Channel-Aware Bayesian Model for Reliable Environmental Monitoring Sensor Networks

Zhenghai Wang, Mao Tian
School of Electronic Information, Wuhan University, China
Yuhao Wang
School of Information Engineering, Nanchang University, China
Email: {zhh.wang.whu}@gmail.com

***Abstract*—We provide a novel statistical-based approach to address the reliable data reception problem for environment monitoring sensor networks. In this paper, a channel-aware Bayesian model is designed to choose the most likely quantization level as the value of the monitored phenomenon at the sink node such that the probability of a correct decision is maximized. Simple methods are presented to formulate both the temporal correlation of the monitored phenomenon and the error information of the data transfer channel between the sensor node which monitors the phenomenon and the sink node as the prior input to this Bayesian device. Evaluation based on real sensor data shows that the proposed model can monitor physical phenomena much more accurately than channel-unaware algorithms.**

***Index Terms*—Bayesian model, reliability, channel-awareness, channel model, environment monitoring, wireless sensor networks**

I. INTRODUCTION

Modern wireless communication and semiconductor technologies have enabled the massive deployments of ultrasmall, cost-efficient and low power sensor nodes, which have led to the blossom of wireless sensor networks (WSNs) in a wide range of applications, such as environment monitoring, battlefield surveillance, etc. A common goal in most WSN applications is to reconstruct the underlying physical phenomenon, e.g., temperature, based on sensor observations [1],[2]. However, sensor observations are exposed to various errors during sampling, quantization, storage, and reporting. First, due to residual calibration errors, sensors may digitalize samples with time-invariant bias known as systematic errors [3]. During storage, the data are again subjected to soft errors caused by on-chip thermal noise and cosmic radiations [4],[5]. Afterward, the data are still suffered from the wireless channel errors during data reporting to the sink node. To provide analytical knowledge about these transient errors on the data transfer channel (DTC), a binary systematic channel (BSC) has been devised to model the noisy logic gates of semiconductors [5]; soft errors on nanometric SRAMs induced by cosmic rays have been tested and analytical models that well fit the experiments have also been given in [5].

To tackle on-chip transient errors and wireless channel errors, various error control codes have been proposed and all of them have turned to introducing redundancies to handle these kinds of errors. For example, error correction codes (ECC) and triple modular redundancy (TMR) have been widely used in semiconductors to control soft errors [6]. Similarly, forward error correction codes (FEC), automatic retransmission requests (ARQ), and the combination of FEC and ARQ (i.e., HARQ) have been studied to correct wireless channel errors [7]. To keep minimal communication overhead, many statistical-based approaches have been used to explore the spatial and temporal correlations of physical phenomena to deal with the problem of reliable data reception [8],[9],[10],[11]. They assume or estimate a statistical model which captures the distribution of a physical phenomenon and evaluate observations with respect to how well they fit the model. E.g., Eiman Elnahrawy et al. [8] use the Bayesian Belief Network (BBN) to capture the spatial-temporal correlations that exist among observations of sensor nodes.

All the aforementioned algorithms address the reliable data reception problem by either introducing redundancies or fitting statistical models of the monitored phenomena, but ignoring the DTC error information. Considering the scarce power supplies and the promise of considerable performance improvements as has been proved in decision fusion [12],[13], target tracking [14] and multiple accesses [15] scenarios, we provide a channel-aware statistical-based approach to address the reliable data reception problem and keep minimal communication overhead. In this paper, we focus on the reliable data reception from a sensor node that quantizes the monitored phenomenon into K levels, which is the basic problem for WSNs. We wish to design a decision device which chooses a quantization level for the current phenomenon such that the probability of a correct decision is maximized. Both the temporal correlation of a physical phenomenon and the error information of the DTC are modeled as prior input to the decision device.

The rest of this paper is organized as follows. Section II introduces the system model of data collection at the sink node from the sensor node that quantizes the monitored

Manuscript received February 15, 2011; revised May 15, 2011; accepted June 15, 2011.
This work was supported by a grant from the National Natural Science Foundation of China (No.60762005) and International Science & Technology Cooperation Project (No.2010DFA70990).

doi:10.4304/jcm.6.5.355-359

physical phenomenon into K levels. Section III models the error-prone DTC by a discrete K-ary input and K-ary output channel. Then, discrete time series models are used to estimate the values of the phenomenon and the estimation errors are modeled as a normal distribution in Section IV. Section V gives some results that demonstrate the efficiency of the proposed algorithms. Finally, conclusions and future extensions are given in Section VI.

II. System Model and Problem Formulation

The problem to be solved including the network and channel-aware Bayesian model is first described formally.

A. Network Model

Fig.1 depicts the model for an end-to-end data transmission in one-hop sensor networks. The sensor node is monitoring a phenomenon $\boldsymbol{\vartheta}$. z^t, where t is the time index, denotes the received data at the sink node. z^t is the error-corrupted version of the digitalized sample θ^t at the sensor node. We make the following definition:

Definition 1: The phenomenon $\boldsymbol{\vartheta}$ is a continuous stochastic process bounded at interval $[l,u]$, and ϑ^t is the realization of $\boldsymbol{\vartheta}$ at time t. The sensor digitalizes the interval $[l,u]$ uniformly into K quantization levels $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_K\}$, where $K = 2^q (q = 1, 2, \ldots)$. Each quantization level represents a subinterval with size $\Delta = (u-l)/K$. Using θ_i^t denotes that $\theta^t = \theta_i (i = 1, 2, \ldots K)$, where θ^t is the digitalized version of ϑ^t at time t.

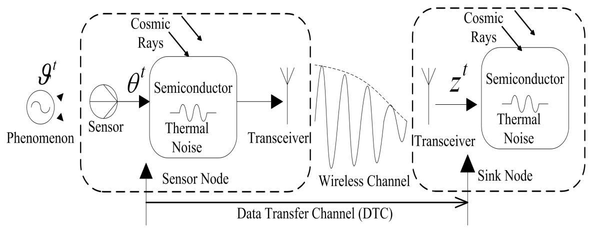


Fig. 1. Model for data transmission in one-hop networks

B. Channel-Aware Bayesian Model

As shown in Fig.1 at time t, the phenomenon $\boldsymbol{\vartheta}$ is first sampled and quantized by the sensor node. After sampling and quantization, the observation is reported to the sink node. After receiving the data at the sink node, we wish to design a decision device that makes a decision on the quantization level based on the received data and the error probability of the DTC such that the probability of a correct decision is maximized. With this goal in mind, using Bayes' rule, the posterior probability is expressed as follows:

$$P(\theta_i^t/z^t) = \frac{P(z^t/\theta_i^t)P(\theta_i^t)}{P(z^t)} \quad (1)$$

where θ_i^t and z^t are the input and output of the DTC at time t, respectively. $P(z^t/\theta_i^t)$ is the conditional probability of the DTC output z^t, given the DTC input θ_i^t. $P(\theta_i^t)$ is the probability of the DTC input θ_i^t. The denominator of Eq.1 can be expressed as follows:

$$P(z^t) = \sum_{k=1}^{K} P(z^t/\theta_i^t)P(\theta_i^t)$$

Since the denominator is always constant, the decision criterion, which selects the quantization level with maximum posterior probability, can be expressed as follows:

$$\begin{aligned}\hat{\theta}^t &= argmax_{\theta_i^t \in \boldsymbol{\theta}} P(\theta_i^t/z^t) \\ &= argmax_{\theta_i^t \in \boldsymbol{\theta}} \Big(P(z^t/\theta_i^t)P(\theta_i^t) \Big) \end{aligned} \quad (2)$$

From Eq.1 and Eq.2, we observe that the computation of the posterior probability $P(\theta_i^t/z^t)$ requires the knowledge of the likelihood function $P(z^t/\theta_i^t)$ and prior probability $P(\theta_i^t)$, which specify the statistical properties of the error-prone DTC and the monitored physical phenomenon at time t, respectively.

III. DTC Model for WSNs

In Section II, we have shown that the posterior probability computation needs the statistical properties of the error-prone DTC. The crossover probabilities of the DTC are determined by errors induced by the cosmic rays, on-chip thermal noise, and wireless channel. Given supply voltage, the above three kinds of errors are independent with each other. The bit error probability of the DTC is as follows:

$$P_b = 1 - (1 - P_{ser})(1 - P_{noise})(1 - P_{hop}) \quad (3)$$

where P_{ser}, P_{noise} and P_{hop} denote the bit error probability induced by the cosmic rays, on-chip thermal noise, and wireless channel, respectively. When the sensor node digitalizes the phenomenon into K quantization levels $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_K\}$ as *Definition 1*, we can use the discrete K-ary input and K-ary output channel shown in Fig.2 to model the DTC shown in Fig.1, and the channel is characterized by the conditional probability as follows:

$$P(\theta_j/\theta_i) = (P_b)^{dis(\theta_j,\theta_i)}(1 - P_b)^{q-dis(\theta_j,\theta_i)} \quad (4)$$

where $dis(\theta_j, \theta_i)$ is the Hamming distance between the binary representations of θ_j and θ_i. In Eq.4, we assume the bit errors on the DTC are random. $P(\theta_j/\theta_i)$ denotes the probability of the crossover from the DTC input θ_i to the DTC output θ_j.

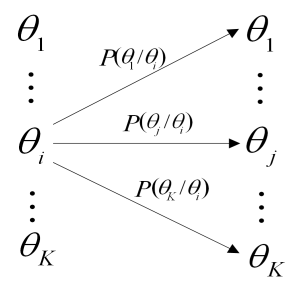


Fig. 2. DTC Model for WSNs

IV. Data Source Model for Physical Phenomena

From Eq.2 in Section II, we have observed that the posterior probability computation requires the knowledge of the distribution of the monitored physical phenomenon as well. In environmental monitoring WSNs, the sensor node needs to continuously sense the phenomenon and report observations to the sink node to reconstruct the phenomenon. In such a field, data are correlated to previous measurements at the same place. Discrete time series models, e.g., autoregressive (AR) [16] and robust Holt-Winters smoothing (RHW) [17], are widely used to predict the future behavior of variables. Herein, we use discrete time series models to estimate of the phenomenon at time t, and make the following definition:

Definition 2: At time t, the value of the monitored phenomenon $\boldsymbol{\vartheta}$, estimated by a discrete time series model, is denoted as $\bar{\theta}^t$, and the estimation error is measured as $\varepsilon_t = \vartheta^t - \bar{\theta}^t$, which obeys the normal distribution $\mathcal{N}(0, \delta^2)$.

From *Definition 1* and *Definition 2*, we observe that $\vartheta^t = \bar{\theta}^t + \varepsilon_t$ is the normal distribution $\mathcal{N}(\bar{\theta}^t, \delta^2)$ and the parameter $P(\theta_i^t)$ in Eq.2 is the cumulative probability of $\theta_i - \Delta/2 < \vartheta^t \leq \theta_i + \Delta/2$ described as follows:

$$\begin{aligned} P(\theta_i^t) =& \frac{1}{\delta\sqrt{2\pi}} \int_{\theta_i - \Delta/2}^{\theta_i + \Delta/2} exp(-(x - \bar{\theta}^t)^2/2\delta^2) dx \\ =& \frac{1}{\sqrt{2\pi}} \int_{(\theta_i - \bar{\theta}^t - \Delta/2)/\delta}^{(\theta_i - \bar{\theta}^t + \Delta/2)/\delta} exp(-y^2/2) dy \\ =& \Phi((\theta_i - \bar{\theta}^t) + \Delta/2)/\delta) - \Phi((\theta_i - \bar{\theta}^t) - \Delta/2)/\delta) \end{aligned} \tag{5}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} exp(-x^2/2) dx$.

The undetermined parameter δ in *Definition 2* and Eq.5 can be updated by observations received at the sink node in the runtime stage as

$$\delta^2 = \frac{1}{w} \sum_{\tau = t-w+1}^{t} \varphi(r^\tau/\xi)\xi \tag{6}$$

where $r^\tau = \left|\bar{\theta}^\tau - \hat{\theta}^\tau\right|^2$; ξ is the median value of $\{r^\tau\}_{\tau=t-w+1}^{t}$;w is the window length less than t; $\hat{\theta}^\tau$ is defined in Eq.2 as the quantization level with maximum posteriori probability at time τ; and φ is taken as the Huber function [18]:

$$\varphi(x) = \begin{cases} x, & |x| \leq b \\ sign(x)b, & otherwise \end{cases} \tag{7}$$

where the positive constant b regulates the identification of outliers [17]. Specifically, if the estimation error is too large, the data is considered as an outlier and replaced by a boundary value b. The choice of b with about 95% confidence is four, assuming a normal distribution of the estimation error [18].

V. Experimental Evaluations

A. Evaluation Setup

The sources of the data, used in our evaluations, include sensor data about water temperature, dissolved oxygen in river water and river stage from the California Data Exchange Center (CDEC) [19]. It is important to note that the data sets differ in terms of autocorrelation properties and degrees of stationarity. In experiments, the data are quantized as 8-bit values with random bit errors. These errors represent transient errors occurring in both the sensor node and wireless channel. We assume the sink node knows perfect bit error probabilities. The residual errors are evaluated in terms of the normalized mean square error (NMSE) as follows:

$$e_{out} = \sqrt{\frac{1}{N} \sum_{\nu=1}^{N} \left| \frac{\vartheta^\nu - \hat{\theta}^\nu}{\vartheta^\nu} \right|^2} \tag{8}$$

In Eq.8, N is the total number of samples. An AR model is first fitted by an offline process with the Yule-Walker method in our evaluations [16]. If the order or prediction error of the fitted AR model is too large, a RHW model [17] will be used instead. In the runtime stage, the parameters of the AR and RHW models are updated by each window of observations with the Yule-Walker method [16] and the method detailed in [17], respectively. The residual errors e_{out} are compared with the NMSE e_{in} of the error-corrupted data received at the sink node as shown in Tab.1, where the bit error probability is 10^{-2}.

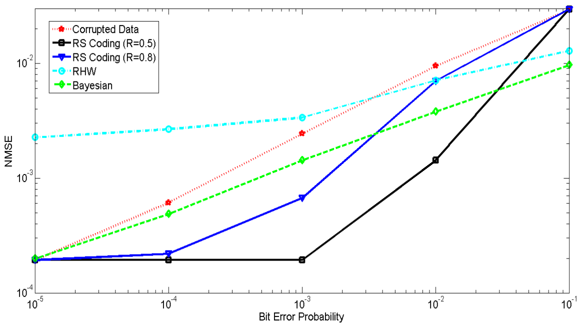
B. Comparisons with Traditional Methods

Fig.3 compares the NMSE of the data corrected by the proposed Bayesian model with the NMSE of the data corrected by Reed-Solomon (RS) coding and estimated by discrete time series models (i.e., RHW or AR models). NMSE floors of the data estimated by discrete time series models can soon be noticed, which are dominated by the estimation lags around relatively fast variations of the monitored physical phenomenon. NMSE floors of the data corrected by Bayesian model and RS coding, and the error-corrupted data received at the sink node are shown as well. However, these NMSE floors are much lower than the error floor of the data estimated by discrete time series models and the common source of these NMSE floors is quantization errors.

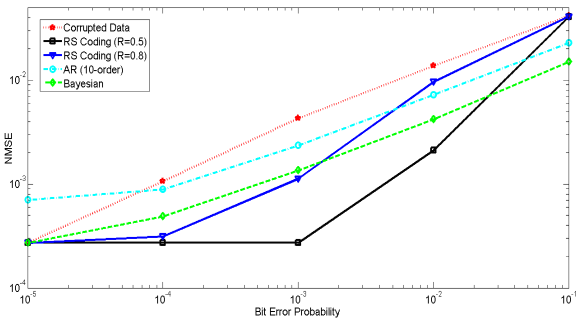
Due to error correction ability, the RS coding exceeds the statistical-based approaches at very low channel error levels. However, this trend is reversed in the presence of high channel error levels, because of the large NMSE incurred by the untreated decoding errors of RS coding. The other side effect of RS coding is high communication overhead. E.g., RS coding with 0.5 code rate (R = 0.5), which can nearly correct all errors at bit error probability 10^{-3}, will increase the communication overhead by $\frac{1}{0.5} - 1 = 100\%$. Even RS coding with R = 0.8, whose environment monitoring performance is close to the channel-aware Bayesian model at bit error probability 10^{-3}, still incurs $\frac{1}{0.8} - 1 = 25\%$ extra communication overhead. Unlike FEC, the overhead of our approach is computation increasing at the sink node, which is caused

TABLE I
DATA SET PARAMETERS AND IMPROVEMENT FACTORS OF THE PROPOSED CHANNEL-AWARE BAYESIAN MODEL

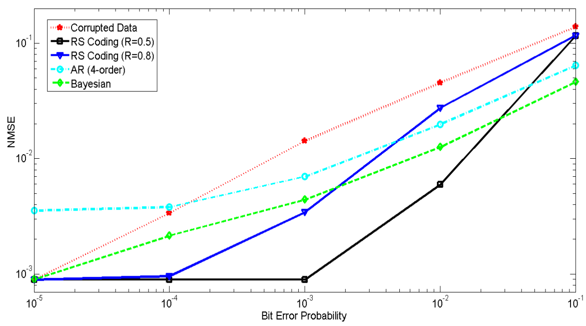
Data Set	Location	Time	Sampling Period	Sensor Type	Number of Samples	Data Source Model	Improvement Factor e_{in}/e_{out}
1	Lewiston	Feb.-Jul., 2010	1 Day	Water Temperature	181	RHW	2.50
2	Balls Ferry Bridge	Sep.-Dec., 2000	1 Hour	Dissolved Oxygen in River	2191	AR, 10-order	3.27
3	Rumsey Bridge	Apr.-Jul., 2010	1 Hour	River Stage	2448	AR, 4-order	3.62



(a) Data Set 1



(b) Data Set 2



(c) Data Set 3

Fig. 3. Performance comparisons between Bayesian model, RS coding and discrete time series models

by probabilities multiplications and can be solved by log-domain algorithms.

Depending on the channel error levels, the channel-aware Bayesian model will weight the quantization level with higher likelihood by more confidence. Particularly, at very low channel error levels (i.e., at bit error probability 10^{-4} and 10^{-5}), the quantization levels with small distances [1] (i.e., 0 or 1) to the channel output will have significant chance to be believed as the channel input by the channel-aware Bayesian model. Hence, at very low channel error levels, the channel-aware Bayesian model performs much better than the AR and RHW models that only explore temporal correlations among observations.

VI. CONCLUSION

In this paper, we have presented a channel-aware Bayesian model to reliably recover data from transient errors on the data transfer channel from the sensor node that monitors a physical phenomenon to the sink node. Using real sensor data, we have shown that a significant gain can be achieved by the proposed model compared with the channel-unaware models based on the correlations of the monitored phenomena. Moreover, our method performs closely to the RS coding with code rate 0.8, which brings 25% extra communication overhead. However, our model has no complexity increasing at the sensor node.

We currently assume the sink node knows perfect bit error probabilities. The model performance with true bit error probabilities still needs our future in-field tests. Moreover, the combinations of the proposed algorithms with distributed data fusion and signal detection algorithms considering noise at the sensor of individual sensor node are the most important extensions.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks,"*IEEE Commun. Mag.*, vol.40, no.8, pp.102-114, 2002.
[2] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: a survey,"*IEEE Commun. Surveys Tuts.*, vol.12, no.2, pp.159-170, 2010.
[3] J. F. Sanford and B. A. Hamilton, "On-line sensor calibration and error modeling using single actuator stimulus,"*in Proc. AERO 2009*, pp.1-11, 2009.
[4] S. M. Jahinuzzaman, M. Sharifkhani, and M. Sachdev, "An analytical model for soft error critical charge of nanometric SRAMs,"*IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol.17, no.9, pp.1187-1195, 2009.
[5] R. Hegde and N. R. Shanbhag, "Toward achieving energy efficiency in presence of deep submicron noise,"*IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol.8, no.4, pp.379-391, 2000.
[6] K. Itoh, M. Horiguchi, and T. Kawahara, "Ultra-low voltage nano-scale embedded RAMs,"*in Proc. ISCAS 2006*, pp.25-28, 2006.

[1] The distance of two quantization levels is defined as the Hamming distance of the binary representations of the both quantization levels.

[7] M.C. Vuran and I.F. Akyildiz, “Error control in wireless sensor networks:a cross layer analysis,”*IEEE/ACM Trans. Netw.*, vol.17, no.4, pp.1186-1199, 2009.
[8] E. Elnahrawy and B. Nath, “Context-aware sensors,” *LNCS*, pp.77-93, 2004.
[9] S. Mukhopadhyay, C. Schurgers, D. Panigrahi, and S. Dey, “Model-based techniques for data reliability in wireless sensor networks,”*IEEE Trans. Mobile Comput.*, vol.8, no.4, pp.528-543, 2009.
[10] T. Zhao and A. Nehorai, “Distributed sequential Bayesian estimation of a diffusive source in wireless sensor networks,”*IEEE Trans. Signal Process.*, vol.55, no.4, pp.1511-1524, 2007.
[11] A. Dogandzic and B. Zhang, “Distributed estimation and detection for sensor networks using hidden markov random filed models,”*IEEE Trans. Signal Process.*, vol.54, no.8, pp.3200-3215, 2006.
[12] B. Chen, L. Tong, and P. K. Varshney, “Channel-aware distributed detection in wireless sensor networks,”*IEEE Trans. Signal Process.*, vol.23, no.4, pp.16-26, 2006.
[13] J. Y. Wu, C. W. Wu, T. Y. Wang, and T. S. Lee, “Channel-aware decision fusion with unknown local sensor detection probability,”*IEEE Trans. Signal Process.*, vol.58, no.3, pp.1457-1463, 2010.
[14] O. Ozdemir, R. Niu, and P. K. Varshney, “Channel aware target location with quantized data in wireless sensor networks,”*IEEE Trans. Signal Process.*, vol.57, no.3, pp.1190-1202, 2009.
[15] H. Jeon, J. Choi, H.Lee, and J. Ha, “Channel-aware energy efficient transmission strategies for large wireless sensor networks,”*IEEE Signal Process.Lett.*, vol.17, no.7, pp.643-646, 2010.
[16] M. H. Hayes, *Statistical digital signal processing and modeling*, Wiley, 1996.
[17] S. Gelper, R. Fried, and C. Croux, “Robust forecasting with exponential and Holt-Winters smoothing,”*J.Forecast.*, vol.29, no.3, pp.285-300, 2010.
[18] R. A. Maronna, D. R. Martin, and V. J. Yohai, *Robust statistics: theory and methods*, Wiley, 2006.
[19] *CDEC: California Data Exchange Center*, California Dept. of Water Resources, http://cdec.water.ca.gov, 2010.