The Fault Probability Estimation and Decision Reliability Improvement in WSNs

Ming-Tsung Hsu*, Frank Yeong-Sung Lin*, Yue-Shan Chang**, Tong-Ying Juang**
* Dept. of IM National Taiwan University, ** Dept. of CSIE National Taipei University
*No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan, **151, University Rd., San Shia, Taipei, 237 Taiwan

***Abstract*-Faults are an essential fact in wireless sensor networking as coupled with a set of constraints. The decision making based on the reports of fragile and fallible sensor nodes might be very unreliable and, therefore, might fail to accomplish the tasks of WSNs. Previously, the reduction in the effects of faults is based on collaborative effort of a large number of sensor nodes. The collaborative work may consume valuable power and may fail as sensor nodes are severely affected by environmental interference. In this paper, instead of the neighborhood communication and pure data fusion, the collaborative effort in decision making is accomplished based on the fault probability estimation, reading variance estimation, and critical value adjustment. The fault probability of sensor nodes is computed by their reports and the *t*-out-of-*n* rule to make reliable decisions. The reading variance is estimated by well-known sample variance to assess the effects of environmental interference. The critical value adjustment is triggered as estimated reading variance changed, high fault probability, or decision quality unsatisfied to reduce the bias of fault probability.**

I. INTRODUCTION

Recently, the wireless sensor networks (WSNs) composed of tiny sensor nodes are broadly deployed for environmental monitoring, automatic controlling, or target tracking [1]. These low-cost battery-powered sensor nodes equipped with radio transceivers are usually left unattended making sensor nodes fault or crash easily, especially when sensor nodes are deployed in harsh environments. Various types of faults which can significantly affect the network operation are known and studied in WSNs [1], [4], [7]. In this paper, the detection fault of sensing systems, such as the missing detection, false alarm, and unusual reading, arisen primarily from environmental interference, is discussed.

In most WSNs, sensor nodes only send detection reports (decisions) to sinks for energy conservation [1], [5], [7]. The detection decision of sensor nodes is made when the sensed energy is higher than the prescribed critical value which is set by the event energy and reading variance. The status of sensor nodes is difficult to evaluate as the sink only receives detection decisions. The decision making based on the fallible sensor nodes might be very unreliable.

Many statistical approaches are previously proposed for improving the quality of distributed detection with multiple sensors [4], [5], [7], [11]. This collaborative work still needs to contemplate the severe energy constraints and environmental interference of WSNs. Unlike previously proposed schemes with less consideration for the effects of environmental interference and the faults produced by them, in this paper, we focus on the fault probability of sensor nodes and the bias of fault probability as sensor nodes are severely affected by the environmental interference.

The remainder of this paper is organized as follows: The related work of the reliability improvement in WSNs is briefly reviewed in Section II. Section III introduces the fault probability and decision quality. The algorithms of reading variance estimation and adaptive critical value adjustment are proposed in Section IV to improve the decision reliability. Section V shows the simulation results of proposed reading variance estimation. Section VI draws our conclusions and future work.

II. RELATED WORK

In unforeseen circumstances of WSN applications, the sensor nodes are redundantly or densely deployed for increasing the reliability. Approaches focused on reliability or fault tolerance are hence designed for how to efficiently use the redundant sensor nodes. For example, several adaptive schemes, which reduce the impact of the failure of sensor nodes by adjusting protocols automatically, are designed based on the usage of redundant sensor nodes [2], [3], [13].

The reliability can also be satisfied by the collaborative effort of sensor nodes. Sun, Chen, Han, and Gerla [10] proposed a simple distributed technique, named CWV, by using neighbor's result and exploiting redundant information to discern local data correctness for improving reliability. Krishnamachari and Iyengar [5] proposed a scheme which an individual sensor node communicates with its neighbors and uses their binary decisions to correct its own decision to detect the event region for increasing fault tolerant capability. Luo, Dong, and Huang [7] enhanced this work by considering measurement error, sensor node fault, and the proper neighborhood size.

The collaborative effort of sensor nodes may cause the consistency problem. Clouqueur, Saluja, and Ramanathan [4] proposed two fusion schemes, value fusion and decision fusion, to solve the Byzantine problem [6] and to accomplish better reliability in the data fusion.

The proactive fault detection schemes focus on whether a sensor node or an entire region is crashed for efficient routing. Staddon, Balfanz, and Durfee [9] proposed a tracing

1-4244-1272-2/07/$25.00 ©2007 IEEE

scheme in continuous WSNs. Ruiz et al. [8] used MANNA to identify the faulty sensor nodes and proposed a management scheme for event-driven sensor networks.

III. Fault Probability and Decision Quality

A. Fault Probability

As mentioned earlier, in most WSNs sensor nodes only send detection decisions (sensor decisions) to sinks for energy conservation. There may have the following possible scenarios of sensor nodes: (a) detect events; (b) fail to detect events; (c) issue fault alarms [4]. The scenarios (b) and (c) are the missing detection and false alarm of sensor nodes, respectively. The fault probability of sensor nodes is then defined in Definition 1 [11], [14].

Definition 1 The fault probability of sensor node k is

$$\begin{aligned} P_{f,k} &= p(A_k \cap \overline{D}_k) + p(\overline{A}_k \cap D_k) \\ &= p(A_k)p(\overline{D}_k \mid A_k) + p(\overline{A}_k)p(D_k \mid \overline{A}_k) \end{aligned} \quad (1)$$

where $P_{f,k}$, D_k, A_k, $\overline{D}_k \mid A_k$, and $D_k \mid \overline{A}_k$ are the fault probability, detection decision, nearby event, missing detection, and false alarm of sensor node k, respectively.

In n-covered WSNs, the t-out-of-n rule of the sensor decisions is proposed to solve the detection problems [14]. The event occurrence can be justified by the t-out-of-n rule of the sensor decisions, e.g., a sensor node correctly detects an event as (t – 1) neighbors also report the same event. The probabilities of near by event, missing detection, and false alarm of sensor node k then can be estimated as follows:

$$\hat{p}(A_k) = \frac{nearby_k}{i} \quad (2)$$

$$\hat{p}(\overline{D}_k \mid A_k) = \frac{miss_k}{i \times p(A_k)} \quad (3)$$

$$\hat{p}(D_k \mid \overline{A}_k) = \frac{false\_alarm_k}{i \times p(\overline{A}_k)} \quad (4)$$

where i is the time slots, and $nearby_k$, $miss_k$ and $false\_alarm_k$ are the number of the nearby event, missing detection and false alarm of sensor node k, respectively. The fault probability of sensor nodes in n-covered WSNs then can be estimated as follows:

$$\hat{P}_{f,k} = \frac{miss_k + false\_alarm_k}{i} \quad (5)$$

For example, in a 5-covered WSN (t = 3), there are 4 events happened in 10 time slots, sensor node k reports 4 detection decisions where 2 events are detected by 2 or more sensor nodes and 2 events are not. The nearby event and fault probabilities of sensor node k are 0.4 and 0.4, respectively.

Instead of using the t-out-of-n rule of the sensor decisions, in target tracking WSNs the number of missing detection and false alarm are observed by the moving path of objects, which is not included in this paper.

B. Decision Quality

The detection problems (decision quality) can be categorized into Bayesian detection and Neyman–Pearson detection problems. The Bayesian detection problem focuses on how to minimize the Bayesian costs and the Neyman–Pearson detection problem focuses on how to maximize the detection probability while the false alarm probability is kept below a prescribed level [11], [14].

In this paper, the error probability of decisions is assumed as the only Bayesian cost [14]. The decision quality of Bayesian detection for each event then can be evaluated by the jointly fault probability of sensor nodes as follows:

Definition 2 The decision quality of Bayesian detection is

$$Q = 1 - p(A)\prod_{m=r}^{n} p(\overline{D}_m \mid A_m) - p(\overline{A})\prod_{f=1}^{r} p(D_f \mid \overline{A}_f) \quad (6)$$

where Q, A, $\overline{D}_m \mid A_m$, and $D_f \mid \overline{A}_f$ are the decision quality, event, missing detection of sensor nodes which do not report the event, and the false alarm of sensor nodes which report the event. The event probability can be estimated by the average of nearby event probabilities of sensor nodes.

IV. Critical Value Adjustment

A sensor node sends a detection decision to the sink when the sensed energy is higher than the prescribed critical value which may be set by the event energy (en_A) and default reading variance (σ_r^2). The probabilities of missing detection and false alarm can be misestimated when the environmental interference affects sensor nodes while the critical value settings are fixed. In this section, based on reading variance estimation, an adaptive critical value adjustment mechanism is proposed to solve this problem.

A. Reading Variance Estimation

The reading variance can be used to estimate the effects of environmental interference as sensor nodes within the same region are sustained the same environmental interference. Equation (7) shows the most used unbiased sample variance where S_r^2 and $\overline{sr}$ are the sample variance and sample mean of readings and sr_k is the reading of sensor node k.

$$S_r^2 = \frac{\sum_{k=1}^{N}(sr_k - \overline{sr})^2}{N-1} \quad (7)$$

The computation of the unbiased estimator is easy, but can be affected by the extreme value easily. Avoiding the extreme value effect, the readings should be examined and the extreme readings should be dropped. There is high probability that extreme readings exist when the test of variance between the prior and latest is rejected. The extreme readings are dropped once a round after compared with the sample median (or mode). The estimated reading variance is recomputed by the prior and latest sample variances after the extreme readings are dropped. The detail steps of reading variance estimation are shown as follows:

Reading Variance Estimation

1. If $S_r^2 = 0 \rightarrow S_r^2 = \sigma_r^2$
2. Compute S_{nr}^2, $\overline{sr}$, and m_r *(median)*
3. If $\frac{(n-1)S_{nr}^2}{S_r^2} > \chi^2(n-1,\alpha)$
   (1) If $(max\{sr_k\} - m_r) > (m_r - min\{sr_k\})$ drop $max\{sr_k\}$
   ELSE drop $min\{sr_k\}$
   (2) Recompute $\overline{sr}, S_{nr}^2$, and m_r
   (3) $S_r^2 = w_1 S_r^2 + (1 - w_1) S_{nr}^2$
   (4) Goto step 3

A reading query for reading variance estimation can be executed periodically, when a certain number of sensor nodes have high fault probability, or when the requirement of decision quality cannot be satisfied.

B. Adaptive Critical Value Adjustment

As mentioned earlier, the critical value of decision making is determined by the event energy and reading variance, e.g., $en_A - 3\sigma_r$, where the reading variance is affected by environmental interference. The missing detection (false alarm) probability might be overestimated if the reading variance increased (decreased). To reduce the effect of environmental interference, the critical value must be adjusted by different reading variances especially in harsh environments.

The critical value adjustment is triggered when the reading variance periodically computed is changed, a certain number of sensor nodes have high missing detection or false alarm probabilities, or the requirement of decision quality cannot be satisfied. The critical value adjustment algorithm is described below.

Adaptive Critical Value Adjustment

1. Compute $\hat{p}(\overline{D}_k \mid A_k)$ and $\hat{p}(D_k \mid \overline{A}_k)$
2. If S_r^2 changed $\rightarrow$ Adjust critical value by S_r^2
3. If x sensor nodes with $\hat{p}(\overline{D}_k \mid A_k) > p_m$
   (1) Query sensor reading and compute S_r^2
   (2) Adjust critical value by S_r^2
4. If y sensor nodes with $\hat{p}(D_k \mid \overline{A}_k) > p_f$
   (1) Query sensor reading and compute S_r^2
   (2) Adjust critical value by S_r^2
5. If $Q < Q_r$
   (1) Query sensor reading and compute S_r^2
   (2) Adjust critical value by S_r^2

V. SIMULATION

The concept of proposed reading variance estimation is validated via simulations by using ns-2 [15]. Compared with the value fusion mechanism proposed by Clouqueur, Saluja, and Ramanathan [4], we show that the reading variance can be well-estimated.

The simulation settings of this comparison are: 100 sensor nodes equipped single sensor deployed in $120 \times 120 m^2$ area, readings are normally distributed from 40 to 50 with variance 4 to 25; the faulty rate is 0.02; and the significance level (α) of the variance test is 0.05; the weight of variance adjustment w_1 is 0.5. The error of transmission is ignored.

Fig. 1 shows that the proposed algorithm can estimate the reading variance better than the value fusion mechanism (with 52%, 17%, and 44% improvement for drop 2, 4, and 6, respectively), where "drop n" means dropping the n/2th smallest and n/2th largest readings. In Fig. 2, the impact of different significance levels (α) is shown to be minor. Fig. 3 shows that the faulty rate will affect proposed algorithm.

The w_1 and Faulty rate are compared in Fig. 4a and 4b. These figures show that the w_1 should be adjusted by the faulty rate, e.g., when the faulty rate is high, the default reading variance should bear more weight to resist the fault.

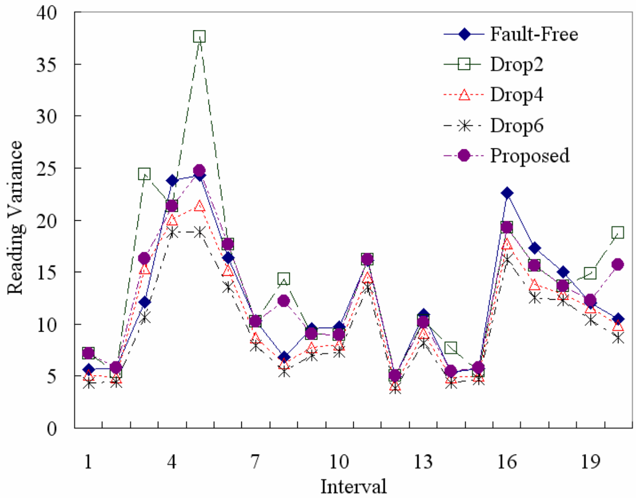

Fig. 1. Reading variance comparison: value fusion vs. proposed algorithm.

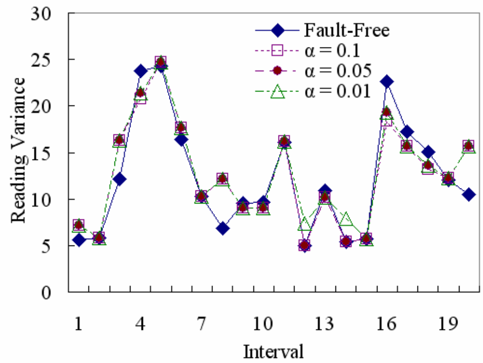

Fig. 2. Reading variance comparison: different significance levels.

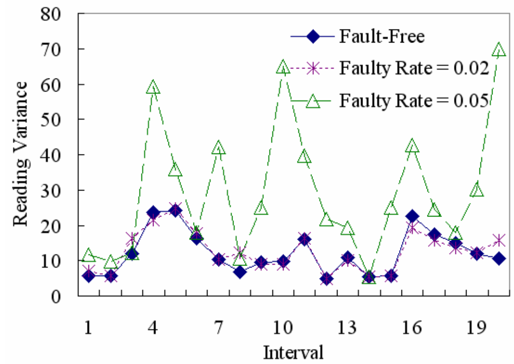

Fig. 3. Reading variance comparison: different faulty rates.

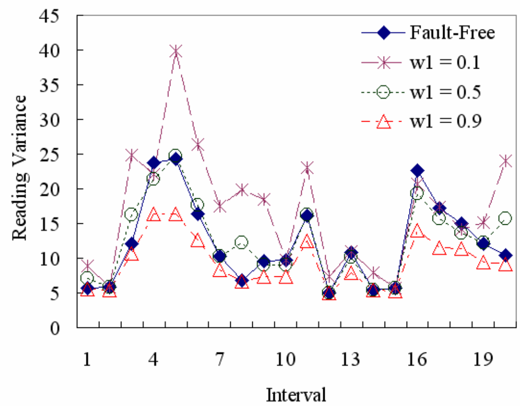

Fig. 4a. Reading variance comparison: different w_Is (faulty rate 0.02).

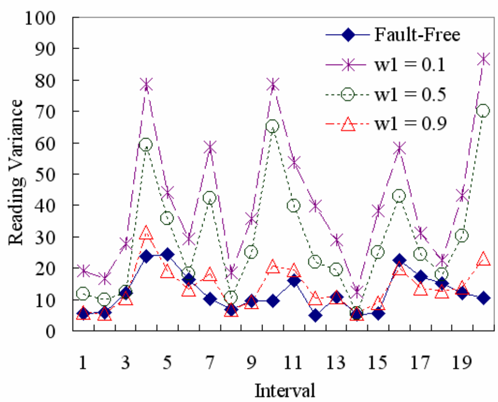

Fig. 4b. Reading variance comparison: different w_Is (faulty rate 0.05).

VI. Conclusions and Future Work

The fault probability estimation, reading variance, and critical value adjustment are proposed to overcome the detection problems of WSNs. The fault probability of sensor nodes is computed by sensor decisions. The environmental interference is estimated by the reading variance. The critical value adjustment is proposed to reduce the bias of fault probability.

In [14], Zhang, Varshney, and Wesel used the prior probabilities of hypotheses and Gaussian noise assumption to find the optimum t for minimizing the probability of error. We enhance this work as the prior probabilities can be estimated by observations and the reading variance estimation is proposed instead of Gaussian noise assumption.

The future work of this research includes: The relationship of w_I and the faulty rate, which can be estimated by the fault probability sensor nodes, should be discussed more detail; the environmental interference and the critical value adjustment policy should also be studied deeply by different real applications.

Reference

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: a survey,” *Computer Networks* 38 (2002) 393-422.

[2] A. Cerpa and D. Estrin, “ASCENT: Adaptive Self-Configuring sEnsor Networks Topologies,” *IEEE Trans. on Mobile Computing,* Special Issue on Mission-Oriented Sensor Networks, Volume 3, Number 3, July-September 2004

[3] Y.-S. Chang, M.-T. Hsu, H.-H. Liu, and T.-Y. Juang, “Dependable Geographical Routing on Wireless Sensor Networks,” *Lecture Notes in Computer Science (LNCS)* of Springer-Verlag, Vol. 4523, May 2007.

[4] T. Clouqueur, K. K. Saluja, and P. Ramanathan, “Fault Tolerance in Collaborative Sensor Networks for Target Detection,” *IEEE Trans. on Computers*, 53(3), pp. 320-333, March 2004.

[5] B. Krishnamachari and S. Iyengar, “Distributed Bayesian Algorithms for Fault-Tolerant Event Region Detection in Wireless Sensor Networks,” *IEEE Trans. on Computers*, 53(3), pp. 241-250, March 2004.

[6] L. Lamport, R. Shostak, and M. Pease, ”The Byzantine Generals Problem,” *ACM Trans. on Programming Languages and Systems*, 4(3), pp. 382-401, July 1982.

[7] X. Luo, M. Dong, and Y. Huang, “On Distributed Fault-Tolerant Detection in Wireless Sensor Networks,” *IEEE Trans. on Computers*, 55(1), Jan. 2006.

[8] L. Ruiz, I. Siqueira, L. e Oliveira, H. Wong, J. Nogueira, and A. Loureiro, “Fault Management in Event-Driven Wireless Sensor Networks,” *MSWiM'04*.

[9] J. Staddon, D. Balfanz, and G. Durfee, “Efficient Tracing of Failed Nodes in Sensor Networks,” *First ACM International Workshop on Wireless Sensor Networks and Applications*, Sep 2002.

[10] T. Sun, L.-J. Chen, C.-C. Han, and M. Gerla, “Reliable Sensor Networks for Planet Exploration,”*The 2005 IEEE International Conference On Networking, Sensing and Control (ICNSC'05).*

[11] P. Varshney, “Distributed Detection and Data Fusion,” *Springer-Verlag*, 1996.

[13] F. Ye, G. Zhong, J. Cheng, S. Lu, and L. Zhang, “PEAS: A Robust Energy Conserving Protocol for Long-lived Sensor Networks,” *Proceedings of the 23rd International Conference on Distributed Computing Systems (ICDCS'03).*

[14] Q. Zhang, P.K. Varshney, and R.D. Wesel, “Optimal Bi-Level Quantization of i.i.d. Sensor Observations for Binary Hypothesis Testing,” *IEEE Trans. Information Theory*, vol. 48, no. 7, 2002.

[15] The Network Simulator- ns-2, http://www.isi.edu/nsnam/ns/, 2007