Poster Abstract: Efficient Decision Fusion with Performance Guarantees in Sensor Networks

Zinaida Benenson
Chair of Computer Science 4
RWTH Aachen University
zina@i4.informatik.rwth-aachen.de

Gernot Fabeck
Chair of Theoretical Information Technology
RWTH Aachen University
fabeck@ti.rwth-aachen.de

ABSTRACT

We consider sensor nodes which make decisions about the state of the observed environment and transmit them to a fusion center for decision combining. We investigate how to ensure pre-specified performance guarantees for the fused decision most efficiently.

1. INTRODUCTION

One important task of wireless sensor networks is the detection of physical phenomena in the observed environment. Consider a set of sensor nodes which all observe the same geographic area, make decisions about the state of the observed environment (e.g., dangerous or safe), and transmit their local decisions to a fusion center (Fig. 1).

Sensor nodes are cheap, so they can be deployed densely [1]. However, their local decisions are fairly unreliable due to their low-cost design and random deployment. The fusion center combines the unreliable local decisions into a reliable fused decision which satisfies some predefined performance measures. For example, the probability of a false alarm (detecting a nonexistent event) at the fusion center is guaranteed to be below a specific value.

As sensors, as well as the communication medium, are unreliable, the fusion center should not have to wait for local decisions of *all* sensors deployed in the area before making the decision: Some sensors may have failed, or some messages may have got lost. Moreover, waiting for more local decisions increases the time of the decision fusion operation. On the other hand, if decision fusion is based on a too small number of local decisions, the fused decision may be too unprecise.

We investigate the relationship between the number of local decisions the fusion center has to wait for, and the quality of the fused decision. Given the individual error probabilities of each sensor, we determine lower bounds on the number of local decisions needed for pre-specified performance guarantees at the fusion center.

2. PRELIMINARIES

2.1 Problem Statement

We consider a binary hypothesis testing problem with hypotheses H_0, H_1 describing the state of the observed environment and their associated prior probabilities π_0, π_1. A set of N sensors take measurements on the environment and make local decisions about the underlying true hypothesis.

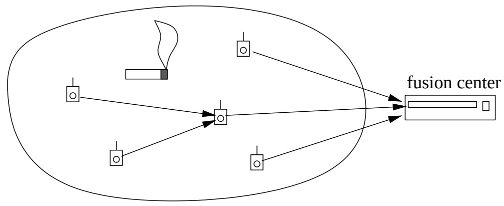

Figure 1: Sensor nodes observe the environment and make unreliable local decisions about its state. Fusion center combines local decisions into a reliable fused decision (e.g., fire/no fire).

Each local decision u_i is interpreted as the realization of a binary random variable U_i, $i = 1, ..., N$, which is characterized by the associated *probability of false alarm* and *probability of miss*:

$$p_f^i := P(U_i = 1 \mid H_0), \quad p_m^i := P(U_i = 0 \mid H_1).$$

The local decisions $u_1, ..., u_N$ form the input to the fusion center which combines them to yield the global decision $u = f(u_1, ..., u_N)$. As in the case of local decisions, the global decision u is interpreted as the realization of a binary random variable U which is characterized by the global probability of false alarm and global probability of miss:

$$p_F := P(U = 1 \mid H_0), \quad p_M := P(U = 0 \mid H_1).$$

The overall *probability of error* p_E at the fusion center is a weighted sum of the false alarm and miss rate:

$$p_E = \pi_0 p_F + \pi_1 p_M.$$

Performance guarantees at the fusion center are determined by upper bounds on its respective error probabilities.

We assume that the error probabilities p_f^i, p_m^i of the sensor nodes are known and that the local decisions U_i are conditionally independent. Our aim is to assess the error probabilities p_F, p_M of the fusion center and give lower bounds on the number of reporting sensors N in order to achieve pre-specified performance guarantees at the fusion center.

2.2 Optimal Fusion Rules

We consider optimal fusion rules in a Bayesian framework [2]. The objective is to determine the fusion rule f that minimizes the overall probability of error p_E. The problem can be viewed as a binary hypothesis testing problem at the fusion center with local decisions being the observations.

According to Chair and Varshney [3], the optimal fusion rule in the case of conditionally independent decisions can be performed by taking a weighted sum of the incoming local decisions and comparing it with a threshold:

$$\sum_{i=1}^{N}\left[\log\frac{(1-p_f^i)(1-p_m^i)}{p_f^i p_m^i}\right]u_i \underset{u=0}{\overset{u=1}{\gtrless}} \log\left[\frac{\pi_0}{\pi_1}\prod_{i=1}^{N}\left(\frac{1-p_f^i}{p_m^i}\right)\right].$$

For reasons of analytical tractability, we will consider equal local error probabilities $p_f^i \equiv p_f$, $p_m^i \equiv p_m$. This yields a simplified fusion rule taking the form

$$\sum_{i=1}^{N} u_i \underset{u=0}{\overset{u=1}{\gtrless}} \vartheta,$$

where the threshold ϑ takes the form $\vartheta = \alpha + \beta N$ with constants α and β.

3. DECISION FUSION WITH PERFORMANCE GUARANTEES

We investigate the error probabilities of the fusion center and give a lower bound on the number N of reporting sensors needed to achieve a pre-specified performance level.

3.1 Exact Expressions for Error Probabilities

In order to compute the error probabilities at the fusion center, we determine the distribution of the random variables

$$V_j := \sum_{i=1}^{N} U_i \,|\, H_j, \quad j = 0, 1,$$

i.e., the total number of "ones" sent by the N sensor nodes under hypothesis H_j true. It is easily shown that the random variables V_0 and V_1 follow a binomial distribution according to $V_0 \sim \mathrm{Bin}(N, p_f)$ and $V_1 \sim \mathrm{Bin}(N, 1-p_m)$.
By using the connection to the beta distribution of the first kind [4], we obtain expressions for the error probabilities $p_F = P(V_0 > \vartheta)$ and $p_M = P(V_1 < \vartheta)$:

$$p_F = 1 - \frac{N!}{\lfloor\vartheta\rfloor!(N-\lfloor\vartheta\rfloor-1)!}\int_{p_f}^{1} x^{\lfloor\vartheta\rfloor}(1-x)^{N-\lfloor\vartheta\rfloor-1}dx,$$

$$p_M = \frac{N!}{(\lceil\vartheta\rceil-1)!(N-\lceil\vartheta\rceil)!}\int_{1-p_m}^{1} x^{\lceil\vartheta\rceil-1}(1-x)^{N-\lceil\vartheta\rceil}dx.$$

3.2 Approximation by Normal Distribution

In typical wireless sensor network scenarios, the number of sensors N reporting to the same fusion center is large enough (e.g., $N = 20$), so that we may apply the Central Limit Theorem [4] to obtain a reasonable approximation for the corresponding error probabilities at the fusion center. Particularly, we obtain the expressions

$$p_F \approx 1 - \Phi\left(\frac{\vartheta - Np_f}{\sqrt{Np_f(1-p_f)}}\right),$$

$$p_M \approx \Phi\left(\frac{\vartheta - N(1-p_m)}{\sqrt{Np_m(1-p_m)}}\right),$$

where Φ is the cumulative distribution function (cdf) of the standard normal distribution.

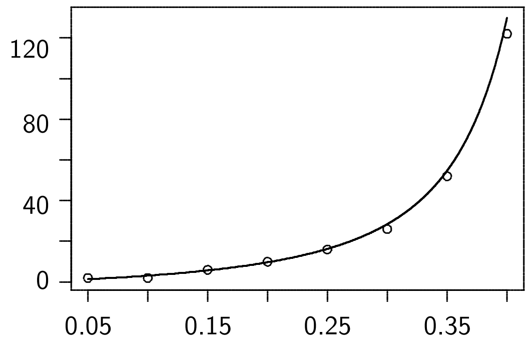

Figure 2: A lower bound on the number of sensors N needed to guarantee the performance level $p_E \le 0.01$ at the fusion center of the symmetric system for various local error probabilities. Circles depict exact values, the line depicts the approximation.

3.3 Special Case: Symmetric System

We illustrate the validity of our approximation by the special case of a *symmetric system*, i.e., $\pi_0 = \pi_1 = \frac{1}{2}$ and $p_f = p_m = p_e$.
In this case, we have $p_F = p_M = p_E$ and thus we can impose a unique lower bound on the number of sensors N to guarantee the restriction $p_E \le \varepsilon_E$ on both the false alarm and the miss rate at the fusion center by the same value ε_E:

$$N \ge \left(\frac{2\Phi^{-1}(\varepsilon_E)\sqrt{p_e(1-p_e)}}{1-2p_e}\right)^2.$$

The relationship between the necessary number of sensors N and the local error probability p_e is depicted in Fig. 2 for the specific performance guarantee $p_E \le \varepsilon_E = 0.01$.

4. FURTHER WORK

In the future, we aim to investigate the influence of correlations between local decisions on the number of sensors needed. Correlations will occur naturally due to dense deployment. Furthermore, we want to consider the case of m-ary hypothesis testing for distributed classification applications involving heterogeneous sensor types. We want to examine the possible advantages of multiple layers in the decision hierarchy. By allocating sensor nodes to fusion centers across multiple layers, we aim to achieve energy and time savings while maintaining performance guarantees on the final decision.

5. REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: A survey. Computer Networks, 38(4):393–422, 2002.

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

[3] Z. Chair and P. K. Varshney. Optimal data fusion in multiple sensor detection systems. In *IEEE Trans. Aerosp. Electron. Syst.*, 22(1):98-101–30, 1986.

[4] A. Papoulis and S. U. Pillai. *Probability, Random Variables and Stochastic Processes*, 4th edition, McGraw-Hill, 2002.