When will my PLC support Mirai? The security economics of large-scale attacks against Internet-connected ICS devices

Michael Dodson
University of Cambridge
Email: md403@cam.ac.uk

Alastair R. Beresford
University of Cambridge
Email: arb33@cam.ac.uk

Daniel R. Thomas
University of Strathclyde
Email: d.thomas@strath.ac.uk

***Abstract*—For nearly a decade, security researchers have highlighted the grave risk presented by Internet-connected Industrial Control Systems (ICS). Predictions of targeted and indiscriminate attacks have yet to materialise despite continued growth of a vulnerable population of devices. We investigate the missing attacks against ICS, focusing on large-scale attacks enabled by Internet-connected populations. We fingerprint and track more than 10 000 devices over four years to confirm that the population is growing, continuously-connected, and unpatched. We also track 150 000 botnet hosts, monitor 120 global ICS honeypots, and sift 70 million underground forum posts to show that the cybercrime community has little competence or interest in the ICS domain. Attackers may be dissuaded by the high cost of entry, the fragmented ICS population, and limited onboard resources; however, this justification is incomplete. We use a series of case studies to develop a security economics model for large-scale attacks against Internet-connected populations in general, and use it to explain both the current lack of interest in ICS and the features of Industry 4.0 that will make the domain more accessible and attractive to attackers.**

I. Introduction

The vulnerability of Internet-connected ICS is a well-studied problem. These devices present open industrial protocol ports to the Internet without any authentication or encryption [1], allowing an attacker to modify the device logic or directly manipulate device behaviour [2]. ICS devices also often expose ports and services to the Internet that are not specific to industry (e.g., webservers) [3], [4], and the slow patch cadence for this domain results in long-term risk from common Information Technology (IT) vulnerabilities (e.g., privilege escalation). Previous studies took snapshots of the Internet-connected ICS device population to identify the number and types of devices and infer risk [1], [3], [4]. These studies concluded that the ICS 'space is in disarray' [1], that the possibility of people with 'criminal intent or a political agenda…act[ing] against these devices is significant' [4], and that malicious actors 'might already be doing this' [3]; however, the predicted consequences have not materialised. We reassess this risk with an end-to-end assessment of the Internet-connected ICS landscape: we track the device population and device owner behaviour over a four-year period and perform a technical and criminological assessment of the adversarial community to understand why an apparently-vulnerable population is largely ignored by cybercriminals.

The Internet-connected ICS device population shares many characteristics with other populations that *have* attracted large-scale criminal activity, such as size, stability, and lack of long-term support. The population is approaching 100 000 devices, of which we tracked approximately 10% over four years. We find that over 60% of tracked devices are continuously connected at the same IP address for years, fewer than 25% ever receive a software update, and that the overall population is growing with the addition of the newest hardware and software from major manufacturers.

Despite these vulnerabilities, we show that the actual risk to ICS devices is low. We provide results from the largest, high-interaction ICS honeypot study to date; monitor scanning malware for knowledge of industrial protocols; and aggregate ICS interest from a database of 70 million hacker forum posts. Collectively, the data demonstrate that the cybercrime community has little competence or interest in ICS.

To explain the limited interest, we introduce a security economics model for characterising and predicting large-scale adversarial interest in a population of Internet-connected devices. The model was developed by studying successful, wormable malware such as Conficker, Mirai, and WannaCry, and identifying aspects of the device population or adversarial community that enabled large-scale exploitation. The model exposes several factors that currently deter large-scale malicious interest in the ICS domain, such as fragmentation in the ICS population, limited compute and memory resources on most ICS devices, and high cost of entry. The model also helps anticipate the effect of changes to the ICS population or adversarial community. We map ongoing and predicted changes to the industrial landscape and show that the ICS community is converging with the Internet-of-Things (IoT) community in terms of homogeneity and connectivity. We conclude that such a convergence makes the ICS domain an attractive target for attackers displaced by increased security or competition within the IoT domain; and even if ICS devices are not directly targeted, such convergence creates a new risk of being swept up in large-scale attacks targeting the billions of non-industrial IoT devices.

We make the following contributions in this paper:

- A security economics model for characterising and predicting the success of large-scale attacks against a device

TABLE I
SAMPLE ETHERNET/IP (PORT 44818) DATA

Property	Value
Host IP	198.51.100.42
Vendor ID	Rockwell Automation/Allen-Bradley
Product / Firmware	1766-L32BXBA / B14.00
Firmware Version	B14.00
Serial number	0xa1b2c3d4

population (Section III) and an empirical demonstration of our model against the ICS device population (Sections IV and V).
- The first longitudinal study of Internet-connected ICS devices (Section IV) and a demonstration that individual ICS devices can be retroactively tracked for years using publicly-available scanning data (Section IV).
- An economic and usability explanation for the growth of the Internet-connected ICS population (Section IV-E).
- A consolidation of existing literature associated with ICS runtimes, Industry 4.0, and Industrial IoT (IIoT), and an analysis of how these changes affect cybercriminal interest in the domain (Section VI).

II. HISTORY OF CONNECTED ICS

ICS are used to command, manage, or regulate physical systems in industry, critical infrastructure and building automation. ICS are often composed of many devices and may be distributed over a large area. Devices communicate using industry-specific protocols, most of which are legacy point-to-point or broadcast protocols designed for use with dedicated cabling; however, many protocols are now layered on top of TCP or UDP, and devices use existing IP-based networks, including the Internet, to communicate. As neither these legacy protocols nor the applications that use them support or implement authentication or encryption, a remote adversary can take control of a device and cause physical damage by sending well-formed packets [1] or modifying programmable logic [5], [6]. Even when securable protocols are used, they are often misconfigured in ways that reduce or eliminate any security benefit [7].

A large number of these ICS devices have their industrial protocol ports directly connected to the Internet. Search engines, such as Shodan and Censys, provide Internet-wide views of devices responding on industrial protocol ports by scanning the IPv4 address space and making response data publicly available [8], [9], [10]. Table I is a sample response for the query 'port:44818', which is conventionally reserved for the Ethernet/IP protocol (IP in this case stands for 'industrial protocol') and is designed for time-critical process automation. Protocols that provide similar detail in the query response include Siemens' S7comm (port 102), typically used to control manufacturing processes, and BACnet (port 47808), designed for building automation [1]. Other protocols, such as Modbus (port 502) and DNP3 (port 20000), provide less information, but still reveal whether a device is alive at a given IP address.

Prior studies quantified the size of the Internet-connected ICS population and demonstrated that the population was growing in absolute terms: studies in 2011, 2014, and 2016 identified approximately 8 000, 13 500 and 60 000 devices, respectively [1], [3], [4]. At the time of writing, the tools used in these studies identify approximately 100 000 devices.

Recent work evaluating Internet exchange traffic indicates that these 100 000 devices may be only a small percentage of the industrial devices using the Internet to communicate via insecure protocols [11] (e.g., because Shodan cannot identify devices behind Network Address Translation (NAT) or firewall device). Further, efforts to shift to secure protocols, or to wrap insecure protocols in a secure protocol, such as OPC UA,

III. SECURITY ECONOMICS MODEL

To better understand the risk of large-scale exploitation to Internet-connected ICS devices, we looked to successful exploitation of other Internet-connected populations. Using the categories in Anderson *et al.*, we pair at least one case study with each large-scale criminal activity: ransomware, cryptomining malware, Distributed Denial-of-Service (DDoS)-for-hire, spam, and disruption/destruction [12]. We use these studies to develop a security economics model for evaluating the risk of large-scale exploitation to an Internet-connected population of devices. Table II summarises the model, comparing the characteristics of the target population and attacker for each case study.

A. Case studies

Here we present case studies of large-scale attacks used to establish our model, focusing on named attacks or campaigns to support reproducibility and traceability. Further, we preferentially select recent attacks or attacks that affected cyber-physical systems, as our primary use for the model is to explain and predict cybercriminal interest in Internet-connected ICS devices. Additional technical details of each case study are compared in the appendix.

There have been several high-profile attacks against ICS devices (e.g., Stuxnet [13], BlackEnergy 3 [14], CRASHOVERRIDE [15]); however, these were targeted attacks and used industrial, Windows-based infrastructure to attack ICS devices that were not directly connected to the Internet; therefore, they are not considered here.

a) Conficker: Conficker emerged in October 2008, using multiple propagation, exploit, and self-update techniques to infect millions of Windows computers. Patches were released by Microsoft in the same month. At its peak, F-Secure reported nearly 9 million infections [16], and at the time of writing their are still 500 000 active infections [17]. The worm took advantage of a large population, known vulnerabilities, slow patch cycles, and a collection of existing exploits, including some from Metasploit [18]. Conficker was originally deployed by Ukrainian cybercriminals, but crime-related payloads only appeared in the fifth variant with limited distribution [17], [19].

TABLE II

LARGE-SCALE CYBERCRIME ENABLERS. (●) DENOTES A FULLY-MET ENABLER, (◐) DENOTES A PARTIALLY-MET ENABLER, AND (○) DENOTES A UN-MET ENABLER. UN-SHADED COLUMNS ARE POPULATED BASED ON THE ENABLERS IDENTIFIED IN THE CASE STUDIES (SECTION III-A). SHADED COLUMNS REPRESENT NEW ANALYSES OF THE CURRENT INTERNET-CONNECTED ICS POPULATION (SECTIONS IV AND V) AND A PREDICTIVE EVALUATION OF THE FUTURE POPULATION BASED ON CURRENT TRENDS IN ICS CONNECTIVITY (SECTION VI).

Category	Enabler	Conficker	Mirai	Brickerbot	WannaCry	NotPetya	UDP amp	Cryptomining	ICS: current	ICS: future
Vulnerable population	Internet-connected	●	●	●	●	●	●	●	●	●
	Large and/or stable	●	●	●	●	●	●	●	●	●
	Slow patching/no support	●	●	●	○	○	●	●	●	●
	Software and/or hardware homogeneity	●	●	●	●	●	●	●	○	◐
	Onboard resources	●	●	●	●	●	●	●	○	◐
	Known vulnerabilities	●	●	●	●	●	◐	●	●	●
	Predictable system/device response	◐	●	●	●	●	●	●	○	◐
Attacker incentives	Financial	◐	●	○	◐	◐	●	●	○	◐
	Ideological	○	○	●	◐	◐	○	○	○	◐
	Low exploitation cost	●	●	●	●	●	◐	◐	○	◐
	Low consequence to attacker	●	◐	●	●	●	○	●	○	◐
Attacker tools and resources	Develops exploits	◐	○	◐	●	●	◐	◐	○	●
	Adapts existing exploits and tools	○	●	○	●	●	◐	●	○	●
	Buys or co-opts access	○	◐	◐	○	○	◐	●	◐	◐

b) Mirai: The Mirai botnet emerged in late 2016 and used aggressive scanning and brute force password searches to infect hundreds of thousands of Linux-based IoT devices, such as routers and webcams. At its peak, an estimated 600 000 hosts were infected [20]. At the time of writing, the Cambridge Cybercrime Centre (CCCC) [21] still observes approximately 150 000 unique, infected IPs per day scanning a monitored /14 network. The success of Mirai has been attributed to the efficient use of Internet-wide scanning, insecurity of the target devices, ease of porting to a large number of device types, and a lack of IoT patching infrastructure. Further, the original source code was publicly released on a hacking community forum, allowing individuals to create botnets without possessing the skill to develop the tool themselves [20].

c) Brickerbot: BrickerBot malware was identified in early 2017 and used Mirai-like techniques to infect Linux-based IoT devices and perform Permanent DoS attacks [22], [23]. The author of the malware claimed he was mitigating the spread of Mirai and that he had 'bricked' over 10 million devices [24]. The malware continued to spread for over a year before being intentionally retired by its author. Even if the number of infections is exaggerated, Brickerbot provides an example of reusing or co-opting an existing botnet.

d) WannaCry and NotPetya: WannaCry and NotPetya are ransomware strains exploiting a vulnerability in Microsoft's implementation of the Server Message Block (SMB) protocol to spread across the Internet and between computers on a local network. The exploit, known as EternalBlue, was allegedly developed by the NSA and publicly released by an organisation called ShadowBrokers [25], [26].

WannaCry appeared in May 2017 and infected hundreds of thousands of Windows machines. Though Microsoft rapidly patched the vulnerability, even for officially-unsupported Windows versions, the attack was finally halted by a security researcher registering a kill-switch domain, preventing the malware from spreading further. The original version of the malware was flawed, and encrypted computers could not be decrypted, though subsequent versions allowed users to pay a ransom and recover data [27].

NotPetya appeared in June 2017 and similarly used EternalBlue to spread between Windows machines [28]. It was a highly modified version of the Petya ransomware strain and it propagated through popular accounting software in Ukraine [29]. The reliance on accounting software for initial distribution limited the direct impact to tens of thousands of victims. NotPetya turned out to be a disruption attack masquerading as ransomware, as the malware provides no means to decrypt infected computers [30], [31].

e) Webstresser: Webstresser was one of the largest DDoS-for-hire, or 'booter', services leveraging UDP amplification to direct traffic at a target. In the six months before its takedown in April 2018, Webstresser executed over 400 000 attacks [32]. UDP amplification attacks use common Internet protocols and services (e.g., DNS, NTP) to create DDoS attacks without a dedicated botnet. The attacks leverage traffic asymmetries with these services, where a client request of a few bytes results in a server response of tens or hundreds of bytes. The attacker spoofs the source IP in the request, so the response from the service is directed at the target. The attacker is not making use of any particular vulnerability, but rather exploiting characteristics of the service that cannot be 'patched' without fundamental changes to the service [33]. Further, booting appears to be a lucrative and easy service to set up, and, until recently, did not result in adverse consequences for providers [32], [34], [35].

f) Illicit cryptomining: Illicit cryptomining computes cryptocurrency hashes using host resources without permission from the owner. With the introduction of cryptocurrencies with

changing proof-of-work (PoW) algorithms (e.g., Monero), binary-based cryptomining with botnets has become profitable. As a result, many cryptomining campaigns use a combination of publicly available mining tools and pay-per-install (PPI) botnets [36]. This ecosystem is designed to adapt when the PoW algorithm changes: miners are updated to the new algorithm and the cryptomining campaigns pay the botnet owners to install the new miner on botnet hosts. Pastrana *et al.* estimate that as of November 2018, such campaigns had illicitly mined over $57 million in Monero [36].

B. Large-scale cybercrime enablers

From these case studies, we identify three, largely-orthogonal categories that are present in each case: a *vulnerable population*, clear *attacker incentives*, and a degree of domain-specific *knowledge, tools, and resources*. Within each of these categories, the case studies reveal several sub-properties, or 'enablers', that characterise the larger category. The categories and constituent enablers are discussed further below. Table II maps the individual case studies into our model, showing the enablers that are fully or partially met for each of the attacks or campaigns. We mark enablers as 'partially-met' if there is uncertainty (e.g., whether the WannaCry incentive was financial or disruptive), if the enabler is possible but not well-documented (e.g., co-opting access to Mirai botnets), or if the enabler is possible but not well-executed (e.g. Conficker spreading faster than intended). We also include columns for Internet-connected ICS devices based on a study of the current population (Section IV) and the anticipated, future population (Section VI). For the latter, we consider enablers to be partially met if they are expected to be possible but are not yet demonstrated.

We argue that enablers from these three categories must be present in high measure for successful, large-scale exploitation of a population. We also show that there is some element of progress; namely, that the existence of a vulnerable population creates incentives that lead to attackers either developing or adapting tools to target that population. Further, without intervention, a positive feedback loop exists, such that a larger attacker population will be attracted to a domain, bringing broader incentives and skills. The following sections describe the categories in more detail, emphasising each *constituent enabler* corresponding to a row of Table II.

1) Vulnerable population: All cited cases involve a *large*, *Internet-connected* population of devices (e.g., Windows computers, IoT devices). While some of these cases could spread via removable media, none primarily relied on such methods for propagation; in all cases, hosts are directly addressable and have open ports. Similarly, these populations all had (or have) *unpatched vulnerabilities* or are fundamentally constructed in such a way that precludes simple mitigation. The populations also have a strong, initial *homogeneity*. To perform the tasks required by the malware authors, devices require sufficient *onboard resources*. In the case of ransomware, this may include sensitive information. For a DDoS botnet, this includes sufficient spare processing and memory resources to avoid impairing normal device function, as the nature of a botnet is to be a parasite. Finally, the device or system response in all cases is *predictable*, though in some cases the malware appears to have spread faster (e.g., Conficker) or wider (e.g., NotPetya) than initially intended.

2) Attacker incentives: The most obvious incentive for cybercrime is *financial*, as demonstrated by ransomware, spam botnets, and DDoS-for-hire services. There may be other incentives, however. So-called 'script kiddies' may seek prestige in their community (e.g., developing DDoS infrastructure to boot rivals from online games [32]). Alternately, *ideology* (e.g., Brickerbot) or grey conflict (e.g., NotPetya) may provide the incentive. An attacker must be able to turn the attack into the capital of the incentive. For example, a ransomware campaign against devices without screens requires an additional step to inform the device owner of the attack and provide a remediation mechanism [37]. Similarly, attackers must ensure gains exceed costs, including the financial *cost to develop and deploy* the attack and the potential *personal consequences* of executing an attack. There is currently little appetite for pursuing low level cybercriminals [38], [12], and even the authors of Mirai avoided jail time, despite a conviction [39]; however, there have been recent cybercrime prosecutions, including those associated with booter services [32], [34], [35], which may be the leading edge of change.

3) Attacker tools and resources: Most perpetrators of cybercrime do not appear to *find and exploit vulnerabilities* themselves, but *rely on and adapt tools* from sources with greater skill and different motivations. WannaCry's reliance on EternalBlue, which has a nation-state pedigree, is a prime example [25], [26]. Similarly, the original Mirai source code was published by an individual with perhaps greater skill than many of the code's subsequent users. The Mirai author even mocked the 'skid[s]' who would use it: 'I know every skid and their mama, it's their…dream to have something besides qbot'[40]. This highlights the fact that even if a given population of devices appears to be too challenging to exploit on a given day, a high-capability tool placed in the hands of low-capability adversaries can change the threat model in a very short period. Finally, in some cases attackers *buy or co-opt* access, such as in binary-based criminal cryptomining, which relies on PPI botnets [36].

C. Security economics model summary

Our model consists of the categories, enablers, and descriptions elaborated above and exercised for several case studies in Table II. It provides a qualitative foundation for security economics assessments of target and adversarial populations to predict the likelihood of large-scale cybercriminal interest in the target population, or to assess the factors driving an existing and well-established interest in a target population. In the following sections we apply the model to Internet-connected ICS devices and potential adversaries.

IV. Longitudinal study of ICS

To apply the model in Section III to Internet-connected ICS devices, we performed a longitudinal study to identify device

characteristics and infer device owner behaviour over a four year period. To identify enablers for large-scale exploitation, we leverage historical data from both Shodan and Censys to track individual devices, allowing us to investigate population growth, stability, and patch cadence.

A. Internet-wide ICS scanning

Shodan and Censys scan dozens of ports over the IPv4 address space and make results publicly available. Both request information from ICS hosts on several protocols. The six most common industrial protocols are Ethernet/IP (port 44818), BACnet (port 47808), S7comm (port 102), Modbus (port 502), DNP3 (port 20000), and Niagara Fox (ports 1911 and 4911). Shodan scans all six protocols, while Censys scans all except Ethernet/IP. We focused on Ethernet/IP, S7comm, and BACnet because devices communicating over these protocols self-report more information than other protocols.

Threat to validity: While ports 44818, 102, and 47808 are the default (and most common) ports used for Ethernet/IP, S7comm, and BACnet devices, respectively, the port number may be configurable and Internet-connectable devices can communicate using these protocols on alternate ports. Gasser *et al.* demonstrated that only about 50% of Internet-connected BACnet devices use port 47808, and the remainder use ports 47809 to 47823 [41]; however, they showed that devices on alternate ports were equally susceptible to protocol-based attacks. Shodan and Censys only scan the default ports (i.e., 44818, 102, and 47808), thereby missing up to half of Internet-connected ICS devices. Because the devices on these alternate ports appear to have the same characteristics as the devices on the default ports, and to minimise our own active scanning (based on ethical considerations discussed at the end of this paper), we chose to limit our datasets to those provided by Shodan and Censys, but consider our conclusions to apply to the broader Internet-connected ICS population.

B. Fingerprinting

Globally-unique, static fingerprints support tracking devices over time to identify changes to the Internet-connected ICS population and to observe user behaviour. Device fingerprinting using historical data relies on identifying device-specific, immutable characteristics in existing datasets. From the three protocols of interest, we develop fingerprints for the manufacturer with the largest number of devices responding on that protocol: Allen-Bradley for Ethernet/IP, Siemens for S7comm, and Tridium for BACnet. Allen-Bradley and Siemens devices provide a serial number that appears to be globally unique, with the exception of honeypots (e.g., default instances of Conpot [42]). Tridium devices do not report serial numbers, but do report several fields of user-configurable data. By evaluating data returned from hundreds of devices, we chose the 'object name' field, as it is generally configured by the user and does not appear to have a default value. Using 'object name' produces globally-unique fingerprints for about 90% of devices, with the exception being devices with general entries such as 'device'. Table III compares the number of ICS devices to the number of devices with unique fingerprints, showing that we are able to uniquely identify a large percentage of the ICS device population to support our longitudinal study.

TABLE III
SUMMARY OF OBSERVED ICS DEVICE POPULATION SIZES AS AN AVERAGE OVER THE DURATION OF THE STUDY.

Source	Manufacturer	Devices per Snapshot	Fingerprints per Snapshot
Censys	Siemens	3 881	1 105
	Tridium	2 117	1 882
Shodan	Siemens	3 227	1 025
	Tridium	2 380	2 079
	Allen-Bradley	6 985	5 674

Threat to validity: Our fingerprints are limited by the lack of ground truth data and reliance on user-configurable data instead of inherent characteristics of each device (e.g., clock skew [43] or factory calibration [44]); however, the stability observed over several years indicates that the fingerprints are sufficient for our study. Additionally, the conclusions drawn in this section are based on the ability to track individual devices to determine their stability, longevity, and patch cadence. For Tridium and Allen-Bradley devices, the ratio of fingerprintable devices to total devices is high (approximately 90% and 80%, respectively). In contrast, the ratio for Siemens devices is low (approximately 30%), introducing potential bias to our analysis. This would be a concern if the lack of a fingerprintable serial number on the remaining 70% of Siemens devices was evidence of greater security awareness amongst device owners; however, the fact that these devices are connected to the Internet at all implies that they are not configured more securely, but that the device type simply does not report a serial number by default. Based on this, we consider conclusions regarding the fingerprintable Siemens devices likely apply to the entire Internet-connected Siemens population, but further work is required to demonstrate this conclusively.

C. Comparison of Shodan and Censys

Other studies have qualitatively demonstrated inconsistencies between Shodan and Censys [1], [45], while our fingerprint allows us to quantify and characterise the difference.

Shodan returns all hosts open on a given port, whether or not they respond correctly to protocol-specific requests. For example, on 10 June 2019, Shodan returned 55 989 results for the query 'port:44818' (Ethernet/IP), of which only 8 501 responses could be parsed as Ethernet/IP data. The remainder are likely incorrectly configured hosts unrelated to the ICS domain. We find all Censys returns to be valid, in that all returns provide parsable, protocol-specific responses, going some way to validating results from studies based on ZGrab, the underlying scanning tool for Censys [1], [10]. In contrast, studies relying solely on Shodan data that do not distinguish between an open port and a correctly-communicating device may over-predict connected devices numbers.

To characterise the difference between Shodan and Censys, we used our fingerprint to compare the set of devices returned

TABLE IV
COMPARISON OF DEVICES IDENTIFIED BY SHODAN AND CENSYS BETWEEN DECEMBER 2018 AND MARCH 2020.

Vendor	Number of Devices		
	Intersection	Shodan Only	Censys Only
Siemens	1 305	468	418
Tridium	3 016	331	206

by each source. Table IV provides a high-level summary of the comparison, which was limited to devices with a unique fingerprint obtained between December 2018 and March 2020.

Table IV shows broad agreement between the two search engines. The larger difference between the two Siemens results may be due to intentionally-limited commands used by Censys [1]. A high-level analysis of the devices identified by one search engine over another did not reveal any obvious reasons for the difference; however, an analysis taking into account the IP blocks from which each search engine scans may be more instructive [45].

We also compared the data for devices identified by both search engines and found nearly perfect agreement. For example, there was 100% agreement for the Siemens model numbers (e.g., 6ES7 315-2EH13-0AB0). Similarly, all but two of the Tridium firmware versions agreed, and the difference appears to be due to a recent software update that had been indexed by Censys but not Shodan. This is further supported by Table III, which shows a consistent fingerprint-to-device ratio between Shodan and Censys.

D. ICS device owner behaviour

While previous studies accurately state that devices are vulnerable to manipulation as long as their industrial protocol port is open to the Internet, they do not attempt to understand why or how these devices are connected. We demonstrate the ability to infer several behavioural characteristics of device owners, including how the devices were connected to the Internet, the software patch or update frequency, and the age of devices being connected.

1) Data collection methodology: Censys provides access to past scanning results via Google BigQuery. We took retroactive, quarterly snapshots from December 2015 to March 2020, resulting in a dataset of all S7comm and BACnet device indexed by Censys on the first day of each quarter. Shodan does not provide direct access to historical scans, but does support historical queries of individual IP addresses. We took daily Shodan snapshots between December 2018 and July 2019, and then occasional snapshots until March 2020, and used the IP addresses in those snapshots to perform historical queries. The resulting dataset includes three to four years of data for ICS devices recently indexed by Shodan; however, any device disconnected before December 2018 would not have been in the list of IPs for which historical queries were performed. To overcome this limitation, IP addresses indexed by Censys between December 2015 and March 2020 were used to supplement the list of IP addresses indexed by Shodan, and this composite list of IP addresses was used to perform Shodan historical queries, surfacing Shodan data for ICS devices that were no longer indexed by Shodan in December 2018. This technique was used for Siemens and Tridium devices only, as Censys does not scan port 44818, used by Allen-Bradley devices.

Threat to validity: The use of IP addresses from Censys in Shodan historical queries biases any further comparison between the two datasets, and we did not use this technique in Section IV-C. In this section, our goal is to maximise the data available from any IP address at which an ICS device has been hosted and we consider the inclusion of IP addresses from Censys the best way to do so, especially in the case of Siemens devices, as only Shodan accurately parses the Siemens firmware information. Similarly, performing quarterly Censys snapshots risks missing devices connected for short periods (i.e., less than 90 days, between quarterly snapshots); however, daily Shodan snapshots showed that almost all devices remained at the same IP address and were connected for extended periods, so we considered more frequent Censys snapshots to be unnecessary.

2) Stability and IP assignment: We used our fingerprint to identify the time each device spent in the population and whether the device's IP address was stable or subject to DHCP churn. Table V documents the percent of devices seen more than once with a *stable presence* (i.e., devices in the population for more than one year) and with a *stable IP* (i.e., devices only observed at one IP address). As discussed above, the Shodan data for Allen-Bradley is skewed toward devices with shorter connection times, as it only incorporates the history of devices connected between December 2018 and March 2020, whereas the remaining rows include all devices connected between December 2015 and March 2020.

This is the first confirmation of the stability of the Internet-connected ICS population, demonstrating that most devices have stable IPs, are continuously connected, and are connected for extended periods. Combined with the lack of patching (see Section IV-D3), this means that an attacker has a nearly unlimited opportunity to reconnoitre, develop, and deploy attacks against this population.

Further, these results provide some insight into the different device environments. For example, nearly 60% of the Allen-Bradley devices are connected via cellular Internet Service Providers (ISPs), such as Verizon or AT&T Wireless, which, combined with the stable IP addresses, suggests they are field devices; whereas fewer than 5% of Siemens devices have cellular ISPs, but double the percentage of dynamic IP addresses, implying they may be inadvertently connected to the Internet, as a dynamic IP has limited value for remote monitoring and control without a service like DynDNS.

Table V shows a higher stable IP address percentage, but a lower stable presence percentage for Shodan compared to Censys. This is an artefact of our data collection methodology. By performing historical Shodan searches on individual IP addresses obtained from Censys, we expect IP address stability for the Shodan results to increase, since the set of

TABLE V
SUMMARY OF ICS DEVICE POPULATION BEHAVIOURS OVER THE COURSE OF THE STUDY. ALL VALUES REPRESENT A PERCENTAGE (%) OF FINGERPRINTABLE DEVICES.

Source	Vendor	Firmware update	Stable IP	Stable presence	Replaced devices
Censys	Siemens	-	58.7	71.3	-
	Tridium	29.6	66.3	64.4	-
Shodan	Siemens	0.5	73.0	40.4	1.3
	Tridium	22.3	82.7	53.7	5.7
	Allen-Bradley[a]	1.8	85.1	63.1	5.0

[a] Historical query based on Shodan IP addresses only

devices is now the union of those identified by Censys and Shodan, whereas the Censys percentages would not include any device only identified by Shodan. Similarly, because we are performing historical searches on IP addresses identified by Censys as early as December 2015, but no longer observed by Shodan in December 2018 (our first Shodan snapshots) we would expect the stable presence percentage to decrease, as the search window is effectively truncated by 15 months.

3) Firmware: Devices that use Ethernet/IP, S7comm, or BACnet over IP provide detailed firmware and/or application software version information in response to a protocol-specific request (e.g., Table I). Nearly every manufacturer of ICS devices with IP communication capability reports such version information. In all cases, the version information provides sufficient granularity to correlate with release notes and Common Vulnerabilities and Exposures (CVEs).

Table V shows the percentage of devices seen more than once for which at least one *firmware update* was observed. We used Censys to track Tridium firmware and Shodan to track Tridium, Siemens, and Allen-Bradley firmware. While Censys does index Siemens data, the firmware version is not correctly parsed. Censys confirmed to us that the data was incorrect and the issue was being tracked, but stated that they did not keep raw historical data for re-processing.

Overall, a stark contrast is evident between building automation controllers (Tridium) and Programmable Logic Controllers (PLCs) (Allen-Bradley and Siemens). The difference may be accounted for by the centrality (logically and physically) of building automation controllers compared to PLCs, which may make it easier to install updates and anticipate or test effects on the rest of the system. Similarly, maintenance contracts for building automation systems may include the entire, integrated system, so the third-party is responsible for and licensed to install updates and test the whole system. Conversely, maintenance contracts for physical processes under PLC control may be limited to the single device or process, so the third-party maintainer is not responsible for the integrated testing that may be required following a software update.

While it may seem that a failure to patch is a secondary problem while the protocol port remains open, these observations raise two immediate concerns. First, vulnerabilities may provide an attacker greater access than the protocol alone. For example, over 70% of the Tridium devices identified

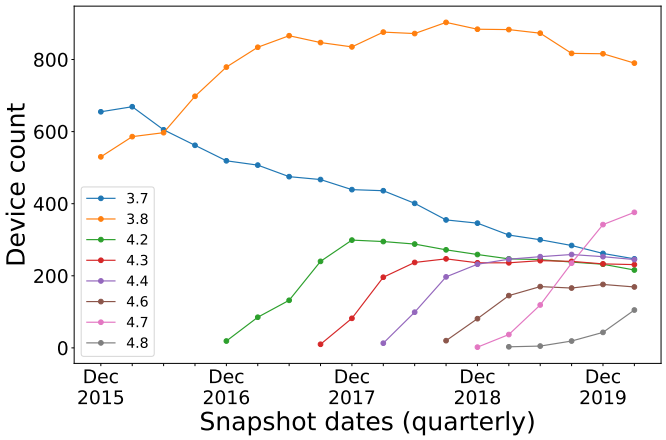

Fig. 1. Tracking firmware versions of Tridium devices running the Niagara framework shows that few devices are updated and that the population includes the newest device hardware and software.

by Shodan report a firmware version against which there is at least one critical-severity CVE allowing an attacker to gain full administrative privileges (CVE-2017-1674). Second, we hypothesise that these results are at least qualitatively applicable to non-Internet-connected ICS devices, as software updates may need to pass through an integrator, be installed by a third party, and require integrated testing after installation, all of which incur delay and engender resistance.

Figure 1 shows frequency plots of firmware versions over time for Tridium devices running the Niagara framework, demonstrating limited, long-term patching behaviour. The figure also shows that connecting ICS devices directly to the Internet is not a legacy problem: even the newest devices with the most recent firmware (version 4.8 was released in 2019) have their industrial protocol ports directly connected to the Internet. It is encouraging to see that manufacturers are releasing updates that are incorporated in new device installations; however, the infrastructure to distribute and install patches in live systems remains inadequate.

Table V also shows the percentage of *replaced devices* observed between December 2015 and March 2020. We tracked this metric separately, but include it in this section because most replacements reported newer firmware. Replacements were identified by fingerprints and timestamps. If one device at an IP address was no longer observed (at any IP address) after the initial observation of a second device at that IP address, the first device was considered replaced. The method is unable to account for DHCP churn and likely under-predicts replacements; however, the majority of devices have stable IP addresses, so the replacement percentages are at least qualitatively instructive and show that device upgrades do not compensate for lack of patching.

E. ICS usability and security economics

The vulnerability of Internet-connected ICS devices raises an obvious question: why are these devices connected to the Internet in the first place?

We are not aware of empirical research explicitly asking this question to device owners; however, we can map existing usable security and security economics research from the IT domain into the industrial domain. Vendor guidance for ICS devices can be summarised as follows [**?**], [**?**], [46]:

1) Do not connect to the Internet
2) If you must connect, use segmentation and isolation
3) Use role-based access control and strong passwords

The guidance is sound but fails to consider user expectations and security models [47], [48], and it does not account for the cost/benefit decisions users are forced to make when determining the resources to invest in security [49], [50]. For example, there is currently no evidence of large-scale criminal interest in Internet-connected ICS devices (see Section V), so the likelihood of an attack against an arbitrary industrial system appears low. Further, the security community has not demonstrated a common-case consequence for attacks against such devices. While the consequence of state-sponsored attacks against high-profile targets are well publicised (e.g., Stuxnet [13], TRITON/TRISIS [51]), it is unclear what attacks and consequences might be expected by smaller, low-profile organisations. With low likelihood and no demonstrable consequence, demanding a wide-spread investment in security is ineffective [49], [50].

The first recommendation does not accord with user expectations, as demonstrated by the fact that many devices are not only intentionally connected, but are packaged with cellular modems for that purpose. Also, given that many devices are equipped with webservers and industrial protocol ports enabled and open out-of-the-box, the recommendation is not enforced by default. Instructions for setting up webservers on many devices fail to mention that the industrial protocol port is open and accessible from the same physical port as the webserver; therefore, users may follow guidance for securing their webserver and inadvertently connect their unprotected, industrial protocol port directly to the Internet. This is highlighted by recent work showing 92% of reachable, Internet-connected OPC UA deployments (a securable industrial protocol) had critical security misconfigurations[7].

The second recommendation generally comes alongside security appliances that the user needs to buy in order to adequately secure the device (e.g., security processors [**?**]). Basic security appliances can cost as much as the ICS device being protected, and generally two are required to secure each end of the communication link. Having purchased a security appliance, the user now has to consider the expertise and expense to maintaining it. Given the high cost and lack of identified risk, users may be making a rational decision not to invest in further security while still benefiting from the connectivity of their new device [49].

The third recommendation has the same problems as the second: because most ICS devices lack host-based access control, implementing a robust access control policy across a facility with a collection of heterogeneous devices generally requires buying and maintaining additional infrastructure.

In summary, the number and growth of Internet-connected ICS devices is not surprising. Some of these devices may be inadvertently connected, due to failures in usable security. Others may be intentionally connected based on either rational, risk-based decision-making or inadequate understanding of security. Regardless, we should expect that the number of connected devices will continue to increase as the number of connectable devices grows.

V. Abusing ICS devices

The Internet-connected ICS population is large, stable, slow to patch, and growing, which meets several of the enablers for large-scale cybercrime (Section III). The *ICS: current* column of Table II provides a comparison of Internet-connected ICS devices against other populations targeted by large-scale attacks, and shows that there are several missing enablers: homogeneity, predictable response, and onboard resources. The model predicts that these missing enablers will inhibit cybercriminal interest in the ICS domain. In this section, we attempt to empirically test that prediction by studying the adversarial population to complete the Table II analysis.

In this section, we are specifically interested in large-scale exploitation of ICS devices directly connected to the Internet. As noted in Section III-A, there have been several high-profile attacks against ICS devices (e.g., Stuxnet [13], BlackEnergy 3 [14], CRASHOVERRIDE [15]); however, these used industrial, Windows-based infrastructure to attack ICS devices that were not directly connected to the Internet. Similarly, there have been high-profile ransomware attacks against industrial organisations (e.g., Norsk Hydro [52]), but, again, these attacks targeted Windows-based infrastructure. While they effectively shut down associated industrial processes, they were not attacks against ICS devices themselves.

First, we collaborated with SecuriOT[1] to evaluate a year of data from a global network of 120 high-interaction ICS honeypots to see whether malicious actors were attempting to modify the behaviour of ICS devices with protocol-specific commands (Section V-A). Second, we used a database of over 70 million cybercrime forum posts to measure interest in and identify any explicit exploitation of ICS devices (Section V-B). Finally, we evaluated whether the ICS devices, or the routers and modems through which they connected to the Internet, were infected with Mirai (Section V-C). This was a proxy to see whether such devices were generally targets for botnets.

These evaluations provide no evidence of cybercriminal interest or expertise in attacking ICS devices at scale, nor that exposed industrial protocol ports are being used as a vector for attacking such devices. We acknowledge the challenges associated with proving a negative, but believe our conclusions are supported by consistent data from largely orthogonal domains (honeypots, cybercrime forums, and malware).

A. Honeypots

There are several studies involving ICS honeypots that identify large amounts of scanning traffic; however, none of

[1]https://www.securiot.se/

these studies identify any instance of malicious manipulation of an ICS device through the industrial protocol [1], [53], [54], [55], though there are indications that this capability exists [56]. These studies were conducted over short periods (weeks or months), had a limited number of devices, were not geographically distributed, or used honeypots that were easily fingerprinted (e.g., they were hosted on Amazon Web Services and used default Conpot [42] configurations); therefore, the lack of interaction may not be surprising.

We collaborated with SecuriOT to overcome some of these shortfalls. SecuriOT develops deception technologies that mimic PLCs, remote terminal units, and firewalls, and they run 120 Internet-connected honeypots in 22 different countries as part of their own intelligence gathering operation. SecuriOT provided us one year of labelled interactions with their honeypot network, covering March 2018 through March 2019. We did not have access to raw packets, so our assessment was limited to packet header data, a fingerprint of the tool used for the interaction, and their analysts' classifications of interactions as either reconnaissance or exploitation.

In one year, SecuriOT recorded approximately 200 000 packets from 80 000 interactions, of which about 1 000 attempted to modify equipment behaviour. All but nine of these were Mirai variants initiating an SSH session with a router honeypot. The nine interactions targeting an ICS protocol used S7comm, Modbus, and IEC-104. According to SecuriOT, these attacks were either DoS or command replay attacks. All exploits were responsibly disclosed.

The results confirm that there is no large-scale effort to indiscriminately compromise ICS devices, though there are many reconnaissance efforts. However, this data may provide the first evidence of efforts to maliciously manipulate Internet-connected ICS device behaviour through the ICS protocol, demonstrating that attackers are interested in protocol-specific attacks and that high-interaction honeypots can successfully deceive skilled attackers [57].

If Internet-connected ICS devices are only subject to targeted attacks, it is not surprising that the number of observed compromises is low, as they likely represent either an attacker testing against a random system or mistaking the honeypot for a specific target. This demonstrates one of the limitations of honeypots as intelligence gathering tools for targeted attacks: while they might reveal an attacker capability or interest, they cannot be used for quantitative measurement in the same way as honeypots targeted by indiscriminate malware like Mirai.

B. Hacker forums

To evaluate cybercrime community interest, we used the CCCC's CrimeBB dataset [58], [21], which contains over 70 million posts from over ten years, scraped from dozens of publicly-accessible cybercrime forums, including Hackforums, where the Mirai source code was originally shared. Analysing hacker forums has been used to characterise and predict cybercrime behaviours in several domains, including cryptomining malware [36], eWhoring [59], and booter services [32]. We searched for ICS-related terms (e.g., 'ICS', 'Shodan', vendor and protocol names), identified relevant posts, and evaluated the entire thread of which the post was a part.

The initial search returned over 13 000 posts, from which we filtered out posts that were either not in English or obviously not applicable.[2] We manually evaluated the remaining posts for relevance before extracting applicable threads. We discarded threads that consisted solely of extracts from news articles or other websites without follow-on posts (e.g., copied Stuxnet articles), threads with hyperbolic claims (e.g., selling an exploit for $1 million), requests for information about hacking without follow-up, and conflations of ICS with IoT (e.g., dozens of posts with links to open webcams).

We identified fewer than 30 relevant threads, including tutorials, discussions of vulnerabilities, demonstrations of compromise, and offers of credible service. There were several Shodan tutorials, scrapers for obtaining Shodan API keys from GitHub, IP addresses of Internet-connected ICS devices, and videos of people manipulating thermostats via web interfaces. Two threads competently discussed how resource constraints and long lifetimes impeded security implementation in the ICS domain. One user claimed to work for a power utility and offered physical access to smart meters, though we found no evidence of anyone offering to pay for his services. Finally, we found two threads where managers of ICS equipment were, apparently inadvertently, providing details of their equipment and security practices; in both cases, they were complaining about the restrictions placed on them by system administrators.

Compared to other cybercrime domains, such as eWhoring, which has over 6 500 tutorial-related posts on Hackforums [59], the ICS picture is one of limited interest and competence. In other domains we observe correlations between posting and actual malicious activity, which is absent in the ICS domain. For example, a spike in Monero-related posts directly corresponds to a demonstrable increase of Monero crypto-mining malware in the wild [36]. Further, these other communities stimulate their own interest and profitability: the distribution of tools (e.g., source code) and knowledge (e.g., tutorials) allow less experienced actors to join and potentially profit, turning niche domains into large-scale attack targets. For ICS, we do not observe even a nascent or incubating community that would support development and distribution of such tools and knowledge.

C. Mirai

Given the cybercrime community's penchant for effective modification and reuse of existing malware, one might expect that initial efforts at large-scale exploitation of Internet-connected ICS devices would be based on modified IoT-targeting malware. Mirai is a prime candidate for such modification based on its flexibility, simplicity, and demonstrated

[2]For example, the search term 'ics' returned 1,033 posts, of which 286 referred to Android Ice Cream Sandwich, 111 referred to Internet connection sharing, and 165 were in Russian (while Russian posts may be relevant, we determined it was not worth translating them based on our evaluation of the English-language posts).

effectiveness [20]. As all known Mirai variants target Linux-based hosts, Mirai would have to be modified for new target hardware and software, as few ICS devices are known to be Linux-based. It is also unclear whether a Mirai-like exploit could infect a device through the industrial protocol port, though proof-of-concept, wormable PLC malware has been demonstrated [37]. Such a worm may also be able to infect and spread via non-industrial protocols on an ICS device (e.g., FTP, HTTP). To detect such modifications to existing malware, one could either directly look for modified Mirai source code in the wild or indirectly look for infected ICS hosts. In this section, we describe the latter method.

Devices infected with Mirai scan for other potential hosts with a distinctive scanning packet that can be used to identify infected hosts [20]. The CCCC [21] collects these packets and makes the data available to researchers. We used a real-time feed of source IP addresses suspected of hosting Mirai and identified approximately 150 000 unique, infected IPs per day. We used ZGrab[3] to scan these Mirai source IP addresses on industrial protocol ports to identify whether any ICS devices were hosted at the infected IP addresses.

Fewer than 1% of the Mirai source IP addresses provided a parsable response to our scans, indicating that there is only a small population of IP addresses hosting an ICS device and a Mirai host. Notably, none of the Mirai hosts were scanning ICS ports, strongly implying that they were not ICS devices, but rather 'normal' Mirai hosts (e.g., routers) sharing an IP address with an ICS device.

Mirai is frequently modified for different hardware and software targets, and its scanning behaviour is unique and easy to detect, making it a good starting point for investigating whether cybercriminals are adapting existing malware for the ICS domain. This negative result is not generalisable to all possible ICS malware; however, the combination of passive honeypots with an active search for infected devices with ICS knowledge (e.g., the ability to parse industrial protocols) may provide leading indicators of a shift in cybercriminal focus.

D. Summary

In Section IV we characterised the large, stable, and vulnerable population of Internet-connected ICS devices; however, our security economics model (Table II) predicted that the population was insufficiently homogeneous, resourced, and predictable to attract attention from the cybercrime community. In this section, we used orthogonal investigative methods to empirically confirm that the cybercrime community has little competence or interest in the ICS domain. While our focus is on largely undefended, Internet-connected ICS devices, our assessment is confirmed by security companies responding to incidents on defended networks [**?**]: ransomware attacks against Windows infrastructure are the only large-scale attacks affecting industrial systems.

[3]https://github.com/zmap/zgrab2

VI. Changing industrial landscape

While the current landscape makes ICS devices unattractive targets for scaled attacks, expected changes to industry are likely to overcome missing enablers in the *ICS: current* column of Table II. In this section we describe these changes, and in the next section we show how these changes may make the future Internet-connected ICS device population an attractive target for cybercrime.

a) ICS runtimes: Industrial runtimes are third-party software that provide common development and execution environments across large numbers of heterogeneous device families. For example, the CODESYS runtime is used in ICS devices and development environments from Schneider Electric, Beckhoff Automation, Bosch, and WAGO, all major ICS vendors [60]. Runtimes are a growing target for security researchers because they create larger, homogeneous populations and broaden attack surfaces by adding complexity and new interfaces. For example, security researchers have developed several exploits in versions 2 and 3 of CODESYS, supported by at least 360 device types [60], [61].

b) Industry 4.0: Industry is trending toward systems with greater connectivity, greater flexibility, and denser information flows, resulting in an increased number of devices in a system and an increased number of *connectable* devices in a system. Data-driven decision making (e.g., condition-based maintenance) requires more sensors providing high resolution data [62], and the desire for reconfigurable manufacturing environments encourages connectable devices, wireless connections, and Software Defined Networks (SDNs) [63]. The net result of these pressures is a greater number of connectable devices in a given industrial environment.

c) Common IoT and IIoT platforms: Technology companies are capitalising on this IIoT market, deploying ready-made solutions for device manufacturers, integrators, and operators. For example, Microsoft's Azure Sphere is a hardware, software, and cloud solution for IoT and IIoT devices [**?**]. Similarly, Huawei has developed open source operating system and cloud solutions for IoT and IIoT [64]. These solutions are designed to make development and testing easy: development boards are inexpensive and readily available, and the software stacks include full, open source operating systems. These products greatly simplify the effort to develop and test malicious software by increasing platform availability and providing higher level and more powerful hardware and software resources.

VII. Economics of attacking ICS

In this section, we attempt to answer two questions using our security economics model (Table II):

1) Why is the current Internet-connected ICS device population not a target of large-scale cybercrime?
2) Will the changing industrial landscape make attacks against the future ICS device population more likely?

Using the *ICS: current* and *ICS: future* columns of Table II we address these questions by looking at each category,

describing the enablers that are currently missing for the ICS population (*Current situation*) and showing how those enablers are or are not satisfied by expected changes to the industrial landscape (*Outlook*).

A. Vulnerable population

a) Current situation: While the Internet-connected ICS population is approaching 100 000, it is fragmented amongst dozens of vendors, with a heterogeneous collection of proprietary hardware and software; even if an attacker developed wormable malware for a particular device family, the vulnerable population may only include a few thousand devices. For example, the largest population of Internet-connected ICS devices from a single vendor (Allen-Bradley) consists of fewer than 10 000 devices. Compared to 2.5 billion Android devices [65], 1.5 billion desktop computers [66], 1.3 billion iOS devices [67], and 1.2 billion in-home IoT devices [68], the ICS population is irrelevant to a large-scale attacker.

Additionally many ICS devices are installed in a unique, physical system, and the response to a manipulation of a given ICS device would be difficult to predict. This is a concern for authors of parasitic malware (e.g., botnet), because unpredictable system response may result in malware being identified and removed. This contrasts with many standalone IT and IoT assets with predicable and easily tested responses. For example, Mirai on a webcam is unlikely to affect the behaviour of any other device on the local network.

b) Outlook: The fragmented nature and limited resources of the existing ICS population are overcome by the growing use of runtimes across manufacturers, the concentration of IIoT hardware and software amongst a limited number of vendors, and the increased hardware and software resources available from those IIoT platforms.

Devices with Linux-based or other widely used, open source operating systems provide the attacker with a large community of people looking for vulnerabilities and developing packaged exploits (e.g., Metasploit). Further, given the slow patch cadence of the Internet-connected ICS population, an attacker can expect vulnerabilities identified in white hat communities to remain exploitable for a long time. Additionally, many IIoT devices (e.g., sensors) are more self-contained than traditional ICS devices (e.g., PLCs), making it easier to predict the device and system response.

A growing population of increasingly homogeneous devices with common runtimes or open-source operating systems satisfies, in part or full, all the missing enablers from the *Vulnerable population* category of Table II. Further, as ICS hardware and software converge with the IoT domain, ICS devices risk being swept up in large-scale attacks targeting the billions of non-industrial IoT devices.

B. Attacker incentives

a) Current situation: Monetising an attack against ICS devices at scale is challenging for several reasons. First, many ICS devices have limited onboard resources, making them unattractive from a cryptomining perspective. Second, ICS devices are unlikely to store high-value data, such as purchase orders, personal details, or financial records, commonly used to support a ransomware campaign. Industrial organisations are generally well-equipped to handle device failures without significant impact to operations, and an encrypted device could simply be treated as a failed device [37].

While the potential financial benefit from exploiting a large population of ICS devices appears to be low, the cost to develop and deploy an exploit appears to be high. The direct cost of buying devices for development and testing may run to thousands of dollars. Further, actually developing a wormable ICS exploit has been shown to require substantial subject matter expertise and time [69], [37], [70].

Finally, there may be personal consequences for the attacker. Authorities have successfully identified and prosecuted those responsible for several cyber attacks against industrial systems [71], [72], [73], [74]. This contrasts sharply with non-industrial cybercrime: while there have been some high-profile cases [34], [35], existing infrastructure to identify and successfully prosecute offenders remains weak [12].

Overall, the high cost to develop exploits, uncertain payoff, and risk of prosecution make Internet-connected ICS devices unattractive targets, when compared with IT and IoT assets. Why spend months developing an exploit for hundreds of devices when off-the-shelf exploits exist for target populations of hundreds of thousands?

b) Outlook: Larger compute and memory resources on IIoT devices make them more attractive hosts for parasitic malware, creating viable financial incentives. Additionally, the full operating systems running on these devices and development board availability simplify malware development and testing. Similarly, these platforms make it easier to modify and test existing malware (e.g., Mirai), reducing uncertainty and limiting the cost to develop an exploit.

While attacks against ICS have been prosecuted in the past, many large-scale cybercrimes do not directly affect the host, and those responsible have not been aggressively pursued [12]. Cybercrime making use of ICS devices may similarly be low risk, provided the parasitic malware does not adversely affect the industrial process.

Devices with greater compute and memory resources and commodity operating systems satisfy several missing enablers in the *Attacker incentives* category of Table II. Given the currently-limited pursuit of cybercriminals targeting IoT devices, we also consider the *low consequence to attacker* enabler to be partially met.

C. Attacker tools and resources

a) Current situation: Wormable ICS malware has yet to be packaged for an unskilled attacker to deploy [69], [37]. Subject matter experts developing such proof-of-concept malware demonstrate that such development is plausible, but admit it is not currently practical. In contrast, off-the-shelf tools exist for attacking IT and IoT at scale. Further, numerous examples show that an adversary interested in attacking industry at scale need not target ICS devices directly, as off-the-shelf malware

for Windows-infrastructure is sufficient to shut down industrial processes for extended periods [52], [75].

b) Outlook: As discussed above, devices with Linux-based or other widely used, open source operating systems provide attackers with a much greater opportunity to develop or use existing exploits, which is exacerbated by the slow patch cadence demonstrated for Internet-connected ICS devices. The increased homogeneity and use of commodity operating systems simplifies the effort to develop or modify exploits targeting ICS devices, satisfying the remaining enablers in the *Attacker tools and resources* category of Table II.

D. Summary

In answer to the questions posed at the beginning of this section, our model demonstrates:

1) The current Internet-connected ICS population is not a target for large-scale cybercrime due to its fragmented nature, the high cost to develop and test exploits, and the unpredictable monetisation and consequences of attack.
2) Our model predicts that adversarial interest in Internet-connected ICS devices will grow along with trends toward more connectable devices; homogeneous hardware, software, and development environments; greater onboard resources; and commodity operating systems.

VIII. Conclusions

In this paper, we presented the first multi-year, longitudinal study of the Internet-connected ICS population. We demonstrated the ability to passively track thousands of ICS devices over many years, showing that the population is growing, that device owners rarely install software updates, and that most devices are continuously connected. Despite these vulnerabilities, we find that the fragmentation of the ICS community, the high cost to develop and test exploits, and the unpredictable monetisation and consequences of attacking ICS make them an unattractive target for the cybercrime community, especially given the continued vulnerability of the larger and more homogeneous IoT population. Our conclusions are supported by technical and criminological analyses of the cybercrime community, using honeypots, malware tracking, and hacker forums to demonstrate that there is currently little competence or interest in the ICS domain amongst cybercriminals.

To explain this limited interest, we introduced a security economics model for characterising and predicting large-scale adversarial interest in Internet-connected populations. We developed the model by studying successful, large-scale attacks and empirically verified it using our end-to-end study of the Internet-connected ICS threat landscape. The model provides a concise explanation for the apparent reluctance of cybercriminals to target ICS devices.

While the current ICS device population may not be an attractive target for large-scale attacks, we surveyed ongoing and expected changes to the industrial environment that will boost the number of connectable devices and will move industry toward more homogeneous hardware, software, and development environments, greater compute and memory resources, and commodity operating systems. Our model predicts that these changes bring the ICS community in line with other, targeted populations, such as IoT; therefore, it is reasonable to expect greater attention from the cybercrime community directed toward ICS. Further, even if ICS devices are not directly targeted, convergence with IoT creates a new risk of being swept up in large-scale attacks targeting the billions of non-industrial IoT devices.

Data availability and ethics

Public release of Shodan, Censys, and honeypot data is prohibited by their respective Terms of Service; however, scripts used to query and process Shodan and Censys data will be available via the CCCC [21]. Similarly, CrimeBB access can be requested through the CCCC [21].

We followed our institution's ethical research policy throughout and obtained IRB approval for searching and processing CrimeBB data as well as scanning IP addresses suspected of hosting Mirai.

For CrimeBB, our primary concerns were violating user privacy (e.g., deanonymising identities) and loss of data control (e.g., allowing another organisation to use the data in a way that violates the CCCC's legal framework [21]). To mitigate these concerns, we restricted our search terms, stored and processed the data on a specified server, and followed strict guidelines to avoid privacy violations, such as specifically avoiding correlations between different forums and avoiding direct quotations from posts.

For scanning, our primary concern was adverse physical effects on ICS (e.g., causing devices to reset or slow down [2]). To mitigate these concerns, we followed the recommended practices in Durumeric *et al.* [76] and used the ZGrab scanner,[4], the tool use by Censys [1], [10]. We reasoned that as Censys performs Internet-wide scanning on a near-daily basis against every IP address and port number we planned to scan, we would not cause any additional adverse effects. Where possible, we used Censys' data directly and only performed our own scans when better-than-daily resolution was necessary to eliminate DHCP churn. We used a superset of the blocklists maintained by Censys, the CCCC, and the OARC-DNS 'don't probe' list [77]. We also hosted a website at the scanner's IP address that provided an explanation of our research and contact information for individuals or organisations to lodge complaints or request to be added do our blocklist. We received no complaints or requests to be added to the blocklist.

Acknowledgment

This work was partly supported by the Gates Cambridge Trust and the EPSRC [grant number EP/M020320/1]. The opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect those of any of the funders. We thank Alexander Vetterl, Ross Anderson, Éireann P. Leverett for their helpful feedback and guidance on earlier drafts of this paper.

[4]https://github.com/zmap/zgrab2

APPENDIX
COMPARING LARGE-SCALE ATTACKS AGAINST INTERNET-CONNECTED DEVICES

Table VI compares the technical details of the case studies in Section III-A. The table summarises the information used to determine which enablers were met when developing the security economics model and populating Table II.

REFERENCES

[1] 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," *Conference on Privacy, Security and Trust (PST)*, 2016.

[2] M. Niedermaier, J.-O. Malchow, F. Fischer, D. Marzin, D. Merli, and V. Roth, "You snooze, you lose: Measuring PLC cycle times under attacks," in *USENIX Workshop on Offensive Technologies (WOOT)*, 2018.

[3] É. P. Leverett, "Quantitatively assessing and visualising industrial system attack surfaces," *University of Cambridge MPhil Thesis*, 2011. [Online]. Available: https://perma.cc/83Z9-Q5J9

[4] B. Radvanovsky and J. Brodsky, "Project SHINE (SHodan INtelligence Extraction)," *Infracritical Technical Report*, 2014. [Online]. Available: https://perma.cc/HA8J-5SNZ

[5] C. D. Schuett, "Programmable logic controller modification attacks for use in detection analysis," *Air Force Institute of Technology Master's Thesis*, 2014. [Online]. Available: https://perma.cc/Q7GZ-8JQM

[6] P. M. Williams, "Distinguishing Internet-facing ICS devices using PLC programming information," *Air Force Institute of Technology Master's Thesis*, 2014. [Online]. Available: https://perma.cc/W7YL-7J7T

[7] M. Dahlmanns, J. Lohmöller, I. B. Fink, J. Pennekamp, K. Wehrle, and M. Henze, "Easing the conscience with OPC UA: An Internet-wide study on insecure deployments," *ACM Internet Measurement Conference IMC)*, Oct. 2020.

[8] Shodan, "Shodan," 2020. [Online]. Available: https://www.shodan.io/

[9] Censys, "Censys," 2020. [Online]. Available: http://censys.io/

[10] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman, "A search engine backed by Internet-wide scanning," *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2015.

[11] G. Barbieri, M. Conti, N. O. Tippenhauer, and F. Turrin, "Sorry, Shodan is not enough! Assessing ICS security via IXP network traffic analysis," *arXiv preprint arXiv:2007.01114*, Jul. 2020. [Online]. Available: http://arxiv.org/abs/2007.01114

[12] R. Anderson, C. Barton, R. Bohme, R. Clayton, C. Ganan, T. Grasso, M. Levi, T. Moore, and M. Vasek, "Measuring the changing cost of cybercrime," *Workshop on the Economics of Information Security (WEIS)*, 2019. [Online]. Available: https://perma.cc/ZKB4-C6Q6

[13] S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security," *Conference of the IEEE Industrial Electronics Society (IECON)*, 2011.

[14] Electricity Information Sharing and Analysis Center, "Analysis of the cyber attack on the Ukrainian power grid," *Electricity Information Sharing and Analysis Center (E-ISAC)*, 2016. [Online]. Available: https://perma.cc/74V7-TN7J

[15] "CRASHOVERRIDE: Analysis of the threat to electric grid operations," *Dragos Inc.*, 2017. [Online]. Available: https://perma.cc/E7K5-9T8M

[16] F-Secure, "News from the lab archive," *F-Secure*, 2009. [Online]. Available: https://perma.cc/A2DK-3JUJ

[17] M. Bowden, "The worm that nearly ate the Internet," *The New York Times*, Jun. 2019. [Online]. Available: https://perma.cc/WU7P-LBAD

[18] B. Nahorney, "The Downadup codex: A comprehensive guide to the threat's mechanics," *Symantec*, 2009. [Online]. Available: https://perma.cc/3ACC-U5R4

[19] G. Keizer, "Conficker cashes in, installs spam bots and scareware," Apr. 2009. [Online]. Available: https://www.computerworld.com/article/2524137/conficker-cashes-in–installs-spam-bots-and-scareware.html

[20] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou, "Understanding the Mirai botnet," *USENIX Security Symposium (USENIX Security)*, 2017. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-antonakakis.pdf

[21] C. C. Centre, "Description of available datasets," 2020. [Online]. Available: https://www.cambridgecybercrime.uk/datasets.html

[22] Cybersecurity and Infrastructure Security Agency, "BrickerBot permanent denial-of-service attack," 2017. [Online]. Available: https://www.us-cert.gov/ics/alerts/ICS-ALERT-17-102-01A

[23] Radware, "BrickerBot results in PDoS (Permanent Denial of Service) Attacks," 2017. [Online]. Available: https://perma.cc/3MTJ-7GWL

[24] C. Cimpanu, "BrickerBot author retires claiming to have bricked over 10 million IoT devices," 2017. [Online]. Available: https://perma.cc/6WUV-99VG

[25] Symantec, "Equation: Advanced cyberespionage group has all the tricks in the book, and more," *Symantec*, 2015. [Online]. Available: https://perma.cc/S9MF-6BK5

[26] E. Nakashima and C. Timberg, "NSA officials worried about the day its potent hacking tool would get loose. Then it did." 2017. [Online]. Available: https://perma.cc/V8D9-GCHS

[27] Symantec, "What you need to know about the WannaCry ransomware," *Symantec*, 2017. [Online]. Available: https://perma.cc/J6HD-HFYR

[28] ——, "Petya ransomware outbreak: Here's what you need to know," *Symantec*, 2017. [Online]. Available: https://perma.cc/7JP5-9D6D

[29] Kaspersky, "New Petya / NotPetya / ExPetr ransomware outbreak," *Kaspersky*, 2017. [Online]. Available: https://perma.cc/BM7P-P2R9

[30] GReAT, "Schroedinger's Pet(ya)," *Kaspersky*, 2017. [Online]. Available: https://perma.cc/U6HH-HDZJ

[31] Cybersecurity and Infrastructure Security Agency, "Malware Initial Findings Report (MIFR) - 10130295," Cybersecurity and Infrastructure Security Agency (CISA), Tech. Rep., 2017. [Online]. Available: https://perma.cc/QV8W-QAKM

[32] B. Collier, D. R. Thomas, R. Clayton, and A. Hutchings, "Booting the booters: Evaluating the effects of police interventions in the market for denial-of-service attacks," *Internet Measurement Conference (IMC)*, 2019.

[33] D. R. Thomas, R. Clayton, and A. R. Beresford, "1000 days of UDP amplification DDoS attacks," *APWG Symposium on Electronic Crime Research (eCrime)*, 2017.

[34] U.S. Department of Justice, "Former operator of illegal booter services pleads guilty to conspiracy to commit computer damage and abuse," Feb. 2019. [Online]. Available: https://perma.cc/HC5G-KYL4

[35] B. Krebs, "DDoS-for-hire service webstresser dismantled," 2018. [Online]. Available: https://perma.cc/586G-TKGG

[36] S. Pastrana and G. Suarez-Tangil, "A first look at the crypto-mining malware ecosystem: A decade of unrestricted wealth," *Internet Measurement Conference (IMC)*, 2019. [Online]. Available: https://dl.acm.org/authorize?N695072

[37] D. Formby, S. Durbha, and R. Beyah, "Out of control: Ransomware for industrial control systems," *RSA conference*, 2017. [Online]. Available: https://perma.cc/XD8V-LP5M

[38] R. Anderson, *Security Engineering*. New York: John Wiley & Sons, 2008.

[39] B. Krebs, "Mirai botnet authors avoid jail time," 2018. [Online]. Available: https://perma.cc/ECB4-T8FC

[40] A. Senpai, "[FREE] World's largest net: Mirai botnet, client, echo loader, CNC source code release," Sep. 2016. [Online]. Available: https://perma.cc/SV6Z-LX92

[41] O. Gasser, Q. Scheitle, B. Rudolph, C. Denis, N. Schricke, and G. Carle, "The amplification threat posed by publicly reachable BACnet devices," *Journal of Cyber Security and Mobility*, vol. 6, no. 1, 2017.

[42] L. Rist, J. Vestergaard, D. Haslinger, A. Pasquale, and J. Smith, "CONPOT ICS/SCADA honeypot," 2019. [Online]. Available: http://conpot.org/

[43] S. J. Murdoch, "Hot or not: Revealing hidden services by their clock skew," *ACM Conference on Computer and Communications Security (CCS)*, 2006.

[44] J. Zhang, A. R. Beresford, and I. Sheret, "SENSORID: Sensor calibration fingerprinting for smartphones," *IEEE Symposium on Security and Privacy (SP)*, 2019.

[45] G. Guo, J. Zhuge, M. Yang, G. Zhou, and Y. Wu, "A survey of industrial control system devices on the Internet," *International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)*, Dec. 2018.

TABLE VI
COMPARISON OF TECHNICAL DETAILS OF CASE STUDIES IN SECTION III-A

		Conficker	Mirai	Brickerbot	WannaCry	NotPetya	Webstresser	Cryptomining
Potential use	Ransomware	✓			✓	✓		
	Cryptojacking		✓					✓
	DDoS		✓				✓	
	Spam and advertising scams	✓	✓					
	Disruption/destruction	✓		✓	✓	✓		
Botnet-based		✓	✓	✓				✓
Method of infection	Exploit functionality						✓	✓
	Exploit vulnerability	✓	✓	✓	✓	✓		✓
	Unnecessarily-open ports		✓	✓				
	Weak passwords	✓	✓	✓				
	Websites/javascript							✓
	Phishing/weaponised files	✓				✓		
Persistence	Host-based payload	✓	✓	✓	✓	✓		✓
	Host-based execution							✓
	Direct manipulation						✓	
	Multi-variant/adaptable	✓	✓		✓		✓	✓

[46] R. Automation, "Industrial Security: Protecting networks and facilities against a fast-changing threat landscape," Tech. Rep., 2016.

[47] A. Adams and M. A. Sasse, "Users are not the enemy," *Communications of the ACM*, vol. 42, no. 12, 1999.

[48] C. Herley, "More is not the answer," *IEEE Symposium on Security and Privacy (SP)*, vol. 12, no. 1, 2014.

[49] ——, "So long, and no thanks for the externalities: The rational rejection of security advice by users," *Workshop on New Security Paradigms (NSPW)*, 2009.

[50] T. Moore and R. Anderson, "Economics and internet security: A survey of recent analytical, empirical and behavioral research," *Harvard University*, vol. TR-03-11, 2011. [Online]. Available: https://perma.cc/AH7R-XRYT

[51] "TRISIS malware: Analysis of safety system targeted malware," *Dragos Inc.*, 2017. [Online]. Available: https://perma.cc/K9EM-CABV

[52] Hydro, "Cyber-attack on Hydro," 2019. [Online]. Available: https://perma.cc/Z4NV-W9WP

[53] A. Vlad, S. Obermeier, and D.-Y. Yu, "ICS threat analysis using a large-scale honeynet," *International Symposium for ICS & SCADA Cyber Security Research*, 2015.

[54] A. Belqruch and A. Maach, "SCADA security using SSH honeypot," *International Conference on Networking, Information Systems & Security*, 2019.

[55] B. Radvanovsky, "Project RUGGEDTRAX SCADA/ICS analysis," *Infracritical Technical Report*, 2015. [Online]. Available: https://perma.cc/7AN4-KR8K

[56] P. Ferretti, M. Pogliani, and S. Zanero, "Characterizing background noise in ICS traffic through a set of low interaction honeypots," *ACM Workshop on Cyber-Physical Systems Security & Privacy (CPS-SPC)*, 2019.

[57] M. Dodson, M. Vingaard, and A. R. Beresford, "Using global honeypot networks to detect targeted ICS attacks," *12th International Conference on Cyber Conflict (CyCon)*, 2020.

[58] S. Pastrana, D. R. Thomas, A. Hutchings, and R. Clayton, "CrimeBB: Enabling cybercrime research on underground forums at scale," *Conference on the World Wide Web (WWW)*, 2018.

[59] A. Hutchings and S. Pastrana, "Understanding eWhoring," *IEEE European Symposium on Security and Privacy (SP)*, 2019.

[60] A. Nochvay, "Security research: CODESYS runtime, a PLC control framework," Kaspersky ICS CERT, Tech. Rep., 2019. [Online]. Available: https://perma.cc/325P-N7AV

[61] Cybersecurity and Infrastructure Security Agency, "3S CoDeSys vulnerabilities," 2013. [Online]. Available: https://perma.cc/F8W4-7H75

[62] D. McFarlane, S. Ratchev, A. Thorne, A. K. Parlikad, L. de Silva, B. Schönfuß, G. Hawkridge, G. Terrazas, and Y. Tlegenov, "Digital manufacturing on a shoestring: Low cost digital solutions for SMEs," *International Workshop on Service Orientation in Holonic and Multi-Agent Manufacturing*, 2019.

[63] M. Wollschlaeger, T. Sauter, and J. Jasperneite, "The future of industrial communication: Automation networks in the era of the internet of things and Industry 4.0," *IEEE Industrial Electronics Magazine*, 2017.

[64] Huawei, "IoT: Driving verticals to digitization," 2019. [Online]. Available: https://www.huawei.com/minisite/iot/en/

[65] E. Protalinski, "Android passes 2.5 billion monthly active devices," May 2019. [Online]. Available: https://perma.cc/2D3D-S7KN

[66] Gartner, "PCs installed base worldwide 2013-2019," 2019. [Online]. Available: https://perma.cc/CSV6-5VF5

[67] J. Clover, "Apple Now Has 1.3 Billion Active Devices Worldwide," 2018. [Online]. Available: https://perma.cc/X599-Q3N8

[68] Gartner, "Global connected IoT devices by type 2017 and 2018," 2019. [Online]. Available: https://perma.cc/6SW2-4LH9

[69] R. Spenneberg, M. Brüggemann, and H. Schwartke, "PLC-Blaster: A worm living solely in the PLC," *Black Hat Asia*, 2016. [Online]. Available: https://perma.cc/XWU5-TZ7L

[70] J. Klick, S. Lau, D. Marzin, J.-O. Malchow, and V. Roth, "Internet-facing PLCs – a new back orifice," *Blackhat USA*, 2015. [Online]. Available: https://perma.cc/XK4N-UPV4

[71] J. Slay and M. Miller, "Lessons learned from the Maroochy water breach," *International Conference on Critical Infrastructure Protection*, vol. 253, 2007.

[72] M. D. Abrams and J. Weiss, "Malicious control system cyber security attack case study – Maroochy water services, Australia," 2008. [Online]. Available: https://perma.cc/CTX9-A673

[73] U.S. Department of Justice, "Former systems administrator sentenced to prison for hacking into industrial facility computer system," Feb. 2017. [Online]. Available: https://perma.cc/PWP7-SPKA

[74] Federal Bureau of Investigation, "Attacks on Arkansas power grid," *Federal Bureau of Investigation (FBI)*, 2015. [Online]. Available: https://perma.cc/DNU6-EM65

[75] L. Frost, N. Tajitsu, E. Auchard, G. Guillaume, and C. Pitas, "Renault-Nissan resumes nearly all production after cyber attack," *Reuters*, 2017. [Online]. Available: https://perma.cc/Y2J7-8PXU

[76] Z. Durumeric, E. Wustrow, and J. A. Halderman, "Zmap: Fast Internet-wide scanning and its security applications," *USENIX Security Symposium (USENIX Security)*, 2013. [Online]. Available: https://zmap.io/paper.pdf

[77] DNS-OARC, "OARC's DNS don't-probe list," 2020. [Online]. Available: https://www.dns-oarc.net/oarc/services/dontprobe