Multi Source Information Matching and Fusion Method in Wireless Network Based on Fuzzy Set

Quan-wei Sheng(✉)

Changsha Medical University, Changsha 410219, China

Abstract. The current wireless network information fusion matching method is based on the calculation of the uncertainty of a large number of data samples, and gets the information fusion results according to the corresponding decision. However, due to the uncertainty of data samples, it is difficult to complete statistics, which leads to low processing accuracy and poor information fusion effect of information fusion matching method. In order to solve the above problems, a multi-source information matching and fusion method based on fuzzy set is studied. After matching with multi-source information, Markov distance is used to allocate the information. According to the probability of supporting information conflict, information fusion rules are formulated, and multi-source fusion is realized by using fuzzy sets. Simulation results show that the proposed information fusion method can reduce the energy consumption of network nodes by 81.4%, and has high fusion accuracy and reliability.

Keywords: Fuzzy set · Wireless network · Multi-source information · Matching fusion

1 Introduction

Multi source information fusion is also called multi-sensor information fusion. It comprehensively analyzes the incomplete environmental information collected in local environment in various forms and ways, and processes the redundant or contradictory information between the collected information in many aspects, so as to achieve the analysis of information complementarity and uncertainty of information, and further to the overall system The lines are expressed in a complete and consistent way [1–3]. The accuracy of planning, decision-making and emergency response can be enhanced by multi-source information fusion, thus reducing the risk brought by decision-making. The multi-source information fusion can improve the robustness and reliability of the system compared with the single source information. The robustness and reliability of a system relying on single information is relatively poor. If the information source fails (sensor failure, transmission delay, signal distortion, etc.), the performance of the whole system will be fatally affected, that is, it can not work normally. With the increasing amount of information obtained by multi-sensor system, it leads to different factors in time domain, space field, credibility, expression mode, information emphasis and practical application, and puts forward new requirements for information processing and

S. Liu and X. Ma (Eds.): ADHIP 2021, LNICST 417, pp. 51–61, 2022.
https://doi.org/10.1007/978-3-030-94554-1_5

work management. At present, no matter in the theoretical method and implementation technology of research, the technology of information fusion needs to be opened up. When multi-source information fusion is combined, the uncertain information provided by multi-sensor system is expressed in probability by Bayes estimation principle. Secondly, the independent decision elements are divided reasonably in a whole sample space. Then, the decision elements can be processed by Bayes conditional probability formula. Finally, the decision elements can be processed by using Bayes conditional probability formula According to the given rules, the system decision is output, and the final decision of the system is generally taken as the maximum posterior probability. The method needs to give the prior probability distribution of target information of each sensor through a large number of experiments, which is difficult to achieve in practice. In addition, when the results of decision-making are increased, the calculation of Bayes estimation will be very complex to affect the real-time performance of the processing system. The robust hypothesis is used to verify whether the test data is consistent and the information matching fusion is realized by using the statistical decision theory. Because of the fusion estimation based on the maximum and minimum decision of Luban, there are some limitations in the fusion of uncertain information. The data fusion based on Markov chain is a linear weighting method by using multiple sensors' collected values. It is necessary to measure the data with common observation values, which is difficult to measure, and affects the accuracy of information matching fusion. In the process of multi-source information collection, because of many reasons, information contains a lot of uncertainty, and fuzzy sets can reflect the uncertainty to the maximum extent in the reasoning process, thus optimizing information matching fusion. Based on the above analysis, this article will study the method of wireless network multi-source information matching and fusion based on fuzzy sets, and verify the feasibility of the method through simulation experiments, and solve the problems in the traditional method.

2 Multi Source Information Matching and Fusion Method in Wireless Network Based on Fuzzy Set

2.1 Reliability Allocation of Multi Source Information in Wireless Networks

In this paper, Markov distance is used as an important basis for reliability allocation, and the corresponding basic reliability distribution function is needed to map the relationship between distance and reliability. The Markov distance between multi-source information in wireless network is defined as follows:

If M is set as n element population (n indexes are investigated in total), the mean vector is $\mu = (\mu_1, \mu_2, \cdots, \mu_n)$ and covariance matrix is $\sum \left(\sigma_{ij}\right)_{n\times n}$, the calculation formula of Markov distance between sample $X = (X_1, X_2, \ldots, X_n)$ and population M is as follows [4]:

$$d^2(X, M) = \frac{(X-\mu)^2}{\sigma^2} \tag{1}$$

Mahalanobis distance has the following characteristics: (1) Mahalanobis distance between two points is independent of measurement unit; (2) Mahalanobis distance between two points calculated by standardized data and centralized data is the same; (3) mutual interference between different variables can be effectively eliminated [5].

The selection of the basic reliability allocation function based on Mahalanobis distance needs to meet the following two main principles

(1) The final distribution result sum of all reliability is 1, and the single distribution result is greater than or equal to 0 and less than or equal to 1;
(2) The distribution function should be a monotonic decreasing function, that is, with the increase of Mahalanobis distance between the target and the known sample, the reliability of the distribution function gradually decreases;

The common distribution functions satisfying the above two basic principles are inverse trigonometric function and exponential function, as shown in the following formula [6]:

$$\begin{array}{l} y = \frac{1}{2} - \frac{1}{\pi}\arctan\left(\frac{x}{a} - b\right), a, b > 0 \\ y = a^{bx}, 0 < a < 1, b > 0 \end{array} \tag{2}$$

In the above formula, a and b are the parameters of the distribution function. Combined with the distribution trend of the basic reliability distribution function of the two functions, the exponential function is selected as the basic reliability distribution function model based on Mahalanobis distance. According to 3 σ principle, the specific values of distribution function parameters a and b are determined. According to the monotonicity of the exponential function, when the Mahalanobis distance between the tested target and the sample set is greater than 3 times of the standard deviation of the known sample data set, the initial reliability allocation result is less than ε value, so the mapping accuracy of the basic reliability allocation function can be adjusted by setting the threshold value ε. Therefore, there are the following relationships:

$$\varepsilon = a^{3\sigma} \tag{3}$$

In order to make the exponential basic reliability distribution function proposed in this paper better adapt to the statistical distribution of different sample data, the distribution function is transformed into the following form:

$$y = \left(\varepsilon^{1/(a\sigma_{\max})}\right)^{x}, 0 < \varepsilon < 1 \tag{4}$$

In the above formula, parameters a and ε are the adjustment coefficients of the basic reliability allocation function; x is the Mahalanobis distance between the target to be tested and each known sample data set; y is the mapped initial reliability allocation result. According to the process shown in the figure below, the reliability of multi-source data in wireless network is allocated [7] (Fig. 1).

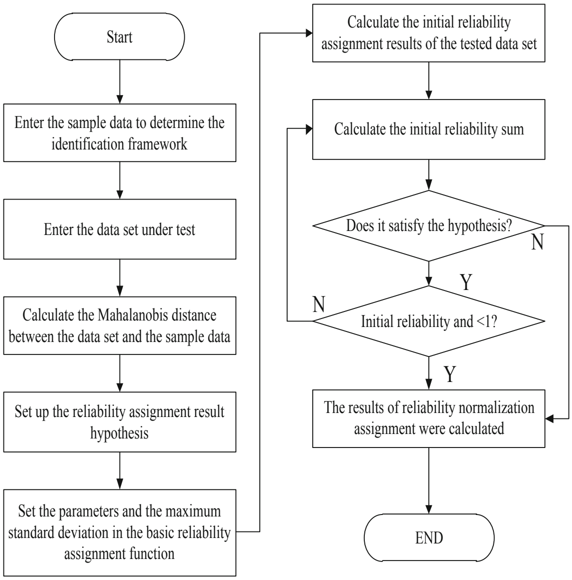


Fig. 1. Reliability allocation process of wireless network information

After the reliability of multi-source information is assigned, the matching relationship between wireless network information and entities is established to facilitate information fusion.

2.2 Multi Source Information and Wireless Network Entity Matching

The process of candidate pair generation based on attribute index can be divided into three parts: inverted index construction, attribute weight calculation and candidate entity pair generation.

Inverted index construction traverses all entity k in multi-source data source, and then traverses attribute p in k. With the attribute name and attribute value as the composite key and the index value as the collection of entity ID, the mapping relationship conforms to the following form [8]:

$$(p.name, p.value) \rightarrow \{ID_1, ID_2, \cdots, ID_n\} \tag{5}$$

In the above formula, $p.name$ is the name of the data source entity property; $p.value$ is the value of the data source entity attribute. The attribute name and the attribute value can reflect the property's discrimination. Define the weight of the attribute as follows:

$$W_{(p.name,p.value)} = \frac{1}{|Svc|} \tag{6}$$

In the above formula, $|Svc|$ represents the number of entities with the same attribute name and corresponding same attribute value; $W_{(p.name,p.value)}$ represents the entity attribute weight. The weight of different attributes can be calculated by inverted index.

For each group of ID set, enumerate any two entity pairs, give a certain weight $W_{(p.name,p.value)}$ or accumulate on the basis of the original weight. Finally, the entity pairs with all weights are traversed. If the accumulated weight of the entity pairs is greater than a certain threshold, they are regarded as candidate entity pairs [9].

The similarity between the two candidate entities is calculated according to the following formula:

$$\begin{aligned} sim(k_1, k_2) = &\omega_1 \times sim(k_1.N, k_2.N) \\ &+ \omega_2 \times sim(k_1.P, k_2.P) + \omega_3 \times sim(k_1.T, k_2.T) \end{aligned} \tag{7}$$

In the above formula, $sim(k_1.N, k_2.N)$ is the similarity of entity name; $sim(k_1.P, k_2.P)$ is the similarity of entity attribute; $sim(k_1.T, k_2.T)$ is the similarity of entity context; ω_1, ω_2 and ω_3 represent the corresponding weights of the three. According to the descending order of the similarity between the candidate entities, the matching relationship between the wireless network multi-source information and the entities is determined. Combined with the matching relationship between entities, the basic synthesis rules of multi-source information fusion are formulated according to the D-S theory.

2.3 Making Rules of Multi-source Information Fusion

In this paper, we use the idea that the probability of supporting information conflict is distributed according to the average support degree of each information, and obtain the following composition rule [10]:

$$\begin{cases} m(A) = \sum\limits_{\substack{ij \\ \cap A_i = A}} \prod\limits_{j=1}^{m} m_j(A_i) + rq(A) \\ \forall A \subset U, A \neq \Phi \\ m(\Phi) = 0 \\ r = \sum\limits_{\cap A_i = \Phi} \prod\limits_{j=1}^{m} m_j(A_i) \\ q(A) = \frac{1}{n} \sum\limits_{i=1}^{n} m_i(A) \end{cases} \tag{8}$$

Where $q(A)$ is the average support degree of proposition A and r is the conflict probability. In order to improve the shortcomings of the above information fusion rules, the correlation coefficient between information is introduced to deal with conflict information, and the correlation coefficient between information can be calculated according to the following formula:

$$d_{12} = \frac{\sum\limits_{A_i \cap B_j = A_k \neq \Phi} m_1(A_i) m_2\left(B_j\right)}{\sqrt{\left(\sum m_1^2(A_i)\right)\left(\sum m_2^2\left(B_j\right)\right)}} \tag{9}$$

In the above formula, m_1 and m_2 are the basic trust allocation functions of two conflict information E_1 and E_2 under the D-S theory recognition framework, and A_i and B_j are the units corresponding to the basic trust allocation functions m_1 and m_2 respectively. Correlation coefficient d_{12} is used to describe the similarity between two information. The greater the value of d_{12}, the better the certainty. When d_{12} is 1, it indicates that the two information E_1 and E_2 are completely similar; when d_{12} is 0, it indicates that the two information E_1 and E_2 are completely conflicting.

It is defined that the amount of information provided by an information is related to the focus element produced by it, then the calculation formula of information capacity provided by information under D-S theory is as follows:

$$C(E) = \sum_{i=1}^{n(A_i)} \frac{m(A_i)}{|A_i|} \tag{10}$$

In the above formula, $|A_i|$ is the cardinality of the focus element; $n(A_i)$ is the number of focus elements. For the focus element A_i, if $m(A_i)$ is 0, the information capacity is 0, indicating that this information does not contain any useful information. If $|A_i|$ is 1, the information capacity is 1, indicating that this information contains the most useful information. Therefore, before information fusion, the information capacity of each information is calculated, and then normalized to get the rule of multi-source information fusion in wireless network. According to the rules of wireless network multi-source information fusion, the fuzzy set theory is used to realize the fusion of multi-source information in wireless network.

2.4 Realize the Multi-source Information Matching and Fusion in Wireless Network

In order to solve the problem of multi-sensor data fusion, the observation data of the sensor needs normalization (fuzziness) processing (fuzziness), which corresponds to the range of [0, 1]. The theory of fuzzy set provides a normalized tool for people, that is, the fuzziness process of sensor observation information can be realized by constructing appropriate membership function.

The membership degree $u \in [0, 1]$ of fuzzy set is introduced to represent the local decision value of sensor. For the measured value x of each sensor, the membership function $u(x)$ can be used to map it to a number u (membership degree) in the interval [0,1]. The size of u reflects the support degree of the sensor for a certain hypothesis, which can also be called basic probability assignment. Here are some membership functions commonly used in fuzzification, but in practical application, we should build practical membership functions according to specific problems.

For the sensor that can determine the mean value $\bar{x}$ and the deviation δ of the parameter, if the measured value is, the membership function can be used:

$$u(z) = \begin{cases} 1 - \frac{|z-\bar{x}|}{2\delta}, |z - \bar{x}| < 2\delta \\ 0, |z - \bar{x}| \geq 2\delta \end{cases} \tag{11}$$

2 For multiple groups of measured values, the maximum value $\max f$ and minimum value $\min f$ can be determined. If the measured value is $f(u_i)$, the membership function can be used as follows:

$$M_f(u_i) = \frac{f(u_i) - \min f}{\max f - \min f} \tag{12}$$

If multiple sensors are used to detect multiple objects, and if the number of sensors in the number of sensors supported for an object can be determined as n, the membership function can be used as follows:

$$A(u_i) = \frac{n}{N} \tag{13}$$

Where u_i is the i object. Simultaneous interpreting the membership function of fuzzy sets according to different sensors' acquisition of multi-source information. The fuzzy similarity matrix is established to calculate the support and reliability of different information. The reliability is taken as the weight, and the average basic trust function of information is obtained by weighted average. According to the information object fusion rules constructed by D-S theory, the final data fusion result is obtained by n-1 fusion operation on N pieces of information. So far, the research of multi-source information matching and fusion method based on fuzzy set is completed.

3 Simulation Experiment

The multiple information sources of wireless network make it more difficult to effectively manage the information of wireless network, and cause a large number of effective data abuse and loss of sensor. In this paper, a multi-source information matching and fusion method based on fuzzy set is proposed. This section will verify the effectiveness of the method.

3.1 Experiment Content

The simulation experiment was carried out in NS 2 test software. In order to reflect the superiority of the above proposed fusion method based on fuzzy sets, the fusion matching method studied above is compared with Bayes estimation based information fusion method and Markov chain based data fusion method. The experiment mainly compares the energy consumption and accuracy of the network nodes when the information is transmitted after fusion. Through the analysis of the experimental data, the information fusion effect of the method is studied.

3.2 Experimental Process

The information fusion matching method based on fuzzy set is named as method 1, the information fusion method based on Bayes estimation is named as method 2, and the data fusion method based on Markov chain is named as method 3. Three different methods are used to fuse the information in the experimental network. The whole process is monitored by the test software and the experimental data are obtained.

The energy consumption of information transmission between nodes is an important index to evaluate the fusion method. The data of each communication node in the network detected by the network communication monitoring software is analyzed, and the final experimental conclusion is obtained.

3.3 Experimental Result

Three kinds of information fusion methods are applied in the experimental network to fuse the information transmitted in the network. The monitoring software detects the energy consumption of network communication nodes, and obtains the comparison results of node energy consumption as shown in the following table (Table 1).

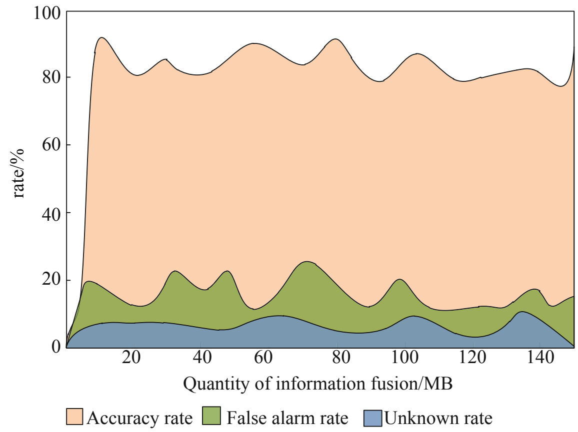
Table 1. Energy consumption comparison of network communication nodes/mJ

Wireless network node	Method 1	Method 2	Method 3
1	1.71	9.41	11.48
2	1.67	8.73	12.30
3	1.73	9.30	10.51
4	1.81	9.22	12.13
5	1.75	8.85	17.96
6	1.68	8.41	14.52
7	1.64	8.83	17.33
8	1.66	9.64	11.49
9	1.79	9.27	9.84
10	1.63	8.96	11.91
11	1.72	9.63	9.93
12	1.69	9.82	10.11
13	1.83	9.81	10.19
14	1.67	9.34	11.88

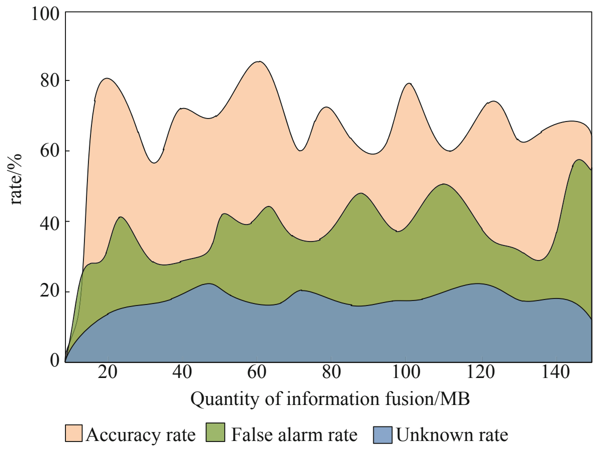
According to the analysis of the above table, the energy consumption between different nodes in method 1 is lower than that in method 2 and method 3. The difference of energy consumption of different nodes in the wireless network of application method 1 is small, while the difference of energy consumption of different nodes in the network of application method 3 is large, and the difference of energy consumption of nodes in the wireless network of application method 2 fluctuates in a certain range.

Calculate the average energy consumption of wireless network nodes in the process of this experiment when transmitting the information fused by different methods. The average energy consumption of wireless network nodes in application method 1 is 1.71 MJ, that in application method 2 is 9.23 MJ, and that in application method 3 is 12.26 MJ. Application method 1 can reduce the energy consumption of 81.4% at least.

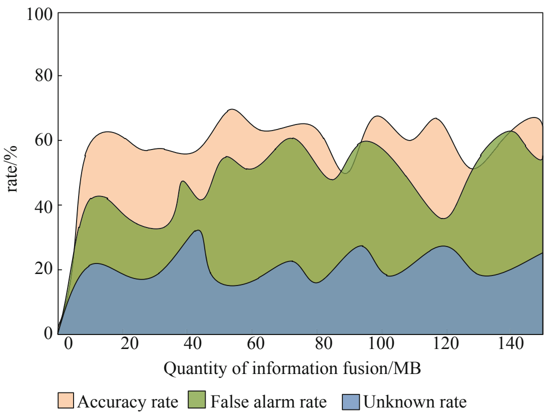
The accuracy comparison of three wireless network information fusion methods in information fusion is shown in Fig. 2.



(a) Method 1 Accuracy



(b) Method 2 Accuracy



(c) Method 3 Accuracy

Fig. 2. Comparison results of information fusion accuracy

Analysis of the above figure shows that the correct rate, false alarm rate and unknown rate of method 1 are lower than those of method 2 and method 3. With the increase of the amount of information to be fused in wireless networks, the accuracy of method 1 fluctuates slightly in a certain range, while method 2 and method 3 fluctuate sharply. It shows that the accuracy and stability of method 1 are better than the other two methods.

To sum up, compared with traditional methods, the multi-source information matching and fusion method based on fuzzy set in this paper can reduce node energy consumption by about 81.4% on average, with high accuracy, strong stability and reliability.

4 Concluding Remarks

With the increasing demand for various kinds of information, multi-sensor information fusion technology has become an effective means of data processing and decision support. Because of the redundancy of information between different kinds of sensors, the system has better fault tolerance ability. Aiming at the problems of the current common information fusion methods, this paper studies the multi-source information matching fusion method based on fuzzy set. After the information is distributed by Mahalanobis distance, the multi-source information and entities are matched. According to the probability of supporting information conflict, information fusion rules are formulated, and multi-source fusion is realized by using fuzzy sets. The feasibility is verified by simulation experiment. In the future research, the experimental parameters will be further changed to improve the performance of the information fusion method.

References

1. Shi, Z., Zhou, X., Li, K., et al.: Cyberspace security monitoring technology based on multi-source information fusion. Comput. Eng. Des. **41**(12), 3361–3367 (2020)
2. You, H., Shi, H., Yang, Y., et al.: Intelligent fusion diagnosis method for multi-source information of power grid fault based on improved D-S evidence theory. Guangdong Electric Power **33**(11), 16–25 (2020)
3. Hu, X., Zhang, W., Xu, J.: A method of multi-source data fusion based on absolut grey fusion algorithm. Electron. Warfare Technol. **35**(05), 37–41 (2020)
4. Ke, J., Lu, X.: Intelligent recommendation algorithm for social networks based on fuzzy perception. Inf. Technol. **44**(02), 130–134 (2020)
5. Liu, Y., Hui, H., Lu, Y., et al.: Multi-source information fusion indoor positioning method based on genetic algorithm to optimize neural network. J. Chin. Inertial Technol. **28**(01), 67–73 (2020)
6. Xiao, L., Luan, X.: Predictive control method of matching deviation quality for car door based on data fusion method. Agric. Equip. Veh. Eng. **57**(12), 29–34 (2019)
7. Ge, X.: Multi-channel information fusion method of wireless network based on grid system. J. Jilin Univ. (Inf. Sci. Ed.) **37**(06), 617–622 (2019)
8. Liu, S., He, T., Dai, J.: A survey of CRF algorithm based knowledge extraction of elementary mathematics in Chinese. Mob. Netw. Appl. (2021). https://doi.org/10.1007/s11036-020-01725-x

9. Liu, S., Pan, Z., Cheng, X.: A novel fast fractal image compression method based on distance clustering in high dimensional sphere surface. Fractals **25**(4), 1740004 (2017)
10. Liu, S., Fu, W., He, L., Zhou, J., Ma, M.: Distribution of primary additional errors in fractal encoding method. Multimed. Tools Appl. **76**(4), 5787–5802 (2014). https://doi.org/10.1007/s11042-014-2408-1