CIMSA 2008 – IEEE International Conference on
Computational Intelligence for Measurement Systems And Applications
Istanbul – Turkey , 14-16 July 2008

Application of Data Approximation and Classification in Measurement Systems - Comparison of "Neural Network" and "Least Squares" Approximation

Amir Jabbari, Reiner Jedermann, Walter Lang
Department of Electrical Engineering, Institute of Micro sensors, Systems and Actuators (IMSAS), University of Bremen
NW1 Building, D-28359 Bremen, Germany
Phone: +49-421-218-7353, Fax: +49-421-218-4774, Email: *ajabbari@imsas.uni-bremen.de*

Abstract *–In measurement systems, environmental conditions are measured based on predefined scenarios. Measured data are then processed in either a decentralized or centralized manner. In advanced systems (especially for distributed data processing), taking artificial intelligence features into consideration could improve measurement performance and reliability. It is assumed as autonomy in measurement system which leads to distributed "intelligent data measurement and processing". In this paper, two different methodologies for "temperature prediction" are compared. A discussion concerning the classification of recorded data is then presented. Both a mathematical approach, the so-called "least squares" approach, and a model-free approach, called back-propagation, are applied and compared for temperature approximation. After approximation, the predicted temperature values are compared with real temperature records for classification purposes. The "classification mechanism" includes signal processing features for improving performance.*

Keywords *– Measurement system, artificial intelligence, temperature approximation and classification, evolutionary computation.*

I. INTRODUCTION

Monitoring and control play very important roles in transportation systems, because the quality of products depends on changing environmental conditions such as temperature and humidity. Therefore, in measurement systems, environmental conditions are measured according to predefined criteria and then the measured data are processed in either a decentralized or centralized method. In centralized systems, measured data are sent for processing through a communication linkage. In distributed systems, depending on the applied mechanism, the measured data are processed or preprocessed at the measurement stage.

In advanced measurement systems, based on recent improvements in artificial intelligence, processing of and decision making about the measured data are applied locally. This procedure is called "intelligent data processing". One of main applications of distributed "intelligent data processing" is "autonomous fault detection and isolation" in measurement systems [1]-[3]. It is necessary to evaluate the reliability of a system during measurement [4], [5], and "artificial intelligence" features could increase the reliability of measurement systems, such as those intended for use in food transportation. In order to process data, the data are first approximated. For this purpose, there are various methodologies for regression (such as support vector machine) and data approximation, which both entail certain advantages and limitations [6], [7]. There are linear and nonlinear techniques to predict the next values of data for approximation purposes [8]. The "least squares" approach is based on minimizing the sum of the squared residuals. The residual symbolizes the difference between the actual observed data and that of the model. The "least squares" approach is used for a "maximum-likelihood" estimation of the parameters [9]. A further solution employs knowledge-based data approximation techniques such as "artificial neural network" that is established through learning and evolutionary computation [10], [11].

In this paper, two different approaches including "least squares" and "back-propagation" are applied and compared to the actual temperature records of a food trading company. Finally, the approximated data is used for data processing using a classification architecture. This leads to evaluation of the reliability of recorded data. The findings of this paper could be used for data approximation and classification for intelligent sensor fusion, intelligent measurement systems, and evolutionary monitoring and control of transportation system, which is described in this application.

II. PROPOSED APPROACH

The "least squares" estimation is one of the most important statistical approaches, which uses either determined or over-determined linear or nonlinear equations according to the relationship between process events for both modeling and prediction. [6].

The artificial neural network (ANN) is a knowledge-based approach with several applications [12]-[14]. There are two main ANN approaches for parameter approximation including "radial basis function" (RBF) and "multi-layer perceptron" (MLP) with "back-propagation" technique [15]. RBF is a local approximator which yields better accuracy for

978-1-4244-2306-4/08/$25.00 ©2008 IEEE

local approximation purposes, while MLP is a better choice for global approximation [16]. Therefore, in this application, the MLP mechanism is used for temperature approximation. For this purpose, two individual mechanisms were designed to approximate data based on the "least squares" and "back-propagation" algorithms.

After approximation, the mechanisms are applied to the temperature records in trucks containing fresh food during transport by Rungis Express, a food trading company in Germany [17]. For this purpose, two trucks are used for food transport and each truck is divided into three compartments. 40 data loggers are attached to different positions (with a maximum error of ±0.5 °C) for recording temperature. The cooling system, called the "reefer unit", is used for ventilation purposes and changes the temperature inside each compartment during its on/off cycles. Every 2.5 minutes the temperature is recorded onto the data loggers.

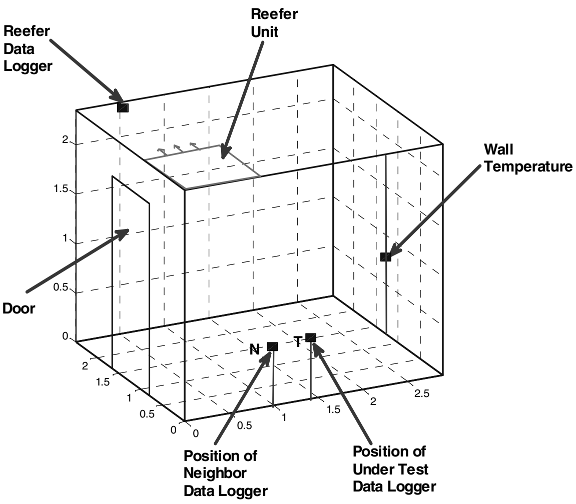

Fig. 1. Positions of the data loggers in the compartment

Three parameters are used for temperature approximation, including ambient temperature (parameter A), reefer temperature (parameter B), and wall temperature (parameter C). To record the reefer temperature, a data logger is positioned opposite the reefer unit. There is a data logger attached to the compartment to record the ambient temperature, which influences the temperature inside the compartment. In addition, the temperature of the wall, which is influenced by both the ambient temperature and the changes inside the compartment, is recorded during transport.

Each data logger in the compartment could be selected as the "under test data logger". In Fig. 1, the position of the "under test data logger" (T) is shown. Also, the position of the neighbor data logger (N) is shown, which is used to improve the classification mechanism.

In this application, the "least squares" approximation relies on the last four equations that shows relationships between the aforementioned parameters and temperature of the "under test data logger".

In (1), current and previous values of parameters A, B, and C are recorded and used to find K_A, K_B and K_C (approximation coefficients), which are calculated by solving the over-determined set.

$$\begin{cases} T(t-3) = K_A \cdot A(t-3) + K_B \cdot B(t-3) \\ + K_C \cdot C(t-3) \\ T(t-2) = K_A \cdot A(t-2) + K_B \cdot B(t-2) \\ + K_C \cdot C(t-2) \\ T(t-1) = K_A \cdot A(t-1) + K_B \cdot B(t-1) \\ + K_C \cdot C(t-1) \\ T(t) = K_A \cdot A(t) + K_B \cdot B(t) + K_C \cdot C(t) \end{cases} \tag{1}$$

By using the new values of parameters A, B, C, and the newly-obtained K_A, K_B, and K_C values, a new value of $T(t+1)$ is approximated. This procedure continues and the coefficients are calculated each time based on the last four equations. This generates over-determined sets each time.

$$\begin{aligned} T(t+1) &= K_A \cdot A(t+1) + K_B \cdot B(t+1) \\ &+ K_C \cdot C(t+1) \end{aligned} \tag{2}$$

Although more equations could be employed to generate over-determined equation sets, it has been verified that in this application, using more equations could decrease the performance of the approximation. In fact, it has been examined that in this application, using the last four equations has optimal performance over using three equations (which leads to determined equations) and also other over-determined sets based on more equations. Therefore, solving the over-determined system according to the last four equations is considered for generating approximation coefficients.

For the MLP network, two main layers (including input and output) are employed to spread the data. In addition, two hidden layers are considered, each including four neurons with a sigmoid activation function. The second hidden layer could improve data mapping performance between the input pattern and related targets.

$$H_{1j} = \sum_{i=1}^{n} w_{ij} x_i \tag{3}$$

In (3), H_{1j} is the weighted input of the j-th neuron in first hidden layer, w_{ij} refers to the weights between the i-th input and the j-th neuron in the first hidden layer, x_i is the i-th input layer element, and n is the number of inputs in the input layer. After calculating H_{1j}, the output of each neuron is obtained according to (4).

$$\Phi(H_{1j}) = (1+e^{(-H_{1j})})^{-1} \quad j=1,...,m \tag{4}$$

$\Phi(H_{1j})$ is the output of each neuron in the first hidden layer and m is the number of neurons in the second layer. The "gradient descent" algorithm is used for training [18], [19].

$$W_{new} = W_c - \alpha_c grad_c \tag{5}$$

In (5), W_{new} and W_c are respectively the new and current weight vectors, $grad_c$ refers to the current gradient based on changes of the "training error", and α_c is the training parameter.

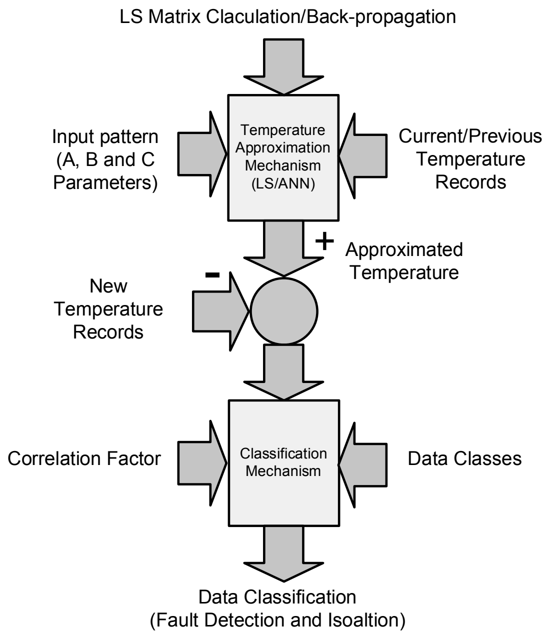

Fig. 2. Temperature approximation (LS/ANN) and data classification in food transportation system

In Fig. 2, the applied mechanism for data approximation is shown, which could be used for either the "least squares" (LS) or "artificial neural network" (ANN) techniques. After approximation, the approximated temperature is compared to the actual temperature records for classification according to defined data classes.

III. RESULTS

To compare the LS and ANN approximation techniques, a similar test period is used (270 minutes). The ANN mechanism includes two phases: the training phase (90 minutes) and the approximation phase (180 minutes), which together total 270 minutes. However, the "least squares" approach could begin approximating data after obtaining just the first four temperature records.

First, the performance of the "least squares" algorithm is presented for temperature approximation using the auxiliary parameters (A, B, and C). For this purpose, after obtaining the first four temperature records, approximation coefficients (K_A, K_B, and K_C) are calculated by recording and using the new values of the auxiliary parameters (A, B, and C). This procedure continues each time, and every 2.5 minutes the last four equations are considered.

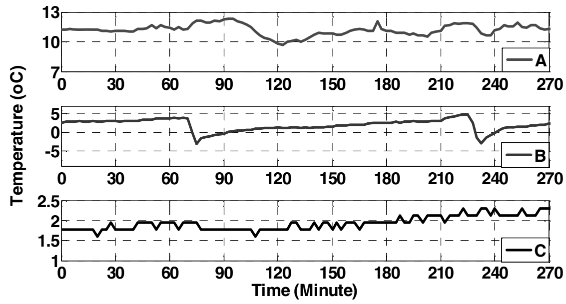

Fig. 3. Values of A, B and C for LS approximation

Fig. 3 shows the auxiliary parameters for temperature approximation. In Fig. 4, actual records and approximated records are shown.

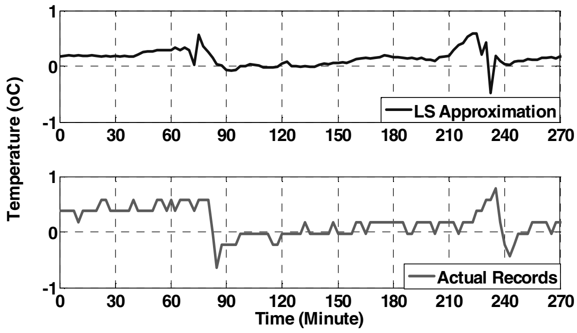

Fig. 4. Temperature prediction using "least squares" technique

In Fig. 5, the obtained approximated residual is shown, which is the calculated difference between the approximated and actual values.

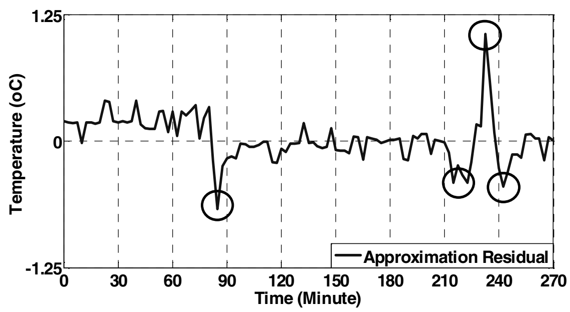

Fig. 5. Difference between actual and predicted values

According to Fig. 5, although the approximation is for the most part accurate, sometimes the approximation differs significantly from the actual recorded value. In this figure, the maximum approximation residual is 1.057 °C and the minimum approximation residual is -0.6674 °C. In some points, the approximation residual is very close to approximation accuracy limits (such as -0.4485 °C and -0.4127 °C in Fig. 5).

At this point, the performance of the "artificial neural network" mechanism for temperature approximation is discussed. For this purpose, the network first starts a training phase. Fig. 6 shows the process of training the network, in which input values are fed to the network. The training time depends on the evolution in the network for starting approximation. In this application, training the ANN approximation mechanism takes about 1.5 hour (0-90 minutes in Fig. 3). This period contains the data range that is considered for the approximation. Afterwards, even during approximation, the training phase is improved. This process is called "evolutionary training". The values of the network parameters for training and approximation (including activation function, epoch, etc.) directly influence the accuracy.

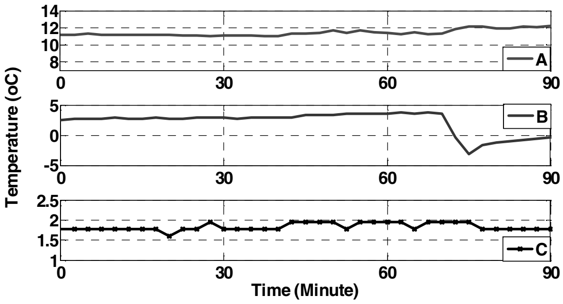

Fig. 6. Training phase for temperature prediction using ANN

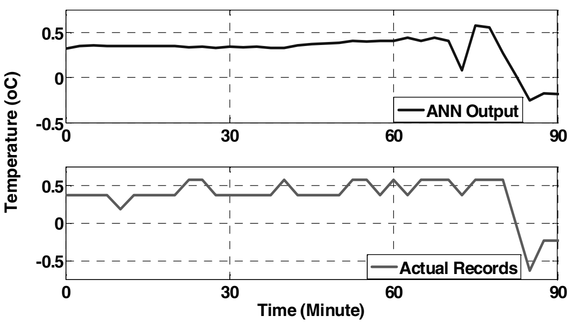

Fig. 7. Comparison of actual and ANN output

In Fig. 7 the performance of the training method is shown. In Fig. 8, the maximum training residual is 0.3059 °C and the minimum value of training residual is -0.3756 °C.

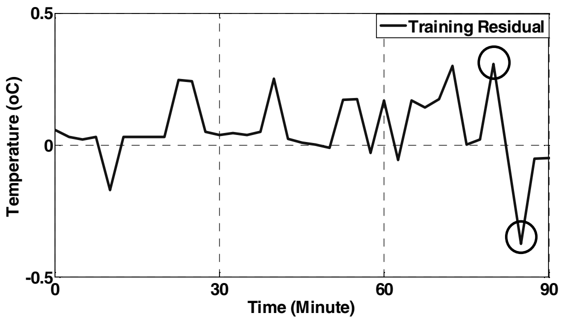

Fig. 8. Evaluation of training performance

Subsequently, new values are used to start the approximation by the network (Fig. 9). Approximation could be compared to the actual records (Fig. 10, 11). For this purpose, the performance of the approximation mechanism is checked over 180 minutes (minutes 90-270 in Fig. 3).

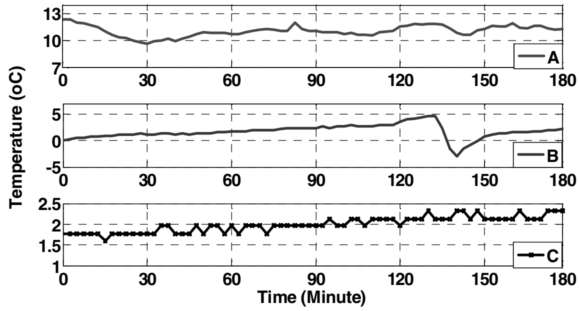

Fig. 9. Feeding A, B, and C for temperature prediction using the artificial neural network

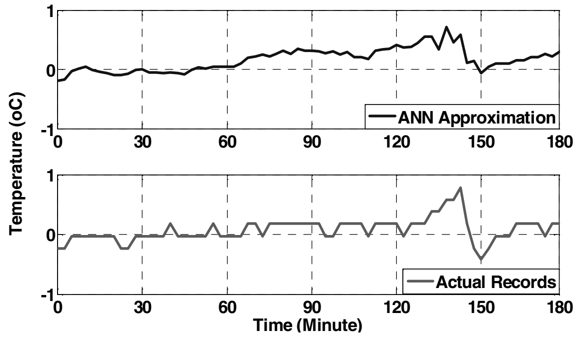

Fig. 10. Difference between actual and predicted values

In Fig. 11, the maximum approximation residual is 0.2338 °C and the minimum approximation residual is -0.3964 °C.

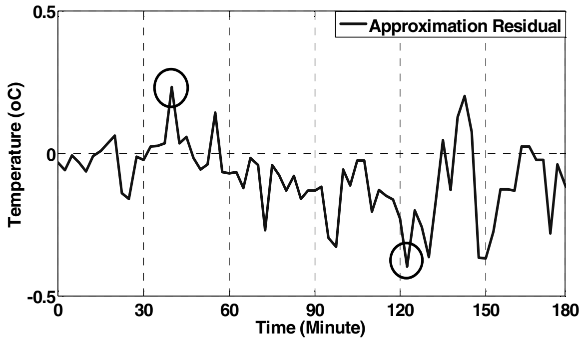

Fig. 11. Evaluation of approximated temperature

These results show that using the "least squares" and "artificial neural network" approaches separately have distinct advantages and disadvantages. Although in both approaches there are different parameters which could directly influence the performance of the approximation, such as using different numbers of equations in LS and application of different training parameters and coefficients for ANN, in this paper optimal parameters have been selected which lead to the most accurate approximation.

The LS approach begins approximation shortly after obtaining the first four records, and it continues approximating. The accuracy of the approximation depends on the relationship between the last four records; therefore, the accuracy could change depending on the last four records, and sometimes the approximation accuracy exceeds the maximum approximation error. However, the ANN approximation technique is established on learning according to an applied training algorithm and a defined architecture. It takes time to be capable of approximation, but after reaching a desired performance by means of evolutionary computation, the network begins approximation and its accuracy is better than the LS approximation technique. Thus, the main advantage of using LS approximation is that early approximation is possible in comparison with using ANN, although afterwards, based on evolution caused by training in ANN, the approximation accuracy is significantly better than the LS technique. By combining the LS approximation technique with the ANN technique, both advantages are obtained. In fact, the LS approximation could first be applied and upon reaching approximation capability in the neural network (based on approximation range), the temperature could be approximated using the LS technique. Subsequently, the ANN architecture will continue temperature approximation.

For classification, first the approximated temperature is compared with the actual value, and an approximation residual (ΔT) is generated.

$$\Delta T = T_{Actual} - T_{Network} \tag{6}$$

In (6), T_{Actual} shows the actual records of the "under test data logger" and $T_{Network}$ is the approximated temperature. If $|\Delta T|$<0.5 (°C), the actual value and approximation have a normal difference and the actual records are reliable. If 0.5<$|\Delta T|$<1.5 (°C), the classification mechanism could evaluate whether this difference comes from fault/failure in the system (such as from a battery problem or data logger defection) or from a weakness in approximation. This occasionally occurs when the LS technique is used for approximation.

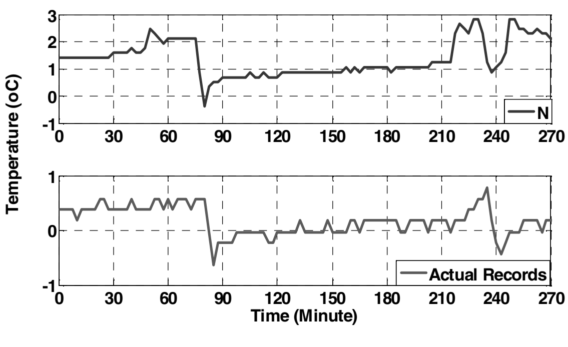

Fig. 12. Improving performance of classification technique

The value of 1.5 °C is derived from the sum of the data logger accuracy (±0.5 °C), the training accuracy (±0.5 °C), and the approximation accuracy (±0.5 °C), which is in the range of ±1.5 °C. To evaluate this case, the value of one of the neighbor data loggers could be used for decision making (Fig. 12).

$$D = |Corr(T, N)| \tag{7}$$

In (7), $Corr(T, N)$ is the correlation factor between the temperature of the "under test data logger" (T) and the records of the neighbor logger (N). D is the absolute value of the correlation factor.

If 0.5<$|\Delta T|$<1.5 (°C) and the value of D is more than 0.5, the correlation factor between the actual records and the values of the neighbor logger is significant and the records are classified in class (1).If 0.5<$|\Delta T|$<1.5 (°C) and the value of D is less than 0.5, the correlation factor between the actual records and values of the neighbor logger is very small, and the records are classified in class (2).

Table 1. Data classification table

Class	Specification	Records
0	$\lvert\Delta T\rvert$<0.5 (°C)	Reliable
1	0.5(°C) <$\lvert\Delta T\rvert$<1.5(°C) 0.5<D	Reliable
2	0.5(°C) <$\lvert\Delta T\rvert$<1.5(°C) D<0.5	Non-reliable
3	1.5(°C) <$\lvert\Delta T\rvert$	Non-reliable

If 1.5(°C) <$|\Delta T|$, the difference between the actual value and approximation is significant and the temperature records are not reliable based on the occurrence of a fault/failure, which is called "class (3)" in this application (Table 1).

IV. CONCLUSION

In this paper, two methods are applied and compared for temperature approximation of a real food transportation system including the "least squares" and "artificial neural network" approaches. The back-propagation algorithm is applied using the "multi layer perceptron" (MLP) architecture and the "least squares" is established based on a coefficient calculation of the last four temperature equations for approximating next values of temperature (over-determined system). In addition, by applying the classification algorithm, the reliability of the temperature records is evaluated.

ACKNOWLEDGMENT

This research was supported by the German Research Foundation (DFG), as part of the Collaborative Research Centre 637 on "Autonomous Cooperating Logistic Processes". We would especially like to thank the CCG Cool Chain Group Holding AG, CCG FRA and Rungis Express AG for providing all test facilities.

REFERENCES

[1] G. Betta, and A. Pietrosanto, ''Instrument Fault Detection and Isolation: State of the Art and New Research Trends,'' IEEE Transactions of Instrumentation and Measurement, 2000.

[2] G. Betta, M. Dell'Isola, C. Liguori, and A. Pietrosanto, "Expert systems for the detection and isolation of faults on low-accuracy sensor systems," in Proc. IEEE Workshop ET&VS-IM/97, Niagara Falls, Ont., Canada, pp. 39–48, May 1997.

[3] G. Betta, M. Dell'Isola, C. Liguori, and A. Pietrosanto, "An artificial intelligence-based instrument fault detection and isolation scheme for air conditioning systems," in Proc. XIV IMEKO World Congr., vol. VII, Tampere, Finland, pp. 88–95, June 1997.

[4] P. Balle, D. Fussel, and O. Hecker, "Detection and isolation of sensor faults on nonlinear processes based on local linear models," in Proc. Amer. Control Conf., Albuquerque, NM, June 1997, pp. 468–472.

[5] P. G. O'Reilly, "Local sensor fault detection using Bayesian decision theory," in UKACC Int. Conf. Control '98, Sept. 1–4, 1998, pp. 247–251.

[6] A. Zaknich, "Principles of adaptive filters and self-learning systems", ISBN 978-1-85233-984-5, 2005.

[7] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines and other kernel-based learning methods", Cambridge University Press, 2000.

[8] J. L. Crassidis and J. L. Junkins, "Optimal Estimation Of Dynamical Systems", Applied Mathematics and Nonlinear Science. Boca Raton, FL: CRC Press, Mar. 2004, vol. 2.

[9] S. Haykin, "Neural Networks", Englewood Cliffs, NJ: Prentice-Hall, 1998.

[10] Madan M. Gupta, Liang Jin, Noriyasu Homma, "Static and Dynamic Neural Networks: From Fundamentals to Advanced Theory", ISBN 0471427950, 2005.

[11] D. F. Specht and H. Romsdahl, "Experience with Adaptive Probabilistic Neural Networks and Adaptive General Regression Neural Networks", Proceedings of the IEEE World Congress on Computational Intelligence, 2, 1203-1208, 1994.

[12] C.P Lim, R. F. Harrison, "Probabilistic Fuzzy ARTMAP: an autonomous neural network architecture for Bayesian probability estimation" Artificial Neural Networks Fourth International Conference, 26-28 Jun 1995.

[13] L. C. Jain and V. R. Venuri, "Industrial Applications of Neural Networks", Boca Raton, FL: CRC Press, 1999.

[14] A. Jain, "Frost prediction using artificial neural networks: A temperature prediction approach," M.Sc. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA, 2003.

[15] Y.-J. Oyang, S.-C. Hwang, Y.-Y. Ou, C.-Y. Chen and Z.-W. Chen, "Data classification with radial basis function networks based on a novel kernel density estimation algorithm," IEEE Trans. Neural Netw., vol. 16, no. 1, pp. 225–236, Jan. 2005.

[16] R. P. Leger, Wm. J. Garland and W. F. S. Poehlman, "Fault detection and diagnosis using statistical control charts and artificial neural networks", Artificial Intelligence in Engineering 12, 1998, pp. 35-41.

[17] R. Jedermann, W. Lang, "Semi-passive RFID and beyond: steps towards automated quality tracing in the food chain", in: International Journal of Radio Frequency Identification Technology and Applications (IJRFITA), 1(2007)3, pp. 247 – 259.

[18] E. Vonk, L. C. Jain, and R. P., Johnson, "Automatic Generation of Neural Network Architecture using Evolutionary Computing", World Scientific Publishing Company, Singapore, 1997.

[19] R. Caruana and A. Niculescu-Mizil, "Predicting good probabilities with supervised learning," in Proc. Amer. Meteorology Conf. (AMS 2005), San Diego, CA, 2005, vol. 1.6.

Amir Jabbari received his M. Sc. in Mechatronics (2006) and B.Sc. in Electrical Engineering (Electronics) (2004). In August 2006, he started his research towards his PhD at the University of Bremen in "Application of Autonomous Fault detection and Isolation Mechanism for Measurement System". His main research activities are in data fusion, neural network and mechatronics.

Reiner Jedermann completed his Diploma in Electrical Engineering in 1990 at the University of Bremen. After working twice on embedded processing of speech and audio signals, he became a research associate in the Department of Electrical Engineering at the University of Bremen in 2004. He is currently writing his Ph.D. thesis on automated systems for freight supervision. His tasks within the CRC 637 research cluster comprise the analysis of application fields and the development of an embedded framework for software agents. Since 2006 he has been a member of the technical committee of the Cool-Chain-Association (see www.coolchain.org).

Walter Lang studied physics at Munich University and received his Diploma in 1982 on Raman spectroscopy of crystals with low symmetry. His Ph.D. in engineering at Munich Technical University dealt with flame-induced vibrations. In 1995 he became the head of the sensors department at the Institute of Micromachining and Information Technology of the Hahn-Schickard Gesellschaft (HSG-IMIT) in Villingen-Schwenningen, Germany, working with microsensors for flow, angular rate and inclination, and sensor testing and modeling. In February 2003 Walter Lang joined the University of Bremen and was subsequently appointed as a professor. He is the head of the Microsystems Center Bremen (MCB) and the head of the Institute of Microsensors, systems and actuators (IMSAS) in Bremen, Germany