Fuzzy and Probabilistic Models of Association Information in Sensor Networks

Leon Reznik
Department of Computer Science
Rochester Institute of Technology
102 Lomb Memorial Drive, Rochester, NY 14623-5608
Email: lr@cs.rit.edu

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso,
El Paso, TX 79968-0518
Email: vladik@cs.utep.edu

Abstract - **The paper considers the problem of improving accuracy and reliability of measurement information acquired by sensor networks. It offers the way of integrating sensor measurement results with association information available or a priori derived at aggregating nodes. The models applied for describing both sensor results and association information are reviewed with consideration given to both neuro-fuzzy and probabilistic models and methods. The information sources, typically available in sensor systems, are classified according to the model (fuzzy or probabilistic), which seems more feasible to be applied. The integration problem is formulized as an optimization problem.**

I. INTRODUCTION

Sensor networks are considered as one of the most promising emerging computing technologies. Distributed sensor networks' foreseeable applications will help protect and monitor military, environmental, safety-critical infrastructure and resources [1]. They are assumed to work primarily on the battlefield, in protection systems, biomedical and other security sensitive areas. Although military applications are the first ones coming into mind, many other opportunities await like applications requiring information gathering in harsh environments, detection of chemical or biological threats, monitoring water ant other resources reservoirs as well as healthcare conditions. Generally, the application of these types will require deployment of heterogeneous sensor networks including video cameras, acoustic equipment as well as sensors measuring different physical variables.

The design of such systems might be based on new computing and biological technologies. The key goal here, as with any other sensor system, would be improving their metrological and reliability characteristics, i.e. evaluation and reduction of the uncertainty of the results produced. This will enable efficient and reliable delivery of relevant data among distributed collection of sensors on a network.

In this paper we propose to investigate the issues for improving the reliability and accuracy of the decisions based on the application of the meta-level models (learning or constraint based), which can be deployed in sensor network environments. The meta-level model represents some sort of relationship or association between different sensors. The key components to develop such a system are:

1. Generation of high-level association models describing the concepts to be learnt. These models can depend on a variety of factors such as an expert opinion, the choice of the data mining technique employed and the type of data collected from the sensors on the network. Subsequently, machine learning techniques like genetic algorithms, neural networks and decision trees can be employed to generate a suitable model between different measurement variables. These models are then reinforced with information integrated from non-sensory external sources.
2. Modification of the models generated in the previous step to make them executable in real-time and applicable for joint processing with sensor data.
3. Application of these modified models for sensor validation, verification of the results and reducing the result uncertainty.

This paper is composed as follows. Section II gives some general information about sensor networks, while section III attempts to describe what we call association information. The variety of sources, where association information may be acquired from, assumes an applicability of different models and methods used in sensor information processing that includes both probabilistic and neuro-fuzzy approaches. Sections IV - VI review and develop further the previous author's works [2-5] on models, which can be applied for describing uncertainty of information sources used in association information gathering. Section VII attempts to provide a formal description of the problem of association information aggregation with measurement results [6-8].

II. SENSOR NETWORKS

Ad-hoc networks, particularly mobile ad-hoc networks most closely resemble the sensor network abstraction [9]. In the sensor network the data should be condensed as much as possible and only relevant data should be sent, this should be decided at the node level, the ad-hoc network sends information as designated by the user, whereas sensor networks send data as designated by the programmer and designer. Sensor networks must also be able to decide when a node is sending inaccurate or corrupted data, and then remove that sensor from the logical network map. Scalability is a large concern for sensor networks as some people visualize sensor networks of tens or hundreds of thousands of nodes. This is another argument for not maintaining identification for each distinct node. Large scale sensor networks employ more complex sink nodes that are distributed among the other sensors. These sink nodes operate kind of like local managers that can aggregate the data from all the local nodes and can process and compress the data into summary messages to send back to the upper network.

Typically sensor networks are constrained in terms of power consumption, computational capacity and reliability of the input data. It makes the measurement results taken from each particular sensor less accurate than with conventional separate sensors. Reliability of result delivery is pretty low also due to a variety of reasons (loss of power, hacker's attack, long route) as a sensor signal may not reach the processing unit.

The natural advantage of combining multiple sensors into networks is the possibility of organizing some form of collaboration between them. Significant improvement in coverage and reliability and a reduced false alarm rate could be achieved by fusing data from multiple sensors. Collaborative signal and information processing over a network is a new area of research and is related to distributed information fusion [9]. Fusion approaches range from simple rules of picking the best result to model-based techniques that consider how the information is generated.

III. Association Information

Association information origin. Association information represents a high level model of the underlying concepts of the object or processes under measurement. The association information could be expressed in a form of functional relationships between measured variables, or order relationship, may be approximate, stochastic or fuzzy. The following examples of possible functional relationships may be given:

1. in the measurement of flow rates in a variety of pipelines, which converge into one, we have the equation $Q_1 + Q_2 + \ldots + Q_n = Q$, where $Q_1, Q_2, \ldots, Q_n$ are the measurements results of the flow rates in converging pipelines and Q is the measured flow rate in the common pipeline.
2. measurement techniques applied in the fault tolerant skewed tolerant inertial measurement units, which are currently used in many aircraft and space systems [10], in which parity residuals indicative of the sensor errors are derived and then compared to calculated thresholds.

Model generation. The model could be provided or derived based upon either of the following or their combination:

1. knowledge of the physical, biological, chemical, mechanical or other natural laws, according to which an underlying object or process operates,
2. knowledge of the design and operational characteristics of the sensors, networks and other equipment
3. expert estimates
4. data mining and knowledge acquisition methods including
   a. statistical methods based on regression analysis and other techniques,
   b. intelligent data-driven methodologies, such as fuzzy logic, neural networks, genetic algorithms

IV. Classification of Uncertainty for Sensor Applications

Table on fig.1 displays a list of uncertainty sources and feasible corresponding models [11,12]. In sensor networks environment, uncertainty of the association information may occur in describing the following information about the object under measurement and/or environment:

- Spatial location of objects
- Attributes of objects (e.g. shape, color, visibility, personal characteristics)
- Attributes of environments (e.g. area, volume, boundary)
- Attribute relationship (e.g. more visible, better looking)
- Spatial relationships (e.g. direction, occlusion, topology)
- Typological attributes (e.g. group of objects, sub-objects)
- Temporal attributes and relationships (when an event occurs, and with respect to other events).

	Uncertainty source	Comments
A	Incomplete definition of the object or value	Incomplete definition may be caused by an impossibility or difficulty to compile an exact functional relationship or by the application of linguistic forms and rules. Fuzzy models may be particularly relevant here, also statistical depending on source.
B	Imperfect realization of a definition	Could be due to a number of factors, statistical models relevant to physical limitations of experiment, fuzzy models relevant to conceptual limitations.
C	Non-representative sampling	The sample may not represent the defined object because of a limited sample size, inhomogeneity of the object under measurement, etc. Sampling implies an underlying distribution and refers primarily to statistical models, but to the degree that the sample may not be representative, may involve fuzzy models.

D	Inadequate knowledge of the effects of the environment	1)the model of environmental conditions may not cover all the influence factors or 2) the model may be made under slightly different conditions because of the environmental changes statistical modeling 3) the model may be based on expert's estimates or guesses, fuzzy and statistical models
E	Personal bias in description	May be difficult to model mathematically as it depends on a particular person, may vary with time, etc. Applicable to both fuzzy and statistical models.
F	Limited accuracy (high inaccuracy) of the information available	Applicable to both statistical and fuzzy models.
G	Inexact values of standards and reference materials	Can be reference sample or reported value from a real underlying distributions (statistical model)
H	Inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm	Usually uncertainty, which is due to this reason, has a relatively small value in comparison to other components. In some cases, it can be single value from literature references (statistical square distribution or fuzzy model).
I	Approximations and assumptions incorporated in the design method and procedure	Difficult to evaluate with statistical methods, here fuzzy modeling approaches may be particularly relevant.

Figure 1. Classification of uncertainty sources and feasible models

V. Representation of uncertain information and uncertainty estimation in measurement science and system design

According to a real-life measurement practice, metrological characteristics can be given either as the limits of the allowed errors (deterministic form) or as the limits of the allowed values for some probabilistic and statistical characteristics (mean value, random component dispersion, confidence intervals). In both cases, a user does not really know how far from the given limits the actual values lay. Moreover, according to a number of standards the limits should be chosen from a given scale. Because of this, the limits are rounded down with the requested values being sometimes significantly lower than actually required. The probabilistic methods pretend to be objective. However, one can see that an application of the probabilistic and statistical methodologies in metrological analysis includes an assignment of a number of values such as the confidence level, the error probability, the significance level, etc. All those values are not calculated but assigned by an expert, who authored the corresponding document, or a person who performed metrological testing. All of these actually mean that real data about error characteristics are fuzzy in their nature. However, a user is recommended to consider those data as having probabilistic characteristics of some general aggregates, which are in turn assigned some probability distributions.

Considering all the information mentioned above the idea of measurement error formulation in terms of fuzzy systems theory looks rather reasonable. Some steps in this direction have been already made. In a number of the international [11,13,14] and national standards, the term "measurement error" has been replaced with the term "measurement uncertainty", which can be considered as more correspondent to fuzzy systems terminology. Publications, criticizing the probabilistic models applied in measurement science, are now followed by a number of works, trying to formulate those models from the fuzzy theory point of view or to combine both theories [15-22]. For example, in [6,7] a priori fuzzy information about the object under measurement is applied to increase the measurement accuracy and/or reliability. In order to apply the fuzzy sets and systems methodology in metrology and measurement practice, one has to prove that this methodology is able to perform mathematical and logical operations with fuzzy values, intervals and functions, typical for measurement science and practice. However, the problem of integrating both statistical and neuro-fuzzy approaches in measurement science and sensor system design still exists.

VI. Mathematical problem formulation

A conventional way of solving the problem of measurement result estimation assumes its definition as a mathematical programming problem and search for the parameter X' estimates by maximizing some criteria

$\hat{X} = \max_{x} F(Y_1, Y_2, ..., Y_n, X)$, where F() is a functional, whose shape is determined by the estimation methods,

Yi (i=1,n) is a set of m_i measurement results of the ith variable.

Let us consider a priori expert's information as a fuzzy constraint for the parameter vector X and given by the set of membership functions μ(f(X)). The methods of an expert's information acquisition and its propagation through are discussed in [20]. In this case the estimation problem with a priori information application can be considered as an optimization problem with fuzzy constraints. By now research of fuzzy constraints has accumulated different methodologies of solving such problems. One of the simplest and the most obvious way is a unification of both functional criteria and constraints into one synergetic criterion and looking for a global solution as the optimization of such criterion. So the problem can be re-formulated as search for the estimate minimizing the synergetic criterion

$$\widetilde{X} = \max_{X} F(Y1, Y2, \ldots, Yn, X) \times \mu(f(X))$$

This problem could be tried with conventional or intelligent methods. We will call the solution of this optimization problem a modified estimate and apply it as an estimate of the measured value modified with the help of expert's information. The method choice should depend on the estimation techniques applied as well as on the membership function shapes.(see [2,4] for more detail).

VII. Conclusion

Sensor networks are a new pervasive computing and measurement technology integrating measurement and processing units into complex network systems. The technological solution and its implementation make the problem of accuracy and reliability improvement in sensor networks much more important but at the same time provide new opportunities. Neuro-fuzzy methods and models may have found their way in this technology further development. The classification of typical models and methods applied for describing information available in sensor networks is given. The models are divided into two big classes: probabilistic and fuzzy. The proper place for an application of each of these groups is demonstrated. The methodology allowing integration of both model types is derived. The methodology is based on formulation of the measurement result processing as an optimization problem and solving it.

References

[1] Wood A.D. and J.A. Stankovich Denial of Service in Sensor networks, IEEE Computer, October 2002, pp. 54-62.

[2] L.Reznik and K.P.Dabke Measurement Models: Application of Intelligent Methods, Measurement, vol. 35, No.1, pp.47 – 58, 2004

[3] G. Solopchenko, V.Kreinovich, L.Reznik Development of Mathematical Structure of the Modern Measurement Science In: Measurement Science - a Discussion, K.Kariya and L. Finkelstein /Eds. Ohmsha Ltd. And IOS Press, Amsterdam, Oxford, Tokyo, Washington, 2000, ISBN 4-274-90398-2 and 1-58603-088-4, p.23-36

[4] L. Reznik What Is Intelligent Measurement? In: Computing with Words in Information /Intelligent Systems, L.A. Zadeh and J.Kacprzyk/Eds. Physica Verlag, Heidelberg – New York, 1999, ISBN 3-7908-1217-X, p.78 - 89

[5] L.Reznik and B.Pham Fuzzy Models in Evaluation of information Uncertainty in Engineering and Technology Applications, The 10th IEEE International Conference on Fuzzy Systems, December 2-5, 2001, Melbourne, Australia, ISBN 0-7803-7294-X

[6] L. Reznik Fuzzy Models of a priori Information in the Measurements' Results Processing, AMSE REVIEW, 1987, v.4, No. 2, p.1-12

[7] L. Reznik, G. Solopchenko Use of a priori Information on Functional Relations between Measured Quantities for Improving Accuracy of Measurement. Measurement, 1986, v.3, No. 2, p.61-69

[8] L.Reznik, V.Kreinovich and S.A.Starks Use of fuzzy expert's information in measurement and what we can gain from its application in geophysics Proceedings of the FUZZ '03, The 12th IEEE International Conference on Fuzzy Systems, St.Louis, May 25- 28, 2003, IEEE, vol. 2, pp. 1026- 1031

[9] C.-Y Chong and Kumar, S.P. Sensor networks: evolution, opportunities, and challenges Proceedings of the IEEE, Volume: 91 , Issue: 8, August 2003, pp. 1247 - 1256

[10] DeAngelis, F.E. A threshold redundancy management algorithm for skewed sensor arrays in: Proceedings., IEEE/AIAA/NASA 9th Digital Avionics Systems Conference, 1990., 10/15/1990 - 10/18/1990 , 15-18 Oct. 1990 pp. 133 – 140

[11] Guide to the Expression of Uncertainty in Measurement. International Organisation for Standardisation, Geneva, 1995.

[12] Reznik L., K.P. Dabke Evaluation of Uncertainty in Measurement: A Proposal for Application of Intelligent Methods in H. Imai/Ed. Measurement to Improve Quality of Life in the 21st Century, IMEKO –XV World Congress, June 13-18, 1999, Osaka, Japan, vol. II, p.93 – 1

[13] International Vocabulary of Basic and General Standard Terms in Metrology.-ISO, Geneva, Switzerland, 1993

[14] Quantifying Uncertainty in Analitical Measurement - Eurachem, 1995

[15] Gonella L. Proposal for a Revision of the Measure Theory and Terminology // Alta Frequenza, vol. XLIV. № 10, 1975.

[16] Destouches J.L., Fevrier P. New Trends in Expressing Results of Measurements // Proc. of IMEKO Coll., Budapest 1980.

[17] Roskrut A. Zapis teorii pomiaru sformulowany w oparciu o teorie zbiorow rozmytych // Postepu technologii maszyn i urzadzen , No.2, 1983 (in Polish).

[18] Otto M., Bandemer H. Calibration with Imprecize Signals and Concentrations Based on Fuzzy Theory // Chemometrics and Intellect. Laboratory Systems, № 1, 1986.

[19] Mari L. Notes on Fuzzy Set Theory as a Tool for the Measurement Theory // Digest of the 12th Triennal World Congress IMEKO, vol. III, Beijing, China.

[20] Mari L. The Role of Determination and Assignment in Measurement , Measurement, vol.21, No.3, pp.79-90, 1997

[21] Reznik L. Towards Fuzzy Logic Application in Measurement Theory 1st International Conference "Soft Computing and Measurement" (SCM'98), St.Petersburg, Russia, June 22-26, 1998

[22] Solopchenko G.N, L.Reznik, W.C.Johnson. Fuzzy Intervals as a Basis for Measurement Theory // Proc. of the First international Joint Conference of the North American Fuzzy Information Processing Society Biannual Conference, the Industrial Fuzzy Control and Intelligent Systems Conference, and the NASA Joint Technology Wofkshop on Neural Networks and Fuzzy Logic, pp. 405 - 406, San Antonio, USA, December 18-21, NASA, 1994

[23] Reznik L.. "Fuzzy controllers", Newnes-Butterworth-Heinemann, Oxford-Boston, 1997, ISBN 0-7506-3429-4, 287 pp.