Scalable PHY-Layer Security for Distributed Detection in Wireless Sensor Networks

Reza Soosahabi, *Student Member, IEEE*, and Mort Naraghi-Pour, *Member, IEEE*

***Abstract*—The problem of binary hypothesis testing is considered in a bandwidth-constrained densely populated low-power wireless sensor network operating over insecure links. Observations of the sensors are quantized and encrypted before transmission. The encryption method maps the output of the quantizer to one of the possible quantizer output levels randomly according to a probability matrix. The intended (ally) fusion center (AFC) is aware of the encryption keys (probabilities) while the unauthorized (third party) fusion center (TPFC) is not. A constrained optimization problem is formulated from the point of view of AFC in order to design its decision rule along with the encryption probabilities. The objective function to be minimized is the error probability of AFC and the constraint is a lower bound on the error probability of TPFC. In the binary case the optimal solution is found and in the nonbinary case a good suboptimal solution is analytically obtained. Numerical results are presented to show that it is possible to degrade the error probability of TPFC significantly and still achieve very low probability of error for AFC. The proposed method which may be considered a PHY-layer security scheme is highly scalable since it does not increase the packet overhead or transmit power of the sensors and has very low computational complexity. A scheme is described to randomize the keys so as to defeat any key space exploration attack.**

***Index Terms*—Decentralized detection, decision fusion rule, information security, soft decision, wireless sensor networks.**

I. Introduction

WIRELESS sensor networks (WSN) have applications in many military and civilian areas including intrusion detection and surveillance, medical monitoring, emergency response, environmental monitoring, target detection and tracking, and battlefield assessment. Providing security in WSNs is a challenging task. In the sensor nodes the resources such as energy supply, processing power, memory size and communication bandwidth are severely limited. Another difficulty arises from the large number of nodes in the network. Future networks are envisioned to consist of hundreds or thousands of nodes to implement ubiquitous networks. To keep the network cost down, as the number of nodes increases, the cost per node must be reduced. Therefore, it is unlikely that in the near future, technological advances will alleviate the scarcity of resources at the nodes. Despite these difficulties in many applications of WSNs security is as important as performance, if not more [1]. This calls for scalable security protocols with minimal resource requirements and low communication overhead.

Several security protocols have been recently proposed for WSNs to combat eavesdropping [1]–[4]. These schemes are mostly independent of the application at hand and adapt the traditional network security protocols using cryptography, authentication, and key management techniques to provide security at the link and network layer, albeit with more efficient implementation and resource utilization. However, the issue of scalability remains since these techniques provide security at the expense of increased energy consumption and bandwidth [5]. In this paper we consider a specific application of WSNs, namely the problem of distributed detection of the state of a phenomenon in an environment, and propose a security scheme which may be considered a physical layer technique since it only randomizes the content of sensor messages. Our method can be used in conjunction with other security protocols at higher layers to enhance the integrity of the network operation.

Distributed detection using WSNs has been extensively investigated [6]. In particular, optimal design of the fusion rule under different conditions of quantization at the individual sensors, topologies of the network, and channel conditions have been investigated in [7]–[11].

Mission-critical applications of WSNs involving distributed detection demand operational security [11]. On the other hand, networks must cope with insecure links. Due to the limited power and low bandwidth, we assume that nodes transmit a quantized version of their observations to their intended (ally) fusion center (AFC). In addition to the AFC, an unauthorized (third-party) fusion center (TPFC) may also be observing the sensor transmissions and attempting to detect the state of the unknown hypothesis. In order to deteriorate the error probability of TPFC, each node uses a simple encryption mechanism whereby it maps the quantizer output level to one of the possible output levels randomly similar to the operation of a discrete memoryless channel (DMC). In the case of a binary quantizer this approach is similar to that in [12] where it is used for secure estimation over insecure links. It is assumed that AFC is aware of the encryption probabilities matrix and can minimize its probability of error accordingly. On the other hand, TPFC does not have access to the encryption matrix. To ensure that the encryption matrix cannot be estimated from the sensor nodes' transmissions, this matrix is (pseudo) randomly selected from a set of designed matrices. Therefore, TPFC has to perform distributed detection without any knowledge of the distributions of the transmitted messages (see Sections VI and VIII for more detail).

Manuscript received November 10, 2011; revised February 16, 2012; accepted March 29, 2012. Date of publication April 13, 2012; date of current version July 09, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Wade Trappe.

The authors are with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803 USA (e-mail: rsoosa1@lsu.edu; naraghi@lsu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2012.2194704

The remainder of this paper is organized as follows. In Section II we consider sensors with binary quantization (hard decision) and formulate the optimization problems from the point of view of AFC and TPFC. Then the optimization problems are solved analytically in Section III. In Section IV we study the case of sensors using M-ary quantization (soft decision). Subsequently, in Section V we derive a suboptimal solution for the AFC. Numerical results are presented in Section VI. Section VIII includes a method to prevent the TPFC from estimating the encryption keys from the sensors' transmitted messages. Finally, concluding remarks are given in Section IX.

II. Problem Statement (Hard Decision)

We consider a network of n sensors observing the state of an unknown hypothesis H where $H \in \{H_0, H_1\}$ and with prior probabilities of H_0 and H_1 being q_0 and q_1, respectively. Let X_i denote the observation of the ith sensor, $i = 1, 2, 3, \ldots, n$. It is assumed that given the hypothesis H_η, (η =0, 1), the observations $X_1, X_2, \cdots, X_n$ are independent and identically distributed. The conditional PDF of X_i under the hypothesis H_η is denoted by $p_\eta(x)$.

Each sensor i makes a decision $u_i \in \{0, 1\}$ regarding the state of the hypothesis H using the likelihood ratio test

$$\frac{p_1(x)}{p_0(x)} \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \lambda \tag{1}$$

where λ is a threshold which is assumed to be identical for all the sensors. The false alarm and detection probabilities of individual sensors, denoted by P_0 and P_1, respectively, are given by

$$P_\eta = P(u_i = 1|H_\eta), \quad \eta = 0, 1. \tag{2}$$

The decisions of individual sensors are to be transmitted to the AFC which must detect the state of H from the received bits. However, the sensors' transmissions may be observed by a third-party (enemy) fusion center (TPFC) who also wishes to detect the state of H. In order to protect the decisions of the sensors from this unauthorized party, we employ the following simple probabilistic cipher. The decision u_i of sensor i is encrypted to obtain z_i, where $P(z_i = 1|u_i = 0) = \pi_0$ and $P(z_i = 0|u_i = 1) = \pi_1$, and where it is assumed that $\pi_0 + \pi_1 \leq 1$. The encrypted bit z_i is then transmitted to the AFC and may also be observed by TPFC. We assume that the channel between the sensors and the fusion centers is error free. The conditional probabilities of z_i given H_0 or H_1 are given by

$$\begin{aligned} \theta_0 &\triangleq P(z_i = 0|H_0) = 1 - P_0 - \pi_0 + (\pi_0 + \pi_1)P_0 \\ \theta_1 &\triangleq P(z_i = 0|H_1) = 1 - P_1 - \pi_0 + (\pi_0 + \pi_1)P_1. \end{aligned} \tag{3}$$

It is assumed that AFC has prior knowledge of the encryption probabilities (keys) π_0 and π_1. On the other hand, TPFC cannot reliably estimate the encryption keys and must perform the distributed detection assuming that it has received the original decisions u_i, $i = 1, 2, \cdots, n$ (see Section VIII).

We consider a Bayesian detection problem where the performance criterion for each of the fusion centers is the probability of error. Specifically, our goal is to design the system parameters so as to minimize P_E^a, the probability of error for AFC, subject to a lower bound on P_E^t, the probability of error for TPFC.

The likelihood ratio test for each of the fusion centers AFC or TPFC leads to a k-out-of-n rule given by [13]

$$\hat{H} = \begin{cases} H_1, & \text{if} \sum_{i=1}^n z_i \geq k \\ H_0, & \text{if} \sum_{i=1}^n z_i < k \end{cases} \tag{4}$$

where k is an integer-valued fusion threshold to be chosen by AFC and TPFC. The error probability for both fusion centers has the same formula given by

$$P_E = q_0 P(\hat{H} = H_1|H_0) + q_1 P(\hat{H} = H_0|H_1). \tag{5}$$

Considering (3) and (4) we can write

$$P_E(k, \theta_0, \theta_1) = q_0 \phi(k, \theta_0) + q_1 (1 - \phi(k, \theta_1)) \tag{6}$$

where $\phi(k, \theta)$ is given by

$$\phi(k, \theta) \triangleq \sum_{i=k}^{n} \binom{n}{i} (1 - \theta)^i (\theta)^{n-i} \tag{7}$$

and where the dependence of P_E on k, θ_0 and θ_1 is shown explicitly.

We would like to note that as evident from (5), the formulas for the false alarm and detection probabilities and the probabilities of error for the two fusion centers are the same. However, these two fusion centers have different views of the network. Consequently, their thresholds [k in (4)] and their performances are different.

Before studying AFC and TPFC error probabilities, we note that θ_0 and θ_1 both depend on P_0 and P_1 whose values depend on the choice of λ. Unlike the cipher parameters which are assigned during message transmission, λ is a built-in parameter of the individual sensors and is usually assigned during the manufacturing. In particular, λ is often chosen to minimize the probability of error in the absence of any encryption [7]. Consequently, hereafter we assume that λ is the fixed parameter calculated accordingly. This implies that P_1 and P_0 are also fixed.

A. Optimization From TPFC's Point of View

As mentioned previously, TPFC is assumed to be unaware of the encryption keys and therefore assumes that $\pi_0 = \pi_1 = 0$. Note that it is not assumed here that TPFC is unaware of the encryption process but only the keys. In light of the discussion in Section VIII this assumption is valid. TPFC is, however, aware of the threshold value λ and chooses its fusion threshold, denoted k^t, to minimize its probability of error. Since λ is also chosen to minimize the probability of error in the absence of any encryption, then the optimal λ and k^t are obtained from the solution of the following problem:

$$\begin{aligned} P1: \quad & \min_{k,\lambda} P_E\left(k, 1 - P_0(\lambda), 1 - P_1(\lambda)\right) \\ & \text{subject to}: \quad 0 \leq k \leq n \end{aligned}$$

where the objective function above is obtained from (6) for $\pi_0 = \pi_1 = 0$ ($\theta_i = 1 - P_i(\lambda)$, $i = 0, 1$). We denote the optimal k and λ obtained from P1 by k^t and λ^*, respectively. The AFC can also solve this problem independently and so it is aware of the values of λ^* and k^t. This problem has been investigated in detail in [7] where the following theorem is proved.

Theorem 1: Given $k, P_E(k, 1-P_0(\lambda), 1-P_1(\lambda))$ is a quasi-convex function of λ and there is a unique λ that minimizes it.

This theorem ensures that the optimal λ can be calculated from gradient-based numerical algorithms. An algorithm is then proposed in [7] in two steps. First, for each k, $0 \le k \le n$, the optimal threshold λ_k which minimizes $P_E(k, 1-P_0(\lambda), 1-P_1(\lambda))$ is computed. Then the optimum k (denoted k^t here) along with the corresponding λ^* are selected which achieve the minimum probability of error.

Note that in the presence of encryption ($\pi_i \neq 0$, and $\theta_i \neq 1-P_i$, $i = 0, 1$), the actual performance of TPFC is given by

$$P_E^t = P_E(k^t, \theta_0, \theta_1). \tag{8}$$

B. Optimization From AFC's Point of View

The allied fusion center must choose its fusion threshold k^a along with the encryption parameters π_0 and π_1 (or equivalently θ_0 and θ_1), so as to minimize its probability of error. In addition it must ensure that the performance of TPFC is degraded through the application of the encryption process. Therefore, AFC attempts to solve the following constrained optimization problem:

$$P2: \quad \min_{k^a, \theta_0, \theta_1} P_E(k^a, \theta_0, \theta_1) \tag{9}$$

$$\text{subject to:} \quad 0 \le k^a \le n \tag{10}$$

$$\theta_1 \le \theta_0 \tag{11}$$

$$e_{\min} \le P_E^t(k^t, \theta_0, \theta_1) \le 0.5 \tag{12}$$

$$\theta_0 - \theta_1 \le \theta_0 P_1 - \theta_1 P_0 \tag{13}$$

$$\theta_0 P_1 - \theta_1 P_0 \le P_1 - P_0. \tag{14}$$

In the above, (11) is due to the fact that in (3), $P_0 \le P_1$. In (12), $e_{\min}$ is a design parameter to ensure a minimum probability of error for TPFC. Moreover, since TPFC makes a binary decision, the case of $P_E^t \ge 0.5$ is of no interest. Finally, (13) and (14) correspond to the fact that $\pi_1 \ge 0$ and $\pi_0 \ge 0$, respectively. These can be simply derived from (3).

Having computed the optimal values of θ_0 and θ_1 from P2, the cipher probabilities π_0 and π_1 can be obtained from (3). In the following we pursue an analytical solution to P2.

III. Optimization for AFC (Hard Decision)

The optimization for AFC is more complicated than the optimization for TPFC due to the additional constraints. A graphical representation of the constraints is provided below which helps us in obtaining the optimal solution analytically. Given k^t and λ^*, the shaded area in Fig. 1 demonstrates the feasible values for θ_0 and θ_1 with respect to the constraints in (11)–(14). As depicted in Fig. 1, the three constraints in (11), (13) and (14) form the triangle ΔOIA in which the set of feasible points (θ_0, θ_1) must reside. The dashed trajectory represents the curve $(1-P_0(\lambda), 1-P_1(\lambda))$ (as λ varies) and the point A corresponds to λ^*. This triangle is always obtuse and resides above the dashed curve due to the concavity of this curve. The three sides OI, OA, and AI correspond to the boundaries of the three constraints in (11) $(\theta_0 = \theta_1)$, (13) $(\pi_1 = 0)$, and (14) $(\pi_0 = 0)$, respectively. In Fig. 1 we have also included the contours of constant P_E^t such that depending on the value of $e_{\min}$ in (12), one of these contours may play an active role on the set of feasible (θ_0, θ_1). A typical example of such a constraint is indicated by the arc MN in Fig. 1. Considering this constraint some portion of ΔOIA around the vertex A is excluded from the feasible set of (θ_0, θ_1). Before proposing an analytical optimization, we can further trim the feasible region through the following lemma.

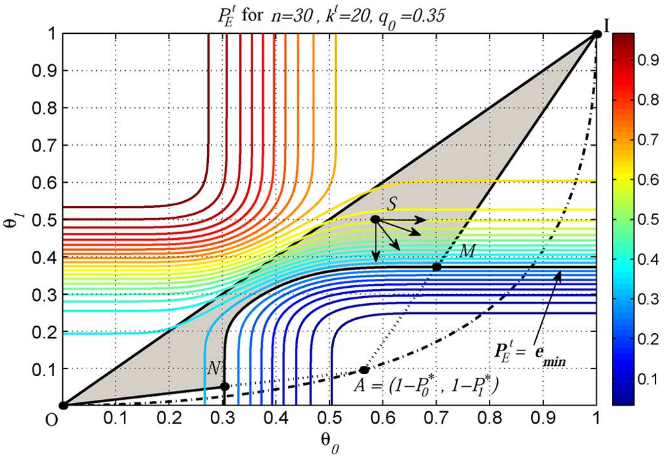


Fig. 1. Feasible region for AFC optimization defined by (11)–(14).

Lemma 1: An optimal pair of (θ_0, θ_1) meets at least one of the three constraints in (12), (13) and (14), with equality.

Proof: Calculating the partial derivatives of $P_E(k, \theta_0, \theta_1)$ with respect to θ_0 and θ_1, one can show that it is a monotone decreasing function of θ_0 and a monotone increasing function of θ_1. Suppose the point S in the shaded region is optimal where all the constraints are met with inequality. Now any change towards south east (see the arrows in Fig. 1) reduces $P_E(k, \theta_0, \theta_1)$. This violates the optimality of S. Clearly such changes are possible unless S satisfies one of the constraints (12)–(14) with equality. ■

The previous lemma limits the optimal solution for (θ_0, θ_1) to reside on the two lines OA and AI and an arc such as MN (e.g., a path like $ONMI$).

The solution to P2 must satisfy the Karush-Khun-Tucker (KKT) conditions. Avoiding trivial solutions and considering Lemma 1, the augmented objective function is written as

$$\begin{aligned} \mathcal{J} = P_E(k^a, \theta_0, \theta_1) &+ \zeta_1 \left(e_{\min} - P_E(k^t, \theta_0, \theta_1)\right) \\ &+ \zeta_2 \left((1-P_1)\theta_0 - (1-P_0)\theta_1\right) \\ &+ \zeta_3 \left((1-\theta_1)P_0 - (1-\theta_0)P_1\right) \end{aligned} \tag{15}$$

where $\zeta_1, \zeta_2, \zeta_3 \ge 0$, are the multipliers corresponding to the constraints in (12)–(14), respectively. From KKT conditions, $\zeta_i = 0$ implies that the optimal solution meets the corresponding constraint with inequality (inactive constraint). Then an optimal pair (θ_0, θ_1) must satisfy the following equations along with the constraints:

$$\nabla_{(\theta_0,\theta_1)} \mathcal{J} = 0 \tag{16}$$

$$\frac{\partial \mathcal{J}}{\partial \zeta_i} = 0 \quad \text{for} \quad \zeta_i \neq 0. \tag{17}$$

Here, the constraint on k^a cannot be included by calculating partial derivatives since it is an integer variable. Thus, a routine approach is to perform the above KKT optimization for a fixed $0 \leq k^a \leq n$. Once the optimal (θ_0, θ_1) are evaluated the optimal k^a can be computed from the MAP rule as discussed in the following.

The following two lemmas and Theorem 2 completely characterize the optimal solution for (θ_0, θ_1). The proofs are provided in the Appendix.

Lemma 2: An optimal (θ_0, θ_1) cannot satisfy only (12) with equality and (13), (14) with inequality, i.e., in (15) we cannot have $\zeta_1 \neq 0$, and $\zeta_2 = \zeta_3 = 0$.

Using the illustration in Fig. 1, Lemma 2 implies that if the optimal solution resides on the arc MN, then it can only be at M or N.

Lemma 3: An optimal (θ_0, θ_1) cannot only satisfy either (13) or (14) with equality and the remaining constraints with inequality, i.e., in (15) we cannot have $\zeta_2 \neq 0, \zeta_1 = \zeta_3 = 0$, or $\zeta_3 \neq 0, \zeta_1 = \zeta_2 = 0$.

Again Lemma 3 implies that if the optimal solution resides on line ON, (respectively, MI), then it must be at the point N (respectively, M). The following theorem summarizes the lemmas and completely characterizes the optimal solution to (9)–(14).

Theorem 2: The optimal solution for (θ_0, θ_1) satisfies (12) and either (13) or (14) with equality.

According to Theorem 2 the optimal solution to P2 lies where $P_E^t = e_{\min}$ contour intersects the lines $\pi_0(\theta_0, \theta_1) = 0$ and $\pi_1(\theta_0, \theta_1) = 0$, i.e., the point M or N in Fig. 1. Depending on the choice of $e_{\min}$ there are one or two such intersection points. Therefore, the optimal solution can be obtained by solving the following two nonlinear equations simultaneously using some efficient numerical method:

$$\begin{cases} P_E^t(k^t, \theta_0, \theta_1) = e_{\min} \\ \pi_0(\theta_0, \theta_1)\pi_1(\theta_0, \theta_1) = 0 \end{cases}. \tag{18}$$

Having obtained the optimal values of θ_0 and θ_1, the value of the optimal k^a is unique and can be calculated from the following equation according to the MAP rule [13]:

$$k^a(\theta_0, \theta_1) = \left\lceil \frac{\ln \Lambda - n \ln \frac{\theta_1}{\theta_0}}{\ln \frac{\theta_0(1-\theta_1)}{\theta_1(1-\theta_0)}} \right\rceil \tag{19}$$

where $\Lambda = q_0/q_1$.

IV. Problem Statement (Soft Decision)

In this section it assumed that sensor i quantizes its observation X_i using an M-level quantizer $\mathcal{Q}$ where $\mathcal{Q}(X_i) \in \mathcal{L} \triangleq \{l_1, l_2, \cdots, l_M\}$ for $i = 1, 2, \cdots, n$. The quantizer uses thresholds $-\infty = t_0 < t_1 < \cdots < t_M = \infty$, such that $\mathcal{Q}(x) = l_j$ if $t_{j-1} < x \leq t_j$, For $j = 1, 2 \cdots, M$ and $\eta = 0, 1$ let

$$a_\eta(l_j) \triangleq P\left(\mathcal{Q}(X_i) = l_j | H_\eta\right) = P(t_{j-1} < X_i \leq t_j | H_\eta). \tag{20}$$

Since the quantization process depends on the sensors' built-in technology, hereafter it is assumed that for $j = 1, 2, \cdots, M$ and $\eta = 0, 1, a_\eta(l_j)$ are fixed and known to both the AFC and TPFC. The optimal selection of the quantizer is investigated in [14].

We assume that the channel between the sensors and the FCs is error free and employ the following simple probabilistic cipher at the sensors where the decision $\mathcal{Q}(X_i)$ of sensor i is randomly encrypted to obtain Y_i, such that:

$$P\left(Y_i = l_k | \mathcal{Q}(X_i) = l_j\right) = \phi_{jk} \quad j, k = 1, 2, \cdots, M \tag{21}$$

for some ϕ_{jk}. The encrypted messages Y_i, $i = 1, 2, \cdots, n$ are then transmitted to AFC over an insecure link. For $\eta = 0, 1$ let $b_\eta(l_j) \triangleq P(Y_i = l_j | H_\eta)$, $j = 1, 2 \cdots, M$. We define

$$\begin{aligned} \boldsymbol{\alpha}_\eta &\triangleq [a_\eta(l_1), a_\eta(l_2), \cdots, a_\eta(l_M)] \\ \boldsymbol{\beta}_\eta &\triangleq [b_\eta(l_1), b_\eta(l_2), \cdots, b_\eta(l_M)]. \end{aligned} \tag{22}$$

Then $\boldsymbol{\beta}_\eta = \boldsymbol{\alpha}_\eta \boldsymbol{\Phi}$ where $\boldsymbol{\Phi} \triangleq [\phi_{ij}]$ is an $M \times M$ matrix. Again it is assumed that AFC has *a priori* knowledge of the encryption matrix $\boldsymbol{\Phi}$, but TPFC is not aware of the value of $\boldsymbol{\Phi}$ and therefore, it can only assume that it has received the original decisions $\mathcal{Q}(X_i)$, $i = 1, 2, \cdots, n$, i.e., it assumes $\boldsymbol{\Phi} = \boldsymbol{I}_{M \times M}$ (see Section VIII). Our goal is to design $\boldsymbol{\Phi}$ so as to minimize P_E^a, the probability of error for AFC, subject to a lower bound on P_E^t, the probability of error for TPFC.

The optimum decision rule for the two fusion centers is given by the log-likelihood ratio test [13], where for a received vector $\mathbf{y} = (y_1, y_2, \cdots, y_n)$

$$T(\mathbf{y}) \triangleq \frac{1}{n} \sum_{i=1}^{n} z_i \underset{H_0}{\overset{H_1}{\gtrless}} \tau \tag{23}$$

where for the AFC

$$\tau = \tau_a, \quad \text{and} \quad z_i \triangleq \log\left(\frac{b_1(y_i)}{b_0(y_i)}\right) \tag{24}$$

and for the TPFC

$$\tau = \tau_r, \quad \text{and} \quad z_i \triangleq \log\left(\frac{a_1(y_i)}{a_0(y_i)}\right). \tag{25}$$

The error probability for the two fusion centers is given by

$$P_E = q_0 P\left(T(\boldsymbol{Y}) \geq \tau | H_0\right) + q_1 P\left(T(\boldsymbol{Y}) < \tau | H_1\right) \tag{26}$$

where AFC and TPFC use their respective decision statistic $T(\mathbf{Y})$ and threshold τ. It can be seen that the values of the quantization levels, $l_j, j = 1, 2, \cdots, M$, do not affect the error probabilities. Invoking the central limit theorem [15] for large n and conditioned on H_η

$$T(\mathbf{Y}) | H_\eta \sim \mathcal{N}\left(\delta_\eta, \frac{\gamma_\eta}{n}\right) \tag{27}$$

where for the AFC

$$\begin{aligned} \delta_\eta &= \mathrm{E}_{\boldsymbol{\beta}_\eta}\left[\log\left(\frac{b_1(Y_i)}{b_0(Y_i)}\right)\right] \triangleq \mu_{a\eta} \\ \gamma_\eta^2 &= \mathrm{Var}_{\boldsymbol{\beta}_\eta}\left[\log\left(\frac{b_1(Y_i)}{b_0(Y_i)}\right)\right] \triangleq \sigma_{a\eta}^2 \end{aligned} \tag{28}$$

and for the TPFC

$$\begin{aligned} \delta_\eta &= \mathrm{E}_{\boldsymbol{\beta}_\eta}\left[\log\left(\frac{a_1(Y_i)}{a_0(Y_i)}\right)\right] \triangleq \mu_{t\eta} \\ \gamma_\eta^2 &= \mathrm{Var}_{\boldsymbol{\beta}_\eta}\left[\log\left(\frac{a_1(Y_i)}{a_0(Y_i)}\right)\right] \triangleq \sigma_{t\eta}^2. \end{aligned} \tag{29}$$

The subscripts for the operators E and Var indicate the distributions under which these are computed. However, note that TPFC does not adjust its fusion rule according to the statistics in (28) nor in (29). In the absence of knowledge of $\boldsymbol{\Phi}$ it has to assume that $\boldsymbol{\Phi} = \boldsymbol{I}_{M\times M}$, and it views (27) with the following statistics:

$$\begin{aligned}\delta_\eta &= \mathrm{E}_{\boldsymbol{\alpha}_\eta}\left[\log\left(\frac{a_1(Y_i)}{a_0(Y_i)}\right)\right] \triangleq \mu_{r\eta}\\ \gamma_\eta^2 &= \mathrm{Var}_{\boldsymbol{\alpha}_\eta}\left[\log\left(\frac{a_1(Y_i)}{a_0(Y_i)}\right)\right] \triangleq \sigma_{r\eta}^2. \end{aligned} \tag{30}$$

For ease of notation let $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_M)$ and $\boldsymbol{\omega} = (\omega_1, \omega_2, \cdots, \omega_M)$ where

$$\xi_i \triangleq \log\left(\frac{a_1(l_i)}{a_0(l_i)}\right), \qquad \omega_i \triangleq \log^2\left(\frac{a_1(l_i)}{a_0(l_i)}\right). \tag{31}$$

Then for $\eta = 0, 1$ we get

$$\mu_{t\eta} = \boldsymbol{\beta}_\eta \boldsymbol{\xi}^T, \quad \nu_{t\eta}^2 \triangleq \sigma_{t\eta}^2 + \mu_{t\eta}^2 = \boldsymbol{\beta}_\eta \boldsymbol{\omega}^T \tag{32}$$

and

$$\mu_{r\eta} = \boldsymbol{\alpha}_\eta \boldsymbol{\xi}^T, \quad \nu_{r\eta}^2 \triangleq \sigma_{r\eta}^2 + \mu_{r\eta}^2 = \boldsymbol{\alpha}_\eta \boldsymbol{\omega}^T. \tag{33}$$

The probability of error for the two fusion centers can be approximated by

$$\begin{aligned}P_E \approx& P_e(\tau, \delta_0, \delta_1, \gamma_0, \gamma_1)\\ =& q_0 Q\left(\frac{\sqrt{n}(\tau-\delta_0)}{\gamma_0}\right) + q_1\left(1 - Q\left(\frac{\sqrt{n}(\tau-\delta_1)}{\gamma_1}\right)\right)\end{aligned} \tag{34}$$

where τ, δ_η and σ_η take on the values corresponding to each fusion center.

A. Optimization from TPFC's Point of View

The TPFC chooses its fusion threshold τ_r to minimize its probability of error. Therefore the optimal τ_r is obtained from the solution of the following problem. Considering (34), optimal threshold τ_r is given by

$$\frac{(\tau_r - \mu_{r0})^2}{2\sigma_{r0}^2} - \frac{(\tau_r - \mu_{r1})^2}{2\sigma_{r1}^2} = \frac{1}{n}\ln\left(\frac{q_0\sigma_{r1}}{q_1\sigma_{r0}}\right). \tag{35}$$

The AFC can also solve this problem independently and so it is aware of the value of τ_r. The actual performance of TPFC, however, is given by

$$P_E^t = P_e(\tau_r, \mu_{t0}, \mu_{t1}, \sigma_{t0}, \sigma_{t1}). \tag{36}$$

The performance of TPFC is degraded since in (36), τ_r is not matched to the mean and variances μ_{t0}, μ_{t1}, σ_{t0} and σ_{t1}.

B. Optimization from AFC's Point of View

The optimization problem for AFC is stated as follows:

$$P1: \min_{\tau, \boldsymbol{\Phi}} P_e(\tau, \mu_{a0}, \mu_{a1}, \sigma_{a0}, \sigma_{a1}) \tag{37}$$

subject to :

$$0 \le \phi_{ij} \le 1 \ \forall i, j, \quad \text{and} \quad \boldsymbol{\Phi}\mathbf{1}_{M\times 1} = \mathbf{1}_{M\times 1} \tag{38}$$

$$e_{\min} \le P_e(\tau_r, \mu_{t0}, \mu_{t1}, \sigma_{t0}, \sigma_{t1}) \le 0.5 \tag{39}$$

where $\mathbf{1}_{M\times 1}$ indicates a column vector of all 1's. Note that the threshold τ in (37) is absent from the constraints. Therefore, for any given values of $(\mu_{a0}, \mu_{a1}, \sigma_{a0}, \sigma_{a1})$ the optimal τ can be calculated from the following:

$$\frac{(\tau_a - \mu_{a0})^2}{2(\sigma_{a0})^2} - \frac{(\tau_a - \mu_{a1})^2}{2(\sigma_{a1})^2} = \frac{1}{n}\ln\left(\frac{q_0\sigma_{a1}}{q_1\sigma_{a0}}\right). \tag{40}$$

Generally the optimization problem $P1$ is not mathematically tractable. In the following section we simplify the cost function and trim the feasible region to obtain a good suboptimal solution.

V. Optimization for AFC (Soft Decision)

The steps that follow will help us to obtain a mathematically tractable suboptimal solution to $P1$.

A. Reducing Number of Variables

From (28), (29), (37), and (39), the given statistics and the error probabilities are all functions of $\boldsymbol{\beta}_\eta$ which is obtained from a linear transformation of ϕ_{ij}. Since $\boldsymbol{\beta}_\eta$ contains fewer variables, it is more convenient to find the optimal $\boldsymbol{\beta}_\eta$ in $P1$ and compute the optimal $\boldsymbol{\Phi}$ from it with no loss in optimality. To this end, we need to express the linear constraints in (38) in terms of $\boldsymbol{\beta}_\eta$. The linear constraints in (38) form a *convex polyhedral set* which can be denoted by

$$\mathcal{P}_\phi \triangleq \{\boldsymbol{\Phi} \mid 0 \le \phi_{ij} \le 1 \ \forall i, j, \text{ and } \boldsymbol{\Phi}\mathbf{1}_{M\times 1} = \mathbf{1}_{M\times 1}\}. \tag{41}$$

According to [16], there is an equivalent (image) convex polyhedral set for $\mathcal{P}_\phi$ in $\boldsymbol{\beta}_\eta$ domain. For $\eta = 0, 1$, let

$$\mathcal{P}_\beta \triangleq \{\boldsymbol{\beta}_\eta \mid \boldsymbol{\beta}_\eta = \boldsymbol{\alpha}_\eta \boldsymbol{\Phi}, \ \boldsymbol{\Phi} \in \mathcal{P}_\phi\} \tag{42}$$

denote the equivalent set. Similar to $\mathcal{P}_\phi$, this is associated with some linear constraints on $\boldsymbol{\beta}_\eta$ which can be simply calculated using the instructions in [16].

B. Simplifying Constraints

For a given $(\mu_{t0}, \mu_{t1}, \sigma_{t0}, \sigma_{t1})$ satisfying (39), $\boldsymbol{\beta}_\eta$ only needs to satisfy the linear equations in (32). Since $\mu_{t\eta}$ and $\nu_{t\eta}$ are linear transformation of $\boldsymbol{\beta}_\eta$, similarly to Section V-A, one can find the convex polyhedral set of $\mu_{t\eta}$ and $\nu_{t\eta}$ corresponding to $\mathcal{P}_\beta$

$$\mathcal{P}_t \triangleq \left\{\mu_{t\eta}, \ \nu_{t\eta} \mid \mu_{t\eta} = \boldsymbol{\beta}_\eta \boldsymbol{\xi}^T, \ \nu_{t\eta}^2 = \boldsymbol{\beta}_\eta \boldsymbol{\omega}^T, \ \boldsymbol{\beta}_\eta \in \mathcal{P}_\beta\right\}. \tag{43}$$

For $\mu_{t\eta}, \nu_{t\eta} \in \mathcal{P}_t$, let us define

$$\boldsymbol{\varepsilon}(\tau, \mu_{t\eta}, \nu_{t\eta}) \triangleq P_e(\tau, \mu_{t0}, \mu_{t1}, \sigma_{t0}, \sigma_{t1}). \tag{44}$$

Thus the constraint in (39) can be replaced with

$$e_{\min} \le \boldsymbol{\varepsilon}(\tau_r, \mu_{t\eta}, \nu_{t\eta}) \le 0.5. \tag{45}$$

Our goal is to select $\mu_{t\eta}$ and $\nu_{t\eta}$ to satisfy (45). However, this selection substitutes the constraint in (39) with a few equality constraints as in (32). As a result the optimal solution to $P1$ may not be achievable. In fact the computed cost function may be far from optimal with these new constraints associated with arbitrary $\mu_{t\eta}$ and $\nu_{t\eta}$. A judicious choice for $\mu_{t\eta}$ and $\nu_{t\eta}$ is needed to ensure that the suboptimal solution computed with the new constraints is close to the optimal. The lemma that follows allows us to formulate an upper bound on the cost function based on $\mu_{t\eta}$ and $\nu_{t\eta}$.

Lemma 4: For any given $(\mu_{a\eta}, \sigma_{a\eta})$ and $(\mu_{t\eta}, \nu_{t\eta})$, $\eta = 0, 1$,

$$P_e(\tau_a, \mu_{a0}, \mu_{a1}, \sigma_{a0}, \sigma_{a1}) \le \boldsymbol{\varepsilon}(\tau^*, \mu_{t\eta}, \nu_{t\eta}) \tag{46}$$

where τ_a is given in (40) and τ^* is obtained from

$$\frac{(\tau^* - \mu_{t0})^2}{2(\sigma_{t0})^2} - \frac{(\tau^* - \mu_{t1})^2}{2(\sigma_{t1})^2} = \frac{1}{n} \ln\left(\frac{q_0 \sigma_{t1}}{q_1 \sigma_{t0}}\right). \tag{47}$$

Proof: For a fixed $\boldsymbol{\beta}_\eta$, the minimum achievable error probability according to the MAP rule is represented by $P_e(\tau_a, \mu_{a0}, \mu_{a1}, \sigma_{a0}, \sigma_{a1})$ for the test statistics described in (23) and (24). On the other hand, $\boldsymbol{\varepsilon}(\tau^*, \mu_{t\eta}, \nu_{t\eta})$ will be the minimum achievable error probability where the terms in the test statistic are given in (25). However, this does not correspond to the MAP rule. Due to the optimality of the MAP rule, the inequality in (46) holds. ■

The judicious choice of $(\mu_{t\eta}, \nu_{t\eta})$ denoted by $(\mu^*_{t\eta}, \nu^*_{t\eta})$ is now computed using the following optimization problem:

$$P1-1: \min_{\mu_{t\eta}, \nu_{t\eta}} \boldsymbol{\varepsilon}(\tau^*, \mu_{t\eta}, \nu_{t\eta}) \tag{48}$$

subject to :

$$\mu_{t\eta}, \nu_{t\eta} \in \mathcal{P}_t \tag{49}$$

$$e_{\min} \le \boldsymbol{\varepsilon}(\tau_r, \mu_{t\eta}, \nu_{t\eta}) \le 0.5. \tag{50}$$

Considering the fact that $\boldsymbol{\varepsilon}(\tau^*, \mu_{t\eta}, \nu_{t\eta})$ is monotonic with respect to $\mu_{t\eta}$ and $\nu_{t\eta}$, the previous problem can be efficiently solved using the KKT method. These optimal values are then used in the optimization problem in the next section.

C. Simplifying Cost Function

For large n, P^a_E is a decreasing function of μ_{a1} and an increasing function of μ_{a0}. It can also be inferred that for large n, the impact of $\sigma_{a\eta}$ becomes small compared to $\mu_{a\eta}$. Thus one is motivated to maximize $\mu_{a1} - \mu_{a1}$ instead of the cost function in $P1$. In [14] the same idea is used to find the optimal quantizer $\mathcal{Q}$ without the security issue. From (28), it can be seen that μ_{a0} and μ_{a1} are associated with Kullback-Leibler divergence

$$\mu_{a0} = -\mathcal{D}(\boldsymbol{\beta}_0 \| \boldsymbol{\beta}_1), \quad \mu_{a1} = \mathcal{D}(\boldsymbol{\beta}_1 \| \boldsymbol{\beta}_0). \tag{51}$$

Then $\mu_{a1} - \mu_{a1}$ can be written in form of J-divergence [14]

$$\mu_{a1} - \mu_{a0} = \mathcal{J}(\boldsymbol{\beta}_1 \| \boldsymbol{\beta}_0). \tag{52}$$

Finally, using the results from Sections V-A and B, the optimization problem is stated as

$$\tilde{P}1: \max_{\boldsymbol{\beta}_0, \boldsymbol{\beta}_1} \mathcal{J}(\boldsymbol{\beta}_1 \| \boldsymbol{\beta}_0) \tag{53}$$

$$\text{subject to :} \tag{54}$$

$$\boldsymbol{\beta}_\eta \in \mathcal{P}_\beta, \quad \eta = 0, 1 \tag{55}$$

$$\boldsymbol{\beta}_\eta \xi^T = \mu^*_{t\eta}, \quad \boldsymbol{\beta}_\eta \omega^T = \left(\nu^*_{t\eta}\right)^2, \quad \eta = 0, 1 \tag{56}$$

Theorem 3: $\mathcal{J}(\boldsymbol{\beta}_1 \| \boldsymbol{\beta}_0)$ is a convex function with respect to $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$.

TABLE I
PERFORMANCE OF PROPOSED METHOD IN HARD-DECISION CASE

n	γ (dB)	q_0	e_{min}	P^a_E	π_0	π_1	k^a
20	3	0.50	0.30	4.38 e-03	0.48	0	16
20	3	0.30	0.30	3.10 e-03	0	0.48	4
20	6	0.50	0.40	1.76 e-03	0.56	0	18
40	0	0.50	0.50	8.45 e-03	0	0.60	7
40	3	0.50	0.30	9.83 e-05	0	0.47	9
40	3	0.50	0.40	2.38 e-04	0.52	0	33
40	3	0.50	0.50	9.61 e-04	0	0.61	6
40	3	0.30	0.30	6.74 e-05	0	0.45	9
40	6	0.50	0.50	6.82 e-05	0	0.62	5
80	-3	0.50	0.50	7.81 e-04	0	0.45	22
80	0	0.50	0.50	3.05 e-05	0.37	0	50

To prove the previous theorem one needs to show that the Hessian matrix $[\nabla^2_{\boldsymbol{\beta}_0, \boldsymbol{\beta}_1} \mathcal{J}(\boldsymbol{\beta}_1 \| \boldsymbol{\beta}_0)]$ is positive definite. It is easy to show that it is a tridiagonal matrix with all positive eigenvalues.

Let $\mathcal{R}$ represent the feasible region for $\boldsymbol{\beta}_\eta$ determined by the constraints in (55) and (56). It is easy to verify that $\mathcal{R}$ is still a convex polyhedral set. Now our goal is to maximize a convex function within a polyhedral region. This maximum must be explored among the extreme points of the region [17]. In other words, we only need to examine the vertices of $\mathcal{R}$ to find the global maximum [18]. In [19], Balinsky has proposed an efficient algorithm to trace all the vertices of a convex polyhedral set. Having computed the optimal $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$, the optimal $\boldsymbol{\Phi}$ can now be calculated accordingly. There are many solutions for $\boldsymbol{\Phi}$ and we choose the one with the fewest number of nonzero elements so as to minimize the storage requirements. Next we obtain τ^a from (40).

VI. NUMERICAL RESULTS AND COMPARISON

We assume that the signal X_i received by sensor i is given by

$$X_i = s + N_i, \quad i = 1, 2, \cdots, n$$

where $s = d$ under H_1, $s = -d$ under H_0, and where $\{N_i\}_{i=1}^n$ are independent identically distributed Gaussian random variables with mean zero and variance σ^2.

A. Hard Decision

In this case each sensor uses its observation to make a decision according to (1) with the preassigned threshold λ^*. Detection and false alarm probabilities for an individual sensor are given by

$$P_0 = Q\left(\frac{\lambda^* + d}{\sigma}\right), \quad P_1 = Q\left(\frac{\lambda^* - d}{\sigma}\right). \tag{57}$$

We define $\gamma = 20\log(d/\sigma)$ as the sensors' signal-to-noise ratio (SNR). Table I shows the performance of the proposed algorithm for several values of n, γ and q_0 in the binary case. It can be seen that using the proposed method, very low error probabilities can be achieved at AFC while imposing high error probabilities on TPFC. For smaller values of $e_{\min}$, the constraint for error probability of TPFC is less stringent and therefore in such cases lower values of P^a_E can be achieved. We note that in the optimal solution for $\boldsymbol{\pi} = (\pi_0, \pi_1)$, only one of the elements of $\boldsymbol{\pi}$ is nonzero.

We have assumed that TPFC is aware of the priors q_0 and q_1. Therefore, the worst case error probability for TPFC is given by

TABLE II
AFC-Optimized Error Performance (Soft Decision Versus Hard Decision)

Case	n	γ (dB)	q_0	e_{min}	$P_E^a(2)$	$P_E^a(4)$	$P_E^a(8)$	$CS(2)$	$CS(4)$	$CS(8)$	$K(4)$	$K(8)$
$c1$	20	0	0.5	0.3	1.52 e-02	3.89 e-04	1.16 e-05	1.8	1.6	0.5	09	12
$c2$	20	0	0.3	0.3	1.37 e-02	1.11 e-04	8.49 e-06	1.7	1.3	0.4	07	15
$c3$	20	3	0.5	0.4	6.66 e-03	1.28 e-06	2.97 e-09	4.1	3.2	1.4	08	12
$c4$	40	-3	0.5	0.5	3.37 e-02	4.79 e-05	9.07 e-06	2.1	0.8	0.3	09	17
$c5$	40	0	0.5	0.3	1.15 e-03	5.59 e-08	4.83 e-10	3.6	2.4	0.6	09	12
$c6$	40	0	0.5	0.4	2.28 e-03	1.28 e-07	3.59 e-10	3.9	2.2	0.3	07	12
$c7$	40	0	0.5	0.5	8.45 e-03	1.14 e-06	1.43 e-09	4.5	3.2	1.0	08	12
$c8$	40	0	0.3	0.3	9.64 e-04	4.14 e-06	9.74 e-10	3.2	3.7	0.8	09	13
$c9$	40	3	0.5	0.5	3.14 e-03	3.18 e-12	3.35 e-17	9.0	6.4	1.6	08	12
$c10$	80	-6	0.5	0.5	1.99 e-02	4.82 e-05	1.40 e-05	2.1	0.7	0.5	08	18
$c11$	80	-3	0.5	0.5	2.15 e-03	3.22 e-08	7.37 e-10	3.0	2.9	0.7	09	17

$P^t_{\max} = \min\{q_0, q_1\}$, which results if TPFC completely ignores the sensors' transmissions and chooses the more likely hypothesis. Table I shows that this worst case scenario can be imposed on TPFC to ensure that $P^t_E \triangleq P^t_{\max}$. This implies that TPFC gains no information from the observation of sensors' transmissions.

B. Soft Decision

Table II shows the error probability of the soft decision quantizers for $M = 4, 8$, denoted by $P^a_E(M)$, for several values of n, γ and q_0. The quantizers have been designed according to [14]. For comparison we have also listed the performance of the hard decision scheme denoted by $P^a_E(2)$. In Table II K denotes the number of nonzero elements of $\mathbf{\Phi}^*$ where the soft decision with M levels is employed, i.e., $K(M)$ can be thought as the hash to store the encryption parameters and is generally smaller than M^2.

We introduce the *cost of security* denoted by $CS(M)$ which indicates the increase in the AFC error probability due to the protection against TPFC for a system with M quantization levels by

$$CS(M) \triangleq \log\left(\frac{P^a_E(M)}{P^{\min}_E(M)}\right) \tag{58}$$

where $P^{\min}_E(M)$ is the minimum achievable error probability with no encryption and M quantization levels. The value of $CS(M)$ is indicated in Table II. It can be seen that applying encryption in the binary case drastically increases the AFC error probability. However, as M increases this increase is diminished. This is better illustrated in Figs. 2 and 3, where the error probabilities of the secure and insecure systems are compared versus SNR and versus the number of sensors n, respectively. It can be seen that the secure system can provide acceptable error probability for most applications even in cases of low SNR or small network sizes.

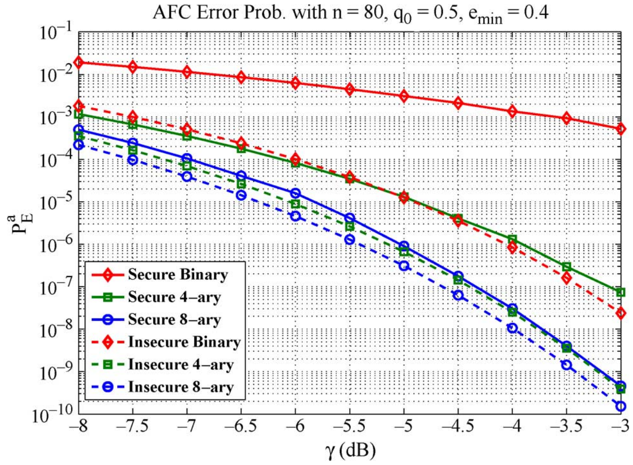


Fig. 2. Comparing AFC error performance versus SNR.

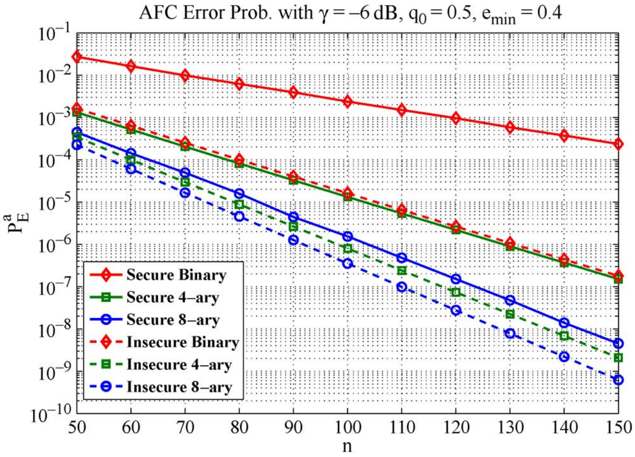


Fig. 3. Comparing AFC error performance versus number of nodes.

VII. Resource Usage and Complexity

To implement the proposed encryption method each sensor requires a random number generator. Several random number generators for low-power sensor networks have been proposed in recent years [20]–[22]. Each sensor node encrypts its quantized decision by comparing the output of the random number generator with the entries of the cipher matrix. This requires fewer than M (the number of quantization levels) comparisons. A table lookup is then used to choose the encrypted message. It can be seen that the increase in processing load due to the proposed method is small.

Additional memory is required to store the cipher matrix (fewer than M^2). The implementation of the random number generator and the encryption algorithm also requires some additional memory. These requirements are also fairly modest and can be easily accommodated given the current state of sensor hardware technology [23], [24].

Of the three main operations of a sensor node, namely sensing, data processing and communication, the latter consumes the maximum energy [23]. Our proposed method does not increase the communication overhead which is the main cause of energy consumption. The increase in energy consumption due to the modest processing and memory requirements is minor.

VIII. Randomization of Encryption Matrix

Given a fixed encryption matrix $\mathbf{\Phi}$, a TPFC with enough resources may attempt an attack through key space exploration by trying to estimate $\mathbf{\Phi}$ and the AFC's corresponding threshold denoted by $\tau^a(\mathbf{\Phi})$ from the transmitted symbols $y_1, y_2, \cdots, y_n$ of all the sensors.

As shown in Section VI, for different values of the error probability of TPFC, $e_{\min}$, our optimization algorithm results in different optimal matrices and thresholds for the AFC. Therefore by selecting two or more values of $e_{\min}$ we can design several matrices and AFC thresholds. Suppose for some integer L we have designed 2^L key/threshold pairs $\mathbf{\Phi}_m$, $\tau^a(\mathbf{\Phi}_m)$, for $m = 1, \cdots, 2^L$. Note that L can be as small as one. Each time the state of H is observed, one matrix is selected (pseudo) randomly by all the nodes and used to transmit the decisions of the sensors to the AFC. AFC is also aware of the matrix being used and will set its thresholds accordingly. This strategy can be implemented using a pseudo noise (PN) sequence generated by a long maximal length (linear feedback) shift register (MLSR) [25]. All the sensor nodes and AFC are equipped with identical MLSRs which start with the same initial state. Every time the state of H is detected, the sensors use L output bits from MLSR to select the matrix and encrypt their decisions before transmission to the AFC. This strategy can also be employed in the binary case.

Applying multiple cipher matrices changes the distributions of the transmitted messages. As a result TPFC has to perform distributed detection without any knowledge of the distributions of the transmitted symbols. For this reason we have assumed that TPFC assumes that no encryption has been used. However, other strategies of TPFC, whereby it may assume an arbitrary distribution will not improve its performance because the optimal $\mathbf{\Phi}$ will be very sensitive to the choice of $e_{\min}$. For example, in Table II, if AFC implements c5, and the TPFC modifies its decision rule suspecting c6 is in use, then the error probability of the TPFC (assuming no key randomization) will be still around 0.26.

IX. Conclusion

The problem of binary hypothesis testing is considered in a bandwidth-constrained low-power wireless sensor network operating over insecure links. Observations of the sensors are quantized and encrypted before transmission. We consider both hard decision (binary quantization) and soft-decision cases (multilevel quantization). The encryption method maps the output of the quantizer to one of the quantizer output levels randomly according to a probability matrix similar to the operation of a discrete memoryless channel. The AFC is aware of the encryption keys (probabilities) and can design its decision rule along with the encryption probabilities so as to impose a high probability of error on the unauthorized TPFC. The fusion rules are derived from the viewpoint of the two fusion centers and the encryption keys are designed so as to achieve a small probability of error for AFC with a lower bound on the error probability of TPFC. It is shown that by appropriate selection of the encryption parameters it is possible to impose a high error probability on TPFC while achieving low error probability for AFC. The proposed method which may be considered a PHY-layer security scheme for distributed detection is highly scalable due to its low computational complexity and no communication overhead.

Appendix

Proof of Lemma 2: Suppose that $P_E(k^t, \theta_0, \theta_1) = e_{\min}$ is the only constraint met with equality. Thus the KKT augmented cost function in (15) is reduced to the following:

$$\mathcal{J} = P_E(k^a, \theta_0, \theta_1) + \zeta_1 e_{\min} - P_E\left(k^t, \theta_0, \theta_1\right)\right). \quad (59)$$

We now set the partial derivatives of $\mathcal{J}$ with respect to θ_0 and θ_1 to zero, as in (16). This yields the following pair of equations:

$$n q_0 \binom{n-1}{k^a-1} (1-\theta_0)^{k^a-1} \theta_0^{n-k^a} = \zeta_1 n q_0 \binom{n-1}{k^t-1} (1-\theta_0)^{k^t-1} \theta_0^{n-k^t}. \quad (60)$$

$$n q_1 \binom{n-1}{k^a-1} (1-\theta_1)^{k^a-1} \theta_1^{n-k^a} = \zeta_1 n q_1 \frac{n-1}{k^t-1} (1-\theta_1)^{k^t-1} \theta_1^{n-k^t}. \quad (61)$$

Then dividing (60) by (61) we get

$$\left(\frac{1-\theta_0}{1-\theta_1}\right)^{k^a-k^t} = \left(\frac{\theta_0}{\theta_1}\right)^{k^a-k^t}. \quad (62)$$

Since $k^a \neq k^t$, the previous equation implies that $\theta_0 = \theta_1$ which, in view of the fact that $\pi_0 + \pi_1 < 1$, cannot hold. Thus it is impossible for the optimal solution to solely meet (12) with equality.

Proof of Lemma 3: Suppose that (13) is the only constraint met with equality implying that $\pi_1(\theta_0, \theta_1) = 0$. Thus, the KKT augmented objective function in (15) is now given by

$$\mathcal{J} = P_E(k^a, \theta_0, \theta_1) + \zeta_2\left((1-P_1^*)\theta_0 - (1-P_0^*)\theta_1\right). \quad (63)$$

Again we calculate the partial derivatives of $\mathcal{J}$ with respect to θ_0 and θ_1 and set them to zero. Dividing the resulting equations, we get

$$\frac{q_0}{q_1}\left(\frac{\theta_0}{\theta_1}\right)^{n-k^a}\left(\frac{1-\theta_0}{1-\theta_1}\right)^{k^a-1} = \frac{1-P_1^*}{1-P_0^*}. \quad (64)$$

Moreover, from $\pi_1(\theta_0, \theta_1) = 0$ we get

$$\frac{1-P_1^*}{1-P_0^*} = \frac{\theta_1}{\theta_0}. \quad (65)$$

Therefore

$$\frac{q_0}{q_1}\left(\frac{\theta_0}{\theta_1}\right)^{n-k^a}\left(\frac{1-\theta_0}{1-\theta_1}\right)^{k^a-1} = \frac{\theta_1}{\theta_0}. \quad (66)$$

This, however, implies that

$$\frac{\ln\frac{q_0}{q_1} - n\ln\frac{\theta_1}{\theta_0}}{\ln\frac{\theta_0(1-\theta_1)}{\theta_1(1-\theta_0)}} = k^a - 1 \quad (67)$$

which contradicts (19). Consequently, the initial assumption is incorrect so (13) cannot be the only constraint met with equality by the optimal solution. A similar argument can be used in the case of (14).

References

[1] Y. Wang, G. Attebury, and B. Ramamurthy, "A survey of security issues in wireless sensor networks," *Proc. IEEE Commun. Surveys Tutorials*, vol. 8, no. 2, pp. 2–23, 2006.

[2] X. Du and H.-H. Chen, "Security in wireless sensor networks," *IEEE Wireless Commun.*, vol. 15, no. 4, pp. 60–66, Aug. 2008.

[3] S. C. Karlof, N. , and D. Wagner, "Tinysec: A link layer security architecture for wireless sensor networks," in *Proc. 2nd Int. Conf. Embedded Networked Sensor Systems*, Nov. 2004, pp. 162–175.

[4] Y.-T. Wang and R. Bagrodia, "Sensec: A scalable and accurate framework for wireless sensor network security evaluation," in *Proc. 31st Int. Conf. Distributed Computing Systems Workshops (ICDCSW)*, Jun. 2011, pp. 230–239.

[5] X. Chen, K. Makki, K. Yen, and N. Pissinou, "Sensor network security: A survey," *IEEE Commun. Surveys*, vol. 11, no. 2, pp. 52–73, Mar. 2009.

[6] R. Tenney and N. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerospace Electronic Syst.*, vol. AES-17, no. 4, pp. 501–510, Jul. 1981.

[7] Q. Zhang, P. Varshney, and R. Wesel, "Optimal bi-level quantization of i.i.d. sensor observations for binary hypothesis testing," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 2105–2111, Jul. 2002.

[8] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerospace Electronic Syst.*, vol. AES-22, no. 1, pp. 98–101, Jan. 1986.

[9] B. Chen, R. Jiang, T. Kasetkasem, and P. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.

[10] R. Niu, B. Chen, and P. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1018–1027, Mar. 2006.

[11] W. Shi, T. Sun, and R. Wesel, "Quasi-convexity and optimal binary fusion for distributed detection with identical sensors in generalized gaussian noise," *IEEE Trans. Inform. Theory*, vol. 47, no. 1, pp. 446–450, Jan. 2001.

[12] T. Aysal and K. Barner, "Sensor data cryptography in wireless sensor networks," *IEEE Trans. Inform. Forensics Security*, vol. 3, no. 2, pp. 273–289, Jun. 2008.

[13] P. Varshney*, Distributed Detection and Data Fusion*. New York: Springer, 1997.

[14] C.-C. Lee and J.-J. Chao, "Optimum local decision space partitioning for distributed detection," *IEEE Trans. Aerospace Electronic Syst.*, vol. 25, no. 4, pp. 536–544, Jul. 1989.

[15] A. Papoulis and S. Pillai*, Probabilty, Random Variables and Stochastic Processes*, 4th ed. New York: McGraw-Hill, 2009.

[16] A. Barvinok, "Lattice points, polyhedra, and complexity," *Geometric Combinatorics, IAS/Park City Math. Ser.* vol. 13, pp. 19–62, 2007 [Online]. Available: http://www.math.lsa.umich.edu/barvinok/lectures.pdf

[17] G. Nash and A. Sofer*, Linear and Nonlinear Programming*, 1st ed. New York: McGraw-Hill, 1995.

[18] A. Boyd and L. Vandenberghe*, Convex Optimization*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[19] M. L. Balinski, "An algorithm for finding all vertices of convex polyhedral sets," *SIAM J. Appl. Math.*, vol. 9, pp. 72–88, 1961.

[20] D. Seetharam and S. Rhee, "An efficient pseudo random number generator for low-power sensor networks [wireless networks]," in *Proc. IEEE Annu. Int. Conf. Local Computer Networks*, Nov. 2004, pp. 560–562.

[21] A. Francillon and C. Castelluccia, "Tinyrng: A cryptographic random number generator for wireless sensors network nodes," in *Proc. 5th Int. Symp. WiOpt Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops*, Apr. 2007, pp. 1–7.

[22] D. Seetharam and S. Rhee, "An efficient pseudo random number generator for low-power sensor networks [wireless networks]," in *Proc. 29th Annu. IEEE Conf. Local Computer Networks*, Nov. 2004, pp. 560–562.

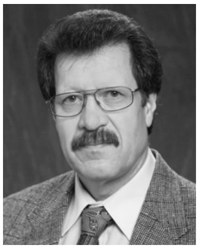
[23] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, pp. 393–422, 2002.

[24] A. Suri, S. Iyengar, and E. Cho, "Ecoinformatics using wireless sensor networks: An overview," in *Proc. 4th Int. Conf. Ecological Informatics*, Nov. 2006, vol. 1, no. 3, pp. 287–293.

[25] S. Golomb*, Shift Register Sequences*. San Francisco, CA: Holden-Day, 1967.

Reza Soosahabi (S'08) was born in June, 1988, in Tehran, Iran. He received the B.S. degree in electrical engineering with distinction from Amirkabir University of Technology, Tehran, Iran, in 2009, and the M.S. degree from Louisiana State University (LSU), Baton Rouge, in 2011.

He was appointed as a Graduate Researcher/Teaching Assistant in the Electrical and Computer Engineering Department, LSU, from 2009 to 2011. His research interests include cognitive radios, wireless sensor networks, statistical signal processing and discrete mathematics.



Mort Naraghi-Pour (S'81–M'87) was born in Tehran, Iran, on May 15, 1954. He received the B.S.E. degree from Tehran University, Tehran, in 1977, and the M.S. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, in 1983 and 1987, respectively.

In 1978, he was a student at the Philips International Institute, Eindhoven, The Netherlands, where he also did research with the Telecommunication Switching Group of the Philips Research Laboratories. Since August 1987, he has been with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, where he is currently an Associate Professor. From June 2000 to January 2002, he was a Senior Member of Technical Staff at Celox Networks, Inc., a network equipment manufacturer in St. Louis, MO. His research and teaching interests include wireless communications, broadband networks, information theory, and coding.

Dr. Naraghi-Pour has served as a Session Organizer, Session Chair, and member of the Technical Program Committee for many international conferences.