

Secure Distributed Detection in Wireless Sensor Networks via Encryption of Sensor Decisions

Thesis
Submitted to the Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

in

Department of Electrical and Computer Engineering

by
Venkata Sriram Siddhardh Nadendla
B.E. in Electronics and Communication Engineering
Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya (Deemed Univ.), India, 2007.
August 2009

To my parents

Acknowledgements

I am delighted to express my sincere gratitude to my major advisor, Dr. Morteza Naraghi-Pour for his exemplary support and guidance for my intellectual progress. He taught me how to approach a problem and inspired me how to be patient in dark times when progress is slow and overcome all the hurdles on the way towards my Masters degree. His role as a major professor was not just restricted to technical advice and has been mentoring me in developing social relations in academia.

I would like to thank my committee members, Dr. Xue-Bin Liang and Dr. Guoxiang Gu for their kind support. I also deeply appreciate Dr. Shaungqing Wei, Dr. Xue-Bin Liang, Dr. Robert Lipton (Department of Mathematics) and Dr. Hsiao-Chun Wu whom I am associated with in my classroom courses. I also thank Dr. Vaidyanathan and Dr. Richardson (Dept. of Mathematics) for their valuable suggestions in my course-plan. Furthermore, I thank the Dept. of Electrical and Computer Engineering for supporting me financially from my first day in LSU, making me concentrate on my research without any deviations.

My deepest gratitude goes to my parents for moulding me as who I am. They patted my back whenever I made a right decision and protected me from the consequences of my wrong decisions. I would also like to thank all my relatives and friends for giving me a wonderful experience during my stay here in LSU. They boosted me with all the energy I need to pursue research, esp. when I was dull and gloomy.

This certainly will prove to be a great starting point for me in my research...

Table of Contents

Acknowledgements	iii
List of Figures	vi
Abstract	viii
1 Introduction	1
1.1 Topologies	1
1.2 Distributed Detection using Sensor Networks	3
1.3 Motivation for this work	5
2 System Model	7
3 Problem Formulation	12
4 Optimal Ally Fusion Rule	13
4.1 Secure Detection in the Presence of Symmetric Noise	16
4.1.1 Gaussian Noise	17
4.1.2 Laplacian Noise	20
4.2 Minimizing the probability of error for AFC	21
4.2.1 Existence of minimum P_e	21
4.3 Numerical Algorithms for Optimal Threshold	24
4.3.1 Secant Method	24
4.3.2 Iterative Method	25
5 Simulation Results	26
5.1 Quasiconvexity of Error Probabilities	26
5.2 Convergence of Numerical Algorithms	30
5.3 Constrained Optimization	32
6 Conclusion and Future Work	34

References	36
Vita	39

List of Figures

1.1	Parallel fusion topology of the sensor network	2
1.2	Serial topology of the sensor network	2
1.3	Tree topology of the sensor network	3
1.4	Censoring sensor network	5
2.1	Sensor network model	8
2.2	Stochastic Encryption of Sensor Decisions	9
4.1	Quasi-convex function	13
4.2	Gaussian Signal Model	18
4.3	Plot of g as a function of τ	18
4.4	Laplacian Signal Model	20
4.5	r as a function of τ for Gaussian and Laplacian signal models for $d = 1$, $q_0 = 0.5$, $p_1 = 0.1$ and $p_2 = 0.1$	23
4.6	ψ as a function of λ	23
5.1	Comparison of performance of AFC and TPFC for $n = 10$, $p_1 = 0.1$, $p_2 = 0.1$ and $d = 1$ in the presence of Gaussian noise	26
5.2	Comparison of performance of AFC and TPFC for $n = 10$, $q_0 = 0.5$ and $d = 1$ in the presence of Gaussian noise	27
5.3	Comparison of performance of AFC and TPFC for $n = 20$, $p_1 = 0.1$, $p_2 = 0.1$ and $d = 1$ in the presence of Gaussian noise	28
5.4	Comparison of performance of AFC and TPFC for $n = 20$, $q_0 = 0.5$ and $d = 1$ in the presence of Gaussian noise	28

5.5	Comparison of performance of AFC and TPFC for $n = 10$, $p_1 = 0.1$, $p_2 = 0.1$ and $d = 1$ in the presence of Laplacian noise	29
5.6	Comparison of performance of AFC and TPFC for $n = 10$, $q_0 = 0.5$ and $d = 1$ in the presence of Laplacian noise	29
5.7	Comparison of convergence of the secant method with the proposed iterative method for $n = 10$, $d = 1$, $p_1 = 0.1$ and $p_2 = 0.1$ in the presence of Gaussian noise	30
5.8	Comparison of convergence of the secant method with the proposed iterative method for $n = 10$, $d = 1$ and $q_0 = 0.5$ in the presence of Gaussian noise	31
5.9	Comparison of convergence of the secant method with the proposed iterative method for $n = 10$, $d = 1$, $p_1 = 0.1$ and $p_2 = 0.1$ in the presence of Laplacian noise	31
5.10	Comparison of convergence of the secant method with the proposed iterative method for $n = 10$, $d = 1$ and $q_0 = 0.5$ in the presence of Laplacian noise	32
5.11	Constrained Optimization of AFC over TPFC in the presence of Gaussian noise for $d = 1$ and $q_0 = 0.5$	32
5.12	Constrained Optimization of AFC over TPFC in the presence of Laplacian noise for $n = 10$, $d = 1$ and $q_0 = 0.5$	33

Abstract

We consider the problem of binary hypothesis testing using a distributed wireless sensor network. Identical binary quantizers are used on the sensor's observations and the outputs are encrypted using a probabilistic cipher. The third party (enemy) fusion centers are unaware of the presence of the probabilistic encipher. We find the optimal (minimum-probability-of-error) fusion rule for the ally (friendly) fusion center subject to a lower bound on the the probability of error for the third-party fusion centers.

To obtain the minimum probability of error, we first prove the quasi-convexity of error probability with respect to the sensor's threshold for a given cipher and show the existence of a unique positive minimum for error probability of the ally fusion center. The threshold corresponding to the minimum error-probability is evaluated numerically and the appropriate cipher that deteriorates the performance of the third-party fusion center below the required limits is obtained.

Our results show that, by adjusting the sensor threshold and the encryption parameters, it is possible to achieve acceptable performance for the ally fusion center while causing significant degradation to the performance of the third party fusion center.

1 Introduction

Recent trends in VLSI and signal processing have led to the emergence of intelligent sensor networks that are capable of improving the sensing performance in multiple dimensions [14]. The motivation for these networks dates back to the implementation of these networks for military surveillance purposes across the borders. Now, this idea is extended to a wide range of domestic applications such as disaster-monitoring, health-monitoring, managing inventory, traffic-control and monitoring product-quality [11].

Distributed sensing came into picture with the emergence of applications where the location of the phenomenon of interest is not known. Sensors when distributed spatially, enhance the line-of-sight and thereby improve SNR, even with the presence of obstructions in the field [10].

Wireless sensor network is a network of sensor nodes that are spatially distributed to monitor an observation space for a physical parameter like temperature, pressure, or motion. These sensor nodes, sometimes also called "intelligent" sensors, consist of sensing, data processing and communicating components. Each network comprises of many such individual sensor nodes densely deployed whose position need not be engineered or predetermined. This allows us to have a random deployment of these sensors which is particularly motivating in regions or situations that are inaccessible. This also means that the design has to involve parameters that have self-organizing capabilities [1].

Intelligent sensors, though limited by their processing abilities and energy-constraints because of the limited battery power, can network themselves and communicate with a central agency (node) called the fusion center. The fusion center, having information from different geographic locations in the total coverage area, can hence give a reliable decision. But the constraints such as limited power supply, limited bandwidth and limited range of radio communication between the sensors and the fusion center makes the design of sensor networks quite challenging.

1.1 Topologies

Wireless sensor networks can be organized in different ways depending on the arrangement of sensors and the fusion center. There are three major topologies used for sensor networks namely parallel, serial and tree topologies [1, 28]. The first is the *parallel* topology, as shown in figure 1.1, wherein sensors do not communicate with each other. They collect

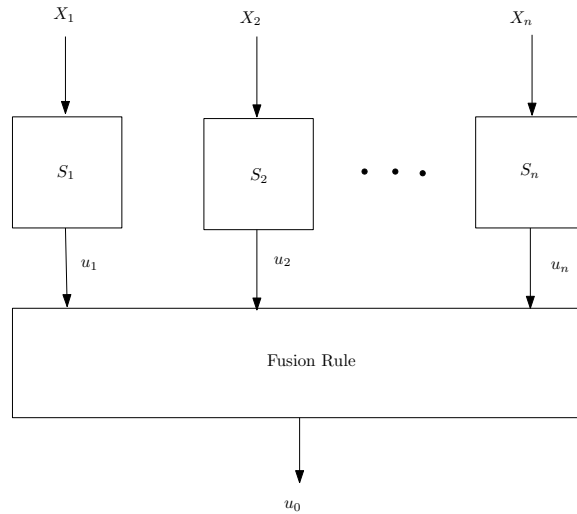


Figure 1.1: Parallel fusion topology of the sensor network

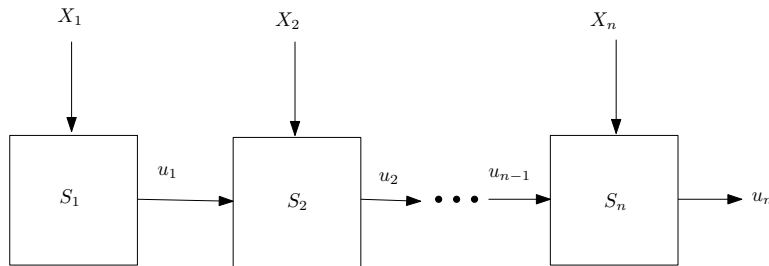


Figure 1.2: Serial topology of the sensor network

the information simultaneously, process it and transmit the partially processed data to the fusion center where final decision is made.

Serial or *Tandem* topology as in figure 1.2 is another topology where the sensors are connected in series and communicate with their immediate neighbors in a serial unidirectional fashion. The first sensor hence preprocesses the data which it received from the surroundings and sends this information to the second sensor. There onwards, the sensors make decisions based on both the sensed data from the observation space and also the data which it received from its predecessor. Notice that there is no need for a fusion center for the design of a serial topology.

In the *tree* topology depicted by figure 1.3, the sensors are arranged in the form of a tree. Sensors are arranged in different stages and the successor stage gets the data both from the sensors of the predecessor stage and the observation space. Sensor at the final stage gives the final decision. This is similar to the serial network in a way that it does not have

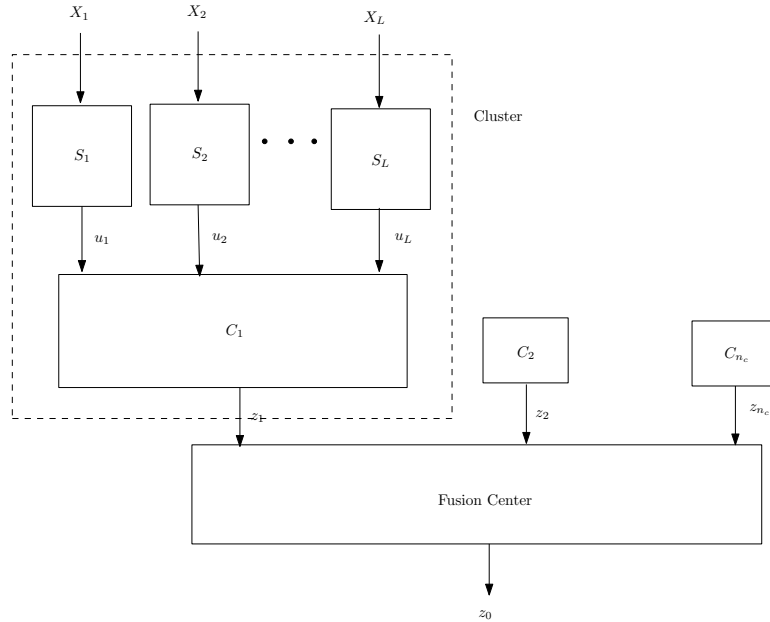


Figure 1.3: Tree topology of the sensor network

fusion center.

1.2 Distributed Detection using Sensor Networks

Potential applications of sensor networks include detection, estimation and tracking of a physical parameter such as temperature, pressure and location. Detection and estimation problems are static in nature as we only consider the present status of the phenomenon of interest. On contrary, tracking is a dynamic problem where we track the changes in the physical parameter in both time and space. Here, we are only interested in the detection problem.

In a conventional sensor network design, the sensors collect information from the environment and transmit it to the fusion center so that a statistical decision can be made. Since all the decision logic is located in one place in the network, this type of detection is known as *centralized detection*. But the raw digital data that is collected occupy a lot of bandwidth and hence we have a bandwidth constraint in the frequency spectrum allocated to the network. In order to eliminate this problem, raw data obtained by the sensing units, is partially processed in the sensor itself resulting in a signal that occupies low bandwidth for transmission across the wireless channel. Hence this type of detection, popularly termed

decentralized detection is preferred over the conventional designs [28].

Distributed detection in sensor networks is one of the research problems that has been extensively covered in the literature [5, 24–26]. Several environments have been considered in proposing an optimal design specific to the constraints posed by them. The following paragraphs give a brief outline of the past research and then fits our work in their context.

Tenney and Sandell, for the first time, extended the problem of classical Bayesian detection to the distributed sensors problem in [24]. The problem was formulated based on a standard hypothesis testing problem was considered and proposed optimal decision rules for the individual sensors. But since they never considered the design of an data fusion scheme, Chair and Varshney proposed an optimal data fusion scheme (k-out-of-n rule) where each sensor's decision is weighted based on the reliability of the local decision rule (binary quantizer) which is later compared to an optimal threshold for the final decision [5]. It is important to note that both these works assume the presence of a noiseless channel.

Later, Tsitsiklis, in his pioneering work [25], proved that the sensors, with i.i.d. observations, can be segregated optimally into $\frac{M(M-1)}{2}$ groups in a M -hypothesis distributed detection problem as the number of sensors tend to infinity and that each group has identical decision rules for all the sensors. Therefore, when a binary hypothesis testing problem is considered, likelihood-ratio quantizers were proposed as an asymptotically optimal decision rule identical to all the sensors.

More recently, focus has been shifted to designing a distributed sensor network in the presence of a noisy channel. Different channel environments had been considered such as the Rayleigh Fading Channel as in [13]. Indeed, the authors in [13] use the same likelihood ratio statistic proposed by Tsitsiklis, as mentioned earlier, in the fading channel environment, but assumes perfect knowledge of both the channel and the performance indices of the local sensor decision rules. Nui *et al.*, also proposed an alternative fusion scheme, namely, the equal-gain combiner which barely requires prior information regarding the sensor or the channel. Chen *et al.* consider the distributed detection problem for non-ideal channels (binary symmetric channel, in short, BSC) and proves that the likelihood ratio test for local sensor decisions are optimal in [9].

Recent trends in the development of wireless sensor network has been directed to the energy efficient schemes which involve intelligent designs like those involving censoring scheme as suggested by Rago *et al.*, [15] which allow only the so-called "informative" decisions to reach the fusion center. In their paper, Rago *et al.*, suggests that there exists a single interval in which conditionally independent sensor information is being censored. Also that, the paper suggests that different multiple intervals can be reduced to a model with single interval.

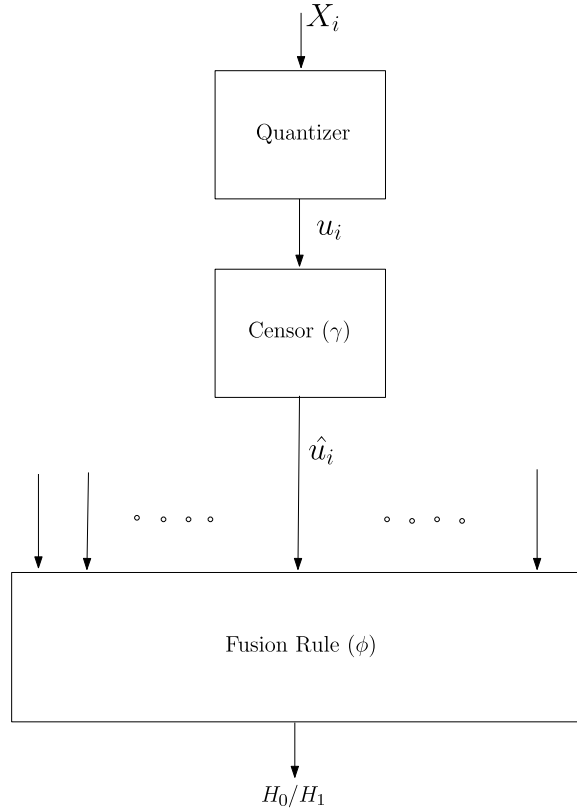


Figure 1.4: Censoring sensor network

After Rago *et al.*, has first presented the idea of 'censoring', there have been many authors proposing wireless sensor networks designs using censoring scheme. Designs with send/no-send transmission scenario has been proposed [2]. Authors like Appadawedula *et al.*, and Tay *et al.*, have considered the problem of optimization over a sequence of detectors especially the case of asymptotic performance [2, 3, 23]. Especially, Tay *et al.*, [23] considered the problem of designing a network of binary sensors where the sensors have access to side information that affects the statistics of the measurements.

1.3 Motivation for this work

In the case of distributed detection problem, the focus was mainly on energy-efficient designs due to the practical constraints. But security is a key issue which has been neglected all these days. All the above mentioned designs do not take into account the possibility of the presence of an eavesdropper (insecure channel) who might try to use the sensor

decisions according to his convenience and distort the channel between the sensors and the fusion center so that no effective decision can be made by our fusion center. Although many security protocols were developed for sensor networks, never was the security issue addressed in distributed detection/estimation problem until, Aysal *et al.*, in 2008, for the first time, proposed a system model with a cipher embedded in the local sensors' design in [4] for a distributed estimation problem. We therefore would consider extending this feature to the distributed detection problem for the same model.

Here we discuss about the performance of distributed detection in a parallel sensor network with an additional dimension of security embedded into its design. Probabilistic enciphers are introduced in the sensor end so that the performance of ally fusion center (AFC) is better than the unauthorized third-party fusion centers (TPFC) that try to seek the information transmitted by the sensors illegally. Hence it is quite reasonable to assume that the probability distribution of the stochastic parameter used in encryption is only known to the AFC which makes the difference in the implementation of the two fusion center designs.

Note that there is always a chance for the TPFC to trace back the cipher parameters and use them to find a design that has a performance similar to the optimal AFC design. In order to get this information, TPFC should have the prior knowledge of the observation model, the sensor decision rules and a large amount of data to statistically compute these probabilistic cipher parameters. The confidence in our model comes from the fact that, even if TPFC has this information, we still have the control of selecting our own sensor thresholds which changes the complete system parameters and thereby, ensuring security back in our design.

2 System Model

In this chapter, we will be describing the model of the sensor network. We assume a parallel-topology configuration of the sensors which transmit the data into the wireless channel. This data, which is nothing but the "intelligent" sensor decisions are available to the ally fusion center. We also assume the presence of an eavesdropper in the neighborhood of these sensors which collects the data, and makes its own decision using a third-party fusion center(TPFC). By making such a decision, the eavesdropper can twist the observations at the reception of the local sensors and thereby mislead our decision. Hence, we assume the presence of a third party fusion center(TPFC) whose performance is deteriorated with the proper design of individual sensors - of course, under a certain constraint on the degradation of ally fusion center's(AFC) performance. Let us start with a detailed description of the sensor network model mentioned earlier in this paragraph.

Consider a system of n sensors observing 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 i th sensor, $i = 1, 2, 3, \dots, n$. It is assumed that given the hypothesis H_η , ($\eta = 0, 1$), the observations X_1, X_2, \dots, X_n are independent and identically distributed. The conditional PDF of X_i under the hypothesis H_η is given by $p_X^\eta(x)$. Each sensor i , $i = 1, 2, \dots, n$, makes a decision $u_i \in \{0, 1\}$ regarding the state of the hypothesis H using the likelihood ratio threshold test

$$\frac{p_X^1(x)}{p_X^0(x)} \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \lambda \quad (2.1)$$

where λ is the identical threshold for all the sensors. In general, identical threshold assumption need not lead us to an optimal solution. However, the complexity of the problem would be prohibitive without such an assumption. Irving *et al.*, proved that the optimality is not lost when identical sensor thresholds are used in the case of a two-sensor system [12]. Furthermore, it is shown in [25] and [8] that this assumption of identical sensors would be optimal asymptotically in the number of sensors, n . Relying on these results and in order to make the problem tractable, we assume identical threshold λ for all the sensors.

The performance of this binary quantizer can be expressed using two quantities - false alarm probability P_F and the detection probability P_D of individual sensors, which are given by

$$\begin{aligned} P_F &= P(u_i = 1 | H_0) \\ P_D &= P(u_i = 1 | H_1) \end{aligned}$$

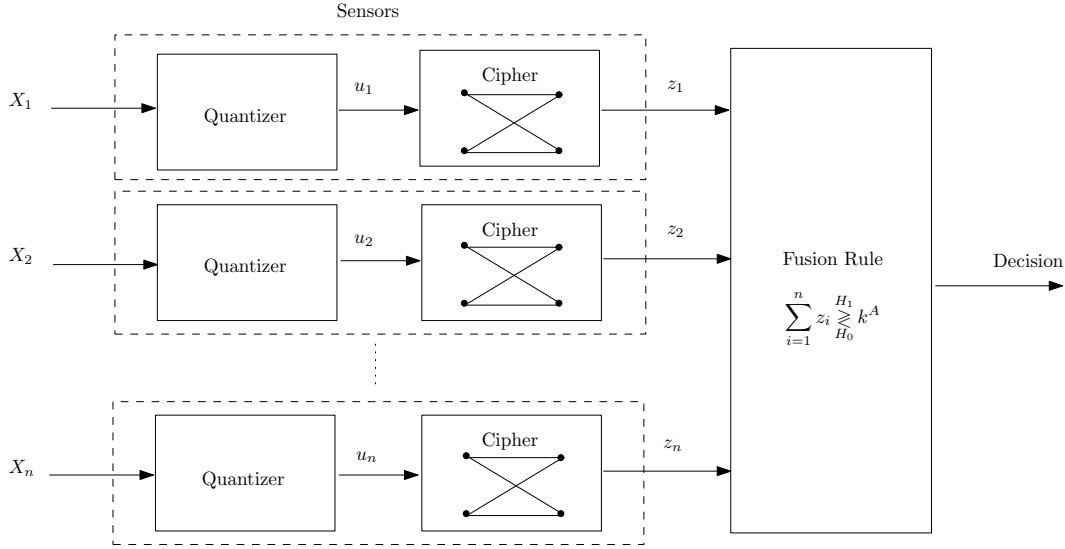


Figure 2.1: Sensor network model

Note that these decisions are transmitted to the fusion center through a wireless channel which can be accessed by many other unauthorized receivers and hence can use this data from the sensors according to their convenience. So there is a need for an encryption system that allows the sensor decisions to be accessible only to the fusion center.

A simple solution to this is to use a fixed cipher to the sensor decisions, but since TPFC has access to the data, there is a very good chance that it may identify the presence of this fixed cipher from the statistics of the data set it received and may change its parameters to get a better performance. So, an appropriate solution is the use of a stochastic cipher whose parametric distribution is known only to AFC. This eliminates the possibility for TPFC to find the existence of a cipher and hence would never have better performance than AFC.

The probabilistic encryption mechanism used in the sensors encrypts the decision u_i of sensor i to obtain z_i such that $P(z_i = 1|u_i = 0) = p_1$ and $P(z_i = 0|u_i = 1) = p_2$. The encrypted binary output z_i is then transmitted to the AFC and may also be observed by the TPFC.

An alternative description of this model can be given as $z_i = u_i \oplus v_i$, where $v_i \in \{0, 1\}$, $\{v_i\}_{i=1}^n$ are independent random variables with $P(v_i = 1|u_i = 0) = p_1$ and $P(v_i = 1|u_i = 1) = p_2$, and where \oplus is modulo-2 addition. It is assumed that the AFC has knowledge of the value of p_1, p_2 but not the actual values of v_1, v_2, \dots, v_n . On the other hand, the TPFC has no knowledge of the existence of cipher and its parameters p_1, p_2 and can only assume

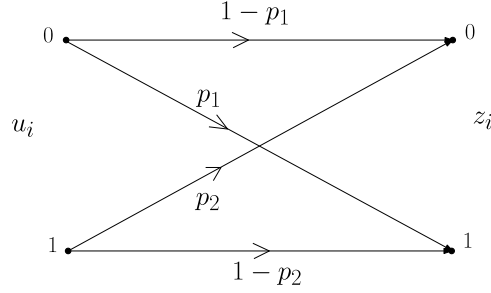


Figure 2.2: Stochastic Encryption of Sensor Decisions

that it received the original decisions u_i , $i = 1, 2, \dots, n$, which corresponds to $p_1 = p_2 = 0$. Thus, AFC takes advantage of this additional information to improve its performance over TPFC. But also note that introducing such a stochastic cipher would degrade the performance of the system as a whole and so the performance of a non-encrypted sensor network is always better than the encrypted design.

Both the fusion centers, AFC and TPFC, receive these encrypted bits z_i and then combine them to make a final decision on the hypotheses. Since both of them act greedily, trying to achieve the best performance possible, the optimum (minimum probability of error) fusion rule for both the AFC and TPFC fusion centers is a k -out-of- n rule is given by Equation 2.4. This can be proved by considering the maximum *a posteriori* probability (MAP) rule which is given by

$$\begin{aligned}
 P(H_1|\mathbf{z}) &\underset{H_0}{\overset{H_1}{\geq}} P(H_0|\mathbf{z}) \\
 \text{or } P(\mathbf{z}|H_1)q_1 &\underset{H_0}{\overset{H_1}{\geq}} P(\mathbf{z}|H_0)q_0 \\
 \text{or } \frac{P(\mathbf{z}|H_1)}{P(\mathbf{z}|H_0)} &\underset{H_0}{\overset{H_1}{\geq}} \frac{q_0}{q_1} (= \Lambda, \text{ in general.})
 \end{aligned}$$

Since the z_i 's are independent of each other, the MAP rule simplifies to

$$\prod_{i=1}^n \frac{P(z_i|H_1)}{P(z_i|H_0)} \underset{H_0}{\overset{H_1}{\geq}} \Lambda$$

Let θ_0 and θ_1 denote the conditional probabilities of $z_i = 0$ given H_0 and H_1 , respectively, i.e.,

$$\begin{aligned}
 \theta_0 &= P(z_i = 0|H_0) = 1 - p_1 - (1 - p_1 - p_2)P_F \\
 \theta_1 &= P(z_i = 0|H_1) = 1 - p_1 - (1 - p_1 - p_2)P_D
 \end{aligned} \tag{2.2}$$

and $\Lambda = \frac{q_0}{q_1}$ in the case of global minimum error-probability criterion, as mentioned earlier.

Hence the likelihood ratio test which is the optimal rule with respect to the probability of error is given by

$$\frac{\theta_1^{n-\#(ones)}(1-\theta_1)^{\#(ones)}}{\theta_0^{n-\#(ones)}(1-\theta_0)^{\#(ones)}} \underset{H_0}{\overset{H_1}{\gtrless}} \Lambda$$

Since $\#(ones) = \sum_{i=1}^n z_i = l$ (say), therefore we have

$$\frac{\theta_1^{n-l}(1-\theta_1)^l}{\theta_0^{n-l}(1-\theta_0)^l} \underset{H_0}{\overset{H_1}{\gtrless}} \Lambda$$

Applying logarithms, we have

$$(n-l) \ln \left(\frac{\theta_1}{\theta_0} \right) + l \ln \left(\frac{1-\theta_1}{1-\theta_0} \right) \underset{H_0}{\overset{H_1}{\gtrless}} \ln \Lambda$$

On simplification, we have the final optimal decision rule as follows.

$$\sum_{i=1}^n z_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\ln \Lambda - n \ln \left(\frac{\theta_1}{\theta_0} \right)}{\ln \left(\frac{(1-\theta_1)\theta_0}{(1-\theta_0)\theta_1} \right)}$$

Hence, the optimal k is given by

$$k^A = \frac{\ln \Lambda - n \ln \left(\frac{\theta_1}{\theta_0} \right)}{\ln \left(\frac{(1-\theta_1)\theta_0}{(1-\theta_0)\theta_1} \right)} \quad (2.3)$$

On the other hand, for the TPFC, the optimum value of k is given by $k^{TP} = k^A(\lambda, 0, 0)$, which the TPFC calculates as given in [29] because of the lack of knowledge about the presence of the stochastic cipher in each sensor.

Hence the fusion rule can be generalized as follows:

$$u_0 = \begin{cases} 1, & \text{if } \sum_{i=1}^n z_i \geq k \\ 0, & \text{if } \sum_{i=1}^n z_i < k \end{cases} \quad (2.4)$$

where $k = k^A(\lambda, p_1, p_2)$ in the case of AFC and $k = k^{TP}(\lambda)$ in the case of TPFC.

In order to find the optimal k , we need to analyze the performance and quality of the fusion centers. Also, note that TPFC does not have the control over the choice of λ, p_1, p_2 as it just uses whatever sensor decisions are released into the wireless channel. Hence we find the metrics that measure the quality of the fusion centers in general and later find the optimal parameters in favorable to AFC. The metrics discussed above are the false alarm probability Q_F and the detection probability Q_D of the fusion centers as a function of λ, p_1, p_2 and k , which are given by

$$Q_F = \sum_{i=k}^n \binom{n}{i} (1 - \theta_0)^i (\theta_0)^{n-i} \quad (2.5a)$$

$$Q_D = \sum_{i=k}^n \binom{n}{i} (1 - \theta_1)^i (\theta_1)^{n-i} \quad (2.5b)$$

Hence the probability of error for fusion centers is given by

$$P_E = q_0 Q_F + q_1 (1 - Q_D) \quad (2.6)$$

While a number of performance criteria may be considered, we are interested in the minimum probability of error and the Bayesian detection problem.

Remark: A remark is in order here. While the formulas for the false alarm and detection probabilities, and the probability of error are the same for the two fusion centers AFC and TPFC, the optimal parameter k is different. Consequently, the optimum performance for the two centers will be different. Subsequently we will use the superscripts AFC or TPFC, respectively, in order to distinguish these quantities for the two centers.

3 Problem Formulation

In the previous chapter, the system model and its performance metrics were introduced. Now, we reached the stage of formulating the problem of finding the optimal (λ, k, p_1, p_2) which, on one hand, minimizes the error probability for AFC and on other hand, would simultaneously deteriorate the performance of TPFC.

As mentioned previously, the Bayesian detection problem is considered. Note that the probability of error is a function of λ, k, p_1 and p_2 . Our goal is to minimize the probability of error for the AFC subject to a lower bound on the probability of error for the TPFC. Equivalently one may consider maximizing the probability of error for the TPFC subject to an upper bound for the AFC. Since the TPFC is assumed to be unaware of the values of p_1 and p_2 used in the sensors, it attempts to minimize the probability of error over λ and k assuming $p_1 = p_2 = 0$.

In the following we consider the former case. This problem can be formulated as a constrained optimization problem as follows.

Problem Statement.

$$\begin{aligned} \arg \min_{\lambda, k, p_1, p_2} & P_E^A(\lambda, k, p_1, p_2) \\ \text{such that} & \\ & 1. P_E^{TP}(\lambda, k, p_1, p_2) \geq \alpha \\ & 2. 1 \leq k \leq n, \\ & 3. 0 \leq p_1, p_2 \leq 1 \end{aligned}$$

Since the TPFC has no idea about the randomization of u_i 's, the optimal k^{TP} as identified by TPFC can be calculated as given in [29]. Also, P_F and P_D are both assumed to be first-order differentiable with respect to λ .

To minimize P_E^A , the optimal λ for each (k, p_1, p_2) is found and then the performance of TPFC (P_E^{TP}) is compared with that of AFC over different values of p_1, p_2 for each k .

4 Optimal Ally Fusion Rule

It is time to solve the problem that we formulated in Chapter 3. The motivation to solve this problem comes from [29] which proves that the error-probability has a unique minimum by proving the quasi-convexity property of P_e with respect to the identical threshold λ in the case of an unsecured sensor network model which is quite similar to our model in terms of the structure of the equations. In fact, we expect the function to be strictly convex so that it has a unique minimum. But since error probability is no longer convex, we try to check for a more relaxed property, i.e. the quasiconvexity property and therefore start investigating the error probability P_e if it is a quasi-convex function of λ , for a given k , p_1 and p_2 which also guarantees a unique minimum. This property of error probability is also corroborated by another work by Shi *et al.*, who proved the error probability as a quasi-convex function of the sensor threshold λ for Gaussian-like distributions [17]. But note that since we are only interested in the optimal design, Equation 2.3 is employed for the value of k which takes care of the problem of optimizing the problem over k , thus reducing the complex notation in the problem.

But before we start, let us know what quasi-convexity is.

Definition 1 (Quasi-convexity). *A function $f(\lambda)$ is quasi-convex if, for some λ^* , $f(\lambda)$ is non-increasing for $\lambda \leq \lambda^*$ and $f(\lambda)$ is non-decreasing for $\lambda \geq \lambda^*$ [29].*

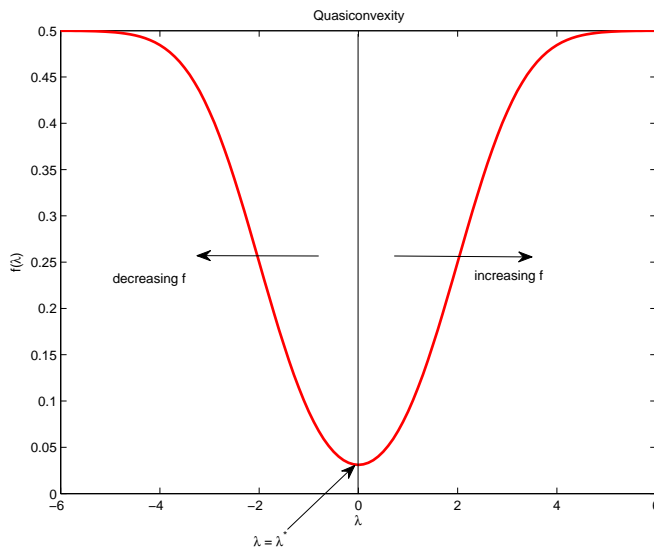


Figure 4.1: Quasi-convex function

In other words, if $\frac{dP_E^A}{d\lambda} \leq 0$ (or $\frac{dP_E^A}{d\lambda} \geq 0$) for all λ , or $\frac{dP_E^A}{d\lambda} \leq 0$ when $\lambda \leq \lambda^*$ and $\frac{dP_E^A}{d\lambda} \geq 0$ when $\lambda \geq \lambda^*$ for some λ^* .

Thus quasi-convexity of P_e guarantees the existence of an optimal solution $\lambda = \lambda^*$ to the problem of minimizing P_e for a fixed k , p_1 and p_2 . So, we start with Lemma 1 which gives the condition for the quasi-convexity of P_e to be satisfied for the system model considered in Chapter 2.

Lemma 1. *Assume that*

$$\frac{d}{d\lambda} \left(\frac{1}{\lambda} \frac{P_D}{P_F} \right) \leq 0 \quad (4.1)$$

Then for the optimal value of k (as given by Equation 2.3) and any fixed value of (p_1, p_2) , when $p_1 + p_2 \leq 1$, $P_E^A(\lambda, k^A(\lambda, p_1, p_2), p_1, p_2)$ is a quasi-convex function of λ .

Proof. In order to check for the quasi-convexity with respect to λ , for a fixed k , p_1 and p_2 , P_E^A (Equation 2.6) is first differentiated with respect to λ , and using Equations (2.5a, 2.5b), and $\frac{dP_D}{dP_F} = \lambda \left(= \frac{d\theta_1}{d\theta_0} \right)$, we have

$$\begin{aligned} \frac{dP_E^A}{d\lambda} &= q_0 \frac{dQ_F^{AFC}}{d\lambda} - q_1 \frac{dQ_D^{AFC}}{d\lambda} \\ &= q_1 \lambda (\theta_0)' n \binom{n-1}{k-1} (1-\theta_1)^{k-1} (\theta_1)^{n-k} \\ &\quad - q_0 (\theta_0)' n \binom{n-1}{k-1} (1-\theta_0)^{k-1} (\theta_0)^{n-k} \end{aligned}$$

where $(\theta_0)' = \frac{d\theta_0}{d\lambda} = -(1-p_1-p_2) \frac{dP_F}{d\lambda} \geq 0$ if $p_1 + p_2 \leq 1$ and $\frac{dP_F}{d\lambda} \leq 0$. [4, 29]

Rewriting the above equation, we have

$$\frac{dP_E^A}{d\lambda} = g(\lambda, k, p_1, p_2) (e^{r(\lambda, k, p_1, p_2)} - 1) \quad (4.2)$$

where

$$g(\lambda, k, p_1, p_2) = n \binom{n-1}{k-1} q_0 (1-\theta_0)^{k-1} (\theta_0)^{n-k} (\theta_0)' \quad (4.3a)$$

and

$$r(\lambda, k, p_1, p_2) = \ln \left(\frac{q_1}{q_0} \right) + \ln \lambda + (k-1) \ln \left(\frac{1-\theta_1}{1-\theta_0} \right) + (n-k) \ln \left(\frac{\theta_1}{\theta_0} \right) \quad (4.3b)$$

We have $g(\lambda, k, p_1, p_2) \geq 0$ which means that the sign of $\frac{dP_E^A}{d\lambda}$ depends on the value of $r(\lambda, k, p_1, p_2)$. In order to complete the proof, $r(\lambda, k, p_1, p_2)$ must be either always positive or negative, or there exists λ^* such that $r(\lambda, k, p_1, p_2) \leq 0$ for all $\lambda \leq \lambda^*$ and $r(\lambda, k, p_1, p_2) \geq 0$ for all $\lambda \geq \lambda^*$. Note that $r(\lambda, k, p_1, p_2)$ being either positive or negative would result in an optimal λ that is either zero or ∞ , which is a trivial solution. We would rather want $r(\lambda, k, p_1, p_2)$ which gives a unique solution to the optimal λ that is positive and finite. So we check if $r(\lambda, k, p_1, p_2)$ is either increasing or decreasing which would guarantee the existence of optimal λ satisfying the equations.

Unfortunately, we were not able to proceed beyond this point without any loss of generality. So, since we are interested in the optimal value of k^A as given by Equation 2.3, we substitute Equation 2.3 in Equation 4.3b and solve the problem only for this special case as follows.

$$r(\lambda, p_1, p_2) = \ln \Lambda + \ln \frac{q_1}{q_0} + \ln \lambda - \ln \left(\frac{1 - \theta_1}{1 - \theta_0} \right) \quad (4.4)$$

and by differentiating $r(\lambda, p_1, p_2)$ with respect to λ , we get

$$\frac{dr(\lambda, p_1, p_2)}{d\lambda} = \frac{1}{\lambda} - \frac{1}{1 - \theta_1} \left[\frac{1 - \theta_1}{1 - \theta_0} - \lambda \right] \frac{d\theta_0}{d\lambda} \quad (4.5)$$

From [29], we get a motivation to check if $r(\lambda, p_1, p_2)$ is increasing, i.e.

$$\frac{dr(\lambda, p_1, p_2)}{d\lambda} \geq 0 \quad (4.6)$$

In other words, we check if

$$\begin{aligned} \frac{1}{\lambda} - \frac{1}{1 - \theta_1} \left[\frac{1 - \theta_1}{1 - \theta_0} - \lambda \right] \frac{d\theta_0}{d\lambda} &\geq 0 \\ \text{or} \\ \frac{1}{\lambda} + \frac{\lambda \frac{d\theta_0}{d\lambda}}{1 - \theta_1} &\geq \frac{1}{1 - \theta_0} \frac{d\theta_0}{d\lambda} \end{aligned}$$

Multiplying $\lambda(1 - \theta_0)$ on both sides, we have

$$(1 - \theta_0) + \left(\frac{1 - \theta_1}{1 - \theta_0} \right) \lambda^2 \frac{d\theta_0}{d\lambda} \geq \lambda \frac{d\theta_0}{d\lambda}$$

Expanding the individual terms θ_0 and θ_1 with Equations 2.2 given in Chapter 2 and dividing with the positive term $(1 - p_1 - p_2)$, we have

$$\frac{p_1}{1 - p_1 - p_2} + P_F + \frac{P_F + \frac{p_1}{1 - p_1 - p_2}}{P_D + \frac{p_1}{1 - p_1 - p_2}} \left(-\lambda^2 \frac{dP_F}{d\lambda} \right) \geq -\lambda \frac{dP_F}{d\lambda}$$

Since $\frac{P_F}{P_D} \leq 1$, we know $\frac{P_F}{P_D} \leq \frac{P_F + \frac{p_1}{1-p_1-p_2}}{P_D + \frac{p_1}{1-p_1-p_2}}$. Therefore, if the inequality given by (4.7) is true, it follows that $\frac{dr(\lambda, p_1, p_2)}{d\lambda} \geq 0$.

$$P_F + \lambda \left(-\lambda \frac{dP_F}{d\lambda} \right) \frac{P_F}{P_D} \geq -\lambda \frac{dP_F}{d\lambda} \quad (4.7)$$

which is equivalent to the condition given by Equation 4.1. \square

Various noise distributions are considered that satisfy the above criterion for a particular model. We start with symmetric noise distributions, obtain a special criterion due to symmetry and then check if this is satisfied for Gaussian and Laplacian distributions. Later we obtain a generalized condition for the model used and search for the distributions that satisfy this condition.

4.1 Secure Detection in the Presence of Symmetric Noise

Now, we have a condition for P_E^A to satisfy the quasi-convexity criterion. But this condition need not be true in general for any noise distribution. Therefore, we start with symmetric noise model in general, obtain a special criterion due to symmetry and then prove the results for both Gaussian and Laplacian noise models. But the condition given by Equation 4.1 can be true even for some special non-symmetric distributions, which is not in the scope of this thesis.

Let us consider the following model for the received signal.

$$X_i = S + N_i \quad (4.8)$$

where $s = d$ under hypothesis H_1 , $S = -d$ under hypothesis H_0 and $N_i \sim p_N(x)$. The local log-likelihood decision rule for the i^{th} sensor is

$$T_i = \ln \frac{p_N(X_i - d)}{p_N(X_i + d)} \underset{H_0}{\overset{H_1}{\gtrless}} \tau \quad (4.9)$$

Let $T_i \sim p_0(t)$ under H_0 and $T_i \sim p_1(t)$ under H_1 , which means that $P_D = \int_{\tau}^{\infty} p_1(t) dt = f_1(t)$ and $P_F = \int_{\tau}^{\infty} p_0(t) dt = f_0(t)$. Hence, Equation 4.1 can be rewritten as

$$\frac{d}{d\lambda} \left(\frac{1}{\lambda} \frac{P_D}{P_F} \right) = e^{-\tau} \frac{d}{d\tau} \left(\frac{1}{e^{\tau}} \frac{f_1(\tau)}{f_0(\tau)} \right) \leq 0 \quad (4.10)$$

Note that since $e^{-\tau} > 0$, it is sufficient if we can show $\frac{d}{d\tau} \left(\frac{1}{e^\tau} \frac{f_1(\tau)}{f_0(\tau)} \right) \leq 0$. Expanding this, we have

$$e^\tau f_0(\tau) \frac{df_1(\tau)}{d\tau} - f_1(\tau) \left[e^\tau f_0(\tau) + e^\tau \frac{df_0(\tau)}{d\tau} \right] \leq 0$$

Since $\frac{df_1(\tau)}{d\tau} = -p_1(\tau)$ and $\frac{df_0(\tau)}{d\tau} = -p_0(\tau)$, and since e^τ is a non-negative quantity, after minimal rearrangements, we have

$$\frac{p_0(\tau)}{f_0(\tau)} - \frac{p_1(\tau)}{f_1(\tau)} \leq 1 \tag{4.11}$$

At this point, we would like to introduce symmetry in the noise distribution as it can eliminate one of the conditional distributions in the above Equation 4.11, making it easy to solve. It is well known that under antipodal signalling, as is the model considered, with symmetric noise, we have [21]

$$p_0(t) = p_1(-t)$$

Therefore, we can rewrite Equation 4.11 as

$$\frac{p_1(-\tau)}{1 - f_1(-\tau)} - \frac{p_1(\tau)}{f_1(\tau)} \leq 1 \tag{4.12}$$

This is particularly useful if we do not have a closed form expressions for $f_1(\tau)$ as in the case of Gaussian distribution. So, we would first start with the Gaussian noise distribution and later, would see the Laplacian noise case where we have closed form expressions for both $p_1(\tau)$ and $f_1(\tau)$.

4.1.1 Gaussian Noise

The first and the foremost noise distribution that comes to anyone's mind is the Gaussian distribution and hence, we would like to continue with the same notion of following the convention. It has some key features like symmetry and strong theory (like Central Limit Theorem) supporting its practical significance which makes it an attractive option.

Coming back to the problem, the following lemma proves that the Gaussian noise distribution in the presence of a stochastic cipher considered in the chapter earlier, satisfies the quasi-convexity property as given in Lemma 1. Since $T_i = 2dX_i \sim \mathcal{N}(2d, 4d^2)$ under hypothesis H_1 , $p_1(-\tau) = p_1(\tau + 4d^2)$ and $1 - f_1(-\tau) = f_1(\tau + 4d^2)$, an equivalent expression for Equation 4.12 can be given by Equation 4.15.

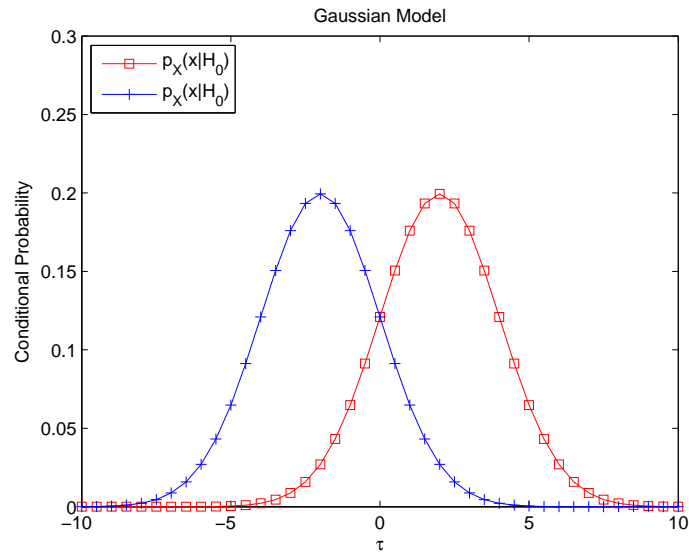


Figure 4.2: Gaussian Signal Model

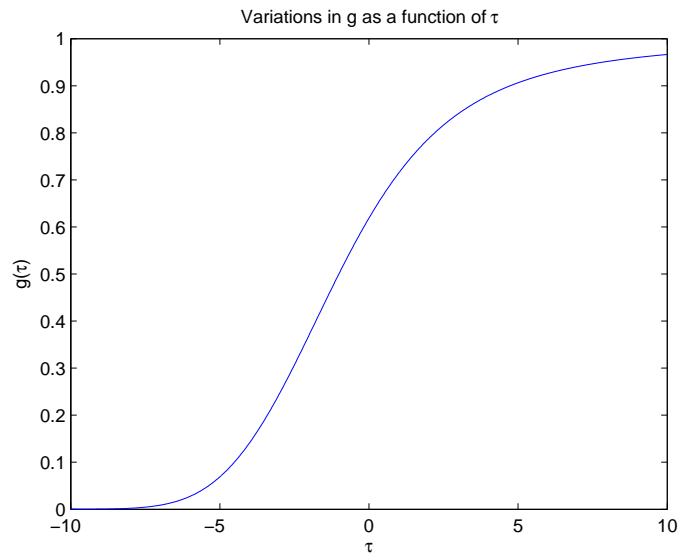


Figure 4.3: Plot of g as a function of τ

Lemma 2. *Suppose $N_i \sim \mathcal{N}(0, 1)$, i.e. N_i is a zero-mean Gaussian random variable with unit variance. Then the condition given by 4.12 is satisfied.*

Proof. Note that in this case, $T_i = 2dX_i$ and therefore,

$$p_1(t) = \frac{1}{2d\sqrt{2\pi}} e^{-\frac{(t-2d^2)^2}{8d^2}} \quad (4.13)$$

and

$$f_1(t) = Q\left(\frac{t-2d^2}{2d}\right) \quad (4.14)$$

Furthermore, $p_0(t) = p_1(t + 4d^2)$. Therefore, an equivalent expression for condition 4.12 would be

$$g(\tau) = \frac{p_1(\tau + 4d^2)}{f_1(\tau + 4d^2)} - \frac{p_1(\tau)}{f_1(\tau)} \leq 1 \quad (4.15)$$

Let $h(\tau, d) = \frac{p_1(\tau)}{f_1(\tau)}$. Hence we need to show that

$$h(\tau + 4d^2, d) - h(\tau, d) \leq 1$$

Observe that

$$\begin{aligned} \lim_{\tau \rightarrow -\infty} h(\tau + 4d^2, d) - h(\tau, d) &= 0 \\ \lim_{\tau \rightarrow \infty} h(\tau + 4d^2, d) - h(\tau, d) &= 1 \end{aligned}$$

Note that $\lim_{\tau \rightarrow \infty} \frac{p_1(\tau)}{f_1(\tau)} = \lim_{\tau \rightarrow \infty} \frac{\tau - 2d^2}{4d^2}$, implying that $g(\tau)$ tends to be linear with slope $\frac{1}{4d^2}$ as τ increases indefinitely and with zero slope as τ shoots to $-\infty$. In light of the above and the mean-value theorem, to prove 4.15, it is sufficient to show the following Lipschitz condition on $h(\tau, d)$:

$$\frac{dh(\tau)}{d\tau} \leq \frac{1}{4d^2} \quad \forall \tau, d. \quad (4.16)$$

Evaluating the derivative of $h(\tau, d)$, 4.16 gets reduced to

$$f_1^2(\tau) - 4d^2 f_1(\tau) p_1'(\tau) \geq 4d^2 p_1^2(\tau)$$

Expanding the terms using Equations 4.13 and 4.14 and substituting x for $\frac{\tau-2d^2}{2d}$, we have

$$\frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} Q(x) + Q^2(x) \geq \frac{1}{2\pi} e^{-\frac{x^2}{2}} \quad (4.17)$$

A strict lower bound on $Q(x)$ was proposed by [18] which is given below

$$Q(x) \geq \frac{2}{x + \sqrt{x^2 + 4}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (4.18)$$

Using 4.18 in the LHS of Equation 4.17, we have

$$\frac{x}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}Q(x) + Q^2(x) \geq \frac{x}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\frac{2}{x + \sqrt{x^2 + 4}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} + \left[\frac{2}{x + \sqrt{x^2 + 4}}\right]^2\frac{1}{2\pi}e^{-x^2}$$

After simplifying the RHS of the above condition, we find that it is equal to $\frac{1}{2\pi}e^{-\frac{x^2}{2}}$ which proves the lemma. \square

Hence, we can conclude that P_E^A is a quasi-convex function of λ in the presence of Gaussian noise.

4.1.2 Laplacian Noise

The next symmetric noise model we would like to consider is the additive Laplacian noise model which has closed form expressions for P_D and P_F making it easy to solve the problem from Equation 4.1 directly. Now let us start proving the quasi-convexity property of P_E^A in the presence of additive Laplacian noise, i.e. $N_i \sim \mathcal{L}(0, 1) = \frac{1}{2}e^{-|t|}$. First, P_D and P_F can

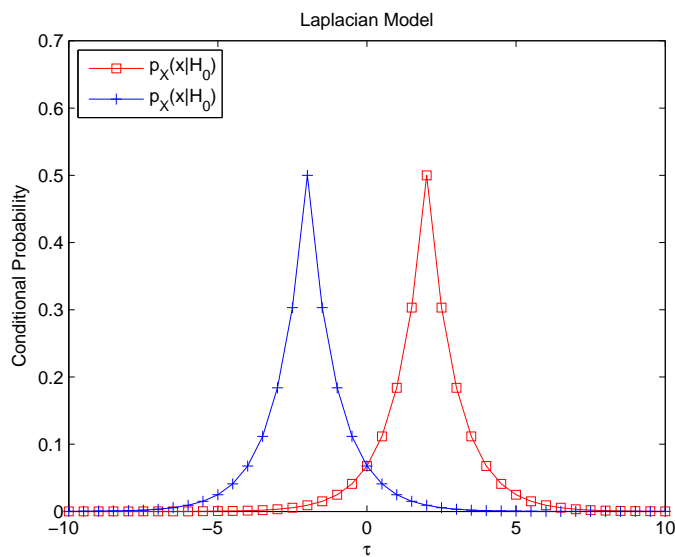


Figure 4.4: Laplacian Signal Model

be expressed as

$$P_D = \begin{cases} \frac{1}{2}e^{-(\tau-d)} & \text{if } \tau \geq d \\ 1 - \frac{1}{2}e^{(\tau-d)} & \text{if } \tau < d \end{cases} \quad (4.19a)$$

$$P_F = \begin{cases} \frac{1}{2}e^{-(\tau+d)} & \text{if } \tau \geq -d \\ 1 - \frac{1}{2}e^{(\tau+d)} & \text{if } \tau < -d \end{cases} \quad (4.19b)$$

These expressions, given by Equations 4.19a and 4.19b, are now used to check if Equation 4.1 is satisfied in Lemma 3 as follows.

Lemma 3. *If $p_N(t) = \frac{1}{2}e^{-|t|}$, then $\frac{d}{d\lambda} \left(\frac{1}{\lambda} \frac{P_D}{P_F} \right) \leq 0$.*

Proof. We evaluate $\frac{1}{\lambda} \frac{P_D}{P_F}$ as a piece-wise function for different values of τ using the Equations (4.19a), (4.19b) and $\lambda = e^\tau$.

CASE-1 ($\tau \geq d$): In this case, $\frac{1}{\lambda} \frac{P_D}{P_F} = e^{-\tau} e^{2d}$ and hence, $\frac{d}{d\lambda} \left(\frac{1}{\lambda} \frac{P_D}{P_F} \right) = -e^{-\tau} e^{2d} \leq 0$.

CASE-2 ($-d \leq \tau < d$): Here, $\frac{1}{\lambda} \frac{P_D}{P_F} = (2 - e^{\tau-d})e^d$. Therefore, $\frac{d}{d\lambda} \left(\frac{1}{\lambda} \frac{P_D}{P_F} \right) = -e^\tau \leq 0$.

CASE-3 ($\tau < -d$): Finally, we have $\frac{1}{\lambda} \frac{P_D}{P_F} = \frac{2e^{-\tau} - e^{-d}}{2 - e^\tau e^d}$ and hence, $\frac{d}{d\lambda} \left(\frac{1}{\lambda} \frac{P_D}{P_F} \right) = -\frac{4(e^{-\tau} - e^d) + e^\tau}{(2 - e^\tau e^d)^2}$. Since $\tau < -d$, the numerator is non-negative and the lemma is proved. \square

Thus, both additive Gaussian and Laplacian noise models support the presence of stochastic ciphers, allowing P_E^A to be a quasi-convex function of λ for optimal value of k . Of course, there are many other distributions waiting to be investigated in this direction, but this thesis only gives a path to follow in the case of distributions with or without closed-form expressions.

4.2 Minimizing the probability of error for AFC

4.2.1 Existence of minimum P_e

Quasi-convexity of $P_E^A(\lambda, p_1, p_2)$ with respect to λ does not guarantee the existence of optimal λ . From 4.2, it is seen that if $r(\lambda^*, p_1, p_2) = 0$ for some λ^* , then $\left. \frac{dP_e^A}{d\lambda} \right|_{\lambda=\lambda^*} = 0$.

If we go back to Chapter 2 in which the system model is described, λ is the threshold for likelihood ratio rule. Likelihood ratios are always non-negative as they are defined as the ratio of two probability distributions. Comparing this ratio to a negative number is trivial as it forces the decision to be always H_0 irrespective of what the observation is. So, when we try to find a non-trivial solution to $r(\lambda^*, p_1, p_2) = 0$ which is positive, we have a reasonable solution. Hence we investigate the conditions under which there exists a root for the equation $r(\lambda, p_1, p_2) = 0$.

Expanding $r(\lambda, p_1, p_2) = 0$, we have

$$\begin{aligned} r(\lambda, p_1, p_2) &= \ln \Lambda + \ln \frac{q_1}{q_0} + \ln \lambda - \ln \left(\frac{1-\theta_1}{1-\theta_0} \right) \\ &= \ln \left(\Lambda \frac{q_1}{q_0} \lambda \frac{1-\theta_0}{1-\theta_1} \right) = 0 \end{aligned}$$

In other words,

$$\Lambda \frac{q_1}{q_0} \lambda \frac{1-\theta_0}{1-\theta_1} = 1$$

or,

$$\lambda = \frac{q_0}{\Lambda q_1} \frac{1-\theta_1}{1-\theta_0} \quad (4.20)$$

Let $\psi(\lambda) = \frac{q_0}{\Lambda q_1} \frac{1-\theta_1}{1-\theta_0}$. Let us consider some properties of $\psi(\lambda)$. It can be verified that $\psi(0) = \psi(\infty) = \frac{q_0}{q_1 \Lambda} > 0$. Now since $\psi(\lambda)$ is continuous, it must intersect the line $y = \lambda$ at some point $\lambda^* > 0$. Therefore $r(\lambda, p_1, p_2) = 0$ has at least one positive solution. Uniqueness follows from the monotonicity of $r(\lambda, p_1, p_2)$.

Consider the ROC curve of individual sensors. Let $\lambda_{(0,0)}$ and $\lambda_{(1,1)}$ denote the slopes of this ROC curve at the points $(0,0)$ and $(1,1)$, respectively.

Theorem 1. *Given (p_1, p_2) such that $0 < p_1, p_2 < 1$ and $p_1 + p_2 < 1$, if $\lambda_{(0,0)} = \infty$, $\lambda_{(1,1)} = 0$, and $q_0, q_1 > 0$, then $r(\lambda, p_1, p_2) = 0$ has a unique positive root.*

Proof. To make sure there exists a positive root for $r(\lambda, p_1, p_2) = 0$, since $r(\lambda, p_1, p_2)$ is a monotonically increasing function, we expect linearity at both ends, i.e. at $\lambda = \pm\infty$. In order to maintain this linearity, the following condition is assumed which is true for a set of sensors that have type-1 ROC curves [22, 29]. If the slope of the ROC ($= \lambda$) at the corners, i.e., at $(0,0)$ and $(1,1)$, are $\lambda_{(0,0)} = \infty$ and $\lambda_{(1,1)} = 0$, then for a given (p_1, p_2) , $r(\lambda, p_1, p_2)$ has a positive root for λ .

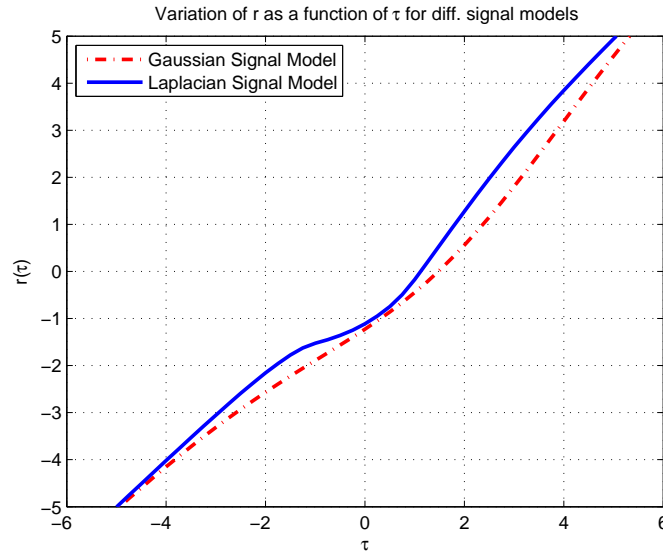


Figure 4.5: r as a function of τ for Gaussian and Laplacian signal models for $d = 1$, $q_0 = 0.5$, $p_1 = 0.1$ and $p_2 = 0.1$

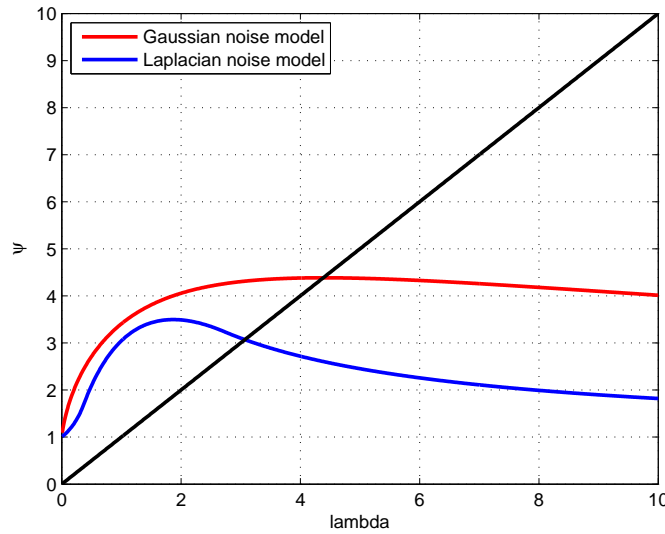


Figure 4.6: ψ as a function of λ

As described above, from the properties of ROC, we know that

$$\lim_{\lambda \rightarrow 0} \frac{\theta_1}{\theta_0} = 1 \quad \lim_{\lambda \rightarrow \infty} \frac{1 - \theta_1}{1 - \theta_0} = 1,$$

$$\lim_{\lambda \rightarrow \infty} \frac{\theta_1}{\theta_0} = 1 \quad \lim_{\lambda \rightarrow 0} \frac{1 - \theta_1}{1 - \theta_0} = 1$$

Defining $\tau = \ln \lambda$ allows us to restrict the domain of r to non-negative real numbers, which guarantees a positive $\lambda = \lambda^*$. Then it follows that

$$\lim_{\tau \rightarrow -\infty} \frac{r(\lambda, p_1, p_2)}{\tau} = 1 \quad \text{and} \quad \lim_{\tau \rightarrow +\infty} \frac{r(\lambda, p_1, p_2)}{\tau} = 1$$

Therefore, $r(\lambda, p_1, p_2) = 0$ is a linear function of τ at $\pm\infty$ and in general, an increasing function of τ . In other words, there is a unique positive root for $r(\lambda, p_1, p_2) = 0$ which assures the optimal threshold λ for the sensors. \square

Corollary 2. *For a given (p_1, p_2) such that $0 < p_1, p_2 < 1$ and $p_1 + p_2 < 1$, there exists a $\lambda = \lambda^*$ such that $P_E^A(\lambda, p_1, p_2)$ is minimized and λ^* satisfies*

$$r(\lambda, p_1, p_2) = \ln \Lambda + \ln \frac{q_1}{q_0} + \ln \lambda - \ln \left(\frac{1 - \theta_1}{1 - \theta_0} \right)$$

Hence, there exists a positive $\lambda = \lambda^*$ such that the probability of error is minimized.

4.3 Numerical Algorithms for Optimal Threshold

In this section, we would like to go a step further and find the optimal λ , p_1 and p_2 that minimizes P_e . First, we find the optimal $\lambda = \lambda^*$ which minimizes P_e^A for a given (p_1, p_2) . Then, we try to find (p_1, p_2) that minimizes P_e under the constraints $P_e^{TP} \geq \alpha$. We start with the description of these numerical algorithms as follows.

Many iterative numerical algorithms can be used to find the solution of $r(\lambda, p_1, p_2) = 0$. We would like to show two such algorithms - one being used earlier in literature [29] for solving a similar problem and an other one which we proposed.

4.3.1 Secant Method

Let us first start with the SECANT method to numerically find the optimal thresholds for the sensors. The following algorithm is used to find the optimal threshold for AFC and also the threshold for TPFC assuming that the TPFC is designed without the knowledge of the stochastic cipher used in AFC model.

- 1: Choose $\epsilon > 0$. Arbitrarily choose τ_1, τ_2 . Let $r_1 = r(\tau_1, p_1, p_2)$, $r_2 = r(\tau_2, p_1, p_2)$ and set $i = 3$.

2: Let

$$\tau_i = \frac{r_{i-1}\tau_{i-2} - r_{i-2}\tau_{i-1}}{\tau_{i-1} - \tau_{i-2}}.$$

3: Let $r_i = r(\tau_i, p_1, p_2)$.

4: If $|r_i| \leq \epsilon$, stop; otherwise, let $i = i + 1$, and go to step 2.

At the end of the above computation process, the optimum $\lambda = e^\tau$ is found for the given (p_1, p_2) .

4.3.2 Iterative Method

This is a more direct method that comes from the proof of Lemma 1. The quasi-convexity property of P_e comes from the result that there exists a unique root for the equation $r(\lambda, p_1, p_2) = 0$. Hence we start from this point to continue further from Equation 4.20 and find the optimal threshold.

As mentioned in Chapter 4, an equivalent expression to $r(\lambda, p_1, p_2) = 0$ is given as

$$\lambda = \psi\lambda \tag{4.21}$$

The iterative algorithm we proposed is based on Equation 4.21 as follows for a given n , d , p_1 and p_2 .

1: Choose $\epsilon > 0$. Arbitrarily choose λ_1 . Let $\psi_1 = \psi(\lambda_1)$ and set $i = 2$.

2: Let

$$\lambda_i = \psi(\lambda_{i-1}).$$

3: Let $\psi_i = \psi(\lambda_i)$.

4: If $|\psi_i - \lambda_i| \leq \epsilon$, stop; otherwise, let $i = i + 1$, and go to step 2.

In the following section, the performance of AFC is compared with that of the TPFC which uses the optimal k^{TP} from [29] over the range of values of p_1 and p_2 .

5 Simulation Results

In this chapter, we will be skimming through the details presented in Chapters 2, 3, and 4 along with the numerical results for the signal model considered in the presence of either Gaussian or Laplacian noise. We started with the system model where the construction of the sensors is described and then proved the quasiconvexity of P_e for the ally fusion center. Later two different numerical algorithms were proposed to find the identical optimal threshold used in the sensors as a function of p_1 and p_2 . Finally, under the constraints placed by TPFC's performance, we find the best cipher that minimizes the P_e^A .

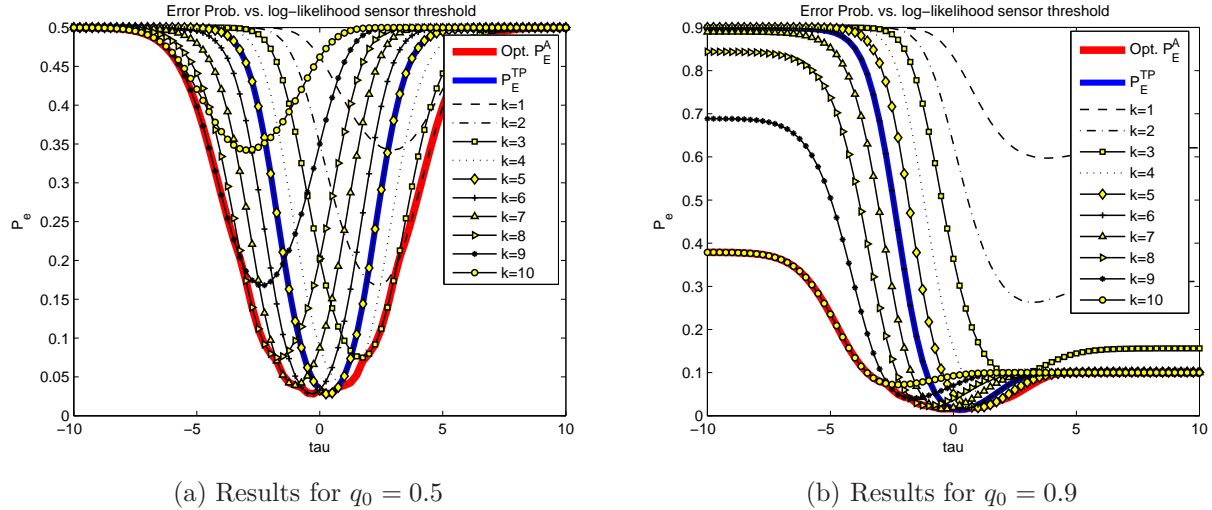


Figure 5.1: Comparison of performance of AFC and TPFC for $n = 10$, $p_1 = 0.1$, $p_2 = 0.1$ and $d = 1$ in the presence of Gaussian noise

Remark: Note that k^A is a continuous function of λ (Equation 2.3). In reality, k^A should be an integer since it is compared to the number of sensors that decide H_1 . So, we assume $k^A = \left\lceil \frac{\ln \Lambda - n \ln \frac{\theta_1}{\theta_0}}{\ln \frac{\theta_0(1-\theta_1)}{\theta_1(1-\theta_0)}} \right\rceil$ in the computations of our results.

5.1 Quasiconvexity of Error Probabilities

We start with the quasiconvexity of P_e^A . Figure 5.1 depicts the performance of P_e^A and P_e^{TP} for 10 sensors with a symmetric cipher using $p_1 = p_2 = 0.1$. These results are produced for

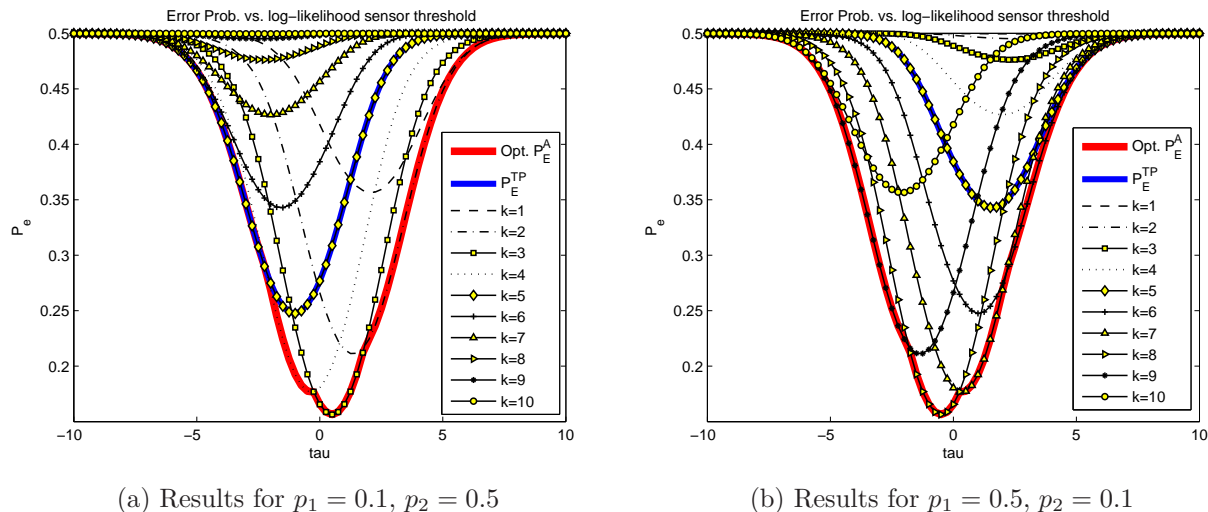


Figure 5.2: Comparison of performance of AFC and TPFC for $n = 10$, $q_0 = 0.5$ and $d = 1$ in the presence of Gaussian noise

the Gaussian noise model with $N_i \sim \mathcal{N}(0, 1)$ and $d = 1$. Multiple graphs (black in color), each one representing P_e for different values of k , are plotted in the same figure because of which, we are able to clearly understand that the optimal P_e^A curve is the lower envelope of all the curves. While the P_e^{TP} curve overlaps with one of the black curves because k^{TP} is found from [29] and is a fixed number which TPFC thinks is optimal for the given environment scenario. Also, one can clearly observe that there is an improvement from $q_0 = 0.5$ case (worst-case scenario) as given in Figure 5.1a to a more practical situation where $q_0 = 0.9$ which is depicted by Figure 5.1b.

Furthermore, we can also observe that the optimal P_e is the same for both AFC and TPFC. Since we want to improve the performance of AFC as we simultaneously deteriorate the TPFC's performance, a skew in the values of p_1 and p_2 is introduced to see if there is an improvement in the performance which is clearly depicted by figure 5.2. Figures 5.2a and 5.2b both refer to ciphers with skewed parameters that are mirror-images to each other which is directly reflected in the plots.

Also, the same set of plots are found for $n = 20$ (Figures 5.3 and 5.4) and we can clearly find that there is a significant increase in the performance of error probabilities of the fusion centers. This is a phenomenon which is expected in a sensor network as the resolution of the observation increases with increase in n .

Similar results are presented in the case of Laplacian noise model in figures 5.5 and 5.6

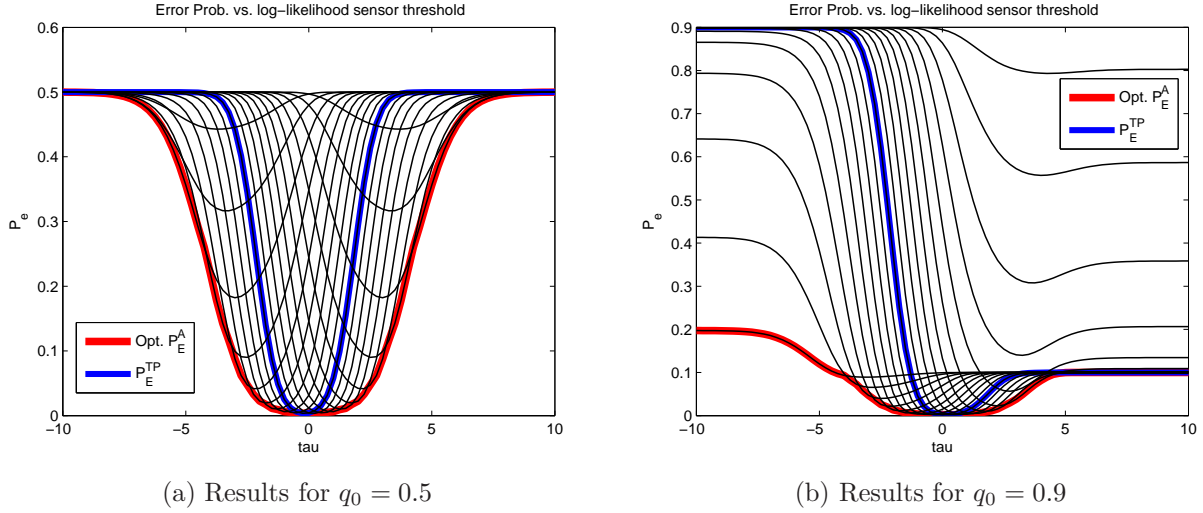


Figure 5.3: Comparison of performance of AFC and TPFC for $n = 20$, $p_1 = 0.1$, $p_2 = 0.1$ and $d = 1$ in the presence of Gaussian noise

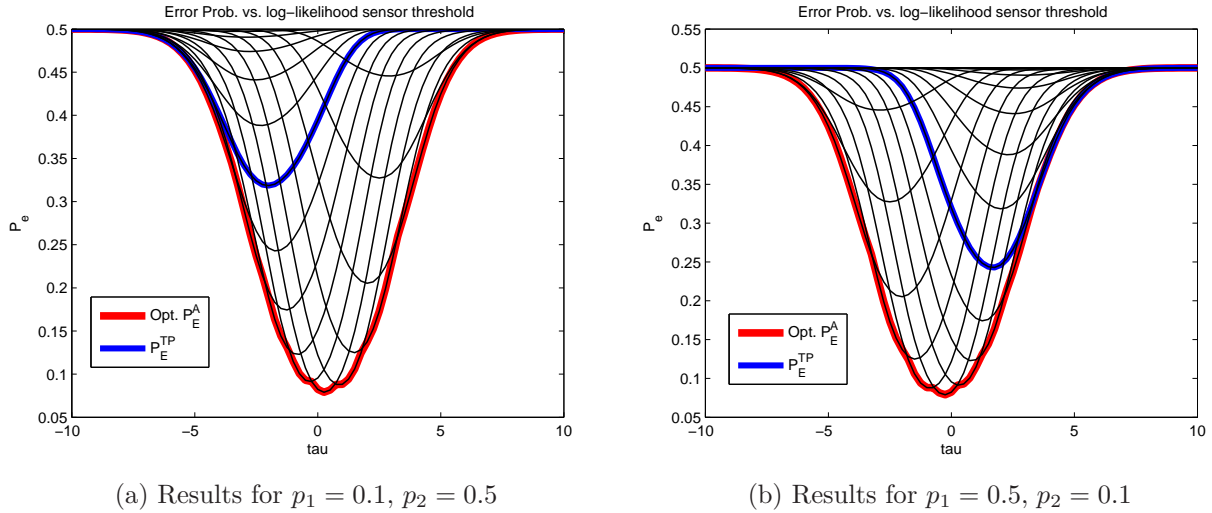
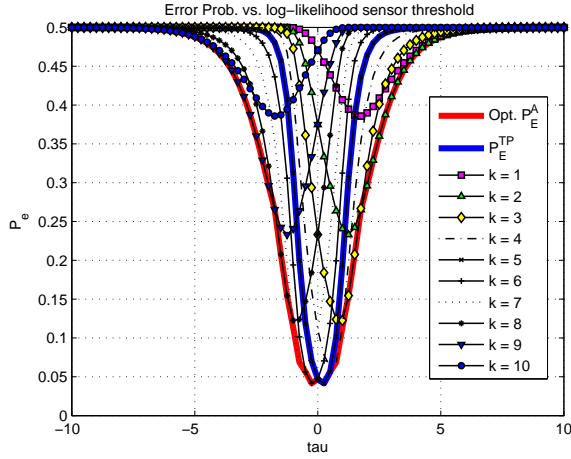
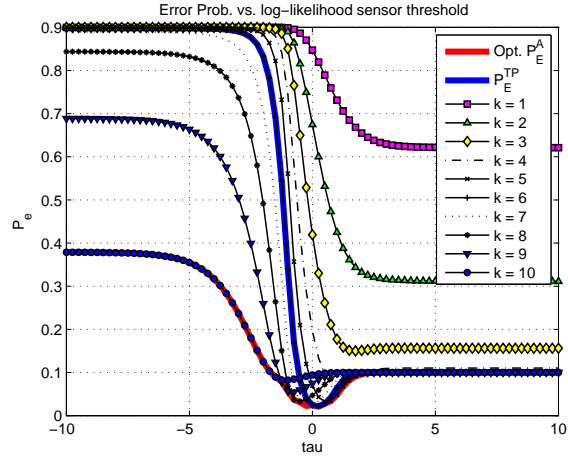


Figure 5.4: Comparison of performance of AFC and TPFC for $n = 20$, $q_0 = 0.5$ and $d = 1$ in the presence of Gaussian noise

and the same arguments can be used to explain these results. The only difference observed between Gaussian noise model and Laplacian noise model is that the curves are more steeper in the case of Laplacian noise model which may be due to the fact that Laplacian distribution has a discontinuity at its mean.

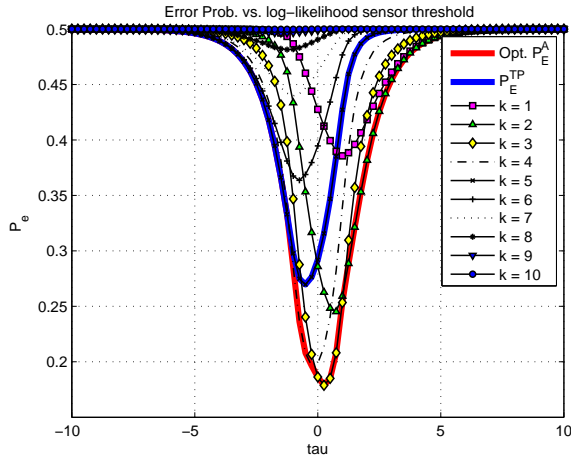


(a) Results for $q_0 = 0.5$

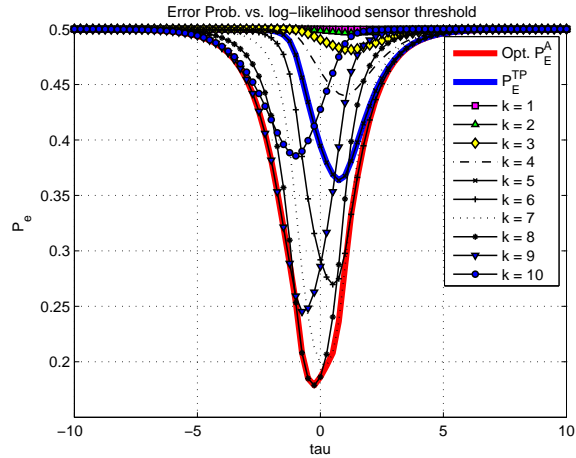


(b) Results for $q_0 = 0.9$

Figure 5.5: Comparison of performance of AFC and TPFC for $n = 10$, $p_1 = 0.1$, $p_2 = 0.1$ and $d = 1$ in the presence of Laplacian noise



(a) Results for $p_1 = 0.1$, $p_2 = 0.5$



(b) Results for $p_1 = 0.5$, $p_2 = 0.1$

Figure 5.6: Comparison of performance of AFC and TPFC for $n = 10$, $q_0 = 0.5$ and $d = 1$ in the presence of Laplacian noise

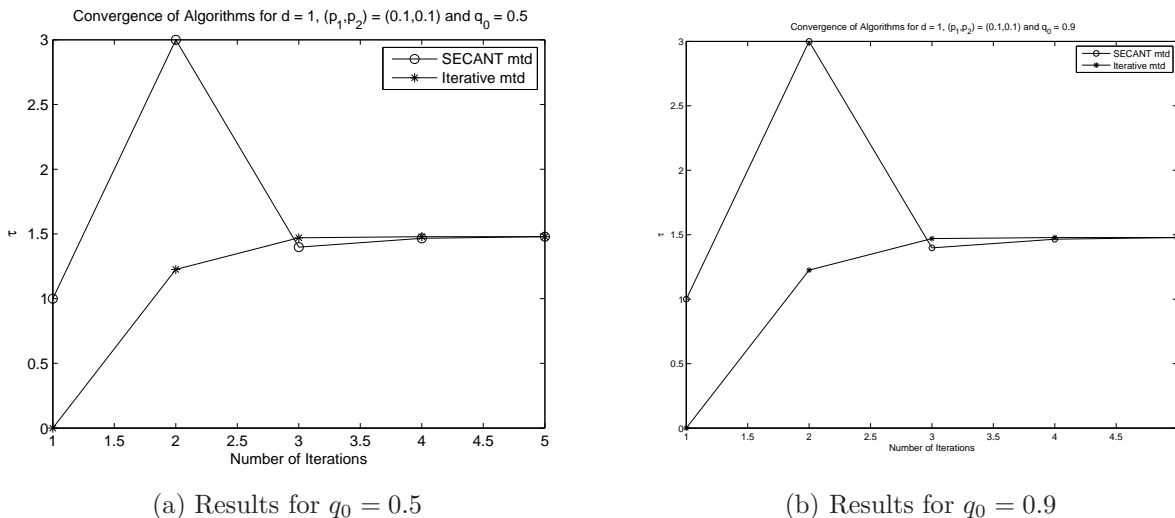


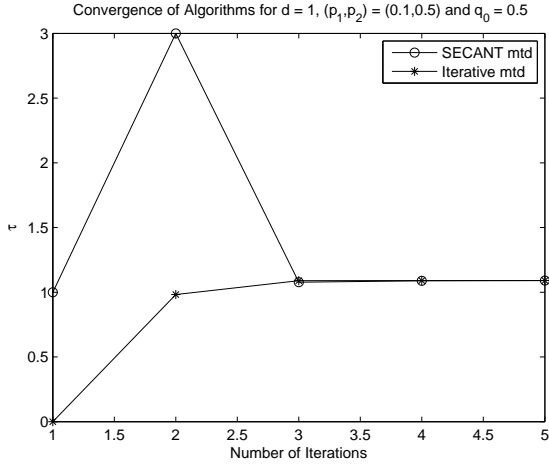
Figure 5.7: Comparison of convergence of the secant method with the proposed iterative method for $n = 10$, $d = 1$, $p_1 = 0.1$ and $p_2 = 0.1$ in the presence of Gaussian noise

5.2 Convergence of Numerical Algorithms

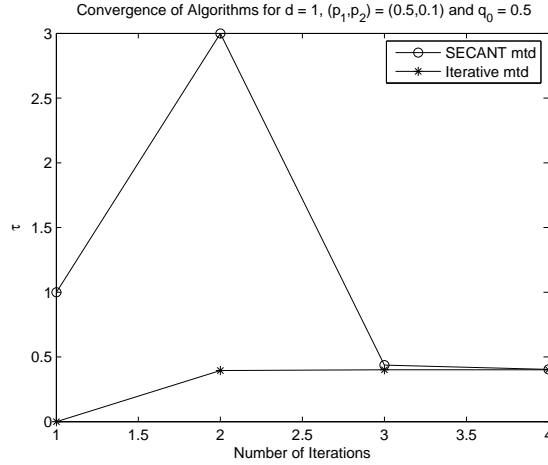
The next stage is the numerical computation of the optimal thresholds. In fact, the numerical computation is performed in the previous results where the optimal P_e^A is depicted. But since we presented two different algorithms to compute the optimal λ , we are more interested in comparing the convergence of the two algorithms. The convergence arguments depend on the initial values, we try to observe this for different initial values and find that both the algorithms converge almost at the same rate, although our iterative algorithm beats the secant method with a little difference which is almost negligible.

Note that in the case of Laplacian distribution, esp. in the case of cipher with skewed parameters as in figure 5.10, convergence is much faster in the case of our iterative method. It results in a solution in the very first iteration. While in the case of secant algorithm, it was continuing to take more than 4 iterations, although the difference is very less. So the difference is explicit only if we go for higher accuracy and precision in finding the roots.

But the one advantage we have with the proposed iterative method is that we only start with one initial value and hence, there are less number of computations to start with, making it a faster algorithm in time.

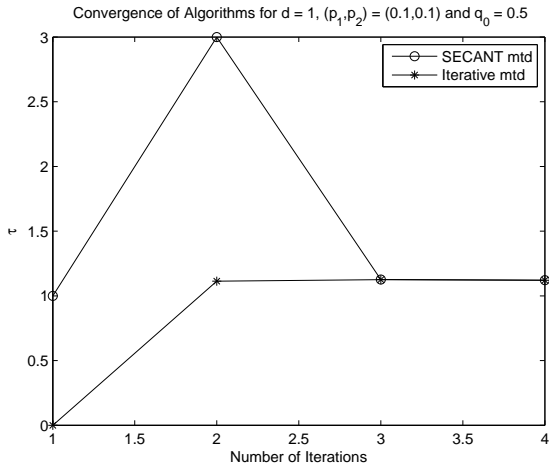


(a) Results for $p_1 = 0.1$ and $p_2 = 0.5$

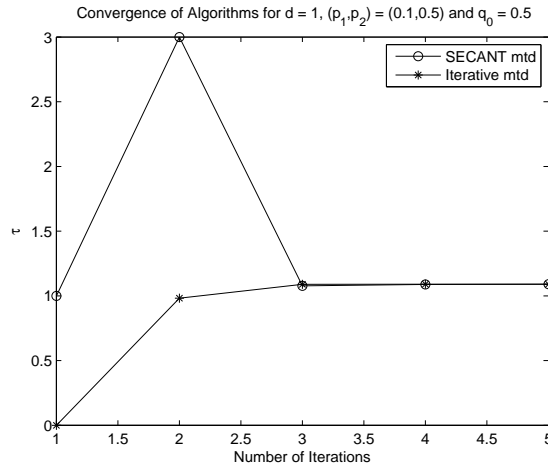


(b) Results for $p_1 = 0.5$ and $p_2 = 0.1$

Figure 5.8: Comparison of convergence of the secant method with the proposed iterative method for $n = 10$, $d = 1$ and $q_0 = 0.5$ in the presence of Gaussian noise

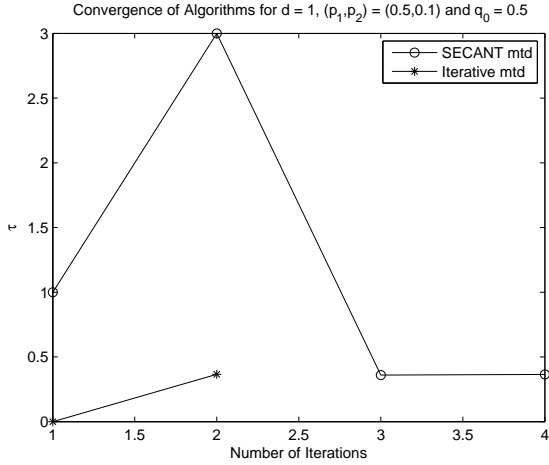


(a) Results for $q_0 = 0.5$

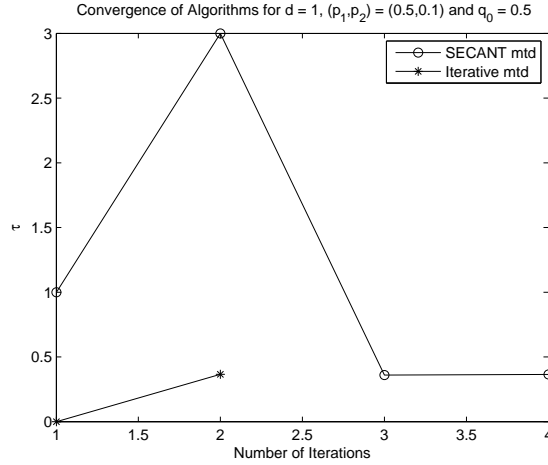


(b) Results for $q_0 = 0.9$

Figure 5.9: Comparison of convergence of the secant method with the proposed iterative method for $n = 10$, $d = 1$, $p_1 = 0.1$ and $p_2 = 0.1$ in the presence of Laplacian noise

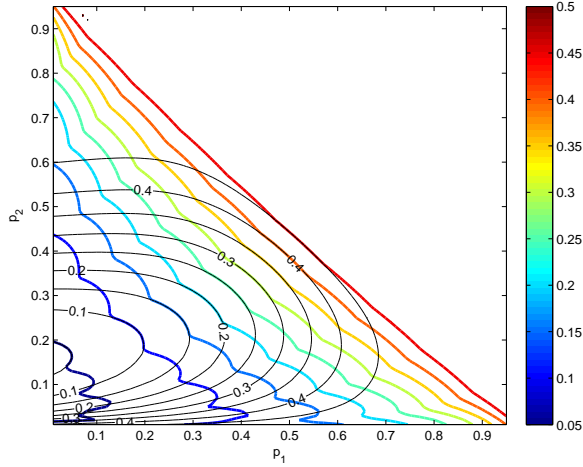


(a) Results for $p_1 = 0.1$ and $p_2 = 0.5$

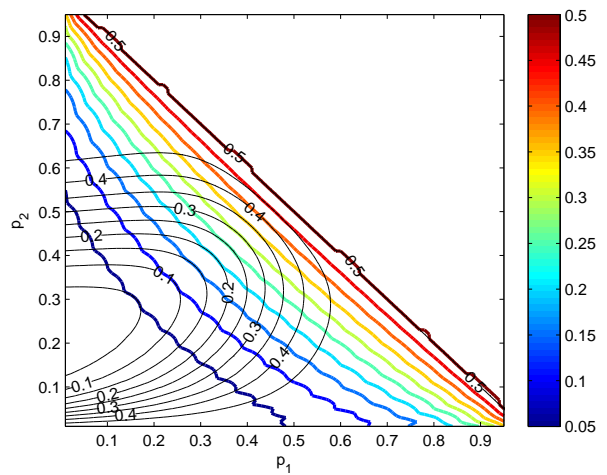


(b) Results for $p_1 = 0.5$ and $p_2 = 0.1$

Figure 5.10: Comparison of convergence of the secant method with the proposed iterative method for $n = 10$, $d = 1$ and $q_0 = 0.5$ in the presence of Laplacian noise



(a) Results for $n = 10$



(b) Results for $n = 20$

Figure 5.11: Constrained Optimization of AFC over TPCF in the presence of Gaussian noise for $d = 1$ and $q_0 = 0.5$

5.3 Constrained Optimization

After the optimal λ is computed numerically for a given (p_1, p_2) as described in the previous section, our next step is to find the best cipher that fits the problem we formulated in

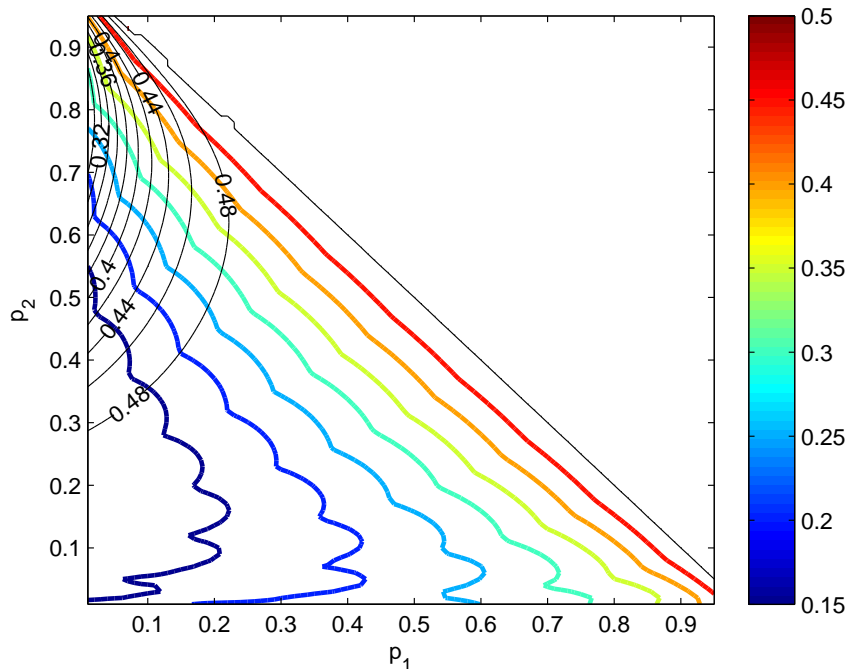


Figure 5.12: Constrained Optimization of AFC over TPFC in the presence of Laplacian noise for $n = 10$, $d = 1$ and $q_0 = 0.5$

Chapter 3. We solve this by constraining the TPFC's error probability with a lower bound on P_e^{TP} and then finding the cipher which minimizes P_e^A , as depicted by figures 5.11 and 5.12 for Gaussian and Laplacian noise models, respectively.

In these figures, the colored contours represent P_e^A while the black contours represent P_e^{TP} . Each color represents a value that is projected in the color-bar shown adjacent to the graph. Say, if we constraint the TPFC's performance as $P_e^{TP} \geq \alpha$, then we find that contour of P_e^A with minimum value which intersects the $P_e^{TP} = \alpha$ contour. Intersection of these two contours gives the values of p_1 and p_2 . This, in turn, gives the optimal value of λ and k , thus completing the task of designing the optimal fusion center.

It is equally important to note that as n increases, there exists an intersection between P_e^A and P_e^{TP} contours even if the difference between them is large. This can be articulated from figure 5.11 where 5.11a has all the colored contours (P_e^A) concentrated close of the unsecured end of the graph, which is represented by the point $(p_1 = 0, p_2 = 0)$. While as n increases, as in figure 5.11b, the contours move towards the diagonal represented by $p_1 + p_2 = 1$, enhancing the difference between P_e^A and P_e^{TP} and making it a more secure system.

6 Conclusion and Future Work

Thus there is a clear picture of the comparison of both the AFC and the TPFC designs. We first found a general condition for the quasi-convexity of P_e and then proved that both additive Gaussian and Laplacian models satisfy the condition. Two fast numerical algorithms were presented to compute the optimal thresholds. A significant improvement in the difference of performance is observed in the presence of a cipher esp. in the case of unequal cipher parameters p_1 and p_2 , even in the worst-case scenario when the hypotheses are equiprobable. Finally, we also presented the improvement in the performance of the design as the number of sensors increases.

Although we could not solve the constrained optimization problem analytically, we provided the numerical results that we achieved from simulations which give us a motivation to adopt this scheme. In other words, security is embedded in the design that allows the AFC design to be more reliable and the information is protected from the other optimal TPFC designs.

”Just when the caterpillar thought the world was over, it became a butterfly.” - Anonymous.

Just as the above proverb quotes, solving this problem raised many new interesting questions. All this started with the distributed estimation problem which was solved by Aysal *et al.* in [4]. This paper motivated us to introduce a similar cipher in a distributed detection problem. Now, we would like to follow the same trends which we find the unsecured distributed detection problem. Let us go through some of the interesting problems one-by-one.

The immediate extension to the problem we worked on is to find a general class of noise distributions that hold the quasiconvexity property for secure sensor networks. It is important to know what makes a distribution eligible to participate in a secure distributed detection problem. Also, we would like to extend this to other network topologies like serial and tree topologies. Furthermore, another interesting extension to this problem is to use different cipher constructions and evaluate the performance of the secure sensor network, in the same way we solved this problem.

Sensor Networks became a very hot topic of research because of the unusual constraints which we do not find in other optimal design problems. One such constraint is the limitation of energy consumption in the individual sensors and hence, energy-efficient schemes were proposed by several authors as discussed in Chapter 1. Therefore, we might be interested

in developing energy-efficient schemes with a new dimension of security embedded into the system model and solve the distributed detection problem for sensor networks.

Another interesting extension to [4] would be to extend the estimation problem to target tracking problem. Also, similar to [19], it is interesting if we can work on a secure distance-based fusion center that exploits the spatial effects of the phenomenon of interest with reference to the location of sensors.

Thus, a problem which we thought is almost dead, now turned to be a very interesting one.

References

- [1] Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E., "A Survey on Sensor Networks", *IEEE Communications Magazine*, pp. 102-114, August 2002.
- [2] Appadwedula, S., Veeravalli, V.V., Jones, D.L., "Energy-Efficient Detection in Sensor Networks", *IEEE J. Selected Areas of Communications*, Vol.23, No.4, pp. 693-702, April 2005.
- [3] Appadwedula, S., "Energy-Efficient Sensor Networks for Detection Applications", Ph.D. thesis, Dept. of Electrical Engineering, Univ. Illinois at Urbana-Champaign, Urbana, Illinois, 2003.
- [4] Tuncer Can Aysal and Kenneth E. Barner, "Sensor Data Cryptography in Wireless Sensor Networks," *IEEE Trans. Info. Forensics and Security*, Vol. 3, No. 2, pp. 273-289, June 2008.
- [5] Z. Chair and P. K. Varshney, "Optimal Data Fusion in Multiple Sensor Detection Systems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES22(1), pp. 98-101, Jan 1986.
- [6] Chamberland, J-F., Veeravalli, V.V., "Decentralized Detection in Sensor Networks," *IEEE Trans. Signal Processing*, Vol. 51, No. 2, pp. 407-416, February 2003.
- [7] Chamberland, J-F., Veeravalli, V.V., "Asymptotic results for Decentralized Detection in Power Constrained Wireless Sensor Networks," *IEEE J. Selec. Areas Comm.*, Vol. 22, No. 6, pp. 1007-1015, August 2004.
- [8] Chen, P-N., Papamarcou, A., "New Asymptotic Results in Parallel Distributed Detection," *IEEE Trans. Info. Theory*, Vol. 39, No. 6, pp. 1847-1863, November 1993.
- [9] Biao Chen, Peter K. Willett, "On the Optimality of the Likelihood-Ratio Test for Local Sensor Decision Rules in the Presence of Nonideal Channels," *Proceedings of*

-
- IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*. , Vol. 4, pp. 2033 - 2036, May 7-11 2001.
- [10] Estrin, D. Girod, L., Pottie, G., Srivastava, M., "Instrumenting the World with Wireless Sensor Networks," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*. , Vol. 4, pp. 2033 - 2036, May 7-11 2001.
- [11] Krishna Kishore Gunturu, "Optimum Energy Allocation for Detection in Wireless Sensor Networks", Masters thesis, Dept. of Electrical and Computer Engineering, Louisiana state University, August 2007.
- [12] Irving, W.W., Tsitsiklis, J.N., "Some Properties of Optimal Thresholds in Decentralized Detection," *IEEE Trans, Automatic Control*, Vol. 39, No. 4, pp. 835-838, April 1994.
- [13] R.Nui, B. Chen, and P. K. Varshney, "Fusion of Decisions Transmitted over Rayleigh Fading Channels in Wireless Sensor Networks," *IEEE Trans. Signal processing*, Vol. 54, No. 3, pp. 1018-1027, March 2006.
- [14] Gregory J. Pottie, "Wireless Sensor Networks," *Inform. Theory Workshop*, Killarney, Ireland, June-22-26, 1998.
- [15] C. Rago, P. Willet, and Y. Bar-shalom, "Censoring Sensors: A Low Communication-Rate Scheme for Distributed Detection," *IEEE trans. Aerosp. Electron. Syst*, Vol. 32, No. 2, pp. 554-568, April 1996.
- [16] Leonard F. Richardson, *Advanced Calculus: An Introduction to Linear Analysis*, John Wiley and Sons, Inc., New Jersey, 2008.
- [17] W. Shi, T.W. Sun, R.D. Wesel, "Quasi-convexity and optimal binary fusion for distributed detection with identical sensors in generalized Gaussian noise," *IEEE Trans. on Information Theory*, Vol. 47, No. 1, pp. 446-450, January 2001.
- [18] Per Ola Börjesson, and Carl-Erik W. Sundberg, "Simple Approximations of the Error Function $Q(x)$ for Communications Applications," *IEEE Trans. Communications*, Vol. COM-27, No. 3, pp. 639-643, March 1979.
- [19] Sung, Y., Tong, L., Swami, A., "Asymptotic Locally Optical Detector for Large-Scale Sensor Networks under the Poisson Regime", *IEEE Trans. Signal Processing*, Vol. 53, No. 6, pp. 2005-2017, June 2005.
- [20] Sung, Y., Tong, L., Swami, A., "Asymptotic Locally Optical Detector for Large-Scale Sensor Networks under the Poisson Regime", Technical Report No. ACSP-TR-10-03-01, Adaptive Communications and Signal Processing Laboratory, Cornell University, Ithaca, NY, October 16, 2003.

- [21] P. F. Swaszek, "On the Performance of Serial Networks in Distributed Detection," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 29, No. 1, pp. 254-260, January 1993.
- [22] Z. B. Tang, K. R. Pattipati, and D. L. Kleinman, "Optimization of Detection Networks: Part I- Tandem Structures," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 21, No. 5, pp. 1044-1059, 1991.
- [23] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "Asymptotic Performance of a Censoring Sensor Network," *IEEE Transactions on Information Theory*, Vol. 53, No. 11, pp. 4191-4209, November 2007.
- [24] R.R. Tenney and N.R. Sandell, "Detection with Distributed Sensors," *IEEE Trans on Aerospace and Electronic Systems*, Vol. 17, No. 4, pp. 501-509, July 1981.
- [25] John Tsitsiklis, "Decentralized Detection by a Large Number of Sensors," *Math. Control, Signals, System*, Vol. 1, No. 2, pp. 167-182, 1988.
- [26] J. N. Tsitsiklis, "Decentralized Detection," *Advances in Signal Processing*, Vol. 2, H. V. Poor and J. B. Thomas, editors, JAI Press, pp. 297-344, 1993.
- [27] Pramod K. Varshney, *Distributed Detection and Data Fusion*, Springer, New York, 1997.
- [28] Viswanathan, R., Varshney, P.K., "Distributed Detection with Multiple Sensors: Part I Fundamentals," *Proceedings of the IEEE*, Vol. 85, No. 1, pp. 54-63, January 1997.
- [29] Qian Zhang, Pramod K. Varshney and Richard D. Wesel, "Optimal Bi-level Quantization of i.i.d. Sensor Observations for Binary Hypothesis Testing," Vol. 48, No. 7, pp. 2105-2111, July 2002.

Vita

Venkata Sriram Siddhardh Nadendla was born in April, 1986 in Rajahmundry, Andhra Pradesh, India. He graduated with his Bachelor of Engineering in Electronics and Communication Engineering from Sri Chandrasekharendra Saraswathi Viswa Maha Vidyalaya, in short, SCSVMV University, Kanchipuram, India in the year 2007 with his bachelor's thesis on "DWRR scheduling and CRC computation of packets in a packet processor for 10G Ethernet Networks using Verilog HDL". He is presently pursuing his Master of Science in Electrical Engineering at Louisiana State University and is expected to graduate in August 2009. His research interests include Digital/Wireless Communications, information theory and coding theory and his present focus is on security issues in distributed detection in wireless sensor networks.

He was awarded a Silver Medal and Dr. S. Subbulakshmi Endowment Cash Prize in SCSVMV University for his cumulative GPA of 9.5 (on the scale of 10) which was the highest among all the graduating candidates in 2007. He was also awarded Dr. S. Suryanarayanan Endowment Cash Prize for standing first in Chemistry in SCSVMV University Examinations, 2007. He was also offered merit-scholarships throughout his undergraduate education by SCSVMV University.

He worked as a teaching assistant in the Department of Electrical and Computer Engineering, Louisiana State University and has taught courses like electric and magnetic fields, signals and systems, random processes-1 and digital logic design lab. He is a graduate student member of IEEE and has volunteered for IEEE Global Communications Conference (GLOBECOM-2008), New Orleans, 2008.