2016 SBESC – Intel® Embedded System Competition

Final Report

Project Name: HydroSys - An online platform for remote monitoring of water distribution networks

Students:	Andréa Brandão Duque, Larissa Lages de Oliveira, Rodolfo Ivo Santos de Andrade
Professor:	Edna Natividade da Silva Barros
University:	Federal University of Pernambuco - UFPE
JEMS ID:	164136

Declaration of Originality

We hereby declare that this thesis and the work reported herein was composed and originated entirely by ourselves. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given in the references.

Professor's Signature: ____________________

____

Name (in Block Letters): ____________________

____

Date: ____________________

HydroSys - An online platform for remote monitoring of water distribution networks

Abstract

Water distribution networks have a water loss rate that exceeds 40% in some Brazilian regions. Most current leak detection methods in Brazilian water distribution networks rely on inspections scheduled by water companies to ensure good conditions of the pipes and to investigate leakage complaints. These manual inspections require from the distribution company employees displacement, which costs time. The proposed system HydroSys measures flow water data in the water pipes and analyzes the data collected, monitoring the leakage occurrence and the network distribution infrastructure remotely from an online platform. Fastening the detection of water loss can decrease the high waste of potable water, contributing to a sustainable world.

Keywords: Water Distribution Networks, Water Leak Detection, Remote Monitoring.

Contents

1 Introduction

1.1 Motivation

Water is the main fluid constituent of most living organisms, vital to most life forms. Farming accounts for around 70% of water used in the world today, responsible for producing the food we eat. It is also used in many industrial processes such as cooling and washing, producing objects that we use on a daily basis, among other goods. Most of the water on planet Earth comes from the world ocean's saline seawater, while freshwater accounts for only 2.5% of the total.

Access to potable water relies on political, geographic, climatic factors and others, but mostly on availability of natural resources. Water scarcity is among the main problems to be faced by our society. According to the UN-Water [13], around 1.2 billion people, or almost one-fifth of the world's population, live in areas of physical scarcity of water. Another 1.6 billion people, or almost one-quarter of the world's population, face economic water shortage, where countries lack the necessary infrastructure to take water from rivers and aquifers. Although there are many areas facing water scarcity, the water use has been growing at more than twice the rate of population increase in the last century [13]. In such context, it is important to come out with solutions which will increase the quantity of fresh water for human consumption.

One of the major challenges facing water distribution is the high level of water loss in distribution networks. Part of the water leaving the treatment station, ready for consumption, does not reach the consumers, and part of the consumed water is not billed by the water company. If a large proportion of water that is supplied is lost, meeting consumer demands is much more difficult. Since this water yields no revenue, heavy losses also make it more difficult to keep water tariffs at a reasonable and affordable level. [4] The adoption of technologies to control losses in water distribution systems is then a necessity of the water companies. [12]

The average water loss rate in the European Union (EU) countries is about 20 %, whereas several countries have water loss rates lower than 10 % [14]. According to the National System of Sanitation Information (SNIS) from Brazil, the Brazilian estimated rate of water loss is 37%, but in some regions, like the north and northeast exceeds 40%. The rate of COMPESA (Companhia Pernambucana de Abastecimento) was 57% in 2010 [11]. COMPESA is the main water distribution company in Pernambuco, a state from the Northeast of Brazil.

The high rate of water loss in Brazil is mainly caused by error measures (defective and uncalibrated hydrometers, frauds), illegal connections and leakage in operational units of the distribution system, mainly in distribution network pipes. The leakage in the pipes is associated with age and quality of the material of the pipe and the absence of monitoring loss mechanisms.

1.2 Challenges

A big challenge is to provide a scalable system, one which can be replicated to monitor several different district meter areas (DMA). A DMA is defined as a discrete area of distribution caused by the closure of valves, in which the quantity of water entering and leaving the area is metered. A permanently monitored DMA is considered to be the most effective tool for reducing the duration of unreported leakage. Monitored DMAs must not be too large to maintain the water data variation within a tolerable range [5].

1.3 Objectives and Results

Most of the methods which detect leaks in Brazil rely on scheduled inspections by the water company. The priority of inspections is in areas of leakage complaints, made by phone calls from the population. In these inspections, the condition of the pipes and leakage complaints are verified, and the repairs are made. In contrast, there has not been much development of monitoring systems capable of detecting anomalous events based on water data analysis, like flow and pressure. It is then necessary the development of a wireless sensor network and a data analysis mechanism of the water data distribution collected by these sensors for remote monitoring of the infrastructure of the water distribution network [3]. The main objective of our proposed system project HydroSys is to provide online monitoring of pipes from a water distribution network, fastening the detection of water loss, therefore, decreasing the waste of potable water.

HydroSys is a system to be installed on pipes, which measures water flow. The data is sent to a gateway capable of analyzing the data and learning a pattern, and then detecting anomalies which indicate a water leak. The result of the data analysis is sent to a web page. In this page, it is possible to visualize the location of all sensor nodes and the pipes on a city map and the analysis results for each area, helping the water operators to guarantee the consumers good conditions of the infrastructure of the distribution network.

1.4 Similar Works

Similar works have been developed in places where water is scarce and expensive, like the system from the company TaKaDu in Israel. Of Israel's total water demand (2.2 billion cubic meters a year), less than one-tenth is supplied by freshwater sources such as the Sea of Galilee. The rest comes from water recycling and from desalination, both are expensive processes [7].

2 Project Description

2.1 Overview

HydroSys is composed of three main modules (Fig. 1):

- The data capture modules, which are the sensor nodes;
- The leaks predictor modules (gateways), which are responsible for controlling a group of nodes and analyzing water data;
- The online monitoring platform, in which is possible to visualize the location in a city map of the gateways and its nodes, the measured flow data by the nodes and the data analysis of the gateways.

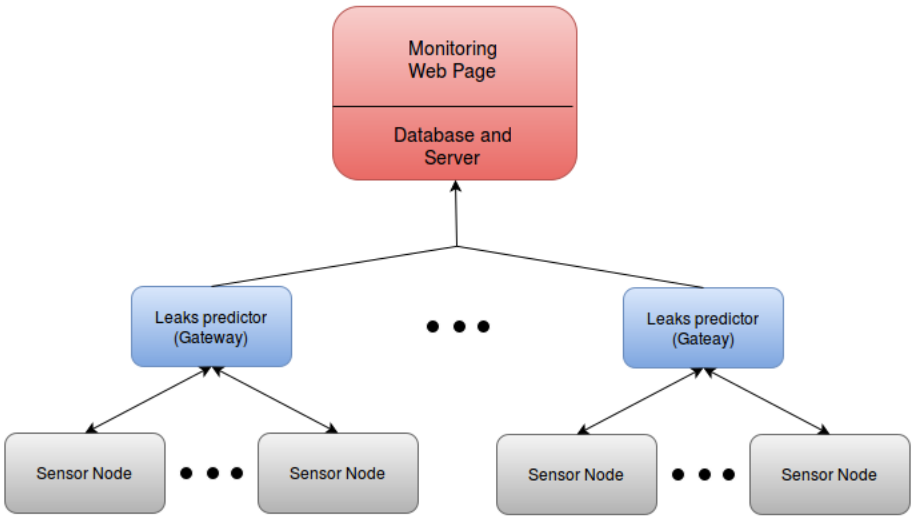


Figure 1: The HydroSys system architecture. Leaks predictor module communicates by radio frequency (RF) with a group of sensor nodes requiring flow data. After data analysis, it communicates with a server by GPRS, sending a sample of the received data flow from the nodes and the result of the analysis. If a leak is detected, a push notification alerts the water operator of the event.

The leaks predictor module reduces the quantity of information to be sent to the internet, lowering the communication cost and not overloading the server.

2.2 Data Capture Module

Measuring water flow data in the pipe that it is located during a period of less flow data variation allows a statistical analysis of the data. This period is called the minimum night flow period (MNF). To estimate real losses, MNF can be an indicator of distribution leakage and consumer wastage. MNF is the measured flow into a DMA of a network during the period with minimum demand, that is, between 1:00 am and 4:00 am. Analysis of DMA flows allows estimation of leakage when the flow into the DMA is at its minimum, which occurs at night when customer demand is at its lowest level, and therefore, leakage component is at its highest value. [5] When a new request is received , the sensor node sends the data by radio frequency (RF) to the leak predictor. While it's not measuring flow, the sensor is on standby to save power (Fig. 2).

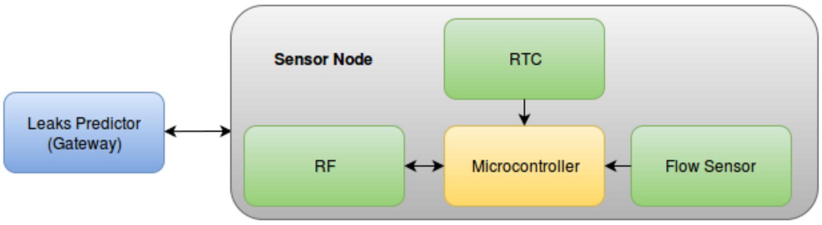


Figure 2: Each sensor node is a microcontroller which controls a radio frequency module and a flow sensor connected to the pipe. When a new request from the leak predictor arrives, the node sends the collected flow data by RF. The RTC generates an alarm interruption to awake the module from stand-by.

2.3 Leak Predictor Module (Gateway)

- Controls the sensor nodes, requesting water flow data;
- For each group of collected data of each sensor, the module performs a statistical analysis with training data and test data, detecting if there is a leak for each sensor. The training data are the data available under normal condition of the system (without leakage), the test data is the data to be analyzed;
- A sample of the data and the analysis result is sent to the web server by GPRS. This technology was chosen so the system may work in locations which do not have an internet connection by Wifi (Fig. 3).

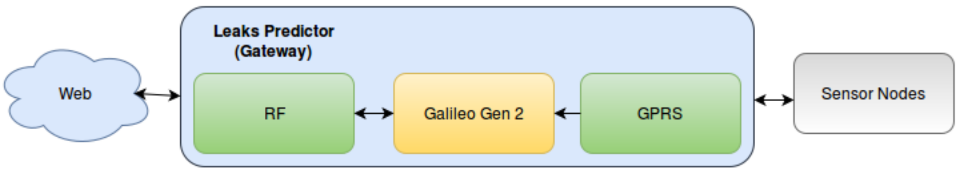


Figure 3: Gateway requests data to sensor nodes. The module recognizes if the test data differ in some respect from the training data. The leak predictor sends those results and sample of the received data to server

2.4 Online Monitoring Platform

A website was designed for easy monitoring of a water distribution infrastructure and for water data visualization. It is an interface that allows the administration of all working sensors and stations and the visualization of the data analysis.

For system administration purposes, the website allows registering leak predictor modules and its nodes. Address of the leak predictor and the nodes are required for the register, also the connections between the nodes. After the register is done, it is possible to visualize in a map points indicating the gateways, and when a gateway is selected, another map opens showing all the nodes a selected gateway is responsible for, also indicated as points, and the pipeline infrastructure.

The website allows the user to be notified as fast as possible of a leakage occurrence. If a leakage is detected, the gateway and the node that detected the leak turn red and a push notification is generated.

A priority for each leakage was calculated for allowing comparison of all occurred leakages by ordering them in a priority list.

The training and test flow dataset for each node are visualized when clicking on the node.

3 Technical description

3.1 Implementation

3.1.1 Data Capture Module

Microcontroller

The microcontroller we used to implement the sensor nodes was the Arduino Uno. It controls the flow sensor, the wireless transceiver and the RTC module.

The Arduino sends to the gateway the data flow it reads from the flow sensor whenever it receives a data solicitation from the gateway. The board is on stand by during the day period, when the water flow patterns are unpredictable due to consumer use hence statistical analysis is difficult. Because of the standby mode, the microcontroller can only be woken up by an external interrupt, for this purpose we used an RTC module. The RTC generates an alarm interruption to wake up the Arduino from standby for the configured night period. The state diagram of the data capture module can be seen in Figure 4.

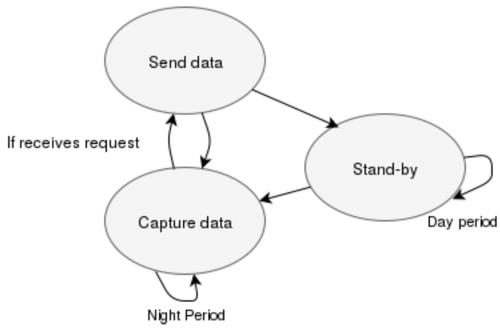


Figure 4: Microcontroller state machine. Arduino waits until receives a request from Galileo. When it receives, reads data from the flow sensor and sends them to Galileo for leak analysis.

Radiofrequency transceiver

We used the wireless RF transceiver module nRF24L01+ to implement the communication between sensor nodes and the gateway. The module operates in 2.4Ghz ISM band and the data transmission frequency is 250Kps. In case the packets delivery fails, the module has up to 15 tries to resend the data. It uses GFSK modulation and the communication protocol Enhanced ShockBurst^{TM}. This protocol supports sending and receiving packets using a buffer, recognition of package and automatic re-transmission. The maximum range distance according to datasheet is 1100m [7]. The channel number 123 is used because is above most wifi channels.

Flow Sensor

This sensor is connected in a pipe, in line with the water line. It contains a pinwheel sensor to measure how much liquid has moved through it. There's an integrated magnetic hall effect sensor that outputs an electrical pulse with every revolution. The tension generates an interruption in the microcontroller, which has a pulse counter. The flow is calculated in liters per hour by the following method: each liter per minutes the sensor generates approximately 7.5 pulses. Dividing the total pulses by 7.5 we obtain the total of liters per minutes, then dividing by 60, results in a total of liter per hour. The measurement error of the sensor is approximately 3%. The maximum flow rate is 60L/min and the maximum supported pressure is of 1.75MPa [8].

3.1.2 Leaks Predictor Module

Intel Galileo Gen 2 (Gateway)

The board controls the GPRS shield SIM900 from Geeetech to communicate with the web server via posts and gets, and the radiofrequency module nRF24L01+ to communicate with the sensor nodes. The Linux distribution used was Yocto, and the image was the IoT-DevKit, pre-compiled by Intel. We used the Python version 2.7 present in the image and the Python's package 'statistics'.

Initially the leak predictor running in the Galileo checks if the GPRS module is working and tests the communication with the nodes. In case it fails to a node at the end of the night, the board sends to the server which node has a communication problem, alerting the user in the website.

After testing communication, it requests data flow to all nodes it is assigned to and the implemented python script will analyze the data pattern for each node, comparing flow samples to the training data. The first time the system executes, it will capture the training data. After this set-up state, the Galileo will continuously request test samples and execute the script, unless it receives a command from the website to retrain the system. In this case, the board will request for data flow and will replace the old training set with the newly received data. This functionality is essential after structural changes in the monitored pipeline, for example when a new connection is added. In this cases, the flow changes, but it is possible to restart the system without restarting the board. Furthermore, the gateway can also receive a command from website to change the frequency of requests that it sends to the nodes.

After the statistical script executes, it will send the results and a sample of the dataset flow that was used in the data analysis to the server. The complete flow dataset was not sent to avoid network congestion. Galileo also sends to the server an approximation of the rate of water lost if a leakage is occurring. This approximation is done calculating the flow difference between the values from the monitored node and the preceding node.

The Galileo performs tests with several test samples it requests during the night period. When the period ends, it sends a command to the nodes putting them on standby. The state diagram of the leak predictor module states is detailed in figure 5.

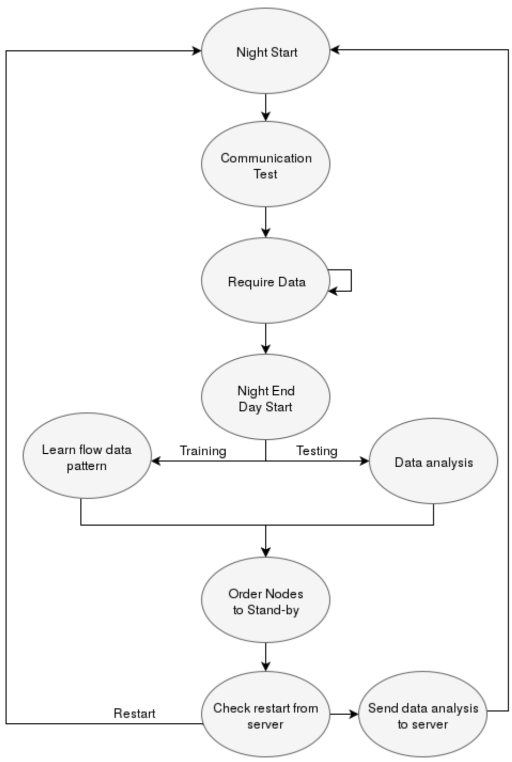


Figure 5: Leak predictor state machine.

GPRS shield

The GPRS Shield is based on SIM900 module from SIMCOM. It has low energy consumption because it enters on standby.

Wireless Transceiver

The transceiver used in the gateway is the same used in the nodes, already described in section 3.1.1.

3.1.3 Data Analysis

The problem of finding leaks using flow values was approached in this work as novelty detection, aiming one-class classification. In this type of classification, one class (the specified normal, "no leakage") has to be distinguished from all other possibilities. The goal is to learn a model of normality from a set of data that is considered "normal", the training set, and the decision process on test data is based on the use of this model [8]. This approach was chosen because leakage data is not available initially and may have unpredictable patterns. Therefore we only collect training data from pipes that are not leaking.

The algorithm compares the training distribution with test distribution to see if they are similar or different. If they are different, a leakage is occurring. If they are similar, then there should be no significant difference between the means from the collected training and test samples, thus a leakage is not occurring.

To aim increase hit rate we implemented three statistical methods to analyze data. These three methods are classifiers which output "leakage" or "no leakage". The outputs of these methods are the input to a voter which chooses the predominant class between these three outputs. The chosen class is the final output of the system. The three statistical methods used were Z Test, T Test and Normality Test. The algorithm is represented in figure 6.

Z test is a statistical test which approximates the distribution by a normal distribution.

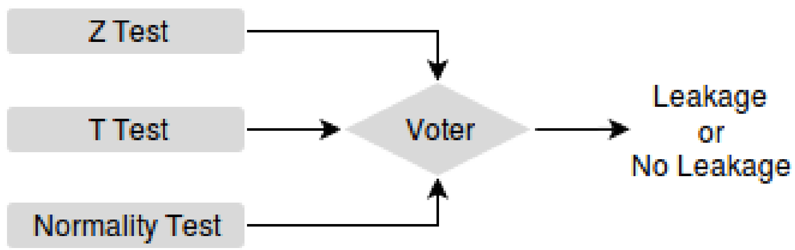


Figure 6: Data analysis algorithm.

We assumed that because of the central limit theorem, which establishes that, for the most commonly studied scenarios, when independent variables are added, their sum tends toward a normal distribution, even if the original variables themselves are not normally distributed. The Z test is useful when it is necessary to do comparisons over time with time series data. We compute the Z statistic, for comparing the two samples (training and test), in the following manner:

$$Z|T = \frac{|\bar{X}_{Test} - \bar{X}_{Training}|}{\sqrt{\sigma^2_{Xtest} + \sigma^2_{Xtraining}}} \tag{1}$$

$$\sigma_X = \frac{\sigma}{\sqrt{n}} \tag{2}$$

$\bar{X}_{Test}$ is the mean value of sample test, $\bar{X}_{Training}$ is the mean value of sample training, σ_{Xtest} is the standard deviation of sample test divided by the square root of the number of data points, and $\sigma_{Xtraining}$ is the same as σ_{Xtest} but for sample training. We empirically found a threshold for which Z values above the threshold perform leaks detection well in our prototype. Above this threshold, training and test samples are significantly different.

The T test used is similar to Z test, but in this case is made a hypothesis test. When there is not a leakage, the sample test and the sample training belong to the same population. Because of that our null hypothesis for the hypothesis test is that the population's average of the two samples are equal, which means there is no leakage. The T calculation is the same as the Z calculation, showed in equation 1. The difference between them is the fact that we compare the calculated value of T with a value called the critical value of the T distribution. The critical value is found in the T-Student Table. If the T value is greater than the critical value the null hypothesis is rejected, therefore a leakage is occurring.

The third method is a method based on normal distribution. We can assume that the flow follows a normal distribution because of the central limit theorem. This theorem states that when the sample size increases, the sampling distribution of their average approaches increasingly a normal distribution. Our system worked throughout all the training period with a high sampling rate, so the samples taken will approach a normal distribution. In this method we verified if the data captured is inside the interval $[\mu - 3\sigma, \mu + 3\sigma]$. If most part of the data is out of range, then the method classifies as a leak.

3.1.4 Online Monitoring Platform

The Online Monitoring platform was developed for users who are responsible for controlling water losses of a water distributed network. Typical users would be governmental water company operators or private pipelines owners.

The index of this website has two buttons (www.hydrosys.org) (Fig. 7). The login button opens a new window asking for user name and password (Fig. 8). After the user writes the required information and clicks on login button, the login window closes and the index page is redirected to the online monitoring platform (www.hydrosys.org/page1.html) (Fig. 9). The second button "Stop the Leaks" takes the user to a public page (www.hydrosys.org/page.html) (Fig. 10). This page has information about the project and its developers. Also, on this page, the general public are able to inform a leakage that is happening.

The platform is composed of seven tabs: "Map", "Search", "Station", "Database", "Setup", "Leakage", "Consumer Notifications", which present different features described in the following sections.

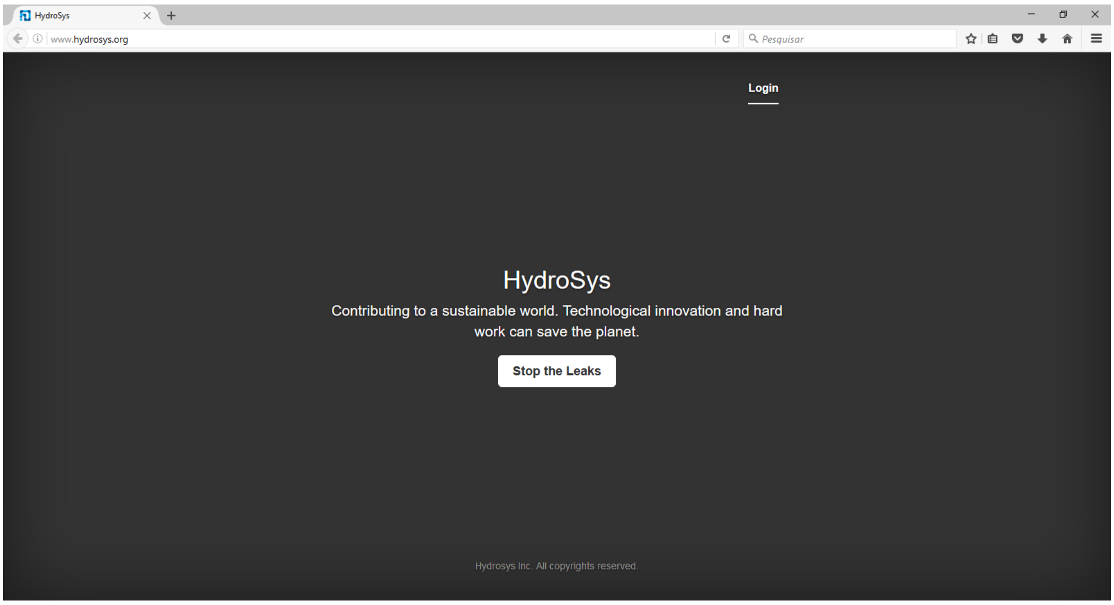


Figure 7: HydroSys index page, login button at the top-right corner and link to HydroSys public page at the center.

Map

The "map" tab is the initial tab. It loads a city map with color filled points indicating the geographic location of the gateways (Fig. 9). The point is red in case there is a leakage detection in one of the gateway's nodes or blue if there is no leakage. When a gateway is selected, clicking on the dot in the map, another map loads on the right side of the previous map (Fig. 9), showing all the nodes a selected gateway is responsible for, also indicated as dots, and the pipeline infrastructure. These nodes also change color, they are blue when there is no leakage nearby or red when there is. Also, the pipeline infrastructure can be viewed in this map. The

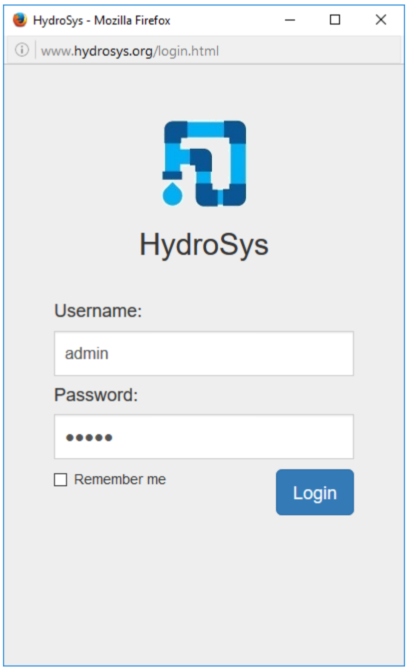


Figure 8: HydroSys login page, user name and password already fulfilled.

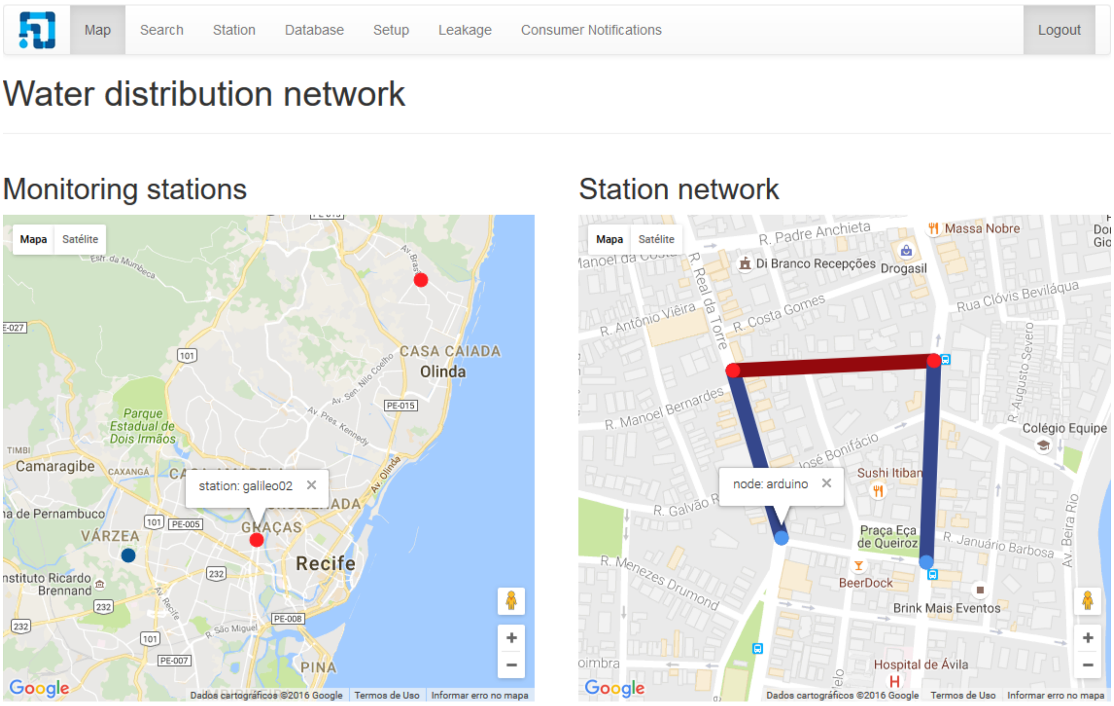


Figure 9: Map tab from HydroSys website. The first map shows points, which represent all the gateways registered in the website. The second map shows all the nodes the selected gateway is responsible for, also represented as points. The pipelines are the connection between the nodes, represented in the map as lines.

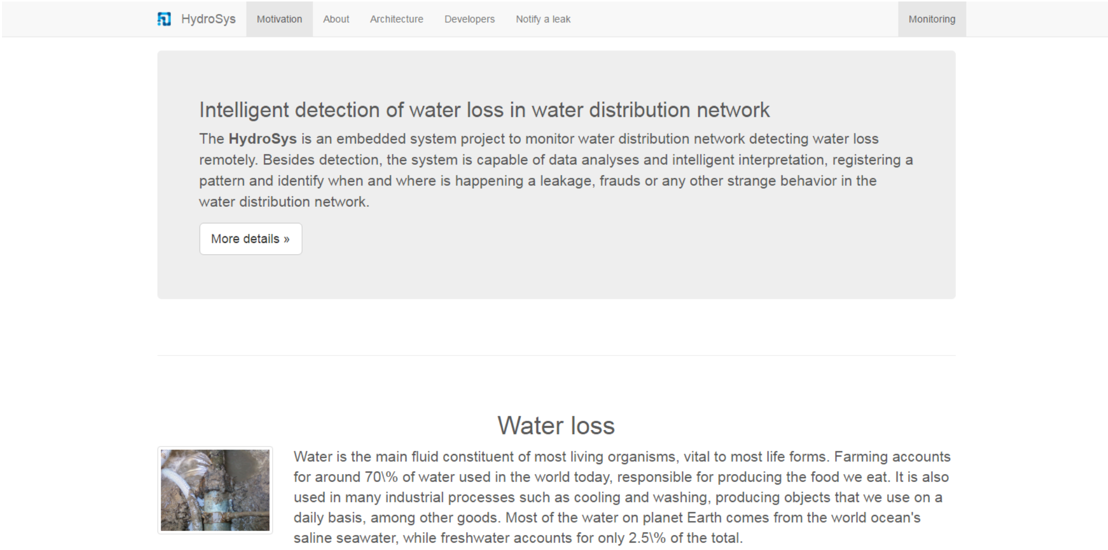


Figure 10: HydroSys public page shows project introduction.

color of the pipeline that connects two nodes is red when these two nodes have a leakage otherwise the color is blue. The map visualization was based in a hydraulic simulator program for water distribution system EPANET [10].

Clicking on a node, the following information shows below both maps: station name, leakage value in liters per hour, node name and the time the flow data sample was sent to the database. Next to the node name there is a label informing the node status, which can be:

- "This node is working properly", if there is no problem with the node;
- "There is a leakage near this node!", if a leakage has been detected by this node;
- Water distribution interrupted or sensor is defective" if flow data captured by this node is zero.

Below this label, there are two different line graphs for data visualization of the flow samples, one graph below another (Fig. 11). This data is a sequence of discrete-time data. There is a set of different samples in the server, each was received in its own timestamp. The user can select one sample in a drop-down menu between all the received samples, which are ordered by the timestamp. The y-axis represents the flow data in liter per hour, while the x-axis represents the ordering of the observations. The graphs are iterative, when the mouse hovers on the graph, the hovered point shows at the top-right corner the numeric value of the flow. The first graph shows the node training data in blue line and test data in red line. Below this graph, there is the second line graph comparing the test data for all station nodes in different line colors.

Finally, below the second graph there is information about the pipeline material where the node was installed: Installation date, Next Revision Date, Brand, Material type, Nominal Pressure, Joint Type, Water Type, Color, Internal Diameter, Thickness, Technical Standard, Telephone Assistance and Fabrication Date.

Search

It is possible to search in this tab a specific station by its location or by ID (Fig. 12). The station search by ID is located below the search by location. In location search, the user has to inform the Country, State, City and Neighborhood of the desired station. In ID search, only the ID station is necessary. In both search types, there are two buttons, "Search" and "Clear". When clicking on "Search", a table appears below the buttons. This table contains the stations that satisfy the search conditions. Each line of this table contains a different station, meanwhile the columns contain station data, such as ID, Name, Country, State, City, Neighborhood, Coordinates, SIM card number and Map. In each map column, there is a button, which opens the tab map and shows in the second map that station nodes. There is also the "Clear" button next to the "Search" one, which removes the table.

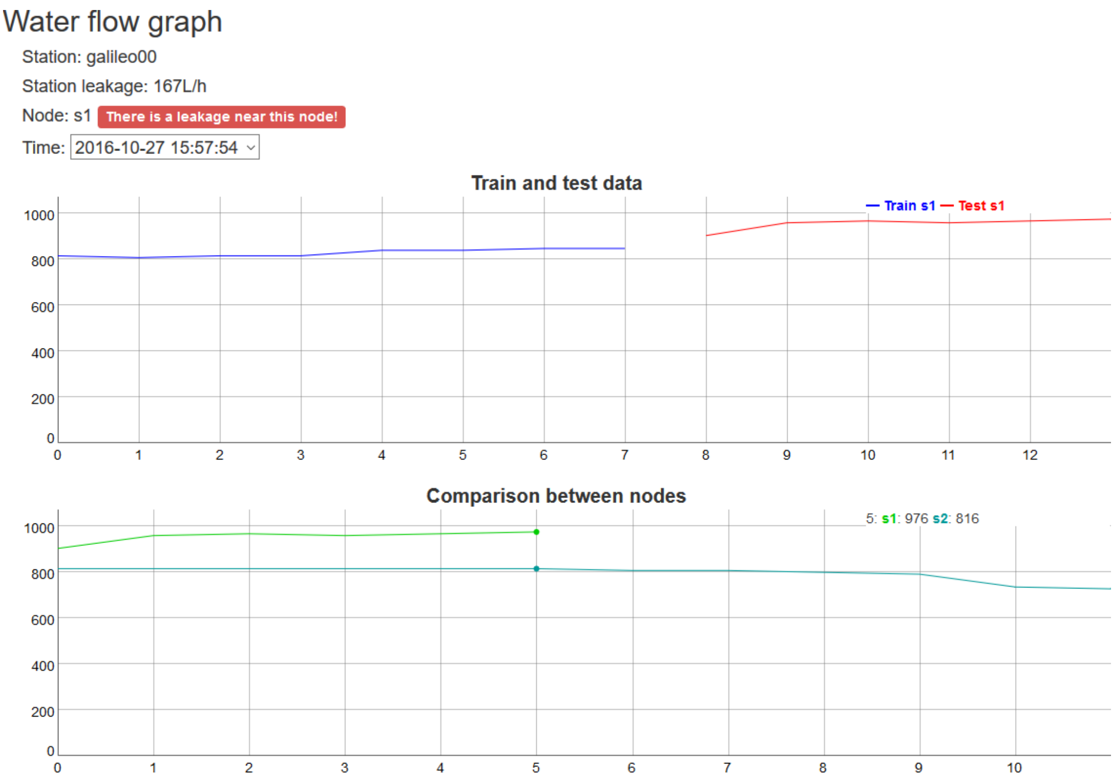


Figure 11: HydroSys page showing station name, station leakage, node label status, node name, data time and graphs showing information from a selected node. The first graph showing training and test data and the second graph showing a comparison between nodes from the same station. For this chosen node, a leakage was detected. There is a visible difference between train line and test data line in the first graph. The comparison graph shows a significant variation between s1 and s2 nodes from galileo00 station.

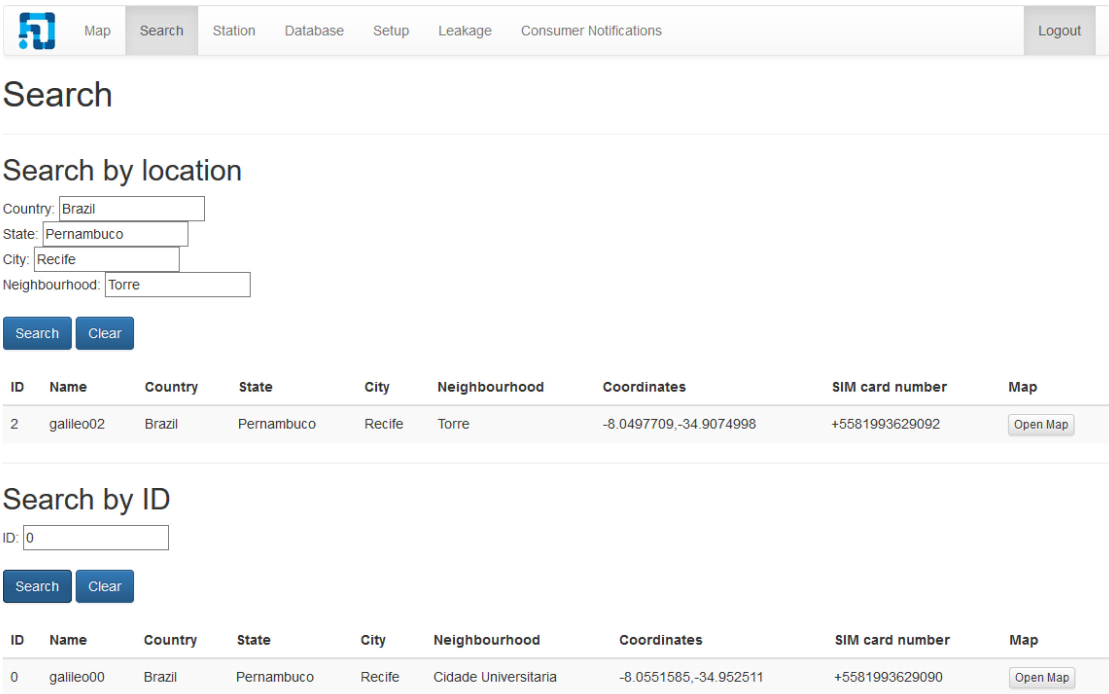


Figure 12: When search tab is selected in HydroSys monitoring page, the search by location and search by id forms are showed. This image shows the two forms fulfilled with its respective results below each one.

Station

It is possible to add a new station, remove a station, add and remove station nodes in the station tab. Each functionality has its own form. They are shown in HydroSys page in vertical order.

Adding a new station requires fulfilling the form "Add new station" with the following parameters: id, name, country, state, city, neighborhood, coordinates, the gateway SIM card number, the quantity of nodes, name of nodes, latitude and longitude of nodes, connection of nodes. After fulfilling the form with information of the new station to be added, and clicking on the "Submit" button, the station is added to the map at the map tab (Fig. 13).

To add a new node to a station requires the selection of the station in a drop down menu, which shows the set of all registered stations, the node name, its geographic coordinates and the station node which the new node will be connected. After that, the new node is added to the selected station clicking on submit (Fig. 14).

To remove a node, in the form "Remove node" the user must select the station, which contains the unwanted node, select the node name and click "Submit".

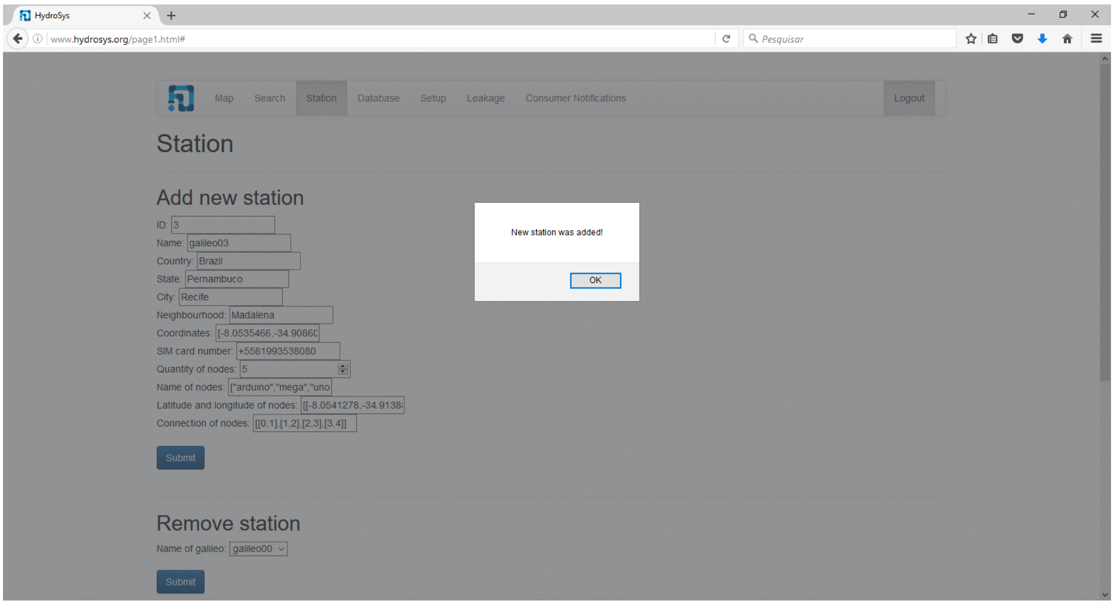


Figure 13: Station tab from HydroSys monitoring page. A confirmation message appears when a new station is added.

Setup

The tab setup is where the operator can send commands to the station gateway. (Fig. 15). At this tab, the operator selects the command and chooses the station that will receive it. There are two possible commands:

- "set capture period", which changes the frequency the data is requested by the gateway. The interval period between two requests is typed in milliseconds by the user in a text box on the left.
- "retrain", which commands the gateway to retrain the analysis.

Leakage

We defined the leakage priority for each station using the estimation already described in section 3.1.2. The gateways are listed by priority in descending order at a table similar to the search table (Fig. 16). The difference from the search table is that there is an extra column on the left informing the leakage value in liter per hour.

One of the water distribution operators problems is which leak repair first when there are a lot of leak reports. The objective of this tab is to solve this problem by providing the user an estimation of the leakage aggravation and a fast way to visualize it.

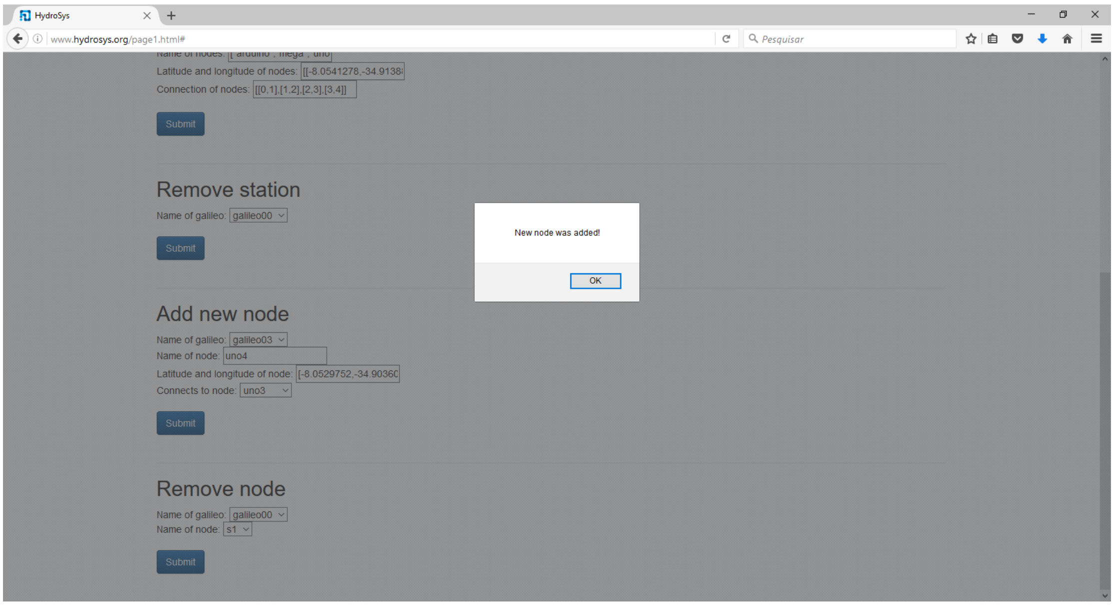


Figure 14: HydroSys Online Monitoring Platform shows station tab, a new node has been added and a alert message appears.

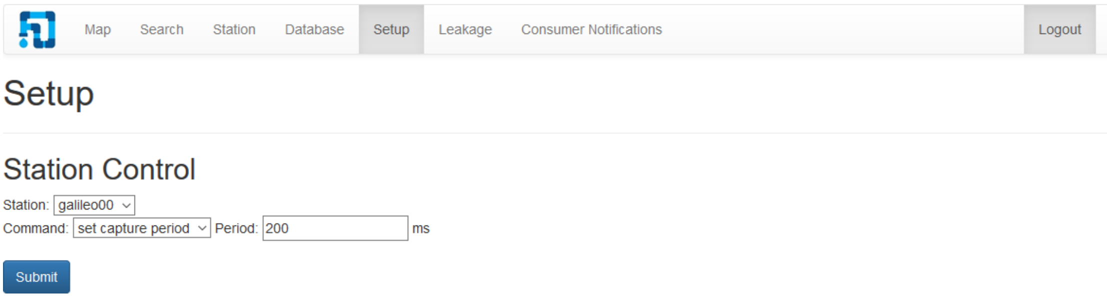
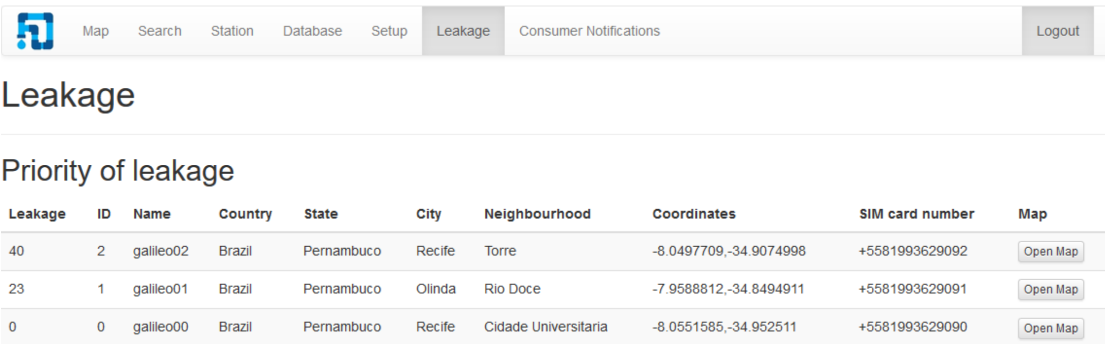


Figure 15: Setup tab from HydroSys page. The command selected sets the interval period in milliseconds to the galileo00 station.



Leakage	ID	Name	Country	State	City	Neighbourhood	Coordinates	SIM card number	Map
40	2	galileo02	Brazil	Pernambuco	Recife	Torre	-8.0497709,-34.9074998	+5581993629092	Open Map
23	1	galileo01	Brazil	Pernambuco	Olinda	Rio Doce	-7.9588812,-34.8494911	+5581993629091	Open Map
0	0	galileo00	Brazil	Pernambuco	Recife	Cidade Universitaria	-8.0551585,-34.952511	+5581993629090	Open Map

Figure 16: Priority tab from HydroSys page. A table of stations ranked by the highest leakage flow to the lowest is available at this tab.

Consumer Notification

Other feature implemented was the "Consumer Notification" tab (Fig. 17). This tab shows a table with the columns: city, neighborhood, address, reference point, road surface and image of the leakage. In this tab are listed all notified leakage by the public through the page (Fig. 10). This public page allows the public to send relevant leakage data. The public can upload a leakage photo to indicate the existence of a leakage, which can be visualized by the website operator.

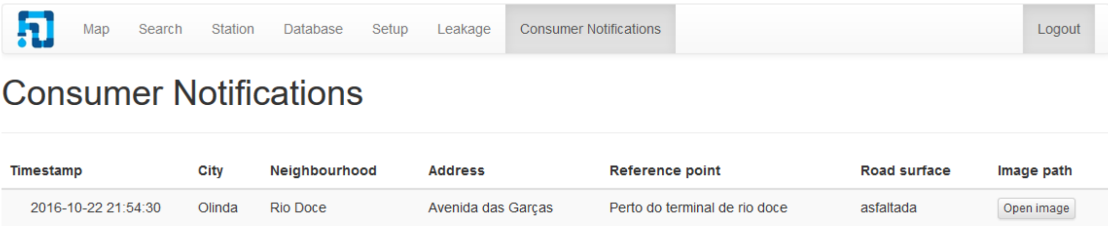


Figure 17: HydroSys Online Monitoring Platform shows Consumer notification tab, it shows a table with data sent by the public. It has a link to visualize a photo proving the existence of a leakage.

Database

At the database tab the user can visualize all the flow values stored in the database grouped by gateways.

Next to the tabs already mentioned, there is a logout button. When the user logs out, the page is redirected to the index page (Fig. 7).

3.1.5 Technologies

City Maps

The city maps were developed using the Google Maps API for JavaScript [1].

Flow data visualization

Both data flow graphics were created using Dygraphs, an open source JavaScript charting library [2].

Push Notification

The notification was generated by Pushetta, which is an API specialized in real-time notifications [9].

Other technologies

The website is hosted by hosting24.com [6]. The programming languages used to implement the features were JavaScript, PHP (server) and SQL (database).

4 Validation

In order to evaluate the proposed system, a prototype of a piping network was built (Fig. 18). A water pump continuously pumps from an open container and delivers the water back to it. The pipe used has 1.875 cm in diameter. We installed two sensor nodes in the pipe, which monitor the water flow continuously. The pipe between the two nodes represents the monitored pipe in our system. To simulate a leak, a tap was positioned between the nodes in order to deviate the path of water flow and generate a water flow variation. The amount of leaked water increases with the tap opening degree.

Fig. 19 shows the two nodes and the Galileo. A LCD (Liquid Crystal Display) was placed in Galileo for debugging purposes, showing the system state (training, testing, sleeping or sending data).

Figure 18: Validation prototype.

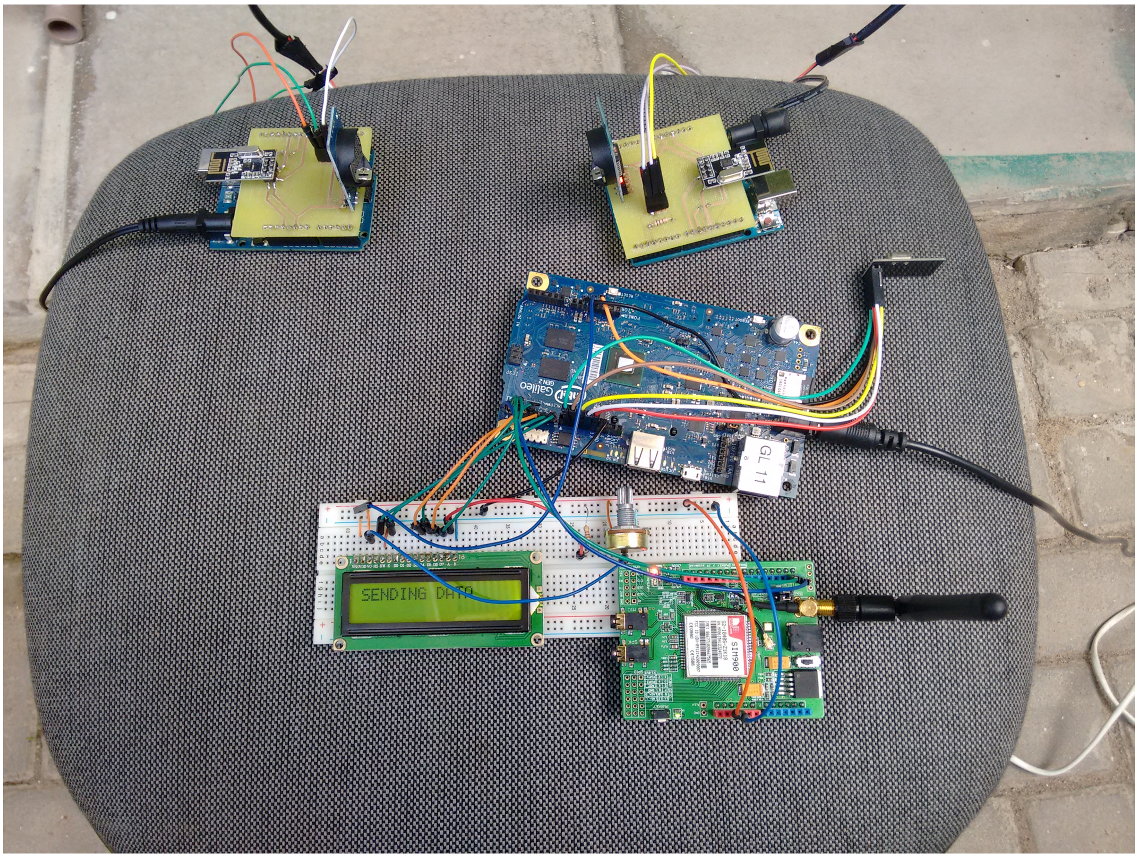


Figure 19: Nodes and Galileo.

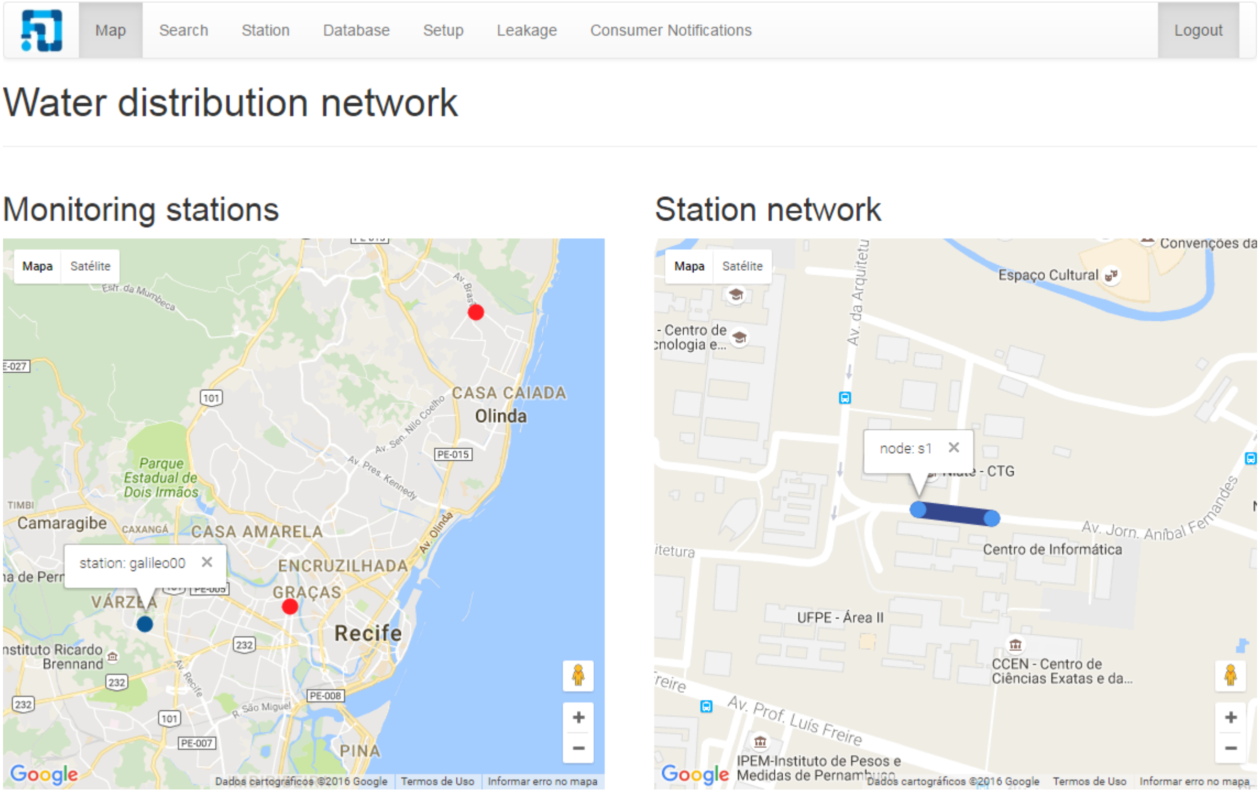


Figure 20: Website shows our simulation as a real station in the left map. In the right map, the location of its two node.

5 Performance Analysis

To evaluate the performance we measure the execution time of the data analysis module and the server arrival time of the data sent by the GPRS, at the end of the night period. The duration of the test, which is the duration of the night, was varied with the objective of analyzing how the system behaves with different amounts of data. The duration of the test was reduced in our prototype to allow a faster analysis.

The table 1 shows the data analysis runtime and the sending data time, for each test period. It also shows the total number of data points in a sample collected by the Gateway at this time.

Table 1: Time Analysis

Duration of the Training	Time Data Analysis	Time Send Data	Data points
2 min	0.01 s	4.44 s	97
5 min	0.01 s	5.34 s	307
15 min	0.01 s	4.70 s	1102

We can see from the table that the data analysis algorithm is very fast. Probably with the period of one night (12 hours), the runtime of this algorithm will increase, but it will continue to be fast enough for our purpose.

The difference of times in GSM is due to the fact that when there is no signal it tries to connect until timeout.

In order to save energy, at the period that the Gateway is not analyzing data, the board sends a command to the nodes to enter in a stand-by state. They are only awakened again when the night period begins. Running on 5V through the +5V pin, the microcontroller draws about 49 mA. On the stand-by mode, it draws 34.5 mA, a savings of approximately 30%.

6 Conclusion

Water distribution networks monitoring present many challenges because of networks complexity. It is necessary then that water companies have mechanisms to monitor their infrastructure in order to perform faster maintenance and repairs, decreasing the water wastage in the network, which is very high in many places.

The HydroSys is a project of an intelligent platform of distribution networks management which aims to help detecting leaks to reduce water losses and mapping pipelines to facilitate the infrastructure maintenance.

Among the challenges, the system must be scalable, covering the most part of the city with a low cost, which will be improved replacing the Arduino for cheaper microcontrollers.

In order to improve the mapping of infrastructure, the website provides other relevant information about the node as brand pipe, material, diameter, and supported pressure by the pipe.

The problem of water losses in the distribution networks and poor management of the network infrastructure is very serious in Brazil, and the HydroSys becomes then very relevant in this scenario.

7 Future Works

For future works, other methods of novelty detection could be used to further improve the performance of data analysis. Furthermore, a hydraulic simulator program, such as EPANET [10], can assist the construction of the normality model used in the novelty detection, improving the accuracy of the distribution used to represent the training dataset.

References

[1] Google Maps Javascript API. <https://developers.google.com/maps/documentation/javascript/?hl=pt-br>. *Access in*, October 2016.

[2] Dygraphs. <http://dygraphs.com/>. *Access in*, October 2016.

[3] Whittle Andrew J. et al. Waterwise@ sg: A testbed for continuous monitoring of the water distribution system in singapore. *Water Distribution System Analysis (WSDA)*, 2010.

[4] Frauendorfer, Rudolf, and Roland Liemberger. The issues and challenges of reducing non-revenue water. *Asian Development Bank*, 2010.

[5] García, Vicente J., Enrique Cabrera, and Enrique Cabrera Jr. The minimum night flow method revisited. *Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium.*, 2006.

[6] Hosting24.com. <https://www.hosting24.com/>. *Access in*, October 2016.

[7] Amanda Little. <http://www.bloomberg.com/news/articles/2015-01-08/takadu-helps-israel-be-a-most-efficient-water-manager>. *Access in*, August 2015.

[8] Marco A. F. Pimentel, David A. Clifton, Lei Clifton, and Lionel Tarassenko. A review of novelty detection. 2014.

[9] Pushetta. <http://www.pushetta.com/>. *Access in*, October 2016.

[10] Lewis A. Rossman. The epanet programmer's toolkit for analysis of water distribution systems. *US Environmental Protection Agency*, 2006.

[11] Danielle Dionisia Santos and Suzana Maria Gico Lima Montenegro. Evaluation of the methodology for control of water losses in distribution network in recife – pe. 2014.

[12] SNIS. <www.snis.gov.br/>. *Access in*, September 2016.

[13] United Nations Water. <http://www.un.org/waterforlifedecade/scarcity.shtml>. *Access in*, September 2016.

[14] Öztürk I, Uyak V, Çakmakci M, and Akça L. Dimension of water loss through distribution system and reduction methods in turkey. *International Congress on River Basin Management*, 2007.