BRNO UNIVERSITY OF TECHNOLOGY

Faculty of Electrical Engineering
and Communication

BACHELOR'S THESIS

Brno, 2023 Anel Sagindykova

BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

DEPARTMENT OF TELECOMMUNICATIONS

SECURITY VERIFICATION TOOL FOR INDUSTRIAL AND ENERGY EQUIPMENTS

BACHELOR'S THESIS
BAKALÁŘSKÁ PRÁCE

AUTHOR
AUTOR PRÁCE
Anel Sagindykova

ADVISOR
VEDOUCÍ PRÁCE
Ing. Petr Blažek

BRNO 2023

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

Bachelor's Thesis

Bachelor's study program **Information Security**

Department of Telecommunications

Student: Anel Sagindykova ***ID:*** 211809

Year of study: 3 ***Academic year:*** 2022/23

TITLE OF THESIS:

Security verification tool for industrial and energy equipments

INSTRUCTION:

The bachelor's thesis aims to implement a testing tool for industrial and power equipment in terms of safety. In the theoretical part, the student will analyze the security of industrial and energy networks, focusing on the security features and describing the attacks that have already been implemented. The student will then compare search tools for Internet-connected devices (Shodan, Zoomeye, etc.) and comment on how these search engines work. In the practical part, the student will compare the different search tools based on their tests. In addition, the student will focus on retrieving information about industrial and energy protocols/devices using the API of the selected search engine(s) and creating a tool to store information from the search engine(s) in a database.
The output of the bachelor thesis will be a tool that will provide security testing on selected devices where at least two communication protocols from the industrial and energy sectors will be represented.

RECOMMENDED LITERATURE:

[1] STOUFFER, Keith; FALCO, Joe; SCARFONE, Karen. Guide to industrial control systems (ICS) security. NIST special publication, 2011, 800.82: 16-16
[2] MATHERLY, John. Complete guide to Shodan. Shodan, LLC (2016-02-25), 2015, 1.

Date of project specification: 6.2.2023 ***Deadline for submission:*** 26.5.2023

Supervisor: Ing. Petr Blažek

doc. Ing. Jan Hajný, Ph.D.
Chair of study program board

ABSTRACT

The aim of this work is to analyze various search engines with the intention to select the most suitable for the creation of a tool that should be able to perform security testing. Comparisons were made between Shodan, ZoomEye, Nexpose, Censys and BinaryEdge. These search engines were compared according to different criteria, namely their perfomance and publicly available information. Shodan and ZoomEye were chosen as the most suitable ones and as a result, web application was created. This application is able to retrieve search results from engines' servers using their API's and is able to store these results in a database. The thesis includes a description of a variety of industrial and energy protocols, as well as industrial and energy networks and their security features. Moreover, a few of the most well-known attacks on these networks were described.

KEYWORDS

Shodan, ZoomEye, BinaryEdge, Censys, Modbus, Siemens S7, DNP3, ICS, PLC, RTU, HMI, SCADA, API , IED, JavaScript, Stuxnet

ABSTRAKT

Cílem této práce je analyzovat různé vyhledávače a vybrat ty nejvhodnější pro vytvoření nástroje k testování průmyslových a energetických zařízení z pohledu bezpečnosti. Bylo provedeno porovnání systémů Shodan, ZoomEye, Nexpose, Censys a BinaryEdge. Tyto vyhledávače byly porovnávány podle různých kritérií, například výkonnosti a jejich dalších veřejně dostupných informací. Jako nejvhodnější byly vybrány Shodan a ZoomEye a pomocí těchto vyhledavačů byla vytvořena webové aplikace. Tato aplikace je schopna načítat výsledky vyhledávání ze serverů pomocí jejich rozhraní API a je schopna je ukládat do databáze. Součástí práce je popis různých průmyslových a energetických protokolů, sítí a jejich bezpečnostních prvků. Součástí práce byl také popis několika nejznámějších útoků na tyto sítě.

KLÍČOVÁ SLOVA

Shodan, ZoomEye, BinaryEdge, Censys, Modbus, Siemens S7, DNP3, ICS, PLC, RTU, HMI, SCADA, API , IED, JavaScript, Stuxnet

ROZŠÍŘENÝ ABSTRAKT

Problematika zabezpečení průmyslových a energetických sítí je stále aktuálním a důležitým tématem v dnešní digitální době. Rostoucí digitalizace těchto sítí a jejich připojení k internetu má za následek vznik nových útoků a hrozeb, které vyžadují zvýšenou pozornost a preventivní opatření.

Teoretická část se zaměřuje na popis industrialních a energetických sítí a jejich bezpečnostní prvky. Práce také zahrnuje popisy nejznámějších útoků, kterým tyto sítě čelily v minulosti. Kromě toho byly popsany známé průmyslové protokoly Modbus, S7 a DNP3, ktere se nasledne pouziji v prakticke casti.

V rámci práce byly porovnány následující vyhledávací nástroje: Shodan, ZoomEye, Censys, BinaryEdge a Nexpose. Tyto vyhledávače narozdíl od obyčejných vyhledávačů jako např. Google se liší tím, že neprohledávájí webové stránky, ale vyhledávají a získávají informace o zařízeních připojených k internetu.

Porovnání bylo rozděleno do teoretické a praktické části. Teoretické porovnání bylo založené na veřejně přístupných informacích a je zaměřeno na různá kritéria, včetně schopnosti pracovat s API a CLI, výhod pro studenty, roku založení, seznamu portů a protokolů s jakými mohou pracovat a v jakém množství. Práce nabízí popis, jak tyto vyhledávače fungují, kolik dat a dotazů mohou poskytnout, a především jejich interakci s protokoly a porty z průmyslového a energetického odvětví.

Pro praktické porovnání byly do školní sítě VUT zapojeny tři zařízení. Úkolem vyhledávačů bylo najít tato zařízení v co nejkratším čase. Ze všech vyhledávačů byly vybrány jen dva - Shodan a ZoomEye z následujících důvodů: Shodan byl nejrychlejší a poskytuje členství pro studenty, ZoomEye byl vybrán neboť našel všechna tři zařízení. Vyhledávač Shodan je unikátní v tom, že uživatel s ním může pracovat prostřednictvím tří možných přístupů: API, CLI nebo webového rozhraní. ZoomEye vyhledávač se ukázal jako spolehlivá druhá volba, která dokázala efektivně nalézt všechna zařízení s vysokou rychlostí a relativně jednoduchou následnou implementací do vytvářené aplikace.

Výsledky tohoto porovnání určily nejvhodnější vyhledávače pro vývoj webové aplikace, která využívá API těchto vyhledávačů a umožňuje uživatelům ukládat a zobrazovat výsledky vyhledávání z tohoto nástroje.

Webová aplikace byla vyvinuta v programovacím jazyce JavaScript ve spojení s Node.js a Express.js, což umožnilo rychlý vývoj a snadnou integraci API vyhledávačů. Aplikace poskytuje uživatelům intuitivní rozhraní pro registraci nebo přihlášení a vyhledávání zařízení. Pro využití aplikací je možné vložit jakýkoli dotaz, podporovaný pravidly API vyhledávačů, kdy se následně vytvoří JSON file a stáhne se na zařízení uživatele. Vytvořený file se také uloží do databáze, kam uživatel ma přístup. Pro pohodlí uživatele je také možné JSON file opětovně z databáze

stáhnout.

Posledním cílem této práce bylo provést bezpečnostní testování nástroje na dvou vybraných zařízeních, ve kterých budou zastoupeny dva komunikační protokoly z průmyslového a energetického odvětví. Pro testování byly vybrány protokoly Modbus a S7. Oba protokoly mají širokou podporu mezi průmyslovými zařízení a výrobci. Jsou jednoduché, spolehlivé a používané v průmyslovém prostředí již mnoho let.

Aplikace byla schopna detekovat obě zařízení a získat informace o těchto zařízeních z obou vyhledávačů. Výsledky byly uloženy do JSON filu a následně do databáze aplikace. JSON file obsahoval veškeré nutné informace včetně protokolu Modbus a S7.

Na konci se práce zabývala možnými způsoby skrývání informací před vyhledávači a obecně o tom jak zlepšit zabezpečení sítě a zařízení.

Author's Declaration

Author:	Anel Sagindykova
Author's ID:	211809
Paper type:	Bachelor's Thesis
Academic year:	2022/23
Topic:	SECURITY VERIFICATION TOOL FOR INDUSTRIAL AND ENERGY EQUIPMENTS

I declare that I have written this paper independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the paper and listed in the comprehensive bibliography at the end of the paper.

As the author, I furthermore declare that, with respect to the creation of this paper, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll. of the Czech Republic, Section 2, Head VI, Part 4.

Brno ....

author's signature*

*The author signs only in the printed version.

ACKNOWLEDGEMENT

I would like to thank the advisor of my thesis, Ing.Petr Blažek for his valuable comments, patience and willingness to continue to work with me.

Contents

List of Figures

List of Tables

Listings

Introduction

Industrial and energy networks are crucial to the functioning of modern society because they supply homes, companies, and significant infrastructure with basic services like power, water, and gas. However, as these networks become more connected to the Internet, there are also new security vulnerabilities that, if ignored, might have serious consequences. These networks frequently rely on old protocols and technology that are simple targets for hackers.

When connecting different devices to the Internet, it is extremely important to have a solid understanding of the kinds of threats that such devices may be vulnerable to. There are numerous different approaches that can be taken to ensure the safety of the equipment. The first chapter 1 includes descriptions of industrial and energy networks, their security features and components. Furthermore, some of the famous attacks on these networks were described. Additionally, there are special search engines that can provide us with the information regarding what is known on the Internet about our gadgets and what kinds of attacks could potentially be launched against them.

One of the primary objective of this work was to investigate and analyze the operation of various search engines, with the goal of selecting the most suitable for the use in the development of an application in the future. In particular, five search engines — Shodan, ZoomEye, BinaryEdge, Censys, and Nexpose — were chosen, described and compared. In contrast to search engines such as Google, these search engines gather information on devices that are connected to the Internet. The comparison of search engines was divided into several parts. In the second chapter 2, a straightforward explanation of each one is given, and in addition to this, a description of the protocols that will be implemented in the subsequent practical section is provided 2.6.

In the first section of the third chapter 3, a comparison of search engines was carried out according to a number of criteria and data which are publicly available. Some of the criteria are as follows: what are the benefits for students, which ports and protocols search engines are able to work with and in what quantity, well-written documentation for future work, the ability to work with both IPv4 and IPv6 addresses. The second aspect of this comparison focused on search results, which refers to how many and how quickly search engines are able to find devices that are connected from the school network 3.2. To choose the option that is best suited, it is necessary to consider all of the comparisons that were made during the process. The final goal, which follow the selection of the most effective search engines, entails the development of a tool that is able to perform security testing on electronic devices that are currently connected from the school's network.

1 Industrial and energy networks

Energy networks are examples of critical infrastructure systems that are necessary for the successful operation of modern society. They include power grids, oil and gas pipelines, and other systems that distribute energy from its source to the end-users.

Because any disturbance in the energy supply can have catastrophic effects, including blackouts, financial losses, and even dangers to public safety, the security of energy networks is very important. Because of the growing use of digital technology in energy networks, these networks are now increasingly sensitive to cyber attacks, which can be initiated from a remote location by cybercriminals.

Same goes for ICS security. With the increasing connectivity of ICS networks to the Internet and other networks, they open themselves up to the possibility of being attacked in cyberspace. In order to cause disruptions, steal sensitive data, or inflict physical harm, hackers or state-sponsored groups may try to exploit vulnerabilities in ICS networks. Strong ICS security policies are necessary for fighting against cyber threats and minimizing the possible impact of attacks. These standards include network segmentation, encryption, access limits, and monitoring. The number of ICS cyber attacks has recently increased significantly compared to the year 1997 as shown in Fig 1.1.

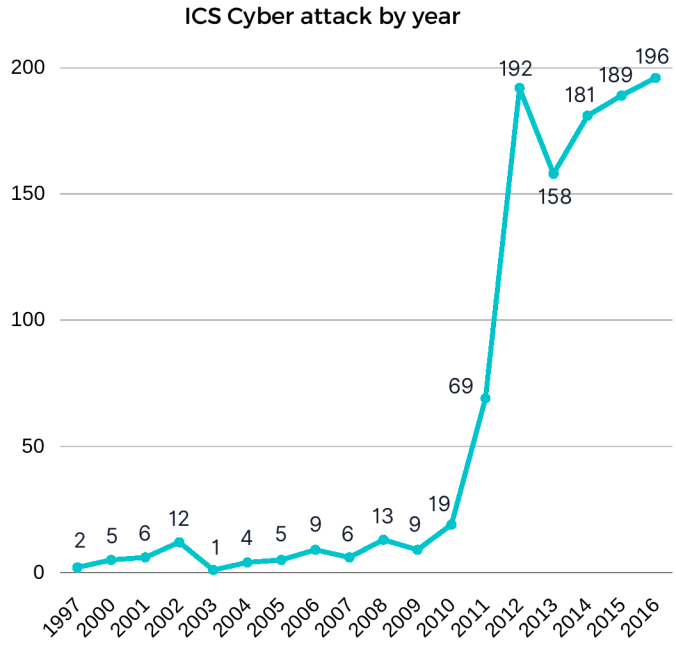

Fig. 1.1: ICS Cyber attacks by year [1]

1.1 ICS Operation and Components

The term "Industrial Control System" (ICS) is a general term that refers to several different kinds of control systems. These control systems include Supervisory Control and Data Acquisition Systems (SCADA), Distributed Control Systems (DCS) and Process Control Systems (PSC) as shown in Fig 1.2. These systems may include Human-Machine Interfaces (HMI), Remote Terminal Unit (RTU), Intelligent Electronic Device (IED) and other control system configurations such as programmable logic controllers (PLC), which can frequently be found in industrial sectors and critical infrastructures. The basic function of an ICS is shown in the Fig 1.3. An ICS contains a wide variety of control loops, human interfaces, as well as remote diagnostics and maintenance tools, all of which are constructed utilizing layered network structures and an array of network protocols [3].

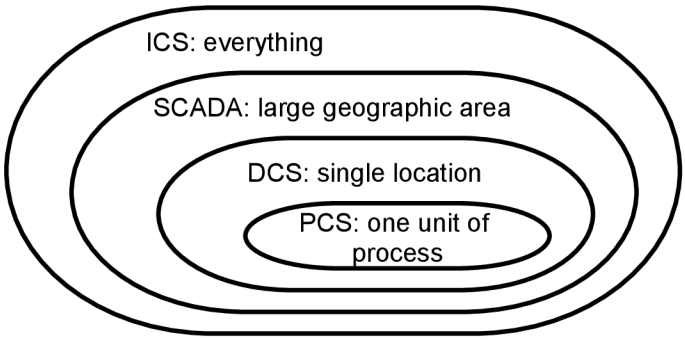

Fig. 1.2: ICS basic parts [2]

The percentage distribution of various types of ICS components is presented in Table 1.1. SCADA and HMI-based systems make the largest portion of the sector's share, followed by PLC and hardware-based systems.

Type	**Global Percentage**
SCADA/HMI	31
SCADA	12
PLC	27
HMI	27
HARDWARE	3

Tab. 1.1: ICS Components Distribution Percentage [1]

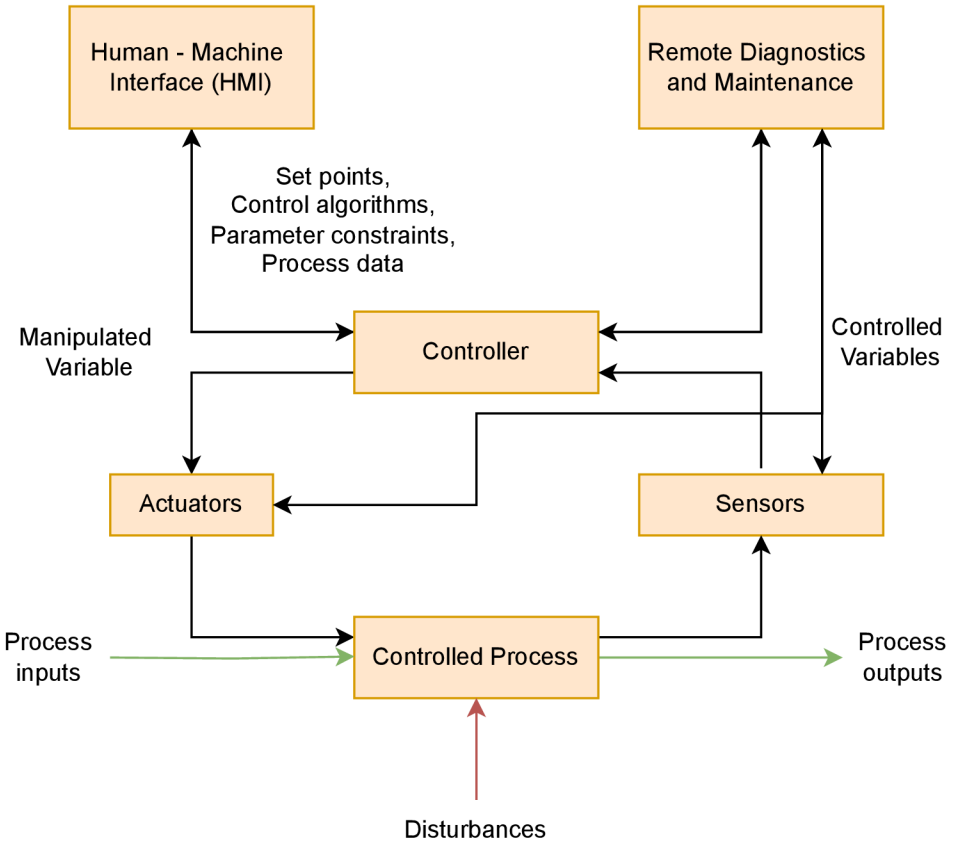

Fig. 1.3: ICS Operation [3]

1.1.1 Supervisory Control And Data Acquisition

SCADA is not a system that is capable of providing complete control on its own. Instead, the focus of its capabilities is placed on the provision of control at the supervisory level. SCADA systems can acquire and send data, and they are coupled with HMI, which offers centralized monitoring and control for a large number of process inputs and outputs.

Utilizing SCADA for long-distance monitoring and management of field sites via a centralized control system is the major goal of doing so. A SCADA system is able to automate the task of workers having to travel large distances in order to execute activities or gather data. Field devices collect data from the sensor systems, monitor the local environment for alert situations, and regulate local operations such as opening and closing valves and breakers. SCADA systems are utilized often in a variety of business sectors, including those concerned with the monitoring and control of pipelines, water treatment centers and distribution, and electrical power transmission and distribution [4].

The components and general configuration of a SCADA system are depicted

in Fig 1.4. A control server and communications routers are located within the control center. Other components of the control center include a LAN-connected HMI, engineering workstations, and a data historian. The control center collects and records data gathered by field sites, displays data to the HMI, and may take action in response to detected events [3]. At this point in time, the majority of SCADA systems are built on a distributed approach; yet, when they first came into being, the central architectural approach was the most common kind used [7].

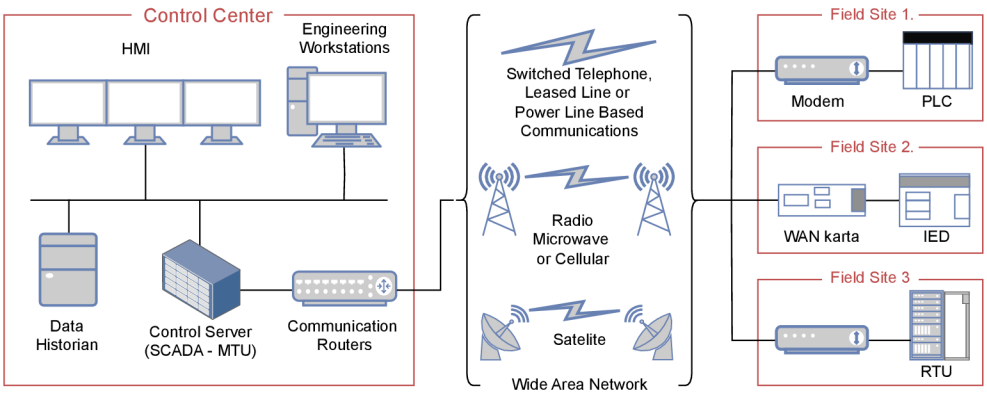

Fig. 1.4: SCADA System General Layout [3]

Some of common threaths faces by SCADA systems are following:

- Malware - If a computer system has a poor security posture, there is an 80% chance that malicious software has already gained access to the system. When it comes to attacks by malware, SCADA systems are no exception. Due to the weak security posture of SCADA systems, these systems are frequently an easy target for attackers.
- DDoS attacks - More than a few DDoS vulnerabilities exist in ICSs, including the use of commercial communication protocols, insecure passwords in internet of things (IoT) devices, and weak passwords in open-source software [6].
- Command injection - In a SCADA system, a high level of authorisation is required for users to be able to send instructions to the system shell. Hackers were able to gain such a high level of control over the target system that they were able to execute arbitrary commands that were capable of changing a variety of parameters. Command injection attacks can be carried out since there is no procedure to validate the data that is supplied by users.
- Networking issues - Misconfigured networks are a common occurrence in SCADA systems.
- Human errors - Employees who click on links in phishing emails let malware spread throughout the company's networks. The usage of simple passwords or

passwords that are too short might also lead to difficulties [5].

1.1.2 Distributed Control System

A digital automated ICS is referred to as a distributed control system (DCS), and it is a type of industrial control system that uses control loops that are geographically dispersed throughout a factory, machine, or control region. Controlling industrial processes in order to improve their safety, cost-effectiveness, and reliability is the purpose of a DCS.

A control system is a collection of mechanical or electrical components that, through the implementation of control loops, are used to regulate the operation of other devices or systems. Control loops are systems that are made up of all of the hardware and software control functions that are required for making the necessary changes and measurements in a certain procedure. Control systems are an essential component of all automation systems, including those used in manufacturing and the process industries [8].

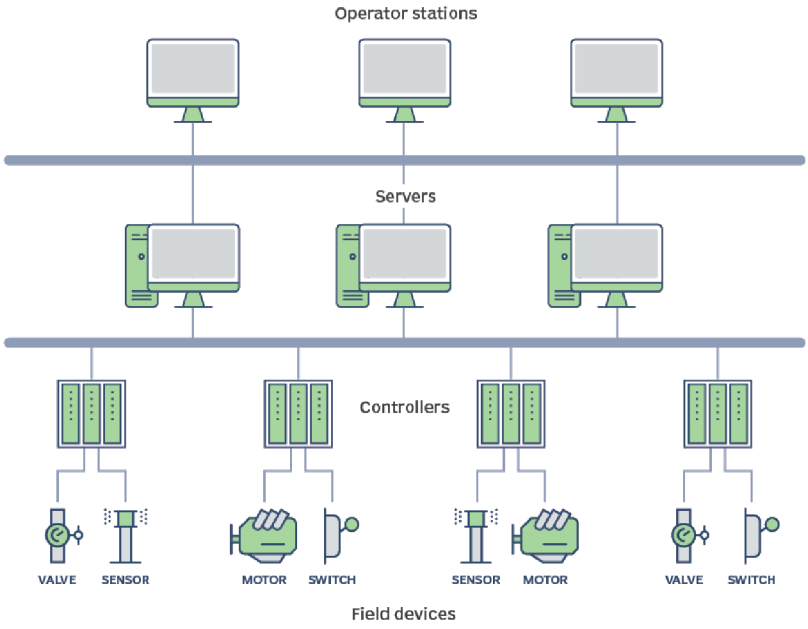

Fig. 1.5: A Distributed control system example [8]

1.1.3 Programmable Logic Controller

PLCs are utilized in both SCADA and DCS systems as control components of an overall hierarchical system to enable local management of processes through feedback control.

PLCs are also used as the main controller in smaller control system designs to offer operational control of discrete processes such as automobile assembly lines and power plant soot blower controls. These topologies are distinct from SCADA and DCS in that, in most cases, they do not include a central control server or an HMI. As a result, their primary function is to offer closed-loop control that does not involve the participation of direct human intervention. PLCs come equipped with a memory that can be programmed by the end user and used to store instructions for the purpose of executing particular operations. These functions include I/O control, logic, timing, counting, three mode proportional-integral-derivative (PID) control, communication, arithmetic, and the processing of data and files [3].

The primary components of the programmable logic controller (PLC) are the central processing unit (CPU), memory, and I/O modules for managing input/output data. PLCs have the fundamental architecture that is shown in Fig 1.6. PLC has four primary parts:

- The program memory - memory space where the program instructions for the logical control sequence are kept.
- The data memory - the state of inputs and outputs, such as switches and interlocks the values of any other relevant working data are saved.
- The input devices - these are the hardware and software inputs that come from the area during the industrial process. Signals may originate from sensors, switches, proximity detectors, interlock settings, and other sources as well. These inputs cause the sequences in the user program to be triggered, which results in the process or output that was requested. For instance, the Emergency stop input is constantly watched by the PLC program. In the event that this switch is activated as a result of an incident or accident, the entire PLC process is immediately paused and brought to a complete standstill.
- The output devices - The most common types of output devices include solenoid valves and pneumatic actuators, motors, heaters, cooling fan motors, alert indicators, and buzzers. These tools rule the industrial operations [9].

ICSs make use of a wide number of different protocols in order to connect with field devices such as sensors, actuators as well as for programming and communication with PLCs in the process network. There are many different kinds of protocols, but among of the most common ones are MODBUS, Ethernet/IP, DNP 317, and ISO-TSAP. Despite the fact that these protocols are quite effective in terms of communication, security was not a primary issue when the protocols were initially developed because there was no need for it in industrial systems. Therefore, the protocols were not designed to ensure security. As a result, these protocols do not provide secrecy, authentication or data integrity [10].

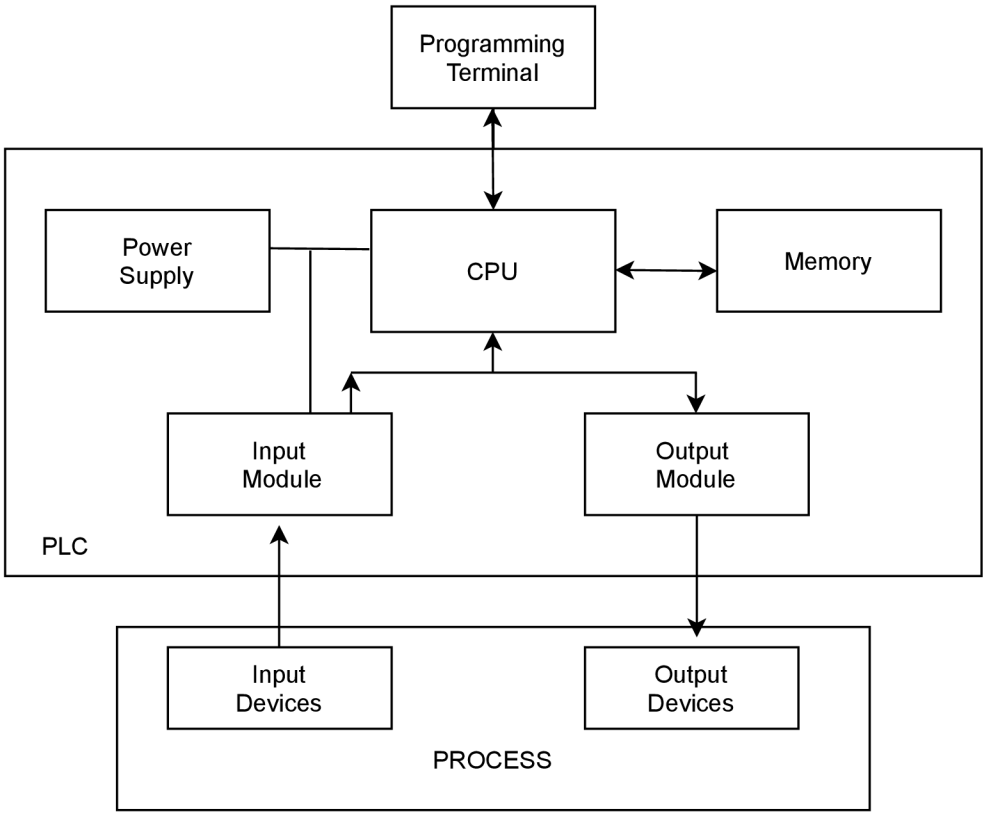

Fig. 1.6: Block diagram of PLC with I/O [9]

A PLC may become the target of a cyberattack in a variety of different ways, including the following:

- Network-based attacks: A PLC that is connected to a network has the potential to be compromised if an attacker is successful in gaining access to the network. To accomplish this, it may be necessary to exploit weaknesses in the network infrastructure or in the PLC device itself.
- Physical attacks: Anyone who can gain physical access to a programmable a PLC may be able to compromise it by connecting to it directly or installing malicious software onto it.
- Malicious software: PLCs are vulnerable to infection with malware through a variety of entry points, including the opening of infected emails or the downloading of infected software updates. [10]

1.1.4 Human-Machine Interfaces

Human-Machine Interfaces (HMIs) is an application that uses a graphical user interface (GUI) and enables interaction between a human operator and the controller hardware. Additionally, it is able to provide current information as well as historical

data that was collected by the devices operating in an ICS environment. In addition to this, it may be used to monitor and modify setpoints, control algorithms, as well as adjust and set parameters in controllers [4].

1.1.5 Remote Terminal Unit

The remote terminal unit (RTU) is an independent data acquisition and control unit. The most important purpose is to regulate and acquire data from process equipment located at the remote location, and then to transmit that data back to a central unit. The data might be in the form of electric parameters such as the RMS value of voltage and current, frequency, active power etc., or they could be in the form of other quantities such as temperatures, oil levels, switch status, and so on. Analog, digital, and control data are the three types of information that can be monitored by RTU [11].

1.1.6 Intelligent Electronic Devices

Intelligent Electronic Devices (IED) are electronic devices that may contain a single microprocessor or numerous microprocessors integrated within it. Transmission or receiving of data or control signals to or from an external device is its primary function [7].

1.2 Energy networks

Energy networks are the lower pressure gas pipes and low, medium, and high voltage electricity lines that carry and distribute gas and electricity from energy transmission systems directly to households, businesses, and industry [12].

There are many distinct kinds of energy networks, each of which is intended to make the transmission and distribution of a certain kind of energy easier. The following are some examples of common kinds of energy networks:

- Electrical Grid - the electrical grid is the complex system that is meant to distribute electricity to customers all the way from the point at which it is generated to the point at which it is used daily by those customers. These networks started out as small local designs and have since expanded to cover thousands of kilometers and connect millions of homes and businesses. Despite the fact that the grid is made up of a limitless amount of complex interactions, it can be broken down into three primary categories: generation, transmission, and distribution of electricity [13].

- Natural Gas Network - is a network of pipelines that carries natural gas, typically methane, from a producing facility such as a refinery or bulk distributor to clients. Gas distribution systems are sometimes referred to by their other names: propane distribution systems and butane distribution systems. It is constructed in such a way as to guarantee that the correct quantity and pressure of gas is delivered to each individual client and to prevent any potentially dangerous leaks in the process. Gas distribution systems are an essential component of our day-to-day lives since they supply gas to commercial establishments all around. They are an essential component of the current infrastructure, which enables the secure and effective distribution of natural gas products with a minimum of adverse effects on the surrounding environment. Gas transportation and distribution are essential components of the energy industry. An extensive network of pipelines and other equipment is required in order to carry this essential resource in a manner that is both secure and effective. This includes the installation of pipelines, which are normally responsible for transporting the gas from the production sites to the processing plants and storage facilities. In addition to this, there must be specialized equipment in order to manage and direct the flow of the gas itself. [14].
- District Heating and Cooling systems - distinct heating also known as heat networks or teleheating, is a system that distributes heat generated in a centralized place through a network of insulated pipes that serve the heating requirements for residential and commercial properties, such as heating the area in which people live and the water used for domestic and commercial purposes. The heat is typically derived from a cogeneration plant that burns fossil fuels or biomass; but, heat-only boiler stations, geothermal heating, heat pumps, and central solar heating are also utilized, in addition to heat waste from companies and heat generated by nuclear power electricity generation. Local boilers may not be able to deliver the same level of efficiency nor the same level of pollution control as district heating plants. District heating with combined heat and power (CHPDH) is the approach that reduces carbon emissions at the lowest cost and has one of the smallest carbon footprints of all fossil production facilities. [15]. District cooling is a method of distributing thermal energy that is obtained through a central cooling plant in the form of chilled water. Both of these models, as well as their combination in heating and cooling solutions, stand out as an alternative system that is helping cities decarbonize their energy consumption and production. In addition, this system is delivering a variety of other benefits that are beneficial to the economy and the environment. [16]
- Water Supply and Distribution Systems - infrastructure for the collection,

transmission, treatment, storage, and distribution of water for residences, business enterprises, industries, and irrigation, as well as for public purposes such as firefighting and street flushing. The provision of drinkable water is widely regarded as being the most important of all municipal services. People are dependent on water for a variety of domestic purposes including drinking, cooking, cleaning etc. In addition to this, water supply systems need to fulfill the requirements for public, commercial, and industrial activities. In every instance, the water must be able to satisfy both the quality and quantity criteria [17].

- Oil Pipeline Network - pipelines are piping systems that are often buried underground and are used to transport and distribute fluids. The fluids that are typically referred to in the context of the energy industry are either crude oil, oil products, or natural gas. Long-distance shipping of crude oil and refined petroleum products is often accomplished through the utilization of oil pipeline networks. These networks are made up of a variety of infrastructure, including pipelines, pumping stations, storage terminals, and other facilities. They are extremely important in the process of transporting crude oil from its point of production to the various refineries and distribution hubs [18].

It is important to keep in mind that the above are but a few instances of energy networks, and that there are more varieties that are unique to the various types of energy sources and geographic locations. The particular configuration and components of each network can differ from one another depending on the geographical region, the sources of energy, and the requirements for the local infrastructure.

1.3 Security features

Due to the seriousness and potential effects of security incidents, it is essential to consider how to secure industrial and energy networks. It is crucial to protect the integrity, availability, and confidentiality of these networks because they are in charge of managing and running energy systems and industrial operations.

The following is a summary of various significant security components necessary for protecting industrial and energy networks:

Network Segmentation and Segregation

One of the most effective architectural principles that a company can use to protect its ICS is network segmentation and segregation. Segmentation creates security domains, or enclaves, which are frequently described as being controlled by the same authority, upholding the same policy, and having the uniform level of trust. The

effects of non-hostile errors and accidents can be contained through segmentation, which can reduce the manner and level of access to critical information, ICS communication, and equipment configuration. Segmentation can also make it substantially more difficult for a malicious cyber adversary.

In order to limit access to sensitive information for systems and users who don't need it and to maintain the organization's ability to function effectively, network segmentation and segregation are used. Depending on the design and configuration of the network, several techniques and technologies can be used to accomplish this.

Network segmentation and segregation are often implemented at the gateway between domains. ICS environments frequently include gateways to non-ICS and less reliable domains like the Internet and corporate LANs in addition to numerous clearly defined domains like operational LANs, control LANs, and operational DMZs. Protecting domain gateways is sensible and deserving of consideration when insider assaults, social engineering, mobile devices, and other vulnerabilities and predisposing factors are taken into account. The process of network segregation entails creating and implementing rules that regulate which communications are allowed to pass over the border. Usually, rules are determined by the identities of the source and destination as well as the kind or substance of the data being sent.

Correct network segmentation and segregation implementation minimizes the way and extent of access to sensitive data. It is possible to accomplish this utilizing a range of tools and techniques. Some of the typical technologies and techniques utilized depend on the design and configuration of a network. They include:

- Physical network separation to entirely prevent any traffic across domains from connecting.
- Network traffic filtering, which can make use of a range of technologies at different network layers to establish security criteria and domains.
- Logical network division enforced by network devices or encryption.

There are four fundamental elements that, regardless of the technology used to accomplish network segmentation and segregation, implement the principle of defense-in-depth by allowing for effective network segmentation and segregation:

1. Applying technologies at more than just the network layer. Every network and system should be divided into separate sections, starting at the data link layer and going all the way up to the application layer.

2. Following the need-to-know and least privilege standards. A system shouldn't be permitted to communicate with another system if it is not required to do so. A system should be restricted if it only needs to communicate with other systems using a particular port or protocol or if it only has to send a small amount of data in a predetermined or designated format.

3. Dividing infrastructure and information according to security needs. This

can entail using various hardware or software platforms according to the various security and risk settings that each system or network segment runs in. More strict separation from other components is necessary for the most important components. Virtualization could be used in addition to network segregation to achieve the necessary isolation.

4. Utilizing a whitelisting strategy rather than a blacklisting strategy, i.e., allow access to the known good instead of restricting access to the known bad. Whitelisting is more useful because the programs that run in ICS are largely static. The ability of an organization to analyze log files will also be enhanced by this [3].

Boundary Protection

In order to protect the ICS against malicious cyber adversaries as well as unintentional mistakes and accidents, boundary protection devices regulate the flow of information between interconnected security domains. It is possible for information transfers between systems representing several security domains to break one or more of the domain security policies. In certain architectural solutions that enforce particular security policies, boundary protection devices play a vital role. Companies are able to isolate ICS and business system components. This isolation prevents illegal information from circulating between system components and offers the option to give some components more protection. The capacity to strengthen component protection and better manage information flows between those components is made possible by dividing system components using border protection techniques.

A few examples of boundary protection measures are gateways, routers, firewalls, guards, network-based malware analysis and virtualization systems, intrusion detection systems (both networked and host-based), encrypted tunnels, controlled interfaces, mail gateways, and unidirectional gateways (such as data diodes). It is frequently done by looking at the data or related metadata that boundary protection devices use to decide whether data transfer is allowed [3].

Physical security

The physical protection system prevents sabotage that damages assets, steals valuable resources, disrupts service, or causes an electric grid cascading failure. The physical protection system detects and evaluates threats using multiple sensing methods and delays threat actions to neutralize or mitigate system loss. Every physical security system needs detection, delay, and response.

- Detection: Some important facilities use passive or active infrared sensors, microwave sensors, or other well-established detection methods and thermal

or infrared camera systems. Control houses and transformer sites may also include motion sensors and fixed cameras on the outside.

- Delay: A perimeter delay is usually accomplished with a chain link fence with vehicle and personnel gates.
- Response: It is essential to make preparations in advance in order to defend against, avoid, or lessen the effects of an attack or incident. [19]

1.4 History of attacks

1.4.1 Stuxnet attack

A crucial component of Iran's nuclear program was targeted by the sophisticated computer worm known as Stuxnet, which was developed jointly by spy agencies in the United States and Israel. It was intended for a facility that was air-gapped, but it unexpectedly expanded to include computer systems outside of the facility, which raises a variety of problems regarding its design and purpose. Stuxnet took advantage of a number of vulnerabilities in Windows that were undiscovered at the time. This explanation ought to make it abundantly evident that the Stuxnet worm was deployed as part of a high-level sabotage operation fought by nation-states against their rivals.

It is now a commonly held belief that the spy agencies of the United States and Israel were responsible for the creation of Stuxnet. The information security community made the discovery that Stuxnet existed in 2010, but it is believed that its development began in 2005. The governments of the United States and Israel designed Stuxnet with the intention of using it as a weapon to thwart or, at the very least, postpone the Iranian drive to develop nuclear weapons.

When Stuxnet infects a computer, it first determines whether or not the computer is linked to certain programmable logic controllers (PLCs) manufactured by Siemens. These PLCs are model-specific. PLCs are the means by which computers communicate with and exert control over various pieces of industrial gear, such as uranium centrifuges. If no PLCs are found, the worm will do nothing; but if it does find PLCs, Stuxnet will change the programming of the PLCs, which will cause the centrifuges to spin erratically, perhaps causing them to become damaged or destroyed. Because of this, it is difficult to discover or diagnose what is going wrong until it is too late. This is because the PLCs are communicating with the controller computer and telling it (incorrectly) that everything is running properly while this is occurring.

Despite the fact that security researchers do not have access to the Stuxnet codebase, they have been able to learn a great deal about it by analyzing it. They

have discovered that it was developed in numerous languages, including C and C++, as well as presumably several other object-oriented languages. This, too, is not typical of malware, which is further evidence of the high amount of complexity that went into its development [20].

1.4.2 Triton Attack

The malicious software known as Triton was used in a cyber attack in 2017 that was carried out by a Russian research institute that was funded by the Russian government against a Middle Eastern petrochemical facility. This was the first time the cybersecurity world had ever seen code that was purposefully meant to put people's lives in danger. According to a warning given by the FBI, malware continues to be a threat to the global energy sector.

The hackers had installed harmful software, sometimes known as malware, which gave them the ability to take control of the plant's safety instrumented systems. These physical controllers and the software that goes along with them make up the very last line of protection against potentially deadly catastrophes. They are designed to take action if they detect dangerous conditions, either bringing the processes down to acceptable levels or terminating them entirely by activating mechanisms such as pressure-relief valves and cutoff valves.

Because of the malware, it was possible to remotely take control of these systems. The consequences may have been devastating if the attackers had deactivated or tampered with them and then used other software to make the machinery at the plant malfunction. Fortunately, a bug in the software exposed the hackers before they could cause any damage. In June 2017, it was the cause of a response from a safety system, which ultimately resulted in the facility coming to a stop. Then in August, numerous further systems failed, which lead to another shutdown.The first outage was incorrectly attributed to a problem with a mechanical glitch; after the second outage, the owners of the factory contacted investigators.

The rogue code had the potential to trigger explosions or a leak of poisonous hydrogen sulfide gas, putting the lives of those both inside the facility and in the surrounding area in danger.

The specifics of how Trisis and Triton communicate with the SIS controllers through this unique protocol are not publicly available. However, researchers and cybersecurity specialists who have examined the virus have formed the hypothesis that the attackers reverse-engineered the unique communication protocol that is used by the SIS controllers of the target system.

The attackers were able to construct an unique implementation of the protocol by reverse-engineering it, which allowed the malware to send commands directly to the

controllers. This allowed the attackers to gain control of the system. This displays a high level of technical competence and resources, in addition to a comprehensive understanding of the architecture and protocols of the target system [21].

1.4.3 NotPetya Attack

NotPetya, which first targeted Ukraine, swiftly spread to more than 60 nations, wiping down the computer systems of thousands of global corporations. More than 10 billion dollars in harm has been estimated to have been done by NotPetya. NotPetya had its name from the ransomware Petya, deployed the previous year.

Both encrypted Windows machines and after that demanded a cryptographic payment in exchange for decryption keys. However, NotPetya did not permit the decryption of the victim's computers after payment, whereas Petya did.

It had been created to make it technically impossible to recover the victim's files, whether they paid the ransom or not. It was designed to look like a typical ransomware program.

Another famous ransomware attack WannaCry also encrypted a user's files in affected computers and mounted disks attached to these computers. Then the now-famous warning screen appeared, requesting payment in Bitcoin to unlock the data.

NotPetya does all this too, it infects a machine via a hached "update" to accounting software from a Ukrainian software provider. And if NotPetya were just pure ransomware, created to get as many ransom payments as possible, it might have ended there. However, NotPetya has the ability to rewrite a hard drive's so-called master boot record, which instructs the computer what operating system to run and where to find it, inflicting a more damaging attack on a system, depending on the level of access it has [22] [23].

1.4.4 Dragonfly 2.0 Attack

In October of 2017, Symantec released a report in which it claimed that the energy industry was being targeted by a sophisticated attack group referred as another version of "Dragonfly". According to the research, this organization possessed a significant amount of resources, including a wide variety of malware tools, and had the ability to carry out attacks via a variety of different entry points. This new behavior using Dragonfly was referred to by Symantec as "Dragonfly 2.0." Dragonfly 2.0 carried out a nasty attack campaign, during which it was successful in compromising a number of ICS equipment providers and infecting their software with a RAT.

The Dragonfly 2.0 campaign demonstrates how attackers may be moving into a new phase, in which new campaigns may provide them with access to operational systems; this access may be utilized for more disruptive objectives in the future. .

According to Symantec, this group looked to be interested in both getting access to operational systems and understanding how energy facilities run. One of the most concerning assertions made in the paper is that Dragonfly 2.0 possesses the capability to either sabotage or take control of ICSs [24].

1.4.5 Shamoon attack

In August of 2012, a group of hackers known as the "Cutting sword of justice" announced that they were responsible for the attack that was carried out against the Saudi Aramco, one of the world's largest oil companies.In a matter of hours, 35,000 computers were partially wiped or totally destroyed. They've created a virus named Shamoon, also known as "Disttrack". Shamoon is a modular computer virus that was discovered in 2012, targeting then-recent 32-bit NT kernel versions of Microsoft Windows. In fact, the virus is composed of malicious code that can destroy data stored on the computer as well as rewrite the Master Boot Record (MBR), leaving the device useless. Because the impacted organizations lacked application and implementation of defense-in-depth security, the virus was frequently able to successfully execute in those enterprises.

The schema looks like following: emails containing spear phishing attacks are sent to the employees of the firm. Attachments in the form of Microsoft Office documents are included in the emails.

- When the emailed attachments are opened, PowerShell is activated, which in turn permits command line access to the machine that has been infected.
- The attacker is now able to communicate with the compromised machines and remotely carry out a variety of commands.
- The attacker adds more malicious software and tools to endpoints or elevates their privileges within the network.
- The attacker does a network analysis by establishing connections to additional systems and determining which servers are most important.
- The perpetrator of the attack activates the Shamoon virus.
- A well-orchestrated Shamoon outbreak occurs, which results in the complete and total deletion of all data stored on the computer's hard drives.

The Shamoon incident brought to light the vital importance of strong cybersecurity security measures for enterprises, in particular those organizations that operate in critical infrastructure sectors like the energy sector. It served as a wake-up call for businesses, alerting them to the necessity of investing in sufficient security measures to protect their systems, data, and operations against sophisticated cyber threats [25] [26].

2 Search Engines and Protocols

When looking for websites, search engines such as Google, Yahoo or Microsoft bing are excellent options. These search engines are one of the most important resources for gathering information, and we make regular use of them. On the other hand, there are other search engines that work in a different way and are able to retrieve a variety of information for us. The following subsections sections (2.1, 2.3, 2.4, 2.2, 2.5) will cover Shodan, ZoomEye, Nexpose, Censys, BinaryEdge search engines and how they operate. These search engines are primarily used to scan and index services and devices that are linked to the internet. They're frequently used for security analysis and research. These engines can be used legitimately in the realm of cybersecurity for things like monitoring the security environment, discovering weak points in systems, and determining how exposed devices and services are to the Internet.

The operation of the Modbus, S7, and DNP3 protocols is described and explained in the subsection 2.6. These protocols were selected to be used later in the work's practical part. They are widely used protocols in industrial automation and control systems because they provide effective and standardized device communication, backward compatibility with older equipment, and support for particular industry standards.

2.1 Shodan

The first engine for description will be Shodan. Shodan is a search engine that provides users with the ability to search for a wide variety of electronic devices (including webcams, routers, servers, smart TVs, refrigerators, traffic lights, heating systems) that are connected to the internet using a number of different filters. If a device is directly connected to the Internet, Shodan will query it for various publicly accessible data. The idea for searching Internet-connected devices was devised in 2003 by computer programmer John Matherly, who released the service with Web User Interface (Fig. 2.1) in 2009 which is available at Shodan website [27, 28].

Besides Web User Interface, Shodan can be accessed through Command Line Interface(CLI) or Application Programming Interface(API). Shodan crawls the web for devices using a global network of computers and servers 24/7. It takes about a month for Shodan to scan 500 million devices [29]. The scanner is capable of scanning devices in IPv4 as well as IPv6 address ranges, and the data collected from scanned devices is what makes up the Shodan database. Shodan's primary data collection unit is the banner. The banner contains textual information about service on a device. The information can differs widely depending on the type of

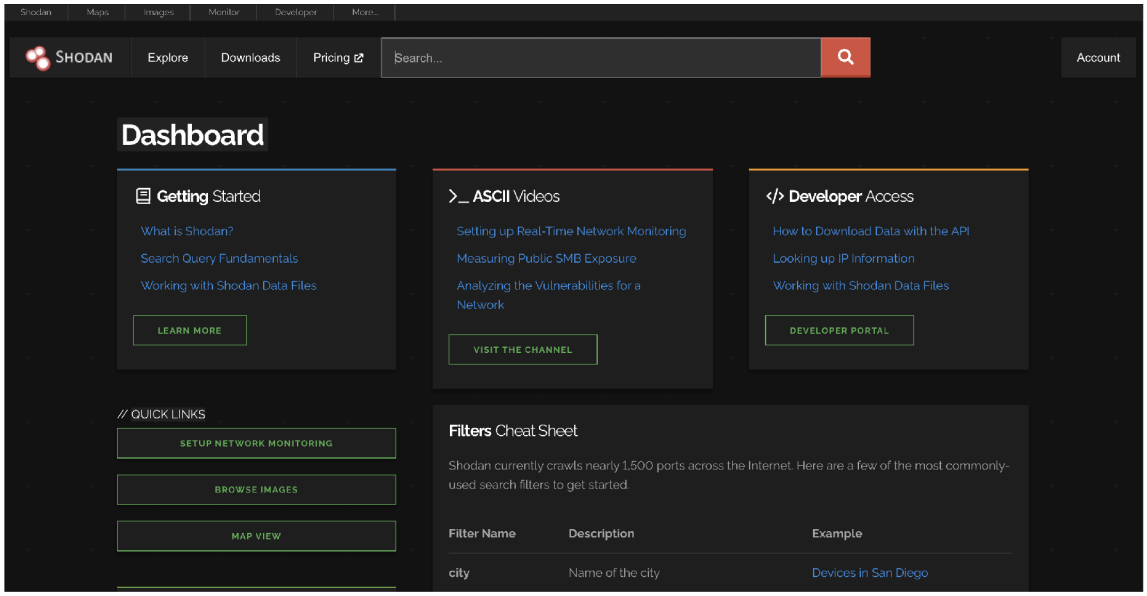

Fig. 2.1: Shodan Homepage

service. For example, on Fig 2.2 on the left side is a classic HTTP banner and on the right Siemens S7 protocol are shown. As can be seen they look differently. In addition to the banner, Shodan also collects meta-data about the device, such as its location, hostname, and operating system [30]. The majority of the meta-data can be searched on the Shodan website, but some fields can be accessed only through the developer API.

```
HTTP/1.1 403 Forbidden
Date: Mon, 07 Nov 2022 22:23:28 GMT
Server: Apache/2.4.38 (Debian)
Content-Length: 276
Content-Type: text/html; charset=iso-8859-1
```

```
Copyright: Original Siemens Equipment
PLC name: S7_Turbine
Module type: CPU 313C
Unknown (129): Boot Loader          A
Module: 6ES7 313-5BG04-0AB0  v.0.2
Basic Firmware: v.3.3.8
Module name: CPU 313C
Serial number of module: S C-D3UM54472013
Plant identification:
Basic Hardware: 6ES7 313-5BG04-...
```

Fig. 2.2: Shodan Banner for HTTP(in the left) and for Siemens S7(in the right)

So how does Shodan work? The way Shodan works can be seen in the Fig 2.3. Shodan functions by initiating connection attempts to every conceivable Internet Protocol (IP) address on the Internet and then indexing the information that is returned in response to those connection attempts [31].

In addition to the traditional Shodan Search, which looks for information about devices, users have access to a variety of different services:

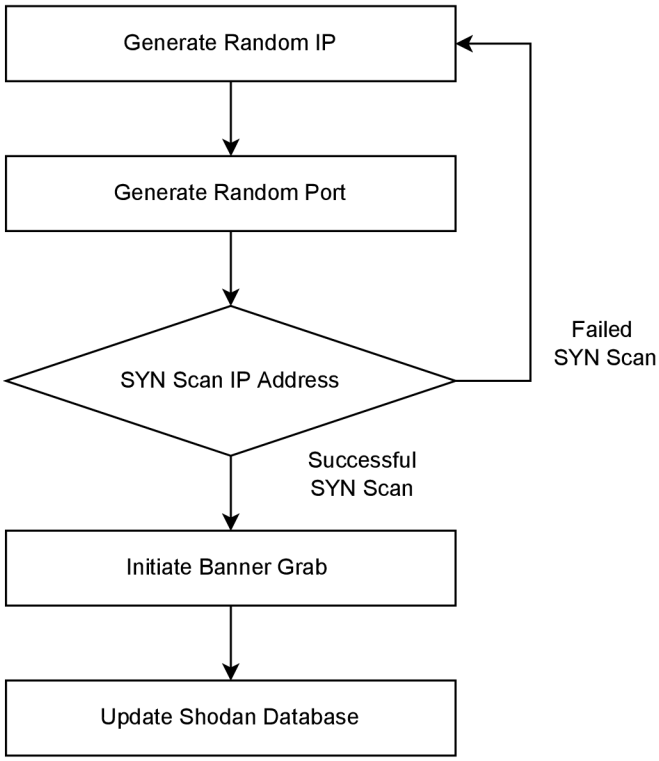

Fig. 2.3: Shodan Scan Method [31]

- Shodan Maps - maps is one of the way to view search results. There is a possibility to zoom in and out to a certain location, filters can also be used.
- Shodan Images - Shodan stores screenshots for many different services.
- Shodan Monitor - helps to keep track of certain devices that could be selected by user's choice. The only requirement is that devices should be connected to the Internet.
- Shodan Developer API - The Shodan application programming interface (API) is the most straightforward method for users to access to the Shodan data. The API grants access to all of the data that is stored in Shodan, enabling to retrieve the specific information that is required.
- Shodan Exploits - a search engine for finding vulnerabilities and exploits from CVE (Common Vulnerabilities and Exposures), ExploitDB and Metasploit [32]

2.2 Censys

Another extremely powerful engine is Censys. Censys was founded by computer scientists as a part of an academic research project at the University of Michigan.

In the same way as Shodan does, Censys keeps an enourmous database of devises exposed on the Internet. Search engine checks approximately 2.1 Billion services daily, scans 107 protocols and more than 3500 ports. Security professionals that use Censys are able to quickly identify weakly protected devices which are open to the Internet [33].

In 2013 a free and open-source security scanner named Zmap was developed. Censys was created due to the large growth of this project after two years of its existence. ZMap can perform a scan of the whole public IPv4 address space on a single computer in less than forty-five minutes when connected to a gigabit network and in only five minutes when a 10 gigabit Ethernet connection is utilized [34]. Censys was published in 2015 with Web User Interface (Fig 2.4.) and is available at Censys website. This search engine, much like Shodan, uses a banner as the fundamental unit for the collection of information; in addition to this information, it also gathers metadata about devices. To access metadata there is a requirement to use filters. The whole list of filters can be found in the bibliography [33].

Fig. 2.4: Censys homepage

How Censys works is shown on the Fig 2.5. Basically Censys operates in the same manner as Shodan does, with the only difference that Censys uses Zmap and ZGrab tools.

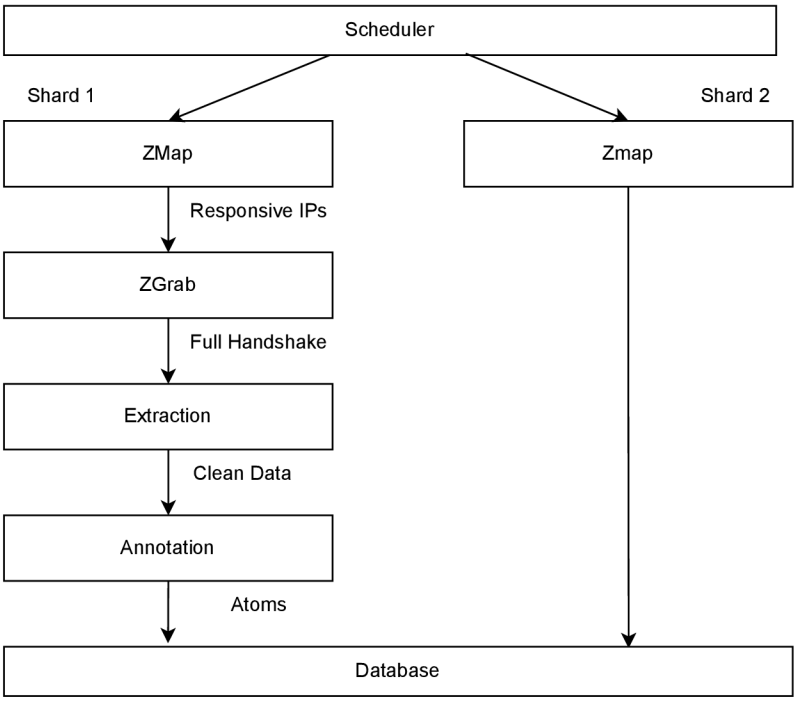

Fig. 2.5: Censys Scan Method [31]

2.3 ZoomEye

The Chinese security company Knownsec Inc. has developed the search engine known as Zoomeye. The first version was released in the year 2013. ZoomEye's fundamental architecture is built on Xmap and Wmap, which are used for data collection from open devices and web services as well as fingerprint analysis [36]. Zoomeye works in the same manner as any other search engine does; to use it, you only need to insert a search for the query. ZoomEye search is available at its website(Fig 2.6) or through API. The search engine supports the detection of multiple port services or protocols for IPv4 and for IPv6 as well [35].

2.4 Nexpose

Rapid7 Nexpose is a vulnerability scanner that intends to handle the entirety of the vulnerability management lifecycle. This includes detection, risk classification, analyses and reporting [37]. Nexpose unlike other search engines can be used only after the registration. There's a free trial for one month which offers to download SW for Windows and Linux only. There is no SW for MacOS users and this is the reason why Nexpose search engine will not be considered for further use nor will be described.

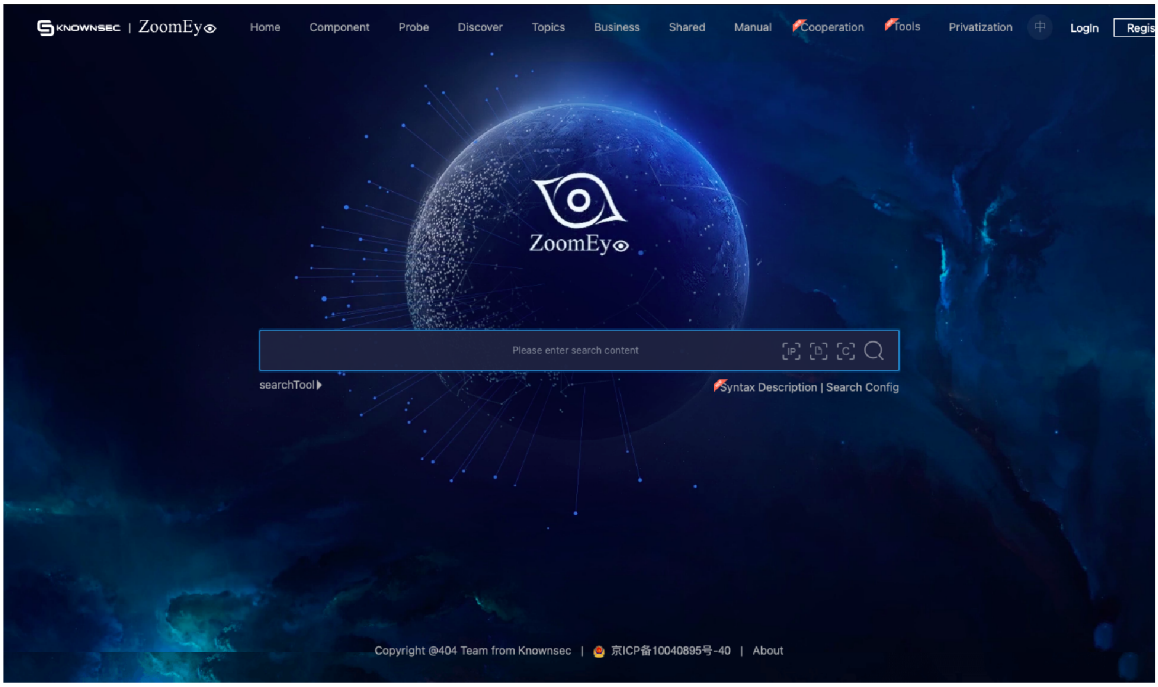

Fig. 2.6: ZoomEye homepage

2.5 BinaryEdge

BinaryEdge is a company with headquarters in Zürich, Switzerland. The company was founded in 2015 and offers many services. This search engine, in contrast to those that were described before, provides additional services to the traditional device search. These services include, for instance, tracking torrent activity and determining whether or not any email address has ever been compromised. Another interesting service provided by BinaryEdge is its own sensors/honeypots summary. The purpose of a honeypot is to demonstrate how attackers operate and investigate the various kinds of threats out there. It's possible to see from which countries and how many incidents engine found for the last month, week or day. BinaryEdge is available at its website (Fig 2.7) or through API as all engines which were mentioned earlier [38].

BinaryEdge continuously collects and analyzes data from internet-accessible devices, which enables enterprises to gain visibility into their attack surface and the vulnerabilities to which they are exposing themselves, namely:

- Ports and Services Exposure
- Misconfigured Network Shares
- Possible Vulnerabilities
- Databases

- Accessible Remote Desktops
- Invalid SSL Certificates [38]

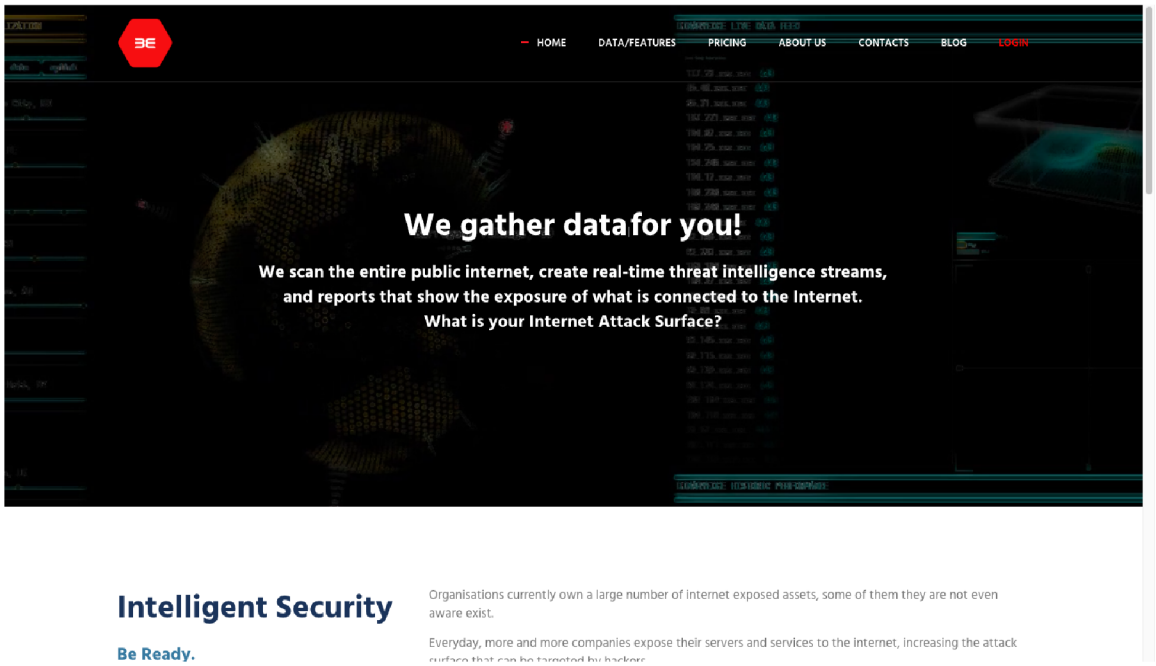

Fig. 2.7: BinaryEdge homepage

It was not possible to find any specific information regarding the method upon which the operation of BinaryEdge is based.

2.6 Protocols

The increasing growth of interconnected devices and the convergence of IT and Operational Technology (OT) have showed that comprehensive network security evaluations are essential. However, this work explores compatibility of industrial protocols like Modbus, DNP3 and S7 with these search engines. The protocols themselves are not the main content of the work and are therefore only described in general terms, but it is important to know how they function and which ports they use in order to carry out practical part.

Modbus

Modbus is the oldest and very popular serial communication protocol. It was published by Modicon (now Schneider Electric) company in 1979. This protocol is very popular not only because it is open source but also because it's royalty-free meaning there is no need to pay subscription or license fees for each use. It widely used

in industry for communication between electronic devices. The Modbus protocol operates at level 7 of the OSI model, which is the application layer and Modbus server device uses port 502 as the default.The transmission of data works on the principle of passing data messages between client and server (master and slave). Typically, the master is a human machine interface (HMI) or Supervisory Control and Data Acquisition (SCADA) system while the slave is a sensor, programmable logic controller (PLC), or programmable automation controller (PAC). In a standard Modbus serial network, there is one master and slaves are limited to 1-255 addresses, all of which have their own individual slave addresses [39, 40]. There are few types of Modbus protocol and the most common used are the following:

- Modbus RTU (Remote Terminal Unit) - The most common implementation that is available for Modbus is known as Modbus RTU. Simply put, Modbus RTU is Modbus that is transmitted over serial communication channels such as RS485, RS422, and RS232.
- Modbus ASCII - The less secured and less used protocol. It's similar to Modbus RTU. The ASCII format uses a checksum that is based on a longitudinal redundancy check.
- Modbus TCP/IP or Modbus TCP - .The new Modbus/TCP Security uses Port 802 while the TCP port for MOdbus is 502. It is not necessary to do a checksum calculation because the layers below it already offer checksum security.
- Modbus Plus (Modbus+, MB+, or MBP) - This protocol differs from others in supporting peer-to-peer communications between multiple clients [41].

Siemens S7

Siemens S7, also known as S7 Communication, is a proprietary communication protocol developed by Siemens for use between programmable logic controllers (PLCs) belonging to the Siemens S7-300/400 family. This protocol was published in 1994 and uses TCP port 102. Programming Programmable Logic Controllers (PLCs), sharing data between PLCs, obtaining PLCs' data from SCADA (supervisory control and data acquisition) and perfoming diagnostics are all examples of uses for this technlogy. The S7 protocol can be executed on either a typical PROFIBUS network or an MPI network [42].

DNP3

Distributed Network Protocol 3 (DNP3) was developed by Harris, Distributed Automation Products in November 1993. It's a collection of communications protocols that are used in process automation systems. Utility businesses, such as those that

provide electricity and water, are its primary users. Utilization in a variety of other fields is not that typical. The DNP3 protocol contains major characteristics that make it more reliable, efficient, and interoperable than previous protocols such as Modbus. These benefits come at the expense of the protocol's higher level of complexity. DNP3 operates at the layer 2 of the OSI model and for communication between devices uses 19999 and 20000 ports [43].

3 Search Engines comparison

There are a few different approaches to take when comparing search engines. In this chapter, search engines will be evaluated by user accounts, which means new accounts will be created for both regular users and students. There is going to be analysis of the following questions: how do search engines differ from one another in using; what are the advantages of using them for students; how much information can be gleaned from each one; what services are provided by each website. All this information is needed in order to understand which search engine is the best for creating application in the future.

Another comparison is based on a more practical method. Certain devices will be connected to the network and according to the information that will be received from each search engine, the best one will be selected. In addition to comparing search engines, at the end of this chapter table will is included which contains different information such as include engines' dates of the occurrence, the tools that engines are based on, which ports and protocols can be used and how many, the possibility to work with API and CLI, whether or not there are benefits for students and how many devices were found in the practical part by each engine **??**.

3.1 Account Comparison

Search engine comparisons of accounts are done for a variety of reasons. Firstly, it's important to find out about capabilities of each one. It's necessary to know how many queries and results these engines allow in order to understand how effective they can be. Secondly, it's good to know what benefits are out there for students or if there is a well written documentation. On top of that it also matters to discuss some limitations or disadvantages of these tools.

3.1.1 Shodan account

Shodan is the oldest and after using a certain time can be considered as the most user friendly engine. Besides having own documentation, Shodan offers videos with tutorials and explanation of how to use it. This engine can be used without the registration but the searching is very limited and filters are not available. Without the registration only one web page is available which means the first 10 results are displayed. To register on Shodan there is a need to create username, password and email. After creating new account with Gmail, Shodan lets using filters and search results are increased to 2 pages meaning the first 20 results are displayed. The big advantage of Shodan is that it gives an ability to use Membership which

normally costs 49 $ but with school email it is for free. Membership can offer many things such as network monitoring, notifications setup, CLI using, searching the images, writing code with Shodan through API. The latter is a huge advantage for this particular work. With network monitoring there is a possibility to monitore 16 IPs and no need to check IPs daily, because notifications can be sent to email, Telegram, or MsTeams, depends what is the best choice for user, as soon as Shodan scans and finds the new information. It's also possible to use all filters except vuln and tag. Searching is limited to 100 scans credits per month, but through CLI there are 65555 scans credits. This engine is definetely the most famous and there is a lot of information out there about Shodan which definitely can be considered as a significant advantage.

3.1.2 ZoomEye account

The search engine can be used only after the registration. ZoomEye as well as Shodan offers video tutorials for using, but unlike Shodan it doesn't have CLI, it's only available through API and its website. Another disadvantage of this search engine is that not everything on its website is translated into English, though the default language is set to English. Even the emails which ZoomEye sends, e.g. when registering, are written in Chinese language which can be very confused. For this reason users cannot understand everything and should translate on their own. To create an account on ZoomEye there is a requirement to have an email. The one distinct thing ZoomEye's website has is that everytime when login there is a need to make a captcha meaning ZoomEye tries to limit spammers and hackers from using it.

ZoomEye doesn't offer any specific benefits or plans for students but anyone can register and be a lifetime member. Lifetime member is free and gives 10000 data every month. However, it's possible to pay 70$ and become an advanced member which allows to have 30000 data per month.

3.1.3 Censys account

Censys' website is very minimalistic and easy to use. The search engine can be used without the registration and unlike Shodan it offers filters for using instantly. Censys doesn't have an extensive documentation but it's possible to find all that is needed on their website.

The biggest disadvantage of this search engine is that Censys offers only one month free trial that includes 250 allowed queries. Censys started as a research project and they claim they'll continue to provide free Internet data to the research community. As a result, they provide a plan for students, but students may only

access it after first receiving information about a project and the reason behind the requirement for their data.

3.1.4 BinaryEdge account

The official website of BinaryEdge provides information indicating that the engine scans 200 ports and monitors 1,000,000 torrents on a monthly basis. BinaryEdge as well as Censys cannot be used without the registration and same as Censys it offers a free account to use. With a free account it's possible to have 250 queries per month with Web and API accesses. Otherwise membership starts with 20 dollars per month which includes another services such as access to historical data, torrents, data leaks, domains and sensors. Unlike other engines BinaryEdge has impressively large and useful API documentation with detailed description and examples.

3.2 Search results evalution

In order to compare how these search engines work in practice and not just in theory, certain devices were connected to the public network from the VUT network. These devices are the following: ABB REC615 (IP - 192.168.64.23), ABB REF615 (IP - 192.168.64.24) and second ABB REF615 (IP - 192.168.64.60). This section will describe how the devices were configured and how they were connected. Comparing the results that will be obtained from each search engines is yet another essential part of this section.

3.2.1 Device connection

At first there was an attempt to find the information about devices only when they were connected inside the VUT network. However, after studying the problem, it turned out that these search engines are unable to search information in private networks. On the Listing 3.1 the result of Shodan is shown about ABB REC615 device. The rest of search engines could not find the information either.

After that there was a need for port forwarding on the router. Different protocols and ports have been chosen for the devices. ABB REC615 communicated by using Modbus protocol which works with port number 502, ABB REF615 with DNP3 protocol through port number 20000, and second ABB REF615 device communicated by using Siemens S7 protocol through port number 102. In the Fig. 3.1 it's possible to see how devices were connected to the Internet. The exact date when ports were configured is 17.11.2022.

Listing 3.1: Failed attemp Shodan

```
(venv) anelsagindykova@Anel-MacBook-Air shodan % shodan
scan submit 192.168.64.23
Error: Private IPs not allowed: 192.164.64.23
```

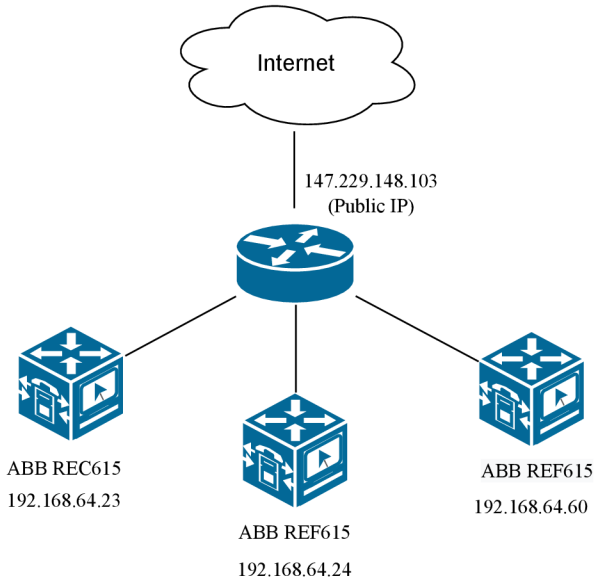

Fig. 3.1: Device Connection

3.2.2 Shodan results

The first engine that has found information about devices was Shodan. Shodan has the advantage that there's no need to track IPs every day because after managing notifications Shodan can send emails when it has found something new. Shodan sent the automatic email after two days since the connection. Also it didn't find all devices, only two of them. Furthermore, the data Shodan provided were sparse compared to the results provided by other search engines which can be seen later on. Shodan has found information about ABB REC615 through 502 (Modbus) and ABB REF615 through 20000 (DNP3). All results Shodan provided are shown on the Fig 3.2. The engine was able to define that this IP is in Czechia, Brno, the name of ISP - Brno University of Technology, domain and hostname.

Surpisingly, after using the Shodan CLI to search for information on devices it did not yield any results (Listing 3.2). The reason why is that is not known.

Listing 3.2: Failed search attemp Shodan

```
(venv) anelsagindykova@Anel-MacBook-Air shodan % shodan
scan submit 147.229.148.103
Starting Shodan scan at 2022-11-25 12:31 - 65524 scan credits
left
No open ports found or the host has been recently crawled
and cant get scanned again so soon.
(venv) anelsagindykova@Anel-MacBook-Air shodan % shodan search
--fields port 'ip:"147.229.220.23"'
Error: No search results found
```

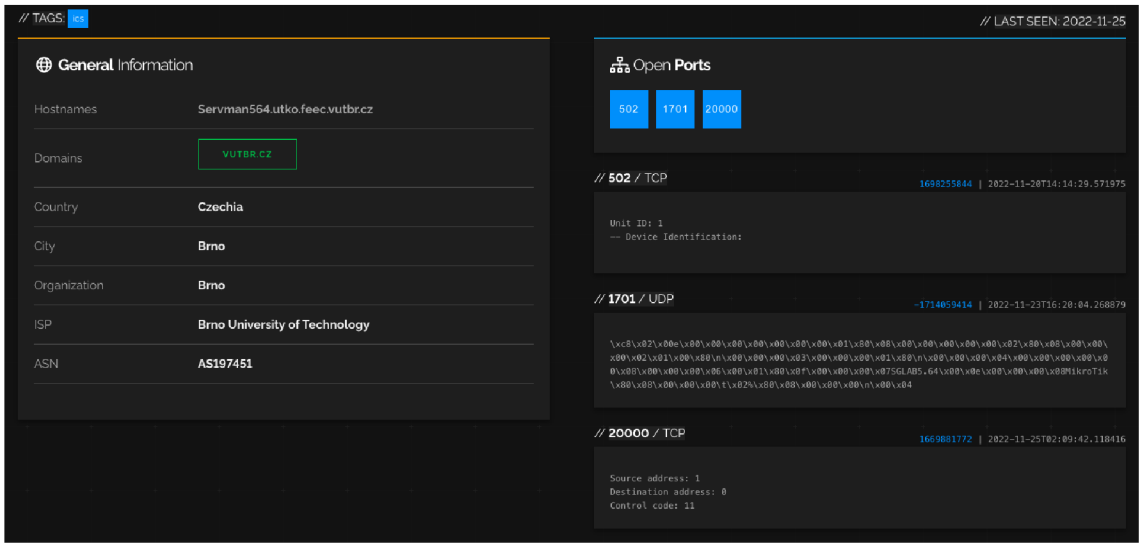

Fig. 3.2: Shodan results

3.2.3 ZoomEye results

ZoomEye was the second engine that has found the devices. Unlike Shodan it has found all three of them after only four days. There is no option to have automatic notifications about specific IP so there was a need to control results manually every day. The information ZoomEye has found can be seen on the Fig 3.3. ZoomEye is the only search engine that found all three devices because it can see that ports 102, 502, 20000 are open. There is a lot of data about the first device. The engine was able to provide the location, system name, module type, serial number. However, there is no so much information about the last two. The information that was provided regarding these two devices is identical to what Shodan provided.

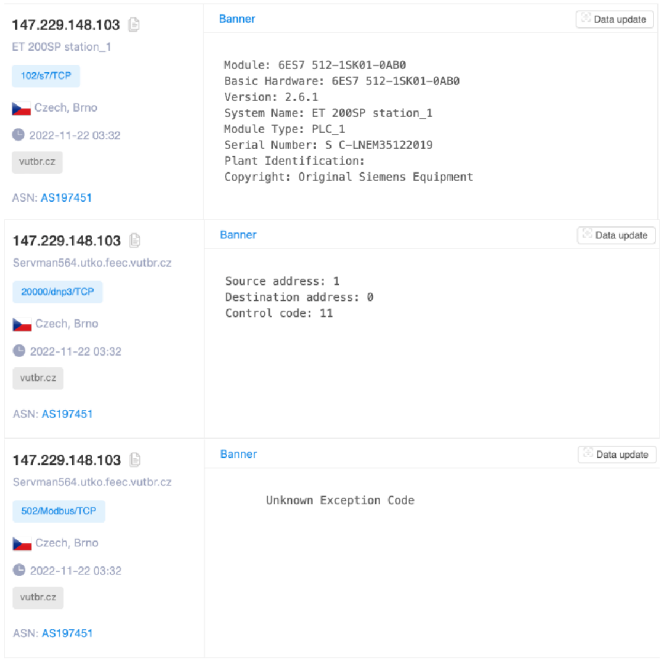

Fig. 3.3: ZoomEye results

3.2.4 Censys results

Censys same as Shodan has found only two devices. It has found two ABB REF615 devices. The first one was through port 102 (Siemens S7) and the second through port 20000 (DNP3). Despite this, the differences in the outcomes about each device are distinct from one another. Regarding the first one, there is a lot of information such as the system name - ET 200SP station_1, module type - PLC, module ID - 6ES7 512-1SK01-0AB0, location - CZ and the rest can be seen on the Fig 3.4. Regarding the second one the is almost no information.

3.2.5 BinaryEdge results

BinaryEdge it took the longest time to find devices, and even then, not all of the gadgets could be found, exactly like with the Censys. BinaryEdge has discovered two devices, namely ABB REC615 and ABB REF615. The most information was discovered regarding the ABB REF615 device, and the data appears to be identical to that which was discovered by the Censys and ZoomEye engines. What exactly BinaryEdge found can be seen on the Fig 3.5.

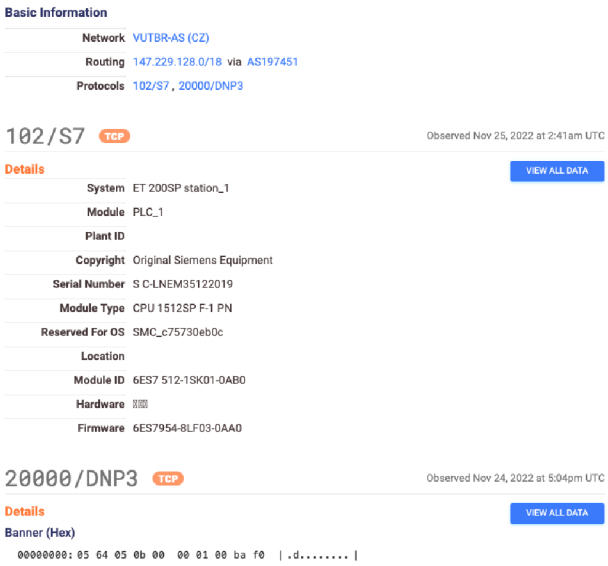

Fig. 3.4: Censys results

Fig. 3.5: BinaryEdge results

3.2.6 Search engine choice

After gathering all of the available data and doing a number of evaluations, the decision was narrowed down to two search engines: Shodan and ZoomEye.

Both BinaryEdge and Censys has found only two devices. ZoomEye as the only one search engine has found all three devices. Shodan was selected as the first suitable engine for this work because it completed searches far more quickly than BinaryEdge and Censys. Moreover, BinaryEdge does not provide any student benefits so it's not free and and it's very limited in allowed data. Other advantages

of using Shodan include the fact that it is the search engine that has been around the longest, and as a result, it has a great community which makes working with this search engine much easier. It is the only engine that has the ability to work through three different ways, specifically via API, CLI, and Web. The fact that ZoomEye was able to identify all three devices was the primary reason that led to choose it as the second engine. Another argument is that it was relatively fast and easy to work with. This engine was capable of providing plenty of information.

For a better overview Tab. 3.1 is added.

Tab. 3.1: Search engine comparison table.

Search engine	**Shodan**	**ZoomEye**	**BinaryEdge**	**Censys**	**Nexpose**
Launching year	2009	2013	2015	2015	2017
Ability to work with IPv6	Yes	Yes	Yes	Yes	Yes
Searching tool	unknown	Xmap, Wmap	unknown	Zmap, ZGrab	unknown
Access to data	Web, CLI, API	Web, API	Web, API	Web, API	its own SW
Number of founded devices	2	3	2	2	wasn't tested
Membership for students	Yes	No	No	Yes	No
Honeypot	No	No	Yes	No	No
Allowed queries per month	100	400	250	250	unknown
Website	[32]	[35]	[38]	[33]	[37]

4 Practical tool development

After deciding upon the most appropriate search engines, one more task associated with this activity is to create a tool. The tool was developed in the form of an web application as web apps are very famous nowadays. The concept and idea of the application will be laid out in detail in the following chapter. The design of the app can be seen on the Fig 4.1.

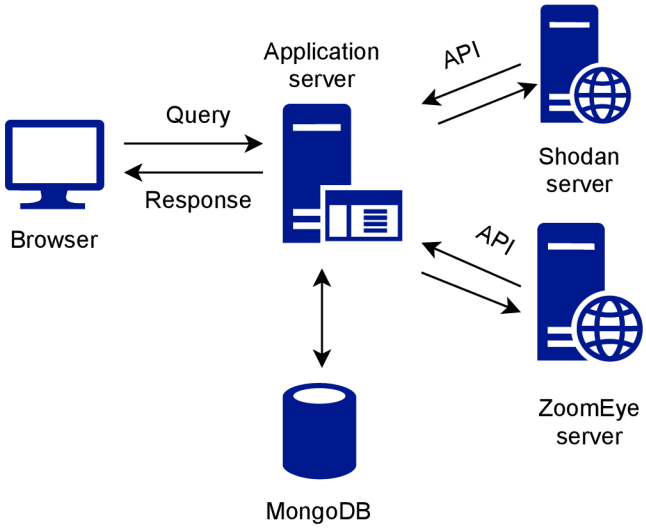

Fig. 4.1: Application Design

The application is developed in the form of a web application. The system applies the client-server model and is divided into two units, namely Backend and Frontend.

4.1 Frontend

The frontend, or often called the client, is the presentation layer of the application that provides the user interface. It is the part that the user interacts with, sees with and manages. Among the most popular technologies for frontend development are HTML, CSS and JavaScript, with which this application is implemented.

4.1.1 JavaScript

JavaScript is a high-level, interpreted language that is executed by web browsers. JavaScript is well-known for its ability to work with HTML and CSS, giving developers the power to dynamically update and change the content of web pages. It

offers a comprehensive selection of built-in APIs and functions that make it possible to do operations like form validation, DOM manipulation, event handling, and AJAX-based asynchronous communication. It's a widely used programming language primarily utilized for web development. For this work it was decided to use JavaScript both for backend and frontend sides.

4.1.2 HTML and CSS

Two of the primary technologies for creating Web pages are HTML (the Hypertext Markup Language) and CSS (Cascading Style Sheets). HTML handles the structure of the page, CSS the (visual and aural) layout, for a variety of devices. Along with graphics and scripting, HTML and CSS together are the basis of building Web pages and Web Applications.

4.2 Backend

The backend, often referred to as the server, is the basis of the application. In the context of web development, the backend typically consists of three main components: the server, the application logic, and the database. It forms a part that is hidden from the user. Its task is to receive and send data for the client through the application interface. This data are processed, stored or modified in the database.

4.2.1 Node.js

Node.js is an open source, cross-platform runtime environment for developing server-side and networking applications. Node.js applications are written in JavaScript, and can be run within the Node.js runtime on OS X, Linux and Microsoft Windows. Additionally, Node.js offers a comprehensive library of different JavaScript modules, which simplifies the development of web applications using Node.js to a great extent [46].

4.2.2 JSON

A JSON (JavaScript Object Notation) file is a simple data exchange format that is easy for both humans and machines to read, write, parse, and generate. Listing 4.1 depicts an example of JSON format. It is frequently used to store and send structured data between servers and clients, and also between various systems.The key-value pairs in a JSON file can be any sort of data, including strings, numbers, booleans, boolean arrays, strings, and other JSON objects. JSON files are often

Listing 4.1: Example of JSON format

```
{"ip": "147.229.148.103",
      "portinfo": {
        "port": 20000,
        "service": "dnp3",
        "title": null,
        "device": "ics",
        }
}
```

used to store API responses which we're going to use in this application because of very large responses from Shodan and ZoomEye.

4.2.3 Express.js

Express.js is a favored option for web development because of its simplicity and flexibility. It provides a simple framework that enables programmers to manage HTTP requests, middleware, and routing with ease. This streamlines and accelerates the development process. Additionally, a sizable ecosystem of libraries and plugins for Express.js provides a wide range of functionality and integration options to improve the creation of web applications [47].

4.2.4 REST API

It's possible to retrieve data from Shodan and ZoomEye by using the REST API. A REST API (also known as RESTful API) is an application programming interface (API or web API) that complies to the limitations of REST architectural style and allows for interaction with RESTful web services. [44]. To execute common database operations like creating, reading, updating, and deleting records (also known as CRUD) within a resource, REST APIs interact via HTTP requests. A POST request is used to create a record, a DELETE request is used to delete one, and a GET request is used to obtain a record from a REST API. Calls to APIs can utilize any HTTP method [45].

4.2.5 Database

In this section, features of the NoSQL database MongoDB and the Mongoose object modeling library will be described.

MongoDB

MongoDB is an open-source, non-relational database based on document-oriented architecture. MongoDB creates collections that consist documents in JSON format rather than using traditional tables and rows like in a relational database. Each document is composed of key-value pairs, and the number of items can vary for each document. The size and content of individual documents can also differ. The structure of a document is similar to the structure that developers create when programming classes and objects [48]. MongoDB was chosen for this work because of its schema flexibility. This flexibility allows to easily modify data structure as application requirements change, making it well-suited for agile development and fast iteration.

Mongoose

Mongoose.js is a Node.js library for object modeling in MongoDB. By transforming database documents into objects, it makes using MongoDB easier. It controls how data relationships are managed, offers typing and schema validation, and is used to represent data models as MongoDB objects. It provides query capabilities for models, enabling data retrieval and database storing. Mongoose uses schemas for modeling and managing data in MongoDB. A schema describes the attributes of individual items that the application works with, such as its data type and whether the item is required or optional [49].

4.3 Application development

This application, which is much like the majority of web applications, has a page for registering and logging in. Either the user can sign up by providing an email address and creating a new password, or can log into an account that already exists. For the logging part bcrypt module was used. With bcrypt user credentials are safely stored. By utilizing a computationally expensive technique and including a different salt for each password, it offers an extra degree of security. As a result, attackers will find it challenging to guess or reverse-engineer passwords.

After the user's authorisation has been processed successfully, he will be taken to the homepage 4.2. The user can directly type their searches on the main page of their respective search engines. Either Shodan or ZoomEye, or perhaps both of them at the same time, are options that can be used. Following the download button of the request, a JSON file will be created and downloaded to the user's machine. This file will contain the results that were received by utilizing the API of the search engines that were chosen. In addition, if a user clicks on the Shodan

or ZoomEye links on the home page, he will be taken to the API reference pages for those respective search engines. It was added so user can find out more about API's and how to use filters.

The user's session terminates if he clicks the Log out button, which will take him back to the registration/login page.

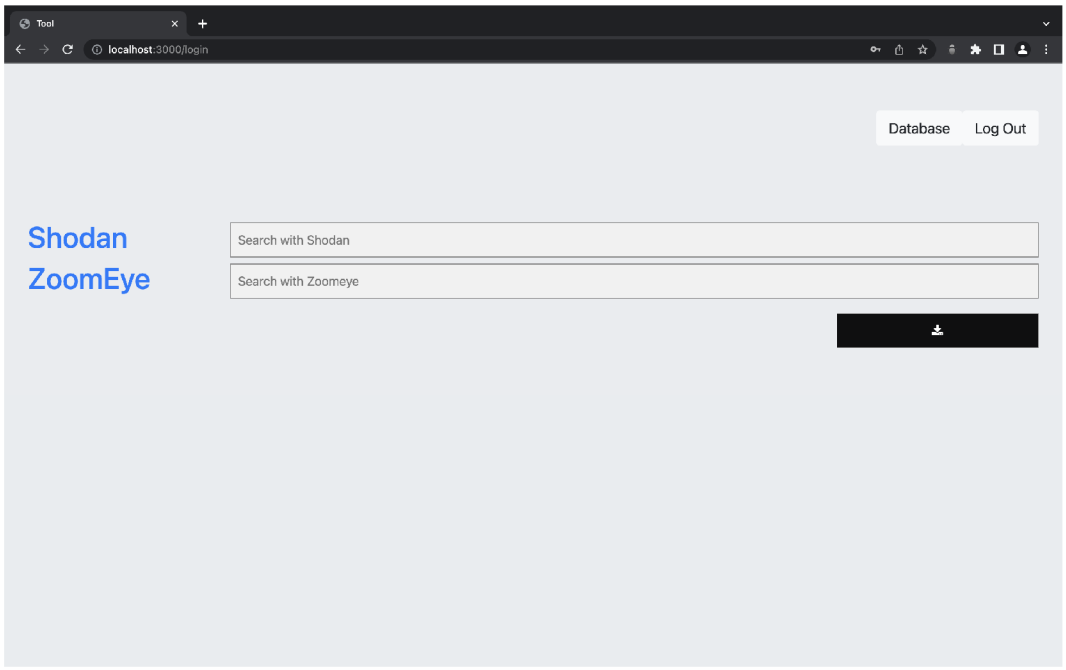

Fig. 4.2: Application main page

For user convenience, a database page was established. On a database page, it's possible to view a number of parameters, such as the name of the file, the query and time when that the file was downloaded. By clicking on a file name, it's possible to download this file once again as shown in Fig 4.3.

As was mentioned earlier, the data for the files and the data for the users are saved in MongoDB as shown in Listing 4.2. The schema for files stores data as following:

- Content - the information that is obtained from search engines.
- Name - is referred to as the name that is formed when the file is downloaded. The file's name is formed with search engines' names and time it was downloaded.
- Time - when file was downloaded.
- Query - what the user submitted in order to search.
- UserID - an ID that is automatically produced by MongoDB. After a user has successfully logged in and downloaded a file, his ID is added to the database

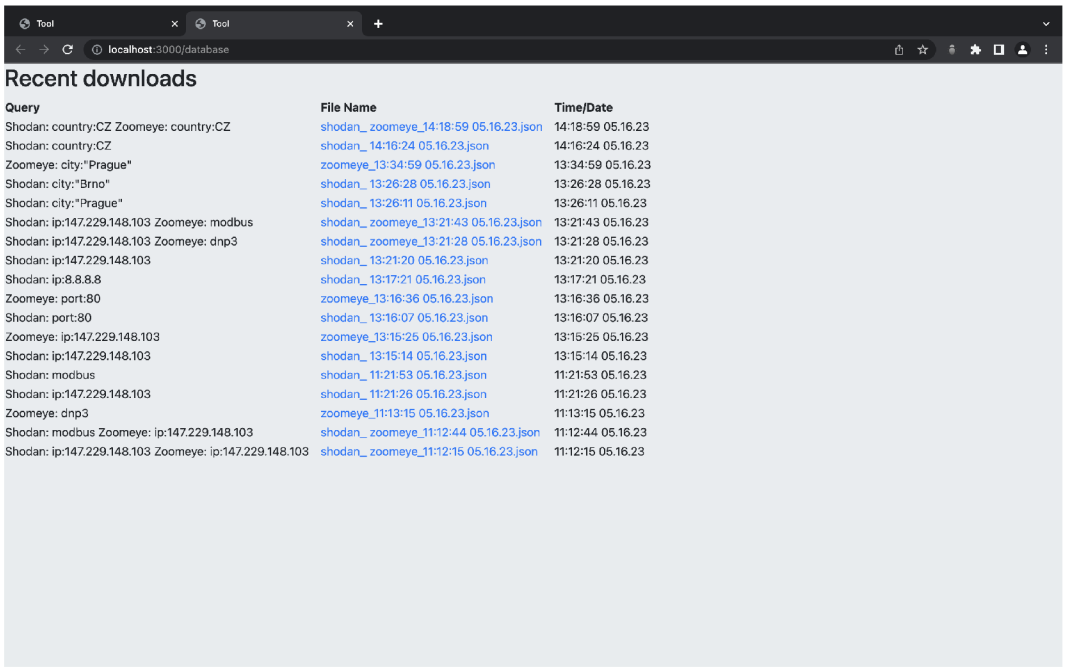

Fig. 4.3: Database page

for that file.

The second schema in this application is a user schema as shown in Listing 4.2. Schema contains just an email and password. The schema is created automatically when new user registers, the real password is not saved, only the hash.

4.4 Application testing

The last goal of this work is to test the newly developed application. Two protocols were chosen for testing, namely Modbus and S7. As was mentioned before, there are three devices that are connected in the school network. Two of them are connected through default ports of these two protocols.

ZoomEye was the first to check. After entering the IP address (147.229.148.103) and port number 102, a JSON file was downloaded. The tool using ZoomEye's API was successful in locating the device as can be seen in the Fig. 4.4. The data provided by ZoomEye about this device are impressive. The search engine was able to System Name, Module Type, Basic Hardware etc.

After the Modbus protocol was tested. Once more, the results were generated after providing an IP address and port number. Again, ZoomEye was able to find the device as is shown in the Fig. 4.5

Unfortunately, Shodan did not have the same level of success as ZoomEye in

Listing 4.2: MongoDB schemas for files and users

```
const fileSchema = new mongoose.Schema({
    content: String,
    name: String,
    time: String,
    query: String,
    userID: String
});

const userSchema = new mongoose.Schema({
    email: String,
    password: String,
});
```

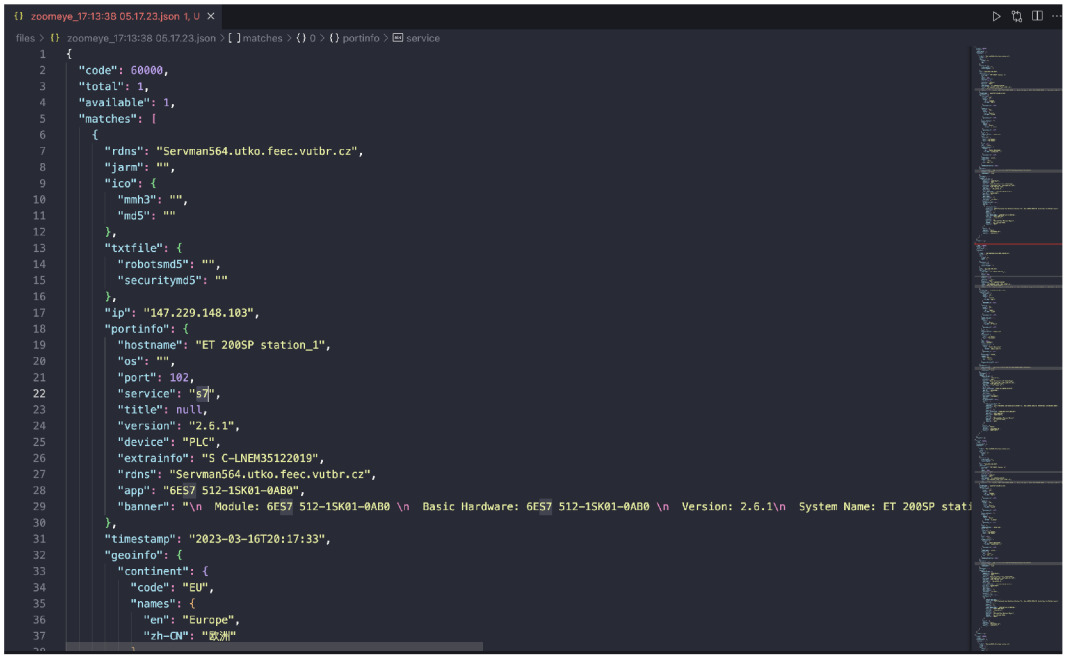

Fig. 4.4: ZoomEye result S7

locating devices. In fact, it was able to only find device connected through 502 port which is the default port of Modbus protocol. The results for Shodan can be seen in the Fig 4.6 and in the Fig 4.7.

Additionally, it's important to note that the inability of Shodan to locate the second device is not due to a problem with the application itself. Even though all three devices had been connected to the Internet for more than half a year, Shodan was unable to locate the other two devices even after the half of the year had passed.

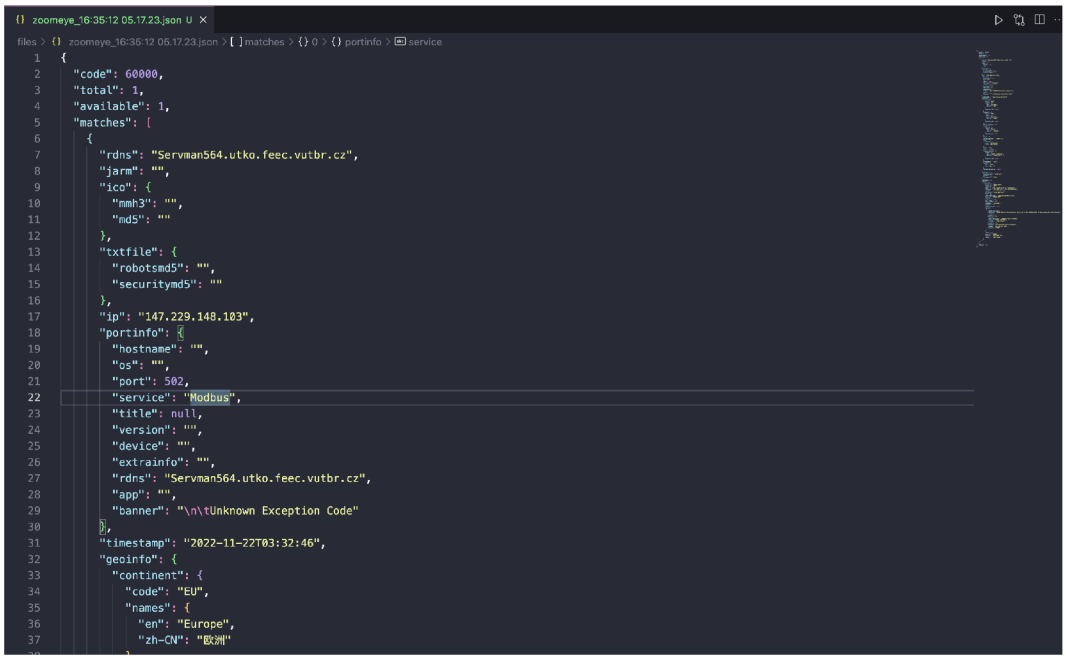

Fig. 4.5: ZoomEye result Modbus

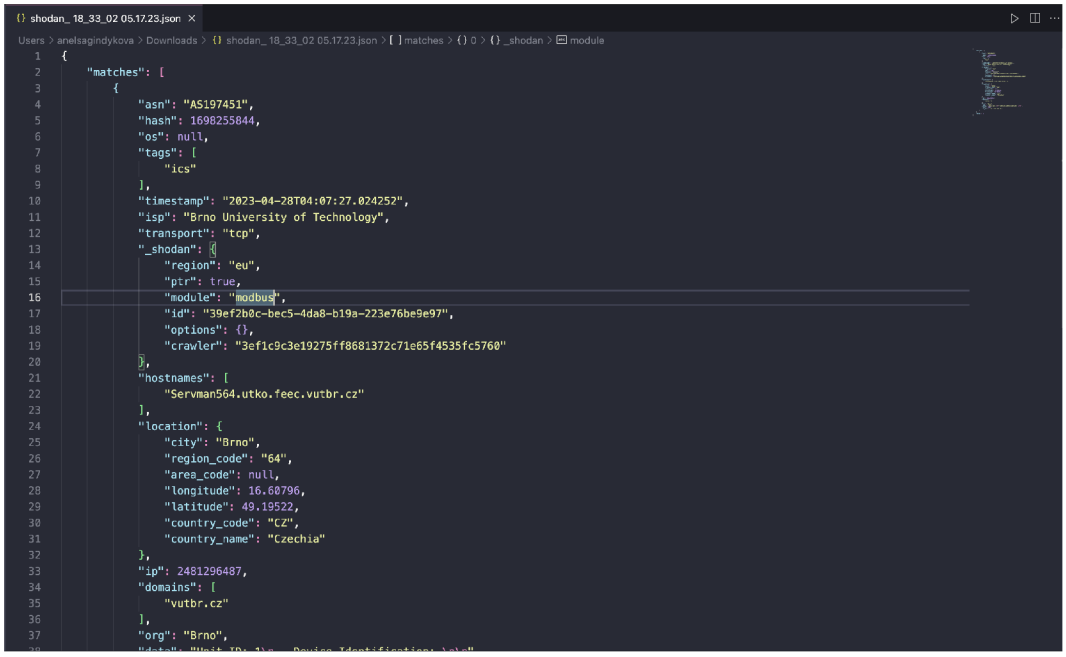

Fig. 4.6: Shodan result Modbus

4.5 Possible ways of concealing information

Keeping information hidden is essential for maintaining security and privacy. It lowers the possibility of being a target of cybercriminals and decreases the potential

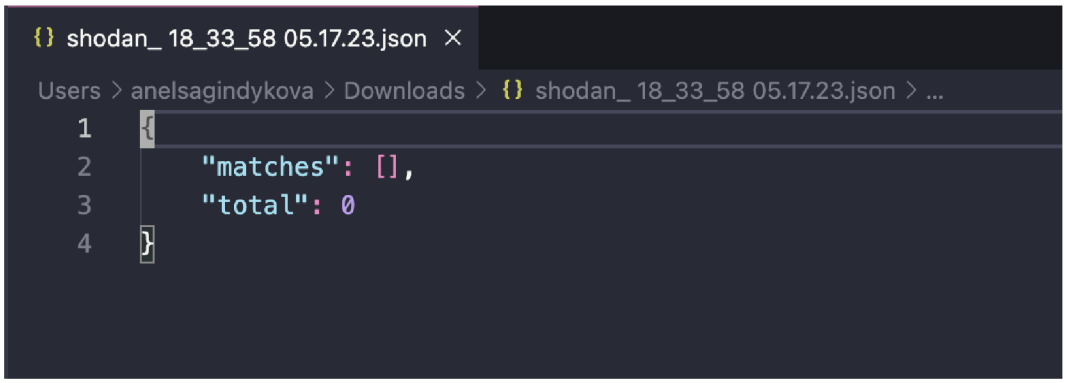

Fig. 4.7: Shodan result S7

damage caused by attacks. Industrial and energy networks should be especially secured because if something happens there can be deadly consequenses for humans and the enviroment. Here are some ways to stay more secure and not be exposed to search engines like Shodan and ZoomEye:

- Disabling unnecessary services: disable or restrict access to any unneeded services or protocols. Disable services that are not necessary for normal operation and close any unused ports. This decreases the attack surface and lowers the possibility of search engines indexing devices.
- Employing Firewall and Access control: Configure firewall rules to restrict external network access to industrial and energy devices
- Applying IP address whitelisting: Set up network equipment such as firewalls or routers to only permit communication from defined trusted IP address ranges. It may prevent illegal access by adding allowed IP addresses to a whitelist.
- Using Virtual Private Networks (VPNs): Implement secure VPN connections to enable remote access to energy and industrial networks. Search engines or unauthorized parties will have more difficulty collecting or accessing the network traffic.
- Employing Intrusion Detection and Prevention Systems (IDPS): Implement IDPS solutions that monitor network traffic and identify potential attempts at unauthorized access or malicious actions. IDPS can aid in detecting and preventing attempts by search engines to scan, providing an additional layer of defense.

Conclusion

One of the primary goals was to analyze and evaluate various search engines. Shodan, Censys, BinaryEdge, Nexpose, and ZoomEye are search tools that were described in the theoretical section. These engines can find information about industrial devices that are connected to a public network. Additionally, famous industrial protocols were described which will later be used in the practical section.

Another part of the theoretical part includes description of energy and industrial control systems, also including security and components of there networks. Moreover, some famous attacks on there networks were described and how they function.

A further component of the theoretical aspect involves the description of energy and industrial control systems, as well as the security features and components of these systems. In addition to that, several well-known attacks on networks and the mechanism behind them were discussed.

The practical part of the work was divided into three sections. The first section is devoted to comparison based on the accessible documentation and information that can be found on websites associated with search engines. It was concluded that the engines which will be later compared in the practical part are Shodan, BinaryEdge, Censys and ZoomEye. Nexpose wasn't considered for further use because it doesn't offer its SW for MacOS users.

The second phase of the comparison was to determine how much data different search engines are capable of collecting in a given amount of time. Shodan was the one that was able to find devices the quickest, while ZoomEye was the only one that has found all the devices connected in the VUT network.

After conducting all of the research and consideration, the decision was made to use two search engines, namely Shodan and ZoomEye. Shodan was chosen because of its ability to create a privileged academic account using a school email address and ZoomEye because of its ability to find all devices. Shodan is also capable to deal with both IP4 and IP6, which is another reason to select it as a tool of choice. In addition to this, it is the only search engine that can work in three different ways, namely through the API, CLI and Web interface. ZoomEye as the second choice was able to find all devices, was fast and relatively easy to implement in the application.

The subsequent goal was to create a tool that would store data obtained through the usage of search engines to a database. As a result, a web application was created that is able to take information from Shodan and ZoomEye search engines using their APIs. The web application was implemented in JavaScript language using Node.js and its famous web application framework Express.js. The application is able to retrieve information from search engines and is able to save data to a database using MongoDB. To be able use the application, it is necessary to either sign up for

an account or sign in with an account that has already created.

As the last part the web application was tested. To perform testing two protocols were chosen: Modbus and S7. The application was able to retrieve data from both search engines in a form of a JSON file. In finding devices the extracted data were the same as when using search engines' websites.

It would be appropriate to connect more devices in the laboratory, run more tests and find out why search engines could not find some of the devices connected to the Internet. As for the web application, more search engines could be added to collect more information. The authorization could be improved, for instance by asking users to create more secure passwords and require users to validate an email by having it redirected to a specific email service that the user has given.

Bibliography

[1] B. Babu, T. Ijyas, Muneer P. and J. Varghese, *Security issues in SCADA based industrial control systems.* 2017 2nd International Conference on Anti-Cyber Crimes (ICACC), Abha, Saudi Arabia, 2017, pp. 47-51, doi: 10.1109/Anti-Cybercrime.2017.7905261.

[2] M. Krotofil and D. Gollmann, *Industrial control systems security: What is happening?,.* 2013 11th IEEE International Conference on Industrial Informatics (INDIN), Bochum, Germany, 2013, pp. 670-675, doi: 10.1109/INDIN.2013.6622964.

[3] STOUFFER, Keith; FALCO, Joe; SCARFONE, Karen. *Guide to industrial control systems (ICS) security* NIST special publication, 2011, 800.82: 16-16

[4] Trendmicro. [online] *Industrial Control System* [cit.2023-04-05] `https://www.trendmicro.com/vinfo/us/security/definition/industrial-control-system`

[5] Security Boulevard *Complete Guide to SCADA Security.* [online] [cit.2023-04-05] `https://securityboulevard.com/2022/09/complete-guide-to-scada-security/`

[6] Dr. James Stanger *Why DDoS attacks are a major threat to industrial control systems.* [online][cit.2023-04-05] `https://www.controleng.com/articles/why-ddos-attacks-are-a-major-threat-to-industrial-control-systems/`

[7] Electronics Coach *SCADA System Architecture.* [online][cit.2023-02-03] `https://electronicscoach.com/scada-system-architecture.html`

[8] Alexander S. Gillis: distributed control system (DCS) [online][cit. 2023-02-03] `https://www.techtarget.com/whatis/definition/distributed-control-system`

[9] Mallikarjun G. Hudedmani,Umayal R M,ShivaKumar Kabberalli,Raghavendra Hittalamani *Programmable Logic Controller (PLC) in Automation* Volume 2, Issue 1, pp. 37-45, July2017 doi:https://doi.org/10.21467/ajgr.2.1.37-45

[10] G. P. H. Sandaruwan, P. S. Ranaweera and V. A. Oleshchuk, *PLC security and critical infrastructure protection,* 2013 IEEE 8th International Conference on Industrial and Information Systems, Peradeniya, Sri Lanka, 2013, pp. 81-85, doi: 10.1109/ICIInfS.2013.6731959.

[11] W. N. S. E. Wan Jusoh, M. A. Mat Hanafiah, M. R. A. Ghani and S. H. Raman, *Remote terminal unit (RTU) hardware design and implementation efficient in different application,* 2013 IEEE 7th International Power Engineering and Optimization Conference (PEOCO), Langkawi, Malaysia, 2013, pp. 570-573, doi: 10.1109/PEOCO.2013.6564612

[12] Who we are | Energy Networks Australia. (n.d.). Energy Networks Australia. URL: https://www.energynetworks.com.au/about/

[13] Electrical grid - Energy Education. (n.d.). Electrical Grid - Energy Education. [online][cit.2023-04-06] URL: https://energyeducation.ca/encyclopedia/Electrical_grid

[14] Gas Distribution Systems: A Beginner's Guide | Dombor Valve. (2023, February 7). Dombor. [online][cit.2023-04-06] URL: https://www.dombor.com/gas-distribution-systems-a-beginners-guide/

[15] District heating - Wikipedia. (2013, May 1). District Heating - Wikipedia. [online][cit.2023-04-05] URL: https://en.wikipedia.org/wiki/District_heating

[16] A. (n.d.). District heating and cooling: advantages of an efficient system. District Heating and Cooling: Advantages of an Efficient System. [online][cit.2023-04-05] URL: https://www.araner.com/blog/district-heating-cooling-advantages

[17] Water supply system | Description, Purification, Distribution, & Water Quality. (n.d.). Encyclopedia Britannica. [online][cit.2023-06-05] URL: https://www.britannica.com/technology/water-supply-system

[18] Pipeline - Energy Education. (n.d.). Pipeline - Energy Education. [online][cit.2023-06-05] URL: https://energyeducation.ca/encyclopedia/Pipeline

[19] Johnson, Hoaglund, Trevizan, & Nguyen. (2020). Chapter 18: Physical Security and Cybersecurity of Energy Storage Systems. ResearchGate.URL:https://www.researchgate.net/publication/348785890_Chapter_18_Physical_Security_and_Cybersecurity_of_Energy_Storage_Systems

[20] Y, R. (2020, November 3). *What are Intelligent Electronic Devices (IED)?* Block Diagram, Hardware and Software Design and Communication Module of IED - Electronics Coach. Electronics Coach. URL: https://electronicscoach.com/intelligent-electronic-devices.html

[21] *Triton is the world's most murderous malware, and it's spreading.* (2019, March 5). MIT Technology Review. [online][cit. 2023-02-03] URL: https://www.technologyreview.com/2019/03/05/103328/cybersecurity-critical-infrastructure-triton-malware/

[22] *NotPetya: the cyberattack that shook the world.* (n.d.). The Economic Times. [online][cit. 2023-03-04] URL: https://economictimes.indiatimes.com/tech/newsletters/ettech-unwrapped/notpetya-the-cyberattack-that-shook-the-world/articleshow/89997076.cms

[23] *'NotPetya': Latest Ransomware is a Warning Note From the Future.* (2017, June 30). IEEE Spectrum. [online][cit. 2023-03-04] URL: https://spectrum.ieee.org/notpetya-latest-ransomware-is-a-warning-note-from-the-future

[24] Hemsley, Kevin E., and E. Fisher, Dr. Ronald. *History of Industrial Control System Cyber Incidents.* United States: N. p., 2018. Web. doi:10.2172/1505628.

[25] T. Alrubaie, W. Elmedany, N. Ababneh, S. Zeadally and K. Curran, "A Cybersecurity Architecture to Mitigate Shamoon Attacks," 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakheer, Bahrain, 2022, pp. 266-277, doi: 10.1109/3ICT56508.2022.9990865.

[26] Pagliery, J. (n.d.). The inside story of the biggest hack in history. CNN-Money. [online][cit. 2023-02-03] URL: https://money.cnn.com/2015/08/05/technology/aramco-hack/index.html

[27] Jelen, S., Top 9 Internet Search Engines Used by Security Researchers. Security Trails [online]. 2022 [cit. 2022-11-15]. [online][cit. 2023-02-03] URL: https://securitytrails.com/blog/hacker-search-engines

[28] *Shodan* [online]. 2016 [cit. 2022-11-12]. URL: https://www.techtarget.com/whatis/definition/Shodan

[29] CHEN, Yongle, Xiaowei LIAN, Dan YU, Shichao LV, Shaochen HAO a Yao MA. *Exploring Shodan From the Perspective of Industrial Control Systems.* IEEE Access. 2020,[cit. 2022-11-15]. 8, 75359-75369. ISSN 2169-3536. Available at: doi:10.1109/ACCESS.2020.2988691

[30] Matherly, J.:*Complete Guide to Shodan.* Austin(Texas): Leanpub, 2016, [cit. 2022-11-25]. URL: https://leanpub.com/shodan

[31] Im, S., Shin, S., Roh, B., Lee, J. (2017, April 30). *Scan Modeling and Performance Analysis for Extensive Terminal Information Identification.* The Journal of Korean Institute of Communications and Information Sciences. Korea Information and Communications Society. URL: https://doi.org/10.7840/kics.2017.42.4.785

[32] Shodan engine. [online], [cit. 2022-12-02]. URL https://www.shodan.io

[33] Censys engine. [online], [cit. 2022-12-02]. URL https://censys.io

[34] Durumeric, Z., Wustrow, E., Halderman, J.A., (August 2013). *ZMap: Fast Internet-Wide Scanning and its Security Applications.* Proceedings of the 22nd USENIX Security Symposium.[online], [cit. 2022-12-02]. URL https://zmap.io

[35] ZoomEye engine. [online], [cit. 2022-12-02]. URL https://www.zoomeye.org

[36] Salame, W., Zoomeye hacker search engine.(September 27, 2020), [online], [cit. 2022-12-05]. URL https://kalitut.com/zoomeye-search-engine/

[37] Rapid7. [online], [cit. 2022-12-02]. URL https://www.rapid7.com/products/nexpose/

[38] BinaryEdge engine. [online], [cit. 2022-12-02]. URL https://www.binaryedge.io

[39] IT Explained: Modbus. [online], [cit. 2022-12-02]. URL https://www.paessler.com/it-explained/modbus

[40] Modbus home page. Modbus. Modbus Organization, Inc. Retrieved 2 August 2013 [online], [cit. 2022-12-02]. URL https://modbus.org

[41] Modbus. In *Wikipedia.* 3 December 2022, [online], [cit. 2022-12-21]. https://en.wikipedia.org/wiki/Modbus

[42] Siemens PROFIBUS/MPI S7 Protocol & Siemens Industrial Ethernet S7 protocol ID:8. Schneider Electric 2022 [cit. 2022-12-02]. URL https://igss.schneider-electric.com/plc-scada-driver-8/

[43] Overview of DNP3 Protocol. [online], ©2022, [cit. 2022-12-12]. URL https://www.dnp.org/About/Overview-of-DNP3-Protocol

[44] What is a REST API? © 2022 Red Hat, Inc. 2020, [online], [cit. 2022-12-25], URL: https://www.redhat.com/en/topics/api/what-is-a-rest-api

[45] What is a REST API? | IBM. (n.d.). What Is a REST API? | IBM. [online], [cit. 2022-12-25]URL: https://www.ibm.com/topics/rest-apis

[46] Node.js - Introduction. (n.d.). Node.js - Introduction. [online], [cit. 2023-5-5] URL: https://www.tutorialspoint.com/nodejs/nodejs_introduction.htm

[47] Express - Node.js web application framework. (n.d.). Express - Node.js Web Application Framework. [online], [cit. 2023-5-5] URL: https://expressjs.com

[48] Why Use MongoDB And When To Use It? (n.d.). MongoDB. [online], [cit. 2023-5-5] URL: https://www.mongodb.com/why-use-mongodb

[49] Mongoose ODM v7.2.1. (n.d.). Mongoose ODM v7.2.1. [online],[cit. 2023-5-5] URL: https://mongoosejs.com

Symbols and abbreviations

API	Application programming interface
CLI	Command Line Interface
CVE	Common Vulnerabilities and Exposures
SCADA	Supervisory Control and Data Acquisition
HTTP	Hypertext Transfer Protocol
RTU	Remote Terminal Unit
TCP	Transmission Control Protocol
DNP3	Distributed Network Protocol 3
PLC	Programmable logic controller
REST	Representational State Transfer
IP	Internet Protocol
ICS	Industrial Control System
IED	Intelligent Electronic Devices
HMI	Human-Machine Interface
GUI	Graphical User Interface
CPU	Central Processing Unit
DCS	Distributed Control Systems
PCS	Process Control Systems