A review of network vulnerabilities scanning tools: types, capabilities and functioning

Andrea Tundis
Technische Universität Darmstadt
Hochschulstrasse 10
Darmstadt, Germany 64289
tundis@tk.tu-darmstadt.de

Wojciech Mazurczyk
Warsaw University of Technology
Nowowiejska 15/19
Warsaw, Poland 00-665
w.mazurczyk@tele.pw.edu.pl

Max Mühlhäuser
Technische Universität Darmstadt
Hochschulstrasse 10
Darmstadt, Germany 64289
max@tk.tu-darmstadt.de

ABSTRACT

The rapid growth of the Internet in the last years has brought many advantages in the modern society in terms of communication and information sharing. Beside that, new and complex issues are emerging due to the network flexibility, openness and systems integration. The vulnerabilities of systems are the basis of these issues. Unfortunately, such vulnerabilities in the Internet can affect not only virtual environments in an isolated way but this can have serious repercussions in the real world. That is why, identifying new system vulnerability represents an important information for malicious parties. Currently, several tools (e.g. Shodan or Censys), which automatically scan the Internet, are available. They first scan the whole IPv4 public address range and ports in a distributed and random manner and then the obtained results are published on the publicly accessible websites. Such information can be later used for the benign or malicious purposes. In the latter case the main advantage for the potential attackers is that they gain reconnaissance data without even directly contacting the targeted device. Additionally, a large list of potential victims sharing the same vulnerability can be rapidly acquired. In this context, this paper aims at providing an overview of various publicly available network vulnerabilities scanning tools. In particular, first the main scanning tools are identified and classified. Then their main features are described and finally their advantages and disadvantages are highlighted.

CCS CONCEPTS

•Social and professional topics → Computer crime; •Security and privacy → *Social aspects of security and privacy;*

KEYWORDS

Scanning Tools, Vulnerability Scanners, Internet, Network Security, Cyber-security, Cyber-crime

ACM Reference format:
Andrea Tundis, Wojciech Mazurczyk, and Max Mühlhäuser. 2018. A review of network vulnerabilities scanning tools: types,
capabilities and functioning. In *Proceedings of International Conference on Availability, Reliability and Security, Hamburg, Germany, August 27–30, 2018 (ARES 2018),* 10 pages.
DOI: 10.1145/3230833.3233287

1 INTRODUCTION

The continuous evolution of the IT technology in the recent years makes new activities possible, ranging from the on-line selling, getting in touch easier with people all over the world, quick information sharing, etc. which contributes to modification of people's lifestyle. Moreover, thanks to the increasing interaction and integration between physical and virtual devices, the gaps between the real and the virtual worlds, where the Internet represents "the glue" among them, are decreasing. Even though this integration brings many advantages, new and complex challenges are emerging due to their vulnerabilities. For example, vulnerabilities in IoT devices and CPS systems not only have an impact bounded in the virtual world but they also can affect the real one with a high risk of compromising even human lives.

The availability of the data on the status of the network, in terms of accessibility, bugs and in general on existing vulnerabilities, represents a profitable source of information for malicious software developers and (cyber-)criminals. Some of them uses such information for entertainment, challenges or bets, whereas others such as criminal organizations or terrorists, exploit this knowledge for nefarious purposes motivated by economic, political or religious reasons such as, to track users or steal their identity sensible data, launch attacks to compromise banking systems, ex-filtrating confidential data, etc. Thus it is obvious that system vulnerabilities represent vital and precious information for malicious parties. The identification of new flaws and weaknesses in a system which is the first step in the attack chain can be a very complicated task to perform. It is mainly caused by a size of the network and typically consumes a lot of time and resources, especially if performed manually. Nowadays different tools are available to support the identification of systems' vulnerabilities by scanning the whole IPv4 address space [12], [28]. They were originally conceived for benevolent purposes, that is, to allow the identification of systems' flaws and weaknesses in order to let the system to be patched for example. However, as this information is publicly available and freely accessible, it is even more often gathered in advance by these criminals, who want to employ it for wicked purposes.

In this context, a review of such network vulnerabilities scanning tools and the capabilities they offer is necessary and in this paper we provide it by classifying such tools into two main groups (Fig.1) based on their operation mode (automatic/manual) and how the

ARES 2018, Hamburg, Germany
© 2018 ACM. 978-1-4503-6448-5/18/08…$15.00
DOI: 10.1145/3230833.3233287

results are generated and presented. The first group consists of *automatic scanning tools with publicly shared results*. The most notable examples of this group are Shodan, Censys, Thingful, PunkSPIDER, and Zoomeye. They that automatically scan the Internet, by providing a first interpretation of the findings and then publish their results publicly e.g. on the freely available website. The second group is composed of *personal interaction-based scanning tools* operated by the user and their results are returned directly and only to the user. The main examples of such solutions are: Nessus, skipfish, Acunetix, IVRE, Vulners, and Vega. These tools are also used to perform vulnerabilities identification and other scanning activities, however, they do not share their results publicly and always require user operator to generate results. For both groups of scanning tools a description, the general goals, as well as the context in which these scanning tools operate are provided by highlighting which kind of data or other information is exploited as well as by discussing pros and cons of each of them. Furthermore, a summary of minor scanning tools is provided.

The rest of the paper is structured as follows. Section II provides an overview regarding the role of scanning tools in the network security, whereas in Section III and IV an overview of the most popular scanning tools is presented. Section V summarizes further minor tools mainly open source. Whereas, a discussion among the main tools based on their distinguishing features is provided in Section VI by highlighting their pros and cons. Finally, the conclusions are drawn in Section VII.

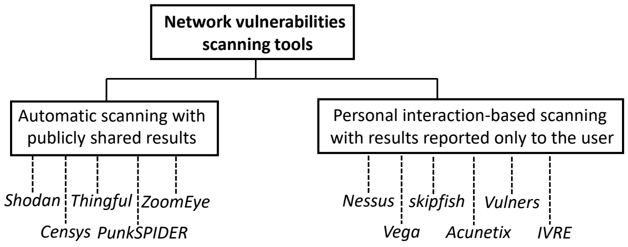

Network vulnerabilities scanning tools
Automatic scanning with publicly shared results
Personal interaction-based scanning with results reported only to the user
Shodan
Censys
Thingful
PunkSPIDER
ZoomEye
Nessus
Vega
skipfish
Acunetix
Vulners
IVRE

Figure 1: Classification of network vulnerabilities scanning tools

2 THE ROLE OF SCANNING TOOLS FOR THE NETWORK SECURITY AND RELATED WORKS

The Internet represents one of the most widespread technology in the modern society. It enables the interoperability among cyber-physical devices and services on the basis of which the online activities take place. In particular, the security, which is defined as a set of policies, rules, protocols, that are accepted by the online community, plays a fundamental role. Indeed, it aims to regulate the use of data and information as well as to ensure access to it in a controlled way.

In this context, the scanning tools analyze the security of the network, by trying to collect useful information as well as to identify potential network vulnerabilities. Indeed, scanning tools can be seen as search engines which seek for specific information. In [22], they are grouped into two main categories: *general purpose* and *subject-specific*. In particular, the first group exemplified by Google, Bing or Yahoo works based on queries that can have different types of input, for example, text, image or audio. These search engines perform their search by analyzing the content of the information source (content-type based research) and then by indexing them. The second group performs the scanning of the Internet for a specific defined subject areas, such as hosted services, SSL/TLS vulnerabilities or specific software or protocol vulnerabilities, such as XSS (Cross Site Scripting) or SQL injection. With respect to the *general purpose* search engines, the *subject specific* ones are mainly focused on the identification of particular vulnerabilities. For example, they take an URL as an input as well as specific attributes in order to discover if certain devices are connected to the Internet and whether they have any vulnerabilities.

Some research efforts devoted to the scanning tools are available in the literature. For example, [35] introduces the most popular open source scanning tool and then it focuses on hands-on ethical hacking and network defense. A vulnerability scanning hands-on labs based on OpenVAS is proposed, as a virtual environment for training people to choose the right scanning tool according to the users' needs. Whereas in [21], the performance evaluation is conducted by comparing Nmap and Nessus scanning tools on different operative systems. In [15], a case study on web application vulnerability scanning tools is described on the basis of the experience gathered by performing vulnerability analysis by using different types of policy configurations. In [25] instead, a vulnerability scanning tool is developed by extending an existing one called Nikto. In particular, it focuses on the detection on session management vulnerabilities, that are related to the session fixation, CSRF, and insufficient cookies attributes.

Considering above as well as the ad-hoc classification of the scanning tools introduced in section 1 (Fig. 1), in the following sections we describe the most notable examples of the automatic scanning tools for the vulnerability network attack analysis that publicly share their results (section 3) and personal interaction-based scanning tools with results presented only to the single user (section 4).

3 AUTOMATIC SCANNING TOOLS WITH PUBLICLY SHARED RESULTS

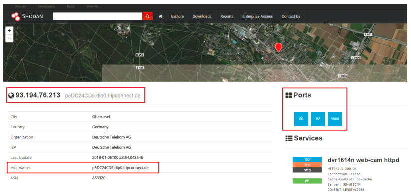
In this section, the description of the most popular automatic scanning tools (i.e. Shodan, Censys, Thingful, PunkSPIDER, and Zoomeye) which share their data publicly is provided, and then a summary regarding their pros, cons and current limitations are highlighted.

Shodan. It is one of the most popular scanning tools in the context of Internet of Things (IoT) regarding, for example, security in buildings, webcams, refrigerators, power plants, etc. It searches continuously for devices connected through the web, which are publicly accessible and which might be potentially attacked by malicious parties, with five main purposes [26]: *(i) IoT exploration*, to discover and track of the devices connected to the Internet and their users; *(ii) Network Security Monitoring*, with the aim of observing the computers available on the network and directly accessible through the Internet; *(iii) Global vision*, by analyzing the network not only considering the classic Internet but also other smart appliances such as smart TVs, refrigerators, power plants etc.; *(iv) Derive Market Advantages*, to monitor the products used by people in order to try to understand the types of products, their location as well as to derive good practices able to drive smartly the marketing (e.g. which countries have the largest number of wind farms, what companies are affected by Heartbleed, etc.); *(v) Cyber Risk Analysis*, to analyze the risk of the on-line connected devices.

Shodan is provided as an online service which, by scanning the IP addresses, tries to identify which service is under execution on which port. On the basis of this search it provides specific information about the services, status of the ports, headers, operative systems, etc. [24].

An example of text-based query is provided in Fig. 2.

```
Server: SQ-WEBCAM
linux upnp avtech
netcam
default password
dreambox country:ES
Server: SQ-WEBCAM country:"US"
admin+1234
category:malware
```



93.194.76.213 p5DC24CD5.dip0.t-ipconnect.de

City	Oberursel
Country	Germany
Organization	Deutsche Telekom AG
ISP	Deutsche Telekom AG
Last Update	2018-01-06T00:23:54.043546
Hostnames	p5DC24CD5.dip0.t-ipconnect.de
ASN	AS3320

Ports

80 82 5060

Services

80 tcp http

dvr1614n web-cam httpd

```
HTTP/1.1 200 OK
Connection: close
Cache-Control: no-cache
Server: SQ-WEBCAM
CONTENT-LENGTH:2936
```

Figure 2: Information Exploitation by Shodan

An important feature of Shodan is the location-based scanning, as shown in Figure 3, that allows to restrict/enlarge the search results on the desired geographical area[2].

Figure 3: Shodan Query Search Example with Shodan Maps

As it is shown in Figure 4 for the search query ”default password”, Shodan is able to generate a summary for statistics. In particular, it reports information about services, ports, headers, that can be employed to identify the owner a specific device, their location and additional data in order to discover potential vulnerabilities that can be used for malicious purposes [18].

The working process of Shodan is based on the combination of the SYN scanning and Banner grabbing that are both behavior-based techniques. In general, SYN scan is adopted to discover the target ports of the devices and then Banner grab is applied in order to get the protocol-specific data. In more detail, SYN scan, which is used in majority of network scanning tools e.g. in *Nmap*, works at TCP level and it is also known as half open scan. It sends a SYN probe segment and it receives a reply, in order to understand whether a port is open or not. A port is considered open when the ACK segment is received, while it is assumed closed when the RST flag is set within the received segment. Whereas Banner grab is an application layer scanning technique which is based on a TCP 3-way hand shaking. Specifically, it is based on the use of text and signatures of services that are displayed when a connection is established in several protocols such as FTP, SMTP, POP3 and telnet. It includes every behavior by gathering information from the open port (i.e. GET method in HTTP). The working process, depicted in Fig. 5 [34], ends by generating a list of the open ports on the devices as well as the services that are running on these ports [24].

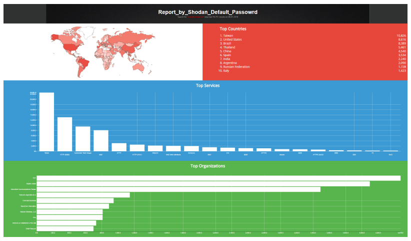


Figure 4: Report generated by Shodan

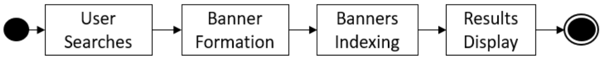


Figure 5: Shodan search functioning

Further features of Shodan include: *(i) Horizontal Scan*, that is focused on the specific categories of targets such as scanning a single port on multiple hosts; *(ii) Distributed Scan* that is devoted to the whole IPv4 address space; *(iii) Target Port* which is the related to a specific port. Moreover, Shodan uses basic filters such as Shodan Maps and Shodan Exploits to get information and it is possible to integrate Shodan with penetration testing tools such as Metasploit, Maltego and Nmap [26], [24]. The main pros and cons of Shodan are summarized in Table 1.

Censys. It is a schedule-based scanning tool, similar to Shodan, that enables users to retrieve information about devices and networks on the Internet. Censys keeps updated dataset with information on the devices exposed in the Internet as well as on the services that they provide. It is based on the SYN Scan and Banner grab techniques. It also utilizes well-known *ping* mechanism, which is exploited in the discovering step by scanning the whole IPv4 address range, is used to identify all occupied IP addresses; then details about applications running on the identified addresses and the potential vulnerabilities of the specific hosts are retrieved. Currently Censys is able to scan the following protocols: HTTP, HTTP Proxies, HTTPS, SMTP(S), IMAP(S), POP3(S), FTP, CWMP, SSSH, Modbus, StartTLS, Heartbleed and SSLv3 [16].

Table 1: Shodan [31], [30]

Pros	Cons
Create reports based on search query	It is not open source. The code and implementation details are unavailable.
Simple search query using words	Due to publicly availability of the data, this search engine can be misused by the attackers to identify and hack into the devices connected to the Internet
Support text-based search and Shodan map	Shodan does not process sensor outputs
Easy and free access to the website. Some of the data is shared for free or for a small cost	Catching everyday objects (capturing live data) on this website is still difficult
Data after search results can be downloaded in different format such as JSON, CSV or XML	-
Knowledge sharing such as a shared search queries the in community	-

Table 2: Censys

Pros	Cons
Easy to use search engine (similar to Google for Internet service querying)	Requires account creation after a certain amount of interaction with the search engine
Reports about security related protocols	Using the API is rate limited by token buckets
Provides an API to get the raw data and use the search engine from the third-party applications	No vulnerability analysis of targets by Censys itself
Allows dedicated querying for hosts, websites or certificates	Data is available, but needs to be interpreted by the user
Provides extensive filtering capabilities (by country, hosts, protocol, Alexa rank, tags, etc...)	Mostly relies on the older data (since the last Censys scan)
Provides detailed information about the search target	Does not support IPv6
Passive scan	The devices behind a NAT cannot be accessed

Through Censys it is possible to get the description of the devices that responded, as well as to visualize technical details related to their software, whether they use encryption, their configuration including also certificates [16]. All the gathered data is stored in a database which is publicly accessible by REST API, Google BigQuery and freely usable, for example, to identify the servers affected by the well-known vulnerabilities such as Heartbleed. From one side, Censys allows the attackers to easily find targets and their weaknesses, but from the other side it can be also used to analyze the Internet and to observe its evolution or discover new trends. Furthermore, the information available on Censys can be also used to prevent large scale attacks by informing the vulnerable devices about the potential dangers.

Censys relies on three main tools based on three main steps: *(i) Internet-Wide Scanning* in which ZMap [17] is used to perform scheduled internet-wide network surveys. With a 10GbE connection ZMap is able to scan the whole IPv4 address space in 5 minutes. It uses UDP and TCP SYN scan and it supports also ICMP, DNS queries, UPnP and BACNET for probing the network; *(ii) Application Scanning* in which ZGrab is used to perform a Banner grab of the services that are identified in the previous step; *(iii) Validation, Extraction, and Annotation* where ZTag is utilized to validate, extract and annotate the raw scanned data with additional meta-data and to transform them into records. [17]. The main pros and cons of Censys are summarized in Table 2.

ZoomEye. It is a cyber-space search engine for the IoT devices and vulnerabilities, alternative or even complementary to Shodan and Censys. It is based on Xmap and Wmap to perform its main activities related to the collection of available data from the public devices and web services and to conduct fingerprint analysis. It uses a keyword-based search criteria focused on the specific search parameters, such as: *(i)* application (e.g. application name and version number); *(ii)* location based on different parameters (e.g. country code, name of the city) through the exploitation of filters and the logical OR-operator; *(iii)* port number, operative system name and service name; *(iv)* hostname and IP address; *(v)* CIDR (Classless Interdomain Routing) segment and domain name as well as the header of the HTTP request; *(vi)* HTML specific tags (e.g. SEO keywords, description and title).

It is very helpful when a specific search string, related to the search topic of interest is known such as device="webcam" in order to retrieve all the webcam devices having an Internet connection; whereas a set of APIs for defining more complex queries are provided. ZoomEye main pros and cons are summarized in Table 3.

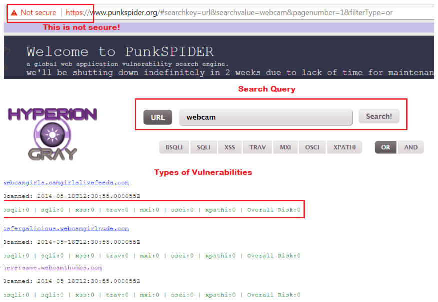
Table 3: ZoomEye

Pros	Cons
Different search criteria	One needs to know the specific word to search
APIs available	Manual search
Customizable search queries	-

PunkSPIDER. It is a vulnerability search engine for the web applications, that allows users to determine the number and the type of vulnerabilities on a website within the Internet quickly, easily, and in a more intuitive way [4], [1]. A list of different types of vulnerabilities that can be identified through PunkSPIDER are as follows [1]: BSQLI (Blind SQL Injection), SQLI (SQL Injection), XSS

(Cross Site Scripting), TRAV (Path Traversal), MXI (Mail Header Injection or Email Injection), OSCI (Operating System Command Injection) and XPATHI (XPath Injection).

Fig. 6 shows an excerpt of the results gathered from a scan search query "webcam", where in red some relevant information is highlighted i.e. insecure website and other results related to its potential vulnerabilities.

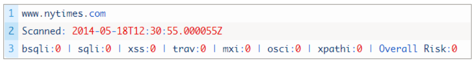


Welcome to PunkSPIDER
a global web application vulnerability search engine.
we'll be shutting down indefinitely in 2 weeks due to lack of time for maintenan

```
webcamgirls.camgirlslivefeeds.com
Scanned: 2014-05-18T12:30:55.000055Z
bsqli:0 | sqli:0 | xss:0 | trav:0 | mxi:0 | osci:0 | xpathi:0 | Overall Risk:0
sfergalicious.webcamgirlnude.com
Scanned: 2014-05-18T12:30:55.000055Z
bsqli:0 | sqli:0 | xss:0 | trav:0 | mxi:0 | osci:0 | xpathi:0 | Overall Risk:0
heversame.webcamthumbs.com
Scanned: 2014-05-18T12:30:55.000055Z
bsqli:0 | sqli:0 | xss:0 | trav:0 | mxi:0 | osci:0 | xpathi:0 | Overall Risk:0
```

Figure 6: PunkSPIDER exemplary results

In particular, after that the scanning is performed the list of the results, as depicted in Figure 7, is provided [14]. Specifically, the first line shows the domain of the result. The timestamp field on line 2 shows the time when the site has been added to the system, whereas the number of vulnerabilities found on the website are shown below. Further information, such as the overall risk, is provided through the risk field, that expresses the risk of visiting this particular website.



```
1 www.nytimes.com
2 Scanned: 2014-05-18T12:30:55.000055Z
3 bsqli:0 | sqli:0 | xss:0 | trav:0 | mxi:0 | osci:0 | xpathi:0 | Overall Risk:0
```

Figure 7: PunkSPIDER scan results

The main pros and cons of the PunkSPIDER are summarized in Table 4.

Table 4: PunkSPIDER

Pros	Cons
It is a global web application vulnerability search engine	Manually need to input the website name
It takes URL as a search query input	It does not really fit in IoT scanning tools category
Helps to determine the vulnerabilities in websites across the Internet	-

Thingful. It is a search engine conceived for the Internet of Things (IoT), which is basically centered on the geo-location. It collects the data coming from connected objects and sensors which generate real-time open data [8]. Through them, it is possible to provide geographical report regarding real-time data regarding energy, radiation, weather, from connected objects around the world such as air quality devices seismographs, ships, aircrafts and even animal trackers [8], [5]. In particular, thanks to sensor nodes, the data collected over the Internet are publicly published and through the portal (as shown in Figure 8a), it is possible to analyze them, for example in order to understand which area of the world is poorly connected.

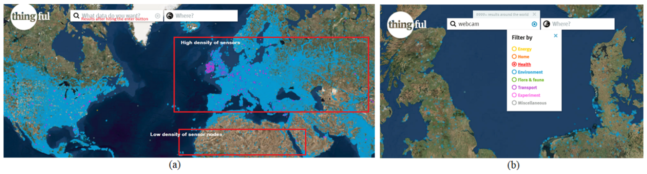


Figure 8: Thingful Sensor Nodes all over the world and its filtering feature

Thingful provides data regarding air quality monitoring, radiation measurement, flood monitoring and so on[8] and in this case, all citizen-centric "smart city" data can be analyzed, visualized, and used for event triggering by sending notifications. Indeed, a huge amount of data is produced and consumed from vehicles such as cars that are connected to the network and which communicate through local and remote services in real-time [33]. From such data useful information can be extracted in terms of drivers' preferences, habits as well as telematics-based insurance, remote diagnostics for maintenance, real-time feedback on driving behavior and vehicle life-cycle management. The urban IoT data can be accessed even through mobile apps using the Thingful APIs [6].

Thingful search is based on a proprietary algorithm for the identification of data (centered on ranking methodology based on patent-pending geo-spatial device data), which is provided on demand or on the basis of a schedule, that indexes the IoT data repositories across the world related to the environment, traffic, health and technology sensors [11], [7]. An example is shown through the Thingful GUI in Fig. 8b, where data about geo-location and time-series are provided, which can help to identify the event-related dynamics around the world. By using a zooming-in/out mechanism, it allows the users to restrict or enlarge the visualization to a specific region or area and access the related information. The main pros and cons of Thingful are summarized in Table 5.

4 PERSONAL INTERACTION-BASED SCANNING TOOLS WITH RESULTS REPORTED ONLY TO THE USER

In this section, the most popular personal interaction-based scanning tools which show their results only to a single user are presented. As previously, their pros and cons are then summarized.

IVRE. Instrument de Veille sur les Réseaux Extérieurs (IVRE) also called DRUNK (Dynamic Recon of UNKnown networks) is an open-source framework for network recognition [23]. Its main purpose is the reconnaissance of the network traffic, by gathering the information through an Internet-wide scanning, and then by investigating the collected information in order to explore potentially vulnerable resources and their activities [23], [20]. It is typically used in the field of digital forensics, information gathering, intrusion detection and identification of vulnerabilities in the network.

Table 5: Thingful [31] [30]

Pros	Cons
Gathers large-scale data and supports vast index of multi-domain data	It is not open source. The code and implementation details are unavailable.
Sensors are used to collect the data	Public limitation on the availability of the collected data
Provides a unique geographical index of real-time data from connected objects around the world	Provides access to its data only via a dedicated UI
Easy and free to access website	Fast expiration of the data due to the highly dynamic nature of the IoT connected devices
-	Unable to tap data where sensors are sparse
-	To collect the temperature data at a particular location on the planet, chances of finding the existing sensors depending on the given access, are limited

The users that are interested in this kind of tools are pentesters, security professionals, and system administrators [20].

IVRE is based on Nmap, Zmap, Masscan, Bro and p0f which allows *(i)* to retrieve data in an intelligent way from the network; *(ii)* to store the data in the MongoDB database, and *(iii)* to analyze the data through specific analysis tools [23]. Its work is based on the satellite imagery in the cyber-space by showing the world map of the Internet-exposed Modbus devices. The IVRE framework allows both active and passive data gathering by performing three main steps [23], [20]. The first is a *Scan and Sniff* step in which Nmap or Masscan are executed on a pre-defined target in terms of a specific network, range of the certain IP addresses or towards a whole country or on the full IPv4 connected address space. Zmap is exploited for a fast pre-scan, and to passively collect information from the network traffic by using Bro, Argus, Nfdump and p0f. In the next *Browse* step, the collected results are browsed using CLI tools, Python API or the available web interface, in order to find specific services or vulnerable versions, within a specific country or network, by using advanced filtering capabilities as shown in Fig. 9a. Finally, in the *Analyze* step the identification of similar hosts and particular cases is enabled by mostly looking for common ports, services or products, and generating a perspective about the space address with the heatmap. Fig. 9b illustrates the Nmap scan results with overview of the address space with a heatmap. Whereas, in Fig. 10 exemplary results coming from the scanning, by using the "heatmap" IP addresses to "zoom" in the address space, are presented.

Additionally, IVRE provides an advanced tool, called IVRE Flow tool to visualize and analyze the network flows among hosts. It is a recon tool for unknown network, that can be used as a cartography tool to retrieve information for a better understanding of a potentially known network and to exploit it as a monitoring tool to spot unwanted flows in a suspected network [23].

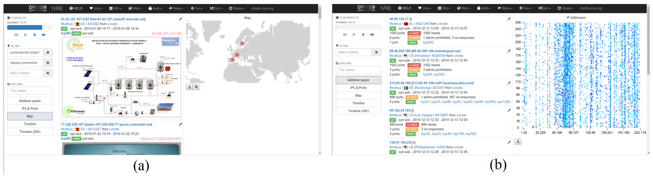
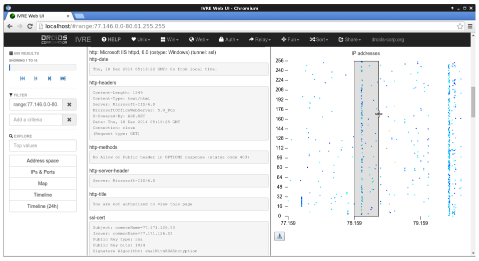


Figure 9: (a) Details of the IVRE scan results: using filters "solar" and map; (b) IVRE Nmap results: Home page with the "heatmap" IP addresses



```
http: Microsoft IIS httpd, 6.0 (ostype: Windows) (tunnel: ssl)
http-date
Thu, 18 Dec 2014 05:16:22 GMT; 0s from local time.
http-headers
Content-Length: 1549
Content-Type: text/html
Server: Microsoft-IIS/6.0
MicrosoftOfficeWebServer: 5.0_Pub
X-Powered-By: ASP.NET
Date: Thu, 18 Dec 2014 05:16:25 GMT
Connection: close
(Request type: GET)
http-methods
No Allow or Public header in OPTIONS response (status code 403)
http-server-header
Server: Microsoft-IIS/6.0
http-title
You are not authorized to view this page
ssl-cert
Subject: commonName=77.171.126.53
Issuer: commonName=77.171.126.53
Public Key type: rsa
Public Key bits: 1024
Signature Algorithm: sha1WithRSAEncryption
```

Figure 10: Details of the IVRE scan results: using the "heatmap" IP addresses to "zoom" in the address space

From a technical perspective, IVRE allows two kinds of working modes based on a Master-slave model: passive network recon and the active network recon which is Nmap-based. It is based on different other tools such as Argus, Bro, Masscan, Nmap, Zmap, and others in order to extract the data which is stored in MongoDB for further analysis. In addition, Neo4j database is also employed in the network flow analysis process. Table 6 summarizes the main pros and cons of IVRE.

Vulners. It is a database related to information security content, based on MongoDB and Elasticsearch databases, which provides a search engine to query for vulnerabilities, exploits, patches, and bug bounty programs. It is based on the aggregation of information coming from more than 70 different sources, by grouping them in six main categories [10]: *(i)* a list of vulnerabilities, a general description and links; *(ii)* various vendors' security bulletins (bug-reports), which are published by software vendors about vulnerabilities in their own products; *(iii)* exploits, which are displayed in a text editor; *(iv)* Nessus plugins for the vulnerabilities detection, which display whether a particular vulnerability can be detected by Nessus; *(v)* bug disclosures for bug bounty programs; *(vi)* publications from hacking resources.

Table 7 summarizes Vulners main characteristics.

Nessus. It is a proprietary vulnerability scanner designed to automate the testing and discovery of the known security problems by scanning ports and services running or listening on these ports. The hosts connected to that network can also be scanned [13]. It is available in three versions which include Nessus Home, Nessus Professional, and Nessus Manager/Nessus Cloud. The main security features of Nessus include: web-based interface, client-server

Table 6: IVRE [23]

Pros	Cons
Free software and open source framework	Requires manual setup. Installation and setup is time consuming
Works with Web UI and command line	Strongly depends on the use, environment and purpose of the implementation
It autoscans the network	IVRE is enumerated and not a simple list
IVRE Agent: meant to run in an environment which is not really controlled	Mostly has external dependencies. It relies on Python, Nmap, Bro, MongoDB, etc. Therefore, difficult to maintain the version and compatibility among the behavior of external programs or dependencies
IVRE Flow: for analyzing the network flows among the hosts	There are some inconsistencies among the IVRE Flow
IVRE web UI: central view is a graph representing the network	Need to improve the customization possibilities in the framework
Supports interaction with graphs	-
Provides data flow filters for analysis purposes	-

Table 7: Vulners

Pros	Cons
When updates arrive for a query, the user gets automatically notified	On-site scanner is not free
Extensive filtering	-
Very detailed description of vulnerabilities	-
Provides auditing to check for vulnerabilities of a certain OS family, version and packages	-
Provides an API for the third-party tools	-
Free to use search engine	-

architecture, remote and local security checks, built-in plugins. Different kinds of plugins and services/functionalities for supporting malware detection, web application scanning, and system configuration check, etc. are available. Whereas its advanced features can be categorized into automated scanning, multi-network scanning, and asset discovery. Nessus is able to perform several kinds of vulnerability scans, such as [32]: *(i) Asset discovery*: which is based on the scanning of a range of IP addresses by trying to understand which system or host is active, which ports are open and which services are listening on these ports and what OS are running on these systems; *(ii) Network Vulnerability*: this scanning mode aims at figuring out which hosts are vulnerable on which ports; *(iii) Patch Auditing*: it is based on the use of credentials and plugins to identify the missing vulnerability patches of an OS on every hosts, as well as missing patches on some third party software; *(iv) Configuration Auditing*: it is based on the exploitation of credentials and it analyzes the system configuration by comparing it to the settings included in an .audit file; *(v) Web Application Fuzz Testing*: it tries to find previously unknown web application vulnerabilities using fuzzy techniques. Table 8 summarizes Nessus main characteristics.

Table 8: Nessus

Pros	Cons
Can run on Windows and Linux (multi-platform)	The free version (Nessus Home) is constrained to scans of 16 IPs only
Uses a plugin based architecture to detect vulnerabilities (extensible)	The professional edition is very expensive
Provides proxy support with authentication	Real-time updates to the scan database require subscription
Targets can be queued and scanned automatically	Limited HTTP authentication support
Client/server architecture allows test automation	Active scanning
Simple graphical front-end	Cannot perform Internet-wide scans
Supports multiple IDS evasion techniques	-
High quality tests	-
Current data	-

Skipfish. It is an active web application security reconnaissance tool. It recursively crawls the targeted website and performs different security checks in the process by generating a report useful for web application security assessments.

It identifies a high number of vulnerabilities, by categorizing them as[27]: *(i) High risk flaws*, which can cause the system to be compromised, such are those related to: Server-side SQL, PHP injection, Explicit SQL-like syntax in GET or POST parameters, Server-side shell command injection, Server-side XML / XPath injection and blind vectors, format string and Integer overflow vulnerabilities, locations accepting HTTP PUT etc.; *(ii) Medium risk flaws*, which can compromise the data, such as those related to: XSS attack types, directory traversal / file inclusion and constrained vectors, MIME vulnerabilities etc.; *(iii) Low risk issues*, that usually have a limited or low impact on the system or on the data. These vulnerabilities are related to directory listing bypass vectors, SSL certificate problems, HTTP form submissions, etc. Further minor vulnerabilities are classified as internal warnings or even not classified at all. Table 9 summarizes Skipfish main characteristics.

Acunetix. It is a scanning tool for the automatic identification of vulnerabilities for websites and applications that are accessible through a web browser. The analysis is not based on the source

Table 9: Skipfish

Pros	Cons
Current data	Only active scanning is supported
Free to use, open source	Not easy to install, since it needs to be compiled from the source
Generate extensive report	Not very user-friendly, because it has no user-interface
-	Only command-line

code of the web application, but on the so called "Black and Gray Box" scanning method [19]. Acunetix is commonly used to perform penetration tests as well as to support the security assessment and vulnerability scanning. By using Acunetix a list of security issues along with specific details on the problem are provided together with possible advices and suggestion to deal with them.

The scanning approach is based on the three step process which ends with the generation of a final report [3]. The first is a *Crawling* step where a DeepScan crawler is employed to analyze the whole website on the basis of its links. Both static and dynamic pages using JavaScript as well as sitemap.xml files can be analyzed. Furthermore, if the AcuSensor is on, additional information can be also extracted from files such as web.config. Moreover, Acunetix can run a back-end crawl, in order to check files, which are not linked through the front-end applications but which may have been placed. The second is the *Vulnerability test* step where different vulnerability checks are performed by emulating a malicious behavior. General attacks, such as XSS, SQL injections are launched, moreover, the input fields are also tested by providing different input combinations to discover possible bugs. Additionally, if the AcuSensor is on, further checks, such as Blind SQL injection, directory traversal, for discovering other vulnerabilities are applied. Finally, in the *Report generation* step a summary is generated based on the information retrieved during the previous steps. Various types of reports such as executive summary and a report for the developers could be generated, which includes not only the identified vulnerabilities found but also recommendations to fix such problems. Two additional kinds of information can be also gathered [19]: *(i)* an audit of the web server, regarding the web applications in execution, by running port scans; *(ii)* a report about network vulnerabilities (by also exploiting OpenVas) that includes information about the operating system and the services.

Table 10 summarizes Acunetix main characteristics.

Vega. It is a free and extensible open source web security scanner and a testing platform that supports multiple scanning modes ranging from fully automated to manual one modality [9]. Its main goal is to identify potential vulnerabilities in a fully automated way or through a semi-automatic interaction with the website, like reflected cross-site scripting, stored cross-site scripting, blind SQL injection, remote file include, shell injection and TLS/SSL configurations. Specifically, for TLS/SSL security settings, it tries to identify how to improve a TLS/SSL security of the website.

Table 10: Acunetix [29]

Pros	Cons
Easy to use	Can only be used for web applications
In depth crawl and analysis	Generates a high number of garbage records in target backends
Generates extensive vulnerability report	Free version is only valid for one year
Current data	Only active scanning supported
-	Vulnerability analysis limited to the web applications

Three main modes of scanning are possible through Vega [9]: *(i) Automated Scanner*: a fully automated crawling is executed in this mode. By providing the user credentials, when necessary, the scanner is also able to automatically login into the system and to work on it; *(ii) Intercepting Proxy*: in this modality the user is able not only to observe the communication between clients and servers, but also to interact with them; *(iii) Proxy Scanner*: in this mode a user-driven semi-automated web security testing is performed, by scanning the target website.

Vega works on Linux, OSX and Windows platform since its detection modules are written in JavaScript and additional modules can be developed by using its APIs [9]. Table 11 summarizes Vega main characteristics.

Table 11: Vega

Pros	Cons
Cross-Platform	Only active scanning is supported
Easy to use graphical user-interface, based on the eclipse platform	No automatic target detection
Automated Scanner	Limited to web applications
Easy to extend by using the rich APIs offered	-
Detailed reports of detected vulnerabilities	-

5 OTHER MINOR SCANNING TOOLS

This section provides a short summary on further additional scanning tools by highlighting their main characteristics [28].

OpenVAS (Open Vulnerability Assessment System) is a scanning tool for network security purposes provided under GNU license. The most important components is the Security Scanner, available only for Linux-based platforms, that can be used to define vulnerability tests by using the Open Vulnerability Assessment Language

(OVAL). Its main options include intelligent custom scanning regarding *(i)* full network scanning; *(ii)* web server and web application scanning, and *(iii)* WordPress vulnerability scanning. Next, *Wireshark* is an open source scanning tool, which runs on Windows, Linux and OSX that supports the analysis of multi-platform network protocols. On the basis a client-server communication it examines the vulnerabilities of the fetched data by showing the flow construction of TCP sessions.

Another open source scanner for web server is *Nikto* and *Angry IP Scanner*. *Nikto* allows to quickly test and recognize possible suspicious behavior on the network, as well as the program which can use the traffic on the network. As main characteristics it provides: *(i)* HTTP proxy support; *(ii)* report generating in XML, HTML and CSV formats; *(iii)* search and check of HTTP servers, web server options and server configurations. *Angry IP Scanner* uses multi-threading mechanisms to make the scanning of both IP addresses and ports more efficient. It collects information by generating a report in CSV, text and/or XML formats regarding hostname, NetBIOS, MAC address, computer name, workgroup information, etc.

Advanced IP Scanner and *Qualys Freescan* are other free and open source network scanning tools. Advanced IP Scanner works in Windows and it aims to identify devices on the Internet by allowing remote access, remote wake-on-LAN and quick shutdown as well as services such as HTTPS, RDP and FTP services on the remote machine. *Qualys Freescan* scans for URLs, IP addresses and local servers to detect security loopholes. Specifically, three kinds of scans are provided: *(i)* vulnerability checks, related to maleware and SSL related issues; *(ii)* OWASP-related web applications security checks; *(iii)* SCAP checks, related to the computer network configuration against security contents.

Next, *Retina Network Security Scanner* is a free vulnerability scanning tool centered on Zenmap with advanced GUI, Ndiff for computer scan results and NPing for Response Analysis. It is based on the credentials provided by the user which allows a user to choose the type of report to be generated. It also supports the risk assessment based on optimal network performance, operating system, and applications, as well as the identification of security issues in the network and mis-configurations. Whereas *Metasploit Framework* is a tool, for performing penetration tests, that can be employed as a scanning tool for the detection of bugs in the network. It was initially open-source, but now it is provided as a commercial tool, even if a limited open-source and free version is still available.

Snort, instead, is an intrusion detection and prevention system, free and open-source which is used to analyze the traffic on the network through IP addresses by performing malware detection, port scan and other network exploits through the analysis of protocols. Whereas *Nexpose* is a commercial scanning tool that is freely available as a community edition which provides scanning capabilities of the network, operating systems, application database, etc. through a web-based GUI for Windows and Linux operating systems. Finally, *Fiddler* is a debugging tool for the web that automatically inspects HTTP traffic. It scans the traffic between specific computers over a network and analyzes the data packets in order to observe the requests and responses between the hosts. It is able to decrypt HTTP traffic and it is also typically utilized to assess system performance and security testing of web applications.

6 OPEN DISCUSSION

As the Internet is becoming even more popular and interactive through the connection of new and more powerful devices, new opportunities for business, social life but also for attackers and criminals, are emerging. One of the major factors is related to the introduction of more complex and hybrid devices such as IoT, CPSs. Indeed, their integration in the Internet by neglecting security and privacy aspects results in an increased growth of systems and devices vulnerabilities. Scanning the Internet in order to identify these devices is of great interest not only for researchers to analyze users' behavior but also for the potential attackers and cyber-criminals to create hardships.

On the basis of the review in the previous sections two main tools emerged for similarities in terms of filtering mechanisms, dealing with a huge set of data, along with the third-party supporting tools based on APIs, that is, Censys and Shodan. However, Shodan provides a more user-friendly front-end as well as it focuses more on presenting the results according to the location as reference parameter, whereas Censys provides more detailed information. Furthermore, Shodan provides an enterprise edition of its platform with enterprise analytical functionalities that Censys does not have [24]. Whereas the main difference between Shodan/Censys and IoT search engines such as Thingful, is that Shodan/Censys are basically designed as search engines for potential attackers [30] as the data is publicly available and can be misused by the intruders to gain unauthorized access to the machines. Moreover, they can identify and thus help to hack into password protected devices which are connected to the Internet. Even, servers, routers as well as other Internet-connected devices have been archived with their IP addresses in its database. Unfortunately, the website itself does not process the sensor outputs so the access to live and updated data is troublesome. Due to its large and broad scope, catching everyday objects on this website is still difficult while servers and network devices constitute the majority of the things in its database.

According to [30], the only working example of the IoT search engine is Thingful and none of the IoT search engines in the literature have been deployed for the real-world or large-scale data. Furthermore, the Thingful initiation itself is still limited, and a significant progress is needed to expand this area. One reason for such limitations is the public availability of the collected data. For example, Thingful provides access to its data only via a specific User Interface. Another example of the limitations is the fast expiration of data due to the highly dynamic nature of IoT devices. Table 12 provides a brief summary of the main reviewed scanning tools.

7 CONCLUSION

The paper presented a review of the most popular on-line scanning tools which are able to extract data, coming from both virtual services and physical/IoT devices. Some of these scanning tool are general purpose and have limited analytical capabilities, whereas other tools are conceived for achieving specific purposes such as the discovering of vulnerabilities and provide more advanced support for analytics.

Two main groups of such scanning tools have been identified: *automatic scanning tools with publicly shared results* which automatically scan the Internet, by providing a first interpretation of

Table 12: A summary of the scanning tools above discussed - Legend: Y (Yes), N (No), P (Partial)

Tools	Free Accessible Data	Manual Installation	Open Source	Proprietary	Free Use	Add-on	Search-query
Shodan	Y	N	N	Y	Y	Y	query syntax, words
Censys	Y	N	P	Y	Y	Y	phrase, words
Thingful	Y	N	N	Y	Y	N	phrase, words
PunkSpider	Y	N	N	Y	Y	N	URL
IVRE	Y	Y	Y	Y	Y	N	phrase, words
Vulners	Y	N	N	Y	Y	Y	phrase, words
Nessus	N	Y	N	Y	P	P	-
Skipfish	N	Y	Y	Y	Y	N	-
Acunetix	N	Y	N	Y	P	P	-
Vega	N	Y	Y	Y	Y	N	-

the findings and then publish their results publicly and *personal interaction-based scanning tools* operated by the user and which results returned directly and only to the user. From this perspective the most notable tools which perform Internet-wide scans and publish these findings publicly have been investigated. For each of them the main objectives, the application context and some distinguishing features have been identified. Furthermore, a short summary of their main advantages and drawbacks has been provided.

We conclude that the existence of these tools has two sides *(i)* the positive aspect is that vulnerable devices could be warned about dangers preemptively, whereas *(ii)* the negative side is that the malicious users can utilize these tools to retrieve information about bugs and vulnerabilities in order to attack and damage systems and/or devices. Future works will be devoted to study and analyze the behavior of users, and specifically of the attackers and cyber-criminals, based on the information provided by these public scanning tools with a particular focus on system vulnerabilities and bugs.

ACKNOWLEDGMENTS

The work was performed in the context of the TAKEDOWN research project, which receives funding from the European Union's Horizon 2020 Research and Innovation Programme award number 700688. The authors would like to thank Mr. Alexander Brakowski and Ms. Manjiri Birajdar for their support in such task.

REFERENCES

[1] *Darknet Blog.* https://www.darknet.org.uk/2016/09/punkspider-web-vulnerability-search-engine/.

[2] *Introducing Shodan Maps.* https://shodanio.wordpress.com/2014/02/18/introducing-shodan-maps/

[3] *Introduction to Acunetix - Why You Need To Secure Your Web Applications.* https://www.acunetix.com/support/docs/introduction/

[4] *PunkSPIDER Official Website.* https://www.punkspider.org/

[5] *Thingful Blog.* http://blog.thingful.net/post/85807476836/welcome-to-thingful

[6] *Thingful Blog london-cambridge cycling finding.* http://blog.thingful.net/post/149362243876/showcase-londoncambridge-cycling-finding

[7] *Thingful Blog Virtual Sensor Project.* http://blog.thingful.net/post/149696551076/virtual-sensors-using-thingful-and-data-science

[8] *Thingful Official Website.* http://umbrellium.co.uk/initiatives/thingful/,https://www.thingful.net/

[9] *Vega Vulnerability Scanner and Web Security Testing Platform.* https://subgraph.com/vega/index.en.html

[10] *Vulnerability Data Base - 'Google' for Hackers.* http://vulners.com/landing

[11] E. Aceves and V. M. Larios. 2015. White paper: Data Visualization for Georeferenced IoT Open Data Flows for a GDL Smart City Pilot. https://smartcities.ieee.org/images/files/pdf/davgdl_iotvisualinterface.pdf

[12] H. Al-Alami, A. Hadi, and H. Al-Bahadili. 2017. Vulnerability scanning of IoT devices in Jordan using Shodan. In *2017 2nd International Conference on the Applications of Information Technology in Developing Renewable Energy Processes Systems (IT-DREPS).* 1–6. DOI:http://dx.doi.org/10.1109/IT-DREPS.2017.8277814

[13] Harry Anderson. *SecurityFocus Printable INFOCUS 1741 - Introduction to Nessus.* http://cryptomex.org/SlidesSeguRedes/TutNessus.pdf

[14] Alejandro Caceres. *PunkSPIDER Search.* https://hyperiongray.atlassian.net/wiki/spaces/PUB/pages/10190871/

[15] N. I. Daud, K. A. A. Bakar, and M. S. M. Hasan. 2014. A case study on web application vulnerability scanning tools. In *2014 Science and Information Conference.* 595–600. DOI:http://dx.doi.org/10.1109/SAI.2014.6918247

[16] Zakir Durumeric, David Adrian, Ariana Mirian, Michael Bailey, and J. Alex Halderman. 2015. A Search Engine Backed by Internet-Wide Scanning. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security.* ACM, New York, NY, USA, 542–553.

[17] Zakir Durumeric, Eric Wustrow, and J Alex Halderman. 2013. ZMap: Fast Internet-wide Scanning and Its Security Applications.. In *Proceedings of the 22nd USENIX Security Symposium.* Washignton, USA.

[18] V. J. Ercolani, M. W. Patton, and H. Chen. 2016. Shodan visualized. In *IEEE Conference on Intelligence and Security Informatics (ISI).* 193–195. DOI:http://dx.doi.org/10.1109/ISI.2016.7745467

[19] Emre Erturk and Angel Rajan. 2017. Web Vulnerability Scanners: A Case Study. *CoRR* (2017). arXiv:1706.08017 http://arxiv.org/abs/1706.08017

[20] Linux Security Expert. *IVRE tool review.* https://linuxsecurity.expert/tools/ivre/

[21] Sun-Young Im, S. H. Shin, Ki Yeol Ryu, and Byeong hee Roh. 2016. Performance evaluation of network scanning tools with operation of firewall. In *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN).* 876–881. DOI:http://dx.doi.org/10.1109/ICUFN.2016.7537162

[22] Simon K., Moucha C., and Keller J. 2017. Contactless Vulnerability Analysis using Google and Shodan. *Journal of Universal Computer Science* 23 (4) (2017).

[23] Pierre Lalet, Florent Monjalet, and Camille Mougey. *IVRE, a network recon framework.* https://ivre.rocks/

[24] S. Lee, S. H. Shin, and B. h. Roh. 2017. Abnormal Behavior-Based Detection of Shodan and Censys-Like Scanning. In *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN).* 1048–1052. DOI:http://dx.doi.org/10.1109/ICUFN.2017.7993960

[25] R. Lukanta, Y. Asnar, and A. I. Kistijantoro. 2014. A vulnerability scanning tool for session management vulnerabilities. In *2014 International Conference on Data and Software Engineering (ICODSE).* 1–6. DOI:http://dx.doi.org/10.1109/ICODSE.2014.7062682

[26] John Matherly. *Shodan Official Website.* https://www.shodan.io/

[27] Zalewski Michal, Heinen Niels, and Roschke Sebastian. *skipfish - web application security scanner.* https://code.google.com/archive/p/skipfish/wikis/SkipfishDoc.wiki

[28] W. Qianqian and L. Xiangjun. 2014. Research and design on Web application vulnerability scanning service. In *2014 IEEE 5th International Conference on Software Engineering and Service Science.* 671–674.

[29] Natasa S., Dragan A., and Aleksandra M. One Unwanted Feature of Many Web Vulnerability Scanners. In *Proceedings of 11th International Conference for Informatics and Information Technology (CIIT 2014).* Bitola, Macedonia.

[30] Ali Shemshadi, Quan Z. Sheng, Wei Emma Zhang, Aixin Sun, Yongrui Qin, and Lina Yao. 2016. Searching for the Internet of Things on the Web: Where It Is and What It Looks Like. *CoRR* abs/1607.06884 (2016). arXiv:1607.06884

[31] M. Sheng, Y. Qin, L. Yao, and B. Benatallah. 2017. *Managing the Web of Things: Linking the Real World to the Web.* Elsevier Science.

[32] tenable. *Vulnerability Reports generated by Nessus.* https://www.tenable.com/products/nessus/sample-reports

[33] Thingful-uh and Khan Usamah. *Thingful Blog connected vehicles leveraging iot data.* http://blog.thingful.net/post/149455464806/showcase-connected-vehicles-leveraging-iot-data

[34] Saja Verma. *Searching Shodan For Fun And Profit.* https://www.exploit-db.com/

[35] Y. Wang and J. Yang. 2017. Ethical Hacking and Network Defense: Choose Your Best Network Vulnerability Scanning Tool. In *2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA).* 110–113. DOI:http://dx.doi.org/10.1109/WAINA.2017.39