Tethys: Sensor-based Aquatic Quality Monitoring in Waterways

Gavriil Tzortzakis
Univ. of Athens
Athens, GR15703
gtzortzakis@di.uoa.gr

Eli Katsiri
Democritus Univ. of Thrace
Xanthi, GR67100
ekatsiri@ee.duth.gr

George Karavokiros
NTUA
Athens, GR15780
george@itia.ntua.gr

Christos Makropoulos
NTUA
Athens, GR15780
c.makropoulos@itia.ntua.gr

Alex Delis
Univ. of Athens
Athens, GR15703
ad@di.uoa.gr

***Abstract*—It is imperative that water is delivered clean to urban centers and towns despite its channelling through waterways, ponds and city aqueducts. Waterways are occasionally polluted by micro-organisms, landslides, pesticides, as well as by human activity and waste. In coordinated efforts to address such problems, water authorities and local governments resort to cleaning facilities whose main task is to filter the water in a timely fashion. In this regard, authorities must be forewarned of imminent or developing pollution issues, so that immediate corrective action can be taken. The installation of sensors that continuously monitor the quality of water passing through specific points in waterways, proves to be an effective way to implement legislative mandates for clean water. We use submerged sensors to gather measurements that can help characterize the quality of water in canals using parameters including temperature, conductivity, turbidity, *PH*, and pressure. Raw sensor-generated measurements turn out to be of limited help when it comes to monitor the overall water quality and by themselves, can be even misleading occasionally. In this paper, we discuss the main features of *Tethys*, a real-time water quality monitoring tool, whose aim is to help the Athens water authorities in their ongoing assessment of water quality. *Tethys* receives as input streams of measurements from stations in the field, detects unexpected events, visualizes the flow of information, and automatically alerts supervisors about potential dangers appearing in waterways. We outline our design choices, filtering mechanisms, and implementation effort in realizing *Tethys* and demonstrate its real-time use.**

Keywords*—*Sensors for Assessing Water Quality, Web Application.

I. Introduction

Water is the epicenter of the biosphere and empowers all forms of life. Societies have been using water for a wide range of critical activities in industry, agriculture, raising livestock and small animals and urban use. Through the ages, waterways or canals have been established to deliver clean water to cities and supply the human activity [1]. Portions of these channels may be natural and vulnerable to pollution. The most common sources for water pollution emanate from land sliding caused by earthquakes and floods, human and industrial waste, high concentrations of micro-organisms, overheating caused by climate change, rusty water pipes as well as fertilizers and pesticides used in agriculture. As various epidemics had been caused and/or spread through water contamination [2], [3], it is evident that the quality of drinking water could not be ascertained exclusively based on human inspection [2]. Around the globe, local and federal governments have taken initiatives and introduced legislature to safeguard the quality of the drinking water. To this end, societies have over the years managed to dramatically prevent treatable stomach diseases by sanitizing the water supply in aqueducts, staging areas as well as in waterways [4].

Cleaning and sanitation facilities are a key element in the distribution network of waterways and are instrumental in maintaining the quality of water provided to consumers. This is a critical task that water authorities and local governments have heavily invested on over the years. It is imperative that such facilities receive early warnings in light of either unexpected or suspicious events taking place along the distribution network canals. Alerting and subsequent preparation to ultimate avert these events have to take place within a *two-hours* notice at most, in order to effectively deploy countermeasures [5]. Ideally, this window should be even within a half-hour period after the appearance of an unusual event in the waterways.

In wide-area hydro-systems such as that of the Athens metropolitan area (Fig. 1), it is imperative to deploy automated means that can detect possible downgrading of the water quality and subsequently alert pertinent authorities if possible in real-time. In this effort, a key aspect is the operation of customized *wireless sensors* [6] used to produce readings for water–state variables and characteristics. Nevertheless, raw measurements of sensors that are continually harnessed by submerged stations are difficult to read and cannot be rapidly interpreted by humans. Such measurements should instead be filtered and analyzed before a clear view of what is actually occurring on the ground can be derived and visualized. In this paper, we present *Tethys*, a real-time automatic water quality monitoring tool using ZigBee–enabled sensors [7] that have been deployed in the field. *Tethys* was developed in the context of the *ALPINE* project[1] and its objectives are to: a) present in a clear and comprehensive manner the state of water's quality in specific sections of the water distribution network, b) enable both the creation and management of noteworthy events that the system continually monitors, c) use data filtering techniques to avoid having outlier measurements partake in the operation of the tool and distort its accuracy, and finally, d) render timely notification to authorities and aqueduct facility operators for events that require not only immediate attention but more importantly, remedial action(s).

II. The Hydro-system of Athens

Tethys focuses on the management of the hydro-system of Athens as depicted in Fig. 1. This is a particularly complex, long-distance conveyance system, bringing water supply from as far to the west as the *Evinos River Basin* in Epirus province to the capital [8]. The *Athens Water Supply and*

[1] http://alpineproject.altec.gr

2375-0324/16 $31.00 © 2016 IEEE
DOI 10.1109/MDM.2016.56

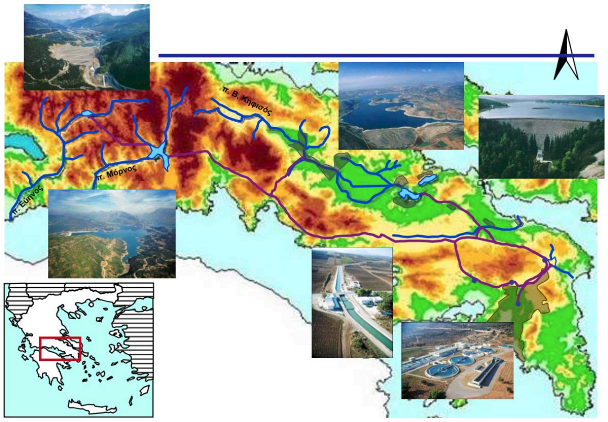

Fig. 1. The hydro–system of Athens

Sewage Company (EYDAP) owns the infrastructure and the company's mandate is to provide high quality water services to a large number of consumers. Transient population and the ever–changing demographics of Athens metropolitan area pose significant challenges for the effective management and sanitation of the water supply. The above goals have to be simultaneously accomplished with a drastic reduction in the energy consumption needed for the operation of the Athens hydro-system. EYDAP system of conveyance, particularly its gravity–fed *Mornos* river canals, necessitates precise knowledge of the state of water flow and depth at multiple locations. The latter helps with the synchronization of opening and closure of sluice gates so that the timely delivery of water to meet demand can be warranted. Moreover, locking out portions or even shutting down segments of the system in light of emergency and in anticipation of accidents ensures its reliability. Although monitoring and control of the operation of the entire system is in place for some time now, EYDAP is set to adopt wide-scale remote sensing and in this way, it can ascertain the quality of its supply en route to the consumer in a holistic way.

Tethys, as a tool, is envisaged to significantly assist both monitoring and control operation of the Athens hydro-system by providing continuous feedback as far as the quality of the channelled water is concerned. Although sensor deployment at this time remains limited into a few locations along the distribution network, *Tethys* deployment yields interesting clues with respect to the water quality and materializes a platform that both enhances visualization of data and assists in taking remedial steps in real-time.

III. Architecture

The overall system architecture is illustrated in Fig. 2. Wireless sensors are submerged in selected geographic locations along the canal structure. These devices continuously transmit the measurements of their sensing to gateway–modules through ZigBee. A gateway is responsible for gathering these measurements and for forwarding acquired values to an industrial `PC` using any available communication interface, which in our case is either *WiFi* or *Ethernet*. In turn, the `PC` maintains Internet connection using *3G* or broadband link and so, it dispatches data to a server operating a dedicated database system. After ingestion, measurement data become available to *Tethys* from the collection server mainly through two approaches: 1) using Web Services launched on demand, 2) issuing `sftp` requests should data be published at regular intervals of time.

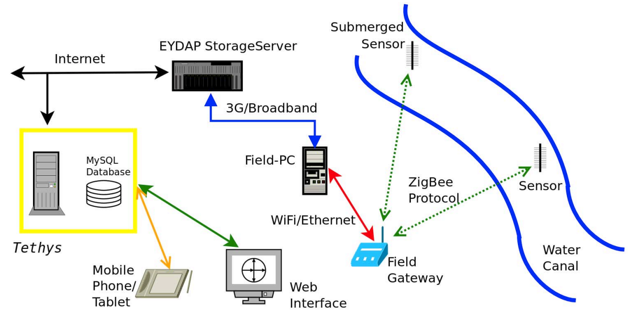

Fig. 2. Overall Architecture of the System

Once measurements become locally available, *Tethys* proceeds as follows:

– measurements are initially parsed using a customized *Java*-based tool and derived data are inserted in the database; the supported input formats are *ASCII*, *XML*, and *JSON*.
– data are then filtered so that apparent outliers are removed for analyses purposes without adversely influencing the monitoring procedure,
– data –including outliers– are stored in a *MySQL* database which is our storage choice for *Tethys*, and finally,
– temporal and value-related conditions are examined and if needed, automatic alerts are dispatched to designated users.

In general, our prototype supports two roles: *users* and *administrators*. On one hand, users are given access to all the monitoring facilities of *Tethys*. On the other, an administrator can check the current status of the system using statistics, graphs, and time-series; she can also handle the nature of events monitored. An administrator may also designate event "listeners" in order to be notified when certain conditions occur.

IV. Data Filtering

Time and again, we have observed that sensors do produce unreliable measurements. For example, our deployed sensors occasionally transmitted unrealistic values such a $-500°$ or temperature values dropped by more than $15°$ within short sampling epochs such as 10 seconds. An analysis of advanced filtering techniques to reject such values, termed *outliers* is discussed in [9]. The current version of the *Tethys* implementation uses Chauvenet's criterion [10] to efficiently identify outliers. The basic concept of this criterion is to find a probability band, centered on the mean of a normal distribution, that will serve as the domain of our acceptable points; points outside this band will be considered outliers and so will be rejected from our dataset. To apply Chauvenet's criterion in practice, we calculate the mean value and the standard deviation of our sample. Then for each outlier candidate, we calculate the probability of finding such a point, if our dataset would follow the corresponding normal distribution.

The latter is accomplished as follows: we compute the distance of a possible outlier from the mean value to determine the probability that a point will occur in this normal distribution [2]. Lastly, we multiply the previously found probability by the size of our sample. If the resulting value is less than 50%, then the point is declared an outlier and is rejected. It is important to note that such "rejected" outliers should not be really purged from the database. Outliers are only removed from the working dataset employed by *Tethys* to create charts and present time-series and other analyses. In our work, we have opted for storing all *raw data* that is sensor-produced in a separate table. This may enable the use of different filtering techniques in the future and may allow for data auditing, if required.

V. Event Detection

While monitoring sensor–produced data, we are keen in identifying "events of interest". The latter entail unexpected occurrences of measured data values that may portray rare or unexpected events. Such events typically affect a subgroup of the collected data rather than an individual data point. *Event detection* [11] refers to the problem of detecting when an event of interest occurs. Before events can be detected, they need to be specified. *Tethys* supports event specification in the following manner: an administrator can write expressions to define events of interest in the hydrological domain. For example, the expression "average temperature of the last 10 minutes in Station A is greater than $10°C$", is defined as: "$average(temperature(stationA, 10mins)) > 10°C$"

In *Tethys*, we express all possible conditions of interest using expressions of the following format:

$$f(v(s,t)) \;\; \Theta \;\; thr \tag{1}$$

where:

- function f can be one of simple aggregate functions including min, max and avg,
- v can be one of the measured entities such as temperature, turbidity, conductivity, *PH* and pressure,
- s designates the mnemonic of the station location in the hydro-systems such as $Station_A$, $Station_B$, etc.,
- t is the threshold of how old the data measurements can be,
- Θ can be either $<$ or $>$,
- *thr* is the user-set threshold value.

When defining events, a user can designate the specific set of data measurements that have to be collected within a time period. At all times, the system finds itself in state whose value can be: *OK, Warning*,and *Danger*. This state is altered according to the triggering of events and is readily visible. Fig. 3 shows a "fast view" of the state of the system through the green tick along with the specifics of the operation of two different measurements coming off sensors from different stations. Event detection is realized through a thread daemon that is responsible for determining whether user-designated conditions are met. Upon successful detection of an event, the portrayed state of the system does change and users in pre-specified lists are alerted via Email or *SMS*.

VI. Interface

We have designed *Tethys* while striving to produce succinct panes of useful information while simultaneously complying

[2]the probability is known by normal distribution tables.

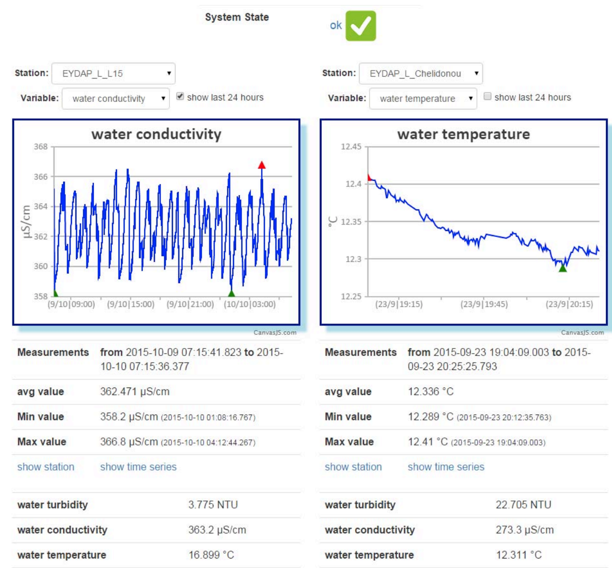

Fig. 3. Fast View

with core usability requirements [12]. By initially presenting an overview of the most important information in a single dashboard, *Tethys* allows the user to become informed at a glance, on key water quality indices (e.g., turbidity, conductivity, pressure, and dissolved oxygen), on the normality of station operation, or any unusual events and potential threats. By combining data from two of more collection stations, the user can develop a better understanding of the hydro-system at a system level. Finally, *Tethys* allows for extensibility with additional/new sensor stations, by providing flexible structures for visualizing the results without cluttering the dashboards with too much information. To this end, we have provided pop-down menus where the quality indices to be displayed can be selected and their format can be indicated through customizable charts.

As a single waterway station can have more that one sensors attached, *Tethys* can display all relevant geographic and operational details of all sensors that are part in the monitoring. A *timeseries* is a sequence of measurements of a specific parameter observed by a sensor. *Tethys Charts* help display time-series data. The user can dynamically observe the changes in water quality indicators using dynamic charts or she can focus on certain time-series and view *average*, *minimum* and *maximum* values. She can also visit a station profile to retrieve all information that relates to that station, including *date of installation, status, list of sensors, manufacturer, statistics* and *current values* of quality indicators. Respectively, there is a profile for each sensor including manufacturer provided technical-specifications, aggregate values, and the latest captured values for the quality indices measured.

Tethys uses a geographic map to indicate current locations of all stations and serves as the means to view the current values measured by corresponding sensors. Events (conditions) do have their own pane where a user can *view, setup, edit, delete, start* or *stop* an event detection.

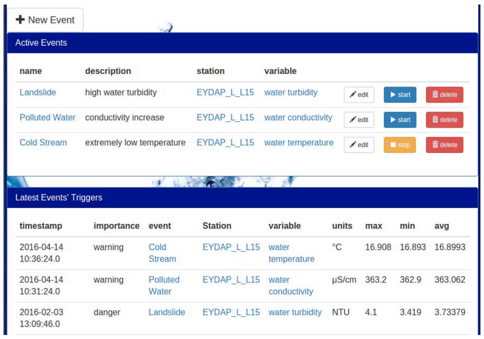

Fig. 4. Events Management

VII. Implementation

We have implemented *Tethys* using the *Spring* v4.0.1 *Model View Controller (MVC)* for it enables the rapid development of flexible, loosely coupled web applications written in *Java*. MVC has three parts: the *Model*, an object storing data in the application, the *View*, the showable representation of the model's data to users and the *Controller*, a module which alters the model [13]. Thank to *Spring*, *Tethys* can be used by the entire range of contemporary devices (PCs, tablets, and smartphones) having Internet access. The interface is enhanced by *Bootstrap v.3.3*, an *HTML*,*CSS*, and *JS* framework for developing responsive front-ends. We use *MySQL v.5.6* and *Tethys* communicates with the database through *Hibernate v4.0.1*, an object-relational mapping framework for *Java*. Our *Tethys* implementation runs on *Ubuntu VM* installation and can use any web server capable of executing *WAR* files including *Apache Tomcat*.

VIII. Demonstration

The goal of our demonstration is to present the system functionalities including *Tethys*' interface as well as event, notification and user management. We anchor our demonstration on the "quick view" of *Tethys*' main pane showing two stations, respective current measurements for utilized quality indicators and the overall system state as Fig. 4 depicts. We showcase how a user with event management rights can create new events, start, stop, modify or even delete events. We discuss how events can be weaved using the available stations/sensors and the types of measurement these device provide. We also demonstrate the automated notification system for *Email* and *SMS* messages to authorized groups of users. Our presentation is driven by a geographic map that distinctly indicates locations of deployed sensors and can easily render the specifications of the deployed devices.

IX. Summary

The automated monitoring of the water quality is an issue of paramount importance to both urban and country population and remains an ongoing concern for water distribution authorities. To address this ongoing challenge, we have created *Tethys*, a monitoring and reaction platform that receives input from sensors deployed in the distribution channels of the Athens hydro-system. *Tethys* follows sensor-derived values that continually measure water quality indicators in the distribution network. In this context, *Tethys* has realized a a comprehensive aquatic sensor monitoring tool [14] that can detect and provide early notification for developing situations and/or undesirable events. We plan to extend *Tethys*' functionality by: a) turning selected sensor(s) on/off in order to lower energy consumption, b) altering the sampling rate according to a quality indicator's current value or resetting the rate due to the encounter of unexpected crucial conditions, c) requesting data on demand and d) enabling online sensor firmware updates. We also plan to extend our system by offering a prediction module whose objective would be to render plausible forecasts in conjunction with other data source such as those provided by meteorological services. We investigate pertinent detection methods [15] as well as clustering and classification techniques [16] to develop multi-granularity prediction models applicable to specific segments of the canals of the hydro-system. Improved predictions of the water quality will inevitably lead to better management of water cleaning facilities and reduction of their relevant power consumption.

Acknowledgments: This work has been partially supported by *GSRT-Alpine* and *EU-Galena* projects.

References

[1] W. Vanneuville and B. Werner, "Water Resources in Europe in the Context of Vulnerability," http://www.eea.europa.eu/publications/water-resources-and-vulnerability, European Environment Agency (EEA), State of Water Assessment, Tech. Rep., 2012.

[2] R. D. Letterman, *Water Quality and Treatment: Handbook of Community Water Supplies*, 5th ed. New York, NY: McGraw Hill, Inc., 1999, american Water Works Association.

[3] A. Prüss, D. Kay, L. Fewtrell, J. Bartram *et al.*, "Estimating the Burden of Disease from Water, Sanitation, and Hygiene at a Global Level," *Environmental Health Perspectives*, vol. 110, no. 5, pp. 537–542, 2002.

[4] U.N., "Water Monitoring Systems," United Nations—Water Task Force on Monitoring, Stockholm, Sweden, Tech. Rep., August 2006.

[5] A. Krause, J. Leskovec, C. Guestrin, J. VanBriesen, and C. Faloutsos, "Efficient Sensor Placement Optimization for Securing Large Water Distribution Networks," *Water Resources Planning & Management*, vol. 134, no. 6, pp. 516–526, 2008.

[6] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless Sensor Network Survey," *Computer Networks*, vol. 52, no. 12, August 2008.

[7] "ZigBee 3.0: The Foundation for the Internet of Things," http://www.zigbee.org/zigbee-for-developers/zigbee3-0/, ZigBee Alliance, Tech. Rep., 2013.

[8] D. Koutsoyiannis, G. Karavokiros, A. Efstratiadis, N. Mamassis, A. Koukouvinos, and A. Christofides, "A Decision Support System for the Management of the Water Resource System of Athens," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 28, no. 1415, pp. 599 – 609, 2003.

[9] M. Gupta, J. Gao, C. Aggarwal, and J. Han, "Outlier detection for temporal data," *Synthesis Lectures on Data Mining and Knowledge Discovery*, vol. 5, no. 1, pp. 1–129, 2014.

[10] W. Chauvenet, *A Manual of Spherical and Practical Astronomy V. II.* Boston, MA: Harvard Univ. Astronomical Tutorial Library, 1863.

[11] D. B. Neill and W.Wong, "Tutorial on Event Detection," in *Proc. of the 15th ACM SIGKDD Conf.*, Paris, France, July 2009.

[12] J. Licklider, "Man-Computer Symbiosis," *IRE Transactions of Human Factors in Electronics*, vol. HFE-1, pp. 4–11, March 1960.

[13] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, *Design Patterns: Elements of Reusable Object-oriented Software*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc., 1995.

[14] E. Katsiri, J. Bacon, and A. Mycroft, "SCAFOS: Linking Sensor Data to Contextaware Applications Using Abstract Events," *Pervasive Computing and Communications*, vol. 3, no. 4, pp. 347–377, 2008.

[15] G. Karavokiros, E. Baltas, and M. Mimikou, "Identification of Extreme Weather Phenomena in Near Real-Time," in *Proc. of 1st EWaS-MED Int. Conf.*, Thessaloniki, Greece, April 2013.

[16] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in *Proc. of the 5th Berkeley Symp. on Math. Stat. and Probability, Volume 1*, Berkeley, CA, 1967, pp. 281–297.