Hydrocontrol IT: an attempt to mitigate water losses in Italy towns

Giuseppe Benvenuto[1], Salvatore Iiritano[1], Luigi Mirto[2], Pasquale Piccione[2], Massimiliano Ruffolo[1], Simone Vizza[1], Manuel Borroto[3] and Francesco Ricca[3]

[1]*Revelis S.r.l, Viale della Resistenza 19/C, 87036 Rende, Italy*
[2]*Polo ICT Pitagora, Largo Zinzi 18, 88100 Catanzaro, Italy*
[3]*University of Calabria, Via Pietro Bucci, 87036 Rende CS, Italy*

Abstract
Respecting and caring for the environment is an aspect that is becoming more important with each passing year. The damage caused by human beings to the planet has triggered a global climate change, favoring the creation of major weather phenomena, increasing the temperatures, and changing the rainfall cycles, which in turn make access to water more complicated. The effects mentioned above are only a part of the true impact of Climate Change, and taking this into account, the European Union has established the European Directive 20-20-20 in an attempt to try to slow it down and mitigate its impact on the planet. In this regard, Italy suffers greatly from water wastage due to leaks in the supply system and a practically non-existent culture of saving water. To address this problem, in this paper, we propose a system based on the Internet of Things (IoT) and Artificial Intelligence (AI) to monitor water consumption in real-time, as well as to detect and prevent possible failures in the water supply system, allowing to save water and economic resources due to maintenance. We evaluate the potential of our system by testing it in a real environment in the province of Catanzaro.

Keywords
Internet of Things, Climate Change, Clustering, Artificial Intelligence, Water Consumption

1. Introduction

Climate Change has caused severe economic and social damage worldwide due to the increase of temperatures, the alteration of rainfall cycles, droughts, very strong storms, among others. To reduce the impact of climate change, the European Directive 20-20-20 establishes the reduction of 20% of greenhouse gas emissions, the achievement of 20% of renewable energies, and the improvement of 20% of energy efficiency by 2020. Italy is going in the right direction for the first two objectives, but there is still much to do on the last point. Concerning the water network, Italy suffers from a serious problem of water wastage, linked both to factors of education in the use of resources by citizens and to leaks in the pipelines due to obsolescence and wear and tear of the pipes, as well as the malfunctioning of the meters. The problems of the distribution network also determine inefficiencies (in particular interruptions in the water supply), which in

AIxIA 2021 Discussion Papers
giuseppe.benvenuto@revelis.eu (G. Benvenuto); salvatore.iiritano@revelis.eu (S. Iiritano); luigi.mirto@dnalab.it (L. Mirto); pasquale.piccione@dnalab.it (P. Piccione); massimiliano.ruffolo@revelis.eu (M. Ruffolo); simone.vizza@revelis.eu (S. Vizza); manuel.borroto@unical.it (M. Borroto); francesco.ricca@unical.it (F. Ricca)

the south of the country occur three times more frequently than in northern regions.

The answer to these issues is provided by new IoT technologies, and in particular by smart meters, electronic devices that can measure electricity, water, and gas consumption and remotely intervene on plants by regulating the exchange of energy and information on their operation to avoid waste. The Italian Authority for Electricity, Gas, and Water announced the nationwide testing of multi-service smart meter solutions for about 60 thousand supply points in Turin, Reggio Emilia, Parma, Modena, Genoa, Verona, Bari, Salerno, Catania, and other smaller municipalities.

Intelligent consumption management is a new frontier in all fields, including the domestic one. In the next few years, it will be possible to control everything in the house with a simple touch on the screen of the smartphone, to program the use of water following a schedule, and to vary it according to certain atmospheric events, such as the humidity of the soil and the outside temperature, to avoid water waste.

In this paper, we discuss the development of *Hydrocontrol IT*, a platform that takes advantage of the development of IoT and AI to reduce the impact of water wastage in the supply network. The project aims to monitor the water supply network for the benefit of the network manager and also for all citizens, who will know in real-time the details of water consumption. The project is divided into three main phases. The first part consists of the installation of all the sensors necessary for the acquisition of the data. That information is received by a platform that allows the storage and processing of the data, which in turn provides the mechanisms for training predictive models based on Deep Learning. Finally, the stakeholders can access all the elaborated information through a web portal that provides monitoring dashboards. We tested the usefulness of the project in a real environment in the province of Catanzaro, in southern Italy.

2. Project description

The project proposes a solution to mitigate the problem of water waste, a common problem in several Italian towns. With the platform, it is possible to monitor, in real-time, the water supply network, which allows knowing the water consumption, as well as the detection and prevention of possible failures.

To create a platform capable of providing a quality service, it is necessary the creation of an infrastructure of last-generation sensors, according to the logic of IoT, which will be installed at the points of adduction of water pipes as well as at the points of withdrawal. This infrastructure will also see the presence of appropriate data stream Big Data [1] transmission systems, which will constantly communicate with the processing platform.

A processing platform provides the functionalities for the acquisition, storage, and processing of large volumes of heterogeneous data, as well as mechanisms for the training of predictive models based on Machine Learning and Deep Learning [2] techniques, in addition to the possibility of combining them with automatic reasoning systems based on Answer Set Programming [3] and Logic Programming [4].

The end-users can access the processed data through the implementation of a web portal that provides monitoring dashboards. Using these dashboards is possible to identify leakage scenarios due to water main failure or water theft. The web portal also offers the possibility of predicting failure situations through the application of predictive maintenance techniques, thus allowing significant savings on maintenance activities and simultaneously minimizing water losses, with the resulting environmental impacts.

2.1. IoT infrastructure

As mentioned above, to monitor the water supply network, it is necessary to install devices that allow us to measure the water flow, as well as the transmission of this information to the platform. The sensors used are listed below:

- **Meters and transmitters**, to be installed in place of the mechanical meters already present in the individual utilities served and in the well from which the pipeline serving the utilities starts.
- **Concentrator device**, for collecting and sending consumption data from users, to be installed on the public lighting pole.

For energy-saving reasons, devices mounted on mechanical meters do not perform complex processing of measured values. However, these devices are able to record and transmit alarms determined by water leaks and/or low battery levels. Complex analysis and computation are all delegated to the application server.

2.2. Analytics platform

In this section, we will analyze in-depth the technical details of the platform in charge of acquiring and processing the data sent by IoT devices, and also, we will mention the platform's main features.

2.2.1. Main features

Previously we have discussed the main idea behind the development of the HidroControlIT project, but to get an idea of what the platform can do, below we list the main features that make it very useful for the water utility and the citizens:

- Real-time monitoring of the water supply on the network and identification of any problems/criticalities
- Prediction of water mains failure situations and identification of possible water tampering/theft scenarios
- Visualization of water consumption readings for individual users or groups and for time intervals parameterizable by the system user through a web dashboard that will be accessible by different types of users with different levels of visibility
- Detection of abnormal situations through real-time reporting of IoT devices in the distribution network that are not transmitting data or providing anomalous readings
- User reporting highlighting personal information and data regarding water readings. The reports will also have to identify any users without readings and will have to provide a series of statistics and graphs to allow network managers to evaluate the trend.

- Geographic localization and map display of IoT monitoring devices
- Monitoring and reporting of abnormal situations, both with reference to the operation of the sensors and concerning the consumption detected

Each citizen will have a dashboard to monitor their water consumption, check for leaks and compare their consumption style with that of other similar users in their area. Figure 1 is a screenshot taken from the web portal that shows some of the features mentioned above.

2.2.2. Logical architecture

In Figure 2 we can see the logical architecture of the proposed solution. The left part of Figure 2 shows the set of possible types of information sources that can interest the analysis processes. In particular, it is possible to store and analyze structured/semi-structured data, unstructured data, and streaming data coming from the IoT sensors.

On the other hand, the central part describes the infrastructure for data storage and processing, possibly hosted on a cloud platform. The processing sub-system is logically divided into a section of Core Capabilities, containing Map Reduce [5] algorithms, Distributed File System Access, In-Memory Data Access, Workload Management, and NoSQL Data Access [6], and a section that offers the real analytical functionalities, consistently with the CRISP-DM [7] methodology. The data access and/or analysis services are shown in the right part of Figure 2, which are available to end-users.

2.2.3. Physical architecture

In the proposed system, the interaction between the sensors and the platform takes place through appropriate Concentrators/Gateways devices that collect the data transmitted by the sensors mounted on volumetric meters and send them to a control server based on LoRa technology [8].

The volumetric counters are classic mechanical counters on which an impulse counter connected to a transceiver unit is mounted. The control unit communicates on LoRa WAN

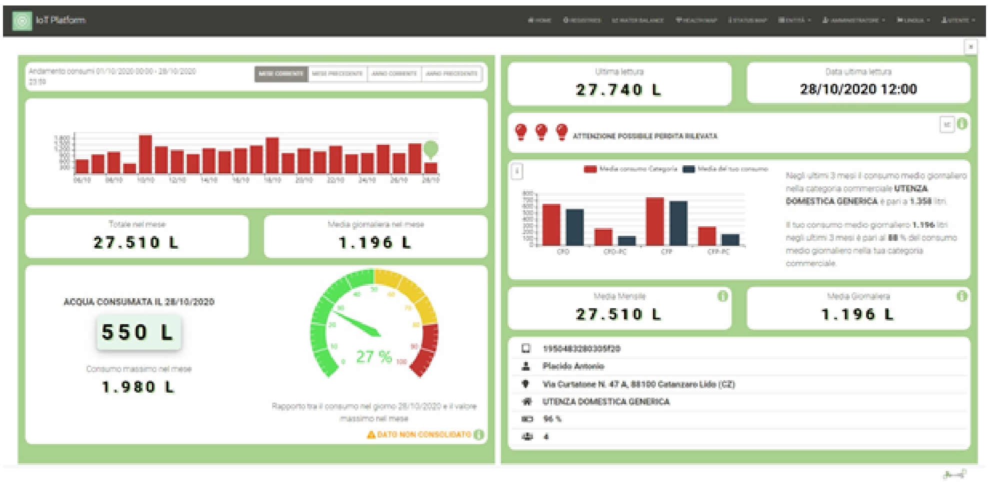

Figure 1: Dashboard to monitoring the incoming data in real-time.

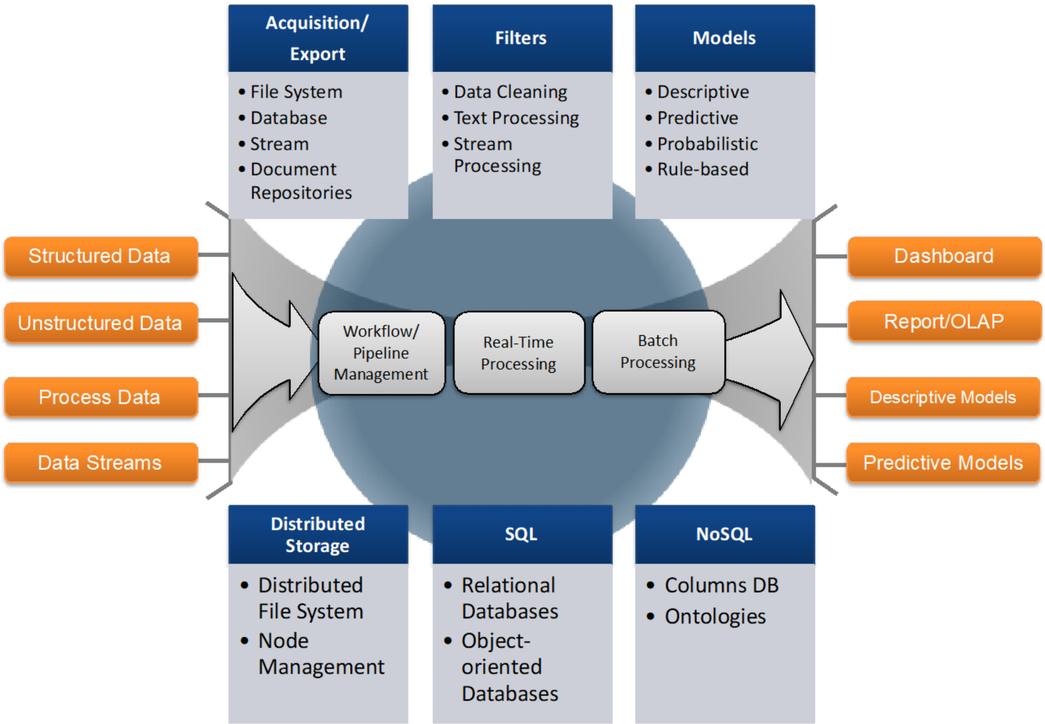

Figure 2: *Hydrocontrol IT* logical architecture.

protocol with a Gateway device. In particular, in the areas subject to experimentation, the meters of the single users are replaced by new meters that are mounted on the joints of the water mains, allowing the calculation of the water balance.

When the LoRa server receives the data transmitted by the Gateway, it forwards it to the application server. The LoRa server also has the task of sending the control commands to the control units mounted on the meters, making it possible to modify the measurements sending frequency or synchronize the internal clock. The sending of these commands is always intermediated by the Gateway device.

The platform is based on a microservice architecture [9], where the services are provided by a Kubernetes cluster hosted in the cloud. There are two main microservices:

- **Gateway + Front End application**: this is a Java + Spring Boot microservice that takes care of forwarding requests to other backend microservices and exposing the Front-End pages written in Angular. This microservice also takes care of security and authorization management.
- **Master Data + Ingestion**: this is a Java + Spring Boot microservice that manages the master data, supporting the whole system and the data ingestion. Data ingestion can take place from different information sources. In particular, the system acquires data from water meters using the LoRa WAN protocol.

Figure 3 is an illustration of the physical architecture of the system. As we can see, all components communicate within a private, secure VPN. End users access the platform by connecting

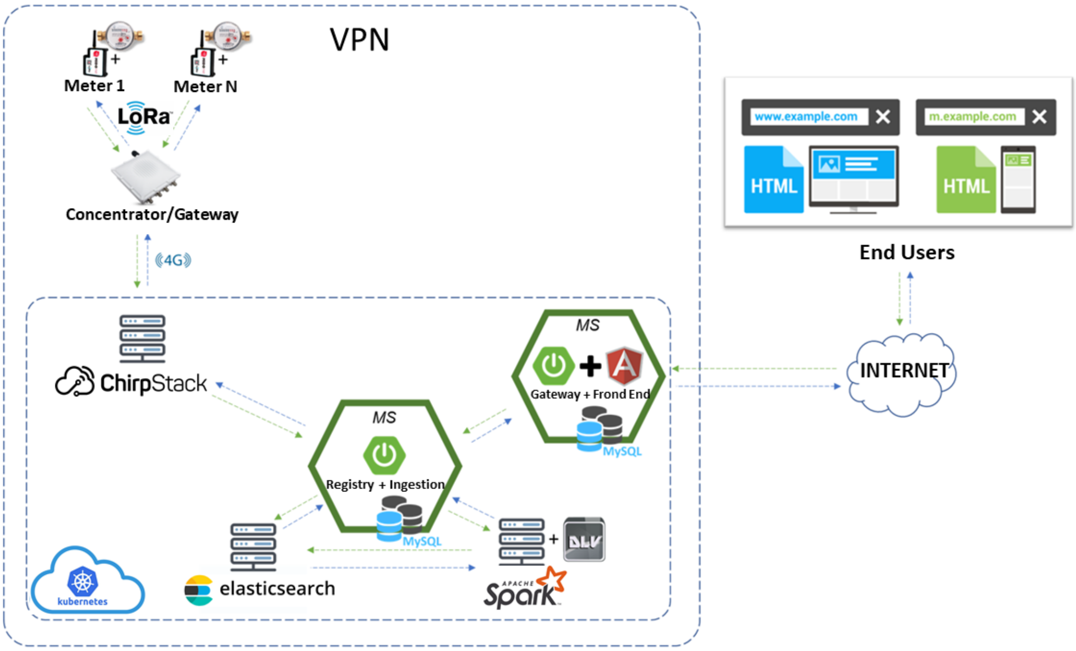

Figure 3: *Hydrocontrol IT* physical architecture.

to the microservice that exposes the front-end.

2.2.4. Predictive models

To detect possible problems in the water supply network, such as leaks or anomalous water consumption, as well as to perform predictive maintenance, it was necessary to apply machine learning algorithms on the data flow obtained from each customer. The data flow is a time series T, expressed as a sequence $[(x_1, z_1), ..., (x_n, z_n)]$ where each pair (x_h, z_h) is composed of a set of data related to the water consumption (x_h) and a timestamp (z_h).

The machine learning algorithm we applied was TS-Part [10], which allows us to perform clustering and detect irregular situations in customer water consumption information. TS-Part is based on a centroid-based partitional method, and as opposed to K-means [11], it does not require the numbers of clusters as a parameter.

TS-Part starts by considering the input dataset as a single cluster. Then, two main steps are iteratively applied until reach convergence. The first step consists in finding the best split for each cluster in the current clustering. The second step recomputes the cluster centroids and reassigns all data according to the actual clustering, similarly to the K-Means algorithm. The convergence of the algorithm is reached when the split procedure does not perform any split. The quality of a candidate cluster is calculated by subtracting the inter-cluster distance (the average pair-wise distance between all the cluster centroids) and the intra-cluster distance (the average distance between all the individual data within the cluster and the corresponding centroid).

To calculate the distance among time series and to be able to compare them, we relied on the

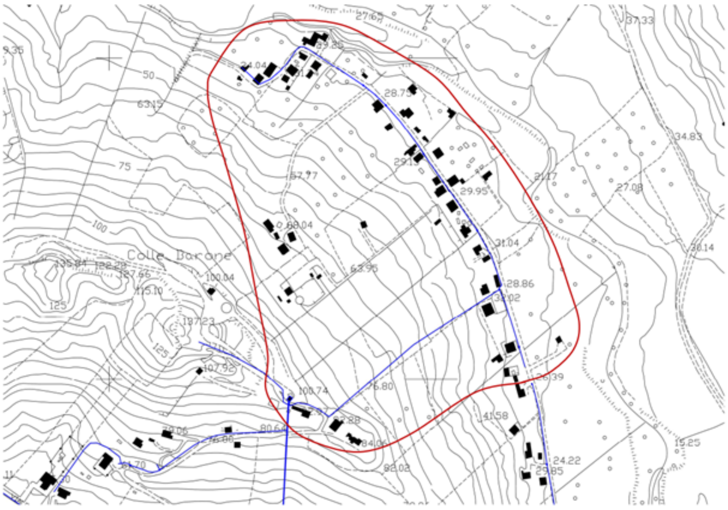

Figure 4: Scale representation of the water supply section related to Curtatone street.

Dynamic Time Warping (DTW) algorithm. The choice of DTW is because it solves the problem of the sensitivity to small distortions (i.e., fluctuations or phase shifts) in the time axis, a factor that affects the Euclidean distance when dealing with time series.

After applying the algorithm, customers are grouped according to the similarity of their water consumption style.

3. Experimentation

To evaluate the usefulness of the project, we decided to experiment in a real environment, and for this purpose, we selected a well-defined stretch of the water network in the municipality of Catanzaro. The water network section that has been chosen for this experimental phase is the one related to the final part of Curtatone street, containing about 80 small private houses with gardens and a cheese factory. Figure 4 shows a scale representation of the mentioned above.

The experimentation lasted approximately seven months, from 2020-10-06 to 2021-04-27. During this time, 8444 m^3 of water were delivered to the network, while the total amount of water measured by the smart meters was 6840 m^3, which indicates that there was a loss of around 18% of the delivered water. Around 15% of the mentioned water loss was due to the poor condition of one of the water pipes. We also identified some water thefts due to garden irrigators connected to the pipe but excluded from the water meters.

In the second part of the experimentation, we applied the TS-Part algorithm to the water-consumption time series of all users in the aforementioned period. As a result, we obtained three different clusters or user profiles, which are detailed below:

- **Cluster 0** contains only one user, the cheese factory, which has a daily average water consumption of 5 m^3

- **Cluster 1** contains about the 25% of domestic users, with a daily average water consumption of 0,7 m^3
- **Cluster 2** contains about the 75% of domestic users, where each user consumes around 0,3 m^3 of water per day

If we compare the centroid curve determined by TS-Part with the water consumption of a specific user, we can determine if the user profile fits the curve or not. In this way, we can detect possible anomalous behaviors.

For example, Figure 5 shows the comparison of the centroid curve of Cluster 1 with a user profile. We can see that in some cases, the curve of the user has the same shape as the centroid curve, but there is a sort of scale factor. This fitting difference probably means there is a water theft because the user reduced the water volume measured by the smart meter by using a pipe branch.

4. Conclusions

As we have shown, the *Hydrocontrol IT* platform can be useful for both customers and the administrator of the water supply network, allowing monitoring of water consumption, and detecting possible anomalies, such as malfunctions in pipes, water thefts, among others. Nonetheless, our system can be extended in many directions, for instance, with an optimized allocation of sensors in the networks to reduce costs, with automatic fault detection algorithms, and/or planners for reducing annoiances for customers during water network repair actions. Thus, with this paper we would like to open a public discussion on the usage of artificial intelligence techniques in the field of application of our system, fostering cooperation with AI researchers and collecting suggestions for future enhancements of our system.

References

[1] S. Sagiroglu, D. Sinanc, Big data: A review, in: 2013 international conference on collaboration technologies and systems (CTS), IEEE, 2013, pp. 42–47.

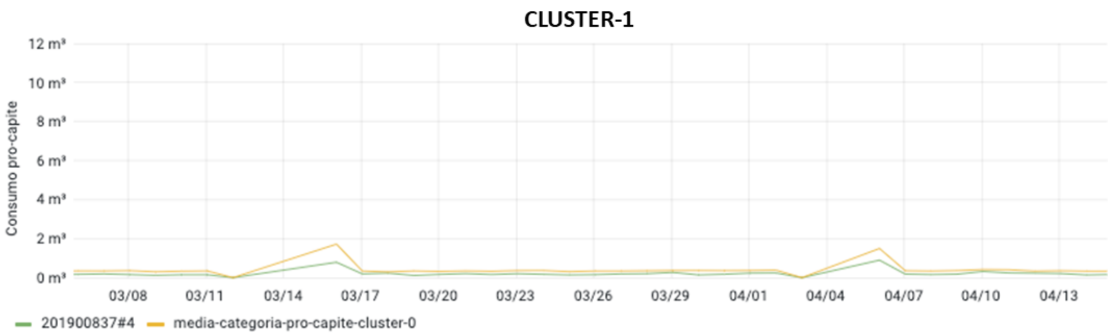

Figure 5: Fit of the Cluster 1 centroid curve and a user water consumption profile.

[2] I. Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT press, 2016.
[3] N. Leone, F. Ricca, Answer set programming: A tour from the basics to advanced development tools and industrial applications, in: Reasoning web international summer school, Springer, 2015, pp. 308–326.
[4] K. R. Apt, Logic programming., Handbook of Theoretical Computer Science, Volume B: Formal Models and Sematics (B) 1990 (1990) 493–574.
[5] J. Dean, S. Ghemawat, Mapreduce: simplified data processing on large clusters, Communications of the ACM 51 (2008) 107–113.
[6] C. Strauch, U.-L. S. Sites, W. Kriha, Nosql databases, Lecture Notes, Stuttgart Media University 20 (2011) 24.
[7] R. Wirth, J. Hipp, Crisp-dm: Towards a standard process model for data mining, in: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining, volume 1, Springer-Verlag London, UK, 2000.
[8] LoRa-Alliance, Lora, 2021. URL: https://lora-alliance.org/.
[9] N. Dragoni, S. Giallorenzo, A. L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin, L. Safina, Microservices: yesterday, today, and tomorrow, Present and ulterior software engineering (2017) 195–216.
[10] F. Gullo, G. Ponti, A. Tagarelli, S. Iiritano, M. Ruffolo, D. Labate, Low-voltage electricity customer profiling based on load data clustering, in: Proceedings of the 2009 International Database Engineering & Applications Symposium, 2009, pp. 330–333.
[11] A. Likas, N. Vlassis, J. J. Verbeek, The global k-means clustering algorithm, Pattern recognition 36 (2003) 451–461.