Bulletin of Electrical Engineering and Informatics
Vol. 10, No. 5, October 2021, pp. 2804~2810
ISSN: 2302-9285, DOI: 10.11591/eei.v10i5.2901

RT-RCT: an online tool for real-time retrieval of connected things

Fatima Zahra Fagroud[1], El Habib Ben Lahmar[2], Hicham Toumi[3], Youssef Baddi[4], Sanaa El Filali[5]
[1,2,5] Laboratory of Information Technology and Modeling Faculty of Sciences Ben M'sik, Hassan II University of Casablanca, BP 7955 Sidi Othman Casablanca, Morocco
[3,4]Higher School of Technology-Sidi Bennour Chouaïb Doukkali University El Jadida, Morocco

Article Info

Article history:

Received Feb 22, 2021
Revised Jun 14, 2021
Accepted Aug 3, 2021

Keywords:

Connected things
IoT
Real-time
Retrieval

ABSTRACT

In recent years, internet of things (IoT) represents a giant and a promoter area in innovation and engineering fields. IoT devices are spread in various fields and offer advanced services which assist their users to monitor and control objects remotely. IoT has a set of special characteristics such as dynamic, variety of data and huge scale which introduces a great challenge in the field of retrieval technologies, more precisely real-time retrieval. This paper addresses the issue of real-time retrieval of connected things and tries to propose an innovative solution which allows the retrieval of these things and their descriptive data. The paper proposes an on-line tool for real-time retrieval of connected things and their descriptive data based on network port scanning technique. The performance of this tool proves to be powerful under normal conditions, however more tests must be implemented in the aim to improve the proposed solution. The tool resulted from this work appears to be promising and can be used as a reference by network administrators and IT security managers for the development of new security mechanisms and security reinforcement.

Corresponding Author:

Fatima Zahra Fagroud
Laboratory of Information Technology and Modeling
Faculty of Sciences Ben M'sik
Hassan II University of Casablanca, BP 7955 Sidi Othman Casablanca, Morocco
Email: fagroudfatimazahra0512@gmail.com

1. INTRODUCTION

Information and communication technologies (ICT) know a significant development especially in terms of hardware miniaturization, cost reduction and energy consumption optimization. This advancement enables the interconnection of a large number of physical objects namely using the Internet, forming what is called the internet of things (IoT). The IoT provides the opportunity to interact with these objects through sensors, actuators and smart applications which may help users in several areas such as transport, logistics, health care, agriculture, etc [1].

IoT represent a static objects that will be intelligent and able to share information and communicate with other devices in an autonomous way [2]. There are many elements used to run the IoT technology which include hardware and software such as sensors, GPS, cameras, applications, and so forth [3]. IoT devices are spread in different areas such as e-tracking, e-commerce, e-home, and e-health, etc. Thus, during the last ten years, the IoT technology has been a research focus [3]. These devices produce a big quantity of information, heterogeneous data, and their state changes very quickly (in a short period of time).

Internet is a popular global information system where users can search relevant information using search engines (SE). SE is a type of software that organizes various content collected from all resource

available in the internet. With SE, users who are wishing to find information only need to enter a keyword about what they had like to see, and the search engine presents the links to the content that resembles what they need [4]. Searching in IoT networks has a different goal than the ones that typical search engines adapt where the users would operate the objects locally or remotely. As a result, this distinguishes between both sides and requires a new design concept for an IoT search engine. This is not simple due to the need to design new techniques of crawling, indexing, storing and querying [5].

Many IoT search engines [6]-[9] are designed to allow the search/retrieve and identification of connected things. Shodan.io is a search engine designed by programmer John Matherly in 2009. It interrogates devices ports and grabs the resulting banners, then indexes the corresponding public IP address and search into an intern databases for futures lookup [10]. Another popular IoT search engine is Censys, which collects all data it can about the connected devices in IPv4 on the net. it use the open source port scanner ZMap specially + ZGrab and stores everything it retrieves in a database, which is then accessible via the web interface, an API or plain text listings to download [11]. Thingful can be described as a "discoverable search engine" which allows its users to have a geographic index of connected objects around the world (https://thingful.net/). As such, Thingful boasts that it can index across multiple IoT networks and infrastructures, because this search engine can locate the geographical position of objects and devices [12]. But these search engines do not perfectly meet the need due to the quick changes of devices state and the complexity of their results, which require the development of a new mechanism for IoT devies retrieval which can respond to the different issues like real-time retrieval, fast response and accurate results.

This work aims to propose an on-line tool for real-time retrieval of connected things in worldwide and descriptive informations related to these devices based on network port scanning technique. The paper starts by introducing the basic concepts related to the development of the proposed tool. In Section 3, the specifaction requirements and the proposed tool are presented. Then, in Section 4 we will present and discuss the results. Lastly, the conclusion and future improvements.

2. BACKGROUND

2.1. Internet of things and connected things

IoT represent a giant infrastructure that enable machine-to-machine communication, remote monitoring and control of objects/devices in many fields and applications such as industry, agriculture, healthcare and education. It represents a network of connected things which are connected to IoT, and able to gather and share information related to the way they are employed and almost the environment around them. IoT represent the main focus of many research works [13]-[18] in latest years.

Connected things refer to smart devices, autonomous electronic devices that may be connected with each others in a network, mobile devices, computing devices which are typically small enough to be handheld. These things are connected by using various wired and wireless networks and protocols (Wi-Fi, 3G and 4G networks... etc.), and are usually monitored and controlled remotely. They are commonly embedded with a set of technologies such as processing chips, software, and sensors.

Things, in the IoT sense, can refer to a wide variety of devices such as heart monitoring implants, biochip transponders on farm animals, cameras streaming live feeds of wild animals in coastal waters, automobiles with built-in sensors, DNA analysis devices for environmental, food, pathogen monitoring, or field operation devices that assist fire fighters in search and rescue operations. Legal scholars suggest regarding "things" as an "inextricable mixture of hardware, software, data and service [19].

2.2. Searching

Web sites, which index and class other web sites according to their keywords, explanations and contents and make it easier and faster to reach obtained site-search results, are called as search engines [20]. Since their appearance in the 90s, they recognize a great success and presents a change in the way of information retrieval. It is a tool based on a set of algorithms which allows its users to search and access to a huge amount of web information in an easy way and also to have well-organized results. These engines become smart due to the integration of new methods like machine learning for results classification task and interpretation of requests.

IoT has a set of special features which present a great challenges for traditional search engines, in order to respond to these issues and continue the success of search engines with the large number of IoT devices joining the Web every day a new evolution of these tools appeared entitled IoT search engines [21]-[23]. It's a solution that allows us to obtain a new search tool able to find connected devices and information about them, and also solve a set of internet of things issues.

With the emergence of the internet of things, challenges relating to network security, devices management, devices status, access control and anomaly detection bring managers and administrators of the IoT infrastructure to think to the design and develop a new support and mechanism. The use of IoT search

engines can alleviate the IoT challenges mentioned [24]-[26] because they have the ability to identify devices and services connected to the Internet as well as vulnerable devices, also they allow learning and search information about IoT.

2.3. Network port scanning

Network analysis represent a technique which scan network ports as a vulnerability analysis, and usually used for security assessment and system maintenance. In addition, it's among necessary ways employed by attackers to assemble their data.

Network scanning consists of network port scanning as well as vulnerability scanning [27]. Network port scanning refers to the way of dispatch information packets via the network to a computing system's given service port numbers (for example, port twenty-three for Telnet, port eighty for protocol so on). This is often to spot the on the world network services on it explicit system. Network port scanning moreover as vulnerability scanning is associate degree information-gathering technique, however once applied by anonymous people, these are viewed as a prelude to associate degree attack. Network scanning processes, like port scans and ping sweeps, come details regarding that information science addresses map to active live hosts and therefore the kind of services they supply [28]. It can be done in an easy way by using the available scanning tools like nmap [29], Angry Ip Scanner [30], Advanced Port Scanner [31].

2.4. Multiprocessing

Multiprocessing or parallel processing is a type of processing which serve to run a set of tasks simultaneously on multiple processors in Figure 1. It aims to get more work done in a shorter period of time and reduce overall processing time than the serial processing. This type is typically used when very high speed is required to process a large volume of data. Multiprocessing serve to distribute a complex and larger tasks into multiple and smaller calculations, when each sub-process will have a dedicated CPU and memory slot. It refers to the ability of a system to support more than one processor at the same time and independently.

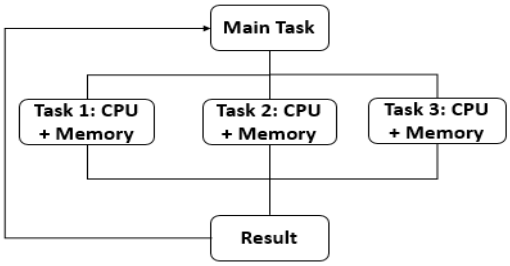

Figure 1. Multiprocessing

Multiprocessing can be used to improve existing version of different proposed solutions by speeding the processing time, like the work presented by Li *et al.* [32]. They develop an efficient guide RNA library designing tool entitled MultiGuideScan. It represents a multi-processing version of GuideScan software (developed to design CRISPR guide RNA libraries, which can be used for genome editing of coding and noncoding genomic regions effectively [32]). Experiments prove that the proposed solution speeds up the design of RNA guide library about 9-12 times by using 32 process than the original GuideScan.

3. RESEARCH METHOD

The main idea of this work is to propose a retrieval tool that provide to users all current available informations of each thing in request with minimum delay possible. The informations about devices in request are collected in real-time by using network port scanning technique especially we used python-nmap library. This data are collected from a set of scans, where each scan is responsible for retrieving a set of specific information which can take a significant time. In the aim to reduce data collection time we elaborate a parallel scans which serve to perform all scan in the same time and as fast as possible.

3.1. Software requirement specification

Requirement specification is the first step to define when developing a tool or application. For that we present in this sub-section the essentials requirements for the development of our solution; a) provide

simple GUI and easy to use, b) provide accurate and understandable results, c) provide maximum of available information related to connected things, d) allow its users to perform real-time retrieval, e) users do not need any technical knowledge, f) free (no registration required), g) unlimited number of searches, h) full access to results.

3.2. Proposed algorithm

The flowchart of the proposed algorithm is shown in Figure 2 which include the following steps:

a. The first step aim to send user query to the server:
   Query aim to specifiy target host Which can be represented by ip adress or hostname
b. Launch a set of scans in parallel in order to find information relating to the device in request:
   Each scan is responsible for collecting specific data
   For scans which take up more time we divide them on sub scans as long as it is possible if not we launch similar scans in parallel
c. Collect the results generated by the performed scans
   Case 1: collect the results from all scans and cobine them then move on to the next step
   Case 2: collect the results of each scan separately then move on to the next step
d. Extract relevant informations to shown from collected data and send them to users
   Case 1: extract information from alll scan combined results
   Case 2: extract information of each scan results received separately
e. Displays search results in an ergonomic and understanvble way on the system interface
f. Due to the dynamic change of information, all scans are relaunshed within a specified period of time and the content of the page is automatically refreshed as long as the user accesses to results interface.

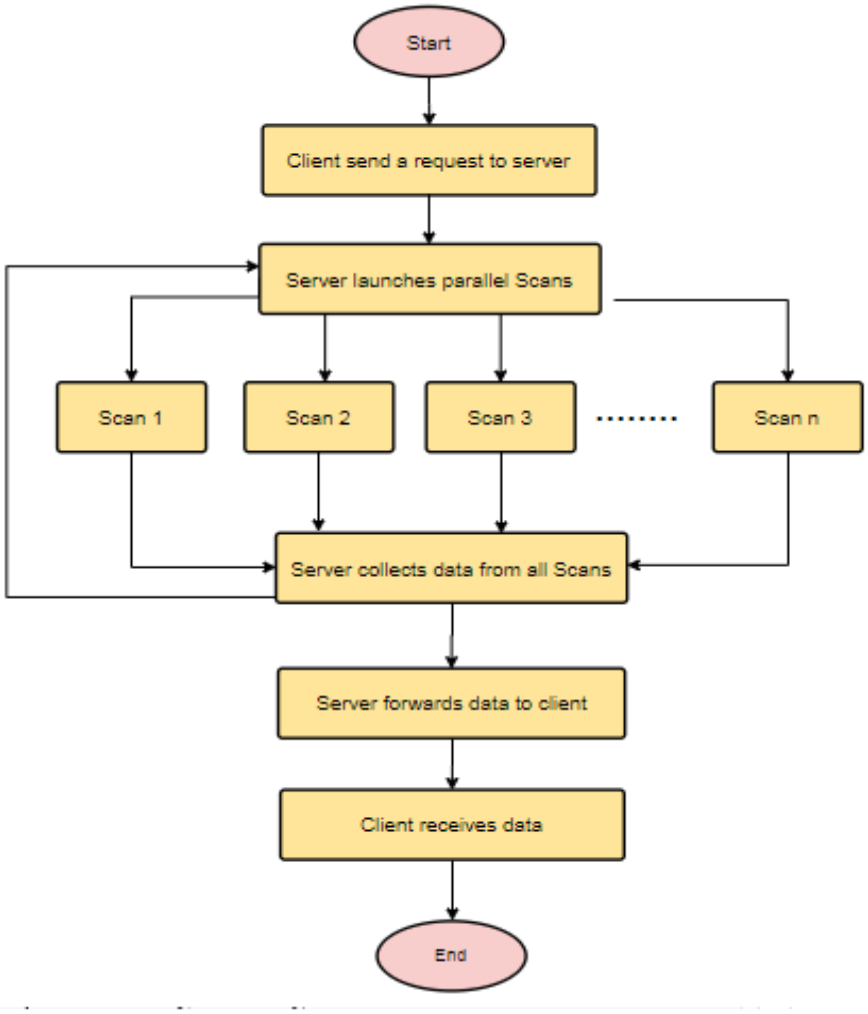

Figure 2. Proposed algorithm for real-time retrieval

4. RESULTS AND DISCUSSION

The proposed retrieval tool is developed and desgnied as a web application to allow an easy way to use this solution and does not require any prior installation, as well as guaranteeing the use of the latest version of this tool. The proposed web application was developed by using open-source micro framework for web development in Python (flask-python [33]) and other technologies and python librairies like python-nmap [34]. This application is based on two main interfaces:

- The first user interface in Figure 3 of this system aim to offer a simple and ergonomic interface that allows users to retrieve current data related to connected things and in an easy way.
- The results interface is displayed in a short time after the user's request. This interface shows the current informations related to state, ports, protocols, os, device type, hostnames and addresses in Figure 4 and Figure 5. It offers a useful and clear visualization of all available data collected in real-time.

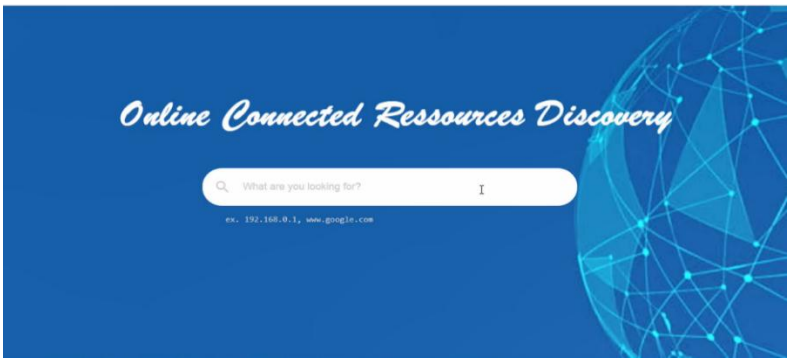

Figure 3. Retrieval GUI

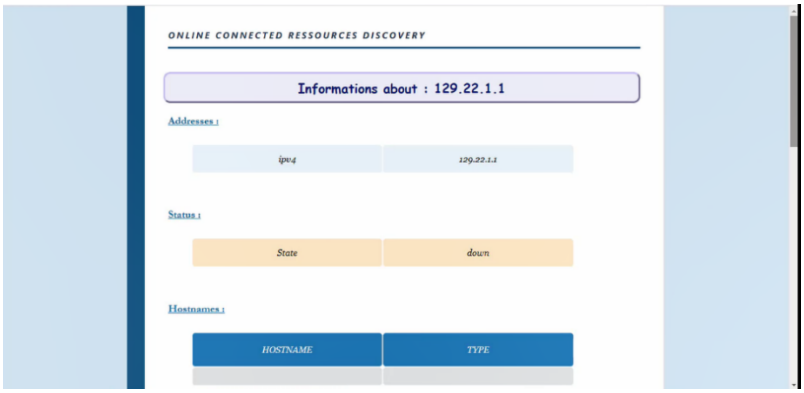

Figure 4. Results GUI: Part 1

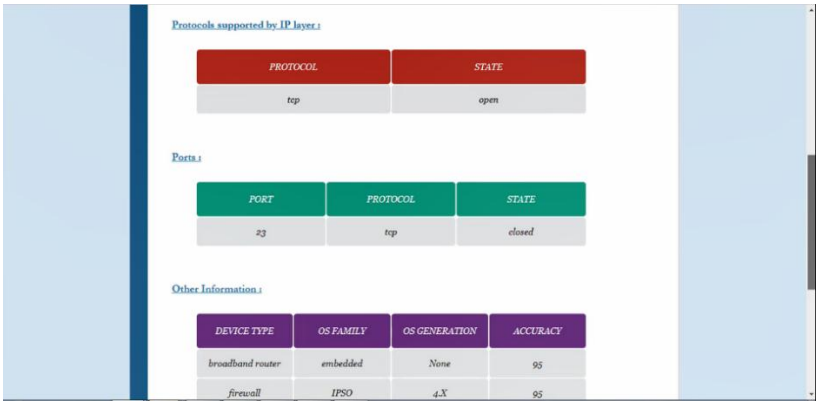

Figure 5. Results GUI: Part 2

5. CONCLUSION

This work has resulted a design of a new tool for real-time retrieval of connected things. The main objective of this tool was to allow real-time and online retrieval of connected devices, using network port scanning which allow collecting data/informations related to these things in real time. The important informations are extracted from the collected data and presented easily to be understandable to all users. For our future works, we will attempt to improve results and evaluate performance of the proposed tool. To this end, we are going to perform a set of tests related to parallel retrieval and response time, use other resources for data collection and improve the security side.

REFERENCES

[1] N. Ismail, E. G. Yassine, and S. Abdelalim, "Towards a semantic web of things framework," *IAES International Journal of Artificial Intelligence*, vol. 8, no. 4, pp. 443, 2019, doi: 10.11591/ijai.v8.i4.pp443-450.

[2] Fatima Zahra Fagroud, Lahbib Ajallouda , El Habib Ben Lahmar, Hicham Toumi and Khadija Achtaich, "IOT search engines: exploratory data analysis," *Procedia Computer Science*, vol. 175, pp. 572-577, 2020, doi: 10.1016/j.procs.2020.07.082.

[3] T. Ghrib, M. Benmohammed, and P. S. Pandey, "Automated diagnosis of attacks in internet of things using machine learning and frequency distribution techniques," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 10, no. 2, pp. 950-961, 2021, doi: 10.11591/eei.v10i2.2766.

[4] A. Shahzad, D. W. Jacob, N. M. Nawi, H. Mahdin, and M. E. Saputri, "The new trend for search engine optimization, tools and techniques," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 18, no. 3, pp. 1568-1583, 2020, doi: 10.11591/ijeecs.v18.i3.pp1568-1583.

[5] M. H. Habaebi, A. Al-Haddad, A. Zyoud, and G. Hijazi, "Micro search engine for IoT: An IoT search engine prototype for private networks," *Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, vol. 11, no. 2, pp. 123-131, 2018, doi: 10.2174/2352096511666180117144450.

[6] M. Faqeeh, M. Al-Ayyoub, M. Wardat, I. Hmeidi, and Y. Jararweh, "Topical search engine for Internet of Things," in *2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA)*, 2014, pp. 829-835, doi: 10.1109/AICCSA.2014.7073287.

[7] X. Feng, Q. Li, H. Wang, L. Sun, "Acquisitional rule-based engine for discovering internet-of-things devices," *27th {USENIX} Security Symposium ({USENIX} Security 18)*, 2018, pp. 327-341.

[8] S. K. Datta, and C. Bonnet, "Search engine based resource discovery framework for Internet of Things," *2015 IEEE 4th Global Conference on Consumer Electronics (GCCE)*, 2015 , pp. 83-85, doi: 10.1109/GCCE.2015.7398707.

[9] A. Shemshadi, Q. Z. Sheng, and Y. Qin, "ThingSeek: A crawler and search engine for the Internet of Things," *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* , 2016, pp. 1149-1152, doi: 10.1145/2911451.2911471.

[10] M. Arnaert, Y. Bertrand, and K. Boudaoud, "Modeling vulnerable internet of things on SHODAN and CENSYS: An ontology for cyber security," in *Proceedings of the Tenth International Conference on Emerging Security Information, Systems and Technologies (SECUREWARE 2016)*, 2016, pp. 299-302.

[11] F. Z. Fagroud, E. H. B. Lahmar, M. Amine, H. Toumi, and S. El Filali, "What does mean search engine for IOT or IOT search engine," *Proceedings of the 4th International Conference on Big Data and Internet of Things*, 2019, pp. 1-7, doi: 10.1145/3372938.3372958.

[12] T. Herman, *Search engines and ethics*, The Stanford Encyclopedia of Philosophy (Fall 2016 Edition), Edward N. Zalta (ed.) 2016. [Online]. Available: https://plato.stanford.edu/archives/fall2016/entries/ethics-search/.

[13] F. Z. Fagroud, E. Ben Lahmar, and S. El Filali, "Internet of things: statistical study on research evolution," *International Journal of Advances in Electronics and Computer Science*, vol. 6, no. 5, 2019.

[14] A. Whitmore, A. Agarwal, and L. D. Xu, "The Internet of Things—A survey of topics and trends," *Information systems frontiers*, vol. 17, no. 2, pp. 261-274, 2015, doi: 10.1007/s10796-014-9489-2.

[15] K. K. Patel, and S. M. Patel, "Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges, "*International journal of engineering science and computing*, vol. 6, no. 5, 2016, doi: 10.4010/2016.1482.

[16] S. N. Ibrahim, A. H. H. Basri, and A. L. Asnawi " Development of web-based surveillance system for Internet of Things (IoT) application," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 8, no. 3, pp. 1108-1116, 2019, doi: 10.11591/eei.v8i3.1520.

[17] J. A. Alvarez-Cedillo, E. Acosta-Gonzaga, M. Aguilar-Fernández, and P. Pérez-Romero, "Internet prospective study," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 6, no. 3, pp. 235-240, 2017, doi: 10.11591/eei.v6i3.653.

[18] F. Khan, R. L. Kumar, S. Kadry, and Y. Nam, "The future of software engineering: Visions of 2025 and beyond, " *International Journal of Electrical & Computer Engineering (IJECE)*, vol. 11, no. 4, pp. 2088-8708, 2021, doi: 10.11591/ijece.v11i4.pp3443-3450.

[19] A. Tiwary, M. Mahato, A. Chidar, M. K. Chandrol, M. Shrivastava, and M. Tripathi, "Internet of Things (IoT): Research, architectures and applications," *International Journal on Future Revolution in Computer Science & Communication Engineering*, vol. 4, no. 3, pp. 23-27, 2018.

[20] N. Yalçın, and U. Köse, "What is search engine optimization: SEO?," *Procedia-Social and Behavioral Sciences*, vol. 9, pp. 487-493, 2010, doi: 10.1016/j.sbspro.2010.12.185.
[21] A. Shemshadi, Q. Z. Sheng, Y. Qin, A. Sun, W. E. Zhang, and L. Yao, "Searching for the internet of things: where it is and what it looks like," *Personal and Ubiquitous Computing*, vol. 21, no. 6, pp. 1097-1112, 2017, doi: 10.1007/s00779-017-1034-0.
[22] F. Liang, C. Qian, W. G. Hatcher, and W. Yu, "Search engine for the internet of things: Lessons from web search, vision, and opportunities," *IEEE Access*, vol. 7, pp. 104673-104691, 2019, doi: 10.1109/ACCESS.2019.2931659.
[23] P. Barnaghi, and A. Sheth, "On searching the Internet of Things: Requirements and challenges, "*IEEE Intelligent Systems*, vol. 31, no. 6, pp. 71-75, 2016, doi: 10.1109/MIS.2016.102.
[24] B. Genge, and C. Enăchescu, "ShoVAT: Shodan- based vulnerability assessment tool for Internet- facing services," *Security and communication networks*, vol. 9, no. 15, pp. 2696-2714, 2016.
[25] R. C. Bodenheim, "Impact of the Shodan computer search engine on internet-facing industrial control system devices," M.Sc thesis, Department of Electrical and Computer Engineering, Air Force Institute of Technology, Air University, U.S., 2014.
[26] S. Otoum, I. A. Ridhawi, and H. Mouftah, "Securing critical IoT infrastructures with blockchain-supported federated learning," *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2021.3088056
[27] A. Rahman, K. R. Kawshik, A. A. Sourav, and A. Gaji, "Advanced network scanning," *American Journal of Engineering Research (AJER)*, vol. 5, no. 6, pp. 38-42, 2016.
[28] F. Z. Fagroud, E. H. Ben Lahmar, H. Toumi, K. Achtaich, and S. El Filali "IOT search engines: study of data collection methods," in *Advances on Smart and Soft Computing*, 2021, pp. 261-272, doi: 10.1007/978-981-15-6048-4_23.
[29] "Nmap," Accessed: Dec. 21, 2019. [Online]. Available: https://nmap.org/
[30] "Angry IP scanner," Accessed: Dec. 2, 2019. [Online]. Available: https://angryip.org/about/.
[31] "Advanced Port Scanner," Accessed: Nov. 10, 2019. [Online]. Available: https://www.advanced-port-scanner.com/fr/
[32] T. Li, S. Wang, F. Luo, F.-X. Wu, and J. Wang, "MultiGuideScan: a multi-processing tool for designing CRISPR guide RNA libraries," *Bioinformatics*, vol. 36, no. 3, pp. 920-921, 2020, doi: 10.1093/bioinformatics/btz616.
[33] "Flask" [Online]. Available: https://www.tutorialspoint.com/flask/index.htm [Accessed: 20-December-2020].
[34] "python-nmap 0.6.1". [Online]. Available: https://pypi.org/project/python-nmap/. [Accessed: 20-March-2020].