Available online at www.sciencedirect.com

ScienceDirect

AASRI Procedia 6 (2014) 59 – 65

AASRI Procedia

www.elsevier.com/locate/procedia

2013 2nd AASRI Conference on Computational Intelligence and Bioinformatics

Optimize on Data Correlation of Sensor Nodes with Adaptive Fault-tolerant Algorithm

Yun Liu, Hongzhi Zhou*

Department of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan, 650500, China

Abstract

A-SMAC protocol based on sensor node data traffic dynamic adjusting node listen and sleep time. Based on A-SMAC protocol, sensor nodes sent out the underlying error decisions .In this paper, we propose an adaptive fault-tolerant algorithm (AFTA) of node. According to the results of fault tolerant judgment, use of A-SMAC protocol dynamic adjusting node listen and sleep time, realize the combination of data judgment and node dynamic listening. Our experimental results show that the AFTA outperforms the Bayesian algorithm and exhibits strong fault-tolerant capabilities and less energy consume.

Keywords: Fault-tolerance, Adaptive, A-SMAC Protocol;

1. Introduction

Event detection is one of the important applications in Wireless sensor network (WSN), the main problem is accuracy affected by environmental noise and energy [1].Existing fault-tolerant for data collection usually

* Corresponding author. Tel.:+86-13888918579;
E-mail address: liuyun@kmust.edu.cn.
The project was supported by the National Nature Science Foundation of China(KKGD201203004)

doi:10.1016/j.aasri.2014.05.009

based on the spatial correlation algorithm, such as distributed Bayesian fault tolerant detection algorithm (BFTD) [2].

However, BFTD frequent communication between adjacent nodes, that leads the energy consume fast .A-SMAC protocol based on data traffic dynamic adjustment to listen and sleep time. A-SMAC protocol does not contribute to the reliability of the underlying data judgment. Usually send out unreliable information, result in wrong decision to Fusion Center.

In this paper, we present an adaptive fault tolerant algorithm (AFTA).Through the Threshold Area Decision Scheme [3], using the data correlation, realize the combination of underlying data fault-tolerant verdict and A-SMAC protocol.

We adopt the simulation shown the values of slow changes in the monitor environments, the AFTA could improve the efficiency of valid data delivery and reduce the unnecessary data to send.

The remainder of this paper is organized as follows. Section 2 presents Time Correlation Fault-tolerant algorithm. Adaptive listening dormancy discussed in Section 3. Section 4 described our simulation results. Section 5 concludes the paper.

2. Fault-tolerant Based on Time Correlation

2.1. System model

When a sensor readings is greater than certain threshold, assumes the sensor monitoring area by events; And when the sensor readings is smaller than certain threshold, it assumes no incident. Because of the sensor is not reliability, sensor readings may greater than a certain threshold method does not necessarily mean the monitoring area has occurred.

We define the node i in t time values is :

$$\chi_i(t) = m_i(t) + Z_i \qquad (2\text{-}1)$$

$m_i(t)$is environment characteristic value at t time.$Z_i(t)$ is Detection of noise(including the effect on the stability of the environment noise and equipment).

Assume that at any time of t,$Z_i(t)$is Gaussian distribution $N(0，\sigma^2)$。We define that t is the system test cycle,this the threshold. Each node periodically send binary decision, get test results:$S_i(KT_s)$

$$S_i(KT_s) = \begin{cases} 1 & \text{if } \chi_i(KT_s) \geq \text{th} \\ 0 & \text{if } \chi_i(KT_s) < th \end{cases} \qquad (2\text{-}2)$$

$S_i(KT_s) = 0$ is at KT_s time test results of the node i is “no event”.

So we know $S_i(KT_s) = 1$ actually as result is “event”. Assume that node i at KT_s time the actual circumstances of the environment in a binary variable $T_i(KT_s)$ to expression:

$$T_i(KT_s) = \begin{cases} 1, & \text{if } m_i(KT_s) \geq \text{th} \\ 0, & \text{if } m_i(KT_s) < th \end{cases} \qquad (2\text{-}3)$$

$T_i(KT_s)$=0 is at time of KT_s situation for the environment without incident.$T_i(KT_s)$=1 is real situation for the incident.

It has four possible situations: $T_i(KT_s) = 0$, $S_i(KT_s) = 0$ (Sensor correctly report without events);$T_i(KT_s) = 1$, $S_i(KT_s) = 1$ (Sensor correctly report events); $T_i(KT_s) = 0$, $S_i(KT_s) = 1$(Sensor error report events); $T_i(KT_s) = 1$, $S_i(KT_s) = 0$(sensor error report without events).

We observe that sensor node have not reliability, $T_i(KT_s) \neq S_i(KT_s)$ may occur, also is to detect errors. Missed detection probabilityP_1 and false alarm probability P_2 are:

$$P_1 = P(S_i(KT_s) = 0|T_i(KT_s) = 1) = P(\chi_i(KT_s) < th|\, m_i(KT_s) \geq th) = Q\left(\frac{m_i(KT_s)-th}{\sigma}\right) \quad (2\text{-}4)$$

$$P_2 = P(S_i(KT_s) = 1|T_i(KT_s) = 0) = P(\chi_i(KT_s) \geq th|\, m_i(KT_s) < th) = Q\left(\frac{th-m_i(KT_s)}{\sigma}\right) \quad (2\text{-}5)$$

Assumptions in ref.[4] the symmetry error probability:$P_1 = P_2$ But from (2-4) and (2-5) can be obtained,

Missed detection probability P_1 and false alarm probability P_2 are all based on environment value$m_i(KT_s)$ and Noise standard variance σ .

We define the time correlation of fault tolerance algorithm, is used to improve the energy efficiency and precision.

2.2. Time correlation data

Under the condition of monitoring the environment change slowly, environment characteristic value$m_i(t)$ in a short period of time considered change slowly.

So we called value at the time of the redundant as time correlation data [5]. We define that node i characteristic value keeps the same situation at $C = [(k-L)T_S, kT_S]$.

So in time t, (L + 1) consecutive redundant information can be used to test results of time error correction. Average values of node i within the time period C is:

$$\overline{X}_{i;C} = \frac{1}{L+1}\sum_{i=0}^{L} x_i\,[(k-l)T_s] = \frac{1}{L+1}\sum_{i=0}^{L} m_i[(k-l)T_s] + \frac{1}{L+1}\sum_{i=0}^{L} z_i\,[(k-l)T_s] = \overline{m}_i + \overline{Z}_i \quad (2\text{-}6)$$

$\overline{m}_i$is the average value of characteristic environment within the time period C.$\overline{Z}_i$ is aGaussian random variable, the value of variance is $\sigma^2/(L+1)$

m_n is the average value without event, such as proposed in [6]. So m_f is average value with event .the binary decision threshold is 0.5(m_n+m_f).

For the actual situation of node i environment to make the following binary estimates$R_i(KT_s)$:

$$R_i(KT_s) = \begin{cases} 0 & \text{if} \quad \overline{X}_{i\prime(k-L)T_S,kT_S} < th \\ 1 & \text{if} \quad \overline{X}_{i\prime(k-L)T_S,kT_S} \geq th \end{cases} \quad (2\text{-}7)$$

Detection error probability can be approximately expressed in the following Q function to:

$$\hat{p} = P(R_i(KT_s) = 0|T_i(KT_s) = 1) = P(R_i(KT_s) = 1|T_i(KT_s) = 0) = Q\left[\frac{(m_f-m_n)\sqrt{L+1}}{2\sigma}\right] \quad (2\text{-}8)$$

The system miss probability$P_M$and false alarm probability P_F are given by

$$P_M = P(R_i(KT_s) = 0|T_i(KT_s) = 1);\; P_F = P(R_i(KT_s) = 1|T_i(KT_s) = 0).$$

If $S_i(KT_s) \neq T_i(KT_s)$ and $R_i(KT_s) = T_i(KT_s)$ that means corrected a mistake. Type (2-8) is shown that probability of miss and probability of false alarm decreases with the increase of L.

2.3. Threshold Area Ruling Model

Using environment characteristic value if there has been a marked change, the node redundancy testing information of fault tolerance, will introduce more mistakes. So we define the node decision threshold area concept (such as those described in [7]).Through the threshold region to estimate changed in the environment.

Here we suppose the threshold area:$[m_f - \alpha\sigma,\ m_n + \alpha\sigma]$, The parameter σ is the standard variance of Environmental noise. a is the parameter.

According to the same node in time of the adjacent value differences, defines the dynamic change of environment. If the same node twice values satisfy the following condition, define that obvious changes have taken place in monitor region.

$$\{\chi_i((k-1)T_s) < m_f - \alpha\sigma \text{ and } \chi_i(kT_s) > m_n + \alpha\sigma\}$$
$$\{\chi_i((k-1)T_s) > m_n + \alpha\sigma \text{ and } \chi_i(kT_s) < m_f - \alpha\sigma\} \tag{2-9}$$

Otherwise, we assume the value of the testing region is not changing significantly.

If the node i characteristic values for environment is m_n then the probability density function of $\chi_i(kT_s)$ is :

$$f_0(\chi_i(kT_s)) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\chi_i(kT_s)-m_n)^2}{2\sigma^2}\right] \tag{2-10}$$

$\chi_i(kT_s)$in the interval $[m_n - \alpha\sigma, m_n + \alpha\sigma]$ Probability as follows:

$$P((m_n - \alpha\sigma) < \chi_i(kT_s) < (m_n + \alpha\sigma)) = \int_{(m_n-\alpha\sigma)}^{(m_n+\alpha\sigma)} f_0(\chi_i(kT_s))d(\chi_i(kT_s)) = 1 - 2Q(\alpha) \tag{2-11}$$

Similarly, if the environmental characteristic value of node i is m_f, Probability of $\chi_i(kT_s)$ as follows:

$$f_1(\chi_i(kT_s)) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\chi_i(kT_s)-m_f)^2}{2\sigma^2}\right] \tag{2-12}$$

$\chi_i(kT_s)$in the interval $[m_f - \alpha\sigma, m_f + \alpha\sigma]$ Probability as follows:

$$P((m_f - \alpha\sigma) < \chi_i(kT_s) < (m_f + \alpha\sigma)) = \int_{(m_f-\alpha\sigma)}^{(m_f+\alpha\sigma)} f_0(\chi_i(kT_s))d(\chi_i(kT_s)) = 1 - 2Q(\alpha) \tag{2-13}$$

If the parameter α is enough big, can approximate thought that $\chi_i(kT_s)$ distribute in the interval $[m_f - \alpha\sigma, m_f + \alpha\sigma]or[m_n - \alpha\sigma, m_n + \alpha\sigma]$. For example, when a = 196, $1 - 2Q(\alpha)$=95%.

When the node values in the interval$[m_f - \alpha\sigma, m_n + \alpha\sigma]$, test results is not clear. Thus, interval as a mixing zone, the aliasing region is defined as the threshold area.

Parameterα shouldsatisfy the inequality:$m_f - \alpha\sigma < th \text{ and } m_n + \alpha\sigma > th$[8]

$$\text{The constraint conditions for: } \alpha \geq \max\{(m_f - th)/\sigma, (th - m_n)/\sigma\} \tag{2-14}$$

Expanding the single threshold for the threshold area, (2-9) formula to judge whether tested environment changed. In order to estimate the duration time of the node and adjust the average values (2-6), and finally make an estimate of the actual situation of the environment. The algorithm implementation step in the following ways:

a) Using the detection threshold area, to determine whether a test environment change obviously.

b) In case of obvious change, the binary decision, binary estimatesR_i and will be the verdict sent three times in a row.

c) If it had no obvious change, determine the time correlation processing period T, to calculate the L average values at period T. reach the χ_i.

d) For binary decision detection average valuesχ_i. For binary estimate T, and then send the verdict.

If this test environment does not change, using the data of time correlation, fault-tolerant processing of test data. If suddenly change, a binary decision sent three times in a row, is to improve the node number of sent packets.

In MAC layer, A-SMAC protocol according to the node throughput dynamic adjustment to listen and sleep time. So it can transmit the information in a timely manner.

3. Adaptive Listening Dormancy Scheme

3.1. A-SMAC protocol principle

A-SMAC protocol is based on the SMAC protocol improved, such as propose in [8,9]. A-SMAC protocol has adaptive listening scheme could according to the value of the node throughput to predict the change of network traffic.

3.2. A-SMAC protocol working process

A-SMAC protocol could base on the node throughput to measure the size of network traffic. Set the parameters used for statistical nodes in each cycle to send and receive packets number. We adopt the two threshold method (respectively set as S, M) can be divided n into three kinds of circumstances.

And in these three cases,D_1, D_2, D_3. The three kinds of duty ratio(D_1<D_2<D_3).Corresponding to the setting as follows:

a) If $n < S$, at this stage the network load is small, adjust the duty ratio, the minimum duty cycleD_1;

b) If $S =< n < M$, at this stage the network load is not big, adjust the duty ratio, the duty cycleD_2;

c) If $n \geq M$, at present stage network load is bigger, use duty cycle D_3.

Duty ratio adjusted, the length of the frame period is not change, just listen time to adjustment, new activity time should be as follows:

$$T_{newlisten} = D_i \times T_{frame}(i=1,2,3,)(3\text{-}1)$$

The D_i is adjusted the value of duty ratio,

Increase the time of data packets, can transmit more data in the effective time. AFTA propose if the environment changed, send three consecutive decisions, so that we can make the node to transmit data traffic increase. Then, according to A-SMAC protocol make duty ratio increases, listen time increases, can make decision information sent out in time, reduce the decision delay data.

4. Results of Simulation Analysis

In order to compare the AFTA and Bayesian algorithm .We use theNS-2.35 emulator as experimental platform [10].

We define the 90 nodes were randomly distributed in the area, between two nodes in turn form pairs of CBR stream transmission link. Adaptive incident detection algorithm adopts the method of sliding window-restart. Determine the length L of the sliding window, binary decision threshold th, and a threshold area parameters.

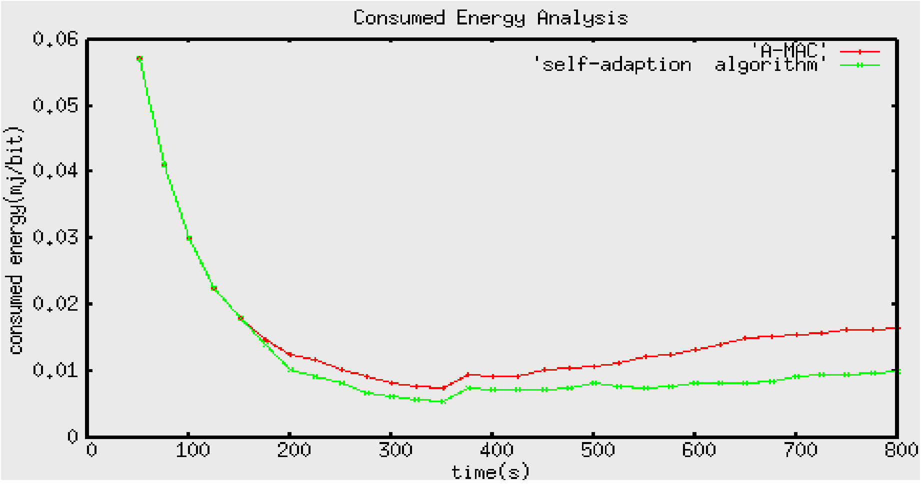

Fig. 4.1. Energy consumption rate

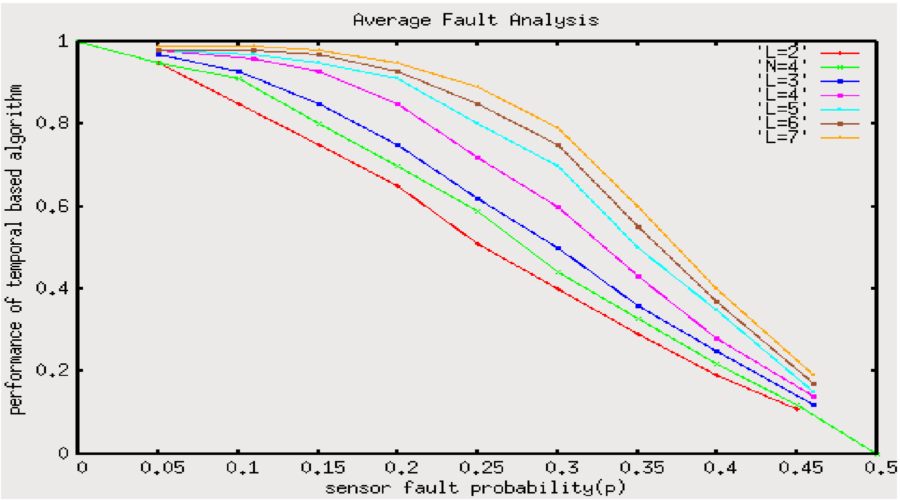

Fig. 4.2. Error performance comparison

Sliding window length L should according to the test to determine changes in the environment condition, general values between 3~5 more appropriate. Binary decision threshold $th = 0.5(m_n + m_f)$ The simulation results as shown in the figure below.

We can see from the graph Fig. 4.1, the algorithm after joined the fault-tolerant judgment relative to A-SMAC protocol, the node energy consumption is lower.

Fig.4.2 shows error reduction performance of AFTA. We can seen from the diagram when data acquisition number L take more than three times, the AFTA is superior to using spatial correlation of (adjacent nodes N = 4)Bayesian fault tolerance detection algorithm at error reduction.

5. Conclusions

Energy consumption and event monitoring has been one of the important indices for wireless sensor network performance .In this paper, AFTA is proposed. This approach require data time correlation to fault tolerance through Threshold area. The proposed approach also determines the combination of algorithm and A - SMAC protocol.

The simulation results demonstrate that this approach can significantly reduce the unnecessary data to send and lead energy consumption is lower than A-SMAC protocol. It is also shown that the AFTA improve the ruling data transmission efficiency effectively, makes the system has a higher real-time performance and reliability is better. As a future extension, AFTA approach can be developed that can be used for even larger networks.

References

[1] Ying Zhang, Shashidhar Gandham, and Qingfeng H. Distributed Minimal time converge cast scheduling for small or sparse data sources. 28th IEEE International Real-Time Systems Symposium, 2007.
[2] Scott C.-H. Huang, Peng-Jun Wan, and Chinh T. Vu. Nearly constant approximation for data aggregation scheduling in wireless sensor networks. IEEE INFOCOM, 2007.
[3] Bo Yu, Jianzhong Li, and Yingshu Li. Distributed data aggregation scheduling in wireless sensor networks. In IEEE INFOCOM, 2009.
[4] Yanwei Wu, Xiang-Yang Li, and YunHao Liu. Energy-efficient wake-up scheduling for data collection and aggregation. IEEE Transactions on Parallel and Distributed Systems, 2010.
[5] Siyuan Chen, Shaojie Tang, and MinsuHuang. Capacity of data collection in arbitrary wireless sensor networks. In IEEE INFOCOM, 2010.
[6] Y.T. Hou, Y. Shi, J. Pan, S.F. Midkiff, Lifetime-optimal data routing in wireless Sensor networks without flows plitting ,in: Workshop on Broadband Advanced Sensor Networks, San Jose, CA, 2004.
[7] Y.T. Hou, Y. Shi, H. Sherali, S.F. Midkiff, On energy provisioning and relay node placement for wireless sensor networks, IEEE Transactions on Wireless Communications 4 (5) (2005) 2579–2590.
[8] [46] J. Suomela, Computational Complexity of Relay Placement in Sensor Networks, SOFSEM 2006.
[9] K. Kalpakis, K. Dasgupta, P. Namjoshi, Maximum lifetime data gathering and aggregation in wireless sensor networks, in: Proceedings of IEEE International Conference on Networking, 2002.
[10] H. Liu, P. Wan, W. Jia, Fault-tolerant relay node placement in wireless sensor networks, in: COCOON, 2005.