ANALYSIS OF INTERVAL DATA FUSION METHODS FOR SENSOR ACCURACY IMPROVEMENT IN WIRELESS SENSOR NETWORKS

Lyudmila V. Galsanova
Scientific adviser – Lyudmila I. Khudonogova
National Research Tomsk Polytechnic University
lvg3@tpu.ru

Introduction

In wireless sensor networks (WSN) sensor nodes can provide measurement data with a large uncertainty due to battery failure and influence of external environmental factors. An important task is to determine the value of the measured quantity with the required accuracy based on inaccurate and/or incomplete data provided by the sensors. One solution to this problem can be fusion of data presented in the form of intervals on the real axis, which centers are measured values and boundaries are specified by the sensor's uncertainty.

Let us consider a set of m closed intervals $\{I_k\}$, $k = 1, \ldots, m$, in the real line. Each interval is characterized by a *lower bound* l_k, *upper bound* u_k and *middle point* x_k, so that $I_k = [l_k, u_k]$; $l_k < x_k$; $x_k = 0{,}5 \cdot (u_k + l_k)$; l_k, u_k, $x_k \in \mathbb{R}$ (Figure 1). Let the measured value of some value of each sensor node in WSN be represented in the form of interval, where the point x_k is measurement result of the sensor of k-th node, the *uncertainty interval* of which is $\left[x_k - 0{,}5 \cdot (u_k - l_k),\ x_k + 0{,}5 \cdot (u_k - l_k)\right] = \left[x_k - \varepsilon_k,\ x_k + \varepsilon_k\right]$, and the value of ε_k is determined by the sensor's uncertainty. Two intervals I_j and I_k, $j \neq k; j,\ k = 1, \ldots, m$, such that $I_j \cap I_k \neq \varnothing$, will be called *inconsistent* [1].

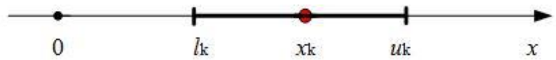

Figure 1 – An interval on the real line

The *interval data fusion procedure* consists in forming an interval that is consistent with maximal number of intervals from $\{I_k\}$ and contains the maximum likelihood value x_r, which can serve as a representative of all middle points x_k with a minimum uncertainty ε_k [2]. Interval measurement data processing in order to obtain the value of the measured value with increased accuracy can be implemented by means of fusion based on voting.

The aim of this work is the development of software for investigation of various algorithms of interval data fusion to determine the most accurate and robust (i.e. stable to the form of the distribution law and possible outliers) algorithm.

1 Data fusion on the basis of approval voting

Voting is one of the promising ways to obtain the correct result from unreliable, incomplete, or inaccurate data. In approval voting, a participant votes for a subset of preferred alternatives, instead of choosing only one "best" alternative. A special case of approval voting is *interval voting*, when an ordered set of preferred alternatives forms interval along the real line, and all values within the interval are considered equally acceptable for the participant [3].

The choice of the voting selection rule has a significant effect on the reliability and rate of the algorithm. The interval data fusion can be based on plurality rule and majority rule.

According to the **plurality rule**, the outcome consists of a set of alternatives that are preferred by the largest number of participants. Consider an ordered list of alternatives $\{a_1, a_2, \ldots, a_n\}$ and a finite set $\{p_1, p_2, \ldots, p_m\}$ of participants of the voting process. The initial data from the sources (voters) are presented as intervals I_k with the upper and lower boundaries [l_k, u_k]. The resulting interval I_r is the intersection of the intervals of the biggest number of participants. Figure 2 shows the example of interval voting with plurality rule. In practice, this voting strategy is not always so simple and a number of complications may arise, associated with the multiplicity of resulting intervals [3].

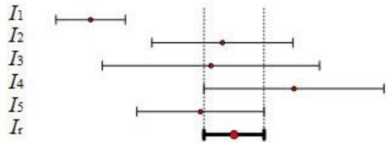

Figure 2 – Interval voting based on plurality rule.

In accordance with the **majority rule**, an alternative needs to get more than half of the votes (minimum 50 % + 1 vote) to win. To implement the rule with regard to interval data, it is necessary to eliminate all nonqualifying intervals, i.e. such that the number of intersections with other intervals is less than the value of y – half of the participants number. The example of interval voting with majority rule is shown in Figure 3. The disadvantage of the majority rule is that the resulting interval is often wider than the initial ones, which leads to increased uncertainty of the result.

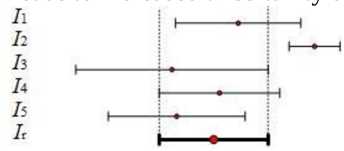

Figure 3 – Interval voting based on majority rule.

2 Software for investigation of interval data fusion algorithms

To research the properties of fusion algorithms based on considered voting rules, a program called "Fusion algorithm" was developed in the NI LabVIEW graphical environment.

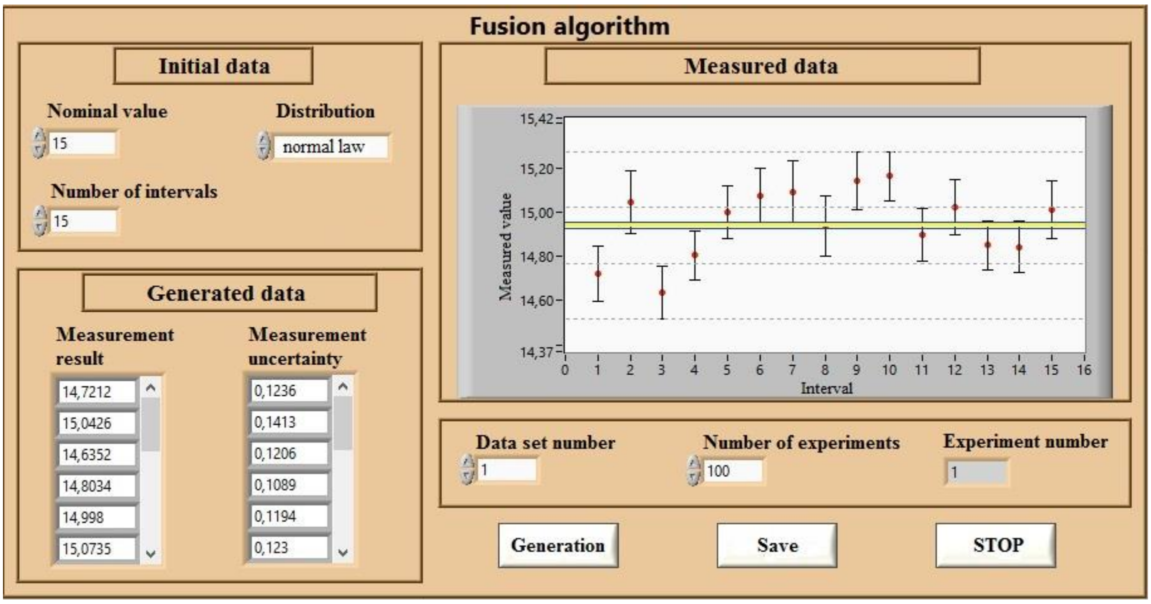

Figure 4 – The front panel of the program "Fusion algorithm"

The program "Fusion algorithm" performs the following functions:

- generation of random interval data on the normal and uniform distribution laws;
- fusion of intervals based on plurality and majority selection rules;
- visualization of results in graphical and numerical forms;
- saving experimental results in MS Excel (*.xls) document format.

The generation of random interval data is implemented using Monte Carlo method [4]. The generated data is fused by means of algorithms based on plurality and majority rules, for which the "PLURALITY" and "MAJORITY" software modules were developed.

Before start on the front panel of the program (Figure 4) the user must specify:

- nominal value of the measured quantity;
- number of intervals;
- distribution law of the quantity;
- number of data set;
- number of experiments.

Next, the user starts the program by pressing the "Generation" button. The front panel displays the generated measured data (measurement results with corresponding uncertainties) in graphical and numerical forms. The program finishes operation after pressing the "STOP" button. When work is finished the user can save the data by clicking the "Save" button. As a result of the "Fusion algorithm" program, the resulting interval $I_r = [x_r \pm \varepsilon_r]$ is found for each fusion algorithm, where the mean point x_r is the result of the fusion, and ε_r is the corresponding uncertainty of the result.

Experimental studies were implemented using the developed program "Fusion algorithm" for considered voting algorithms with plurality and majority rules on the basis of the same initial data.

Conclusion

To solve the problem of finding the value of the measured value on the basis of inaccurate data provided by the sensors in WSN, algorithms for interval data fusion on the basis of voting with plurality and majority rules were researched. A "Fusion algorithm" program was developed for experimental investigations of algorithms in LabVIEW graphical programming environment. The developed program allows to estimate such properties of algorithms as accuracy and robustness as well as their time complexity.

References

1. Muravyov S. V., Khudonogova L. I., Marinushkina I. A. Representation of interval data by weak orders yields robustness of the data fusion outcomes // Journal of Physics: Conference Series. – 2016. – N 1. – P. 360-365.

2. Muravyov S. V., Khudonogova L. I. Multisensor accuracy enhancement on the base of interval voting in form of preference aggregation in WSN for ecological monitoring // Proceedings of the ICUMT 2015. – Czech Republic, 2015. – P. 293-297.

3. Parhami B. Distributed Interval Voting with Node Failures of Various Types // Proc 12th IEEE Workshop on Dependable Parallel, Distributed and Network-Centric Systems DPDNS '07 Long Beach, California USA. – 2007. – P. 1-7.

4. ГОСТ Р 54500.3.1-2011/Руководство ИСО/МЭК 98-3:2008/Дополнение 1:2008. Неопределенность измерения. Часть 3. Руководство по выражению неопределенности измерения. Дополнение 1. Трансформирование распределений с использованием метода Монте-Карло официальное издание. – М.: Стандартинформ, 2012. – 82 с.