Communication-efficient decentralized sequential detection

by

Hongfei Wang

A thesis submitted to the
Department of Electrical and Computer Engineering
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
October 2016

Abstract

The problem of decentralized sequential detection is studied in this thesis, where local sensors are memoryless, receive independent observations, and no feedback from the fusion center. In addition to traditional criteria of detection delay and error probability, we introduce a new constraint: the number of communications between local sensors and the fusion center. This metric is able to reflect both the cost of establishing communication links as well as overall energy consumption over time. A new formulation for communication-efficient decentralized sequential detection is proposed where the overall detection delay is minimized with constraints on both error probabilities and the communication cost.

Two types of problems are investigated based on the communication-efficient formulation: decentralized hypothesis testing and decentralized change detection. In the former case, an asymptotically person-by-person optimum detection framework is developed, where the fusion center performs a sequential probability ratio test based on dependent observations. The proposed algorithm utilizes not only reported statistics from local sensors, but also the reporting times. The asymptotically relative efficiency of proposed algorithm with respect to the centralized strategy is expressed in closed form. When the probabilities of false alarm and missed detection are close to one another, a reduced-complexity algorithm is proposed based on a Poisson arrival approximation.

In addition, decentralized change detection with a communication cost constraint is also investigated. A person-by-person optimum change detection algorithm is proposed, where transmissions of sensing reports are modeled as a Poisson process. The optimum threshold value is obtained through dynamic programming. An alternative method with a simpler fusion rule is also proposed, where the threshold values in the algorithm are determined by a combination of sequential detection analysis and constrained optimization. In both decentralized hypothesis testing and change detection problems, tradeoffs in parameter choices are investigated through Monte Carlo simulations.

Acknowledgments

I would like to thank my supervisors Prof. Steven D. Blostein for his support, suggestions, and encouragement throughout my PhD program. He educated me with his abundant engineering knowledge and guided me patiently all the time. Without his persistent help, this thesis would not have been possible.

I would also like to thank Prof. Rakesh Bansal from IIT-Kanpur India, Prof. Saeed Gazor, and Prof. Serdar Yuksel for serving on my thesis committee. Special thanks go to Prof. Serdar Yuksel for his suggestions and insightful ideas during my comprehensive exams.

My sincere thanks also go to all members and staff of the Electrical and Computer Engineering Department, which provided a friendly and resourceful environment for conducting research work. I have been fortunate to work in such pleasant atmosphere. Thanks especially to Debra Fraser, who provided me valuable suggestions and helped me patiently all the time.

There are too many friends to mention individually, who have assisted me in so many ways during my study and life. I would like to thank all the my lab mates: Ye Li, Yu Cao, Majid Bavand, Phillip Oni, James Falt, Nidhi Sharma and Hassan Abou Saleh for all their support.

Last, and most important, I would like to thank my family. I cannot find words to express my gratitude to my dear wife Valerie Zhang for all her love and encouragement. I also owe a lot to my parents for their continuous support and endless love. I dedicate my thesis to them.

Contents

List of Tables

List of Figures

Acronyms

ARE Asymptotically Relative Efficiency.

CDF Cumulative distribution function.

CUSUM Cumulative Sum.

DS Data subsequence.

FSS Fixed sample size.

IID Independent and Identically Distributed.

KKT Karush-Kuhn-Tucker conditions.

PDF Probabiliy Density Function.

PMF Probability Mass Function.

SPRT Sequential Probability Ratio Test.

TS Tail subsequence.

List of Important Symbols

M The number of local sensors in the sensing system.

P_{FAC} False alarm probability of the fusion center test.

P_{FAL} False alarm probability of the local sensor test.

P_{MDC} Missed detection probability of the fusion center test.

P_{MDL} Missed detection probability of the local sensor test.

R_{FAC} False alarm rate at the fusion center.

R_{FAL} False alarm rate at the local sensor.

Γ Change time of the observed distribution.

α_C Fusion center false alarm probability constraint.

α_L Local sensor false alarm probability constraint.

β_C Fusion center missed detection probability constraint.

β_L Local sensor missed detection probability constraint.

γ_C Fusion center false alarm rate constraint.

γ_L Local sensor false alarm rate constraint.

κ Communication cost constraint.

$\mathbb{R}$ The set of real numbers.

$\mathcal{R}$ Distribution of the reported information.

$\mathcal{T}$ Distribution of the report-generating delay.

τ Stopping time of the detection task.

g Local sensor test statistic.

r Reporting information.

v Detection delay of the local sensor test.

Chapter 1

Introduction

1.1 Sequential detection

The problem of sequential detection is of considerable practical importance in a variety of applications including cognitive radio networks, quality control engineering, biomedical signal processing and surveillance systems, etc. [1]-[2]. Unlike fixed sample size detection schemes [3]-[14], where the number of collected measurements is pre-specified, sequential detection allows the flexibility of stopping the test when observations are informative enough to guarantee a desired performance usually expressed in terms of error probability. As decisions are made as quickly as possible, sequential detection requires, on average, fewer observations than that of fixed sample size detection.

Existing research work conducted on sequential detection may be grouped into two broad categories: sequential hypothesis testing [16]-[18] and sequential change detection [19]-[23]. The problem of sequential hypothesis testing is to decide between two possible statistical models based on sequentially observed random variables. It is shown that Wald's sequential probability ratio test (SPRT) offers optimal performance in sense that a minimum expected number of observations is needed to achieve predefined error probabilities [16]. On the other hand, the change detection problem is to find, as quickly as possible, an abrupt change in the statistical behavior of stochastic observations [15]. It is of interest to perform detection of a change in a way that minimizes the delay between the time that a change occurs and the time it is detected. The optimality of change detection strategies is studied in [24]-[26].

In wireless communication problems caused by multipath fading, shadowing and hidden terminal phenomena, single node, or centralized based sequential detection encounters many limitations [27]-[28]. As a result, there has been increasing interest in decentralized formulations of sequential detection [29]-[48]. In a decentralized system, information about the sensed target is available through measurements taken by a set of geographically distributed local sensors. A central entity (fusion center) makes final decision based only on the local decisions or summarized statistics reported by these sensors.

Previous investigations of decentralized detection apply both hard combining and soft combining strategies at the fusion center. In hard combining strategies, local sensors make decisions based on their own observations and subsequently forward binary information to the fusion center [29], [43]-[44]. Based on reported local sensing results, the fusion center can apply AND, OR, or majority rule for decision making. However, with a hard combining strategy, the fusion center cannot make a decision until local sensors have made decisions. This means that detection delay is not significantly reduced even though multiple sensors are used. For soft combining strategies

[31]-[37] [47]-[49], a local sensor sends a sequence of quantized observations or summarized information to the fusion center. Reported information by local sensors are viewed as observations of the fusion center, based on which a sequential test is carried out to make the final decision.

1.2 Related work

1.2.1 Decentralized hypothesis testing

A hard combining based decentralized framework is proposed in [29], where local sensors make their own decisions on the hypothesis, and the fusion rule applies either AND, OR, or Majority rule. A decentralized Wald's problem is investigated in [30]. Although each detector makes its own final decision, the detection policy is designed to minimize the cost of the entire detection system. In [31], local sensing information is quantized through likelihood ratio tests. Based on achieved summarized messages, a sequential test is performed at the fusion center. Under the setup that local sensors have full feedback from the fusion center and local memories are restricted to past decisions, an optimal decentralized sequential detection can be found. In [33], an alternative configuration with full or limited local memory but no fusion center feedback is considered. Though an optimal solution for such a setup is intractable, an asymptotically optimal solution can be achieved. In the case that local memory is limited, each local sensor summarizes its previous observations; together with its current observations, quantized information is send to the fusion center. This configuration is also considered in [34], where asymptotically optimal schemes are suggested in both discrete and continuous time scenarios. In the system considered in [35], neither feedback from the fusion center nor the local memory is available at local sensors. It is shown that person-by-person optimal sensor decision functions for such configuration are likelihood ratio tests. However, the optimal thresholds, which satisfy a set of coupled equations, are almost impossible to solve numerically. Moreover, asynchronous communication between local sensors and the fusion center is considered in [34]. That is, rather than transmit sensing reports at common, deterministic, equidistant times [31]-[32], local sensors communicate with the fusion center when a certain local test's stopping condition is satisfied. A different asynchronous scheme is proposed in [36], where local decisions are generated according to a Poisson process at each local sensor. In [37], the communication noise between local sensors and the fusion center is considered. A DualSPRT algorithm is proposed with local memory and synchronous communication in [38]. It is assumed that the false alarm and missed detection probabilities are same at local sensors, and that the absolute values of lower and upper thresholds for local log-likelihood ratio tests are the same as well. This setup is also assumed in [37] for performance analysis. Under such conditions, both error probabilities and the delay of local sequential tests can be computed from the theory of random walks. A sequential detection framework of Gaussian binary hypothesis is developed in [39], where the communication between adjacent local sensors is available.

Most existing decentralized hypothesis testing approaches are restricted to test two simple hypotheses. The problem of sequential multiple hypothesis testing, with more than two hypotheses is more challenging; even in the classical centralized case, the solution is very complex in general [40]. A person-by-person optimal decision rule is developed in [41], where the observations are independent and uniformly distributed. A two-stage test is proposed for multiple hypothesis detection in [42], and the asymptotic optimality of such a test is established.

1.2.2 Decentralized change detection

In [43], a one-shot cumulative sum (CUSUM) scheme is introduced. Each sensor runs a CUSUM algorithm and communicates with the fusion center only when it believes that a change has occurred. In this approach, the fusion center can either apply a so-called *minimal* or *maximal* strategy for decision making. In the former case, the fusion center signals a change to have occurred when any of the sensors reports a change. In the latter case, the fusion center signals a change when all sensors report a change. In [44], a one-shot CUSUM scheme with a *minimal* fusion strategy is applied. Moreover, it is assumed that change times at local sensors may be different. An asymptotically Min-Max optimal algorithm is proposed in [55], by combining CUSUM test and a *maximal* fusion strategy. A main drawback of [43] and [44] is that the fusion center cannot make a decision until local sensors have made decisions.

A Bayesian formulation of the decentralized change detection problem with energy constraints is considered in [45], and the formulated problem is solved though dynamic programming. Another work that takes energy constraints into account is proposed in [46], where local sensors do not have full feedback. In [47], local sensors with memory are considered that record their past local decisions and also utilize full feedback from the fusion center. At each sensing period, reports generated from current local observations and past decisions from all the sensors are forwarded to the fusion center. The fusion center receives sensing reports sequentially and stops sensing when it is able to make a decision. However, such continuous communication between local sensors and the fusion center may be too costly for some applications. For example, remote sensors with energy constraints may not afford such continuous communication with the fusion center. A data-efficient quickest change detection scheme is developed in [48] and [49] based on a minimax formulation, where the cost in taking samples is controlled through a censoring technique.

One of the major problems in existing soft combining based algorithms is that unlimited local memory is required to store the sensing history [48]-[49], i.e., all previous observed information and local decisions. In [47], full feedback from the fusion center is additionally assumed to obtain the optimal solution. These conditions on sensing devices are strong and therefore restrict practical applications. Moreover, in [47]-[48], the decision policy at each local sensor requires sensing information from other local sensors as well. Thus, each local sensor must continuously communicate with the fusion center to update its local test statistic. Such communication is costly, especially when sensors have energy constraints and are sparsely located.

1.3 Thesis objective

Existing works make a variety of assumptions on the functionalities of sensing devices, including local memory, full feedback from the fusion center, etc. However, these requirements on the sensing devices may not always be satisfied and therefore restrict real applications. In practice, the communication between local sensors and the fusion center is costly.Not only does this introduce extra overhead but also results in considerable local sensor energy consumption, especially when sensors are distantly located.

The objective of this thesis is to develop decentralized sequential detection frameworks for both hypothesis testing and change detection problems, where local sensors are memoryless, without full feedback from the fusion center, and have energy constraints. Moreover, in addition to detection delay and the error probability, we introduce a new communication cost metric, which is able to control local sensor energy consumption and transmission overhead in the link between local sensors and the fusion center. The new formulated optimization problem aims to minimize the overall detection delay with constraints on both error probability and communication cost. Detection methodologies to solve formulated problems are studied in different contexts.

1.4 Thesis organization

The rest of the thesis is organized as follows.

In Chapter 2, preliminary background material to the thesis results are introduced.

In Chapter 3, the problem of decentralized hypothesis testing is investigated. A communication-efficient formulation is proposed, where average detection delay is minimized with constraints on error probabilities and on communication cost. An asymptotically person-by-person optimal algorithm is developed to solve the formulated problem, and the asymptotically relative efficiency (ARE) of the developed algorithm with respect to the centralized strategy is proposed as a performance metric.

In Chapter 4, a more practical method to solve the formulated hypothesis testing problem is developed, where sensing report transmissions from local sensors to the fusion center are modeled as a Poisson arrival process. A communication-efficient decentralized hypothesis detection algorithm is proposed, where both local sensors and the fusion center apply the sequential probability ratio test (SPRT). The ARE performance is also investigated.

The decentralized change detection problem is investigated in Chapters 5 and 6. In Chapter 5, the communication cost is formulated in two different ways, and threshold values for the optimum stopping times are investigated in both cases. Following the same model of a Poisson arrival process as in Chapter 4, a person-by-person communication-efficient decentralized change detection algorithm is proposed.

In Chapter 6, an alternative way to solve the formulated change detection problem is studied, where synchronization is not required and the fusion center applies a simpler fusing strategy. The methodology to joint threshold value design at local sensors and the fusion center is investigated.

The conclusions and suggestions for future work are provided in Chapter 7.

1.5 Contributions

The main contributions of the thesis are summarized as follows:

- New formulations for both decentralized hypothesis testing and decentralized change detection problems are presented, where local sensors are memoryless, receive independent observations, and no feedback from the fusion center.

- In addition to average detection delay and the error probability, which have been previously used for system design and performance assessment, we introduce a new constraint: the number of communications between local sensors and the fusion center. This metric is able to reflect both the cost of establishing communication links as well as overall energy consumption over time.

- An asymptotically person-by-person optimal decentralized hypothesis testing algorithm is developed to solve proposed optimization problem that exploits time dependency of locally transmitted decisions. The fusion center test utilizes not only the statistic contained in local sensing reports, but also the times at which these reports are received.

- The performance of the proposed hypothesis testing algorithm with respect to the centralized detection scheme is quantified via the asymptotic relative efficiency (ARE). We derive closed form expressions for ARE as the function of the communication cost and error probabilities at both local sensors and fusion center tests.

- A practical methodology is developed to solve formulated hypothesis testing problem, by fusing only the reported statistic at the fusion center. As the dependency of the report-generating delay is ignored, the arrival times of sensing reports are assumed memoryless, and therefore the transmissions of local reports are modeled as a Poisson process. Situations where this approximation is accurate are determined. The ARE of the proposed algorithm with respect to the centralized strategy is also investigated.

- A person-by-person optimal decentralized change detection algorithm is proposed, based on a Poisson arrival model. The communication cost is formulated in two different ways, and threshold values for the optimum stopping rule are obtained through dynamic programming.

- An alternative change detection algorithm with a simpler fusion rule is developed. In the proposed algorithm, thresholds at both local sensors and the fusion center ought to be jointly optimized. It is shown the optimal choice of thresholds in the algorithm can be obtained through one dimensional search. Precise synchronization of sensing devices can be avoided, however, at a price of performance degradation.

Chapter 2

Background

2.1 Introduction

In this chapter, preliminary background material to the thesis results are introduced. We first discuss sequential detection, which includes the problems of hypothesis testing and change detection. Then, the decentralized sensing system model is introduced. Here, rather than providing a complete and balanced review we put emphasis on what we will need in later chapters. For a more general setting and background, we refer interested readers to [15], [31], and [50].

2.2 Sequential hypothesis testing

Consider a sequence of independent and identically distributed (IID) random variables, $X_1, X_2, ...$, which has either the common probability density function (PDF) f_0 or f_1, i.e.,

$$\begin{aligned} H_0 &: X_1,\ X_2, ... \sim f_0, \\ H_1 &: X_1,\ X_2, ... \sim f_1, \end{aligned} \tag{2.1}$$

The criteria for the choice of sequential tests are the error probabilities and the average detection delay. There are two types of errors in the hypothesis testing problem,

(i) False alarm probability, $P_0(\text{Accept } H_1)$; and

(ii) Missed detection probability, $P_1(\text{Reject } H_1)$.

Smaller values of the error probabilities indicate more reliable detection results. In sequential tests, it is required that $P_0(\text{Accept } H_1) \leq \alpha$ and $P_1(\text{Reject } H_1) \leq \beta$, where α and β are constraints on the false alarm and missed detection probabilities, respectively. The detection delay, or the number of observations, required by a sequential test, is a random variable and usually characterized by its expected value, i.e.,

(i) $E_1\{v\}$; and

(ii) $E_0\{v\}$,

where v denotes the delay of the sequential test and $E_\theta\{\bullet\}$ is the conditional expectation under H_θ, $\{0, 1\}$.

It is desirable that the expected delay of the test is as short as possible, while the error probability constraints are satisfied. A classic hypothesis testing procedure is to track the likelihood ratios of the sequential observations, i.e.,

$$g_k = \sum_{i=1}^{k} \log \frac{f_1(X_i)}{f_0(X_i)}, \tag{2.2}$$

where g_k is the test statistic after taking k observations. The cumulated statistic is compared to the predefined test threshold values $a < 0 < b$, and the stopping time of the test is defined as

$$v = \inf\{v \geq 1 : g_k \notin (a, b)\}. \tag{2.3}$$

The final decision of the test is then given by

$$\begin{aligned} \text{Accept } H_1 \quad \text{if} \quad g_v \geq b, \\ \text{Reject } H_1 \quad \text{if} \quad g_v \leq a. \end{aligned} \tag{2.4}$$

The above test is the well-known sequential probability ratio test (SPRT), which is developed by Wald in 1948 [16].

Theorem 2.1 (**Wald optimality [16]**). *Among all tests (sequential or non-sequential) for which*

$$P_0(\textit{Accept } H_1) \leq \alpha \textit{ and } P_1(\textit{Reject } H_1) \leq \beta,$$

and for which $E_\theta\{v\}$, $\theta \in \{0, 1\}$, *is finite, the sequential probability ratio test with error probabilities* α *and* β *minimizes both* $E_1\{v\}$ *and* $E_0\{v\}$ *simultaneously.*

Remark 2.1. *It is further shown by Lorden that the condition of finite* $E_1\{v\}$ *and* $E_0\{v\}$ *is not needed [51].*

Proposition 2.1 (**The test threshold [16]**). *The threshold values and error probabilities are related via*

$$a \leq \log \frac{\beta}{1-\alpha} \textit{ and } b \geq \log \frac{1-\beta}{\alpha}. \tag{2.5}$$

A further pair of inequalities can be used to relate the error probabilities to the expected detection delay.

Proposition 2.2 (**The detection delay [16]**). *Suppose the random variable* $\log \frac{f_1(X_1)}{f_0(X_0)}$

has finite means $\mathfrak{d}_0$ *and* $\mathfrak{d}_1$*, respectively, under* H_θ*,* $\theta \in \{0,1\}$*. Then*

$$\begin{aligned} E_0\{v\} &\geq \mathfrak{d}_0^{-1}\left[\alpha \log \frac{1-\beta}{\alpha} + (1-\alpha)\log\frac{\beta}{1-\alpha}\right], \\ E_1\{v\} &\geq \mathfrak{d}_1^{-1}\left[(1-\beta) \log \frac{1-\beta}{\alpha} + \beta\log\frac{\beta}{1-\alpha}\right]. \end{aligned} \tag{2.6}$$

Remark 2.2. *Based on Theorem 2.1, the inequalities of Proposition 2.2 hold for all tests (sequential or non-sequential) with error probabilities* α *and* β*.*

In general, the conditions for equality in Proposition 2.1 and 2.2 will not be met. However, if the excess over boundary parts, i.e., $a - g_v$ or $g_v - b$, are negligible, the inequalities can be considered as approximate equalities, which are known as Wald's approximations. As the error probability decrease, a greater upper boundary threshold, b, and a smaller lower boundary threshold value, a, are required. Then the quantity of likelihood ratio $\log f_1(X_1)/f_0(X_0)$ becomes less significant compared to the threshold values. As a result, the effect of ignoring the excess over boundary parts is smaller. Thus, the accuracy of Wald's approximations increases as the error probabilities decrease. Further more, the bounds in Proposition 2.1 and 2.2 can be viewed as asymptotically closed form expressions for test thresholds and delays, as $\alpha, \beta \to 0$.

2.2.1 Relative efficiency

It is of interest to compare the performance of sequential tests versus that of fixed sample sized tests. The best fixed sample size (FSS) is given by

$$\sum_{i=1}^{v_F} \log \frac{f_1(X_i)}{f_0(X_i)} \underset{H_0}{\overset{H_1}{\lessgtr}} T_F, \tag{2.7}$$

where the threshold T_F and the fixed sample size v_F are chosen such that the test has false alarm and missed detection probabilities equal α and β, respectively.

Definition 2.1 (**Relative efficiency** [**52**])**.** *The relative efficiency of the detector DT_1 with respect to the detector DT_2 is*

$$RE_{DT_1,DT_2} = \frac{v_2}{v_1}, \tag{2.8}$$

where v_1 and v_2 are the expected delay of detectors DT_1 and DT_2, respectively.

From (2.6), we have the relative efficiency of FSS test with respect to SPRT given by

$$RE_{FSS,SPRT} = \frac{E_\theta\{v\}}{v_F}, \theta \in \{0,1\}. \tag{2.9}$$

The above relative efficiency depends on α and β and, in general, on the H_θ, $\theta \in \{0,1\}$. The asymptotic relative efficiency (ARE) is the limiting value of the relative efficiency as H_1 approaches H_0.

Proposition 2.3 (**Asymptotic relative efficiency** [**53**])**.** *The asymptotic relative efficiency of the FSS test with respect to the SPRT is given by*

$$\begin{aligned} ARE_0 &= -2\left(\alpha \log \frac{1-\beta}{\alpha} + (1-\alpha)\log\frac{\beta}{1-\alpha}\right)\left(\mathfrak{N}^{-1}(\alpha) + \mathfrak{N}^{-1}(\beta)\right)^{-2}, && (2.10)\\ ARE_1 &= 2\left((1-\beta)\log\frac{1-\beta}{\alpha} + \beta\log\frac{\beta}{1-\alpha}\right)\left(\mathfrak{N}^{-1}(\alpha) + \mathfrak{N}^{-1}(\beta)\right)^{-2}, && (2.11) \end{aligned}$$

where ARE_θ denotes the asymptotic relative efficiencies under H_θ, $\theta \in \{0,1\}$, and $\mathfrak{N}(\bullet)$ is the standard normal distribution.

2.2.2 Dependent observations

In case that the observations $X_1, X_2, ...$ are statistically dependent, the test statistic, g_k, is no longer a sum of IID random variables. The test statistic of SPRT is the ratio of joint probability density functions of observed information, i.e.,

$$\frac{f_{d,1}(X_1, X_2, ...)}{f_{d,0}(X_1, X_2, ...)} \begin{cases} \leq a_k, & H_0 \text{ is true}, \\ \geq b_k, & H_1 \text{ is true}, \\ otherwise, & \text{continue taking observations.} \end{cases} \tag{2.12}$$

Theorem 2.2 (**Optimality with dependent Gaussian variables [69]**). *For dependent Gaussian random variables, the optimum detector has the form of* (2.12).

However, the test threshold values in (2.12), $a_k < 0 < b_k$, are time varying, and the design of which is still an open problem. If observations are not Gaussian random variables, no optimum tests have been established. There are relatively few sequential detection results developed using dependent observations, most of which use constant threshold values $a < 0 < b$.

Proposition 2.4. *The Wald's SPRT in* (2.12) *still preserves the inequalities* (2.5).

It is further shown in [73] that with constant threshold values $a < 0 < b$, the SPRT test in (2.12) is asymptotically optimum as $\alpha + \beta \to 0$, in certain cases.

2.3 Sequential Change detection

We now discuss another type of sequential detection: sequential change detection. Consider a sequence of observations $X_1, X_2, ...$ which initially has common probability density function (PDF) f_0, and starts to follow common PDF f_1 after a random change point Γ. Such a change point is often modeled as geometrically distributed. Let $0 < \rho < 1$ denote the distribution parameter, and $0 \leq \pi_0 < 1$ denote the prior probability that a change happened before the test. We have

$$P(\Gamma = k) = \pi_0 \mathrm{I}_{\{k=0\}} + (1 - \pi_0)\rho(1 - \rho)^{k-1}\mathrm{I}_{\{k \geq 1\}}, \tag{2.13}$$

where $\mathrm{I}_{\{\bullet\}}$ is the indicator function, and π_0 represents the probability of the change having occurred before the observations are taken. The change detection problem can be then expressed as the following multiple hypothesis testing problem:

$$\begin{aligned} H_0 &: X_1, ..., X_{\Gamma-1} \sim f_0, \\ H_1 &: X_\Gamma, X_{\Gamma+1}, ... \sim f_1, \end{aligned} \tag{2.14}$$

where each integer Γ indicates a different hypothesized change time.

During the change detection procedure, after taking a new observation, the detector either stops and claims a change or continues taking a new observation. It is desirable that the test stops as quickly as possible when a change has occurred. Let τ denote the stopping time of a certain sequential test. The criteria of the change detection can be then expressed as

(i) The expected detection delay, $E\{(\tau - \Gamma)^+\}$; and

(ii) False alarm probability, $P(\tau < \Gamma)$,

where $(x)^+ = \max\{x, 0\}$. The stopping time τ of a certain sequential test is determined to achieve optimal trade-offs between the average detection delay and false alarm probability. This can be implemented by solving following optimization problem

$$\inf_{\tau}[P(\tau < \Gamma) + cE\{(\tau - \Gamma)^{+}\}], \tag{2.15}$$

where $c > 0$ is a constant controlling the relative importance of the two performance metrics. The formulation of (2.15) for the change detection problem is first introduced in [26].

Definition 2.2. *The posterior probability that a change has occurred at the kth time slot,*

$$\pi_k = P(\Gamma \leq k | F_k), \tag{2.16}$$

where $F_k = \sigma(X_1, X_2, ..., X_k)$ are the sigma algebras generated by the observed information up to the kth time slot.

Remark 2.3. *Under the geometric prior* (2.13)*, the sequence of posteriors $\{\pi_k\}$ can be calculated recursively, i.e.,*

$$\pi_k = \frac{P_1(X_k)\,[\pi_{k-1} + \rho(1 - \pi_{k-1})]}{P_1(X_k)\,[\pi_{k-1} + \rho(1 - \pi_{k-1})] + (1 - \rho)(1 - \pi_{k-1})}. \tag{2.17}$$

Theorem 2.3 (**Kolmogorov and Shiryaev [26]**)**.** *For appropriately chosen threshold $\pi^* \in \{0, 1\}$, the stopping time*

$$\tau_S = \inf\{k \geq 0 | \pi_k \geq \pi^*\} \tag{2.18}$$

is optimum in solving (2.15) *with the prior* (2.13)*. Moreover, if $c \geq 1$, then $\pi^* = 0$.*

So far we assume that the distribution of the change point is known in prior (2.13). However, there are some situations that the pre-existing statistic model for the change occurrence is unavailable. In such cases, the change detection problem is formulated alternatively, and the "worst case" detection delay is evaluated.

Definition 2.3 (**Lorden [70]**). *The "worst case" detection delay of a change detection problem is given by*

$$E_\Gamma\{v\} = \sup_{\Gamma \geq 1} \left(ess \sup E_\Gamma\{(\tau - \Gamma)^+ | X_1, ..., X_{\Gamma-1}\} \right). \tag{2.19}$$

Remark 2.4. *If $E_\Gamma\{v\}$ is finite, then we will have a false alarm with probability one even if there are no changes. The criterion of the false alarm can also be measured by the false alarm rate, i.e.,*

$$R_f = \frac{1}{E_\infty\{v\}} \leq \gamma. \tag{2.20}$$

Here $E_\infty\{\bullet\}$ is the conditional expectation when there is no change and γ is the false alarm rate constraint. A good decision policy for change detection should have $E_\infty\{v\}$ large and at the same time obtain small $E_\Gamma\{v\}$. The well known cumulative sum test (CUSUM) is first introduced by Page [19], where the detector tracks likelihood ratios for all possible change hypotheses given by

$$g_k = \max_{1 \leq j \leq k} \sum_{i=j}^{k} \log \frac{f_1(X_i)}{f_0(X_i)}. \tag{2.21}$$

At each time slot the test statistic is compared to a pre-defined threshold, h, the value of which is set to satisfy the false alarm constraint. The stopping time can be then expressed as

$$\tau_h = \inf\{k \geq 1 | g_k \geq h\}. \tag{2.22}$$

Theorem 2.4 (**Moustakides [24]**). *τ_h minimizes the worst case detection delay*

(2.19) *among all stopping times satisfying* $\gamma = 1/E_\infty\{\tau_h\}$.

From (2.2) and (2.21), the CUSUM test can be expressed as a sequence of sequential ratio probability tests (SPRT) with boundaries $(0, h_l)$ and initial statistic zero [19]. Each time the lower boundary is exceeded, the test resets its statistic to zero and continues taking observations. Such procedure is repeated until the upper boundary is reached, and then a change is claimed.

2.4 Decentralized sequential detection

Sequential detection with a single sensing device encounters many limitations. As a result, there has been increasing interest in the decentralized formulations of sequential detection. In practice, the configurations of the decentralized system varies according to the requirement of the applications, and therefore lead to different formulates of the decentralized sequential detection problem. We start with a general form of the decentralized system model, and then discuss possible modifications to it.

As in Figure 1.1, multiple sensing devices work cooperatively to finish the detection task. Local sensors S_1 and S_2 take observations individually and transmit sensing reports to a central unit, fusion center F, for decision making. During the detection progress, local sensors may exchange information to improve the detection performance, and the fusion center is able to send feedback to local sensors to adjust local sensing policies. In general, the fusion center makes the final decision based only on the sensing reports transmitted from local sensors. The reported information is received sequentially, based on which a sequential test can be performed at the fusion center.

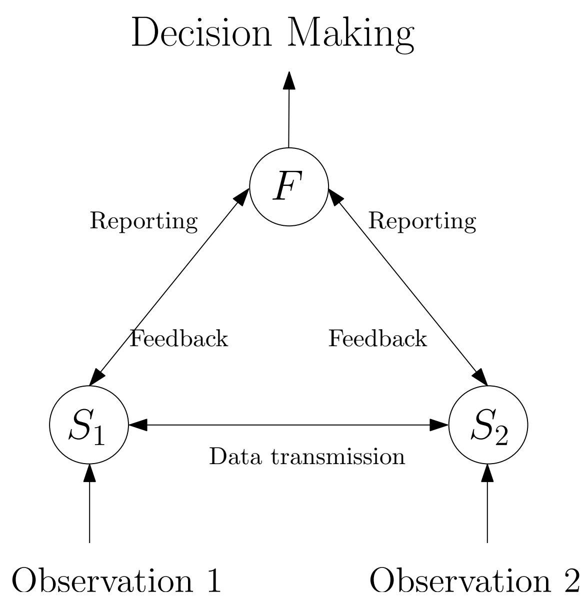

Figure 2.1: Decentralized sensing system.

In some applications, the fusion center may take observations itself, and combine these observations with the reported information to make a final decision [54]. It is also possible that the role of the fusion center being replaced by local sensors. That is, each sensor takes both local observations and sensing reports from other sensors. Any sensor in the system is able to make a finial decision, as long as its stopping condition is satisfied [55]-[57].

Two important functionalities are usually discussed in the design of decentralized sensing systems: local sensor memory and feedback. If local memory is available at each sensor, the observations taken during the detection task can all be recorded, and the sensing reports generated by a local sensor is based on the observation taken at the current time slot and all past observations recorded in the memory. The feedback from the fusion center provides a local sensor with the reported information of other local sensors. A sensing system has full feedback if local sensors are able to obtain the complete reporting histories of each other through fusion center feedback. Let Υ^m be the local decision policy and $\mathfrak{J}_k^m$ denote the reported information at time k from the mth local sensor. We have the following five different formulations of the decentralized sensing system.

(i) No local memory, no feedback [58].

Local decision making is based only on the currently observed sample, i.e.,

$$\mathfrak{J}_k^m = \Upsilon_k^m(X_k^m). \tag{2.23}$$

It is shown that the person-by-person optimal sensor decision functions are likelihood ratio tests. A set of decision functions is said to be person-by-person optimal if it is not possible to improve the corresponding team performance by unilaterally changing any one of the decision functions [31].

(ii) With local memory, no feedback [59].

$$\mathfrak{J}_k^m = \Upsilon_k^m(X_1^m, ..., X_k^m). \tag{2.24}$$

No optimum solution is proved.

(iii) No feedback, local memory restricted to past decisions [60].

$$\mathfrak{J}_k^m = \Upsilon_k^m(X_1^m, \mathfrak{J}_1^m, ..., \mathfrak{J}_{k-1}^m). \tag{2.25}$$

It is shown in [60] that likelihood ratio tests are optimal.

(iv) Full feedback, full local memory.

$$\mathfrak{I}_k^m = \Upsilon_k^m(X_{[1,k]}^m, \mathfrak{I}_{[1,k-1]}^m, ..., \mathfrak{I}_{[1,k-1]}^m). \tag{2.26}$$

No optimum solution is proved.

(v) Full feedback, local memory restricted to past decisions [31].

$$\mathfrak{I}_k^m = \Upsilon_k^m(X_k^m, \mathfrak{I}_{[1,k-1]}^m, ..., \mathfrak{I}_{[1,k-1]}^m). \tag{2.27}$$

Optimum solution is given in [31] using dynamic programming arguments [61]-[62].

We are interested in the first case where local sensors are memoryless and without feedback from the fusion center, since this is the most practical model which can be implemented with simple sensing devices. Moreover, the transmission of sensing reports between local sensors and the fusion center is costly in some applications. This transmission not only introduces extra overhead but also results in considerable local sensor energy consumption, especially when sensors are distantly located. Thus, in addition to the error probability and detection delay, we would like to control such cost as well. In this thesis, communication-efficient sequential detection schemes are developed in various contexts.

Chapter 3

Communication-efficient decentralized hypothesis testing

3.1 Introduction

In this chapter, we investigate the decentralized hypothesis testing problem. We first introduce the system model, where sensing devices are energy constrained, memoryless, and without full feedback as discussed in the previous chapter. Since sensing devices are geographically separated and are usually located at different distances from the fusion center, the transmission costs of different local sensors are modeled as non-identical. An optimization problem is then formulated which minimizes the overall detection delay with constraints on both error probabilities and the communication cost. An asymptotically person-by-person optimal algorithm is developed to solve the proposed optimization problem, and asymptotic optimality is established. In order to evaluate the performance of the proposed algorithm, we further investigate its asymptotic relative efficiency (ARE) with respect to the centralized detection algorithm, corresponding to the ideal scenario of zero transmission cost.

3.2 System model

We consider a distributed sensing system with M geographically separated sensors, $S_1, ..., S_M$, and fusion center, F, that comprise the sensing system shown in Figure 1. The fusion center can only gain access to the data of the sensed target through each local sensor. Each sensor takes observations sequentially and forwards sensing reports to the fusion center where the final decision is made. Let X_k denote the observation obtained by a certain sensor at time slot k. We consider a binary hypothesis problem where X_k has either the common probability density function (PDF) f_0 or f_1, i.e.,

$$\begin{aligned} H_0 &: X_1,\ X_2, ... \sim f_0, \\ H_1 &: X_1,\ X_2, ... \sim f_1, \end{aligned} \tag{3.1}$$

with known prior probabilities

$$P(H_0) = p_0 \text{ and } P(H_1) = 1 - p_0. \tag{3.2}$$

The cost to transmit a single sensing report from sensor m to the fusion center is denoted by w^m, $m = 1, 2, ..., M$. This assumed cost structure may represent local energy consumption and transmission overhead. The value of such cost controls the relative importance of different control links, and encourages more transmissions on link(s) with low cost. Thus, we can also relate w^m, $m = 1, ..., M$ to the network layer reliability of the sensing system, e.g., packet loss probability.

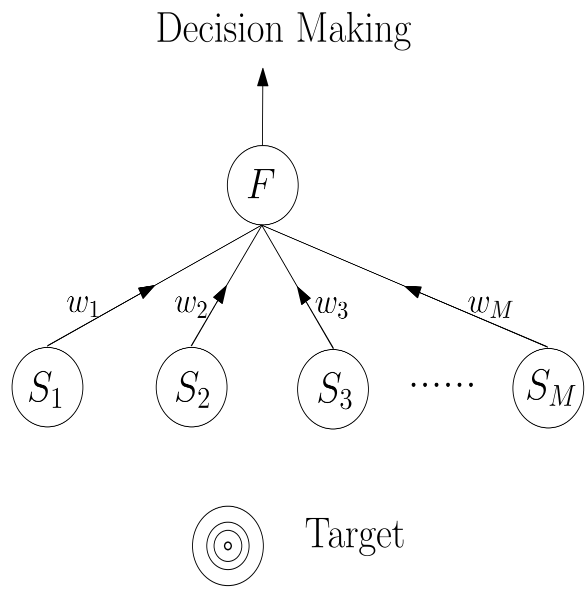

Figure 3.1: Detection system with non-identical control links.

Local sensors are able to communicate with the fusion center through a control link, but communications among local sensors are not permitted. We assume that the communication channel between sensors and fusion center is error-free and extremely limited two-way communication is possible: the fusion center only provides local sensors simple acknowledgments to stop sensing when it is able to make a decision; otherwise, there is no feedback from the fusion center. The observations across sensors are independent and identically distributed conditioned on the hypothesis H_0 or H_1. This assumption reflects the fact that sensor noise is local.

3.3 Problem formulation

After taking each observation, a local sensor decides whether it is worth sending a report to the fusion center. Let $T_k \in \{0, 1\}$ denote the reporting decision indicator of a certain sensor at time k, defined as follows:

$$T_k = \eta_k(g_{k-1}, X_k), \quad k = 1, 2... \tag{3.3}$$

where g_{k-1} is the cumulative statistic based on observations before time k, i.e., $\{X_1, ..., X_{k-1}\}$, and η_k is the local decision policy at time k. If $T_k = 1$, the local sensor forwards a local sensor report to the fusion center. If $T_k = 0$, the local sensor does not send a report but continues taking observations. Reported information by the local sensor is the accumulated statistic at time k and takes the form

$$r_k = \begin{cases} 0, & T_k = 0, \\ g_k, & T_k = 1, \end{cases} \tag{3.4}$$

where $r_k = 0$ means that no report is transmitted from sensor m to the fusion center. After transmitting a sensing report to the fusion center, each local sensor test statistic is reinitialized to zero. It is not necessary for a local sensor to draw a conclusion on a hypothesis. Only if a local sensor makes a decision in favor of either hypothesis does it forward a report to the fusion center.

At the fusion center, sensing reports from local sensors are received sequentially, and comprise the observations at the fusion center. The cumulative statistic at time slot k at the fusion center is updated based on reported sensed information via

$$u_k = \psi(u_{k-1}, \underline{r_k}), \quad k = 1, 2..., \tag{3.5}$$

where ψ represents the update function. $\underline{r_k} = [r_k^1, r_k^2, ..., r_k^M]$ is the M-dimensional vector of reported sensed information at the kth time slot, with components r_k^m,

$m \in \{1, ..., M\}$ representing the information reported by the m-th local sensor. Local sensors do not send reports when $r_k^m = 0$, and the fusion center can infer this information by knowing that a report is missing. Thus, the total communication cost for reports transmission received by the fusion center in (3.8) during the detection process is given by

$$R = \sum_{k=1}^{\tau} \sum_{m=1}^{M} w^m \times \mathrm{I}_{\{r_k^m \neq 0\}}, \tag{3.6}$$

where $\mathrm{I}_{\{\bullet\}}$ is the indicator function and w^m is the cost of transmitting a single report to the fusion center through sensor m. Based on the statistic u_k, the fusion center makes decisions on whether to stop and make a final decision on H_0 or H_1 or to continue receiving sensing reports, i.e.,

$$D_k = \phi(u_k), \quad n = 1, 2..., \tag{3.7}$$

where ϕ is the decision policy. $D_k \in \{0, 1\}$ denotes the stopping decision of the fusion center at time k. If $D_k = 1$, the fusion center stops and claims H_0 or H_1 when lower or upper threshold is exceeded, respectively. If $D_k = 0$, the detection task continues.

We are interested in a detection strategy that minimizes the overall expected detection delay under error probability and communication cost constraints, i.e.,

$$\begin{aligned} \underset{\eta \;\; \phi}{\text{minimize}} \qquad & E\{\tau\}, \\ \text{subject to} \qquad & P_{FAC} \leq \alpha_C, \; P_{MDC} \leq \beta_C, \\ & \text{and} \;\; E\{R\} \leq \kappa, \end{aligned} \tag{3.8}$$

where τ represents the stopping time of the fusion center, P_{FAC} and P_{MDC} denote false alarm and missed detection probabilities, constrained to α_C and β_C, respectively. R is the total transmission cost for local sensors to send reports to the fusion center during the detection task and κ is a given communication cost constraint. The proposed communication metric is able to reflect the cost of energy consumption and transmission overhead at local sensors. In contrast to other formulations, the proposed communication cost restricts channel uses rather than bandwidth of the control link. As the communication constraint increases, local sensing devices are able to send more reports to the fusion center.

In case that κ is greater than the average number of samples required by the centralized detection scheme, the communication cost constraint vanishes. Moreover, the value of κ should be set to support at least one transmission from each local sensor to the fusion center. Therefore κ should satisfy

$$\sum_{m=1}^{M} w^m < \kappa < \frac{M}{d_\theta}\left[\beta_\theta \log\frac{1-\beta_C}{\alpha_C} + (1-\beta_\theta)\log\frac{\beta_C}{1-\alpha_C}\right], \tag{3.9}$$

with

$$d_\theta = E_\theta\left\{\log\frac{f_1(X_1^1,...,X_1^M)}{f_0(X_1^1,...,X_1^M)}\right\}, \ \theta \in \{0,1\}, \tag{3.10}$$

where $f_\theta(X^1,...,X^M)$ denotes the joint PDF of local observations under H_θ, $\theta \in \{0,1\}$. The rightmost term in (3.9) represents a lower bound on the average number of samples transmitted to the fusion center for the centralized detection scheme using the optimal sequential probability ratio test (SPRT) . In the rest of this paper, the optimization problem (3.8) applies to a communication cost constraint which satisfies (3.9); otherwise the formulated problem degenerates either to the Or-rule or to the centralized scheme.

Remark 3.1. *In case that the communication cost constraint is large enough, the local error probability constraints loosen to the point where all local observations are reported to the fusion center directly, and the optimal solution to the formulated problem is the centralized scheme. That is, the local SPRT thresholds converge to each other and communication between local sensors and the fusion center is not limited.*

Remark 3.2. *If the communication constraint is at the most restricted extreme, each of the local sensors performs a SPRT with error probability constraints equal to the overall error probability constraints. The fusion center makes a final decision when the first local sensing report is received. This leads to an Or-rule scheme.*

3.4 Decentralized sensing strategy

In the sensing system, local sensors cannot communicate with one another, and the fusion center only provides minimum feedback when a final decision is made. Thus, sensing devices cannot optimize their strategies cooperatively; the best they can do is to individually optimize their strategies based on local information. Such a sub-optimal approach is referred as person-by-person optimization [31], where it is not possible to improve overall team performance by unilaterally changing any of the decision functions. More specifically, we first investigate the optimal sensing strategy at each local sensor. Following this, we develop the optimal fusion rule for fixed local sensing policies. The proposed detection strategy is based on the following assumptions:

Assumption 3.1. *The distribution of observations at all local sensors follow either f_0 or f_1. The PDFs under H_θ, $\theta \in \{0,1\}$ are known by all local sensors and by the fusion center.*

Assumption 3.2. *Each local sensor makes decisions based only on its own observations. No feedback information from the fusion center is made available to local sensors.*

Assumption 3.3. *The fusion center makes its final decision based only on the reported information from local sensors, including the absence of reporting.*

Remark 3.3 (Person-by-person optimal strategy [35]). *Under Assumptions 3.2, and 3.3, the person-by-person optimal decision strategy which minimizes the overall detection delay subject to error probabilities is to apply likelihood-ratio tests at both local sensors and fusion center. Together with Assumption 3.1, it follows that such likelihood-ratio based tests are SPRTs.*

In contrast to the work in [35], we additionally introduce communication cost constraint, which can be satisfied by adjusting the relationship between local sensor and fusion center threshold values.

3.4.1 Local sensing strategy

A final decision is made at the fusion center based on received sensing reports from local sensors. In order to minimize the overall detection delay, each sensing report ought to be generated as quickly as possible. Thus, the optimization problem local to each sensor is of the form

$$\begin{aligned} \underset{\eta}{\text{minimize}} \quad & E\{v\}, \\ \text{subject to} \quad & P_{FAL} \leq \alpha_L \text{ and } P_{MDL} \leq \beta_L, \end{aligned} \tag{3.11}$$

where v represents the delay for a local sensor to generate a sensing report. P_{FAL} and P_{MDL} are the local false alarm and missed detection probabilities, constrained to α_L and β_L, respectively.

Proposition 3.1. *Under Assumptions 3.1, 3.2, and 3.3, the SPRT is locally person-by-person optimal whose thresholds are determined from the global problem constraints.*

Proof. Consider the mth local sensor, with P^m_{FAL}, P^m_{MDL}, $1 \leq m \leq M$ incurred by the overall constraints, and suppose that the decision rules of all the other local sensors, as well as that of the fusion center are fixed. According to Assumptions 3.1 and 3.2, as well as Theorem 2.1, among all tests which achieve prescribed error probability constraints, the average delay of SPRT is minimum. In the case that sensor m applies a local test other than the SPRT, a greater local detection delay is obtained, and therefore from Assumption 3.3, the cost of the overall detection system increases. □

Thus, the optimal solution to (3.11) can be achieved by tracking the likelihood ratio statistics at each local sensor via

$$g_k = \sum_{i=1}^{k} \log \frac{f_1(X_i)}{f_0(X_i)} \tag{3.12}$$

and forming a likelihood ratio test: a local sensor forwards a report to the fusion center whenever the test statistic g_k is significant, i.e., exceeds predefined threshold. Let a local sensor decision rule at time k be given by

$$T_k = \begin{cases} 0, & g_k \in (a, b), \\ 1, & g_k \notin (a, b), \end{cases} \tag{3.13}$$

where $a < 0 < b$ are local thresholds. When $T_k = 1$ the local statistic is reported to the fusion center; while when $T_k = 0$, local sensing task continues without sending a report to the fusion center. After a sensing report is forwarded to the fusion center, the local memoryless sensor reinitiates and repeats its sequential test until the final decision at the fusion center is made. With predefined local sensing threshold values $a < 0 < b$, data transmitted to the fusion center through each sensing report is independent and identically distributed (IID) random variables conditioned on each hypothesis, i.e.,

$$\begin{aligned} H_0 &: \mathfrak{R}_1, \mathfrak{R}_2, \mathfrak{R}_3 \ldots \sim \mathcal{R}_0, \\ H_1 &: \mathfrak{R}_1, \mathfrak{R}_2, \mathfrak{R}_3 \ldots \sim \mathcal{R}_1, \end{aligned} \tag{3.14}$$

where $\mathfrak{R}_k = g_k|_{T_k \neq 0}$ and $\mathcal{R}_\theta$, $\theta \in \{0, 1\}$ is the conditional probability density function (PDF) of information contained in a sensing report under H_θ, $\theta \in \{0, 1\}$.

Let $\mathfrak{t}_n$ denote the time slot that a local sensor sends the n-th sensing report to the fusion center. We have incurred a delay to generate the n-th local sensing report, i.e., the delay of the n-th local sequential test is given by

$$\mathfrak{T}_n = \mathfrak{t}_n - \mathfrak{t}_{n-1}, \tag{3.15}$$

with $\mathfrak{t}_0 = 0$. Since a local sensor applies its local test repeatedly using fixed threshold values, $a < 0 < b$, and the test statistic is reset to zero when test thresholds are reached (when a local report is forwarded), we can straightforwardly determine that the delays $\mathfrak{T}_n$ to generate reports at any certain local sensor are IID random variables conditioned on each hypothesis, i.e.,

$$\begin{aligned} H_0 &: \mathfrak{T}_1, \mathfrak{T}_2, \mathfrak{T}_3 ... \sim \mathcal{T}_0, \\ H_1 &: \mathfrak{T}_1, \mathfrak{T}_2, \mathfrak{T}_3 ... \sim \mathcal{T}_1, \end{aligned} \tag{3.16}$$

where $\mathcal{T}_\theta$, $\theta \in \{0, 1\}$ is the conditional probability mass function (PMF) of the report-generating delay at a local sensor, which is determined by f_θ, $\theta \in \{0, 1\}$ and the values of a and b. In general, $\mathcal{T}_\theta$, $\theta \in \{0, 1\}$ is not a memoryless discrete time distribution, i.e., $\{\mathfrak{T}_k\}$ is not geometrically distributed. Let $\mathcal{G}_{\theta,n}$ denote the distribution of the local test statistic after n observations are taken. Then

$$\mathcal{G}_{\theta,n}(y) = \int_a^b \mathcal{G}_{\theta,n-1}(s) z_\theta(y-s) ds, \ n > 1, \tag{3.17}$$

$$\mathcal{T}_{\theta,u}(n) = \int_a^b \mathcal{G}_{\theta,n-1}(s)(1 - Z_\theta(b-s)) ds, \ n > 1, \tag{3.18}$$

$$\mathcal{T}_{\theta,l}(n) = \int_a^b \mathcal{G}_{\theta,n-1}(s) Z_\theta(a-s) ds, \ n > 1, \tag{3.19}$$

with

$$\mathcal{G}_{\theta,1}(y) = z_\theta(y), \ \mathcal{T}_{\theta,u}(1) = 1 - Z(b), \text{ and } \mathcal{T}_{\theta,l}(1) = Z(a), \tag{3.20}$$

where $z_\theta(s)$ and $Z_\theta(s)$ are the probability density function (PDF) and cumulative distribution function (CDF) of $\log \frac{f_1(X_1)}{f_0(X_1)}$ under H_θ, $\theta \in \{0, 1\}$, respectively. In (3.16) and (3.17), $\mathcal{T}_{\theta,u}(n)$ and $\mathcal{T}_{\theta,l}(n)$ represent the probability that the local test delay equals n time slots and the local statistic exceeds upper and lower boundaries, respectively. Thus, the probability that local report-generating delay is equal to n can be expressed as,

$$\mathcal{T}_\theta(n) = \mathcal{T}_{\theta,u}(n) + \mathcal{T}_{\theta,l}(n). \tag{3.21}$$

It is straightforward that the data contained in each sensing report can be used to imply the true hypothesis, since the statistical properties of the reported data when H_0 or H_1 is true are different. Moreover, the distribution of the report-generating delay usually varies under different hypotheses as well, and therefore can also be applied to infer the true hypothesis. Thus, the reported information by local sensors consists of the transmitted data and the reporting time, both of which ought to be utilized in the fusion center test.

3.4.2 Fusion center strategy

Following the person-by-person optimization methodology, we have so far established the optimal hypothesis testing strategy at each local sensor, independently of the fusion center and other local sensors. We now investigate the optimal solution to the fusion center test when local sensing threshold values are fixed. As is shown previously, when local detection strategies are fixed, the sensing reports received by the fusion center are time dependent. The sequence $\{\underline{r_1}, ..., \underline{r_k}\}$ can be therefore viewed as dependent observations at the fusion center. As discussed in Chapter 2, the SPRT based on dependent observations is of the form

$$D_k = \begin{cases} 0, & u_k \in (A_k, B_k), \\ 1, & u_k \notin (A_k, B_k), \end{cases} \tag{3.22}$$

with the test statistic given by

$$u_k = \log \frac{\mathcal{P}_1(\underline{r_1}, ..., \underline{r_k})}{\mathcal{P}_0(\underline{r_1}, ..., \underline{r_k})}, \tag{3.23}$$

where $\mathcal{P}_\theta$, $\theta \in \{0, 1\}$ is the joint PDF of received sensing information at the fusion center up to time k and $A_k < 0 < B_k$ are time varying threshold values. Since reported information across sensors are independent conditioned on each hypothesis, we have

$$u_k = \sum_{m=1}^{M} \log \frac{\mathcal{P}_1^m(r_1^m, r_2^m, ..., r_k^m)}{\mathcal{P}_0^m(r_1^m, r_2^m, ..., r_k^m)}, \tag{3.24}$$

where $\mathcal{P}_\theta^m$, $\theta \in \{0, 1\}$ is the joint PDF of reported information from sensor m. Let $\mathcal{N}_k^m$ denote the number of reports already sent to the fusion center by sensor m at time k. The sequence $\{r_1^m, ..., r_k^m\}$ from sensor m can be partitioned into $\mathcal{N}_k^m + 1$ subsequences, i.e., $\{r_1^m, ..., r_{t_1^m}^m\}$, $\{r_{t_1^m+1}^m, ..., r_{t_2^m}^m\}$, ..., $\{r_{t_{\mathcal{N}_k^m}+1}^m, ..., r_k^m\}$. We define the first $\mathcal{N}_k^m$ subsequences as the *data subsequences* (DSs), and the left over subsequence after removing DSs from the sequence, $\{r_1^m, ..., r_k^m\}$, as the *tail subsequence* (TS). Note that if $r_k^m = 0$, there is no TS and all subsequences are of type DS.

Since the local test statistic is reset each time a report is transmitted, the reported statistics of these subsequences are mutually independent conditioned on each hypothesis, i.e.,

$$\begin{aligned} \mathcal{P}_\theta^m(r_1^m, ..., r_k^m) &= \mathcal{P}_\theta^m(r_1^m, ..., r_{t_1^m}^m) \times \mathcal{P}_\theta^m(r_{t_1^m+1}^m, ..., r_{t_2^m}^m) \\ &\times \; ... \times \mathcal{P}_\theta^m(r_{t_{\mathcal{N}_k^m}+1}^m, ..., r_k^m). \end{aligned} \tag{3.25}$$

By definition, we notice that only the last element in each DS is non-zero, which is the information contained in the sensing report. Thus, the joint PDF of each DS in (3.25) can be expressed as

$$\begin{aligned}
\mathcal{G}_\theta^m(1) &\triangleq \mathcal{P}_\theta^m(r_1^m, ..., r_{t_1^m}^m) = P_\theta(\mathfrak{T}_1^m = t_1^m)P_\theta(\mathfrak{R}_1^m = r_{t_1^m}^m) \\
&= \mathcal{T}_\theta(\mathfrak{T} = t_1^m)\mathcal{R}_\theta(\mathfrak{R} = r_{t_1^m}^m), \\
\mathcal{G}_\theta^m(2) &\triangleq \mathcal{P}_\theta^m(r_{t_1^m+1}^m, ..., r_{t_2^m}^m) = P_\theta(\mathfrak{T}_2^m = t_2^m - t_1^m)P_\theta(\mathfrak{R}_2^m = r_{t_2^m}^m) \\
&= \mathcal{T}_\theta(\mathfrak{T} = t_2^m - t_1^m)\mathcal{R}_\theta(\mathfrak{R} = r_{t_2^m}^m), \\
&\cdots \\
\mathcal{G}_\theta^m(\mathcal{N}_k^m) &\triangleq \mathcal{P}_\theta^m(r_{t_{\mathcal{N}_k^m-1}^m+1}^m, ..., r_{t_{\mathcal{N}_k^m}^m}^m) = P_\theta(\mathfrak{T}_{\mathcal{N}_k^m}^m = t_{\mathcal{N}_k^m}^m - t_{\mathcal{N}_k^m-1}^m)P_\theta(\mathfrak{R}_{\mathcal{N}_k^m}^m = r_{t_{\mathcal{N}_k^m}^m}^m) \\
&= \mathcal{T}_\theta(\mathfrak{T} = t_{\mathcal{N}_k^m}^m - t_{\mathcal{N}_k^m-1}^m)\mathcal{R}_\theta(\mathfrak{R} = r_{t_{\mathcal{N}_k^m}^m}^m),
\end{aligned} \tag{3.26}$$

where P_θ is the probability conditioned on θ and the last equality in each term of (3.26) is obtained by using the PDFs from (3.14) and (3.16). By substituting (3.25) and (3.26) into (3.23), we obtain fusion center test statistic at time slot k as

$$u_k = \sum_{m=1}^{M} \sum_{n=1}^{\mathcal{N}_k^m} \log \frac{\mathcal{G}_1^m(n)}{\mathcal{G}_0^m(n)} + \log \frac{\mathcal{T}_1(\mathfrak{T} > k - t_{\mathcal{N}_k^m})}{\mathcal{T}_0(\mathfrak{T} > k - t_{\mathcal{N}_k^m})}. \tag{3.27}$$

From (3.26) and (3.27), one can notice that the fusion center test utilizes not only the data contained in sensing reports, but also the times when these reports are received.

The fusion center test statistic (3.27) is updated at each time slot and compared with the fusion center thresholds. Since Wald's original equations are valid with dependent observations [72], the fusion center error probability constraints can be satisfied by applying the following constant thresholds:

$$A \leq \log \frac{\beta_C}{1-\alpha_C} \text{ and } B \geq \log \frac{1-\beta_C}{\alpha_C}, \tag{3.28}$$

where $A < 0 < B$ are constants. We now prove that the proposed fusion center sequential test which computes the statistic (3.27) and compares it to the constant threshold values (3.28), is asymptotically optimal in the following sense:

Theorem 3.1. *Let $\mathscr{F}(\alpha_C, \beta_C)$ denote the class of all fusion center tests based on $\{\underline{r_1}, \underline{r_2}, ..., \underline{r_k}\}$ satisfying $P_{FAC} \leq \alpha_C$ and $P_{MDC} \leq \beta_C$, and let $(\tau, \delta) \in \mathscr{F}(\alpha_C, \beta_C)$ denote a test with stopping time τ and decision rule δ. The stopping time of fusion center sequential test with statistic* (3.27) *and thresholds* (3.28) *can be expressed as*

$$\tau_{A,B} = \inf\{k \geq 1 : u_k \leq A \text{ or } u_k \geq B\}, \tag{3.29}$$

and such a test is asymptotically optimum in the sense that for all $0 < \epsilon < 1$,

$$\lim_{\alpha_C+\beta_C \to 0} \inf_{(\tau,\delta)\in\mathscr{F}(\alpha_C,\beta_C)} P_\theta(\tau > \epsilon\tau_{A,B}) = 1, \;\; \theta \in \{0, 1\}, \tag{3.30}$$

where $\tau_{A,B}$ is the stopping time of the fusion center test with statistic u_k and thresholds $A < 0 < B$.

Proof. Let $u_{k,\theta}$ denote the fusion center test statistic at time k when H_θ, $\theta \in \{0, 1\}$, is true. With predefined local test thresholds, $\mathcal{G}_\theta^m(n)$, $n = 1, 2, ...,$ defined in (3.26) have common mean and variance, and by the Weak Law of Large Numbers, the fusion center statistic, u_k, is governed by

$$\frac{u_{k,\theta}}{k} \xrightarrow{i.p.} \lambda_\theta, \theta \in \{0,1\}, \text{ when } k \to \infty. \tag{3.31}$$

where $\lambda_0 < 0 < \lambda_1$ are constants. Let $\epsilon' > 1$ such that $\epsilon\epsilon' < 1$, and τ_N is the greatest integer which satisfies

$$\tau_N \leq \epsilon \min\{|\log \alpha_C/\lambda_1|, |\log \beta_C/\lambda_0|\}. \tag{3.32}$$

The reported information $\{\underline{r_1}, \underline{r_2}, ..., \underline{r_k}\}$ can be viewed as a sequence of random variables defined on the same underlying measurable space $(\Omega, \mathscr{F})$. Let P denote the probability measure on $(\Omega, \mathscr{F})$. For any $(\tau, \delta) \in \mathscr{F}(\alpha_C, \beta_C)$, we have

$$\begin{aligned}
\alpha_C &= \int_{\tau<\infty,\ (\tau,\delta) \text{ rejects } H_0} \exp(-u_{\tau,0})dP, \\
&\geq \int_{\tau\leq\tau_N,\ u_{\tau,0}\leq\epsilon'\lambda_1\tau_N,\ (\tau,\delta) \text{ rejects } H_0} \exp(-u_{\tau,0})dP, \\
&\geq \exp(-\epsilon'\lambda_1\tau_N)P(\tau \leq \tau_N, u_{\tau,0} \leq \epsilon'\lambda_1\tau_N,\ (\tau,\delta) \text{ rejects } H_0).
\end{aligned} \tag{3.33}$$

From (3.32) we have $\epsilon'\lambda_1\tau_N \leq \epsilon\epsilon'|\log\alpha_C|$. As $\log\alpha_C < 0$, we can further have $-\epsilon'\lambda_1\tau_N \geq \epsilon\epsilon'\log\alpha_C$, which by substituting into (3.33), one can further obtain

$$\begin{aligned}
\alpha_C &\geq \exp(\epsilon\epsilon'\log\alpha_C)P(\tau \leq \tau_N, u_{\tau,0} \leq \epsilon'\lambda_1\tau_N,\ (\tau,\delta) \text{ rejects } H_0), \\
&= \alpha_C^{\epsilon\epsilon'}P(\tau \leq \tau_N, u_{\tau,0} \leq \epsilon'\lambda_1\tau_N,\ (\tau,\delta) \text{ rejects } H_0),
\end{aligned} \tag{3.34}$$

from which follows

$$\frac{\alpha_C}{\alpha_C^{\epsilon\epsilon'}} = \alpha_C^{1-\epsilon\epsilon'} \geq P(\tau \leq \tau_N, u_{\tau,0} \leq \epsilon'\lambda_1\tau_N,\ (\tau,\delta) \text{ rejects } H_0). \tag{3.35}$$

By adding $P(\tau \leq \tau_N, u_{\tau,0} > \epsilon'\lambda_1\tau_N, \text{ rejects } H_0)$ to both sides of (3.35), we have

$$\begin{aligned} P(\tau \leq \tau_N, (\tau,\delta) \text{ rejects } H_0) &\leq \alpha_C^{1-\epsilon\epsilon'} + P(\tau \leq \tau_N, u_{\tau,0} > \epsilon'\lambda_1\tau_N, \text{ rejects } H_0), \\ &\leq \alpha_C^{1-\epsilon\epsilon'} + P(\tau \leq \tau_N, u_{\tau,0} > \epsilon'\lambda_1\tau_N), \\ &\leq \alpha_C^{1-\epsilon\epsilon'} + P(\max_{j\leq\tau_N} u_{j,0} > \epsilon'\lambda_1\tau_N). \end{aligned} \tag{3.36}$$

Using a similar argument, we can also obtain that

$$P(\tau \leq \tau_N, (\tau,\delta) \text{ rejects } H_1) \leq \beta_C^{1-\epsilon\epsilon'} + P(\max_{j\leq\tau_N} u_{j,1} > \epsilon'\lambda_0\tau_N). \tag{3.37}$$

By combining (3.36) and (3.37), which are mutually exclusive, it follows that

$$\begin{aligned} \sup_{(\tau,\delta)\in\mathscr{F}(\alpha_C,\beta_C)} P(\tau &\leq \tau_N) \leq \alpha_C^{1-\epsilon\epsilon'} + \beta_C^{1-\epsilon\epsilon'} + P(\max_{j\leq\tau_N} u_{j,0} > \epsilon'\lambda_1\tau_N) \\ &+ P(\max_{j\leq\tau_N} u_{j,1} > \epsilon'\lambda_0\tau_N). \end{aligned} \tag{3.38}$$

As $\alpha_C + \beta_C \to 0$, we have $\tau_N \to \infty$ and therefore obtain $u_{\tau_N,0} \to \lambda_0\tau_N$ and $u_{\tau_N,1} \to \lambda_1\tau_N$ from (3.31), which by substituting into (3.38), we obtain

$$\sup_{(\tau,\delta)\in\mathscr{F}(\alpha_C,\beta_C)} P(\tau \leq \tau_N) \to 0. \tag{3.39}$$

From (3.32) and (3.39), we have that as $\alpha_C + \beta_C \to 0$,

$$\inf_{(\tau,\delta)\in\mathscr{F}(\alpha_C,\beta_C)} P(\tau > \epsilon \min\{|\log\alpha_C/\lambda_1|, |\log\beta_C/\lambda_0|\}) \to 1. \tag{3.40}$$

Moreover, as $\alpha_C + \beta_C \to 0$, the fusion center thresholds (3.28) converge to $A = \log\beta_C$ and $B = \log 1/\alpha_C$. Thus, together with (3.31) we also have that $\tau_{A,B}$ converges to

$\min\{|\log \alpha_C/\lambda_1|, |\log \beta_C/\lambda_0|\}$ in probability, i.e.,

$$\frac{\tau_{A,B}}{\min\{|\log \alpha_C/\lambda_1|, |\log \beta_C/\lambda_0|\}} \xrightarrow{i.p.} 1, \text{ when } \alpha_C + \beta_C \to 0, \tag{3.41}$$

and (3.30) follows from (3.40) and (3.41). □

Similar arguments to the proof of Theorem 3.1 also appear in a different context in [73]. However, the proof in [73] that the generalized SPRT with constant thresholds (3.28) is asymptotically optimal assumes stability of the test statistic. In Theorem 3.1, stability of the test statistic is established.

3.4.3 Determination of local thresholds

With constant fusion center threshold values $A < 0 < B$, we can further investigate the detection delay of the sensing system. Let u_τ denote the fusion center statistic when a final decision is made. We have from (3.27), the expected fusion center test statistic under H_θ

$$\begin{aligned}
E_\theta\{u_\tau\} &= E_\theta\left[\sum_{m=1}^{M}\sum_{n=1}^{\mathcal{N}_\tau^m}\log\frac{\mathcal{G}_1^m(n)}{\mathcal{G}_0^m(n)} + \log\frac{\mathcal{T}_1(\mathfrak{T} > \tau - t_{\mathcal{N}_\tau^m})}{\mathcal{T}_0(\mathfrak{T} > \tau - t_{\mathcal{N}_\tau^m})}\right], \\
&= \sum_{m=1}^{M} E_\theta\{\mathcal{N}_\tau^m\}E_\theta\{\log\frac{\mathcal{G}_1^m(1)}{\mathcal{G}_0^m(1)}\} + E_\theta\{\log\frac{\mathcal{T}_1^m(\mathfrak{T} > \tau - t_{\mathcal{N}_\tau^m})}{\mathcal{T}_0^m(\mathfrak{T} > \tau - t_{\mathcal{N}_\tau^m})}\}, \\
&= \sum_{m=1}^{M} E_\theta\{\mathcal{N}_\tau^m\}\left(E_\theta\{\log\frac{\mathcal{T}_1^m(\mathfrak{T} = t_1^m)}{\mathcal{T}_0^m(\mathfrak{T} = t_1^m)}\} + E_\theta\{\log\frac{\mathcal{R}_1^m(\mathfrak{R} = r_{t_1^m}^m)}{\mathcal{R}_0^m(\mathfrak{R} = r_{t_1^m}^m)}\}\right), \\
&+ E_\theta\{\log\frac{\mathcal{T}_1^m(\mathfrak{T} > \tau - t_{\mathcal{N}_\tau^m})}{\mathcal{T}_0^m(\mathfrak{T} > \tau - t_{\mathcal{N}_\tau^m})}\}, \\
&= \sum_{m=1}^{M} E_\theta\{\mathcal{N}_\tau^m\}\mathfrak{C}_\theta^m + \mathfrak{D}_\theta^m,\ \ \theta \in \{0,1\},\ \text{ with} \\
\mathfrak{C}_\theta^m &\triangleq \mathfrak{C}_{\theta,t}^m + \mathfrak{C}_{\theta,d}^m, \\
\mathfrak{C}_{\theta,t}^m &\triangleq \sum_{n=1}^{\infty}\mathcal{T}_\theta^m(n)\log\frac{\mathcal{T}_1^m(n)}{\mathcal{T}_0^m(n)}, \\
\mathfrak{C}_{\theta,d} &\triangleq \int_{-\infty}^{\infty}\mathcal{R}_\theta^m(y)\log\frac{\mathcal{R}_1^m(y)}{\mathcal{R}_0^m(y)}dy, \\
\mathfrak{D}_\theta^m &\triangleq \sum_{n=1}^{\infty}\mathcal{T}_\theta^m(\mathfrak{T} = n)\log\frac{\mathcal{T}_1^m(\mathfrak{T} > n)}{\mathcal{T}_0^m(\mathfrak{T} > n)},
\end{aligned} \tag{3.42}$$

where $\mathfrak{C}_\theta^m$ is the expected contribution of likelihood ratio to the fusion center test statistic after a DS is received from sensor m. The contribution of $\mathfrak{C}_{\theta,t}^m$ is attributed to the time dependency of the reporting time, while the contribution of $\mathfrak{C}_{\theta,d}^m$ is due to the data contained in the sensing report. $\mathfrak{D}_\theta^m$ represents the expected contribution of the TS of sensor m. Let $E_\theta\{\tau\}$ denote the expected delay of the fusion center test under H_θ. We have

$$E_\theta\{\tau\} = E_\theta\{v^m\}E_\theta\{\mathcal{N}_\tau^m\}. \tag{3.43}$$

We further assume that when the fusion center test stops, the fusion center test statistic reaches the threshold values with equality. Then,

$$E_\theta\{u_\tau\} \;=\; (1-\beta_\theta)A + \beta_\theta B, \tag{3.44}$$

where $\beta_0 = \alpha_C$ and $\beta_1 = 1 - \beta_C$. By substituting (3.43) and (3.44) into (3.42), we can further have the expected delay of the fusion center test expressed as

$$E_\theta\{\tau\} = \left(E_\theta\{u_\tau\} - \sum_{m=1}^{M} \mathfrak{D}_\theta^m\right)\left(\sum_{m=1}^{M} \frac{\mathfrak{C}_\theta^m}{E_\theta\{v^m\}}\right)^{-1}, \;\theta \in \{0,1\}. \tag{3.45}$$

The total communication cost constraint can be then expressed as

$$E\{R\} = \sum_{m=1}^{M} E_\theta\{\mathcal{N}_\tau^m\} w^m = \sum_{m=1}^{M} \frac{E_\theta\{\tau\}}{E_\theta\{v^m\}} w^m \leq \kappa. \tag{3.46}$$

By substituting (3.45) into (3.46), we have

$$\begin{aligned}\sum_{m=1}^{M} \frac{w^m}{E_\theta\{v^m\}} &\leq \kappa \left(E_\theta\{u_\tau\} - \sum_{m=1}^{M} \mathfrak{D}_\theta^m\right)^{-1} \sum_{m=1}^{M} \frac{\mathfrak{C}_\theta^m}{E_\theta\{v^m\}}, \\ &\leq \frac{\kappa}{E_\theta\{u_\tau\}} \sum_{m=1}^{M} \frac{\mathfrak{C}_\theta^m}{E_\theta\{v^m\}}.\end{aligned} \tag{3.47}$$

When H_1 is true, by definition we have $0 < \mathfrak{D}_1^m << \mathfrak{C}_1^m$, and the detection delay $E_1\{\tau\}$ decreases when $\sum_{m=1}^{M} \mathfrak{C}_1^m / E_1\{v^m\}$ increases. Combining this property with

(3.47), the target optimization problem expressed as

$$\begin{aligned} \underset{\alpha_L^m \ \beta_L^m}{\text{maximize}} \qquad & \sum_{m=1}^{M} \frac{\mathfrak{C}_1^m}{E_1\{v^m\}}, \\ \text{subject to} \qquad & \sum_{m=1}^{M} \frac{w^m}{E_1\{v^m\}} \leq \frac{\kappa}{E_1\{u_\tau\}} \sum_{m=1}^{M} \frac{\mathfrak{C}_1^m}{E_1\{v^m\}}. \end{aligned} \tag{3.48}$$

As local sensors cannot communicate one another, and local detection strategies are optimized individually, the local threshold values of each local sensor are set to satisfy

$$\begin{aligned} \underset{\alpha_L^m \ \beta_L^m}{\text{maximize}} \qquad & \frac{\mathfrak{C}_1^m}{E_1\{v^m\}}, \\ \text{subject to} \qquad & \frac{w^m}{E_1\{v^m\}} \leq \frac{\kappa}{E_1\{u_\tau\}} \frac{\mathfrak{C}_1^m}{E_1\{v^m\}}. \end{aligned} \tag{3.49}$$

Following a similar argument when H_0 is true and noting that the $\mathfrak{C}_0^m$ term defined in (3.42) is negative, we obtain the optimization problem

$$\begin{aligned} \underset{\alpha_L^m \ \beta_L^m}{\text{maximize}} \qquad & -\frac{\mathfrak{C}_0^m}{E_0\{v^m\}}, \\ \text{subject to} \qquad & \frac{w^m}{E_0\{v^m\}} \leq \frac{\kappa}{E_0\{u_\tau\}} \frac{\mathfrak{C}_0^m}{E_0\{v^m\}}. \end{aligned} \tag{3.50}$$

By combining (3.49) and (3.50), and substituting (3.44), we have

$$\begin{aligned} \underset{\alpha_L^m \ \beta_L^m}{\text{maximize}} \qquad & \left| \frac{\mathfrak{C}_\theta^m}{E_\theta\{v^m\}} \right|, \\ \text{subject to} \qquad & \frac{w^m}{E_\theta\{v^m\}} \leq \frac{\kappa}{(1-\beta_\theta)A + \beta_\theta B} \frac{\mathfrak{C}_\theta^m}{E_\theta\{v^m\}}. \end{aligned} \tag{3.51}$$

The optimum threshold values which solve (3.51) can be found with the help of the following:

Lemma 3.1. *$\mathfrak{C}_\theta$ and $E_\theta\{v\}$, $\theta \in \{0,1\}$, are monotone functions of each of a_L and β_L. As α_L (or β_L) decreases, $|\mathfrak{C}_\theta|$ increases, while $|\mathfrak{C}_\theta/E_\theta\{v\}|$ decreases.*

Proof. $\mathfrak{C}_\theta$ and $E_\theta\{v\}$ can be determined by local threshold values $a < 0 < b$. Since a and b can be expressed by α_L and β_L using

$$a \leq \log \frac{\beta_L}{1-\alpha_L} \text{ and } b \geq \log \frac{1-\beta_L}{\alpha_L}, \tag{3.52}$$

$\mathfrak{C}_\theta$ and $E_\theta\{v\}$ can be expressed as functions of α_L and β_L.

Let $\mathfrak{a}_\theta$ and $\mathfrak{b}_\theta$ denote the conditional PDF of transmitted data when local statistic reaches the lower and upper boundary, respectively. We have

$$\begin{aligned} \mathcal{R}_0 &= \alpha_L \mathfrak{b}_0 + (1-\alpha_L)\mathfrak{a}_0, \\ \mathcal{R}_1 &= \beta_L \mathfrak{a}_1 + (1-\beta_L)\mathfrak{b}_1. \end{aligned} \tag{3.53}$$

By substituting (3.53) into (3.42), we have

$$\begin{aligned} \mathfrak{C}_{0,d} &= \int_{-\infty}^{\infty} [\alpha_L \mathfrak{b}_0(y) + (1-\alpha_L)\mathfrak{a}_0(y)] \log \frac{\beta_L \mathfrak{a}_1(y) + (1-\beta_L)\mathfrak{b}_1(y)}{\alpha_L \mathfrak{b}_0(y) + (1-\alpha_L)\mathfrak{a}_0(y)} dy, \\ &= \int_{-\infty}^{\infty} [\alpha_L \mathfrak{b}_0(y) \log(1-\beta_L)\mathfrak{b}_1(y) + (1-\alpha_L)\mathfrak{a}_0(y) \log \beta_L \mathfrak{a}_1(y) \\ &- \alpha_L \mathfrak{b}_0(y) \log \alpha_L \mathfrak{b}_0(y) - (1-\alpha_L)\mathfrak{a}_0(y) \log(1-\alpha_L)\mathfrak{a}_0(y)] dy, \\ &= \alpha_L \log \frac{1-\beta_L}{\alpha_L} + (1-\alpha_L) \log \frac{\beta_L}{1-\alpha_L} + (1-\alpha_L)\mathfrak{A}_0 + \alpha_L \mathfrak{B}_0, \end{aligned} \tag{3.54}$$

with

$$\mathfrak{A}_0 = \int_{-\infty}^{a} \mathfrak{a}_0(y) \log \frac{\mathfrak{a}_1(y)}{\mathfrak{a}_0(y)} dy \text{ and } \mathfrak{B}_0 = \int_{b}^{\infty} \mathfrak{b}_0(y) \log \frac{\mathfrak{b}_1(y)}{\mathfrak{b}_0(y)} dy, \tag{3.55}$$

where the second equality in (3.54) is based on the fact that $\mathfrak{a}_\theta(y) = 0$ when $y > a$

and $\mathfrak{b}_\theta(y) = 0$ when $y < b$. As the quantity $\mathfrak{C}_{0,d}$ is dominated by the first two terms in the last equality of (3.54), it is straightforward to obtain that $|\mathfrak{C}_{0,d}|$ increases with decreasing a_L and β_L. Moreover, as $E_0\{v\}$ is lower bounded by

$$E_0\{v\} \geq \frac{1}{c_0}\left[\alpha_L \log\frac{1-\beta_L}{\alpha_L} + (1-\alpha_L)\log\frac{\beta_L}{1-\alpha_L}\right], \tag{3.56}$$

where $c_0 = E_0\{\log\frac{f_1(X_1)}{f_0(X_1)}\}$. With (3.54) and (3.56), we obtain

$$\left|\frac{\mathfrak{C}_{0,d}}{E_0\{v\}}\right| \geq -c_0\left|1 + \frac{(1-\alpha_L)\mathfrak{A}_0 + \alpha_L\mathfrak{B}_0}{E_0\{v\}}\right|. \tag{3.57}$$

Thus, when α_L and β_L decrease, $E_0\{v\}$ increases, and therefore decreasing $|\mathfrak{C}_{0,d}/E_\theta\{v\}|$. Using a similar argument, when α_L and β_L decrease, $|\mathfrak{C}_{1,d}|$ increases and $|\mathfrak{C}_{1,d}/E_\theta\{v\}|$ decreases. Lemma 3.1 then follows from the fact that $\mathfrak{C}_\theta$ is dominated by $\mathfrak{C}_{\theta,d}$. □

Theorem 3.2. *The optimal local threshold values a^m and b^m, $m = 1, ..., M$ satisfy*

$$a^m \leq \log\frac{\beta_L^m}{1-\alpha_L^m} \text{ and } b^m \geq \log\frac{1-\beta_L^m}{\alpha_L^m}, \tag{3.58}$$

and solve (3.51) *by satisfying*

$$\mathfrak{C}_\theta^m = [(1-\beta_\theta)A + \beta_\theta B]\, w^m \kappa^{-1}, \theta \in \{0,1\}, \tag{3.59}$$

with $\beta_0 = \alpha_C$ and $\beta_1 = 1 - \beta_C$.

Proof. From Lemma 3.1, $|\mathfrak{C}_\theta^m/E_\theta\{v^m\}|$ increases with α_L^m and β_L^m, and the maximum value of $|\mathfrak{C}_\theta^m/E_\theta\{v^m\}|$ is obtained when α_L^m and β_L^m are maximized. Since $|\mathfrak{C}_\theta^m|$ decreases with α_L^m and β_L^m, the maximum values of α_L^m and β_L^m are obtained when (3.59) is satisfied with equality. □

Summarizing the threshold determination procedure, the local thresholds a^m and b^m are calculated by (3.58), where α_L^m and β_L^m are obtained by solving M sets of nonlinear equations (3.59), each in the two variables α_L^m and β_L^m. $\mathfrak{C}_\theta$ is calculated from (3.42), where the distributions of report-generating delay, $\mathcal{T}_\theta$, and reported data, $\mathcal{R}_\theta$, are required. $\mathcal{T}_\theta$ can be obtained from either recursively calculating the convolution in (3.17) or applying off-line Monte Carlo experiments. $\mathcal{R}_\theta$ can also be obtained from off-line Monte Carlo experiments. The proposed decentralized hypothesis testing procedure is summarized by Algorithm 1.

Algorithm 1 Communication-efficient decentralized detection.

procedure
  $g^m \leftarrow 0,\ m = 1, 2, ..., M.$
  while $u \in (A, B)$ **do**
    $r^m \leftarrow 0,\ m = 1, 2, ..., M.$
    for $m = 1$ to M **do** ▷ Local sensing strategy.
      if $g_k^m \notin (a^m, b^m)$ **then** $r_k^m \leftarrow g_k^m,\ g_k^m \leftarrow 0.$
      else $g^m \leftarrow g^m + \log \frac{f_1^m(X_k^m)}{f_0^m(X_k^m)}.$
      end if
    end for
    $\mathfrak{D}^m \leftarrow 0,\ m = 1, 2, ...M.$
    for $m = 1$ to M **do** ▷ Fusion center strategy.
      if $r^m \neq 0$ **then** $\mathfrak{C}^m \leftarrow \mathfrak{C}^m + \log \frac{\mathcal{R}_1^m(r^m)}{\mathcal{R}_0^m(r^m)} + \log \frac{\mathcal{T}_1^m(t^m)}{\mathcal{T}_0^m(t^m)}.$
      else $t^m \leftarrow t^m + 1,\ \mathfrak{D}^m \leftarrow \mathfrak{D}^m + \log \frac{\mathcal{T}_1^m(t>t^m)}{\mathcal{T}_0^m(t>t^m)}.$
      end if
    end for
    $u \leftarrow \mathfrak{C}^m + \mathfrak{D}^m.$
    $k \leftarrow k + 1.$
  end while
  if $u \geq B$ **then** H_1 is true., ▷ Decision making.
  else H_0 is true.
  end if
end procedure

3.5 Performance analysis

In the proposed formulation, the number of communications between local sensors and the fusion center is constrained. Unlike a centralized detection scheme, where observed information is directly sent to the fusion center, the final decision of the proposed algorithm is made based on a limited number of sensing reports. As introduced in Chapter 2, we investigate the delay performance of the proposed algorithm through the asymptotic relative efficiency (ARE) [67]. The ARE of the proposed detection scheme with respect to the centralized scheme can be expressed as

$$\mathrm{ARE}_\theta(\alpha_C, \beta_C, \kappa) = \lim_{H_1 \to H_0} \frac{E_\theta^C\{\tau\}}{E_\theta\{\tau\}},\ \theta \in \{0,1\}, \tag{3.60}$$

where $E_\theta^C\{\tau\}$ denotes the average detection delay of the sequential probability ratio test, which is the delay-optimum centralized scheme,

$$E_\theta^C\{\tau\} = \frac{1}{d_\theta}\left[\beta_\theta \log\frac{1-\beta_C}{\alpha_C} + (1-\beta_\theta)\log\frac{\beta_C}{1-\alpha_C}\right], \tag{3.61}$$

with

$$d_\theta = E_\theta\{\log\frac{f_1(X_1^1, ..., X_1^M)}{f_0(X_1^1, ..., X_1^M)}\},\ \theta \in \{0,1\}, \tag{3.62}$$

where $f_\theta(X^1, ..., X^M)$ denotes the joint PDF of local observations under H_θ, $\theta \in \{0,1\}$. Since local observations across sensors are independent conditioned on each hypothesis,

$$d_\theta = E_\theta\{\log\frac{f_1^1(X_1^1)}{f_0^1(X_1^1)} + ... + \frac{f_1^M(X_1^M)}{f_0^M(X_1^M)}\} = \sum_{m=1}^{M} d_\theta^m. \tag{3.63}$$

It should be noted that as H_1 approaches H_0, the detection delay of the local test approaches infinity, and the excess value over test thresholds vanishes. Therefore the use of (3.61) in the calculation of (3.60) does not rely on the assumption that excess over threshold values are negligible. By substituting (3.59) into (3.45), we can further express the detection delay of the proposed algorithm

$$E_\theta\{\tau\} = \kappa \left(\sum_{m=1}^{M} \frac{w^m}{E\{v^m\}} \right)^{-1} \tag{3.64}$$

By substituting (3.61) and (3.64) into (3.60), we obtain

$$\begin{aligned} \mathrm{ARE}_\theta(\alpha_C, \beta_C, \kappa) = \kappa^{-1} d_\theta^{-1} & \left[\beta_\theta \log \tfrac{1-\beta_C}{\alpha_C} + (1-\beta_\theta) \log \tfrac{\beta_C}{1-\alpha_C} \right] \\ \times \textstyle\sum_{m=1}^{M} w^m d_\theta^m & \left[\alpha_\theta^m \log \tfrac{1-\beta_L^m}{\alpha_L^m} + (1-\alpha_\theta^m) \log \tfrac{\beta_L^m}{1-\alpha_L^m} \right]^{-1}, \end{aligned} \tag{3.65}$$

where $\alpha_0^m = \alpha_L^m$ and $\alpha_1^m = 1-\beta_L^m$. An alternative ARE analysis method is introduced by Chernoff [68], when error probabilities tend to zero, i.e., $\alpha_C + \beta_C \to 0$. However, as local false alarm and missed detection probabilities of the proposed algorithm are usually not related in closed form, we may not able to simplify the ARE expression to enable analysis of this problem.

3.6 Numerical results

Simulation results were conducted to illustrate the performance tradeoffs. We consider observations at each local sensor to be IID Gaussian sequences with mean 0 and variance 1 under H_0, and mean 0.1 and variance 1 under H_1. Without much loss in generality, only a single example of H_0 vs H_1 PDFs are used throughout the results since the fusion center test threshold values are independent of these PDFs.

However local sensor threshold values are solutions to nonlinear equations that are functions of the PDFs, though the methodology is the same. Other than that, it is only the likelihood ratio test statistics that need to be modified for different examples. The prior probability that H_0 is true is $\pi_0 = 0.5$, and $\alpha_C = \beta_C$ is considered. Note that for a different setup of local observation distributions, the fusion center test threshold values stay the same, as they depend only on the predefined false alarm and missed detection constraints. However, local test threshold values may change as nonlinear equations (3.59) are changed. We assume that local sensors transmit binary information to the fusion center through sensing reports, and unit cost to transmit a single report from different sensors, i.e., $w^m = 1$, $m = 1, ..., M$. The false alarm and missed detection probabilities, average detection delay, and the distribution of report-generating delay are all calculated based on 10^5 Monte Carlo trials.

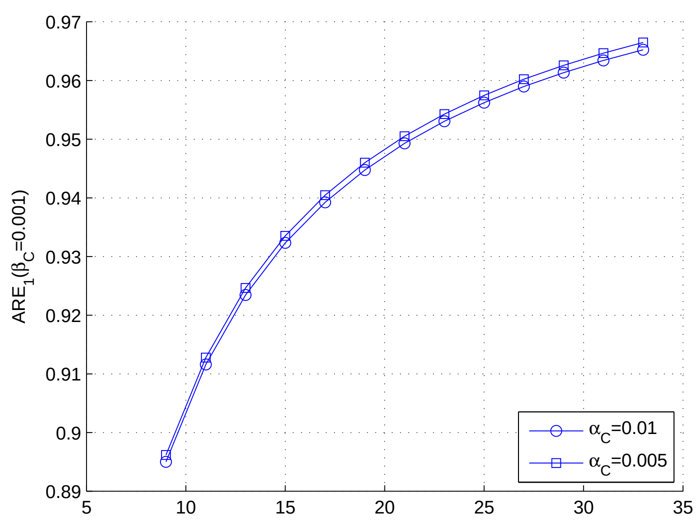

Figure 3.2: ARE_1 of Algorithm 1 with respect to the centralized scheme versus communication cost constraint, κ, with $M = 2$ local sensors.

Figure 3.2 shows ARE versus κ when H_1 is true and $\beta_C = 0.001$ when $M = 2$

local sensors are applied. The ARE of the Algorithm 1 increases as the constraint on communications is relaxed, i.e., κ is increased. This is expected since allowing more communications between local sensors and the fusion center reduces information loss during local sensing report transmission. We also investigate the ARE of Algorithm 1 when different numbers of local sensors are applied. As in Figure 3.3, the ARE of Algorithm 1 with respect to the centralized scheme is shown. One can observe that ARE is reduced as more local sensors are utilized under the same constraint, κ, i.e., more observations in the overall system need to be processed and quantized to obtain the same performance.

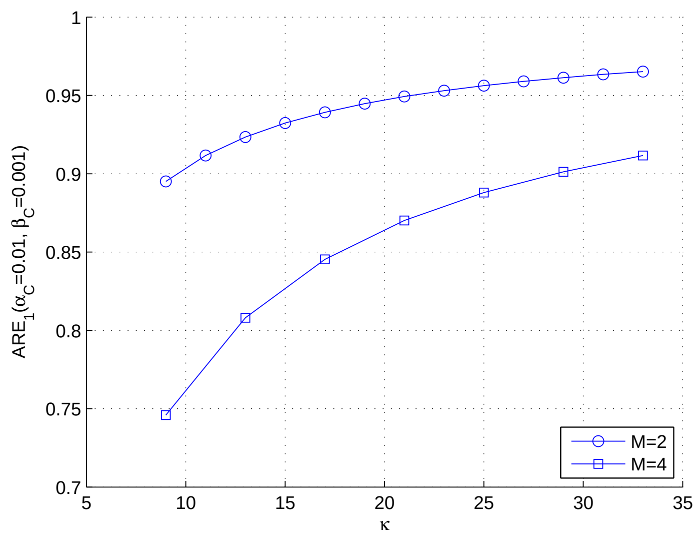

Figure 3.3: ARE_1 of Algorithm 1 with respect to the centralized scheme versus communication cost constraint, when $\alpha_C = 0.005$ and $\beta_C = 0.001$.

In Figure 3.4, the effect of false alarm probability on ARE is investigated when the number of local sensors, missed detection probability and communication cost constraints are held constant, i.e., $M = 2$, $\beta_C = 0.001$, $\kappa = 10$. As shown, the ARE of Algorithm 1 decreases as the false alarm probability increases. This is expected: when the false alarm probability at the fusion center increases, the false alarm probabilities of local sensors' sequential tests increase as well. Since the transmitted sensing reports become less reliable, the ARE of Algorithm 1 decreases.

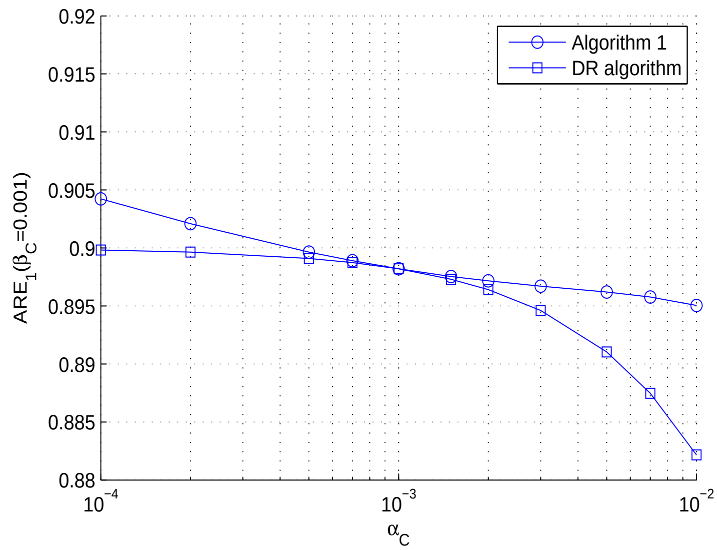

Figure 3.4: ARE_1 of Algorithm 1 with respect to the centralized scheme versus false alarm probability, when $M = 2$, $\beta_C = 0.001$.

In most existing approaches to decentralized hypothesis testing, the fusion center only utilizes the data contained in sensing reports for the decision making. However, the times at which these sensing reports are received can also infer the true hypothesis. In the proposed work, the information from both the statistics and arrival times of sensing reports are applied at the fusion center. To further investigate the attribution in the reporting time information, we compare the performance of Algorithm 1 with that using a simplified fusion statistic, i.e.,

$$u_k = \sum_{m=1}^{M} \sum_{n=1}^{\mathcal{N}_k^m} \log \frac{P_1(\mathfrak{R}_n^m = r_{t_n^m}^m)}{P_0(\mathfrak{R}_n^m = r_{t_n^m}^m)}. \tag{3.66}$$

Comparing (3.66) with (3.27), one can notice that the statistic above only fuses the data contained in sensing reports. We refer to this strategy as the dependency removed (DR) algorithm. The DR algorithm does not necessarily preserve asymptotic person-by-person optimality, and is only used for the comparison with the proposed algorithm. Through such comparison, the contribution of the time dependency terms to the detection performance can be shown in different situations.

In Figure 3.4, the ARE performance of both Algorithm 1 and the DR algorithm are presented. One can notice that by taking the time dependency into account, the ARE of Algorithm 1 increases. Moreover, the performance gain from the time dependency contribution is more significant when the differences between α_C and β_C are greater. This phenomenon occurs because when values of α_C and β_C are further apart, the distributions of report-generating delay under different hypotheses are more distinguishable.

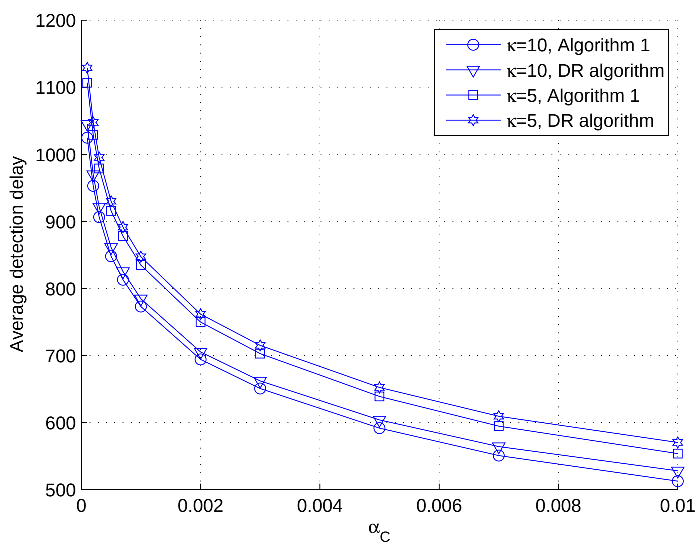

Figure 3.5: Average detection delay versus false alarm constraint, using different communication cost constraints.

In Figure 3.5, the relationship between false alarm probability and expected detection delay is illustrated, where the missed detection probability is fixed to $\beta_C = 0.001$ and $M = 2$ local sensors are applied. As the false alarm probability increases, the overall detection delay decreases. One can also notice that a shorter overall detection delay is achieved when the communication constraint κ is increased, since local sensor error probability constraints are relaxed, leading to a shorter delay for local sensor tests. The contribution of time dependency to the delay performance can be observed.

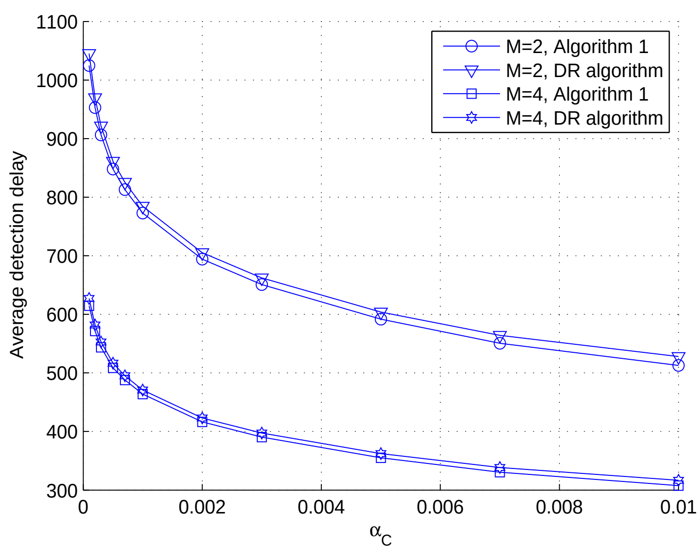

Figure 3.6: Average detection delay versus false alarm constraint, using different numbers of local sensors.

Figure 3.6 investigates the delay performance when different numbers of local sensors are deployed, i.e., $M = 2$ and 4. Both the results with and without time dependency are presented. Although shorter delay is achieved when deploying more local sensors, performance improvement is not proportional to the number of applied sensors. More specifically, the average detection delay achieved using four sensors is greater than half of the average delay using two sensors, which is consistent with the ARE of proposed algorithm decreasing as M increases.

3.7 Summary

A decentralized hypothesis testing framework is proposed. In addition to the criterion of detection delay with constrained error probabilities, we introduce a new communication cost constraint. Based on the new formulation, the energy consumption and transmission overhead of sensing devices can be controlled. An asymptotically person-by-person optimum algorithm is developed to solve the formulated optimization problem, where the fusion center utilizes not only the data contained in local sensing reports, but also the times at which these reports are received. Asymptotic relative efficiency of proposed algorithm with respect to the centralized detection scheme provides insights to various trade-offs encountered in the design and performance of this sensor network.

Chapter 4

Decentralized hypothesis testing with Poisson model

4.1 Introduction

In the previous chapter, it is shown that when local sensors apply SPRTs and local threshold values are fixed, the observations at the fusion center are time dependent. The proposed fusion test not only applies the reported statistics sent by local sensors but also utilizes the reporting times when a report is forwarded by local sensors. From the simulation results in the previous chapter, one can notice that the performance gain by exploiting the reporting times is not significant, especially when false alarm and missed detection probabilities are close to one another. Moreover, the distribution of the report-generating delay can be obtained from either recursively calculating the convolution in (3.17) or applying off-line Monte Carlo experiments, both of which consume considerable computational resource of the sensing system. This motivates the idea of a more conventional form of the fusion center strategy, which only fuses the statistics contained in the local sensor reports. As the dependency of the report-generating delay is ignored, the arrival time of sensing reports is assumed memoryless, and therefore the transmissions of local reports is modeled as a Poisson process.

4.2 Problem formulation

We consider the same distributed sensing system in Figure 3.1, with unit cost to transmit a single report from different sensors, i.e., $w^m = 1$, for all $m = 1, ..., M$. Each sensor takes observations sequentially and forwards sensing reports to the fusion center where the final decision is made. Let X_k denote the observation obtained by a certain local sensor at time slot k. We consider a binary hypothesis problem where X_k has either the common probability density function (PDF) f_0 or f_1, i.e.,

$$\begin{aligned} H_0 &: X_1,\ X_2, ... \sim f_0, \\ H_1 &: X_1,\ X_2, ... \sim f_1, \end{aligned} \tag{4.1}$$

with known prior probabilities

$$P(H_0) = \pi_0 \text{ and } P(H_1) = 1 - \pi_0. \tag{4.2}$$

After taking each observation, a local sensor decides whether it is worth sending a report to the fusion center. Let $Y_k^m \in \{0, 1\}$ denote the reporting decision indicator of a certain local sensor at time k.

$$Y_k = \varsigma_k(g_{k-1}, X_k), \quad k = 1, 2... \tag{4.3}$$

where g_{k-1} is the cumulative statistic based on observations before time k, i.e. $\{X_1, ...,$ $X_{k-1}\}$, and ς_k is the local decision policy at time k. If $Y_k = 1$, the local sensor forwards a local sensing report to the fusion center for a final decision. If $Y_k = 0$, the local sensor does not send a report but continues taking observations. Reported information by the local sensor is the accumulated statistic at time k and takes the form

$$r_k = \begin{cases} 0, & Y_k = 0, \\ g_k^m, & Y_k = 1, \end{cases} \tag{4.4}$$

where $r_k = 0$ means that no report is transmitted from the local sensor to the fusion center. After transmitting a sensing report to the fusion center, each local sensor test statistic is reinitialized to zero. It is not necessary for a local sensor to draw a conclusion on a hypothesis. Only if a local sensor makes a decision in favor of either hypothesis does it forward a report to the fusion center.

At the fusion center, sensing reports from local sensors are received sequentially, and can be viewed as the observations of the fusion center. The cumulative statistic at the fusion center is updated based on reported sensed information,

$$l_k = \xi(l_{k-1}, r_k), \quad k = 1, 2..., \tag{4.5}$$

with ξ represents the statistic update function and r_k is the reported sensing information at the kth time slot. Based on the statistic l_k, the fusion center makes decisions on whether to stop and make a final decision on H_0 or H_1 or to continue receiving sensing reports, i.e.,

$$G_k = \varrho(l_k), \quad n = 1, 2..., \tag{4.6}$$

where ϱ is the decision policy. $G_k \in \{0, 1\}$ denotes the stopping decision of the fusion center at time k. If $G_k = 1$, the fusion center stops and claims H_0 or H_1. If $G_k = 0$, the detection task continues.

We are interested in the detection strategy that minimizes the overall expected detection delay under error probability and communication cost constraints, i.e.,

$$\begin{aligned} \underset{\varsigma \quad \varrho}{\text{minimize}} \qquad & E\{\tau\}, \\ \text{subject to} \qquad & P_{FAC} \leq \alpha_C,\ P_{MDC} \leq \beta_C, \\ & \text{and} \ \ E\{R\} \leq \kappa, \end{aligned} \tag{4.7}$$

where τ represents the stopping time of the detection rule, P_{FAC} and P_{MDC} denote false alarm and missed detection probability, constrained to α_C and β_C, respectively. R is the number of sensing reports received by the fusion center, i.e., number of communications, during the detection task and κ is an integer communication cost constraint.

4.3 Proposed algorithm

In the sensing system, local sensors cannot communicate with one another, and the fusion center only provides minimum feedback when a final decision is made. Thus, sensing devices cannot optimize their strategies cooperatively; the best they can do is to individually optimize their strategies based on local information. Thus, the optimization problem local to each sensor is of the form

$$\begin{aligned} \underset{\eta}{\text{minimize}} \qquad & E\{v\}, \\ \text{subject to} \qquad & P_{FAL} \leq \alpha_L \text{ and } P_{MDL} \leq \beta_L, \end{aligned} \tag{4.8}$$

where v represents the delay for a local sensor to generate a sensing report. P_{FAL} and P_{MDL} are the local false alarm and missed detection probability, constrained to α_L and β_L, respectively. Here, we also apply Assumptions 3.1, 3.2, and 3.3 in Chapter 3, and from Proposition 3.1, we have that the optimal solution to (4.8) can be achieved by tracking the likelihood ratio statistics at each local sensor via

$$g_k = \sum_{i=1}^{k} s_i \quad \text{with} \quad s_i = \log \frac{f_1(X_i)}{f_0(X_i)}. \tag{4.9}$$

A local sensor forwards a report to the fusion center whenever the test statistic g_k is significant, i.e., exceeds predefined thresholds,

$$Y_k = \begin{cases} 0, & g_k \in \{a, b\}, \\ 1, & g_k \notin \{a, b\}, \end{cases} \tag{4.10}$$

where $a < 0 < b$ are local thresholds. After a sensing report is forwarded to the fusion center, the local sensor repeats its sequential test until the final decision at the fusion center is made. We assume that when the local statistic exceeds the thresholds a and b, the excess value over the threshold is negligible. As discussed in Chapter 2, this assumption is accurate when the number of observations taken at local sensors is large. Using the Proposition 2.1 in Chapter 2, the local threshold values and error probabilities are related by

$$a = \log \frac{\beta_L}{1 - \alpha_L} \text{ and } b = \log \frac{1 - \beta_L}{\alpha_L}. \tag{4.11}$$

We consider the simpler case where error probability constraints are identical at each local sensor, so that the sensing results reported from local sensors are equally reliable. Following Proposition 2.2 in Chapter 2, the average delay to generate each local sensing report can be expressed as

$$E_\theta\{v\} = \frac{1}{d_\theta}\left[\alpha_\theta \log\frac{1-\beta_L}{\alpha_L} + (1-\alpha_\theta)\log\frac{\beta_L}{1-\alpha_L}\right], \tag{4.12}$$

where $\alpha_0 = \alpha_L$ and $\alpha_1 = 1 - \beta_L$. $E_\theta\{\bullet\}$ denotes conditional expectation under the hypothesis H_θ, $\theta \in \{0, 1\}$, and d_θ is given by

$$d_\theta = E_\theta\{\log\frac{f_1(X_1)}{f_0(X_1)}\}, \quad \theta \in \{0, 1\}. \tag{4.13}$$

The average rate that a fusion center receives sensing reports is given by

$$\lambda_\theta = \frac{M}{E_\theta\{v\}}. \tag{4.14}$$

As established in the previous chapter, the performance gain by applying the dependency of local sensor reporting times is limited, especially when the false alarm and missed detection probability constraints are close to one another. Thus, we consider a fusion center strategy with lower computational complexity which only fuses the statistics contained in the local sensor reports.

Assumption 4.1. *The time dependence of reported information is negligible, as in the case where the false alarm and missed probability constraints are close to each other.*

As the time dependence of the reporting times is neglected, therefore the arrival time of sensing reports is assumed memoryless.

Proposition 4.1. *The transmission of local reports is a Poisson process with average arrival rate λ_θ under H_θ, $\theta \in \{0, 1\}$.*

The probability that the fusion center receives k sensing reports in a certain time slot can be then calculated as

$$P_\theta(n = k) = e^{-\lambda_\theta} \frac{\lambda_\theta^k}{k!}, \tag{4.15}$$

with

$$\sum_{k=0}^{\infty} P_\theta(n = k) = e^{-\lambda_\theta} + e^{-\lambda_\theta}\lambda_\theta + o(\lambda_\theta) = 1. \tag{4.16}$$

In (4.16), we neglect the probability that the fusion center receives more than one sensing report in a single time slot, i.e., we consider that the fusion center receives either no report or only one report in each time slot. As the excess value over thresholds a and b is neglected, the reported statistic at a local sensor is given by

$$r_k = \begin{cases} a, & g_k \leq a, \\ 0, & a < g_k < b, \\ b, & g_k \geq b. \end{cases} \tag{4.17}$$

If local false alarm and missed detection constraints are achieved with equality, we can further express the distribution of received sensing information as

$$\begin{aligned} P(r_k = 0|H_\theta) &= e^{-\lambda_\theta}, \\ P(r_k = a|H_\theta) &= \lambda_\theta e^{-\lambda_\theta}(1-\alpha_\theta), \\ P(r_k = b|H_\theta) &= \lambda_\theta e^{-\lambda_\theta}\alpha_\theta. \end{aligned} \tag{4.18}$$

As (4.18) differs under H_0 and H_1, the optimal stopping problem at the fusion center can be formulated as the following binary hypothesis problem:

$$\begin{aligned} H_0 &: r_1,\ r_2, \dots \sim P_0(r), \\ H_1 &: r_1,\ r_2, \dots \sim P_1(r), \end{aligned} \tag{4.19}$$

where $P_\theta(r)$, $\theta \in \{0,1\}$, is the probability mass function of r_k under hypothesis H_θ, $\theta \in \{0,1\}$.

Proposition 4.2. *Under Assumptions 3.2, 3.3, and 4.1, the person-by-person optimal decision strategy at the fusion center is to apply SPRT.*

Proof. For any fixed decision rules of local sensors, the fusion center is faced with a classical sequential detection problem, the observations of which are reported information by local sensors. Following Proposition 4.1, these observations received by the fusion center are independent and identically distributed. Thus, Proposition 4.2 follows from Theorem 2.1. □

Thus, the fusion center decision rule tracks likelihood ratio statistic

$$l_k = \sum_{i=1}^{k} \log \frac{P_1(r_i)}{P_0(r_i)}. \tag{4.20}$$

The fusion center stops and makes a final decision G_k at time k whenever the test statistic l_k is significant, predefined thresholds, A and B are exceeded, i.e.,

$$G_k = \begin{cases} 0, & l_k \in (A, B), \\ 1, & l_k \notin (A, B). \end{cases} \tag{4.21}$$

where $A < 0 < B$ are fusion center thresholds, which can be determined based on the fusion center false alarm and missed detection probabilities computed from (4.7) [15]

$$A = \log \frac{\beta_C}{1-\alpha_C} \text{ and } B = \log \frac{1-\beta_C}{\alpha_C}. \tag{4.22}$$

Based on (4.11) and (4.18), the expected contribution of the likelihood ratio to the fusion center statistic can be calculated as

$$\begin{aligned} E_\theta\{\log \frac{P_1(r_i)}{P_0(r_i)}\} &= e^{-\lambda_\theta} \log \frac{e^{-\lambda_1}}{e^{-\lambda_0}} + \lambda_\theta e^{-\lambda_\theta} \alpha_\theta \log \frac{\lambda_1 e^{-\lambda_1} \alpha_1}{\lambda_0 e^{-\lambda_0} \alpha_0} \\ &\quad + \lambda_\theta e^{-\lambda_\theta} (1-\alpha_\theta) \log \frac{\lambda_1 e^{-\lambda_1} (1-\alpha_1)}{\lambda_0 e^{-\lambda_0} (1-\alpha_0)} \\ &= c_{\lambda_\theta} + \lambda_\theta e^{-\lambda_\theta} (1-\alpha_\theta) a + \lambda_\theta e^{-\lambda_\theta} \alpha_\theta b, \end{aligned} \tag{4.23}$$

with

$$c_{\lambda_\theta} = e^{-\lambda_\theta} \log \frac{e^{-\lambda_1}}{e^{-\lambda_0}} + \lambda_\theta e^{-\lambda_\theta} \log \frac{\lambda_1 e^{-\lambda_1}}{\lambda_0 e^{-\lambda_0}}. \tag{4.24}$$

If the absolute value of fusion center test thresholds are large, based on (4.18), (4.20), and (4.22), the expected number of sensing reports transmitted to the fusion center is given by

$$
\begin{aligned}
E_0\{R\} &= A\lambda_0 e^{-\lambda_0}/E_0\{\log\frac{P_1(r_i)}{P_0(r_i)}\},\\
E_1\{R\} &= B\lambda_1 e^{-\lambda_1}/E_1\{\log\frac{P_1(r_i)}{P_0(r_i)}\}.
\end{aligned}
\tag{4.25}
$$

Above set of equnations can be further expressed using (4.23), i.e,

$$
\begin{aligned}
E_0\{R\} &= \frac{A}{c_{\lambda_\theta}e^{\lambda_\theta}/\lambda_\theta + (1-\alpha_\theta)a + \alpha_\theta b},\\
E_1\{R\} &= \frac{B}{c_{\lambda_\theta}e^{\lambda_\theta}/\lambda_\theta + (1-\alpha_\theta)a + \alpha_\theta b}.
\end{aligned}
\tag{4.26}
$$

If the communication cost constraint is achieved with equality, the relationship between local sensor and fusion center threshold values can be expressed as

$$
\begin{aligned}
c_{\lambda_0} + \lambda_0 e^{-\lambda_0}((1-\alpha_L)a + \alpha_L b) &= A\lambda_0 e^{-\lambda_0}/\kappa,\\
c_{\lambda_1} + \lambda_1 e^{-\lambda_1}(\beta_L a + (1-\beta_L)b) &= B\lambda_1 e^{-\lambda_1}/\kappa.
\end{aligned}
\tag{4.27}
$$

From hypotheses H_0 and H_1, the two nonlinear equations in (4.27) can be solved to determine local sensor thresholds a and b in terms of $A < 0 < B$ and κ.

The proposed procedure is summarized by Algorithm 2.

4.4 Performance analysis

Similar to the analysis in the previous chapter, we investigate the asymptotic relative efficiency (ARE) [67] of the proposed detection scheme with respect to the centralized

Algorithm 2 Decentralized detection with Poisson arrival model.

procedure
  $g^m \leftarrow 0,\ m = 1, 2, ..., M.$
  while $l \in (A, B)$ **do**
    $r^m \leftarrow 0,\ m = 1, 2, ..., M.$
    for $m = 1$ to M **do** ▷ Local sensing strategy.
      if $g_k^m \notin (a, b)$ **then** $r_k^m \leftarrow g_k^m,\ g_k^m \leftarrow 0.$
      else $g^m \leftarrow g^m + \log \frac{f_1^m(X_k^m)}{f_0^m(X_k^m)}.$
      end if
    end for
    for $m = 1$ to M **do** ▷ Fusion center strategy.
      if $r^m \neq 0$ **then** $l \leftarrow l + \log \frac{P_1(r_k^m)}{P_0(r_k^m)}.$
      end if
    end for
  end while
  if $l \geq B$ **then** H_1 is true., ▷ Decision making.
  else H_0 is true.
  end if
end procedure

scheme, i.e.,

$$\text{ARE}_\theta(\alpha_C, \beta_C, \kappa) = \lim_{H_1 \to H_0} \frac{E_\theta^C\{\tau\}}{E_\theta\{\tau\}},\ \theta \in \{0, 1\}, \tag{4.28}$$

where $E_\theta^C\{\tau\}$ denotes the average detection delay of the sequential probability ratio test, which is the delay-optimum centralized scheme and is given by

$$E_\theta^C\{\tau\} = \frac{1}{d_\theta}\left[\beta_\theta \log \frac{1-\beta_C}{\alpha_C} + (1-\beta_\theta)\log \frac{\beta_C}{1-\alpha_C}\right], \tag{4.29}$$

where $\beta_0 = \alpha_C$ and $\beta_1 = 1 - \beta_C$, and

$$d_\theta = E_\theta\{\log \frac{f_1(X_1^1, ..., X_1^M)}{f_0(X_1^1, ..., X_1^M)}\},\ \theta \in \{0, 1\}, \tag{4.30}$$

with $f_\theta(X^1, ..., X^M)$ denotes the joint pdf of local observations under H_θ, $\theta \in \{0, 1\}$. Since local observations across sensors are independent conditioned on each hypothesis,

$$d_\theta = E_\theta\{\log \frac{f_1^1(X_1^1)}{f_0^1(X_1^1)} + ... + \frac{f_1^M(X_1^M)}{f_0^M(X_1^M)}\} = \sum_{m=1}^{M} d_\theta^m. \tag{4.31}$$

It should be noted that as H_1 approaches H_0, the detection delay of the local test approaches infinity, and the excess value over test thresholds vanishes. Therefore the use of (4.29) in the calculation of (4.28) does not rely on the assumption that excess over threshold values are negligible. For the case of identical local sensors

$$d_\theta^m = \frac{d_\theta}{M} = E_\theta\{\log \frac{f_1^1(X_1^1)}{f_0^1(X_1^1)}\},\ m = 1, ..., M. \tag{4.32}$$

If R sensing reports are received by the fusion center over the course of the detection task, the detection delay of proposed algorithm can be expressed as

$$E_\theta\{\tau\} = \frac{E\{R\}}{M} E_\theta^m\{v\}. \tag{4.33}$$

By substituting (4.12), (4.29), and (4.33) into (4.28), we obtain

$$\begin{aligned} \mathrm{ARE}_\theta(\alpha_C, \beta_C, \kappa) = &\left[\beta_\theta \log\tfrac{1-\beta_C}{\alpha_C} + (1-\beta_\theta)\log\tfrac{\beta_C}{1-\alpha_C}\right] \\ &\times \left[\alpha_\theta \log\tfrac{1-\beta_L}{\alpha_L} + (1-\alpha_\theta)\log\tfrac{\beta_L}{1-\alpha_L}\right]^{-1} \kappa^{-1}. \end{aligned} \tag{4.34}$$

4.5 Numerical results

A scenario is considered where observations at each local sensor are IID Gaussian sequences with mean 0 and variance 1 under H_0, and mean 0.1 and variance 1 under H_1. For similar reasons as explained in the first paragraph of Section 3.6, without much loss in generality, only a single set of example PDFs are used. The prior probability that H_0 is true is $p_0 = 0.5$, and $\alpha_C = \beta_C$ is considered. Note that for a different setup of local observation distributions, the fusion center test threshold values stay the same, as they depend only on the predefined false alarm and missed detection constraints. However, different local test threshold values may be obtained, as nonlinear equations (4.27) are changed. The false alarm and missed detection probabilities and the average detection delay are calculated based on 10^5 Monte Carlo trials.

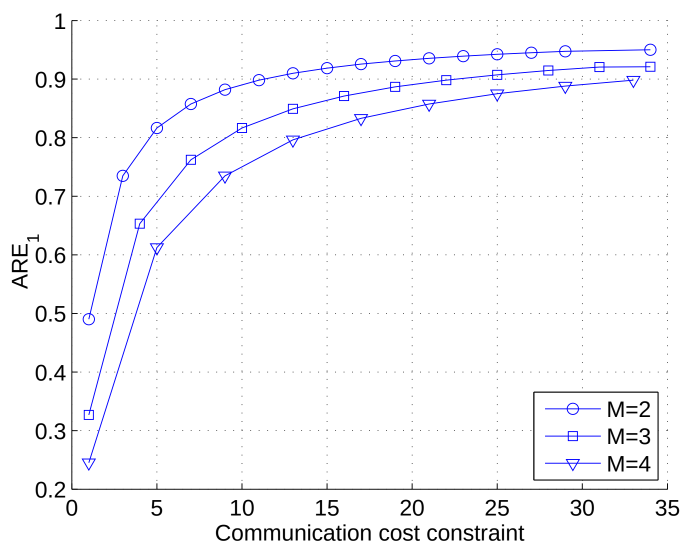

Figure 4.1: ARE_1 of Algorithm 2 with respect to the centralized scheme versus communication cost constraint, κ.

Figure 4.1 shows ARE versus κ when H_0 is true and $\alpha_C = \beta_C = 0.01$. The ARE of Algorithm 2 increases as the communication cost constraint increases. This is expected since allowing more communications between local sensors and the fusion center reduces the information loss during local sensing report generation. We observe that ARE is reduced as more local sensors are utilized under the same constraint, κ, i.e., more observations in the overall system need to be processed and quantized to obtain the same performance.

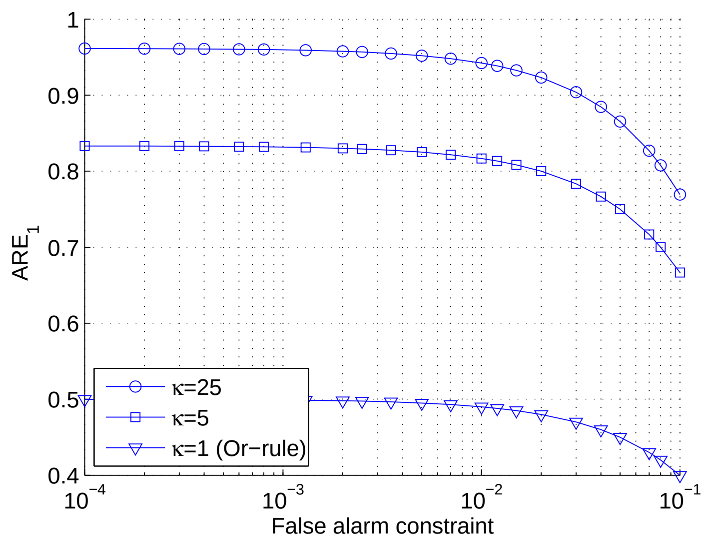

Figure 4.2: ARE_1 of Algorithm 2 with respect to the centralized scheme versus false alarm constraint.

In Figure 4.2, the effect of false alarm probability on the ARE is investigated. As shown, the ARE of Algorithm 2 decreases as the false alarm probability increases. This is expected: when the false alarm probability at the fusion center increases, the false alarm probabilities of local sensors' sequential tests increase as well. Since the transmitted sensing reports become less reliable, the ARE of Algorithm 2 decreases. An extreme case of the proposed test with $\kappa = 1$ (Or-rule) is also considered for comparison. In summary, one can observe the gain in ARE by enabling more communications.

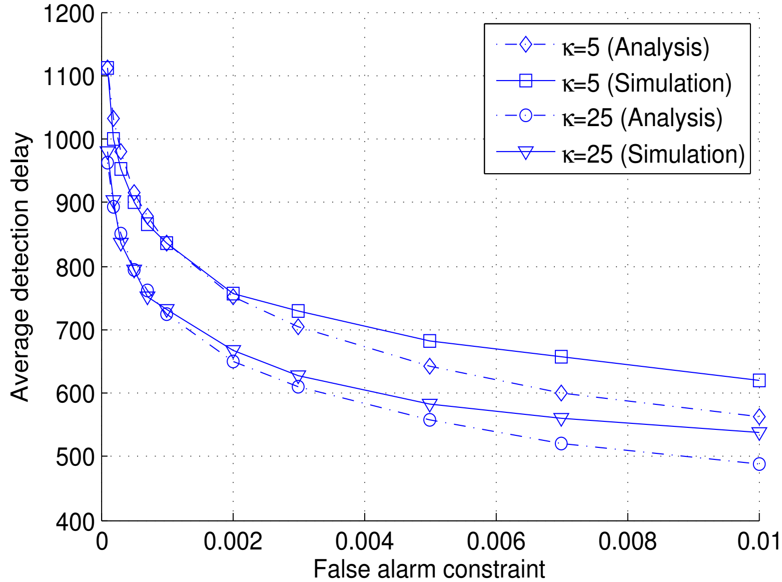

Figure 4.3: Average detection delay versus false alarm constraint, with different communication constraints.

In Figure 4.3, the relationship between false alarm probability and expected detection delay is illustrated, where the analytical performance is calculated based on (4.28). Average detection delay obtained from analysis and simulation are in closer agreement when the false alarm constraint is small. The gap is due to the fact that assumptions that neglect excess values over thresholds are becoming less accurate. One can also notice that a shorter overall detection delay is achieved when κ is increased, since local sensor error probability constraints are relaxed, leading to a shorter delay for local sensor tests.

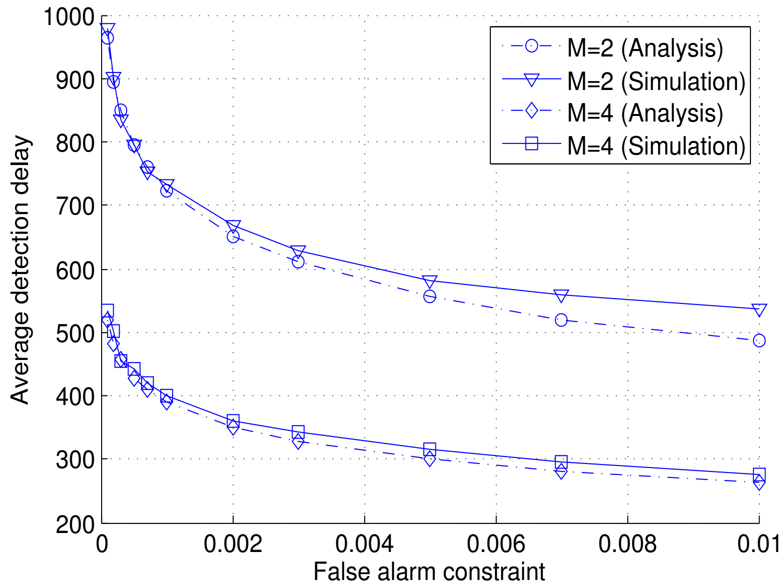

Figure 4.4: Average detection delay versus false alarm constraint, with different numbers of local sensors.

Figure 4.4 investigates performance when different numbers of local sensors are deployed. Both analytical and simulation results are presented. Although shorter delay is achieved when deploying more local sensors, performance improvement is not in proportion to the number of applied sensors. More specifically, the average detection delay achieved using four sensors is greater than half of the average delay using two sensors, which is consistent with ARE of proposed algorithm decreasing as M increases.

In Figure 4.5, the detection performance of Algorithm 1 from Chapter 3 is considered for comparison. The relationship between false alarm probability and expected detection delay is illustrated, when the missed detection probability is fixed to $\beta_C = 0.002$, and $M = 2$ and 4. As is shown in the figure, the performance loss by using the Poisson arrival approximation is relatively small, especially when false alarm and missed detection probabilities are close to one another. As predicted, when

$\alpha_C = \beta_C$ the two algorithms have nearly identical performance.

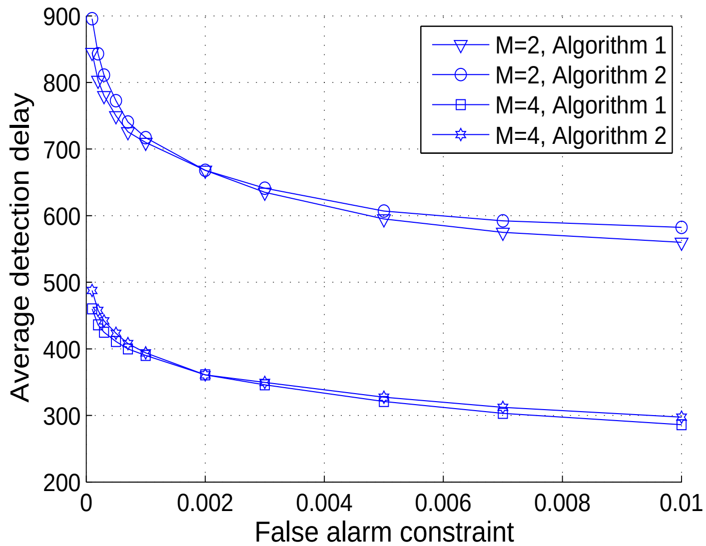

Figure 4.5: Performance comparison between Algorithm 2 and Algorithm 1 in Chapter 3, with different numbers of local sensors.

4.6 Summary

In this chapter, a simplified decentralized hypothesis testing algorithm is proposed using the same system model as in the previous chapter. The transmission of local sensing reports is modeled as a Poisson process, and a person-by-person optimum strategy is applied in solving the formulated problem. The fusion center test statistic is updated based only on the data contained in each sensing report. Compared with the algorithm developed in the previous chapter, the computational complexity is reduced as the numerical integration calculation of the report-generating distribution is avoided. Asymptotically relative efficiency of proposed algorithm with respect to the centralized detection scheme is investigated.

Chapter 5

Decentralized change detection with Poisson model

5.1 Introduction

In this chapter, we investigate the decentralized change detection problem. The considered system model and problem formulation are introduced. The formulated change detection problem aims to minimize the average detection delay, subject to the false alarm rate and the communication cost. Following the assumption in the previous chapter, the transmission of local sensing reports is modeled as a Poisson process. A person-by-person optimum strategy is applied to solve the proposed problem.

We investigate two different formulations for the communication cost constraint. One is to limit the expected number of reports received by the fusion center during the detection task. The other formulation is more restrictive, where the exact number of communications is not allowed to exceed a predefined integer. In the former formulation, the fusion center threshold is a constant, which can be achieved through dynamic programming. In the latter case, the value of fusion center threshold is time varying.

5.2 System model

We consider a distributed sensing system with M geographically separated sensors, $S_1, ..., S_M$, and a fusion center, F, that comprise the sensing system shown in Figure 5.1. The fusion center can only gain access to the data of the sensed target through each local sensor.

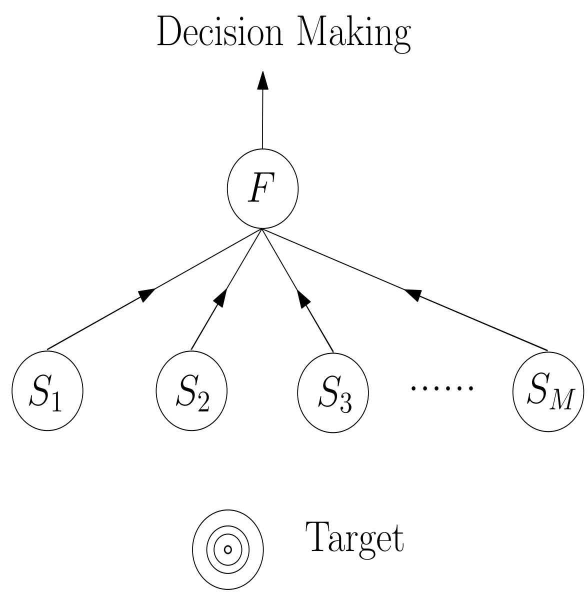

Figure 5.1: System model of change detection.

Each sensor takes observations sequentially and forwards sensing reports to the fusion center where a final decision is made. Let X_k^m denote the observation obtained by the mth sensor at time slot k. At local sensor m, $1 \leq m \leq M$, observed sequence $\{X_1^m, ..., X_k^m\}$ has common probability density function (PDF) f_0^m before a random change point Γ, and has common PDF f_1^m after Γ. The change point Γ is modeled as geometrically distributed motivated by the fact that it is the only discrete-valued distribution with the memoryless property. Let $0 < \rho < 1$ denote the distribution parameter, and $0 \leq \pi_0 < 1$ denote the prior probability that a change happened before the test. We have

$$P(\Gamma = k) = \pi_0 \mathrm{I}_{\{k=0\}} + (1 - \pi_0)\rho(1-\rho)^{k-1}\mathrm{I}_{\{k \geq 1\}}, \tag{5.1}$$

where $\mathrm{I}_{\{\bullet\}}$ is the indicator function, and π_0 represents the probability of the change having occurred before the observations are taken. The change detection problem at each local sensor can be expressed as the following multiple hypothesis testing problem:

$$\begin{aligned} H_0 &: \quad X_1, ..., X_{\Gamma-1} \sim f_0, \\ H_1 &: \quad X_\Gamma, X_{\Gamma+1}, ... \sim f_1, \end{aligned} \tag{5.2}$$

where each integer Γ indicates a different hypothesized change time. It is assumed that statistical properties of the sensors' observations change at the same time, and observations across sensors are nonidentical but independent conditioned on hypothesis H_0 or H_1.

Local sensors are able to communicate with the fusion center, but communications among local sensors are not permitted. We assume that the communication channel between sensors and the fusion center is error-free and extremely limited two-way communication is possible: the fusion center only provides local sensors simple acknowledgments to stop sensing when it is able to make a decision; otherwise, there is no feedback from the fusion center.

5.3 Problem formulation

After taking each observation, a local sensor decides whether it is worth sending a report to the fusion center. Let $V_k \in \{0, 1\}$ denote the reporting decision indicator of a certain local sensor at time k by

$$V_k = \vartheta_k(g_{k-1}, X_k), \quad k = 1, 2..., \tag{5.3}$$

where g_{k-1} is the cumulative statistic based on observations before time k, i.e. $\{X_1, ..., X_{k-1}\}$, and ϑ_k is local decision policy of the local sensor at time k. If $V_k = 0$, the local sensor continues taking observations. If $V_k = 1$, the sensor generates a local report, r_k, and sends it to the fusion center. It is not necessary for a local sensor to draw a conclusion on whether a change has occurred before sending its local report. As long as a local sensor believes that a change is likely to occur it ought to report to the fusion center.

At the fusion center, sensing reports from local sensors are received sequentially. A cumulative statistic at the fusion center, q_k, is updated at each time slot based on the reported information at current time slot k via

$$q_k = \Phi(q_{k-1}, r_k), \quad k = 1, 2..., \tag{5.4}$$

where Φ represents the statistic update function and r_k is the reported information at time k which may be sent from any local sensor. If r_k equals zero, no sensing report is received at time k. Based on the statistic q_k, the fusion center makes decisions on whether to stop or continue taking sensing reports, i.e.,

$$U_k = \varphi(q_k), \quad k = 1, 2..., \tag{5.5}$$

where φ is the fusion center decision policy and $U_k \in \{0, 1\}$ denotes the stopping decision of the fusion center at time k: if $U_k = 0$, the fusion center continues receiving sensing reports. If $U_k = 1$, the sensing task stops and claims that a change has occurred. The detection problem of the sensing system can be formulated as

$$\begin{aligned} \underset{\vartheta \quad \varphi}{\text{minimize}} \qquad & E\{(\tau - \Gamma)^+\}, \\ \text{subject to} \qquad & R_{FAC} \leq \gamma_C \ \text{ and } \ N \leq \kappa, \end{aligned} \tag{5.6}$$

where $(x)^+ = \max\{x, 0\}$ and τ is the stopping time of the sensing system. R_{FAC} is the false alarm rate at the fusion center, which is given by

$$R_{FAC} = \frac{1}{E_\infty\{\tau\}}, \tag{5.7}$$

with $E_\infty\{\bullet\}$ representing the conditional expectation given no change occurs. In (5.6), N is the number of reports received by the fusion center, i.e., number of communications before the decision is made and is given by

$$N = \sum_{m=1}^{M} \sum_{k=1}^{\tau} V_k^m. \tag{5.8}$$

In (5.6), γ_C and κ are constraints for false alarm rate and communication cost, respectively. In addition to traditional metrics of average detection delay and false alarm rate a new metric defined as the number of communications between local sensors and the fusion center is introduced. This constraint is able to reflect the cost of energy consumption and transmission overhead at local sensors. In contrast to other formulations, the proposed communication cost restricts channel uses rather than the bandwidth of the control link.

In the sensing system considered, local sensors cannot communicate with one another, and the fusion center only provides minimum feedback when a final decision is made. Thus, sensing devices cannot optimize their strategies cooperatively. The best they can do is to individually optimize their strategies based on local information. Such a suboptimal strategy is known as person-by-person optimality, i.e., it is not possible to improve the corresponding team performance by unilaterally changing any of the decision functions [31]. More specifically, we first investigate the optimal sensing strategy at each local sensor, and then develop the optimal fusion rule for fixed local sensing policies. The proposed detection strategy follows Assumptions 3.2, 3.3, and 4.1 in previous chapters, and we additionally apply:

Assumption 5.1. *The observations at a each local sensor is IID under f_0, and IID under f_1 after a random change point in time. The PDFs of f_0 and f_1 are known by all local sensors and the fusion center.*

5.4 Local sensing strategy

A final decision is made at the fusion center based on received sensing reports from local sensors. In order to minimize the overall detection delay, each sensing report ought to be generated as quickly as possible, given a constraint on the local decision reliability. Thus, the optimization problem at each local sensor can be expressed as

$$\begin{aligned} \underset{\vartheta}{\text{minimize}} \quad & E\{v\}, \\ \text{subject to} \quad & R_{FAL} \le \gamma_L, \end{aligned} \tag{5.9}$$

where v represents the delay for a local sensor to generate a sensing report and $R_{FAL} = 1/E_{\infty}\{v\}$ is the local average false alarm rate, with constraint of γ_L. Following Assumption 5.1 and Theorem 2.4, we can obtain

Proposition 5.1. *Under Assumptions 3.2, 3.3 and 5.1, the CUSUM test is locally person-by-person optimal with a set of precomputed thresholds determined globally.*

Proof. Consider the mth local sensor, with false alarm rate R_{FAL}^m, $1 \le m \le M$ incurred by the system, and suppose that the decision rules of all the other local sensors, as well as the fusion center are fixed. According to Assumptions 3.2 and 5.1, as well as Theorem 2.4, among all tests which achieve prescribed false alarm constraint, the average delay of CUSUM is minimum. In case that sensor m applies local tests other than CUSUM, a greater local detection delay is obtained, and therefore from Assumption 3.3, the cost of the overall detection system increases. □

Remark 5.1. *If the communication constraint is at the most restricted extreme, each of the local sensors performs a CUSUM. The fusion center makes a final decision when the first local sensing report is received. This leads to the One-shot CUSUM scheme [44].*

From Proposition 5.1, the optimal solution to (5.9) can be achieved by tracking likelihood ratios for all possible change hypotheses given by (5.2) corresponding to all possible change times

$$g_k = \max_{1 \le j \le k} \sum_{i=j}^{k} \log \frac{f_1(X_i)}{f_0(X_i)}. \tag{5.10}$$

A local sensor forwards a sensing report to the fusion center whenever the test statistic g_k^m is significant, i.e., exceeds a predefined threshold,

$$V_k = \begin{cases} 1, & g_k \ge h_l, \\ 0, & g_k < h_l, \end{cases} \tag{5.11}$$

where h_l is the threshold for local sequential test, which is set to satisfy the false alarm constraint γ_L. The optimality of such strategy in solving (5.9) is shown in [24].

From (5.1), we can imply that a change will occur with probability one. Due to the communication cost constraint, the fusion center cannot take more than κ sensing reports. As a result, once κ reports have been received, the detection delay can only be minimized if the decision taken is H_1. Otherwise, if we decide H_0 after the last report is received, there will be infinite delay with probability one as no communication is allowed. Under H_0, a false report is generated with rate R_{FAL} at each local sensor. Since κ false sensing reports trigger a false alarm at the fusion center, we have

$$R_{FAC} = R_{FAL} \frac{M}{\kappa}. \tag{5.12}$$

Thus, if the false alarm rate of each local sensor is set to $R_{FAL} \le \gamma_L$, we have that the false alarm rate at the fusion center must satisfy

$$R_{FAC} = \frac{m}{\kappa} R_{FAL} \le \gamma_C. \tag{5.13}$$

From (5.9) and (5.13), the local false alarm rate must satisfy

$$\gamma_L \leq \frac{\kappa}{m}\gamma_C. \tag{5.14}$$

Since the delay for generating each sensing report increases with the local false alarm rate, we set the value of γ_L to satisfy (5.14) with equality to achieve minimum local detection delay.

5.5 Fusion center strategy

At the fusion center, sensing reports are received sequentially in time. According to Theorem 2.3 in Chapter 2, if the report generation policy at the local sensor is fixed, the fusion center is faced with an optimal stopping problem, which can be solved by tracking the posterior probability [26]. Let π_k be the posterior probability that a change has occurred at the kth time slot, which is given by

$$\pi_k = P(\Gamma \leq k | F_k), \tag{5.15}$$

where $F_k = \sigma(r_1, r_2, ..., r_k)$ are the sigma algebras generated by the observed information up to the kth time slot and $r_k = 0$ if no report is received at time k. Based on Bayes' rule, π_k can be shown to satisfy the recursive formula

$$\pi_{k+1} = \begin{cases} \Phi^{(0)}(\pi_k), & \text{if } r_{k+1} = 0, \\ \\ \Phi^{(1)}(\pi_k, r_{k+1}), & \text{if } r_{k+1} \neq 0, \end{cases} \tag{5.16}$$

with

$$\Phi^{(0)}(\pi_k) = \pi_k + (1 - \pi_k)\rho \tag{5.17}$$

and

$$\begin{aligned}\Phi^{(1)}(\pi_k, r_{k+1}) &= \frac{P(H_{1,k+1})P(r_{k+1}|H_{1,k+1})}{P(r_{k+1})} \\ &= \frac{P(H_{1,k+1})P(r_{k+1}|H_{1,k+1})}{P(r_{k+1}, H_{1,k+1}) + P(r_{k+1}, H_{0,k+1})} \\ &= \frac{\Phi^{(0)}(\pi_k)P(r_{k+1}|H_1)}{\Phi^{(0)}(\pi_k)P(r_{k+1}|H_1) + (1 - \Phi^{(0)}(\pi_k))P(r_{k+1}|H_0)},\end{aligned} \tag{5.18}$$

where $P(H_{\theta,k+1}) = \Phi^{(0)}(\pi_k)$, $\theta \in \{0, 1\}$ is the prior probability of H_θ at the $(k+1)$th time slot. The third equality in (5.18) is based on the memoryless property of the Poisson model, i.e., $P(r_{k+1}|H_{1,k+1}) = P(r_{k+1}|H_1)$. The average rate that a local sensor reports to the fusion center is given by

$$\lambda_\theta^m = \frac{1}{E_\theta^m\{v\}}, m = 1, 2, ..., M, \tag{5.19}$$

where $E_\theta\{\bullet\}$ is the conditional expectation under H_θ, $\theta \in \{0, 1\}$. Thus, the average rate that a fusion center receives sensing reports is the summation of local reporting rates. That is,

$$\lambda_\theta = \sum_{m=1}^{M} \frac{1}{E_\theta^m\{v\}}. \tag{5.20}$$

Following the same argument as in the previous chapter, the transmission of sensing reports to the fusion center can be modeled as a Poisson arrival process with average arrival rate λ_θ under H_θ, $\theta \in \{0, 1\}$. The probability that the fusion center receives k sensing reports in a certain time slot under H_θ, $\theta \in \{0, 1\}$ can be calculated as (4.15) and (4.16). Moreover, the probability that the fusion center receives more than one sensing report in a single time slot is neglected. Thus, we consider that the fusion center receives either no report or only one report in each time slot. The probability of $P(r_{k+1}|H_\theta)$ can be then calculated as

$$P(r_{k+1}|H_\theta) = \lambda_\theta e^{-\lambda_\theta}, \theta \in \{0, 1\}. \tag{5.21}$$

5.5.1 Expected-communication-cost-constraint optimization

We first consider the solution to the problem with a constraint on the expected number of received sensing reports, i.e.

$$\begin{aligned} \underset{\ell}{\text{minimize}} \quad & E\{(\tau - \Gamma)^+\}, \\ \text{subject to} \quad & R_{FAC} \leq \gamma_C \ \text{ and } \ E\{N\} \leq \kappa, \end{aligned} \tag{5.22}$$

and then extend the result to solve (5.6). The total expected cost of (5.22) can be expressed as

$$R(\ell) = P_\Gamma(\tau < \Gamma) + c_d E\{(\tau - \Gamma)^+\} + c_n E\{N\}. \tag{5.23}$$

where ℓ denotes the stopping policy at the fusion center, and c_d and c_n are nonnegative constants that represent the cost of taking one sample and one report respectively. The values of c_d and c_n control the relative importance of the three performance indices in the optimization problem. At each time slot, the cost incurred if the fusion center stops is given by

$$J_{stop}(\pi_k) = 1 - \pi_k, \tag{5.24}$$

and the expected cost that the fusion center will incur if it continues the test can be expressed as

$$J_{cont}(\pi_k) = c_d + c_n P(r_{k+1}) + E\{J_{cont}(\pi_{k+1})\}, \tag{5.25}$$

with

$$P(r_{k+1}) = \pi_k P(r_{k+1}|H_1) + (1 - \pi_k) P(r_{k+1}|H_0), \tag{5.26}$$

and using (5.17) and (5.18),

$$E\{J_{cont}(\pi_{k+1})\} = P(r_{k+1}) J_{cont}(\Phi^{(0)}(\pi_k)) + (1 - P(r_{k+1})) E\{J_{cont}(\Phi^{(1)}(\pi_k, r_{k+1}))\}, \tag{5.27}$$

It can be shown that both $J_{stop}(\pi_k)$ and $J_{cont}(\pi_k)$ are nonnegative concave functions on the interval $[0, 1]$, with $J_{stop}(1) = J_{cont}(1) = 0$ [64].

Proposition 5.2. *Under Assumptions 3.2, 3.3, 4.1, and 5.1, the person-by-person optimal decision strategy at the fusion center which solves* (5.22) *is given by*

$$\tau^* = inf\{k \geq 1 | \pi_k > \pi^*\}. \tag{5.28}$$

The threshold of the fusion center test, π^*, is the solution to

$$J_{stop}(\pi) = J_{cont}(\pi). \tag{5.29}$$

Proof. For any fixed decision rules at local sensors, the fusion center is faced with a classical sequential detection problem, the observations of which are reported information by local sensors. Following Proposition 4.1, these observations received by the fusion center are independent and identically distributed. Thus, Proposition 4.2 follows from Theorem 2.3. □

In Proposition 5.2, the cost function $J_{cont}(\pi)$ is computed recursively using (5.25). Figure 5.2 shows an example of $J_{cont}(\pi)$ and $J_{stop}(\pi)$, and the optimal solution π^* is found when two curves intersect. More accurate $J_{cont}(\pi)$ approximation can be obtained by increasing the number of iterations and points on the π-axis [65].

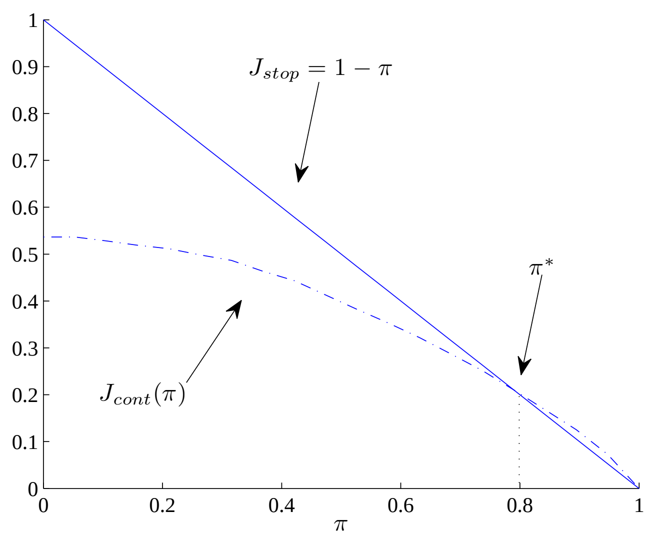

Figure 5.2: Cost function to stop and continue at the fusion center, using the expected communication cost constraint.

5.5.2 Communication-cost-constraint optimization

When we have a constraint on the expected number of received sensing reports, the optimization problem (5.22) can be solved by tracking the posterior probability and comparing it to a constant threshold. The number of received sensing reports during the detection process, N, is a random variable. It is possible that after receiving N reports, the posterior probability is less than the test threshold. In such a situation, the fusion center still needs to stop and make a decision, as no more communication is allowed. We refer to such a stop as an early termination. Since the test statistic does not reach the desired threshold, the fusion center decision in an early termination state is less reliable. In order to avoid this situation, we add a new term to the cost function, i.e.,

$$R(\varphi) = P_\Gamma(\tau < \Gamma) + c_d E\{(\tau - \Gamma)^+\} + c_n E\{N\} + c_{tn} P(\kappa < \sum_{i=1}^{\tau} r_i), \qquad (5.30)$$

where c_{tn} and $P(\kappa < \sum_{i=1}^{\tau} r_i)$ are the cost and probability of early termination, respectively. The value of c_{tn} is set to be greater than 0.5 so that the penalty is never less than simply guessing the result. This avoids situations in which the fusion center prefers to exceed the communication cost constraint. At a certain time slot, if the fusion center stops to make a decision, it incurs a cost of error. However, by stopping the test, the fusion center avoids the cost of termination. At a certain time slot, the probability of early termination is related to the applied test threshold and the number of sensing reports already received by the fusion center. Let $P_{tn}(n, \pi^L)$ denote the early termination probability using the test threshold π^L when n sensing reports (out of κ) is received at the fusion center. We can express the cost function to stop before the first sensing report is received given by

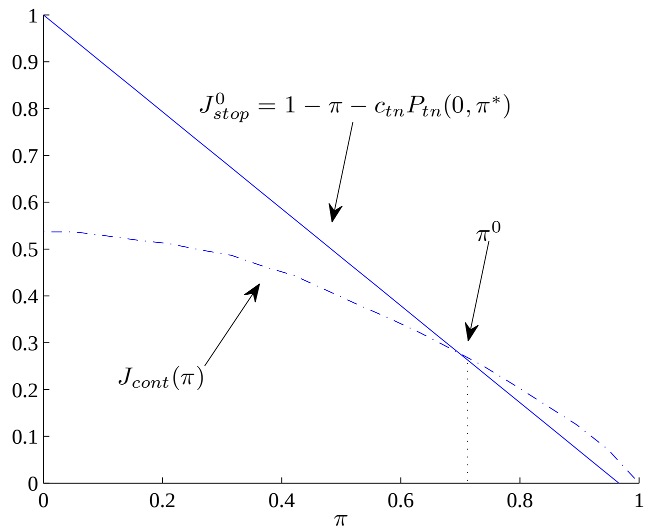

Figure 5.3: Cost function to stop and continue at the fusion center, using the communication cost constraint.

$$J^0_{stop}(\pi_k) = 1 - \pi_k - c_{tn} P_{tn}(0, \pi^*). \tag{5.31}$$

The value of $P_{tn}(0, \pi^*)$ can be computed numerically and we can then express J^0_{stop} as a function of π_k. The cost function of continuing can be expressed in the same way as (5.25), which can be approximated through dynamic programming. An initial threshold value π^0 is the solution to

$$J^0_{stop}(\pi) = J_{cont}(\pi). \tag{5.32}$$

Figure 5.3 shows an example of curves $J^0_{stop}(\pi_k)$ and $J_{cont}(\pi)$, and the solution to (5.32) is found at their intersection.

Every time the fusion center receives a new sensing report, the probability of early termination changes as the number of communications approaches the allowed limit κ. We need to calculate this new probability and update the value of the cost function to stop. Let $J^n_{stop}(\pi)$ denote the cost function to stop when n sensing reports are received. We have

$$\begin{aligned} J^1_{stop}(\pi_k) &= 1 - \pi_k - c_{tn} P_{tn}(1, \pi^0), \\ J^n_{stop}(\pi_k) &= 1 - \pi_k - c_{tn} P_{tn}(n, \pi^{n-1}),\ n \geq 2, \end{aligned} \tag{5.33}$$

where π^{n-1} is the solution to

$$J_{cont}(\pi_k) = 1 - \pi_k - c_{tn} P_{tn}(n-1, \pi^{n-2}). \tag{5.34}$$

As the fusion center continues taking sensing reports, we obtain a sequence of thresholds, π^n, $n = 0, 1, ...$, which are obtained recursively based on (5.33). Using a similar argument as in the proof of Proposition 5.2, we have

Proposition 5.3. *Under Assumptions 3.2, 3.3, 4.1, and 5.1, the person-by-person optimal decision strategy at the fusion center which solves* (5.6) *is given by*

$$\tau^* = inf\{k \geq 1 | \pi_k > \pi^n, where\ n = \sum_{i=1}^{k} r_i\}. \tag{5.35}$$

The proposed procedure is summarized by Algorithm 3.

5.6 Numerical results

For the numerical results presented in this section, we assume that the observations $\{X_n\}$ are i.i.d. Gaussian sequences with mean 0 and variance 1 under H_0, and mean 0.1 and variance 1 under H_1. Here, since the PDFs only affect the likelihood ratio

Algorithm 3 Decentralized change detection with Poisson transmission model.

procedure
  $i \leftarrow 1$, $\pi \leftarrow 0$, $\pi^* \leftarrow [\pi^0, ..., \pi^{\kappa-1}]$, $g^m \leftarrow 0$, $m = 1, 2, ..., M$.
  while $\pi < \pi^*(i)$ **do**
    $r \leftarrow 0$
    for $m = 1$ to M **do** ▷ Local sensing strategy.
      if $g_k^m < 0$ **then** $g_k^m \leftarrow 0$
      end if
      if $g_k^m < h_l$ **then** $g^m \leftarrow g^m + \log \frac{f_1^m(X_k^m)}{f_0^m(X_k^m)}$.
      else $r \leftarrow 1$, $g_k^m \leftarrow 0$.
      end if
    end for
    if $r == 1$ **then** $\pi \leftarrow \Phi^{(1)}(\pi)$, $i \leftarrow i + 1$. ▷ Fusion center strategy.
    else $\pi \leftarrow \Phi^{(0)}(\pi)$.
    end if
  end while
  Declare a change
end procedure

term in Algorithm 3, only a single set of example PDFs are used. We also assume the change time is geometrically distributed with parameter $\rho = 0.01$ and prior probability $\pi_0 = 0$. For simplicity, we consider a sensing system with two local sensors, $M = 2$. The false alarm rate, average detection delay, and the early termination probability are calculated based on 100000 Monte Carlo trials. For the optimal threshold computation, we consider 5000 points on π-axis and 2000 iterations.

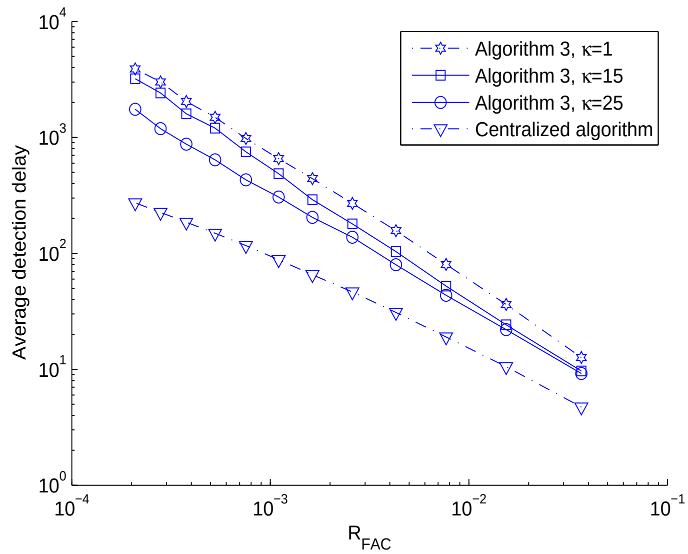

Figure 5.4: Average detection delay versus false alarm rate.

In Figure 5.4, the relationship between false alarm probability and average detection delay is illustrated when the number of communications is constrained to be 15 and 25. As noted in Figure 5.4, lower false alarm rate results in longer detection delay. This is due to local sensors requiring more time to collect enough observations to improve the detection reliability. The detection performance of one-shot CUSUM algorithm with *minimal* strategy for fusion rule is considered for comparison. It can be observed that shorter detection delay is achieved with Algorithm 3. We also investigate the performance of a centralized detection algorithm, which can be considered as the performance upper bound of Algorithm 3. In the centralized algorithm, raw data collected from local sensors are directly forwarded to the fusion center, where a CUSUM test is applied.

In Figure 5.5, we keep the false alarm rate constraint unchanged, $R_{FAC} = 0.001$ and 0.002, and investigate the detection delay under different communication cost constraints. As shown in the figure, detection delay decreases when the communication cost constraint is increased. Such behavior can be explained from (5.13). If κ is increased, local false alarm rate increases as well. This leads to a shorter delay for local sensors to generate each sensing report, and therefore reduce the detection delay of the sensing system.

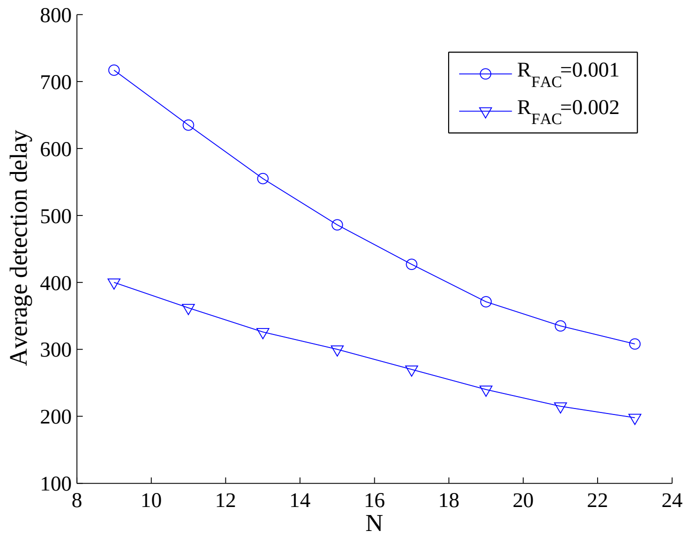

Figure 5.5: Average detection delay versus communication cost constraint.

5.7 Summary

A decentralized change detection algorithm is proposed, where the transmissions of local sensing reports are modeled as a Poisson process. The proposed optimization problem can be solved by a person-by-person strategy. That is, each local sensor optimizes its sensing policy individually and the fusion center applies a sequential test to make the final decision. Optimal threshold values for the proposed algorithm are obtained through dynamic programming. Simulation results are provided to illustrate features of the proposed algorithm.

Chapter 6

Communication-efficient decentralized change detection

6.1 Introduction

In this chapter, an alternative method to solve decentralized change detection problem is proposed. Unlike the framework in Chapter 5, the local sensor test applies the false alarm probability constraint, and the fusion center test tracks the sum of reported statistics from local sensors. A two-threshold based detection algorithm is developed, and the performance of the algorithm is optimized through the joint design of threshold values at both local sensors and the fusion center. The proposed change detection framework is much easier to implement compared with the one in the previous chapter, as a simple fusion rule is applied and coupled threshold values can be obtained through one-dimensional search. In practice, the proposed algorithm will not require the decision-epochs of the local detectors to be precisely synchronized.

6.2 System model and problem formulation

We consider the same decentralized sensing system as in Chapter 5, where local sensors are memoryless, receive independent observations, and without full feedback from the fusion center. At each time slot k, a sensor decides whether it is worth sending a report to the fusion center. Let $W_k \in \{0, 1\}$ denote the reporting decision indicator of a certian local sensor at time k.

$$W_k = \mu_k(g_{k-1}, X_k), \quad k = 1, 2... \tag{6.1}$$

where X_k is the local observation taken at time k and g_{k-1} is the cumulative statistic based on observations before time k, i.e. $\{X_1, ..., X_{k-1}\}$, and μ is local decision policy of a certain sensor. If $W_k = 0$, the local sensor continues taking observations. If $W_k = 1$, the sensor generates a local report, r_k, based on already accumulated information and sends it to the fusion center, i.e.,

$$r_k = \begin{cases} g_k, & W_k = 1, \\ 0, & W_k = 0. \end{cases} \tag{6.2}$$

Note that it is not necessary for a local sensor to draw a conclusion on whether a change has occurred before sending its local report. If the communication cost is affordable, as long as a local sensor believes that a change is likely to occur it ought to report to the fusion center. Let σ denote the decision making policy of the fusion center. The detection problem can be formulated as

$$\begin{aligned} \underset{\mu \quad \sigma}{\text{minimize}} \qquad & E\{(\tau-\Gamma)^+\}, \\ \text{subject to} \qquad & P_{FAC} \leq \alpha_C \ \text{ and } \ E\{N\} \leq \kappa, \end{aligned} \tag{6.3}$$

where $(x)^+ = \max\{x, 0\}$. In (6.3), N represents the number of reports received by the fusion center, i.e., number of communications, before the decision is made and is given by

$$N = \sum_{m=1}^{M} \sum_{k=1}^{\tau} W_k^m, \tag{6.4}$$

P_{FAC} is the probability of false alarm that the fusion center claims a change has occurred when there is no change, which is given by

$$P_{FAC} = P_\Gamma(\tau < \Gamma), \tag{6.5}$$

where P_Γ represents the conditional probability given the change occurs at time Γ. α_C and κ are constraints of false alarm probability and communication cost, respectively.

A global decision is made at the fusion center based only on the reports received from local sensors. Thus, the global performance is crucially affected by the delay and reliability of reports generated by local sensors. That is, in order to achieve minimum the average detection delay, local sensors need to forward reliable sensing reports to the fusion center as quickly as possible. This function can be achieved by applying the following strategy at local sensors:

$$\begin{aligned} \underset{\mu}{\text{minimize}} \qquad & E\{v\}, \\ \text{subject to} \qquad & P_{FAL} \leq \alpha_L, \end{aligned} \tag{6.6}$$

where υ is the delay for a local sensor to generate each sensing report, P_{FAL} is the probability of false alarm for a local sensor detecting a change, i.e.,

$$P_{FAL} = P(W_k = 1|H_0), \tag{6.7}$$

and α_L is the local sensing false alarm constraint.

6.3 Proposed algorithm

The optimal solution for the formulation of (6.3) requires joint design of policies at both local sensors and the fusion center. However, except for final acknowledgement, the communication between the local sensor and the fusion center is unidirectional and local sensors cannot communicate with one another. Thus, the desired strategy is to optimize the sensing task (6.6) locally, and then find the optimal policy at the fusion center. Here, we also apply Assumptions 3.2, 3.3, and 5.1, as in Chapter 5. Following the same argument in Chapter 5, one obtains

Proposition 6.1. *Under Assumptions 3.2, 3.3 and 5.1, the CUSUM test is locally person-by-person optimal with a set of precomputed thresholds determined globally.*

The optimal local policy can be achieved by tracking the CUSUM likelihood ratio statistic as follows:

$$g_k = \max_{1\le j\le k} \sum_{i=j}^{k} s_i \quad \text{with} \quad s_i = \log\frac{f_1(X_i)}{f_0(X_i)}, \tag{6.8}$$

where g_k represents the test statistic of the local sensor at time k, f_0 and f_1 are the probability density functions. A local sensor forwards a sensing report to the fusion center whenever the test statistic g_k is significant, i.e., exceeds a predefined positive threshold h_l. That is

$$W_k = \mathrm{I}_{\{g_k \geq h_l\}}, \tag{6.9}$$

where $\mathrm{I}_{\{\bullet\}}$ is the indicator function. The information contained in the local sensing report is the value of g_k,

$$r_k = \begin{cases} g_k, & g_k \geq h_l, \\ 0, & g_k < h_l. \end{cases} \tag{6.10}$$

In (6.10), $r_k = 0$ represents the event that no report is forwarded to the fusion center. The fusion center collects local sensing reports sequentially. The sensing task stops and claims a change occurs as soon as the test statistic at the fusion sensor exceeds a threshold, i.e.,

$$\Lambda = \sum_{m=1}^{M} \sum_{k=1}^{\tau} r_k \geq h_f, \tag{6.11}$$

where Λ is the fusion center test statistic and h_f is the fusion center threshold. It is clear from (6.10) and (6.11) that $h_f \geq h_l$.

Remark 6.1. *The proposed fusion rule* (6.11) *does not retain any optimal properties. However, it does not rely on time information of local sensing reports and therefore does not require precise synchronization of sensing devices in the system.*

The proposed procedure is summarized by Algorithm 4. In the previous chapter, the fusion center test statistic (5.16) is updated at each time slot, no matter whether a sensing report is received or not; and the statistic is updated based on the state of all local sensors at each time slot. Thus, the implementation of the algorithm in Chapter 5 requires precise synchronization of all sensing devices in the system. On the other hand, the proposed fusion center test statistic (6.11) is updated only when a sensing report is received; and as long as the summed reported statistic exceeds h_f, a final decision can be made. This enables a lower complexity detection method, where precisely synchronized devices are not needed. However, the price to be paid for avoiding synchronization is significant, which is shown in the following section.

Algorithm 4 Decentralized change detection.

procedure
  $g^m \leftarrow 0$, $m = 1, 2, ..., M$.
  while $\Lambda \in (0, h_f)$ **do**
    for $m = 1$ to M **do**
      if $g_k^m < 0$ **then** $g_k^m \leftarrow 0$
      end if
      if $g_k^m < h_l$ **then** $g^m \leftarrow g^m + \log \frac{f_1^m(X_k^m)}{f_0^m(X_k^m)}$.
      else $z \leftarrow z + g_k^m$, $g_k^m \leftarrow 0$.
      end if
    end for
  end while
  Declare a change
end procedure

The performance of the proposed algorithm can be optimized by suitable selections of threshold values h_l and h_f. For this setup, the problem formulation of (6.3) becomes

$$\begin{aligned} \underset{h_l\ h_f}{\text{minimize}} \quad & E\{(\tau - \Gamma)^+\}, \\ \text{subject to} \quad & P_{FAC} \leq \alpha_C \ \text{ and } \ E\{N\} \leq \kappa. \end{aligned} \tag{6.12}$$

The constrained optimization problem in (6.12) can be also expressed in the following manner:

$$\underset{h_l\ h_f}{\text{minimize}} \quad \alpha_C + c_1 E\{(\tau - \Gamma)^+\} + c_2 \kappa. \tag{6.13}$$

From a Bayesian view point, c_1 and c_2 are nonnegative constants that represent the cost of taking one sample and one report, respectively. The values of c_1 and c_2

control the relative importance of the three performance indices in the optimization problem. According to the well-known Karush-Kuhn-Tucker (KKT) conditions [63], the solutions to (6.12) and (6.13) can be made equivalent by proper choices of c_1 and c_2. More specifically, for arbitrary values of α_C and κ, there exists constants $c_1(\alpha_C, \kappa)$ and $c_2(\alpha_C, \kappa)$ such that the solution to (6.13) is the solution to (6.12). In order to solve (6.13), we define the cost function λ as

$$\lambda = \alpha_C + c_1 E\{(\tau - \Gamma)^+\} + c_2\kappa. \tag{6.14}$$

As in (6.10), each time a report r_k is sent to the fusion center, we have $r_k \geq h_l > 0$. Based on the fusion strategy (6.11), the expected number of reports received at the fusion center before a decision is made is lower bounded by h_f/h_l, that is

$$\kappa \leq \frac{h_f}{h_l}. \tag{6.15}$$

We next assume that when the local statistic reaches the threshold h_l, the excess value over h_l is negligible. This assumption is accurate when the number of observations taken at local sensors is large. Under this assumption, (6.15) becomes approximate equality, i.e.,

$$\kappa \cong \frac{h_f}{h_l}. \tag{6.16}$$

Average delay: The fusion center is able to make a decision after receiving κ, on average, number of reports. Thus, the average detection delay can be obtained by calculating the average delay for local sensors to generate κ reports. Let L denote the expected number of samples required by a local sensor before generating a report. It is straightforward that the average delay for a single local sensor to generate κ reports is κL. One can subsequently obtain that the average detection delay by applying M independent identical local sensors is given by

$$E\{(\tau - \Gamma)^+\} = \frac{L}{M}\kappa. \quad (6.17)$$

Using the approximation in (6.16), the average detection delay can be further expressed as

$$E\{(\tau - \Gamma)^+\} \cong \frac{h_f L}{h_l M}. \quad (6.18)$$

False alarm: If a local sensor falsely detects a change, it still forwards a statistic $r_k \cong h_l$ to the fusion center. Since the fusion strategy calculates the sum of reported local statistics, after receiving h_f/h_l mistaken reports, false alarm occurs at the fusion center. That is, h_f/h_l local false alarms trigger a global false alarm. Thus, the false alarm probability of the fusion center is related to that of the local sensors is given by

$$\alpha_C \cong \alpha_L^{\frac{h_f}{h_l}}. \quad (6.19)$$

By substituting the expressions (6.16), (6.18), and (6.19) in (6.14), we have

$$\lambda \cong \alpha_C^{\frac{h_f}{h_l}} + c_1 \frac{h_f L}{h_l M} + c_2 \frac{h_f}{h_l}. \quad (6.20)$$

From (6.8), one can notice that the statistic applied at the local sensor is the well-known CUSUM test. That is, the local sensor strategy can be viewed as a CUSUM with the threshold value h_l. It is well known that any CUSUM test can be expressed as a sequence of sequential ratio probability tests (SPRT) with boundaries $(0, h_l)$ with initial statistic zero [19]. For each SPRT, let P_0 denote the probability that the test ends on the lower boundary, 0, and $1 - P_0$ denote the probability that the test ends on the upper boundary, h_l. Also, we use L_0 and L_1 to represent the average number of samples required for the SPRT test to end on boundary 0 and h_l, respectively. The expected number of samples at the local sensor required to generate a report is therefore given by

$$\begin{aligned} L &= L_0 \sum_{n=1}^{\infty} n P_0^n (1 - P_0) + L_1 \\ &= \frac{P_0}{1 - P_0} L_0 + L_1 \\ &= \frac{P_0}{1 - P_0} L_0 + \frac{1 - P_0}{1 - P_0} L_1 \\ &= \frac{S(0, h_l)}{1 - P_0}, \end{aligned} \tag{6.21}$$

where $S(0, h_l)$ is the expected number of samples required by an SPRT to stop with boundaries $(0, h_l)$, and the value of $S(0, h_l)$ is a function of h_f and the PDF H_θ, $\theta \in \{0, 1\}$. From (6.20), the communication cost and false alarm terms do not depend on the number of local sensors, M. This means that increasing the number of sensors does not increase the delay cost nor false alarm cost. Moreover, for any constant values of c_1 and c_2, we note that by calculating the second derivative of (6.20) with respect to h_f,

$$\frac{\partial^2 \lambda}{\partial h_f^2} \cong \frac{\ln^2 \alpha_L}{h_l^2} \alpha_L^{\frac{h_f}{h_l}} > 0, \;\; h_f \geq h_l > 0. \tag{6.22}$$

Thus, (6.20) is a convex function of h_f for $h_l \leq h_f < \infty$ with the minimum value given by the solution to

$$\frac{\partial \lambda}{\partial h_f} \cong 0. \tag{6.23}$$

By solving (6.23), the relationship between the optimal value of h_l and h_f that minimizes the expected cost is given by

$$h_f \cong \frac{h_l}{\ln\alpha_L}\ln\frac{c_1 L + c_2 M}{M\ln\frac{1}{\alpha_L}}. \tag{6.24}$$

From (6.19), we have

$$h_f \cong h_l \ln\alpha_C / \ln\alpha_L. \tag{6.25}$$

By equating (6.24) and (6.25) and simplifying, we have

$$c_1 L + c_2 M - M\alpha_C \ln\frac{1}{\alpha_L} \cong 0. \tag{6.26}$$

From (6.21), $L = S(0, h_l)/(1 - P_0)$ and α_L are solely determined by h_l. In addition, c_1 and c_2 are constants determined by α_C and κ. Therefore, (6.26) can be expressed in the form of

$$f_h(h_l, \alpha_C, \kappa) \cong 0. \tag{6.27}$$

Here, f_h is a scalar function which can be uniquely determined given h_l, α_C, and κ. Thus, for any predefined α_C and κ, the optimal value of h_l which minimizes the expected total cost can be obtained by solving (6.27). The optimal h_f is then determined using (6.24) or (6.25). It is a well known result in sequential analysis that the SPRT test with boundary value of $(0, h_l)$ can be approximated as [15]

$$S(0, h_l) \cong d_1^{-1} \ln\frac{1}{\alpha_L}, \tag{6.28}$$

where d_1 is the expected value of $\log\frac{f_1(X_1)}{f_0(X_1)}$, under hypotheses H_1. The above approximation is achieved by assuming that when the test statistic in the SPRT crosses a boundary, the excess over the boundary is negligible. The approximation is consistent with that in (6.16) and will be accurate if the number of observed samples is relatively large on average. By substituting the value of $S(0, h_l)$ in (6.26), we have

$$f_h(h_l, \alpha_C) \cong (c_1 d_1^{-1}(1 - P_0)^{-1} - M\alpha_C)\ln\frac{1}{\alpha_L} + c_2 M. \tag{6.29}$$

Remark 6.2. *For local detection policy* (6.8) *and fusion rule* (6.11)*, the threshold values which minimize the detection delay subject to error probability and communication cost constraints satisfy* (6.24) *and* (6.29)*.*

In (6.29), P_0 increases with h_l whereas α_L decreases with h_l. It can easily be shown that $f_h(h_l, \alpha_C)$ is a monotonically increasing function of h_l. Thus, for an arbitrary value of α_C, the solution to (6.27) can be found through one dimensional search. After obtaining h_l, one can subsequently calculate h_f through (6.24) or (6.25).

6.4 Numerical results

For the numerical results presented in this section, we assume that the observations $\{X_n\}$ are i.i.d. Gaussian sequences with mean 0 and variance 1 under H_0, and mean 0.1 and variance 1 under H_1. Since PDFs of H_0 and H_1 only affect the likelihood ratio term in Algorithm 4, a single set of example PDFs are used. We also assume the change time is geometrically distributed with parameter $\rho = 0.01$. For simplicity, we first consider two local sensors, $M = 2$. The false alarm probability and average detection delay are calculated based on 10^5 Monte Carlo trials.

In Tables 6.1 and 6.2, the constraints of the false alarm probability α_C and the average number of communications κ are listed. Based on (6.17), the average detection

Table 6.1: Optimal solutions of thresholds under different constraints on the false alarm probability and the on average number of communications.

α_C	κ	$E\{(\tau-\Gamma)^+\}$	h_l	h_f
10^{-1}	8	73	0.5746	4.5968
10^{-2}	8	202	0.8789	7.0312
10^{-3}	8	316	1.0623	8.4984
10^{-1}	12	81	0.4777	5.7324
10^{-2}	12	196	0.6930	8.3160
10^{-3}	12	323	0.8359	10.031
10^{-1}	24	72	0.3199	7.6776
10^{-2}	24	193	0.4981	11.954
10^{-3}	24	360	0.5791	13.898

Table 6.2: Performance comparisons between theoretical and simulated results.

α_C	$\hat{\alpha}_C$	κ	$\hat{\kappa}$	$E\{(\tau-\Gamma)^+\}$	$\hat{E}\{(\tau-\Gamma)^+\}$
10^{-1}	0.1007	8	7.9006	73	98
10^{-2}	0.0092	8	7.9641	202	246
10^{-3}	0.0012	8	7.9945	316	353
10^{-1}	0.0994	12	11.549	81	105
10^{-2}	0.0094	12	11.635	196	242
10^{-3}	0.0013	12	11.868	323	354
10^{-1}	0.1030	24	20.829	72	107
10^{-2}	0.0082	24	21.975	193	227
10^{-3}	0.0011	24	22.236	360	362

delay approximations using various constraints are calculated in the third column. The threshold values h_l and h_f that correspond to these constraints are obtained through the proposed optimization methodology. In Table 6.2, the proposed Algorithm 4 is simulated based on the obtained thresholds h_l and h_f. With such threshold values, the simulated false alarm probability, $\hat{\alpha}_C$, and the average number of communications, $\hat{\kappa}$, are close to the predefined constraints (see Table 6.2). This means the false alarm probability and the communication cost can be well controlled to the desired values at the same time. However, the simulated average detection delays,

$\hat{E}\{(\tau - \Gamma)^+\}$, are greater than their approximated values calculated. This is due to the approximation in (6.16). It is noteworthy that when false alarm probability is relatively low, the achieved detection delay is close to the approximated value.

In Figure 6.1, the relationship between false alarm probability and average detection delay is illustrated when the average number of communications is constrained to be 10. As noted in Figure 6.1, the average detection delay increases as the false alarm probability decreases, since longer detection delay allows local sensors to collect more observations which improves the detection reliability. We investigated the cases where two and three local sensors are used. For any false alarm probability, it is apparent that shorter detection delay is obtained by using more local sensors as expected.

Also in Figure 6.1, the delay performance of Algorithm 4 is compared with that of Algorithm 3 in Chapter 5. Unlike the fusion center strategy in Chapter 5 which tracks the posterior probability that a change has occurred, the proposed fusion center test in this chapter is not person-by-person optimal. Thus, a longer detection delay is required by the proposed algorithm to achieve the same false alarm probability and communication cost constraints as in Chapter 5.

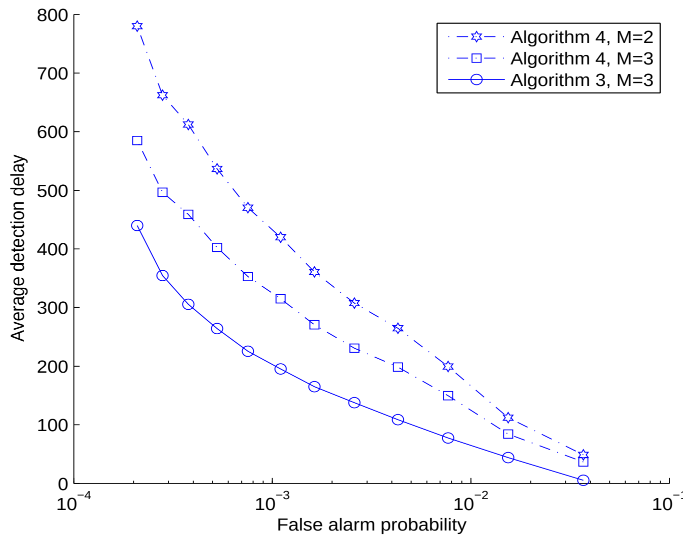

Figure 6.1: Average detection delay versus false alarm probability with communication constraint of $\kappa = 10$ for different numbers of local sensors.

6.5 Summary

In this chapter, we develop a two-threshold based communication-efficient change detection algorithm. It is shown the optimal choice of thresholds in the algorithm can be obtained through one dimensional search. Performance is investigated for different scenarios through both analysis and simulation, where the effects of approximations used are evaluated. The proposed algorithm is easy to implement and does not require synchronization.

Chapter 7

Conclusions and future work

7.1 Conclusions

In this thesis, the problem of sequential detection is studied in a decentralized sensing system, where local sensors are memoryless, energy constrained, receive independent observations, and without full feedback from the fusion center. In addition to traditional criteria of detection delay and the error probability, we introduce a new constraint: the number of communications between local sensors and the fusion center. This metric is able to reflect both the cost of establishing communication links as well as overall energy consumption over time. The communication-efficient formulations for both decentralized hypothesis testing and decentralized change detection problems are proposed. The sensing system aims to minimize the overall detection delay with constraints on both the error probability and communication cost.

In Chapter 3, an asymptotically person-by-person optimum detection framework is developed. It is shown that when local sensing strategies are fixed, the reported information to the fusion center is time dependent. The performance gain from such time dependency is investigated, and the asymptotic optimality of the fusion center test is established. The asymptotically relative efficiency of proposed algorithm with respect to the centralized strategy is expressed in closed form.

In Chapter 4, a lower complexity decentralized hypothesis testing algorithm is proposed, where the transmission of local sensing reports are modeled as a Poisson arrival process. Although the performance gain from the time dependency of reported information is lost, the computational complexity is reduced. Note that for both Chapter 3 and 4, the optimality of SPRT requires that the observations must obey the exact distributions of H_0 and H_1, i.e., simple binary hypothesis testing. Thus, the proposed algorithms may not perform optimally or even reasonably well in composite hypothesis testing situations when observations are distributed as neither H_0 nor H_1.

In Chapter 5, the decentralized change detection problem with the a communication cost constraint is investigated. Following the same Poisson process transmission model, a person-by-person optimum change detection algorithm is proposed. The optimal threshold value is obtained through dynamic programming.

In Chapter 6, a two-threshold based decentralized change detection algorithm is developed. The choices of threshold values in the algorithm are determined by a combination of sequential detection analysis and constrained optimization. The proposed sensing scheme does not require synchronization, and the fusion strategy is easy to implement.

In each chapter, simulation results are investigated to explore the tradeoffs in parameter choices of the proposed algorithm.

7.2 Future work

- In the decentralized hypothesis testing problem formulated in Chapter 3, the transmission costs of different local sensors are modeled as non-identical. Follow this formulation, the proposed decentralized change detection schemes in Chapters 5 and 6 can be extended.

- In Chapter 4, we assume that when the local test statistic exceeds the thresholds, the excess value over the threshold is negligible. Such an assumption results in over-design of test threshold values. A more accurate expression for the SPRT test thresholds, which takes the overshoot statistic into account may be of interest in future work. A possible method to tighten the excess over the boundary, S_v, would be to assume that it is uniformly distributed, and its mean value would be

$$E_\theta\{S_v\} = \frac{1}{2} E_\theta\{\log \frac{f_1(X_1)}{f_0(X_0)}\}, \theta \in \{0, 1\}. \tag{7.1}$$

  More accurate SPRT threshold values may be achieved by reducing this mean value from currently applied ones, i.e.,

$$a = \log \frac{\beta}{1-\alpha} - E_0\{S_v\} \text{ and } b = \log \frac{1-\beta}{\alpha} - E_1\{S_v\}. \tag{7.2}$$

- In Chapters 4 and 5, the transmission of local sensing reports are modeled as Poisson arrival processes. It is assumed that the fusion center receives either zero or one report at each time slot. A more general case is of interest, where multiple reports can be received at the same time.

- It is shown in Chapter 3 that when the local sensing policy is fixed, the reported information forwarded to the fusion center is time dependent. Such time dependency in the decentralized hypothesis testing problem is investigated, and the performance gain is illustrated through simulations. However, in the algorithms developed Chapter 5 and 6, the time dependent property of reported information is not utilized. In the future research, it is of interest to investigate the effect of the time dependency of reported information in the decentralized change detection problem.

- In current research work, we assume that statistics at local sensors can be directly transmitted to the fusion center. For example, in the algorithm proposed in Chapter 3, the reported information of a certain local center at time k is defined as (3.4). However, in some applications, the control link between local sensors and the fusion center is bandwidth limited. In such case, the sensing information needs to be appropriately quantized by the local sensor before sending to the fusion center, i.e.,

$$r_k = \begin{cases} 0, & T_k = 0, \\ Q(g_k), & T_k = \pm 1, \end{cases} \tag{7.3}$$

with

$$Q(y) \in \{0, 1, ..., \mathcal{L} - 1\}, y \in \mathbb{R}, \tag{7.4}$$

where $Q(\bullet)$ denotes the quantization scheme at the local sensor. $\{0, 1, ..., \mathcal{L}-1\}$ is the available alphabet for data transmission and $\mathbb{R}$ represents the set of real numbers. The statistic loss during the sensing report transmission and the information structure of the fusion center observations are related to the local quantization scheme. Thus, the design of $Q(\bullet)$ is crucial when control link is bandwidth limited.

- In the system model considered in this thesis, the fusion center makes final decision based only on the reported information from local sensors. In some applications, the fusion device can also take measurements of the sensing target itself, and the fusion center test statistic is updated based on both reported and observed information, i.e.,

$$\varpi_k = \partial_k(\varpi_{k-1}, \underline{r_k}, X_k), \tag{7.5}$$

where ϖ_k is the fusion center test statistic at time k and ∂_k is the fusion strategy. $\underline{r_k} = [r_k^1, ..., r_k^M]$ is reported information from local sensors and X_k is the observation of the fusion center taken at time slot k. This configuration can be viewed as a special case of our proposed system model, when there exists one local sensor whose cost to transmit sensing reports equals zero. Using this system model, a new communication-efficient decentralized detection formulation can be achieved. It would be of interest to find detection strategies with such formulation.

References

[1] D. Siegmund, *Sequential Analysis, Tests and Confidence Intervals*. New York: Springer, 1985.

[2] B. K. Ghosh and P. K. Sen, Handbook of Sequential Analysis. New York: Marcel Dekker, 1991.

[3] R. R. Tenney and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, no. 4, pp. 501-510, Jul. 1981.

[4] J. Zhang and R. S. Blum, "Asymptotically optimal truncated multivariate Gaussian hypothesis testing with application to consensus algorithms," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 431-442, 2014.

[5] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 1917-1932, 2011.

[6] F. Cattivelli and A. H. Sayed, "Diffusion lms-based distributed detection over adaptive networks," *in Conf Rec. IEEE 43rd Asilomar Conf. Signals, Syst. Comput.*, pp. 171-175, 2009.

[7] Q. Zhou, S. Kar, L. Huie, and S. Cui, "Distributed estimation in sensor networks with imperfect model information: an adaptive learning-based approach," *in Proc. IEEE Int. Conf. Acoust., Speech Signal Processing (ICASSP)*, pp. 3109-3112, 2012.

[8] P. Braca, S. Marano, and V. Matta, "Enforcing consensus while monitoring the environment in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3375-3380, 2008.

[9] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks based on diffusion estimation schemes," *in Proc. IEEE 10th Workshop on Signal Process. Adv. Wireless Commun. (SPAWC'09)*, pp. 61-65, 2009.

[10] D. Jakovetic, J. M. F. Moura, and J. Xavier, "Distributed detection over noisy networks: Large deviations analysis," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4306-4320, 2012.

[11] Q. Zhou, S. Kar, L. Huie, H. V. Poor, and S. Cui, "Robust distributed least-squares estimation in sensor networks with node failures," *in Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, pp. 1-6, 2011.

[12] P. Braca, S. Marano, V. Matta, and P. Willett, "Asymptotic optimality of running consensus in testing binary hypotheses," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 814-825, 2010.

[13] D. Bajovic, D. Jakovetic, J. Xavier, B. Sinopoli, and J. M. Moura, "Distributed detection via Gaussian running consensus: Large deviations asymptotic analysis," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4381-4396, 2011.

[14] S. Kar, R. Tandon, H. V. Poor, and S. Cui, "Distributed detection in noisy sensor networks," *in Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2856-2860, 2011.

[15] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge University Press, Cambridge UK, 2008.

[16] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. Math. Statist.*, vol. 19, pp. 326-329, 1948.

[17] V. P. Dragalin, A. G. Tartakovsky, and V. Veeravalli, "Multihypothesis sequential probability ratio testsPart I: Asymptotic optimality, *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2448-2461, Nov. 1999.

[18] V. P. Dragalin, A. G. Tartakovsky, and V. Veeravalli, "Multihypothesis sequential probability ratio testsPart II: Accurate asymptotic expansions for the expected sample size," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 1366-1383, Jul. 2000.

[19] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100-115, 1954.

[20] D. Yeelin and Y. T. Su, "A sequential test based cooperative spectrum sensing scheme for cognitive radios," in *Proc. IEEE 19th International Symposium on PIMRC,* pp. 1-5, Sept. 2008

[21] L. Lai, Y. Fan, and H. V. Poor, "Quickest detection in cognitive radio: A sequential Change detection framwork," in *Proc. IEEE GLOBEOM 2008*, pp. 1-5, Nov. 2008.

[22] L. Lai, H. V. Poor, X. Yan, and G. Georgiadis, "Quickest search over multiple sequences," *IEEE Trans. Inf. Theory,* vol. 57, no. 8, pp. 5375-5386, Aug. 2011.

[23] Q. Zhao and J. Ye, "Quickest detection in multiple on-off process," *IEEE Trans. Signal process.,* vol. 58, no. 12, pp. 5994-6006, Dec. 2010.

[24] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Ann.Stat.*, vol. 14 pp. 1379-1387, Dec. 1986.

[25] Y. Ritov, "Decision theoretic optimality of the CUSUM procedure," *Ann. Stat.*, vol. 18, pp. 1464-1469, Nov. 1990.

[26] A. N. Shiryaev, "On optimum method in quickest detection problems," *Theory Prob. Appl.*, vol. 8, pp. 22-46, Jun. 1963.

[27] Z. Quan, S. Cui, H.V. Poor, and A. Sayed, "Collaborative wideband sensing for cognitive radios," *IEEE Signal Processing Mag.*, vol. 25, pp. 60 - 73, Nov. 2008.

[28] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proc. IEEE* vol. 97, no. 5, pp. 849-877, May 2009.

[29] J. Lunden, V. Koivunen, A. Huttunen, and H. V. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Trans. Signal Processing*, vol. 57, pp. 4182-4195, Nov. 2009

[30] D. Teneketzis and Y. Ho, "The decentralized Wald problem," *Information and Computation (formerly Information and Control)*, vol. 73, no. 1, pp. 2344, Apr. 1987.

[31] V. V. Veeravalli, T. Basar, and H. V. Poor, "Decentralized sequential detection with a fusion center performing the sequential test," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 433-442, Mar. 1993.

[32] V.V. Veeravalli, "Sequential Decision Fusion: Theory and Applications," *J. Franklin Inst.*, vol. 336, pp. 301-322, 1999.

[33] Y. Mei, "Asymptotic optimality theory for decentralized sequential hypothesis testing in sensor networks," *IEEE Trans. Inf. Theory* vol. 54, pp. 2072-2089, May. 2008.

[34] G. Fellouris and G. V. Moustakides, "Decentralized sequential hypothesis testing using asynchronous communication," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 534-548, Jan. 2011.

[35] J. N. Tsitsiklis, "On threshold rules in decentralized detection," in *Proc. 25th IEEE Conf. Decision Contr.*, Athens, Greece, 1986, pp. 232-236.

[36] V. N. S. Samarasooriya and P. K. Varshney, "Sequential approach to asynchronous decision fusion," *Opt. Eng.*, vol. 35, no. 3, pp. 625-633, 1996.

[37] J. K. Sreedharan and V. Sharma, "Spectrum sensing using distributed sequential detection via noisy reporting mac," in *Signal Processing*, vol. 106, pp. 159-173, January 2015.

[38] A. M. Hussain, "Multisensor distributed sequential detection," *IEEE Trans. Aerosp. Elect. Syst.*, vol. 30, no. 3, pp. 698-708, Jul. 1994.

[39] A. K. Sahu and S. Kar, " Distributed sequential detection for Gaussian shift-in-mean hypothesis testing," *IEEE Trans. Signal Process.*, vol. 64, pp. 89-103, Sept. 2015.

[40] S. Zacks, *The Theory of Statistical Inference.* New York, NY: Wiley, 1971.

[41] J.-W. Lee and G. E. Dullerud, "A dynamic decentralized sequential multi-hypothesis testing problem under uniformly distributed nonstationary observations," *Proceedings of the 43rd IEEE Conference on Decision and Control*, vol. 2, pp. 1976–1981, 2004.

[42] Y. Wang and Y. Mei, "Asymptotic optimality theory for decentralized sequential multihypothesis testing problems," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 7068-7083, Oct. 2011.

[43] G. V. Moustakides, "Decentralized CUSUM change detection," *in Proc. 9th Int. Conf. Information Fusion (ICIF)*, Florence, Italy, 2006, pp. 1-6.

[44] O. Hadjiliadis, H. Zhang, and H. Vincent Poor, "One shot schemes for decentralized quickest change detection," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3346-3359, July 2009.

[45] L. Zacharias and R. Sundaresan, "Decentralized sequential change detection using physical layer fusion," *in Proc. ISIT*, France, Jun. 2007.

[46] T. Banerjee, V. Sharma, V. Kavitha, and A. K. Jayaprakasam, "Generalized analysis of a distributed energy efficient algorithm for change detection," *IEEE Trans. wireless commun.*, vol. 10, no. 1, pp. 91-101, Jan. 2011.

[47] V. V. Veeravalli, "Decentralized quickest change detection," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1657-1665, May 2001.

[48] T. Banerjee, V. V. Veeravalli, and A. Tartakovsky, "Decentralized data-efficient quickest change detection," *In Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, pp. 2587-2591, Jul. 2013.

[49] T. Banerjee and V. V. Veeravalli, "Data-efficient quickest change detection in minimax settings," *ArXiv e-prints*, Nov. 2012. http://arxiv.org/abs/1211.3729.

[50] T. L. Lai, "Sequential analysis: some classical problems and new challenges," *Statist. Sinica*, vol. 11, no. 2, pp. 303-408, Apr. 2001.

[51] G. Lorden, "Structure of sequential tests minimizing an expected sample size," *Z. Wahrsch. Verw. Gebiete*, vol. 51(2), pp. 291-302, Jan. 1980.

[52] S. Tantaratana and J. B. Thomas, "Relative efficiency of the sequential probability ratio test in signal detection," *IEEE Trans. Inform. Theory*, vol. 24, pp. 22-31, Jan. 1978.

[53] E. Paulson, "A note on the efficiency of the Wald sequential test," *Ann. Math. Stat.*, vol. 18, pp. 447-450, 1947.

[54] A. Nayyar and D. Teneketzis, "Sequential problems in decentralized detection with communication," *IEEE Trans. Inf. Theory*, vol. 57, pp. 6196-6214, Sept. 2011.

[55] Y. Mei, "Information bounds and quickest change detection in decentralized decision systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2669-2681, 2005.

[56] G. V. Moustakides, "Decentralized CUSUM change detection, " *Proc. of the 9th Int. Conf. Information Fusion (ICIF)*, Florence, Italy, Jul. 2006, pp. 1-6.

[57] R. W. Crow and S. C. Schwartz, "Quickest detection for sequential decentralized decision systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 1, pp. 267-283, 1996.

[58] J. N. Tsitsiklis, "On threshold rules in decentralized detection," *in Proc. 25th IEEE Conf. Decision Contr.*, Athens, Greece, 1986, pp. 232-236.

[59] H. R. Hashmi and I. B. Rhodes, "Decentralized sequential detection," *IEEE Trans. Inf. Theory*, vol. 35, pp. 509-520, May 1989.

[60] V. V. Veeravalli, " Comments on 'Decentralized sequential detection'," *IEEE Trans. Inf. Theory*, vol. 38, pp. 1428-1429, Jul. 1992.

[61] T. Basar and J. B. Cruz, " Concepts and methods in multi-person coordination and control, " in *Optimization and Control of Dynamic Operational Research Models*, S. G. Tzafestas, Ed, Amsterdam: North-Holland, 1982, pp. 351-394.

[62] H. S. Witsenhausen, "Separation of estimation and control for discrete-time systems," *Proc. IEEE*, vol. 59, pp. 1557-1566, 1971.

[63] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[64] K. Premkumar and A. Kumar, "Optimal sleep/wake scheduling for quickest intrusion detection using sensor networks," *IEEE INFOCOM,* Arizona, USA, 2008.

[65] D. P. Bertsekas, "Dynamic programming and optimal control," vol. I and II (3rd ed.). Athena Scientific. 2007.

[66] A. Wald and J. Wolfowitz, "Optimal character of the sequential probability ratio test," *Ann. Math. Statist.*, vol. 19, pp. 326-339, 1948.

[67] J. Thomas, "Nonparametric detection," *Proc. IEEE*, vol. 58, no. 5, pp. 623-631, May 1970.

[68] R.J. Serfling, *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, New York, 1980

[69] J. Cochlar and I. Vrana, "On the optimum sequential test of two hypotheses for statistically dependent observations," *Kybernetika*, vol. 14, pp. 57-69, 1978.

[70] G. Lorden, "Procedure for reacting to a change in distribution," *Ann. Math. Statist*, vol. 42, pp. 1897-1908, 1971.

[71] W. D. Ewan and K. W. Kemp, "Sampling inspection of continuous processes with no autocorrelation between successive results, " *Biometrika*, vol. 47, pp. 363-380, 1960.

[72] I. F. Blake and H. V. Poor, Communications and Networks: A Survey of Recent Results. New York: Springer-Verlag, 1986.

[73] T. L. Lai, "Asymptotic optimality of invariant sequential probability ratio tests," *Ann. Statist.*, vol. 9, pp. 318-333, 1981.