Paper presented at *IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*
Montreal, Canada, Oct. 8-13, 2017

A First Empirical Look on Internet-scale Exploitations of IoT Devices

Mario Galluscio*, Nataliia Neshenko*, Elias Bou-Harb*, Yongliang Huang*
Nasir Ghani†, Jorge Crichigno‡, Georges Kaddoum§
*Cyber Threat Intelligence Lab, Florida Atlantic University, United States
†Florida Center for Cybersecurity and University of South Florida, United States
‡Northern New Mexico College, United States
§École de Technologie Supérieure (ÉTS), Université du Québec, Canada

***Abstract*—Technological advances and innovative business models led to the modernization of the cyber-physical concept with the realization of the Internet of Things (IoT). While IoT envisions a plethora of high impact benefits in both, the consumer as well as the control automation markets, unfortunately, security concerns continue to be an afterthought. Several technical challenges impedes addressing such security requirements, including, lack of empirical data related to various IoT devices in addition to the shortage of actionable attack signatures.**

In this paper, we present what we believe is a first attempt ever to comprehend the severity of IoT maliciousness by empirically characterizing the magnitude of Internet-scale IoT exploitations. We draw upon unique and extensive darknet (passive) data and develop an algorithm to infer unsolicited IoT devices which have been compromised and are attempting to exploit other Internet hosts. We further perform correlations by leveraging active Internet-wide scanning to identify and report on such IoT devices and their hosting environments. The generated results indicate a staggering 11 thousand exploited IoT devices that are currently in the wild. Moreover, the outcome pinpoints that IoT devices embedded deep in operational Cyber-Physical Systems (CPS) such as manufacturing plants and power utilities are the most compromised. We concur that such results highlight the wide-spread insecurities of the IoT paradigm, while the actionable generated inferences are postulated to be leveraged for prompt mitigation as well as to facilitate IoT forensic investigations using real empirical data.

I. Introduction

The Internet of Things (IoT) paradigm represents advances in computing power, electronics miniaturization and network bandwidth interconnection, which have armed physical objects with the ability to collect, process and act upon various types of information. The widespread deployment of IoT devices ranging from refrigerators to light bulbs to Internet-controlled insulin pumps and mining equipment brings forward a plethora of benefits in an effort to improve various aspects of our everyday life [1]. In fact, home automation, powered by the IoT, have provided essential support for elderly or disabled residents, while wearable health monitoring devices have increased the quality of medical service and enabled the real-time provisioning of medication [2]. Further, hazardous industrial plants, which might pose various health and safety risks to their employees, have indeed adopted the IoT to notify their personnel, in near real-time, about critical incidents to avoid [3]. Additionally, many case studies report the significant impact of IoT on disaster and crime prevention, reduction in traffic congestion and parking time, and improvement of emergency services [4, 5]. The IoT paradigm also addresses the issue of water and energy consumption, thus reducing the cost for homeowners, companies, and entire cities [6]. Indeed, this prominent notion holds a commitment to transform the majority of business models, and to improve efficiency, service levels, and customer satisfaction.

Indisputable benefits proposed by the IoT paradigm, in both, consumer environments as well as Cyber-Physical Systems (CPS) realms (i.e., manufacturing plants, power utilities, building automation, etc.), are nevertheless coupled with serious security flaws [7]. Time-to-market and cost considerations along with the scarcity of related legislation have stimulated manufacturers to design and produce potentially insecure IoT devices, leaving an open door for future exploitation. This practice continues to enable exposure of user-centric information and data such as unprotected video streaming of baby monitors [8] and sensitive cryptographic primitives [9]. Moreover, poorly designed devices can quickly be recruited into malicious botnets by allowing the execution of arbitrary commands or re-programming of device firmwares [10]. Given the large-scale deployment of IoT devices, such vulnerabilities could affect the security and the resiliency of the entire Internet space. The latter has recently received tremendous media exposure. For instance, an IoT smart refrigerator have participated in launching massive phishing campaigns [11], while the Mirai malware [12] and its extension Hajime [13] have demonstrated how unsecured IoT devices can serve as entry points for conducting orchestrated Denial of Service (DoS) attacks and other malicious misdemeanors. Undoubtedly, such and other security breaches largely challenge the trust level in the IoT paradigm, hindering its wide-spread implementation in various sectors and critical infrastructure.

While the security and networking research and operational communities are undertaking several steps towards the goal of immuning the IoT paradigm from participating in or being the target of debilitating cyber threats, significant security weaknesses continue to exist, impeding IoT's ability to achieve its intended goals. In fact, the preliminary task of Internet-scale characterization of the magnitude of IoT exploitations is still relatively obscured. This is especially factual when attempting to assess the maliciousness of the IoT paradigm deployed in critical CPS. Indeed, one can not devise effective IoT-relevant detection capabilities without initially understanding the scale of this issue. Nevertheless, several obstacles hampers the realization of the latter task, including, lack of visibility into local IoT realms, which prevents accessing and analyzing IoT-relevant empirical data and the general insufficiency of IoT-specific attack signatures [14].

In this paper, we take a first step towards comprehending the severity and scale of IoT maliciousness. To this end, we scrutinize Internet-scale empirical unsolicited traffic to identify exploited IoT devices and their hosting environments. In summary, we frame the contributions of this paper as follows:

- Proposing an innovative approach which fuses extensive passive measurements with results from Internet-wide scanning to shed the light on compromised IoT devices. In this context, it is important to note that the approach is also capable of generating IoT-relevant malicious empirical data and attack signatures, which we hope to be shared with the research community at large to facilitate further forensic investigation and advanced IoT data analytics.

- Developing and operating a probing inference algorithm that is capable of fingerprinting malicious activities which are generated from exploited IoT devices.

- Evaluating the proposed approach by employing 130 GB of real darknet data and reporting on more than 11 thousand exploited IoT devices that are currently in the wild. Such results could indeed be leveraged for immediate mitigation by local IoT operators.

The road-map of this paper is as follows. In the next section, we review related works on various concerned IoT security topics and demonstrate the added value of the proposed approach. In Section III, we present the proposed methods and techniques for inferring Internet-scale unsolicited IoT devices. In Section IV, we discuss the obtained results, while in Section V, we summarize our contributions and pinpoint several research topics that aim at paving the way for future work in this impactful IoT security research area.

II. Related Work

In this section, we review the recent literature on various concerned topics, including, IoT vulnerabilities, empirical characterization of devices and IoT data capturing initiatives.

The majority of IoT security research work has been dedicated to addressing IoT security flaws. To this end, Ur et al. [15] analyzed IoT access control by studying numerous types of home automation devices. The authors investigated ownership rules, roles, and monitoring capabilities, which unveiled various issues such as revoked access permissions, insufficient auditing capabilities, and usability flaws. In alternative work, Ho et al. [16] explored IoT smart lock systems and demonstrated how network architectures, trust models, and malicious replay traffic could unlock doors, allowing unauthorized physical access. Further, the authors noted that most of the investigated devices lacked access to elaborative logging procedures, rendering it impossible for users to know who have accessed their devices and what type of traffic was generated. Moreover, Ronen and Shamir [9] illustrated information leakage attempts by simulating an attack on a set of smart LED light bulbs. The authors were able to extract sensitive information from an air-gapped infrastructure, including encryption primitives and passwords. In a different work, Sachidananda et al. [17] conducted port scanning, process enumeration, and vulnerability scans of numerous IoT devices. Their investigations unveiled that a plethora of devices have open ports, allowing attackers to obtain information related to their vulnerabilities by means of fingerprinting their deployed operating systems and device types/firmwares. In this work, we extend such research efforts by uniquely leveraging and correlating empirical data to understand, macroscopically, the magnitude of Internet-scale IoT exploitations.

In the area of empirical measurements for device characterization, Cui and Stolfo [18] executed a large-scale active probing of the Internet space to uncover close to half a million vulnerable embedded devices. In more recent works, Costin et al. [19] statically analyzed more than 30 thousand firmware images derived from embedded devices to shed the light on their insecurities, while Fachkha et al. [20] conducted passive measurements to analyze attackers' intentions when targeting protocols of Internet-facing CPS. In a similar work, Bodenheim et al. [21] evaluated the Shodan service, a search engine for Internet-connected devices, in its capability to scan and index online industrial control systems. In contrast, in this work, we particularly address the IoT paradigm and develop an algorithm operating an passive darknet data to not only infer compromised IoT devices but also pinpoint their activities as they attempt to exploit other Internet hosts.

In the context of IoT data capturing initiatives, the first IoT tailored honeypot, namely IoTPOT, was designed and deployed by Pa et al. [22]. IoTPOT emulates telnet services of various IoT devices running on different CPU architectures. IoTPOT demonstrated its capability to capture various types of malware samples for the sake of performing in-depth analysis of IoT targeted attacks. In alternative work, Guarnizo et al. [23] presented the Scalable high-Interaction Honeypot (SIPHON) platform for IoT devices. The authors demonstrated how by leveraging worldwide wormholes and few physical devices, they were able to mimic various IoT devices on the Internet and to attract massive malicious traffic. The authors further characterized such traffic by elaborating on attackers' frequency and their employed protocols. Auxiliary, several attempts to fingerprint IoT devices were executed. For instance, very recently, Meidan et al. [24] leveraged network traffic analysis to classify IoT devices connected to an organization's network, by applying techniques rooted in machine learning supervised data classification. In contrary, our analysis draws upon pure passive darknet data to present a first look on compromised IoT devices in both, consumer and CPS environments. Further, the proposed approach is envisioned to be leveraged to extract raw empirical data related to various IoT devices to support supplementary forensic investigations as well as capture tangible IoT attacks signatures. The latter two artifacts are currently lacking in the academic as well as the security operations communities.

III. Proposed Approach

In this section, we detail our proposed approach which aims at leveraging passive empirical measurements to infer unsolicited Internet-scale IoT devices.

A. Exploiting Darknet Data

Having access to empirical IoT data is indeed quite challenging. Several hurdles confirm the latter, including, the lack of visibility into local IoT realms due to logistic and privacy concerns, the general scarcity of malicious empirical data related to unsolicited IoT devices [25], and the lack of tangible IoT-specific attack signatures [14]. To this end, complementary methods ought to be explored; without access to tangible IoT empirical data, the notion of maliciousness in this context can not be elaborated. In this work, we uniquely exploit passive measurements rendered by analyzing darknet data to achieve the latter task. A darknet (also commonly referred to as a network telescope) is a set of routable and allocated yet unused IP addresses [26, 27]. It represents a partial view of the entire Internet address space. From a design perspective, a darknet is transparent and indistinguishable compared with the rest of the Internet space. From a deployment perspective, it is rendered by network sensors that are implemented and dispersed on numerous strategic points throughout the Internet. Such sensors are often distributed and are typically hosted by various global entities, including Internet Service Provides (ISPs), academic and research facilities, and backbone networks. The aim of a darknet is to provide a lens on Internet-wide unsolicited traffic; since darknet IP addresses are unused, any traffic targeting them represents anomalous traffic. Such traffic (i.e., darknet data) could be leveraged to generate various cyber threat intelligence, including inferences and insights related to probing activities [27]. Such events are indeed the very first signs of infections and propagation [28]. In this context, a darknet is capable of capturing some of the probes of Internet-scale infected hosts. Recall, that the probing machine, while spraying its probes, can not avoid the darknet as it does not have any knowledge about its existence. Further, it is known that it is extremely rare if not impossible for a probing source to have any capability dedicated to such avoidance [29]. Thus, the proposed approach endeavors to scrutinize darknet data to infer probing activities which are generated from Internet-scale unsolicited IoT devices as an indicator of their exploitations. We are fortunate to have access to real-time darknet data from a /8 network telescope through our collaboration with the Center for Applied Internet Data Analysis (CAIDA)[1].

B. Probing Inference

To infer probing activities from darknet data, we present Algorithm 1, which exploits flow-based parameters.

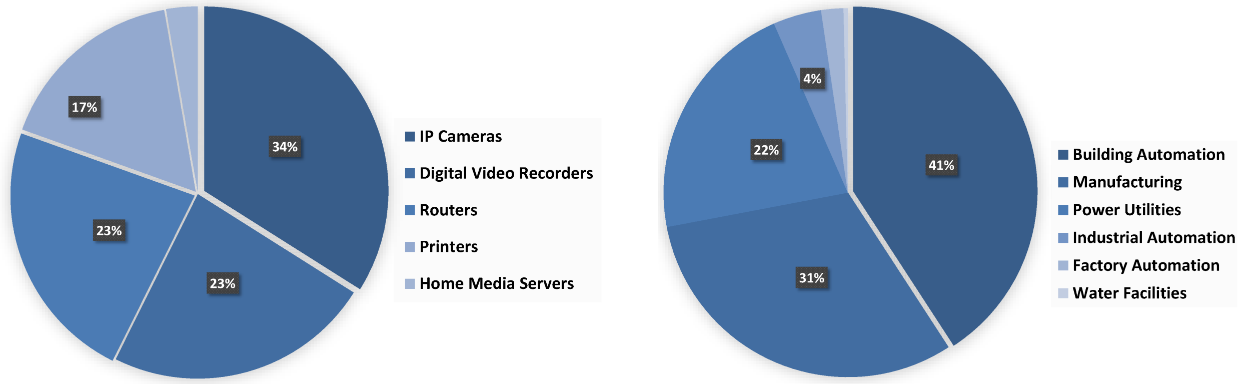
Algorithm 1 Probing Inference Algorithm

```
Input: A set (F) of unique darknet flows (f),
Each flow f contains packet count (pkt_cnt) and rate (rate)
Tw: Time window
Pth: Packet threshold
Rth: Rate threshold,
Tn: Time of packet number n in a flow
pkt: Packet
Output: Probing flag, Pr_flag

for Each f in F do
    pkt_cnt ← 0
    T1 ← pkt_gettime()
    Tf ← T1 + Tw
    while pkt in f do
        Tn= pkt_gettime()
        if Tn < Tf then
            pkt_cnt ← pkt_cnt + 1
        end if
    end while
    rate ← $\frac{pkt\_cnt}{Tw}$
    if pkt_cnt > Pth & rate > Rth then
        Pr_flag() ← 1
    end if
end for
```

Algorithm 1 operates on darknet flows, which are defined by a series of consecutive packets sharing the same source IP address. The algorithm counts the number of packets per flow to measure the rate of the suspicious activities within a certain time window (*Tw*). If the flow packet count (*pkt_cnt*) is beyond a specific threshold, the flow is deemed as a probe. To this end, we employ the packet count threshold from [30], defined by 64 probed darknet addresses on the same port. Please note, that typically, the probing engine would have also required and established a rate threshold (*Rth*). Nevertheless, we do not enforce one here, to enable the algorithm to infer very low rate, possible stealthy activities. Indeed, the

[1] http://www.caida.org/home/



(a) Consumer Sectors

(b) CPS Sectors

Fig. 1: Distribution of IoT devices deployed in consumer and CPS realms

approach embedded in Algorithm 1 would fingerprint Internet-scale probing traces. Please note that from a performance perspective, when implemented "on the fly" on the darknet data stream using the `C libpcap` library [31], the developed inference algorithm can process close to 10,000 flows in approximately 1 minute (average throughput of 150 flows/sec).

C. Correlation with Active Measurements

To infer probes that have been specifically generated from exploited IoT devices, one needs to fingerprint IoT generated traffic. Indeed, such task is currently an open research problem and very few endeavors (if not nil) have addressed it. While we continue to investigate this problem using real empirical data, in this work, we approach this issue from a different perspective by leveraging active measurements. This entitles executing Internet-wide scanning, capturing the results and filtering the replies from the destinations based on their nature. Fortunately, the Shodan service [32] performs the latter and indexes online IoT devices. To this end, we leverage Shodan's available database of IoT devices, which are deployed in both, consumer environments as well as in CPS realms. In total, we retrieve more than 900 thousand online IoT devices and correlate them using their source IP information with IP data retrieved by conducting probing analysis of the darknet data as previously mentioned. Thus, one core outcome of such proposed approach are inferred Internet-scale exploited IoT devices, which are attempting to scan other Internet hosts (to fingerprint or exploit them). Auxiliary outcomes, which are currently work in progress, include (*i*) accessing IoT malicious empirical data which can be extracted from darknet data and shared at large with the research community to facilitate forensic investigations of IoT-relevant data, and (*ii*) generating tangible IoT-specific attack signatures, using tools such as `ssdeep` [33], which can be deployed at local IoT realms to aid with the task of prompt mitigation.

IV. Empirical Evaluation

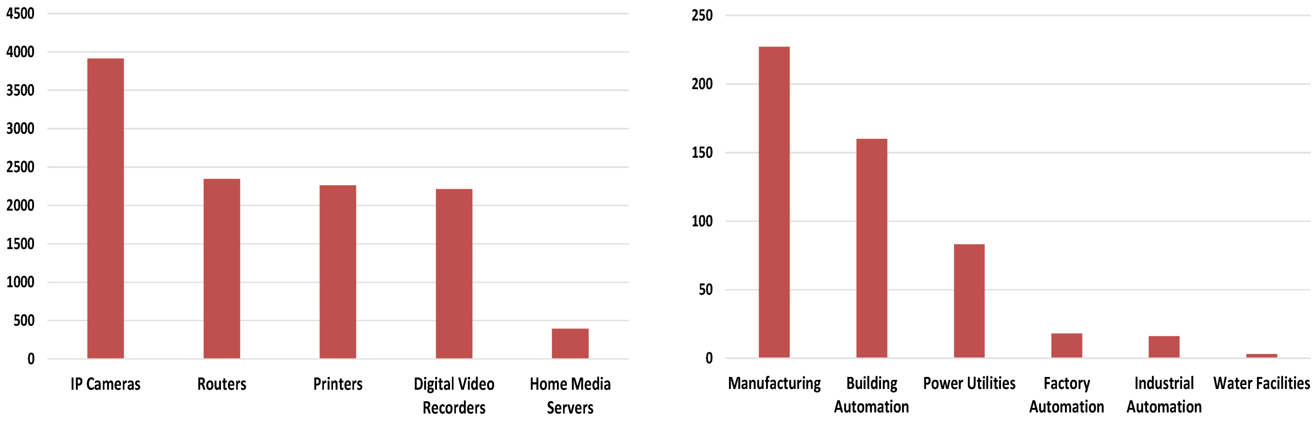
In this section, we employ the proposed approach of Section III to elaborate on the generated insights and inferences. We exploit close to 130 GB of darknet data that was recently retrieved in the month of June 2017. We executed queries using the Shodan service to index online IoT devices, which are deployed in both, consumer and CPS environments. On one hand, for the IoT consumer market, we focused on 5 categories, namely, IoT cameras, Digital Video Recorders (DVRs), routers, printers and home media servers. We chose the latter as they seemed to be widely deployed and well adopted in addition to showing a history of exploitation (as in the case of the Mirai malware abusing IP cameras and DVRs). In total, we have indexed 862,014 IoT consumer devices that were online at the time of writing of this paper. On the other hand, from the CPS perspective, we focused on 6 sectors as summarized in Table I.

CPS Sector	Protocol
Building Automation	`BACnet, Tridium`
Factory Automation	`CoDeSys`
Industrial Automation	`Red Lion Controls, Siemens-S7, MELSEC-Q`
Manufacturing	`OMRON, EtherNet/IP`
Power Utilities	`Modbus`
Water Facilities	`DNP3`

TABLE I: IoT devices related to various CPS deployments

In total, we were able to index 72,554 IoT devices which have been deployed and operated in those CPS sectors. Figures 1(a) and 1(b) illustrate the distribution of such IoT devices in their numerous corresponding realms. It can be extracted that IoT devices related to IP cameras, DVRs and routers are quite well deployed. Further, it can be inferred that IoT devices in building automation facilities, manufacturing plants and power utilities render the majority of the IoT deployments.

We proceed by invoking the inference algorithm and the correlation procedure as briefed in Sections III-B and III-C. The outcome uncovers Internet-scale compromised IoT devices in various sectors. Overall, we were able to infer 11,122 exploited IoT devices related to the consumer sector, while the results further disclosed 510 vulnerable IoT devices in critical CPS sectors. Figures 2(a) and 2(b) illustrate the distribution



(a) Consumer Sectors

(b) CPS Sectors

Fig. 2: Distribution of **exploited** IoT devices deployed in consumer and CPS realms

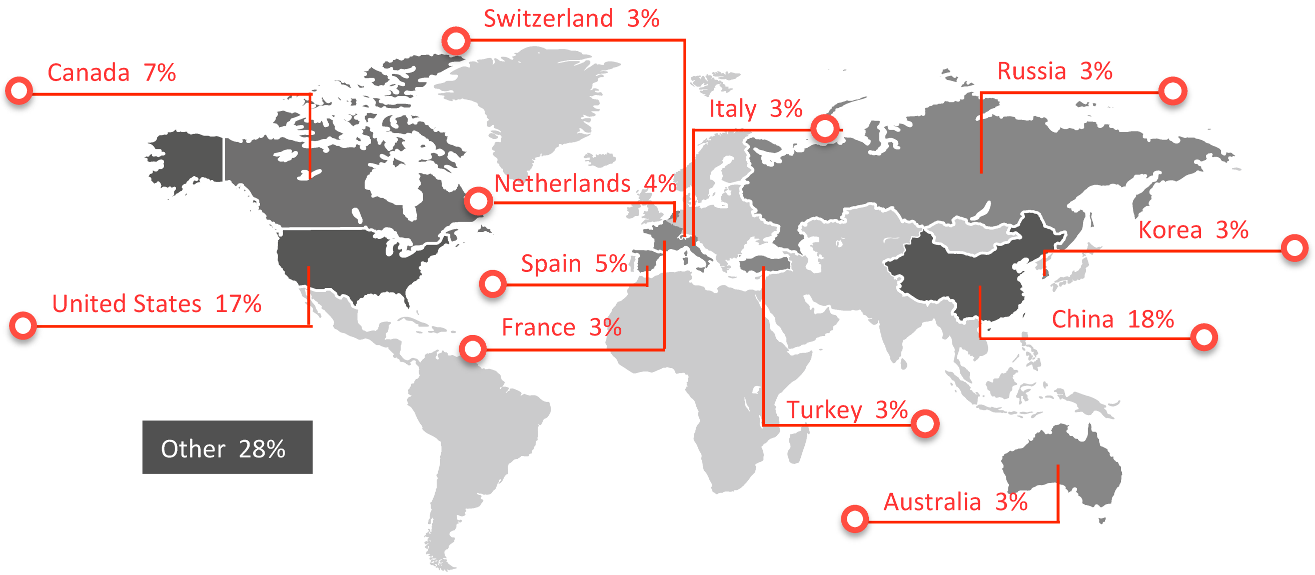


Fig. 3: Internet-scale distribution of **exploited** IoT devices deployed in CPS realms

of such exploitations within their corresponding categories. While the exploitation of IoT cameras is a reasonable outcome, DVRs, which have been exploited earlier this year by the Mirai malware, do not seem to be on top of the list of most exploited. In fact, IoT routers and printers appear to be more heavily compromised. This questions the fact if such devices will soon be leveraged as new bots within numerous botnets to launch similarly devastating attacks towards high priority Internet assets. More alarming, the results also demonstrate that IoT devices in manufacturing plants, building automation facilities and power utilities are the most exploited. This is indeed quite worrying, given that such vulnerabilities not only could lead to theft of highly sensitive and possibly classified intellectual property, but can also cause issues to the power infrastructure of nations and even endanger human life. By performing geo-location procedures using `maxmind`[2], we were able to attribute such IoT exploitations deployed in various CPS realms to their hosting environments (i.e., ISPs and countries). Please note that since we are exploiting probing intelligence as indicators of exploitation, the sources render real, non-spoofed IP addresses [28]. Figure 3 reveals that China, the United States, Canada and Spain host the top most IoT exploitations while Figure 4 shows the top 6 corresponding ISPs hosting these compromised IoT devices. To the best of our knowledge, the generated results herein render a first attempt ever to shed the light on Internet-scale IoT maliciousness. Indeed, empowered with such cyber threat intelligence, one can share such information with local IoT realms which are hosting these compromised IoT devices for prompt eradication, thus providing effective IoT mitigation.

[2] https://www.maxmind.com/en/geoip2-databases

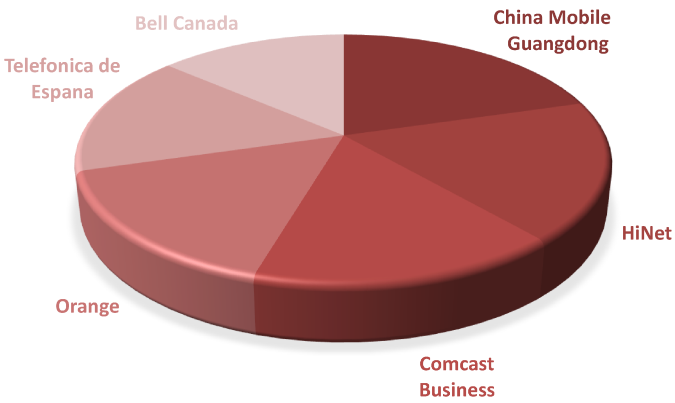


Fig. 4: Top 6 ISPs hosting **compromised** IoT devices in CPS realms

It is noteworthy to mention that the overall ratio between sampled (i.e., deployed) IoT devices and inferred exploited IoT devices in both, the consumer market and CPS realms, is computed to be around 2%. While this number seem to be small, one should note that IoT projections for 2020 is expected to reach 50 billion online IoT device, thus ominously causing IoT exploitations to develop into a momentous cyber security distress, to say the least.

V. Concluding Remarks

The Internet of Things (IoT) is an emerging paradigm of technical, social, and economic significance. Projections for the impact of IoT on the Internet and economy are impressive, with plethora of enterprises and analysts anticipating billions of connected IoT devices and a global economic impact of more than $11 trillion by 2025. While IoT deployments in the consumer sector have been receiving much hype, their corresponding implementations in CPS settings will undoubtedly provide massive benefits in terms of increased efficiency and cost reduction. Nevertheless, the initial priorities of IoT vendors have been focused on providing novel functionality, getting products to market, and making IoT devices more accessible and easier to use. Unfortunately, security concerns have not received as much attention. Motivated by the lack of IoT-relevant empirical data, this paper strived to present a first look on the scale and magnitude of IoT maliciousness. By fusing Internet-scale unsolicited darknet data with the results of active measurements, in addition to leveraging a probing inference technique, this paper shed the light on exploited world-wide IoT devices that have been deployed in consumer as well CPS environments. Some of the outcomes suggested the wide-spread compromise of IoT cameras and routers as well as the alarming exploitations of IoT devices in manufacturing and building automation facilities. This work indeed presents a solid foundation, in which future efforts, in this imperative IoT empirical security research area, are currently being planned and pursued. Foremost, a large-scale thorough characterization and analysis ought to be executed to precisely determine the scope of this issue. Further, we are currently investigating various IoT malware samples to perform correlations with data that is extracted from this work in an attempt to analyze the orchestration behavior of such malicious IoT devices. Additionally, we are presently developing a data sharing facility, where researchers can have access to IoT-relevant empirical data to strongly support IoT security research.

References

[1] Daniele Miorandi, Sabrina Sicari, Francesco De Pellegrini, and Imrich Chlamtac. Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, 10(7):1497–1516, 2012.

[2] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions. *Future Generation Computer Systems*, 29(7):1645–1660, 2013.

[3] Internet of things applications part 2: The mining industry.

[4] Smart cities - international case studies.

[5] Pavel Masek, Jan Masek, Petr Frantik, Radek Fujdiak, Aleksandr Ometov, Jiri Hosek, Sergey Andreev, Petr Mlynek, and Jiri Misurec. A harmonized perspective on transportation management in smart cities: the novel iot-driven environment for road traffic modeling. *Sensors*, 16(11):1872, 2016.

[6] Charith Perera, Arkady Zaslavsky, Peter Christen, and Dimitrios Georgakopoulos. Sensing as a service model for smart cities supported by internet of things. *Transactions on Emerging Telecommunications Technologies*, 25(1):81–93, 2014.

[7] Jun Zhou, Zhenfu Cao, Xiaolei Dong, and Athanasios V Vasilakos. Security and privacy for cloud-based iot: challenges. *IEEE Communications Magazine*, 55(1):26–33, 2017.

[8] Mark Stanislav and Tod Beardsley. Hacking iot: A case study on baby monitor exposures and vulnerabilities. *Rapid 7*, 2015.

[9] Eyal Ronen and Adi Shamir. Extended functionality attacks on iot devices: The case of smart lights. In *Security and Privacy (EuroS&P), 2016 IEEE European Symposium on*, pages 3–12. IEEE, 2016.

[10] Elisa Bertino and Nayeem Islam. Botnets and internet of things security. *Computer*, 50(2):76–79, 2017.

[11] Brandon Griggs. Connected tvs, fridge help launch global cyberattack. http://www.cnn.com/2014/01/17/tech/gaming-gadgets/attack-appliances-fridge/. 2015.

[12] Mirai: what you need to know about the botnet behind recent major ddos attacks.

[13] Hajime worm battles mirai for control of the internet of things.

[14] Tianlong Yu, Vyas Sekar, Srinivasan Seshan, Yuvraj Agarwal, and Chenren Xu. Handling a trillion (unfixable) flaws on a billion devices: Rethinking network security for the internet-of-things. In *Proceedings of the 14th*

ACM Workshop on Hot Topics in Networks, page 5. ACM, 2015.

[15] Blase Ur, Jaeyeon Jung, and Stuart Schechter. The current state of access control for smart devices in homes. In *Workshop on Home Usable Privacy and Security (HUPS)*. HUPS 2014, 2013.

[16] Grant Ho, Derek Leung, Pratyush Mishra, Ashkan Hosseini, Dawn Song, and David Wagner. Smart locks: Lessons for securing commodity internet of things devices. In *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*, pages 461–472. ACM, 2016.

[17] Vinay Sachidananda, Shachar Siboni, Asaf Shabtai, Jinghui Toh, Suhas Bhairav, and Yuval Elovici. Let the cat out of the bag: A holistic approach towards security analysis of the internet of things. In *Proceedings of the 3rd ACM International Workshop on IoT Privacy, Trust, and Security*, pages 3–10. ACM, 2017.

[18] Ang Cui and Salvatore J Stolfo. A quantitative analysis of the insecurity of embedded network devices: results of a wide-area scan. In *Proceedings of the 26th Annual Computer Security Applications Conference*, pages 97–106. ACM, 2010.

[19] Andrei Costin, Jonas Zaddach, Aurélien Francillon, Davide Balzarotti, and Sophia Antipolis. A large-scale analysis of the security of embedded firmwares. In *USENIX Security*, pages 95–110, 2014.

[20] Claude Fachkha, Elias Bou-Harb, Anastasis Keliris, Nasir Memon, and Mustaquel Ahamad. Internet-scale probing of cps: Inference, characterization and orchestration analysis. In *The Network and Distributed System Security Symposium (NDSS)*, To appear, 2017. http://public.eng.fau.edu/ebouharb/ndss-149.pdf.

[21] Roland Bodenheim, Jonathan Butts, Stephen Dunlap, and Barry Mullins. Evaluation of the ability of the shodan search engine to identify internet-facing industrial control devices. *International Journal of Critical Infrastructure Protection*, 7(2):114–123, 2014.

[22] Yin Minn Pa Pa, Shogo Suzuki, Katsunari Yoshioka, Tsutomu Matsumoto, Takahiro Kasama, and Christian Rossow. Iotpot: A novel honeypot for revealing current iot threats. *Journal of Information Processing*, 24(3):522–533, 2016.

[23] Juan Guarnizo, Amit Tambe, Suman Sankar Bunia, Martín Ochoa, Nils Tippenhauer, Asaf Shabtai, and Yuval Elovici. Siphon: Towards scalable high-interaction physical honeypots. *arXiv preprint arXiv:1701.02446*, 2017.

[24] Yair Meidan, Michael Bohadana, Asaf Shabtai, Juan David Guarnizo, Martın Ochoa, Nils Ole Tippenhauer, and Yuval Elovici. Profiliot: A machine learning approach for iot device identification based on network traffic analysis. 2017.

[25] Dina Hadžiosmanović, Robin Sommer, Emmanuele Zambon, and Pieter H Hartel. Through the eye of the plc: semantic security monitoring for industrial processes. In *Proceedings of the 30th Annual Computer Security Applications Conference*, pages 126–135. ACM, 2014.

[26] David Moore, Colleen Shannon, Geoffrey M Voelker, and Stefan Savage. *Network telescopes: Technical report*. Department of Computer Science and Engineering, University of California, San Diego, 2004.

[27] Claude Fachkha and Mourad Debbabi. Darknet as a source of cyber intelligence: Survey, taxonomy, and characterization. *IEEE Communications Surveys and Tutorials*, 18(2):1197–1227, 2016.

[28] Bou-Harb, Elias and Debbabi, Mourad and Assi, Chadi. Cyber scanning: a comprehensive survey. *IEEE Communications Surveys & Tutorials*, 16(3):1496–1519, 2014.

[29] Evan Cooke, Michael Bailey, Farnam Jahanian, and Richard Mortier. The dark oracle: Perspective-aware unused and unreachable address discovery. In *NSDI*, volume 6, pages 8–8, 2006.

[30] Lukas Krämer, Johannes Krupp, Daisuke Makita, Tomomi Nishizoe, Takashi Koide, Katsunari Yoshioka, and Christian Rossow. Amppot: Monitoring and defending against amplification ddos attacks. In *Research in Attacks, Intrusions, and Defenses*, pages 615–636. Springer, 2015.

[31] Luis Martin Garcia. Programming with libpcap-sniffing the network from our own application. *Hakin9-Computer Security Magazine*, pages 2–2008, 2008.

[32] J Matherly. Shodan search engine. [Online]: https://www.shodan.io.

[33] Jesse Kornblum. Identifying almost identical files using context triggered piecewise hashing. *Digital investigation*, 3:91–97, 2006.