SMaRT-OnlineWDN D5.5: Development of demand forecast software

Mathias Braun, Thomas Bernard, Idel Montalvo Arango, Jochen Deuerlein, Olivier Piller

Deliverable 5.5

Development of demand forecast software

Dissemination level: Public

WP5
Online Simulation Module

27th May 2014

SMaRT-Online WDN
Online Security Management and Reliability Toolkit for Water Distribution Networks

ANR reference project: ANR-11-SECU-006
BMBF reference project: 13N12180

Contact persons
Fereshte SEDEHIZADE Fereshte.Sedehizade@bwb.de
Olivier PILLER Olivier.Piller@irstea.fr

WP 5 – Online Simulation

D5.5 Development of demand forecast software

List of Deliverable 5.5 contributors:

From 3S [WP5 leader]
Jochen Deuerlein (deuerlein@3sconsult.de)
Idel Montalvo (montalvo@3sconsult.de)

From Irstea
Olivier Piller (olivier.piller@irstea.fr)

From IOSB
Thomas Bernard (Thomas.Bernard@iosb.fraunhofer.de)
Mathias Braun (Mathias.Braun@iosb.fraunhofer.de)

Work package number	5.5		Start date:	01/10/2012
Contributors	3S Consult	Irstea	IOSB	
Person-months per partner	2	1	1	
Keywords Time series analysis – Computational Intelligence – Demand Prediction				
Objectives Description of the development and implementation of the demand prediction module as part of the online simulation module.				

Summary

The main objective of the SMaRT-Online^{WDN} project is the development of an online security management toolkit for water distribution networks which is based on sensor measurements of water quality as well as water quantity. The development of a demand prediction module is one objective of the "online simulation" project work package 5 of. The task is the development of a prediction module that allows for the prediction of a single future water demand and the determination of a complete demand pattern for the following 24 hour cycle.

The objective of this deliverable is to give a report of the development and implementation of the demand prediction module. Initially, section 1 gives the context of the demand prediction module in the project. Section 2 explains the objectives of the module and how they are integrated in the workflow of the SMaRT-Online^{WDN} process. Section 3 gives the details on Prediction algorithm. Section 4 proceeds with details on the software implementation. A conclusion on the development is given in section 5.

List of Figures:

Contents:

1. Background

The SMaRT-Online^{WDN} security management toolkit is based on sensor measurements (hydraulic and water quality sensors) and online hydraulic and transport models and will allow an estimation of the localization of the contamination source and the simulation of short-time future scenarios in order to estimate the impact of a contamination source. In order to make an accurate short-term prediction for contamination scenarios a good prediction for the hydraulic state of the network is needed. The hydraulic state is strongly influenced by the water demand in the observed area what poses the need for an accurate prediction and calibration of the water demand (Piller *et al.,* 2014 and Preis *et al.*, 2010). Since water demand is subject to high seasonality it is prudent to use the tools of time series analysis to model the water demand and predict the consumption in a district or consumption nodes of the network.

In time series analysis autoregressive integrate moving average (ARIMA) models have been used for decades and due to their success they have been used in a wide variety of scientific applications. In recent years a growing popularity of machine learning algorithms like the artificial neural network (ANN) and support vector machine (SVM) have led to new approaches in time series analysis.

2. Aims and Workflow of the Demand Prediction Module

There are two basic aims for the demand prediction module called the One-Step-Prediction and the Pattern-Prediction. The Objective of the One-Step-Prediction is to predict the water demand for an area or node of the network for a predetermined time step (i.e. fifteen minutes) into the future. The objective of the Pattern-Prediction is to predict the water demand in an area or node of the network over a predetermined timespan in a discrete number of steps (i.e. during the next 24 hours in hourly steps).

These two modules are integrated into the SMaRT-Online^{WDN} toolkit in two different workflows. The results of One-Step-Prediction are used as initial estimation by the online demand calibration. Calibration of the demand for the last observation period consists of seeking demande estimates that minimize the sum of squared residuals between the observed and hydraulic and transport model values. Then the initial demand prediction are corrected thanks to the last measurements (flow rates, pressures, concentrations….). To do this the customers in the network model are divided into a number of customer groups. Based on the adapted prediction model and historic demand data the water consumption for each customer is first predicted by the One-Step-Prediction tool then corrected (adjusted) thanks to additional information in the measurement and the simulation models. The workflow of the Online-Calibration is shown in Figure 1.

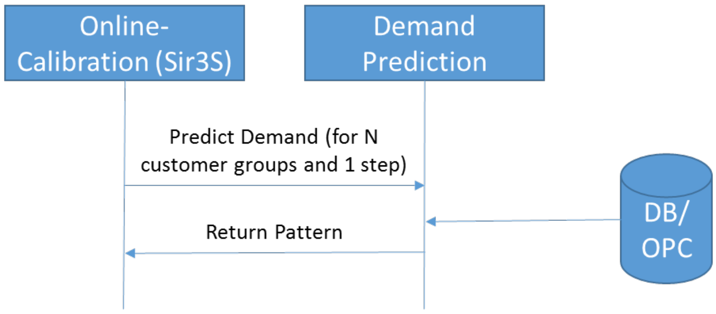

Figure 1: Workflow for One Step Demand Prediction in Online Calibration

The second one is the Look-Ahead simulation shown in Figure 2. In case of a contamination event detected by the alarm generation the SMaRT Online process starts a Look-Ahead simulation to determine the propagation of the contamination throughout the network. For this task the hydraulic state of the network has to be determined based on the prediction of the water consumption.

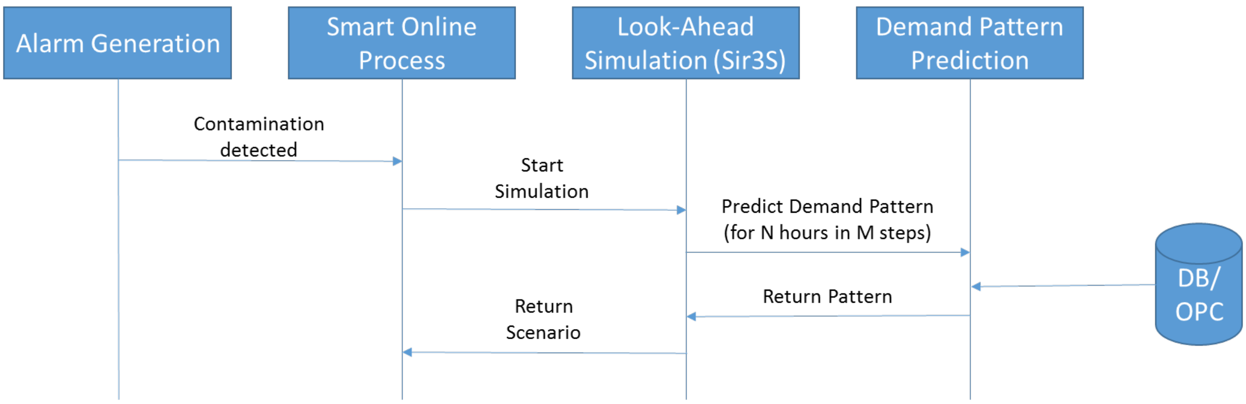

Figure 2: Workflow for Demand Pattern Prediction in SMaRT Online

3. Algorithm of the Demand Prediction Module

For the development of the prediction module different algorithms have been researched and evaluated. Namely the classes of autoregressive integrated moving average (ARIMA) models and support vector regression (SVR). Due to their superior ability in modeling non-linear effects the SVR is chosen as basis for the demand forecasting module.

Support Vector Machines are supervised learning algorithms from statistical learning theory that are used to recognize patterns in data for classification. Support Vector machines for function estimation are also referred to a Support Vector Regression (SVR). For the demand prediction algorithm the ε-SVR is used. The goal of ε-SVR is to find a function that has a maximum distance of ε from the actual measurement points y_i and at the same time is as flat as possible. For points that do not fall within the margin ε an additional cost is added proportional to its distance from the margin. These points define the support vectors. The model has two independent parameters C and ε that have to be adjusted for fitting the model to the data. The cost factor $C > 0$ which determines the trade-off between the flatness of the regression line and the distance to the margin ε that is tolerated. The SVR is linear process, however by applying kernel functions the feature space can be transformed and the same method can be used for nonlinear problems. The kernel function can be defined by polynomial functions, Gaussian radial basis, hyperbolic tangent and many more. In the following the radial basis Gaussian kernel will be used.

The inputs to the model are given by past demand values with a time lag of $t - lag_{min}$, $t - 24$, $t - 25$, $t - 168$ and $t - 169$. This pays respect to how far into the future the prediction goes, the daily and the weekly cycles in water demand. Additional information is given by the time of day, the day of the week and the month. Based on these factors and the model the water demand is predicted. More detailed information on the algorithm selection and application are given in (Braun *et al.*, 2014)

4. Software-Implementation and functionality

The software is implemented using C# and the .NET Framework. For the implementation the Accord.NET Framework is used under the GNU Lesser General Public License (LGPL) license.

The core of the demand prediction is implemented in the DemandForecasting.dll. It contains the class *CForecastingProblem* that contains the kernel support vector machine and a list of normalized historic data elements which contain the information defined in the previous chapter. The central methods are given by *DoTrain* and *DoPredictSingle*. The first method trains the kernel support vector machine with the input data. The second method predicts the demand for a single normalized input data element. Three further functions are given by *DoPredictMany* which predicts the demand for a list of input elements, *Save* to save the instance of the *CForecastingProblem* in a binary data format and *Load* to load a particular instance of *CForecastingProblem*.

For testing and demonstration purposes the DemandForecasting.dll has been integrated into a simple GUI application shown in Figure 3. It contains two tabs called “Training Data” and “Testing Data”. Both windows are designed similarly, showing the input data elements in the table on the left and the visualization of the time series on the upper right. For training the SVM, a prediction span is chosen from the combo box (here reaching from 1 up to 14 hours) a set of training data is selected on the button “Load Training Data” and the training is started by the button “Train”. Afterwards the SVM parameter estimation is saved. For testing, the SVM is selected by choosing the time span from the combo box. A testing set different from the training set is selected by clicking on “Load Testing Data”. The prediction is performed by clicking “Test”. The results are shown in the Visualization together with the correct data and the relative error.

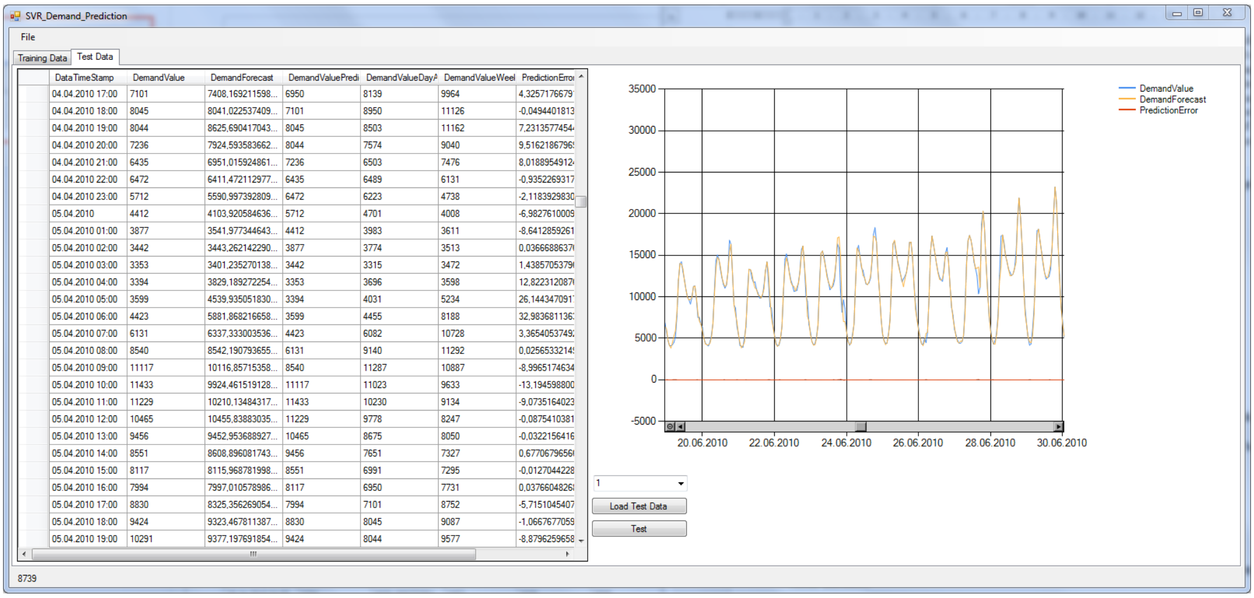

Figure 3: Demand Prediction GUI using the DemandForecasting.dll

5. Conclusion

The following list summarizes the different steps that were performed in this study with the provided resources:

- Comparison of different time series modelling approaches with SARIMA and SVM based on the consumption data from Berlin östl. Hochstadt. It was found that the SVM framework has better goodness-of –fit.
- Definition of two workflows for the online simulation and look-ahead simulation using “One Step Prediction” and “Pattern Prediction”.
- Implementation of a SVM-based demand prediction module in C# using Accord.NET resulting in DemandForecasting.dll
- Implementation of the Demand Prediction GUI for demonstration and testing purpose.

6. Referred bibliography

Piller, O, Montalvo, I., Deuerlein, J., Braun, M., & Ung, H. (2014). A gradient-type method for real-time state estimation of water distribution networks. International Conference on Hydroinformatics.

Preis, A., Whittle, A. J., Ostfeld, A., & Perelman, L. (2010). Efficient hydraulic state estimation technique using reduced models of urban water networks. Journal of Water Resources Planning and Management, 137(4), 343-351.

Braun, M., Bernard, T., Piller, O., & Sedehizade, F. (2014): "24-hours demand forecasting based on SARIMA and support vector machines. Water Distribution Systems Analysis (WDSA) 2014