2018, pp. 94–106.
- [7] K.-S. Hong, H.-B. Kim, D.-H. Kim, and J.-T. Seo, "Detection of replay attack traffic in ICS network," in *Proc. Int. Conf. Appl. Comput. Inf. Technol.* Cham, Switzerland: Springer, 2018, pp. 124–136.
- [8] A. Scarfò, "The cyber security challenges in the IoT era," in *Security and Resilience in Intelligent Data-Centric Systems and Communication Networks*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 53–76.
- [9] S. Lee, S.-H. Shin, and B.-H. Roh, "Abnormal behavior-based detection of shodan and census-like scanning," in *Proc. IEEE 9th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2017, pp. 1048–1052.
- [10] J. Matherly, *The Complete Guide to Shodan*. Leanpub, 2017.
- [11] R. Bodenheimer, J. Butts, S. Dunlap, and B. Mullins, "Evaluation of the ability of the shodan search engine to identify Internet-facing industrial control devices," *Int. J. Crit. Infrastruct. Protection*, vol. 7, no. 2, pp. 114–123, 2014.
- [12] M. Krzywinski, "Port knocking: Network authentication across closed ports," *SysAdmin Mag.*, vol. 12, no. 6, 2003.
- [13] J. Vinet. *Knockd—A Port-Knocking Server*. [Online]. Available: <http://www.zeroflux.org/projects/knock>
- [14] A. Gurtov, M. Liyanage, and D. Korzun, "Secure communication and data processing challenges in the industrial Internet," *Baltic J. Mod. Comput.*, vol. 4, no. 4, pp. 1058–1073, 2016.
- [15] M. K. Ferst, H. F. M. de Figueiredo, G. Denardin, and J. Lopes, "Implementation of secure communication with modbus and transport layer security protocols," in *Proc. 13th IEEE Int. Conf. Ind. Appl.*, Nov. 2018, pp. 155–162.
- [16] INCIBE. (2015). *Protocols and Network Security in ICS Infrastructures*. Spanish National Cybersecurity Institute. [Online]. Available: https://www.incibe.es/extfrontinteco/img/File/intecocert/ManualesGuias/incibe_protocol_net_security_ics.pdf
- [17] P. Huitsing, R. Chandia, M. Papa, and S. Shenoi, "Attack taxonomies for the modbus protocols," *Int. J. Crit. Infrastruct. Protection*, vol. 1, pp. 37–44, Dec. 2008.
- [18] MQTT. *Frequently Asked Questions*. [Online]. Available: <http://mqtt.org/faq>
- [19] Tridium. (2005). *Niagara Security Overview*. [Online]. Available: https://www.vykon.com/library/white-papers/Niagara_AX_Security_Overview.pdf
- [20] CyberInfo. (2018). *Fortfarande låg IPv6-Utbredning i Sverige*. [Online]. Available: <https://www.cyberinfo.se/arkiv/fortfarande-lag-ipv6-utbredning-i-sverige/>
- [21] S. Tiilikainen and J. Manner. (2013). *Suomen Automaatioverkköjen Haavoittuvuus*. Aalto University. [Online]. Available: <https://research.comnet.aalto.fi/public/Aalto-Shodan-Raportti-julkinen.pdf>
- [22] S. Tiilikainen, "Improving the national cyber-security by finding vulnerable industrial control systems from the Internet," M.S. thesis, School Elect. Eng., Aalto Univ., Espoo, Finland, 2014. [Online]. Available: https://aaltodoc.aalto.fi/bitstream/handle/123456789/12918/master_Tiilikainen_Seppo_2014.pdf
- [23] A. Mirian, Z. Ma, D. Adrian, M. Tischer, T. Chuenchujit, T. Yardley, R. Berthier, J. Mason, Z. Durumeric, J. A. Halderman, and M. Bailey, "An Internet-wide view of ICS devices," in *Proc. 14th Annu. Conf. Privacy, Secur. Trust (PST)*, Dec. 2016, pp. 96–103.
- [24] B. Wang, X. Li, L. P. de Aguiar, D. S. Menasche, and Z. Shafiq, "Characterizing and modeling patching practices of industrial control systems," in *Proc. ACM Meas. Anal. Comput. Syst.*, 2017, Art. no. 18.
- [25] A. Hansson, M. Khodari, and A. Gurtov, "Analyzing Internet-connected industrial equipment," in *Proc. Int. Conf. Signals Syst.*, 2018, pp. 29–35.
- [26] World Bank Group. *GDP (Current US Dollars)*. [Online]. Available: <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?end=2017&start=1960&view=chart>

¹https://www.theregister.co.uk/2019/02/12/russia_disconnect_internet_intranet/

- [27] Z. Durumeric, E. Wustrow, and J. A. Halderman, "ZMap: Fast Internet-wide scanning and its security applications," in *Proc. 22nd USENIX Conf. Secur.*, 2013, pp. 605–620.
- [28] Shodan. (2018). *On-Demand Scanning*. [Online]. Available: <https://help.shodan.io/the-basics/on-demand-scanning>



DAVID HASSELQUIST received the B.S. degree in computer science and engineering from Linköping University, Sweden, in 2018, where he is currently pursuing the M.S. degree. He is also an Assistant Lecturer with the Division for Software and Systems, Department of Computer and Information Science, Linköping University, where his main teaching subject is programming. His main research interests include networks and secure systems.



ABHIMANYU RAWAT received the master's degree in computer science from the Birla Institute of Technology and Science, Pilani Campus, in 2017. He started his career as a Protocols Engineer with DellEMC working on SMB protocol design and development. He spent a brief time at Linköping University, Sweden, where he assisted Prof. A. Gurtov. His main research interests include distributed systems, security, and networks.



ANDREI GURTOV received the M.Sc. and Ph.D. degrees in computer science from the University of Helsinki, Finland, in 2000 and 2004, respectively.

He was with the University of Oulu for three years and with Aalto University for six years and visited the International Computer Science Institute, Berkeley, multiple times. He is currently a Professor of computer science with Linköping University, Sweden. He has coauthored over 200 publications, including four books, five IETF RFCs, and over 60 journals and 110 conference papers. He holds six patents. He supervised 15 Ph.D. theses. His research interests include network protocols, security of vehicular, airborne, industrial systems, mobile, wireless and the IoT networks, and smartgrids. He is an ACM Distinguished Scientist, the IEEE ComSoc Distinguished Lecturer, and the Vice-Chair of the IEEE Sweden Section. He received the Best Paper Awards at the IEEE CSCN'17 and the IEEE Globecom'11 and was a Co-Adviser of the Best Doctoral Thesis in CS in Finland, in 2017. He served on numerous journal editorial boards and conference program committees, including the IEEE INTERNET OF THINGS JOURNAL, *MDPI Sensors*, the IEEE ICNP, ACM MSWiM, and IFIP Networking. For more information, visit <http://gurtov.com>.

• • •