Optimization of Linear Wireless Sensor Networks for Serial Distributed Detection Applications

Gernot Fabeck, Rudolf Mathar
Institute for Theoretical Information Technology
RWTH Aachen University
D-52056 Aachen, Germany
Email: {fabeck, mathar}@ti.rwth-aachen.de

***Abstract*—Typical applications of wireless sensor networks include infrastructure monitoring and surveillance, where in many cases the geometry of the monitored object determines the topology of the deployed network. For example, important applications like pipeline monitoring and border surveillance feature a linear arrangement of wireless sensors. In this paper, we address the scalable optimization of linear sensor networks for serial distributed detection applications. In serial distributed detection, signal detection is performed collaboratively by multiple sensors arranged in serial until a final detection result is reached. By locally maximizing the Chernoff information at each sensor in the serial network, scalable solutions are obtained which only rely on local information. By considering the problem of detecting a deterministic signal in the presence of Gaussian noise, a detailed numerical study reveals interesting trade-offs and dependencies between communication constraints and detection performance.**

I. Introduction

Infrastructure monitoring and surveillance are among the important applications of wireless sensor networks. Very often, the geometry of the monitored object determines the topology of the wireless sensor network deployed, e.g., pipeline monitoring advocates a linear arrangement of sensors embedded in the outer surface of the pipeline [1], [2]. Since the detection of leaks, bursts and other anomalies is one of the core functions of pipeline monitoring, the adaptation of detection algorithms to the linear structure of the network is recommended or even necessary. Additionally, wireless sensors typically operate on limited energy budgets and are consequently subject to communication constraints. This recommends compression of observations at the sensors and transmission of quantized observations or local decisions [3], [4].

In serial distributed detection, an initial sensor starts the detection process by making a decision and transmitting it to its neighbor (see Fig. 1). Based on the received decisions form their predecessors, the succeeding sensors again form decisions and transmit them until a final sensor or gateway node is reached. The gateway node is responsible for the final detection result. The main problem is to design the local sensor decision rules and the gateway node decision rule with respect to an overall performance criterion. In principal, the globally optimal solution for the local sensor decision rules can be obtained by examining all the solutions of a system of coupled nonlinear equations which is very hard to solve [5]. Besides the fact that the computational complexity of the system of equations increases exponentially with the number of sensors, the corresponding solutions are not scalable, i.e., if a new sensor is added to the network, the decision rules of all the sensors have to be adapted. Both aspects imply the infeasibility of globally optimal design in realistic scenarios.

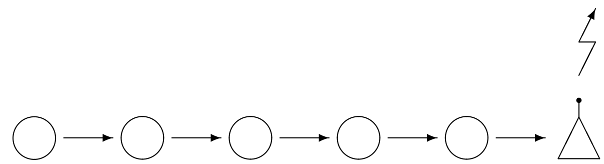

Fig. 1. Linear wireless sensor network with gateway node.

Suboptimal design of serial distributed detection systems is investigated in [6] and [7]. The authors consider local optimization of binary local detectors, i.e., every sensor in the network minimizes its own probability of error. The main drawback of this approach is that it is only applicable to binary quantization at the sensors and that the restriction to binary quantization is very disadvantageous for serial networks.

In this paper, we consider the optimization of serial distributed detection systems with an arbitrary number of quantization levels based on a local performance metric. An appropriate metric is derived from the asymptotic error exponents in binary hypothesis testing and is given by the Chernoff information. By locally maximizing the Chernoff information between the stochastic vectors of quantization probabilities, the optimization procedure only relies on local information. After the local sensor decision rules have been determined, the optimal decision rule of the gateway node can be derived. As the numerical results show, the presented approach enables efficient and scalable design of serial distributed detection systems and reveals interesting trade-offs between communication constraints and detection performance.

The remainder of this paper is organized as follows. In Section II, the problem of serial distributed detection with M-ary quantization at the local sensors is stated. The Chernoff information-based optimization procedure is presented in Section III. In Section IV, a numerical analysis of the proposed approach is given. Finally, we conclude in Section V.

978-1-4244-2519-8/10/$26.00 ©2010 IEEE

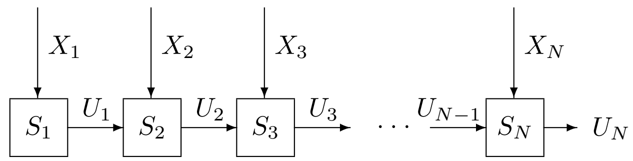

Fig. 2. Serial distributed detection system.

II. Serial Distributed Detection

The problem of serial distributed detection with M-ary quantization at the local sensors can be stated as follows (see Fig. 2). We consider a binary hypothesis testing problem with hypotheses H_0 and H_1 indicating the state of the monitored environment. The associated prior probabilities are $\pi_0 = P(H_0)$ and $\pi_1 = P(H_1)$. In order to detect the true state of nature, a network of N sensors $S_1, \ldots, S_N$ obtains random observations

$$(X_1, \ldots, X_N)' \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_N, \tag{1}$$

which are generated according to either H_0 or H_1. The random observations $X_1, \ldots, X_N$ are assumed to be conditionally independent across sensors given the underlying hypothesis, i.e., the joint conditional probability density functions of the observations factorize according to

$$f(x_1, \ldots, x_N | H_k) = \prod_{j=1}^{N} f_j(x_j | H_k), \quad k = 0, 1. \tag{2}$$

In the serial topology, the first sensor S_1 makes a decision $U_1 = \delta_1(X_1)$ which only depends on its own observation and subsequently transmits it to its neighbor (see Fig. 2). The succeeding sensors $S_2, \ldots, S_N$ form decisions

$$U_j = \delta_j(U_{j-1}, X_j), \quad j = 2, \ldots, N, \tag{3}$$

which depend on the received decision U_{j-1} of the preceding sensor as well as the own observation X_j.

A. Local sensor decision rules

In the general case of M-ary quantization at the sensors, the local sensor decision rules δ_j are mappings

$$\delta_1 \colon \mathcal{X}_1 \to \{1, \ldots, M\}, \tag{4}$$

$$\delta_j \colon \{1, \ldots, M\} \times \mathcal{X}_j \to \{1, \ldots, M\}, \quad 1 < j < N. \tag{5}$$

Warren and Willett have shown that local sensor decision rules leading to jointly optimal configurations under the minimum probability of error criterion are monotone log-likelihood ratio quantizers, provided that the observations are conditionally independent across sensors [8]. Hence, in the design of distributed detection systems under the assumption of conditional independence, it is necessary only to consider sensor decision rules δ_j that can be parameterized by a set of real quantization thresholds $(\tau_j^{(1)}, \ldots, \tau_j^{(M-1)})$, where $\tau_j^{(0)} = -\infty$, $\tau_j^{(M)} = \infty$, and $\tau_j^{(k)} \leq \tau_j^{(k+1)}$. In this way, the sensors $S_1, \ldots, S_{N-1}$ are characterized by the conditional quantization probabilities

$$\alpha_j^{(k)} = P(U_j = k | H_0) = P(\tau_j^{(k-1)} < L_j \leq \tau_j^{(k)} | H_0), \tag{6}$$

$$\beta_j^{(k)} = P(U_j = k | H_1) = P(\tau_j^{(k-1)} < L_j \leq \tau_j^{(k)} | H_1), \tag{7}$$

where the local log-likelihood ratios are given by

$$L_1 = \log \left(\frac{f_1(X_1 | H_1)}{f_1(X_1 | H_0)} \right) \tag{8}$$

for the first sensor S_1, and by

$$L_j = \log \left(\frac{P(U_{j-1} | H_1) f_j(X_j | H_1)}{P(U_{j-1} | H_0) f_j(X_j | H_0)} \right) \tag{9}$$

for the sensors $S_2, \ldots, S_{N-1}$. In the latter case, L_j is the combined log-likelihood ratio of the received decision U_{j-1} and the local observation X_j. The stochastic vectors of quantization probabilities

$$\boldsymbol{\alpha}_j = (\alpha_j^{(1)}, \ldots, \alpha_j^{(M)})', \tag{10}$$

$$\boldsymbol{\beta}_j = (\beta_j^{(1)}, \ldots, \beta_j^{(M)})' \tag{11}$$

are computable given the conditional distributions of the log-likelihood ratio L_j and the quantization thresholds $\tau_j^{(1)}, \ldots, \tau_j^{(M-1)}$ for each $j = 1, \ldots, N-1$.

B. Final sensor decision rule

At the final sensor S_N, the received decision U_{N-1} and the local observation X_N are combined to the final detection result $U_N = \delta_N(U_{N-1}, X_N) \in \{0, 1\}$. The global performance metric of the serial network is the probability of error P_e of the final sensor S_N according to

$$P_e = \pi_0 P(U_N = 1 | H_0) + \pi_1 P(U_0 = 0 | H_1) \tag{12}$$

$$= \pi_0 P_f + \pi_1 P_m \tag{13}$$

which can be written as a weighted sum of the probability of false alarm $P_f = P(U_N = 1 | H_0)$ and the probability of miss $P_m = P(U_N = 0 | H_1)$ of sensor S_N. The optimal decision rule at the final sensor S_N under the minimum probability of error criterion can be performed by evaluating a log-likelihood ratio test with a variable threshold according to

$$L_N \underset{U_N = 0}{\overset{U_N = 1}{\gtrless}} \tau_N^{(k)}, \tag{14}$$

where

$$L_N = \log \left(\frac{f_N(X_N | H_1)}{f_N(X_N | H_0)} \right) \tag{15}$$

is the log-likelihood ratio of the local observation X_N and

$$\tau_N^{(k)} = \log \left(\frac{\pi_0 P(U_{N-1} = k | H_0)}{\pi_1 P(U_{N-1} = k | H_1)} \right) = \log \left(\frac{\pi_0 \alpha_{N-1}^{(k)}}{\pi_1 \beta_{N-1}^{(k)}} \right) \tag{16}$$

is the decision threshold for the received decision $U_{N-1} = k$.

C. Detection error probabilities

For the optimal decision rule at the final sensor S_N according to (14), the probability of false alarm P_f and the probability of miss P_m can be calculated by applying the theorem of total probability. After some calculations, we obtain the expressions

$$P_f = \sum_{k=1}^{M} \alpha_{N-1}^{(k)} \Big(1 - F_{L_N}(\tau_N^{(k)}|H_0)\Big), \tag{17}$$

$$P_m = \sum_{k=1}^{M} \beta_{N-1}^{(k)} F_{L_N}(\tau_N^{(k)}|H_1) \tag{18}$$

for the probability of false alarm P_f and the probability of miss P_m, where $F_{L_N}(\cdot|H_k)$ denotes the conditional cumulative distribution function of the log-likelihood ratio L_N under hypothesis H_k, $k = 0, 1$.

III. Chernoff Information-Based Optimization of Local Sensor Decision Rules

In this section, we motivate and present the Chernoff information-based optimization procedure for the local sensor decision rules. The rationale behind this approach is that the Chernoff information arises as asymptotic error exponent in Bayesian hypothesis testing [9].

A. Hypothesis testing and Chernoff information

If we assume conditionally independent and identically distributed (i.i.d.) sensor observations $X_1, \ldots, X_N$, the local conditional probability density functions $f_j(\cdot|H_k)$ are the same for all $j = 1, \ldots, N$, and we can write

$$H_0 \colon X_j \sim f_0, \tag{19}$$

$$H_1 \colon X_j \sim f_1, \tag{20}$$

where f_k is the conditional probability density function under hypothesis H_k for all sensors. The Chernoff information between the two distributions f_0 and f_1 is defined as

$$D^* = C(f_0, f_1) = -\min_{0 \le t \le 1} \log \int f_0(x)^t f_1(x)^{1-t} dx. \tag{21}$$

If a centralized observer has access to the unquantized observations $X_1, \ldots, X_N$ and uses the Bayes optimal decision rule, for the probability of error P_e asymptotically it holds that

$$\lim_{N \to \infty} \frac{\log P_e}{N} = -D^*. \tag{22}$$

In other words, for large N we obtain

$$P_e \approx \exp(-ND^*), \tag{23}$$

i.e., the Chernoff information D^* is the asymptotic error exponent in minimum probability of error hypothesis testing. Intuitively, in the unquantized case considered above, every sensor contributes with the full Chernoff information D^* to the exponent in (23). The higher the contributed Chernoff information D^*, the lower the probability of error P_e at the centralized observer. This motivates the approach that in the case of M-ary quantization of observations at the sensors, the quantization thresholds at each sensor should be chosen in such a way that the Chernoff information between the stochastic vectors of quantization probabilities (10) and (11) is maximized throughout the network.

B. Chernoff information-based optimization

Analogously to definition (21), the Chernoff information between two stochastic vectors $\boldsymbol{p} = (p_1, \ldots, p_M)'$ and $\boldsymbol{q} = (q_1, \ldots, q_M)'$ in $\mathbb{R}^M$ is given by

$$D^* = C(\boldsymbol{p}, \boldsymbol{q}) = -\min_{0 \le t \le 1} \log \sum_{k=1}^{M} p_k^t q_k^{1-t}. \tag{24}$$

Starting with the first sensor S_1, the succeeding sensors $S_2, \ldots, S_{N-1}$ maximize the Chernoff information $C(\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j)$ between the stochastic vectors of quantization probabilities $\boldsymbol{\alpha}_j$ and $\boldsymbol{\beta}_j$. Thereby, it is assumed that every sensor S_j has knowledge of its own observation statistics given by the conditional marginal probability density functions $f_j(\cdot|H_k)$, $k = 0, 1$, and that it has knowledge of the stochastic vectors of quantization probabilities $\boldsymbol{\alpha}_{j-1}$ and $\boldsymbol{\beta}_{j-1}$ of its predecessor. The observation statistics or quantization probabilities of the other sensors do not have to be available at sensor S_j. Furthermore, the knowledge of the prior probabilities π_0 and π_1 is only necessary at the final sensor S_N.

Based on the knowledge available locally at the sensors, the quantization thresholds of sensor S_j are optimized in such a way that the Chernoff information $C(\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j)$ between the corresponding stochastic vectors of quantization probabilities is maximized, i.e.

$$(\tau_{j,opt}^{(1)}, \ldots, \tau_{j,opt}^{(M-1)}) = \underset{(\tau_j^{(1)}, \ldots, \tau_j^{(M-1)})}{\operatorname{argmax}} C(\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j). \tag{25}$$

Since no closed-form solution of (25) exists, the optimization is done numerically.

IV. Numerical Results

In the following, we provide a detailed numerical analysis of the Chernoff information-based optimization procedure. Thereby, we consider the problem of detecting a deterministic signal in the presence of Gaussian noise. First, we show how the maximum Chernoff information evolves as a function of the number of sensors in the serial network for different values of the local observation signal-to-noise ratio (SNR). After that, we compare the performance of serial distributed detection systems for a varying number of quantization levels at the local sensors and study the influence of the local observation SNR. This reveals interesting trade-offs and dependencies between communication constraints and detection performance for different sensing scenarios.

A. Distribution of sensor observations

As an illustrative example, we consider the problem of detecting the presence or absence of a deterministic signal in

Gaussian noise, i.e., we assume that the random observations $X_1, \ldots, X_N$ at the local sensors are distributed according to

$$H_0\colon X_j \sim \mathcal{N}(0, \sigma_j^2), \tag{26}$$

$$H_1\colon X_j \sim \mathcal{N}(\mu_j, \sigma_j^2), \tag{27}$$

for $j = 1, \ldots, N$. The variance σ_j^2 describes the Gaussian background noise and the mean μ_j indicates the deterministic signal component under hypothesis H_1 at sensor S_j. Accordingly, the local observation SNR at sensor S_j is given by

$$\mathrm{SNR}_j = 10 \log_{10} \left(\frac{\mu_j^2}{\sigma_j^2} \right) \quad [\mathrm{dB}]. \tag{28}$$

The log-likelihood ratio L_1 of the first sensor is a Gaussian random variable with conditional distributions according to

$$H_0\colon L_1 \sim \mathcal{N}\Big(-\frac{\mu_1^2}{2\sigma_1^2}, \frac{\mu_1^2}{\sigma_1^2} \Big), \tag{29}$$

$$H_1\colon L_1 \sim \mathcal{N}\Big(\frac{\mu_1^2}{2\sigma_1^2}, \frac{\mu_1^2}{\sigma_1^2} \Big). \tag{30}$$

The combined log-likelihood ratios $L_2, \ldots, L_{N-1}$ of the succeeding sensors $S_2, \ldots, S_{N-1}$ are essentially convolutions of discrete and continuous random variables and accordingly described by Gaussian mixture distributions given by

$$H_0\colon L_j \sim \sum_{k=1}^{M} \alpha_{j-1}^{(k)} \mathcal{N}\Big(\log \frac{\beta_{j-1}^{(k)}}{\alpha_{j-1}^{(k)}} - \frac{\mu_j^2}{2\sigma_j^2}, \frac{\mu_j^2}{\sigma_j^2} \Big), \tag{31}$$

$$H_1\colon L_j \sim \sum_{k=1}^{M} \beta_{j-1}^{(k)} \mathcal{N}\Big(\log \frac{\beta_{j-1}^{(k)}}{\alpha_{j-1}^{(k)}} + \frac{\mu_j^2}{2\sigma_j^2}, \frac{\mu_j^2}{\sigma_j^2} \Big). \tag{32}$$

The log-likelihood ratio L_N of the final sensor S_N is conditionally distributed according to

$$H_0\colon L_N \sim \mathcal{N}\Big(-\frac{\mu_N^2}{2\sigma_N^2}, \frac{\mu_N^2}{\sigma_N^2} \Big), \tag{33}$$

$$H_1\colon L_N \sim \mathcal{N}\Big(\frac{\mu_N^2}{2\sigma_N^2}, \frac{\mu_N^2}{\sigma_N^2} \Big), \tag{34}$$

i.e., it is again a Gaussian random variable. For the ease of presentation, we assume in the following that all sensors $S_1, \ldots, S_N$ have the same local observation SNR.

B. Performance of serial distributed detection systems

First, we illustrate how the maximum Chernoff information evolves as a function of the number of sensors in the serial network for different values of the local observation SNR. Fig. 3 shows the results for serial networks consisting of up to 50 quaternary sensors, i.e., sensors with $M = 4$ quantization levels. For low and medium observation SNR of 0 and 2 dB, the slope of the curve is low, whereas for high SNR of 6 dB the slope increases significantly. In all three considered cases, the slope of the curve decreases with the number of sensors. This indicates that the benefit of additional sensors will be particularly large for small and medium-sized networks.

Fig. 4 illustrates the numerical results of the Chernoff information-based optimization procedure at low observation SNR. The probability of error P_e is evaluated for serial networks consisting of $N = 2, \ldots, 50$ binary ($M = 2$), ternary ($M = 3$), and quaternary ($M = 4$) sensors at a local observation SNR of 0 dB. The prior probabilities are assumed to be $\pi_0 = \pi_1 = 0.5$. The probability or error of serial distributed detection systems with ternary and quaternary sensors is considerably smaller compared to serial distributed detection systems with binary sensors. For example, in order to obtain a probability of error of $P_e \approx 0.1$, one needs either 28 binary sensors, 12 ternary sensors or 9 quaternary sensors.

Fig. 5 illustrates the numerical results at a medium observation SNR of 2 dB. Again, the probability of error is evaluated for serial networks consisting of $N = 2, \ldots, 50$ binary, ternary and quaternary sensors and the prior probabilities are assumed to be equal. In order to obtain a probability of error of $P_e \approx 0.05$, one needs either 34 binary sensors, 13 ternary sensors or 9 quaternary sensors.

Finally, Fig. 6 illustrates the numerical results at a high observation SNR of 6 dB. In order to obtain a probability of error of $P_e \approx 0.01$, one needs either 20 binary sensors, 9 ternary sensors or 7 quaternary sensors. It is very interesting to note the large gap between binary and ternary sensors in all SNR regimes, an effect which occurs when the number of quantization levels is increased just from $M = 2$ to $M = 3$.

V. Conclusions

In this paper, we have presented a scalable approach to the optimization of linear sensor networks for serial distributed detection that only relies on local information. The approach is based on the local maximization of the Chernoff information between the stochastic vectors of quantization probabilities at the sensors. By considering the problem of detecting a deterministic signal in the presence of Gaussian noise, numerical results are obtained which reveal interesting trade-offs and dependencies between communication constraints and detection performance. In particular, the effect of binary, ternary and quaternary quantization of sensor observations on the probability of error of serial distributed detection systems is illustrated.

Acknowledgment

This work was partly supported by the UMIC excellence cluster of RWTH Aachen University.

References

[1] I. Jawhar, N. Mohamed, and K. Shuaib, "A framework for pipeline infrastructure monitoring using wireless sensor networks," in *Proceedings of WTS'07*, Pomona, CA, April 2007.

[2] I. Stoianov, L. Nachman, S. Madden, and T. Tokmouline, "PIPENET: A wireless sensor network for pipeline monitoring," in *Proceedings of IPSN'07*, Cambridge, MA, April 2007.

[3] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407-416, Feb. 2003.

[4] J.-F. Chamberland and V. V. Veeravalli, "Wireless sensors in distributed detection applications," in *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 16-25, May 2007.

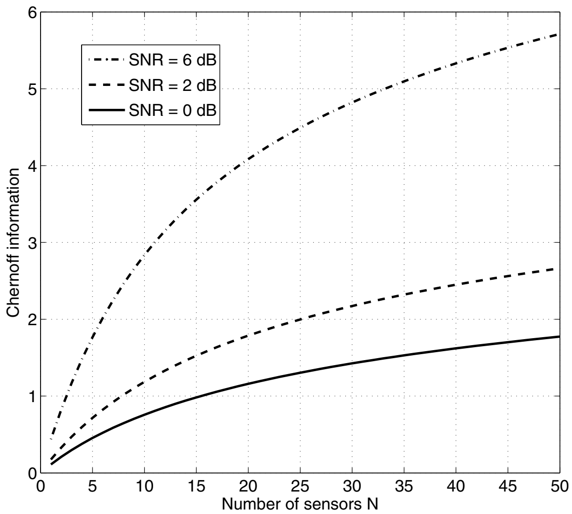

Fig. 3. Chernoff information in serial networks consisting of quaternary sensors as a function of the number of sensors for varying observation SNR.

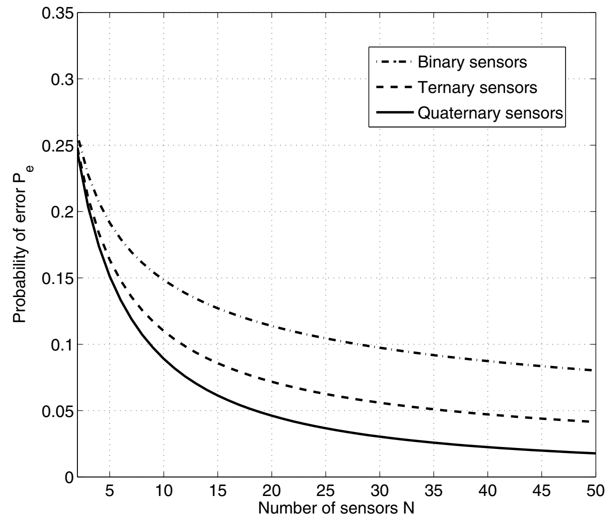

Fig. 4. Probability of error P_e of Chernoff information-based serial networks with binary, ternary, and quaternary sensors at observation SNR of 0 dB.

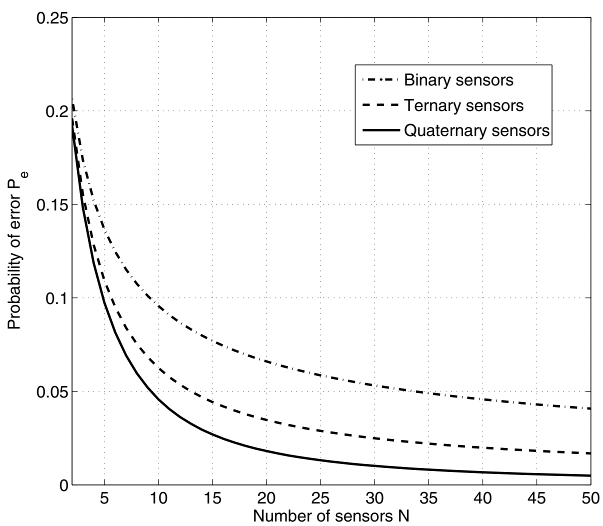

Fig. 5. Probability of error P_e of Chernoff information-based serial networks with binary, ternary, and quaternary sensors at observation SNR of 2 dB.

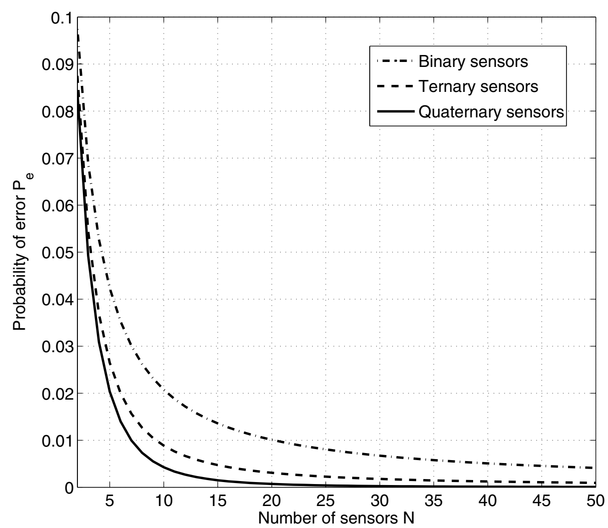

Fig. 6. Probability of error P_e of Chernoff information-based serial networks with binary, ternary, and quaternary sensors at observation SNR of 6 dB.

[5] P. K. Varshney, *Distributed Detection and Data Fusion*, Springer, 1997.

[6] J. D. Papastravou and M. Athans, "Distributed detection by a large team of sensors in tandem," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 639-653, July 1992.

[7] P. F. Swaszek, "On the performance of serial networks in distributed detection," in *IEEE Transactions on on Aerospace and Electronic Systems*, vol. 29, no. 1, pp. 254-260, January 1993.

[8] D. Warren and P. Willett, "Optimum quantization for detector fusion: Some proofs, examples, and pathologies," *Journal of the Franklin Institute*, vol. 336, pp. 323-359, 1999.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.