Sapienza
Università di Roma

Data Collection, Storage and Processing for Water Monitoring based on IoT and Blockchain Technologies

Facoltà di Ingegneria dell' Informazione, Informatica e Statistica
Corso di Laurea Magistrale in Engineering in Computer Science

Candidate
Marco Zecchini
ID number 1596071

Thesis Advisor
Prof. Ioannis Chatzigiannakis

Co-Advisor
Prof. Andrea Vitaletti

Academic Year 2017/2018

Data Collection, Storage and Processing for Water Monitoring based on IoT and Blockchain Technologies

Master thesis. Sapienza – University of Rome

This thesis has been typeset by LATEX and the Sapthesis class.

Author's email: zecchini.1596071@studenti.uniroma1.it

Acknowledgments

I would like to express my sincere thanks to Professor Ioannis Chatzigiannakis and Professor Andrea Vitaletti for having guided me with experience and competence in the realization of this work. I hope that our collaboration can continue in the future. I thank my parents, Fabrizio and Carmela, for their constant support during these years of study, and my sister, Chiara, for having dedicated the right attention to my work, taking care of its form. Finally, I would like to thank Agnese and Lorenzo, to whom I hope to have been able to explain what a blockchain is, for their critical ideas and, Paolo and Aurelia, because with interest they have always been present in these months of research.

Abstract

Wastewater might represent an interesting resource in terms of water reuse and preservation of the environment. However, this is not always well purified, resulting instead in a source of pollution with both environmental and economic consequences. The following thesis proposes a new approach to achieve a form of monitoring based on IoT technologies and addresses a careful analysis in the acquisition, transmission and storage of data and, in the latter phase, reflecting on when it is appropriate, depending on who collects the data, to use a system based on blockchain or centralized server.

Finally, on the basis of the data collected, a system of taxation is supported if the quality thresholds of the water discharged into the environment are not respected in a way similar to the carbon tax. On the top of this, the thesis theorizes a "quality-credit" trading scheme also modeled in game theory in a Stackelberg scheme.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The widespread problem of water pollution is affecting our health and our environment. Unsafe water kills more people each year than war and all the other forms of violence combined. Moreover, drinkable water sources are finite: less than 1 percent of earth's freshwater is accessible to us[2]. Over the future years, without actions, the situation will get worse when global demand for freshwater is expected to be one-third higher than it is now. Used waters, also known as wastewater, are part of this scenario. Source of wastewater includes homes, farms, hospitals and businesses. Sewers, usually, collect both sewage and stormwater runoff from streets, lawns, farms, and other land areas. So wastewater can include any debris from roads and waste oils, pesticides, fertilizers, and wastes from humans and animals. Much of wastewater, treated or untreated, eventually ends up in rivers, lakes and oceans; when the untreated one reaches water used as a drinking water source for the community there can be a significant health risk[1]. Besides, an interesting opportunity would be to recycle water used more than once: treating industries or sewers wastewater can be an opportunity to reuse water for agriculture, irrigation, bath drains, replenish a groundwater system and even drinking water. This means that there is a massive opportunity for using recycled water too.

Although it is a massive scale problem, it is not properly considered by the State authorities. For example, in a country like Italy, a peninsula in the Mediterranean Sea that is part of the EU, it is registered a very negative data: according to a recent survey by "Legambiente"[3], an Italian non-profit organization for the environment protection, the 48%[4] of Italian coasts and lakes are polluted; more precisely the 39% of the monitored waters are highly contaminated, the 9% are polluted. The classification was realized according to the results obtained by microbiological analysis related to the number of fecal bacteria in the sampled water. Legambiente's report covered the entire expanse of Italy's coastline focusing also on the so-called "abandoned waters", meaning 170-kilometers of coast where 556 watercourses run into the sea but are not regularly monitored by Italy's Ministry of Health. Also, the same report states that out of 149 monitored river mouths, 106 (71%) are polluted. It could be most probably related to the fact that, in Italy, most of the big cities rise along the rivers. The figure below summarizes the situation described upon here.

The green pins refer to the monitored waters that respect the quality limits, while yellow pins represent those that are polluted and red those that are highly polluted. The picture above shows that half of the Italian coasts are dirty.

Wastewater pollution also affects the U.S., for example, where 3.5 million Americans get sick each year after swimming, boating, fishing, or otherwise touching water they thought was safe and where each year, more than 860 billion gallons of used water escapes sewer system across the country.[5] Water pollution represents an issue

Figure 1.1. Water quality of Italian coasts and lakes

not only for the health but also for the financial resource of countries:

- Italy was fined to pay 25 million euros plus 30 million euros per semester by European Union[6] due to the situation described above and for its deficiency in appropriate sewage systems. To solve the whole situation at least half a billion euros should be spent.
- Freshwater pollution costs to the U.S. 4.3 billion dollars per year. [7]

The inadequacy of good infrastructures, which are very often lacking, is one of the primary cause of this situation: wastewater may exceed the capacity of the treatment plant due to stronger rainfall and therefore this water is directly discharged into rivers or seas. Moreover, some purifiers do not carry out sufficient purification treatments. [4] But, it is important to flag that over the 80% of wastewater around the world returns to the environment without any treatment.[8]

The only resolutive mechanisms adopted by the world community concern mainly to regulatory policies and awareness. In the "World Water Development Report", UN requests an extra effort in data monitoring systems because wastewater collection and treatment data are sparse, particularly in developing countries. Only 55 out of 181 analyzed countries have reliable statistical information on generation, treatment, and use of wastewater, 69 countries have data on one or two aspects, and 57 countries have no information at all. Moreover, data from approximately two thirds (63%) of the countries were over five years old. [8]

Some actors may be involved to address the situation described above: the purpose of this thesis is to provide an analysis of how these different actors are involved in the collection of sensitive data, how they are involved in storing it so this data remains safe and non-corruptible. This thesis has also the aim of supporting a concrete proposal, based on an approach similar to carbon taxation, for data processing and analysis. Finally, this paper will indicate the future work that will be pursued to propose a possible resolutive approach to this issue.

In Chapter 2 the actors addressed before are briefly introduced as well as data collection possibilities, considering both their benefits and their drawbacks, are presented.

In Chapter 3 storage issues are faced and which data structures should be used depending on the considered actor in order to ensure integrity, no repudiability and, when needed, confidentiality.

In Chapter 4 it is described and supported an approach, similar to the carbon tax and its emission market, to increase the average of the quality of wastewater returned to the environment over the time.

Chapter 2

Data collection analysis

Monitoring provides factual information about water quality and can confirm that licence conditions are being met. The information obtained through monitoring provides the basis for making water quality decisions. The regulations of the permitted quality of water are enforced by the state authorities, which have the task of viewing the correct monitoring of treated and untreated wastewater and possibly fining those who do not meet the permitted parameters. The monitoring process can be managed in a variety of ways: it can be carried out directly by the state; it can be carried out by industries that discharge directly into watercourses and by those that deal with the treatment of urban, agricultural and, if necessary, industrial wastewater. However, monitoring activities done ordinary people should be also taken into consideration. Indeed, those who have sufficient computer and electronic skills, applying a *Do-It-Yourself* approach, can create prototypes to monitor the quality of wastewater. These people are called as *Makers*.

In the following sections will be presented the different actions of each *actor* mentioned above and presented the benefits and problems for each of them in the monitoring action.

2.1 The role of Central Government in data collection

Before proceeding with the analysis of the government or state as the main actor in the monitoring process, it is right to set the motivation for this type of action: trivially, since the government and the state authorities are elected by the citizens, it is reasonable to think that they, acting in the name of people, care about the environment: so, this represents the main reason that drives it.

The management and installation of the monitoring systems are carried out by an authoritative body that protects the environment by ensuring that all updated data is collected for all points deemed important.

It is reasonable to think that the design of the system and the quality of the sensors make the data accurate and precise since the resources available, financial and staff ones, allow this. The power supply of the monitoring stations is not even a problem: the powerful state resources allow to create stations that do not have this issue, designing solutions that are directly linked to the electric grid. The same goes for data transmission that can use reliable and efficient mechanisms based on Ethernet and 3G/4G.

However, it has a great limit: it is a system in which not everyone can participate, where only the points considered important are monitored, leaving room for the possibility of the creation of “abandoned waters“, as for Italy as described in Chapter

	Benefit	Problem
Central Govern	Good Accuracy	Not accessible to everybody
	Reliable transmission mechanisms	Abandoned water
	Data for all important points (purifiers,...)	If State is corrupted, whole system too

Table 2.1. Benefits and problems of using Government as actor

1. Finally, it is an approach that is based on a full trust in the government: if it is corrupted the whole system will suffer.

The above table summarizes the benefits and drawbacks of this approach.

2.2 The role of Purifier Companies in data collection

The motivation that drives the companies of the purifiers or the industrial ones, that discharge directly in the watercourses, to monitor the quality of the water poured in the environment differs from that already introduced for the Central Government: these actors would participate in this system either because they are obliged to collect information from a third entity or because spontaneously they have an interest in showing their good conduct.

Obviously, the third party wants to create a standard that everyone must follow both in the design of monitoring stations and in the positioning of them. The case of "Online Source Water Quality Monitoring"[9] is analyzed as an example, although concerning the source water, it is perfectly suitable to the case of wastewater. It defines, using the same basic functional block diagram shown below, how a source water monitoring station is realized that defines the facilities as follows:

- *Instrumentation*. Providing the means to measure the selected parameters.
- *Sampling*. Placing the sensors in contact with the source water and, as necessary, disposing of the waste stream.
- *Power Supply and Distribution*. Supplying sufficient power to the energized equipment in the station.
- *Communications*. Providing the means to transfer the data collected by the station to a control center and transfer instructions from the control center to the station.
- *Packaging*. Providing a structure for mounting and protecting the instrumentation and ancillary equipment both from the environment and potential tampering.

As the case of government, having big financial and technological resources the accuracy and the precision of the samples represent a benefit for this approach like the reliability and efficiency in data transmission. A power supply can be achieved easily by the companies since they have direct access to the electrical grid.

The same report then shows a way to choose the location where to place the stations: this should be chosen with a strategical position close to potentially threatening points in a way similar to the (2.2) figure.

Leaving it up to everyone to participate and adding a monitoring station is once again complicated: this could recreate the problem of "abandoned waters". Moreover, it can be easier for a company to corrupt the data collected by the station having both physical access and the chance to move the probes and tamper the firmware.

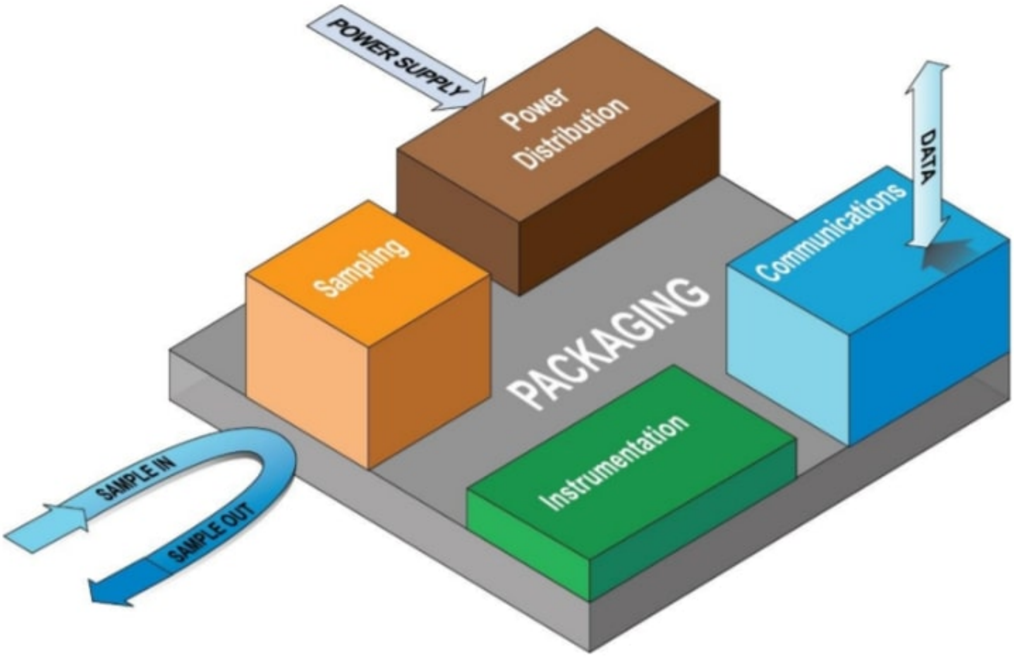

Figure 2.1. Block diagram of a water monitoring station

Figure 2.2. Water monitoring station location strategy

	Benefit	Problem
Company	Good Accuracy	Not accessible to everybody
	Reliable transmission mechanism	Abandoned water
	Data for all important points (purifiers,...)	Easier to corrupt
	Standard to respect	

Table 2.2. Benefits and problems of using Company as actor

2.3 The role of Makers in data collection

To solve the problem of free participation and consequently minimize the "abandoned waters" a possible solution is represented by that class of people who are both interested in the environment and the subject under discussion but also have the necessary skills to create a prototype that monitors the wastewater independently. However, given this total independence in the points to be monitored, there is a risk that some points that need to be checked do not receive the right attention from the makers.

With the access to open-source hardware devices, the *DIY* (*Do-it-yourself*) population has started realizing any electronic device. Since they do not have extensive financial and technological resources, their culture often involves open-source projects and solutions in both software and hardware within large communities working together.

Due to their scarce resources, they also need low-cost and easy-to-use sensors and products; this, however, is detrimental to the accuracy of sensitive measurements because the sensors are less accurate and their calibration deviates over time.

The same applies to the power supply to the monitoring stations. These can be placed anywhere, therefore, access to the electricity grid is not available and devices must be powered by battery: this requires the design and development of specific functions for energy saving. The device needs some sleep periods where consumption is reduced to a minimum and to use communication protocols that use little battery. For the latter, a possible solution is the LPWAN (Low-power wide area network) that is a type of wireless telecommunication wide area network designed to allow long-range communications at a low bit rate among things (connected objects), such as sensors operated on a battery. It is explained more in details in the Section 2.3.2.

Security is a critical topic in this scenario; being low-cost devices, there are not great resources to make them secure in terms of integrity with particular reference to updating the firmware by someone not authorized. A secure boot can avoid this problem: in order to do so it might be used a *Trusted Platform Module* (TPM) that enhances the security of general purpose computer systems by authenticating the platform at boot time and extending a chain-of-trust, based on next-boot-layer digest, until the conclusion of the boot process. However, working with TPM is complex and not all boards have one. So being challenging to work with, a possible solution is to work with SW-TPM which can be executed within protected or isolated execution domains that are increasingly provided by embedded CPUs (e.g., ARM TrustZone) and can utilize on-chip storage in order to provide a reasonable degree of tamper-resistance.[14] However, not all microcontrollers in the makers market are equipped with this feature making this aspect of integrity reasonably critical. Like the other cases, a table summarizes benefits and drawbacks for this actor.

	Benefit	Problem
Makers	Low Cost	Low accuracy
	Accessible to everybody	Scarce energy resource
	No more abandoned water	Integrity security drawbacks
	Easy-to-realize	No data certainty for all important points

Table 2.3. Benefits and problems of using DIY populations as actor

2.3.1 Low-Cost water quality sensors

As presented above in this scenario it is very important to choose the sensors carefully so everybody can use them at a relatively low-cost and are easy to use too. In this section several sensors that respect these requirements are presented: there is a temperature one, a turbidity one, a pH one, a dissolved oxygen one, a TDS one.

Waterproof DS18B20 Digital Temperature Sensor

This is a waterproofed version of the DS18B20 Temperature sensor produced by "DFRobot". Useful when it is necessary to measure in wet conditions, it supports a temperature range between -55℃ and 125℃ with an accuracy of ±0.5℃ from -10℃ to +85℃. This sensor requires a 4.7K Ohm resistor between the voltage and Signal pin but DFRobot offers a pluggable terminal sensor adapter to help in making this connection secure.

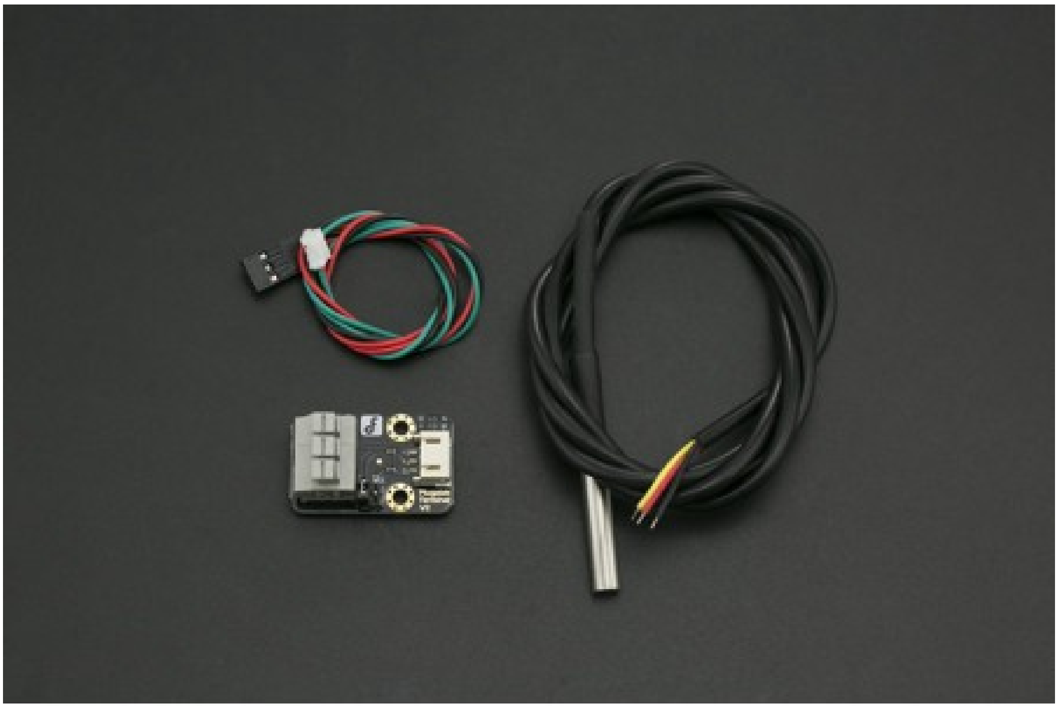

Figure 2.3. Waterproof DS18B20 Digital Temperature Sensor

The DS18B20 provides 9 to 12-bit (configurable) temperature readings over a 1-Wire interface so that only one wire (and ground) needs to be connected from a central microprocessor. Since it communicates using the 1-Wire protocol, it requires only one digital pin for communication and multiple sensors can share the same pin because each of them is identified by a unique 64 bit ID burned into the chip. The interaction with an easy-to-use microcontroller, like Arduino, results simple using one of the many shared libraries.

Analog TDS Sensor

TDS (Total Dissolved Solids) indicates the number of milligrams of soluble solids dissolved in one liter of water. The higher the TDS value, the more soluble solids dissolved in water, and the less clean the water is. Therefore, the TDS value can be used as one of the references for reflecting the cleanliness of the water.

Its TDS measurement range is between 0 and 1000 ppm with an accuracy of ±10%. Included in the kit, there is also a board that ensures correctness in the connection between the microcontroller and the probe.

Figure 2.4. Analog TDS Sensor

Turbidity Sensor

The turbidity sensor detects water quality by measuring the levels of turbidity. It measures the light transmittance and the scattering rate which change with the amount of total suspended solids (TSS). As the TSS increases, the liquid turbidity level increases giving the idea of the quality of water.

Figure 2.5. Turbidity Sensor

This liquid sensor provides analog and digital signal output modes. The threshold is adjustable when in digital signal mode through a potentiometer on the auxiliary board where is also located a switch to pass from analog to digital mode and vice versa.

Analog pH Sensor

This cheap sensor offers the possibility of measuring this significant property. With an accuracy of ± 0.1 pH and a response time less or equal to one minute, it embeds an electrode that has as reference solution the 3NKCL one.

The sensor needs to be calibrated before the long-term use and to do so two standard solutions are necessary, one with pH value around 7.00 and one around 4.00. These are used to note the two reference voltages related with these two standard solutions and, finally, computes a line to calculate the future pH values.

A notable precaution is that the electrode used for the first or long set without re-use should be immersed in the 3NKCL solution activated eight hours.

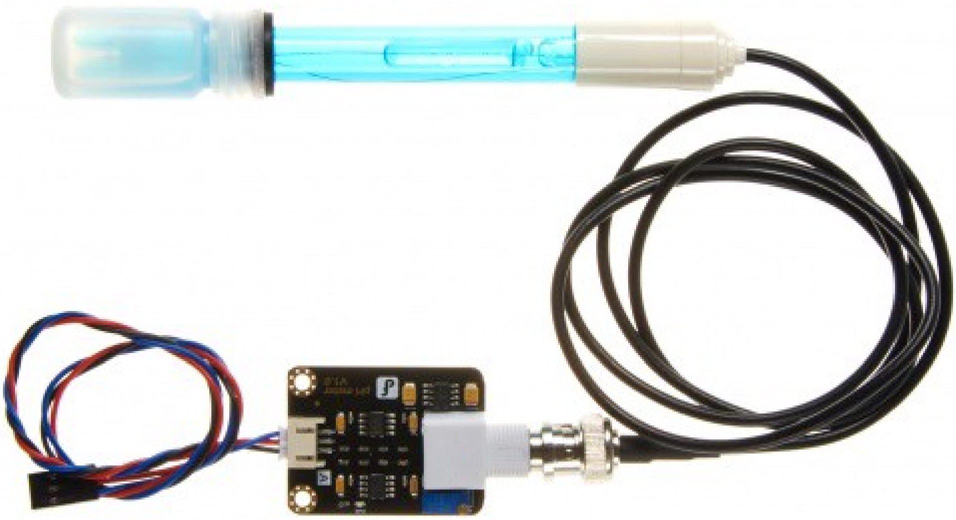

Figure 2.6. Analog pH Sensor

Analog Dissolved Oxygen Sensor

Good water quality is essential to the aquatic organisms. Dissolved oxygen is one of the important parameters to reflect the water quality. Low dissolved oxygen in the water will lead to difficulty in breathing for aquatic organisms, which may threaten the lives of aquatic organisms their lives. It is widely applied in many water quality applications, such as aquaculture, environment monitoring, natural science and so on. This sensor kit helps to build a dissolved oxygen detector quickly.

It has a response time within 90 seconds and it has a detection range of 0 and

Figure 2.7. Analog Dissolved Oxygen Sensor

20 mg/L. To be maintained correctly its membrane cap needs to be replaced within 1 2 months (in muddy water) or 4 5 months (in clean water).

The probe must be filled carefully using a 0.5 mol/L NaOH solution that should be added into the membrane cap. This process is necessary to calibrate the sensor properly.

Example: usage interaction with TDS sensor

Interacting with this class of sensors, as already mentioned, is very simple. In this section will be shown how to easily interact with the TDS sensor by means of an Arduino Uno board.

```
#define TdsSensorPin A1
// analog reference voltage(Volt) of the ADC
#define VREF 5.0
// sum of sample point
#define SCOUNT  30
// store the analog value in the array, read from ADC
int analogBuffer[SCOUNT];
int analogBufferTemp[SCOUNT];
int analogBufferIndex = 0,copyIndex = 0;
float averageVoltage = 0,tdsValue = 0,temperature = 25;

void setup()
{
    Serial.begin(115200);
```

```
    pinMode(TdsSensorPin,INPUT);
}

void loop()
{
   static unsigned long analogSampleTimepoint = millis();
   //every 40 milliseconds,read the analog value from the ADC
   if(millis()-analogSampleTimepoint > 40U)
   {
     analogSampleTimepoint = millis();
      //read the analog value and store into the buffer
     analogBuffer[analogBufferIndex] = analogRead(TdsSensorPin);
     analogBufferIndex++;
     if(analogBufferIndex == SCOUNT)
         analogBufferIndex = 0;
   }
   static unsigned long printTimepoint = millis();
   if(millis()-printTimepoint > 800U)
   {
      printTimepoint = millis();
      for(copyIndex=0;copyIndex<SCOUNT;copyIndex++)
        analogBufferTemp[copyIndex]= analogBuffer[copyIndex];

      // read the analog value more stable by the median filtering
      // algorithm and convert to voltage value

      averageVoltage = getMedianNum(analogBufferTemp,SCOUNT) *
                                  (float)VREF / 1024.0;
      //temperature compensation
      float compensationCoefficient=1.0+0.02*(temperature-25.0);
      float compensationVolatge=averageVoltage/
                                  compensationCoefficient;
       //convert voltage value to tds value
      tdsValue=(133.42*compensationVolatge*compensationVolatge
                                *compensationVolatge -
            255.86*compensationVolatge*compensationVolatge +
            857.39*compensationVolatge)*0.5;
      Serial.print("TDS Value:");
      Serial.print(tdsValue,0);
      Serial.println("ppm");
   }
}
int getMedianNum(int bArray[], int iFilterLen)
{
      int bTab[iFilterLen];
      for (byte i = 0; i<iFilterLen; i++)
      bTab[i] = bArray[i];
      int i, j, bTemp;
      for (j = 0; j < iFilterLen - 1; j++)
      {
      for (i = 0; i < iFilterLen - j - 1; i++)
          {
```

```
        if (bTab[i] > bTab[i + 1])
            {
        bTemp = bTab[i];
            bTab[i] = bTab[i + 1];
        bTab[i + 1] = bTemp;
         }
      }
      }
      if ((iFilterLen & 1) > 0)
    bTemp = bTab[(iFilterLen - 1) / 2];
      else
    bTemp = (bTab[iFilterLen / 2] + bTab[iFilterLen / 2 - 1])/2;
      return bTemp;
}
```

The main function is *loop*, a method that executes infinitely, that every 40 milliseconds reads through the ADC channel, invoked through the *analogRead* function, and saves the obtained value, between 0 and 1023, in an array. The latter is used every 800 milliseconds to calculate the median value of the last measurements, which is then normalized to a value between 0 and 5.0V, because of this voltage powers it, and then converted, in the end, into the standard TDS value using a simple mathematical formula.

2.3.2 Data Transmission: LPWAN and The Things Network

A low-power wide-area network (LPWAN) is a type of wireless telecommunication wide area network designed to allow long-range communications at a low bit rate among things. The low power, low bit rate and intended use distinguish this type of network from a wireless WAN that is designed to connect users or businesses, and carry more data, using more power. The LPWAN data rate ranges from 0.3 kbit/s to 50 kbit/s per channel. [10]

There are many competing standards and vendors in the LPWAN space:

- *Sigfox* is a French global network operator founded in 2009 that employs the differential binary phase-shift keying (DBPSK) and the Gaussian frequency shift keying (GFSK) that enables communication using the Industrial, Scientific and Medical ISM radio band which uses 868MHz in Europe and 902MHz in the US.

- *NB-Fi Protocol* employs a proprietary Ultra Narrow Band technology that enables communication using the Industrial, Scientific and Medical ISM radio band (and in other parts of sub-GHz license-free spectrum as well).

- *LoRa* uses license-free sub-gigahertz radio frequency bands like 169 MHz, 433 MHz, 868 MHz (Europe) and 915 MHz (North America) that employs a Chirp Spread Spectrum (CSS) modulation. The technology has two parts, the physical layer and LoRaWAN. LoRaWAN is a media access control (MAC) layer protocol but mainly is a network layer protocol for managing communication between LPWAN gateways and end-node devices as a routing protocol, maintained by the LoRa Alliance.

LoRaWAN defines the communication protocol and system architecture for the network, while the LoRa physical layer enables the long-range communication link.

LoRaWAN is also responsible for managing the communication frequencies, data rate, and power for all devices. Devices in the network are asynchronous and transmit when they have data available to send. Data transmitted by an end-node device is received by multiple gateways, which forward the data packets to a centralized network server. The network server filters duplicate packets, it performs security checks and manages the network. Data is then forwarded to application servers. The technology shows high reliability for the moderate load, even if, it has some performance issues related to sending acknowledgements.[11]

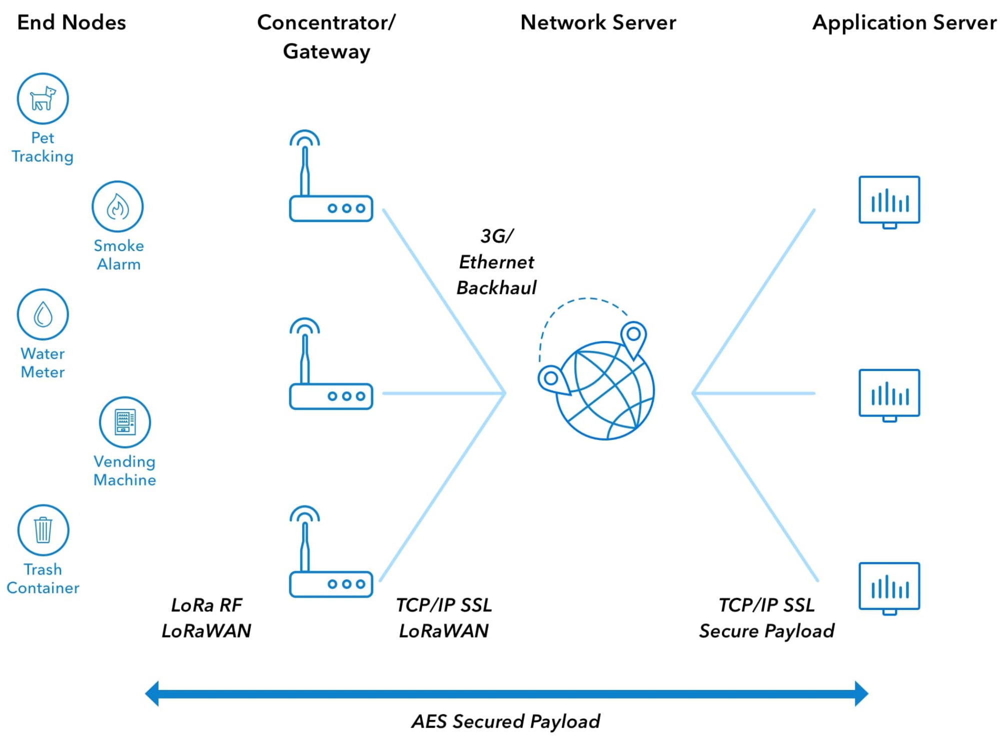

Figure 2.8. LoRaWAN functioning architecture

The Things Network is building a globally distributed, crowd-sourced and open IoT data network, owned and operated by its users. Using low power, long range technologies, The Things Network provides an end-to-end stack: from nodes to gateways, network server, device management and integrations with major cloud providers and IoT platforms. All fully secure and configurable by the end user.[13]

However, the architecture described in the picture (2.8) might have some scalability limitations because the number of connected devices should increase in the future years exceeding LPWAN transmission capacity. Organizing the network using hierarchical structures reduces the number of the LPWAN enabled or equivalent devices within a certain area and consequently reduce the total number of devices accessing the sub-GHz spectrum.[29] Indeed, devices are increasingly equipped with two or more different networking interfaces. One network interface provides short-range high-rate communication capabilities by exploiting IEEE 802.15.4 technologies. Another network interface provides long-range low-rate communication based on recently introduced LPWAN technologies. End-level IoT devices, (e.g., inside a house) can communicate locally using non-LPWAN technologies to exchange data and choose a nearby one that is capable of accessing the LPWAN network to relay their information to the gateway.

Security in LPWAN Transmission Networks

To realize secure transmission, the application key *AppKey* is only known by the device and by the application. When a device joins the network (this is called a join or activation), an application session key *AppSKey* and a network session key *NwkSKey* are generated. The *NwkSKey* is shared with the network, while the *AppSKey* is kept private. These session keys will be used for the entire duration of the session.[12]

The algorithm used for this is AES-128, similar to the algorithm used in the 802.15.4 standard. The NwkSKey is used to validate the integrity of each message by its Message Integrity Code (MIC). This MIC is similar to a checksum, except that it prevents intentional tampering with a message. For this, LoRaWAN uses AES-CMAC. The AppSKey is used for encryption of the application payload.[12]

The definition of the NwsKey and of the AppSKey can be realized either by *Activation-by-Personalization* (ABP) or by *Over-The-Air Activation* (OTAA). In the ABP case, the authentication data are hard-coded into the device before it begins the communication with the network; hence, no join procedure is required. These data tie the end-device to a particular LoRaWAN network, as they contain, also, the LoRaWAN Network Identifier (NwkId), the Network Address (NwkAddr) of the device and the cryptographic session keys. Thus, the end-device is not allowed to communicate with other LoRaWAN networks while these values remain the same. [25]

During the OTAA, which is the usually recommended scheme, a message exchange including cryptographic material is performed between the end-device and the network server. By using this material and the AppKey, both the NwkSKey and the AppSKey are derived. After this step, communications are secured by using the derived keys. This process is carried out when the device connects for the first time to the network or in case of connectivity loss. [25]

However, the connection structured in this way has some limitations regarding the absence of periodic updates of such keys and the risk that if an attacker succeeds in obtaining these session keys, all the encrypted information exchanged between the end-device and the network server will be accessible during a long period of time until the attack is detected. The work in [25] describes several protocols that may be considered as potential alternatives to design a suitable key management approach. In particular, it assesses the use of IKE[26], that is a component of IPsec used for performing mutual authentication and establishing and maintaining Security Associations (SAs), DTLS[27], which provides communications privacy for datagram protocols, and COSE[28], which describes how to create and process signatures, message authentication codes, and encryption using CBOR for serialization. Finally, it designs an architecture to exploit better all these protocols combined.

Chapter 3

Data storage analysis

The data collected from the measurements of the monitoring stations must be stored in data structures which ensure a certain level of security and which respect specific properties. The use of a centralized system may have some limitations about the below-mentioned characteristics.

- *Public Verifiability* allows anyone to verify the correctness of the state of the system. In a centralized system, observers need to trust the central entity to provide them with the correct state; they might not be able to verify all state transitions were executed correctly.[15]
- *Integrity* of information ensures that data is protected by unauthorized modifications. It is linked with "Public Verifiability"; if the system provides this latter, anyone might verify the integrity of the data. A central system ensures it if it is not compromised otherwise the whole storage system would be too. [15]
- *Transparency* of the data and the process of updating the state is a requirement for public verifiability.[15]
- *Redundancy* of the data is important for many use cases. In a centralized system, it is achieved through replication on di�erent physical servers and backups.[15]
- *Fault-tolerance and Availability* are important to ensure that at any moment is possible for the data to flows into a storage system. In a centralized system, if the server would not be online for a while, the whole storage system will su�er.[15]

A centralized system may not be the ideal solution if you do not want to agree on a trusted third party or when di�erent entities do not trust each other; in this case, *Integrity* and *Public Verifiability* acquire a primary role. The choice of a distributed approach, like the Blockchains, might represent the optimal one.

3.1 Preliminaries on Blockchain technologies

The name Blockchain comes from its technical structure - a chain of blocks. A Blockchain is a shared, trusted, public ledger of transactions, that everyone can inspect but which no single user controls.[16] Transactions are created and exchanged by peers of the blockchain network and modify the state of the blockchain. As such transactions, for example, can exchange monetary amount and also allow to execute arbitrary code within so-called smart contracts. The ledger is built using a linked list, or chain of blocks, where each block contains a certain number of transactions that were validated by the network in a given timespan. The crypto-economic rulesets of the blockchain protocol (consensus layer) regulate the behavioral rulesets and incentive mechanism of all stakeholders in the network.[16]

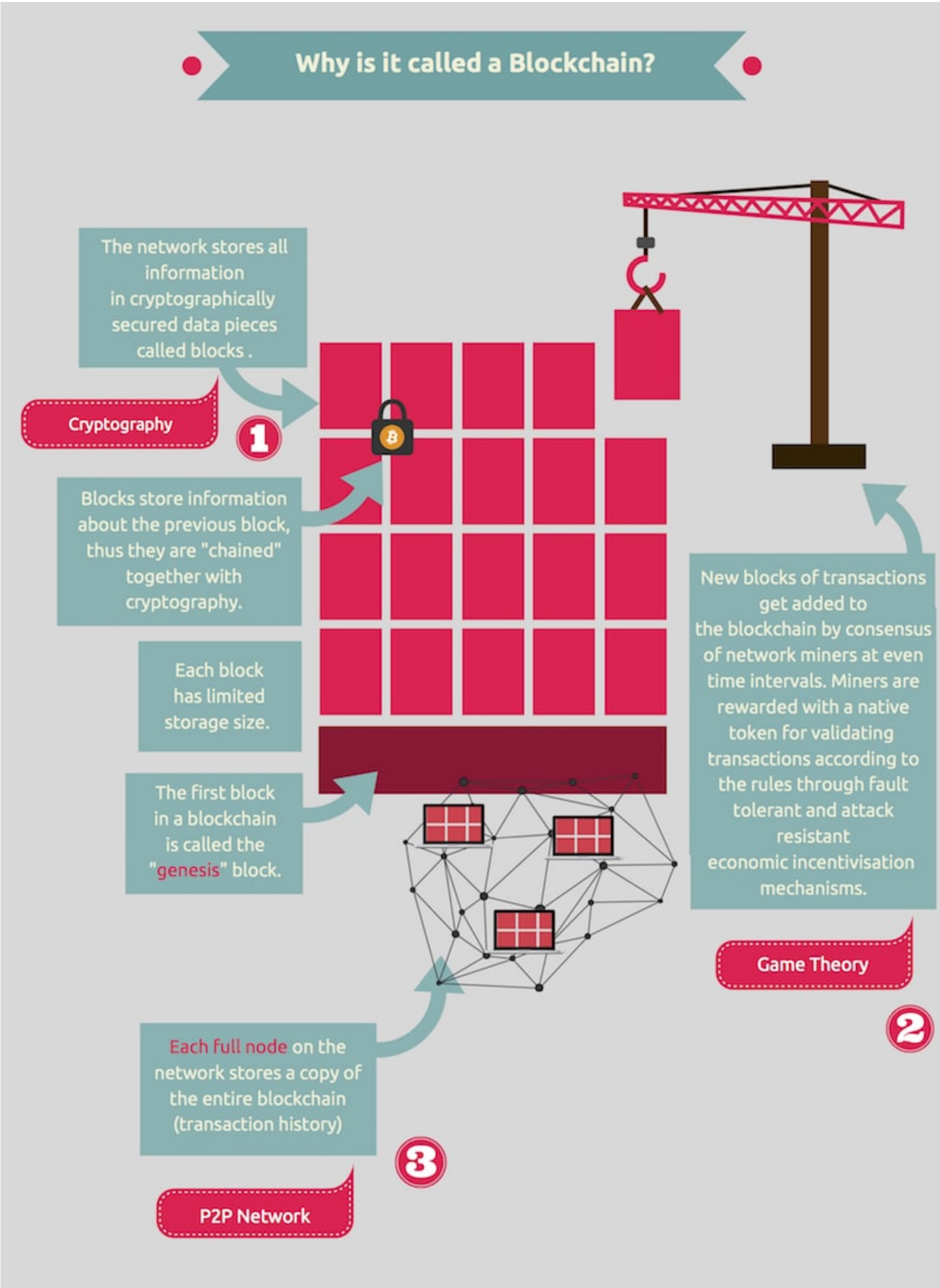

Figure 3.1. Blockchain architecture

There are two types of participants for these networks. It is denoted as *writer* any entity which is involved in the consensus protocol and helps grow the blockchain.

As such, a writer can accumulate transactions within a block and append this to the blockchain. A *reader* is an entity which does not extend the blockchain but participates either creating transactions or simply reading and auditing the blockchain. [15]

On the top of these assumptions also two types of blockchain may be described:

- *Permissionless Blockchains* are open and decentralized. Any peer can join or leave the network as reader and writer at any time and there is no central entity which can add or ban illegitimate readers or writers. This structure implies that any transaction on the blockchain is visible by everybody. Examples of this kind of blockchain are Bitcoin [17] and Ethereum.[18]
- *Permissioned Blockchains* only authorize a limited set of readers and writers. A central entity decides the peers that can write and read in the blockchain. An example of permissioned blockchain is Hyperledger. [19]

3.2 When to use blockchain

Generally it is not always convenient to use a blockchain as a data structure to store information; in fact, the use of this technology has limitations in terms of performance due to its relatively low scalability as it requires for each transaction to execute the consensus among writers. For this reason, the choice of the blockchain must be justified in the design phase. It makes sense to use it when multiple entities do not trust each other and want to interact without the involvement of a trusted third party.

Wurst and Gervais in their paper "Do you need a blockchain?"[15] have schematized a technique to evaluate whether the use of blockchain makes sense or not, summarized in the (3.2) figure.

If no data needs to be saved, then it is useless that we pose the problem both to use a database and to use the blockchain. Similarly, if it is only an entity that has to store the data, then write to the system, then the use of blockchains do not provide any additional advantage over a central database, which instead is more performing in terms of latency and throughput. If there is a trusted third party (TTP), two are the possible options. Option one, that it is always online and then you can delegate the task of verifying the transactions of di�erent writers using a centralized approach. Otherwise, option 2 that it is often o�ine then acts as an authoritative body that allows some entities to be writers for a permissioned blockchain.

If, we assume that no writer is malicious and they are all trusted then it makes no sense to use a blockchain while if the number of writers is known but there is no mutual trust then the choice of permissioned blockchain is the best option. It can be public, if public verifiability is allowed to all, or private if a minimum level of privacy must be guaranteed and then the set of readers is restricted.

Finally, if the number of participants is not known and there is no mutual trust, as in the case of cryptocurrency like Bitcoin, then the choice of permissionless blockchain is the best.

Once again the focus is on the tradeo� between decentralisation and scalability of the system. An accurate evaluation of this factor must be also taken into account.

3.3 Data storage analysis of actors

In this section, the relationship between the actors seen so far, government, companies and DIY people, and storage technologies, permissionless blockchain,

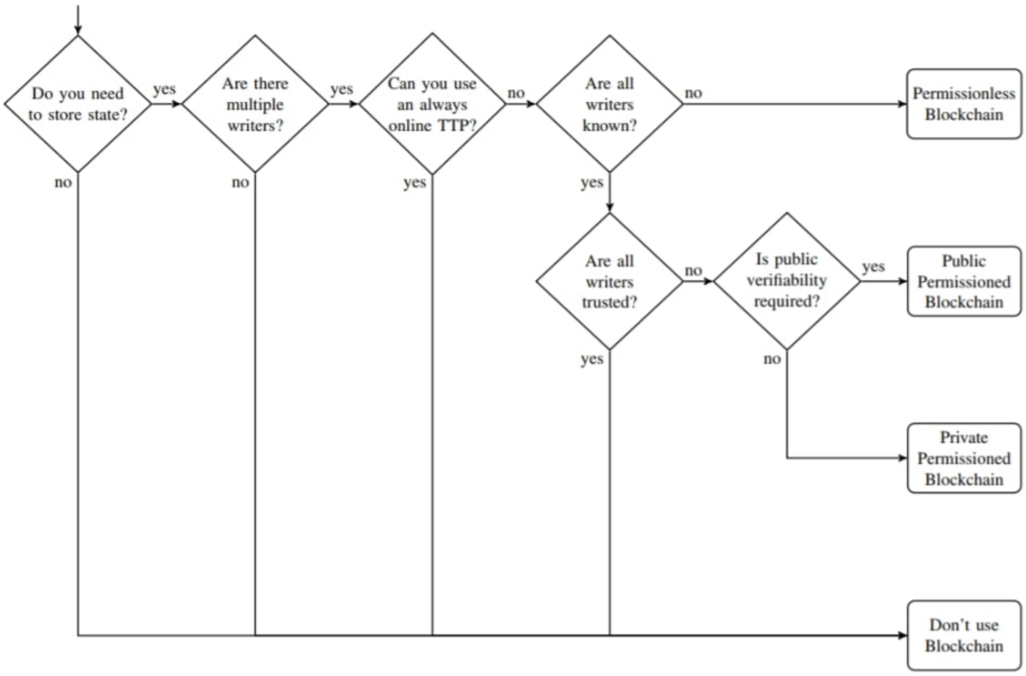

Figure 3.2. When to use a blockchain scheme

permissioned blockchain and central database, will be analyzed, following the method described above. It is not always trivial to choose suitable technology.

3.3.1 Government or State data storage analysis

In case the Government or the State is the only writer who can perform writing operations the use of a blockchain is meaningless because there is only one writer the problem of trust between multiple entities does not arise. The best solution is, therefore, the use of a central server that also in terms of throughput of transitions provides high-performance results. However, in this circumstance, you lose any opportunity to provide public verifiability by placing complete trust in the State and then losing the ability for someone to verify the correctness of the state of the system.

	Govern / State
Permissionless Blockchain	Since there is just one writer it doesn't make sense to use it
Permissioned Blockchain	Since there is just one writer it doesn't make sense to use it
Central Server	Best solution but it requires total trust in the State

Table 3.1. Summary table on data storage analysis for the State

3.3.2 Company data storage analysis

Now let's consider companies as the main players. In their case there are many writers, there is no trust between the different entities and you can not count on a third trusted party always online: using the blockchain makes sense and seems to be a good solution. But which one? Permissionless or permissioned? There are two situations to consider.

- A company participates in this monitoring spontaneously. It makes the number of writers unknown so the best solution would be to use a permissionless blockchain. However, two issues arise: the first is to develop a method for identifying companies; the second is that it is necessary to design an incentive to participate in this kind of so structured monitoring otherwise it would be difficult for a company to participate.
- A company participates because it is obliged to do so. Also, in this case, there are multiple writers, no third trusted parties always online and not even in this case the various entities trust each other. However, being known the number of companies that write in the system is correct to use a permissioned blockchain. This must be made "public" because it must ensure public verifiability and everyone must have the opportunity to check the correctness.

Note that if it is wanted to apply a solution with a centralized server in this scenario it would need a trusted online third party that would have the task of verifying the correctness of the data.

	Company
Permissionless Blockchain	It would make sense if companies would participate spontaneously. Then, the number of writers is not known.
Permissioned Blockchain	It would make sense if very company has to participate. Since there isn't a TTP always online, this solution might work
Central Server	TTP alway online needed

Table 3.2. Summary table on data storage analysis for the Company

3.3.3 Makers data storage analysis

It is the case of greater openness to participation and of course as such more malicious entities can infiltrate to corrupt the data: there is a need, therefore, to use a technology where the various entities can collaborate even not trusting each other as the blockchain. This makes sense to be also used because no one can act as a third trusted party always online.

Of course, even the number of participants cannot be known to use the most general solution as the permissionless blockchain is the right one.

	Makers
Permissionless Blockchain	It would make sense: it is a trustless system and the number of writers is not known at priori
Permissioned Blockchain	Not the appropriate solution. It allows a trustless application but the problem of participants is not known
Central Server	TTP needed

Table 3.3. Summary table on data storage analysis for the Makers

However, this system will su er the slow throughput due to the consensus algorithm process allowing it to manage at most 10 transactions per second instead of a centralized system which reaches thousands of transactions per second.

	Govern/State	Company	Makers
Permissionless Blockchain	Since there is just one writer it doesn't make sense to use it	It would make sense if companies would participate spontaneously. Then, the number of writers is not known.	It would make sense: it is a trustless system and the number of writers is not known at beginning
Permissioned Blockchain	Since there is just one writer it doesn't make sense to use it	It would make sense if the companies have to participate. Since there isn't a TTP always online, this solution might work.	Not the appropriate solution. It allows a trustless application but the problem of participants is not known
Central Server	Best solution but it requires total trust in the State	TTP always online needed	TTP needed

Table 3.4. Summary table on data storage analysis

Chapter 4

Data processing analysis

It has been described so far how data is taken from the real world in relation to the person taking it and how this data is stored in order to ensure integrity and public verifiability. All this is only a first step in resolving the problem described in the beginning.

A further step is represented by strongly taxing those purifiers that do not respect certain parameters of water quality that grows over the years following a model similar to that of the carbon tax.

The carbon tax[20] is a tax on energy resources that emit CO_2 into the atmosphere. It is an example of an ecotax, which has been proposed by economists as preferable because it taxes something "bad" rather than "good". It is a fiscal policy instrument according to which every tonne of carbon dioxide pollution released by fossil fuels will be subject to a rate set by the government. The European Union discussed a carbon tax for its member states, as well as an emissions trading on carbon that started in January 2005.

On the top of this, this paper supports the creation of a "quality-credit" market where there are two players:

- *Virtuous companies* that can acquire, then eventually resell, credits (tokens) because they meet the quality thresholds.
- *Incorrect companies* that do not respect the quality thresholds but that can buy from virtuous companies at most a percentage (which decreases over time) of credits to reach the quality threshold set for that year (or period) otherwise, most likely, they run into the steep fine.

This approach aims to increase the average quality of water over the time by giving the opportunity to participating companies to gradually invest over time and possibly earn for their good behavior.

Company 1	*Good Measurements Ratio*	0.8	0.79	0.76	0.73	0.90
	$	500	1300	800	-200	0
	Credits bought	0	0	0	2	0
	DO value	4.0	4.5	6.5	8.0	9.0
	Percentage credit bought allowed	20	15	10	5	3
Company 2	*Good Measurements Ratio*	0.70	0.62	0.67	0.78	0.90
	$	-500	-1300	-800	200	0
	Credits bought	5	13	8	0	0

Table 4.1. Example of quality-credit market execution

Figure 4.1. Summary diagram of the "quality-credit" market functioning

The table (4.1) shows a little example to see better how this market works. Suppose a Govern wants to improve the "Dissolved Oxygen" values (mg/L) of its wastewater returned to the environment. It would like to achieve it in 5 years and two companies are involved in this program: each of them must achieve 75% of measurements that meet the DO threshold in returned wastewater. If this is lower than 75%, a company might buy 1% credits for $100 each one.

It is possible to note how the average of dissolved oxygen values increase over the time passing from 4.0 mg/L to 9.0 mg/L in five years. Moreover, it is advantageous for both the companies: "Company 2" is not fined anymore, actually at the last period would get paid by "Company 1", and this latter from the beginning would earn because of its good behavior.

In the following sections is presented how normal carbon taxes and emission trade might be represented in game theory to optimize both companies profits and environmental protection. Then, this is transformed into a new model that fits better for our purposes and a brief description of the technologies implied is provided.

4.1 A Stackelberg game model for an emission trading scheme

Other examples of game theory analysis models already exist in the literature, but one of interesting relevance is the Stackelberg model provided by Zhaofu Hong, Chengbin Chu, Linda L. Zhang, Yugang Yu, *Optimazing an emission trading scheme for local governments: A Stackelberg game mode and hybrid algorithm*.

It claims that the emission trading scheme involves a regulator (or local government), which implements emission regulations through the allowance trading scheme, and M firms in the same sector indexed by i with $i = 1, ..., M$, which have to design properly their financial plan in order to respect the regulations and maximize their profit. Each firm might use either *green technologies*, which cost more but emit less CO_2, and *regular technologies*, which do the opposite. Let:

- c_{ir} and c_{ig} be the unit production cost using regular and green technologies with $c_{ir} < c_{ig}$.
- s_{ir} and s_{ig} be the setup costs using regular and green technologies with $s_{ir} < s_{ig}$
- e_{ir} and e_{ig} be the emissions generated from using regular and green technologies with $e_{ir} > e_{ig}$

In the allowance trading scheme, the government for each period defines an emission reduction target φ, with $0 < \varphi < 1$ which is a reduction ratio that indicates how much the government would reduce the emissions and that is used to calculate the initial emission allowed for each firm. This is denoted with IE_i as the initial emission allowance allocated to firm i and that has this formula:

$$IE_i(\varphi) = (1-\varphi)\cdot\frac{1}{M}\sum_{i=1}^{M} e_{ir}\cdot\sum_{t=1}^{T} d_{it},$$

where the production period is indexed by $t = 1, ..., T$ and T is the number of periods and also the planning cycle of the emission reduction target. d_{it} is firm i's demand at period t; $(1-\varphi)\cdot\frac{1}{M}\sum_{i=1}^{M} e_{ir}$ is the number of emissions for each unit of product; finally, $\sum_{t=1}^{T} d_{it}$ is the market share of firm i in the planning horizon T.

The initial emission value needs to respect the environmental bearing capacity which is a feature of the ecosystem and that represents a threshold of emission level that a region can endure. This is denoted with Ω_i and the following relationship must be ensured: $IE_i(\varphi) < \Omega_i$.

The problem is formulated within a Stackelberg game where the government acts as leader and can fixes the emission reduction target φ, while the firms are the followers making production planning decisions in response to φ.

4.1.1 Firm's decision problem

In accordance with the emission trading scheme, firms optimize their production to maximize their profits respecting the constraints of emission allowances. They can buy or sell emission allowances from or to other firms in the region at an allowance trading price, denoted as y. The trading price is defined under the Cournot competition game framework. We first analyze the production planning decision problem and then we investigate the allowance trading price.

Let $x_{itr} \geq 0$ and $x_{itg} \geq 0$ be firm i's production quantities using the regular and the green technologies at period t, respectively. $f_i(x_{itr}, x_{itg})$ represents the production cost at period t which has this form:

$$f_i(x_{itr}, x_{itg}) = \sum_{t=1}^{T}(s_{ir}\delta(x_{itr}) + c_{ir}x_{itr} + s_{ig}\delta(x_{itg}) + c_{ig}x_{itg}),$$

where $\delta(x_{itr}) = 1$ if $x_{itr} > 0$ and $\delta(x_{itg}) = 1$ if $x_{itg} > 0$.
The amount of actual emissions of firm i, AE_i, is indicated by

$$AE_i = \sum_{t=1}^{T}(e_{ir}x_{itr} + e_{ig}x_{itg}).$$

Using $\mathbf{x}_{ir}$ and $\mathbf{x}_{ig}$ to indicate the vectors of x_{ir} and x_{ig} and p_{it} as the product price of firm i in period t, d_{it} firm i's customer demand in period and y fixed, the firm decision problem might be formulated as follows

Model F

$$\text{Max } \pi_i(\mathbf{x}_{ir}, \mathbf{x}_{ig}|y) = \sum_{t=1}^{T} (p_{it}d_{it} - f_{it}(x_i tr, x_i tg) - h_{it}I_{it}) - (AE_i - IE_i)y, \quad (4.1)$$

$$\text{s.t. } I_{it} + d_{it} = I_{i(t-1)} + x_{itr} + x_{itg}, \quad (4.2)$$

$$e_{ir}x_{itr} + e_{ig}x_{itg} \le \Omega_i, \quad (4.3)$$

$$I_{it} \ge 0, i = 1, ..., M, t = 1, ..., T \quad (4.4)$$

where the objective function (4.1) consists of two terms. The first term includes the production, and inventory costs; the second represents either the cost or the revenue of either buying or selling emission allowances. (4.2) ensures that the inventory is balanced and (4.4) ensures that the inventory level is nonnegative. Finally, (4.3) limits the emission released by i to less than (at most equal to) the region's bearing capacity in each period.

The initial emission allowances are tradable among firms. They can either buy them if they pollute too much or sell them if, having invested in green technologies, respect the emission limits. So they all are peers in the allowances trading market; each of them may define the proper price for an allowance that maximizes their profits. Therefore, we can define a Cournot competition game model where each firm aims to maximize its profit by optimally determining $\mathbf{x}_{ir}$ and $\mathbf{x}_{ig}$, according to the allowance trading price. Use this model helps to obtain the equilibrium allowance trading price. This and firms' production policies affect each other: if firms choose to invest in green technology the price decrease; instead, if companies don't invest in green technology the allowance trading price raises. The model is formulated as follows:

Model P

$$\begin{cases} AE_1 : max\ \pi_1(\mathbf{x}_{1r}, \mathbf{x}_{1t}|y), \\ AE_2 : max\ \pi_2(\mathbf{x}_{2r}, \mathbf{x}_{2t}|y), \\ \dots \\ AE_M : max\ \pi_M(\mathbf{x}_{Mr}, \mathbf{x}_{Mt}|y) \end{cases}$$

The equilibrium allowance trading price is achieved when the total emissions (defined as $TAE = \sum_{i=1}^{M} AE_i$) are equal to the total initial emission allowances (defined as $TIE = \sum_{i=1}^{M} IE_i$), i.e., $TAE = TIE$. In this way each firm obtain the maximal profit.

4.1.2 Government's decision problem

The government, as Stackelberg leader, sets the emission reduction target in consideration of the firms' responses to maximize the social welfare of the region. In the literature, the social welfare consists of three components named economic benefits, consumer surplus, and negative environmental impact.[22]. Since the prices of the products are exogenous the consumer surplus can be omitted from the social welfare function.

The economic benefits are the profits of all firms in the region.

$$\sum_{i=1}^{M} \pi_i(x_{ir}, x_{ig}) - \sum_{i=1}^{M} (AE_i - IE_i)\gamma. \tag{4.5}$$

The environmental impact is the emission released by firm i as ξAE_i, where ξ is a monetary term indicating the cost of dispelling the impacts of the emissions. Therefore, the total environmental impact cost is

$$\sum_{i=1}^{M} \xi AE_i \tag{4.6}$$

Finally, combining (4.1), (4.5) and (4.6) the government's decision problem is

Model G

$$\text{Max } W_g(\varphi) = \alpha_{ec} \sum_{i=1}^{M} \sum_{t=1}^{T} (p_{it}d_{it} - f_{it}(x_{itr}, x_{itg}) - h_{it}I_{it}) - \beta_{ev} \sum_{i=1}^{M} \xi AE_i \tag{4.7}$$

$$\text{s.t. } \varphi > 0, \tag{4.8}$$

$$IE_i(\varphi) < \Omega_i, \tag{4.9}$$

where α_{ec} and β_{ev} are the coeﬃcients reflecting the government's preferences for economic and environmental benefits, respectively.

4.2 A Stackelberg game model for a wastewater "quality-credit" trading scheme

On the basis of the previous model, a similar model will be created for the "quality-credit" trading scheme supported in this chapter. Let i ($i = 1, ..., M$) be the index of the purifier company and M be the number of companies. In our case, we don't treat with regular and green technologies but since in this model the quality threshold of the water increases over the time, it can be diﬀered among: *bad-purified* water, the wastewater that doesn't respect the quality threshold, and *good-purified* water, the wastewater that respects the quality threshold. To assess a fixed amount of water this can be realized estimating the quantity of water that flows among two measurements of the monitoring station and referring to it as *unit*. c_{ib} and c_{ig} are the unit costs of the bad purified and well purified water, respectively; s_{ib} and s_{ig} are the setup costs for the bad purification and well purification; e_{ib} and e_{ig} are the measures that indicate how much wastewater still contaminates after the purification process. The following rules always stand: $c_{ib} < c_{ig}, s_{ib} < s_{ig}, e_{ib} > e_{ig}$.

As the previous model, the government draws up a reduction target and allocates tradable initial allowances to each company in the light of the target: the initial allowances represent the number of bad measurements allowed. The polluted wastewater reduction target, denoted with φ, is a ratio that needs to fall in the interval $0 < \varphi < 1$ that is used for each company i to calculate the initial bad allowances, representing the maximum number of allowed measurements detecting non-compliant treatment and denoted with IB_i, in this way:

$$IB_i(\varphi) = (1 - \varphi)\frac{1}{M}\sum_{i=1}^{M} e_{ib} \cdot \sum_{t=1}^{T} d_{it}$$

where the period is indexed by $t = 1, \dots, T$ and T is the number of periods and also the planning cycle of the emission reduction target. d_{it} is the estimation of the company i's demand of wastewater to be purified at period t; $(1 - \varphi) \cdot \frac{1}{M} \sum_{i=1}^{M} e_{ib}$ is the quantity of contaminators for each unit of output wastewater; finally, $\sum_{t=1}^{T} d_{it}$ is the market share of company i in the planning horizon T. The self-purification of waters is defined as the totality of biological, physical and chemical processes which take place within the water reservoir and which lead to the diminution of pollutants concentration to a level that is harmless for ecosystems functionality. [23] This process may be realized if the contaminants respect the environmental bearing capacity, denoted with Ω_i, of wastewater self-purification, therefore, the initial bad allowances need to always be less than the environmental bearing capacity, i.e. $IB_i(\varphi) < \Omega_i$, otherwise the self-purification process will not return the water to its original state.

One of the requirements of the quality-credit trading scheme scenario is that the amount of credits that can be purchased is reduced over time to prevent those who have more money but do not invest in new technologies for better purification from continuing to buy credits instead of modernizing. For this reason the variable (where $0 < \quad < 1$) is introduced for which a quantity $IB_i({}_t)$ can be calculated that respects this property: $IB_i(\varphi) + IB_i({}_t) < \Omega_i$. Since the percentage that can be purchased must decrease over time, the following property also applies:

$$t \quad 1, \dots, T, k \quad I : {}_{t} > {}_{t+k}$$

4.2.1 Companies' decision problem

In the "quality-credit" trading scheme, the companies (Stackelberg followers) plan their production and trade credits among each other to comply with the regulation from the government.
Let $x_{itb} \geq 0$ and $x_{itg} \geq 0$ be the output of water that is badly or well purified for company i at period t. The amount of total non-compliant measurements for company i is formulated as:

$$AB_i = \sum_{t=1}^{T} e_{ib} x_{itb}$$

Company i's production cost is formulated as:

$$f_i(x_{itb}, x_{itg}) = \sum_{t=1}^{T} (s_{ib}\delta(x_{itb}) + c_{ib}x_{itb} + s_{ig}\delta(x_{itg}) + c_{ig}x_{itg}),$$

where $\delta(x_{itb}) = 1$ if $x_{itb} > 0$ and $\delta(x_{itg}) = 1$ if $x_{itg} > 0$.

As in the previous section, models can be defined both for the companies' decision problem and for that of Cournot to create a credit sale price that maximizes the profits of all the companies.

Model C

$$\text{Max } \pi_i(x_{ib}, x_{ig}|\gamma) = \sum_{t=1}^{T} (p_{it}d_{it} - f_{it}(x_{itb}, x_{itg}) - h_{it}I_{it}) - (AB_i - IB_i)\gamma, \quad (4.10)$$

$$\text{s.t. } I_{it} + d_{it} = I_{i(t-1)} + x_{itb} + x_{itg}, \quad (4.11)$$

$$e_{ib}x_{itb} + e_{ig}x_{itg} \leq \Omega_i, \quad (4.12)$$

$$AB_i \leq IB_i(\varphi) + IB_i(\ ) \quad (4.13)$$

$$I_{it} \geq 0, i = 1, ..., M, t = 1, ..., T \quad (4.14)$$

Model P

$$\begin{cases} AB_1 : max\ \pi_1(x_{1b}, x_{1t}|\gamma), \\ AB_2 : max\ \pi_2(x_{2b}, x_{2t}|\gamma), \\ ... \\ AB_M : max\ \pi_M(x_{Mb}, x_{Mt}|\gamma) \end{cases}$$

The equilibrium allowance trading price is achieved when the total non-compliant measurements (defined as $TAB = \sum_{i=1}^{M} AB_i$) are equal to the maximum number of non-compliant allowed measurements (defined as $TIB = \sum_{i=1}^{M} IB_i$), i.e., $TAB = TIB$. In this way each company obtain the maximal profit.

4.2.2 Government's decision problem

The government (Stackelberg leader) sets the polluted wastewater reduction target in consideration of the companies' responses to maximize the social welfare in the region. The components of the social are the same described in the previous model and it makes sense to consider for this case just the economic benefits, the profits of all firms in the region, and the environmental impact, the negative utility in the social welfare corresponding to the environmental damage caused by wastewater pollution, that are formulate as the following:

$$\sum_{i=1}^{M} \pi_i(x_{ib}, x_{ig}) - \sum_{i=1}^{M} (AB_i - IB_i)\gamma. \quad (4.15)$$

$$\sum_{i=1}^{M} \xi AB_i \quad (4.16)$$

where ξ represents the monetary cost of dispelling the contaminated wastewater into the environment.

Once again based on (4.10), (4.14), (4.15) is possible to define the government's decision problem that maximize the social welfare.

Model G

$$\text{Max } W_g(\varphi, \ ) = \alpha_{ec} \sum_{i=1}^{M} \sum_{t=1}^{T} (p_{it} d_{it} - f_{it}(x_{itb}, x_{itg}) - h_{it} I_{it}) - \beta_{ev} \sum_{i=1}^{M} \xi AB_i \quad (4.17)$$

$$\text{s.t. } \varphi > 0, \quad (4.18)$$

$$IB_i(\varphi) + IB_i(\ ) < \Omega_i, \quad (4.19)$$

where α_{ec} and β_{ev} are the coefficients reflecting the government's preferences for economic and environmental benefits, respectively.

4.2.3 Useful technologies suitable for the "quality-credit" trading scheme"

The acquisition, sale and purchase of quality-credits can be carried out automatically by developing techniques that optimize the profit and timing of companies: smart contracts are one of those. A smart contract is a computer code running on top of a blockchain containing a set of rules under which the parties to that smart contract agree to interact with each other. If and when the pre-defined rules are met, the agreement is automatically enforced. The smart contract code facilitates, verifies, and enforces the negotiation or performance of an agreement or transaction. It is the simplest form of decentralized automation. Smart contracts radically reduce transaction costs. Auto enforceable code – whether on the protocol level or on the application level – standardizes transaction rules, thus reducing the transaction costs of: reaching an agreement, formalization, and enforcement. A smart contract can formalize the relationships between people, institutions and the assets they own. The transaction rulesets (agreement) of the smart contract define the conditions – rights and obligations – to which the parties of a protocol or smart contract consent. It is often predefined, and agreement is reached by simple opt-in actions. This transaction rule set is formalized in digital form, in machine-readable code (formalization). These rights and obligations established in the smart contract can now be automatically executed by a computer or a network of computers as soon as the parties have come to an agreement and met the conditions of the agreement (enforcement).[24]

Smart Contracts are: [24]

- *Self-verifying*, when an event occurs the smart contract checks whether a condition is met.
- *Self-executing*, when an event occurs and a condition is met, the smart contract takes an action as consequence.
- *Tamper resistant*, since it is based on blockchains public verifiability and transparency ensure it.

In the example of the "quality-credit" trading scheme, smart contracts allow to automate, among entities that do not trust each other, the control system of quality parameters, the reception of initial bad credits and the purchase of the remaining credits or the possible sale, the choice of the sale price of the credits.

The figure (4.2) shows how a smart contract would work in a car sale among two entities that don't trust each other and it gives an idea on how a smart contract works.

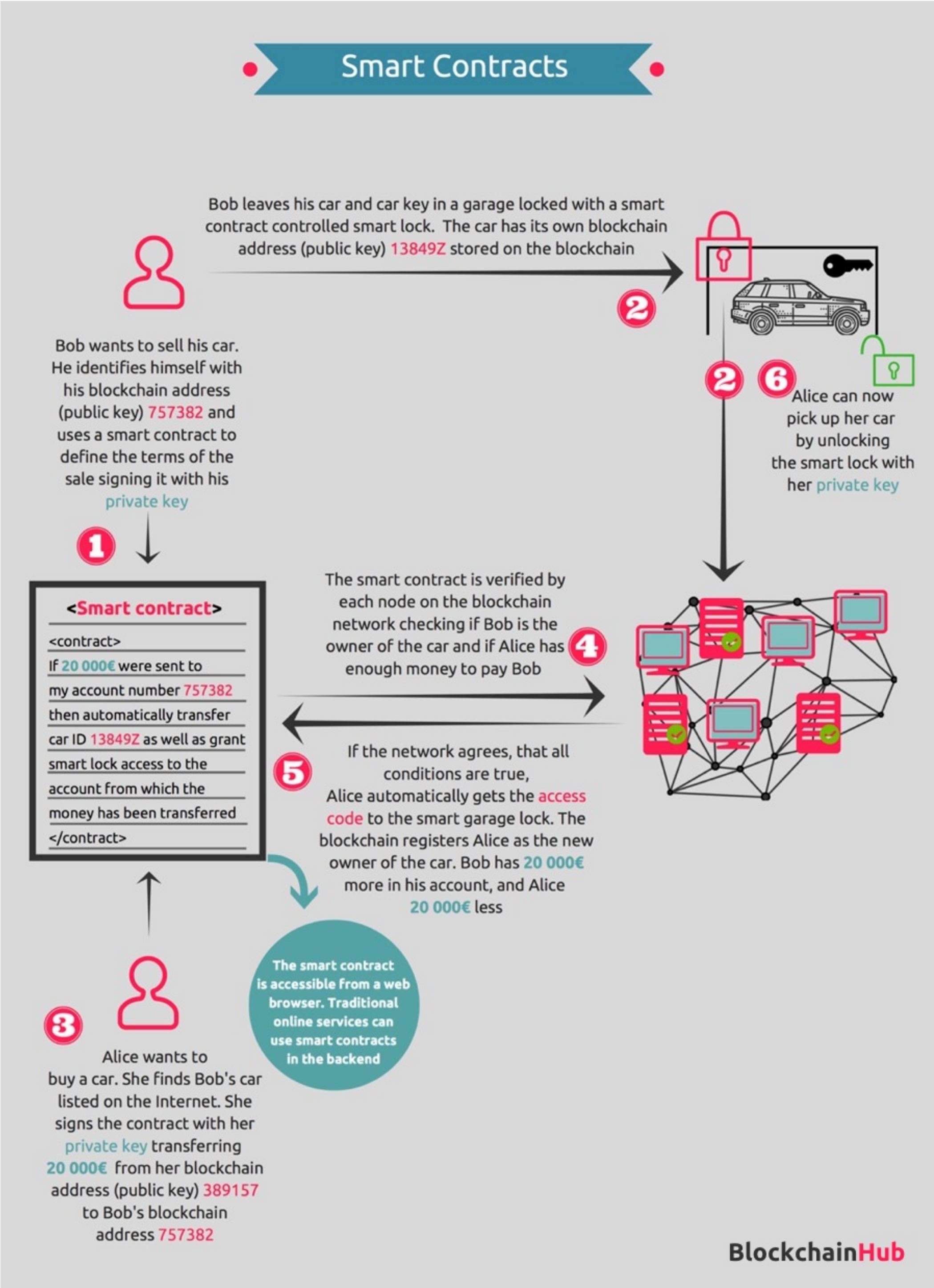

Figure 4.2. Smart contract example

Chapter 5

Conclusion and future works

In this thesis, an overview has been given of how wastewater data is collected, transmitted, stored and finally processed by analyzing the various needs and limitations depending on the main actor that exercises the monitoring. If there is the will to create a monitoring system open to all, which can be a significant advantage, this is possible with the use of low-cost sensors, a transmission of data with low energy consumption and long range as LPWAN where finally the data is stored in blockchain that renouncing performance advantages offers an excellent solution in trustless systems. All this approach has limitations in terms of measurement accuracy and safety of stations whose firmware may be corrupted to falsify the detected data.

The use of the blockchain, moreover, is perfectly suited to the possibility of developing smart contracts that manage a "quality-credit" market that offers not only economic incentives for the participating entities but also generates an advantage for the environment where the average quality of water increases over time.

Note also that the application of smart contracts, for the creation of the market scheme, can also be exploited by a scenario not open to all but where the main players are the purifier companies because it makes sense to use a permissioned blockchain to store data that, in addition, supports smart contracts.

Note also that the application of smart contracts, for the creation of the market scheme, can also be exploited by a scenario not open to all but where the main players are the purifier companies because it makes sense to use a permissioned blockchain to store data that, in addition, supports smart contracts.

Other works will follow this thesis, just after another revision of the technological requirements, like the development of some open-source prototypes both Arduino based, which is very popular and easy-to-use technology but that has lacks in power efficiency and security, and STMicroelettroncics based, which is less popular and easy-to-use but still guarantees a Do-It-Yourself approach and professional performances in terms of power efficiency and security.

It follows an evaluation phase where experiments on power supply of the prototypes are realized, analyzing in detail the consumption during the sleep phase, encryption phase, sensor data collection phase and transmit phase, and security issues. It will end with a real-world deployment, realizing an experience of the duration of a month or a couple of months.

On the other hand, the server should satisfy some technical requirements: developing a blockchain where the makers write, as it is a permissionless blockchain, the scalability of the system must be studied in an appropriate way and then perform stress tests on the system to see how it reacts; creating smart contracts that execute an algorithm that maximizes the profits of both players in the Stackelberg game illustrated in Chapter 4.

Bibliography

[1] C. Taylor, J. Yahner, D. Jones, A. Dunn, *"On-Site Wastewater Disposal and Public Health"*, Indiana State Department of Health, Summer 1996.

[2] NRDC, *Water Pollution: Everything You Need to Know*, May 14, 2018, https://www.nrdc.org/stories/water-pollution-everything-you-need-know

[3] *Legambiente*, https://www.legambiente.it/

[4] *Press release relating "Goletta Verde" mission 2018,* https://www.legambiente.it/sites/default/files/docs/comunicati/0813golettaverdefinale2018.pdf

[5] *HOW SEWAGE POLLUTION ENDS UP IN RIVERS*, American Rivers, https://www.americanrivers.org/threats-solutions/clean-water/sewage-pollution/

[6] *EU's top court fines Italy for failure to treat sewage*, https://www.thelocal.it/20180531/eus-top-court-fines-italy-for-failure-to-treat-sewage

[7] Kansas State University, *Freshwater Pollution Costs US At Least $4.3 Billion A Year*, November 2008, https://www.sciencedaily.com/releases/2008/11/081112124418.htm

[8] United Nation, *UN World Water Development Report, Wastewater: The Untapped Resource*, 2017.

[9] EPA (United States Environment Protection Agency), *Online Source Water Quality Monitoring*, September 2016.

[10] Ferran Adelantado, Xavier Vilajosana, Pere Tuset-Peiro, Borja Martinez, Joan Melià-Seguí, Thomas Watteyne, *Understanding the Limits of LoRaWAN*, January 2017.

[11] LoRa Alliance, *LoRaWAN For Developers*. www.lora-alliance.org, November 2018.

[12] The Things Network, *LoRaWAN Overview*, https://www.thethingsnetwork.org/docs/lorawan/, December 2018.

[13] The Things Network, *The Things Network*, https://www.thethingsnetwork.org/, December 2018.

[14] Chris Echard, *Ensuring Software Integrity in IoT Devices*, 2017

[15] Karl Wust, Arthur Gervais, *Do you need a blockchain?*

[16] Blockchain Hub, https://blockchainhub.net/

[17] Satoshi Nakatomo, *Bitcoin: a peer-to-peer electronic cash system*, 2009.

[18] Gavin Wood, *Ethereum: A secure decentralized generalized transaction ledger*, 2014

[19] Linux Foundation, *Hyperledger*, https://www.hyperledger.org/, December 2018

[20] Hoeller, P.; M. Wallin (1991). *OECD Economic Studies No. 17, Autumn 1991. Energy Prices, Taxes and Carbon Monoxide Emissions.*

[21] Zhaofu Hong, Chengbin Chu, Linda L. Zhang, Yugang Yu, *Optimazing an emission trading scheme for local governments: A Stackelberg game mode and hybrid algorithm*,2017.

[22] D. Krass, T. Nedorezov, A. Ovchinnikov, *Environmental taxes and the choice of green technology*, 2013

[23] Gheorghe Duca, Elena Bunduchi, Viorica Gladchi, Lidia Romanciuc, Nelli Goreaceva, *ESTIMATION OF THE NATURAL WATER SELF-PURIFICATION CAPACITY FROM THE KINETIC STANDPOINT*, 2008.

[24] Blockchain Hub, *Smart Contracts*, https://blockchainhub.net/smart-contracts/, January 2019.

[25] Ramon Sanchez-Iborra 1, Jesús Sánchez-Gómez, Salvador Pérez, Pedro J. Fernández, José Santa 1OrcID, José L. Hernández-Ramos, Antonio F. Skarmeta, *Enhancing LoRaWAN Security through a Lightweight and Authenticated Key Management Approach*

[26] Kaufman, C.; Ho□man, P.; Nir, Y.; Eronen, P.; Kivinen, T. *Internet Key Exchange Protocol Version 2*, Accessed 2019

[27] Rescorla, E.; Modadugu, N. *Datagram Transport Layer Security Version 1.2.*, Accessed 2019

[28] Schaad, J. *CBOR Object Signing and Encryption (COSE)*, Accessed 2019.

[29] Dimitrios Amaxilatis and Ioannis Chatzigiannakis, *Design and Analysis of Adaptive Hierarchical Low-Power Long-Range Networks*, 2018.