

Research on The Credibility of Wireless Sensor Networks in Multi Dimensional Data

WANG Guotao, SHI Zhenguo, WU Xiaojing

School of Computing , Nantong University, Nantong , Jiangsu 226019, China

290292149@qq.com

Abstract— The application of wireless sensor networks is almost involved in all industries. However, due to the limitations of technology, the wireless sensor has many weaknesses. In the process of application, the sensor nodes can't get accurate data from the environment because of the interference of various factors, then the real fact of the environment won't be reflected, the whole system can't work properly. In this paper, I propose a method called MDRE to evaluate the credibility of sensor nodes which based on the modified D-S evidence theory. Lots of experiments make it clear that the method can evaluate the credibility of the node more accurately.

Keywords—WSN ; credibility ; Modified D-S Evidence Theory; Deviation degree; Multidimensional data; Basic distribution function

I. INTRODUCTION

The rapid development of wireless sensor network (WSN) [1][2] has promoted the application of wireless sensor networks in various fields. However, the data monitored by sensor network is often inaccurate because of the vulnerability of individual sensor node [3]. In this case, the research on the credibility of sensor network nodes is definitely important. Credibility is a collective term used to express the availability, reliability, maintainability, security, and sometimes in the general description of non quantitative terms. The credibility problem originates from social science which is used to describe the trust between people and people, and then is introduced into the economic field [4] [5], Marsh proposes the calculation model of credibility firstly [6]. Since then the credibility problem has been applied to the Internet industry, such as the mutual trust between buyers and sellers in e-commerce, and researchers propose a peer's calculation model of network[7][8]. With the continuous development of wireless sensor networks, the research on the credibility of the wireless sensor network is getting more and more attention.

Most of these schemes which are proposed on the credibility of sensor network nodes are based on evaluated by the analysis of the communication between nodes and the behavior of the node itself, so it is easy to bring about the network overload, and consume a large amount of energy and storage resources. At present, most of the Fault Diagnosis of sensor network nodes are based on Sink node. By introducing the related protocol, the nodes will send its own state information periodically to the Sink node [9] [10]. However, this method increases the burden of network traffic. Another

method is detecting the running state of the nodes by inserting the inspection points into the sensor nodes [11]. There is a better method: self diagnosis, which is used to detect the faults of the system [12]. This method effectively reduces the bandwidth of network communication between nodes, and prevents the loss of information from the process of Sink node. However, it has brought a lot of tests for the duration of the sensor nodes.

This paper proposes a new method to evaluate the credibility: modified D-S evidence theory which based on the deviation degree, named MDRE, it can greatly improve the performance of the sensor network through evaluating the credibility of nodes by the research of the multi dimensional data collected by the sensor network nodes. This paper uses the D-S evidence theory to measure the credibility of the nodes because the Dempster-Shafer evidence theory [13]-[15] has a unique advantage in the process of dealing with uncertain information. However, the traditional D-S evidence theory is insufficient in the combination rule and the decision-making stage, so it needed to be improved. The modified evidence theory is combined with the historical data, the correlation of the data from each dimension, and the basic probability distribution which is calculated based on the deviation degree. At the same time, the calculation of weight of the credibility of the nodes are based on the deviation degree. These optimizations can greatly reduce the influence of the deviation of the large dimension data.

II. DESCRIPTION OF THE PROBLEM

Sensor network is widely used, its structure diagram as shown in figure 1. A sensor node usually collected different data, and the collected data is transmitted to the management center.

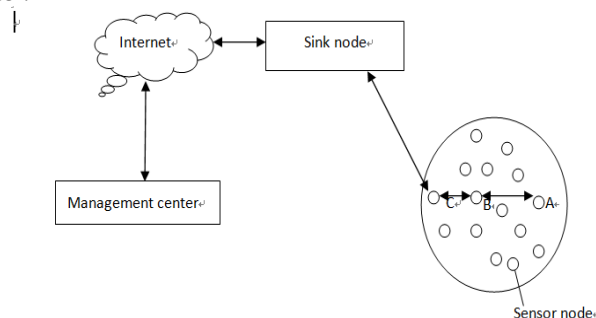


Figure 1. Picture of sensor network

Suppose that a wireless sensor network is arranged in a real environment, and the number of nodes is N, each node is denoted as $\{A_1, A_2, \dots, A_n\}$, these nodes can get M type data from environment, such as node A_h get the data $\{A_{h1}, A_{h2}, \dots, A_{hm}\}$. The nodes in sensor network are report the collected data to the remote monitoring center at all times. When the monitoring center finds that the data is abnormal, it will measure the credibility of the node, which can help us to find the presence of an abnormal node.

III. ALGORITHM: MDRE

A. Calculate Deviation Degree

Suppose that all nodes in the sensor network can collect the data from the i dimension, each data is denoted as $\{A_{1i}, A_{2i}, \dots, A_{ni}\}$, such as get the temperature in the environment, these data reflect a property in the region, so they are very similar. The mean value of the data on the i

dimension is denoted as $\bar{A}_i = \frac{1}{n} \sum_{k=1}^n A_{ki}$, $\alpha = \frac{|A_{ki} - \bar{A}_i|}{\bar{A}_i}$ denotes as the deviation degree of node K on i dimension data.

B. Modified D-S Evidence Theory

D-S evidence theory is proposed by Dempster in 1967. he uses the definition of upper and lower bounds of probability, and then his student Shafer promote it, make it systematic and theoretical, and to be a kind of theory of uncertainty reasoning, that is D-S theory. The core of this paper point out its shortcomings in dealing with the problem of the credibility of sensor nodes in multi dimensional data environment, and improve it.

Suppose that $\{A_{k1}, A_{k2}, \dots, A_{km}\}$ is the m dimension data detected by node K. Assuming A_{ki} is a suspicious data (which is also possible for several data), then the node K is considered as a suspicious node

C. The Correction of D-S Theory

Proposition p: Suppose node K is trusted, then the sub propositions of p $\{P_1, P_2, \dots, P_m\}$ respectively represent data $A_{k1}, A_{k2}, \dots, A_{km}$ is credible, when evaluate belief function(Bel_k) of node K by suspicious data A_{ki} , its basic probability distribution function $m_{ki}(B)$ ($B \in 2^\Theta$) on the sub proposition space($\Theta = \{P_1, P_2, \dots, P_m\}$) not only need consider the deviation degree but also need consider the

combination of data collected by the node and correlation in each dimensions.

According to the method of determining the basic distribution function $m_{ki}(B)$, to calculate the belief function $Bel_k(P) = Bel_k(\{P_1, P_2, \dots, P_m\})$ (the lower bound of the credibility of the node K) and the likelihood function $Pl_k(P) = 1 - Bel_k(\bar{P})$ (the upper bound of the credibility of the node K) by A_{ki} .

According to the method of the determination of the basic distribution function, respectively calculate the value of the lower bound function $\{Bel_1(P), Bel_2(P), \dots, Bel_n(P)\}$ and the value of the lower bound function $\{Pl_1(P), Pl_2(P), \dots, Pl_n(P)\}$

Last, the lower bound of the overall credibility of the node

$T_k \downarrow = \frac{1}{n} \sum_{1 \leq i \leq m} \lambda_i \alpha_{ki} Bel_i(P)$, the upper bound value is $T_k \uparrow = \frac{1}{n} \sum_{1 \leq i \leq m} \mu_i \alpha_{ki} Pl_i(P)$. $T_k \downarrow, T_k \uparrow \in [0, 1]$, λ_i, μ_i are distribution factors.

D. Process of Algorithm MDRE

(1) By calculating the degree of deviation of the data from each node in the sensor network to determine the suspicious node (when the deviation degree is above the threshold, the node will be determined as the suspicious node).

(2) To calculate the upper and lower bound of the overall credibility of the node based the modified D-S theory

IV. SIMULATION AND ANALYSIS

A. simulation

In order to guarantee the authority of the simulation experiment data, this paper get data from the national climate data center. These meteorological sensor nodes collect the daily maximum temperature, daily minimum temperature and rainfall of three environmental data. Since these nodes are normal, this paper will set up 3 experimental scenarios

Experimental scene 1: select 10 nodes in the sensor network, and the number of 1-10 for 10 nodes. To evaluate the credibility of the nodes by the 3D data collected by these nodes, and to calculate the upper and lower bounds of the overall credibility of each node. The results are as shown in figure 2.

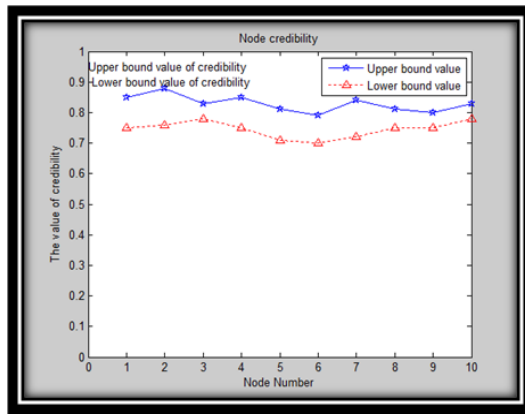


Figure 2. Nodes without interference

Experimental scene 2: select 10 nodes in the sensor network, and the number of 1-10 for 10 nodes. The interference of one dimensional data on each node is made artificially, and the interference of the nodes is gradually enhanced with the increase of the number of the nodes, and to calculate the upper and lower bounds of the overall credibility of each node. The results are as shown in figure 3.

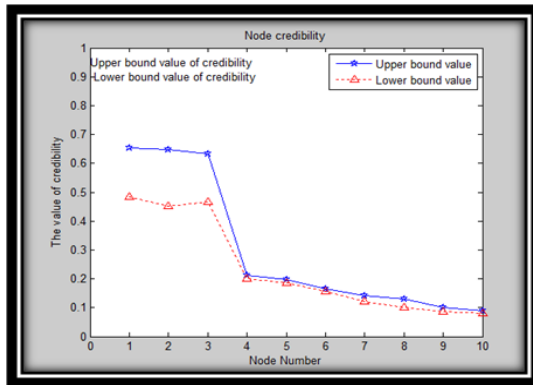


Figure 3. One dimension data is interfered gradually

Experimental scene 3: Select 10 nodes in the sensor network, and the number of 1-10 for 10 nodes. The interference of two dimensional data on each node is made artificially, and to calculate the upper and lower bounds of the overall credibility of each node. The results are as follows.

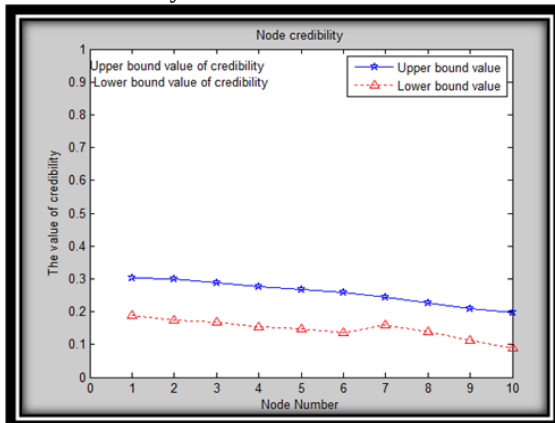


Figure 4. Two dimension data are interfered

B. experimental result analysis

According to the experimental results, When the sensor node are affected by the same number of dimensions, the upper and lower bounds of the overall credibility of the node are decreased with the increase of the data. On the other hand, with the increase of the number of the interference dimensions, the upper and lower bounds of the overall credibility of the node are also reduced. The results of the experiments show that the new algorithm based on the deviation degree is reasonable.

V. CONCLUDING REMARKS

This paper focuses on how to calculate the credibility of sensor network nodes, and proposes a new method which named MDRE, The main idea of this method is to calculate the mean value of each dimension which is collected from all the nodes in the network firstly, Then according to the difference between value and mean value in each dimension to determine whether the node is a suspicious node. Once again according to the relationship between the data in the dimension and the collected historical data to determine the basic probability distribution function, which can effectively reduce the influence of suspicious data on basic probability distribution function. At the same time, to apply the comprehensive result of the distribution factor and the degree of deviation as the weight of the credibility will lead to more scientific and accurate results. The experimental results show that this method can evaluate the reliability of the node effectively.

REFERENCES

- [1] Akyildiz I F, Su W, Sankarasubramaniam Y, et al. Wireless sensor networks: a survey[J]. Computer Networks, 2002,38(4): 393- 422
- [2] Sun Limin, Li Jianzhong, Chen Yu. Wireless sensor network [M]. Beijing: Tsinghua University press,2005:3- 5.
- [3] S.Liu, G. Xing, H. Zhang, J. Wang, J. Huang, M. Sha, and L. Huang. Passive interference measurement in wireless sensor networks. In Proceedings of IEEE ICNP, Kyoto, Japan, 2010
- [4] S. Ba and P. A. Pavlou. Evidence of the effect of trust building technology in electronic markets: price premiums and buyers behavior. MIS Quarterly, vol.26, 2002
- [5] P. Dasgupta, "Trust as a commodity," in n Gambetta, Diego (ed.) Trust: Making and Breaking Cooperative Relations, vol. electronic edition, D. Ingram, Ed.: Department of Sociology, University of Oxford, pp. 49-72, 2002.
- [6] S. Marsh. Formalizing Trust as a Computational Concept. In Department of Computer Science and Mathematics, vol. PhD: University of Stirling, 1994
- [7] P. Resnick and R. Zeckhauser. Trust among strangers in internet transactions: empirical analysis of eBays reputation system. Presented at NBER: workshop on empirical studies of electronic commerce, 2000.
- [8] L. Xiong and L. Liu A reputation-based trust model for peer-to-peer e-commerce communities. Presented at IEEE conference on E-commerce, 2003
- [9] L. Girod, J. Elson, A. Cerpa, T. Stathopoulos, N. Ramanathan, and D. Estrin. Emstar. A software environment for developing and deploying wireless sensor networks. In Proceedings of the USENIX Annual Technical Conference, Boston, MA, 2004

- [10] N. Ramanathan, K. Chang, R. Kapur, L. Girod, E.Kohler, and D. Estrin. Sympathy for the sensor network debugger. In Proceedings of ACM SenSys, San Diego, USA, 2005
- [11] Q. Cao, T. Abdelzaher, J. Stankovic, K. Whitehouse, and L. Luo. Declarative trace points: a programmable and application independent debugging System for wireless sensor networks. In proceedings of ACM SenSys, Raleigh, NC, USA, 2008
- [12] Kebin Liu, Qiang Ma, Xibin Zhao and Yunhao Liu. Self-Diagnosis for Large Scale Wireless Sensor Networks. In Proceedings of IEEE INFOCOM, Shanghai, China, 2011
- [13] Dempster A P. Upper and lower probabilities induced by multivalued mapping. *Annals of Mathematical Statistics*, vol.38(2): pp. 325-339, 1967
- [14] Dempster A P. A Generalization of Bayesian Inference. *Journal of the Royal Statistical Society*. 1968, Series B.30:205-247
- [15] Shafer and Glenn. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.